

# The Three Faces of Riboviral Spontaneous Mutation: Spectrum, Mode of Genome Replication, and Mutation Rate

Libertad García-Villada, John W. Drake\*

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, United States of America

## Abstract

Riboviruses (RNA viruses without DNA replication intermediates) are the most abundant pathogens infecting animals and plants. Only a few riboviral infections can be controlled with antiviral drugs, mainly because of the rapid appearance of resistance mutations. Little reliable information is available concerning i) kinds and relative frequencies of mutations (the mutational spectrum), ii) mode of genome replication and mutation accumulation, and iii) rates of spontaneous mutation. To illuminate these issues, we developed a model *in vivo* system based on phage Q $\beta$  infecting its natural host, *Escherichia coli*. The Q $\beta$  RT gene encoding the Read-Through protein was used as a mutation reporter. To reduce uncertainties in mutation frequencies due to selection, the experimental Q $\beta$  populations were established after a single cycle of infection and selection against RT<sup>-</sup> mutants during phage growth was ameliorated by plasmid-based RT complementation *in trans*. The dynamics of Q $\beta$  genome replication were confirmed to reflect the linear process of iterative copying (the stamping-machine mode). A total of 32 RT mutants were detected among 7,517 Q $\beta$  isolates. Sequencing analysis of 45 RT mutations revealed a spectrum dominated by 39 transitions, plus 4 transversions and 2 indels. A clear template-primer mismatch bias was observed: A•C>C•A>U•G>G•U> transversion mismatches. The average mutation rate per base replication was  $\approx 9.1 \times 10^{-6}$  for base substitutions and  $\approx 2.3 \times 10^{-7}$  for indels. The estimated mutation rate per genome replication,  $\mu_g$ , was  $\approx 0.04$  (or, per phage generation,  $\approx 0.08$ ), although secondary RT mutations arose during the growth of some RT mutants at a rate about 7-fold higher, signaling the possible impact of transitory bouts of hypermutation. These results are contrasted with those previously reported for other riboviruses to depict the current state of the art in riboviral mutagenesis.

**Citation:** García-Villada L, Drake JW (2012) The Three Faces of Riboviral Spontaneous Mutation: Spectrum, Mode of Genome Replication, and Mutation Rate. *PLoS Genet* 8(7): e1002832. doi:10.1371/journal.pgen.1002832

**Editor:** Diarmaid Hughes, Uppsala University, Sweden

**Received:** March 23, 2012; **Accepted:** May 31, 2012; **Published:** July 26, 2012

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** This research was supported by funds allocated to project number Z01ES065016 of the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: drake@niehs.nih.gov

## Introduction

Riboviruses (RNA viruses with no DNA replication intermediates) infect organisms from prokaryotes to higher eukaryotes and frequently cause deadly diseases. The mortality, morbidity, and economic burden of ribovirus-borne diseases strongly impact human society, especially in developing countries where neither sanitation nor treatment may be adequate [1]. Although extensive efforts have focused on developing countermeasures to prevent or treat riboviral diseases, only a few of these diseases can be effectively controlled by vaccination or antiviral drugs. In addition, control or eradication of riboviral diseases is soon balanced by the emergence of new riboviral pathogens or treatment-resistant strains of old ones (reviewed in [2]). Thus, we seek to understand which special features of these viruses contribute to their success. One key feature is their high mutation rate (reviewed in [3]).

Although the evolutionary forces driving high riboviral mutation rates remain unclear (reviewed in [1]), three mechanistic factors play important roles: the higher error-insertion rates of RNA replicases compared to DNA replicases, the lack of proofreading activity in RNA replicases, and the nonexistence of post-replicative RNA mismatch repair. The estimated mean rate per infection cycle is about 1.3 for several common single-stranded

RNA (ssRNA) human pathogens [4], roughly 0.1 for ssRNA tobacco viruses [5], and 0.03 for the double-stranded RNA (dsRNA) bacteriophage  $\phi 6$  [6]. Unfortunately, most of these estimates were based on studies in which small, potentially unrepresentative sequences were used as mutation reporters. In some cases, estimated rates in excess of 1 per infection cycle are probably incompatible with viability [7]. A further problem is the scarce information on the mode (linear, exponential, or mixed) by which riboviruses replicate their genomes within the host cell. Distinct modes of genome replication impact the pattern of intracellular mutation accumulation in the riboviral genome (and hence the mutation rate per infection cycle) differently. The only two empirical studies published to date on riboviral replication strategy, one conducted with the phage  $\phi 6$  [8] and the other with the ssRNA turnip mosaic virus [9], suggest that riboviruses replicate their genome mostly in a linear fashion, but further results are needed based on other riboviral systems.

In addition, there are limited data on the kinds and relative frequencies of spontaneous mutations (the mutation spectrum) in riboviruses, again a reflection of mutation reporters that do not sufficiently sample the genome. Only three spontaneous mutation spectra based on a cognate riboviral gene of adequate size are available and, unfortunately, none seems to be fully illustrative.

## Author Summary

Viral disease is a subject of major concern in public health. Diseases produced by riboviruses (RNA viruses *sensu stricto*) represent a special urgency, because these viruses display an exceptional capability to generate resistance mutations against antiviral drugs. Unfortunately, little is known about the rate and nature of spontaneous mutation in riboviruses. Thus, characterization of their mutation process may be helpful in the development of improved ways to counteract riboviral diseases. In this study, we investigated the mutation process *in vivo* of a model ribovirus, the bacteriophage Q $\beta$ , focusing on three key aspects: i) the kinds and relative frequencies of mutations, ii) the mode of genome replication, and iii) the rate of spontaneous mutation. Our results, combined with other information about riboviral mutagenesis, depict a ribovirus mutation spectrum largely dominated by transitions, a predominantly linear mode of genome replication, and a mutation rate per genome replication on the order of 0.04 for bacteriophages and plant viruses but perhaps an order of magnitude higher for mammalian riboviruses.

The tobacco mosaic virus rate and spectrum [10] were derived under conditions of multiple sequential infections. The tobacco etch potyvirus spectrum [11] probably contains a large fraction of mutations resulting from methodological manipulations rather than from virus replication errors. Finally, the phage  $\phi 6$  spectrum [6] was obtained from a mutation-accumulation experiment in the absence of gene complementation *in trans*, which tends to discriminate against strongly deleterious mutations.

A complete portrait of spontaneous mutagenesis in riboviruses is important not only for understanding their prevalence but also for improving ways to prevent and to treat riboviral diseases. For instance, accurate information on riboviral mutation kinds and rates may facilitate the creation of more stable attenuated vaccines [12]. Similarly, it seems likely that antiviral treatments based on mutagenic base analogs may prove to be more effective if the base analogs specifically increase the rate of those errors that riboviral replicases already generate most frequently. Although pathway-directed mutagenesis is unlikely to prevent the appearance of riboviral resistance to specific base analogs, it may enlighten the development of more efficient combinatory therapies [13] and at least slow disease progression, thus enhancing the immune response.

The main aims of the present study were to characterize the mutation spectrum, to determine the mode of genome replication, and to estimate the spontaneous mutation rate of a ribovirus using the bacteriophage Q $\beta$  as an experimental model. Q $\beta$  has been well characterized physiologically [e.g., 14–17], physiochemically [e.g., 18–21], structurally [e.g., 22–27], and molecularly [e.g., 28–32]. It is a linear (+)-strand ssRNA phage whose natural host is *Escherichia coli*, although it can also propagate in other gram-negative bacteria with an F pilus. Its 4217-nt long genome is organized in three cistrons that encode (from 5' to 3') the A2 or Maturation protein, which mediates both the binding of Q $\beta$  to the host and post-replicative host lysis; the Coat protein and its elongated A1 or Read-Through (RT) protein, which is required for Q $\beta$  capsid assembly and for host infection; and the catalytic  $\beta$  subunit of the Q $\beta$  replicase. (RT is translated when a ribosome incorporates tryptophan at the natural UGA stop codon of the Coat-coding gene at a frequency of  $\approx 3\%$  [33].) Q $\beta$ 's life cycle may be summarized as follows: i) the phage attaches to the F pilus of *E. coli*

and the genome enters the cytoplasm; ii) cellular components translate the  $\beta$  subunit of the phage replicase, which then polymerizes with four host subunits (the ribosomal protein S1, the translation elongation factors EF-Tu and EF-Ts, and the host factor HF) and binds the Q $\beta$  genome; iii) the  $\beta$  subunit copies the (+)-strand genome to produce a (–)-strand RNA that in turn is used as template to produce more (+) strands; iv) (+) strands serve as templates for the production of the phage proteins; v) 40–60 minutes after infection, by which time the host cell is filled with phage particles, partially assembled virions, and phage-specific side products, the cell lyses, releasing  $(10\text{--}40)\times 10^3$  particles of which only 10–50% are infectious (reviewed in [14,34]).

Here, we used the gene encoding the RT protein (excluding the portion that encodes the Coat protein) as an *in vivo* mutation reporter. Selection against RT<sup>–</sup> mutants was ameliorated by using a complementing system *in trans* based on a plasmid that encodes the entire Coat/RT mRNA with the natural UGA stop codon replaced with a TGG tryptophan codon [35]. To further reduce the effect of selection, the experimental Q $\beta$  populations were established after a single cycle of infection. We assessed the Q $\beta$  genome mutation rate ( $\mu_g$ ) in three different ways: i) a forward-mutation test in which mutants carrying phenotypically detectable RT mutations were isolated and sequenced and  $\mu_g$  was estimated from the frequency of observed nonsense mutations and indels; ii) single-burst reversion tests in which two different RT<sup>–</sup> mutants were employed (one carrying a single-base substitution and the other a four-base insertion) and  $\mu_g$  was estimated from the corresponding reversion rates; and iii) a phenotype-blind forward-mutation test in which some first-generation progeny of the RT mutants detected by the first method were isolated and sequenced and  $\mu_g$  was estimated from the frequency of all secondary RT mutations generated *de novo*. The distributions of RT<sup>+</sup> revertants observed in the reversion tests were used to infer the mode in which Q $\beta$  replicates its genome, and the spontaneous Q $\beta$  mutation spectrum was obtained from the RT mutations collected through the forward-mutation tests.

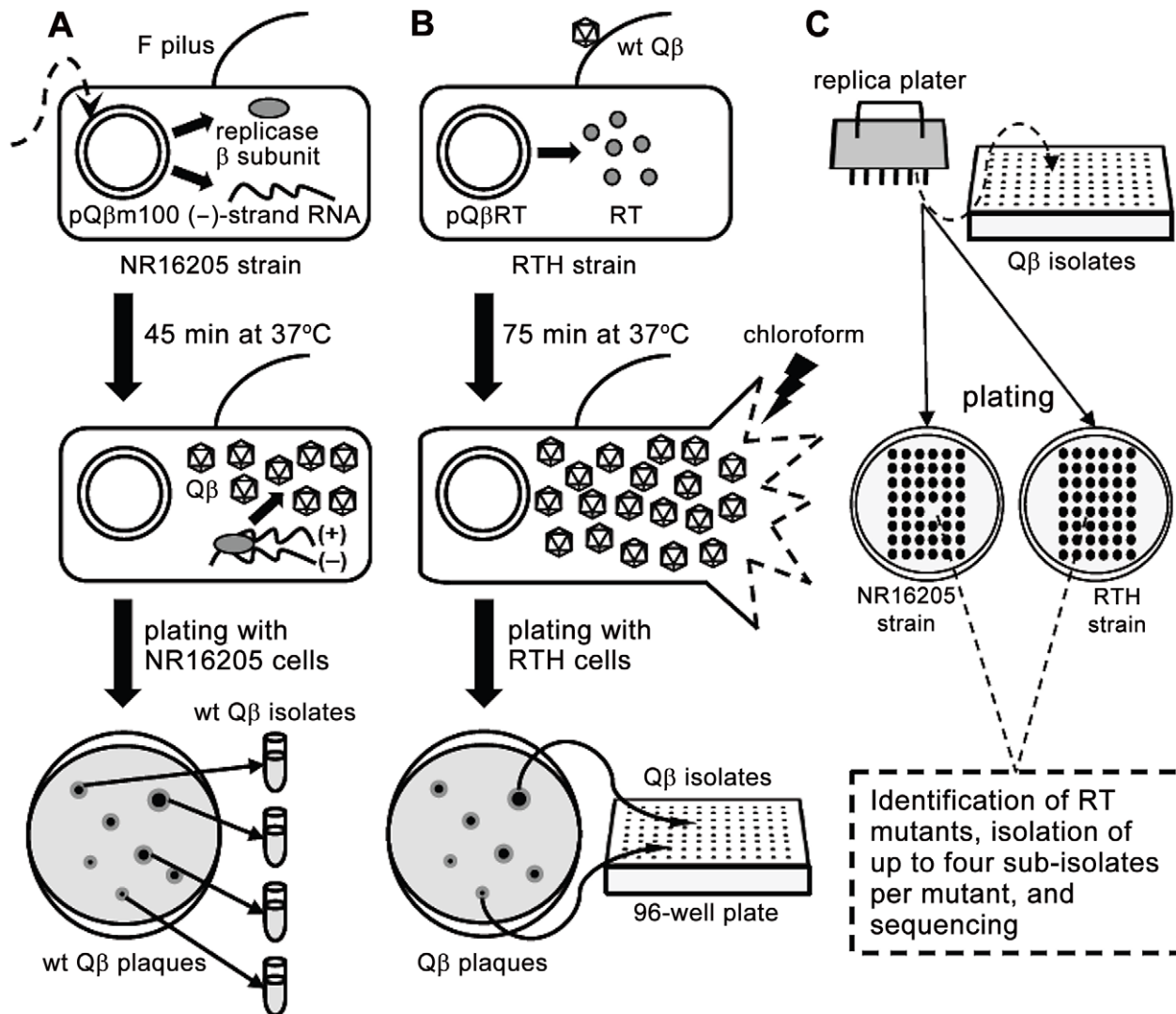
## Results

### Description of the System

The basics of the experimental system and the strains used in this study are described in Figure 1 and Table 1, respectively.

### The Mutation Spectrum

Mutations arising in a mutation-reporter (target) sequence can be of two types. “Detectable” mutations are those that display the mutant phenotype when present as a single mutation. “Undetectable” mutations lack the mutant phenotype when present as a single mutation but may nevertheless be observed when they arise in the presence of a detectable mutation, in which case they are sometimes called “hitchhiker” mutations (and their detectable partner may be called a “driver” mutation). Sometimes, especially with mutants with equivocal phenotypes, no mutation is found in the target, reflecting either some imperfection in the screening method or a mutation elsewhere in the genome whose effect mimics that of the reference mutation; such isolates are thereafter included in the non-mutant total. Another distinction is often relevant: some mutations produce a fully mutant phenotype but others produce an intermediate phenotype (and are therefore often called “leaky” mutations or are said to produce a “weak” mutant phenotype). In this study, yet another dimension is added. Each Q $\beta$  mutant originally isolated as requiring a helper host to generate a plaque or each of a number of non-mutant control plaques was re-plated and up to four next-generation plaques were



**Figure 1. Schematic representation of the experimental system used to isolate *RT* mutants.** A. *E. coli* NR16205 cells were transformed with pQ $\beta$ m100, which constitutively expresses low levels of Q $\beta$  (-) strands. Some  $\beta$  subunits of the Q $\beta$  replicase are also transcribed from pQ $\beta$ m100 and, once translated, they replicate the (-) strands yielding (+) strands, which are infectious (*i.e.*, they produce Q $\beta$  virions). Transformants were plated with NR16205 cells and incubated overnight to generate wt Q $\beta$  plaques, which were individually isolated. B. *E. coli* RTH cells, which carry pQ $\beta$ RT that expresses RT constitutively, were infected with wt Q $\beta$ . After one cycle of infection (75 min), chloroform was added to the experimental cultures to prevent further infections. Samples of the resulting lysates were plated with RTH cells, rendering Q $\beta$  plaques that were independently harvested in 96-well plates. The use of RTH cells during these steps assures that any *RT*<sup>-</sup> mutant generated during the replication of wt Q $\beta$  would be as able to produce a plaque as its wild-type siblings. C. Using a replica plater, the isolated Q $\beta$  plaques were tested for the *RT*<sup>-</sup> phenotype (impaired growth on non-complementing NR16205 cells but normal growth on RTH cells). Putative *RT* mutants were identified and sequenced, after which up to four sub-isolates per verified *RT* mutant were isolated and sequenced.  
doi:10.1371/journal.pgen.1002832.g001

harvested and sequenced. When all members of such a family contain the same mutation, we call it the “primary” mutation, and if some of the next-generation plaques contain additional mutations, we call them “secondary” mutations, which may arise when mutation rates are sufficiently high.

One-step growth curves of wild-type (wt) Q $\beta$  in RT-helper (RTH) cells, which complement *RT*<sup>-</sup> mutations, indicated that Q $\beta$  requires  $\approx$ 75 min to lyse an infected RTH cell (Figure 2). Thus, to limit the number of infection cycles to one before seeking *RT* mutants, RTH lysates were generated by adding chloroform 75 min after infection with wt Q $\beta$ . Samples of these lysates were plated with RTH cells and the resulting plaques were harvested and tested for the *RT*<sup>-</sup> phenotype (impaired growth on non-complementing NR16205 cells but normal growth on RTH cells).

Among 7517 plaques tested in four independent experiments, 47 candidate *RT* mutants were recovered and sequenced. Of these, 30 contained at least one primary *RT* mutation (Table 2). (The 17 candidates with no primary *RT* mutation may have carried an *RT*<sup>-</sup>-mimicking mutation elsewhere in the genome or, because Q $\beta$  grows better on RTH cells than on NR16205 cells, might have carried weak non-*RT* mutations and showed enhanced growth on RTH cells.) Most of the primary mutations were missense but two (one in mutant RT23 and one in RT37) were indels consisting of single-base additions. Two mutants each carried two primary mutations; in RT18, both were missense; in RT41, one was missense and the other was a synonym. Three mutants (RT10, RT40 and RT46) each carried a nonsense mutation that generated a stop codon; RT40 is a special case because it

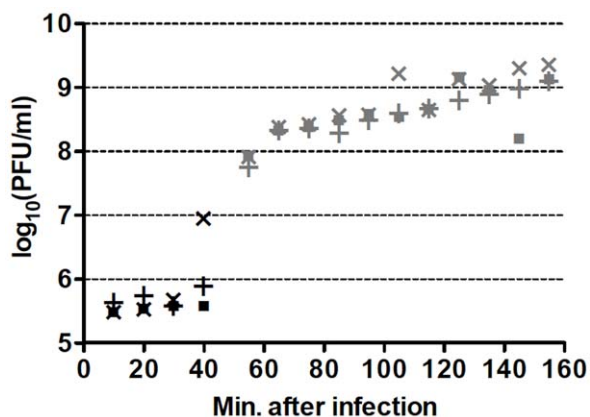
**Table 1. Strains.**

System	Designation	Relevant properties	Source
Plasmids	pQ $\beta$ m100	Expresses wt Q $\beta$	DM
	pQ $\beta$ m100(RT2157/fs)	Expresses RT <sub>IN</sub>	DM
	pQ $\beta$ BRT	Expresses RT and TMP <sup>r</sup>	DM
<i>E. coli</i>	NR10836	<i>ara thi <math>\Delta</math>prolac F'prolac</i>	RS
	NR16205	NR10836 Tet <sup>r</sup>	RS
	RTH	NR10836+pQ $\beta$ BRT	LGV
Q $\beta$	wt		pQ $\beta$ m100
	RT <sub>IN</sub>	+UUAA at 2157	pQ $\beta$ m100
	RT <sub>SUB</sub>	G1990A	LGV

Sources: DM = Donald Mills (SUNY Downstate Medical Center, Brooklyn, NY); RS = Roel Schaaper (National Institute of Environmental Health Sciences) [65]; LGV = this report.

doi:10.1371/journal.pgen.1002832.t001

converted the leaky UGA codon that terminates the Q $\beta$  Coat protein to a far less leaky UAA stop codon. In two cases, the primary mutation displayed at most a very weak phenotype upon re-plating: the primary mutation of RT27 was a synonym and that of RT33 was missense. These mutants are included in Table 2 and dependent calculations because their mutations could in principle produce a deleterious effect but would have no significant impact if disregarded. The 13 RT secondary mutations (Table 2) presumably arose sufficiently early during the growth of the screened plaques on RTH lawns to be detected among the next-generation progeny. They include 6 missense mutations and 7 synonyms, a ratio that deviates from the approximately 3.3:1 ratio expected from the set of RT codons. Applying the binomial distribution, finding 6 missense among a total of 13 mutations has  $P=0.014$  and finding  $\leq 6$  has  $P=0.018$ . This result presumably signals selection against RT mutations with strong effect during plaque growth on RTH lawns, consistent with the smaller burst sizes of RT<sub>IN</sub> (a mutant Q $\beta$  strain carrying a four-bases insertion in RT; see Table 1) than wt Q $\beta$  in helper cells. (The average burst sizes of RT<sub>IN</sub> and wt Q $\beta$  are 328 and 847, respectively, estimated from three different one-step curves per phage type.)



**Figure 2. Representative one-step growth curves of wt Q $\beta$  in RTH cells.** Individual cultures are represented by different symbols while black and gray colors represent two different culture dilutions. PFU = plaque-forming units. Similar curves were obtained for each of the four wt Q $\beta$  isolates used to generate one-step growth lysates. doi:10.1371/journal.pgen.1002832.g002

Table 3 lists the kinds of mutations in the entire set of 45. The 2 indels are strikingly less frequent than the 43 single-base substitutions. The general expectation that frameshifting indels generate a detectable mutant phenotype when arising in a protein-coding sequence reduces the chances of having missed other indels during the scoring of mutants. In addition, pQ $\beta$ BRT, the RT-expressing plasmid used in this study, can complement RT deletions comprising up to 447 nt [36], which reduces the probability of having missed indels  $>1$ -nt long. The 39 transitions were almost 10-fold more frequent than the 4 transversions; if both transversions and transitions were to arise at equal frequencies among base-substitution pathways, the expected ratio would be 1:0.5 (each site being able to generate two kinds of transversions and one kind of transition), a 20-fold difference from the observed ratio. Transitions, when ranked in decreasing order of observed numbers, were U $\rightarrow$ C (16) $>$ G $\rightarrow$ A (10) $>$ A $\rightarrow$ G (8) $>$ C $\rightarrow$ U (5). The numbers of the four bases in the target decrease in the same order, U(175) $>$ G(147) $>$ C(139) $>$ A(130), but this trend cannot quantitatively explain the normalized frequencies of mutated bases, which is 0.091 $>$ 0.068 $>$ 0.058 $>$ 0.038. Thus, the intrinsic mutability of the four bases, presumably reflecting the error propensities of the Q $\beta$  replicase, is likely to be the main determinant of the relative frequencies of observed mutations.

The mutations were widely distributed over the target (Figure 3). Because only 4 RT positions out of the observed 38 hosted more than one substitution, the spectrum is clearly far from saturation. Both indels arose within short homopolymeric runs, a common pattern in mutation spectra that presumably reflects misaligned primer-templates [37,38]. The substitutions showed no correlation with their nearest neighbors either individually or as purines *versus* pyrimidines (analyses not shown). However, because a tendency towards enhanced mutability of any base within a G/C-rich sequence has been observed in both *E. coli* [39,40] and the T-even coliphage RB69 [41], we also examined the base composition (G+C versus A+T) of the local sequence environments where substitutions were observed. G•C base pairs are more stable than A•U pairs, so that G/C-rich sequences might help to stabilize secondary structures containing hairpin loops, where unpaired bases may be more sensitive to oxidative damage. In addition, duplexes richer in G•C pairs may be slower to unwind, which might render replication more error-prone in currently unknown but perhaps generally applicable ways. A recent description of the structure of the Q $\beta$  replicase [31] suggested that the replicating Q $\beta$  genome (template+complement) forms a 6–7 base-pair duplex in the internal cavity of the replicase before both the single-stranded product and template exit the enzyme. Accordingly, we analyzed the base composition of the sequences six and seven bases upstream of the observed substitutions (Figure 4). Both the 6-mers and the 7-mers contain more (G+C) than expected from the target content of bases. The difference for the 6-mers has  $P=0.059$  and for the 7-mers has  $P=0.034$  (replicated G-test for goodness-of-fit,  $P$ -values for “pooled G”,  $G_p$ , 1 df). Nevertheless, these small differences, combined with the homogeneity in base composition of the analyzed sequences, made the “total G” ( $G_T$ ) non-significant in both analysis ( $G_T=0.409$ , 6 df, and  $G_T=0.420$ , 7 df, for 6- and 7-mers, respectively). Overall, a larger sample of mutations would probably indicate more clearly the existence (or absence) of any effect of the G/C content of the local sequence on Q $\beta$ -replicase error tendencies.

### The Mode of Genome Replication

To determine how mutations accumulate in the Q $\beta$  (+)-strand progeny during replication and thus to estimate the rate of spontaneous mutation per genome replication in Q $\beta$ , it is

**Table 2.** Spontaneous RT mutants and mutations.

Mutant	Mutation(s)	Position	Impact
RT1	G→U	383	Gly→Val
	(A→G)	(13)	(Ile→Val)
RT3	A→G	449	Asp→Gly
RT4	U→C	122	Leu→Pro
	(G→A)	(473)	(Gly→Asp)
RT5	G→A	383	Gly→Asp
	(U→G)	(283)	(Cys→Gly)
RT6	U→C	260	Leu→Pro
RT7	G→U	526	Asp→Tyr
RT10	G→A	249	Trp→Stop
RT14	U→C	388	Phe→Leu
	(U→C)	(411)	(Ile→Ile)
RT16	G→A	430	Ala→Thr
RT18	C→U	119	Ser→Phe
	A→G	527	Asp→Gly
RT20	U→C	460	Tyr→His
RT21	A→G	170	Tyr→Cys
RT23	+T	221–224	Frameshift
	(A→U)	(85)	(Thr→Ser)
RT25	A→G	449	Asp→Gly
RT26	U→C	11	Leu→Pro
RT27	U→C	261	Leu→Leu
	(A→G)	(478)	(Ile→Val)
RT29	U→C	389	Phe→Ser
RT31	G→A	367	Glu→Lys
	(U→C)	(21)	(Gly→Gly)
RT32	U→C	479	Ile→Thr
	(C→U)	(105)	(Pro→Pro)
RT33	G→A	382	Gly→Ser
	(C→U)	(442)	(Leu→Phe)
RT36	U→C	439	Ser→Pro
RT37	+T	106–107	Frameshift
RT40	G→A	2	Stop→Stop
	(C→U)	(213)	(Leu→Leu)
	(A→G)	(216)	(Lys→Lys)
RT41	U→C	24	Gly→Gly
	G→A	565	Ala→Thr
RT42	U→C	68	Ile→Thr
RT43	G→A	430	Ala→Thr
RT44	U→C	283	Cys→Arg
RT45	U→C	308	Leu→Pro
RT46	G→A	248	Trp→Stop
RT47	A→G	449	Asp→Gly

The 30 mutants were recovered from 7517 screened plaques and include two isolates (RT27 and RT33) in which the primary mutation had no RT<sup>-</sup> phenotype, although their secondary mutation (or the combination of both) did. None of the other 17 candidates in the series RT1-RT47 displayed primary mutation, but two isolates, RT24 and RT35, had one secondary mutation each (C→U at position 18 and U→C at position 294, respectively). Altogether, 148 mutants were sequenced: [28 RT mutants carrying primary detectable mutations]+[4 next-generation isolates per each of these]+[2 next-generation isolates per each of four RT mutants (RT24, RT27, RT33 and RT35) that lacked detectable primary

mutations]. The mutants were obtained in four different screens yielding 8, 5, 6 and 11 mutants in the order shown. Mutations in parentheses are secondary mutations. Mutant positions refer to the target sequence (1 through 591 including the UAG stop codon) that corresponds to Q $\beta$  genomic positions 1743 through 2333.

doi:10.1371/journal.pgen.1002832.t002

necessary to know the mode by which Q $\beta$  produces its progeny during cell infection. Two distinct modes are possible. One is linear, wherein the infecting (+)-strand genome is used repeatedly as a template and then at least some of the resulting (-)-strand RNAs are each used repeatedly as templates; consequently, at the end of the infection cycle, each of the many (+)-strand progeny has experienced only two replications, from (+) to (-) and from (-) to (+). In this model, due to the many (+)-strand progeny contributed by the fewer (-)-strand templates, most replication errors will produce a single mutant during the second round of replication and only a small fraction of errors will generate a clone of mutants when a replication error occurs in the first round of replication and is further repeatedly copied in the second round [4]. The other mode is classical exponential replication, in which case the numbers of mutants recovered from single viral bursts display an exponential distribution [42]. Intermediate models combining linear and exponential replication in different proportions are also conceivable.

To determine which model best fits the distribution of mutants in Q $\beta$ , two separate single-burst reversion tests were conducted (see Materials and Methods), one using the mutant RT<sub>IN</sub> and the other using RT<sub>SUB</sub> (described in Table 1). The tests involved plating cultures, containing bursts from infected RTH cells, onto NR16205 cells, aiming to deliver roughly 1 revertant-yielding burst on each of many plates. With RT<sub>IN</sub>, among the 250 cultures plated, six did not form plaques well and were discarded, and five had the following number of Q $\beta$  plaques: 1254, 585, 345, 342 and 105. Because the plating efficiency of wt Q $\beta$  with NR16205 is

**Table 3.** Sequence changes in RT.

Mutation type	No.	Mispair <sup>a</sup>
Indels		
-1	2	
Transitions		
A→G	8	<u>U</u> •G
G→A	10	<u>C</u> •A
U→C	16	<u>A</u> •C
C→U	5	<u>G</u> •U
Transversions		
A→U	1	<u>U</u> •U
A→C	0	<u>U</u> •C
G→U	2	<u>C</u> •U
G→C	0	<u>C</u> •C
U→A	0	<u>A</u> •A
U→G	1	<u>A</u> •G
C→A	0	<u>G</u> •A
C→G	0	<u>G</u> •G
Total	45	

<sup>a</sup>Assuming mispairing at the second round of genome replication; the template base is underlined.

doi:10.1371/journal.pgen.1002832.t003



have occurred along the available single-base-substitution pathways (up to 8 for a UAG stop codon, depending on the functional competence of the encoded amino acids), each of which differs between replications and which might therefore have displayed large rate asymmetries, which can easily exceed 100-fold in the case of DNA genomes [e.g., 43]. The second cause is that the total number of copying events probably differ between (–)-strand and (+)-strand synthesis during cell infection; in Q $\beta$ , for instance, the number of accumulated (+) strands was estimated to be about 10 times greater than the number of (–) strands [14,44], so that revertant bursts of size 1 from the (+)-strand synthesis would then be more frequent than the larger bursts from the first rounds of replication. Notably, however, these larger bursts, once they appear, are expected to exhibit variable sizes that depend on, among other factors, the growth conditions [45], and that might therefore impact the observed distribution.

### The Rate of Spontaneous Mutation

To estimate the rate of spontaneous mutation per genome replication ( $\mu_g$ ) for a ribovirus, it is necessary to know (i) the mutation frequency  $f$ , (ii) the number of infection cycles  $c$  that elapse between the initial infection and the scoring of mutants, (iii) the average number of times  $n$  that each genome is replicated per infection cycle, (iv) the number of detectably mutable bases in the mutational target ( $T$ ), and (v) the genome size ( $G$ ). In the present case,  $G=4217$  nt,  $c=1$ , and, from our results,  $n\approx 2$ . Although  $T=591$  RT bases for estimating the indel mutation rate ( $\mu_I$ ), that number cannot be used when estimating the corresponding base-

substitution rate ( $\mu_{SUB}$ ) because, while nearly all indels are detectable, many substitutions fail to produce a mutant phenotype. Instead,  $\mu_{SUB}$  may be estimated from the number of substitutions that generate a stop codon (nonsense mutation) because, like indels, nonsense mutations are generally detectable. When considering nonsense mutations,  $T$  equals one-third of the number of paths in the mutational target that may generate a stop codon (one-third because each base can mutate by three different paths) [46]. In this study, 3 nonsense mutations were found among 7517 Q $\beta$  isolates, and  $T=66$  paths leading to a stop codon in the RT target. Thus,  $f_{path}=3/(7517)(66)=6.047\times 10^{-6}$ ,  $f_{SUB}=3f_{path}=1.814\times 10^{-5}$  per base, and  $\mu_{SUB}=f_{SUB}/cn=9.0704\times 10^{-6}$ . Because 2 indels were found,  $f_I=2/(7517)(591)=4.502\times 10^{-7}$ , and  $\mu_I=f_I/cn=2.251\times 10^{-7}$ . Hence,  $\mu_g=(\mu_I+\mu_{SUB})G=0.039$ .

In addition to the primary mutations detected by their phenotypes, some hitchhiking mutations were found. These secondary mutations may be used for an independent estimate of  $\mu_g$ , in which case  $T=591$  bases. A total of 9 secondary mutations (all base substitutions) were detected among 112 sequenced sub-isolates. (The remaining secondary mutations from the 13 described in Table 2 were observed in RT<sup>–</sup> isolates lacking any detectable primary RT mutation and thus were excluded from these calculations.) Thus,  $f_{SUB}=9/(112)(591)=1.36\times 10^{-4}$ ,  $\mu_{SUB}=f_{SUB}/cn=6.80\times 10^{-5}$ , and  $\mu_{SUBg}=\mu_{SUB}G=0.287$ . This value is greater than the corresponding value from the nonsense-mutation method by 7.4-fold and may, as discussed later, signal the impact of transient hypermutation.

Mutation rates can also be estimated for the reversion of the mutants RT<sub>IN</sub> and RT<sub>SUB</sub> using the results of the single-burst reversion tests. First, some definitions are needed: the number of cultures =  $C$ ; the average number of infected cells per tube =  $N$ ; the average burst size =  $B$ ; the number of initial [(+)-strand→(–)-strand] copies =  $c_1$  with an error rate  $\mu_1$  per copy; the number of succeeding [(–)-strand→(+)-strand] copies =  $c_2$  with an error rate  $\mu_2$  per copy and a burst size  $B=c_2$  that ignores unpackaged genomes; and there are  $n=2$  two rounds of replication per infection. Then the average total number of mutational events per infected cell will be  $c_1\mu_1+c_2\mu_2$ ; however, these components cannot be disentangled with our data, so we will assume that  $c_2\mu_2\gg c_1\mu_1$  (e.g., most of the mutations are generated in the second replication, as indicated by the results from the single-burst reversion tests), in which case the average total number of mutations per infected cell will be  $c_2\mu_2=B\mu_2$ .

For a set of cultures of which some contain 0 mutants, the fraction of null tubes is  $e^{-m}$  where  $m$  is the average number of mutational events per culture [47]. The total number of replication events per culture  $\approx NB$ , whence  $\mu_2\approx\mu=m/NB$ . For RT<sub>IN</sub>, the fraction of null tubes was 31/239,  $m=2.04$ ,  $N\approx 4560$  infected cells per tube, and  $B=328\pm 93$  (mean  $\pm$  SD,  $n=3$  one-step curves), so that  $\mu(\text{RT}_{IN})=(1.37\pm 0.39)\times 10^{-6}$ . For RT<sub>SUB</sub>, the fraction of null tubes was 238/498,  $m=0.738$ ,  $N\approx 35$ ,  $B=859\pm 165$  (mean  $\pm$  SD,  $n=3$  one-step curves), and  $\mu(\text{RT}_{SUB})=(2.46\pm 0.47)\times 10^{-5}$ . The ratio  $\mu(\text{RT}_{SUB}):\mu(\text{RT}_{IN})\approx 18$  which, given the indel sample size of 2, agrees well with the corresponding ratio of the two kinds of mutations (substitutions and indels) in the spectrum (43:2 $\approx$ 22) or by rate ( $9.07\times 10^{-6}):(2.25\times 10^{-7})\approx 40$ ).

Another way to estimate these reversion rates is to use  $\mu=f/2$  but, as directly above, to assume that all detected mutations arose in the second round of replication, those arising in the first round being too infrequent to be readily observed, in which special case,  $\mu=f$  as above. Here,  $f$  is simply the sum of all observed RT<sup>+</sup> revertants divided by all the Q $\beta$  progeny in all tubes,  $NBC$ . For RT<sub>IN</sub>, the total number of revertants was 510 (Table 4), so that  $\mu(\text{RT}_{IN})=510/$

**Table 4.** Observed and expected distributions of RT<sup>+</sup> revertants in RT<sub>IN</sub> and RT<sub>SUB</sub> bursts.

RT <sup>+</sup> per plate	RT <sub>IN</sub> <sup>a,b</sup>		RT <sub>SUB</sub> <sup>a,b</sup>	
	Observed	Expected <sup>c</sup>	Observed	Expected <sup>c</sup>
0	31	31.00	238	238.01
1	63	63.32	150	175.72
2	55	64.66	66	64.87
3	49	44.02	27	15.96
4	24	22.48	10	2.95
5	11	9.18	2	0.44
6	4	3.13	3	0.05
7	1	0.91	1	0.1
8	1	0.23	1	0
9	0	0.05	0	0
10	0	0.01	0	0
Total plates	239	239	498	498
Total plaques	510	488.13	446	367.67
$P^c$	0.825		<0.001	

<sup>a</sup>Sets of 10 and 5 independent experiments were conducted with RT<sub>SUB</sub> and RT<sub>IN</sub>, respectively. Because no significant differences were found in the observed distributions of revertants among the experiments of each set (G-test of independence,  $P=0.3414$  and  $P=0.1103$ , with 16 and 27 df for the set of experiments conducted with RT<sub>IN</sub> and RT<sub>SUB</sub>, respectively), the results of all the experiments comprising each set were pooled.

<sup>b</sup>Expected distributions of revertants were inferred using the  $p_0$  method, based on the fraction of plates displaying no revertants [47].

<sup>c</sup> $P$  values represent the probability that the observed distributions fit the Poisson distributions as determined by G-tests for goodness-of-fit. See text for details.

doi:10.1371/journal.pgen.1002832.t004

$(4560)(328 \pm 93)(239) = (1.43 \pm 0.40) \times 10^{-6}$ , a value close to the null-class value because of the excellent agreement between the observed distribution and the Poisson expectations (Table 4). For RT<sub>SUB</sub>, the total number of revertants was 446 (Table 4), so that  $\mu(\text{RT}_{\text{SUB}}) = 446/(35)(859 \pm 165)(498) = (2.98 \pm 0.58) \times 10^{-5}$ , again a value close to the null-class value but slightly higher due to the occurrence of a small excess of plates with larger numbers of revertants compared to the expectations of the Poisson distribution (Table 4). Because the number of paths in which the RT<sub>SUB</sub> mutated codon (UAG) may change producing an RT<sup>+</sup> revertant is not known, the estimated  $\mu(\text{RT}_{\text{SUB}})$  is an upper limit corresponding to 8 paths or  $2^{2/3}$  substitutions.

## Discussion

### The Mutation Spectrum

We have obtained a spontaneous mutation spectrum for the RNA coliphage Q $\beta$  using a cognate mutational target, the RT-coding gene minus the portion encoding the Coat protein. This 591-nt target generously samples the 4217-nt Q $\beta$  genome, and the RT and genome base compositions are indistinguishable (G-test of independence,  $P = 0.9719$ , 3 df). The spectrum, based on 45 single-base changes, is a mixture of 32 primary mutations plus 11 secondary mutations found hitchhiking on some primary mutations, plus 2 single synonymous mutations (at target sites 18 and 294) arising during sequencing that showed no primary mutation. This spectrum has three defining characteristics. One is its strikingly low frequency of indels, only 2 among 30 RT mutants and 45 mutations, thus representing only about 4% of the total mutations, while in spectra from several DNA-based microbes (phages  $\lambda$  and T4, *E. coli*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*), indels comprise about 40% of the mutations (average 41%, range = 27–59% [39,46,48–51]). Another characteristic is its unusually high transition:transversion ratio (39:4 = 9.75) compared to a random expectation of 1:2 = 0.5. This transition bias contrasts with the transition:transversion ratios observed for the same DNA-based microbes mentioned above (mean 0.87, range 0.08–1.67). Finally, normalized to target-base frequencies, the spectrum reveals a biased mutation tendency consisting of U $\rightarrow$ C>G $\rightarrow$ A>A $\rightarrow$ G>C $\rightarrow$ U> all transversions. Taking into account the dynamics of Q $\beta$  genome replication with most mutations arising during the second round of replication, this mutation bias reflects a mismatch formation/extension bias in the template $\rightarrow$ progeny sense of A $\rightarrow$ C>C $\rightarrow$ A>U $\rightarrow$ G>G $\rightarrow$ U> transversions mismatches. This bias does not seem to reflect either cytosine deamination (which promotes C $\rightarrow$ U) or guanine oxidation (which promotes G $\rightarrow$ U), but rather the insertion of ionized, tautomerized, wobbled or *syn*-conformation bases.

Several other spectra of spontaneous riboviral mutations have been described previously.

(i) The first, using (+)-strand tobacco mosaic virus (TMV) and a target-complementation system *in trans*, reported a notable preponderance (24:11) of indels over substitutions, similar numbers of transversions (6) and transitions (5), and a remarkably high frequency (9/23) of mutants with multiple mutations (“multiples”) [10]. These multiples may have arisen either because of transient hypermutability (which may also have been observed with Q $\beta$ , as described below) or because the mutants were recovered after multiple sequential cell infections, estimated at  $c \approx 6$ . In this study, only mutants carrying lethal mutations were recovered. Thus, if many of the TMV base-substitutions were leaky, then, given  $c \approx 6$ , the high frequency of indels

might have reflected their fully mutant phenotypes in contrast to the often leaky substitutions; however, although the TMV target sequence contained 115 paths to stop codons, no nonsense mutations were recovered in the small sample of substitutions. Perhaps some other factor impinged on this system, such as photodynamic mutagenesis.

- (ii) Another spectrum, obtained in a mutation-accumulation experiment using almost the entire genome as a mutational target for the dsRNA phage  $\phi 6$ , also exhibited a strong transition:transversion bias (46:5 = 9.2) and only one indel [6], although selection may have reduced the recovery of indels. In this phage, the parental genome can be written as [+/-]. Upon infection, the (-) strand is copied repeatedly into (+) strands, some of which are translated and others of which are encapsulated. Within the nascent viroid, the (+) strand is copied once to produce a [+/-] progeny particle. In this pathway, a mutation such as U $\rightarrow$ C can arise by an A $\rightarrow$ C mismatch in the first round of replication, or by a U $\rightarrow$ G mismatch in the second round. Both paths would produce clones of size 1 in a single-burst reversion test, the predominantly observed result [8], but the culpable mismatch remains unknown. The observed mutation bias was A $\rightarrow$ G(14)  $\approx$  C $\rightarrow$ U(13)  $\geq$  G $\rightarrow$ A(10)  $\approx$  U $\rightarrow$ C(9) > all transversions, or, when normalized to the numbers of target bases, A $\rightarrow$ G(0.0052) > C $\rightarrow$ U(0.0034)  $\geq$  G $\rightarrow$ A(0.0027)  $\approx$  U $\rightarrow$ C(0.0028). While this is approximately the opposite of the Q $\beta$  bias, it could turn out to be very similar in terms of preferred mismatches depending on which mismatches generated the  $\phi 6$  mutations.
- (iii) A (+)-strand tobacco etch potyvirus spectrum [11], obtained by means of a target-complementation system *in trans*, was probably contaminated with mutations arising during reverse transcription and PCR amplification of the isolated virus, with no way to sort out which mutations were of viral and which of processing origin because the existence of any relation between the detected mutations and a mutant phenotype was not examined.
- (iv) For the (+)-strand hepatitis C virus, mutation sampling *in vivo* based on deep sequencing of plasma samples from untreated patients revealed an approximately 75-fold ratio of transitions to transversions with all four transitions present at similar frequencies, whereas kinetic studies *in vitro* of the viral replicase revealed a strong bias in favor of G $\rightarrow$ U and U $\rightarrow$ G mismatches [52], although mismatch extension efficiencies were not much explored. This contrast resembles that seen in numerous reversion systems, which are typically limited to small numbers of templates that tend towards often strongly site-specific mismatching biases.

Taken together, the informative parts of these spectra indicate that riboviral mutation spectra differ from those characteristic of DNA viruses and cellular organisms in displaying many more transitions than transversions and an even smaller proportion of indels.

### The Mode of Genome Replication

With the aims of determining the way in which mutations accumulate during riboviral replication and estimating the rate of spontaneous mutation per genome replication, we investigated the mode in which Q $\beta$  replicates its genome. Results from two independent single-burst reversion tests indicated that this mode is essentially linear, with the genome of each Q $\beta$  progeny resulting



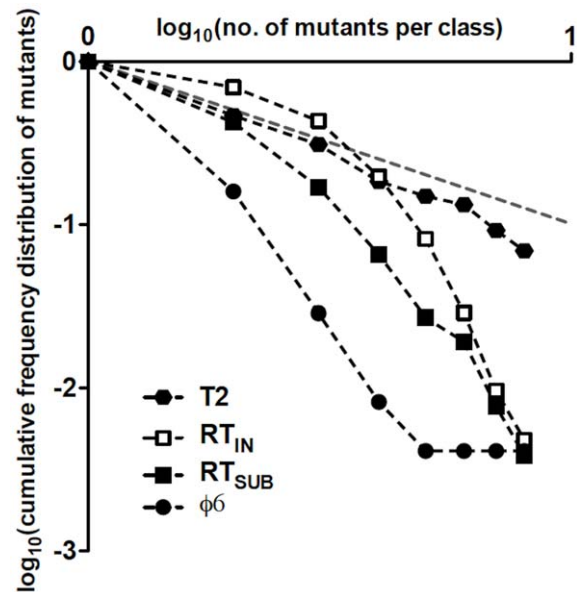
from only two replications: from the original parental (+) strand to a (-) strand and then to a new (+) strand. Our results further suggest that most replication errors occur during the second round of replication, which in turn reveals the specific mismatches that produced the substitutions in the mutational spectrum.

An interesting result is that the distribution of  $RT^+$  revertants deviated significantly from the expected Poisson distribution for  $RT_{SUB}$  but not for  $RT_{IN}$ . With no reason to suspect that the two strains replicate their genomes differently, this discrepancy may reflect intrinsic differences between their reversion targets. The reversion target in  $RT_{IN}$  is the same in both rounds of replication, while reversion in  $RT_{SUB}$  may occur through up to 8 different single-base-substitution pathways in each round of replication. Thus, reversion rate asymmetries between the two rounds of replication may be anticipated for  $RT_{SUB}$ , allowing some reversion to occur during the first round of replication and thus producing some revertant clones of size  $>1$  during the second round.

The mode of genome replication in riboviruses has been addressed in only a few instances. Using the single-burst reversion test, a predominantly linear mode was reported for the phage  $\phi 6$  [8]. In that study, however, the observed distribution of mutants (*i.e.*, revertants) differed somewhat from the expected Poisson for a linear mode of replication, suggesting an exponential component in the replication dynamics that was estimated to generate  $\approx 1\%$  of the total progeny [8]. Such discrepancies between observed and Poisson distributions may occur because of sampling errors or the presence of a small exponential component. A way to discriminate among these is to plot the logarithm of the cumulative frequency distribution of observed mutants against the logarithm of the sizes of the mutant classes [53]. Figure 5 shows such plot for T2,  $\phi 6$ ,  $RT_{IN}$ , and  $RT_{SUB}$ . In a log-log plot, exponential replication will display a linear relationship between the cumulative distribution of mutants and the number of mutants per class, with a slope close to  $-1$ . In linear replication, however, the plot will not be linear and the slope for the cumulative distribution will be steeper because most mutants arise in clones of size 1. In agreement with this reasoning, the data for T2, which replicates exponentially [42], exhibits a linear relationship with slope  $-1.20 \pm 0.03$ , based on the sum of the  $r$  and the  $w$  mutants and excluding all classes containing  $\geq 16$  mutants (*i.e.*, classes starting to approach the T2 burst size), while the plots for  $\phi 6$ ,  $RT_{IN}$ , and  $RT_{SUB}$  display nonlinear relationships.

In a recent report [9], the dynamics of (+)-strand and (-)-strand accumulation during cell infection were quantitatively analyzed for the (+)-strand RNA turnip mosaic virus using strand-specific quantitative real-time PCR. The results indicated that the virus replicates its genome in a mostly linear mode, in agreement with other quantitative results from *in silico* modeling of the optimal riboviral replication strategy in response to the error rate and the availability of resources, among other parameters [45,54]. However, the continuous accumulation of turnip mosaic virus (-) strands throughout infection suggests that a purely linear mode of replication may have been unlikely; indeed, the occurrence of a trace of exponential replication was reported. While we cannot exclude a trace of exponential replication in the case of QB, our results suggest that the  $RT_{SUB}$  revertant distribution may depart from a Poisson distribution mostly due to asymmetries in the reversion rates at the first and the second rounds of replication.

Overall, the empirical data gathered to date on the riboviral mode of replication indicate that, regardless of the single- or double-stranded genome structure of the virus, the strategy of preference is mainly linear. The advantages that this mode of replication may confer to riboviruses over an exponential mode have been evaluated previously (*e.g.*, [8]).



**Figure 5. Double-logarithmic plots of the relative cumulative frequency distributions of mutants for T2,  $\phi 6$ ,  $RT_{IN}$ , and  $RT_{SUB}$ .** The relative cumulative frequencies were calculated from Table 3 in [42] for T2 and from Table 1 in [8] for  $\phi 6$ . The plot of a cumulative frequency distribution of phage mutants from a single-burst reversion test has a slope whose steepness indicates the mode of replication of the phage; a steep slope resulting from large differences in frequency between adjacent mutant classes indicates a linear mode of replication, while a shallow slope indicates an exponential mode of replication. In a log-log plot, where the relative cumulative frequencies of mutants and the clone-size classes have a similar range of values, a slope close to  $-1$  (represented as a gray dashed line) will reflect an exponential mode of replication, while a linear mode of replication will show a slope  $< -1$ . doi:10.1371/journal.pgen.1002832.g005

### The Spontaneous Mutation Rate

Our results provide several independent estimates of the spontaneous QB mutation rate per genome replication ( $\mu_g$ ). The first is based mainly on the small set of three nonsense mutations detected among 30  $RT$  mutants. Because many base substitutions do not produce a detectable phenotype, the estimation of the  $\mu_{SUB}$  fraction of  $\mu_g = (\mu_{SUB} + \mu_{IN})G$  from the frequency of nonsense mutations is a preferred method because nonsense mutations are highly detectable and their target size is easy to determine from the codon composition of the mutation target. However, this method has two drawbacks: nonsense mutations are typically a small fraction of all substitutions, so that sufficient mutants must be harvested and sequenced for a reliable estimate [46]; and no substitutions to C (in the coding strand) can generate a stop codon, so that the average rate per base from the nonsense-generating pathways must be assumed to apply to all pathways. Using the nonsense-mutation method and adding the small component due to indel mutagenesis, the QB genomic rate was estimated to be  $\mu_g = 0.039$  per replication or about 0.08 per infection cycle. For this nonsense-based estimate of  $\mu_g$ ,  $RT$  mutants were collected after one-step growth of wt QB, so that  $c = 1$  in the calculations. While some prior  $RT$  mutations may have arisen during the growth of the wt QB stocks in non-complementing NR16205 lawns, lethal indels and nonsense mutations would have been subjected to strong negative selection. In some riboviruses, mutants bearing lethal mutations can grow in the presence of complementation *in trans* provided by a plasmid (*e.g.*, this study) or by a gene inserted into a host chromosome (*e.g.*, [10]), and thus it

may be assumed that complementation can also be provided by a co-infecting wild-type phage, which means that even *de novo* Q $\beta$  mutants carrying an *RT*<sup>-</sup> mutation may have expanded during the growth of the original wt Q $\beta$  stocks, rendering  $1 < c \leq 3$ . In the case where only one *RT*<sup>-</sup> mutant and one wild-type co-infect the same host cell and up to 50% of the resulting progeny are *RT*<sup>-</sup> mutants, the consequent selection coefficient (*s*) of the *RT*<sup>-</sup> mutant will be 0.50 per infection cycle. Applying the method described in Burch *et al.* [6] to estimate the effect of selection within the plaque and considering  $\mu_g = 0.039$ , the probability of loss of an *RT*<sup>-</sup> mutant with  $s \geq 0.50$  arising in the first infection cycle of wt Q $\beta$  on a host lawn would be  $\geq 42\%$  at the end of the growth phase (Figure S1). However, previous reports [22,55] indicate that co-infection by distinct Q $\beta$  mutants and consequent complementation occur at low to undetectable frequencies. Even if any *RT*<sup>-</sup> mutation arose during the last cycle of growth in NR16205 lawns, the small fraction of each wt Q $\beta$  isolate used to establish the one-step lysates further reduced the frequency of preexisting *RT*<sup>-</sup> mutations in our starting wt Q $\beta$  populations.

The second method for estimating  $\mu_g$  is based on the single-burst reversion tests. Here, the mutation rate is based on the null-class method [47]. For RT<sub>IN</sub>,  $\mu_{\text{target}} = 1.37 \times 10^{-6}$  and the number of mutating bases in the four-base duplication may be taken as 4 (although it may be argued to be 1); then  $\mu_{\text{INg}} = (4217/4)(1.37 \times 10^{-6}) = 0.0014$ . For RT<sub>SUB</sub>,  $\mu_{\text{target}} = 2.46 \times 10^{-5}$  and the potential number of mutating bases is 3; then  $\mu_{\text{SUBg}} = (4217/3)(2.46 \times 10^{-5}) = 0.035$ . The sum of the indel and the substitution rates is  $\mu_g = 0.036$  (0.041 if the indel reversion target size is taken as 1), a value (perhaps deceptively) close to that of  $\mu_g = 0.039$  calculated from the spectrum.

The third method for estimating  $\mu_g$  applies only in cases where the mutations are not required to produce a mutant phenotype, which can arise when a target is sequenced without regard to phenotype (provided the mutation is not a dominant lethal) or when, as in the present case, hitchhiker mutations arise secondarily to and in combination with a driver mutation, the target then consisting of the entire sequence of the mutation reporter. Hitchhikers could arise during any of the roughly 3 infection cycles that generate a Q $\beta$  plaque but, in order to be detected, most would have to arise in the first cycle with  $\mu_g = 0.287$ . This value may be an underestimate because the low ratio (6/7) of missense mutations to synonymous mutations among the secondary *RT* mutations suggests significant selection against missense mutations during plaque growth even in RTH lawns; if the hitchhikers were a random set, then missense mutations would comprise about  $\frac{3}{4}$ , or 9.75, of 13 substitutions. In fact, RT<sub>IN</sub> has an average burst size in RTH cells that is 2.9-fold smaller than that of wt Q $\beta$ , a difference that implies a selection coefficient of  $s \approx 0.65$  per infection cycle. An *RT* mutant with  $s \geq 0.65$  arising with  $\mu_g = 0.287$  in the first infection cycle on a RTH lawn would have a  $\geq 40\%$  probability of being lost by the end of the growth phase (Figure S1). Thus,  $\mu_g = 0.287$  estimated from hitchhikers might be significantly underestimated.

The presence of more mutants with multiple mutations than expected from a random distribution is remarkably widespread among DNA and RNA genomes and is probably more often due to transient hypermutation caused by some temporary perturbation of replication-fidelity factors than due to mutator mutations [56,57]. A notable example is the considerably higher frequency of mutations than expected among mutants already carrying a driver mutation produced by the replicase of the DNA phage RB69 [56]. However, because the Q $\beta$  replicase gene occupies 42% of the genome and the estimated  $\mu_g$  is high, we considered that some mutants might have arisen in a mutator background

and then gone on to produce hitchhikers at an increased frequency. Therefore, we examined whether the gene encoding the  $\beta$  subunit of the Q $\beta$  replicase harbored mutations in the 28 *RT* mutants carrying detectable primary mutations and their four parental wild-types. We observed three T $\rightarrow$ C substitutions (at  $\beta$ -subunit position 75 of mutant RT32, position 550 of RT42, and position 1668 of RT20) but all were synonyms, so that replicase mutators were apparently not impacting our set of *RT* mutations. Instead, the excess of secondary mutations among our *RT* mutants may have arisen by the action of an abnormal replicase  $\beta$  subunit produced by an error of translation or protein conformation.

Among the three values, our best estimate of  $\mu_g$  was obtained by the nonsense-mutation method. While the rate obtained from the reversion tests was similar, its accuracy depends on the extent to which the two mutants fairly sample the whole genome, and the similarity may have been fortuitous. Both selection and transient hypermutation may have played an important role in the production of the *RT* mutations considered in the third method. Unfortunately, even our favored  $\mu_g$  estimate is based on small samples of mutations (3 nonsense and 2 indels), which enlarges the margin of potential sampling error. When our first and second  $\mu_g$  estimates are combined with the fraction (0.4) of random mutations that are lethal for the (-)-strand-RNA vesicular stomatitis virus [58], (0.075 mutations per infection cycle  $\times$  591 nt per target  $\times$  7517 targets tested  $\times$  0.4 of mutations detectable)/4217 nt per genome = 32 *RT* mutants, in close agreement with the 30 observed and providing modest further support for  $\mu_g \approx 0.04$ .

While the mutation rates per genome replication estimated here and reported for TMV [10] and phage  $\phi 6$  [6] are all in the neighborhood of 0.04, rates for mammalian riboviruses center around 0.7 and display a wide range [7]. However, the latter rates were based on tiny targets often consisting of a single base or pathway and may have been reported because they were large and thus more easily measured; alternatively, as has been frequently suggested, immune surveillance in mammals may drive higher mutation rates. It is interesting that while the mutation frequency can be increased over the background with nitrous acid by up to 80-fold in tobacco mosaic virus with retention of some viability [59], it can be increased only about 2.5-fold in poliovirus and vesicular stomatitis virus before extinction begins [60], suggesting that mammalian riboviruses do indeed sustain mutation rates substantially higher than those of phage and plant riboviruses. Finally, although co-infection and complementation do not seem to occur at a detectable frequency with Q $\beta$ , it may occur with other riboviruses, perhaps somewhat elevating mutant frequencies and thus causing mutation rates to be overestimated.

## Materials and Methods

### Plasmids, Bacterial Strains, and Growth Media

Plasmids and bacterial strains are listed in Table 1. All three pQ $\beta$  plasmids express the indicated Q $\beta$  components constitutively and have been described [35,36,61]. The RT<sub>IN</sub> mutant carries a tandem duplication of 2158-UUAA-2161 that corresponds to 416–419 in the target sequence. Cell transformations with the plasmids were performed using CaCl<sub>2</sub> [62]. Unless otherwise indicated, RTH cells were grown in Luria-Bertani medium (LB) supplemented with 2 mM CaCl<sub>2</sub> and 100  $\mu$ g/ml trimethoprim (TMP), while NR16205 cells were grown in LB containing 15  $\mu$ g/ml tetracycline. Cells and phages were plated using LB bottom agar with 2.0% Bacto agar. The top agar was always made up in

distilled water. For counting plaques or scoring mutants, the top agar contained 0.4% Sigma-Aldrich Noble agar; for other uses, it contained 0.8% Bacto agar. All growth was at 37°C.

### One-Step Growth

NR16205 cells were transformed with pQ $\beta$ m100 (which expresses wt Q $\beta$ ) and plated on NR16205 lawns to yield wt Q $\beta$  plaques, which were independently harvested into tubes containing 1 ml D broth (0.2% Bacto tryptone, 0.5% NaCl) and 25  $\mu$ l of chloroform. For one-step growth curves in RTH cells, 10  $\mu$ l of phage suspension from a wt Q $\beta$  isolate was mixed with 1 ml of cells at OD<sub>600</sub>  $\approx$  0.5 (10<sup>8</sup> cells/ml) at a multiplicity of infection (MOI)  $\approx$  0.01 for 20 min at room temperature, centrifuged to remove non-adsorbed phages, resuspended, and serially diluted in LB+TMP. Samples diluted 10<sup>3</sup>- and 10<sup>5</sup>-fold were held for 3 h at 37°C with gentle shaking, and 100- $\mu$ l aliquots were removed from each dilution every 10 min and plated with RTH cells. Plates were incubated overnight and the follow-on titers were used to estimate Q $\beta$  densities over time. Three one-step growth experiments were conducted in parallel for each wt Q $\beta$  isolate used to generate one-step lysates. Visual inspection of the resulting curves sufficed to determine the time ( $\approx$ 75 min) for Q $\beta$  to complete one infection cycle in RTH cells. These one-step curves were also used to estimate the burst size of wt Q $\beta$  in RTH cells according to the protocol detailed for RT<sub>IN</sub> and RT<sub>SUB</sub> (see Table S1).

### Q $\beta$ Replication Mode

The distribution of RT<sup>+</sup> revertants among RT<sup>-</sup> bursts was monitored as in a previous study [8] using two different RT<sup>-</sup> mutants (RT<sub>IN</sub> and RT<sub>SUB</sub> as described in Table 1). Preliminary measurements provided their burst sizes (Figure S2, Table S1) and revertant frequencies, which are needed to conduct the burst experiments. Ten and five independent experiments were carried out with RT<sub>SUB</sub> and RT<sub>IN</sub>, respectively, and  $\approx$ 500 RT<sup>+</sup> revertants were scored per mutant. In each experiment,  $\approx$ 10<sup>6</sup> phages were added to 1 ml of RTH cells at OD<sub>600</sub>  $\approx$  0.5. After 20 min of adsorption at room temperature, the mixture was centrifuged for 1 min at 8,000 *g* and the pellet was resuspended in 1 ml LB broth. From the supernatant, 100  $\mu$ l were collected to estimate the amount of non-adsorbed phages. The resuspended pellet was further diluted and 50 aliquots of 100  $\mu$ l each were distributed into individual tubes, where infection was allowed to continue for  $\approx$ 75 min at 37°C and then stopped with 15  $\mu$ l of dichloromethane. Lysates were aerated for 30 min at 37°C to allow the dichloromethane to evaporate and their entire volumes were then independently plated on NR16205 lawns. The observed distributions of RT<sup>+</sup> revertants were compared to the expected Poisson distributions using G-tests for goodness-of-fit.

### Isolating and Sequencing Spontaneous RT Mutants

To limit the number of infection cycles to one before seeking spontaneous mutants, RT mutants were scored among the progeny of one-step growth of wt Q $\beta$  in RTH cells. RTH cells were infected with wt Q $\beta$  (MOI  $\approx$  0.01) as above and one-step lysates were recovered by adding chloroform after 75 min of growth. Samples from the lysates were plated on RTH lawns at  $\approx$ 70 plaques per plate and well-isolated plaques were independently sampled into 96-well plates containing 0.6 ml D-broth per well (reserving six un-inoculated wells as cross-contamination controls). For each of four independent lysates, three different rounds of 630 isolations each were performed. In each round, a control plate containing 8 wt and 82 RT<sub>IN</sub> isolates was also established to confirm the ability of RTH cells to complement RT<sup>-</sup> mutants and

the inability of any RTH cells remaining in the isolates to grow in LB supplemented with tetracycline. Isolates were spotted in parallel on lawns of NR16205 and RTH cells using a 6 $\times$ 8-array replica plater. After a few losses, a total of 7517 plaques were tested. Isolates that grew poorly in NR16205 cells were re-tested in both bacterial strains and the RT-coding genes of two independent sub-isolates per putative RT mutant were sequenced. After this first round of sequencing, two additional sub-isolates as well as the original isolate were sequenced for each verified RT mutant. The original wt Q $\beta$  isolate used to develop each lysate and two sub-isolates of it were also sequenced.

Plate lysates were prepared from RTH cells (0.25 ml at OD<sub>600</sub>  $\approx$  0.5) mixed with phages at MOI  $\approx$  0.1 in Noble top agar. After overnight incubation, the plates were covered with 7 ml of SM buffer with gelatin [62] and were gently rocked for 30 min. The SM buffer was recovered and 100  $\mu$ l of chloroform were added to each sample. Cell debris was removed by centrifugation at 12,000 *g* for 10 min. The supernatant was supplemented with polyethylene glycol (PEG 8000) to 10% w/v and NaCl to 1 M, incubated for 1 h on ice, and centrifuged at 3,000 *g* for 15 min at 4°C [63]. The pellets were resuspended in 2 ml of 10 mM MgSO<sub>4</sub>, 10 mM Tris-HCl, pH 8, and the resulting concentrated phages were used as sources for RNA purification. Phage RNA was isolated using the QIAamp Viral RNA Mini Kit. From the extracted RNAs, 10  $\mu$ g were then treated with DNase I (New England BioLabs) to degrade residual host DNA. The DNase-treated product was purified using the RNeasy Mini Kit. From the purified RNA, 1  $\mu$ g was subjected to reverse transcription with the Omniscript RT Kit and about 25 ng of the RT product was amplified with PfuTurbo DNA polymerase (Stratagene). PCR products were confirmed by agarose gel electrophoresis, purified with the QIAquick PCR Purification Kit, and sequenced using BigDye Terminator v3.1 (Applied Biosystems). All kits were purchased from Qiagen and were used according to the manufacturer's recommendations. Sub-isolates showing secondary mutations were subjected to a second round of RT, amplification and sequencing. The primers utilized in the RT, PCR, and sequencing reactions and the PCR cycling parameters are listed in Table S2.

### Statistical Analyses

The RT and genome base compositions were compared using the G-test of independence. This test was also applied to compare the observed distributions of RT<sup>+</sup> revertants among the single-burst reversion tests conducted with each of two different RT<sup>-</sup> mutants, RT<sub>SUB</sub> and RT<sub>IN</sub>. The G-test for goodness of fit was used to compare the observed and expected Poisson distributions of RT<sup>+</sup> revertants among RT<sup>-</sup> single-bursts, and the replicated G-test for goodness of fit was applied to compare the G+C content of the local sequence environment (six to seven bases upstream) of the base substitutions observed in RT with the expected content according to the base composition of the whole gene. When applying this last test, each upstream position (from +1 to +6 or +7) was considered as an independent replicate. All tests were performed as per Sokal and Rohlf [64].

### Supporting Information

**Figure S1** Probability that a new RT mutation with selection coefficient *s* is lost during plaque growth in a host lawn. The black line represents the estimations for  $\mu_g = 0.039$  while the gray line represents the estimations for  $\mu_g = 0.287$ . (TIF)

**Figure S2** One-step curves for RT<sub>IN</sub> and RT<sub>SUB</sub>. Three independent curves were obtained per RT mutant. In all cases, the first bursts appeared between 30 and 40 min after infection. If the 20 min allowed

for adsorption are also considered, it means that Q $\beta$  requires a minimum of 50 to 60 min to complete an infection cycle in RTH cells. Thus, if any of the progeny released in the first bursts would have immediately infected a new cell, the first second-generation bursts might have been expected 100 min after the first infection. Indeed, the curves show a slight increase of Q $\beta$  density at that time. For this reason, values collected 100 min after infection (empty symbols) were not considered in the curve-fitting analyses conducted to characterize RT<sub>IN</sub> and RT<sub>SUB</sub> single-burst dynamics. Squares represent outliers that were automatically excluded from the analyses, and the resulting fitted curves are shown as dashed curves. PFU = plaque-forming units. See Table S1 for additional information.

(TIF)

**Table S1** Parameters of the sigmoidal curves fitted to the RT<sub>IN</sub> and RT<sub>SUB</sub> one-step curves shown in Figure S2.

(DOC)

## References

- Holmes EC (2009) The evolution and emergence of RNA viruses. Oxford University Press. 254 p.
- Domingo E (2010) Mechanisms of viral emergence. *Vet Res* 41: 38.
- Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 51: 151–178.
- Drake JW, Holland JJ (1999) Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* 96: 13910–13913.
- Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R (2010) Viral mutation rates. *J Virol* 84: 9733–9748.
- Burch CL, Guyader S, Samarov D, Shen H (2007) Experimental estimate of the abundance and effects of nearly neutral mutations in the RNA virus  $\phi$ 6. *Genetics* 176: 467–476.
- Drake JW (1993) Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci U S A* 90: 4171–4175.
- Chao L, Rang CU, Wong LE (2002) Distribution of spontaneous mutants and the inferences about the replication mode of the RNA bacteriophage  $\phi$ 6. *J Virol* 76: 3276–3281.
- Sardanyés J, Martínez F, Daròs J-A, Elena SF (2011) Dynamics of a plant RNA virus intracellular accumulation: stamping machine *vs.* geometric replication. *Genetics* 188: 637–646.
- Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, et al. (2002) The rate and character of spontaneous mutation in an RNA virus. *Genetics* 162: 1505–1511.
- Tromas N, Elena SF (2010) The rate and spectrum of spontaneous mutations in a plant RNA virus. *Genetics* 185: 983–989.
- Das AT, Berkhout B (2010) HIV-1 evolution: frustrating therapies, but disclosing molecular mechanisms. *Phil Trans R Soc B* 365: 1965–1973.
- Anderson JP, Daifuku R, Loeb LA (2004) Viral error catastrophe by mutagenic nucleosides. *Ann Rev Microbiol* 58: 183–205.
- Weissmann C (1974) The making of a phage. *FEBS Lett* 40: S10–S18.
- Woody MA, Cliver DO (1995) Effects of temperature and host cell growth phase on replication of F-specific RNA coliphage Q $\beta$ . *Appl Environ Microbiol* 61: 1520–1526.
- Woody MA, Cliver DO (1997) Replication of coliphage Q $\beta$  as affected by host cell number, nutrition, competition from insusceptible cells and non-FRNA coliphages. *J Appl Microbiol* 82: 431–440.
- Tsukada K, Okazaki M, Kita H, Inokuchi Y, Urabe I, et al. (2009) Quantitative analysis of the bacteriophage Q $\beta$  infection cycle. *Biochim Biophys Acta* 1790: 65–70.
- Garwes D, Sillero A, Ochoa S (1969) Virus-specific proteins in *Escherichia coli* infected with phage Q $\beta$ . *Biochim Biophys Acta* 186: 166–172.
- Radloff RJ, Kaesberg P (1973) Electrophoretic and other properties of bacteriophage Q $\beta$ : the effect of a variable number of Read-Through proteins. *J Virol* 11: 116–128.
- Blumenthal T, Carmichael GG (1979) RNA replication: function and structure of Q $\beta$ -replicase. *Ann Rev Biochem* 48: 525–548.
- Eigen M, Biebricher CK, Giebinoga M, Gardiner WC (1991) The hypercycle. Coupling of RNA and protein biosynthesis in the infection cycle of an RNA bacteriophage. *Biochemistry* 30: 11005–11018.
- Horiuchi K, Matsuhashi S (1970) Three cistrons in bacteriophage Q $\beta$ . *Virology* 42: 49–60.
- Vollenweider HJ, Koller T, Weber H, Weissmann C (1976) Physical mapping of Q $\beta$  replicase binding sites on Q $\beta$  RNA. *J Mol Biol* 101: 367–377.
- Edlind TD, Bassel AR (1977) Secondary structure of RNA from bacteriophages  $\phi$ 2, Q $\beta$  and PP7. *J Virol* 24: 135–141.
- Takamatsu H, Iso K (1982) Chemical evidence for the capsomeric structure of phage Q $\beta$ . *Nature* 298: 819–824.
- Schuppli D, Barrera I, Weber H (1994) Identification of recognition elements on bacteriophage Q $\beta$  minus strand RNA that are essential for template activity with Q $\beta$  replicase. *J Mol Biol* 243: 811–815.
- Golmohammadi R, Fridborg K, Bundule M, Valegård K, Liljas L (1996) The crystal structure of bacteriophage Q $\beta$  at 3.5 Å resolution. *Structure* 4: 543–554.
- Beekwilder MJ, Nieuwenhuizen R, van Duin J (1995) Secondary structure model for the last two domains of single-stranded RNA phage Q $\beta$ . *J Mol Biol* 247: 903–917.
- Beekwilder J, Nieuwenhuizen R, Poot R, van Duin J (1996) Secondary structure model for the first three domains of Q $\beta$  RNA. *J Mol Biol* 256: 8–19.
- Inokuchi Y, Kajitani M (1997) Deletion analysis of Q $\beta$  replicase. *J Biol Chem* 272: 15339–15345.
- Kidmose RT, Vasiliev NN, Chetverin AB, Rom Andersen G, Knudsen CR (2010) Structure of the Q $\beta$  replicase, an RNA-dependent RNA polymerase consisting of viral and host proteins. *Proc Natl Acad Sci U S A* 107: 10884–10889.
- Takehita D, Tomita K (2010) Assembly of Q $\beta$  viral RNA polymerase with host translational elongation factors EF-Tu and -Ts. *Proc Natl Acad Sci U S A* 107: 15733–15738.
- Hofstetter H, Monstein H-J, Weissmann C (1974) The readthrough protein A1 is essential for the formation of viable Q $\beta$  particles. *Biochim Biophys Acta* 374: 238–251.
- Van Duin J, Tsareva N (2006) Single-stranded RNA phages. In: Calendar R, editor. *The bacteriophages*. 2<sup>nd</sup> edition. Oxford University Press. pp 175–196.
- Priano C, Arora R, Butke J, Mills DR (1995) A complete plasmid-based complementation system for RNA coliphage Q $\beta$ : three proteins of bacteriophages Q $\beta$  (Group III) and SP (Group IV) can be interchanged. *J Mol Biol* 249: 283–297.
- Arora R, Priano C, Jacobson AB, Mills DR (1996) *ais*-Acting elements within an RNA coliphage genome: fold as you please, but fold you must!! *J Mol Biol* 258: 433–446.
- Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, et al. (1968) Frameshift mutations and the genetic code. *Cold Spring Harbor Symp Quant Biol* 31: 77–84.
- Bebenek K, Kunkel TA (2000) Streisinger revisited: DNA synthesis errors mediated by substrate misalignments. *Cold Spring Harbor Symp Quant Biol* 65: 81–92.
- Halliday JA, Glickman BW (1991) Mechanisms of spontaneous mutation in DNA repair-proficient *Escherichia coli*. *Mutat Res* 250: 55–71.
- Drake JW (2012) Contrasting mutation rates from specific-locus and long-term mutation-accumulation procedures. *G3* 2: 483–485.
- Bebenek A, Dressman HK, Carver GT, Ng S, Petrov V, et al. (2001) Interacting fidelity defects in the replicative DNA polymerase of bacteriophage RB69. *J Biol Chem* 276: 10387–10397.
- Luria SE (1951) The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. *Cold Spring Harbor Symp Quant Biol* 16: 463–470.
- Ronen A, Rahat A (1976) Mutagen specific and position effects on mutation in *T4rII* nonsense sites. *Mut Res* 34: 21–34.
- Billeter MA, Libonati M, Viñuela E, Weissmann C (1966) Replication of viral ribonucleic acid. X. Turnover of virus-specific double-stranded ribonucleic acid during replication of phage MS2 in *Escherichia coli*. *J Biol Chem* 241: 4750–4757.
- Thébaud G, Chadoeuf J, Morelli MJ, McCauley JW, Haydon DT (2010) The relationship between mutation frequency and replication strategy in positive-sense single-stranded RNA viruses. *Proc R Soc B* 277: 809–817.
- Drake JW (2009) Avoiding dangerous missense: thermophiles display especially low mutation rates. *PLoS Genet* 5: e1000520. doi:10.1371/journal.pgen.1000520
- Luria SE, Delbrück M (1943) Mutations in bacteria from virus sensitivity to virus resistance. *Genetics* 28: 491–511.
- Wagner J, Nohmi T (2000) *Escherichia coli* DNA polymerase IV mutator activity: genetic requirements and mutational specificity. *J Bacteriol* 182: 4587–4595.

**Table S2** Primers and PCR cycling parameters for amplifying and sequencing. (DOC)

## Acknowledgments

We heartily thank Don Mills for supplying the Q $\beta$  plasmids and for insights into their properties. Roel Schaaper provided the *E. coli* strains. We thank Katarzyna Bebenek and Roel Schaaper for critical readings of the manuscript and Stanislav Kozmin for advice on the manipulation of the bacterial strains used in this study.

## Author Contributions

Conceived and designed the experiments: LG-V JWD. Performed the experiments: LG-V. Analyzed the data: LG-V JWD. Contributed reagents/materials/analysis tools: LG-V. Wrote the paper: LG-V JWD.

49. Schultz GE Jr, Carver GT, Drake JW (2006) A role for replication repair in the genesis of templated mutations. *J Mol Biol* 358: 963–973.
50. Lang GI, Murray AW (2008) Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 178: 67–82.
51. Fraser JLA, Neill E, Davey S (2003) Fission yeast Uve1 and Apn2 function in distinct oxidative damage repair pathways in vivo. *DNA Repair* 2: 1253–1267.
52. Powdrill MH, Tchesnokov EP, Kozak RA, Russell RS, Martin R, et al. (2011) Contribution of a mutational bias in hepatitis C virus replication to the genetic barrier in the development of drug resistance. *Proc Natl Acad Sci U S A* 108: 20509–20513.
53. Denhardt DT, Silver RB (1966) An analysis of the clone size distribution of  $\phi$ X174 mutants and recombinants. *Virology* 30: 10–19.
54. Regoes RR, Crotty S, Antia R, Tanaka MM (2005) Optimal replication of poliovirus within cells. *Am Nat* 165: 364–373.
55. Ling CM, Hung PP, Overby LR (1970) Independent assembly of Q $\beta$  and MS2 phages in doubly infected *Escherichia coli*. *Virology* 40: 920–929.
56. Drake JW, Bebenek A, Kissling GE, Peddada S (2005) Clusters of mutations from transient hypermutability. *Proc Natl Acad Sci U S A* 102: 12849–12854.
57. Drake JW (2007) Mutations in clusters and showers. *Proc Natl Acad Sci U S A* 104: 8203–8204.
58. Sanjuán R, Moya A, Elena SF (2004) The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc Natl Acad Sci U S A* 101: 8396–8401.
59. Gierer A, Mudry KW (1958) Production of mutants of tobacco mosaic virus by chemical alteration of its ribonucleic acid in vitro. *Nature* 182: 1457–1458.
60. Holland JJ, Domingo E, de la Torre JC, Steinhauer DA (1990) Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J Virol* 64: 3960–3962.
61. Mills DR, Priano C, Merz PA, Binderow BD (1990) Q $\beta$  RNA bacteriophage: mapping *cis*-acting elements within an RNA genome. *J Virol* 64: 3872–3881.
62. Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*. 3<sup>rd</sup> edition. Cold Spring Harbor (New York): Cold Spring Harbor Laboratory Press. 999 p.
63. Yamamoto KR, Alberts BM, Bezing R, Lawhorne L, Treiber G (1970) Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* 40: 734–744.
64. Sokal RR, Rohlf FJ (1995) *Biometry*. 3<sup>rd</sup> edition. W.H. Freeman and Company. 887 p.
65. Kozmin SG, Pavlov YI, Dunn RL, Schaaper RM (2000) Hypersensitivity of *Escherichia coli*  $\Delta$ (*uvrB-bio*) mutants to 6-hydroxylaminopurine and other base analogs is due to a defect in molybdenum cofactor biosynthesis. *J Bacteriol* 182: 3361–3367.