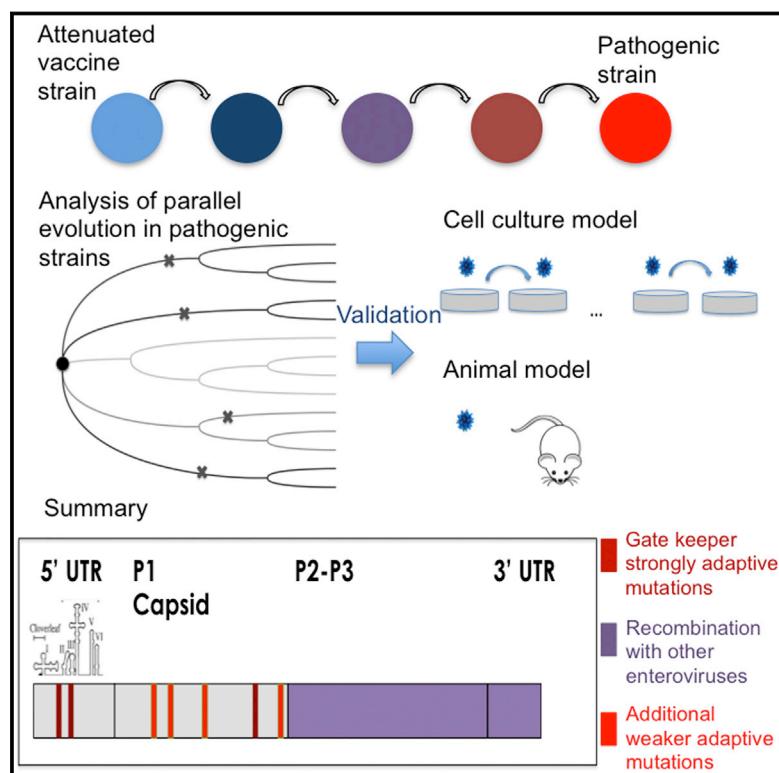


The Evolutionary Pathway to Virulence of an RNA Virus

Graphical Abstract



Authors

Adi Stern, Ming Te Yeh, Tal Zinger, ..., Rasmus Nielsen, Andrew Macadam, Raul Andino

Correspondence

sternadi@post.tau.ac.il (A.S.), raul.andino@ucsf.edu (R.A.)

In Brief

Understanding how an attenuated strain of polio evolved to become fully virulent provides a new framework for rational design of safer vaccines.

Highlights

- Mutations contributing to evolution of virulence in an attenuated RNA virus are inferred
- An experimental evolution approach in cell culture recapitulates the key mutations
- Contribution to virulence is validated in vitro and in an animal model of infection

The Evolutionary Pathway to Virulence of an RNA Virus

Adi Stern,^{1,6,*} Ming Te Yeh,^{2,6} Tal Zinger,¹ Matt Smith,³ Caroline Wright,² Guy Ling,¹ Rasmus Nielsen,^{4,5} Andrew Macadam,³ and Raul Andino^{2,7,*}

¹Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel-Aviv 6997801, Israel

²Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94158, USA

³Division of Virology, National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire, EN6 3QG, UK

⁴Department of Integrative Biology

⁵Department of Statistics

University of California, Berkeley, Berkeley, CA 94720, USA

⁶Co-first author

⁷Lead contact

*Correspondence: sternadi@post.tau.ac.il (A.S.), raul.andino@ucsf.edu (R.A.)

<http://dx.doi.org/10.1016/j.cell.2017.03.013>

SUMMARY

Paralytic polio once afflicted almost half a million children each year. The attenuated oral polio vaccine (OPV) has enabled world-wide vaccination efforts, which resulted in nearly complete control of the disease. However, poliovirus eradication is hampered globally by epidemics of vaccine-derived polio. Here, we describe a combined theoretical and experimental strategy that describes the molecular events leading from OPV to virulent strains. We discover that similar evolutionary events occur in most epidemics. The mutations and the evolutionary trajectories driving these epidemics are replicated using a simple cell-based experimental setup where the rate of evolution is intentionally accelerated. Furthermore, mutations accumulating during epidemics increase the replication fitness of the virus in cell culture and increase virulence in an animal model. Our study uncovers the evolutionary strategies by which vaccine strains become pathogenic and provides a powerful framework for rational design of safer vaccine strains and for forecasting virulence of viruses.

INTRODUCTION

Predicting the course of evolution is one of the most challenging and important areas in biology. One key strategy of this endeavor involves assessing how often evolutionary trajectories occur repeatedly, and if they do, using this to predict future evolutionary trajectories. Many laboratory experiments with various microbial systems have shown a remarkable degree of parallelism (e.g., Cuevas et al., 2002; Meyer et al., 2012). Similarly, the parallel (or convergent) emergence of traits such as echolocation, toxin resistance, and pigmentation is being increasingly inferred in natural life settings (Arendt and Reznick, 2008; Feldman et al., 2012;

Teeling et al., 2002). However, it is often difficult to identify parallel gain-of-function from a set of ancestors that are shrouded in a distant past. At the molecular level, detecting convergence is even more challenging in the absence of the ancestral sequence. Here, we present evidence of extensive parallel substitution and recombination events occurring in repeated epidemics of vaccine-derived polioviruses, where the ancestral sequence is a known entity with the defined sequence of the vaccine strain.

Global eradication of poliovirus is based on immunization with the oral poliovirus vaccine (OPV), composed of three strains (1, 2, and 3), each of which provides protection from the corresponding wild-type (WT) viral serotypes. These highly attenuated strains replicate in the recipient's gut, mimicking natural infection. Vaccination with OPV elicits both humoral and mucosal immunity, thus providing robust and long-lasting protection from infection with WT polioviruses. Using this effective vaccine, WT poliovirus has been brought to the brink of global eradication. However, despite the remarkable success of the polio eradication campaign, in regions with low vaccine coverage, dozens of poliomyelitis outbreaks associated with circulating vaccine-derived polioviruses (cVPVs) have been observed (Kew et al., 2005). The largest number of these incidents is due to OPV type 2 (OPV2). Following an extensive period of circulation, an evolved form of OPV2 gains both elevated virulence and the ability to effectively circulate in human populations.

To date, limited analyses of localized cVPVs (Burns et al., 2013; Endegue-Zanga et al., 2015; Famulare et al., 2015; Hovi et al., 2013; Tao et al., 2013) have been undertaken, and many aspects of the evolutionary and epidemiological dynamics of cVPVs remain unclear (Duintjer Tebbens et al., 2013). Understanding the evolutionary path(s) that leads from the attenuated and non-hazardous form of the virus into a circulating and pathogenic virus is central to the global poliovirus eradication efforts, for prediction of future epidemics, to understand the risks of current strategies, and for designing new and safer vaccines.

Notably, the study of cVPVs allows a unique scrutiny of evolution as it occurs in real-life settings. From an evolutionary perspective, several key features of cVPV epidemics are exceptional: first, as opposed to natural life setting, the precise

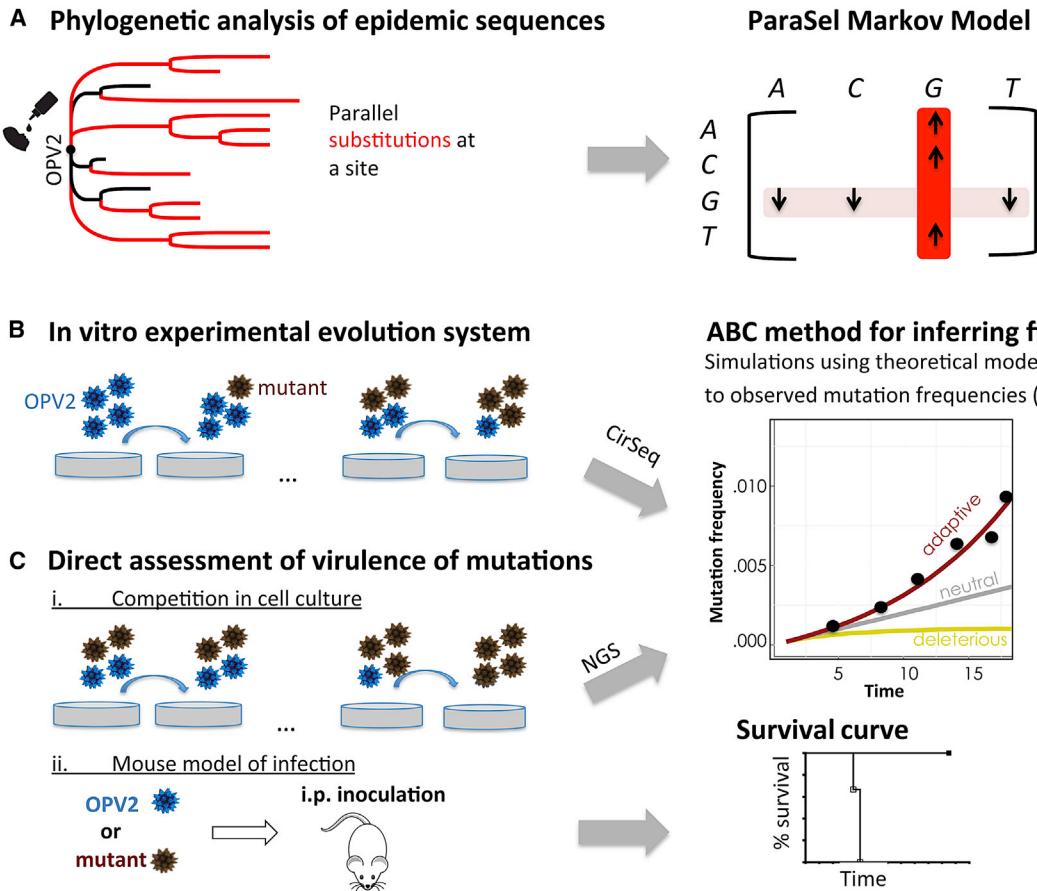


Figure 1. Schematic of Analyses and Experiments Performed in the Manuscript

Experiments and data are illustrated on the left and computational approaches for data analysis on the right.

(A) A phylogenetic analysis of cVDPV consensus sequences from epidemics across the globe was undertaken with the aim of detecting parallel substitutions under selection (in red) as compared to the ancestral OPV2 state (in black). A novel Markov model denoted ParaSel was developed, which assumes that selection will lead to an increase in the rate of substitutions into a certain nucleotide, and decrease in the rate of loss of this nucleotide (exemplified as the “G”). The model then allows calculating the likelihood of the data and performing model choice, followed by inference of specific sites where selection led to parallel substitutions.

(B) An experimental evolution approach was used to monitor the emergence of mutations in OPV2 conferring an evolutionary fitness advantage during growth in cell culture at elevated body temperature.

(C) Direct assessment of the effect of mutations inferred in (A) and (B) on viral virulence was performed using (i) competition assays between selected mutants and OPV2 at high temperature, and (ii) a mouse model of infection that allowed comparing the virulence of selected mutants versus OPV2. An ABC approach ([STAR Methods](#)) was used to infer whether an increase in the frequency of a mutation over time represents significant adaptive evolution. The method compares theoretical simulations of allele frequencies to the empirical data. Accordingly, deleterious alleles are expected to be present at low frequencies throughout the experiment due to purifying selection, whereas adaptive alleles will increase rapidly in frequency. Neutral alleles will accumulate at a rate equal to the mutation rate.

ancestor of all type 2 cVDPVs is known—the original OPV2 strain. Second, global nationwide surveillance has led to hundreds of available cVDPV sequences, allowing the reconstruction of a fine-resolution phylogeny of cVDPVs ([Figure S1A](#)). Finally, the extremely high rate of evolution of poliovirus ([Jorba et al., 2008](#)) drives rapid divergence of cVDPVs from their ancestral founder. Thus, analysis of cVDPV sequences allows a unique study of repeated independent evolutionary processes that often begin and end similarly at the phenotypic level: vaccination with a highly attenuated viral strain ends tragically in cases of paralytic poliomyelitis, indistinguishable from those associated with wild-type poliovirus.

To examine cVDPV evolution in detail, we developed a novel evolutionary framework that allowed us to determine the key substitutions that drive loss of attenuation and regain of virulence in OPV2 and offer a coherent model of evolution for cVDPV. We then used a simple in vitro experimental evolution paradigm to identify OPV2 adaptive mutations during controlled infection of OPV2 in cell culture ([Figure 1](#)). Notably, the key substitutions inferred to be adaptive in the epidemics were also observed in the in vitro system. We then validated the contribution to viral virulence of these key mutations in competition assays and in an animal model. Thus, we propose that our experimental evolution paradigm, which is based on highly accurate

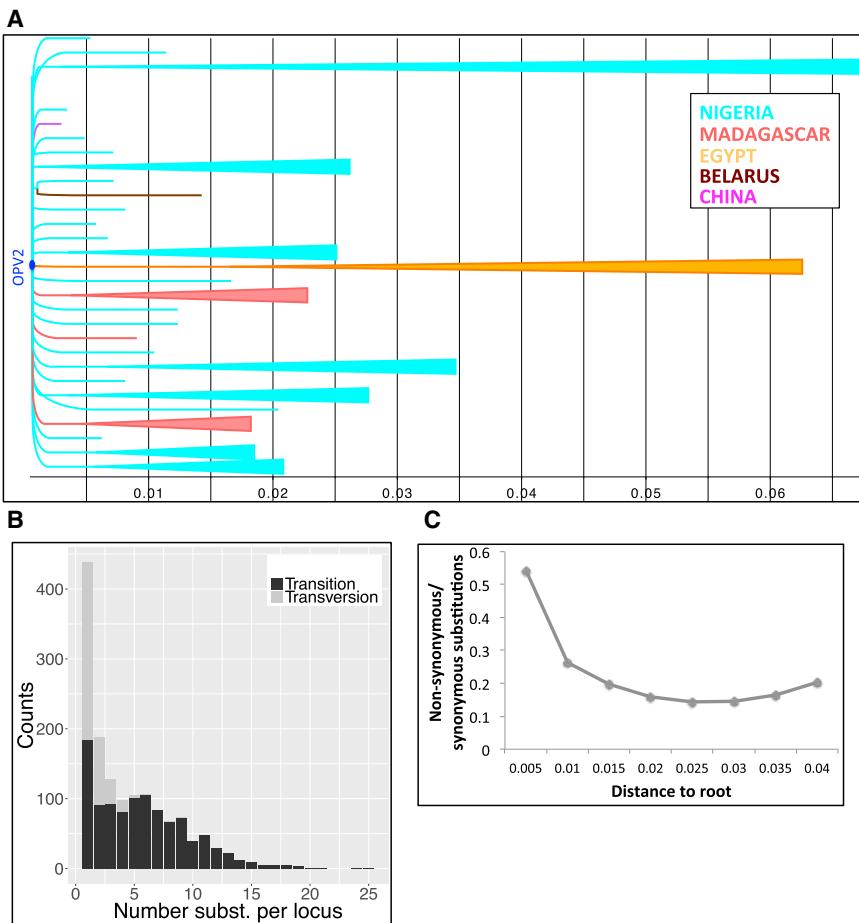


Figure 2. Phylogenetic Analysis of cVDPV2 Sequences

(A) Maximum-likelihood phylogenetic tree for 424 cVDPV sequences, based on synonymous polymorphisms in the non-recombinant capsid region. The x axis represents substitutions per site, where 0.011 substitutions are equivalent to one year of viral circulation based on the PV molecular clock (Jorba et al., 2008). Clades were collapsed for visualization.

(B) The distribution of the number of independently occurring substitutions inferred across the phylogeny. Transitions are colored in black and transversions are in gray. Notably the transition/transversion ratio is much higher in events mapped two or more times.

(C) The ratio of the number of non-synonymous to synonymous substitutions across time shows that strong adaptation occurs early on in the epidemics, which is relaxed later on, and that incomplete purifying selection likely occurs during recent evolution.

(A) through (C) all refer to the non-recombinant capsid region. See also Figure S1 and Table S1.

sequencing and computational inference, can be a powerful tool to understand and predict adaptation of viruses to their host in real-life setting.

RESULTS

Abundant Parallel Substitutions Are Detected throughout cVDPV2 Evolution

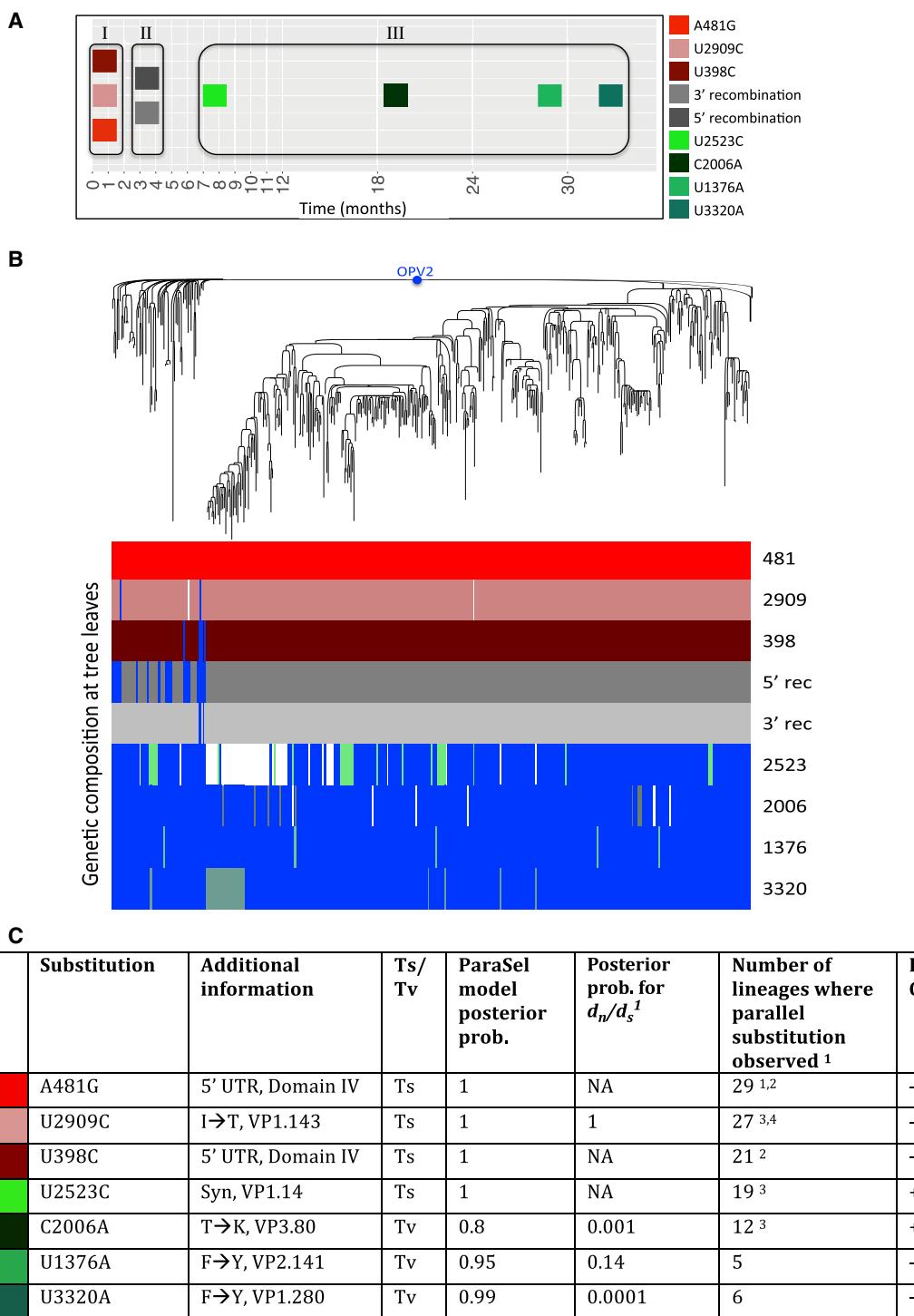
An extensive dataset of 424 full-genome consensus sequences of verified type 2 cVDPVs (defined as more than 0.5% divergence in the VP1 gene that forms part of the capsid) was compiled, with sequences from Belarus, China, Egypt, Madagascar, and Nigeria (Figure 2A; Figure S1A). Our aim was to use these genomic sequences to uncover the evolutionary pathways that lead from the non-virulent OPV2 strain into a highly virulent cVDPV2 strain, capable of causing acute flaccid paralysis. We used the unique phylogenetic nature of these cVDPV2 sequences, whereby the tree topology is described by one known ancestral sequence that subsequently branches out into several independent epidemics (illustrated in Figure 1; Figure 2A). Thus, the tree topology has a known ancestral sequence at the root, and is characterized by multifurcation at the root (i.e., multiple branches diverge from the root), both of which do not typically occur in natural evolutionary settings.

Sequence comparison demonstrated a high level of identity (typically 95%–99% identity; Figure S1B) between all cVDPV2 sequences and the OPV2 sequence along the region encoding the capsid (P1) portion of the genome. In all but three of the 424 genomic sequences, recombination was identified, either in the P2 or P3 region of the genome.

Furthermore in 397 of 424 sequences an additional recombination event was observed leading to replacement of the OPV2 5' untranslated region (UTR). The recombinant sequences were derived from co-circulating HEV-C strains, most often a coxsackievirus strain, but in some cases a circulating PV strain (Table S1) (Burns et al., 2013).

We hypothesized that evolution from OPV2 to virulent poliovirus strains involves a limited number of genetic events that are similar in the different parallel epidemics. OPV2 is an avirulent strain, with reduced replicative fitness at body temperature in human cells. This potentially leads to smaller viral population sizes, which in turn reduces neurovirulence and probability of transmission. Accordingly positive selection in human hosts could lead to fitter viruses over time, i.e., viruses that replicate better in human tissues. These viruses will eventually lead to neurovirulence (as exhibited by all cVDPV sequences analyzed here) and potentially increased probability of virus transmission. An analysis of the phylogeny of the largest epidemic occurring in Nigeria supported the fact that peaks in transmission correspond to gain of function events described below and to gaps in vaccination (Figure S1C).

To identify the evolutionary changes contributing to cVDPV regain of virulence, we searched for substitutions that occurred repeatedly and independently along the phylogeny of the cVDPV sequences (i.e., with no shared ancestry). To this end



¹ A total of 29 epidemics (lineages coming out of the OPV2 sequence) were analyzed here.

² In several lineages, substitution is obtained via recombination rather than via mutation.

³ Alternative substitutions found at this locus are counted as well (e.g., U2523A).

⁴ In two additional lineages an A2908G substitution led to a replacement of I→V at VP1.143.

(legend on next page)

we focused on non-recombinant regions of the genomes, and used a phylogenetic mapping approach to reconstruct the history of substitutions of each of the cVDPV sequences (**STAR Methods**). Two key substitutions have been previously implicated in the loss of attenuation of OPV2: A481G within the 5' UTR of the genome, and a Thr to Ile (or Asn) substitution at amino acid 143 of the capsid protein VP1 corresponding to genomic loci 2908-2909 (Macadam et al., 1991; Ren et al., 1991). Indeed, in each of the independent outbreak lineages analyzed, these substitutions occurred early on in the outbreaks, i.e., are observed in the nearest branch from the OPV2 root. Surprisingly, we also deduced 841 loci in the capsid where additional parallel substitutions occurred. These events occurred repeatedly in two or more different lineages of the phylogeny, with greater than six parallel substitutions on average per locus.

While past studies have assumed that parallel substitutions typically represent the fixation of positively selected mutations (but see Zou and Zhang, 2015), the huge number of substitutions observed in parallel lineages seems improbable and challenges this assumption. Instead we propose that several factors characteristic of RNA virus evolution lead to an unusual large number of parallel substitutions, which are not necessarily under positive selection. First, polioviruses possess an exceptionally high substitution rate of ~ 0.01 substitutions/site/year (Jorba et al., 2008). Given that several of the epidemics examined here spanned 5 or more years (Burns et al., 2013; Yang et al., 2003), it is quite possible that two independent substitutions at the same locus occurred by chance (a simple calculation yields a probability of ~ 0.85 for two independent substitutions; **STAR Methods**). Second, the transition/transversion ratio in polioviruses exceeds ten (Acevedo et al., 2014; Crotty et al., 2001; Freistadt et al., 2007), implying that parallel substitutions involving transitions have an even higher probability. Indeed, most of the parallel substitutions inferred from our analysis correspond to transition events (Figure 2B). Third, certain mutations are expected to be under relaxed selection, such as synonymous mutations. Thus for example, codons with 2-fold degeneracy may lead to multiple parallel substitution events. Finally, it has been demonstrated that viruses undergo relaxed purifying selection near the tips of the phylogeny (also termed incomplete purifying selection), leading to accumulation of slightly deleterious substitutions (Park et al., 2015; Pybus et al., 2007). A hallmark of this effect is an increase in the ratio of non-synonymous (d_n) to synonymous (d_s) substitutions toward the tips of the phylogeny, as indeed we observe in this study as well (Figure 2C). All these effects could

thus lead to parallel substitutions not driven by selection, but rather occurring by chance.

A Probabilistic Evolutionary Model Detects a Small Number of Parallel Substitutions under Selection

To identify among the parallel substitutions those that are under positive selection, we designed a phylogenetic probabilistic Markov model named ParaSel, which infers sites where parallel substitutions are driven by positive selection rather than by random genetic drift (**STAR Methods**; Figure 1). Essentially, this model is similar to phylogenetic codon models that search for an increase in the d_n/d_s ratio (Muse and Gaut, 1994; Nielsen and Yang, 1998). Here, we search for an increase in the substitution rate as compared to a baseline rate that can be loosely defined as the average rate of substitution observed across all sequences. This strategy allows the method to detect selection at non-coding regions and at synonymous sites, which d_n/d_s signals are inherently incapable of examining. In order to model parallel selection, the ParaSel model assumes that a specific allele will experience positive directional selection across all branches of the phylogeny. This is modeled as an increase in the background substitution rate by a factor of $S/(1 - e^{-S})$, the rate of fixation of a non-neutral mutation (Kimura, 1962; Nielsen and Yang, 2003), where $S = Ns$ is the population-size (N)-scaled selection coefficient. Similar to most phylogenetic Markov models, ParaSel takes into account ancestry via the topology of the phylogeny, a high rate of evolution is accounted by the branch lengths of the phylogeny, the transition/transversion rate is a parameter of the model, and among-site rate variation (ASRV) allows a higher rate of evolution for some sites (such as synonymous sites). We also uniquely model incomplete purifying selection by changing the ASRV distribution toward the tips of the tree (**STAR Methods**). If the ParaSel model provides significantly better fit to the data than a null model that assumes no selection, the posterior probability for parallel selection is calculated at each site (**STAR Methods**). Testing ParaSel on simulated data demonstrated high levels of sensitivity and specificity of the method (Table S2).

Applying ParaSel to the sequences of cVDPVs, we obtained strong statistical support for parallel positive selection acting on a small group of substitutions (Table S3). We inferred seven sites with a high posterior probability of selection for parallel substitutions (Figure 3). Using a similar rationale to the ParaSel model we further extrapolated that the recombination events replacing the 5' UTR and 3' regions of OPV2 with a HEV-C

Figure 3. Results of ParaSel Phylogenetic Model Allow Reconstructing the Regain of Virulence of cVDPVs from Independent Epidemics across the Globe

Seven substitutions and two recombination events are inferred to be under parallel positive selection.

(A) The timeline shows the approximate inferred timing of each event, illustrating a sequential process of increase in fitness. Three “waves” of events are observed, inferred based on the estimated fixation time of each event. The relative timing of each event within each wave could not be inferred consistently and is shown based on one of the epidemics.

(B) A projection of each of the nine events (bottom) onto the phylogenetic tree of the cVDPVs (top). A thin line colored in red/gray/green, corresponding to the colors in (A), represents a substitution or recombination event under parallel selection present in the sequence corresponding to the leaf of the tree on the top. Blue represents the ancestral OPV2 state, whereas white represents a third alternative.

(C) A table summarizing the properties of the substitutions detected by ParaSel, with colors as in (A). Ts refers to transition, whereas Tv refers to transversion. Additional information specifies the location of the substitution on the accepted RNA structure of the 5' UTR (Andino et al., 1990; Pilipenko et al., 1989), or the amino-acid replacement and its location in the capsid forming genes VP1 through VP4.

See also Figure S2 and Table S3.

Table 1. Frequencies of Gate-Keeper Mutants in Day 14 Stool Samples of 11 Individuals Vaccinated with Trivalent OPV

Nucleotide	Change	Codon	Amino Acid Change	Vaccinee										
				S3	S5	S6	S8	S10	S14	S17	S18	S20	S29	S30
398	U → C			0.03	0.33	0.06	0.02	0.06	0.04	0.12	0.09	0.22	0.20	0.07
481	A → G			0.99	1.00	0.94	0.99	0.99	0.95	0.94	0.99	0.98	1.00	0.98
2908-10	AUU → ACU, AAU, VP1-143 GUU, AGU		I → T, N, V, S	0.89	0.67	0.54	0.32	0.81	0.31	0.09	0.71	0.44	0.73	0.65
Recombinant ^a (R)/ non-recombinant (NR)				NR	R	R	NR	R	NR	NR	R	NR	R	NR

^aOPV2/OPV1, in P2 or P3.

sequence were themselves under positive selection (STAR Methods, Figure S2), with the latter potentially endowing the viruses with more fit non-structural proteins. Our inference results overlapped the results reported by the site-specific d_n/d_s signal for positive selection only at VP1.143 (Figure 3C). Thus, ParaSel is a method complementary to d_n/d_s based methods, since d_n/d_s only captures diversifying positive selection at amino-acid sites.

By scrutinizing the pattern of substitutions using phylogenetic mapping, we were able to reconstruct the order of emergence of the different parallel selection events during viral adaptation. These substitutions could be assigned into three “waves” of events (Figure 3A). First, three substitution events occurred almost invariantly in all lineages deriving from OPV2. These substitutions include the known A481G and U2909C attenuation reversions, as well as an additional reversion at U398C in the 5' UTR (Famulare et al., 2015; Macadam et al., 1991; Muzychenco et al., 1991). In all lineages with enough resolution (i.e., enough branching), these mutations preceded all other events observed in the second two “waves.” We thus propose these three mutations serve as “gate-keeper” mutations, which are akin to “driver” mutations in cancer evolution (Krogan et al., 2015).

The second wave of events included one to two recombination events with a HEV-C partner. Interestingly, in the 27 sequences where there was no recombination in the 5' UTR, the gate-keeper and additional mutations led to a sequence similar to a HEV-C sequence, presumably endowing the virus with the same beneficial alterations obtained by the recombination. During the third wave of events, the virus continued to slowly revert to sequences that are conserved across WT poliovirus. We estimate that fine-tuning adaptation of the virus continues even 30 months after vaccine administration.

Next Generation Sequencing of Vaccines Confirms Rapid Emergence of Gate-Keeper Mutations

We next sought to confirm how fast mutations accumulate post-vaccination. Next generation sequencing analysis of OPV2 viruses excreted by 11 individuals, 14 days after vaccination with trivalent OPV (Table 1), supports a model in which increase in fitness follows a defined evolutionary pathway as all three “gate-keeper” mutations were observed at high frequencies in viruses from all vaccinees studied. The mutant frequencies at day 14 further suggest that the order of emergence is usually A481G followed by VP1-143X, then U398C. At day 14 five samples contained type 2 viruses that were recombinants with OPV1 in the P2 or P3 regions. While not proving definitively

that subsequent evolutionary events could not occur without the gatekeeper mutations these findings suggest that alternative evolutionary trajectories are less common.

An In Vitro Model of Experimental Evolution in Cell Culture Recapitulates Several of the Phylogenetic Findings

We hypothesized that the increase in virus virulence and circulation in human populations results from a general replicative fitness gain that could be monitored in a simpler system such as cell culture. To examine the possibility of establishing a link between real-life evolution of viral sequences and this simple experimental paradigm, we serially passaged OPV2 strain in HeLa S3 cells for 14 replication cycles, at either 33°C or 39.5°C (Figure 1B). High temperature was used in order to create strong selection pressure that could accelerate gain of fitness (Manor et al., 1999), whereas 33°C, the vaccine manufacturing temperature, was designed as a control. Importantly, we used large viral population sizes for our experiments: during each passage, we ensured an input of 10⁶ particle forming units (PFUs) with a burst size of ~10⁸ particles. Based on population genetics theory, these large population sizes are expected to accelerate adaptation (Rouzine et al., 2001), especially as compared to real-life setting, where transmission bottlenecks severely slow down the rate at which a population can adapt due to the effects of genetic drift.

Our results showed that the OPV2 consensus sequence replicating in cell culture remained essentially identical during serial passage at high temperature (Figure S3). However, using a highly accurate sequencing approach (CirSeq (Acevedo and Andino, 2014; Acevedo et al., 2014)), we uncovered a large number of minor alleles accumulating over the passages (Figure S3B). We detected on average more than 18,400 variant alleles per passage with an average coverage of more than 240,000 reads per position (Table S5). The vast majority of variants were homogeneously distributed at low frequencies between 10⁻⁵ and 10⁻³, with only a few allele frequencies higher than 10⁻³. Using this powerful sequencing strategy, we were able to (1) estimate mutation rates, and (2) track the frequency of each minor allele across time in order to estimate its relative fitness (Figure 1B). Changes in allele frequency along time are driven by the fitness of each mutation and by stochastic effects, especially when allele counts are low. We hence used an approximate Bayesian computation (ABC) method (Foll et al., 2014; Foll et al., 2015) (STAR Methods), which compares simulations based on the Wright-Fisher model to the empirical allele frequency data.

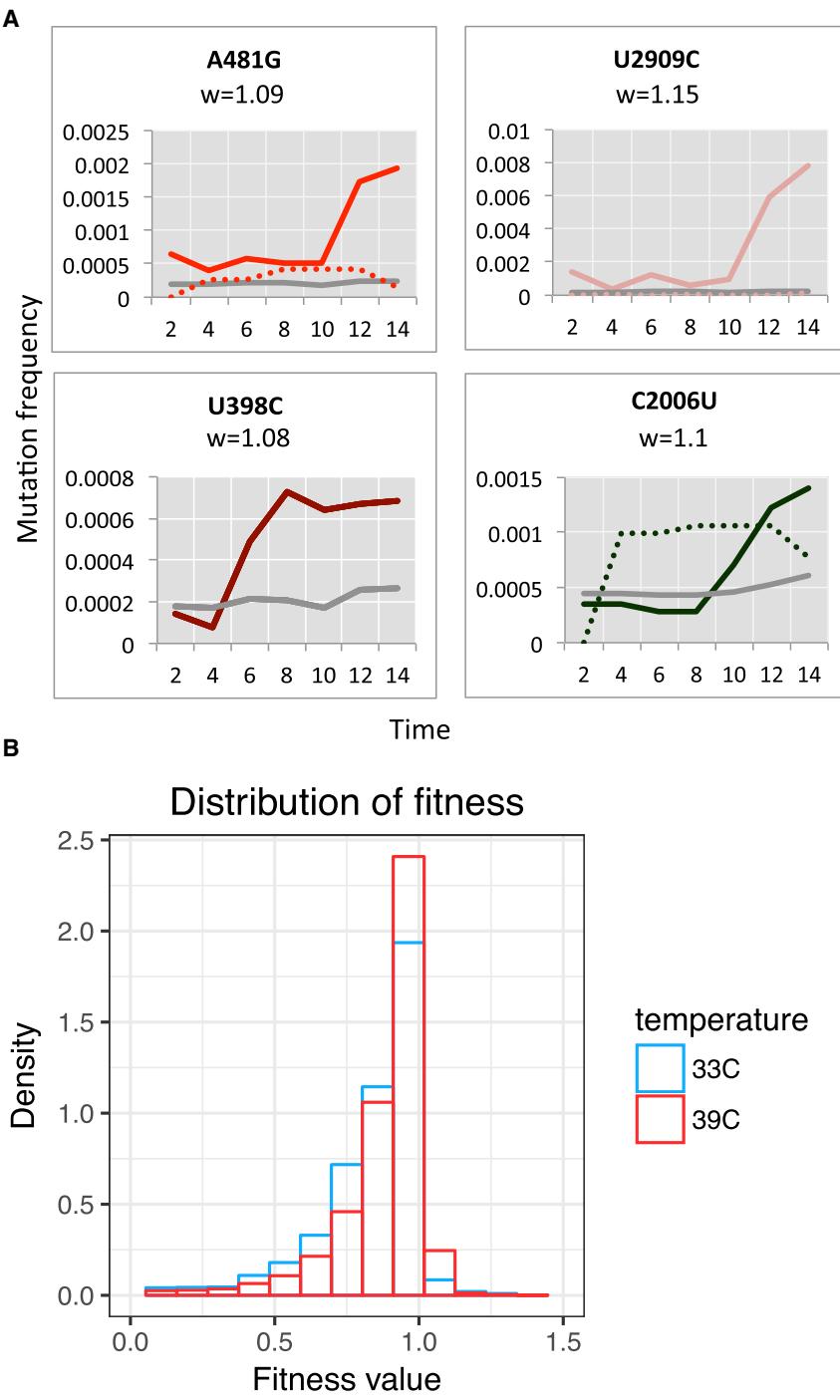


Figure 4. Results of In Vitro Experimental Evolution at 33°C and 39.5°C

Viruses were serially passaged in HeLa cells for seven passages (corresponding to 14 generations) at both temperatures. Passages were sequenced using highly accurate CirSeq sequencing (Acevedo et al., 2014).

(A) Four mutations predicted with ParaSel were validated using the experimental evolution approach. Time-series trajectories are shown in solid lines (39.5°C) and dashed lines (33°C), with colors corresponding to those in Figure 3. For U398C, lack of coverage precluded inferring reliable allele frequencies at 33°C. The gray line in each box represents the neutral allele behavior over time, based on the mean behavior of synonymous mutations in the relevant class of mutation, excluding CpG and UpA sites.

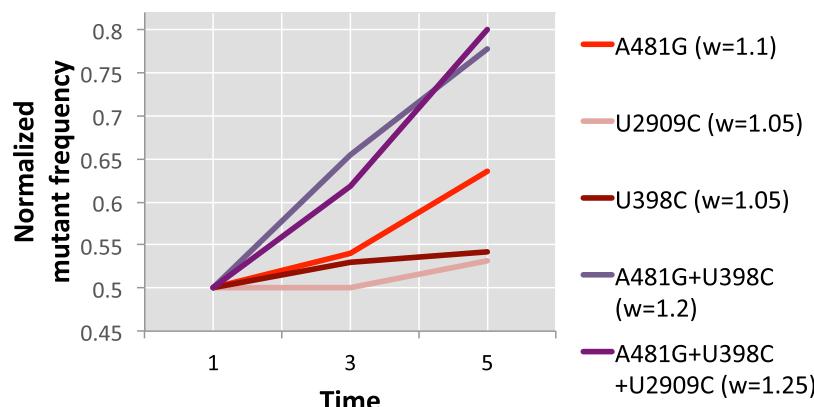
(B) Distributions of genome-wide fitness values obtained at 33°C (blue) and 39.5°C (red) show that the 39.5°C distribution is shifted to the right, indicating more adaptation at elevated temperature. See also Figures S3, S4 and S5 and Tables S5.

Then, the simulations yielding the best fit to the empirical data form the posterior probability distribution for the fitness of each allele. This posterior distribution allows inferring whether an allele is under significant positive or purifying selection, and allows estimating the magnitude of the fitness effect. Validation of our approach is shown in Figure S4, which shows a high accuracy of inferring mutations rate and a very high specificity in inferring positive selection from simulated data.

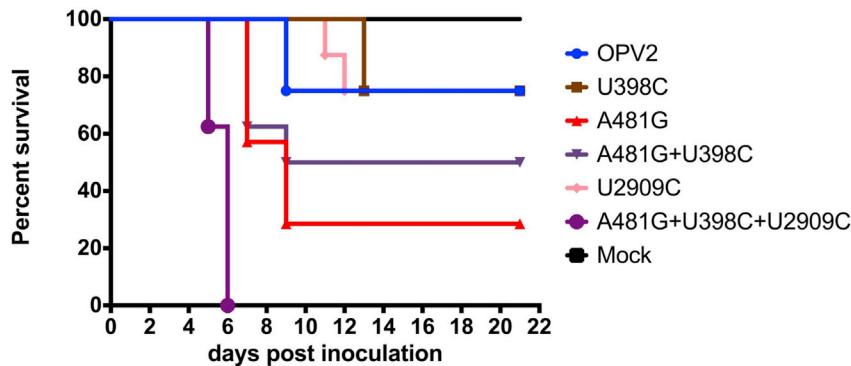
We considered that most of the parallel selection mutations inferred in the epidemic sequences would not arise in the tissue culture experiments, since (1) tissue culture lacks critical features present in a natural infection (acquired immunity, elements of innate immunity, and the complex environment inside the human host), (2) our experimental paradigm lacked the effects of recombination with other viruses, and (3) the length of our experiment was dramatically shorter than the time cVPDVs have been adapting in nature. Nevertheless, four of the seven mutations inferred under parallel selection in the outbreak sequences were identified as positively selected in cell culture at 39.5°C (Figure 4A). Of the remaining mutations, two were inferred to be neutral and one as deleterious at 39.5°C (Figure S5). Conversely, all seven mutations were inferred to be deleterious at 33°C. In general the genome-wide distribution of fitness effects was shifted to the right at 39.5°C as compared to 33°C (Figure 4B) ($p < 2.2e-16$, two-sample K-S test). This reflects the fact that more mutations are neutral

or adaptive at high temperature whereas at the lower temperature more mutations are deleterious, confirming the increase in selection pressure at high temperature. We note that both distributions are relatively devoid of lethal alleles, since low confidence (most often low frequency) mutations were filtered out (STAR Methods). The four verified adaptive mutations at high temperature included all three gate-keeper mutations, and an additional non-synonymous mutation (Figure 4A; Figure S5).

A Competition in cell culture, “gate-keeper” mutants



B Virulence of gate-keeper mutations in susceptible mice



Experimental Measurements of Virus Fitness Confirm the Role of “Gate-Keeper” Mutations in Increasing OPV2 Virulence

We next sought to directly verify the effect of the gate-keeper mutations on viral fitness and virulence. To this end we performed two sets of complementary experiments: (1) competition assay of OPV2 versus mutants in tissue culture and (2) estimation of OPV2 mouse model of infection, where replicative fitness as well as adaptive immunity and intercellular interactions are all critical for OPV2 replication. We began by cloning the gate-keeper mutations onto the backbone of OPV2, either as single mutations or in combinations.

We performed direct competition assays between the mutant clones and OPV2, by initially combining the two at equal proportions, and infecting HeLa cells at high temperature (Figure 1C). At the end of each round of infection virus was collected (representing an unknown mix of OPV2 and the mutant) and used to reinfect cells, and thus serially passaged in a similar manner five times. We used next generation sequencing to measure the frequencies of OPV2 versus the mutant over time, and used our ABC approach to measure the fitness of the cloned mutant/s compared to OPV2. Our results confirmed our first analysis: all gate-keeper mutations were found to be adaptive (Figure 5A). Interestingly, we discovered that the combinations of mutants were fitter than the sum of each mutant's effect on its own, sug-

Figure 5. Direct Assessment of Virulence of Mutations

(A) Results of competition assay of the three gate-keeper mutations reveal that all three mutations outcompete the OPV2 strain, on their own or in combination. Mutant frequencies varied between 0.3 and 0.6 at the first passage and were normalized for purpose of presentation to begin at 0.5.

(B) Survival analysis of susceptible mice infected with OPV2 gate-keeper mutants. The analysis reveals that the A481G mutation leads to increased virulence as compared to OPV2, with an even stronger effect observed for the combination of all three mutants. Potential epistatic interactions are observed in (A) and (B) for combinations of mutations (see text).

See also Figure S6.

gesting a synergistic epistatic interaction between all three mutations.

Next, we examined how gate-keeper mutations affect virulence in susceptible mice (STAR Methods; Figure 1). To this end, we first established a type I interferon receptor knockout (IFNR-ko) mouse model (Ida-Hosonuma et al., 2005). This IFNR-ko mouse model is susceptible to OPV2 infection (as opposed to wild-type mice) and thus supports determination of viral dissemination and virulence (Figure S6). Using this model, we determined virulence of OPV2 carrying gate-keeper mutations by measuring survival rates following intra-

peritoneal infection. Similar to the competition assay results, the strongest effect was obtained when combining the three gate-keepers together, leading to increased and accelerated mortality compared to OPV2 and to A481G on its own (log-rank test; $p < 0.0005$). Only A481G on its own showed a significant increase in virulence as compared to OPV2 (75% survival versus 28.5% 9 days post-infection; log-rank test; $p < 0.05$); both U398C and U2909C showed a non-significant delay in mortality (Figure 5B). Conversely the combination of A481G and U398C led to a slight reduction in virulence as compared to A481G on its own (log-rank test; $p < 0.05$), suggesting an antagonistic epistatic effect. Thus, these results suggest epistatic effects between the three gate-keeper mutations, which may determine the “order-of-addition” in the OPV2 evolutionary trajectory to increased virulence.

OPV2 Fitness Increase Appears to Rely on Synonymous Substitutions that Disrupt CpG or UpA Dinucleotides

Intriguingly, one of the mutations contributing to OPV2 evolution to virulence in the phylogenetic analysis was a synonymous substitution (Figure 3C). In fact, when lowering the posterior probability threshold of the ParaSel analysis, many more synonymous substitutions were identified (Table S3), and these extra substitutions were enriched for substitutions disrupting a CpG or UpA dinucleotide (Fisher exact test, $p = 0.006$). Similarly, the cell culture experimental evolution analysis identified a number of positively

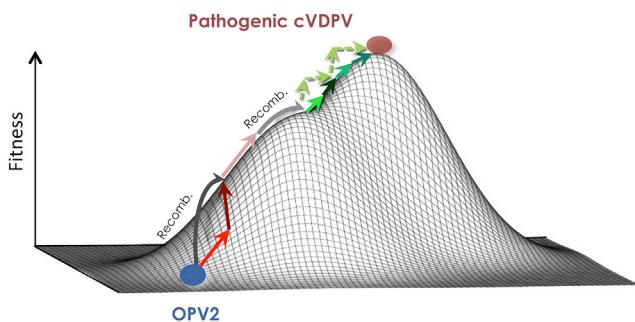


Figure 6. A Model Summarizing the Proposed Path/s to Virulence of OPV2

Illustrative fitness landscape of OPV2, which starts off as poorly adapted to replication in human cells at high temperature and is thus illustrated at the bottom of the landscape. Colored arrows (corresponding to Figure 3) represent substitutions that increase the fitness of the virus, while the gray arrows represent recombination with a HEV-C sequence. Substitutions disrupting CpG/UpA are illustrated by dashed arrows, with potentially several different options leading to the same fitness altitude in the landscape.

See also Table S4.

selected synonymous mutations. While there was no overlap between the ParaSel and cell culture results for these mutations, the cell culture analysis at high temperature revealed a higher fitness for mutations disrupting a CpG or a UpA dinucleotide (t test; $p < 2e-16$, $p < 1e-15$, respectively). Furthermore, we observed a depletion in the relative ratio of CpG or UpA dinucleotides in cVPDV genomes, and an abundance of CpA and UpGs, both of which represent one transition away from CpG or UpA (Figure S7A). We also observed an overall decrease in the expected dinucleotide frequencies over the first few years of cVPDV evolution (Figures S7B–S7E), and a more complex pattern for UpA observed over the entire 6 years. Several recent studies have indeed indicated that an increase in CpG or UpA dinucleotides lead to attenuation of various RNA viruses (Atkinson et al., 2014; Burns et al., 2009; Cheng et al., 2013; Tulloch et al., 2014), and an as-yet unidentified cellular pathway may recognize and target viral genomes containing CpG or UpA sites. Accordingly, our data supports the hypothesis that cVPDVs are under selection to reduce the frequency of CpGs and to some extent also of UpAs in their genomes.

CONCLUSIONS

Our analysis provides a model describing the evolutionary steps sufficient for the OPV2 strain to lose its attenuation and become virulent. These include a series of initial substitutions, most of which are transitions, notably a relatively “easy” event for polioviruses. In particular, the gate-keeper transition mutations appear to endow the highest initial fitness, which drives them to rapid fixation from absent or extremely low levels present in the vaccine stock (Neverov and Chumakov, 2010). This is supported by the fact that all three mutations are already present in viruses excreted by vaccinees 14 days after vaccination, and have been observed previously in viruses from primary vaccinees or in sewage surveillance (Dedepsidis et al., 2006; Macadam et al., 1991; Nakamura et al., 2015). Following this, we propose that repli-

cation becomes more efficient, allowing a larger viral population size, which increases the probability of transmission, co-infection and recombination with prevalent HEV-C strains (Jegouic et al., 2009). The common denominator of all these different recombination events is that they almost always supply the OPV2 strains with the region encoding the viral protease (3C gene) and the RNA-dependent RNA polymerase (3D gene), as well as the 3' UTR. This region is frequently replaced with that of OPV1 within weeks of trivalent OPV vaccination, underlining the selection against the OPV2-derived genes in the human gut. The predominance of HEV-C P3 sequences in cVPDV isolates then suggests OPV1-derived genes are also sub-optimal compared to those obtained from circulating HEV-C viruses. We noted several nucleotide and amino-acid variants at 3C and 3D that almost invariantly “arrived” with the HEV-C recombination (Table S4). However, we cannot tease apart whether these are positions conserved in HEV-C sequences or rather that selection favored these variants in the context of an OPV2 capsid. Finally, the last so-called “wave” likely contributes a smaller adaptive value to the virus. It is likely that only a critical mass of CpG/UpA disruptions has a significant effect on fitness. The additional non-synonymous mutations detected are not detected in antigenic sites (Minor, 1986); these mutations, as well as the CpG/UpA disruptions, merit further investigation.

Previous research has shown that epistatic interactions may dictate the order in which mutations are fixed (Meyer et al., 2012; Weinreich et al., 2006), since the fitness advantage of one mutation is dependent on the presence of previous ones. To test whether such epistatic interactions limited the evolution of Sabin in the capsid region, we calculated the time shared across the phylogeny by all pairs of mutations (Figure S1D). Accordingly, we tested whether pairs of mutations fixed consecutively and were maintained together along the phylogeny. Our phylogenetic analysis found no support for such epistatic mutations in the capsid region on its own. On the other hand, our experimental results supported an epistatic interaction between locus 398 and 481 in the 5' UTR. Interestingly, the epistatic effect was opposed between the human cell culture model and the mouse model, which suggests that these loci interact with other cellular components that differ between mice and human. Indeed, on passage of OPV2 in mouse L cells expressing the human poliovirus receptor the A481G mutation is selected rapidly whereas the U398C mutation has never been observed (A.M., unpublished data). Based on the cell culture results, this epistatic interaction may explain why most sequences are recombinant in the 5' UTR, since the HEV-C recombination supplies these mutations together with others at once. Together, these lines of evidence lead us to propose a potentially rugged fitness landscape leading from OPV2 to cVPDV (Figure 6). Additional support for this rugged model is that (1) multiple RNA structures exist in the 5' UTR region and in the P2/P3/3'UTR regions (Burrill et al., 2013b), where recombination is observed. Such structures may impose epistatic interactions, (2) chimeric sequences reconstructed from OPV2 and a HEV-C sequence displayed elevated fitness at higher temperatures (Jegouic et al., 2009). Thus, we suggest that a relatively small number of events, composed mainly of transition mutations and recombination, can lead OPV2 to a fitness peak similar in altitude to WT.

Our approach combines phylogenetic analysis of sequences from ongoing epidemics based on parallel selection events,

and an experimental evolution approach of virus replication in tissue culture. The differences between the two systems is enormous: the evolutionary dynamics of viruses during an epidemic include establishment of an initial infection, immune evasion, navigation among different tissues, and transmission dynamics, all of which are lacking in a tissue culture setup. Nevertheless, our results indicate that some of the key evolutionary features of viral replication can be inferred in the relatively simple experimental evolution setup in tissue culture, which spanned a very short time-frame. We note that our setup was based on HeLa cells; while these cells do not necessarily best mimic gut epithelial cells or neuronal cells where PV strains replicate, we emphasize their key utility: by supporting large virus population sizes, we were able to capture small differences in mutation frequencies as they arose. Other cell lines that we tested (e.g., Vero cells; data not shown) did not allow the experimental design used here, mainly in terms of population size and sequencing. Importantly, using CirSeq and our ABC approach allow us to infer increase in fitness at a much higher resolution and accuracy than previous approaches used to predict loss of attenuation in tissue culture (Taffs et al., 1995). To summarize, we propose that vaccine design can in the future be guided by a similar approach, which will allow predicting in a short time whether or not live attenuated vaccine candidates can revert to pathogenic forms. The combined approaches like that described here, may also facilitate evolutionary forecasting toward the prediction of future viral epidemics.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Mice
 - Cells and Virus
- METHOD DETAILS
 - Phylogenetic analysis
 - Experimental validation
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Inferring mutational fitness from time-series allele frequencies
 - Inferring mutation rates
 - Inference of beneficial mutations
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2017.03.013>.

A video abstract is available at <http://dx.doi.org/10.1016/j.cell.2017.03.013#mmc3>.

AUTHOR CONTRIBUTION

A.S., M.T.Y., A.M., and R.A. designed the project. A.S., T.Z., and R.N. developed the theoretical and computational methodologies. M.T.Y., M.S., and

C.W. performed all of the experiments. A.S., T.Z., G.L., and R.A. analyzed the results. A.S. and R.A. wrote the manuscript, which all authors reviewed.

ACKNOWLEDGMENTS

We thank Cara Burns and Olen Kew from the CDC for sharing preliminary sequence data from the Nigerian outbreaks. We further thank many colleagues at UCSF and TAU for helpful discussions and comments on the manuscript, in particular Tzachi Hagai, and Mor Geva for initial analyses. A.S. was supported by a WIS post-doctoral award and by a Rothschild post-doctoral award. T.Z. and G.L. are supported by the Edmond J. Safra Center for Bioinformatics at Tel-Aviv University. This work was supported in part by NSF (CMMI-0941355), NIH (R01 GM097115, AI36178, AI40085, P01 AI091575) and the University of California (CCADD), DARPA Prophecy, and Bill and Melinda Gates Foundation.

Received: April 5, 2016

Revised: January 3, 2017

Accepted: March 6, 2017

Published: March 23, 2017

REFERENCES

- Acevedo, A., and Andino, R. (2014). Library preparation for highly accurate population sequencing of RNA viruses. *Nat. Protoc.* 9, 1760–1769.
- Acevedo, A., Brodsky, L., and Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505, 686–690.
- Andino, R., Rieckhof, G.E., and Baltimore, D. (1990). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* 63, 369–380.
- Arendt, J., and Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23, 26–32.
- Atkinson, N.J., Witteveldt, J., Evans, D.J., and Simmonds, P. (2014). The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids Res.* 42, 4527–4545.
- Brent, R.P. (1971). An algorithm with guaranteed convergence for finding a zero of a function. *Comput. J.* 14, 422–425.
- Bull, J.J., Badgett, M.R., Wichman, H.A., Huisken, J.P., Hillis, D.M., Gulati, A., Ho, C., and Molineux, I.J. (1997). Exceptional convergent evolution in a virus. *Genetics* 147, 1497–1507.
- Burge, C., Campbell, A.M., and Karlin, S. (1992). Over- and under-representation of short oligonucleotides in DNA sequences. *Proc. Natl. Acad. Sci. USA* 89, 1358–1362.
- Burns, C.C., Campagnoli, R., Shaw, J., Vincent, A., Jorba, J., and Kew, O. (2009). Genetic inactivation of poliovirus infectivity by increasing the frequencies of CpG and UpA dinucleotides within and across synonymous capsid region codons. *J. Virol.* 83, 9957–9969.
- Burns, C.C., Shaw, J., Jorba, J., Bulkbuk, D., Adu, F., Gumede, N., Pate, M.A., Abanida, E.A., Gasasira, A., Iber, J., et al. (2013). Multiple independent emergences of type 2 vaccine-derived polioviruses during a large outbreak in northern Nigeria. *J. Virol.* 87, 4907–4922.
- Burrill, C.P., Strings, V.R., and Andino, R. (2013a). Poliovirus: generation, quantification, propagation, purification, and storage. Current protocols in microbiology Chapter 15, Unit 15H 11.
- Burrill, C.P., Westesson, O., Schulte, M.B., Strings, V.R., Segal, M., and Andino, R. (2013b). Global RNA structure analysis of poliovirus identifies a conserved RNA structure involved in viral replication and infectivity. *J. Virol.* 87, 11670–11683.
- Byrd, R.H., Lu, P.H., Nocedal, J., and Zhu, C.Y. (1995). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J. Sci. Comput.* 16, 1190–1208.

- Cheng, X., Virk, N., Chen, W., Ji, S., Ji, S., Sun, Y., and Wu, X. (2013). CpG usage in RNA viruses: data and hypotheses. *PLoS ONE* 8, e74109.
- Crotty, S., Cameron, C.E., and Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl. Acad. Sci. USA* 98, 6895–6900.
- Cuevas, J.M., Elena, S.F., and Moya, A. (2002). Molecular basis of adaptive convergence in experimental populations of RNA viruses. *Genetics* 162, 533–542.
- Dedepsidis, E., Karakasiliotis, I., Paximadi, E., Kyriakopoulou, Z., Komiotis, D., and Markoulatos, P. (2006). Detection of unusual mutation within the VP1 region of different re-isolates of poliovirus Sabin vaccine. *Virus Genes* 33, 183–191.
- Duintjer Tebbens, R.J., Pallansch, M.A., Kim, J.H., Burns, C.C., Kew, O.M., Oberste, M.S., Diop, O.M., Wassilak, S.G., Cochi, S.L., and Thompson, K.M. (2013). Oral poliovirus vaccine evolution and insights relevant to modeling the risks of circulating vaccine-derived polioviruses (cVPVs). *Risk Anal.* 33, 680–702.
- Dunn, G., Begg, N.T., Cammack, N., and Minor, P.D. (1990). Virus excretion and mutation by infants following primary vaccination with live oral poliovaccine from two sources. *J. Med. Virol.* 32, 92–95.
- Endegue-Zanga, M.C., Sadeuh-Mba, S.A., Iber, J., Burns, C., Nimpa-Mengouo, M., Demanou, M., Vernet, G., Etoa, F.X., and Njouom, R. (2015). Circulating vaccine-derived polioviruses in the Extreme North region of Cameroon. *J. Clin. Virol.* 62, 80–83.
- Famulare, M., Chang, S., Iber, J., Zhao, K., Adeniji, J.A., Bukbuk, D., Baba, M., Behrend, M., Burns, C.C., and Oberste, M.S. (2015). Sabin Vaccine Reversion in the Field: a Comprehensive Analysis of Sabin-Like Poliovirus Isolates in Nigeria. *J. Virol.* 90, 317–331.
- Feldman, C.R., Brodie, E.D., Jr., Brodie, E.D., 3rd, and Pfrenger, M.E. (2012). Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. USA* 109, 4556–4561.
- Foll, M., Poh, Y.P., Renzette, N., Ferrer-Admetlla, A., Bank, C., Shim, H., Maspas, A.S., Ewing, G., Liu, P., Wegmann, D., et al. (2014). Influenza virus drug resistance: a time-sampled population genetics perspective. *PLoS Genet.* 10, e1004185.
- Foll, M., Shim, H., and Jensen, J.D. (2015). WFABC: a Wright-Fisher ABC-based approach for inferring effective population sizes and selection coefficients from time-sampled data. *Mol. Ecol. Resour.* 15, 87–98.
- Freistadt, M.S., Vaccaro, J.A., and Eberle, K.E. (2007). Biochemical characterization of the fidelity of poliovirus RNA-dependent RNA polymerase. *Virol. J.* 4, 44.
- Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174.
- Hovi, T., Paananen, A., Blomqvist, S., Savolainen-Kopra, C., Al-Hello, H., Smura, T., Shimizu, H., Nadova, K., Sobotova, Z., Gavrilin, E., and Roivainen, M. (2013). Characteristics of an environmentally monitored prolonged type 2 vaccine derived poliovirus shedding episode that stopped without intervention. *PLoS ONE* 8, e66849.
- Huelsnbeck, J.P., Nielsen, R., and Bollback, J.P. (2003). Stochastic mapping of morphological characters. *Syst. Biol.* 52, 131–158.
- Ida-Hosonuma, M., Iwasaki, T., Yoshikawa, T., Nagata, N., Sato, Y., Sata, T., Yoneyama, M., Fujita, T., Taya, C., Yonekawa, H., and Koike, S. (2005). The alpha/beta interferon response controls tissue tropism and pathogenicity of poliovirus. *J. Virol.* 79, 4460–4469.
- Jegouic, S., Joffret, M.L., Blanchard, C., Riquet, F.B., Perret, C., Pelletier, I., Colbere-Garapin, F., Rakoto-Andrianarivo, M., and Delpeyroux, F. (2009). Recombination between polioviruses and co-circulating Coxsackie A viruses: role in the emergence of pathogenic vaccine-derived polioviruses. *PLoS Pathog.* 5, e1000412.
- Jorba, J., Campagnoli, R., De, L., and Kew, O. (2008). Calibration of multiple poliovirus molecular clocks covering an extended evolutionary range. *J. Virol.* 82, 4429–4440.
- Kew, O.M., Sutter, R.W., de Gourville, E.M., Dowdle, W.R., and Pallansch, M.A. (2005). Vaccine-derived polioviruses and the endgame strategy for global polio eradication. *Annu. Rev. Microbiol.* 59, 587–635.
- Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* 47, 713–719.
- Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature* 217, 624–626.
- Kirkegaard, K., and Baltimore, D. (1986). The mechanism of RNA recombination in poliovirus. *Cell* 47, 433–443.
- Kosakovsky Pond, S.L., Poon, A.F., Leigh Brown, A.J., and Frost, S.D. (2008). A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. *Mol. Biol. Evol.* 25, 1809–1824.
- Krogan, N.J., Lippman, S., Agard, D.A., Ashworth, A., and Ideker, T. (2015). The cancer cell map initiative: defining the hallmark networks of cancer. *Mol. Cell* 58, 690–698.
- Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingwersen, R., Sheppard, H.W., and Ray, S.C. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160.
- Lorenzo-Redondo, R., Fryer, H.R., Bedford, T., Kim, E.Y., Archer, J., Kosakovsky Pond, S.L., Chung, Y.S., Penugonda, S., Chipman, J.G., Fletcher, C.V., et al. (2016). Persistent HIV-1 replication maintains the tissue reservoir during therapy. *Nature* 530, 51–56.
- Lou, D.I., Hussmann, J.A., McBee, R.M., Acevedo, A., Andino, R., Press, W.H., and Sawyer, S.L. (2013). High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl. Acad. Sci. USA* 110, 19872–19877.
- Macadam, A.J., Pollard, S.R., Ferguson, G., Dunn, G., Skuce, R., Almond, J.W., and Minor, P.D. (1991). The 5' noncoding region of the type 2 poliovirus vaccine strain contains determinants of attenuation and temperature sensitivity. *Virology* 181, 451–458.
- Manor, Y., Handsher, R., Halmut, T., Neuman, M., Bobrov, A., Rudich, H., Vonsover, A., Shulman, L., Kew, O., and Mendelson, E. (1999). Detection of poliovirus circulation by environmental surveillance in the absence of clinical cases in Israel and the Palestinian authority. *J. Clin. Microbiol.* 37, 1670–1675.
- Mee, E.T., Minor, P.D., and Martin, J. (2015). High resolution identity testing of inactivated poliovirus vaccines. *Vaccine* 33, 3533–3541.
- Meyer, J.R., Dobias, D.T., Weitz, J.S., Barrick, J.E., Quick, R.T., and Lenski, R.E. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* 335, 428–432.
- Minor, P.D. (1986). Antigenic structure of poliovirus. *Microbiol. Sci.* 3, 141–144.
- Muse, S.V., and Gaut, B.S. (1994). A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Mol. Biol. Evol.* 11, 715–724.
- Muzychenko, A.R., Lipskaya GYu, Maslova, S.V., Svitkin, Y.V., Pilipenko, E.V., Nottay, B.K., Kew, O.M., and Agol, V.I. (1991). Coupled mutations in the 5'-untranslated region of the Sabin poliovirus strains during *in vivo* passages: structural and functional implications. *Virus Res.* 21, 111–122.
- Nakamura, T., Hamasaki, M., Yoshitomi, H., Ishibashi, T., Yoshiyama, C., Maeda, E., Sera, N., and Yoshida, H. (2015). Environmental surveillance of poliovirus in sewage water around the introduction period for inactivated polio vaccine in Japan. *Appl. Environ. Microbiol.* 81, 1859–1864.
- Neverov, A., and Chumakov, K. (2010). Massively parallel sequencing for monitoring genetic consistency and quality control of live viral vaccines. *Proc. Natl. Acad. Sci. USA* 107, 20063–20068.
- Nielsen, R., and Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148, 929–936.

- Nielsen, R., and Yang, Z. (2003). Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Mol. Biol. Evol.* 20, 1231–1239.
- Park, D.J., Dudas, G., Wohl, S., Goba, A., Whitmer, S.L., Andersen, K.G., Sealton, R.S., Ladner, J.T., Kugelman, J.R., Matranga, C.B., et al. (2015). Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell* 161, 1516–1526.
- Pilipenko, E.V., Blinov, V.M., Romanova, L.I., Sinyakov, A.N., Maslova, S.V., and Agol, V.I. (1989). Conserved structural domains in the 5'-untranslated region of picornaviral genomes: an analysis of the segment controlling translation and neurovirulence. *Virology* 168, 201–209.
- Pybus, O.G., Rambaut, A., Belshaw, R., Freckleton, R.P., Drummond, A.J., and Holmes, E.C. (2007). Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. *Mol. Biol. Evol.* 24, 845–852.
- Ren, R.B., Moss, E.G., and Racaniello, V.R. (1991). Identification of two determinants that attenuate vaccine-related type 2 poliovirus. *J. Virol.* 65, 1377–1382.
- Rouzine, I.M., Rodrigo, A., and Coffin, J.M. (2001). Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiol. Mol. Biol. Rev.* 65, 151–185.
- Runckel, C., Westesson, O., Andino, R., and DeRisi, J.L. (2013). Identification and manipulation of the molecular determinants influencing poliovirus recombination. *PLoS Pathog.* 9, e1003164.
- Seoighe, C., Ketwaroo, F., Pillay, V., Scheffler, K., Wood, N., Duffet, R., Zveibil, M., Martinson, N., McIntyre, J., Morris, L., and Hide, W. (2007). A model of directional selection applied to the evolution of drug resistance in HIV-1. *Mol. Biol. Evol.* 24, 1025–1031.
- Stern, A., Doron-Faigenboim, A., Erez, E., Martz, E., Bacharach, E., and Pupko, T. (2007). Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res.* 35, W506–11.
- Stern, A., Mayrose, I., Penn, O., Shaul, S., Gophna, U., and Pupko, T. (2010). An evolutionary analysis of lateral gene transfer in thymidylate synthase enzymes. *Syst. Biol.* 59, 212–225.
- Stern, A., Bianco, S., Yeh, M.T., Wright, C., Butcher, K., Tang, C., Nielsen, R., and Andino, R. (2014). Costs and benefits of mutational robustness in RNA viruses. *Cell Rep.* 8, 1026–1036.
- Sunnåker, M., Busetto, A.G., Numminen, E., Corander, J., Foll, M., and Dessimoz, C. (2013). Approximate Bayesian computation. *PLoS Comput. Biol.* 9, e1002803.
- Taffs, R.E., Chumakov, K.M., Rezapkin, G.V., Lu, Z., Douthitt, M., Dragunsky, E.M., and Levenbook, I.S. (1995). Genetic stability and mutant selection in Sabin 2 strain of oral poliovirus vaccine grown under different cell culture conditions. *Virology* 209, 366–373.
- Tao, Z., Zhang, Y., Liu, Y., Xu, A., Lin, X., Yoshida, H., Xiong, P., Zhu, S., Wang, S., Yan, D., et al. (2013). Isolation and characterization of a type 2 vaccine-derived poliovirus from environmental surveillance in China, 2012. *PLoS ONE* 8, e83975.
- Teeling, E.C., Madsen, O., Van den Bussche, R.A., de Jong, W.W., Stanhope, M.J., and Springer, M.S. (2002). Microbat paraphyly and the convergent evolution of a key innovation in Old World rhinolophoid microbats. *Proc. Natl. Acad. Sci. USA* 99, 1431–1436.
- Tulloch, F., Atkinson, N.J., Evans, D.J., Ryan, M.D., and Simmonds, P. (2014). RNA virus attenuation by codon pair deoptimisation is an artefact of increases in CpG/UpA dinucleotide frequencies. *eLife* 3, e04531.
- Weinreich, D.M., Delaney, N.F., Depristo, M.A., and Hartl, D.L. (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312, 111–114.
- Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39, 306–314.
- Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.M.K. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155, 431–449.
- Yang, C.F., Naguib, T., Yang, S.J., Nasr, E., Jorba, J., Ahmed, N., Campagnoli, R., van der Avoort, H., Shimizu, H., Yoneyama, T., et al. (2003). Circulation of endemic type 2 vaccine-derived poliovirus in Egypt from 1983 to 1993. *J. Virol.* 77, 8366–8377.
- Zou, Z., and Zhang, J. (2015). Are Convergent and Parallel Amino Acid Substitutions in Protein Evolution More Prevalent Than Neutral Expectations? *Mol. Biol. Evol.* 32, 2085–2096.

STAR★METHODS**KEY RESOURCES TABLE**

Bacterial and Virus Strains		
Oral Poliovaccine Sabin Type 2 reference strain (OPV2)	NIBSC	01/530
OPV2-U398C	This paper	pRA-Sabin2-398
OPV2-A481G	This paper	pRA-Sabin2-481
OPV2-U398C/A481G	This paper	pRA-Sabin2-398/481
OPV2-U2909C	This paper	pRA-Sabin2-2909
OPV2-U398C/A481G/U2909C	This paper	pRA-Sabin2-398/481/2909
Critical Commercial Assays		
KAPA Stranded mRNA-Seq Kit with KAPA mRNA Capture Beads	KAPA Biosystems	Cat# KK8420
HiSeq Rapid Cluster Kit v2	Illumina	Cat# GD-402-4002
HiSeq Rapid SBS Kit v2	Illumina	Cat# FC-402-4022
Miseq Reagent Kit v2 (300 cycle)	Illumina	Cat# MS-102-2002
ZR Viral RNA Kit	ZYMO Research	Cat# R1035
SuperScript III One-Step RT-PCR System with Platinum Taq High Fidelity DNA Polymerase	Invitrogen	Cat# 12574030
AMPure XP magnetic beads	Beckman Coulter	Cat# A63880
Qubit High Sensitivity dsDNA assay	Life Technologies	Cat# Q32854
Deposited Data		
CirSeq passages of OPV2	This paper	SRA: PRJNA313030
OPV2 Gate Keeper clones competition assays	This paper	SRA: SRP098858
Poliovirus shed from OPV vaccinees	This paper	SRA: PRJNA369470
Experimental Models: Cell Lines		
Human: HeLa S3 cells	ATCC	CCL2.2
Human: Hep 2C cells	NIBSC	740502
Experimental Models: Organisms/Strains		
Mouse: PVRTg21-IFNR-ko	Gift from Satoshi Koike, Tokyo Metropolitan Institute of Medical Science	N/A
Oligonucleotides		
NEXTflex RNA-Seq Barcodes	BIOO Scientific	Cat# 512914
Primer 398F (5'-CGCCATAGGACGCTAGATGTGAACAAGGTGTG AAGAGC-3')	This paper	N/A
Primer 398R (5'-CTTGTTCACATCTAGCGTCCTATGGCGTTAGCC ATAGGTAGG-3')	This paper	N/A
Primer 481F (5'- CCTAACACGGAGCAGGCAGTCGCGAACCA TGACTGG-3')	This paper	N/A
Primer 481R (5'- CGCGACCGCCTGCTCCGTGGTTAGGATTAGCC GCATTC-3')	This paper	N/A
Primer 2909F (5'- CCTCAAACACTACGTGCAAATAACGGACAT GCATTG-3')	This paper	N/A
Primer 2909R (5'- GTTATTTGCATCAGTGTAGTTGAGGTGACCAC AAAAGTG-3')	This paper	N/A
PCR F (5'-AGAGGCCACGTGGCGGCTAG-3')	This paper	N/A
PCR 3' (5'-CCGAATTAAAGAAAATTACCCCTACA-3')	This paper	N/A
Recombinant DNA		
Plasmid: pRA-Sabin2	This paper	pRA-Sabin2

(Continued on next page)

Continued

Plasmid: pRA-Sabin2-U398C	This paper	pRA-Sabin2-398
Plasmid: pRA-Sabin2-A481G	This paper	pRA-Sabin2-481
Plasmid: pRA-Sabin2-U398C/A481G	This paper	pRA-Sabin2-398/481
Plasmid: pRA-Sabin2-U2909C	This paper	pRA-Sabin2-2909
Plasmid: pRA-Sabin2-U398C/A481G/U2909C	This paper	pRA-Sabin2-398/481/2909
Software and Algorithms		
Prism	GraphPad Software	Version 5
R statistical package	https://www.r-project.org	V3.3.3
MAFFT	http://mafft.cbrc.jp/alignment/software/	V6.684b
PhyML	Guindon and Gascuel, 2003	V3.0
Simplot	Lole et al., 1999	V3.5.1
Geneious	Biomatters	R7
Other		
ParaSel software	This paper; www.sternadi.com/parasel	N/A
Sequence alignment and phylogeny	This paper; www.sternadi.com/cell2017	Alignment/phylogeny
Fitness of viral mutations passaged at 33°C	This paper; www.sternadi.com/cell2017	33C
Fitness of viral mutations passaged at 39.5°C	This paper; www.sternadi.com/cell2017	39.5C

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Raul Andino (raul.andino@ucsf.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Mice**

PVRTg21-IFNR-ko mice were obtained from Dr. Satoshi Koike, then bred and maintained in the AAALAC-certified animal facility at UCSF. Mice were specific pathogen free and maintained under a 12 hr light/dark cycle with standard chow diet provided. Both male and female mice were used for all experiments. All animal experiments were conducted in accordance with the guidelines of Laboratory Animal Center of National Institutes of Health. The Institutional Animal Care and Use Committee of University of California of San Francisco approved all animal protocols. (Approved protocol No. AN128674-01A).

Cells and Virus

HeLa S3 and Hep2C cells were maintained in Dulbecco's modified Eagle's medium: Nutrient mixture F12 (DMEM/F12, UCSF Cell Culture Facility) supplemented with 10% fetal bovine serum (FBS) and 1X Penicillin/Streptomycin (Invitrogen) at 37°C with 5% CO₂. Oral poliovirus vaccine type 2 (OPV2) SO+2 was obtained from National Institute for Biological Standards and Control (NIBSC, UK). OPV2 was propagated and titered in HeLa S3 cells at 33°C with 5% CO₂ by standard plaque forming assay.

METHOD DETAILS**Phylogenetic analysis****cVDPV dataset**

Genomic consensus sequences were obtained by mining the literature for circulating vaccine derived polioviruses, and by searching the GenBank repository for sequences labeled as such. All sequences were manually validated, and laboratory manipulated sequences and sequences from immune-compromised individuals and atypical cVDPVs (less than 0.5% divergence compared to the OPV2 sequence, or annotation as such) were removed. This yielded a total of 424 full-length genomic sequences. All sequence alignments were performed using MAFFT version 6.684b and are available at www.sternadi.com/cell2017.

Phylogeny reconstruction

Initial reconstruction was performed based on the amino-acid sequence of 424 amino-acid sequences of the non-recombinant capsid region. However, the resulting phylogeny supported only one emergence of type 2 cVDPV, followed by subsequent spread from country to country across highly distinct and distant locations across the globe. This scenario was highly incompatible with the epidemiological data that supported multiple emergences of the virus ([Burns et al., 2013](#)). Previous research has shown that

phylogenetic reconstruction will fail to recover the true history when parallel advantageous substitutions occur on the background of limited genetic diversity (Bull et al., 1997). We thus used the synonymous sites of the non-recombinant portion of the capsid, spanning genomic coordinates 1222 till 3384 (the coordinates were chosen to ensure non-recombinant regions, see below).

Our analysis relied on the fact that on average, selection at synonymous sites would be weaker than selection at non-synonymous sites. The tree topology and branch lengths were estimated using the maximum-likelihood methodology implemented in PhyML version 3.0 (Guindon and Gascuel, 2003), with 100 bootstrap replicates generated to assess confidence of splits (branches). The tree was rooted at the OPV2 sequence (accession AY184220). The resulting tree was bifurcating, as imposed by the reconstruction algorithm of PhyML. Epidemiological evidence strongly supports several independent emergence events across disparate locations (e.g., Egypt and China), and it is thus unlikely that this bifurcation at the root is real. Indeed, the ancestral branches surrounding the OPV2 sequence were all near zero and bootstrap support was extremely low. Hence when cumulative branch length from the root to a node was less than $\epsilon = 6 \times 10^{-5}$ the intermediate branches were collapsed. This value was chosen at it reflects an expectation of one substitution only, suggesting that values lower than it are indicative of zero shared substitutions. The resulting tree is available as supplemental data in Figure S1A.

Recombination analysis

Since the recombination donor sequence (OPV2) was well established, identity versus OPV2 allowed detecting recombination breakpoints. Simplot software version 3.5.1 (Lole et al., 1999) was run on all full-length cVPDVs with OPV2 as a reference sequence. A window size of 200 bp and a step size of 20 were used (Figure S1B). Recombination breakpoints were inferred where a sharp drop in % identity occurred, i.e., when sequence identity dropped below more than 2.5 standard deviations of the median identity over the capsid region. This allowed constructing several datasets of non-recombinant regions. The largest such dataset was of the P1 capsid region, where all of the 424 sequences were of OPV2 origin from position 1222 (middle of VP2) through 3384 (the end of VP1). To explore parallel selection outside the capsid region, we further constructed six additional datasets with non-recombinant sequences: (a) 27 sequences from positions 1-1222, (b) 44 sequences from positions 748-1222, (c) 25 sequences from positions 3385-4383, (d) 8 sequences from positions 3385-5382, (e) 4 sequences from 3385-6381, and (f) 3 sequences from 3385-7439. The smaller number of sequences in datasets (c) through (f) precluded statistical significance. We report results in Table S3 from the P1 dataset and from dataset (a) (which incorporates part of (b)). The phylogeny for dataset (a) is available at www.sternadi.com/cell2017.

To study the process of recombination, we next projected the recombination breakpoint on to the cVPDV phylogeny. Notably, this assumes that sequences with similar recombination breakpoints are all derived from one shared ancestor where recombination occurred. Different recombination breakpoints indicate dissimilar ancestry with regards to the recombination. We first classified each cVPDV sequence based on the inferred breakpoint, and labeled breakpoints in windows of 100 bps (e.g., a breakpoint inferred at base 4736 and a breakpoint inferred at base 4760 would both be labeled as a breakpoint at position 4700). Figure S2 displays the pattern of recombination in the (A) 5' UTR region, and (B) 3' P2-P3 region. Interestingly, this shows that some clades may have undergone at least two recombination events during evolution, such as the clade marked by an arrow in Figure S2.

Substitution mapping

In order to infer the mutational history of a site we mapped substitution events to specific branches of the phylogeny. At each site we calculated the joint posterior probability of two characters a, b (where $a, b \in \{A, C, G, T\}$) populating the nodes along each branch of the tree: $P(a, b | X) = P(a, b)/P(X)$ where X is the phylogeny and sequence alignment, using the HKY model (Hasegawa et al., 1985). Full details of this calculation were described previously (Stern et al., 2010). Only sites and branches where $P(a, b | X) > 0.6$ were reported.

Co-evolution of pairs of sites

In order to test for correlated evolution between pairs of sites in the non-recombinant region, we measured the fraction of shared time $O(i, j)$ when site 1 is in state i and site 2 is in state j (Huelsenbeck et al., 2003). This fraction was then compared to the expected fraction of time these two states should share together $E(i, j)$, based on the marginal probabilities of finding each character on its own. The metric for assessing correlated evolution is then the difference:

$$d_{ij}(\tau, v, h_1, h_2) = O(i, j) - E(i, j) \quad (1)$$

where τ is the tree topology, v is a set of branch lengths representing the time, and h_1 and h_2 are character mappings for characters 1 and 2 based on the mapping described above using joint posterior probabilities. A key difference in our approach here as compared to (Huelsenbeck et al., 2003) is that we do not integrate over different trees and branch lengths but rather use only one such realization. We compared values of d_{ij} to a set of values obtained by simulating over the same set of tree topology and branch lengths, with a null model of evolution where directional selection is absent and sites are simulated independently. Our results (Figure S1D) show no difference between the simulated and real data; thus we could not detect pairs of sites displaying a significant pattern of co-evolution in the capsid region.

Calculating the ad hoc probability of parallel substitutions in two or more lineages of the phylogeny

Let p denote the probability of a given substitution during one year of viral evolution. Based on the molecular clock calibration of PVs whereby the expected number of substitutions per site per year is 0.011 (Jorba et al., 2008), let us set $p = 0.011$. The sum of branch lengths in our phylogeny is 3.40761, which can be translated to an accumulated 310 years of viral evolution (which occurred during

overlapping time frames in independent infections). We can now use the binomial distribution to estimate the probability of two or more substitutions at the same locus throughout the phylogeny as:

$$P(X \geq 2) = 1 - P(X = 1) - P(X = 0) =$$

$$= 1 - \binom{310}{1} p \cdot (1-p)^{309} - \binom{310}{0} (1-p)^{310} \approx 0.85 \quad (2)$$

Notably, this is a highly simplistic estimate that ignores rate variation, transition/transversion rates, and other biological properties of viral replication. However, it still captures the fact that the probability of parallel substitutions is extremely high in polioviruses in the scenario described herein.

An evolutionary Markov model for detecting parallel directional selection

Under strong parallel (directional) selection, we expect one allele to be highly preferred over all other alleles (Kosakovsky Pond et al., 2008; Seoighe et al., 2007). Hence the rate of a substitution leading to this allele along the phylogeny will be higher, in proportion to the selective advantage it confers, and correspondingly substitutions that change this allele to an allele of lower fitness will have a reduced rate. This is captured by assuming a continuous time Markov process, defined by an instantaneous rate matrix Q with elements $Q(i,j)$, for $i \neq j$, given by:

$$Q(i, j) = \begin{cases} H(i, j) & i, j \text{ are not the preferred allele } k \\ H(i, j) \cdot \frac{S}{1 - e^{-S}} & j=k, \text{ directional selection for allele } k \\ H(i, j) \cdot \frac{S}{e^S - 1} & i=k, \text{ negative selection against losing allele } k \end{cases} \quad (3)$$

where H is any standard rate matrix over any alphabet (e.g., nucleotides or amino-acids). $S = Ns$ is the genomic population size (N) scaled selection coefficient, and $S/(1 - e^{-S})$ is the rate of fixation of a non-neutral mutation (Kimura, 1962; Nielsen and Yang, 2003). The diagonal elements of Q are defined so that the sum of entries in each row is zero.

Since we do not know a priori which allele is preferred, the likelihood function was calculated using a mixture model in which all possible assignments of the preferred allele were given equal prior weight. We denote the model in which allele k , $k \in \{1, 2, \dots, K\}$, is preferred as M_k and the standard baseline model without selection by M_{base} . Further, we assume a proportion P_{DS} of sites (to be estimated) are under directional selection. Thus the total likelihood function, given an alignment of sequences and a phylogeny (X), is

$$P(X) = \sum_{k=1}^K \frac{P_{DS}}{K} P(X | M_k) + (1 - P_{DS}) \cdot P(X | M_{base}) \quad (4)$$

We further assume among-site rate variation (ASRV), where rates are sampled from a discrete approximation of the gamma distribution with a mean of one (Yang, 1994).

Finally, to model incomplete selection near the tips of the tree, we introduced the parameter β , which effectively increases the branch lengths at the tips of the tree by rescaling the among-site rate variation distribution. Site-specific evolutionary rates are directly linked to branch lengths, and a site evolving with rate r can be modeled by multiplying all branch lengths in a tree by r . We model incomplete purifying selection by modifying r only at the tips of the tree as follows: $r_{relax} = r + (r_{max} - r) \cdot \beta$, where r_{max} is the maximal rate category of the discretized gamma distribution. This modification was applied to the edges leading to leafs of the tree at all sites of the alignment, and shifts the ASRV distribution toward values representing less purifying selection without affecting the maximal rate. Importantly this captures incomplete purifying selection that may occur during shorter timescales in viral evolution, due to smaller effective population sizes (Pybus et al., 2007).

Thus, parameters in the model are: S , P_{DS} , α (shape parameter of the gamma distribution), parameters of the baseline model (here the transition/transversion rate K in the HKY model (Hasegawa et al., 1985)), relaxation parameter β , and a scaling factor for the branch lengths of the phylogeny (γ). All parameters were optimized under maximum likelihood using the Brent optimization scheme (Brent, 1971) and the BFGS optimization scheme (Byrd et al., 1995), with the likelihood function initialized at multiple starting points to reduce the chance of the algorithm being trapped in a local, but not global, maximum. Similar results were obtained with both optimization methods (data not shown).

To test whether the data exhibited significant support for directional selection, we used model selection based on the corrected Akaike Information Criterion (AIC). This approach was necessary since none of the asymptotic methods are applicable for using the χ^2 approximation for a likelihood ratio test, typically used when comparing two phylogenetic models. This pathology occurs since S is only estimated in the alternative model and not in the null model.

Testing for site-specific directional selection

Given significant support for the selection model, as specified by the AIC comparison, we calculated the posterior probability that a site i with site pattern X_i , experienced directional selection by calculating the posterior probability of a site evolving under each of the K directional selection models. This is given by $P(M_k | X_i) = ((P_{DS}/K)P(X_i | M_k)/P(X_i))$.

We note that our model, similar to most evolutionary models, assumes that sites in the alignment are independent. This assumption hence means our method fails to take into account different forms of linkage among sites. However, polioviruses possess extremely high rates of recombination (Kirkegaard and Baltimore, 1986; Runckel et al., 2013), suggesting that linkage among sites may be frequently broken up in these viruses.

Testing ParaSel on simulated data

We began by simulating 100 datasets of nucleotide sequences under a neutral model of evolution (HKY). Sequences were simulated using the OPV2 root sequence along the phylogeny of the cVDPVs as obtained above. All parameters were inspired by those inferred above for the cVDPV data. For computational reasons, the sequence length was truncated to 200.

We next sought to test the effect of an epistatic interaction on ParaSel inference. To this end, 100 datasets were simulated similar to above with an HKY model; however, only sequences which maintained a paired base composition at a given pair of sites that allows RNA base-pairing (G:C, G:U or A:U) were retained, thus creating simulated data with purifying selection against non-base pairs.

Finally, we simulated 100 datasets with 10% of sites under directional selection for "A," and 90% of sites under a neutral HKY model. Furthermore, we tested the specificity and sensitivity of the method to detect the specific sites under selection, and used a receiver-operator-characteristic curve to explore different thresholds of the posterior probability.

We then ran ParaSel on the simulated dataset and tested in how many datasets the method showed significantly lower AIC scores under the ParaSel model as compared to the null model. All results are summarized in Table S2.

Recombination as an adaptive event

We wished to infer whether the recombination events with HEV-C sequences are under positive parallel selection as well. Sequences were classified for presence or absence of a recombination breakpoint. Notably all but three of the cVPDVs were recombinant at the 3' region, whereas in the 5' UTR region all but twenty-seven of the cVDPVs underwent recombination. This allowed us to map the recombination events onto the phylogeny. Indeed, this showed that almost all of the lineages coming out of the ancestral OPV2 sequence were recombinant (Figure S2). We could not utilize ParaSel for inferring selection of these recombination events since we wanted to avoid inference based on one character only. However, the rate of substitutions is likely much higher than the combined rate of co-infection and the rate of strand displacement (that leads to recombinant genomes for polioviruses). Thus, given our inferences of positive selection for many of the point mutations, the exceptional parallelism of the recombination events strongly suggests that they are under positive selection as well. This is line with experimental data confirming this conjecture (Jegouic et al., 2009).

Inferring site-specific diversifying selection using dn/ds

The "selecton" web-server (Stern et al., 2007) was used to run the M8 and M8a models (Yang et al., 2000) in order to test for the existence of diversifying positive selection. A likelihood ratio test between the two models yielded strong significant support for positive selection ($p < 10^{-8}$). Next, site-specific values of d_n/d_s were inferred based on the M8 model using the posterior probability distribution of d_n/d_s values at each site. Sites were considered to be under positive diversifying selection when the lower bound of the 95% credible interval of this distribution was higher than one.

Calculating odds ratios of dinucleotides in cVDPV genomes

Odds ratios for dinucleotides were calculated as $R(XpY) = (f_{XY}/f_X \cdot f_Y)$ where X and Y stand for single nucleotides, f_{XY} stands for the joint frequency of a dinucleotide and $f_X \cdot f_Y$ is the expected frequency based on the product of the two bases' frequencies. As proposed by Burge et al. (Burge et al., 1992), values below or above the range of 0.81–1.19 were considered low or high abundance, respectively. Indeed, overall levels of both CpG and UpA dinucleotides were consistently lower than 0.81 (Figure S7A). The CpG levels in cVDPV strains were found to decrease constantly over time, with a linear regression model explaining 15% of the variance (Figure S7B and S7D). A 2nd degree polynomial model led to a non-significant fit, suggesting that selection operates in a linear fashion against the accumulation of CpG sites.

On the other hand, UpA levels showed a more complex pattern. In the first 2–3 years, levels of UpA decrease over time (Figure S7C), particularly in the capsid region. Surprisingly, after that levels of UpA increase till reaching what appear to be a steady state at 0.76 O/E (Figure S7E), which still reflects lower than expected dinucleotide frequencies. Quadratic models with increasingly high order polynomials indeed yielded a significant fit ($p < 0.000142$, $p < 0.001370$, $p < 0.004$ for 2nd, 3rd, and 4th degree polynomials, respectively), suggesting a more complex model of evolution with potential interaction with other as yet unknown factors. All in all, it appears that the dinucleotide levels in the OPV2 genome have an optimal steady-state level that the genomes converge upon.

Experimental validation

Next generation sequencing of excreted viruses

Stool samples were collected from primary vaccinees 14 days after vaccination with trivalent OPV (Dunn et al., 1990) and processed by adding 1g to 10 mL PBS containing 1g of glass beads and 1 mL chloroform in a 50mL Falcon tube. Tubes were shaken vigorously at 4°C for 20 min using a mechanical shaker then centrifuged for 20 min at 1500 x g in a refrigerated centrifuge. RNA was purified from stool extracts using Roche High Pure viral RNA kits. Water only controls were extracted, amplified, and sequenced in parallel with each set of samples. Almost full-length poliovirus genomes were amplified in duplicate by one-step RT-PCR using a

SuperScript III One-Step RT-PCR System with Platinum Taq High Fidelity DNA Polymerase (Invitrogen) and primers PCR F (5'-AGA GGC CCA CGT GGC GGC TAG -3') and PVR 3' (5'-CCG AAT TAA AGA AAA ATT TAC CCC TAC A -3'). Products were purified using AMPure XP magnetic beads (Beckman Coulter), quantified using Qubit High Sensitivity dsDNA assay (Life Technologies) and diluted to 0.2 ng/μl in molecular grade 10 mM Tris- EDTA, pH8.0.

Sequencing libraries were prepared using Nextera XT reagents (Illumina) and the manufacturer's protocol, and sequenced on a MiSeq using a 2 × 251 paired-end v2 Flow Cell (Illumina). Quality trimming and assembly were carried out as in [Mee et al. \(2015\)](#). Reads were then mapped to an OPV2 reference sequence using Geneious R7 (Biomatters) software and SNPs present at ≥ 1.0% identified. Only those SNPs present in both replica amplicons were retained.

Recombinant type 2 virus genomes were identified by deep sequencing type 2 viruses and mapping reads to OPV1, 2 & 3 reference sequences as above. Type 2 viruses were obtained by incubation of stool extracts with high titer polyclonal antisera against type 1 and type 3 poliovirus followed by infection of HEp2c cells and incubation at 35°C until full cytopathic effect. Capsid sequences of these viruses mapped to OPV2 alone.

Viral serial passaging in tissue culture

To test adaptation of OPV2 to high temperature, we passaged oral polio vaccine (OPV) type 2 in cell culture. Originally we performed serial passaging in Vero cells at 37°C. Notably this led to very poor growth of the virus, and due to the low virus yield we lacked the right population size for accurate next generation sequencing. We next passaged OPV2 at HeLa S3 (ATCC, CCL2.2) cells, which are more supportive of viral replication, at both 33°C (control condition) and 39.5°C. For both conditions, 10⁷ HeLa S3 cells were seeded the day before the experiment and infected with OPV2 at multiplicity of infection (MOI) of 0.1 for one hour to allow virus adsorption, and then replaced with virus culture medium. The infected cells were maintained for 24 hr, allowing two replication cycles per passage, leading to a total of seven passages and fourteen replication cycles. Cells were then harvested by freezing at -80°C. After three freeze-thaw cycles, virus suspension was clarified by centrifugation at 3,500 xg for 10 min at 4°C, and stored at -80°C for future passages. Plaque assays were performed to determine virus titer of each passage at an MOI of 0.1 for subsequent passages.

Viral population deep sequencing and inference of mutation frequencies

Highly accurate sequencing of viral populations was obtained by our newly developed CirSeq approach ([Acevedo and Andino, 2014](#); [Acevedo et al., 2014](#)). Briefly, this approach is based on the conversion of short fragments of the viral RNAs to circular molecules. When copied with reverse transcriptase, tandemly repeated cDNAs are produced. Mutations in the original viral RNA are shared by all tandem repeats, while errors produced during reverse transcription, PCR or sequencing will be randomly distributed along the tandem repeats. Subsequent computational mapping can thus reduce sequencing error to a point that is much lower than the estimated mutation rate of an RNA virus.

Viral populations were amplified once in HeLa S3 to increase the percentage of viral RNA for sequencing. Library preparation and sequencing were performed as described in ref. ([Acevedo et al., 2014](#); [Stern et al., 2014](#)) on a Hiseq (Illumina). Initially passages 2,4, and 7 were sequenced for both temperatures. Subsequently passages 1,3,5,6,7, were sequenced at 39.5°C, and passages 2 and 4 were discarded for the 39.5°C analysis due to batch effects. Reads were mapped to the OPV2 reference genome (GenBank accession AY184220) using Q23 as a quality threshold, yielding a Cirseq error rate of approximately 10⁻⁷. We then estimated allele frequencies as in ([Stern et al., 2014](#)), yielding values for minor allele frequencies that typically ranged between 10⁻⁶ and 10⁻² ([Figure S3B](#)). We further assessed the reliability of allele frequencies. Based on the geometric distribution, in order to guarantee that a viral template occurring at frequency f is detected with probability p or better, it is necessary to sequence at least $\log(1-p)/\log(1-f) - 1$ templates ([Lorenzo-Redondo et al., 2016](#)). We estimated the number of templates sequences as the coverage at a certain locus, allowing us to calculate p for each mutation. Only allele frequencies where $p > 0.95$ were retained for analysis of fitness.

Generation of gatekeeper mutant viruses

Gatekeeper mutations were introduced into the OPV2 infectious cDNA clone (pRA-Sabin2) with PCR-directed mutagenesis. PCR was performed using pRA-Sabin2 as a template with primers containing gatekeeper mutations (primer sequences are provided in [Key Resources Table](#)). The PCR product was digested with Dpn I (NEB) at 37°C for 1 hr, and transformed into One shot Top10 competent cells (Invitrogen). Plasmids were verified with Sanger sequencing for the introduced gatekeeper mutation. Protocols for virus generation from infectious cDNA clone and quantification have been previously described ([Burrill et al., 2013a](#)).

Competition assays

HeLa S3 cells were coinfecting with OPV2 and gatekeeper mutant virus at a MOI of 0.05 for 10 hr. Cells were harvested with 3 freeze-thaw cycles and cleared at 3500 rpm for 10 min at 4°C. The viral supernatant was titrated by plaque assay and passaged further at low MOI for 10 hr. This process was repeated to get a total of 5 passages. Viral RNA was purified from supernatants of passages 1, 3, and 5 using the ZR Viral RNA Kit (Zymo Research). The proportions of OPV2 and gatekeeper mutant virus in each viral RNA sample were determined by standard next generation sequencing. RNA libraries were prepared with the KAPA Stranded mRNA-Seq Kit (KAPA Biosystems) and NEXTflex RNA-Seq Barcodes (BIOO Scientific), and sequenced on a MiSeq (Illumina).

Virulence of gatekeeper mutant viruses in mice

A novel mouse model of infection for OPV2 was established by using poliovirus receptor (PVR) transgenic Tg21 mice deficient in the alpha/beta interferon (IFN) receptor gene ([Ida-Hosonuma et al., 2005](#)). To determine viral dissemination, a total of 18 10-day-old IFNR-ko mice were intraperitoneally (i.p.) inoculated with 10⁵ plaque-forming unit (PFU) of OPV2. Various tissues were harvested

from 3 mice everyday till day 6 post-inoculation. The harvested tissues were weighted, homogenized in 1 mL of culture medium and cleared at 10,000 xg for 10 min at 4°C. Virus titer in cleared supernatant was determined by plaque assay, and expressed as log PFU per milligram of tissue.

For virulence testing, mice were infected with either OPV2 or gatekeeper mutant virus. I.p. infections were performed using 10-day-old mice with 1×10^6 PFU per mouse (7–8 mice per virus strain, 100 µl per mouse). An additional group of mice were inoculated the same way with 100 µl of viral medium as a control. Mice were monitored daily for signs of paralysis. Mice were euthanized upon appearance of dual hind limb paralysis, a sign of imminent death, and death was recorded for the following day.

QUANTIFICATION AND STATISTICAL ANALYSIS

Inferring mutational fitness from time-series allele frequencies

Mutational fitness values w of different alleles (mutations) were inferred based on the time-series of allele frequencies. We elaborated on a novel approximate Bayesian computation (ABC) method (([Foll et al., 2014](#)), T.Z. and A.S., unpublished data) that takes into account stochastic effects of viral replication and library preparation. Importantly, this model accounts for genetic drift that affects alleles present at low counts during early passages.

Our approach uses as a baseline the methodology described previously ([Foll et al., 2014](#)), which bypasses the need for computationally intensive likelihood calculations, and instead relies on simulations using parameters sampled from the prior distribution of the model parameters. Based on the distance between a simulated trajectory and the real data, a rejection-based scheme is used to create the posterior distribution of the parameter at each site. Here, we use a Wright-Fisher model with a binomial sampling step that further accounts for novel mutations and back mutations at each generation ([Stern et al., 2014](#)) to simulate trajectories of allele frequencies, and assume a uniform prior distribution over our mutational specific parameter w (mutational fitness) spanning [0,2]. At each locus, we retain the top 1% out of 1,000,000 simulations to be used as a posterior distribution for a rejection ABC algorithm. Our approach differs from that described in [Foll et al. \(2014\)](#) on three levels: first, we infer the base-by-base mutation rates directly from the data, rather than assuming a predetermined mutation rate value. This is made feasible due to the high precision of the CirSeq approach ([Acevedo and Andino, 2014; Acevedo et al., 2014](#)) (see section on Inferring mutation rates). Second, we do not estimate the effective population size N_e . Our experimental setup was designed for very high population sizes (10^6 PFUs at each passage), and furthermore we obtain coverage ranging typically from 10^5 – 10^6 reads per locus ([Figure S3A](#)). We hence assume a large population size in our simulations (10^5). Finally, when testing the distance between simulations (sim) to the data (d), we calculate the ℓ_1 distance between the curves defined by $\sum_{p=1}^7 |sim_p - d_p|$ where $p \in \{1 \dots 7\}$ are the passage numbers. Notably, we would like to emphasize that since we are looking at rare new mutations, this implies that mutations are approximately independent.

The ABC approach described herein has been tested extensively validated and found to yield very good estimates of fitness ([Foll et al., 2015](#)). We further tested our approach on simulated data ([Figure S4A](#) and S4B) using a population size and mutation rate similar to that used in this study. Reassuringly the results show that our method is highly successful in recapitulating the true fitness value used in the simulations.

Inferring mutation rates

We infer mutation rates based on the change in frequencies of synonymous mutations across time. Since mutations at neutral sites accumulate freely, we expect the rate of divergence of neutral sites to be precisely equal to the mutation rate ([Kimura, 1968](#)). We thus use synonymous mutations, excluding those that disrupt or create CpG or UpA sites, as a proxy for neutral mutations, and divide the mutations into the four types of transitions ([Figure S4A](#) and S4D). We used linear regression to estimate the slope at each independent trajectory at each site. The mean value across sites was used as an estimate of the mutation rate. We tested our approach on 100 simulated datasets using a Wright-Fisher model as described above, assuming a set of one hundred loci all evolving neutrally ($w = 1$) and a mutation rate of 10^{-5} mutations/base/generation. Two approaches were tested: once, linear regression was used to fit all passages, and second we used the last five time-points only. The latter was done to test the effects of random genetic drift operating at earlier passages when allele counts are low. We further tested a combination of two different scenarios: population size of $N = 100,000$ as used here and a lower population size ($N = 10,000$). Finally, we also tested a noisier scenario where 10% of the hundred simulated loci were deleterious ($w = 0.5$). The results of the simulations are presented in [Figure S4E](#). The results showed excellent recovery of the mutation rate, apart for the scenario when $N = 10,000$ and only the last time-points are used. Based on these results we chose to estimate the mutation rates based on all time-points ([Figure S4C](#)).

We further compared our mutation rates estimates to those obtained using our previous approach that relies on the frequencies of lethal mutations ([Acevedo et al., 2014](#)). The neutral regression approach yielded lower mutation rates for most types of mutations ([Figure S4D](#)). We hypothesize that this occurs due to the experimental protocol of expanding the viral populations before sequencing at high multiplicity of infection (MOI), and since during the second replication cycle at each passage high MOI will occur as well. Under this scenario, genomes with deleterious mutations may be compensated for by other co-infecting functional genomes, leading to a general elevation in frequencies of deleterious mutations ([Stern et al., 2014](#)). Furthermore, we noticed a general elevation of C→T mutations, possibly consistent with our previous report on elevation of this type of mutation due to library preparation ([Lou et al., 2013](#)). This strengthens the use of an approach relying on changes in neutral mutation frequencies across time, which is expected to overcome this effect since it assumes any bias in library preparation will not increase over serial passages.

Inference of beneficial mutations

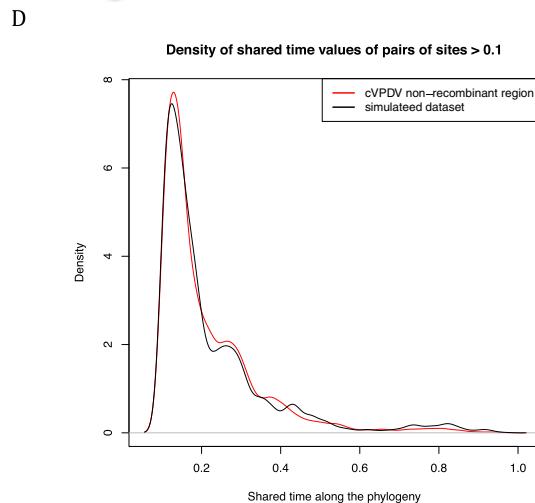
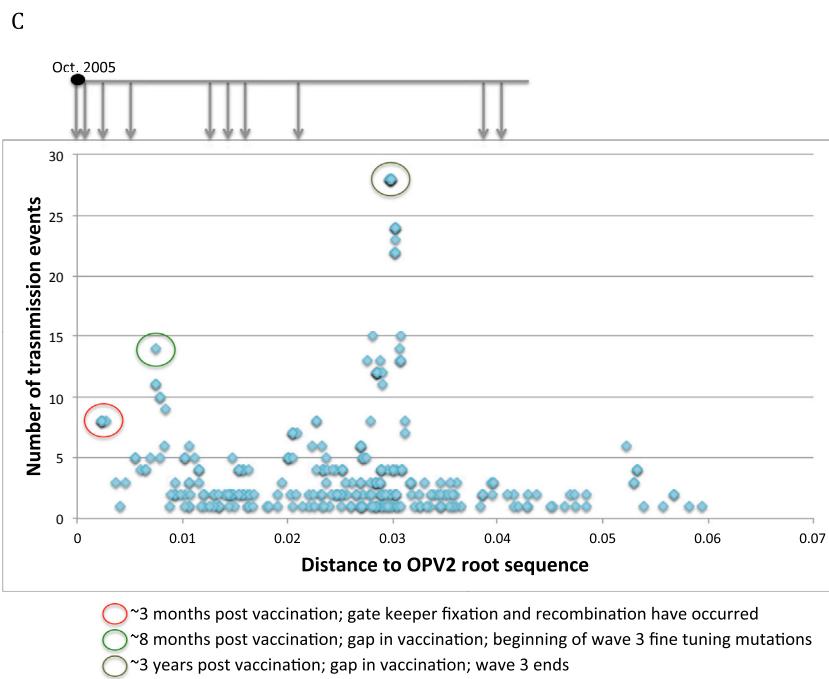
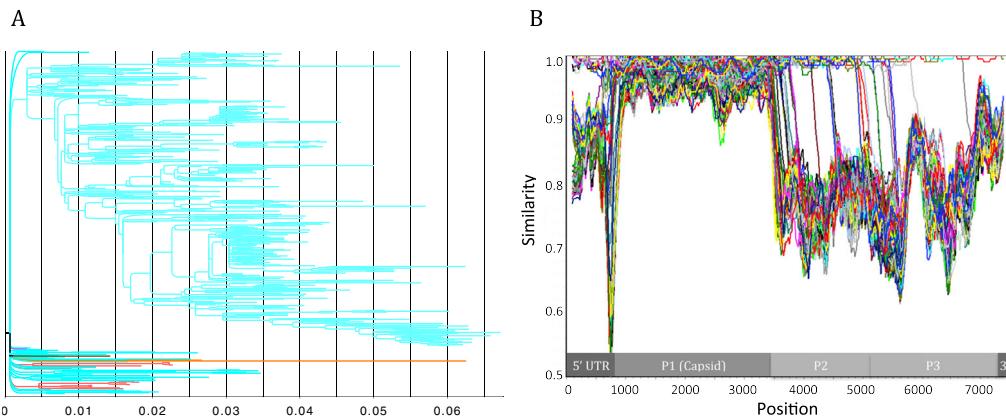
After inferring the mutation rates, we next set out to infer whether the set of mutations found using ParaSel are under positive selection. To this end, similar to (Foll et al., 2014; Sunnåker et al., 2013), we assumed a uniform prior over w , spanning the interval [0,2]. We retained the best 1% of simulation as defined above to create posterior distributions (Figure S5) and median values of these distributions. We report mutations under positive selection based on Bayesian “p values” ($P(w > 1 | \text{data})$) and thus a mutation is considered significant if at least 95% of the posterior distribution exceeds one. We thus considered a mutation to be adaptive at high temperature if it adhered to the following criteria: (a) it was found to be under positive selection in the 39.5°C experiment, and (b) it was not found to be under positive selection in the control 33°C experiment.

DATA AND SOFTWARE AVAILABILITY

All sequence alignments and phylogenies used in this study, and fitness of viral mutations passaged at 33 and 39.5°C are available at www.sternadi.com/cell2017. The ParaSel software is available at www.sternadi.com/parasel. Read files for all sequencing data were deposited in the NCBI Sequence Read Archive (SRA) as listed in the [Key Resources Table](#), SRA: PRJNA313030, SRP098858, and PRJNA369470.

Supplemental Figures

Cell



(legend on next page)

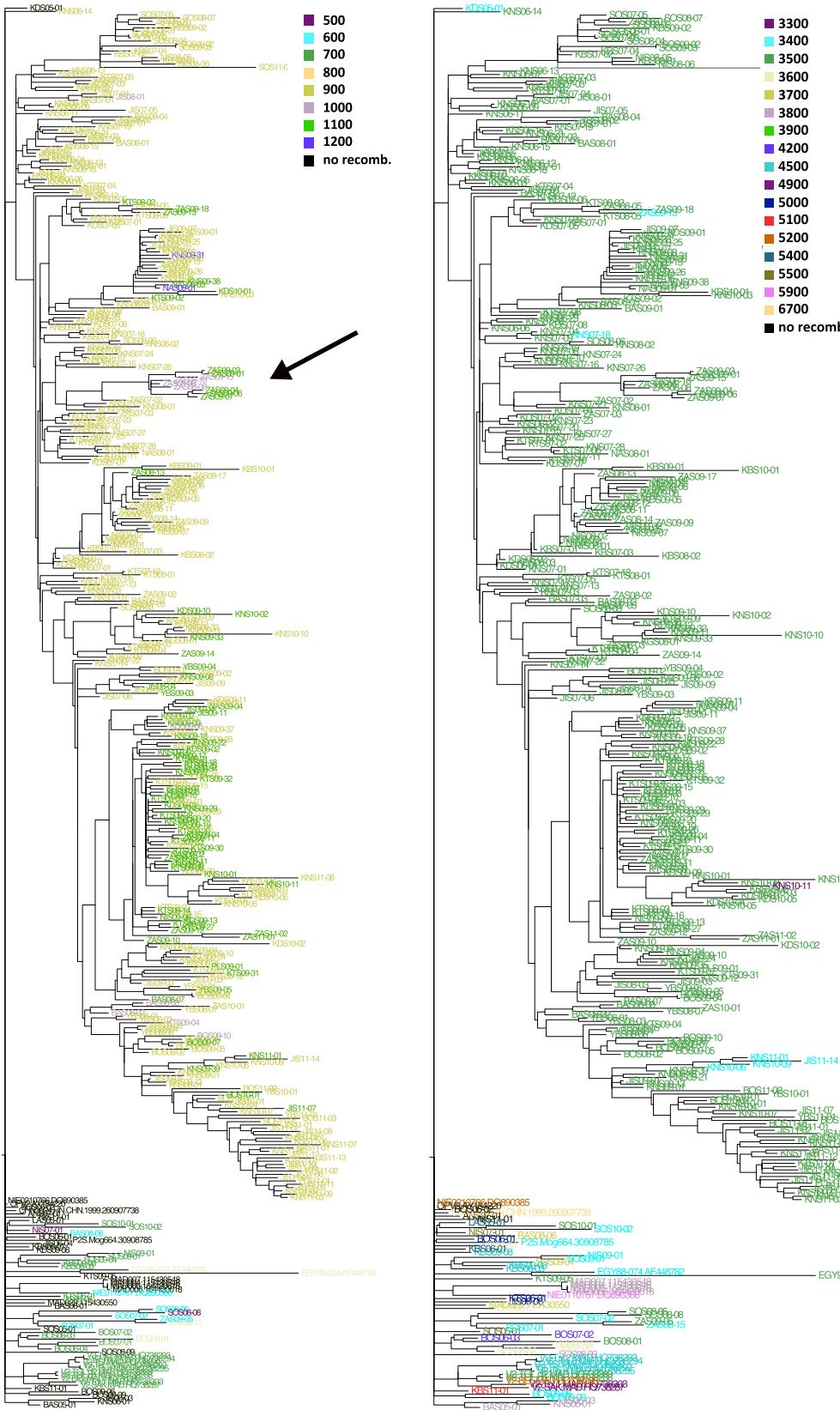
Figure S1. Analysis of cVDPV2 sequences. Related to Figure 2

(A) Full Maximum-likelihood phylogenetic tree based on synonymous polymorphisms for 424 cVDPV sequences based on the non-recombinant capsid region. Related to [Figure 2](#). The x axis represents substitutions per site, where 0.011 substitutions are equivalent to one year of viral circulation based on the PV molecular clock ([Jorba et al., 2008](#)). Colors correspond to the legend in [Figure 2](#).

(B) Similarity plot between cVDPV and OPV2 sequences. Percent similarity between each cVDPV2 sequence (colored line) and the OPV2 sequence, plotted along the genome using Simplot version 1.3. A simple genome schematic is shown at the bottom. Sharp drops in sequence identity represent a recombination breakpoint. Potential recombination partners are in [Table S1](#).

(C) Number of estimated transmission events plotted against time since epidemic initiation, calculated for the largest outbreak in Nigeria. The number of transmissions is estimated based on number of nodes spawning from each father node in the phylogeny ([STAR Methods](#)). Arrows at the top represent mass vaccination campaigns that took place in Nigeria; the black oval at the top represents the estimated date of the initiating vaccine dose for this epidemic (data taken from ([Burns et al., 2013](#))). Encircled peaks in transmission correlate well both with gaps in vaccination, and with gain of function events predicted herein.

(D) Testing for co-evolution in the capsid for pairs of sites. Density plot of values of shared time along the phylogenies across pairs of sites, plotted for values large than 0.1 ([STAR Methods](#))). For each pair of sites in the alignment, the time shared along the phylogeny during certain states was measured (e.g., the time shared along the phylogeny when locus 1 encodes for "A" and locus 2 encodes for "G"). The density plots of both the empirical data and the simulated data (where sites were simulated independently) appear identical. Thus, the null hypothesis whereby there is no co-evolution of pairs of sites cannot be ruled out, suggesting that shared time of sites in the cVDPV capsid is driven mainly by stochastic effects. Due to recombination in the non-capsid region it is impossible to use this approach to test for co-evolution of sites there.



(legend on next page)

Figure S2. Tips of the cVDPV2 Phylogeny Are Color-Coded Based on the Inferred Recombination Breakpoint at the 5' UTR (left) and 3' region (right), Related to Figure 2

The legend denotes the approximate location of the breakpoint in intervals of 100 nucleotides. An arrow marks an example clade where multiple recombination events may have occurred sequentially.

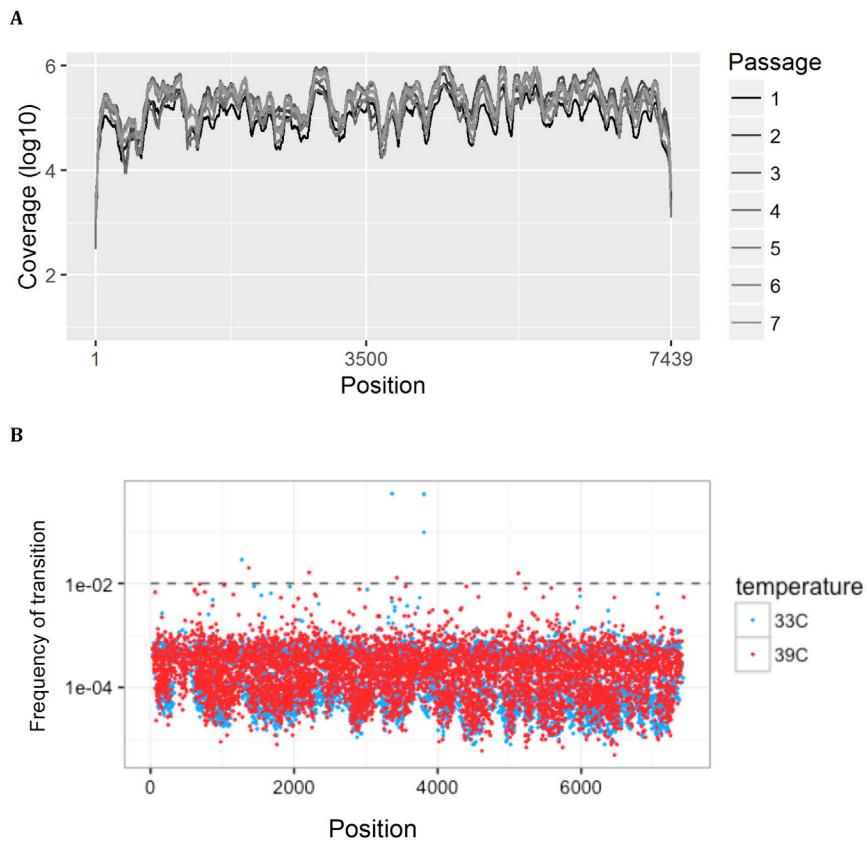
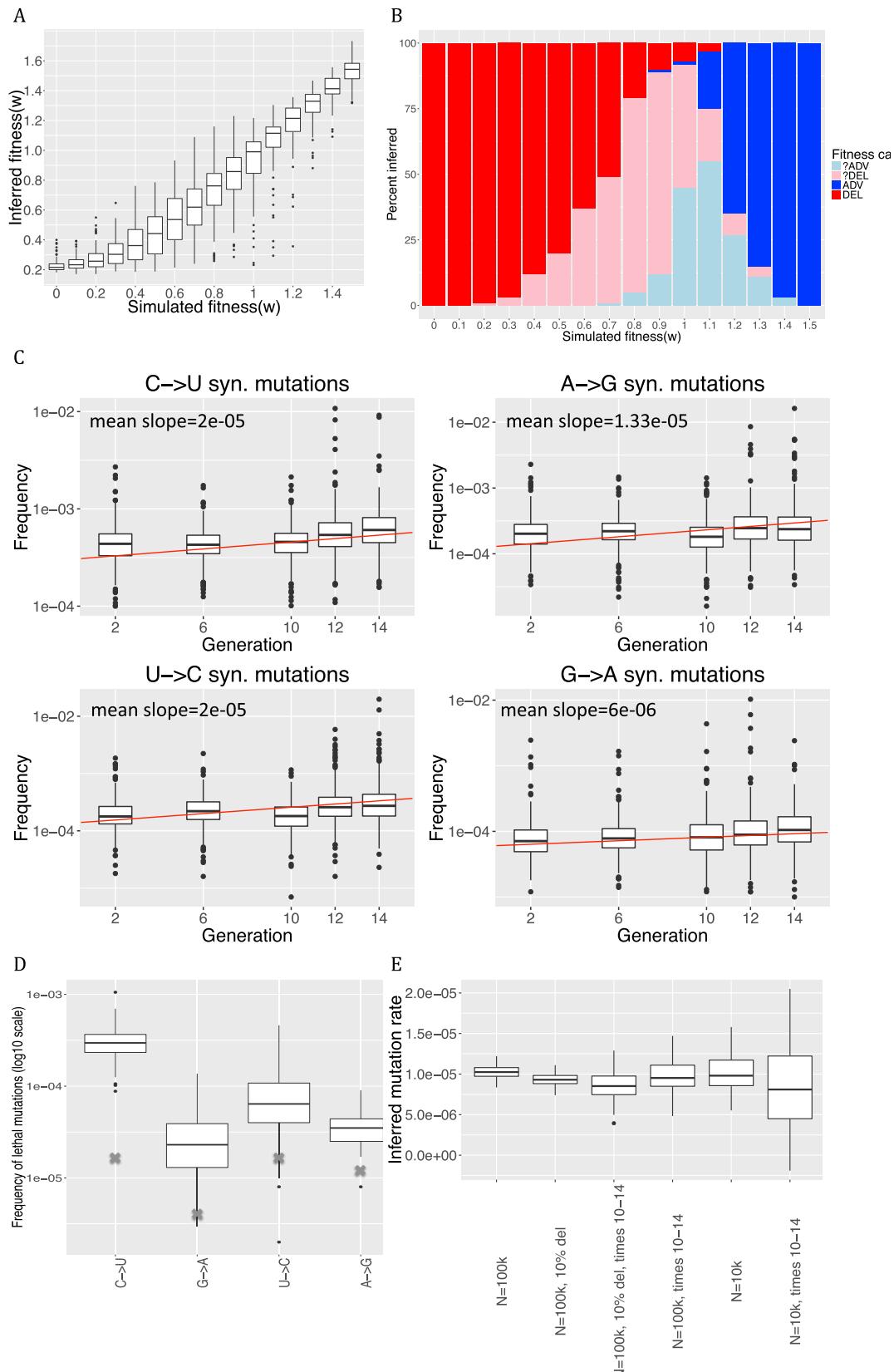


Figure S3. Summary of Data from CirSeq Sequenced Passages, Related to Figure 4

(A) Genome coverage per base for each passage sequenced at 39.5°C. The coverage is shown to be highly consistent among passages, with fluctuations typical of RNA-seq studies.

(B) The mutational spectrum of passage 7 for 33°C and 39.5°C. Frequencies of transition mutations are plotted along the genome. The figure illustrates the fact that the vast majority of mutations are at very low frequency and hence (i) OPV2 is mostly stable at both temperatures, and (b) classical next generation sequencing (NGS) strategies are unable to detect most mutations, as depicted by the gray dashed line showing the typical threshold of detection for NGS.



(legend on next page)

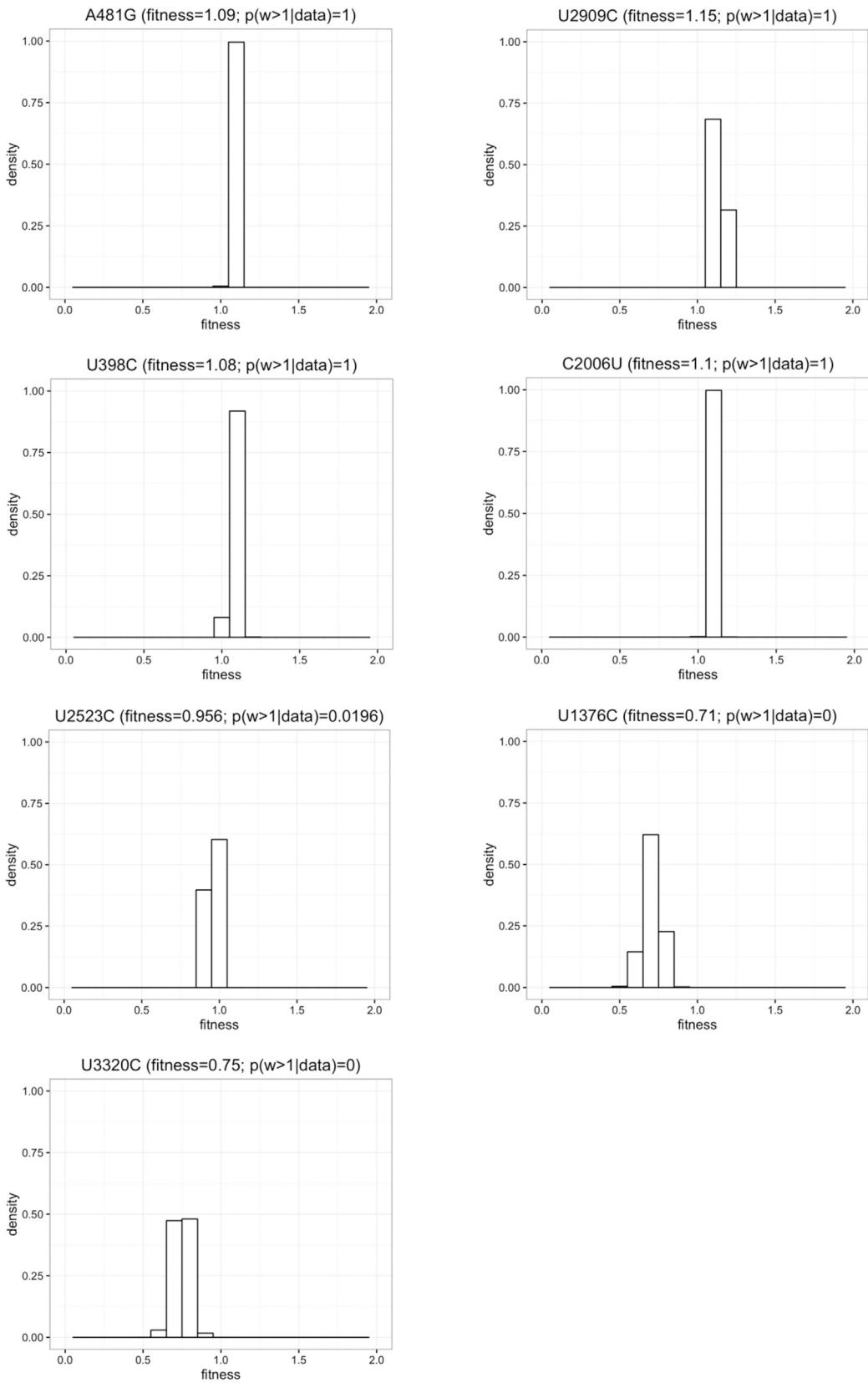
Figure S4. Results of ABC Approach Run on 100 Sets of Simulated Allele Frequencies Reveal a High Level of Accuracy for the Method, Related to STAR Methods

(A and B) Testing fitness inference. Simulations were run assuming a population size of $N = 10^5$ and a mutation (and back mutation) rate of 10^{-5} similar to the empirical data in this study. For each fitness value, 100 sets of allele frequencies were simulated based on a Wright-Fisher model with selection and with back-mutation (Stern et al., 2014) (STAR Methods). (A) Inferred fitness refers to the median of the posterior distribution inferred by the ABC approach. Boxplots show the 25th and 75th percentiles of inferred fitness values, and black lines represent the median value across all datasets. The results display high accuracy for most values of w ; for 82% of the datasets the difference between the simulated and inferred w was less than 0.2. Lethal alleles are slightly over-estimated, likely due to the strong effects of genetic drift that lead to low allele counts and confounding of low w values. On the other hand the fitness of slightly advantageous alleles, as found in the empirical data ($w = 1.1$), tends to be underestimated. (B) The inferred category across all simulated datasets, based on Bayesian p values: ADV refers to $P(w > 1 | data) > 0.95$, DEL refers to $P(w < 1 | data) > 0.95$, ?ADV refers to $0.95 > P(w > 1 | data) > 0.5$ and ?DEL refers to $0.95 > P(w < 1 | data) > 0.5$. Importantly, the results show that our false positive rate for inferring an advantageous allele at a Bayesian p value of 0.05 is very low ($2/1100 = 0.0018$).

(C) Inference of mutation rates. Frequencies of synonymous mutations across time for 39.5°C passages, shown at transition sites only. Mutations disrupting or forming CpG/UpA dinucleotides were removed. Generations 4 and 8 (corresponding to passages 2 and 4) were removed due to batch effects, particularly pronounced for C → U mutations. Linear regression was used to assess the mutation rates, the mean slope shown is the mean coefficient across all individual regression slopes (STAR Methods).

(D) Frequencies of lethal mutations across all passages, defined as mutations creating premature stop-codons or non-synonymous mutations at protein active sites (defined in (Acevedo et al., 2014)). Due to a paucity of data for U->C lethal mutations, we also added on mutations at highly conserved (100% identity) U sites across an alignment of forty-eight available wild-type poliovirus sequences of all three serotypes, and considered U->C mutations as lethal at those sites. Mutation rate estimates based on the synonymous sites regression (A) are marked with an “x.” Higher mutation frequencies of lethal mutations are likely due to complementation of deleterious genomes at high MOI occurring during the protocol. See STAR Methods for more details.

(E) Testing mutation rate inference. 100 datasets of 100 sites were simulated under two different population sizes ($N = 100,000$, $N = 10,000$), under two conditions (100% of sites were neutral versus 90% of neutral sites and 10% deleterious sites), and under two inference conditions (all times points were used for inference, only times points 10-14 were used for inference). Results show that the “true” value of mutation simulated ($1E-05$) is inferred accurately in the scenarios simulated. The largest variance is obtained when $N = 10,000$ and only limited time points are used.



(legend on next page)

Figure S5. Posterior Probability Distributions for Mutational *w* Fitness of All Seven Mutations Inferred as Adaptive by ParaSel, Related to Figure 4

The title of each figure shows the inferred fitness values (the median of the posterior distribution) as well as the Bayesian p values for positive selection ([STAR Methods](#)).

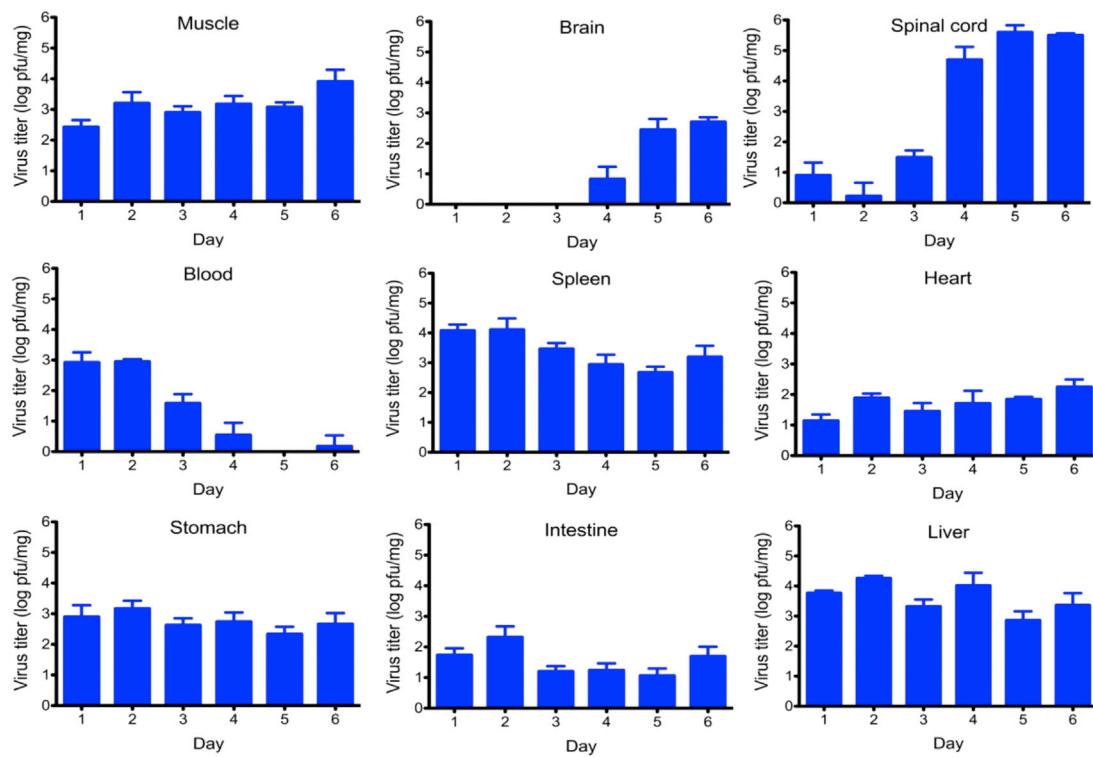
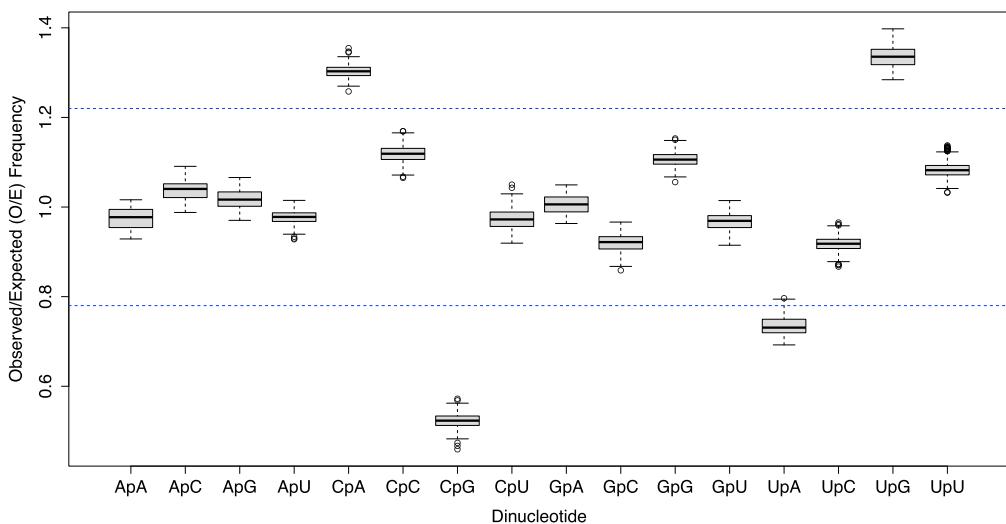


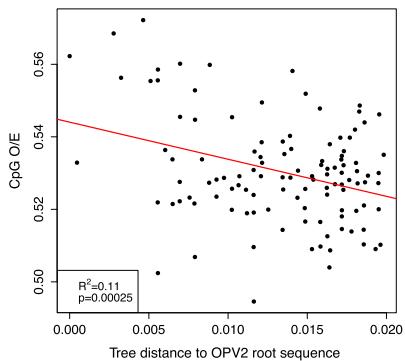
Figure S6. Comparison of OPV2 Titers in Tissues of PVR-Tg21-IFN-Knockout Mice, Related to Figure 5

Typical of PV, the virus spreads to the spinal cord and brain a few days post-infection. Virus titers were determined 7 days post-infection, and values represent the mean virus titer + standard deviation of eighteen mice.

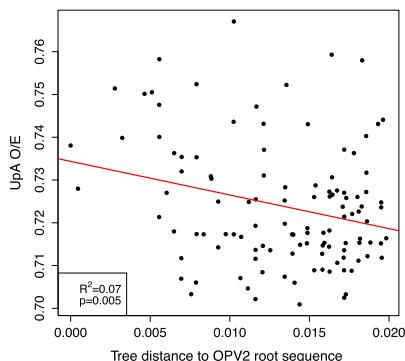
A Dinucleotide distributions in cVDPV sequences



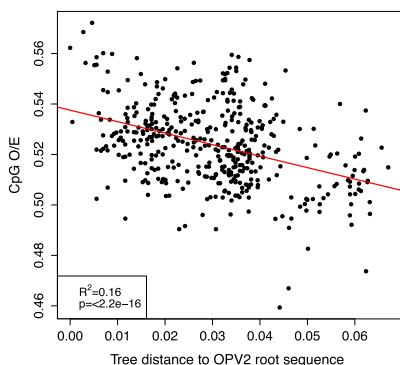
B CpG over first two years



C UpA over first two years



D CpG over six years



E UpA over six years

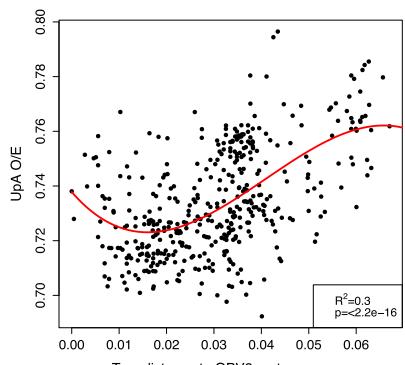


Figure S7. Related to Figure 6

(A) The odds ratio of observed/expected (O/E) frequencies of all dinucleotides in cVDPV. Dashed blue lines reflect significance cutoffs based on (Cheng et al., 2013). (B–E) The CpG and UpA dinucleotides O/E in the first two years (C and D, respectively) and over more than six years (D and E respectively), plotted against the distance of each sequence from the OPV2 root sequence. A linear regression model best fit the CpG data over both time periods whereas for the UpA sites a linear model best fit the first two years whereas a third degree polynomial best fit the entire time frame of over six years.