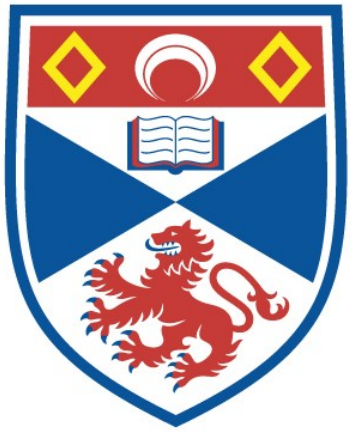
From the field to the laboratory, investigating the effects of Deformed wing virus on honey bee health

Luke Woodford

A thesis submitted for the degree of PhD
at the
University of St Andrews



2022

Abstract

Varroa destructor is an ectoparasitic mite associated with significant losses of honey bee colonies globally. The mite vectors a range of pathogenic viruses, most notably Deformed wing virus (DWV). DWV is transmitted orally between generations of bees and by this route rarely causes symptomatic infection, but the introduction of *Varroa* alters the transmission route of the virus and leads to highly elevated virus titres. Annual overwintering colony losses of ~25% are associated with high levels of *Varroa*-DWV infestation. Effective miticide treatments are available to control *Varroa*. However, the absence of coordinated treatment means environmental transmission of mites continues unchecked. One of the aims of this study was to determine whether rational, coordinated treatment is beneficial to colony health. Over a period of three years, all colonies managed on the island of Arran were treated in unison. Changes in the mite levels and DWV levels and diversity were measured as indicators of changing colony health. Across the three years of analysis, total mite numbers decreased by 58% and the number of managed colonies increased by 50%. In parallel to the Arran study, a colony management technique – a shook swarm – combined with appropriate miticide treatments was shown to be highly effective in reducing mite infestations and DWV titres. The impact of no mite management was investigated by placing healthy colonies in a high-mite apiary and measuring their rates of mite uptake. This study indicated that mites rapidly enter the colonies, altering the DWV dynamic and typically resulting in colony death. A mite-virus-bee *in-silico* model, simulating the dynamic changes in the virus population across multiple generations of pupae was designed and recombination between two DWV variants was investigated to examine recombination junctions. These studies shed light on how *Varroa*-control and colony management techniques impact DWV levels and diversity and honey bee colony health.

Table of Contents

List of figures

List of tables

Acknowledgements

I would firstly like to thank Professor David J Evans for his support and guidance throughout the five years of this PhD project and for giving me the freedom to explore interesting avenues of research. In my time here I have been offered lots of opportunities to investigate different topics, work with other great scientists and travel to meet collaborators, so thank you. The odd jar of honey was also very welcome!

A big thank you to Fiona Highet and Graeme Sharpe for all their work on the Arran project, it would not have been possible without both of you. I had great fun roaming across the island with you hunting out those colonies! Thank you also to Fiona for providing me with the opportunity to do a 3-month work placement at SASA, which was very enjoyable. Thank you to Dr Alan Bowman at the University of Aberdeen and Professor Mike Ritchie at the University of St Andrews for additional support and guidance. I am grateful to the BBSRC and EastBIO for funding.

Thank you to all the Evans lab members past and present, it has been great fun working with you all. Special shoutouts to Kirsten Bentley and Alex (Olesya) Gusachenko for your endless advice, support and kindness. I really enjoyed working on all the projects we did together, and I hope our paths cross again in the future! Thank you to the Bowman lab in Aberdeen, especially Craig Christie, Ewan Campbell and Amy Cooper, for all their help with fieldwork, for providing lab space for me after the BMS fire and for all the good nights out!

Thank you to everyone in the department – what was the BMS and is now the WRL – particularly Ashley Pearson and Dan Young for always being there to lend a hand, fix things or show me where things are for the umpteenth time! Thank you to the office staff, Jean Johnson, Margaret Wilson and Val Milne for all your help with my endless admin questions and for fuelling this PhD with all the biscuits!

Thank you to my friends and family for all their support during all the ups and downs. Big shoutout to Chris and Pete for all the advice and laughs, you guys made the whole process easier! Thank you to my parents for all their love and support over the years, I would not have been able to do this without you both. Thank you to my grandparents on my father's side who have supported me throughout my time at university and on my mother's side who are sadly no longer with us – I wish I could talk to you both about this too. Thank you also to my high school chemistry teacher, Ms Cleary, who phoned my parents after a few dodgy prelim grades! I rather brashly told her I'd get a PhD one day…

Finally, thank you to Laura. You have probably endured this PhD more than I have! Thank you for always being there, for keeping me grounded and for being the support I needed when I needed it. I could not have done this without you!

Funding

This work was supported by the BBSRC (Biotechnology and Biological Sciences Research Council) through an EastBIO funded PhD scholarship [Grant number: BB/M010996/1].

Author’s declaration

Candidate's declaration

I, Luke Woodford, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 60,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2016.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 08/02/22 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date 08/02/22 Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Luke Woodford, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

Embargo on all of electronic copy for a period of 1 year on the following ground(s):

- Publication would preclude future publication

Supporting statement for electronic embargo request

Manuscripts containing data generated and presented in this thesis are being prepared and publication of this thesis prior to the completion of those manuscripts would be detrimental to that work.

Title and Abstract

- I agree to the title and abstract being published.

Date 08/02/22 Signature of candidate

Date 08/02/22 Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Luke Woodford, understand that by declaring that I have original research data or digital outputs, I should make every effort in meeting the University's and research funders' requirements on the deposit and sharing of research data or research digital outputs.

Date 08/02/22 Signature of candidate

Permission for publication of underpinning research data or digital outputs

We understand that for any original research data or digital outputs which are deposited, we are giving permission for them to be made available for use in accordance with the requirements of the University and research funders, for the time being in force.

We also understand that the title and the description will be published, and that the underpinning research data or digital outputs will be electronically accessible for use in accordance with the license specified at the point of deposit, unless exempt by award of an embargo as requested below.

The following is an agreed request by candidate and supervisor regarding the publication of underpinning research data or digital outputs:

No embargo on underpinning research data or digital outputs.

Date 08/02/22 Signature of candidate

Date 08/02/22 Signature of supervisor

List of abbreviations

ABPV - Acute bee paralysis virus
BAM - Binary alignment map
BLAST - Basic local alignment search tool
CBPV - Chronic bee paralysis virus
cDNA - Complementary deoxyribonucleic acid
COI - Cytochrome oxidase I
DDD - Reverse genetic construct of Type A DWV
DNA - Deoxyribonucleic acid
dNTP - Deoxyribonucleotide triphosphate
DWV - Deformed wing virus
E. coli - *Escherichia coli*
EDTA - Ethylenediaminetetraacetic acid
FP - Forward primer
GE - Genome equivalent
IAPV - Israeli acute paralysis virus
IRES - Internal ribosome entry site
LB - Luria broth
NCBI - National Centre for Biotechnology Information
NGS – Next-generation sequencing
nt - Nucleotide
ORF - Open reading frame
PCR - Polymerase chain reaction
qPCR - Quantitative polymerase chain reaction
RdRp – RNA-dependent RNA polymerase
RNA - Ribonucleic acid
RP - Reverse primer
RT - Reverse transcription
SAM - Sequence alignment map
SASA – Science and Advice for Scottish Agriculture
SBV - Sacbrood virus
ShoRAH - Short Reads Assembly into Haplotypes
SNP - Single nucleotide polymorphism
SNV - Single nucleotide variant
SRUC – Scotland's Rural College
TAE – Tris-acetate-EDTA
TBE - Tris-borate-EDTA
TE - Tris-EDTA
UTR – Untranslated region
UV – Ultraviolet
VDD – Reverse genetic construct of recombinant DWV
ViReMa – Viral Recombination Mapper
VVD – Reverse genetic construct of recombinant DWV
VVV – Reverse genetic construct of Type B DWV

Definitions

Variant – Major types of DWV, typically referred to as Type A, Type B and recombinants thereof are classified as variants throughout this thesis

Haplotype – a unique sequence as defined by probabilistic alignments from ShoRAH
Cluster – a group of sequences with identical or highly similar sequence identity, which group together during phylogenetic analysis
Site – a region of the Isle of Arran assigned based on the proximity of honey bee colonies, typically each town with bees made up a ‘Site’.

1 Introduction

"A ~~towel~~ ***thesis****, [The Hitchhiker's Guide to the Galaxy] says, is about the most massively useful thing an interstellar ~~hitchhiker~~* ***researcher*** *can have. Partly it has great practical value. You can wrap it around you for warmth as you bound across the cold moons of Jaglan Beta; you can lie on it on the brilliant marble-sanded beaches of Santraginus V, inhaling the heady sea vapors; you can sleep under it beneath the stars which shine so redly on the desert world of Kakrafoon; use it to sail a miniraft down the slow heavy River Moth; wet it for use in hand-to-hand-combat; wrap it round your head to ward off noxious fumes or avoid the gaze of the Ravenous Bugblatter Beast of Traal (such a mind-boggingly stupid animal, it assumes that if you can't see it, it can't see you); you can wave your ~~towel~~* ***thesis*** *in emergencies as a distress signal, and of course dry yourself off with it if it still seems to be clean enough."*
— Douglas Adams, The Hitchhiker's Guide to the Galaxy.

Note: *some of this chapter has been published as Woodford & Evans (2020) – see* ***Appendix 8-7***

1.1 Threats to honey bees

1.1.1 Are key pollinators in decline?

The evolutionary origin of the European honey bee (*Apis mellifera*) is unclear, with evidence supporting expansion from either Asia or Africa (Whitfield *et al.*, 2006; Han, Wallberg and Webster, 2012). What is clear is that subsequent global expansion was undoubtedly aided by human intervention, initially by provision of nesting sites and later by direct colony management, with evidence for recognisable 'beekeeping' activity dating back several thousand years. This anthropogenic expansion was driven by the demand for honey, a highly prized natural source of glucose. More recently, the movement of honey bees within the UK and globally has expanded in order to increase pollination of a large number of economically important agricultural crops. Although local pollinating insects play a key role in the pollination of crops, the honey bee is considered essential for meeting the growing demands of some sectors of agriculture. For example, every year approximately two thirds of all colonies in the USA are moved to California, where they pollinate almond crops that yield up to 80% of the total world almond harvest (Beaurepaire *et al.*, 2020). In the UK there are 39 insect-pollinated agricultural crops, the vast majority of which are pollinated by honey bees, and in the USA there are over 130 agricultural plants pollinated by honey bees and bumble bees (Delaplane and Mayer, 2000). Despite the obvious need for honey bees, the global stock of farmed bees is not growing at the rate required by the demands of the industry (Aizen and Harder, 2009). Additionally, although the positive impact of farmed honey bees on agricultural production is well established, their impact on natural ecosystems remains a point of contention and needs further study (Costanza *et al.*, 1997), with some suggesting it may be detrimental (Lindström *et al.*, 2016).

Despite the number of managed colonies increasing globally by 45% in the last half century (Aizen and Harder, 2009), there have been large regional declines (*e.g.* 59% decline in the period 1947–2005 in the USA (Potts *et al.*, 2010), average 28% annual US

colony losses 2006–2020 (Bruckner *et al.*, 2020) and up to 32% overwinter losses in parts of Europe in 2018–19 (Gray *et al.*, 2020)). These losses have not been attributed to a single cause and are believed to reflect a combination of anthropogenic factors. These include increased pesticide use (Chauzat *et al.*, 2009), the stress induced by long-distance movement of colonies to aid crop pollination (Morimoto *et al.*, 2011), and the increased presence and transmission of pathogens. In addition, monoculture farming practices lead to a substantial decrease in suitable habitat (Donaldson-Matasci and Dornhaus, 2012). These habitat losses have been deemed one of the largest causes of total insect population declines (Hallmann *et al.*, 2017) and increased habitat fragmentation is a key factor for honey bee survival too (Brown and Paxton, 2009). Following very large-scale honey bee colony losses in the USA in 2006, the term 'colony collapse disorder' (CCD) became widely used to describe the phenomenon and particular features of lost hives (Hayes, 2007; Watanabe, 2008; vanEngelsdorp *et al.*, 2009). Although the term CCD remains a convenient acronym to define colony losses, losses of a similar magnitude and type have not been routinely reported subsequently. The causes of CCD remain unclear and it is thought to be a combination of some or all of these anthropogenic factors (Cox-foster *et al.*, 2007). However, the magnitude of the losses refocused attention on the increasing problem of pathogens in managed honey bee colonies.

The spread of parasites and pathogens is a significant factor in honey bee health. Honey bees have evolved to control pathogens that spread within the colony and naturally restrict colony-to-colony transmission by existing at low nest densities in the environment (Seeley, 2007; Seeley *et al.*, 2015). In contrast, managed honey bees hives are co-located in apiaries, are routinely moved for crop pollination and are kept at a significantly higher overall density than natural colonies. This increases the risk of disease transmission between colonies or apiaries, through the natural processes of drifting, where bees relocate to a hive other than their natal colony (Pfeiffer *et al*., 1998; Nolan *et al*., 2016), and during the robbing of weak colonies by nearby strong colonies (Peck and Seeley, 2019). Due to these processes, many pathogens have been able to establish and spread rapidly through the global honey bee population (Dietemann, Ellis and Neumann, 2013).

1.2 *Varroa destructor* infestations and treatments

1.2.1 The history of *Varroa destructor*

Originally an ectoparasitic mite of the Asian honey bee *Apis cerana*, *Varroa destructor* jumped host to the European honey bee, *Apis mellifera*, in Asia. The mite subsequently spread globally as honey bee colonies were transported from country to country, reaching Europe in the 1980s and the UK in 1992 (Oldroyd, 1999). Originally known as *Varroa jacobsoni*, molecular methods subsequently showed that *Varroa destructor* was a distinct species (Anderson and Trueman, 2000).

Female *Varroa,* known as foundress mites, enter unsealed brood cells in honey bee colonies, where they lay eggs and feed on the developing pupa, transmitting viruses in the process. The progeny, usually 2–3 females and a male, hatch and mate with one another and feed on the pupa (Traynor *et al.*, 2020). The mature female mites emerge when the honey bee brood reaches maturation, and climb onto the backs of nurse bees, where they either remain phoretic or move to the next brood cell, typically preferring drone brood due to their increased reproductive potential (Fuchs, 1992).

Varroa are efficient and effective invaders of honey bee colonies, moving between colonies in apiaries and across neighbouring sites. Colonies that are left untreated quickly succumb to mite infestation. This, coupled with their ability to vector a range of viruses, altering the viral landscape (Wilfert, *et al.*, 2016), makes *Varroa* a significant threat to beekeeping and honey bee health. Together, the virus and mite are the single largest cause of overwintering colony losses (Fries, Imdorf and Rosenkranz, 2006).

Figure 1-1 – A mite on a honey bee with deformed wings. *A phoretic* Varroa *mite and a newly emerged honey bee exhibiting symptoms of Deformed wing virus.*

1.2.2 Drifting and robbing between colonies

The presence of *Varroa* within one colony also potentially impacts the health of neighbouring colonies. Modelled data and microsatellite analysis have indicated as many as 42% of workers in a colony could be alien due to drifting between colonies (Pfeiffer and Crailsheim, 1998; Forfert *et al.*, 2015). Drifting is most common during the orientation flights of newly foraging bees and workers entering neighbouring colonies do not experience any antagonism, regardless of their age (Šekulja, Pechhacker and Licek, 2014). It is inevitable that the drifting of bees between colonies leads to the spread of pathogens. Phoretic mites are known to spread between colonies on worker bees, particularly in tightly packed apiaries (**Figure 1-2**), meaning that highly infested hives can transfer mites to neighbouring colonies (Fries *et al.,* 2001). Infestation can be rapid, with rates of ~76 mites a day recorded in mite-free colonies from neighbouring mite-infested colonies (Greatti *et al.*, 1992).

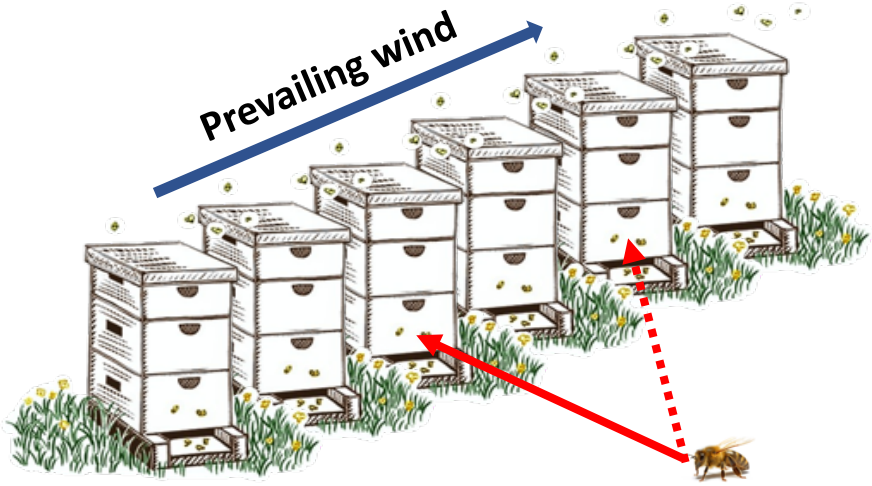


Figure 1-2 – Honey bees drift between colonies. *An example of closely packed colonies, where worker bees will drift between hives when they return from foraging and a weak colony may be prone to robbing by neighbouring colonies. The blue arrow indicates the prevailing wind direction. The solid red arrow indicates the bee's intended destination and the dashed arrow indicates the colony the bee drifts to.*

In addition to drifting bees, robbers are also associated with *Varroa* infestation and mites are known to climb onto robber bees to enter a new hive (Greatti, Milani and Nazzi, 1992; Fries and Camazine, 2001; Loftus, Smith and Seeley, 2016). Weak colonies, potentially debilitated by high levels of mite infestation, are robbed by nearby strong colonies who, in turn, acquire an increased level of mites and the novel virus populations they harbour (Peck and Seeley, 2019). This is further exacerbated by colonies with high *Varroa* infestation showing an increased acceptance of drifting workers, allowing for the potential uptake of other associated pests and pathogens (Forfert *et al.*, 2015).

Swarming of large colonies has been shown to reduce the level of mites associated with the colony, because the swarm carries a proportion of the mites with it, up to 35%, and increases the chances of the colony surviving over winter; however, this disease-limiting effect is negated when drifting and robbing occur in overcrowded colonies (Seeley and Smith, 2015). Feral colonies, presumed to originate from swarms lost from managed hives, have pathogen levels – including DWV and *Varroa* – that are as high as those in unmanaged hived colonies (Thompson *et al.*, 2014). This indicates that feral colonies may act as reservoirs for re-infestation within the environment, and a potential source of disease for non-*Apis* pollinators.

1.2.3 *Varroa* control and miticide treatments

For several reasons, *Varroa* represent a unique problem for beekeepers in the UK. They were introduced to the country relatively recently (1992), meaning experience among beekeepers for treating infestations is limited, and they are a relatively new parasite to the western honey bee, *Apis mellifera*, meaning that a balanced host–parasite relationship does not exist (Rosenkranz, Aumeier and Ziegelmann, 2010). It is clear that *Varroa destructor* is having a large impact on beekeeping practice and is a significant contributor to annual losses of honey bees in Europe (De La Rua *et al.*, 2009).

There are a variety of chemical treatments available to beekeepers for *Varroa* infestations in the UK, including oxalic acid, formic acid, pyrethroid-based miticides and thymols. Additionally, there are beekeeping practices that can help reduce and control *Varroa* spread, such as a shook swarm, queen trapping and drone brood uncapping (Rosenkranz, Aumeier and Ziegelmann, 2010). When *Varroa* is first detected in a new geographical location, the rate at which beekeepers develop their practices and start treating their colonies for *Varroa* has a large bearing on the persistence and spread of the parasite (Rosenkranz, Aumeier and Ziegelmann, 2010) and without standard protocols to implement such measures, the mite will rapidly spread and infest colonies.

Several 'hard' miticides, such as coumaphos, pyrethroids such as Apistan and Amitraz, have been used extensively by beekeepers for over 20 years (Milani and Iob, 1998). The widely used fluvalinate compound Apistan in has been used in hives for decades, however resistance has been widely reported in Italy in 1992 and in the UK in 2001 (Toufailia and Ratnieks, 2016) and use of this miticide is now discouraged. When effective, Apistan kills >99% of the mites present in a hive and is only mildly toxic to honey bees and beekeepers; however, when resistance is observed the miticide only kills 30% of the *Varroa* present (MJ Gracia-Salinas *et al.*, 2006) and the resistant mites have been shown to be able to stand up to 20 times the normal dose in some cases (Thompson *et al.*, 2002). This highlights that, while hard miticides can be very effective, improper use and overexposure at the wrong dosage can be detrimental to the colony. Hard miticides can also potentially contaminate honey stocks and leave residues in the wax foundation (Rosenkranz, Aumeier and Ziegelmann, 2010), a fact that is very likely to discourage beekeepers from applying them.

As an alternative to the hard miticides, some soft organic compounds are also effective, such as oxalic acid, formic acid, lactic acid and thymol. These compounds are less harmful to the honey bee colonies, they are often water soluble, they are less likely to result in resistance and in many cases they already contain a natural ingredient of honey (Rosenkranz, Aumeier and Ziegelmann, 2010). Oxalic acid has been shown to be as effective as hard-chemical treatments like Apistan: as little as 2.25g of oxalic acid administered to the hive by sublimation can kill 97% of the phoretic mites present without causing any harm to the colony of bees (Toufailia *et al.,* 2014). However, there are some disadvantages: many of the compounds, oxalic acid included, must be applied when the colony is brood-less, requiring accurate treating and good management. The difference between an effective miticide treatment and a dose that is toxic to the colonies is small and thus the climate and conditions within the hive need to be understood before treatment (Martín-Hernandez *et al.*, 2007). Due to these issues the effects of soft chemicals can be more varied than those of registered hard acaricides.

An additional challenge for honey bee health and potential miticide resistance is the general competence, education and experience of the beekeepers maintaining colonies. Jacques *et al* (2017) found that overwinter colony losses were twice as high in amateur beekeeper apiaries as in professional beekeeper apiaries. Their study, encompassing 5,798 sites across two years, showed clear correlations between the levels of experience and education of the beekeepers and the colony losses observed. Additionally, they found no disease present in the colonies maintained by the professional beekeepers, but observed bacterial infections and high *Varroa* loads in the amateur beekeepers' hives. Educating beekeepers is clearly important to improving miticide applications. Knowing

when to apply the correct treatments, at which concentration, and for what length of time, is essential for maintaining the effectiveness of the treatment regimes.

1.2.4 Is coordination key? – treatment attempts and benefits

Despite the range of treatments available there is no single miticide that is implemented on a widespread and regulated scale that successfully removes all the *Varroa* from a colony. In the UK there is limited regulation currently in place governing the use of miticides, with no coordinated timing of treatment between apiaries, meaning that whilst some beekeepers treat appropriately others are not obliged to and often do not. This lack of coordination can have severe consequences for colony health.

The extent to which coordination on a landscape scale would impact mite levels in colonies has not been fully examined. Two studies have attempted to eradicate *Varroa* from highly infested and geographically isolated locations (islands, to which bees will not fly from neighbouring landmasses). These were on Jersey (Sampson and Martin 1999), an island in the English Channel, and on Gorgona (Giusti *et al.*, 2016), an island off the north-east coast of Italy, which had limited success. Despite multiple treatments of the colonies on Jersey, the mites were still detected in the apiaries examined; one of the major causes of failure was thought to be the size of the island, with unknown and infested colonies outside the study probably facilitating the persistence of *Varroa*. Gorgona is approximately 50 times smaller than Jersey and the eradication protocol focused on the only apiary on the island. The eradication process was effective: a single treatment with Apistan was performed in 2009, followed by monitoring of the hives between 2010 and 2014, which showed no signs of *Varroa* infestation. The study also examined the levels of Deformed wing virus (DWV), one of the viruses transmitted by *Varroa* mites, detected in the honey bees post-treatment and found, with the exception of one testing point in 2014, that the viral titre was low in all bees free from *Varroa* infestation (Giusti *et al.*, 2016). Although this was an effective process of mite removal, it does not properly consider the impacts of coordination on a larger scale. Further studies are required to determine the effect of long-term miticide coordination across multiple apiaries on mite levels and colony health, through the examination of mite numbers and transmitted viruses, such as DWV, post-treatment.

1.3 Deformed wing virus

1.3.1 The biology, prevalence and distribution of DWV

The long-held understanding that *Varroa* feed on the haemolymph of developing pupae has recently been questioned by studies that suggest it is the fat bodies that are the major target (Ramsey *et al.*, 2018). Notwithstanding this, *Varroa* ingests and subsequently transmits a cocktail of viruses while feeding on honey bee pupae (Tentcheva *et al.*, 2004). The most important of these is the single-stranded positive-sense RNA virus, Deformed wing virus (DWV). The introduction and dissemination of *Varroa destructor* in the global honey bee population added a fundamentally new transmission route for DWV. In the absence of the mite, DWV is transmitted vertically within the colony as infected nurse bees transmit the virus to developing larvae during feeding and the virus is detectable in the gut of most managed honey bee populations (Tentcheva *et al.*, 2004). However, in the presence of *Varroa*, DWV is able to bypass the normal routes of transmission and is directly injected into the developing honey bee pupae when the mites feed on them (Yue and Genersch, 2005). The virus subsequently increases in titre dramatically and spreads

through most of the bee's body (Fujiyuki *et al.*, 2004, 2005; Zioni *et al.*, 2011; Gusachenko *et al.*, 2020a).

DWV is a member of the *Iflavirus* genus of viruses, family *Iflaviridae*, order Picornavirales. The name Deformed wing virus encompasses at least three genetically similar viruses that were independently isolated. DWV was originally isolated from honey bees in the UK (Bailey and Ball, 1991) and a very similar virus, designated VDV-1, was isolated from *Varroa* parasitising *A. mellifera* in the Netherlands (Ongus *et al.*, 2004). A third closely related virus, designated Kakugo ('ready to attack' in Japanese), was isolated from the brains of honey bees exhibiting aggressive behaviours in Japan (Fujiyuki, *et al.*, 2004). In this overview the name DWV will be used to encompass all three, unless specific reference to VDV-1 or Kakugo is required. The genetic relationships between DWV strains are addressed below.

The genome organisation of DWV is characteristic of members of the Picornavirales. The RNA genome of DWV (**Figure 1-3**) contains a single large open reading frame (ORF) which encodes both structural and non-structural proteins of the virus (Ongus *et al.*, 2004). The ORF is flanked by a 1,144-nucleotide 5' non-translated leader sequence and a 317-nucleotide 3' non-translated region terminated by a poly(A) tail (Lanzi *et al.*, 2006). The 2,894-amino acid ORF is translated via an internal ribosomal entry site (IRES) located in the 5' non-translated region. The structural proteins, VP1 – 4, that form the icosahedral viral capsid (Organtini *et al.*, 2017; Škubník *et al.*, 2017) are located towards the N-terminal end of the polyprotein, preceded by the leader (L) polypeptide, the function of which has yet to be defined. The remainder of the polyprotein carries the non-structural proteins responsible for genome replication and the manipulation of the cellular milieu. These include a recognisable RNA helicase, a chymotrypsin-like 3C protease, and an RNA-dependent RNA polymerase (Lanzi *et al.*, 2006). The region of the genome encoding the L protein is a hotspot for variation between different viral variants (Lanzi *et al.*, 2006) and contributes 26.2%–33.3% of all amino acid variation found between DWV isolates despite comprising only 7.3% of the polyprotein sequence. This large disparity in the diversity of the L-protein and the remainder of the genome has been observed in other Picornaviruses where the L protein has a number of different functions including the stimulation of IRES activity or the inhibition of host cell mRNA translation (Glaser, Cencic and Skern, 2001; Hinton *et al.*, 2002).

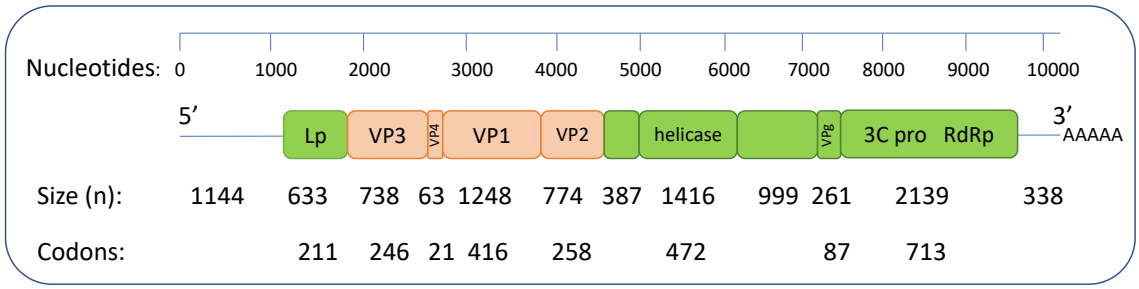


Figure 1-3 – The structure of the Deformed wing virus genome. *DWV is a positive-sense single-stranded RNA virus encoded by a single ORF, spanning approximately 10 kb. Genes encoding the structural proteins are shown in orange and genes encoding the non-structural proteins are shown in green. The size of each region is shown below in nucleotides (n) and in codons. The 5' and 3' non-transcribed regions are shown by a line at either end of the ORF. The figure is adapted from Dalmon et al. 2017.*

Evidence suggests that DWV is widespread in the honey bee population, even in geographic areas with no current or historical *Varroa* infestation. The exception may be Australia, which is reported to lack both the mite and DWV (Roberts *et al*. 2017). DWV has been detected in honey bee populations in isolated regions with no *Varroa* presence, such as on the island of Colonsay, Scotland and the Isle of Man, UK (Fürst *et al.*, 2014; Ryabov *et al.*, 2014; McMahon *et al.*, 2015), indicating that the virus is probably naturally endogenous in honey bees and suggesting that reports of its absence may reflect poor assay sensitivity or sampling techniques. If DWV is an endogenous viral parasite of honey bees it is unclear how Australian bees remain uninfected, as honey bees were introduced there from Europe in the 1820s. Despite the widespread distribution of DWV, there is compelling evidence that the anthropogenic dissemination of *Varroa* through the global honey bee population has also resulted in the transmission of particular DWV variants (Wilfert, *et al.*, 2016).

Figures quoted for DWV prevalence have gradually increased as assay methods have improved, with data from 32 countries suggesting an average prevalence of 55% (10%-100%) (Martin and Brettell, 2019). In the absence of *Varroa*, viral levels in individual adult workers can be very low (<1000 genome equivalents per microgram (GE/µg) of total RNA) but, as discussed below, can increase 1–100 million times after transmission by *Varroa*. This has fundamental implications for prevalence studies; pooled worker bee or pupal samples from colonies with even low levels of *Varroa* infestation may well exceed the detection threshold for DWV. In contrast, analysis of individual workers or pupae using poorly designed or insensitive assays may fail to detect DWV in the sample, particularly in colonies with low or zero mite levels.

1.3.2 DWV infections altered by *Varroa* mite transmission

How is a potentially ubiquitous virus largely or completely asymptomatic in the absence of *Varroa*? This appears to be related to the viral titre, potentially to the diversity of the virus population, and to the honey bee tissues to which the virus has access. In healthy honey bee colonies the virus population is highly diverse and the viral loads are low, typically 10^4–10^6 copies per bee with 5–20 variants reported (Martin et al., 2012; Mondet, et al. 2014; Ryabov et al., 2014), and bees that harbour such low titres are asymptomatic. During horizontal transmission, for example during trophallaxis, the bee is exposed to DWV via the gut, an organ shaped by evolution to provide protection from environmental pathogens (Mikonranta *et al.*, 2014). Altering the transmission route leads to significant

changes in the virus population, including a large increase in the viral load (titres increase by ~10^6-fold per bee), which is often accompanied by a marked reduction in virus diversity (Martin *et al.*, 2012; Nazzi *et al.*, 2012; Ryabov *et al.*, 2014).

Virus transmission by mites is causally associated with the characteristic symptoms of DWV infection, including crumpled and poorly developed wings, general paralysis, discoloration and abdominal bloating (**Figure 1-4**), as well as reduced longevity of worker bees (Highfield *et al.*, 2009; Möckel, Gisder and Genersch, 2011; Dainat *et al.*, 2012). However, it does not invariably cause overt symptoms. Honey bees with low DWV can still rarely develop crippled wings, possibly independently of the virus. In addition, up to 25% of bees with high levels of DWV (>10^{10} copies/bee) by injection can develop normal wings (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a). These apparently morphologically normal workers may still have impaired foraging abilities, loss of cognitive function and a reduced lifespan despite appearing healthy (Iqbal and Mueller, 2007; Dainat *et al.*, 2012; Campbell *et al.*, 2016; Gisder *et al.*, 2018). Highly infected honey bee drones are still able to compete reproductively and transmit virus, as shown when highly DWV-infected drones transmit virus when successfully mating with queens, resulting in infections of >10^7 copies in some mated queens (Amiri, Meixner and Kryger, 2016).

Figure 1-4 – Honey bees displaying the symptoms of Deformed wing virus infection.

1.3.3 The genetic diversity of Deformed wing virus

Replication errors by the RNA-dependent RNA polymerase of the single-stranded positive-sense RNA viruses mean that these viruses exist as a ‘cloud’ of very closely related variants termed a quasispecies (Domingo et al., 2005). In addition to this variation, there are identifiably divergent DWV isolates that have variously been termed DWV or Type A DWV (Lanzi *et al.*, 2006) and VDV-1 or Type B DWV (Ongus *et al.*, 2004). There is an additional, though rarely reported, Type C variant (Mordecai *et al.*, 2015; de Souza *et al.*, 2019; Kevill *et al.*, 2019). It should be noted that the Type C virus is distinct from Kakugo, which is identifiably a Type A strain of the virus.

These three variants exhibit very significant genetic relatedness, sharing over 80% nucleotide identity and >90% amino acid identity across the genome. For comparison, the three serotypes of the distantly related poliovirus (the prototype picornavirus), all of

which exhibit similar tropism and pathogenesis, are 71% identical at the nucleotide level and 69–88% at the amino acid level, depending on the region (Toyoda *et al.*, 1984).

A comparison of amino acid and nucleotide identity between DWV Type A (NCBI accession no. AJ489744.2) and Type B (NCBI accession no. AY251269.2) revealed a high degree of sequence similarity, with only the variable Leader protein region showing less than 83% nucleotide and 96.8% amino acid similarity (**Table 1-1**). Phylogenetic analysis of the nucleotide and amino acid differences between DWV Type A, DWV Type B and KV (Kakugo virus) indicate they are small enough (~84% average nucleotide identity) to be considered strains of the same virus (Lanzi *et al.*, 2006; Ryabov *et al.*, 2014).

Table 1-1 - Sequence similarities between two well characterised DWV variants. DWV Type A (NCBI accession no. AJ489744.2) and DWV Type B (NCBI accession no. AY251269.2)

Region of genome	Genome location	Sequence identity (%)	Amino acid identity (%)
leader protein	1132-1764	74.24	83.4
VP3	1765-2523	84.04	96.8
VP1	2587-3834	84.46	97.6
VP2	3835-4608	84.9	98.1
RNA helicase	4996-6411	88.69	98.1
protease	7681-8274	87.69	97.5
RdRp	8551-9615	83.29	97.5

In the VP1 region of the DWV genome (**Figure 1-3**), the Type-A and -B DWV variants exhibit 84.5% nucleotide and 98% amino acid identity. VP1 is an extensively studied region for measuring diversity between variants of human enteroviruses, a distantly related group of viruses with a similar genome organisation (Oberste *et al.*, 1999). Phylogenetic analysis of poliovirus and human enterovirus species C have shown that the variants within these groups are highly similar (differing by 14–16% between variants) and some strains have been reclassified based on their high amino acid similarity (>96%) and phenotypically similar infections (Brown *et al.*, 2003). Sequence analysis of the VP1 region of human enteroviruses showed intraserotypic divergence of about 25% in the nucleotide sequence and 12% in the amino acid sequence (Oberste *et al.*, 1999). Although antibody selection will have driven some variation in the vertebrate-infecting enteroviruses, it is likely that RNAi may act similarly in honey bees (Ryabov *et al.*, 2014), making the comparison of sequence variation and virus/strain classification valid. Considered in this context, the sequence divergence of even the most divergent DWV variants would place them as a single virus group. Initial reports of phenotypic differences between the variants (Natsopoulou *et al.*, 2017) have since been disputed (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a), further calling into question the classification of DWV into distinct variants.

It should be noted that even sequences with very limited genetic divergence can generate viruses with fundamentally different phenotypes. For example, a single nucleotide substitution in poliovirus determines the majority of the neurovirulence phenotype of the virus (Almond *et al.*, 1987). It remains possible that the genetic differences reported between the type A, B and C strains of DWV might obscure similar fundamental determinants that influence the resultant phenotype. The recent availability of reverse genetic systems will, as they did with poliovirus, allow this to be determined in the future.

A comparison of all DWV sequences >9500bp uploaded on the NCBI database revealed a close genetic similarity (**Figure 1-5**). Using a neighbour-joining tree, all sequences clustered into two distinct clades, which have historically been referred to as 'DWV' or Type A and 'VDV-1' or Type B, but the two clades are in fact closely related, indicating a strong similarity between all DWV variants. Furthermore, and as elaborated upon in section 1.3.4, several laboratories have identified genetic recombinants between Type A and B DWV with virulent phenotypes, all of which cluster tightly with the Type B variants during phylogenetic analysis.

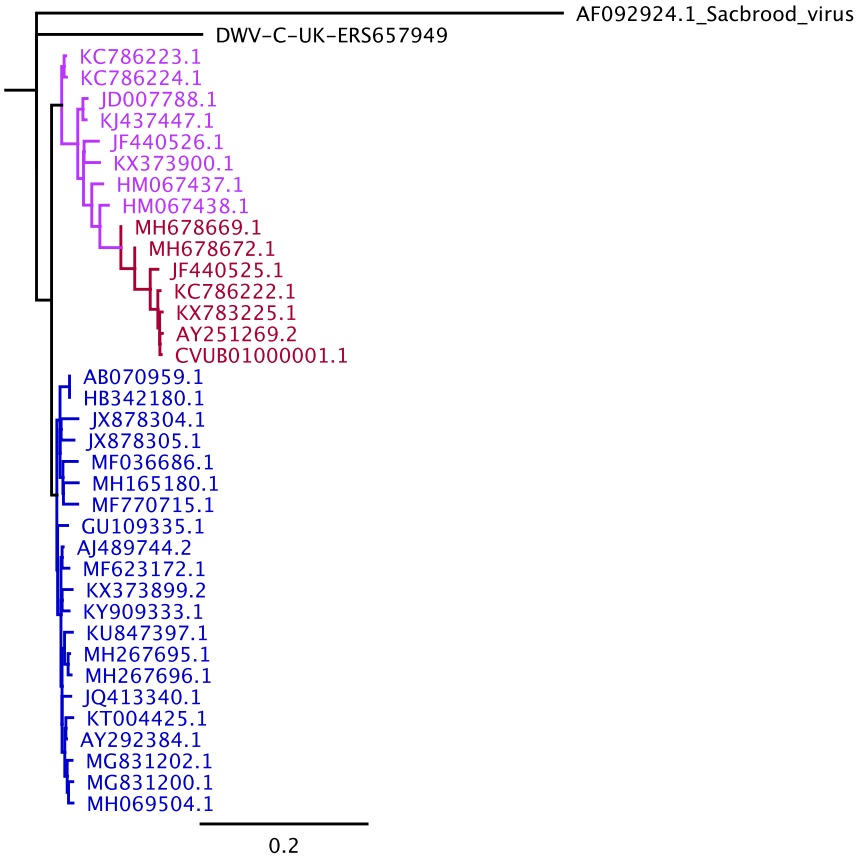


Figure 1-5 – Neighbour-joining tree of all DWV sequences >9500bp available from the NCBI database. *Red sequences are classified on the database as Type B or VDV-1. Pink sequences are classified as recombinant forms of DWV and blue sequences are classified as Type A or 'DWV'. The Type B sequences have clustered at the bottom of the red branch (sequences MH678669.1 – CVUB01000001.1), whilst the genomes showing closer sequence homology to the Type A sequences are reported as recombinants (for example, KC786223.1). The only complete Type C sequence is included and a complete sequence of Sacbrood virus is used as an outlier. 500 bootstrap iterations were run to generate the final tree.*

1.3.4 Deformed wing virus variants, co-infections and recombination

For reasons associated with their original isolation and the differences in their distribution and prevalence, many studies have focused on the distinction in infectivity and pathogenesis of DWV Type A and Type B. Natsopoulou *et al.* (2017) observed a strong association between colonies with higher overwinter mortality rates and elevated levels of the Type B strain in the individual workers analysed. This may be related to the reported larger reduction in worker bee longevity when infected with Type B when compared with the Type A (53.5% and 38% respectively) (McMahon *et al.*, 2016) and with the faster replication of this strain of the virus (McMahon *et al.*, 2016), which may explain why Type B establishes in the host population more readily. Research in apiaries in the USA suggests that Type B has emerged more recently, with findings of only 2.7% prevalence in 2010 increasing to 66% in 2016 (Ryabov *et al.*, 2017), although DWV Type A was still the most widespread variant found there (in 89% of samples) and these differences may be dependent on the sensitivity of the assay used.

However, other studies have produced contradictory results. Despite apparently replicating more slowly than Type B, the Type A virus has been reported as the more pathogenic variant in both the UK and USA (Kevill *et al.*, 2019). The global spread of Type A was strongly associated with the spread of *Varroa*, although Type B was not investigated in this study (Wilfert, *et al.*, 2016). Field-isolate stocks of Type A and B DWV have been used to demonstrate that both viruses exhibit similar patterns of infectivity and mortality rates post-injection (of pupae), strongly suggesting that there is little phenotypic difference between these two variants (Tehel *et al.*, 2019). Despite all injected pupae showing high viral titres, only 60% of the eclosing workers had deformed wings and pupal mortality rates were low (~20%). In contrast, field-isolate stock injections in pupae of Type A and Type B to *Varroa*-naïve colonies in Australia demonstrated that Type A initially replicated faster in pupae, but by 48 hours Type B infections had reached higher titres and remained higher for the remainder of the study (Norton *et al.*, 2020).

Recombination during virus replication (discussed further in Chapter 6) has been observed between DWV Type A and Type B strains, and in some instances these recombinant forms have been reported to predominate in a population over the strains they originate from (Moore *et al.*, 2011; Ryabov *et al.*, 2014; McMahon *et al.*, 2016). Understanding the dynamics of recombination in viral co-infections will aid our understanding of recombination in DWV and whether these forms of the virus have greater replicative capacity or higher virulence compared with their parental variants. The predominance of a particular virus population within a *Varroa*-infested colony is a consequence of the original viruses present, the passage history of the population and the acquisition of exogenous viruses from neighbouring hives or, possibly, through environmental contamination. If recombinant viruses are more transmissible, either vertically or horizontally from bee to bee or via *Varroa*, it will inevitably influence the resulting virus population. This is an inherently noisy system in which to reach definitive conclusions about relative virulence of different virus strains.

The differences observed in these studies, which all rely on quantification of viral loads from field samples, may be due to fundamental differences in the characteristics of the variants. Alternatively, the phenotypes may be due to differences in the levels of the variants present, as well as external factors, such as the strain of honey bee infected or the mite infestation rates of the colonies sampled. Additionally, it is often not obvious

whether the differences observed between the variants (for example in distribution or replication rate) result in measurable effects on colony losses or overwintering success. There is therefore a disconnect between our understanding of the prevalence and incidence of DWV and the impact the virus strain has on colony losses. The relationship between DWV variants and the consequences for the colony remains poorly understood and, because of the importance of DWV to overwintering colony losses, deserves further study.

1.3.5 DWV replication and tropism in *Varroa destructor*

A variable that needs to be taken into account is the ability of DWV to replicate in *Varroa*. Recently a recombinant form of DWV was shown to replicate in *Varroa* mites that were kept on an in vitro feed-packet system that contained no honey bee-derived material. All subsequently screened mites contained negative-strand RNA and a proportion of this could be digested with the restriction endonuclease specific for the unique genetic tag engineered into the recombinant DWV genome, indicating viral replication was taking place in the mites (Gusachenko *et al.*, 2020a). Attempts to measure Type A replication in *Varroa* concluded that the mites transmitted the virus in a non-propagative manner, with no evidence of viral replication in the mites observed (Posada-Florez *et al.*, 2019). Based on these findings it can be assumed that viral replication in mites is at least partially responsible for the highly elevated viral loads of some variants found in parasitized pupae and could alter the dynamics of the virus population in pupae if particular variants replicate while others do not.

Where in the mite DWV replication occurs and how this influences the transmissibility of the virus to a new host remains unknown. These factors are known to have a significant influence in certain arboviruses; for example, a major bottleneck in dengue virus infections is observed in the mosquito midgut, where as few as 5-42 founder viruses can initiate infection (Lequime *et al.*, 2016). If there are strain-specific differences in the ability of DWV to replicate in *Varroa,* there are likely to be subtle (if amplification is limited) to profound (if certain strains replicate very well) influences on the DWV variants passed to the next pupa, leading to potential virus population bottleneck events. Furthermore, such differences may inform our understanding of changes of the DWV population at the landscape scale as mites continue to infest colonies (Wilfert *et al.*, 2016).

1.3.6 Changes in virus replication and mite-related bottlenecks

The change in DWV transmission route arising from the introduction of *Varroa* mites to colonies causes a rapid elevation of the levels of virus in parasitised bees (Martin and Gunn, 1999; Martin, 2001). In laboratory studies of individual pupae, the DWV titre increases from ~10^4 to >10^8–10^{13} GE per bee within 24–48 hours of virus acquisition (Gusachenko *et al.*, 2020a). Subsequent analysis of the virus population shows a shift from a low-level, highly diverse population to a very high-level, near-clonal one. It is the latter sort of population that causes symptomatic infections and contributes significantly to colony losses (Chen et al., 2005; Martin et al., 2012; Ryabov et al., 2014). In individual pupae this switch from a low-titre, diverse population to a high-titre, near-clonal population can occur within hours of virus inoculation (Ryabov *et al.*, 2014). However, at the colony level this is recapitulated over months or years, for example following the introduction of *Varroa* to naïve honey bees on Hawaii (Martin *et al.*, 2012). This suggests that this transition in the characteristics of the virus population seen at the

colony/landscape scale reflects events that occur at the level of individual mite-exposed pupae within colonies.

Several studies have demonstrated that different variants predominate in mite-infested colonies; in the USA high levels of Type A variants (Kevill *et al.*, 2019) and VDV-1 (Type B) variants (Ryabov *et al.*, 2017) have been reported and in the UK different Type A/B recombinants (Moore *et al.*, 2011; Ryabov *et al.*, 2014) have been observed to dominate the population after mite transmission. What drives the amplification of particular variants from a mixed inoculum (Ryabov *et al.*, 2014) is unclear. One possibility is that certain viruses within the population contain virulence determinants that enable them to preferentially replicate rapidly after direct inoculation by the mite (Posada-Florez *et al.*, 2019; Gusachenko *et al.*, 2020a), essentially outcompeting other co-inoculated or endogenous viruses. In this scenario it would be expected that there would be some shared molecular characteristics between the viruses that end up dominating the population.

An alternative interpretation is that most or all DWV variants have the ability replicate to very high levels after mite transmission, but that the inoculum acquired or administered by the mite is extremely small. This results in a subset of the virus population having the opportunity to expand after a population bottleneck (**Figure 1-6A**), as observed in some hepatitis C virus infections (Bull *et al.*, 2011).

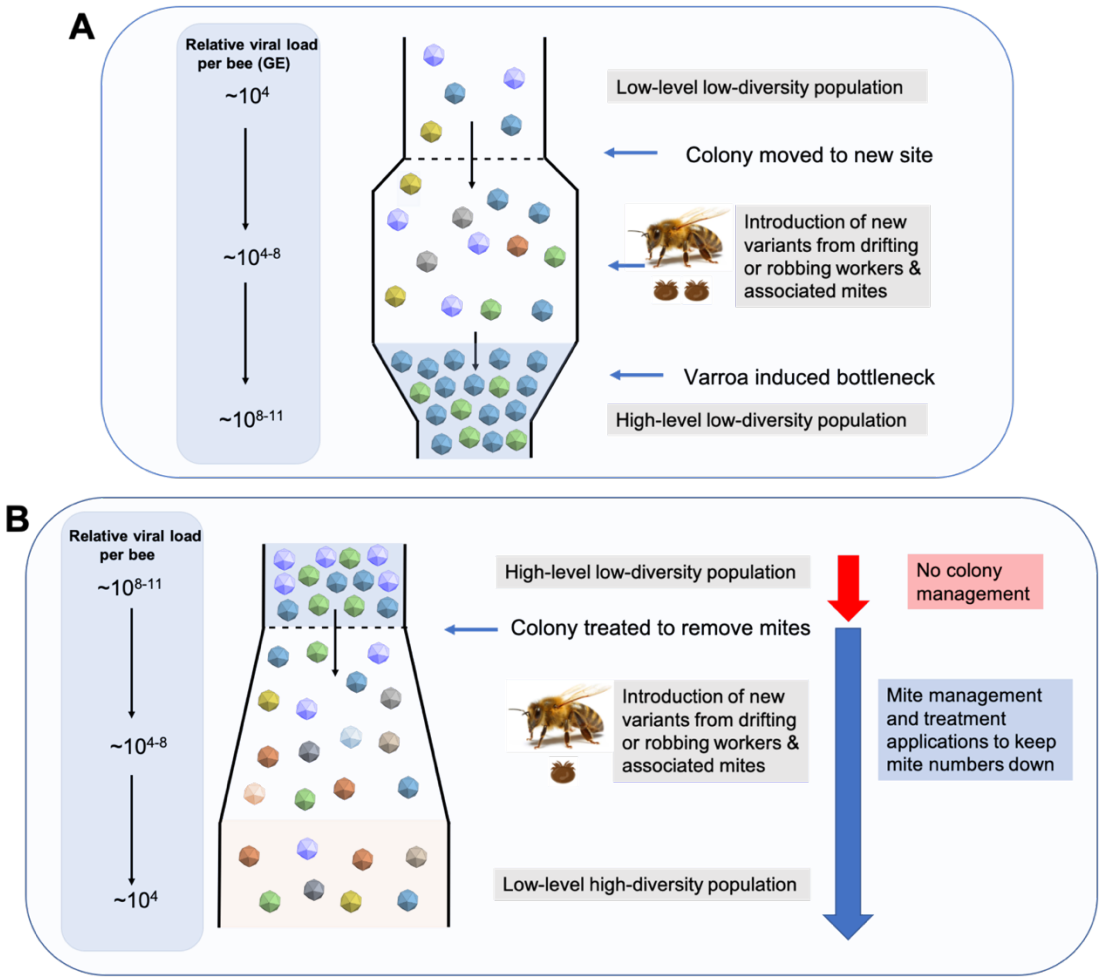


Figure 1-6 - Changes to the Deformed wing virus population in managed honey bee colonies. ***A.*** *The effect on the virus population when an otherwise healthy honey bee colony is placed in close proximity to infested hives. Robbing and drifting bees bring phoretic mites which alter the virus population over time.* ***B.*** *The effects of colony management and mite removal on the virus population. The dashed lines indicate 'hard' changes to the colony, e.g., movement or treatment application.*

Although one or a limited range of variants may be numerically dominant after mite transmission, in recent analyses there remains an underlying diversity in the DWV population in *Varroa*-exposed pupae (Annoscia *et al.*, 2019; Ryabov *et al.*, 2019). This presumably reflects the combination of the pre-*Varroa* viral population overlaid with the *Varroa*-transmitted viruses, some or all of which have undergone a very rapid elevation in titre. Annoscia *et al* (2019) examined virus diversity in honey bees with low and high viral titres after mite-feeding and found that some individual worker bees still contained diverse DWV populations regardless of the viral titre observed. Ryabov *et al* (2019) evaluated DWV infections in honey bee populations in the USA and demonstrated similar levels of population diversity in bees with high- or low-level viral titres. The introduction of *Varroa*-free colonies with low, asymptomatic DWV populations to mite-infested apiaries results in a rapid infestation of those colonies and a shift in viral load and in some instances diversity, but after several months of infestation some viral diversity is still observed in highly infected workers (Martin *et al.*, 2012). Infected neighbouring colonies

continue to display viral diversity at high infection levels despite 5+ years of mite infestation and no mite management techniques.

The diversity in the virus population can be viewed at the scale of the individual honey bee, of the colony, or of the landscape. It is clear from numerous studies that the introduction of *Varroa destructor* to honey bees has fundamentally altered the DWV population at all levels. In the same way that the amplification of a very small initial inoculum from *Varroa* can be considered a bottleneck event that can influence the virus population in the recipient pupa, the removal of *Varroa* from a population could also be considered as a potential bottleneck influencing the resulting DWV population (**Figure 1-6B**).

Used properly, hive management techniques and chemical miticides can significantly reduce the mite population in a colony. Treatments are typically applied at the end of the summer season and in midwinter and under field conditions, reductions of 95% are achievable (Rinkevich, 2020). Since only a subset of mites survive these interventions, the virus population they harbour will have been similarly restricted as host mites perish. This represents a potential bottleneck that will influence the population of viruses that survive and are amplified in the colony the following season (**Figure 1-6B**). There are only limited studies of changes in the virus population following miticide treatments. Locke *et al.*, (2017) demonstrated 1000-fold reductions in DWV levels within the colony after miticide treatment but did not investigate changes in population diversity or predominant genotypes before and after treatment. Their study indicates DWV titres could increase again without colony intervention, possibly due to drifting mites re-infesting the colonies. These mites could be introducing different DWV variants into the virus population, with consequences for the diversity of the population and the dynamics of infection. However, without analysing the make-up of the virus population it is difficult to decipher the causes of these changes.

1.4 Measuring virus population diversity

1.4.1 What causes virus diversity and how can it be measured?

Key factors in a virus's ability to sustain infection in a host are the high mutation rates, large population sizes and relatively short replication times required once a virus has entered a host (Duffy, Shackelton and Holmes, 2008). As a result, viruses are able to coexist in a large pool of highly similar variants often referred to as a virus quasispecies (Domingo *et al.*, 2005). High mutation and selection rates lead to the evolution of diverse RNA viruses, this is primarily driven by a lack of proof-reading capabilities by the RNA polymerase (the RNA dependent RNA polymerase in DWV), leading to high mutation rates (Duffy, Shackelton and Holmes, 2008). As a result of these high error rates, new viral strains are produced during almost every replication cycle by either insertions, deletions or single nucleotide polymorphisms (Posada-Cespedes, Seifert and Beerenwinkel, 2017). Changes in the environment also influence the diversity of a virus population. This selective pressure forces viruses to adapt quickly to an altered or new environment, sometimes resulting in fewer (or only one) dominating viral variants (Bonhoeffer *et al.*, 1997). This has been proposed as a reason for the differences observed in DWV population diversity. Low levels of highly diverse viral variants are found in healthy honey bees, but when parasitized by *Varroa* the passage of viral entry into the host shifts (from oral transmission to entry via the *Varroa* bite-site) and the make-up of

the virus population reflects this, often through highly elevated viral loads and decreased diversity.

The changes and persistence of viral variants could have a major impact on the way a disease progresses, persists and causes pathogenesis (Vignuzzi *et al.*, 2006). Therefore, identifying diversity or changes in a virus population's make-up is of clinical relevance as many infectious diseases that remain a threat to human health are caused by RNA viruses (HIV, Zika virus, dengue virus, influenza and Ebola virus). Therefore, having the appropriate tools to rapidly and accurately identify changes in virus populations is extremely important.

1.4.2 Tools for measuring virus population diversity

Recent advances in next-generation sequencing (NGS) have led to the development of a new suite of tools and opportunities for in-depth virus population analysis. Traditional sequence analysis of viruses was performed using Sanger sequencing; however this method lacks sensitivity and typically generates a consensus sequence of a population, so is unable to determine mutations or low-level variants that could be linked to emerging pathogenesis (Zagordi *et al.*, 2010). To get around this, individual viruses could be cloned and sequenced, but this is labour-intensive and only gives a small subset of the potential variants present in a population. Next-generation sequencing offers increased sensitivity and the ability to produce very large volumes of data relatively easily. This opens up the possibility to study the less abundant variants within a population, which might be lost using traditional methods, and to characterise low-frequency mutations, which could be responsible for virus adaptation (Metzner *et al.*, 2009).

In the context of Deformed wing virus infections, the use of next-generation sequencing tools is key to determining virus population structure as it is hypothesized that asymptomatic viral infections in healthy colonies exist as a diverse mixture of low-level viruses, which are difficult to detect by standard sequencing methods. NGS can be used to obtain sequence data on these variants and determine their sequence similarities, as has been illustrated in several studies of DWV diversity (Moore *et al.*, 2011; Ryabov *et al.*, 2014, 2019; Annoscia *et al.*, 2019). Various studies have used NGS to determine the presence of recombinant forms of DWV, including McMahon *et al* (2016), who used Illumina data to infer the presence of recombinant viruses in the population, but did not determine the sequences of these variants or construct full genome sequences of the variants. However, they did show recombinants with an A-B fusion from 5' to 3' on the viral genome and vice versa, B-A fusion. Typically, the recombination events occurred most frequently around the 5000bp position where the sequence encoding the non-structural proteins begins, as previously suggested (Ryabov *et al.*, 2014), and in the region encoding the RdRp (~8000-9500 bp).

Additionally, in symptomatic infections, where it has been shown that a single near-clonal variant dominates the population (Ryabov *et al.*, 2014), NGS methods could be used to prove this, and to determine if a quasispecies of low-level minor variants still persists in these populations below the limits of detection by standard molecular biology techniques, as suggested in some recent work (Annoscia *et al.*, 2019; Ryabov *et al.*, 2019). There are technical difficulties with this, however, as many of the low-level variants present in virus populations contain single-nucleotide variants (SNVs), which exist at very low frequencies, and the technical error inherent in some NGS platforms remains high,

leading to an inability to distinguish between genuine sequence changes in the virus and sequencing noise.

Analysis software that can read Illumina short-read datasets of mixed populations and infer haplotype frequency has been developed and is of benefit to studies such as those mentioned above. ShoRAH (Short Read Assembly into Haplotypes) is software that uses cluster sizes of reads to determine the frequency with which each variant is present in the virus population (Zagordi *et al.*, 2010; McElroy *et al.*, 2013). ShoRAH can accurately reconstruct local haplotypes by running multiple iterations where all sequences within a window of the genome are assigned with a certain probability to a particular cluster. This can then be used to determine how many unique clusters exist within that window of the genome as well as an estimate of the approximate percentage of the population that they make up. The clusters can be displayed as a list of the potential viral variants in the population, which can indicate whether a population shows high diversity or if a dominant virus is present with multiple low-level variants still present in the sample.

1.5 Aims and objectives

The primary objective of this PhD was to contribute to improving honey bee health by implementing effective control methods for *Varroa destructor* infestation on an apiary scale and a landscape scale. Specifically, this was done by analysing the influence of coordinated, rational *Varroa* control on DWV levels and diversity within colonies and by measuring changes to virus diversity and levels over time when colonies are treated or left to become infested. Additionally, a model to evaluate the dynamic changes in the virus population across multiple generations of pupae and mites was generated and an examination of the formation of recombinant forms of DWV was examined using infectious clones of known variants. The purpose of these experiments was to highlight the detrimental impact of *Varroa* and DWV on untreated honey bee colonies and to determine whether there are significant disease control impacts on beekeeping from using coordinated control methods.

The aims of the project were as follows:

1. To significantly reduce or eradicate *Varroa* from a geographically isolated environment and measure the changes in colony health using DWV levels and diversity.
2. To measure both the rate of mite infestation in low-*Varroa* colonies when placed in a highly infested apiary and the changes in DWV levels and diversity.
3. To ‘rescue’ highly infested colonies using management methods and miticides before measuring changes in the colony health post-treatment.
4. To design and optimise a model to measure DWV diversity changes over time in parasitised honey bees.
5. To examine recombination between infectious clones of DWV variants using next-generation sequencing technologies.

Published manuscripts from parts of this thesis are available in **Appendices 8-7** and **8-8**. Additional projects were carried out throughout the course of this PhD, some of which feed into the experiments presented here, including the use of infectious clones of DWV variants. These works have been published and the full manuscripts are presented in **Appendices 8-9** and **8-10**.

2 Materials and Methods

"We demand rigidly defined areas of doubt and uncertainty!"
— Douglas Adams, The Hitchhiker's Guide to the Galaxy.

2.1 Fieldwork and sample collection

2.1.1 Sampling and set-up for colony exchange experiments

The colony exchange experiments (Chapter 3) were run in conjunction with the Bowman lab at the University of Aberdeen. Colonies were placed in an apiary in Newburgh, north of Aberdeen, where the group works with highly mite-infested colonies. Basic maintenance of the colonies placed on the Newburgh site was carried out by the Bowman group throughout winter. The site is never treated with miticides and the large number of colonies on site (15–20) are maintained through hive management techniques. The site is a large field with hedgerow shielding the north side, which most colonies line in a u-shape. No routine mite counts are performed as the group harvests many of the mites from the colonies throughout the season for experimental work.

Each time the colonies were sampled, a small square of honeycomb (approximately 3cm × 3cm) containing late-stage pupae was cut from the comb and placed in a 30°C incubator until all worker bees had emerged. The workers were then snap-frozen using liquid nitrogen as they emerged from the comb and stored at -80°C until further analysis. The experimental set-ups, treatment regimes, sampling and differences year-on-year are described in more detail in Chapter 3.

2.1.2 Sampling honey bees on the Isle of Arran

To determine the viral load at each site on the Isle of Arran (Chapter 4), adult worker bees were sampled from every colony on each visit. Unlike the colony exchange projects, in this study cutting comb and sampling emerging brood was not possible due to the number of colonies sampled, the logistics of storing the comb after sampling and the disruption it would cause to the colonies, all of which belonged to local amateur beekeepers.

Samples were collected in August 2017, May 2018, May 2019 and August 2019. On each occasion small tubs with a ~1 cm^3 of fondant were prepared and ~30–50 adult bees were scooped into the tubs directly from the frames, causing the minimal amount of disruption to the colonies (**Figure 2-1**). Where possible, nurse bees were selected to try to maximise the number of bees collected that belonged to the colony in question, rather than drifting or robbing bees from neighbouring colonies. The tubs were sealed with tape and labelled with the hive ID and stored in a cool dry place until back in the laboratory, where all bees were snap-frozen and stored at -80°C. All honey bee samples were stored as per section 2.2.1.

Figure 2-1 - Sampling adult worker bees on the Isle of Arran. *A - Small plastic tubs with holes in lid and small block of fondant for feeding collected workers; B - adults scooped off frames of colonies; C - sealed and labelled accordingly. Photos taken by Fiona Highet during fieldwork and used with permission.*

2.1.3 Treating colonies on the Isle of Arran

Miticide treatments on the Isle of Arran were carried out after beekeepers had harvested their honey stocks but before the colonies reduced in size for winter survival. This typically falls between late August and September in Scotland. The beekeepers were supplied with Apivar (supplied by Veto-pharma): two plastic strips, 210 × 40 mm wide, which slot between the comb. The strips are coated with Amitraz, a chemical which paralyses *Varroa* and causes them to drop off the comb. All beekeepers placed the strips in their colonies at the same time and left them for the stated dose exposure time of a minimum of six weeks.

As the treatments were placed in each hive a mite collection tray was placed on the floor under the area which contained the largest number of bees. These trays consist of a paper envelope and a plastic mesh net (**Figure 2-2**). As the mites were exposed to the miticide they dropped off and collected in the tray on the hive floor. After one week of treatment the trays were removed, sealed up and collected. The mites were counted and recorded for each hive. Although this does not account for all the mites present in the colony it allows for a rough approximation of the total mite numbers.

Figure 2-2 – Mesh net and envelopes for catching mites. *The packet is placed on the floor of the hive during treatment application, then removed and carefully sealed in the brown envelope.*

Over winter *Varroa* are phoretic and live on the backs of adult bees while the colony is broodless. This creates an opportunity to treat the colonies and expose ~100% of the mites to a miticide. On Arran, we encouraged the beekeepers to use a mid-winter treatment of Api-Bioxal (an oxalic acid-based chemical treatment). As the bees do not fly in mid-winter the treatments did not need to be coordinated and beekeepers could choose the most appropriate time for treatment to be applied to their colonies.

Batches of the treatment in doses sufficient for 11 colonies were prepared by weighing 25 g of Api-Bioxal powder and mixing it with 490 ml of 1:1 sugar solution. Batches of 11 were made up based on the proximity of hives to one another, meaning beekeepers with hives close together could share the solutions. These were kept in the dark at 4°C until administered to the colonies. Before application the solution was warmed as the Api-Bioxal needs to be between 20-25°C when used. During a spell of moderately warmer weather (~10°C) after a prolonged spell of cold weather (several days of <5°C) the Api-Bioxal solution was dribbled between the frames of each hive where the bees had clustered for winter. The hives were sealed up again and left.

The two treatment regimens were repeated in the autumn and winter of 2018 and 2019, with the *Varroa* drop collected and recorded after each autumn treatment.

2.2 Virus extraction and analysis

2.2.1 Tissue storage and RNA extraction

Honey bees collected from the field or extracted from brood comb were kept in a 30°C incubator (70% humidity). Individual samples were placed in 2ml screwcap tubes with 5–10 ceramic beads and snap-frozen in liquid nitrogen or placed immediately into a -80°C freezer for long term storage.

Individual honey bees were homogenized using the Precellys Evolution tissue homogenizer (Stretton Scientific Ltd, UK) using 3 × 15 seconds at 10,000 rpm cycle with 20 second pauses between each pulse. The supernatant was collected by pulse centrifugation and immediately placed on dry ice before RNA extraction. RNA was extracted using the GeneJET RNA purification kit (Thermo Fisher, UK) as per manufacturer's instructions, with a final elution volume of 50 µl into 1.5 ml Eppendorf tubes. Total RNA was quantified using the Nanodrop-1000 and stored at -80°C.

Pooled honey bees (×30) were homogenized using a mortar and pestle: the pool of bees was placed in the mortar and submerged in liquid nitrogen, then immediately ground up using the pestle until a fine powder. The powder was collected, with a sub-sample placed in a 1.5ml Eppendorf tube and processed using the GeneJET RNA purification kit using the same process used for individual samples. Remaining powder from the pooled bees was stored at -80°C.

2.2.2 Reverse transcription

2.2.2.1 Random oligo cDNA synthesis

Total cDNA was synthesized from extracted RNA using the qScript cDNA synthesis kit (Quanta Biosciences, UK) following the manufacturer's protocol using oligo(dT) and random primers. Reactions were set up using 1µg of RNA, 5× qScript RT reaction mix, 1 µl of qScript reverse transcriptase and made up to 20 µl using nuclease-free water. Reverse transcription incubation conditions were 22°C for 5 minutes, 42°C for 30 minutes and 85°C for 5 minutes. Samples were stored at -20°C until required for further experiments.

2.2.2.2 Strand-specific cDNA synthesis

When a strand-specific cDNA product was required, reverse transcription with Superscript III (Invitrogen, UK) using a gene-specific reverse primer (see **Appendix 1** – primer list) was carried out. The reaction mix was prepared using 1µg RNA, 0.5mM primer and 0.4 mM dNTPs. Reactions were made up to 13 µl and incubated at 65°C for 5 minutes, then placed on ice for 1 minute. 1U of Superscript III, 5x RT buffer and 0.1 M Dithiothreitol (DTT) were added to make the reactions up to 20 µl. Reactions were then incubated at 50°C for 1 hour, 75°C for 15 minutes. Samples were stored at -20°C if not used immediately afterwards.

2.2.3 Primer design

All primers were designed in Snapgene (V4.08) using aligned reference sequences from the NCBI database for the most common variants of DWV and reverse genetics-derived molecular clones (Gusachenko *et al.*, 2020a; Gusachenko *et al.*, 2020b). Primers were selected based on highly variable regions of interest in the genome, such as the leader protein or the RdRp, and conserved sequences within these regions were selected so that the primers would amplify the largest number of DWV variants possible. All primers were synthesized by IDT (USA) and supplied as a 100µM lyophilised stock, which were resuspended in nuclease-free water and diluted to the working concentration of 10 µM. All primers used in this study are listed in **Appendix 1**.

2.2.4 PCR amplification of target DNA

Polymerase chain reaction (PCR) set-up was dependent on the application, but was generally carried out using Hot Start *Taq* polymerase (New England Biolabs, UK). Standard assays were set up in 25 µl reactions containing 10x PCR buffer, 0.2 mM dNTPs, 0.4 µM forward and reverse primer (**Appendix 1**), 1U *Taq* DNA polymerase (New England Biolabs Ltd) and 100 ng of cDNA template. An initial denaturation step at 94°C for 20 seconds was followed by 30 cycles of 94°C for 15 seconds, 57°C for 20 seconds and 68°C for 2 minutes. A final 10-minute extension of 68°C was carried out and reactions were then held at 10°C until further processing.

2.2.5 DNA gel electrophoresis analysis

PCR samples were analysed on a 1% agarose gel stained with ethidium bromide (Thermo Fisher, UK). A 1KB+ DNA ladder (Thermo Fisher, UK) was used to measure fragment size and all samples were mixed with a 6x loading buffer (Thermo Fisher, UK). Agarose gels were made using 1g of Agarose powder (Thermo Fisher, UK) and 100 ml of TAE (Tris-acetate EDTA) buffer. The solution was heated continuously until all powder had fully dissolved, then left to cool for 5 minutes, before 3µl of ethidium bromide (10 mg/ml) was added to stain the gel. Agarose gel results were imaged using a Chemi-doc UV gel imaging system.

2.2.6 qPCR analysis of viral load

To determine the approximate viral load present in honey bees in each study a qPCR assay with known standards was used to determine the DWV copy number present in analysed worker bees. For absolute viral load, qPCRs were performed for DWV using primers which spanned a known region of recombination (see **Appendix 1**) in a C1000 Thermal Cycler (Bio-Rad, UK). Reactions were set up using 1× Luna Universal qRT-PCR master mix (New England Biolabs, UK), 0.5 µM forward and reverse primers and 100 ng of cDNA in a final volume of 20 µl. Amplification was performed using the following thermal profile: 1 minute at 95°C, followed by 40 cycles of 15 seconds at 95°C and 30 seconds at 60°C. Post amplification melting curve analysis was used to check for non-specific amplification (60–95°C with 5 second measurements). Negative template controls and positive standards were included in each run.

The positive DWV standards were prepared by measuring the concentration of a cDNA standard or reverse genetics-derived clone (Gusachenko *et al.*, 2020a) using a Nanodrop 1000 spectrophotometer (Thermo Fisher, UK). Copy number was then calculated using the formula:

$$\textit{Copy number} = \frac{\textit{DNA concentration (ng/µL)} \times 6.02 \times 10^{23}\ \textit{(copies/mol)}}{\textit{length (bp)} \times 6.6 \times 10^{11}\ \textit{(ng/mol)}}$$

In which 6.6×10^{11} ng/mol is the average molecular mass of one base pair and 6.022×10^{23} copies/mol is Avogadro's constant. Linear standard curves were then generated using target plasmid DNA of 10^3–10^9 copy numbers per reaction, spanning the range of expected concentrations of the majority of infected honey bees (some highly infected bees can exceed this titre). This dilution series was then used to calculate the copy number of DWV per 1 µg of RNA for each sample analysed and could also be used to calculate the approximate total virus per bee.

2.2.7 DWV negative strand detection and restriction digest

To test for DWV replication in the host, detection of the negative-sense RNA strand is a standard technique which indicates that the positive-sense RNA virus is replicating. When working with infectious clones this method was applied to determine if inoculated stocks were replicating in the host.

Gene specific cDNA synthesis using Superscript III (Fisher Scientific, UK) was carried out using a negative strand specific primer (see **Appendix 8-1**). The reaction mix was prepared using 1 µg cDNA, 0.5 mM primer and 0.4 mM dNTPs. Reactions were made up to 13 µl and incubated at 65°C for 5 minutes, then placed on ice for 1 minute. 1 U of Superscript III (Invitrogen, UK), 5× RT buffer and 0.1M Dithiothreitol (DTT) were added, and the reactions were made up to 20 µl. Reactions were incubated at 50°C for 1 hour, 75°C for 15 minutes. The qRT-PCR set-up was performed using the Luna Universal qPCR master mix (New England Biolabs, UK) following the method in 2.2.6.

If the target was a cDNA generated from a reverse genetic clone with silent restriction sites inserted in the genome, an additional screening using restriction digest could be used to distinguish the tagged virus from the wild-type population of DWV. Samples were digested with the restriction enzymes unique to the reverse genetic constructs (Gusachenko *et al.*, 2020a), to ensure the negative strand detected was from the injected viral stocks. Typically, 1 µl enzyme was added to 500 ng product and 5 x buffer and incubated at 37°C for 1 hour before the digested product was visualised on a 1% agarose gel to check for the appropriately sized bands.

2.2.8 Preparing new plasmid stocks from *E. coli*

E. coli competent cells transformed with DNA plasmids containing the two reverse genetic constructs (VVD-4 and VVD-5) were used throughout this project. VVD-4 contains a *HpaI* silent restriction site (SRS) and has been transformed in DH-5α *E. coli* cells. VVD-5 contains a *SalI* SRS and has been transformed in DH-10β *E. coli* cells. NEB DH-5α and DH-10β *E. coli* cells can be grown on 30 µg/ml Kanamycin selection. 60 µl of a 50 mg/ml Kanamycin stock was added to 100 ml of LB broth to give a final concentration of 30 µg/ml. Starter cultures of 5 ml LB + Kanamycin were setup with 50 µl of each *E. coli*/plasmid stock. These were left at 37°C for several hours and then split into 5x 10 ml stocks, each containing 1ml of the original started culture. The 10 ml stocks were left overnight at 37°C.

500 ul of culture was mixed with 500 µl 50% glycerol and stored at -80°C for both stocks. 4.5 ml of each flask was used for mini-prep extraction of the plasmid using the Fisher Genejet miniprep kit, as per manufacturer's protocol. The final eluate of the different samples was pooled together and quantified using a Nanodrop-1000.

2.2.9 Generating virus stocks from infectious cDNAs

Viral clones of the major DWV variants (Type A – VDD, Type B – VVV and a recombinant – VVD) (Gusachenko *et al.*, 2020a; Gusachenko *et al.*, 2020b) were used in various experiments throughout this study as internal controls for PCR, qPCR or injection experiments. The clones are stored as plasmid DNA and new stocks had to be generated when required.

Plasmid DNA was transformed in *E. coli* as per 2.2.8 to increase the stock concentrations. The plasmids were checked by double restriction digest to ensure they contained the expected inserted DWV genomes. The extracted plasmid DNA was then linearised using

PmeI which linearises the plasmid by cutting near the 3'-end of the viral insert. The samples were digested for 90 minutes at 37°C and inactivated at 65°C for 20 minutes.

RNA was transcribed from the linearised cDNA insert using a RiboMax (Promega, UK) kit as per manufacturer's instructions. The samples were incubated for 4 hours at 37°C and then treated with DNase for 20 minutes at 37°C. RNA extractions of the treated samples were then carried out using the Genejet RNA extraction kit (Thermo Fisher, UK) as per manufacturer's instructions.

The extracted RNA was then used to generate cDNA standards for qPCR as per section 2.2.1 where 1 µg RNA was used to generate approximately 1 µg cDNA. RNA was injected into honey bee pupae (~2/3 µg RNA total) using an insulin syringe and a 30 G needle. Bees were placed at 30°C for three days and then harvested as per 2.2.1. The homogenized sample was mixed with equal volume PBS, thoroughly vortexed and centrifuged for 10 minutes at 4°C. The supernatant was removed and passed through a 0.22-micron filter. The stock was then treated with RNAse for 10 minutes at room temperature before being aliquoted and snap-frozen at -80°C.

2.2.10 Statistical analysis and modelling

The statistical significance of DWV titre changes in Chapter 3 and Chapter 4 was determined using a Mann-Whitney test to calculate the statistical significance of the quantified DWV yields at different time points in the study. As the data is not normally distributed (log differences between workers) a non-parametric test was performed for each colony (Mann-Whitney, instead of a t-test). Multiple comparisons of DWV titres in Chapter 4 were carried out using a Kruskal-Wallis one-way test for variance. Wilcoxon t-tests were used in Chapter 4 to determine changes in *Varroa* over time. All statistical tests were carried out using Graphpad Prism (V9).

2.3 Sequencing preparation and analysis

2.3.1 Standard Sanger sequencing

PCR products were column purified using a Wizard SV gel and clean up kit (Promega) following the manufacturer's recommended protocol. Samples were eluted on the column in 20 µl of nuclease-free water and the concentration was determined by Nanodrop-1000 measurement.

Where appropriate, positive samples were prepared for Sanger sequencing in a 10 µl total volume, using a 5 mM primer stock and 20–100 ng of purified PCR product, or 50–200 ng of purified plasmid DNA. Samples were diluted with molecular grade water if required. All Sanger sequencing was performed by GATC in Germany.

2.3.2 Analysis of virus diversity using SangerseqR

Due to the large number of samples to be analysed in this study and the prohibitive cost of next-generation sequencing methods, a cost-effective and reliable alternative was investigated. Using chromatograms obtained from Sanger sequences of PCR products, a

binary analysis of virus population diversity was achieved using an R studio package called SangerseqR (Hill *et al.*, 2014; Hill, 2017). A chromatogram where each nucleotide peak is well defined would indicate a clonal population of a single DWV variant. Mixed nucleotide peaks, where more than one base is called in a single position in the sequence, would potentially indicate a mixed population of viral variants within a sample (**Figure 2-3**).

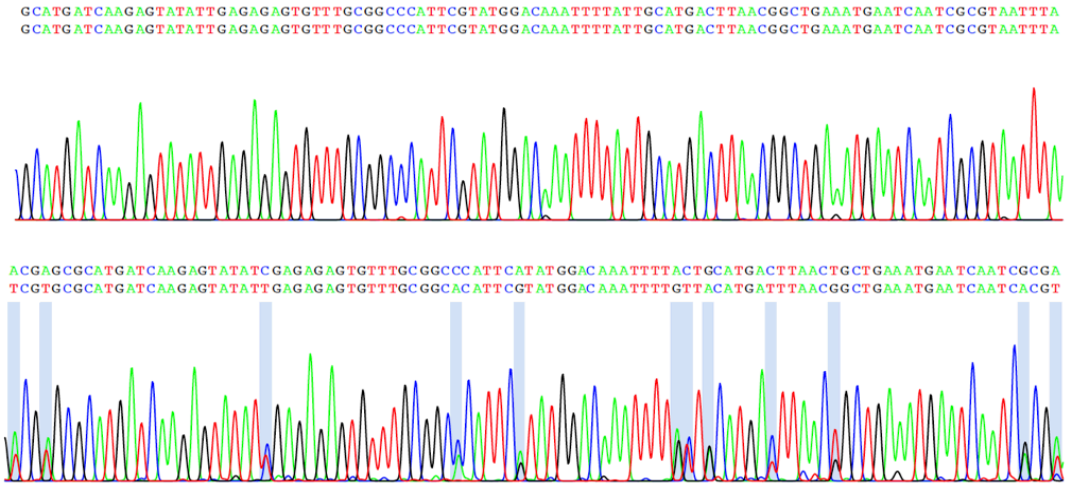


Figure 2-3 – Sanger sequence chromatograms for a clonal sample (top) and a mixed sample (bottom). *Using the R program SangerseqR two sequences can be extracted from ab1 Sanger sequence files, a 'top' match and a 'second' match. Ambiguous base positions are highlighted in navy blue on the lower sequence strand, indicating two or more possible sequence variants are present in the sample. In the top sequence the chromatogram is clean with each position clearly showing a single base, indicating no diversity in the sample.*

The SangerseqR software calls ambiguous bases based on the percentage of the peak for the secondary base compared to the primary basecall at a given position in the sequence and in this context can be used to determine whether a virus population is clonal or diverse. The degree of diversity within the sample cannot be established by this process, but it is a binary indicator of potential diversity. The R script produces an output of a chromatogram with each ambiguous base position highlighted in blue (**Figure 2-3**) and a .csv file with the relative peak intensity of each base call.

As part of the output, the package produces two DNA sequences, one each for the primary and secondary basecalls, which can then be compared using an NCBI BLAST search. Samples which show no diversity produce two identical sequences, whilst diverse samples produce two varying sequences when compared by NCBI BLAST. Sequences that produced poorly resolved peaks at the beginning and end of sequence reads had to be trimmed before being processed. Obtained sequences could then be used to compare diversity across different honey bee populations in a cost-effective and rapid manner.

2.3.3 Validation of SangerseqR methodology

Before using the SangerseqR package on experimental data, a test experiment was performed where two DWV clones were spiked into samples at known ratios and

sequenced to determine the relative sensitivity of the method for determining diversity in samples.

To test this method, two control constructs that have a unique silent restriction site (SRS) in different places in the genome were used (**Table 2-1**). A *HpaI* (5276bp) and *SalI* (5143bp) site is present in two otherwise unique VVD-9 reverse genetic construct sequences.

Table 2-1 - Restriction site differences between the two DWV infectious clones.

Template	***SalI***	***HpaI***
VVD9-4_2U	GTAGAT (no *SalI*)	GTTAAC
VVD9-5	GTCGAC	GTGAAT (no *HpaI*)

Purified plasmid DNA, prepared as per 2.2.8, was amplified by PCR using primers that flanked the SRS (see primers SRS FP/RP in **Appendix 1**). Each template was amplified as per section 2.2.4, using 10ng of plasmid DNA per reaction. Samples were checked by gel electrophoresis (**Figure 2-4**).

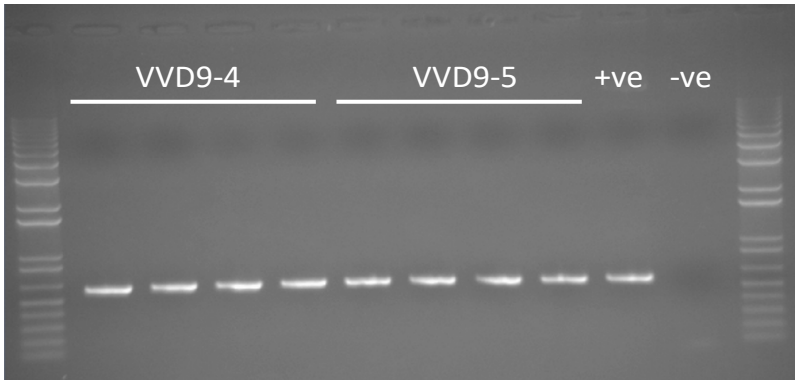


Figure 2-4 - Gel electrophoresis analysis of purified PCR fragments*. +ve and -ve indicate positive and negative controls, DNA ladder is a 1KB+ marker.*

PCR reactions were purified using the Promega Wizard SV clean-up kit (Promega, UK). Samples were pooled prior to purification and each sample was then quantified by Nanodrop1000 (VVD9-4 – 80.8 ng/µl / VVD9-5 – 73.2 ng/µl). Following purification, ratios of the products were mixed together as per **Table 2-2**.

Table 2-2 - Ratios of the two purified samples used for SangerseqR analysis. *In each ratio, VVD9-4 is the 1.*

Ratio	**Concentration ratio**	**Volume added (based on plasmid conc.)**
1:1	200 ng / 200 ng	2.5 µl / 2.7 µl
1:2	130 ng / 260 ng	1.6 µl / 3.6 µl
1:5	66 ng / 333 ng	1 µl / 4.4 µl
1:10	40 ng / 400 ng	0.5 µl / 5.0 µl

Samples were analysed using SangerseqR in R (See **Appendix 2** for the annotated R script) to determine the difference between the respective chromatogram peaks for the base changes between VVD9-5 and VVD9-4. The values reported at each nucleotide indicate the intensity of that particular nucleotide in the sample mix (**Table 2-3**). These were then used to calculate a % of each called base and plotted accordingly.

Table 2-3 - Peak intensity at each unique base position. *SalI site (GTAGAT or GTCGAC) and HpaI site (GTGAAT or GTTAAC).*

	Base call in 1:1	**Base call in 1:2**	**Base call in 1:5**	**Base call in 1:10**
SalI (A / C)	A = 286 (43%) C = 373 (57%)	A = 172 (20%) C = 665 (80%)	A = 88 (9%) C = 896 (91%)	A = 22 (3%) C = 779 (97%)
SalI (T / C)	T = 153 (36%) C = 267 (67%)	T = 72 (13%) C = 468 (87%)	T = 28 (4%) C = 644 (96%)	T = 0 (0%) C = 560 (100%)
HpaI (G /T)	G = 236 (59%) T = 105 (41%)	G = 141 (75%) T = 48 (25%)	G = 316 (97%) T = 11(3%)	G = 454 (100%) T = 0 (0%)
HpaI (T / C)	T = 186 (51.5%) C = 173 (48.5%)	T = 249 (72%) C = 99 (28%)	T = 283 (84%) C = 54 (16%)	T = 400 (100%) C = 1 (0%)

The calculated % from each single base call was then averaged to give a % for *SalI* and *HpaI* for both targets to show the ratio between VVD9-4 and VVD9-5 (**Figure 2-5**).

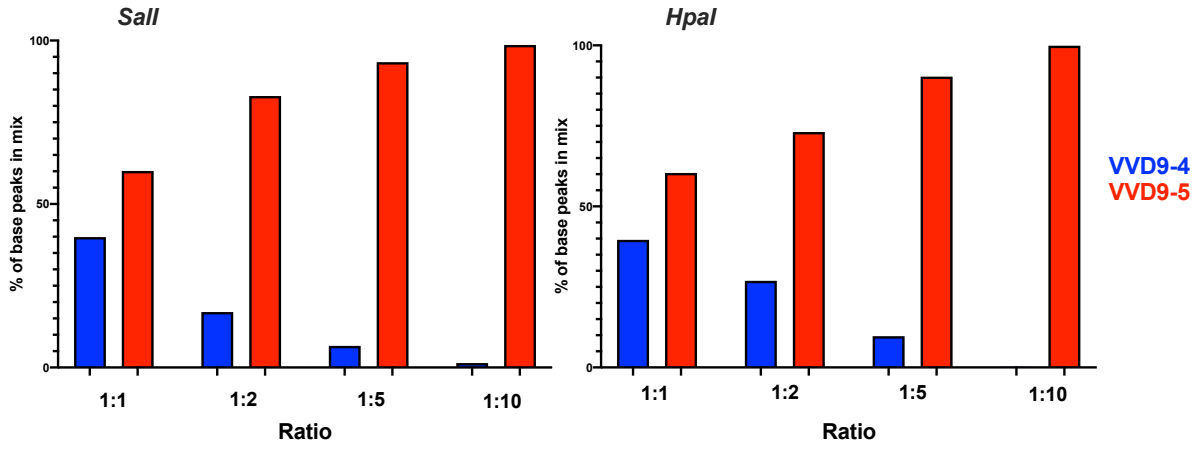


Figure 2-5 - Ratios of base peak calls from mixed sanger sequence samples of VVD9-4 and VVD9-5*. SalI restriction site ratios shown on the left and HpaI ratios on the right.*

The sequence results are not evenly distributed as per the PCR input, but the low-level variant was detectable in the samples up to the 1:5 ratio. Therefore this method was

deemed acceptable for use as an initial binary indicator of virus diversity in samples where we expect to find a mix of equally, or close to equally distributed viruses. When one variant is dominant in a population the trace levels of low variants are only detectable by next-generation sequencing methods.

2.3.4 Preparing amplicons for Illumina analysis

As part of the characterisation of virus diversity, sequences were desired which covered the majority of the DWV genome and amplified all known major variants. Primers were designed to amplify a large proportion of the DWV genome and all of the ORF (111bp - 10086bp based on VVD9-4_2u) selecting conserved regions close to the terminal 5' and 3' ends of the viral genome. Several common DWV variants (NCBI accession codes: KJ437447.1, KU847397.1, KX783225.1, HM067438.1) and DWV infectious clones VVD9-4 and VVD9-5 (Gusachenko, *et al.*, 2020a) were aligned in Snapgene and primer design was based on conservation between these variants.

Strand-specific cDNA synthesis was carried out as per section 2.2.2.2 using primer DWV FG RP1 (10068-10086) (see **Appendix 1** – primers) and 1 µg of total RNA.

Amplicons were generated using LongAmp *taq* polymerase (New England Biolabs, UK). PCR assays were carried out in 25 µl reactions containing 5 × LongAmp reaction buffer, 0.2 mM dNTPs, 0.4 µM forward and reverse primer, 2.5 units LongAmp *Taq* DNA polymerase (New England Biolabs, UK) and 2 µl (100 ng) of cDNA template. PCRs were carried out on a Veriti Proflex. An initial denaturation step of 94°C was applied for 30 seconds, followed by 35 cycles of 94°C for 15 seconds, 53°C for 30 seconds and 65°C for 8 minutes. A final extension of 65°C was run for 10 minutes and all reactions were held at 10 until further processing. Samples were analysed by gel electrophoresis as per section 2.2.5 and column purified using a Promega SV Wizard gel and PCR clean-up kit as per manufacturer's instructions with a final elution of 20 µl. All purified samples were quantified by Nanodrop-1000 and stored at -20°C.

Purified PCR products were processed externally, but tagged libraries were prepared and run on an Illumina Hi-seq. A purified amplicon library of (2x 300 bp) paired-end reads were processed at the Department of Biomedical Sciences, University of St Andrews in 2018 and at the Centre for Genome-Enabled Biology and Medicine at the University of Aberdeen in 2020. Sequence data was generated as Fastq files for each barcoded sample.

2.4 Bioinformatics and analysis pipelines

2.4.1 Analysis of virus diversity using Illumina reads

Illumina sequence reads were stored as paired-end reads in .fastq files labelled as R1 and R2. The barcoded sequences were converted to fasta format, extracted and trimmed using Geneious (v.2019.1.3) (Kearse *et al.*, 2012).

To visualise the coverage of reads across the viral genome for individual samples the software Interactive Genomics Viewer (IGV, version 2.8.2) was used. A reference

genome was indexed using the prompt 'index' with the alignment software Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009):

```
BWA index ./reference_genome.fa
```

In order to process virus population data, the sequences needed to be mapped to the indexed reference genome using BWA and Samtools (v1.10) (Li *et al.*, 2009). This created an indexed binary alignment map (bam file) output using the following command prompt:

```
BWA mem ./reference_genome.fa ./Illumina.reads.fa | samtools sort -o
./Illumina.reads.sorted.bam -T reads.tmp
```

BWA maps the Illumina sequencing reads against a chosen reference genome and the second command, after the |, Samtools sort creates an indexed bam file as the output. This output can be indexed and then visualised in IGV to compare coverage of the sequence reads against the full-length DWV reference genome.

2.4.2 Determining variants in virus populations using ShoRAH

Key to this study was the development of a method to accurately and sensitively measure changes in DWV diversity in individual honey bees. Determining virus population diversity in a single, mixed population sample is a key issue with next-generation sequence data. A series of tools have been designed that can elucidate virus diversity in mixed population samples (Domingo et al., 2012; Posada-Cespedes et al., 2017). One such tool, ShoRAH, was employed as part of this study.

ShoRAH (**Sho**rt **R**ead **A**ssembly into **H**aplotypes) is software developed to determine variants within a virus population, which is also capable of determining the percentage of a given target within the population, correcting for sequence error and producing sequence outputs of each variant identified (Zagordi *et al.*, 2011; McElroy *et al.*, 2013). The software uses a series of Python scripts to analyse a given window of a genome and as such can be used to select known regions of variability or conservation in sequences.

To use ShoRAH analysis for local haplotype variant calling the following basic command is used:

```
shorah.py -b illumina_dataset_sorted.bam -f reference_seq.fa -r SeqID:5000-6000
```

-b calls an indexed bam file of reads
-f calls the reference genome sequence
-r calls the window within the reference sequence to be analysed
- the SeqID is the name in the reference fasta file.

The process is handled by running the shorah.py script, which acts as a wrapper for a series of scripts which carry out alignment, error correction (local haplotype reconstruction), global haplotype reconstruction and SNV frequency estimation. The first of these scripts, diri_sampler.py, estimates the diversity at a local level of the sequence alignment within a given window of the genome. As global haplotype reconstruction is increasingly discouraged when using Illumina short-read sequence data due to the

difficulty and ambiguity of phasing variation at different sites, only local haplotype reconstruction was employed in this study (Zagordi *et al.*, 2011). The diri_sampler.py script uses a model-based probabilistic clustering algorithm to correct for errors and infer haplotypes and their frequencies (Zagordi *et al.*, 2011). It then measures the coverage across the selected window and uses a series of overlapping read windows to determine the diversity within the sample set, with each position in the sequence being covered a minimum of three times under the default settings.

The frequency is estimated using 'FreqEst' and the output produces each of the identified haplotypes and the percentage of the sequence population attributed to them (**Figure 2-6**). For this project, windows of the DWV genome were analysed, including known regions of recombination (Ryabov *et al.*, 2014), the leader protein and the RdRp (see **Figure 1-3** – viral genome) were selected and examined. The leader protein and RdRp were selected as known regions of high variability, while the region in the middle of the genome would show any potential recombination events.

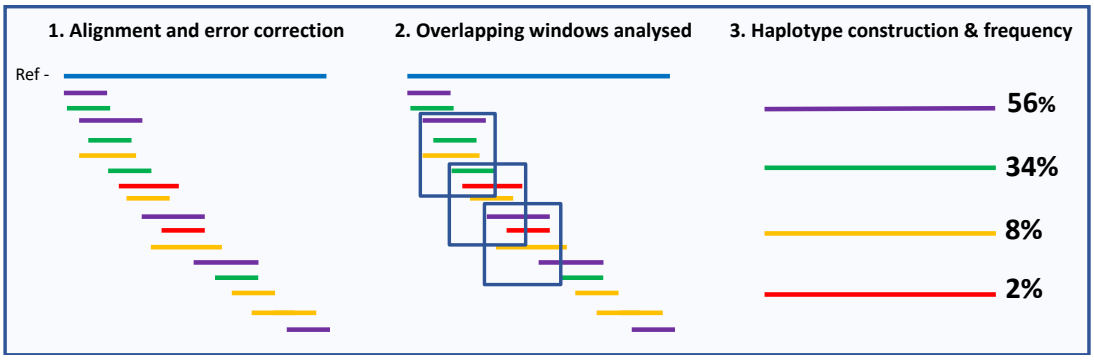


***Figure 2-6 - Schematic of ShoRAH analysis**. Short reads are trimmed and aligned to a reference genome, indicated by the blue line (1), then a region of interest is selected within the genome and overlapping windows are analysed to ensure coverage (2), finally haplotype sequences are constructed based on their frequency within the selected region (3).*

2.4.3 Phylogenetic analysis of population variants

ShoRAH output produced single fasta files containing all possible haplotypes from each sample. Based on a control sample, a VVD9-4 clonal sample amplified and sequenced in the same manner as all test samples, the threshold for viable haplotypes was determined to be >3% of the population match. All haplotypes below this threshold were discarded from further processing and clustered as "other sequences" in further data analysis.

Sequences above the 3% threshold were aligned using MUSCLE (Edgar, 2004) and neighbour-joining phylogenetic trees were generated in Geneious (v.2019.1.3) (Kearse *et al.*, 2012) or Mega (v7.0) (Kumar, Stecher and Tamura, 2016) depending on the project. 1000 bootstrap iterations were run for each neighbour-joining tree and, when using Mega, model selection was always determined before selecting and running the best-fitting model for the phylogeny. Reference sequences for all known major variants and commonly reported recombinant sequences were included in analysis.

2.4.4 SNV calling and sequence data validation

Statistical modelling of the data set was carried out to evaluate the accuracy of the sequence diversity analysis performed using the full genome amplification and ShoRAH analysis protocols. Single nucleotide variants (SNVs) were called using statistical modelling at a frequency lower than the error rate of sequencing. ShoRAH analysis includes a statistical test of strand bias by utilising a Fisher's exact test and a Benjamini-Hochberg correction process, which rejects any p-values <0.05 in the final analysis of SNVs, indicating any base changes with a value lower than 0.05 are error-driven and not the result of a substitution in some of the samples of a population. A Fisher's exact test score close to 1.0 for a sample indicated that the majority of sequences in that sample had undergone that particular base-change compared to the reference genome and as such was an indicator of a dominant viral sequence.

Based on this method, combined with the probabilistic clustering method implemented by ShoRAH, SNVs can be called from deep sequenced viral populations with high accuracy (McElroy *et al.*, 2013). This method was utilised to evaluate the sequence diversity observed in the different studies, including the Arran samples (Chapter 4) and the colony exchange samples (Chapter 3). The SNV values were sorted so that only values recorded in all three iterations of the modelling were included in analysis and any values <0.05 were discounted from the final dataset.

3 Investigating the effect of infestation or removal of *Varroa destructor* from honey bee colonies on the Deformed wing virus population

"The best model of a cat is a cat" – Norbert Weiner

3.1 Introduction and aim

The introduction of *Varroa* to honey bee colonies is responsible for a shift in the dynamics of Deformed wing virus infections, as the mites alter the transmission route, leading to highly elevated, often clonal, viral populations which cause symptomatic infections superseding the low-level diverse viral populations usually passed *per os* from worker to larvae (Martin *et al.*, 2012; Ryabov *et al.*, 2014). Due to the changes caused to DWV transmission and pathogenicity, and the subsequent impact on colony health, *Varroa* is considered a major threat to European honey bee health on a global level (Wilfert *et al.*, 2016). Additionally, it has been shown that symptomatic DWV infections, causing colony death, can still occur after *Varroa* have been removed from a colony (Highfield *et al.*, 2009). Due to the association between *Varroa* and DWV it is believed that controlling *Varroa* infestation could significantly reduce the rate of highly pathogenic DWV accumulation and reduce colony losses.

Phoretic mites are known to spread between colonies on the bodies of worker bees as they drift between hives, meaning that highly infested hives can transfer mites to neighbouring colonies, even if they have been treated (Fries and Camazine, 2001). Modelled data has suggested as much as 42% of workers in a colony could be alien due to drifting when colonies are placed alongside each other in close proximity (Pfeiffer and Crailsheim, 1998). These high drifting rates also result in increased mite uptake, with re-infestation rates of ~76 mites a day recorded in mite-free colonies over 200m away from the nearest infested hives (Greatti, Milani and Nazzi, 1992). Drifting rates between colonies in close proximity have been shown to be as high as 30% for foraging workers (Forfert *et al.*, 2015) and mite uptake has been observed to occur rapidly and consistently over anything up to a 100m distance in some apiaries (Nolan and Delaplane, 2016).Weak colonies, potentially debilitated by high levels of mite infestation, may be robbed by neighbouring strong colonies who, in turn, acquire an increased level of mites and the novel virus populations they harbour (Peck and Seeley, 2019). Mite transmission is further exacerbated by colonies with high *Varroa* infestation showing an increased acceptance of drifting workers, allowing uptake of other associated pests and pathogens to occur (Forfert *et al.*, 2015).

A range of miticide treatments are available to combat *Varroa* infestation, but if they aren't used correctly or in unison with other neighbouring colonies, infestations can persist through drifting or robbing (see section 1.2.2). The use of Apistan to treat mite-infested colonies with high DWV load resulted in a 1000-fold reduction in DWV titres compared with neighbouring untreated colonies (Locke *et al.*, 2017), but DWV persisted and even increased again by the end of the study, highlighting the need to continuously monitor and treat colonies where appropriate. As DWV changes manifest through changes in *Varroa* infestations, our understanding of these changes over time is critical

for improving colony health and emphasizes the importance of developing effective and affordable means to control *Varroa*. As discussed in section 1.3.4, DWV exists as two commonly reported variants, Type A and Type B, with some dispute about differences in infectivity, replication in mites and competition during co-infection (Mordecai, Brettell, *et al.*, 2015; McMahon *et al.*, 2016; Tehel *et al.*, 2019; Gisder and Genersch, 2020; Gusachenko *et al.*, 2020a). These differences may prove significant for the population dynamics of DWV and determine which viruses become dominant as the population shifts over time.

The aims of this chapter were to measure changes in the DWV populations in two unique colony monitoring experiments. Firstly, the impact on healthy *Varroa*-naïve colonies when they are placed near highly mite-infested colonies was measured by studying changes in virus titre and diversity over time. It was hypothesised that mites would quickly begin to accumulate in these colonies, entering with adult workers from neighbouring colonies through known processes like drifting and robbing, and establishing in the brood of the naïve colonies. It was hypothesized that this would subsequently cause DWV titres in the developing brood of the naïve colonies to increase significantly, resulting in colony losses (Martin, 2001). Samples were collected monthly throughout each season and the viral load of emerging brood was quantified from each colony. NGS methods were used to analyse changes in the virus populations over time.

Secondly, the application of a 'shook swarm' beekeeping technique, combined with effective miticide treatments, was investigated to measure whether it effectively reduced mite infestations and as a result reduced levels of Deformed wing virus to those of 'healthy' colonies, and how quickly this took place after treatments. As most mites are found in brood cells (~80% depending on the time of year), a shook swarm technique was considered the most efficient way to remove mites from the colonies as it involves discarding all brood comb. The changes in virus level and diversity were examined before the treatment method was applied and in subsequent brood cycles afterwards, extending to the end of each "bee season" to determine whether the viral load decreased significantly and remained low post-treatment.

3.2 Healthy colonies exposed to *Varroa* in a mite-infested apiary: how does this impact the colonies' DWV population?

3.2.1 Colony preparation and virus screening

Four packages of European honey bees (*Apis mellifera*) were purchased from an importer (original source Poland) and placed in new Abelo national hives at the University of St Andrews in May 2017. Eleven frames of foundation were placed in each hive and 1/1 sugar solution was provided to encourage the bees to quickly draw up the frames, feed, create stores and allow the queen to start laying. The hives were left for at least 1 month to allow the colonies to build up strength. In 2018 bees were purchased from Italy but prepared in the same manner as 2017, and in 2019 colonies from the St Andrews apiary were selected for use.

Once ready, the colonies were treated with an appropriate miticide to remove any *Varroa* present. Honey bee colonies were required to be mite-free, or as close as possible, before being deemed acceptable for this experiment. Therefore, in 2017 (the first year of the

study), selected colonies were treated three times with Api-Bioxal (an oxalic acid-based product) by sublimation over the course of one brood cycle. This ensured any mites contained in capped cells with brood should have emerged and been exposed during at least one of the treatment applications. 1.6g of Api-Bioxal was administered to each hive and the vaporizer was left in place until the oxalic acid vapor had clearly dispersed through the hive (vapor is partially visible from the gaps in the hive sides, despite the entrance being sealed up). The number of *Varroa* present in each hive was measured 24 hours after treatments by recording the number of dead mites found on the collection tray on the floor of each hive (**Figure 3-1**). The first dose of Api-Bioxal treatment killed 60+ mites in three of the colonies in the first 24 hours. Dead mite levels were very low after the two subsequent treatments in all remaining counts and by the final treatment all colonies were close to mite-free.

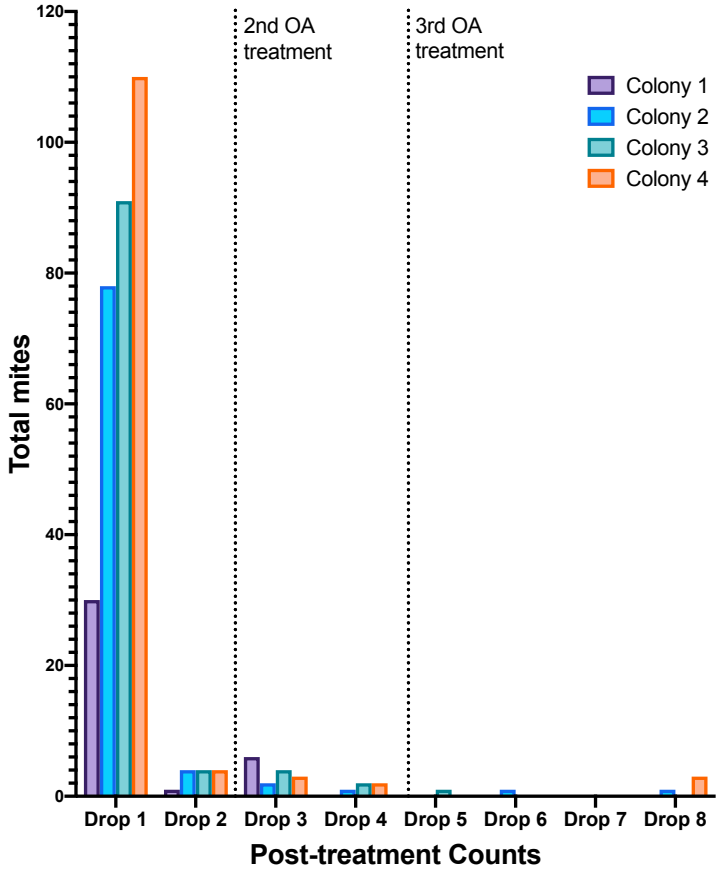


Figure 3-1 – Mite drop after three administrations of Api-Bioxal by sublimation to four colonies in the St Andrews apiary. *The treatments were applied 3 times over 18 days and mite drops were recorded 24 hours after each treatment and immediately before the next treatment. Dotted vertical lines indicate the application of the second and third doses of Api-Bioxal.*

In the second year of the study, the colonies were given a 6-week dose of Apivar, a plastic strip coated in 500mg of Amitraz, which splits in two and is placed in the hives two or three frames from the ends to maximise chemical distribution. The mite drop was very low post-treatment for all hives (data not shown).

In the third year, seven colonies in the St Andrews apiary were screened for DWV and four with very low viral titres were selected. All colonies had received Apivar treatment

the previous autumn and oxalic acid sublimation treatments during mid-winter so were not treated again prior to being moved to Aberdeen. No mites were observed in the colonies or on the floors during the testing process and the DWV titre was sufficiently low (<10^6 GE/µg RNA) in all colonies screened (**Figure 3-2**). As all the colonies screened had DWV titres which fell within the range of a 'healthy' colony, they were selected based on their strength and temperament with colonies 27 and 34 being larger/calmer than 8 and 42.

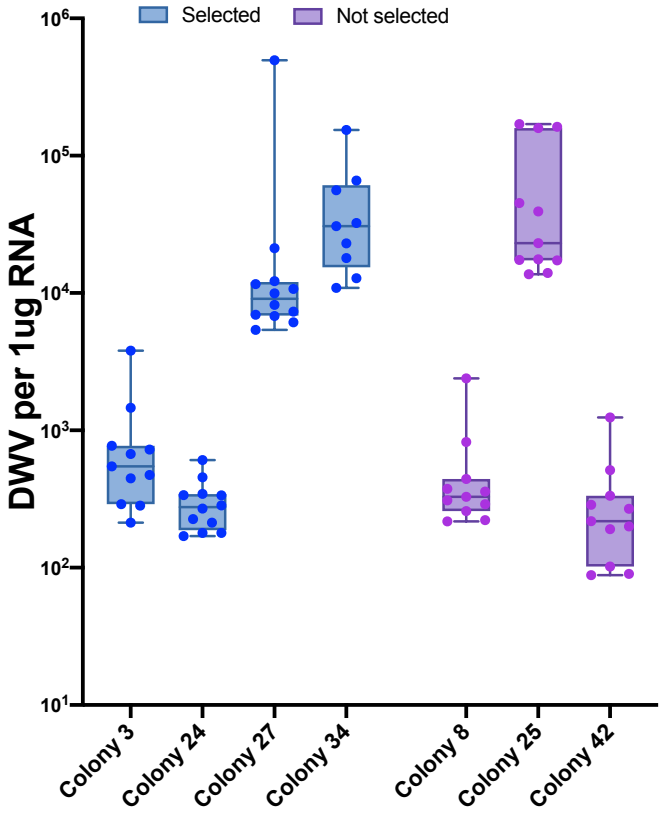


Figure 3-2 - Initial screening of St Andrews apiary colonies for DWV titre. *Seven colonies in the apiary were screened for DWV in May 2019 and four (in blue) were selected for use in this experiment. A minimum of 10 emerging brood were quantified from each colony by qPCR.*

3.2.2 Apiary positions and sampling methods

Hives were sealed and transported to an infested apiary containing 15–20 colonies with known 'high' *Varroa* mite abundance, where colonies are never treated with any miticides. The site was located north of Aberdeen (approximate coordinates 57.3°, -2.0°). The transported colonies were placed in the centre of the infested apiary in the first two years, approximately 10 metres from all the neighbouring colonies (**Figure 3-3A**) and distributed amongst the existing colonies in the third year (**Figure 3-3B**). The positioning of colonies in the third year differed as the four hives from 2018 remained on a pallet in the middle of the apiary, therefore the 2019 colonies were placed on the hive stands amongst the infested colonies (**Figure 3-3B)**. The different positioning of the colonies to the infested Aberdeen colonies should not significantly alter the mite uptake between

seasons as all colonies were within 10 metres of the infested colonies and thus prone to robbing and drifting at a high rate.

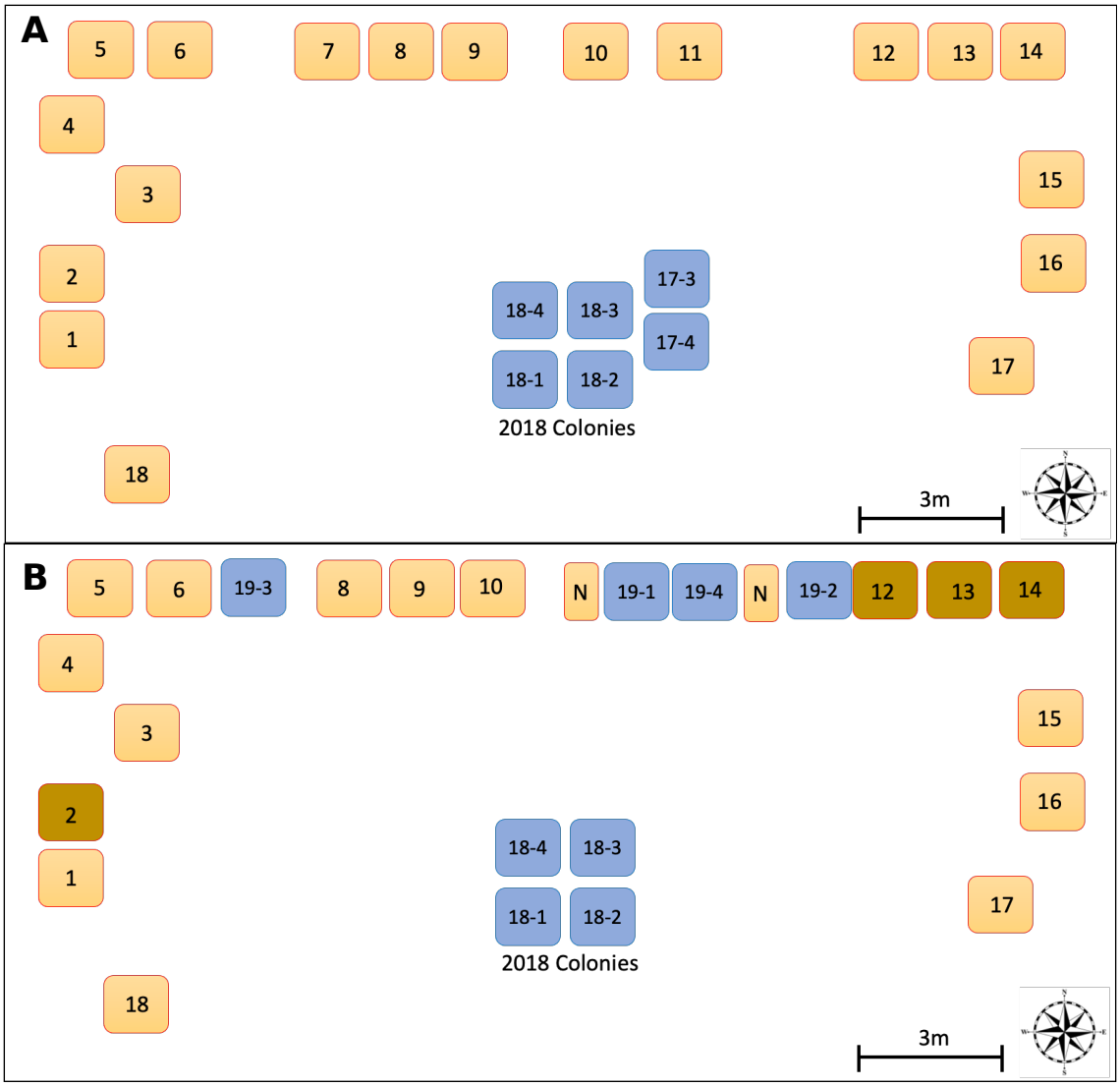


Figure 3-3 – Aerial view visualisation of the infested apiary. *Each square represents a colony on the site, the yellow colonies are pre-existing highly-infested colonies.* ***A*** *- In the first two years of this experiment, the healthy colonies were left on a pallet in the middle of the site, within a 10m range of all the colonies. The four blue tiles, labelled 18-1 – 18-4 represent the 2018 colonies, tiles 17-3 and 17-4 represent the two remaining colonies from 2017. Those that survived the winter of 2017 were placed alongside these new colonies.* ***B*** *- As four hives from the 2018 sample set were still viable at the start of 2019, the 2019 hives (labelled as 19-1 – 19-4 in blue), were put in available spaces between infested hives. The dark shaded hives (labelled 2, 12, 13 and 14) were the colonies selected for use in the 2019 shook swarm experiment (described in section 3.3). N = nucleus hive. Some of the Aberdeen colony numbers changed over time as colonies failed and/or were merged, but the positions remained largely similar.*

When sampling bees from the colonies, emerging brood, rather than adult workers, were used, to ensure all bees originated from the hive being sampled. All colonies were screened for DWV before being moved to the infested apiary. All initial viral loads are reported as the first time point in each year's analysis. Colonies were sampled every

month thereafter whilst brood was available. A small square of honeycomb (approximately 5cm × 5cm) containing late-stage (purple-eyed) pupae was cut from each of the four hives and placed in a 35°C incubator. Emerging workers were snap-frozen in liquid nitrogen as they emerged from the comb and stored at -80°C until further analysis.

3.2.3 Analysis of DWV titres in year-one colonies

From July 2017 (17/7) onwards emerging brood were sampled from the four colonies (**Table 3-1**). The RNA was extracted, cDNA generated and the relative DWV titre of individual bees was calculated by qPCR using known quantified standards. By October (17/10) of 2017, colony 2 had died with the change in viral titre increasing significantly from June to September (non-parametric Mann-Whitney test of statistical significance p-value = 0.0147) (**Figure 3-4**). The viral titre in the other three colonies significantly increased between 17/7 and 17/10 from ~10^4 GE/µg RNA to between 10^6-10^{10} GE/µg RNA (Mann-Whitney p-values of 0.0029, <0.0001 and <0.0001 for colonies 17-1, 17-3 and 17-4 respectively).

Table 3-1 - Colony sampling time points. *A tick indicates brood samples were collected, an X indicates no brood was available that month. The numbers by the months are the numerical identifiers for each sample collection.*

Relative Viral load table

	Timepoint		Hive set 1 (2017)				Hive set 2 (2018)				Hive set 3 (2019)			
			17-1	17-2	17-3	17-4	18-1	18-2	18-3	18-4	19-1	19-2	19-3	19-4
2017	July	17/7	✓	✓	✓	✓								
	August	17/8	✓	✓	✓	✓								
	September	17/9	✓	✓	✓	✓								
	October	17/10	✓		✓	✓								
2018	May/June	18/6			✓	✓	✓	✓	✓	✓				
	July	18/7			✓	✓	✓	✓	✓	✓				
	August	18/8				✓	✓	✓	✓	✓				
	September	18/9				✓	✓	x	x	x				
2019	May	19/5					✓	✓	✓	✓				
	June	19/6									✓	✓	✓	✓
	July	19/7					✓	x	✓	✓	✓	✓	✓	✓
	August	19/8					✓	✓	✓		✓	✓	✓	✓
	September	19/9					✓	✓	✓		✓	✓	✓	✓
	October	19/10					x	✓	✓		✓	✓	✓	✓

Despite the three surviving colonies displaying significantly increased viral loads on average by the end of the first season, not all emerging workers had very high DWV levels. DWV yields in several samples from 17-1 were comparable to earlier months and no samples in 17-4 exceeded 10^8 GE/µg RNA. Colony IN (infested neighbouring), a local colony, had higher DWV titres than any of the new colonies, with the lowest DWV yield observed >10^8 GE/µg RNA.

The two surviving colonies (17-3 and 17-4) were sampled again from Spring 2018 (18/5) onwards (**Figure 3-4**). DWV titres in 17-4 were comparable to the spring of 2017, whilst 17-3 was lower on average than the samples observed at the end of 2017. DWV titres from 17-3/18/7 were increasing (>10^5 GE/µg RNA with one sample >10^{11} GE/µg RNA), but the colony died before any further sampling could be carried out. Colony 17-4, however, did not show significant increases in viral titre (as it had done in 2017), and by the final collection point (18/9) both it and colony IN had viral loads of ~5×10^4 GE/µg RNA with the exception of one sample (~10^{10} GE/µg RNA).

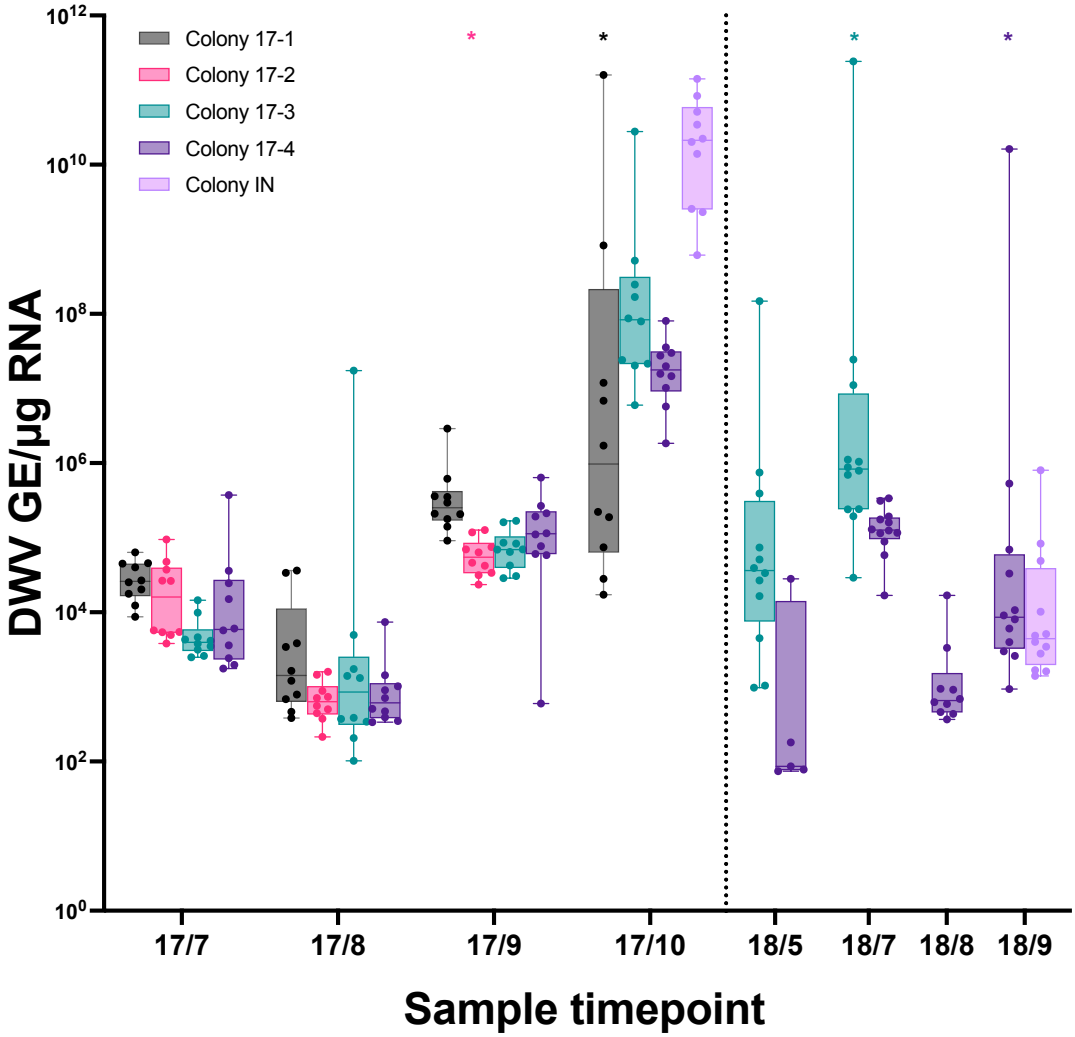


Figure 3-4 - Analysis of DWV titres in the four colonies moved to the mite-infestation apiary in 2017. *The hives which survived the first season (Colony 3 and 4) were measured throughout 2018 whilst they still produced brood. Samples were taken at the end of each season from neighbouring mite-infested colonies to compare relative DWV loads. The dashed line indicates the end of the sampling season when the colonies have no brood. The asterisks shown above the sample values indicate the last brood sample before the colony died. Colony IN is a neighbouring colony in the apiary.*

3.2.4 Analysis of DWV titres in year-two colonies

In the second year all colonies (18-1 – 18-4) had very low DWV titres during the initial screening and were placed in the apiary in July (18/7). No colonies had any brood with DWV titres >10^5 GE/µg RNA throughout the season and the local infested colony (colony IN), although slightly higher, didn't have any samples with DWV titres >10^6 GE/µg RNA (**Figure 3-5**). Only Colony 18-1 and Colony IN had brood samples by 18/9.

In 2019, DWV titres began increasing significantly from July (19/7) onwards, with colony 18-1 >10^8 GE/µg RNA for around half the samples processed and colony 18-3 titres all >10^6 GE/µg RNA (**Figure 3-5**, 19/7). The colonies had patchy brood and as such

no emerging brood samples were available from 18-2 by 19/7 and 18-4 had a drone-laying queen meaning the colony was dead by 19/8. 18-1 had high mite levels by 19/7 (observed when checking frames and brood) and this was reflected in the high DWV titres, recorded between 10^6 and 10^{12} GE/µg RNA for all samples, with only 1 sample at 10^6 GE/µg. 18-1 had died by 19/9 and the remaining sealed brood cells were sampled from an otherwise empty hive. Three colonies (18-1 – 18-3) showed statistically significant changes in viral titre from 18/6 to 19/9 (Mann-Whitney p-value of <0.0001) and colony 18-4 was statistically significant between 18/6 and 19/7 (p-value - 0.0127). Samples of adult workers were collected on 19/10 from colonies 18-2 and 18-3 and colony IN, as all colonies were broodless following a prolonged spell of bad weather.

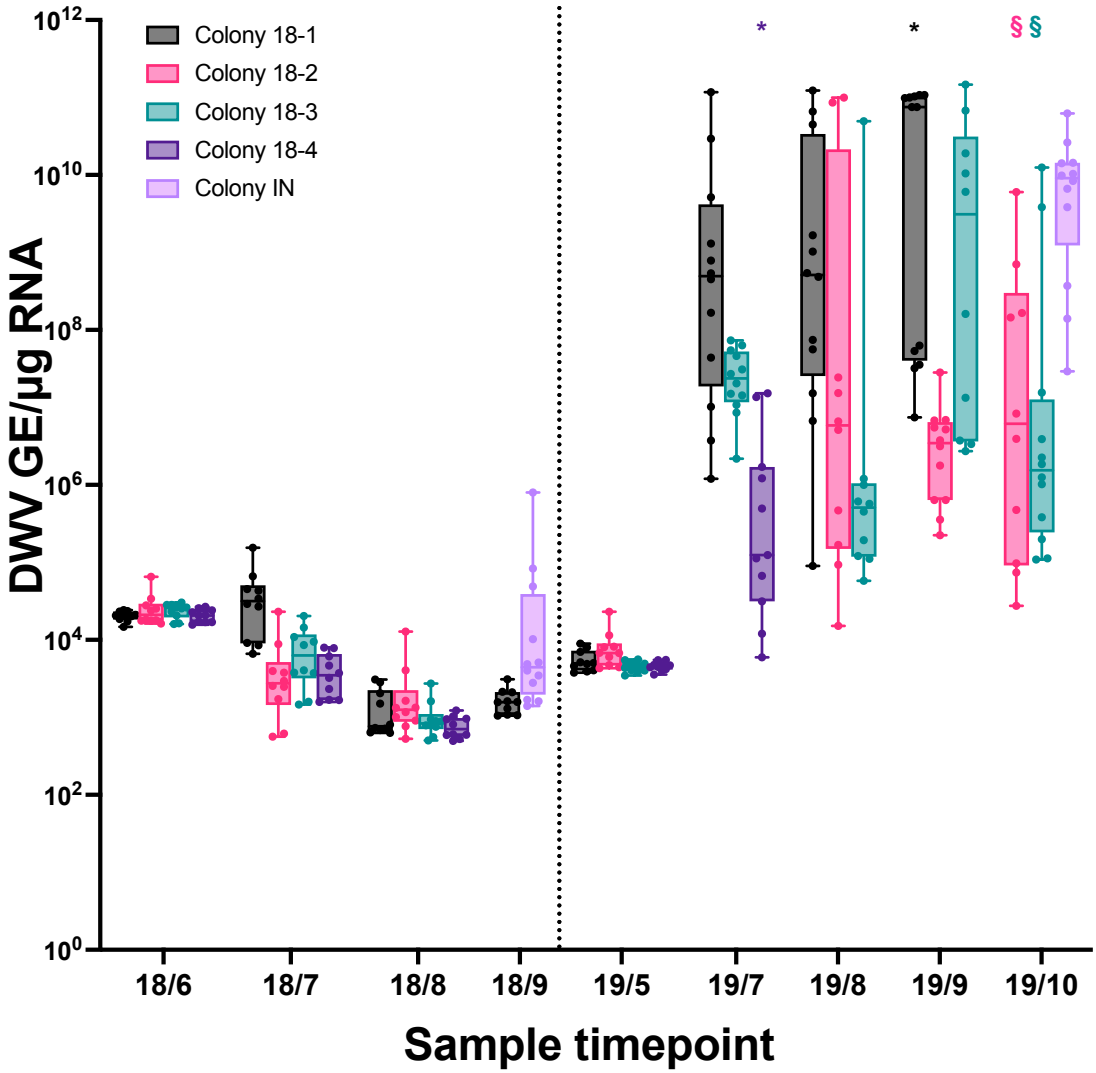


Figure 3-5 - Analysis of DWV titres from the four colonies moved to the mite-infested apiary in 2018. *The colonies were moved to the infested site in 2018 and sampled until the end of 2019. The dashed line indicates the end of one bee season. Asterisks above two colonies indicate the point where the last sample was taken before it died. The § indicates samples of adult workers taken when no emerging brood was available.*

3.2.5 Analysis of DWV titres in year-three colonies

The four colonies placed in the mite-infested apiary in 2019 (19-1 – 19-4) showed a rapid increase in DWV titre (**Figure 3-6**). By the first sample point after colony movement

(19/7) the virus levels had already increased significantly in 19-1 (all >10^6 GE/µg RNA) and were higher in 19-2 and 19-4 (10^5–10^8 GE/µg RNA). Viral titres in 19-3 did not change significantly between 19/6 and 19/7 and the majority of samples taken from this colony ranged between 10^4 and 10^8 GE/µg RNA (Mann-Whitney p-value = 0.1673). 19-1 and 19-4 had DWV-symptomatic bees emerging from brood cells by 19/9 and these bees had very high viral loads (~10^{11} GE/µg RNA), which had increased significantly from 19/6 (Mann-Whitney p-value <0.0001 for both colonies). No brood was available from 19-2 at 19/9 or any of the colonies in 19/10, so viral titres were quantified from adult workers instead. This is reflected in the lower average DWV titre and the increased variation observed between workers sampled at 19/10. The samples taken from the local colony (Colony IN) indicated viral loads were typically high on the site by 19/10 (~10^{10} GE/µg RNA), similar to the data observed in 2017. Further testing of these colonies in 2020 was not possible due to Covid-19 restrictions.

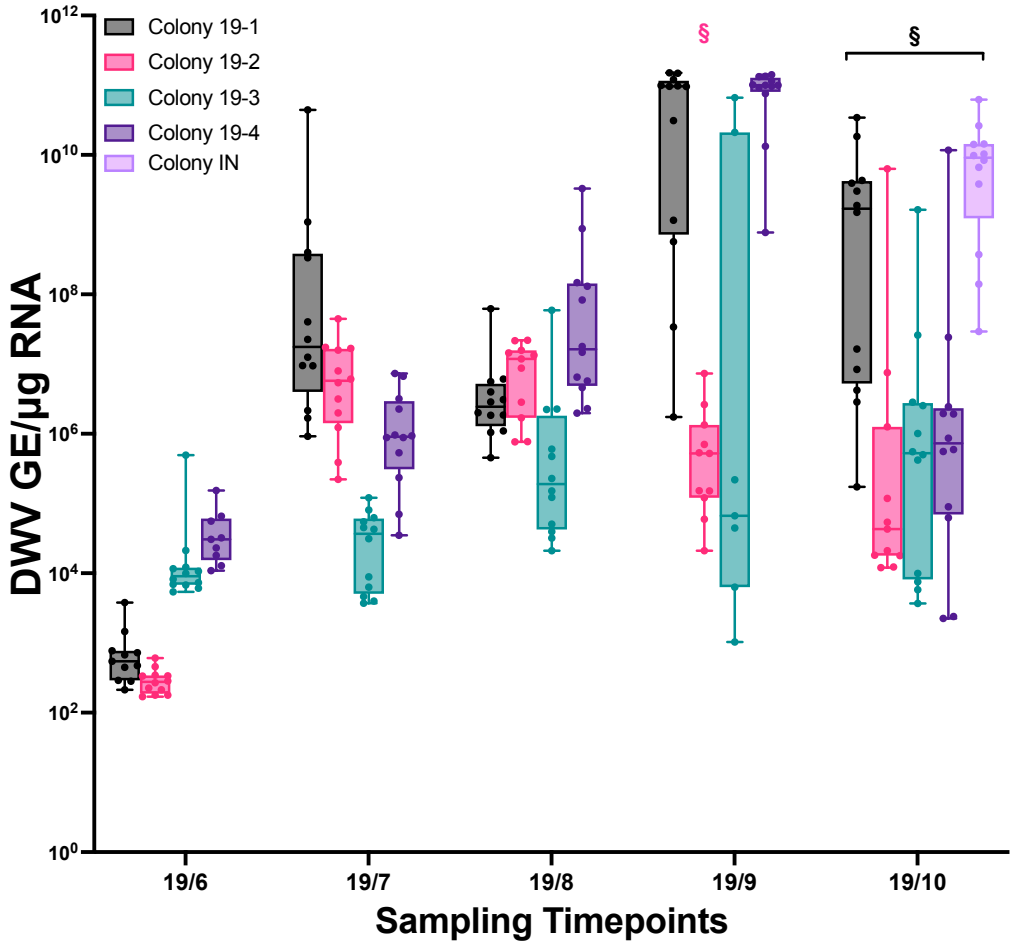


Figure 3-6 - Analysis of DWV titres in the four colonies moved to the mite-infestation apiary in 2019. *The colonies were moved to the infested site in June '19 (19/6) and emerging brood was sampled every month until the queen stopped laying towards the end of the season. The § indicates the viral load of adult workers which were taken when no sealed brood was available.*

3.2.6 Comparison of DWV titre changes across all 3 years

The average DWV titre at each sampling point was calculated from the 12 individual workers sampled and quantified. These averages were plotted as a single point per month to illustrate the changes in viral load over the three years (**Figure 3-7**). By plotting all

twelve colonies together it is clear that the colonies placed on site in 2018 were not infested as quickly as those in 2017 or 2019 and this is reflected in the DWV titres recorded that year. Despite this, those colonies from 2018 still displayed significantly higher DWV titres by the middle of 2019 (Mann-Whitney p-value <0.0001, **Table 3-2**) and two of the colonies died (**Table 3-3**). In 2017 and 2019 the DWV titres of the colonies freshly placed on site had significantly changed between the first time point and the last (Mann-Whitney p-value <0.0001) and all the 2017 colonies died within 18 months of placement (**Table 3-3**).

Samples were taken from neighbouring infested colonies (IN) to determine the base level of DWV on the site. The purple points represent the average DWV titre at the end of each season from a neighbouring colony (**Figure 3-7**). The DWV titres from 2017 and 2019 are like those obtained from the *Varroa*-naïve colonies, whilst the 2018 viral load was much lower, reflecting the overall pattern of low DWV titres for that season. One possible explanation for the year-to-year variation is the difference in average temperature observed throughout the study. **Figure 3-8** illustrates that the average temperature between November 2017 and March 2018 was colder than during the same period in 2018-19. This prolonged colder spell may have reduced the number of phoretic mites able to survive through winter and infest new colonies the following spring, and may also have reduced the number of reproductive cycles possible. The year-to-year differences in DWV titres may reflect that temperature plays a significant role in the transmission of *Varroa* from one colony to another. For example, warmer temperatures might increase the number of days over which drifting could occur, or a warmer autumn could increase brood rearing in a colony when mite levels are very high.

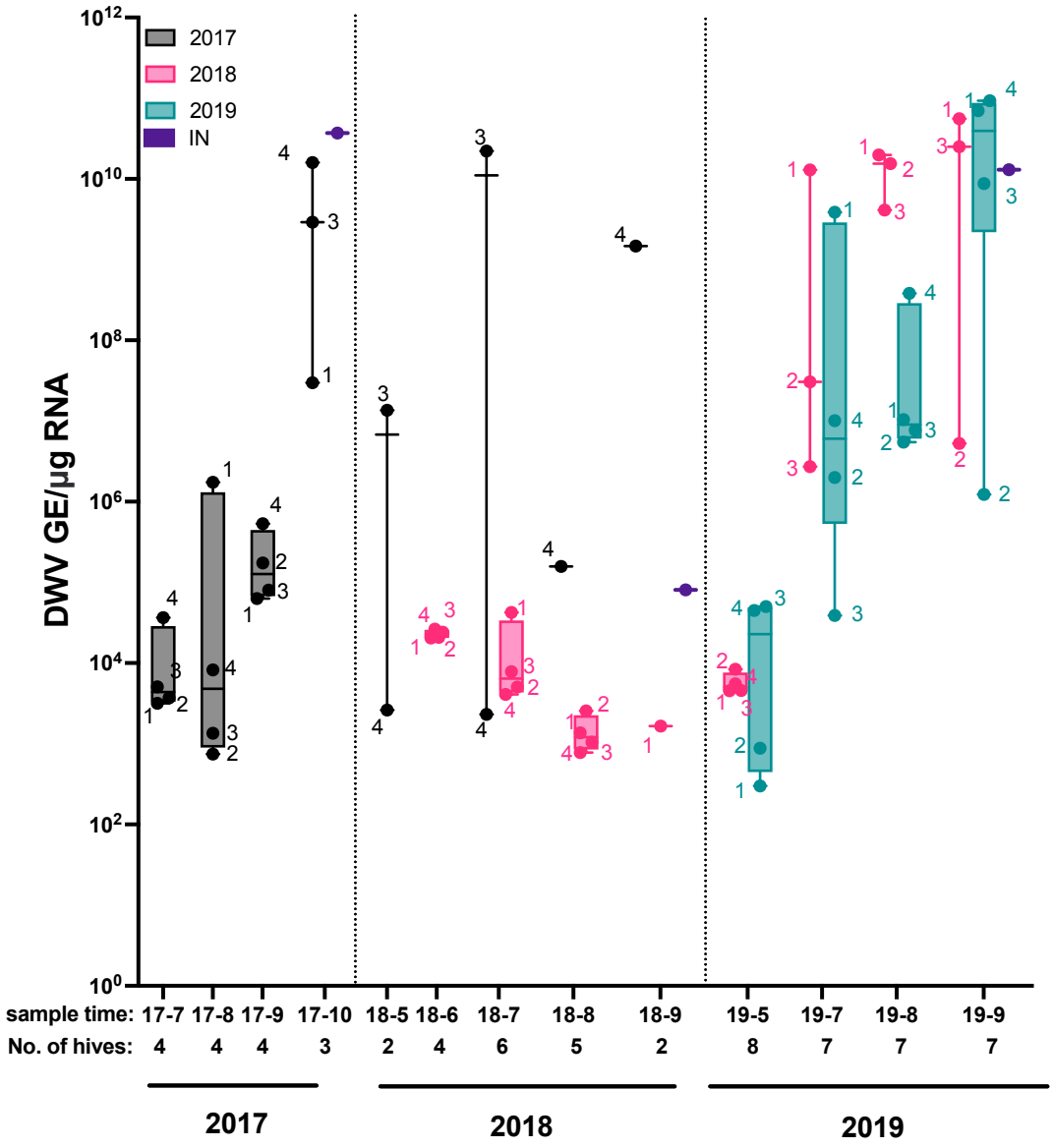


Figure 3-7 - Analysis of DWV titres in twelve colonies from three seasons when placed in the mite-infested apiary without treatment. *Across three seasons colonies with very low* Varroa *and DWV titres were placed in a highly mite-infested apiary and monitored. Each point represents the average quantified DWV load from 12 emerging workers from each colony placed in the apiary. The violin plots represent all the colonies from each year for each time point. The purple dots indicate average DWV titres for neighbouring infested colonies. Only viral titres from emerging brood are shown in this figure.*

Table 3-2 - Mann-Whitney unpaired t-test summary data. *The statistical test was used to compare the DWV titre reported from the brood samples taken prior to colony movement and at the last available brood sample*

Colony	Sample time point	P-value	P-value summary
17 - 1	17/7 – 17/10	0.0029	**
17 - 2	17-7 – 17/9	0.0147	*
17 - 3	17/7 – 17/10	<0.0001	****
17 - 4	17/7 – 17/10	<0.0001	****
18 - 1	18/5 - 19/9	<0.0001	****
18 - 2	18/5 - 19/9	<0.0001	****
18 - 3	18/5 - 19/9	<0.0001	****
18 - 4	18/5 - 19/6	0.0127	*
19 - 1	19/5 - 19/9	<0.0001	****
19 - 2	19/5 - 19/9	<0.0001	****
19 - 3	19/5 - 19/9	0.1673	ns
19 - 4	19/5 - 19/9	<0.0001	****

Table 3-3 - Relative viral loads of colonies left in highly infested apiary over time. *Colour scale represents increasing viral load from yellow to red. Black bars represent points where colonies have died. White boxes indicate where no brood samples were available in the colonies. Only viral titres from brood samples are shown in this table.*

	Timepoint		Hive set 1 (2017)				Hive set 2 (2018)				Hive set 3 (2019)			
			17-1	17-2	17-3	17-4	18-1	18-2	18-3	18-4	19-1	19-2	19-3	19-4
2017	July	17/7	3.7E+04	5.1E+03	3.7E+03	3.2E+03								
	August	17/8	8.2E+03	7.5E+02	1.7E+06	1.4E+03								
	September	17/9	5.4E+05	6.3E+04	8.0E+04	1.7E+05								
	October	17/10	1.6E+10		2.9E+09	3.0E+07								
2018	May/June	18/6			1.4E+07	2.6E+03	2.04E+04	2.64E+04	2.41E+04	2.08E+04				
	July	18/7			2.2E+10	2.3E+03	4.24E+04	5.04E+03	7.83E+03	4.08E+03				
	August	18/8				1.6E+05	1.37E+03	2.56E+03	1.05E+03	7.80E+02				
	September	18/9				1.5E+09	1.7E+03	No brood available						
2019	May	19/5					1.55E+03	3.80E+03	2.20E+04	1.80E+04	8.81E+02	3.00E+02	5.01E+04	4.49E+04
	July	19/7					1.29E+10	NB	3.06E+07		3.86E+09	1.01E+07	3.89E+04	2.00E+06
	August	19/8					1.98E+10	1.55E+10	4.11E+09		7.70E+06	1.04E+07	5.46E+06	3.83E+09
	September	19/9					5.57E+10	5.27E+06	2.51E+10		7.02E+10	1.23E+06	8.76E+09	9.33E+10

Key
- $<10^6$
- 10^6-10^8
- $>10^8$
- Dead

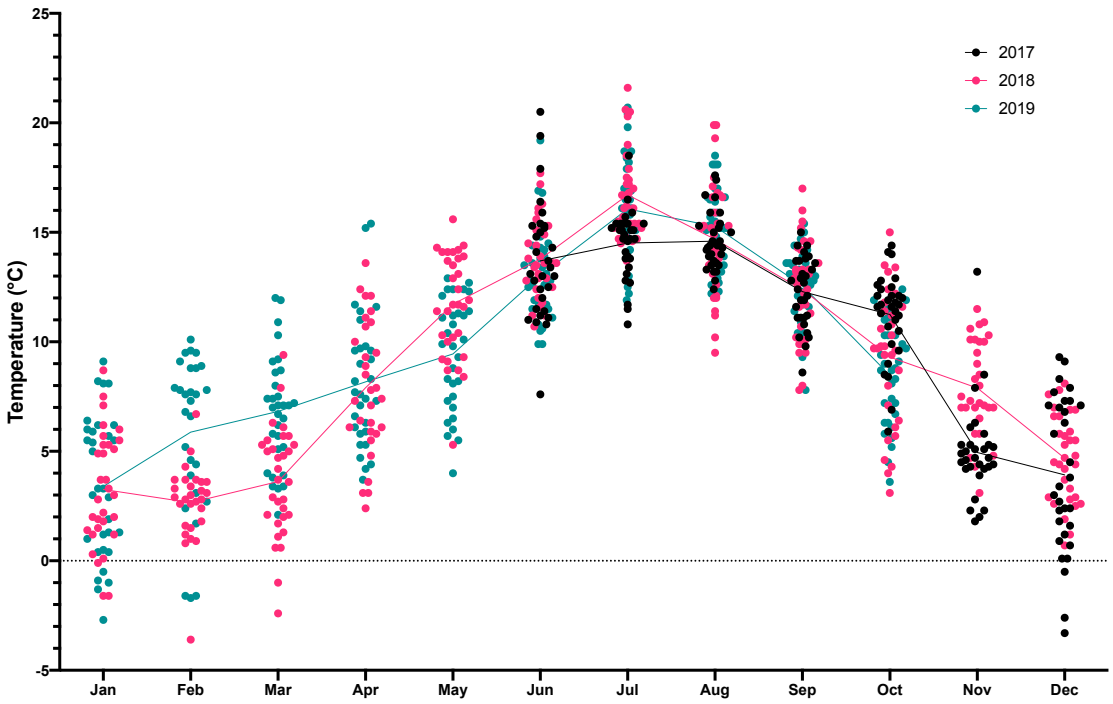


Figure 3-8 - Mean daily temperature for the period in which colonies were placed in the mite-infested apiary (June 2017–October 2019). *The mean daily temperature (sourced from Met Office data by email request) is shown for each month, each data point represents the mean for one day and the line represents the mean for each month. Colonies were placed on site in June 2017 and data is shown here until October 2019 when the final sample of the year was collected.*

3.2.7 Analysis of Sanger sequences of virus populations

Diversity in the viral population is considered a metric of colony health, with decreased diversity associated with increased viral load (Martin *et al.*, 2012; Ryabov *et al.*, 2014). To evaluate viral diversity in the honey bees sampled over time, Sanger sequencing on uncloned PCR products was carried out on known regions of variability in the DWV genome. The point between the structural and non-structural proteins is a known region of recombination (Ryabov *et al.*, 2014) and the leader protein is the most variable region of the DWV genome (Lanzi *et al.*, 2006) so both regions were tested here. The RNA-dependent RNA polymerase (RdRp) was also tested based on previous analyses of DWV diversity (Wilfert *et al.*, 2016).

An initial examination of the DWV population from the 2017 colonies was carried out using the R package SangerseqR (Hill, 2017) to determine whether samples had mixed populations or were near-clonal. A region of known recombination in the middle of the DWV genome (4900–5500bp based on the VVD DWV construct, NCBI accession number: MT415950) was amplified, however sensitivity was poor, so only the dominant variant in each sample is reported. These were used to generate neighbour-joining phylogeny of the changing virus populations over time for each colony (**Figure 3-9**). Over time the dominant virus in the populations shifts from a 'Type A'-like sequence to a 'Type B'-like sequence for 17-1 (panel A) and 17-4 (D), regardless of the viral titre (samples between 10^4-10^{11} GE/µg RNA). However, colony 17-3 (C) still aligned with

known Type A reference genomes by 17/10. Colony 17-2 (B) died by 17/7 but showed a mixed population of Type A and Type B like sequences at early time points. Samples from a neighbouring colony are shown in bold and all align with Type A-like sequences.

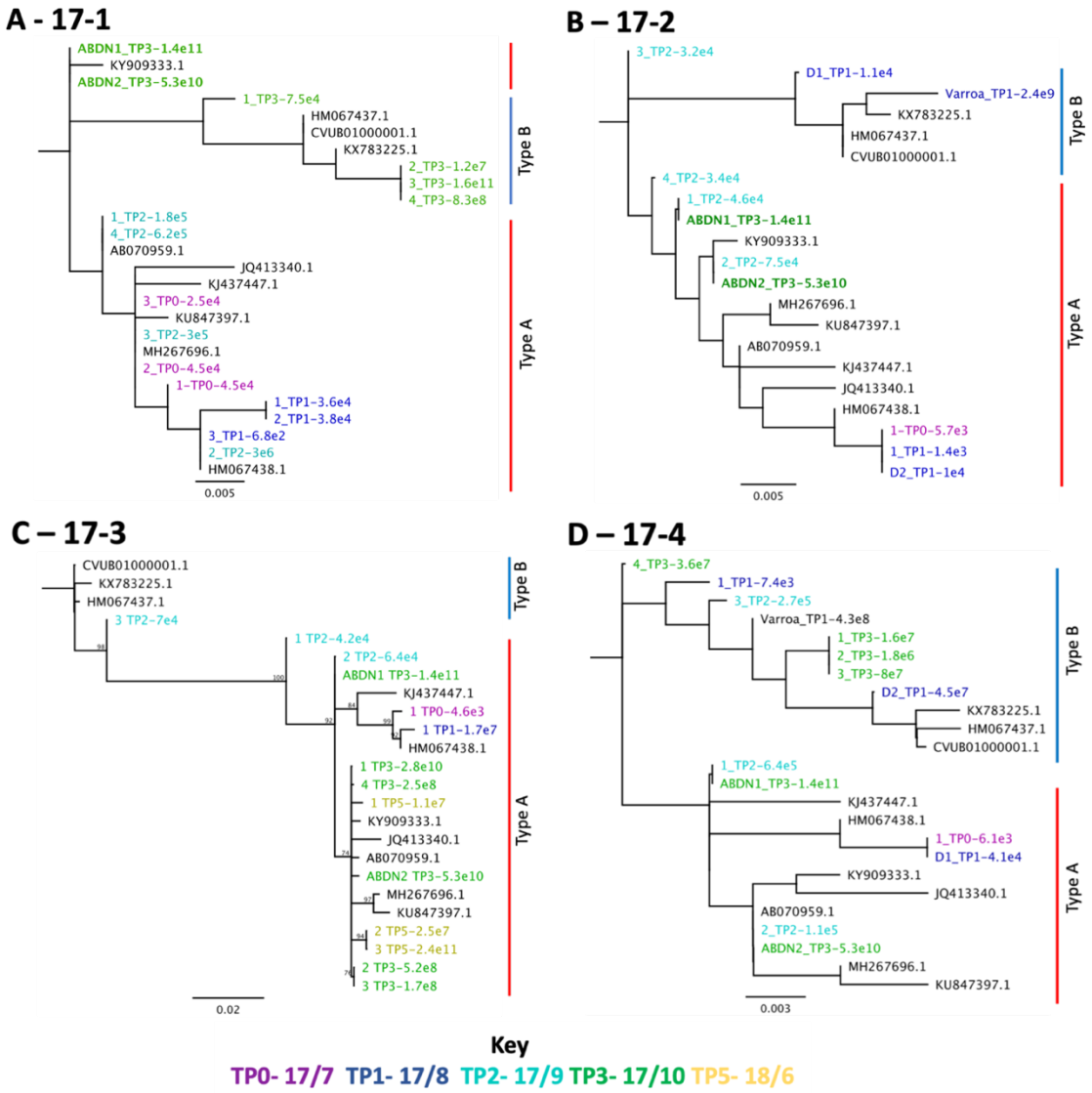


Figure 3-9 - Neighbour-joining phylogenetic analysis of Sanger sequences from four colonies placed in an infested apiary in 2017. *Sequence alignments for each colony shown on separate phylogenetic trees (panels A-D). All samples are coloured by time point including neighbouring infested colonies. Reference sequences are shown by their NCBI accession numbers. The red and blue lines indicate sequences clustering with known 'Type A' and 'Type B' reference genomes respectively. The number at the end of each sequence name indicates the quantified viral load (GE/μg), determined by qPCR. Reference sequences are shown in black by their NCBI accession numbers. The genome position is based on primers designed using a DWV reference sequence (NCBI accession number: MT415950).*

Samples from two colonies placed in the mite-infested apiary in 2019 were amplified and sequenced in the same way as the 2017 samples. **Figure 3-10** shows the phylogenetic diversity of DWV variants over time in two colonies, 19-3 (panel A) and 19-4 (panel B)

from **Figure 3-6**. As with the 2017 colonies, the variants change over time on a colony level. Colony 19-3 viruses clustered with Type-A variants during the first two months of sampling, but the dominant variants from time point 2 (19/8) onwards were more closely related to Type B variants. 19-4 remained tightly clustered to the Type A variants throughout the time course, with only one honey bee showing a phylogenetically distinct variant by time point 3. The samples taken at each time point clustered closely to one another, time point 0 (purple) and time point 1 (indigo) show high sequence similarity to one another, indicating DWV diversity changes are occurring on a colony level rather than on an individual worker level over time. Three sequences from a neighbouring colony clustered in the Type B clade, like the 2017 samples, but one sample clustered with the Type A variants.

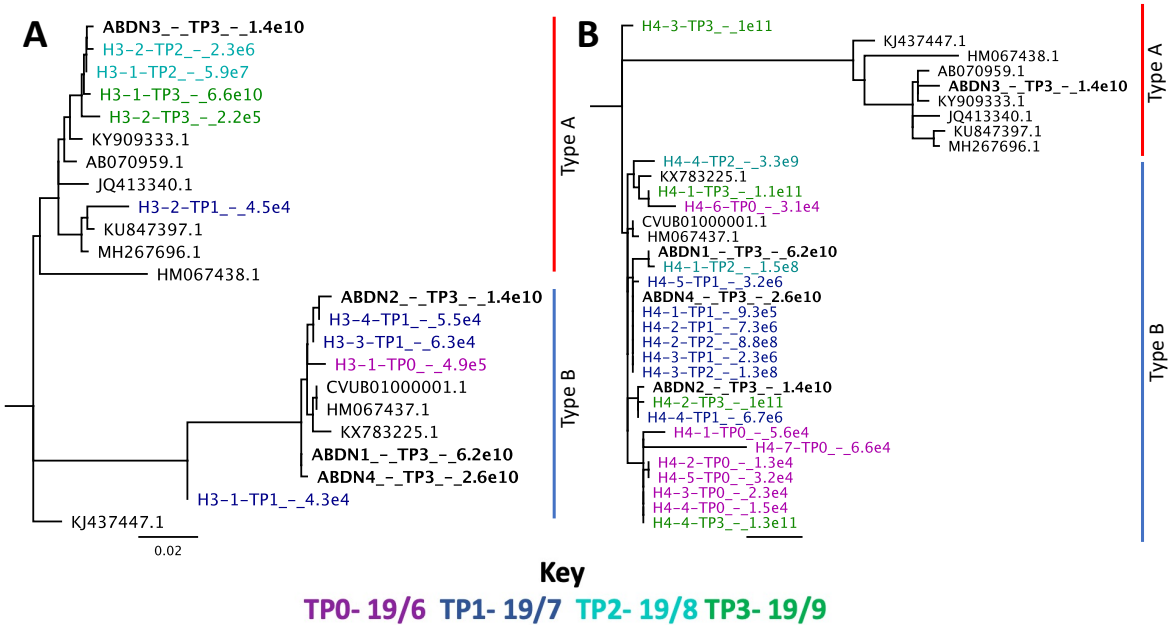


Figure 3-10 - Neighbour-joining phylogenetic analysis of DWV sequences from two colonies when placed in a mite infested apiary in 2019. *The phylogenetic diversity of the virus population from colony 19-3 (**A**) and colony 19-4 (**B**) placed in the infested site in 2019 is shown with each time point coloured. Samples taken from neighbouring infested sites are highlighted in bold and reference sequences are shown by their NCBI accession numbers. The red and blue vertical lines indicate branches corresponding to 'Type A' and 'Type B' sequences. The number at the end of each sequence name indicates the quantified viral load for that bee, determined by qPCR analysis and recorded as GE/µg RNA.*

3.2.8 Next-generation sequencing analysis of DWV diversity changes

The R-based SangerseqR analysis was insufficient for determining sequence diversity in the viral populations as only a single variant was detected by standard endpoint-PCR and chromatogram analysis failed to detect any secondary variants. Therefore, a selection of samples were analysed using next-generation sequencing methods. Samples were selected from colonies 17-1 and 17-3, as they survived the first few months in the infested apiary and showed, based on Sanger sequences, distinctive viral variants from one another over time in 2017 (**Figure 3-9A** and **C**). As colony 17-3 had very low diversity and 17-1 appeared to have mixed populations these two colonies were selected.

Full genome amplicons from individual workers from 17-1 and 17-3, and an infested neighbouring colony, were generated using Long-amp PCR, selecting samples from every time point in the first year, and sequenced using an Illumina MiSeq. Viral diversity in the individual samples was determined based on assigned haplotype clusters using ShoRAH (Short Read Assembly into Haplotypes) (Zagordi *et al.*, 2011). Briefly, ShoRAH generates haplotypes by detecting SNP (single nucleotide polymorphisms) variants using a set of overlapping windows across a specified window of the genome to measure all the mapped sequence reads, the SNPs which occur across multiple windows are then used to compile unique sequence variants and the resulting haplotypes are then assigned a percentage of the total make-up of that sample's population. Using all the Illumina-generated sequence reads, the RdRp region of the DWV genome was selected to measure diversity in the population. The RdRp was selected as it is the region in the genome with the second highest sequence divergence between variants (Woodford and Evans, 2020) and because the sequence-depth for the leader protein was insufficient across all samples.

All haplotypes that represented a >3% assignment of the population (see section 2.4.3 for rationale) were aligned using Muscle (Edgar, 2004) and neighbour-joining phylogenetic trees were generated to identify the genetic diversity in the samples**.** From the resulting phylogeny (**Figure 3-11**), samples were assigned a 'cluster' ID based on the clades they aligned to. Some sequences were placed into one clade for simplicity even if they showed some genetic distinction from neighbouring sequences, because these reads still clustered more tightly to one another than to any other matches (cluster 13). All sequences were assigned to sample collection time points for simplicity, where TP0-3 are the sample collection points of 17/7-17/10.

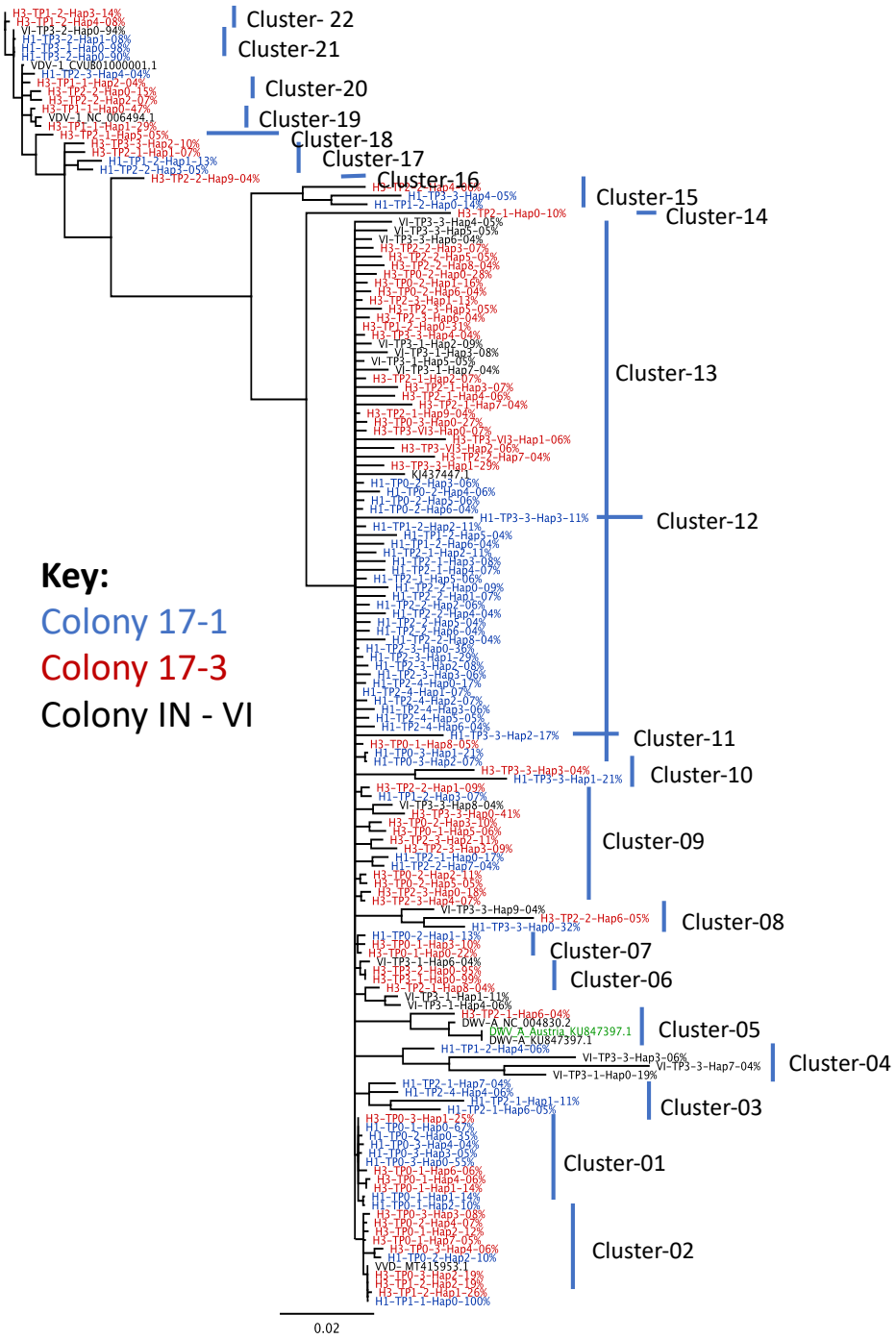


Figure 3-11 - Neighbour-joining phylogenetic analysis of all ShoRAH haplotypes for the RdRp of colonies 1 and 3 and a neighbouring infested colony in 2017. *All haplotypes assigned with >3% of the population were included in analysis. For simplicity, sequences were assigned to 'clusters' based on where they aligned to determine large changes in virus population diversity over time. The HAPx-xx% represents the haplotype number assigned by ShoRAH and the percentage of the population of that bee that it accounts for. Sequences from a neighbouring colony are included in the analysis and are denoted by VI-TP3-x-Hap-xx. Reference genomes are shown in black and represented by their NCBI accession number. The sequence shown in green was used for the sequence alignment. 1000 bootstrap iterations were run and genetic-distance was calculated using a Tamura-Nei model* (Tamura and Nei, 1993).

Probabilistic clustering of the sequences combined with statistical tests for strand bias for each time point was carried out using the ShoRAH single nucleotide variant (SNV) caller (McElroy *et al.*, 2013). This method verified that the SNVs determining sequence diversity were being called with high accuracy and were not false positives due to low sequence coverage or systematic errors during variant calling. The p-values of each SNV were used to generate plots for each time point (**Figure 3-12**). Low sequence diversity is indicated by tight clustering of all the p-value scores towards 1.0, indicating the majority of sequences in that sample had the same SNV present and therefore little diversity (for example, 17-3, TP1, blue sample), whilst colonies which contained a large range of SNVs produced p-value scores from 0.05 – 1.0 indicating several distinct variants present within the population (infested colony, black and blue samples). Both colonies show some range in SNV p-values, indicative of diversity, prior to colony movement (TP0), with a slight reduction after movement and then a large increase in SNV variability by TP2, before reducing in diversity again as DWV titre increased. The final SNVs reported in TP3 correlate with the results observed in the neighbouring mite-infested colony (shown at the bottom of **Figure 3-12**).

The phylogenetic clusters from **Figure 3-11** were sorted to display the percentage frequency of the most common DWV variants across time (**Figure 3-13**). An initial viral variant present at time point 0 (Cluster-01, blue bars) in both colonies is lost in the population over time, possibly as drifting workers or robber bees enter the hives and the colony becomes infested with mites and the viruses they carry. Over the course of the season, new viral variants are observed and by the end of the season, 4 of the 6 samples analysed are dominated by a single viral variant, although distinctive to each colony (**Figure 3-13, TP3**). The majority of the sequences aligned with Type A-like reference genomes and by the end of the season all the samples from colony 3 were predominantly infected with Type A-like variants, in agreement with the previous Sanger sequence analysis (**Figure 3-9**). The infested neighbouring colony contained a mix of A and B variants, with one sample containing a near-clonal level of a Type B variant (cluster-21) and this same variant is found at a similarly clonal level in 2/3 of the samples from 17-1 in time point 3, also similar to the Sanger sequence results. The remaining two samples from the infested colony contain a mix of Type A-like variants despite the high viral titres and symptomatic infections observed. Figure **3-14** shows all the haplotypes for each time point for all samples as pie charts, with the chart size determined by the titre of virus observed in the workers. The large change in size over time emphasizes the dramatic changes in viral load that occurred during the time course.

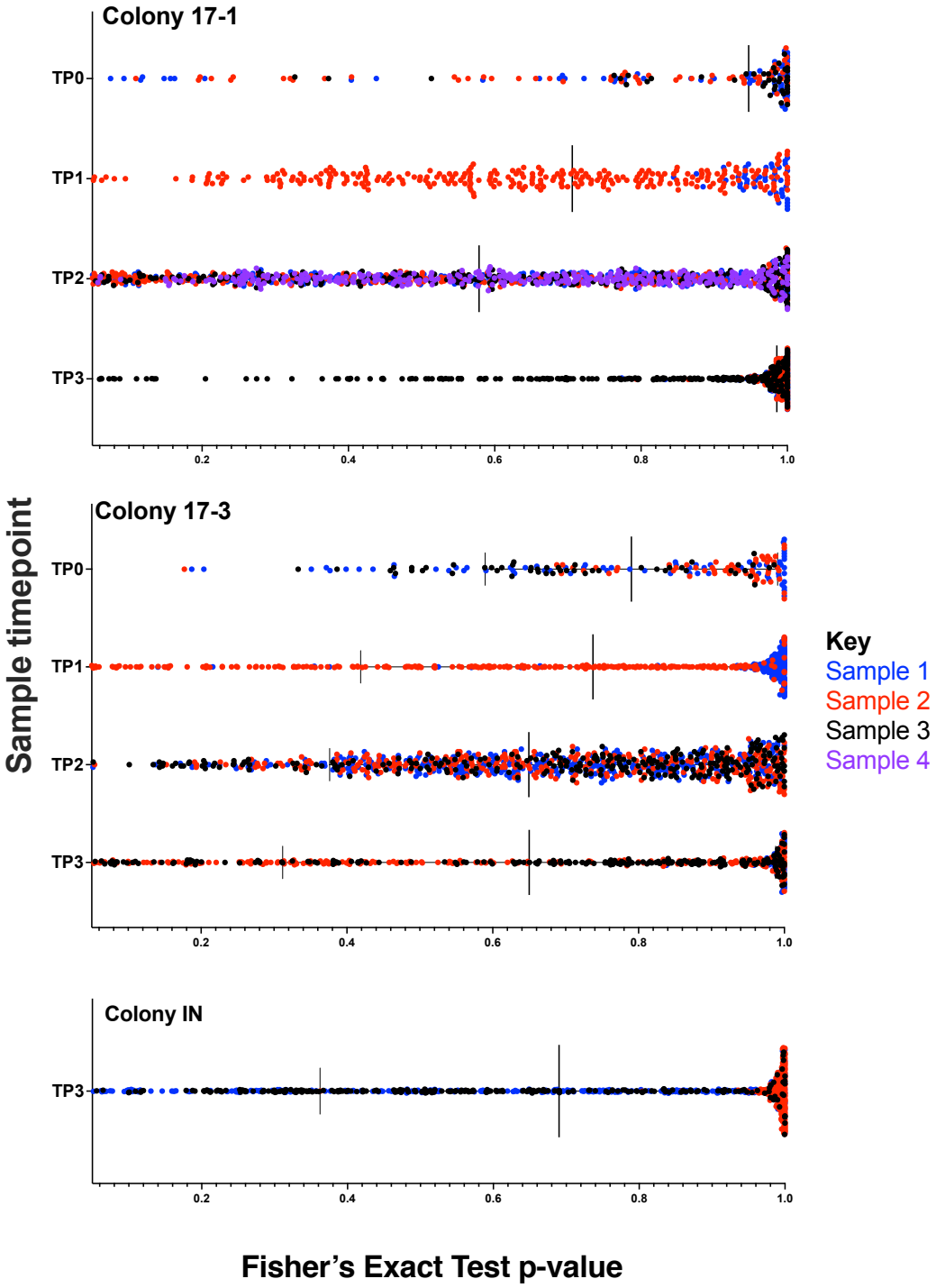


Figure 3-12 - Single nucleotide variants (SNVs) called for each sample from ShoRAH analysis of the RdRp region of DWV genome. *SNVs from ShoRAH analysis for all samples from colonies 17-1,17- 3 and IN were called if present in 3/3 iterations of the modelling. SNV p-values close to 1.0 represent a variant in the majority of sequences in that dataset differing from the reference sequence. Samples with SNVs with lower p-value scores therefore have a great amount of sequence variation within the sample. Anything with a p-value <0.05 was excluded as the threshold for error based on McElroy et al (2013) modelling data. The different colours at each time point represent the SNVs for different honey bee samples sequenced.*

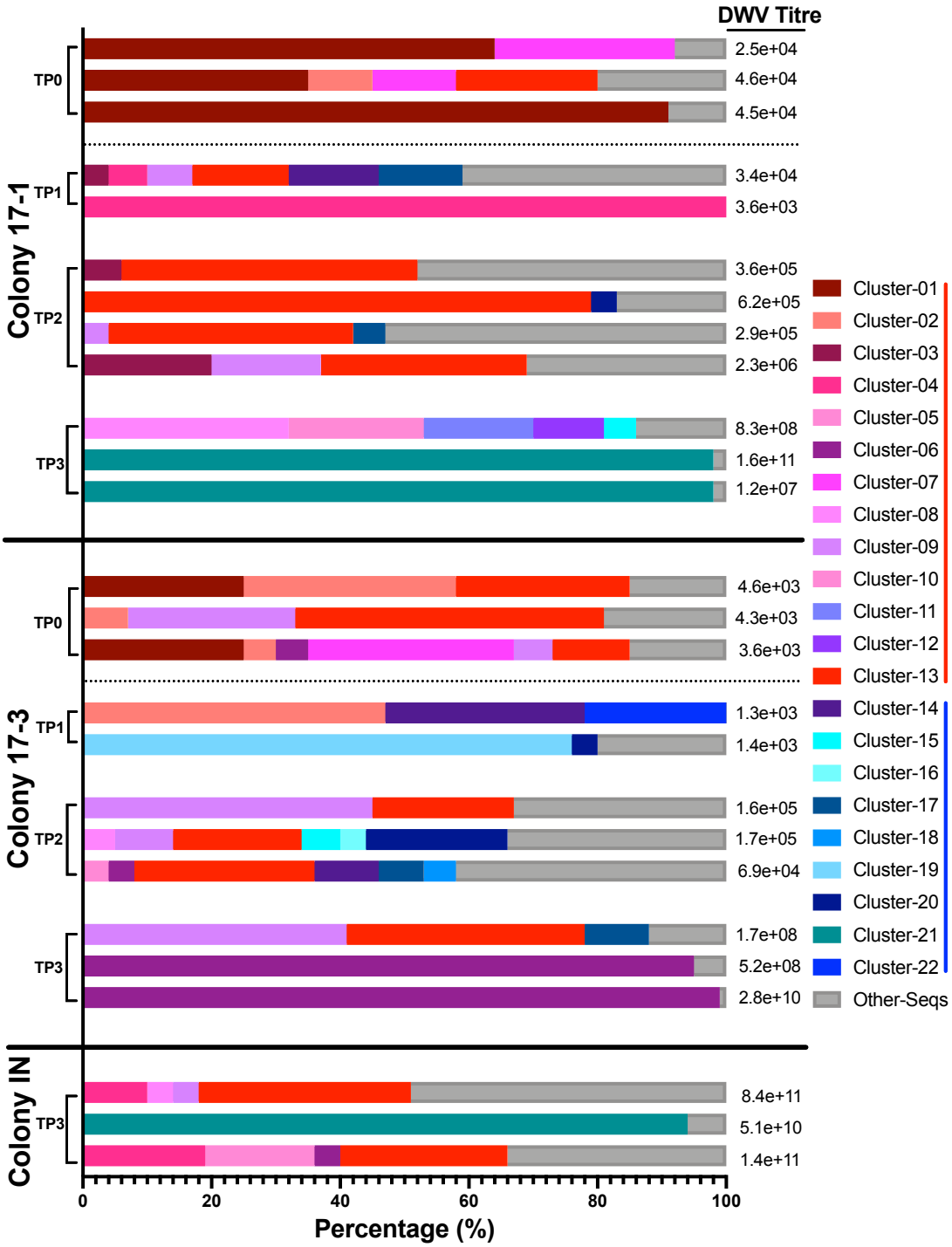


Figure 3-13 – Percentages of DWV haplotype sequence variants in two colonies and neighbouring infested colony. *Haplotypes were generated from next-generation sequence data and assigned to phylogenetic clusters for both the examined colonies. The assigned clusters were coloured shades of red and blue if they aligned with known Type A and Type B reference genomes respectively, as shown by the key on the right. The quantified viral load of each honey bee is shown beside each bar. The grey bars indicate sequences assigned as 'other' during haplotype sorting because they fell below the threshold for sequence identity (>3% of the population).*

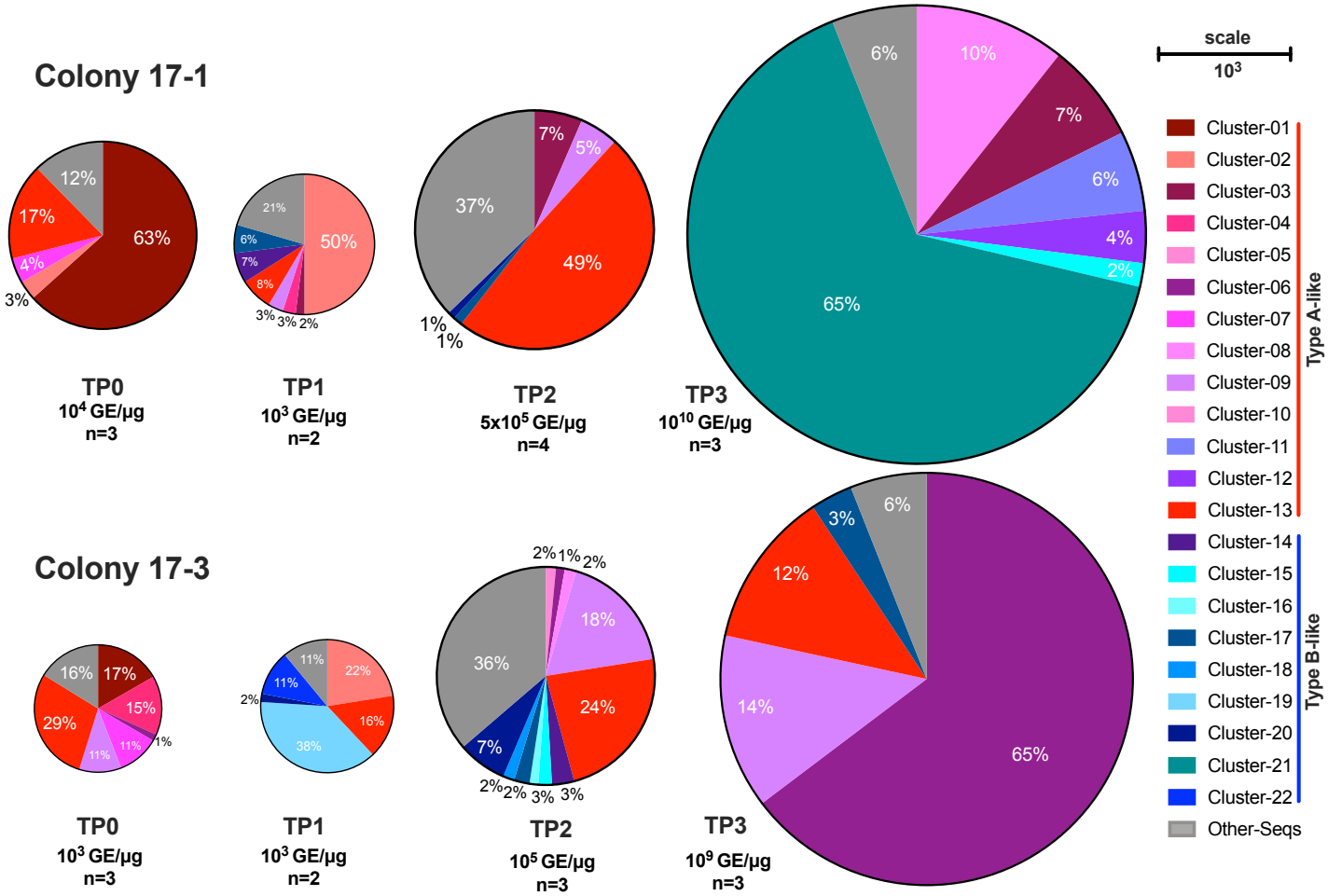


Figure 3-14 – Pie charts of haplotype diversity for each time point for Colony 17-1 and 17-3. *Each pie chart represents a single time point for one colony, with all haplotype variants from multiple bees shown. The size of the pie is determined by the DWV titre, shown below each pie, scaled using the circle diameter. The percentage that each haplotype makes up is shown in each segment. The number of samples at each time point are shown below the titre as n=x.*

3.3 Using a ‘shook swarm’ to reduce *Varroa* and DWV in highly infested honeybee colonies

3.3.1 Colony preparation and movement

In the second part of this study, the aim was to use a combination of a ‘shook swarm’ technique and miticide treatments to remove *Varroa* and then to measure the resulting changes in the DWV population, a marker of colony health. The rationale for using a shook swarm method is that the vast majority of mites reside in brood cells (~80% depending on the time of year), with the remaining mites found phoretic in the colony, therefore removing the brood removes a large percentage of the mites and a subsequent miticide treatment removes a large proportion of the remaining mites. All brood frames were removed from the colonies along with any stores they had accumulated. Therefore, the timing of this treatment and the strength of the colonies selected is critical. If the shook swarm process was performed at a time of year when the colony could not rapidly replenish its stores and the queen start laying eggs again, the colony would fail. In both years of this study the shook swarm process was performed at the start of the summer when the colonies had already had 2–3 months of mild weather in the spring to build up colony strength.

In the first year of the study (2018) four colonies were chosen (**Figure 3-15**, panel 1) by collaborators from a highly infested apiary located north of Aberdeen. This was done without screening the viral loads of the colonies. Instead, the selection was based on the number of mites observed in brood cells during early season sampling. The rationale for this was that all the colonies kept on the site are highly infested with mites and as such all should carry high levels of DWV. The selected four colonies were all reported to have high mite infestation levels, although no counts were done at this site. In the second year, 12 colonies in the apiary were screened for DWV prior to selection (see section 3.3.2).

Once colonies were selected, the queens were all marked, captured and sealed in queen-cages in each hive. This ensures the queen can be safely moved from the old, infested hive to the new, clean hive without risk of harming her or leaving her behind.

New poly-hives were taken to the mite-infested apiary to shake the adult bees into. Hives were moved off their stand and replaced with new poly-hives (**Figure 3-15**, panel 2). All frames of worker bees were taken to the new hives and shaken into the empty boxes (**Figure 3-15**, panel 3). Fresh, undrawn comb was placed in the boxes, with Apivar strips, and the hives were closed up and left until late in the afternoon to allow any foraging workers to return to the position of their hive. Colonies were then sealed and transported to a new apiary in St Andrews, approximately 20 metres from any other colonies, where they were opened and supplemented with a 1:1 syrup solution to allow the workers to quickly build up stores and the queen to start laying.

The comb containing all the original sealed brood was left in the original infested colonies on site or added to existing colonies. A small selection of this sealed brood was cut from the comb and placed in an incubator for further analysis once the workers emerged. These samples formed the first time point of DWV analysis for each colony. The first samples of the colonies post-‘shook swarm’ were taken as soon as sealed brood was available. The brood were removed from the colonies roughly once every brood cycle until the end of the season.

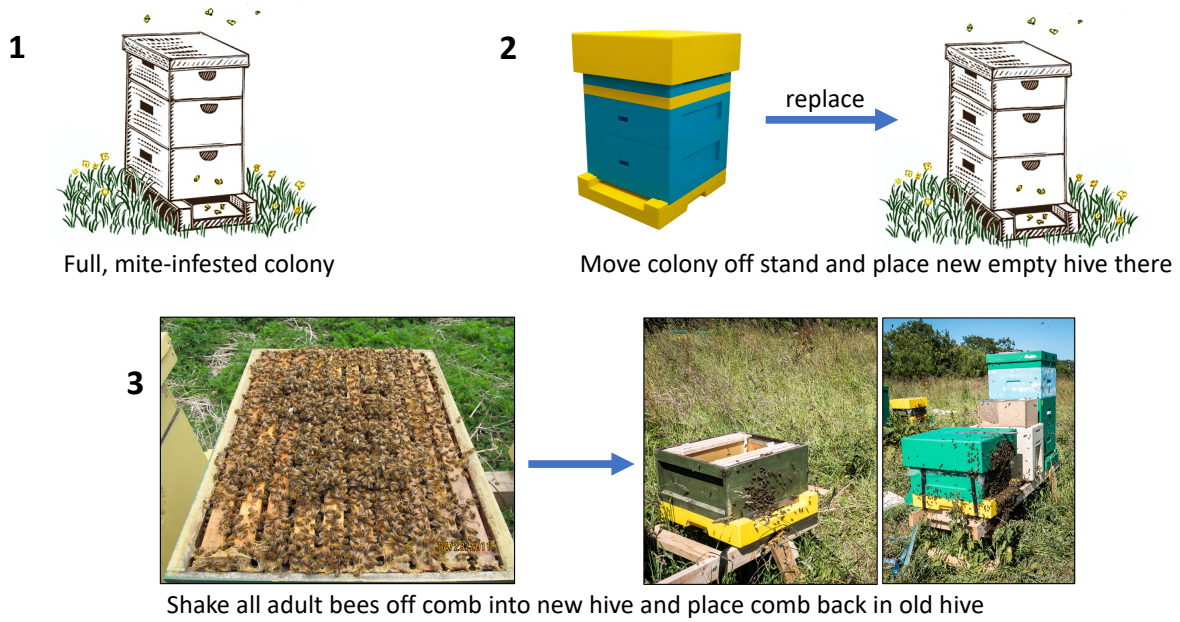


Figure 3-15 - The three step 'shook swarm' process. *Panel 1 - the selected infested colony. Panel 2 - move the infested colony from its position and put a clean, empty hive in its place. Panel 3 - take the frames of worker bees and shake them into the new hive, place the sealed brood back in the old, infested hive, wait for all foragers to return, then seal up the new colony and move to a new site.*

3.3.2 Year two screening and preparation

In the second season a selection of colonies in the mite-infested apiary were screened prior to the shook swarm, unlike in year one, to determine which had the highest DWV levels. This was to avoid a scenario similar to that observed in Colony 4 in 2018, where the DWV levels were very low at the start of the experiment and did not change significantly throughout (see **Figure 3-17B**).

In late April, emerging brood was extracted from 12 colonies and screened for DWV using standard RNA extraction, cDNA synthesis and qPCR methods. In late May the chosen colonies (the four red boxes in **Figure 3-16**) were placed into new hives using the same 'shook swarm' method as the first season. Colony 14 was selected ahead of colony 9 or colony 6 as it had a higher average viral load across all the screened individuals. Colony 2 failed shortly after moving to St Andrews, most likely due to the relatively small size of the colony and the harsh nature of the treatment method; it is therefore not included in any of the analysis shown here.

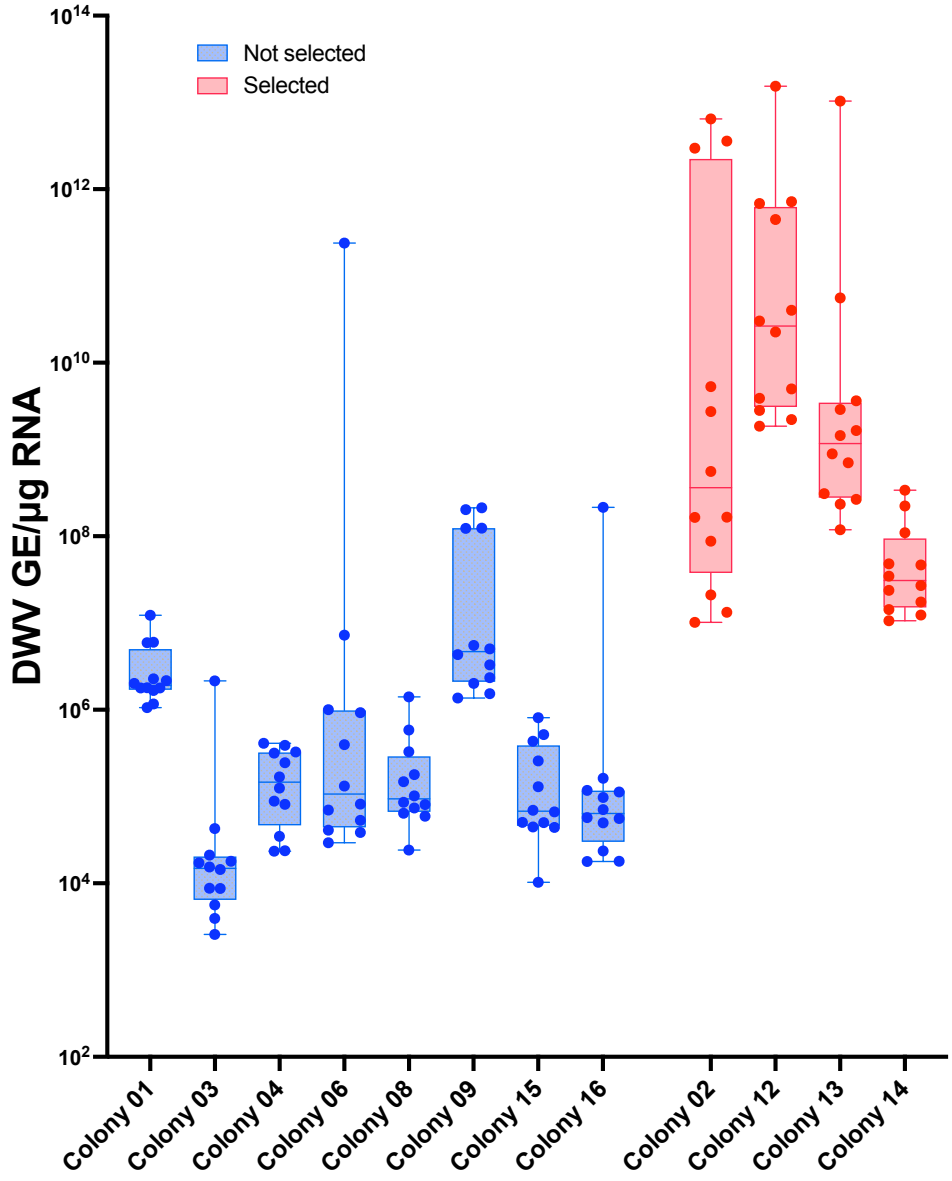


Figure 3-16 - Screening colonies from the infested apiary for base DWV titres before the shook swarm experiment. *Selected colonies shown in red and non-selected colonies shown in blue. Worker brood was sampled from each hive and 12 per colony were screened for DWV using qPCR methods. The colonies selected were chosen as they had the consistently highest level of DWV between individual workers.*

3.3.3 DWV and mite changes post shook swarm

Following the shook swarm and miticide treatments in early June, mite counts were carried out 24 hours after application and subsequently throughout each season whilst the colonies still produced brood. The initial mite drop after treatment application was high in both years (**Figure 3-17A**) as the phoretic mites transferred with the adult workers were killed by the Apivar. In 2018 Colony 16 had >400 mites in the first count, colony 114 had >200, and colony 6 and 13 had ~100 mites. This was likely only 10–20% of the total mites in these colonies at the infested apiary, given that most mites were likely in the sealed brood. By the third count (10 days post-treatment) no mites were observed on the floor of any of the colonies. In 2019,

colonies 13 and 14 had very high levels of dead mites (>2000), and colony 12 had ~450 mites. The number of dead mites observed was considerably lower across all colonies in 2018 (between 60 and 450) than 2019 (between 450 and 2500). Despite these differences, within a few days of treatment in 2019 the mite levels had dropped sharply in all colonies. Unlike in 2018, in 2019 a small number of mites (<10) were observed in the colonies throughout the season. Therefore, a second dose of Apivar was administered to the colonies in late August. This was partly due to the mites observed, but also due to the close proximity of the colonies to other colonies on a shared apiary site, which had an unknown disease history and could potentially infest the colonies by drifting/robbing. Unlike in the first year, these colonies were not moved to the isolated University apiary until after sampling was complete for the season.

Pre-treatment screening of colony DWV levels in May of each year revealed high virus titres in 3/3 2019 colonies and 3/4 2018 colonies, with 6/7 colonies having an average viral yield of $>10^6$ GE/µg RNA (**Figure 3-17B**). Colony 114 in 2018 had low virus titres (~10^3 GE/µg RNA) despite having one of the higher mite-drops for that season. Post-treatment, DWV titres of all emerging brood were immediately at levels considered healthy for a colony (<10^6 GE/µg RNA) and none of the brood had the typical symptoms of DWV – crippled wings or discoloration. The average DWV titres in 2019 remained higher than those of the 2018 colonies, possibly due to the low levels of mites which persisted in these colonies. This may also have been influenced by the variation in meteorological conditions between the two years, as highlighted in section 3.2.6.

The change in viral load from pre-treatment to 1st post-treatment was statistically significant in the 6 colonies with initial DWV titres >10^6 GE/µg RNA (p-value <0.0001, Mann-Whitney test, **Table 3-4**). All emerging brood sampled subsequently had low virus titres and no bees exhibiting typical DWV symptoms were observed for the remainder of the study. Total DWV titres across all colonies for each month were combined into single columns and analysed using a Kruskal-Wallis one-way test for variance (**Figure 3-18**). The virus titre for each month post shook swarm was found to be statistically significant for all sample collection time points.

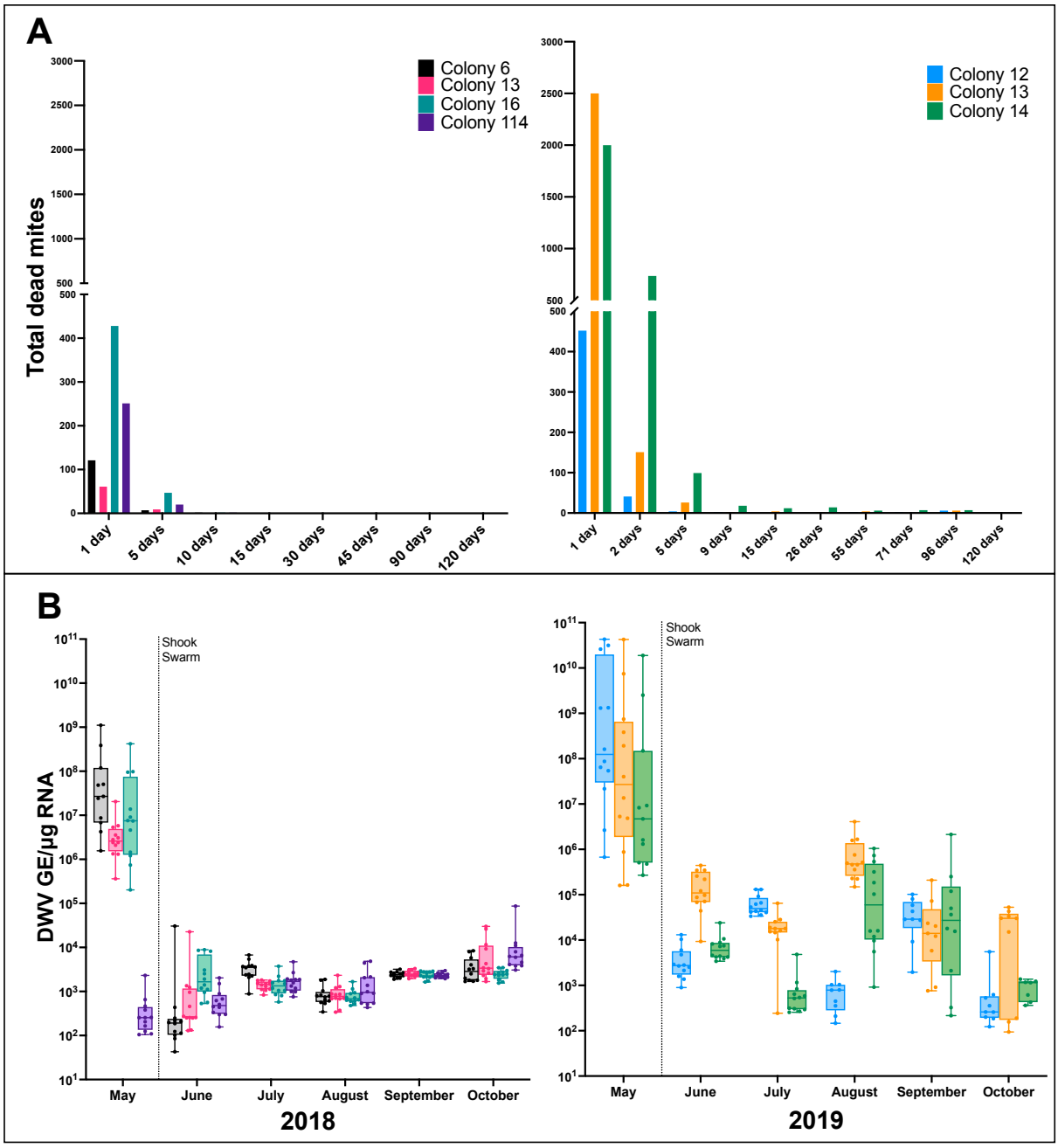


Figure 3-17 - Varroa drop post-treatment and quantified DWV titre in emerging brood before and after shook swarm and Apivar treatments. ***A****. Varroa drop was counted every few days in 2018 (top left) and 2019 (top right) after the shook swarm and Apivar application. Additional time points where no Varroa was observed omitted for simplicity* ***B.*** *Quantified DWV load in workers was measured every month after treatment. Each bar represents the average viral load of 12 workers from a single colony. The dotted line represents the point when the shook swarm was carried out and the Apivar treatment was applied to the colonies.*

Table 3-4 - Shook swarm Mann-Whitney unpaired t-test summary data. *The statistical test was used to compare the DWV titre reported from the brood samples taken before the treatments (May) and in the first brood after treatments (June).*

Colony	Sample time point	P-value	P-value significance
2018 - 6	May-June	<0.0001	****
2018 - 16	May-June	<0.0001	****
2018 - 13	May-June	<0.0001	****
2018 - 114	May-June	0.3164	ns
2019 - 12	May-June	<0.0001	****
2019 - 13	May-June	0.0001	***
2019 - 14	May-June	<0.0001	****

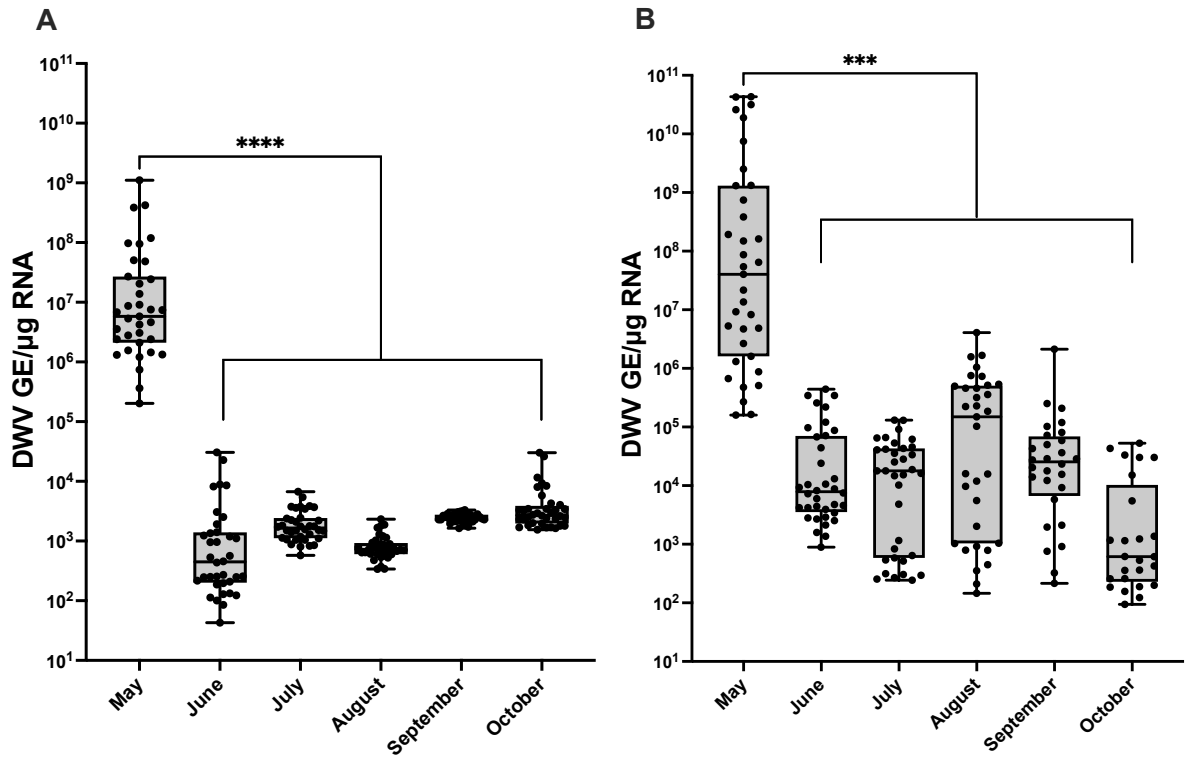


Figure 3-18 – Multiple comparison of DWV titres using a Kruskal-Wallis one-way analysis of variance. One-way Anova Kruskal-Wallis analysis of all DWV samples for each month in 2018 (A) and 2019 (B), using a non-parametric test with samples from May used as the control group. For simplicity only the significance values for the first and last month post-treatment are shown, all other values in the interim months gave the same p-value scores. *** = <0.001, **** = <0.0001.

3.3.4 Sanger sequencing of the virus population

Analysis of DWV diversity was carried out to determine whether the viral variant that was dominant in the colonies when they were highly mite-infested remained dominant, despite the removal of the mites and the subsequent decrease in viral load, or another variant or mix of variants were present.

Initially, sequences from the leader protein region of the DWV genome were generated from May and September of 2019 using standard PCR methods and analysed using the R program SangerseqR (Hill, 2017) to assess diversity in the individual workers. Each sequenced sample produced two sequence outputs, a 'primary' and a 'secondary' sequence match based on the nucleotide peaks produced by the chromatogram. Any position with two peaks would produce two different nucleotides, a dominant base and a secondary one, and these would be assigned to two sequences at the end of the analysis. The output from this analysis indicates whether samples contain a single clonal population or potentially contain sequence diversity. The reads were aligned with reference genomes and a neighbour-joining phylogenetic tree was generated (**Figure 3-19**). In the first time point (orange) each secondary sequence aligned closely to the primary sequence, potentially an indicator of low diversity, but secondary samples from the later time point (shown in light blue), aligned on a unique branch from the primary sequences.

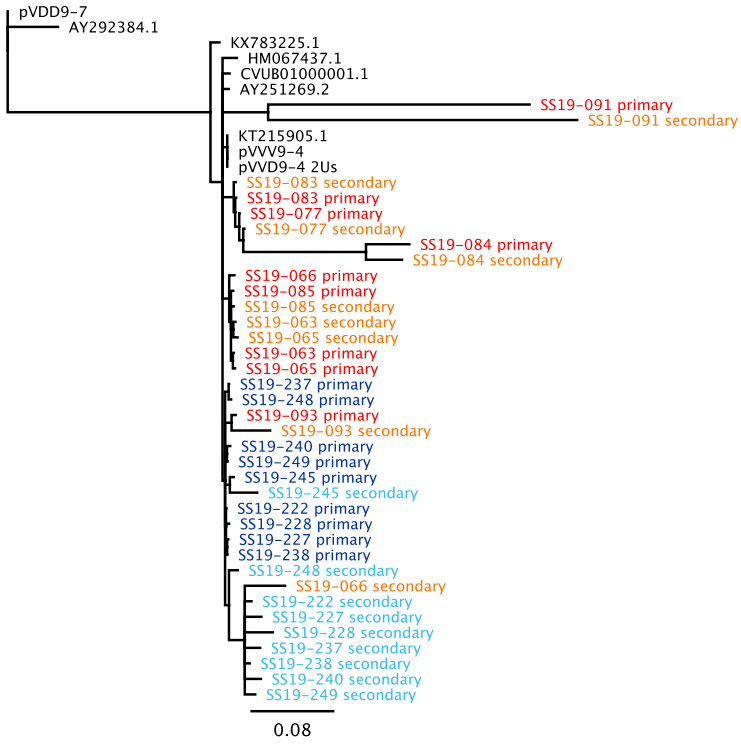


Figure 3-19 - Phylogenetic analysis of leader protein sequences derived from SangerseqR analysis of the 2019 shook swarm experiment. *Samples from May and September were sequenced and analysed, with the primary and secondary sequences shown for each time point and a selection of the closest matching reference sequences from the NCBI database (shown by their accession number). Red and orange indicate primary and secondary sequences from May and navy and light blue indicate primary and secondary sequences from September. The number at the end of each sequence ID refers to the sample ID from all brood samples analysed, e.g., SS19-222 was sample 222 from the 2019 sampling. A 500 bootstrap maximum-likelihood tree was designed using the Tamura-Nei model* (Tamura and Nei, 1993).

Several samples, particularly among those from 2018, had very low viral titres (**Figure 3-17B**, left panel), meaning that generating sufficient sequence data was difficult. Therefore, phylogenetic analysis of only the dominant variants was carried out for a larger number of samples. Sequences were aligned with several DWV reference genomes from the NCBI database. The best-fit model for phylogenetic analysis was determined in Mega7 as the Tamura-Nei model (Tamura and Nei, 1993) and a maximum likelihood phylogenetic tree was generated from 500 bootstrap iterations. From the phylogenetic analysis of 2018 sequences encoding the leader protein region of the DWV genome (**Figure 3-20**), sequences from individual workers from two colonies (6 and 16) in May (TP0) predominantly aligned with Type A-like reference sequences, but sequences from the third highly infected colony (13) aligned with Type B-like reference sequences. The fourth colony (114), which had a low viral titre initially, had a mixed population of low-level variants in the different workers sampled in TP0 (**Figure 3-20A,** shown in green). By September (TP5), workers from colonies 6 and 114 had a mixed population of low-level DWV variants, but sequences from 116 and 13 clustered on a single node, indicating the dominant virus had changed over time in these colonies, but little diversity existed between individual workers from the same colony (**Figure 3-20A,** shown in purple).

Similar analysis of the 2019 samples revealed closer sequence conservation amongst workers from the same colony over time than 2018. All samples clustered in a clade with known reference sequences of Type B variants of the virus, but there was still a change in the dominant viral sequence observed in all colonies over time (**Figure 3-20B**). The sequences from May (navy blue) formed two clades, one for colony 12 variants and another for colony 13 and 14 variants. Pools of *Varroa* (~100 mites per pool) from each colony (shown with asterisks and in bold) were included in the 2019 analysis and the virus sequences from the mites clustered with those from the associated colony workers. The sequences obtained from emerging workers from all three colonies sampled in TP5 clustered on a single clade (orange) distinct from the TP1 variants, but still within the Type-B branch of the tree. Sequences from the same colony typically clustered together, indicating that at a colony level the DWV population contained low diversity.

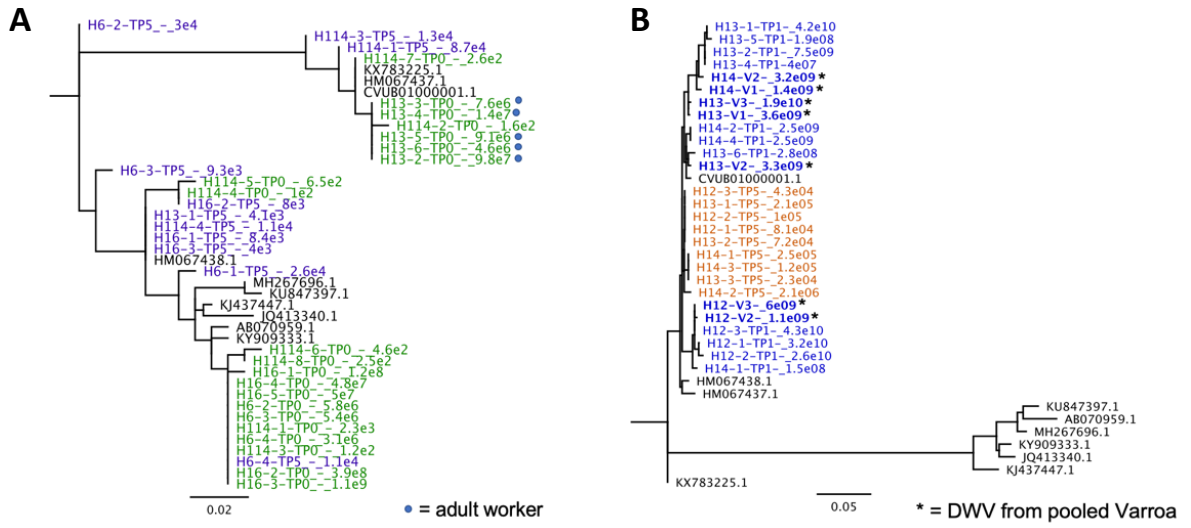


Figure 3-20 - Neighbour-joining phylogenetic analysis of leader protein sanger sequences from the 2018 and 2019 shook swarm experiment. *Samples from time point 1 (TP0/1) and time point 5 (TP5) in 2018 (**A**) and 2019 (**B**) were sequenced and analysed with reference sequences from the NCBI database (shown by their accession number). Blue dots in panel A represent sequences from adult workers where no brood was available. Asterisks in panel B represent virus sequences obtained from pools of dead* Varroa *mites collected after the shook swarm and Apivar application. The number at the end of each sequence name is the quantified viral load by qPCR for that sample (GE/µg RNA). Each neighbour-joining tree was generated from a 500-bootstrap iteration predicted using the Tamura-Nei model (Tamura and Nei, 1993)*

3.3.5 Using ShoRAH to determine sequence diversity changes post-treatment

As with section 3.2.8, Illumina sequencing was used to determine sequence diversity in samples pre- and post-treatments in the colonies from 2019. Close to the full DWV genome was amplified for a selection of samples from TP1 (before treatment), TP2 (first brood post-treatment) and TP5 (late-season brood). 150bp paired-end reads were generated using an Illumina Hi-Seq and samples were analysed to determine the virus population based on unique generated haplotypes using ShoRAH (Short Read Assembly into Haplotypes) (Zagordi *et al.*, 2011). Unlike in section 3.2.8, the 5' end of the sequences did not generate high-quality data for enough of the samples, so the leader protein was not analysed. Additionally, changes to ShoRAH meant smaller windows in the genome had to be analysed (150bp). Regardless of these restrictions, three regions of the genome were analysed: the RdRp, the protease, and a region of known recombination found between the regions of the genome encoding the structural and non-structural proteins.

Figure 3-21 shows neighbour-joining phylogenetic analysis of the haplotypes generated by ShoRAH for the three regions mentioned. Changes in viral diversity are colony-specific, as shown in **Figure 3-21**, where all colony 1 sequences across all three regions cluster with Type B sequences, whilst colony 2 and 3 sequences align with Type A and Type B reference sequences. Colonies 2 and 3, which have haplotypes aligning with Type A and Type B sequences initially, are composed of only Type B sequences by TP5.

Plotting the assigned clusters from **Figure 3-21** as bar plots visualises the changes in the viral diversity in the three colonies after the treatments are applied. **Figure 3-22** shows all Type A-aligned sequences in shades of red and Type B sequences in shades of blue. Colony 1 contained little viral diversity over time, whilst colony 2 showed an initial increase in viral diversity at TP2, before a clear reduction by TP5, something which was also observed in colony 3. Sequenced mite pools (V in each plot) complemented the diversity observed in the colonies: mixed populations of A and B variants in colonies 2 and 3, but only Type B variants observed in colony 1.

Analysis of the SNVs (single nucleotide variants) generated from the ShoRAH analysis indicated that diversity increased with time in all the colonies, as shown in **Figure 3-23** The largest change in diversity was observed in colony 2, similar to the bar plot analysis, and the virus diversity in the mites was similar to the worker bee virus diversity from TP1 for all colonies. Colony 1, indicated by the black dots, showed much higher SNV conservation across all time points than colonies 2 and 3, but still showed an increase in diversity by TP5.

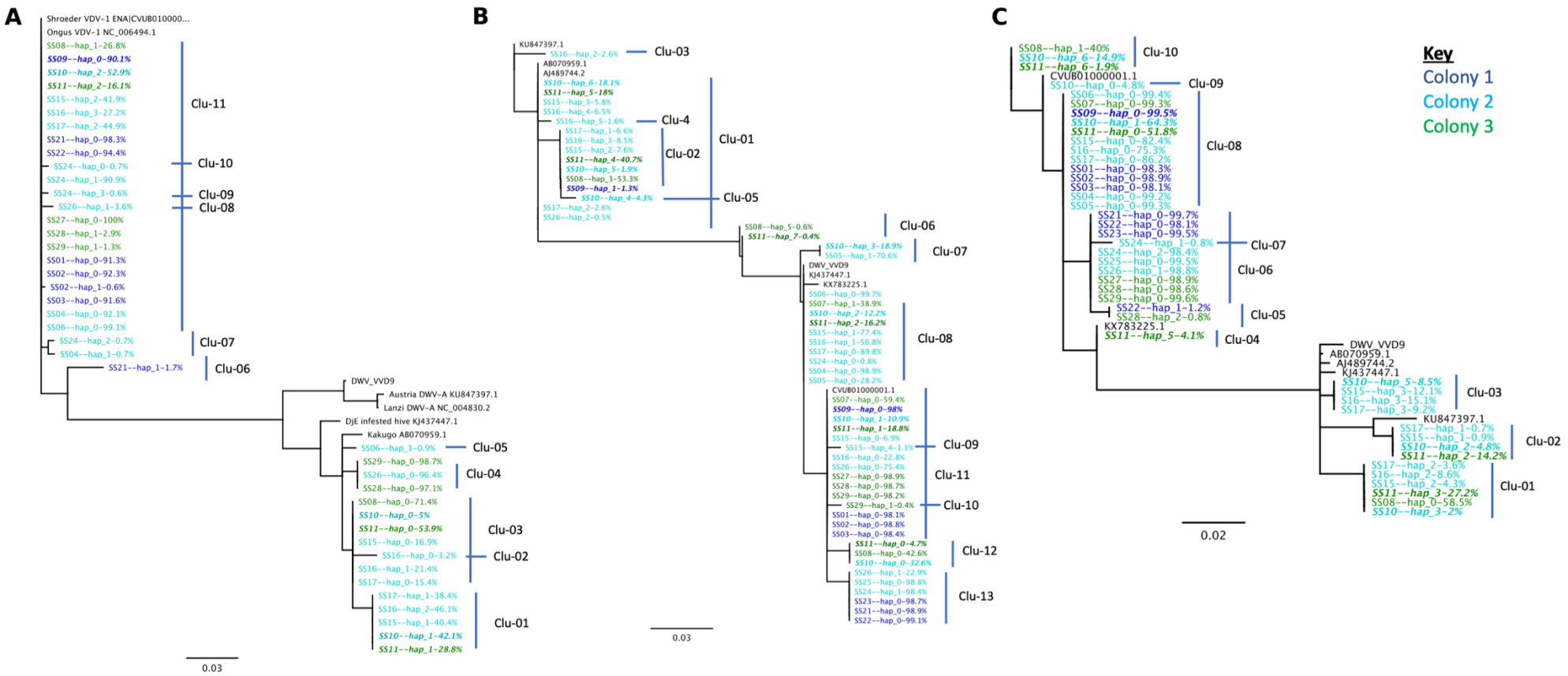


Figure 3-21 - Neighbour joining phylogenetic analysis of ShoRAH haplotypes from 2019 shook swarm colonies. *The (**A**) RdRp (9000–9300-relative position to aligned reference sequence), (**B**) helicase (4850–5300) and (**C**) the protease (8000–8300) were analysed. Colours of sequence determine which colony they were sampled from. Bold sequences are Varroa pools. The hap_x_xx% indicates the haplotype number from ShoRAH and the percentage of the population the sample corresponds to. The blue bars indicate the sequence cluster assignment used for Figure 3-22. Reference genomes are shown in black and represented by their NCBI accession number. The phylogeny were generated from 1000 bootstrap iterations. Sequences were aligned using Geneious.*

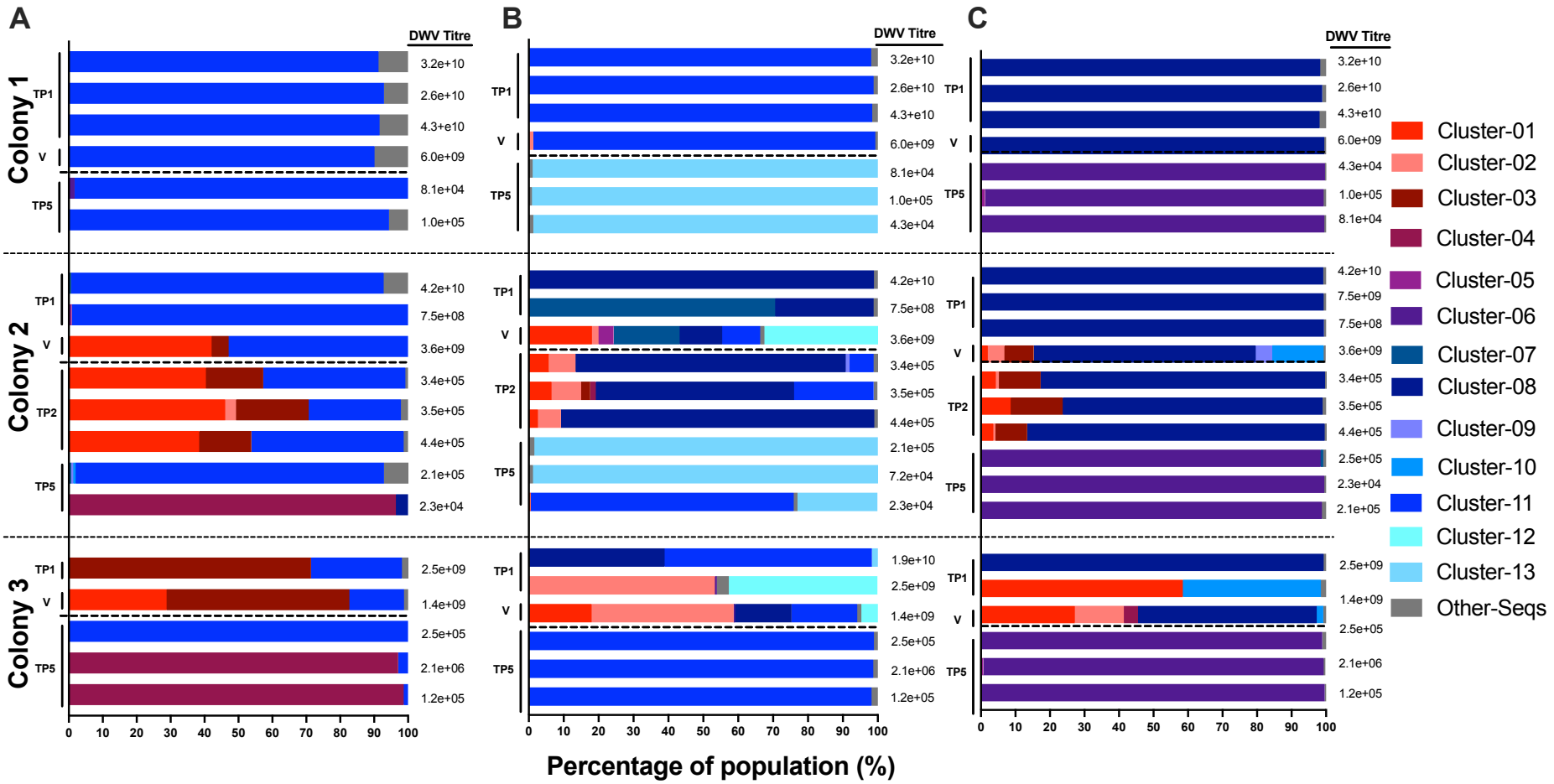


Figure 3-22 - Virus sequence diversity in all 2019 shook swarm colonies. *Haplotypes were generated from next-generation sequence data an assigned to phylogenetic clusters for the (**A**) RdRp, (**B**) helicase and (**C**) protease. The bars are coloured red if they aligned with Type A-like sequences and blue if they aligned with Type B-like sequences in Figure 3-21. The quantified viral load of each honey bee is shown beside each bar.*

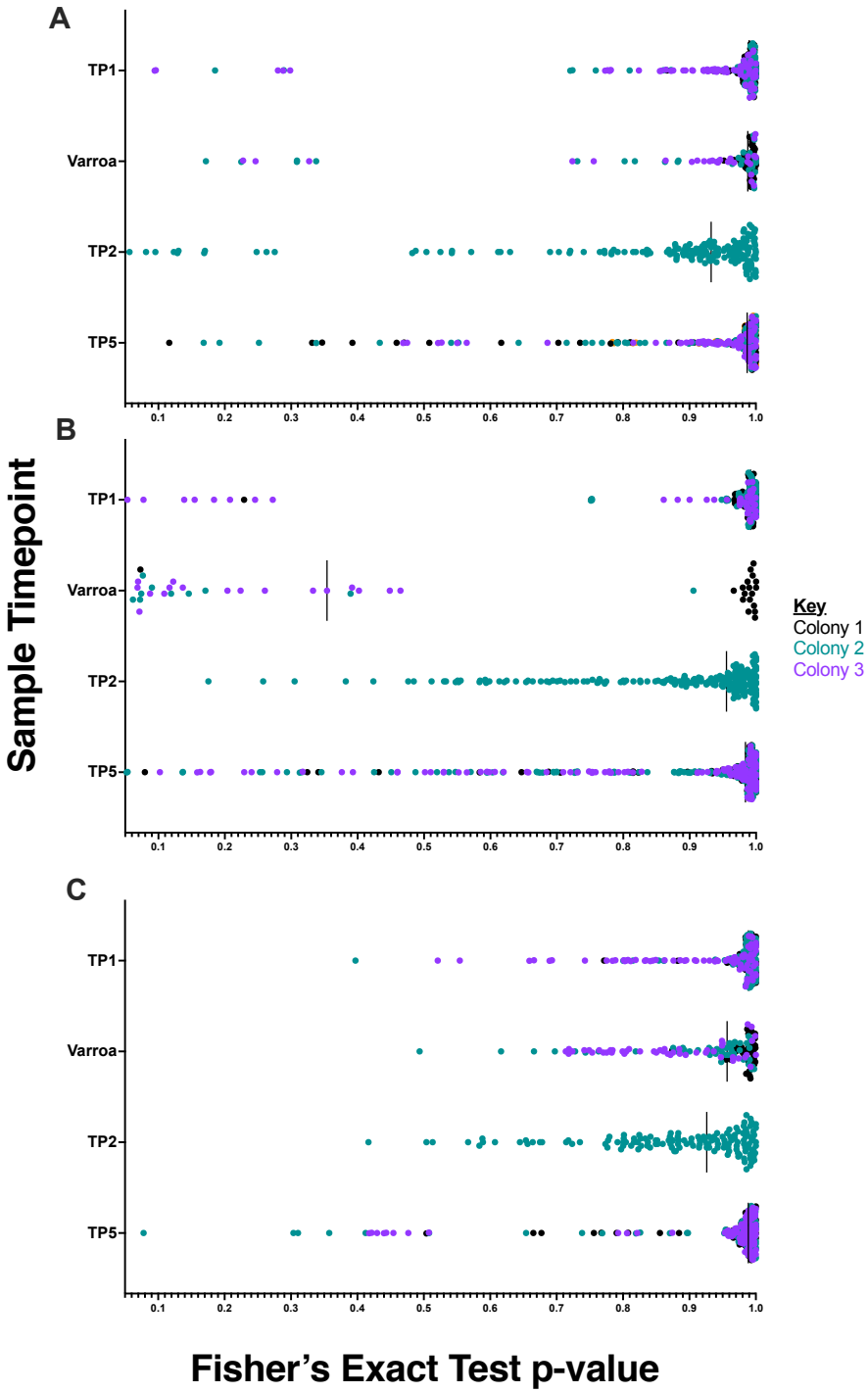


Figure 3-23 - Single nucleotide variants (SNVs) called for samples from ShoRAH analysis of different regions of DWV genome. *SNVs from ShoRAH analysis for all samples from the (**A**) RdRp, (**B**) helicase and (**C**) the protease were analysed if present in 3 iterations of the modelling. SNV p-values close to 1.0 represent a nucleotide variant present in the majority of sequences in that dataset differing from the reference sequence. Samples with SNVs with lower p-value scores therefore have a greater amount of sequence variation within the sample. Anything with a p-value <0.05 was excluded based on threshold for error* (McElroy *et al.*, 2013). *The different colours at each time point represent the SNVs for each colony. Varroa pools are shown separately from the TP1 honey bee samples.*

3.4 Discussion

Since 1992 most honey bee colonies in the UK have been infested with *Varroa destructor,* the ectoparasitic mite which feeds on the haemolymph and fat body (Ramsey *et al.*, 2018) of developing bees and transmits a cocktail of viruses to the bee in the process. It is well known that the mite is responsible for the transmission of a range of viruses, most notably Deformed wing virus (DWV) (Dainat *et al.*, 2012), and that between them, DWV and *Varroa* are responsible for a large number of the overwintering colony losses suffered every year (van Dooremalen *et al.*, 2012). Only by controlling mite infestations will we be able to control DWV infections, which are a key cause of colony losses in infested apiaries. However, although drifting and robbing between colonies were known to occur and to cause the transfer of mites and associated diseases into colonies, the resulting changes in DWV levels and speed of colony failures were unclear. In this chapter there were two aims: firstly to establish the speed of changes in viral titre when mites infest colonies and the effect this has on colony survival, and secondly to measure what happens when the mite population is controlled, by examining whether high DWV titres revert to low-level highly diverse populations of viruses, which rarely cause symptomatic infections.

In the first part of this study, the changes in average DWV titre and colony diversity were analysed over time when well-managed colonies with very low mites and viral loads were placed in highly mite-infested apiaries and left without treatments. In the first and third years of the study viral load increased rapidly over the season and colonies from the first year all died within 18 months. In the second year, DWV titres remained low throughout the season in the colonies moved to the infested apiary in late May (**Figure 3-7**). The two surviving colonies from 2017 both showed increased viral loads as the second season progressed and both died out before October, but samples from a neighbouring infested colony showed similarly low viral titres to those of the new colonies. Based on these observations and the general observation that lower mites were observed across the whole apiary in 2018 (based on the findings from the researchers using the apiary) it is thought that the lower DWV titres reflect the lower abundance of mites in 2018 and therefore a reduced capacity to infest new colonies and transmit virus, perhaps emphasising the importance of coordination between colonies. Further sampling from the local colonies in the apiary, possibly at every collection point, would have given a clearer indication of these changes. The two colonies from 2017 (17-3 and 17-4), possibly weakened over the winter from the high virus/*Varroa* levels observed at the end of 2017 and with surviving phoretic mites already present in the colonies, showed elevated viral titres in the summer of 2018 and died. Colony size is a reliable predictor of overwintering survival (Döke *et al.*, 2019) and the 2017 colonies were smaller than the 2018 or 2019 colonies, increasing their susceptibility to collapse. Some colonies with existing mite infestations can show greater tolerance to further mite infestation by drifting (Forfert *et al.*, 2015), so this may also be a factor in the apparent discrepancy between the 2017 and 2018 colonies in the second year.

The difference in average temperature year on year (**Figure 3-8**) provides some insight into colony differences, as a prolonged and colder winter between 2017 and 2018 may have killed a lot of the phoretic mites on the overwintering adult bees. It is known that mite infestation is positively correlated with higher colony temperature (Hou *et al.*, 2016). The warmer winter between 2018 and 2019 was followed by high mite abundance, as shown by the large difference in mite drop after the 'shook swarm' method between the two years (**Figure 3-19a**). However, as no mite counts were performed at the mite-infested apiary throughout this study the changes in abundance are unknown. Future experiments monitoring colonies over multiple seasons

could use techniques like a sugar roll or ethanol wash of bees to count phoretic mites from the positioned colonies and the local infested colonies to get approximate numbers of mites at the site over time.

The warmer weather and potential boost to mite survival was reflected in a rapid increase in DWV titres in the 2018 colonies during their second season on site and the newly placed 2019 colonies in the infested apiary. We can therefore deduce that a weather-induced reduction in mite abundance may prolong colony survival, but that as soon as mite abundance does increase the colonies will be highly susceptible to phoretic mite infestation and subsequent increased viral load (Greatti, Milani and Nazzi, 1992; Fries and Camazine, 2001). This study used changes in DWV titre as the measure of colony health, therefore the absence of mite data does not reduce the validity of the results, as sharp increases in DWV titre were typically followed by colony deaths.

In the second part of this study, we aimed to show that with a combination of the 'shook swarm' beekeeping technique and an effective miticidal treatment, Apivar, we could rapidly remove *Varroa* from highly infested colonies and measure the subsequent changes to the DWV population. The results from 2018 and 2019 (**Figure 3-17**) showed that the combination of Apivar and a shook swarm significantly reduced the number of mites present in the colony and reduced the viral loads to those of a well-managed, low-mite colony by the first brood reared post-treatment ($<10^6$ GE/µg RNA). Given the differences in mite levels and virus titres between colonies in 2018 and 2019 in the first part of this study, it could be assumed a similar effect was likely here, but other than the single colony in 2018 which had very low DWV titres throughout and a consistent presence of mites at a low level throughout the season in 2019, there was no significant difference in the outcome between the two years and seven colonies.

In both years, all colonies which had highly elevated virus loads prior to treatment had low-level DWV titres by the first brood cycle post-treatment, indicating that the surviving adult bees from pre-treatment did not pass on high viral loads to the developing larvae by oral transmission. Therefore, we can conclude that the viral titre passed from the infected adults to the developing larvae was below the approximate titre required to cause infection in larvae ($>5\times10^6$ GE), based on previous work using infectious clones of different variants to study DWV thresholds during larvae feeding (Gusachenko *et al.*, 2020a). This is perhaps surprising given that the bees analysed prior to treatment had viral titres between 10^7 and 10^{10} GE/µg RNA and therefore it would be expected that the workers that fed the first generation of pupae would pass on relatively high viral titres to some of the developing brood. This highlights the differences in the tropism of the viruses between different transmission routes, with parasitised bees likely to have DWV throughout their body (Fujiyuki *et al.*, 2004; Yue *et al.*, 2007; Gusachenko *et al.*, 2020a). Despite the adult workers carrying high viral loads, it is possible that the virus in their guts was sufficiently low to not cause elevated virus in the next generation, or that the bees with the highest viral titres were not involved in brood rearing. The DWV titres in all seven colonies remained low for the remainder of the sampling process, indicating that the mite removal from the colony was highly effective and removing this transmission route was a sufficient means for controlling DWV titres.

ShoRAH analysis of the next-generation sequencing data from the 2019 shook swarm (**Figure 3-21/22**) showed that the viruses present prior to the treatments were not present in the bees sampled by the end of the season in 2 of the 3 colonies. The virus populations present in the mites revealed a mix of Type A and Type B variants in colonies 2 and 3 which mirrored the diversity in the bees. Interestingly, a mixed population of variants was observed in colony 2 in

the first brood sampled post-treatments, indicating that low-level variants transmitted by feeding are presumably present to some degree in the workers reporting high levels of near-clonal virus. As the mites were pooled for analysis it cannot be determined whether the diversity observed is unique to individual mites or a reflection of different mites carrying different variants.

Reports of differences in viral replication rate between DWV variants have suggested that Type B may initially replicate faster than Type A variants during viral infections, before both reach similar titres (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a) and in some instances Type B remains higher (Norton *et al.*, 2020). Here, we observed both variants were present in the infested neighbouring colonies and the different variants were present across different years when mites infested the *Varroa*-naïve colonies (**Figure 3-9** and **Figure 3-10**). The presence of both variants in the infested colonies indicates that survival or tolerance of honey bee colonies does not appear to be variant-specific (Mordecai *et al.*, 2015), as these colonies are never treated with miticides and across multiple years both variants were observed. Interestingly, the virus type changed across sample time points (**Figure 3-9**) in some colonies regardless of titre, with a range of Type B samples observed (10^4–10^{11} copies), perhaps indicating that a faster replication rate is not the only factor involved in this virus population shift.

The data across the three seasons of mite-exposure suggests that the virus population is constantly in flux, as a mix of Type A and Type B were found in 2017 (**Figure 3-9**) and in late 2019 (**Figure 3-10**) but in early 2019 the dominant variant was Type B (**Figure 3-20B**), perhaps supporting the theory of Gusachenko *et al* (2020) and Tehel *et al* (2019) that Type B replicates faster initially. It is likely that the dominance of one variant over another is a temporal and spatial issue, where the first injected virus type passed from mites to pupae can become highly elevated in pupae and dominates the virus population before other variants can become established (Gusachenko *et al.*, 2021). Over time this dynamic can change dependent on the mites infesting the colony, the types of variants they carry and the selection pressures the viruses undergo. This seems particularly relevant given the shift towards Type B in some colonies in 2017, despite the low titres, indicating a potential selection advantage for this variant in the mites. Recent reports that Type A does not replicate in *Varroa* (Posada-Florez *et al.*, 2019), but recombinants can (Gusachenko *et al.*, 2020a), could explain this change.

High titres of DWV have been observed in honey bees which display no phenotypic signs of infection (Gusachenko *et al.*, 2020a), although they may have a reduced lifespan (Dainat *et al.*, 2012). We investigated whether the bees observed in this study with highly elevated DWV titres all contained the same virus type and found that this was not a factor in determining the phenotype of infection. When using ShoRAH to examine viral diversity we observed bees from the local infested colony which contained a mixed population of viral variants despite carrying very high viral titres (10^{10} GE/µg RNA) (**Figure 3-11** and **Figure 3-13, Colony IN**), contrary to work suggesting mite infestation and elevated viral loads culminate in near-clonal virus populations (Martin *et al.*, 2012; Ryabov *et al.*, 2014). Although the majority of samples from the colonies placed on site had near-clonal viral populations by the end of the season (5/6 samples), one of these also contained a mixed population of variants (**Figure 3-13**, **17-1 – 17/9**), findings supported by recent research examining the changes in DWV diversity between high and low infections in mite-parasitized honey bees (Annoscia *et al.*, 2019) and an examination of diversity in wild-type infected honey bees in USA (Ryabov *et al.*, 2019), where both studies found bees with high viral titres still contained mixed DWV populations. Additionally, some of the individual workers and pooled mites analysed prior to carrying out the ‘shook swarm’ contained mixed populations of Type A and Type B variants (**Figure 3-22**),

despite the high viral titres reported by qPCR (**Figure 3-16**). Whilst it is possible that mites only contained individual variants, as they were analysed as a pool, multiple mites feeding on individual pupae may transmit different variants which can establish in the host in co-infections. Type A and Type B variants have been shown to replicate to equally high titres over time during *in-vitro* injection co-infection experiments (Gusachenko *et al.*, 2021), so it is possible that mites are injecting variants of both types which can establish in pupae, resulting in mixed population infections and potential recombination events.

It could be that viral diversity, even at a higher viral titre, is a factor in colony survival, as in the colonies permanently located in the infested apiary which survive with no treatments and minimal hive management. Historically, DWV diversity has been examined at a colony level, by examining pools of workers (Martin *et al.*, 2012), but the analysis from the infested colonies in 2017 (**Figure 3-13**) indicates that some workers will have mixed viral populations and some will contain near-clonal variants, so whether a particular variant or a mixed population is the cause of colony death is not known. However, not enough samples from the infested colonies were analysed during this study to fully determine this.

The findings from the ShoRAH analysis, where a large percentage of sequence data was assigned to 'other' variants that occurred at a low level in the samples, support the theory that mutant clouds of viral variants are present during infections, with low-frequency mutants present at one point in a population increasing in frequency over time and potentially influencing the pathogenesis of the virus quasispecies (Domingo and Perales, 2019). Some of these low-level variants, as observed at TP0 in both colonies, persist throughout the season in the viral population. Other variants, which appear at certain time points but are then seemingly lost in later viral populations, may have reduced viral fitness or belong to defective genomes which are capable of exploiting host interactions, but are not sustained over time (Rezelj, Levi and Vignuzzi, 2018). These changes over time may be due to a selective sweep, whereby an introduced favourable mutation increases in frequency and becomes fixed in the population (Elena, Cooper and Lenski, 1996); this may become more common during recombination events (Ryabov *et al.*, 2019). There is evidence of the presence of a recombinant or recombination occurring between existing variants in the shook swarm virus populations, even when the titres are low, and the virus is presumably isolated to the gut of the workers. RdRp sequence alignments in the final time point typically clustered with Type A sequences for two colonies (**Figure 3-20**), whilst other regions of the genome aligned with Type B sequences.

These fluctuating viral variants are indicative of a population undergoing selective pressure, possibly passing through a bottleneck, whereby an initial population increases as new variants are introduced (during mite feeding, or *per os* from workers), before elevation of a selection of these variants through further mite feeding or injection results in a decrease in diversity, potentially to near-clonality (Ryabov *et al.*, 2014). The SNV analysis of the infested colonies suggests an increase in viral diversity by TP2 in both colonies, but by TP3, as the titre increases further, the diversity is reduced (**Figure 3-12**). This could be due to a large number of new variants introduced by mites before selective pressure and competition results in a reduction in diversity. These shifts in virus populations can result in random drift in evolutionary outcomes and potentially negative fitness effects for the virus population (Chao, 1990; Bull *et al.*, 2011). Additionally, biological constraints can impose negative selection on the generation of a lot of new mutants after a bottleneck event, resulting in a general maintenance of the overall virus population identity (Domingo, Sheldon and Perales, 2012; Agol and Gmyl, 2018). This may explain the near-clonal populations observed in individual workers from the newly infested colonies by the end of the season, as new mutants are lost over time (**Figure 3-15, TP3**) and

potentially explains the lack of variation observed in colony 1 after the shook swarm (**Figure 3-22**). This suggests that the transmission route of worker to larva *per os* does not increase DWV diversity and that the removal of the mite transmission route may also remove any bottlenecks imposed on the diversity of the virus population, allowing the dominant variant to persist.

The next-generation sequencing results show that standard sequence analysis of a consensus sequence, or a dominant variant, is not sufficient when analysing a DWV population over time. Equally, pooled samples of workers from colonies may mask the true diversity of the virus population in a colony, resulting in inaccurate assessments of colony health. A range of viral mutants will define the consensus sequence generated, but the consensus sequence itself is an abstraction and may not even exist as a viable virus in the population of variants (Domingo and Perales, 2019). Previous reports of experiments using 'wild-type' isolated Type A or Type B inoculants to test virulence in developing honey bees (Tehel *et al.*, 2019; Norton *et al.*, 2020) define these wild-type populations as a single DWV-type and compare them as such, when they will almost certainly contain a subgroup of variants not detected by standard sequencing techniques, which may either inhibit or increase the virulence of the inoculum. There are many examples of pathogenesis and evolution of viruses due to low-level variants or mutant modifications that cannot be correctly interpreted based on the consensus sequence of the population (Domingo *et al.*, 1978; Figlerowicz *et al.*, 2003; Domingo, Sheldon and Perales, 2012; Lowry *et al.*, 2014; Xiao *et al.*, 2016; Moreno *et al.*, 2017).

Additionally, high viral titres of mixed virus populations may be more capable of adapting to a new host as different variants within a virus quasispecies may have the adaptive changes required to adjust to a new host, such as increased transmission or evading the immune response (Longdon *et al.*, 2014). In this context these findings may shed some light on reports of DWV infections in other insect species which share a niche with honey bees but are not parasitized by *Varroa*, such as bumble bees (Fürst *et al.*, 2014; McMahon *et al.*, 2015) or hornets (Dalmon *et al.*, 2019). Comparison of virus infections that evolved in different host species found that parallel genetic changes were more likely to occur if the two host species were closely related (Longdon *et al.*, 2018). This is particularly of interest as studies using reverse genetics clones to infect bumble bees failed to reproduce the classic phenotype of DWV infection, crippled wings, and DWV was only infectious in injected bumble bees at very high doses ($>10^8$ GE/µg RNA) (Gusachenko, *et al*., 2020b), making it unclear at present how DWV is able to infect other pollinators sharing the same ecosystems as honey bees.

These studies highlight that controlling the mite is essential when colonies are close to other infested colonies, as mite uptake is rapid and the DWV titre and population shifts dramatically within a single season, killing colonies. These findings are not surprising but highlight the necessity of managing and treating colonies appropriately. However, effective beekeeping practices shown here, such as the shook swarm technique combined with appropriate miticide treatments, can very efficiently reduce *Varroa* infestations. This consequently returns the virus population to a low level, like that seen in colonies that are well managed, overwinter successfully and do not show signs of DWV pathogenesis. This is a clear practical solution to high mite infestations which can be performed by any beekeeper.

4 Investigating the impact of coordinated treatments for *Varroa destructor* on honey bee colonies on the Isle of Arran

*'To know fully one field or one land (*or one island!*) is a lifetime's experience' – Patrick Kavanagh.*

4.1 Introduction and aim

Coordinated treatment programmes, whereby a treatment or a technique is coordinated over a geographically defined area to the benefit of the individuals within that area, have a long history of use for successful disease control and eradication. In 1906 the Cattle fever tick eradication programme (CFTEP) was established in the USA to control the two ticks *R. microplus* and *R. annulatus*, and by 1943 the US was declared tick free (Pérez de Leon *et al.*, 2012). The coordinated miticide effort, dependent on a coumaphos treatment regime, was partly responsible for the subsequent increased development and productivity of the US cattle industry.

More recently, Zika virus, a pathogen vectored by mosquitoes, which can cause microcephaly when transmitted from mother to unborn child, was controlled in a coordinated manner in Puerto Rico using an integrative vector management (IVM) system to increase public awareness, trap mosquitoes and perform larviciding (killing the mosquito larvae). This coordinated effort was based on communicating information to the public, rather than applying treatments, but resulted in a significant decrease in mosquito density in these areas, reducing the risk of disease transmission (Barrera *et al.*, 2019).

One of the best-known coordinated treatment programmes in the UK is the treatment of sea-lice in managed salmon fish farms. The sea-lice have free-living larvae that are capable of moving between farms via currents, and as such coordination of treatments between the fish farms is critical to control the lice (Murray and Salama, 2016). Modelling data has indicated that coordinated treatments reduce the number of treatments required to keep the lice population under control on the farms, which reduces costs and improves treatment efficacy.

These examples indicate that there is a precedent for using coordinated management techniques to control the spread of pathogens and vector-transmitted diseases, regardless of the organisms involved. These coordinated programmes can take different forms: applying a chemical to multiple sites in unison, as with the fish farms, or dissemination of information to the public to improve understanding, as with the mosquito control programme. Previous attempts to coordinate treatments for *Varroa* infestations of honey bee colonies have been unsuccessful or carried out on too small a scale to be realistic and have a meaningful impact. These were an attempt on the isle of Jersey (Sampson and Martin 1999) and on Gorgona, a small island off the north-east coast of Italy (Giusti *et al.*, 2016). In Jersey an apiary was treated, but continued to be mite-infested throughout the study, possibly due to other feral or managed colonies on the island, and on Gorgona there was only a single apiary, making mite eradication a considerably simpler process. Therefore, a thorough long-term study of coordinated treatments of colonies, across multiple sites in an isolated environment of an appropriate size is required.

The aim of this study was to use a coordinated miticide treatment to remove or reduce *Varroa destructor* infestations in all honey bee colonies in a geographically isolated environment and to measure the effectiveness of this coordination through changes in Deformed wing virus (DWV) population over time. This is based on our understanding that *Varroa* infestations result

in significant increases in DWV titre and decreases in DWV diversity and consequently increased colony losses (Highfield *et al.*, 2009; Martin *et al.*, 2012).

The hypothesis was that using coordinated treatments in an isolated environment would significantly reduce mite infestation and improve overall colony health. The Isle of Arran, off the west coast of Scotland, provides an ideal site for testing coordinated treatments of *Varroa* infested colonies. The island is at least 3 miles away from the mainland at all points and, because bees will not fly across such a large body of water, there was no chance of bees drifting from the mainland into the island's apiaries. The island also contains a suitable number of colonies to make the experiment viable but not so many as to make it unmanageable. The intention was to treat all colonies on the island in unison, to reduce the likelihood of mite transfer from highly infested colonies to healthier, neighbouring colonies post-treatment. The study was carried out over three years, with late-summer and winter miticide treatments applied to all known honey bee colonies in collaboration with the local beekeepers, who would benefit from a better understanding of disease transmission and control in honey bee colonies.

The beekeepers on the Isle of Arran lost all their colonies to a suspected combination of high *Varroa* infestations and very high DWV titres in 2013. After extensive conversations regarding improving bee health with Fiona Highet from SASA (Science and Advice for Scottish Agriculture) and Graeme Sharpe from SRUC (Scotland's Rural College), the beekeepers were keen to participate in a study which could improve the health of their honey bees and their beekeeping practices. One condition of the study was that no packages or hives of bees were brought on to the island once the coordinated treatments had begun as it would compromise the control of the experiment. The beekeepers agreed that the only imports would be queens if they were needed for making new colonies, otherwise they would just split the colonies they had on the island.

4.2 Colony movement and mite abundance

4.2.1 Site visits, treatments and sample collection

Visits to the island were carried out when the appropriate people were available to facilitate field work and at a time which was suitable for the experiments and the beekeepers. Visits to collect samples were typically coordinated with meetings of the beekeepers' association (known as the Arran Bee Group), to allow the sharing of information and updates about the project. **Table 4-1** outlines when the various aspects of the project were carried out. Apivar treatments were placed in the colonies by the beekeepers after the removal of honey. The beekeepers agreed a date for this treatment among themselves and communicated it to us. The mid-winter treatment did not require coordination as bees are not foraging, and therefore not drifting/robbing, in mid-winter.

Table 4-1 - Site visits, treatments and sample collection on the Isle of Arran.

Year	Honey bee collection	Treatment application (Apivar)	Mite collection	Treatment application (Api-bioxal)
2017	August	1st week - September	7-8 days post-treatment	December
2018	June	1st week - September		December
2019	May & August	1st week - September		December/January 2020

4.2.2 Colony movements throughout the three years

As no new colonies were brought to the island for the duration of the project, the beekeepers split strong colonies to create nucleus colonies (small-scale colonies typically containing 5 frames of bees) for new beekeepers or to build up their own stocks. Consequently, they moved colonies around the island quite frequently. On the first visit to the island, apiaries were assigned to 'Sites' and coloured accordingly, as shown in **Figure 4-1**. The Site assignment, where a 'Site' was deemed a geographically distinct area, was based on the relative distances between the apiaries and was typically each settlement on the island with bees. Most apiaries were managed and occupied by a single beekeeper, so some site-to-site variation will be explained by differences in beekeeping skills. In 2017 there were 6 sites with 17 apiaries containing a total of 57 colonies, but a large number of the colonies (29/57) died by the spring of 2018 (see **Figure 4-2**). Colony movements were then required to allow beekeepers to recover lost colonies. The movements in 2018 and 2019 are indicated by the coloured arrows in **Figure 4-1**. The colour of the arrow indicates the site the colony was moved from. In 2018 the Arran Bee Group apiary at Site 3 had several surviving colonies and as a result a lot of stocks were prepared from there and distributed for other beekeepers (yellow arrows). Additional colonies were split from another group apiary located at Site 5 (dark red arrows). The following year the same apiary at Site 5 was used to supply colonies to other apiaries, typically made up from swarms caught in bait-hives throughout the season. In one case, an apiary located on the southern edge of Site 5 was moved to a new location, as indicated by the dark red arrow in the south-east in 2019.

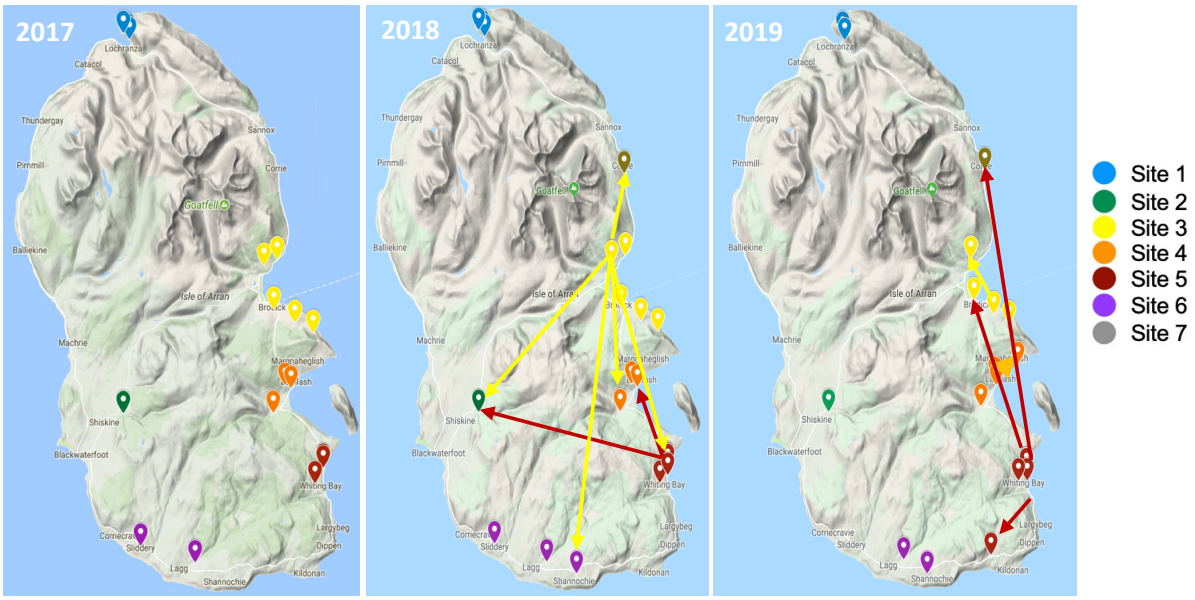


***Figure 4-1 - Hive locations and movements throughout duration of study**. From left to right, apiary locations in 2017, 2018 and 2019 on the Isle of Arran. Each marker indicates a single apiary and the colours indicate the 'Sites' they belong to. The arrows on the 2018 and 2019 maps indicate the movement of colonies from one site to another, with the colour of the arrow indicating where the hive has been moved from. The area of the island is 428km^2 and it is approximately 16km from east to west and 32km from north to south.*

Figure 4-2 indicates the change in the total number of colonies at each apiary during each sampling trip on the island. Each block in the bars represents a single apiary and the 'site' is coloured as per the maps of the island in **Figure 4-1**. An initially large number of colonies in 2017 (57) were reduced to 28 in June of 2018, after the overwinter loss of colonies (n=29).

Colony numbers increased to 54 by August of 2018, 69 by May 2019 and 84 by August 2019 as the beekeepers utilised splitting their colonies and capturing swarms to increase the number of colonies on the island. Several new beekeepers joined the group throughout the study, increasing the club size from 17 in 2017 to 22 by 2019. To the best of the Arran Bee Group's knowledge, all beekeepers on the island were members.

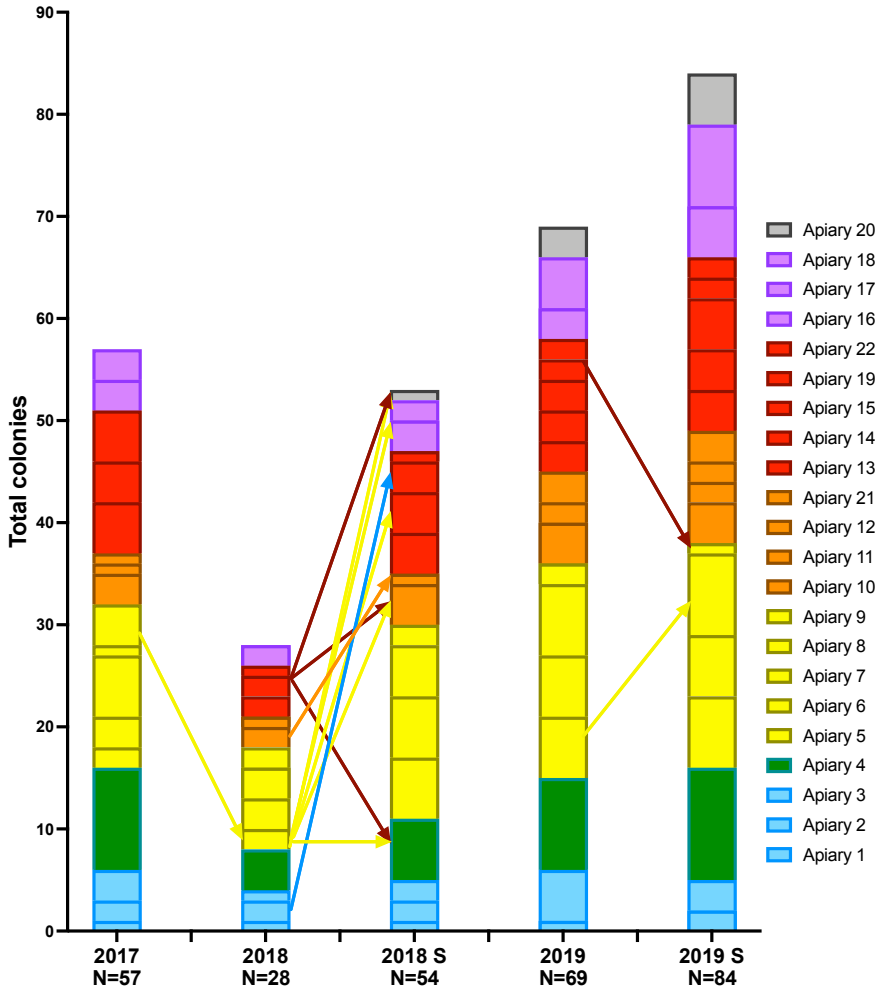


Figure 4-2 - Colony numbers at each honey bee sample collection point between 2017 and 2019. *Each block represents an apiary, all are coloured by 'Site'. Arrows indicate colonies or splits moved from one apiary to another between visits, coloured by the original apiary 'Site'. The total number of colonies at each point is shown below the x-axis as N = x. The S after 2018 and 2019 indicates the second visit in August.*

4.2.3 Measuring mite drop changes after treatments over three years

In September of each year (2017–2019) all colonies were treated with Apivar for six weeks (see section 2.1.3). Briefly, plastic strips coated with the miticide were placed in all the colonies by the beekeepers once the honey harvest had been removed and all within a 24-hour window. A mesh tray was placed on the floor of each hive at the same time as the miticide application and the dead mites dropped onto the trays as the treatment took effect. At the end of the first week the beekeepers collected these trays and sealed them up. The mites were then counted for each hive. Mite counts over 1000 were estimated for some colonies by dividing the collecting tray into sections and multiplying the count from sub-sections. Based on the effectiveness of

the miticide on *Varroa* (Rinkevich, 2020), the ease of application for a group of beekeepers with a range of abilities, and the effects of the treatments observed on mite numbers following the shook swarm experiment (Chapter 3), the treatment and duration were considered the most effective method for this experiment.

Figure 4-3 shows the mite drop for each Site in each of the three years of sampling. The mite numbers recorded only represent a proportion of the total mites in the colonies and are not standardised based on colony size and strength, which varied throughout the apiaries. Site 1 showed a gradual reduction in mite abundance, with one colony in 2017 reporting >1500 mites, in 2018 all colonies were below 1000 but some remained high, but in 2019 no mites were recorded in any colonies. It is unlikely that this means no *Varroa* are present at this site, but does indicate a substantial decrease in the mite population. Several colonies died in 2018 and the new colonies split from the survivors remained healthy. One beekeeper stopped keeping bees between 2018 and 2019 and no other beekeepers joined the area to the best of our knowledge. Site 2 consisted of a single large apiary which reported very high mite levels in most hives in 2017 (4 colonies had >2000 mites in 2017) and had high colony losses over the 17/18 winter (8/10 colonies died). In 2018 the mite abundance reported was considerably lower, but following some colony movement in late 2018 (**Figure 4-4**) the mite levels were high in a number of colonies in 2019 (1 colony >1500 mites). Some of the colonies in 2019 had enough mites to suggest the colonies would have very high virus loads and be at risk of collapsing over the winter.

Site 3, location of the Arran Bee Group apiary, which is used to split and generate new colonies for member beekeepers, reported very low *Varroa* levels in 2017 and 2018 (below 250 mites in all colonies), but an increase in 2019 (3 colonies with >500 mites). Site 4 and 5 contained a large number of apiaries and colonies, many of which died during the study resulting in replacements being sourced and necessitating colony movements. These sites also had several new beekeepers joining the group during the study, requiring new colonies. Both Sites contained relatively high mite numbers throughout the study, particularly Site 5. Both Site 4 and 5 are also relatively close to Site 3 (Site 4 is <5 km from Site 3) and may account for the observed changes in mite abundance there in 2019 (**Figure 4-1**). Site 5 had the highest mite abundance in all three years and ended 2019 with a higher mite drop than 2018, with three colonies reporting over 1500 mites within a week of treatment.

In stark contrast to those heavily populated sites, Site 6 contained 3 apiaries spread over a large area (~15km between two furthest apart apiaries) which was well isolated from the rest of the Sites. These apiaries all reported very low *Varroa* prevalence throughout the study. Site 7 was a new apiary, so data was limited, however one colony had >300 mites by the end of 2019.

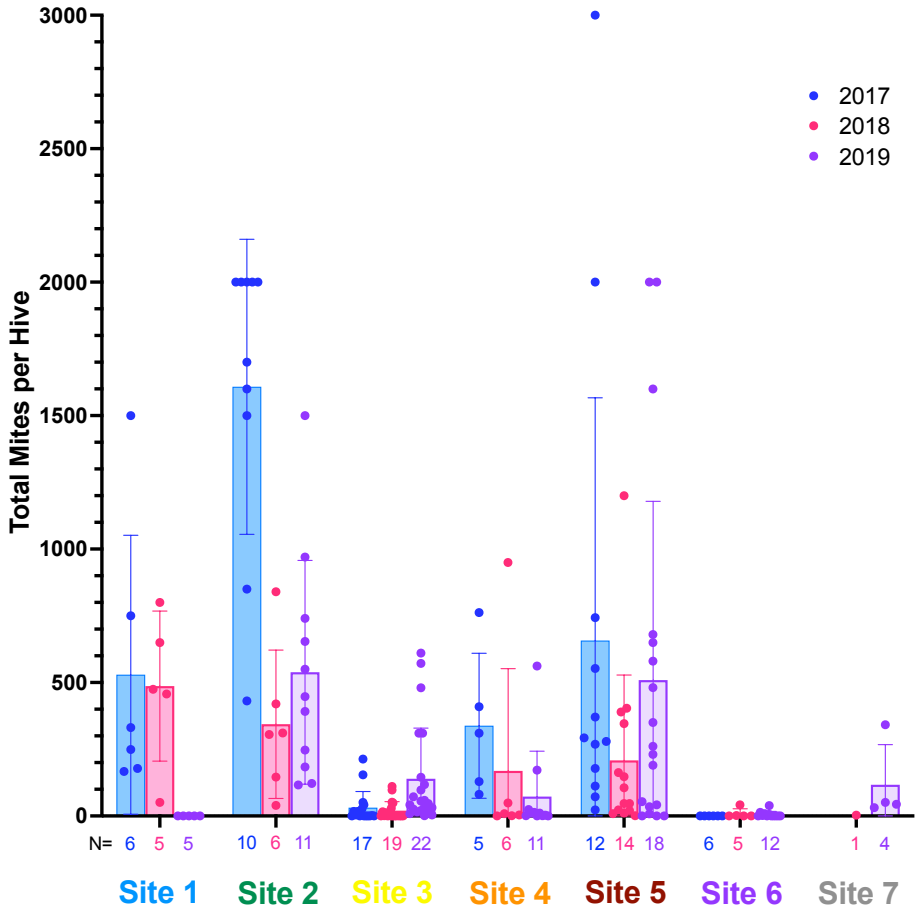


Figure 4-3 – Mite drop 7 days post-treatment for each Site. Each value shown here is the number of dead mites observed in a single colony post Apivar treatment. The error bars represent the standard deviation of the mean. No data shown for Site 7 in 2017 as they joined the study in late 2018. N= x below the x-axis indicates the number of colonies at each Site when the treatments were applied, coloured by year.

The percentage mite drop for each Site as a proportion of the total drop on the island was calculated for each year (**Figure 4-4**). A relatively even distribution of mites between Sites 1, 2, 4 and 5 in 2017 shifted to a much higher percentage of mites observed at Site 5 by 2019 (over 50% of all mites observed in 2019). Despite the increase in mite abundance in 2019 from 2018, the total number is considerably lower than observed in 2017 (18,506 compared to 29,371) and the total number of managed colonies increased significantly (57 colonies in 2017 to 84 colonies in 2019). **Table 4-2** shows the average mite drop per colony each year and by the final sample point the average number of *Varroa* observed in each colony drop had decreased from 544 in 2017 to 228 in 2019, a 58% reduction.

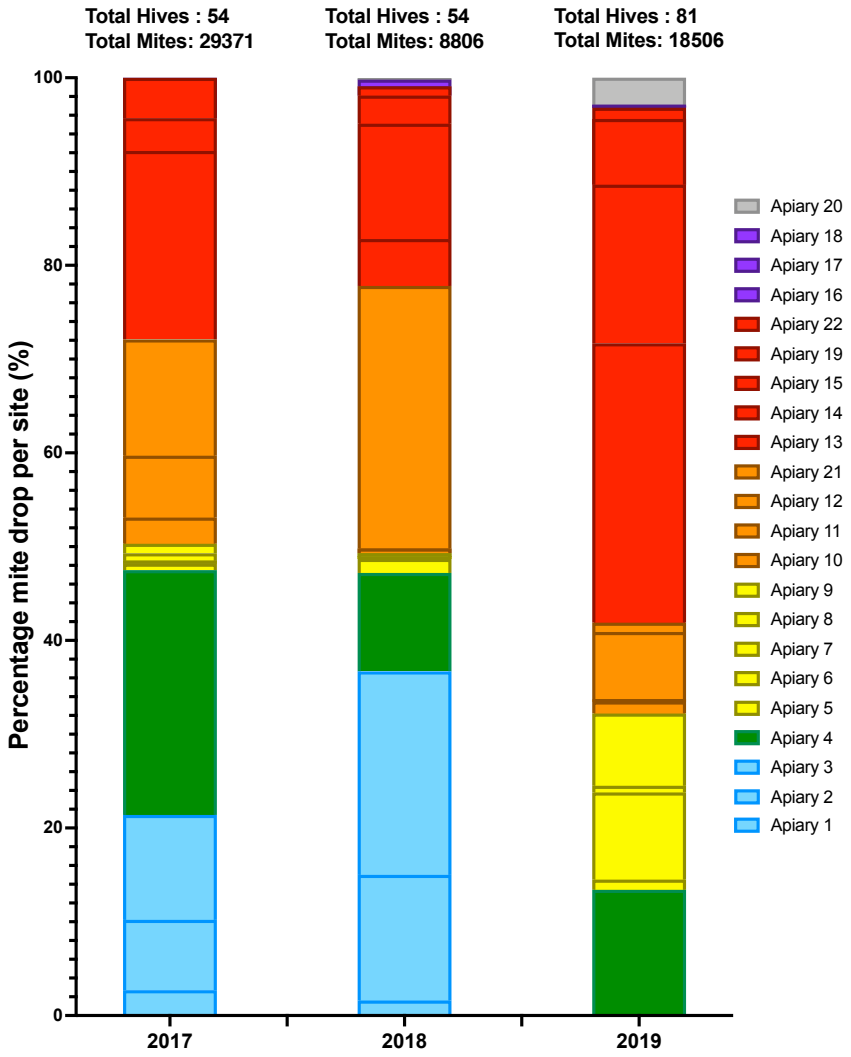


Figure 4-4 - Mite drop as a percentage of the total each year. *Each block represents one apiary. When blocks are not shown it indicates colonies did not record any mites post-treatment. Note – hive numbers differ from sample points in Figure 4-2, as colony numbers changed between these two points, for example 3 colonies were merged with others in 2017, meaning only 54/57 were treated.*

Table 4-2 - Changes in mite abundance over 3 years across all Sites.

Year	*Total colonies*	*Total mites*	*Average per colony*
2017	*54*	*29,371*	*544*
2018	*54*	*8,806*	*163*
2019	*81*	*18,506*	*228*

Figure 4-5 shows the average mite drop for each year with the individual colony values shown by the colour of the Site they are in. Using an unpaired t-test the statistical significance between each year was calculated. Changes from 2017 to 2018 (p-value <0.0001) and 2017 to 2019 (p-value <0.001) were statistically significant, but changes between 2018 and 2019 were not (p value - 0.26). In 2018 and 2019 the majority of colonies had fewer than 1000 mites, with only 1 colony in 2018 and 4 colonies in 2019 having over 1000 mites.

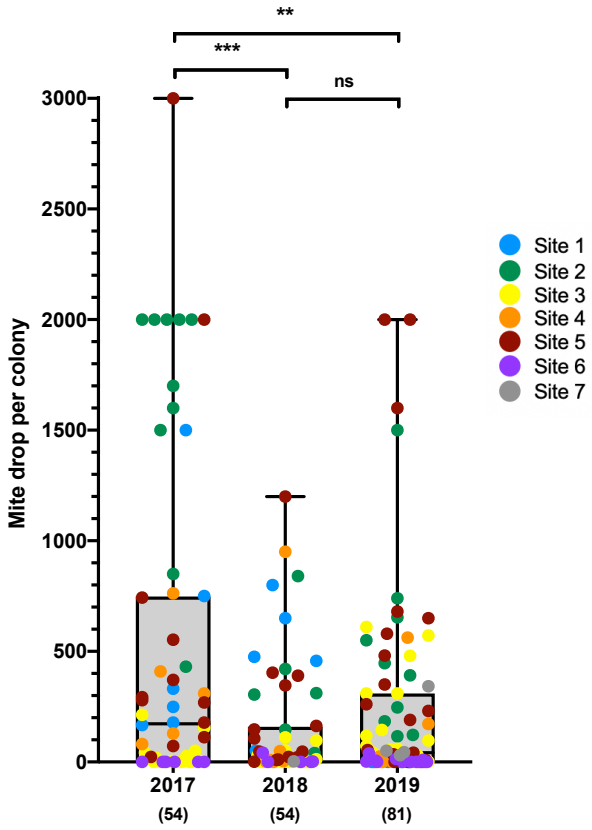


Figure 4-5 - Mite drop average for each year of treatment. *Colonies are coloured by site with all individual values shown. The grey boxplots represent the average mite drop with the quartile range, and the whiskers show maximum and minimum values. The total number of colonies is shown for each year under the x-axis. The bars at the top indicate the statistical significance between the years (p value <0.001 = **, <0.0001 = ***, ns = not significant).*

4.3 DWV population analysis in individual honey bee workers

4.3.1 Measuring DWV titre of individual workers

In 2017, honey bee samples were collected from every site in August, before application of the first round of miticide treatments, which was carried out in September. This was to determine the DWV titres in the population prior to any coordinated miticide treatment. Initially, analysis was carried out using individual worker bees from every colony, rather than pooled workers per colony or per apiary, so as to analyse changes in the viral population over time at an individual worker level in each apiary.

Adult workers were collected from every colony across the 6 Sites in 2017 (**Figure 4-6A**). Every colony was sampled individually and approximately ~30 adult workers were scooped off the comb into small plastic tubs supplemented with fondant (see section 2.1.2 for details). The bees were stored in these tubs until they were frozen at -80°C for further processing. The decision to sample adult workers, rather than emerging brood as in Chapter 3, was made because the volume of colonies sampled (57 during the first visit) would have made sampling brood comb and storing it appropriately until bees emerged impractical. Sampling adults also minimised the disruption to, and amount of time spent in, the colonies, which belonged to members of the public.

Five workers were selected from every colony and processed for DWV quantification using the methods detailed in sections 2.2.1, 2.2.2 and 2.2.6. The quantified viral load for each colony was calculated as genome equivalents (GE) /µg RNA from five sampled adult workers and an average is shown in **Figure 4-6B**. The Sites had different average viral titres, with Sites 2, 4 and 5 showing high average DWV titres typically associated with symptomatic infection (>10^6 GE/µg RNA) and Sites 1, 3 and 6 showing relatively 'healthy' DWV titres which rarely result in symptomatic infections (<10^6 GE/µg RNA).

The DWV titres for each colony were plotted against the mite drop (**Figure 4-7**) to determine whether there was a correlation between mite infestations and DWV titres. Typically, higher viral titres in honey bees were associated with higher mite abundance in the respective colonies (Site 4 for example), however some colonies with high mite counts had lower average DWV titres (>10^6 GE/µg RNA) than might be expected for infested colonies (such as Site 1). The Pearson's R analysis of correlation between DWV and *Varroa* indicated a weak positive correlation (R = 0.29). Following these observed differences between Sites, samples were selected from each Site for next-generation sequencing to determine whether these differences in viral titre were accompanied by differences in the virus population diversity.

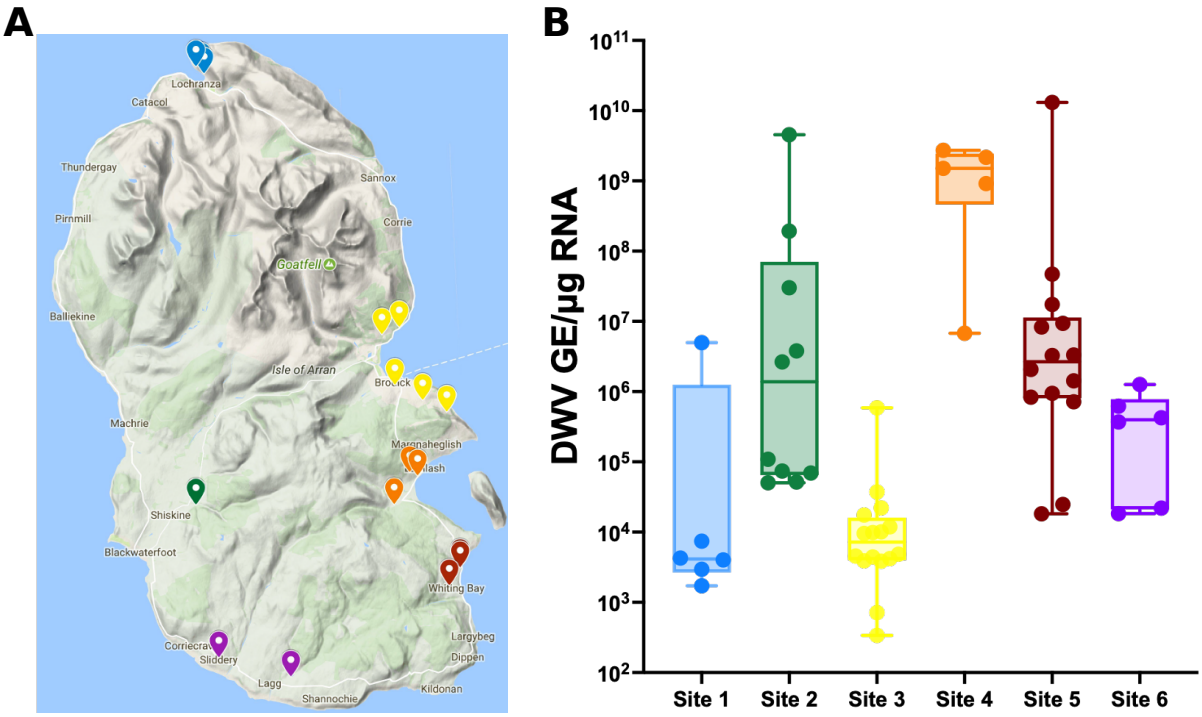


Figure 4-6 - Map of all apiaries on the Isle of Arran in 2017. *A - Each apiary is coloured according to which of the six site areas it falls within. B – the DWV titre for individual workers sampled from all six sites based on qPCR analysis. Each point in the plot is the average of five workers from a single apiary. The box plot shows the mean with the 25% and 75% interquartile distance with the whiskers spanning the minimum and maximum values, calculated using the Tukey method.*

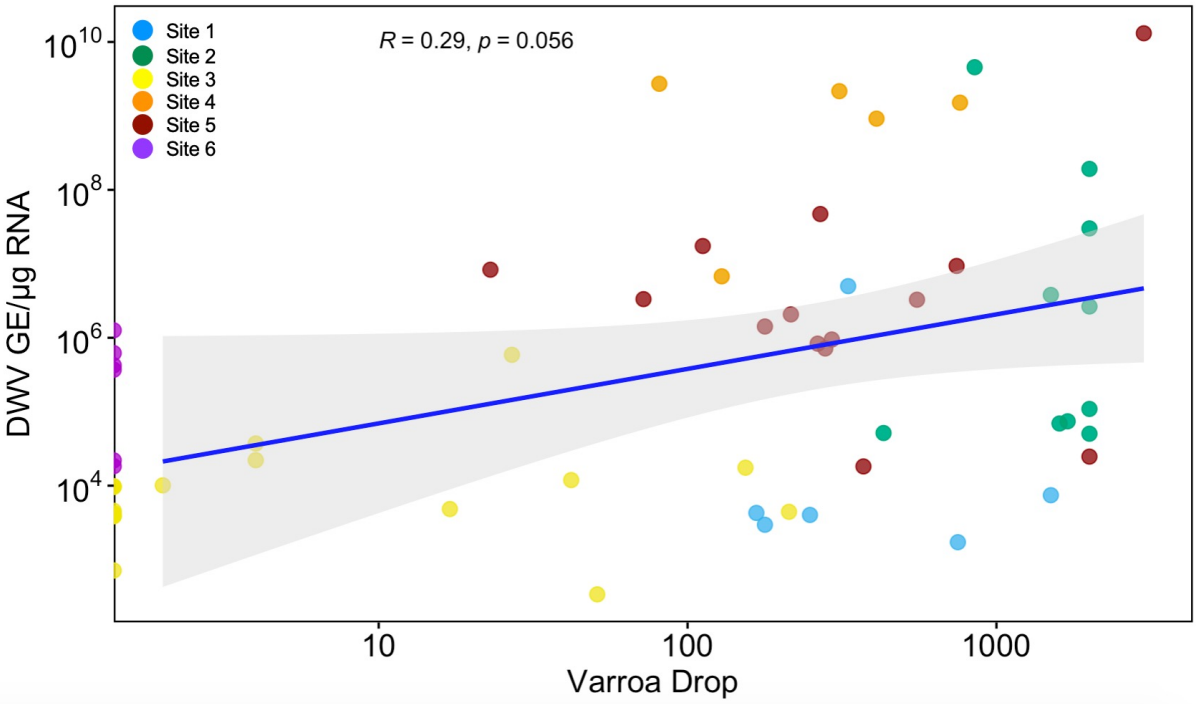


Figure 4-7 – Comparison of the Varroa mite drop and quantified DWV of each colony for all sites measured in August of 2017. *Each point represents the average quantified DWV GE/µg RNA of five individual adult workers per colony and the total Varroa drop from that colony after one week of Apivar application. The points are coloured based on geographic location of the colonies. Both axes are plotted on a log 10 scale. The blue line is the Pearson's correlation coefficient (Pearson's R) between DWV and Varroa, with the p-value and R score indicated at the top of the box.*

4.3.2 Measuring DWV population diversity using sequencing analysis

To determine the diversity of the DWV population, samples from each Site were selected for Sanger and next-generation sequencing (NGS). A known region of high nucleotide variability in the DWV genome, the leader protein (see **Table 1-1**), was Sanger sequenced and analysed using the SangerseqR analysis method in R (Hill, 2017) to look for peak diversity within samples (see section 2.3.2 and Appendix 2).

To prepare samples for Illumina sequencing, large amplicons (~10 kb) of the DWV genome were amplified using a specifically designed long-amp PCR (see section 2.3.4 for PCR conditions and primer information). Universal primers were designed using several published DWV full-genome sequences for reference, in order to detect and sequence as many of the 'common variants' as possible. The products were amplified, checked by gel electrophoresis (**Figure 4-8**) and purified as per section 2.3.4, prior to Illumina paired-end sequencing. A PCR amplification step was carried out prior to sequencing due to the very low virus titres recorded in some of the samples. Illumina sequencing was performed on a Mi-Seq and 300bp paired-end reads were generated using a multiplex library kit at the University of St Andrews. Sequenced samples were provided as fastq files for further processing.

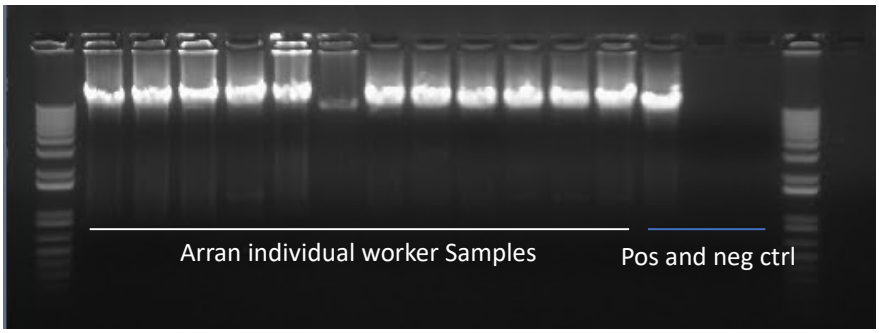


Figure 4-8 – Gel electrophoresis of PCR products spanning the full DWV genome. *A selection of DWV amplicons for Illumina sequencing were amplified using primers designed to amplify the majority of known major DWV variants. Positive control is the DWV clone pVVD9-4* (Gusachenko, *et al.*, 2020a).

To measure virus diversity, reads were extracted from fastq format and trimmed using Geneious (v2019.1.3) (Kearse *et al.*, 2012) and then aligned to a reference genome using Samtools (v1.10) (Li *et al.*, 2009). The aligned reads were then analysed using ShoRAH (see section 2.4.2 for full methods) and sorted haplotypes were analysed using the same method outlined in section 3.2.8. A cut-off of 3% was applied for haplotype matches (see section 2.4.3), so any reads below this were assigned to the 'Cluster-other' grouping when sequences were grouped.

To confirm that the sequences produced using this method were not a result of amplicon bias caused by the long-amp PCR method or the ShoRAH analysis, Sanger sequences of shorter fragments were also included in the phylogenetic analysis (**Figure 4-9**). The same honey bee samples were reverse transcribed from RNA using a random oligo cDNA synthesis kit and the leader protein region of the genome was amplified using universal primers (see Appendix 1). The Sanger sequences were aligned with the ShoRAH sequence output and the same reference genomes and all the Sanger and Illumina sequences clustered together on the same branch of the phylogenetic tree (**Figure 4-9**). These results indicated very limited virus diversity was present across the majority of Sites on the Isle of Arran during the summer of 2017, despite variation in the viral loads found in these workers and the variation in *Varroa* drop in each apiary. This perhaps reflects that all the colonies were obtained from a single source in Ayr in 2015 and any subsequent variation in mite levels could reflect colony management. When amplified by PCR, a near-clonal variant will also likely mask low-level variants (1–2% of the virus population) generated by polymerase errors typical of RNA viruses.

The sequences obtained from Sites 2, 4, 5 and 6 differed only by one or two single nucleotide polymorphisms (SNPs) across the whole region analysed. The Illumina data obtained from Site 1 revealed that one low-level DWV sample was composed of a mixed population of 8 unique DWV variants, four of which clustered closely with the samples from the other Sites, two clustered with Type A-like reference sequences and two others formed a distinct clade. Other samples with very low DWV titres from Site 1 and Site 3 did not have sufficient read-depth coverage of the leader protein region to apply ShoRAH analysis, so were omitted from this analysis. No sequences aligned with the DWV Type B-like reference genomes.

Figure 4-9 – Maximum-likelihood phylogenetic alignment of the top ShoRAH haplotypes and Sanger sequences for individual workers on the Isle of Arran. *Sanger sequences of individual workers were aligned against the top Illumina matches and the phylogenetic diversity of the samples was estimated. Several of the Illumina samples failed to produce clean results by Sanger sequence, so are not included here. Haplotypes are coloured by Site and the percentage of each sample they represent is indicated by the final number (e.g., ARR035-98% – sample 35, 98%). Sanger sequences are identified by sample number only (ARR17-061 – sample 61). The reference genomes are indicated by their NCBI accession number.*

Due to insufficient read depth at the 5' end of the genome for ShoRAH analysis, two other regions of the DWV genome where greater sequence depth was achieved were analysed. The RNA-dependent RNA polymerase (RdRp) found near the 3' end of the DWV genome, and the region of the genome including the helicase and spanning the junction between the structural

and non-structural proteins (~4500b–5500bp) were selected. The latter was selected because it is a known region of viral recombination between common variants (Ryabov *et al.*, 2014). Coverage of these regions was sufficient for analysis, therefore all the samples sequenced from Sites 1 and 3 were included, as well as some low-level samples from Site 2.

Haplotypes accounting for >3% of the viral population (see section 2.4.3 for rationale) in each sample were aligned with known reference sequences and coloured by site for the helicase and the RdRp region (**Figure 4-10A and B** respectively). Haplotypes from the Helicase (**Fig 4-10A**) show greater sequence conservation than the RdRp sequences, similar to the results of the leader-protein analysis (**Figure 4-9**). Sequence homology between variants is higher in the helicase than the RdRp at the nucleotide level (88.7% to 83.3% respectively), perhaps accounting for these differences (Woodford and Evans, 2020). The sequences all cluster on a unique clade from any of the reference genomes used in this analysis, but the closest branch of analysis suggests they are similar to published Type A-like sequences, similar to the leader-protein alignments, albeit further diverged. This is visualised in the cluster assignment bar plots (**Figure 4-10C**). Two samples from Site 1 contained mixed populations and two samples from Site 3 formed a unique cluster from the other sequences, perhaps reflecting the lower DWV titres observed at these sites. Most samples across Sites 2, 4, 5 and 6 all had >90% haplotype identity within the same cluster, with any remaining sequences classed as 'Cluster-other' due to a low percentage match (<3%).

The haplotypes from the region encoding the RdRp showed greater sequence diversity than the leader protein sequences (**Figure 4-10B**), both between sites and within samples. ShoRAH haplotypes were assigned to 20 clusters based on where they aligned during phylogenetic analysis. The majority of sequences were closely related, much like the leader protein analysis, however several of the haplotypes showed some sequence divergence, typically samples from Sites 1 and 3, which had low DWV titres (**Figure 4-6B**). Some haplotypes from Site 1 aligned on the same clade as known Type-B reference sequences. The higher number of unique haplotype clusters observed by phylogenetic analysis is reflected in the bar plots comparing percentage diversity at each Site (**Figure 4-10D**). Some haplotypes from Sites 5 and 6 formed a closely related clade, where a low number of highly similar variants are reported, all from Cluster 01, 02 and 03 and differing by a few SNPs.

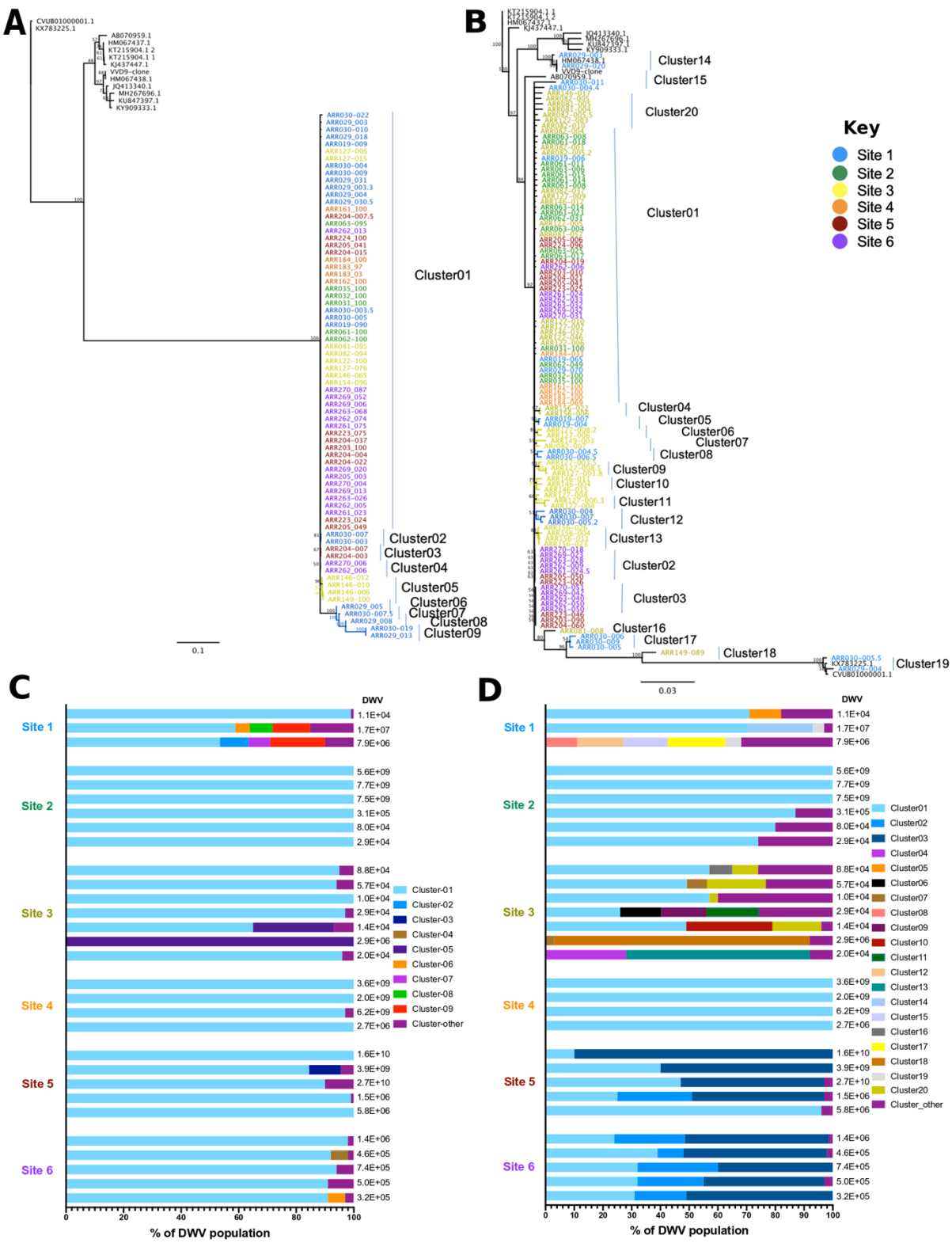

A
B
C
D
Cluster01
Cluster02
Cluster03
Cluster04
Cluster05
Cluster06
Cluster07
Cluster08
Cluster09
Cluster10
Cluster11
Cluster12
Cluster13
Cluster14
Cluster15
Cluster16
Cluster17
Cluster18
Cluster19
Cluster20
Key
Site 1
Site 2
Site 3
Site 4
Site 5
Site 6
DWV
% of DWV population
Cluster-other
Cluster_other

Figure 4-10 - ShoRAH phylogenetic analysis and assigned haplotype clusters for two regions of the DWV genome. *Panels A and B show the neighbour-joining phylogenetic analysis of all ShoRAH- sorted unique DWV haplotypes. The reads were assigned using analysis windows of the helicase region of the DWV genome (**A**) and the RdRp (**B**). The samples are coloured by Site location. Each sample is numbered per sampling ID and the last number in the name is the % of the sample that each haplotype represents in the population for, i.e. ARR030-005 is a sample haplotype comprising 5% of sample 30's virus population. Samples were assigned to clusters based on their genetic relatedness. In panels **C** and **D** the assigned haplotype clusters are displayed for each Site on Arran for the helicase and RdRp respectively. Each bar represents 100% of the virus population for a particular sample. The quantified viral load of each sample is shown at the end of each bar. Any sequences which did not produce a haplotype above the threshold of sensitivity (>3%) were assigned to 'Cluster_other'. Unlike in other plots, the colours assigned to these clusters do not correspond to virus identity and are randomly assigned.*

To determine the accuracy of the reported haplotypes from ShoRAH, a verification process was performed using probabilistic clustering of the sequences combined with statistical tests for strand bias for each Site on the island. This was carried out using the ShoRAH single nucleotide variant (SNV) caller package (McElroy et al., 2013). This method verified that the SNVs determining sequence diversity were being called with high accuracy and were not false positives due to low sequence coverage or systematic errors during variant calling. Low sequence diversity is indicated by a tight clustering of all the Fisher exact test p-value scores towards 1.0, indicating the majority of sequences in that sample have the same SNV change from the reference sequence present and therefore little within-sample diversity, whilst colonies which contained a large range of SNVs produced p-value scores from 0.05 – 1.0 indicating several distinct DWV variants within the population (**Figure 4-11**). This analysis was performed on the haplotypes identified from the helicase (**Figure 4-11A**) and the RdRp (**Figure 4-11B**), and both sample sets indicated that Site 1 had high sequence diversity, whilst the RdRp data indicated Site 3 also contained a mixed population, whilst all other sites contained little to no sequence diversity within samples. This reflects the results of the haplotype cluster alignments shown in **Figure 4-10C and D**.

Due to the large volume of samples gathered over the three years of field work and the prohibitive costs of processing and sequencing all samples individually at a high enough number to produce meaningful information, it was decided that all remaining samples would be analysed as pools of workers from each apiary. A sample batch from every apiary collected in 2017 was also processed in this analysis, this would act as an internal control for virus diversity in the pooled analysis.

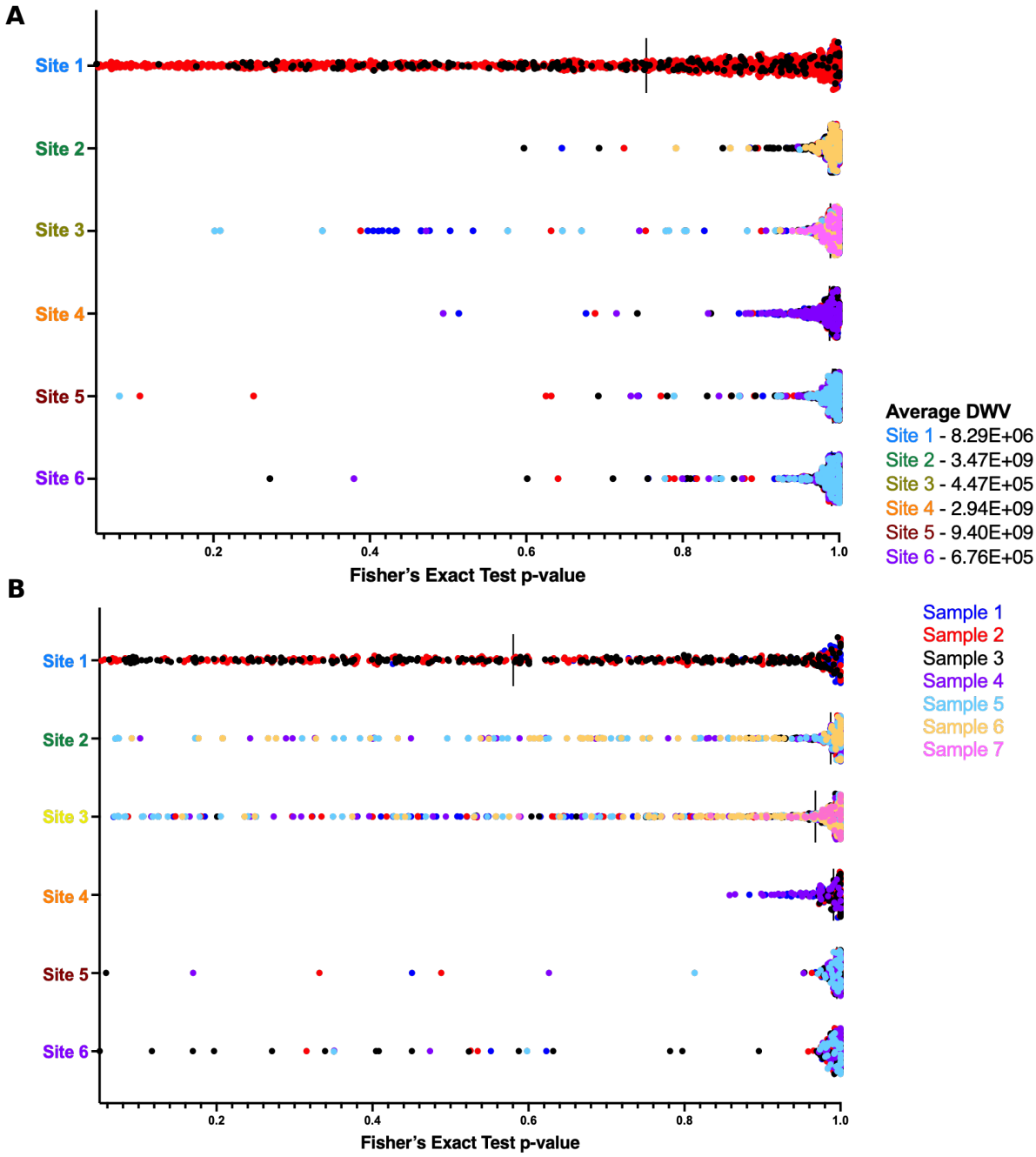


Figure 4-11 - Single Nucleotide Variant (SNV) Analysis of ShoRAH assorted Haplotypes. *Reads from the (**A**) helicase and (**B**) RdRp were analysed and all SNVs that conferred a greater than 0.05 p-value were assigned as a true sequence change and assigned Fisher's Exact Test p-value. SNV p-values close to 1.0 represent a variant in the majority of sequences in that dataset differing from the reference sequence they were aligned against. Samples with SNVs with lower p-value scores therefore have a great amount of sequence variation within the sample. Anything with a p-value <0.05 was excluded as the threshold for error based on McElroy et al (2013) modelling data. The different colours at each time point represent the SNVs for different honey bee samples sequenced and the average DWV titre for each Site is shown.*

4.4 DWV population analysis from pooled worker samples – years 1–3

4.4.1 Analysis of DWV titre changes over three years

Pooled samples of a minimum of 30 workers per apiary were crushed with a mortar and pestle, extracted and analysed for each collection time point. Where an apiary had multiple colonies, individuals were taken from all colonies and placed in a single pool. **Figure 4-12** shows the DWV titre for each apiary for each of the four sample time points, coloured by 'Site'. Site 1 has low DWV titres throughout the study, whilst Site 5 remains high. Site 2 decreases year on year (>10^8 GE/µg RNA in 2017 to ~10^3 GE/µg RNA by 2019), whilst Site 3 shows a moderate increase in early 2019, before decreasing to a low level typically associated with asymptomatic infections (<10^6 GE/µg RNA) by late 2019. **Figure 4-13** shows the reported *Varroa* drop per apiary plotted against the DWV titre for each year.

The pooled samples from 2017 produced similar viral titres to the individual workers (**Figure 4-7**), with Sites 1, 3 and 6 producing low DWV yields and Sites 2, 4 and 5 producing high titres. In 2018 the honey bee samples were gathered in May and, perhaps because of this, most apiaries had lower DWV titres than 2017, however apiaries from Site 5 still contained very high viral titres and mite numbers collected in September were high in several apiaries from Sites 1, 2 and 5. Many colonies perished over the winter between 2017 and 2018 and as a result there were fewer colonies to sample in the summer of 2018. By 2019 the beekeepers had substantially increased the numbers of colonies (**Figure 4-2**) from 28 in early 2018 to 84 by late 2019. In 2019 (**Figure 4-13C**), Site 5 still contained high DWV titres and *Varroa* counts in four of the five apiaries, and Site 2 had relatively high mite numbers, but much lower DWV titres. *Varroa* drop increased at Site 3 for the first time in the study and this was reflected in a higher DWV titre in one of the apiaries. No mites were recorded at Site 1 in 2019 and the lower DWV titres reflected this.

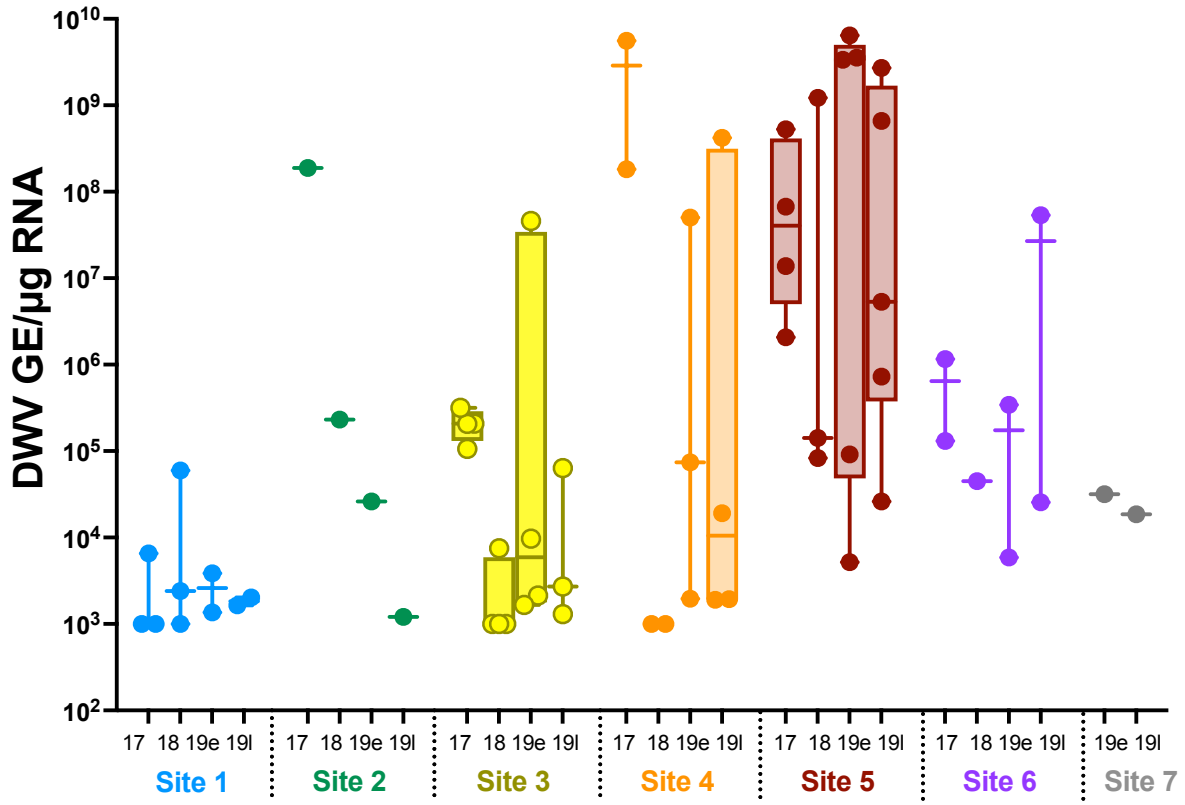


Figure 4-12 - The DWV titre for each apiary pool. *The DWV titre calculated using qPCR for a pool of 30 workers from each apiary over the four sample collection points. Each point represents a single apiary. The numbers along the x axis represent the sample collection times, 2017, 2018, May 2019 (19e) and August 2019 (19l). The plots show the mean with the 25% and 75% interquartile distance where applicable and the whiskers span the minimum and maximum values, calculated using the Tukey method; where there are insufficient samples to form a box the values are shown by a line spanning the distance.*

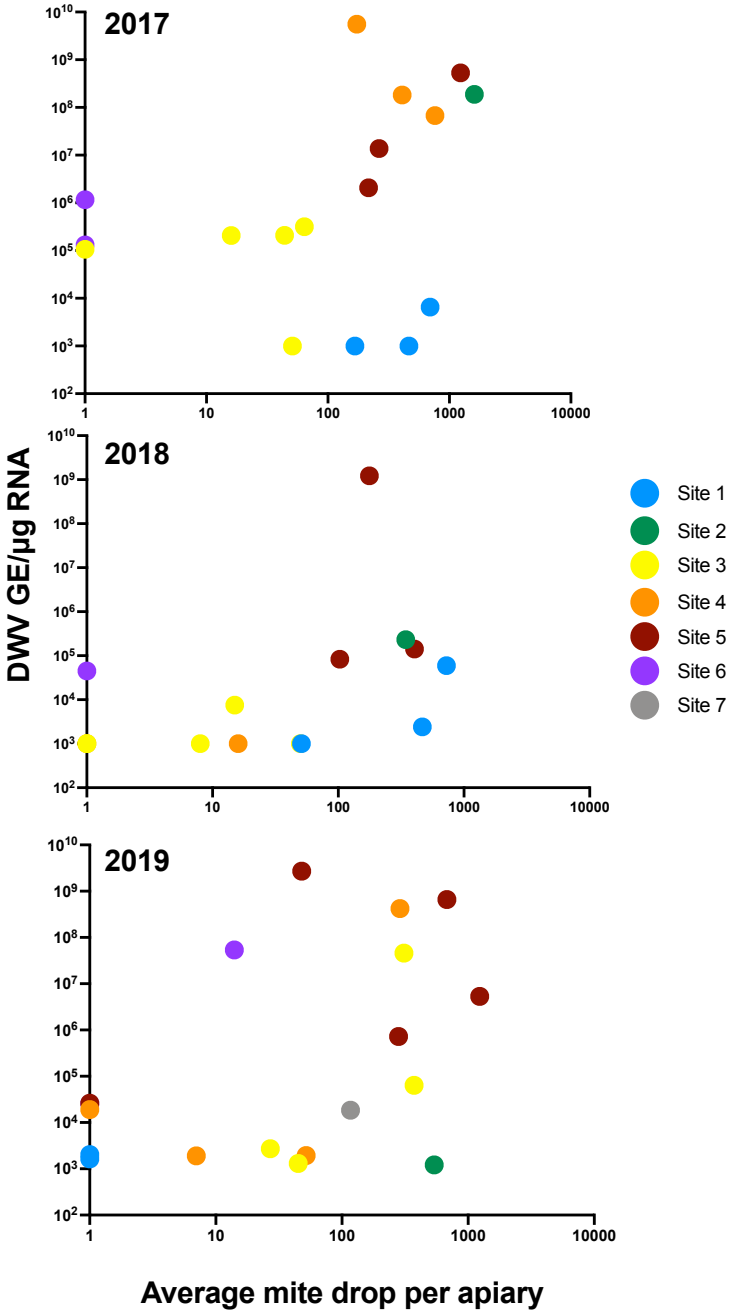


Figure 4-13 - The Average mite drop per apiary and the pooled DWV titre for each apiary in the three years of sampling. *Pools of ~30 adult worker bees were analysed by qPCR to determine the DWV titre for each apiary. The average Varroa drop from each apiary was calculated from the drop for each colony.*

Figure 4-14 shows the changes on a Site-by-Site basis over the course of the three years of treatments and sampling. Each black dot represents the average DWV titre from one apiary and the pink squares represent the average *Varroa* drop for the whole site. The plots show the improvements at Sites 1 and 2 over the three years, where *Varroa* numbers dropped significantly and DWV either remained low (Site 1) or decreased (Site 2). The figure also shows Sites where little has changed, such as Sites 5 and 6, and Sites which have become slightly more infested over time, such as Site 3. This figure highlights that the observed changes

are occurring on a site-by-site basis instead of at an island level, perhaps reflecting the fractured distribution of colonies on the island.

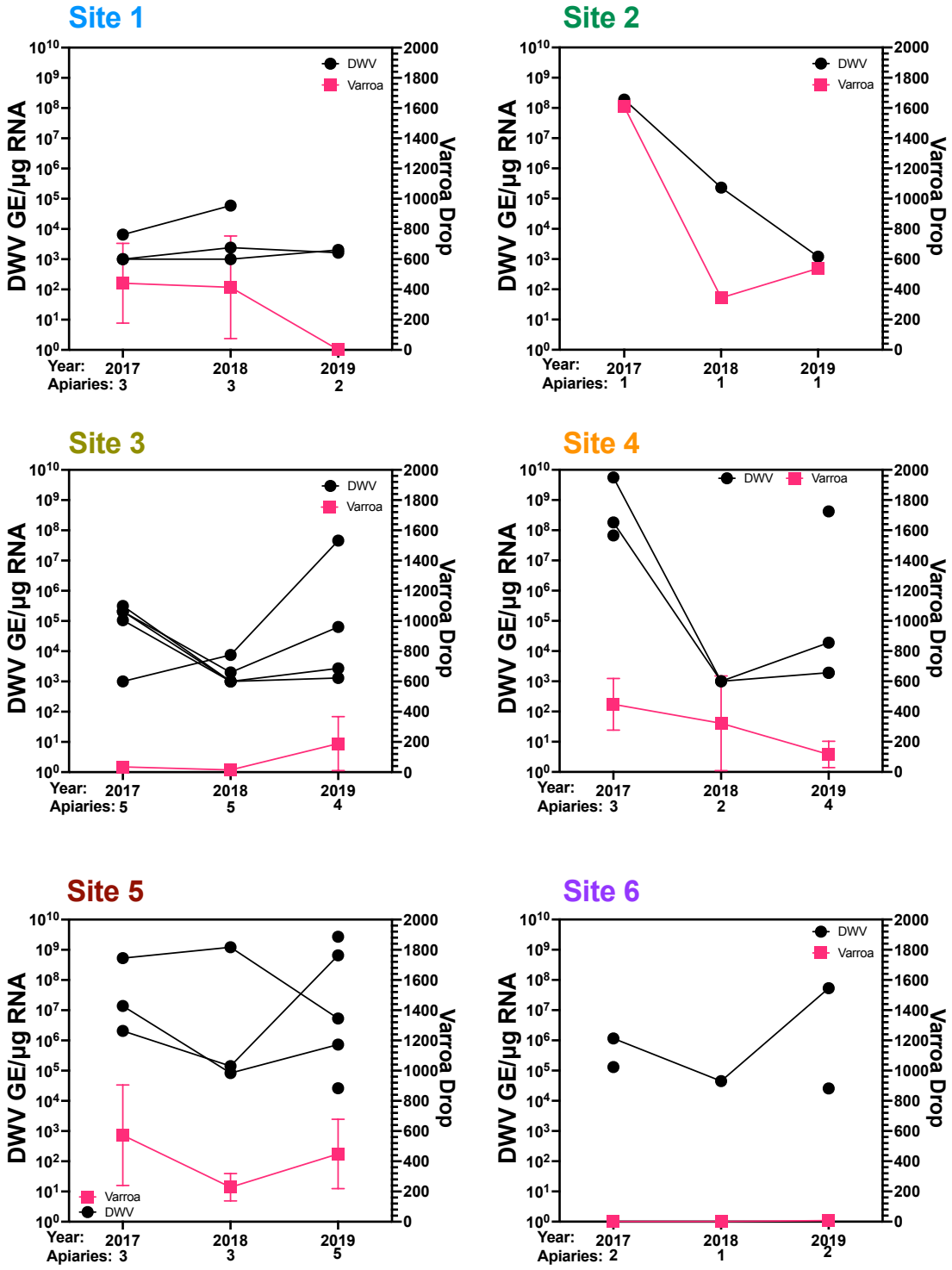


Figure 4-14 - Site by Site DWV and mite drop averages across 3 years. Each black circle represents the average DWV for each apiary located at the 6 sites. Varroa drop is shown as an average of all apiaries at each site by a pink square. Apiaries are linked year on year by lines to indicate changes over time, where apiaries are not linked it indicates all the colonies perished (Site 6 – 2017) or a new keeper joined (Site 5 – 2019). The number of apiaries at each Site is shown below the x-axis.

4.4.2 Sanger Sequence Analysis of Virus Pools

Following the observation of changes in viral titre on a site-by-site basis over the three years of sampling, further work was then carried out to determine changes in viral diversity. Initially, changes in the virus population were investigated by Sanger sequencing, using the pooled samples to determine which variants were dominant in each apiary. Sequencing of DWV in individual honey bees (section 4.3) revealed very low diversity in the 2017 samples; here, pools were analysed for all three years to determine if this changed over time. Standard end-point PCR for the presence of DWV was carried out, targeting the leader protein region of the DWV genome. The amplified samples for all three years were analysed by gel electrophoresis and the 2019 samples are shown below (**Figure 4-15**). Additional PCRs of the 2017 and 2018 pooled samples were amplified and the leader protein region of each was sent for Sanger sequencing (2017 and 2018 gels not shown).

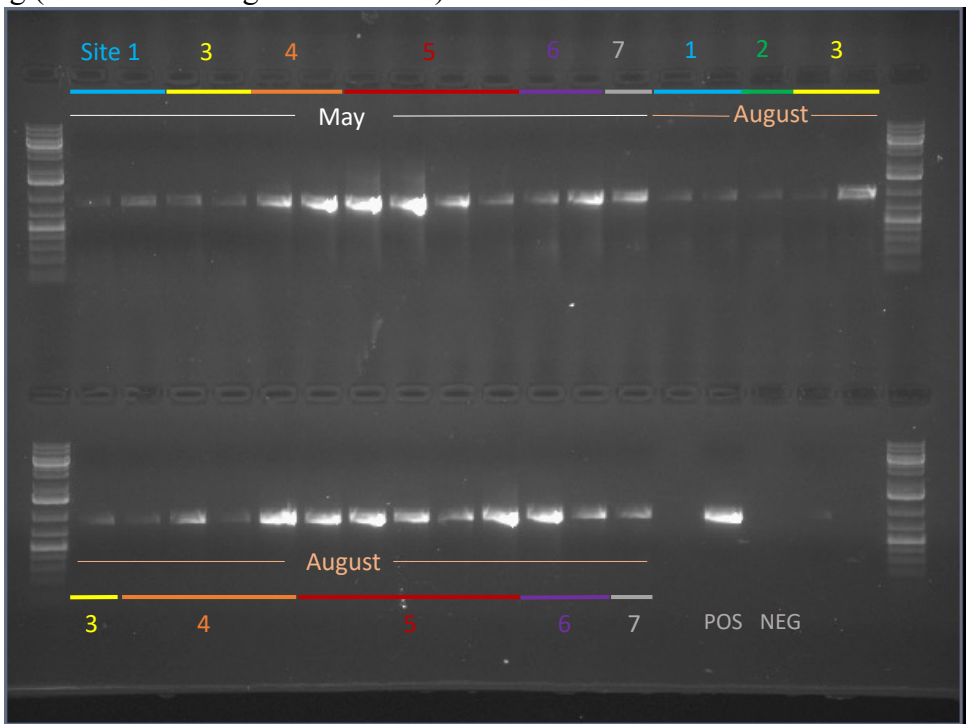


Figure 4-15 - PCR products from pooled samples amplifying the DWV leader protein region from May and August of 2019. *Each band represents a pool of bees from a single apiary, the colours and numbers indicate the 7 sites across the island. A pVVD plasmid DNA sample* (Gusachenko *et al.*, 2020a) *was used as a positive control.*

Sanger sequences from the 3 years of pooled samples were used to generate phylogenetic trees to determine the dominant variant in each pool. **Figure 4-16** shows the neighbour-joining phylogenetic analysis of all the samples that produced good quality Sanger sequences (several low-DWV-titre samples failed to amplify, particularly from 2018). The sequences were aligned with published sequences of known Type A and Type B variants. The phylogenetic tree shows the available sequences coloured by year of collection. From the sequences obtained, the 2017 samples, similar to the ShoRAH analysis of individual workers, are highly conserved and align with Type A-like reference sequences (red band on right). However, in 2018 and 2019 there are more sequences aligning with Type B-like reference sequences (blue band), particularly from Site 5. Several samples are missing from 2018 and 2019, due to low DWV titres failing to produce PCR products with adequate concentrations for sequencing.

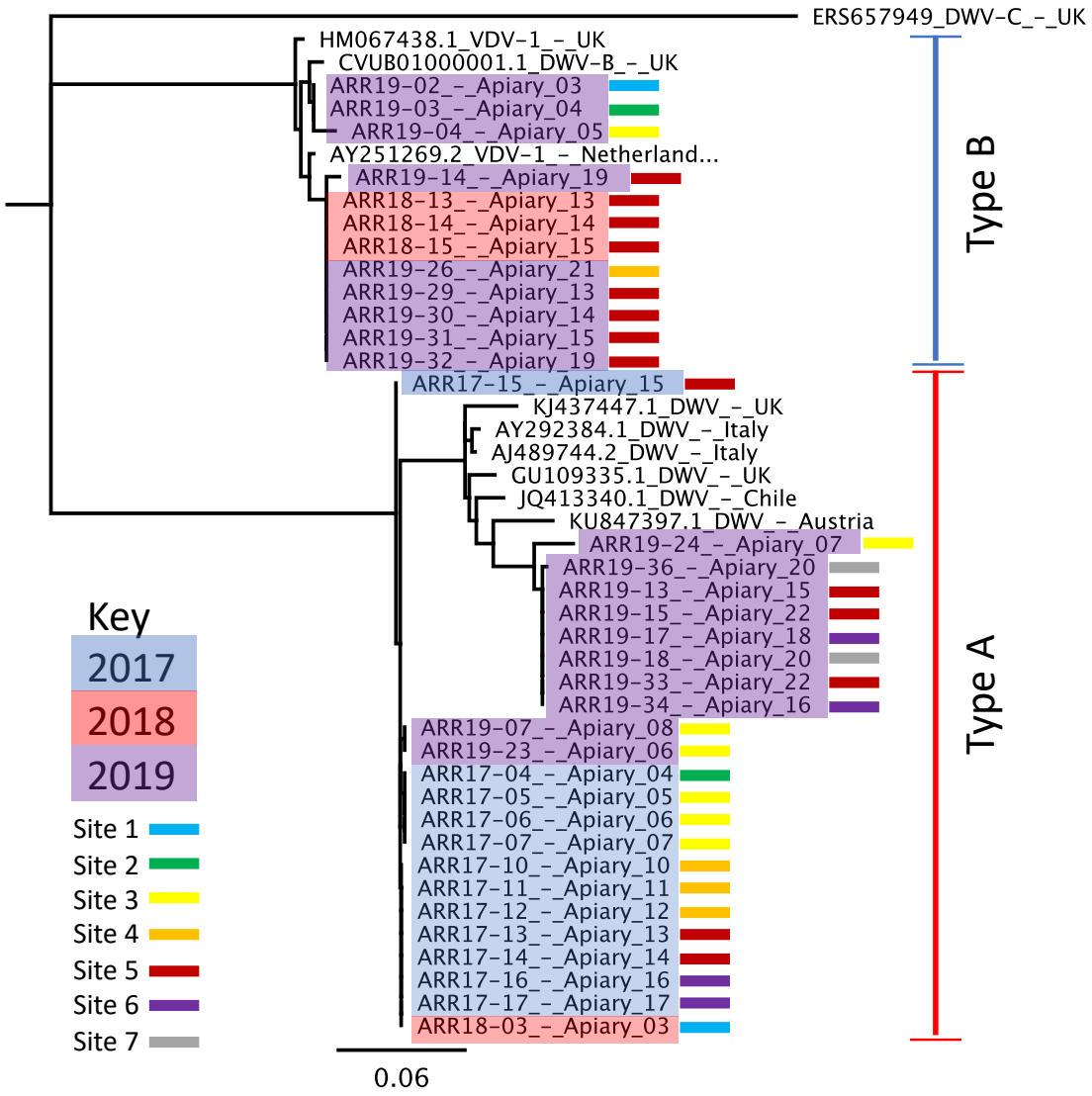


Figure 4-16 - Neighbour-joining phylogenetic tree of leader protein Sanger sequences from pooled samples. *The amplified pools from sites from 2017, 2018 and 2019 are aligned with published DWV reference genomes (shown by their NCBI number and location of origin). Each apiary is coloured by the year of sample collection, the bar at the end of each sequence name is the colour of the Site. The analysis uses a neighbour-joining tree with a Tamura-Nei model* (Tamura and Nei, 1993) *and 1000 bootstrap iterations to compile.*

4.4.3 Next-generation sequence analysis of worker bee pools

As with the DWV population analysis carried out in Chapter 3 and section 4.3, Illumina sequencing was used to determine sequence diversity obscured by standard Sanger methods or to examine samples where Sanger sequences were not obtained. In Chapter 3 and section 4.3 this was done using individual workers, however due to the large number of samples and prohibitive costs of processing, pooled samples were used for the remainder of this study.

Full-length genome DWV amplicons were generated by long-amp PCR and purified. The purified fragments were then processed using an Illumina Hi-Seq. 150bp paired-end reads were generated and barcoded libraries for every pool were generated. The samples were processed as per section 2.4.2 and analysed using ShoRAH (Short Read Assembly into Haplotypes)

(Zagordi *et al.,* 2011) to determine the virus population based on the unique haplotypes in each sample. A combination of low DWV titres, poor quality RNA (extractions were performed on 3-year-old samples in some instances) and a different sequencing platform from previous analysis meant full length genomic sequences were difficult to generate. Consequently, the analysis focused on 3'-end fragments including the RdRp and the helicase, but not the leader protein. As mentioned in Chapter 3, changes to ShoRAH meant smaller windows in the genome (150bp maximum) had to be analysed compared to previous ShoRAH analyses (section 4.3).

Figure 4-17A and **Figure 4-18B** show the neighbour-joining phylogenetic analysis for the RdRp and the helicase respectively, with samples coloured by year and clusters assigned to sequences based on their clade alignments. As with the Sanger sequence and NGS analysis of individual workers (**Figure 4-10** and **Figure 4-16**), sequences from 2017 show little diversity, with the dominant variant in each pool aligning with Type A at the RdRp (**Figure 4-17A**) and Type B (**Figure 4-18A**) at the helicase, indicating a recombinant form of DWV. The sequences obtained from 2018 show greater sequence diversity in the RdRp region, but only form two distinct clades in the helicase region. The bar plots in **Figure 4-17B** highlight the increased sequence diversity in 2018 compared to 2017 across some sites, perhaps reflecting the lower DWV titres observed. The haplotypes from 2019 produced greater sequence diversity, for both regions of the DWV genome, compared to the first two years. Analysis of RdRp samples (**Figure 4-17B**) from late 2019 at Sites 1, 2, 3 and 4 showed all contained a mix of Type A-like and Type B-like sequences and all six of the Sites were Type B-dominant by the end of sampling, although a low percentage of Type A-like sequences were still detectable. The 2019 sequences from the helicase region were more conserved than the RdRp, but a similar pattern of Type B dominance, with some low-level Type A variants at Site 1, 2 and 4, was observed (**Figure 4-18B**).

Analysis of the SNVs (single nucleotide variants) generated from the ShoRAH analysis was also performed to validate the haplotype diversity (**Figure 4-19**). The results confirm those found through phylogenetic analysis, with little diversity in 2017 except for Site 1, increasing diversity in 2018 across some sites, and much greater diversity by both 2019 sampling points.

Figure 4-20 shows a map of the island with the changes in DWV diversity in the RdRp genomic region alongside the changes in DWV titre and *Varroa* drop, as previously shown in **Figure 4-14**. Side-by-side, the changes emphasise that the dominant DWV variant changed regardless of geography or increases and decreases in DWV titre or *Varroa* drop. The arrows indicate colony movements and notably movement from Site 5 to Site 3 in 2019 appears to coincide with a subsequent increase in *Varroa* drop and DWV titre at Site 3.

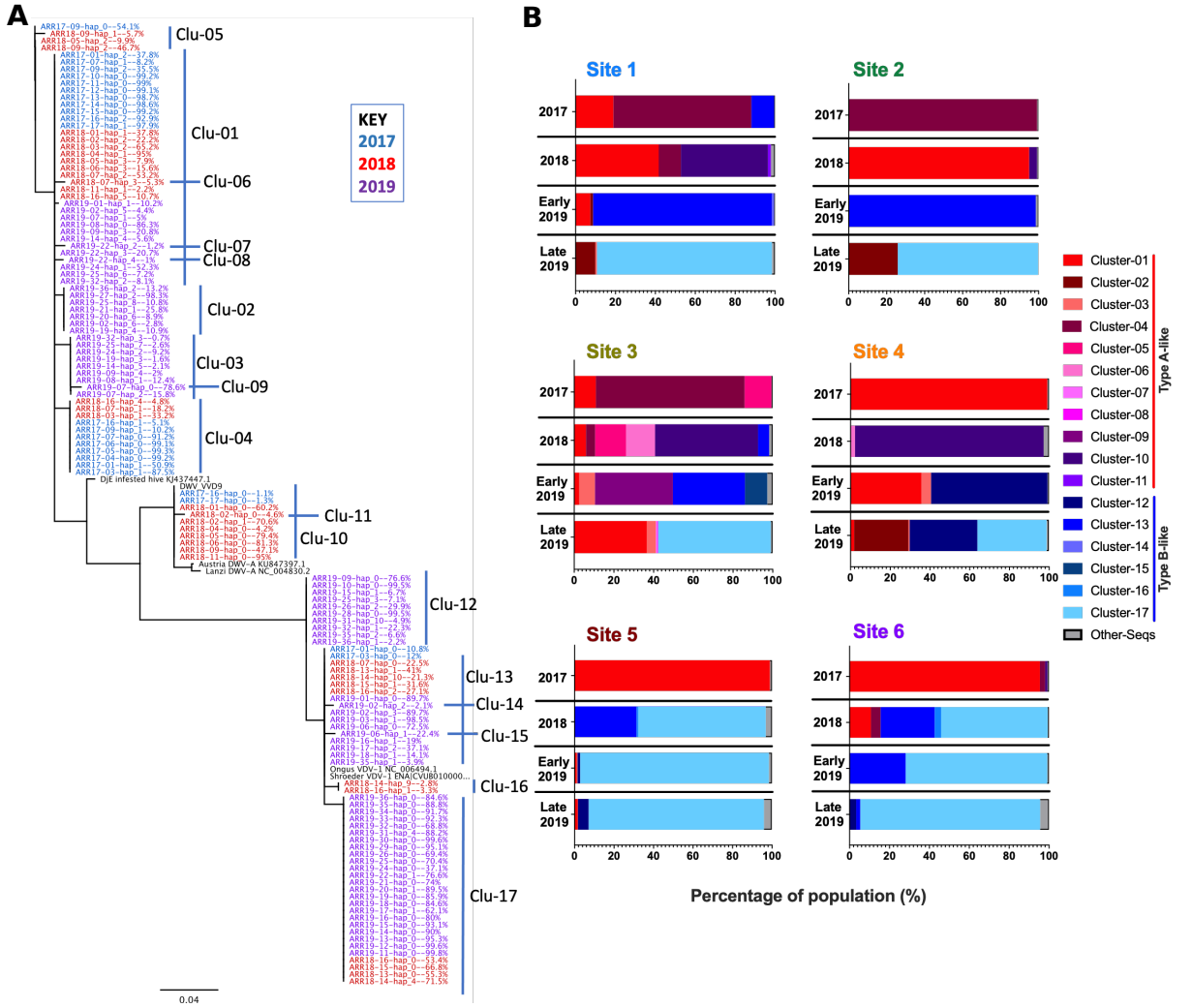


Figure 4-17 - ShoRAH analysis of RdRp sequences from pooled samples from all sites and years. ***A*** *– phylogenetic analysis of all haplotypes generated by ShoRAH in a window spanning 9000–9150 of the DWV genome (based on the pVVD – Accession no. MT415950). Samples are coloured by year, blue for 2017 pools, red for 2018 and purple for 2019. Each sample is labelled with the year, the sample number and the percentage of the sample that haplotype represents. Clusters (Clu-xx) are assigned based on sequence similarity and reference genomes are shown in black. The analysis uses a neighbour-joining tree with a Tamura-Nei model* (Tamura and Nei, 1993) *and 1000 bootstrap iterations to compile.* ***B*** *– The assigned clusters from panel A are used to generate bar plots for each site and year. The clusters are coloured shades of red for Type A-like sequences and shades of blue for Type B-like sequences. Sequences that make up a percentage below the QC limit are classed as 'Other-Seqs'.*

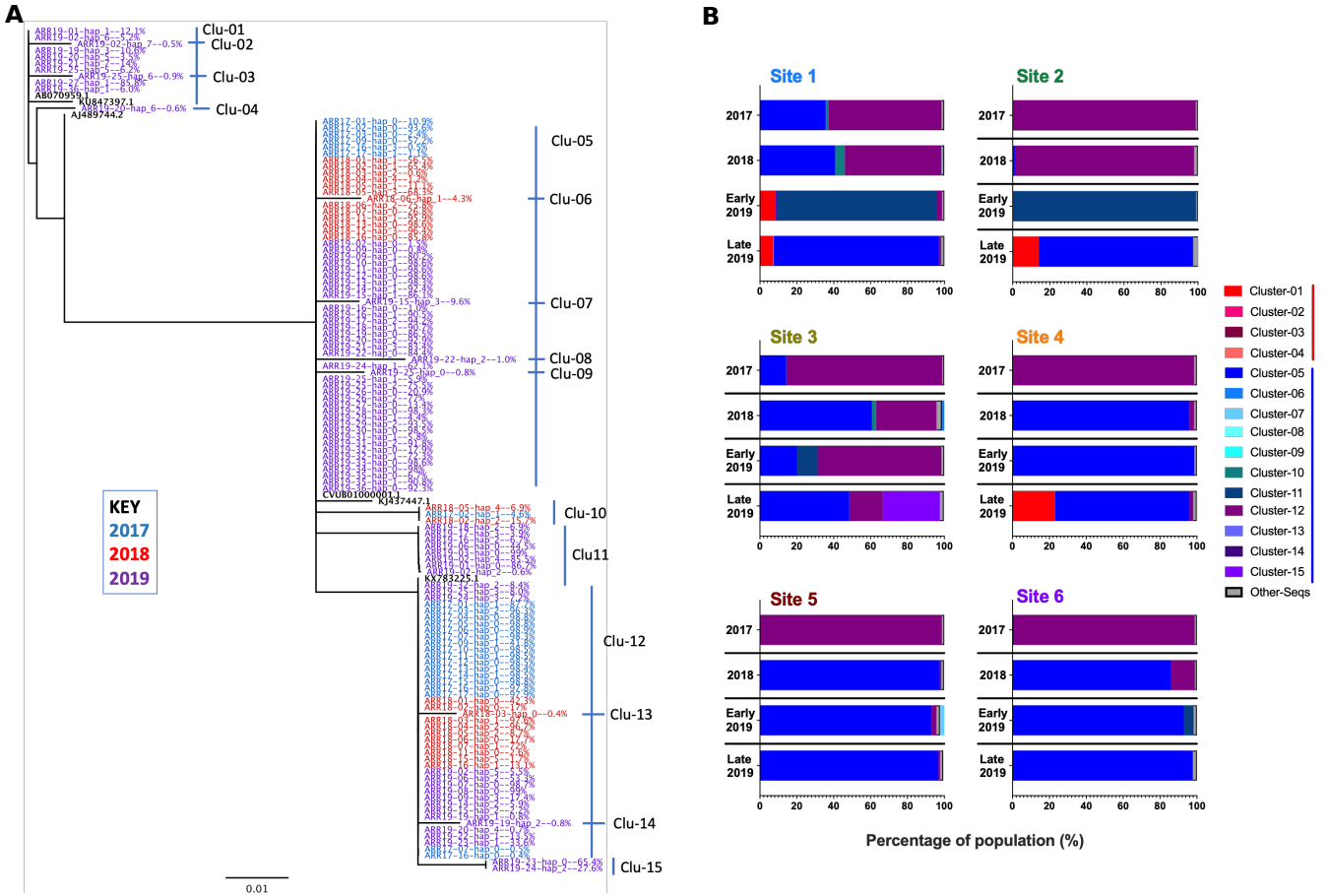


Figure 4-18 - ShoRAH analysis of helicase sequences from pooled samples from all sites and years. ***A*** *– phylogenetic analysis of all haplotypes generated by ShoRAH in a window spanning 5150–5300 of the DWV genome (based on pVVD - Accession no. MT415950). Sample format is the same as Figure X. Clusters (Clu-xx) are assigned based on sequence similarity and reference genomes are shown in black. The analysis uses a neighbour-joining tree with a Tamura-Nei model* (Tamura and Nei, 1993) *and 1000 bootstrap iterations to compile.* ***B*** *– The assigned clusters from panel A are used to generate bar plots for each site and year. The clusters are coloured shades of red for Type A-like sequences and shades of blue and purple for Type B-like sequences. Sequences that make up a percentage below the QC limit are classed as 'Other-Seqs'.*

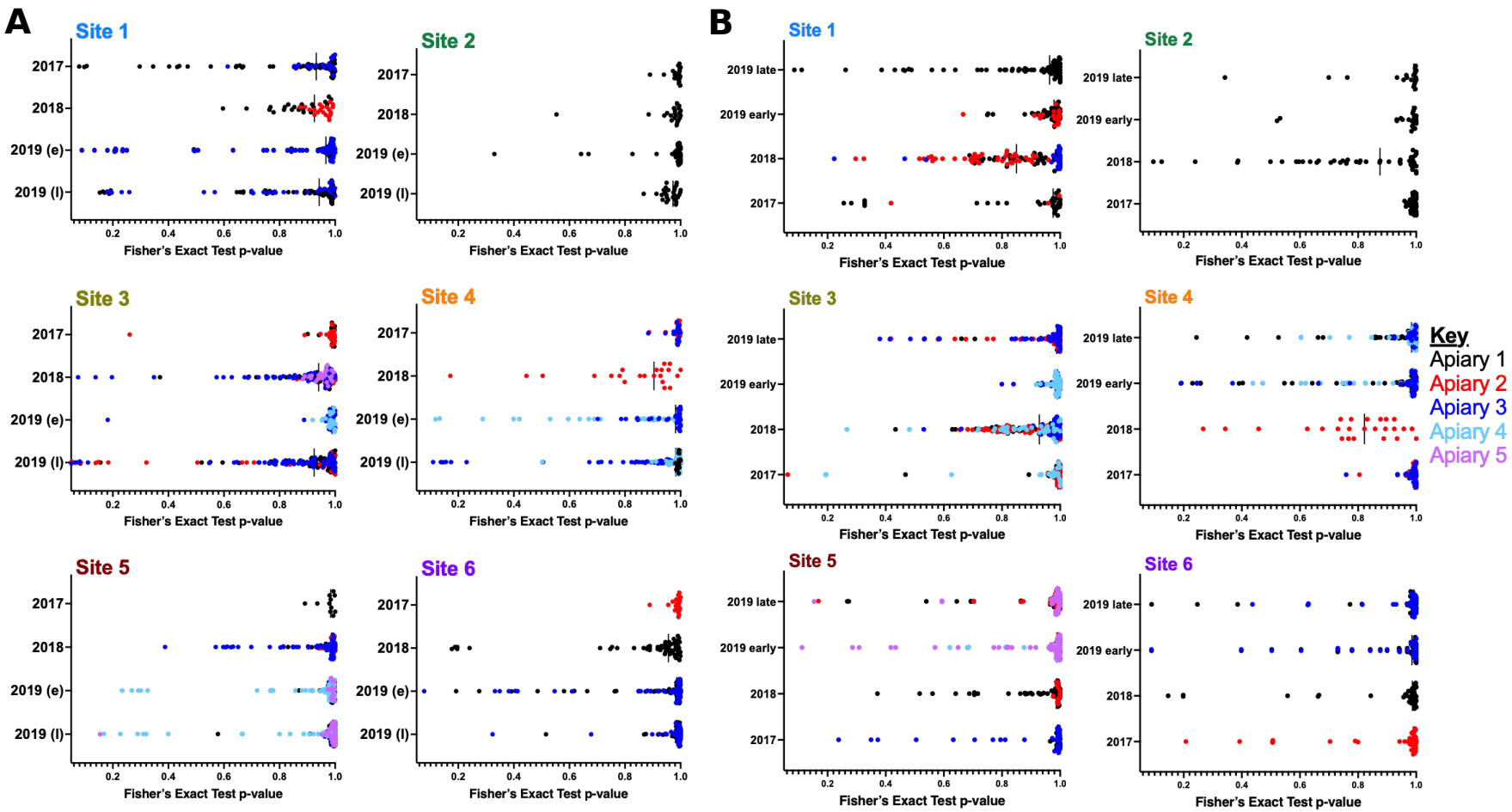


Figure 4-19 - Single nucleotide variants (SNVs) called for each sample from ShoRAH analysis. *SNVs from ShoRAH analysis for all samples from Arran Sites for the RdRp (**A**) and the helicase (**B**) were called if present in 3/3 iterations of the modelling. SNV p-values close to 1.0 represent a variant in the majority of sequences in that dataset differing from the reference sequence. Samples with SNVs with lower p-value scores therefore have a greater amount of sequence variation within the sample. Anything with a p-value <0.05 was excluded as the threshold for error based on McElroy et al (2013) modelling data. The different colours at each time point represent the SNVs for different apiaries within each 'Site' sequence.*

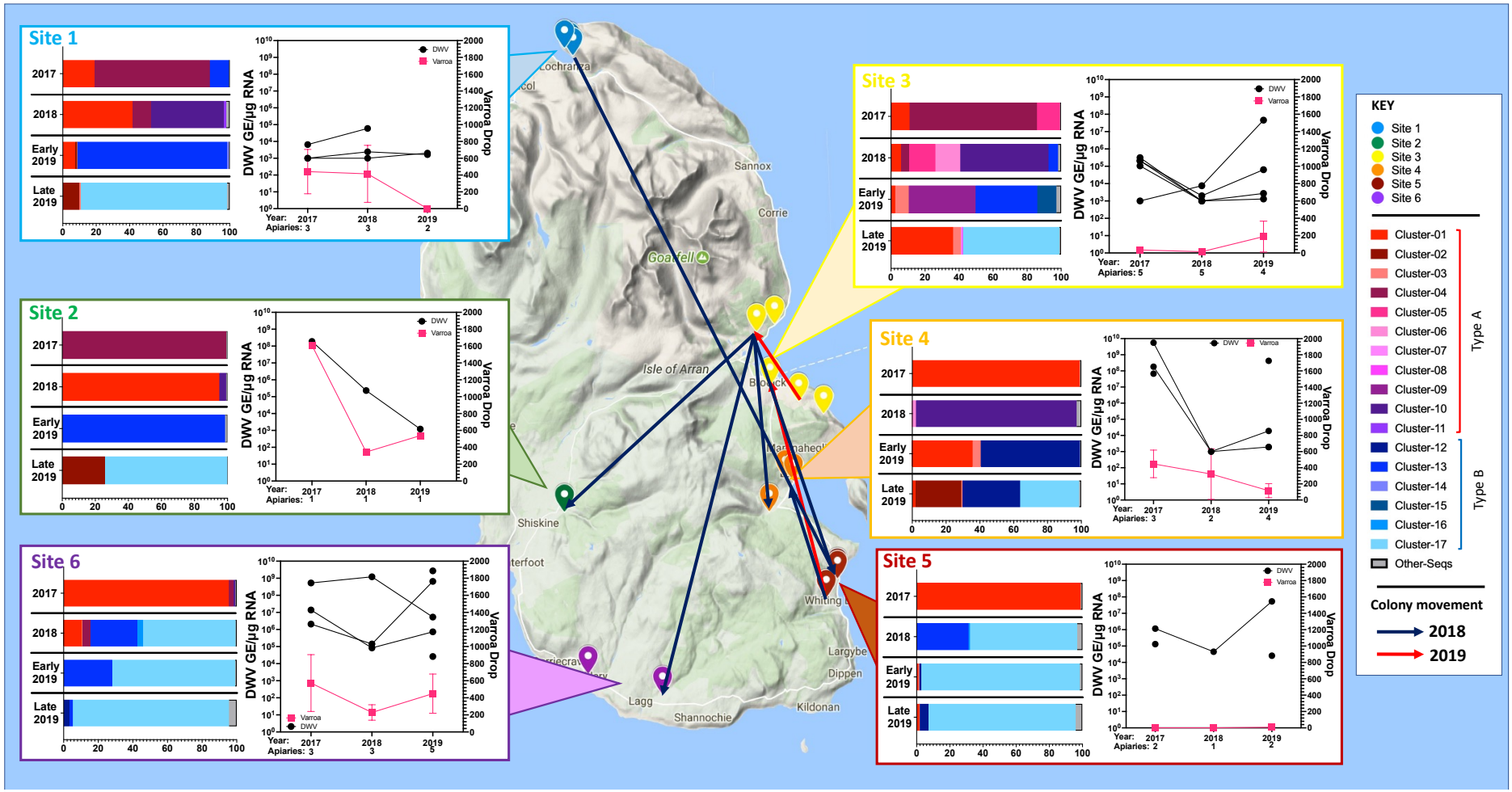


Figure 4-20 - Changes to DWV titre, mite drop and DWV diversity on a site-by-site basis over the duration of the study. *Each box represents a Site on the island, the bar charts show changes in DWV diversity over time for the RdRp region, the line graphs indicate changes in DWV titre and mite drop. The arrows on the map indicate colony movements, coloured by year.*

4.4.4 Correlation matrix analysis of virus population changes

A large number of variables influenced the data analysis throughout the study on Arran. A correlation matrix was generated to investigate the relatedness between the key variables observed during sample collection and analysis. This analysis was performed to summarise the large amount of information obtained in this study and to identify patterns between variables in the data. The matrix was generated using the data from the pooled honey bee samples from all apiaries and the NGS data generated from the RdRp region of the DWV genome. To run the correlation matrix several factorial variables in the model were converted to numeric values and some variables were assigned binary values, such as the classification of the dominant variant which was given a binary option of Type A or Type B based on phylogenetic alignment (**Figure 4-17**). Samples were classified as clonal if they only contained a single ShoRAH-assigned haplotype. Sacbrood virus (SBV) levels were determined based on the information in section 4.5.1. **Table 4-3** indicates which variables were converted and the values assigned to the variants within these variables for the model.

Table 4-3 - Correlation matrix converted variables and their assigned numeric values.

Variable	Variants	Assigned values
Clonality	mixed	0
	clonal	1
Colony_movement	No movement	0
	Colonies moved to site	1
Dominant_Variant	Type A	0
	Type B	1
Year	2017	0
	2018	1
	2019 early	2
	2019 late	3
Sacbrood virus (SBV)	Negative	0
	Low	1
	Medium	2
	High	3

The analysis was performed in R-Studio (v1.3.0) using the packages 'Hmisc' and 'Corrplot' (Wei *et al.*, 2017; Harrell, 2020). The Hmisc package computes the significance levels for a Spearman and Pearson correlation and returns both as correlation coefficients. It also generates a p-value for all possible pairs of correlations (full script in **Appendix 3**). The significance levels were then processed using Corrplot to generate a plot (**Figure 4-21A**) which shows the positive (blue) or negative (red) correlation between all sets of variables and scores them between -1 and 1. In **Figure 4-21B** the correlation coefficients are illustrated by the intensity of the colour based on the scale shown on the right and non-significant correlations with a p-value of >0.01 are shown with a cross through them. All p-values are shown in **Table 4-4**.

The analysis shows strong positive correlations between the clonality of the population and DWV titre (0.43), as well as with *Varroa* levels (0.38), but only a weak positive correlation between DWV titre and *Varroa* level (0.13). This weak correlation perhaps reflects the nature of sampling adults rather than brood in the colonies, where some Sites recorded low DWV despite reporting high mite drops, perhaps indicating that the bees carrying higher DWV titres perished or were not sampled. The 'dominant variant' is positively correlated with Site, apiary and year, with a particularly strong correlation with year (0.72), which fits with the sequence

data, which showed a shift from Type A to Type B over time on the island. Haplotypes, an indicator of diversity, are negatively correlated with DWV titre (-0.34), *Varroa* (-0.34) and SBV (-0.13), an expected result as DWV diversity is lost as mite levels increase in colonies (Martin *et al.*, 2012a). SBV was weakly positively correlated with *Varroa* (0.25), but weakly negatively correlated with DWV titre (-0.04) and significantly negatively correlated with 'dominant variant' (-0.41) and year, Site and apiary (-0.43, -0.35, -0.35 respectively). The strongest correlations were artificial, between Site and apiary, and between Haplotypes and Clonality, but acted as internal controls for the reliability of the analysis. The correlation analysis highlights the complexity of a data set with a large number of variables, but still produces the expected correlations between some variables, such as DWV clonality and high *Varroa* counts or year of sampling and dominant DWV variant. The model highlights that a lot of the factors involved in the analysis did not have a significant relatedness to the DWV titres or diversity changes throughout the study.

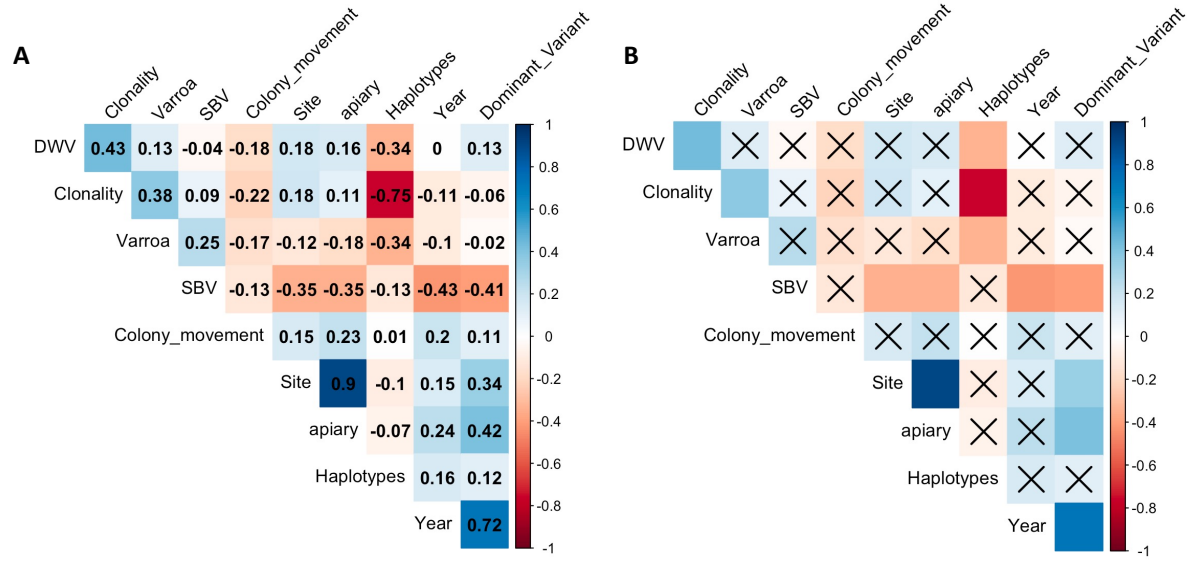


Figure 4-21 - Correlation matrix for Arran pooled samples using NGS data analysis of the DWV RdRp. *Correlation matrix for Arran data using RdRp NGS data. Positive correlations are displayed in blue and negative correlations in red. The colour intensity is proportional to the correlation coefficients. On the right side of the correlogram, the legend colour shows the correlation coefficients and the corresponding colours.* ***A*** *indicates the correlation coefficient value and* ***B*** *indicates which correlations are significant based on p-value scores (non-significant correlations indicated with a cross).*

Table 4-4 - P-value scores for Pearson correlation coefficient matrix. *The table shows the p-value scores for each correlation shown in* ***Figure 4-20B****. Cells shaded in grey had non-significant p-value scores for the Pearson correlation coefficient, cells in white had significant p-values. Values below 0.0001 are show as <0.0001 for simplicity.*

	Year	Site	apiary	DWV	Varroa	Haplotypes	Dominant Variant	Clonality	Colony movement	SBV
Year	0.0000	0.2489	0.0618	0.9854	0.4281	0.2270	<0.0001	0.4131	0.1139	0.0006
Site	0.2489	0.0000	<0.0001	0.1693	0.3408	0.4467	0.0083	0.1638	0.2397	0.0059
Apiary	0.0618	<0.0001	0.0000	0.2320	0.1750	0.5956	0.0008	0.4124	0.0751	0.0063
DWV	0.9854	0.1693	0.2320	0.0000	0.3342	0.0078	0.3226	0.0005	0.1718	0.7701
Varroa	0.4281	0.3408	0.1750	0.3342	0.0000	0.0076	0.8497	0.0025	0.1971	0.0512
Haplotypes	0.2270	0.4467	0.5956	0.0078	0.0076	0.0000	0.3762	<0.0001	0.9628	0.3224
Dominant_V	<0.0001	0.0083	0.0008	0.3226	0.8497	0.3762	0.0000	0.6601	0.3920	0.0012
Clonality	0.4131	0.1638	0.4124	0.0005	0.0025	<0.0001	0.6601	0.0000	0.0960	0.5087
Colony_movement	0.1139	0.2397	0.0751	0.1718	0.1971	0.9628	0.3920	0.0960	0.0000	0.3241
SBV	0.0006	0.0059	0.0063	0.7701	0.0512	0.3224	0.0012	0.5087	0.3241	0.0000

4.5 Other viruses and honey bee sub-species screening

4.5.1 PCR screen for Sacbrood virus

Although this study focuses on DWV and *Varroa destructor*, a range of other viruses infect honey bee colonies, including Sacbrood virus (SBV), a virus which infects the developing larvae in a colony (Bailey, 1969). There is a possible relationship between mite activity and the presence of SBV in honey bee colonies (Giuffre, Lubkin and Tarpy, 2019), so a screen was performed on all pooled samples to determine the presence of the virus in the colonies on the island. PCRs were carried out on all pooled Arran samples using primers from Grabensteiner et al., (2001) and samples were visualised by standard gel electrophoresis. **Figure 4-22A and B** shows the results of the SBV PCR amplification for all pooled samples. Of the 2017 pools, 11/16 samples were positive, 13/15 were positive in 2018, 4/18 in early 2019 and 3/19 in late 2019. The positive bands on the gel were scored as 'high', 'medium' or 'low' based on band strength and plotted as a percentage of all the samples for that year (**Figure 4-22C**). The number of SBV-positive samples and the 'intensity' of the bands decreases from 2017 to 2019.

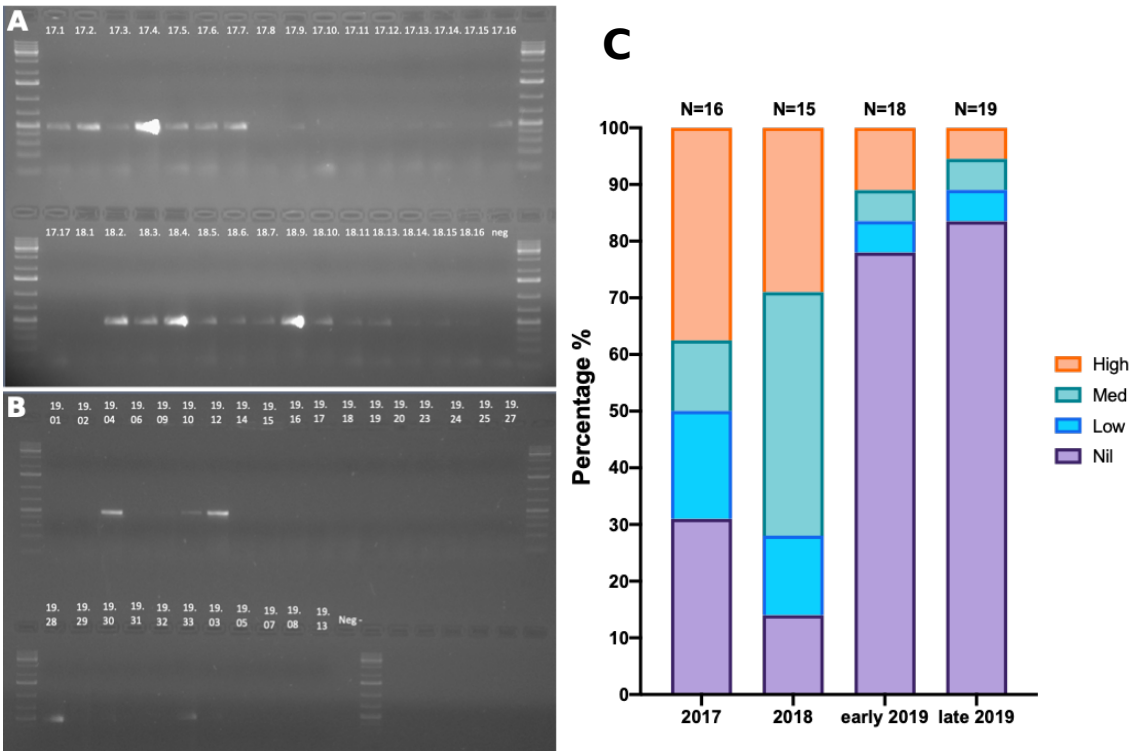


*Figure 4-22 - PCR analysis of Sacbrood virus (SBV) in honey bee pools on Arran. Panel **A** and **B** indicate PCR positive samples by gel electrophoresis visualisation. Each sample is labelled above the well. Panel **C** shows the percentage of samples which contain high, medium, low and no SBV at each of the four sample time points based on the intensity of the bands from panels A and B. The N= above each bar indicates the total number of samples from each time point.*

The relative levels of SBV were then compared to the DWV titres calculated for each apiary and are shown in **Figure 4-23**. In 2017 the apiaries with the highest average DWV titres typically had lower SBV presence or no detectable SBV by PCR. In 2018, DWV titres were low across most of the samples and the SBV presence and absence is evenly distributed amongst the samples. In 2019 most samples (n=13/18) are free from SBV, but the three samples that contained the highest DWV titres were all SBV-positive too.

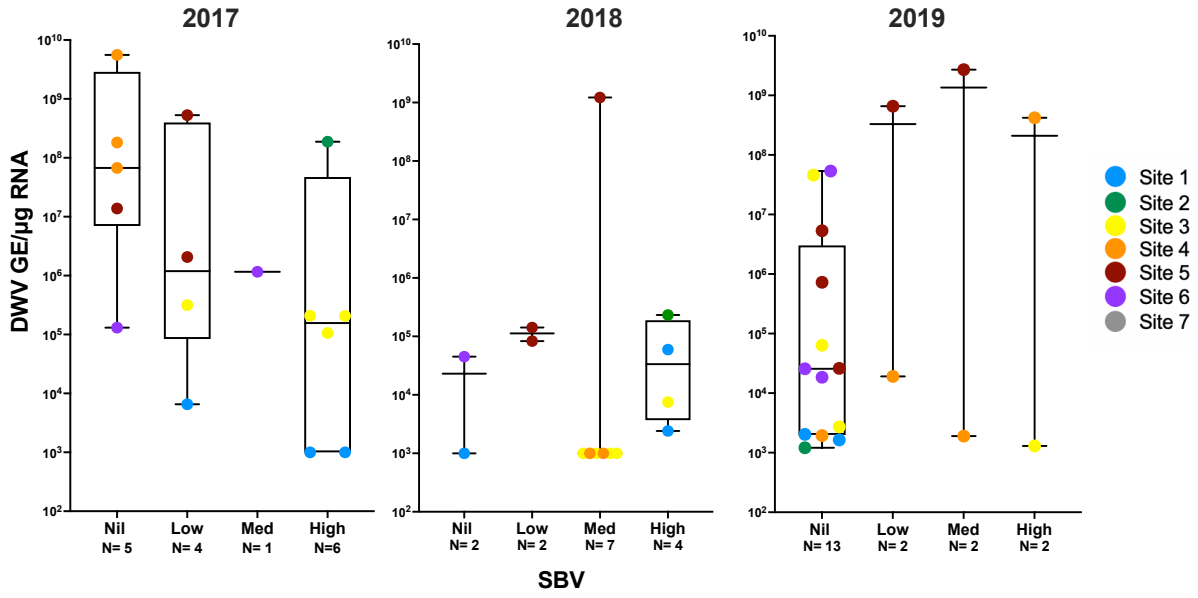


Figure 4-23 - Sacbrood virus presence in pooled apiary samples plotted against the DWV titres for each apiary. *Each apiary average is coloured by Site and the total number of apiaries in each SBV category are shown along the bottom of the x-axis. The samples from 2019 compare SBV presence with the late DWV averages only.*

4.5.2 Screening honey bee sub-species by PCR

Despite all honey bees being purchased from the same supplier in 2015, following the loss of all colonies managed on the Isle of Arran, there is a possibility that some bees remained as feral colonies/swarms on the island and could potentially rob or drift into the newly purchased colonies. Therefore, a screen of all the managed bees on the island was performed to determine their relatedness and the degree of sub-species variation, if any, that existed on the island. This could also determine whether there was any pattern between sub-species and *Varroa*/DWV susceptibility. Using a published method for honey bee sub-species identification (Syromyatnikov, Borodachev and Kokina, 2018), the Cytochrome oxidase I (COI) was targeted for each honey bee pool.

The COI of each pooled sample from 2017, 2018 and 2019 was amplified and sequenced. A neighbour-joining phylogenetic tree was generated from the Sanger sequences using Geneious (**Figure 4-24**). The sequence analysis indicated that most sites on the island have very closely related sub-species of honey bee, as expected based on the purchasing history and colony splitting efforts. This is also reflective of the limited amount of information obtained from sequencing only one honey bee gene for these samples, where the number of SNPs observed between samples was low. Typically, the colonies contained sequences which aligned with the sub-species *Apis mellifera caucasia*, with a small set showing a small divergence and aligning more closely with *A. m carnica*. However, the Site 2 colonies (and colonies split from Site 2 for Site 7) clustered on a distinct branch from the majority of the other Arran samples, aligning closely with reference sequences for *A. m. capensis*, *A. m. scutellata, A. m. lamarckii and A.m. intermissa*. Site 1 apiary 1 also aligned with these sub-species in 2017 and 2018, however the colonies died in late 2018 and the replacement stock, sourced from apiary 3 at Site 1 is a different sub-species (19.01 and 19.19). There was no observed correlation between the honey bee sub-species sequenced and the number of mites observed in each apiary, as most apiaries contained the same sub-species. It should be noted that, given that only one gene was sequenced

and the number of observed SNPs between samples was low, the differences between the samples only indicate a small difference between the bees at Site 2 and associated Sites and the bees elsewhere on the island, and do not suggest that species such as *A. m. capensis* are actually present on Arran.

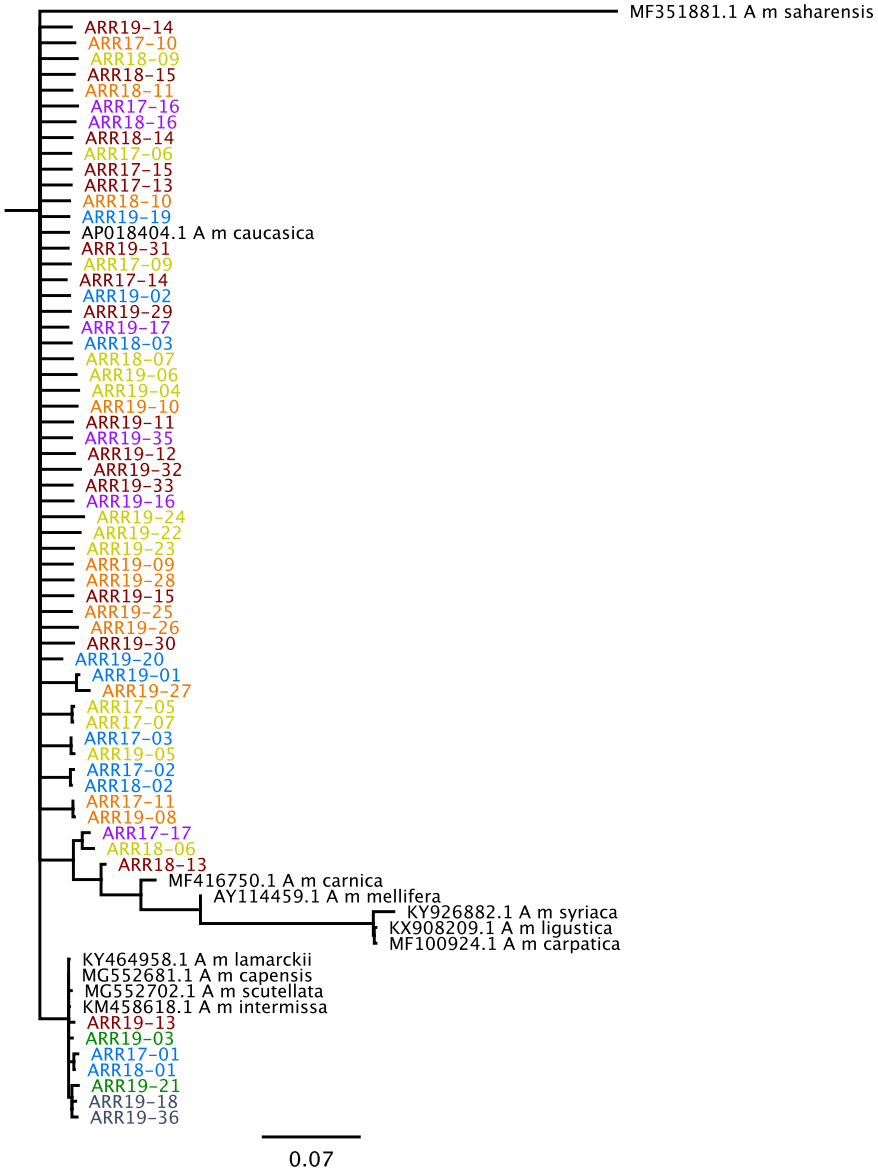


Figure 4-24 - Maximum likelihood phylogenetic analysis of honey bee species ID sequences from all pooled samples from 2017, 2018 and 2019. *Honeybee COI sequences were aligned to the most common Apis mellifera sub-species matches by BLAST. Apis mellifera saharensis was used as the outgroup. A 1000 bootstrap iteration was run using the Tamura-Nei model* (Tamura and Nei, 1993). *Sequences are coloured by Site on the island and are labelled by year and apiary (e.g. ARR17-01, is a pool from 2017 and apiary 1).*

4.6 Discussion

The aim of the research was to determine whether coordinated treatments in an isolated environment reduced *Varroa* infestations and improved honey bee health across all the apiaries. Using an island for the study meant that the colonies under the treatment regime should not be exposed to drifting or robbing bees from neighbouring untreated colonies. The Isle of Arran is 3 miles at any point from the mainland of Scotland and as such bees are unlikely to fly from the mainland to the island. The engagement of the beekeepers on the island was key to this project and, as they had recent experience of losing all their colonies to *Varroa* infestation and subsequent DWV infections, they were suitably motivated to take part. The island provided a good environment for a large-scale coordinated treatment programme, with beekeepers with a range of different experience levels and a large number of managed colonies. Over the three years of the study the *Varroa* drop post-treatment reduced on average across all sites (58% from 2017 to 2019, see **Figure 4-5**) and the number of managed colonies increased from 54 to 81. Together, these data provide clear evidence that coordinated treatments are improving colony health on the island, however high mite numbers persisted at some Sites.

At the start of the study there were 17 beekeepers in the Arran Bee Group, split across 6 'Sites' and managing 54 hives between them. Over time the number of beekeepers, the total number of hives and even the sites where bees were kept changed (**Figure 4-1/4-2**). An essential requirement of the study was that no new colonies were introduced to the island during the study, so as not to potentially bring new infestations or viruses to the island, and as such the beekeepers agreed to not bring any new colonies of bees on to the island and instead replaced lost colonies by splitting existing strong hives. This method has both positive and negative effects. Splitting from a known colony means the strength, disease history and mite load should all be better understood. The drawback of splitting from one of the island's hives is that one with high mites, DWV or both, could be selected and the split takes a proportion of these mites and viruses with it to a new site, disseminating mites around the island and facilitating further disease transmission. This is evident with some of the colony movements from Site 5, a site with high mite infestations throughout the study, which contained one of the 'association apiaries' responsible for providing new beekeepers with bees. Colonies were moved from Site 5 to Site 2, Site 4 and Site 7 in 2018 to compensate for high overwinter losses, and contributed to increased mite loads and DWV titres at these locations (**Figure 4-4 /4-13**). Additionally, splitting from a small original batch of colonies in an isolated environment can form a breeding bottleneck, in both the bees, and in the virus populations they carry.

From the mite drop and DWV titre data after three years of coordinated treatments, mites were still present in high numbers in some colonies (**Figure 4-12, Figure 4-13**), however total mite numbers were significantly reduced across the island between 2017 and 2018 and remained lower in 2019, indicating that coordination was reducing *Varroa* infestations on an island-wide scale (**Figure 4-5**). An extension of this monitored coordination for a couple more years could have led to further improvements. Individual sites showed significant changes over time (**Figure 4-14**). Site 1 reported no *Varroa* by the end of 2019 and 'healthy' levels of DWV in all apiaries ($<10^6$ GE). The Site, located in the north of the island, is cut off from the rest of the beekeepers by a large range of hills and as such probably benefitted from this additional isolation, although it is unlikely that the report of 0 mites means *Varroa* are eradicated from the Site, rather that the numbers are very low and manageable. Additionally, the keepers in this location stopped moving colonies to and from the Site during the study. Site 2, in a similar manner, benefitted from its more isolated location, and showed improvements across the three years, with both mite levels and DWV titre reducing each year.

Given the location of Site 2 (green marker in **Figure 4-6A**), the high mite numbers are not likely to be a result of drifting from known colonies at other Sites, meaning that either the beekeeper has not managed the colonies sufficiently well, there are other unknown beekeepers in close proximity to Site 2, or there are feral bees located within close proximity which are infesting the site through drifting and robbing. Swarms may have emerged from Site 2 during the first, heavily mite infested, season. Swarming reduces mite levels in colonies, and colonies which do not swarm can have higher mite numbers (Fries et al., 2003), but if the colonies are overcrowded, regardless of whether they swarm or not, they are at high risk of mite transmission via robbing and drifting and at risk of dying over winter (Seeley and Smith, 2015). Given that Site 2 had 11 colonies in a small apiary with high mite abundance this seems a probable outcome. Additionally, surviving swarms which left the apiary in 2017 could act as 'mite bombs' for the managed apiary (Peck and Seeley, 2019), reinfesting the treated colonies. Feral colonies have been shown to have higher DWV levels than managed colonies, but untreated managed colonies can quickly reach similar levels to their feral neighbours (Thompson et al., 2014). The bees kept at Site 2 also appeared to be a different sub-species of honey bee compared to most of the other sites on the island (**Figure 4-23**), another indicator that swarms or drifting workers from an unknown site may have entered this apiary, particularly as all the island's colonies were purchased from a single source in 2015 and therefore likely to all be a similar sub-species. However, this information was obtained from a limited number of SNPs when evaluating a single honey bee gene, the COI, and its significance should not be overstated. The high mite and DWV burden, coupled with poor colony management, are more likely to account for the high colony losses observed at Site 2 in the first year of the study.

Given the size of the Isle of Arran (432 km²) and the size of the human population (~5000), it is possible that unknown colonies are present, or beekeepers not associated with the Arran Bee Group are managing colonies on the island without the group's awareness. Both are likely to contribute to mite reintroduction and disease transmission. Unknown swarms or beekeepers may also provide the background for the different sub-species observed at Site 2, as all the bees in the association were obtained from a single source in 2015. If a swarm or two were captured at Site 2 of a sub-species with similar COI to *A. m. lamarckii* (as defined by Sanger sequencing) they may have become the dominant sub-species on that site. This may also explain the persistently high mite levels observed at Site 5, which is in the most populous area of the island and had a high density of beekeepers. Swarms were reported by a couple of beekeepers during the study and others may have occurred and become established without their knowing.

Due to the scale of the project and the number of variables which may have impacted these results it was not realistic to expect the coordinated treatments to eradicate *Varroa* on the island, however the combination of coordinating the treatments for the first time and using treatments at the optimal time, including a mid-winter treatment for the first time, was hypothesized to reduce mite levels. Although reductions on average (**Figure 4-5**) were observed across the island, several Sites still had high or unchanged mite levels by the end of the study. Swarms, feral colonies and unknown beekeepers are the most likely reason why mites persisted at relatively high levels at some Sites on the island. However, miticide resistance and poor beekeeping practices can also play a significant role in *Varroa* infestations. Amitraz, used throughout this study as the commercially available Apivar, had not previously reported mite-resistance (Rinkevich, Danka and Healy, 2017), unlike other miticides such as Apistan (MJ Gracia-Salinas *et al.*, 2006). However, resistance has been anecdotally discussed amongst beekeepers and a recent report from the USA (Rinkevich, 2020) indicated that

although it was still an effective miticide treatment, there was a range of sensitivities to Apivar, with some mite-infested colonies completely resistant to it. In this study, while mites in some colonies within an apiary potentially showed miticide resistance, the apiary remained sensitive to the treatment overall.

On Arran, as the number of colonies is small and no new colonies were introduced, an environmental bottleneck has been introduced, which could inadvertently select for miticide-resistant mites. The treatments used in this study are unlikely to have increased the risk of resistance, given the applications were applied for the correct duration, at the most effective time of year and coordinated across all colonies on the island, however the beekeepers had used Apivar in an uncoordinated manner prior to the study beginning. This may provide an explanation for the persistently high mite numbers at locations such as Site 5, however this would need considerable further study, including sampling of mites from these colonies before and after treatments were carried out.

A marker of honey bee colony health is the diversity of the DWV population. Colonies with high mite infestations tend to lose the high viral diversity over time, as particular variants are elevated in titre by the altered transmission route, resulting in near-clonal variants dominating colonies (Martin *et al.*, 2012; Ryabov *et al.*, 2014). This study used DWV diversity as a measure of the colony health throughout the miticide treatment regime, with the hypothesis that, if colony health improved through coordinated treatments, the diversity of the DWV population would increase and the titre would decrease. As we showed in Chapter 3 with the shook swarm experiment, post-treatment and removal of mites the diversity of the DWV population can be 'rescued' in a honey bee colony (**Figure 3-22 and 3-23**).

The initial NGS analysis of individual workers indicated little diversity was present in the population at an individual worker level (**Figure 4-9 and 4-10**) regardless of the viral titre (**Figure 4-6B**). This was potentially because all colonies originated from a single source and there was a relatively high level of mite infestation throughout the island. Both factors likely contributed to a bottleneck event in the viral population (Woodford and Evans, 2020), where diversity was lost over time regardless of the titres recorded at the time of sampling. Additionally, as the mite and DWV history of the colonies prior to their movement to the island is unknown, it is possible that even after the removal/reduction of mites the virus diversity would have remained low. Due to cost and time restrictions, from the second season onwards samples were processed as pools from each apiary, rather than individual workers, however the pools from 2017 also showed the same pattern of low diversity or near-clonality across most sites, with the exception of Sites 1 and 3 (**Figure 4-16, 4-17 and 4-18**). In 2018 and 2019 there was a shift in the dominant variant present in the colonies and in the diversity observed across the sites. The samples from 2017 aligned to Type B reference sequences in the helicase region (**Figure 4-18**) and Type A-like sequences in RdRp region (**Figure 4-17**), indicating that the dominant variant was a recombinant form of DWV, potentially with Type B structural proteins and at least partial Type A non-structural proteins, similar to other recombinants observed in the UK (Moore *et al.*, 2011; Ryabov *et al.*, 2014). By 2018 Sites 5 and 6 were predominately infected with Type B variants, based on sequence alignments in both the RdRp and helicase regions of the DWV genome.

By 2019 all sites contained a larger percentage of Type B-like variants than Type A, although Type A-like variants did persist at lower levels in the RdRp region for Sites 1, 2, 3 and 4. The shift from 2018 to 2019 was probably facilitated by the number of colonies moved from Site 5 to other locations on the island, including Sites 2, 3 and 4. Why one variant can supersede

another in the DWV population is unclear, however various reports have shown that Type B-like variants are increasingly common (Natsopoulou *et al.*, 2017; Ryabov *et al.*, 2017) and are equally as infectious as Type A-like variants (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a), with some research suggesting they can replicate faster than Type A under in-vitro conditions (Norton et al., 2020). Additionally, DWV Type A is passaged by *Varroa* in a non-propagative manner (Posada-Florez *et al.*, 2019), whilst Type B and recombinants are able to replicate (Gisder and Genersch, 2020; Gusachenko *et al.*, 2020a). These recent studies provide a body of evidence that Type B-like variants, and recombinants thereof, may have selection advantages over Type A-like variants, which could explain the shift from Type A dominant to Type B dominant over the course of this study.

Survival of colonies may also be determined by the DWV variant dominant in the honey bee population. There have been contrary reports on the implications of DWV Type B infections, with reports in Europe demonstrating that Type B was found in higher titres in colonies and using modelling it was shown that those colonies collapsed faster (McMahon *et al.*, 2016). In the USA, colonies that died had a higher prevalence of Type A-like variants (89%) and a higher viral load (10^8–10^{11} copies) than colonies which survived (49% and 10^6–10^{10} respectively) (Kevill *et al.*, 2019). However, Type B-like variants were found at both a lower prevalence and lower viral load in colonies that perished (56% and $<10^{10}$ copies). Surviving colonies were also routinely found to contain high titres of Type B-like variants, perhaps indicating that Type A variants pose a greater threat to colony survival. Applying this theory to Arran, where Type A-like variants dominated in the first year only for a very high mortality rate over the winter of 2017/2018 to diminish colony numbers (**Figure 4-2** – 29 of 57 colonies perished), it is possible that the improved colony survival rate in subsequent years was due to the emergence of Type B-like variants across the island. The correlation matrix determined that year of sample and DWV dominant variant were positively correlated (0.72), indicating a link between later sampling and increased Type B presence on the island (**Figure 4-21**), which supports the sequence data. By the end of the study the beekeepers had 84 colonies, compared to an initial 57 and a low of 28 in Spring of 2018 (**Figure 4-2**). There are several possible reasons for this, including improved beekeeping practices and favourable weather conditions, but the shift in viral variant dominance may also have had an impact.

As with Chapter 3, the NGS data highlights the importance of understanding the make-up of the virus population and not relying on consensus sequence generation or less expensive Sanger sequence analysis to make inferences about the virus population. From the initial Sanger sequence analysis and the limited first ShoRAH analysis of individual samples, the 2017 samples appeared to be almost entirely clonal, but the pooled NGS analysis revealed the presence of low-level Type B variants at Site 1 and Site 3 (**Figure 4-17** and **Figure 4-18**), offering an explanation for the origin of the much larger presence of Type B-like variants in subsequent years. Despite this, poor assay sensitivity could explain a lack of diversity at some Sites in the first year of sampling, perhaps missing low-level Type B variants in the population. As all sequencing was carried out using amplicons there will be a bias in the variants sequenced, as some variants may not be selected by the primers and some variants may amplify more efficiently. A large selection of the samples also had very low DWV titres, meaning only partial genomes could be amplified and sequenced, perhaps masking the presence of recombinants or low-level variants in the populations.

Therefore, analysis of multiple regions of the DWV genome or indeed full genomes would be beneficial as DWV recombinants can form during viral replication (Moore *et al.*, 2011; Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). Despite not having complete genomes for all samples, the

data from the helicase and RdRp regions indicated a recombinant sequence was dominant in the first year, however this changed over time as RdRp sequences were more commonly reported as Type B- than Type A-like by the final year. It is possible selection pressure during viral replication either in the host or the mite caused a reduction of the recombinant form.

The scale of this project and the number of experimental variables indicate this was always a difficult experiment to control. The smaller scale experiments in Chapter 3 serve to highlight both the effectiveness of colony management techniques (such as a shook swarm and miticide treatment) and the impact of doing nothing to colonies in infested areas. These experiments were performed as 'controls' for the Arran study and highlight some of the shortfalls of this work. As we were unable to rule out the presence of other colonies or swarms it is possible that the mite numbers observed at several sites were a result of drifting and robbing of neighbouring mite-infested colonies, as observed in Chapter 3. Greater sampling frequency across the three years and more control of treatment application may have also benefitted the analysis and improved *Varroa* control, particularly the mid-winter applications of Apibioxal, which many of the beekeepers were doing for the first time, unsupervised.

Despite the continued observations of *Varroa* infestations at several Sites on the island there were many positives from the study. Most notably, the engagement of the beekeepers with the science of disease management and the large increase in colony numbers in the association. The Arran Bee Group consisted of 17 members with 57 colonies in the first year and despite the large losses at the end of 2017, they had 86 colonies between 22 beekeepers by the end of 2019. As suggested, there are several reasons why the number of colonies may have increased, including weather conditions and the potential difference in virulence between DWV variants, but an overall improvement in colony management was a significant factor, as evidenced by the group's ability to split and rear new colonies without the need to bring new bees onto the island for the whole three years of the study. *Varroa* continue to infest the colonies, but the beekeepers will persist with coordinated treatment routines now the study has finished. Some Sites have significantly reduced *Varroa* (Site 1, 2 and 4) or maintained low *Varroa* counts throughout (Site 3 and Site 6), so there is a strong base for further improvements across the island. One of the most significant hurdles in this project was cooperation and coordination between the beekeepers, as without their understanding of the science or engagement in the project, it would not have been possible; however the beekeepers were engaged, enthusiastic and open-minded to the challenges of the experiment, and ultimately benefitted from the project.

5 Examining DWV population dynamics using a mite/bee/virus model

"All models are wrong, but some models are useful." - George E. P. Box

5.1 Introduction and aims

Varroa infestations alter the dynamics of Deformed wing virus infections by introducing a new route of transmission, resulting in significant elevation of viral titres and overt symptomatic infections. Various studies have reported changes in the diversity of the virus population as a consequence of the altered infection route, with many reports finding that the introduction of *Varroa* results in near-clonality of a single variant in the infected honey bees on a colony level (Martin *et al.*, 2012) and at an individual level (Ryabov *et al.*, 2014). There are likely several reasons why one DWV variant reaches clonality in the population. It is assumed that only certain variants are taken up by the mites during feeding, or that the number of viruses taken up is low, similar to observations in mosquitoes when vectoring dengue virus, where a stringent bottleneck event results in only 5–42 virus copies moving from host to vector (Lequime *et al.*, 2016). Certain DWV variants are known to replicate in the mites, such as DWV Type B and recombinants thereof (Gisder and Genersch, 2020; Gusachenko *et al.*, 2020a), but DWV Type A has been shown to be transmitted between mites and bees in a non-propagative manner, where the virus does not replicate in the mite (Posada-Florez *et al.*, 2019). This creates an obvious selection advantage for Type B-like variants if they pass from the gut to the salivary glands and are infectious after replication in the mites.

As the viruses in healthy honey bees exist at low levels ($<10^6$ GE – see Chapter 3), detecting them can depend on assay sensitivity and poor methodology, including in tissue handling, RNA extraction, cDNA synthesis and qPCR sensitivity, can jeopardise this. Even when samples are extracted and detected at low level, this often only gives us a superficial indication of total viral load and not population diversity. Therefore, expensive and time-consuming techniques, typically using next-generation sequencing, have to be used to determine virus population diversity. Improvements in sequencing techniques and reductions in costs have made this process easier in recent years, but it remains an expensive and time-consuming method that is out of reach for many studies.

An alternative, and potentially informative, method is to simulate virus diversity changes in the colony using an *in silico* model. This is both cost effective and can be informative for future experimental designs. Modelling tools, such as BEEHAVE (Becher *et al.*, 2014), exist to study the impacts of a range of environmental variables on honey bee colonies. These variables include *Varroa* infestation and *Varroa*-transmitted viruses, including the capacity to have infected or uninfected bees in colonies. However, these tools operate at a colony level and do not specifically focus on virus population dynamics or virus diversity. As the dynamics of DWV diversity are likely observable at an individual worker bee level, a more sensitive and focused model would be beneficial.

The aim of this chapter was to make a model that could predict changes in virus diversity over time in a honey bee colony at an individual pupa level, where mites infest and alter the virus dynamics across multiple generations of honey bee pupae. The model uses a range of variables which may alter transmission rate, amplification, and virus diversity, including potential bottleneck events and selective variant bias.

5.2 Methods

5.2.1 Model parameters and rationale

The model was designed to include a list of fixed and variable parameters that influence DWV diversity (see **Table 5-1**), based on findings from previously published literature and from our own experimental findings. Due to the large number of variables in the model, where possible a variable was fixed or kept within a narrow range. Known variable parameters, such as *Varroa* bite volume (Bowen-Walker and Gunn, 2001), were analysed across a range in some instances.

The model generates a seeded virus population in the initial pupa of 1000 copies of 6-12 DWV variants, assigned identifiers A–F or A-L. The virus titre is based on evidence observed in Chapter 3 and published work (Gusachenko *et al.*, 2020a) that healthy honey bees contain approximately $<10^6$ GE/µg RNA of DWV and that low-level DWV populations tend to consist of a mix of variants (Ryabov *et al.*, 2014). The number of variants has been shown to range between 3 or 4 (see Chapter 3) and 8 in some populations (Annoscia *et al.*, 2019), and up to 19 variants in pooled samples of 30 workers (Martin et al 2012). As this model works on a bee-by-bee basis, pooled samples were discounted, and for mathematical simplicity we restricted this model to 6 unique variants in the initial population.

Varroa feeding on a developing pupa remove on average 0.67 µl of haemolymph over a 24-hour period. This accounts for approximately 3% of the bee's water weight (any extra water being held in the body) and over 165% of the mite's body weight (Bowen-Walker and Gunn, 2001). Based on the likelihood that this bite volume varies from mite to mite, one of the ranges applied to the model was a shifting-scale bite volume of between 1% and 5% of the bee's water weight, referred to as the 'bite limit'. The bite limit was fixed for each run but could be changed between runs to examine the impact taking a higher or lower number of variants had on virus amplification and diversity. Analysis of the mosquito/dengue vector–virus system has indicated that the volume of viruses obtained by the vector during feeding acts as the 'pinch point' in the virus population, where mosquitoes take between 5 and 42 dengue virus variants with their blood meal (Lequime et al., 2016). This causes a bottleneck in the virus population. In this model, a similar restriction was imposed on the DWV population as it moves from pupae to mite, with a tested range of between 10–100 viruses in each bite. A higher number of viruses (1000) was also examined.

Virus amplification in the gut of the mite, post feeding, was set at ×100,000. This was based on evidence of viral replication of at least one variant in *Varroa* (Gusachenko *et al.*, 2020a) and evidence of DWV replication in honey bee pupae from 10 copies to 10^7 within 24 hours post-injection. Injection volume was defined based on an assumption that only a percentage of the replicating virus population found in *Varroa* will move from the gut to the salivary glands and then enter a new pupa when the mite feeds. This was created as a sliding scale of between 1% and 10% of the 100,000 copies of DWV found in the mite and was based on unpublished data which found at least 10% of the gut virus concentration was present in the salivary glands of mites. A large volume of saliva was expelled by the mites when feeding on an artificial system, suggesting a high percentage of the viruses in the salivary gland would be injected into a pupa during feeding. The passing of amplified variants back to the pupa represents another potential bottleneck in the virus population as it is possible that a small number of variants enter the new pupae or that a selective advantage occurs here, whereby certain variants are able to avoid the bee's immune response or replicate faster in the host.

Amplification of the viruses injected into the pupae was fixed at 1,000,000 virus copies based on 24-hour amplification rates of 10^6 GE per bee from Gusachenko *et al* 2020a, when injecting pupae using low copy number reverse genetic constructs. Each developing pupa is fed by a nurse bee in the colony, so will also receive a low-level mixed DWV population via this transmission route, and as such the 1000 virus copies of 6 variants are reintroduced in each cycle of the model. The model was run for 30 cycles to reflect 30 feeding events over a significant period of time within a single bee season. This number was selected to reflect that *Varroa* feed on different pupae during their lifecycle and viruses can be transferred from pupa to pupa via this route. For simplicity this model examines a single feeding event per pupa, which is unlikely to reflect the nature of mite feeding in the colony (shown in **Figure 5-1**), where the mother will feed on a pupa and her progeny will then also feed on it, consequently extracting some of the amplified viruses the mother has injected into the pupa, before all the females vacate the cell and move to other pupae. The model was intended to infer changes in viral diversity over time by examining potential bottleneck events and for this reason was simplified to only examine a single feeding event per pupa (see **Table 5-1** parameters), rather than multi-generational dynamics involving brood and multiple feeding events which would have required a much more complex system of analysis.

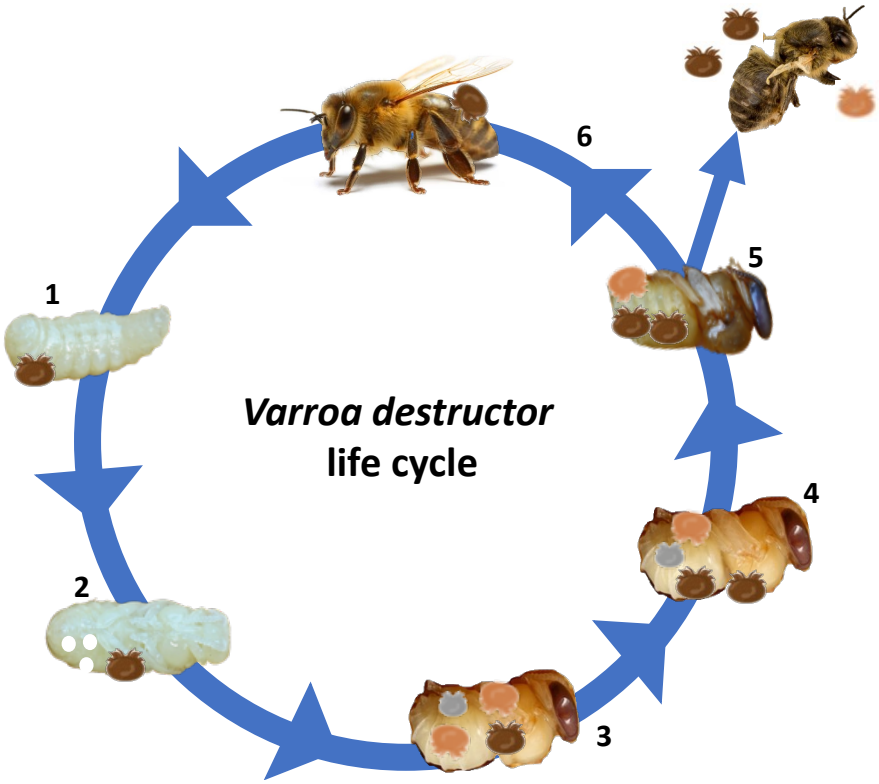


Figure 5-1 - The* Varroa *mite reproduction cycle. *1 - A mite enters an open brood cell with a larva. 2 - The mite lays its eggs in the cell. 3 - Juvenile mites hatch, the male (grey) reaches maturity before the females (orange). 4 - The male reproduces with the female juveniles and all mites feed on the developing pupa. 5 - The bee ecloses and the adult female and 1–2 juvenile females emerge with it. Some bees emerge with deformed wings. 6 - The female mites leave on the back of the adult worker and spend a week on their back before entering a new brood cell.*

Table 5-1 - Parameters of the mite model. *Each parameter in the model, whether that parameter is fixed or variable and where the information for the criteria originates from. The number in parenthesis indicates where each parameter occurs in the schematic* ***Figure 5-2****.*

Parameter	Fixed or Variable	Reason
Viruses in bee gut (1)	Fixed – 1000	Based on qPCR analysis of healthy bees (Ryabov *et al.*, 2014)
Virus diversity in bee gut (1)	6 – 12 variants	Various literature (see text for details)
Varroa bite volume (1)	3% – 0.7 µl	(Bowen-Walker and Gunn, 2001)
Varroa bite limit (viruses) (2)	10 – 100	5-42 dengue copies in mosquitoes (Lequime et al., 2016)
Virus amplification in mite (2)	100,000-fold	qPCR data (this project, not shown)
Virus injection by mite (3)	10% population (salivary gland)	Approximation, based on unpublished experiments
Virus amplification in next bee (3)	1,000,000-fold	Based on pupal injections (Gusachenko et al 2020a)
Virus bias in mite (3)	% of certain variant	Based on DWV-A vs DWV-B replication in mites (Posada *et al.*, 2019, Gusachenko *et al.,* 2020a)

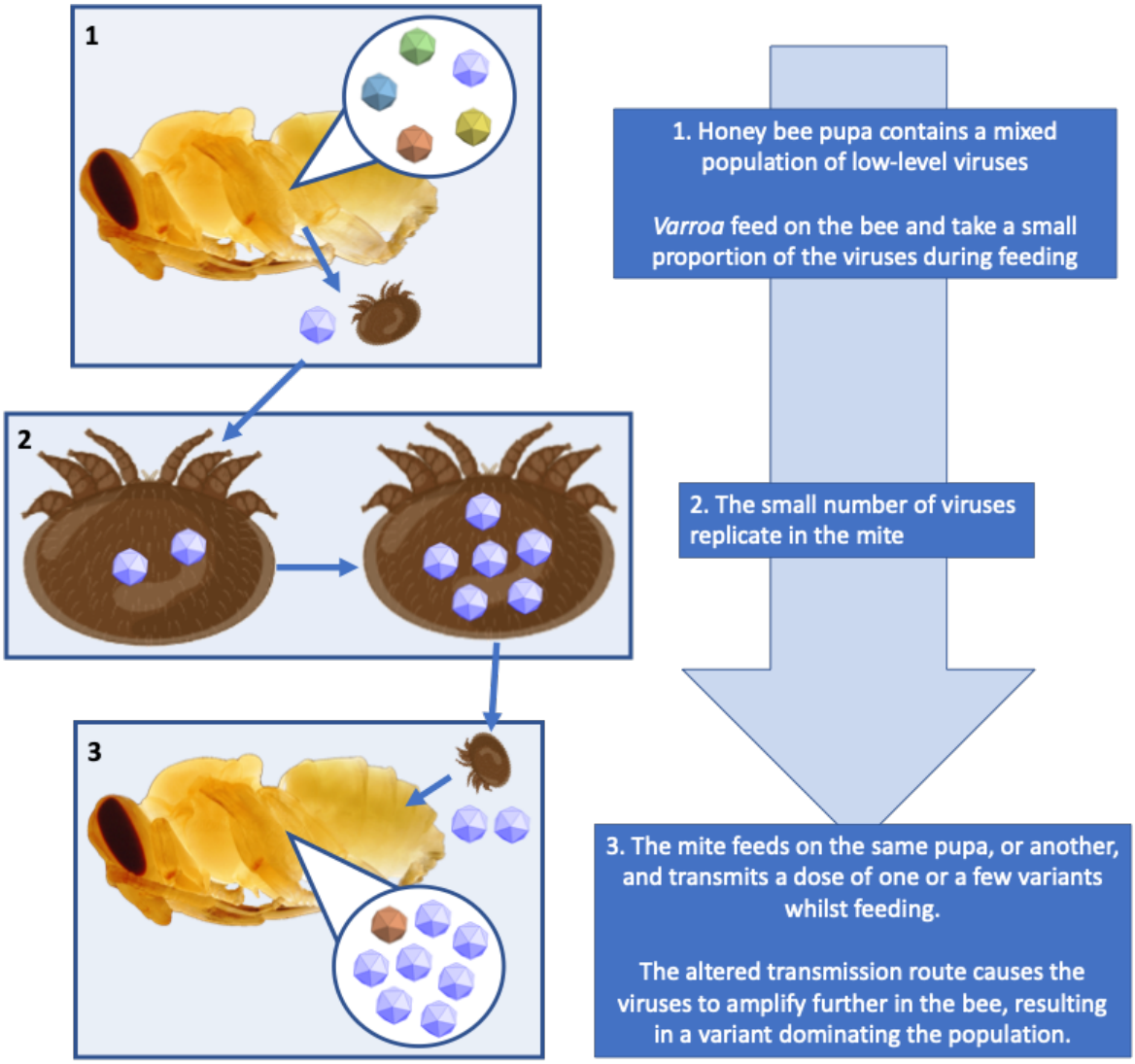


Figure 5-2 - Schematic representation of virus transmission from a developing honey bee pupa to a parasitic mite and then on to subsequent pupae in the honey bee colony. *An initial low-level diverse virus population goes through a bottleneck when mites acquire the viruses. These then amplify in the mites before being injected into the pupa, where the elevated viruses then go through another amplification step.*

5.2.2 R script process and defined parameters

A set of variables based on the parameters shown in **Table 5-1** were defined in the R model prior to generating a loop to run the 30 cycles (**Table 5-2**). To generate the model in R a dataset of 6 variants, simply defined as 'A–F' were created and defined as 'strains'. The R script (see **Appendix 4**) was set up using a 'for loop', with 30 cycles as the defined number of 'feeding cycles'.

Table 5-2 - Variables contained in the model. *Before starting the for loop, the following variables were defined. These variables were determined and fixed for the full 30 cycles of each run, except the last two which were contained within the for loop and could therefore be randomised in every cycle or fixed at a set value for the full 30cycle run.*

Model Variable	Value	Reason
VirusLoad_L	1000	The number of viruses in the first pupa
VirusLoad_M	100,000	The amplified virus in the mite gut
VirusLoad_H	1,000,000	The amplified virus post-injection in pupae
InjectLimit	10,000	A cap on the total number of viruses the mite injects (10% of VirusLoad_M)
BiteLimit	10–100	Maximum number of viruses mite takes during a bite
Strains	A–F or A-L	The number of different variants in the population
StrainFitness	1–100%	The bias of any variant in the population (16.7% each for no-bias)
The following variables were embedded in the for loop		
Varroabite	1–5%	A sliding variable for the % of the virus population sampled during feeding
Varroainject	%	A sliding variable for the % of virus injected into the next pupae

Briefly, the for loop takes the 1000 viruses generated from Strains and VirusLoad_L (**Table 5-2**) and randomly samples a small number of them based on the BiteLimit cap, to create ref_virusPop_bite. The selected variants are then amplified in the mite (ref_mitePop_tmp) using VirusLoad_M as the maximum cap. An if-else loop then determines if the viruses in the mite have reached the limit imposed by InjectLimit, or not. If not, it takes a percentage defined as 'injectSize'. InjectSize or InjectLimit then define the number of viruses passed to the next pupa, with a maximum amount defined outside the loop (100,000).

A new pupa is generated with 6 or 12 virus variants, totalling 1000 copies (VirusLoad_L + new_beepop) and the capped VirusLoad_M is added (ref_mite_inject). This new mix is then amplified (ref_bee_amplifier) using VirusLoad_H to cap the amplification at 1,000,000. The resulting virus population reflects the virus amplified in the newly injected pupae and the low-level mixed population passed *per os*. This process repeats for 29 further cycles before stopping. The viruses in the population are calculated as a percentage of the total virus population and are determined at the point where the amplified virus is mixed with the VirusLoad_L in the new pupae (ref_virusPop_tmp), therefore after each cycle a breakdown of the percentage of each of the 6 virus variants (A–F) in the overall population is reported.

The for loop is set up within another for loop, which runs the entire process 10 times and bins the output data into a .csv file, resulting in 10 iterations of the 30-cycle loop compiled in a

single datasheet. These data were then used to generate the averages for the viruses shown in the figures in the results section below. A single run of the model typically took ~2 hours to generate 10 iterations.

5.3 Results

5.3.1 Running the model with no virus bias

In the first instance the model was run with no bias set for a particular variant in the original 1000 virus population, so all 6 variants were amplified at an equal proportion. This implies there is no selective advantage during the initial amplification of the viruses and that acquisition of variants by the mite during feeding is simply by chance. Whichever virus is taken up during feeding then elevates in the mite gut. There was no bias set towards a particular variant in the virus population and the bite volume was set at 3% and the volume of virus the mite injected into the next generation of pupae was set at 10% (see section 5.2.1). One variable was changed on each run initially, this was the bite limit of the mite as it samples the virus population. This was set as a sliding window of between 10 and 50 viruses based on published findings in other vector–virus systems. One variable at a time was analysed for simplicity and although this does not reflect an *in vivo* system, it meant different variables could be examined for their potential influence on the diversity of the virus population over time.

In the first run, 10 iterations of 30 cycles, the *Varroa* bite limit was set to 10 virus copies and in each iteration one variant reached clonality (>99.9% of the virus population) in the virus population by a maximum of 19 cycles (**Figure 5-3**). Each virus that became elevated above the rest in the population over the course of the 30 cycles was labelled as Virus A for simplicity.

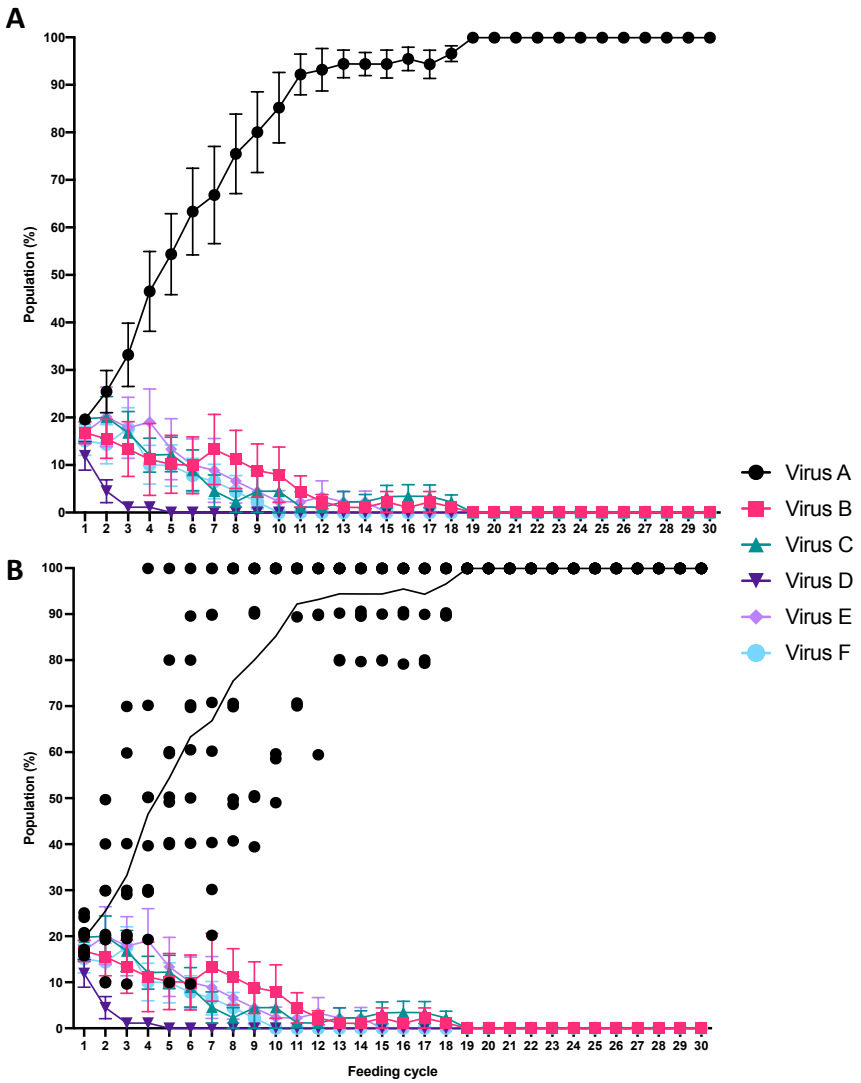


Figure 5-3 - Varroa bite limit of 10 viruses with a subsequent 10% injection volume. *This iteration leads to a rapid clonality in the virus population. All iterations had a single clonal virus (>99.9%) by 20 cycles or fewer. Plot (**A**) shows the averages of 10 iterations of the 30-cycle model with the error bars indicating the SEM, plot (**B**) shows the same figure, but with all iterations of Virus A shown.*

For the next two runs the *Varroa* bite limit was increased to 25 and 50 viruses, with all other parameters remaining fixed. In these instances, the number of iterations with a variant reaching near-clonality decreased as the number of viruses acquired by the bite increased, with clonality only reached in 3/10 and 2/10 iterations when 25 and 50 viruses were acquired respectively (**Figure 5-4**). This shift indicates clonality may be determined by the number of variants of DWV the mite acquires during its feeding and their subsequent amplification; however one variant was still dominant (>50% of all viruses) in each population set by the end of the 30 cycles. Increasing the bite limit further to 100 viruses resulted in no variants reaching clonality (99.9%) in the population over the 30 cycles, and no variants reached a level whereby they were dominant in the population, with Virus A reaching ~40% of the population after 30 cycles (**Figure 5-5**).

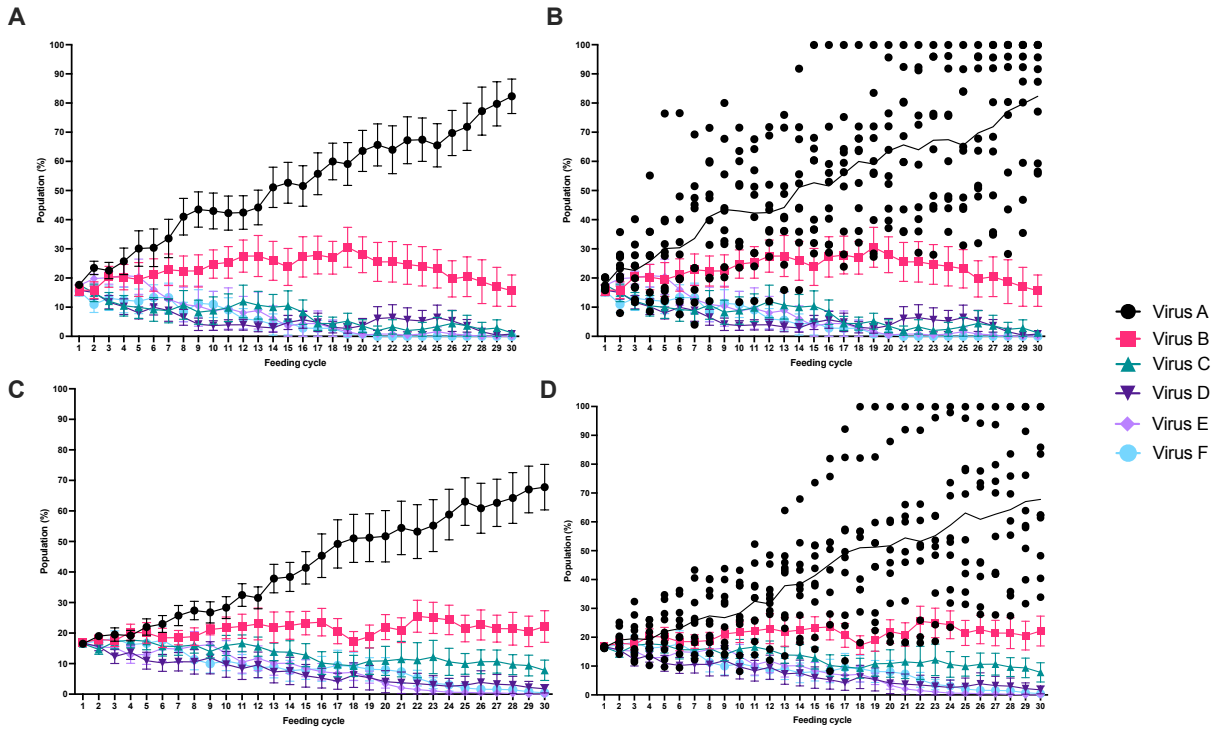


Figure 5-4 - Increasing the number of viruses obtained by the mite alters the dynamic. *With a 25-virus bite limit (**A** & **B**) the population reaches clonality in 30 cycles in 3/10 iterations and the population average for Virus A is >80%. With 50 viruses (**C** & **D**) obtained in each bite the dominant variant only reaches clonality in 2/10 iterations, however the bottleneck still causes one variant to consistently increase over time in all iterations and the final population average for Virus A is ~70%.*

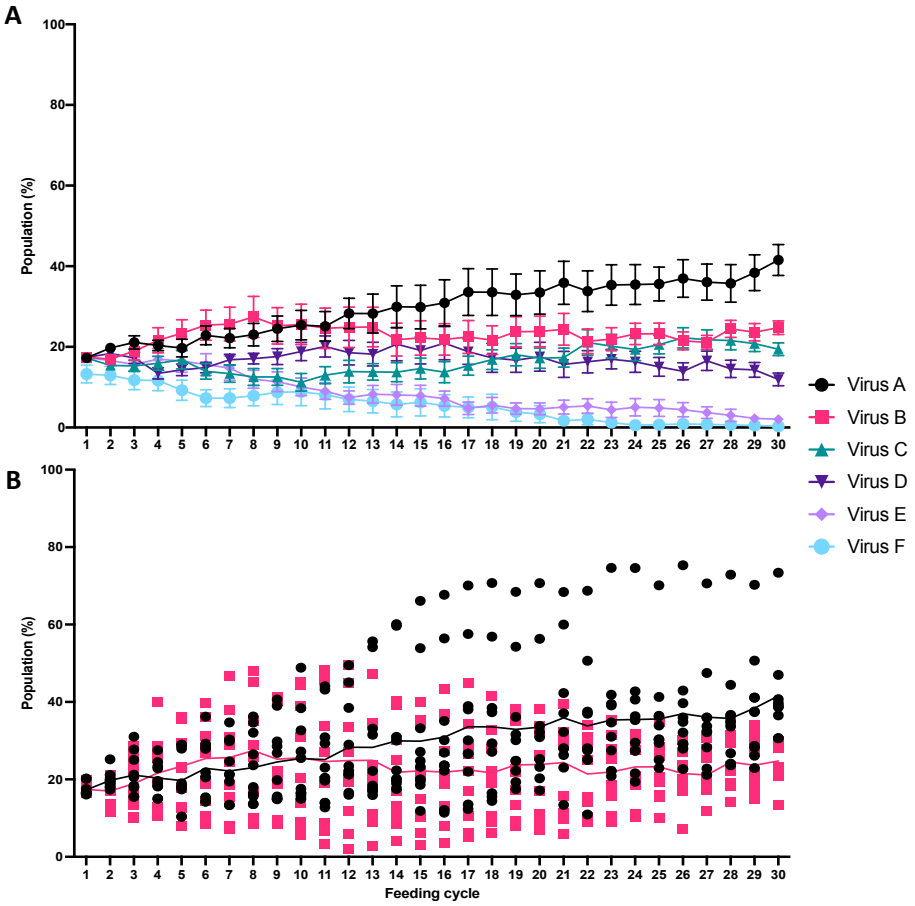


Figure 5-5 - Changes in the virus population when the mite bite volume is set to take 100 viruses per pupa feeding. ***A*** *- When the bite limit is increased to 100 no virus emerges to dominate the population over the 30 cycles.* ***B*** *- Virus A and Virus B, all iterations shown for 100-virus bite limit plot, with viruses C–F omitted. In no iterations did a virus reach clonality and on average the most abundant virus accounted for less than 40% of the total virus population. As per the other plots the virus which accounted for the highest percentage of the population at the end of the 30 cycles was labelled Virus A.*

As the number of DWV variants found in low-level asymptomatic honey bees has been reported across a dynamic range (3–19 variants – see section 5.2.1) the model was also set to examine whether clonality was possible when 12 viruses were present in the first honey bee pupa. From **Figure 5-6A** it is clear that increasing the number of viruses found in the first pupa slows the rate at which a variant reaches clonality from 19 cycles to 30 cycles when the bite limit remains low (10), but all iterations still reached clonality within the 30 feeding cycles. As with the 6-virus population, increasing the bite volume to 25 reduced the likelihood of a virus reaching clonality in the population with clonality only observed in 2/10 iterations of the model (**Figure 5-6B**). The dominant virus did make up >90% of the virus population by cycle 30 in 5/10 iterations, with other iterations resulting in a 2-virus co-infection of roughly even split.

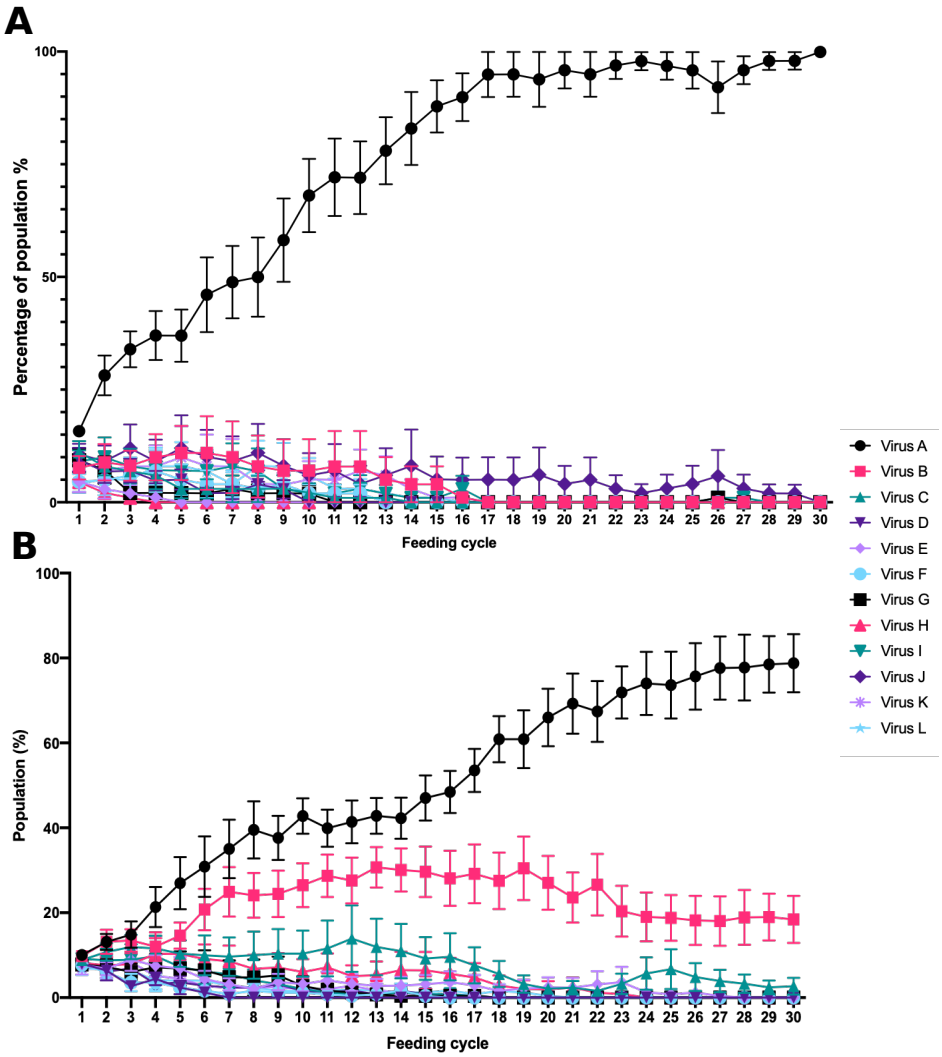


Figure 5-6 - Effect of changing 'Strain' to 12 viruses with a bite limit of 10 (A) and a bite limit of 25 (B) and 10% injection limit imposed. *The dominant virus reached clonality in all iterations when the bite limit was set to 10 but, as with the 6-virus population, the number of clonal iterations was significantly lower when the bite limit increased to 25 – just 2/10 iterations.*

From these results, increasing the bite volume – and thereby increasing the number and potential diversity of the viruses in the inocula – rapidly reduces the frequency with which a single virus reaches clonality in the population, regardless of the starting population of viruses. The volume of virus that the mite injects into the next pupa it feeds on will also have a significant bearing on the amplification and selection of the virus population. As can be seen in **Figure 5-7**, altering the injection volume (the volume of viruses the mite puts back into a pupa), but leaving the bite volume unaltered, shifts the virus population dynamic (see **Table 5-2**). Clonality (>99.9%) still occurs but takes longer to occur across the feeding cycles with 6 variants, typically happening >25 feeding cycles into the model run**.** Despite the bite volume remaining at 10 viruses clonality is only achieved across 8/10 iterations, regardless of whether the injection volume is set to 20% (**Figure 5-7 A & B**) or 30% (**Figure 5-7 C & D**).

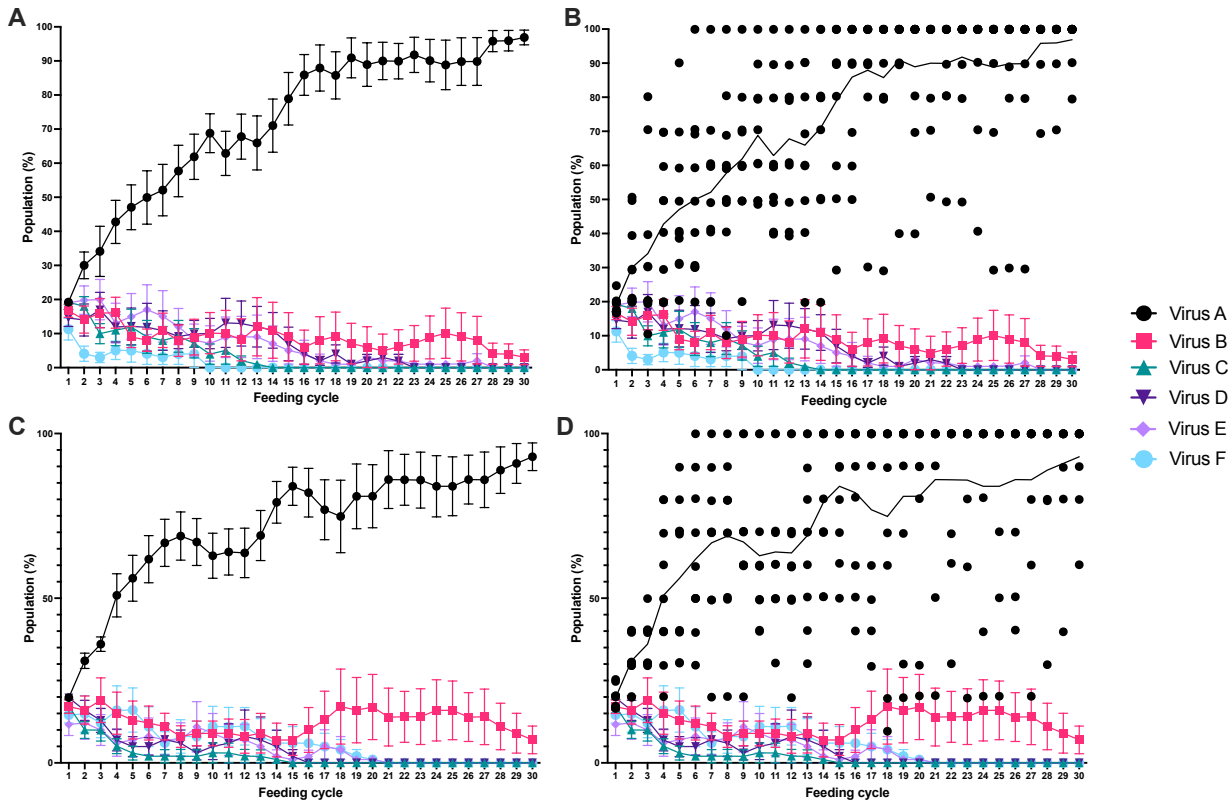


Figure 5-7 - Bite limit fixed at 10, but modifications to the injection percentage when mite next feeds. *The mite bite limit was set to 10 viruses but a 20% (**A** and **B**), and 30% (**C** and **D**) injection volume was set when the mites fed on the next pupae in the colony, an increase from 10% for the previous models. Plots A and C show the average with the SEM, plots B and D show all iterations for the dominant virus (Virus A). Clonality is achieved by one virus in 8/10 iterations with 20% injection volume and 7/10 iterations with 30% injection volume.*

5.3.2 Introducing a DWV variant bias into the initial mixed population

In section 5.3.1, all models were run without a preference towards a particular variant in the first virus population. The variant which ended the 30 cycles with the highest percentage of the population was assigned to 'Virus A' during data collation. Given recent reports that certain variants of DWV can replicate in *Varroa* mites (Gusachenko *et al.*, 2020a) whilst others may simply be vectored in a non-propagative manner (Posada-Florez *et al.*, 2019) it is clear that some viruses may have a selective advantage in the cycle between mite and host. With this in mind, a selective advantage was conferred upon single or multiple variants in the simple A–F mixed virus population.

In the first run one virus comprised 50% of the initial population, whilst the other five viruses accounted for 10% each. The same fixed parameters were used to run the model as per section 5.3.1 (10 virus bite limit, 10% injection into next pupae, 10^6 virus amplification in pupae), but this method resulted in very few instances of clonality in the population for the biased variant (3/10). The other viruses reached clonality in 5/10 occasions despite the imposed bias and starting from an initial 10% introduction, with the remaining 2/10 iterations containing mixed samples (**Figure 5-8**). So, 8/10 iterations reached clonality for the total virus population, likely due to the bite limit imposed on the model.

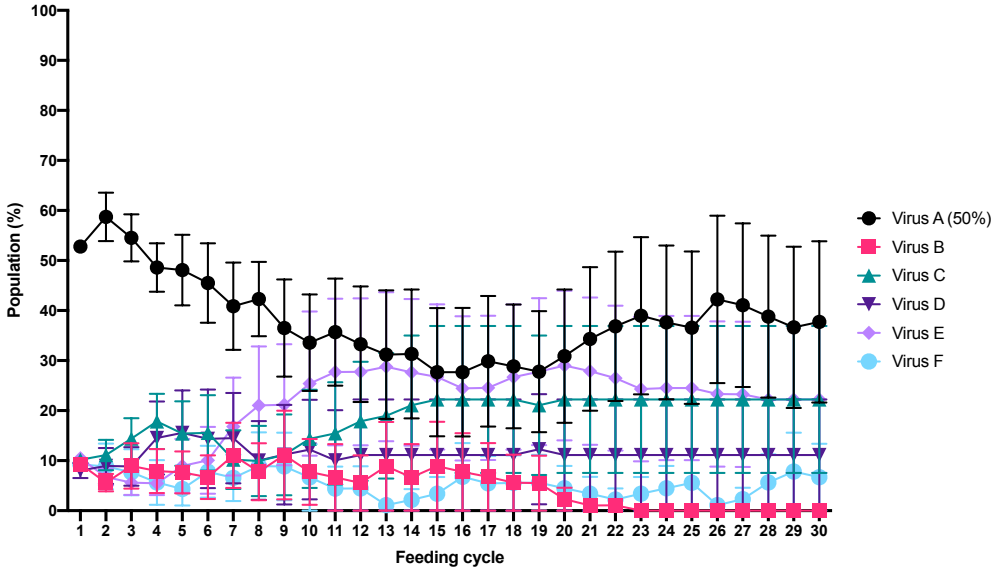


Figure 5-8 - Virus A accounts for 50% of the 1000 viruses in the initial population. *8/10 runs reach clonality, but only 3/10 were for virus A. 2/10 populations finished the 30-feeding cycle run with a mixed population despite the stringent bite and injection limits imposed on the model.*

In the next run, Virus A accounted for 75% of the population, with the remaining 5 viruses accounting for 5% each of the population. This resulted in 9/10 iterations reaching clonality, all for Virus A (**Figure 5-9A**). The one run which didn't reach clonality resulted in a 2-virus population with a split of 60%/40%, with Virus A still being the most prevalent in that instance. The selective advantage for a single virus was further examined by increasing Virus A to account for 90% of the initial population. However, this did not result in a significant shift in viral dominance when compared with 75%. In this instance Virus A reached clonality in 8/10 iterations within 16 cycles and the remaining 2/10 resulted in clonality for a different virus in the mix, despite them each starting at just 2% of the initial virus mix in the first pupa (**Figure 5-9C**). The large margins of error between the 75% and 90% populations and the clonality of 2/10 iterations of low-level variants in the 90% bias run highlight the stochastic nature of the system. **Figure 5-9B** and **5-9D** show the iterations of Virus A that despite starting with a high bias, eventually finished the cycle making up a low percentage of the overall population (20% in 5-9B and <1% in 5-9D). This is particularly apparent when examining a small sample size population, whereby individual viruses can gain dominance in a population not because of a dominant mechanism but because of a chance event. If such chance events happen early on in hive infestation, the virus which comes to dominate might well be widely propagated.

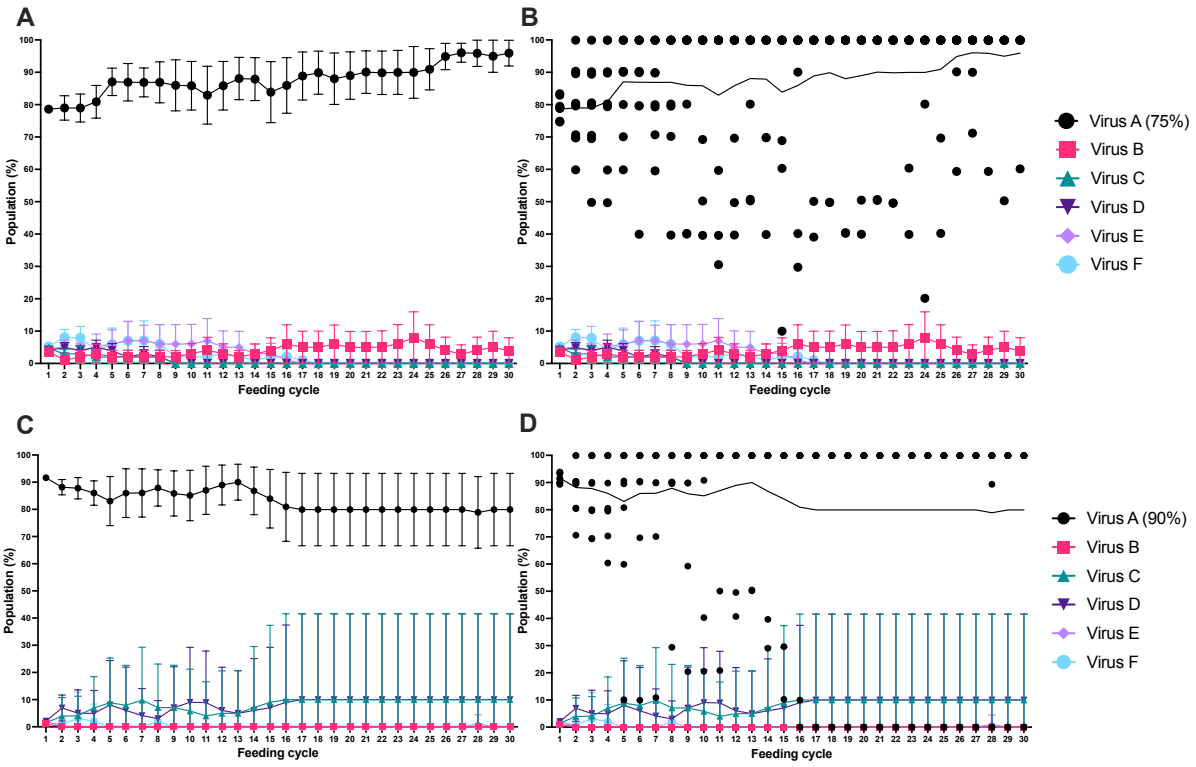


Figure 5-9 - Virus A bias in the 1000 virus population. Virus A accounted for 75% (A) and 90% (B) of the initial virus population in the first pupae. *When 75% of the population was virus A (**A & B**) 9/10 iterations of the model reach clonality by cycle 30, with the tenth iteration resulting in a 60/40 split between Virus A and another virus. When Virus A accounted for 90% of the population (**C & D**) 10/10 iterations of the model reach clonality by cycle 16, but two of the ten iterations produced clonality of a secondary virus, starting from a much lower percentage of the viral population.*

5.3.3 Competition between two elevated viruses

As DWV Type A and Type B are often detected in honey bee colonies, it can be assumed that co-infections are common, most obviously evidenced by the frequent detection of recombinants (Moore *et al.*, 2011; Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). To examine this, the model was set with a weighted bias towards two variants in the initial virus population, so that they could theoretically compete with one another in the vector. In **Figure 5-10A** a 40% weight was given to Virus A and Virus B with the remaining variants on 5% each. All other variables were fixed, with a 10-virus bite limit and a 10% injection limit applied throughout. Clonality was reached in 10/10 iterations, 3/10 for Virus A, 5/10 for Virus B and 2/10 for the low-level variants (those starting at 5% of the population), therefore the dominant variant was one of the weighted variants 80% of the time when they made up 80% of the initial population. This was tested further by increasing the number of viruses acquired by the mite during feeding to 25 (**Figure 5-10B**). This produced a significantly more 'random' amplification of certain variants. 30% of runs reached clonality with Virus A, 20% with virus B and 10% with another low-level variant. Even with the initial bias in the data set, 4/10 iterations did not reach clonality. This result of 6/10 iterations reaching clonality is higher than the observed clonality of just 3/10 iterations when 25 viruses were samples at each bite with 6 equally proportioned variants (**Figure 5-4A)** and highlights that with a bias for two variants clonality may occur more frequently with a larger bite volume.

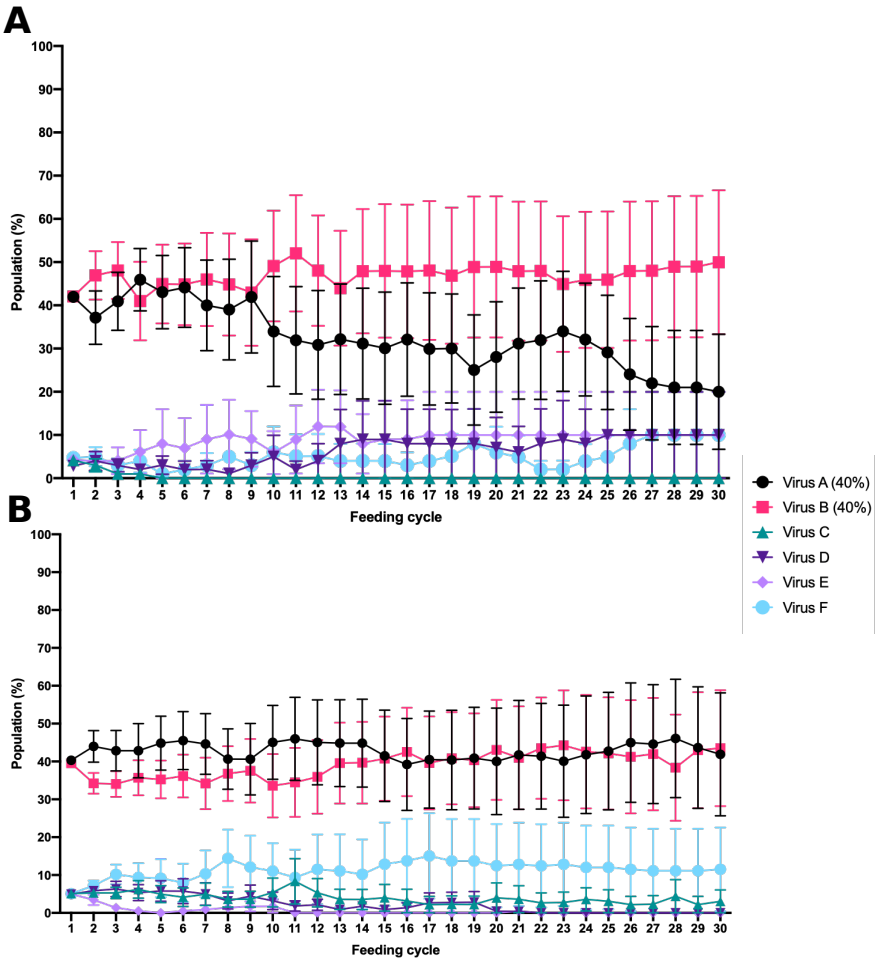


Figure 5-10 - Virus A and Virus B account for 40% each of the initial virus population. *In this run the bite limit was set to 10 viruses (**A**) and 25 viruses (**B**) and the injection percentage was left at 10%. With 10 viruses per bite, the population reached clonality in all 10 iterations, with 30% for Virus A, 50% for virus B and 20% for other viruses. With 25 viruses per bite clonality was reached in 30% of iterations by Virus A, 20% by Virus B and 10% by one of the low-level variants, with the remaining 4/10 iterations not reaching clonality over 30 feeding cycles.*

5.4 Discussion

To better understand the dynamics of the DWV population and the influence *Varroa* has on it we need to know the relative virulence of different DWV variants. Tools such as NGS analysis can be used to quantify and qualify virus presence, but they remain a time-consuming, technically challenging and expensive means of analysis. Generating simple models to measure changes in honey bee colonies can be a useful low-cost alternative. Existing models for measuring *Varroa* infestation rates, amongst other things, such as BEEHAVE exist (Becher *et al.*, 2014), but at present no model exists for measuring DWV population changes and the factors which may influence these changes. Here, we have shown that simplistic modelling of the variables which shape the DWV population can allow us to demonstrate the significant changes in virus diversity observed when *Varroa* feed on developing honey bee pupae. Some iterations of the model indicate that a near-clonal high level accumulation of DWV can occur within a few generations of honey bee pupae from an initially very limited viral inocula.

In these models the reduction in viral diversity is achieved without any assumptions about particular variants within the virus population (**Figure 5-3**). Instead, it assumes stochastic changes in virus load can occur at multiple 'pinch points' in the cycle between vector and host. Mosquitoes transmitting dengue virus can take as few as 5–40 virus copies during a bloodmeal and this imposes a severe selection pressure on and reshuffling of the virus diversity as they replicate in the mosquito midgut (Lequime *et al.*, 2016). A similar logic was applied to the mite–pupae model here, based on this research and the reported small volume of haemolymph taken by mites during feeding (0.67µl) (Bowen-Walker and Gunn, 2001). The effect on viral diversity of reducing the number of viruses taken by the mite was stark, with 10 viruses resulting in a single variant reaching clonality in each run in under 20 cycles. But as the number of viruses increased, the number of iterations reaching clonality decreased, with a 25-virus and 50-virus bite limit reaching clonality in 3/10 and 2/10 iterations respectively (**Figure 5-4 & 5-5**) and when 100 viruses are taken in a single bite no variant reaches clonality in any of the runs (**Figure 5-6**).

The bite limit was the most significant bottleneck in the amplification of a particular viral variant, which is unsurprising given it was one of the smallest quantities in the model. Increasing the number of viruses in the initial pupa did not significantly alter the dynamic changes in virus diversity, with the bite limit of 10 still being the significant factor in a variant reaching clonality. Changes to injection limits (when the mite feeds on the next pupa and injects virus as it feeds) between 10 and 30% had little influence on the dynamics of the virus population, although the log amplification of the virus was at the lower ends of estimates in the models as higher amplification levels significantly slowed the model down. As recent studies have found that DWV Type B and recombinants thereof replicate in mites (Gisder and Genersch, 2020; Gusachenko *et al.*, 2020a), whilst DWV Type A appears to be transferred in a non-propagative manner (Posada-Florez *et al.*, 2019), a bottleneck at this stage followed by the subsequent selection advantage for Type B-like variants would significantly alter the dynamics of the viral population. When a bias towards a particular variant prior to mite feeding was created, as shown in section 5.3.2, with the already imposed bite limit it was assumed that the weighted virus would reach clonality. With an initial 90% bias towards a single variant in cycle 1 clonality was reached in 100% of 30-cycle iterations, however in 2/10 of those iterations the final clonal virus was not the 90% weighted variant (**Figure 5-9**), indicating that it is the dynamics of the bottleneck event and subsequent amplification that shape the population as well as the starting population. Some of the stochastic changes in the virus models, such as the switch between Virus A and Virus B during co-infections (**Figure 5-10A**), may also explain some reports using qPCR detection methods which find infections with one variant in one season and then a different variant by the following season in some colonies (Kevill *et al.*, 2019).

Based on the results of this model and the body of research indicating Type B-like variants replicate in *Varroa,* it could be assumed that a population of Type B variants would replace Type A viruses in most mite-infested honey bee populations, however DWV Type A continues to be found in infested colonies long after mite infestations (see Section 3.2.8). This suggests that DWV Type A has a selection advantage over DWV Type B not considered in this model, such as a further bottleneck event, possibly in the salivary glands of the mites, or a better means of avoiding an immune response in the pupae post-injection. Research comparing injections of Type A and Type B reverse genetic stocks in pupae indicates the variants replicate to similarly high titres over a 48-hour period (Gusachenko *et al.*, 2020a) and in some cases a wild-type Type B appears to replicate faster (Norton et al., 2020). This implies DWV Type A probably does not have a selection advantage in the honey bee host. Conclusive proof that Type A does

or does not replicate in *Varroa* has yet to be shown, so it remains a possibility that at least some variants of Type A could replicate in mites, or that they are better able to avoid a potential mite immune response or navigate to the salivary glands of the mite better than Type B variants, possibly giving them a selection advantage prior to re-injection into the host.

There are clear limitations to this model: given the large number of variables and unknowns, some assumptions had to be made prior to running it and several factors had to be simplified, most notably the log change in viral load in pupae as infected pupae can typically have between 10^{10} and 10^{13} genome equivalents of DWV (Ryabov *et al.*, 2014; Gusachenko *et al.*, 2020a). Attempts to run the model with increases of 10^{10} or 10^{11} copies in pupae invariably caused the model to crash (due to computational limitations). *Varroa* express a salivary protein which is toxic to *Apis cerana* worker larvae, but not drone brood, indicating a preferential bias for drone brood may exist (Zhang and Han, 2018). The preferential movement of mites to certain brood cells was not considered in this model, although it is understood they preferentially enter drone brood cells in *Apis mellifera* colonies too (Boot et al., 1995). There will also be natural variation in mite feeding time, volume and subsequent injection rates in pupae. In the case of ticks, the tick-borne pathogens are affected by temperature and tick feeding rates year on year (Rosà et al., 2007), so it is likely that a similar variation exists in *Varroa* feeding and virus transmission.

This model effectively shows the stochastic effects of a small virus–vector–host system on population changes at the individual pupa level. To better reflect the effects of mite infestation on honey bees at the colony level this model could be expanded to account for several other factors. A model where additional *Varroa* are introduced would reflect the reproductive cycle of the mite (**Figure 5-1**). These additional *Varroa* would obtain viruses already amplified by the foundress mite, before leaving the cell and infecting another pupa, resulting in a more accurate reflection of the rate at which a particular variant was able to spread at the colony level. An expanded model could also include delays to transmission, reflecting the period of time that mites spend phoretic on adult workers between entering cells. The variable nature of the virus population observed in this model indicates that even in a scenario where one variant is higher than the others but does not reach clonality (for example the first 7–8 cycles of **Figure 5-8** where Virus A had a 50% bias initially), it could spread to dominate the virus population of the colony if the daughter mites feed at the right time on an infected pupa and subsequently transmit higher doses of that variant.

Despite its limitations, this model demonstrates that DWV variants can rapidly elevate and reach near-clonality in a vector-host system under certain set parameters. Further optimisation of the model could involve a refinement of the virus uptake during the 'bite limit' to include one variant which can replicate in the mite and another which is vectored in a non-propagative manner and a widening of the parameters such as virus elevation in the mite and the injection limit. The model also highlights the seemingly stochastic nature of virus dynamic changes in a system with bottlenecks and constraints applied, giving one possible explanation for some of the reported dynamic shifts in DWV populations. Many of the factors determining how DWV variants attain selection advantages over other viruses in the population remain unknown, but it is clear that bottlenecks influence variant selection and replication. This model can be used as a basic system to extrapolate information about DWV populations in *Varroa* and honey bees at a low cost.

6 Mapping DWV recombinants in co-infection experiments

'A solid foundation in (reverse) genetics is increasingly important for everyone.' - Anne Wojcicki, sort of.

This chapter was published as part of Gusachenko et al., (2021) – see ***Appendix 8-8***

6.1 Introduction and aim

Genetic recombination – the formation of hybrid or chimeric genomes – is an intrinsic consequence of the replication mechanism of all single-stranded positive-sense RNA viruses. Evidence currently suggests that it is an evolutionarily retained mechanism (as non-recombinogenic viruses can be engineered (Woodman *et al.*, 2016)) that compensates for the inherently error-prone polymerase these viruses possess, by 'rescuing' damaged genomes, and actively contributes to the generation of quasispecies diversity. There are numerous examples where recombination leads to the development of novel phenotypes, in both plant and animal viruses (Bentley and Evans, 2018). The polymerase switches templates during replication, probably in a negative-strand sequence-independent manner (Kirkegaard and Baltimore, 1986; Lowry *et al.*, 2014) to generate a hybrid genome (**Figure 6-1**), the viability of which then determines whether it is further propagated or disappears. Replication-competent recombinants may undergo or acquire further adaptive changes, the result being a chimeric genome with breakpoints marking the junctions between the two or more parental sequences from which it was derived. Knowing where the breakpoints are in viral recombination events can provide insight into the biochemical and mechanistic properties of viral infection (Rohayem, Munch and Rethwilm, 2005; Simon-Loriere *et al.*, 2009). The modular nature of the genomes of all single-stranded positive-sense RNA viruses contributes to the process of recombination, with the junctions usually located between the modules for the structural proteins, the non-structural proteins or the non-coding regions and the coding regions (Lukashev *et al.*, 2003; Lukashev, 2005; Lowry *et al.*, 2014; Bentley and Evans, 2018). During the process of recombination viral genomes can also form with large deletions or insertions in the sequence, when the polymerase jumps to a different position on the second viral strand (**Figure 6-1**).

Figure 6-1 - Schematic of recombination events. *Recombination between the purple and yellow strands can occur at a similarly positioned junction, or it can jump along the donor strand to create insertions (a larger viral genome) or deletions (a truncated viral genome). The red lines indicate where the two strands are recombining, and the blue arrow indicates the final sequence(s).*

Recombination between known DWV variants has been widely reported (Moore *et al.*, 2011; Ryabov *et al.*, 2014; Dalmon *et al.*, 2017; Natsopoulou *et al.*, 2017) and in some instances the recombinant form has been reported to predominate in the population over the strains it originated from (Ryabov *et al.*, 2014), although whether this is through a selection advantage or through stochastic events during transmission or replication is unknown. Moore *et al* (2011) identified two DWV recombinants: VDV-$1_{DVD,}$ comprising the DWV Type A 5'-UTR (untranslated region), a VDV-1 (Type B) structural gene and DWV Type A non-structural genes, and VDV-$1_{VVD,}$ comprising the 5'-UTR of VDV-1, a VDV-1 structural gene and DWV Type A non-structural genes. An observed correlation between high VDV-1_{DVD} titres in infected pupae and their associated mites suggested this recombinant may have been better adapted to transmission between *Varroa* and honey bees than its parental variants. McMahon *et al* (2016) found that Type A and Type B viral variants recombined readily *in vivo*, indicating non-independence of the two viruses among foraging honey bees. The use of reverse genetics-derived cDNAs of specific recombinant forms of DWV allows experimental comparison of these forms of the virus with parental sequences (Ryabov *et al.*, 2019; Gusachenko *et al.*, 2020a).

Recombination 'hotspots' refer to areas in the virus genome where recombination predominantly occurs. One such hotspot, observed by Ryabov *et al* (2014) and in related studies by Moore *et al.* (2011), is between the structural and non-structural proteins. Additionally, observations of recombination events occurring in the leader protein region (a 211-codon region towards the 5' end of the viral genome) between VDV-1 (Type B) and DWV Type A have been observed in recent analyses of samples in the USA (Ryabov *et al.*, 2017). The leader protein is the most variable region of the genome, and therefore an ideal region for measuring sequence diversity in a virus population, with a nucleotide identity match of only

~74% between DWV and VDV-1 (Dalmon *et al.*, 2017). Further 'hotspots' were independently discovered in the 5' end of the genome (Cornman, 2017; Natsopoulou *et al.*, 2017). All these identified recombination junctions occur between functional 'modules' of the virus genome (see **Figure 1-3** in Chapter 1).

The Type A/B recombinant viruses observed by Ryabov *et al* (2014) predominated in a colony over the parental Type A or Type B strains which were also present. This suggests that recombination may occur due to fitness selection or as the result of a bottleneck event (see section 5.1). Without evidence to the contrary it is likely that DWV (like other single-stranded positive-sense RNA viruses) recombines using a mechanism similar to that observed during poliovirus recombination (Kirkegaard and Baltimore, 1986). This is a biphasic process in which the initial recombinant product, which may have acquired sequence duplications during recombination, undergoes a subsequent resolution process involving the loss of sequence duplication and the optimisation of virus fitness (Lowry *et al.*, 2014). Multiple recombination events within a single DWV genome have been observed. For example, Dalmon *et al* (2017) reported a recombinant DWV genome which contained a Type B 5' untranslated region, a Type A leader protein sequence, the remaining structural proteins from Type B and the non-structural proteins from Type A, implying that at least three independent recombination events had occurred. It is notable that most of the recombinant strains reported in that study derived their structural proteins from Type B sequences (Dalmon *et al.*, 2017), indicating that this variant may have a selective advantage either during the recombination process *per se* or in the fitness of the resulting virus when compared with recombinants carrying Type A-like structural coding regions, a finding also reported by Ryabov *et al* (2014).

Recombination across the full DWV genome has been widely reported now, in part due to improvements in the tools used for analysis. Using next-generation sequencing tools, recombination was recently observed readily throughout the entire DWV genome using mixed populations of infectious clones (Ryabov *et al.*, 2019) and co-infected wild types stocks of Type A and B (McMahon *et al.*, 2016). Reciprocal recombinants between co-infecting strains have been identified throughout the viral genome, resulting in the widespread generation of novel variants. However, the previously reported 'hotspots' of recombination (Ryabov *et al.*, 2014) still produce the highest abundance of recombinant reads and most recombination events occur at the borders between the functional blocks forming the structural and non-structural coding regions of the genome, emphasising the modular nature of this genome (Ryabov *et al.*, 2019). The viability or fitness of the breadth of recombinants identified in whole-cell RNA deep sequencing studies remains to be determined. Comparison of this population with viruses transmitted experimentally or by *Varroa* would provide significant insights into the role of recombination in the tropism and pathogenesis of DWV. In addition, the use of reverse genetics allows the fitness or otherwise of individual recombinants to be readily determined after building or synthesising a suitable cDNA using standard molecular biology approaches.

This experiment was part of a broader study (Gusachenko *et al.*, 2021) to examine co-infection and superinfection exclusion between infectious clones of different DWV variants. As part of that study, co-infections of DWV variants were examined to determine whether recombination occurs at 'hotspots' in the genome or across the whole viral genome. This was observed using next-generation sequencing of DWV amplicons. We investigated whether we see a directional bias during recombination based on which variant is injected first or if the effect is stochastic. We also aimed to investigate the potential formation of defective RNAs during the replication process of the two variants during co-infection.

6.2 Methods

6.2.1 Sample preparation and injections

Pupae were extracted from comb from a low-level DWV and *Varroa* colony at the University of St Andrews apiary. New virus stocks of the infectious clones 'VVV' and 'VDD' were thawed on ice and prepared for injections. On the first day pupae were injected with 10^2 GE (total concentration in 5µl) of one variant and the following day pupae were injected with 10^6 GE of the second variant. This injection process was performed to test the superinfection exclusion principle for different DWV variants, where it has been proposed that a Type B variant (in this instance VVV) can effectively block the successful establishment of a Type A variant (VDD) (Mordecai *et al.*, 2015). Here we injected pupae with one variant, then the other, in both orientations (VVV→VDD and VDD→VVV) to test this theory. Once injected, pupae were placed in a 30°C incubator for 5 to 7 days and left to develop. On the fifth or seventh day the samples were snap-frozen in liquid nitrogen and placed at -80°C until RNA extractions were carried out.

RNA extractions and cDNA synthesis were carried out as per Materials and methods sections 2.2.1 and 2.2.2. cDNAs were amplified using the LongAmp PCR protocol as described in section 2.3.4 and purified using the Wizard PCR clean up kit (Promega, UK).

6.2.2 Recombination analysis

The purified PCR products were barcoded and sequenced using an Illumina Hi-seq. Sequences of 2 × 300 bp paired-end reads were processed at the University of St Andrews. Illumina fastq reads were extracted, trimmed and paired-end reads were matched using Geneious (V2019.1.3). All paired sequences were extracted as single fasta files for each injection sample.

Recombinants were analysed using ViReMa (Viral-Recombination Mapper, Version 0.15). For analysis a reference genome file was made using both injected DWV-variant sequences (VVV and VDD) with an A-tail terminal pad added to both sequences to maximise alignment sensitivity (Routh and Johnson, 2014). The reference file was indexed using Bowtie-build (Version 0.12.9) (Langmead *et al.*, 2009) and the Illumina reads were mapped to the reference file using the ViReMa recombinant mapping algorithm. ViReMa output generated matches between the two variants (e.g. VVV → VDD), matches within the same sequence (e.g. VVV → VVV) and reverse matches (e.g. VVVrev_strand → VDDrev_strand). The sense and anti-sense reads were merged for this analysis (e.g. recombination events between VVV → VDD and VVVrev_strand → VDD rev_strand). In total 16 different potential pairings of these reads were produced per sequence file.

To examine the locations of recombination junctions between variants only the sense and anti-sense reads between the two sequences were examined. These reported recombinant sequences were compiled as a text file and reorganised into a .csv file for data analysis using a custom perl script (**Appendix 4**). The .csv files were analysed using ggpubr v2.3 (Kassambara, 2020) in R Studio (Team, 2013).

To determine whether viral replication was producing intratypic insertions or deletions the reads discarded from the initial recombination analysis were recompiled and analysed separately using the same process as above. This involved all recombination events marked by

ViReMa as VVV > VVV, VVVrev_strand > VVVrev_strand, VDD > VDD and VDDrev_strand > VDDrev_strand.

6.2.3 Generation of synthetic library of Illumina reads

During analysis of the intratypic and intertypic recombination events, a significant number of large deletion events were observed (see section 6.3.3). To determine whether these ViReMa sequence events between variants were true deletion events or a product of an alignment or sequence analysis error, a library of 'Illumina-like' reads was generated using InSilicoSeq (Gourlé *et al.*, 2019). An in-frame section of the DWV open reading frame between 3666 and 5668 nucleotides of the VVV9-4 viral construct was deleted to form a truncated, in-frame viral genome. This was then used with InSilicoSeq to generate a database of 250 bp paired-end reads spanning the whole of the shortened genome. The following command was used to generate the reads, where *--genome* is the sequence of interest, *-n* is the number of generated reads required, *--model* is the sequencing platform and *--output* is the directory for the generated reads.

```
$ iss generate --genomes /genome_of_interest.fasta --model miseq -n 10K --output /genome_of-interest_miseq_reads
```

Two iterations of the database were generated, one yielding 10,000 reads and another yielding 100,000 reads, with 10,000 being a typical number of reads for a common recombinant in the sequence data set and 100,000 reads at the upper limit for a recombinant in the dataset. The synthetic databases were then merged with one of the Illumina datasets (a pupae injected with VVV and then VDD) and the combined datasets were analysed with ViReMa in the same way as the original datasets described above in order to determine if the spiked synthetic reads would be detected by ViReMa using the standard method used to analyse the injection dataset.

6.3 Results

6.3.1 Co-infection and qPCR analysis of VVV and VDD variants

Superinfection exclusion samples were analysed by qPCR to determine the viral titres of the different variants 5 and 7 days post-second injection. The qPCR analysis of the superinfection exclusion experiments indicated that the co-infected samples of VVV and VDD all had highly elevated DWV levels and that both variants reached $>10^9$ GE/µg RNA regardless of which variant was injected first (**Figure 6-2**). Samples infected with the recombinant variant (VVD) showed reduced accumulation of the second injected variant regardless of which way around the inoculations were performed, with VVD accumulating to titres as low as 10^4 GE/µg RNA in some samples (**Figure 6-2**, purple points).

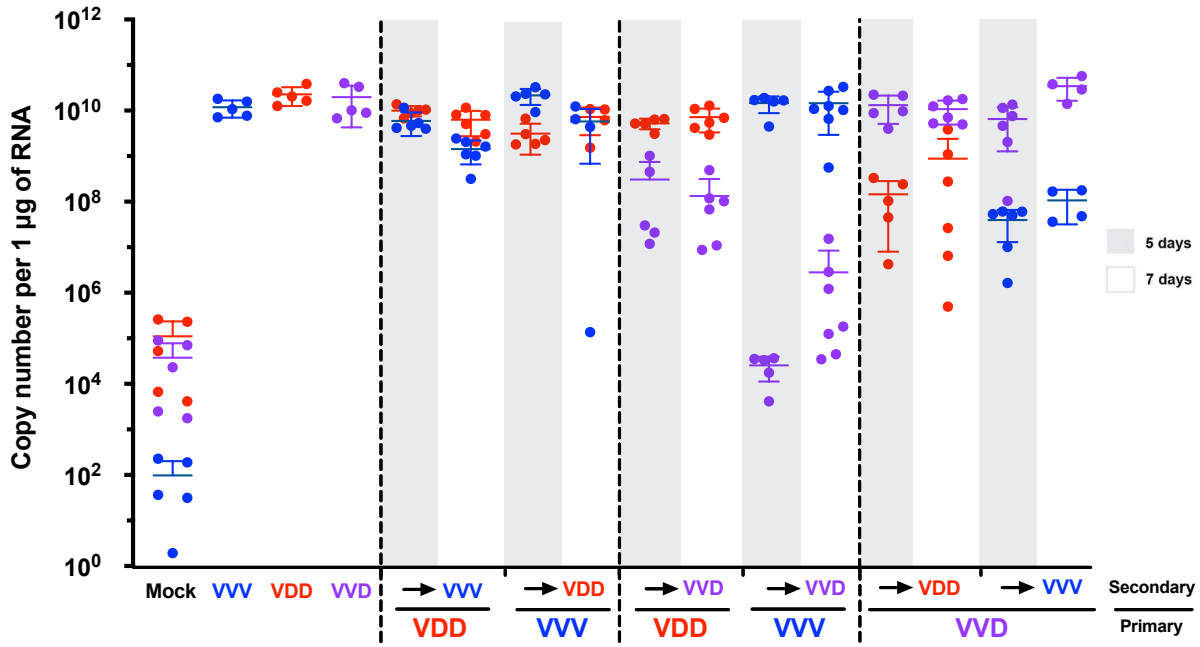


Figure 6-2 - Analysis of DWV accumulation of super-infecting virus variants by qPCR. VDD, VVD and VVV DWV accumulation was measured in honey bee pupae after superinfection with one of the virus variants and analysed after 5 (grey bars) and 7 days (no bars) incubation. Primer sets specifically targeting the "Type A" and "Type B" RNA-dependent RNA polymerase (RdRp) and a region in the structural proteins (StPr) were used to detect the DDD and VVV polymerase and structural protein encoding sequences respectively.

6.3.2 Analysis of recombination between variants

Illumina paired reads were generated from PCR amplicons of a 10 Kb fragment of the DWV genome targeting samples initially infected with VDD and challenged with VVV and vice versa (samples "VVV→VDD" and "VDD→VVV" at 5 or 7 days after superinfection in **Figure 6-2**). Illumina reads were analysed using ViReMa as per section 6.2.3 and the recombination locations observed were visualised using bubble plots generated in R. Reads were filtered with a cut-off of >3 reads per recombination junction before being plotted. Recombination was observed across the full DWV genome, spanning structural and non-structural protein sequences in all samples analysed (example – **Figure 6-3**, all plots – **Appendix 5**). **Table 6-1** shows the summary data for the total Illumina reads generated, the number of viral recombinant events observed, and the intertypic recombination events observed after the cut-off threshold was applied. Viral recombination comprised 2.8% to 5.8% of the total Illumina reads and intertypic recombination was observed in 1–1.9% of all reads.

Table 6-1 - Summary of ViReMa recombination output. *The data shows the total Illumina reads, the viral recombination events and the intertypic events. The percentage of the total reads for viral recombination and intertypic are shown in brackets. The unique mapped junctions are shown for all intertypic recombination events. 5d/7d refers to pupa incubation period as per* ***Figure 6-2****. V and D in the first column refer to VVV and VDD respectively.*

Sample	Total Illumina reads	Viral recombination events	Intertypic recombinants			Unique junctions		
			VVV> VDD	VDD> VVV	Total	VVV> VDD	VDD> VVV	Total
V>D 5d #2	278944	16121 (5.8%)	2501	2871	5372 (1.9%)	259	254	513
V>D 7d # 2	329434	16897 (5.1%)	2836	3068	5904 (1.8%)	263	283	546
V>D 7d # 3	182202	8334 (4.6%)	1194	1339	2533 (1.4%)	160	172	332
D>V 5d # 1	261208	11020 (4.2%)	1900	2002	3902 (1.5%)	216	223	439
D>V 5d # 3	372898	10604 (2.8%)	1639	1829	3468 (0.9%)	202	206	408
D>V 5d # 4	231382	7741 (3.3%)	1107	1191	2298 (1%)	159	151	310
D>V 7d # 1	245568	11610 (4.7%)	1711	2001	3712 (1.5%)	203	216	419
D>V 7d # 2	161506	7427 (4.6%)	1080	1274	2354 (1.5%)	150	157	307
D>V 7d # 3	218282	11617 (5.3%)	1901	2063	3964 (1.8%)	202	215	417
D>V 7d # 4	180096	7123 (4%)	1110	1136	2246 (1.2%)	154	161	315

Hotspots of recombination were observed throughout the genome, indicated by the larger bubble sizes in **Figure 6-3**, including some previously reported known regions of recombination (Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). Directionally-specific recombination was occurring in some samples at different positions across the genome, for example at nucleotide position 4731 in the genome 271 reads were observed for a recombination event from VDD5' → VVV3', but 0 reads were observed at this position for VVV5' → VDD3' (**Figure 6-4**). The top 'hotspot 'matches were examined from all samples (**Table 6-2**), with the most abundant selected for each (all with >200 reads shown). Four recombination hotspots observed in both directions are shown by the black lines in **Figure 6-4**, and hotspots in which the 5' was only ever derived from one variant or the other are shown in blue (VDD 5' → VVV 3') and red (VVV 5' → VDD 3'). These were also observed regardless of which viral variant was injected first during experimental set-up. These recombination points were predominantly observed in the non-structural proteins (10/14 of the top recombination hotspots).

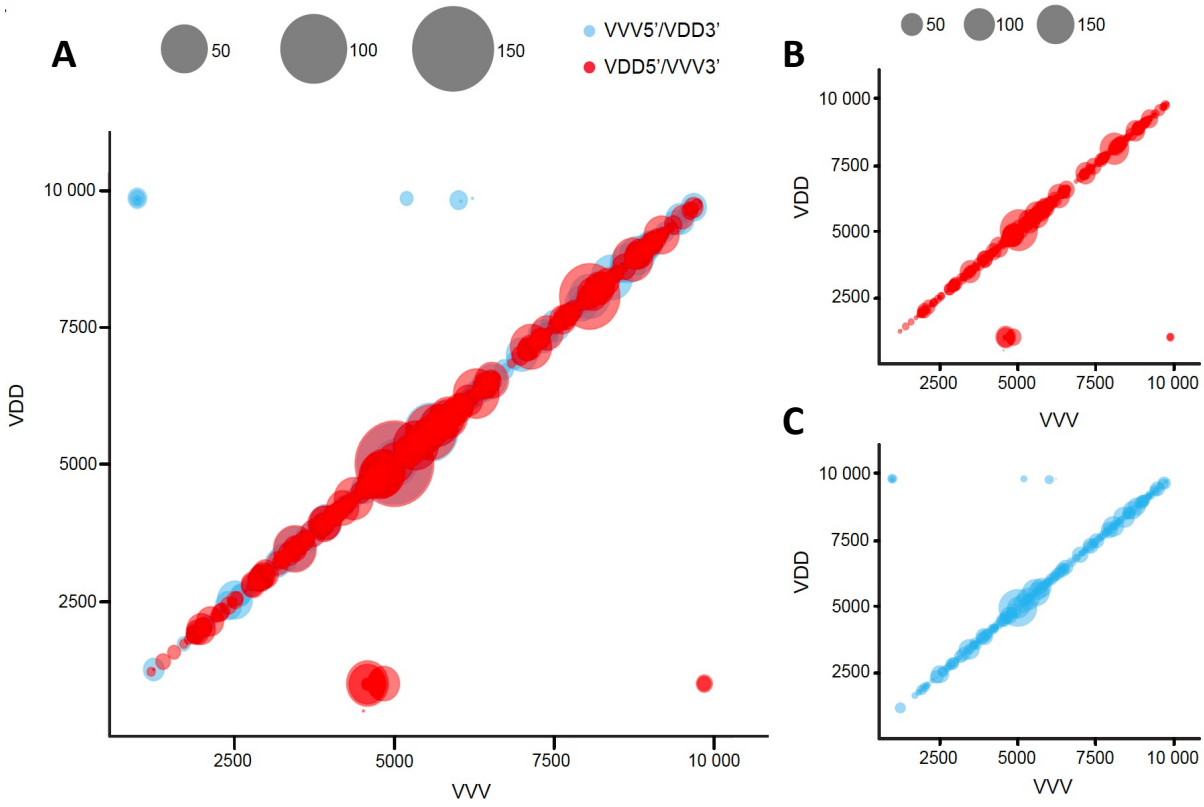


Figure 6-3 - Mapped recombination events observed during superinfection exclusion between VVV and VDD viral variants in a single honey bee pupa. *Each mapped recombination event occurring along the full DWV genome, with VDD shown on the Y-axis and VVV shown on the X-axis.* ***A*** *– all recombination events,* ***B*** *- recombination events VDD 5' → VVV 3'and* ***C*** *- recombination events VVV 5' → VDD 3'. Each bubble represents a unique recombination event and bubble size is determined by the number of mapped and aligned reads obtained using ViReMa analysis. In this example the honey bee was injected with the VVV template prior to the VDD template. The colour of the bubbles indicates the direction of recombination, with blue representing VVV as the acceptor sequence and red representing VDD as the acceptor sequence. Only one sample is shown for simplicity, all other co-infection bubble plots are shown in Appendix 5.*

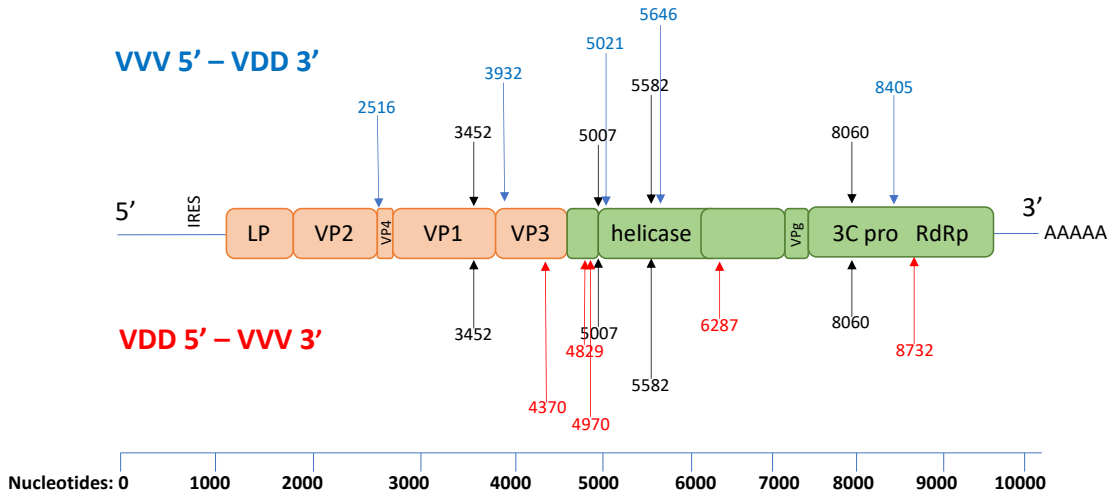


Figure 6-4 - Recombination 'hotspot' junctions shown at the point of recombination on the genome between VVV and VDD. *The points shown in black occurred with similar frequency in both directions. Those shown in blue occurred predominantly with VVV as 5'-acceptor and VDD as 3'-donor and those in red with VDD as 5'-acceptor and VVV as 3'-donor respectively.*

Table 6-2 - Top recombination hotspots from all samples. *Most common recombination junctions observed between VVV and VDD DWV genomes when recombining in both directions, with VVV as acceptor and VDD as donor predominantly, or VDD as acceptor and VVV as donor only. All events with >200 reads shown.*

Recombinant formed	Recombination point	Total mapped reads
	5' nucleotide	
Both directions	3452	600
	5007	1930
	5582	1179
	8060	789
Predominantly VVV5'/VDD3'	2516	246
	3932	257
	5021	326
	5646	266
	8405	242
Only VDD5'/VVV3'	4371	271
	4830	365
	4971	273
	6288	362
	8733	221

Although the percentage of mapped reads corresponding to recombination junctions varied in individual pupae from 1 to 1.9% of all mapped reads (**Figure 6-5**), in all cases approximately equal proportions of recombinants were detected with VVV or VDD as the 5' acceptor partner (terminology assumes that recombination occurs during negative strand synthesis (Kirkegaard and Baltimore, 1986; Lowry *et al.*, 2014)). A number of recombinants were observed throughout in which the junction was identical regardless of whether the 5' was derived from VVV or VDD, but in comparison some junctions were observed in which the 5' was only ever derived from one variant or the other. The implications of this for recombination formation remain unclear. These studies demonstrate that although superinfecting viruses recombine readily with an established virus population, the overall level of recombinants within the resulting virus population is probably ~1–2% of the total, similar to findings observed during poliovirus recombination (Alnaji *et al.*, 2020).

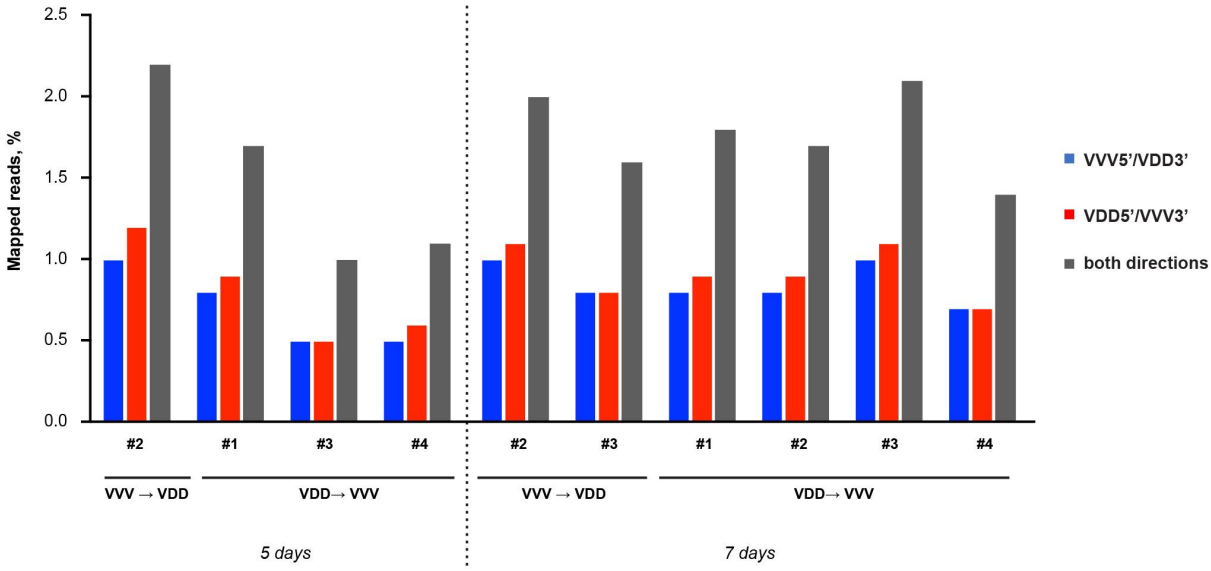


Figure 6-5 - Percentage of recombination reads in each sample sequenced. *Each bar shows the percentage of recombination reads for each direction in each sample analysed using ViReMa. Blue bars represent VVV 5'/VDD 3' and red bars represent VDD 5'/VVV 3'. The percentage was calculated as a part of the total number of Illumina reads obtained for each sample post-trimming.*

6.3.3 Truncated RNA sequences detected during recombination events

Analysis of the recombination junctions revealed the apparent formation of truncated RNA sequences caused by large deletions during viral replication, possibly forming defective DWV genomes during the process of viral replication. During DWV replication multiple reads aligned to produce fragments where the viral polymerase had inaccurately switched strand during replication, causing large deletions of genomic sequence between the VDD 5' and the VVV genome (**Figure 6-3)**, forming truncated and potentially non-viable viruses. This reciprocal recombination occurred across multiple samples but was observed at higher frequency during the formation of VDD 5' to VVV 3' recombinants, typically skipping large regions of the genome which encode the structural proteins. Additionally, intertypic formation of these RNAs was observed frequently in all samples for both variants. **Figure 6-6** shows all the observed truncated sequences that occurred between the two different viral variants (352 reads from 49 unique events). A large number of deletions occurred from ~5000 bp of VDD to

the 5' end of VVV, forming a truncated sequence encoding none of the structural proteins and therefore not likely to be viable. **Figure 6-7** shows all the intratypic deletions and insertions observed, with VVV → VDD in **Figure 6-7A** and VDD → VVV in **Figure 6-7B**. The insertions and deletions occur across the genome, but clusters are observed around the region between the structural and non-structural proteins (~5000 bp) and in the 3' non-transcribed region (>10000 bp). Most VDD insertions and deletions were short and mapped along the central line of **Figure 6-7**, but many of the VVV insertions and deletions involved large jumps along the genome, as shown by the large blue bubbles in both plots. Deletions and insertions occurred most frequently in the template which was added first in the co-infections (**Table 6-3**), so when VVV was injected first a great number of VVV deletions and insertions were observed and vice versa when VDD was added first. In all samples, insertions occurred more frequently in terms of total reads (49.3% and 48% compared to 45.7% and 46.7% for deletions).

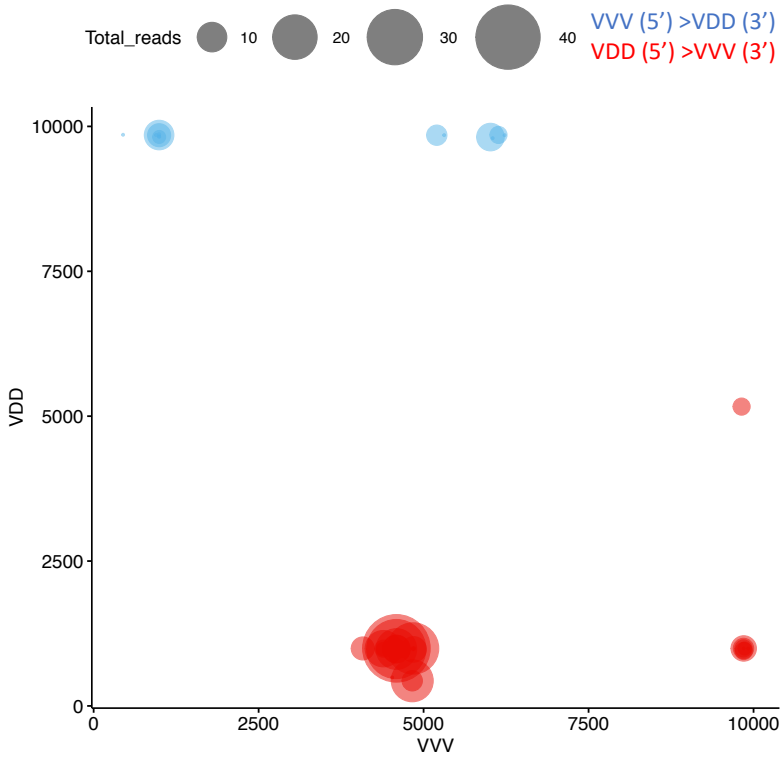


Figure 6-6 - Potential intertypic deletant RNAs observed between VVV and VDD. *Data shows all dRNAs detected from the samples processed by Illumina and ViReMa analysis (10 samples). Sequences which have jumped from VVV 5' to VDD 3' are shown in blue and samples which have jumped from VDD 5' to VVV 3' are shown in red.*

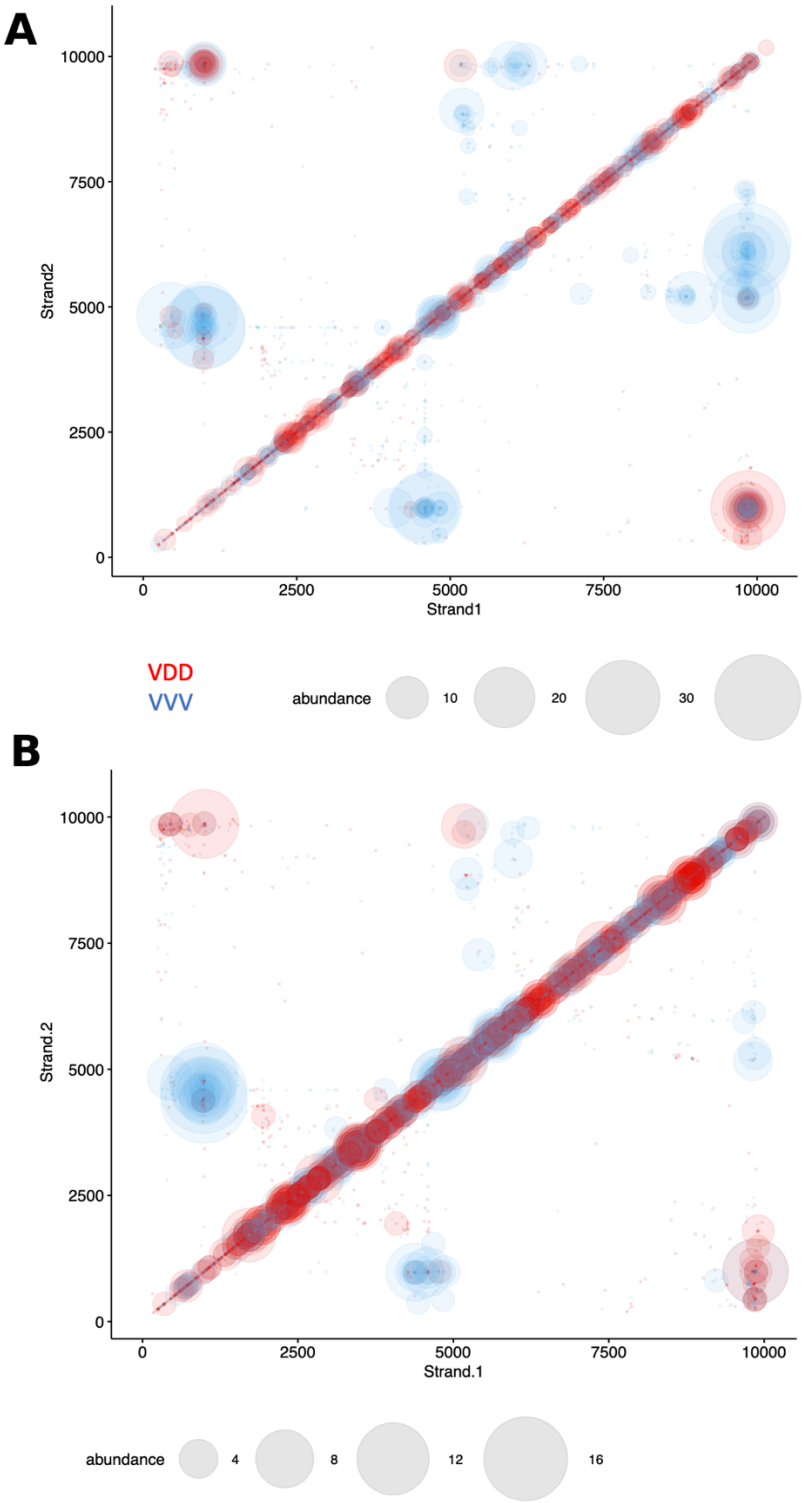


Figure 6-7 - Potential intratypic defective RNA formation during virus co-infections. ***A*** *- Recombination bubble plot showing all intratypic insertions and deletions for samples infected initially with VVV and subsequently VDD (3 samples).* ***B*** *- Recombination bubble plot showing all intratypic insertions and deletions in all samples infected with VDD and subsequently VVV (7 samples). All intertypic recombinant sequences are removed from both data sets so the plots show only the intratypic recombination and deletion events. The blue bubbles are VVV intrartypic sequences and the red bubbles are VDD intratypic sequences.*

Table 6-3 - Summary of unique recombination events and total reads for intratypic insertion and deletion events for all samples. *The samples are split for VVV → VDD and VDD → VVV and only show the intratypic recombination events reported.*

	VVV → VDD (3 samples)			VDD → VVV (7 samples)		
	Unique events	Total reads	% of reads	Unique events	Total reads	% of reads
All	4793	6192	100	8658	10240	100
Deletions	2233	2830	45.7	3975	4780	46.7
VVV deletions	1146	1520	24.5	1825	2194	21.4
VDD deletions	1087	1310	21.2	2150	2586	25.3
Insertions	2231	3050	49.3	4203	4913	48
VVV insertions	1243	1691	27.3	1991	2309	22.5
VDD insertions	1068	1359	22	2212	2604	25.5
zeros	249	312	5	442	509	5.3

6.3.4 Comparison of a synthetic library with raw Illumina reads

Some of the truncated recombinant RNA sequences (**Figure 6-6**) were extracted from the dataset and aligned against the RG sequences in Snapgene. Alignment of these deletant genomic sequences to the parental genomes revealed none of the selected truncated sequences were in frame and therefore they were unable to form viable replicating viruses (data not shown), however whether their presence during viral replication benefits or inhibits viral replication remains unclear.

An artificial sequence dataset was generated to determine whether the large deletions observed during viral co-infections, and subsequently truncated RNA sequences, were artefacts of the sequence analysis tools used or real sequences generated during viral replication. A large region of the DWV genome for one of the sequences (VDD) was removed (5' end--3666xxxxxxxxxx5668--3' end). Artificial reads of 10,000 and 100,000 Illumina paired-end 300bp reads were generated using InSilicoSeq (Gourlé *et al.*, 2019) and inserted into the original sequence files. These spiked samples were then analysed by ViReMa in the same manner as the original samples. The aligned and sorted reads were indexed against a reference file and visualised in IGV (**Figure 6-8**). In panel A the 10,000 artificial reads generate a sequence with a deletion in the expected region of the genome and can be observed aligning to the reference genome correctly. When 100,000 spiked reads were inserted into the dataset (panel B) the ratio of 'true' reads to artificial reads was much closer and this was reflected by the presence of several aligned deletion events against the reference genome for the fake reads. This confirmed that the process of alignment and analysis using ViReMa was valid for this dataset and that the truncated RNA sequences observed are not a product of analysis.

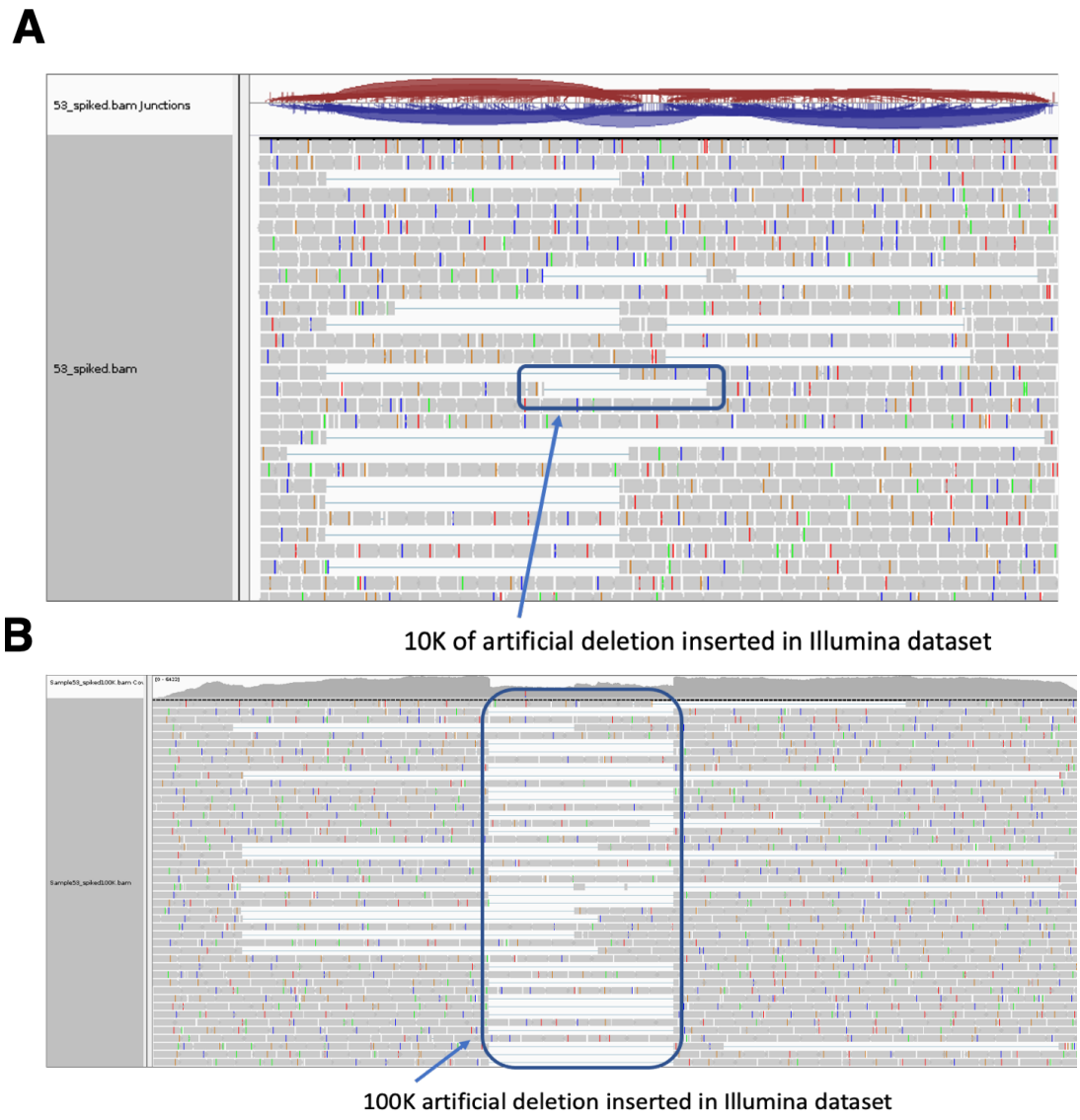


Figure 6-8 - IGV windows of aligned Illumina reads with 10,000 and 100,000 spiked samples. *Bam files of Illumina reads, aligned with ViReMa and converted from a Sam file to bam file using Samtools, are visualised against the reference genome. The large deletion artificially inserted into the genome is visible amongst the aligned reads when 10,000 (**A**) and 100,000 (**B**) reads are generated and inserted into the original data set.*

6.4 Discussion

To examine the frequency and extent of recombination during co-infection experiments, samples of VVV and VDD were selected for sequencing. Evidence of widespread recombination was clear in all samples (**Appendix 5**) and junctions were observed between the variants across the full viral genome in both directions and regardless of which viral variant was the primary infectant (**Figure 6-3**). 'Hotspots' of recombination with significantly higher read counts were also detected across the genome, like those previously reported (Moore *et al.*, 2011; Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). Most notably, junctions were observed between the regions encoding the structural and non-structural proteins, possibly indicating a fitness advantage for viruses with a particular set of capsid proteins. If Type A variants readily recombine to form a recombinant with a Type B capsid sequence, such as the VVD variant

observed by Ryabov *et al.* (2014), it may indicate a selection advantage to this configuration, particularly if such a variant is also able to replicate in *Varroa*.

The modular nature of the DWV genome (see **Figure 1-3**) might mean that 'hotspots' are more likely to occur during viral recombination and the hotspots observed in this study typically occurred at the ends of these 'modules'. As the variants used in this study were built using reverse genetics and each variant was built in a modular fashion, with blocks of the genome cut or inserted during design (for example VVD to VDD involved the removal of a Type B non-structural region and insertion of a Type A non-structural sequence), it is possible that the hotspots observed during recombination here and in other studies using similar methods (Ryabov *et al.*, 2019) are an artefact of this technique. However, recombination can still occur in the un-transcribed 5' end of the genome and throughout the region encoding the structural proteins (Dalmon *et al.*, 2017) and extensive recombination events were also observed in this study in the region encoding the non-structural proteins (**Figure 6-3 and 6-4**). This indicates that recombination between DWV variants is not limited to the capsid proteins or the modular junctions between parts of the genome.

As well as the clear evidence of precise recombination between the variants, imprecise recombinants were detected intratypically and intertypically (**Figure 6-6 and 6-7**). Intratypic recombination, recombination between two viral genomes of the same variant, produced a large number of insertions and deletions. These also appear to form in a modular fashion (**Figure 6-7**), although clouds of low-level events were observed around the more abundant forms, suggesting a degree of stochasticity in their formation. Whether these variants form functional viral particles or are a snapshot in time during viral replication is unclear. It is also possible these variants would undergo a process of resolution whereby the inserted repeated region is deleted from the genome during further rounds of replication (Bentley *et al.*, 2021).

Truncated RNAs, often referred to as deletants (see **Figure 6-1**), were observed in most samples, in both intertypic and intratypic recombination events. These truncated forms of the viral genome typically contain the 5' and 3' ends of the genome and in other viral infections are known to form in the presence of high viral titres when the polymerase drops off the template and reattaches further downstream (Wilhelm *et al.*, 2014; Salas-Benito and De Nova-Ocampo, 2015). The presence of large deletion events could result in the formation of defective RNAs, which may play a role in viral infection. The role of defective RNAs in viral persistence in natural infections is not well understood, but some observations *in vitro* suggest that alternating cycles of defective interfering particle formation and viral replication can promote virus persistence (Salas-Benito and De Nova-Ocampo, 2015). All co-infected samples tested here reached very high viral levels and the presence of potential defective RNA could be aiding the infectivity of these clonal variants, even though the vast majority of the screened variants were not in frame.

It is possible that these truncated RNAs serve no function in viral replication or the dynamics of virus competition. However, it is also possible that they were a product of assembly or sequencing error during the generation of recombination reads. Therefore, the presence of these truncated RNA sequences as a potential sequencing artefact was investigated by generating a library of artificial reads and inserting them in to one of our Illumina datasets, before aligning and analysing the samples with ViReMa and searching for the defective RNA they would generate in the final dataset. In all iterations examined we found the inserted reads (**Figure 6-8**), indicating that the other defective RNA sequences observed were not a product of sequencing error or alignment issues during ViReMa analysis. However, it remains a

possibility that these sequences are a result of replication error or a biological variant only observed at very low levels in the population. To investigate this further, total RNA samples could be screened by northern blot or one of the fragments could be synthesized and used to replicate infection in a developing pupa.

Further examination of these potential defective RNAs revealed that many of them were out of frame and therefore could not form viable replicating viruses, but most likely existed at a snapshot in time of this virus population. If any of these defective RNAs were found to be in frame they could be of potential use as a molecular biology tool, for example a sub-genomic replicon, like that observed in poliovirus (Barclay *et al.*, 1998), or a reporter with a GFP-encoding insert could be designed and examined in a similar method to that utilised for the full DWV genome recently (Gusachenko *et al.*, 2020a; Ryabov *et al.*, 2020).

This work has shown that recombination is a random process which occurs across the full DWV genome, but the modular nature of the DWV genome results in high-prevalence 'hotspots'. As with other positive-strand RNA viruses, DWV recombination appears to result in the formation of some variants with large insertions or deletions which may be resolved through future viral replication and may play a role in the virus' ability to sustain infection in the host. Understanding the role these variants play may help elucidate how DWV is able to rapidly elevate in the host and overcome the host immune response.

7 Discussion

"The Vermin only teaze and pinch
Their Foes superior by an Inch.
So, Nat'ralists observe, a Flea
Hath smaller Fleas that on him prey,
And these have smaller yet to bite 'em,
And so proceed ad infinitum:
Thus ev'ry ~~Poet~~ PhD, in his Kind
Is bit by him that comes behind."

Johnathan Swift, On ~~Poetry~~ a PhD: A Rhapsody (1733).

7.1 Changes in colony health and viral diversity in the presence and absence of Varroa

Honey bees are key pollinators of wild flowers and a range of agriculturally important crop species. They face a range of threats to their survival, including increased mono-culture crop growth resulting in habitat loss, increased pesticide use and susceptibility to emerging diseases. The most notable disease threat comes from the combination of Deformed wing virus and the parasite which can vector the virus, *Varroa destructor*. Combined, they are responsible for large overwinter colony losses of honey bees every year, and thus advancing our understanding of the dynamics of the host-vector-virus interaction is key to improving the health of honey bees.

The introduction of *Varroa destructor* to Western (European) honey bee (*Apis mellifera*) colonies irrevocably altered the dynamic between the host and a range of pathogenic viruses, in particular Deformed wing virus (DWV). By introducing an alternative transmission route, potentially allowing the virus to bypass host immune barriers, the mite shifted the balance between the virus and the host. The study outlined in Chapter 4 aimed to investigate coordination of miticide treatments between neighbouring apiaries in an isolated environment as a means of improving honey bee health, measuring changes to the DWV population and mite abundance over time. In tandem, two smaller-scale experiments were carried out (see Chapter 3) which investigated the impacts of mite infestation and their removal using practical beekeeping techniques. The studies highlighted the positive impact on colony health of miticide treatments combined with these techniques, and the negative impacts of not treating colonies appropriately in infested apiaries.

Using a combination of appropriate miticide treatments and a 'shook swarm' colony management technique it was possible to rapidly reduce *Varroa* in highly infested colonies. The sealed brood of a honey bee colony can contain as much as ~80% of the mites depending on the time of year. Immediately after the shook swarm, the colonies were broodless, so the remaining ~20% of mites were exposed to the miticide treatments applied. This allowed the miticide to effectively remove the majority of the mites in the colonies. This change was also reflected in the DWV population, which decreased from high titres (~10^{10} GE/µg RNA) to a healthy level (<10^{6} GE/µg RNA) by the first brood cycle and remained low throughout all colonies across both years of the experiment. A healthy level of DWV in this instance is based on experimental data which showed the threshold for symptomatic infection to be 10^{6} GE/µg RNA by oral transmission to larvae (Gusachenko *et al.*, 2020a), and routine screening of the

St Andrews apiary colonies, which showed levels of 10^3–10^6 GE/µg RNA and no symptomatic infections (data not shown).

Previous experimental work examining the impacts of mite removal on DWV titres did not combine treatments with beekeeping techniques or study DWV diversity changes (Locke *et al.*, 2017). Although the treatment application reduced DWV in that study, the viral levels remained $>10^6$ copies per bee and began to increase again before the end of the experimental time course, most likely due to mite-reinfestation from neighbouring colonies in the apiary. That study also did not examine the impacts of the miticide treatments on the diversity of the virus population, screening the DWV by qPCR only, perhaps missing shifts from clonal to mixed populations over time.

Although some beekeepers consider the method applied in this study to be harsh on the colony, and it can only be applied to relatively strong colonies and during conditions which allow comb building, it was very effective across all 6 colonies over the 2 years (**Chapter 3, Figure 3-18**). The level of mites was expected to be very high initially and low-level mites may have persisted after the shook swarm. Therefore, Apivar, a miticide which is placed in the colony for an extended period, was preferred over treatments like oxalic acid sublimation, which is applied in one-off doses. However, the mite levels dropped rapidly in all colonies in both years and remained low throughout, so it is possible that a shorter treatment such as oxalic acid may have worked effectively too.

A screen of the DWV population before treatments, in the first brood cycle after treatments and a further sample set from the end of the season revealed colony-to-colony differences in dominant variants and diversity. The mites killed by the Apivar treatment were screened in pools and revealed mixed populations of Type A-like and Type B-like DWV variants in colonies 2 and 3 (**Figure 3-22**). These mites would have fed on the adult workers which would then feed the first generation of bees reared post-treatment. This resulted in mixed populations of A and B variants in colony 2 at time point 2, but by the end of the season (TP5 – September) the workers all produced near-clonal virus populations. However, the individual workers often contained different variants and if these samples had been analysed as a pool they would have produced a low viral titre population of mixed variants on a colony level. This would have masked the true virus dynamics occurring on an individual worker level, which highlights the importance of analysing DWV populations at the individual level.

In the other section of Chapter 3 (see section 3.2), where colonies with low DWV titres and low mite levels were placed near highly mite-infested colonies, virus levels rapidly increased in two of the three years, reflecting high mite acquisition in the introduced colonies from drifting or robbing workers. Despite the rapid change in virus levels, similar results were observed in terms of virus diversity at an individual worker level in this part of the study. As shown in **Figure 3-15,** it was observed that some individuals with low level infections had clonal virus populations, whilst others which had very high titres ($>10^9$ GE/µg RNA) contained mixed virus populations. Previous reports of the almost binary differences between DWV populations in the presence and absence of *Varroa* have either been based on pooled sampling (Martin *et al.*, 2012) or on relatively small sample sizes (Ryabov *et al.*, 2014). More recently, mixed populations of DWV have been observed at higher titres when analysing individuals (Annoscia *et al.*, 2019; Ryabov *et al.*, 2019), indicating that variants can coexist at high titres. The results observed in this study, and those in recent publications, highlight that DWV probably exists as a dynamic group of variants which are ever-changing as they mutate and recombine, and bees containing certain variants die from other causes or because of high

infections. Despite this, many reports of DWV focus on Type A vs Type B infections. Using near-clonal virus strains recovered using reverse genetics, co-infections of these variants have found both are capable of replicating to very high titres simultaneously, despite competing for resources in the host (Gusachenko *et al.*, 2021). Therefore, it may be posited that observations of clonality at high titres could be a result of stochastic differences induced by virus bottlenecks or selection pressures, such as certain variants of the virus replicating in the mite (Posada-Florez *et al.*, 2019; Gisder and Genersch, 2020; Gusachenko *et al.*, 2020a), or of a lack of adequate sensitivity during sequence analysis, particularly using pooled data or inferences about virus populations from consensus sequences.

The techniques used to prepare samples and analyse data will also influence the conclusions, particularly as RNA viruses can rapidly degrade if not stored correctly. Poor tissue extraction techniques or unsuitable molecular biology techniques can result in low-level variants not being detected by standard methods. The use of next-generation sequencing of samples can provide significant improvements in sensitivity and the new suite of tools for analysing virus populations means a great deal more information can now be obtained (Posada-Cespedes, Seifert and Beerenwinkel, 2017). In this study all samples were analysed as amplicons, rather than using RNAseq analysis, as it was believed that by following good RNA handling techniques the low-level variants could be amplified to generate sufficient coverage of DWV with enhanced specificity. The drawback of this is the inherent bias that PCR amplification will introduce to a sample set, based on better primer specificity for particular targets and more divergent variants perhaps being omitted.

The tool ShoRAH (short read alignment to haplotype) (Zagordi *et al.*, 2011; McElroy *et al.*, 2013) was used to accurately calculate not just the diversity within a sample population, but the percentage of the population each variant comprised. In instances where a variant constitutes the majority of the virus population standard molecular biology techniques or consensus sequence calling would define it as the only variant present, but the presence of low-level variants can have a significant influence on virus infections, sometimes causing drug resistance and later becoming the dominant variant (Metzner *et al.*, 2009). In the studies presented in this thesis, low-level (<10%) variants in early time points became the dominant variant at later time points (Chapter 3 and Chapter 4).

Tools like ShoRAH are undoubtedly useful for determining virus population diversity and abundance, but when populations contain recombinant variants of the dominant viruses it becomes difficult to decipher diversity when analysis consists of small windows (150 bp) in the genome, which cannot distinguish between a particular variant and a recombinant thereof. If a recombination 'hotspot' is already known prior to analysis, such as the region between the structural and non-structural proteins in the DWV genome (Moore *et al.*, 2011; Ryabov *et al.*, 2014; Dalmon *et al.*, 2017) then a window spanning this area can be selected and used to determine the population. But if the mixed samples contain potential recombinants forming at unknown points in the genome, this form of sequence analysis is liable to miss variants in the population, particularly as they make up a low percentage of the population (1.5%) (Chapter 6). Initial reports of recombination between DWV variants focused on the hotspots, such as that between the structural and non-structural proteins, possibly as a result of the analysis used (Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). However, recombination has now been observed across the whole genome to some degree, with some areas showing a higher frequency (Chapter 6, (Ryabov *et al.*, 2019). Multiple windows of analysis with a tool such as ShoRAH would be needed to comprehend the true diversity of the DWV population. Alternatively, to fully determine the presence of recombinants in the sample pool, a combination of short-read

sequencing and long-read sequencing could be used. Long reads could be sequenced from the 10 kb amplicons generated in this study (see section 2.3.4) using Nanopore or PacBio sequencing and used to scaffold Illumina sequence data, therefore generating sequence depth and revealing potential novel recombinant forms of DWV. At the start of this project, a Nanopore approach was tested but at the time the SNP error rate from sequencing was approximately equal to the SNP differences observed between common DWV variants, so it was deemed unsuitable (data not shown). Improvements in the technology over the last five years mean this may now be a viable option and could be used to elucidate some of the dynamic changes observed in the virus populations in this study.

7.2 Coordinated treatments on the Isle of Arran

The study of coordinated treatments on the Isle of Arran used pooled honey bee samples, rather than individual workers like Chapter 3, due to the large sample size and prohibitive costs of analysing individuals. As the colonies all belonged to amateur beekeepers, disruption needed to be kept to a minimum and for this reason, as well as the number of colonies included in the study, it was not practical to sample brood from every colony. Therefore, the bees sampled on the island were all adults found on the comb in each hive. This is a limitation as the bees carrying the largest disease burden are likely to emerge with deformed wings and be ejected from the colony by healthy workers, therefore those bees which are sampled will always present a slightly misleading picture of the overall health of the colony. An improvement to this study would have been to select colonies of interest across the island, for example one per Site, and sample a small patch of brood from each to compare the DWV titres from a limited but representative number of colonies. Additionally, different variants from those observed may have been present in those bees which produced symptomatic infections; however this seems quite unlikely given the general lack of virus diversity observed across the island in the first year and no evidence at present to suggest one variant is more likely to cause symptomatic infections (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a).

There were many external variables in the study, which made controlling and understanding the changes on a colony-to-colony basis difficult. Swarms were frequently caught in bait hives, but it is equally likely that other swarms were not caught and subsequently established as feral colonies at unknown locations near the apiaries on the island. This creates an opportunity for reinfestation by drifting and robbing from the untreated swarms to the treated colonies. A possible way to measure this would have been to mark foragers and try to hunt for swarms by tracking the flight paths of these workers (Visscher and Seeley, 1989). Miticide resistance may also have been a factor in the sustained *Varroa* presence, as all colonies on the island received the same treatment type for three years and the beekeepers had already been using Apivar prior to the study, albeit not in a coordinated manner. Resistance to Amitraz, the active ingredient in Apivar, is rarely reported, but recent research has indicated it may be becoming an issue (Rinkevich, 2020). If resistance has occurred it may explain the persistently high mite levels in some colonies, such as at Site 5. The beekeepers were also given treatments to apply in mid-winter, by Api-bioxal dribbling, a method that many of them had not used before. Due to logistical issues, the beekeepers applied these treatments themselves, but it may have benefitted the study if we had done so. Better still, we could have applied an oxalic acid sublimation as the mid-winter treatment, a method that more effectively disseminates the treatment through the colony. The beekeepers' lack of experience and an inability to report the outcomes from this treatment, will have introduced inconsistencies across the island. Additionally, the timing of the treatment is critical: if treatments are applied when the queen has begun laying again and brood cells are capped, mites will inevitably escape. On Arran, where the climate is temperate,

a prolonged period of cold weather (<10°C) was required to ensure a broodless period and this was when beekeepers were instructed to apply the miticide. As this was left to the judgement of the individual beekeepers, there will undoubtably have been some variation in application success.

From the September mite drop data, there was clear evidence for site-to-site variation in mite numbers after treatments (**Chapter 4, Figure 4-3**). Site 1, in the north of the island, reported no mites by the third year of the study (a very low level likely persisted), whilst other sites continued to have high mite abundances throughout the study. Site 2, isolated on the west coast of the island, had high colony losses and mite levels early in the study, possibly due to swarms or feral colonies in close proximity. Several colonies died and the isolation of this site may play a factor in that. Genetic diversity in a honey bee colony improves foraging ability and general health (Mattila and Seeley, 2007) so if the queens of the Site 2 colonies were only able to breed with drones from the same site it may have created a genetic bottleneck, which then impacted colony health. Despite the significant differences between sites, the total number of mites collected post-treatment decreased significantly from the first year to the third (58% fewer mites), despite a 50% increase in the total number of colonies being managed on the island (**Figure 4-5, Table 4-1**), indicating that the coordinated treatments were having a positive impact on the mite infestations. The beekeepers on the island had a range of skill and experience levels, including some who became beekeepers during the study, and this will have had an influence on the management of the colonies, including disease awareness and swarm management. Measuring the impacts of coordination over a longer period of time may have shown further reductions in mite abundance across the island.

7.3 Virus diversity and DWV Type A vs Type B

As mentioned above, the samples of worker bees on Arran were analysed as pools to examine virus diversity. As shown in Chapter 3, individual workers within the same colony with similar viral loads can contain clonal or mixed virus populations and these differences are lost when samples are pooled. However, practical considerations made it difficult to examine the Arran samples as individuals, and the initial individual analysis of the samples from 2017 indicated very little virus diversity in most sites, hence the decision was taken to analyse pools of adults (**Figure 4-9 and 4-10**). If pools had been taken on a colony-by-colony basis, rather than per apiary, it may have been possible to determine changes happening within an apiary. Over time all the Sites shifted from Type A-dominant virus populations to Type B-dominant, regardless of viral titre or mite abundance (**Figure 4-14**). Knowing the virus populations of individual colonies may have allowed us to better understand changes in an apiary, such as whether colonies with high titres of Type A variants died over winter while colonies containing high titres of Type B survived, which would be one explanation for the significant shift from one variant to the other at Site level.

Recent studies using field-isolated stocks of Type A and B DWV have demonstrated that both viruses exhibit similar patterns of infectivity and mortality rates post-injection (of pupae), strongly suggesting that there is little phenotypic difference between these two variants (Tehel *et al.*, 2019). Therefore, the shift from Type A to Type B observed on the Isle of Arran may just represent a snapshot of a stochastically varied population that is constantly in flux. However, in contrasting research to Tehel *et al* (2019), field-isolate stock injections of Type A and Type B to *Varroa*-naïve colonies in Australia demonstrated that Type A initially replicated faster in pupae, but by 48 hours Type B infections had reached higher titres and remained higher for the remainder of the study (Norton *et al.*, 2020).

The inherent contradiction of these studies, coupled with high background viral titres in the control bees (>10^6 GE of Type B in Tehel *et al* (2019)) and therefore presumably the injected bees, indicates that wild-type virus stocks may not provide a robust system for these types of studies. The differences observed in these studies, which often rely on quantification of viral loads from field samples, may be due to fundamental differences in the characteristics of the viruses. Alternatively, the phenotypes may be due to differences in the levels of the variants present and external factors, such as the strain of honey bee infected or the mite infestation rates of the colonies sampled. Additionally, it is often not obvious whether the differences observed in distribution or replication rate for example, result in measurable differences in colony losses or overwintering success. There is therefore a disconnect between our understanding of the prevalence and incidence of DWV and the impact the virus strain has on colony losses. As mentioned, monitoring of colonies on Arran, rather than apiaries, may have provided valuable insight into the differences between these variants. The relationship between DWV sequence, strain variation and the consequences for the colony remains poorly understood and, because of the importance of DWV in overwintering colony losses, deserves further study.

Recent studies have used engineered cDNAs designed in a modular fashion, each containing unique restriction sites that allowed easy discrimination both from co-injected virus and from any endogenous DWV. Pupal injections with reverse genetic-derived viral clones of three variants (Type B (VVV), Type A coding region (VDD) and a recombinant (VVD)) caused similar levels of morbidity in eclosing workers and no observable differences in virulence (Gusachenko *et al.*, 2020a). DWV copy number was found to elevate from 10 injected copies to levels typically observed in overt infections (10^8–10^{10} GE/μg RNA) within 24 hours of inoculation. In these side-by-side comparisons all three variants achieved a similar titre after 48 hours.

Using NGS analysis, co-infections between Type A and Type B were examined to determine the extent of recombination between the variants. Recombination was observed widely across the genome, but with a larger number of reads observed at widely reported recombination junctions between DWV variants (Ryabov *et al.*, 2014; Dalmon *et al.*, 2017). A small number of the recombination junctions (~350 of 35750 unique junctions mapped) plotted as outliers from the diagonal of genome-length recombinants (**Figure 6-2** and **6-5**). Analysis of these sequences showed that the majority were out of frame deletions and so incapable of replicating. Further recombination junctions indicating large deletions were observed in intertypic recombinants but were also mostly out of frame (**Figure 6-6**). Research using analogous approaches in other RNA viruses has shown that these types of aberrant products are not unusual and reflect the random nature of the molecular mechanism of recombination (Lowry *et al.*, 2014; Alnaji *et al.*, 2020).

If individual variants (Type A, Type B and their recombinants) of DWV are equally virulent (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a) does the diversity within the population of a specific variant influence virus pathogenesis? This question is relevant because population diversity is known to influence virus fitness in other viruses, and has been demonstrated to influence interhost transmission and differ according to the tissue the virus replicates in. Viral fitness, for example defined by the ability of poliovirus to cross the blood-brain barrier, is reduced in low-diversity populations (Pfeiffer and Kirkegaard, 2006; Xiao *et al.*, 2017). Although this is of little relevance for poliovirus transmission within the population, it has a fundamental influence on poliovirus pathogenesis. More broadly, it probably reflects the ability

of the virus to overcome innate barriers to spread within and between tissues. Related studies on Chikungunya virus demonstrate that the passage history between host and vector influences population diversity and could increase diversity and adaptability to novel selection pressures (Coffey and Vignuzzi, 2011), such as the virus might encounter in a new host or when transmitted via a different route. Once spread within the host, the virus can evolve separately within tissues or organs, as has been demonstrated for HIV, where different virus populations are present in the blood and brain of infected individuals (Pillai *et al.*, 2006; Schnell *et al.*, 2009; Poirier and Vignuzzi, 2017). This in turn may influence the ability of the virus to escape those tissues or to transmit to a new host.

Unfortunately, studies into the influence of DWV population diversity are in their infancy and are hampered by the absence of a suitable cell culture system and the ubiquitous nature of DWV in the honey bee population. However, the progress made in other host–vector–virus systems give a good indication of what might be expected, and of the types of studies that will be needed to understand the specifics for DWV. Analysis of *Ixodes sp*. tick-transmitted Powassan virus in mice has demonstrated the influence of genetic bottlenecks due to the amount of virus transmitted and how purifying selection in ticks counteracts the role of host RNAi in driving virus variation (Grubaugh, Rückert, *et al.*, 2016). The kinetics of this series of events is markedly different from the evolution of mosquito-transmitted viruses such as West Nile Virus (WNV), reflecting the different pattern and frequency of vector feeding on the host through the season. Powassan virus diversity is greater in the host than the tick vector, but in the mosquito/bird/WNV system, diversity expands in the vector and is reduced in the vertebrate host (Grubaugh, Weger-Lucarelli, *et al.*, 2016).

Whether DWV fits either of these two contrasting patterns or has its own distinct characteristics has yet to be determined. Recent studies have demonstrated that mite feeding, with consequent virus transmission, does not result in the complete loss of DWV diversity (Annoscia *et al.*, 2019; Ryabov *et al.*, 2019). This is perhaps surprising considering the striking amplification (10^6-fold) of very limited amounts of virus in the 24–48 hours post-inoculation (Gusachenko *et al.*, 2020a). This may reflect the importance of the direct bee-to-bee transmission routes during trophallaxis and feeding, or vertically in the population. Alternatively, it may indicate that *in vitro* inoculation of DWV in the laboratory – either in these particular studies or more generally – does not properly replicate the events following mite transmission to a naïve pupa. Other studies have produced conflicting results. In early studies of changes in population diversity and viral load before and after *in vitro* DWV transmission, there was a striking switch from a low-level/high-diversity population to a high-level/low-diversity population (Wood *et al.*, 2014). In this study, the diversity in the amplified virus population after injection was at or near to the lower limit of detection of the next-generation sequencing strategy used to quantify variation. In contrast, there was no apparent restriction in virus diversity in isolates with high viral loads compared with those with low levels of virus in samples from the USA between 2015 and 2017 (Ryabov *et al.*, 2019), although this was not an equivalent study as it was based upon analysing viral load and diversity in individual field samples.

7.4 Modelling DWV diversity changes in individual workers

Simplistic modelling of the DWV transmission route between pupae and mites over a cycle of multiple pupal generations can be used to infer changes in diversity and viral titre between populations of variants. Significant changes in virus diversity, similar to those seen in an infested colony, can be accomplished within a few generations from very limited viral inocula that rapidly amplify in naïve pupae (See Chapter 5). In these models the reduction in viral

diversity is achieved without any assumptions about differential virulence of variants within the virus population, but with other impositions placed on virus replication, such as a 'bite volume limit' for the mite, whereby it takes a small number of viruses from the pupae as it feeds. These changes to the system had rapid and significant effects on the diversity of the virus population (**Figure 5-2**), essentially forcing the virus population through a diversity bottleneck.

Generation-to-generation changes in diversity have also been examined experimentally where serial passaging of relatively high titres of Type A DWV (2µl of the original 10^7 DWV/bee stock) across multiple generations of pupae was used in an attempt to simulate the transmission of the virus by *Varroa*. This showed that by the time the workers reached eclosion similar variants would dominate the population regardless of the passage number or whether the infection was symptomatic or asymptomatic. This was despite an initial increase in diversity post-injection and indicates that the selective pressure that favours certain variants may be occurring in a dose-dependent manner (Yañez *et al.*, 2020). In HIV infections mixed variant infections reduce the host response and subsequently undergo a less stringent transmission bottleneck compared to clonal infections. This then allows less-fit variants to establish in the population (Macharia *et al.* 2020). A similar effect may be observed in mixed populations of DWV. To properly examine the relative virulence of dominant genotypes and mixed populations, reverse genetic methods could be used to build recombinants and to replicate mixed virus populations observed during natural infections.

7.5 Conclusions and future work

In summary, the data in this thesis has demonstrated the positive impacts of using beekeeping techniques with appropriate miticide treatments to effectively reduce *Varroa* infestations from honey bee colonies, restoring the DWV population to a low level with no symptomatic infections, reflected in the very rapid reduction in the viral load within the population. It also demonstrated the impact of poor management in highly infested apiaries, whereby 'clean' colonies with low levels of both mites and viruses can rapidly become infested with mites, followed by sharp increases in DWV titres and, typically, colony death. The coordinated treatments of all known honey bee colonies on the Isle of Arran resulted in reduced *Varroa* mites across the island, including isolated sites which reported no *Varroa* in the final year of sampling, and a large increase in honey bee colonies, indicating that large scale management techniques like this are both plausible and beneficial. This project also showed that recombination occurs across the full DWV genome with hotspots of recombination detected when using viral clones isolated using reverse genetics methods. Additionally, a simple model was generated to highlight the variables that influence shifts in viral diversity in a mite–bee–virus system, which can be used to make inferences about the shifting dynamics of DWV infections.

Regarding DWV variants, changes in virus diversity, and clonality vs mixed populations of variants, the data presented in this thesis highlights that there are many questions remaining about DWV infection of honey bees. The use of high-titre stocks of near-clonal genetically tagged viruses will allow a number of questions regarding DWV pathogenesis and virus population dynamics to be addressed. For example, a mixed population of defined clonal variants generated using reverse genetics could be used to experimentally inoculate pupae at different ratios to better examine the dynamic shift in DWV populations. Recent published results have shown that only 80% of experimentally inoculated eclosing workers exhibit overt deformed wings, regardless of the variant inoculated (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a). The remainder appear phenotypically normal, despite having the same high level of

DWV. It could be that the presence of other low-level variants has a bearing on the development of symptomatic infections, either through evading immune responses or through selective pressures. Additional questions regarding Type A and Type B remain, as shown on Arran where there was a complete shift in dominance from one variant to the other over 3 years, even at locations seemingly isolated from the rest of the island.

Do physiologically distinct bees, such as workers produced mid-season and the winter bees produced in early autumn, show differential susceptibility? More fundamentally, what is different about the workers that emerge without overt symptoms, such as those sampled on Arran? Is the virus population in these bees different? It is possible that by only selecting 'healthy' adults on Arran the most pathogenic strains of the virus were missed. Finally, how important is overt disease anyway? Perhaps those with deformed wings are just the most susceptible bees in the population. The number of naturally infected pupae which emerge with symptomatic infections is not known, but it is possible the 80% reported via pupae injection (Tehel *et al.*, 2019; Gusachenko *et al.*, 2020a) is much higher than a natural infection ever produces. If very high levels of DWV always reduce the longevity of overwintering worker bees – the primary cause of winter colony losses – then the influence of strain variation, virus diversity and genotype on this characteristic is of paramount importance.

Answers to these questions will not directly or immediately improve honey bee health. They will fill in some of the numerous gaps we have in our understanding of the honey bee–virus–*Varroa* interaction, and in doing so may provide important insights to improve current treatment or management methods, or help define the markers required for selective breeding strategies. The recent development of tractable reverse genetic systems, if applied appropriately, will go a long way to improving understanding of the most important global threat to honey bee health, and establish the experimental framework in which new and emerging threats to their health are studied.

8 Appendices

8.1 Appendix 1 - Primers

Table 8-1 - primers used, including sequences, target and references.

Primer	Sequence (5' – 3')	Target	Source
DWVqPCR_F	ATATCACTTGGCGACGCAAC	VVD9-4_2u (4949-4968)	(Gusachenko *et al.*, 2020a)
DWVqPCR_R	CCAATCTTTAAATTGTTTCGGTTTTTGAGC	VVD9-4_2u (5097-5126)	
VDD_qPCR_F	GTCAAGAAGCAGGCGAATGTA	qPCR VDD	
VDD_qPCR_R	GCATAGGGGATCTAGAACACATAG		
VVV_qPCR_F	GAAAAGACGCGGGTGAGTTCG	qPCR VVV	
VVV_qPCR_R	AACATTGGCGATCGATTACAAACG		
Adapter-VDV-1 CP FP (389)	CTTGGTTAGCTGTGTTGCAGTTGCTGTAGTTAAGCGGTTATTAGAA	VDV-1 negative strand (4890-4912)	
Adapter, FP (388)	CTTGGTTAGCTGTGTTGCAGTTG	5' adapter on primer 389	
VDV-1 CP RP (1382)	GGTGCTTCTGGAACAGCGGAA	VDV-1 negative strand (4986-4966)	
155 DWV/VDV-1 FP	CAGTAGCTTGGGCGATTGTTTCG	VVD9 (4829-4851)	(Ryabov *et al.*, 2014)
156 DWV/VDV-1 RP	CGCACTTAACACACGCAAATTATC	VVD9 (6715-6738)	
153 DWV only RP	CTTGGAGCTTGAGGCTCTGCA	VVD9 (6513-6533)	
154 VDV-1 only RP	CTGAAGTACTAATCTCTGAG	VVV9-4 (6333-6352)	
RdRp FP	TGGAATACTAGTGCTGGTTTTCC	VVD9 (8608-8630)	This study
RdRp RP	CGACCCAATCCTCGAGCAT	VVD9 (9620-9638)	
L-protein FP	ACGGTACGTTACGTTCGCA	VVD9 (877-895)	
L-protein RP	CACTCTAACTCATATTCACGCTGTT	VVD9 (1691-1715)	
Inner L-protein FP	GGCCTTTAGTTGTGGAACTC	VVD9 (1133-1154)	
VDV-1 RdRp FP (inner)	TATCTTCATTAAAACCGCCAGGCT	VVV9-4 (8653-8677)	(McMahon *et al.*, 2016)
VDV-1 RdRp RP inner	CATCAGAATCTAACCCAGGACC	VVV9-4 (9082-9103)	
DWV-A RdRp FP inner	TGTCTTCATTAAAGCCACCT	VVD9 (8633-8652)	This study
DWV-A RdRp RP inner	CATCGGAATCTAACCCAGGACC	VVD9 (9061-9082)	
Sal1/Hpa1 FP	CAGGAATCCGCCAATTCAGGTAC	VVD9 (4951-4973)	(Gusachenko, *et al.*, 2020a)
VDV9 SRS R (6)	TTCCAGATGCACCACACATGC	VVD9 (5537-5557)	
DWV Full Genome FP1	CGTAGCATGAAGCGCATGCTTG	VVD9 (111bp - 132bp)	This study
DWV Full Genome FP2	TAGTCGTTTGTGGTTCAAG	VVD9 (211-229bp)	
DWV Full Genome FP3	GACCACTGCAGTATCGAGTAGAG	VVD9 (505bp - 527bp)	
DWV Full Genome RP1	TACGCGAGTAACACCTAAC	VVD9 (10068-10086bp)	
DWV Full Genome RP2	CCTAGAGTACCACTTTAATCG	VVD9 (10048-10068bp)	

DWV Full Genome RP3	GTCTTTAGAGGCAACTCAACC	VVD9 (9998-10018bp)	
DWV Full genome Survey1 FP4	GCGAATTACGGTGCAACTAAC	VVD9 (19-39bp)	
DWV Full genome Survey1 RP4	CGAGTAACACCTAACCTAGAG	VVD9 (10062-10082bp)	
Apis Actin FP	AGGGTGTGATGGTCGGTATGG	Apis Control	
Apis Action RP	GAGATCCACATCTGTTGGAAGG		
LepR1- RP	TAAACTTCTGGATGTCCAAAAAATCA	Apis sub-species ID	(Syromyatnikov, Borodachev and Kokina, 2018)
AmCarp FP	GAATATGAGCCGGAATAGTAGGA		
BQCV qPCR FP	AGGTTTACGCTCCAAGATCG	BQCV - 112bp	(Remnant *et al.*, 2019)
BQCV qPCR RP	TTTGTTCAGCAGGTAAATTGTTC		
ABPV FP	CATATTGGCGAGCCACTATG	ABPV – 398bp	(Bakonyi *et al.*, 2002)
ABPV RP	CCACTTCCACACAACTATCG		
SBV Grab FP	ACCAACCGATTCCTCAGTAG	SBV – 469bp	(Grabenstein er *et al.*, 2001)
SBV Grab RP	CCTTGGAACTCTGCTGTGTA		

8.2 Appendix 2 - SangerseqR R-script

This is an annotated R-script for Sanger sequence diversity analysis (see Chapter 2.3.2).

```
# Using SangerseqR to analyse Chromatogram sequence reads
# 11 July 17

# This is the analysis of sanger sequences for VVD9-4 and VVD9-5 in mixed
pools.
# Sanger sequences were generated using mixed ratios of both sequences and
their restriction sites will
# be used to distinguish their differences.
# Here, we try to plot the Chromatograms and show the sequence differences
in the pools.

# User manual for SangerseqR
browseVignettes("sangerseqR")

library(knitr, quietly=TRUE)
library(sangerseqR, quietly=TRUE)

library(Biostrings, quietly=TRUE)
opts_chunk$set(tidy=TRUE)

#modified from:
#https://github.com/yihui/knitr-examples/blob/master/077-wrap-output.md
hook_output = knit_hooks$get("output")
knit_hooks$set(output = function(x, options) {
  n <- 90
  x <- knitr:::split_lines(x)
  # any lines wider than n should be wrapped
  if (any(nchar(x) > n)) {
    x <- gsub(sprintf('(.{%d})', n), "\\1\n## ", x)
  }

  hook_output(x, options)
})

#Load a sequence file in to directory
file33F165 <- read.abif("33FI65.ab1")
file33F165

#Load from a sequence file object
homosangerseq <- sangerseq(file33F165)
str(homosangerseq

## ------------------------------------------------------------------------
---------------
#default is to return a DNAString object
Seq1 <- primarySeq(homosangerseq)
reverseComplement(Seq1)

#can return as string
primarySeq(homosangerseq, string=TRUE
#Build a chromatogram of the sequences in mix
```

```
## ------------------------------------------------------------------------
---------------
chromatogram(homosangerseq, width=100, height=2, trim5=50, trim3=100,
             showcalls='both', filename="chromatogram33FI65.pdf")

## ------------------------------------------------------------------------
---------------
# The makeBaseCalls function essentially divides the sequence into a serie
s of basecall windows
# and identifies the tallest peak for each fluorescence channel within the
window. These peaks are
# converted to signal ratio to the tallest peak. A cutoff ratio is then ap
plied to determine if a
# peak is signal or noise. Peaks below this ratio are ignored. Remaining p
eaks in each window are
# used to make primary and secondary basecalls.

homocalls <- makeBaseCalls(homosangerseq, ratio=0.33)
homocalls

## ------------------------------------------------------------------------
---------------
# This plot will show the highest peak and the secondary peak in the same
Chromatogram file
# The second peak will only be present if it is above the cut off ratio (i
n this case, set above as 0.33)

chromatogram(homocalls, width=100, height=2, trim5=50, trim3=100,
             showcalls='both', filename="chromatogram33FI65_basecall.pdf")

## ------------------------------------------------------------------------
---------------
# Here we need to set a reference sequence, however as these two seqs are
so similar, the first seq can
# can be used instead. If there were indels a ref would be needed to span
the gaps.

ref <- subseq(primarySeq(homosangerseq, string=TRUE), start=30, width=500)
homoseqalleles <- setAllelePhase(homocalls, ref, trim5=50, trim3=500)

## ------------------------------------------------------------------------
---------------
# At this point, we could plot the chromatogram again, but it is more info
rmative
# to align the resulting sequences to see how the alleles differ.
# We use Pairwise alignment for this

pa <- pairwiseAlignment(primarySeq(homoseqalleles)[1:570],
                        secondarySeq(homoseqalleles)[1:570],
                        type="global-local")
writePairwiseAlignments(pa)

# print A numerical matrix containing the maximum peak amplitudes for each
base within each
```

```
# Basecall window. If no peak was detected for a given base in a given win
dow, then 0.
# Column order = A,C,G,T
peakAmpMatrix(homosangerseq)

# shows the peak measurement for each base where more than one is called b
y BaseCall function
peakAmpMatrix(homocalls)

post<- peakAmpMatrix(homocalls)

# Forward sequence read for 1:1 ratio
# SalI site (GTCGAC or GTAGAT)
# position 140 - A = 286, C = 373
# position 143 - C = 267, T = 153

# HpaI site (GTTAAC or GTGAAT)
# position 272 - G = 236, T = 105
# position 275 - C = 173, T = 186
```

8.3 Appendix 3 - Correlation coefficient matrix

This is an R Markdown document for the Correlation Coefficient Matrix script from Chapter 4, including the plot details.

```
#-----------------------------------------

# Correlation Matrix for Arran

# -----------------------------------------

rm(list=ls())

#all factors converted to numeric values
#clonality - 1 = clonal, 0 = mixed
#movement - 0 - no movement, 1 = movement
# type -  typeA = 0, Type B = 1
# year - 2017 = 0, 2-18 = 1, 2019e = 2, 2019l = 3

#install.packages("Hmisc")
#install.packages("corrplot")
library("Hmisc")

library("corrplot")

#----------------------------------

Arr_core2 <- read.csv("Arr_correlation_matrix_withSBV.csv")
attach(Arr_core2)

#transform factor to numeric

Arr_core2$Clonality <- as.numeric(as.character(Arr_core2$Clonality))
Arr_core2$Dominant_Variant <- as.numeric(as.character(Arr_core2$Dominant_V
ariant))
Arr_core2$Colony_movement <- as.numeric(as.character(Arr_core2$Colony_move
ment))
Arr_core2$Year <- as.numeric(as.character(Arr_core2$Year))
Arr_core2$Site <- as.numeric(as.character(Arr_core2$Site))
Arr_core2$SBV <- as.numeric(as.character(Arr_core2$SBV))

head(Arr_core2, 10)

res1 <- cor(Arr_core2)
round(res1, 2)

res2 <- rcorr(as.matrix(Arr_core2))
res2

# Extract the correlation coefficients
res2$r

# Extract p-values
res2$P
```

```r
corrplot(res1, type = "upper", order = "hclust", method = 'color',
         tl.col = "black", tl.srt = 45, addCoef.col = "black")

corrplot(res1, type = "upper", order = "hclust", method = 'number',
         tl.col = "black", tl.srt = 45)

# Insignificant correlations are left blank
corrplot(res2$r, type="upper", order="hclust", method = 'color', tl.col =
"black", tl.srt = 45,
         p.mat = res2$P, sig.level = 0.01, insig = "blank")

# pmat : matrix of the correlation p-values
flattenCorrMatrix <- function(cormat, pmat) {
  ut <- upper.tri(cormat)
  data.frame(
    row = rownames(cormat)[row(cormat)[ut]],
    column = rownames(cormat)[col(cormat)[ut]],
    cor  =(cormat)[ut],
    p = pmat[ut]
  )
}

library(Hmisc)
res2<-rcorr(as.matrix(Arr_core2[,1:10]))
flattened_table<- flattenCorrMatrix(res2$r, res2$P)

write.table(flattened_table, file = "Arr_cor_flattened_table2.xls", sep =
"\t", quote = FALSE, row.names = T)
#creates a 4 column table of data as so
#Column 1 : row names (variable 1 for the correlation test)
#Column 2 : column names (variable 2 for the correlation test)
#Column 3 : the correlation coefficients
#Column 4 : the p-values of the correlations

#comute the p-value of the correlation matrix
cor.mtest <- function(res1, ...) {
  res1 <- as.matrix(res1)
  n <- ncol(res1)
  p.mat<- matrix(NA, n, n)
  diag(p.mat) <- 0
  for (i in 1:(n - 1)) {
    for (j in (i + 1):n) {
      tmp <- cor.test(res1[, i], res1[, j], ...)
      p.mat[i, j] <- p.mat[j, i] <- tmp$p.value
    }
  }
  colnames(p.mat) <- rownames(p.mat) <- colnames(res1)
  p.mat
}

# matrix of the p-value of the correlation
p.mat <- cor.mtest(Arr_core2)
head(p.mat[, 1:10])
```

```r
# Specialized the insignificant value according to the significant level
corrplot(res1, type="upper", order="hclust", method = 'color',
         tl.col = "black", tl.srt = 45,
         p.mat = p.mat, sig.level = 0.01)

#col <- colorRampPalette(c("#BB4444", "#EE9988", "#FFFFFF", "#77AADD", "#4
477AA"))
corrplot(res1, method="color", #col=col(200),
         type="upper", order="hclust",
         addCoef.col = "black", # Add coefficient of correlation
         tl.col="black", tl.srt=45, #Text label color and rotation
         # Combine with significance
         #p.mat = p.mat, sig.level = 0.01, insig = "blank",
         # hide correlation coefficient on the principal diagonal
         diag=FALSE
)

corrplot(res1, method="color", #col=col(200),
         type="upper", order="hclust",
         #addCoef.col = "black", # Add coefficient of correlation
         tl.col="black", tl.srt=45, #Text label color and rotation
         # Combine with significance
         p.mat = p.mat, sig.level = 0.01, #insig = "blank",
         # hide correlation coefficient on the principal diagonal
         diag=FALSE
)

#install.packages("PerformanceAnalytics")
library("PerformanceAnalytics")

my_data <- Arr_core2[, c(1,3,4,5,6,7, 8, 9, 10)]
chart.Correlation(Arr_core2, histogram=TRUE, pch=19)

#///////////////////////////////////////////////////////////////////
```

8.4 Appendix 4 - DWV population model R Script

Below is the annotated R script for the mite-bee-DWV population model generated in R, as discussed in Chapter 5.

```
rm(list=ls())
library(stringr)

data2<-c() # Create empty vector

#big loop to run the whole thing 10 times and catch all data together
for (i in 1:10) {

  #here are all the model variables
  virusLoad_L = 1000 # initial viral pop in bee
  virusLoad_M = 100000 # amplified virus in mite
  virusLoad_H = 1000000 # amplified virus in new bee
  # varroabite = 0.5 #fixed varroabite variable
  #varroabite = round(runif(1, 1.0, 1.0),2) # random % of virus mite takes
from bee between 0.1/ 2%
  #varroainject = round(runif(1, 1.0, 10.0),1) # % of virus mite injects i
nto bee
  BiteLimit =10 # maximum amount of virus mite can take from bee
  InjectLimit = 100000 # the amount the varroa injects into a new bee (10%
of virus pop ammplified)
  strains=c("A","B","C","D","E","F") # calling strains
  strainFitness=c(0.18, 0.18, 0.18, 0.18, 0.18, 0.18) # probability of a s
pecific strain being amplified


  # Xstrain takes the 6 strains and using prob calls different proportions
of each
  # Xstrain<- sample(c("A","B","C","D", "E", "F"), size = 50, replace = TR
UE, prob = c(.75, .05, .05, .05, .05, .05))

  ref_virusPop_tmp <- sample(strains, virusLoad_L, replace = TRUE)


  # here is the loop
  for (i in 1:30) {
    varroabite = round(runif(1, 3.0, 3.0),2) # moved variable into loop -
randomized
    varroainject = round(runif(1,10.0, 10.0),1) # moved variable into loop
- randomized

    # make an ifelse to say take % of viruses or a max number in the bite
    if ((length(ref_virusPop_tmp) * (varroabite/100)) > BiteLimit) {
      biteSize <- BiteLimit
    } else {
      biteSize <- length(ref_virusPop_tmp) * (varroabite/100)
    }

    # this is the viruses in the bite
    ref_virusPop_bite <- sample(ref_virusPop_tmp, biteSize, replace=TRUE)
```

```
    # this is the viruses amplified in the mite, including probability par
ameter after certain number of iterations
    # so we can use this to distort one virus at one particular cycle of f
eeding
    #if(!(x > 0)){
    if (i != 1) {
      ref_mitePop_tmp <- sample(ref_virusPop_bite, virusLoad_M, replace =
TRUE)
    } else {
      # here is the probability part First, make a copy of the ref_virusPo
p_bite data frame
      probNumbs<-ref_virusPop_bite

      # Convert all A's to a probability.
      # The conversion in question is the first entry in the "strains" var
iable, in this case A - command is strains[idiot] to specify the A.
      # The probability is the first number in the strainsFitness variable
. This is what the strainsFitness[idiot] refers to
      # Then that number is divided by the total number of A's in the ref_
virusPop_bite data frame

      # idiot here (me), calls each letter sequentially in the loop and ap
plies strainfitness value to them
      for (idiot in 1:length(strains)){
        probNumbs[probNumbs==strains[idiot]]<-(strainFitness[idiot]/sum(st
r_count(probNumbs,strains[idiot])))
      }

      # Now calculate new mite population including these probabilities
      ref_mitePop_tmp <- sample(ref_virusPop_bite, virusLoad_M, replace =
TRUE, prob=probNumbs)
    }

    # this ifelse takes % of those viruses or a max number (defined at top
) for injection
    if ((length(ref_mitePop_tmp) * (varroainject/100)) > InjectLimit) {
      injectSize <- InjectLimit
    } else {
      injectSize <- length(ref_mitePop_tmp) * (varroainject/100)
    }

    # Here is the new bee with its 1000 virus gut pop
    new_beepop <- sample(strains, virusLoad_L, replace = TRUE)

    # this injects the viruses into a new bee (10% or 10,000)
    ref_mite_inject <- sample(ref_mitePop_tmp, injectSize, replace = TRUE)

    # now amplify the injected viruses in the bee
    ref_bee_amplifier <- sample(ref_mite_inject, virusLoad_H, replace = TR
UE)

    # here is the virus population in the bitten bee
    ref_virusPop_tmp<-c(ref_bee_amplifier, new_beepop)
```

```
    A<-(sum(str_count(ref_virusPop_tmp,"A"))/length(ref_virusPop_tmp))*100
    B<-(sum(str_count(ref_virusPop_tmp,"B"))/length(ref_virusPop_tmp))*100
    C<-(sum(str_count(ref_virusPop_tmp,"C"))/length(ref_virusPop_tmp))*100
    D<-(sum(str_count(ref_virusPop_tmp,"D"))/length(ref_virusPop_tmp))*100
    E<-(sum(str_count(ref_virusPop_tmp,"E"))/length(ref_virusPop_tmp))*100
    F<-(sum(str_count(ref_virusPop_tmp,"F"))/length(ref_virusPop_tmp))*100

    data<-c(i,A,B,C,D,E,F, varroabite, varroainject) # create vector with
percentages of current iteration
    data2<-cbind(data2,data) # bind previous data and current vector toget
her

    print(paste0("Varroa bite vol =", varroabite))
    print(paste0("Varroa inject % =", varroainject))
    print(paste0("Iteration ", i))
    print(paste0("Strain A: ", A, "%"))
    print(paste0("Strain B: ", B, "%"))
    print(paste0("Strain C: ", C, "%"))
    print(paste0("Strain D: ", D, "%"))
    print(paste0("Strain E: ", E, "%"))
    print(paste0("Strain F: ", F, "%"))
    cat("\n")
    cat(paste0("\n\n"))

  }


  #save the data as a .csv file and make a basic plot
  data3<-t(data2) # transpose the dataset
  colnames(data3)<-c("iteration","A","B","C","D","E","F", "Varroa bite vol
", "Varroa inject %") # add column names
  data3<-as.data.frame(data3, row.names = 1) # convert to data frame
  write.csv(data3, 'data3.csv')

}
```

8.5 Appendix 5 - Perl Script for ViReMa data compiling

This is an annotated perl script for sorting ViReMa file output (see Chapter 6).

```
#!/usr/bin/perl
# v2
my $flattenedOutput = 'yes'; # set to anything but 'yes' to produce multi-
line output for each junction found
my $file = shift @ARGV or die "I need a file to work with: $!";
open (FILE, "<", $file) or die "Can't open $file: $!";
my $match;
my $matchfull;
while (<FILE>) {
    chomp;
    if ($_ =~ /NewLibrary:\s+(.*)/) {
        $match = $1;
        my ($donor,$acceptor) = split (/_to_/, $match);
        foreach my $parent ($donor,$acceptor) {
            my ($id,$strand) = split (/_/, $parent);
            unless ($strand =~ /RevStrand/) { $strand = 'ForStrand'}
            $matchfull .= $id . ',' . $strand . ',';
        }
    } elsif ($_ =~ /EndofLibrary/) {
        $match = '';
        $matchfull = '';
        next;
    } else {
        my @counts = split (/\t+/, $_);
        foreach my $c (@counts) {
            my ($start,$to,$end,$hash,$number) = split (/_/, $c);
            unless ($flattenedOutput eq 'yes') {
                for (1 .. $number) {
#                    print join (",", $matchfull, $start, $end, "\n");
                    print join (",", $match, $start, $end, "\n");
                }
            } else {
                print join (",", $match, $start, $end, $number, "\n");
#                print join (",", $matchfull, $start, $end, $number, "\n")
;
            }
        }
    }
}

close FILE;
exit ();
```

8.6 Appendix 6 - Bubble-plots for ViReMa Recombination events

These are the recombination plots for each superinfection sample analysed using ViReMa (see Chapter 6).

Table 8-2 - All samples and total precise recombination reads.

Experimental set up	VVV>VDD	VDD>VVV	Total
V>D 5d # 2	2501	2871	5372
D>V 5d # 1	1900	2002	3902
D>V 5d # 3	1639	1829	3468
D>V 5d # 4	1107	1191	2298
V>D 7d # 2	2836	3068	5904
V>D 7d # 3	1194	1339	2533
D>V 7d # 1	1711	2001	3712
D>V 7d # 2 (6)	1080	1274	2354
D>V 7d # 3 (7)	1901	2063	3964
D>V 7d # 4	1110	1136	2246
Total			35753

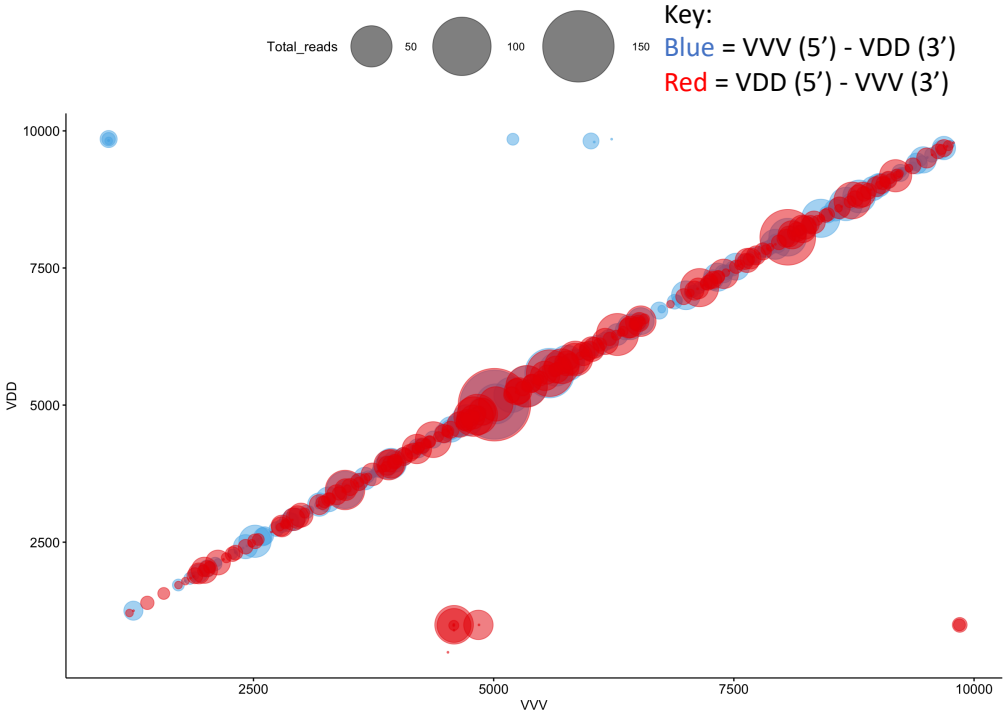


Figure 8-1 - Bubble plot for sample V>D #2 - 5 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

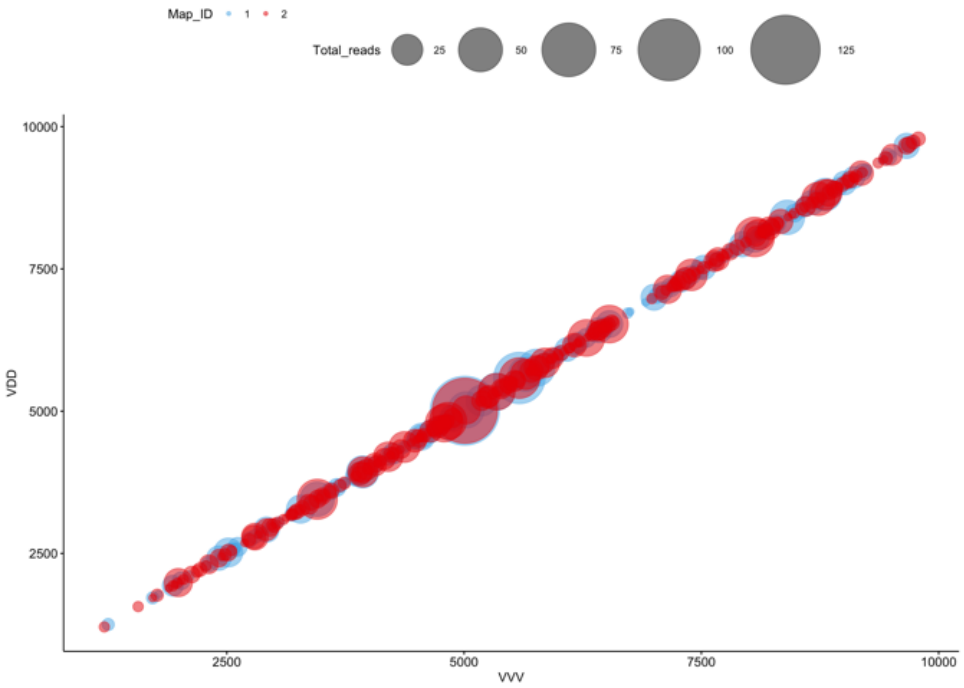


Figure 8-2 - Bubble plot for sample D>V #1 - 5 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

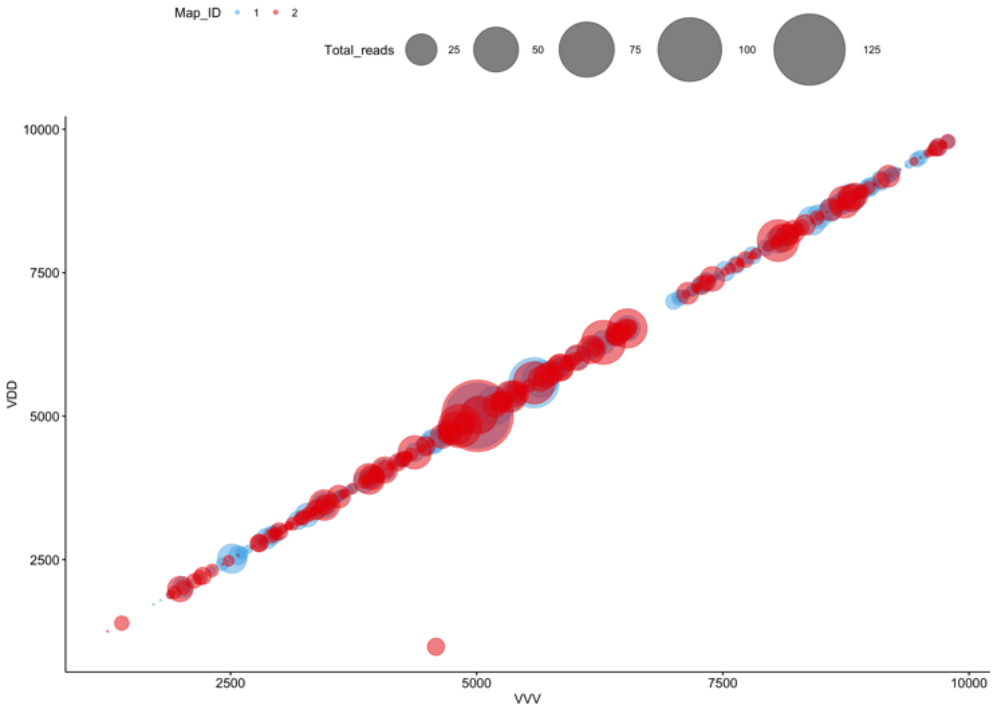


Figure 8-3 - Bubble plot for sample D>V #3 - 5 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

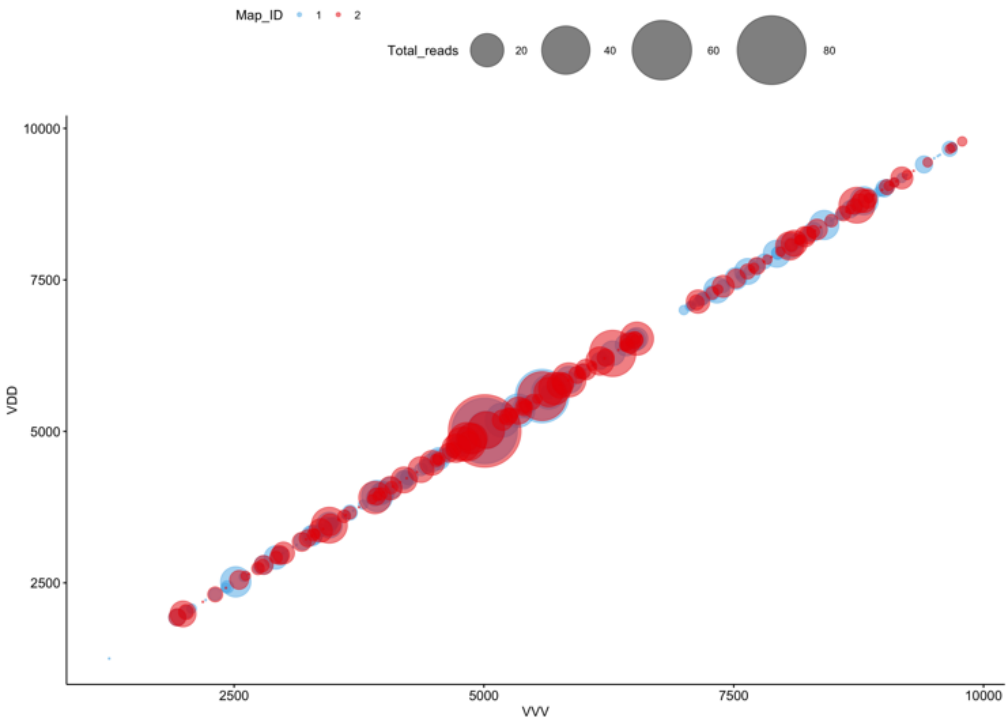


Figure 8-4 - Bubble plot for sample D>V #4 - 5 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

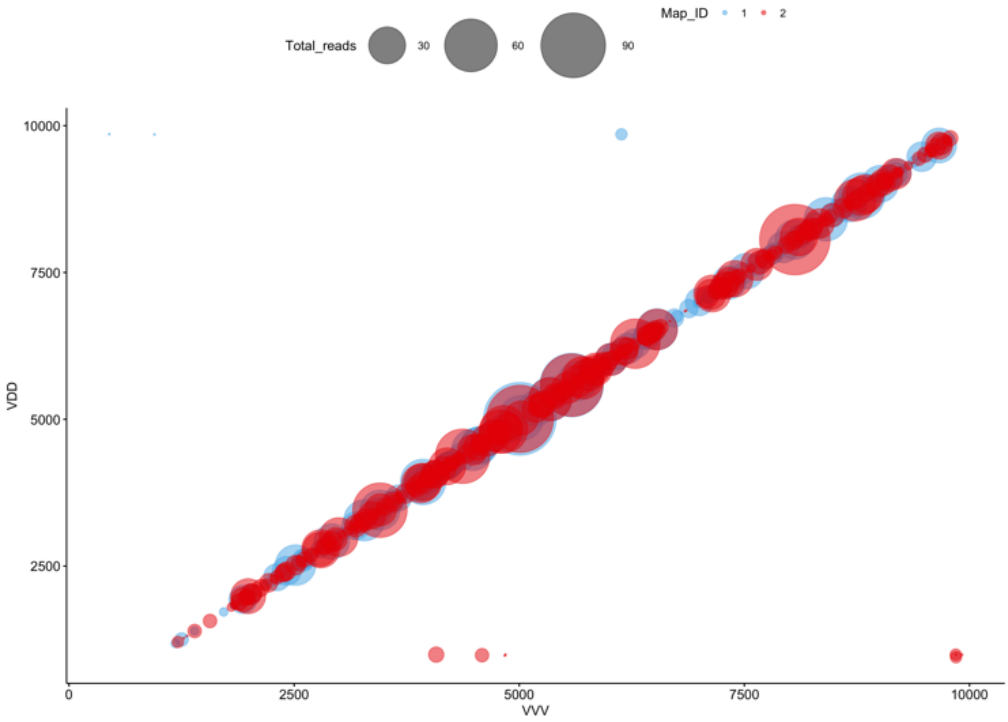


Figure 8-5 - Bubble plot for sample V>D #2 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

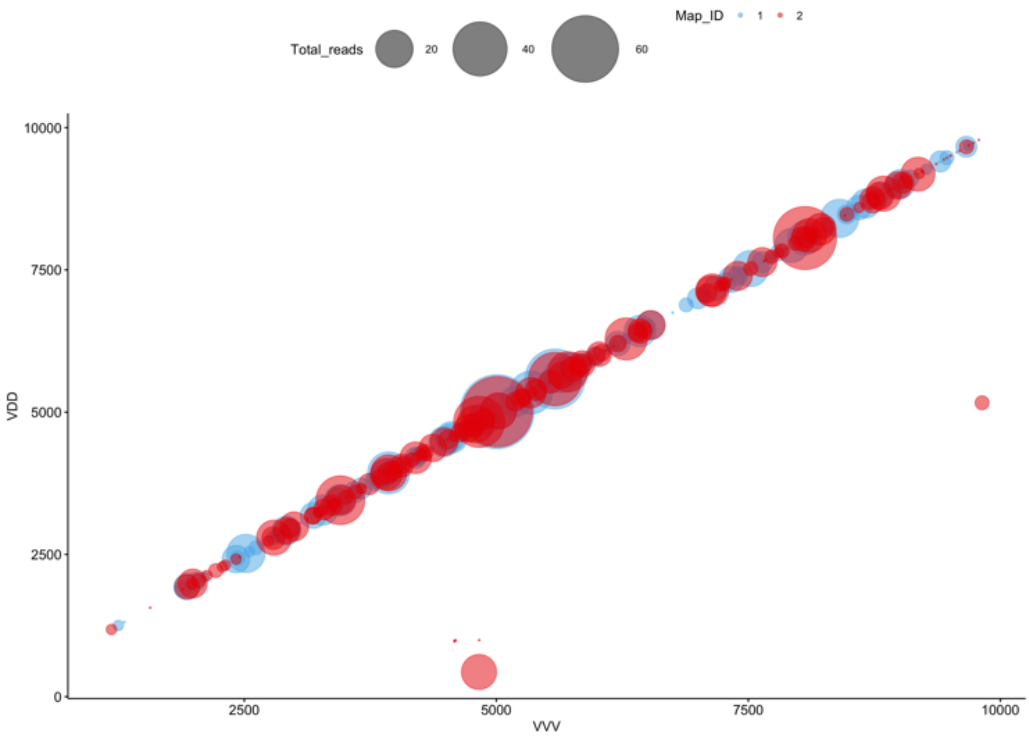


Figure 8-6 - Bubble plot for sample V>D #3 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

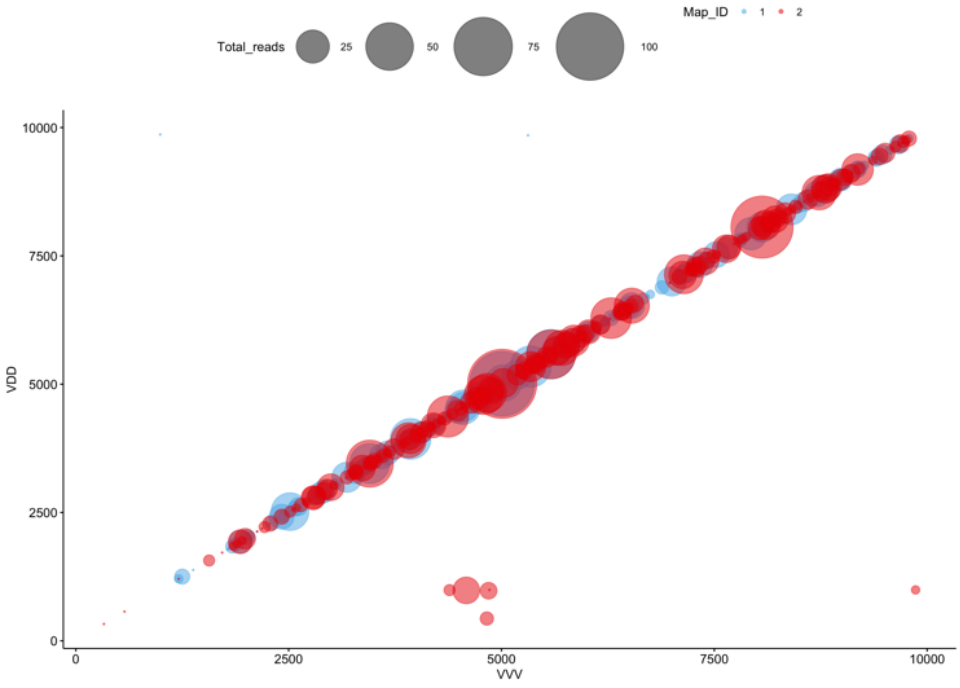


Figure 8-7 - Bubble plot for sample D>V #1 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

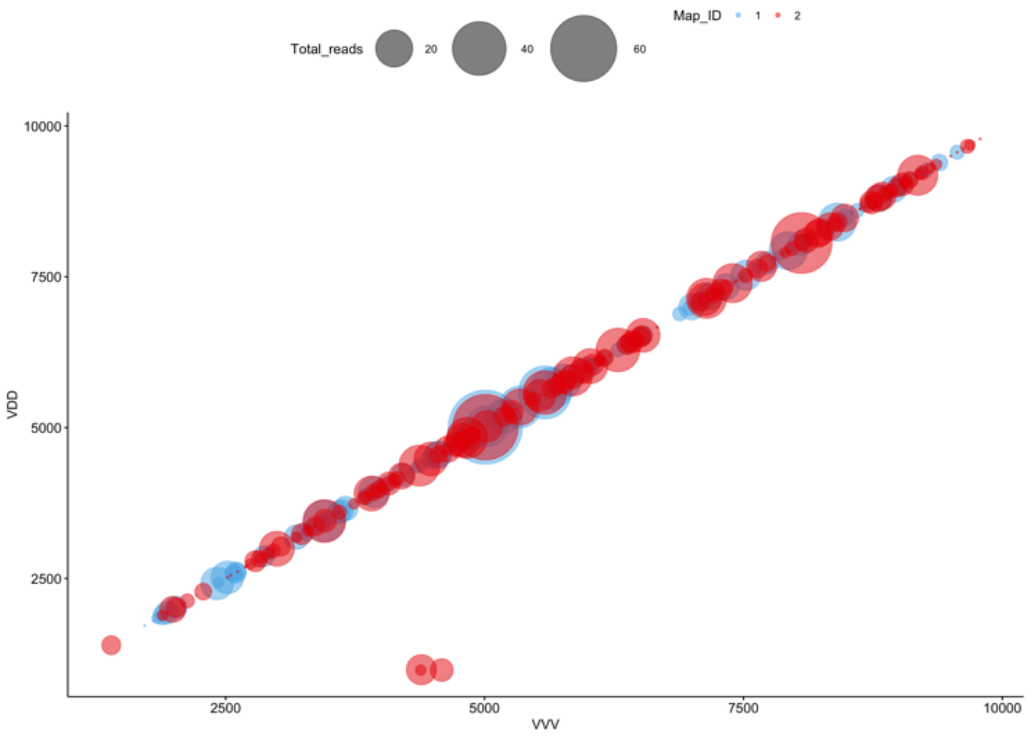


Figure 8-8 - Bubble plot for sample D>V #2 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

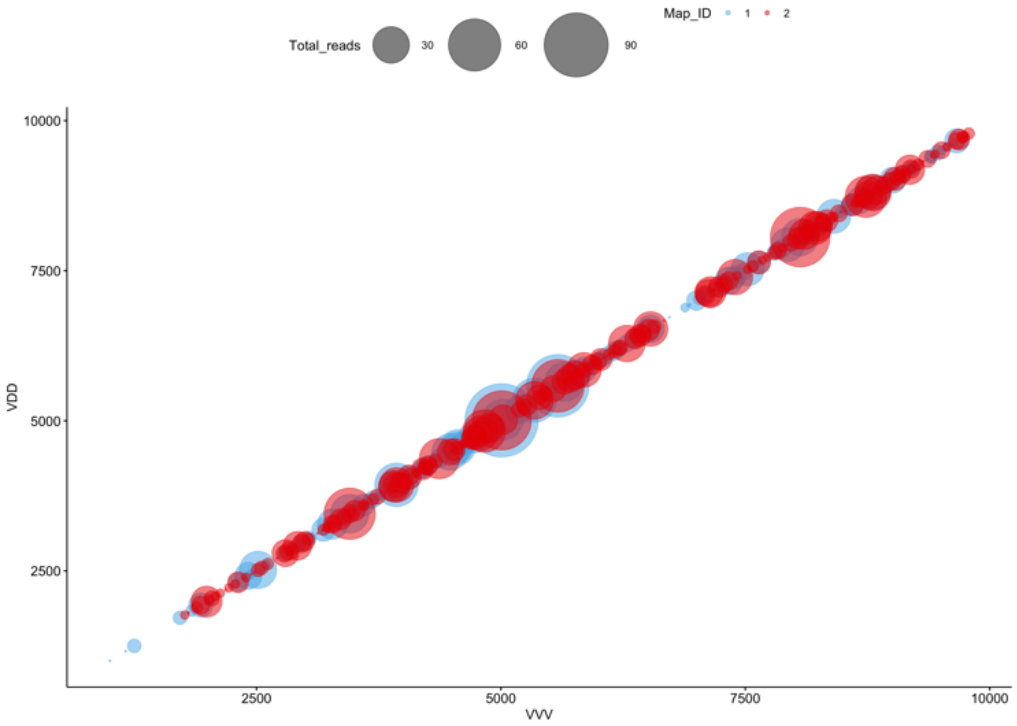


Figure 8-9 - Bubble plot for sample D>V #3 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

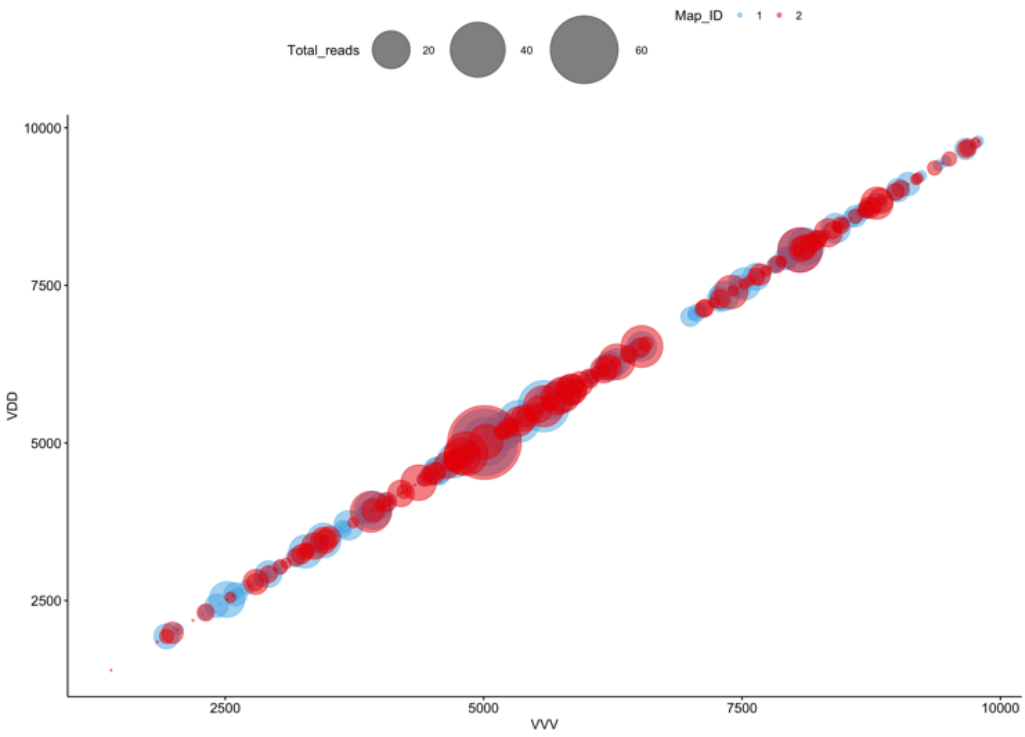


Figure 8-10 - Bubble plot for sample D>V #4 - 7 days. Blue bubbles = VVV (acceptor) >VDD (donor) and red bubbles = VDD (acceptor) >VVV (donor).

8.7 Appendix 7 - Published manuscript – Woodford & Evans (2020), FEMS.

FEMS Microbiology Reviews, fuaa070, 1–20

doi: 10.1093/femsre/fuaa070
Advance Access Publication Date: 15 December 2020
Review Article

REVIEW ARTICLE

Deformed wing virus: using reverse genetics to tackle unanswered questions about the most important viral pathogen of honey bees

Luke Woodford*,† and David J. Evans

Biomedical Sciences Research Complex, University of St Andrews, St Andrews, KY16 9ST, UK

*Corresponding author: Biomedical Sciences Research Complex, University of St Andrews, St Andrews, Fife, KY16 9ST. Tel: +44 (0)1334 463396; E-mail: lw86@st-andrews.ac.uk

One sentence summary: *Deformed wing virus* is the most important viral pathogen of honey bees and along with its vector, *Varroa destructor*, is responsible for large-scale annual colony losses; recent advances in molecular biology have offered new insights into the biology of this honey bee virus, which address the determinants of tropism and pathogenesis, the importance of strain variation and the significance of possible pathogen spillover to other pollinating insects.

Editor: Blossom Damania

†Luke Woodford, http://orcid.org/0000-0003-2530-2120

ABSTRACT

Deformed wing virus (DWV) is the most important viral pathogen of honey bees. It usually causes asymptomatic infections but, when vectored by the ectoparasitic mite *Varroa destructor*, it is responsible for the majority of overwintering colony losses globally. Although DWV was discovered four decades ago, research has been hampered by the absence of an *in vitro* cell culture system or the ability to culture pure stocks of the virus. The recent developments of reverse genetic systems for DWV go some way to addressing these limitations. They will allow the investigation of specific questions about strain variation, host tropism and pathogenesis to be answered, and are already being exploited to study tissue tropism and replication in *Varroa* and non-*Apis* pollinators. Three areas neatly illustrate the advances possible with reverse genetic approaches: (i) strain variation and recombination, in which reverse genetics has highlighted similarities rather than differences between virus strains; (ii) analysis of replication kinetics in both honey bees and *Varroa*, in studies that likely explain the near clonality of virus populations often reported; and (iii) pathogen spillover to non-*Apis* pollinators, using genetically tagged viruses to accurately monitor replication and infection.

Keywords: *Deformed wing virus*; *Varroa destructor*; reverse genetics; honey bee; virus bottleneck; pathogen spillover

GENERAL OVERVIEW

The European honey bee, *Apis mellifera*, is an economically important insect managed globally for honey production, commercial pollination of agricultural crops and wildflower pollination. Terrestrial insect biomass is declining at an alarming rate (van Klink *et al.* 2020) and there is an increasing focus on the factors causing this. One suggestion is that the pests and pathogens of managed honey bee colonies, present at high levels in some environments, 'spill over' to non-*Apis* pollinators, so accounting for some of these losses (Fürst *et al.* 2014). *Deformed wing virus* (DWV) is the most important and ubiquitous viral pathogen of honey bees. [This review will refer to DWV and VDV-1, the first identified and hence prototype viruses, as genotypes. The use of terms 'variant' or 'strain' are generic terms that refer to sequences closely related to these two genotypes. The use of Type A and Type B have become a convenient and well-used shorthand way to refer to DWV and VDV-1, respectively, in the published literature and as such are used throughout this manuscript. In this review, where a specific genotype is

Received: 6 October 2020; **Accepted:** 11 December 2020

being referred to this will be made clear.] DWV usually causes asymptomatic infections but, when vectored by the ectoparasitic mite *Varroa destructor*, it is responsible for the majority of overwintering honey bee colony losses. Although the virus was discovered four decades ago, research has been limited by the absence of an immortalised cell line to propagate the virus *in vitro*. In addition, without a reverse genetic system for the virus, there has been no easy way to address fundamental questions relating genotype and phenotype. Although a cell culture system is still needed, reverse genetic systems have been independently developed in recent years. These allow specific questions to be addressed about virus strain variation, host and tissue tropism and pathogenesis. These and future advances in the understanding of DWV disease in honey bees should allow the role of the virus in the proposed decline of pollinating species to be more readily investigated. As the *Varroa*/DWV combination is of paramount importance in honey bee health, it has been extensively reviewed, including recent discussions on the structure of the virus particle, honey bee viruses, the life cycle of the ectoparasitic mite *Varroa destructor* and host immunity to DWV (Nazzi and Le Conte 2016; Nazzi and Pennacchio 2018; Grozinger and Flenniken 2019; Martin and Brettell 2019; Traynor *et al.* 2020). In contrast, the technology and opportunities offered by a reverse genetics system for DWV have not been reviewed. We focus on the key features of reverse genetics and its application to strain variation, virus competition, replication kinetics and potential pathogen spillover to non-*Apis* pollinators.

INTRODUCTION—THREATS TO HONEY BEES

Key pollinators in decline?

The evolutionary origin of the European honey bee (*Apis mellifera*) is unclear, with evidence supporting expansion from either Asia or Africa (Whitfield *et al.* 2006; Han, Wallberg and Webster 2012). What is clear is that subsequent global expansion was undoubtedly aided by human intervention, initially by provision of nesting sites and subsequently by direct colony management, with evidence for recognisable 'beekeeping' activity dating back several thousand years. This anthropogenic expansion was driven by the demand for honey, a delicious and valuable natural source of sugars. More recently, the movement of honey bees, nationally and globally, has expanded in order to increase pollination of a large number of economically important agricultural crops. Although local pollinating insects play a key role in the pollination of crops and wildflowers, the honey bee is considered essential for meeting the growing demands of some sectors of agriculture. For example, every year approximately two thirds of all colonies in the United States are moved to California, where they pollinate almond crops that yield up to 80% of the total world almond harvest (Beaurepaire *et al.* 2020). In the UK, there are 39 insect pollinated agricultural crops, the vast majority of which are pollinated by honey bees, and in the United States there are over 130 agricultural plants pollinated by honey bees and bumble bees (Delaplane and Mayer 2000). Despite the obvious need for honey bees, and the increased pollinator dependence of modern agriculture (Aizen *et al.* 2019), the global stock of farmed bees is not growing at the rate required by the demands of the industry (Aizen and Harder 2009). Additionally, although the positive impact of farmed honey bees on agricultural production is well established, there are suggestions they may be detrimental to ecosystems (Lindstro¨m *et al.* 2016). Meta-analysis indicates that, whilst *A. mellifera* is the most frequent floral visitor in natural ecosystems, a large proportion of plants were never visited and so were reliant on other pollinators; further studies will be required to better determine the importance of honey bees on natural habitat pollination (Hung *et al.* 2018).

Despite the number of managed colonies increasing globally by 45%, there have been large annual regional declines [e.g. 59% decline in the period 1947–2005 in the United States (Potts *et al.* 2010), average 28% annual US colony losses 2006–2020 (Bruckner *et al.* 2020) and up to 32% overwinter losses in parts of Europe between 2018 and 2019 (Gray *et al.* 2020)]. These losses have not been attributed to a single cause and are believed to reflect a combination of anthropogenic factors. These include increased pesticide use (Chauzat *et al.* 2009), the induced stress of long distance movement of colonies to aid selected crop pollination (Morimoto *et al.* 2011), and the increased presence and transmission of pathogens. In addition, monoculture farming practices lead to a substantial decrease in suitable habitat (Donaldson-Matasci and Dornhaus 2012). These habitat losses have been deemed one of the largest causes of total insect population declines (Hallmann *et al.* 2017) and increased habitat fragmentation is a key factor for honey bee survival too (Brown and Paxton 2009). Following very large-scale honey bee colony losses in the United States in 2006, the term 'colony collapse disorder' (CCD) became widely used to describe the phenomenon and particular features of lost hives (Hayes 2007; Watanabe 2008; vanEngelsdorp *et al.* 2009). Although the term CCD remains a convenient acronym to define colony losses, losses of a similar magnitude and type have not been routinely reported subsequently. The causes of CCD remain unclear and it is thought to be a combination of some or all of these anthropogenic factors (Cox-Foster *et al.* 2007). However, the magnitude of the losses refocused attention on the increasing issues of pathogens in managed honey bee colonies.

A significant factor in honey bee health is the spread of parasites and pathogens. Honey bees have evolved to control pathogens that spread within the colony and, where honey bees nest at low densities, colony to colony transmission has been naturally restricted (Seeley 2007; Seeley *et al.* 2015). In contrast, managed honey bee hives are co-located in apiaries, are routinely moved for crop pollination and are kept at a significantly higher overall density than natural colonies. This increases the risk of disease transmission within and between colonies or apiaries, through the natural processes of drifting, where bees relocate to a hive other than their natal colony (Pfeiffer and Crailsheim 1998; Nolan and Delaplane 2016), and during the robbing of weak colonies by nearby strong colonies (Peck and Seeley 2019). Due to these processes many pathogens have been able to establish and spread rapidly through the global honey bee population (Dietemann, Ellis and Neumann 2013).

Honey bees are exposed to a host of different pathogens, from fungi such as *Nosema* spp. or *Ascosphaera apis* (Aronstein and Murray 2010) to bacterial infections like American and European foulbroods (*Paenibacillus larvae* and *Melissococcus plutonius*) (Bailey 1983; Lauro *et al.* 2003) and a range of viruses, including *Deformed wing virus*, *Chronic bee paralysis virus*, *Black queen cell virus* and *Sacbrood virus* (Yan˜ ez *et al.* 2020b). Additionally, parasites such as *Acarapis woodi*, *Tropilaelaps* spp. and, most notably, *Varroa destructor* infest honey bee colonies. These may cause disease directly and/or act as vectors for other pathogens (Jong, Morse and Eickwort 1982).

Varroa destructor and virus transmission

Originally an ectoparasitic mite of the Asian honey bee *Apis cerana*, *Varroa destructor* jumped host to the European honey bee, *Apis mellifera*, in Asia. The mite subsequently spread globally as honey bee colonies were transported from country to country, reaching Europe in the 1960s (Traynor *et al.* 2020) and the UK in 1992 (Oldroyd 1999). Originally known as *Varroa jacobsoni*, molecular methods subsequently showed that *Varroa destructor* was a distinct species (Anderson and Trueman 2000).

The long-held understanding that *Varroa* feed on the haemolymph of developing pupae has recently been questioned by studies that suggest it is the fat bodies that are the major target (Ramsey *et al.* 2018). Notwithstanding this, *Varroa* ingests and subsequently transmits a cocktail of viruses while feeding on honey bee pupae. The most important of these is the single-stranded positive-sense RNA virus, DWV. The introduction to and dissemination of *Varroa destructor* in the global honey bee population added a fundamentally new transmission route for DWV.

In the absence of the mite, DWV is transmitted horizontally and vertically within the colony. Infected nurse bees transmit the virus to developing larvae during feeding and the virus is detectable in the gut of most managed honey bee populations (Tentcheva *et al.* 2004). The virus is also transmitted vertically via eggs and sperm (Yue *et al.* 2007) from the queen and drones to progeny. In the presence of *Varroa*, DWV is able to bypass the normal routes of transmission and is directly injected into the developing honey bee when the mite feeds on pupae (Yue and Genersch 2005).

The presence of *Varroa* within one colony also potentially impacts the health of neighbouring colonies. Modelled data and microsatellite analysis has indicated as much as 42% of workers in a colony could be alien due to drifting between colonies in the same apiary (Pfeiffer and Crailsheim 1998; Forfert *et al.* 2015). Phoretic mites are known to spread between colonies on worker bees as they drift between hives, meaning that highly infested hives can transfer mites to neighbouring colonies (Fries and Camazine 2001). Studies have demonstrated re-infestation rates of ~76 mites a day recorded in mite free colonies from mite-infested hives in the same environment (Greatti *et al.* 1992). In addition, weak colonies, potentially debilitated by high levels of mite infestation, are robbed by neighbouring strong colonies that, in turn, acquire an increased level of mites and the novel virus populations they harbour (Peck and Seeley 2019; Woodford unpublished data). This is further exacerbated by colonies with high *Varroa* infestation showing an increased acceptance of drifting workers, allowing for the potential uptake of other associated pests and pathogens (Forfert *et al.* 2015). Feral colonies, presumed to originate from swarms lost from managed hives, have pathogen levels—including of DWV and *Varroa*—that are as high as those in unmanaged hived colonies (Thompson *et al.* 2014). This indicates that feral colonies may act as reservoirs for re-infestation within the environment, and a potential source of disease for non-*Apis* pollinators.

DEFORMED WING VIRUS

The biology, prevalence and distribution of *Deformed wing virus*

DWV is a member of the *Iflavirus* genus of viruses, family *Iflaviridae*, order *Picornavirales*, which also includes the honey bee pathogen sac brood virus (SBV) (Ghosh *et al.* 1999). The name *Deformed wing virus* encompasses at least three genetically similar viruses that were independently isolated. DWV was originally isolated from honey bees in the UK (Bailey and Ball 1991) and a very similar virus, designated VDV-1, was isolated from *Varroa* parasitising *A. mellifera* in the Netherlands (Ongus *et al.* 2004). A third closely related virus, designated Kakugo ('ready to attack' in Japanese), was isolated from the brains of honey bee exhibiting aggressive behaviours (Fujiyuki *et al.* 2004). For the purpose of this overview the name DWV will be used, unless specific reference to VDV-1 or Kakugo is required. The genetic relationships between DWV strains is addressed below.

The genome organisation of DWV is characteristic of members of the *Picornavirales*. The RNA genome of DWV contains a single large open reading frame (ORF) that encodes both structural and non-structural proteins of the virus (Ongus *et al.* 2004). The ORF is flanked by a 1144-nucleotide 5′ non-translated leader sequence and a 317-nucleotide 3′ non-translated region terminated by a poly(A) tail (Lanzi *et al.* 2006). The 2894-amino acid ORF is translated via an internal ribosomal entry site (IRES) located in the 5′ non-translated region. The structural proteins, VP1–4, that form the icosahedral viral capsid (Organtini *et al.* 2017; S˘kubnı´k *et al.* 2017) are located towards the N-terminal end of the polyprotein, preceded by the leader (L) polypeptide, the function of which has yet to be defined. The remainder of the polyprotein carries the non-structural proteins responsible for genome replication and the manipulation of the cellular milieu. These include a recognisable RNA helicase, a chymotrypsin-like 3C protease and an RNA-dependent RNA polymerase (Lanzi *et al.* 2006). The region of the genome encoding the L protein is a hotspot for variation between different viral variants (Lanzi *et al.* 2006) and contributes 26.2–33.3% of all amino acid variation found between DWV isolates despite comprising only 7.3% of the polyprotein sequence. This large disparity in the diversity of the L-protein and the remainder of the genome has been observed in other picornaviruses where the L protein has a number of different functions including the stimulation of IRES activity or the inhibition of host cell mRNA translation (Glaser, Cencic and Skern 2001; Hinton *et al.* 2002).

Evidence suggests that DWV is widespread in the honey bee population, even in geographic areas with no current or historical mite infestation. The exception may be Australia, which is reported to lack both the mite and DWV (Roberts *et al.* 2017). DWV has been detected in honey bee populations in isolated regions with no *Varroa* presence, such as on the island of Colonsay, Scotland and the Isle of Man, UK (Fu¨ rst *et al.* 2014; Ryabov *et al.* 2014; McMahon *et al.* 2015), indicating that the virus is probably naturally endogenous in honey bees and suggesting that reports of its absence may reflect poor assay sensitivity or sampling techniques. If DWV is an endogenous viral parasite of honey bees it is unclear how Australian bees remain uninfected, as honey bees were introduced there from Europe in the 1820s. Despite the widespread distribution of DWV, there is compelling evidence that the anthropogenic dissemination of *Varroa* through the global honey bee population has also resulted in the transmission of particular DWV variants (Wilfert *et al.* 2016).

Figures quoted for DWV prevalence have gradually increased as assay methods have improved, with data from 32 countries suggesting an average prevalence of 55% (10–100%) (Martin and Brettell 2019). In the absence of *Varroa*, viral levels in individual adult workers can be very low [<1000 genome equivalents per microgram (GE/μg) of total RNA] but, as discussed below, can increase 1–100 million times after transmission by *Varroa*. This has fundamental implications for prevalence studies; pooled worker bee or pupal samples from colonies with even low levels

of *Varroa* infestation may well exceed the detection threshold for DWV. In contrast, analysis of individual workers or pupae using poorly designed or insensitive assays may fail to detect DWV in the sample, particularly in colonies with low mite levels or of unparasitised bees.

DWV infections altered by *Varroa* mite transmission

How is a potentially ubiquitous virus largely or completely asymptomatic in the absence of *Varroa*? This appears to be related to the viral titre, potentially to the diversity of the virus population, and to the honey bee tissues to which the virus has access after *Varroa* transmission. In healthy honey bee colonies the virus population is highly diverse and the viral loads are low, typically 10^4–10^6 copies per bee with 5–20 variants reported (Martin *et al.* 2012; Mondet *et al.* 2014; Ryabov *et al.* 2014), and bees that harbour such low titres are asymptomatic. During horizontal transmission, for example during trophallaxis, the bee is exposed to DWV via the gut, an organ shaped by evolution to provide protection from environmental pathogens (Mikonranta *et al.* 2014). Altering the transmission route leads to significant changes in the virus population, including a large increase in the viral load (titres increase by ~10^6-fold per bee) that is often accompanied by a marked reduction in virus diversity (Martin *et al.* 2012; Nazzi *et al.* 2012; Ryabov *et al.* 2014).

Virus transmission by mites is causally associated with the characteristic symptoms of DWV infection, including crumpled and poorly developed wings, general paralysis, discoloration and abdominal bloating, as well as reduced longevity of worker bees (Highfield *et al.* 2009; Mo¨ckel, Gisder and Genersch 2011; Dainat *et al.* 2012; Natsopoulou, McMahon and Paxton 2016; Benaets *et al.* 2017). However, it does not invariably cause overt symptoms. Honey bees with low DWV can still rarely develop crippled wings, possibly independently of the virus. In addition, up to 25% of bees with high levels of DWV (>10^{10} copies/bee) can develop normal wings (Tehel *et al.* 2019; Gusachenko *et al.* 2020a). These apparently morphologically normal workers may still have impaired foraging abilities, loss of cognitive function and a reduced lifespan despite appearing healthy (Iqbal and Mueller 2007; Dainat *et al.* 2012; Campbell *et al.* 2016; Gisder *et al.* 2018). Highly infected honey bees are still able to compete reproductively and transmit virus, as shown when highly DWV-infected drones transmit virus when successfully mating with queens, resulting in infections of >10^7 copies in some mated queens (Amiri, Meixner and Kryger 2016).

Investigating DWV tissue tropism using western blotting of VP1 from infected adult bees showed that the virus was concentrated in the head and abdomen, with the thorax containing lower levels of viral antigens (Lanzi *et al.* 2006). Kakugo virus was preferentially located in the brains of aggressive worker bees (Fujiyuki *et al.* 2004, 2005) and VDV-1 has also been reported to replicate in the heads of worker honey bees (Zioni, Soroker and Chejanovsky 2011). By RT-PCR (reverse transcription polymerase chain reaction), the virus was shown to be distributed throughout the body of symptomatic workers, with elevated levels of virus antigen in all tissues, whilst asymptomatic workers only showed positive results for DWV in the thorax and abdomen, not the head (Yue and Genersch 2005).

Until recently virus localisation studies were conducted using western blot on isolated tissues, or by a variety of *in situ* hybridisation studies. Sensitivity of these assays varies and can be influenced by the detection reagents and/or the sequence variation of the virus. With the advent of reverse genetic technology (see the section 'Using reverse genetics to investigate the biology of DWV') it is possible to engineer reporter genes directly into the genome to facilitate detection (Gusachenko *et al.* 2020a; Ryabov *et al.* 2020). Live imaging of larvae fed with very high doses of an engineered DWV expressing the green fluorescent protein (GFP) showed virus was localised throughout the body, with significantly elevated foci of infection in the wing rudiments of the thoracic segments, the spiracle openings, the developing ovaries and in the head. In pupae infected with a similar GFP-expressing DWV by injection, virus fluorescent signal was detected in the head tissues, developing wings and throughout the digestive tract (Gusachenko *et al.* 2020a). These studies demonstrate that infectious virus spreads throughout the developing bees readily and can affect many organs.

The genetic diversity of *Deformed wing virus*

Replication errors by the RNA-dependent RNA polymerase of the single-stranded positive-sense RNA viruses mean that these viruses exist as a 'cloud' of very closely related variants termed a quasispecies (Domingo *et al.* 2005). In addition to this variation, there are identifiably divergent DWV isolates that have variously been termed DWV or Type A DWV (Lanzi *et al.* 2006) and VDV-1 or Type B DWV (Ongus *et al.* 2004). There is an additional, though rarely reported, Type C variant (Mordecai *et al.* 2015b; de Souza *et al.* 2019; Kevill *et al.* 2019). It should be noted that the Type C virus is distinct from Kakugo, which is identifiably a Type A strain of the virus.

These three variants exhibit very significant genetic relatedness, sharing over 80% nucleotide identity and >90% amino acid identity across the genome. For comparison, the three serotypes of the distantly related poliovirus (the prototype picornavirus), all of which exhibit similar tropism and pathogenesis, are 71% identical at the nucleotide level and 69–88% at the amino acid level, depending on the region (Toyoda *et al.* 1984).

A comparison of amino acid and nucleotide identity between DWV Type A (NCBI accession no. AJ489744.2) and Type B (NCBI accession no. AY251269.2) revealed a high degree of sequence similarity, with only the variable Leader protein region showing <83% nucleotide and 96.8% amino acid similarity (Table 1). Phylogenetic analysis of the nucleotide and amino acid differences between DWV Type A, DWV Type B and KV (Kakugo virus) indicate they are small enough (~84% average nucleotide identity) to be considered variants of the same virus (Lanzi *et al.* 2006; Ryabov *et al.* 2014).

In the VP1 region of the DWV genome, the Type-A and -B DWV variants exhibit an 84.5% nucleotide and 98% amino acid identity. VP1 is an extensively studied region for measuring diversity between variants of human enteroviruses, a distantly related group of viruses with a similar genome organisation (Oberste *et al.* 1999). Phylogenetic analysis of poliovirus and human enterovirus species C have shown that the variants within these groups are highly similar (differing by 14–16% between variants) and some strains have been reclassified based on their high amino acid similarity (>96%) and phenotypically similar infections (Brown *et al.* 2003). Sequence analysis of the VP1 region of human enteroviruses showed intraserotypic divergence of ~25% in the nucleotide sequence and 12% in the amino acid sequence (Oberste *et al.* 1999). Although antibody selection will have driven some variation in the vertebrate-infecting enteroviruses, it is likely that RNAi may act similarly in honey bees (Ryabov *et al.* 2014), making the comparison of sequence variation and virus/strain classification valid. Considered in this context, the sequence divergence of the even the most divergent DWV variants would place them as a single virus group.

Table 1. Sequence similarities between two well-characterised DWV variants—DWV Type A (NCBI accession no. AJ489744.2) and DWV Type B (NCBI accession no. AY251269.2).

Region of genome	Genome location	Sequence identity (%)	Amino acid identity (%)
Leader protein	1132–1764	74.24	83.4
VP3	1765–2523	84.04	96.8
VP1	2587–3834	84.46	97.6
VP2	3835–4608	84.9	98.1
RNA helicase	4996–6411	88.69	98.1
Protease	7681–8274	87.69	97.5
RdRp	8551–9615	83.29	97.5

It should be noted that even sequences with very limited genetic divergence can generate viruses with fundamentally different phenotypes. For example, a single nucleotide substitution in poliovirus determines the majority of the neurovirulence phenotype of the virus (Almond *et al.* 1987). It remains possible that the genetic differences reported between the type A, B and C strains of DWV might obscure similar fundamental determinants that influence the resultant phenotype. The recent availability of reverse genetic systems will, as they did with poliovirus, allow this to be determined.

A comparison of all DWV sequences >9500 bp uploaded on the NCBI database revealed a close genetic similarity (Fig. 1). Using a neighbour-joining tree, all sequences clustered into two distinct clades, which have historically been referred to as 'DWV' or Type A and 'VDV-1' or Type B, but the two clades are in fact closely related, indicating a strong similarity between all DWV variants. Furthermore, and as elaborated upon in the section 'Does DWV diversity matter for pathogenesis?', several laboratories have identified genetic recombinants between Type A and B DWV with virulent phenotypes, which have all clustered tightly with the Type B variants during phylogenetic analysis.

For reasons associated with their original isolation and the differences in their distribution and prevalence, many studies have focused upon the differences in infectivity and pathogenesis of DWV Type A and Type B. Natsopoulou *et al.* (2017) observed a strong association between colonies with higher overwinter mortality rates and elevated levels of the Type B strain in the individual workers analysed. This may be related to the reported reduction in worker bee longevity when infected with Type B when compared with the Type A (53.5–38%, respectively) (McMahon *et al.* 2016) and with the faster replication of this strain of the virus (McMahon *et al.* 2016), which may explain why Type B establishes in the host population more readily. Research in apiaries in the United States suggests that Type B has emerged more recently, with findings of only 2.7% prevalence increasing to 66% from 2010 to 2016 (Ryabov *et al.* 2017), although DWV Type A was still the most widespread variant found there (in 89% of samples) and these differences may be dependent on the sensitivity of the assay used. However, other studies have produced contradictory results. Despite apparently replicating more slowly than Type B, the Type A virus has been reported as the more pathogenic variant in both the UK and United States (Kevill *et al.* 2019). The global spread of Type A was strongly associated with the spread of *Varroa*, although Type B was not investigated in this study (Wilfert *et al.* 2016).

The differences observed in these studies, which all rely on quantification of viral loads from field samples, may be due to fundamental differences in the characteristics of the viruses. Alternatively, the phenotypes may be due to differences in the levels of the variants present and external factors, such as the strain of honey bee infected or the mite infestation rates of the colonies sampled. Additionally, it is often not obvious whether the differences observed in distribution or replication rate for example result in measurable differences in colony losses or overwintering success. There is therefore a disconnect between our understanding of the prevalence and incidence of DWV and the impact the virus strain has on colony losses. Clearly the relationship between DWV sequence, strain variation and the consequences for the colony remains poorly understood and, because of the importance of DWV to overwintering colony losses, deserves further study.

Using reverse genetics to investigate the biology of DWV

Understanding the relationship between the type and strain of the virus associated with disease, or the influence of different replication kinetics for example, usually involves the isolation and characterisation of the virus *in vitro*, often as a precursor to comparative infection studies *in vivo*. However, attempts to develop a viable cell culture system of *Apis mellifera* cells have been unsuccessful (Genersch *et al.* 2013; Goblirschi, Spivak and Kurtti 2013). A more recent report using Lepidopteran P1 cell lines has shown limited DWV replication (Erez and Chejanovsky 2020). However, the replication observed in the latter is orders of magnitude less than seen in honey bee pupae and its relevance and suitability for DWV studies remains to be determined.

Therefore, in the absence of a suitable cell line, studies of the fundamental characteristics of the virus have to be conducted *in vivo*, typically in the presence of the endogenous DWV that is ubiquitously present. Such *in vivo* studies present additional problems; these include strain variation of the host honey bees and the inability to propagate or purify clonal virus populations from mixed infections, together meaning that the results obtained may confound subsequent analysis. Recombination, discussed further in the section 'Are recombinant forms of DWV more pathogenic?', between the injected clonal virus and the endogenous virus population further complicates investigation by this route.

The use of reverse genetics to recover a virus and investigate the relationship between genotype and phenotype has become a standard method in virology for four decades (Taniguchi, Palmieri and Weissmann 1978; Racaniello and Baltimore 1981). This approach facilitates the study of virus replication and pathogenesis and the recent application of this strategy to further our understanding of DWV is an important development.

Following elucidation of the near-full-length genome sequence (Lanzi *et al.* 2006) a series of reverse genetics systems

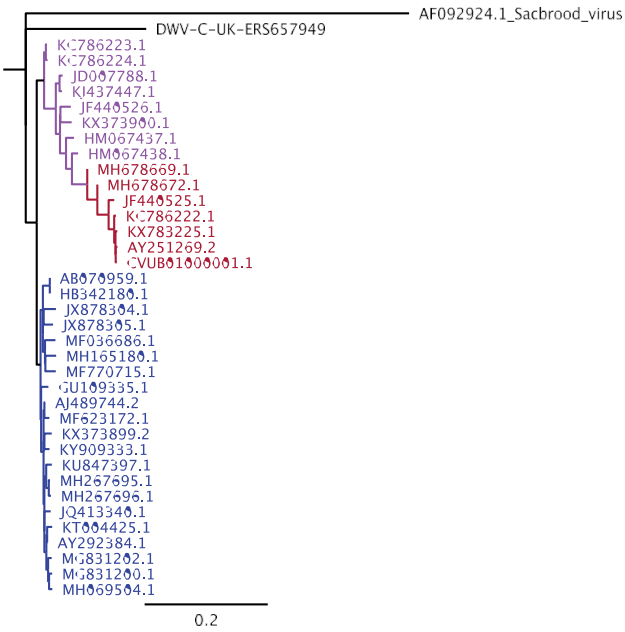


Figure 1. Neighbour-joining tree of all DWV sequences >9500 bp available from the NCBI database. Red sequences are classified on the database as Type B or VDV-1. Pink sequences are classified as recombinant forms of DWV and blue sequences are classified as Type A or DWV. The Type B sequences have clustered at the bottom of the red branch (sequences MH678669.1–CVUB01000001.1), whilst the genomes showing closer sequence homology to the Type A sequences are reported as recombinants (for example, KC786223.1). The only complete Type C sequence is included and a complete sequence of Sacbrood virus is used as an outgroup. Using Geneious, a Tamura–Nei genetic distancing model was applied to generate a neighbour-joining phylogenetic tree. Following sequence alignment, 1000 bootstrap iterations were run to generate the final tree.

have been developed for DWV, based upon Type A, Type B or Type A/B recombinants seen to predominate in some colonies (Lamp *et al.* 2016; Ryabov *et al.* 2019; Gusachenko *et al.* 2020a). These reverse genetics tools at least partially circumvent the need for virus isolation and allow comparisons to be made between different infectious clones of DWV variants.

This technology is well established but new to the study of DWV. Briefly, a full-length cDNA (complementary DNA) is maintained in a plasmid vector, typically bearing a DNA-dependent RNA (Ribonucleic acid) polymerase promoter (e.g. bacteriophage T7) allowing *in vitro* transcription of the sense strand of the viral genome. The RNA produced *in vitro* is directly injected to honey bee pupae harvested from a colony and maintained in the laboratory (Fig. 2). In analogous cell culture systems, the input RNA is translated and then subsequently replicates. Presumably the same processes occur in injected honey bee pupae, though details of the cell types the RNA is taken up by, or the mechanism by which this occurs, remains unknown.

Importantly, the input RNA is by definition representative of the virus positive strand. Therefore, the subsequent occurrence of complementary negative strand RNA is indicative of virus replication. To distinguish the input RNA from any pre-existing or co-replicating endogenous DWV the viral cDNA is engineered to contain unique restriction endonuclease sites, typically by exploiting synonymous codon variation, that allow the unambiguous identification of the input genome. By coupling this unique restriction endonuclease 'tag' with a negative strand-specific PCR (polymerase chain reaction) assay (Gusachenko *et al.* 2020a), it is possible to definitively demonstrate the replication of the virus genome. A good example of the further exploitation of this approach is the engineering of reporter genes into the virus genome that enable the direct visualisation of sites of virus replication (Gusachenko *et al.* 2020a; Ryabov *et al.* 2020).

DNA dependent RNA polymerases have low error rates meaning that there is little variation in the input RNA (<0.5 errors per genome). Although the error rate of the DWV polymerase is not known, it is likely to be approximately an order of magnitude higher (~10^{-3}–10^{-4} errors per nucleotide) based upon studies of other picornaviruses (Ward, Stokes and Flanegan 1988). Notwithstanding the viral polymerase error rate, the resulting virus population arising from the input RNA is near clonal. In addition, because of the very rapid amplification of directly inoculated virus arising from injected RNA (see the section 'Changes in virus replication and mite-related bottlenecks') the resulting harvested and purified DWV is likely to be almost exclusively (>99.99%) derived from the input *in vitro*-synthesised RNA (Gusachenko *et al.* 2020a).

It has recently been demonstrated that DWV replicates in the buff-tailed bumble bee, *Bombus terrestris* (see the section 'Pathogen spillover—is DWV an environmental threat?'). As part of these studies, DWV RNA injected to *Bombus* pupae allowed the recovery of DWV, albeit with lower efficiency than achieved by injection in honey bee pupae (Gusachenko *et al.* 2020b). Since

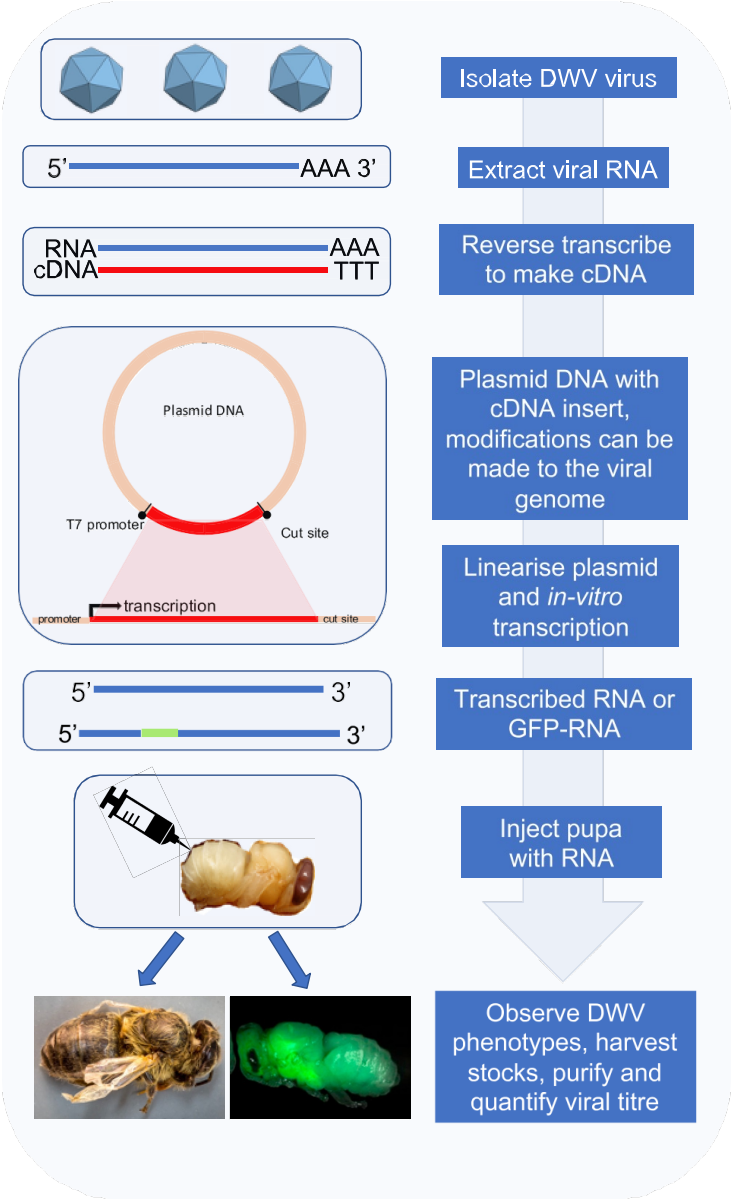


Figure 2. Schematic diagram of the generation of reverse genetics-derived clones of DWV. Deformed bee photo adapted with permission from theapiarist.org; green bee image adapted with permission from Gusachenko *et al.* (2020a).

was not as pronounced when the recombinant was the first virus administered. Further studies are required to determine the molecular mechanism that accounts for the inability of the recombinant to compete efficiently in the presence of a genetically related pre-established virus infection. The sequence identity between the established and incoming viruses may prime the RNAi-based immune response and suppress the replication of the second virus. Studies of differences in the RNAi response, a proportion of which is directed against the capsid-coding sequences (Ryabov *et al.* 2014), will be needed to investigate this further. If supported, this could help develop an assay to identify RNAi suppressors that would be expected to be present within the DWV genome (Gammon and Mello 2015). The proposed superinfection exclusion of Type A by Type B (Mordecai *et al.* 2015a) deserves further study to determine whether it can be recapitulated with other wild-type virus stock preparations, and to determine how sensitive it is to the presence of subpopulations of recombinant viruses within the inocula.

Are recombinant forms of DWV more pathogenic?

Genetic recombination—the formation of hybrid or chimeric genomes—is an intrinsic consequence of the replication mechanism of most single-stranded positive-sense RNA viruses. Evidence currently suggests that it is an evolutionarily retained mechanism (as non-recombinogenic viruses can be engineered; Woodman *et al.* 2016) that either compensates for the inherently error-prone polymerase these viruses possess, by 'rescuing' damaged genomes, or actively contributes to the generation of quasispecies diversity. There are numerous examples where recombination leads to the development of novel phenotypes, in both plant and animal viruses (Bentley and Evans 2018). The polymerase switches templates during replication, probably in a negative-strand sequence-independent manner (Kirkegaard and Baltimore 1986; Lowry *et al.* 2014) to generate a hybrid genome, the viability of which then determines whether it is further propagated or disappears. Replication-competent recombinants may undergo or acquire further adaptive changes, the final result being a chimeric genome with breakpoints marking the junction between the two (or more) parental sequences from which it was derived. Knowing where the breakpoints are in viral recombination events can provide insight into the biochemical and mechanistic properties of viral infection (Rohayem, Munch and Rethwilm 2005; Simon-Loriere *et al.* 2009). The modular nature of the genomes of all single-stranded positive-sense RNA viruses contributes to the process of recombination, with the junctions usually located between the modules for the structural proteins, the non-structural proteins or the non-coding regions and the coding regions (Lukashev *et al.* 2003; Lukashev 2005; Lowry *et al.* 2014; Bentley and Evans 2018).

Recombination has been observed between DWV Type A and Type B strains, and in some instances these recombinant forms have been reported to predominate in a population over the strains they originate from (Moore *et al.* 2011; McMahon *et al.* 2016). Moore *et al.* (2011) identified two DWV recombinants, VDV-1$_{DVD}$ [comprising the DWV Type A 5′-UTR (untranslated region), a VDV-1 structural gene and DWV Type A non-structural genes] and VDV-1$_{VVD}$ (comprising the 5′-UTR of VDV-1, a VDV-1 structural gene and DWV Type A non-structural genes). An observed correlation between high VDV-1$_{DVD}$ titres in infected pupae and their associated mites suggested this recombinant may have been better adapted to transmission between *Varroa* and honey bees than its parental variants. McMahon *et al.* (2016) found that Type A and Type B viral variants recombined readily *in vivo*, indicating non-independence of the two viruses among foraging honey bees. In laboratory studies with wild-type (i.e. not derived by reverse genetics) purified virus preparations, a type B/A recombinant (DWV VVD) had a selective advantage over non-recombinant stocks when injected into pupae (Ryabov *et al.* 2014).

The predominance of a particular virus population within a *Varroa*-infested colony is a consequence of the original viruses present, the passage history of the population and the acquisition of exogenous viruses from neighbouring hives or, possibly, through environmental contamination. If recombinant viruses are more transmissible, either vertically or horizontally from bee to bee or via *Varroa*, it will inevitably influence the resulting virus population. This is an inherently noisy system in which to reach definitive conclusions about relative virulence of different virus strains. In contrast, cDNA-derived viruses directly inoculated to honey bee pupae in the laboratory can be relatively easily compared in virulence and replicative capacity. Under these conditions, recombinants (at least those tested to date), do not appear to out-replicate the parental viruses to which they are related (Gusachenko *et al.* 2020a).

Are recombination events sequence specific?

'Hotspots' in the viral genome are where recombination preferably occurs. One such hotspot, observed by Ryabov *et al.* (2014) and Moore *et al.* (2011), is between the structural and non-structural proteins. The most variable region of the genome, and therefore an ideal candidate region for measuring the extent of sequence diversity in a virus population, is the leader protein (a 211-codon region towards the 5′ end of the viral genome) with a nucleotide identity match of only ~74% between DWV and VDV-1 (Dalmon *et al.* 2017). Observations of recombination events occurring in the leader protein region between VDV-1 and DWV-A have been observed in recent analyses of samples in the United States (Ryabov *et al.* 2017). Additionally, hotspots were independently discovered in the 5′ end of the genome (McMahon *et al.* 2016; Cornman 2017). All these identified recombination junctions occur between functional 'modules' of the virus genome.

The Type A/B recombinant viruses observed by Ryabov *et al.* (2014) predominated in a colony when compared with the parental Type A or Type B strains that were also present. This suggests that recombination may be due to fitness selection or the result of a bottleneck event (see the section 'Changes in virus replication and mite-related bottlenecks'). Without evidence to the contrary it is likely that DWV (like other single-stranded positive-sense RNA viruses) recombines using a mechanism similar to that observed during poliovirus recombination (Kirkegaard and Baltimore 1986). This is a biphasic process in which the initial recombinant product, which may have acquired sequence duplications during recombination, undergoes a subsequent resolution process involving the loss of sequence duplication and the optimisation of virus fitness (Lowry *et al.* 2014). Multiple recombination events within a single DWV genome have been observed. For example, Dalmon *et al.* (2017) reported a recombinant DWV genome that contained a Type B 5′-UTR, a Type A leader protein sequence, the remaining structural proteins from Type B and the non-structural proteins from Type A, implying that at least three independent recombination events had occurred. It is notable that the majority of recombinant strains reported in that study derived their structural proteins from Type B sequences (Dalmon *et al.* 2017), indicating that this variant may have a selective advantage either

was not as pronounced when the recombinant was the first virus administered. Further studies are required to determine the molecular mechanism that accounts for the inability of the recombinant to compete efficiently in the presence of a genetically related pre-established virus infection. The sequence identity between the established and incoming viruses may prime the RNAi-based immune response and suppress the replication of the second virus. Studies of differences in the RNAi response, a proportion of which is directed against the capsid-coding sequences (Ryabov *et al.* 2014), will be needed to investigate this further. If supported, this could help develop an assay to identify RNAi suppressors that would be expected to be present within the DWV genome (Gammon and Mello 2015). The proposed superinfection exclusion of Type A by Type B (Mordecai *et al.* 2015a) deserves further study to determine whether it can be recapitulated with other wild-type virus stock preparations, and to determine how sensitive it is to the presence of subpopulations of recombinant viruses within the inocula.

Are recombinant forms of DWV more pathogenic?

Genetic recombination—the formation of hybrid or chimeric genomes—is an intrinsic consequence of the replication mechanism of most single-stranded positive-sense RNA viruses. Evidence currently suggests that it is an evolutionarily retained mechanism (as non-recombinogenic viruses can be engineered; Woodman *et al.* 2016) that either compensates for the inherently error-prone polymerase these viruses possess, by 'rescuing' damaged genomes, or actively contributes to the generation of quasispecies diversity. There are numerous examples where recombination leads to the development of novel phenotypes, in both plant and animal viruses (Bentley and Evans 2018). The polymerase switches templates during replication, probably in a negative-strand sequence-independent manner (Kirkegaard and Baltimore 1986; Lowry *et al.* 2014) to generate a hybrid genome, the viability of which then determines whether it is further propagated or disappears. Replication-competent recombinants may undergo or acquire further adaptive changes, the final result being a chimeric genome with breakpoints marking the junction between the two (or more) parental sequences from which it was derived. Knowing where the breakpoints are in viral recombination events can provide insight into the biochemical and mechanistic properties of viral infection (Rohayem, Munch and Rethwilm 2005; Simon-Loriere *et al.* 2009). The modular nature of the genomes of all single-stranded positive-sense RNA viruses contributes to the process of recombination, with the junctions usually located between the modules for the structural proteins, the non-structural proteins or the non-coding regions and the coding regions (Lukashev *et al.* 2003; Lukashev 2005; Lowry *et al.* 2014; Bentley and Evans 2018).

Recombination has been observed between DWV Type A and Type B strains, and in some instances these recombinant forms have been reported to predominate in a population over the strains they originate from (Moore *et al.* 2011; McMahon *et al.* 2016). Moore *et al.* (2011) identified two DWV recombinants, VDV-1$_{DVD}$ [comprising the DWV Type A 5′-UTR (untranslated region), a VDV-1 structural gene and DWV Type A non-structural genes] and VDV-1$_{VVD}$ (comprising the 5′-UTR of VDV-1, a VDV-1 structural gene and DWV Type A non-structural genes). An observed correlation between high VDV-1$_{DVD}$ titres in infected pupae and their associated mites suggested this recombinant may have been better adapted to transmission between *Varroa* and honey bees than its parental variants. McMahon *et al.* (2016) found that Type A and Type B viral variants recombined readily *in vivo*, indicating non-independence of the two viruses among foraging honey bees. In laboratory studies with wild-type (i.e. not derived by reverse genetics) purified virus preparations, a type B/A recombinant (DWV VVD) had a selective advantage over non-recombinant stocks when injected into pupae (Ryabov *et al.* 2014).

The predominance of a particular virus population within a *Varroa*-infested colony is a consequence of the original viruses present, the passage history of the population and the acquisition of exogenous viruses from neighbouring hives or, possibly, through environmental contamination. If recombinant viruses are more transmissible, either vertically or horizontally from bee to bee or via *Varroa*, it will inevitably influence the resulting virus population. This is an inherently noisy system in which to reach definitive conclusions about relative virulence of different virus strains. In contrast, cDNA-derived viruses directly inoculated to honey bee pupae in the laboratory can be relatively easily compared in virulence and replicative capacity. Under these conditions, recombinants (at least those tested to date), do not appear to out-replicate the parental viruses to which they are related (Gusachenko *et al.* 2020a).

Are recombination events sequence specific?

'Hotspots' in the viral genome are where recombination preferably occurs. One such hotspot, observed by Ryabov *et al.* (2014) and Moore *et al.* (2011), is between the structural and non-structural proteins. The most variable region of the genome, and therefore an ideal candidate region for measuring the extent of sequence diversity in a virus population, is the leader protein (a 211-codon region towards the 5′ end of the viral genome) with a nucleotide identity match of only ~74% between DWV and VDV-1 (Dalmon *et al.* 2017). Observations of recombination events occurring in the leader protein region between VDV-1 and DWV-A have been observed in recent analyses of samples in the United States (Ryabov *et al.* 2017). Additionally, hotspots were independently discovered in the 5′ end of the genome (McMahon *et al.* 2016; Cornman 2017). All these identified recombination junctions occur between functional 'modules' of the virus genome.

The Type A/B recombinant viruses observed by Ryabov *et al.* (2014) predominated in a colony when compared with the parental Type A or Type B strains that were also present. This suggests that recombination may be due to fitness selection or the result of a bottleneck event (see the section 'Changes in virus replication and mite-related bottlenecks'). Without evidence to the contrary it is likely that DWV (like other single-stranded positive-sense RNA viruses) recombines using a mechanism similar to that observed during poliovirus recombination (Kirkegaard and Baltimore 1986). This is a biphasic process in which the initial recombinant product, which may have acquired sequence duplications during recombination, undergoes a subsequent resolution process involving the loss of sequence duplication and the optimisation of virus fitness (Lowry *et al.* 2014). Multiple recombination events within a single DWV genome have been observed. For example, Dalmon *et al.* (2017) reported a recombinant DWV genome that contained a Type B 5′-UTR, a Type A leader protein sequence, the remaining structural proteins from Type B and the non-structural proteins from Type A, implying that at least three independent recombination events had occurred. It is notable that the majority of recombinant strains reported in that study derived their structural proteins from Type B sequences (Dalmon *et al.* 2017), indicating that this variant may have a selective advantage either

was not as pronounced when the recombinant was the first virus administered. Further studies are required to determine the molecular mechanism that accounts for the inability of the recombinant to compete efficiently in the presence of a genetically related pre-established virus infection. The sequence identity between the established and incoming viruses may prime the RNAi-based immune response and suppress the replication of the second virus. Studies of differences in the RNAi response, a proportion of which is directed against the capsid-coding sequences (Ryabov *et al.* 2014), will be needed to investigate this further. If supported, this could help develop an assay to identify RNAi suppressors that would be expected to be present within the DWV genome (Gammon and Mello 2015). The proposed superinfection exclusion of Type A by Type B (Mordecai *et al.* 2015a) deserves further study to determine whether it can be recapitulated with other wild-type virus stock preparations, and to determine how sensitive it is to the presence of subpopulations of recombinant viruses within the inocula.

Are recombinant forms of DWV more pathogenic?

Genetic recombination—the formation of hybrid or chimeric genomes—is an intrinsic consequence of the replication mechanism of most single-stranded positive-sense RNA viruses. Evidence currently suggests that it is an evolutionarily retained mechanism (as non-recombinogenic viruses can be engineered; Woodman *et al.* 2016) that either compensates for the inherently error-prone polymerase these viruses possess, by 'rescuing' damaged genomes, or actively contributes to the generation of quasispecies diversity. There are numerous examples where recombination leads to the development of novel phenotypes, in both plant and animal viruses (Bentley and Evans 2018). The polymerase switches templates during replication, probably in a negative-strand sequence-independent manner (Kirkegaard and Baltimore 1986; Lowry *et al.* 2014) to generate a hybrid genome, the viability of which then determines whether it is further propagated or disappears. Replication-competent recombinants may undergo or acquire further adaptive changes, the final result being a chimeric genome with breakpoints marking the junction between the two (or more) parental sequences from which it was derived. Knowing where the breakpoints are in viral recombination events can provide insight into the biochemical and mechanistic properties of viral infection (Rohayem, Munch and Rethwilm 2005; Simon-Loriere *et al.* 2009). The modular nature of the genomes of all single-stranded positive-sense RNA viruses contributes to the process of recombination, with the junctions usually located between the modules for the structural proteins, the non-structural proteins or the non-coding regions and the coding regions (Lukashev *et al.* 2003; Lukashev 2005; Lowry *et al.* 2014; Bentley and Evans 2018).

Recombination has been observed between DWV Type A and Type B strains, and in some instances these recombinant forms have been reported to predominate in a population over the strains they originate from (Moore *et al.* 2011; McMahon *et al.* 2016). Moore *et al.* (2011) identified two DWV recombinants, VDV-1$_{DVD}$ [comprising the DWV Type A 5′-UTR (untranslated region), a VDV-1 structural gene and DWV Type A non-structural genes] and VDV-1$_{VVD}$ (comprising the 5′-UTR of VDV-1, a VDV-1 structural gene and DWV Type A non-structural genes). An observed correlation between high VDV-1$_{DVD}$ titres in infected pupae and their associated mites suggested this recombinant may have been better adapted to transmission between *Varroa* and honey bees than its parental variants. McMahon *et al.* (2016) found that Type A and Type B viral variants recombined readily *in vivo*, indicating non-independence of the two viruses among foraging honey bees. In laboratory studies with wild-type (i.e. not derived by reverse genetics) purified virus preparations, a type B/A recombinant (DWV VVD) had a selective advantage over non-recombinant stocks when injected into pupae (Ryabov *et al.* 2014).

The predominance of a particular virus population within a *Varroa*-infested colony is a consequence of the original viruses present, the passage history of the population and the acquisition of exogenous viruses from neighbouring hives or, possibly, through environmental contamination. If recombinant viruses are more transmissible, either vertically or horizontally from bee to bee or via *Varroa*, it will inevitably influence the resulting virus population. This is an inherently noisy system in which to reach definitive conclusions about relative virulence of different virus strains. In contrast, cDNA-derived viruses directly inoculated to honey bee pupae in the laboratory can be relatively easily compared in virulence and replicative capacity. Under these conditions, recombinants (at least those tested to date), do not appear to out-replicate the parental viruses to which they are related (Gusachenko *et al.* 2020a).

Are recombination events sequence specific?

'Hotspots' in the viral genome are where recombination preferably occurs. One such hotspot, observed by Ryabov *et al.* (2014) and Moore *et al.* (2011), is between the structural and non-structural proteins. The most variable region of the genome, and therefore an ideal candidate region for measuring the extent of sequence diversity in a virus population, is the leader protein (a 211-codon region towards the 5′ end of the viral genome) with a nucleotide identity match of only ~74% between DWV and VDV-1 (Dalmon *et al.* 2017). Observations of recombination events occurring in the leader protein region between VDV-1 and DWV-A have been observed in recent analyses of samples in the United States (Ryabov *et al.* 2017). Additionally, hotspots were independently discovered in the 5′ end of the genome (McMahon *et al.* 2016; Cornman 2017). All these identified recombination junctions occur between functional 'modules' of the virus genome.

The Type A/B recombinant viruses observed by Ryabov *et al.* (2014) predominated in a colony when compared with the parental Type A or Type B strains that were also present. This suggests that recombination may be due to fitness selection or the result of a bottleneck event (see the section 'Changes in virus replication and mite-related bottlenecks'). Without evidence to the contrary it is likely that DWV (like other single-stranded positive-sense RNA viruses) recombines using a mechanism similar to that observed during poliovirus recombination (Kirkegaard and Baltimore 1986). This is a biphasic process in which the initial recombinant product, which may have acquired sequence duplications during recombination, undergoes a subsequent resolution process involving the loss of sequence duplication and the optimisation of virus fitness (Lowry *et al.* 2014). Multiple recombination events within a single DWV genome have been observed. For example, Dalmon *et al.* (2017) reported a recombinant DWV genome that contained a Type B 5′-UTR, a Type A leader protein sequence, the remaining structural proteins from Type B and the non-structural proteins from Type A, implying that at least three independent recombination events had occurred. It is notable that the majority of recombinant strains reported in that study derived their structural proteins from Type B sequences (Dalmon *et al.* 2017), indicating that this variant may have a selective advantage either

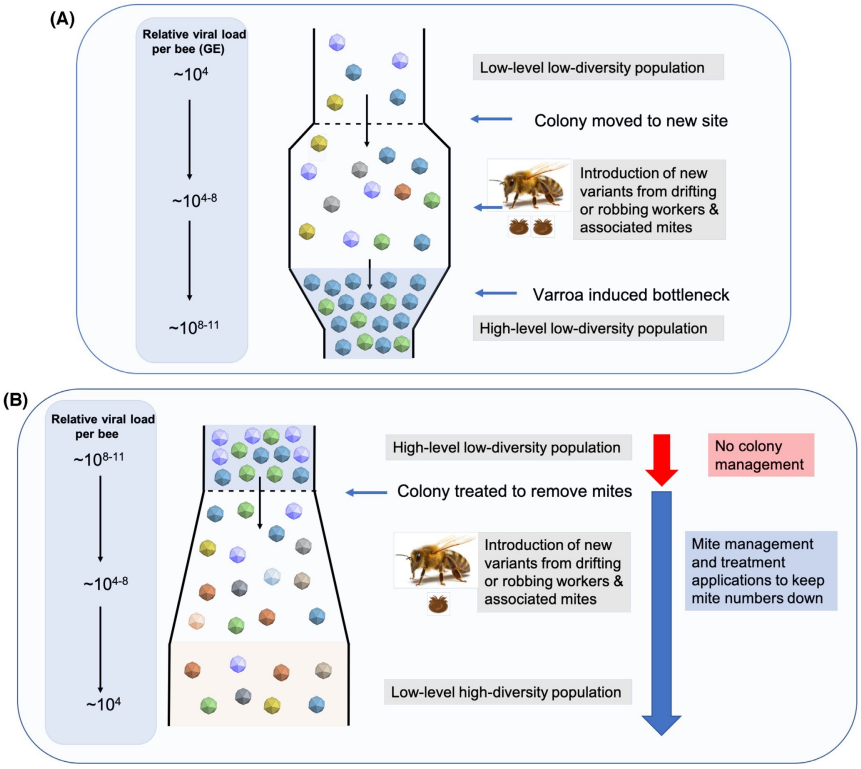


Figure 3. Changes to the DWV population in managed honey bee colonies. **(A)** The effect on the virus population when an otherwise healthy honey bee colony is placed in close proximity to infested hives. Robbing and drifting bees bring phoretic mites that alter the virus population over time. **(B)** The effects of colony management and mite removal on the virus population. The dashed lines indicate 'hard' changes to the colony, e.g. movement or treatment application.

titre observed. Ryabov *et al.* (2019) evaluated DWV infections in honey bee populations in the United States and demonstrated similar levels of population diversity in bees with high- or low-level viral titres. The introduction of *Varroa*-free colonies with low asymptomatic DWV populations to mite-infested apiaries results in a rapid infestation of colonies and a shift in viral load and in some instances diversity, but after several months of infestation some viral diversity is still observed in highly infected workers (Martin *et al.* 2012; Woodford in preparation). Infected neighbouring colonies continue to display viral diversity at high infection levels despite 5+ years of mite infestation and no mite management techniques.

The diversity in the virus population can be viewed at the level of the individual honey bee, the colony or at the landscape scale. It is clear from numerous studies that the introduction of *Varroa destructor* to honey bees has fundamentally altered the DWV population at all levels, in the same way that the amplification of a very small initial inocula from *Varroa* can be considered a bottleneck event that can influence the virus population in the recipient pupa. The removal of *Varroa* from a population could also be considered as a potential bottleneck influencing the resulting DWV population (Fig. 3B).

Used properly, hive management techniques and chemical miticides can significantly reduce the mite population within a colony. Under field conditions reductions of 95% are achievable (Rinkevich 2020), and are typically applied at the end of the summer season and in midwinter. Since only a subset of mites survive these interventions, the virus population they harbour will have been similarly restricted as host mites perish. As a consequence, this represents a potential bottleneck that will influence the virus population that survive and are amplified in the colony the following season (Fig. 3B). There are only limited studies of changes in the virus population following miticide treatments. Locke *et al.* (2017) demonstrated 1000-fold reductions in DWV levels within the colony after miticide treatment, but did not investigate changes in the population diversity or genotypes that predominated before and after treatment. In recent studies, involving a combination of brood removal and the timely application of miticides, we have demonstrated a sustained reduction of virus levels in newly eclosed individual workers of at least 10^5 GE/μg RNA (Woodford in preparation) and are currently characterising the virus population diversity pre- and post-treatment.

To better understand the dynamics of the DWV population and the influence *Varroa* has on it we need to know the relative virulence of different types of DWV and investigate whether all

variants of a particular type have a similar virulence. Simplistic modelling of the DWV population demonstrates gross changes in virus diversity, recapitulating that seen in an infested colony, can be accomplished within a few generations solely resulting from very limited viral inocula that rapidly amplify in naïve pupae (Woodford & Evans unpublished). In these models the reduction in viral diversity is achieved without any assumptions about differential virulence of variants within the virus population. This has also been examined experimentally where serial passaging of relatively high titres of DWV-A (2μl of the original 10^7 DWV/bee stock) across multiple generations of pupae simulated the effect of *Varroa* transmission (Yañez *et al.* 2020a) and showed that by the time the workers reached eclosion similar variants would dominate the population regardless of the passage number or whether the infection was symptomatic or asymptomatic. This was despite an initial increase in diversity post-injection and indicates that the selective pressure that favours certain variants may be occurring in a dose-dependent manner (Yañez *et al.* 2020a). In HIV infection bottlenecks, compared with clonal infections, mixed variant infections reduce the host response and subsequently undergo a less stringent transmission bottleneck that allows less fit variants to establish in the population (Macharia *et al.* 2020). A similar effect may be observed in mixed populations of DWV. To properly examine the relative virulence of dominant genotypes and mixed populations, reverse genetics methods can be utilised to build recombinants and to replicate mixed virus populations observed during natural infections.

An additional variable that needs to be taken into account in these *in silico* studies is the ability of DWV to replicate in *Varroa*. If it does, and particularly if only certain strains of DWV can replicate, it may significantly influence the subsequent viral inocula and the resulting virus variants that are amplified following transmission. Very recent studies, including some using reverse genetics, are addressing the replication of DWV in *Varroa* as described in the following section.

DWV replication and tropism in *Varroa destructor*

A factor that may influence the development of high virus levels in developing honey bees, or could bias changes in the virus population diversity, is replication of the virus in the ectoparasite *Varroa destructor*. If this occurs, or if only certain variants replicate in the mite, there may be preferential amplification of a subset of the virus population prior to transmission. It was proposed that the virus taken up by *Varroa* during feeding replicates in the vector and thus when it is passed to the next developing honey bee the virus is transmitted at a higher viral titre, thus aiding its ability to reach levels that induce a symptomatic infection in the host (Yue and Genersch 2005). There have been limited and somewhat contradictory studies of the virus population in *Varroa* mites. VDV-1 (DWV Type B) was originally isolated from *Varroa* and high levels of virus were detected in mites that were associated with pupae (Ongus *et al.* 2004). In addition, negative strand DWV RNA—assumed to be an indicator of replicating virus—was detected in mites (Yue and Genersch 2005; Gisder, Aumeier and Genersch 2009) without formally demonstrating it had been generated in the mite. Ryabov *et al.* (2014) demonstrated there was a high diversity of viruses present in the mite but that this diversity was not representative of the near-clonal virus population present in the mite-exposed pupae from the same colony. This, and earlier work on the same virus population (Moore *et al.* 2011), suggested that mites may form a selection bottleneck responsible for the limited diversity in subsequently parasitised pupae. However, these studies did not address whether the virus replicated in the mite and, formally, a selection bottleneck does not need to involve replication *per se* in the mite. In contrast to these studies in which replication in the mite was assumed, Erban *et al.* (2015) investigated the presence of proteins associated with virus replication in the mite. Using mass spectroscopy, the authors failed to detect the non-structural proteins of DWV that are essential for virus replication. By definition, these non-structural proteins are not components of the DWV virus particle, but are only present in cells in which the virus replicates. The conclusion from this study was that DWV therefore does not replicate in *Varroa*.

Two recent studies have applied reverse genetics approaches to try and resolve whether DWV replicates in mites. The advantage of using a reverse genetic approach is that the identity of the input virus is defined by its sequence or the presence of unique genetic tags engineered into the genome, allowing unequivocal discrimination of the virus from the pre-existing DWV population. Despite the application of reverse genetic with the concomitant ability to track viruses within a population, these two studies still appear to produce contradictory results. Posada-Florez *et al.* (2019) serially passaged mites on bees initially injected with a reverse genetics-derived DWV Type A. They demonstrated virus accumulation in mites and subsequent transmission of the genetically distinct virus by the mite. However, they also attributed the presence of negative strand RNA in the mites to the ingestion of cells from the pupae within which DWV was or had replicated, concluding that DWV Type A is transmitted by *Varroa* in a non-propagative manner (Posada-Florez *et al.* 2019).

In the second study, a stock of purified DWV derived from a genetically tagged recombinant DWV (VVD) was prepared following injection of pupae *in vitro*. During purification this stock was filtered to remove cellular material and treated with ribonucleases to degrade any residual unencapsidated RNA. The absence of negative strand RNA was confirmed using a sensitive strand-specific PCR assay. It is well established that the picornalike viruses only package positive strands of the viral genome (Schwartz *et al.* 2002), and there is no reason to suspect that DWV will not be equally selective in this regard. *Varroa* were maintained on an artificial feed-packet system *in vitro* that contained no honey bee-derived components (Christie *et al.* in preparation). All subsequently screened mites contained negative strand RNA and a proportion of this could be digested with the restriction endonuclease specific for the unique genetic tag engineered into the recombinant DWV genome, indicating viral replication was taking place in the mites (Gusachenko *et al.* 2020a).

This study neatly demonstrates the power of a reverse genetic strategy for the analysis of the biology of DWV. It exploits the ease with which highly purified high titre stocks can be prepared of a virus containing unique genetic markers that allow it to be detected, even in the presence of excess amounts of endogenous virus. Based on these findings it can be assumed that viral replication in mites is at least partially responsible for the highly elevated viral loads of some variants found in parasitised pupae and a cause of virus population bottlenecks (Fig. 3). Further studies, using similar strategies, are required to determine whether the reported differences between Type A and Type B viruses are still observed. If they are, the availability of a range of molecular clones of Type A/Type B recombinants will facilitate determining the molecular basis for the phenotypic differences.

Where in the mite DWV replication occurs and how this influences the transmissibility of the virus to a new host remain

unanswered. This is known to be a significant influence in certain arboviruses; for example, a major bottleneck in dengue virus infections is observed in the mosquito midgut, where as few as 5–42 founder viruses can initiate infection (Lequime *et al.* 2016). If there are strain-specific differences in the ability of DWV to replicate in *Varroa*, there are likely to be subtle (if amplification is limited) to profound (if certain strains replicate very well) influences on the DWV variants passed to the next pupa. Furthermore, such differences may inform our understanding of changes of the DWV population at the landscape scale (Wilfert *et al.* 2016).

Does DWV diversity matter for pathogenesis?

If individual variants (Type A, Type B and their recombinants) of DWV are equally pathogenic (see the section '*Deformed wing virus*') does the diversity within the population of a specific variant influence virus pathogenesis? This question is relevant both in terms of understanding the biology of DWV and in determining the significance of studies using virus recovered by reverse genetics. Our virus quantification analysis suggests that cDNA-recovered viruses are near-clonal (>99.99% representative of the cDNA sequence) when recovered in honey bee pupae with a low level (~10^3–10^4 GE/μg) of endogenous DWV, and clonal when recovered in *Bombus sp.* (see below) in which DWV is not naturally present (Gusachenko *et al.* 2020b).

This question is also relevant as population diversity is known to influence virus fitness in other viruses, and has been demonstrated to influence interhost transmission and differ according to the tissue the virus replicates in. Viral fitness, for example defined by the ability of poliovirus to cross the blood-brain barrier, is reduced in low diversity populations (Pfeiffer and Kirkegaard 2006; Xiao *et al.* 2017). Although this is of little relevance for poliovirus transmission within the population it has a fundamental influence on poliovirus pathogenesis. More broadly, it probably reflects the ability of the virus to overcome innate barriers to spread within and between tissues. Related studies on Chikungunya virus demonstrate that the passage history between host and vector influences population diversity and could increase diversity and adaptability to novel selection pressures (Coffey and Vignuzzi 2011), such as the virus might encounter in a new host or when transmitted via a different route. Once spread within the host, the virus can evolve separately within tissues or organs, as has been demonstrated for HIV where different populations of virus are present in the blood and brain of infected individuals (Pillai *et al.* 2006; Schnell *et al.* 2009; Poirier and Vignuzzi 2017). This, in turn, may influence the ability of the virus to escape those tissues or to transmit to a new host. These examples are of virus evolution in vertebrate hosts with a B- and T-cell mediated acquired immune system. However, in studies of West Nile virus in the mosquito, it has been reported that the RNAi-mediated immune response of insects is also a driver of virus variation (Brackney *et al.* 2015) and it remains to be determined whether the honey bee RNAi response to DWV may function in a similar manner.

Unfortunately, studies on the influence of DWV population diversity are in their infancy and are again hampered by the absence of a suitable cell culture system and the ubiquitous nature of DWV in the honey bee population. However, the progress made in other host-vector-virus systems give a good indication of what might be expected, and of the types of studies that will be needed to understand the specifics for DWV. Analysis of *Ixodes* sp. tick-transmitted Powassan virus in mice has demonstrated the influence of genetic bottlenecks due to the amount of virus transmitted and how purifying selection in ticks counteracts the role of host RNAi in driving virus variation (Grubaugh *et al.* 2016a). The kinetics of this series of events is markedly different from the evolution of mosquito-transmitted viruses such as West Nile Virus (WNV), reflecting the different pattern and frequency of vector feeding on the host through the season. Powassan virus diversity is greater in the host than the tick vector, but in the mosquito/bird/WNV system diversity expands in the vector and is reduced in the vertebrate host (Grubaugh *et al.* 2016b).

Whether DWV fits either of these two contrasting patterns or has its own distinct characteristics have yet to be determined. Recent studies has demonstrated that mite feeding, with consequent virus transmission, does not result in the complete loss of DWV diversity (Annoscia *et al.* 2019; Ryabov *et al.* 2019). This is perhaps surprising considering the striking amplification (10^6-fold) of very limited amounts of virus in the 24–48 h post-inoculation (Gusachenko *et al.* 2020a). This may reflect the importance of the direct bee to bee transmission routes during trophallaxis and feeding, or vertically in the population. Alternatively, it may indicate that *in vitro* inoculation of DWV in the laboratory—either in these particular studies or more generally—does not properly recapitulate the events following mite transmission to a naïve pupa. Other studies have produced conflicting results. In early studies of changes in population diversity and viral load before and after *in vitro* virus transmission, there was a striking switch from a low-level/high-diversity population to a high-level/low-diversity population (Wood *et al.* 2014). In this study, the diversity in the amplified virus population after injection was at or near to the lower-limit of detection of the next-generation sequencing strategy used to quantify variation. In contrast, although not an equivalent study as it was based upon analysing viral load and diversity in individual field samples, there was no apparent restriction in virus diversity in isolates with high viral loads compared with those with low levels of virus in samples from the United States between 2015 and 2017 (Ryabov *et al.* 2019).

It is possible that recent reports of high viral diversity in highly infected workers (Annoscia *et al.* 2019; Ryabov *et al.* 2019) are a consequence of the viruses evading the host immune responses. Some signalling pathways, such as Toll in *Drosophila*, have been associated with controlling RNA-viral infections (Merkling and van Rij 2013), and changes in Toll pathways are seen in DWV-infected honey bee pupae (Ryabov *et al.* 2014). If these are also antiviral in honey bees there are probably virus variants able to subvert or avoid the response. Likewise, the expression of RNAi has been associated with controlling the persistence of viral-RNA infections in insects, such as *Drosophila* or mosquitoes (Wang *et al.* 2006; Goic *et al.* 2013). RNAi suppressors are detected in a variety of different types of insect and plant viruses where host RNAi responses are a predominant protection mechanism. It would be expected that they also occur in DWV. Understanding the complex interaction of potential host-generated diversity, viral subversion of immune responses and the influence of multiple transmission bottlenecks in both host and vector, is unlikely to be straightforward. Strategies using *in vitro* generated populations of DWV with varying levels of diversity, for example following synthesis using high- and low-fidelity T7 polymerases, or by direct modification of the DWV polymerase to recapitulate the 'HiFi' and 'LoFi' polymerase mutations created in other picornaviruses (Woodman *et al.* 2016), offer exciting and challenging approaches to investigate this intriguing aspect of the biology of DWV.

PATHOGEN SPILLOVER—IS DWV AN ENVIRONMENTAL THREAT?

As honey bees exist in complex ecosystems and may interact frequently with other pollinating insects—e.g. during shared flower visits or attempted colony robbing—there is a risk that pathogens could transmit from one host species to another. This is essentially what happened when *Varroa destructor* crossed from *Apis cerana* to *A. mellifera*. The subsequent devastating consequences for the global honey bee population emphasise the potential impact and damage such interspecies spillover events can have, and is repeatedly seen with viruses, e.g. HIV transmission from chimpanzees to humans (Sharp and Hahn 2011), and Ebola and SARS-type coronaviruses from bats to humans (Li *et al.* 2005). As mentioned above, the introduction of *Varroa* has resulted in changes to both the level and variants of DWV circulating in the honey bee population. It therefore now poses a potential risk to insects that inhabit the same or overlapping environmental niches as honey bees (Manley *et al.* 2015). Depending upon the geographic area this could include other *Apis* species, some free-living colonial (e.g. *Bombus* sp.) and solitary bees, other Hymenoptera such as wasps and ants and even other pollinators such as Lepidopteran butterflies and moths. As a consequence there have been several surveys of honey bee pathogens in these and other species (see Martin and Brettell 2019). For the purpose of this overview, and to highlight the contribution reverse genetics can make to understanding the relevance of these studies, we will only consider suspected interspecies transmission ('spillover') of DWV and highlight how reverse genetics may help discriminate between correlation and causation in studies of disease attributed to DWV in non-*Apis* species.

Although initially reported as being present at low prevalence (Li *et al.* 2012), recent studies have indicated that over 80% of *Apis cerana* samples tested in China are DWV positive (Hassanyar *et al.* 2019). Considering that *A. cerana* and *A. mellifera* are often both managed in similar environments this is perhaps not surprising. However, wild-living *Apis* species including isolates of *A. florea* and *A. dorsata* have also been reported with a DWV prevalence of 11% and 15%, respectively (Zhang *et al.* 2012). In addition, DWV has also been identified in several solitary bee species, including *Osmia bicornis*, *Osmia cornuta* and several *Andrena* sp. (Radzevic˘iu˙te˙ *et al.* 2017), although of the *Andrena* sp. that tested positive for DWV, only a single *Andrena haemorrhoa* contained the DWV negative-sense RNA, indicative of viral replication.

Some of the most extensively surveyed species are wild and managed *Bombus* sp. 'bumble bees'. In particular, *B. terrestris* is widespread in the environment, easy to survey, and is the only other routinely commercially reared pollinating insect. Since some *B. terrestris* rearing methods may involve use of callow honey bee workers to stimulate brood production (Velthuis and Van Doorn 2006), there may also be ample opportunities for pathogen transmission during commercial rearing, and following subsequent release to the environment. Numerous studies have identified the presence of DWV in *Bombus* sp., including *B. terrestris* (Genersch *et al.* 2006; Fu¨rst *et al.* 2014; Manley *et al.* 2015; McMahon *et al.* 2015; Manley *et al.* 2019). In some of these studies an association was made between the presence of DWV and DWV-like symptoms in bumble bees (Genersch *et al.* 2006). In others, a correlation between *Varroa* presence in honey bee colonies and the increased prevalence of DWV in neighbouring wild bumble bee populations was observed (Manley *et al.* 2019).

In many of these studies the sampling involves collecting *Bombus* sp. from flowers and as such these may represent the healthier proportion of the population and those with higher viral titres are not sampled. It should be noted that there is no evidence that *Varroa* parasitise any species other than *Apis mellifera* and *A. cerana*.

In reports that DWV does replicate in bumble bees the supporting evidence is often correlative and lacks evidence of causation if symptoms are present. For example, in the case of DWV-like symptoms in *Bombus terrestris*, the virus was detected by PCR for the positive-sense genomic strand only (Genersch *et al.* 2006). Without negative-sense RNA detection or the re-inoculation of purified isolated virus to naı¨ve *Bombus terrestris*, evidence for DWV replication that is a requirement for causing disease, is not possible. In other studies, negative strand assays are inconclusive. For example, Fu¨rst *et al.* (2014) reported DWV prevalence of 11% in *Bombus* sp. of which 38% contained negative strands with no report of symptomatic infections. It is possible that the low percentage of *Bombus* samples with replicating virus indicated an assay operating near the limit of detection, that virus may only replicate in some (ages, sexes or strains of) bumble bees, or that there may be alternative explanations unrelated to DWV replication.

Even when negative strands of the virus are present, there remains the possibility that it was present in the inocula or the last meal eaten by the insect species rather than being synthesised *de novo*. This is exemplified in recent studies of ant populations feeding on the corpses of DWV-positive honey bees (Schla¨ppi *et al.* 2019). In these studies, negative-sense strands of the DWV were detected in ants after feeding, but their detection subsequently decreased over time. This strongly suggested that the DWV had been ingested, but that the virus was unable to replicate or sustain infection and was eventually lost.

To address DWV replication in *Bombus* sp. directly, genetically tagged DWV was used to infect bumble bees using a broadly similar strategy to that used to study the replication of DWV in *Varroa* (Gusachenko *et al.* 2020a). Fed or injected DWV preparations containing unique tags were used to inoculate individual adult *B. terrestris* workers, pupae or larvae maintained *in vitro*. Bumble bee pupae inoculated with 10^3 or 10^6 GE of DWV showed 100–1000-fold increases in viral load after 48 h, indicative of viral replication, but no overt 'deformed wings' in subsequently eclosed workers (Gusachenko *et al.*2020b). First and second instar larvae were fed a diet containing 10^8 GE of DWV and although the viral yields were similar to that in the fed diet, negative strand RNA was detected in 50% of samples 5 days after inoculation. In further studies, adult workers were injected with 10^4 or 10^8 GE of DWV to determine if viral infections could be established by this route. Of the injected workers 40% and 100% of samples tested positive by RT-PCR, respectively, and negative strand RNA was detected, indicative of virus replication. Since the *B. terrestris* colonies used contained no detectable endogenous DWV it is possible to use this strategy to prepare and obtain concentrated stocks of pure clonal DWV for further experiments with *Bombus*, *Apis* or other pollinators. In a complimentary study using well-characterised virus inocula of Type A and Type B DWV it was shown that injected and fed virus replicates in adult bumble bees when administered at high doses (10^7 and 10^9 GE, respectively), but does not negatively impact worker mortality (Tehel *et al.* 2020).

As *Varroa* do not parasitise bumble bees, virus transmission to larvae is only possible by feeding. Individual adult bees were fed DWV-laced (10^8 GE/adult bee) syrup but exhibited no evi-

dence of infection, with detected virus titres 100–200-fold lower 7 days after feeding (Gusachenko *et al.* 2020b). If DWV does not replicate in fed adults, how are cases of DWV-positive field samples possible? One possible explanation is that the virus is transmitted to larvae by adults when being fed in the nest. To test this, continuous feeding of *Bombus terrestris* colonies with a diet supplemented with 2×10^8 GE/adult bee/day of 3 DWV variants was performed over a 4–6-week period, followed by sensitive PCR-based screening of all larvae, pupae and adults in the nest. However, despite feeding significantly more DWV than is reported in environmental contamination studies (Mazzei *et al.* 2014), there was no evidence of DWV infection and replication in fed *B. terrestris* colonies (Gusachenko *et al.* 2020b). This possibly indicates that there is threshold of susceptibility for successful DWV infections in fed larvae.

Together, these results strongly suggest that adult bumble bees, at least *B. terrestris*, are not susceptible to infection through feeding and are unable to transmit infectious virus when feeding developing larvae. However, *Bombus* are susceptible to DWV via injections in adults and pupae and high dose feeding in larvae, though none of these routes appear to cause symptomatic infections *in vitro*. These studies highlight the importance of formally identifying the negative-strand intermediates in virus replication—a relatively straightforward task using genetically tagged input viruses—and testing whether Koch's postulates are fulfilled when interpreting data on the occurrence and consequences of environmental pathogen spillover. If *Bombus* sp. are naturally infected and replicate DWV, the route by which they acquire the virus remains unclear.

These findings, coupled with reports of DWV presence in a host of different important pollinator species, highlight the need for robust and reproducible methods for virus detection. The advent of sensitive molecular methods that are capable of quantifying intermediate products in the replication cycle of viruses will allow a re-evaluation of the actual risks to the environment due to pathogen spillover from managed honey bee populations. By using purified, RNAse-treated, genetically tagged input virus preparations it is possible to quickly and sensitively discriminate replicating virus from even high levels of background contaminating non-replicating virus. Further studies of this type will help understand the likely occurrence and consequences of pathogen spillover from managed honey bees to a diverse range of other pollinating species that coexist in the environment.

CONCLUSION

Over the last four decades, our understanding of the molecular biology of single-stranded positive-sense RNA viruses has increased rapidly. This largely reflects three things: their importance as agents of human and animal disease, the development of reverse genetic systems enabling the dissection of genotype and phenotype, and the availability of immortalised cell lines that allow pure cultures of the virus to be propagated *in vitro*.

In the case of DWV, which, when transmitted by its vector the ectoparasitic mite *Varroa destructor*, is the most significant pathogen to threaten honey bee health globally, progress has been more limited. Despite being discovered almost four decades ago, the near ubiquitous presence of the virus in the honey bee population and the lack of a suitable cell culture system for propagation *in vitro* have hampered progress. The recent development of reverse genetic systems for DWV will go some way to overcoming these limitations.

Although these studies are still in their infancy, it is already clear that reverse genetic approaches allow discrimination between endogenous virus and virus recovered from cDNA. That the latter can be recovered in pure culture from bumble bees (*Bombus terrestris*) neatly avoids contamination of virus stocks with low levels of endogenous DWV. The benefits arising from manipulation of the virus genome include the incorporation of unique genetic tags to distinguish the input virus from anything already present and, in very recent studies, the ability to localise replicating virus *in vivo* by inclusion of fluorescent reporter genes.

Whilst these studies demonstrate that encouraging progress is possible, there remain a number of questions and contradictions in the literature that will need to be resolved. Many of these have been introduced in the preceding text—the influence of virus type and strain variation on replication and pathogenesis, the ability of the virus to replicate in *Varroa destructor* and in other, non-*Apis*, pollinators. Our recent studies have provided evidence that DWV replicates in *Bombus terrestris*. However, confoundingly, we see no evidence for overt pathogenesis, or a mechanism by which infection could occur through natural exposure. Further studies will be needed to investigate this interaction, and to determine whether the numerous other *Apis* and non-*Apis* pollinators previously reported to be DWV-positive actually support replication of the virus or simply reflect environmental contamination due to the ubiquitous nature of DWV in the honey bee population.

Numerous questions remain about pathogenesis of DWV in *Apis mellifera*. Up to 75% of experimentally inoculated eclosing workers exhibit overt deformed wings. The remainder appear phenotypically normal, despite having the same high level of DWV. The ability to produce high titre stocks of near-clonal genetically tagged viruses allows a number of questions about pathogenesis to be readily addressed: do all strains of honey bees exhibit a similar division of overtly diseased/normal workers, or are some strains more or less susceptible? Do physiologically distinct bees, such as workers produced mid-season and the winter bees produced in early autumn, show differential susceptibility? More fundamentally, what is different about the workers that emerge without overt symptoms? Is the virus population in these bees different? Does the distribution of the virus within the bee differ, and does this reflect variation in the host immune response, potentially protecting key organs or tissues? Finally, how important is overt disease anyway? Perhaps these are just the most susceptible bees in the population. If very high levels of DWV always reduce the longevity of overwintering worker bees—the primary cause of winter colony losses—then the influence of strain variation, virus diversity and genotype on this characteristic is of paramount importance.

Answers to these questions will not directly or immediately improve honey bee health. They will fill in some of the numerous gaps we have in our understanding of the honey bee/virus/*Varroa* interaction, and in doing so may provide important insights to improve current treatment or management methods, or help define the markers required for selective breeding strategies. They will additionally inform our understanding of the real threat honey bees pose as a source of environmental pathogen spillover. The recent development of tractable reverse genetic systems, if applied appropriately, will go a long way to understanding the most important global threat to honey bee health, and establish the experimental framework in which new and emerging threats to their health are studied.

8.8 Appendix 8 - Published manuscript – Gusachenko *et al.*, (2021), ISME.

ISME

www.nature.com/ismej

ARTICLE OPEN

Check for updates

First come, first served: superinfection exclusion in Deformed wing virus is dependent upon sequence identity and not the order of virus acquisition

Olesya N. Gusachenko[1], Luke Woodford[1], Katharin Balbirnie-Cumming[1] and David J. Evans[1]

Deformed wing virus (DWV) is the most important globally distributed pathogen of honey bees and, when vectored by the ectoparasite *Varroa destructor*, is associated with high levels of colony losses. Divergent DWV types may differ in their pathogenicity and are reported to exhibit superinfection exclusion upon sequential infections, an inevitability in a *Varroa*-infested colony. We used a reverse genetic approach to investigate competition and interactions between genetically distinct or related virus strains, analysing viral load over time, tissue distribution with reporter gene-expressing viruses and recombination between virus variants. Transient competition occurred irrespective of the order of virus acquisition, indicating no directionality or dominance. Over longer periods, the ability to compete with a pre-existing infection correlated with the genetic divergence of the inoculae. Genetic recombination was observed throughout the DWV genome with recombinants accounting for ~2% of the population as determined by deep sequencing. We propose that superinfection exclusion, if it occurs at all, is a consequence of a cross-reactive RNAi response to the viruses involved, explaining the lack of dominance of one virus type over another. A better understanding of the consequences of dual- and superinfection will inform development of cross-protective honey bee vaccines and landscape-scale DWV transmission and evolution.

The ISME Journal; https://doi.org/10.1038/s41396-021-01043-4

INTRODUCTION

Honey bees (*Apis mellifera*) are globally important pollinators of wild flowers and agricultural crops, and the source of honey, with annual global production worth in excess of $7bn [1]. Both honey production and pollination services require strong, healthy colonies, which are threatened by a range of factors, but most significantly by disease. One of the major viral pathogens of honey bees is Deformed wing virus (DWV). When transmitted by the parasitic mite *Varroa destructor*, DWV is responsible for high overwinter colony losses, which can exceed 37% annually [2]. Improvements to honey bee health, through direct control of virus transmission or replication, require a better understanding of how the virus propagates within and between bees.

The historical identification and naming of DWV-like viruses imply a greater genetic divergence than subsequent molecular analysis has demonstrated. In 2004–2006 several picorna-like viruses with high levels of sequence identity were reported [3–5]. These viruses were initially named according to their origins; the virus from honey bees with characteristic wing deformities was termed DWV [4], a similar virus found in aggressive workers in Japan was designated Kakugo virus [3, 5] and analysis of *Varroa* mites yielded Varroa destructor virus type 1 (VDV-1) [3]. Limited genetic divergence (~84–97% genomic RNA identity), similar infectivity in honey bees, and demonstrated ability to freely recombine during coinfections [6–9] resulted in them now being considered as different variants of DWV [6, 7, 10], albeit occupying two genetic branches (VDV-1-like and DWV-like) of the same phylogenetic tree [11]. To distinguish between these branches the terminology 'type A' and 'type B' has been adopted for DWV-like and VDV-1-like variants respectively. Evidence for the existence of a third type named DWV type C has also been reported [12].

DWV is ubiquitous in honey bees [13–15], with the possible exception of Australian colonies [16]. In the absence of *Varroa* the virus is transmitted horizontally, *per os*, and vertically from the infected queen and the drones [17]. With subsequent *Varroa* mite transmission it is therefore inevitable that the virus enters a host already harbouring one or multiple DWV variants. Current studies suggest that DWV infection can occur with several variants cocirculating in the same apiary, colony, or individual honey bee host [18–21]. Although the type A and B variants appear to be differentially distributed, with type A frequently reported in the US and type B being commonly detected in European colonies [8, 13, 22], direct competition may occur where they cocirculate. If this competition has a directionality, it will influence the distribution and future spread of DWV at the landscape scale. While some studies of mixed DWV infections demonstrate no predominance of one variant over another [18, 23], others show possible competition between the variants and higher accumulation of DWV B in infected bees [24]. In addition, superinfection exclusion (SIE) has been proposed, in which a pre-existing type B virus prevents the establishment of a type A infection at the colony level [10].

[1]Biomedical Sciences Research Complex, University of St. Andrews, North Haugh, St. Andrews, UK. email: olesya.gusachenko@gmail.com

Received: 22 March 2021 Revised: 14 June 2021 Accepted: 15 June 2021
Published online: 30 June 2021

A recently developed reverse genetics (RG) system comprising a set of genetically tagged DWV variants and reporter gene-expressing viruses provides an opportunity to investigate coinfection kinetics and competition between DWV types [25]. Since SIE is a widely observed virological phenomenon [26–35], we extended these studies to assay dominance of one variant over another during sequential infection. Using reporter gene-expressing DWV we additionally investigated the influence of competition on tissue distribution of infection. We show that where competition is observed, manifested as reduced virus levels, it is reflected in reduced reporter gene expression at the cellular level. Notably we show that DWV accumulation during superinfection is influenced by the genetic identity between the viruses, rather than by a directionality of competition. Genetically divergent DWV variants (such as those representing type A and type B) exhibit transient competition, whilst viruses with greater identity (e.g. type A/B recombinants with either type A or type B) demonstrate a distinctly more pronounced effect. We also analysed the occurrence and identity of recombinants during mixed infections and confirmed that these are present with junctions widely distributed throughout the genome. These studies provide further insights into the biology of DWV. In particular they address the consequences of co- and superinfection, an important consideration when transmitted by the ectoparasite *Varroa*. Our results indicate that genome identity is the determinant that defines the outcome of dual infections; this will inform studies of population transmission at the landscape scale and possible future developments of 'vaccines' to protect honey bees from viral disease [36].

MATERIALS AND METHODS

RG DWV clones preparation

VDD, VVD and VVV RG constructs used in this study were described earlier [25], DDD RG cDNA was prepared by modification of the VDD RG system with a DWV type A parental sequence insert, which was based on published data [37] and obtained by custom gene synthesis (IDT, Leuven, Belgium). Enhanced green fluorescent protein (EGFP) and mCherry-expressing chimeric DWV genomes were built via incorporation of the reporter-encoding sequence into DWV cDNA as described previously [25]. All plasmid sequences were verified by Sanger sequencing. cDNA sequences of DDD, VVV_{mC}, and VVD_{S} are shown in Text S1, other RG cDNAs are available online (GenBank accession numbers: DWV-VDD - MT415949, DWV-VVD (VVD_{H}) - MT415950, DWV-VVV - MT415952, DWV-VDD-eGFP - MT415948, DWV-VVD-eGFP - MT415953).

Viral RNA and siRNA synthesis

DWV RNA was synthesised from linearised plasmid templates with T7 RiboMAX Express Large Scale RNA Production System (Promega, Southampton, UK), and purified with the GeneJet RNA Purification Kit (Thermo Fisher Scientific) as described in [25].

siRNA strands were prepared using Express Large Scale RNA Production System (Promega) according to the manufacturer's protocol with double stranded DNA templates annealed from synthetic oligonucleotide pairs containing T7 RNA polymerase promoter sequence (Table S1).

Viruses

Infectious DWV was prepared from honey bee pupae injected with in vitro generated RNA as previously described [25, 38]. For quantification RNA was extracted from 100 µl of virus preparation using the RNeasy kit (Qiagen, Manchester, UK) and analysed by reverse transcription and quantitative PCR (qPCR).

Honey bees and bumble bees

All honey bee (*Apis mellifera*) brood in this study was obtained from the University of St Andrews research apiary. Colonies were managed to reduce *Varroa* levels and endogenous DWV levels were regularly tested. Honey bee larvae and both honey and bumble bee pupae (*Bombus terrestris audax*, Biobest, Belgium) were maintained and fed as described previously [25, 38].

Virus inoculations

Virus injections of pupae were performed with insulin syringes (BD Micro Fine Plus, 1 ml, 30 G, Becton Dickinson, Oxford, UK) as described in [25, 38]. Oral larval infection was carried out by single DWV feeding according to the previously described procedure [25].

RNA extraction, reverse transcription and PCR (RT-PCR)

RT-PCR and qPCR analysis of individual pupae samples was performed as previously described [25]. Sequences of primers are shown in Table S1. When required, PCR products were subjected to restriction digest prior to loading on the 1% agarose gels stained with ethidium bromide. DWV titres were calculated by relating the resulting Ct value to the standard curve generated from a serial dilution of the cDNA obtained from the viral RNA used for virus stock preparation.

Microscopy

Imaging was conducted using a Leica TCS SP8 confocal microscope with 10 × HC PL FLUOTAR objective. For dissected pupae analysis samples were mounted in a drop of PBS under the microscope cover slides and observed by microscopy within 1 h after the dissection.

Sample libraries for next generation sequencing

RNA was reverse transcribed using Superscript III polymerase (Invitrogen, Thermo Fisher Scientific) with DWV FG RP1 primer (Table S1) using 1 µg of total RNA in a 20 µl final reaction volume and following the manufacturer's protocol. Reactions were incubated at 50 °C for 1 h, 75 °C for 15 min.

The transcribed cDNA was amplified using LongAmp Taq polymerase (New England Biolabs) to produce a ~10 Kb PCR fragment. The reactions were carried out according to the manufacturer's protocol with the following thermal profile: 30 s at 95 °C, 30 cycles of 95 °C for 15 s, 53 °C for 30 s, and 65 °C for 8 min, with a final extension at 65 °C for 10 min.

Recombination analysis

Purified amplicons were sequenced using an HiSeq (Illumina, Cabmridge, UK) at the University of St Andrews, producing 2 × 300 bp paired-end reads. The sequences were converted to FASTA format, extracted and trimmed using Geneious (v.2019.1.3). A reference genome file was made using VVV and VDD cDNA sequences with a terminal pad of A-tails added to maximise sensitivity [39]. The reference file was indexed using Bowtie Build (Version 0.12.9) and the reads were mapped to the reference file using the recombinant-mapping algorithm, ViReMa (Viral-Recombination Mapper, Version 0.15). The recombinant sequences were compiled as a text file and analysed using ggpubr (v2.3) in RStudio.

RESULTS

Modular RG system design for DWV

To compare the virulence and competitiveness of DWV types and their recombinants a set of cDNA clones were prepared. By exploiting the modular organisation of the DWV genome [9, 21] we have previously constructed infectious cDNAs for several distinct genetic variants of DWV [25]. For convenience these are referred to as follows: VDD (DWV type A coding sequence, GenBank MT415949), VVD (a type B/A recombinant, GenBank MT415950), and VVV (DWV type B, GenBank MT415952). In addition, we constructed a cDNA for a complete type A DWV, designated DDD, using a similar gene synthesis and module replacement strategy [25] to incorporate the DWV type A 5′-untranslated region (5′-UTR; DWV-A 1414, GenBank KU847397 used as a reference—Fig. S1). VDD, VVV, and VVD DWV variants were previously shown to be infectious and cause symptomatic disease in honey bees [25]. Infectivity of the DDD virus was verified by analysis of DWV accumulation in injected pupae and was indistinguishable from the VDD virus (Fig. S2a). Derivatives of VDD, VVD, and VVV, expressing EGFP or mCherry, were generated as previously described [25] (Fig. S1) and their replication verified following inoculation of pupae (for example, Fig. S2b).

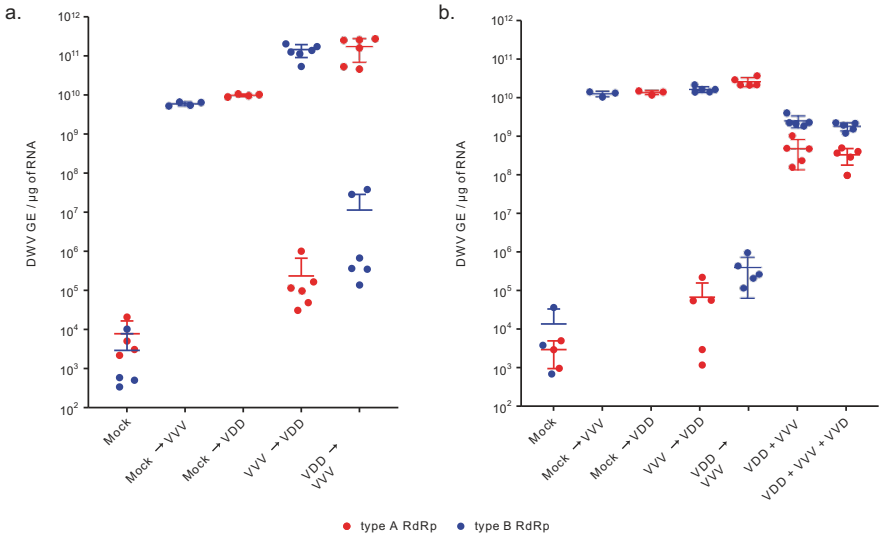


Fig. 1 Coinfection and superinfection of honey bee brood with DWV. a Superinfection of honey bee pupae preliminarily infected *per os* at larval stage and then via injection at pupal stage. Quantified viral titres of VV and VDD DWV in pupae 24 h post-injection of the superinfecting virus are shown (second virus inoculated by injection after primary infection by feeding with the reciprocal DWV variant is indicated by "→", e.g. VV (fed) → VDD (injected)). b DWV accumulation in honey bee pupae in coinfected (mixed virus population indicated by "+") or superinfected samples (second injection 24 h after primary infection by injection is indicated by "→"). Primer pairs for type A or type B RdRp amplifying variant-specific fragments of virus polymerase encoding region were used to distinguish between the administered variants. Data points represent DWV levels in individual pupae with two points of different colour corresponding to different virus variants (red for type A RdRp and blue for type B RdRp respectively) in the same pupa (or in individual pupae for VV or VDD only injected samples). Error bars show mean ± SD for each virus variant in each injection group, GE—genome equivalents. ANOVA: $p < 0.05$ for type A accumulation in "Mock→VDD" vs. "VV → VDD", for type B in "Mock→VV" vs. "VDD → VV" groups, for type A vs. type B in "VV + VDD" and "VV + VD + VDD" groups.

Superinfection and coinfection studies

Varroa delivers DWV to developing honey bee pupae by direct injection when feeding. Pupae will already contain previously acquired DWV and the mite may contain one or more DWV variants. We investigated the consequences of coinfection and superinfection on accumulation of distinct DWV variants in honey bee pupae under laboratory conditions. Primary infection was achieved by feeding first instar larvae (0–1 day old) with a diet containing 10^7 genome equivalents (GE) of either VDD or VV DWV, followed by secondary inoculation by injection (10^3 GE) with the reciprocal virus variant 10 days later at the white-eyed pupal stage. The viral load in individual pupae was analysed by qPCR 24 h post-injection using DWV type-specific primers for the RNA-dependent RNA polymerase (RdRp) coding region. In the absence of superinfection, larvae fed either VDD or VV showed accumulation of the virus to levels of ~10^{10} GE/µg RNA by the time of pupation (Fig. S2c). Pupae initially infected by larval feeding and subsequently superinfected by injection showed a markedly reduced accumulation of the injected DWV variant when compared to the same virus in pupae which were not fed DWV as larvae (Fig. 1a).

Reduced accumulation of a superinfecting virus was also observed when white-eyed pupae were initially injected with VDD or VV 24 h prior to introduction of the reciprocal virus variant (first injection—10^2 GE, superinfection—10^6 GE, Fig. 1b). In contrast, simultaneous infection with two or three (VDD, VV and VD) DWV variants (10^2 GE in total virus injected corresponding to 0.5×10^2 or 0.33×10^2 GE of each variant for two- and three-component infections respectively) resulted in nearly equivalent virus loads, although the VDD variant accumulated to slightly lower (~0.5 $\log_{10}$) titres at 24 h post-injection.

Dynamics of DWV accumulation in superinfection conditions

We extended these studies to determine whether the apparent competitive disadvantage for the second virus remained after an extended incubation period. Pupal injections were repeated as before and viral loads quantified 5 and 7 days after superinfection. A recombinant type B/A variant (VD) was additionally included both as primary and superinfecting virus. In reciprocal infections using VDD and VV both the initial and the superinfecting virus reached nearly equivalent levels within the incubation period (Fig. 2). In contrast, in virus pairings with a greater sequence identity between the genomes the superinfecting virus exhibited a reduced accumulation even after prolonged incubation. In the "VDD → VD", "VD → VDD", and "VD → VV" groups the super-infecting virus levels were ~2 $\log_{10}$ lower than the initial inoculum at 5–7 days post-injection. For the "VV → VD" pairing this was more marked, with the superinfecting virus still ~4 $\log_{10}$ lower after 7 days. In control pupae infected with VDD, VD or VV individually all three viruses reached high titres 7 days post-injection (Fig. 2 and Fig. S3). In addition, virus accumulation was monitored after coinfection of equal amounts of each combination of VDD, VD and VV over time. In these studies, all coinfecting variants achieved similar titres 5 days post-inoculation (Fig. S3).

We recently demonstrated that bumble bees are susceptible to DWV infection when directly injected at high doses [38]. We therefore investigated the influence of the host environment on

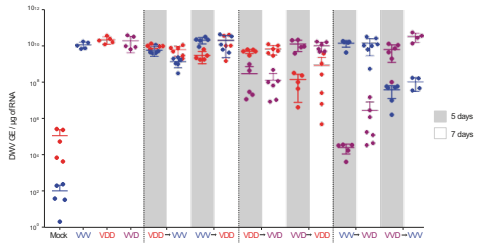


Fig. 2 qPCR analysis of DWV accumulation in superinfection conditions. Honey bee pupae received a primary injection with one DWV variant (VV, VDD or VVD) and a secondary injection (superinfection) with a different variant 24 h later. DWV accumulation was quantified 5 (grey shading) and 7 (no shading) days after the second injection. Primer sets specifically targeting the viral polymerase or a structural protein encoding region of DWV type A or type B were used to detect accumulation of each of the injected variants. Data points represent DWV levels in individual samples with two points of different colour corresponding to different virus variants (red for VDD, blue for VV, and purple for VVD respectively) in the same pupa (or in individual pupae for VV, VDD or VVD only injected samples). Error bars show mean ± SD for each virus variant in each injection group, GE—genome equivalents. ANOVA: $p < 0.05$ for "Mock→VDD" vs. "VVD → VDD", "Mock→VVD" vs. "VDD → VVD", "Mock→VVD" vs. "VV → VVD", "Mock→VV" vs. "VVD → VV" at 7 days time point.

the DWV superinfection by conducting similar experiments in bumble bee pupae. At 48 h post superinfection the levels of the second virus administered were lower than that of the primary virus inoculated but—with the exception of the "VVD → VV" combination—had achieved similar levels by 6 days post-injection (Fig. S4).

These results suggest that a superinfecting virus experiences an initial competitive disadvantage, but that this disadvantage is overcome after 5–7 days unless the viruses exhibit more extensive sequence identity. To investigate this further we studied superinfection with essentially identical viruses, using two VVD variants distinguishable solely by unique genetic tags—VVD_S and VVD_H, tagged with a *SalI* or *HpaI* restriction site respectively (Fig. S1, Text S1)—which differ by just four nucleotides. Honey bee pupae injected with VVD_H were challenged 24 h later with VVD_S and analysed by end point PCR and restriction assay 1, 3 and 6 days after superinfection. No VVD_S was detectable in superinfected pupae at any time point analysed (Fig. S5) suggesting a complete or near-complete block of the superinfecting genome amplification. Control injections of VVD_S into pupae, which did not receive VVD_H virus, allowed detection of *SalI*-tagged cDNA 24 h post-inoculation.

Tissue localisation studies using reporter-encoding DWV

Total RNA levels analysis allows the quantification of DWV to be determined, but it obscures details of the relative distribution and tissue tropism of individual virus variants. Previously we developed an EGFP-encoding RG system for DWV [25] based upon the VDD genome and designated DWV_E (for convenience here renamed to VDD_E). We used VDD_E to define whether the primary infection also affects the distribution of the superinfecting virus. Furthermore, we constructed a full length DWV type A genome, designated DDD (Fig. S1), and similarly investigated superinfection of DDD infected pupae. Pupae that had received an initial injection of 10^2 GE of DDD, VDD, VVD or VV were inoculated 24 h later with 10^6 GE of VDD_E. Live pupae were analysed by confocal microscopy for the presence of the EGFP signal (Fig. 3). Three regions of each pupa were visualised—the head, the developing wing and the abdomen—as we have previously demonstrated significant virus accumulation in these locations [25].

Injection of VDD_E in the absence of a primary infection ("Mock→VDD_E" group) resulted in efficient expression of EGFP throughout the pupa 24 h post-inoculation (Fig. 3a–c). In the case of superinfection, the EGFP signal could be seen 24 h later only in pupae where VV was used as a primary infecting genotype (Fig. 3h and i). In these pupae, the number of fluorescent foci was lower when compared to the "Mock → VDD_E" group infected for the same 24 h period (Fig. 3, panels a–c vs. g–i). No EGFP signal was visible upon superinfection with VDD_E after 24 h in pupae first injected with VVD, VDD and DDD (data not shown). At 4–5 days post-inoculation with VDD_E there were also differences observed in the levels and distribution of the reporter protein. For example, no EGFP signal was found in the wings after primary inoculation with VVD or DDD (Fig. 3, panels n and t vs. e and k). Visible EGFP expression was detected in the head and abdomen in the pupae from these injection groups after 4–5 days but the extent and number of fluorescent foci was reduced when compared to the "Mock → VDD_E" and "VV → VDD_E" pupae (Fig. 3, panels m, o, s and u vs. panels d, f, j and l). In contrast to the "Mock→VDD_E" and "VV → VDD_E" samples, only a fraction of pupae in "VVD → VDD_E" and "DDD → VVD_E" groups exhibited detectable EGFP signal in each of the body sites under analysis (Table S2). Finally, pupae initially injected with VDD did not show any detectable EGFP signal even 6 days after superinfection with VDD_E (Fig. 3p–r, Fig. S6), suggesting again that greater sequence identity restricts the activity of the superinfecting virus.

To confirm that the external analysis of the intact living pupae was representative, selected samples were dissected. Tissue samples, including parts of the digestive tract, wing rudiments, thoracal muscle tissue, brain and cephalic glands were visualised by confocal microscope (Fig. S6). This analysis recapitulated the pattern of fluorescence observed by previous visualisation of intact pupae. To complement the microscopy data, we quantified DWV RNA in selected pupae by qPCR at 24 h and 5 days post superinfection (Fig. S7). Analysis by qPCR revealed that a proportion of the VDD_E population has lost the non-essential EGFP-coding sequences evidenced by the EGFP/RdRp ratio in the "Mock→VDD_E" group. This instability of EGFP-expressing viruses has previously been reported for DWV [40] and poliovirus [41]. However, sufficient EGFP-expressing reporter viruses remained in the population to allow their detection by microscopy and RNA quantification 5 days post-injection. The results of qPCR have shown that there was a good agreement between the amount of genomic RNA and the observed level of fluorescence.

Localisation of DWV in coinfected and superinfected pupae using two-colour microscopy

In order to visualise the distribution of infection with different DWV variants we used EGFP- and mCherry-expressing viruses, VDD_E, VVD_E, and VV_{mC} (with subscript E and mC indicating the EGFP or mCherry reporter respectively, Fig. S1). For coinfection, pupae were injected with equimolar mixtures of VVD_E or VDD_E and VV_{mC} and analysed under the confocal microscope 1–5 days post-inoculation. We could readily detect red and green fluorescent signals present in the same tissues of virus-injected pupae as previously described [25], including multiple tissues of the digestive tract, wings and head tissues. The reporter gene expression sites appeared as individual punctate foci of either red or green fluorescence, with only a few displaying dual fluorescence for both reporters (Fig. 4a and Fig. S8). The analysis of VDD_E–infected pupae superinfected with VV_{mC} and visualised by microscopy after a further 24 h revealed a similar distribution of the fluorescent signal as in coinfected samples (Fig. 4b).

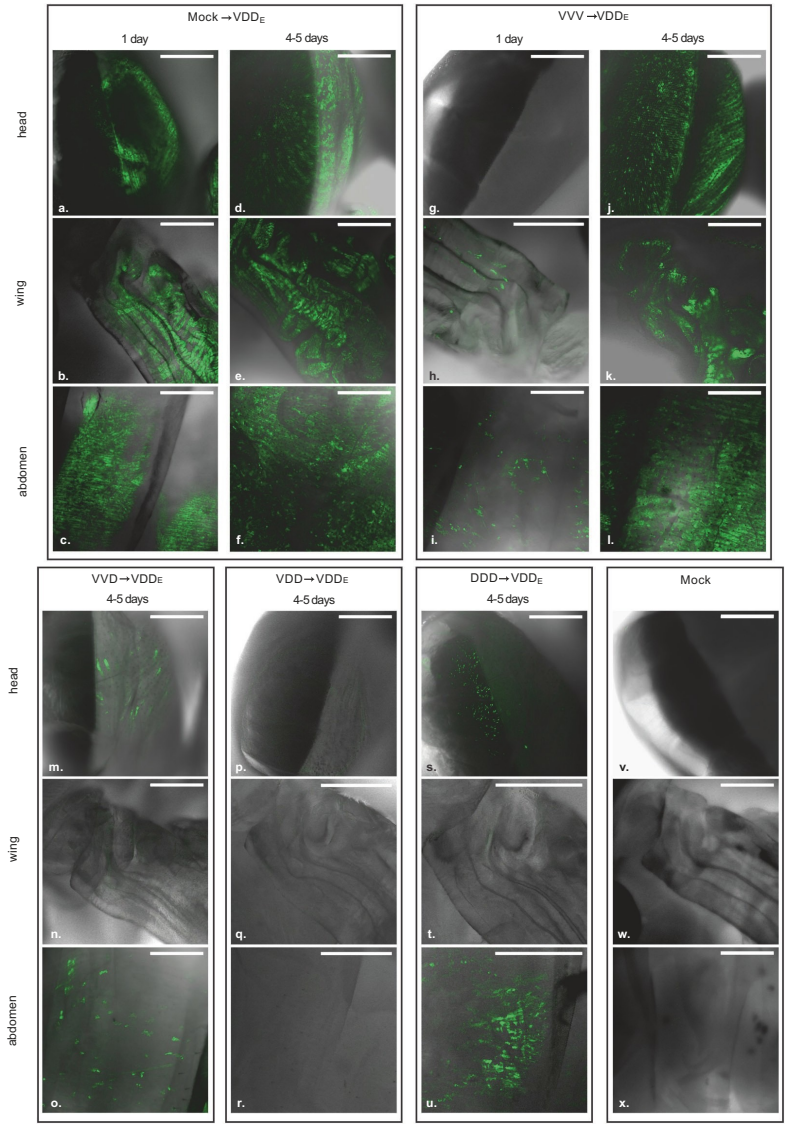


Fig. 3 EGFP signal localisation in VDDE injected honey bee pupae analysed by confocal microscopy. Combined white-field and fluorescent images are shown for convenience of interpretation. Pupae were analysed 1 and 4–5 days after the second injection. For VVD, VDD and DDD primary infection groups only samples incubated for 4–5 days are shown, as no EGFP was detected after 1 day. Scale bars correspond to 500 µm.

Recombination between VDD and VVV DWV

The interpretation of the superinfection studies is based upon sequence-specific quantification of particular regions of the virus genome by qPCR. This interpretation could be confounded by extensive levels of genetic recombination, a natural consequence of coinfection with related viruses [42]. Genetic recombination of RNA viruses requires that both parental genomes are present within an individual cell [43]. Since our microscopy analysis had detected only limited numbers of apparently dually infected foci during mixed infections, we conducted further analysis to investigate the presence and identity of viral recombinants, and the influence of the order of virus acquisition on recombination, using next generation sequencing. Illumina paired-reads were generated from PCR amplicons of a 10 Kb fragment of the DWV genome targeting pupae initially

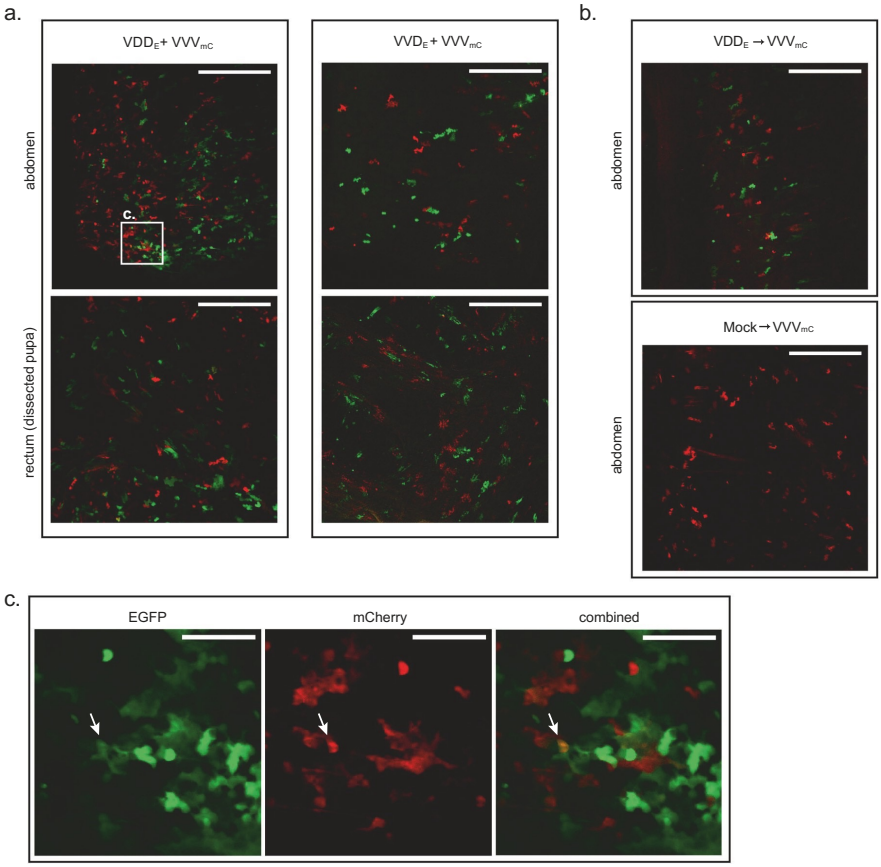


Fig. 4 Confocal microscopy analysis of honey bee pupae coinfected or superinfected with DWV variants encoding EGFP and mCherry. a Coinfection: "VDD$_E$+ VVV$_{mC}$" or "VVD$_E$+ VVV$_{mC}$" panels show red (mCherry) and green (EGFP) fluorescence signals detected in the abdomen of intact pupae (upper row) or in the dissected tissues of the digestive tract (rectum tissue shown as an example). Scale bars correspond to 500 µm. b Superinfection: upper panel—abdomen of an intact pupa initially infected with VDD$_E$ and superinfected 24 h later with VVV$_{mC}$ analysed 24 h after the second injection; lower panel—abdomen of an intact pupa injected with VVV$_{mC}$ only and analysed 24 h post-inoculation. Scale bars correspond to 500 µm c Magnified image of highlighted region (from a) of pupa coinfected with VDD$_E$ + VVV$_{mC}$. Individual images for EGFP and mCherry signals, and a combined image for both fluorophores are shown. Arrows indicate individual foci of infection exhibiting both EGFP and mCherry expression. Scale bars correspond to 100 µm.

infected with VDD and challenged with VVV, or vice versa (samples "VVV → VDD" and "VDD → VVV" at 5 or 7 days after superinfection). Recombination junctions were detected across the entirety of the DWV genome in all samples analysed with 'hotspots' of recombination denoted by an increased number of aligned reads identified at numerous points in the genome (Fig. 5b, Table S3), including some previously reported [21]. The percentage of reads corresponding to recombination junctions varied in individual pupae from 1–2.2% of all mapped reads (Fig. S9). In all cases approximately equal proportions of recombinants were detected with VVV or VDD as the 5′-acceptor partners (terminology assumes that recombination occurs during negative strand synthesis [42, 44]). In several instances we detected the same recombination junction with both VVV and VDD as the 5′-acceptor. Our analysis also revealed recombination sites in which the 5′ was only ever derived from one variant or the other (red and blue points in Fig. 5b, Table S3). These results demonstrate that although superinfecting virus recombines readily with an established variant, the recombinant population remains a minor component of the total virus population, and is well below the level expected to confound our analysis of competition between extant and superinfecting viruses.

DISCUSSION

The global distribution and ubiquitous nature of DWV [14, 19], transmitted vertically and horizontally in honey bees [15, 45, 46], inevitably means that when vectored by *Varroa* it is introduced to the host as a superinfecting virus. As such, there is the potential for competition for cellular resources in coinfected tissues, or the possibility of a pre-existing infection retarding or inhibiting superinfection through molecular mechanisms including SIE or the immune responses induced by the initial

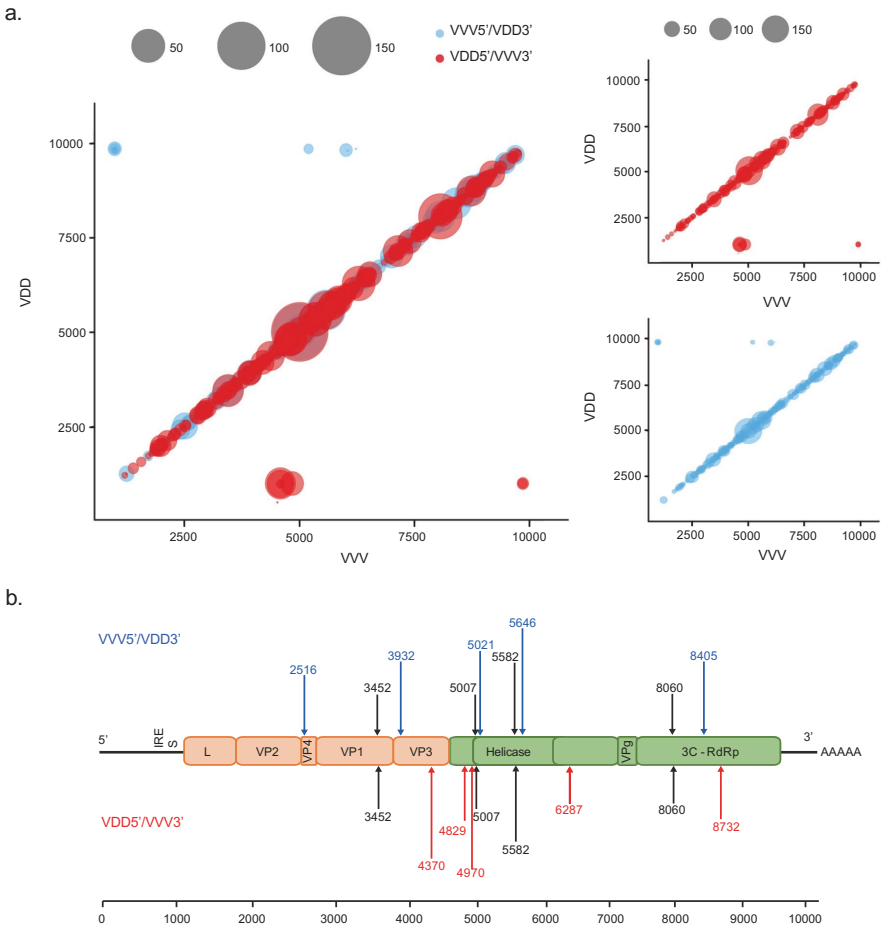


Fig. 5 Genomic recombination events observed between VVV and VDD DWV variants in a superinfected honey bee pupa. a Mapped recombination events in a honey bee pupa initially infected with VVV and superinfected with VDD DWV. The plot shows recombination events occurring along the full length of the DWV genome, with the VDD genome length shown on the *Y*-axis and VVV shown on the *X*-axis. Each bubble represents a unique recombination site and bubble size is determined by the number of mapped and aligned reads for this site obtained using ViReMa analysis. The colour of the bubble indicates the direction of recombination, with blue representing VVV as 5′-acceptor and VDD as 3′-donor and those in red with VDD as 5′-acceptor and VVV as 3′-donor. b The most frequently observed recombination junctions in all pupal samples analysed, shown at the recombination junction between VVV and VDD genomes. The junctions shown in black occurred with similar frequency in both directions. Those shown in blue occurred predominantly with VVV as 5′-acceptor and VDD as 3′-donor and those in red with VDD as 5′-acceptor and VVV as 3′-donor sequences.

virus. There are at least two distinct types of DWV circulating globally—type A and type B—with documented differences in their distribution [8, 18, 22] and, perhaps, pathogenesis [13, 19, 23, 25]. If the outcome of superinfection always favoured one virus type it would influence transmission of DWV variants potentially accounting for their geographic distribution and—if associated with differences in virulence—the impact on the honey bees.

SIE has been reported for DWV, with the suggestion that bees bearing a type B virus were protected from subsequent type A transmitted from infesting *Varroa* mites [10]. SIE is described for several human, animal and plant viruses [26–35], and may operate via a number of molecular mechanisms [26, 27, 30, 31, 47–50]. Precedents already exist in plants with milder forms of a virus providing protection against more virulent strains [51, 52] and the recent spread of DWV type B in the USA [8, 22] could be interpreted as an indirect consequence of SIE, with bees harbouring this virus less susceptible to infection by DWV type A. However, there are other potential differences between DWV types such as the ability of variants with type B capsid to replicate in *Varroa* [25, 53], which may enhance its spread over the non-propagative transmission reported for type A [54].

The availability of a RG system allowed us to investigate the consequences of coinfection and superinfection with DWV type A

and B in individual honey bees. We found that when coinfected DWV type A and B (VDD and VVV variants) demonstrate broadly similar levels of replication (Fig. 1b). In contrast, in sequential infections, either of virus-fed larvae or injected pupae, superinfecting DWV variants showed delayed replication (Fig. 1). This delay was dependent upon the genetic similarity of the primary and secondary viruses and appeared transient in certain pairings. In genetically divergent pairings (e.g. "VDD → VVV" and "VVV → VDD") high levels of both viruses were reached after a prolonged incubation period. In contrast, where the extent of genetic identity between the primary and secondary virus was greater, the superinfecting virus failed to 'catch up', even after 7 days (Fig. 2). This was most dramatically demonstrated using two genomes that differed by just four nucleotides (VVD_S and VVD_H variants), in which case the superinfecting virus remained undetectable after 6 days incubation (Fig. S5). In addition, we found that delayed accumulation of the genetically similar superinfecting DWV variants is not specific to honey bee host and was also observed in bumble bees, a species susceptible to DWV infection when directly injected (Fig. S4) [38].

We extended our analysis in honey bee pupae using reporter gene-expressing viruses and demonstrated that replication, characterised by the expression of the fluorescent protein, was inversely related to the level of genetic identity between the primary and superinfecting viruses (Fig. 3). For example, VDD_E replicated extensively, albeit somewhat delayed when compared with VDD_E-only infected pupae, in pupae that had received VVV as the primary virus (Fig. 3j–l), but was undetectable in pupae initially inoculated with VDD (Fig. 3p–s). Notably, in each case where the superinfecting virus showed reduced replication after extended incubation, dominance in the replication showed no directionality according to virus type and was due solely to the order of addition. Based on this data it is likely that sequential infection with DWV type A and B will result in both viruses replicating to maximal levels before eclosion of either worker or drone brood pupae (which pupate for ~12 or ~14 days respectively). It remains to be determined whether the delay we demonstrate is sufficient to influence the colony-level virus population, or that carried and transmitted by *Varroa*.

Where cellular coinfection occurs, viruses have the opportunity to genetically recombine. This is a widespread phenomenon in the single-stranded positive-sense RNA viruses [55, 56] and has previously been documented in DWV [6–9]. Our microscopy analysis of honey bee pupae infected with two reporter-expressing DWV variants predominantly demonstrated non-colocalised expression of the fluorescent signal. However, small numbers of dual-infection foci were detected, directly implying that the opportunity for recombination exists (Fig. 4c and Fig. S8). Using next generation sequencing we confirmed the formation of recombinants and characterised the recombination products by analysis of the viral RNA in pupae reciprocally superinfected with VDD and VVV. 1–2.2% of mapped reads spanned recombination junctions, with no evidence for any bias in their directionality (VDD/VVV or VVV/VDD; Fig. S9). Although these junctions mapped throughout the DWV genome, the greatest number were concentrated in the region of the genome encoding the junction of the structural and non-structural proteins (Fig. 5a). This observation matches that found for other picornaviruses and reflects the mix'n'match modular nature of the *Picornavirales* genome [6, 7, 9, 18, 21]. In this, functional capsid-coding modules can, through recombination, be juxtaposed with non-structural coding modules from a different parental genome [57]. A small number of recombination junctions (~350 of 35750 unique junctions mapped) plotted as outliers from the diagonal of genome-length recombinants. Analysis of these sequences showed that the majority were out of frame deletions (Woodford, unpublished), and so incapable of replicating. Our studies using analogous approaches in other RNA viruses show that these types of aberrant products are not unusual and reflect the random nature of the molecular mechanism of recombination [44, 58].

The competition we demonstrate in sequential DWV infections appears to be guided by the amount of genetic identity between the viruses. This suggests it is most likely mediated via RNA interference (RNAi). In arthropods antiviral RNAi response acts via generation of short double stranded RNAs (siRNA) from virus RNA replication intermediates through cleavage by the enzyme Dicer. These are further used by the RNA induced silencing complex to target the destruction of complementary sequences [59, 60]. Hence viral RNA genomes exhibiting greater identity are likely to generate higher numbers of cross-reactive siRNAs while the differences between these siRNA and the target virus genome might have a significant impact [18]. Previous analysis of the RNAi population in DWV-infected honey bees demonstrated that 75% of DWV-specific short RNA are 21/22 mers [21]. Although DWV type A and B exhibit ~85% genetic identity it is not contiguous (Fig. S1), but is instead distributed in ~1350 short regions of 1–389 nucleotides. Less than 4% of these identical regions are 21 nucleotides or longer, and therefore capable of generating perfectly complementary siRNAs. Recalculation of the identity between genomes having excluded sequences under 21 nucleotides demonstrates that there is only 34% genetic identity between DDD and VVV (Table S4). Comparing the figures from this analysis and the quantification of DWV accumulation in superinfected pupae suggests a clear relationship between the extent of the competition observed and the genetic identity of contiguous sequences. It is already known that exogenous RNAi can control DWV and other RNA viruses of honey bees [61–63], and in our preliminary studies we have shown that RNAi-mediated suppression of Dicer leads to both increased pathogenesis and viral loads in DWV-infected bees (Fig. S10). Further research will be required to determine the impact of RNAi in competition between superinfecting DWV variants and its potential exploitation in studies to develop cross-reactive vaccines against DWV [36]. These future studies will need to take account of the disrupted complementarity between the genomes (Table S4), the uneven distribution of the RNAi response mapped to the genomic RNA of the infecting virus [21], and both the variation acceptable within the RNAi seed sequence and the RNA structure of the target. Sequential or simultaneous infections with DWV is an inevitable consequence of the multiple routes by which the virus can be acquired—vertically, horizontally *per os* and vectored by *Varroa*. Our studies demonstrate that it is the order of acquisition, not the specific type of DWV, that determines the outcome of sequential or superinfection. The role of genetic identity in competition between DWV variants at an individual and landscape scale is likely to be a fruitful area of research, and may allow the future development of rationally designed vaccines against DWV that exploit the conserved RNAi response of the infected host.

REFERENCES

1. Honey: market value worldwide 2007–2016. https://www.statista.com/statistics/933928/global-market-value-of-honey/. Accessed Nov 2020.
2. Highfield AC, El Nagar A, Mackinder LCM, Noël LM-LJ, Hall MJ, Martin SJ, et al. Deformed wing virus implicated in overwintering honeybee colony losses. Appl Environ Microbiol. 2009;75:7212–20.
3. Ongus JR, Peters D, Bonmatin JM, Bengsch E, Vlak JM, van Oers MM. Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite *Varroa destructor*. J Gen Virol. 2004;85:3747–55.
4. Lanzi G, Miranda JRD, Boniotti MB, Cameron CE, Lavazza A, Capucci L, et al. Molecular and biological characterization of Deformed wing virus of honeybees (*Apis mellifera* L.). J Virol. 2006;80:4998–5009.
5. Fujiyuki T, Takeuchi H, Ono M, Ohka S, Sasaki T, Nomoto A, et al. Kakugo virus from brains of aggressive worker honeybees. Adv Virus Res. 2005;65:1–27.
6. Dalmon A, Desbiez C, Coulon M, Thomasson M, Le Conte Y, Alaux C, et al. Evidence for positive selection and recombination hotspots in *Deformed wing virus* (DWV). Sci Rep. 2017;7:41045.

7. Zioni N, Soroker V, Chejanovsky N. Replication of Varroa destructor virus 1 (VDV-1) and a Varroa destructor virus 1–deformed wing virus recombinant (VDV- 1–DWV) in the head of the honey bee. Virology. 2011;417:106–12.
8. Ryabov EV, Childers AK, Chen Y, Madella S, Nessa A, Vanengelsdorp D, et al. Recent spread of Varroa destructor virus - 1, a honey bee pathogen, in the United States. Sci Rep. 2017;7:17447.
9. Moore J, Jironkin A, Chandler D, Burroughs N, Evans DJ, Ryabov EV. Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies. J Gen Virol. 2011;92:156–61.
10. Mordecai GJ, Brettell LE, Martin SJ, Dixon D, Jones IM, Schroeder DC. Super-infection exclusion and the long-term survival of honey bees in Varroa-infested colonies. ISME J. 2015;10:1182–91.
11. Woodford L, Evans DJ. Deformed wing virus: using reverse genetics to tackle unanswered questions about the most important viral pathogen of honey bees. FEMS Microbiol Rev. 2020; fuaa070, https://doi.org/10.1093/femsre/fuaa070.
12. Mordecai GJ, Wilfert L, Martin SJ, Jones IM, Schroeder DC. Diversity in a honey bee pathogen: first report of a third master variant of the Deformed Wing Virus quasispecies. ISME J. 2016;10:1264–73.
13. McMahon DP, Natsopoulou ME, Doublet V, Fürst M, Weging S, Brown MJF, et al. Elevated virulence of an emerging viral genotype as a driver of honeybee loss. Proc Biol Sci. 2016;283:443–9.
14. Wilfert L, Long G, Leggett HC, Schmid-Hempel P, Butlin R, Martin SJM, et al. Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites. Science. 2016;351:594–7.
15. de Miranda JR, Genersch E. Deformed wing virus. J Invertebr Pathol. 2010;103: S48–S61.
16. Roberts JMK, Anderson DL, Durr PA. Absence of deformed wing virus and Varroa destructor in Australia provides unique perspectives on honeybee viral landscapes and colony losses. Sci Rep. 2017;7:6925.
17. Yue C, Schröder M, Gisder S, Genersch E. Vertical-transmission routes for deformed wing virus of honeybees (Apis mellifera). J Gen Virol. 2007;88:2329–36.
18. Ryabov EV, Childers AK, Lopez D, Grubbs K, Posada-Florez F, Weaver D, et al. Dynamic evolution in the key honey bee pathogen deformed wing virus: novel insights into virulence and competition using reverse genetics. PLoS Biol. 2019; 17; https://doi.org/10.1371/journal.pbio.3000502.
19. Martin SJ, Highfield AC, Brettell L, Villalobos EM, Budge GE, Powell M, et al. Global honey bee viral landscape altered by a parasitic mite. Science. 2012;336:1304–6.
20. Loope KJ, Baty JW, Lester PJ, Wilson Rankin EE. Pathogen shifts in a honeybee predator following the arrival of the Varroa mite. Proc Biol Sci. 2019;286:20182499.
21. Ryabov EV, Wood GR, Fannon JM, Moore JD, Bull JC, Chandler D, et al. A virulent strain of deformed wing virus (DWV) of honeybees (Apis mellifera) prevails after Varroa destructor-mediated, or in vitro, transmission. PLoS Pathog. 2014;10: e1004230.
22. Kevill JL, de Souza FS, Sharples C, Oliver R, Schroeder DC, Martin SJ. DWV-A lethal to honey bees (*Apis mellifera*): a colony level survey of DWV variants (A, B, and C) in England, Wales, and 32 States across the US. Viruses. 2019;11:426.
23. Tehel A, Vu Q, Bigot D, Gogol-Döring A, Koch P, Jenkins C, et al. The two prevalent genotypes of an emerging infectious disease, *Deformed wing virus*, cause equally low pupal mortality and equally high wing deformities in host honey bees. Viruses. 2019;11:114.
24. Norton AM, Remnant EJ, Buchmann G, Beekman M. Accumulation and competition amongst Deformed wing virus genotypes in naïve Australian honeybees provides insight Into the increasing global prevalence of genotype B. Front Microbiol. 2020;11:620.
25. Gusachenko ON, Woodford L, Balbirnie-Cumming K, Campbell EM, Christie CR, Bowman AS, et al. Green bees: reverse genetic analysis of Deformed wing virus transmission, replication, and tropism. Viruses. 2020;12:532.
26. Steck FT, Rubin H. The mechanism of interference between an avian leukosis virus and Rous sarcoma virus. II. Early steps of infection by RSV of cells under conditions of interference. Virology. 1966;29:642–53.
27. Adams RH, Brown DT. BHK cells expressing Sindbis virus-induced homologous interference allow the translation of nonstructural genes of superinfecting virus. J Virol. 1985;54:351–7.
28. Strauss JH, Strauss EG. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev. 1994;58:491–562.
29. Karpf AR, Lenches E, Strauss EG, Strauss JH, Brown DT. Superinfection exclusion of alphaviruses in three mosquito cell lines persistently infected with Sindbis virus. J Virol. 1997;71:7119–23.
30. Singh IR, Suomalainen M, Varadarajan S, Garoff H, Helenius A. Multiple mechanisms for the inhibition of entry and uncoating of superinfecting Semliki Forest virus. Virology. 1997;231:59–71.
31. Geib T, Sauder C, Venturelli S, Hässler C, Staeheli P, Schwemmle M. Selective virus resistance conferred by expression of Borna disease virus nucleocapsid components. J Virol. 2003;77:4283–90.
32. Edwards MC, Bragg J, Jackson AO. Natural resistance mechanisms to viruses in barley. In: Loebenstein G and Carr JP, editors. Natural Resistance Mechanisms of Plants to Viruses. Dordrecht, The Netherlands: Springer; 2006. p. 465–501.
33. Bergua M, Zwart MP, El-Mohtar C, Shilts T, Elena SF, Folimonova SY. A viral protein mediates superinfection exclusion at the whole-organism level but Is not required for exclusion at the cellular Level. J Virol. 2014;88:11327–38.
34. Michel N, Allespach I, Venzke S, Fackler OT, Keppler OT. The Nef protein of human immunodeficiency virus establishes superinfection immunity by a dual strategy to downregulate cell-surface CCR5 and CD4. Curr Biol. 2005;15:714–23.
35. Tscherne DM, Evans MJ, von Hahn T, Jones CT, Stamataki Z, McKeating JA, et al. Superinfection exclusion in cells infected with hepatitis C virus. J Virol. 2007;81:3693–703.
36. Leonard SP, Powell JE, Perutka J, Geng P, Heckmann LC, Horak RD, et al. Engineered symbionts activate honey bee immunity and limit pathogens. Science. 2020;367:573–6.
37. Lamp B, Url A, Seitz K, Rgen Eichhorn J, Riedel C, Sinn LJ, et al. Construction and rescue of a molecular clone of Deformed wing virus (DWV). PLoS ONE. 2016;11: e0164639.
38. Gusachenko ON, Woodford L, Balbirnie-Cumming K, Ryabov EV, Evans DJ. Evidence for and against deformed wing virus spillover from honey bees to bumble bees: a reverse genetic analysis. Sci Rep. 2020;10:16847.
39. Routh A, Johnson JE. Discovery of functional genomic motifs in viruses with ViReMa – a Virus Recombination Mapper – for analysis of next-generation sequencing data. Nucleic Acids Res. 2014;42:e11.
40. Ryabov EV, Christmon K, Heerman MC, Posada-Florez F, Harrison RL, Chen Y, et al. Development of a honey bee RNA virus vector based on the genome of a Deformed wing virus. Viruses. 2020;12:374.
41. Mueller S, Wimmer E. Expression of foreign proteins by poliovirus polyprotein fusion: analysis of genetic stability reveals rapid deletions and formation of cardioviruslike open reading frames. J Virol. 1998;72:20–31.
42. Kirkegaard K, Baltimore D. The mechanism of RNA recombination in poliovirus. Cell. 1986;47:433–43.
43. Egger D, Bienz K. Recombination of poliovirus RNA proceeds in mixed replication complexes originating from distinct replication start sites. J Virol. 2002;76:10960–71.
44. Lowry K, Woodman A, Cook J, Evans DJ. Recombination in enteroviruses is a biphasic replicative process involving the generation of greater-than genome length 'imprecise' Intermediates. PLoS Pathog. 2014;10; https://doi.org/10.1371/journal.ppat.1004191.
45. de Miranda JR, Fries I. Venereal and vertical transmission of deformed wing virus in honeybees (Apis mellifera L.). J Invertebr Pathol. 2008;98:184–9.
46. Yañez O, Jaffé R, Jarosch A, Fries I, Robin FAM, Robert JP, et al. Deformed wing virus and drone mating flights in the honey bee (*Apis mellifera*): Implications for sexual transmission of a major honey bee virus. Apidologie. 2012;43:17–30.
47. Simon KO, Cardamone JJ Jr, Whitaker-Dowling PA, Youngner JS, Widnell CC. Cellular mechanisms in the superinfection exclusion of vesicular stomatitis virus. Virology. 1990;177:375–9.
48. Stevenson M, Meier C, Mann AM, Chapman N, Wasiak A. Envelope glycoprotein of HIV induces interference and cytolysis resistance in CD4+ cells: mechanism for persistence in AIDS. Cell. 1988;53:483–96.
49. Bratt MA, Rubin H.Specific interference among strains of Newcastle disease virus. II. Comparison of interference by active and inactive virus.Virology. 1968;35:381–94.
50. Zou G, Zhang B, Lim P-Y, Yuan Z, Bernard KA, Shi P-Y. Exclusion of West Nile virus superinfection through RNA replication. J Virol. 2009;83:11765–76.
51. Ziebell H, Carr JP. Cross-protection: a century of mystery. Adv Virus Res. 2010;76:211–64.
52. Folimonova SY. Developing an understanding of cross-protection by Citrus tristeza virus. Front Microbiol. 2013;4; https://doi.org/10.3389/fmicb.2013.00076.
53. Gisder S, Genersch E. Direct evidence for infection of mites with the bee-pathogenic Deformed wing virus variant B – but not variant A – via fluorescence-hybridization analysis. J Virol. 2021;95:e01786–20.
54. Posada-Florez F, Childers AK, Heerman MC, Egekwu NI, Cook SC, Chen Y, et al. Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite Varroa destructor in a non-propagative manner. Sci Rep. 2019;9:12445.
55. Barr JN, Fearns R. How RNA viruses maintain their genome integrity. J Gen Virol. 2010;91:1373–87.
56. Bentley K, Evans DJ. Mechanisms and consequences of positive-strand RNA virus recombination. J Gen Virol. 2018;99:1345–56.
57. Muslin C, Mac Kain A, Bessaud M, Blondel B, Delpeyroux F. Recombination in enteroviruses, a multi-step modular evolutionary process. Viruses. 2019;11:859.
58. Alnaji FG, Bentley K, Pearson A, Woodman A, Moore JD, Fox H, et al. Recombination in enteroviruses is a ubiquitous event independent of sequence homology and RNA structure. 2020; preprint at bioRxiv; https://doi.org/10.1101/2020.09.29.319285.

59. Brutscher LM, Flenniken ML. RNAi and antiviral defense in the honey bee. J Immunol Res. 2015;2015:941897.
60. Chejanovsky N, Ophir R, Schwager MS, Slabezki Y, Grossman S, Cox-Foster D. Characterization of viral siRNA populations in honey bee colony collapse disorder. Virology. 2014;454-5:176–83.
61. Desai SD, Eu YJ, Whyard S, Currie RW. Reduction in deformed wing virus infection in larval and adult honey bees (Apis mellifera L.) by double-stranded RNA ingestion. Insect Mol Biol. 2012;21:446–55.
62. Hunter W, Ellis J, Vanengelsdorp D, Hayes J, Westervelt D, Glick E, et al. Large-scale field application of RNAi technology reducing Israeli acute paralysis virus disease in honey bees (Apis mellifera, hymenoptera: Apidae). PLoS Pathog. 2010;6:e1001160.
63. Maori E, Paldi N, Shafir S, Kalev H, Tsur E, Glick E, et al. IAPV, a bee-affecting virus associated with colony collapse disorder can be silenced by dsRNA ingestion. Insect Mol Biol. 2009;18:55–60.

ACKNOWLEDGEMENTS

We express our gratitude to Dr Marcus Bischoff and Gill McVee (University of St Andrews) for helping with microscopic imaging, Ashley Pearson (University of St Andrews) for assistance in molecular biology assays. This work was supported by grant funding from BBSRC BB/M00337X/2 and BB/I000828/1.

DECLARATIONS (ETHICS):

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-021-01043-4.

Correspondence and requests for materials should be addressed to O.N.G.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

8.9 Appendix 9 - Published manuscript – Gusachenko *et al.*, (2020), Viruses.

viruses

MDPI

Article

Green Bees: Reverse Genetic Analysis of Deformed Wing Virus Transmission, Replication, and Tropism

Olesya N. Gusachenko [1,*], Luke Woodford [1], Katharin Balbirnie-Cumming [2], Ewan M. Campbell [3], Craig R. Christie [3], Alan S. Bowman [3] and David J. Evans [1]

1 Biomedical Sciences Research Complex, University of St. Andrews, St. Andrews KY16 9ST, UK; lw86@st-andrews.ac.uk (L.W.); d.j.evans@st-andrews.ac.uk (D.J.E.)
2 Centre for Inflammation Research, Queen's Medical Research Institute, University of Edinburgh, Edinburgh EH16 4TJ, UK; kebalbirnie@outlook.com
3 Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, Aberdeen AB24 3FX, UK; e.m.campbell@abdn.ac.uk (E.M.C.); craig.christie@abdn.ac.uk (C.R.C.); a.bowman@abdn.ac.uk (A.S.B.)
* Correspondence: olesya.gusachenko@gmail.com

Received: 31 March 2020; Accepted: 8 May 2020; Published: 12 May 2020

check for updates

Abstract: Environmental and agricultural pollination services by honey bees, *Apis mellifera*, and honey production are compromised by high levels of annual colony losses globally. The majority are associated with disease caused by deformed wing virus (DWV), a positive-strand RNA virus, exacerbated by the ectoparasitic mite *Varroa destructor*. To improve honey bee health, a better understanding of virus transmission and pathogenesis is needed which requires the development of tools to study virus replication, transmission, and localisation. We report the use of reverse genetic (RG) systems for the predominant genetically distinct variants of DWV to address these questions. All RG-recovered viruses replicate within 24 h post-inoculation of pupae and could recapitulate the characteristic symptoms of DWV disease upon eclosion. Larvae were significantly less susceptible but could be infected orally and subsequently developed disease. Using genetically tagged RG DWV and an *in vitro Varroa* feeding system, we demonstrate virus replication in the mite by accumulation of tagged negative-strand viral replication intermediates. We additionally apply a modified DWV genome expressing a fluorescent reporter protein for direct *in vivo* observation of virus distribution in injected pupae or fed larvae. Using this, we demonstrate extensive sites of virus replication in a range of pupal tissues and organs and in the nascent wing buds in larvae fed high levels of virus, indicative of a direct association between virus replication and pathogenesis. These studies provide insights into virus replication kinetics, tropism, transmission, and pathogenesis, and produce new tools to help develop the understanding needed to control DWV-mediated colony losses.

Keywords: insect viruses; honey bee; pollination; virus vector; *Varroa*; RNA viruses; DWV; reverse genetics

1. Introduction

Deformed Wing Virus (DWV) is arguably—in concert with its vector the ectoparasitic mite *Varroa destructor*—the most important pathogen of the European honey bee (*Apis mellifera*). DWV has a near-global distribution (excluding Australia, where it is either absent or present at lower levels without *Varroa* transmission [1]) and, in the absence of *Varroa*, persists at low levels and is rarely pathogenic [2]. In contrast, when transmitted by mites, DWV titres become highly elevated, infested pupae may develop characteristic symptoms and significant levels of overwintering colony losses occur [3–6], attributable to the reduction of honey bee longevity [7]. Evidence on the selective evolution of DWV

through vector transmission bottlenecking [8,9], selection of *Varroa* propagating DWV variants [10], and synergistic action of the mite and the virus on the host [11] have been reported. However, there remain conflicting studies on the ability of the mite to support DWV replication with some indicating biological [10] and others favouring non-propagative transmission routes [12]. Although there is a clear correlation between the mite-borne transmission and symptomatic outcome of the DWV infection the underlying mechanisms of the cooperative action of the two pathogens needs further clarification. A better understanding of DWV pathogenesis is needed to further develop intervention strategies to prevent and control disease.

DWV is a picorna-like virus from the *Iflaviridae* family [13,14]. The single-stranded positive-sense RNA genome encodes a polyprotein flanked by 5′- and 3′-untranslated regions (UTR). Based upon our understanding of related viruses, the polyprotein is processed by viral and/or cellular enzymes into the structural and non-structural proteins required to complete the virus life cycle. The structural proteins form the virus capsid [14], whereas the non-structural proteins modify the cellular milieu and replicate the genome. Like other RNA viruses, DWV is genetically diverse, with a related complex of viruses divided into two or three groups sharing ~84–97% genetic identity. DWV A [14] and Kakugo virus [13] exhibit 97% identity in their RNA sequences and form the type A subgroup. Another master variant of DWV was initially isolated from *Varroa* and named Varroa Destructor Virus type 1 (VDV-1) [15]. As a consequence of its high sequence similarity (84/95% identity at the RNA and protein levels respectively to DWV A) [16] and its ability to infect the same host (honey bee), it is often referred to as DWV type B [17,18]. A third master variant of the virus designated as DWV C has also been reported [18]. A range of differences in host preference, tissue tropism, morbidity, and pathogenicity have been suggested for the two master variants [9,10,19–22]. For example, the predominance of DWV A in a landscape-scale study on Hawaii following the introduction of *Varroa* to naïve colonies with a diverse virus population was interpreted as an indication that this variant was more virulent [17,22]. Conversely, in side-by-side studies in laboratory experiments, DWV A had a less pronounced effect on adult honey bee survival compared to DWV B or a mixture of both variants [20]. Further studies using field sourced inoculates of DWV A and B showed that they were equally virulent and generated similar levels of morbidity in emerged adult bees [23]. In addition to these so-called master variants, a range of recombinants between DWV A and B have been reported [21,24–26]. For example, VDV-1_{DVD} (GenBank HM067437) and VDV-1_{VVD} (VDV-1-DWV-No-9, GenBank HM067438), both bearing the DWV A capsid proteins coding region and DWV B non-structural coding region [26]. In some studies, these accumulated to a higher level in infected honey bees than the parental strains and it has been suggested that evolution of the DWV quasispecies is driven by *Varroa* transmission toward the emergence of variants with enhanced virulence [21,26]. All of these reports are based on virus field isolates, and it remains unclear whether the DWV master variants and recombinants fundamentally differ in their phenotypes or if the differences reported reflect local strain variation or the experimental system used [9,17]. Therefore, further studies are required to associate the virulence with a particular genotype. A direct way to address this, and one that allows the propagation of near-clonal viral stocks for analysis, is to generate viruses using a reverse genetic (RG) system.

In virology, RG involves the manipulation of the genotype, the recovery of the virus, and the investigation of the phenotype. Over almost four decades, it has become the *de facto* standard approach to address questions about virus replication, virulence and pathogenesis [27,28]. To facilitate these studies, a range of genome modifications (e.g. reporter genes) have been used to allow the sensitive quantification and localization of the virus [29]. Molecular cloning of individual genetic variants of DWV is required to establish a direct connection between infection, and the observed symptoms. RG systems for type A DWV have been reported [12,30]. In this study, we exploit an extended RG toolbox for both DWV A and B master variants and a type B/A recombinant to investigate their comparative transmission, tropism and pathogenesis in honey bees. Using these resources, we also provide direct evidence of DWV replication in *Varroa destructor*. Finally, by introducing a reporter encoding sequence to the virus genome we visualise the *in vivo* tissue distribution of DWV in infected

honey bee brood. The results of this study provide new insights into understanding the nature of DWV infection and introduce new molecular tools for honey bee research.

2. Materials and Methods

1. *RG System for Three DWV Variants*

RG constructs used in this study were based on the cDNA clone of a recombinant DWV variant VDV-1-DWV-No-9 (GenBank HM067438.1). VDV-1-DWV-No-9 sequence was rescued using samples from a *Varroa*-infested honey bee colony from Warwick-HRI apiary (2014), and a cloned full-length cDNA was incorporated into a plasmid vector containing all required elements for the transcription of the viral RNA. VVD RG clone is identical to the source VDV-1-DWV-No-9, with the exception of two nucleotide substitutions (positions 5277 and 5280 in the VVD clone cDNA, which corresponds to positions 5124 and 5127 in the GenBank HM067438.1 sequence lacking the very 5'-end of the virus genome), resulting in creation of an *HpaI* restriction site. In order to obtain VDD and VVV constructs bases 2727–4888 (capsid proteins encoding region) and 4885-9783 (non-structural proteins encoding region) were replaced with corresponding DWV A and B fragments respectively (Figure 1, supplementary text S1). Inserts encoding DWV A capsid proteins and DWV B non-structural proteins were based on published data [14,15] and obtained by custom gene synthesis (IDT, Leuven, Belgium). The replacement sequences were amplified with High Fidelity Phusion DNA polymerase (Thermo Fisher Scientific, Winsford, UK) and incorporated into the initial construct using NEBuilder Hifi assembly reaction (New England Biolabs, Hitchin, UK). The nomenclature used for the corresponding virus clones in this study is as follows: "VDD" —type A DWV (protein coding part), "VVV"—type B, "VVD"—B/A recombinant (see Figure 1 for details). New restriction sites were introduced into each plasmid either by standard site-directed mutagenesis or by including the modification into the synthetic sequence: *HpaI* (5275-5280 nt in VVD variant), *Kpn2*I and *AvrII* (2751-2756 and 4884-4889 nt in VDD variant), and *AvrII*, *PflFI* and *BglII* (4884-4889, 6087-6095, and 9783-9788 nt, respectively, in VVV variant).

EGFP-encoding RG constructs (DWV_E) were obtained by NEBuilder Hifi assembly (New England Biolabs, Hitchin, UK) reaction applied to previously made RG constructs. EGFP sequence flanked by 32 additional nucleotides (ATGGATAACCCT from 5' and GCAAAACCAGAG from 3' end) resulting in duplication of the protease cleavage site (amino acid composition of the site is AKPEMDNP [14]) was inserted between nucleotides 1785–1786 of DWV cDNA. All plasmids were subjected to Sanger sequencing and the results were aligned with data from *in silico* cloning simulation. Full cDNA sequences of DWV clones used in this study are available in Text S1 (Supplementary Materials) and online (GenBank accession numbers: DWV-VDD - MT415949, DWV-VVD - MT415950, DWV-VVD_truncated - MT415951, DWV-VVV - MT415952, DWV-VDD-eGFP - MT415948, DWV-VVD-eGFP - MT415953).

2. *Viral RNA Synthesis*

DWV RNA was synthesized using linearized plasmid templates. Full length and truncated templates were linearised with *Pme* I (cutting at the end of the sequence encoding the poly-A tail) or *Nru* I (nt 9231 located within the sequence encoding the viral polymerase) respectively. Linearized DNA was purified by phenol-chloroform extraction and ethanol precipitation and T7 transcription was performed with T7 RiboMAX™ Express Large Scale RNA Production System (Promega, Southampton, UK) according to the manufacturer's protocol. In order to account for the stabilization effect of the poly-A tail truncated transcripts were subjected to an additional polyadenylation step with poly(A) Tailing Kit (Thermo Fisher Scientific, Winsford, UK). RNA transcripts were purified with GeneJet RNA Purification Kit (Thermo Fisher Scientific, Winsford, UK) using clean up protocol and eluted in RNAse free H_2O. All transcripts were analysed for integrity by gel electrophoresis and stored at −80 °C.

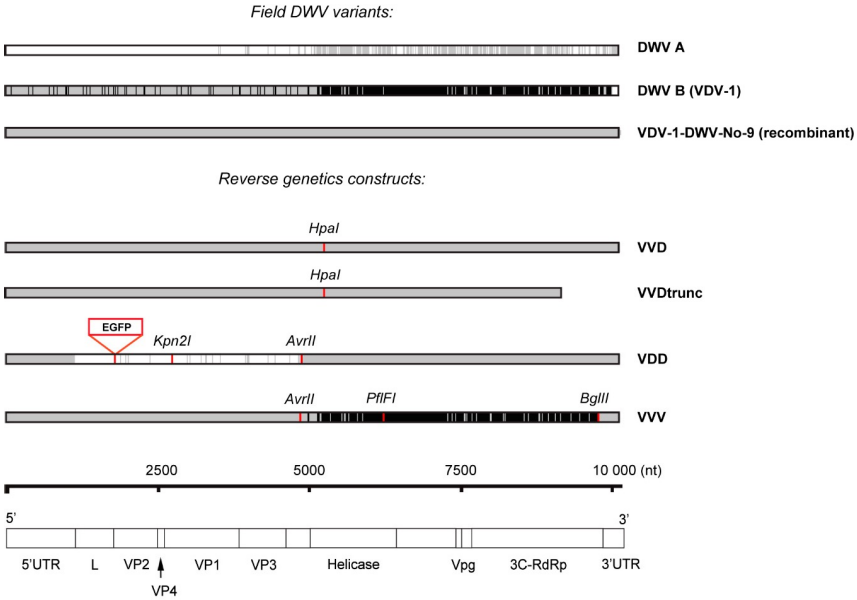


Figure 1. Sequence homology between field DWV type A (NC_004830.2), DWV type B (GenBank: AY251269.2), and recombinant VDV-1-DWV-No-9 (GenBank HM067438.1) variants and RG virus clones. VVD RG clone was prepared using full-length cDNA of VDV-1-DWV-No-9. VDD and VVV RG clones were obtained by replacing 5' and 3' parts of protein-encoding sequence of VVD construct with synthetic fragments homologous to corresponding parts of DWV A and DWV B genomes. DWV A and B specific sequences are shown in white and black respectively, genome regions identical to the recombinant clone sequence (VDV-1-DWV-No-9) are shaded in grey. Unique synonymous mutations introducing new restriction sites are indicated in red, as well as the position at the junction between L- and VP2 proteins coding sequence where EGFP insert was incorporated in order to obtain reporter-expressing DWV. Genomic RNA organization of DWV and encoded proteins are shown below with numbers (nt) indicating nucleotide position along the genome.

3. *Virus Stocks*

DWV stocks were prepared from honey bee pupae injected with *in vitro* transcribed RNA. Homogenized tissue was diluted with sterile PBS in 1:1 (*w:v*) ratio and centrifuged at 13 000× *g*, 4 °C for 10 min. The supernatant was sterilized by passing through 0.22 µM PES filter (Merck Millipore, Watford, UK) and treated with RNase A to destroy all non-encapsidated RNA. RNA was extracted from 100 µL of the virus stock using RNeasy kit (Qiagen, Manchester, UK) and analysed by RT-qPCR.

4. *Honey Bees*

All honey bee (*Apis mellifera*) brood in this study was obtained from the research apiary of the University of St Andrews. Hives used for sampling were regularly treated for *Varroa* with an appropriate miticide and routinely screened for DWV. DWV level in bees obtained from these hives was found to range within 10^2–10^6 genome equivalents (GE) per 1 µg of total RNA. No phenotypic evidence for other honey bee viruses being present was found within the colonies throughout the course of the studies. Pupae and larvae of the required age were collected from the comb and transferred to the incubator set at 34.5 °C and 90% relative humidity. Pupae were kept on folded sheets of filter paper, and larvae were transferred into 96-well plates with round bottom wells containing feeding diet (6% (*w/v*) glucose, 6% (*w/v*) fructose, 1% (*w/v*) yeast extract, and 50% (*w/v*) royal jelly in H_2O).

5. Varroa Feeding

Varroa mites were collected from the brood frames of infested colonies kept at the University of Aberdeen. Mites were placed in Petri dishes in groups of 10 with a single honey bee pupa and kept overnight in the incubator (34.5 °C, 90% humidity). On the next day, feeding packets containing artificial feeding diet and supplemented with DWV stock or equal amount of PBS were prepared. The feeding packets were prepared by wrapping a 175 µL drop of the diet containing RNase A treated DWV inoculate (prepared as described in Section 2.3). The diet is 75:25 holidic:locust haemolymph (the detailed description of the feeding diet to be published separately [31]). To prepare the diet the locust (fourth instar *Locusta migratoria*) haemolymph was frozen-thawed, heat-treated, and centrifuged for 30 s at 7500× *g*.

Mites were placed in groups of 5–10 mites per feeding packet with four replicates for each treatment and kept in the incubator for four days. At the end of incubation live mites were collected, snap-frozen in liquid nitrogen and stored at −80 °C until further analysis. Total RNA was extracted from pooled mite samples using the standard TriReagent protocol (Thermo Fisher Scientific, Winsford, UK).

6. RNA Injections and Virus Inoculations

RNA and virus injections into pupae were performed with insulin syringes (BD Micro Fine Plus, 1 mL, 30 G (Becton Dickinson, Oxford, UK)). 5 µL injections were introduced by inserting the syringe needle between the second and third abdominal segments of a pupa. All pupae were maintained on a warm heating plate during the injections.

0.5 and 5 µg (equivalent to 8.7×10^{10} and 8.7×10^{11} of DWV RNA copies) of *in vitro* transcribed RNA was injected individually into white-eyed honey bee pupae. Truncated VVD transcript injections were used as a negative control in RNA transcript injections. RNA-injected pupae were analysed at 72 h post-injection.

Injections of pupae with virus stocks were performed using the same technique as for RNA transcripts. Serial dilutions of DWV stocks in sterile PBS were prepared immediately before the injections. Mock control groups were injected with PBS only, while non-injected controls were left intact throughout the duration of the experiment.

For oral infection, larvae were placed into 96 well plates with a diet supplemented with DWV. Fresh diet without virus was added after 24 h or when all virus-supplemented diet was consumed.

2.7. RNA Extraction, Reverse Transcription and PCR

Total RNA was extracted individually from all bee samples, *Varroa* mites were analysed in pools of 5–10 according to the treatment group. Samples were homogenized using a Precellys Evolution instrument (Bertin Instruments, Montigny-le-Bretonneux, France). Total RNA was extracted with GeneJet RNA Purification Kit (Thermo Fisher Scientific, Winsford, UK) using a protocol adapted for vacuum manifold application. cDNA was prepared with qScript cDNA Synthesis Kit (Quanta Biosciences, VWR International Ltd, Lutterworth, UK) from 1 µg of total RNA following the manufacturer's protocol with both oligo dT and random hexamer primers included in the reaction mixture in the reaction volume of 20 µL.

All sequences of primers used in this study are shown in Table S1. Detection of DWV and honey bee *actin*, used as an internal RNA quality control, was carried out by end-point PCR with *Taq* DNA polymerase (New England Biolabs, Hitchin, UK) and 2 µL of cDNA. DWV_RTPCR primers were designed to detect all three DWV variants under study. To amplify the cDNA regions containing restriction site tags in VDD and VVV virus variants Kpn2I_F/R and PflFI_F/R primer sets were used respectively. PCR cycling conditions were 30 cycles of 95 °C (15 s), 55 °C (15 s), 68 °C (2 min) with an initial 95 °C step (30 s), and a final extension at 68 °C (5 min). PCR samples were analysed on a 1% agarose gel stained with ethidium bromide. When required, DWV PCR products were subjected to restriction digest prior to loading on the gel.

The quantification of DWV genome copies was performed by SYBR-Green Real-Time Quantitative PCR (qPCR). Reactions were carried out in a C1000 Thermal Cycler (Bio-Rad Laboratories, Deeside, UK) using Luna Universal qPCR master mix (New England Biolabs, Hitchin, UK), 0.25 µM forward and reverse DWV_qPCR primers, and 2 µL of cDNA with the following thermal profile: 1 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 30 s at 60 °C with a final post-amplification melting curve analysis step. Accumulation of DWV_E RNA was assayed by a set of primers amplifying a junction region between EGFP and viral sequence (primers EGFP_qPCR_F and VDD_VP2qPCR_RP or VVV_VP2qPCR_RP depending on the virus variant used). DWV and DWV_E titers were calculated by relating the resulting Ct value to the standard curve generated by performing qPCR from a serial dilution of the cDNA obtained from 1 µg of VVD or DWV_E RNA transcript respectively.

8. Negative Strand Assay

Strand-specific detection of DWV RNA was performed as described earlier [10]. Briefly, 1 µg of total RNA was used in reverse transcription reaction carried out with Superscript III reverse transcriptase (Invitrogen, Thermo Fisher Scientific, Winsford, UK) and the adapter extended primer DWV(-RNA)_RT designed to anneal to the negative strand RNA of DWV. The PCR step was carried out by *Taq* DNA polymerase (New England Biolabs, Hitchin, UK) using a forward primer identical to the adapter sequence (primer 388) and DWV(-RNA)_RT R primer. PCR was run for 35 cycles in the same conditions as described above.

9. Microscopy

All imaging to detect EGFP signal in samples infected with DWV_E was done using Leica TCS SP8 confocal microscope with 10× HC PL FLUOTAR objective (Leica Biosystems, Newcastle, UK).

10. Cryosection of Larvae Samples

Live larvae were washed with increasing concentrations of aqueous ethanol solution and fixed in 4% paraformaldehyde in PBS for 2 h at 4 °C. Fixed larvae were allowed to sink in 30% sucrose in PBS, mounted in NEG-50 Frozen Section Medium (Thermo Fisher Scientific, Winsford, UK) and subjected to microsectioning on CM1860 Cryostat (Leica Biosystems, Newcastle, UK). Sections of 50–80 µm thickness were placed on Superfrost Plus microscope slides (Thermo Fisher Scientific, Winsford, UK), mounted in ProLong Gold antifade medium, and analysed by microscopy.

3. Results

1. Infectivity of Parental and Recombinant DWV Variants in Honey bee Brood

To examine the biology and pathogenesis of different strains of DWV we prepared near-clonal virus stocks from honey bee pupae injected with *in vitro* transcribed viral RNA from relevant cDNAs. All RG templates were derived from an infectious cDNA of a recombinant DWV variant (GenBank HM067438.1) [26]. The 5′ UTR, capsid-coding, and 5′ part of the presumed helicase-coding region of this genome are 99% identical to VDV-1 (DWV B, AY251269.2) with the remainder of the genome ~97% identical to DWV-A (NC_004830.2) henceforth this variant was designated VVD. The VVD cDNA was modified by replacing the capsid-coding region with the analogous sequence from DWV A to generate a VDD cDNA, and by replacement of the 3′ non-structural protein coding region with DWV B sequence to generate VVV cDNA (Figure 1) [32]. Each cDNA carried unique synonymous genetic tags (restriction sites) allowing their unambiguous identification. The protein coding part of the resulting VDD and VVV constructs was 98.51% and 99.44% identical to the reference DWV A and DWV B field genomes respectively (99.65% and 99.79% identity at the encoded protein level).

Direct inoculation of white-eyed honey bee pupae with *in vitro* synthesised RNA enabled recovery of genetically tagged virus at high efficiency. Using qPCR analysis, the average viral load in RNA-injected pupae was shown to be 1.5×10^{10} DWV GE per 1 µg of total RNA (3×10^{12} GE per

pupa, Figure S2) compared to ≤10^5 DWV GE/µg of RNA in mock-inoculated pupae. Therefore at least 99.999% of the virus preparations from injected pupae originated from the cDNA-derived injected RNA. In each case, the identity of the virus was verified by RT-PCR and the presence of the unique restriction endonuclease site engineered into the cDNA was confirmed.

We investigated the infectivity of RG-recovered VDD, VVV and VVD variants by inoculation of white-eyed honey bee pupae using serial dilutions of virus stocks and analysed all samples individually 24 h post-inoculation by RT-qPCR (Figure 2a and Figure S3). Under these conditions, we observed robust accumulation of virus following inoculation of just 1 GE, with increasing yields of virus when 10–1000 GE were injected. At the lowest level of inoculum VDD appeared to accumulate slightly more slowly (Tukey's multiple comparisons test, p = 0.0319 for VVD vs VVD in 1 GE injections groups), but there was no discernible difference in yield of the three variants when ≥10 GE were inoculated (p > 0.05). Incubation of pupae for a further 24 h resulted in virus yields of ~10^{10} GE/µg of RNA with no notable difference between the three DWV variants tested. This shows that RG-generated DWV is infectious and the kinetics of virus replication is very rapid, with the genome being amplified ~10^8–10^{13} times within 48 h of injection of pupae (minimum yield of RNA/pupa was 200 µg).

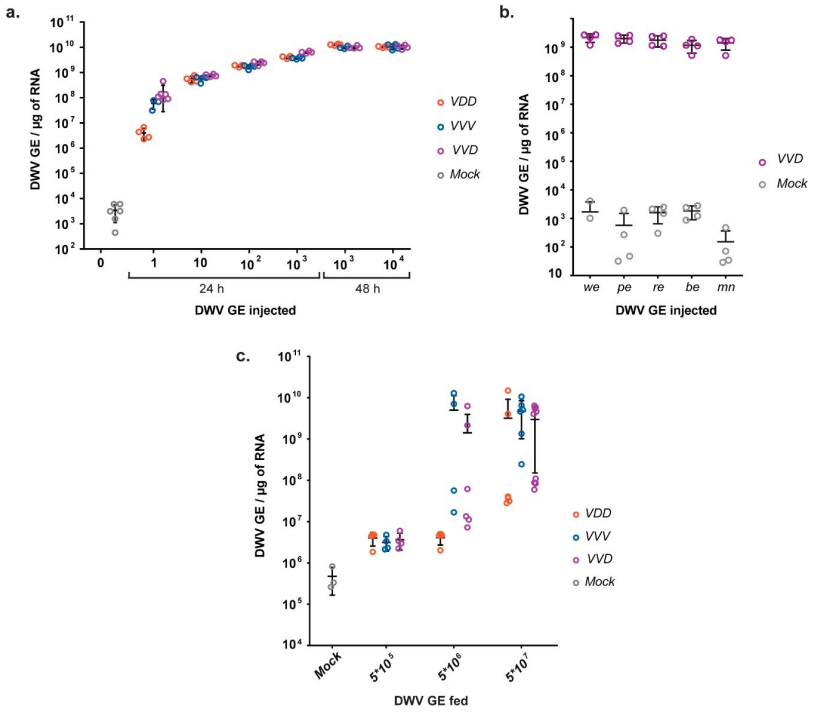


Figure 2. Inoculation of the honey bee brood with RG-DWV. (**a**) RT-qPCR analysis of DWV level in honey bee pupae (injected at white-eyed stage) 24 h and 48 h post-injection with different amounts of DWV variants, "Mock"—uninoculated pupae; (**b**). RT-qPCR analysis of honey bee pupae, injected with 10^2 GE of VVD at different stages of pupal development: white-eyed ("we"), pink-eyed ("pe"), red-eyed ("re"), blue-eyed ("be"), and at the start of full melanisation ("mn"). Pupae were sacrificed 24 h post-inoculation; (**c**) RT-qPCR analysis of DWV accumulation levels in honey bee larvae 48 h after feeding with DWV variants at different concentrations. Each value corresponds to an individual pupa or larva analysed, and error bars show mean ±SD.

Varroa feed throughout the development of the pupa and possibly during the phoretic phase of the life cycle. We investigated whether pupal development phases were partially or totally refractory to DWV replication by inoculating morphologically distinct stages with 10^2GE of VVD DWV (Figure 2b). In each instance DWV levels reached ~10^9GE/µg within 24 h post-injection, indicating that all stages of pupal development are apparently equally capable of supporting DWV replication.

Having demonstrated pupal susceptibility to directly inoculated virus we then examined virus transmission to developing larvae via the oral route. First instar larvae were individually fed a diet containing dilutions of RG-derived DWV, and the virus was detected and quantified after 48 h. Although RG-derived DWV was detectable when 5×10^5 GE were fed, virus replication to levels distinctly higher than the inoculum required an input of at least 5×10^6GE, with VDD again slightly slower than the other variants tested (Figure 2c). Uninoculated control larvae contained ~10^5DWV GE/µg of RNA, presumably reflecting virus acquired during initial feeding in the hive or vertically from the queen. All larvae in DWV-fed groups contained elevated levels of viral RNA compared to the control group 48 h after inoculation (Tukey's multiple comparisons test, $p = 0.0372$ for VVD vs mock, $p = 0.0205$ for VDD vs mock, for 5×10^5GE feeding groups). High DWV levels (up to 10^{10}GE/µg of RNA) were detected in 5×10^6GE fed groups for VVV and VVD inoculates and for all three variants in 5×10^7GE fed larvae, clearly indicating virus replication. It was notable that a proportion of larvae showed little or no amplification of DWV over the fed amount suggesting, in comparison to inoculation via injection, a significant barrier to the development of a productive infection may exist for virus acquired *per os*.

3.2. Replication of RG DWV in Varroa Destructor

There are conflicting reports on the replication of DWV in *Varroa*. A confounding issue in some studies is the pre-existing presence of DWV within the mite, and the potential presence of contaminating DWV genomic negative strands (a marker of replication) in either the mite or the inoculum. To address this, we investigated the replication of RG-derived genetically tagged VVD in *Varroa* using an *in vitro* feeding system [31] containing no honey bee-derived material (Figure 3a).

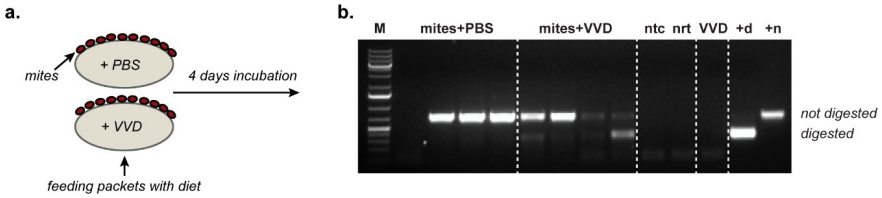


Figure 3. DWV replication in *Varroa* mites. (**a**) *Varroa* feeding experimental setup using feed packets with artificial diet supplemented with RG VVD DWV in PBS or PBS only (control); (**b**) 1% agarose gel with *HpaI*-digested PCR products amplified using a DWV negative strand-specific RT-PCR assay. Specific (-)RNA PCR products from *Varroa* mites fed with artificial diet supplemented with VVD virus stock ("mites+VVD") or PBS ("mites+PBS"); four groups of 10 mites were used for each feeding condition - each lane corresponds to a pooled sample of mites from one group (fed on the same diet packet), "ntc" and "nrt"—no template and no RT controls, respectively, "VVD"—negative strand-specific RT-PCR analysis of the virus stock used for mite feeding, " + d" and "+n"—digested and undigested positive PCR controls, "M"—molecular weight DNA marker.

VVD RG-derived virus stock was amplified in honey bee pupae, extracted and treated with ribonuclease A to remove any non-encapsidated viral or cellular RNA. *Varroa* were maintained on artificial feed packets supplemented with 1.75×10^9GE VVD (10^7GE/µl).

After four days, mites were harvested, total RNA was extracted and screened using a DWV negative strand-specific RT-PCR assay. All pooled mite samples from DWV-fed groups produced a DWV-specific product; three out of four mite pools from the mock-inoculated groups also produced a

product as expected since these were harvested from hives with high *Varroa*, and consequently high DWV, levels. Upon digestion with the relevant restriction enzyme (*HpaI*), this product was partially cleaved in the VVD groups only, indicating that at least some of it originated from the negative-strand replication intermediate of VVD (Figure 3b). Importantly, no negative strand RNA of DWV was detected in the RNAse-treated input virus stock used for these feeding experiments. The absence of negative RNA strands in the viral input combined with the specific detection of RG-tagged DWV negative strand RNA in mites that have fed provides evidence for the replication of VVD DWV in *Varroa destructor*.

3.3. Pathogenicity of DWV Variants in Honey Bees

DWV infection in honey bees is known to result in a range of developmental abnormalities, the most prominent of which is malformation or arrested unfolding of wings [14,33–35]. Since RG-recovered DWV replicated to high levels in injected pupae, we went on to determine whether pupae incubated until eclosion also exhibited characteristic developmental defects. DWV inoculation at the white-eyed pupal stage, mimicking the route and timing of *Varroa* transmission [9,30,36], resulted in 75–80% of bees developing overtly deformed wings (Figure 4b). Analysis by qPCR showed no difference in the final DWV titers between highly deformed and apparently phenotypically normal eclosed workers (Figure 4a). In the same experiment we observed no differences in either virus levels or the proportions of normal vs deformed workers with the three different RG DWV variants tested. These findings indicate that RG-derived DWV is pathogenic when directly inoculated into developing worker pupae and results in symptoms and virus titres that are similar to those seen in *Varroa*-exposed pupae [9]. In addition, this study demonstrated that phenotypically normal eclosed workers can have virus levels indistinguishable from those with deformed wings. This confirmes that wing deformities are not an inevitable consequence of high levels of DWV replication [23,30].

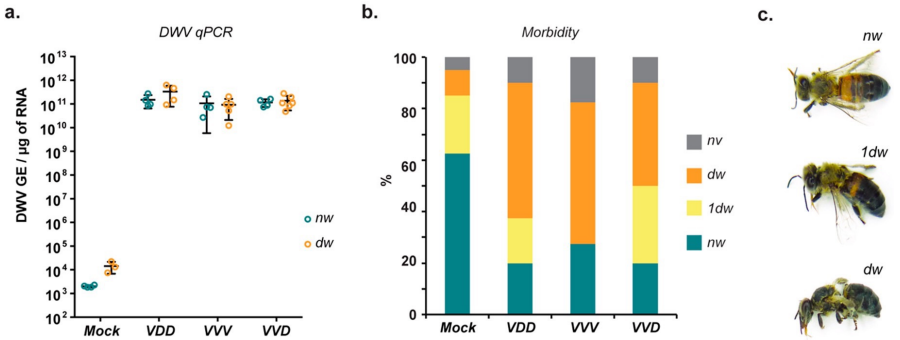


Figure 4. Morbidity of DWV variants in honey bee pupae. (**a**) RT-qPCR analysis of DWV levels in honey bees developed from pupae injected with the indicated virus variants and displaying normal ("nw") or deformed wing ("dw)" phenotype; no significant difference in DWV levels was found between deformed and non-deformed bees (Sidak's multiple comparisons test, $p = 0.1307$ (VDD), $p = 0.9975$ (VVV), $p = 0.9911$ (VVD)); (**b**) Percentage of visually normal ("*nw*"), partially deformed ("1dw"), deformed ("*dw*") and non-viable ("*nv*") honey bees eclosed from pupae injected at the white-eyed stage with 10^2 GE of the indicated DWV variants ($n = 40$ for each group); (**c**) Newly emerged honey bees with different phenotypes of morbidity: "nw"—non-deformed bee with fully unfolded wings; "1dw"—bee with only one normal wing; "dw"—bee with deformities in both wings.

In the field non-*Varroa* transmission of DWV (*per os* or vertically) results in predominantly asymptomatic (also referred to as covert) infection, with no apparent phenotypic deformities. We tested the morbidity in honey bees infected orally with clonal DWV by feeding larvae with

a virus-supplemented diet, allowing pupation and eclosion and scoring viability and the presence of overt symptoms. Although lower levels of virus ($\leq 10^5$ GE) in the diet again failed to establish robust replication in the experimental groups, the feeding of 2×10^7 GE of RG-derived DWV to honey bee larvae (delivered as single time feeding at the first instar larval stage) resulted in high morbidity rates (Figure S4). In larval groups infected with DWV within 24 h of hatching from eggs, all eclosed adult honey bees had deformities ranging from malformation of one or both wings, abdominal bloating, discoloration, and dwarfism. A proportion also exhibited arrested development at the pupal stage and failed to eclose. The majority of the developed bees revealed high virus levels of up to 10^{12} GE/µg of RNA (Figure S4). Notably, the laboratory-reared honey bee brood revealed relatively high levels of wing deformities present in mock-inoculated pupae and PBS-fed larvae upon eclosion (Figures 4b and S4). It is known that alterations in incubation conditions of pupae at key stages of development can also result in wing deformities [30], suggesting that wing morphogenesis is a particularly sensitive stage of development.

3.4. DWV Localization in Infected Honey Bee Brood

The tropism and pathogenesis of DWV remains poorly understood. Relatively little is known about the sites of virus replication and whether differences in tissue tropism after oral or mite-mediated transmission account for the appearance of symptomatic and asymptomatic forms of disease. To facilitate virus localization in experimentally inoculated bees, we developed an EGFP-expressing chimeric DWV (designated DWV_E), analogous to that previously reported for DWV A [37]. Briefly, the cDNA encoding the viral polypeptide was modified by the insertion of the EGFP cDNA at the junction between the Leader and VP2 capsid protein coding regions similarly to the system developed earlier for the poliovirus [38]. The in-frame EGFP sequence was flanked with partial duplications of the predicted $3C^{pro}$ proteolytic cleavage site [14] with the intention of co-translational processing and release of the EGFP protein (Figure 5a).

Injection of *in vitro*–synthesised DWV_E RNA into white-eyed honey bee pupae resulted in distinct green fluorescence in the head, thorax and abdomen observable within 20 h post inoculation (Figure 5b). The fluorescence signal could be readily detected throughout inoculated pupae at least seven days post-injection (Figure S5 and Supplementary Video 6), implying that this represents a robust experimental system to investigate virus localization *in vivo*. In preliminary studies we went on to investigate the tissue tropism of reporter-encoding derivates of VDD, VVV and VVD variants of DWV in inoculated pupae and observed no discernible differences; therefore, subsequent analysis was performed using the EGFP derivative of VDD.

Confocal microscopy allowed the visualisation of individual foci of virus replication, identified as distinct punctate fluorescence, in a range of tissues throughout the inoculated developing pupae. At 22 h post inoculation the fluorescent signal was apparent in the head, parts of the gut (the crop, ventriculus, small gut and rectum), but was largely absent from the Malphigian tubules and thoracic muscles (Figure 5c). Since overt DWV is predominantly associated with wing deformities we looked in detail at the developing wings in inoculated pupae. At 22 h post-inoculation punctate fluorescence was clearly visible in the wings. White-eyed pupae inoculated with DWV_E and incubated until eclosion were also examined. As before (Figure 4), all had significant levels of virus replication irrespective of the presence of overtly deformed wings. The wings of 90% of all injected pupae and honey bees developed from these pupae, regardless of their deformity status, were shown to express EGFP upon eclosion indicative of sites of DWV_E replication (Figure S6).

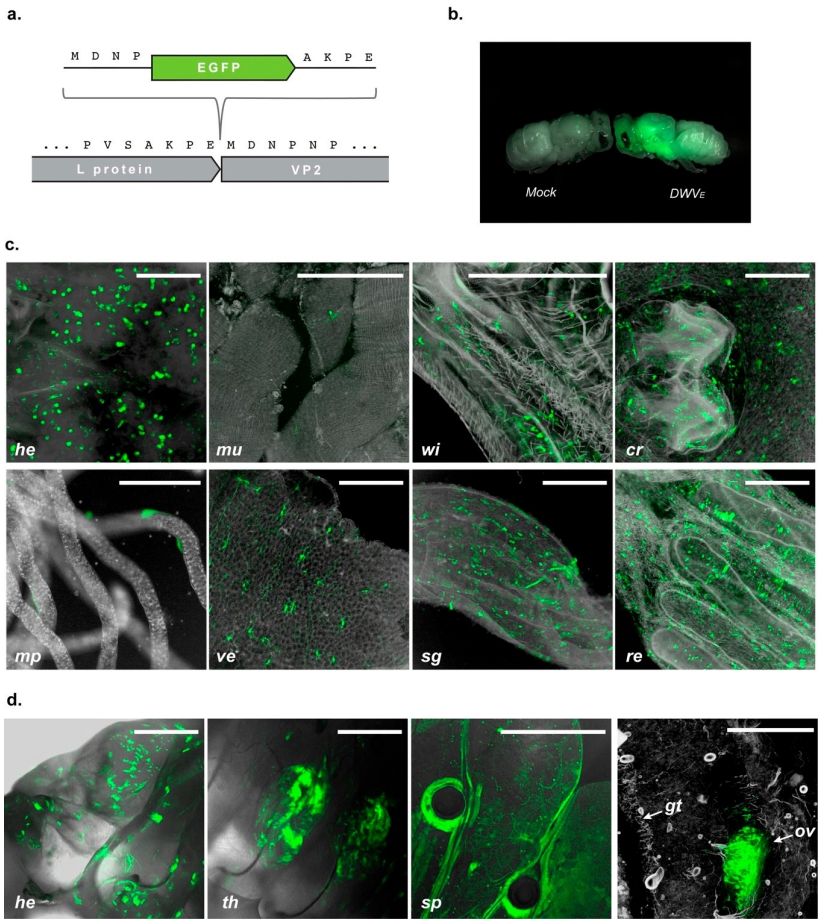


Figure 5. Localization analysis of the DWV-produced EGFP signal in infected honey bee brood. (**a**) Schematic representation of DWV_E RG construct design; only part of the viral genome, where the insert was placed, is shown; amino acid sequences are presented using single-letter code. (**b**) Combined image of fluorescent and white-field photo showing not-injected ("Mock") and DWV_E injected pupae. (**c**) Confocal microscopy analysis of EGFP signal localisation in DWV_E infected honey bee pupae 22 h post-inoculation: dorsal side of the head with removed cuticle ("he"), thorax muscles ("mu"), wing rudiment ("wi"), crop ("cr"), Malpighian tubules ("mp"), ventriculus ("ve"), small gut ("sg"), and rectum ("re"); composite of fluorescent signal z-stack and inverted white-field image are shown for convenience of interpretation; scale bars correspond to 500 µm on all panels except "mp", where 200 µm is shown. (**d**) Confocal microscopy analysis of living larva (first three panels) and cryosection of infected larva (right panel) sampled 6 days after feeding with 5×10^7 GE of DWV_E inoculate: head ("he"), wing rudiments in thoracic segments ("th"), spiracle openings ("sp"), sagittal section of larva showing rudimental ovary ("ov") and part of the midgut wall ("gt"); composite of fluorescent signal z-stack and white-field image (inverted on the cryosection panel) are shown; scale bars correspond to 500 µm.

To trace DWV distribution following oral acquisition of virus we fed developing larvae with 5×10^7 GE of DWV_E purified from previously inoculated pupae and analysed larval fluorescence 4 and 6 days later. Samples infected with DWV_E exhibited extensive fluorescence; external observation localized this to the larval head, two thoracic segments containing wing rudiments, the caudal part of the body and to cells surrounding the spiracle openings (Figure 5d, Supplementary Videos 1–5). Additional sites of EGFP localization were apparent after cryosectioning samples, confirming DWV_E infection in the larval head, tissues of thoracal and caudal abdominal segments (Figure S7), and in the developing reproductive system (Figure 5d).

4. Discussion

The inexorable rise in global honey bee colony numbers masks increasing levels of colony losses, which are regularly reported to exceed 30% per annum and predominantly occur during the winter in temperate regions [39–41]. For almost a decade the major cause of these losses, the ectoparasitic mite *Varroa destructor* and the smorgasbord of viruses it transmits, has been well known [7]. Of these, the most important virus associated with overwintering colony loss is DWV. Although DWV is a single stranded, positive sense RNA virus—a group that includes poliovirus which has been dissected at the molecular level for almost four decades—molecular methods to study its biology have developed slowly. One of the major reasons for this is the absence of a usable *in vitro* cell culture system enabling virus propagation [42]. The recent development of an RG system [30] allowing the recovery of infectious virus from a cDNA for the type A variant of DWV provided the first tractable approach to a better understanding of the biology and pathogenesis of DWV.

We report here the extension of the genetic tools to study the biology of DWV including the type B variant (also designated VDV-1) and a recombinant that has previously been reported to predominate in *Varroa*-infested colonies [9,26]. We additionally demonstrate how the RG approach can be exploited to study host-pathogen and host-vector-pathogen interactions, and—with the development of reporter gene-expressing variants of the virus—to study tissue distribution in developing honey bee larvae, pupae, and eclosed adults.

Using standard molecular cloning techniques combined with *in vitro* gene synthesis, we constructed infectious cDNAs for type A (protein coding sequence), type B and a recombinant variant of DWV, recovered molecularly tagged virus after RNA inoculation and investigated the kinetics of virus replication in honey bee pupae. Replication was rapid, amplifying to ~10^8GE/µg of RNA within 24 h, and plateauing at ~10^{10}GE per µg of total cellular RNA within 48 h. There were no major differences between the replication rates of the three DWV variants under study, and all stages of pupal development tested appeared equally susceptible to virus infection (Figure 2). This implies that, although pupae are only naturally exposed to mite-borne virus upon capping, even very low amounts of virus inocula introduced by the mite have the capacity to replicate to very high levels before eclosion. The rapid kinetics of DWV replication also means that all progeny mites produced by a single pupa are likely to almost exclusively carry the virus population representative of that introduced by the foundress mite, as these (and not the endogenous virus population) are what are amplified following direct injection.

Horizontal virus transmission within the colony also occurs during larval feeding by nurse bees. We show that larvae exhibit higher resistance to infection with clonal DWV inoculates *per os* when compared with the susceptibility of pupae to injected virus. Greater than 10^6-10^7 GE of virus was needed to reliably infect more than 50% of larvae with any DWV variants tested (Figure 2c). This result indicates that high levels of DWV present in the colony due to amplification by *Varroa* parasitized individuals can boost the development of a symptomatic infection state both through pupal infestation by mites and feeding of larvae by diseased nurse bees. A better understanding of this will require the levels of virus transmitted orally, or present in bee bread fed to developing larvae, to be determined. Interestingly, inoculation with VDD clone containing DWV type A structural proteins sequence required higher virus concentrations to achieve efficient infection in larvae. This result may

indicate that the 5' part of the protein encoding sequence of DWV B genome enables higher infectivity either due to enhanced binding and penetration into the target host cells or via other interactions caused by clone-specific secondary or tertiary structures of viral RNA.

As honey bee pupae appear very sensitive to infection by direct injection (recapitulating transmission by *Varroa*), any virus replication in the mite would likely guarantee the transmitted dose would significantly exceed the ID_{50}. The viral load in mites has been reported to be high [43] though this may reflect the level of virus in the preceding host pupa rather than replication *per se*. Previous analysis of virus replication in *Varroa* has produced conflicting results. Mass spectroscopy studies failed to detect viral non-structural proteins implying that DWV may not replicate in the mite [44]. Conversely, although detection of viral negative strand replication intermediates [15,45] in mites may indicate replication, they may also reflect carry-over from the previous pupal feed. To address this, we maintained mites *in vitro* on a diet containing no honey bee-derived components. The only DWV virus (and viral negative strand RNA) present would therefore be derived from the last pupa the mites had fed on, together with any subsequent replication in the mite. We supplemented the diet with the RG-derived VVD variant of DWV carrying a unique genetic tag allowing its unambiguous identification. The presence of genetically tagged negative strand RNA of DWV in the pooled mites samples fed on this diet demonstrates that VVD does replicate in *Varroa* and that newly acquired virus can replicate in the presence of a pre-existing virus population in the mite (Figure 3). The latter point is significant as it suggests that the virus population in the mite reflects its historical diet from successive infested pupae, potentially influenced by any differential virus replication in the mite.

Although mite transmission of DWV is associated with overt symptoms such as wing deformities these are not the inevitable consequence of high viral loads [23,30]. Approximately 25% of inoculated pupae had normal wings on eclosion, similar to another recent report [23], despite having viral loads in excess of 10^{10} GE/µg RNA. While apparently developmentally normal, evidence indicates that workers with high viral loads are impaired in foraging ability, cognitive functions, and die prematurely [7,10,46,47]. In contrast to the pupal injections where a significant proportion eclosed and appeared developmentally normal despite high levels of virus replication, no virus-fed larvae which eclosed exhibited normal wings, although >50% reached the adult stage (Figure S4). These observations suggest that the earlier stages of honey bee development may be more susceptible to DWV-mediated damage. Additionally, this signifies that the occurrence of high DWV levels in bees with normal wings in the field may be due to *Varroa*-mediated transmission, while the outcome of oral infection in larvae depends on the amount of virus ingested and results in either benign asymptomatic infection or in high morbidity due to high levels of DWV replication and consequential developmental damage. Apart from horizontal transmission addressed in this study, DWV can be transmitted vertically from an infected queen [48]. Further studies are required to elucidate the exact impact of this route on the phenotypic outcome of the infection.

It remains unclear what determines whether virus exposure results in overt disease or asymptomatic infection [43] although results presented here suggest that the timing of infection (larvae *vs*. pupae) is probably critical. Analysis of whether infection of all stages of pupal development are as likely to result in wing deformities may be informative in this regard, though *Varroa*-mediated virus transmission will initially occur when the foundress mite feeds on the just-capped pre-pupa. One of the factors which can potentially influence the development of overt disease are the sites of virus replication in the developing honey bee. This may result in direct cytopathicity and tissue damage or indirectly by dysregulation leading to damage at remote locations. To address the tissue tropism of DWV, we constructed a modified virus genome co-translationally expressing the green fluorescent protein. Inoculation of pupae, or ingestion by larvae, resulted in distinct fluorescence in a range of organs and tissues. In fed larvae, the EGFP signal accumulated in thoracic segments where the developing wing buds are located [49], and infection of these presumably accounts for the characteristic symptoms of DWV infection. It was notable that larvae fed with large amounts of DWV invariably developed with malformed wings. The developing ovary of worker larvae also appears

to be a site of virus replication. Worker ovaries—other than in laying workers—never produce eggs, so further studies will be required to determine whether the ovary of developing queens also shows evidence of virus replication. This, together with the known presence of DWV in drone semen [50], would presumably explain the vertical transmission of virus.

Our observations indicate that DWV selectively targets organs most favourable for its oral and vertical transmission. Sites of DWV replication colocalize with exocrine glands, which are responsible for the secretion of larval diet components, and digestive tract tissues, from which it can presumably be transmitted in faeces [51] and, following regurgitation, orally. These results support previous studies using anti-capsid antibodies or riboprobes targeting the virus genome [30,35] where DWV-specific signals were found in honey bee brains, exocrine glands, midgut, fat bodies, and reproductive organs [30,35,52,53]. Since the emergence of *Varroa* as a vector the oral route no longer plays the decisive role in DWV spread, but the observed tropism of DWV infection indicates the initial evolutionary trait used by the virus. Robust virus replication in the primary tissues that enable subsequent transmission often results in spillover to secondary sites that are permissive for virus infection, but that offer no further route to a new host, e.g., the neurotropism of the faecal-orally transmitted poliovirus [54]. It is therefore unsurprising that DWV infection was additionally found in a range of sites in the developing larva or pupa, including the spiracles and nascent wing tissues. The availability of EGFP-markers for sites of virus replication will enable further *in vivo* time course studies of virus dissemination and facilitate host-vector transmission studies. More generally, the ability to introduce a 'payload' to the virus genome may also be exploited by using genetically modified DWV as a virus-based gene delivery system [37]. Our results clearly demonstrate that DWV is abundant in many tissues in honey bees and a EGFP, or similar reporter, would enable tissue-specific RNAi responses to be quantified and optimized. Likewise, reporter-expressing derivatives of DWV have potential in tropism and transmission studies in other species in which the virus is known or suspected of replicating, including *Varroa* or *Bombus* [32] or, more speculatively, as gene delivery vectors for mite control.

In many single stranded positive sense RNA viruses, the development of reverse genetic systems was facilitated by a good understanding of virus replication *in vitro*. With no system for propagating DWV *in vitro* these studies have had to be conducted *in vivo*. We show that the availability of a reverse genetic system allows the kinetics of virus replication in larvae and pupae to be examined. Comparative studies demonstrate that the two predominant variants of DWV essentially replicate equivalently, and that the symptomatic outcome of DWV infection is not linked to a particular genotype of the virus nor to the route of transmission. Unique genetic tags introduced to the genome enabled discrimination of input from the endogenous virus population and provide unequivocal evidence for virus replication in *Varroa* following detection of *de novo* synthesised negative-sense viral RNA. A viable GFP-tagged DWV genome supplies the basis for a better understanding of virus tropism and pathogenesis in infected brood and will provide a useful tool for elucidating the mechanisms behind symptomatic DWV infection in developing honey bees.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4915/12/5/532/s1, Figure S1. Accumulation of field DWV in honey bee samples depending on performed treatment; Figure S2. DWV accumulation in honey bee pupae injected with RG RNA transcripts; Figure S3. Inoculation of honey bee pupae with highly diluted RG-DWV.; Figure S4. Morbidity of DWV variants in honey bees developed from infected larvae; Figure S5. EGFP expression in honey bee infected with DWV_E; Figure S6. DWV localisation in honey bee wings; Figure S7. EGFP signal localisation in honey bee larva infected with DWV_E; Table S1. PCR primers used in this study, Videos S1–S6: Confocal microscopy imaging of live honey bees infected with DWV_E. Text S1: Full cDNA sequences of VDD, VVV, VVD, and VDD-EGFP (DWV_E) reverse genetics clones.

Author Contributions: Conceptualization, D.J.E., A.S.B., and O.N.G.; methodology, O.N.G., L.W., K.B.-C., E.M.C., and C.R.C.; validation, O.N.G and L.W.; formal analysis, O.N.G. and L.W.; investigation, O.N.G., L.W., and K.B.-C.; resources, D.J.E; data curation, D.J.E., O.N.G., and L.W.; writing—original draft preparation, O.N.G. and D.J.E.; writing—review and editing, D.J.E., L.W. E.M.C., and A.S.B.; visualization, O.N.G.; supervision, D.J.E. and A.S.B.; project administration, D.J.E.; funding acquisition, D.J.E., A.S.B., E.M.C., and C.R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grant funding from Biotechnology and Biological Sciences Research Council (BBSRC): BBSRC BB/M00337X/2 and BB/I000828/1. C.R.C. was supported by a KTN BBSRC CASE studentship BB/M503526/1 (http://www.bbsrc.ac.uk). C.R.C was part-funded by the Scottish Beekeeping Association (https://www.scottishbeekeepers.org.uk/) and the Animal Health and Welfare program by the Scottish Government. E.M.C. was supported by the Veterinary Medicines Directorate, Department for Environment Food & Rural Affairs (Project #VM0517) (https://www.gov.uk/government/organisations/veterinary-medicines-directorate).

Acknowledgments: We would like to express our gratitude to Robert J Paxton (Martin Luther University Halle-Wittenberg) for providing a reference sequence for the construction of DWV-B RG system, Marcus Bischoff and Gill McVee (University of St Andrews) for helping with microscopic imaging set up, Javier Tello (University of St Andrews) for help and advice on sample cryosectioning, Ivan Gusachenko (University of St Andrews) for obtaining a photo of EGFP-expressing pupa, Ashley Pearson (University of St Andrews) for assistance in molecular biology assays, and Kirsten Bentley (University of St Andrews) for the high practical and moral support throughout the project.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Roberts, J.M.K.; Anderson, D.L.; Durr, P.A. Absence of deformed wing virus and Varroa destructor in Australia provides unique perspectives on honeybee viral landscapes and colony losses. *Sci. Rep.* **2017**, *7*, 6925. [CrossRef] [PubMed]
2. Beaurepaire, A.; Piot, N.; Doublet, V.; Antunez, K.; Campbell, E.; Chantawannakul, P.; Chejanovsky, N.; Gajda, A.; Heerman, M.; Panziera, D.; et al. Diversity and Global Distribution of Viruses of the Western Honey Bee, Apis mellifera. *Insects* **2020**, *11*, 239. [CrossRef]
3. Neumann, P.; Carreck, N.L. Honey bee colony losses. *J. Apic. Res.* **2010**, *49*, 1–6. [CrossRef]
4. Van Der Zee, R.; Gray, A.; Pisa, L.; De Rijk, T. An observational study of honey bee colony winter losses and their association with Varroa destructor, neonicotinoids and other risk factors. *PLoS ONE* **2015**, *10*, 1–25. [CrossRef] [PubMed]
5. Martin, S.J. The role of varroa and viral pathogens in the collapse of honeybee colonies: A modelling approach. *J. Appl. Ecol.* **2001**, *38*, 1082–1093. [CrossRef]
6. Highfield, A.C.; El Nagar, A.; Mackinder, L.C.M.; Noël, L.M.L.J.; Hall, M.J.; Martin, S.J.; Schroeder, D.C. Deformed wing virus implicated in overwintering honeybee colony losses. *Appl. Environ. Microbiol.* **2009**, *75*, 7212–7220. [CrossRef]
7. Dainat, B.; Evans, J.D.; Chen, Y.P.; Gauthier, L.; Neumanna, P. Dead or alive: Deformed wing virus and varroa destructor reduce the life span of winter honeybee. *Appl. Environ. Microbiol.* **2012**, *78*, 981–987. [CrossRef]
8. Ryabov, E.V.; Childers, A.K.; Lopez, D.; Grubbs, K.; Posada-Florez, F.; Weaver, D.; Girten, W.; vanEngelsdorp, D.; Chen, Y.; Evans, J.D. Dynamic evolution in the key honey bee pathogen Deformed wing virus. *PLoS Biol.* **2019**, *17*, e3000502. [CrossRef]
9. Ryabov, E.V.; Wood, G.R.; Fannon, J.M.; Moore, J.D.; Bull, J.C.; Chandler, D.; Mead, A.; Burroughs, N.; Evans, D.J. A Virulent Strain of Deformed Wing Virus (DWV) of Honeybees (Apis mellifera) Prevails after Varroa destructor-Mediated, or In Vitro, Transmission. *PLoS Pathog.* **2014**, *10*, e1004230. [CrossRef]
10. Gisder, S.; Möckel, N.; Eisenhardt, D.; Genersch, E. In vivo evolution of viral virulence: Switching of deformed wing virus between hosts results in virulence changes and sequence shifts. *Environ. Microbiol.* **2018**, *20*, 4612–4628. [CrossRef]
11. Nazzi, F.; Brown, S.P.; Annoscia, D.; Del Piccolo, F.; Di Prisco, G.; Varricchio, P.; Vedova, G.D.; Cattonaro, F.; Caprio, E.; Pennacchio, F. Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies. *PLoS Pathog.* **2012**, *8*, e1002735. [CrossRef] [PubMed]
12. Posada-Florez, F.; Childers, A.K.; Heerman, M.C.; Egekwu, N.I.; Cook, S.C.; Chen, Y.; Evans, J.D.; Ryabov, E.V. Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite Varroa destructor in a non-propagative manner. *Sci. Rep.* **2019**, *9*, 12445. [CrossRef] [PubMed]
13. Fujiyuki, T.; Ohka, S.; Takeuchi, H.; Ono, M.; Nomoto, A.; Kubo, T. Prevalence and Phylogeny of Kakugo Virus, a Novel Insect Picorna-Like Virus That Infects the Honeybee (Apis mellifera L.), under Various Colony Conditions. *J. Virol.* **2006**, *80*, 11528–11538. [CrossRef] [PubMed]

14. Lanzi, G.; Miranda, J.R.D.; Boniotti, M.B.; Cameron, C.E.; Lavazza, A.; Capucci, L.; Camazine, S.M.; Rossi, C.; de Miranda, J.R. Molecular and Biological Characterization of Deformed Wing Virus of Honeybees (*Apis mellifera L.*). *Society* **2006**, *80*, 4998–5009. [CrossRef] [PubMed]
15. Ongus, J.R.; Peters, D.; Bonmatin, J.M.; Bengsch, E.; Vlak, J.M.; van Oers, M.M. Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite Varroa destructor. *J. Gen. Virol.* **2004**, *85*, 3747–3755. [CrossRef] [PubMed]
16. Baker, A.C.; Schroeder, D.C. The use of RNA-dependent RNA polymerase for the taxonomic assignment of Picorna-like viruses (order Picornavirales) infecting Apis mellifera L. populations. *Virol. J.* **2008**, *5*, 10. [CrossRef]
17. Martin, S.J.; Highfield, A.C.; Brettell, L.; Villalobos, E.M.; Budge, G.E.; Powell, M.; Nikaido, S.; Schroeder, D.C. Global Honey Bee Viral Landscape Altered by a Parasitic Mite. *Science* **2012**, *336*, 1304–1306. [CrossRef]
18. Mordecai, G.J.; Wilfert, L.; Martin, S.J.; Jones, I.M.; Schroeder, D.C. Diversity in a honey bee pathogen: First report of a third master variant of the Deformed Wing Virus quasispecies. *ISME J.* **2016**, *10*, 1264–1273. [CrossRef]
19. Genersch, E.; Yue, C.; Fries, I.; De Miranda, J.R. Detection of Deformed wing virus, a honey bee viral pathogen, in bumble bees (Bombus terrestris and Bombus pascuorum) with wing deformities. *J. Invertebr. Pathol.* **2006**, *91*, 61–63. [CrossRef]
20. McMahon, D.P.; Natsopoulou, M.E.; Doublet, V.; Fürst, M.; Weging, S.; Brown, M.J.F.; Gogol-Döring, A.; Paxton, R.J. Elevated virulence of an emerging viral genotype as a driver of honeybee loss. *Proc. Biol. Sci.* **2016**, *283*, 443–449. [CrossRef]
21. Dalmon, A.; Desbiez, C.; Coulon, M.; Thomasson, M.; Le Conte, Y.; Alaux, C.; Vallon, J.; Moury, B. Evidence for positive selection and recombination hotspots in Deformed wing virus (DWV). *Sci. Rep.* **2017**, *7*, 41045. [CrossRef] [PubMed]
22. Mordecai, G.J.; Brettell, L.E.; Martin, S.J.; Dixon, D.; Jones, I.M.; Schroeder, D.C. Superinfection exclusion and the long-term survival of honey bees in Varroa-infested colonies. *ISME J.* **2016**, *10*, 1182–1191. [CrossRef] [PubMed]
23. Tehel, A.; Vu, Q.; Bigot, D.; Gogol-Döring, A.; Koch, P.; Jenkins, C.; Doublet, V.; Theodorou, P.; Paxton, R. The Two Prevalent Genotypes of an Emerging Infectious Disease, Cause Equally Low Pupal Mortality and Equally High Wing Deformities in Host Honey Bees. *Viruses* **2019**, *11*, 114. [CrossRef] [PubMed]
24. Zioni, N.; Soroker, V.; Chejanovsky, N. Replication of Varroa destructor virus 1 (VDV-1) and a Varroa destructor virus 1–deformed wing virus recombinant (VDV-1–DWV) in the head of the honey bee. *Virology* **2011**, *417*, 106–112. [CrossRef]
25. Ryabov, E.V.; Childers, A.K.; Chen, Y.; Madella, S.; Nessa, A.; Vanengelsdorp, D.; Evans, J.D. Recent spread of Varroa destructor virus—1, a honey bee pathogen, in the United States. *Sci. Rep.* **2017**, *7*, 17447. [CrossRef]
26. Moore, J.; Jironkin, A.; Chandler, D.; Burroughs, N.; Evans, D.J.; Ryabov, E.V. Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies. *J. Gen. Virol.* **2011**, *92*, 156–161. [CrossRef]
27. Racaniello, V.; Baltimore, D. Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* **1981**, *214*, 916–919. [CrossRef]
28. Taniguchi, T.; Palmieri, M.; Weissmann, C. Qβ DNA-containing hybrid plasmids giving rise to Qβ phage formation in the bacterial host. *Nature* **1978**, *274*, 223–228. [CrossRef]
29. Stobart, C.C.; Moore, M.L. RNA virus reverse genetics and vaccine design. *Viruses* **2014**, *6*, 2531–2550. [CrossRef]
30. Lamp, B.; Url, A.; Seitz, K.; Rgen Eichhorn, J.; Riedel, C.; Sinn, L.J.; Indik, S.; Kö Glberger, H.; Rü Menapf, T. Construction and Rescue of a Molecular Clone of Deformed Wing Virus (DWV). *PLoS ONE* **2016**, *11*, e0164639. [CrossRef]
31. Christie, C.R.; Bradford, E.L.; Budge, G.E.; Campbell, E.M.; Bowman, A.S. A novel artificial feeding system for Varroa destructor: Demonstration of utility for *in vitro* varroacide testing, pathogen transmission, and delivery of double-stranded RNA. In Preparation. 2020.
32. Gusachenko, O.N.; Woodford, L.; Balbirnie-Cumming, K.; Ryabov, E.V.; Evans, D.J. Deformed Wing Virus spillover from honey bees to bumble bees: A reverse genetic study. *bioRxiv* **2020**. [CrossRef]
33. Bailey, L.; Ball, B.V. *Honey Bee Pathology*; Academic Press: Cambridge, MA, USA, 1991; ISBN 9780120734818.

34. Bowen-Walker, P.; Martin, S.; Gunn, A. The transmission of deformed wing virus between honeybees (Apis mellifera L.) by the ectoparasitic mite varroa jacobsoni Oud. *J. Invertebr. Pathol.* **1999**, *73*, 101–106. [CrossRef] [PubMed]
35. Fievet, J.; Tentcheva, D.; Gauthier, L.; de Miranda, J.; Cousserans, F.; Colin, M.E.; Bergoin, M. Localization of deformed wing virus infection in queen and drone *Apis mellifera* L. *Virol. J.* **2006**, *3*, 16. [CrossRef] [PubMed]
36. Möckel, N.; Gisder, S.; Genersch, E. Horizontal transmission of deformed wing virus: Pathological consequences in adult bees (*Apis mellifera*) depend on the transmission route. *J. Gen. Virol.* **2011**, *92*, 370–377. [CrossRef]
37. Ryabov, E.V.; Christmon, K.; Heerman, M.C.; Posada-Florez, F.; Harrison, R.L.; Chen, Y.; Evans, J.D. Development of a Honey Bee RNA Virus Vector Based on the Genome of a Deformed Wing Virus. *Viruses* **2020**, *12*, 374. [CrossRef]
38. Mueller, S.; Wimmer, E. Expression of foreign proteins by poliovirus polyprotein fusion: Analysis of genetic stability reveals rapid deletions and formation of cardioviruslike open reading frames. *J. Virol.* **1998**, *72*, 20–31. [CrossRef]
39. Kulhanek, K.; Steinhauer, N.; Rennich, K.; Caron, D.M.; Sagili, R.R.; Pettis, J.S.; Ellis, J.D.; Wilson, M.E.; Wilkes, J.T.; Tarpy, D.R.; et al. A national survey of managed honey bee 2015–2016 annual colony losses in the USA. *J. Apic. Res.* **2017**, *56*, 328–340. [CrossRef]
40. Lee, K.V.; Steinhauer, N.; Rennich, K.; Wilson, M.E.; Tarpy, D.R.; Caron, D.M.; Rose, R.; Delaplane, K.S.; Baylis, K.; Lengerich, E.J.; et al. A national survey of managed honey bee 2013–2014 annual colony losses in the USA. *Apidologie* **2015**, *46*, 292–305.
41. Barron, A.B. Death of the bee hive: Understanding the failure of an insect society. *Curr. Opin. Insect Sci.* **2015**, *10*, 45–50. [CrossRef]
42. Carrillo-Tripp, J.; Dolezal, A.G.; Goblirsch, M.J.; Miller, W.A.; Toth, A.L.; Bonning, B.C. In vivo and in vitro infection dynamics of honey bee viruses. *Sci. Rep.* **2016**, *6*, 22265. [CrossRef]
43. Gisder, S.; Aumeier, P.; Genersch, E. Deformed wing virus: Replication and viral load in mites (Varroa destructor). *J. Gen. Virol.* **2009**, *90*, 463–467. [CrossRef]
44. Erban, T.; Harant, K.; Hubalek, M.; Vitamvas, P.; Kamler, M.; Poltronieri, P.; Tyl, J.; Markovic, M.; Titera, D. In-depth proteomic analysis of Varroa destructor: Detection of DWV-complex, ABPV, VdMLV and honeybee proteins in the mite. *Sci. Rep.* **2015**, *5*, 13907. [CrossRef] [PubMed]
45. Campbell, E.M.; Budge, G.E.; Watkins, M.; Bowman, A.S. Transcriptome analysis of the synganglion from the honey bee mite, Varroa destructor and RNAi knockdown of neural peptide targets. *Insect Biochem. Mol. Biol.* **2016**, *70*, 116–126. [CrossRef] [PubMed]
46. Iqbal, J.; Mueller, U. Virus infection causes specific learning deficits in honeybee foragers. *Proc. Biol. Sci.* **2007**, *274*, 1517–1521. [CrossRef]
47. Dainat, B.; Neumann, P. Clinical signs of deformed wing virus infection are predictive markers for honey bee colony losses. *J. Invertebr. Pathol.* **2013**, *112*, 278–280. [CrossRef]
48. Amiri, E.; Kryger, P.; Meixner, M.D.; Strand, M.K.; Tarpy, D.R.; Rueppell, O. Quantitative patterns of vertical transmission of deformed wing virus in honey bees. *PLoS ONE* **2018**, *13*, e0195283. [CrossRef]
49. Myser, W.C. The Larval and Pupal Development of the Honey Bee Apis Mellifera Linnaeus1. *Ann. Entomol. Soc. Am.* **1954**, *47*, 683–711. [CrossRef]
50. Amiri, E.; Meixner, M.D.; Kryger, P. Deformed wing virus can be transmitted during natural mating in honey bees and infect the queens. *Sci. Rep.* **2016**, *6*, 1–7. [CrossRef]
51. Chen, Y.P.; Pettis, J.S.; Collins, A.; Feldlaufer, M.F. Prevalence and transmission of honeybee viruses. *Appl. Environ. Microbiol.* **2006**, *72*, 606–611. [CrossRef]
52. Fujiyuki, T.; Takeuchi, H.; Ono, M.; Ohka, S.; Sasaki, T.; Nomoto, A.; Kubo, T. Kakugo Virus from Brains of Aggressive Worker Honeybees. *Adv. Virus Res.* **2005**, *65*, 1–27.
53. Shah, K.S.; Evans, E.C.; Pizzorno, M.C. Localization of deformed wing virus (DWV) in the brains of the honeybee, Apis mellifera Linnaeus. *Virol. J.* **2009**, *6*, 182. [CrossRef]
54. Racaniello, V.R. One hundred years of poliovirus pathogenesis. *Virology* **2006**, *344*, 9–16. [CrossRef]

8.10 Appendix 10 - Published manuscript – Gusachenko *et al.*, (2020), Scientific Reports.

www.nature.com/scientificreports

SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

Evidence for and against deformed wing virus spillover from honey bees to bumble bees: a reverse genetic analysis

Olesya N. Gusachenko[1]✉, Luke Woodford[1], Katharin Balbirnie-Cumming[1], Eugene V. Ryabov[2] & David J. Evans[1]✉

Deformed wing virus (DWV) is a persistent pathogen of European honey bees and the major contributor to overwintering colony losses. The prevalence of DWV in honey bees has led to significant concerns about spillover of the virus to other pollinating species. Bumble bees are both a major group of wild and commercially-reared pollinators. Several studies have reported pathogen spillover of DWV from honey bees to bumble bees, but evidence of a sustained viral infection characterized by virus replication and accumulation has yet to be demonstrated. Here we investigate the infectivity and transmission of DWV in bumble bees using the buff-tailed bumble bee *Bombus terrestris* as a model. We apply a reverse genetics approach combined with controlled laboratory conditions to detect and monitor DWV infection. A novel reverse genetics system for three representative DWV variants, including the two master variants of DWV—type A and B—was used. Our results directly confirm DWV replication in bumble bees but also demonstrate striking resistance to infection by certain transmission routes. Bumble bees may support DWV replication but it is not clear how infection could occur under natural environmental conditions.

Deformed wing virus (DWV) is a widely established pathogen of the European honey bee, *Apis mellifera*. In synergistic action with its vector—the parasitic mite *Varroa destructor*—it has had a devastating impact on the health of honey bee colonies globally[1,2]. As the primary managed insect pollinator, honey bees are of high ecological and economic value and contribute an estimated 30–50% of mobile pollination activity. This is carried out by seasonal transportation of honey bee hives between agricultural areas requiring pollination services[3,4]. For example, two-thirds of all colonies in the USA (~ 1.6 million hives) are transported to California in February/March for almond pollination[5]. Inevitably, transporting bees also transports their pathogens. This, coupled with the local pathogen density associated with ~ 50 000 bees in a single hive, has raised concerns about pathogen spillover from managed honey bees to other pollinators[6]. DWV was found as a frequent component of pollen pellets[7] and is also present in bee faeces[8], suggesting honey bee foragers and colonies could facilitate horizontal virus transmission to the wider pollinator community. Significantly, DWV RNA has been detected in many insects sharing the environment with managed honey bees, including Asian bee species, solitary bees, bumble bees, wasps, cockroaches and ants[6,7,9–21]. Due to their extended activity at lower temperatures (compared to honey bees) bumble bees are considered particularly important pollinators in temperate and subarctic climates[22,23] and are also reared and managed for commercial-scale pollination[23]. As a consequence, the potentially negative impact of extensive honey bee management and failing pathogen control on co-located *Bombus* species has received significant attention.

Following a report that DWV was detected in symptomatic *Bombus terrestris* and *Bombus pascuorum* individuals with deformed wings[9] there have been several studies of DWV prevalence in a wide range of *Bombus* species[6,7,13,14,24–27]. In *Varroa*-infested honey bee colonies DWV levels can exceed 10^{11} genome copies per bee[28], with considerable potential for environmental contamination. DWV-positive *Bombus* sp. have been shown to correlate to areas with high DWV prevalence in *Apis*[6,25–27]. The majority of screening studies have used end-point

[1]Biomedical Sciences Research Complex, University of St. Andrews, North Haugh, St. Andrews KY16 9ST, UK. [2]USDA-ARS Bee Research Laboratory, Beltsville Agricultural Research Center, Beltsville, MD 20705, USA. ✉email: olesya.gusachenko@gmail.com; d.j.evans@st-andrews.ac.uk

polymerase chain reaction (PCR) for detecting DWV RNA in environmental bumble bee samples (reviewed in[29]). However, the near ubiquitous presence of managed hives, the honey bee—and consequently pathogen—density around hives, and the foraging range of *Apis* mean that DWV is likely widespread[2]. Although detection of the negative strand intermediate of replication is regarded as a marker of DWV replication, quantitative analysis of virus amplification in the infected tissues is required to unequivocally demonstrate infection and replication.

The name DWV is currently attributed to an evolving complex of closely related viruses, which includes DWV-A[30], Kakugo virus[31], Varroa destructor virus-1 (VDV-1; also referred to as DWV-B[1,32,33]) and a range of recombinants between DWV-A and -B[34–37]. All viruses exhibit at least 84% identity at the nucleotide level and 95% identity at the protein level[34,37,38]. The sensitivity of current diagnostic methods means DWV detection in environmental samples is regularly reported, with different DWV variants identified in *Bombus*[6]. Far less frequent are studies investigating potential routes of transmission from *Apis* to *Bombus*, or the subsequent replication of DWV in bumble bees. Due to the absence of a suitable cell line for in vitro propagation, laboratory-based assays have been limited to the application of field-sourced virus. Infectivity of DWV obtained from field honey bee samples was tested via inoculations of adult *Bombus terrestris*[6]. It was reported that a DWV complex containing both DWV-A and -B is infectious when fed at high concentrations—10^9 GE (genome equivalents) of virus per bee.

We have used a reverse genetic (RG) approach to generate near-clonal populations of genetically tagged DWV-A, -B and a B/A recombinant after transfection of honey bees with in vitro transcribed RNA. A similar system has recently been reported for DWV-A[39–41]. Using RG-derived DWV inocula we address the question of DWV pathogenesis and likely transmission routes in *Bombus terrestris* at both the individual and colony level. Using this strategy we provide direct evidence of DWV replication in bumble bees via virus feeding and injection. Importantly, adult *Bombus terrestris* appear resistant to productive infection by DWV orally and do not exhibit the developmental defects characteristic of DWV infection and replication in honey bees.

Results

Infectivity of DWV RNA and virus in injected honey bee and bumble bee pupae. In order to test infectivity of DWV variants in controlled laboratory experiments we used RG systems for DWV-A and -B master variants and a recombinant B/A variant reported earlier by our group[42]. Full-length viral RNA was prepared in vitro and directly injected into honey bee pupae from which near-clonal infectious virus was recovered. Incorporation of synonymous mutations, which create new restriction sites in each RG DWV genome, allow unambiguous distinction from wild type virus genomes. The following nomenclature was used for the generated viruses: "VDD"—type A DWV (type A polyprotein-coding part only), "VVV"—type B, "VVD"—B/A recombinant (Fig. 1a). Full sequences of DWV cDNAs used in this study are available online (GenBank accession numbers: DWV-VDD—MT415949, DWV-VVD—MT415950, DWV-VVD_truncated—MT415951, DWV-VVV—MT415952)[42].

In temperate climates honey bee brood is only available for ~ 50% of the year and in vivo research is of necessity seasonal. Although the internal ribosome entry site of DWV retains partial activity in at least one cell line of non-honey bee origin (*Lymanthria dispar* LD652Y cells)[43], attempted infection of those cells with DWV or transfection of in vitro-generated DWV RNA does not result in virus replication (Gusachenko, unpublished data). This impediment to DWV studies prompted us to investigate the recovery of clonal stocks of DWV in bumble bees (*Bombus terrestris audax*) injected with viral RNA. Commercially farmed bumble bee colonies are available year-round and, in our preliminary studies, are free of DWV RNA (data not shown). Four of six RNA-injected bumble bee pupae tested positive for DWV, while all honey bee pupae injected with full-length RNA were shown to be DWV-positive (Fig. 1b,c). The RG origin of DWV in injected samples was confirmed by restriction digest of PCR products using endonucleases specific for the introduced genetic tag (Fig. 1b,c—restriction digest). The remaining tissue from RNA-injected samples was used to prepare crude DWV stocks. Quantitative PCR (qPCR) analysis showed that the *Apis*-derived stocks contained between 10^7–10^9 GE of DWV/µl, while *Bombus*-derived inoculum contained 3.75×10^5 GE/µl. For comparison, DWV levels in mock-injected *Apis* pupae did not exceed 10^5 GE DWV/µg of RNA (representing the endogenous viral load) and *Bombus* pupae had no detectable DWV sequences in mock-inoculated samples.

DWV extracted from *Bombus* was infectious when re-inoculated by injection to *Bombus* and *Apis* pupae. We observed no differences in the infectivity of *Apis*- or *Bombus*-derived DWV (Fig. S1) and, at the genome level, no obvious signs of adaptation following comparison of the parental cDNA sequence with next generation sequencing (NGS) data from the second passage of *Bombus*-derived DWV (Fig. S2).

In honey bees, both DWV-A and -B are pathogenic when inoculated, though recent studies have produced contradictory results when comparing their relative virulence[1,42,44–46]. We therefore tested infectivity of VDD, VVV and VVD DWV in bumble bee brood and adults. As our primary interest was to investigate the potential for DWV spillover from infected honey bees we used DWV inocula prepared from RNA-injected honey bee pupae for all further experiments.

White-eyed (P0-P1 pupa stage according to the published classification on bumble bee pupae morphology[47]) or brown-eyed (P7-P8) bumble bee pupae were injected with 10^3 or 10^6 GE of each DWV variant, and virus levels quantified 48 h post-inoculation. In all cases we observed a 10^2–10^4 increase in total DWV load compared to the innocula, providing clear evidence for replication (Fig. 2a). More DWV accumulated in older pupae but this was only statistically significant for 10^6 GE of the VVV variant (Tukey's multiple comparisons test, $P = 0.03$).

***Bombus* larvae can be infected with DWV *per os*.** Since bumble bee pupae are not parasitised by *Varroa* naturally-infected pupae must acquire the virus by prior exposure of larvae. We therefore investigated virus infection after feeding DWV to 1st and 2nd instar *Bombus* larvae (age group 1 and 2 on Fig. 2b respectively).

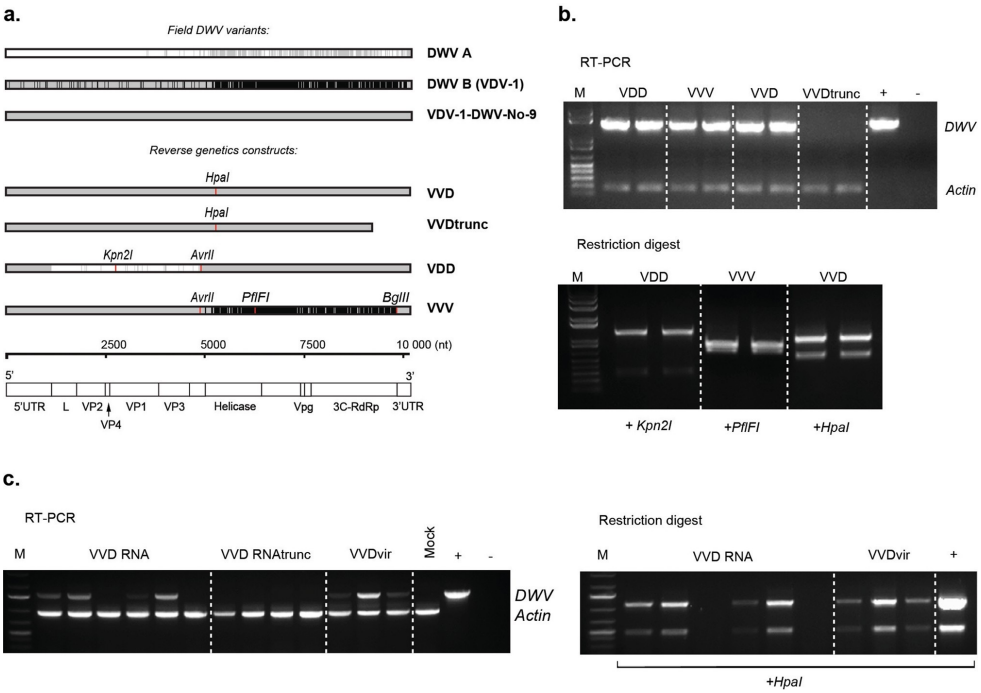


Figure 1. Reverse genetics (RG) system for three DWV variants. (**a**) Diagram showing identical regions between genomic RNA sequences of field DWV variants (DWV-A—NC_004830.2, DWV-B—GenBank AY251269.2, B/A recombinant VDV-1-DWV-No-9—GenBank HM067438.1) and three RG clones (VVV, VDD and VVD); DWV-A and DWV-B specific regions are shown in white and black respectively, regions identical to the VDV-1-DWV-No-9 recombinant are shaded in grey, location of restriction sites introduced into cDNA of each variant is indicated by red marks, DWV genomic RNA organization is presented below to help interpretation. (**b**) Detection of RG DWV RNA by end-point PCR in honey bee pupae injected with in vitro synthesized RNA transcripts: "VDD", "VVV", and "VVD"—PCR products from pupae injected with corresponding full-length RNA, "VVDtrunc"—PCR from pupae, injected with truncated VVD RNA, "+" and "–" —PCR controls, "M"—molecular weight DNA marker; restriction digest—verification of the RG origin of detected DWV by the digest of the PCR products at artificially introduced restriction sites. (**c**) Detection of RG DWV in bumble bee pupae injected with VVD RNA and with virus stock obtained from RNA-injected bumble bee pupae ("VVDvir"); "VVD RNAtrunc"—PCR from pupae injected with truncated RNA, "Mock"—PBS-injected pupae, "+"—positive PCR control for DWV, "M"—molecular weight DNA marker; RG origin of the PCR products for all DWV-positive samples was confirmed by digest with *HpaI* restriction enzyme, amplification of the actin mRNA product was used as an indicator of RNA integrity and loading control.

Each larva individually received a single dose of diet containing 10^8 GE of DWV on the first day of the experiment. DWV was detectable in all fed larvae on the 5th day post inoculation, although virus levels were comparable to the amount initially administered. When analysed ~ 50% of fed larvae had detectable levels of DWV negative strand RNA, absent from the inocula and indicative of virus replication (Fig. 2b).

Investigation of colony scale transmission of DWV and adult infection *per os*. In honey bees, horizontal transmission of DWV occurs *per os* in larvae or adults, or when *Varroa* feed on pupae and adults. Whilst we show here that DWV can replicate in *Bombus* pupae after direct injection, and in virus fed larvae, the route by which adult bumble bees could become infected remains unclear. We reasoned that two likely routes would be via direct oral exposure of adult *Bombus* to virus in the environment or indirectly following larval infection with virus carried by adult bees.

Groups of adult bumble bees received 10^7 or 10^8 DWV GE per bee via feeding with virus-supplemented sucrose solution. All control group bees remained viable during the experiment, and only one dead bee was found in each of the virus fed groups. Bees were analyzed 7 days after DWV feeding and none of the virus-fed samples tested positive for DWV in end-point PCR (Fig. 3). According to qPCR analysis of DWV-fed bees no sample contained greater than 10^5 GE DWV per 1 µg of total RNA. Average RNA yield per bee did not exceed 50 µg

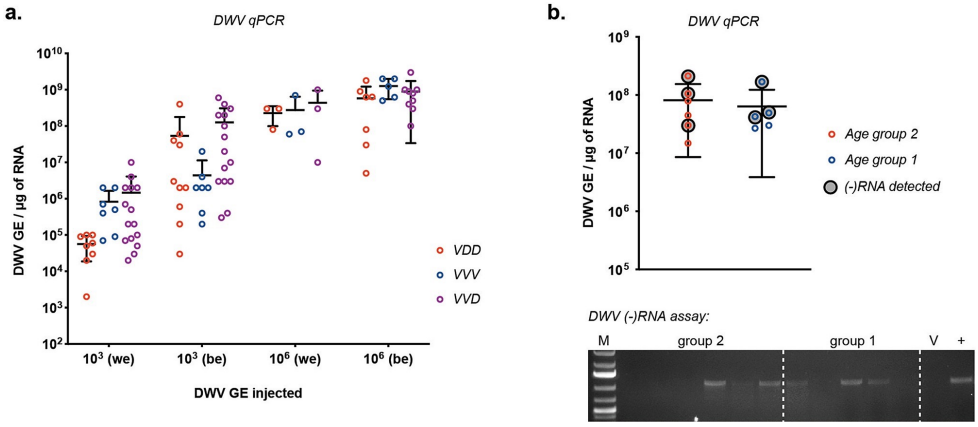


Figure 2. Inoculation of bumble bee pupae and larvae with RG-DWV. (**a**) qPCR analysis of DWV accumulation in bumble bee pupae injected with VVV, VDD and VVD DWV. Pupae were injected at white-eyed (P0-P1)—"we"—and brown-eyed (P7-P8)—"be"—stages and analyzed 48 h post injection. Each value corresponds to an individual sample analyzed, black lines show mean ± SD, "*GE*"—genome equivalents. (**b**) Detection of DWV RNA in bumble bee larvae from two different age groups fed with 10^8 GE VVD DWV: qPCR analysis of DWV levels in individual larvae samples, black-circled values correspond to samples which produced a positive result in (−)RNA assay; each value corresponds to an individual sample analyzed, error bars show mean ± SD, "*GE*"—genome equivalents. DWV (-)RNA assay—PCR products obtained after strand-specific reverse transcription and run in 1% agarose gel, "+"—positive PCR control for DWV, "M"—molecular weight DNA marker, "V"—PCR from the DWV inoculum used for larvae feeding.

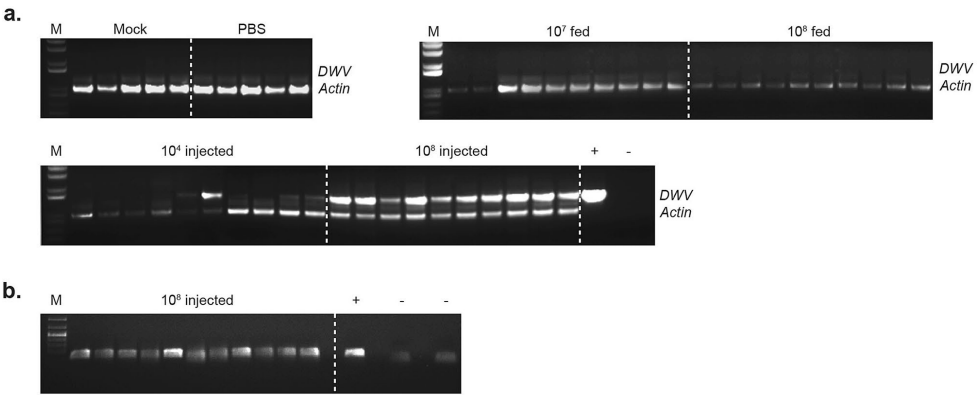


Figure 3. Detection of DWV in adult bumble bees. (**a**) Detection of DWV RNA by end-point PCR in adult bumble bees after inoculation with VVD DWV via feeding or injections (10 samples from each group shown): "Mock" – non-treated bumble bees from the same colony, "PBS"—PBS-injected group (no virus), "10^7 and 10^8 fed"—bumble bees fed with sucrose solution containing 10^7 or 10^8 DWV GE of DWV per bee, "10^4 and 10^8 injected"—bumble bees injected with 10^4 or 10^8 DWV GE of DWV respectively. (**b**) Detection of the DWV (-) RNA strand in adult bumble bees injected with 10^8 DWV GE. "+" and "-"—positive and negative PCR controls, "M"—molecular weight DNA marker. Detection of the bumble bee actin RNA was used to assay the quality of extracted RNA. RG origin of the PCR products for all DWV-positive samples was confirmed by restriction enzyme digest.

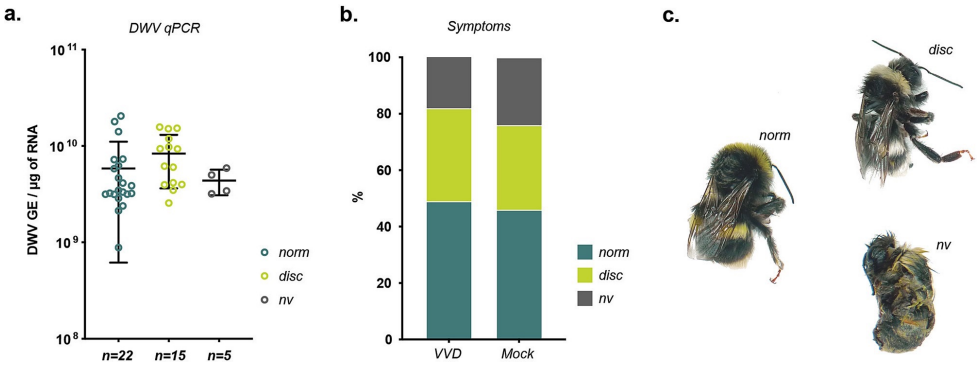


Figure 4. Morbidity of DWV in bumble bees. (**a**) qPCR analysis of DWV level in developed bumble bees injected with 10^6 GE of VVD DWV at white-eyed (P0-P1) pupa stage. Individual values for each sample are shown with dots, lines represent mean ± SD, "GE"—genome equivalents. (**b**) Percentage of visually normal ("*norm*"), discolored ("*disc*") and non-viable ("*nv*") bumble bees developed from pupae in VVD DWV-injected (n = 43) and PBS-injected ("Mock", n = 37) groups. (**c**) Phenotype of bumble bees developed from pupae in the incubator.

and therefore the level of virus detected was 100–200-fold lower than the amount administered. Hence, despite evidence that bumble bee larvae displayed susceptibility to DWV infection *per os*, feeding of adult bumble bees up to 10^8 GE DWV per bee did not lead to infection or detectable replication. Since bumble bees, like many host species, may exhibit differential age-related susceptibility to pathogens we additionally tested the oral infection of newly eclosed workers only, and a mixed group of randomly selected older bees collected from declining nests. However, in no cases were we able to demonstrate infection by the oral route (data not shown).

We extended this study to investigate whole-nest inoculation with DWV. Three individual bumble bee nests were fed for 4–6 weeks with a sucrose solution supplemented with 2×10^8 GE/adult bee/day of VVD, VVV or VDD DWV. All brood (pooled egg samples, 30 larvae, 47 pupae) and 30 randomly selected adults from virus-exposed nests were screened for DWV using end-point PCR assay and showed no positive results (data not shown).

Direct inoculation of adult bumble bees. With no evidence that adult bumble bees could be infected when fed a DWV-supplemented diet, or that they were able to transfer virus to larvae, we performed direct virus injections of adult bumble bees to determine whether adults could support DWV replication. Two groups of adult bumble bees were intra-abdominally injected with 10^4 or 10^8 DWV GE per bee. Envisaging a possible impact of injection on viability, an additional group of bumble bees was injected with phosphate buffered saline (PBS) only. 20–30% of injected bees died before termination of the experiment in each group. The remaining bumble bees were screened for DWV RNA 8 days post-inoculation. End-point PCR analysis indicated that 40% and 100% were DWV positive in the 10^4-injected and 10^8-injected groups respectively (Fig. 3a), with qPCR analysis showing that the virus levels in these groups ranged between 2.7×10^4–1.4×10^7 and 4.8×10^6–2.1×10^8 GE/µg of total RNA respectively. Accumulation of DWV (-)RNA in DWV-positive samples was confirmed by strand-specific PCR assay (Fig. 3b). This demonstrates that DWV can replicate in adult *Bombus terrestris* after direct injection of 10^8 GE per bee. We therefore proceeded to investigate if there were any pathogenic consequences of virus infection and replication.

Pathogenicity of DWV in *Bombus terrestris*. DWV produces characteristic pathogenicity in honey bees and similar symptoms have been reported in bumble bees[9]. We injected 43 white-eyed (P0–P1) bumble bee pupae with 10^6 GE of DWV and maintained them through development in an incubator. In parallel, a group of 37 similarly-aged pupae were injected with PBS. All fully-developed bumble bees could be classified into one of three groups: phenotypically normal, discolored with normal wings or non-viable, which did not eclose (Fig. 4). Analysis by qPCR demonstrated that all DWV-inoculated bumble bees contained high levels of virus (8.8×10^8–2.2×10^{10} GE/µg of RNA) with no significant differences in viral load between the three phenotypic groups (ANOVA, P = 0.21) (Fig. 4). DWV levels were 1–2 $\log_{10}$ greater than in pupae analysed 48 h post-inoculation (Fig. 2a) reflecting the additional time the virus had to replicate. This further supports the conclusion that pupae support productive infection with DWV following direct inoculation. Strikingly, none of the eclosed bumble bees showed any signs of the wing deformities that are characteristic of DWV infection of honey bees. Inspection of PBS-injected pupae showed that they could be separated into the same phenotypical groups; normal, discoloured and dead. There was no statistical difference between the proportions in each group of virus- or PBS-injected bumble bees (ANOVA, P > 0.999). Upon analysis, none of the mock inoculated group

samples showed evidence of DWV infection. Injection of honey bee pupae with the same virus variant caused characteristic wing deformities in eclosed bees[42].

Discussion

The term "pathogen spillover" describes the transmission of a pathogen from a reservoir host to a different species in a shared environment[48]. To achieve this, several conditions must be fulfilled; the reservoir host must be sufficiently abundant to guarantee exposure, pathogen prevalence must be high enough to ensure direct or environmental transmission, and the recipient host must be susceptible to infection via direct or indirect transmission[49].

Managed honey bee colonies are near-ubiquitous in many environments, including both rural and urban locations. On agricultural crops that require commercial pollination very high colony densities are achieved through migratory beekeeping. The combination of colony movements following the introduction of *Varroa destructor* resulted in the near-global distribution of the ectoparasite, with the concomitant spread of a range of honey bee viruses that are detrimental to colony survival[1]. Most important of these viruses is DWV which likely accounts for the majority of overwintering colony losses[50]. DWV causes overt and lethal developmental defects after pupal inoculation and reduces the longevity of bees that do successfully emerge. Colonies that collapse due to high *Varroa*/DWV levels may be robbed-out by other insects including bumble bees, ants and wasps.

In addition to the cocktail of *Varroa* and DWV, honey bees are increasingly subject to stresses through a combination of limited dietary variation, repeated transportation to new pollination sites and exposure to agrochemicals, all of which are associated with increased susceptibility to pathogens and potential colony failure[51]. As a consequence, honey bees readily fulfill two of the requirements needed for pathogen spillover *i.e.* abundance in the environment and pathogen prevalence. In addition, due to the extensive foraging range of honey bees and the excretion of viable DWV in faeces[8], a high density of hives ensures widespread environmental contamination.

Notwithstanding the likely exposure of other species to the pathogen-laden environment, spillover also requires susceptibility of the species exposed and pathogen infectivity via a relevant transmission route. We exploited commercially available colonies of *Bombus terrestris* to determine if and how they could support infection by DWV. The buff-tailed bumble bee is a relevant model system in which to explore pathogen spillover; it is naturally found in the same environment as honey bees[6], and there are reports of DWV infection of wild-caught *B. terrestris* and several related species[6,7,13,14,24,27].

We developed a RG system for the two prototype strains of DWV (type A and B) and a hybrid previously reported to predominate in *Varroa*-infested colonies[42]. The synthesized virus genomes contain silent restriction sites that unambiguously distinguish the three types and—since they are unique to the RG sequences—allow discrimination from endogenous DWV. Direct inoculation of honey bee pupae with in vitro generated RNA resulted in DWV replication (Fig. 2). Virus purified from these pupae caused symptomatic infection when reinoculated into honey bee pupae and accumulated to 10^{10} GE/µg of RNA[42], a level similar to that observed in *Varroa*-exposed pupae[28].

Using a similar strategy we demonstrated that bumble bee pupae could be infected when inoculated with in vitro generated DWV RNA. The resulting virus was purified, shown to be infectious for naïve bumble bee pupae when re-injected and deep sequencing of the virus population indicated no significant sequence changes from the originating cDNA (Fig. S2). Infected pupae reached up to ~ 10^9 GE of DWV/µg of RNA two days post-injection (Fig. 2a) and accumulation of the negative strand DWV RNA was confirmed by strand-specific qPCR (Fig. S3). Therefore DWV is infectious for both honey bee and bumble bee pupae and adaptive changes are unlikely to create a bottleneck in any potential transmission between the species.

Direct recovery and amplification of DWV in bumble bees offers advantages to virologists and entomologists attempting to determine the significance of the limited differences between the reported strains. Since essentially all honey bee pupae are infected with DWV, perhaps with the exclusion of those from Australia[52], it is impossible to obtain truly clonal virus preparations. By recovering virus after RNA inoculation in bumble bees pure populations of DWV strains can be prepared for subsequent studies of virus pathogenesis and evolution.

Since *Varroa* does not parasitise bumble bee pupae it is difficult to rationalise direct injection as a potential transmission route. Therefore, considering robbing by bumble bees of collapsing honey bee colonies and the likely widespread DWV contamination in environments with large numbers of honey bees, we reasoned that oral transmission was a more likely route for virus acquisition. We therefore investigated oral susceptibility of larvae and adult bumble bees, fed directly or by extended exposure of bumble bee nests to virus-supplemented diet.

Direct feeding with 10^8 GE of DWV enabled the detection of the negative strand of DWV, absent from the input virus preparation and indicative of virus replication, in ~ 50% of larvae (Fig. 2). In contrast, DWV fed adult bumble bees, or larvae and pupae harvested from nests supplemented with diet containing 2×10^8 GE of DWV per bee, failed to provide any evidence for virus acquisition and transmission *per os*. Therefore, whilst larvae exhibit susceptibility to infection by orally acquired DWV, it is clear that the threshold for infection may be high and that it was not achieved with continuous feeding by adult bees with high levels of input virus. Further studies will be required to determine whether this was due to virus inactivation after ingestion by nurse bees, viable virus not being fed to the larvae, or some other undefined cause.

Although we found no evidence for oral infection of adult bumble bees with DWV, we were able to demonstrate the presence of DWV in adult *Bombus* after direct virus injections. Adult bumble bees are able to support replication of DWV though, as with pupal inoculation with RNA, there may be a threshold (exceeded at 10^8 GE) needed to achieve 100% infection.

In this study we allowed pupae to complete development and scored them phenotypically upon eclosion. Bumble bee pupae are susceptible to handling and survival rates (~ 50%) were similar in virus- or mock-inoculated samples (Fig. 4). Amongst the three phenotypically-distinct groups there was no difference in viral load in virus inoculated samples. Strikingly, the same three groups and the proportion of each were present in the

mock-inoculated samples. No wing deformities were detected in eclosed bumble bees from either group. It has previously been reported that DWV-positive field isolates of both *B. terrestris* and *B. pascuorum* have been identified with wing deformities characteristic of those seen in DWV-infected honey bees[9]. However, these studies did not demonstrate DWV replication or reproduce symptoms by DWV inoculation. In honey bees apart from DWV infection, wing malformation can be caused during development by injury, hormonal disorders, intoxication or absence of cocoon[39].

While laboratory inoculations represent the gold standard for virus infectivity assays this approach probably does not recapitulate an environmentally meaningful infection route. Compared to their independent effect a combination of stressors is suggested to introduce a greater threat to wild pollinators in their native environment[53]. For example, exposure to clothianidin—a neonicotinoid insecticide—was found to have a negative impact on honey bee immune status and promote DWV infection[54]. In *Bombus terrestris* condition mediated virulence of Slow bee paralysis virus upon starvation has been demonstrated[55]. Our study uses a reverse genetic approach to investigate pathogen spillover from honey bees to bumble bees. Whilst clear evidence is obtained to confirm DWV replication in bumble bee larvae, pupae and adults, we were unable to demonstrate a compelling route by which transmission would likely occur in the natural environment. The levels of virus required to orally infect bumble bee larvae are significantly higher than have been reported in environmental pollen samples[56]. Indeed, the levels required are likely higher than larvae are ever exposed to. In contrast, adult bumble bees may experience very high virus levels in collapsing honey bee colonies while robbing. However, adult bumble bees feeding on virus-supplemented syrup remain uninfected and—importantly—were unable to transmit infectious virus to the developing larvae. Further studies will be required to determine whether the gut environment of adult bumble bees is sufficiently hostile to DWV that the virus is inactivated e.g. by gut proteases, or if there are other factors that restrict infection and replication of DWV in bumble bees.

The results obtained from this study provide a strong indication that oral acquisition of virus from a contaminated environment does not represent an effective DWV transmission route from *Apis* to otherwise healthy *Bombus* individuals. Bumble bees are known to carry their own mite parasites, such as *Locustacarus buchneri*, which infest the air sacs of adult bumble bees and feed via piercing the tracheal wall[57], however, these mites do not parasitise on honey bees. Other mites found on *Bombus* lack feeding behaviour similar to *Varroa* and hence the capacity for virus vectoring during feeding. Non-Acari parasites of bumble bees such as protozoans *Apicystis bombi* and *Crithidia bombi* were shown to be present in honey bee collected pollen[58], but no evidence of their ability to transmit viruses between bee species has been reported. Therefore, a route for productive DWV spillover to bumble bees from infected honey bees remains to be determined. Previous reports have emphasised the haplotype identity of DWV between honey bee and bumble bee populations[6,26,27]. We propose that the detection of DWV RNA in geographically co-located bumble bees may reflect environmental contamination from the abundant honey bee population without necessarily supporting DWV replication in the bumble bee population. Further studies on defining the stressors that may account for DWV infection of bumble bees are required in order to estimate the actual impact of DWV on this important group of pollinators.

Materials and methods

Honey bees. All honey bee (*Apis mellifera*) brood were collected from the University of St Andrews research apiary. All hives used for sampling were routinely treated for *Varroa* with an appropriate miticide. Pupae were extracted from the comb and maintained in the incubator set at 34.5 °C with 90% relative humidity.

Bumble bees. Bumble bees *(Bombus terrestris audax*; Biobest, Belgium) were maintained in the laboratory in isolated nests supplemented with feeders containing 50% aqueous sucrose solution. Nests were regularly fed with bee pollen from a DWV-free region (Saxonbee, Australia). For adult inoculation experiments newly eclosed (harvested directly on emergence from sealed pupal cells) or mixed-age workers were selected at random from established nests in groups of 20–25 and maintained at RT in separate cages with *ad libidum* feeding on pollen/syrup. Prior to inoculation each group was left to recover for 24 h to account for any mortality caused by the handling. Pupae and larvae were extracted from the brood cells and transferred into the incubator set at 30.5 °C with 90% relative humidity. Larvae were fed a diet consisting of 25% (v:v) ground pollen and 25% sucrose solution (w:v) in H_2O.

RG system and in vitro RNA synthesis. All RG constructs in this study were based on a cDNA clone of a recombinant DWV variant VDV-1-DWV-No-9 (GenBank HM067438.1) and described in[42].

Viral RNA was generated in vitro from linearised plasmid templates, using the T7 RiboMAX Express Large Scale RNA Production System (Promega). Full length and truncated templates were linearised with *Pme* I (cuts at the end of poly-A tract) or *Nru* I (cuts at nucleotide position 9231 located within the sequence encoding the virus polymerase) respectively. Truncated RNA transcripts were polyadenylated using the poly(A) Tailing Kit (Thermo Fisher Scientific) to ensure their stability. After purification, RNA transcripts were eluted in RNAse free H_2O, their integrity confirmed by gel electrophoresis, and stored at − 80 °C.

RNA and virus injections. RNA and virus injections of 5 and 10 µl were performed with insulin syringes (BD Micro Fine Plus, 1 ml, 30 G, Becton Dickinson) into honey bee and bumble bee pupae respectively. Up to 10 µg (equivalent to 1.7×10^{12} of DWV RNA copies) of in vitro transcribed RNA was injected individually into pupae. Truncated VVD transcript injections were used as a negative control. All RNA-injected pupae were analyzed at 72 h post injection. Injections of pupae or low temperature-immobilized adult bees with virus stocks diluted in PBS were performed using the same technique as for RNA transcripts.

Virus stocks. DWV stocks were prepared from pupae injected with in vitro transcribed RNA as described earlier[42]. Homogenized tissue was diluted with sterile PBS in 1:1 (w:v) ratio and centrifuged at 13 000 *g*, 4 °C for 10 min. Supernatant was filter sterilized with 0.22 µM filter (PES, Merck Millipore) and treated with RNase A to destroy all unencapsidated RNA. RNA was extracted from 100 µl of the virus stock using RNeasy kit (Qiagen) and analyzed by reverse transcription followed by qPCR. Apart from the virus inoculum used for testing the infectivity of the *Bombus*-derived DWV, all virus stocks were prepared from RNA-injected honey bee pupae.

Oral inoculation and bumble bee nest feeding. Individual bumble bee larvae were placed into 96 well plates containing 2 µl of pollen/syrup mixed with DWV inoculum. DWV feeding was delivered once on the first day of the experiment. Fresh diet without virus was added individually after virus-containing food had been consumed. On the fifth day larvae were snap-frozen in liquid nitrogen and stored at − 80 °C.

Adult bumble bees were fed with virus in groups of 20 on the first and on the second day of the experiment. Total amount of DWV supplied with each of the two virus feedings was 2.2×10^9 and 2.2×10^8 DWV GE for 10^8 and 10^7 groups respectively. Bumble bees were snap-frozen in liquid nitrogen and analyzed for the presence of DWV RNA 7 days after DWV feeding.

For the colony-scale inoculations three bumble bee nests each containing a mated queen and 120–125 adult bees were used. Prior to the start of the experiment all brood was removed from the nests. Virus feeding was delivered daily by replacing the nest feeder with a tube containing 2 ml of sucrose solution supplemented with DWV, previously confirmed to be infectious for bumble bee pupae by direct injection. Each nest received a daily dose of the virus corresponding to 2.4×10^{10} DWV GE. The virus-containing solution was fully consumed (< 3–4 h) before replacement with the regular sucrose feeder. DWV feeding continued for 4–6 weeks and stopped when the first group of larvae developed during the virus-feeding period approached pupation. Upon termination of the experiment all brood (eggs, larvae and pupae) and 10 adult workers were sampled from each nest and tested for the presence of DWV.

RNA extraction, reverse transcription and PCR. Samples were homogenized with a Precellys Evolution homogenizer (Bertin Instruments). Total RNA was extracted using the GeneJet RNA Purification Kit (Thermo Fisher Scientific). cDNA was prepared with qScript cDNA Synthesis Kit (Quanta Biosciences) from 1 µg of total RNA following the manufacturer's protocol. Detection of DWV and actin mRNA, used as an internal RNA quality control, was carried out by end-point PCR with *Taq* DNA polymerase (New England Biolabs) and 2 µl of cDNA. DWV_RTPCR primers were designed to detect all three DWV variants under study (Table S1). To amplify the regions containing restriction site tags in VDD and VVV Kpn2I_F/R and PflFI_F/R primer pairs were used respectively. PCR cycling conditions were 30 cycles of 95 °C (15 s), 55 °C (15 s), 68 °C (2 min) with an initial 95 °C step (30 s) and a final extension at 68 °C (5 min). PCR samples were analyzed on a 1% agarose gel and when required, PCR products were subjected to restriction digest prior to loading on the gel.

The quantification of DWV genome copies was performed by SYBR-Green Real-Time Quantitative PCR using Luna Universal qPCR master mix (New England Biolabs), 0.25 µM forward and reverse DWV_qPCR primers and 2 µl of cDNA. The following thermal profile was used: 1 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 30 s at 60 °C with a final post-amplification melting curve analysis step. DWV titres were calculated by comparison of the resulting Ct value to the standard curve generated from a serial dilution of the VVD cDNA qPCR standard prepared by reverse transcribing the RNA transcript.

Negative strand assay. Strand-specific detection of DWV RNA was performed as described earlier. Briefly, 1 µg of total RNA was used in reverse transcription reaction carried out with Superscript III reverse transcriptase (Invitrogen) and the adapter extended primer designed to anneal to the negative strand RNA of DWV. The PCR step was carried out by *Taq* DNA polymerase (New England Biolabs) using forward primer 388 identical to the adapter sequence (Table S1) and DWV_RTPCR_R primer. PCR was run for 35 cycles in the same conditions as described above.

Next generation sequencing sample preparation and analysis. Gene specific cDNA was generated from selected samples using SSIII reverse transcriptase (Invitrogen) and DWV FG RP1 primer (Table S1). The cDNAs were amplified using LongAmp *Taq* polymerase (New England Biolabs) following standard protocols and DWV FG RP1 and FP4 primers to produce a single ~ 10 Kb PCR fragment spanning the DWV genome. Following PCR purification (Wizard PCR cleanup kit, Promega) samples were sequenced using an Illumina Hiseq. Paired-end reads were processed at the University of St Andrews and sequences were analyzed for genetic diversity using ShoRAH[59,60]. Phylogenetic analysis was performed using Geneious Prime 2019.1.3.

Received: 5 February 2020; Accepted: 22 September 2020
Published online: 08 October 2020

References

1. Martin, S. J. *et al.* Global honey bee viral landscape altered by a parasitic mite. *Science* **336**, 1304–1306. https://doi.org/10.1126/science.1220941 (2012).
2. Wilfert, L. *et al.* Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites. *Science* **351**, 594–597. https://doi.org/10.1126/science.aac9976 (2016).

3. Gallai, N., Salles, J. M., Settele, J. & Vaissière, B. E. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecol. Econom.* **68**, 810–821. https://doi.org/10.1016/j.ecolecon.2008.06.014 (2009).
4. Potts, S. G. *et al.* Safeguarding pollinators and their values to human well-being. *Nature* **540**, 220–229. https://doi.org/10.1038/nature20588 (2016).
5. Cavigli, I. *et al.* Pathogen prevalence and abundance in honey bee colonies involved in almond pollination. *Apidologie* **47**, 251–266. https://doi.org/10.1007/s13592-015-0395-5 (2016).
6. Fürst, M. A., McMahon, D. P., Osborne, J. L., Paxton, R. J. & Brown, M. J. F. Disease associations between honeybees and bumblebees as a threat to wild pollinators. *Nature* **506**, 364–366. https://doi.org/10.1038/nature12977 (2014).
7. Singh, R. *et al.* RNA viruses in hymenopteran pollinators: Evidence of inter-taxa virus transmission via pollen and potential impact on non-Apis hymenopteran species. *PLoS ONE* https://doi.org/10.1371/journal.pone.0014357 (2010).
8. Chen, Y. P., Pettis, J. S., Collins, A. & Feldlaufer, M. F. Prevalence and transmission of honeybee viruses. *Appl. Environ. Microbiol.* **72**, 606–611. https://doi.org/10.1128/AEM.72.1.606-611.2006 (2006).
9. Genersch, E., Yue, C., Fries, I. & De Miranda, J. R. Detection of Deformed wing virus, a honey bee viral pathogen, in bumble bees (*Bombus terrestris* and *Bombus pascuorum*) with wing deformities. *J. Invertebr. Pathol.* **91**, 61–63. https://doi.org/10.1016/j.jip.2005.10.002 (2006).
10. Loope, K. J., Baty, J. W., Lester, P. J. & Wilson Rankin, E. E. Pathogen shifts in a honeybee predator following the arrival of the *Varroa* mite. *Proc. Biol. Sci.* **286**, 20182499. https://doi.org/10.1098/rspb.2018.2499 (2019).
11. Zhang, X. *et al.* New evidence that deformed wing virus and black queen cell virus are multi-host pathogens. *J. Invertebr. Pathol.* **109**, 156–159. https://doi.org/10.1016/j.jip.2011.09.010 (2012).
12. Forsgren, E. *et al.* Preliminary observations on possible pathogen spill-over from Apis *mellifera* to *Apis cerana*. *Apidologie* **46**, 265–275. https://doi.org/10.1007/s13592-014-0320-3 (2015).
13. Levitt, A. L. *et al.* Cross-species transmission of honey bee viruses in associated arthropods. *Virus Res.* **176**, 232–240. https://doi.org/10.1016/j.virusres.2013.06.013 (2013).
14. Reynaldi, F. J., Sguazza, G. H., Albicoro, F. J., Pecoraro, M. R. & Galosi, C. M. First molecular detection of co-infection of honey bee viruses in asymptomatic *Bombus atratus* in South America. *Braz. J. Biol.* **73**, 797–800. https://doi.org/10.1590/s1519-69842013000400016 (2013).
15. Gamboa, V. *et al.* Bee pathogens found in Bombus atratus from Colombia: A case study. *J. Invertebr. Pathol.* **129**, 36–39. https://doi.org/10.1016/j.jip.2015.05.013 (2015).
16. Graystock, P., Meeus, I., Smagghe, G., Goulson, D. & Hughes, W. O. H. The effects of single and mixed infections of *Apicystis bombi* and deformed wing virus in *Bombus terrestris*. *Parasitology* **143**, 358–365. https://doi.org/10.1017/s0031182015001614 (2016).
17. Sachman-Ruiz, B., Narváez-Padilla, V. & Reynaud, E. Commercial *Bombus impatiens* as reservoirs of emerging infectious diseases in central México. *Biol. Invasions* **17**, 2043–2053. https://doi.org/10.1007/s10530-015-0859-6 (2015).
18. Sébastien, A. *et al.* Invasive ants carry novel viruses in their new range and form reservoirs for a honeybee pathogen. *Biol. Lett.* **11**(9), 20150610. https://doi.org/10.1098/rsbl.2015.0610 (2015).
19. Guzman-Novoa, E. *et al.* First detection of honey bee viruses in stingless bees in North America. *J. of Apicultural Res.* **54**, 93–95. https://doi.org/10.1080/00218839.2015.1100154 (2015).
20. Ravoet, J. *et al.* Widespread occurrence of honey bee pathogens in solitary bees. *J. Invertebr. Pathol.* **122**, 55–58. https://doi.org/10.1016/j.jip.2014.08.007 (2014).
21. Lucia, M., Reynaldi, F. J., Sguazza, G. H. & Abrahamovich, A. H. First detection of deformed wing virus in Xylocopa augusti larvae (Hymenoptera: Apidae) in Argentina. *J. Apic. Res.* **53**, 466–468. https://doi.org/10.3896/IBRA.1.53.4.11 (2014).
22. Bingham, R. A. & Orthner, A. R. Efficient pollination of alpine plants. *Nature* **391**, 238–239. https://doi.org/10.1038/34564 (1998).
23. Velthuis, H. H. W. & van Doorn, A. A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. *Apidologie* **37**, 421–451. https://doi.org/10.1051/apido:2006019 (2006).
24. Li, J. *et al.* Cross-species infection of deformed wing virus poses a new threat to pollinator conservation. *J. Econ. Entomol.* **104**, 732–739. https://doi.org/10.1603/ec10355 (2011).
25. Radzevičiūtė, R. *et al.* Replication of honey bee-associated RNA viruses across multiple bee species in apple orchards of Georgia Germany and Kyrgyzstan. *J. Invertebr. Pathol.* **146**, 14–23. https://doi.org/10.1016/j.jip.2017.04.002 (2017).
26. Alger, S. A., Burnham, P. A., Boncristiani, H. F. & Brody, A. K. RNA virus spillover from managed honeybees (*Apis mellifera*) to wild bumblebees (*Bombus* spp.). *PLoS ONE* **14**, e0217822. https://doi.org/10.1371/journal.pone.0217822 (2019).
27. Manley, R. *et al.* Knock-on community impacts of a novel vector: spillover of emerging DWV-B from Varroa-infested honeybees to wild bumblebees. *Ecol. Lett.* **22**, 1306–1315. https://doi.org/10.1111/ele.13323 (2019).
28. Ryabov, E. V. *et al.* A Virulent strain of deformed wing virus (DWV) of honeybees (*Apis mellifera*) prevails after *Varroa destructor*-mediated, or in vitro, transmission. *PLoS Path.* **10**(6), e1004230. https://doi.org/10.1371/journal.ppat.1004230 (2014).
29. Tehel, A., Brown, M. J. F. & Paxton, R. J. Impact of managed honey bee viruses on wild bees. *Curr. Opin. Virol.* **19**, 16–22. https://doi.org/10.1016/j.coviro.2016.06.006 (2016).
30. Lanzi, G. *et al.* Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *J. Virol.* **80**, 4998–5009. https://doi.org/10.1128/jvi.80.10.4998-5009.2006 (2006).
31. Fujiyuki, T. *et al.* Prevalence and phylogeny of Kakugo virus, a novel insect picorna-like virus that infects the honeybee (*Apis mellifera* L.), under various colony conditions. *J. Virol.* **80**, 11528–11538. https://doi.org/10.1128/jvi.00754-06 (2006).
32. Mordecai, G. J., Wilfert, L., Martin, S. J., Jones, I. M. & Schroeder, D. C. Diversity in a honey bee pathogen: First report of a third master variant of the Deformed Wing Virus quasispecies. *ISME J.* **10**, 1–10. https://doi.org/10.1038/ismej.2015.178 (2015).
33. Ongus, J. R. *et al.* Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite Varroa destructor. *J. Gen. Virol.* **85**, 3747–3755. https://doi.org/10.1099/vir.0.80470-0 (2004).
34. Moore, J. *et al.* Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in *Varroa destructor*-infested honeybee colonies. *J. Gen. Virol.* **92**, 156–161. https://doi.org/10.1099/vir.0.025965-0 (2011).
35. Zioni, N., Soroker, V. & Chejanovsky, N. Replication of varroa destructor virus 1 (VDV-1) and a varroa destructor virus 1-deformed wing virus recombinant (VDV-1-DWV) in the head of the honey bee. *Virology* **417**, 106–112. https://doi.org/10.1016/j.virol.2011.05.009 (2011).
36. Ryabov, E. V. *et al.* Recent spread of Varroa destructor virus-1, a honey bee pathogen, in the United States. *Sci. Rep.* **7**, 17447. https://doi.org/10.1038/s41598-017-17802-3 (2017).
37. Dalmon, A. *et al.* Evidence for positive selection and recombination hotspots in Deformed wing virus (DWV). *Sci. Rep.* **7**, 41045. https://doi.org/10.1038/srep41045 (2017).
38. Baker, A. C. & Schroeder, D. C. The use of RNA-dependent RNA polymerase for the taxonomic assignment of Picorna-like viruses (order Picornavirales) infecting *Apis mellifera* L. populations. *Virol. J.* **5**, 10. https://doi.org/10.1186/1743-422x-5-10 (2008).
39. Lamp, B. *et al.* Construction and rescue of a molecular clone of deformed wing virus (DWV). *PLoS ONE* **11**(11), e0164639. https://doi.org/10.1371/journal.pone.0164639 (2016).
40. Posada-Florez, F. *et al.* Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite Varroa destructor in a non-propagative manner. *Sci. Rep.* **9**, 12445. https://doi.org/10.1038/s41598-019-47447-3 (2019).
41. Ryabov, E. V. *et al.* Dynamic evolution in the key honey bee pathogen deformed wing virus: Novel insights into virulence and competition using reverse genetics. *PLoS Biol.* **17**, e3000502. https://doi.org/10.1371/journal.pbio.3000502 (2019).

42. Gusachenko, O. N. *et al.* Green Bees: reverse genetic analysis of deformed wing virus transmission, replication, and tropism. *Viruses* **12**, 532. https://doi.org/10.3390/v12050532 (2020).
43. Ongus, J. R., Roode, E. C., Pleij, C. W. A., Vlak, J. M. & van Oers, M. M. The 5' non-translated region of Varroa destructor virus 1 (genus *Iflavirus*): Structure prediction and IRES activity in *Lymantria dispar* cells. *J. Gen. Virol.* **87**, 3397–3407. https://doi.org/10.1099/vir.0.82122-0 (2006).
44. McMahon, D. P. *et al.* Elevated virulence of an emerging viral genotype as a driver of honeybee loss. *Proc. Biol. Sci.* **283**, 443–449. https://doi.org/10.1098/rspb.2016.0811 (2016).
45. Mordecai, G. J. *et al.* Superinfection exclusion and the long-term survival of honey bees in Varroa-infested colonies. *ISME J.* **10**, 1–10. https://doi.org/10.1038/ismej.2015.186 (2015).
46. Tehel, A. *et al.* The two prevalent genotypes of an emerging infectious disease cause equally low pupal mortality and equally high wing deformities in host honey bees. *Viruses* **11**, 114. https://doi.org/10.3390/v11020114 (2019).
47. Tian, L. & Hines, H. M. Morphological characterization and staging of bumble bee pupae. *PeerJ* **6**, e6089. https://doi.org/10.7717/peerj.6089 (2018).
48. Becker, D. J. *et al.* Dynamic and integrative approaches to understanding pathogen spillover. *Trans. R. Soc. Lond. B Biol. Sci Philos* https://doi.org/10.1098/rstb.2019.0014 (2019).
49. Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510. https://doi.org/10.1038/nrmicro.2017.45 (2017).
50. Dainat, B., Evans, J. D., Chen, Y. P., Gauthier, L. & Neumanna, P. Dead or alive: Deformed wing virus and varroa destructor reduce the life span of winter honeybees. *Appl. Environ. Microbiol.* **78**, 981–987. https://doi.org/10.1128/aem.06537-11 (2012).
51. Goulson, D., Nicholls, E., Botías, C. & Rotheray, E. L. Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science* **347**, 1255957. https://doi.org/10.1126/science.1255957 (2015).
52. Roberts, J. M. K., Anderson, D. L. & Durr, P. A. Absence of deformed wing virus and Varroa destructor in Australia provides unique perspectives on honeybee viral landscapes and colony losses. *Sci. Rep.* **7**, 6925. https://doi.org/10.1038/s41598-017-07290-w (2017).
53. Belsky, J. & Joshi, N. K. Impact of biotic and abiotic stressors on managed and feral bees. *Insects* **10**, 233. https://doi.org/10.3390/insects10080233 (2019).
54. Di Prisco, G. *et al.* Neonicotinoid clothianidin adversely affects insect immunity and promotes replication of a viral pathogen in honey bees. *Proc. Nat. Acad. Sci.* **110**, 18466–18471. https://doi.org/10.1073/pnas.1314923110 (2013).
55. Manley, R., Boots, M. & Wilfert, L. Condition-dependent virulence of slow bee paralysis virus in *Bombus terrestris*: are the impacts of honeybee viruses in wild pollinators underestimated? *Oecologia* **184**, 305–315. https://doi.org/10.1007/s00442-017-3851-2 (2017).
56. Mazzei, M. *et al.* Infectivity of DWV associated to flower pollen: Experimental evidence of a horizontal transmission route. *PLoS ONE* **9**, e113448. https://doi.org/10.1371/journal.pone.0113448 (2014).
57. Yoneda, M., Furuta, H., Tsuchida, K., Okabe, K. & Goka, K. Commercial colonies of *Bombus terrestris* (Hymenoptera: Apidae) are reservoirs of the tracheal mite *Locustacarus buchneri* (Acari: Podapolipidae). *Appl. Entomol. and Zool.* **43**, 73–76. https://doi.org/10.1303/aez.2008.73 (2008).
58. de Pereira, S., Meeus, I. & Smagghe, G. Honey bee-collected pollen is a potential source of Ascosphaera apis infection in managed bumble bees. *Sci. Rep.* **9**, 4241. https://doi.org/10.1038/s41598-019-40804-2 (2019).
59. Zagordi, O., Bhattacharya, A., Eriksson, N. & Beerenwinkel, N. ShoRAH: Estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinform.* **12**, 119. https://doi.org/10.1186/1471-2105-12-119 (2011).
60. McElroy, K., Zagordi, O., Bull, R., Luciani, F. & Beerenwinkel, N. Accurate single nucleotide variant detection in viral populations by combining probabilistic clustering with a statistical test of strand bias. *BMC Genom.* **14**, 501. https://doi.org/10.1186/1471-2164-14-501 (2013).

Acknowledgements

This work was supported by Grant funding from BBSRC BB/M00337X/2 and BB/I000828/1. This research was also supported by the United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) Grant 2017-06481 (EVR).

Author contributions

D.J.E. and O.N.G. designed the study. O.N.G., L.W. and K.B.-C. performed and analyzed experiments. E.V.R. developed a reverse genetics system for VVD DWV. D.J.E. supervised the study. O.N.G. wrote the manuscript. D.J.E. and L.W. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-73809-3.

Correspondence and requests for materials should be addressed to O.N.G. or D.J.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

9 References

Agol, V. I. and Gmyl, A. P. (2018) 'Emergency Services of Viral RNAs: Repair and Remodeling', *Microbiology and Molecular Biology Reviews*. doi: 10.1128/mmbr.00067-17.

Aizen, Marcelo A and Harder, L. D. (2009) 'Report The Global Stock of Domesticated Honey Bees Is Growing Slower Than Agricultural Demand for Pollination', *Current Biology*. Elsevier Ltd, 19(11), pp. 915–918. doi: 10.1016/j.cub.2009.03.071.

Aizen, Marcelo A. and Harder, L. D. (2009) 'The Global Stock of Domesticated Honey Bees Is Growing Slower Than Agricultural Demand for Pollination', *Current Biology*. doi: 10.1016/j.cub.2009.03.071.

Almond, J. W. *et al.* (1987) 'Studies on the attenuation of the Sabin type 3 oral polio vaccine', *Journal of Virological Methods*. doi: 10.1016/0166-0934(87)90081-4.

Alnaji, F. G. *et al.* (2020) 'Recombination in enteroviruses is a ubiquitous event independent of sequence homology and RNA structure', *bioRxiv*. doi: 10.1101/2020.09.29.319285.

Amiri, E., Meixner, M. D. and Kryger, P. (2016) 'Deformed wing virus can be transmitted during natural mating in honey bees and infect the queens', *Scientific Reports*. doi: 10.1038/srep33065.

Anderson, D. L. and Trueman, J. W. H. (2000) 'Varroa jacobsonii is more than one species', *Experimental and Appl ied Acarology*.

Annoscia, D. *et al.* (2019) 'Haemolymph removal by Varroa mite destabilizes the dynamical interaction between immune effectors and virus in bees, as predicted by Volterra's model', *Proceedings of the Royal Society B: Biological Sciences*, 286(1901). doi: 10.1098/rspb.2019.0331.

Bailey, L. (1969) 'The multiplication and spread of sacbrood virus of bees', *Annals of Applied Biology*. doi: 10.1111/j.1744-7348.1969.tb02844.x.

Bailey, L. and Ball, B. V. (1991) 'The Honey Bee', in *Honey Bee Pathology*. doi: 10.1016/b978-0-12-073481-8.50006-0.

Bakonyi, T. *et al.* (2002) 'Detection of acute bee paralysis virus by RT-PCR in honey bee and Varroa destructor field samples: Rapid screening of representative Hungarian apiaries', *Apidologie*. doi: 10.1051/apido:2001004.

Barclay, W. *et al.* (1998) 'Encapsidation studies of poliovirus subgenomic replicons', *Journal of General Virology*. doi: 10.1099/0022-1317-79-7-1725.

Barrera, R. *et al.* (2019) 'Citywide Control of Aedes aegypti (Diptera: Culicidae) during the 2016 Zika Epidemic by Integrating Community Awareness, Education, Source Reduction, Larvicides, and Mass Mosquito Trapping', *Journal of Medical Entomology*. doi: 10.1093/jme/tjz009.

Beaurepaire, A. *et al.* (2020) 'Diversity and global distribution of viruses of the western honey bee, Apis mellifera', pp. 1–25. doi: 10.3390/insects11040239.

Becher, M. A. *et al.* (2014) 'BEEHAVE: A systems model of honeybee colony dynamics and foraging to explore multifactorial causes of colony failure', *Journal of Applied Ecology*, 51(2), pp. 470–482. doi: 10.1111/1365-2664.12222.

Bentley, K. *et al.* (2021) 'Imprecise recombinant viruses evolve via a fitness-driven, iterative process of polymerase template-switching events', *bioRxiv*.

Bentley, K. and Evans, D. J. (2018) 'Mechanisms and consequences of positive-strand RNA virus recombination', *Journal of General Virology*, 99(10), pp. 1345–1356. doi: 10.1099/jgv.0.001142.

Bonhoeffer, S. *et al.* (1997) 'Virus dynamics and drug therapy', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.94.13.6971.

Boot, W. J. *et al.* (1995) 'Invasion of Varroa jacobsoni into drone brood cells of the honey bee, Apis mellifera', *Apidologie*. doi: 10.1051/apido:19950204.

Bowen-Walker, P. L. and Gunn, A. (2001) 'The effect of the ectoparasitic mite, Varroa destructor on adult worker honeybee (Apis mellifera) emergence weights, water, protein, carbohydrate, and lipid levels', *Entomologia Experimentalis et Applicata*, 101(3), pp. 207–217. doi: 10.1046/j.1570-7458.2001.00905.x.

Brown, B. *et al.* (2003) 'Complete Genomic Sequencing Shows that Polioviruses and Members of Human Enterovirus Species C Are Closely Related in the Noncapsid Coding Region', *Journal of Virology*, 77(16), pp. 8973–8984. doi: 10.1128/jvi.77.16.8973-8984.2003.

Brown, M. J. F. and Paxton, R. J. (2009) 'Review article The conservation of bees : a global perspective *', *Apidologie*, 40, pp. 410–416.

Bruckner, S. *et al.* (2020) '2019-2020 Honey Bee Colony Losses in the United States: Preliminary Results Embargoed until Monday, June 22', pp. 3–7.

Bull, R. A. *et al.* (2011) 'Sequential bottlenecks drive viral evolution in early acute hepatitis c virus infection', *PLoS Pathogens*. doi: 10.1371/journal.ppat.1002243.

Campbell, E. M. *et al.* (2016) 'Transcriptome analysis of the synganglion from the honey bee mite, Varroa destructor and RNAi knockdown of neural peptide targets', *Insect Biochemistry and Molecular Biology*. doi: 10.1016/j.ibmb.2015.12.007.

Chao, L. (1990) 'Fitness of RNA virus decreased by Muller's ratchet', *Nature*. doi: 10.1038/348454a0.

Chauzat, M.-P. *et al.* (2009) 'Influence of pesticide residues on honey bee (Hymenoptera: Apidae) colony health in France.', *Environmental entomology*, 38(3), pp. 514–23. doi: 10.1603/022.038.0302.

Chen, Y. P., Higgins, J. A. and Feldlaufer, M. F. (2005) 'Quantitative real-time reverse transcription-PCR analysis of deformed wing virus infection in the honeybee (Apis mellifera L.)', *Appl Environ Microbiol*, 71(1), pp. 436–441. doi: 10.1128/AEM.71.1.436.

Coffey, L. L. and Vignuzzi, M. (2011) 'Host Alternation of Chikungunya Virus Increases Fitness while Restricting Population Diversity and Adaptability to Novel Selective Pressures', *Journal of Virology*. doi: 10.1128/jvi.01918-10.

Cornman, R. S. (2017) 'Relative abundance of deformed wing virus, Varroa destructor virus 1, and their recombinants in honey bees (Apis mellifera) assessed by kmer analysis of public RNA-Seq data', *Journal of Invertebrate Pathology*. Elsevier, 149(July), pp. 44–50. doi: 10.1016/j.jip.2017.07.005.

Costanza, R. *et al.* (1997) 'The value of the world's ecosystem services and natural capital', *Nature*. doi: 10.1038/387253a0.

Cox-foster, D. L. *et al.* (2007) 'A Metagenomic Survey of Collapse Disorder', *October*, 318(October), pp. 283–287. doi: 10.1126/science.1146498.

Dainat, B. *et al.* (2012) 'Dead or alive: Deformed wing virus and varroa destructor reduce the life span of winter honeybees', *Applied and Environmental Microbiology*, 78(4), pp. 981–987. doi: 10.1128/AEM.06537-11.

Dalmon, A. *et al.* (2017) 'Evidence for positive selection and recombination hotspots in Deformed wing virus (DWV)', *Scientific Reports*. Nature Publishing Group, (December 2016), pp. 1–12. doi: 10.1038/srep41045.

Dalmon, A. *et al.* (2019) 'Viruses in the Invasive Hornet Vespa velutina', *Viruses*, 11(1041), pp. 1–22.

Delaplane, K. S. and Mayer, D. F. (2000) *Crop pollination by bees.*, *Crop pollination by bees.* doi: 10.1079/9780851994482.0000.

Dietemann, V., Ellis, J. D. and Neumann, P. (2013) 'The COLOSS BEEBOOK Volume II, Standard methods for Apis mellifera pest and pathogen research: Introduction', *Journal of*

Apicultural Research. doi: 10.3896/IBRA.1.52.4.16.

Döke, M. A. *et al.* (2019) 'Colony Size, Rather Than Geographic Origin of Stocks, Predicts Overwintering Success in Honey Bees (Hymenoptera: Apidae) in the Northeastern United States', *Journal of Economic Entomology*. doi: 10.1093/jee/toy377.

Domingo, E. *et al.* (1978) 'Nucleotide sequence heterogeneity of an RNA phage population', *Cell*. doi: 10.1016/0092-8674(78)90223-4.

Domingo, E. *et al.* (2005) 'Quasispecies dynamics and RNA virus extinction', *Virus Research*. doi: 10.1016/j.virusres.2004.11.003.

Domingo, E. and Perales, C. (2019) 'Viral quasispecies', *PLoS Genetics*, 15(10), pp. 1–20. doi: 10.1371/journal.pgen.1008271.

Domingo, E., Sheldon, J. and Perales, C. (2012) 'Viral Quasispecies Evolution', *Microbiology and Molecular Biology Reviews*, 76(2), pp. 159–216. doi: 10.1128/MMBR.05023-11.

Donaldson-Matasci, M. C. and Dornhaus, A. (2012) 'How habitat affects the benefits of communication in collectively foraging honey bees', *Behavioral Ecology and Sociobiology*. doi: 10.1007/s00265-011-1306-z.

van Dooremalen, C. *et al.* (2012) 'Winter survival of individual honey bees and honey bee colonies depends on level of varroa destructor infestation', *PLoS ONE*, 7(4). doi: 10.1371/journal.pone.0036285.

Duffy, S., Shackelton, L. A. and Holmes, E. C. (2008) 'Rates of evolutionary change in viruses: Patterns and determinants', *Nature Reviews Genetics*. doi: 10.1038/nrg2323.

Edgar, R. C. (2004) 'MUSCLE: Multiple sequence alignment with high accuracy and high throughput', *Nucleic Acids Research*. doi: 10.1093/nar/gkh340.

Elena, S. F., Cooper, V. S. and Lenski, R. E. (1996) 'Punctuated evolution caused by selection of rare beneficial mutations', *Science*. doi: 10.1126/science.272.5269.1802.

Figlerowicz, Magdalena *et al.* (2003) 'Genetic variability: The key problem in the prevention and therapy of RNA-based virus infections', *Medicinal Research Reviews*. doi: 10.1002/med.10045.

Forfert, N. *et al.* (2015) 'Parasites and pathogens of the honeybee (Apis mellifera) and their influence on inter-colonial transmission', *PLoS ONE*, 10(10), pp. 1–14. doi: 10.1371/journal.pone.0140337.

Fries, I. *et al.* (2003) 'Swarming in honey bees (Apis mellifera) and Varroa destructor population development in Sweden', *Apidologie*. doi: 10.1051/apido:2003032.

Fries, I. and Camazine, S. (2001) 'Implications of horizontal and vertical pathogen transmission for honey bee epidemiology', *Apidologie*, 32(3), pp. 199–214. doi: 10.1051/apido:2001122.

Fries, I., Imdorf, A. and Rosenkranz, P. (2006) 'Survival of mite infested (Varroa destructor) honey bee (Apis mellifera) colonies in a Nordic climate', *Apidologie*, 37, pp. 564–570. doi: 10.1051/apido.

Fuchs, S. (1992) 'Choice in Varroa jacobsoni Oud. between honey bee drone or workerbrood cells for reproduction', *Behavioral Ecology and Sociobiology*. doi: 10.1007/BF00170610.

Fujiyuki, T., Takeuchi, H., Ono, M., Ohka, S., Sasaki, T., Nomoto, a., *et al.* (2004) 'Novel Insect Picorna-Like Virus Identified in the Brains of Aggressive Worker Honeybees', *Journal of Virology*, 78(3), pp. 1093–1100. doi: 10.1128/JVI.78.3.1093-1100.2004.

Fujiyuki, T., Takeuchi, H., Ono, M., Ohka, S., Sasaki, T., Nomoto, A., *et al.* (2004) 'Novel Insect Picorna-Like Virus Identified in the Brains of Aggressive Worker Honeybees', *Journal of Virology*, 78(3), pp. 1093–1100. doi: 10.1128/jvi.78.3.1093-1100.2004.

Fujiyuki, T. *et al.* (2005) 'Kakugo Virus from Brains of Aggressive Worker Honeybees', *Advances in Virus Research*. doi: 10.1016/S0065-3527(05)65001-4.

Fürst, M. A. *et al.* (2014) 'Disease associations between honeybees and bumblebees as a threat to wild pollinators', *Nature*, 506(7488), pp. 364–366. doi: 10.1038/nature12977.

Gisder, S. *et al.* (2018) 'In vivo evolution of viral virulence: switching of deformed wing virus between hosts results in virulence changes and sequence shifts', *Environmental Microbiology*, 20(12), pp. 4612–4628. doi: 10.1111/1462-2920.14481.

Gisder, S. and Genersch, E. (2020) ' Direct Evidence for Infection of Varroa destructor Mites with the Bee-Pathogenic Deformed Wing Virus Variant B - but Not Variant A - via Fluorescence- in situ -Hybridization Analysis. ', *Journal of Virology*, (December). doi: 10.1128/jvi.01786-20.

Giuffre, C., Lubkin, S. R. and Tarpy, D. R. (2019) 'Does viral load alter behavior of the bee parasite Varroa destructor?', *PLoS ONE*. doi: 10.1371/journal.pone.0217975.

Giusti, M. *et al.* (2016) 'Scientific note: varroa mite eradication, the strange case of Gorgona Island', *Apidologie*, 47(5), pp. 688–690. doi: 10.1007/s13592-015-0417-3.

Gladys N Macharia *et al.* (2020) 'Infection with multiple HIV-1 founder variants is associated with lower viral replicative capacity , faster CD4 + T cell decline and increased immune activation during acute infection', *PLoS Pathogens*, pp. 1–22. doi: 10.1371/journal.ppat.1008853.

Glaser, W., Cencic, R. and Skern, T. (2001) 'Foot-and-mouth disease virus leader proteinase: Involvement of C-terminal residues in self-processing and cleavage of eIF4GI', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M104192200.

Gourlé, H. *et al.* (2019) 'Simulating Illumina metagenomic data with InSilicoSeq', *Bioinformatics*. doi: 10.1093/bioinformatics/bty630.

Grabensteiner, E. *et al.* (2001) 'Sacbrood virus of the honeybee (Apis mellifera): Rapid identification and phylogenetic analysis using reverse transcription-PCR', *Clinical and Diagnostic Laboratory Immunology*. doi: 10.1128/CDLI.8.1.93-104.2001.

Gray, A. *et al.* (2020) 'Honey bee colony winter loss rates for 35 countries participating in the COLOSS survey for winter 2018–2019, and the effects of a new queen on the risk of colony winter loss', *Journal of Apicultural Research*. Taylor & Francis, 0(0), pp. 1–8. doi: 10.1080/00218839.2020.1797272.

Greatti, M., Milani, N. and Nazzi, F. (1992) 'Reinfestation of an acaricide-treated apiary by Varroa jacobsoni Oud.', *Experimental and Applied Acarology*, 16(4), pp. 279–286. doi: 10.1007/BF01218569.

Grubaugh, N. D., Weger-Lucarelli, J., *et al.* (2016) 'Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching', *Cell Host and Microbe*. doi: 10.1016/j.chom.2016.03.002.

Grubaugh, N. D., Rückert, C., *et al.* (2016) 'Transmission bottlenecks and RNAi collectively influence tick-borne flavivirus evolution', *Virus Evolution*, 2(2), p. vew033. doi: 10.1093/ve/vew033.

Gusachenko, O. N., Woodford, L., Balbirnie-Cumming, K., Ryabov, E. V., *et al.* (2020) 'Evidence for and against deformed wing virus spillover from honey bees to bumble bees : a reverse genetic analysis', *Scientific Reports*. Nature Publishing Group UK, pp. 1–10. doi: 10.1038/s41598-020-73809-3.

Gusachenko, O. N., Woodford, L., Balbirnie-Cumming, K., Campbell, E. M., *et al.* (2020) 'Green Bees: Reverse Genetic Analysis of Deformed Wing Virus Transmission, Replication and Tropism', *Viruses*, 12(532). doi: 10.3390/v12050532.

Gusachenko, O. N. *et al.* (2021) 'First come, first served: Superinfection exclusion in Deformed wing virus is dependent upon sequence identity and not the order of virus acquisition', *ISME Journal*. Springer US, (June), pp. 1–10. doi: 10.1038/s41396-021-01043-4.

Hallmann, C. A. *et al.* (2017) 'More than 75 percent decline over 27 years in total flying

insect biomass in protected areas', *PLoS ONE*, 12(10). doi: 10.1371/journal.pone.0185809.

Han, F., Wallberg, A. and Webster, M. T. (2012) 'From where did the western honeybee (Apis mellifera) originate?', *Ecology and Evolution*, 2(8), pp. 1949–1957. doi: 10.1002/ece3.312.

Harrell, F. E. (2020) 'Hmisc: Harrell miscellaneous', *R package version 4.4-0*.

Hayes, J. (2007) 'Colony collapse disorder', *American Bee Journal*. doi: 10.1007/978-3-662-43978-4_4587.

Highfield, A. C. *et al.* (2009) 'Deformed wing virus implicated in overwintering honeybee colony losses', *Applied and Environmental Microbiology*, 75(22), pp. 7212–7220. doi: 10.1128/AEM.02227-09.

Hill, J. T. *et al.* (2014) 'Poly Peak Parser: Method and software for identification of unknown indels using Sanger Sequencing of PCR products Jonathon', *Developmental Dynamics*, 2(2), pp. 1632–1636. doi: 10.14440/jbm.2015.54.A.

Hill, J. T. (2017) 'Walkthrough for using the sangerseqR package', pp. 1–12.

Hinton, T. M. *et al.* (2002) 'Conservation of L and 3C proteinase activities across distantly related aphthoviruses', *Journal of General Virology*. doi: 10.1099/0022-1317-83-12-3111.

Hou, C. S. *et al.* (2016) 'Effects of varroa destructor on temperature and humidity conditions and expression of energy metabolism genes in infested honeybee colonies', *Genetics and Molecular Research*. doi: 10.4238/gmr.15038997.

Iqbal, J. and Mueller, U. (2007) 'Virus infection causes specific learning deficits in honeybee foragers', *Proceedings of the Royal Society B: Biological Sciences*. doi: 10.1098/rspb.2007.0022.

Kassambara, A. (2020) '"ggpubr": "ggplot2" Based Publication Ready Plots', *R package version 0.2.5.*

Kearse, M. *et al.* (2012) 'Geneious', *Bioinformatics (Oxford, England)*. doi: 10.1093/bioinformatics/bts199.

Kevill, J. L. *et al.* (2019) 'DWV-A lethal to honey bees (apis mellifera): A colony level survey of DWV Variants (A, B, and C) in England, Wales, and 32 States across the US', *Viruses*, 11(5). doi: 10.3390/v11050426.

Kirkegaard, K. and Baltimore, D. (1986) 'The mechanism of RNA recombination in poliovirus', *Cell*. doi: 10.1016/0092-8674(86)90600-8.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets.', *Molecular biology and evolution*, p. msw054. doi: 10.1093/molbev/msw054.

De La Rua, P. *et al.* (2009) 'Biodiversity, conservation and current threats to European honeybees', *Apidologie*, 40(December), pp. 263–284. doi: 10.1051/apido/2009027.

Langmead, B. *et al.* (2009) 'Ultrafast and memory-efficient alignment of short DNA sequences to the human genome', *Genome Biology*. doi: 10.1186/gb-2009-10-3-r25.

Lanzi, G. *et al.* (2006) 'Molecular and Biological Characterization of Deformed Wing Virus of Honeybees (Apis mellifera L .)', *Society*, 80(10), pp. 4998–5009. doi: 10.1128/JVI.80.10.4998.

Lequime, S. *et al.* (2016) 'Genetic Drift, Purifying Selection and Vector Genotype Shape Dengue Virus Intra-host Genetic Diversity in Mosquitoes', *PLoS Genetics*. doi: 10.1371/journal.pgen.1006111.

Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*. doi: 10.1093/bioinformatics/btp352.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*. doi: 10.1093/bioinformatics/btp324.

Lindström, S. A. M. *et al.* (2016) 'Experimental evidence that honeybees depress wild insect densities in a flowering crop', *Proceedings of the Royal Society B: Biological Sciences*.

doi: 10.1098/rspb.2016.1641.

Locke, B. *et al.* (2017) 'Persistence of subclinical deformed wing virus infections in honeybees following Varroa mite removal and a bee population turnover', pp. 1–10. doi: 10.1371/journal.pone.0180910.

Loftus, J. C., Smith, M. L. and Seeley, T. D. (2016) 'How honey bee colonies survive in the wild: Testing the importance of small nests and frequent swarming', *PLoS ONE*, 11(3), pp. 1–11. doi: 10.1371/journal.pone.0150362.

Longdon, B. *et al.* (2014) 'The Evolution and Genetics of Virus Host Shifts', *PLoS Pathogens*. doi: 10.1371/journal.ppat.1004395.

Longdon, B. *et al.* (2018) 'Host shifts result in parallel genetic changes when viruses evolve in closely related species', *PLoS Pathogens*, 14(4), pp. 1–14. doi: 10.1371/journal.ppat.1006951.

Lowry, K. *et al.* (2014) 'Recombination in Enteroviruses Is a Biphasic Replicative Process Involving the Generation of Greater-than Genome Length "Imprecise" Intermediates', *PLoS Pathogens*, 10(6). doi: 10.1371/journal.ppat.1004191.

Lukashev, A. N. *et al.* (2003) 'Recombination in Circulating Enteroviruses', *Journal of Virology*. doi: 10.1128/jvi.77.19.10423-10431.2003.

Lukashev, A. N. (2005) 'Role of recombination in evolution of enteroviruses', *Reviews in Medical Virology*. doi: 10.1002/rmv.457.

Martín-Hernandez, R. *et al.* (2007) 'Short term negative effect of oxalic acid in Apis mellifera iberisensis', *Spanish Journal of Agricultural Research*, 5(4), pp. 474–480. doi: 10.5424/sjar/2007054-270.

Martin, S. J. (2001) 'The role of *Varroa* and viral pathogens in the collapse of honeybee colonies: A modelling approach', *Journal of Applied Ecology*, 38(Allen 1960), pp. 1082–1093.

Martin, S. J. *et al.* (2012a) 'Global Honey Bee Viral Landscape Altered by a Parasitic Mite', *Science*, 336, pp. 1304–1306. doi: 10.1126/science.1220941.

Martin, S. J. *et al.* (2012b) 'Global honey bee viral landscape altered by a parasitic mite - supplementary material', *Science*, 336(6086), pp. 1304–1306. doi: 10.1126/science.1220941.

Martin, S. J. and Brettell, L. E. (2019) 'Deformed Wing Virus in Honeybees and Other Insects', *Annu. Rev. Virol.*, 6, pp. x–x. doi: https://doi.org/10.1146/annurev-virology-092818-015700.

Martin, S. J. and Gunn, A. (1999) 'The Transmission of Deformed Wing Virus between Honeybees (Apis mellifera L .) by the Ectoparasitic Mite Varroa jacobsoni Oud', *Journal of Invertebrate Pathology*, 106(1999), pp. 101–106.

Mattila, H. R. and Seeley, T. D. (2007) 'Genetic diversity in honey bee colonies enhances productivity and fitness', *Science*. doi: 10.1126/science.1143046.

McElroy, K. *et al.* (2013) 'Accurate single nucleotide variant detection in viral populations by combining probabilistic clustering with a statistical test of strand bias', *BMC Genomics*, 14(1), p. 501. doi: 10.1186/1471-2164-14-501.

McMahon, D. P. *et al.* (2015) 'A sting in the spit: Widespread cross-infection of multiple RNA viruses across wild and managed bees', *Journal of Animal Ecology*, 84(3), pp. 615–624. doi: 10.1111/1365-2656.12345.

McMahon, D. P. *et al.* (2016) 'Elevated virulence of an emerging viral genotype as a driver of honeybee loss', *Proceedings of the Royal Society B: Biological Sciences*, 283.

Metzner, K. J. *et al.* (2009) 'Minority quasispecies of drug-resistant HIV-1 that lead to early therapy failure in treatment-naive and -adherent patients', *Clinical Infectious Diseases*. doi: 10.1086/595703.

Mikonranta, L. *et al.* (2014) 'Insect immunity: Oral exposure to a bacterial pathogen elicits free radical response and protects from a recurring infection', *Frontiers in Zoology*.

doi: 10.1186/1742-9994-11-23.

Milani, N. and Iob, M. (1998) 'Plastic strips containing organophosphorous acaricides to control Varroa jacobsoni: a preliminary experiment', *American Bee Journal*, 138(8), pp. 612–615.

MJ Gracia-Salinas *et al.* (2006) 'Detection of fluvalinate resistance in Varroa destructor in Spanish apiaries', *Journal of Apicultural Research*, 45(3), pp. 101–105. doi: 10.1080/00218839.2006.11101326.

Möckel, N., Gisder, S. and Genersch, E. (2011) 'Horizontal transmission of deformed wing virus: Pathological consequences in adult bees (Apis mellifera) depend on the transmission route', *Journal of General Virology*, 92(2), pp. 370–377. doi: 10.1099/vir.0.025940-0.

Mondet, F. *et al.* (2014) 'On the Front Line: Quantitative Virus Dynamics in Honeybee (Apis mellifera L.) Colonies along a New Expansion Front of the Parasite Varroa destructor', *PLoS Pathogens*, 10(8). doi: 10.1371/journal.ppat.1004323.

Moore, J. *et al.* (2011) 'Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies', *Journal of General Virology*, 92(1), pp. 156–161. doi: 10.1099/vir.0.025965-0.

Mordecai, G. J., Wilfert, L., *et al.* (2015) 'Diversity in a honey bee pathogen: first report of a third master variant of the Deformed Wing Virus quasispecies', *The ISME Journal*. Nature Publishing Group, pp. 1–10. doi: 10.1038/ismej.2015.178.

Mordecai, G. J., Brettell, L. E., *et al.* (2015) 'Superinfection exclusion and the long-term survival of honey bees in Varroa-infested colonies.', *The ISME journal*. Nature Publishing Group, pp. 1–10. doi: 10.1038/ismej.2015.186.

Moreno, E. *et al.* (2017) 'Internal Disequilibria and Phenotypic Diversification during Replication of Hepatitis C Virus in a Noncoevolving Cellular Environment', *Journal of Virology*. doi: 10.1128/jvi.02505-16.

Morimoto, T. *et al.* (2011) 'The habitat disruption induces immune-suppression and oxidative stress in honey bees', *Ecology and Evolution*. doi: 10.1002/ece3.21.

Murray, A. G. and Salama, N. K. G. (2016) 'A simple model of the role of area management in the control of sea lice', *Ecological Modelling*. doi: 10.1016/j.ecolmodel.2016.06.007.

Natsopoulou, M. E. *et al.* (2017) 'The virulent, emerging genotype B of Deformed wing virus is closely linked to overwinter honeybee worker loss', *Scientific Reports*. Springer US, 7(1), p. 5242. doi: 10.1038/s41598-017-05596-3.

Nazzi, F. *et al.* (2012) 'Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies', *PLoS Pathogens*, 8(6). doi: 10.1371/journal.ppat.1002735.

Nolan, M. P. and Delaplane, K. S. (2016) 'Distance between honey bee Apis mellifera colonies regulates populations of Varroa destructor at a landscape scale', *Apidologie*. Apidologie, (Free 1958). doi: 10.1007/s13592-016-0443-9.

Norton, A. M. *et al.* (2020) 'Accumulation and Competition Amongst Deformed Wing Virus Genotypes in Naïve Australian Honeybees Provides Insight Into the Increasing Global Prevalence of Genotype B', *Frontiers in Microbiology*, 11(April), pp. 1–14. doi: 10.3389/fmicb.2020.00620.

Oberste, M. S. *et al.* (1999) 'Molecular Evolution of the Human Enteroviruses: Correlation of Serotype with VP1 Sequence and Application to Picornavirus Classification', *Journal of Virology*, 73(3), pp. 1941–1948. doi: 10.1128/jvi.73.3.1941-1948.1999.

Oldroyd, B. P. (1999) 'Coevolution while you wait: Varroa jacobsoni, a new parasite of western honeybees', *Trends in Ecology and Evolution*. doi: 10.1016/S0169-5347(99)01613-4.

Ongus, J. R. *et al.* (2004) 'Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite Varroa destructor', *Journal of General Virology*, 85(12), pp. 3747–3755. doi: 10.1099/vir.0.80470-0.

Organtini, L. J. *et al.* (2017) 'Honey Bee Deformed Wing Virus structures reveal that conformational changes accompany genome release', *American Society for Microbiology*, 91(2), pp. 10–13.

Peck, D. T. and Seeley, T. D. (2019) 'Mite bombs or robber lures? The roles of drifting and robbing in Varroa destructor transmission from collapsing honey bee colonies to their neighbors', *PLoS ONE*, 14(6), pp. 1–14. doi: 10.1371/journal.pone.0218392.

Pérez de Leon, A. A. *et al.* (2012) 'Integrated strategy for sustainable cattle fever tick eradication in USA is required to mitigate the impact of global change', *Frontiers in Physiology*. doi: 10.3389/fphys.2012.00195.

Pfeiffer, J. K. and Kirkegaard, K. (2006) 'Bottleneck-mediated quasispecies restriction during spread of an RNA virus from inoculation site to brain', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0600834103.

Pfeiffer, K. J. and Crailsheim, K. (1998) 'Drifting of honeybees', *Insectes Sociaux*, 45(2), pp. 151–167. doi: 10.1007/s000400050076.

Pillai, S. K. *et al.* (2006) 'Genetic attributes of cerebrospinal fluid-derived HIV-1 env', *Brain*. doi: 10.1093/brain/awl136.

Poirier, E. Z. and Vignuzzi, M. (2017) 'Virus population dynamics during infection', *Current Opinion in Virology*. Elsevier B.V., 23, pp. 82–87. doi: 10.1016/j.coviro.2017.03.013.

Posada-Cespedes, S., Seifert, D. and Beerenwinkel, N. (2017) 'Recent advances in inferring viral diversity from high-throughput sequencing data', *Virus Research*. Elsevier B.V., 239, pp. 17–32. doi: 10.1016/j.virusres.2016.09.016.

Posada-Florez, F. *et al.* (2019) 'Deformed wing virus type A, a major honey bee pathogen, is vectored by the mite Varroa destructor in a non-propagative manner', *Scientific Reports*, 9(1), p. 12445. doi: 10.1038/s41598-019-47447-3.

Potts, S. G. *et al.* (2010) 'Global pollinator declines: Trends, impacts and drivers', *Trends in Ecology and Evolution*. doi: 10.1016/j.tree.2010.01.007.

Ramsey, S. D. *et al.* (2018) 'Varroa destructor feeds primarily on honey bee fat body tissue and not hemolymph', *PNAS*. doi: 10.1073/pnas.1818371116.

Remnant, E. J. *et al.* (2019) 'Direct transmission by injection affects competition among RNA viruses in honeybees-', *Proceedings of the Royal Society B: Biological Sciences*, 91, pp. 399–404.

Rezelj, V. V., Levi, L. I. and Vignuzzi, M. (2018) 'The defective component of viral populations', *Current Opinion in Virology*. doi: 10.1016/j.coviro.2018.07.014.

Rinkevich, F. D. (2020) 'Detection of amitraz resistance and reduced treatment efficacy in the Varroa Mite, Varroa destructor, within commercial beekeeping operations', *PLoS ONE*. doi: 10.1371/journal.pone.0227264.

Rinkevich, F. D., Danka, R. G. and Healy, K. B. (2017) 'Influence of varroa mite (Varroa destructor) management practices on insecticide sensitivity in the honey bee (Apis mellifera)', *Insects*. doi: 10.3390/insects8010009.

Roberts, J. M. K., Anderson, D. L. and Durr, P. A. (2017) 'Absence of deformed wing virus and Varroa destructor in Australia provides unique perspectives on honeybee viral landscapes and colony losses', *Scientific Reports*. Springer US, 7(1), p. 6925. doi: 10.1038/s41598-017-07290-w.

Rohayem, J., Munch, J. and Rethwilm, A. (2005) 'Evidence of Recombination in the Norovirus Capsid Gene', *Journal of Virology*. doi: 10.1128/jvi.79.8.4977-4990.2005.

Rosà, R. *et al.* (2007) 'Temporal variation of Ixodes ricinus intensity on the rodent host

Apodemus flavicollis in relation to local climate and host dynamics', *Vector-Borne and Zoonotic Diseases*. doi: 10.1089/vbz.2006.0607.

Rosenkranz, P., Aumeier, P. and Ziegelmann, B. (2010) 'Biology and control of Varroa destructor', *Journal of Invertebrate Pathology*, 103(SUPPL. 1), pp. S96–S119. doi: 10.1016/j.jip.2009.07.016.

Routh, A. and Johnson, J. E. (2014) 'Discovery of functional genomic motifs in viruses with ViReMa-a virus recombination mapper-for analysis of next-generation sequencing data', *Nucleic Acids Research*, 42(2), pp. 1–10. doi: 10.1093/nar/gkt916.

Ryabov, E. V. *et al.* (2017) 'Recent spread of Varroa destructor virus-1, a honey bee pathogen, in the United States', *Scientific Reports*. Springer US, 7(1), p. 17447. doi: 10.1038/s41598-017-17802-3.

Ryabov, E. V *et al.* (2014) 'A virulent strain of deformed wing virus (DWV) of honeybees (Apis mellifera) prevails after Varroa destructor-mediated, or in vitro, transmission.', *PLoS pathogens*, 10(6), p. e1004230. doi: 10.1371/journal.ppat.1004230.

Ryabov, E. V *et al.* (2019) 'Dynamic evolution in the key honey bee pathogen deformed wing virus: Novel insights into virulence and competition using reverse genetics.', *PLoS biology*, 17(10), p. e3000502. doi: 10.1371/journal.pbio.3000502.

Ryabov, E. V *et al.* (2020) 'Development of a honey bee RNA virus vector based on the genome of Deformed wing virus'.

Salas-Benito, J. S. and De Nova-Ocampo, M. (2015) 'Viral interference and persistence in mosquito-borne flaviviruses', *Journal of Immunology Research*. doi: 10.1155/2015/873404.

Schnell, G. *et al.* (2009) 'Compartmentalized Human Immunodeficiency Virus Type 1 Originates from Long-Lived Cells in Some Subjects with HIV-1-Associated Dementia', *PLoS Pathogens*. doi: 10.1371/journal.ppat.1000395.

Seeley, T. D. (2007) 'Honey bees of the Arnot Forest: A population of feral colonies persisting with Varroa destructor in the northeastern United States', *Apidologie*. doi: 10.1051/apido:2006055.

Seeley, T. D. *et al.* (2015) 'A survivor population of wild colonies of European honeybees in the northeastern United States: investigating its genetic structure', *Apidologie*. doi: 10.1007/s13592-015-0355-0.

Seeley, T. D. and Smith, M. L. (2015) 'Crowding honeybee colonies in apiaries can increase their vulnerability to the deadly ectoparasite Varroa destructor', *Apidologie*, 46(6), pp. 716–727. doi: 10.1007/s13592-015-0361-2.

Šekulja, D., Pechhacker, H. and Licek, E. (2014) 'Drifting behaviour of honey bees (Apis mellifera carnica Pollman, 1879) in the epidemiology of American foulbrood', *Zbornik sveučilišta u Rijeci*, 2(1), pp. 345–358.

Simon-Loriere, E. *et al.* (2009) 'Molecular mechanisms of recombination restriction in the envelope gene of the human immunodeficiency virus', *PLoS Pathogens*. doi: 10.1371/journal.ppat.1000418.

Škubník, K. *et al.* (2017) 'Structure of deformed wing virus, a major honey bee pathogen', *Proceedings of the National Academy of Sciences*, 114(12), pp. 3210–3215. doi: 10.1073/pnas.1615695114.

de Souza, F. S. *et al.* (2019) 'Occurrence of deformed wing virus variants in the stingless bee melipona subnitida and honey bee Apis mellifera populations in Brazil', *Journal of General Virology*. doi: 10.1099/jgv.0.001206.

Syromyatnikov, M. Y., Borodachev, A. V and Kokina, A. V (2018) 'A Molecular Method for the Identification of Honey', pp. 1–12. doi: 10.3390/insects9010010.

Tamura, K. and Nei, M. (1993) 'Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees', *Mol Biol Evol*, 10(3), pp. 512–526. doi: 10.1093/molbev/msl149.

Tehel, A. *et al.* (2019) 'The Two Prevalent Genotypes of an Emerging Equally Low Pupal Mortality and Equally High Wing Deformities in Host Honey Bees', *Viruses*, 11(114), pp. 1–18. doi: 10.3390/v11020114.

Tentcheva, D. *et al.* (2004) 'Prevalence and Seasonal Variations of Six Bee Viruses in *Apis mellifera* L . and *Varroa destructor* Mite Populations in France', *Applied and Environmental Microbiology*, 70(12), pp. 7185–7191. doi: 10.1128/AEM.70.12.7185.

Thompson, C. E. *et al.* (2014) 'Parasite pressures on feral honey bees (Apis mellifera sp.)', *PLoS ONE*, 9(8), pp. 1–8. doi: 10.1371/journal.pone.0105164.

Thompson, H. M. *et al.* (2002) 'First report of *Varroa destructor* resistance to pyrethroids in the UK', *Apidologie*, 33(4), pp. 357–366. doi: 10.1051/apido:2002027.

Toufailia, H. M. Al *et al.* (2014) 'Towards integrated control of varroa: Effect of variation in hygienic behaviour among honey bee colonies on mite population increase and deformed wing virus incidence', *Journal of Apicultural Research*, 53(5), pp. 555–562. doi: 10.3896/IBRA.1.53.5.10.

Toufailia, H. M. Al and Ratnieks, F. L. W. (2016) 'How Effective is Apistan at Killing Varroa ?', *Bee Craft*, 98(2), pp. 2013–2016.

Toyoda, H. *et al.* (1984) 'Complete nucleotide sequences of all three poliovirus serotype genomes', *Journal of Molecular Biology*. doi: 10.1016/0022-2836(84)90084-6.

Traynor, K. S. *et al.* (2020) 'Varroa destructor: A Complex Parasite, Crippling Honey Bees Worldwide', *Trends in Parasitology*. The Authors, xx(xx), pp. 1–15. doi: 10.1016/j.pt.2020.04.004.

vanEngelsdorp, D. *et al.* (2009) 'Colony collapse disorder: A descriptive study', *PLoS ONE*. doi: 10.1371/journal.pone.0006481.

Vignuzzi, M. *et al.* (2006) 'Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population', *Nature*. doi: 10.1038/nature04388.

Visscher, P. and Seeley, T. (1989) 'Bee-lining as a research technique in ecological studies of honey bees.', *American Bee Journal*, (October), pp. 536–539. Available at: http://www.cabdirect.org/abstracts/19900228410.html.

Watanabe, M. E. (2008) 'Colony collapse disorder: Many suspects, no smoking gun', *BioScience*. doi: 10.1641/B580503.

Wei, T. *et al.* (2017) 'R package "corrplot": Visualization of a Correlation Matrix', *Statistician*.

Whitfield, C. W. *et al.* (2006) 'Thrice out of Africa: Ancient and recent expansions of the honey bee, Apis mellifera', *Science*. doi: 10.1126/science.1132772.

Wilfert, L. *et al.* (2016) 'Deformed wing virus is a recent global epidemic in honeybees driven by Varroa mites.', *Science (New York, N.Y.)*, 351(6273), pp. 594–7. doi: 10.1126/science.aac9976.

Wilhelm, T. *et al.* (2014) 'Spontaneous slow replication fork progression elicits mitosis alterations in homologous recombination-deficient mammalian cells', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1311520111.

Wood, G. R. *et al.* (2014) 'MosaicSolver: A tool for determining recombinants of viral genomes from pileup data', *Nucleic Acids Research*, 42(16). doi: 10.1093/nar/gku524.

Woodford, L. and Evans, D. J. (2020) ' Deformed wing virus : using reverse genetics to tackle unanswered questions about the most important viral pathogen of honey bees ', *FEMS Microbiology Reviews*. Oxford University Press, (December), pp. 1–20. doi: 10.1093/femsre/fuaa070.

Woodman, A. *et al.* (2016) 'Biochemical and genetic analysis of the role of the viral polymerase in enterovirus recombination', *Nucleic Acids Research*. doi: 10.1093/nar/gkw567.

Xiao, Y. *et al.* (2016) 'RNA Recombination Enhances Adaptability and Is Required for Virus Spread and Virulence', *Cell Host and Microbe*. doi: 10.1016/j.chom.2016.03.009.

Xiao, Y. *et al.* (2017) 'Poliovirus intrahost evolution is required to overcome tissue-specific innate immune responses', *Nature Communications*. doi: 10.1038/s41467-017-00354-5.

Yañez, O. *et al.* (2020) 'The honeybee (Apis mellifera) developmental state shapes the genetic composition of the deformed wing virus-A quasispecies during serial transmission', *Scientific Reports*, pp. 1–12. doi: 10.1038/s41598-020-62673-w.

Yue, C. *et al.* (2007) 'Vertical-transmission routes for deformed wing virus of honeybees (Apis mellifera)', *Journal of General Virology*. doi: 10.1099/vir.0.83101-0.

Yue, C. and Genersch, E. (2005) 'RT-PCR analysis of Deformed wing virus in honeybees (Apis mellifera) and mites (Varroa destructor)', *Journal of General Virology*, 86(12), pp. 3419–3424. doi: 10.1099/vir.0.81401-0.

Zagordi, O. *et al.* (2010) 'Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies', *Nucleic Acids Research*. doi: 10.1093/nar/gkq655.

Zagordi, O. *et al.* (2011) 'ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data', *BMC Bioinformatics*, 12(1), p. 119. doi: 10.1186/1471-2105-12-119.

Zhang, Y. and Han, R. (2018) 'A Saliva Protein of Varroa Mites Contributes to the Toxicity toward Apis cerana and the DWV Elevation in A. mellifera', *Scientific Reports*. Springer US, 8(1), pp. 1–9. doi: 10.1038/s41598-018-21736-9.

Zioni, N., Soroker, V. and Chejanovsky, N. (2011) 'Replication of varroa destructor virus 1 (VDV-1) and a varroa destructor virus 1-deformed wing virus recombinant (VDV-1-DWV) in the head of the honey bee', *Virology*. Elsevier Inc., 417(1), pp. 106–112. doi: 10.1016/j.virol.2011.05.009.