PHILOSOPHICAL TRANSACTIONS B

rstb.royalsocietypublishing.org

Review

Cite this article: Pesce D, Lehman N, de Visser JAGM. 2016 Sex in a test tube: testing the benefits of *in vitro* recombination. *Phil. Trans. R. Soc. B* **371**: 20150529.
http://dx.doi.org/10.1098/rstb.2015.0529

Accepted: 13 May 2016

One contribution of 15 to a theme issue 'Weird sex: the underappreciated diversity of sexual reproduction'.

Subject Areas:
evolution

Keywords:
evolution of sex, *in vitro* recombination, fitness landscape, laboratory directed evolution, origin of life

Author for correspondence:
Diego Pesce
e-mail: diego.pesce@wur.nl

Sex in a test tube: testing the benefits of *in vitro* recombination

Diego Pesce[1], Niles Lehman[2] and J. Arjan G. M. de Visser[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, The Netherlands
[2]Department of Chemistry, Portland State University, Portland, OR 97207, USA

DP, 0000-0003-4623-3558

The origin and evolution of sex, and the associated role of recombination, present a major problem in biology. Sex typically involves recombination of closely related DNA or RNA sequences, which is fundamentally a random process that creates but also breaks up beneficial allele combinations. Directed evolution experiments, which combine *in vitro* mutation and recombination protocols with *in vitro* or *in vivo* selection, have proved to be an effective approach for improving functionality of nucleic acids and enzymes. As this approach allows extreme control over evolutionary conditions and parameters, it also facilitates the detection of small or position-specific recombination benefits and benefits associated with recombination between highly divergent genotypes. Yet, *in vitro* approaches have been largely exploratory and motivated by obtaining improved end products rather than testing hypotheses of recombination benefits. Here, we review the various experimental systems and approaches used by *in vitro* studies of recombination, discuss what they say about the evolutionary role of recombination, and sketch their potential for addressing extant questions about the evolutionary role of sex and recombination, in particular on complex fitness landscapes. We also review recent insights into the role of 'extracellular recombination' during the origin of life.

This article is part of the themed issue 'Weird sex: the underappreciated diversity of sexual reproduction'.

1. Introduction

The question why sex is widespread in nature has no simple answer, even after more than a century of investigation [1–3]. Sex typically involves recombination of closely related DNA or RNA sequences, which is fundamentally a random process that may create novel beneficial allele combinations, but may also break up existing favourable combinations. At present, at least two issues are not well understood. First, for random recombination to have a short-term adaptive benefit, theory requires weak or no epistasis among alleles [4]. Yet, evidence accumulates that epistasis may be strong and even affect the sign of the fitness effects of alleles, changing deleterious into beneficial alleles and vice versa, depending on the genetic background [5,6]. Such strong interactions among mutations cause fitness landscapes to be complex (figure 1). Despite lack of theory about the effect of recombination on complex fitness landscapes, a few recent studies predict primarily negative effects and the dependence of occasional benefits on landscape topography [7–10]. In light of the reports of recombination benefits [11–13], does this mean that theoretical work has explored the wrong conditions or has the empirical work yielded a biased view of the topography of fitness landscapes? As we still know little about the topographies of actual fitness landscapes, the problem of sex and recombination is largely an empirical problem.

Second, almost all theory on the evolution of sex and recombination is about short-term effects on adaptation, whereas a higher longer-term evolvability of sexual relative to asexual lineages may also be important [1]. However, as is true for the topography of real fitness landscapes, long-term benefits of recombination are typically outside the realm of empirical investigation. In this regard, the approach of laboratory-directed evolution presents a possible solution, because it

THE ROYAL SOCIETY PUBLISHING

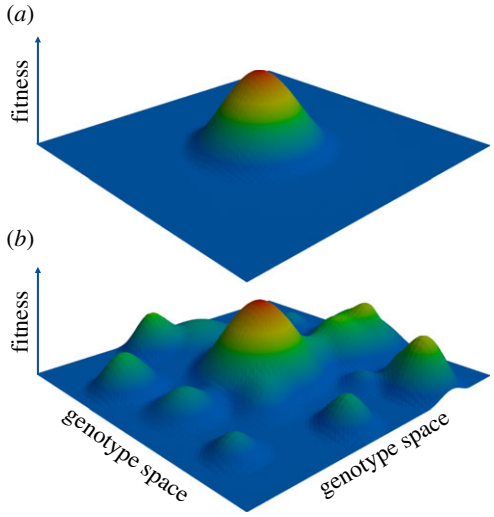

Figure 1. Fitness landscapes: (*a*) smooth single peak and (*b*) rugged multi-peak fitness landscape. Zones of higher elevation are highlighted by heat map colouration and represent genotypes with higher fitness.

allows for extreme control over the parameters and conditions and a substantial acceleration of evolutionary rates [14].

Directed evolution experiments, which combine *in vitro* mutation and recombination protocols with *in vitro* or *in vivo* selection, are an effective approach for improving or altering the functionality of biomacromolecules, from nucleic acids to enzymes. During a directed evolution experiment, evolutionary conditions and parameters are under extreme control of the experimentalist. For example, mutation rates for the target gene can be used that would be far beyond what a living organism could tolerate. Therefore, this approach allows the identification of possible small benefits associated with recombination, which are revealed only over long evolutionary periods.

In this review, we summarize efforts in directed evolution of biomacromolecules using *in vitro* recombination, often for practical applications. We include only those experiments in which the genetic material has been evolved in test tubes outside the biological context, although often selection happened in living cells (*in vivo*). We thus exclude all work on *in vivo* recombination, often in microbial evolution experiments, which is extensively reviewed elsewhere [1]. We also exclude work in computer science on the effect of recombination, where it is studied in the context of evolutionary computation algorithms [15,16]. Finally, we propose that directed evolution experiments provide a highly suited, yet underexplored, approach to test for the evolutionary role of recombination across timescales and conditions while considering the role of the topography of the fitness landscape.

2. Sequence space and fitness landscape

John Maynard Smith [17] described protein evolution as a walk from one functional protein to another in the space of possible sequences, leading to the concept of sequence space. For each sequence a 'fitness' value can be assigned, reflecting the ability of the 'owner' of the sequence to reproduce and persist in a given environment (organismal fitness). In the case of a single biomacromolecule, the concept has been extended to the level of activity or stability (molecular [18] or protein fitness [19,20]). Thus, the evolution of a biomacromolecule can be represented as an exploratory walk in this high-dimensional fitness landscape, in which regions of higher elevation represent adaptive phenotypes (figure 1).

In natural evolution, organisms explore the fitness landscape by generating genetic diversity. This diversity is the raw material upon which natural selection operates via the consequences it has for reproduction and survival, and can be distinguished in two main categories (figure 2*a*): single point mutations and recombination (depending on the mechanism, insertions and deletions may be part of both categories). Point mutations are changes in the identities of single positions in the nucleotide sequence that are 'vertically' passed to the next generation after replication. By contrast, recombination is the 'horizontal' exchange of genetic information between two or more genotypes in the same generation. Recombination may either be homologous, causing the reciprocal swapping of contiguous nucleotide sequences, or non-homologous, leading to sequences that are different in length from that of the parents (figure 2*b*). Depending on the number and position of the crossovers, homologous recombination can be symmetric, asymmetric or with multiple crossovers (figure 2*b*). These different mechanisms of generating genetic variation will affect the way in which the fitness landscape is explored (figure 2*c*). For instance, recombination, particularly non-homologous recombination, may cause populations to take larger 'jumps' than point mutations. The question about the benefit of recombination then boils down to which variation-generating mechanism, mutation or recombination, more effectively explores the adaptive landscape and converges on genotypes of high fitness (figure 2*c*).

According to population genetic theory, recombination only provides a short-term adaptive benefit when alleles with similar effect on fitness (beneficial or deleterious) present at different loci, are in negative linkage disequilibrium [2,21]. This means that these alleles should be present together in the same genotype less often than expected by chance. If so, the net result of random recombination is to increase the frequency of genotypes with multiple high fitness (and genotypes with multiple low-fitness alleles), increasing the genetic variation in fitness and allowing a faster response to natural selection. In populations adapting to a novel environment, this would mean faster adaptation (known as the Fisher–Muller effect) [22,23]; in populations experiencing mainly deleterious mutations, recombination should counter the accumulation of deleterious mutations (Muller's ratchet [24] or deterministic mutation hypothesis [25]).

Two mechanisms are thought to cause negative linkage disequilibrium—the prerequisite for a short-term recombination benefit: (i) weak negative epistasis among fitness-affecting alleles, and (ii) genetic drift and selection in the absence of epistasis [2,26]. Significant experimental effort has been devoted to finding weak negative epistasis, particularly among deleterious mutations, but without prevailing support [1,27]. Instead, there is growing support for the prevalence of much stronger epistasis affecting the sign of the fitness effects of alleles [5,6]. We have, therefore, reasons to believe that real fitness landscapes may be rugged, and contain multiple peaks and valleys [28].

Yet, little or no theory exists on the effect of recombination across timescales on fitness landscape with a rugged topography. Theoretical work has been limited to simple models and low-dimensional fitness landscapes [9]. These studies found opposite results for the short and longer-term effects of

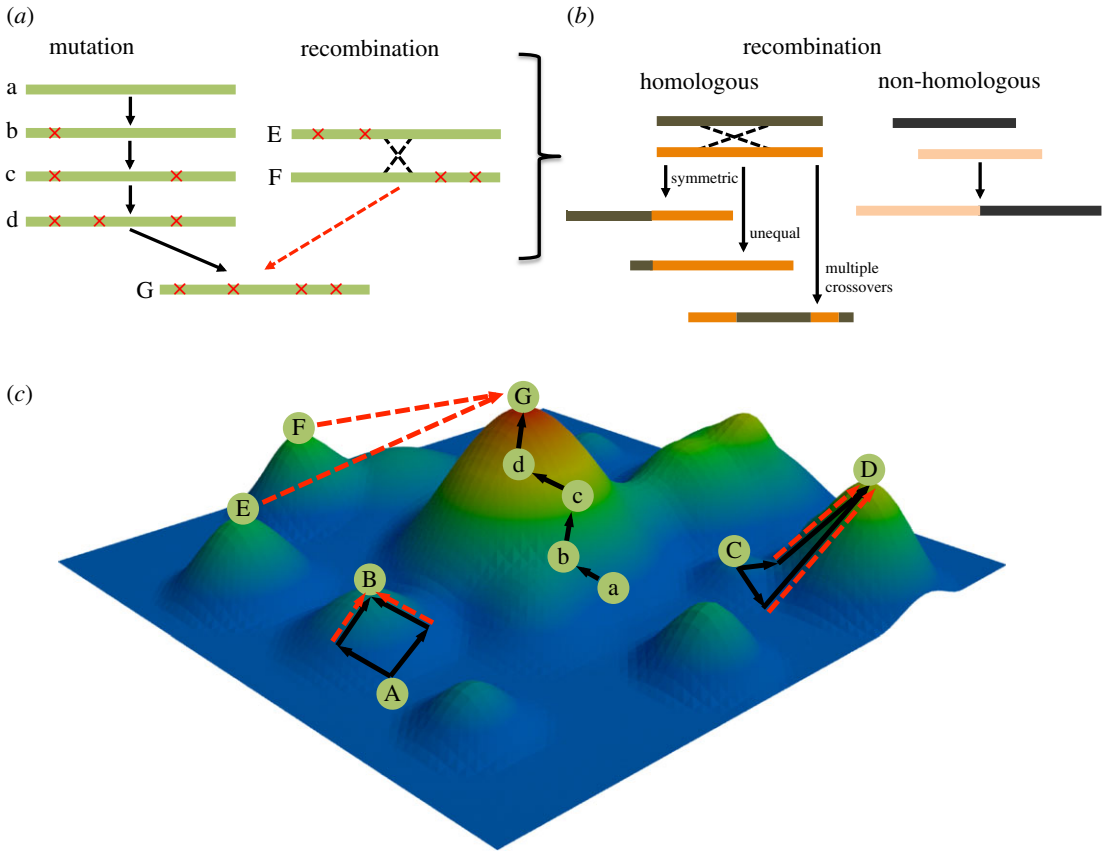

Figure 2. Genotype space and fitness landscape exploration: (*a*) schematic representation of the sequence space exploration during the evolution of a gene by mutation alone (asexual) and mutation and recombination (sexual). In (*b*) different kinds of recombination are depicted: homologous and non-homologous. Homologous recombination can be symmetric, asymmetric and with multiple crossovers. (*c*) The correspondent modes of exploration (black solid arrows for asexual and red dashed arrows for sexual mode) of a hypothetical rugged fitness landscape with local fitness peaks (e.g. B, C, D, E and F) and global peak G. The asexual uphill climbing of the global peak G by sequential mutations is represented by the transitions along the genotypes a, b, c and d (black solid arrows). In comparison, recombination between the two local peak genotypes E and F (shown by red dashed arrows) might speed up the ascent of the global peak G. Two additional contrasting scenarios for the short-term effects of recombination are shown, where low-fitness genotypes A or C adapt via two sequential mutations (indicated by solid black arrows) into higher-fitness genotypes B and D, respectively, and recombination (indicated by dashed red arrows) may or may not accelerate the generation of the double mutant.

recombination: positive short-term effects of recombination have been found associated with the ascent of a local fitness peak (figure 2*c*), which may be accounted for by existing theory [9]. However, negative long-term effects have been found during the escape phase when populations are trapped on local fitness peaks and further adaptation requires the escape from the local and ascent of a higher peak [9] (figure 2*c*). Thus, much about the effects of recombination on rugged fitness landscapes remains unexplored. For instance, recombination benefits may exist even when sign epistasis prevails, provided that sign epistasis occurs mostly between genetically linked loci, i.e. in genomic modules such as genes, and is weak between modules [29,30]. Interestingly, recombination itself is predicted to affect the epistatic architecture of the genome [31], perhaps even in ways that may have assisted its own evolution. Finally, long-term recombination benefits may arise also on rugged fitness landscapes when environmental conditions change sufficiently rapidly or populations are polymorphic and occupy multiple local peaks [8,32].

In light of these developments, a better understanding of the role of recombination on complex fitness landscapes is imperative. However, empirical studies of the role of the fitness landscape in cases of observed recombination effects are challenging, due to our limited ability to manipulate and observe small recombination benefits and obtain information about the topography of relevant parts of the fitness landscape. We propose that the use of *in vitro* recombination protocols in the directed evolution of biomacromolecules, provides a powerful alternative to the conventional use of microbes in experimental evolution [1].

3. Laboratory-directed evolution

Laboratory directed evolution is not much different from the breeding technique of artificial selection in agriculture and food production that mankind has carried out for millennia, and that helped Darwin in deciphering the mechanism of evolution and origin of species *'by Means of Natural Selection . . .'*. Indeed, Darwinian evolution involves the concerted operation of three processes that may also occur at the level of informational biomacromolecules: amplification, mutation and selection, where the phenotype under selection may be a property of the biomacromolecule itself or—more conventionally—of the organism expressing it (figure 3*a*). By repeated mutation, selection and amplification, directed evolution mimics Darwinian evolution in a test tube on a much shorter timescale. For this reason, directed evolution has been also termed 'evolutionary protein engineering' and has been successfully used to create 'artificial' biomacromolecules (i.e. DNAs, RNAs and proteins) with novel or altered functions.

In a typical directed evolution experiment, diverse alleles encoding a macromolecule of interest are generated *in vitro*,

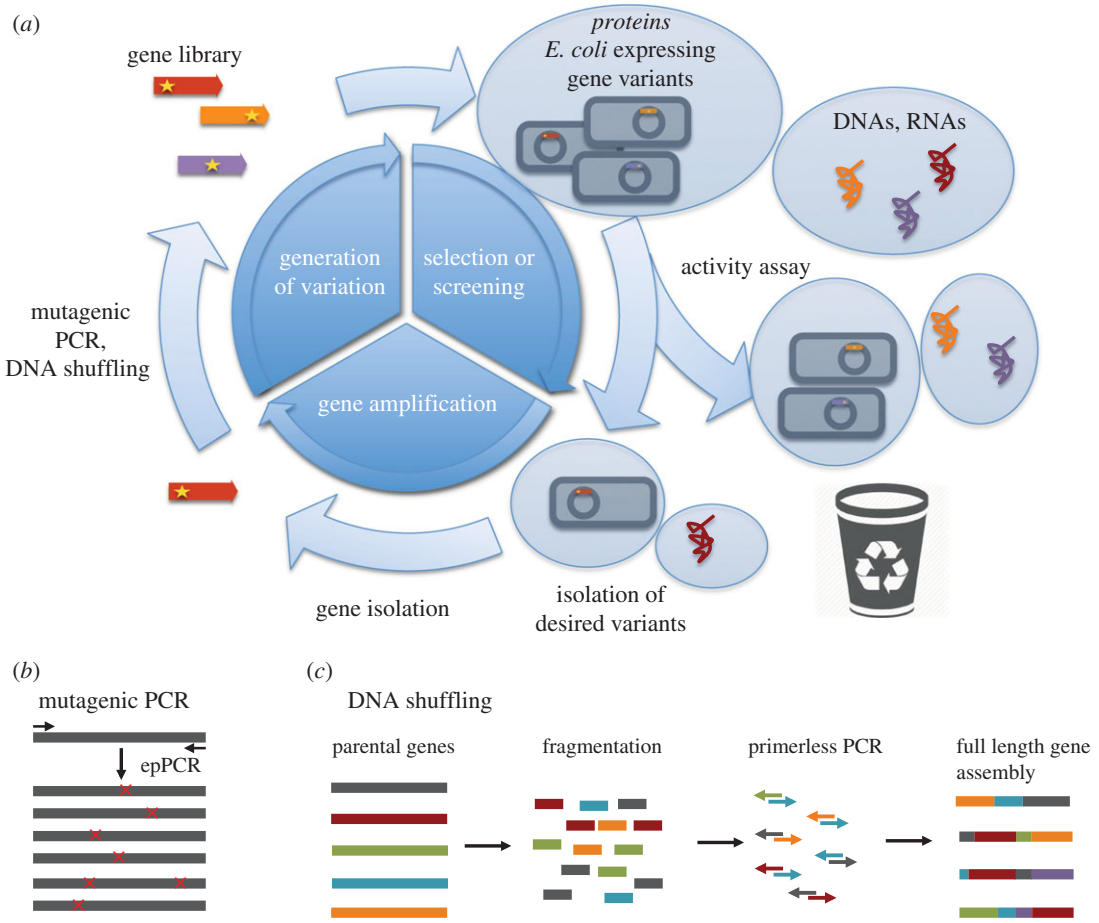

Figure 3. Laboratory directed evolution and *in vitro* diversity generation. (*a*) Schematic representation of the laboratory directed evolution method of biomacromolecules: proteins, DNAs and RNAs. Iterative cycles of generation of variation, selection/screening and amplification mimic Darwinian evolution *in vitro*. Genetic diversity is generated as follows: (*b*) mutagenic PCR or error-prone PCR (epPCR) introduces single or multiple mutations after replication of the target genetic material, whereas (*c*) DNA shuffling generates recombinant variants of the parental gene by fragmentation and primerless PCR followed by full-length reassembly of the gene.

either with mutagenesis or by a combination of mutagenesis and recombination, and appropriate screening or selection methods are used to identify mutants that possess particular properties (figure 3*a*). Selected properties include the ability to bind to a specific target, the ability to catalyse a desired chemical reaction, or increased stability in specific conditions [33]. Methods for generating diversity and advances in selection and screening methods for enzymes have been extensively reviewed elsewhere [34]. Diversity generation strategies in directed evolution are inspired by similar mechanisms occurring in nature: (i) random or targeted mutagenesis (the random or site-directed replacement of single residues; figure 3*b*), and (ii) recombination (gene sequences are mixed in a random or site-directed manner; figure 3*c*).

Mutations are generally introduced by error-prone PCR (epPCR) or other mutagenesis protocols to introduce point mutations. As its name suggests, epPCR copies the parent genes while introducing mutations as errors, mimicking naturally occurring, imperfect, DNA replication. During this process, iterative rounds of random mutagenesis via epPCR are performed and the best sequences are identified at each round of evolution and used as a template for additional mutagenesis (figure 3*a*). It is important to mention that in the biological context the mutation rate is often much lower than that used in laboratory-directed evolution experiments, where it can be fine-tuned and modulated by the experimenter, and generally allows for a much faster evolutionary response. Iterative rounds of random mutagenesis and selection have been the most frequently used method in directed evolution experiments. When mutation rates are not too high, this process mimics the 'step by step' exploration of the fitness landscape where asexual populations move 'uphill' towards a fitness peak (figure 2*c*).

4. *In vitro* sexual evolution

In 1994, Pim Stemmer introduced *in vitro* recombination of DNA sequences in the form of 'DNA shuffling' for the directed evolution of proteins [35]. Inspired by the theoretical advantage of sex in natural populations and in artificial breeding, he intentionally introduced a sexual mechanism in the evolutionary design strategy (unintentional recombination can take place by template switching during PCR [36]). For this reason, the method is also known as 'sexual PCR' or 'molecular breeding'. Ever since, the Stemmer method has been the most used *in vitro* recombination protocol in directed evolution of proteins. It involves generally the following steps (figure 3*c*): (i) selection of the parental sequences, (ii) fragmentation of the parental sequences by enzymatic digestion, (iii) reassembly of the fragments by PCR cycles, and (iv) amplification by PCR of the full-length reassembled sequences.

In the past 20 years, a broad array of recombination methods have been reported (extensively reviewed in [34,37]), such as family shuffling [38], staggered extension PCR (StEP) [39] and RACHITT (random chimeragenesis on transient

templates) [40]. Those methods were developed with the aim of obtaining: (i) higher fidelity recombination (as the original procedure was introducing both mutations and recombination), (ii) better control over the number and position of crossovers, (iii) recombination independent from sequence homology in order to mimic different kinds of naturally occurring recombination, and (iv) easier and faster protocols.

In vitro recombination experiments can start from two (pairwise) or multiple (pool-wise) parental sequences with either high, low or no sequence homology (figure 3*c*). When homology is high, the procedure resembles the natural process of homologous recombination (figure 2*b*). Methods such as DNA shuffling, StEP, family shuffling or RACHITT simulate homologous recombination and thus largely preserve gene structure in the recombinants. If there is little or no sequence homology (e.g. SHIPREC [41], ITCHY [42], NRR [43]), the process is similar to naturally occurring non-homologous recombination, such as during the horizontal transfer of new genes or the non-homologous DNA end joining used by eukaryotic cells for repairing DNA breaks [44]. Those recombination methods permit shuffling of disparate gene elements. In this regard, it is important to consider that virtually all evolution of sex theories only deal with homologous recombination, despite the frequent occurrence of non-homologous recombination in nature. Moreover, DNA shuffling can be performed with a controlled number of crossing overs, ranging from a single to several (either in a homologous or non-homologous fashion), simply by varying the experimental conditions.

In conclusion, virtually all forms of recombination occurring in nature can be performed in controlled ways in directed evolution experiments (figure 2*a*). These methods have been shown to be effective in evolving new functions. However, recombination has not been systematically compared with asexual protocols (e.g. epPCR) in order to test for benefits of recombination, let alone to test population genetic theory on the advantages of sex. Indeed, there are very few studies that have directly compared asexual and sexual evolution protocols.

5. Comparison of *in vitro* asexual and sexual evolution

Although asexual directed evolution is simple to execute, it has the significant drawback of discarding numerous beneficial mutations at each round of evolution. Hence, these useful variants must be rediscovered in future generations in order to become incorporated. On the other hand, sex and recombination offer the possibility to save those beneficial mutations and introduce them on successful sequences. The assortment of mutations to produce beneficial combinations is a crucial component of biological evolution and some of the benefits can be captured in the test tube. By recombining parental genes to produce libraries of sequences containing different combinations of mutations, genetic engineers may efficiently separate beneficial from deleterious mutations, and allow a faster response to selection—although such a net benefit depends on the existing distribution of these mutations across genotypes, called linkage disequilibrium (see §2). Moreover, in a final step of directed evolution the use of backcross DNA shuffling, a method that crosses the mutant of interest with a parent, allows one to remove non-contributing mutations [45,46].

Table 1 represents a comprehensive list of directed evolution experiments where *in vitro* recombination methods have been compared to a certain extent with asexual ones and benefits of recombination have been experimentally identified. In his early work, Stemmer compared three mutagenesis strategies (regular PCR, epPCR and DNA shuffling) for the directed evolution of a phage display antibody library [48]. DNA shuffling showed a larger improvement compared with the other two methods. However, the original protocol for DNA shuffling was also mutagenic, and thus the observed advantages may have come from the additional mutations introduced during *in vitro* recombination. The same group later used DNA 'family' shuffling to recombine four cephalosporinases that showed between 18 and 42% divergence at the DNA level. The experiment yielded a chimeric cephalosporinase with a 270-fold increase in resistance to the cephalosporin moxalactam, when compared with the parental sequences. By comparison, the highest improvement achieved via single sequence DNA shuffling (i.e. effectively through mutagenesis and selection only) was only eightfold [38].

Matsumura and co-workers [49] reported a more direct comparison of sexual and asexual *in vitro* evolution. The group employed the same model system (a β-glucuronidase) in order to test the efficiencies of different directed evolution techniques: recursive random mutagenesis, DNA shuffling and combinatorial cassette mutagenesis (which is a high-frequency recombination method) (sexual). Their results showed that recursive random mutagenesis produced asexual populations, where beneficial mutations drove each other extinct via clonal interference, while DNA shuffling and combinatorial cassette mutagenesis led to the accumulation of beneficial mutations within a single allele. These results therefore support a clear benefit of *in vitro* recombination [49].

Another example where effective purging of deleterious mutations was observed is the directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment [50]. Analysis of the evolutionary intermediates shows that negative mutations were effectively removed, while beneficial mutations accumulated. Using repeated rounds of low-frequency mutations (two error-prone PCRs followed by DNA shuffling) with repeated selection, Melzer *et al.* [54] changed the pH activity range and product specificity of the cyclodextrin glucanotransferase (from *Bacillus* sp.). Variants of the enzyme with increased activity in comparison with the wild-type enzyme were identified. Using high mutagenesis, only non-functional variants were obtained. Arnold and co-workers [55] showed that, among β-lactamase variants containing the same number of amino acid substitutions, variants created by recombination retain their original function (ability to confer resistance to ampicillin) with a higher probability than those generated by random mutagenesis. These results directly demonstrated that recombination of structurally related proteins preserves function with a higher probability than does random mutation introducing comparable genetic changes [55].

6. Recombination in nucleic acid evolution

In vitro recombination has most often been applied for the directed evolution of proteins, but its potential utility for the *in vitro* evolution of nucleic acids has been recently

Table 1. Directed evolution experiments using *in vitro* recombination protocols. Only studies which compare the benefits of sex and recombination with asexual evolution, are reported.

biomacro molecule	type of recombination	method	function/activity	evolved function	observed benefit	notes	reference
protein	homologous	DNA shuffling	β-lactamase, interleukin 1β	improved catalysis	n.a.	n.a.	[35,47]
	homologous	DNA shuffling	antibody-phage	binding	n.a.	comparison ePCR/shuffling	[48]
	homologous	StEP	5 subtilisin E variants	improved catalysis	n.a.	n.a.	[39]
	homologous	family shuffling	4 cephalosporinases	improved (broaden) catalysis	sampling larger sequence space	comparison pool-wise rec and single seq shuffling	[38]
	homologous	RACHITT	monooxygenase	improved catalysis	jumps in sequence space	comparison with other random mutagenesis	[40]
	homologous	DNA shuffling/combinatorial cassette mutagenesis	β-glucuronidase	improved β-galactosidase activity	clonal interference	comparison of DNA shuffling and recursive random mutagenesis	[49]
	homologous	DNA shuffling	β-lactamase	from functional unrelated sequence	purging deleterious mutations	analysis of evolutionary intermediates	[50]
RNA	homologous	synthetic shuffling	kinase ribozyme	improved	combination advantageous mutations	n.a.	[51]
	homologous	StEP	trans-splicing variant of the group I intron ribozyme from *Tetrahymena*	mRNA mediating antibiotic resistance	removing deleterious mutations	comparison of mutagenic and recombinative PCR	[52]
	homologous	StEP	RNA polymerase ribozyme	improved	escaping from fitness maximum	comparison with previous experiments	[53]
DNA	non-homologous	non-homologous random (NNR)	DNA aptamer	binding		comparison of directed evolution methods	[43]

highlighted [56–58]. Ever since the pioneering work of Spiegelman [59] and the discovery of the catalytic potential of RNA [60,61], a plethora of nucleic acids with new or improved binding and catalytic properties has been discovered through laboratory evolution, mostly by iterative mutation selection protocols. However, most of the work on RNA has been relegated to the experimental validation of the 'RNA world' hypothesis in the context of the origin of life [62]. A few studies used *in vitro* recombination of nucleic acids, both homologous [51,63,64] and non-homologous recombination [65–68], mediated by DNA shuffling either at the DNA level or directly catalysed by RNA [69–71].

In some of these studies, benefits of recombination have been observed. For example, several *in vitro* evolution attempts to improve the activity of an RNA polymerase ribozyme [72] resulted in relatively modest gains [73]. This led some to speculate that the ribozyme occupied a fitness maximum in sequence space and therefore could not be further improved [57,74]. However, Holliger and co-workers [53] used StEP (homologous recombination) and succeeded in further improving it (capable of synthesizing RNAs of up to 95 nucleotides in length). In another study, Muller and co-workers [52] evolved a trans-splicing variant of the group I intron ribozyme from *Tetrahymena* over 21 cycles of *in vitro* evolution either by epPCR or recombinative PCR (StEP) followed by selection in *E. coli* cells. The results suggested that recombination played an important role in removing deleterious mutations from the evolving ribozyme population. Also, Liu and co-workers [43] compared the effectiveness of non-homologous random recombination (NRR) with that of epPCR in the evolution of a DNA aptamer using identical starting sequences and selection conditions. They obtained an RNA aptamer with several fold higher activity than the ones without recombination. Their findings suggest that NRR evolves nucleic acids substantially more effectively than pure SELEX or SELEX combined with epPCR [43].

Benefits for non-homologous recombination have no asexual reference, but may still exist. For instance, *in vitro* evolution experiments have shown that the ligation and exchange of RNA structural domains through recombination can generate new functional RNAs [65,75]. Non-enzymatic non-homologous recombination of RNA can also be a powerful way to rapidly promote oligomer diversity [76,77], complexity [65,78] and new functionality [79]. The non-homologous recombination of preformed, short modular elements is a relevant mechanism for the natural evolution of RNA, as observed in current functional RNA molecules [80], viroids [81], and RNA virus genomes [82]. Moreover, it is important to mention that RNA has been reported to be prone to enzymatic-free recombination [64,67,83,84]. Most probably, mutation and recombination together constituted the basic evolutionary mechanisms allowing the appearance of RNA variants further selectable and replicable in a progressively successful RNA world [85]. This holds especially true when considering that primordial genomes had the necessity to acquire substantial length and to escape from Muller's ratchet and consequent error catastrophes caused by the very low fidelity of prebiotic RNA replication. These observations suggest that recombination is probably older than LUCA (last universal common ancestor) and hence of life as we know it. As initially noted by Stemmer:

> It is of interest to note that the crossover, the essence of sexual recombination and a presumably complex behavior, actually occurs in naked DNA. Since sexuality appears to be a behavior that is inherent in the building blocks, recombination and sexuality are probably as old as DNA itself. [35].

Indeed, even though recombination is often assumed to be a mode of reproduction that evolved long after asexual life forms, it is possible that it was instead an evolutionary development as ancient as the origins of life itself [62,76,86]. At the molecular level, recombination is effectively the transfer of energetically equivalent chemical moieties from one substrate to another. As this can proceed with very little free energy change, it is a facile reaction that can occur with a variety of prebiotically relevant polymers (nucleic acids, polypeptides and lipids) under a wide range of environmental conditions. Consequently, the utility of such reactions in promoting informational diversity at the precipice of life could have been extreme. In addition, there have been laboratory demonstrations, at least with RNA, that recombination can be effective in creating new, and potentially self-reproducing, sets of molecules [79,86,87]. It is of course conceivable that life might have evolved without recombination at the outset, but Occam's razor would suggest that its prevalence today is an indication of its existence at the outset.

7. What can we learn from *in vitro* evolution experiments?

Typically, in directed evolution studies, researchers are only interested in the best variants that are selected and often ignore the evolutionary paths taken by the evolving biomacromolecules. Although directed evolution experiments illustrate the power of recombination in its various forms and identify molecules with novel properties, they have not been designed to examine the evolutionary significance of recombination: if and when recombination accelerates the production of better-adapted genotypes relative to asexual evolution. Another limitation of existing *in vitro* recombination studies is that often, limited conditions are explored. Most experiments are designed to produce improved variants in a continuous search uphill (even if some have combined neutral drift and selection [88], and *in vivo* experiments have sometimes found long-term adaptive benefits of such regimes [89,90]). When strong selective pressures are applied, possible recombination benefits apparent under weaker selection pressure, when more rugged parts of the fitness landscape may be traversed, are obscured.

According to Fisher's fundamental theorem of natural selection, the response to selection is proportional to the genetic variation in fitness that is present in the breeding population [91]. However, Fisher's theorem assumes infinite population size, and may not generally hold for finite populations, particularly on rugged landscapes. Two recent evolution experiments with asexual bacterial populations of varying population size [89] and population structure [90], suggest that conditions leading to inefficient selection may sometimes provide adaptive benefits, which were ascribed to the ruggedness of the fitness landscapes involved. Strong selection was found to result in short-term adaptive benefits, as predicted by Fisher's theorem, but these immediate benefits were traded off against longer-term disadvantages. Whether recombination benefits on rugged fitness landscapes are apparent under conditions of maximal selection, as assumed by Fisher's theorem, or rather under more modest selective pressures, is not known.

Yet, the approach of *in vitro* recombination provides the experimental handles to address this question. Note, however, that finding conditions allowing maximal long-term selection responses is not only relevant for gaining fundamental insights in the evolutionary benefits of recombination, but also for maximizing responses in biotechnology.

8. New pipeline for *in vitro* studies of benefit of sex and recombination

New technologies such as next generation sequencing (NGS) are turning Darwinian engineers into 'evolutionary biochemists', with the ability to analyse high-resolution experimental fitness landscapes [92]. NGS is being exploited to determine frequency changes of microbes expressing different alleles from mutant libraries in bulk fitness assays. When such analyses are performed on systematically produced mutant libraries containing all individual mutations across entire genes, full information about the local fitness landscape is obtained. Some of these analyses have already shown how complex experimental fitness landscapes can be [6,93].

So far none of these studies involve recombination, but the combination of directed evolution experiments with and without *in vitro* recombination with the bulk analysis of their fitness holds promise to shed light also on the paradox of the evolution of recombination. *In vitro* recombination studies provide two clear advantages for studying the evolutionary significance of recombination relative to other approaches. First, the relative role of mutation and recombination can be compared under tightly controlled conditions, including a wide range of mutation and recombination rates. This allows for the detection of small effects of recombination, as well as a systematic comparison of mutation and recombination-based evolution schemes, mimicking different evolutionary timescales. Second, observed recombination benefits can be correlated with information about the fitness landscape, by reconstructing relevant parts of it, possibly followed by replay evolution experiments to test new hypotheses about the role of certain genotypes or conditions in the originally observed effects [94].

In vitro studies with proteins form a well-established playground for evolutionary biochemists, although naked nucleic acids are also a good target for studying recombination benefits in conjunction with the fitness landscape. Model systems such as antibiotic resistance enzymes (e.g. TEM β-lactamase [95]) and RNA aptamers (e.g. GTP binding aptamer [93]), which have been experimentally characterized to have complex fitness landscapes, represent good candidates for such studies, because their key properties are characterized relatively easily.

9. Outlook and challenges

The extreme experimental control of evolutionary parameters and conditions allowed by the approach of laboratory-directed evolution provides substantial promise for future *in vitro* studies of the benefits of recombination. Given the limitless abilities to manipulate mutation rates, such studies will allow the detection of possible adaptive recombination benefits across different timescales. Moreover, the approach of directed evolution will make it possible to vary the rate and position of crossovers in order to examine the influence of genetic architecture. Crucially also, observed recombination effects can be correlated with the topography of the local fitness landscape to examine the origin of linkage disequilibria underlying observed effects.

Technological improvements will enhance further possibilities with directed evolution to study evolutionary benefits of sex and recombination. For example, further improvement of sequencing technologies may increase the maximum length of single sequences that can be analysed as haplotypes, which is essential for detecting recombination in genes of typical sequence length. Another limitation is imposed by the size of combinatorial libraries that can be generated for the exploration of sequence space. Presently, relatively small sequence spaces can be investigated, which severely limits the region of the fitness landscape that can be analysed. The scale at which fitness landscapes can be investigated is not only limited by restrictions in generating mutants, but also in analysing their fitness. However, rapidly developing microfluidic and millifluidic [96,97] technologies and decreasing sequencing costs hold the promise for much higher throughput fitness assays in the near future.

One of the possible limitations in the *in vitro* study of recombination is represented by the mutagenic nature of some recombination methods. Improvements have been made [98], but further efforts are needed in order to minimize mutagenic effects of DNA shuffling methods. Another future goal will be to move from the study of a single, monomeric gene with a single function to more complex allosteric proteins and small genetic networks and operons to study the benefit of recombination in increasingly complex, yet also more biologically interesting systems. Finally, existing theories on sex consider only effects from homologous recombination, while non-homologous recombination occurs frequently in nature. It would be very informative to use the same pipeline we are proposing for comparing the effect of homologous versus non-homologous recombination, in similarly controlled ways (e.g. using the same mutant library).

Competing interests. We declare no competing interests.

Funding. D.P. and J.A.G.M.d.V. acknowledge financial support from the NWO. N.L. acknowledges partial financial support from NASA.

References

1. de Visser JAGM, Elena SF. 2007 The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat. Rev. Genet.* **8**, 139–149. (doi:10.1038/nrg1985)
2. Otto SP. 2009 The evolutionary enigma of sex. *Am. Nat.* **174**, S1–S14. (doi:10.1086/599084)
3. Weismann A. 1886 *Die Bedeutung der sexuellen Fortpflanzung für die Selektions-Theorie*. Jena, Germany: Gustav Fischer.
4. Barton NH. 1995 A general model for the evolution of recombination. *Genet. Res.* **65**, 123–144. (doi:10.1017/s0016672300033140)
5. Weinreich DM, Watson RA, Chao L. 2005 Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59**, 1165–1174. (doi:10.1111/j.0014-3820.2005.tb01768.x)
6. de Visser JAGM, Krug J. 2014 Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490. (doi:10.1038/nrg3744)
7. Wright S. 1932 The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proc. 6th Intl Congr. Genetics, Ithaca, NY*, vol. 1 (ed. DF Jones), pp. 356–366. See http://www.esp.org/

books/6th-congress/facsimile/contents/6th-cong-p356-wright.pdf.

8. de Visser JAGM, Park SC, Krug J. 2009 Exploring the effect of sex on empirical fitness landscapes. *Am. Nat*. **174**, S15–S30. (doi:10.1086/599081)
9. Nowak S, Neidhart J, Szendro IG, Krug J. 2014 Multidimensional epistasis and the transitory advantage of sex. *PLoS Comput. Biol*. **10**, e1003836. (doi:10.1371/journal.pcbi.1003836)
10. Feldman MW, Otto SP, Christiansen FB. 1996 Population genetic perspectives on the evolution of recombination. *Annu. Rev. Genet*. **30**, 261–295. (doi:10.1146/annurev.genet.30.1.261)
11. Goddard MR, Godfray H, Burt A. 2005 Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* **434**, 636–640. (doi:10.1038/nature03405)
12. Cooper TF. 2007 Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS Biol*. **5**, e225. (doi:10.1371/journal.pbio.0050225.st003)
13. McDonald MJ, Rice DP, Desai MM. 2016 Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* **531**, 233–236. (doi:10.1038/nature17143)
14. Joyce GF. 2007 Forty years of in vitro evolution. *Angew. Chem. Int. Ed. Engl*. **46**, 6420–6436. (doi:10.1002/anie.200701369)
15. Holland JH. 1992 Adaptation in natural and artificial systems. *SIAM Rev*. **18**, 529–530. (doi:10.1137/1018105)
16. Watson RA. 2006 *Compositional evolution*. Cambridge, MA: MIT Press.
17. Smith JM. 1970 Natural selection and the concept of a protein space. *Nature* **225**, 563–564. (doi:10.1038/225563a0)
18. Kauffman SA, Macready WG. 1995 Search strategies for applied molecular evolution. *J. Theor. Biol*. **173**, 427–440. (doi:10.1006/jtbi.1995.0074)
19. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik DS. 2006 Robustness–epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature* **444**, 929–932. (doi:10.1038/nature05385)
20. Romero PA, Arnold FH. 2009 Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell. Biol*. **10**, 866–876. (doi:10.1038/nrm2805)
21. Otto SP, Barton NH. 1997 The evolution of recombination: removing the limits to natural selection. *Genetics* **147**, 879.
22. Muller HJ. 1932 Some genetic aspects of sex. *Am. Nat*. **66**, 118–138. (doi:10.1086/280418)
23. Fisher RA. 1930 *The genetical theory of natural selection*. Oxford, UK: Oxford University Press.
24. Muller HJ. 1964 The relation of recombination to mutational advance. *Mutat. Res*. **106**, 2–9. (doi:10.1016/0027-5107(64)90047-8)
25. Kondrashov AS. 1988 Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440. (doi:10.1038/336435a0)
26. Barton NH. 1995 Linkage and the limits to natural selection. *Genetics* **140**, 821–841. (doi:10.1098/rstb.2010.0106)
27. Kouyos RD, Silander OK, Bonhoeffer S. 2007 Epistasis between deleterious mutations and the evolution of recombination. *Trends Ecol. Evol. (Amst.)* **22**, 308–315. (doi:10.1016/j.tree.2007.02.014)
28. Poelwijk FJ, Tănase-Nicola S, Kiviet DJ, Tans SJ. 2011 Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol*. **272**, 141–144. (doi:10.1016/j.jtbi.2010.12.015)
29. Wagner A. 2011 The low cost of recombination in creating novel phenotypes. *Bioessays* **33**, 636–646. (doi:10.1002/bies.201100027)
30. Watson RA, Weinreich DM, Wakeley J. 2010 Genome structure and the benefit of sex. *Evolution* **65**, 523–536. (doi:10.1111/j.1558-5646.2010.01144.x)
31. Neher RA, Shraiman BI. 2009 Competition between recombination and epistasis can cause a transition from allele to genotype selection. *Proc. Natl Acad. Sci. USA* **106**, 6866–6871. (doi:10.1073/pnas.0812560106)
32. Mustonen V, Lässig M. 2009 From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet*. **25**, 111–119. (doi:10.1016/j.tig.2009.01.002)
33. Joyce GF. 1989 Amplification, mutation and selection of catalytic RNA. *Gene* **82**, 83–87. (doi:10.1016/0378-1119(89)90033-4)
34. Packer MS, Liu DR. 2015 Methods for the directed evolution of proteins. *Nat. Rev. Genet*. **16**, 379–394. (doi:10.1038/nrg3927)
35. Stemmer WP. 1994 DNA shuffling by random fragmentation and reassembly: *in vitro* recombination for molecular evolution. *Proc. Natl Acad. Sci. USA* **91**, 10 747–10 751. (doi:10.1073/pnas.91.22.10747)
36. Meyerhans A, Vartanian JP, Wain-Hobson S. 1990 DNA recombination during PCR. *Nucleic Acids Res*. **18**, 1687–1691. (doi: 10.1093/nar/18.7.1687)
37. Ruff AJ, Dennig A, Schwaneberg U. 2013 To get what we aim for—progress in diversity generation methods. *FEBS J*. **280**, 2961–2978. (doi:10.1111/febs.12325)
38. Crameri A, Raillard SA, Bermudez E, Stemmer WP. 1998 DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291. (doi:10.1038/34663)
39. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH. 1998 Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol*. **16**, 258–261. (doi:10.1038/nbt0398-258)
40. Coco WM, Levinson WE, Crist MJ, Hektor HJ, Darzins A, Pienkos PT, Squires CH, Monticello DJ. 2001 DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat. Biotechnol*. **19**, 354–359. (doi:10.1038/86744)
41. Sieber V, Martinez CA, Arnold FH. 2001 Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol*. **19**, 456–460. (doi:10.1038/88129)
42. Lutz S, Ostermeier M, Moore GL, Maranas CD, Benkovic SJ. 2001 Creating multiple-crossover DNA libraries independent of sequence identity. *Proc. Natl Acad. Sci. USA* **98**, 11 248–11 253. (doi:10.1073/pnas.201413698)
43. Bittker JA, Le BV, Liu DR. 2002 Nucleic acid evolution and minimization by nonhomologous random recombination. *Nat. Biotechnol*. **20**, 1024–1029. (doi:10.1038/nbt736)
44. Lieber MR. 2010 The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem*. **79**, 181–211. (doi:10.1146/annurev.biochem.052308.093131)
45. Kurtzman AL, Govindarajan S, Vahle K, Jones JT, Heinrichs V, Patten PA. 2001 Advances in directed protein evolution by recursive genetic recombination: applications to therapeutic proteins. *Curr. Opin. Biotech*. **12**, 361–370. (doi:10.1016/S0958-1669(00)00228-7)
46. Stoop AA, Jespers L, Lasters I, Eldering E, Pannekoek H. 2000 High-density mutagenesis by combined DNA shuffling and phage display to assign essential amino acid residues in protein-protein interactions: application to study structure-function of plasminogen activation inhibitor 1 (PAI-1). *J. Mol. Biol*. **301**, 1135–1147. (doi:10.1006/jmbi.2000.4035)
47. Stemmer WP. 1994 Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389–391. (doi:10.1038/370389a0)
48. Crameri A, Cwirla S, Stemmer WP. 1996 Construction and evolution of antibody-phage libraries by DNA shuffling. *Nat. Med*. **2**, 100–102. (doi:10.1038/nm0196-100)
49. Rowe LA, Geddie ML, Alexander OB, Matsumura I. 2003 A comparison of directed evolution approaches using the β-glucuronidase model system. *J. Mol. Biol*. **332**, 851–860. (doi:10.1016/S0022-2836(03)00972-0)
50. Yano T, Kagamiyama H. 2001 Directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment: a laboratory model of molecular evolution. *Proc. Natl Acad. Sci. USA* **98**, 903–907. (doi:10.1073/pnas.031442298)
51. Curtis EA, Bartel DP. 2013 Synthetic shuffling and in vitro selection reveal the rugged adaptive fitness landscape of a kinase ribozyme. *RNA* **19**, 1116–1128. (doi:10.1261/rna.037572.112)
52. Olson KE, Dolan GF, Müller UF. 2014 In vivo evolution of a catalytic RNA couples trans-splicing to translation. *PLoS ONE* **9**, e86473. (doi:10.1371/journal.pone.0086473)
53. Wochner A, Attwater J, Coulson A, Holliger P. 2011 Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332**, 209–212. (doi:10.1126/science.1200752)
54. Melzer S, Sonnendecker C, Föllner C, Zimmermann W. 2015 Stepwise error-prone PCR and DNA shuffling changed the pH activity range and product specificity of the cyclodextrin glucanotransferase from an alkaliphilic *Bacillus* sp. *FEBS Open Bio*. **5**, 528–534. (doi:10.1016/j.fob.2015.06.002)

55. Drummond DA, Silberg JJ, Meyer MM, Wilke CO, Arnold FH. 2005 On the conservative nature of intragenic recombination. *Proc. Natl Acad. Sci. USA* **102**, 5380–5385. (doi:10.1073/pnas.0500729102)
56. Lehman N, Unrau PJ. 2005 Recombination during in vitro evolution. *J. Mol. Evol*. **61**, 245–252. (doi:10.1007/s00239-004-0373-4)
57. Lehman N, Arenas CD, White WA, Schmidt FJ. 2010 Complexity through recombination: from chemistry to biology. *Entropy* **13**, 17–37. (doi:10.3390/e13010017)
58. Chetverin AB. 2004 Replicable and recombinogenic RNAs. *FEBS Lett*. **567**, 35–41. (doi:10.1016/j.febslet.2004.03.066)
59. Mills DR, Peterson RL, Spiegelman S. 1967 An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl Acad. Sci. USA* **58**, 217–224. (doi:10.1073/pnas.58.1.217)
60. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. 1982 Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147–157. (doi:10.1016/0092-8674(82)90414-7)
61. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983 The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**, 849–857. (doi:10.1016/0092-8674(83)90117-4)
62. Gilbert W. 1986 Origin of life: the RNA world. *Nature* **319**, 618. (doi:10.1038/319618a0)
63. Gmyl AP, Korshenko SA, Belousov EV, Khitrina EV, Agol VI. 2003 Nonreplicative homologous RNA recombination: promiscuous joining of RNA pieces? *RNA* **9**, 1221–1231. (doi:10.1261/rna.5111803)
64. Nechaev SY, Lutay AV, Vlassov VV, Zenkova MA. 2009 Non-enzymatic template-directed recombination of RNAs. *Int. J. Mol. Sci*. **10**, 1788–1807. (doi:10.3390/ijms10041788)
65. Burke DH, Willis JH. 1998 Recombination, RNA evolution, and bifunctional RNA molecules isolated through chimeric SELEX. *RNA* **4**, 1165. (doi:10.1017/S1355838298980542)
66. Wang QS, Unrau PJ. 2005 Ribozyme motif structure mapped using random recombination and selection. *RNA* **11**, 404–411. (doi:10.1261/rna.7238705)
67. Lutay AV, Zenkova MA, Vlassov VV. 2007 Nonenzymatic recombination of RNA: possible mechanism for the formation of novel sequences. *Chem. Biodivers*. **4**, 762–767. (doi:10.1002/cbdv.200790062)
68. Vaidya N, Walker SI, Lehman N. 2013 Recycling of informational units leads to selection of replicators in a prebiotic soup. *Chem. Biol*. **20**, 241–252. (doi:10.1016/j.chembiol.2013.01.007)
69. Mörl M, Schmelzer C. 1990 Group II intron RNA-catalyzed recombination of RNA *in vitro*. *Nucleic Acids Res*. **18**, 6545–6551. (doi:10.1093/nar/18.22.6545)
70. Riley CA, Lehman N. 2003 Generalized RNA-directed recombination of RNA. *Chem. Biol*. **10**, 1233–1243. (doi:10.1016/j.chembiol.2003.11.015)
71. Hayden EJ, Riley CA, Burton AS, Lehman N. 2005 RNA-directed construction of structurally complex and active ligase ribozymes through recombination. *RNA* **11**, 1678–1687. (doi:10.1261/rna.2125305)
72. Johnston WK, Unrau PJ, Lawrence MS. 2001 RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292**, 1319–1325. (doi:10.1126/science.1060786)
73. Zaher HS, Unrau PJ. 2007 Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. *RNA* **13**, 1017–1026. (doi:10.1261/rna.548807)
74. Joyce GF. 2007 A glimpse of biology's first enzyme. *Science* **315**, 1507–1508. (doi:10.1126/science.1140736)
75. Joyce GF. 2004 Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem*. **73**, 791–836. (doi:10.1146/annurev.biochem.73.011303.073717)
76. Lehman N. 2003 A case for the extreme antiquity of recombination. *J. Mol. Evol*. **56**, 770–777. (doi:10.1007/s00239-003-2454-1)
77. Lehman N. 2015 The RNA world: 4,000,000,050 years old. *Life* **5**, 1583–1586. (doi:10.3390/life5041583)
78. Pino S, Sponer J, Costanzo G, Saladino R, di Mauro E. 2015 From formamide to RNA, the path is tenuous but continuous. *Life* **5**, 372–384. (doi:10.3390/life5010372)
79. Mutschler H, Wochner A, Holliger P. 2015 Freeze–thaw cycles as drivers of complex ribozyme assembly. *Nat. Chem*. **7**, 502–508. (doi:10.1038/nchem.2251)
80. Ancel LW, Fontana W. 2000 Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool*. **288**, 242–283. (doi:10.1002/1097-010X(20001015)288:3<242::AID-JEZ5>3.0.CO;2-O)
81. Flores R, Delgado S, Gas ME, Carbonell A, Molina D, Gago S, la Peña De M. 2004 Viroids: the minimal non-coding RNAs with autonomous replication. *FEBS Lett*. **567**, 42–48. (doi:10.1016/j.febslet.2004.03.118)
82. Cheng A, Zhang W, Xie Y, Jiang W, Arnold E, Sarafianos SG, Ding J. 2005 Expression, purification, and characterization of SARS coronavirus RNA polymerase. *Virology* **335**, 165–176. (doi:10.1016/j.virol.2005.02.017)
83. Chetverin AB, Chetverina HV, Demidenko AA, Ugarov VI. 1997 Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. *Cell* **88**, 503–513. (doi:10.1016/S0092-8674(00)81890-5)
84. Chetverina HV, Demidenko AA, Ugarov VI, Chetverin AB. 1999 Spontaneous rearrangements in RNA sequences. *FEBS Lett*. **450**, 89–94. (doi:10.1016/S0014-5793(99)00469-X)
85. Briones C, Stich M, Manrubia SC. 2009 The dawn of the RNA World: toward functional complexity through ligation of random RNA oligomers. *RNA* **15**, 743–749. (doi:10.1261/rna.1488609)
86. Lehman N. 2008 A recombination-based model for the origin and early evolution of genetic information. *Chem. Biodivers*. **5**, 1707–1717. (doi:10.1002/cbdv.200890159)
87. Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N. 2012 Spontaneous network formation among cooperative RNA replicators. *Nature* **491**, 72–77. (doi:10.1038/nature11549)
88. Bershtein S, Tawfik DS. 2008 Ohno's model revisited: measuring the frequency of potentially adaptive mutations under various mutational drifts. *Mol. Biol. Evol*. **25**, 2311–2318. (doi:10.1093/molbev/msn174)
89. Rozen DE, Habets MGJL, Handel A, de Visser JAGM. 2008 Heterogeneous adaptive trajectories of small populations on complex fitness landscapes. *PLoS ONE* **3**, e1715. (doi:10.1371/journal.pone.0001715)
90. Nahum JR, Godfrey-Smith P. 2015 A tortoise–hare pattern seen in adapting structured and unstructured populations suggests a rugged fitness landscape in bacteria. *Proc. Natl Acad. Sci. USA* **112**, 7530–7535. (doi:10.1073/pnas.1410631112)
91. Fisher RA. 1919 The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edin*. **52**, 399–433. (doi:10.1017/s0080456800012163)
92. Boucher JI, Cote P, Flynn J, Jiang L, Laban A, Mishra P, Roscoe BP, Bolon DNA. 2014 Viewing protein fitness landscapes through a next-gen lens. *Genetics* **198**, 461–471. (doi:10.1534/genetics.114.168351)
93. Jiménez JI, Xulvi-Brunet R, Campbell GW, Turk-MacLeod R, Chen IA. 2013 Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc. Natl Acad. Sci. USA* **110**, 14 984–14 989. (doi:10.1073/pnas.1307604110)
94. Salverda MLM, Dellus E, Gorter FA, Debets AJM, van der Oost J, Hoekstra RF, Tawfik DS, de Visser JAGM. 2011 Initial mutations direct alternative pathways of protein evolution. *PLoS Genet*. **7**, e1001321. (doi:10.1371/journal.pgen.1001321)
95. Schenk MF, Szendro IG, Krug J, de Visser JAGM. 2012 Quantifying the adaptive potential of an antibiotic resistance enzyme. *PLoS Genet*. **8**, e1002783. (doi:10.1371/journal.pgen.1002783)
96. Agresti JJ *et al.* 2010 Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl Acad. Sci. USA* **107**, 4004–4009. (doi:10.1073/pnas.0910781107)
97. Cottinet D, Condamine F, Bremond N, Griffiths AD, Rainey PB, de Visser JAGM, Baudry J, Bibette J. 2016 Lineage tracking for probing heritable phenotypes at single-cell resolution. *PLoS ONE* **11**, e0152395. (doi:10.1371/journal.pone.0152395)
98. Zhao H, Arnold FH. 1997 Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res*. **25**, 1307–1308. (doi:10.1093/nar/25.6.1307)