Research article

Phenotypic error threshold; additivity and epistasis in RNA evolution

Nobuto Takeuchi*, Petrus H Poorthuis and Paulien Hogeweg

Address: Theoretical Biology/Bioinformatics Group, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

Email: Nobuto Takeuchi* - takeuchi_nobuto@yahoo.co.jp; Petrus H Poorthuis - Pieter@porthome.nl;
Paulien Hogeweg - P.Hogeweg@bio.uu.nl

* Corresponding author

Abstract

Background: The error threshold puts a limit on the amount of information maintainable in Darwinian evolution. The error threshold was first formulated in terms of genotypes. However, if a genotype-phenotype map involves redundancy ("mutational neutrality"), the error threshold should be formulated in terms of phenotypes since there is no unique fittest genotype. A previous study formulated the error threshold in terms of phenotypes, and their results showed that a rather low degree of mutational neutrality can increase the error threshold unlimitedly.

Results: We obtain an analytical formulation of the phenotypic error threshold by considering the "additive assumption", in which base substitutions do not influence each other (no epistasis). Our formulation shows that an increase of the error threshold due to mutational neutrality is limited. Computer simulations of RNA evolution are conducted to verify our formulation, and the results show a good agreement between the analytical prediction and the simulations. The comparison with the previous formulation illustrates that it is important for the prediction of the error threshold to consider that the number of base substitutions per replication is rather large near the error threshold. To examine the additive assumption, a detailed analysis of additivity and epistasis in RNA folding of a particular sequence is performed. The results show a high degree of epistasis in RNA folding; furthermore, the analysis also elucidates the reason of the success of the additive assumption.

Conclusions: We conclude that an increase of the error threshold by mutational neutrality is limited, and that the additive assumption achieves a good prediction of the error threshold in spite of a high degree of epistasis in RNA folding because the average number of base substitutions of sequences retaining the phenotype per replication is sufficiently small to avoid of the effect of epistasis.

Background

The error threshold is a limit on the permissible mutation rate for which "survival of the fittest" holds in Darwinian evolution [1]. The error threshold can be seen as a limit on the amount of information maintainable in evolutionary systems (information threshold) since an increase in sequence length results in an increase in error rate. The information threshold leads to a paradox in prebiotic evolution [2]. Suppose that to increase the maintainable amount of information, an evolving system must acquire

a more complex molecular mechanism to reduce the mutation rate. However, to have such a complex molecular mechanism the system must maintain a longer sequence in the first place. Thus, the system will encounter a barrier in the evolution of complexity (cf. [3]).

The error threshold was first formulated in terms of genotypes. However if some changes in genotype do not alter the phenotype or the fitness (mutational neutrality), there is no unique genotype which can be stably maintained. Instead, the survival of a phenotype should be considered, and thus the error threshold should be formulated in terms of phenotypes [4].

In this paper, first we formulate the phenotypic error threshold analytically by employing the additive assumption, in which base substitutions do not influence each other. Under the additive assumption, we obtain the probability that a replication does not alter the phenotype (neutral replication) as a function of the number of base substitutions and the fraction of "neutral substitutions". Our results show a qualitative difference from the previous formulation [5]. Second, we analyze epistasis in RNA folding of a particular RNA sequence to examine the additive assumption.

Results and discussion
Phenotypic error threshold
Analytical formulation

The quasispecies equation describes (prebiotic) replicator dynamics in well-mixed systems [1]. We transform the equation in two ways: (1) describing the abundance of phenotypes instead of that of genotypes by denoting the population of genotypes which share the same phenotype by one variable (see [5] for mathematical details); (2) distinguishing only two classes of phenotypes, the focal phenotype (denoted by x) and the others called mutants (denoted by y). The population size is assumed to be large enough to express the abundance of the phenotypes by normalized concentration. The population dynamics of the phenotypes is described as

$$dx/dt = \sigma Q x + \sigma \Lambda (1 - Q)x - Dx - \Phi x,$$

$$dy/dt = y + \sigma (1 - \Lambda)(1 - Q)x - Dy - \Phi y \quad (1)$$

where σ (> 1) is the replication rate of x (that of y is normalized to 1); Q is the replication accuracy of x. Λ is the fraction of neutral mutants of x. D is the degradation rate (or the death rate) assumed to be uniform over the phenotypes. $\Phi = (\sigma - D)x + (1 - D)y$ is the excess production (or the mean fitness). The terms $-\Phi x$ and $-\Phi y$ induce a selection pressure. We neglect back mutation from y to x (this simplification will be discussed in the next section). From Eq. 1, we obtain the survival condition of x as $Q + \Lambda$

$(1 - Q) > \sigma^1$, for which the stationary value of x is larger than zero. From this inequality, we will deduce the phenotypic error threshold, i.e., the maximum error rate of replication for which x can be stably maintained in the system.

We distinguish two classes of single base substitutions: neutral and deleterious substitutions. The class of a substitution is determined by the effect of the substitution on the phenotype when there are no other substitutions in the genotype: A substitution which retains the phenotype is a neutral substitution; otherwise it is a deleterious substitution. A beneficial substitution is not considered since the focus of the study is on the maintenance of x. [A replicator is thought of as a polymer in our study. We refer to a monomer as a base, having RNA in mind. The formulation of the phenotypic error threshold itself is independent of this terminology.]

To calculate the effective replication accuracy $Q_e = Q + \Lambda(1 - Q)$, we assume that a mutant is neutral *iff* there is no deleterious substitution: The effect of a substitution on the phenotype is independent of the other substitutions; i.e., no epistasis is assumed (the additive assumption). Let λ denote the fraction of neutral substitutions in all possible single substitutions, and let d denote the number of substitutions per replication. Then, the probability that d substitutions are all neutral substitutions (thus neutral replication) is approximated by λ^d by assuming that the number of neutral substitutions in d substitutions follows the binomial distribution (the binomial approximation). This approximation is valid if the probability of correct replication per base (denoted by q) is sufficiently large so that d is small. Denoting the sequence length of replicators by N, the effective replication accuracy (Q_e) is obtained as

$$Q_e = \sum_{d=0}^{N} \binom{N}{d} q^{N-d} (1-q)^d \lambda^d$$
$$= \left\{ q + (1-q)\lambda \right\}^N, \quad (2)$$

by assuming that q and λ are uniform among the genotypes in x, that q is invariable over sequence positions, and that N is the same among the populations. [A similar formula was obtained in [6] as $Q_e = (q + v(1 - q))^N$, where v is a parameter to be tuned to match the formula to the observed value of Q_e. Therefore v in [6] implicitly involves both the additive effect and epistasis.]

The minimum q for which x can survive is derived from $Q + \Lambda(1 - Q) > \sigma^1$ and Eq. 2 as

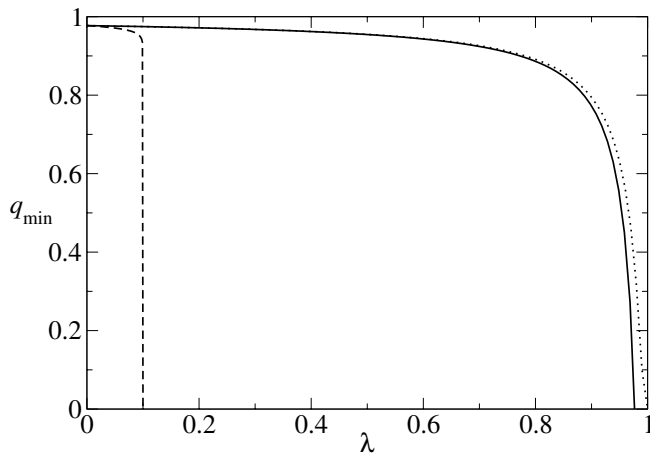$$q_{\min} = (\sigma^{1/N} - \lambda)/(1 - \lambda). \quad (3)$$

Figure 1
Error threshold The minimum permissible replication accuracy per base (q_{min}) is plotted against λ for three different ways of the calculation. The solid line is obtained from the additive assumption and the binomial approximation (Eq. 3). The binomial approximation is threatened at a high error rate. To examine this, the error threshold is calculated without the binomial approximation (but with the additive assumption) in the extreme example, where the binomial approximation deviates most (see the text). Let N_δ be the sum of the sequence length of the parts where all single mutations are deleterious. Then q_{mim} is calculated as

σ^{-1/N_δ} . The dotted line represents the so calculated error threshold in this extreme example. The *x*-axis for the dotted line (i.e., λ) is calculated as $(N - N_\delta)/N$. The dashed line is obtained from the formulation of Reidys *et al.* [5] (Eq. 5). In all cases, $N = 100$ and $\sigma = 10$. (The same values of N and of σ as those used in [5] are chosen for a comparison purpose.)

The phenotypic error threshold is $1 - q_{min}$. As seen in Fig. 1 (the solid line), the increase in the error threshold is limited for almost all values of λ. This is because if q decreases, the number of substitutions per replication (d) increases; hence the probability of neutral replication (λ^d) decreases (cf. [5] and Eq. 5). At a large λ (= $\sigma^{1/N}$), there is a singularity such that q_{min} becomes zero. However, this singularity is not plausible in two ways: (1) Such a large λ is not realistic (see below); (2) q at the singularity is so small that the binomial approximation is threatened. We studied the validity of the binomial approximation, and found that the inaccuracy in the binomial approximation is largest if in some positions of the sequence all possible single substitutions are neutral, but in the rest of the positions all possible single substitutions are deleterious; i.e., the distributions of neutral and deleterious substitutions over the sequence positions are completely separated. By taking this extreme example, q_{min} is calculated with the

additive assumption but without the binomial approximation. As Fig. 1 (the dotted line) shows, for a wide range of λ the binomial approximation is valid. q_{min} is underestimated by the binomial approximation only around the singularity ($\lambda > 0.8$), and thus the singularity is actually located at higher λ, which makes the singularity even less plausible. We conclude that the increase in the error threshold due to mutational neutrality is limited.

From Eq. 3, we obtain the information threshold, i.e., the maximum permissible sequence length as

$$N_{max} = \ln(\sigma^{-1})/\ln(q + (1 - q)\lambda). \quad (4)$$

As Fig. 2 (the solid lines) shows, the increase in the information threshold is limited for plausible values of λ. (N_{max} reaches infinity only when λ increases to one.) However, this result does not mean that a longer sequence in fact can have a larger λ, and thus it can be maintained. We studied the relationship between λ and the sequence length by utilizing RNA folding, which is a well-studied prototype of genotype-phenotype map, where the genotype is the primary structure of an RNA sequence and the phenotype is the minimum free energy secondary structure of the RNA sequence. We utilized Vienna RNA package [7] to fold RNA (The default parameters are used in the all occasions in the study). The average λ for different sequence length was obtained through comparing the secondary structure of randomly created RNA sequences with that of all possible mutants with only one substitution. A substitution which retains the original secondary structure is considered as neutral; otherwise deleterious. As Fig. 2 (the filled circles) shows, the average λ is a decreasing function of the sequence length. This relationship further limits an increase of information threshold due to mutational neutrality.

Comparison between the analytical prediction and computer simulations
We compare our analytical prediction with computer simulations. Our computer program simulates the evolution of RNA replicators in a well-mixed flow reactor (e.g. see [8]). In the simulations, each RNA sequence replicates and/or is diluted (be taken out from the reactor) with a certain probability in every time step. RNA folding is utilized again as a genotype-phenotype map (computed by [7]): The fitness of an RNA sequence depends on the secondary structure (i.e., the phenotype) of the RNA sequence. The fittest phenotype is set to the secondary structure of a yeast tRNAphe (the clover leaf structure, $N = 76$). RNA sequences which have the fittest phenotype replicate with the probability 0.01 per time step; all the other RNA sequences (mutants) replicate with the probability 0.001 per time step (thus $\sigma = 10$). The replication introduces mutations with a certain probability. Back
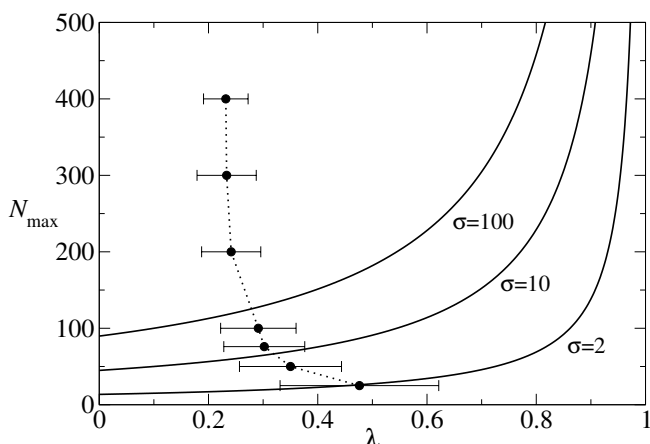
Figure 2
Information threshold The solid lines represent the maximum maintainable sequence length (N_{max}) plotted against λ, where N_{max} is calculated by using Eq. 4 where σ = 2, 10 and 100 (as indicated within the figure) and q = 0.95. The value of q was chosen to be plausible for ribozyme polymerization [14]. These solid lines show the dependence of N_{max} on λ; on the other hand, the dependence of λ on N (the length of sequence) is examined by calculating the average λ in RNA folding for various values of N. The filled circles represent so obtained λ values as a function of N. In obtaining λ, a neutral substitution is defined as a single base substitution which does not alter the secondary structure of a focal sequence. For each sequence length, a hundred randomly generated RNA sequences are examined. In RNA folding program [7], the default parameters are used (in the all occasions in the study).
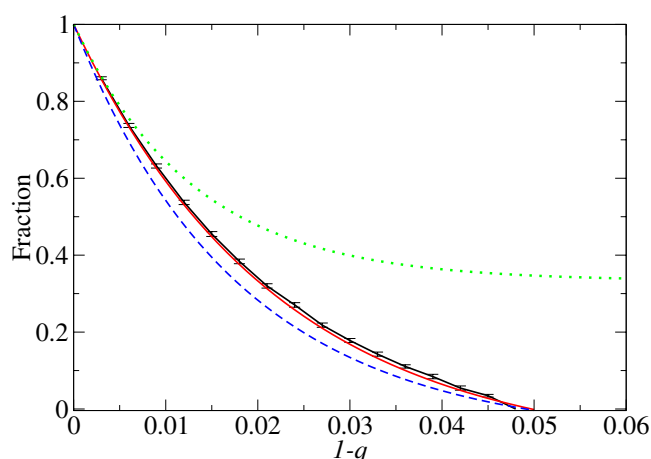

Figure 3
Comparison of population structure between analytical predictions and computer simulations The fraction of the fittest phenotype population is plotted against the error rate per base (1 - q). The black line (the solid line with error bars) is obtained from computer simulations. The red line (the solid line without error bars) is calculated with the additive assumption from Eq. 1 and Eq. 2 with σ = 10 and N = 76 as in the simulations. λ is set to be 0.4 which is the time averaged representative λ value observed in the simulations after the population evolves (see the Methods section – Non-uniform distribution of λ – for the definition of the representative λ). The green line (the dotted line) is obtained from the formulation of Reidys *et al.* [5] with the same parameters as the above. The blue line (the dashed line) is calculated with the four λ approximation by using λ values reported in [5] (λ = 0.2489 on average) and σ = 10 and N = 76 as in our simulations.

mutations are not allowed to occur – the effect of back mutations is negligible if the sequence length is large enough [9] (this was confirmed by the simulations which were the same as the above except for allowing back mutations [data not shown]). The dilution probability (Φ) is calculated as the average probability of replication divided by the target population size. The target population size is set to 10000. All the simulations start with 10000 yeast tRNAphe sequences. The degradation of sequences is ignored (D = 0).

The results of the computer simulations showed that the "representative λ" (defined in Methods section – Non-uniform distribution of λ) of the fittest sequences increased from 0.307 to ca. 0.40 for the examined values of the error rate (data not shown). (This is also true for the population average of λ [op. cit.].) The value of the representative λ fluctuates over the time (st. dev. = 0.01 at 1 - q = 0.0475).

The equilibrium fraction of the fittest sequences of the computer simulations is compared to that of the analytical prediction over the different error rate in Fig. 3. The analytical prediction is calculated under the additive assumption from Eq. 1 and Eq. 2 by using the time averaged representative λ observed in the simulations after the evolution (λ = 0.4). As Fig. 3 shows, the calculation (the [red] solid line without error bars) closely predicts the result obtained from the computer simulations (the [black] solid line with error bars). The predicted error threshold (0.05) is slightly higher than that observed (between 0.045 and 0.048) probably due to the assumption of infinite population in Eq. 1 (see [3,10]).

Comparison with a previous formulation
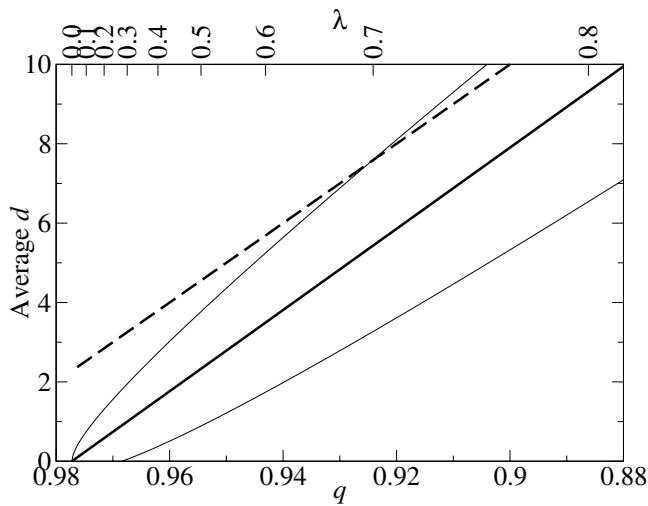Reidys *et al.* [5] derived the phenotypic error threshold as

Figure 4
Number of substitutions per replication in mutants
The y-axis is the number of base substitutions (d) per replication (or per neutral replication) at the error threshold. The thick solid line represents the average d per neutral replication (i.e., the average d of the sequences which retain the master phenotype per replication): $d = Np/(q_{mim} + p)$ where $p = \lambda (1 - q_{min})$ and $\lambda = (\sigma^{1/N} - q_{min})/(1 - q_{min})$. The thin solid line represents the standard deviation of it, i.e., $\pm (q_{mim} + p)$

$Np \sqrt{(N-1)/N + (q_{min} + p)/(Np) - 1}$. The dashed line represents the average d per replication, which is $N(1 - q_{min})$. N is 100 and σ is 10. The lines are plotted against q_{min} (the lower x-axis), and the corresponding λ is shown in the upper x-axis.

$$q_{min} = \left(\frac{1 - \lambda\sigma}{(1 - \lambda)\sigma} \right)^{1/N} \qquad (5)$$

from $Q_e = Q + \lambda (1 - Q)$. This equation shows an unlimited increase in the error threshold for $\lambda \geq \sigma^1$ (see the dashed line in Fig. 1 and the [green] dotted line in Fig. 3). However, $Q_e = Q + \lambda (1 - Q)$ is valid only if either (1) a neutral set is uniformly distributed over the genotype space [a neutral set is a set of genotypes where all genotypes map to the same phenotype], or (2) q is so large that most mutants have d = 1. The uniform distribution of neutral sets in the genotype space is not applicable in RNA folding as shown later. The latter possibility is discussed next.

Studies of replicator dynamics on a neutral network often consider a very large value of q so that most mutants have d = 1 (e.g., [11]). [A neutral network is a neutral set, or its

Table 1: Definition of positive and negative epistasis.

| | mutants are | |
	neutral	deleterious
$\delta = 0$	additive neutral	negative epistasis
$\delta > 0$	positive epistasis	additive deleterious

δ is the number of deleterious base substitutions.

subset, where every genotype is connected to at least one genotype of the set by one or two base substitutions.] However, if the error rate (1 - q) is close to the error threshold, mutants can have on average d > 1 even if $\lambda = 0$, for which the error threshold is at the lowest error rate (see the dashed line in Fig. 4). The average d of the neutral sequences (i.e., the exact copies and the neutral mutants) per replication (this will be later called the average d per neutral replication) is lower than the average d per replication. However at the error threshold, even the average d per neutral replication is larger than one for $\lambda > 0.32$ (see Fig. 4, the solid lines). Above consideration asserts that the error threshold will be substantially overestimated if one considers only a single mutation.

Reidys et al. [5] obtained an extension of Eq. 5, the so called "four λ approximation". This extension divides a sequence in four sub-sequences in order to take into account the fact that the fraction of neutral substitutions varies over the sequence position. This extension still overestimates Q_e though less so than Eq. 5 because the approximation now permits four substitutions per replication as a side effect of the subdivision. Note that this extension makes a fairly good prediction on the error threshold (see the [blue] dashed line Fig. 3) because the use of a small non-evolved λ value coincidentally cancels out the overestimation.

In conclusion, it is crucial for the calculation of the error threshold to consider that the number of substitutions per replication is large near the error threshold.

Epistasis in RNA folding
The rather impressive success of the additive assumption is counter-intuitive in view of RNA folding, in which many interactions occur between bases. In the next part of the paper, we study a particular RNA sequence, namely the yeast tRNAphe (which comprises the initial population of the previously described RNA evolution simulations), in terms of additivity and epistasis. The objective of this study is to understand how the additive assumption achieves a good prediction in spite of a high degree of nonlinearity in RNA folding [12,13].
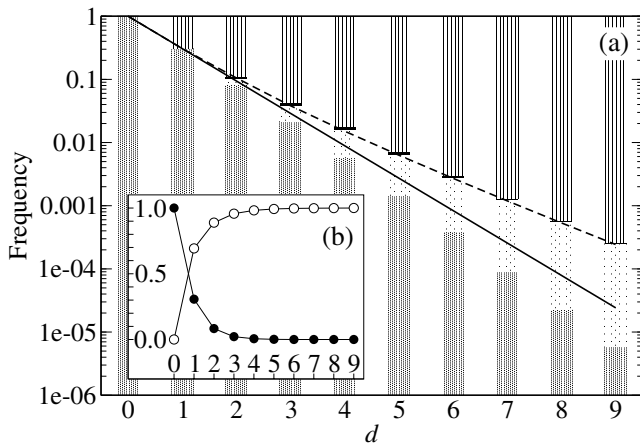
Figure 5
Additivity and epistasis in RNA folding The frequency of mutant classes is plotted against the number of base substitutions (*d*). **(a)** Log. plot. The patterns in the bars indicate the mutant classes: (from bottom) mesh, additive neutral; dots, positive epistasis; black, negative epistasis; stripes, additive deleterious (see Table 1 for the definition). The data were generated by RNA folding (by using [7]) with a *S. cerevisiae* tRNAphe sequence as a reference sequence: GCGGAU-UUACCUCAGUUGGGAGAGGGCCAGACUGAACAUCUGGAGGUCCGGCGCGCGAUACGCCGAAUUCG-CACCA (each non-RNA is converted to RNA). We examined all possible mutants at *d* = 1, 2 and the subsets of mutants for other *d* values (2<*d*<10, the portion of examined mutants is respectively, 10, 1, 0.1, 0.01, 0.3×10^{-3}, 0.1×10^{-3}, 0.4×10^{-5}%). These observations in RNA folding are compared with the following two analytical predictions. The solid line is the probability of neutral replication estimated under the additive assumption (λ^d, λ = 0.307). The dashed line is the probability of neutral replication estimated with epistasis

($\tilde{\lambda}$ (*d*), see Methods section – Probabilistic approach). **(b)** Linear plot. Symbols: ● the frequency of the neutral mutants (additive neutral and positive epistasis); ○ the frequency of the deleterious mutants (additive deleterious and negative epistasis).

We compare the secondary structure of randomly sampled mutants to that of the tRNAphe. Similar to the previous section, a mutant is neutral if its secondary structure is the same as that of the tRNAphe; otherwise, it is deleterious. To evaluate the deviation from the additive assumption, we categorize mutants into four classes as shown in Table 1. Negative epistasis refers to a mutant which is predicted to be neutral under the additive assumption, but turns out to be deleterious due to the interaction of the base substitutions. Positive epistasis refers to the reverse case.

As Fig. 5 shows, the additive assumption underestimates the degree of mutational neutrality. The same conclusion was drawn differently in [13], where the additive neutral mutant is defined as a neutral mutant which lies in the same neutral network as that of the original sequence. Our results show that positive epistasis occurs more frequently than negative epistasis in total. What actually happens is as follows. If mutants with $\delta > 0$ are only considered, positive epistasis occurs very rarely compared to additive deleterious case: No more than 0.5% of the mutants are neutral at *d* = 5 if they carry at least one deleterious substitution. If mutants with $\delta = 0$ are only considered, negative epistasis is rather frequent relative to additive neutral case: As much as 35% of the mutants are deleterious at *d* = 5 even if they carry only neutral substitutions. However, replication with $\delta > 0$ occurs far more frequently than replication with $\delta = 0$: As much as 99.7% of the replication contains at least one deleterious substitution at *d* = 5 and λ = 0.307. Therefore, the relative frequency of epistasis is flipped around. Consequently, the additive assumption underestimates the degree of mutational neutrality. (Note that in Fig. 3 the additive assumption predicts the fraction of the fittest sequences always slightly smaller than that of the computer simulations.)

The effect of epistasis is already noticeable when *d* > 2 as seen in the comparison between the probability of neutral replication under the additive assumption (λ^d) and that observed in RNA folding (see Fig. 6a). Since the average *d* per replication is more than 3 close to the error threshold in our simulations, Fig. 6a may seem to suggest that the additive assumption would substantially underestimate the effective replication accuracy (Q_e) near the error threshold.

We calculate the effective replication accuracy (Q_e) including the effect of epistasis in order to compare it with Q_e calculated under the additive assumption. The first trial was to include a "trivial" epistasis in base paired regions (helices) as a part of the additive effect (see Methods section – Trivial epistasis). However, the analysis showed that epistasis occurs mainly in a "non-trivial" way (data not shown), and thus it is not sufficient for our sake to include a trivial epistasis. We next took a probabilistic approach to calculate Q_e with epistasis (see Methods section – Probabilistic approach). The results of this method agree with the observation (see the dashed line in Fig. 5).

We compare Q_e calculated under the additive assumption to that calculated with epistasis as shown in Fig. 6b (the solid line). As the comparison shows, the additive assumption indeed underestimates Q_e; however, the underestimation becomes prominent only if the error rate is higher than the error threshold (1 - $q_{\min}$ = 0.05). As Fig. 6b (the dashed line) shows, the average *d* per neutral
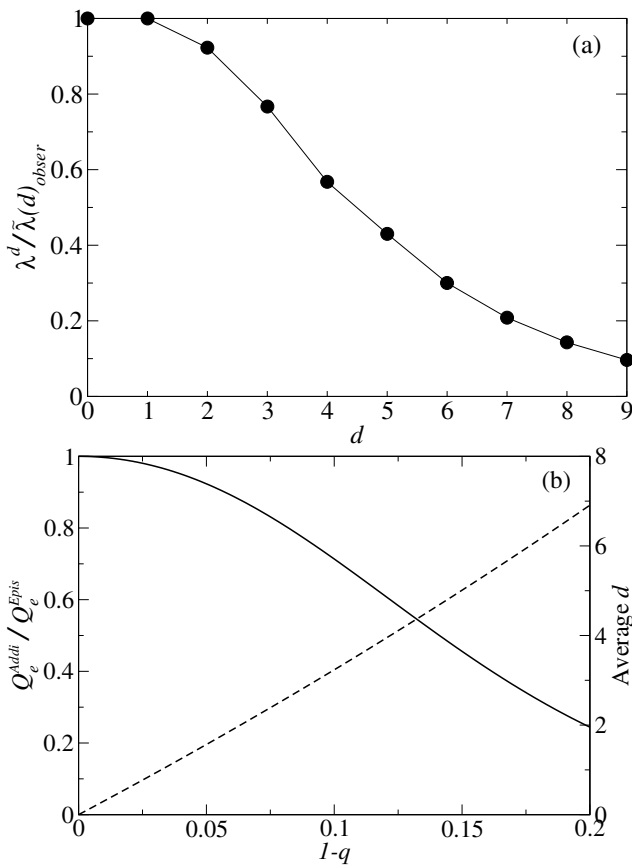
Figure 6
Comparison between additivity and epistasis in RNA folding (a) The *relative* probability of neutral replication under the additive assumption (λ^d) is plotted against the number of base substitutions (*d*), where the probability of neutral replication with epistasis (i.e. the fraction of neutral mutants observed in the yeast tRNAphe folding) is set to be 1 for each *d* as a reference. It can be seen that the effect of epistasis on the probability of neutral replication becomes larger as the number of substitution (*d*) increases. The same data as that of Fig. 5 are used. **(b)** The solid line (the left *y*-axis) represents the *relative* effective replication accuracy (Q_e) under the additive assumption plotted against the error rate (1 - *q*), where Q_e calculated with epistasis is set to be 1 as reference (see Methods section – Probabilistic approach – for details). It can be seen that the effect of epistasis on Q_e increases as the error rate (1 - *q*) increases. The shape of the curve is in a similar manner as that of the curve in Fig 6a. Although the *x*-axis of Fig 6b is different from that of Fig 6a, one can relate the two graphs via the average *d* per neutral replication, with which the different *x*-axes can be transformed to each other. The average *d* per neutral replication is represented by the dashed line (the right *y*-axis plotted against 1 - *q*). Its value is calculated under the additive assumption as $d = Np/(q + p)$ where $p = \lambda (1 - q)$, $\lambda = 0.4$ and $N = 76$.

replication is ca. 1.5 at the error threshold, which is much smaller than the average *d* per replication (ca. 3.8). This means that the main contribution to Q_e under the additive assumption is from the mutants of *d* = 1 or 2 at the error threshold. According to Fig. 6a the additive assumption is a good approximation at *d* = 1 or 2. Therefore, the additive assumption accurately estimates Q_e and thus the error threshold. When the average *d* per neutral replication reaches 3, the additive assumption substantially underestimates Q_e (see Fig. 6b), which is consistent with Fig. 6a. [However, note that this analysis does not imply that the average number of substitutions per replication is safely assumed to be near one. On the contrary, in our calculation it was more than one – actually 3.8 – at the error threshold.]

In the above examination of the additive assumption, there are two points which must be examined further: (1) The analysis of epistasis was performed on a yeast tRNA-phe, which comprises the initial population of the RNA evolution simulations, but the results may differ if the analysis is done for a sequence which appears later in the RNA evolution simulations. Thus, we performed the same analysis to a sequence which was chosen from the population of the fittest sequences after the evolution in the simulations (at the 20000th time step). The results, however, did not change our conclusion (data not shown). (2) If the length of sequences is larger, the average *d* per neutral replication may increase, and thus the additive assumption may break down before the error threshold. However, it turns out from the analytical calculation that the average *d* per neutral replication at the error threshold decreases as *N* increases when λ is invariant (cf. the caption of Fig. 6b). Furthermore, λ decrease as *N* increases (see the filled circles in Fig. 2). Therefore, if the sequence length is larger, the average *d* per neutral replication will be actually smaller. We also conducted computer simulations of RNA evolution with a longer sequence length (200 bases). The results showed that the average *d* per neutral replication (calculated under the additive assumption) at the error threshold was indeed smaller (ca. 1.2 substitutions with $\lambda \approx 0.35$) than in the previous case of the shorter sequence length (ca. 1.5 substitutions with $\lambda \approx 0.4$), and the additive assumption still predicts the results of the simulations closely (data not shown).

Conclusions
• The phenotypic error threshold was formulated under the additive assumption. The formulation asserted that mutational neutrality increases the error threshold but the increase is limited.

• The importance of considering multiple substitutions per replication at the error threshold was illustrated.

• The comparison with the computer simulations and the analysis of epistasis showed that the additive assumption correctly estimates the effective replication accuracy (Q_e) and thus the error threshold.

• The reason why the additive assumption achieves a good prediction of the error threshold in spite of a high degree of (non-trivial) epistasis in RNA folding is that the average number of substitutions per *neutral* replication is small enough to avoid of the effect of epistasis.

Methods

Non-uniform distribution of λ

If λ is not uniform over the genotypes sharing the same phenotype, the effective replication accuracy (Q_e) depends on the distribution of the genotypes in the population. In this case, Q_e is calculated under the additive assumption as

$$Q_e = \sum_{i \in S_I} \left\{ q + \lambda_i (1-q) \right\}^N x_i / X_I \qquad (6)$$

where X_I denotes the population of the focal phenotype (I), and x_i (resp. λ_i) is the population (resp. the fraction of neutral substitutions) of the genotype i. The set S_I denotes the set of genotypes which have the phenotype I. If x_i and λ_i are known, the representative λ of the phenotype can be calculated from the following equation as

$$\sum_{i \in S_I} \left\{ q + \lambda_i (1-q) \right\}^N x_i / X_I = \left\{ q + \lambda (1-q) \right\}^N. \qquad (7)$$

The difference between the representative λ and the population average of λ was very small in the computer simulations. (The population average was always slightly smaller [ca. 99%] than the representative λ unless the distribution of λ in the fittest population is completely homogeneous [data not shown].)

Calculation of Q_e with epistasis

Trivial epistasis in RNA folding

It is trivial that epistasis occurs between bases which make a pair (hydrogen bond) in the reference secondary structure. Our first trial to include epistasis in the calculation of Q_e was to include this epistasis as a part of the additive effects of mutations as described in [5]. In this approach, the reference sequence is subdivided into non-paired regions and paired regions; paired regions are treated as strings of base pairs (one pair of bases is considered as one character); a substitution of a base pair is considered as an elementary step of mutations in paired regions. Following this procedure, the epistasis occurring between bases in a pair is now treated as an additive effect. [For example, two mutations – GC→GG and GC→CC – occurring in paired base must be deleterious because the

bases can not make a pair any more. Given that the combined mutation – GC→CG – is neutral, it will be a case of positive epistasis in the previous procedure. However, in the new procedure it will be a case of additive neutral because the combined mutation is treated as one substitution of a base pair.] We categorized the mutants into the previously defined four groups of mutants (i.e. additive neutral, additive deleterious, positive epistasis and negative epistasis) using the same data as that of Fig. 5. However, the result did not differ much from that shown in Fig. 5 (data not shown). We conclude that epistasis occurs mainly in a non-trivial way, and thus this approach is not effective for our purpose.

Probabilistic approach

Since (non-trivial) epistasis makes it difficult to predict what happens to the phenotype given a specific change in genotype, we take the following probabilistic approach: We assume that a mutant is neutral with a certain probability (denoted by $\mu(v, \delta)$), which depends on the number of neutral base substitutions (denoted by v) and on that of deleterious base substitutions (denoted by δ). Then, the probability of neutral replication is obtained (by using the binomial approximation) as

$$\tilde{\lambda}(d) = \sum_{v=0}^{d} \binom{d}{v} \lambda^v (1-\lambda)^{\delta} \mu(v, \delta) \qquad (8)$$

where $d = v + \delta$. Q_e is thereupon derived as

$$Q_e = \sum_{d=0}^{N} \binom{N}{d} q^{N-d} (1-q)^d \tilde{\lambda}(d). \qquad (9)$$

We measured $\mu(v, \delta)$ in the tRNAphe folding as shown in Fig. 7ab. When $\delta = 0$ and $v > 0$, μ declines a little slower than exponentially as v increases due to negative epistasis (Fig. 7a). When $v = 0$ and $\delta > 1$, μ is not zero due to positive epistasis, and μ decreases slower than exponentially as δ increases (Fig. 7b). When $v > 0$ and $\delta > 0$, $\mu(v, \delta)$ increases, saturates, and finally decreases as v increases (Fig. 7a): neutral substitutions can compensate deleterious substitutions. We express the above observations as follows:

$$\mu(v, \delta) =
\begin{cases}
1 & (v = 0, \delta = 0) \\
(1 - \varepsilon_n)^{(v-1)^\alpha} & (v > 0, \delta = 0) \\
0 & (v = 0, \delta = 1) \\
\varepsilon_d^{\delta^\eta} & (v = 0, \delta > 1) \\
(1 - \varepsilon_n)^{(v-1)^\alpha} (1 + \varepsilon_{nd})^{\frac{\beta v}{\gamma + v}} \varepsilon_d^{\delta^\eta} & (v > 0, \delta > 0)
\end{cases} \qquad (10)$$

where ε_n, ε_d and ε_{nd} are the epistatic parameters of the interactions among neutral substitutions, among
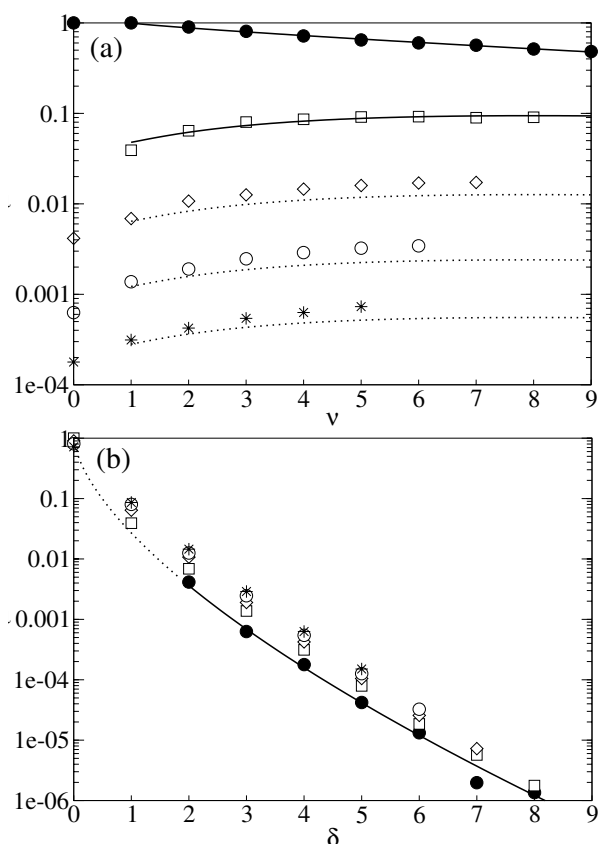
Figure 7
Probabilistic approach in calculating the effective replication accuracy with epistasis (a) The probability that a mutant is neutral with v neutral substitutions and δ deleterious substitutions (i.e., $\mu(v, \delta)$) is plotted against the number of neutral substitutions (v). Symbols: ● $\delta = 0$; □ $\delta = 1$; ◇ $\delta = 2$; ○ $\delta = 3$; * $\delta = 4$. The plots were obtained from the same data set as that of Fig. 5. The solid lines represent the results of curve fitting. We used Eq. 10 ($v>0$, $\delta = 0$) to the $\delta = 0$ data set, and Eq. 10 ($v>0$, $\delta>0$) to the $\delta = 1$ data set. The second fitting was done after we obtained α and ε_n from the first fitting, and ε_d and η from the fitting in Fig. 7b. The dotted lines are the estimation made with the obtained parameters (listed below). **(b)** $\mu(v, \delta)$ plotted against the number of deleterious substitutions (δ). Symbols: ● $v = 0$; □ $v = 1$; ◇ $v = 2$; ○ $v = 3$; * $v = 4$. The solid part of the line represents the curve fitting; the dotted part is an exception, i.e., $\mu(0, 1) = 0$. We used Eq. 10 ($v = 0$, $\delta>0$) toward the $v = 0$ data set in the fitting. All the fitting was done after transforming both the equations and the data sets to logarithmic scale to reduce the biased importance of the points in small d. The obtained parameters are as follows: $\varepsilon_n = 0.1190$, $\alpha = 0.8483$, $\varepsilon_{nd} = 2.418$, $\beta = 2.333$, $\gamma = 3.996$, $\varepsilon_d = 0.02697$ and $\eta = 0.6380$.

deleterious substitutions, and between neutral and deleterious substitutions, respectively. Note that in the additive assumption, all epistatic parameters are zero. α and η represent non-exponential decay. To express the compensation by neutral substitutions, we arbitrarily used a saturation function $\beta v/(\gamma + v)$ where β and γ are parameters. To obtain the parameters, we fitted Eq. 10 to the data in Fig. 7ab (the solid lines) as explained in the caption. As shown in Fig. 7a (the dotted lines), the theoretical estimation turns out to be a slight underestimation.

$\tilde{\lambda}$ (d) was calculated from the above obtained parameters, and the calculated values match the observed ones (see the dashed line in Fig. 5).

Authors' contributions
NT contributed to the entire part of the study. PHP contributed to the computer programing of the RNA evolution simulation. PH contributed to the conceptual development of the study and the manuscript preparation as the supervisor.

References

1. Eigen M: **Selforganization of Matter and the Evolution of Biological Macromolecules.** *Naturwissenschaften* 1971, **58:**465-523.
2. Maynard Smith J, Szathmary E: **Chaper 4.** In *The Major Transitions in Evolution* reprint edition. New York: Oxford University Press; 1997:41-58.
3. van Nimwegen E, Crutchfield JP: **Metastable Evolutionary Dynamics: Crossing Fitness Barriers or Escaping via Neutral Paths?** *Bull Math Biol* 2000, **62:**799-848.
4. Huynen MA, Stadler PF, Fontana W: **Smoothness within ruggedness: the role of neutrality in adaptation.** *Proc Natl Acad Sci USA* 1996, **93:**397-401.
5. Reidys C, Forst CV, Schuster P: **Replication and mutation on neutral networks.** *Bull Math Biol* 2001, **63:**57-94.
6. Wilke CO: **Selection for fitness versus selection for robustness in RNA secondary structure folding.** *Evolution* 2001, **55:**2412-2420.
7. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P: **Fast Folding and Comparison of RNA Secondary Structures.** *Monatsh Chem* 1994, **125:**167-188.
8. Fontana W, Schuster P: **A computer model of evolutionary optimization.** *Biophys Chem* 1987, **26:**123-47.
9. Eigen M, McCaskill J, Schuster P: **The molecular quasi-species.** *Adv Chem Phys* 1989, **75:**149-263.
10. Nowak M, Schuster P: **Error Thresholds of Replication in Finite Populations Mutation Frequencies and the Onset of Muller's Ratchet.** *J Theor Biol* 1989, **137:**375-395.
11. van Nimwegen E, Crutchfield JP, Huynen M: **Neutral evolution of mutational robustness.** *Proc Natl Acad Sci USA* 1999, **96:**9716-9720.
12. Huynen MA, Konings DA, Hogeweg P: **Multiple coding and the evolutionary properties of RNA secondary structure.** *J Theor Biol* 1993, **165:**251-267.
13. Wilke CO, Lenski RE, Adami C: **Compensatory mutations cause excess of antagonistic epistasis in RNA secondary structure folding.** *BMC Evol Biol* 2003, **3:**3.
14. Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP: **RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension.** *Science* 2001, **292:**1319-1325.