Employing Transposon Mutagenesis to Investigate Foot-and-Mouth Disease Virus Replication

Running title: Transposon mutagenesis of the FMDV genome

Morgan R. Herod¹, Eleni-Anna Loundras¹, Joseph C. Ward¹, Fiona Tulloch², David J. Rowlands¹ and Nicola J. Stonehouse*¹

¹School of Molecular and Cellular Biology, Faculty of Biological Sciences and Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, United Kingdom.
²Biomedical Sciences Research Complex (BSRC), School of Biology, University of St. Andrews, North Haugh, St. Andrews, Fife, KY16 9ST, Scotland.

*N.J.Stonehouse@leeds.ac.uk

Key Words: FMDV, picornavirus, 3A trans-complementation, transposon, mutagenesis

Abstract word count: 197

Word count: 4799

Summary

Probing the molecular interactions within the foot-and-mouth disease virus (FMDV) RNA replication complex has been restricted in part to the lack of suitable reagents. Random insertional mutagenesis has proven an excellent method to reveal domains of proteins essential for viral replication as well as locations that can tolerate small genetic insertions. Such insertion sites can be subsequently adapted by the incorporation of commonly used epitope tags and so facilitate their detection with commercial available reagents. In this study, we use random transposon-mediated mutagenesis to produce a library of 15 nucleotide insertions in the FMDV non-structural polyprotein. Using a replicon-based assay we isolated multiple replication-competent as well as replication-defective insertions. We have adapted the replication-competent insertion sites for the successful incorporation of epitope tags within FMDV non-structural proteins, for the use in a variety of downstream assays. Additionally, we show that replication of some of the replication-defective insertion mutants can be rescued by co-
transfection of a ‘helper’ replicon, demonstrating a novel use of random mutagenesis to identify inter-genomic trans-complementation. Both the epitope tags and replication-defective insertions identified here will be valuable tools for probing interactions within picornaviral replication complexes.

Introduction

Foot-and-mouth disease is an acute systemic disease of cloven-hoofed animals, outbreaks of which in domestic livestock have significant economic consequences for the agricultural and tourism industries. The causative agent, foot-and-mouth disease virus (FMDV), is endemic in wide areas of Asia, Southern America, Africa and the Middle East and has the potential to cause major epidemics globally. Difficulties arising from the control of the spread of disease stem primarily from the high infectivity and transmissibility of the virus and the asymptomatic carrier state the virus can adopt.

FMDV is a member of the *Picornaviridae* family of single-stranded positive sense RNA viruses. The genome is translated as a single open reading frame, flanked by both 5′ and 3′ untranslated regions (UTR) and a 3′ poly(A) tail (Carrillo *et al.*, 2005). The long and highly secondary structured 5′ UTR contains at least 5 discrete domains, including the type II viral IRES (Belsham & Brangwyn, 1990; Lopez de Quinto & Martinez-Salas, 1997; Lopez de Quinto *et al.*, 2002) and *cre* or *cis*-acting replicative elements (Mason *et al.*, 2002) in addition to 3 elements of unknown function; varying copies of pseudoknots, a polypyrimidine (pC) tract of variable length and a predicted large 5′ stem loop or S-fragment (Carrillo *et al.*, 2005; Clarke *et al.*, 1987; Escarmis *et al.*, 1995; Mason *et al.*, 2003; Rowlands *et al.*, 1978). The comparatively smaller 3′ UTR contains two stem-loop structures both of which appear to have roles in viral RNA replication (Rodriguez Pulido *et al.*, 2009; Saiz *et al.*, 2001).

Following translation the FMDV polyprotein is processed co- and post-translationally to produce 4 primary products; mature L\(^{pro}\) (self-processed *in cis* at its own C-terminus), and the precursors P1-2A. 2BC and P3 (3AB\(_{1-3}\)CD). Processing at the 2A2B boundary occurs co-translationally through a
ribosome skipping mechanism to release P1-2A from the rest of the polyprotein (de Felipe et al., 2003; Donnelly et al., 2001; Ryan & Drew, 1994). The P1-2A primary product is subsequently processed by the 3C/3CD protease to generate three structural proteins, 1AB, 1C, and 1D (1AB being cleaved by an unknown mechanism into 1A and 1B in a final virus maturation step), whereas the 2BC and P3 precursors undergo 3C/3CD mediated proteolysis to generate the mature viral RNA replication proteins (reviewed by Ryan & Flint, 1997). In poliovirus and FMDV, non-structural (NS) proteins 2B and 2C along with their precursor 2BC are involved in disrupting endoplasmic reticulum to Golgi transport to inhibit the cellular secretory pathway (Doedens & Kirkegaard, 1995; Moffat et al., 2005; Moffat et al., 2007) via a PI4K-independent mechanism in FMDV (Loundras, Herod, Harris and Stonehouse, manuscript in preparation). The 2C protein from poliovirus also demonstrates ATPase activity and is likely to play a direct role in replicating the viral genome (Rodriguez & Carrasco, 1993; Xia et al., 2015).

The P3 precursor undergoes proteolysis, likely through both major and minor pathways, to generate four mature viral non-structural proteins; the viral RNA-dependant-RNA polymerase 3D<sup>pol</sup> (Ferrer-Orta et al., 2009; Ferrer-Orta et al., 2006; Ferrer-Orta et al., 2004), the major viral protease 3C<sup>pro</sup> (Birtley et al., 2005; Grubman et al., 1995), three non-identical tandem repeats of the primer polypeptide 3B (3B<sub>1-3</sub>) (Forss & Schaller, 1982; King et al., 1980; Nayak et al., 2005; Paul et al., 1998; Paul et al., 2003) and the transmembrane protein 3A (Gonzalez-Magaldi et al., 2014; Gonzalez-Magaldi et al., 2012). Replication of the viral genome requires expression of all viral NS proteins in addition to cis-acting RNA elements, which are thought to localise to membrane-associated replication compartments where viral RNA synthesis occurs. To date, defining the molecular interactions in these replication compartments has largely remained elusive, in part due to the limited reagents available for studying such interactions.

Random transposon-mediated mutagenesis has been extensively exploited for the functional profiling of viral proteins and identifying essential protein functional domains as well as characterising cis-acting RNA replication elements (Brune et al., 1999; McMahon et al., 1998; Mohl et al., 2010;
Remenyi et al., 2014; Teterina et al., 2011a; Thorne et al., 2012). This powerful technique allows for the simultaneous screening of large numbers of insertional mutations on viral replication to identify locations within proteins which can tolerate small genetic insertions. Once identified, these insertion sites can be utilised for the genetic incorporation of epitope tags facilitating further downstream study such as co-immunoprecipitation and immunofluorescence (Teterina et al., 2010; Teterina et al., 2011b; Thorne et al., 2012).

In this study, our aim was to use transposon mutagenesis on the FMDV P2P3 polyprotein to identify locations within NS proteins which could tolerate small genetic insertions. For this work, we employed a FMDV replicon, a self-replicating mini viral genome in which the viral structural proteins have been removed and replaced by a reporter transgene allowing for quantification of viral replication in the absence of virion production (Tulloch et al., 2014). Using this replicon-based reporter assay, we have identified functionally permissive insertion sites in addition to replication-defective insertions. Additionally, we have adapted well tolerated insertion sites by the incorporation of epitope tags. Finally, we demonstrate that replication defective insertion sites within the 3A NS protein can be complemented in trans by co-transfection with a ‘helper’ replicon construct.

Results

Transposon mutagenesis of the FMDV NS polyprotein

Multiple studies have demonstrated the use of transposon-mediated random mutagenesis to probe the genomes of positive-strand RNA viruses for sites that can tolerate the insertion of small exogenous sequences (McMahon et al., 1998; Remenyi et al., 2014; Teterina et al., 2011a; Thorne et al., 2012) and we have used this method to identify regions of the FMDV NS polyprotein which can tolerate genetic insertions. The FMDV replicon plasmid, pGFP-PAC (Tulloch et al., 2014) was used to generate a replicon library containing insertions solely in the NS polyprotein (2A though to 3Dpol). To exclude irrelevant insertions sites within the vector backbone, the XmaI – BamHI fragment of the FMDV NS polyprotein was subcloned into XmaI – BamHI digested pUC18 to generate pUC-2A-3D. Mutagenesis was conducted on pUC-2A-3D prior to replacement of the XmaI - BamHI fragment into
pGFP-PAC and the transposon removed from the library by digestion with NotI. This resulted in a library of 15 nucleotide insertions spanning the FMDV 2A-3D region. Each 15 nucleotide insertion consisted of 10 nucleotides remaining from the transposon insertion (including a unique NotI restriction enzyme site) and 5 nucleotides originating from the target DNA directly upstream of the insertion site which is duplicated during the transposition event.

To identify insertion sites in vitro, transcribed RNA from the transposon-mutated replicon library was transfected into BHK-21 cells along with appropriate controls. Transfection was performed with either 0.3 µg or 1 µg of in vitro-transcribed RNA per well. Total RNA was extracted at 8 hours post-transfection when GFP expression was maximal (Fig. S1), replicon genomes amplified by RT-PCR (Fig. S2) and subcloned into a plasmid vector. To identify a limited selection of different transposon insertions, a total of 38 individual clones were selected at random and the location of transposon insertion determined by DNA sequencing. The name of each transposon insertion derived from the nucleotide of the corresponding NS protein after which insertion occurred (Tables 1 and 2).

Of the 16 clones isolated from transfection of 0.3 µg of RNA, 11 insertions were located in the C-terminal region of 3A, downstream of the predicted transmembrane region. Two insertion sites were within the multiple copies of 3B (one within 3B₁ and one within 3B₂), with two insertions in 2C and one within the 3D⁺pol coding region. Similarly, of the 22 clones isolated after transfection of 1 µg of the transposon library, 7 insertions were in 3A, with one insertion in 3B₁ and 3B₃, respectively. In addition, using this higher RNA concentration, 4 and 9 insertions were identified in 3C⁺pro and 3D⁺pol, respectively. Only two common insertions were isolated after transfection with either RNA concentration, all within the 3A C-terminal region (3A229 and 3A339).

To broadly screen a selection of the identified transposon insertions for replication competence, eleven 3A transposon insertions, one insertion in each of 3B₁₃ and all the isolated 2C, 3C⁺pro and 3D⁺pol insertions were introduced individually into the pGFP-PAC replicon. RNA transcripts were generated and transfected into BHK-21 cells along with a wild-type positive control, and a replication-defective
negative control construct bearing a large deletion to the polymerase gene (Δ3Dpol), and replication monitored by GFP expression hourly over a 24 hour period using an IncuCyte Dual Colour Zoom® FLR. Relative replication is shown at maximal GFP expression at 8 hours post-transfection (Fig. 1).

All except one of the replicon constructs containing transposon insertions in 3A demonstrated replication above that of the replication-defective polymerase knockout control (indicating the level of input translation), the exception being the insertion at nucleotide 3A341 which was replication-incompetent. All 3B transposon insertions tested also demonstrated replication almost equivalent to the wild-type construct. The single 3Dpol insertion identified after transfection of 0.3 µg of RNA from the transposon mutated replicon library (3D450) replicated to wild-type levels, whereas none of the 3Cpro or 3Dpol insertions identified after transfection of 1 µg of library RNA were replication-competent, and only expressed GFP equivalent to the polymerase knockout control construct (Δ3Dpol).

Neither of the isolated 2C transposon insertions showed replication in BHK-21 cells. Notably, a limited number of the replicon constructs containing transposon insertions demonstrated GFP expression below that of the negative control, polymerase knockout replicon, in particular insertions 3A341 and 3D747 as well as possibly both 2C insertions.

Modifications of the C-terminal portion of FMDV 3A have been shown to limit replication of the virus in bovine cells (Beard & Mason, 2000; Li et al., 2010; Pacheco et al., 2013; Pacheco et al., 2003). To investigate whether the identified replication-competent transposon insertions could replicate in bovine cells, replicons bearing replication-competent insertions were transfected into MDBK cells along with controls and GFP expression monitored over 24 hours (Fig. 2).

The majority of 3A insertions completely abrogated or severely impaired replication in MDBK cells, with exception of the 3A339, 3A358 and 3A387 insertions which maintained over 50% replication. The identified transposon insertions in 3B1 and 3B3 were essentially replication-incompetent in MDBK cells, whereas the 3B2 insertion tested maintained replication equal to the wild-type replicon.
The replication-competent 3D\textsuperscript{pol} insertion (3D450) maintained good replication in MDBK cells, with GFP expression approximately 50% of that of the wild-type replicon.

**Generation of tagged FMDV replicons**

Identification of replication-competent insertion sites suggests potential locations for insertion of alternative exogenous sequences such as epitope tags. Two frequently used small epitope tags, FLAG (DYKDDDDK) and HA (YPYDVPDYA), were chosen for insertion into two different location in 3A (after nucleotides 303 and 358) and the one functional 3D\textsuperscript{pol} insertion site (nucleotide 450) (Fig. 3a), all three insertion sites demonstrating high levels of replication in BHK-21 cells. Insertion site 3A303 was selected for epitope labelling, based on the inserted transposon sequence at this location (DCGRTDDDK) which partially resembled a FLAG epitope and 3A358 was selected since this demonstrated moderate replication in MDBK cells. Replication of the epitope tagged replicon constructs was assessed in both BHK-21 and MDBK cells along with the relevant controls (Fig. 3b and 3c).

Replicons bearing either a FLAG or HA tag in either of the 3A insertion sites tested showed levels of replication equivalent to the wild-type replicon in BHK-21 cells, but little or no replication in the bovine cell line, MDBK. In contrast, insertion of either epitope tag in the 3D450 insertion site completely abrogated replication, even in BHK-21 cells.

Western blot analysis of BHK-21 cells transfected with FLAG and HA labelled 3A replicons with anti-FLAG and anti-HA primary antibodies detected epitope labelled 3A and 3A precursors as expected (Fig. 3d). Probing with an anti-3A monoclonal antibody also detected epitope-labelled 3A303, but failed to detect either 3A358FLAG or 3A358HA, possibly because genetic insertions in this position disrupt the epitope recognised by this monoclonal antibody or insertion of epitope tags disrupts the native folding of 3A when introduced into this position.
To demonstrate that the epitope tagged replicons could be used for immunofluorescent detection of
3A, BHK-21 cells were transfected with the 3A303FLAG or 3A358FLAG replicon, fixed at 8 hours
post-transfection, probed with an anti-FLAG antibody and analysed by confocal microscopy (Fig. 3e).
As would be anticipated, FLAG staining was clearly detected only in cells expressing the GFP
transgene, used as a marker for replicon replication, and demonstrated a diffuse punctate staining
concordant to that previously described for FMDV 3A (Garcia-Briones et al., 2006; Gonzalez-Magaldi et al., 2014; Gonzalez-Magaldi et al., 2012; O'Donnell et al., 2001).

Replication-defective 3A mutations can be complemented in trans
Previous studies with FMDV and other picornaviruses have demonstrated that certain non-structural
protein functions can be rescued in trans by co-expression with a replication competent helper virus
or replicon (Garcia-Arriaza et al., 2005; Giachetti et al., 1992; Teterina et al., 1995; Tiley et al., 2003;
Towner et al., 1998). All of the 3A transposon insertions identified in this study were located within
the 3A C-terminus, downstream of the predicted transmembrane domain, and the majority were
replication-competent in BHK-21 cells. The one exception was the single replication-defective 3A
transposon insertion site identified, 3A341, which was isolated after transfection with the higher
amount of the mutant replicon library, along with multiple other replication-defective 3Cpro and 3Dpol
insertions. We hypothesised that selection using this higher concentration of transposon library was
fostering an environment favourable for replication-defective insertions to be replicated in trans due
to co-transfection with replication competent genomes in the library.

To investigate this possibility the replication-defective 3A341 transposon insertion was introduced
into a replicon construct in which the GFP-PAC reporter cassette had been replaced by mCherry red
fluorescent protein (Tulloch, Luke, Nicholson and Ryan, manuscript in preparation). In addition, two
negative control replicons were generated containing either a double point mutation to the 3Dpol active
site GDD motif (3DpolGNN) or the same 3Dpol deletion as used previously (Δ3Dpol). Equivalent
‘helper’ replicon constructs were generated containing the GFP reporter gene from Ptilosarcus
(ptGFP) in place of mCherry (Tulloch, Luke, Nicholson and Ryan, manuscript in preparation),
allowing for discrimination of replication between the replication-defective mCherry construct and the ‘helper’ replicon. We theorised that the use of ‘helper’ replicons would allow for expression of wild-type 3A from the native precursor(s) with the level and timing of protein expression balanced to that of the mutant replicon.

The ptGFP replicons were subsequently assessed for their ability to support replication of the mCherry replicons bearing replication-defective mutations (Fig. 4). For the replicon bearing the replication-defective mutation 3A341, trans-complementation was observed when this was co-transfected with the wild-type ptGFP ‘helper’ construct, with an approximate 2-fold significant increase in mCherry expression. Neither of the replication-defective polymerase constructs (ptGFP-3DpolGNN or ptGFP-Δ3Dpol) were able to rescue the 3A mutant. However the 3Dpol active site point mutant construct could itself be efficiently rescued in trans by co-transfection with a wild-type helper replicon. Interestingly, however, the mutant construct bearing a replication defective 3Dpol deletion (mCherry-Δ3Dpol) was not rescued by any of the ‘helper’ replicons, with only a small but non-significant decrease in mCherry observed upon co-transfection with the wild-type ptGFP replicon. Thus this data would suggest that whereas both 3A and 3Dpol can be complemented in trans, not all functions of 3Dpol can be rescued by co-transfection indicating some cis preferential functions of certain NS proteins.

Discussion

Understanding the replication of positive strand RNA viruses is key to the development of novel therapeutic strategies. Despite over 50 years of intense study, relatively little is known about the molecular interactions within the picornavirus RNA replication complex and the functions of some viral proteins have not yet been fully elucidated.

Here, we have conducted transposon insertional mutagenesis of the FMDV non-structural polyprotein to find some of the locations within individual NS proteins which could accept insertions of epitope
tags whilst maintaining replicon replication. Selection from the transposon-mutated replicon pool was conducted at both low and 3x higher concentrations of replicon library RNA. At both concentrations, replication-competent insertions sites were readily identified in 3A, with 16 separate 3A insertions identified across the two concentrations. Notably, all insertion sites were located within the C-terminal unstructured region, with the most N-terminal insertion occurring immediately following the terminal residue of the predicted transmembrane domain (Fig. S3). It is remarkable that for a protein of 459 nucleotides, no insertions were identified in the first 229 nucleotides, strongly implicating the N-terminal half of the protein as essential for viral RNA replication. Furthermore, transposon mutagenesis of the poliovirus genome by Teterina et al, only isolated functional 3A insertion sites within the first 11 amino acid residues of the protein and not within the N-terminal alpha helical and hydrophobic domains, further highlighting the importance of these structures in picornaviral replication. The N-terminal portion of FMDV 3A is predicted to contain two alpha helices involved in 3A homo-dimerisation followed by a hydrophobic transmembrane domain spanning residues ~59 – 76 (Gonzalez-Magaldi et al., 2012). In comparison to the N-terminal region, the C-terminal portion of FMDV 3A is relatively non-conserved and is extended by some 60 amino acids compared to most other picornaviruses. Mutations, insertions and deletions within the C-terminal region of 3A have implicated its importance for pathogenicity and host range with natural viral isolates with large deletions in this region having been identified (such as residues 85-102, 93-102 and 133-143) (Knowles et al., 2001; O'Donnell et al., 2001; Pacheco et al., 2013; Pacheco et al., 2003). Deletion of amino acids 93-102 of 3A has been observed in natural isolates of FMDV and this correlates with an attenuated phenotype in bovines, but normal porcine pathogenicity and replication both in vitro and in vivo (Beard & Mason, 2000; Knowles et al., 2001; Li et al., 2010; 2011). However, it is not clear whether the 3A deletion alone is responsible for this phenotype and the molecular basis for the host specific restriction of replication is unknown. It has been demonstrated by immunofluorescent and FRAP studies that deletions within either the N- or C-terminal regions of 3A can increase protein mobility and alter the cellular distribution in the absence of viral replication, an observation potentially related to the interaction of 3A with the cellular protein DCTN3, which has been implicated in the pathogenicity phenotype in cattle (Gladue et al., 2014; Gonzalez-Magaldi et al.,
Host-cell restriction of 3A is of particular significance, since some of the 3A insertions tested allow for replication in BHK-21 cells, while restricting the replication in the MDBK bovine cell line. Further investigation is warranted into the 3A-mediated host cell phenotype, in particular the C-terminal unstructured half of the protein and particular care must be taken when choosing the cellular context for understanding the molecular basis of viral replication.

Transposon insertions were identified within all three copies of 3B. However previous studies have suggested the possibility of redundancy within the repeated 3B proteins and it is not known at present whether 3Bs bearing insertions retain the function of the protein (Arias et al., 2010; Falk et al., 1992). A tagged poliovirus has been generated containing an 8 amino acid HA tag after residue 17 of 3B which displayed wild-type growth properties, suggesting some of the 3B insertion sites identified here may retain protein function. Interestingly however the insertions tested in both 3B1 and 3B3 abolish replication in the bovine cell line MDBK, whilst maintaining good replication in BHK-21 cells. In contrast, the insertion in 3B2 demonstrated good replication in both cell lines, potentially suggesting a role of the various 3B proteins in regulating host-cell restriction.

One striking observation was the difference in the numbers of 3Cpro and 3Dpol insertions identified using low vs high concentration of mutated library for selection. At the lower library concentration no insertional sites were identified within 3Cpro and a single transposon insertion was identified within 3Dpol (3D450), which replicated efficiently in both BHK-21 and MDBK cells. This insertion is located in an unstructured region at the end of α5 on the outside of the finger domain (Fig. S4), which despite the conservation between picornavirus 3Dpol polymerases, was not identified in the transposon mutagenesis study of poliovirus (Teterina et al., 2011a). In contrast, selection at the higher RNA concentration readily yielded 3Cpro and 3Dpol insertions, all of which were found to completely abrogate replicon replication in isolation, as did the only replication-defective 3A insertion site identified (3A341). However, it must be noted that the observed preference for 3Cpro and 3Dpol insertions (at the higher concentration of library RNA) may be an unintentional consequence of the
non-exhaustive screening approach used in selecting a limited number of colonies at each RNA concentration from a relative large transposon library.

The frequencies at which replication-defective insertions were identified using a high RNA concentration led us to hypothesise that transfection of greater amounts of replicon RNA provided conditions in which replication-defective genomes were being maintained or replicated in trans, by co- incidental co-transfection with genomes containing replication-permissive insertions. In accordance with this hypothesis, replication of a construct bearing the lethal replication defective 3A insertion was rescued by simultaneous co-transfection of a wild-type ‘helper’ replicon. Trans-complementation of replication-defective NS protein mutations has been described within some NS proteins of picornaviruses including poliovirus 3A (Giachetti et al., 1992; Teterina et al., 1995; Towner et al., 1998) and some FMDV proteins (Garcia-Arriaza et al., 2005; Tiley et al., 2003), however to our knowledge, this is the first study to demonstrate rescue in trans of a replication-defective FMDV 3A mutation.

Due to the limited structural information on the C-terminal domain of 3A it is hard to speculate why the 3A341 insertion renders the replicon replication-defective. The observation that this replication-defective lesion can be complemented in trans and that insertions at positions 3A339 and 3A342 were tolerated in both hamster and bovine cells suggests a disruption in 3A protein function, as opposed to effects at the RNA level. Further characterisation of the 3A341 insertion may be valuable in yielding information as to the function of the FMDV 3A C-terminal region.

Having identified functional transposon insertion sites, we generated replicon constructs containing epitope tags in either 3A or 3Dpol. FLAG or HA tags were successfully inserted into 2 separate locations of 3A to yield replication-competent replicons that could be characterised by Western blotting and immunofluorescence analysis. Incorporation of either tag into the 3A358 position abolished recognition by the anti-3A 2C2 mono-clonal antibody, possibly due to disruption of the monoclonal antibody epitope or disruption of the native 3A C-terminal folding. Incorporation of either tag into the 3A303 position allowed for maintained recognition for the antibody, however
resulted in slight changes of 3A mobility by SDS-PAGE, particularly in the case of the FLAG epitope, possibly due to the change in the local charge environment in this unstructured region. Insertions of functional epitope tags have previously been reported in the N-terminal region of poliovirus 3A, albeit with some deleterious effects on replication, and within the C-terminal region of FMDV 3A at similar locations as described here (Li et al., 2012; Ma et al., 2015; Teterina et al., 2011b). However, despite the success of epitope tagging 3A in this study, 3D<sup>pol</sup> did not tolerate the insertion of epitope tags at the single replication-competent transposon insertion site identified, presumably due to the nature of the sequence of the epitope insertion at this location.

In comparison to P3, few insertions were identified within P2, either in this study which only identified 2 replication-defective insertions in 2C, or in the previous transposon mutagenesis study of poliovirus where only one replication-competent insertion was identified at the N-terminus of 2B (Teterina et al., 2011a). Together these data would suggest P2 is relatively less amenable for mutation or modification when compared to P3 and a more focused study using transposon mutation of P2 alone may be required to discover replication-competent insertions in FMDV 2B or 2C non-structural proteins. However, the methodologies used to identify insertions in both this and the previous study by Teterina et al were non-exhaustive and it is therefore possible that tolerated P2 insertions could be identified using alternative methodologies, such as next generation sequencing.

In conclusion, we have used random transposon-mediated mutagenesis to identify replication tolerant insertion sites within the P3 region of the FMDV NS polyprotein and have exploited these sites for the incorporation of epitope tags which will be invaluable for downstream studies. Furthermore, selection using high concentrations of mutagenised replicon RNA enabled the identification of replication-defective insertions which could be rescued in trans. Further investigation of such replication-defective mutations is ongoing and may yield insights into the mechanisms of picornaviral RNA replication.

**Materials and Methods**
Cells lines

BHK-21 and MDBK cells were obtained from the ATCC (LGC Standard) and maintained in Dulbecco modified Eagle Medium with glutamine (Sigma-Aldrich) supplemented with 10% foetal calf serum (FCS), 50 U/ml penicillin and 50 µg/ml streptomycin.

Plasmid constructs

The FMDV replicon plasmid constructs pGFP-PAC and pGFP-PAC-Δ3Dpol polymerase knockout control have been previously described (Tulloch et al., 2014).

Generation of the transposon-mutated replicon library first required transfer of the FMDV non-structural polyprotein coding region into a sub-cloning vector for mutagenesis. Therefore, the XmaI - BamHI fragment from pGFP-PAC was transferred to XmaI - BamHI digested pUC18 (Invitrogen) to regenerate pUC-2A-3D. The transposon mutagenesis system Mutation Generation System Kit (Life technologies) was employed following manufacturer’s instruction, for transposition of a chloramphenicol resistant transposon, on construct pUC-2A-3D to generate the plasmid library pUC-2A-3D-TnC, in which each plasmid contained on average a single chloramphenicol-resistant transposon insertion. Mutagenised clones were transformed in ElecoTen Ultracompetent cells (Stratagene), selected for resistance to chloramphenicol plus kanamycin and total colonies collected. The library was estimated to contain over 20,000 clones. This library was subsequently digested with XmaI and BamHI and the resulting approximately 4.7 kb fragment cloned back into pGFP-PAC and selected against chloramphenicol and ampicillin to remove wild-type replicon and so create the replicon library pGFP-PAC-TnC. The library pGFP-PAC-TnC was digested with NotI to remove the chloramphenicol resistance cassette, and religated to make pGFP-PAC-Tn, a replicon library containing 15 nucleotide insertions randomly located across the FMDV NS polyprotein coding region. To introduce individual transposon insertions into the replicon plasmid, the XmaI - BamHI fragment from pGFP-PAC were replaced by an equivalent XmaI - BamHI fragment obtained from the cloned products derived from the initial transposon selection experiment.

In vitro transcription
Replicon plasmid DNA (5 µg) was linearised with *Hpa*I (NEB) or *Asc*I (NEB), as appropriate, purified by phenol-chloroform extraction, ethanol precipitated and redissolved in RNase-free water. The linear DNA was used in a 50 µl *in vitro* transcription reaction containing transcription buffer and BSA, treated with RNAsecure reagent (Ambion) following manufacturers recommendation, before the addition of 40 units T7 polymerase (NEB), 50 units RNaseOut (Invitrogen) and 8 mM rNTPs (Roche). The *in vitro* transcription reaction was incubated at 32°C for 4 hours after which 2.5 units of RQ1 DNase (Promega) were added, followed by incubation at 37°C for 30 minutes before the RNA was recovered with RNA Clean & Concentrator-25 spin columns (Zymo Research) following manufacturer’s instructions. Transcript integrity was assessed by MOPS-formaldehyde gel electrophoresis prior to transfection.

**Cell transfection and fluorescent reporter assays**

BHK-21 and MDBK cells were seeded into tissue culture plates at 5x10⁴ cells/cm² and 6.25x10⁴ cells/cm², respectively and allowed to adhere for 16 hours. Immediately prior to transfection cells were washed briefly in PBS and media replace with 100 µl/cm² of Minimal Essential Medium (Invitrogen) supplemented with 10 % FCS, 50 U/ml penicillin and 50 µg/ml streptomycin, 1 X non-essential amino acids, and 2 mM glutamine. Duplicate wells were transfected with replicon transcripts using Escort I transfection reagent (Sigma) or Lipofectin (Life technologies) as indicated, following manufacturer’s instructions, using 0.25 µg/cm² or 0.5 µg/cm² total RNA, respectively. For co-transfections, equal amounts of the two RNA transcripts were transfected simultaneously.

Fluorescent protein expression and live cell imaging was analysed using an IncuCyte Dual Colour Zoom® FLR (Essen BioScience) within a 37°C humidified CO₂ incubator scanning hourly up to 24-hours post-transfection collecting multiple images per well. Images were analysed using the associated Zoom® software with the integrated algorithm measuring fluorescent object counts per well as previously described (Forrest *et al.*, 2014; Tulloch *et al.*, 2014). Data is presented to show GFP expression at 8 hours post-transfection as a measure of maximum replication.

**Isolation of transposon insertions**
Following BHK-21 cell transfection of *in vitro* transcribed replicon RNA or replicon library RNA, cells were detached by trypsin and washed once in ice cold PBS. Total RNA was extracted from cell pellet using TRIzol® reagent (Life Technologies) following manufacturer’s protocol. Total RNA was treated with RQ1-DNase (Promega) and FMDV cDNA amplified using Superscript® II (Life Technologies) following manufacturer’s protocol. FMDV genomes were amplified using Phusion® High-Fidelity DNA polymerase (NEB) and blunt-end ligated into pCRBlunt. Individual colonies were isolated and the location of transposon insertions identified by DNA sequencing.

**Western blotting**

Immunoblotting was carried out as previously described (Forrest *et al.*, 2014). Briefly, cells were washed in PBS, detached by trypsin, washed in PBS before lysis in radioimmunoprecipitation assay buffer (0.1 % sodium dodecyl sulfate [SDS], 0.5 % sodium deoxycholate, 1 % NP-40, 150 mM sodium chloride, 50 mM Tris pH 8.0, 1 mM EDTA) supplemented with 2X cOmplete® protease inhibitor (Roche) and incubated on ice for 5 minutes before clarification by centrifugation. Cell lysates were separated by SDS-PAGE using miniProtean gel system (Biorad), followed by transfer to PVDF membrane (Bio-Rad) using XCell SureLock® Mini-Cell wet transfer apparatus (Life technologies). Membranes were blocked with 10 % dried milk, 0.1 % Tween-20 (Sigma-Aldrich) in Tris-buffered saline. Primary antibodies used were rabbit anti-3D 397 polyclonal, mouse anti-3A 2C2 monoclonal (Prof. Francisco Sobrino, Centro De Biologia Molecular Severo Ochoa, Madrid, Spain), mouse anti-FLAG M2 (Sigma-Aldrich) and mouse anti-HA (Sigma-Aldrich) and detected with anti-mouse-HRP or anti-rabbit-HRP (Sigma-Aldrich), as appropriate.

**Confocal microscopy**

BHK-21 cells seeded onto glass coverslips were transfected with *in vitro* transcripts, fixed at indicated time points with 4 % paraformaldehyde, washed in PBS and permeabilised in saponin buffer (0.1 % saponin, 10 % FCS, 0.1 % sodium azide) for 1 hour at 4°C. Primary and secondary antibodies were incubated in saponin buffer for 2 hours at room temperature with three washes in saponin buffer between steps. Primary antibody anti-FLAG M2 (Sigma-Aldrich) was detected with anti-mouse-
Alexa568 (Life technologies) secondary antibodies. Following a final wash in PBS, coverslips were mounted in VECTASHIELD mounting medium with DAPI (Vectorlabs) and images captured using a Zeiss LSM-700 confocal microscope.

Acknowledgement

We would like to thank Professor Sobrino at the Centro De Biologia Molecular in Madrid for the generous gift of antibodies, Dr Martin Stacey plus colleagues at the University of Leeds and Dr Christopher McCormick at the University of Southampton for helpful comments on this manuscript. This work was funded by the BBSRC (Grant BB/K003801/1).
References


### Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Library Isolation</th>
<th>Insertion Site</th>
<th>Nucleotide insertion</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2C17</td>
<td>0.3 µg p1</td>
<td>CATCAA</td>
<td>TCAACTGC GG GCC GCATC</td>
<td>KARDINCGRINDIFA</td>
</tr>
<tr>
<td>2C842</td>
<td>0.3 µg p1</td>
<td>CAAGAC</td>
<td>AGACTGC GG GCC GAAGAG</td>
<td>KRMQDDCGRKDMFKP</td>
</tr>
<tr>
<td>3A229</td>
<td>0.3 µg p1</td>
<td>CCAGGC</td>
<td>CGTTGGCG GCC GC CATCCG</td>
<td>IVIMIRCGRIRETRK</td>
</tr>
<tr>
<td>3A233</td>
<td>0.3 µg p1</td>
<td>TGAGAC</td>
<td>ATGCCGGCC GCACGTGAG</td>
<td>VIMIDAAARETRKR</td>
</tr>
<tr>
<td>3A266</td>
<td>0.3 µg p1</td>
<td>TGAGCT</td>
<td>CAGTGGCG GCC GCATGC</td>
<td>MVDDAVRPHAVNEY</td>
</tr>
<tr>
<td>3A274</td>
<td>0.3 µg p1</td>
<td>ATGACT</td>
<td>GACTGCG GCC CAT GAG</td>
<td>DAVNELRPHEYIEKA</td>
</tr>
<tr>
<td>3A282</td>
<td>0.3 µg p1</td>
<td>TTAGAG</td>
<td>GATGCCGG GCC GA ATTGA</td>
<td>VNEYIDAAAIEKANI</td>
</tr>
<tr>
<td>3A339</td>
<td>0.3 µg p1</td>
<td>CCTCTA</td>
<td>TCTAT GC GCC GC CACTC</td>
<td>AEKSPLCGRRTLTSIG</td>
</tr>
<tr>
<td>3A342</td>
<td>0.3 µg p1</td>
<td>TAGAGA</td>
<td>GAGATGC GG GCC GAAGA</td>
<td>EKSPLEMRPQETSIG</td>
</tr>
<tr>
<td>3A350</td>
<td>0.3 µg p1</td>
<td>AGCCGC</td>
<td>CGCGCTGC GCC GCAGCG</td>
<td>SPLETSGCGRSGAST</td>
</tr>
<tr>
<td>3A354</td>
<td>0.3 µg p1</td>
<td>GCCGCA</td>
<td>GCCAGCTGC GCC GCACC</td>
<td>ETSAGSCGRST SVGF</td>
</tr>
<tr>
<td>3A387</td>
<td>0.3 µg p1</td>
<td>TCAGGC</td>
<td>CCAATGC GC GG GCCGACC</td>
<td>RERTLPVPHPGQKA</td>
</tr>
<tr>
<td>3A442</td>
<td>0.3 µg p1</td>
<td>AGGAGC</td>
<td>GTGCCG GCC CATGAGGA</td>
<td>QPVEVRPHEEQPO</td>
</tr>
<tr>
<td>3B7</td>
<td>0.3 µg p1</td>
<td>CCTAGC</td>
<td>TACCGTGC GCC GCC GTA</td>
<td>EGPYAAAYAGPLE</td>
</tr>
<tr>
<td>3B25</td>
<td>0.3 µg p1</td>
<td>GAGACA</td>
<td>GACTGCCGCC GCAGAGA</td>
<td>AGPMERLPQORQKL</td>
</tr>
<tr>
<td>3D450</td>
<td>0.3 µg p1</td>
<td>ATGGAG</td>
<td>GATGCCGG GCC GAAATGG</td>
<td>ALKLMDDAAMEKREY</td>
</tr>
</tbody>
</table>

**Table 1.** Transposon insertions identified after selection following transfection with 0.3 µg of replicon library RNA. The location of each insertion identified is named with the number indicating the nucleotide residues of the corresponding FMDV NS protein after which insertion occurred. The dinucleotide at which insertion occurred is shown in bold. The inserted nucleotide sequence and amino acid translation is underlined.
<table>
<thead>
<tr>
<th>Name</th>
<th>Library Isolation</th>
<th>Insertion Site</th>
<th>Nucleotide Insertion</th>
<th>Amino Acid Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3A229</td>
<td>1 µg p1</td>
<td>TCCTGA</td>
<td>CTTGCGGCCGCATCAG</td>
<td>IVIMIRCGRIRETRK</td>
</tr>
<tr>
<td>3A234</td>
<td>1 µg p1</td>
<td>TGAGAG</td>
<td>AGTGCAGGCCGCAGTGA</td>
<td>IMIRECFRSETRKRM</td>
</tr>
<tr>
<td>3A301</td>
<td>1 µg p1</td>
<td>CACAGA</td>
<td>CAGATTGCAGGCCGCACAG</td>
<td>ANITTDCEGRDSDKT</td>
</tr>
<tr>
<td>3A303</td>
<td>1 µg p1</td>
<td>CCTCTA</td>
<td>TCTATGGCGGCGCGCACTC</td>
<td>AEKSLCGRLETSG</td>
</tr>
<tr>
<td>3A339</td>
<td>1 µg p1</td>
<td>CTCTAG</td>
<td>CTAGTGCGGGCCGCACTCA</td>
<td>AEKSLVRPHLETSG</td>
</tr>
<tr>
<td>3A358</td>
<td>1 µg p1</td>
<td>CCAGCA</td>
<td>AGTGCAGGCCGCAAGAG</td>
<td>TGSASAAASTVGFH</td>
</tr>
<tr>
<td>3B11</td>
<td>1 µg p1</td>
<td>CGCCGG</td>
<td>CGGATGCAGGCCGCAAGAG</td>
<td>EGPYAGCGRTPLELR</td>
</tr>
<tr>
<td>3B11</td>
<td>1 µg p1</td>
<td>GAGGGA</td>
<td>GGTGCAGGCCGCAAGAG</td>
<td>EGPYEGAAAEGPVKK</td>
</tr>
<tr>
<td>3C1</td>
<td>1 µg p1</td>
<td>TGAGAG</td>
<td>AGTGCAGGCCGCACTGA</td>
<td>KNLIVTECGRTESGA</td>
</tr>
<tr>
<td>3C57</td>
<td>1 µg p1</td>
<td>GCTGAG</td>
<td>TGATGCAGGCCGCAATTG</td>
<td>NTKPVECGRIELLD</td>
</tr>
<tr>
<td>3C196</td>
<td>1 µg p1</td>
<td>AGTGAC</td>
<td>TGATGCAGGCCGCAATG</td>
<td>AMTDSDCGRSDYVF</td>
</tr>
<tr>
<td>3C196</td>
<td>1 µg p1</td>
<td>ACAGCA</td>
<td>AGCTGCAGGCCGCAACAG</td>
<td>KHFDTAAAATARMKK</td>
</tr>
<tr>
<td>3D306</td>
<td>1 µg p1</td>
<td>ATCAAG</td>
<td>CAAGTGCGGGCCGCACTCA</td>
<td>IYEAIKCGHIKGVDG</td>
</tr>
<tr>
<td>3D489</td>
<td>1 µg p1</td>
<td>CTGAAG</td>
<td>GAATGCAGGCCGCACTGA</td>
<td>COTFLNAALKDEIR</td>
</tr>
<tr>
<td>3D525</td>
<td>1 µg p1</td>
<td>CCGGTA</td>
<td>GGTGCAGGCCGCAAGAGG</td>
<td>EKVRAGAAAAGKTRI</td>
</tr>
<tr>
<td>3D551</td>
<td>1 µg p1</td>
<td>TGCAAC</td>
<td>CACTGCAGGCCGCAAGCA</td>
<td>SAVGCNCGRSNOVD</td>
</tr>
<tr>
<td>3D557</td>
<td>1 µg p1</td>
<td>CCTGCA</td>
<td>CTATGCAGGCCGCACTA</td>
<td>VGNPDCRGRTDVWQ</td>
</tr>
<tr>
<td>3D716</td>
<td>1 µg p1</td>
<td>GACTA</td>
<td>ATGCAGGCCGCAATGCA</td>
<td>WDVDAAAVDYSADF</td>
</tr>
<tr>
<td>3D738</td>
<td>1 µg p1</td>
<td>GTTAAT</td>
<td>TATGCAGGCCGCACTA</td>
<td>SAFDANAAAHNCSD</td>
</tr>
<tr>
<td>3D747</td>
<td>1 µg p1</td>
<td>TAGTGA</td>
<td>GTGCGGCCGCACTA</td>
<td>DNHCSGCRSSDAMN</td>
</tr>
<tr>
<td>3D750</td>
<td>1 µg p1</td>
<td>TAGTGA</td>
<td>GTGCGGCCGCACTA</td>
<td>ANHCSGCRSSDAMN</td>
</tr>
</tbody>
</table>

Table 2. Transposon insertions identified after selection following transfection with 1 µg of replicon library. The location of each insertion identified is named with the number indicating the nucleotide residues of the corresponding FMDV non-structural protein after which insertion occurred. The dinucleotide at which insertion occurred is shown in bold. The inserted nucleotide sequence and amino acid translation is underlined.
**Figure Legends**

**Figure 1. Replication of individual transposon insertions in BHK-21 cells.** (a) Cartoon of the FMDV replicon genome showing the location of transposon insertions chosen for analysis. Number indicating the nucleotide residues of the corresponding FMDV non-structural protein after which insertion occurred. (b) BHK-21 cells seeded into 12-well plates were allowed to adhere for 16 hours before transfection with replicon transcripts containing individual transposon insertions using Escort reagent. Wild-type GFP-PAC replicon (wt) and polymerase knockout (Δ3Dpol) constructs were included, the latter as a negative control for input translation. GFP expression was monitored hourly using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents mean GFP positive cells per well at 8 hours post-transfection (n = 2, ± SEM).

**Figure 2 Replication of individual transposon insertions in MDBK cells.** MDBK cells seeded into 12-well plates were allowed to adhere for 16 hours before transfection with replicon transcripts containing individual transposon insertions after indicated nucleotide using Escort reagent. Wild-type GFP-PAC replicon (wt) and polymerase knockout (Δ3Dpol) constructs were included as controls. GFP expression was monitored hourly using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents mean GFP positive cells per well at 8 hours post-transfection (n = 2, ± SEM).

**Figure 3. Replication of epitope tagged FMDV replicons.** (a) Cartoon of the FMDV replicon genome showing the positions of the inserted 3A or 3D epitope tags. (b) BHK-21 or (c) MDBK cells seeded into 24-well plates were allowed to adhere for 16 hours before transfection with replicon transcripts containing epitope tags using Lipofectin reagent. Wild-type GFP-PAC replicon (wt) and polymerase knockout (Δ3Dpol) constructs were included as controls. GFP expression was monitored hourly using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents mean GFP positive cells per well at 8 hours post-transfection (n = 2, ± SEM). (d) BHK-21 cells transfected with epitope tagged 3A and 3D constructs in addition to controls, protein lysates prepared at 8 hours post-transfection and probed by Western blot for FLAG and GAPDH loading control. (e) Simultaneously BHK-21 cells seeded onto glass coverslips were fixed in formaldehyde at 8 hours post-transfection before being stained for anti-FLAG (red), GFP expression (green) and cell nuclei counterstained with DAPI (blue). Images were captured by confocal microscopy. Scale bar is 50 µm.

**Figure 4. Replication-defective 3A insertions can be complemented in trans.** BHK-21 cells seeded into 24-well plates were allowed to adhere for 16 hours before co-transfection using Lipofectin reagent with mCherry replicons containing either a 3Dpol or 3A replication-defective mutation or wild-type control, and a wild-type (ptGFP) or polymerase knockout helper replicons expressing ptGFP (ptGFP 3DpolGNN and ptGFP Δ3Dpol) or yeast tRNA as a negative control. Both mCherry and ptGFP expression were monitored hourly using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents mean GFP positive cells per well at 8 hours post-transfection (n = 3, ± SEM, * = p < 0.05).
**Figure S1. Replication of the transposon-mutated replicon library.** BHK-21 cells seeded into 12-well plates were allowed to adhere for 16 hours before transfection with transposon-mutated replicon library using Escort reagent. Wild-type GFP-PAC (wt) and polymerase knockout (Δ3Dpol) constructs were also included, the latter as a negative control for input translation. GFP expression was monitored hourly over 24 hours using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents typical GFP positive cells per well. Relative replication is at the maximal GFP expression approximately 8-9 hours post-transfection.

**Figure S2. RT-PCR from transposon-mutated library.** Total RNA was extracted from BHK-21 cells transfected with indicated transcripts at 8 hours post transfection. 2 µg of each extracted was used in reverse transcription reaction before replicon genomes were amplified by PCR. Control reactions containing no reverse transcriptase (RT) were set up in parallel. Amplified products were analysed by 1 % agarose gel electrophoresis. Samples were as follows: (1) positive control, (2) DNA ladder, (3) wt + RT, (4) Rep Tn library + RT, (5) Δ3Dpol + RT, (6) wt – RT, (7) Rep Tn library – RT, (8) Δ3Dpol – RT.

**Figure S3. Location of 3A transposon insertions.** Cartoon schematic of the predicted 3A structure showing the two predicted alpha helices and the transmembrane region, with the amino acid sequence shown below. Arrows indicate the location of transposon insertions with the number indicating the nucleotide position after which insertion occurred.

**Figure S4. Location of the replication-competent 3D450 insertion.** Cartoon of the 3Dpol crystal structure showing the conventional right-hand front view (a) and rotated ~ 180° (b). Motif C containing the active site motif and motif A are shown in green and blue, respectively. The replication-competent 3D450 insertion is situated between amino acids M149 and E150, highlighted in stick representation in hot pink positioned at the end of helix α5.
Figure

Click here to download Figure: Figure 3_revised.tif
Figure

Click here to download Figure: Figure 4.tif
Figure S1

Figure S1. Replication of the transposon-mutated replicon library. BHK-21 cells seeded into 12-well plates were allowed to adhere for 16 hours before transfection with transposon-mutated replicon library using Escort reagent. Wild-type GFP-PAC (wt) and polymerase knockout (∆3D<sub>pol</sub>) constructs were also included, the latter as a negative control for input translation. GFP expression was monitored hourly over 24 hours using an IncuCyte Zoom® Dual Colour FLR and analysed using the integrated software. Data shown represents typical GFP positive cells per well. Relative replication is at the maximal GFP expression approximately 8-9 hours post-transfection.
Figure S2

**Figure S2. RT-PCR from transposon-mutated library.** Total RNA was extracted from BHK-21 cells transfected with indicated transcripts at 8 hours post transfection. 2 µg of each extracted was used in reverse transcription reaction before replicon genomes were amplified by PCR. Control reactions containing no reverse transcriptase (RT) were set up in parallel. Amplified products were analysed by 1 % agarose gel electrophoresis. Samples were as follows: (1) positive control, (2) DNA ladder, (3) wt + RT, (4) Rep Tn library + RT, (5) Δ3Dpol + RT, (6) wt – RT, (7) Rep Tn library – RT, (8) Δ3Dpol – RT.
**Figure S3.** Location of 3A transposon insertions. Cartoon schematic of the predicted 3A structure showing the two predicted alpha helices and the transmembrane region, with the amino acid sequence shown below. Arrows indicate the location of transposon insertions with the number indicating the nucleotide position after which insertion occurred.
Figure S4. **Location of the replication-competent 3D450 insertion.** Cartoon of the 3D<sup>pol</sup> crystal structure showing the conventional right-hand front view (a) and rotated ~ 180° (b). Motif C containing the active site motif and motif A are shown in green and blue, respectively. The replication-competent 3D450 insertion is situated between amino acids M149 and E150, highlighted in stick representation in hot pink positioned at the end of helix α5.