

## Fold or hold: experimental evolution *in vitro*

S. COLLINS\*, A. RAMBAUT\* & S. J. BRIDGETT†

\*Ashworth Laboratories, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

†Ashworth Laboratories, Institute of Evolutionary Biology, The Gene Pool Genomics Facility, University of Edinburgh, Edinburgh, UK

### Keywords:

adaptation;  
aptamer;  
DNA hairpin;  
experimental evolution;  
polymerase chain reaction.

### Abstract

We introduce a system for experimental evolution consisting of populations of short oligonucleotides (Oli populations) evolving in a modified quantitative polymerase chain reaction (qPCR). It is tractable at the genetic, genomic, phenotypic and fitness levels. The Oli system uses DNA hairpins designed to form structures that self-prime under defined conditions. Selection acts on the phenotype of self-priming, after which differences in fitness are amplified and quantified using qPCR. We outline the methodological and bioinformatics tools for the Oli system here and demonstrate that it can be used as a conventional experimental evolution model system by test-driving it in an experiment investigating adaptive evolution under different rates of environmental change.

A central goal of evolutionary biology is to explain adaptation seamlessly from gene to ecosystem – to identify genetic changes within a population, quantify the action of natural selection on genetic variation, link this genetic variation to phenotypic variation and to variation in fitness and finally to relate organismal changes to adaptation to some aspect of the environment. To do this comprehensively in a natural population in a natural ecosystem requires persistence and luck. The handful of successful cases study traits with a simple genetic basis underlying a phenotype with a clear and easily measured adaptive value, or involve research lifetimes devoted to the careful study of a single natural system (Rainey & Travisano, 1998; Abzhanov *et al.*, 2006; Schluter *et al.*, 2010). One alternative to such luck and devotion is to bring evolution into the laboratory, where biologists can carefully control environmental conditions and use genetically tractable model organisms to study traits that have clear links to fitness particular environmental drivers. This approach has been successful with systems such as *Pseudomonas*, where adaptive radiation can be reliably induced under laboratory conditions (Spiers *et al.*, 2002; McDonald *et al.*, 2009).

However, we are still unable to predict or interpret most of the genetic variation that occurs during laboratory selection experiments, even in our best-studied model organisms evolved in monocultures in simple environments, and many of our insights stem from case studies of exceptions rather than systematic surveys (Barrick *et al.*, 2009). Although high-throughput sequencing can give us a comprehensive picture of genetic changes, it cannot help us link these genetic changes to changes in phenotype or to particular environmental drivers. In short, published studies have provided us with a sample of what can happen in adapting populations, but have left open the question of what, on average, does happen most of the time.

Much of our inability to explain adaptation from gene to ecosystem stems from our lack of data on the distributions of effects that contribute to fitness gain. Concretely, it would be useful to know the distributions of fitness effects of beneficial mutations, of fixed mutations and of epistatic interactions between fixed mutations. For example, knowing the distribution of epistatic interactions between fixed mutations would allow us to quantify epistatic constraints on evolutionary trajectories and formalize how epistasis and pleiotropy contribute to the repeatability of adaptive outcomes.

Here, we present a new *in vitro* system based on a modified quantitative polymerase chain reaction (qPCR) for carrying out laboratory selection experi-

Correspondence: Sinead Collins, Ashworth Laboratories, Institute of Evolutionary Biology, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK.  
Tel.: +44 (0) 131 650 8657; fax: +44 (0)131 650 5455;  
e-mail: s.collins@ed.ac.uk

ments to build empirical genotype–phenotype–fitness maps and to measure the distributions of fitness effects of mutations and epistatic interactions for adapting populations of DNA hairpins. This system is a halfway-point between computer simulations, which are limited by our understanding of how biology works, and simple viruses or virions, which, although simple relative to cellular organisms, are often still too complex to be completely tractable (Peris *et al.*, 2010; Sanjuán, 2010). Our system is conceptually similar to RNA-based *in vitro* systems such as the Q $\beta$  system (Orgel, 1979; Joyce, 2007), where short molecules evolve under carefully defined chemical conditions and can be studied at the level of sequence change or phenotypic change.

Although RNA-based or aptamer systems have been used to demonstrate that natural selection operates *in vitro* (Orgel, 1979), shed light on origin of life chemistry and the evolution of the specific biology of short nucleotides (Ellington, 2009), they have not been widely adopted as a model system for experimental evolution studying fundamental questions on the distributions of fitness effects of mutations, epistatic interactions or determinants of lineage success or extinction in the same way or to the same extent that *in silico*, viral and microbial systems have, and are now used primarily by and for chemistry researchers. We hope that by introducing a PCR-based, easy-to-use system that depends on skills and equipment that most experimental evolutionary biologists already possess, *in vitro* evolution will become more widely used. Indeed, our system could be used alongside one like Q $\beta$  to gain power and generality, in much the same way *Pseudomonas* is used alongside *Escherichia coli*.

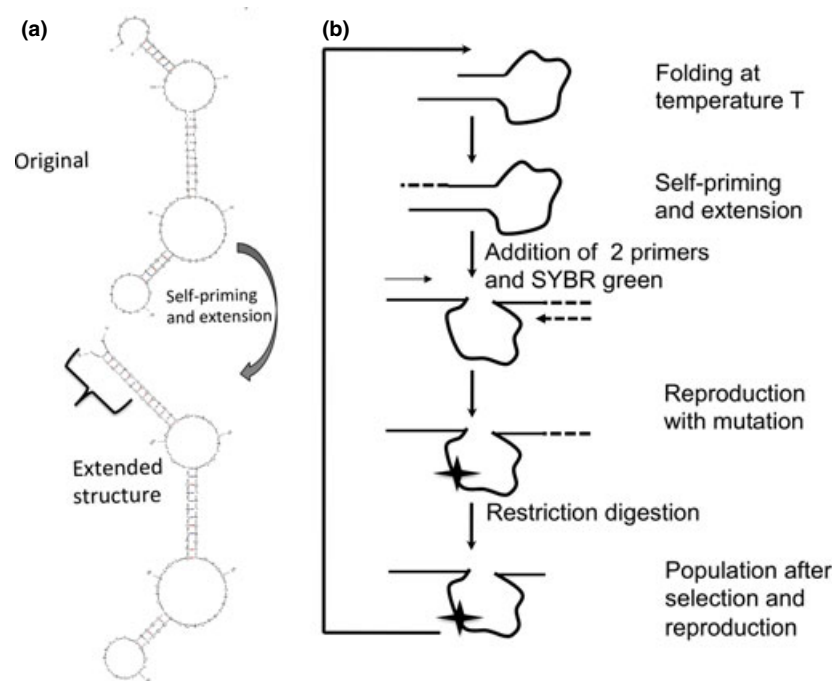
The *in vitro* system presented here, like other nucleotide-based systems for *in vitro* evolution, has shortcomings, one of which is obviously that it is far simpler than any real organism. However, it is intended to fill in gaps in our current arsenal of experimental and digital model systems used to ask general questions using experimental evolution, not to replace any of them. For this proof-of-principle study, we use this novel *in vitro* model system to explore determinants of lineage persistence in changing environments and outline other questions that could be answered from the same experiment.

## The model system

The ‘organism’ is called Oli, which is short for ‘short oligonucleotide’. Oli is a single-stranded DNA hairpin designed to fold stably at 45 °C in its ancestral state into the predicted hairpin structure shown in Fig. 1a, where it is the only structure predicted to have a negative free energy. This predicted hairpin remains the only predicted structure with a negative free energy at between 45 and 55 °C, gradually becoming less stable as the folding temperature increases. Above 55 °C, there are no folded structures predicted with a

negative free energy of folding, so that Oli is unlikely to form a hairpin where the two ends manage to pair over the section where they are complimentary. The hairpin is still predicted to form with low likelihood with a negative free energy at 55 °C. The upper temperature limit for self-priming was determined empirically by quantitative PCR, and is consistent with biomolecular folding predictions. Details of Oli design and structure can be found in the Materials and Methods. As Oli has a genotype (sequence) and a phenotype (predicted most likely structure) that is sensitive to environmental conditions, such as ion concentrations and temperature, Oli populations can evolve if they can be copied with a mutation rate. Oli does not reproduce by being alive, but is instead copied by being supplied with the building blocks for new individuals (deoxynucleotides) and the means to use them (Taq polymerase) under defined chemical and physical conditions that can be easily monitored (qPCRs). Mutagenic Taq with a defined mutation rate and spectrum can be used to control the rate of evolution of Oli populations.

Oli populations can be used for laboratory selection experiments as outlined in Fig. 1b. Here, the trait under selection is the ability to form a structure that pairs the complementary sections of the 5' and 3' ends of Oli at increasing temperatures. A single round of selection and reproduction with mutation consists of the following five steps: (1) In a solution supplied with deoxynucleotides and Taq, initially dilute Oli populations are given a short time to form any structure that allows them to self-prime at a given temperature. Those that are able to self-prime can extend their 3' end (indicated by bracket in Fig. 1a and a dashed section in 1b). Because populations are dilute at this step, it is more effective to self-prime than to use another Oli molecule as a primer, although using other molecules as primers probably happens at low frequencies. Note that no primers are present in solution at this step. (2) Two unique primers are added to the solution. As Oli has nonidentical 5' and 3' ends when 20 nucleotide (or longer) primers are used, molecules that have successfully self-primed and extended in step 1 are able to bind both primers, whereas molecules that have failed can only bind one primer. (3) Oli populations expand in a PCR. Molecules able to bind both primers found lineages that increase exponentially, whereas molecules that only bind one primer found lineages that increase linearly, giving a selective advantage to molecules that self-primed in step 1. This results in ‘fold or hold’ selection on phenotype. In this step, primer binding temperatures are kept constant, and the formation of a secondary structure allowing self-priming is no longer needed. Overall population growth rates can be monitored in real time using SYBR green. Populations are primer-limited. Details of the PCR reagents and program are given in the Materials and Methods. (4) The



**Fig. 1** Schematic of the founding Oli molecule and ‘fold or hold’ selection. (a) Self-priming and extension in the founding Oli molecule. The most likely structure of Oli for folding and self-priming at the beginning of the experiment is shown. See Materials and Methods for details of structure prediction. (b) A single cycle of ‘fold or hold’ selection. Oli molecules are challenged to fold at a given temperature into a structure where they can self-prime and extend. Populations of extended molecules grow exponentially in the quantitative polymerase chain reaction (qPCR) step, whereas populations of unextended molecules grow linearly and mutate. Mutations are shown as stars. Following qPCR, the extended ends of Oli molecules are cleaved. These cleaved populations are diluted, and a subsample is used for the next cycle of ‘fold or hold’ selection.

extended (3′) ends of the Oli molecules are cleaved with two restriction enzymes. Restriction sites are designed so that escaping cleavage results in being unable to bind one primer, which selects against that strategy. (5) Approximately  $10^5$  individuals (1  $\mu$ L) from the diluted and digested population are used for the next round, where steps 1–5 are repeated. In principle, it is also possible to use media composition as a selective environment, but it is logistically simpler to manipulate reaction temperatures. Note that there is no purification step, and that cut pieces of DNA will be present at very low concentrations in the inoculum for the next round of selection and qPCR. In theory, these very short oligos, which would have the same sequence as the 3′ end primer, could reduce the effective concentration of this primer in the reaction. However, note that the 5′ end primer and not the 3′ end primer, is limiting, so that this is unlikely to have affected the results.

At the end of an experiment, populations from some or all of the time points can be sequenced using 454 sequencing. Oli is short enough so that no assemblies are required, and sequencing runs can be of near-entire populations, rather than small samples of populations,

allowing population composition to be calculated directly from sequence data taken from a large proportion of the population rather than estimated from a small subsample. This allows rare or transient genotypes to be detected. Fitness for entire populations is measured during the selection experiment as qPCR output, and fitness for individual lineages/genotypes can be measured by synthesizing the derived Oli sequence of interest and subjecting them to steps 1–3 above. Oli is short enough that it may be ordered and synthesized as a standard oligonucleotide in most cases. When genome expansion occurs, Oli must be ordered as a synthetic gene.

This paper presents a ‘proof-of-principle’ experiment to demonstrate how the system may be used. Populations of Oli were selected at increasing temperatures for self-priming, where the temperature was raised either suddenly or gradually. This experiment is analogous to previous studies that investigate evolutionary responses to different rates of environmental change in digital (Collins *et al.*, 2007; Kopp & Hermisson, 2009; Uecker & Hermisson, 2011) and cellular (Perron *et al.*, 2008; Collins & de Meaux, 2009; Bell & Gonzalez, 2011) model systems.

## Materials and methods

### Oli design and structure determination

The unextended Oli molecule is a 107-bp oligonucleotide with the sequence 5'-AAGCAGTGG**GagaCCGAGTT**AtC CAaGtataGAGgccgACCcaacaaCAGCAGGCTGCTccattaac ccCAGGCTCAGATCTGggTCTAtACTtcCttc**GGTCTCC**-3'. Oli is based on a published DNA hairpin sequence from HIV-1 (Driscoll & Hughes, 2000). Upper-case letters indicate positions where Oli has the same sequence as the original hairpin, and lower-case letters indicate positions where the Oli sequence differs from the original hairpin. To form a structure capable of self-priming, Oli molecules must pair the 3' end indicated in bold with the complementary tract near the 5' end indicated in bold. Extension then produces a new 3' end ... cCtt **cGGTCTC**\*C\*CACTGCTT-3', capable of binding a primer supplied later in the protocol, and containing recognition sites for the two restriction enzymes BsaI and HpyAV (New England Biolabs, Ipswich, MA, USA) that cleave at the sites indicated by the asterisks. The recognition sequences of BsaI and HpyAV are GGTCTCN and CCTTC (N<sub>6</sub>), respectively. To escape restriction digestion, three or more independent point mutations are needed. We verified that qPCR primers could not reliably bind and produce product with two or more mismatches under the stringent conditions used in our experiment. This selects against restriction digestion escape during the experiment. This is further confirmed by the failure of genotypes showing mutations that would allow 'digestion escape' to show up in the sequenced populations, even at later time points. qPCR primer sequences are 5'-aag cag tgg gag acc gaa gg-3' ( $T_m = 62$  °C under reaction conditions used) and 5'-aag cag tgg gag acg tag tta tcc a-3' ( $T_m = 63$  °C under reaction conditions used).

The most likely structure for the ancestral Oli was determined using mfold (<http://mfold.rna.albany.edu/>) with the following settings: single-stranded DNA, 200 foldings maximum, temperature = 70 °C. Other folding programs give the same results for this system. Sequences as well as their reverse complements were folded. Structures were not empirically verified, but the decreased ability of the original Oli molecules to self-prime and extend at temperatures above 55 °C was verified by sequencing the end products of PCRs that were carried out at increasing annealing temperatures for the self-priming step, which is the key feature needed for this system. The upper temperature limit for self-priming was also confirmed functionally using qPCR. PCRs were set up with a known number of starting molecules and subjected to a single round of selection (steps up to but not including restriction digestion) at self-priming temperatures from 45 to 65 °C followed by the addition of primers and SYBR green, then 35 cycles of standard qPCR. The number of starting molecules was then calculated from the qPCR data. When

nearly all of the molecules are able to self-prime, the qPCR estimate of initial number of molecules matches the known initial number of molecules, as all molecules will be able to recruit both primers in all but the first cycle of the qPCR program. When few of the molecules self-prime, the starting population size estimated by qPCR is lower than the known starting population size of Oli. An extended Oli was synthesized and used as a positive control for the population growth rate expected if all of the molecules were able to self-prime.

### Selection experiment

All populations were started from a single sample of Oli ordered from Eurofins MWG Operon (Ebersberg, Germany), diluted to a known concentration. The initial population was sequenced using conventional Sanger sequencing. Sequencing reads did not show polymorphisms, so initial diversity was presumed to be low relative to the diversity generated by the error-prone Taq during the experiment. All starting populations consisted of 1 µL of 0.05 pmol stock (about 30 000 molecules). Fifty-six independent replicate populations were used for each of the following four selection regimes: (i) control (no change in self-priming temperature); (ii) sudden environmental change (a single increase in self-priming temperature of 15 °C); (iii) intermediate environmental change (five sequential increases in self-priming temperature of 3 °C each); (iv) slow environmental change (ten sequential increases in self-priming temperature of 1.5 °C each). In the three treatments where temperature increased, the total increase in temperature over the entire selection experiment is the same (15 °C, from 55 to 70 °C); only the rate differs. All populations in all treatments were subjected to the same number of PCR cycles and thus had the same opportunity for mutation. See Table S1 for a timeline of self-priming temperatures.

### Polymerase chain reaction

A quantitative mutagenic PCR kit was assembled from the components of the Brilliant II SYBR Green QRT-PCR kit, replacing the standard Taq with Mutazyme II from the GeneMorph II Random Mutagenesis Kit (Stratagene, California). Betaine was added to the reactions so that hairpins would melt. Mutation rate is 1–16 bp per kb of DNA. Information on the mutation rate and spectrum of Mutazyme II can be found at (<http://www.genomics.agilent.com/>). The PCR protocol consisted of one cycle of self-priming and extension, followed by the addition of primers and SYBR green to the reactions, followed by 40 cycles of quantitative PCR not requiring self-priming. The self-priming reaction used 1 µL of (restriction digested) Oli, 7.08 µL H<sub>2</sub>O, 1.5 µL 5 M betaine and the following volumes of reagents supplied in the Brilliant II SYBR Green

QRT-PCR kit: 1.5  $\mu\text{L}$  buffer, 0.0932  $\mu\text{L}$   $\text{MgCl}_2$ , 0.6  $\mu\text{L}$  dNTPs, 2.4  $\mu\text{L}$  glycerol, 0.45  $\mu\text{L}$  DMSO, 0.375  $\mu\text{L}$  Mutazyme II, for a total volume of 15  $\mu\text{L}$  per reaction. Program: melting at 98 °C for 3 min, hairpin folding temperature for 10 s, extension at 72 °C for 2 min. Following the self-priming and extension step, the reactions were cooled to 4 °C in the PCR machine, moved to ice, and a cocktail made up of the following reagents was added to each well: 0.5  $\mu\text{L}$  5 M betaine, 0.5  $\mu\text{L}$  buffer, 0.0313  $\mu\text{L}$   $\text{MgCl}_2$ , 0.2  $\mu\text{L}$  dNTPs,  $5 \times 10^6$  molecules of each primer, 0.8  $\mu\text{L}$  glycerol, 0.15  $\mu\text{L}$  of (1 : 1500 diluted) SYBR green, 0.2  $\mu\text{L}$  Mutazyme II, bringing the total reaction volume to 20  $\mu\text{L}$ . Plates were spun briefly and returned to the machine. Program: initial melting at 98 °C for 1 min followed by 40 cycles of the following: melting at 98 °C for 1 min, primer annealing at 67 °C for 1 min, extension at 72 °C for 2 min. A melting curve was plotted for each reaction following this. At an annealing temperature of 67 °C, a single mismatch is enough to impair primer binding. Except for betaine and Mutazyme II, all reagents were used in the concentrations supplied in the Brilliant II SYBR Green QRT-PCR kit. Providing each reaction with  $5 \times 10^6$  molecules of each primer results in transfer population sizes of Oli of about  $10^5$  individuals.

For qPCRs, water was used as a negative control. Standards for qPCR quantification were six different known concentrations of pre-extended Oli molecules that did not require self-priming, so that every member of the population could recruit both PCR primers. In fitness measurements, all growth rates and population sizes are calculated relative to standards.

### Restriction digestion

Following PCR, populations were digested for 1 h with BsaI, followed by digestion for 1 h with HpyAV, followed by heat inactivation at 65 °C for 25 min. One microlitre of this diluted and digested Oli population (on the order of  $10^5$  individuals) was used for the next round of PCR. As populations are upstream (5' end) primer-limited, total population sizes are smaller than expected based on the number of PCR cycles that they have gone through. Maximum population size can be controlled by the amount of primer supplied and remains roughly constant over the selection experiment.

### Fitness measurements and population size determination

The initial total population size is calculated fluorometrically using the Quant-iT DNA Assay kit, high sensitivity (Invitrogen) and black FluoroNunc 96-well flat-bottomed polystyrene microplates (Fisher). The standard Quant-iT protocol (20  $\mu\text{L}$  sample added to 180  $\mu\text{L}$  working solution) produced weak signals, possibly because

folded and/or short molecules take up the fluorophore unevenly. We modified the protocol by using 30  $\mu\text{L}$  sample added to 190  $\mu\text{L}$  working solution and a finer-scale linear standard curve (0–1.0 ng/ $\mu\text{L}$  DNA), which produced good results, as verified by measuring known concentrations of the control Oli molecules. The agreement between this measurement, which is based on the total amount of DNA present, and the initial number of molecules calculated by qPCR was used to calculate the proportion of the initial population that was able to fold and self-prime.

Population-level fitness (increase in population size per PCR cycle) was calculated manually from the linear section of the qPCR curve and standardized to the rate of increase in population size per PCR cycle of the standards on that plate.

### Sequencing and bioinformatics

Populations at time points 1, 3, 7 and 11 were pooled by treatment. Inadvertently, replicate populations were not tagged, so information that was expected to be available was lost at this step. Although unfortunate for the analysis of this particular experiment, it does not affect the proof-of-principle that the Oli system and 'fold or hold' selection may be used for this type of analysis. Because the data shown are from pooled samples, we limit our interpretation of the patterns of diversification and extinction patterns shown. They are meant only to illustrate what the system can do. In principle, individual populations should be tagged, allowing for proper reconstruction of networks below.

Sequencing and bioinformatics were carried out at The GenePool Genomics Facility at the University of Edinburgh. Scripts used are available on request. The pipeline used was as follows:

1 Roche 454 Pyro-sequencing: Sequencing libraries were prepared as per the Roche 454 Titanium library preparation protocols (Roche, 2009), using the low molecular weight (LMW) library method. Each of the sixteen samples was sequenced in separate lanes of 1/16 or 1/8 of a picotiter plate on the Roche 454-FLX Titanium sequencing platform, using in total just over three plates. Signal processing and base-calling were performed using the Roche shotgun signal-processing software, gsRunProcessor version 2.5.3. Run quality was checked, and statistics for the 454 reads were calculated using custom Perl scripts. The first sample (SC02) had been sequenced and reads randomly sampling (e.g. at 10 000, 20 000, 30 000 reads) to see at what read numbers new novel sequences stop appearing, to obtain an estimate the number reads required to maintain the same numbers of jMOTU clusters, then the remaining fifteen samples were sequenced to obtain at least this number of reads.

- 2 Extraction of reads: The sff files of reads were converted to fasta file format using the Roche's 'sffinfo' program, with the '-notrim' option to maximize the read length by *not* quality trimming. The number of 454 reads generated for each sample is in Table S2.
- 3 Clustering of reads: Clustering of reads was performed using two different methods of comparison: (i) jMOTU and (ii) a Perl script. First, the jMOTU program (Jones *et al.*, 2011) was used to cluster reads within each sample for sequence differences of 1–30 mismatches. The clustering parameters were 'no minimum length', 'low BLAST identity filter' set to 97%, 'percentage of minimum sequence length' set to 87. To run jMOTU, eight gigabytes of memory were allocated to the Java virtual machine using 'java -Xmx8000 m'. jMOTU first uses MegaBLAST (NCBI) to identify pairs of reads with high similarity, then carries out a Needleman–Wunsch alignment to calculate the exact distances between pairs of reads. 'MOTU' is an acronym for molecular operational taxonomic units. The second method used custom Perl scripts to filter reads allowing limited mismatches in the start and end primers; identify all unique sequences by collecting together reads that exactly match including checking the reverse complement of the sequence; then ordering the sequences by the most abundant within each sample and across all the samples. The final number of unique sequences for each sample is listed in the right column of Table S2. Code is posted at <https://github.com/rambaut/OliGraph>.
- 4 Restriction-site digestion: In the experiment, the sequences were restriction digested before folding, using two enzymes (same restriction site): *Bsa1* which cuts at 5'GGTCTC(N)3' and 3'CCAGAG(NNNNN)5'; and *HpyAV* which cuts at 5'CCTTC(NNNNN)3' and 3'GGAAG(NNNNN)5'. The 454 reads were similarly cut using a Perl script with regular expressions to trim the reads at these sites.
- 5 Folding prediction: The UNAFold software (Markham & Zuker, 2008) was used to predict the secondary structure of the folded nucleic acid sequence for the twenty most abundant sequences within each sample. UNAFold replaces the earlier mfold software and includes DINAMelt. For its predictions, UNAFold combines free energy minimization, partition function calculations and stochastic sampling. The predicted structures were plotted for comparison. The parameters used for UNAFold were as follows:

```
UNAFold.pl - NA = DNA - temp = 55 - max
              = 200 sequenceFile
```

### Sequence alignments and network analysis

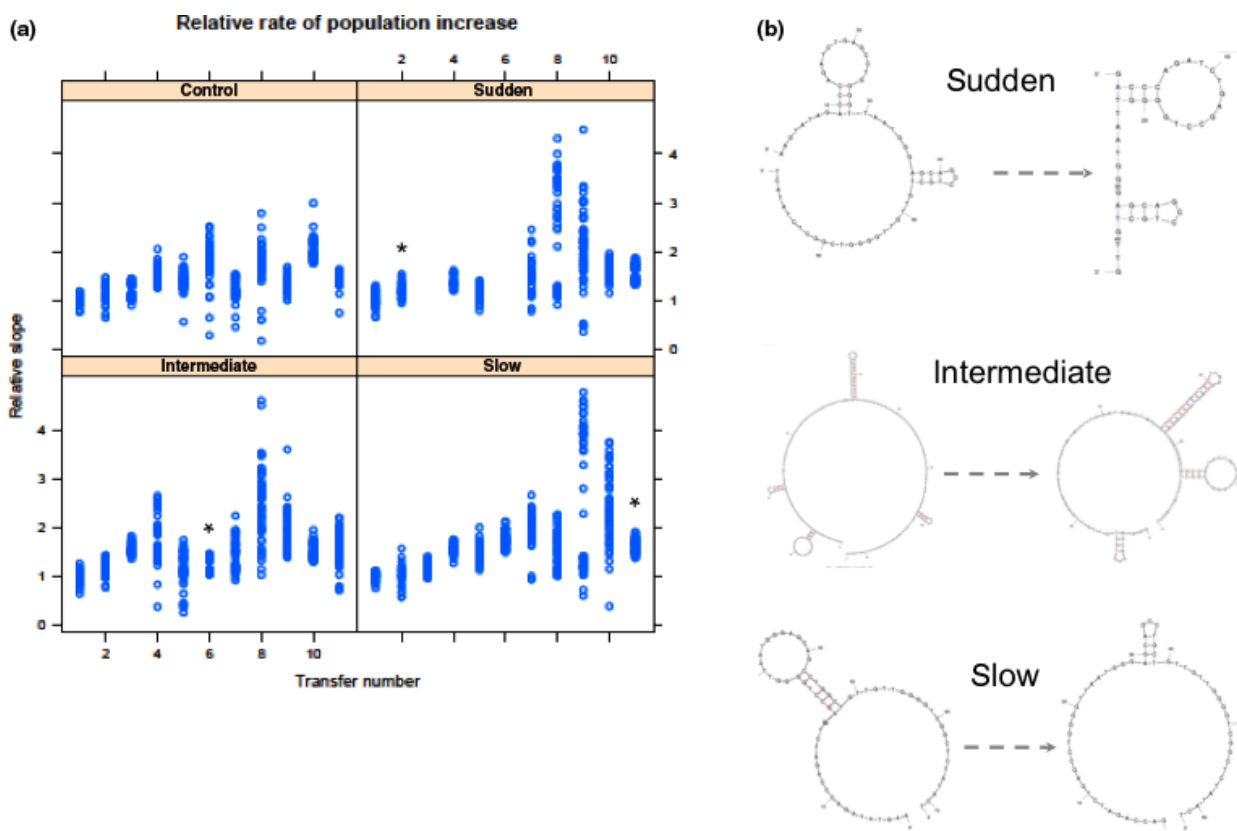
Sequence alignments and network analysis were carried out on the sequences from the pooled samples above. For each treatment, each unique sequence at the most

recent generation was pairwise aligned to each in the previous generation using the Smith–Waterman alignment algorithm (gap opening penalty of 20, extension penalty of 1). The code is available on request. The closest match (with ties broken at random) was recorded for each sequence as its putative ancestor. This was then repeated for each previous generation with the exception of the first. To construct a graphical representation of the population, each unique sequence was placed as vertex with edges provided by the inferred ancestor-descendent relationships. Vertices were annotated with generation number and sequence count and edges with the hamming distances. These graphs were visualized using Cytoscape (Shannon *et al.*, 2003). Here, predictors of lineage success were analysed using an ANOVA, where the total number of descendents was used as a measure of lineage success.

Sequencing and clustering errors were dealt with in three main ways. First, during basecalling, reads that do not have the expected four-base key sequence at their start are discarded. The reads are trimmed when the quality within a sliding window falls below preset thresholds. Second, the reads were filtered after basecalling using a custom Perl scrip to exclude reads that failed to match both primer sequences. Low-quality reads have less accurate base-calls so are less likely to match the primers so are filtered out. We used only exact primer matches for our analysis. Finally, jMOTU uses MegaBLAST to identify similar sequences, which are then merged into clusters. This clustering tends to reduce random errors in the sequences, as these random errors tend to cancel out when averaged in the cluster.

## Results

The results presented here show the range of data types that can be collected from a 'fold or hold' experiment with Oli populations. Most importantly, Oli populations adapt in response to selection for replication at increasing temperatures. This is shown in Fig. 2a. Adaptation is similar to results seen in standard microbial selection experiments and can be analysed in the same ways. Here, the population-level fitness at any time point is the rate at which the population increases, which is equal to the slope of the qPCR curve during log-linear growth. In all rising temperature treatments, the fitness of the evolved populations is higher than the populations in the control treatment ( $F_{3,219} = 9.57$ ,  $P \ll 0.0001$ ) (end relative slopes as grand means for 56 populations  $\pm$  SD for each treatment: control =  $1.41 \pm 0.0024$ , one step =  $1.54 \pm 0.0031$ , five steps =  $1.55 \pm 0.0059$ , ten steps =  $1.62 \pm 0.0024$ ), with slower rates of environmental change reaching higher end fitnesses ( $F_{1,221} = 19.45$ ,  $P \ll 0.0001$ ). This is consistent with previous results obtained in digital (Collins *et al.*, 2007) and cellular (Collins & de Meaux, 2009) experiments. The



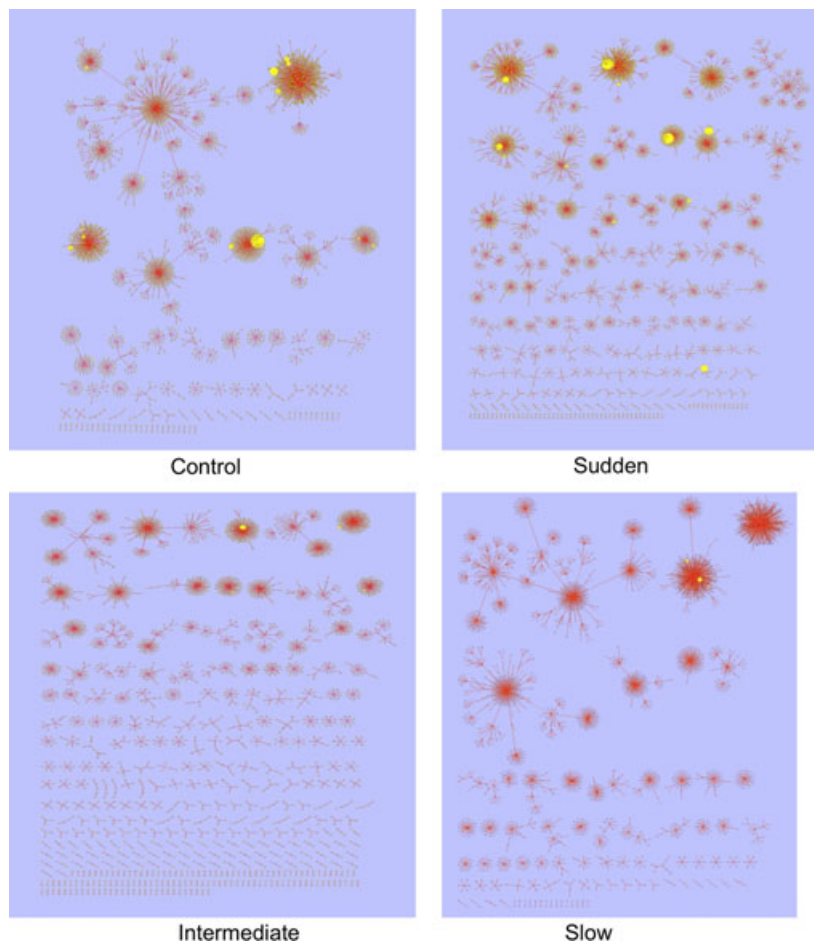
**Fig. 2** Fitness trajectories and phenotypic evolution. (a) Population growth rates, shown as relative slope, at each cycle of ‘fold or hold’ selection. Each dot represents an independent replicate population. The transfer number where the populations reach the maximum temperature of 70 °C is indicated with an asterisk. For temperatures corresponding to individual time points in individual treatments, see Table 1. Missing points are missing data rather than low slopes. (b) Phenotypic evolution. The most likely structure at 70 °C of the most numerous evolved Oli sequence at the end of the experiment (right hand structure) and its ancestor from transfer 1 (left hand structure), for each treatment. The end of the experiment corresponds to transfer 11 in panel a.

most likely phenotypes at 70 °C predicted by mfold of the most successful lineages in each treatment at the beginning of the experiment, along with their most common descendants, are shown in Fig. 2b. As programs such as mfold ignore three-dimensional structure, and as the folded structure of oligonucleotides can be dynamic, the most likely structure is not necessarily the one used for self-priming during the experiment. Indeed, the structure used for self-priming may be unlikely and/or transient. Here, visualizing the most likely structure here under the same conditions demonstrates only that phenotypic evolution has occurred and shows one way in which it can be quantified using either a structural comparison or differences in folding free energy.

The proportion within populations of individuals that fold and self-prime, as measured by the discrepancy between the total amount of DNA present and the total number of molecules detected by qPCR at each time point, varies over time and between treatments. However, it does not remain high in any of the treatments,

despite occasional bouts of wild success in some replicate populations. This can be seen in Fig. 2a. Where the among-population variance is high (such as time point 9, in the 10-step treatment), it is because of variation in the proportion of individuals within a population that fold. This is not surprising, as the mutation rate in this system is high, and so mutational load is also high. A similar phenomenon is seen in viruses with high mutation rates (Codoner *et al.*, 2006). In subsequent experiments, we recommend using a Taq polymerase with a substantially lower mutation rate to avoid this.

Figure 3 shows network diagrams depicting radiations and extinctions over time in the different treatments. Lineage success is measured by the number of descendants that a lineage has produced by the end of the experiment. Lineages at time point 1 are referred to as ‘founders’. Note that these founders are one round of selection and amplification (Fig. 1) removed from the original Oli sequence, and are all derived from the same Oli ancestor shown in Fig. 1.



**Fig. 3** Network diagrams showing radiations and extinctions in each treatment. Diagrams were made using transfers 1, 3, 7 and 11. Nodes are colour-coded by time so that transfer 1 = red, transfer 3 = yellow, transfer 7 = green and transfer 11 = violet. Sizes of nodes are proportional to the number of individuals of a particular unique sequence in the pooled populations. Distances between nodes are proportional to differences in sequence. As networks were made using populations pooled by treatment, they show average dynamics and illustrate one type of data that the Oli system can produce. See the Materials and Methods section for a detailed description of how network diagrams were produced.

For illustration, these alignments can be used to understand predictors of lineage persistence. Recall that the particular results and statistics here are for pooled populations. There is no effect of treatment on lineage success ( $F_{3,3} = 0.75$ ,  $P = 0.52$ ). Here, founders with the most descendants are also the ones with the most diverse descendants (highest number of unique sequences at the end of the experiment) ( $F_{1,1} = 68.37$ ,  $P < 0.0001$ ), and this effect varied by treatment ( $F_{3,3} = 13.99$ ,  $P < 0.0001$ ). In addition, the number of individuals present in a founding lineage does not predict the eventual success of that lineage in any treatment ( $F_{1,1} = 0.029$ ,  $P = 0.86$ ). Taken together, this suggests that the best strategy for persistence while adapting with a high mutation rate is to produce diverse offspring, some of which will survive both the high mutation rate and the environmental degradation, rather than to produce many similar offspring that then try to crowd out other lineages.

Conversely, extinction rates can be calculated, as shown in Table 1. By the end of the experiment, the control treatment has a population that derives from fewer ancestors than do the high-temperature

treatments. This appears to be due to bursts of diversity early on in adaptation in the face of environmental change, after which the per-lineage extinction rate is about the same at any time point regardless of treatment. Although the treatments have different absolute numbers of lineages, the extinction rate per lineage is the same. Here, we have demonstrated how persistence and extinction may be studied with the Oli system. However, other conventional metrics, such as nucleotide diversity, could also be used to study how variability is expected to change under different selection regimes.

One of the striking outcomes of this experiment was the repeated occurrence of drastic changes in genome length. This is summarized in Fig. 4. Changes in genome size resulted in an important hurdle to producing the networks discussed above, which was aligning palindromic sequences with many insertions/deletions. This could be eased in future experiments by using Taq with a lower error rate and subjecting populations to fewer PCR cycles in between rounds of selection for self-priming. Providing fewer limiting primers and sequencing populations at more tightly spaced time



**Table 1** Summary of extinction events pooled by selection regime. Time refers to the transfer number. The number of ancestors is the total number of unique sequences from time 1 that are still present within a given selection regime at the time stated. The number of extinctions is the number of unique sequences that have gone extinct between a time point and the time point immediately preceding it. Cumulative extinctions are the total number of unique sequences that have gone extinct between time 1 and the time point of interest. Scaled cumulative extinctions reflect the total proportion of unique sequences present at time point 1 that have gone extinct by the time point of interest. No errors are reported as the entire populations, minus the amount used for propagation and measurements during the selection experiment, were surveyed.

Treatment	Time	Number of ancestors	Number of extinctions	Cumulative extinctions	Scaled cumulative extinctions
Control	1	257	0	0	0
Control	3	175	82	82	0.319
Control	7	27	148	230	0.894
Control	11	8	19	249	0.968
Sudden	1	1074	0	0	0
Sudden	3	888	186	186	0.173
Sudden	7	106	782	968	0.901
Sudden	11	45	61	1029	0.958
Intermediate	1	1676	0	0	0
Intermediate	3	1305	371	371	0.221
Intermediate	7	113	1194	1565	0.933
Intermediate	11	36	77	1642	0.979
Slow	1	667	0	0	0
Slow	3	473	194	194	0.290
Slow	7	29	444	638	0.956
Slow	11	16	13	651	0.976

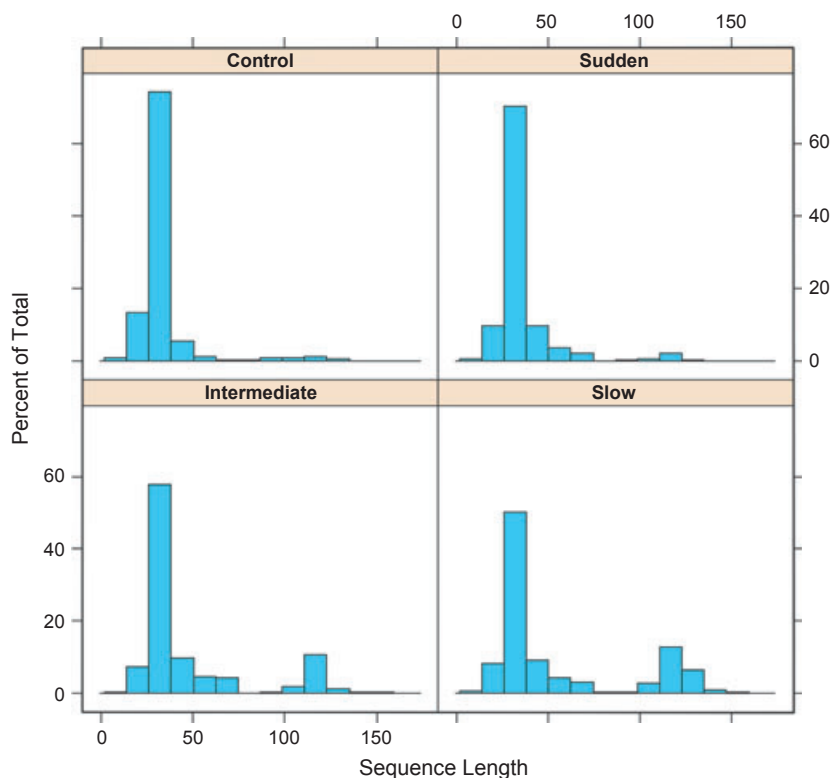
points could further help analyses by providing more intermediate sequences. The evolution of short sequences is expected in selection experiments using short oligonucleotides, and the ‘tyranny of short motifs’ has been observed before (Ellington, 2009). Interestingly, although the end populations in our experiment are dominated by sequences shorter than the ancestor, longer sequences are common in the subset of sequences at intermediate time points that leave descendants through to the end of the experiment. For example, for each treatment, the most common sequence in the final evolved populations has ancestors where the length at time 1 is 175 base pairs (control), 70 base pairs (one step), 185 base pairs (five steps) and 181 base pairs (ten steps). The same pattern continues when these successful lineages are tracked from time point 1 through to the end of the experiment. This may be because, although short sequences replicate faster during the qPCR, they may also be at a disadvantage during the fold or hold step of selection. In addition, different selection regimes result in different proportions of long and short molecules evolving, where populations evolved under gradual rates of environmental change have a higher proportion of long molecules

than either the control populations or those evolved under sudden environmental change (proportion of sequences over 75 base pairs in end populations for each treatment: control = 0.042, one step = 0.038, five steps = 0.183, ten steps = 0.324). This suggests that the extent of the ‘tyranny of short motifs’ that emerges during experiments using oligonucleotides depends on the strength of selection present.

## Discussion

The main result presented here is that the Oli system and ‘fold or hold’ selection can be used for experimental evolution in much the same way as viral or bacterial populations, using resources and expertise available in most microbial evolution laboratories. The system has both advantages and limitations relative to living systems. The main advantages of this system are that the populations are completely tractable genetically, phenotypically and in terms of fitness. Moreover, any individual sequence, including rare or extinct ones, or even ones that did not occur, can be easily synthesized and its fitness measured. The ability to create individuals that never occurred during highly replicated adaptation events allows us to directly test hypotheses about the presence and nature of evolutionary constraints, whereas the possibility of re-creating intermediate types in any desired order allows us to directly measure the fitness distributions consequences sign epistasis between sequential mutations. Distributions of fitness effects (mutation, epistatic, pleiotropic, genome size) can be determined empirically rather than estimated from a subsample that excludes rare types or events. This means that the Oli system can produce data that will improve the distributions of fitness effects that we use to inform our fundamental ideas about the way adaptation works, including the importance of rare events. Although not carried out here, Oli populations could also be used to investigate the emergence of cooperation (using others to prime and extend) or to explore questions about the evolution of recombination rates. Here, we used populations that were initially dilute to minimize the advantages of cooperation and recombination, but there is no reason that others could not change this. When Oli is used in conjunction with other (live) systems, these many advantages outweigh the limits and faults of the system, and these advantages should increase as improvements to the system are made.

All experimental systems have limitations, and Oli is no different. Most obviously, Oli is not alive and lacks the complexity of even the simplest living system. However, the point of experimental evolution is rarely to reproduce natural systems in all their glorious complexity, but rather to simplify them enough so that we can understand the general rules that evolution plays by. A second limitation is that we have not yet been able to repress cooperation



**Fig. 4** Distributions of genome lengths in evolved Oli populations at the end of the experiment, pooled by treatment. Each panel corresponds to one treatment. For reference, the length of the ancestral Oli genome is 107 base pairs long.

and recombination in this system (nor has it been quantified in this experiment – this is not an inherent limit of the system, but rather a decision about what to include in the first experiment with it). As such, Oli populations are not, and at present cannot be, completely asexual. One cheeky way around this problem would be to turn it on its head and simply use Oli populations to study the evolution of cooperation and the evolution of recombination rates. Finally, Oli populations are tricky to sequence. As fold-or-hold selection results in ever more stable hairpins, and because the precise outcome of evolution is unknown, sequencing protocols must be modified and tested at the end of the experiment. This requires a good working relationship with a sequencing centre. However, as Oli hairpins must melt so that individuals can reproduce in the PCRs, the hairpins are unlikely to evolve stability that makes sequencing impossible.

We expect to make, and hope that others will make, improvements to the Oli system. The first and most pressing improvement would be to use a polymerase with a lower mutation rate, so that more intermediate genomes are available during adaptive walks. However, polymerases with lower mutation rates do not have even mutation spectra, which introduces a mutational bias into the experiment. This could be accounted for at the end. A second and more ambitious improvement would be to estimate and eventually control the recombination and concatamerization rates of Oli

molecules in the PCRs. Although recombination seems unlikely early in the PCR program when populations are dilute, it becomes increasingly likely as population density increases. As Oli sequences must be partially palindromic and self-similar to survive fold-or-hold selection, recombination becomes likely, as does the use of other Oli molecules as primers. This results in concatamerization and changes in genome size by duplicating or deleting entire segments and may also introduce an element of genome ‘modularity’ into the system. A final and powerful improvement would be the design of several different starting Oli sequences to allow for true replication in experiments, which is necessary to generalize results. Although most microbial experimental evolution uses a single genetic starting point and still produces generalizable results, the outcome of experiments is always at least partially contingent on the genetic starting point, especially when selection is weak (Travisano *et al.*, 1995; Collins *et al.*, 2006). One aspect of the Oli system that remains to be tested is how well the realized structure of the folded molecules matches the structures predicted by folding programs. Oli is based on a structure where hairpin formation was empirically tested, and must be able to at least function as if they can fold, but it is unknown how well the most likely structure predicted by folding programs predicts the actual phenotypes present in this system.

The idea of using short oligonucleotides to investigate evolutionary phenomena is not new. It is usually applied to the evolution of hypercycles to address variations on origin of life questions laid out by Smith (1979) [But see (Wlotzka & McCaskill, 1997) for an interesting investigation of predator–prey systems using self-sustained sequence replication]. Ribozymes have also been used to investigate the role of cryptic variation in adaptation (Hayden *et al.*, 2011). In these systems, fitness and phenotype (catalytic ability) are the same trait, so that while genotype–fitness or genotype–phenotype maps may be made, connections between the three cannot be made. In contrast, the Oli system provides a novel tool for investigating evolution at the genetic, genomic, phenotypic and fitness level in populations under conditions that are simpler and better-defined than those needed for viruses or cellular organisms. Right now, accessible high-throughput sequencing is giving us more data than ever before on the genetic and genomic changes that occur during evolution, but it does not let us map those changes to the changes in fitness that natural selection ultimately acts on or to see how natural selection triages variation that is produced. The Oli system provides a way to construct genotype–phenotype–fitness maps under a range of selection and demographic regimes. One apt comment on the growing use of new technologies such as high-throughput sequencing is that ‘the techniques have galloped ahead of the concepts. We have moved away from studying the complexity of the organism; from processes and organization to composition’ [in Andrew Jack, ‘An Acute Talent for Innovation’, Financial Times (1 February 2009)]. The new ‘normal’ problem in evolutionary biology is a giant mountain of sequence data, from which biologists are supposed to extract meaningful insights about the world outside our computers and laboratories. The Oli system uses the tools that are burying us in data about composition, and subverts them to learn about process, which is not only useful, but also very satisfying indeed.

## Acknowledgments

We thank Heidi Kuehne for laboratory assistance; Petra Schneider and Andrea Betancourt for discussions on how best to subvert qPCR; Mark Blaxter, Karim Gharbi, Anna Montazam and Denis Cleven at The GenePool Genomics Facility for experimental design, library preparation and sequencing, and Martin Jones for advice on jMOTU. This work was supported by NERC small project Grant NE/G00904X/1 to SC.

## Author contributions

SC conceived of, designed, and tested the Oli system and fold-or-hold selection, carried out the experiments, analysed the data and wrote the manuscript. AR designed

and carried out the sequence alignment and network production. SB carried out the bioinformatics.

## References

- Abzhanov, A., Kuo, W.P., Hartmann, C., Grant, B.R., Grant, P.R. & Tabin, C.J. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin’s finches. *Nature* **442**: 563–567. Nature Publishing Group.
- Barrick, J.E., Yu, D.S., Yoon, S.H., Jeong, H., Oh, T.K., Schneider, D. *et al.* 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* **461**: 1243–1247.
- Bell, G. & Gonzalez, A. 2011. Adaptation and evolutionary rescue in metapopulations experiencing environmental deterioration. *Science* **332**: 1327–1330.
- Codoner, F.M., Daros, J.-A., Sole, R.V. & Elena, S.F. 2006. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS Pathog.* **2**: e136.
- Collins, S. & de Meaux, J. 2009. Adaptation to different rates of environmental change in *Chlamydomonas*. *Evol. Biol.* **63**: 2952–2965.
- Collins, S., Sültemeyer, D. & Bell, G. 2006. Rewinding the tape: selection of algae adapted to high CO<sub>2</sub> at current and pleistocene levels of CO<sub>2</sub>. *Evol. Biol.* **60**: 1392–1401.
- Collins, S., de Meaux, J. & Acquisti, C. 2007. Adaptive walks toward a moving optimum. *Genetics* **176**: 1089–1099.
- Driscoll, M.D. & Hughes, S.H. 2000. Human immunodeficiency virus type 1 nucleocapsid protein can prevent self-priming of minus-strand strong stop DNA by promoting the annealing of short oligonucleotides to hairpin sequences. *J. Virol.* **74**: 8785–8792.
- Ellington, A.D., Chen, X., Robertson, M. & Syrett, A. 2009. Evolutionary origins and directed evolution of RNA. *Int. J. Biochem. Cell Biol.* **41**: 254–265.
- Hayden, E.J., Ferrada, E. & Wagner, A. 2011. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**: 92–95.
- Jones, M., Ghoorah, A. & Blaxter, M. 2011. jMOTU and Taxonomer: Turning DNA Barcode Sequences into Annotated Operational Taxonomic Units. *PLoS ONE* **6**: e19259.
- Joyce, G.F. 2007. Forty years of *in vitro* evolution. *Angew. Chem.* **46**: 6420–6436.
- Kopp, M. & Hermisson, J. 2009. The genetic basis of phenotypic adaptation I: fixation of beneficial mutations in the moving optimum model. *Genetics* **182**: 233–249.
- Markham, N.R. & Zuker, M. 2008. UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **453**: 3–31.
- McDonald, M.J., Gehrig, S.M., Meintjes, P.L., Zhang, X.X. & Rainey, P.B. 2009. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. IV. Genetic constraints guide evolutionary trajectories in a parallel adaptive radiation. *Genetics* **183**: 1041–1053.
- Orgel, L.E. 1979. Selection *in vitro*. *Proc. Biol. Sci.* **205**: 435–442.
- Peris, J.B., Davis, P., Cuevas, J.M., Nebot, M.R. & Sanjuán, R. 2010. Distribution of fitness effects caused by single-nucleotide substitutions in bacteriophage  $\phi$ 1. *Genetics* **185**: 603–609.
- Perron, G.G., Gonzalez, A. & Buckling, A. 2008. The rate of environmental change drives adaptation to an antibiotic sink. *J. Evol. Biol.* **21**: 1724–1731.

- Rainey, P.B.P. & Travisano, M.M. 1998. Adaptive radiation in a heterogeneous environment. *Nature* **394**: 69–72.
- Roche 2009. *GS FLX Titanium General Library Preparation Method Manual*. Roche, Mannheim.
- Sanjuán, R. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**: 1975–1982.
- Schluter, D., Marchinko, K.B., Barrett, R.D.H. & Rogers, S.M. 2010. Natural selection and the genetics of adaptation in threespine stickleback. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**: 2479–2486.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. *et al.* 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**: 2498–2504.
- Smith, J.M. 1979. Hypercycles and the origin of life. *Nature* **280**: 445–446.
- Spiers, A.J.A., Kahn, S.G.S., Bohannon, J.J., Travisano, M.M. & Rainey, P.B.P. 2002. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. *Genetics* **161**: 33–46.
- Travisano, M., Mongold, J.A. & Bennett, A.F. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**: 87–90.
- Uecker, H. & Hermisson, J. 2011. On the fixation process of a beneficial mutation in a variable environment. *Genetics* **188**: 915–930.
- Wlotzka, B. & McCaskill, J.S. 1997. A molecular predator and its prey: coupled isothermal amplification of nucleic acids. *Chem. Biol.* **4**: 25–33.

## Supporting information

Additional Supporting Information may be found in the online version of this article:

**Table S1** Temperatures in °C used for the self-priming step at each transfer for each selection regime.

**Table S2** Number of 454 reads generated for each sample.

Data deposited at Dryad: doi:10.5061/dryad.1nt11

Received 19 April 2013; revised 21 July 2013; accepted 29 July 2013