VIRUS EVOLUTION

Virus Evolution, 2016, 2(1): vew008

doi: 10.1093/ve/vew008
Research Article

Sendai virus intra-host population dynamics and host immunocompetence influence viral virulence during *in vivo* passage

José Peña,[1] Haiyin Chen-Harris,[1,†] Jonathan E. Allen,[1] Mona Hwang,[1] Maher Elsheikh,[1] Shalini Mabery,[1] Helle Bielefeldt-Ohmann,[2] Adam T. Zemla,[1] Richard A. Bowen,[3] Monica K. Borucki[1,*]

[1]Lawrence Livermore National Laboratory, Livermore, CA, USA; [2]Australian Infectious Diseases Research Centre, University of Queensland , Brisbane, Australia; and [3]Department of Biomedical Sciences, Colorado State University, Fort Collins, CO, USA

*Corresponding author: E-mail: borucki2@llnl.gov

†Present address: Haiyin Chen-Harris, Genentech, Inc., South San Francisco, CA, USA.

Abstract

In vivo serial passage of non-pathogenic viruses has been shown to lead to increased viral virulence, and although the precise mechanism(s) are not clear, it is known that both host and viral factors are associated with increased pathogenicity. Under- or overnutrition leads to a decreased or dysregulated immune response and can increase viral mutant spectrum diversity and virulence. The objective of this study was to identify the role of viral mutant spectra dynamics and host immunocompetence in the development of pathogenicity during *in vivo* passage. Because the nutritional status of the host has been shown to affect the development of viral virulence, the diet of animal model reflected two extremes of diets which exist in the global population, malnutrition and obesity. Sendai virus was serially passaged in groups of mice with differing nutritional status followed by transmission of the passaged virus to a second host species, guinea pigs. Viral population dynamics were characterized using deep sequence analysis and computational modeling. Histopathology, viral titer and cytokine assays were used to characterize viral virulence. Viral virulence increased with passage and the virulent phenotype persisted upon passage to a second host species. Additionally, nutritional status of mice during passage influenced the phenotype. Sequencing revealed the presence of several non-synonymous changes in the consensus sequence associated with passage, a majority of which occurred in the hemagglutinin-neuraminidase and polymerase genes, as well as the presence of persistent high frequency variants in the viral population. In particular, an N1124D change in the consensus sequences of the polymerase gene was detected by passage 10 in a majority of the animals. *In vivo* comparison of an 1124D plaque isolate to a clone with 1124N genotype indicated that 1124D was associated with increased virulence.

1. Introduction

Novel human viral pathogens are emerging with increasing frequency, however the mechanisms of viral adaptation to a new host are not well understood (Morse, et al. 2012; Anthony et al. 2013). The majority of these pathogens have an animal reservoir and an RNA genome. RNA viruses have a highly error-

prone polymerase used for genome replication, and individual RNA genomes within a species generally differ by many nucleotides (nts) from the consensus (average) genome sequence. Because of this, the term 'quasispecies' is often used to describe the mutant spectra of viral genomes that constitute an RNA viral species (Domingo and Holland 1997). Recent studies indicate that intra-host viral population diversity may be a prerequisite for a virulent infection (Vignuzzi, et al. 2006; Kuss, Etheredge, and Pfeiffer, 2008; Gutiérrez, Michalakis, and Blanc 2012) although the nature of the infection (acute versus chronic) and viral lifecycle (i.e. vector-borne viruses) also influence the role of intra-host viral diversity in infection outcome (Weaver, Rico-Hesse, and Scott 1992; Jerzak, et al. 2007; Andersen et al. 2015).

Serial passage of a virus in laboratory animals is often used to increase pathogenicity of the virus for that host in order to generate a laboratory model of the disease (Aronson, Herzog, and Jerrells 1995; Tiwari et al. 2006). For example, passage of the SARS virus in rats generated a viral variant with a point mutation in the spike protein, and this mutation increased the efficiency of interaction of the spike protein and the viral receptor on the rat cells (Nagata, et al. 2007). Interestingly, data from this same study showed that both host and viral factors are associated with increased virulence, as virulence was associated with increased age of the rodent as well as mutations in the viral genome.

The precise mechanism(s) by which *in vivo* passage leads to increased virulence are not clear and may differ between viral families, genera, and even species. *In vivo* passage outcome is shaped by numerous factors including viral strain, genotypic diversity, size of the inoculum, route of exposure, and by host factors such as onset and strength of immune response (Kuss, Etheredge, and Pfeiffer, 2008; Gutiérrez, Michalakis, and Blanc 2012; Morelli, et al. 2013). Decreased immunocompetence may allow infecting viruses to cause prolonged infections which may increase the size and depth of genetic diversity present in the intra-host population of RNA viruses (Domingo 1997). Immunocompetence of a host is determined by many factors such as age, pregnancy, and nutrient intake (Levander 1997; Beck, Levander, and Handy 2003; Beck 2007). Recent studies by Beck and others suggest that both malnutrition and obesity adversely affect the ability of a host to mount a robust innate and adaptive immune response to viral infections (Beck, Handy, and Levander 2004; Zaki, Akuta, and Akaike 2005; Karlsson and Beck 2009). For example, obese mice infected with influenza virus had a significantly higher mortality and increased lung pathology as compared with lean mice (Smith et al. 2007). Similarly, infection of selenium deficient mice with an avirulent strain of Coxsackievirus B3 resulted in evolution of the virus into a virulent form that caused cardiac inflammation in both selenium deficient mice and normal control mice (Beck et al. 1995; Beck, Williams-Toone, and Levander 2003).

Host nutritional status may influence the generation of variant viruses and thus the viral genotype by a number of mechanisms. Both selenium and vitamin E are antioxidants and function to control oxidative stress that is induced as part of the host response to infection. Although production of reactive oxygen species such as nitric oxide may play a role in microbial clearance, it may also act as a mutagen and therefore may expand the diversity of the mutant spectra by acting on the fidelity of the viral polymerases (Domingo 1997; Akaike et al. 2000). The effect of oxidative stress on viral evolution has been demonstrated in a variety of viruses (influenza, coxsackie, polio, Sendai, and human immunodeficiency virus) and hosts (mouse, human) (Akaike et al. 2000; Beck et al. 2001; Beck, Williams-Toone, and Levander 2003; Beck, Handy, and Levander 2004). Selenium deficiency has also been shown to impair the cellular and humoral immune response. Similarly, obesity leads to increased oxidative stress and immune dysregulation although through a different mechanism driven by chronic inflammation (Monteiro and Azevedo 2010; Kanneganti and Dixit 2012; Aroor and DeMarco 2014). In particular, obesity has been shown to reduce cytokine production and NK cell cytotoxicity (Akaike et al. 2000; Beck et al. 2001; Beck, Williams-Toone, and Levander 2003; Beck, Handy, and Levander 2004; Smith et al. 2007). The presence of excess dietary iron, which has a pro-oxidant effect *in vivo*, may also increase viral titer and virulence (Beck, Handy, and Levander 2004).

Despite these studies the role that nutritional status plays on driving viral evolution within host under various nutritional statuses and the ability to infect a different host species have not been addressed. In this study, we address this gap in knowledge by serial passage of Sendai virus (SeV) (genus *Respirovirus*, *Paramyxoviridae* family) in mice under three different dietary conditions: normal diet (ND), selenium-deficient diet (SD), and high fat diet (FD). Following the last serial passage in mice, individual bronchial alveolar lavage (BAL) containing SeV from the three dietary conditions was used to infect guinea pigs under normal dietary conditions. Lung histopathology, viral titers, and cytokine profiles were used to characterize viral virulence within the hosts.

Viral population dynamics were characterized using Illumina ultra-deep sequence analysis and computational modeling. SeV has a single-stranded, negative-sense RNA genome ~15–16 kb in length which codes for six structural proteins: two glycoproteins hemagglutinin-neuraminidase (HN) and fusion protein (F), a matrix protein (M), and three nucleocapsid proteins, Nucleoprotein (N), Phosphoprotein (P), and Large polymerase protein (L). The P gene which codes for the P protein also codes for an additional set of up to seven different nested polyproteins (P, V, W, C', C, Y1, and Y2) and may be referred to as 'P/V/C' (Lamb and Parks 2007). The HN glycoprotein plays critical role in mediating the virus entry into host cell and structural modeling was used to assess the impact mutations in this gene may have on protein function.

2. Methods and materials

2.1 Virus and cell lines

SeV strain 52 was obtained from ATCC (Manassas, VA; VR-105). Seed virus was prepared by inoculation of 10-day embyonated chicken eggs, harvesting allantoic fluid 2 days later and storing aliquots at −80°C.

2.2 SeV serial passage in vivo

Groups of 6-week-old C57BL/6 mice (Charles River; $n = 5$ per experimental group, $n = 4$ for mock infected group for each diet) were fed a ND, or selenium deficient diet for 7 weeks, or a high FD 15 weeks prior to infection. The ND chow, Global 18% Protein Rodent Diet 2018 (Teklad Diets, Madison, WI) contained 0.23 mg/kg Se, and 18% kcal from fat (soybean oil). The selenium deficient diet, Teklad Custom Diet, Selenium deficient TD.92163, contained 14.3 % kcal from fat (corn oil). The high FD (D12331, Research Diets, New Brunswick, NJ) contained 58 kcal% fat (coconut and soybean oils). Note, the selenium status was not assessed via glutathione peroxidase activity in any of the diet

groups prior to infection, and the diets varied in ingredients such as fat and carbohydrate source. Diets were staggered for each infection group so all mice were of the same approximate age when infected.

Mice were anesthetized with a mixture of ketamine and xylazine, and inoculated intranasally with 25 μl (1,000 plaque forming unit (PFU)) of stock virus (passage 1) or tracheal lavage fluid from infected mice. Mice were monitored daily for weight loss and visually inspected twice daily for signs of disease such as lethargy, ruffled fur, and hunching. Mice were euthanized 5 days post infection and virus was obtained from infected lung tissue via BAL (Aronson, Herzog, and Jerrells 1995). The infection protocol of intranasal inoculation and infection of the recipient group with pooled donor BAL (versus homogenized lung tissue) was designed to simulate of viral transmission via the respiratory route. Lavage samples were pooled within each treatment group with equal quantities from each mouse (100 μl) then diluted according to titration results to a concentration of $\sim 5 \times 10^3$ PFU/25 μl for inoculation of the next group. Thus the viral dose was held constant for subsequent *in vivo* passage for a total of ten serial passages. BAL samples were assayed for viral titer via plaque assay (described below), and lungs were fixed in buffered formalin for histopathology.

After mouse passage 10, virus was obtained from infected mouse lung tissue via BAL and used to infect three guinea pigs intranasally per diet group. The guinea pigs diets did not vary (all on ND) and were infected intranasally with 0.25 ml containing 1×10^4 PFU. At day 7-post infection the guinea pigs were euthanized and lung lavage and lung homogenate were collected.

2.3 Histopathology

Formalin-fixed tissues were embedded in paraffin, sectioned at 5 μm and sections stained with hematoxylin and eosin. An experienced veterinary pathologist evaluated tissues and score for pathology. Tissues were scored as follows: 0 = no change; 1 = minimal change (or possibly non-specific background or appearance compounded by lung collapse, i.e. not inflated); 2 = mild inflammation and/or pneumocyte hypertrophy; 3 = moderate inflammation and/or pneumocyte hypertrophy; 4 = marked inflammation and/or pneumocyte hypertrophy; 5 = severe inflammation and/or pneumocyte hypertrophy affecting <50% of lung lobe; 6 = severe inflammation and/or pneumocyte hypertrophy affecting >50% of lung lobe.

2.4 Plaque assays

Infectivity titers for both the seed virus and material recovered from infected mice were determined by plaque titration on Vero E6 or LLC-MK2 cells using an overlay supplemented with 1 μg TCPK trypsin/ml and expressed as plaque-forming units. SeV serially diluted in sterile phosphate-buffered saline (PBS) was inoculated onto the wells and incubated at 37°C for 45 min with agitation every 15 min. Wells were overlaid with 3 ml of 2× MEM (Life Technologies, Gibco) with 100 IU/ml penicillin/100 μg/ml streptomycin (Gibco), 2% BSA (Gibco), 2 μg/ml of TCPK trypsin (Sigma), and 1.6% agarose. Plaques were visualized at day 3 by staining with crystal violet in 20% methanol processed as described earlier.

For SeV single plaque amplification, LLC-MK2 cell monolayers were seeded in six-well plates and infected with 100 ul of BAL from infected mice and incubated at 37°C for 1 h. Wells were overlaid with 3 ml of 2× MEM (Life Technologies, Gibco) with 100 IU/ml Penicillin/100 μg/ml Streptomycin (Gibco), 2% BSA (Gibco), 2 μg/ml of TPCK trypsin (Sigma), and 2% agarose. Live plaques were visualized by adding 300 μl MTT solution (1/10 volume of agarose plug: 5 mg/ml tetrazolium sallt, MTT (Sigma) in PBS filtered. Single plaques were collected and resuspended in 1 ml of unsupplemented DMEM (Life Technologies) overnight. Isolated plaques were then amplified in LLC-MK2 cells and supernatant harvested and stored at −80°C on day 7.

2.5 Primer design

Approximately 15.1 kb of the virus genome was amplified using degenerate primers (Supplementary Table S1). Primers were designed to be as sensitive to target strain variants as possible, while still being specific enough to not cross-react with non-targets as described previously in Slezak et al. (2003) and Borucki et al. (2013). The Taqman primers/probe used for quantification of viral genomes was SeV511F: CAGAATGGCTGTTTGGACCT, SeV640R: CTTCACCAGCACAATCCAGA, SeV604probe: GGTAT CCTGCATGCCTAGGA.

2.6 Preparation and sequencing of SeV RNA

Total RNA from BAL or single plaque isolated virus supernatant was extracted using TRIzol LS Reagent (Invitrogen) following the manufacturer's protocol. Reverse transcription was performed using random hexamers and the Superscript III RT reverse transcriptase kit (Invitrogen). Viral cDNA templates were amplified using the Phusion polymerase kit (New England BioLabs, Ipswich, MA), following manufacturer's instructions. PCR conditions consisted of 98°C for 30 s, followed by 40 cycles of 98°C for 15 s, 64°C for 20 s, and 72°C for 1.2 min. The final cycle was 72°C for 10 min. PCR products were prepared for sequencing using the QIAquick PCR Purification kit (Qiagen, Valencia, CA).

The pooled PCR products from each sample were sequenced using Illumina paired-end sequencing in which the two read pairs overlap and only the overlapping regions of the read-pairs were used to make variant calls (Chen-Harris et al. 2013). The samples were sequenced in multiplex (four lanes, nine to ten samples per lane, 150 PE) using the Illumina Hi Seq platform at the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley.

2.7 Determination of processing error rate

Sequencing control plasmids were generated to determine the error rate of the PCR and Illumina sequencing as described previously in Chen-Harris et al. (2013). To gauge sensitivity of the protocol, dilutions of control plasmids containing inserts from SeV, rabies virus (RABV) and bovine coronavirus (BCoV) were combined at following ratios 100,000 SeV: 1,000 RABV: 10 BCoV. Plasmid copy number for each plasmid control was estimated via spectrophotometer and Taqman assay (Borucki et al. 2013a,b) Viral inserts present in the plasmid mixtures were amplified using M13 primers and Phusion PCR as described previously. This process was done in replicate to generate two independent samples for use as Illumina sequencing controls.

Detecting BCoV seqeuences in the mixture would mean a variant detection sensitivity of 1 in 10,000 copies (0.01%). However, amplification efficiency was affected by slight differences in plasmid insert size between the controls (SeV 877 bp, RABV 1138 bp, and BCoV 1003 bp) causing the insert sequence from the smaller inserts to have slightly more coverage as compared with the larger insert (i.e. RABV). In the resulting

sequencing data, the SeV plasmid insert had ~1,000× more coverage as compared with the RABV insert, about 10× more than expected ratio of 100×. However the ratio of RABV to BCoV insert was about 88.5×, much closer to the expected value of 100×. These data indicate that very rare sequences (BCoV insert was present at 0.01% in the mixture) are readily detectable using this ultradeep sequencing (UDS) methodology.

2.8 Sequencing data analysis

Read mapping and rare variant detection was performed as described previously in Chen-Harris et al. (2013) and Borucki et al. (2013). The open source read mapping software SHRiMP2 (David et al. 2011), was used to map reads using GenBank SeV Z strain (GI: 546225824) as a reference sequence which was shown to have high read mapping sensitivity (Holtgrewe et al. 2011) and was chosen for the tool's ability to map as many reads as possible in the face of individual errors within each read. To differentiate rare variants from sequencing errors, methodologies were developed to measure and control for sequencing and PCR errors and described previously in Chen-Harris et al. (2013). Error models have been refined for Illumina sequencing platform to allow detection of ultra-rare mutations within the viral population (Chen-Harris et al. 2013; Borucki et al. 2013). The Illumina reads were submitted to the GenBank Sequence Read Archive (BioProject PRJNA314501).

Software was created to link Illumina reads aligned to regions nearby on the reference and construct viral haplotypes. The resulting haplotypes were experimentally verified using cloning and Sanger sequencing. The haplotype building software linked non-overlapping read pairs from Illumina data and thus allowed linkage of mutations that are separated by as much as the length of both read pairs (300 bp). Fisher's exact test was used to determine statistical significance for each linkage.

The average per position entropy of the HN gene was calculated to track changes in intra-host viral diversity. Entropy is calculated as $\sum_j \sum_{i=1}^{4} f_i * \log(f_i)/N$ where entropy is calculated for each position j in the gene and f_i is the relative frequency of each of the 4 nts (A,C,T,G).

2.9 Protein modeling

To help characterize an effect of identified mutations on virus phenotypic changes a homology-based structural model of HN protein from SeV strain Z (GI:546225829) was constructed using the AS2TS system (Zemla et al. 2005). The HN protein of human parainfluenza 3 virus (hPIV3) was identified as the closest Protein Data Bank (PDB) structural template for the modeling. It was experimentally solved at the resolution of 1.65 angstroms (Å) (PDB entry: 4MZA) (Xu et al. 2013), and the level of sequence identity between HN from Sendai and hPIV3 is 55%. Examples of other identified structural homologs that were leveraged during the modeling process include: HN from Newcastle Disease virus (NDV) (PDB entry: 1E8U; resolution: 2.00 Å; sequence identity: 27%) (Connaris et al. 2002), HN from Simian virus 5 (Sv5) (4JF7; 2.50 Å; 24%) (Welch et al. 2013), Nipah virus (NiV) (3D11; 2.31; 22%), Hendra virus (HeV) (2VSK; 2.00; 20%), Measles virus (MeV) (2ZB6; 2.60; 16%). Using available templates a set of initial structural models was created. To assess regions of sequence-structure conservation or variability in the modeled Sendai HN protein an extensive search for similar fragments in other proteins from PDB was performed using the StralSV algorithm (Zemla et al. 2011), which identifies protein fragments that exhibit structural similarities despite low primary amino acid sequence similarity. Results from these searches were used for modeling of missing loop or termini regions in the constructed preliminary models. For the construction of the final structural model the coordinates from the PDB chain 4mza_B were used as a primary template. The conformation of side-chain atoms was predicted using SCWRL (Krivov, Shapovalov, and Dunbrack 2009) when residue-residue correspondences did not match. Residues that were identical in the template and HN protein were copied from the template onto the model. The structural and stereochemical quality of the model was checked using a contact-dot algorithm in the MolProbity software package (Chen et al. 2010), and the final constructed model was finished with relaxation using UCSF Chimera (Pettersen et al. 2004).

Sequence variability/conservation scores were calculated from amino acids frequencies per positions in sequence and structure-based alignments using Sum-of-Pairs algorithm with BLOSUM62 as an amino acids substitution matrix. Structure variability/conservation analysis was performed by StralSV (Zemla et al. 2011) applied to all sixty-one currently available variants (PDB chains) of structural models of paramyxovirus HNs deposited in PDB (date: 4 August 2015). Conserved sialic acid binding site residues were identified based on results published by Connaris et al. (2002) and from alignments calculated between HN proteins.

To assess the possibility that the mutation position could interfere with interfaces of the tetramer formation, we evaluated all three possible dimer-of-dimers assemblies of heads conformations relative to the four-helix bundle stalk ('four-heads-up', 'four-heads-down', and 'two-heads-up/two-heads-down') identified by Welch et al. (2013). The regions of interfaces between subunits were defined by identifying amino acids that are from different chains but have atoms in a distance closer than 3 Å.

2.10 Cytokine/chemokine expression

Quantitative PCR was used to measure expression of host cytokine genes in lung homogenate using Taqman Gene Expression Assays (Life Technologies) for the mouse cytokines and QuantiTect SYBR Green PCR (Invitrogen) for the guinea pig cytokines. Total RNA was extracted using TRIzol LS Reagent (Invitrogen) following the manufacturer's protocol. Reverse transcription was performed using Oligo dT primers and the Superscript III RT reverse transcriptase kit (Invitrogen) following the manufacturer's protocol. Primer sequences for amplification of guinea pig mRNA were obtained from Scott and Aronson (2008) and Jeevan et al. (2011), GAPDH was used as endogenous control for both the mouse and guinea pig cytokine measurements.

Cytokine expression was also measured using a cytokine multiplex assay. Lung homogenates were processed according to manufacturer instructions and then analyzed using a Bio-Plex 200 system (Bio-Rad, Hercules, CA). Briefly, tissue homogenates were centrifuged for 10 min at 1,000 rpm at 4°C to remove cellular debris. Aliquots of the cleared supernatants were transferred into 96-well round bottom plates and processed for use on the Bio-Plex system. The cytokines/chemokines were coupled to cytokine/chemokine specific multi-plex beads (R&D Systems, INC., Minneapolis, MN). This panel measures the concentrations of cytokines and chemokines IL-1β, IL-4, IL-6, IL-10, IFN-γ, CCL2 (MCP-1), CCL3 (MIP1α), CCL4 (MIP1β), CCL5 (RANTES), CXCL10 (IP-10), and tumor necrosis factor-α (TNF-α).

3. Identification of clones containing variant genotypes

BAL samples in which high frequency variants were detected by sequencing were further analyzed using plaque assays and Sanger sequencing. For each sample plaqued, ten plaques were picked for Sanger sequencing of ten genome sites containing the high-frequency mutations. Following RNA extraction, variable regions derived from each plaque were amplified by RT-PCR using PCR primers sets spanning the mutation sites.

3.1 Infection of mice with variant genotypes

Six week old C57BL/6 mice (ND) were infected intranasal inoculation with 1×10^4 PFU/ml virus. Inoculum consisted of either virus derived from a single plaque/clone, a mixture of clones in equal ratios, seed stock, or no virus (media only). A group of six mice were infected with each inoculum. Mice were euthanized on day 7 post infection and lung tissue was collected for histopathology.

3.2 Statistical analysis

All chemokine/cytokine expression analysis was carried out using an unpaired Student's *t*-test and changes in expression levels were consider significant if $P < 0.05$. nt/amino acid substitution analysis was performed using Mann-Whitney U test and changes in the frequency of nt substitution or amino acid change was consider significant if $P < 0.05$. Statistical analysis was performed using Prism version 6.0 g (GraphPad Software, La Jolla, CA).

4. Results

4.1 In vivo serial passage of SeV

SeV was serially passaged ten times in C57BL/six mice under three dietary conditions; ND, FD, or selenium deficient diet (SD). Mouse weight loss was determined as a measure of morbidity following intranasal infection with each serial passage. Infection with SeV led to statistically significant weight loss under all dietary conditions (Fig. 1), although the percent body weight loss between the diet groups was not significantly different.

Lung tissues were collected on day 7 and histopathology scores were determined as a measure of virulence in the lungs of mice infected with SeV passage 1 (P1), P4, P7, and P10. Within diet groups, only mice receiving a ND were shown to have a statistically significant increase in histopathology scores, while histopathology scores of mice under FD and SD diet conditions trended up (Fig. 2A). Analysis of histopathology scores between diet groups demonstrated a statistically significant increase between P1 of SD mice compared with P1 of FD mice. The increase in histopathology scores was not associated with increased in viral loads following serial passage. The viral titer in BAL for each diet group averaged $\sim 1.6 \times 10^5$ PFU/ml and there were no significant difference between groups. However, viral titers within diet groups trended lower with passage and there was a statistically significant decrease in SeV titers between P1 and P10 in SD mice (Fig. 2B).

In the secondary host transmission study, unpassaged SeV seed stock, PBS (Mock) and the SeV virus obtained from P10 BAL of mice in each diet group, were used to infect a different host, guinea pigs, under ND conditions (three animals per treatment group). None of the guinea pigs lost weight or displayed any signs of clinical infection; however, guinea pigs infected with

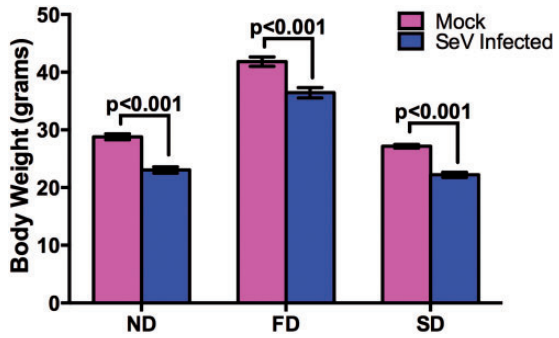

Figure 1. Weight loss following SeV infection under different dietary conditions. C57BL/6 mice were fed different diets for 7 weeks (ND and SD) or 15 weeks (FD) before infection with SeV. Mice were euthanized 5 days post infection.

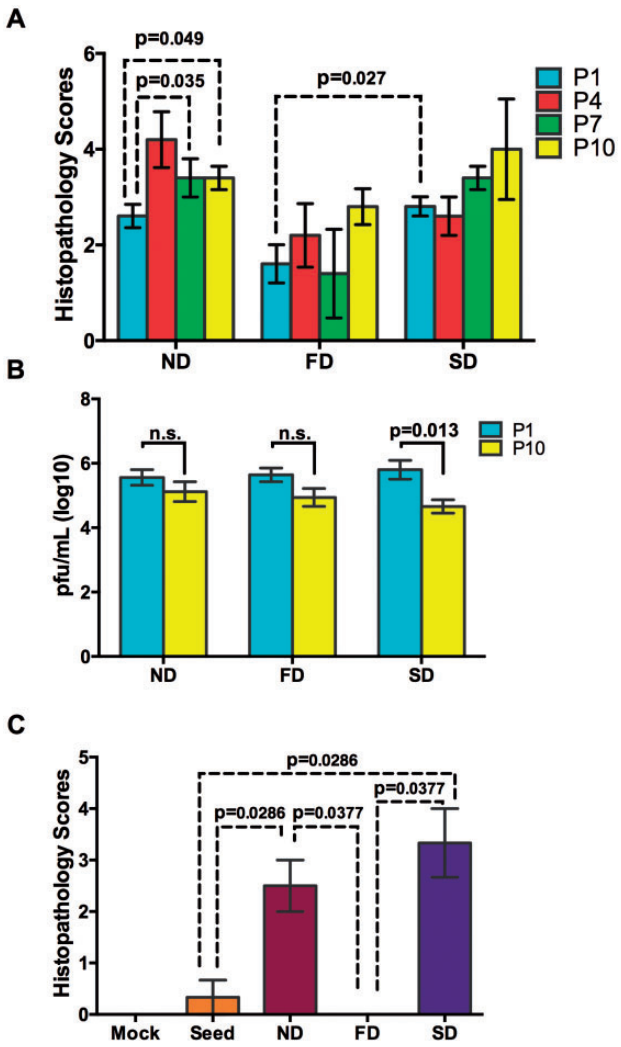

Figure 2. Intra- and Inter-host lung histopathology scores and viral titers following serial passage of SeV. **(A)** Lung histopathology scores for mice infected with passage 1, 4, 7, and 10 for all diet groups ($n = 5$/group). **(B)** Viral titers from mouse BAL collected from P1 and P10 were determined by plaque assay. **(C)** Guinea pig lung histopathology scores following inoculation with seed virus or P10 BAL from ND, FD or SD mice ($n = 6$/group). Values are means ± SEM *P*-values were determined by unpaired *t*-test and $P < 0.05$ were considered significant.

P10 BAL from the ND and SD mouse groups had significantly higher lung histopathology scores than those infected with BAL from the FD mouse group or seed stock, which showed virtually no histopathology changes (Fig. 2C).

4.2 Altered innate immune response in mice infected with serially passaged SeV under different dietary conditions

Lung homogenates from all three diet groups were assayed for changes in cytokine gene expression according to diet group and passage number. Gene expression dynamics of cytokines *Ccl3, Ifng, Il4, Il6*, and *Tnfα* was determined using qRT-PCR. As shown in Figure 3, basal levels of gene expression in uninfected mice were altered between FD and SD mice compared with ND mice (Fig. 3A). Following SeV infection gene expression of *Ccl3, Ifng, Il4, Il6*, and *Tnfα* was suppressed in mice receiving a fat or selenium deficient diet compared with mice under ND conditions. Although most genes were suppressed under the various diet conditions, SD mice had a several fold increase in the expression of *Il4*, a Th2 response gene, compared with ND mice (Fig. 3B–D). Additionally, analysis of gene expression changes in guinea pigs infected with P10 virus from the various diet groups yielded little to no differences in the expression levels of *Ccl2, Ifng*, or *Il-1b* compared with seed virus (Supplementary Fig. S1). However, the expression levels of *Tnfα* were between 1.5- and 2.0-fold higher in guinea pigs infected with the P10 virus inoculum from the different mouse diets groups compared with seed virus.

Cytokine/chemokine protein expression was assayed to determine whether the observed changes in innate immune response gene expression correlated with protein expression. An eleven-plex, Luminex bead-based screening assay was used to determine protein levels of CCL2, CCL3, CCL5, CXCL2, CXCL10, IL-1b, IL-4, IL6, IL10, IFNγ, and TNFα in lung homogenates of mice infected with SeV P1 or P10. Unlike the gene expression analysis there were no pronounced changes in protein expression levels within diet groups (Fig. 4). Moreover, similar expression profiles were observed within diet groups between SeV P1 and P10. However significant changes in the protein expression profiles of P1 and P10 were detected between diet groups (Supplementary Fig. S2). SD mice infected with SeV P1 demonstrated a statistically significant lower expression level of CCL2, CCL3, CXCL2, CXCL10, IL-6, IL-10, IFNγ, and TNFα compared with P1 of ND or FD mice. Infection of SD mice with P10 on the other hand resulted in statistically lower expression of CCL2 and IL-6 compared with P10 from ND or FD mice.

4.3 PCR amplification of viral cDNA

Although SeV primers were designed to be as sensitive to variant SeV genomes as possible, amplification of some regions of viral genome proved to be extremely difficult for many of the samples, with selenium deficient diet passage 1 samples showing the poorest amplification results. Many different combinations of primers were used in order to amplify as much of the viral genome as possible (Supplementary Table S1). To determine if the poor PCR amplification results were due to low numbers of viral genomes present in each sample, viral titers were determined by both plaque assay and Taqman PCR. The average number of positive PCR results per region (best possible score is two primers per region) was compared with the viral titer (PFU/ml) and to the C_t (cycle threshold) average of two replicates. The average C_t had a negative correlation with positive PCR results ($r = -0.97$), however there was no correlation between PCR outcome and viral titer as determined by plaque assay ($r = -0.35$).

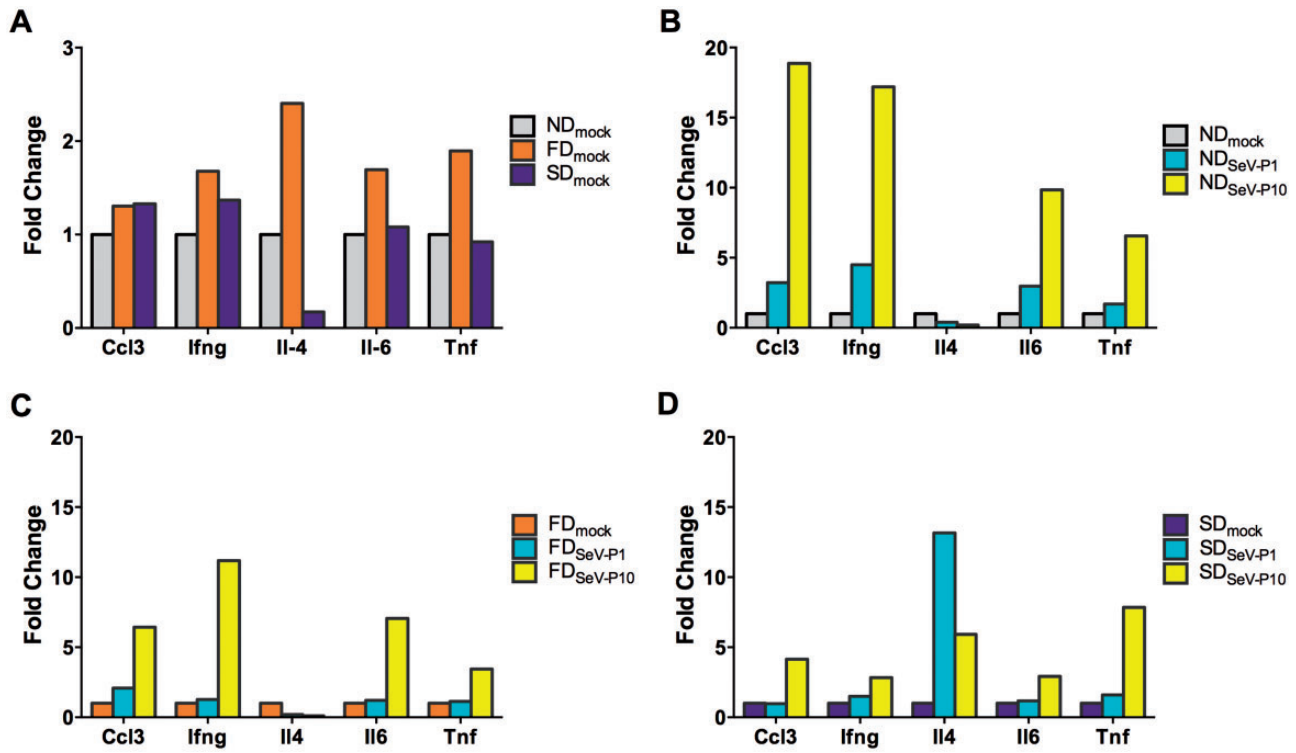

Figure 3. Lung cytokine gene expression profile according to diet following serial passage in mice. Lung homogenates from ND, FD, and SD mice were analyzed for Ccl3, Ifng, IL-4, IL-6, and TNF-α gene expression based on **(A)** Mock Infected, **(B)** ND, **(C)** FD, **(D)** SD diets infected with indicated passage. ND, normal diet; FD, fat diet; SD, selenium deficient diet.

Number of viral genomes per PCR reaction was estimated using standard curve derived from control SeV plasmid dilutions and values ranged from ~800 genomes (sample SD4.P10, PCR score of 0) per PCR reaction to 25 million genomes (FD2.P10, PCR score of 2). These data indicate that lack of PCR amplification most likely resulted from a low number of total viral genomes, present in the sample (both infectious and non-infectious).

4.4 Illumina sequencing

Illumina deep sequencing of the SeV genome was used to determine whether changes at the consensus or subconsensus level occurred. Sequence data was obtained for up to 99% of the SeV genome (nts 50-15315; numbering according to SeV Z strain (GI: 546225824)). Depth of coverage for most regions was between 10,000× and 100,000×. Prior to analyzing variants in the SeV samples, we verified that the UDS allowed detection of very rare variants. To control for PCR error, three control plasmids containing viral sequence insertions from SeV, RABV or BCoV were mixed at the ratio of 100,000:1,000:10 (SeV:RABV:BCoV), PCR amplified and sequenced.

4.5 Consensus sequence changes

Despite the changes in phenotype from P1 and P10, across the SeV genome only 20 nt mutations occurring at consensus sequence level were detected in any of the samples. Among these, ten mutations led to an amino acid substitution and two occurred in a non-coding region (Table 1, Supplementary Table S2, and Supplementary Fig. S3). A majority of these consensus mutations occurred in the polymerase protein L and the glycoprotein HN.

4.6 Variant detection

The presence of high frequency variants were detected at sites where consensus changes occurred. For example, at nt 2159 of the P gene, a L106V residue change was detected in guinea pig No. 2 (GP2) infected with FD mouse inoculum (FD.GP2) in 95% of the reads. nt variants at nt 2159, (C to G or C to A) were detected in most of the other samples at <10% frequency, except for FD2.P10 with 14% G and SD2.P10 with 11% G detected. Many samples had both G and A present as variants.

High-frequency variants were also detected at nt 8,053, 8,073, and 8,265 of the HN gene (Figs. 5A–C and 6A–C, Supplementary Table S2). An A454V consensus mutation was detected in the BAL sample from SD.GP2. At this same site, nt

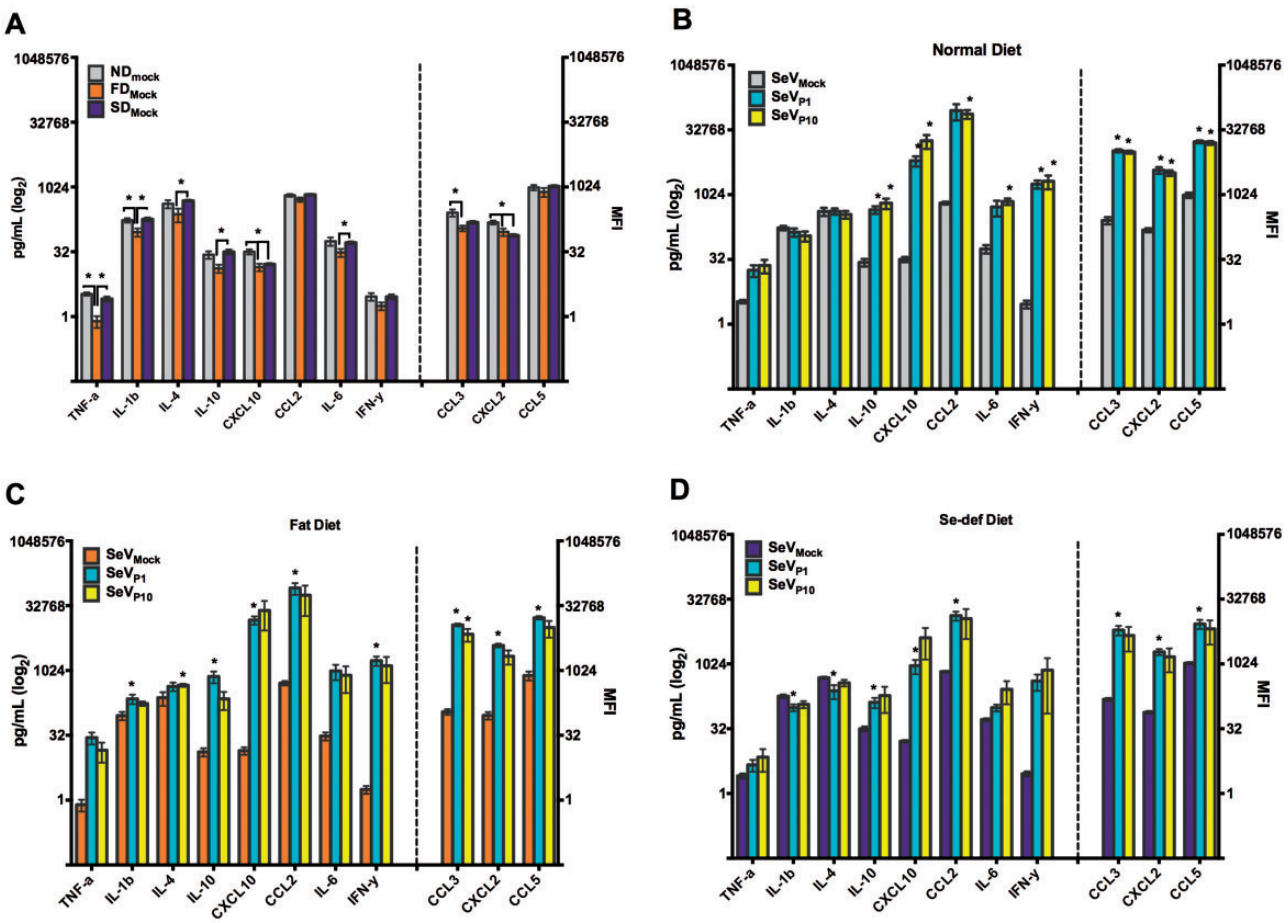

Figure 4. Lung cytokine protein expression profiles according to diet group following serial passage in mice. Cytokine expression profiles were analyzed from lung homogenates using an 11-plex detection assay as indicated in the graph. Expression levels were analyzed under **(A)** Mock infected, **(B)** ND, **(C)** FD, **(D)** SD diets infected with indicated passage. Values are means ± SEM; $n = 4$ for mock infected and $n = 5$ for ND, SD, and FD. Statistical significance was determined by unpaired *t*-test and considered significant if $P < 0.05$. Cytokines to the left of the hash line correspond to pg/ml y-axis and cytokines to the right of the hash line correspond to the MFI y-axis. ND, normal diet; FD, fat diet; SD, selenium deficient; MFI, mean fluorescence intensity.

Table 1. Mutations resulting in an amino acid change or a change in a non-coding region

Gene	Nt	AA	Sample
NC	G1746C	–	SD4.P1, SD5.P1
P	C2159G	L106V	FD.GP2
NC	C3665A	–	SD4.P1[a]
HN	C8053T	A454V	SD.GP3
	G8073A	E461K	FD.GP1, SD.GP3
	A8265C	K525Q	FD.GP1,SD.GP3
L	A9368C	E271D	SD5.P1
	C10811G	I752M	ND3.P1
	A11925G	N1124D[b]	ND3-5.P1, FD5.P1, SD4.P1, ND1-5.P10, FD1,2,3/4[c],5.P10, ND.GP1-3, FD.GP1[a], SD.GP2
	C12001A	T1149N	SD15.P10
	G13366A	R1604K	ND4.P1
	T15136G	M2194R	ND5.P1

[a]Sequence data limited to 1 sample from this treatment group/passage.
[b]Seed stock has both iSNVs present at approximately equal amounts.
[c]Due to limited PCR products for FD3.P10 and FD4.P10, data from the samples were combined.
AA, amino acid; NC, non-coding region; Sample labels are according to diet (ND, FD, SD), mouse or guinea pig No. and passage (P) No. (not applicable for guinea pig samples).

8053, two variants were identified at different frequencies. The more prevalent change was a C to T with a range of 3–31% was detected across all diet groups in the P1, P10, and guinea pig samples. This mutation was detected at a 6.5% in the seed inoculum, thus it appears to have been selected for during passage.

Similarly, at nt 8,073 a G to A substitution resulting in a non-synonymous change, E461K, occurred between 3 and 22% in the P1 population across all ND and FD samples compared with 8.5% in the seed inoculum. By P10 all mouse samples had the change at detected 16–41%, and the mutation was detected at 26–91% in the guinea pig samples (consensus change for ND.GP3 and SD.GP3).

The third high-frequency variant present in the HN gene was detected at nt 8,265 and resulted in a K525Q non-synonymous consensus mutation for the same two guinea pig samples as E461K, samples ND.GP3 and SD.GP3. This A to C nt substitution was detected at 11% in the seed inoculum and ranged between 3 and 21% in the P1 samples, 16–41% in P10 samples, and 27–91% in the guinea pig samples. Thus it appears that all three mutations in the HN gene were enriched during *in vivo* passage.

Multiple consensus changes were detected in the polymerase gene. In sample SD5.P1, nt 9,368 had an A to C change that resulted in a non-synonymous change, E271D, and was detected at 91%. This substitution was present as a rare variant in sample SD4.P1 at 1.4% but was not detected in any of the other samples. Similarly, nt 10,811 had a C to G change that resulted in a non-synonymous change, I752M in a single sample ND3.P1, and was detected at 59%.

A consensus change in the polymerase gene at nt 11,925, an A to G transition, resulted in amino acid residue 1,124 changing from an asparagine to aspartic acid (N1124D) and appeared to be selected for during *in vivo* passage. Interestingly, the egg-passaged seed stock had 50:50 ratio of each amino acid, although the Illumina data analysis called this as an asparagine. There was a trend toward an A to G transition for all diet groups (Figs. 5D and 6D). The frequency of residue 1124D in the P1 sample ranged from 1.2 to 76% while the P10 samples ranged from 13 to 71% with the average frequency of G increasing between P1 and P10 from 40 to 62%. In guinea pig samples this A to G substitution ranged between 49 and 79%.

4.7 HN linkage and entropy analysis

Three amino acid substitutions, A454V, E461K, and K525Q, occurred in the glycoprotein HN and were detected either at the consensus sequence or minority variant level in a majority of the BAL samples. The respective underlying nt changes were C8053T, G8073A, and A8265C. Although all three substitutions already existed at the minority level in the seed stock, their frequencies were far higher in the mice P10 samples (all diet groups) and the guinea pigs. Notably, the frequency of these three amino acid substitutions was highly correlated across the samples (Fig. 5A–C), prompting us to carry out a closer examination of the linkage relationship at the three positions based on the sequencing reads. Indeed, the linkage association between the alleles at these three positions was highly significant. For residues 454 and 461, linkage between the seed stock WT alleles A454 and E461 and the linkage between mutant alleles V454 and K461 were significant at P-value $<10^{-300}$ for all samples sequenced (Fisher's exact test, Supplementary Table S3). At residues 461 and 525, the E461-K525 seed stock WT alleles linkage and the K461-Q525 mutant alleles linkage were significant at P-values $<10^{-28}$ (Supplementary Table S3) for all samples sequenced.

Due to the diversity detected in the HN gene sequence, the average per position entropy was calculated to track changes in intra-host viral diversity. Plots are given for the three diets in Supplementary Figure S4 and showed similar mean entropy for each diet (0.009–0.01), which was a significant increase over the seed stock (0.002), assuming the entropy is normally distributed.

4.8 Structural modeling of SeV HN gene mutations 454, 461, and 525

Structural comparisons of HN proteins of *Paramyxoviridae* (hPIV3, Simian PIV5, NDV, Nipah, Hendra) showed strong conformational similarities (above 60% of structure similarity measured by LGA_S score (Zemla 2003)) even for proteins with sequence identity as low as 21% (Fig. 7)). The strongest structural similarity was seen in conserved regions forming the neuraminidases six-bladed beta-propeller fold, with some deviations observed in more exposed loop regions. The receptor-binding site is located in the center of the propeller.

Constructed structural models allowed us to characterize location of identified mutation points (A454V, E461K, K525Q) within functional complex of HN protein of SeV as well as among HN proteins from other related viruses. Results showed that these positions are located on the protein surface on the top of the receptor-binding domain next to sialic acid-binding regions and don't overlap with interfaces of the tetramer formation (Fig. 8). Additionally, the mutation positions are located in the regions identified as variable in sequence according to calculated multiple sequence alignments between HN proteins of paramyxovirus sequences from UniProt. Results from calculated structural alignments between structural models of HN proteins currently available in PDB suggest that two mutation positions (454, 461) are located in the regions conserved in structure (conserved backbone conformation), while 525 is located in the structurally variable region.

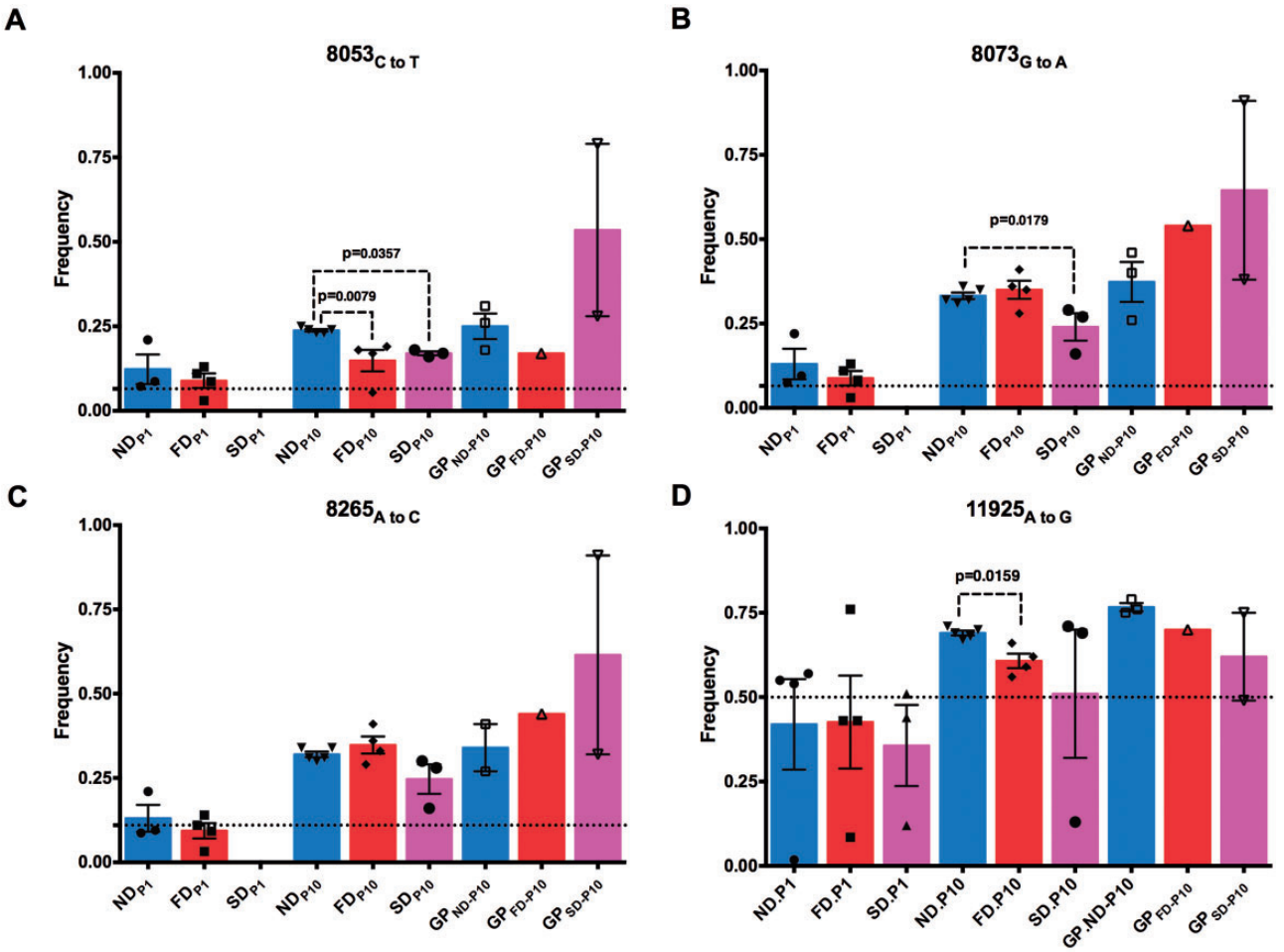

Figure 5. Consensus sequence changes in the coding region of the glycoprotein and polymerase gene following intra- and inter-host passage. Glycoprotein mutations at nt **(A)** nt 8,053, **(B)** nt 8,073, **(C)** nt 8,265, or in the polymerase gene **(D)** nt 11,925 as determine by deep sequencing analysis. Dotted denotes frequency of each variant in the seed inoculum. Values are means ± SEM. Statistical significance was determined by Mann-Whitney test and considered significant if $P < 0.05$. ND, normal diet; FD, fat diet; SD, selenium deficient; GP, Guinea pig.

4.9 The effect of viral diversity on histopathology

Plaque assays, RT-PCR, and Sanger sequencing of a variable regions of the HN and polymerase genes and variable non–coding regions were used to obtain isolates with each of the high-frequency variant genotypes (i.e. nt sites 8,053, 8,073, 8,265, and 11,925). Also included in the analysis was the leader region, including nt sites 20 and 24 which have been associated with virulence in mice (Fujii et al. 2002) but were not included in the Illumina sequence analysis. Analysis of the plaque genotypes showed that nt 20 of the leader region was mutated from G to A in 27% of the clones, with most of the mutations occurring later in passage. Seventeen percent of clones from P1 had an A present (10/58), whereas by P10 an A was detected in 48% of clones (13/27).

After plaque selection and genotyping, twelve clones with sequence that represented high-frequency variants were used to construct populations of single and mixed genotypes for testing in mice (Table 2, Supplementary Fig. S5). Eight groups of mice were infected with homogeneous or heterogeneous mixtures of clones (six mixtures were assayed). The egg-passaged seed stock used to infect mice for P1 served as a 'high diversity' control because it contains a full mutant spectrum versus the other inocula that contained one to six clones. Virulence of each inoculum was gauged using weight loss and lung histopathology scores. Infection with the high-diversity seed stock displayed the greatest virulence as compared with the clone mixtures (Table 2, Supplementary Fig. S5). An inoculum representing a homogenous preparation of one of the two major genotypes present in the seed stock also had higher virulence as compared with the other clone preparations, although the difference statistically significant for only two of the other inoculums. This clone contained the wild type genotype with an Asp residue in the polymerase gene that was selected for during serial passage (N1124D).

5. Discussion

The diversity of viral genomes present in an intra-host population provides an enormous pool of genetic and phenotypic variants and this increases the likelihood of emergence of a variant with a high virulence phenotype. Host nutritional status may influence the generation of variant viruses by altering the selective pressure induced by the host immune response (Domingo and Holland 1997) or by acting on the fidelity of the viral polymerases (Nelson et al. 2001; Beck, Handy, and Levander 2004; Yoshitake et al. 2004; Zaki, Akuta, and Akaike 2005). In this study, we tested the hypothesis that undernutrition (selenium deficiency) and/or overnutrition (high FD) leads

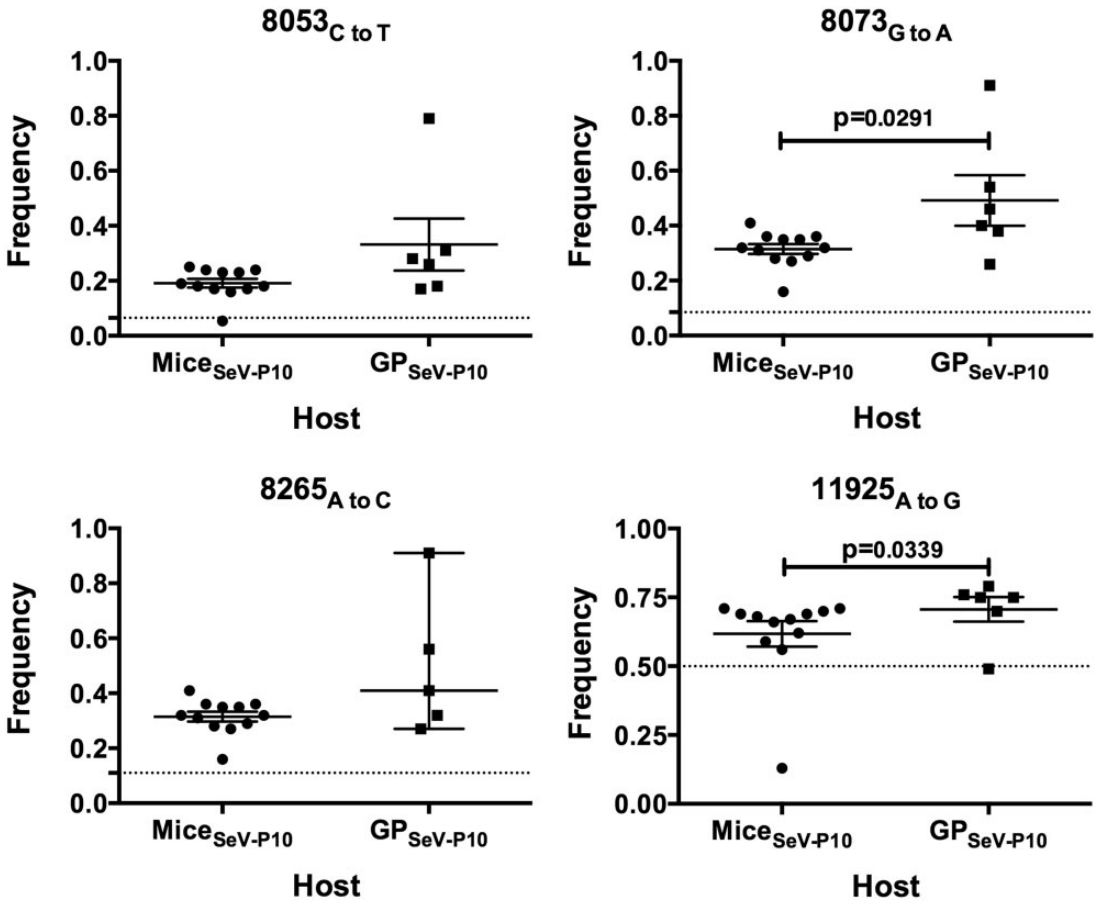

Figure 6. Inter-host selection of a new consensus sequence from low-frequency variants present in initial inoculum. Glycoprotein mutations at nt positions **(A)** 8,053, **(B)** 8,073, **(C)** 8,265, and **(D)** 11,925 of the polymerase gene following passage between different host. Frequency of seed inoculum amplified in eggs is denoted by dotted line. Values are means ± SEM. Statistical significance was determined by Mann-Whitney test and considered significant if $P < 0.05$.

to decreased immune response, increased mutant spectrum variability, and evolution of increased virulence during *in vivo* passage. Deep Illumina sequencing was used to define the viral mutant spectra during serial passage and a host-jump from mice to guinea pigs. Virus phenotype (lung pathology, viral titer) and host response (weight loss, cytokine expression) were characterized.

Analysis of these data shows that viral virulence increased with passage in mice and that the virulent phenotype may also persist upon passage to a second host species, although nutritional status of mice during passage influenced the phenotype. Although deep sequencing revealed the presence of persistent high-frequency variant genotypes in the viral population, only one non-synonymous change in the consensus sequences was detected by passage 10 in a majority of the animals regardless of diet group, polymerase residue N1124D. Thus, it seems likely that shifts in variant populations at the subconsensus level had an effect on virulence. This phenomenon has been observed in other RNA viruses (Domingo 1997; Akaike et al. 2000) indicating the importance of viral population dynamics in understanding mechanisms of viral emergence.

When changes in consensus occur they often involve discrete regions of the genome such as the HN protein and were usually preceded by increased frequency of the variant across diet groups (Supplementary Table S2). Our data show enrichment for three amino acid changes in the HN protein during passage, including residue 461, which is known to affect protein function (Takahashi, Ryan, and Portner 1992; Baumann and Neubert 2010) and virulence (Kiyotani et al. 2001; Fujii et al. 2002). The presence of Lys at 461 rather than Glu is known to decrease HN neuraminidase activity without affecting binding of the HN protein to the cell receptor. For example, experiments described by Takahashi, Ryan, and Portner (1992) showed that while HN protein of SeV is heat-stable with Lys at 461, it becomes heat-labile when substitutions (Gly or Gln) occur at this position. Position 461 is also associated with virulence in mice, and mutates from a negatively charged residue, Glu, to a nonpolar residue, Gly, in response to passage of an egg-propagated isolate in mouse cells (Kiyotani et al. 2001; Itoh et al. 1992). Our data show a mutation of Glu, to a positively charged residue, Lys, in response to passage of an egg-propagated isolate in mice. These data indicate that HN 461 is not only a variable residue but also important for host tropism.

Similar to our data, a mutation at HN residue 525 (Q525R) was reported by Itoh et al. (1992). The R525 mutation appeared after passage in eggs and was associated with conformational changes in the receptor binding site and increased neuraminidase activity, but was not associated with mouse pathogenicity (Itoh et al. 1992). In this study, the egg-passaged seed stock had a positively charged residue, Lys, as consensus with a Gln present as a variant. This mutation, K525Q, was detected with increasing frequency in P10 and guinea pig passages as compared with the seed and P1 samples. Genotype Q525 became the

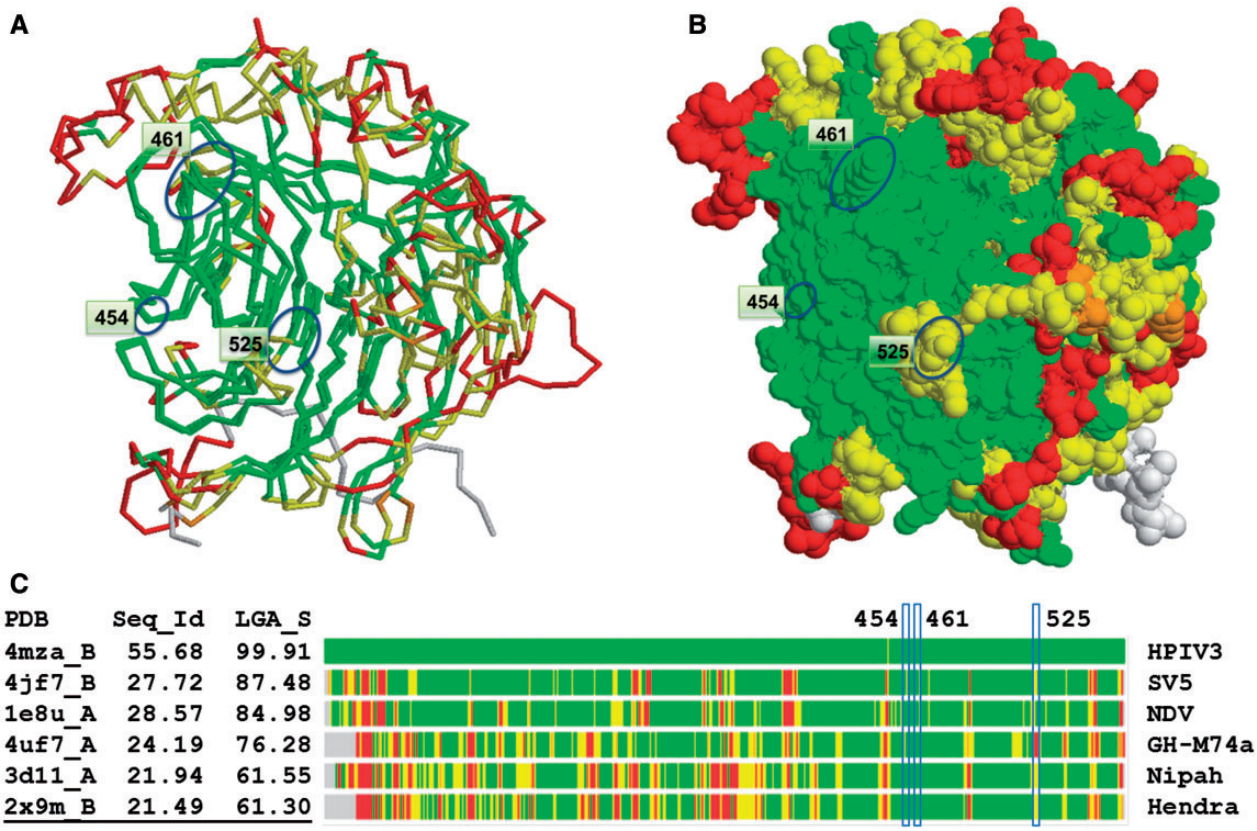

Figure 7. Results from the structure conservation analysis. Homology model of HN protein from SeV is structurally compared with HN proteins from HPIV3, Sv5, NDV, Bat Paramyxovirus (GH-M74a), Nipah and Hendra viruses. Regions identified as structurally conserved are colored from green when Calpha–Calpha deviations in corresponding residues are <2 Å to red (>6 Å), with gray indicating no match. Pairwise structural superposition between homology model of HN protein from SeV and PDB structure (2 × 9 m_B) of HN protein from Hendra virus is shown in Figure 7A (which corresponds to the colored bar representation of calculated superpositions shown in the bottom of Fig. 7C) In Figure 7B mutation positions 454, 461, and 525 are marked by blue ovals. Two of them (454 and 461) are located in the regions observed as structurally conserved between paramyxovirus species (colored in green), and the position 525 is located in the structurally variable region (colored in yellow). Calculated structural alignments show strong conformational similarities (structure similarity measured by LGA_S is above 61%) even for proteins with sequence identity (Seq_Id) as low as 21%.

consensus sequence by guinea pig passage in two of the three guinea pigs from the FD and SD groups.

Data from Fujii et al. (2002) using a pathogenic field isolate of SeV, the Hamamatsu strain, showed that two regions of the Sendai leader sequence, nts 20 and 24, mutated during serial passage in embryonated eggs (U20A and U24A), with an A at both sites resulting in an attenuated phenotype in mice (Fujii et al. 2002). Our data obtained from Sanger sequencing of clone isolates of the lab-adapted strain 52 showed a trend of G to A mutation for nt site 20 as egg-passaged seed virus was passaged serially in mice. In fact, there was an increase in the percentage of clones from P10 having an A as compared with P1 (17% of the clones from P1 had an A, whereas by P10 an A was detected in 48% of clones). Analysis of about eight clones per sample revealed that most samples (6/11) had a mixture of nts at nt 20. However, all of the 85 clones that were sequenced in this study had an A at nt 24. Thus, our data confirm the variability of nt site 20 as described by Fujii et al. and illustrate how consensus sequence data alone may hide the complexities of the viral population diversity.

HN plays a critical role in mediating the virus entry into the host cell thus structural modeling was used to assess the potential impact of the mutations on the viral phenotype. HN glycoproteins initiate virus-cell fusion by binding to the sialic acids found on the surface of the host cell (hemagglutinin activity), trigger a structural refolding of a trimeric fusion protein F needed for merging of viral and cell membranes, and facilitate viral assembly and budding by cleaving the glycosidic linkages with sialic acids via neuraminidase activity. The ectodomain part of the HN proteins is composed of a helical stalk with a large globular head, and the functional form of the molecule (biological assembly) is a tetramer.

Our analysis of the potential impact to HN protein structure indicates that none of the mutations observed at residues 454, 461, or 525 should interfere with receptor binding. Based on structural modeling we conclude the following characteristics of identified mutation positions: (1) they are located on the surface, but should not interfere with critical functional regions such as the sialic acid-binding site, interfaces of HN tetramer formations, and interfaces within F-HN complex (Fig. 8). (2) These positions are in regions of high sequence variability among strains, serotypes or other related viruses. (3) Position 525 is located in the structurally variable region, while the two other positions, 454 and 461, can be considered as structurally conserved on the backbone level among number of closely related viruses. (4) Observations suggest that these positions are not involved in some critical function of the protein where conservation in both structure and sequence would be expected but they may play an important role to help virus adjustment to particular conditions or escape immune response. For example,

results from the antigenic sites predictions on HN proteins characterize the region 451–464 (overlapping with two of our mutation positions) as an epitope with good scores 0.6 and 0.4 by BepiPred and BEOracle prediction algorithms, respectively (Larsen, Lund, and Nielsen 2006; Wang et al. 2011).

Perhaps the most surprising result from this study is the lack of histopathologic lesions in any of the three guinea pigs that were infected with FD.P10 inoculum. Although limited sequencing data were available from these animals, no changes were observed that differentiated FD.GP samples or the FD.P10 from the corresponding samples from the ND and SD groups. However, the combination of amino acids (residues 454, 461, 525) at these three sites is unique for FD.GP1 (the only FD.GP with sequence data for the HN gene). FD.GP1 had HN genotype AKQ (residues 454, 461, 525, respectively), as compared with genotypes AEK and VKQ for the ND.GP and SD.GP samples. Thus, it is possible that it is the unique combination of AKQ that attenuates the virus in guinea pigs. Alternatively other parameters such as temporal aspects of host response may have affected histopathology scores. For example, obesity may result in a delayed pro-inflammatory response to influenza virus infection (Smith et al. 2007), if this response is also delayed in Sendai infected animals, this may have resulted in decreased histopathology scores in the FD group (Fig. 2C).

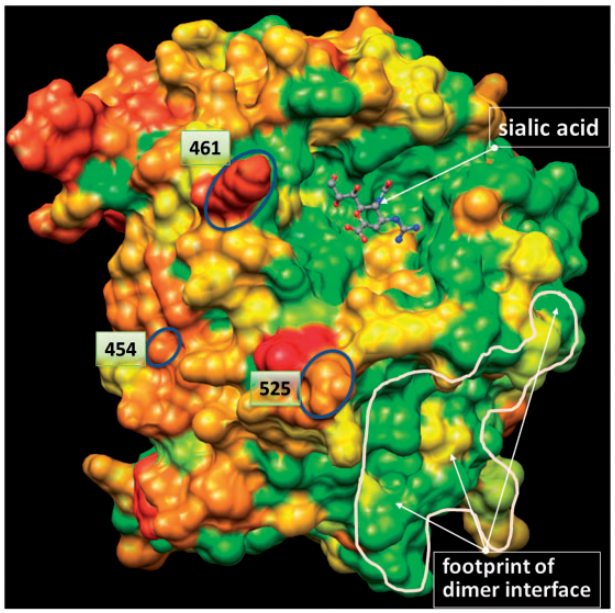

Figure 8. Results from the sequence conservation analysis. Structural model of HN protein from SeV strain Z (GI:546225829) is shown as a surface plot. Coloring scheme reflects sequence conservation calculated using Sum-of-Pairs algorithm. Regions identified as sequence conserved are colored in green and variable in red. Mutation positions 454, 461, and 525 are marked by blue ovals and are located on the protein surface in the regions observed as sequence variable between paramyxovirus species. Positions within the sialic acid-binding pocket and the region of dimer interface are mostly conserved in sequence.

To ascertain the effect of discrete mutations on the viral phenotype, clones were selected which represented the predominant genotypes present in the mutant spectra and were tested for virulence via mouse infections. Despite the fact that clones represented genotypes that increased in frequency during mouse and guinea pig passage, only the genetically diverse seed inoculum significantly surpassed all clones in causing mouse weight loss and lung histopathology. Interestingly, the clone with polymerase mutation D1124 (Genotype 2) was the only clone that resulted in significant weight loss for some of the mice and increased histopathology as compared with other clones (Supplementary Fig. S5, Table 2).

The nutritional status of the host and its relation to infection and immunity has been well documented (Beck 1996; Beisel 1996; Chandra 1997; Keusch 2003; Beck, Handy, and Levander 2004; Schaible and Kaufmann 2007; Alice, Tang, and Semba 2013). These studies have primarily focused on malnutrition of the host; however, more recent studies have begun to look at the role obesity plays in infection and immunity (Nieman et al. 1999; Ritz and Gardner 2006; Kanneganti and Dixit 2012). In this study, we focused on the role malnutrition (in the form of selenium deficiency) and diet induced obesity had in mice infected with SeV to understand the effects these two dietary conditions have on the viral mutant spectra and infection dynamics. Diet-induced obesity has been demonstrated to alter cytokine profiles in mice (Mito et al. 2000). Consistent with those studies, our studies found altered cytokine profiles in uninfected FD mice compared with SD and ND mice (Supplementary Fig. S2). In SeV-infected mice there were no differences found in cytokine expression profiles associated with *in vivo* passage within

Table 2. Clone genotypes used for mouse infections and infection outcome

Genotype No.	Plaque No.	Leader nt 20[a]	NC nt 3665	HN nt 8053	HN nt 8073	HN nt 8265	Poly 11925	Poly.HN AA seq	Average Histo. Score	Average % Weight Loss[b]
1	SD1.P1.4[c]	G	C	C	G	A	A	N.AEK	1.8	0.4
2	SD1.P1.6[c]	G	C	C	G	A	G	D.AEK	2.7	8.0
3	SD2.P1.3[c]	A	C	C	G	A	A	N.AEK	ND	ND
4	SD1.P10.10	A	C	C	G	A	G	D.AEK	ND	ND
5	SD4.P1.9[c]	G	A	C	G	A	A	N.AEK	2.0	2.0
6	FD5.P10.5[c]	A	C	C	A	C	G	D.AKQ	ND	ND
7	SD1.P1.3[c]	G	C	C	A	C	G	D.AKQ	1.2	2.1
8	*SD1.P10.6[c]	A	A	C	G	A	A	N.AEK	ND	ND
9	SD1.P10.8	A	C	T	A	C	G	D.VKQ	ND	ND
10	FD2.P10.5[c]	G	C	T	A	C	G	D.VKQ	0.7	−0.1
11	Mixture ([c])	G/A	C/A	C/T	G/A	A/C	A/G	NA	0.3	2.0
12	Seed stock	G/A	C/A	C/T	G/A	A/C	A/G	NA	5.3	23.3
13	Mock (PBS)	NA	NA	NA	NA	NA	NA	NA	0.3	0

[a]Leader sequence was derived from clone and Sanger sequence data

[b]The average was calculated using the maximum weight loss over the course of the infection for each mouse infected

[c]Mixture of clones consisted of clones representing genotypes 1, 2, 3, 5, 6, 7, 8, and 10. The remaining genotypes did not replicate well enough to allow adequate titer.

dietary groups; however, further analysis of cytokine profiles between the dietary groups found SD mice infected with SeV had a suppressed immune response compared with ND and FD mice infected with SeV (Supplementary Fig. S2). Similarly, mice infected with influenza virus receiving a selenium deficient diet have been shown to have suppressed cytokine profiles compared with mice receiving a ND (Beck et al. 2001).

6. Conclusion

The high mutation rate of RNA viruses increases their ability to adapt to diverse hosts and cause novel human and veterinary diseases. It is likely that humans and animals swap viruses on a regular basis, with the vast majority of these transmission events going undetected. When these host jumps are detected it is hard to predict the likelihood that this new virus is capable of causing significant disease in the human population. Deep sequencing of viral populations provides important genetic information at a high level of resolution (Wanget al. 2007; Fishman and Branch 2009; Borucki et al. 2013; Morelli, et al. 2013; Andersen, et al. 2015) thus allowing detection of subconsensus mutations that may influence the infection outcome. Additionally, data from this study provide further evidence for the role of host nutritional status in the evolution of viral genotype and phenotype (Beck et al. 2001; Beck, Williams-Toone, and Levander 2003; Beck 2007; Smith et al. 2007; Karlsson, Sheridan, and Beck 2010), and indicate that these changes may impact cross-species transmission events. In-depth data analysis, for both the viral population genotypes and the host response, is likely to be critical for understanding the mechanisms of viral host jumps and the role of the viral mutant spectra in vaccine efficacy and reversion.

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgments

We thank Clinton Torres for SeV PCR primer design. This work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Funding

This work was funded by a grant from the Defense Threat Reduction Agency, HDTRA1-23-4-9701.

Conflict of interest: None declared.

References

Akaike, T., Fujii, S., Kato, A., et al. (2000) 'Viral Mutation Accelerated by Nitric Oxide Production During Infection In Vivo', *The FASEB Journal*, 14: 1447–54.

Alice, M., Tang, E. S., and Semba, R. D. (2013) 'Nutrition and Infection', in: E., Kenrad C.M.W., Nelson (eds.) *Infectious Disease Epidemiology: Theory and Practice*, 3rd edn. Burlington, MA: Jones & Bartlett Learning.

Andersen, K. G., Shapiro, B. J., Matranga Christian, B., et al. (2015) 'Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus', *Cell*, 162: 738–50.

Anthony, S. J., Epstein, J. H., Murray, K.A., et al. (2013) 'A Strategy To Estimate Unknown Viral Diversity in Mammals', *MBiology*, 4: e00598–13.

Aronson, J. F., Herzog, N. K., and Jerrells, T. R. (1995) 'Tumor Necrosis Factor and the Pathogenesis of Pichinde Virus Infection in Guinea Pigs', The *American Journal of Tropical Medicine and Hygiene*, 52: 262–9.

Aroor, A. R., and DeMarco, V. G. (2014) 'Oxidative Stress and Obesity: The Chicken or the Egg?' *Diabetes*, 63: 2216–8.

Baumann, C., and Neubert, W. (2010) 'Neuraminidase-Deficient Sendai Virus HN Mutants Provide Protection from Homologous Superinfection', *Archives of Virology*, 155: 217–27.

Beck, M. A. (1996) 'The Role of Nutrition in Viral Disease', *The Journal of Nutritional Biochemistry*, 7: 683–90.

—— (2007) 'Selenium and Vitamin E Status: Impact on Viral Pathogenicity', *Journal of Nutrition*, 137: 1338–40.

——, Handy, J., and Levander, O. A. (2004) 'Host Nutritional Status: The Neglected Virulence Factor', *Trends in Microbiology*, 12: 417–23.

——, Levander, O. A., and Handy, J. (2003) 'Selenium Deficiency and Viral Infection', *Jourrnal of Nutrition*, 133: 1463S–7S.

——, Nelson, H. K., Shi, Q., et al. (2001) 'Selenium Deficiency Increases the Pathology of an Influenza Virus Infection', *The FASEB Journal*, 15: 1481–3.

——, Shi, Q., Morris, V. C., et al. (1995) 'Rapid Genomic Evolution of a Non-Virulent Coxsackievirus B3 in Selenium-Deficient Mice Results in Selection of Identical Virulent Isolates', *Nature Medicine*, 1: 433–6.

——, Williams-Toone, D., and Levander, O. A. (2003) 'Coxsackievirus B3-Resistant Mice Become Susceptible in Se/Vitamin E Deficiency', *Free Radical Biology and Medicine*, 34: 1263–70.

Beisel, W. R. (1996) 'Nutrition in Pediatric HIV Infection: Setting the Research Agenda. Nutrition and Immune Function: Overview', *Journal of Nutrition*, 126: 2611S–5S.

Borucki, M. K., Allen, J. E., Chen-Harris, H., et al. (2013a) 'The Role of Viral Population Diversity in Adaptation of Bovine Coronavirus to New Host Environments', *PLoS One*, 8: e52752.

——, Chen-Harris, H., Lao, V., et al. (2013b) 'Ultra-Deep Sequencing of Intra-Host Rabies Virus Populations During Cross-Species Transmission', *PLoS Neglected Tropical Diseases*, 7: e2555.

Chandra, R. K. (1997) 'Nutrition and The Immune System: An Introduction', *American Journal of Clinical Nutrition*, 66: 460S–3S.

Chen, V. B., Arendall, W. B., III, Headd, J. J., et al. (2010) 'MolProbity: All-Atom Structure Validation for Macromolecular Crystallography', *Acta Crystallographica Section D*, 66: 12–21.

Chen-Harris, H., Borucki, M., Torres, C., et al. (2013) 'Ultra-Deep Mutant Spectrum Profiling: Improving Sequencing Accuracy Using Overlapping Read Pairs', *BMC Genomics*, 14: 96.

Connaris, H., Takimoto, T., Russell, R., et al. (2002) 'Probing the Sialic Acid Binding Site of the Hemagglutinin-Neuraminidase of Newcastle Disease Virus: Identification of Key Amino Acids Involved in Cell Binding, Catalysis, and Fusion', *Journal of Virology*, 76: 1816–24.

David, M., Dzamba, M., Lister, D., et al. (2011) 'SHRiMP2: Sensitive Yet Practical Short Read Mapping', *Bioinformatics*, 27: 1011–2

Domingo, E. (1997) 'RNA Virus Evolution, Population Dynamics, and Nutritional Status', *Biological Trace Elements Research*, 56: 23–30.

—— and Holland, J. J. (1997) 'RNA Virus Mutations and Fitness for Survival', *Annual Review of Microbiology*, 51: 151–78.

Fishman, S. L., and Branch, A. D. (2009) 'The Quasispecies Nature and Biological Implications of the Hepatitis C Virus', *Infection, Genetics and Evolution*, 9: 1158–67.

Fujii, Y., Sakaguchi, T., Kiyotani, K., et al. (2002) 'Identification of Mutations Associated with Attenuation of Virulence of a Field Sendai Virus Isolate by Egg Passage', *Virus Genes*, 25: 189–93.

——, ——, ——, et al. (2002) 'Involvement of the Leader Sequence in Sendai Virus Pathogenesis Revealed by Recovery of a Pathogenic Field Isolate from cDNA', *Journal of Virology*, 76: 8540–7.

Gutiérrez, S., Michalakis, Y., and Blanc, S. (2012) 'Virus Population Bottlenecks During Within-Host Progression and Host-To-Host Transmission', *Current Opinion in Virology*, 2: 546–55.

Holtgrewe, M., Emde, A.-K., Weese, D., et al. (2011) 'A Novel and Well-Defined Benchmarking Method for Second Generation Read Mapping', *BMC Bioinformatics* 12: 210.

Itoh, M., Wang, X.-L., Suzuki, Y., et al. (1992) 'Mutation of the HANA Protein of Sendai Virus by Passage in Eggs', *Virology*, 190: 356–64.

Jeevan, A., Yoshimura, T., Ly, L. H., et al. (2011) 'Cloning of Guinea Pig IL-4: Reduced IL-4 mRNA After Vaccination or Mycobacterium Tuberculosis Infection', *Tuberculosis* 91: 47–56.

Jerzak, G. V., Bernard, K., Kramer, L. D., et al. (2007) 'The West Nile Virus Mutant Spectrum is Host-Dependant and a Determinant of Mortality in Mice', *Virology*, 360: 469–76.

Kanneganti, T. D., and Dixit, V. D. (2012) 'Immunological Complications of Obesity', *Nat Immunol*, 13: 707–12.

—— and —— (2012) 'Immunological Complications of Obesity', *Nature Immunology*, 13: 707–12.

Karlsson, E. A., and Beck, M. A. (2009) 'Diet-Induced Obesity Impairs The T Cell Memory Response to Influenza Virus Infection', *The FASEB Journal*, 23: 110–3.

——, Sheridan, P. A., and Beck, M. A. (2010) 'Diet-Induced Obesity Impairs the T Cell Memory Response to Influenza Virus Infection', *The Journal of Immunology*, 184: 3127–33.

Keusch, G. T. (2003) 'The History of Nutrition: Malnutrition, Infection and Immunity', *Journal of Nutrition*, 133: 336S–40S.

Kiyotani, K., Sakaguchi, T., Fujii, Y., et al. (2001) 'Attenuation of a Field Sendai Virus Isolate Through Egg-Passages is Associated with an Impediment of Viral Genome Replication in Mouse Respiratory Cells', *Archives of Virology*, 146: 893–908.

Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. (2009) 'Improved Prediction of Protein Side-Chain Conformations with SCWRL4', *Proteins: Structure, Function, and Bioinformatics*, 77: 778–95.

Kuss, S. K., Etheredge, C. A., and Pfeiffer, J. K. (2008) 'Multiple Host Barriers Restrict Poliovirus Trafficking in Mice', *PLoS Pathogens*, 4: e1000082.

Lamb, R., and Parks, G. (2007) 'Paramyxoviridae: The Viruses and Their Replication', in: D., Knipe, P., Howley (eds.). *Fields Virology*. Philadelphia, PA: Lippincott, Williams & Wilkins.

Larsen, J.E.P., Lund, O., and Nielsen, M. (2006) 'Improved Method for Predicting Linear B-Cell Epitopes', *Immunome Research*, 2: 2.

Levander, O. A. (1997) 'Nutrition and Newly Emerging Viral Diseases: An Overview', *Journal of Nutrition*, 127: 948S–50S.

Mito, N., Hosoda, T., Kato, C., et al. (2000) 'Change of Cytokine Balance in Diet-Induced Obese Mice', *Metabolism*, 49: 1295–300.

Monteiro, R., and Azevedo, I. (2010) 'Chronic Inflammation in Obesity and the Metabolic Syndrome', *Mediators of Inflammation*, 2010, pii:289645.

Morelli, M., Wright, C., Knowles, N., et al. (2013) 'Evolution of Foot-and-Mouth Disease Virus Intra-Sample Sequence Diversity During Serial Transmission in Bovine Hosts', *Veterinary Research*, 44: 12.

Morse, S. S., Mazet, J.A.K., Woolhouse, M., et al. (2012) 'Prediction and Prevention of the Next Pandemic Zoonosis', *The Lancet*, 380: 1956–65.

Nagata, N., Iwata, N., Hasegawa, H., et al. (2007) Participation of both host and virus factors in induction of severe acute respiratory syndrome (SARS) in F344 rats infected with SARS coronavirus. *Journal of Virology*, 81: 1848–57.

Nelson, H. K., Shi, Q., Van Dael, P., et al. (2001) 'Host Nutritional Selenium Status as a Driving Force for Influenza Virus Mutations', *The FASEB Journal*, 15: 1846–8. 10.1096/fj. 01-0115fje.

Nieman, D. C., Henson, D. A., Nehlsen-Cannarella, S. L., et al. (1999) 'Influence of Obesity on Immune Function', *Journal of the American Dietetic Association*, 99: 294–9.

Pettersen, E. F., Goddard, T. D., Huang, C. C., et al. (2004) 'UCSF Chimera—A Visualization System for Exploratory Research and Analysis', *Journal of Computational Chemistry*, 25: 1605–12.

Ritz, B. W., and Gardner, E. M. (2006) 'Malnutrition and Energy Restriction Differentially Affect Viral Immunity', *Journal of Nutrition*, 136: 1141–4.

Schaible, U. E., and Kaufmann, S. H. (2007) 'Malnutrition and Infection: Complex Mechanisms and Global Impacts', *PLoS Medicine*, 4: e115.

Scott, E. P., and Aronson, J. F. (2008) 'Cytokine Patterns in a Comparative Model of Arenavirus Haemorrhagic Fever in Guinea Pigs', *Journal of General Virology* 89: 2569–79.

Slezak, T., Kuczmarski, T., Ott, L., et al. (2003) 'Comparative Genomics Tools Applied to Bioterrorism Defence', *Briefings in Bioinformatics*, 4: 133–49.

Smith, A. G., Sheridan, P. A., Harp, J. B., et al. (2007) 'Diet-Induced Obese Mice have Increased Mortality and Altered Immune Responses when Infected with Influenza Virus', *Journal of Nutrition*, 37: 1236–43.

Takahashi, T., Ryan, K. W., and Portner, A. (1992) 'Expression of cDNA Encoding the Sendai Virus Hemagglutinin-Neuraminidase Gene: Characterization of Wild-Type and Mutant Gene Products', *Virology*, 187: 837–40.

Tiwari, A., Patnayak, D. P., Chander, Y., et al. (2006) 'Permissibility of Different Cell Types for the Growth of Avian Metapneumovirus', *Journal of Virological Methods*, 138: 80–4.

Vignuzzi, M., Stone, J. K., Arnold, J. J., et al. (2006) 'Quasispecies Diversity Determines Pathogenesis Through Cooperative Interactions in a Viral Population', *Nature*, 439: 344–8.

Wang, C., Mitsuya, Y., Gharizadeh, B., et al. (2007) 'Characterization of Mutation Spectra with Ultra-Deep Pyrosequencing: Application to HIV-1 Drug Resistance', *Genome Research*, 17: 1195–201.

Wang, Y., Wu, W., Negre, N. N., et al. (2011) 'Determinants of Antigenicity and Specificity in Immune Response for Protein Sequences', *BMC Bioinformatics*, 12: 1–13.

Weaver, S. C., Rico-Hesse, R., and Scott, T. W. (1992) 'Genetic Diversity and Slow Rates of Evolution in New World Alphaviruses', *Current Topics in Microbiology and Immunology*, 176: 99–117.

Welch, B. D., Yuan, P., Bose, S., et al. (2013) 'Structure of the Parainfluenza Virus 5 (PIV5) Hemagglutinin-Neuraminidase (HN) Ectodomain', *PLoS Pathogens*, 9: e1003534.

Xu, R., Palmer, S. G., Porotto, M., et al. (2013) 'Interaction Between the Hemagglutinin-Neuraminidase and Fusion Glycoproteins of Human Parainfluenza Virus Type III Regulates Viral Growth In Vivo', *MBiology*, 4: e00803–13.

Yoshitake, J., Akaike, T., Akuta, T., et al. (2004) 'Nitric Oxide as an Endogenous Mutagen for Sendai Virus Without Antiviral Activity', *Journal of Virology*, 78: 8709–19.

Zaki, M. H., Akuta, T., and Akaike, T. (2005) 'Nitric Oxide-Induced Nitrative Stress Involved in Microbial Pathogenesis', *Journal of Pharmacological Science*, 98: 117–29.

Zemla, A. (2003) 'LGA: A Method for Finding 3D Similarities in Protein Structures', *Nucleic Acids Research*, 31: 3370–4.

——, Lang, D., Kostova, T., et al. (2011) 'StralSV: Assessment of Sequence Variability Within Similar 3D Structures and Application to Polio RNA-Dependent RNA Polymerase', *BMC Bioinformatics*, 12: 226.

——, Zhou, C. E., Slezak, T., et al. (2005) 'AS2TS System for Protein Structure Modeling and Analysis', *Nucleic Acids Research*, 33: W111–5.