

7-1-2000

## Replicability and Recurrence in the Experimental Evolution of a Group I Ribozyme

Martin M. Hanczyc  
*Yale University*

Robert L. Dorit  
*Yale University*, rdorit@smith.edu

Follow this and additional works at: [https://scholarworks.smith.edu/bio\\_facpubs](https://scholarworks.smith.edu/bio_facpubs)



Part of the [Biology Commons](#)

---

### Recommended Citation

Hanczyc, Martin M. and Dorit, Robert L., "Replicability and Recurrence in the Experimental Evolution of a Group I Ribozyme" (2000). Biological Sciences: Faculty Publications, Smith College, Northampton, MA. [https://scholarworks.smith.edu/bio\\_facpubs/70](https://scholarworks.smith.edu/bio_facpubs/70)

This Article has been accepted for inclusion in Biological Sciences: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact [scholarworks@smith.edu](mailto:scholarworks@smith.edu)

# Replicability and Recurrence in the Experimental Evolution of a Group I Ribozyme

Martin M. Hanczyc\* and Robert L. Dorit†

\*Department of Genetics and †Department of Ecology and Evolutionary Biology, Yale University

In order to explore the variety of possible responses available to a ribozyme population evolving a novel phenotype, five *Tetrahymena thermophila* group I intron ribozyme pools were evolved in parallel for cleavage of a DNA oligonucleotide. These ribozyme populations were propagated under identical conditions and characterized when they reached apparent phenotypic plateaus; the populations that reached the highest plateau showed a near 100-fold improvement in DNA cleavage activity. A detailed characterization of the evolved response in these populations reveals at least two distinct phenotypic trajectories emerging as a result of the imposed selection. Not only do these distinct solutions exhibit differential DNA cleavage activity, but they also exhibit a very different correlation with a related, but unselected, phenotype: RNA cleavage activity. In turn, each of these trajectories is underwritten by differing genotypic profiles. This study underscores the complex network of possible trajectories through sequence space available to an evolving population and uncovers the diversity of solutions that result when the process of experimental evolution is repeated multiple times in a simple, engineered system.

## Introduction

The daunting task facing evolutionary theory is to simultaneously account for the two salient features of the natural world: adaptation and diversity. In the years since the publication of *The Origin of Species* (Darwin 1859) biologists have relied on detailed, comparative examination of organisms to explain the multiplicity of forms inhabiting the planet and the apparent close fit between these forms and their environments. As powerful as these retrospective approaches have been, they are necessarily limited to the examination of realized forms (the “actual,” in Jacob’s [1982] description of the domain of biology as “the possible and the actual”).

The frustration with retrospective approaches, coupled with the advent of molecular tools in evolutionary biology, has led to an increasing number of experimental studies of microevolution using both viral (e.g., Bull et al. 1993, 1997; Cunningham et al. 1997) and bacterial systems (e.g., Lenski and Travisano 1994; Nakatsu et al. 1998; de Visser et al. 1999).

This study takes an explicitly experimental approach to understanding molecular adaptation and diversification. It belongs to the expanding body of work that uses simplified molecular systems to address evolutionary questions (Ellinger, Ehrlich, and McCaskill 1998; Hanczyc and Dorit 1998). Using a deliberately simplified experimental system consisting of catalytic RNA molecules (ribozymes) evolving in controlled in vitro environments, we probed the trajectories that lead to novel catalytic function. Our previous work dealt with the emergence of complexity and divergent function in a similar experimental system (Hanczyc and Dorit 1998). Here, we focus on the recurrence of particular outcomes when replicate molecular populations, identi-

cal at the outset, are selected in parallel for the same catalytic function.

In our system, evolving catalytic RNA molecules acquire the ability to cleave a novel substrate (DNA) in response to the experimental conditions we impose. Those ribozymes best able to efficiently cleave this DNA substrate are preferentially replicated. By repeatedly supplying variation to the system and reiterating the selection cycle, we improve the mean DNA cleavage activity of the population. As selectable variance is exhausted, the rate of improvement of the population for the selected phenotype slows and eventually asymptotes: the population reaches an activity plateau and shows no further improvement in the selected activity. In principle, only those ribozymes best able to cleave the DNA substrate will be represented in the population. The individual ribozyme variants present at the end of the experiment embody an adaptive solution to the imposed catalytic challenge.

The evolving entity in this study is the L-21 ribozyme, derived from the self-excising group I intron located in the 23S rRNA gene of *Tetrahymena thermophila*. This ribozyme readily cleaves an RNA oligonucleotide that is partially complementary to the 5' end of the ribozyme (Zaug and Cech 1986;  $k_{\text{cat}}/K_m = 1.6 \times 10^7/\text{M}/\text{min}$ ; see *Materials and Methods*). This same ribozyme can also cleave a DNA oligonucleotide counterpart, but it does so one million times less efficiently (Herschlag and Cech 1990;  $k_{\text{cat}}/K_m = 3.6 \times 10^1/\text{M}/\text{min}$ ). This low but latent ability of the L-21 ribozyme to cleave DNA, coupled with the results of previous studies that demonstrated the selectability of the DNA-cleaving phenotype (Beaudry and Joyce 1992; Tsang and Joyce 1994), made this ribozyme the ideal choice for this study.

This study explores the replicability and variety of adaptive responses that evolve under well-defined and concerted directional selection. We established replicate, independent populations drawn from the same initial variable ribozyme pool and selected these populations in parallel for improved DNA cleavage. Single solutions were thus likely to be fixed in each of the populations,

Key words: experimental evolution, group I ribozyme, catalytic RNA, directed evolution, adaptive landscape, evolution in vitro.

Address for correspondence and reprints: Robert L. Dorit, Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, Connecticut 06511. E-mail: rdorit@beagle.biology.yale.edu.

*Mol. Biol. Evol.* 17(7):1050–1060. 2000

© 2000 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

but these endpoints—and the trajectories that give rise to them—could then be compared among populations. In effect, each of the five evolving populations undertakes an excursion along the functional landscape. The extent to which these independent populations converge on a single outcome, or arrive instead at different functional solutions, sheds light on the structure and topography of the underlying functional landscape.

## Materials and Methods

### Mutagenesis of L-21 Ribozyme

We amplified 20 fmol of the pT7L-21 plasmid containing the L-21 ribozyme sequence (courtesy of T. Cech) via mutagenic PCR (Cadwell and Joyce 1992) with an upstream primer (TAS2.1) containing a T7 promoter and the first 20 nt of the ribozyme (positions 22–41: 5'-CTGCAGAATTCTAATACGACTCACTATAGGAGGGAAAAGTTATCAGG-3') and a downstream primer (T13) complementary to the 3' end of the ribozyme (positions 378–414: 5'-CGAGTACTCCAAAAC-TAATCAATATACTTTTCGCATAC-3'). PCR was performed for 30 cycles under the following conditions: 90°C for 1 min, 60°C for 30 s, and 72°C for 1 min. Initial mutagenesis was carried out in five consecutive rounds of amplification. Resultant molecules were purified by standard phenol/chloroform extraction and subsequent ethanol precipitation after each round of amplification. Approximately 1/1,000 of the resulting PCR product was used as template for the next round of amplification. Amplification products were gel-purified between the third and fourth rounds of amplification as follows: 200 ng of third-round mutant PCR product was transcribed overnight in the presence of 25 U of T7 RNA polymerase at 37°C in a reaction containing 30 mM Tris (pH 8.0), 25 mM MgCl<sub>2</sub>, 20 mM spermidine, 25 mM DTT, 0.1% Triton X-100, and 2 mM each NTP; the mutant transcripts were resolved on a 5% acrylamide/8 M urea gel, and the full-length transcripts were cut out of the gel. The ribozymes were eluted from the gel fragments overnight by immersing the crushed gel slice in a buffer containing 200 mM NaCl<sub>2</sub>, 10 mM Tris (pH 7.5), and 0.5 mM EDTA and subsequently recovered by ethanol precipitation. Then, 340 fmol of the purified transcripts were reverse-transcribed with 5 U of AMV reverse transcriptase at 42°C for 45 min in a reaction containing 25 mM Tris (pH 7.5), 5 mM MgCl<sub>2</sub>, and 25 mM DTT and subjected to two additional rounds of mutagenic PCR. After the last round of mutagenesis, the mutant pool was amplified by PCR with TAS2.1 and T13 primers under nonmutagenic conditions as described in Hanczyc and Dorit (1998). Three additional rounds of mutagenic PCR, using the above procedure, were carried out after the fourth generation of evolution *in vitro*.

### Construction of Independently Evolving Ribozyme Lines

The variant pool of ribozymes created by mutagenic PCR was transcribed by T7 RNA polymerase, and the transcripts were gel-purified as described above. From

this pool of mutant transcripts, five samples, each containing approximately 10<sup>13</sup> RNA molecules, were used to initiate the five independent evolving lines (lines A, B, C, D, and E).

### Evolution *In Vitro*

The evolution *in vitro* protocol is outlined below and is based on a procedure originally devised by Joyce and coworkers (Beaudry and Joyce 1992; Lehman and Joyce 1993) and subsequently modified as described in Hanczyc and Dorit (1998) with the following changes: Ribozymes were transcribed *in vitro* by T7 RNA polymerase from PCR templates, gel-purified, and quantitated by spectrophotometry, and approximately 10<sup>13</sup> ribozyme molecules were added to the cleavage reactions. Two biotin molecules (“B”) are covalently linked to the 3' end of the DNA substrate (STC) (5'-GGCCCTCTAAATAAATAAATAAATAAACAAACAAA-BB-3'). Upon cleavage, the 3' moiety of the DNA oligonucleotide substrate (5'-AAATAAATAAATAAATAAACAAACAAA-BB-3') covalently attaches to the ribozyme. Catalytically active ribozymes thus end up as tagged, biotinylated RNA-DNA chimeras which can be captured on streptavidin-coated magnetic beads. A second stage of selection used a reverse-transcription primer, complementary to the tag-ribozyme junction, further ensuring that only active ribozyme variants remained in the evolving population.

The cleavage reaction (1 μM ribozyme, 10 μM DNA oligonucleotide substrate, 10 mM MgCl<sub>2</sub>, and 30 mM EPPS [pH 7.5]) proceeded at 37°C for 180 min in generations 1–2. This cleavage time was reduced to 10 min in generations 3–6, and to 5 min in generations 7–15, effectively increasing the intensity of functional selection for improved cleavage. Following cleavage, the reaction products were incubated (80°C for 30 s) with streptavidin-coated Dynabeads (M-280, Dynal) in order to bind and retain the successful ribozymes. The beads and bound ribozymes were then washed twice with heated (80°C) SDS buffer (10 mM Tris [pH 7.5], 0.15 M LiCl, 1 mM EDTA, 0.3% SDS) and twice with heated (80°C) reverse-transcriptase buffer (25 mM Tris [pH 7.5], 5 mM MgCl<sub>2</sub>, 25 mM DTT) in order to remove the unbound ribozymes and prepare the bound ribozymes for reverse transcription. The retained ribozymes were amplified from the beads with the primer (TAG1) 5'-TTTGTGTTTATTATTATTTA-TTTATTTC-3' and AMV reverse transcriptase. The resultant cDNAs were amplified by PCR. No isothermal amplification was required.

### Construction and Propagation of the Control Line

L-21 ribozymes were transcribed from PCR product and gel-purified as described above. The control line, begun with approximately 10<sup>13</sup> molecules of L-21 ribozyme transcripts, was propagated through the steps of evolution *in vitro* as outlined above, with two exceptions: no mutagenic PCR was performed on the control line, and the selection step was modified such that successful DNA cleavage was not a precondition for prop-

agation. Specifically,  $10^{13}$  molecules of the control line were added to a DNA cleavage reaction, and cleavage was allowed to proceed. Following the cleavage reaction, ribozymes were added directly to the reverse-transcription step and subsequently amplified in a reaction containing both T13 and TAG1 primers. The molecules of the control line are therefore replicated independently of their ability to cleave the DNA substrate.

#### Cloning of Individual Ribozymes

PCR products to be further characterized were cloned into linear T-tailed pBlueScript vector by T4 DNA ligase under standard conditions. The ligated plasmids were transformed into competent cells and grown on ampicillin plates containing X-gal, as described in Hanczyc and Dorit (1998). Overnight cultures were prepared from a random sampling of single white transformant colonies. Ligated plasmids containing a PCR insert were isolated from overnight cultures using Qiagen miniprep protocols.

#### Sequencing of Individual Ribozymes

Ribozyme variants were sequenced from plasmid templates by dye-terminator cycle sequencing according to Perkin Elmer protocols, and the genotypes were determined (with a minimum twofold coverage) using a 373A ABI DNA sequencer and software.

#### DNA Cleavage Assays

The overall DNA cleavage activities of the evolving and control populations were assayed at each generation. In addition, the activity of selected clones was further characterized as follows: 0.5  $\mu\text{M}$  ribozyme was preincubated at 50°C in 50 mM  $\text{MgCl}_2$  and 150 mM EPPS (pH 7.5) in a 20- $\mu\text{l}$  reaction. 0.1  $\mu\text{M}$  5'  $^{32}\text{P}$ -labeled STC substrate was added to the reaction and incubated at 37°C for 60 min. The products of the reaction were resolved on a 16% acrylamide, 8 M urea gel, visualized using a phosphorimager (Fujix BAS 2000), and quantitated using MacBAS, version 2.0.

#### RNA Cleavage Assays

The RNA-cleaving activity of the evolved ribozyme populations was determined using a 5'  $^{32}\text{P}$ -labeled substrate (RSTC) (5'-GGCCCUCUAAAUAAAUAAAUAAAUAAACAA-ACAAA-BB-3'). Then, 4 nM ribozyme was preincubated at 50°C in 50 mM  $\text{MgCl}_2$  and 150 mM EPPS (pH 7.5) in a 20- $\mu\text{l}$  reaction. An equivalent of 1 nM labeled RNA substrate was added to the reaction and incubated at 37°C for 1 min. The reaction was stopped with 2 volumes of 7 M urea and 20 mM EDTA with 0.1% xylene cyanol and 0.1% bromophenol blue, and the reaction products were resolved in a 16% acrylamide/8 M urea gel and quantitated as described above.

#### Determination of $k_{\text{cat}}$ and $K_m$ Values

Assays with the 5' end-labeled RNA substrate were carried out under single-turnover conditions with a sub-

strate concentration at least 10-fold below  $K_m$ . These assays also involved using four different ribozyme concentrations, which spanned the  $K_m$  and always exceeded the concentration of substrate. Ribozymes were preincubated at 50°C in 50 mM  $\text{MgCl}_2$  and 150 mM EPPS (pH 7.5), and the labeled RNA substrate was subsequently added to the reaction. Aliquots (typically 2  $\mu\text{l}$  each) were taken at three time points during the linear portion of the reaction and quenched with 2 volumes of 7 M urea and 20 mM EDTA with 0.1% xylene cyanol and 0.1% bromophenol blue. Reaction products were resolved and quantitated as described above. The  $k_{\text{cat}}/K_m$  values were determined from  $k_{\text{obs}}$  values according to the equation  $k_{\text{obs}} = (k_{\text{cat}}/K_m)[E]$ .

Assays with the 5' end-labeled DNA substrate were performed under multiple-turnover conditions using five different excess substrate concentrations. The ribozymes were preincubated at 50°C in 50 mM  $\text{MgCl}_2$  and 150 mM EPPS (pH 7.5), and the labeled DNA substrate was subsequently added to the reaction. Aliquots were extracted at four time points during the linear portion of the reaction and were resolved and quantitated as described above. Rate constants ( $K_m$  and  $k_{\text{cat}}$ ) were determined using Lineweaver-Burk plots.

#### Determination of $k_{\text{obs}}$

The kinetic parameter  $k_{\text{obs}}$  was estimated for both DNA and RNA substrates from a sampling of individual ribozyme clones in single-turnover reactions. The conditions were modeled after those used above for the determination of  $k_{\text{cat}}$  and  $K_m$ . Time points were taken during the linear portion of the reaction, and  $k_{\text{obs}}$  values were estimated as the absolute value of the slope of the natural log (ln) of substrate remaining over time.

## Results

### Phenotypic Response of Evolved Populations

The responses of all populations to the imposed selection for DNA substrate catalysis are shown in figure 1. All evolving populations show marked increases in their ability to cleave the DNA oligonucleotide, with the most active lines (B, C, D, and E) showing a 100-fold improvement in DNA cleavage relative to the wild-type ribozyme.

After an initial response to the selection (generations 1–4; see fig. 1), the pools were subjected to three more rounds of mutagenic PCR (estimated error rate of 5.4% per position). As these variant pools began responding to selection, the amount of time allotted to the cleavage reaction was reduced from 180 min to 10 min and, finally, to 5 min. This marked increase in the intensity of selection reduced the noise in the experiment and insured that only the highest-activity variants were selected. The experiment was terminated after generation 15, when three of the five pools had reached phenotypic plateaus (no phenotypic improvement for at least three consecutive generations).

The dynamics of the selective response differ from one population of molecules to the next (fig. 1). Line D exhibits the fastest response, followed closely by line B,

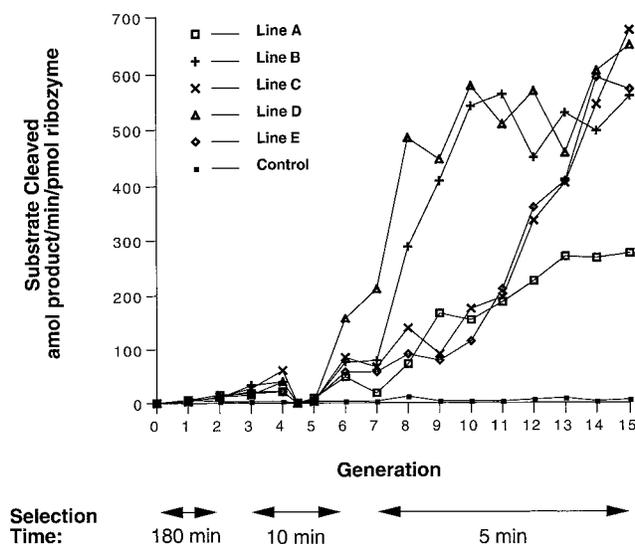


FIG. 1.—Phenotypic response of replicate populations over 15 generations of evolution in vitro. The graph depicts DNA substrate cleaved (expressed as amol cleaved product/min/pmol ribozyme) at every generation. The point located between generations 4 and 5 marks the performance of the lines following additional mutagenesis. The selection intensity, corresponding to the time allowed for cleavage to occur under selection, is shown along the bottom of the figure. All assays were carried out as described in *Materials and Methods*.

with both lines reaching a cleavage plateau at approximately 600 amol of DNA substrate cleaved/min/pmol ribozyme. After generation 11, lines C and E eventually evolve to reach the same level of catalytic activity as lines B and D. The response of lines C and E subsequent to generation 15 has not been determined; they may well continue to improve in subsequent generations. Line A

shows no response after generation 13, reaching a plateau at 300 amol of DNA substrate cleaved/min/pmol ribozyme. As expected, in the absence of selection the WT control shows no improvement in DNA cleavage activity.

Our replicate runs of experimental evolution thus reveal heterogeneity both in the line-specific responses and in the resultant phenotypic plateaus. We obtained a variety of outcomes: similar trajectories (lines B and D), differing trajectories (line B vs. line C, for example), convergent endpoints (lines B and D), and differing endpoints (line A vs. lines B, C, D, and E).

#### Correlations with RNA Catalysis

The evolved ribozymes were further characterized by examining the extent to which improvements in the evolved activity (cleavage of DNA) correlated with cleavage of an RNA substrate (the “ancestral” activity). Previous studies had shown a positive correlation between these two activities in evolved variants of the L-21 ribozyme (Tsang and Joyce 1996), suggesting that increases in DNA activity stemmed largely from overall improvements in nucleic acid cleavage chemistry. A similar pattern emerges in our evolved lines B, C, D, and E: a 100-fold improvement in DNA catalysis is accompanied by a twofold increase in RNA catalysis (fig. 2). In contrast, line A shows an approximately twofold decrease in RNA catalysis relative to the starting L-21 (wild-type) ribozyme, indicating that line A ribozymes have lost some RNA cleavage ability while evolving enhanced DNA cleavage activity. The association between DNA and RNA catalysis was explored in detail by assaying eight clones (sampled independently of their

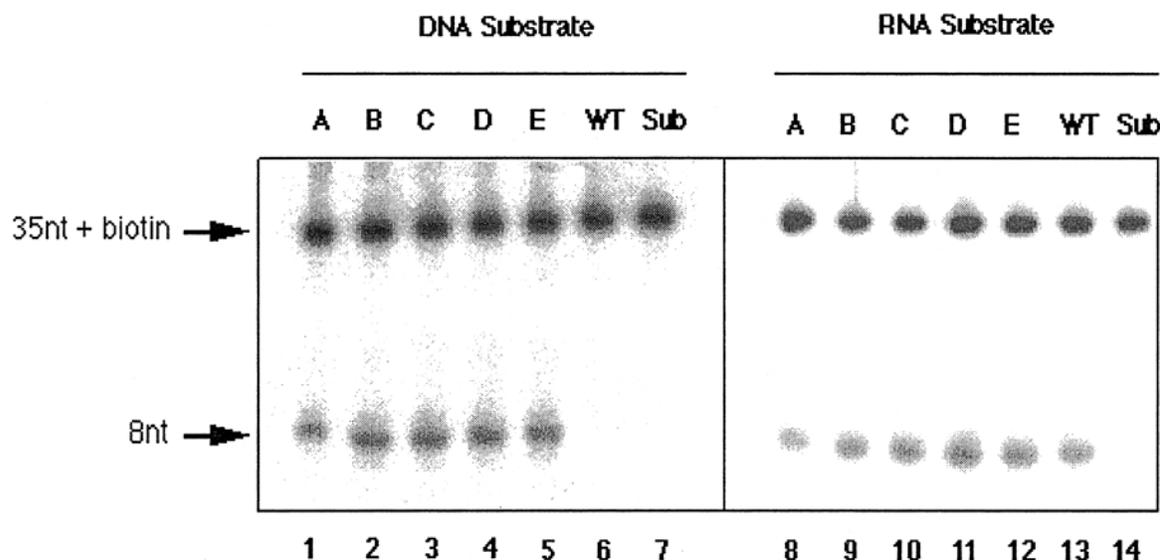


FIG. 2.—Activity of evolved lines on DNA and RNA substrates. Populations of ribozymes after 15 generations of evolution in vitro were incubated with a radioactively labeled, biotinylated DNA or RNA substrate as described in *Materials and Methods*. The products of the reactions were resolved on a denaturing gel and evaluated by phosphorimaging. The upper band represents the 35-nt uncleaved, biotinylated substrate, and the lower band represents the 8-nt product of the cleavage reaction. Lanes 1–7 show the products of cleavage reactions using a DNA substrate incubated with the evolved lines A, B, C, D, and E (lanes 1–5) or the L-21 (WT) ribozyme (lane 6). A substrate-only DNA control (Sub) is shown in lane 7. Lanes 8–14 show the products of cleavage reactions using an RNA substrate incubated with the evolved lines A, B, C, D, and E (lanes 8–12) or the L-21 (WT) ribozyme (lane 13). A substrate-only RNA control (Sub) is shown in lane 14.

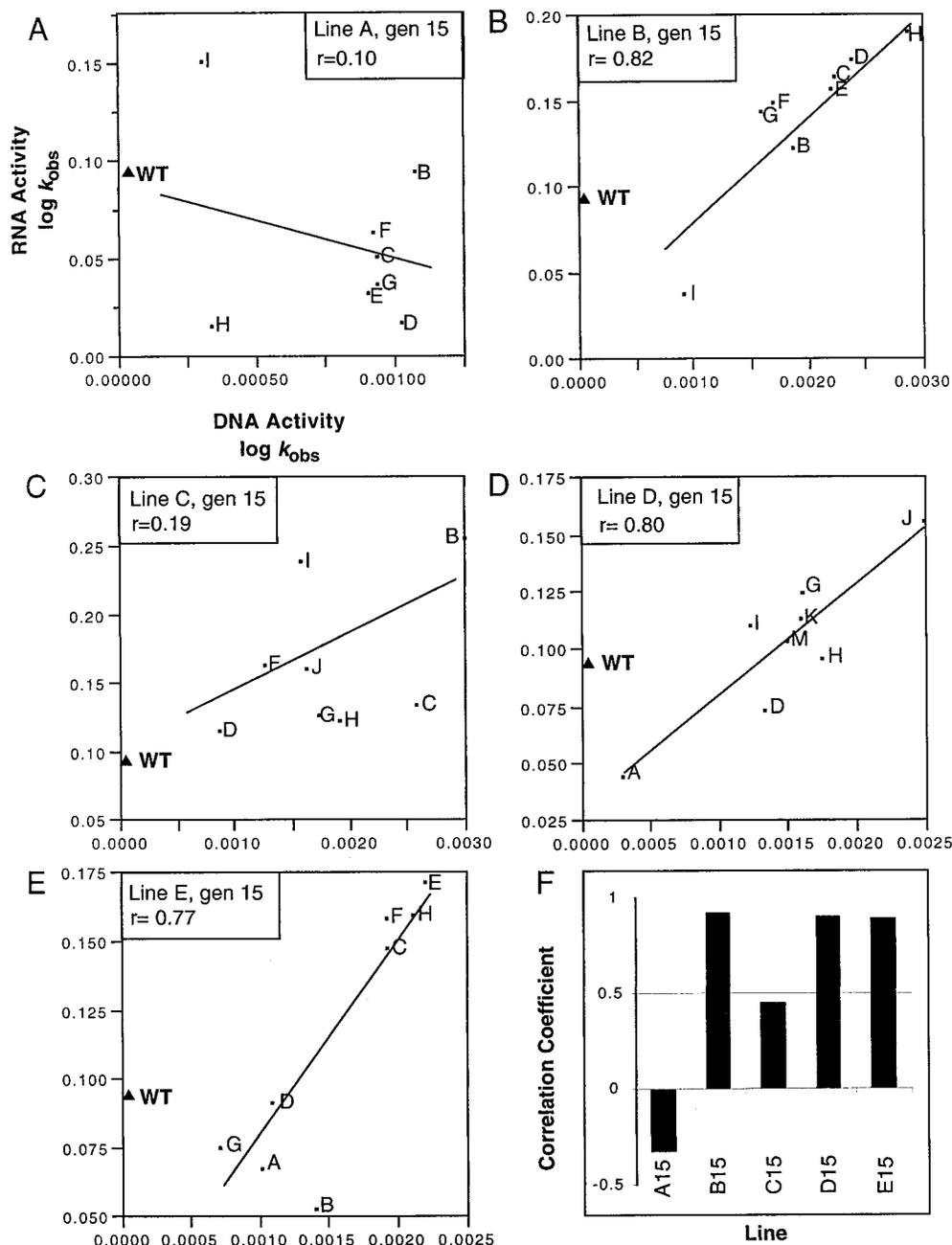


FIG. 3.—Correlation plots of the DNA and RNA cleavage activities of individual ribozymes. In panels A–E, each point represents the activity of an individual ribozyme. The X-axis represents the log-transformed DNA-cleaving activity ( $k_{obs}$ ); the Y-axis is the log-transformed RNA-cleaving activity ( $k_{obs}$ ). The activity of the wild-type (L-21) ribozyme is shown, for comparison only, as a closed triangle. Each line represents a regression across all of the evolved ribozymes sampled within a population, with the fit shown as the  $r$  value (the WT values were not included in the regression or correlation analyses). Panel F summarizes the Pearson product-moment correlation coefficient between DNA and RNA cleavage activity for each of the lines based on the activity of their representative ribozymes.

cleavage phenotype) from each of the five evolved populations. The DNA- and RNA-cleaving abilities of each of these clones are shown in figure 3A–E. As expected from population assays (fig. 2), clone activity on DNA and RNA substrates is correlated in several of the clones. Clones from lines B, D, and E show a high degree of correlation between DNA and RNA catalysis (Pearson product-moment analysis correlations of 0.90, 0.89, and 0.88 respectively; fig 3F). Line C clones still show a weak positive correlation (0.43). In contrast, line

A clones show a weak negative correlation between DNA and RNA catalysis ( $-0.31$ ).

#### Genotypic Characterization of the Evolved Populations

The initial variability present in the ribozyme pool following five consecutive rounds of mutagenic PCR was quantified by sequencing a sample of 25 molecules (data not shown). The resulting error rate of the starting

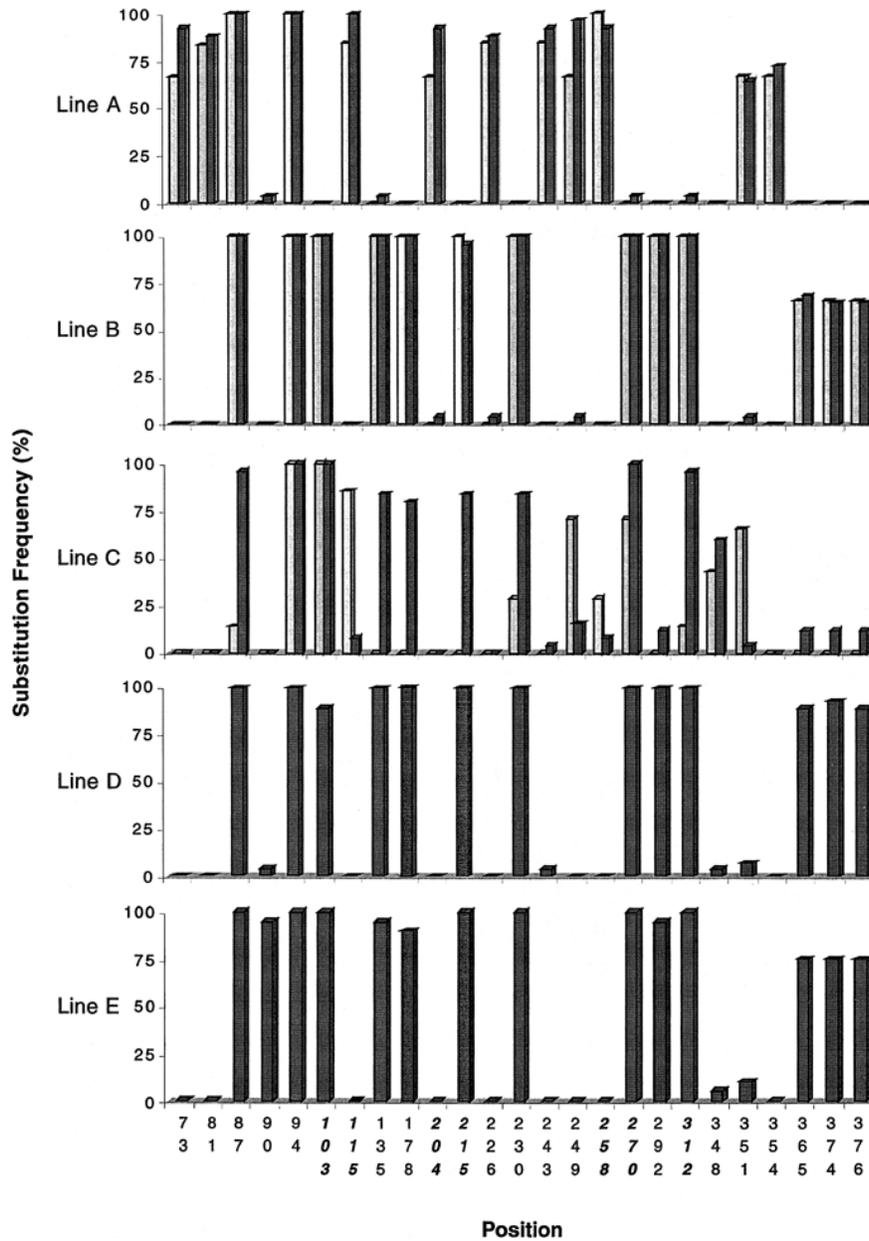


FIG. 4.—The frequencies and positions of prevalent mutations found in each line following 10 (stippled bars, lines A–C) and 15 generations of evolution in vitro (black bars, lines A–E). The sites of prevalent mutations (sites at which more than 50% of the ribozymes sampled from any one line differ from the L-21 ribozyme) are shown along the bottom axis. The Y-axis shows the frequency of a prevalent mutation for each of the lines. Those positions located in the catalytic core of the ribozyme are shown in italicized boldface type. A compilation of the prevalent mutations for the sequenced genotypes is available (<http://www.eeb.yale.edu/faculty/dorit>).

pool was estimated as 9.1% per position, with an average of 30.7 mutations per molecule.

After 15 generations of evolution in vitro, we characterized the five evolved lines by sampling 25 individual clones from each line, with no a priori knowledge of their genotypic or phenotypic characteristics. The resulting genotypes were compared with the ancestral L-21 sequence; we define a genotypic difference as a “prevalent mutation” if more than 50% of the sampled sequences from any one line differ from the L-21 sequence at that position. We identify 25 such prevalent mutations in our evolving lines. The identities of these prevalent mutations and their frequencies across all lines

at generation 15 are summarized in figure 4. In addition, we sampled individual clones from lines A, B, and C at generation 10, when the lines appear most divergent in their response to selection (fig. 4A–C).

By generation 15, lines B, C, D, and E share nine of the prevalent mutations and clearly constitute a genotypic cluster. Lines B and D are genotypically virtually indistinguishable and share 13 of the prevalent mutations. Their genotypic similarity correlates with the observation that both lines B and D responded similarly to in vitro selection and demonstrate comparable phenotypic activities (see fig. 1). Line E shares all 13 of the prevalent mutations observed in lines B and D but is

distinguished by an additional adenine-to-guanine mutation at position 90 ( $f > 0.9$ ). While Line C shares nine prevalent mutations with lines B, D, and E, it is set apart by a common change at position 348, as well as by infrequent mutations at position 292. Line C appears to be the most variable population after 15 generations of evolution: only 3 of the 10 prevalent mutations are represented in 100% of the clones from this line (fig. 4). In contrast, 9 of 13 prevalent mutations were fixed in line B by generation 15.

Line A shares only two prevalent mutations (positions 87 and 94) with the remaining lines; the other 10 prevalent mutations are unique to line A. The distinct character of line A is evident by generation 10 (fig. 4B), when the phenotypic divergence between lines is at its maximum (fig. 1). The eventual genotypic profile of all lines is foreshadowed well before the activity plateaus are reached.

In all five lines, prevalent mutations occur at positions 87 and 94 (see fig. 4). The change at position 87, a deletion of a single adenine in stem P2.1, results in the elimination of a bulged nucleotide. The change at position 94 involves an A→U change in a putatively single-stranded linker region between stems P2.1 and P3. Of the 125 evolved clones sampled in this study, all exhibit the A→U change at position 94, and all but 2 (123/125) have lost the bulged nucleotide at position 87. Both changes have been observed in previous studies (Beaudry and Joyce 1992; Tsang and Joyce 1994, 1996) in which they appear to be associated with bound substrate orientation.

The sequence differences accumulated by the end of the experiment in lines A and B (relative to L-21) are also shown as relative-frequency histograms superimposed on the secondary structure of this ribozyme (fig. 5). The differences occur in both the conserved catalytic core of the ribozyme and the periphery; both position and frequency of changes differ significantly between lines A and B.

### The Control Line

The control line was maintained for two reasons. First, characterization of the control line allows the estimation of the inherent mutation rate of the system due to replication errors. Second, characterization of the control line allows us to distinguish mutations that simply confer a replicative advantage in this experimental system from those emerging in response to the catalytic challenge (Wright and Joyce 1997; Hanczyc and Dorit 1998).

Twenty-four clones from the wild-type control line were analyzed after 15 generations to test for the distribution of mutations along the molecule. The molecule was divided into five equivalent (67-nt) intervals, which insured that no interval exhibited fewer than five mutations in our sample. A test of the observed distribution failed to reject the null hypothesis of equal numbers of mutations (including indels) across all intervals ( $G = 0.769$ ,  $df = 4$ ,  $P > 0.9$ ). We noted a significant excess of transitions (54) over transversions (11) and a signif-

icant number of insertions (13) and deletions (2), as expected from amplification and reverse-transcription errors. We calculated an inherent mutation rate of 0.066% per position per generation (average = 3.3 mutations per clone; range = 0–9) and did not detect any pattern of molecular changes evolving solely in response to the experimental design of the system.

### Kinetic Analyses

Kinetic analyses quantitate the changes in the kinetic parameters that are responsible for the enhanced phenotype seen in figure 1. Kinetic dissection can also reveal further disparities among the evolved lines, since similar improvements in catalytic activity (measured as  $k_{\text{obs}}$ ) can obscure different underlying causes (e.g., increased  $k_{\text{cat}}$  or decreased  $K_m$ ).

Representative molecules from each of the five populations were chosen to insure that the set of characteristic changes for each population was examined. Clone A15.G, for instance, contains all 12 prevalent mutations characteristic of line A. Given the similarity between lines B and D, an additional clone from line D (D15.I), exhibiting a variant 3' end, was also included in this study. The results of the kinetic analysis for all five clones and the wild-type ribozyme are shown in table 1. Both  $K_m$  and  $k_{\text{cat}}$  values changed in response to the selection for DNA catalysis when compared with the L-21 (WT) ribozyme.  $K_m$  values for most of the clones show about a 10-fold decrease (clone A15.G shows only a fourfold decrease in  $K_m$ ). All clones show 20- to 40-fold increases in  $k_{\text{cat}}$  over L-21 (WT). Upon comparing  $k_{\text{cat}}/K_m$  values, clone C15.H shows the highest efficiency, a 330-fold increase over L-21 (WT).

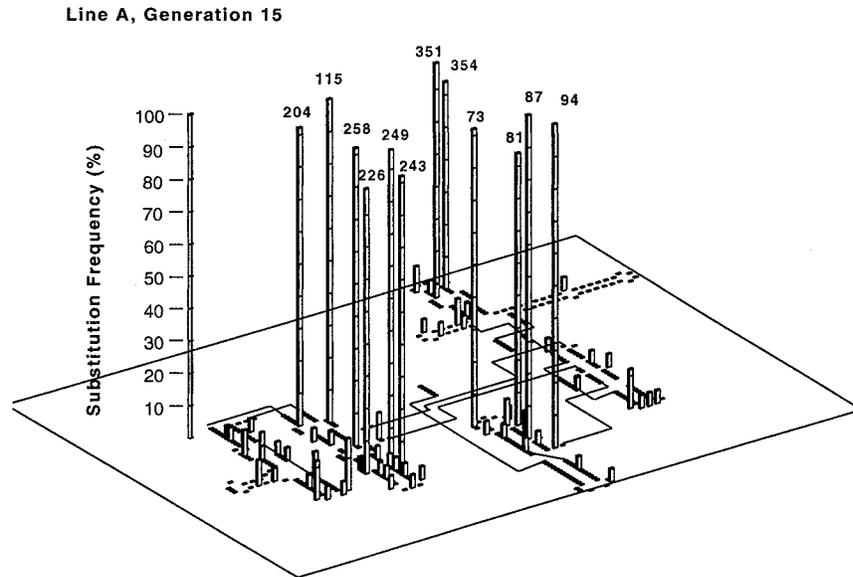
These levels of catalytic improvement are comparable to those in previous studies (e.g., see Tsang and Joyce 1994), despite some significant differences in the design and selective conditions used. Both the  $k_{\text{cat}}$  and the  $K_m$  components are modified by selection.

### Discussion

Parallel populations of ribozymes evolved for improvement in DNA cleavage activity exhibit different phenotypic trajectories and reach different endpoints. These differing endpoints in phenotypic activity (fig. 1) are underwritten by correspondingly different genotypes (fig. 4). The L-21 ribozyme variants explore a minimum of two solution classes in response to the challenge of DNA catalysis—one embodied by line A, the other by lines B, C, D, and E.

We delineate these two solution classes based on the following criteria: (1) the different levels of DNA cleavage activity at which the lines reach a plateau; (2) the differences exhibited by the lines in their RNA cleavage capacities (a related but distinct and unselected phenotype); and (3) the overlapping mutational spectrum of lines B, C, D, and E, in contrast to the distinct mutational profile of line A. Ultimately, the independent character of the two solutions requires an exploration of the genotype space separating them and a demonstration

**A**



**B**

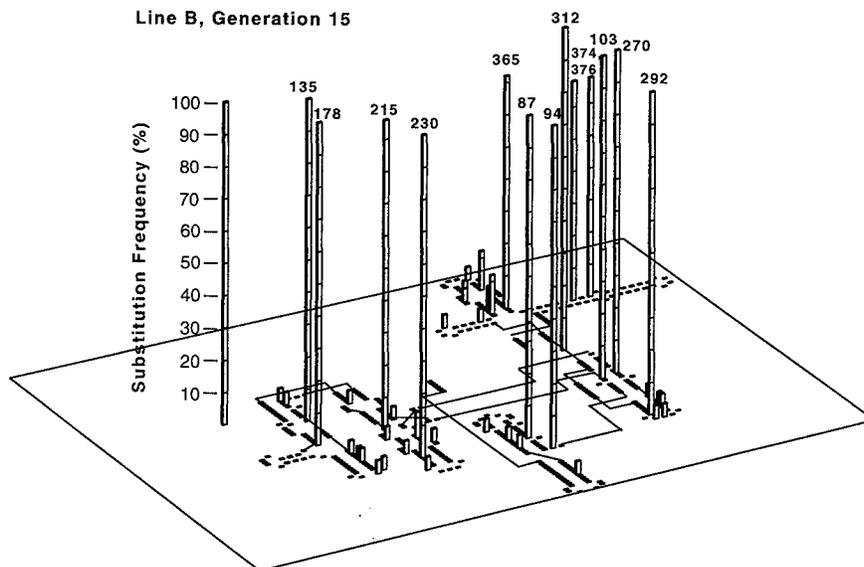


FIG. 5.—Histograms of sequence changes observed in the evolving populations at generation 15. Twenty-five clones were randomly sampled and sequenced from lines A and B at generation 15. Sequence changes (relative to the WT L-21 sequence) which occurred during the course of evolution are shown as vertical bars, superimposed on the secondary structure of the L-21 ribozyme. The height of the bar, ranging from 0% to 100% of sampled ribozymes, denotes the frequency of changes at that position. The number immediately above the bar indicates the positions of the most frequent changes. Panel A shows the frequency and distribution of changes that occurred in line A after 15 generations; Panel B shows comparable information for line B.

**Table 1**  
**Kinetic Analysis of DNA Catalytic Activity for Representative Ribozymes**

Ribozyme	$K_m$ ( $\mu\text{M}$ )	$k_{\text{cat}}$ ( $10^{-2}/\text{min}$ )	$k_{\text{cat}}/K_m$ ( $10^4/\text{M}/\text{min}$ )
L-21 (WT) . . . .	$4.50 \pm 0.80$	$0.16 \pm 0.02$	$0.0037 \pm 0.0003$
A15.G . . . . .	$1.12 \pm 0.53$	$3.0 \pm 0.66$	$0.31 \pm 0.12$
B15.E . . . . .	$0.59 \pm 0.17$	$5.8 \pm 1.4$	$0.99 \pm 0.08$
C15.H . . . . .	$0.58 \pm 0.11$	$6.7 \pm 0.83$	$1.2 \pm 0.11$
D15.I . . . . .	$0.46 \pm 0.16$	$3.0 \pm 0.98$	$0.73 \pm 0.32$
E15.C . . . . .	$0.38 \pm 0.19$	$4.2 \pm 0.7$	$1.2 \pm 0.29$

of the incompatible nature of the mutations that define the two solutions (e.g., see Gavrillets and Gravner 1997).

### The Genotypic Basis of Phenotypic Divergence

The solution class that displays the greatest improvement in the selected phenotype evolved more than once (lines B, C, D, and E). These lines have converged phenotypically and genotypically. In contrast, line A also shows enhanced DNA cleavage activity but embodies a different catalytic solution. The phenotypic divergence of line A is accompanied by a comparable divergence at the genotypic level, already evident by generation 10. Despite the heterogeneity in the phenotypic responses at that point, the eventual distribution of prevalent mutations that distinguishes line A from lines B, C, D, and E takes shape early in the experiment.

The presence of distinct solution classes reveals the interplay of stochastic and deterministic forces. Particular mutations responsible for the high activity of lines B, C, D, and E may not have arisen in line A. Eight prevalent mutations present in lines B, C, D, and E are not observed in line A; these mutations could account for the difference in the catalytic activities of the different lines.

Stochastic effects may also be manifested in the differential accumulation of mildly deleterious mutations in the various lines. The activity of line A may be compromised by line-specific deleterious mutations that diminish catalytic activity. In the absence of recombination, hitchhiking of mildly deleterious mutations will continue until mutation-selection equilibrium is reached. Candidates for this mildly deleterious designation—prevalent mutations unique to line A (positions 73, 81, and 354; see fig. 4)—show no obvious correlation with reduced DNA cleavage activity in the representative clones. Despite the large effective population sizes ( $>10^5$ ) in these experiments, mutations of small deleterious effect ( $s \cong 1/N_e$ ) are nevertheless unlikely to be eliminated from the population.

Deterministic explanations for the existence of divergent solutions postulate epistatic interactions among the observed point mutations. Because catalytic activity depends on both the sequence and the three-dimensional structure of the ribozyme, such epistatic interactions can be reasonably postulated. The enhanced DNA cleavage activity of lines B, C, D, and E may depend on a combination of mutations that has not arisen in line A. Positions 135, 270, and 312 may represent such an interaction. These individual mutations manifest in line A

variants (see fig. 4), yet they do not co-occur in a single ribozyme molecule from line A. The absence of recombination again retards the formation of such favorable epistatic combinations, since multiple mutations must arise on the same molecule before a fitness advantage obtains. Finally, the fate of new mutations may depend primarily on the genetic background in which the mutations occur. This context-dependent fate, arising from the type of epistatic interactions described above, sets the stage for the emergence of divergent genotypic and phenotypic classes. As line A and lines B, C, D, and E begin to define distinct genotypic contexts early in the experiment, the fate of any subsequent new point mutation, even if it appears in all lines, is unlikely to be the same. Every new fixation defines a novel context for subsequent mutational events. By generation 10, line A has acquired a suite of line-specific mutations that differ from those seen in lines B, C, D, and E. The two genotypic solutions seen in this experiment are separated by a large Hamming distance, suggesting that line A is not simply an intermediate stage in the evolution of a single convergent solution embodied by lines B, C, D, and E.

### Correlated Response: Uncoupling DNA and RNA Catalysis

The line-specific correlation between DNA and RNA cleavage activities reveals further heterogeneity in the phenotypic responses of the replicate lines. In our experiments, there is no selection for retaining activity on an RNA substrate or any counterselection for reduced activity on an RNA target. Despite this, lines B, D, and E show a tight positive correlation between the two cleavage activities, a result expected from previous studies (Tsang and Joyce 1994, 1996). Line C shows a weaker, but still positive, correlation. In contrast, DNA cleavage appears uncoupled from RNA cleavage in line A, as reflected by the weak negative correlation shown in figure 3F. This lack of correlation between DNA- and RNA-cleavage activities in line A stands in contrast to previous studies where the correlation between DNA and RNA catalysis could only be weakened by direct counterselection (Tsang and Joyce 1996). Line A has evolved a distinct solution with greater specificity for the DNA substrate: DNA and RNA cleavage are no longer correlated in this line.

### Evolving DNA Cleavage: Necessary and Contingent Changes

Two prevalent mutations, at positions 87 and 94, recurred in every line (fig. 4). In pursuit of the same objective—enhanced cleavage of a DNA substrate—but utilizing different experimental methods, previous studies also report changes at positions 87 and 94 in response to *in vitro* selection (Beaudry and Joyce 1992; Tsang and Joyce 1994). We note that these changes are not seen in our unselected control line and are thus not a bias or artifact of the evolution *in vitro* system.

All of the evolved clones from generation 15 sampled in this study (125) exhibit the change at position

94; all but two clones show the deletion at position 87. The prevalence of these two changes does not per se mean that they are the basis for improved DNA cleavage activity. Clone C15.G, for instance, lacks the change at position 87, yet still shows enhanced DNA cleavage (fig. 3). This deletion is therefore not necessary for improved DNA cleavage. Only one of the 25 prevalent changes described in figure 4, the A→U change at position 94, is truly ubiquitous in our experiment and has previously been observed in similar selections (see Beaudry and Joyce 1992; Tsang and Joyce 1994). While the role played by these mutations has yet to be established, we suspect that the changes at positions 87 and 94 are neither necessary nor sufficient to confer enhanced DNA cleavage activity. They may instead represent easily accessible steps in the evolution of enhanced DNA cleavage, and their fixation may set the stage for the emergence of subsequent functional mutations. The diversity of evolved molecules emerging in these experiments suggests that multiple combinations of genotypic changes can confer an enhancement in DNA-cleaving activity.

The multiple evolutionary trajectories seen in this experiment underscore the many-to-one relationship between ribozyme genotypes and catalytic phenotypes. This relationship is reflected in the phylogenetic diversity of many catalytic RNAs (e.g., group I [Michel and Westhof 1990] or group II introns [Michel and Ferat 1995]). Such phylogenetic comparisons, however, take place over large Hamming distances and compare very divergent entities from varying contexts.

The results presented here suggest that even on a local scale, where common ancestry is recent and still visible, we are dealing with a rugged adaptive landscape where multiple genotypes give rise to a single convergent function. Such ruggedness is a prediction of a number of theoretical models and arises from assumptions about the role of stochastic events in small populations (Wright 1948, 1949), the nature of the mutational spectrum (Gillespie 1984), the prevalence of epistatic interactions (Kaufmann 1995), the character of the genotype-to-phenotype map (Gavrilets 1997), and the nature of RNA free-energy landscapes (Schuster et al. 1994; Hunyen and Fontana 1996; Fontana and Schuster 1998).

This ruggedness has clear implications for the applied *in vitro* evolution of novel molecules. Most evolution *in vitro* is not likely to preserve multiple alternative solutions to a catalytic challenge (but see Eklund, Szostak, and Bartel 1995; Hager and Szostak 1997). In the present case, the solution represented by line A (with its lower overall activity) would not have been recovered had it arisen along with the alternate solution in a single, evolving population. While this would seem to argue for the power of single *in vitro* runs to produce the “optimal” molecule, it is worth recalling that other aspects of the evolved line A molecules, such as their increased specificity for DNA substrates, may subsequently prove desirable. Replicate *in vitro* evolution experiments reduce the probability that alternate and potentially desirable solutions will remain undiscovered.

## Conclusions

This simplified experimental system provides a detailed look at molecular populations evolving under selection. Populations were found to follow different phenotypic trajectories and probed different regions of sequence space in which catalytically active solutions were to be found. Each of these solutions resulted from a unique array of mutations which presumably gives rise to a unique three-dimensional ribozyme architecture.

The multiple solutions exhibited by our evolving lines shed light on a fundamental question in evolutionary biology, namely, the relationship between realized solutions and the universe of possible solutions. Our observations suggest that for a single functional phenotype under selection, multiple pathways and multiple endpoints exist and can be reached even on a local scale. Any one of the solutions we have uncovered likely represents an element of a much larger ensemble of functional answers to the catalytic challenge. There is no reason to suspect that a similar relationship between the possible and the actual does not apply for any given novel molecular phenotype we choose to explore.

## Acknowledgments

We wish to thank Sid Altman, Ron Breaker, Susan Baserga, Kyle Cole, Jeffrey Huckaby, Gerald Joyce, Kenneth Kidd, Junhyong Kim, Richard Lenski, Margaret Riley, Rajni Sethi, and our anonymous reviewers for suggestions and comments that greatly improved this manuscript. This work was supported by the Yale Science Development Fund, by an NSF Dissertation Improvement Grant to M.M.H. (DEB 9623303), and by NASA grant NAG5-8813.

## LITERATURE CITED

- BEAUDRY, A. A., and G. F. JOYCE. 1992. Directed evolution of an RNA enzyme. *Science* **257**:635–641.
- BULL, J. J., M. R. BADGETT, H. A. WICHMAN, J. P. HUELSENBECK, D. M. HILLIS, A. GULATI, C. HO, and I. J. MOLINEUX. 1997. Exceptional convergent evolution in a virus. *Genetics* **147**:1497–1507.
- BULL, J. J., C. W. CUNNINGHAM, I. J. MOLINEUX, M. R. BADGETT, and D. M. HILLIS. 1993. Experimental molecular evolution of bacteriophage T7. *Evolution* **47**:993–1007.
- CADWELL, R. C., and G. F. JOYCE. 1992. Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**:28–33.
- CUNNINGHAM, C. W., K. JENG, J. HUSTI, M. BADGETT, I. J. MOLINEUX, D. M. HILLIS, and J. J. BULL. 1997. Parallel molecular evolution of deletions and nonsense mutations in bacteriophage T7. *Mol. Biol. Evol.* **14**:113–116.
- DARWIN, C. 1859. The origin of species by means of natural selection. Republished *in* B. A. CERF and D. S. KLOPFER, eds. The origin of species and the descent of man. Random House, New York.
- DE VISSER, J. A. G. M., C. W. ZEYL, P. J. GERRISH, J. L. BLANCHARD, and R. E. LENSKI. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* **283**:404–406.
- EKLUND, E. H., J. W. SZOSTAK, and D. P. BARTEL. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269**:364–370.

- ELLINGER, T., R. EHRLICH, and J. S. MCCASKILL. 1998. In vitro evolution of molecular cooperation in CATCH, a cooperatively coupled amplification system. *Chem. Biol.* **5**:729–741.
- FONTANA, W., and P. SCHUSTER. 1998. Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. *J. Theor. Biol.* **194**: 491–515.
- GAVRILETS, S. 1997. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* **12**:307–312.
- GAVRILETS, S., and J. GRAVNER. 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. *J. Theor. Biol.* **184**:51–64.
- GILLESPIE, J. H. 1984. Molecular evolution over the mutational landscape. *Evolution* **28**:1116–1129.
- HAGER, A. J., and J. W. SZOSTAK. 1997. Isolation of novel ribozymes that ligate AMP-activated RNA substrates. *Chem. Biol.* **4**:607–617.
- HANCZYC, M. M., and R. L. DORIT. 1998. Experimental evolution of complexity: *in vitro* emergence of intermolecular ribozyme interactions. *RNA* **4**:268–275.
- HERSCHLAG, D., and T. R. CECH. 1990. DNA cleavage catalyzed by the ribozyme from *Tetrahymena*. *Nature* **344**:405–409.
- HUYNEN, M., and W. FONTANA. 1996. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc. Natl. Acad. Sci. USA* **93**:397–401.
- JACOB, F. 1982. *The possible and the actual*. Pantheon Books, New York.
- KAUFMANN, S. A. 1995. *The origins of order: self-organization and selection in evolution*. Oxford University Press, New York.
- LEHMAN, N., and G. F. JOYCE. 1993. Evolution *in vitro* of an RNA enzyme with altered metal dependence. *Nature* **361**: 182–185.
- LENSKI, R. E., and M. TRAVISANO. 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc. Natl. Acad. Sci. USA* **91**: 6808–6814.
- MICHEL, F., and J. L. FERAT. 1995. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64**:435–461.
- MICHEL, F., and E. WESTHOF. 1990. Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**:585–610.
- NAKATSU, C. H., R. KORONA, R. E. LENSKI, F. J. DEBRUIJN, T. L. MARSH, and L. J. FORNEY. 1998. Parallel and divergent genotypic evolution in experimental populations of *Ralstonia* sp. *J. Bacteriol.* **180**:4325–4331.
- SCHUSTER, P., W. FONTANA, P. F. STADLER, and I. HOFACKER. 1994. From sequences to shapes and back: A case study in RNA secondary structures. *Proc. R. Soc. Lond. B Biol. Sci.* **255**:279–284.
- TSANG, J., and G. F. JOYCE. 1994. Evolutionary optimization of the catalytic properties of a DNA-cleaving ribozyme. *Biochemistry* **33**:5966–5973.
- . 1996. Specialization of the DNA-cleaving activity of a group I ribozyme through *in vitro* evolution. *J. Mol. Biol.* **262**:31–42.
- WRIGHT, S. 1948. On the role of directed and random changes in gene frequencies in the genetics of populations. *Evolution* **2**:279–294.
- . 1949. Adaptation and selection. Pp. 365–389 *in* G. L. JEPSON, G. G. SIMPSON, and E. MAYR, eds. *Genetics, paleontology and evolution*. Princeton University Press, Princeton, N.J.
- WRIGHT, M. C., and G. F. JOYCE. 1997. Continuous *in vitro* evolution of catalytic function. *Science* **276**:614–617.
- ZAUG, A. J., and T. R. CECH. 1986. The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science* **231**:470–475.
- . 1997. Peptide bond formation by *in vitro* selected ribozymes. *Nature* **390**:96–100.

JULIAN ADAMS, reviewing editor

Accepted March 14, 2000