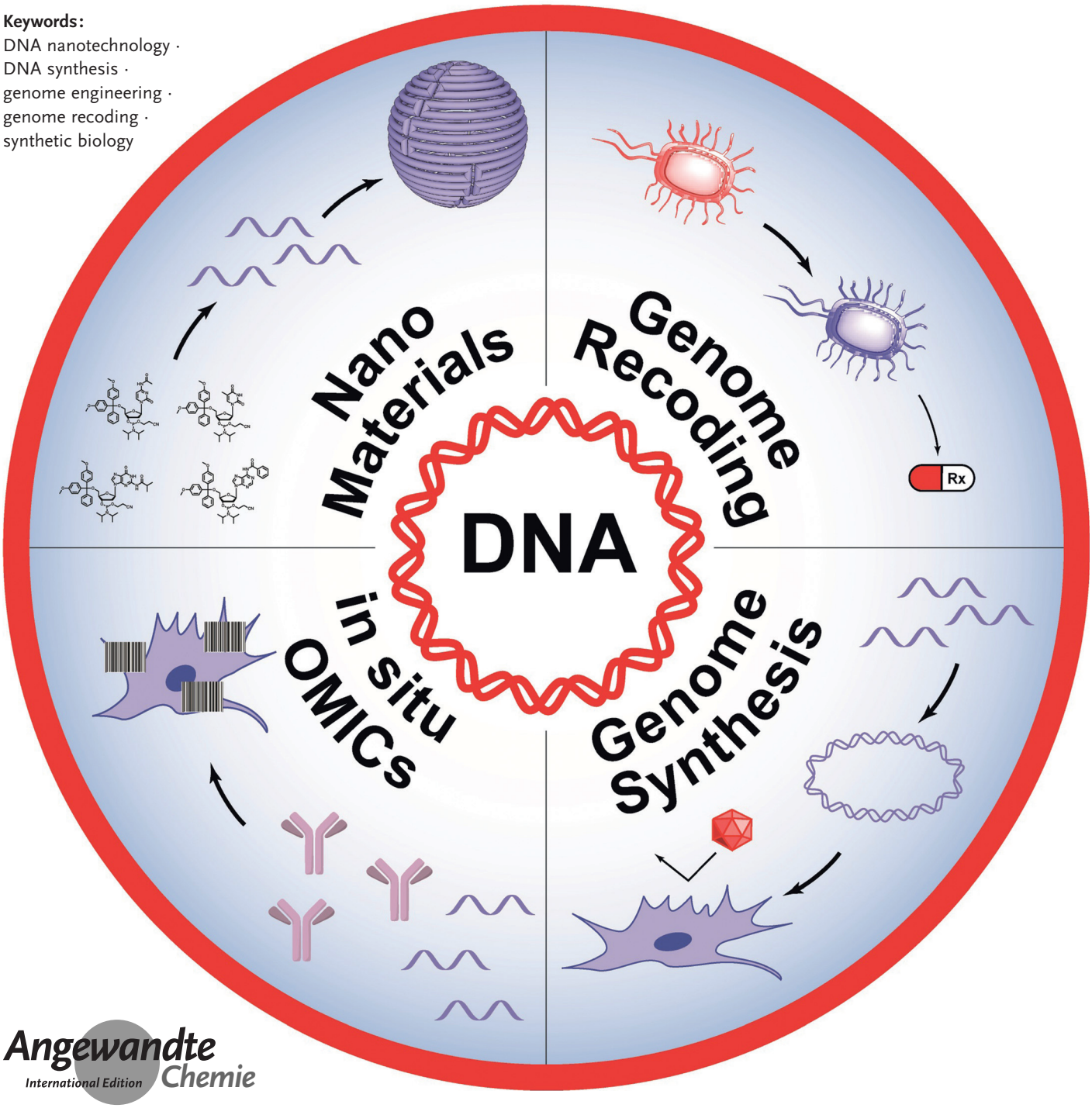
DNA Nanotechnology

International Edition: DOI: 10.1002/anie.201707976
German Edition: DOI: 10.1002/ange.201707976

From Designing the Molecules of Life to Designing Life: Future Applications Derived from Advances in DNA Technologies

*Richie E. Kohman[+], Aditya M. Kunjapur[+], Eriona Hysolli[+], Yu Wang[+], and George M. Church**

Keywords:
DNA nanotechnology · DNA synthesis · genome engineering · genome recoding · synthetic biology



Since the elucidation of its structure, DNA has been at the forefront of biological research. In the past half century, an explosion of DNA-based technology development has occurred with the most rapid advances being made for DNA sequencing. In parallel, dramatic improvements have also been made in the synthesis and editing of DNA from the oligonucleotide to the genome scale. In this Review, we will summarize four different subfields relating to DNA technologies following this trajectory of smaller to larger scale. We begin by talking about building materials out of DNA which in turn can act as delivery vehicles in vivo. We then discuss how altering microbial genomes can lead to novel methods of production for industrial biologics. Next, we talk about the future of writing whole genomes as a method of studying evolution. Lastly, we highlight the ways in which barcoding biological systems will allow for their three-dimensional analysis in a highly multiplexed fashion.

From the Contents

1. Introduction

Technology to read, write, and edit DNA is central to all fields of current biological investigation. For decades, continuing effort has been made to improve upon these three DNA-based technologies. The most rapid advances have been made in the ability to sequence DNA, as the cost per base had dropped by one million-fold in three decades. This massive technological improvement has truly changed the way biology is investigated with sequencing being a commonplace tool. In parallel, technologies to write and edit DNA have seen similar advances in recent years. Due to improvements in instrumentation, the cost of DNA synthesis continues to decrease at rates faster than predicted, while our ability to edit genomes has been dramatically enhanced due to key discoveries in the field of endonucleases.

The technologies enabled by these advances are numerous, and continue to flourish. Due to the decrease in the cost of DNA synthesis and our understanding of the functional properties of various genetic elements, the behavior of entire cells can now be programmed to carry out specific operations. Additionally, DNA can now be used in a non-canonical fashion to influence and study biological systems through the use of multiplexed DNA tags or through the construction of DNA-based materials.

In this Review, we will highlight recent advances in several DNA-based technologies and discuss what their future applications will be. We will follow the trajectory of increasing scale and complexity, beginning with engineering DNA nanomaterials and continuing towards engineering genomes and organisms.

2. How Advancements in DNA Synthesis Will Effect Nanotechnology

The availability of synthetic DNA has allowed researchers to utilize it as a nanofabrication substrate rather than just the substrate for biological information storage.[1] Over several decades, numerous methods for building nanostructures out of DNA have been demonstrated. As expected, progress in DNA nanotechnology is closely influenced by DNA synthesis technology. Here we entertain the question of how this field would be effected if longer and cheaper DNA was readily available. Although relevant for all DNA nanotechnology,[2] here we indulge upon how this technological leap will specifically effect applications of DNA origami-based nanotechnology for in vivo delivery.

DNA origami is a technique developed in 2006 that involves the self-assembly of a large strand of DNA (the scaffold) with many small pieces of DNA (the staples) to generate well-defined nanostructures (Figure 1 A).[3] In a relatively short period of time, numerous advances have been made to the original methodology[4] such as methods to create three-dimensional structures with and without curvature[5] using the assembly of two-dimensional sheets,[6] packing of multiple sheets at a variety of densities,[7] or using a wireframe approach.[8] Design software was created to drastically

[*] Dr. A. M. Kunjapur,[+] Dr. E. Hysolli,[+] Y. Wang,[+] Prof. G. M. Church
Department of Genetics, Harvard Medical School
Boston, MA 02115 (USA)

Dr. R. E. Kohman,[+] Y. Wang,[+] Prof. G. M. Church
Wyss Institute for Biologically Inspired Engineering
Harvard University
Boston, MA 02115 (USA)

[+] These authors contributed equally to this work.

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/anie.201707976.

This article is part of the Special Issue to commemorate the 350th anniversary of Merck KGaA, Darmstadt, Germany. More articles can be found at http://doi.wiley.com/10.1002/anie.v57.16.

simplify the design process,[9] and techniques were developed that improved both folding[10] and purification[11] of the nanostructures. It is now possible to design, create, and characterize a novel DNA nanostructure within a few weeks.[12]

DNA origami nanostructures have a variety of unique properties that make them ideal for biological applications such as drug delivery, enzymatic nanoreactors, and bioanalysis:[4] 1) The design space is large, as it is possible to create uniform structures of arbitrary two- and three-dimensional shapes; 2) nanostructure creation is relatively straightforward with most designs folding successfully in one thermal annealing step over the course of 12 to 48 hours; 3) DNA origami nanostructures are highly addressable as functional handles can be positioned at precise locations along their surfaces; and 4) structures can possess dynamic properties such as the ability to undergo conformational changes when exposed to intrinsic, such as endogenous intracellular biomolecules or pH changes, or extrinsic triggers, such as exposure to chemicals[13] or light.[14] Taken together, these properties make DNA nanostructures optimal candidates to address many of the unmet challenges in bio-nanotechnology, such as the in vivo targeting and delivery of therapeutics and diagnostics. Progress has already been made toward this end. Chemotherapeutic intercalators have been bound to packed origami helices and delivered to cancer cells after being internalized (Figure 1B).[15] Small molecules neurotransmitters as well as full proteins have been released from the cavities of DNA nanostructures with light-triggered uncaging (Figure 1C).[16] Also, we have recently designed cancer-killing nano-robots which underwent structural conformations in the presence of environmental cues to expose cell-killing antibodies previously hidden within their interior (Figure 1D).[17] Lastly, to address stability in in vivo environments, surface modifications using lipid bilayers (Figure 1E)[18] and hydrophilic polymers[19] have been developed.

Despite these advances, several issues need to be addressed before DNA origami nanostructures can effectively scale for many biological applications. The cost of the synthetic oligonucleotides used for the staple DNA is still substantial compared to synthetic materials and must be lowered to be able to create affordable therapies and diagnostics. This will most likely require substantial improvements in DNA synthesis technology, such as using biosynthetic methods as opposed to purely chemical ones.[20] An additional issue is that the scaffold DNA exceeds the length that can be chemically synthesized and thus prevents

Richie E. Kohman is a research scientist and member of the Advanced Technology Team at the Wyss Institute for Biologically Inspired Engineering at Harvard University in the laboratory of Professor George Church. His reseach interests lie in nucleic acid chemistry, nanotechnology, and neurotechnology. He obtained his B.S. in Chemistry from Santa Clara University and his Ph.D. in Chemistry from the University of Illinois Urbana-Champaign. He then pursued postdoctoral research in the Department of Biomedical Engineering at Boston University and as an affliate in the Media Lab at the Massachusetts Institute of Technology. Prior to working at the Wyss Institute, he served as Group Leader at Expansion Technologies, Inc.

Aditya M. Kunjapur is a Postdoctoral Research Fellow in the laboratory of Professor George Church in the Department of Genetics at Harvard Medical School. He is focused on developing high-throughput methods to selectively introduce novel functional groups in metabolic and protein engineering. He obtained his B.S. in Chemical Engineering from the University of Texas at Austin in 2010 and his Ph.D. in Chemical Engineering from the Massachusetts Institute of Technology in 2015. At MIT, Aditya worked under the supervision of Professor Kristala Prather as an NSF GRFP Fellow. He was recently awarded a Synthetic Biology Biosecurity Fellowship in recognition of his work on biocontainment.

Eriona Hysolli is a Postdoctoral Researcher in the laboratory of Professor George Church at Harvard Medical School. Her research interests include site-specific DNA integration, DNA delivery methods, and mammoth de-extinction. She received a B.A. in Cell Biology/Neuroscience, Chemistry, and German from Rutgers University in 2008, and a Ph.D. from Yale University in 2015 studying the role of microRNAs in the reprogramming of human induced pluripotent cells.

Yu Wang is a Ph.D. student in the laboratory of Professor George Church in the Department of Genetics at Harvard Medical School and Wyss Institute for Biologically Inspired Engineering at Harvard University. His research interests are focused on highly multiplexed in situ molecular analysis with DNA barcoding systems and antibody engineering. He obtained a dual-bachelor degree from Nanjing University and the University of Southampton and is currently pursuing his PhD degree in biological and biomedical science at Harvard Medical School.

George M. Church is the Robert Winthrop Professor of Genetics at Harvard Medical School and a founding member of the Wyss Institute for Biologically Inspired Engineering at Harvard. He is widely regarded as one of the founders of the field of synthetic biology and has made pioneering discoveries in many aspects of genomic science. He developed methods used for the first genome sequence and helped initiate the Human Genome Project, the Personal Genome Project, and the BRAIN Initiative. He has received numerous awards including the Bower Award and Prize for Achievement in Science and election to the National Academy of Sciences and Engineering.

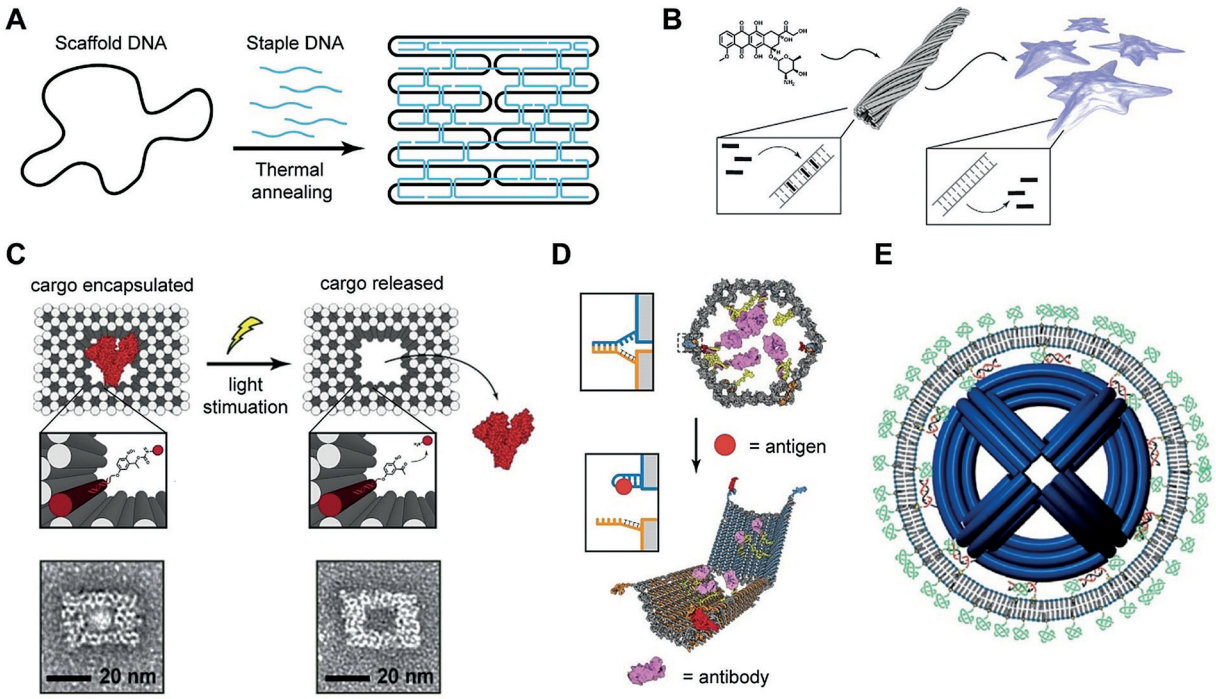


Figure 1. Controlled release from DNA origami. A) Schematic depiction of the self-assembly of staple and scaffold DNA to produce a DNA origami nanostructure. B) A twisted DNA origami tube can bind the DNA-intercalating drug doxorubicin and release it in the presence of cancer cells. C) Light-triggered release of molecular cargo from the interior of a DNA nanocage. D) A cancer-killing DNA nano-robot. Binding to the aptamer handles holding the structure closed triggers the exposure of cell-killing antibodies. E) The coating of a DNA nanostructure with a lipid bilayer protects against nuclease digestion and in vivo immune activation. Figures for (B)–(E) from references [15a, 16–18], respectively.

researchers from easily obtaining custom sequences for a particular DNA origami design. Currently, (with some exceptions)[21] most researchers utilize the genome of the m13 bacteriophage[22] as a scaffold due to its ease of availability. The lack of readily available long, inexpensive, and custom DNA prevents researchers from having the tools to discover the fundamental folding rules for complex nanostructures and prohibits one's ability to predict how well a new structure will fold a priori. With the availability of custom sequences, nanostructure folding could be optimized by screening the melting temperatures within subdomains of the structure. This would allow for new structures to be created at high yields and eliminate the costly step of purifying desired product away from improperly folded structures.

The ability to create monodisperse DNA nanostructures at a large scale will lead to advances in numerous in vivo therapeutic and diagnostic applications. It has been a long-standing challenge to create a "magic-bullet" delivery vehicle which could evade immune detection and target specific cell types of interest in the body. This is especially true with delivery to the brain due to the difficulty in bypassing the blood–brain barrier. Although safe, surgical procedures towards this end have been developed,[23] systemic delivery would be advantageous due to their non-invasiveness.[24] Some success has been obtained towards creating brain-targeting nanostructures by functionalizing polymer or lipid nanoparticles with natural or synthetic peptide signals capable of inducing brain entry[25] as well selectively targeting neurons.[26] These attempts suffer from the general lack of control in the three-dimensional placement of bioactive motifs. DNA nanostructures have the potential to make for more efficient and effective targeting agents because the number and location of attached motifs can be precisely controlled. Figure 2 outlines the implementation for one such brain-targeting nanocarrier. A DNA nanostructure containing polyvalent, spatially patterned bioactive cues (such as virus-derived peptides)[27] could be designed to contain encapsulated therapeutic or sensing molecules. Coated with inert polymers[19] or lipids,[18] these structures would remain protected in the bloodstream after injection until they reach the brain vasculature and penetrate into the brain tissue via receptor-mediated transcytosis through the vascular endothelial cell layer.[29] Because of the addressability of DNA origami nanostructures, multiple targeting groups can be patterned on the particles' surface allowing additional cellular targeting signals to also be attached.[26a] Once at their target intracellular destination, triggered conformational changes can release therapeutics and/or expose sensing molecules. Future technologies such as these remain on the horizon but will only be achievable with advancements in fundamental DNA synthesis technology.

3. The Future of Microbial Genome Recoding and Synthesis

Along with significant decreases in DNA synthesis cost, advancements in multiplexed genome engineering and edit-

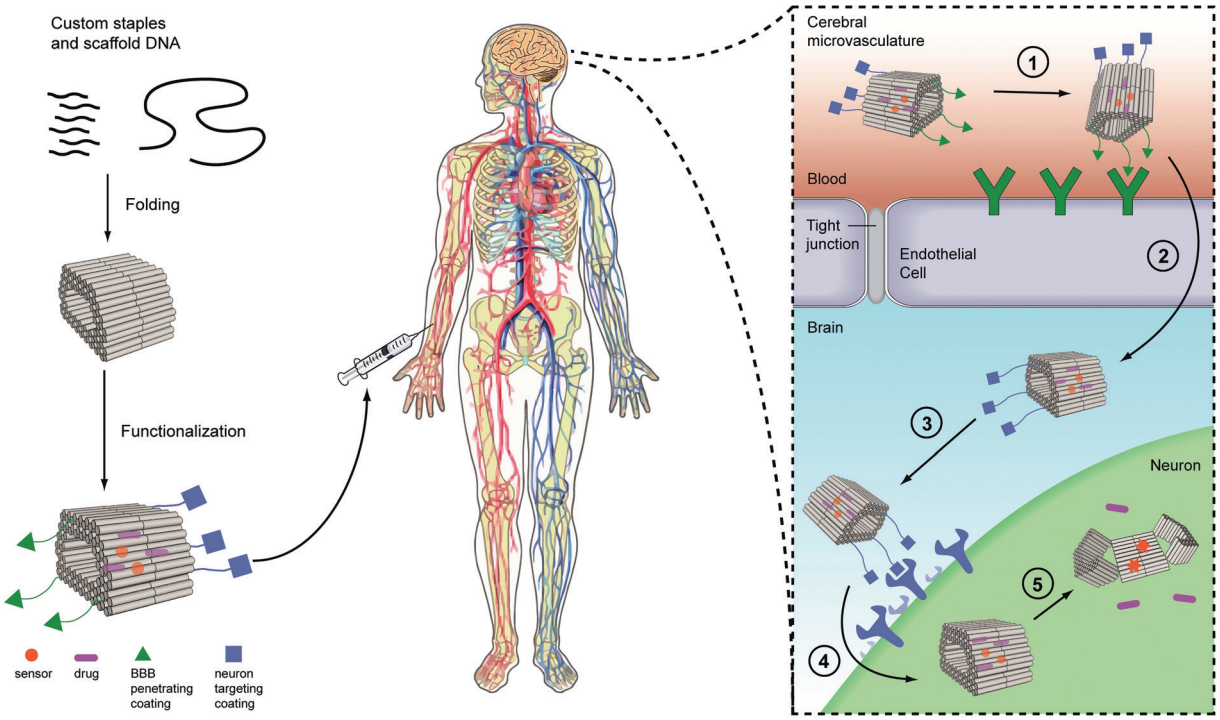


Figure 2. The construction of a DNA nano-robot capable of penetrating into the brain and delivering cargo to targeted neurons. Custom designed scaffold and staple DNA are synthesized and optimized for a desired nanostructure design. Structures are then modified to contain encapsulated drugs and sensors as well as a protective coating containing blood–brain barrier penetrating and neuron targeting motifs. Systemic injection of the nanostructures enables them to access the brain microvasculature where they can bind to vascular endothelial cells (1) and enter into the brain through receptor-mediated transcytosis (2). Once in the brain tissue, nanostructures will bind to target neurons (3), enter cells (4), and release/expose cargo (5) as the result of a structural transformation triggered by endogenous signals within the cell.

ing technologies enabled the construction of recoded genomes,[30] and parallel advancements in genome design, synthesis, and assembly have ushered in an era of synthetic microbial genomes spanning bacteria[31] and yeast.[32] These organisms have the potential to enhance industrial production of chemicals and biologics given distinct advantages that can be endowed through large-scale genetic alterations. In this section, we discuss some of these advantages, such as multi-virus resistance, biocontainment by synthetic auxotrophy, and non-standard amino acid (NSAA) incorporation.

Recoding is the process of substituting codons for synonymous codons that encode the same amino acid. Because the genetic code is degenerate and nearly universally conserved, a codon can become unassigned by systematic substitution of all instances of the codon across an entire genome and removal of associated translation machinery. One major industrially relevant advantage of genomically recoded organisms (GROs) is their decreased likelihood of viral infection given the reliance of viral replication on host machinery for all 64 codons.[30c,33] Viral contamination of industrial facilities has led to multi-million dollar losses in the past.[34] Although experiments in environments outside of the lab have not been conducted, virus resistance should confer an advantage outside of the lab. Experiments from our lab and the Isaacs lab have demonstrated that recoding increases resistance to at least five different viruses.[30c,33] The ability of GROs to resist horizontal gene transfer and maintain virus resistance in small microbial communities has also been investigated.[33] We expect even greater multi-virus resistance for an *E. coli* strain containing a 57-codon genome under construction. This genome is being assembled completely from synthetic DNA and the following codons have been removed by synonymous replacement: UAG, AGC, AGU, UUG, UUA, UGC, AGA, and AGG.

Based on these observations, it is prudent to require biocontainment for GROs as a safeguard against unchecked proliferation. Although GROs warrant increased consideration for biocontainment, recoding enables new biocontainment strategies. A highly effective strategy is synthetic auxotrophy, wherein an organism is engineered to be dependent upon the presence of a synthetic NSAA for proper expression and function of essential proteins (Figure 3A).[35] Synthetic auxotrophy has successfully limited the ability of an engineered organism to grow in non-permissive media conditions (i.e., media lacking NSAA) to below 1 in 10^{12} cells.[35a] Further experiments demonstrated that media containing bacterial lysate permitted escape for non-synthetic but not synthetic auxotrophs, and that conjugation involving synthetic auxotrophs containing multiple NSAA-dependent genes at worst resulted in inviable cells or cells with their entire genomes overwritten, thereby preserving the contained population.[35a] Thus, synthetic auxotrophy may enable GRO use in a variety of contexts requiring spatiotemporal control of cell proliferation. For example, engineered probiotics would benefit from biocontainment by synthetic auxotrophy to prevent propagation within the gut and after excretion.

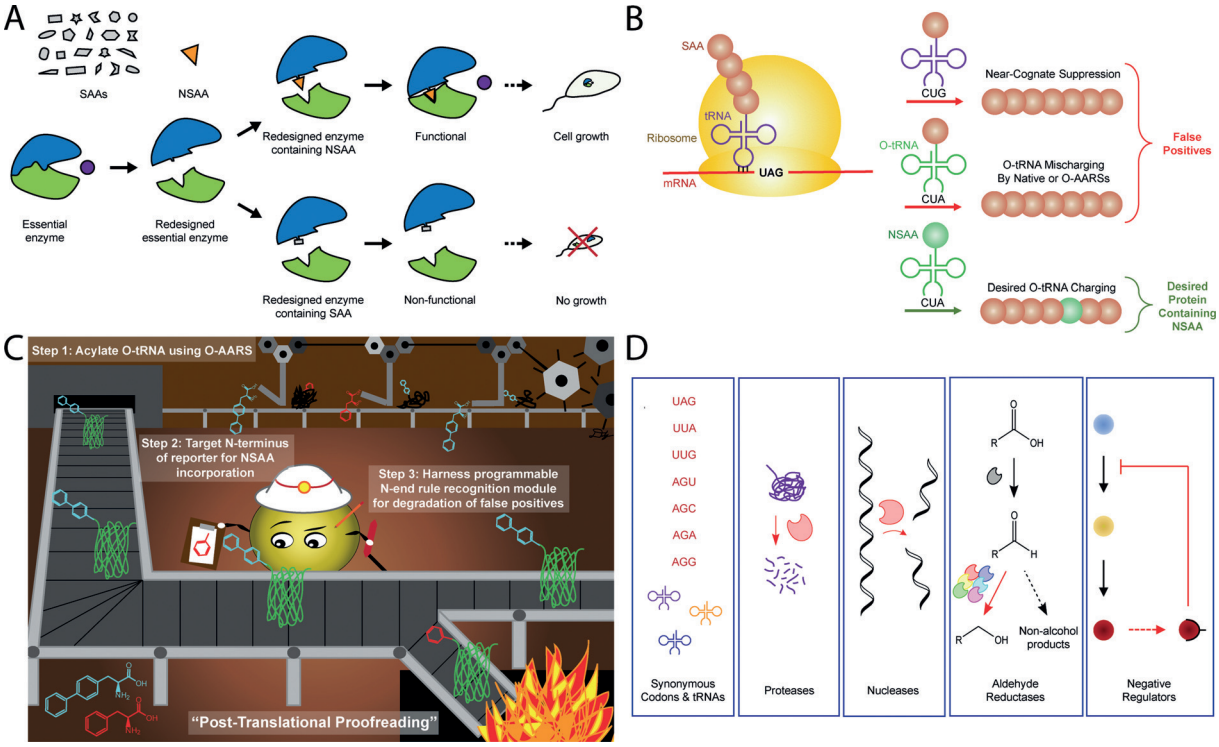


Figure 3. Recoded and synthetic microbial genomes present new engineering opportunities. A) Conceptual depiction of biocontainment by synthetic auxotrophy strategy in which essential protein function is dependent upon incorporation of a non-standard amino acid (NSAA). While this cartoon depicts a split protein for simplicity, published strategies include burying NSAAs inside protein cores. B) Translational outcomes upon ribosomal encounter with repurposed codons. Target proteins produced without NSAAs, or false positives, are possible due to near-cognate suppression, mischarging by orthogonal aminoacyl tRNA synthetases, or mischarging of orthogonal tRNA by endogenous aminoacyl tRNA synthetases. C) Cartoon illustrating the development of a synthetic quality control strategy ("post-translational proofreading") that results in degradation of proteins containing misincorporation events at the N-terminus and facilitates evolution for selective orthogonal translation machinery. Cartoon kindly edited by Dr. Jeffrey E. Ting (U. Chicago). D) Undesired and often redundant genomic elements for potential removal in future genomes of industrial strains to aid metabolic and protein engineering.

Similarly, synthetic auxotrophy could serve as an added safeguard for limited environmental deployment of engineered strains for use in remediation or as sentinels.

Common to initial demonstrations of genomic recoding, synthetic auxotrophy, and synthetic genomes are undesired decreases in organismal fitness. Fitness differences are exacerbated in defined media, which is frequently used in industrial contexts. Lower fitness from genome engineering errors can be addressed by further genome engineering aided by novel modeling strategies.[36] Another exciting approach that simultaneously improves fitness and assesses the robustness of genome design is adaptive laboratory evolution. We recently evolved GROs for the first time and for more than 1000 generations. In independent sequenced clones, we observed selective mutations that compensated for intentional genome design alterations as well as unintentional off-target errors that occurred during engineering.[37] Similar evolution of GROs and synthetic organisms in the future would provide unique insights into genome pliability.

A major opportunity offered by genome recoding and synthesis is the expansion of the genetic code to include dedicated codons for NSAAs.[38] Besides enabling biocontainment by synthetic auxotrophy, NSAAs expand the structural and functional diversity of proteins for basic and applied research in academic and commercial settings. A growing share of all medicines are biologics, and numerous start-ups are currently striving to incorporate NSAAs in their products (Ambrx,[39] GRO Biosciences,[40] Sutro,[41] and Synthorx[42]). These companies and their academic labs of origin have complementary technologies that enable NSAAs in antibody–drug conjugates,[43] bispecific antibodies,[44] and therapeutic proteins that exhibit improved affinity and pharmacodynamics.[45] Additionally, NSAAs can improve the catalytic properties of industrially relevant enzymes.[46] In these three studies, enzyme sites were investigated for improved catalytic efficiency or selectivity using all 20 standard amino acids (SAAs) or NSAAs. In each report, substitution using the best NSAA resulted in better performance than using any of the 20 SAAs. Furthermore, Ugwumba and co-workers claimed that prior directed evolution efforts that surveyed the use of SAAs more broadly (resulting in several hundred thousand variants) could not achieve the improvement obtained using the best NSAA.[46b]

These examples only reveal the tip of the iceberg of possibilities enabled by NSAAs. As genomes are more thoroughly recoded, or as they are constructed with alternative base pairs beyond A–T and G–C,[47] multiple codons will become available for simultaneous use of several different NSAAs. We foresee that the ability to incorporate multiple types of NSAAs into proteins will lead to useful

hybrids of several kinds of biological macromolecules. However, we have recently observed that a major limitation is the promiscuity of previously engineered orthogonal translation systems, which has been discussed in the literature[48] and is perhaps as great a hurdle as the availability of new dedicated codons. These systems often incorporate structurally similar SAAs as well as numerous NSAAs if supplied, leading to the presence of false positive protein production (Figure 3B). Low fidelity of orthogonal translation machinery also poses a hurdle for efforts to metabolically engineer NSAA biosynthetic pathways given their likelihood of generating structurally similar NSAA precursors. Yet, the use of metabolic engineering for NSAA biosynthesis can circumvent industrially relevant issues of supplementation cost and limited transport across cell membranes. A biosynthetic pathway has been engineered for the NSAA *p*-aminophenylalanine.[49]

To improve the selectivity of orthogonal translation systems, we recently engineered a synthetic quality control process that enriches for proteins containing NSAAs (Figure 3C).[50] Our new post-translational proofreading method harnesses the N-end rule of protein degradation, which is a natural protein regulatory and quality control pathway conserved across prokaryotes and eukaryotes.[51] The N-end rule states that the half-life of a protein is determined by its N-terminal residue. Because bulky hydrophobic standard amino acids appearing at the N-terminus are specifically recognized by pathway components as substrates for degradation (and are therefore "N-end destabilizing"), we hypothesized that many NSAAs that contain minor deviations from these structures may not be recognized (and may instead be "N-end stabilizing"). We classified the N-end rule status of numerous NSAAs and altered NSAA recognition by rationally engineering the *E. coli* ClpS adaptor protein. This proofreading strategy can discriminate incorporation of desired NSAAs from both related SAAs and undesired NSAAs, thus minimizing the false positive protein production that can result from traditional amber suppression methods. To ensure high-fidelity incorporation for applications that feature NSAAs at positions besides the N-terminus, we demonstrated how proofreading can aid screens for directed evolution of orthogonal translation systems. After evolution, the more selective version of the orthogonal translation system used for biocontainment lowered escape frequency while increasing fitness of all tested biocontained *E. coli* strains.[50] We expect that proofreading will dramatically facilitate efforts to engineer systems for simultaneous incorporation of multiple NSAAs, including D-amino acids.

In addition to facilitating genetic code expansion, genome synthesis can include other genome modifications relevant for chemical and pharmaceutical production (Figure 3D). Protein and metabolic engineers may benefit from a blank canvas genome after removal of undesired and often redundant endogenous activities such as proteases,[52] nucleases,[53] aldehyde reductases,[54] and negative regulatory elements.[55] Genomes could be "defragmented" or refactored such that genes encoding similar organism functions were clustered and independently controlled.[56] Alternatively, pending the development of numerous independently and tightly regulated expression systems, it may be most valuable for a microbial strain to have a maximal genome. Such a generalist workhorse strain could be called "*E. pluri*", where *E. coli* meets "*E pluribus unum*" ("out of many, one").[57] This maximal genome concept is a counterpoint of the minimal genome, a concept whose practical or industrial advantage has not been thoroughly demonstrated. A minimal genome completely relies on its few genes, which has negative consequences for its evolution. Despite having an additional DNA synthesis burden, a larger genome can contain more functionality, and by containing these functionalities as well as redundancies it can bypass or accelerate evolution. Notably, new gene functions often emerge after gene duplication.[58]

Eventually we expect synthetic recoded genomes and engineered translation machinery to facilitate the construction of organisms exhibiting mixed or complete opposite chirality from natural life (Figure 4). Nature has nearly uniformly employed D-sugars and L-amino acids in polysaccharides and proteins, respectively. D-amino acids (D-AAs) exist naturally in some secreted polypeptides[59] and in bacterial peptidoglycan. Additionally, they have been shown to make proteins less susceptible to degradation[60] and have been used to advance bacterial cell labeling and microscopy.[61] Early work in this area demonstrated that D-AA incorporation by the ribosome could be improved by engineering ribosomal RNA.[62] D-AA compatibility with translation was later reevaluated, with the finding of three classes of D-AA incorporation efficiencies. Attempts toward

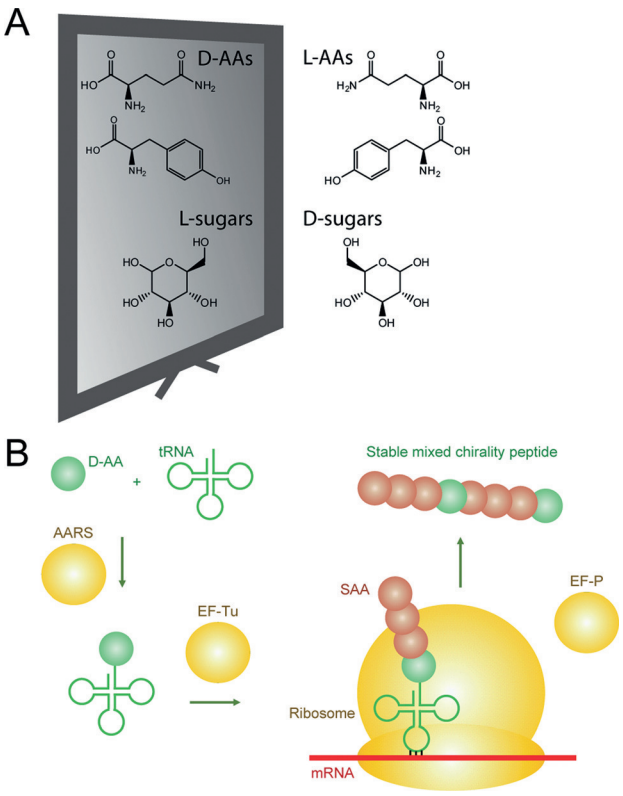


Figure 4. Mixed chirality biosynthesis. A) Cartoon depicting natural forms of amino acids (L) and sugars (D) and molecules of opposite chirality shown in mirror. B) Components of translation machinery that can be targeted to improve ribosomal synthesis of peptides containing D-amino acids.

consecutive incorporation of two D-AAs failed whereas double-incorporation efficiency was restored when two or three L-AAs were inserted between the D-AAs.[63] Besides the ribosome, other translation factors influence D-AA incorporation. An in vitro translation system using Flexizymes to misacylate tRNAGly with D-AAs enabled single incorporation of 17 out of 18 tested D-AAs into a polypeptide, remarkably bypassing known preferences for thermodynamic compensation in elongation factor EF-Tu.[64] The observation that elongation factor EF-P alleviates ribosome peptidyl transferase stalling at polyproline stretches on a growing ribosomal peptide chain[65] motivated our lab to use EF-P to alleviate stalling that occurs between two consecutive D-amino acids.[66] Recent work has demonstrated consecutive elongation of D-AAs in translation using complementary strategies.[67] The ability of fully D-AA-containing enzymes to act on opposite chirality has been demonstrated using a D-AA polymerase on an L-DNA template, which rendered a functionally active L-DNAzyme.[68] Taken together, efforts targeting the ribosome, elongation factors, and fully D-AA-containing enzymes have improved our ability to incorporate D-AAs and demonstrated the utility of enzymes exhibiting opposite chirality. Further advancements are required to achieve the kind of "mirror life" organism that would be capable of complete genetic isolation and production of enantioselective catalysts and highly stable therapeutics because of opposite chiralities that nature is unlikely to have evolved proteases for.

4. Writing Genomes to Explore Evolution

The launch of Genome Project–Write (GP-Write) in 2016 is beginning to address the next research phase of genomics with pilot projects extending beyond sequencing to writing complete genomes. Precedents have been set with the synthesis and assembly of recoded *E. coli* fragments, the complete small genome of *Mycobacterium genitalium* and the full set of yeast chromosomes (Sc2.0)—collective work that will be used as examples and scaffolds for writing larger and more complex genomes.[69] Synthesis and an extended repertoire of genome editing technologies like CRISPR/Cas9 with its derivatives and alternatives, recombinases, meganucleases, zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) and delivery methods such as physical/chemical, virus-mediated transduction, yeast spheroplasts and cell–cell fusion will make genome writing as routine as sequencing (Figure 5A).

The drawbacks of these technologies include lack of high-efficiency editing, delivery and assembly, and low multiplexability, all of which are required to tackle giga base-pair size mammalian genomes. A need to adapt, evolve or emulate strategies like MAGE (multiplexed automated genome engineering) and CAGE (conjugative assembly genome engineering) arises. These tools work well for making a large number of changes simultaneously, and assembling the edited DNA fragments in the fraction of the time in prokaryotic genomes.[70] Another example of incorporating a large number of changes in prokaryotes is the λ-Red phage system; it enables a double-stranded DNA fragment to be digested into single-stranded DNA and incorporated into the lagging strand during replication, thus creating a large number of mutations.[71] This recombineering technology is an exciting prospect that has been shown to work in conjunction with CRISPR/Cas9 to replace large modified DNA fragments in prokaryotic genomes.[72] The large modified DNA in the form of human and mammalian artificial chromosomes (HACs and MACs) can be delivered to human cells via microcell-mediated chromosome transfer (MMCT). This is a challenging but useful technique which has been recently and significantly improved via expression of viral envelope proteins.[73] Successful adaptation and use of these or similar systems in eukaryotes may be essential to accelerate the fast manipulation of large genomes beyond synthesis.

Why engineer a genome? Among the goals of GP-Write are ultrasafe cell lines resistant to cancer and viruses, multipurpose cell lines to manufacture therapeutics, xenotransplantation, a prototrophic human cell and disease modeling among others that will have enormous implications in human health.[74] Endowing genomes with new functionalities, new circuits to produce biologics, or generating transplantable organs and organoids with universal immune compatibility but without germ cell development as a "fail-safe" measure all belong to the medicine of the near future (Figure 5B–D). But almost three decades after the human genome project (HGP) what can synthesizing and assembling a genome teach us about the building principles of genomes and how they evolve over time?

There are an estimated 21000 genes in the 3.3 Gb human genome. There is a staggering amount of DNA that makes up the non-coding portion of the genome and most of it is still of unknown function. Design of a smaller functional eukaryotic genome could be invaluable to discover the principles behind the expansion of eukaryotic genomes and the new roles the non-coding elements acquire or permutate for gene regulation and shaping genome architecture. Eliminating DNA- and retrotransposons can be an important first step to elucidate the role of these elements in genome stability and cell identity. Our lab and others are targeting active repetitive elements like Alu and long interspersed nuclear elements (LINE1) (10% and 17% of genome, respectively) genome-wide to understand their implication in cancer, development and aging. Successful simultaneous inactivation of a family of repetitive elements in the pig genome via CRISPR/Cas9 was reported from our lab in a significant step towards xeno-transplantation efforts.[75] However, a complete removal of both active and inactive elements from the genome is an alternative path to understanding their function as these genomic sequences contain transcription factor binding motifs or resemble enhancers among other yet-to-be understood functions.[76] A smaller genome can also be conducive to understanding the minimal amount of genes needed for propagation of a mammalian cell. The minimal genome of *M. mycoides* revealed an additional 79 essential genes with no assigned roles that are being currently investigated and a similar strategy applied to a mammalian genome could shed light on its fundamental organization.[69b]

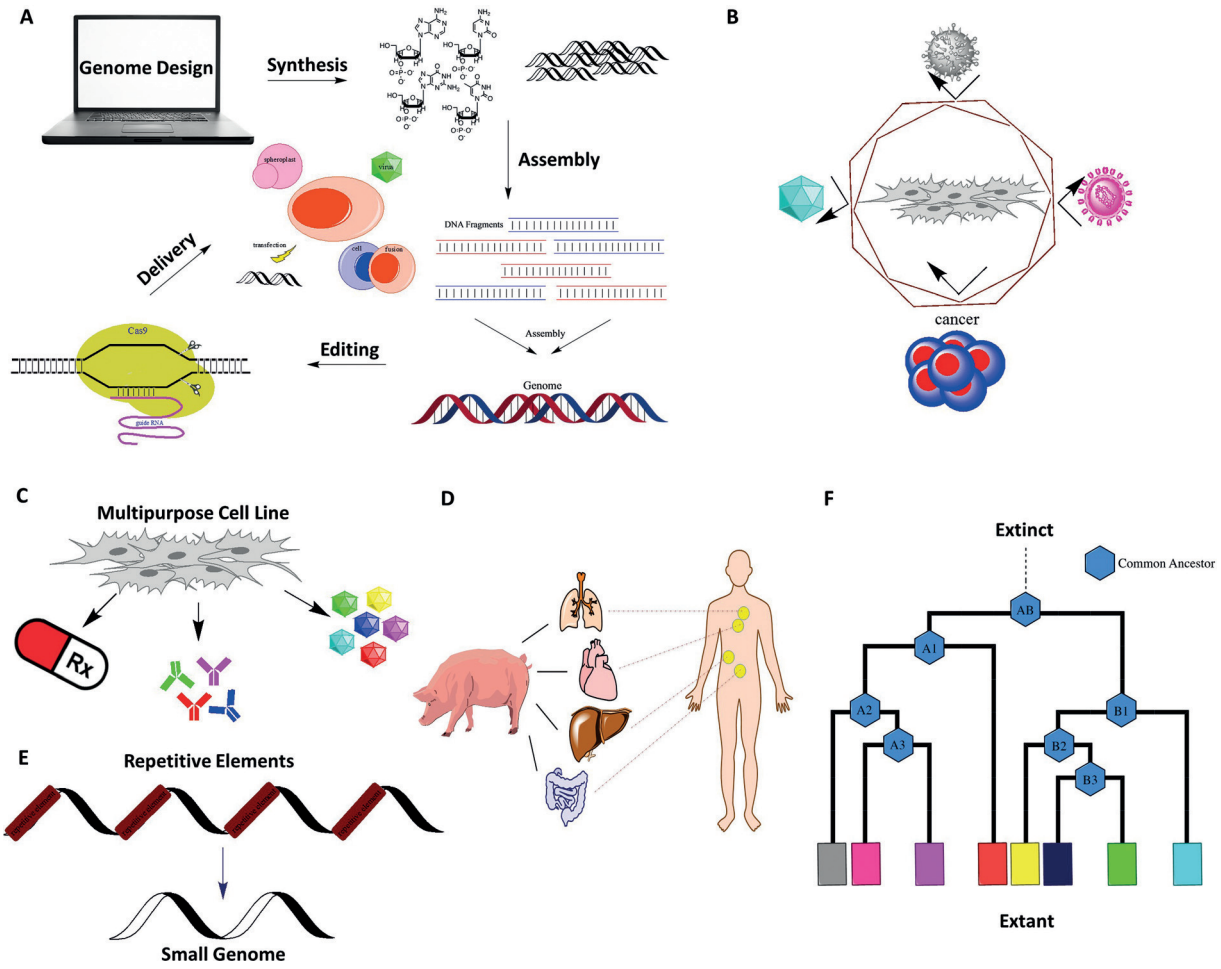


Figure 5. Genome engineering in health and biology. A) Genome design in silico coupled with DNA synthesis precedes genome assembly, editing and delivery of novel genomes. B–D) GP-Write pilot projects. B) Building cell lines resistant to virus and cancer. C) A multipurpose cell line to produce various biologics. D) Humanizing the pig genome for xenotransplantation of pig organs to humans. E,F) Engineering genomes to elucidate evolution. E) Designing genomes free of repetitive elements to elucidate their role in development and evolution of eukaryotic genomes. F) Resurrecting and synthesizing ancient genomes and common ancestor sequences to study evolution.

Genome synthesis can help us further our understanding on how they evolve by resurrecting ancient genomes or realizing the in-silico reconstruction of ancient genomes beyond physical DNA recovery (Figure 5F). Our lab is involved in approaches to de-extinct the woolly mammoth and therefore shed light on the phenotypic evolution of the elephantids. Resurrecting ancient genes in order to explain the evolution of protein families like the steroid hormone receptor is an important field in molecular evolution buoyed by great advances in phylogenetics.[77] Moving from gene to genome resurrection could help elucidate not only how proteins evolve, but also how they co-evolve in a context of a protein interactome.

Not much is known on the role of segmental duplications throughout mammalian genomes, ultraconserved elements (UCEs) of 200 bp or longer that are found in identical sequences in more than one other species, or the human accelerated regions (HARs), conserved DNA segments highly mutated in humans compared to other species and with implications on the evolution of human-specific traits.[78] Furthermore, being able to build and study ancestral genomes derived by phylogenetic tree analysis of species will accelerate our understanding of evolution and selection. Building segments of the chimpanzee–human ancestral sequence and elucidating the biology in the context of early stem cell differentiation or organoids can have powerful implications on interpreting evolutionary branching and what accounts for species-specific differences. We can also generate genomes with combinatorial *cis* regulatory elements found in other species or totally novel sequences to understand gene regulatory networks that are important for evolution of the species from ancestral genomes, or future evolution of genomes with novel regulatory sequences. This has begun to be comprehensively addressed in prokaryotes where libraries of promoter sequences have been synthesized with all possible sequence combinations. These synthetic promoters then drive the expression of a fluorescent reporter, and all the sequence mutations are directly correlated to changes in protein translation.[79] Additionally, given that enhancers play an important role in eukaryotic gene expression, it would be interesting to learn whether shuffling genes and non-coding regions will still result in the same cell type or organism.

Another question that can be addressed by generating hybrid sequences of unicellular with simple multicellular organisms is the evolution of single-cell organisms to complex life forms. Work on unicellular and colonial choanoflagellates and their relationship with the bacterial life forms is one of the directions being explored as to how multicellularity arose.[80] However, hybrid genome synthesis and assembly could shed more light on the division of cellular functions. Taken together these research directions taken in parallel with significant ethical discussions and considerations, will usher in the transition from mostly molecular to testable organismic evolution.

Recoding genomes, or reducing the redundancy of the genetic code has been shown to render bacteria resistant to bacteriophages.[81] However, synthesizing genomes with non-redundant amino acid codes can also elucidate the importance of a redundant code. A recent attempt to recode arginine resulted in instances of recalcitrant mutations due to changes in mRNA structure and interference with the RBS site.[82] Furthermore, many organisms use 3 instances of stop codons but only one start codon. The ambush hypothesis assigns the need for the existence of more than one stop codon to the higher incidence of stop codons after frameshift mutations, which can save the cell energy and time by aborting translation of aberrant proteins.[83] Generating a complete bacterium with one triplet code (remove redundancy) per amino acid and evolving it over many generations can start to address the flexibility of the genetic code.

As we consolidate pathways for cell functions like metabolism, cell cycle, development etc., into distinct independent entities and synthetic gene circuits, and as we move towards evolving stable genomes in organisms with incorporated synthetic base pairs,[84] new protein machinery[35a] or different codon variations (i.e. 4 nucleotides rather than 3),[85] we can envision transitioning from learning the principles of putting together a genome to designing pipelines of building any genome of interest, whether modified or novel and with new functionalities.

5. The Use of DNA as Information Carriers for In Situ "Omics" Analysis

With the capability of engineering and generating more and more complex organisms, we need better tools to investigate and understand the biological processes underlying their functions. Both natural and synthetic biological systems typically consist of a large number of distinct molecular species that coordinate with each other to realize certain biological functions. By investigating only a portion of the system, limited information can be acquired, which may be inconclusive or misleading in some cases. Technological advances, especially in the fields of DNA sequencing and mass spectrometry, have enabled us to measure different molecular species from cells in an unbiased and comprehensive manner.[86] The resulted system-level "omics" data have greatly expanded our understanding of complex biological systems.[87] However, most currently available technologies can only perform in-bulk analysis.[88] Recently developed single-cell DNA/RNA sequencing can work for cells in suspension that are obtained naturally (e.g. blood cells), or from mechanical/enzymatic dissociation.[88,89] Given that complex biological systems are spatially well organized, the loss of positional information of individual cells makes it challenging to trace the data back to their original location in order to generate physiologically relevant interpretation. Therefore, developing technologies for spatially resolved omics or so-called in situ omics is of great importance and broad interest. Here, we summarize some past and ongoing work from our and other labs on using DNA for in situ omics analysis.

Beyond its role as the building blocks for genome, DNA can be viewed simply as a string of 4-letter code. It has enormously large design space with its diversity increasing exponentially (4^N) as a function of the length of DNA. This information-rich property of DNA has been utilized to archive digital data, such as books and movies.[90] DNA can also be repurposed to serve as molecular barcodes. We envision that if we can label each molecule species with a specific DNA barcode in situ, we can obtain the positional and quantitative information of each targeted molecule by simply reading out the corresponding DNA barcode (Figure 6). A library of orthogonal barcodes can be curated and used to label diverse types of molecules (DNA, RNA, protein, metabolites etc.) to generate the in situ omics.

A variety of methods can be used for decoding DNA in situ, and we herein will focus on two most widely used methods, fluorescence in situ sequencing (FISSEQ) and sequential hybridization with complementary fluorescent DNA probes (Figure 6). Both methods utilize fluorescence microscopy as the readout platform. FISSEQ, previously developed by our lab, is an in situ DNA sequencing method, wherein DNA sequences can be directly read out in biological samples such as tissue sections without DNA/RNA extraction.[91] Unbiased in situ RNA sequencing with FISSEQ has been demonstrated in fixed human fibroblasts.[91b] In a simulated injury model FISSEQ could successfully capture the gene expression profile change between fibroblasts close to and further from the wound sites.[91b] Nilsson et al. used FISSEQ to perform highly multiplexed targeted RNA detection.[92] In their study, 39 different mRNA species were simultaneously measured in fresh-frozen breast cancer tissue sections, and the result revealed heterogeneity in the localization pattern across the tissue and the number of signals for different mRNA transcripts. The second strategy for decoding DNA sequences in situ is to perform sequential hybridization with complementary fluorescent DNA probes. Compared to FISSEQ, the technical requirement for sequential hybridization is lower; however the multiplexing capability of this method also decreases. By reiterative probe hybridization and signal removal using photobleaching, Guo et al. demonstrated seven RNA targets imaging in single HeLa cells.[93] Combinatorial DNA barcoding strategy was introduced to enhance the multiplexing capability of sequential hybridization methods.[94] Zhuang and co-workers have developed multiplexed error-robust fluorescence in situ hybridization (MERFISH) to detect hundreds of different RNA species in situ.[94c] In MERFISH, each RNA species is labeled with

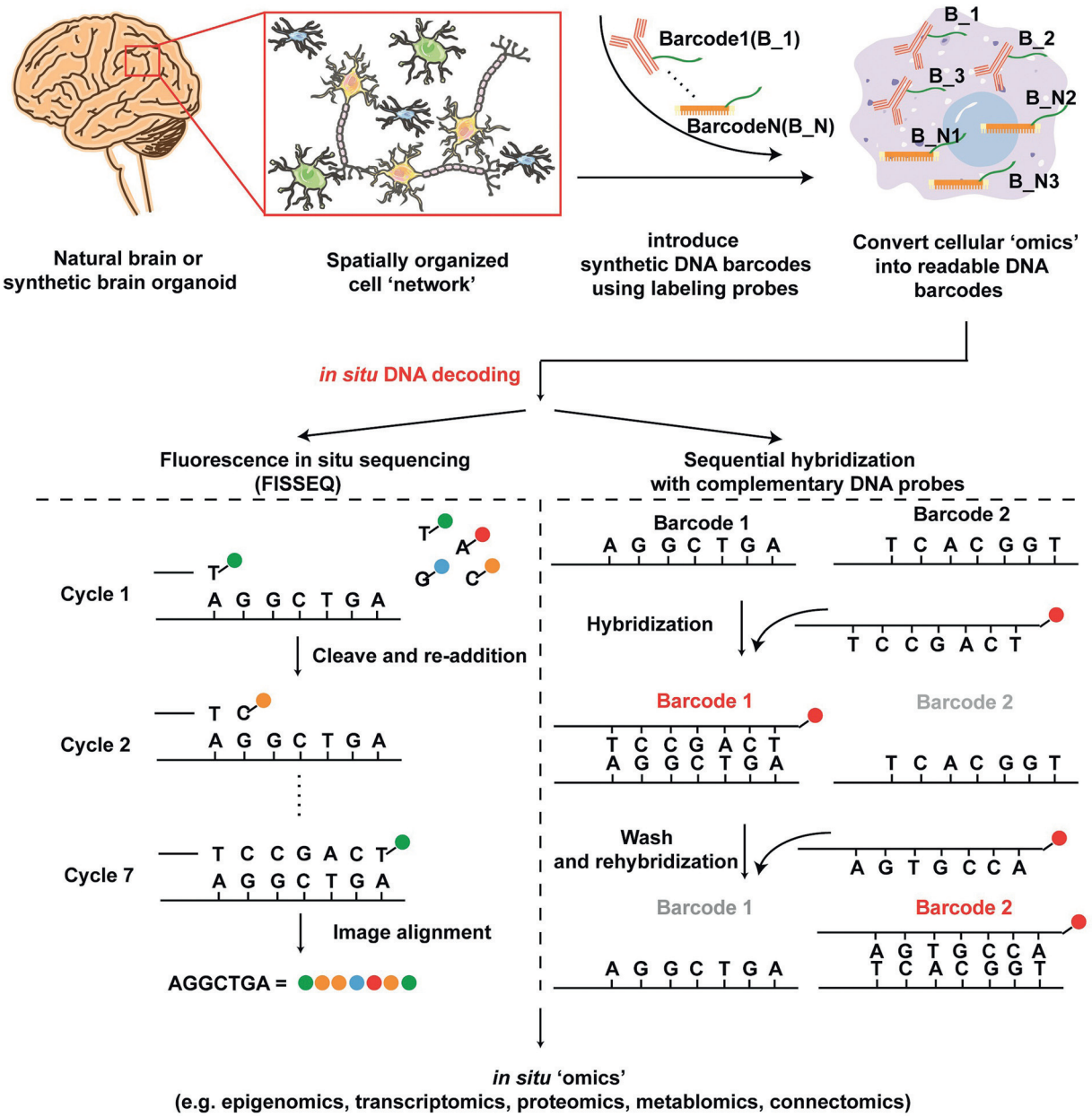


Figure 6. The use of DNA as information carrier for in situ omics analysis. Complex biological systems, such as natural brain or synthetic brain organoid, are built from a spatially well-organized network of different types of cells. Labeling probes, such as antibodies and fluorescence in situ hybridization (FISH) probes, that are conjugated with synthetic DNA barcodes are introduced to biological samples to convert their molecular information to readable DNA barcodes. It is followed by in situ DNA decoding using either fluorescence in situ sequencing (FISSEQ) or sequential hybridization with complementary DNA probes. In FISSEQ, each type of four nucleotides (A,T,C,G) is labeled with a spectrally distinct fluorophore. Iterative addition and cleavage of fluorophore-conjugated nucleotides is performed in conjunction with fluorescent image acquisition. The DNA sequence can be obtained by computational image alignment to register the fluorescent signal back to the position of corresponding DNA barcode. Sequential hybridization method uses iterative hybridization and removal of fluorophore-labeled complementary DNA probes. The DNA sequences are designed to be orthogonal to minimize crosstalk between different barcodes.

a combination of multiple (N) readout DNA sequences, resulting in a total of $2N-1$ barcodes. The actual barcode library is smaller than $2N-1$ for the purpose of error detection and correction. MERFISH has been applied to detect RNA expression profile in both fibroblast cell cultures and mouse brain samples.[94c,95] A key requirement for in situ detection is sufficient signal-to-noise (S/N) ratio. In FISSEQ, rolling circle amplification is used to generate hundreds of copies of the same signal in order to achieve a reasonable S/N ratio.[91] MERFISH relies on labeling the same RNA molecule with multiple FISH probes with the same DNA barcodes. As a result, it sets the limit for the minimal length of RNA can be detected using MERFISH (i.e. small RNAs like microRNA cannot be detected using this method). To increase S/N ratio, Cai et al. integrated a signal amplification method called hybridization chain reaction (HCR) with multiplexed

RNA FISH.[96] This allows the researchers to perform highly multiplexed RNA detection in thick tissue samples.[96e,f]

The same strategy can also be applied to DNA and protein to achieve highly multiplexed DNA or protein analysis in situ. Yin et al. have developed DNA-Exchange-Imaging, a highly multiplexed in situ protein detection method in which DNA barcode-conjugated antibodies are used to label different protein targets inside biological samples.[97] By simple buffer exchange to introduce and remove fluorophore-conjugated complementary DNA probes, the researchers have shown rapid multiplexed protein detection in both cell cultures and tissue samples. More interestingly, by tuning the binding affinity between fluorophore-conjugated DNA probes to the DNA barcodes on antibodies, the researchers can achieve highly multiplexed super-resolution imaging that is critical to resolve fine cellular structures (e.g. neuronal synapses).[97a,b,d] A related work employing a different sequential hybridization strategy (i.e. toehold displacement rather than buffer exchange in DNA-Exchange-Imaging) also proved that highly multiplexed protein detection could be readily accomplished using DNA-barcoded antibodies.[98] More recently, DNA barcodes have also been introduced into neuron projections and individual synapses, which can provide a potential strategy for high-throughput neuron projectome and connectome analysis.[99] We anticipate that using DNA barcodes for in situ omics can be integrated into an even broader number of systems to study questions that cannot be easily tackled with before.

A few challenges, however, still remain to be solved in order to achieve a more comprehensive in situ omics analysis. One is spatial crowding when a large number of molecules are labeled and imaged. A potential solution is to combine it with expansion microscopy, a technique that physically expands samples using swellable hydrogel.[100] Another challenge is lack of specific and high-affinity labeling probes to target proteins and metabolites. Commercially available antibodies are costly and unable to be engineered (e.g. to specific formulation for DNA conjugation). We have recently initiated a multi-institute effort to generate an open-access antibody database that deposits DNA sequences of antibodies obtained from in-house selection or hybridoma sequencing. We believe this open-access antibody database will greatly benefit the research community beyond the use for in situ omics analysis we present here.

6. Summary and Outlook

In this Review we have highlighted current progress on DNA-based technologies with an emphasis on DNA writing and editing while also describing future technologies that lie on the horizon. The recent progress in gene editing has been staggering and we anticipate a similar trend in the near future for DNA synthesis. The consequence of these advances will be far reaching and effect several fields of investigation. The availability of inexpensive, long stretches of DNA will immediately effect fields that use DNA for nanomaterials applications but will rapidly effect the engineering of biology at larger scales. Advances in DNA nanotechnology will allow for the design and creation of complex nanostructures that can target specific locations in the human body to deliver therapeutics or localize diagnostics. Microbial genome recoding will allow the engineering of organisms that have the potential to revolutionize the industrial production of chemicals and biologics such as those containing non-standard amino acids or D-amino acids. Whole genome writing and engineering will allow for the production of cell lines that are resistant to viral infection and cancer, the humanization of animal organs for transplantation, and the resurrection of extinct organisms. Lastly, the high-resolution, three-dimensional structure of biological systems will become uncovered through the highly multiplexed barcoding and imaging of healthy and diseased tissues and cells. It is possible that applications such as these are only the tip of the iceberg of what is to come from the exponential growth of current DNA-based technologies. From designing the molecules of life to designing life itself, the ever-advancing technologies that have enabled the reading and writing DNA with greater ease will continue to influence many aspects of the life sciences and ideally translate to an improvement in human health and quality of life.

Conflict of interest

The authors declare no conflict of interest.

How to cite: *Angew. Chem. Int. Ed.* **2018**, *57*, 4313 – 4328
Angew. Chem. **2018**, *130*, 4395 – 4411

[1] a) N. C. Seeman, *J. Theor. Biol.* **1982**, *99*, 237 – 247; b) N. C. Seeman, *Annu. Rev. Biochem.* **2010**, *79*, 65 – 87.
[2] Y. Ke, L. L. Ong, W. M. Shih, P. Yin, *Science* **2012**, *338*, 1177 – 1183.
[3] P. W. K. Rothemund, *Nature* **2006**, *440*, 297 – 302.
[4] F. Hong, F. Zhang, Y. Liu, H. Yan, *Chem. Rev.* **2017**, *117*, 12584 – 12640.
[5] a) H. Dietz, S. M. Douglas, W. M. Shih, *Science* **2009**, *325*, 725 – 730; b) D. Han, S. Pal, J. Nangreave, Z. Deng, Y. Liu, H. Yan, *Science* **2011**, *332*, 342 – 346.
[6] a) E. S. Andersen, M. Dong, M. M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M. M. Golas, B. Sander, H. Stark, C. L. P. Oliveira, J. S. Pedersen, V. Birkedal, F. Besenbacher, K. V. Gothelf, J. Kjems, *Nature* **2009**, *459*, 73 – 76; b) A. Kuzuya, M. Komiyama, *Chem. Commun.* **2009**, 4182 – 4184; c) Y. Ke, J. Sharma, M. Liu, K. Jahn, Y. Liu, H. Yan, *Nano Lett.* **2009**, *9*, 2445 – 2447.
[7] a) S. M. Douglas, H. Dietz, T. Liedl, B. Hogberg, F. Graf, W. M. Shih, *Nature* **2009**, *459*, 414 – 418; b) Y. Ke, S. M. Douglas, M. Liu, J. Sharma, A. Cheng, A. Leung, Y. Liu, W. M. Shih, H. Yan, *J. Am. Chem. Soc.* **2009**, *131*, 15903 – 15908; c) Y. Ke, N. V. Voigt, K. V. Gothelf, W. M. Shih, *J. Am. Chem. Soc.* **2012**, *134*, 1770 – 1774.
[8] a) D. Han, S. Pal, Y. Yang, S. Jiang, J. Nangreave, Y. Liu, H. Yan, *Science* **2013**, *339*, 1412 – 1415; b) E. Benson, A. Mohammed, J. Gardell, S. Masich, E. Czeizler, P. Orponen, B. Hogberg, *Nature* **2015**, *523*, 441 – 444; c) F. Zhang, S. Jiang, S. Wu, Y. Li, C. Mao, Y. Liu, H. Yan, *Nat. Nanotechnol.* **2015**, *10*, 779 – 784; d) R. Veneziano, S. Ratanalert, K. Zhang, F. Zhang, H. Yan, W. Chiu, M. Bathe, *Science* **2016**, *352*, 1534; e) M. Matthies, N. P. Agarwal, T. L. Schmidt, *Nano Lett.* **2016**, *16*, 2108 – 2113.

[9] a) S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, W. M. Shih, *Nucleic Acids Res.* **2009**, *37*, 5001 – 5006; b) D.-N. Kim, F. Kilchherr, H. Dietz, M. Bathe, *Nucleic Acids Res.* **2012**, *40*, 2862 – 2868.

[10] a) T. G. Martin, H. Dietz, *Nat. Commun.* **2012**, *3*, 1103; b) J.-P. J. Sobczak, T. G. Martin, T. Gerling, H. Dietz, *Science* **2012**, *338*, 1458 – 1461; c) Z. Zhang, J. Song, F. Besenbacher, M. Dong, K. V. Gothelf, *Angew. Chem. Int. Ed.* **2013**, *52*, 9219 – 9223; *Angew. Chem.* **2013**, *125*, 9389 – 9393.

[11] a) G. Bellot, M. A. McClintock, C. Lin, W. M. Shih, *Nat. Methods* **2011**, *8*, 192 – 194; b) C. Lin, S. D. Perrault, M. Kwak, F. Graf, W. M. Shih, *Nucleic Acids Res.* **2013**, *41*, e40; c) E. Stahl, T. G. Martin, F. Praetorius, H. Dietz, *Angew. Chem. Int. Ed.* **2014**, *53*, 12735 – 12740; *Angew. Chem.* **2014**, *126*, 12949 – 12954; d) A. Shaw, E. Benson, B. Högberg, *ACS Nano* **2015**, *9*, 4968 – 4975; e) C. Timm, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2015**, *54*, 6745 – 6750; *Angew. Chem.* **2015**, *127*, 6849 – 6854.

[12] C. E. Castro, F. Kilchherr, D.-N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe, H. Dietz, *Nat. Methods* **2011**, *8*, 221 – 229.

[13] D. Han, S. Pal, Y. Liu, H. Yan, *Nat. Nanotechnol.* **2010**, *5*, 712 – 717.

[14] R. E. Kohman, X. Han, *Chem. Commun.* **2015**, *51*, 5747 – 5750.

[15] a) Y.-X. Zhao, A. Shaw, X. Zeng, E. Benson, A. M. Nyström, B. Högberg, *ACS Nano* **2012**, *6*, 8684 – 8691; b) Q. Jiang, C. Song, J. Nangreave, X. Liu, L. Lin, D. Qiu, Z.-G. Wang, G. Zou, X. Liang, H. Yan, B. Ding, *J. Am. Chem. Soc.* **2012**, *134*, 13396 – 13403; c) Q. Zhang, Q. Jiang, N. Li, L. Dai, Q. Liu, L. Song, J. Wang, Y. Li, J. Tian, B. Ding, Y. Du, *ACS Nano* **2014**, *8*, 6633 – 6643; d) P. D. Halley, C. R. Lucas, E. M. McWilliams, M. J. Webber, R. A. Patton, C. Kural, D. M. Lucas, J. C. Byrd, C. E. Castro, *Small* **2016**, *12*, 308 – 320.

[16] R. E. Kohman, S. S. Cha, H.-Y. Man, X. Han, *Nano Lett.* **2016**, *16*, 2781 – 2785.

[17] S. M. Douglas, I. Bachelet, G. M. Church, *Science* **2012**, *335*, 831 – 834.

[18] S. D. Perrault, W. M. Shih, *ACS Nano* **2014**, *8*, 5132 – 5140.

[19] N. Ponnuswamy, M. M. C. Bastings, B. Nathwani, J. H. Ryu, L. Y. T. Chou, M. Vinther, W. A. Li, F. M. Anastassacos, D. J. Mooney, W. M. Shih, *Nat. Commun.* **2017**, *8*, 15654.

[20] S. Roy, M. Caruthers, *Molecules* **2013**, *18*, 14268.

[21] A. R. Chandrasekaran, M. Pushpanathan, K. Halvorsen, *Mater. Lett.* **2016**, *170*, 221 – 224.

[22] B. Kick, F. Praetorius, H. Dietz, D. Weuster-Botz, *Nano Lett.* **2015**, *15*, 4672 – 4676.

[23] a) C. Marianecci, F. Rinaldi, P. N. Hanieh, L. Di Marzio, D. Paolino, M. Carafa, *Drug Des. Dev. Ther.* **2017**, *11*, 325 – 335; b) B. S. Bleier, R. E. Kohman, R. E. Feldman, S. Ramanlal, X. Han, *PLoS One* **2013**, *8*, e61694; c) B. S. Bleier, R. E. Kohman, K. Guerra, A. L. Nocera, S. Ramanlal, A. H. Kocharyan, W. T. Curry, X. Han, *Neurosurgery* **2016**, *78*, 448 – 457.

[24] Y. Chen, L. Liu, *Adv. Drug Delivery Rev.* **2012**, *64*, 640 – 665.

[25] a) Y. Liu, R. Huang, L. Han, W. Ke, K. Shao, L. Ye, J. Lou, C. Jiang, *Biomaterials* **2009**, *30*, 4195 – 4202; b) Y. Liu, J. Li, K. Shao, R. Huang, L. Ye, J. Lou, C. Jiang, *Biomaterials* **2010**, *31*, 5246 – 5257; c) I. van Rooy, E. Mastrobattista, G. Storm, W. E. Hennink, R. M. Schiffelers, *J. Controlled Release* **2011**, *150*, 30 – 36; d) H. Xin, X. Sha, X. Jiang, L. Chen, K. Law, J. Gu, Y. Chen, X. Wang, X. Fang, *Biomaterials* **2012**, *33*, 1673 – 1681; e) Y. Qian, Y. Zha, B. Feng, Z. Pang, B. Zhang, X. Sun, J. Ren, C. Zhang, X. Shao, Q. Zhang, X. Jiang, *Biomaterials* **2013**, *34*, 2117 – 2129.

[26] a) J. K. Liu, Q. Teng, M. Garrity-Moses, T. Federici, D. Tanase, M. J. Imperiale, N. M. Boulis, *Neurobiol. Dis.* **2005**, *19*, 407 – 418; b) S. A. Townsend, G. D. Evrony, F. X. Gu, M. P. Schulz, R. H. Brown, Jr., R. Langer, *Biomaterials* **2007**, *28*, 5176 – 5184.

[27] M. Yang, K. Sunderland, C. Mao, *Chem. Rev.* **2017**, *117*, 10377 – 10402.

[28] N. Ponnuswamy, M. M. C. Bastings, B. Nathwani, J. H. Ryu, L. Y. T. Chou, M. Vinther, W. A. Li, F. M. Anastassacos, D. J. Mooney, W. M. Shih, *Nat. Commun.* **2017**, *8*, 15654.

[29] a) H. Yang, *Pharm. Res.* **2010**, *27*, 1759 – 1771; b) M. Demeule, J.-C. Currie, Y. Bertrand, C. Ché, T. Nguyen, A. Régina, R. Gabathuler, J.-P. Castaigne, R. Béliveau, *J. Neurochem.* **2008**, *106*, 1534 – 1544.

[30] a) H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest, G. M. Church, *Nature* **2009**, *460*, 894 – 898; b) F. J. Isaacs, P. A. Carr, H. H. Wang, M. J. Lajoie, B. Sterling, L. Kraal, A. C. Tolonen, T. A. Gianoulis, D. B. Goodman, N. B. Reppas, *Science* **2011**, *333*, 348 – 353; c) M. J. Lajoie, A. J. Rovner, D. B. Goodman, H.-R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church, F. J. Isaacs, *Science* **2013**, *342*, 357 – 360.

[31] a) D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison, H. O. Smith, *Science* **2008**, *319*, 1215 – 1220; b) D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z.-Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, H. O. Smith, J. C. Venter, *Science* **2010**, *329*, 52 – 56; c) N. Ostrov, M. Landon, M. Guell, G. Kuznetsov, J. Teramoto, N. Cervantes, M. Zhou, K. Singh, M. G. Napolitano, M. Moosburner, E. Shrock, B. W. Pruitt, N. Conway, D. B. Goodman, C. L. Gardner, G. Tyree, A. Gonzales, B. L. Wanner, J. E. Norville, M. J. Lajoie, G. M. Church, *Science* **2016**, *353*, 819 – 822.

[32] a) N. Annaluru, H. Muller, L. A. Mitchell, S. Ramalingam, G. Stracquadanio, S. M. Richardson, J. S. Dymond, Z. Kuang, L. Z. Scheifele, E. M. Cooper, Y. Cai, K. Zeller, N. Agmon, J. S. Han, M. Hadjithomas, J. Tullman, K. Caravelli, K. Cirelli, Z. Guo, V. London, A. Yeluru, S. Murugan, K. Kandavelou, N. Agier, G. Fischer, K. Yang, J. A. Martin, M. Bilgel, P. Bohutskyi, K. M. Boulier, B. J. Capaldo, J. Chang, K. Charoen, W. J. Choi, P. Deng, J. E. DiCarlo, J. Doong, J. Dunn, J. I. Feinberg, C. Fernandez, C. E. Floria, D. Gladowski, P. Hadidi, I. Ishizuka, J. Jabbari, C. Y. L. Lau, P. A. Lee, S. Li, D. Lin, M. E. Linder, J. Ling, J. J. Liu, J. J. Liu, M. London, H. Ma, J. Mao, J. E. McDade, A. McMillan, A. M. Moore, W. C. Oh, Y. Ouyang, R. Patel, M. Paul, L. C. Paulsen, J. Qiu, A. Rhee, M. G. Rubashkin, I. Y. Soh, N. E. Sotuyo, V. Srinivas, A. Suarez, A. Wong, R. Wong, W. R. Xie, Y. Xu, A. T. Yu, R. Koszul, J. S. Bader, J. D. Boeke, S. Chandrasegaran, *Science* **2014**, *344*, 55 – 58; b) S. M. Richardson, L. A. Mitchell, G. Stracquadanio, K. Yang, J. S. Dymond, J. E. DiCarlo, D. Lee, C. L. V. Huang, S. Chandrasegaran, Y. Cai, J. D. Boeke, J. S. Bader, *Science* **2017**, *355*, 104 – 1044; c) L. A. Mitchell, A. Wang, G. Stracquadanio, Z. Kuang, X. Wang, K. Yang, S. Richardson, J. A. Martin, Y. Zhao, R. Walker, Y. Luo, H. Dai, K. Dong, Z. Tang, Y. Yang, Y. Cai, A. Heguy, B. Ueberheide, D. Fenyö, J. Dai, J. S. Bader, J. D. Boeke, *Science* **2017**, *355*, eaaf4831; d) G. Mercy, J. Mozziconacci, V. F. Scolari, K. Yang, G. Zhao, A. Thierry, Y. Luo, L. A. Mitchell, M. Shen, Y. Shen, R. Walker, W. Zhang, Y. Wu, Z.-X. Xie, Z. Luo, Y. Cai, J. Dai, H. Yang, Y.-J. Yuan, J. D. Boeke, J. S. Bader, H. Muller, R. Koszul, *Science* **2017**, *355*, eaaf4597; e) Y. Shen, Y. Wang, T. Chen, F. Gao, J. Gong, D. Abramczyk, R. Walker, H. Zhao, S. Chen, W. Liu, Y. Luo, C. A. Müller, A. Paul-Dubois-Taine, B. Alver, G. Stracquadanio, L. A. Mitchell, Z. Luo, Y. Fan, B. Zhou, B. Wen, F. Tan, Y. Wang, J. Zi, Z. Xie, B. Li, K. Yang, S. M. Richardson, H. Jiang, C. E. French, C. A. Nieduszynski, R. Koszul, A. L. Marston, Y. Yuan, J. Wang, J. S.

Bader, J. Dai, J. D. Boeke, X. Xu, Y. Cai, H. Yang, *Science* **2017**, *355*, eaaf4791; f) Y. Wu, B.-Z. Li, M. Zhao, L. A. Mitchell, Z.-X. Xie, Q.-H. Lin, X. Wang, W.-H. Xiao, Y. Wang, X. Zhou, H. Liu, X. Li, M.-Z. Ding, D. Liu, L. Zhang, B.-L. Liu, X.-L. Wu, F.-F. Li, X.-T. Dong, B. Jia, W.-Z. Zhang, G.-Z. Jiang, Y. Liu, X. Bai, T.-Q. Song, Y. Chen, S.-J. Zhou, R.-Y. Zhu, F. Gao, Z. Kuang, X. Wang, M. Shen, K. Yang, G. Stracquadanio, S. M. Richardson, Y. Lin, L. Wang, R. Walker, Y. Luo, P.-S. Ma, H. Yang, Y. Cai, J. Dai, J. S. Bader, J. D. Boeke, Y.-J. Yuan, *Science* **2017**, *355*, eaaf4706; g) Z.-X. Xie, B.-Z. Li, L. A. Mitchell, Y. Wu, X. Qi, Z. Jin, B. Jia, X. Wang, B.-X. Zeng, H.-M. Liu, X.-L. Wu, Q. Feng, W.-Z. Zhang, W. Liu, M.-Z. Ding, X. Li, G.-R. Zhao, J.-J. Qiao, J.-S. Cheng, M. Zhao, Z. Kuang, X. Wang, J. A. Martin, G. Stracquadanio, K. Yang, X. Bai, J. Zhao, M.-L. Hu, Q.-H. Lin, W.-Q. Zhang, M.-H. Shen, S. Chen, W. Su, E.-X. Wang, R. Guo, F. Zhai, X.-J. Guo, H.-X. Du, J.-Q. Zhu, T.-Q. Song, J.-J. Dai, F.-F. Li, G.-Z. Jiang, S.-L. Han, S.-Y. Liu, Z.-C. Yu, X.-N. Yang, K. Chen, C. Hu, D.-S. Li, N. Jia, Y. Liu, L.-T. Wang, S. Wang, X.-T. Wei, M.-Q. Fu, L.-M. Qu, S.-Y. Xin, T. Liu, K.-R. Tian, X.-N. Li, J.-H. Zhang, L.-X. Song, J.-G. Liu, J.-F. Lv, H. Xu, R. Tao, Y. Wang, T.-T. Zhang, Y.-X. Deng, Y.-R. Wang, T. Li, G.-X. Ye, X.-R. Xu, Z.-B. Xia, W. Zhang, S.-L. Yang, Y.-L. Liu, W.-Q. Ding, Z.-N. Liu, J.-Q. Zhu, N.-Z. Liu, R. Walker, Y. Luo, Y. Wang, Y. Shen, H. Yang, Y. Cai, P.-S. Ma, C.-T. Zhang, J. S. Bader, J. D. Boeke, Y.-J. Yuan, *Science* **2017**, *355*, eaaf4704; h) W. Zhang, G. Zhao, Z. Luo, Y. Lin, L. Wang, Y. Guo, A. Wang, S. Jiang, Q. Jiang, J. Gong, Y. Wang, S. Hou, J. Huang, T. Li, Y. Qin, J. Dong, Q. Qin, J. Zhang, X. Zou, X. He, L. Zhao, Y. Xiao, M. Xu, E. Cheng, N. Huang, T. Zhou, Y. Shen, R. Walker, Y. Luo, Z. Kuang, L. A. Mitchell, K. Yang, S. M. Richardson, Y. Wu, B.-Z. Li, Y.-J. Yuan, H. Yang, J. Lin, G.-Q. Chen, Q. Wu, J. S. Bader, Y. Cai, J. D. Boeke, J. Dai, *Science* **2017**, *355*, eaaf3981.

[33] N. J. Ma, F. J. Isaacs, *Cell Syst.* **2016**, *3*, 199 – 207.

[34] V. Bethencourt, *Nat. Biotechnol.* **2009**, *27*, 681.

[35] a) D. J. Mandell, M. J. Lajoie, M. T. Mee, R. Takeuchi, G. Kuznetsov, J. E. Norville, C. J. Gregg, B. L. Stoddard, G. M. Church, *Nature* **2015**, *518*, 55 – 60; b) A. J. Rovner, A. D. Haimovich, S. R. Katz, Z. Li, M. W. Grome, B. M. Gassaway, M. Amiram, J. R. Patel, R. R. Gallagher, J. Rinehart, F. J. Isaacs, *Nature* **2015**, *518*, 89 – 93.

[36] G. Kuznetsov, D. B. Goodman, G. T. Filsinger, M. Landon, N. Rohland, J. Aach, M. J. Lajoie, G. M. Church, *Genome Biol.* **2017**, *18*, 100.

[37] T. M. Wannier, A. M. Kunjapur, D. P. Rice, M. J. McDonald, M. M. Desai, G. M. Church, *Proc. Natl. Acad. Sci. USA* **2018**, https://doi.org/10.1073/pnas.1715530115.

[38] a) L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498 – 500; b) J. W. Chin, *Annu. Rev. Biochem.* **2014**, *83*, 379 – 408; c) A. Dumas, L. Lercher, C. D. Spicer, B. G. Davis, *Chem. Sci.* **2015**, *6*, 50 – 69.

[39] http://ambrx.com/.

[40] http://harvardlifelab.com/ventures/team/gro-biosciences/.

[41] http://www.sutrobio.com/.

[42] http://synthorx.com/.

[43] E. S. Zimmerman, T. H. Heibeck, A. Gill, X. Li, C. J. Murray, M. R. Madlansacay, C. Tran, N. T. Uter, G. Yin, P. J. Rivers, A. Y. Yam, W. D. Wang, A. R. Steiner, S. U. Bajad, K. Penta, W. Yang, T. J. Hallam, C. D. Thanos, A. K. Sato, *Bioconjugate Chem.* **2014**, *25*, 351 – 361.

[44] S. A. Kularatne, V. Deshmukh, M. Gymnopoulos, S. L. Biroc, J. Xia, S. Srinagesh, Y. Sun, N. Zou, M. Shimazu, J. Pinkstaff, S. Ensari, N. Knudsen, A. Manibusan, J. Y. Axup, C. H. Kim, V. V. Smider, T. Javahishvili, P. G. Schultz, *Angew. Chem. Int. Ed.* **2013**, *52*, 12101 – 12104; *Angew. Chem.* **2013**, *125*, 12323 – 12326.

[45] a) C. C. Liu, P. G. Schultz, *Nat. Biotechnol.* **2006**, *24*, 1436 – 1440; b) C. C. Liu, H. Choe, M. Farzan, V. V. Smider, P. G. Schultz, *Biochemistry* **2009**, *48*, 8891 – 8898; c) H. Cho, T. Daniel, Y. J. Buechler, D. C. Litzinger, Z. Maio, A.-M. H. Putnam, V. S. Kraynov, B.-C. Sim, S. Bussell, T. Javahishvili, S. Kaphle, G. Viramontes, M. Ong, S. Chu, B. Gc, R. Lieu, N. Knudsen, P. Castiglioni, T. C. Norman, D. W. Axelrod, A. R. Hoffman, P. G. Schultz, R. D. DiMarchi, B. E. Kimmel, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9060 – 9065.

[46] a) J. C. Jackson, S. P. Duffy, K. R. Hess, R. A. Mehl, *J. Am. Chem. Soc.* **2006**, *128*, 11124 – 11127; b) I. N. Ugwumba, K. Ozawa, Z.-Q. Xu, F. Ely, J.-L. Foo, A. J. Herlt, C. Coppin, S. Brown, M. C. Taylor, D. L. Ollis, L. N. Mander, G. Schenk, N. E. Dixon, G. Otting, J. G. Oakeshott, C. J. Jackson, *J. Am. Chem. Soc.* **2011**, *133*, 326 – 333; c) C. L. Windle, K. J. Simmons, J. R. Ault, C. H. Trinh, A. Nelson, A. R. Pearson, A. Berry, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2610 – 2615.

[47] D. A. Malyshev, K. Dhami, T. Lavergne, T. Chen, N. Dai, J. M. Foster, I. R. Corrêa, F. E. Romesberg, *Nature* **2014**, *509*, 385 – 388.

[48] a) K. Oki, K. Sakamoto, T. Kobayashi, H. M. Sasaki, S. Yokoyama, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13298 – 13303; b) T. S. Young, I. Ahmad, J. A. Yin, P. G. Schultz, *J. Mol. Biol.* **2010**, *395*, 361 – 374; c) A. K. Antonczak, Z. Simova, I. T. Yonemoto, M. Bochtler, A. Piasecka, H. Czapinska, A. Brancale, E. M. Tippmann, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1320 – 1325; d) S. Nehring, N. Budisa, B. Wiltschi, M. Oliveberg, N. Budisa, *PLoS One* **2012**, *7*, e31992; e) J. W. Monk, S. P. Leonard, C. W. Brown, M. J. Hammerling, C. Mortensen, A. E. Gutierrez, N. Y. Shin, E. Watkins, D. M. Mishler, J. E. Barrick, *ACS Synth. Biol.* **2017**, *6*, 45 – 54.

[49] R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, A. David, M. Horn, P. G. Schultz, *J. Am. Chem. Soc.* **2003**, *125*, 935 – 939.

[50] A. M. Kunjapur, D. A. Stork, E. Kuru, O. Vargas-Rodriguez, M. Landon, D. Söll, G. M. Church, *Proc. Natl. Acad. Sci. USA*, **2018**, *115*, 619 – 624.

[51] a) A. Bachmair, D. Finley, A. Varshavsky, *Science* **1986**, *234*, 179 – 186; b) J. W. Tobias, T. E. Shrader, G. Rocap, A. Varshavsky, *Science* **1991**, *254*, 1374 – 1377; c) T. Tasaki, S. M. Sriram, K. S. Park, Y. T. Kwon, *Annu. Rev. Biochem.* **2012**, *81*, 261 – 289.

[52] a) F. Baneyx, G. Georgiou, *J. Bacteriol.* **1990**, *172*, 491 – 494; b) L. SaiSree, M. Reddy, J. Gowrishankar, *J. Bacteriol.* **2001**, *183*, 6943 – 6946.

[53] a) R. Zilhão, F. Cairrão, P. Régnier, C. M. Arraiano, *Mol. Microbiol.* **1996**, *20*, 1033 – 1042; b) P. J. Lopez, I. Marchand, S. A. Joyce, M. Dreyfus, *Mol. Microbiol.* **1999**, *33*, 188 – 199.

[54] a) A. M. Kunjapur, Y. Tarasova, K. L. J. Prather, *J. Am. Chem. Soc.* **2014**, *136*, 11644 – 11654; b) G. M. Rodriguez, S. Atsumi, *Metab. Eng.* **2014**, *25*, 227 – 237.

[55] A. M. Kunjapur, J. C. Hyun, K. L. J. Prather, *Microb. Cell Fact.* **2016**, *15*, 1.

[56] K. Temme, D. Zhao, C. A. Voigt, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7085 – 7090.

[57] G. Church, E. Regis, *Regenesis: How Synthetic Biology Will Reinvent Nature and Ourselves*, Basic Books, **2014**.

[58] J. Z. Zhang, *Trends Ecol. Evol.* **2003**, *18*, 292 – 298.

[59] C. Ollivaux, D. Soyez, J. Y. Toullec, *J. Pept. Sci.* **2014**, *20*, 595 – 612.

[60] R. Tugyi, K. Uray, D. Iván, E. Fellinger, A. Perkins, F. Hudecz, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 413 – 418.

[61] E. Kuru, H. V. Hughes, P. J. Brown, E. Hall, S. Tekkam, F. Cava, M. A. de Pedro, Y. V. Brun, M. S. VanNieuwenhze, *Angew. Chem. Int. Ed.* **2012**, *51*, 12519 – 12523; *Angew. Chem.* **2012**, *124*, 12687 – 12691.

[62] L. M. Dedkova, N. E. Fahmi, S. Y. Golovine, S. M. Hecht, *J. Am. Chem. Soc.* **2003**, *125*, 6616 – 6617.

[63] T. Fujino, Y. Goto, H. Suga, H. Murakami, *J. Am. Chem. Soc.* **2013**, *135*, 1830 – 1837.

[64] J. Achenbach, M. Jahnz, L. Bethge, K. Paal, M. Jung, M. Schuster, R. Albrecht, F. Jarosch, K. H. Nierhaus, S. Klussmann, *Nucleic Acids Res.* **2015**, *43*, 5687 – 5698.
[65] S. Ude, J. Lassak, A. L. Starosta, T. Kraxenberger, D. N. Wilson, K. Jung, *Science* **2013**, *339*, 82 – 85.
[66] P.-Y. Huang, F. Wang, K. Narasimhan, K. Chatman, J. Aach, S. A. Trauger, R. Spoering, G. M. Church, *bioRxiv* **2017**, https://doi.org/10.1101/125930.
[67] T. Katoh, K. Tajima, H. Suga, *Cell Chem. Biol.* **2017**, *24*, 46 – 54.
[68] Z. Wang, W. Xu, L. Liu, T. F. Zhu, *Nat. Chem.* **2016**, *8*, 698 – 704.
[69] a) D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison III, H. O. Smith, *Science* **2008**, *319*, 1215 – 1220; b) C. A. Hutchison III, R. Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z. Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson, J. C. Venter, *Science* **2016**, *351*, aad6253; c) L. A. Mitchell, A. Wang, G. Stracquadanio, Z. Kuang, X. Wang, K. Yang, S. Richardson, J. A. Martin, Y. Zhao, R. Walker, Y. Luo, H. Dai, K. Dong, Z. Tang, Y. Yang, Y. Cai, A. Heguy, B. Ueberheide, D. Fenyo, J. Dai, J. S. Bader, J. D. Boeke, *Science* **2017**, *355*, eaaf4831; d) S. M. Richardson, L. A. Mitchell, G. Stracquadanio, K. Yang, J. S. Dymond, J. E. DiCarlo, D. Lee, C. L. Huang, S. Chandrasegaran, Y. Cai, J. D. Boeke, J. S. Bader, *Science* **2017**, *355*, 1040 – 1044.
[70] F. J. Isaacs, P. A. Carr, H. H. Wang, M. J. Lajoie, B. Sterling, L. Kraal, A. C. Tolonen, T. A. Gianoulis, D. B. Goodman, N. B. Reppas, C. J. Emig, D. Bang, S. J. Hwang, M. C. Jewett, J. M. Jacobson, G. M. Church, *Science* **2011**, *333*, 348 – 353.
[71] G. Pines, E. F. Freed, J. D. Winkler, R. T. Gill, *ACS Synth. Biol.* **2015**, *4*, 1176 – 1185.
[72] K. Wang, J. Fredens, S. F. Brunner, S. H. Kim, T. Chia, J. W. Chin, *Nature* **2016**, *539*, 59 – 64.
[73] T. Suzuki, Y. Kazuki, M. Oshimura, T. Hara, *PLoS One* **2016**, *11*, e0157187.
[74] J. D. Boeke, G. Church, A. Hessel, N. J. Kelley, A. Arkin, Y. Cai, R. Carlson, A. Chakravarti, V. W. Cornish, L. Holt, F. J. Isaacs, T. Kuiken, M. Lajoie, T. Lessor, J. Lunshof, M. T. Maurano, L. A. Mitchell, J. Rine, S. Rosser, N. E. Sanjana, P. A. Silver, D. Valle, H. Wang, J. C. Way, L. Yang, *Science* **2016**, *353*, 126 – 127.
[75] L. Yang, M. Guell, D. Niu, H. George, E. Lesha, D. Grishin, J. Aach, E. Shrock, W. Xu, J. Poci, R. Cortazio, R. A. Wilkinson, J. A. Fishman, G. Church, *Science* **2015**, *350*, 1101 – 1104.
[76] a) M. Su, D. Han, J. Boyd-Kirkup, X. Yu, J. D. Han, *Cell Rep.* **2014**, *7*, 376 – 385; b) J. N. Athanikar, R. M. Badge, J. V. Moran, *Nucleic Acids Res.* **2004**, *32*, 3846 – 3855.
[77] a) J. W. Thornton, *Nat. Rev. Genet.* **2004**, *5*, 366 – 375; b) J. W. Thornton, E. Need, D. Crews, *Science* **2003**, *301*, 1714 – 1717.
[78] a) E. S. Lander, *Nature* **2011**, *470*, 187 – 197; b) G. Bejerano, M. Pheasant, I. Makunin, S. Stephen, W. J. Kent, J. S. Mattick, D. Haussler, *Science* **2004**, *304*, 1321 – 1325; c) M. J. Hubisz, K. S. Pollard, *Curr. Opin. Genet. Dev.* **2014**, *29*, 15 – 21.
[79] S. Kosuri, D. B. Goodman, G. Cambray, V. K. Mutalik, Y. Gao, A. P. Arkin, D. Endy, G. M. Church, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14024 – 14029.
[80] J. Carpenter, *Science* **2012**, *337*, 510.
[81] M. J. Lajoie, A. J. Rovner, D. B. Goodman, H. R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church, F. J. Isaacs, *Science* **2013**, *342*, 357 – 360.
[82] M. G. Napolitano, M. Landon, C. J. Gregg, M. J. Lajoie, L. Govindarajan, J. A. Mosberg, G. Kuznetsov, D. B. Goodman, O. Vargas-Rodriguez, F. J. Isaacs, D. Soll, G. M. Church, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5588 – E5597.
[83] T. R. Singh, K. R. Pardasani, *Comput. Biol. Chem.* **2009**, *33*, 239 – 244.
[84] D. A. Malyshev, K. Dhami, T. Lavergne, T. Chen, N. Dai, J. M. Foster, I. R. Correa, Jr., F. E. Romesberg, *Nature* **2014**, *509*, 385 – 388.
[85] J. W. Chin, *Nature* **2017**, *550*, 53.
[86] R. P. Horgan, L. C. Kenny, *Obstet. Gynaecol.* **2011**, *13*, 189 – 195.
[87] H. Ge, A. Walhout, M. Vidal, *Trends Genet.* **2003**, *19*, 551 – 560.
[88] N. Crosetto, M. Bienko, A. van Oudenaarden, *Nat. Rev. Genet.* **2015**, *16*, 57 – 66.
[89] Y. Wang, N. E. Navin, *Mol. Cell* **2015**, *58*, 598 – 609.
[90] a) G. M. Church, Y. Gao, S. Kosuri, *Science* **2012**, *337*, 1628; b) N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E. M. LeProust, B. Sipos, E. Birney, *Nature* **2013**, *494*, 77 – 80; c) R. N. Grass, R. Heckel, M. Puddu, D. Paunescu, W. J. Stark, *Angew. Chem. Int. Ed.* **2015**, *54*, 2552 – 2555; *Angew. Chem.* **2015**, *127*, 2582 – 2586; d) Y. Erlich, D. Zielinski, *Science* **2017**, *355*, 950 – 954; e) S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Nature* **2017**, https://doi.org/10.1038/nature23017.
[91] a) J. Shendure, G. J. Porreca, N. B. Reppas, X. Lin, J. P. McCutcheon, A. M. Rosenbaum, M. D. Wang, K. Zhang, R. D. Mitra, G. M. Church, *Science* **2005**, *309*, 1728 – 1732; b) J. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, J. L. Yang, T. C. Ferrante, R. Terry, S. S. F. Jeanty, C. Li, R. Amamoto, D. T. Peters, B. M. Turczyk, A. H. Marblestone, S. A. Inverso, A. Bernard, P. Mali, X. Rios, J. Aach, G. M. Church, *Science* **2014**, *343*, 1360 – 1363; c) J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, G. M. Church, *Nat. Protoc.* **2015**, *10*, 442 – 458.
[92] R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, M. Nilsson, *Nat. Methods* **2013**, *10*, 857 – 860.
[93] L. Xiao, J. Guo, *Anal. Methods* **2015**, *7*, 7290 – 7295.
[94] a) E. Lubeck, L. Cai, *Nat. Methods* **2012**, *9*, 743 – 748; b) E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, L. Cai, *Nat. Methods* **2014**, *11*, 360 – 361; c) K. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **2015**, *348*, aaa6090; d) A. F. Coskun, L. Cai, *Nat. Methods* **2016**, *13*, 657 – 660.
[95] a) J. R. Moffitt, J. Hao, D. Bambah-Mukku, T. Lu, C. Dulac, X. Zhuang, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 14456 – 14461; b) J. R. Moffitt, J. Hao, G. Wang, K. H. Chen, H. P. Babcock, X. Zhuang, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 11046 – 11051.
[96] a) R. M. Dirks, N. A. Pierce, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15275 – 15278; b) H. M. Choi, J. Y. Chang, L. A. Trinh, J. E. Padilla, S. E. Fraser, N. A. Pierce, *Nat. Biotechnol.* **2010**, *28*, 1208 – 1212; c) H. M. Choi, V. A. Beck, N. A. Pierce, *ACS Nano* **2014**, *8*, 4284 – 4294; d) H. M. Choi, C. R. Calvert, N. Husain, D. Huss, J. C. Barsi, B. E. Deverman, R. C. Hunter, M. Kato, S. M. Lee, A. C. Abelin, A. Z. Rosenthal, O. S. Akbari, Y. Li, B. A. Hay, P. W. Sternberg, P. H. Patterson, E. H. Davidson, S. K. Mazmanian, D. A. Prober, M. van de Rijn, J. R. Leadbetter, D. K. Newman, C. Readhead, M. E. Bronner, B. Wold, R. Lansford, T. Sauka-Spengler, S. E. Fraser, N. A. Pierce, *Development* **2016**, *143*, 3632 – 3637; e) S. Shah, E. Lubeck, M. Schwarzkopf, T.-F. F. He, A. Greenbaum, C. H. Sohn, A. Lignell, H. M. Choi, V. Gradinaru, N. A. Pierce, L. Cai, *Development* **2016**, *143*, 2862 – 2867; f) S. Shah, E. Lubeck, W. Zhou, L. Cai, *Neuron* **2016**, *92*, 342 – 357.
[97] a) R. Jungmann, M. S. Avendano, J. B. Woehrstein, M. Dai, W. M. Shih, P. Yin, *Nat. Methods* **2014**, *11*, 313 – 318; b) S. S. Agasti, Y. Wang, F. Schueder, A. Sukumar, R. Jungmann, P. Yin, *Chem. Sci.* **2017**, *8*, 3080 – 3091; c) F. Schueder, M. T. Strauss, D. Hoerl, J. Schnitzbauer, T. Schlichthaerle, S. Strauss, P. Yin, H. Harz, H. Leonhardt, R. Jungmann, *Angew. Chem. Int.*

Ed. **2017**, *56*, 4052 – 4055; *Angew. Chem.* **2017**, *129*, 4111 – 4114; d) Y. Wang, J. B. Woehrstein, N. Donoghue, M. Dai, M. S. Avendano, R. C. J. Schackmann, J. Zoeller, S. S. H. Wang, P. W. Tillberg, D. Park, S. W. Lapan, E. S. Boyden, J. S. Brugge, P. S. Kaeser, G. Church, S. S. Agasti, R. Jungmann, P. Yin, *Nano Lett.* **2017**, https://doi.org/10.1021/acs.nanolett.7b02716.

[98] R. M. Schweller, J. Zimak, D. Y. Duose, A. A. Qutub, W. N. Hittelman, M. R. Diehl, *Angew. Chem. Int. Ed.* **2012**, *51*, 9292 – 9296; *Angew. Chem.* **2012**, *124*, 9426 – 9430.

[99] a) A. H. Marblestone, E. R. Daugharthy, R. Kalhor, I. D. Peikon, J. M. Kebschull, S. L. Shipman, Y. Mishchenko, J. H. Lee, K. P. Kording, E. S. Boyden, *arXiv*, 1404.5103, **2014**; b) J. M. Kebschull, P. G. da Silva, A. P. Reid, I. D. Peikon, D. F. Albeanu, A. M. Zador, *Neuron* **2016**, *91*, 975 – 987; c) I. D. Peikon, J. M. Kebschull, V. V. Vagin, D. I. Ravens, Y.-C. C. Sun, E. Brouzes, I. R. Corrêa, D. Bressan, A. M. Zador, *Nucleic Acids Res.* **2017**, https://doi.org/10.1093/nar/gkx292.

[100] a) Y. Zhao, O. Bucur, H. Irshad, F. Chen, A. Weins, A. L. Stancu, E. Y. Oh, M. DiStasio, V. Torous, B. Glass, I. E. Stillman, S. J. Schnitt, A. H. Beck, E. S. Boyden, *Nat. Biotechnol.* **2017**, https://doi.org/10.1038/nbt.3892; b) J. B. Chang, F. Chen, Y. G. Yoon, E. E. Jung, H. Babcock, J. S. Kang, S. Asano, H. J. Suk, N. Pak, P. W. Tillberg, A. T. Wassie, D. Cai, E. S. Boyden, *Nat. Methods* **2017**, *14*, 593 – 599; c) P. W. Tillberg, F. Chen, K. D. Piatkevich, Y. Zhao, C. C. Yu, B. P. English, L. Gao, A. Martorell, H. J. Suk, F. Yoshida, E. M. DeGennaro, D. H. Roossien, G. Gong, U. Seneviratne, S. R. Tannenbaum, R. Desimone, D. Cai, E. S. Boyden, *Nat. Biotechnol.* **2016**, *34*, 987 – 992; d) F. Chen, A. T. Wassie, A. J. Cote, A. Sinha, S. Alon, S. Asano, E. R. Daugharthy, J. B. Chang, A. Marblestone, G. M. Church, A. Raj, E. S. Boyden, *Nat. Methods* **2016**, *13*, 679 – 684; e) F. Chen, P. W. Tillberg, E. S. Boyden, *Science* **2015**, *347*, 543 – 548.

Manuscript received: August 4, 2017
Revised manuscript received: November 14, 2017
Accepted manuscript online: January 9, 2018
Version of record online: February 23, 2018

