CELL TROPISM PREDICTS LONG-TERM NUCLEOTIDE SUBSTITUTION RATES OF MAMMALIAN RNA VIRUSES

By

ALLISON HICKS

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Microbial Biology

written under the direction of

Dr. Siobain Duffy

and approved by

_____

_____

_____

New Brunswick, New Jersey

October, 2013

ABSTRACT OF THE THESIS

CELL TROPISM PREDICTS LONG-TERM NUCLEOTIDE SUBSTITUTION RATES OF MAMMALIAN RNA VIRUSES

by ALLISON HICKS

Thesis Director: Dr. Siobain Duffy

The high rates of RNA virus evolution are generally attributed to replication with error-prone RNA-dependent RNA polymerases. However, these long-term nucleotide substitution rates span three orders of magnitude and do not correlate well with mutation rates or selection pressures. This substitution rate variation may be explained by differences in virus ecology or intrinsic genomic properties. We generated long-term nucleotide substitution rate estimates for mammalian RNA viruses and compiled comparable published rates, yielding a dataset of 118 substitution rates of structural genes from 51 different species, as well as 40 rates of non-structural genes from 28 species. Through multiple regression analyses, we evaluated the relationships between these rates and four ecological factors: target cell, transmission route, host range, infection duration; and three genomic properties: genome length, genome sense, genome segmentation**.** Of these seven factors, we found target cells to be the only significant predictors of viral substitution rates, with tropisms for epithelial cells ($P<2x10^{-5}$ for the structural genes) or neurons ($P<3x10^{-7}$ and $P<0.01$ for the structural genes and non-structural genes, respectively) as the most significant predictors. Further, one-tailed t-tests showed that viruses primarily infecting epithelial cells evolve significantly faster than neurotropic viruses ($P=1.83x10^{-10}$ and $P=6.30x10^{-4}$ for

the structural genes and non-structural genes, respectively). These results provide strong evidence that the fastest evolving mammalian RNA viruses infect cells with the highest turnover rates: the highly proliferative epithelial cells. Estimated viral generation times suggest that epithelial-infecting viruses replicate more quickly than viruses with different cell tropisms. Our results indicate that cell tropism is a key factor in viral evolvability.

ACKNOWLEDGEMENTS

I would like to thank my principal investigator, co-author, and mentor, Dr. Siobain Duffy, for all of her support and guidance throughout this project. Dr. Duffy has played an invaluable role in my development as a student and scientist.

I would also like to thank Yee Mey Seah, Preshita Gadkari, Dan Cardinale, Jenn McConnell, Eric Ho, Lily Li, and Shannon Loelius, as well as Dr. Bradley Hillman and Dr. Peter Kahn for their support.

Finally, I would like to thank the National Institute of Allergy and Infectious Diseases at the National Institutes of Health (grant number 1R03AI096265-01) and the Rutgers School of Environmental and Biological Sciences for funding this study.

Table of Contents

LIST OF TABLES

1. INTRODUCTION

1.1. Public Health and the Evolutionary Dynamics of RNA Viruses

RNA viruses are responsible for a disproportionate number of emerging human diseases, including influenza, ebola hemorrhagic fever, hantavirus pulmonary syndrome, Middle East respiratory syndrome, and rotavirus-associated diarrhea, which place tremendous health and economic burdens on both the developing and developed world [1,2]. In 2008, rotavirus and measles virus caused the deaths of 570,000 children under the age of five, making them two of the leading killers of children worldwide [3]. In 2009, it was estimated that rotavirus infections alone result in $325 million in medical treatment costs and $423 million in societal costs each year [4]. Further, the implementation of many intervention strategies has either failed or been delayed as a result of the evolutionary dynamics of these pathogens [1,5,6,7,8,9].

1.2. Variation Among Evolutionary Rates of RNA Viruses

Differences in viral evolutionary dynamics, such as rates of evolution, can explain why certain viruses have the capacity to adapt to new host species, increase in virulence, or develop resistance to antivirals [7,8,9,10,11]. Therefore, understanding why some RNA viruses evolve more quickly can facilitate better prediction of their pathogenic and epidemiological potential [8,10,11,12]. Though extremely high nucleotide substitution rates are a defining feature of RNA virus

evolution [1,13,14,15], there have been few attempts to comprehensively examine the driving genomic and ecological factors behind these rates.

Differences in the strength and direction of selection pressures on these viruses result in variation among their substitution rates [1,5,13]. However, while some general patterns have been observed in selection pressures, such as enhanced purifying selection on the structural proteins of arboviruses [16], there have been no attempts to quantify the relationship between selection pressures and long-term viral substitution rates.

The high rates of RNA virus evolution are often attributed to their replication with error-prone RNA-dependent RNA polymerases (RdRps) [1,17], but these nucleotide substitution rates are known to span at least three orders of magnitude [5,17] and do not correlate well with experimentally measured viral mutation rates [5]. Further, the substitution rates of some DNA viruses, which replicate with high-fidelity DNA polymerases, are comparable to the high substitution rates of RNA viruses [13]. Therefore, the polymerase error rate cannot explain the substitution rate variation in RNA viruses.

Along with mutation rate, viral replication frequency directly impacts the rate at which mutations can be introduced, and ultimately fixed as substitutions [13]. Replication frequencies could be influenced by a variety of factors related to viral genomic architecture or ecology [13]. For example, weak negative correlations between viral genome lengths and substitution rates have been attributed to enhanced replication frequencies of viruses with smaller genomes [15,17,18]. It has also been suggested that different transmission and infection

modes result in differences in generation time, ultimately causing variation among per-year rates of synonymous substitution of RNA virus structural genes [5].

1.3. Objectives

In this modern survey of mammalian RNA virus evolution rates, we generated and compiled published substitution rates of structural and non-structural genes produced by Bayesian coalescent analyses [19]. We analyzed these rates as a function of seven factors related to virus genomic architecture (*i.e.*, genome length, genome sense, and whether or not the genome is segmented) and virus ecology (*i.e.*, target cell, transmission mode, host range, and whether the infection is acute or persistent). We also evaluated the relationships of viral substitution rates with estimated ratios of nonsynonymous to synonymous evolutionary changes (dN/dS), experimentally measured mutation rates, and estimated generation times. Though recombination undeniably plays a role in shaping viral evolutionary dynamics and could inflate substitution rate estimates [20,21], we conservatively removed any potential recombinants from our datasets prior to analysis. Through this broad meta-analysis, we were able to demonstrate that cell tropism, and its impact on viral generation time, has the greatest influence on rates of mammalian RNA virus evolution.

2. MATERIALS AND METHODS

2.1. Published Nucleotide Substitution Rates

Long-term nucleotide substitution rates of mammalian RNA viruses were collected from the literature, with a focus on finding rates for the outer structural gene containing the major antigenic site(s) and non-structural (preferably the RdRp) genes. While the RdRp genes of the (-)ssRNA and dsRNA viruses are classified as structural, or virion-associated, genes [22], they are generally thought to be more conserved and under very different selection pressures than the structural genes that interact with the host immune system [23,24]. We excluded retroviruses from analysis because they are known to have highly variable substitution rates due to time spent integrated into DNA genomes, where they evolve at the rate of their hosts' genome [13,25]. Viruses that predominately infect non-mammals, with mammals serving as incidental, dead-end hosts, were also excluded. Only rates estimated for individual viral species or strains were used, not those that aggregated multiple species into one analysis. Similarly, only rates from single gene analyses were included, not those based on full genomes or multiple gene alignments. In order to minimize any rate discrepancies that could result from variations among datasets (*e.g.*, number of taxa, temporal range, portion of gene analyzed) and/or subtle methodological variations [26,27,28,29,30,31], only rates produced by Bayesian coalescent analyses of datasets composed of at least 30 taxa, isolated over a minimum range of 15 years and spanning at least 40% of the analyzed gene were included. Data

regarding genomic architecture and ecology were obtained for all viruses with published substitution rates that met these criteria. We included multiple rates for a given virus when available, except when a single study examined multiple lineages and summarized the results in a single rate [32,33,34,35]. Corresponding dN/dS estimates were collected when available.

2.2. Sequence Data for Novel Rate Analyses

These published substitution rates were supplemented with novel BEAST [19] rate analyses based on the sequence data available in GenBank (accessed through Taxonomy Browser, http://www.ncbi.nlm.nih.gov/Taxonomy). Sequences for structural and non-structural genes with years of isolation available in GenBank or the literature were manually aligned using Se-Al v2.0a11 [36]. Sequences with GenBank or published information that indicated they were genetically manipulated or extensively passaged in the lab prior to sequencing were eliminated from further analysis. The final datasets also adhered to the conservative criteria described for published datasets, above.

2.3. Substitution Rate and Selection Analyses

As recombination events can lead to over-estimation of nucleotide substitution rates, each dataset was scanned for recombination using seven different algorithms (RDP, GENECONV, Bootscan, MaxChi, Chimaera, SiScan, and 3seq) implemented in RDP v3.44 [37]. Sequences implicated as recombinant by two or more algorithms were excluded from further analysis.

Modeltest v3.7 [38] was used to determine the best-fit model of nucleotide substitution for each dataset (by AIC).

Long-term nucleotide substitution rates were estimated using BEAST v1.5.4 [19]. Each dataset was run for at least 50 million generations and until all parameters had stabilized (effective sampling size > 200). Each dataset was run with two different clock models (strict and uncorrelated lognormal) and three different demographic models (constant, exponential, and Bayesian skyline). The best-fitting clock/demographic model combination for each dataset was determined using Bayes factors as implemented in Tracer v1.5 [39]. For each best set of priors, two independent runs were performed to ensure that the results were replicable, and a control analysis was run without the dataset to ensure that the priors were not controlling the outcome of the analysis.

The Single Likelihood Ancestor Counting (SLAC), codon-based maximum likelihood method available in the HYPHY package on the Datamonkey web server [40] was used to evaluate the strength of selection pressure on these datasets.

2.4. Statistical Analyses

In order to determine which factors most significantly predict substitution rates of mammalian RNA viruses, multiple regression analyses were run using SPSS Statistics v21 (IBM) with log-transformed mean substitution rates as the dependent variable and seven overarching predictor variables (target cell, transmission route, whether the infection is acute or persistent, host range,

genome length, genome sense, and whether or not the genome is segmented). Categorical predictor variables (*i.e.*, target cell, transmission route, host range, infection mode, genome sense, and whether or not the genome is segmented) were dummy coded into *n-1* dichotomous variables to account for *n* levels. For each variable, different base levels were tested to ensure that the chosen base level did not significantly influence the results. Collinearity among the variables was also assessed, with variance inflation factors (VIF) greater than 10 indicating redundancy among variables. Separate multiple regression analyses were run on the structural and non-structural gene datasets. As there were multiple published rates for some viral species and strains, additional analyses were run for each dataset with only one substitution rate per virus species, using the average rate in the case of multiple rates for a given virus species.

One-tailed t-tests were subsequently run in R v2.14.1 [41] to provide an additional measure of significant directional variation among the log-transformed mean rates of different levels for any categorical variable that was found to be a significant rate predictor ($P<0.01$) in the multiple regression analyses. Additional t-tests were also conducted using the control datasets with one rate per virus species.

Additionally, though there were no dN/dS or mutation rate estimates available for all viruses used in this study, the available data for each variable were compared to corresponding log-transformed mean substitution rate estimates using Spearman rank correlation (for dN/dS) or Pearson correlation coefficient (for mutation rates). Structural and non-structural gene rate estimates

were also compared using Pearson correlation coefficient. All correlation analyses were performed in SPSS Statistics v21.

3. RESULTS AND DISCUSSION

3.1. Datasets

A literature review yielded 92 published Bayesian nucleotide substitution rate estimates for the structural genes of 35 different mammalian RNA viral species, and 21 published Bayesian rates for RdRps or a non-structural gene of 14 different viral species (referred to collectively as “non-structural,” Appendix 1). These rates were supplemented with 26 novel Bayesian substitution rates of structural genes of 19 different viral species, and 19 novel Bayesian rates of non-structural genes of 16 different viral species (Appendix 2). Collectively, these rates span three orders of magnitude, ranging from $3.0x10^{-5}$ to $1.5x10^{-2}$ nucleotide substitutions per site per year (ns/s/y) and $2.0x10^{-5}$ to $1.3x10^{-2}$ ns/s/y for the structural genes and non-structural genes, respectively (Appendix 1).

Plotting the levels of each variable by increasing mean substitution rate revealed similar patterns for both the structural (S) and non-structural (NS) datasets in three of these variables, excepting transmission route. Viral substitution rates sorted by different target cells (panels 1A and 1B), transmission routes (panels 1C and 1D), infection type (panels 1E and 1F), and host ranges (panels 1G and 1H) are shown in Figure 1.

Substitution rates were also sorted by viral genomic architecture (genome sense/strandedness, Figure 2A and 2B, and genome segmentation, Figure 2C and 2D) and plotted against viral genome length (Figure 2E and 2F). In the case of segmented genomes, the lengths of the individual segments were added

together. There were no apparent relationships between genomic properties and substitution rates (Figure 2), including no linear relationship between substitution rates and genome lengths in either dataset (coefficient of determination, S: R^2=0.06, NS: R^2=0.08). Examining viruses with segmented and non-segmented genomes separately also revealed no linear relationship between substitution rate and genome length (S: R^2=0.15 for non-segmented only and R^2=0.06 for segmented only, NS: R^2=0.29 for non-segmented only, and R^2=0.00 for segmented only).

dN/dS estimates calculated in this study were compiled with published estimates also calculated using the Single Likelihood Ancestor Counting (SLAC) method (56 structural gene dN/dS estimates, 33 non-structural gene dN/dS estimates total, Appendix 1).

3.2. Statistical Analyses

Multiple regression analyses were performed separately on the S and NS gene datasets to determine which, if any, of seven factors (target cell, transmission route, infection mode, host range, genome length, genome sense, and genome segmentation) significantly predict the long-term nucleotide substitution rates of mammalian RNA viruses. To explore the many dummy-coded categorical variables, three analyses were run using different variable levels as the base levels (Tables 1 and 2). For all of the regression analyses, the adjusted coefficient of determination ($\overline{R}^2$) was $\geq$ 0.73, indicating that over 70% of the substitution rate variability can be explained by the predictor variables

included in this study. Standardized residual plots identified only six potential outliers of the 118 structural gene rates and one potential outlier of the 40 non-structural gene rates (data not shown), indicating that the data are normally distributed and therefore amenable to a linear regression model.

Table 1: Significant predictors of viral structural gene substitution rates. For each multiple regression analysis, the overall adjusted R^2 ($\overline{R}^2$) of the model is given along with significant predictor variables (P<0.01) and their standardized coefficients (β) with 95% confidence intervals (CIs). In the first regression, the base levels were epithelial target cells, fecal-oral/respiratory transmission route, acute/persistent infection, species-specific host range, and dsRNA genome architecture. In the second regression, the base levels were neural target cells, bites/scratches transmission route, persistent infection, order-specific host range, and (-)ssRNA genome architecture. In the third regression, the base levels were leukocyte target cells, respiratory/vertical transmission route, acute infection, family-specific host range, and (+)ssRNA genome architecture.

	$\overline{R}^2$	Predictor	β (95% CI)	Significance
1	0.73	Neurons	-0.80 (-1.00, -0.59)	1.24×10^{-13}
		Leukocytes	-0.54 (-0.74, -0.33)	4.63×10^{-7}
		Hepatocytes	-0.24 (-0.40, -0.08)	4.73×10^{-4}
		Endothelial cells	-0.18 (-0.32, -0.05)	4.54×10^{-3}
2	0.74	Epithelial cells	1.05 (0.77, 1.33)	9.40×10^{-12}
		Leukocytes	0.54 (0.34, 0.80)	5.07×10^{-7}
3	0.77	Neurons	-0.93 (-0.53, -0.25)	2.15×10^{-7}
		Epithelial cells	0.59 (0.36, 0.82)	1.15×10^{-5}

Regardless of the base levels, target cells were the only significant predictors of log-transformed substitution rates for both structural and non-structural genes (Tables 1 and 2). Targeting epithelial cells or neurons was found to be the most significant predictor of structural gene rates in each analysis where these were not the base levels (P<2×10^{-5}, Table 1, Figure 3), while targeting neurons was found to be the only significant predictor of substitution rates for the smaller non-structural gene dataset (P<0.01, Table 2, Figure 3). Further, there was a high correlation between each viral species' estimated

structural gene substitution rate and its corresponding non-structural gene rate (33 viruses, Pearson r=0.87, P=3.11x10^{-11}). This suggests that if it were possible to calculate more non-structural rates, we would likely see results similar to those from the structural gene dataset.

Table 2: Significant predictors of viral non-structural gene substitution rates. For each multiple regression analysis, the overall adjusted R^2 ($\overline{R}^2$) of the model is given along with the significant predictor variable (P<0.01) and its standardized coefficients (β) with 95% confidence intervals (CIs). In the first regression, the base levels were epithelial target cells, fecal-oral/respiratory transmission route, acute/persistent infection, species-specific host range, and dsRNA genome architecture. In the second regression, the base levels were neural target cells, bites/scratches transmission route, acute infection, order-specific host range, and (-)ssRNA genome architecture. In the third regression, the base levels were leukocyte target cells, respiratory/vertical transmission route, acute infection, family-specific host range, and (+)ssRNA genome architecture.

	$\overline{R}^2$	Predictor	β (95% CI)	Significance
1	0.77	-	-	-
2	0.77	-	-	-
3	0.77	Neurons	-0.77 (-1.33, -0.20)	9.56x10^{-3}

To minimize any potential bias introduced by using multiple published rates for a single viral strain or species, we conducted control analyses using datasets with only one rate per species. We calculated the average log substitution rate when there were multiple rates for a species. These data also had a linear relationship (data not shown), but the $\overline{R}^2$ for these analyses were slightly lower than for the full datasets (S: $\overline{R}^2$≥0.64, NS: $\overline{R}^2$=0.76, Tables 3 and 4). These control results were consistent with those from the full dataset analyses: tropisms for epithelial cells or neurons were the most significant substitution rate predictors (Tables 3 and 4, Figure 4).

Table 3: Significant predictors of viral structural gene substitution rates using one rate per viral species. For each multiple regression analysis, the overall adjusted R^2 ($\overline{R}^2$) of the model is given along with significant predictor variables (P<0.01) and their standardized coefficients (β) with 95% confidence intervals (CIs). In the first regression, the base levels were epithelial target cells, fecal-oral/respiratory transmission route, acute/persistent infection, species-specific host range, and dsRNA genome architecture. In the second regression, the base levels were neural target cells, bites/scratches transmission route, persistent infection, order-specific host range, and (-)ssRNA genome architecture. In the third regression, the base levels were leukocyte target cells, respiratory/vertical transmission route, acute infection, family-specific host range, and (+)ssRNA genome architecture.

	$\overline{R}^2$	Predictor	β (95% CI)	Significance
1	0.65	Neurons	-0.85 (-1.18, -0.51)	2.57×10^{-5}
		Leukocytes	-0.51 (-0.84, -0.18)	3.57×10^{-3}
2	0.66	Epithelial cells	0.95 (0.58, 1.39)	1.61×10^{-5}
		Leukocytes	0.50 (0.20, 0.77)	1.47×10^{-3}
3	0.64	Neurons	-0.41 (-0.66, -0.15)	2.65×10^{-3}
		Epithelial cells	0.49 (0.15, 0.84)	6.50×10^{-3}

Table 4: Significant predictors of viral non-structural gene substitution rates using one rate per viral species. For each multiple regression analysis, the overall adjusted R^2 ($\overline{R}^2$) of the model is given along with significant predictor variables (P<0.01) and their standardized coefficients (β) with 95% confidence intervals (CIs). In the first regression, the base levels were epithelial target cells, fecal-oral/respiratory transmission route, acute/persistent infection, species-specific host range, and dsRNA genome architecture. No factors were significant in this analysis. In the second regression, the base levels were neural target cells, bites/scratches transmission route, acute infection, order-specific host range, and (-)ssRNA genome architecture. No factors were significant in this analysis. In the third regression, the base levels were leukocyte target cells, respiratory/vertical transmission route, acute infection, family-specific host range, and (+)ssRNA genome architecture.

	$\overline{R}^2$	Predictor	β (95% CI)	Significance
1	0.76	-	-	-
2	0.76	-	-	-
3	0.76	Neurons	-0.53 (-0.58, -0.09)	5.58×10^{-3}

To ensure that any substitution rate variability attributed to a given predictor variable was not significantly dependent on other predictor variables,

we examined collinearity in all datasets. With the exception of the persistent infection variable, which was nested with the endothelial target cell variable and thus excluded, the regression analyses for the structural gene rate datasets showed no significant collinearity (no variance inflation factors (VIF) were greater than 10). For the non-structural gene rate datasets, many different predictor variables had VIF>10. However, subsequent analyses where each individual variable was removed did not significantly reduce collinearity in these datasets (data not shown). Due to the consistent results between the structural and non-structural gene datasets, we concluded that correlations among independent variables did not significantly impact our results.

Since target cells were found to be the only significant predictors of substitution rates, a series of one-tailed t-tests was used to confirm which cell tropisms are associated with higher viral substitution rates than others. Viruses that target epithelial cells were found to have significantly higher structural gene substitution rates than viruses that target neurons, endothelial cells, or leukocytes (Table 5, $P<9\text{x}10^{-4}$, Figure 5). Similarly, viruses that target epithelial cells were found to have significantly higher non-structural gene substitution rates than viruses that target neurons, hepatocytes, or leukocytes (Table 6, $P<7\text{x}10^{-4}$, Figure 5). These results were recapitulated in the control datasets with one rate per viral species (Tables 7 and 8). It should be noted, however, that most of the viruses in this study that are classified as targeting leukocytes ultimately cause systemic infections and infect a wide variety of cell types.

Consequently, viruses in the leukocyte target cell category had the most rate variation of all the target cell categories (Figure 1).

Table 5: Structural gene substitution rate variation among viruses with different cell tropisms. The significance of viruses with each target cell in the left column having higher log-scale mean substitution rates than the viruses with each target cell in the top row is designated with a p-value from a one-tailed t-test. The threshold for statistical significance (P<0.01) was Bonferroni-corrected to account for multiple comparisons (P<1x10^{-3}). N=neurons, En=endothelial cells, L=leukocytes, H=hepatocytes, Ep=epithelial cells.

	N	En	L	H	Ep
N	-	0.97	1.00	1.00	1.00
En	0.03	-	0.98	1.00	1.00
L	**2.99x10^{-5}**	0.02	-	0.99	1.00
H	**9.54x10^{-6}**	0.001	0.006	-	0.98
Ep	**1.83x10^{-10}**	**8.05x10^{-4}**	**4.14x10^{-16}**	0.03	-

Table 6: Non-structural gene substitution rate variation among viruses with different cell tropisms. The significance of viruses with each target cell in the left column having higher log scale mean substitution rates than the viruses with each target cell in the top row is designated with a p-value from a one-tailed t-test. The threshold for statistical significance (P<0.01) was Bonferroni-corrected to account for multiple comparisons (P<2x10^{-3}). N=neurons, L=leukocytes, H=hepatocytes, Ep=epithelial cells.

	N	L	H	Ep
N	-	0.99	0.99	1.00
L	7.20x10^{-3}	-	0.56	1.00
H	8.77x10^{-3}	0.44	-	1.00
Ep	**6.30x10^{-4}**	**1.01x10^{-4}**	**1.09x10^{-4}**	-

Table 7: Structural gene substitution rate variation among viruses with different cell tropisms (control datasets). Based on the control datasets with one substitution rate per viral species. The significance of viruses with each target cell in the left column having higher log scale mean substitution rates than the viruses with each target cell in the top row is designated with a p-value from a one-tailed t-test. The threshold for statistical significance (P<0.01) was Bonferroni-corrected to account for multiple comparisons (P=1x10^{-3}). N=neurons, En=endothelial cells, L=leukocytes, H=hepatocytes, Ep=epithelial cells.

	N	En	L	H	Ep
N	-	0.98	1.00	1.00	1.00
En	0.02	-	0.98	0.98	1.00
L	**2.28x10^{-4}**	0.02	-	0.91	1.00
H	1.17x10^{-3}	0.02	0.09	-	0.92
Ep	**2.27x10^{-7}**	**7.16x10^{-5}**	**1.31x10^{-4}**	0.08	-

Table 8: Non-structural gene substitution rate variation among viruses with different cell tropisms (control datasets). Based on the control datasets with one substitution rate per viral species. The significance of viruses with each target cell in the left column having higher log scale mean substitution rates than the viruses with each target cell in the top row is designated with a p-value from a one-tailed t-test. The threshold for statistical significance (P<0.01) was Bonferroni-corrected to account for multiple comparisons (P<$2x10^{-3}$). N=neurons, L=leukocytes, H=hepatocytes, Ep=epithelial cells.

	N	L	H	Ep
N	-	0.99	0.99	1.00
L	$6.34x10^{-3}$	-	0.43	1.00
H	0.01	0.57	-	0.99
Ep	**$2.40x10^{-4}$**	**$1.82x10^{-3}$**	$5.21x10^{-3}$	-

We also tested for linear relationships between viral substitution rates and other evolutionary parameters for which only smaller subsets of our datasets could be analyzed. Reliable experimentally measured mutation rates estimated as mutations per base per infectious cycle were only available for four different viruses included in this study (poliovirus 1 [11,42,43], hepatitis C virus [44], influenza A virus [45,46,47], influenza B virus [45]). Mutation rates measured as mutations per base per strand replication were only available for three viruses included in this study (poliovirus 1 [48], measles virus [49,50], and influenza A virus [51]). These mutation rates were not significantly correlated with their corresponding substitution rate estimates (r=0.69, P=0.31 and r=-0.93, P=0.25, for mutation rates measured as mutations per base per infection and mutation rates measured as mutations per base per replication, respectively). Similarly, there were no significant correlations between the estimated substitution rates and dN/dS estimates (ρ=-0.02, P=0.88 and ρ=-0.07, P=0.68, for the limited structural gene and non-structural gene datasets, respectively).

Since epithelial cells and neurons have some of the highest and lowest turnover rates, respectively, of all mammalian cells [52,53,54,55], we sought to determine if there were any associations between host cell turnover rate and viral generation time. Using the model proposed by Sanjuán (2012) that relates the long-term substitution rate, K, to the mutation rate, μ, correcting for transient deleterious mutations, we were able to estimate generation times for the few viruses with reliable mutation rate estimates. This model, $K=a\mu e^{-b_\mu G}$, with $a=g\alpha\cong 0.27$, $b=(1-\alpha)/s_H\cong 3.744$ (G=genome length, g=generation time, s_H=harmonic mean of the selection coefficient) [15], confirmed that influenza A virus, influenza B virus, and poliovirus, which target epithelial cells, have substantially shorter generation times (<40 hours) than hepatitis C virus, which targets hepatocytes (>200 hours). Shorter average generation times lead to more rounds of replication per year, neatly explaining higher per-year substitution rates.

3.3. Ecological and Genomic Driving Forces Behind RNA Virus Evolution

A variety of intrinsic and ecological factors could plausibly alter the tempo of virus evolution by influencing the rate at which genetic diversity is generated, maintained, and fixed within viral populations. Others have focused on genomic properties as drivers of substitution rate variation [14,15,17,18], demonstrating a weak negative correlation between the genome lengths and substitution rates of RNA viruses [15,17] or suggesting that ssRNA viruses evolve faster than dsRNA viruses [15]. However, we did not find any significant relationship between

genomic properties and substitution rates (Figures 2 and 3). While some have conducted more limited studies on the influence of ecological factors [5,56], we performed a comprehensive analysis that revealed that cell tropism is a key factor in understanding mammalian RNA viral substitution rates.

It has been proposed that persistent viruses evolve more slowly than those that produce acute infections [1,5,15,57]. Unfortunately, with the exception of latent viruses, which are most commonly retro- or DNA viruses and thus not within the scope in this study, it can be difficult to classify viruses as acute or persistent. The duration of persistence can vary; most persistent viral infections begin with an acute phase and may occasionally be resolved after only this acute phase, and many viruses that predominantly result in acute infections occasionally persist [58,59]**.** By classifying the viruses in this study as accurately as possible, we found no significant association between infection mode and substitution rate. However, only three viruses in this study, all endothelial-infecting hantaviruses, were classified as strictly persistent. This causes the nesting of the persistent level with tropism for endothelial cells, and the persistent infection variable was therefore excluded from our analyses. Infection duration could be a factor explaining substitution rate variation across the Baltimore classifications of viruses, but there is no evidence that it affects mammalian RNA virus substitution rates.

Transmission mode and, less explicitly, host range are frequently invoked as determinants of viral substitution rates [5,60]. Specifically, plant or animal viruses that primarily rely on arthropod vectors for transmission, and therefore

obligately infect very diverse hosts, are thought to evolve more slowly than viruses with other transmission modes [5,60,61,62]. Surprisingly, we found no significant relationship between substitution rate and transmission mode or host range.

The seven genomic and ecological factors examined are not necessarily independent. For example, many arboviruses are neurotropic (Appendix 1). Therefore, the suggestion that vector-borne viruses tend to evolve more slowly is qualitatively consistent with our results. Previous studies have also indicated that phylogenetic relationships are predictive – that sister taxa have similar rates of evolution [5]. However, when we initially included virus families as predictor variables in our analyses, we had to discard them due to high collinearity. Once the virus families were discarded, there was no significant collinearity within the structural gene dataset, indicating that, of the factors considered in this study, cell tropism is unambiguously the best predictor of viral substitution rates. The smaller non-structural gene dataset, on the other hand, had significant collinearity among predictor variables that could not be resolved. The NS dataset also had only 1/3 of the taxa, inherently reducing its statistical power. It is not possible to expand the mammalian RNA virus NS dataset at this time; our novel rate analyses increased the number of reliable rates by 40% by exhaustively searching the available sequences in GenBank. Despite these limitations, target cells were still the only significant predictor variables. We consider this qualitative support for our more robust S dataset results.

Though previously unexplored, cell tropism could influence viral substitution rates by the same mechanisms that have been suggested for the other ecological factors described above [63]. Infection of different host cells could expose viruses to different selection pressures, which could influence the rates at which mutations are fixed as substitutions. Additionally, it is possible that cell tropism influences the rate at which genetic diversity is generated by affecting viral mutation rates or generation times.

3.4. Selection Pressures Do Not Predict Substitution Rates

Variation in strength and/or direction of selection has frequently been invoked as a determinant of viral substitution rates [12,13,20]. While positive selection can certainly result in variation among very short-term substitution rates, purifying selection tends to dominate over longer timescales [20,28,64,65]. However, variation is observed in the strength of purifying selection due to differences in host ranges. For instance, as previously mentioned, viruses vectored by arthropods have unique evolutionary constraints placed on them by their host diversity [60,61,62,66]. While previous studies found that arboviruses are under stronger purifying selection than non-arboviruses [1,60,67], we found that the dN/dS estimates based on structural genes of arboviruses were not significantly lower than those for non-arboviruses (P=0.19). The dN/dS estimates based on non-structural genes were only moderately lower than those for non-arboviruses (P=0.04). Further, we found no significant correlation between the estimated dN/dS and substitution rates, suggesting that detectable differences in

selection pressures do not explain the variation in long-term substitution rates of mammalian RNA viruses. To date, there are no data supporting a link between cell tropism and sustained differences in selection pressures.

3.5. Mutation and Substitution Rates are Uncorrelated

Compared to the slower evolution of DNA viruses, the evolution of RNA viruses is dominated by their high mutation rates [1,13,15]. Weak negative correlations between genome lengths and viral substitution rates have been attributed to a relationship between mutation rate and substitution rate, as smaller genomes could in theory withstand higher mutation rates than larger genomes [13,15,68]. However, while differences in spontaneous mutation rates appear to be significantly correlated to the long-term substitution rates of DNA viruses [15], this linear relationship disappears past a certain mutation rate threshold: around 10^{-6} mutations per site per infectious cycle, the lower end of the mutation rate range of RNA viruses [13,15]. It is, therefore, not surprising that we found no significant correlation between long-term substitution rates and the available, reliable mutation rate estimates. Additionally, a recent study of the retrovirus HIV-1 found that infection of different cell types did not lead to differences in mutation rate [69], providing some evidence that mutation rate is not correlated with cell tropism. Together, these data suggest that mutation rate variation among different cell types is not driving higher substitution rates in epithelial-infecting mammalian RNA viruses.

3.6. Cell Tropism May Influence Substitution Rates through Generation Time

Ruling out selection, mutation rates, and recombination frequencies as drivers of RNA virus substitution rates implies that the rate variation is largely the result of variation in replication dynamics [5,13]. Enhanced replication frequencies (shorter generation times) have been used to explain a variety of the previously suggested links between virus ecology and substitution rate. For example, viruses in the acute phase of an infection generally replicate more frequently than those in a persistent infection, and viruses in a latent phase do not replicate at all [58]. Further, as an alternative to differential selection pressures, the argument that transmission mode drives viral substitution rates assumes that viruses that can be transmitted more rapidly will have shorter generation times (*e.g.*, horizontal transmission vs. vertical transmission [5,70,71]).

DNA viruses have shorter generation times in faster dividing cells [72,73], but the associations between cell tropism and RNA virus generation time are less obvious, as RNA viruses do not depend on cellular replication machinery. However, there is evidence that some RNA viruses infecting rapidly dividing cells can replicate faster and more efficiently [74,75,76,77,78]. For example, it has been repeatedly demonstrated that hepatitis C virus produces a more robust infection in proliferating cells, perhaps due to enhanced levels of available nucleotides [77], or because of higher levels of viral protein synthesis facilitated by nuclear translation initiation factors that only become available in the

cytoplasm during cell division [76]. Similar dependence on cell proliferation for viral replication efficiency has been demonstrated in a number of picornaviruses [75,78,79,80]. Further, using the model proposed by Sanjuán (2012), we were able to show that viruses that infect epithelial cells have generation times that are between 5- and 40-fold shorter than a virus that infects non-epithelial cells. This offers a possible mechanistic basis for our finding that viruses that target the fastest-dividing cells in the body (intestinal and respiratory epithelial cells [53,54,55,81]) have higher substitution rates than viruses that infect cells that turnover at very low rates, if at all (neurons [52,54,82]).

4. CONCLUSIONS

We are the first to provide statistical evidence that cell tropism predicts rates of mammalian RNA virus evolution, likely through its influence on virus generation time. These results offer a new perspective on why it has been difficult to create effective vaccines for viruses that infect epithelial tissue, such as rotavirus and enterovirus 71 [83,84]. Further, as it has been shown that higher rates of viral evolution can result in increased genetic diversity and higher epidemiological fitness [45,85,86], the higher substitution rates of epithelial-infecting viruses predict increased evolvability and epidemic potential.

5. APPENDICES

5.1. Appendix 1. Nucleotide substitution rates and characteristics of all viruses used in this study.

5.2. Appendix 2. Dataset and analysis information for novel substitution rates produced in this study. Abbreviations for viruses and genes are as in Table S1. Nucleotide substitution models shown general time reversible (GTR), Tamura-Nei (TrN), transition (TIM), transversion (TVM), transversion with equal frequencies (TVMef), Kimura 3-parameter with unequal frequencies (K81uf), and Hasegawa-Kishino-Yano (HKY); corrections for invariant sites (+i) and a gamma distribution of rate heterogeneity (+G) were also included in some models.

5.3. Appendix 3. Figures.

5.3.1. Figure 1: Nucleotide substitution rates and ecological properties of mammalian RNA viruses. Log scale mean substitution rate (log_{10}(nucleotide substitutions/site/year, NS/S/Y)) estimates for different target cells (A and B), transmission routes (C and D), infection modes (E and F), and host ranges (G and H). Plots on the left show rates based on structural genes, while the plots on the right show those of non-structural genes. Each black bar indicates the mean of each level, and the levels of each variable are sorted by increasing mean substitution rate.

5.3.2. Figure 2: Nucleotide substitution rates and genomic properties of mammalian RNA viruses. Log scale mean substitution rate (log_{10}(nucleotide substitutions/site/year, NS/S/Y)) estimates for different genomic architectures (sense/strandedness, A and B, and whether or not the genome is segmented, C and D) and plotted against genome lengths (E and F). The plots on the left show rates based on structural genes, while the plots on the right show those of non-structural genes. Each black bar in A-D indicates the mean of each level, and the levels of each of these variables are sorted by increasing mean substitution rate. The line of best fit is shown in E and F. The coefficients of determination (R^2) for the linear regression models of genome lengths vs. substitution rates were 0.06 for the structural gene dataset and 0.08 for the non-structural gene dataset.

5.3.3. Figure 3: Standardized regression coefficients for predictors of viral substitution rates. Standardized coefficients with 95% confidence intervals for the different predictor variables of structural (left) and non-structural (right) gene substitution rates. A and B show the coefficients from the first regression analysis, C and D show coefficients from the second regression analysis, and E and F show coefficients from the third regression analysis. Coefficients are indicated by the same symbols used in Figures 1 and 2. Dark

coefficients correspond to significant substitution rate predictors (*P*<0.01: neural, leukocyte, hepatocyte, and epithelial target cells in A, leukocyte and epithelial target cells in C, neural and epithelial target cells in E, and neural target cells in F), while the other coefficients are shown in gray.

5.3.4. Figure 4: Standardized regression coefficients for predictors of viral substitution rates based on analyses of control datasets. Standardized coefficients with 95% confidence intervals for the different predictor variables of structural (left) and non-structural (right) gene substitution rates, using the datasets with one rate per viral species. A and B show the coefficients from the first regression analysis, C and D show coefficients from the second regression analysis, and E and F show coefficients from the third regression analysis. Coefficients are indicated by the same symbols used in Figures 1 and 2. Dark coefficients correspond to significant substitution rate predictors (*P*<0.01, epithelial, leukocyte, hepatocyte, and epithelial target cells in A, leukocyte and epithelial target cells in C, neural and epithelial target cells in E, and neural target cells in F), while the other coefficients are shown in gray.

5.3.5. Figure 5: Nucleotide substitution rates and principle target cells of mammalian RNA viruses. Log scale mean nucleotide

substitution rates ($\log_{10}$(nucleotide substitutions per site per year, NS/S/Y)) of all RNA viruses included in this study with 95% credibility intervals. Credibility intervals that are not visible are eclipsed by the symbol or, in three cases (NoV GII.b, HEV, and TBEV), were not available from the published source. Rates based on structural genes are shown in the left panel, and rates based on non-structural genes are shown in the right panel. Sources of the rates are given in Appendix 1.

5.1. Appendix 1. Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)[1]	Principal Transmission Route(s)	Infection Duration	Host Range[2]	Nucleotide Substitution Rate (x10^{-3})[3]	dN/dS[4]	Gene[5]	Source of Rate, dN/dS
Human astrovirus (hAstV)	(+)ssRNA	*Astroviridae*	7.3	Non-segmented	Intestinal epithelial cells [42]	Fecal-oral [42]	Acute, persistent [42]	Species (*Homo sapiens*) [42]	2.38 (1.43-3.49)	0.07	S (p)	This study
Equine arteritis virus (EAV)	(+)ssRNA	*Arteriviridae*	12.7	Non-segmented	Leukocytes (systemic) [43,44]	Respiratory, vertical [43,44]	Acute, persistent [43,44]	Genus (*Equus*) [44]	2.70 (2.09-3.28)	0.27	S	This study
Porcine reproductive and respiratory syndrome virus (PRRSV)	(+)ssRNA	*Arteriviridae*	15.4	Non-segmented	Leukocytes (systemic) [43,45,46,47]	Respiratory, vertical [43,45]	Acute, persistent [43,45]	Species (*Sus scrofa*) [43,45]	5.19 (1.97-8.23)	0.14	NS	This study
Porcine reproductive and respiratory syndrome virus type 2 (PRRSV-2)	(+)ssRNA	*Arteriviridae*	15.4	Non-segmented	Leukocytes (systemic) [43,45,46,47]	Respiratory, vertical [43,45]	Acute, persistent [43,45]	Species (*Sus scrofa*) [43,45]	9.60 (8.70-11.00)	N/A	S	[48]
Norwalk virus GII.b (NoV GII.b)	(+)ssRNA	*Caliciviridae*	7.7	Non-segmented	Intestinal epithelial cells [49,50]	Fecal-oral [50,51,52,53]	Acute, persistent [50]	Class (Mammalia) [54]	6.12	N/A	NS	[55]
Norwalk virus GII.3 (NoV GII.3)	(+)ssRNA	*Caliciviridae*	7.7	Non-segmented	Intestinal epithelial cells [49,50]	Fecal-oral [50,51,52,53]	Acute, persistent [50]	Class (Mammalia) [54]	5.80 (4.40-7.40)	N/A	S	[56]
									5.54 (4.43-6.74)	N/A	S	[55]
									3.99	N/A	NS	[55]
Norwalk virus GII.4 (NoV GII.4)	(+)ssRNA	*Caliciviridae*	7.7	Non-segmented	Intestinal epithelial cells [49,50]	Fecal-oral [50,51,52,53]	Acute, persistent [50]	Class (Mammalia) [54]	5.33 (4.62-6.02)	N/A	S	[57]
									5.10 (4.40-6.00)	N/A	S	[58]
									5.63 (3.62-7.71)	0.07	NS	This study
Rabbit hemorrhagic disease virus (RHDV)	(+)ssRNA	*Caliciviridae*	7.4	Non-segmented	Leukocytes (systemic) [59,60]	Respiratory [61,62,63]	Acute, persistent [62,63,64]	Family (Leporidae) [61,64,65]	1.91 (1.50-2.34)	0.09	S	[31]
									1.92 (1.24-2.49)	0.07	NS	[31]
Bovine coronavirus (bCoV)	(+)ssRNA	*Coronaviridae*	31.0	Non-segmented	Intestinal, respiratory epithelial cells [66,67,68,69]	Fecal-oral, Respiratory [69,70]	Acute, persistent [71]	Family (Bovidae) [67,68,69]	0.83 (0.52-1.16)	0.32	S (p)	This study

[1] Principal target cell(s) refers to the cell(s) that is/are most commonly targeted by the virus. Viruses that infect leukocytes as well as a variety of different cell types (with no clear principal target cell) are classified as "leukocytes (systemic)".

[2] Some host ranges classified as class-wide (infecting mammals from different orders) may be broader (*i.e.*, also infect non-mammalian vertebrate hosts, or, in the case of arboviruses, arthropod hosts).

[3] Substitution rates are given in nucleotide substitutions/site/year; 95% HPD shown in parentheses, when available.

[4] N/A indicates that dN/dS ratios were not calculated, were calculated by methods other than the Single Likelihood Ancestor Counting method, or were only reported for individual codon positions and/or individual lineages within the full dataset.

[5] Unless other otherwise indicated, S refers to the outer structural protein with the major antigenic site(s), and NS refers to the RNA-dependent RNA polymerase. (p) indicates partial gene sequences.

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Dengue virus (DENV)	(+)ssRNA	*Flaviviridae*	10.7	Non-segmented	Leukocytes (systemic) [72,73]	Arthropod vector [72]	Acute [72,73]	Order (Primates) [72,74]	0.76 (0.66-0.87)	0.06-0.07	S	[75]
									0.70 (0.60-0.80)	0.07	S	[76]
									0.78 (0.65-0.91)	N/A	S	[77]
									0.86 (0.65-1.10)	N/A	S	[78]
Dengue virus type 2 (DENV-2)	(+)ssRNA	*Flaviviridae*	10.7	Non-segmented	Leukocytes (systemic) [72,73]	Arthropod vector [72]	Acute [72,73]	Order (Primates) [72,74]	0.71 (0.60-0.82)	0.06	S	[76]
									0.75 (0.63-0.87)	0.25	S	[79]
									0.65 (0.41-0.87)	N/A	S	[80]
									0.83 (0.66-0.98)	0.05	S	[81]
									0.80 (0.66-0.95)	N/A	S	[82]
									0.86 (0.74-0.98)	N/A	S	[77]
									0.85 (0.70-1.00)	N/A	S	[78]
									0.69 (0.59-0.79)	0.05	NS	[81]
Dengue virus type 3 (DENV-3)	(+)ssRNA	*Flaviviridae*	10.7	Non-segmented	Leukocytes (systemic) [72,73]	Arthropod vector [72]	Acute [72,73]	Order (Primates) [72,74]	0.87 (0.76-0.98)	0.08	S	[76]
									0.90 (0.69-1.00)	N/A	S	[83]
									1.10 (0.83-1.38)	N/A	S	[84]
									0.95 (0.73-1.04)	N/A	S	[77]
									0.87 (0.67-1.08)	N/A	S	[78]
									0.89 (0.79-1.00)	N/A	S	[33]
Dengue virus type 4 (DENV-4)	(+)ssRNA	*Flaviviridae*	10.7	Non-segmented	Leukocytes (systemic) [72,73]	Arthropod vector [72]	Acute [72,73]	Order (Primates) [72,74]	0.69 (0.41-1.00)	N/A	S	[83]
									0.97 (0.79-1.06)	N/A	S	[77]
									0.06 (0.05-0.08)	N/A	S	[78]
									0.72 (0.58-0.88)	0.07	S	[76]
									0.83 (0.68-1.00)	N/A	S	[82]
Japanese encephalitis virus (JEV)	(+)ssRNA	*Flaviviridae*	11.0	Non-segmented	Neurons [72,85,86]	Arthropod vector [72]	Acute, persistent [72,85,86]	Class (Mammalia) [72]	0.14 (0.09-0.20)	0.16	S	[87]
									1.18 (0.72-1.40)	0.02	S	[87]
									0.15 (0.07-0.24)	0.05	NS	This study

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Tick-borne encephalitis virus (TBEV)	(+)ssRNA	*Flaviviridae*	11.1	Non-segmented	Neurons [72,88]	Arthropod vector [72]	Acute, persistent [72]	Order (Rodentia) [72,74]	0.14	0.05	S	[89]
									0.79 (0.41-1.12)	N/A	S	[90]
									0.03 (0.01-0.07)	N/A	S (p)	[91]
									0.02 (0.00-0.04)	0.06	NS	This study
Powassan virus (POWV)	(+)ssRNA	*Flaviviridae*	10.8	Non-segmented	Neurons [72,92]	Arthropod vector [72]	Acute [72,92]	Class (Mammalia) [72,74,92]	0.34 (0.09-0.68)	0.09	S	This study
Yellow fever virus (YFV)	(+)ssRNA	*Flaviviridae*	10.9	Non-segmented	Leukocytes (systemic) [72,93]	Arthropod vector [72]	Acute [72,93]	Order (Primates) [72,74]	0.21 (0.10-0.33)	0.04	S	[76]
									0.13 (0.05-0.21)	0.05	NS	This study
Hepatitis C virus type 1a (HCV1a)	(+)ssRNA	*Flaviviridae*	9.5	Non-segmented	Hepatocytes [94]	Parenteral [94]	Acute, persistent [94]	Species (*Homo sapiens*) [94]	2.45 (1.67-3.21)	0.27	S	[95][6]
									3.41 (2.54-4.32)	N/A	S (p)	[96]
									0.71 (0.38-1.07)	0.11	NS	[95]
Hepatitis C virus type 1b (HCV1b)	(+)ssRNA	*Flaviviridae*	9.5	Non-segmented	Hepatocytes [94]	Parenteral [94]	Acute, persistent [94]	Species (*Homo sapiens*) [94]	2.72 (1.71-3.75)	0.26	S	[95]
									0.57 (0.22-0.96)	0.12	NS	[95]
Hepatitis E virus (HEV)	(+)ssRNA	*Hepeviridae*	7.2	Non-segmented	Hepatocytes [97]	Fecal-oral [97,98]	Acute [97]	Class (Mammalia) [97]	1.13	N/A	S	[99]
Foot-and-mouth disease virus (FMDV)	(+)ssRNA	*Picornaviridae*	8.2	Non-segmented	Many epithelial cells [100]	Respiratory [101]	Acute, persistent [101,102]	Order (Artiodactyla) [101]	2.48 (1.69-3.31)	N/A	S	[35]
Foot-and-mouth disease virus type A (FMDV-A)	(+)ssRNA	*Picornaviridae*	8.2	Non-segmented	Many epithelial cells [100]	Respiratory [101]	Acute, persistent [101,102]	Order (Artiodactyla) [101]	5.77 (4.81-6.74)	N/A	S	[103]
									1.45 (0.07-2.24)	0.05	NS	[104]
Foot-and-mouth disease virus type O (FMDV-O)	(+)ssRNA	*Picornaviridae*	8.2	Non-segmented	Many epithelial cells [100]	Respiratory [101]	Acute, persistent [101,102]	Order (Artiodactyla) [101]	4.81 (4.04-5.46)	N/A	S	[103]
Foot-and-mouth disease virus type SAT2 (FMDV-SAT2)	(+)ssRNA	*Picornaviridae*	8.2	Non-segmented	Many epithelial cells [100]	Respiratory [101]	Acute, persistent [101,102]	Order (Artiodactyla) [101]	2.42 (1.75-3.12)	N/A	S	[105]
Coxsackievirus A16 (CVA16)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Many epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	5.38 (3.96-6.85)	0.04	S	This study
									6.23 (4.46-8.11)	0.05	NS (p)	This study
Enterovirus 71 (EV71)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Many epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	5.53 (4.29-6.67)	0.05	NS	[104]

[6]Rates in this paper were reported for small partitions of the full genome that did not coincide with gene boundaries. The single-gene rates (and dN/dS ratios) reported here were produced in this study using the exact datasets from the paper cited.

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Enterovirus 71 type B (EV71-B)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Many epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	4.50 (4.30-4.70)	N/A	S	[112]
Enterovirus 71 type C (EV71-C)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Many epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	3.66 (3.25-4.05)	N/A	S	[113]
									4.20 (4.00-4.40)	N/A	S	[112]
Coxsackievirus B3 (CVB3)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute, persistent [106,114,115,116,117]	Species (*Homo sapiens*) [106,109]	4.80 (3.80-5.80)	N/A	S	[118]
Coxsackievirus B4 (CVB4)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute, persistent [106]	Species (*Homo sapiens*) [106,109]	4.95 (4.17-5.83)	0.02	S	This study
Coxsackievirus B5 (CVB5)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute, persistent [106]	Species (*Homo sapiens*) [106,109]	4.20 (3.30-5.20)	N/A	S	[119]
Echovirus 6 (E6)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	6.42 (5.04-7.84)	0.04	S	This study
Echovirus 9 (E9)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	5.80 (3.70-8.10)	N/A	S	[120]
									9.37 (5.17-14.34)	0.03	NS (p)	This study
Echovirus 11 (E11)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	4.80 (3.60-6.10)	N/A	S	[120]
									4.30 (1.66-7.42)	0.03	NS (p)	This study
Echovirus 13 (E13)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	15.01 (7.52-24.98)	0.04	S	This study
Echovirus 30 (E30)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	4.38 (3.95-4.83)	0.06	S	This study
									4.30 (1.66-7.42)	0.03	NS (p)	This study
Echovirus 33 (E33)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal, respiratory epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	10.71 (5.48-15.57)	0.04	S	This study
Swine vesicular disease virus (SVDV)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Many epithelial cells [121]	Fecal-oral [122]	Acute, persistent [122,123]	Genus (*Sus*) [121]	3.49 (2.44-4.56)	0.10	S	This study
Coxsackievirus A24 (CVA24)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute, persistent [106]	Species (*Homo sapiens*) [106,109]	11.82 (9.23-14.57)	0.07	S	This study
									11.61 (8.73-14.44)	0.06	NS (3C)	This study

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Poliovirus type 1 (PV1)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute, persistent [106]	Species (*Homo sapiens*) [106,109]	6.56 (5.96-7.15)	0.04	S	This study
									11.60 (3.52-19.87)	0.02	NS (p)	This study
Enterovirus 68 (EV68)	(+)ssRNA	*Picornaviridae*	7.4	Non-segmented	Intestinal epithelial cells [106,107,108,109,110]	Fecal-oral [106,108,109,111]	Acute [106]	Species (*Homo sapiens*) [106,109]	6.20 (5.40-7.10)	N/A	S	[124]
									4.93 (4.01-5.85)	0.09	S (p)	[125]
Hepatitis A virus (HAV)	(+)ssRNA	*Picornaviridae*	7.5	Non-segmented	Hepatocytes [126]	Fecal-oral [126]	Acute [126]	Order (Primates) [126]	1.08 (0.80-1.34)	0.05	S	This study
									0.36 (0.21-0.52)	0.04	NS	This study
Human parechovirus (HPeV)	(+)ssRNA	*Picornaviridae*	7.3	Non-segmented	Intestinal, respiratory epithelial cells [127]	Fecal-oral, respiratory [127,128]	Acute [127,128]	Species (*Homo sapiens*) [127,128]	2.79 (2.05-3.66)	N/A	S	[129]
									2.96 (1.88-3.92)	0.04	NS (p)	[104]
Porcine teschovirus (PTV)	(+)ssRNA	*Picornaviridae*	7.1	Non-segmented	Leukocytes (systemic) [121,130,131]	Fecal-oral [132]	Acute, persistent [133]	Genus (*Sus*) [132]	2.46 (2.03-2.95)	N/A	S	[132]
									1.60 (1.34-1.85)	N/A	S	[134]
									1.62 (0.63-2.75)	0.10	S (p)	[104]
Chikungunya virus (CHIKV)	(+)ssRNA	*Togaviridae*	11.5	Non-segmented	Leukocytes (systemic) [135,136]	Arthropod vector [137]	Acute, persistent [135,136]	Order (Primates) [137]	0.84 (0.50-1.23)	N/A	S (p)	[138]
									0.66 (0.47-0.87)	0.07	NS	This study
Ross River virus (RRV)	(+)ssRNA	*Togaviridae*	11.7	Non-segmented	Leukocytes (systemic) [137,139]	Arthropod vector [137]	Acute, persistent [137,139]	Class (Mammalia) [137,139]	0.49 (0.24-0.75)	0.38	S	This study
Venezuelan equine encephalitis virus (VEEV)	(+)ssRNA	*Togaviridae*	11.4	Non-segmented	Neurons [137,140,141]	Arthropod vector [137]	Acute [137,142]	Class (Mammalia) [137,143,144,145]	0.07 (0.00-0.21)	0.07	S	This study
									0.12 (0.04-0.21)	0.04	NS	This study
Western equine encephalitis virus (WEEV)	(+)ssRNA	*Togaviridae*	11.5	Non-segmented	Neurons [137,146]	Arthropod vector [137]	Acute, persistent [137,147]	Class (Mammalia) [137,143,148]	0.17 (0.10-0.26)	0.23	S	This study
Rubella virus (RuV)	(+)ssRNA	*Togaviridae*	9.8	Non-segmented	Leukocytes (systemic) [149,150,151,152]	Respiratory, vertical [153]	Acute, persistent [153]	Species (*Homo sapiens*) [153]	0.82 (0.68-0.97)	0.05	S	This study
Lassa virus (LasV)	(-)ssRNA	*Arenaviridae*	10.7	Segmented	Leukocytes (systemic) [152,154,155]	Respiratory, vertical [154]	Acute, persistent [154]	Genus (*Mastomys*) [154]	2.88 (1.41-4.52)	0.07	S (p)	This study
Lymphocytic choriomeningitis virus (LCMV)	(-)ssRNA	*Arenaviridae*	10.1	Segmented	Leukocytes (systemic) [152,154]	Respiratory, vertical [154]	Acute, persistent [154]	Genus (*Mus*) [154]	0.33 (0.14-0.52)	N/A	S	[156]
Borna disease virus (BDV)	(-)ssRNA	*Bornaviridae*	8.9	Non-segmented	Neurons [157,158]	Respiratory [157,159,160,161]	Acute, persistent [157]	Class (Mammalia) [157,159]	0.08 (0.01-0.15)	0.03	S (N)	This study
Dobrava-Belgrade virus (DBV)	(-)ssRNA	*Bunyaviridae*	11.8	Segmented	Endothelial cells [162,163]	Fecal-oral, respiratory [163]	Persistent [163]	Genus (*Apodemus*) [163,164]	0.28 (0.01-0.68)	N/A	S (p) (N)	[165]

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Puumala virus (PUUV)	(-)ssRNA	*Bunyaviridae*	12.1	Segmented	Endothelial cells [162,163]	Fecal-oral, respiratory [163]	Persistent [163]	Family (Cricetidae) [163,164]	0.54 (0.07-0.98)	N/A	S (N)	[165]
Seoul virus (SEOV)	(-)ssRNA	*Bunyaviridae*	12.0	Segmented	Endothelial cells [162,163]	Fecal-oral, respiratory [163]	Persistent [163]	Genus (*Rattus*) [163,164]	0.41 (0.02-1.05)	0.06	S (N)	This study
Crimean-Congo hemorrhagic fever virus (CCHFV)	(-)ssRNA	*Bunyaviridae*	19.1	Segmented	Leukocytes (systemic) [166]	Arthropod vector [163]	Acute [163]	Class (Mammalia) [163]	0.15 (0.06-0.24)	N/A	S	[167]
									0.07 (0.03-0.11)	0.05	NS	This study
Rift Valley fever virus (RVFV)	(-)ssRNA	*Bunyaviridae*	12.0	Segmented	Leukocytes (systemic) [163,168,169]	Arthropod vector [163,169]	Acute, persistent [163]	Class (Mammalia) [163,169]	0.24 (0.18-0.30)	0.04-0.07	S	[170]
									0.36 (0.26-0.46)	N/A	S	[171]
									0.28 (0.20-0.35)	0.03	NS	[170]
									0.28 (0.18-0.39)	N/A	NS	[171]
Toscana virus (TosV)	(-)ssRNA	*Bunyaviridae*	12.5	Segmented	Neurons	Arthropod vector	Acute	Class (Mammalia)	0.09 (0.00-0.25)	N/A	S	[172]
									0.25 (0.03-0.54)	N/A	S (p)	[173]
Influenza A virus (FLUAV)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	3.92 (2.43-5.40)	0.13	S	[32]
									2.23 (1.98-2.49)	N/A	S	[176]
									3.00 (2.50-3.40)	0.03	NS[7]	[177]
									2.86 (1.93-3.75)	0.05	NS	[32]
									2.10 (1.90-2.31)	N/A	NS	[176]
									2.60 (2.29-2.92)	N/A	NS	[178]
									2.59 (2.13-3.04)	N/A	NS	[179]
Influenza A virus H1 (FLUAV H1)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	3.67 (3.41-3.92)	N/A	S	[178]
Influenza A virus H4 (FLUAV H4)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	2.50 (2.00-3.10)	0.09	S	[177]
Influenza A virus H5 (FLUAV H5)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	4.77 (3.88-5.74)	0.27	S	[179]
Influenza A virus H6 (FLUAV H6)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	4.20 (3.30-5.00)	0.15	S	[177]
Influenza B virus (FLUBV)	(-)ssRNA	*Orthomyxoviridae*	13.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	2.15 (1.85-2.46)	0.22	S	[180]
									0.27 (0.12-0.41)	0.04	NS	[180]

[7]The PB1 polymerase gene was used for all of the influenza viruses.

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate (x10^{-3})	dN/dS	Gene	Source of Rate, dN/dS
Influenza C virus (FLUCV)	(-)ssRNA	*Orthomyxoviridae*	12.6	Segmented	Respiratory epithelial cells [174,175]	Respiratory [174,175]	Acute [174,175]	Class (Mammalia) [174,175]	0.49 (0.41-0.57)	N/A	S	[181]
									0.68 (0.48-0.89)	N/A	NS	[181]
Canine distemper virus (CDV)	(-)ssRNA	*Paramyxoviridae*	15.7	Non-segmented	Leukocytes (systemic) [152,182,183]	Respiratory [182,183]	Acute, persistent [182,184]	Class (Mammalia) [182,183]	1.05 (0.52-1.60)	0.26	S	[185]
Measles virus (MeV)	(-)ssRNA	*Paramyxoviridae*	15.9	Non-segmented	Leukocytes (systemic) [152,183,186]	Respiratory [183,186]	Acute, persistent [183,186]	Species (*Homo sapiens*) [183,186]	0.66 (0.48-0.83)	0.23	S	[187]
									0.56 (0.45-0.68)	0.02	S (p)	[188]
									0.42 (0.27-0.56)	0.52	NS (P/C/V)	This study
Human parainfluenza virus 1 (HPiV-1)	(-)ssRNA	*Paramyxoviridae*	15.6	Non-segmented	Respiratory epithelial cells [189]	Respiratory [189]	Acute [189]	Species (*Homo sapiens*) [189]	1.37 (1.16-1.59)	N/A	S	[190]
Mumps virus (MuV)	(-)ssRNA	*Paramyxoviridae*	15.4	Non-segmented	Leukocytes (systemic) [191,192]	Respiratory [192]	Acute, persistent [192]	Species (*Homo sapiens*) [192]	0.41 (0.30-0.51)	0.16	S	This study
Human metapneumovirus (HMPV)	(-)ssRNA	*Paramyxoviridae*	13.3	Non-segmented	Respiratory epithelial cells [193,194]	Respiratory [194]	Acute, persistent [193]	Species (*Homo sapiens*) [194]	7.40 (5.72-9.15)	N/A	S	[195]
									3.50 (2.30-4.80)	0.52	S	[196]
									5.18 (3.76-6.78)	N/A	S	[34]
									6.49 (4.60-8.44)	N/A	S	[34]
Human respiratory syncytial virus type A (HRSV-A)	(-)ssRNA	*Paramyxoviridae*	15.2	Non-segmented	Respiratory epithelial cells [194,197]	Respiratory [194]	Acute, persistent [194]	Species (*Homo sapiens*) [194]	2.22 (1.93-2.56)	N/A	S	[198]
Bovine ephemeral fever virus (BEFV)	(-)ssRNA	*Rhabdoviridae*	14.9	Non-segmented	Leukocytes (systemic) [199,200,201,202]	Arthropod vector [202]	Acute [199,201,202]	Family (Bovinae) [201,202]	0.87 (0.48-1.28)	0.10	S	This study
Rabies virus (RabV)	(-)ssRNA	*Rhabdoviridae*	11.9	Non-segmented	Neurons [202]	Bites, scratches [202]	Acute [202]	Class (Mammalia) [202]	0.39 (0.12-0.65)	N/A	S	[203]
									0.40 (0.21-0.60)	N/A	S	[204]
									0.33 (0.22-0.43)	0.08	S	[205]
									0.32 (0.22-0.44)	N/A	S	[206]
									0.63 (0.33-1.10)	N/A	S	[207]
									0.09 (0.00-0.20)	0.04	NS (p)	This study
European Bat Lyssavirus 1 (EBLV-1)	(-)ssRNA	*Rhabdoviridae*	12.0	Non-segmented	Neurons [208,209]	Bites, scratches [202,210]	Acute [208,209]	Order (Chiroptera) [208,209]	0.05 (0.00-0.09)	N/A	S	[211]
Bluetongue virus (BTV)	dsRNA	*Reoviridae*	19.2	Segmented	Leukocytes (systemic) [212,213]	Arthropod vector [212,213]	Acute, persistent [212,213]	Class (Mammalia) [212,213]	0.49 (0.20-0.81)	0.12	S	[214]
									0.69 (0.34-1.07)	0.06	NS (NS3)	[214]

Appendix 1 (continued). Nucleotide substitution rates and characteristics of all viruses used in this study.

Virus	Genome Architecture	Family	Genome Length (kb)	Segmented or Non-segmented	Principal Target Cell(s)	Principal Transmission Route(s)	Infection Duration	Host Range	Nucleotide Substitution Rate ($x10^{-3}$)	dN/dS	Gene	Source of Rate, dN/dS
Epizootic hemorrhagic disease virus type 2 (EHDV-2)	dsRNA	*Reoviridae*	19.4	Segmented	Leukocytes (systemic) [213,215,216]	Arthropod vector [213,217]	Acute, persistent [213,215]	Class (Mammalia) [213,217]	0.48 (0.31-0.66)	N/A	NS (p) (NS3)	[217]
Rotavirus A (RVA)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	2.49 (1.00-4.08)	0.03	NS	This study
Rotavirus A G1 (RVA G1)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.41 (1.03-1.79)	N/A	S	[219]
Rotavirus A G2 (RVA G2)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.53 (0.80-2.38)	0.17	S	This study
Rotavirus A G3 (RVA G3)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.95 (1.55-2.39)	0.19	S	This study
Rotavirus A G9 (RVA G9)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.87 (1.45-2.27)	N/A	S	[220]
Rotavirus A G12 (RVA G12)	dsRNA	*Reoviridae*	18.6	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.66 (1.13-2.32)	N/A	S	[220]
Rotavirus B (RVB)	dsRNA	*Reoviridae*	18.0	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	1.38 (0.36-2.72)	0.10	S	[221]
									1.91 (0.59-3.41)	0.32-0.23	NS (NSP1)	[221]
Rotavirus C (RVC)	dsRNA	*Reoviridae*	17.9	Segmented	Intestinal epithelial cells [218]	Fecal-oral [218]	Acute, persistent [218]	Class (Mammalia) [218]	10.24 (7.36-13.05)	0.09	S	This study
									13.04 (7.15-36.03)	0.13	NS (p) (NSP4)	This study

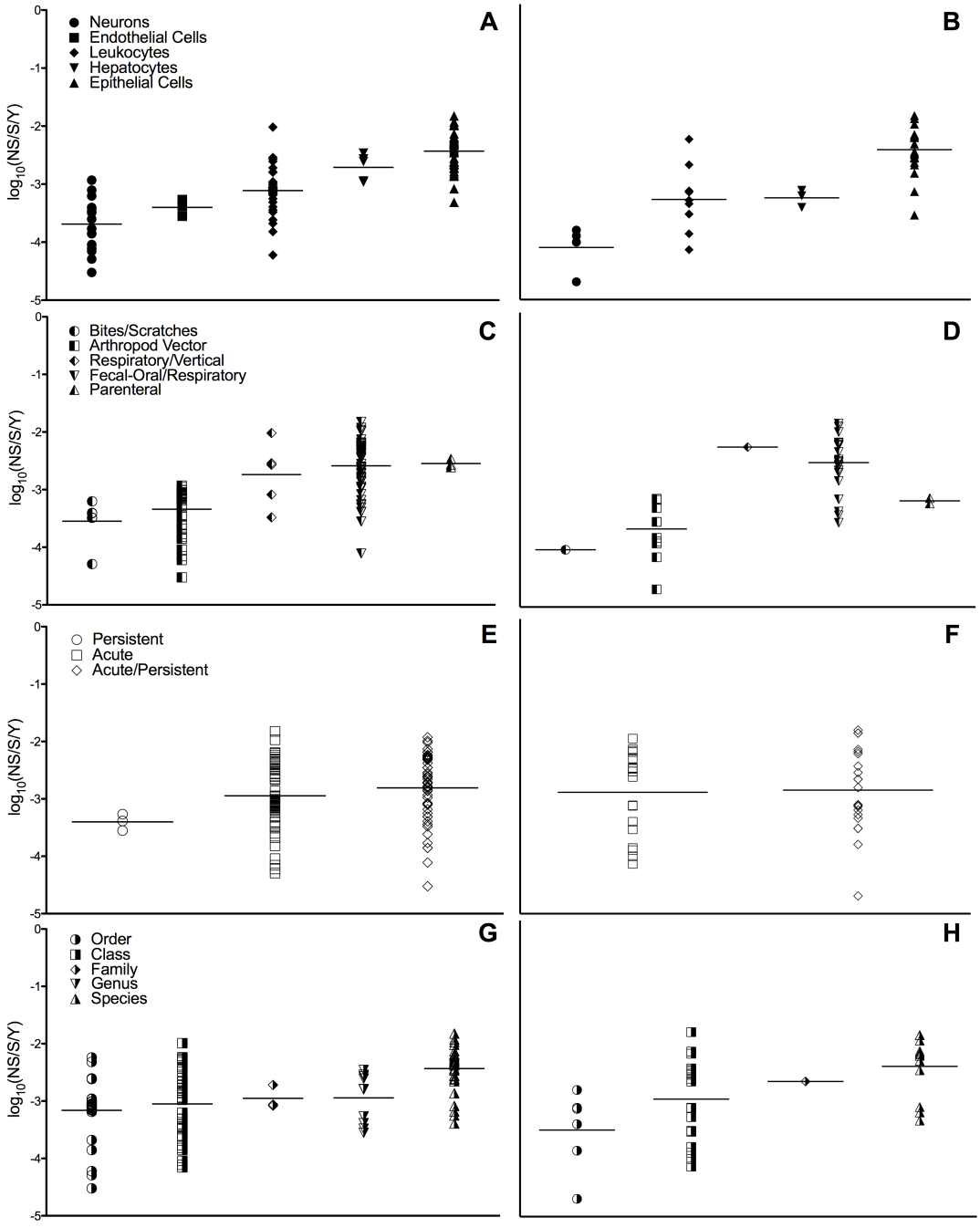
5.2. Appendix 2. Dataset and analysis information for novel substitution rates produced in this study.

Virus	Gene	$N_{nucleotides}$	N_{taxa}	Date Range	Nucleotide Substitution Model	Clock Model	Demographic Model
hAstV	S (p)	1071	86	1990-2010	TrN+i+G	Uncorrelated lognormal	Bayesian skyline
EAV	S	768	155	1954-2008	TVM+i+G	Uncorrelated lognormal	Constant
PRRSV	NS	1872	122	1990-2008	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
NoV GII.4	NS	1524	332	1971-2010	GTR+i+G	Uncorrelated lognormal	Constant
bCoV	S (p)	2703	45	1993-2011	GTR+i+G	Uncorrelated lognormal	Exponential
JEV	NS	2673	111	1935-2009	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
POWV	S	660	66	1952-2011	TIM+G	Uncorrelated lognormal	Bayesian skyline
TBEV	NS	2709	65	1952-2010	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
YFV	NS	2724	36	1973-2010	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
CVA16	S	891	394	1981-2010	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
CVA16	NS (p)	708	69	1997-2011	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
CVB4	S	807	100	1959-2010	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
E6	S	807	185	1991-2010	GTR+G	Uncorrelated lognormal	Exponential
E9	NS (p)	549	83	1995-2010	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
E11	NS (p)	549	127	1982-2008	GTR+i+G	Uncorrelated lognormal	Exponential
E13	S	807	64	1991-2006	TrN+i+G	Uncorrelated lognormal	Constant
E30	S	810	421	1959-2010	GTR+G	Uncorrelated lognormal	Bayesian skyline
E30	NS (p)	561	64	1981-2005	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
E33	S	807	44	1983-2005	GTR+G	Uncorrelated lognormal	Exponential
SVDV	S	897	51	1970-1999	TrN+i+G	Uncorrelated lognormal	Bayesian skyline
CVA24	S	915	121	1963-2010	GTR+i+G	Uncorrelated lognormal	Constant
CVA24	NS	549	236	1963-2010	TrN+i+G	Uncorrelated lognormal	Constant
PV1	S	906	478	1959-2010	GTR+G	Uncorrelated lognormal	Constant
PV1	NS (p)	792	35	1989-2008	GTR+i+G	Uncorrelated lognormal	Constant
HAV	S	900	213	1957-2010	GTR+i+G	Uncorrelated lognormal	Exponential
HAV	NS	1467	34	1957-2010	GTR+i+G	Uncorrelated lognormal	Exponential
CHIKV	NS	1833	151	1953-2011	GTR+G	Uncorrelated lognormal	Exponential
RRV	S	1266	137	1959-2009	GTR+i	Uncorrelated lognormal	Constant
VEEV	S	1269	40	1953-2008	GTR+i+G	Uncorrelated lognormal	Constant
VEEV	NS	1809	32	1943-2001	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
WEEV	S	801	45	1930-2005	GTR+G	Uncorrelated lognormal	Constant
RuV	S	1416	194	1967-2007	GTR+G	Uncorrelated lognormal	Bayesian skyline
LasV	S (p)	771	50	1974-2009	GTR+i+G	Uncorrelated lognormal	Exponential
BDV	S	1110	44	1984-2008	TVMef+i+G	Strict	Constant

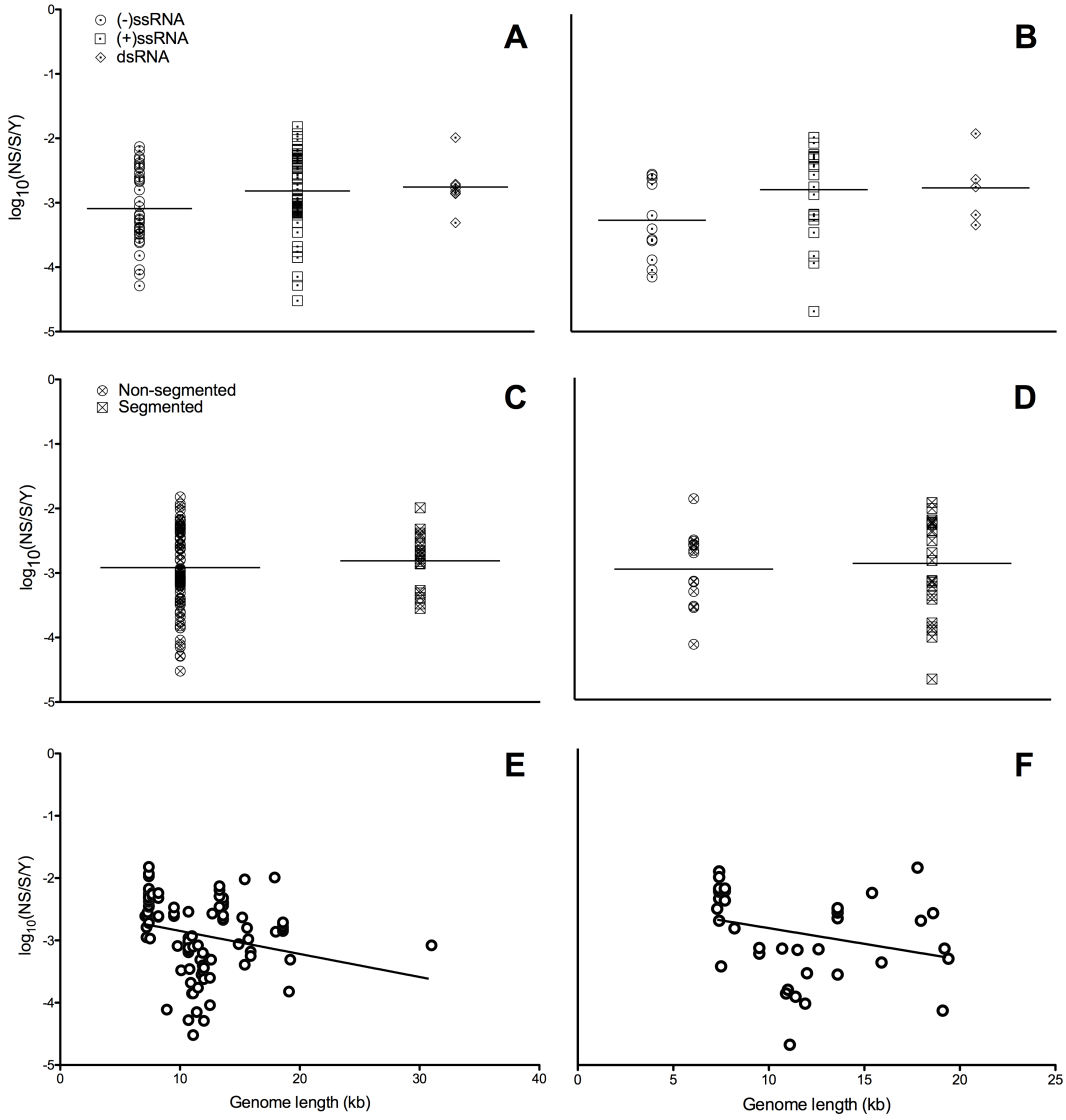
5.2. Appendix 2 (continued). Dataset and analysis information for novel substitution rates produced in this study.

Virus	Gene	$N_{nucleotides}$	N_{taxa}	Date Range	Nucleotide Substitution Model	Clock Model	Demographic Model
CCHFV	NS	11838	33	1956-2011	GTR+i+G	Uncorrelated lognormal	Exponential
SEOV	S (p)	1290	53	1995-2009	TrN+i+G	Uncorrelated lognormal	Bayesian skyline
MeV	NS	1524	94	1979-2009	TIM+G	Uncorrelated lognormal	Bayesian skyline
MuV	S	1746	51	1969-2009	TVM+G	Uncorrelated lognormal	Bayesian skyline
BEFV	S	1872	39	1966-2008	GTR+G	Uncorrelated lognormal	Constant
RabV	NS (p)	5979	35	1983-2009	GTR+i+G	Uncorrelated lognormal	Bayesian skyline
RVA	NS	3264	58	1974-2010	GTR+i+G	Uncorrelated lognormal	Constant
RVA G2	S	972	117	1976-2007	K81uf+G	Uncorrelated lognormal	Bayesian skyline
RVA G3	S	972	130	1974-2008	HKY+G	Uncorrelated lognormal	Bayesian skyline
RVC	S	1008	70	1986-2011	GTR+i+G	Uncorrelated lognormal	Exponential
RVC	NS (p)	450	46	1986-2009	TIM+G	Uncorrelated lognormal	Constant

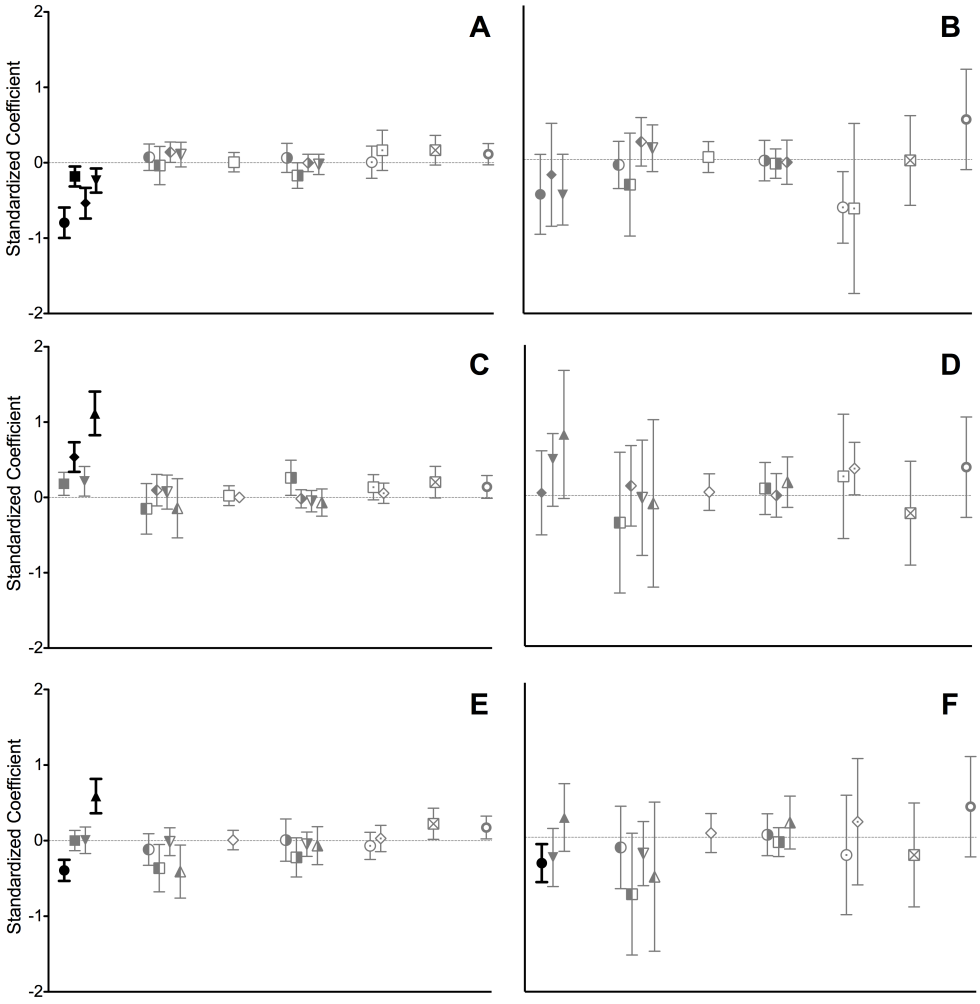
5.3.1. Appendix 3.1. Figure 1: Substitution rates and ecological properties of mammalian RNA viruses.



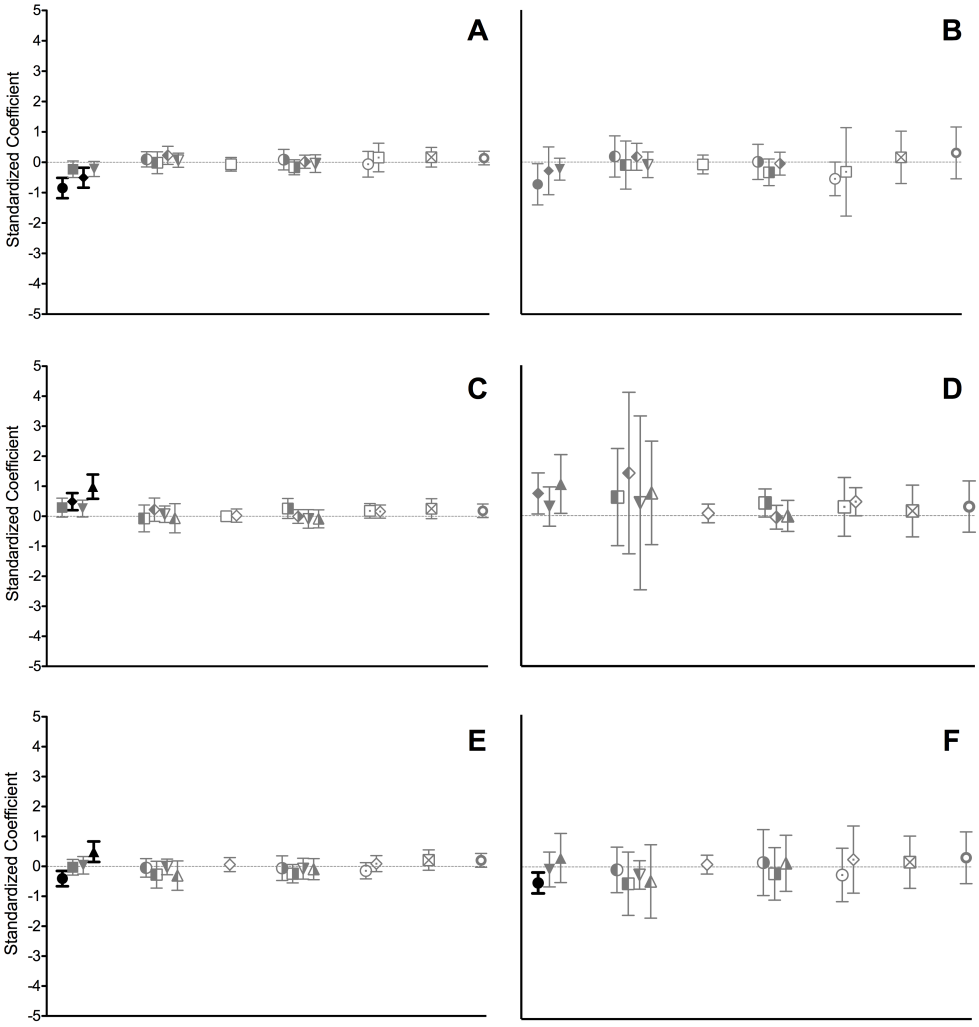
5.3.2. Appendix 3.2. Figure 2: Nucleotide substitution rates and genomic properties of mammalian RNA viruses.



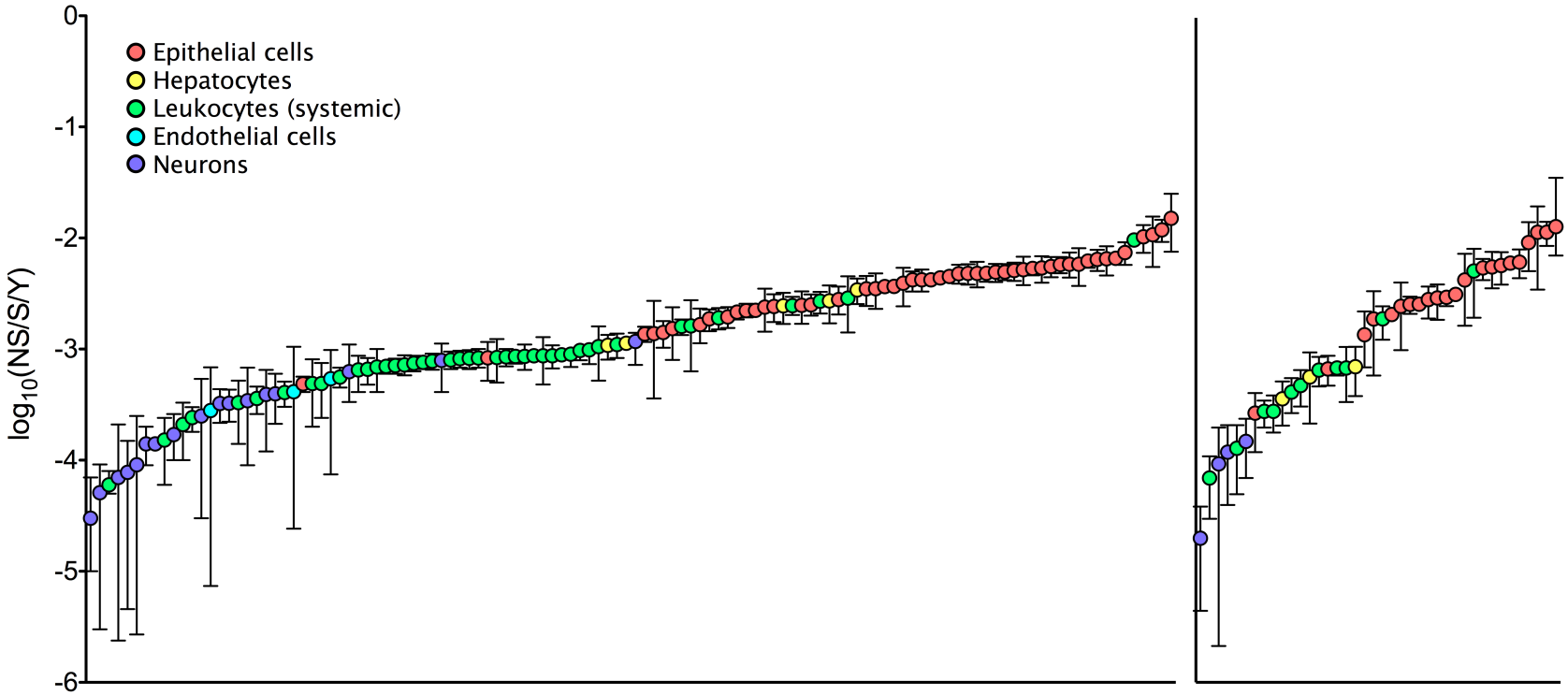
5.3.3. Appendix 3.3. Figure 3: Standardized regression coefficients for predictors of viral substitution rates.



5.3.4. Appendix 3.4. Figure 4: Standardized regression coefficients for predictors of viral substitution rates based on analyses of control datasets



5.3.5. Appendix 3.5. Figure 5: Nucleotide substitution rates and principle target cells of mammalian RNA viruses.



6. REFERENCES

1. Holmes EC (2009) The evolution and emergence of RNA viruses. Oxford: Oxford University Press. 254 p.
2. Peters CJ (2007) Emerging viral diseases. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 605-625.
3. World Health Organization (2012) Vaccine-preventable diseases. Available: http://www.who.int/immunization_monitoring/diseases/en/. Accessed 20 September 2012.
4. Rheingans RD, Antil L, Dreibelbis R, Podewils LJ, Bresee JS, et al. (2009) Economic costs of rotavirus gastroenteritis and cost-effectiveness of vaccination in developing countries. J Infect Dis 200 Suppl 1: S16-27.
5. Hanada K, Suzuki Y, Gojobori T (2004) A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. Molecular Biology and Evolution 21: 1074-1080.
6. Perelson AS (2002) Modelling viral and immune system dynamics. Nat Rev Immunol 2: 28-36.
7. Gerrish PJ, Garcia-Lerma JG (2003) Mutation rate and the efficacy of antimicrobial drug treatment. Lancet Infect Dis 3: 28-32.
8. Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, et al. (2006) Viruses as quasispecies: Biological implications. Quasispecies: Concept and Implications for Virology 299: 51-82.
9. Lauring AS, Andino R (2010) Quasispecies Theory and the Behavior of RNA Viruses. Plos Pathogens 6.
10. Moya A, Holmes EC, Gonzalez-Candelas F (2004) The population genetics and evolutionary epidemiology of RNA viruses. Nat Rev Microbiol 2: 279-288.
11. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344-348.
12. Pybus OG, Rambaut A (2009) Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet 10: 540-550.
13. Duffy S, Shackelton LA, Holmes EC (2008) Rates of evolutionary change in viruses: patterns and determinants. Nat Rev Genet 9: 267-276.
14. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R (2010) Viral Mutation Rates. Journal of Virology 84: 9733-9748.
15. Sanjuan R (2012) From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. Plos Pathogens 8: e1002685.
16. Chare ER, Holmes EC (2004) Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission. Journal of General Virology 85: 3149-3157.

17. Jenkins GM, Rambaut A, Pybus OG, Holmes EC (2002) Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54: 156-165.
18. Eigen M (1993) The origin of genetic information: viruses as models. Gene 135: 37-47.
19. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.
20. Holmes EC (2009) The Evolutionary Genetics of Emerging Viruses. Annual Review of Ecology Evolution and Systematics 40: 353-372.
21. Hughes GJ, Orciari LA, Rupprecht CE (2005) Evolutionary timescale of rabies virus adaptation to North American bats inferred from the substitution rate of the nucleoprotein gene. Journal of General Virology 86: 1467-1474.
22. Ng KK, Arnold JJ, Cameron CE (2008) Structure-function relationships among RNA-dependent RNA polymerases. Curr Top Microbiol Immunol 320: 137-156.
23. Chong YL, Padhi A, Hudson PJ, Poss M (2010) The effect of vaccination on the evolution and population dynamics of avian paramyxovirus-1. Plos Pathogens 6: e1000872.
24. Firth C, Tokarz R, Simith DB, Nunes MR, Bhat M, et al. (2012) Diversity and distribution of hantaviruses in South America. Journal of Virology 86: 13756-13766.
25. Switzer WM, Salemi M, Shanmugam V, Gao F, Cong ME, et al. (2005) Ancient co-speciation of simian foamy viruses and primates. Nature 434: 376-380.
26. Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test to taxonomic sampling. Molecular Biology and Evolution 15: 1091-1098.
27. Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. Journal of Systematics and Evolution 46: 239-257.
28. Ho SYW, Lanfear R, Bromham L, Phillips MJ, Soubrier J, et al. (2011) Time-dependent rates of molecular evolution. Molecular Ecology 20: 3087-3101.
29. Duffy S, Holmes EC (2009) Validation of high rates of nucleotide substitution in geminiviruses: phylogenetic evidence from East African cassava mosaic viruses. Journal of General Virology 90: 1539-1547.
30. Firth C, Kitchen A, Shapiro B, Suchard MA, Holmes EC, et al. (2010) Using Time-Structured Data to Estimate Evolutionary Rates of Double-Stranded DNA Viruses. Molecular Biology and Evolution 27: 2038-2051.
31. Hicks AL, Duffy S (2012) One misdated sequence of rabbit hemorrhagic disease virus prevents accurate estimation of its nucleotide substitution rate. Bmc Evolutionary Biology 12: 74.
32. Chen RB, Holmes EC (2006) Avian influenza virus exhibits rapid evolutionary dynamics. Molecular Biology and Evolution 23: 2336-2341.
33. Araujo JMG, Nogueira RMR, Schatzmayr HG, Zanotto PMD, Bello G (2009) Phylogeography and evolutionary history of dengue virus type 3. Infection Genetics and Evolution 9: 716-725.

34. Padhi A, Verghese B (2008) Positive natural selection in the evolution of human metapneumovirus attachment glycoprotein. Virus Res 131: 121-131.
35. Tully DC, Fares MA (2008) The tale of a modern animal plague: Tracing the evolutionary history and determining the time-scale for foot and mouth disease virus. Virology 382: 250-256.
36. Rambaut A (2002) Se-Al: sequence alignment editor, version 2.0a11. Available: http://tree.bio.ed.ac.uk/software/seal/. Accessed 12 September 2009.
37. Martin DP, Lemey P, Lott M, Moulton V, Posada D, et al. (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26: 2462-2463.
38. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.
39. Rambaut A, Drummond AJ (2009) Tracer, version 1.5: MCMC trace analyses tool. Available: http://beast.bio.ed.ac.uk/Tracer. Accessed 1 December 2009.
40. Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21: 2531-2533.
41. R Development Core Team (2011) R: A language and environment for statistical computing, version 2.14.1. Vienna, AT: R Foundation for Statistical Computing.
42. Mendez E, Arias CF (2007) Astroviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 981-1000.
43. Snijder EJ, Spaan WJ (2007) Arteriviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1337-1355.
44. Balasuriya UB, Hedges JF, Smalley VL, Navarrette A, McCollum WH, et al. (2004) Genetic characterization of equine arteritis virus during persistent infection of stallions. Journal of General Virology 85: 379-390.
45. Zimmerman JJ, Yoon KJ, Pirtle EC, Wills RW, Sanderson TJ, et al. (1997) Studies of porcine reproductive and respiratory syndrome (PRRS) virus infection in avian species. Veterinary Microbiology 55: 329-336.
46. Duan X, Nauwynck HJ, Pensaert MB (1997) Virus quantification and identification of cellular targets in the lungs and lymphoid tissues of pigs at different time intervals after inoculation with porcine reproductive and respiratory syndrome virus (PRRSV). Veterinary Microbiology 56: 9-19.
47. Rowland RR, Lawson S, Rossow K, Benfield DA (2003) Lymphoid tissue tropism of porcine reproductive and respiratory syndrome virus replication during persistent infection of pigs originally exposed to virus in utero. Veterinary Microbiology 96: 219-235.
48. Shi M, Lam TTY, Hon CC, Murtaugh MP, Davies PR, et al. (2010) Phylogeny-Based Evolutionary, Demographical, and Geographical Dissection of North American Type 2 Porcine Reproductive and Respiratory Syndrome Viruses. Journal of Virology 84: 8700-8711.

49. Karst SM (2010) Pathogenesis of noroviruses, emerging RNA viruses. Viruses 2: 748-781.
50. Green KY (2007) *Caliciviridae*: The noroviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 949-979.
51. Donaldson EF, Lindesmith LC, Lobue AD, Baric RS (2008) Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations. Immunol Rev 225: 190-211.
52. Moe CL (2009) Preventing norovirus transmission: how should we handle food handlers? Clin Infect Dis 48: 38-40.
53. Parashar U, Quiroz ES, Mounts AW, Monroe SS, Fankhauser RL, et al. (2001) "Norwalk-like viruses". Public health consequences and outbreak management. MMWR Recomm Rep 50: 1-17.
54. Widdowson MA, Monroe SS, Glass RI (2005) Are noroviruses emerging? Emerg Infect Dis 11: 735-737.
55. Mahar JE, Bok K, Green KY, Kirkwood CD (2013) The Importance of Intergenic Recombination in Norovirus GII.3 Evolution. Journal of Virology 87: 3687-3698.
56. Boon D, Mahar JE, Abente EJ, Kirkwood CD, Purcell RH, et al. (2011) Comparative Evolution of GII.3 and GII.4 Norovirus over a 31-Year Period. Journal of Virology 85: 8656-8666.
57. Siebenga JJ, Lemey P, Pond SLK, Rambaut A, Vennema H, et al. (2010) Phylodynamic Reconstruction Reveals Norovirus GII. 4 Epidemic Expansions and their Molecular Determinants. Plos Pathogens 6.
58. Bok K, Prikhodko VG, Green KY, Sosnovtsev SV (2009) Apoptosis in Murine Norovirus-Infected RAW264.7 Cells Is Associated with Downregulation of Survivin. Journal of Virology 83: 3647-3656.
59. Ramiro-Ibanez F, Martin-Alonso JM, Garcia Palencia P, Parra F, Alonso C (1999) Macrophage tropism of rabbit hemorrhagic disease virus is associated with vascular pathology. Virus Res 60: 21-28.
60. Kimura T, Mitsui I, Okada Y, Furuya T, Ochiai K, et al. (2001) Distribution of rabbit haemorrhagic disease virus RNA in experimentally infected rabbits. J Comp Pathol 124: 134-141.
61. Le Gall G, Arnauld C, Boilletot E, Morisse JP, Rasschaert D (1998) Molecular epidemiology of rabbit haemorrhagic disease virus outbreaks in France during 1988 to 1995. Journal of General Virology 79 (Pt 1): 11-16.
62. Cooke BD, Fenner F (2002) Rabbit haemorrhagic disease and the biological control of wild rabbits, Oryctolagus cuniculus, in Australia and New Zealand. Wildlife Research 29: 689-706.
63. Cooke BD (2002) Rabbit haemorrhagic disease: field epidemiology and the management of wild rabbit populations. Rev Sci Tech 21: 347-358.
64. Forrester NL, Boag B, Moss SR, Turner SL, Trout RC, et al. (2003) Long-term survival of New Zealand rabbit haemorrhagic disease virus RNA in wild rabbits, revealed by RT-PCR and phylogenetic analysis. Journal of General Virology 84: 3079-3086.

65. Kerr PJ, Kitchen A, Holmes EC (2009) Origin and phylodynamics of rabbit hemorrhagic disease virus. Journal of Virology 83: 12129-12138.
66. Navas-Martin SR, Weiss S (2004) Coronavirus replication and pathogenesis: Implications for the recent outbreak of severe acute respiratory syndrome (SARS), and the challenge for vaccine development. J Neurovirol 10: 75-85.
67. Weiss SR, Navas-Martin S (2005) Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. Microbiol Mol Biol Rev 69: 635-664.
68. Lai MMC, Perlman S, Anderson LJ (2007) *Coronaviridae*. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1305-1335.
69. Park SJ, Kim GY, Choy HE, Hong YJ, Saif LJ, et al. (2007) Dual enteric and respiratory tropisms of winter dysentery bovine coronavirus in calves. Archives of Virology 152: 1885-1900.
70. Tsunemitsu H, Yonemichi H, Hirai T, Kudo T, Onoe S, et al. (1991) Isolation of bovine coronavirus from feces and nasal swabs of calves with diarrhea. J Vet Med Sci 53: 433-437.
71. Rekik MR, Dea S (1994) Comparative sequence analysis of a polymorphic region of the spike glycoprotein S1 subunit of enteric bovine coronavirus isolates. Archives of Virology 135: 319-331.
72. Gubler DJ, Kuno G, Markoff L (2007) Flaviviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1153-1252.
73. Martina BE, Koraka P, Osterhaus AD (2009) Dengue virus pathogenesis: an integrated view. Clin Microbiol Rev 22: 564-581.
74. Gould EA, Solomon T (2008) Pathogenic flaviviruses. Lancet 371: 500-509.
75. Costa RL, Voloch CM, Schrago CG (2012) Comparative evolutionary epidemiology of dengue virus serotypes. Infect Genet Evol 12: 309-314.
76. Sall AA, Faye O, Diallo M, Firth C, Kitchen A, et al. (2010) Yellow Fever Virus Exhibits Slower Evolutionary Dynamics than Dengue Virus. Journal of Virology 84: 765-772.
77. Allicock OM, Lemey P, Tatem AJ, Pybus OG, Bennett SN, et al. (2012) Phylogeography and population dynamics of dengue viruses in the Americas. Molecular Biology and Evolution 29: 1533-1543.
78. Chen SP (2012) Molecular evolution and epidemiology of four serotypes of dengue virus in Thailand from 1973 to 2007. Epidemiol Infect: 1-6.
79. Anez G, Morales-Betoulle ME, Rios M (2011) Circulation of different lineages of dengue virus type 2 in Central America, their evolutionary time-scale and selection pressure analysis. Plos One 6: e27459.
80. Kumar SR, Patil JA, Cecilia D, Cherian SS, Barde PV, et al. (2010) Evolution, dispersal and replacement of American genotype dengue type 2 viruses in India (1956-2005): selection pressure and molecular clock analyses. Journal of General Virology 91: 707-720.

81. Zhang C, Mammen MP, Jr., Chinnawirotpisan P, Klungthong C, Rodpradit P, et al. (2006) Structure and age of genetic diversity of dengue virus type 2 in Thailand. Journal of General Virology 87: 873-883.
82. Carrington CV, Foster JE, Pybus OG, Bennett SN, Holmes EC (2005) Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. Journal of Virology 79: 14680-14687.
83. Patil JA, Cherian S, Walimbe AM, Bhagat A, Vallentyne J, et al. (2012) Influence of evolutionary events on the Indian subcontinent on the phylogeography of dengue type 3 and 4 viruses. Infect Genet Evol 12: 1759-1769.
84. de Araujo JM, Bello G, Romero H, Nogueira RM (2012) Origin and evolution of dengue virus type 3 in Brazil. PLoS Negl Trop Dis 6: e1784.
85. Solomon T, Vaughn DW (2002) Pathogenesis and clinical features of Japanese encephalitis and West Nile virus infections. Curr Top Microbiol Immunol 267: 171-194.
86. Johnson RT (1987) The pathogenesis of acute viral encephalitis and postinfectious encephalomyelitis. J Infect Dis 155: 359-364.
87. Chen SP (2012) Molecular phylogenetic and evolutionary analysis of Japanese encephalitis virus in China. Epidemiol Infect 140: 1637-1643.
88. Mansfield KL, Johnson N, Phipps LP, Stephenson JR, Fooks AR, et al. (2009) Tick-borne encephalitis virus - a review of an emerging zoonosis. Journal of General Virology 90: 1781-1794.
89. Subbotina EL, Loktev VB (2012) Molecular Evolution of the Tick-Borne Encephalitis and Powassan Viruses. Molecular Biology 46: 75-84.
90. Weidmann M, Ruzek D, Krivanec K, Zoller G, Essbauer S, et al. (2011) Relation of genetic phylogeny and geographical distance of tick-borne encephalitis virus in central Europe. Journal of General Virology 92: 1906-1916.
91. Uzcategui NY, Sironen T, Golovljova I, Jaaskelainen AE, Valimaa H, et al. (2012) Rate of evolution and molecular epidemiology of tick-borne encephalitis virus in Europe, including two isolations from the same focus 44 years apart. Journal of General Virology 93: 786-796.
92. Ebel GD (2010) Update on Powassan virus: emergence of a North American tick-borne flavivirus. Annu Rev Entomol 55: 95-110.
93. Monath TP, Barrett AD (2003) Pathogenesis and pathophysiology of yellow fever. Advances in Virus Research, Vol 57 60: 343-395.
94. Lemon SM, Walker C, Alter M, Yi M (2007) Hepatitis C virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1253-1304.
95. Gray RR, Parker J, Lemey P, Salemi M, Katzourakis A, et al. (2011) The mode and tempo of hepatitis C virus evolution within and among hosts. Bmc Evolutionary Biology 11.
96. Culasso AC, Elizalde M, Campos RH, Barbini L (2012) Molecular survey of hepatitis C virus in the touristic city of Mar del Plata, Argentina. Plos One 7: e44757.

97. Emerson SU, Purcell RH (2007) Hepatitis E virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 3048-3058.
98. Chauhan A, Jameel S, Dilawari JB, Chawla YK, Kaur U, et al. (1993) Hepatitis E virus transmission to a volunteer. Lancet 341: 149-150.
99. Purdy MA, Khudyakov YE (2010) Evolutionary History and Population Dynamics of Hepatitis E Virus. Plos One 5.
100. Mason PW, Grubman MJ, Baxt B (2003) Molecular basis of pathogenesis of FMDV. Virus Res 91: 9-32.
101. Saiz M, Nunez JI, Jimenez-Clavero MA, Baranowski E, Sobrino F (2002) Foot-and-mouth disease virus: biology and prospects for disease control. Microbes Infect 4: 1183-1192.
102. Whitton JL, Cornell CT, Feuer R (2005) Host and virus determinants of picornavirus pathogenesis and tropism. Nat Rev Microbiol 3: 765-776.
103. Yoon SH, Lee KN, Park JH, Kim H (2011) Molecular epidemiology of foot-and-mouth disease virus serotypes A and O with emphasis on Korean isolates: temporal and spatial dynamics. Archives of Virology 156: 817-826.
104. Hicks AL, Duffy S (2011) Genus-specific substitution rate variability among picornaviruses. Journal of Virology 85: 7942-7947.
105. Sangula AK, Belsham GJ, Muwanika VB, Heller R, Balinda SN, et al. (2010) Co-circulation of two extremely divergent serotype SAT 2 lineages in Kenya highlights challenges to foot-and-mouth disease control. Archives of Virology 155: 1625-1630.
106. Pallansch M, Roos R (2007) Enteroviruses, polioviruses, coxsackieviruses, echoviruses, and newer enteroviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 839-893.
107. Brown EH (1973) Enterovirus infections. Br Med J 2: 169-171.
108. Racaniello VR (1995) Early events in infection: Receptor binding and cell entry. In: Rotbart HA, editor. Human Enterovirus Infections. Washington, DC: ASM Press. pp. 73-93.
109. Morens D, Pallanasch M (1995) Epidemiology. In: Rotbart HA, editor. Human Enterovirus Infections. Washington, DC: ASM Press. pp. 2-23.
110. Smura T, Ylipaasto P, Klemola P, Kaijalainen S, Kyllonen L, et al. (2010) Cellular tropism of human enterovirus D species serotypes EV-94, EV-70, and EV-68 in vitro: implications for pathogenesis. J Med Virol 82: 1940-1949.
111. Jacques J, Moret H, Minette D, Leveque N, Jovenin N, et al. (2008) Epidemiological, molecular, and clinical features of enterovirus respiratory infections in French children between 1999 and 2005. Journal of Clinical Microbiology 46: 206-213.
112. Tee KK, Lam TTY, Chan YF, Bible JM, Kamarulzaman A, et al. (2010) Evolutionary Genetics of Human Enterovirus 71: Origin, Population Dynamics, Natural Selection, and Seasonal Periodicity of the VP1 Gene. Journal of Virology 84: 3339-3350.

113. Mirand A, Schuffenecker I, Henquell C, Billaud G, Jugie G, et al. (2010) Phylogenetic evidence for a recent spread of two populations of human enterovirus 71 in European countries. Journal of General Virology 91: 2263-2277.
114. Pinkert S, Klingel K, Lindig V, Dorner A, Zeichhardt H, et al. (2011) Virus-host coevolution in a persistently coxsackievirus B3-infected cardiomyocyte cell line. Journal of Virology 85: 13409-13419.
115. Klingel K, Kandolf R (1993) The role of enterovirus replication in the development of acute and chronic heart muscle disease in different immunocompetent mouse strains. Scand J Infect Dis Suppl 88: 79-85.
116. Kandolf R, Klingel K, Zell R, Selinka HC, Raab U, et al. (1993) Molecular pathogenesis of enterovirus-induced myocarditis: virus persistence and chronic inflammation. Intervirology 35: 140-151.
117. Chapman NM, Kim KS (2008) Persistent coxsackievirus infection: enterovirus persistence in chronic myocarditis and dilated cardiomyopathy. Curr Top Microbiol Immunol 323: 275-292.
118. Chu PY, Ke GM, Chen YS, Lu PL, Chen HL, et al. (2010) Molecular epidemiology of Coxsackievirus B3. Infection Genetics and Evolution 10: 777-784.
119. Gullberg M, Tolf C, Jonsson N, Mulders MN, Savolainen-Kopra C, et al. (2010) Characterization of a Putative Ancestor of Coxsackievirus B5. Journal of Virology 84: 9695-9708.
120. Leitch ECM, Cabrerizo M, Cardosa J, Harvala H, Ivanova OE, et al. (2010) Evolutionary Dynamics and Temporal/Geographical Correlates of Recombination in the Human Enterovirus Echovirus Types 9, 11, and 30. Journal of Virology 84: 9292-9300.
121. Knowles NJ, Hovi T, Hyypia T, King AM, Lindberg AM, et al. (2012) *Picornaviridae*. In: King AM, Adams MJ, Carstens EB, Lefkowitz EJ, editors. Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. Oxford, UK: Elsevier. pp. 855-880.
122. Escribano-Romero E, Jimenez-Clavero MA, Ley V (2000) Swine vesicular disease virus. Pathology of the disease and molecular characteristics of the virion. Anim Health Res Rev 1: 119-126.
123. Lin F, Mackay DK, Knowles NJ (1998) The persistence of swine vesicular disease virus infection in pigs. Epidemiol Infect 121: 459-472.
124. Tokarz R, Firth C, Madhi SA, Howie SR, Wu W, et al. (2012) Worldwide emergence of multiple clades of enterovirus 68. Journal of General Virology 93: 1952-1958.
125. Linsuwanon P, Puenpa J, Suwannakarn K, Auksornkitti V, Vichiwattana P, et al. (2012) Molecular epidemiology and evolution of human enterovirus serotype 68 in Thailand, 2006-2011. Plos One 7: e35190.
126. Hollinger F, Emerson SU (2007) Hepatitis A virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 911-947.

127. Harvala H, Robertson I, McWilliam Leitch EC, Benschop K, Wolthers KC, et al. (2008) Epidemiology and clinical associations of human parechovirus respiratory infections. Journal of Clinical Microbiology 46: 3446-3453.
128. Calvert J, Chieochansin T, Benschop KS, McWilliam Leitch EC, Drexler JF, et al. (2010) Recombination dynamics of human parechoviruses: investigation of type-specific differences in frequency and epidemiological correlates. Journal of General Virology 91: 1229-1238.
129. Faria NR, de Vries M, van Hemert FJ, Benschop K, van der Hoek L (2009) Rooting human parechovirus evolution in time. Bmc Evolutionary Biology 9.
130. Bangari DS, Pogranichniy RM, Gillespie T, Stevenson GW (2010) Genotyping of Porcine teschovirus from nervous tissue of pigs with and without polioencephalomyelitis in Indiana. J Vet Diagn Invest 22: 594-597.
131. Chiu SC, Hu SC, Chang CC, Chang CY, Huang CC, et al. (2012) The role of porcine teschovirus in causing diseases in endemically infected pigs. Veterinary Microbiology 161: 88-95.
132. Cano-Gomez C, Palero F, Buitrago MD, Garcia-Casado MA, Fernandez-Pinero J, et al. (2011) Analyzing the genetic diversity of teschoviruses in Spanish pig populations using complete VP1 sequences. Infection Genetics and Evolution 11: 2144-2150.
133. Wang B, Tian ZJ, Gong DQ, Li DY, Wang Y, et al. (2010) Isolation of serotype 2 porcine teschovirus in China: evidence of natural recombination. Veterinary Microbiology 146: 138-143.
134. Cano-Gomez C, Garcia-Casado MA, Soriguer R, Palero F, Jimenez-Clavero MA (2012) Teschoviruses and sapeloviruses in faecal samples from wild boar in Spain. Veterinary Microbiology.
135. Schwartz O, Albert ML (2010) Biology and pathogenesis of chikungunya virus. Nat Rev Microbiol 8: 491-500.
136. Ozden S, Huerre M, Riviere JP, Coffey LL, Afonso PV, et al. (2007) Human muscle satellite cells as targets of Chikungunya virus infection. Plos One 2: e527.
137. Griffin DE (2007) Alphaviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1023-1067.
138. Cherian SS, Walimbe AM, Jadhav SM, Gandhe SS, Hundekar SL, et al. (2009) Evolutionary rates and timescale comparison of Chikungunya viruses inferred from the whole genome/E1 gene with special reference to the 2005-07 outbreak in the Indian subcontinent. Infect Genet Evol 9: 16-23.
139. Russell RC (2002) Ross River virus: ecology and distribution. Annu Rev Entomol 47: 1-31.
140. Charles PC, Walters E, Margolis F, Johnston RE (1995) Mechanism of neuroinvasion of Venezuelan equine encephalitis virus in the mouse. Virology 208: 662-671.

141. Jackson AC, SenGupta SK, Smith JF (1991) Pathogenesis of Venezuelan equine encephalitis virus infection in mice and hamsters. Vet Pathol 28: 410-418.
142. Weaver SC, Ferro C, Barrera R, Boshell J, Navarro JC (2004) Venezuelan equine encephalitis. Annu Rev Entomol 49: 141-174.
143. Weaver SC, Rico-Hesse R, Scott TW (1992) Genetic diversity and slow rates of evolution in New World alphaviruses. Curr Top Microbiol Immunol 176: 99-117.
144. Weaver SC, Barrett ADT (2004) Transmission cycles, host range, evolution and emergence of arboviral disease. Nature Reviews Microbiology 2: 789-801.
145. Kubes V, Rios FA (1939) The Causative Agent of Infectious Equine Encephalomyelitis in Venezuela. Science 90: 20-21.
146. Liu C, Voth DW, Rodina P, Shauf LR, Gonzalez G (1970) A comparative study of the pathogenesis of western equine and eastern equine encephalomyelitis viral infections in mice by intracerebral and subcutaneous inoculations. J Infect Dis 122: 53-63.
147. Lockart RZ, Jr. (1963) Production of an Interferon by L Cells Infected with Western Equine Encephalomyelitis Virus. J Bacteriol 85: 556-566.
148. Meyer KF, Haring CM, Howitt B (1931) The Etiology of Epizootic Encephalomyelitis of Horses in the San Joaquin Valley, 1930. Science 74: 227-228.
149. Dent PB, Olson GB, Good RA, Rawls WE, South MA, et al. (1968) Rubella-virus/leucocyte interaction and its role in the pathogenesis of congenital syndrome. The Lancet 291: 291-293.
150. Chantler JK, Tingle AJ, Petty RE (1985) Persistent rubella virus infection associated with chronic arthritis in children. N Engl J Med 313: 1117-1123.
151. Chantler JK, da Roza DM, Bonnie ME, Reid GD, Ford DK (1985) Sequential studies on synovial lymphocyte stimulation by rubella antigen, and rubella virus isolation in an adult with persistent arthritis. Ann Rheum Dis 44: 564-568.
152. Nathanson N, Griffin DE (2007) Virus-induced immunosuppression. In: Nathanson N, Ahmed R, Biron CA, Brinton MA, Gonzalez-Scarano F et al., editors. Viral Pathogenesis and Immunity. 2nd ed. Oxford, UK: Elsevier. pp. 99-110.
153. Hobman T, Chantler J (2007) Rubella virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1069-1100.
154. Buchmeier MJ, de la Torre JC, Peters CJ (2007) *Arenaviridae*: The viruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1791-1827.
155. Russier M, Reynard S, Tordo N, Baize S (2012) NK cells are strongly activated by Lassa and Mopeia virus-infected human macrophages in vitro but do not mediate virus suppression. Eur J Immunol 42: 1822-1832.

156. Albarino CG, Palacios G, Khristova ML, Erickson BR, Carroll SA, et al. (2010) High diversity and ancient common ancestry of lymphocytic choriomeningitis virus. Emerg Infect Dis 16: 1093-1100.
157. Lipkin WI, Briese T (2007) *Bornaviridae*. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1829-1851.
158. Lipkin WI, Briese T, Hornig M (2011) Borna disease virus - fact and fantasy. Virus Res 162: 162-172.
159. Staeheli P, Sauder C, Hausmann J, Ehrensperger F, Schwemmle M (2000) Epidemiology of Borna disease virus. Journal of General Virology 81: 2123-2135.
160. Rackova S, Janu L, Kabickova H (2010) Borna disease virus (BDV) circulating immunocomplex positivity in addicted patients in the Czech Republic: a prospective cohort analysis. BMC Psychiatry 10: 70.
161. Richt JA, Pfeuffer I, Christ M, Frese K, Bechter K, et al. (1997) Borna disease virus infection in animals and humans. Emerg Infect Dis 3: 343-352.
162. Valbuena G, Walker DH (2006) The endothelium as a target for infections. Annu Rev Pathol 1: 171-198.
163. Schmaljohn CS, Nichol ST (2007) *Bunyaviridae*. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1741-1789.
164. Klingstrom J, Heyman P, Escutenaire S, Sjolander KB, De Jaegere F, et al. (2002) Rodent host specificity of European hantaviruses: evidence of Puumala virus interspecific spillover. J Med Virol 68: 581-588.
165. Ramsden C, Melo FL, Figueiredo LM, Holmes EC, Zanotto PM (2008) High rates of molecular evolution in hantaviruses. Molecular Biology and Evolution 25: 1488-1492.
166. Connolly-Andersen AM, Douagi I, Kraus AA, Mirazimi A (2009) Crimean Congo hemorrhagic fever virus infects human monocyte-derived dendritic cells. Virology 390: 157-162.
167. Carroll SA, Bird BH, Rollin PE, Nichol ST (2010) Ancient common ancestry of Crimean-Congo hemorrhagic fever virus. Molecular Phylogenetics and Evolution 55: 1103-1110.
168. Pepin M, Bouloy M, Bird BH, Kemp A, Paweska J (2010) Rift Valley fever virus(Bunyaviridae: Phlebovirus): an update on pathogenesis, molecular epidemiology, vectors, diagnostics and prevention. Vet Res 41: 61.
169. Gommet C, Billecocq A, Jouvion G, Hasan M, Zaverucha do Valle T, et al. (2011) Tissue tropism and target cells of NSs-deleted rift valley fever virus in live immunodeficient mice. PLoS Negl Trop Dis 5: e1421.
170. Bird BH, Khristova ML, Rollin PE, Ksiazek TG, Nichol ST (2007) Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. Journal of Virology 81: 2805-2816.

171. Bird BH, Githinji JW, Macharia JM, Kasiiti JL, Muriithi RM, et al. (2008) Multiple virus lineages sharing recent common ancestry were associated with a Large Rift Valley fever outbreak among livestock in Kenya during 2006-2007. Journal of Virology 82: 11152-11166.
172. Collao X, Palacios G, Sanbonmatsu-Gamez S, Perez-Ruiz M, Negredo AI, et al. (2009) Genetic diversity of Toscana virus. Emerg Infect Dis 15: 574-577.
173. Zehender G, Bernini F, Delogu M, Cusi MG, Rezza G, et al. (2009) Bayesian skyline plot inference of the Toscana virus epidemic: a decline in the effective number of infections over the last 30 years. Infect Genet Evol 9: 562-566.
174. Palese P, Shaw ML (2007) *Orthomyxoviridae*: The viruses and their replication. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1647-1689.
175. Wright PF, Neumann G, Kawaoka Y (2007) Orthomyxoviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1691-1740.
176. Chen R, Holmes EC (2010) Hitchhiking and the population genetic structure of avian influenza virus. J Mol Evol 70: 98-105.
177. Bahl J, Vijaykrishna D, Holmes EC, Smith GJ, Guan Y (2009) Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. Virology 390: 289-297.
178. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, et al. (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459: 1122-1125.
179. Vijaykrishna D, Bahl J, Riley S, Duan L, Zhang JX, et al. (2008) Evolutionary dynamics and emergence of panzootic H5N1 influenza viruses. Plos Pathogens 4: e1000161.
180. Chen RB, Holmes EC (2008) The evolutionary dynamics of human influenza B virus. Journal of Molecular Evolution 66: 655-663.
181. Gatherer D (2010) Tempo and mode in the molecular evolution of influenza C. PLoS Curr 2: RRN1199.
182. Beineke A, Puff C, Seehusen F, Baumgartner W (2009) Pathogenesis and immunopathology of systemic and nervous canine distemper. Vet Immunol Immunopathol 127: 1-18.
183. Griffin DE (2007) Measles virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1551-1585.
184. Harder TC, Kenter M, Vos H, Siebelink K, Huisman W, et al. (1996) Canine distemper virus from diseased large felids: biological properties and phylogenetic relationships. Journal of General Virology 77 (Pt 3): 397-405.

185. Monne I, Fusaro A, Valastro V, Citterio C, Pozza MD, et al. (2011) A distinct CDV genotype causing a major epidemic in Alpine wildlife. Veterinary Microbiology 150: 63-69.
186. Rima BK, Duprex WP (2006) Morbilliviruses and human disease. J Pathol 208: 199-214.
187. Pomeroy LW, Bjornstad ON, Holmes EC (2008) The evolutionary and epidemiological dynamics of the paramyxoviridae. Journal of Molecular Evolution 66: 98-106.
188. Saitoh M, Takeda M, Gotoh K, Takeuchi F, Sekizuka T, et al. (2012) Molecular evolution of hemagglutinin (H) gene in measles virus genotypes D3, D5, D9, and H1. Plos One 7: e50660.
189. Karron R, Collins P (2007) Parainfluenza viruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1497-1526.
190. Beck ET, He J, Nelson MI, Bose ME, Fan J, et al. (2012) Genome sequencing and phylogenetic analysis of 39 human parainfluenza virus type 1 strains isolated from 1997-2010. Plos One 7: e46048.
191. Hviid A, Rubin S, Muhlemann K (2008) Mumps. Lancet 371: 932-944.
192. Carbone K, Rubin S (2007) Mumps virus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1527-1550.
193. Papenburg J, Boivin G (2010) The distinguishing features of human metapneumovirus and respiratory syncytial virus. Rev Med Virol 20: 245-260.
194. Collins PL, Crowe Jr. JE (2007) Respiratory syncytial virus and Metapneumovirus. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1601-1646.
195. Li J, Ren L, Guo L, Xiang Z, Paranhos-Baccala G, et al. (2012) Evolutionary dynamics analysis of human metapneumovirus subtype A2: genetic evidence for its dominant epidemic. Plos One 7: e34544.
196. de Graaf M, Osterhaus ADME, Fouchier RAM, Holmes EC (2008) Evolutionary dynamics of human and avian metapneumoviruses. Journal of General Virology 89: 2933-2942.
197. Easton AJ, Domachowske JB, Rosenberg HF (2004) Animal pneumoviruses: molecular genetics and pathogenesis. Clin Microbiol Rev 17: 390-412.
198. Tan L, Lemey P, Houspie L, Viveen MC, Jansen NJG, et al. (2012) Genetic Variability among Complete Human Respiratory Syncytial Virus Subgroup A Genomes: Bridging Molecular Evolutionary Dynamics and Epidemiology. Plos One 7.
199. van der Westhuizen B (1967) Studies on bovine ephemeral fever. I. Isolation and preliminary characterization of a virus from natural and experimentally produced cases of bovine ephemeral fever. Onderstepoort J Vet Res 34: 29-40.

200. St George TD (1988) Bovine ephemeral fever: a review. Trop Anim Health Prod 20: 194-202.
201. Nandi S, Negi BS (1999) Bovine ephemeral fever: a review. Comp Immunol Microbiol Infect Dis 22: 81-91.
202. Lyles D, Rupprecht C (2007) *Rhabdoviridae*. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1363-1408.
203. Bourhy H, Reynes JM, Dunham EJ, Dacheux L, Larrous F, et al. (2008) The origin and phylogeography of dog rabies virus. Journal of General Virology 89: 2673-2681.
204. Ming P, Yan J, Rayner S, Meng S, Xu G, et al. (2010) A history estimate and evolutionary analysis of rabies virus variants in China. Journal of General Virology 91: 759-764.
205. Talbi C, Holmes EC, de Benedictis P, Faye O, Nakoune E, et al. (2009) Evolutionary history and dynamics of dog rabies virus in western and central Africa. Journal of General Virology 90: 783-791.
206. Gong W, Jiang Y, Za Y, Zeng Z, Shao M, et al. (2010) Temporal and spatial dynamics of rabies viruses in China and Southeast Asia. Virus Res 150: 111-118.
207. Kobayashi Y, Suzuki Y, Itou T, Ito FH, Sakai T, et al. (2011) Evolutionary history of dog rabies in Brazil. Journal of General Virology 92: 85-90.
208. Fooks AR, Brookes SM, Johnson N, McElhinney LM, Hutson AM (2003) European bat lyssaviruses: an emerging zoonosis. Epidemiol Infect 131: 1029-1039.
209. Fooks AR, McElhinney LM, Pounder DJ, Finnegan CJ, Mansfield K, et al. (2003) Case report: isolation of a European bat lyssavirus type 2a from a fatal human case of rabies encephalitis. J Med Virol 71: 281-289.
210. Amengual B, Bourhy H, Lopez-Roig M, Serra-Cobo J (2007) Temporal dynamics of European bat Lyssavirus type 1 and survival of Myotis myotis bats in natural colonies. Plos One 2: e566.
211. Davis PL, Holmes EC, Larrous F, Van der Poel WH, Tjornehoj K, et al. (2005) Phylogeography, population dynamics, and molecular evolution of European bat lyssaviruses. Journal of Virology 79: 10487-10497.
212. Schwartz-Cornil I, Mertens PP, Contreras V, Hemati B, Pascale F, et al. (2008) Bluetongue virus: virology, pathogenesis and immunity. Vet Res 39: 46.
213. Roy P (2007) Orbiviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1975-1997.
214. Carpi G, Holmes EC, Kitchen A (2010) The evolutionary dynamics of bluetongue virus. J Mol Evol 70: 583-592.
215. Tsai K, Karstad L (1973) The pathogenesis of epizootic hemorrhagic disease of deer: an electron microscopic study. Am J Pathol 70: 379-400.
216. Fletch AL, Karstad LH (1971) Studies on the pathogenesis of experimental epizootic hemorrhagic disease of white-tailed deer. Can J Comp Med 35: 224-229.

217. Biek R (2007) Evolutionary dynamics and spatial genetic structure of epizootic hemorrhagic disease virus in the eastern United States. Infect Genet Evol 7: 651-655.
218. Estes MK, Kapikian AZ (2007) Rotaviruses. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 1917-1974.
219. Nagaoka Y, Tatsumi M, Tsugawa T, Yoto Y, Tsutsumi H (2012) Phylogenetic and computational structural analysis of VP7 gene of group a human rotavirus G1P[8] strains obtained in Sapporo, Japan from 1987 to 2000. J Med Virol 84: 832-838.
220. Matthijnssens J, Heylen E, Zeller M, Rahman M, Lemey P, et al. (2010) Phylodynamic Analyses of Rotavirus Genotypes G9 and G12 Underscore Their Potential for Swift Global Spread. Molecular Biology and Evolution 27: 2431-2436.
221. Lahon A, Walimbe AM, Chitambar SD (2012) Full genome analysis of group B rotaviruses from western India: Genetic relatedness and evolution. Journal of General Virology.
222. de la Torre JC, Giachetti C, Semler BL, Holland JJ (1992) High frequency of single-base transitions and extreme frequency of precise multiple-base reversion mutations in poliovirus. Proc Natl Acad Sci U S A 89: 2531-2535.
223. de la Torre JC, Wimmer E, Holland JJ (1990) Very high frequency of reversion to guanidine resistance in clonal pools of guanidine-dependent type 1 poliovirus. Journal of Virology 64: 664-671.
224. Cuevas JM, Gonzalez-Candelas F, Moya A, Sanjuan R (2009) Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo. Journal of Virology 83: 5760-5764.
225. Nobusawa E, Sato K (2006) Comparison of the mutation rates of human influenza A and B viruses. Journal of Virology 80: 3675-3678.
226. Parvin JD, Moscona A, Pan WT, Leider JM, Palese P (1986) Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. Journal of Virology 59: 377-383.
227. Stech J, Xiong X, Scholtissek C, Webster RG (1999) Independence of evolutionary and mutational rates after transmission of avian influenza viruses to swine. Journal of Virology 73: 1878-1884.
228. Sedivy JM, Capone JP, RajBhandary UL, Sharp PA (1987) An inducible mammalian amber suppressor: propagation of a poliovirus mutant. Cell 50: 379-389.
229. Zhang X, Rennick LJ, Duprex WP, Rima BK (2013) Determination of Spontaneous Mutation Frequencies in Measles Virus under Nonselective Conditions. Journal of Virology 87: 2686-2692.
230. Schrag SJ, Rota PA, Bellini WJ (1999) Spontaneous mutation rate of measles virus: direct estimation based on mutations conferring monoclonal antibody resistance. Journal of Virology 73: 51-54.

231. Suarez P, Valcarcel J, Ortin J (1992) Heterogeneity of the mutation rates of influenza A viruses: isolation of mutator mutants. Journal of Virology 66: 2491-2494.
232. Qian XM, Shen Q, Goderie SK, He WL, Capela A, et al. (2000) Timing of CNS cell generation: A programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. Neuron 28: 69-80.
233. van der Flier LG, Clevers H (2009) Stem cells, self-renewal, and differentiation in the intestinal epithelium. Annu Rev Physiol 71: 241-260.
234. Savage VM, Allen AP, Brown JH, Gillooly JF, Herman AB, et al. (2007) Scaling of number, size, and metabolic rate of cells with body size in mammals. Proc Natl Acad Sci U S A 104: 4718-4723.
235. Shorter RG, Titus JL, Divertie MB (1964) Cell Turnover in the Respiratory Tract. Dis Chest 46: 138-142.
236. Streicker DG, Lemey P, Velasco-Villa A, Rupprecht CE (2012) Rates of viral evolution are linked to host geography in bat rabies. Plos Pathogens 8: e1002720.
237. Maljkovic Berry I, Ribeiro R, Kothari M, Athreya G, Daniels M, et al. (2007) Unequal evolutionary rates in the human immunodeficiency virus type 1 (HIV-1) pandemic: the evolutionary rate of HIV-1 slows down when the epidemic rate increases. Journal of Virology 81: 10625-10635.
238. Virgin S (2007) Pathogenesis of viral infection. In: Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA et al., editors. Fields Virology. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins. pp. 327-388.
239. Woelk CH, Holmes EC (2002) Reduced positive selection in vector-borne RNA viruses. Molecular Biology and Evolution 19: 2333-2336.
240. Coffey LL, Vasilakis N, Brault AC, Powers AM, Tripet F, et al. (2008) Arbovirus evolution in vivo is constrained by host alternation. Proc Natl Acad Sci U S A 105: 6970-6975.
241. Weaver SC, Brault AC, Kang W, Holland JJ (1999) Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. Journal of Virology 73: 4316-4326.
242. Pybus OG, Rambaut A, Belshaw R, Freckleton RP, Drummond AJ, et al. (2007) Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. Molecular Biology and Evolution 24: 845-852.
243. Wertheim JO, Kosakovsky Pond SL (2011) Purifying selection can obscure the ancient age of viral lineages. Molecular Biology and Evolution 28: 3355-3365.
244. Greene IP, Wang E, Deardorff ER, Milleron R, Domingo E, et al. (2005) Effect of alternating passage on adaptation of sindbis virus to vertebrate and invertebrate cells. Journal of Virology 79: 14253-14260.
245. Holmes EC (2003) Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. Journal of Virology 77: 11296-11298.

246. Frederico LA, Kunkel TA, Shaw BR (1990) A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. Biochemistry 29: 2532-2537.
247. Holtz CM, Mansky LM (2013) Variation of HIV-1 mutation spectra among cell types. Journal of Virology 87: 5296-5299.
248. Salemi M, Lewis M, Egan JF, Hall WW, Desmyter J, et al. (1999) Different population dynamics of human T cell lymphotropic virus type II in intravenous drug users compared with endemically infected tribes. Proc Natl Acad Sci U S A 96: 13253-13258.
249. Vandamme AM, Bertazzoni U, Salemi M (2000) Evolutionary strategies of human T-cell lymphotropic virus type II. Gene 261: 171-180.
250. Middelboe M (2000) Bacterial Growth Rate and Marine Virus-Host Dynamics. Microb Ecol 40: 114-124.
251. Rabinovitch A, Fishov I, Hadas H, Einav M, Zaritsky A (2002) Bacteriophage T4 development in Escherichia coli is growth rate dependent. J Theor Biol 216: 1-4.
252. Scholle F, Li K, Bodola F, Ikeda M, Luxon BA, et al. (2004) Virus-host cell interactions during hepatitis C virus RNA replication: impact of polyprotein expression on the cellular transcriptome and cell cycle association with viral RNA synthesis. Journal of Virology 78: 1513-1524.
253. Feuer R, Whitton JL (2008) Preferential coxsackievirus replication in proliferating/activated cells: implications for virus tropism, persistence, and pathogenesis. Curr Top Microbiol Immunol 323: 149-173.
254. Honda M, Kaneko S, Matsushita E, Kobayashi K, Abell GA, et al. (2000) Cell cycle regulation of hepatitis C virus internal ribosomal entry site-directed translation. Gastroenterology 118: 152-162.
255. Nelson HB, Tang H (2006) Effect of cell growth on hepatitis C virus (HCV) replication and a mechanism of cell confluence-based inhibition of HCV RNA and protein expression. Journal of Virology 80: 1181-1190.
256. Kusov YY, Gosert R, Gauss-Muller V (2005) Replication and in vivo repair of the hepatitis A virus genome lacking the poly(A) tail. Journal of General Virology 86: 1363-1368.
257. Feuer R, Mena I, Pagarigan R, Slifka MK, Whitton JL (2002) Cell cycle status affects coxsackievirus replication, persistence, and reactivation in vitro. Journal of Virology 76: 4430-4440.
258. Kaminski A, Hunt SL, Patton JG, Jackson RJ (1995) Direct evidence that polypyrimidine tract binding protein (PTB) is essential for internal initiation of translation of encephalomyocarditis virus RNA. RNA 1: 924-938.
259. Marshman E, Booth C, Potten CS (2002) The intestinal epithelial stem cell. Bioessays 24: 91-98.
260. Bhardwaj RD, Curtis MA, Spalding KL, Buchholz BA, Fink D, et al. (2006) Neocortical neurogenesis in humans is restricted to development. Proceedings of the National Academy of Sciences of the United States of America 103: 12564-12568.

261. Ciarlet M, Schodel F (2009) Development of a rotavirus vaccine: clinical safety, immunogenicity, and efficacy of the pentavalent rotavirus vaccine, RotaTeq. Vaccine 27 Suppl 6: G72-81.
262. Zhang D, Lu J (2010) Enterovirus 71 vaccine: close but still far. Int J Infect Dis 14: e739-743.
263. Hay AJ, Gregory V, Douglas AR, Lin YP (2001) The evolution of human influenza viruses. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 356: 1861-1870.
264. Bull RA, Eden JS, Rawlinson WD, White PA (2010) Rapid evolution of pandemic noroviruses of the GII.4 lineage. Plos Pathogens 6: e1000831.

7. PUBLICATION ACKNOWLEDGMENT

This study has been submitted for publication at PLOS Pathogens with Siobain Duffy as a co-author. Both Siobain and I conceived and designed the experiments. I performed all of the experiments and analyses. I wrote the first draft of this manuscript and worked extensively with Siobain on manuscript revisions.