

# Molecular Biological Techniques and Genome Sequencing/Bioinformatics Analyses Converge To Identify RNA Secondary Structures in the Genome of Foot-and-Mouth Disease Virus Essential for Replication

Martin D Ryan\*, Garry A Luke

Department of Biomedical Sciences Research Complex (BSRC), School of Biology, University of St. Andrews, St. Andrews, United Kingdom

## ABSTRACT

Foot-and-Mouth Disease Virus (FMDV) causes perennial infections of both domestic and wild cloven-hoofed animals around the globe causing severe economic damage and restrictions on World trade. Economies of the developing World are disproportionately affected by FMDV outbreaks. Present FMDV vaccines are 'killed': large quantities of highly virulent virus are grown in bulk, particles purified, then chemically inactivated. This requires expensive, high-containment, production facilities with attendant risks of breaches in biosecurity. The RNA structures we, and others, have identified could be partially weakened, or destabilised, rather than be completely disrupted, to produce attenuated FMDV strains. These could serve to either (i) enhance the biosecurity of conventional inactivated vaccine production methods, or, (ii) serve as a basis for the rational design of a new generation of live-attenuated vaccines.

**Keywords:** FMDV sequencing; FMDV replicons; Live-cell imaging; RNA structure; Synthetic biology

## INTRODUCTION

Foot-and-Mouth Disease Virus (FMDV) is a picornavirus within the same virus family as the human pathogens polio-, entero- and hepatitis A viruses. FMDV infects both domesticated and wild cloven-hoofed animals (e.g. cattle/buffalo/pigs/sheep/goats) in many regions of the World, causing substantial, year-upon-year, economic losses for the agricultural sector. The genomes of picornaviruses are single-stranded RNA of positive (mRNA) sense. The virus RNA (vRNA) possesses the same architecture as cellular mRNAs: a 5' Non-Coding (untranslated) Region (5'NCR), a long open reading frame (ORF), a 3' NCR region and a poly (A) tail (Figure 1, Panel A). The tools available to manipulate RNA directly are, however, virtually non-existent, but cloning such RNA genomes into complementary DNA (cDNA) enables one to utilise the manifold techniques of recombinant DNA technology including synthetic biology, to modify the genome by insertions/deletions and mutagenesis. Over two decades ago it was shown that a full-length cDNA copy of the poliovirus genome cloned into a plasmid was 'infectious': when such plasmid DNA was transfected into mammalian tissue-culture cells live poliovirus could be rescued [1]. Subsequently

such an 'infectious (cDNA) copy' of the FMDV RNA genome (~8,300nts) was created [2]. FMDV is a highly significant animal pathogen and is amongst the most contagious mammalian viruses: research on both viruses and infectious copies is, therefore, restricted to licensed high-containment facilities such as the UK Pirbright Institute.

## FMDV REPLICONS

Long-standing observations showed that passage of picornaviruses at high multiplicity of infection (moi) produced 'defective-interfering' particles in which replication competent genomes, with naturally-occurring in-frame deletions of sequences encoding capsid proteins, were packaged into particles by co-infecting virus with full-length genomes (packaging in trans). Modern molecular biological techniques enabled such genomes to be created by design: notably, sequences encoding capsid proteins which were deleted could be replaced with sequences encoding 'reporter' proteins maintaining the long ORF. With permission from the UK regulatory authorities, at St Andrews we constructed a 'replicon' form of the genome in which the vast majority of sequences encoding the capsid

**Correspondence to:** Martin Ryan, Department of Biomedical Sciences Research Complex (BSRC), School of Biology, University of St. Andrews, St. Andrews, Fife, KY16 9ST, United Kingdom, E-mail: martin.ryan@st-andrews.ac.uk

**Received:** April 23, 2021; **Accepted:** May 07, 2021; **Published:** May 14, 2021

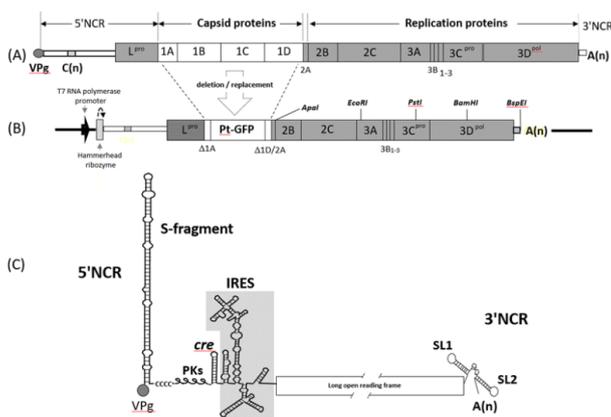
**Citation:** Ryan M, Luke GA (2021) Molecular Biological Techniques and Genome Sequencing/Bioinformatics Analyses Converge To Identify RNA Secondary Structures in the Genome of Foot-and-Mouth Disease Virus Essential for Replication. *J Infect Dis Prev Med.* 9.223.

**Copyright:** © (2021) Ryan M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

proteins (termed 1A, 1B, 1C and 1D) were deleted and replaced with sequences encoding Ptilosarcus green fluorescent protein (PtGFP; Figure 1, Panel B) all replication proteins being present. When transcribed into RNA and transcripts transfected into mammalian (BHK21) cells, replicon RNA replicated exactly like virus RNA, but could not form infectious virus particles: a completely biosecure procedure. Crucially, the FMDV replicon system we developed enabled RNA replication to be monitored in real time, in live cells, by GFP fluorescence [3]. This allowed us to study the effect of genome modifications upon replication competence: (i) without recourse to Reverse Transcription Polymerase Chain Reaction (RT-PCR) analyses of multiple samples taken throughout the replication cycle and, crucially, (ii) outwith high disease secure facilities.

## RNA STRUCTURES WITHIN THE FMDV GENOME

One function of the RNA genome is, naturally, to encode the virus polyprotein. RNA possesses, however, a notably higher propensity than DNA to form secondary structures: such features in the vRNA may interact with other vRNA structures, virus and/or host-cell proteins. A combination of bioinformatics and RNA structure probing techniques showed the 5'NCR of picornaviruses comprised numerous RNA structural features: in the case of FMDV, the 5'NCR comprises ~14% of the entire genome. The oligopeptide VPg (viral protein genome-linked) is covalently attached to the 5' end of the genome followed by a large stem-loop structure the 'S-fragment'. A poly(C) tract is followed by a (strain-variable) number of RNA pseudoknots and a stem-loop structure-the cis-acting replication element (cre), involved in RNA replication. A series of structures comprise an Internal Ribosome Entry Site (IRES) which recruits the translation machinery at this internal site-conferring a (7meG) cap-independent mode of translation. The short 3'NCR region comprises two stem-loop structures (SL1, SL2) which stimulate IRES activity, with recent data strongly suggesting a direct, long distance, interaction between SL1/2 and the IRES [4]. The poly (A) tail is templated by the genome, rather than being added by cellular poly (A) polymerase (Figure 1, Panel C).



**Figure 1:** Architecture of the FMDV genome. A small oligopeptide (VPg) is covalently bound to the 5' terminus of the vRNA. The 5' non-coding region precedes the long open reading frame comprising the L proteinase (Lpro), the capsid proteins 1A-1D and the RNA replication proteins 2B to the 3D

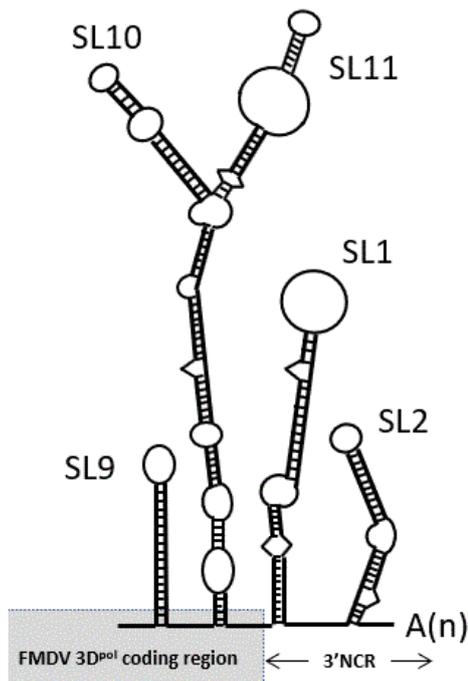
RNA-dependent RNA polymerase (3Dpol). The short 3'NCR precedes the poly (A) tail templated by the virus genome (Panel A). A complementary DNA infectious copy of the genome was modified such that the vast majority of the region encoding the capsid proteins was deleted and replaced by Ptilosarcus GFP (PtGFP; Panel B). The replicon genome is transcribed from the T7 RNA polymerase and the hammerhead ribosome self-cleaves to produce an authentic (virus) 5' terminus for the RNA transcript RNA. Previously identified RNA structures within the 5' and 3' NCRs are shown: the function of the S-fragment and the RNA pseudoknots (PKs) are unknown, the cis-acting element (cre) is involved in priming vRNA replication whilst the internal ribosome entry site (IRES) binds the translational machinery to initiate a cap-independent mode of translation. Stem-loops 1 and 2 within the 3'NCR interact with the IRES and serve as translational enhancers (Panel C)

## IDENTIFICATION OF FURTHER RNA STRUCTURES

Alongside characterising vRNA structures and their role in replication, a major goal of this research collaboration (Uni. St. Andrews/Pirbright Institute/Uni. Oxford) was to determine (i) whether mutation of such structures could attenuate the virus, and (ii) to identify any further RNA structures within the region encoding the polyprotein: in addition identifying any constraints upon introducing mutations within this region. Our strategy was two-fold: firstly, the group at the Pirbright Institute expanded the FMDV sequence database by sequencing and, together with Prof. Simmonds (Uni. Oxford) performed comprehensive bioinformatics analyses to identify RNA structures conserved amongst the seven FMDV serotypes (O, A, Asia 1, C and SATs 1/2/3) [5]. Secondly, at St Andrews we individually introduced into the FMDV replicon short (334-687 bp) synthetic gene blocks covering the entire tract encoding the 2A through to 3D<sup>pol</sup> replication proteins (Figure 1, Panel B). These gene blocks were designed by Prof. Simmonds such that potential RNA structures were disrupted whilst maintaining the amino acid coding, native dinucleotide frequencies and codon usage identical to that of wild-type sequences (generically referred-to as 'CDLR' mutants). The effect upon replication could rapidly be assessed by transfecting cells with each mutated FMDV replicon, live-cell imaging and quantifying the PtGFP fluorescence (and hence RNA replication) arising from each construct [5].

Bioinformatics analyses of the expanded sequence database predicted all of the known RNA secondary structural features shown in Figure 1, Panel C. An additional 45 stem-loops (counting each RNA hairpin individually, although located within a single branched structure) were conserved among the various serotypes within the region encoding the replication proteins - interestingly, 17 occurring within the region encoding 3D<sup>pol</sup>. In parallel, data from the analyses of the CDLR mutated replicons showed no, or very little, effect upon replication of CDLR mutant forms throughout the region encoding 2A through to 3C<sup>pro</sup> plus the region encoding the N-terminal ~50% of 3D<sup>pol</sup>. However, introduction of the CDLR mutated gene block encoding the C-terminal ~50% of 3D<sup>pol</sup> (Figure 1, Panel B) into

the FMDV replicon abrogated replication. These data are consistent with studies which showed that in the genome of another picornavirus, Poliovirus, RNA stem-loops within the coding region of 3D<sup>pol</sup> are required for viral RNA synthesis [6,7]. A detailed mutagenetic analysis of stem loop structures present with the FMDV genome encoding the C-terminal region of 3D<sup>pol</sup> (within the *Bam*HI to *Bsp*EI restriction enzyme sites in the replicon cDNA, Figure 1, Panel B) identified three stem-loop structures (SL9, SL10 and SL11) (Figure 2) whose disruption had major effects upon replication. Combining the disruption of stem-loops 9, 10 and 11 abrogated RNA replication [5].



**Figure 2:** RNA Structures Within the 3' region of the genome. The RNA stem-loop structures SL9, SL10 and SL11 were newly identified and shown to be required for RNA replication. These structures are formed by sequences encoding the C-terminus of the 3D polymerase. Stem-loop structures SL1 and SL2 (formed by the 3'NCR) were previously identified and shown to interact with the IRES and function as translational enhancers.

## CONCLUSION

Europe and N. America do not routinely vaccinate animals against FMDV. Disease control in these regions is based upon mass-slaughter of infected and surrounding, potentially susceptible, animals and bans on animal movements. Vaccine

stocks are, however, produced and maintained for use in emergency 'ring-vaccination' if culling has not controlled the spread of disease. The expensive high disease security facilities required to produce the killed vaccine naturally increases the cost-per-dose of FMDV vaccines. The conundrum facing FMDV vaccine producers is, therefore, that the markets that could afford to meet these higher costs do not routinely vaccinate! Farmers in Africa, India and Asia cannot afford to routinely vaccinate against FMDV. The FMDV serotypes O, A, Asia 1, are predominantly Indo-European viruses whilst the South African serotypes SATs 1-3 are confined to Africa: note, therefore, that full protection requires coverage against multiple, immunologically distinct, serotypes. The study outlined above increases the opportunities for the rational, large-scale, attenuation of a wide range of FMDV serotypes/strains. These could serve to either (i) enhance the biosecurity of conventional inactivated vaccine production methods, or, (ii) serve as a basis for the rational design of a new generation of live-attenuated vaccines-reducing the crucial cost-per-dose of FMDV vaccines.

## CONFLICT OF INTEREST

The authors report no conflict of interests.

## REFERENCES

1. Racaniello VR, Baltimore D. Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science*. 1981; 214:916-919.
2. Zibert A, Maass G, Strebel K, Falk MM, Beck E. Infectious foot-and-mouth disease virus derived from a cloned full-length cDNA. *J Virol*. 1990; 64(4):2467-2473.
3. Tulloch F, Pathania U, Luke GA, Nicholson J, Stonehouse NJ, et al. FMDV replicons encoding green fluorescent protein are replication competent. *J. Virol. Methods*. 2014;209:35-40.
4. Diaz-Toledano R, Lozano, G, Martinez-Salas, E. In-cell SHAPE uncovers dynamic interactions between the untranslated regions of the foot-and-mouth disease virus RNA. *Nuc. Acids Res*. 2017; 45(3):1416-1432.
5. Lasecka-Dykes L, Tulloch F, Simmonds P, Luke GA, Ribeca P, et al. Mutagenesis mapping of RNA structures within the Foot-and-Mouth Disease Virus genome reveals functional elements localised in the polymerase (3Dpol) encoding region. *mSphere* (in the press).
6. Burrill CP, Westesson O, Schulte MB, Strings VR, Segal M, et al. Global RNA structure analysis of poliovirus identifies a conserved RNA structure involved in viral replication and infectivity. *J. Virol*. 2013; 87(21):11670-11683.
7. Song Y, Liu Y, Ward CB, Mueller S, Futcher B, et al. Identification of two functionally redundant RNA elements in the coding sequence of poliovirus using computer-generated design. *Proc. Natl. Acad. Sci. USA*. 2012;109(36):14301-14307.