Reduction-to-synthesis: the dominant approach to genome-scale synthetic biology

Kim, Kangsan; Choe, Donghui; Cho, Suhyung; Palsson, Bernhard; Cho, Byung Kwan

Published in:
Trends in Biotechnology

Link to article, DOI:
10.1016/j.tibtech.2024.02.008

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

CellPress
OPEN ACCESS

Trends in
Biotechnology

Review

Reduction-to-synthesis: the dominant approach to genome-scale synthetic biology

Kangsan Kim [1,2,7], Donghui Choe [3,7], Suhyung Cho[2], Bernhard Palsson [3,4], and Byung-Kwan Cho [1,2,5,6,*]

Advances in systems and synthetic biology have propelled the construction of reduced bacterial genomes. Genome reduction was initially focused on exploring properties of minimal genomes, but more recently it has been deployed as an engineering strategy to enhance strain performance. This review provides the latest updates on reduced genomes, focusing on dual-track approaches of top-down reduction and bottom-up synthesis for their construction. Using cases from studies that are based on established industrial workhorse strains, we discuss the construction of a series of synthetic phenotypes that are candidates for biotechnological applications. Finally, we address the possible uses of reduced genomes for biotechnological applications and the needed future research directions that may ultimately lead to the total synthesis of rationally designed genomes.

Highlights

Top-down and bottom-up genome streamlining (or synthesis) are the two distinct modalities for synthetic genome construction.

Adaptive laboratory evolution (ALE) plays multiple roles, including debugging system abnormalities, exploring emergent properties of minimal genomes, and providing design principles for genome-scale engineering.

The genome 'reductionist' approach is an engineering method that markedly improves strain performance.

Advances in genome synthesis technology have opened up new possibilities for the chemical synthesis of customized designer genomes with specialized functions.

The synthetic minimal genome JCVI-syn3.0 and its derivatives serve as conducive platforms for studying fundamental aspects of genome composition and organization.

The concept of minimal genomes

The quest for minimalistic organisms containing only the bare minimum genetic material essential for survival has remained an active research avenue since its inception in the 20th century [1]. The goal of generating a genome smaller than the smallest extant *Mycoplasma genitalium* has been addressed through the chemical synthesis of self-replicating JCVI-syn3.0 in the past decade [2]. Since then, the simplest genome with minimal genome complexity and biological unknowns has been endowed with 'platform chassis' status to investigate fundamental aspects of biology. Undoubtedly, the endeavor towards **minimal genome** (see Glossary) synthesis has enabled exciting advances in the fields of **systems biology** and **synthetic biology**. Among the numerous important milestones is the convergence of omics-driven systems inspection of genome properties and DNA technologies that build and rewrite the genome, establishing the design–build–test–learn (DBTL) concept, which is arguably the most common strain engineering approach in biological engineering [3].

Earlier genome minimization efforts focused on stripping down as many genes as possible to characterize the effects and consequences of genome reduction on strain phenotypes [4]. The concept of the **reduced genome** is increasingly being adopted as an engineering method to enhance strain performance. Through rationally designed selective streamlining of unwanted genotypes from an intact genome, this approach has resulted in improved genomic stability, biomass yield, and commodity production in industrial chassis. Minimized genomes also translate into greater amenability of genetic constituents. In particular, the total synthesis of genome-reduced *Escherichia coli* MDS42 enabled the genome-wide rewriting of its natural genetic codes, opening up novel modalities of viral defense and *in vivo* biosynthesis of noncanonical polypeptides [5–9]. As such, the emergent utility of synthetic bacterial genomes is becoming more prevalent in biotechnology, as evident by the growing list of genome-streamlined (non-)model organisms for various bioprocess applications (Table 1).

[1]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea
[2]KI for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea
[3]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA
[4]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet, Kongens, Lyngby, Denmark
[5]Graduate School of Engineering Biology, Korea Advanced Institute of Science and Technology, Daejeon, 34141, Republic of Korea

 https://doi.org/10.1016/j.tibtech.2024.02.008

Herein we review the recently developed genome-streamlined chassis strains and synthetic genomes, covering aspects of top-down and bottom-up tracks of genome synthesis, implications, and possible avenues of synthetic reduced genomes in biotechnology.

Navigating top-down and bottom-up paths in minimal genome construction

Reduced genomes are modified genomes with fewer genes or DNA sequences than their natural counterparts. The construction of a reduced genome aims to streamline the genetic makeup of an organism into its essential components, providing insights into the fundamental understanding of genomes and fostering applications in biotechnology [4,10]. This reductionist approach entails the systematic removal of nonessential DNA segments from the genome of an organism to discern the minimal set of genetic elements necessary for basic cellular functions. Although various genome recombineering techniques are currently available, identifying essential genomic regions remains difficult, necessitating a combination of computational and experimental strategies [11–15] (Figure 1, Key figure).

Over the past decade, advances in next-generation sequencing have facilitated the sequencing of an extensive number of bacterial genomes. This has significantly enhanced our understanding of the minimal genomic content achieved through comparisons among hundreds of related species [15,16]. Despite this extensive sequence information, the direct integration of reduced genome construction has been limited. The current approach relies predominantly on rational design strategies that hinge on the elimination of nonessential cellular functions, including motility, prophages, secondary metabolism, and horizontally acquired genomic islands (Table 1). Although genome-wide high-resolution essentiality surveys using recent techniques have not yet been implemented, reduced genomes designed rationally based on an early comparative genomics approach exhibit advantageous traits. However, these engineered genomes are not without challenges, manifesting in instances of growth retardation and unforeseen abnormal behaviors [17,18]. Analyses of these reduced genomes have revealed a gradual decline in fitness, even in cases where no essential cellular functions have been removed [19,20]. Factors contributing to this decline have not been fully elucidated, highlighting the complexity of cellular systems and the gaps between our updated genomic understanding and design implementation. In response to these challenges, a reassessment was conducted on the essential genetic composition and methodologies employed, revealing inaccuracies in both the identified essential gene set and investigative methods applied [21–23]. Contributors to these effects include metabolic imbalance [24], improper levels of the protein quality control system [20], the influence of quasi-essential genes [2], and synthetic lethal interactions resulting from combinations of gene deletions that could not be detected in single knockout studies [25].

An alternative approach involves the accumulation of iterative random deletions [26,27]. This deletion strategy is selection driven, avoiding clones with undesirable fitness defects. Interestingly, randomized deletions also highlighted the competitive advantages of genomes with deletions in prophages and flagella, supporting the validity of rational design. A more comprehensive understanding and implementation of precise high-throughput approaches are imperative to navigate the complexities associated with genome reduction.

Following the advances and maturation of technologies that 'read' and 'analyze' genomes, DNA synthesis technologies enabling genome 'writing' and 'editing' have followed suit. Collective progress in systems and synthetic biology has paved the way for the design, synthesis, and assembly of customized genomes (Figure 1). This has led the research community to shift toward bottom-up construction methodologies to understand and engineer genomes.

[6]http://cholab.or.kr/
[7]These authors contributed equally to this work

*Correspondence:
bcho@kaist.ac.kr (B.-K. Cho).

Completion of the *Mycoplasma mycoides* minimal genome JCVI-syn3.0 project prompted researchers to envision rewriting the entire genome by design. Syn61 is the first-of-its-kind designer genome chemically synthesized using the genome-reduced *E. coli* MDS42 as a template [5,28], in which the native genome was replaced with 100 kb synthetic fragments at a time using the CRISPR/Cas9-assisted, λ-red recombineering method REXER [29]. Syn61 differs from MDS42 in that two serine codons and amber stop codons (TAG) are substituted with synonymous codons. This represents an important milestone in synthetic biology in that the genome-wide compression of the two serine codons, TCA and TCG, has further expanded the capacity of bacteria to incorporate more than one type of noncanonical amino acid (ncAA) and build more stringent genetic containment measures [6–8]. Such engineering features have broad implications for bioproduction, viral resistance, and biosafety, which are discussed in the next section. The genome-wide recoding cost Syn61 an observable reduction in fitness, growing at a rate 1.6-fold slower than MDS42 [5]. Similar to what was observed in genome-streamlined strains, this indicates perturbed cellular processes in synthetic backgrounds, calling for additional optimization efforts to configure the synthetic genotypes.

Adaptive laboratory evolution as a strategy for genome optimization

Adaptive laboratory evolution (ALE) is at the forefront of genomic optimization strategies, offering a dynamic and controlled approach to shape and refine the genetic landscape of microorganisms [30]. This methodology involves subjecting microbial populations to repeated cycles of selective pressure in laboratory environments, allowing for spontaneous adaptation and evolution of their genomes. It harnesses the inherent ability of microbes to adapt to diverse stressors, without requiring an in-depth understanding of the biology of complex systems. An intriguing aspect of this paradigm is the adaptability of minimal genomes to laboratory settings. Interestingly, despite concerns about potential limitations – such as a reduced gene repertoire that constrains evolution and diminished insertion sequence (IS)-mediated mutagenesis – studies have demonstrated that reduced genomes evolve effectively [24,31], though there are conflicting reports on spontaneous mutation rates in reduced genomes [25,31]. Another notable milestone involved the demonstration of the evolutionary dynamics of minimal synthetic cells, resulting in a 37% increase in fitness compared with their nonminimal counterparts [32–34].

The primary objective of ALE is to restore fitness defects in genome-streamlined organisms, with minimal genomes tending to recover their growth to levels comparable with those of their wild-type (WT) counterparts. Contributors to these effects include metabolic rebalancing [24,35], an elevated protein quality control system [20], and genes related to cell division and morphology [32], which are not observed in a nonminimal background (Box 1). This reflects a distinct adaptive strategy of the minimal genome, which likely stems from an imbalance in various cellular functions that cannot be reshaped by endogenous regulatory mechanisms.

The presence of several unknown genes, participating in diverse cellular processes, further complicates the understanding of the mechanistic basis of fitness restoration. Through complex systematic analysis, enabled by integrative multi-omics analytics, we can get a glimpse of the molecular basis of fitness restoration. For instance, the molecular basis of fitness restoration in eMS57, a reduced-genome *E. coli*, has been partially revealed through integrative analysis of different omics tools that included genome sequencing and mutation calling, transcriptomes, translatomes, and interactomes [24]. Briefly, characterization of causal mutations in RNA polymerase subunits of *rpoS*, *rpoD*, and *rpoC*, led to the identification of distinct promoter binding profiles and expression changes in downstream genes associated with central metabolic pathways and redox potentials. Although direct causation of each mutant on strain

Glossary

Adaptive laboratory evolution (ALE): the experimental evolution of bacterial populations in the presence of a specific, user-defined selection pressure with the final goal of obtaining mutant clones with desirable (adaptive) phenotypes.
Biofoundry: an automated, synthetic, biology-focused infrastructure that integrates robotics, analytical equipment, software, and digitized processing units to generate, collect, process, and analyze biodata in a multiplexed and high-throughput manner, thereby accelerating the pace of DBTL.
Genome recoding: customization of natural genetic codes (codons) within the genome of an organism to incorporate amino acids that are otherwise nonorthogonal in nature.
Genome-scale models (GEMs) of metabolism: an exhaustive mathematical reconstruction of biochemical reactions in an organism in a computational format that enables the prediction and analysis of metabolic flux distributions (such as rates of biomass generation and metabolite secretion) of the reconstructed metabolic network.
Horizontal gene transfer (HGT): the transfer of genetic material between unrelated species, facilitating bidirectional spread and the acquisition of novel genetic (and phenotypic) traits.
Minimal genome: a genome retaining the minimal set of genetic elements necessary and sufficient for the survival and self-replication of an organism.
Reduced genome: a genome that has undergone the deliberate elimination of genomic constituents that are not essential for desired cellular functions. This concept differs from the minimal genome in the selective removal of genetic elements by necessity, and other redundant elements may still be retained.
Synthetic biology: an interdisciplinary field that aims to augment natural biological systems and create new biological entities using standardized bioparts, modules, chassis, and systems.
Systems biology: an interdisciplinary field that combines computational analysis and mathematical modeling frameworks aimed at understanding biological systems in terms of interacting networks instead of focusing on individual components in isolation.
Whole-cell model (WCM): a comprehensive *in silico* framework

fitness has not been validated, the analytic workflow facilitated facile comprehension of the underlying adaptive traits.

simulating the molecular, kinetic, and physiological processes and interactions occurring in a dividing cell in spatiotemporal resolution. In contrast to GEM, the WCM is currently exclusive to JCVI-syn3.0A.

As such, ALE is an effective method to overcome these challenges as it comprehensively affects cellular processes, including metabolism, transcription, and stress responses. In contrast to synthetic biology approaches, ALE works in the absence of an *a priori* understanding of biological systems. ALE can restore fitness for deployment in biotechnological applications, as well as provide fundamental insights into the consequences of genome-scale engineering. Identifying mutations in endpoints retroactively reveals the changes that occur during ALE. It thus provides a rapid and robust approach for improving the functionality of minimal genomes and exploring design principles at an unprecedented scale.

Genome-streamlining as an engineering tool to improve the bioproduction capacity of chassis strains

One of the main advantages of genome reduction engineering is its effect on bioproduction. The proposed benefits are as follows: (i) genome reduction engineering provides more efficient and higher production of valuable biochemicals and peptides by saving cellular resources required for genome maintenance, (ii) higher DNA yield from the same amounts of nutrients due to the lower genome maintenance cost, and (iii) a stable and predictable genetic background from attenuated mutagenic potentials (Figure 2A). Reduced genomes exhibit several advantages over their parental counterparts, as evidenced by their higher biomass yields, accelerated growth rates, and enhanced energy and redox capacities (Table 1). In particular, numerous reduced genomes exhibit high productivity for the synthesis of valuable biochemicals such as bioplastics, vitamins, and antibiotics [36–40].

One notable industrial application lies in the domain of antibiotic production in *Streptomyces* spp., renowned for their ability to synthesize diverse bioactive secondary metabolites, including antibiotics [41]. The engineered *Streptomyces albus* strain Del14, devoid of 15 endogenous secondary metabolite biosynthetic gene clusters (smBGCs), exhibits increased production of polyketides, phosphoglycolipids, nucleosides, and antibiotics [40]. Similarly, *Streptomyces chattanoogensis*, which lacks smBGCs, displayed increased yields of the nonribosomal peptide indigoidine [39]. In another study, *Streptomyces lividans* ΔYA11, with 11 endogenous smBGCs removed, exhibited a fourfold increase in lanthipeptide deoxycinnamycin production compared with its parental counterpart [42]. Interestingly, a comparative analysis of two reduced genome strains, Del14 and ΔYA11, revealed distinct metabolic characteristics, with ΔYA11 excelling in deoxycinnamycin productivity, whereas Del14 performed better in the production of the nucleoside antibiotic tunicamycin [42]. Reduced genomes of *Streptomyces* spp. seemingly have different metabolic characteristics, such that the strains could specialize in certain types of secondary metabolites. The removal of smBGCs from the genome consistently results in a higher production capacity of heterologous secondary metabolites, highlighting the potential for clean secondary metabolite profiles and the discovery of unknown smBGCs [43]. In addition to *Streptomyces* spp., the reduced genomes of the industrially relevant bacteria *Corynebacterium glutamicum*, *Pseudomonas putida*, and *E. coli* exhibited better production of diverse biochemicals such as *N*-methylanthranilate, polyhydroxyalkanoate, and *myo*-inositol (vitamin B_8) [36,37,44].

The reduced genomes of industrially relevant bacteria illustrate the rerouting of cellular resources to other cellular processes, which could result in distinct metabolic traits in the reduced genomes of different species. Without heterologous metabolism, the reduced genomes exhibited distinct metabolic traits that affected the energy or redox equivalents. *Pseudomonas mendocina*

Table 1. Major milestones of recently constructed reduced genomes[a]

Ancestor	Milestone strain	Original genome	Reduction (fraction)	Genetic alterations (removal of)	Phenotypic characteristics						Miscellaneous characteristics	Refs
					BY	GR	GS	HPE	CP	TE		
Escherichia coli K-12 MG1655	H16	4.64 Mb	245 kb (5.3%)	Nine random genomic regions including mobile elements, prophage, restriction modification systems	+	+			+	+		[27]
	CDΔ3456		313 kb (6.8%)	Random regions between transposon insertions		C						[76]
	MDS12, MDS42		376 kb (8.1%), 663 kb (14.3%)	12 K-islands, prophages, transposable elements	+	C	+		+	+		[28,77–79]
	MDS69		939 kb (20.2%)	Descendent of MDS42; surface structure, mobile genetic elements	–	–					Increased acid tolerance	[80]
	MS56		1.07 Mb (23.0%)	Descendent of MDS42; motility, LPS biosynthesis	C	C	+	+		+		[51]
	eMS57		1.09 Mb (23.5%)	Laboratory evolved MS56, spontaneous deletion of 21 kb region containing mismatch repair protein, and stress sigma factor	C	C		+	+		ALE applied to recover fitness; increased plasmid yield, antibiotic sensitivity, redox ratios, and translation efficiency; decreased energy capacity (ATP)	[24,36,49]
	Δ16		1.38 Mb (29.7%)	Nonessential genes		–					Abnormal morphology and nucleoid development	[17]
	Δ41		2.04 Mb (43.9%)	Descendent of Δ16; sequential deletion of nonessential genes		–					ALE applied to recover fitness	[20]
E. coli K-12 W3110	MGF-01	4.65 Mb	1.03 Mb (22.2%)	Nonessential regions	+				+			[81,82]
	DGF-298		1.66 Mb (35.6%)	Descendent of MGF-01; IS elements		C/+		C				[83,84]
Bacillus subtilis 168	MBG874	4.22 Mb	814 kb (19.3%)	Prophages, antibiotic biosynthesis, misc. nonessential genes	+	–			+	–	Decreased sporulation	[85]
	MG1M		991 kb (23.5%)	Nonessential genes, sporulation		C		C				[86]
	MGP254		1.49 Mb (35.3%)	Prophages and nonessential genes		–	+				Increased mitomycin C (DNA-damaging agent) resistance	[25]

	PG10, PS38		1.46 (34.6%), 1.54 Mb (36.5%)	Sporulation, antibiotic biosynthesis, motility, secondary carbon metabolism, hypothetical genes	–	–			+		Cell elongation	[18,38]
	IIG-Bs27-47-24		1.32 Mb (30.1%)	Extracellular proteases, prophages, and key sporulation genes		–		+			Altered thiol–disulfide oxidoreductase requirements for disulfide bonding	[87]
Bacillus amyloliquefaciens LL3	GR167	4.00 Mb	168 kb (4.2%)	Prophages, unknown genes, antibiotic biosynthesis	+	+		+	+	+	Increased mitomycin C (DNA-damaging agent) resistance	[47]
Corynebacterium glutamicum ATCC13032	MB001	3.28 Mb	205 kb (6.2%)	Prophages	C	C	C	+		+	Competitive growth advantage over the ancestor	[35,88]
	C1*		439 kb (13.4%)	Descendent of MB001; nonessential genes	C	C			+		Long-term growth stability	[44,89,90]
Halomonas bluephagenesis TD10	ΔPS1234Δ50k	4.14 Mb	121 kb (2.9%)	Flagellum, exopolysaccharide biosynthesis, O-antigen	–				C	+	Became electrotransformable	[54]
Lactococcus lactis NZ9000	9k-4	2.53 Mb	72 kb (2.8%)	Large nonessential regions, cryptic phages	C/+	+		+			Higher energy capacity (ATP)	[46]
L. lactis N8	N8-8	2.57 Mb	176 kb (6.8%)	Prophages and genomic islands	C	+					Nisin production unchanged	[91]
Magnetospirillum gryphiswaldense	ΔTZ-17	4.16 Mb	228 kb (5.5%)	Secondary metabolite biosynthetic clusters, mobile elements, prophages		C	+					[52]
Pseudomonas chlororaphis GP72	MDS22	6.66 Mb	686 kb (10.3%)	22 species-specific genome regions (5 smBGCs and 525 nonessential genes)		C			+			[92]
Pseudomonas mendocina NK-01	NKU421	5.43 Mb	418 kb (7.7%)	Prophages, genomic island, hypothetical genes	+	C		C	+	+	Increased energy capacity (ATP); comparable redox ratios	[45]
Pseudomonas putida KT2440	EM383	6.18 Mb	266 kb (4.3%)	Prophages, transposable elements, restriction modification, endonucleases, flagellum	+			+		+/–	Short lag phase; increased energy capacity (ATP), redox ratios, antibiotic sensitivity, cellular viability, plasmid yield, and chromosomal gene integration efficiency	[48]

(continued on next page)

Table 1. (continued)

Ancestor	Milestone strain	Original genome	Reduction (fraction)	Genetic alterations (removal of)	Phenotypic characteristics						Miscellaneous characteristics	Refs
					BY	GR	GS	HPE	CP	TE		
	KTU-U13		255 kb (4.1%)	13 genomic islands with GC content distinct from the genome		C	+	+	+	+		[37]
	KTU-U27		516 kb (8.4%)	Descendent of KTU-U13; 10 genomic islands	+	+			+	+		[93]
Pseudomonas taiwanensis VLB120	GRC1	5.97 Mb	640 kb (10.7%)	Plasmid, prophages, flagellum, biofilm	+	+			+			[94]
Schlegelella brevitalea DSM 7029	DT10	6.48 Mb	316 kb (4.9%)	Transposons, prophages, biosynthetic gene clusters	+	+			+	+/–	Increased cellular viability	[95]
Streptomyces avermitilis K139	SUKA17	9.03 Mb	1.67 Mb (18.5%)	(secondary metabolite biosynthetic gene clusters) smBGC, IS elements, non-essential genes in subtelomeric regions	+	C			+			[96,97]
Streptomyces albus J1074	Del14	6.84 Mb	500 kb (7.3%)	smBGCs		C			+			[40]
Streptomyces chattanoogensis L10	L321	Approx. 9.1 Mb	736 kb (approx. 8.1%)	smBGCs, IS elements	C	C	+	+	+	C	Increased energy capacity (ATP) and redox ratios	[39]
Streptomyces coelicolor A3 (2)	ZM12	8.67 Mb	1.60 Mb (18.5%)	smBGCs, subtelomeric regions					+			[98]
Streptomyces lividans TK24	ΔYA11	8.35 Mb	229 kb (2.7%)	smBGCs	+	C			+			[42]

[a]Abbreviations: +, increase; –, decrease; BY, biomass yield; C, comparable with the parental counterpart; CP, chemical production; GR, growth rate; GS, genome stability, represented by the mutation rate or a gene inactivation; HPE, heterologous protein expression; smBGCs, secondary metabolite biosynthetic gene clusters; TE, transformation efficiency.

Key figure

An overview of analytic tools and techniques involved in the synthesis of reduced and synthetic minimal genomes

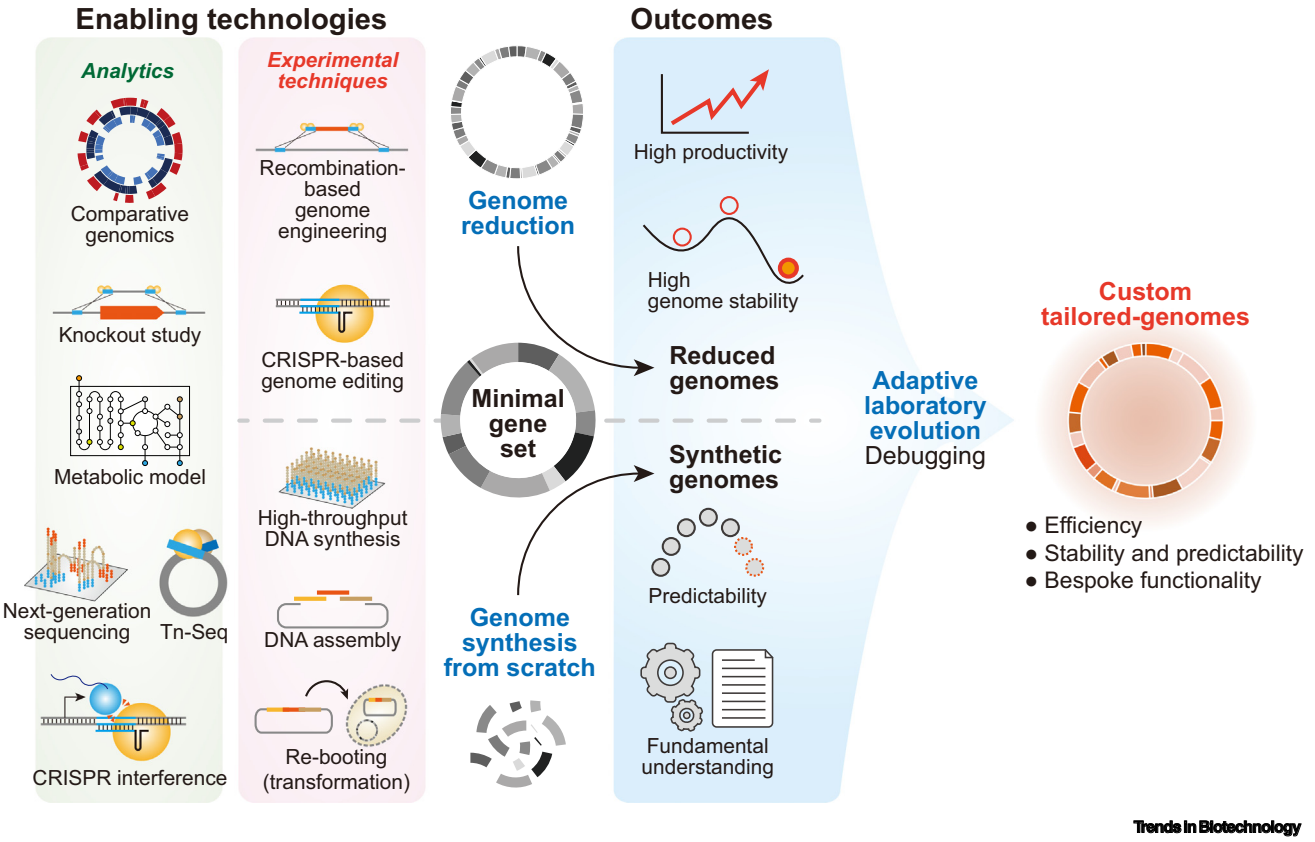

Figure 1. Analytics: analytic technologies enable mapping of genotype–phenotype relationships. Comparative genomics facilitates the identification of essential core genes conserved across species. Knockout studies provide direct genotype–phenotype mapping of gene targets with high confidence, but it is limited in throughput and may be labor-intensive. *In silico* metabolic model simulations and next-generation sequencing technologies provide inference of genome properties (i.e., gene essentiality) in a high-throughput manner. However, their heuristic nature may necessitate additional experimental validation steps. Experimental technique: homology-based recombineering and CRISPR-mediated genome editing techniques enable site-specific modification of target genomic regions, representing two key methods for top-down genome streamlining. By contrast, bottom-up genome synthesis requires the assembly of phosphoramidite-based designer DNA fragments into larger intermediates. Gibson cloning, bacterial artificial chromosome (BAC) and yeast artificial chromosome (YAC) assembly, and appropriate genome re-booting technologies are the key prerequisites for any bottom-up synthesis project. Outcomes: reduced genetic complexity and redundancies in synthetic reduced genomes hold significance in biotechnology and advances fundamental understanding of basic biological principles.

NKU421 and *Lactococcus lactis* 9k-4 l demonstrated 11-fold and 1.2-fold higher ATP content, respectively [45,46], correlating with increased heterologous protein production (Figure 2A). This resulted in higher heterologous protein production, which was not unexpected, as the rate of translation is dependent on the number of ribosomes regulated by ATP availability. Interestingly, the adaptively evolved reduced genome *E. coli* eMS57 displayed lower ATP content while achieving higher protein productivity, indicating unique translational efficiency in handling varying transcript levels [24]. This highlights the versatility of reduced genomes, not only in metabolite production but also in protein and peptide synthesis. Although the exact underlying cellular

Box 1. Evolvability of minimal genomes

One notable milestone in understanding the biology of the minimal genomes involves a demonstration of evolutionary dynamics of the minimal synthetic cells [32–34]. The primary objective was to restore the fitness defects that often inflict genome-streamlined organisms, such as JCVI-syn3.0B, which showed a 57% reduction in strain fitness [32]. After 2000 generations of serial passaging, the evolved minimal genome regained fitness comparable with that of the nonminimal cell. The rate of fitness gain was 37% faster than the evolved nonminimal counterpart. This presented an interesting discovery in evolutionary biology, where it had been posited that streamlined sequence space and genetic redundancies in genome-minimized organisms would constrain evolution and reduce tolerance toward accumulating mutational loads. The authors also identified mutant phenotypes that disproportionately affected cell size in different genomic contexts: where a nonsense mutation on *ftsZ* (regulating cell division and morphology) increased cell size in a nonminimal background but not in the minimal genome. This seemed to reflect a distinct adaptive strategy of the minimal genome which retains a smaller number of transmembrane proteins and transporter proteins for metabolite exchange in and out of cytoplasm [32]. Independent ALE experiments on JCVI-syn3.0 derivatives have demonstrated divergent evolutionary trajectories across replicate ALE lineages, implications of uncharacterized (mutant) genes on fitness [33], and a striking capacity of the minimal genome to develop resistance against a spectrum of antibiotics [34] Taken together, this reveals that the genome that retains the bare minimal genetic material nonetheless retains the capacity to adapt, evolve, and acquire novel phenotypes.

mechanisms that increase ATP content remain unclear, transcriptome analysis of eMS57 revealed different uses of the two glycolytic pathways, leading to a higher redox capacity as a trade-off for reduced ATP yield [24]. Similarly, the reduced genome of *Bacillus amyloliquefaciens* GR167 exhibited a higher redox capacity, affirming a metabolic shift with ATP and a redox balance distinct from its WT counterpart [47].

Moreover, *P. putida* reduced genome EM383 and *E. coli* eMS57 exhibited increased plasmid DNA (pDNA) yield, ranging from 20% to 200% [48,49]. Notably, eMS57 demonstrates over a threefold increase in pVAX1 plasmid production, a backbone vector for pDNA vaccines, compared with the commercial cloning strain *E. coli* DH5α. The amplification of pDNA is important in the DNA and mRNA vaccine industries [50]. Consequently, reduced genomes with enhanced pDNA yields show considerable potential in the biopharmaceutical sector (Figure 2A). In general, these genome reductions show significant promise for diverse industrial applications owing to their improved characteristics. These improvements are attributed to the alleviated burden of genome maintenance, resulting in more efficient DNA, mRNA, and protein production, and an overall streamlined metabolic load within the cell.

Notably, the deliberate removal of transposable elements, including IS elements and transposons, significantly contributes to enhanced genetic stability, rendering reduced genomes particularly well suited as heterologous hosts. In particular, IS-free *E. coli* derivatives MDS42 and MS56 were able to propagate recombinant vectors more stably over longer durations than their WT counterparts [28,51]. Similarly, the sustained expression of recombinant proteins in other IS-free reduced genomes of *Magnetospirillum gryphiswaldense* and *P. putida* [37,52] underscores their resilience against inactivating mutations typically originating from the insertion of transposable elements [53]. Interestingly, genome-reduced *Bacillus subtilis* lacking >30% of its original genome, MGP229, exhibited a 25-fold lower spontaneous mutation rate, independent of transposable elements, possibly owing to a more efficient mismatch repair system [25]. These emerging characteristics highlight the utility of reduced genome sequences as attractive hosts for producing heterologous proteins and biosynthetic enzymes.

Finally, additional emergent properties have been observed, including increased transformation efficiency [54] and antibiotic sensitivity [49]. Although the precise mechanisms underlying these enhancements remain elusive, current hypotheses suggest that augmented nucleic acid yield, alterations in membrane proteins, lipopolysaccharide (LPS) biosynthesis, and

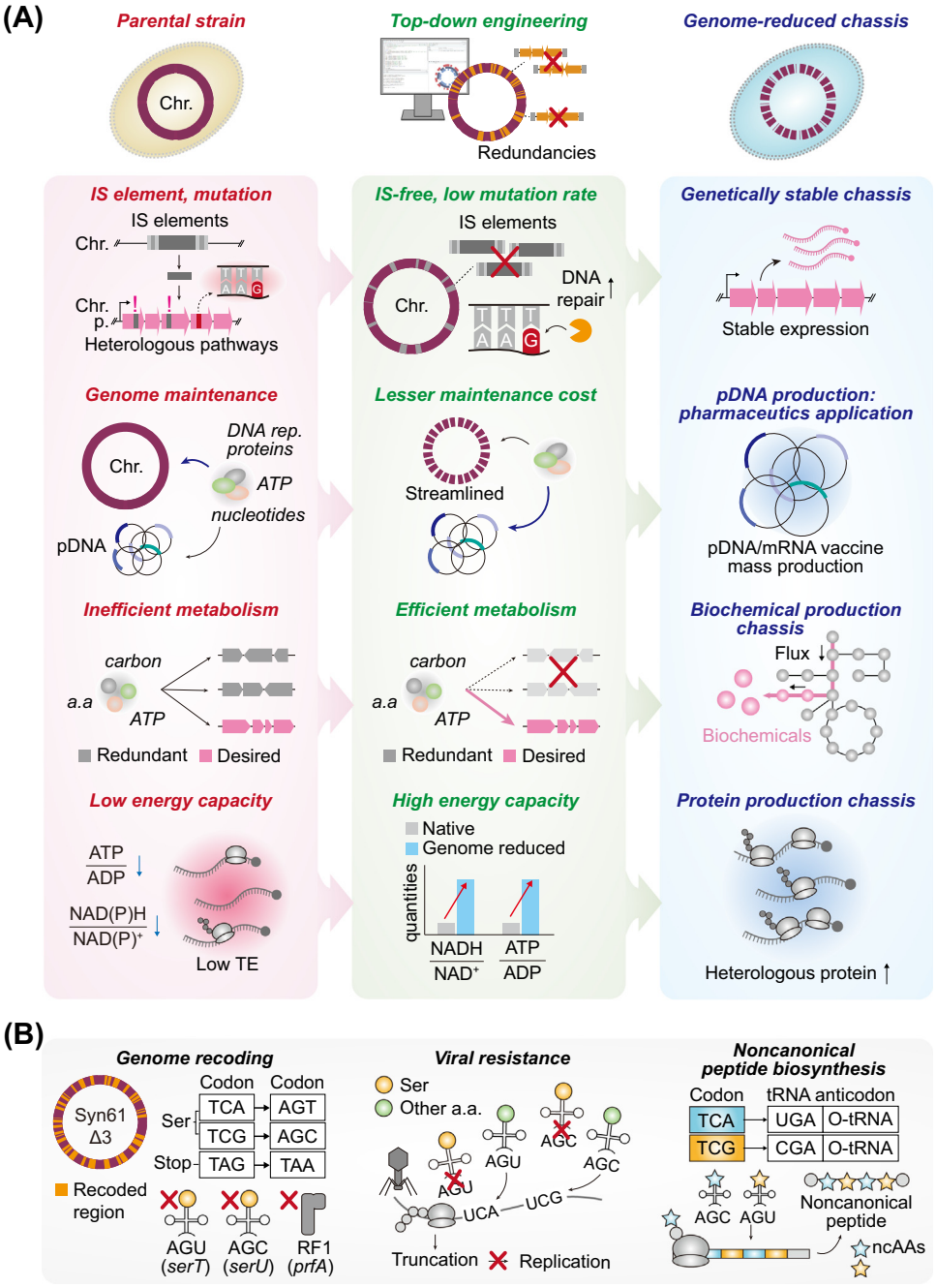

Figure 2. Biotechnologically important features in synthetic genome-streamlined organisms. (A) A key summary of top-down genome engineering aimed at improving biotechnologically important phenotypes. (i) The elimination of mobile insertion elements on a genome-wide scale enhances genome stability and DNA replication fidelity, yielding a genetically robust platform capable of supporting stable heterologous gene expression. (ii) The genome-scale trimming of functionally redundant elements leads to reduced genome maintenance cost. The substrates conserved in the process can be redirected toward plasmid DNA (pDNA) maintenance, thus enhancing the total yield of pDNA. As described in the main text, this feature is significant in the DNA and mRNA vaccine industry. (iii) In line with point (ii), conserved cellular substrates can be redirected toward biosynthesis of desired compounds, enabling construction of a robust biochemical production

(Figure legend continued at the bottom of the next page.)

modifications to the cell wall collectively influence the physical properties of the cell. An illustrative example is the reduced genome of the previously non-electrotransformable *Halomonas bluephagensis*, which attained electrotransformability after the removal of exopolysaccharides and the O-antigen [54]. These findings highlight the multifaceted impact of genome reduction on cellular traits and reveal novel characteristics that extend beyond traditional considerations.

Novel biotechnological modalities enabled by synthetic genomes

The **genome recoding** of the two serine and amber codons in Syn61 offers opportunities to test novel modalities of viral resistance and biosynthesis of non-canonical peptides, which represent avenues that synthetic biology has long sought to address: viral resistance and biosynthesis of non-canonical heteropolymers *in vivo* [55] (Figure 2B and Table 2).

First, it was hypothesized that the removal of native codons and cognate tRNAs would prevent viral genome replication. In a previous study, an amber codon-recoded *E. coli* lacking the cognate release factor (RF) displayed attenuated but incomplete resistance to an invading bacteriophage [55]. This was ascribed to the rare use of amber codons in translation termination and the unaffected capacity to translate full-length viral proteins. Instead, the compression of sense codons, which are abundant and almost ubiquitous throughout coding regions, is expected to confer more stringent immunity against viruses. For instance, a Syn61 derivative lacking tRNAs and RF cognate to the recoded codons (Syn61Δ3) demonstrated a broad-range phage resistance, to which Syn61 lacking only RF remained susceptible (Table 2) [6]. Although this has marked a promising development in synthetic viral resistance, **horizontal gene transfer (HGT)** of foreign tRNAs can still complement the artificially deleted capacity to read certain genetic codes. In this regard, engineering bacteria to read amino acid-swapped genetic codes using chimeric tRNAs may help avoid the unwanted restoration of TCA and TCG codon functionalities and establish a 'genetic firewall' that prohibits the HGT of mobile genetic elements and viral infection [7,8] (Table 2). The biocontainment approach implemented for Syn61 derivatives presents a novel modality for multiviral resistance in otherwise susceptible workhorse strains. In addition, a bidirectional genetic containment strategy that restricts HGT between organisms should motivate synthetic genome recoding as an engineering framework in genetically modified organism research and biotech industries.

Another emergent property of genome-recoded organisms is programmable cell-based biosynthesis of noncanonical heteropolymers in bacteria [5,6,9]. The translational incorporation of ncAAs into recombinant biomolecules allows the engineering of new functionalities with applications in probing and imaging [56]. The expression of TCR-suppressing aminoacyl-tRNA synthetase (aaRS)/tRNA pairs orthogonal to ncAAs in a Syn61Δ3 derivative enabled the co-translational incorporation of a desired ncAA into the codon-reassigned sequences, which is otherwise toxic in nonrecoded organisms [5]. Another challenge is the engineering and discovery of appropriate aaRS/tRNA pairs that incorporate the desired ncAAs into the intervening sequence [57]. In this regard, Syn61Δ3 derivatives serve as a testbed to define and

chassis. (iv) The higher energy capacity observed in genome-reduced derivatives can inadvertently lead to an increased capacity for heterologous protein production. (B) The biotechnologically relevant utility of synthetic recoded *Escherichia coli* Syn61Δ3. The expression of amino acid swapped, chimeric aminoacyl-tRNA synthetase (aaRS)/tRNA pairs that decode TCA and TCG in the Syn61Δ3 background induces mistranslation of nascent viral proteomes, thus preventing replication of invading viruses. Similarly, the expression of engineered, mutually orthogonal aaRS/tRNA pairs that recognize noncanonical amino acids (ncAAs) enables translational incorporation of ncAAs. This holds the potential to effectively reprogram Syn61Δ3 as a cell factory for diverse non-canonical peptides. Abbreviations: a.a., amino acid; IS, insertion sequence; TE, transformation efficiency.

Table 2. Summary of biotechnologically relevant applications of Syn61 derivatives

Base strain	Genetic modifications/heterologous expressions	Engineered phenotypes	Refs
Syn61	Directed evolution of fast-growing Syn61 through random mutagenesis, generating Syn61(ev2) Deletion of Δ*serT* ($tRNA^{Ser}_{UGA}$), Δ*serU* ($tRNA^{Ser}_{CGA}$), Δ*prfA* in Syn61(ev2), generating Syn61Δ3	Resistance against a cocktail of five bacteriophages Incorporation of up to nine ncAAs in a single polypeptide, and biosynthesis of noncanonical heteropolymers by reassigning TCR or TAG using orthogonal aaRS/tRNA pairs recognizing distinct ncAAs	[6]
Syn61Δ3	Expression of chimeric tRNAs for alanine, histidine, leucine, and proline that decode TCA and TCG codons	Reassignment of serine TCR codons to alanine, histidine, proline, or leucine HGT of deleted serine tRNAs from wild-type into Syn61Δ3 no longer interrupts the engineered orthogonality of TCA and TCG codons	[7]
Syn61Δ3	Directed evolution of TCA and TCG suppressors from the endogenous *Escherichia coli* and bacteriophage-derived leuC tRNA.	Serine-to-leucine swapped genetic code (TCA and TCG codons) HGT of serine tRNAs from viral mobile tRNA pools into Syn61Δ3 no longer abolishes viral resistance	[8]
Syn61Δ3(ev5)	Directed evolution of fast-growing Syn61Δ3 through three sequential rounds of random mutagenesis, generating Syn61(ev5) Heterologous expression of engineered aaRS/tRNA pairs that recognize distinct ncAAs and decodes TCR	Programmed incorporation of ncAAs into the codon-reassigned TCR over the length of depsipeptide-encoding genes *In vivo* biosynthesis of 12 nonnatural depsipeptide macrocycles	[9]

assess the orthogonality and the efficiency of engineered aaRS/tRNA pairs that incorporate distinct ncAAs at defined codons [9]. Utilizing this characteristic, dozens of noncanonical peptides were biosynthesized in the Syn61Δ3 background (Table 2) [9]. This approach has presented exciting advances in several areas. First, the translational incorporation of more than one type of ncAA simultaneously greatly expands the variety of noncanonical heteropolymers that can be manufactured *in vivo*, as opposed to the previous approach, which is dependent solely on amber codon suppression [55]. Second, cell-based noncanonical heteropolymer biosynthesis may enable the random shuffling of different types of ncAAs at integration sites over the length of peptides to probe for novel functional discoveries [56]. The recently showcased tRNA display technology is expected to enable incorporation of an even broader repertoire of noncanonical monomers in the cellular proteomes [58]. Furthermore, the amenability to fermentation opens up the possibility of scalable production of the designed noncanonical heteropolymers.

Collectively, these findings represent landmark milestones in advances in synthetic biology that have brought about novel engineering modalities with significant relevance to biotechnology. The augmentation of viral resistance, bidirectional genetic containment, and capacity to incorporate multiple ncAAs shown in Syn61 represent the first few utilities of this emerging technology. What awaits remains an open question.

It is also important to note that genome-reduced strains facilitate the implementation of such synthetic resistance against viruses and HGT. This is primarily due to the reduced cost and labor associated with the total synthesis of recoded genomes, which is proportional to the genome size. Additionally, the beneficial properties of reduced genomes, such as enhanced transformation efficiency and genome stability, could further expedite the efficiency and speed of genome engineering. Reduced genomes also harbor features that are coveted for industrial applications, including, but not limited to, improved bioproduction yield of non-native chemicals, recombinant proteins, and vectors. Collectively, the beneficial traits resulting from genome streamlining endow genome-reduced strain a status of 'next-generation production chassis'. In this respect, the bottom-up synthesis of Syn61 presents a milestone advance, embracing inherent features and advantages that a genome-reduced chassis could provide.

Concluding remarks and future perspectives

Decade-long efforts to construct genomes that are smaller and simpler than those found in nature have culminated in the total chemical synthesis of minimal genomes capable of self-replication. Two different approaches have driven genome reduction: top-down and bottom-up. The first approach aims to improve the performance of chassis strains, in which the reduction in genome redundancies, genome instability, and unpredictability has translated to better and more amenable strain performance. To this end, top-down genome streamlining presents a practical solution for optimizing a (potential) industrial chassis. The second approach entails exploring the fundamental properties of living organisms, particularly the minimal genetic components that constitute life [59]. Years after the completion of the JCVI-syn3.0 synthesis project, the properties of JCVI-syn3.0 and its derivatives are still being actively explored in terms of protein function prediction [60], functional characterization of (non)-native genes [61–63], evolutionary dynamics [32–34], and *in silico* model-based simulations of metabolism, gene essentiality, and whole-cell life processes [64,65] (Box 2). The bottom-up approach is more fundamental and more likely to be our long-term goal.

JCVI's bottom-up genome synthesis project has also enabled significant advances across synthetic biology disciplines. In particular, the DBTL-driven modular genome design, assembly, and testing approach represents the most standardized and effective strain development pipeline that also facilitates the total synthesis of the recoded *E. coli* genome [5]. With the completion of the Syn61 project, synthetic biology aims to expand its reach to the biotechnology industry, providing novel solutions to challenging problems beyond the scope of existing methodologies (see Outstanding questions). The standardization of bottom-up genome synthesis methods is expected to accelerate the pace and throughput of customized genome design and synthesis, as witnessed in **biofoundry** applications [66]. Recent upgrades in REXER technology present a promising development in this regard. Standardization of the REXER iteration design through the incorporation of universal Cas9 spacer arrays has massively accelerated iterative genome replacement (termed CONEXER), enabling the implementation of 0.5 Mb synthetic DNA within 10 days [67].

Bottom-up genome synthesis is still in its infancy, and the current synthetic genome design is largely confined to replicating a 'synthetic twin' of a naturally occurring genome. We envision

Box 2. *In silico* modeling of minimal genomes

One notable advance stemming from synthetic minimal genomes involves fine-grained *in silico* modeling, which aims to simulate and explain cellular processes using mathematical models [64] or kinetic simulations [65]. The **genome-scale model (GEM) of metabolism for** JCVI-syn3.0A, representing an *in silico* network of an exhaustive list of metabolic reactions in the cell, was reconstructed to map the genotype–phenotype relationship and to perform predictive modeling of cellular phenotypes in response to genetic and environmental perturbations [64]. JCVI-syn3.0A is particularly useful in mathematical modeling because of the minimal number of growth-supporting reactions and biological uncertainties, which would translate to superior predictive performance of the resulting model. The flux balance analysis of the GEM yielded close approximations of the experimental growth rate and proteome abundance, suggesting that the simulated metabolism closely recapitulated the actual metabolic network. In essence, the JCVI-syn3.0A GEM serves as a valuable community resource for investigating the nature of its constituents in relation to growth, motivating future efforts to further minimize the minimal genome [64], and may facilitate mission-oriented strain design (such as virulence-attenuated *Mycoplasma* derivatives) for medical research applications [62,63]. A more recently developed, fully kinetic 3D **whole-cell model (WCM)** that simulates the cell cycle in JCVI-syn3.0A is worthy of mention [65]. This differs from a previous WCM for *E. coli* [74] in the continuous, time-course simulation of cellular constituents between different subcellular networks, which is enabled by the availability of precise cellular compositions in a relatively small cell [64,75]. The current utility of WCM is aimed largely at recapitulating emergent behaviors in cell division for intellectual merit [65]. The application of WCMs is expected to extend to the prediction of pleiotropic effects of subcellular and environmental perturbations in strain design applications in the future. Questions remain regarding whether WCMs will be utilized as a platform for rational engineering design, similar to GEMs. The current pace and agenda of WCM development seems to endorse this idea (www.wholecell.org). Extension of the current WCM to strain optimization problems will likely motivate future WCM efforts for other prominent workhorse organisms.

Outstanding questions

Currently, the construction of genome-reduced microbes, even with technically less challenging top-down approaches, is significantly time-consuming and expensive. Whether we wish to pursue this type of engineering approach or not comes down to need and cost. What aspects of reduced genome traits make them appealing for industrial and biotechnological applications? Are they sufficient to outweigh the liabilities of time and cost?

Despite the increasing number of genome-reduced microbes and their functional superiority over their natural counterparts, we have not witnessed their practical use in industry. What limits the transfer of genome-reduced organisms from laboratory to industry?

The genome-recorded Syn61 derivatives demonstrated groundbreaking advances in bidirectional firewalls against genetic containment (HGT) and *in vivo* biosynthesis of noncanonical peptides. What other specialized functions could synthetic genomics provide us with?

Significant efforts from across disciplines, including, but not limited to, systems and synthetic, structural, molecular, and computational biology, have been made to decipher the remaining biological uncertainties in JCVI-syn3.0. However, dozens of 'seemingly essential' genes require functional annotations. What are we missing here and what other options/approaches/alternatives can we take? Can whole-cell modeling and machine learning help connect the missing dots in essential biological processes?

What is the ultimate minimum number of genes? If only one exists, it is desirable for practical applications. It may be sensitive to unpredictable fluctuations, such as environmental stresses.

that the emergence of enabling technologies will open up possibilities for *de novo* genome design. For instance, one study took REXER a step further, to transfer a genomic segment into another position on the genome [68]. This method, termed chromosomal fission and fusion, enabled the programmed reorganization of genomic segments. The capability to 'create and shuffle' different versions of reorganized genomes is expected to facilitate the exploration of genome modularity and assembly of synthetic genomes with novel genomic architectures [68,69].

Parallel to the advances in synthetic engineering technologies, machine-learning-based analytics now enable an elucidation of the topology of transcription regulatory networks [70]. This advance has led to the discovery of co-regulated gene groups that can be functionally delineated into distinct 'modules' across different bacterial genomes [71], leading one to envision deliberate horizontal transfer of modularized strain-specific genotypes that give rise to particular phenotypes from one species into another. Together, it is speculated that the convergence of emerging technologies in the fields of systems biology and synthetic biology will culminate in customized, new-to-nature synthetic genomes in the foreseeable future.

Although minimal genomes obtained in a top-down manner offer considerable benefits, they are not without limitations, which often manifest in the form of growth retardation or emergence of abnormal properties. The design rationality in the minimal genome may lead to the unintended deletion of essential genes, the functions of which were previously unknown. Notably, the pleiotropic effects associated with large-scale genome reduction, including synthetic lethality, warrant the re-evaluation of reduced genomes using contemporary genome-scale gene essentiality assessments. Revisions based on an updated understanding and the application of advanced techniques – such as CRISPR interference, genome-scale metabolic models, and deep-learning-based essentiality assessments – open up possibilities for novel top-down minimal genome designs [72,73]. Moreover, previous minimal genomes have focused primarily on industrial applications, emphasizing bioprocess phenotypic characteristics, and thus general systematic analyses of them is currently available for only a small fraction of milestone strains. Thus, to advance the field, comprehensive analyses covering a broad spectrum of strains and elucidated design principles are required for a more nuanced understanding of genome functions.

Acknowledgments

This work was supported by the Korea Bio Grand Challenge (2018M3A9H3024759 to B-K.C.) and the Bio and Medical Technology Development Program (2021M3A9I4024308 to B-K.C.) through the National Research Foundation (NRF) of Korea funded by the Ministry of Science and Information and Communication Technology (ICT), and the Basic Science Research Program through the NRF funded by the Ministry of Education (RS-2023-00246928 to S.C.). This work was also funded by a Novo Nordisk Foundation grant (NNF20CC0035580 to B.P.).

Declaration of interests

The authors declare no competing interests.

References

1. Fraser, C.M. *et al.* (1995) The minimal gene complement of *Mycoplasma genitalium*. *Science* 270, 397–403
2. Hutchison 3rd, C.A. *et al.* (2016) Design and synthesis of a minimal bacterial genome. *Science* 351, aad6253
3. Dahabieh, M.S. *et al.* (2020) Multimodal microorganism development: integrating top-down biological engineering with bottom-up rational design. *Trends Biotechnol.* 38, 241–253
4. Choe, D. *et al.* (2016) Minimal genome: worthwhile or worthless efforts toward being smaller? *Biotechnol. J.* 11, 199–211
5. Fredens, J. *et al.* (2019) Total synthesis of *Escherichia coli* with a recoded genome. *Nature* 569, 514–518
6. Robertson, W.E. *et al.* (2021) Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science* 372, 1057–1062
7. Zürcher, J.F. *et al.* (2022) Refactored genetic codes enable bidirectional genetic isolation. *Science* 378, 516–523
8. Nyerges, A. *et al.* (2023) A swapped genetic code prevents viral infections and gene transfer. *Nature* 615, 720–727
9. Spinck, M. *et al.* (2023) Genetically programmed cell-based synthesis of non-natural peptide and depsipeptide macrocycles. *Nat. Chem.* 15, 61–69
10. Hwang, S. *et al.* (2023) Recent advances in non-model bacterial chassis construction. *Curr. Opin. Syst. Biol.* 36, 100471

11. Rousset, F. *et al.* (2018) Genome-wide CRISPR-dCas9 screens in *E. coli* identify essential genes and phage host factors. *PLoS Genet.* 14, e1007749
12. Ryu, J.Y. *et al.* (2019) Deep learning enables high-quality and high-throughput prediction of enzyme commission numbers. *Proc. Natl. Acad. Sci. USA* 116, 13996–14001
13. Harris, C.D. *et al.* (2021) CoreCruncher: fast and robust construction of core genomes in large prokaryotic data sets. *Mol. Biol. Evol.* 38, 727–734
14. Bleem, A. *et al.* (2023) Multiplexed fitness profiling by RB-TnSeq elucidates pathways for lignin-related aromatic catabolism in *Sphingobium* sp. SYK-6. *Cell Rep.* 42, 112847
15. Rosconi, F. *et al.* (2022) A bacterial pan-genome makes gene essentiality strain-dependent and evolvable. *Nat. Microbiol.* 7, 1580–1592
16. Bu, Q.-T. *et al.* (2020) Comprehensive dissection of dispensable genomic regions in *Streptomyces* based on comparative analysis approach. *Microb. Cell Factories* 19, 99
17. Hashimoto, M. *et al.* (2005) Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Mol. Microbiol.* 55, 137–149
18. Reuß, D.R. *et al.* (2017) Large-scale reduction of the *Bacillus subtilis* genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* 27, 289–299
19. Kurokawa, M. and Ying, B.-W. (2019) Experimental challenges for reduced genomes: the cell model. *Microorganisms* 8, 3
20. Kotaka, Y. *et al.* (2023) Mutations identified in engineered *Escherichia coli* with a reduced genome. *Front. Microbiol.* 14, 1189877
21. Choe, D. *et al.* (2023) Revealing causes for false-positive and false-negative calling of gene essentiality in *Escherichia coli* using transposon insertion sequencing. *mSystems* 8, e0089622
22. Commichau, F.M. *et al.* (2013) Essential genes in *Bacillus subtilis*: a re-evaluation after ten years. *Mol. BioSyst.* 9, 1068–1075
23. Goodall, E.C.A. *et al.* (2018) The essential genome of *Escherichia coli* K-12. *MBio* 9, e02096-17
24. Choe, D. *et al.* (2019) Adaptive laboratory evolution of a genome-reduced *Escherichia coli*. *Nat. Commun.* 10, 935
25. Dervyn, E. *et al.* (2023) Greedy reduction of *Bacillus subtilis* genome yields emergent phenotypes of high resistance to a DNA damaging agent and low evolvability. *Nucleic Acids Res.* 51, 2974–2992
26. Vernyik, V. *et al.* (2020) Exploring the fitness benefits of genome reduction in *Escherichia coli* by a selection-driven approach. *Sci. Rep.* 10, 7345
27. Ma, S. *et al.* (2022) Random genome reduction coupled with polyhydroxybutyrate biosynthesis to facilitate its accumulation in *Escherichia coli*. *Front. Bioeng. Biotechnol.* 10, 978211
28. Pósfai, G. *et al.* (2006) Emergent properties of reduced-genome *Escherichia coli*. *Science* 312, 1044–1046
29. Wang, K. *et al.* (2016) Defining synonymous codon compression schemes by genome recoding. *Nature* 539, 59–64
30. Sandberg, T.E. *et al.* (2019) The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab. Eng.* 56, 1–16
31. Nishimura, I. *et al.* (2017) Coordinated changes in mutation and growth rates induced by genome reduction. *MBio* 8, e00676-17
32. Moger-Reischer, R.Z. *et al.* (2023) Evolution of a minimal cell. *Nature* 620, 122–127
33. Sandberg, T.E. *et al.* (2023) Adaptive evolution of a minimal organism with a synthetic genome. *iScience* 26, 107500
34. Hossain, T. *et al.* (2021) Antibiotic tolerance, persistence, and resistance of the evolved minimal cell, *Mycoplasma mycoides* JCVI-Syn3B. *iScience* 24, 102391
35. Pfeifer, E. *et al.* (2017) Adaptive laboratory evolution of *Corynebacterium glutamicum* towards higher growth rates on glucose minimal medium. *Sci. Rep.* 7, 16780
36. Choe, D. *et al.* (2022) Synthetic 3′-UTR valves for optimal metabolic flux control in *Escherichia coli*. *Nucleic Acids Res.* 50, 4171–4186
37. Liang, P. *et al.* (2020) Deletion of genomic islands in the *Pseudomonas putida* KT2440 genome can create an optimal chassis for synthetic biology applications. *Microb. Cell Factories* 19, 70
38. van Tilburg, A.Y. *et al.* (2020) Mini*Bacillus* PG10 as a convenient and effective production host for lantibiotics. *ACS Synth. Biol.* 9, 1833–1842
39. Bu, Q.-T. *et al.* (2019) Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb. Cell Factories* 18, 16
40. Myronovskyi, M. *et al.* (2018) Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* 49, 316–324
41. Nikolaidis, M. *et al.* (2023) A panoramic view of the genomic landscape of the genus *Streptomyces*. *Microb. Genom.* 9, mgen001028
42. Ahmed, Y. *et al.* (2020) Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb. Cell Factories* 19, 5
43. Peng, Q. *et al.* (2018) Engineered strains for optimal identification and expression of cryptic biosynthetic gene clusters. *Front. Microbiol.* 9, 3042
44. Walter, T. *et al.* (2020) Fermentative N-methylanthranilate production by engineered *Corynebacterium glutamicum*. *Microorganisms* 8, 866
45. Fan, X. *et al.* (2020) Genome reduction enhances production of polyhydroxyalkanoate and alginate oligosaccharide in *Pseudomonas mendocina*. *Int. J. Biol. Macromol.* 163, 2023–2031
46. Zhu, D. *et al.* (2017) Enhanced heterologous protein productivity by genome reduction in *Lactococcus lactis* NZ9000. *Microb. Cell Factories* 16, 1
47. Zhang, F. *et al.* (2020) Engineering of a genome-reduced strain *Bacillus amyloliquefaciens* for enhancing surfactin production. *Microb. Cell Factories* 19, 223
48. Martínez-García, E. *et al.* (2014) *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microb. Cell Factories* 13, 159
49. Nguyen, T.T. *et al.* (2023) Exploring the potential of a genome-reduced *Escherichia coli* strain for plasmid DNA production. *Int. J. Mol. Sci.* 24, 11749
50. Ohlson, J. (2020) Plasmid manufacture is the bottleneck of the genetic medicine revolution. *Drug Discov. Today* 25, 1891–1893
51. Park, M.K. *et al.* (2014) Enhancing recombinant protein production with an *Escherichia coli* host strain lacking insertion sequences. *Appl. Microbiol. Biotechnol.* 98, 6701–6713
52. Zwiener, T. *et al.* (2021) Towards a 'chassis' for bacterial magnetosome biosynthesis: genome streamlining of *Magnetospirillum gryphiswaldense* by multiple deletions. *Microb. Cell Factories* 20, 35
53. Consuegra, J. *et al.* (2021) Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. *Nat. Commun.* 12, 980
54. Xu, T. *et al.* (2022) Deficiency of exopolysaccharides and O-antigen makes *Halomonas bluephagenesis* self-flocculating and amenable to electrotransformation. *Commun. Biol.* 5, 623
55. Lajoie, M.J. *et al.* (2013) Genomically recoded organisms expand biological functions. *Science* 342, 357–360
56. Sisila, V. *et al.* (2023) Building biomaterials through genetic code expansion. *Trends Biotechnol.* 41, 165–183
57. Beattie, A.T. *et al.* (2023) Quintuply orthogonal pyrrolysyl-tRNA synthetase/tRNAPyl pairs. *Nat. Chem.* 15, 948–959
58. Dunkelmann, D.L. *et al.* (2024) Adding α,α-disubstituted and β-linked monomers to the genetic code of an organism. *Nature* 625, 603–610
59. Venter, J.C. *et al.* (2022) Synthetic chromosomes, genomes, viruses, and cells. *Cell* 185, 2708–2724
60. Antczak, M. *et al.* (2019) Environmental conditions shape the nature of a minimal bacterial genome. *Nat. Commun.* 10, 3100
61. Pelletier, J.F. *et al.* (2021) Genetic requirements for cell division in a genomically minimal cell. *Cell* 184, 2430–2440.e16
62. Jores, J. *et al.* (2019) Removal of a subset of non-essential genes fully attenuates a highly virulent *Mycoplasma* strain. *Front. Microbiol.* 10, 664
63. Nishiumi, F. *et al.* (2021) Blockade of endoplasmic reticulum stress-induced cell death by *Ureaplasma parvum* vacuolating factor. *Cell. Microbiol.* 23, e13392
64. Breuer, M. *et al.* (2019) Essential metabolism for a minimal cell. *Elife* 8, e36842
65. Thornburg, Z.R. *et al.* (2022) Fundamental behaviors emerge from simulations of a living minimal cell. *Cell* 185, 345–360.e28

66. Tellechea-Luzardo, J. *et al.* (2022) Fast biofoundries: coping with the challenges of biomanufacturing. *Trends Biotechnol.* 40, 831–842
67. Zürcher, J.F. *et al.* (2023) Continuous synthesis of *E. coli* genome sections and Mb-scale human DNA assembly. *Nature* 619, 555–562
68. Wang, K. *et al.* (2019) Programmed chromosome fission and fusion enable precise large-scale genome rearrangement and assembly. *Science* 365, 922–926
69. Yoneji, T. *et al.* (2021) Grand scale genome manipulation via chromosome swapping in *Escherichia coli* programmed by three one megabase chromosomes. *Nucleic Acids Res.* 49, 8407–8418
70. Sastry, A.V. *et al.* (2019) The *Escherichia coli* transcriptome mostly consists of independently regulated modules. *Nat. Commun.* 10, 5536
71. Rychel, K. *et al.* (2021) iModulonDB: a knowledgebase of microbial transcriptional regulation derived from machine learning. *Nucleic Acids Res.* 49, D112–D120
72. Hasan, M.A. and Lonardi, S. (2020) DeeplyEssential: a deep neural network for predicting essential genes in microbes. *BMC Bioinformatics* 21, 367
73. Peters, J.M. *et al.* (2016) A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* 165, 1493–1506
74. Macklin, D.N. *et al.* (2020) Simultaneous cross-evaluation of heterogeneous *E. coli* datasets via mechanistic simulation. *Science* 369, eaav3751
75. Stevens, J.A. *et al.* (2023) Molecular dynamics simulation of an entire cell. *Front. Chem.* 11, 1106495
76. Yu, B.J. *et al.* (2002) Minimization of the *Escherichia coli* genome using a Tn5-targeted Cre/loxP excision system. *Nat. Biotechnol.* 20, 1018–1023
77. Umenhoffer, K. *et al.* (2010) Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microb. Cell Factories* 9, 38
78. Kolisnychenko, V. *et al.* (2002) Engineering a reduced *Escherichia coli* genome. *Genome Res.* 12, 640–647
79. Lee, J.H. *et al.* (2009) Metabolic engineering of a reduced-genome strain of *Escherichia coli* for L-threonine production. *Microb. Cell Factories* 8, 2
80. Karcagi, I. *et al.* (2016) Indispensability of horizontally transferred genes and its impact on bacterial genome streamlining. *Mol. Biol. Evol.* 33, 1257–1269
81. Mizoguchi, H. *et al.* (2007) *Escherichia coli* minimum genome factory. *Biotechnol. Appl. Biochem.* 46, 157–167
82. Mizoguchi, H. *et al.* (2008) Superpositioning of deletions promotes growth of *Escherichia coli* with a reduced genome. *DNA Res.* 15, 277–284
83. Matteau, D. *et al.* (2023) Complete sequence of the genome-reduced *Escherichia coli* DGF-298. *Microbiol. Resour. Announc.* 12, e0066523
84. Hirokawa, Y. *et al.* (2013) Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *J. Biosci. Bioeng.* 116, 52–58
85. Li, Y. *et al.* (2016) Characterization of genome-reduced *Bacillus subtilis* strains and their application for the production of guanosine and thymidine. *Microb. Cell Factories* 15, 94
86. Ara, K. *et al.* (2007) *Bacillus* minimum genome factory: effective utilization of microbial genome information. *Biotechnol. Appl. Biochem.* 46, 169–178
87. Schilling, T. *et al.* (2023) Let there be light: genome reduction enables *Bacillus subtilis* to produce disulfide-bonded *Gaussia* luciferase. *ACS Synth. Biol.* 12, 3656–3668
88. Baumgart, M. *et al.* (2013) Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl. Environ. Microbiol.* 79, 6006–6015
89. Baumgart, M. *et al.* (2018) *Corynebacterium glutamicum* chassis C1*: building and testing a novel platform host for synthetic biology and industrial biotechnology. *ACS Synth. Biol.* 7, 132–144
90. Ferrer, L. *et al.* (2022) Fermentative indole production via bacterial tryptophan synthase alpha subunit and plant indole-3-glycerol phosphate lyase enzymes. *J. Agric. Food Chem.* 70, 5634–5645
91. Qiao, W. *et al.* (2021) Genomic features and construction of streamlined genome chassis of nisin Z producer *Lactococcus lactis* N8. *Microorganisms* 10, 47
92. Shen, X. *et al.* (2017) Developing genome-reduced *Pseudomonas chlororaphis* strains for the production of secondary metabolites. *BMC Genomics* 18, 715
93. Liu, H. *et al.* (2022) Enhanced production of polyhydroxyalkanoates in *Pseudomonas putida* KT2440 by a combination of genome streamlining and promoter engineering. *Int. J. Biol. Macromol.* 209, 117–124
94. Wynands, B. *et al.* (2019) Streamlined *Pseudomonas taiwanensis* VLB120 chassis strains with improved bioprocess features. *ACS Synth. Biol.* 8, 2036–2050
95. Liu, J. *et al.* (2021) Rational construction of genome-reduced *Burkholderiales* chassis facilitates efficient heterologous production of natural products from proteobacteria. *Nat. Commun.* 12, 4347
96. Komatsu, M. *et al.* (2010) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. USA* 107, 2646–2651
97. Komatsu, M. *et al.* (2013) Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* 2, 384–396
98. Zhou, M. *et al.* (2012) Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of *Streptomyces coelicolor*. *FEMS Microbiol. Lett.* 333, 169–179