Title
Systems for in vivo hypermutation: a quest for scale and depth in directed evolution

Authors
Rix, Gordon
Liu, Chang C

Publication Date
2021-10-01

DOI
10.1016/j.cbpa.2021.02.008

Peer reviewed

HHS Public Access

Author manuscript

Curr Opin Chem Biol. Author manuscript; available in PMC 2022 October 01.

Published in final edited form as:

Curr Opin Chem Biol. 2021 October ; 64: 20–26. doi:10.1016/j.cbpa.2021.02.008.

Systems for *in vivo* hypermutation: a quest for scale and depth in directed evolution

Gordon Rix[1,2], **Chang C. Liu**[1,2,3,4,a]

[1]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697

[2]Center for Synthetic Biology, The Henry Samueli School of Engineering, University of California, Irvine, CA 92697

[3]Department of Biomedical Engineering, University of California, Irvine, CA 92697

[4]Department of Chemistry, University of California, Irvine, CA 92697

Abstract

Traditional approaches to the directed evolution of genes of interest (GOIs) place constraints on the scale of experimentation and depth of evolutionary search reasonably achieved. Engineered genetic systems that dramatically elevate the mutation of target GOIs *in vivo* relieve these constraints by enabling continuous evolution, affording new strategies in the exploration of sequence space and fitness landscapes for GOIs. We describe various *in vivo* hypermutation systems for continuous evolution, discuss how different architectures for *in vivo* hypermutation facilitate evolutionary search scale and depth in their application to problems in protein evolution and engineering, and outline future opportunities for the field.

Introduction

Evolution is the ultimate bioengineer. Yet from the perspective of any individual gene in a modern organism's genome, evolution acts very slowly. This is not just an empirical observation but rather a basic outcome of growing complexity in self-evolving systems. As organisms became more complex over the history of life, genomes gained and relied on more genes. Consequently, the mutation rate per gene had to decrease, because there were more and more things to break. Modern life has thus reached a point where organisms are complex, genomes are large, and the genomic mutation rate must stay low [1–4], too low to rapidly drive the extensive evolution of any particular gene.

Corresponding author: Liu, Chang C. (ccl@uci.edu).
[a]Lead contact.

Conflict of interest statement

C.C.L. is a co-founder of K2 Biotechnologies, Inc., which focuses on the use of continuous evolution technologies applied to protein therapeutics.

To an evolutionary biologist, this may not be a problem, as genomes and many-gene systems can still evolve at a high rate, giving much to observe. But to a protein engineer, this is a major problem. *If the gene encoding a protein must obey the low mutational speed limit of large genomes, how can we watch protein evolution in action? And how can we exploit the tried-and-true power of evolution to make proteins carry out ambitious new functions on laboratory timescales?* In this essay, we discuss a quickly growing area of research that aims to design and apply genetic systems that selectively hypermutate user-defined genes of interest (GOIs) within living cells [5,6]. Such systems bypass the low mutational speed limits of genomes in order to drive the rapid continuous evolution of GOIs and the proteins they encode simply as cells are passaged under selection.

Motivation

It is useful first to consider why there is a need for hypermutation and continuous evolution systems when the mature field of directed evolution has already made it possible to evolve GOIs on laboratory timescales. In classical directed evolution, researchers subject one or more GOIs to cycles of *in vitro* diversification (*e.g.* error-prone PCR), transformation of the diversified GOI library into cells, and screening or selection for desired functions. This process can be viewed as a manual bypass of genomic mutational speed limits: one mutates GOIs in a test tube to access the high rates of diversification that large genomes disallow but then transforms GOI variants into cells where they can express and be subjected to functional selection. Yet by manually staging the steps of evolution, classical directed evolution largely forfeits what may be the two most defining features of natural evolution: *scale* and *depth* (Figure 1).

First, scale. Because manual stages of diversification, transformation, and selection are labor-intensive and technically challenging, one can only classically run a few independent evolution experiments at a time, hindering exploration of powerful ideas requiring evolutionary search scale. Such ideas include exploiting spatial structure during protein evolution to escape local optima (*e.g.* the division of a single well-mixed population into many physically separated populations), maximizing diversity of outcomes by running evolution experiments in hundreds if not thousands of independent replicates, obtaining statistical information on evolutionary pathways (*e.g.* understanding drug resistance by mapping rugged fitness landscapes), and evolving a parent protein into families of variants that have different functions (*e.g.* creating versatile collections of antibodies, biosensors, or biosynthetic enzymes).

Second, depth. Classically, one can only take short “walks” on the fitness landscape of a GOI, because a single manual cycle of diversification, transformation, and selection can take days to weeks. Tantalizing ideas requiring evolutionary search depth (*i.e.* long mutational walks) are thus restricted. Such ideas include probing or exploiting the relationship between complex selection histories and adaptation (*e.g.* employing drift to escape local optima during enzyme evolution or using alternating selections to evolve “evolvability”) [7] and attaining ambitious protein functions (*e.g.* novel enzyme activities, catalytic activity from *de novo* designed structures, custom function from multi-gene metabolic pathways, or intricate protein-protein or protein-nucleic acid interactions) [8,9] that, by definition, require long

mutational pathways to reach. All of these privileges of evolutionary scale and depth are found in abundance in the natural evolution of species – they are responsible for what Darwin described as the “endless forms most beautiful and most wonderful” around us – but have not been captured for laboratory application to engineering GOIs for user-defined functions.

In our view, the key motivation for building genetic systems that achieve targeted hypermutation of GOIs *in vivo* is to transform classical directed evolution into an autonomous and continuous process so that evolutionary search scale and depth become readily available in protein and GOI engineering experiments. This transformation may usher in an era of protein biology and engineering where we can study the process of protein evolution and the underlying sequence-function relationships behind proteins with newfound statistical power, where the evolution of previously difficult protein and GOI functions becomes facile and high-throughput, and where truly novel protein functions fall into the reach of protein evolution and design.

Categorization of *in vivo* hypermutation and continuous evolution systems

The critical task for achieving *in vivo* hypermutation is figuring out how to target rapid mutational accumulation only to GOIs. This targeting is what allows us to bypass the low mutational speed limit of a cell’s genome without resorting to the manual staging of diversification and selection in classical directed evolution that restricts scale and depth in evolutionary search. So far, targeting has been achieved through three general architectures that define existing systems: architecture 1 – propagation of viral genomes through non-propagating hosts, architecture 2 – recruitment of mutagenesis machinery to specific DNA loci, and architecture 3 – orthogonal DNA replication. For reasons that will become clear, we call architecture 1 *viral*, architecture 2 *epi-hypermutation*, and architecture 3 *direct hypermutation* (Figure 2). Given our view that the main significance of continuous evolution systems is their admission of scale and depth in protein evolution, we survey these architectures with a focus on their ability to enable evolutionary scale and depth.

Viral architectures

In viral architectures for continuous evolution, GOIs are encoded on the genome of a virus where they are induced to hypermutate when propagating through host cells (Figure 2). High rates of mutation can be durably maintained over time because viral genomes are sufficiently small. By coupling the desired activity of the GOI to the ability to make virus, the hypermutating GOI continuously and rapidly evolves as the virus passes through successive hosts.

This is the basic strategy behind phage-assisted continuous evolution (PACE), which is the most mature example of the viral architecture [5,6,10,11]. In PACE, GOIs are encoded on the M13 bacteriophage genome and the phage are propagated through *Escherichia coli* host cells engineered to inducibly mutate at high rates. Although these high mutation rates in the range of 10^{-5} to 10^{-4} substitutions per base (s.p.b.) are deleterious or lethal to *E. coli*, this is not a problem for the system because the *E. coli* are constantly flowed into and

out of a reservoir containing phage; the flow rate is such that only the phage population persists, resulting in the continuous accumulation of mutations only in the phage genome and GOI. To select for desired function from the GOI, the M13 genome is engineered to lack the essential gIII gene whose protein product, pIII, is required for phage packaging and infection. Instead, gIII is encoded in the *E. coli* host where its expression can be made to depend on the desired GOI function. GOI variants with improved function induce higher expression of pIII and more phage descendants, thereby effecting the continuous evolution of the GOI. In this manner, PACE has been used to evolve RNA polymerases [10], biosensors [12], insecticidal proteins [13], base editors [14,15], metabolic enzymes [16], proteases [17], and more [18–21].

The PACE concept has been extended to mammalian cell hosts as well. Recent work has shown that it is possible to use either an engineered error-prone version of adenovirus or a naturally error-prone RNA virus, sindbis, as vehicles for GOI hypermutation [22,23]. As in PACE, GOIs are encoded on the genome of a virus engineered to lack the essential proteins required for viral production. Instead, the essential proteins are supplied by host mammalian cells where their expression is linked to the GOI's activity. In this manner, adenovirus and sindbis have been used to rapidly evolve transcription factors, GPCRs, and conformationspecific anti-GPCR nanobodies by passaging virus through successive cultures of host cells [22,23].

It has become clear, particularly with PACE, that the viral architecture can be highly durable, enabling prolonged continuous or semi-continuous GOI evolution experiments that result in depth of evolutionary search. The scale of evolutionary search using the viral architecture can also be high, but this is an area of ongoing development [21]. The viral architecture is naturally limited by the fact that viruses are not autonomously propagating agents and instead depend on a constant supply of new host cells. This can introduce the requirement for bioreactors, as in PACE where the rate of fresh cell supply demands precise external control to prevent the persistence of host mutations, or technical steps such as centrifugation and filtration to physically separate viral-encoded GOI evolution from host evolution in passaging steps. Another challenge in the viral architecture is that the unit of selection is defined by viral production. Although clever genetic circuits can link many GOI functions to viral synthesis [9,20,24], functions that occur on timescales beyond the viral lifecycle, functions that involve complex host biology, or GOI functions that are meant to change a host cell's (or even a multicellular organism's) physiology are not directly selectable through the viral architecture. To achieve greater evolutionary search scale and expand the types of functions that can be evolved, there is room in the ecosystem of continuous evolution systems for fully *in vivo* architectures.

Epi-hypermutation architectures

Since cells are autonomously propagating agents, continuous evolution of GOIs fully inside cells supports extensive evolutionary search scale – culturing cells is easy to parallelize into many independent replicates, distinct experiments, or spatially structured populations – as well as depth – culturing cells for many generations is straightforward. Of course, cells have very low mutation rates in order to properly maintain the large information content

in their megabase to gigabase genomes, and such low mutation rates (typically 10^{-9}–10^{-10} s.p.b.) are not enough to rapidly sample diversity at the level of an individual GOI. (A mutation rate of ~10^{-10} s.p.b. would sample only a single mutation every million times a kilobase-sized GOI was replicated.) To exploit the scale and depth of evolution that culturing cells should afford, one must devise systems for targeted hypermutation *in vivo* to speed up GOI evolution.

A popular strategy for building targeted hypermutation systems is to fuse DNA mutating enzymes to site-specific DNA binding proteins. A prominent example is the fusion of nucleotide deaminases to T7 RNA polymerases (RNAPs), where the deaminase induces mutagenesis and the T7 RNAP proteins provide targeting [25–28]. A particular advantage of T7 RNAP is its processivity, which acts to drag the deaminase across an entire GOI or section therein. Likewise, fusion of an error-prone DNA polymerase (DNAP) to a nickase-Cas9 can cause mutagenesis across a stretch of DNA near the Cas9 nick site, as is the basis for the EvolvR system [29,30]. These systems have been used evolve model proteins such as antibiotic resistance genes and cancer drug targets through the simple serial culturing of bacteria, yeast, or mammalian cells under selection [25–30].

We categorize these systems involving the fusion of a DNA mutating enzyme to a DNA binding protein, along with other systems such as CRISPR-X, ICE and TaGTEAM [31–33], as epi-hypermutation architectures to emphasize that the GOI targeted for hypermutation is not replicated by the hypermutation system itself. Instead, hypermutation is *epi* to the GOI's propagation, which occurs independently through the high-fidelity replication systems of the host (Figure 2). This feature may have implications on durability and evolutionary search depth achievable. For example, when a GOI is targeted for hypermutation, the very sequence responsible for recruiting hypermutation machinery may become corrupted via hypermutation. If the GOI under selection can still propagate and express independently of hypermutation, we speculate that hypermutation may slow over time. Current epi-hypermutation systems also have notable off-target activity, resulting in elevations in the mutation rate of the host genome. This potentially increases the chance of evolution outside the GOI, including selection circuits that link the desired GOI function to cell survival as well as the genes encoding the hypermutation machinery itself. The ongoing application of epi-hypermutation systems will reveal whether these issues affect the goal of enabling scale and depth in the evolution of GOIs.

Direct hypermutation architectures

Another way to achieve targeted hypermutation of GOIs *in vivo* is to give cells a separate DNA replication system dedicated to the propagation of GOIs. Such an orthogonal DNA replication system would consist of a special DNAP that only replicates a cognate plasmid encoding GOIs. Host DNAPs would replicate the host genome but not the special plasmid, completing orthogonality [34]. If the orthogonal DNAP is then made to be error-prone, the system enforces continuous hypermutation of plasmid-encoded GOIs, driving their rapid evolution *in vivo*.

We categorize orthogonal replication as a direct hypermutation architecture because mutation occurs during replication. This distinction from epi-hypermutation, where mutation of a GOI occurs independently of its replication, may be important for achieving search depth in continuous evolution experiments. For example, if the orthogonal plasmid is hypermutated in a way that prevents its recognition by the orthogonal DNAP, the plasmid isn't replicated, ensuring that only continuously mutating GOIs persist in an experiment. An additional advantage of having an altogether separate replication system for GOIs is that targeting of hypermutation can be more effectively achieved, for example through spatial separation between the orthogonal and the genomic replication systems.

Early work towards an orthogonal replication system was carried out by Camps *et al.* through the establishment of an error-prone Pol I DNAP that targets ColE1 plasmids [35]. However, since Pol I is also essential for genomic DNA replication, full orthogonality was not achieved. Over the past few years, our lab has developed a fully orthogonal DNA replication (OrthoRep) system by adapting an autonomously replicating cytoplasmic plasmid element found in certain strains of yeast and engineering error-prone variants of the DNAP responsible for replicating the cytoplasmic plasmid [34]. Our current OrthoRep systems drive the hypermutation of GOIs at a mutation rate of ~10^{-5} s.p.b., while keeping the host *Saccharomyces cerevisiae* genomic mutation rate unchanged at ~10^{-10} s.p.b. [4]. This ~100,000-fold increase in the *in vivo* mutation rate of GOIs has allowed us to evolve several enzymes and proteins with exceptional scale and depth. For example, OrthoRep was used to evolve the malarial drug target, dihydrofolate reductase, in over 100 independent experiments to map mutational pathways leading to drug resistance [4]. OrthoRep was also used to evolve the *Thermatoga maritima* TrpB enzyme in over a dozen replicate experiments to yield a collection of 60 diverse TrpB variants whose distinct promiscuous activities enabled the efficient biosynthesis of several valuable tryptophan analogs [36]. More recently, OrthoRep was used to drive the evolution of yeast surface-displayed antibodies in multiple parallel experiments that yielded potent neutralizers of SARS-CoV-2 pseudovirus and conformationally-selective high-affinity nanobodies against a GPCR [37]. In these and other examples of OrthoRep-based GOI evolution [38], the sheer number of evolution experiments performed and the complexity of the evolved GOI sequences, which contain 10–20 mutations in some cases, highlight the scale and depth of evolutionary search possible with the direct hypermutation architecture.

OrthoRep is not without its limitations. New versions of the orthogonal DNAP are needed to reach higher mutation that can speed up GOI evolution even further. New DNAPs are needed to lower the bias for transition mutations that current error-prone orthogonal DNAPs exhibit. Stronger promoters for hypermutating GOIs on the orthogonal plasmid and new strategies for expression control are desired to improve the range of GOIs and properties that can be evolved [39]. Novel selection strategies that couple arbitrary desired GOI functions to cell fitness with a level of durability that matches the prolonged hypermutation power of OrthoRep are also needed. Additionally, whether OrthoRep can be ported into organisms beyond yeast remains to be seen, in contrast to epi-hypermutations systems, which have already been established in bacteria, yeast, and mammalian cells. We expect that the unique architectural advantages of OrthoRep should provide sufficient motivation for continued development in these directions.

Opportunities for the future

Throughout this essay, we have emphasized that the distinguishing value of *in vivo* hypermutation systems is their ability to access scale and depth in the evolution of GOIs, because those are the two properties that we believe can extend the field of directed evolution into new categories of problems. Such problems include understanding the rules of evolution with greater statistical power, mapping the fitness landscapes of proteins at higher resolution, and evolving biomolecules to achieve increasingly ambitious functions. A more fundamental problem that scale and depth can address is the challenge of escaping local optima in the evolution of GOIs to reach better and better function. Spatially structured populations that contain thousands of semi-independently evolving subpopulations (requiring scale) and fluctuating selection environments that introduce periods of drift (requiring depth) can suppress the dominance of local optima in the outcomes of evolution, and the exploitation of these two evolutionary strategies is an opportunity made realistic by the *in vivo* hypermutation systems described. Finally, the powerful combination of directed evolution experiments with machine learning should further elevate the value of evolutionary search scale and depth enabled by *in vivo* hypermutation systems [40,41]. In the future, it may be possible to use systems such as OrthoRep to evolve thousands of highly diverse variants of thousands of proteins that achieve thousands of different functions. As the resulting datasets are used to train models that predict even more diverse functional sequences, a virtuous cycle may initiate where the sequence-function relationships governing all proteins are captured over time. We believe these exciting opportunities motivate the continued development of *in vivo* hypermutation systems with deliberate focus on architectures that maximize scale and depth in evolutionary search.

Acknowledgements

This work was funded by NIH NIGMS 1R35GM139513.

References

[1]. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. Proc Natl Acad Sci U S A 1991;88:7160–4. 10.1073/pnas.88.16.7160. [PubMed: 1831267]

[2]. Biebricher CK, Eigen M. What is a quasispecies? Curr Top Microbiol Immunol 2006;299:1–31. 10.1007/3-540-26397-7_1. [PubMed: 16568894]

[3]. Herr AJ, Ogawa M, Lawrence NA, Williams LN, Eggington JM, Singh M, et al. Mutator suppression and escape from replication error-induced extinction in yeast. PLoS Genet 2011;7. 10.1371/journal.pgen.1002282.

[4]. Ravikumar A, Arzumanyan GA, Obadi MKA, Javanpour AA, Liu CC. Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. Cell 2018;175:1946–1957.e13. 10.1016/j.cell.2018.10.021. [PubMed: 30415839] ** A highly error-prone variant of an orthogonal DNA polymerase was engineered in order to drive the rapid evolution of GOIs on an orthogonal plasmid. This system was applied toward the replicate evolution of *Plasmodium falciparum* DHFR to understand common and rare pathways to resistance against an antimalarial drug. This application highlights the benefit of scalability where many replicate evolution experiments were necessary to obtain a broader understanding of mutational pathways leading to resistance.

[5]. Simon AJ, d'Oelsnitz S, Ellington AD. Synthetic evolution. Nat Biotechnol 2019. 10.1038/s41587-019-0157-4.

[6]. Morrison MS, Podracky CJ, Liu DR. The developing toolkit of continuous directed evolution. Nat Chem Biol 2020;16:610–9. 10.1038/s41589-020-0532-y. [PubMed: 32444838]

[7]. Aharoni A, Gaidukov L, Khersonsky O, Gould SMQ, Roodveldt C, Tawfik DS. The "evolvability" of promiscuous protein functions. Nat Genet 2005;37:73–6. 10.1038/ng1482. [PubMed: 15568024]

[8]. Chen K, Arnold FH. Engineering new catalytic activities in enzymes. Nat Catal 2020;3:203–13. 10.1038/s41929-019-0385-5.

[9]. Zinkus-Boltz J, Devalk C, Dickinson BC. A Phage-Assisted Continuous Selection Approach for Deep Mutational Scanning of Protein-Protein Interactions. ACS Chem Biol 2019;14:2757–67. 10.1021/acschembio.9b00669. [PubMed: 31808666]

[10]. Esvelt KM, Carlson JC, Liu DR. A system for the continuous directed evolution of biomolecules. Nature 2011;472:499–503. 10.1038/nature09929. [PubMed: 21478873]

[11]. Miller SM, Wang T, Liu DR. Phage-assisted continuous and non-continuous evolution. Nat Protoc 2020;15:4101–27. 10.1038/s41596-020-00410-3. [PubMed: 33199872]

[12]. Pu J, Zinkus-Boltz J, Dickinson BC. Evolution of a split RNA polymerase as a versatile biosensor platform. Nat Chem Biol 2017;13:432–8. 10.1038/nchembio.2299. [PubMed: 28192413]

[13]. Badran AH, Guzov VM, Huai Q, Kemp MM, Vishwanath P, Kain W, et al. Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. Nature 2016;533:58–63. 10.1038/nature17938. [PubMed: 27120167] * Phage assisted continuous evolution involving a bacterial two-hybrid selection for protein-protein interactions was used to evolve an insecticide that largely overcomes overcomes insect resistance. The ~10 mutations relative to wild-type present in evolved variants demonstrate the high mutational depth that is achievable with PACE.

[14]. Thuronyi BW, Koblan LW, Levy JM, Yeh WH, Zheng C, Newby GA, et al. Continuous evolution of base editors with expanded target compatibility and improved activity. Nat Biotechnol 2019;37. 10.1038/s41587-019-0193-0.

[15]. Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol 2020;38:883–91. 10.1038/s41587-020-0453-z. [PubMed: 32433547]

[16]. Roth TB, Woolston BM, Stephanopoulos G, Liu DR. Phage-Assisted Evolution of Bacillus methanolicus Methanol Dehydrogenase 2. ACS Synth Biol 2019;8:796–806. 10.1021/acssynbio.8b00481. [PubMed: 30856338]

[17]. Packer MS, Rees HA, Liu DR. Phage-assisted continuous evolution of proteases with altered substrate specificity. Nat Commun 2017;8:956. 10.1038/s41467-017-01055-9. [PubMed: 29038472]

[18]. Brödel AK, Rodrigues R, Jaramillo A, Jaramillo A, Jaramillo A, Isalan M. Accelerated evolution of a minimal 63-amino acid dual transcription factor. Sci Adv 2020;6:1–10. 10.1126/sciadv.aba2728.

[19]. Ye X, Tu M, Piao M, Yang L, Zhou Z, Li Z, et al. Using phage-assisted continuous evolution (PACE) to evolve human PD1. Exp Cell Res 2020;396:112244. 10.1016/j.yexcr.2020.112244. [PubMed: 32860814]

[20]. Bryson DI, Fan C, Guo L-T, Miller C, Söll D, Liu DR. Continuous directed evolution of aminoacyl-tRNA synthetases to alter amino acid specificity and enhance activity. Nat Chem Biol 2017. 10.1038/nchembio.2474.

[21]. DeBenedictis EA, Chory EJ, Gretton D, Wang B, Esvelt K. A high-throughput platform for feedback-controlled directed evolution. BioRxiv 2020:2020.04.01.021022. 10.1101/2020.04.01.021022.** Authors use a liquid handling robotics platform to automate PACE and assays of evolving populations in 96-well replicates, enabling a scale of experimentation that is normally prohibited by the equipment demands of continuous dilution.

[22]. Berman CM, Papa LJ, Hendel SJ, Moore CL, Suen PH, Weickhardt AF, et al. An Adaptable Platform for Directed Evolution in Human Cells. J Am Chem Soc 2018;140:18093–103. 10.1021/jacs.8b10937. [PubMed: 30427676]

[23]. English JG, Olsen RHJ, Lansu K, Patel M, White K, Cockrell AS, et al. VEGAS as a Platform for Facile Directed Evolution in Mammalian Cells. Cell 2019;178:748–761.e17. 10.1016/j.cell.2019.05.051. [PubMed: 31280962] ** Capitalizing on the naturally high error

rate of the RNA alphavirus Sindbis, this viral hypermutation system enables rapid evolution of GOIs encoded on the viral genome in a mammalian cell context. The high error rate enabled evolution of drug resistance in a transcription factor, constitutive activity in GPCRs, and active state-specific GPCR binding nanobodies, each within a week.

[24]. Wang T, Badran AH, Huang TP, Liu DR. Continuous directed evolution of proteins with improved soluble expression. Nat Chem Biol 2018. 10.1038/s41589-018-0121-5.

[25]. Moore CL, Papa LJ, Shoulders MD. A Processive Protein Chimera Introduces Mutations across Defined DNA Regions in Vivo. J Am Chem Soc 2018;140:11560–4. 10.1021/jacs.8b04001. [PubMed: 29991261] * A T7 RNAP fused to a cytosine deaminase demonstrated elevated mutagenesis of user-selected sequences or genes encoded between a T7 promoter and T7 terminators in *E. coli*.

[26]. Chen H, Liu S, Padula S, Lesman D, Griswold K, Lin A, et al. Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. Nat Biotechnol 2019. 10.1038/s41587-019-0331-8.* A T7 RNAP fused to a cytosine deaminase demonstrated elevated mutagenesis of user-selected sequences or genes encoded downstream of a T7 promoter in mammalian cells. Multiple T7 RNAP mutants with variable processivity and multiple cytosine deaminases were evaluated to achieve a highly optimized system.

[27]. Álvarez B, Mencía M, de Lorenzo V, Fernández LÁ. In vivo diversification of target genomic sites using processive T7 RNA polymerase-base deaminase fusions blocked by RNA-guided dCas9. Nat Commun 2020;11:6436. 10.1101/850974. [PubMed: 33353963] * A T7 RNAP fused to either a cytosine or adenine deaminase demonstrated elevated mutagenesis downstream of a T7 promoter in *E. coli*. In addition to the use of T7 terminators to reduce off target mutation, dCas9 was also used to block RNAP, allowing for targeting to be restricted to a specific segment of a gene.

[28]. Park H, Kim S. Gene-specific mutagenesis enables rapid continuous evolution of enzymes in vivo. Nucleic Acids Res 2021:1–10. 10.1093/nar/gkaa1231.* A T7 RNAP fused to a cytosine deaminase demonstrated elevated mutagenesis of user-selected sequences or genes encoded between a T7 promoter and a T7 terminator in *E. coli*. A highly efficient cytosine deaminase was used, demonstrating markedly higher rates of mutation than other similar systems.

[29]. Halperin SO, Tou CJ, Wong EB, Modavi C, Schaffer DV., Dueber JE. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. Nature 2018;560:248–52. 10.1038/s41586-018-0384-8. [PubMed: 30069054] ** A fusion between nickase-Cas9 and an error-prone DNA polymerase produced highly elevated mutation rates in a 50-bp window adjacent to the Cas9 targeting site.

[30]. Tou CJ, Schaffer DV., Dueber JE. Targeted Diversification in the S. cerevisiae Genome with CRISPR-Guided DNA Polymerase i. ACS Synth Biol 2020;9:1911–6. 10.1021/acssynbio.0c00149. [PubMed: 32485105]

[31]. Hess GT, Frésard L, Han K, Lee CH, Li A, Cimprich KA, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat Methods 2016;13:1036–42. 10.1038/nmeth.4038. [PubMed: 27798611]

[32]. Crook N, Abatemarco J, Sun J, Wagner JM, Schmitz A, Alper HS. In vivo continuous evolution of genes and pathways in yeast. Nat Commun 2016;7. 10.1038/ncomms13051.

[33]. Finney-Manchester SP, Maheshri N. Harnessing mutagenic homologous recombination for targeted mutagenesis in vivo by TaGTEAM. Nucleic Acids Res 2013;41:1–10. 10.1093/nar/gkt150. [PubMed: 23143271]

[34]. Ravikumar A, Arrieta A, Liu CC. An orthogonal DNA replication system in yeast. Nat Chem Biol 2014;10:175–7. 10.1038/nchembio.1439. [PubMed: 24487693]

[35]. Camps M, Naukkarinen J, Johnson BP, Loeb LA. Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I. Proc Natl Acad Sci U S A 2003;100:9727–32. 10.1073/pnas.1333928100. [PubMed: 12909725]

[36]. Rix G, Watkins-Dulaney EJ, Almhjell PJ, Boville CE, Arnold FH, Liu CC. Scalable continuous evolution for the generation of diverse enzyme variants encompassing promiscuous activities. Nat Commun 2020;11:1–11. 10.1038/s41467-020-19539-6. [PubMed: 31911652] ** The OrthoRep direct hypermutation system was applied toward evolution of a thermophile homolog of tryptophan synthase toward activity in yeast in 40 replicate populations. This resulted in

highly active diverse variants of the enzyme that demonstrated substantial substrate promiscuity, achieving high % yields in production of several tryptophan analogs. The many independently evolved multi-mutation variants of tryptophan synthase highlight scale and depth in evolutionary search.

[37]. Wellner A, McMahon C, Gilman MSA, Clements JR, Clark S, Nguyen KM, et al. Rapid generation of potent antibodies by autonomous hypermutation in yeast. BioRxiv 2020. 10.1101/2020.11.11.378778.* OrthoRep was combined with yeast surface display and selection via fluorescence activated cell sorting to evolve camelid antibodies (nanobodies) to several targets in as many as 8 independent replicates, yielding high affinity nanobodies with ~5 mutations each. This provides a platform for rapid affinity maturation that supports scale and depth in antibody evolution.

[38]. Zhong Z, Wong BG, Ravikumar A, Arzumanyan GA, Khalil AS, Liu CC. Automated continuous evolution of proteins in vivo. ACS Synth Biol 2020:2020.02.21.960328. 10.1021/acssynbio.0c00135.

[39]. Zhong Z, Ravikumar A, Liu CC. Tunable Expression Systems for Orthogonal DNA Replication. ACS Synth Biol 2018;7:2930–4. 10.1021/acssynbio.8b00400. [PubMed: 30408954]

[40]. Yang KK, Wu Z, Arnold FH. Machine-learning-guided directed evolution for protein engineering. Nat Methods 2019;16:687–94. 10.1038/s41592-019-0496-6. [PubMed: 31308553]

[41]. Stiffler MA, Poelwijk FJ, Brock KP, Stein RR, Riesselman A, Teyra J, et al. Protein Structure from Experimental Evolution. Cell Syst 2020;10:15–24.e5. 10.1016/j.cels.2019.11.008. [PubMed: 31838147]

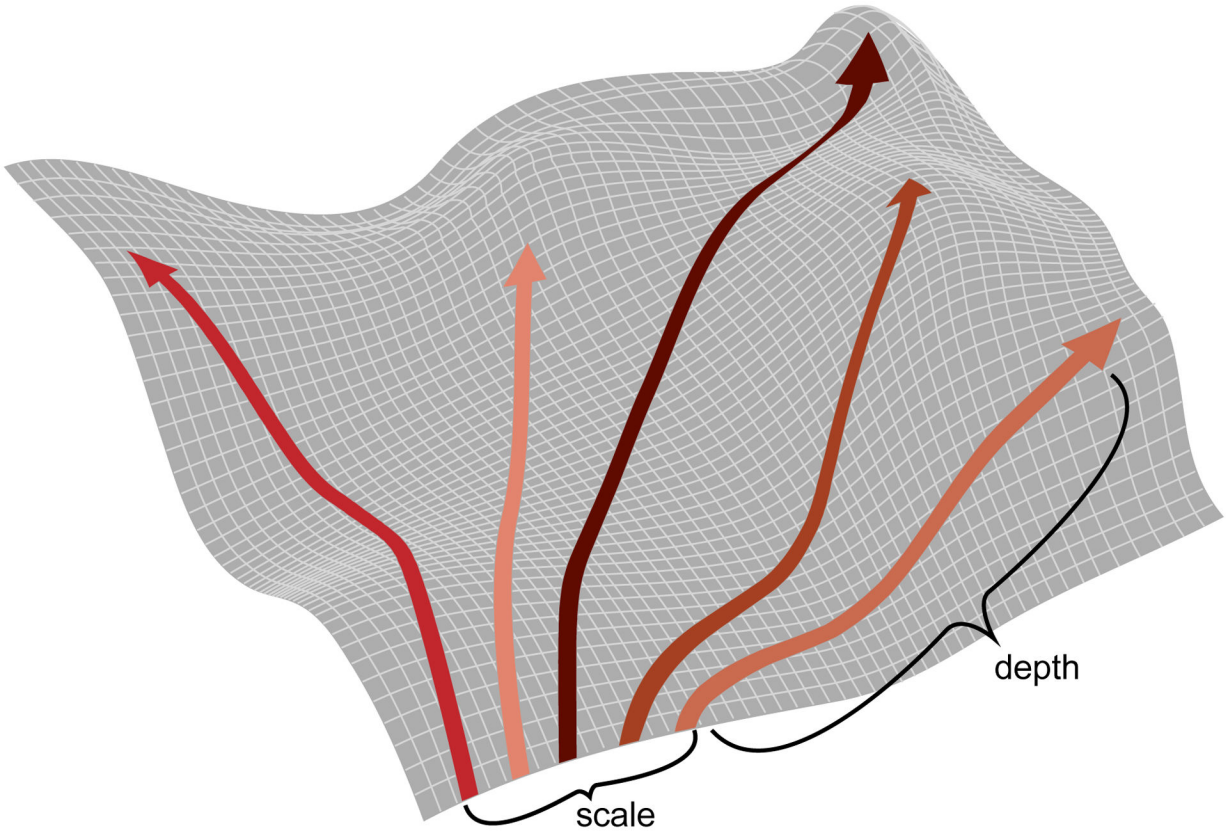

Figure 1. Evolutionary scale and depth.

Separate evolving populations traverse a fitness landscape in which the x- and y-axes represent genotype, and the z-axis represents fitness. A greater number of individual populations (scale) enables exploration of more distinct areas of sequence space, while the capacity to traverse longer mutational pathways (depth) enables exploration of areas of sequence space that are further away from the starting point.

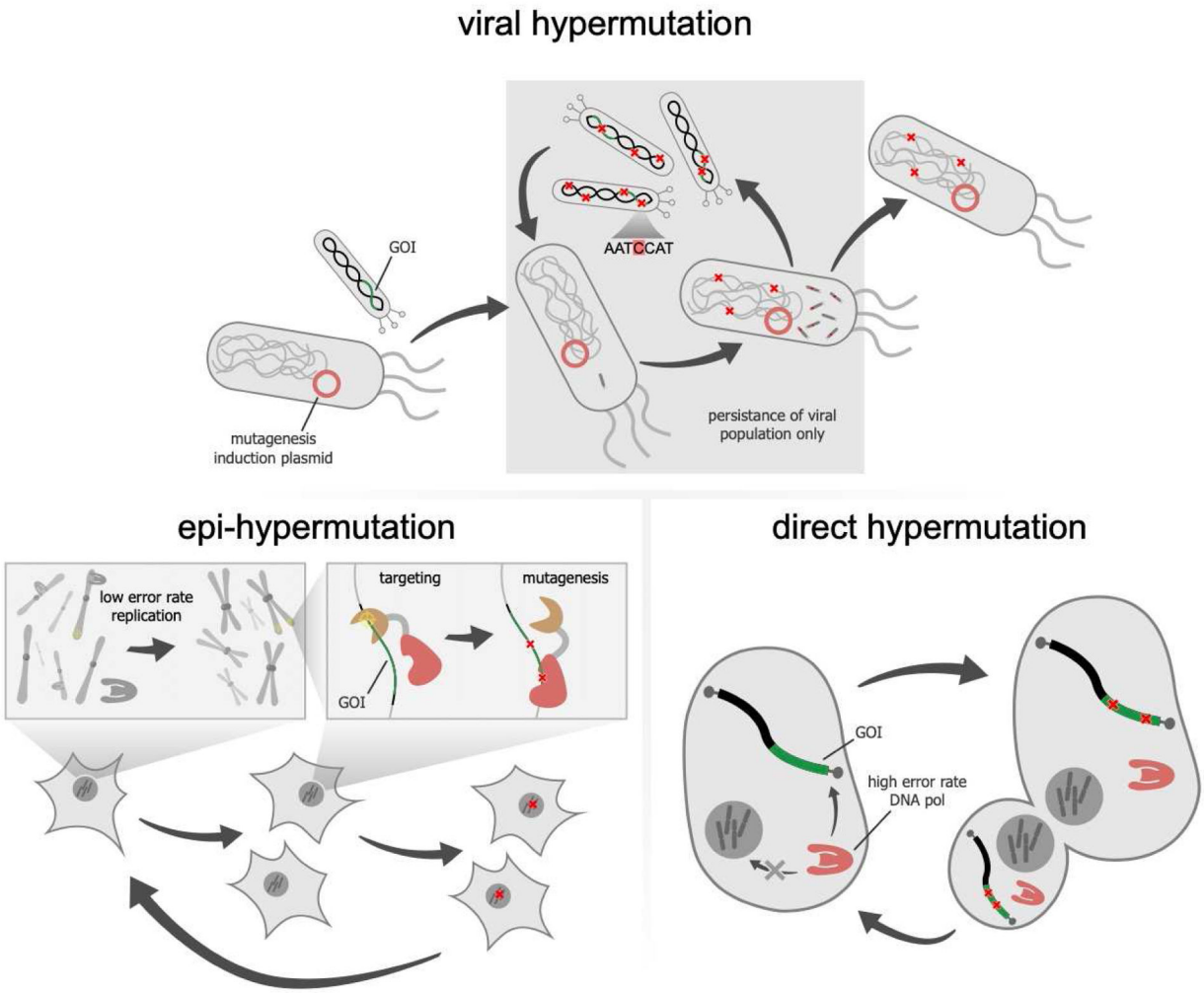

Figure 2.
Categories of *in vivo* hypermutation architectures.