Opinion

Evolution, ecology and the engineered organism: lessons for synthetic biology

Jeffrey M Skerker*†, Julius B Lucks*† and Adam P Arkin*†

Addresses: *Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA. †Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

Correspondence: Adam P Arkin. Email: aparkin@lbl.gov

Abstract

As the scope and complexity of synthetic biology grows, an understanding of evolution and ecology will be critical to its success.

One of the most powerful and controversial aspects of engineering living organisms is that they reproduce, evolve, and interact with their environment. Humans have been engineering plants and animals since the advent of agriculture approximately 12,000 years ago through breeding and artificial selection for their domestication [1]. The evolution of corn from the small grass teosinte [2], or the transformation of the wolf into 'man's best friend' (the dog) [1] are testaments to the success of this approach. We have even 'domesticated' microorganisms, using yeast and bacteria for the production of beer, wine, cheese and yogurt as well as numerous other products we consume every day [3,4].

Although powerful, genetic engineering by classical breeding and selection is slow, and results in a large number of unknown genetic changes that are hard to reconcile and may have unintended secondary effects. What we need is a rational approach to the engineering of biological systems that makes the process fast, cheap and safe, to solve problems in energy, health, agriculture and the environment. First steps towards realizing this aim began with the advent of recombinant DNA technology in the latter half of the 20th century, which created visions of a new era of 'synthetic biology' where novel genes could be designed and constructed for useful purposes [5-7]. Since then we have made incredible advances in our ability to manipulate genes, genomes and organisms, and this has led to a renewed interest in making synthetic biology a reality [8].

A number of recent reviews have been written on the principles and practice of synthetic biology [8-11], but here we focus on the interplay between synthetic biology, evolution and ecology. Evolution teaches us about what solutions nature has evolved for biological problems, how to evolve them further, and how robust they are to change. Ecology gives us information on how our engineered systems will perform once they leave the laboratory and enter an industrial bioreactor (a vessel or tank used for the controlled growth of microorganisms) or the natural environment. As the scope and complexity of synthetic biology grows, we argue that an understanding of evolution and ecology is critical to its success. We have explored some of these ideas in the past [12-14], but here we focus on four practical lessons that serve as a starting point for integrating evolutionary and ecological concepts into synthetic biology research and practice (Figure 1).

Lesson 1: Evolution is a source of functional diversity and modularity

One of the central goals of synthetic biology is to develop genetic elements with encapsulated functions, such as regulatory circuits or environmental sensors, that can be combined to create new pathways with predictable behaviours. Despite our ability to synthesize genes and even genomes [15], we still lack the sophistication to design *de novo* those genetic elements needed for advanced synthetic biology applications. Fortunately, evolution has already provided us with an immense diversity of biomolecular functions that can be used individually or combined by putting together natural functional modules.

Bacteria and archaea represent perhaps the largest reservoirs of new genes and new biochemical functions that can be harnessed by the synthetic biologist. Current estimates of the number of bacterial species range from 1 million to as many as 1 billion [16,17], each representing a unique genetic solution to the environmental challenges posed by diverse ecological niches. This incredible diversity of species in turn encodes a vast universe of protein functions. As of October 2009, there were 11,912 protein families in the Pfam database alone [18,19]. Despite this large number, our sampling of protein function is still incomplete, and many new activities still remain to be discovered in nature [20]. In addition, there is probably a vast array of non-coding RNA functions and DNA regulatory sequences that would serve as useful genetic elements for synthetic biology but which are difficult to detect by typical sequencing methods because of their fast rate of evolution.
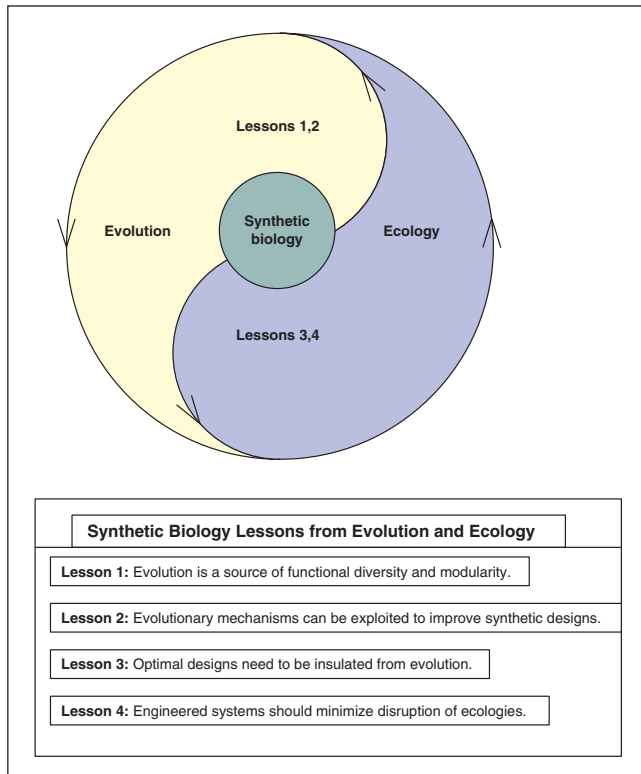
Figure 1

Ecological forces drive evolution, which in turn influences ecologies. This cycle creates a diverse array of functions that can be used in synthetic designs. Individual functions may be combined and evolved in the laboratory to create new synthetic systems that may ultimately enter natural ecologies.

This plethora of gene functions derived from evolution has not gone unnoticed, and it has been standard practice in genetic engineering to mix and match genes from many organisms. One driving force behind this has been to make bacteria such as *Escherichia coli* into 'chemical factories' for the production of drugs, fine chemicals and other commercially important compounds. Recent successes include the production of amorpha-4,11-diene (a precursor of the antimalarial drug artemisinin) [21], the production of putrescine (used for the production of the plastic nylon-4,6) [22] and the production of 4-hydroxyvalerate (which can be converted into polyesters and other plastics) [23]. Along with other examples, such as the production of the amino acids L-valine [24] and L-threonine [25] from engineered bacteria, such successes have founded a field of metabolic engineering that strives to leverage the metabolic flexibility of microbes to convert simple inputs such as sugars to desirable complex compounds [11,26]. For many applications, the gene function or enzymatic chemistry is already available in nature, but if not, there are experimental strategies that can circumvent this problem (see Lesson 2).

Even if a gene function exists in nature, our ability to use it to engineer complex biological systems with new composite functions relies on the modularity inherent in naturally evolved systems. Modular biological systems are composed of functional domains that can be individually swapped or altered to change the overall characteristics of the system. Examples of modularity in biology abound at nearly all scales, and include basic gene regulation elements (promoters and binding sites for transcription factors), protein domains, macromolecular protein complexes, and cellular regulatory networks [27-31]. A number of compelling studies have demonstrated that modularity in biological systems arises under selection in a changeable environment [32,33], and modularity seems to have been selected because it makes 'rewiring' on an evolutionary timescale more effective [34]. The ability to rewire natural biological systems makes nature a vast source of modular 'parts' for the synthetic biologist. However, we must be careful to obey the rules of modularity and domain boundaries that nature uses. Understanding these rules, at both the molecular and organismal levels, is currently an active area of research [35-37].

Lesson 2: Evolutionary mechanisms can be exploited to improve synthetic designs

As discussed above, evolution has provided a vast universe of genes and factorable modules that can be harnessed by the synthetic biologist to engineer new biological systems. In the simplest scenario, the desired function can be used 'as is' without any further modification. However, many synthetic designs require that we modify or tweak a gene function, such as altering an enzyme activity or changing a regulatory element. In extreme cases we need a gene function or activity that does not actually exist in nature. For example, incorporation of unnatural amino acids (for example, *p*-boronophenylalanine) into proteins is now possible using tRNA synthetases created in the laboratory and this enables the site-specific modification of proteins using boronate-based chemistry [38]. Enzymes that catalyze the Kemp elimination reaction have been produced by using a combination of computational protein design and molecular evolution (see below) [39].

Fortunately, a suite of experimental techniques exists that can create new gene function in the laboratory on the basis of a deep understanding of the fundamental mechanisms of evolutionary change - variation by mutation and recombination, differential reproduction and heredity. These so-called '*in vitro* evolution' methods have been applied successfully to DNA, RNA and proteins [40-44]. Like classical breeding and artificial selection, they are iterative processes, involving rounds of library creation, screening and selection. Here, we focus on the library-creation step because it has benefited most from our knowledge of evolutionary mechanisms. The traditional approach to library creation involves generating random

variation, for which there are a number of standard methods such as random DNA synthesis, error-prone PCR, chemical mutagenesis or the use of mutator strains.

Random mutagenesis by itself is inefficient, and computer simulations of evolution have demonstrated that a low level of point mutation plus recombination is an optimal strategy for creating diversity [45]. This observation led to the development of gene shuffling, which is a powerful technique for the rapid evolution of protein function [44]. In this process, random DNA fragmentation and reassembly by PCR is used to simulate recombination in the laboratory. Gene shuffling has been used to increase enzyme activity [46], alter substrate specificity [47] and improve the properties of green fluorescent protein [48].

Gene shuffling has been further expanded to genome shuffling, which combines mutagenesis with protoplast fusion to rapidly evolve microbes for the purpose of strain improvement [49]. Because multiple advantageous mutations may be combined during each round of mutagenesis and protoplast fusion, genome shuffling has proved superior to classical methods for strain improvement (that is, mutagenesis plus selection); however, it still suffers from the limitation that the genetic basis for the improvement is not known. Most recently, a method for rapid genome engineering in bacteria has been developed, called multiplexed automated genome engineering (MAGE), that allows at least 20 directed genomic mutations at once by using mutagenic oligos [50]. The combination of MAGE, genome shuffling and the means to vary the selection pressure to enable bouts of random mutation without selection (that is, neutral evolution) [51] might be a powerful approach to the more rapid evolution of strains with desired characteristics. This method could be applied to developing strains with increased metabolic flux through an engineered pathway, or to improve tolerance to environmental stresses, such as pH or temperature. The take-home lesson is that evolutionary mechanisms have provided a powerful set of experimental tools for the rapid engineering of biological function. As we continue to understand how natural systems evolve, we can further exploit these processes for engineering genes and genomes in the laboratory.

Ultimately, even laboratory evolution is not sufficient for the engineering of complex biological systems. As designs become more complex, directed evolution at multiple genetic loci starts to resemble classical breeding and selection - where we do not understand the connection between genotype and phenotype. Furthermore, these evolution-based strategies require that we have selections or screens for the desired traits, which rapidly becomes too difficult as we move beyond the simplest applications. We envisage that synthetic biologists will use a hybrid approach starting with rational design using modular parts

(Lesson 1), followed by organism-level evolution around the designed genetic architecture of the system for final optimization [52].

Lesson 3: Optimal designs need to be insulated from evolution

Even though we may use evolution as a tool to create novel function and optimize designs, we must be aware that its driving force for change does not stop when we deploy a system in a bioreactor or in the environment. Once a system is ready for use we would like to halt evolution, or at least minimize it, so that our system can perform without diverging from its original specifications. All the mechanisms of evolutionary change that were exploited to develop our system now need to be counteracted. This is quite a challenge and requires a focus on the two main sources of evolutionary change in nature - horizontal gene transfer (HGT) and random mutation.

One strategy for minimizing evolution is to prevent HGT. HGT can occur in three ways: by conjugation, transduction or transformation [53]. Conjugation is the transfer of genetic material (often a plasmid) between bacteria through direct cell-to-cell contact. Many plasmids encode their own mobilization and transfer functions and can move between bacteria by conjugation. In the early days of recombinant DNA research it was recognized that these plasmid sequences could be deleted, thus preventing their transfer [54]. In addition, cell-envelope proteins that are necessary for conjugation can be mutated.

By contrast, transduction and transformation enable transfer of DNA without cell contact. Transduction is mediated by bacteriophages whereas transformation is the uptake of free DNA from the environment. Transduction can be prevented by mutating a wide-range of bacteriophage receptors to give phage-resistant strains. Ideally, we could develop broad-range phage resistance, and there is evidence that such mutations exist. In one example, three mutants of *Streptococcus thermophilus* were identified that were resistant to 14 phages after screening for resistance to just one lytic phage, Sfi19 [55]. Other strategies for broad-range phage resistance could include engineering the CRISPR (clustered, regularly interspaced, short palindromic repeat) genes, which have recently been hypothesized to be a bacterial 'immune system' that targets the degradation and silencing of foreign DNA [56].

The third mechanism of HGT involves natural transformation, and one strategy to prevent this is to mutate *com* genes and thus prevent uptake of DNA from the environment [57]. Competence (*com*) genes encode a set of proteins that are localized in the bacterial cell envelope and are critical for processing and uptake of DNA. If all else fails and foreign DNA does get inside the cell of an engineered strain it could be prevented from integrating

into the genome by using a *rec-* strain background or by installing a strong restriction/modification system. Recombination (*rec*) genes are essential for homologous recombination, so a *rec-* strain would not be able to recombine the foreign DNA into its chromosome. Restriction/modification systems degrade incoming DNA that is not specifically 'marked' by methylation by the host bacterium, and so would block HGT before the recombination step.

A second strategy for minimizing evolution is to modulate the mutation rate. Defects in the mismatch repair system, for example, dramatically increase the mutation rate. The mismatch repair system recognizes mispaired nucleotides that arise during errors in DNA replication and recombination and recruits the necessary enzymes to repair the mistake. Many of these genes were first identified as mutator (*mut*) genes, which led to an increase in mutation frequency when deleted. For example, loss of function of *mutS* or *mutL* leads to a 10^2- to 10^3-fold increase in the frequency of transition and frameshift mutations [58]. By contrast, overexpression of MutS or MutL leads to a decrease in the mutation frequency, and this could be one strategy for minimizing evolution [59]. This study suggested that other genes might exist that increase the mutation rate when overexpressed. In this regard, a multicopy genetic screen in *E. coli* identified 15 loci that when overexpressed led to a mutator-like phenotype, and 12 of these were previously uncharacterized [60]. In theory, every mechanism that nature uses to increase the mutation rate could be reversed by overexpression or deletion of the appropriate genes, although this general idea remains to be tested.

Lesson 4: Engineered systems should minimize disruption of ecologies

At present, the cutting-edge of genetic manipulation is in metabolic engineering [21,22,50]. The bacterium *E. coli* has long been a workhorse in this field, largely because of its ease of genetic manipulation and the large amount of knowledge and resources accumulated. However, when we start to consider applications of synthetic biology beyond the bioreactor, such as bioremediation or therapeutic use in the human body, we must consider the complex nature of these environments. In particular, we must ensure that our engineered biological system works to specification without unintended disruptions to the natural ecology of the environment or human host, and that it can be easily identified and removed if necessary.

Bioremediation is the use of living organisms to return an ecosystem to its natural state after toxic contamination. Ever since the advent of recombinant DNA technology, the use of genetically modified (GM) organisms for bioremediation has been a holy grail. Unfortunately, most attempts at using GM bacteria for bioremediation have failed because the engineered strain had reduced fitness

and competed poorly with indigenous microbial communities [61]. Although *E. coli* is a natural choice for use as a chemical factory in a laboratory bioreactor, it makes no sense to engineer a bacterium that normally resides in the human gut for bioremediation of a toxic-waste site. It is more appropriate to engineer organisms that are derived directly from the target ecology. This is not without its challenges, however.

The industrial chemical 2-chlorotoluene is produced in large amounts and is used in a variety of consumer products. It is toxic to aquatic environments and humans, is inert to chemical hydrolysis in environmental conditions, and is therefore an interesting target for microbial bioremediation. Initial attempts at engineering soil-derived *Pseudomonas* species for 2-chlorotoluene degradation [62] failed because of the complex nature of environmental influences on gene regulation [61]. Given the tools of synthetic and systems biology, there is renewed hope that such problems, which are due to strong coupling of engineered organisms to target ecologies, can now be overcome.

One of the principal areas that needs development is the characterization of organisms for use in different bioremediation applications. This will mean identifying the key organisms responsible for the biotransformation process of interest, isolating and culturing their communities in the laboratory so they can be engineered for enhanced bioremediation and ecological stabilization, and then reintroducing them into the environment. Although there will be many difficulties in implementing this strategy, metagenomic techniques have greatly advanced the identification of the complex microbial communities that exist in the environment [63]. Recent work also shows that we now have the technology to manipulate previously genetically intractable systems: the complete genome of *Mycoplasma mycoides* was transferred into yeast, altered using yeast genetic tools, and then transplanted back into a Mycoplasma cell to yield a new *M. mycoides* strain [64].

When considering the 'real-world' applications of synthetic biology such as bioremediation the environmental impact and safety of the engineered organisms are important considerations. Introducing an engineered organism into a bioremediation site can be thought of as purposefully introducing an invasive species. Whether it is successful and competes with the native organisms depends on its relative fitness and its ability to evolve and adapt to its environment [65]. Even though these engineered strains may be less fit and perhaps even less effective than the native species, they have the advantage that they can be engineered with a 'leash' or other system to prevent their unwanted spread. Such safeguards have been in place since the beginning of recombinant DNA research, and have been further developed over the years [66-68].

One worry is that engineered strains will evolve around introduced safeguards, and Lesson 3 highlights ways in which we might address this possibility. Even so, the DNA of the engineered system could still be released after cell death and could be taken up by other bacteria in the ecosystem by natural transformation. How can we prevent the spread of engineered DNA by this route? If we could engineer strains that use an alternative genetic code, then even if the DNA gets transferred into other bacteria, translation would produce a functionless protein. This would similarly prevent 'natural' DNA accidentally imported into the engineered organism from being expressed. Alternative genetic codes exist in mitochondria and ciliates [69], and there are many examples of artificial alternative codes based on the tRNA synthetase system first developed by Schultz and co-workers [70]. There are even translation systems that work orthogonally to the natural host system, and that would not function in bacteria that did not have the correct ribosomal apparatus [71].

The interplay between synthetic biology, evolution and ecology

Whatever the strategy we choose to follow to prevent unwanted spread, understanding the interplay between ecology and synthetic biology is critical to predicting how an engineered system might evolve in and interact with a natural environment. Once we take our engineered system out of the laboratory, whether into an industrial fermentation tank, the environment (for example, bioremediation) or a human host (for example, a therapeutic organism), we need to understand how our design will evolve according to the selective pressures of its environment, and how it will affect the ecology of its environment. The synthetic biologist is constantly in a state of tension - on one hand, exploiting the mechanisms of evolution to engineer more complex biological systems, and on the other trying to keep the design robust to evolution once it is released. Once introduced into the environment, the engineered biological system also needs to 'play well with others' and not adversely disrupt the natural ecology. There are complex considerations, both ethical and legal, in releasing genetically modified bacteria into the environment for study or application [72] or even in disclosing the technology that enables the engineering of organisms able to survive in the outside world. However, having a deeper understanding of the interplay between evolution, ecology and synthetic biology will be critical in moving our designs 'beyond the bioreactor' into the real world where they can safely and effectively benefit society.

Acknowledgements

References

1. Driscoll CA, Macdonald DW, O'Brien SJ: **From wild animals to domestic pets, an evolutionary view of domestication.** *Proc Natl Acad Sci USA* 2009, **106 Suppl 1:**9971-9978.

2. White S, Doebley J: **Of genes and genomes and the origin of maize.** *Trends Genet* 1998, **14:**327-332.

3. Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goltsman E, Mazur M, Pusch GD, Fonstein M, Overbeek R, Kyprides N, Purnelle B, Prozzi D, Ngui K, Masuy D, Hancy F, Burteau S, Boutry M, Delcour J, Goffeau A, Hols P: **Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*.** *Nat Biotechnol* 2004, **22:**1554-1558.

4. Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, Shimonaga T, Hattori M, Ashikari T: **Genome sequence of the lager brewing yeast, an interspecies hybrid.** *DNA Res* 2009, **16:**115-129.

5. Lobban PE, Kaiser AD: **Enzymatic end-to end joining of DNA molecules.** *J Mol Biol* 1973, **78:**453-471.

6. Jackson DA, Symons RH, Berg P: **Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*.** *Proc Natl Acad Sci USA* 1972, **69:**2904-2909.

7. Cohen SN, Chang AC, Boyer HW, Helling RB: **Construction of biologically functional bacterial plasmids *in vitro*.** *Proc Natl Acad Sci USA* 1973, **70:**3240-3244.

8. Purnick PE, Weiss R: **The second wave of synthetic biology: from modules to systems.** *Nat Rev Mol Cell Biol* 2009, **10:**410-422.

9. Salis H, Tamsir A, Voigt C: **Engineering bacterial signals and sensors.** *Contrib Microbiol* 2009, **16:**194-225.

10. Lucks JB, Qi L, Whitaker WR, Arkin AP: **Toward scalable parts families for predictable design of biological circuits.** *Curr Opin Microbiol* 2008, **11:**567-573.

11. Keasling JD: **Synthetic biology for synthetic chemistry.** *ACS Chem Biol* 2008, **3:**64-76.

12. Wolf DM, Arkin AP: **Motifs, modules and games in bacteria.** *Curr Opin Microbiol* 2003, **6:**125-134.

13. McAdams HH, Srinivasan B, Arkin AP: **The evolution of genetic regulatory systems in bacteria.** *Nat Rev Genet* 2004, **5:**169-178.

14. Arkin AP, Fletcher DA: **Fast, cheap and somewhat in control.** *Genome Biol* 2006, **7:**114.

15. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA 3rd, Smith HO: **Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome.** *Science* 2008, **319:**1215-1220.

16. Quince C, Curtis TP, Sloan WT: **The rational exploration of microbial diversity.** *ISME J* 2008, **2:**997-1006.

17. Gans J, Wolinsky M, Dunbar J: **Computational improvements reveal great bacterial diversity and high metal toxicity in soil.** *Science* 2005, **309:**1387-1390.

18. **Pfam database** [http://pfam.sanger.ac.uk]

19. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A: **The Pfam protein families database.** *Nucleic Acids Res* 2008, **36:**D281-D288.

20. Yooseph S, Sutton G, Rusch DB, Halpern AL, Williamson SJ, Remington K, Eisen JA, Heidelberg KB, Manning G, Li W, Jaroszewski L, Cieplak P, Miller CS, Li H, Mashiyama ST, Joachimiak MP, van Belle C, Chandonia JM, Soergel DA, Zhai Y, Natarajan K, Lee S, Raphael BJ, Bafna V, Friedman R, Brenner SE, Godzik A, Eisenberg D, Dixon JE, Taylor SS, *et al.*: **The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families.** *PLoS Biol* 2007, **5:**e16.

21. Tsuruta H, Paddon CJ, Eng D, Lenihan JR, Horning T, Anthony LC, Regentin R, Keasling JD, Renninger NS, Newman JD: **High-level production of amorpha-4,11-diene, a precursor**

of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* 2009, **4:**e4489.

22. Qian ZG, Xia XX, Lee SY: **Metabolic engineering of *Escherichia coli* for the production of putrescine: a four carbon diamine.** *Biotechnol Bioeng* 2009, **104:**651-662.

23. Martin CH, Prather KL: **High-titer production of monomeric hydroxyvalerates from levulinic acid in *Pseudomonas putida*.** *J Biotechnol* 2009, **139:**61-67.

24. Park JH, Lee KH, Kim TY, Lee SY: **Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation.** *Proc Natl Acad Sci USA* 2007, **104:**7797-7802.

25. Lee KH, Park JH, Kim TY, Kim HU, Lee SY: **Systems metabolic engineering of *Escherichia coli* for L-threonine production.** *Mol Syst Biol* 2007, **3:**149.

26. Khosla C, Keasling JD: **Metabolic engineering for drug discovery and development.** *Nat Rev Drug Discov* 2003, **2:**1019-1025.

27. Price MN, Dehal PS, Arkin AP: **Horizontal gene transfer and the evolution of transcriptional regulation in *Escherichia coli*.** *Genome Biol* 2008, **9:**R4.

28. Gordley RM, Gersbach CA, Barbas CF 3rd: **Synthesis of programmable integrases.** *Proc Natl Acad Sci USA* 2009, **106:** 5053-5058.

29. Park SH, Zarrinpar A, Lim WA: **Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms.** *Science* 2003, **299:**1061-1064.

30. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KL, Keasling JD: **Synthetic protein scaffolds provide modular control over metabolic flux.** *Nat Biotechnol* 2009, **27:**753-759.

31. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U: **Network motifs: simple building blocks of complex networks.** *Science* 2002, **298:**824-827.

32. Hintze A, Adami C: **Evolution of complex modular biological networks.** *PLoS Comput Biol* 2008, **4:**e23.

33. Kashtan N, Mayo AE, Kalisky T, Alon U: **An analytically solvable model for rapid evolution of modular structure.** *PLoS Comput Biol* 2009, **5:**e1000355.

34. Wagner GP, Pavlicev M, Cheverud JM: **The road to modularity.** *Nat Rev Genet* 2007, **8:**921-931.

35. Skerker JM, Perchuk BS, Siryaporn A, Lubin EA, Ashenberg O, Goulian M, Laub MT: **Rewiring the specificity of two-component signal transduction systems.** *Cell* 2008, **133:**1043-1054.

36. Zarrinpar A, Park SH, Lim WA: **Optimization of specificity in a cellular protein interaction network by negative selection.** *Nature* 2003, **426:**676-680.

37. Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM: **Genome-wide experimental determination of barriers to horizontal gene transfer.** *Science* 2007, **318:**1449-1452.

38. Brustad E, Bushey ML, Lee JW, Groff D, Liu W, Schultz PG: **A genetically encoded boronate-containing amino acid.** *Angew Chem Int Ed Engl* 2008, **47:**8220-8223.

39. Röthlisberger D, Khersonsky O, Wollacott AM, Jiang L, DeChancie J, Betker J, Gallaher JL, Althoff EA, Zanghellini A, Dym O, Albeck S, Houk KN, Tawfik DS, Baker D: **Kemp elimination catalysts by computational enzyme design.** *Nature* 2008, **453:**190-195.

40. Ellington AD, Szostak JW: **In vitro selection of RNA molecules that bind specific ligands.** *Nature* 1990, **346:**818-822.

41. Ellington AD, Szostak JW: **Selection *in vitro* of single-stranded DNA molecules that fold into specific ligand-binding structures.** *Nature* 1992, **355:**850-852.

42. Roberts RW, Szostak JW: **RNA-peptide fusions for the *in vitro* selection of peptides and proteins.** *Proc Natl Acad Sci USA* 1997, **94:**12297-12302.

43. Tuerk C, Gold L: **Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** *Science* 1990, **249:**505-510.

44. Stemmer WP: **Rapid evolution of a protein *in vitro* by DNA shuffling.** *Nature* 1994, **370:**389-391.

45. Forrest S: **Genetic algorithms: principles of natural selection applied to computation.** *Science* 1993, **261:**872-878.

46. Songsiriritthigul C, Pesatcha P, Eijsink VG, Yamabhai M: **Directed evolution of a *Bacillus* chitinase.** *Biotechnol J* 2009, **4:**501-509.

47. Zhang JH, Dawes G, Stemmer WP: **Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening.** *Proc Natl Acad Sci USA* 1997, **94:**4504-4509.

48. Crameri A, Whitehorn EA, Tate E, Stemmer WP: **Improved green fluorescent protein by molecular evolution using DNA shuffling.** *Nat Biotechnol* 1996, **14:**315-319.

49. Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WP, del Cardayre SB: **Genome shuffling leads to rapid phenotypic improvement in bacteria.** *Nature* 2002, **415:**644-646.

50. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM: **Programming cells by multiplex genome engineering and accelerated evolution.** *Nature* 2009, **460:**894-898.

51. Gupta RD, Tawfik DS: **Directed enzyme evolution via small and effective neutral drift libraries.** *Nat Methods* 2008, **5:** 939-942.

52. Yokobayashi Y, Weiss R, Arnold FH: **Directed evolution of a genetic circuit.** *Proc Natl Acad Sci USA* 2002, **99:**16587-16591.

53. Norman A, Hansen LH, Sorensen SJ: **Conjugative plasmids: vessels of the communal gene pool.** *Philos Trans R Soc Lond B Biol Sci* 2009, **364:**2275-2289.

54. Berg P, Baltimore D, Brenner S, Roblin RO, Singer MF: **Summary statement of the Asilomar conference on recombinant DNA molecules.** *Proc Natl Acad Sci USA* 1975, **72:** 1981-1984.

55. Lucchini S, Sidoti J, Brussow H: **Broad-range bacteriophage resistance in *Streptococcus thermophilus* by insertional mutagenesis.** *Virology* 2000, **275:**267-277.

56. Marraffini LA, Sontheimer EJ: **CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.** *Science* 2008, **322:**1843-1845.

57. Dubnau D: **DNA uptake in bacteria.** *Annu Rev Microbiol* 1999, **53:**217-244.

58. Saint-Ruf C, Matic I: **Environmental tuning of mutation rates.** *Environ Microbiol* 2006, **8:**193-199.

59. Zhao J, Winkler ME: **Reduction of GC → TA transversion mutation by overexpression of MutS in *Escherichia coli* K-12.** *J Bacteriol* 2000, **182:**5025-5028.

60. Yang H, Wolff E, Kim M, Diep A, Miller JH: **Identification of mutator genes and mutational pathways in *Escherichia coli* using a multicopy cloning approach.** *Mol Microbiol* 2004, **53:**283-295.

61. de Lorenzo V: **Recombinant bacteria for environmental release: what went wrong and what we have learnt from it.** *Clin Microbiol Infect* 2009, **15 Suppl 1:**63-65.

62. Haro MA, de Lorenzo V: **Metabolic engineering of bacteria for environmental applications: construction of *Pseudomonas* strains for biodegradation of 2-chlorotoluene.** *J Biotechnol* 2001, **85:**103-113.

63. García Martín H, Ivanova N, Kunin V, Warnecke F, Barry KW, McHardy AC, Yeates C, He S, Salamov AA, Szeto E, Dalin E, Putnam NH, Shapiro HJ, Pangilinan JL, Rigoutsos I, Kyrpides NC, Blackall LL, McMahon KD, Hugenholtz P: **Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities.** *Nat Biotechnol* 2006, **24:**1263-1269.

64. Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, Noskov VN, Denisova EA, Gibson DG, Assad-Garcia N, Alperovich N, Thomas DW, Merryman C, Hutchison CA 3rd, Smith HO, Venter JC, Glass JI: **Creating bacterial strains from genomes that have been cloned and engineered in yeast.** *Science* 2009, **325:**1693-1696.

65. Strayer DL, Eviner VT, Jeschke JM, Pace ML: **Understanding the long-term effects of species invasions.** *Trends Ecol Evol* 2006, **21:**645-651.

66. Knudsen SM, Karlstrom OH: **Development of efficient suicide mechanisms for biological containment of bacteria.** *Appl Environ Microbiol* 1991, **57:**85-92.

67. Schweder T, Hofmann K, Hecker M: ***Escherichia coli* K12 relA strains as safe hosts for expression of recombinant DNA.** *Appl Microbiol Biotechnol* 1995, **42:**718-723.

68. Steidler L, Neirynck S, Huyghebaert N, Snoeck V, Vermeire A, Goddeeris B, Cox E, Remon JP, Remaut E: **Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10.** *Nat Biotechnol* 2003, **21:**785-789.

69. Salas-Marco J, Fan-Minogue H, Kallmeyer AK, Klobutcher LA, Farabaugh PJ, Bedwell DM: **Distinct paths to stop codon reassignment by the variant-code organisms *Tetrahymena* and *Euplotes*.** *Mol Cell Biol* 2006, **26:**438-447.

70. Liu DR, Magliery TJ, Schultz PG: **Characterization of an 'orthogonal' suppressor tRNA derived from *E. coli* tRNA2(Gln).** *Chem Biol* 1997, **4:**685-691.

71. Rackham O, Chin JW: **A network of orthogonal ribosome x mRNA pairs.** *Nat Chem Biol* 2005, **1:**159-166.

72. Tiedje JM, Colwell RK, Grossman YL, Hodson RE, Lenski RE, Mack RN, Regal PJ: **The planned introduction of genetically engineered organisms: ecological considerations and recommendations.** *Ecology* 1989, **70:**298-315.