INTERFACE FOCUS

royalsocietypublishing.org/journal/rsfs

Research

**Cite this article:** Orton RJ, Wright CF, King DP, Haydon DT. 2019 Estimating viral bottleneck sizes for FMDV transmission within and between hosts and implications for the rate of viral evolution. *Interface Focus* **10**: 20190066.
http://dx.doi.org/10.1098/rsfs.2019.0066

Accepted: 30 October 2019

One contribution of 9 to a theme issue 'Multi-scale dynamics of infectious diseases'.

**Subject Areas:**
environmental science, systems biology

**Keywords:**
foot-and-mouth disease virus, transmission, bottleneck, molecular clock

**Author for correspondence:**
Daniel T. Haydon
e-mail: daniel.haydon@glasgow.ac.uk



# Estimating viral bottleneck sizes for FMDV transmission within and between hosts and implications for the rate of viral evolution

Richard J. Orton[1,2,†], Caroline F. Wright[3,†], Donald P. King[3] and Daniel T. Haydon[1]

[1]Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK
[2]MRC-University of Glasgow Centre for Virus Research, University of Glasgow, Sir Michael Stoker Building, 464 Bearsden Road, Glasgow G61 1QH, UK
[3]The Pirbright Institute, Ash Road, Pirbright GU24 0NF, UK

RJO, 0000-0002-3389-4325; CFW, 0000-0002-1770-4325; DPK, 0000-0002-6959-2708; DTH, 0000-0002-1240-1886

RNA viruses exist as populations of closely related genomes, characterized by a high diversity of low-frequency variants. As viral genomes from one population disperse to establish new sites of replication, the fate of these low-frequency variants depends to a large extent on the size of the founding population. Focusing on foot-and-mouth disease virus (FMDV) we conjecture that variants are more likely to be transmitted through wide bottlenecks, but more likely to approach fixation in new populations following narrow bottlenecks; therefore, the longer-term rate of accumulation of 'nearly neutral' variants at high frequencies is likely to be inversely related to the bottleneck size. We examine this conjecture *in vivo* by estimating bottleneck sizes relating 'parent' and 'daughter' populations observed at different scales ranging from within host to between host (within the same herd, and in different herds) using a previously established method. Within hosts, we find bottleneck sizes to range from 5 to 20 viral genomes between populations transmitted from the pharynx to the serum, and from 4 to 54 between serum and lesion populations. Between hosts, we find bottleneck sizes to range from 2 to 39, suggesting inter-host bottlenecks are of a similar size to intra-host bottlenecks. We establish a statistically significant negative relationship between the probability of genomic consensus level change and bottleneck size, and present a simple sampling model that captures this empirical relationship. We also present a novel *in vitro* experiment to investigate the impact of bottleneck size on the frequency of mutations within FMDV populations, demonstrate that variant frequency in a population increases more rapidly during small population passages, and provide evidence for positive selection during the passage of large populations.

## 1. Background

Foot-and-mouth disease virus (FMDV) is a positive-sense RNA virus, a member of the *Picornaviridae* family [1], and the causative agent of the highly contagious foot-and-mouth disease (FMD). The genome of FMDV is a single sequence of positive-sense RNA, length approximately 8400 nucleotides. RNA viruses such as FMDV exist as populations of closely related genomes as a consequence of the poor proofreading ability of their RNA dependent RNA polymerase, with the error rate reported to be of the order of $10^{-3}$–$10^{-5}$ mutations per nucleotide copied [2–5]. This confers a high degree of genetic heterogeneity

[†]These authors contributed equally to this work.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.4728779.

THE ROYAL SOCIETY PUBLISHING

on RNA virus populations, which is thought to favour adaptability to different environments, hosts and drug treatments. High throughput sequencing (HTS) is now routinely used to interrogate FMDV-positive samples and investigate the diversity contained within these viral populations [6–8].

An important factor that determines the infectivity of animal viruses is the presence of suitable cellular surface receptors for attachment and internalization. The epithelial cell expressed heterodimer, integrin, has been shown to be the cellular receptor for FMDV *in vivo* [9–11]. However, Jackson *et al*. [12] observed that the glycosaminoglycan, heparan sulfate (HS), could mediate the interaction of FMDV serotype O with cells in culture. FMDV infected cattle typically show clinical signs 2–6 days post exposure that include vesicles on the coronary bands of the feet, in the mouth, on the tongue and teats [13]. Although alternative primary sites of replication have been studied [14], in cattle, rapid dissemination of FMDV from host entry most likely follows initial replication in the pharyngeal area [15], passing into the systemic circulation [16–18], where the virus is not thought to replicate. From there, the virus is transported to other distant, non-contiguous epithelia, including those of the feet, where the virus can once again replicate [19]. As a consequence of this establishment of new local foci, the viral population passes through intra-host bottlenecks. Subsequent transmission of the virus to a naive host (via inter-host bottlenecks) most frequently occurs shortly after the appearance of clinical signs [20] when an infected individual can secrete large amounts of viral particles into the environment.

Much of the genetic variation within FMDV populations is thought to be under nearly neutral selection or varying levels of purifying selection, with evidence for positive selection observed in only a small fraction of capsid codons perhaps in response to interaction with the host immune system [21]. However, novel variants continually arise, increasing from very low frequencies to frequencies greater than 0.5 in HTS alignments (consensus level mutations) with an average of about three such consensus mutations reported to arise between individuals sequentially infected with FMDV [22]. The potential for novel mutations within a viral population to spread both within and between hosts is heavily dependent on the size of the transmission bottleneck [15,23–25]. If a transmission bottleneck is wide, numerous low-frequency variants present in the 'parent' population can be transmitted to the 'daughter' population, facilitating their subsequent spread through the host population [25,26]. By contrast, if a transmission bottleneck is narrow, few low-frequency variants in the parent will pass to the daughter, but through virtue of the narrow bottleneck, those that do will have a higher chance of representation in the daughter population at a disproportionately higher frequency. Indeed, variants that are rare in the parent population may achieve consensus level in daughter populations founded through narrow bottlenecks. Thus, bottleneck size is likely to have a bearing on the rate at which the molecular clocks tick, with potentially important implications for the rate of evolution of RNA viruses.

One of the best studied viruses in relation to the quantification of transmission bottlenecks is the influenza A virus (IAV). Narrow transmission bottlenecks have been reported in experimental infections of ferrets [27,28], while wide bottlenecks have been reported in horses enabling the transmission of numerous low-frequency variants [29,30]. Recently, novel computational methods have been developed to quantify the size of transmission bottlenecks using viral HTS data and the variants present in parent and daughter populations [25,31]. In humans, inter-host IAV bottleneck reports are slightly conflicting, with some studies estimating wide bottlenecks of between 200 and 250 viral genomes [31,32], while another study estimated narrow bottlenecks of between 1 and 2 viral genomes [25], albeit on a different dataset; however, re-analysis of the wider bottleneck dataset has indicated some issues with the data [33]. HTS has also been used to provide initial evidence for bottleneck driven diversification of norovirus populations [34], as well as for determining a narrow inter-host bottleneck of only 1–3 viral genomes for hepatitis C virus [35,36], and human immunodeficiency virus (HIV) [37] transmission events. However, few studies have investigated bottlenecks in FMDV [38,39], and to date, there has been little work specifically quantifying the size of FMDV bottlenecks, nor on comparing bottlenecks across the various scales of transmission. Bottlenecks control how much genetic information from one population founds another, and quantification of bottlenecks is, therefore, essential for our understanding of how FMDV transmits and evolves.

There have been numerous studies investigating the effect of bottleneck size on viral fitness and population diversity *in vitro*, particularly using plaque-to-plaque transfers. As a result of genetic drift, mutations may more rapidly approach fixation in small founding populations by chance and, if deleterious, can lead to the decline in the replicative ability of the daughter population. Bottleneck associated decline in viral population fitness has been demonstrated experimentally, using this technique, for a range of RNA viruses [40–43]. This fitness loss induced by bottleneck effects (termed Muller's ratchet) has also been reported during serial contact transmission of FMDV in pigs [38]. Furthermore, it has been demonstrated for a range of viruses that the serial passage of small heterogeneous populations through sequential bottlenecks results in reduced diversity [30,44–47]. By contrast, the serial passage of relatively large populations of heterogeneous viruses is generally accompanied by an overall gain in population fitness [39,48,49], possibly explained by competitive optimization between different mutants within the viral swarm [49]. In some studies, bottleneck size is controlled experimentally by varying the multiplicity of infection (MOI—the ratio of viruses to cells in the initial culture) between 0.01–0.1 (low) and 1–10 (high) [39,49]. But tighter experimental control is maintained by varying the size of the founding populations while holding MOI constant and requires using flasks of an appropriate size for each population size [50]. Positive selection has also been observed during large population passages but not during narrower more bottlenecked transfers [39].

In this study, we first analyse existing HTS data from controlled and natural FMDV transmission chains to quantify the transmission bottlenecks arising at different scales. Parent–daughter populations were examined within single hosts, and between hosts that were in the same herd (infected under controlled conditions) and in different herds (infected during the UK 2001 outbreak). We also present a simple model to investigate how bottleneck size can affect observed consensus level changes and thus in principle the observed molecular clock rate. We then present the results of a novel *in vitro* experiment, that uses HTS to dissect and compare the populations of FMDV passaged through serial narrow bottlenecks to FMDV populations passaged through serial wide

bottlenecks; importantly, MOI was kept constant between the two passages. The aim of this study was to use these *in vivo* and *in vitro* experiments to explore the hypothesis that mutations will reach consensus level more rapidly in populations during more severely bottlenecked transmissions.

# 2. Methods

## 2.1. Data from an experimental transmission chain

HTS data generated from experimental studies in cattle [6,7,51] were used to investigate both intra-host and inter-host transmission. For full details see [6,7], but briefly, 21 samples from a chain of sequentially infected animals were ultra-deep sequenced on the Illumina Genome Analyzer platform. Animal 1 (A1) was inoculated intradermolingually with FMDV ($O_1$BFS 1860 at $10^{5.7}$ $TCID_{50}$). The inoculum was derived from a bovine tongue vesicle specimen that had been passaged extensively in cell culture [52]. The previous analysis showed that the virus quickly reverts from using HS as a cellular receptor in the first animal in the chain, with dNdS ratios also falling below one after the first animal [7]. In addition, entropy analysis showed an immediate increase from the initial inoculum, followed by variation between samples from the time of first sampling onwards, with no temporal trend apparent [7], and with diversity levels comparable to samples from the UK 2001 outbreak (see electronic supplementary material). Twenty-four hours post needle-challenge, A1 was used to challenge naive calf A2 by direct contact for a total of 4 days. The transmission chain was subsequently maintained for a further two cycles by first placing infected A2 in direct contact with a naive calf (A3) for 24 h, and then placing infected A3 in direct contact with a further naive calf (A5) for 14 days. FMDV-positive samples were normalized to $10^6$ copies of FMDV RNA per microliter prior to RT-PCR amplification to create two long overlapping fragments of 4065 bp (PCR1) and 4033 bp (PCR2). Samples were then sequenced twice on an Illumina Genome Analyzer IIx to generate single-end reads of 70–73 nt. For the analyses presented here, reads were filtered using an average quality of Q30, and the first and last five bases of reads were removed due to elevated mismatches caused by indels [6,7]. Reads were then aligned to the FMDV reference genome $O_1$BFS1860 (GenBank accession EU448369) using bwa-mem [53], variants were identified using DiversiTools (http://josephhughes.github.io/DiversiTools/) and only variants observed above a 0.5% frequency threshold, based on RT-PCR clone control data [54], and outside of primer regions were included in the analyses. Although samples were previously sequenced in duplicate, for this analysis we only consider the first replicate of each sample to facilitate comparison with data from other scales which are only sequenced once.

FMDV infection within the host is believed to transmit from the pharynx, through the blood, and to the epithelia. Although simplistic, intra-host samples were split into parent–daughter pairs to reflect this (probang-to-serum and serum-to-lesion), as well as linking multiple probang (probang-to-probang) and serum (serum-to-serum) samples from the same animal into temporal parent–daughter pairs. For probang-to-serum pairs, for each potential daughter serum sample in an individual animal, the nearest previous probang sample (in days) from the same animal is designated the parent. For serum-to-lesion samples, for each daughter lesion sample in an animal, the serum sample with the largest estimated bottleneck to the lesion sample is designated the parent. For inter-host parent–daughter pairs, the earliest probang sample from the receiving animal is designated the daughter, while the parent sample is identified from the donor animal as the sample with the largest estimated bottleneck to the daughter sample.

## 2.2. Data from the UK 2001 outbreak

Five premises from the Darlington cluster of the 2001 outbreak of FMDV in the UK were deep sequenced. The five infected premises (IPs) are labelled A, B, C, K and N, as designated in [52], and constitute three inferred farm-to-farm transmission events (K-C, K-B and A-N) previously determined [52,55]; all samples had their genomes Sanger sequenced previously [52]. Although at a different scale, a farm-to-farm transmission event is essentially an inter-host transmission event. These samples were sequenced using the same standardization, reverse transcription and PCR protocols described above in §2.1. However, although both PCR fragments were attempted, only PCR1 was successful and sent for sequencing on an Illumina MiSeq to generate paired end reads of 151 nt. Reads were processed and analysed as described in §2.1.

## 2.3. Bottleneck quantification

We used the beta-binomial sampling method developed by Sobel *et al*. [31], and applied in [25], to infer bottleneck sizes between parent and daughter samples (see the electronic supplementary material and [31] for full details). It is important to note that the method does not determine the number of viral genomes passing between parent and daughter, but the number of viral genomes that pass into the daughter and contribute genetically to the viral population that is being sequenced. The method is similar to a standard presence/absence model, but allows variant frequencies in the daughter host to change between the time of founding and the time of sampling (to account for the stochasticity of viral replication dynamics in the early stages of infection). The method is insensitive to the time interval between transmission and sampling, but determined by the size of the founding population ($N_b$) and the number of variant genomes present in it. It is assumed that the variant frequency in the parent remains constant between sampling and transmission [25]. We report the maximum-likelihood estimates (MLE) of bottlenecks with associated 95% confidence intervals.

## 2.4. Relating bottleneck size to consensus mutations

To investigate whether the bottleneck size will have an influence on consensus level differences between a parent and daughter, we use the binomial distribution. The probability that a variant at site $i$ in the parent is represented at a consensus frequency of 50% or more in the bottleneck population ($p_{i,c}$) is modelled as

$$p_{i,c} = \sum_{j=N_b/2}^{N_b} \binom{N_b}{j} (1 - V_{P,i})^{N_b - j} V_{P,i}^{j}, \tag{2.1}$$

where $V_{P,i}$ is the variant frequency at genome position $i$. The expected number of consensus level differences between parent and daughter populations ($D$) is predicted by $\sum_{i=1}^{S} p_{i,c}$, where $S$ is the number of nucleotide positions considered.

The number of consensus level differences between parent and daughter populations was related to estimated bottleneck size using a generalized linear model (glm) with Poisson error in R (v.3.3.3). Significance was determined using a likelihood ratio test comparing to the intercept only model.

## 2.5. *In vitro* bottleneck experiment

FMDV was subjected to two different passage series *in vitro*. One passage series was propagated through three serial narrow bottleneck transmissions while the other was through three wide bottleneck transmissions. Bottleneck size was controlled by varying the magnitude of the viral population transferred at each cell-culture passage. However, a requirement of generating comparable results is that the MOI is maintained at a constant level regardless of bottleneck size, and to achieve this the number of

cells available for infection was manipulated by using different sized cell-culture vessels: T175 flasks (175 $cm^2$, $2 \times 10^7$ cells) for the large (L) populations with wide bottlenecks, and 96 well, flat bottomed ELISA plates (0.3 $cm^2$, $3.4 \times 10^4$ cells) for the small (S) populations with narrow bottlenecks. The virus used was rescued from a plasmid containing full-length FMDV O1Kaufbeuren cDNA (pT7S3) [56]. This infectious copy was a cell-culture-adapted B64 strain of the O1Kaufbeuren virus (O1 K B64) with the ability to use cell surface HS molecules as receptors [57]. After transcription from the plasmid, the transcribed RNA was electroporated into BHK cells and allowed to replicate for 6 h. The rescued viral progeny was then used to infect a fresh monolayer of BHKs and allowed to replicate for 24 h; the resultant viral progeny was the starting input for both the L and S passages. The cell line used was a fetal goat tongue epithelium cell line (ZZ-R 127), from the Friedrich-Loeffler-Institute collection of cell lines in veterinary medicine, which expresses the αvβ6 integrin receptor [58]; medium for this cell line was DMEM/F-12 supplemented with 1% fetal bovine serum, 1% L-glutamine and 1% pen/strep. A concern was that a large number of passages would be required before mutation frequency reached levels above background sequence noise (0.5% [54]). Therefore, both the virus and the cell line used were selected in order to induce a degree of selective pressure at known sites associated with a reversion in cellular receptor usage from HS to integrin. All *in vitro* infections were conducted at a constant MOI, as calculated by viral RNA copy number quantification by qRT-PCR, with a fixed incubation period of 12 h at 37°C. Briefly, viral RNA copy number was quantified within the starting input virus for both L and S passages by qRT-PCR and diluted appropriately for a relative MOI of 0.01 (the PFU:RNA copy number ratio was calculated according to PFU titres and viral RNA copy number measurements, in duplicate, for the starting input virus grown in BHK cells). Where possible, the volume of input virus was kept constant for both L and S passages (2 ml for L passages and 10 µl for S passages). This *in vitro* bottleneck experiment was not replicated.

All viral populations from both passage series were ultra-deep sequenced on the Illumina Genome Analyzer platform. In brief, total RNA extraction, RT, PCR and product visualization were as described [7]. The FMDV genomic region anticipated to come under selective pressure during the cell-culture passage was incorporated within amplification fragment 1 (PCR1), so this PCR assay alone was performed. qRT-PCR was as described in [7]. Samples were then sequenced on the Illumina Genome Analyzer IIx to generate single-end reads of 73 nt; reads were processed and analysed as described in §2.1.

# 3. Results

## 3.1. Intra-host bottlenecks

Previously, intra-host diversity of FMDV samples from different tissues in a chain of sequentially infected cattle was measured using HTS [6,7]. Here, using the methodology developed by Sobel *et al*. [31], we quantified the intra- and inter-host bottleneck population sizes within this transmission chain, determining the MLE of the number of viral genomes passing from the initial pharyngeal infection into the blood circulatory system, and subsequently seeding infection in terminal foot lesions (table 1). We observe a median intra-host bottleneck of 5.5 viral genomes (range 2–20) between pharyngeal and serum samples, suggesting a relatively narrow bottleneck from the initial site of infection into the rest of the body. Perhaps surprisingly, sequential probang samples from the same animal displayed a similarly small bottleneck size, median 8 (range 4–16), but probang scraping samples may only capture a small proportion of the viral diversity in the pharynx due to the numerous lesions present, each potentially containing a distinct FMDV population. We observe a much wider median bottleneck of 57 genomes (range 37–174) between sequential serum samples from the same animal indicative of a more continuously mixed population. Intra-host bottlenecks between the serum and terminal foot lesions were observed to have a median of 18 genomes, but with a range of between 4 and 54, highlighting variability in the lesion founder process.

Overall, these results correlate well with the observed consensus level differences between samples, with sequential serum samples maintaining the same consensus sequence (with one exception involving a 54.4% variant), while between 0 and 3 consensus level mutations across the genome are typically observed between probang–serum and serum–lesion samples [7].

## 3.2. Inter-host bottlenecks

Quantification of the three inter-host bottlenecks in the transmission chain led to MLE bottleneck sizes of 4, 10 and 39 viral genomes (table 2), respectively, suggesting inter-host bottlenecks are of a similar size to intra-host bottlenecks (except the wide serum-to-serum bottlenecks), with relatively few viral genomes from the parent population founding infection in the daughter. This again correlates well with consensus level differences with comparable inter- and intra-host consensus level mutations observed [7,22]. The three inter-host bottlenecks from the farm-to-farm transmission events of the 2001 FMDV outbreak generated MLEs of 2, 4 and 9, respectively. Inter-farm bottlenecks could well be expected to be narrower than inter-host events on the same farm, due to aerial or fomite transfer of small numbers of virions over large geographical distances. However, our results suggest that inter-farm bottlenecks are not substantially different from the other inter-host bottlenecks analysed, although the data are limited. Again, bottleneck size appears to correlate with consensus level distances, with the most distant samples in terms of consensus (IP-K, IP-C) having the narrowest bottleneck and the closest consensus sequences (IP-A, IP-N) having the widest.

## 3.3. Relating bottleneck size to consensus mutations

Equation (2.1) was used to generate a contour surface describing the probability that variants present at minority frequencies might be disproportionately represented as consensus level variants in bottlenecks of different sizes (figure 1*a*). While variants are clearly more likely to be passed on through wide bottlenecks, it is only through narrow bottlenecks that they are likely to appear in the consensus sequences of daughter populations.

The observed number of consensus level differences between parent and daughter populations from the intra- and inter-host levels was related to MLE bottleneck size using a glm which revealed a strong negative relationship between the number of consensus level differences between parent and daughter populations and the bottleneck size linking the two (figure 1*b*, $\chi^2 = 23.0$, d.f. = 1, $p < 0.0001$).

## 3.4. *In vitro* bottlenecks

FMDV was subjected to two serial passages *in vitro*. One passage series was propagated through serial narrow bottleneck

**Table 1.** Intra-host bottleneck estimates between different types of host samples. Bottlenecks were determined using the methodology of Sobel *et al*. [31], with the MLE of bottleneck size shown, with 95% confidence intervals in square brackets. Sample labels contain three elements: (1) the animal number in the transmission chain: A2, A3 or A5; (2) the number of days post first contact (DPC); and (3) the sample type: PB (probang), SR (serum) and front/back (F/B) left/right (L/R) foot (F) lesion.

| sample category | parent | daughter | bottleneck [viral genomes] | consensus differences |
|---|---|---|---|---|
| probang-to-probang | A2_2DPC_PB | A2_4DPC_PB | 6 [3–10] | 3 |
| | A2_4DPC_PB | A2_6DPC_PB | 16 [9–26] | 4 |
| | A3_1DPC_PB | A3_3DPC_PB | 8 [4–13] | 1 |
| | A3_3DPC_PB | A3_5DPC_PB | 4 [2–7] | 5 |
| | A5_5DPC_PB | A5_7DPC_PB | 14 [7–27] | 2 |
| probang-to-serum | A2_2DPC_PB | A2_3DPC_SR | 10 [6–15] | 3 |
| | A2_2DPC_PB | A2_4DPC_SR | 7 [4–12] | 2 |
| | A2_4DPC_PB | A2_5DPC_SR | 20 [10–36] | 1 |
| | A3_1DPC_PB | A3_3DPC_SR | 3 [1–5] | 1 |
| | A3_3DPC_PB | A3_4DPC_SR | 4 [2–6] | 5 |
| | A3_3DPC_PB | A3_5DPC_SR | 2 [1–4] | 5 |
| serum-to-serum | A2_3DPC_SR | A2_4DPC_SR | 56 [35–86] | 1 |
| | A2_4DPC_SR | A2_5DPC_SR | 37 [22–60] | 0 |
| | A3_3DPC_SR | A3_4DPC_SR | 174 [66–401] | 0 |
| | A3_4DPC_SR | A3_5DPC_SR | 58 [23–126] | 0 |
| serum-to-lesion | A2_3DPC_SR | A2_6DPC_BRF | 18 [10–39] | 1 |
| | A2_3DPC_SR | A2_6DPC_FLF | 4 [2–7] | 3 |
| | A2_3DPC_SR | A2_6DPC_FRF | 10 [5–17] | 4 |
| | A3_3DPC_SR | A3_5DPC_BLF | 54 [19–131] | 0 |
| | A3_4DPC_SR | A3_5DPC_BLF | 42 [15–101] | 0 |

**Table 2.** Inter-host bottleneck estimated from the transmission chain. Transmission chain samples are labelled as in table 1; samples from the UK 2001 outbreak are labelled with the IP letter as in [52]. Bottlenecks were determined using the methodology of Sobel *et al*. [31], with the MLE of bottleneck size shown, with 95% confidence intervals in square brackets.

| sample category | parent | daughter | bottleneck [viral genomes] | consensus differences |
|---|---|---|---|---|
| transmission chain | A1_2DPC_FLF | A2_2DPC_PB | 4 [2–8] | 3 |
| | A2_5DPC_SR | A3_1DPC_PB | 10 [5–17] | 2 |
| | A3_4DPC_SR | A5_5DPC_PB | 39 [20–70] | 0 |
| 2001 outbreak | IP-K | IP-C | 2 [1–3] | 10 |
| | IP-K | IP-B | 4 [2–7] | 7 |
| | IP-A | IP-N | 9 [5–13] | 6 |

transmissions while the other was through wide bottleneck transmissions. Bottleneck size was controlled by varying the size of the viral population transferred at each cell-culture passage, but MOI was kept constant by using different sized cell-culture vessels with a large population (L) for wide bottlenecks and a small population (S) for narrow bottlenecks. The aim of this study was to test the hypothesis that mutations, irrespective of selective value, approach fixation more rapidly in a population during more severe bottleneck transmissions. The rate of viral replication was measured by increases in viral RNA copy number, quantified by qRT-PCR [59]. Although the incubation period was constant (12 h), the extent of viral replication decreased over sequential passages for both series (figure 2); however, this decrease was more pronounced within population S compared to population L. Inoculation volume remained constant within population L but, was increased in the last passage of population S to compensate for a decrease in RNA copy number (figure 2). These results reveal a bottleneck associated decline in viral population fitness, which has previously been demonstrated experimentally *in vitro* for a range of RNA viruses [40–43].

A frequency threshold of 0.5% was used to identify mutations of interest above the expected error rate, with clear differences between the pattern of variants observed in the L and S populations (figure 3). In the first passage (P1) of population S (narrow bottlenecks), one variant was observed above 10%, and six variants were observed below

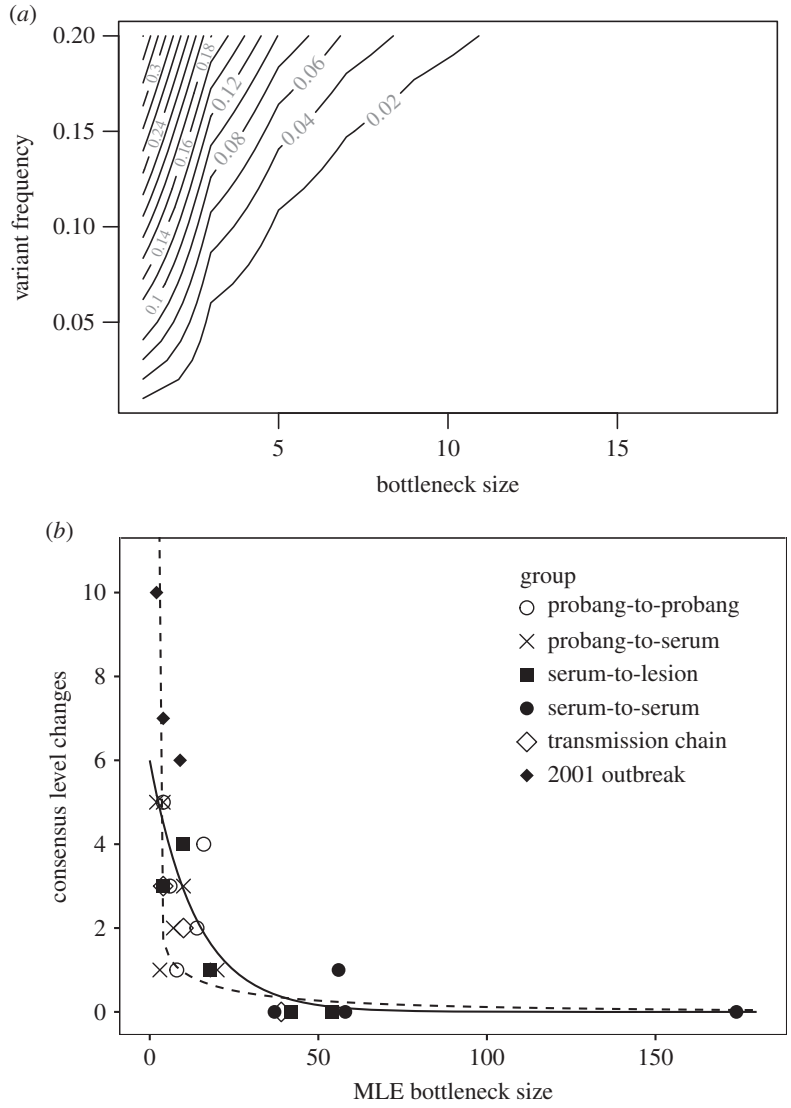


**Figure 1.** (*a*) Contour surface describing the probability (contours) that a variant present at different (minority) frequencies (*y*-axis) will be represented at consensus level in populations founded through bottlenecks of different sizes (*x*-axis). (*b*) The observed relationship between the number of differences (*D*) by which consensus sequences of parent and daughter populations differed and the MLE of the bottleneck size ($N_b$) linking the two populations. The solid line indicates the fit from the generalized linear model: $\ln(D) = 1.79–0.072 \times N_b$. The dashed line indicates the expected number of changes given by the model: $\sum_{i=1}^{5} p_{i,c}$ (conditional on $p_{i,c} > 0.005$, $V_{P,i}$ values taken from the A2_3DPC_SR sample variants). The 2001 outbreak samples were sequenced with PCR1 only.

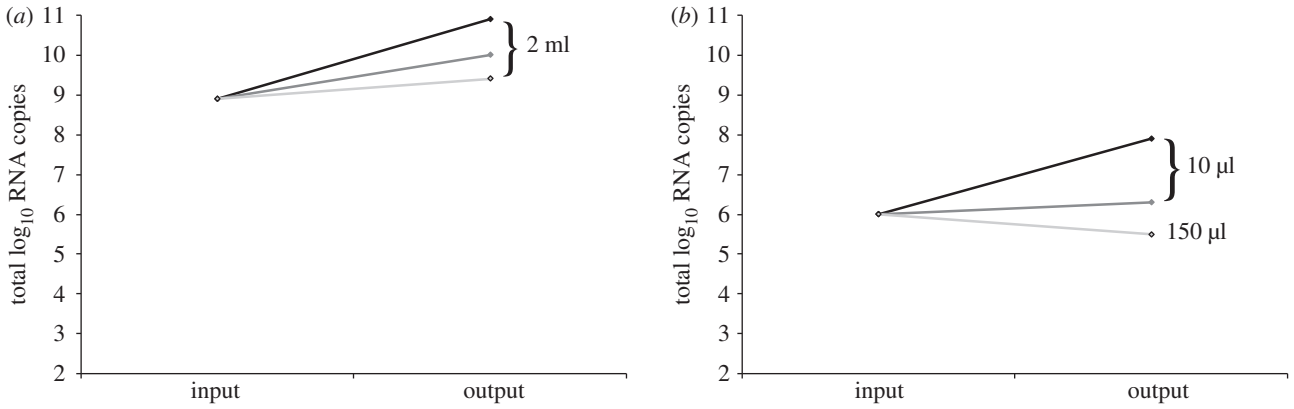


**Figure 2.** Quantification of O1 K B64 RNA copy number by qRT-PCR across three passages in cell culture for (*a*) population L and (*b*) population S. Both populations were infected at an MOI of 0.01. Passage 1 (P1: black line), P2 (dark grey line) and P3 (light grey line). Inoculation volumes are indicated.

10% but above 1%. However, the overall pattern in population S was relatively inconsistent, with variants increasing, decreasing or maintained at the same frequency through the passages. By contrast, the pattern observed in population L (wide bottlenecks) was markedly different. There were relatively few changes in P1 and although variants were observed above the 0.5% threshold, it took until P2 for variants to be observed greater than 1%, and until P3 for any frequency to reach the 10% observed in population S. In addition, the variants observed in population L share

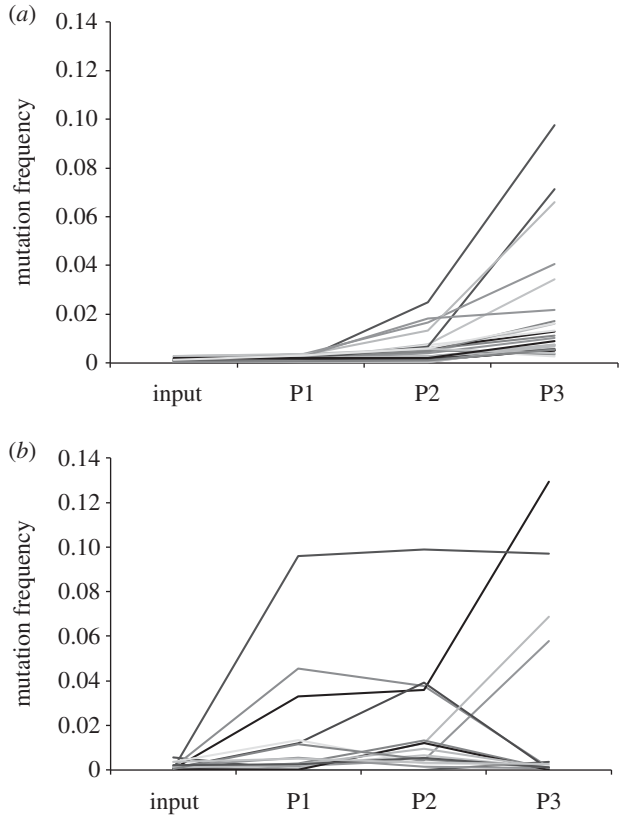


**Figure 3.** Variant frequencies across three passages for population L (*a*) and population S (*b*). The frequency of all variants observed above the 0.5% frequency threshold is shown across all three passages, for both populations L and S.

consistent patterns, with variant frequencies typically increasing with passage number. Overall, this experiment demonstrates that variants can increase in frequency more rapidly in a population during more severe bottleneck transmissions; however, as our simple model predicts, with such wide bottlenecks, consensus level changes do not occur even when selection is expected to be acting on some sites.

There was evidence for positive selection in population L in the form of increasing frequencies of non-synonymous mutations associated with the subtype $O_1$ FMDV-HS receptor complex (table 3). However, there was no evidence for positive selection in population S, which contained none of the observed variants associated with the switch to integrin receptor usage. Currently, there are nine amino acids associated with the subtype $O_1$ FMDV-HS receptor complex at residues (discussed in [60]). Seven of these nine variants were observed in population L. The C → T mutation at $VP3^{56}$ (thought of as the most critical motif in terms of virus/cell receptor recognition), only reached a frequency of 1.3% by P3. A potentially novel variant linked to O1 FMDV-HS receptor complex was observed at $VP3^{61}$ with a variant frequency of 7.3% (Val → Glu); $VP3^{61}$ is in close proximity to $VP3^{59}$ and $VP3^{60}$ and is exposed on the capsid surface (data not shown).

We applied the bottleneck quantification method of Sobel *et al*. [31] to the HTS data generated for populations L and S over the passage series; however, we note the method was validated on data derived from naturally occurring viruses, rather than data derived from rescued clones. As expected, the bottlenecks between passages were estimated to be very wide compared to cow-to-cow transmissions (electronic supplementary material, table S1). Although population S always has narrower bottlenecks than population L, both chains display a marked decrease in bottleneck size as the passage progresses. This could be linked to decrease in viral fitness/infectivity (as observed in population S), positive selection resulting in smaller founding populations (as observed in population L) or random drift.

## 4. Discussion

We have used HTS data to estimate the viral bottlenecks through which FMDV passes when transmitted within and between hosts. Within hosts, we estimated the median bottleneck size connecting pharyngeal and serum viral populations to be 5.5 viral genomes (range 2–20), and between serum and lesion populations to be 18 viral genomes (range 4–54). Between hosts estimated bottleneck sizes were 4, 10 and 39 making inter- and intra-host bottlenecks a similar size, and between herds we estimated bottlenecks to be 2, 4 and 9. A simplistic sampling model captures the nature of the negative relationship between bottleneck size and consensus level differences between parent and daughter populations, and is a reasonable match to the observed empirical relationship. Our results suggest that while all variants are more likely to be passed on through wider bottlenecks, it is only following narrow bottlenecks that variants are likely to appear in consensus sequences in daughter populations (although such narrow bottlenecks are likely to result in an initial loss of genetic variation).

We have presented the results of a novel *in vitro* cell-culture experiment in which bottleneck size was experimentally controlled to be wide or narrow (while maintaining constant MOI). Although each of the three passages was undertaken only once, the results show clear differences between the two populations. Passaging a small viral population through narrow bottlenecks resulted in immediate and substantial increases in observed variant frequencies, but the observed variants did not appear to be positively selected. By contrast, passaging a large viral population through wide bottlenecks resulted in a slower, more gradual increase in observed variant frequencies. As variants were observed at seven out of the nine codon positions known to be associated with the FMDV-HS receptor complex, there was strong evidence for positive selection in the large population, and suggesting that in the presence of positive selection, variants will accumulate at higher frequencies (and therefore potentially appear in consensus level sequences of HTS alignments) faster following wider bottlenecks.

With the advent of deep sequencing, within-host FMDV populations are now frequently characterized although it is often challenging to distinguish between low-frequency polymorphisms and RT-PCR error and sequencing artefacts. We have previously estimated [6] that a small fraction of the sites displays no variability, most sites exhibit a small amount of polymorphism, and a very small fraction of the sites (0.2–0.4%) present variation at a level above 1%. The deep sequences analysed in the current study exhibited comparable levels of variability. If these parent populations found daughter populations through wide bottlenecks this variation is extremely unlikely to appear at consensus level in the absence of selection. Therefore, the small bottleneck transmission events have a critical role to play in the rate that

**Table 3.** Details of the non-synonymous mutations associated with the subtype O1 FMDV-HS receptor complex in population L.

| genome position | mutation | frequency (%) | AA location | AA change |
|---|---|---|---|---|
| 2334 | A → C | 9.8 | VP2–134 | Lys[K] → Gln[Q] |
| 2346 | T → C | 3.4 | VP2–138 | Tyr[Y] → His[H] |
| 2754 | C → T | 1.3 | VP3–56 | Arg[R] → Cys[C] |
| 2764 | G → C | 0.7 | VP3–59 | Gly[G] → Ala[A] |
| 2767 | G → A | 1.7 | VP3–60 | Gly[G] → Asp[D] |
| 2851 | A → G | 4 | VP3–88 | Asn[N] → Ser[S] |
| 3832 | A → C | 2.2 | VP1–195 | His[H] → Pro[P] |

minority neutral variants become established in the consensus sequence of daughter populations, and in the longer term, since sites under positive selection are rare [21], the rate at which the molecular clock 'ticks'.

While there is a large body of theory that has established that the rate of fixation of mutations in finite populations slows as population size increases, and that the process of intra-clone competition (clonal interference [61,62]) can slow this process further, and increase the genetic diversity, we regard this as a distinct process from those we propose here. We think of the bottleneck—the size of a sample of the parent population that founds the daughter population—as a non-replicating 'package' of viruses, transported to the site of the new population, after which the new population expands rapidly in size. Consequently, the changes in variant frequency caused by bottlenecking result purely from a random sampling process, and not through any replicative or competition process [63]. Furthermore, this sampling process is of the greatest importance to 'nearly neutral' variants. However, we cannot fully disentangle the influences of these two processes. We suggest that existing variants under positive selection will increase in frequency quickly, most likely regardless of their initial frequency, while variants under purifying selection will not long remain common even if disproportionately over represented in a bottleneck population. However, since the number of variants under positive selection is few [21] the bulk of virus evolutionary change could be argued to comprise these 'nearly neutral' variants.

The method of *Sobel et al*. [31] is tested in their original paper using computer simulations and found to perform better than pre-existing alternative approaches. Their beta-binomial method accounts for likely false-negative variants not called as a present due to variant calling thresholds, and for changes in variant frequencies arising from stochastic replication dynamics early in infection. Their method, however, makes a number of assumptions that while almost certainly violated, may have only a modest influence on bottleneck size estimation. First (and perhaps most importantly) that variants are independent of each other, second, that donor-identified variants do not originate in any recipient hosts de novo, third, that variants are biallelic and fourth, that variants are not under positive or purifying selection. In our *in vivo* transmission study, it is not possible to know the true parent–daughter relationships between intra-host or inter-host samples due to the complexities of the biology. We, therefore, used the method of Sobel *et al*. [31] to evaluate all temporally possible parent–daughter relationships between relevant samples and selected the one with the largest bottleneck (i.e. the most number of viral genomes passing between the two) as the most likely parent–daughter relationship.

Following infection, FMDV establishes primary replication sites in the nasopharyngeal mucosa [15,64], leading to viremia, systemic dissemination of infection and the lesions characteristic of clinical disease. Bottleneck sizes between populations in the nasopharyngeal mucosa and serum, and serum and lesion sites are probably more indicative of a more continuous level of connectivity than discrete population founding events. Nonetheless, this analysis shows this connectivity to be characterized by narrow bottlenecks, while the connectivity indicated by a sequential daily sampling of serum is much higher and indicative of much better mixed population. In these experiments, probang samples collected in the acute phase of FMD (1–5 days post-contact) were used to define FMDVs present at the primary site of replication. However, we recognize that this sample type does not necessarily represent a precise anatomical site and in later stages of infection, probang samples may contain heterogeneous populations derived from primary sites of replication, as well as viruses derived from secondary sites (including those seeded via FMDV in the blood). Bottleneck sizes associated with between-host transmission seem to be variable and dependent on the nature of transmission. While the sample size is low, our data suggest that one 'one-time' infection typical of airborne aerosol spread would be characterized by narrow bottlenecks. During the UK 2001 outbreak, only one sample was taken from each herd (the individual with the oldest looking lesions); it is, therefore, possible that this individual was not directly infected from the farm suspected to be the source of infection. Were this to be the case, we consider it likely that the bottleneck size would be underestimated.

Although wide bottlenecks of 200–250 viral genomes were reported for IAV, re-analysis has suggested there was an erroneous mixing of sample reads [33]. Other studies have reported relatively narrow inter-host bottlenecks of 1–3 viral genomes for IAV [25], hepatitis C virus [35,36] and HIV [37]. Our results suggest FMDV intra-host bottlenecks can be an order of magnitude higher at both the inter- and intra-host level.

Data accessibility. Sequence FASTQ files from the *in vitro* bottleneck samples, as well as the 2001 inter-host outbreak samples, have been submitted to the European Nucleotide Archive (ENA) under project accession number PRJEB34979. Previously published data from the

experimental transmission chain are available under project number PRJEB3331 at ENA.

Authors' contributions. R.J.O. undertook the bioinformatic and bottleneck computations. C.F.W. conducted all the cell-culture experimental work supervised by D.T.H. and D.P.K. D.T.H. contributed the mathematical modelling. The paper was drafted and edited by all authors.

Competing interests. We declare we have no competing interests.

Funding. Work undertaken at The Pirbright Institute was supported by the Biotechnology and Biological Sciences Research Council of the United Kingdom (projects BBS/E/I/00007035 and BBS/E/I/00007036). Work undertaken at the University of Glasgow was supported by the Biotechnology and Biological Sciences Research Council, UK via a DTA PhD studentship (project BB/E018505/1) and BBSRC standard grant (projects BB/I013784/1 and BB/I014314/1). R.J.O. is currently funded by the Medical Research Council (MC_UU_12014/12).

Acknowledgements. We thank Katia Koelle (Emory University) and Tyler Smith for providing their R code for estimating bottleneck sizes from HTS data, Veronica Fowler (Pirbright Institute) for providing FMDV $O_1$Kaufbeuren cDNA, Valerie Mioulet for providing the FMDVs from the 2001 epidemic, Marco Morelli for useful discussions and reviewers for particularly helpful observations.

## References

1. Zell R *et al.* 2017 ICTV virus taxonomy profile: picornaviridae. *J. Gen. Virol.* **98**, 2421–2422. (doi:10.1099/jgv.0.000911)
2. Batschelet E, Domingo E, Weissmann C. 1976 The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* **1**, 27–32. (doi:10.1016/0378-1119(76)90004-4)
3. Drake JW. 1993 Rates of spontaneous mutation among RNA viruses. *Proc. Natl Acad. Sci. USA* **90**, 4171–4175. (doi:10.1073/pnas.90.9.4171)
4. Domingo E, Holland JJ. 1997 RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**, 151–178. (doi:10.1146/annurev.micro.51.1.151)
5. Drake JW, Holland JJ. 1999 Mutation rates among RNA viruses. *Proc. Natl Acad. Sci. USA* **96**, 13 910–13 913. (doi:10.1073/pnas.96.24.13910)
6. Wright CF, Morelli MJ, Thebaud G, Knowles NJ, Herzyk P, Paton DJ, Haydon DT, King DP. 2011 Beyond the consensus: dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing. *J. Virol.* **85**, 2266–2275. (doi:10.1128/JVI.01396-10)
7. Morelli MJ, Wright CF, Knowles NJ, Juleff N, Paton DJ, King DP, Haydon DT. 2013 Evolution of foot-and-mouth disease virus intra-sample sequence diversity during serial transmission in bovine hosts. *Vet. Res.* **44**, 12. (doi:10.1186/1297-9716-44-12)
8. King DJ, Freimanis GL, Orton RJ, Waters RA, Haydon DT, King DP. 2016 Investigating intra-host and intra-herd sequence diversity of foot-and-mouth disease virus. *Infect. Genet. Evol.* **44**, 286–292. (doi:10.1016/j.meegid.2016.07.010)
9. Jackson T, Clark S, Berryman S, Burman A, Cambier S, Mu D, Nishimura S, King AMQ. 2004 Integrin alphavbeta8 functions as a receptor for foot-and-mouth disease virus: role of the beta-chain cytodomain in integrin-mediated infection. *J. Virol.* **78**, 4533–4540. (doi:10.1128/JVI.78.9.4533-4540.2004)
10. Monaghan P *et al.* 2005 The alpha(v)beta6 integrin receptor for foot-and-mouth disease virus is expressed constitutively on the epithelial cells targeted in cattle. *J. Gen. Virol.* **86**(Pt 10), 2769–2780. (doi:10.1099/vir.0.81172-0)
11. O'Donnell V, Pacheco JM, Gregg D, Baxt B. 2009 Analysis of foot-and-mouth disease virus integrin receptor expression in tissues from naive and infected cattle. *J. Comp. Pathol.* **141**, 98–112. (doi:10.1016/j.jcpa.2008.09.008)
12. Jackson T *et al.* 1996 Efficient infection of cells in culture by type O foot-and-mouth disease virus requires binding to cell surface heparan sulfate. *J. Virol.* **70**, 5282–5287.
13. Alexandersen S, Oleksiewicz MB, Donaldson AI. 2001 The early pathogenesis of foot-and-mouth disease in pigs infected by contact: a quantitative time-course study using TaqMan RT-PCR. *J. Gen. Virol.* **82**(Pt 4), 747–755. (doi:10.1099/0022-1317-82-4-747)
14. Arzt J, Juleff N, Zhang Z, Rodriguez LL. 2011 The pathogenesis of foot-and-mouth disease I: viral pathways in cattle. *Transbound. Emerg. Dis.* **58**, 291–304. (doi:10.1111/j.1865-1682.2011.01204.x)
15. Stenfeldt C, Eschbaumer M, Pacheco JM, Rekant SI, Rodriguez LL, Arzt J. 2015 Pathogenesis of primary foot-and-mouth disease virus infection in the nasopharynx of vaccinated and non-vaccinated cattle. *PLoS ONE* **10**, e0143666. (doi:10.1371/journal.pone.0143666)
16. Burrows R, Mann JA, Garland AJ, Greig A, Goodridge D. 1981 The pathogenesis of natural and simulated natural foot-and-mouth disease infection in cattle. *J. Comp. Pathol.* **91**, 599–609. (doi:10.1016/0021-9975(81)90089-X)
17. Alexandersen S, Zhang Z, Reid SM, Hutchings GH, Donaldson AI. 2002 Quantities of infectious virus and viral RNA recovered from sheep and cattle experimentally infected with foot-and-mouth disease virus O UK 2001. *J. Gen. Virol.* **83**(Pt 8), 1915–1923. (doi:10.1099/0022-1317-83-8-1915)
18. Alexandersen S, Zhang Z, Donaldson AI, Garland AJ. 2003 The pathogenesis and diagnosis of foot-and-mouth disease. *J. Comp. Pathol.* **129**, 1–36. (doi:10.1016/S0021-9975(03)00041-0)
19. Stenfeldt C, Hartwig EJ, Smoliga GR, Palinski R, Silva EB, Bertram MR, Fish IH, Pauszek SJ, Arzt J. 2018 Contact challenge of cattle with foot-and-mouth disease virus validates the role of the nasopharyngeal epithelium as the site of primary and persistent infection. *mSphere* **3**, e00493-18. (doi:10.1128/mSphere.00493-18)
20. Charleston B *et al.* 2011 Relationship between clinical signs and transmission of an infectious disease and the implications for control. *Science* **332**, 726–729. (doi:10.1126/science.1199884)
21. Haydon DT, Bastos AD, Knowles NJ, Samuel AR. 2001 Evidence for positive selection in foot-and-mouth disease virus capsid genes from field isolates. *Genetics* **157**, 7–15.
22. Orton RJ *et al.* 2013 Observing micro-evolutionary processes of viral populations at multiple scales. *Phil. Trans. R. Soc. B* **368**, 20120203. (doi:10.1098/rstb.2012.0203)
23. Alizon S, Luciani F, Regoes RR. 2011 Epidemiological and clinical consequences of within-host evolution. *Trends Microbiol.* **19**, 24–32. (doi:10.1016/j.tim.2010.09.005)
24. Zwart MP, Elena SF. 2015 Matters of size: genetic bottlenecks in virus infection and their potential impact on evolution. *Annu. Rev. Virol.* **2**, 161–179. (doi:10.1146/annurev-virology-100114-055135)
25. McCrone JT, Woods RJ, Martin ET, Malosh RE, Monto AS, Lauring AS. 2018 Stochastic processes constrain the within and between host evolution of influenza virus. *eLife* **7**, e35962. (doi:10.7554/eLife.35962)
26. Geoghegan JL, Senior AM, Holmes EC. 2016 Pathogen population bottlenecks and adaptive landscapes: overcoming the barriers to disease emergence. *Proc. R. Soc. B* **283,** 20160727. (doi:10.1098/rspb.2016.0727)
27. Varble A, Albrecht RA, Backes S, Crumiller M, Bouvier NM, Sachs D, GarcãA-Sastre A, Tenoever B. 2014 Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell Host Microbe* **16**, 691–700. (doi:10.1016/j.chom.2014.09.020)
28. Wilker PR *et al.* 2013 Selection on haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1 influenza viruses. *Nat. Commun.* **4**, 2636. (doi:10.1038/ncomms3636)
29. Hughes J *et al.* 2012 Transmission of equine influenza virus during an outbreak is characterized by frequent mixed infections and loose transmission bottlenecks. *PLoS Pathog.* **8**, e1003081. (doi:10.1371/journal.ppat.1003081)
30. Murcia PR *et al.* 2010 Intra- and interhost evolutionary dynamics of equine influenza virus. *J. Virol.* **84**, 6943–6954. (doi:10.1128/JVI.00112-10)
31. Sobel LA, Weissman DB, Greenbaum B, Ghedin E, Koelle K. 2017 Transmission bottleneck size estimation from pathogen deep-sequencing data, with an application to human influenza A virus. *J. Virol.* **91**, e00171-17. (doi:10.1128/JVI.00171-17)

32. Poon LL *et al.* 2016 Quantifying influenza virus diversity and transmission in humans. *Nat. Genet.* **48**, 195–200. (doi:10.1038/ng.3479)
33. Xue KS, Bloom JD. 2019 Reconciling disparate estimates of viral genetic diversity during human influenza infections. *Nat. Genet.* **51**, 1298–1301. (doi:10.1038/s41588-019-0349-3)
34. Bull RA, Eden JS, Luciani F, McElroy K, Rawlinson WD, White PA. 2012 Contribution of intra- and interhost dynamics to norovirus evolution. *J. Virol.* **86**, 3219–3229. (doi:10.1128/JVI.06712-11)
35. Wang GP, Sherrill-Mix SA, Chang KM, Quince C, Bushman FD. 2010 Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J. Virol.* **84**, 6218–6228. (doi:10.1128/JVI.02271-09)
36. Bull RA *et al.* 2011 Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. *PLoS Pathog.* **7**, e1002243. (doi:10.1371/journal.ppat.1002243)
37. Fischer W *et al.* 2010 Transmission of single HIV-1 genomes and dynamics of early immune escape revealed by ultra-deep sequencing. *PLoS ONE* **5**, e12303. (doi:10.1371/journal.pone.0012303)
38. Carrillo C, Lu Z, Borca MV, Vagnozzi A, Kutish GF, Rock DL. 2007 Genetic and phenotypic variation of foot-and-mouth disease virus during serial passages in a natural host. *J. Virol.* **81**, 11 341–11 351. (doi:10.1128/JVI.00930-07)
39. Escarmis C, Lazaro E, Arias A, Domingo E. 2008 Repeated bottleneck transfers can lead to non-cytocidal forms of a cytopathic virus: implications for viral extinction. *J. Mol. Biol.* **376**, 367–379. (doi:10.1016/j.jmb.2007.11.042)
40. Duarte E, Clarke D, Moya A, Domingo E, Holland J. 1992 Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc. Natl Acad. Sci. USA* **89**, 6015–6019. (doi:10.1073/pnas.89.13.6015)
41. Novella IS, Elena SF, Moya A, Domingo E, Holland JJ. 1995 Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *J. Virol.* **69**, 2869–2872.
42. Escarmis C, Davila M, Charpentier N, Bracho A, Moya A, Domingo E. 1996 Genetic lesions associated with Muller's ratchet in an RNA virus. *J. Mol. Biol.* **264**, 255–267. (doi:10.1006/jmbi.1996.0639)
43. Yuste E, Sanchez-Palomino S, Casado C, Domingo E, Lopez-Galindez C. 1999 Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. *J. Virol.* **73**, 2745–2751.
44. Li H, Roossinck MJ. 2004 Genetic bottlenecks reduce population variation in an experimental RNA virus population. *J. Virol.* **78**, 10 582–10 587. (doi:10.1128/JVI.78.19.10582-10587.2004)
45. Ali A, Li H, Schneider WL, Sherman DJ, Gray S, Smith D, Roossinck MJ. 2006 Analysis of genetic bottlenecks during horizontal transmission of *cucumber mosaic virus*. *J. Virol.* **80**, 8345–8350. (doi:10.1128/JVI.00568-06)
46. Jridi C, Martin JF, Marie-Jeanne V, Labonne G, Blanc S. 2006 Distinct viral populations differentiate and evolve independently in a single perennial host plant. *J. Virol.* **80**, 2349–2357. (doi:10.1128/JVI.80.5.2349-2357.2006)
47. Boeras DI *et al.* 2011 Role of donor genital tract HIV-1 diversity in the transmission bottleneck. *Proc. Natl Acad. Sci. USA* **108**, E1156–E1163. (doi:10.1073/pnas.1103764108)
48. Novella IS, Duarte EA, Elena SF, Moya A, Domingo E, Holland JJ. 1995 Exponential increases of RNA virus fitness during large population transmissions. *Proc. Natl Acad. Sci. USA* **92**, 5841–5844. (doi:10.1073/pnas.92.13.5841)
49. Escarmis C, Davila M, Domingo E. 1999 Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. *J. Mol. Biol.* **285**, 495–505. (doi:10.1006/jmbi.1998.2366)
50. Novella IS, Dutta RN, Wilke CO. 2008 A linear relationship between fitness and the logarithm of the critical bottleneck size in vesicular stomatitis virus populations. *J. Virol.* **82**, 12 589–12 590. (doi:10.1128/JVI.01394-08)
51. Juleff N, Valdazo-Gonzalez B, Wadsworth J, Wright CF, Charleston B, Paton DJ, King DP, Knowles NJ. 2013 Accumulation of nucleotide substitutions occurring during experimental transmission of foot-and-mouth disease virus. *J. Gen. Virol.* **94**(Pt 1), 108–119. (doi:10.1099/vir.0.046029-0)
52. Cottam EM *et al.* 2008 Transmission pathways of foot-and-mouth disease virus in the United Kingdom in 2007. *PLoS Pathog.* **4**, e1000050. (doi:10.1371/journal.ppat.1000050)
53. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760. (doi:10.1093/bioinformatics/btp324)
54. Orton RJ, Wright CF, Morelli MJ, King DJ, Paton DJ, King DP, Haydon DT. 2015 Distinguishing low frequency mutations from RT-PCR and sequence errors in viral deep sequencing data. *BMC Genomics* **16**, 229. (doi:10.1186/s12864-015-1456-x)
55. Morelli MJ, Thebaud G, Chadoeuf J, King DP, Haydon DT, Soubeyrand S. 2012 A Bayesian inference framework to reconstruct transmission trees using epidemiological and genetic data. *PLoS Comput. Biol.* **8**, e1002768. (doi:10.1371/journal.pcbi.1002768)
56. Ellard FM, Drew J, Blakemore WE, Stuart DI, King AM. 1999 Evidence for the role of His-142 of protein 1C in the acid-induced disassembly of foot-and-mouth disease virus capsids. *J. Gen. Virol.* **80**(Pt 8), 1911–1918. (doi:10.1099/0022-1317-80-8-1911)
57. Botner A, Kakker NK, Barbezange C, Berryman S, Jackson T, Belsham GJ. 2011 Capsid proteins from field strains of foot-and-mouth disease virus confer a pathogenic phenotype in cattle on an attenuated, cell-culture-adapted virus. *J. Gen. Virol.* **92**(Pt 5), 1141–1151. (doi:10.1099/vir.0.029710-0)
58. Brehm KE, Ferris NP, Lenk M, Riebe R, Haas B. 2009 Highly sensitive fetal goat tongue cell line for detection and isolation of foot-and-mouth disease virus. *J. Clin. Microbiol.* **47**, 3156–3160. (doi:10.1128/JCM.00510-09)
59. Shaw AE, Reid SM, Ebert K, Hutchings GH, Ferris NP, King DP. 2007 Implementation of a one-step real-time RT-PCR protocol for diagnosis of foot-and-mouth disease. *J. Virol. Methods* **143**, 81–85. (doi:10.1016/j.jviromet.2007.02.009)
60. Fry EE *et al.* 1999 The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex. *EMBO J.* **18**, 543–554. (doi:10.1093/emboj/18.3.543)
61. Gerrish PJ, Lenski RE. 1998 The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**, 127–144. (doi:10.1023/A:1017067816551)
62. Miralles R, Gerrish PJ, Moya A, Elena SF. 1999 Clonal interference and the evolution of RNA viruses. *Science* **285**, 1745–1747. (doi:10.1126/science.285.5434.1745)
63. Mayr E. 1942 *Systematics and the origin of species*. New York, NY: Columbia University Press.
64. Arzt J, Pacheco JM, Rodriguez LL. 2010 The early pathogenesis of foot-and-mouth disease in cattle after aerosol inoculation. Identification of the nasopharynx as the primary site of infection. *Vet. Pathol.* **47**, 1048–1063. (doi:10.1177/0300985810372509)