Addressing evolutionary questions with synthetic biology

Florian Baier and Yolanda Schaerli

Department of Fundamental Microbiology, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland

Correspondence: florian.baier@unil.ch; yolanda.schaerli@unil.ch

Abstract

Synthetic biology emerged as an engineering discipline to design and construct artificial biological systems. Synthetic biological designs aim to achieve specific biological behavior, which can be exploited for biotechnological, medical and industrial purposes. In addition, mimicking natural systems using well-characterized biological parts also provides powerful experimental systems to study evolution at the molecular and systems level. A strength of synthetic biology is to go beyond nature's toolkit, to test alternative versions and to study a particular biological system and its phenotype in isolation and in a quantitative manner. Here, we review recent work that implemented synthetic systems, ranging from simple regulatory circuits, rewired cellular networks to artificial genomes and viruses, to study fundamental evolutionary concepts. In particular, engineering, perturbing or subjecting these synthetic systems to experimental laboratory evolution provides a mechanistic understanding on important evolutionary questions, such as: Why did particular regulatory networks topologies evolve and not others? What happens if we rewire regulatory networks? Could an expanded genetic code provide an evolutionary advantage? How important is the structure of genome and number of chromosomes? Although the field of evolutionary synthetic biology is still in its teens, further advances in synthetic biology provide exciting technologies and novel systems that promise to yield fundamental insights into evolutionary principles in the near future.

1. Introduction

Evolutionary biology traditionally studies past or present organisms to reconstruct past evolutionary events with the aim to explain and predict their evolution. However, understanding evolution and why life evolved the way it did might require going beyond solutions found in nature. As the evolutionist J. M. Smith states in 1992 in respect to the development of computational models of artificial evolving systems: '*So far, we have been able to study only one evolving system, and we cannot wait for interstellar flight to provide us with a second. If we want to discover generalizations about evolving systems, we will have to look at artificial ones.*' (J. M. Smith 1992). Now, with advances in synthetic, systems and computational biology it is actually possible to design, create and study artificial (synthetic) biological systems (Hutchison et al. 2016; Cameron et al. 2014). The interdisciplinary field of synthetic biology essentially started with the publication of the first synthetic regulatory networks in 2000 (Elowitz & Leibler 2000; Gardner et al. 2000; Becskei & Serrano 2000) and since then expanded to the design and construction of more complex genetic circuits (Bashor & Collins 2018), organelles (Lee et al. 2018) and even whole genomes (Hutchison et al. 2016) and cells (Blain & Szostak 2014). The design and construction of synthetic biological systems is usually focused on new and desirable metabolic, sensory, regulatory and physical capabilities, which are of particular interest for biotechnological and medical applications (Xie & Fussenegger 2019; Nielsen & Keasling 2016; Tang et al. 2020). Synthetic systems are built by combining, co-opting and modifying biological parts that are implemented in a biological host "chassis" (Y.-H. Wang et al. 2013). A desirable feature of synthetic systems is their (at least partial) orthogonality, meaning that their functionality is not affecting the host's regulation and fitness (C. C. Liu et al. 2018). This allows manipulating, tuning and recording the function of synthetic biological systems independently and without causing much undesirable side effects. In contrast, studying natural systems is often challenging because many biological phenotypes are difficult to disentangle, quantify and characterize. Even for the well-studied single-cell model organisms such as *E. coli* and yeast, we often poorly understand how genes are functionally interconnected and contribute to particular phenotypes (Paaby & Rockman 2013).

The ability to design and build synthetic biological systems that achieve a specific desired phenotype already demonstrates a significant knowledge about its functionality, which is summarized by Richard Feynman's famous quote: "*What I cannot create, I do not understand*" (written on his blackboard at the time of his death in February 1988). One step further is to not only to build a system with a particular purpose in mind, but also to study it and learn how it behaves when it is perturbed and/or evolving. Directed evolution mimics the process of diversification and natural selection that resembles Darwinian evolution under well-defined conditions and is frequently used to optimize proteins and enzymes towards specific functions (Zeymer & Hilvert 2018). Beyond its applied side, the ability to control each parameter in directed evolution experiments has substantially contributed

towards a more fundamental understanding of the dynamics and constraints of molecular evolution (Kaltenbach & Tokuriki 2014; Arnold 2010; Dean & Thornton 2007). Similarly, the well-defined properties of synthetic biological systems and the ability to manipulate and report the system's behavior have motivated researchers to explore fundamental biological and evolutionary questions with synthetic biology (Peisajovich 2012; Davidson et al. 2012; de Lorenzo 2018; Bashor & Collins 2018; Simon et al. 2019). Although the exciting synthesis of evolutionary and synthetic biology is still young, it benefits from existing technologies that have been used to debug and optimize imperfect designs of synthetic systems (Yokobayashi et al. 2002; Cobb et al. 2012; Haseltine & Arnold 2007).
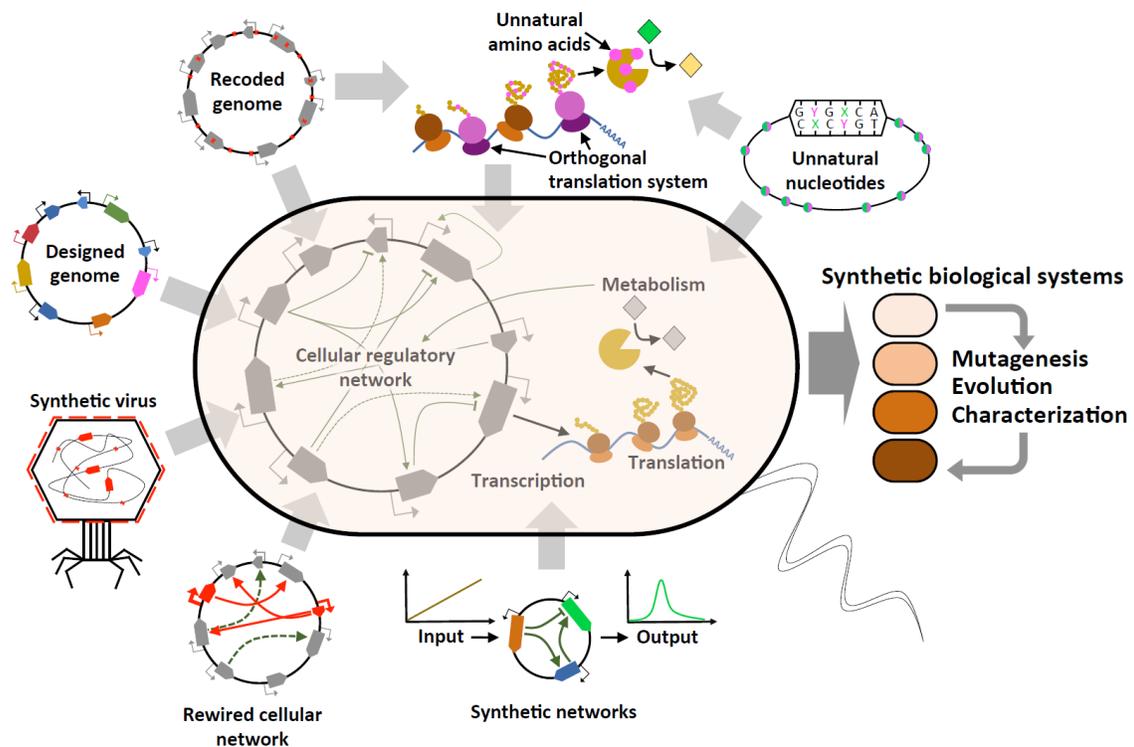
Figure 1. Overview of synthetic biological systems discussed in this chapter. Bottom-up designed and built synthetic biological systems are implemented into a host cell (center) that executes their function using its internal cellular machinery and resources. Synthetic biological systems are amenable for characterization, perturbations and evolution.

In the following, we will highlight a selection of studies that apply synthetic biology tools to understand evolutionary dynamics at the systems level. In particular, we focus on examples of synthetic systems, ranging from simple synthetic regulatory networks (section 2), rewired gene regulatory networks (section 3) and extended genetic codes to synthetic viruses and genomes (section 4) (summarized in **Figure 1** as an graphical overview). In section 5, we end with a discussion on how further advances in synthetic biology will pave

the way for a deeper understanding of life and evolutionary principles and how synthetic biology and evolutionary systems biology can benefit from each other in the future.

2. Synthetic Regulatory Networks

Regulatory networks control the spatial and temporal expression of downstream genes through interactions between DNA, RNA, proteins and/or metabolites (Rockman & Kruglyak 2006; Hill et al. 2020). The first synthetic gene regulatory networks (synGRNs) that were constructed are the repressilator and the toggle switch in 2000 (Gardner et al. 2000; Elowitz & Leibler 2000). Since then, many diverse operational behaviors have been achieved with synGRNs, such as logic gates (Guet et al. 2002), counting events (Friedland et al. 2009), cellular memory (Ajo-Franklin et al. 2007) pattern formation (Schaerli et al. 2014; Barbier et al. 2020; Santos-Moreno et al. 2020), cell polarization (Chau et al. 2012) and light-sensing (Tabor et al. 2009) (reviewed in (Xie & Fussenegger 2019; Santos-Moreno & Schaerli 2018; Bashor & Collins 2018)). Synthetic RNs have been successfully implemented in prokaryotic and eukaryotic systems (Xie & Fussenegger 2019). In addition to the numerous potential applications of synGRNs in biotechnology, biomedicine and other fields (Xie & Fussenegger 2019; Ruder et al. 2011; Weber & Fussenegger 2011), synGRNs are also great model systems to study function and evolution of cellular regulation (Crocker & Ilsley 2017; Bashor & Collins 2018; Davies 2017; Santos-Moreno & Schaerli 2018). In the following, we will highlight several studies that explored design-principles and evolutionary dynamics with synGRNs.

2.1 Exploring network design space with synthetic regulatory networks

A certain biological function can be achieved with different networks varying in their topologies, i.e. in their type and number of regulatory nodes and interactions (Cotterell & Sharpe 2010). Explorations of distinct networks topologies can be highly insightful in regard to their functional properties and evolutionary potential. Chau *et al.* explored the design space of simple gene regulatory networks that achieve cell polarization (Chau et al. 2012). Spatial organization within cells through polarization is crucial for many cellular behaviors such as motility, asymmetric cell division and establishing polarity in epithelial cells and neurons (Raman et al. 2018). The team first computationally predicted all one- and two-node network topologies capable of cell polarization. All functional solutions contained one of three minimal motifs: positive feedback, mutual inhibition, or inhibition with positive feedback. Combinations of two or three minimal motifs increased the robustness of cell polarization, *i.e.* it was achieved over a larger range of parameters. To test their predictions experimentally, Chau *et al.* built an elegant synthetic system to study cell polarization in budding yeast using a toolkit of well-characterized biological parts including promoters, kinases, phosphatases and localization tags, most of which are not naturally found in budding yeast. Using this strategy, the authors were able to generate the three minimal motifs predicted to provide cell polarization and explored the parameter

range under which the networks are functional. In agreement with their theoretical predictions, the minimal motifs alone gave rise to cell polarization, although within a limited parameter range. However, when the minimal motifs were combined into more complex networks, robust cell polarization could be achieved over a wide parameter range. Thus, in this case, network robustness can be achieved by combining multiple minimal motifs, which might explain why combinations of multiple network motifs occur frequently in nature (Chau et al. 2012). In summary, this work not only demonstrated that it is possible to predict, design and synthetically build cell polarization networks, but also how multiple interactions contribute to network robustness and consequently evolutionary dynamics.

In a similar approach, Schaerli and colleagues explored the design space of three-node gene regulatory networks that translate a morphogen concentration gradient into a stripe-like gene expression pattern, *i.e.* a low-high-low gene expression along a morphogen gradient (Schaerli et al. 2014). The ability of RNs to convert a gradient input into spatial information is crucial during development, for example during axial patterning of the *Drosophila* embryo (Wolpert 1969). Similar to the study above, Schaerli *et al.* first explored the design space computationally. Interestingly, the identified networks can be divided into four groups and the simplest network of each group corresponds to one of the four types of incoherent feed-forward loops (Mangan & Alon 2003). Each of the four groups employs a distinct dynamical mechanism (spatiotemporal course of gene expression) to form a stripe (Cotterell & Sharpe 2010). The four simple networks were constructed by combining well-characterized regulatory components with a fluorescent reporter and expressed in *E. coli*. Indeed, a stripe pattern was experimentally achieved with all four network topologies. Based on the experimental and model results, they also designed and built a two-node stripe-forming network, representing the archetype of the four minimal three-node network topologies. Conclusively, this study experimentally demonstrated that stripe formation can be achieved with various network topologies and dynamical mechanisms, some of which have not yet been discovered in nature.

2.2 Exploring evolutionary dynamics with synthetic regulatory networks

In a follow-up study, Schaerli *et al.* used two of the stripe-forming synthetic networks (the incoherent feed-forward loops type 2 (I2) and 3 (I3)) to investigate whether and how the underlying dynamical regulatory mechanism of a network biases and affects its evolutionary potential (**Figure 2a)** (Schaerli et al. 2018). To this end, they introduced random mutations into each network and used a combination of experimental measurements, DNA sequencing and mathematical modeling to understand how mutations affect their ability to evolve novel phenotypes, i.e. phenotypes that are different from a stripe (**Figure 2a**). Remarkably, each network could only access a limited set of novel phenotypes and the accessible phenotypes differed for each network. The study provides thus the first empirical evidence that the underlying regulatory mechanisms of a RN can

cause constrained variation, as was previously proposed (Jiménez et al. 2015). Consequently, RNs with the same phenotype, but different underlying topology or regulatory mechanism, may not be equally evolvable and may constrain an organism's ability to evolve innovative and/or adaptive properties.

Using a similar approach, the Guet lab aimed at understanding how mutations in multiple components of a regulatory system interact and potentially yield new expression phenotypes (Lagator et al. 2017). The group built a simple synthetic regulatory system in *E. coli* consisting of three interacting molecular components: two *trans*-regulatory elements, namely a repressor (cI from lambda phage) and an endogenous RNA polymerase, and a *cis*-regulatory element consisting of the promoter and the overlapping repressor binding site (**Figure 2b**). The three components regulate the expression of a fluorescent protein that can be quantitatively measured. To understand how mutations in two components alter the regulatory behavior and function of the system, they introduced mutations in the *cis*-regulatory element and in the repressor. They analyzed the effect of the mutations independently and in combination. Surprisingly, when introducing mutations in both components the regulatory system produced gene expression phenotypes that were not observed when mutating only one component. The authors attribute these emerging phenotypes to epistatic interactions between the transcription factor and its DNA binding site. Epistasis means that the combined functional effect of two or more mutations differs from the expected value based on the individual effects (Lehner 2011; Nghe et al. 2020). In this case, epistasis increased phenotypic variation that selection can act on and thus might facilitate subsequent adaptive evolution. This seems to be true also for other transcriptional regulatory systems, such as for the above-described stripe-forming three-gene synGRN of Schaerli *et al.* (Schaerli et al. 2018). Here, mutations in multiple network nodes gave rise to a wider spectrum of phenotypes compared to mutations in only one of three network nodes.

Phenotypic heterogeneity, e.g. due to stochasticity in gene expression, can be a beneficial property for microorganisms in a changing or fluctuating environment by providing some individuals within a population with a survival advantage (reviewed in (Ackermann 2015; Payne & Wagner 2019). How does phenotypic heterogeneity influence evolutionary dynamics? To address this question experimentally the Pál lab used an elegant combination of synthetic biology and experimental evolution by designing and implementing two synGRNs that control the expression of an antifungal resistance gene in *S. cerevisiae* with different degrees of gene expression heterogeneity (Bódi et al. 2017). They found that synGRNs with higher heterogeneity not only provided a higher initial resistance to the antifungal drug, but also allowed the yeast cells to evolve a higher resistance after several rounds of evolution under gradually increasing concentrations of the antifungal drug. Also, beneficial mutations in the synGRN with high heterogeneity were contingent on this high gene expression stochasticity, meaning that their adaptive effects were substantially reduced in a background with low gene expression stochasticity.

Thus, gene expression stochasticity can influence evolutionary trajectories by widening the spectrum of available beneficial mutations during evolutionary adaptation. Remarkably, in the synGRN with initial low stochasticity elevated phenotypic heterogeneity evolved as a direct response to the antifungal stress. However, at the same time, the benefit of high phenotypic heterogeneity trades-off with reduced fitness in the drug-free medium. Thus, gene expression stochasticity might be an evolvable trait that is selected for in fluctuating and changing environments (Sánchez-Romero & Casadesús 2013; Arnoldini et al. 2014; Holland et al. 2014; Salathé et al. 2009; Acar et al. 2008; Kuwahara & Soyer 2012; Sato et al. 2003).
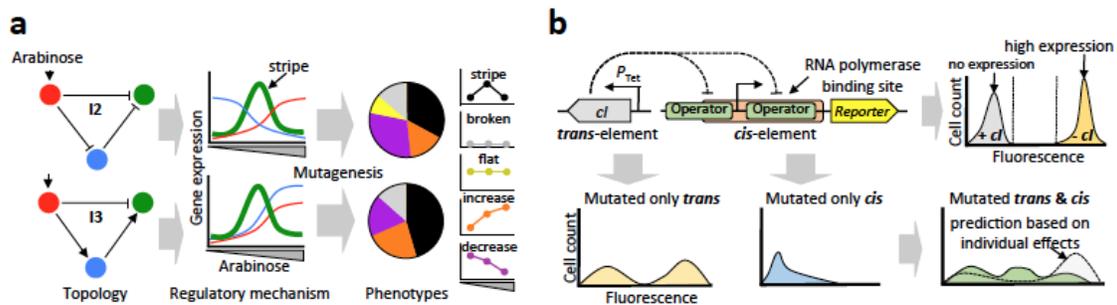

Figure 2. Using synthetic circuits to understand evolutionary constraints and epistasis in gene regulatory networks. (**a**) Investigating evolutionary biases of two synthetic networks (incoherent feed-forward loops type 2 (I2, top) and 3 (I3, bottom)) that achieve stripe formation with distinct topologies and regulatory mechanisms (Schaerli et al. 2018). The networks were implemented in *E. coli* cells with a "morphogen" (arabinose) input receiver gene (red), an intermediate loop gene (blue) and an output gene (green) with GFP expression as readout. Arrows represent activation and bars indicate repression. Mutations were introduced into the regulatory regions of each node of both networks and resulting phenotypic changes were characterized at three arabinose concentrations (low, medium and high). The two circuits produce a different spectrum of novel gene expression phenotypes, *e.g.* only I2 achieves a flat phenotype. (**b**) A synthetic transcriptional regulatory system to study how mutations in single and multiple components affect gene expression phenotypes (Lagator et al. 2017). In this system, an inducible *trans*-element (lambda repressor protein *cI*) represses expression of a fluorescent reporter gene by binding to the operator sites in a *cis*-element that overlaps with the promoter region (RNA polymerase binding site). Presence of the *trans*-element (+cI) results in a low fluorescence phenotype, whereas its absence (-cI) results in a high fluorescence phenotype. Introducing mutations only in *trans* yields a bimodal fluorescence distribution, whereas mutations only in *cis* yield low and intermediate fluorescence phenotypes. Combined mutagenesis of *trans* and *cis* sequences produces more intermediate fluorescence phenotypes than expected from each separately.

3. Rewired regulatory networks

Understanding why natural RNs evolved a particular topology and which components and connections of a network are either necessary or dispensable for functionality is a difficult question to address. Classical genetic approaches usually delete individual network components or interactions to decipher their functional role within the network. However, alternative topologies with new and rewired interactions are difficult to study in this way. Instead, with a synthetic biology approach network components and interactions cannot only be deleted but also added, rewired and fine-tuned (Mukherji & van Oudenaarden 2009; Bashor et al. 2010). Additionally, a networks' output, which might be difficult to observe and quantify, can be linked to an additional measurable output such as expression of a fluorescent protein.

An interesting question is why a particular topology was favored during evolution over alternative topologies. A study by the Elowitz lab investigated the network topology of the Hedgehog signaling pathway by reconstituting a developmental morphogen gradient *in vitro*, with a tunable synthetic rewiring of regulatory interactions and a fluorescent readout in a cell-culture model, combined with *in silico* modeling (**Figure 3 a**) (Li et al. 2018). The Hedgehog pathway is crucial in establishing positional information for proper patterning during embryonic development and is composed of a double-negative regulatory logic and an additional negative feedback (Briscoe & Thérond 2013). The combination of a semi-synthetic system and *in silico* modeling revealed that the natural negative feedback architecture shows the most robust behavior in length scale and amplitude of the hedgehog signaling gradient compared to alternative architectures. In addition, it reaches steady state more rapidly and over a wider range of signaling molecule concentration than alternative topologies (**Figure 3b**). However, it remains an open question of whether the rapid response and robustness of the system to changes in the rate of morphogen production has been directly selected for during evolution or rather resulted as a by-product of other evolutionary forces (Kaneko 2007; Ciliberti et al. 2007).
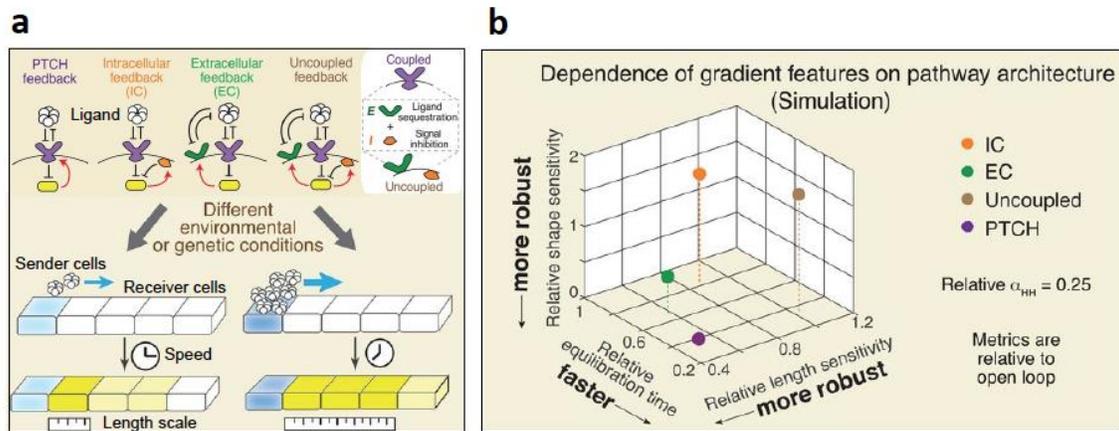
Figure 3. Rewiring of cellular circuits to understand design principles. (**a**) Hedgehog pathway topologies studied by Li *et al.* The natural bifunctional PTCH receptor provides intracellular (IC) and extracellular (EC) feedback simultaneously. The IC feedback inhibits the signaling pathway only intracellular, whereas the EC feedback does so only extracellular by sequestering the ligand. The PTCH feedback was uncoupled through separating the functional parts of the PTCH receptor. The cell culture system with sender and receiver cells allowed quantitative analysis of the spatiotemporal patterning dynamics resulting from the different circuit wirings. (**b**) Summary of robustness and speed properties of different feedback architectures for Hedgehog (HH) signaling gradient formation. Values are based on *in silico* simulations and relative to that of the open loop model at relative HH (relative ligand production rate) = 0.25, which has a value of 1 in each dimension. Reproduced with permission (Li et al. 2018).

Isalan and colleagues went beyond individual networks and instead rewired large parts of the *E. coli* genome (Isalan et al. 2008). The authors rewired a set of transcription and σ-factor genes with different unrelated regulatory regions, thus creating almost 600 reconnected networks on top of an otherwise unchanged *E. coli* genome. Surprisingly, only a few of the synthetic rewirings did considerably affect growth, although a follow-up study showed that the perturbations of gene expressions span up to four orders of magnitude and changed up to ∼ 70% of the transcriptome (Baumstark et al. 2015). In fact, some of the rewirings were actually beneficial under stressful conditions such as heat-shock and prolonged stationary phase. Thus, substantial rewiring of transcriptional networks, at least in *E. coli*, is tolerable to some extent and may even be advantageous under stressful environmental conditions.

4. Synthetic genomics
The ability to large-scale recoding and the emergence of computationally (re-)designed synthetic genes and genomes opened many possibilities for applied and fundamental research. Synthetic genomics advanced drastically since the first synthetic gene was

synthesized in 1970 (Agarwal et al. 1970). In particular the cost, quality and speed of *de novo* DNA synthesis improved significantly, as well as assembly technologies that allow the *de novo* synthesis of whole chromosomes and genomes (H. Wang et al. 2016; Haimovich et al. 2015). Here, we will focus on the evolutionary perspective of recoded and designed synthetic genomes (see also (Pál et al. 2014)) and refer to other excellent reviews that cover technological aspects and potential applications of this exciting topic (Mukai et al. 2017; Chari & Church 2017; Haimovich et al. 2015; W. Zhang et al. 2020).

4.1 Extending the alphabet of life

In the early 2000s, researchers first demonstrated that the genetic code can be engineered and expanded to incorporate non-natural amino acids (nnAA) with distinct chemical and structural properties (Chin et al. 2003; L. Wang et al. 2001) (reviewed in (C. C. Liu & Schultz 2010; Chin 2014; Santos-Moreno & Schaerli 2020a)). To encode a nnAA within a gene or genome, an unassigned free codon (usually the rare TAG amber stop codon) is reassigned to encode a nnAA (Chin 2017). Second, incorporation of a nnAA into proteins requires an orthogonal tRNA-synthetase pair that specifically recognizes the assigned codon and the nnAA, but not any other codon or AA (Chin 2017). Protein engineers have used nnAAs to generate proteins and enzymes with new biophysical and chemical properties that would not be possible with standard AAs (W. H. Zhang et al. 2013; C. C. Liu & Schultz 2010). Now, with the development of genome-wide editing tools researchers have recoded whole genomes in order to allow incorporation of nnAAs into the proteome (Arranz-Gibert et al. 2018; Fredens et al. 2019). The Isaacs and Church labs engineered an *E. coli* strain (named C321.ΔA) that has all 321 TAG stop codons replaced with the synonymous TAA stop codon (Lajoie et al. 2013). The complete codon substitution allowed the deletion of the release factor 1 (RF1, which terminates translation at UAG and UAA). Therefore, UAG codons are unassigned in C321.ΔA while RF2 still terminates translation at UGA and UAA. The evolutionary consequence of the complete codon substitution and RF1 deletion is that horizontally acquired DNA containing TAG codons cannot correctly be translated, meaning that the organism is genetically isolated (Lajoie et al. 2013). On the one hand, this provides resistance to phages that contain TAG codons but it also hinders the acquisition of potentially beneficial DNA, such as plasmids containing antibiotic resistance genes. Jing Ma and Isaacs performed an interesting evolution experiment with phage populations infecting a mixture of standard and recoded *E. coli* strains at various ratios (N. J. Ma & Isaacs 2016). The study showed that phages adapted towards C321.ΔA by reducing their TAG codon usage. This experiment provides compelling evidence that phages and viruses adapt rapidly their genetic code to achieve compatibility with their host.

Another intriguing question is how an expanded genetic code beyond the generic 20 amino acids alters the evolution of an organism. Does the availability of new chemistry promote novel opportunities for beneficial mutations? A study by the Barrick lab addressed

this question experimentally by evolving T7 bacteriophages with a host *E. coli* in which the amber stop codon (TAG) was reassigned to incorporate 3-iodotyrosine (IdioY) as a 21st amino acid (**Figure 4a**) (Hammerling et al. 2014). The T7 phage populations improved their lysis times (fitness proxy) during several rounds of laboratory evolution and adapted specifically to the IdioY incorporating *E. coli* host. At the end of the experiment several improved phage mutants also incorporated IdioY into their genome and their improved fitness was dependent on the unnatural amino acid. For example, a Tyr39-to-IodoY mutation in the T7 type II holin gene was more beneficial than having a Tyr or Trp at this position. Although other mutants in this study evolved to have the same fitness using only the generic 20 amino acids, the readiness to which the phage population adapted to an alternative genetic code and incorporated a non-natural amino acid in its proteins is quite remarkable. Since incorporation of nnAA can generate proteins with novel functionalities and properties (see reviews (Chin 2014) and(C. C. Liu & Schultz 2010)), an expanded genetic code could potentially facilitate evolution towards higher fitness beyond what is feasible with the universal 20 amino acids.

Recently, systems have been developed that allow the incorporation of more than one nnAA at the same time in one organism. This is achieved by freeing up and reassigning existing triplet codons (Fredens et al. 2019) or by using quadruplet codons (K. Wang et al. 2012), both of which require modified and orthogonal translation systems in the host cell. For example, the Chin lab recoded the entire *E. coli* genome to use only 61 (instead of 64) codons, which will eventually allow the reassignment of three codons to distinct nnAAs (Fredens et al. 2019). To encode even more nnAAs simultaneously, orthogonal quadruplet-decoding ribosome and tRNA-synthase have been evolved in the laboratory, which theoretically allows the incorporation of more than 200 nnAAs combinations in recombinant proteins or recoded organisms (K. Wang et al. 2012). However, the approach is challenging and so far only a few studies were able to encode nnAAs with quadruplet codons (Neumann et al. 2010; Niu et al. 2013).

Not only have nnAAs been incorporated into proteins, but also non-natural nucleotides into the DNA and RNA of living organisms (Pinheiro & Holliger 2012; Y. Zhang & Romesberg 2018). In 2014 the Romesberg lab reported the first successful incorporation of an unnatural base pair (UBP) dNaM-dTPT3 into a *E. coli* plasmid (Malyshev et al. 2014). The UBP forms through hydrophobic interactions instead of hydrogen bonding present in the natural base pairs. However, cells with an expanded genetic code grew poorly and very easily lost the UBP. The stable maintenance of the six-letter/three-base-pair code required tuning of the nucleotide uptake system and the implementation of an elegant Cas9 based control system that eliminates DNA that had lost the UBP (Y. Zhang, Lamb, et al. 2017). Subsequently, the team extended their system to code for nnAAs with the UBP (Y. Zhang, Ptacin, et al. 2017). This required transcription by T7 RNA polymerase and translation involving a tRNA containing the unnatural anticodon. As a proof-of-principle, the Romesberg lab successfully incorporated a nnAA

into a fluorescent protein using two codons that contain the UBP (Y. Zhang, Ptacin, et al. 2017). In a recent follow up study, the Romesberg lab systematically generated and studied the functionality of unnatural codons using the dNaM-dTPT3 UBP (Fischer et al. 2020). Out of the 152 theoretical possible codons they identified nine functional unnatural codons that are stably integrated in DNA, transcribed into mRNA and tRNA, and finally can be decoded into nnAAs in proteins. Out of the nine, three unnatural codons also function orthogonal and can be used for simultaneously incorporating of multiple nnAAs into a single protein, thus allowing the decoding of 67 codons in a living semi-synthetic organism (Fischer et al. 2020).

Together, this body of work demonstrates that the central dogma of life is not limited to four DNA bases and 20 amino acids. This in turn may allow us to design and construct new synthetic life forms that are different from natural ones at the molecular level. The work on synthetic genetic polymers (XNA) with different backbone chemistries than DNA and RNA also brings us closer to this goal (Pinheiro & Holliger 2012; Pinheiro et al. 2012; Anosova et al. 2016; Hoshika et al. 2019). Not only can such synthetic life forms be exploited for applications such as biocontainment, therapeutics and novel chemistry (Sun et al. 2014), but also to better understand the evolutionary constraints and benefits of the natural genetic code and the canonical nucleic and amino acids (Bacher et al. 2004; Koonin & Novozhilov 2017).

4.2 Synthetic karyotyping

The number of chromosomes varies widely in eukaryotic species. What happens if the number of chromosomes changes? Comparative studies between related species with different chromosome numbers are difficult to interpret due to the simultaneous presence of other changes in the genome, such as sequence divergence and genomic rearrangements. This makes it difficult to assign a given phenotypic feature to the difference in chromosome number. Synthetic biology allows us to change only the chromosome number, while maintaining the genetic content. Two studies successfully fused the 16 chromosomes of the budding yeast in successive rounds of chromosome fusion down to two chromosomes (n=2) (J. Luo et al. 2018) or a single giant chromosome (n=1) (Shao et al. 2019) (**Figure 4b**). Both groups used the CRISPR–Cas9 technology to remove the telomers and centromeres of the chromosomes and took advantage of the endogenous DNA repair machinery for chromosome fusion. Surprisingly, chromosome fusion had little impact on cell fitness, with only small fitness defects becoming apparent at n=2 and n=1. Also, the mating efficiency of strains with the same number of chromosomes was unaffected, e.g. n=2 x n=2 were not different in sporulation efficiency compared to n=16 x n=16. However, the larger the difference in chromosome number between the two mating partners, the fewer viable spores resulted from crossing. For example, n=16 could only produce viable spores with strains having at least 10 chromosomes, but not less. Therefore, despite having identical sequences, a reduction to eight chromosomes is enough to isolate strains

reproductively. This is interesting because the phenomenon of "reproductive isolation" is usually associated with sequence divergence rather than with the number of chromosomes (Greig 2009). One avenue for future work with these strains might be to perform laboratory evolution experiments in order to investigate how they adapt to a reduced number of chromosomes (Liti 2018).

4.3 Synthetic viruses

Viruses and phages are one of the simplest biological systems in which genotype and phenotype are directly linked through encapsulation. However, they require a host to proliferate. Given the small genomes of viruses, it is not surprising that the first genomes to be chemically synthesized were that of viruses, namely that of the poliovirus in 2002 (Cello et al. 2002) and of the bacteriophage φX174 in 2003 (H. O. Smith et al. 2003). Building a synthetic virus only requires chemically synthesizing its genome and injecting it into the right host cells, which will then produce viral particles that are infectious. The synthesis of viral genomes has allowed reconstructing and characterizing viruses from past and current pandemics, such as the Spanish Flu virus from 1918 (Tumpey et al. 2005) and SARS-CoV-2 (Thao et al. 2020), and studying the effect of genome modifications on pathology as well as to test potential vaccine candidates (Wimmer & Paul 2011).

The emergence of protein cages capable of encapsulating its DNA or RNA genome was probably the critical step in the evolution of viruses. Two studies describe the design, construction and evolution of viral-like capsids from non-viral proteins that encapsulate their own genetic information (Butterfield et al. 2017; Terasaka et al. 2018) (**Figure 4c**). In both cases, the genome packaging and protection properties of the starting capsids were improved by carrying out several rounds of *in vitro* directed evolution. This approach quickly yielded mutants that could compete with recombinant virus vectors and thus established simple evolutionary pathways by which virus-like genome packaging can emerge. In future, such synthetic capsids might be endowed with further properties such as cell recognition and infection, unloading and even self-replication (Lemire et al. 2018). Moreover, synthetic viruses and phages offer interesting alternatives to natural viruses and phages as vectors in drug delivery in therapeutics and vaccines or as platforms for phage display (Citorik et al. 2014).
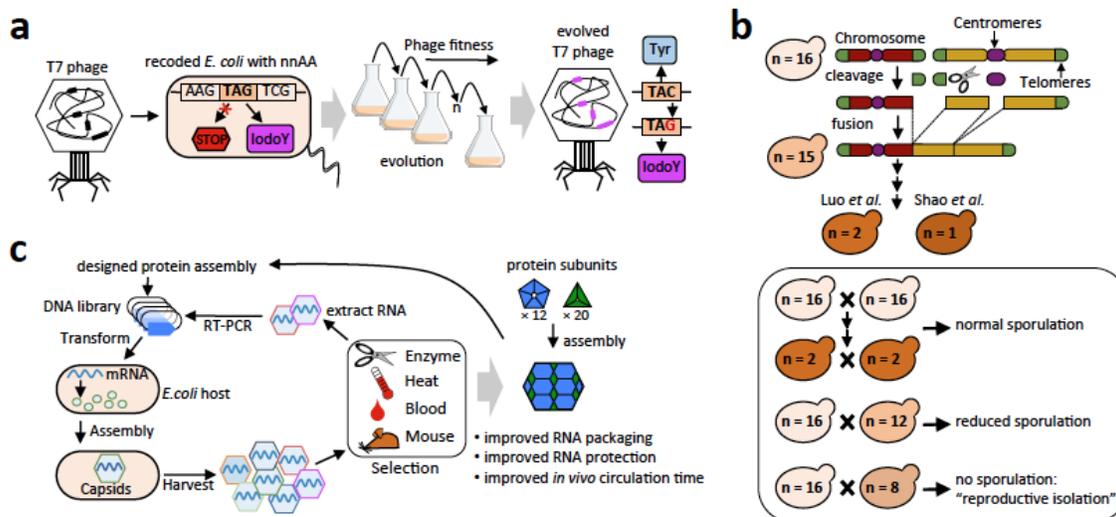
Figure 4. Evolutionary insights with modified, recoded and designed synthetic genomes. (**a**) Adaptive evolution experiment of T7 bacteriophages infecting a recoded *E. coli* strain with an expanded genetic code that incorporates 3-idiotyrosine (IodoY) at the amber stop codon TAG (Hammerling et al. 2014). Phages adapt by incorporating IdioY in their proteome, which provides higher fitness than other amino acids at the same positions. (**b**) Two independent studies explore the consequences of step-wise fusing the 16 chromosomes (n=16) in *S. cerevisiae* down to n=2 (J. Luo et al. 2018) and n=1 (Shao et al. 2019). (**Box**) Mating strains with equal numbers of chromosomes produces viable spores. Increasing the difference in chromosome numbers between mating strains results in reduced spore viable and eventually in reproductive isolation. (**c**) Evolution of a computationally designed protein capsid that encapsulates its own RNA (Butterfield et al. 2017; Terasaka et al. 2018). A DNA mutant library is transformed into *E. coli* hosts that produce the encoding capsids, which are subsequently harvested and selected for improved properties, such as RNase and heat stability as well as blood and mouse circulation times (small box). Isolated RNA is reverse-transcribed to DNA with RT-PCR (reverse-transcription polymerase chain reaction) to start a new round of evolution. Several rounds of evolution resulted in improved RNA packaging and stability within the capsid.

4.4 Designing a synthetic minimal genome

The minimum number of genes to sustain life is a fundamental question in biology. Researchers have approached this question theoretically by hypothesizing that a common set of genes shared between species with small genomes might be a good approximation of how many genes are essential for life (Mushegian & Koonin 1996). Complementary, scientist performed transposon mutagenesis on the *Mycoplasma genitalium* genome, which has the smallest genome of any organism that can be grown in pure culture, to identify essential genes for bacterial growth under laboratory conditions (Glass et al. 2006). Synthetic biology now offers tools to address this question by designing and synthesizing a minimal synthetic genome. In 2010, the team of Craig Venter generated the first bacterial cell controlled by a chemically synthesized *M. mycoides* genome (1077 kb), named JCVI-syn1.0 (Gibson et al. 2010). Although the JCVI-syn1.0 genome is nearly identical to the

natural *M. mycoides* genome, it was an important technical milestone towards bottom-up genome designs. In 2016, the Venter team released JCVI-syn3.0, a reduced version of the JCVI-syn1.0 synthetic genome: JCVI-syn3.0 contains only 473 genes encoded on only 531 kb, which makes it the smallest genome of any autonomously replicating cell known so far. In addition to essential genes, it also contains quasi-essential genes that are required for robust growth with a doubling time of around 180 min (Hutchison et al. 2016). Thus, JCVI-syn3.0 is an approximation of a minimal cellular genome. Intriguingly, even in this simple organism the cellular function of 149 genes ($\approx$31%) is still unknown. Thus, much has still to be learnt about what functions are required for life. Once the functions of all these genes are known, the genome size may potentially even be further reduced. In addition, knowing all the biochemical, structural and cellular functions essential for supporting life, one can start to design a new organisms from these basic principles and learn much about the origin and evolution of life (Göpfrich et al. 2018; Forster & Church 2006).

4.5 Synthetic self-replicating systems in cell like compartments

Studying the factors that facilitated the emergence of life from chemical molecules is an exciting, but obviously also a challenging task. In the last decade, several groups have built simple self-replicating systems based on RNA molecules, which are hypothesized to have stored genetic information and at the same time catalyzed chemical reactions in primitive cells (Joyce & Szostak 2018). A particular interesting question is how compartmentalization facilitated the early evolution of self-replicating molecules such as RNA and dealt with the emergence of parasitic mutants, which are replicated but themselves are not replicating and thus would cause eventual collapse of the systems (Ichihashi et al. 2013). Matsumara *et al.* (Matsumura et al. 2016) evolved self-replicating RNA molecules in a scenario of repeated mixing and compartmentalization in non-biological material, using a droplet-based microfluidic system, which provided protection from emerging parasitic mutants. The study supports the hypothesis that transient compartmentalization, e.g. in aerosols, microcompartments in hydrothermal vents or on mineral surfaces, has facilitated evolution before the first cell-like structures emerged (Ichihashi et al. 2013). In another study Mizuuchi & Ichihashi extended a synthetic RNA self-replicating system to depend on cooperation, which is a necessary requirement for the evolution of higher complexity (Mizuuchi & Ichihashi 2018). Emerging parasitic RNA molecules plagued self-replication and cooperation, but compartmentalization protected the system and supported cooperation. Interestingly, evolutionary adaption through mutations towards higher replication efficiency was characterized by improved 'selfish' replication and at the same time coevolution of the cooperators.

5. Outlook

Since its foundation in 2000, synthetic biology has undergone dramatic growth to a vibrant research discipline that is poised to provide fundamental insights into biological questions as well as to revolutionize many aspects of our lives, for example by producing smart materials, sustainable biofuels and personalized therapeutics (Cameron et al. 2014; Purnick & Weiss 2009; Tang et al. 2020). This progress was enabled by improved and novel technologies, including better computational models, cheaper DNA sequencing, improved DNA synthesis, high-precision DNA editing tools such as CRISPR, novel tools for gene expression control, microfluidic devices, as well as high-throughput assembly and screening methods (Gach et al. 2017; H. Wang et al. 2016; Cobb et al. 2013; Santos-Moreno & Schaerli 2020b). As reviewed here, synthetic biological systems already started to improve our understanding of fundamental evolutionary concepts. We predict that this process will continue, and we highlight here some research avenues where we expect interesting results in near future.

The improvement in quality and speed of *de novo* DNA synthesis and assembly technologies and the accompanying reduced costs allowed the *de novo* synthesis of whole chromosomes and genomes, as discussed above for bacterial genomes. The Synthetic Yeast 2.0 consortium (www.syntheticyeast.org) is on its way to build the first synthetic eukaryotic genome, a synthetic version of *S. cerevisiae* genome, called Sc2.0. So far, the synthesis of six (out of 16) synthetic Sc2.0 chromosomes has been published and we expect soon the publication of the whole genome (Richardson et al. 2017; Kannan & Gibson 2017). The next version of a synthetic yeast genome (Sc3.0) is also already in planning (Dai et al. 2020) with the aim for further compacting the synthetic chromosomes (Z. Luo et al. 2020). For now, the Sc2.0 features the following changes compared to the natural yeast genome: all TAG stop codons are changed to TAA, loxP sites are introduced after nonessential genes to allow increased evolutionary diversification using SCRaMbLE (synthetic chromosome rearrangement and modification by loxP-mediated evolution), and enhanced genome stability is achieved through removal of repeat elements, introns and relocation of all transfer RNAs to a new chromosome (Richardson et al. 2017). These design features not only increase its potential for biotechnology applications, but will also allow us to address fundamental questions. For example, the SCRaMbLE technology, with its different versions, such as L-SCRaMbLE (Hochrein et al. 2018) (light-inducible and completely reversible), *in vitro* DNA SCRaMbLE (Wu et al. 2018), SCRaMbLE-in (W. Liu et al. 2018) (combination of *in vitro* and *in vivo* recombination) and ReSCuES (Z. Luo et al. 2018) (reporter of SCRaMbLEd cells using efficient selection), allows large-scale genome reshuffling to expand the evolutionary potential of budding yeast (Blount et al. 2018; L. Ma et al. 2019; Wightman et al. 2020). Evolution experiments with SCRaMbLEd synthetic genomes might provide crucial insights into speciation, minimal genome requirements and genome evolution (Szymanski & Calvert 2018). The Genome Project-write (http://engineeringbiologycenter.org), aiming to synthesize the human and other

genomes, is also poised to significantly advance genome-engineering technologies and provide new exciting platforms to study evolutionary questions (Boeke et al. 2016).

Another area that promises to yield fundamental insights into evolutionary systems biology questions is the building of a synthetic cell. Several interdisciplinary consortia, such as BaSyC (Building a Synthetic Cell, www.basyc.nl), MaxSynBio (www.maxsynbio.mpg.de) and Synthetic Cell Initiative (www.syntheticcell.eu) started to work on the ambitious goal to create a completely synthetic cell-like system that has characteristics of life, such as reproduction, metabolism, growth, compartmentalization, homeostasis, heredity, adaptation and communication. Although we are still far away from a synthetic cell that is truly "alive", simple cell-like systems that exhibit some of these characteristics have already been built (Vogele et al. 2018; van Nies et al. 2018; Lavickova et al. 2020), also reviewed in (Buddingh & van Hest 2017).

So far, we discussed how evolutionary systems biology can benefit from synthetic biology. However, the combination of the two disciplines is not a one-way road. Despite impressive progress, building synthetic systems rationally is often still not straightforward. Therefore, synthetic biologists increasingly choose selecting or screening a library of different variants to obtain systems that function as desired (de Lorenzo 2018; Szymanski & Calvert 2018). By adopting the powerful method of directed evolution, which originates from protein engineering (Packer & D. R. Liu 2015), functional synthetic systems can be obtained through screening or selecting from randomized or combinatorial libraries (Schaerli & Isalan 2013; Duarte et al. 2017; Cobb et al. 2013). In fact, initial designs of computationally designed proteins (Giger et al. 2013; Blomberg et al. 2013), synthetic circuits (Yokobayashi et al. 2002), metabolic pathways (Bachmann 2016), synthetic genomes (Wannier et al. 2018) and virus-like nucleocapsids (Butterfield et al. 2017; Terasaka et al. 2018) have been optimized with directed evolution. In many of these cases, rational design and modeling alone could not have identified the necessary modifications to optimize the systems. Indeed, various techniques based on evolutionary principles have been developed such as MAGE (multiplex automated genome engineering) (H. H. Wang et al. 2009), SEER (serial enrichment for efficient recombineering) (Wannier 2020) DIvERGE (directed evolution with random genomic mutations) (Nyerges et al. 2018), CAGE (conjugative assembly genome engineering) (Isaacs et al. 2011), PACE (phage-assisted continuous evolution) (Esvelt et al. 2011) and eVOLVER (automated high-throughput growth experiments) (Wong et al. 2018) and help to achieve ambitious goals in synthetic biology.

In conclusion, the combination of synthetic and evolutionary systems biology is proving to be a successful partnership. As showcased in this book chapter, the application of synthetic biology to address evolutionary questions has already produced promising results. In the future, the combination of synthetic systems and evolutionary experiments promises to deliver further exciting fundamental insights into the principles of molecular evolution.

References

Acar, M., Mettetal, J.T. & van Oudenaarden, A., 2008. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics*, 40(4), pp.471–475.

Ackermann, M., 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology*, 13(8), pp.497–508.

Agarwal, K.L. et al., 1970. Total synthesis of the gene for an alanine transfer ribonucleic acid from yeast. *Nature*, 227(5253), pp.27–34.

Ajo-Franklin, C.M. et al., 2007. Rational design of memory in eukaryotic cells. *Genes & Development*, 21(18), pp.2271–2276.

Anosova, I. et al., 2016. The structural diversity of artificial genetic polymers. *Nucleic Acids Research*, 44(3), pp.1007–1021.

Arnold, F.H., 2010. How Proteins Adapt: Lessons from Directed Evolution. *Cold Spring Harbor Symposia on Quantitative Biology*, 74(0), pp.41–46.

Arnoldini, M. et al., 2014. Bistable expression of virulence genes in salmonella leads to the formation of an antibiotic-tolerant subpopulation. *PLoS Biology*, 12(8), p.e1001928.

Arranz-Gibert, P., Vanderschuren, K. & Isaacs, F.J., 2018. Next-generation genetic code expansion. *Current Opinion in Chemical Biology*, 46, pp.203–211.

Bacher, J.M. et al., 2004. Evolving new genetic codes. *Trends in ecology & evolution*, 19(2), pp.69–75.

Bachmann, B.O., 2016. Applied evolutionary theories for engineering of secondary metabolic pathways. *Current Opinion in Chemical Biology*, 35, pp.133–141.

Barbier, I., Perez-Carrasco, R. & Schaerli, Y., 2020. Controlling spatiotemporal pattern formation in a concentration gradient with a synthetic toggle switch. *Molecular Systems Biology*, 16(6), p.611.

Bashor, C.J. & Collins, J.J., 2018. Understanding Biological Regulation Through Synthetic Biology. *Annual Review of Biophysics*, 47, pp.399–423.

Bashor, C.J. et al., 2010. Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. *Annual Review of Biophysics*, 39, pp.515–537.

Baumstark, R. et al., 2015. The propagation of perturbations in rewired bacterial gene networks. *Nature Communications*, 6, p.10105.

Becskei, A. & Serrano, L., 2000. Engineering stability in gene networks by autoregulation. *Nature*, 405, pp.590–593.

Blain, J.C. & Szostak, J.W., 2014. Progress toward synthetic cells. *Annual Review of Biochemistry*, 83, pp.615–640.

Blomberg, R. et al., 2013. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature*, 503(7476), pp.418–421.

Blount, B.A. et al., 2018. Rapid host strain improvement by in vivo rearrangement of a synthetic yeast chromosome. *Nature Communications*, 9(1), p.1932.

Boeke, J.D. et al., 2016. The Genome Project-Write. *Science*, 353(6295), pp.126–127.

Bódi, Z. et al., 2017. Phenotypic heterogeneity promotes adaptive evolution. *PLoS Biology*, 15(5), p.e2000644.

Briscoe, J. & Thérond, P.P., 2013. The mechanisms of Hedgehog signalling and its roles in development and disease. *Nature Reviews Molecular Cell Biology*, 14(7), pp.416–429.

Buddingh, B.C. & van Hest, J.C.M., 2017. Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Accounts of chemical research*, 50(4), pp.769–777.

Butterfield, G.L. et al., 2017. Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature*, 552(7685), pp.415–420.

Cameron, D.E., Bashor, C.J. & Collins, J.J., 2014. A brief history of synthetic biology. *Nature Reviews Microbiology*, 12(5), pp.381–390.

Cello, J., Paul, A.V. & Wimmer, E., 2002. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science*, 297(5583), pp.1016–1018.

Chari, R. & Church, G.M., 2017. Beyond editing to writing large genomes. *Nature Reviews Genetics*, 18(12), pp.749–760.

Chau, A.H. et al., 2012. Designing synthetic regulatory networks capable of self-organizing cell polarization. *Cell*, 151(2), pp.320–332.

Chin, J.W., 2014. Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annual Review of Biochemistry*, 83(1), pp.379–408.

Chin, J.W., 2017. Expanding and reprogramming the genetic code. *Nature*, 550(7674), pp.53–60.

Chin, J.W. et al., 2003. An expanded eukaryotic genetic code. *Science*, 301(5635), pp.964–967.

Ciliberti, S., Martin, O.C. & Wagner, A., 2007. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Computational Biology*, 3(2), p.e15.

Citorik, R.J., Mimee, M. & Lu, T.K., 2014. Bacteriophage-based synthetic biology for the study of infectious diseases. *Current Opinion in Microbiology*, 19, pp.59–69.

Cobb, R.E., Si, T. & Zhao, H., 2012. Directed evolution: an evolving and enabling synthetic biology tool. *Current Opinion in Chemical Biology*, 16(3-4), pp.285–291.

Cobb, R.E., Sun, N. & Zhao, H., 2013. Directed evolution as a powerful synthetic biology tool. *Methods*, 60(1), pp.81–90.

Cotterell, J. & Sharpe, J., 2010. An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular Systems Biology*, 6, pp.1–14.

Crocker, J. & Ilsley, G.R., 2017. Using synthetic biology to study gene regulatory evolution. *Current opinion in genetics & development*, 47, pp.91–101.

Dai, J. et. al., 2020. Sc3.0: revamping and minimizing the yeast genome. *Genome Biology,* 21(1), pp.205.

Davidson, E.A., Windram, O.P.F. & Bayer, T.S., 2012. Building synthetic systems to learn nature's design principles. *Advances in experimental medicine and biology*, 751, pp.411–429.

Davies, J., 2017. Using synthetic biology to explore principles of development.

Development, 144(7), pp.1146–1158.

de Lorenzo, V., 2018. Evolutionary tinkering vs. rational engineering in the times of synthetic biology. *Life sciences, society and policy*, 14(1), p.18.

Dean, A.M. & Thornton, J.W., 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nature Reviews Genetics*, 8(9), pp.675–688.

Duarte, J.M., Barbier, I. & Schaerli, Y., 2017. Bacterial Microcolonies in Gel Beads for High-Throughput Screening of Libraries in Synthetic Biology. *ACS Synthetic Biology*, 6(11), pp.1988–1995.

Elowitz, M.B. & Leibler, S., 2000. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767), pp.335–338.

Esvelt, K.M., Carlson, J.C. & Liu, D.R., 2011. A system for the continuous directed evolution of biomolecules. *Nature*, 472(7344), pp.499–503.

Fischer, E.C. et al., 2020. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nature Chemical Biology*, 16, pp. 570–576.

Forster, A.C. & Church, G.M., 2006. Towards synthesis of a minimal cell. *Molecular Systems Biology*, 2, p.45.

Fredens, J. et al., 2019. Total synthesis of Escherichia coli with a recoded genome. *Nature*, 569(7757), pp.514–518.

Friedland, A.E. et al., 2009. Synthetic gene networks that count. *Science*, 324(5931), pp.1199–1202.

Gach, P.C. et al., 2017. Droplet microfluidics for synthetic biology. *Lab on a chip*, 17(20), pp.3388–3400.

Gardner, T.S., Cantor, C.R. & Collins, J.J., 2000. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767), pp.339–342.

Gibson, D.G. et al., 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987), pp.52–56.

Giger, L. et al., 2013. Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nature Chemical Biology*, 9(8), pp.494–498.

Glass, J.I. et al., 2006. Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences*, 103(2), pp.425–430.

Göpfrich, K., Platzman, I. & Spatz, J.P., 2018. Mastering Complexity: Towards Bottom-up Construction of Multifunctional Eukaryotic Synthetic Cells. *Trends in Biotechnology*, 36(9), pp.938–951.

Greig, D., 2009. Reproductive isolation in Saccharomyces. *Heredity*, 102(1), pp.39–44.

Guet, C.C. et al., 2002. Combinatorial synthesis of genetic networks. *Science*, 296(5572), pp.1466–1470.

Haimovich, A.D., Muir, P. & Isaacs, F.J., 2015. Genomes by design. *Nature Reviews Genetics*, 16(9), pp.501–516.

Hammerling, M.J. et al., 2014. Bacteriophages use an expanded genetic code on evolutionary paths to higher fitness. *Nature Chemical Biology*, 10(3), pp.178–180.

Haseltine, E.L. & Arnold, F.H., 2007. Synthetic Gene Circuits: Design with Directed Evolution. *Annual Review of Biophysics and Biomolecular Structure*, 36(1), pp.1–19.

Hill, M.S., Zande, P.X.T.V. & Wittkopp, P.J., 2020. Molecular and evolutionary processes generating variation in gene expression. *Nature Reviews Genetics*, pp.1–13.

Hochrein, L. et al., 2018. L-SCRaMbLE as a tool for light-controlled Cre-mediated

recombination in yeast. *Nature Communications*, 9(1), p.1931.

Holland, S.L. et al., 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environmental Microbiology*, 16(6), pp.1729–1740.

Hoshika, S. et al., 2019. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science*, (363), pp.884–887.

Hutchison, C.A. et al., 2016. Design and synthesis of a minimal bacterial genome. *Science*, 351(6280), p.aad6253.

Ichihashi, N. et al., 2013. Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nature Communications*, 4, p.2494.

Isaacs, F.J. et al., 2011. Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science*, 333(6040), pp.348–353.

Isalan, M. et al., 2008. Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, 452(7189), pp.840–845.

Jiménez, A. et al., 2015. Dynamics of gene circuits shapes evolvability. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), pp.2103–2108.

Joyce, G.F. & Szostak, J.W., 2018. Protocells and RNA Self-Replication. *Cold Spring Harbor perspectives in biology*, 10(9).

Kaltenbach, M. & Tokuriki, N., 2014. Dynamics and constraints of enzyme evolution. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 322(7), pp.468–487.

Kaneko, K., 2007. Evolution of robustness to noise and mutation in gene expression dynamics. *PLoS ONE*, 2(5), p.e434.

Kannan, K. & Gibson, D.G., 2017. Yeast genome, by design. *Science*, 355(6329), pp.1024–1025.

Koonin, E.V. & Novozhilov, A.S., 2017. Origin and Evolution of the Universal Genetic Code. *Annual review of genetics*, 51(1), pp.45–62.

Kuwahara, H. & Soyer, O.S., 2012. Bistability in feedback circuits as a byproduct of evolution of evolvability. *Molecular Systems Biology*, 8, p.564.

Lagator, M. et al., 2017. Regulatory network structure determines patterns of intermolecular epistasis. *eLife*, 6, pp.1–22.

Lajoie, M.J. et al., 2013. Genomically recoded organisms expand biological functions. *Science*, 342(6156), pp.357–360.

Lavickova, B., Laohakunakorn, N. & Maerkl, S.J., 2020. A partially self-regenerating synthetic cell. *Nature Communications*, pp.1–11.

Lee, K.Y. et al., 2018. Photosynthetic artificial organelles sustain and control ATP-dependent reactions in a protocellular system. *Nature Biotechnology*, 36(6), pp.530–535.

Lehner, B., 2011. Molecular mechanisms of epistasis within and between genes. *Trends in genetics : TIG*, 27(8), pp.323–331.

Lemire, S., Yehl, K.M. & Lu, T.K., 2018. Phage-Based Applications in Synthetic Biology. *Annual Review of Virology*, 5(1), pp.453–476.

Li, P. et al., 2018. Morphogen gradient reconstitution reveals Hedgehog pathway design principles. *Science*, 360(6388), pp.543–548.

Liti, G., 2018. Yeast chromosome numbers minimized using genome editing. *Nature*,

560(7718), pp.317–318.

Liu, C.C. & Schultz, P.G., 2010. Adding new chemistries to the genetic code. *Annual Review of Biochemistry*, 79, pp.413–444.

Liu, C.C. et al., 2018. Toward an orthogonal central dogma. *Nature Chemical Biology*, 14(2), pp.103–106.

Liu, W. et al., 2018. Rapid pathway prototyping and engineering using in vitro and in vivo synthetic genome SCRaMbLE-in methods. *Nature Communications*, pp.1–12.

Luo, J. et al., 2018. Karyotype engineering by chromosome fusion leads to reproductive isolation in yeast. *Nature*, 560(7718), pp.392–396.

Luo, Z. et al., 2020. Compacting a synthetic yeast chromosome arm. pp.1–18.

Luo, Z. et al., 2018. Identifying and characterizing SCRaMbLEd synthetic yeast using ReSCuES. *Nature Communications*, 9(1), p.1930.

Ma, L. et al., 2019. SCRaMbLE generates evolved yeasts with increased alkali tolerance. *Microbial Cell Factories*, pp.1–11.

Ma, N.J. & Isaacs, F.J., 2016. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell systems*, 3(2), pp.199–207.

Malyshev, D.A. et al., 2014. A semi-synthetic organism with an expanded genetic alphabet. *Nature*, 509(7500), pp.385–388.

Mangan, S. & Alon, U., 2003. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, 100(21), pp.11980–11985.

Matsumura, S. et al., 2016. Transient compartmentalization of RNA replicators prevents extinction due to parasites. *Science*, 354(6317), pp.1293–1296.

Mizuuchi, R. & Ichihashi, N., 2018. Sustainable replication and coevolution of cooperative RNAs in an artificial cell-like system. *Nature Ecology & Evolution*, 2(10), pp.1654–1660.

Mukai, T. et al., 2017. Rewriting the Genetic Code. *Annual review of Microbiology*, 71, pp.557–577.

Mukherji, S. & van Oudenaarden, A., 2009. Synthetic biology: understanding biological design from synthetic circuits. *Nature Reviews Genetics*, 10(12), pp.859–871.

Mushegian, A.R. & Koonin, E.V., 1996. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proceedings of the National Academy of Sciences*, 93(19), pp.10268–10273.

Neumann, H. et al., 2010. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature*, 464(7287), pp.441–444.

Nghe, P. et al., 2020. Predicting Evolution Using Regulatory Architecture. *Annual Review of Biophysics*, 49(1), pp.181-197.

Nielsen, J. & Keasling, J.D., 2016. Engineering Cellular Metabolism. *Cell*, 164(6), pp.1185–1197.

Niu, W., Schultz, P.G. & Guo, J., 2013. An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS chemical biology*, 8(7), pp.1640–1645.

Nyerges, Á. et al., 2018. Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. *Proceedings of the National Academy of Sciences*, 115(25), pp.E5726–E5735.

Paaby, A.B. & Rockman, M.V., 2013. The many faces of pleiotropy. *Trends in Genetics*, 29(2), pp.66–73.

Packer, M.S. & Liu, D.R., 2015. Methods for the directed evolution of proteins. *Nature*

Reviews Genetics, 16(7), pp.379–394.

Payne, J.L. & Wagner, A., 2019. The causes of evolvability and their evolution. *Nature Reviews Genetics*, 20(1), pp.24–38.

Pál, C., Papp, B. & Pósfai, G., 2014. The dawn of evolutionary genome engineering. *Nature Reviews Genetics*, 15(7), pp.504–512.

Peisajovich, S.G., 2012. Evolutionary synthetic biology. *ACS Synthetic Biology*, 1(6), pp.199–210.

Pinheiro, V.B. & Holliger, P., 2012. The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Current Opinion in Chemical Biology*, 16(3-4), pp.245–252.

Pinheiro, V.B. et al., 2012. Synthetic genetic polymers capable of heredity and evolution. *Science*, 336(6079), pp.341–344.

Purnick, P.E.M. & Weiss, R., 2009. The second wave of synthetic biology:from modules to systems. *Nature Reviews Molecular Cell Biology*, 10(6), pp.410–422.

Raman, R., Pinto, C.S. & Sonawane, M., 2018. Polarized Organization of the Cytoskeleton: Regulation by Cell Polarity Proteins. *Journal of Molecular Biology*, 430(19), pp.3565–3584.

Richardson, S.M. et al., 2017. Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Science*, 355(6329), pp.1040–1044.

Rockman, M.V. & Kruglyak, L., 2006. Genetics of global gene expression. *Nature Reviews Genetics*, 7(11), pp.862–872.

Ruder, W.C., Lu, T. & Collins, J.J., 2011. Synthetic Biology Moving into the Clinic. *Science*, 333(6047), pp.1248–1252.

Salathé, M., Van Cleve, J. & Feldman, M.W., 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics*, 182(4), pp.1159–1164.

Santos-Moreno, J. & Schaerli, Y., 2020a. Changing the biological Rosetta stone: the (commercial) potential of recoded microbes. *Microbial Biotechnology*, 13(1), pp.11–13.

Santos-Moreno, J. & Schaerli, Y., 2020b. CRISPR-based gene expression control for synthetic gene circuits. *Biochemical Society Transactions*, 48(5), pp.1979–1993.

Santos-Moreno, J. & Schaerli, Y., 2018. Using Synthetic Biology to Engineer Spatial Patterns. *Advanced Biosystems*, 12, p.1800280.

Santos-Moreno, J. et al., 2020. Multistable and dynamic CRISPRi-based synthetic circuits. *Nature Communications*, pp.1–8.

Sato, K. et al., 2003. On the relation between fluctuation and response in biological systems. *Proceedings of the National Academy of Sciences*, 100(24), pp.14086–14090.

Sánchez-Romero, M.A. & Casadesús, J., 2013. Contribution of phenotypic heterogeneity to adaptive antibiotic resistance. *Proceedings of the National Academy of Sciences*, 111(1), pp.355–360.

Schaerli, Y. & Isalan, M., 2013. Building synthetic gene circuits from combinatorial libraries: screening and selection strategies. *Molecular bioSystems*, 9(7), pp.1559–1567.

Schaerli, Y. et al., 2014. A unified design space of synthetic stripe-forming networks. *Nature Communications*, 5, p.4905.

Schaerli, Y. et al., 2018. Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Molecular Systems Biology*, 14(9), p.e8102.

Shao, Y. et al., 2019. Creating a functional single-chromosome yeast. *Nature*, 560(7718), pp.331–335.

Simon, A.J., d'Oelsnitz, S. & Ellington, A.D., 2019. Synthetic evolution. *Nature Biotechnology*, 37(7), pp.730–743.

Smith, H.O. et al., 2003. Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proceedings of the National Academy of Sciences*, 100(26), pp.15440–15445.

Smith, J.M., 1992. Evolutionary Biology - Byte-Sized Evolution. *Nature*, 355(6363), pp.772–773.

Sun, S.B., Schultz, P.G. & Kim, C.H., 2014. Therapeutic applications of an expanded genetic code. *Chembiochem*, 15(12), pp.1721–1729.

Szymanski, E. & Calvert, J., 2018. Designing with living systems in the synthetic yeast project. *Nature Communications*, 9(1), p.2950.

Tabor, J.J. et al., 2009. A synthetic genetic edge detection program. *Cell*, 137(7), pp.1272–1281.

Tang, T.-C. et al., 2020. Materials design by synthetic biology. *Nature Reviews Materials*, pp.1–19.

Terasaka, N., Azuma, Y. & Hilvert, D., 2018. Laboratory evolution of virus-like nucleocapsids from nonviral protein cages. *Proceedings of the National Academy of Sciences*, 115(21), pp.5432–5437.

Thao, T. et al., 2020. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature*, pp.1–24.

Tumpey, T.M. et al., 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*, 310(5745), pp.77–80.

van Nies, P. et al., 2018. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nature Communications*, 9(1), p.1583.

Vogele, K. et al., 2018. Towards synthetic cells using peptide-based reaction compartments. *Nature Communications*, 9(1), p.3862.

Wang, H., La Russa, M. & Qi, L.S., 2016. CRISPR/Cas9 in Genome Editing and Beyond. *Annual Review of Biochemistry*, 85, pp.227–264.

Wang, H.H. et al., 2009. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460(7257), pp.894–898.

Wang, K., Schmied, W.H. & Chin, J.W., 2012. Reprogramming the genetic code: from triplet to quadruplet codes. *Angewandte Chemie (International ed. in English)*, 51(10), pp.2288–2297.

Wang, L. et al., 2001. Expanding the Genetic Code ofEscherichia coli. pp.1–4.

Wang, Y.-H., Wei, K.Y. & Smolke, C.D., 2013. Synthetic Biology: Advancing the Design of Diverse Genetic Systems. *Annual Review of Chemical and Biomolecular Engineering*, 4(1), pp.69–102.

Wannier, T.M et. al., 2020. Improved bacterial recombineering by parallelized protein discovery. *Proceedings of the National Academy of Sciences*, 117(24):13689-13698.

Wannier, T.M. et. al., 2018. Adaptive evolution of genomically recoded Escherichia coli. *Proceedings of the National Academy of Sciences*, 115(12), pp.3090–3095.

Weber, W. & Fussenegger, M., 2011. Emerging biomedical applications of synthetic

biology. *Nature Reviews Genetics*, 13(1), pp.21–35.

Wightman, E.L.I. et al., 2020. Rapid optimisation of cellulolytic enzymes ratios in. *Biotechnology for biofuels*, pp.1–10.

Wimmer, E. & Paul, A.V., 2011. Synthetic Poliovirus and Other Designer Viruses: What Have We Learned from Them? *Annual review of microbiology*, 65(1), pp.583–609.

Wolpert, L., 1969. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1), pp.1–47.

Wong, B.G. et al., 2018. Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nature Biotechnology*, 36(7), pp.614–623.

Wu, Y. et al., 2018. In vitro DNA SCRaMbLE. *Nature Communications*, 9(1), p.1935.

Xie, M. & Fussenegger, M., 2019. Designing cell function: assembly of synthetic gene circuits for cell biology applications. *Nature Reviews Molecular Cell Biology*, pp.1–19.

Yokobayashi, Y., Weiss, R. & Arnold, F.H., 2002. Directed evolution of a genetic circuit. *Proceedings of the National Academy of Sciences*, 99(26), pp.16587–16591.

Zeymer, C. & Hilvert, D., 2018. Directed Evolution of Protein Catalysts. *Annual Review of Biochemistry*, 87, pp.131–157.

Zhang, W., Mitchell, L.A., Bader & J.S., Boeke, J.D., 2020. Synthetic Genomes. *Annual Review of Biochemistry*, 89(1), pp.77–101.

Zhang, W.H., Otting, G. & Jackson, C.J., 2013. Protein engineering with unnatural amino acids. *Current Opinion in Structural Biology*, 23(4), pp.581–587.

Zhang, Y. & Romesberg, F.E., 2018. Semisynthetic Organisms with Expanded Genetic Codes. *Biochemistry*, 57(15), pp.2177–2178.

Zhang, Y. et. al., 2017. A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proceedings of the National Academy*, 114(6), pp.1317–1322.

Zhang, Y. et. al., 2017. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature*, 551(7682), pp.644–647.