Trends in Biotechnology

CellPress
REVIEWS

Feature Review

Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering

Kyeong Rok Choi,[1] Woo Dae Jang,[1] Dongsoo Yang,[1] Jae Sung Cho,[1] Dahyeon Park,[1] and Sang Yup Lee[1,2,3,*]

Metabolic engineering allows development of microbial strains efficiently producing chemicals and materials, but it requires much time, effort, and cost to make the strains industrially competitive. Systems metabolic engineering, which integrates tools and strategies of systems biology, synthetic biology, and evolutionary engineering with traditional metabolic engineering, has recently been used to facilitate development of high-performance strains. The past decade has witnessed this interdisciplinary strategy continuously being improved toward the development of industrially competitive overproducer strains. In this article, current trends in systems metabolic engineering including tools and strategies are reviewed, focusing on recent developments in selection of host strains, metabolic pathway reconstruction, tolerance enhancement, and metabolic flux optimization. Also, future challenges and prospects are discussed.

Highlights

Systems metabolic engineering, which integrated systems biology, synthetic biology, and evolutionary engineering with traditional metabolic engineering, is facilitating the development of high performance strains.

More diverse microorganisms are being used as production host strains, supported by the new genetic tools and strategies.

Recent advances in biosynthetic/semi-synthetic design strategies are expanding the portfolio of products that can be produced biologically.

Evolutionary engineering tools and strategies are facilitating the improvement of strain and enzyme performances.

Advances in tools and strategies of omics, *in silico* metabolic simulation, genetic and genomic engineering, and high-throughput screening are accelerating optimization of metabolic fluxes for the enhanced production of target bioproducts.

Emergence of Systems Metabolic Engineering

Since its formal recognition in 1991 [1], **metabolic engineering** (see Glossary) has demonstrated its capability in developing microbial strains for the production of hundreds of different chemicals from renewable raw materials [2–4]. The number of bioproducts actually reaching production on industrial scales, however, is small, mainly due to the high overall production costs comprising costs of raw materials, production (fermentation), and recovery/purification. Although tremendous advances have been made in the field of metabolic engineering, it is still necessary to improve strain performance to make bioprocesses competitive with the corresponding petrochemical processes. Also, metabolic engineering projects at their initial stages often fail to consider the practical issues arising in industrial-scale production [5]. Furthermore, repeated trial-and-error type development cycles to improve strain performance, which results from incomplete understanding of the genetic and metabolic mechanisms of the host organisms in laboratory, pilot, demonstration, and actual production scales, further increases the amount of manpower (e.g., 50–300 person–years) and investments (hundreds of millions of US dollars) required to develop strains [6].

The emergence of **systems metabolic engineering** – which integrates **systems biology**, **synthetic biology**, and **evolutionary engineering** with traditional metabolic engineering – has expedited the development of industrially competitive strains, as exemplified by initial works on developing *Escherichia coli* strains to overproduce L-valine [7] and L-threonine [8] in 10 person–years. Systems metabolic engineering also considers the midstream (fermentation) and downstream (recovery and purification) processes during the upstream (strain

[1]Metabolic and Biomolecular Engineering National Research Laboratory, Systems Metabolic Engineering and Systems Healthcare Cross Generation Collaborative Laboratory, Department of Chemical and Biomolecular Engineering (BK21 Plus Program), Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea
[2]BioProcess Engineering Research

development) process in designing a new project (Figure 1, Key Figure), facilitating the scale-up production of the target bioproduct at the end of the project [5]. Although systems metabolic engineering is not yet a universally used term, it captures all the advances that have been made to progress the field of metabolic engineering.

Recent advances in the fields of omics, genome-scale metabolic simulation, genetic engineering, and evolutionary engineering have expanded the tools and strategies of systems metabolic engineering, enabling increasingly massive, parallel, and systematic engineering of strains toward achieving their outstanding performances. Previous reviews on systems metabolic engineering have focused on either describing systems metabolic engineering strategies in the order of developing overproducer strains with several specific examples [5,9], or summarizing tools and strategies of systems biology, synthetic biology, and evolutionary engineering useful for conducting systems metabolic engineering research [2,10–12]. This article combines recent trends in systems metabolic engineering with tools and strategies useful at each key step of systems metabolic engineering research [5] and introduces new tools and strategies that are expected to further expand the power of systems metabolic engineering. In addition, current limitations and suggestions for future advancements are discussed. Two recent reviews [2,5] and references cited therein can provide the detailed information and knowledge on the tools and strategies of systems metabolic engineering previously developed.

Project Design

In response to the market and societal demands, biobased production of various chemicals ranging from bulk to specialty chemicals has attracted much attention (Figure 1A). For the biobased production of bulk chemicals – from biofuels and solvents [13–18], to organic acids [19,20], to polymers and their monomers [19,21–24], to food and animal feedstock supplements [22,25,26] – consumed in large quantities at low prices, it is critical to achieve high titer ($> \sim 100$ g/L and preferably even $> \sim 200$ g/L) of the product with high yield and productivity. Titer (product concentration), yield, and productivity are important performance metrics to assess the competitiveness of a bioprocess. The titer is the final concentration of the product at the end of fermentation, while the yield is gram (or mole) of product formed per gram (or mole) of substrate (usually carbon substrate) consumed. Productivity can be defined as either specific productivity or volumetric productivity. The specific productivity refers to the amount (gram or mole) of product produced per cell per unit time, while the volumetric productivity refers to the amount (gram or mole) of product produced per volume per unit time. Achieving high values for all of these performance metrics is important, although one parameter might be emphasized more than the others depending on the product type and overall bioprocess economics. For example, a high titer is usually helpful in reducing the costs and simplifying technical difficulties in separation and purification processes. When producing bulk chemicals, achieving a high yield becomes important as the major cost in the overall production costs comes from carbon substrate. Productivity is tightly linked to the overall operation cost of the bioprocess, as it determines the sizes of the fermentor and other operating units in the whole bioprocess, which in turn affects annual equipment depreciation costs as well as initial direct fixed capital costs. Also, elimination of byproduct formation is important to reduce recovery and purification costs, while further increasing the product yield [27].

As consumers are increasingly advocating the use of environmentally friendly products, while governments worldwide are enforcing some regulations preferring biobased products (e.g., banning of one-time use petroleum-driven plastics), biobased chemicals and materials will be increasingly adopted if they possess the same or similar properties, even at slightly higher production costs. Obviously, high-value products including biobased polymers with special

Center, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea
[3]BioInformatics Research Center, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

*Correspondence: leesy@kaist.ac.kr (S.Y. Lee).

properties and functions [e.g., poly(lactic-co-glycolate) [21] and bacterial cellulose [23]], pharmaceuticals [28–31], neutraceuticals [32–34], fragrance and flavoring chemicals [35–37], functional proteins [38,39], and other specialty chemicals and materials have higher potential to more easily penetrate into the market. These high-value products are attractive targets for contemporary metabolic engineers. Although the current production titers of such bioproducts are not as high as those of biobased bulk chemicals, the high prices due to scarcity from natural sources [28,30,37,40], ethical/political concerns on the current production [30,34], or absence/inefficiency of the chemical synthesis methods [30–32,38] contribute to the competitiveness of their biobased production.

In addition to the production of pharmaceuticals and nutraceuticals, engineered microorganisms themselves can also be used as live therapeutics. With the recent recognition of the profound impact of the microbiome on human health [41], engineering commensals is gaining significant attention to achieve desired health effects. Metabolically engineered commensals have been exploited for gut infection prevention [42], live diagnostics [43], live therapeutics [44], drug delivery [45], and real-time diagnostics [46].

Degradation of environmental pollutants (and xenobiotics) is another important goal of systems metabolic engineering (Figure 1A) [47–49]. While some strains are engineered toward efficient degradation of target compounds [48,49], other strains are also able to produce valuable compounds by consuming the pollutants [47]. This fascinating branch of research promising the detoxification of environmental pollutants such as micro/nanoplastics [50] and spilled oil [51] has recently been on the rise.

A majority of the metabolic engineering projects have focused on the *de novo* (i.e., from renewable carbon sources, usually from glucose) biosynthesis of target bioproducts. Biotransformation (bioconversion), either single or multiple steps, from precursor materials, however, have also been frequently devised when the upstream biosynthetic pathways (i.e., from the renewable carbon sources to certain points in the middle of the biosynthetic pathways) are not complete or inefficient while the precursors are available at low prices (Figure 1A) [52]. In addition, integrating traditional bioproduction strategies to versatile and efficient (and more preferably, environmentally friendly) chemical conversion strategies have further diversified the portfolio of biobased products (Figure 1A) [29,53,54].

The choice of renewable carbon sources is also expanding. Although the use of typical sugars, such as glucose and sucrose in the form of hydrolyzed starch and raw sugar, respectively, is most common in producing target chemicals, the use of lignocellulosics has also been actively examined in recent decades (Figure 1A) [13,55,56]. **C1 chemicals** are recently emerging carbon sources (Figure 1A) [57]. Although current bioproduct formation via C1 carbon assimilation/utilization is not as efficient as that using conventional carbon sources, the use of methane [58], methanol [59], formic acid [60], and carbon dioxide [61] is on the rise. Strategies to design systems metabolic engineering projects will continue to evolve to meet the rapidly changing societal, economic, and political situations of the world.

Selection of a Host Strain

Because their metabolism and physiology are best understood, with correspondingly well-developed engineering tools, model microorganisms such as *E. coli* and *Saccharomyces cerevisiae* are still the most widely used organisms for the biobased production of diverse products [4,13,32,33,62]. However, some chemicals can be more efficiently produced by natural overproducers, such as *Clostridium* sp. for acetone and butanol [16,63],

Glossary

Adaptive laboratory evolution (ALE): a process of adapting microorganisms in specified cultivation conditions with a selection pressure for a prolonged period, often hundreds or thousands of generations, to obtain strains possessing preferred phenotypes (e.g., improved growth or tolerance to toxic chemicals).

Bio-big data: whole datasets obtained from biological and biotechnological systems. Examples include conventional omics data from high-throughput experiments, *in silico* simulation results (e.g., fluxome), spatiotemporal operation data obtained from fermentation, recovery, and purification processes.

C1 chemicals: single carbon atom chemicals, such as carbon dioxide, formic acid, methanol, and methane. Many C1 chemicals, such as carbon dioxide, chemicals catalytically converted from carbon dioxide (e.g., formic acid), and methane from biogas, are considered renewable. In contrast, methane from shale gas and carbon monoxide from fossil resources are not considered renewable, yet largely available in certain regions at low costs.

CRISPR/Cas: a recently exploited genetic engineering system derived from bacterial adaptive immune system called clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system.

Directed evolution: a process that mimics the mechanisms of evolution to more quickly obtain biological parts (e.g., genes, proteins, and strains) with desired characteristics, which includes steps for selecting desirable phenotypes from the libraries generated during the process.

Evolutionary engineering: a discipline of engineering that applies evolutionary power, rather than rational manipulation, to achieve certain phenotypes.

Metabolic engineering: a discipline of engineering that genetically modulates living cells or organisms to overproduce desired products.

Natural products: compounds found in nature in general, but are often used in narrower meaning as

Corynebacterium sp. for amino acids [22,26,64], *Mannheimia succiniciproducens* for succinic acid [65], *Rhodococcus opacus* and *Yarrowia lipolytica* for lipids, fatty acids, and derivatives [18,66,67], and actinomycetes for antibiotics and polyketides (Figure 1B) [68]. Thus, the metabolic pathways of these microorganisms can be tweaked to produce chemicals sharing the same biosynthetic pathways with their naturally overproduced metabolites [69,70]. Efficient engineering tools for these important platform strains are continuously being upgraded to facilitate metabolic engineering [68,69,71].

Accompanied by the development of genetic engineering tools readily adaptable to diverse organisms, less explored organisms with attractive properties have been actively exploited as host strains as well (Figure 1B). In response to public concerns about using possibly unsafe microorganisms to produce chemicals and materials that humans eat or use, organisms generally recognized as safe (GRAS) are increasingly being considered as the host strains to overproduce bioproducts for direct human uses. In addition to the well-studied GRAS strains (e.g., *S. cerevisiae* [28–30], *Bacillus subtilis* [72], and *Corynebacterium glutamicum* [22,26]), lactic acid bacteria [73], *Pseudomonas putida* KT2440 [74], and other less explored GRAS strains are receiving attention as potential host strains. Moreover, cyanobacteria, microalgae, and methanotrophic bacteria are increasingly spotlighted as engineering hosts consuming C1 chemicals as raw materials [58,75]. Thermophilic bacteria have also been considered as hosts for metabolic engineering as their fermentation processes at elevated temperatures can reduce risk of contamination – by both microorganisms and phages – and are more compatible with various existing industrial chemical processes [76]. Similarly, halophiles allow open-field fermentation based on seawater, minimizing the concern of contamination while saving fresh water for other purposes [77].

As our capabilities to engineer organisms advance, engineering hosts are no longer limited to microorganisms and are being expanded to higher eukaryotes, including plant cells [78] and mammalian cells [79], and even intact living multicellular organisms, such as living insects [80] and plants [81]. Further exploration toward diverse organisms with attractive advantages for biobased production of certain products will likely be increasingly pursued.

Metabolic Pathway Reconstruction

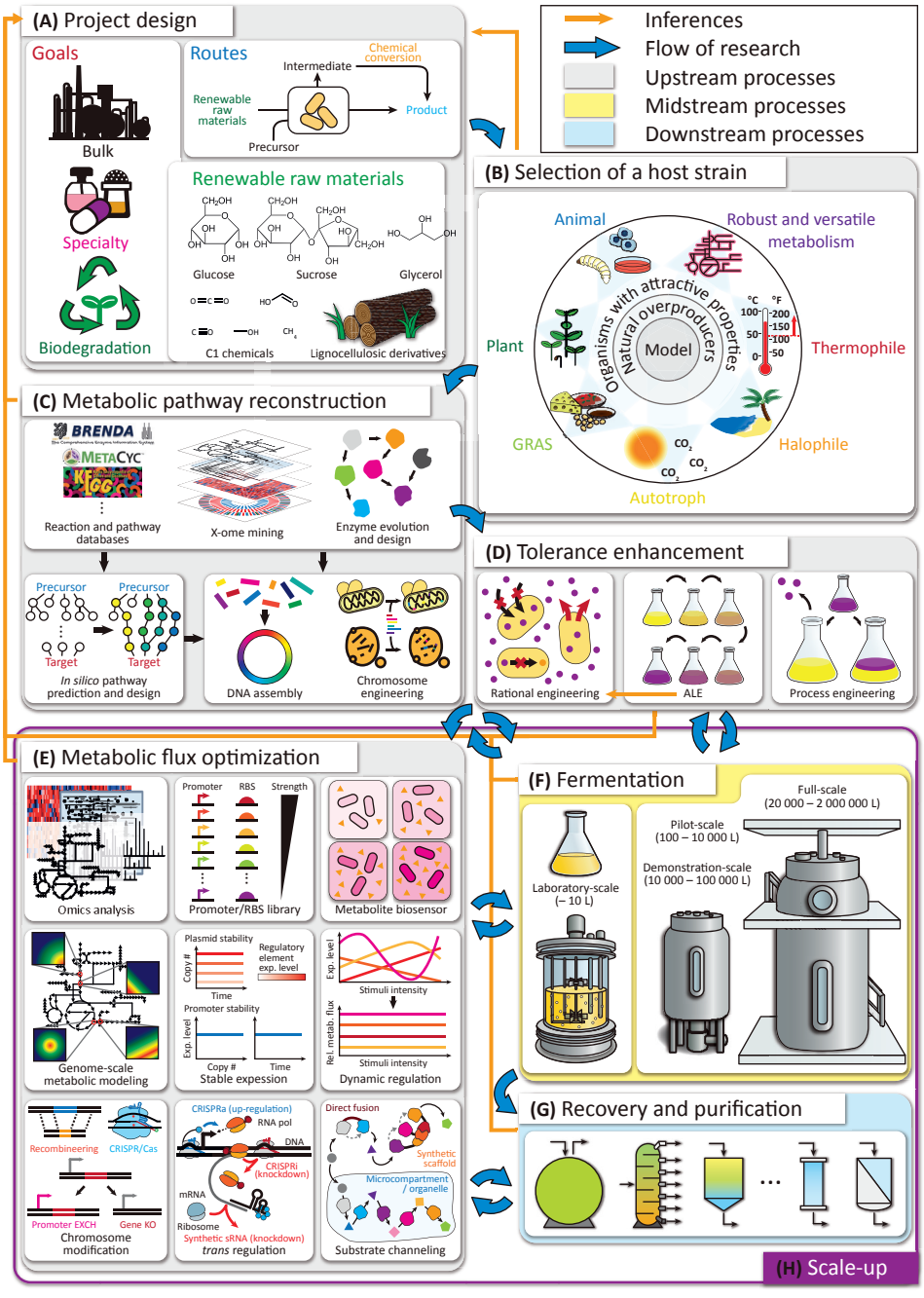
In contrast to the traditional focuses on optimizing endogenous pathways or reconstituting heterologous yet natural pathways to produce canonical metabolites [7,22], recent advances in synthetic biology and computational biology have enabled designing novel and specific metabolic pathways for desired chemicals [29,52,82,83]. In addition to the typical strategies of reconstructing heterologous metabolic pathways through the combination of known metabolic reactions deposited in metabolic reaction/pathway databases such as KEGG (http://www.genome.jp/kegg/), MetaCyc (https://metacyc.org/), and BRENDA (https://www.brenda-enzymes.org/), the use of **reaction rules** generated based on the chemical structures of substrates and products in enzymatic reactions can help streamline the design processes (Figure 1C and Table 1) [84–86]. Exploiting substrate promiscuity of enzymes further accelerates designing novel and efficient biosynthetic pathways toward the target compounds [82,83]. Moreover, directed evolution (Box 1) and *de novo* design of enzymes have contributed to expanding the spectrum of natural and non-natural chemicals that can be produced biologically (Figure 1C and Table 2) [87–92].

There have been continued efforts to discover new enzymes catalyzing new reactions through X-ome mining with the help of screening tools while activating silent biosynthetic genes and clusters (Figure 1C) [87,93–95]. The integration of chemical approaches with biosynthetic approaches also broadens the profile of producible chemicals further [29,53]. Artificial

functional secondary metabolites found in microorganisms and plants.
Reaction rules: descriptors for the change of bonding patterns in the reactants during its transformation to the products.
Recombineering: a strategy in genetic manipulation that uses bacteriophage-derived recombinases to facilitate genetic engineering based on sequence homologies.
RNA interference (RNAi): an antisense RNA-based gene expression knockdown tool where a double-stranded RNA is cleaved by Dicer proteins to yield small guide RNAs. These small guide RNAs then bind to target mRNAs with the help of Argonaute proteins to inhibit translation and/or to degrade transcripts.
Synthetic biology: a discipline of biology and engineering that aims to design, construct, and modify biological parts and modules ranging from genes to enzymes to pathways to organisms, pursuing their standardization, modularization, and systems-level analyses and uses.
Small regulatory RNA: a *trans*-acting short RNA which binds to a target mRNA by base pairing, generally leading to alteration in the target gene expression level.
Systems biology: a discipline of biology that tries to analyze, understand, model, and simulate complicated biological systems at a systems level with the help of mathematical and computational methods.
Systems metabolic engineering: an interdisciplinary research strategy integrating systems biology, synthetic biology, and evolutionary engineering with classical metabolic engineering.

Key Figure

Scheme of Systems Metabolic Engineering Research with Current Tools and Strategies



(See figure legend on the bottom of the next page.)

intelligence is anticipated to play increasingly important roles in predicting feasible chemical synthesis routes [96] for the production of desired chemicals and materials.

Advances in genetic engineering tools and strategies have accelerated successful introduction of the designed metabolic pathways into actual producer strains. Novel and efficient DNA assembly tools, such as BioBrick assembly [97], Gibson assembly [98], Golden Gate assembly [99], ligase cycling reaction [100], single strand assembly [101], transformation-associated recombination (TAR) cloning [102], and uracil specific excision reagent (USER) cloning [103], have facilitated the assembly of multicomponent and large-sized gene clusters and consequent expression of the assembled metabolic pathway genes based on plasmids (Figure 1C and Table 2). In addition, advances in oligonucleotide and gene synthesis technologies at decreasing synthesis costs [104] have been making it more feasible to construct metabolic pathway genes optimal for their expression in each host strain (e.g., codon optimized) and to test many combinatorially rearranged gene clusters and expression modules [105]. Indeed, such new technologies are accelerating the design–build–test cycles of systems metabolic engineering research.

Despite the convenience of reconstructing and expressing the metabolic pathway gene clusters based on plasmids, concerns on plasmid instability and copy number fluctuation undesirable in an industrial setting have been urging researchers to develop methods for stable chromosomal integration (Figure 1C). As the traditional homologous recombination systems using counter selection markers (e.g., *pyrF*, *sacB*, and *upp* genes) are often time-consuming and laborious in introducing genes of interest to target genomic loci, site-specific recombination systems [106], and transposon-based random insertion systems [107] are often used to facilitate these procedures (Table 2). However, the choice of gene integration sites is limited compared with the homologous recombination system due to the intrinsic nature of the site-specific recombinases and transposons (i.e., too site specific and too random, respectively). Recently, **recombineering** technology was demonstrated to expand the integration size limit with the help of a donor plasmid system [108], while overcoming the aforementioned limitations of the site-specific recombination and transposon-based systems. **CRISPR/Cas**-induced double-/single-strand breaks followed by either homologous recombination or recombineering is also frequently employed for the chromosomal integration of the genes (Table 2) [69,109,110].

The tools and strategies mentioned above have facilitated the reconstitution of long and complicated biosynthetic pathways for the heterologous production of complex chemicals, especially the **natural products**. For example, complete opioids biosynthetic pathways consisting of 21 and 23 enzymatic reactions from plants, mammals, yeasts, and bacteria were integrated into the

Figure 1. Systems metabolic engineering streamlines developing industrial overproducers by considering midstream (fermentation) and downstream (recovery and purification) processes during upstream process (strain development) in designing new projects. (A) Projects are designed with diverse goals, production routes, and renewable raw materials. (B) The portfolio of host strain is expanding beyond model organisms (Model) and natural overproducers to include diverse organisms with attractive properties. (C) Reaction/pathway databases are used to design metabolic pathways by manual curation or *in silico* pathway prediction/design strategies. X-ome mining, directed evolution, and synthetic design of enzymes expand reaction pools. Emerging DNA assembly and chromosome engineering tools are expediting actual construction of the designed pathways in production hosts. (D) Tolerance against target chemicals is enhanced rationally or through adaptive laboratory evolution (ALE). Tolerant strains isolated from ALE may provide clues to further enhance tolerance rationally. Process engineering approaches can also reduce effective concentration of toxic chemicals. (E) Systems biology, synthetic biology, and evolutionary engineering tools accelerate metabolic flux optimization to maximize target chemical production. (F) Fermentation goes in parallel with strain development (C–E), providing useful feedback. (G) Recovery and purification of products are critical factors that must be considered to develop overproducer strains. (H) Metabolic fluxes are iteratively optimized based on the fermentation and recovery/purification performances to facilitate scale-up from laboratory scale to full-scale. Abbreviations: EXCH, exchange; exp., expression; GRAS, generally recognized as safe; rel. metab. flux, relative metabolic flux with respect to those in the whole biosynthetic pathway; KO, knockout; RBS, ribosome binding site.

Table 1. Systems Biology Tools and Strategies for Systems Metabolic Engineering

Categories	Tool and strategies	Description	Refs
Omics/multiomics	Genomics	Analysis of genetic structure of host organisms and identification of genetic changes associated with desired cell properties by whole-genome sequencing.	[88]
	Transcriptomics	Analysis of transcriptome to understand the mechanism of cellular metabolism on the level of gene expression patterns and identify transcriptional regulators for the overproduction of target products.	[69,129]
	Proteomics	Understanding cellular metabolism based on the protein expression profiles and revealing the role of post-translational modifications.	[133]
	Metabolomics	Discovering changes in metabolism upon engineering and identifying rate-limiting steps for production of chemicals to find additional engineering targets.	[134]
	Fluxomics	Quantification of intracellular metabolic fluxes by ^{13}C metabolic flux analysis enabling identification of potential targets for engineering targets.	[64,136]
	Genomics + transcriptomics	Identification of changes in genetic composition and global gene expression during adaptive laboratory evolution to understand the mechanism of acquiring desired traits and identify further engineering targets.	[182]
	Genomics + transcriptomics + phenomics	Characterization of strain-specific genetic and physiological differences for comprehensive understanding of phenotypes.	[142]
	Proteomics + metabolomics	Combined analysis of proteome and metabolome to understand the changes in cell physiology upon different experimental conditions and identify promising metabolic pathways and engineering targets.	[183]
X-ome mining	antiSMASH	Computational resource for the genome mining of biosynthetic gene clusters from bacteria, fungi, and plants.	[93]
	PRISM	Computational approach for identifying biosynthetic genes and predicting chemical structures of nonribosomal peptides and type I and II polyketides based on the identified biosynthetic genes.	[94]
	RODEO	Genome mining tool for identifying biosynthetic gene clusters and predicting ribosomally synthesized and post-translationally modified peptides by combining hidden-Markov-model-based analysis and heuristic scoring.	[184]
	GE-PrISM	Strategy to discover expressed gene clusters using proteomics-based approach.	[95]
Pathway prediction and design	BNICE	Computational tool to discover novel metabolic pathways using generalized reaction rules.	[185]
	rePrime/novoStoic	Computational framework for metabolic pathway prediction including generation of reaction rules (rePrime) and pathway design algorithm (novoStoic).	[86]
	RetroPath	Computational tool for pathway prediction based on generalized reaction rules and retrosynthesis scheme. The effectiveness of the tool was experimentally verified by constructing a pinocembrin overproducer strain.	[83,186]
	RetroRules	Reaction rule database extracted from public databases of metabolic pathway prediction.	[85]
Genome-scale metabolic modeling	AGORA	Resource of semiautomatically generated genome-scale metabolic models for 773 human gut bacteria, which is compatible with human genome-scale metabolic models for analysis of host–microbiome interactions.	[138]
	merlin	Computational framework for the reconstruction of genome-scale metabolic models, which provides graphical interface for manual curation with visualized metabolic pathways.	[187]
	RAVEN	Computational framework for homology-based genome-scale metabolic model reconstruction using template models. RAVEN provides several tools for model curation and simulation.	[188]
Omics-integrated genome-scale metabolic modeling	ME-model	Computational framework for constructing and simulating genome-scale models of metabolism and gene expression, enabling computation of optimal proteome abundances for a given condition.	[143]
	ssbio	Computational framework for genome-scale metabolic models integrated with protein structure information from useful third-party structural bioinformatics tools such as DSSP, I-TASSER, and FATCAT.	[189]
	Thermodynamics-based flux analysis (TFA)	Thermodynamics-based framework integrating the metabolome data to assign the reaction directionality and thermodynamically constrain genome-scale metabolic models.	[190]

Table 1. (continued)

Categories	Tool and strategies	Description	Refs
Gene knockout and overexpression target identification	CAMEO	Python-based library for *in silico* metabolic modeling containing state-of-the-art algorithms for identifying gene knockout and overexpression targets (e.g., FSEOF, MOMA, and ROOM).	[191]
	COBRA	Computational framework for quantitative prediction of biochemically feasible phenotypic states using constraint-based modeling.	[192]
Enzyme engineering	Allosteric regulation engineering	Engineering of allosteric binding sites to deregulate native feedback inhibitions.	[193]
	Post-translational modification engineering	Engineering of the regulation of metabolic pathways and cell phenotypes through the modification of post-translational modification systems.	[194]

chromosome of *S. cerevisiae*, successfully producing the opioids thebaine and hydrocodone, respectively [28]. In another study, the opioid biosynthetic pathways were allocated to different *E. coli* strains, and stepwise fermentation of these strains allowed production of the opioids from simple carbon sources [111]. Similarly, allocating the biosynthetic pathway of oxygenated taxanes, precursors of an anticancer agent paclitaxel (also known as the brand name Taxol), into *E. coli* and *S. cerevisiae* and subsequent coculture of the two microbial consortia was demonstrated to produce oxygenated taxanes [30]. It will be interesting to see if bioprocesses using the mixed engineered strains developed by such 'divide-and-conquer' type approaches can actually be used in large industrial-scale fermentation for the production of complex molecules. By contrast, production of the antibiotic erythromycin by engineered *E. coli* harboring a reconstructed 50-kb-long biosynthetic gene cluster after testing various genetic constructs and plasmid combinations has been reported [112].

Increasing Host Tolerance against Target Chemicals

The tolerance of host strains against the target chemicals can be rationally enhanced if the molecular mechanisms of the toxicity are understood. For example, reducing the entry of toxic chemicals to the cells [113], preventing toxic conversion and incorporation to cell biomass [114], and active exporting of the toxic compounds to extracellular spaces [7,8,115,116] have successfully enhanced the tolerance to toxic chemicals (Figure 1D). Readers are encouraged to read a thorough review on the rational tolerance engineering [117].

Box 1. Evolutionary Engineering Tools and Strategies

Evolutionary engineering has been a common approach to improve phenotypes of strains or biological components (e.g., protein) mimicking natural evolution process, but at higher rate and efficiency [165,166]. Evolutionary engineering covers both directed evolution and adaptive laboratory evolution (ALE). The principle behind both techniques is to select a strain or protein with desired phenotype or characteristics from randomly generated variants. Directed evolution mainly deals with proteins, while ALE is applied to cells. Engineering proteins to enhance catalytic activity, substrate specificity, thermal tolerance, and stability of enzymes by applying artificial selection pressure is frequently called directed evolution [91,92]. The key principles of evolutionary engineering are massive genetic diversification followed by rapid screening of the generated library. Various novel tools and strategies have been developed to facilitate evolutionary engineering (Figure I). ALE is a powerful approach for evolutionary engineering (Figure IB) [167]. Automated serial culture can also be performed in smaller multimicroplates enabling massive and parallel experiments [121,122]. Furthermore, microbioreactors allowing fine control of the culture conditions such as pH, temperature, dissolved oxygen, and nutrients are also available (Figure IB) [120]. Since the success of evolutionary engineering depends on population heterogeneity generated by mutation, a variety of genome-wide evolution approaches have been applied to increase genetic diversity of the mutant library, such as genome shuffling [168–170], global transcription machinery engineering (gTME) [170,171], multiplex automated genome engineering (MAGE) [127,172], transcription activator-like effector nucleases (TALENs) [173,174], and CRISPR/Cas systems [175] (Figure IA). In addition to the traditional evolutionary engineering approaches, biosensor-assisted adaptive evolution, which is a growth-uncoupled method, enables the effective laboratory evolution of strains to possess desired phenotypes (Figure IC) [176]. Recent advances in whole-genome sequencing, omics technologies, and computational biology also enable better understanding of the genetic basis of evolved populations at molecular levels. In particular, evolutionary engineering can compensate rational and/or other metabolic engineering tools to achieve certain phenotypes, such as tolerance to a product, which is difficult to find a way to engineer. Numerous metabolic engineering studies have used evolutionary engineering approaches to improve the tolerance [118] and performance [22,177] of the strains.

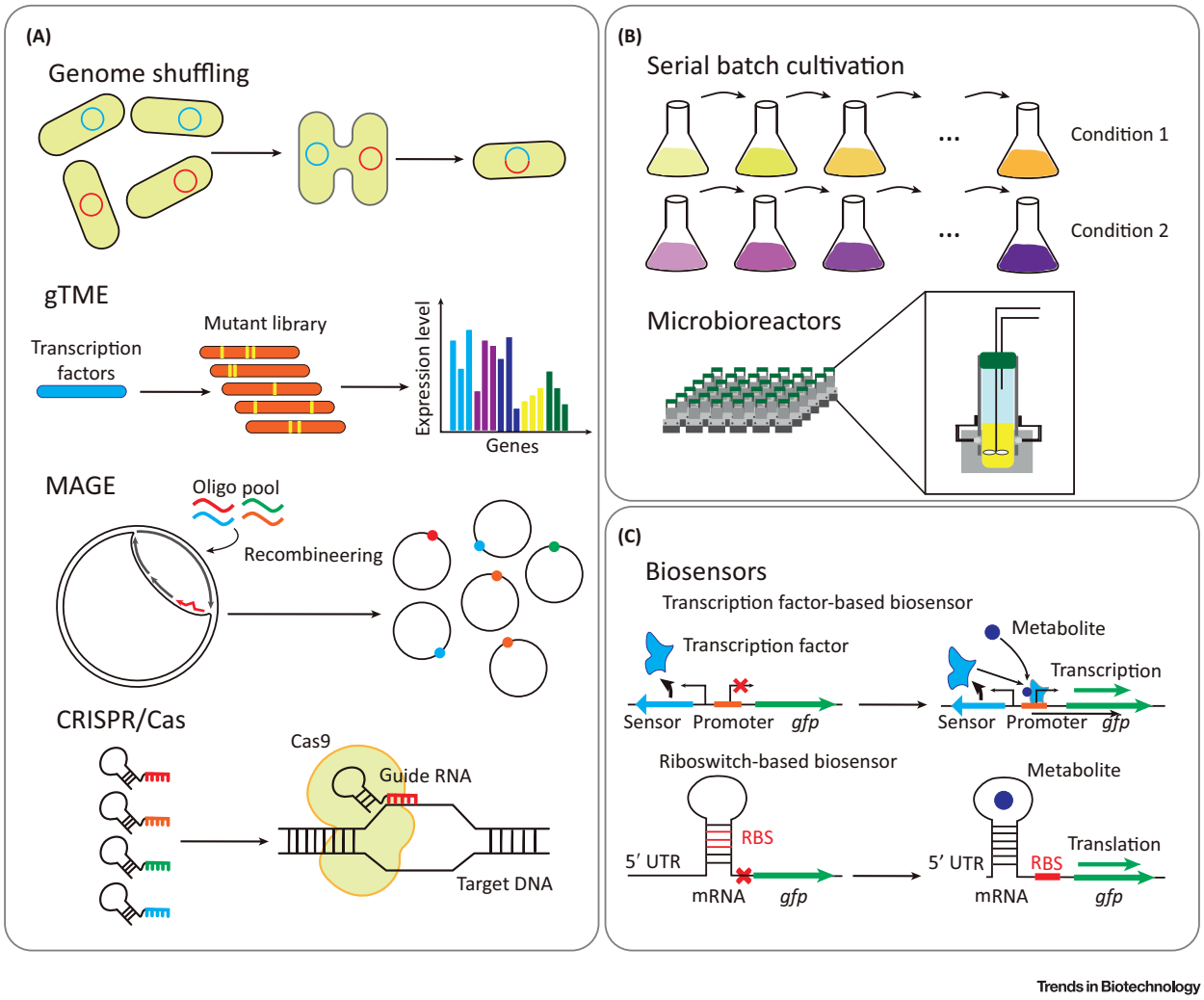


Figure I. Evolutionary Engineering Tools and Strategies for Systems Metabolic Engineering. (A) Genome-wide evolution strategies to increase genetic diversity. Genome shuffling conducted through protoplast fusion introduces multiple mutations throughout the chromosome by recombination-based DNA shuffling. Global transcription machinery engineering (gTME) randomly mutates global transcription factors and modifies global transcription profile. The resulting mutant library with altered phenotypes is screened for desired traits. Multiplex automated genome engineering (MAGE) is a recombineering strategy allowing parallel, continuous, and accelerated evolution of cells. MAGE can generate combinatorial mutations on multiple target sites through recursive rounds of single-stranded DNA recombineering. CRISPR/Cas system allows sequence-specific cleavage of target DNA directed by guide RNAs. Use of multiple guide RNAs allows multiplexed engineering. (B) Serial batch cultivation can be performed in shake flasks to obtain superior variants under selective pressures (e.g., temperature, carbon sources, and toxic compounds). High-throughput microbioreactors enable parallel screening under controlled culture conditions. (C) Biosensors can accelerate ALE enabling high-throughput screening of strains with desired properties. Transcription-factor-based biosensors use allosterically regulated transcription factors that induce transcription of reporter genes (e.g., *gfp*) only in the presence of target metabolites. Riboswitch-based biosensors exploit riboswitches (i.e., RNA stem-loops that change the secondary structure upon the binding of specific ligands). Ribosome binding site (RBS) in the riboswitch forms the stem structure in the absence of target metabolite, preventing translation of reporter mRNA. Binding of the target metabolite to riboswitch induces change in the secondary structure and exposes RBS, leading to the translation of reporter mRNA.

If the toxicity mechanism has yet to be reported, **adaptive laboratory evolution (ALE)** is a useful strategy to isolate strains resistant to the target compound (Box 1 and Figure 1D) [118]. Subsequent systematic analyses of the isolated strains might reveal molecular mechanisms of the resistance [119] and allow additional rational engineering of the host strains, possibly

Table 2. Synthetic Biology Tools and Strategies for Systems Metabolic Engineering

Categories	Tools and strategies	Description	Refs
DNA assembly	BioBrick assembly	Idempotent assembly strategy using isocaudomers (i.e., restriction enzymes pairs that generate the same overhangs). The number and relative positions of the isocaudomer recognition sites in the assembly products are maintained after each round of cloning, allowing simple iterative cloning of DNA fragments.	[97]
	Golden Gate assembly	Seamless, scarless method to assemble multiple DNA molecules using type IIS restriction enzymes and DNA ligase.	[99]
	Gibson assembly	Isothermal, scarless, one-step method for assembling multiple overlapping DNA molecules. 5′ Exonuclease generates 5′ single-stranded DNA overhangs, which anneal to their complementary pairs, DNA polymerase fills the gaps, and DNA ligase repairs the nick.	[98]
	Single strand assembly (SSA)	Gibson assembly-derived method using multiple short overlapping single-stranded DNA oligos for assembly.	[101]
	Ligase cycling reaction (LCR)	One-step, scarless assembly method using single-stranded bridging oligos complementary to the ends of DNA fragments. Denaturation–annealing–ligation cycles controlled by thermocycling are used for assembly.	[100]
	Uracil specific excision reagent (USER) cloning	DNA assembly method employing compatible overhangs generated by uracil-excision. Each DNA fragment is prepared by PCR amplification using primers containing uracil residues and flexible assembly sequence tags.	[103]
	Transformation-associated recombination (TAR) cloning	PCR-independent *in vivo* assembly strategy of capturing, refactoring, and expressing biosynthetic gene clusters of large sizes by exploiting endogenous homologous recombination of *S. cerevisiae*.	[102]
Chromosome engineering	Transposon-mediated integration	Gene integration method exploiting the nature of transposons jumping to random positions on DNA.	[107]
	Site-specific integration	Integration tool employing sequence-specific recombination systems to insert genes of interest to specific genetic elements on chromosome (either endogenous or artificially introduced).	[106]
	Recombineering	Homology-based recombination system derived from bacteriophages. Use of the donor plasmid system was demonstrated to facilitate the integration of large genes to chromosomes. Both chromosome and plasmids can be engineered using this system.	[108]
	Multiplex automated genome engineering (MAGE)	Recombineering-based genome engineering strategy that uses multiple short oligonucleotides for simultaneous modification on multiple target sites. Other genetic engineering tools, such as coselection markers, site-specific recombinases, and CRISPR/Cas systems can be used together to further increase the efficiency of this strategy.	[195]
	CRISPR/Cas technologies	RNA-guided target specific DNA cleavage system originated from bacterial adaptive immune system. This system has many variations for diverse engineering purposes.	[68,69,109,110]
Trans-acting gene expression modulation	Synthetic small regulatory RNA (synthetic sRNA)	Gene expression knockdown tool based on synthetically designed sRNAs that complementarily bind to target mRNAs and block translation.	[148]
	Small transcription activating RNAs (STAR)	Gene expression activation tool that uses synthetic sRNAs binding upstream of target genes to disrupt the formation of stem-loop structures and derepresses the gene expression.	[153]
	CRISPRi	Gene expression knockdown tool using catalytically inactivate effector Cas proteins and their guide RNAs to block transcription of target genes.	[150,151]
	CRISPRa	Gene overexpression tool that uses catalytically inactive effector Cas proteins fused to transcription activators to enhance the expression of target genes.	[154]
	CRISPR/Cas-based DNA methylation editing	CRISPR/Cas-based tool that sequence-specifically edits DNA methylation profile. Catalytically inactive effector Cas proteins are fused to methyltransferases or a dioxygenases to methylate or demethylate target DNA sites.	[155]

Table 2. (continued)

Categories	Tools and strategies	Description	Refs
Stable gene expression	Stabilized promoters with incoherent feedforward loop (iFFL)	Strategy to stabilize the gene expression level of promoters regardless of copy number. TALEs are used to form incoherent feedforward loop.	[147]
	Plasmid addiction system	Post-segregational killing system used to maintain plasmids stably without antibiotics. A toxin-antitoxin system with stable toxin and less stable antitoxin (e.g., *hok/sok* system) is frequently used for this purpose.	[33]
	Stable and tunable plasmid (STAPL) system	A variation of plasmid addiction system for stable and tunable maintenance of plasmid copy number. An essential gene is incorporated to the plasmid and expressed in different levels to modulate the plasmid copy number.	[146]
Substrate channeling	Direct fusion of enzymes	Direct connection of enzymes via short and flexible protein linkers (e.g., glycine–serine linker)	[160]
	Synthetic protein scaffold	Enzyme colocalization strategy using synthetic protein scaffolds and affinity tags to recruit enzymes of interest.	[196]
	Synthetic DNA scaffold	Enzyme colocalizing strategy using zinc finger domains to recruit target enzymes to synthetic DNA motifs used as scaffolds.	[158]
	Synthetic lipid-containing scaffold (SLS)	Enzyme colocalizing strategy exploiting the coassembly of lipids and bacteriophage ϕ6 membrane proteins P9 and P12.	[159]
	Bacterial microcompartment (BMC)	Self-assembling microstructures consisting of selectively permeable protein shells that allow the encapsulation of target enzymes.	[161]
	Sequestration to eukaryotic organelles	Spatial regulation of metabolic pathways employing translocation tags to localize target enzymes to a specific organelle.	[36]
Biosensors	Enzyme-coupled biosensor	Enzyme-based reporter system that converts target metabolites into easy-to-detect compound to estimate the level of target metabolites.	[157]
	Transcription factor-based biosensor	Reporter system based on target metabolite-specific transcription factor that induces the expression of reporter genes, allowing the estimation of target metabolite level based on the intensity of the signals from the reporter system.	[24]
Synthetic genetic circuits	Cello	Genetic circuit design automation strategy using electronic design automation.	[197]
	CHOMP	Synthetic protein-level genetic circuit that consists of viral proteases capable of constructing protein-protein regulation system.	[198]
	Deadman and Passcode	Biocontainment system with circuit-based microbial kill switches. Multiple environmental signals are considered to induce cell death.	[164]
Enzyme engineering	*De novo* enzyme design	Protein engineering strategy using computational models to predict the structure of a desirable protein and create stable and customized proteins.	[39]
	Modular polyketide synthase (PKS) engineering	Protein engineering strategy that exploits different modules of multiple PKSs to create recombinant PKSs with novel and desired catalytic activities.	[89,90]

leading to further tolerance enhancement (Figure 1D) [22]. Recently invented tools and strategies for serial cultivation and subsequent screening – for example, eVOLVER [120], Mini Pilot Plant [121,122], PALE ALE (R. A. LaCorix, PhD thesis, UC San Diego, 2016), and TALE [123] – and for increasing genetic diversity – for example, PACE [124], YOGE [125], ICE [126], and MAGE [127] – are expected to facilitate isolating strains with high resistance against the target products (Box 1). Based on the selection strategies used, the ALE approach will also facilitate identification of strains possessing better performance in terms of titer and productivity.

The toxicity issues of the target chemicals can also be circumvented by process engineering strategies. *In situ* recovery of the chemicals with toxic effects [128] and sequestration of toxic products or precursors to another phase during the fermentation [52,114] have been proved to reduce the toxicity and enhance the performance of the strain (Figure 1D). Such processes might provide further advantages of product recovery and purification in the downstream processes.

Metabolic Flux Optimization

Upon the construction of base strains capable of producing desired products with the reconstructed biosynthetic pathways, various tools and strategies are available to maximize the metabolic flux toward the target products. Systems-level analyses of the host metabolic network and other omics data have played pivotal roles to this end, as exemplified by the initial systems metabolic engineering studies that significantly enhanced the production of L-valine [7] and L-threonine [8]. Recent technological breakthroughs in DNA/RNA sequencing, mass spectrometry, and other omics strategies allow acquisition of more data (**bio-big data**) related to cellular physiology and metabolism, providing clues to optimize the production strains (Figure 1E and Table 1).

Powered by high-throughput DNA/RNA sequencing techniques, transcriptomic profiling has been applied to design and optimize strains and processes by revealing the metabolic processes activated under certain conditions [65,129,130]. However, transcriptomic data should be carefully inspected to understand the metabolism, as the metabolic activities are not directly associated to the transcript levels. Proteome analysis, which more closely represents the actual metabolic activities than the transcriptome analysis, has been facilitated by the recent advancement of mass spectrometry and has contributed to more comprehensive understanding of cell physiology [131,132]. Recent advances in proteomics have also enabled the understanding of post-translational modifications and their impacts on metabolic engineering [133]. Further considerations on protein–protein interactions are expected to provide deeper insights for systems metabolic engineering. It should however be noted that the levels of proteins (enzymes in particular) also do not necessarily match their enzyme activities. Metabolomics, which directly observe the levels of metabolites under particular conditions, can complement our understanding, and thus has also been used to optimize strain performances [134]. Nevertheless, the quantification of metabolites at low concentrations is still a challenge. Fluxomic analysis, which provides the closest description of the cell metabolism among all omics studies, has been widely used to develop industrial strains for various chemicals [64,135,136]. Fluxomic analysis is expected to present blueprints to systems metabolic engineering in spite of the remaining technical difficulties in accurately determining the flux values.

Along with the omics tools and strategies, various *in silico* genome-scale metabolic models (GEMs) [137,138] and associated simulation methods [139,140] have been successfully applied to developing diverse overproducer strains (Figure 1E and Table 1) [33]. Integrating omics data such as transcriptome, proteome, and fluxome information into GEMs to obtain a more comprehensive perspective of the cell metabolism is a recent trend in this field [141,142]. Metabolism and expression (ME) models, which integrate the gene expression and protein synthesis information extracted from the quantitative proteomics data to GEMs, are one example [143]. Similarly, another modeling method called GECKO provides a mechanistic approach to computing cellular conditions by incorporating kinetic parameters of enzymes [144].

Parallel to the advances in omics and *in silico* modeling/simulation, rapid advances in synthetic biology tools and strategies (Table 2) further expedite construction of overproducer strains.

While conventional yet efficient genome-engineering tools including recombineering systems [108] are actively used for engineering, the recent emergence of CRISPR/Cas technologies has shifted the engineering paradigm (Figure 1E) [68,69,109,110]. To precisely control target gene expression and balance metabolic fluxes toward the target product, promoter and ribosome binding site (RBS) libraries have been generated to find optimal expression levels of target genes maximizing the formation of desired products (Figure 1E) [145]. Moreover, incorporation of essential genes was exploited for tunable and stable control over plasmid copy number [146], and transcription activator-like effector (TALE)-coupled promoters were devised to fix the expression level independent of gene copy numbers (Figure 1E and Table 2) [147].

The development of *trans*-acting gene expression knockdown tools, including synthetic **small regulatory RNA (sRNA)**, **RNA interference (RNAi)**, and CRISPR interference (CRISPRi), has allowed the rapid system-wide screening of gene expression modification targets (Figure 1E and Table 2). Downregulating the translation of selected target mRNAs in bacteria by using synthetic sRNAs, strains overproducing chemicals could be easily and efficiently developed [148]. For eukaryotes, RNAi has been successfully used to knockdown target gene expression by translational inhibition or transcript degradation, resulting in enhanced target chemical production [149]. In contrast, CRISPRi, which blocks the transcription of the target gene [150], also enables efficient screening of beneficial gene knockdown targets and subsequent enhanced production of target chemicals [151,152]. Further exploitation of antisense RNAs and CRISPR/Cas systems also enables target gene derepression [153], overexpression [154], and methylation systems (Figure 1E and Table 2) [155]. Introduction of an aptamer, a *cis*-acting gene regulatory element that changes the secondary/tertiary structures upon binding of specific ligands, to the 5′ end of the target gene is another strategy to fine-tune the expression level of target genes based on the concentrations of the specific ligands [156]. These gene-expression-modulating tools can be coupled with metabolite biosensors to identify engineering targets by high-throughput screening (Figure 1E, Figure IC in Box 1, and Table 2), as exemplified by a recent study that coupled a malonyl-CoA biosensor and an *E. coli* genome-scale sRNA library [157]. Further expansion of metabolite biosensor portfolio [24] is expected to facilitate high-throughput screening of beneficial engineering targets for developing high performance strains suitable for industrial-scale biobased production of chemicals.

In addition to the genetic engineering strategies, substrate channeling is an alternative strategy to enhance metabolic fluxes (Figure 1E and Table 2). In addition to spatially colocalizing relevant enzymes using synthetic scaffolds [158,159] or directly fusing the modules [160], target enzymes can be sequestered and colocalized to synthetic microcompartments in bacterial cells [161] or specific organelles in eukaryotic cells [36]. These strategies allow securing intermediates from competing pathways, isolating toxic intermediates, and increasing the local concentrations of intermediates and enzymes.

Scaling up to Industrial Production

An overproducer strain that is engineered to perform well at the laboratory scale might show suboptimal performances as the fermentation is scaled up to the pilot or demonstration scale, and eventually to full industrial-scale production for commercialization (Figure 1F). For example, the changes in oxygen transfer profiles at different fermentation scales are critical to the strain performances in aerobic fermentations. Also, as the fermentor size increases, there is higher chance of local substrate concentration difference due to imperfect mixing, causing suboptimal performance of the developed strain. Many typical issues during scale-up processes can be predicted and considered at the project design step of systems metabolic engineering research to generate strains resistant to such changes in fermentation conditions

Box 2. Current State of Bioproduct Commercialization

There have been increasing numbers of bioproducts commercialized by using metabolically engineered microbes. These bioproducts include monomers, polymers, natural products, amino acids, and fuels among others (Figure I). Diols used as monomers include 1,2-propanediol, 1,3-propanediol, and 1,4-butanediol and were commercialized by Cargill, DuPont, and Genomatica/Novamont. In addition, engineered microorganisms and bioprocesses for the fermentative production of one or more of the monomers used in manufacturing bio-based polymers have also been commercialized by many companies, including Arkema, BASF, Braskem, Cathay industrial biotech, Coca-Cola, DSM, Cristal Union, DuPont, Evonik, GFBiochemicals, Global Bioenergies, Goodyear, India Glycols Ltd., Itaconix, JBF Industries Ltd., NatureWorks, Purac, and Toray.

Natural products that include fragrance and flavoring agents have also been commercialized. Vanillin is a popularly used flavor compound with many industrial applications, where 85% of the total vanillin produced has traditionally been sourced from petrochemicals (labeled 'artificial'), 15% from woody biomass, and less than 1% extracted naturally from vanilla orchids (labeled 'natural'). As a step toward increasing natural vanillin, a metabolic engineering effort has been exerted to introduce heterologous biosynthetic pathways to produce vanillin from 3-dehydroshikimic acid in microbes [178–180]. The Swiss company Evolva successfully commercialized vanillin production using a metabolically engineered microbe, and in 2011 partnered with International Flavor and Fragrances (IFF) with the product being brought to market in mid-2014 (www.evolva.com/vanillin). Evolva also commercially produces resveratrol using engineered microbes (www.veriteresveratrol.com). The antimalarial drug artemisinin, produced by engineered yeast, is one of the best examples of commercialized natural products produced by metabolic engineering. It will be of interest to see how complex economic issues hampering continued large-scale production associated with this process will be resolved [181].

Amino acids are another group of chemicals that are extensively produced by engineered microbes such as *C. glutamicum* and *E. coli*. One recent success example is L-methionine, a highly demanded essential amino acid for animal feed supplement with a global market size of US$5 billion dollars. Conventional technologies using petroleum-based methods had dominated the market and had been only able to offer DL-methionine. More recently, Metabolic Explorer and CJ Cheil Jedang independently developed and established biofermentation processes to produce stereospecific L-methionine, which was once infeasible with conventional production methods (https://www.metabolic-explorer.com/2016/11/03/metabolic-explorer-has-received-from-evonik-a-binding-offer-for-its-methionine-technology-platform/; https://redox.com/News/product-focus-cj-start-production-of-l-methionine).

Chemicals that can be used as biofuels such as isobutanol and farnesene have also been successfully commercialized by companies Gevo and Amyris, respectively. Other chemicals that are being commercially produced are docosahexaenoic acid (Corbion), β-carotene (BASF), and lipids (Terravia).

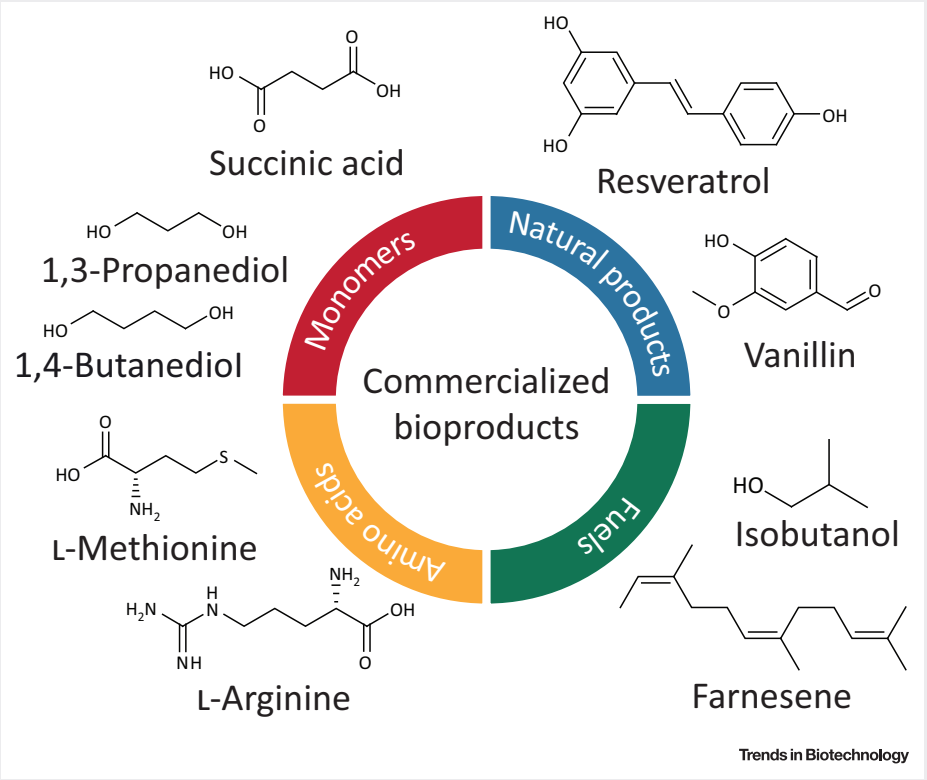
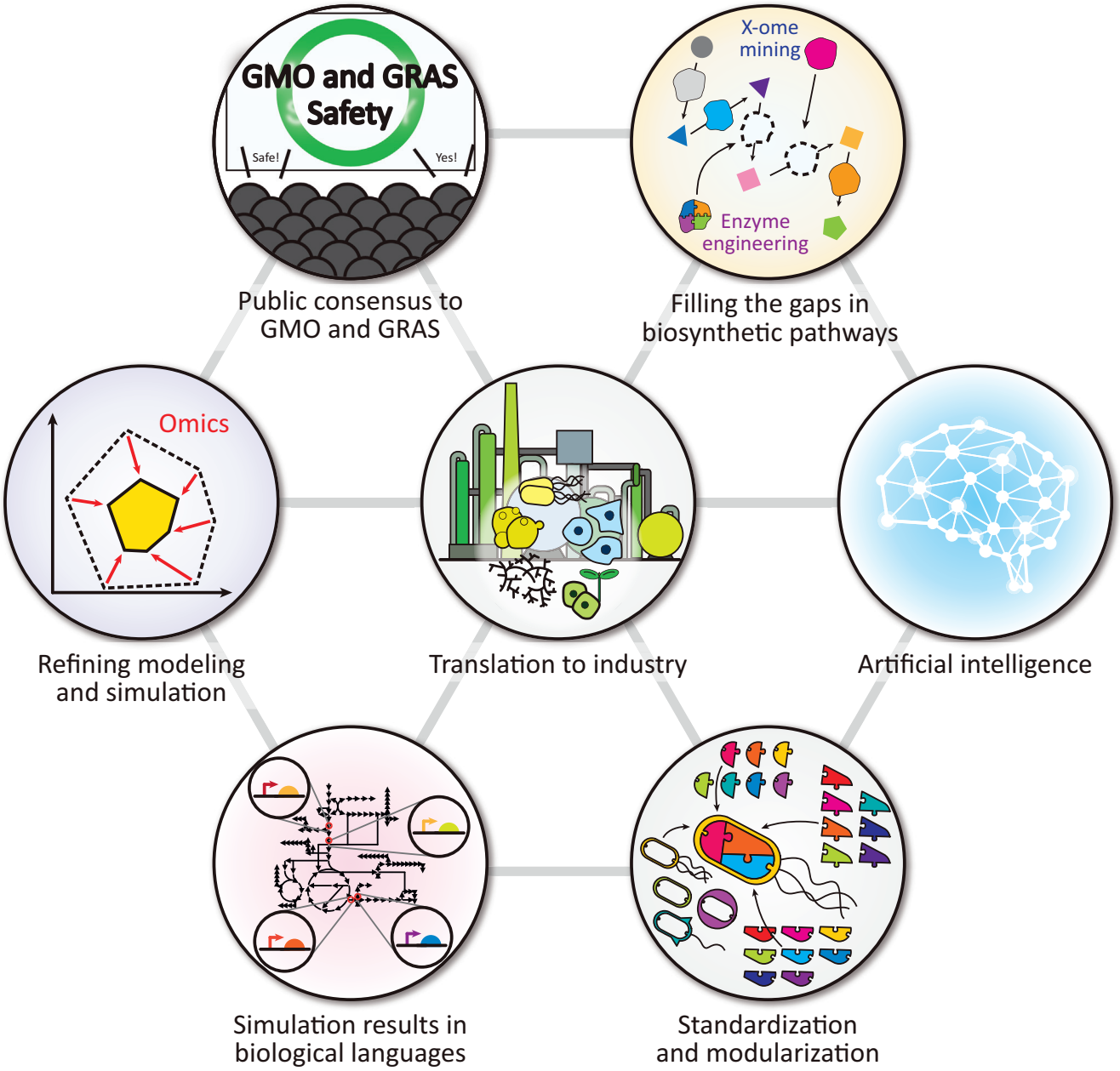


Figure I. Representative Examples of Commercialized Bioproducts. Representative bioproducts that have been commercialized through industrial-scale fermentations using metabolically engineered microorganisms.



Trends in Biotechnology

Figure 2. Future Directions of Systems Metabolic Engineering. Public consensus about the safety of genetically modified organisms (GMOs), especially engineered generally recognized as safe (GRAS) strains and their products, plays an important role in promoting the translation of high-performing production strains to industries. Although more than 16 000 metabolic reactions have been reported, biosynthetic pathways of many important chemicals (e.g., natural products) are still unknown. Continued efforts, such as X-ome mining and enzyme engineering, are required to fill the gaps. The *in silico* modeling/simulation of production host metabolism can be refined to provide more precise results with the integration of omics information such as transcriptome, proteome, metabolome, fluxome, regulome, or even enzyme kinetics. Moreover, current *in silico* modeling/simulation returns numerical values of optimal fluxes that cannot be directly translated to biological expression components [e.g., promoter and ribosome binding site (RBS)], requiring screening experiments to identify the best expression conditions. *In silico* tools that provide modeling/simulation results in biological languages (e.g., suggestions for promoters and RBSs) will reduce the gap between *in silico* and wet experiments, thus streamlining strain development. Standardization and modularization of biological parts are also urgent assignments to be resolved as modest changes in the context (e.g., medium composition, temperature, expression platform, and biological components used together) often cause significant fluctuations in the performances. *In silico* tools that predict contextual effects and suggest appropriate biological parts to be used can be an alternative solution for the performance variation issue. Recent advances in artificial intelligence are expected to upgrade multiple aspects of systems metabolic engineering, including aforementioned issues.

(Figure 1). In addition, systems metabolic engineering helps resolve unexpected issues that arise during the scale-up by re-entering the strain engineering cycles to improve the performance of the strains during larger scale fermentations (Figure 1). Further iteration of the systems metabolic engineering cycles that further upgrade strain performances based on the feedback obtained from bioprocess operation (including fermentation, product recovery, and purification) allows successful establishment of industrial-scale production. In addition, plasmid-based overexpression of genes often carries the risk of plasmid instability, particularly when metabolic burden exists [33]. Such genetic instability issues of production strains harboring plasmids can be overcome by integrating genetic changes to the chromosomes [22,26] as described earlier. While the use of antibiotics is discouraged, the problem of contaminating microorganisms affecting fermentation industries is a serious problem that requires much attention. Recent approaches demonstrated improvement in competitive fitness of production hosts against contaminating organisms by introducing xenobiotic nutrient utilization pathways to the production hosts [162]. Instead of using common compounds for the nitrogen source of the host that is readily used also by contaminating organisms, rare xenobiotic nutrient compounds (e.g., melamine) that are only catabolized by the production host harboring the utilization pathway can be used as the nitrogen sources. Such innovative strategies can address the problem of contaminating microorganisms growing in the absence of antibiotics [163].

Biocontainment of the engineered stains is another important issue to be considered to develop an engineered industrial production strain – not only to prevent unwanted environmental dissemination but also to secure the strain. While most of such biocontainment systems explore the auxotrophic strains that require exogenous or nonnatural metabolites for growth, recent advances in synthetic biology have constructed synthetic biological circuits that allow cell growth only when complex combination of nutrients have been met (Table 2) [164]. Such tools and strategies are believed to further expand the current profile of outstanding overproduction strains and bioprocesses successfully translated to the industry (Box 2).

Concluding Remarks and Future Perspectives

For more than a decade, systems metabolic engineering has demonstrated its potential to streamline the overall processes from the initial strain design to scaling up to full-scale industrial production. This interdisciplinary field continues to evolve rapidly with advances in the fields of systems biology (Table 1), synthetic biology (Table 2), and evolutionary engineering (Box 1). New tools and strategies are continuously developed in these three disciplines and are accelerating the design and engineering of superior producer strains optimized for actual industrial applications.

As mentioned earlier, only a few producer strains and processes developed have actually been translated to industrial production (Box 2; see Outstanding Questions). Although many different microbial strains capable of overproducing diverse chemicals and materials have been successfully developed by systems metabolic engineering, there are still even more important chemicals – especially natural products of high medical and nutritional significance – awaiting production. Development of engineered strains for producing many such natural products is still difficult, mainly due to the lack of complete knowledge on the corresponding biosynthetic pathways/enzymes and generally long/complex biosynthetic pathways involving some difficult-to-express genes of plant or animal origin. Continued genome mining and enzyme engineering/evolution will expand the pool of efficient biosynthetic pathways and enzymes (Figure 2), while better tools still need to be developed for expressing genes that have so far been difficult to express.

Outstanding Questions

How can we accelerate the translation of engineered overproducer strains to industries?

How can we improve public perception to metabolically engineered microbial strains?

How can we accelerate the expansion of biosynthetic pathways and enzyme pools? What are the new tools and strategies in genome mining and relevant omics studies that will expedite the identification of metabolic pathways and enzymes for bioproducts of interest, in particular natural products? How can we best apply enzyme engineering and evolution strategies to develop missing enzymes for the biosynthesis of bioproducts including natural products?

How can we more accurately determine solution spaces computed from *in silico* metabolic simulation for more precise prediction and design? How can data from omics studies and systems biology strategies be better integrated to enhance precision of modeling and simulation?

How can we make direct links between *in silico* simulation results and practical parameters in strain engineering? Can *in silico* simulation provide specific biological parts in the library, rather than the numerical data on optimal fluxes, for more intuitive and straightforward engineering?

How can we standardize and modularize biological parts for more accurate plug-and-play uses during strain engineering? Are perfectly modular biological parts (i.e., standardized performance and interference in any biological context) realizable to some extent? If it is unrealistic, could *in silico* strategies be used to consider context-dependent fluctuations during the modeling/simulation and to provide adjusted outputs for practical engineering purposes?

How can we incorporate recent advances in artificial intelligence and machine learning to upgrade the current tools and strategies of systems metabolic engineering? How can we make biological data to be collected as bio-big data more reliable and universally usable?

Refining the solution spaces computationally determined by *in silico* metabolic simulation is also urgently needed to enable higher engineering predictability (Figure 2). Integration of omics data on the transcription, translation, and catalytic rates of metabolic enzymes as well as relevant proteome, metabolome, fluxome, and regulome information are expected to improve the precision of modeling and simulation. Better understanding and developing the quantitative links between the *in silico* simulation results (e.g., numerical flux data) and actual components and parameters in strain development (e.g., promoter/RBS and their strengths) will further facilitate the application of the modeling and simulation results to actual construction of strains (Figure 2).

Moreover, biological parts used for strain development often show significant fluctuations in different biological contexts and interference to each other. Standardization and modularization of the strains/biological parts or the development of *in silico* tools that predict the inevitable interferences and provide corrected designs for engineering will minimize unnecessary trial-and-error type experiments (Figure 2). It is expected that still rapidly increasing amounts of data will become available. As such bio-big data become available, tools and strategies of data science and artificial intelligence will be increasingly used for extracting new knowledge and information and also for suggesting better engineering strategies to upgrade systems metabolic engineering (Figure 2).

Acknowledgments

This work was supported by the Technology Development Program to Solve Climate Changes on Systems Metabolic Engineering for Biorefineries (NRF-2012M1A2A2026556, NRF-2012M1A2A2026557) from the Ministry of Science and ICT through the National Research Foundation (NRF) of Korea. This work was further supported by Hanwha Chemical through the KAIST-Hanwha Future Technology Institute.

References

1. Bailey, J.E. (1991) Toward a science of metabolic engineering. *Science* 252, 1668–1675
2. Choi, K.R. *et al.* (2016) Systems metabolic engineering of *Escherichia coli*. *EcoSal Plus* 7, 1–56
3. Park, S.Y. *et al.* (2018) Metabolic engineering of microorganisms for the production of natural compounds. *Adv. Biosyst.* 2, 1700190
4. Pontrelli, S. *et al.* (2018) *Escherichia coli* as a host for metabolic engineering. *Metab. Eng.* 50, 16–46
5. Lee, S.Y. and Kim, H.U. (2015) Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072
6. Hong, K.K. and Nielsen, J. (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.* 69, 2671–2690
7. Park, J.H. *et al.* (2007) Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 7797–7802
8. Lee, K.H. *et al.* (2007) Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* 3, 149
9. Gustavsson, M. and Lee, S.Y. (2016) Prospects of microbial cell factories developed through systems metabolic engineering. *Microb. Biotechnol.* 9, 610–617
10. Chae, T.U. *et al.* (2017) Recent advances in systems metabolic engineering tools and strategies. *Curr. Opin. Biotechnol.* 47, 67–82
11. Lee, J.W. *et al.* (2011) Systems metabolic engineering for chemicals and materials. *Trends Biotechnol.* 29, 370–378
12. Lee, J.W. *et al.* (2012) Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* 8, 536–546
13. d'Espaux, L. *et al.* (2017) Engineering high-level production of fatty alcohols by *Saccharomyces cerevisiae* from lignocellulosic feedstocks. *Metab. Eng.* 42, 115–125
14. Wang, J. *et al.* (2017) Rational engineering of diol dehydratase enables 1,4-butanediol biosynthesis from xylose. *Metab. Eng.* 40, 148–156
15. Chen, Z. *et al.* (2015) Protein design and engineering of a *de novo* pathway for microbial production of 1,3-propanediol from glucose. *Biotechnol. J.* 10, 284–289
16. Kim, S. *et al.* (2015) Redox-switch regulatory mechanism of thiolase from *Clostridium acetobutylicum*. *Nat. Commun.* 6, 8410
17. Liao, J.C. *et al.* (2016) Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* 14, 288–304
18. Qiao, K. *et al.* (2017) Lipid production in *Yarrowia lipolytica* is maximized by engineering cytosolic redox metabolism. *Nat. Biotechnol.* 35, 173–177
19. Choi, S. *et al.* (2016) Highly selective production of succinic acid by metabolically engineered *Mannheimia succiniciproducens* and its efficient purification. *Biotechnol. Bioeng.* 113, 2168–2177
20. Rohles, C.M. *et al.* (2018) A bio-based route to the carbon-5 chemical glutaric acid and to bionylon-6,5 using metabolically engineered *Corynebacterium glutamicum*. *Green Chem.* 20, 4662–4674
21. Choi, S.Y. *et al.* (2016) One-step fermentative production of poly (lactate-*co*-glycolate) from carbohydrates in *Escherichia coli*. *Nat. Biotechnol.* 34, 435–440
22. Park, S.H. *et al.* (2014) Metabolic engineering of *Corynebacterium glutamicum* for L-arginine production. *Nat. Commun.* 5, 4618

23. Jang, W.D. *et al.* (2017) Bacterial cellulose as an example product for sustainable production and consumption. *Microb. Biotechnol.* 10, 1181–1185
24. Rogers, J.K. and Church, G.M. (2016) Genetically encoded sensors enable real-time observation of metabolite production. *Proc. Natl. Acad. Sci. U. S. A.* 113, 2388–2393
25. Brazeau, B. *et al*. CJ CheilJedang Corp. Compositions and methods of producing methionine, US8551742B2
26. Becker, J. *et al.* (2011) From zero to hero – design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metab. Eng.* 13, 159–168
27. Kim, T.Y. *et al.* (2015) Design of homo-organic acid producing strains using multi-objective optimization. *Metab. Eng.* 28, 63–73
28. Galanie, S. *et al.* (2015) Complete biosynthesis of opioids in yeast. *Science* 349, 1095–1100
29. Paddon, C.J. *et al.* (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532
30. Zhou, K. *et al.* (2015) Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* 33, 377–383
31. Shomar, H. *et al.* (2018) Metabolic engineering of a carbapenem antibiotic synthesis pathway in *Escherichia coli*. *Nat. Chem. Biol.* 14, 794–800
32. Zhao, X.R. *et al.* (2018) Metabolic engineering of *Escherichia coli* for secretory production of free haem. *Nat. Catal.* 1, 720–728
33. Park, S.Y. *et al.* (2018) Metabolic engineering of *Escherichia coli* for high-level astaxanthin production with high productivity. *Metab. Eng.* 49, 105–115
34. Katabami, A. *et al.* (2015) Production of squalene by squalene synthases and their truncated mutants in *Escherichia coli*. *J. Biosci. Bioeng.* 119, 165–171
35. Wriessnegger, T. *et al.* (2014) Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of *Pichia pastoris*. *Metab. Eng.* 24, 18–29
36. Li, S. *et al.* (2015) Compartmentalizing metabolic pathway in *Candida glabrata* for acetoin production. *Metab. Eng.* 28, 1–7
37. Fleige, C. *et al.* (2016) Metabolic engineering of the *Actinomycete amycolatopsis* sp. strain ATCC 39116 towards enhanced production of natural vanillin. *Appl. Environ. Microbiol.* 82, 3410–3419
38. Lee, S.Y. *et al*. Korea Advanced Institute of Science and Technology. Method for synthesizing protein containing high content of specific amino acid through simultaneous expression with tRNA of the specific amino acid, EP2330186B1
39. Chevalier, A. *et al.* (2017) Massively parallel *de novo* protein design for targeted therapeutics. *Nature* 550, 74–79
40. Paddon, C.J. and Keasling, J.D. (2014) Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat. Rev. Microbiol.* 12, 355–367
41. Yano, J.M. *et al.* (2015) Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* 161, 264–276
42. Hwang, I.Y. *et al.* (2017) Engineered probiotic *Escherichia coli* can eliminate and prevent *Pseudomonas aeruginosa* gut infection in animal models. *Nat. Commun.* 8, 15028
43. Riglar, D.T. *et al.* (2017) Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* 35, 653–658
44. Isabella, V.M. *et al.* (2018) Development of a synthetic live bacterial therapeutic for the human metabolic disease phenylketonuria. *Nat. Biotechnol.* 36, 857–864
45. Din, M.O. *et al.* (2016) Synchronized cycles of bacterial lysis for *in vivo* delivery. *Nature* 536, 81–85
46. Mimee, M. *et al.* (2018) An ingestible bacterial-electronic system to monitor gastrointestinal health. *Science* 360, 915–918
47. Choi, K.Y. *et al.* (2014) Consolidated conversion of protein waste into biofuels and ammonia using *Bacillus subtilis*. *Metab. Eng.* 23, 53–61
48. Joo, S. *et al.* (2018) Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat. Commun.* 9, 382
49. Yang, Y. *et al.* (2015) Biodegradation and mineralization of polystyrene by plastic-eating mealworms: part 2. Role of gut microorganisms. *Environ. Sci. Technol.* 49, 12087–12093
50. Urbanek, A.K. *et al.* (2018) Degradation of plastics and plastic-degrading bacteria in cold marine habitats. *Appl. Microbiol. Biotechnol.* http://dx.doi.org/10.1007/s00253-018-9195-y
51. Nie, Y. *et al.* (2014) Diverse alkane hydroxylase genes in microorganisms and environments. *Sci. Rep.* 4, 4968
52. Luo, Z.W. and Lee, S.Y. (2017) Biotransformation of *p*-xylene into terephthalic acid by engineered *Escherichia coli*. *Nat. Commun.* 8, 15689
53. Cheong, S. *et al.* (2016) Energy- and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. *Nat. Biotechnol.* 34, 556–561
54. Karp, E.M. *et al.* (2017) Renewable acrylonitrile production. *Science* 358, 1307–1310
55. Chae, T.U. *et al.* (2018) Production of ethylene glycol from xylose by metabolically engineered *Escherichia coli*. *AIChE J.* 64, 4193–4200
56. Shin, H.D. *et al.* (2014) Comparative engineering of *Escherichia coli* for cellobiose utilization: hydrolysis versus phosphorolysis. *Metab. Eng.* 24, 9–17
57. Clomburg, J.M. *et al.* (2017) Industrial biomanufacturing: the future of chemical production. *Science* 355
58. Kalyuzhnaya, M.G. *et al.* (2015) Metabolic engineering in methanotrophic bacteria. *Metab. Eng.* 29, 142–152
59. Yu, H. and Liao, J.C. (2018) A modified serine cycle in *Escherichia coli* coverts methanol and CO_2 to two-carbon compounds. *Nat. Commun.* 9, 3992
60. Bang, J. and Lee, S.Y. (2018) Assimilation of formic acid and CO_2 by engineered *Escherichia coli* equipped with reconstructed one-carbon assimilation pathways. *Proc. Natl. Acad. Sci. U. S. A.* 115, E9271–E9279
61. Oliver, J.W.K. and Atsumi, S. (2015) A carbon sink pathway increases carbon productivity in cyanobacteria. *Metab. Eng.* 29, 106–112
62. Choi, Y.J. and Lee, S.Y. (2013) Microbial production of short-chain alkanes. *Nature* 502, 571–574
63. Qi, F. *et al.* (2018) Improvement of butanol production in *Clostridium acetobutylicum* through enhancement of NAD(P)H availability. *J. Ind. Microbiol. Biotechnol.* 45, 993–1002
64. Hoffmann, S.L. *et al.* (2018) Lysine production from the sugar alcohol mannitol: design of the cell factory *Corynebacterium glutamicum* SEA-3 through integrated analysis and engineering of metabolic pathway fluxes. *Metab. Eng.* 47, 475–487
65. Lee, J.W. *et al.* (2016) Homo-succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. *Metab. Eng.* 38, 409–417
66. Kurosawa, K. *et al.* (2015) Engineering L-arabinose metabolism in triacylglycerol-producing *Rhodococcus opacus* for lignocellulosic fuel production. *Metab. Eng.* 30, 89–95
67. Sagnak, R. *et al.* (2018) Modulation of the glycerol phosphate availability led to concomitant reduction in the citric acid excretion and increase in lipid content and yield in *Yarrowia lipolytica*. *J. Biotechnol.* 265, 40–45
68. Tong, Y. *et al.* (2015) CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS Synth. Biol.* 4, 1020–1029
69. Cho, J.S. *et al.* (2017) CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. *Metab. Eng.* 42, 157–167
70. Yu, J. *et al.* (2018) Bioengineering triacetic acid lactone production in *Yarrowia lipolytica* for pogostone synthesis. *Biotechnol. Bioeng.* 115, 2383–2388
71. Cho, C. and Lee, S.Y. (2017) Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. *Biotechnol. Bioeng.* 114, 374–383

72. Zhang, X. *et al.* (2014) The rebalanced pathway significantly enhances acetoin production by disruption of acetoin reductase gene and moderate-expression of a new water-forming NADH oxidase in *Bacillus subtilis*. *Metab. Eng.* 23, 34–41
73. Kalyanasundram, J. *et al.* (2015) Surface display of glycosylated tyrosinase related protein-2 (TRP-2) tumour antigen on *Lactococcus lactis*. *BMC Biotechnol.* 15, 113
74. Graf, N. and Altenbuchner, J. (2014) Genetic engineering of *Pseudomonas putida* KT2440 for rapid and high-yield production of vanillin from ferulic acid. *Appl. Microbiol. Biotechnol.* 98, 137–149
75. Kanno, M. and Atsumi, S. (2017) Engineering an obligate photoautotrophic *Cyanobacterium* to utilize glycerol for growth and chemical production. *ACS Synth. Biol.* 6, 69–75
76. Zeldes, B.M. *et al.* (2015) Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Front. Microbiol.* 6, 1209
77. Fu, X.Z. *et al.* (2014) Development of *Halomonas* TD01 as a host for open production of chemicals. *Metab. Eng.* 23, 78–91
78. Kim, Y.K. *et al.* (2014) Enhanced triterpene accumulation in *Panax ginseng* hairy roots overexpressing mevalonate-5-pyrophosphate decarboxylase and farnesyl pyrophosphate synthase. *ACS Synth. Biol.* 3, 773–779
79. Wang, W. *et al.* (2018) Enhanced biosynthesis performance of heterologous proteins in CHO-K1 cells using CRISPR-Cas9. *ACS Synth. Biol.* 7, 1259–1268
80. Xu, J. *et al.* (2018) Mass spider silk production through targeted gene replacement in *Bombyx mori*. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8757–8762
81. Gutensohn, M. *et al.* (2014) Metabolic engineering of monoterpene biosynthesis in tomato fruits via introduction of the non-canonical substrate neryl diphosphate. *Metab. Eng.* 24, 107–116
82. Yim, H. *et al.* (2011) Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* 7, 445–452
83. Feher, T. *et al.* (2014) Validation of RetroPath, a computer-aided design tool for metabolic pathway engineering. *Biotechnol. J.* 9, 1446–1457
84. Hadadi, N. *et al.* (2016) ATLAS of biochemistry: a repository of all possible biochemical reactions for synthetic biology and metabolic engineering studies. *ACS Synth. Biol.* 5, 1155–1166
85. Duigou, T. *et al.* (2018) RetroRules: a database of reaction rules for engineering biology. *Nucleic Acids Res.* 47, D1229–D1235
86. Kumar, A. *et al.* (2018) Pathway design using *de novo* steps through uncharted biochemical spaces. *Nat. Commun.* 9, 184
87. Kan, S.B. *et al.* (2016) Directed evolution of cytochrome *c* for carbon-silicon bond formation: bringing silicon to life. *Science* 354, 1048–1051
88. Siegel, J.B. *et al.* (2015) Computational protein design enables a novel one-carbon assimilation pathway. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3704–3709
89. Hagen, A. *et al.* (2016) Engineering a polyketide synthase for *in vitro* production of adipic acid. *ACS Synth. Biol.* 5, 21–27
90. Yuzawa, S. *et al.* (2018) Short-chain ketone production by engineered polyketide synthases in *Streptomyces albus*. *Nat. Commun.* 9, 4569
91. Chen, K. *et al.* (2018) Enzymatic construction of highly strained carbocycles. *Science* 360, 71–75
92. Kan, S.B.J. *et al.* (2017) Genetically programmed chiral organoborane synthesis. *Nature* 552, 132–136
93. Blin, K. *et al.* (2017) antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 45, W36–W41
94. Skinnider, M.A. *et al.* (2017) PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 45, W49–W54
95. Albright, J.C. *et al.* (2014) Strain-specific proteogenomics accelerates the discovery of natural products via their biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* 41, 451–459
96. Segler, M.H.S. *et al.* (2018) Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* 555, 604–610
97. Storch, M. *et al.* (2015) BASIC: a new biopart assembly standard for idempotent cloning provides accurate, single-tier DNA assembly for synthetic biology. *ACS Synth. Biol.* 4, 781–787
98. Casini, A. *et al.* (2014) One-pot DNA construction for synthetic biology: the Modular Overlap-Directed Assembly with Linkers (MODAL) strategy. *Nucleic Acids Res.* 42, e7
99. Potapov, V. *et al.* (2018) Comprehensive profiling of four base overhang ligation fidelity by T4 DNA ligase and application to DNA assembly. *ACS Synth. Biol.* 7, 2665–2674
100. de Kok, S. *et al.* (2014) Rapid and reliable DNA assembly via ligase cycling reaction. *ACS Synth. Biol.* 3, 97–106
101. Coussement, P. *et al.* (2014) One step DNA assembly for combinatorial metabolic engineering. *Metab. Eng.* 23, 70–77
102. Ross, A.C. *et al.* (2015) Targeted capture and heterologous expression of the *Pseudoalteromonas* alterochromide gene cluster in *Escherichia coli* represents a promising natural product exploratory platform. *ACS Synth. Biol.* 4, 414–420
103. Lund, A.M. *et al.* (2014) A versatile system for USER cloning-based assembly of expression vectors for mammalian cell engineering. *PLoS One* 9, e96693
104. Hughes, R.A. and Ellington, A.D. (2017) Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* 9, a023812
105. Smanski, M.J. *et al.* (2014) Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* 32, 1241–1249
106. Petersen, K.V. *et al.* (2013) Repetitive, marker-free, site-specific integration as a novel tool for multiple chromosomal integration of DNA. *Appl. Environ. Microbiol.* 79, 3563–3569
107. Domrose, A. *et al.* (2017) Rapid generation of recombinant *Pseudomonas putida* secondary metabolite producers using yTREX. *Synth. Syst. Biotechnol.* 2, 310–319
108. Choi, K.R. *et al.* (2018) Markerless gene knockout and integration to express heterologous biosynthetic gene clusters in *Pseudomonas putida*. *Metab. Eng.* 47, 463–474
109. Cong, L. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823
110. Choi, K.R. and Lee, S.Y. (2016) CRISPR technologies for bacterial systems: current achievements and future directions. *Biotechnol. Adv.* 34, 1180–1209
111. Nakagawa, A. *et al.* (2016) Total biosynthesis of opiates by stepwise fermentation using engineered *Escherichia coli*. *Nat. Commun.* 7, 10390
112. Fang, L. *et al.* (2018) Heterologous erythromycin production across strain and plasmid construction. *Biotechnol. Prog.* 34, 271–276
113. Tan, Z. *et al.* (2016) Membrane engineering via *trans* unsaturated fatty acids production improves *Escherichia coli* robustness and production of biorenewables. *Metab. Eng.* 35, 105–113
114. Sherkhanov, S. *et al.* (2014) Improving the tolerance of *Escherichia coli* to medium-chain fatty acid production. *Metab. Eng.* 25, 1–7
115. Fisher, M.A. *et al.* (2014) Enhancing tolerance to short-chain alcohols by engineering the *Escherichia coli* AcrB efflux pump to secrete the non-native substrate n-butanol. *ACS Synth. Biol.* 3, 30–40
116. Foo, J.L. *et al.* (2014) Improving microbial biogasoline production in *Escherichia coli* using tolerance engineering. *mBio* 5, e01932
117. Mukhopadhyay, A. (2015) Tolerance engineering in bacteria for the production of advanced biofuels and chemicals. *Trends Microbiol.* 23, 498–508
118. Caspeta, L. *et al.* (2014) Biofuels. Altered sterol composition renders yeast thermotolerant. *Science* 346, 75–78
119. Royce, L.A. *et al.* (2015) Evolution for exogenous octanoic acid tolerance improves carboxylic acid production and membrane integrity. *Metab. Eng.* 29, 180–188

120. Wong, B.G. *et al.* (2018) Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nat. Biotechnol.* 36, 614–623
121. Radek, A. *et al.* (2017) Miniaturized and automated adaptive laboratory evolution: evolving *Corynebacterium glutamicum* towards an improved D-xylose utilization. *Bioresour. Technol.* 245, 1377–1385
122. Unthan, S. *et al.* (2015) Bioprocess automation on a Mini Pilot Plant enables fast quantitative microbial phenotyping. *Microb. Cell Fact.* 14, 32
123. Mohamed, E.T. *et al.* (2017) Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution. *Microb. Cell Fact.* 16, 204
124. Esvelt, K.M. *et al.* (2011) A system for the continuous directed evolution of biomolecules. *Nature* 472, 499–503
125. DiCarlo, J.E. *et al.* (2013) Yeast oligo-mediated genome engineering (YOGE). *ACS Synth. Biol.* 2, 741–749
126. Crook, N. *et al.* (2016) *In vivo* continuous evolution of genes and pathways in yeast. *Nat. Commun.* 7, 13051
127. Wang, H.H. *et al.* (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature* 460, 894–898
128. Jang, Y.S. *et al.* (2013) Acetone-butanol-ethanol production with high productivity using *Clostridium acetobutylicum* BKM19. *Biotechnol. Bioeng.* 110, 1646–1653
129. Ajjawi, I. *et al.* (2017) Lipid production in *Nannochloropsis gaditana* is doubled by decreasing expression of a single transcriptional regulator. *Nat. Biotechnol.* 35, 647–652
130. Chung, S.C. *et al.* (2017) Improvement of succinate production by release of end-product inhibition in *Corynebacterium glutamicum*. *Metab. Eng.* 40, 157–164
131. Schmidt, A. *et al.* (2016) The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* 34, 104–110
132. Basan, M. *et al.* (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* 528, 99–104
133. Xu, J.Y. *et al.* (2018) Protein acylation affects the artificial biosynthetic pathway for pinosylvin production in engineered *E. coli*. *ACS Chem. Biol.* 13, 1200–1208
134. Ohtake, T. *et al.* (2017) Metabolomics-driven approach to solving a CoA imbalance for improved 1-butanol production in *Escherichia coli*. *Metab. Eng.* 41, 135–143
135. Lange, A. *et al.* (2017) Bio-based succinate from sucrose: high-resolution ^{13}C metabolic flux analysis and metabolic engineering of the rumen bacterium *Basfia succiniciproducens*. *Metab. Eng.* 44, 198–212
136. Schwechheimer, S.K. *et al.* (2018) Improved riboflavin production with *Ashbya gossypii* from vegetable oil based on ^{13}C metabolic network analysis with combined labeling analysis by GC/MS, LC/MS, 1D, and 2D NMR. *Metab. Eng.* 47, 357–373
137. Monk, J.M. *et al.* (2017) iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* 35, 904–908
138. Magnusdottir, S. *et al.* (2017) Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35, 81–89
139. Choi, H.S. *et al.* (2010) *In silico* identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.* 76, 3097–3105
140. Park, J.M. *et al.* (2012) Flux variability scanning based on enforced objective flux for identifying gene amplification targets. *BMC Syst. Biol.* 6, 106
141. Kim, M. *et al.* (2016) Multi-omics integration accurately predicts cellular state in unexplored conditions for *Escherichia coli*. *Nat. Commun.* 7, 13090
142. Monk, J.M. *et al.* (2016) Multi-omics quantification of species variation of *Escherichia coli* links molecular features with strain phenotypes. *Cell Syst.* 3, 238–251.e212
143. Lloyd, C.J. *et al.* (2018) COBRAme: a computational framework for genome-scale models of metabolism and gene expression. *PLoS Comput. Biol.* 14, e1006302
144. Sanchez, B.J. *et al.* (2017) Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* 13, 935
145. Elmore, J.R. *et al.* (2017) Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. *Metab. Eng. Commun.* 5, 1–8
146. Kang, C.W. *et al.* (2018) Synthetic auxotrophs for stable and tunable maintenance of plasmid copy number. *Metab. Eng.* 48, 121–128
147. Segall-Shapiro, T.H. *et al.* (2018) Engineered promoters enable constant gene expression at any copy number in bacteria. *Nat. Biotechnol.* 36, 352–358
148. Na, D. *et al.* (2013) Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat. Biotechnol.* 31, 170–174
149. Crook, N.C. *et al.* (2014) Optimization of a yeast RNA interference system for controlling gene expression and enabling rapid metabolic engineering. *ACS Synth. Biol.* 3, 307–313
150. Li, Q. *et al.* (2013) Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* 152, 633–641
151. Wu, J. *et al.* (2017) Efficient *de novo* synthesis of resveratrol by metabolically engineered *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* 44, 1083–1095
152. Tao, S. *et al.* (2018) Regulation of ATP levels in *Escherichia coli* using CRISPR interference for enhanced pinocembrin production. *Microb. Cell Fact.* 17, 147
153. Chappell, J. *et al.* (2015) Creating small transcription activating RNAs. *Nat. Chem. Biol.* 11, 214–220
154. Gilbert, L.A. *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451
155. Liu, X.S. *et al.* (2016) Editing DNA methylation in the mammalian genome. *Cell* 167, 233–247.e217
156. Tapsin, S. *et al.* (2018) Genome-wide identification of natural RNA aptamers in prokaryotes and eukaryotes. *Nat. Commun.* 9, 1289
157. Yang, D. *et al.* (2018) Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 115, 9835–9844
158. Liu, Y. *et al.* (2014) Spatial modulation of key pathway enzymes by DNA-guided scaffold system and respiration chain engineering for improved *N*-acetylglucosamine production by *Bacillus subtilis*. *Metab. Eng.* 24, 61–69
159. Myhrvold, C. *et al.* (2016) Synthetic lipid-containing scaffolds enhance production by colocalizing enzymes. *ACS Synth. Biol.* 5, 1396–1403
160. Lewicka, A.J. *et al.* (2014) Fusion of pyruvate decarboxylase and alcohol dehydrogenase increases ethanol production in *Escherichia coli*. *ACS Synth. Biol.* 3, 976–978
161. Kerfeld, C.A. *et al.* (2018) Bacterial microcompartments. *Nat. Rev. Microbiol.* 16, 277–290
162. Shaw, A.J. *et al.* (2016) Metabolic engineering of microbial competitive advantage for industrial fermentation processes. *Science* 353, 583–586
163. Lennen, R.M. (2016) Benefits of selective feeding. *Science* 353, 542–543
164. Chan, C.T. *et al.* (2016) 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nat. Chem. Biol.* 12, 82–86
165. Mans, R. *et al.* (2018) Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production. *Curr. Opin. Biotechnol.* 50, 47–56
166. Zhu, Z. *et al.* (2018) Evolutionary engineering of industrial microorganisms-strategies and applications. *Appl. Microbiol. Biotechnol.* 102, 4615–4627
167. McCloskey, D. *et al.* (2018) Adaptive laboratory evolution resolves energy depletion to maintain high aromatic metabolite phenotypes in *Escherichia coli* strains lacking the phosphotransferase system. *Metab. Eng.* 48, 233–242

168. Luna-Flores, C.H. *et al.* (2017) Improved production of propionic acid using genome shuffling. *Biotechnol. J.* 12, 1600120
169. Zhang, Y.X. *et al.* (2002) Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* 415, 644–646
170. Alper, H. *et al.* (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* 314, 1565–1568
171. Tan, F. *et al.* (2016) Using global transcription machinery engineering (gTME) to improve ethanol tolerance of *Zymomonas mobilis*. *Microb. Cell Fact.* 15, 4
172. Amiram, M. *et al.* (2015) Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat. Biotechnol.* 33, 1272–1279
173. Chao, R. *et al.* (2017) Fully automated one-step synthesis of single-transcript TALEN pairs using a biological foundry. *ACS Synth. Biol.* 6, 678–685
174. Sun, N. and Zhao, H. (2013) Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.* 110, 1811–1821
175. Garst, A.D. *et al.* (2017) Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.* 35, 48–55
176. Mahr, R. *et al.* (2015) Biosensor-driven adaptive laboratory evolution of L-valine production in *Corynebacterium glutamicum*. *Metab. Eng.* 32, 184–194
177. Mundhada, H. *et al.* (2017) Increased production of L-serine in *Escherichia coli* through Adaptive Laboratory Evolution. *Metab. Eng.* 39, 141–150
178. Li, K. and Frost, J.W. (1998) Synthesis of vanillin from glucose. *J. Am. Chem. Soc.* 120, 10545–10546
179. Hansen, E.H. *et al.* (2009) *De novo* biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). *Appl. Environ. Microbiol.* 75, 2765–2774
180. Brochado, A.R. *et al.* (2010) Improved vanillin production in baker's yeast through *in silico* design. *Microb. Cell Fact.* 9, 84
181. Peplow, M. (2016) Synthetic biology's first malaria drug meets market resistance. *Nature* 530, 389–390
182. Horinouchi, T. *et al.* (2017) Improvement of isopropanol tolerance of *Escherichia coli* using adaptive laboratory evolution and omics technologies. *J. Biotechnol.* 255, 47–56
183. Thomson, N.M. *et al.* (2018) Efficient 3-hydroxybutyrate production by quiescent *Escherichia coli* microbial cell factories is facilitated by indole-induced proteomic and metabolomic changes. *Biotechnol. J.* 13, e1700571
184. Tietz, J.I. *et al.* (2017) A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* 13, 470–478
185. Islam, M.A. *et al.* (2017) Exploring biochemical pathways for mono-ethylene glycol (MEG) synthesis from synthesis gas. *Metab. Eng.* 41, 173–181
186. Delepine, B. *et al.* (2018) RetroPath2.0: a retrosynthesis workflow for metabolic engineers. *Metab. Eng.* 45, 158–170
187. Dias, O. *et al.* (2015) Reconstructing genome-scale metabolic models with *merlin*. *Nucleic Acids Res.* 43, 3899–3910
188. Wang, H. *et al.* (2018) RAVEN 2.0: a versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *PLoS Comput. Biol.* 14, e1006541
189. Mih, N. *et al.* (2018) ssbio: a Python framework for structural systems biology. *Bioinformatics* 34, 2155–2157
190. Salvy, P. *et al.* (2018) pyTFA and matTFA: a Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. *Bioinformatics* http://dx.doi.org/10.1093/bioinformatics/bty499
191. Cardoso, J.G.R. *et al.* (2018) Cameo: a python library for computer aided metabolic engineering and optimization of cell factories. *ACS Synth. Biol.* 7, 1163–1166
192. Heirendt, L. *et al*. (2018) Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3.0. *arXiv.org* arXiv:1710.04038v2 [q-bio.QM]
193. Chen, Z. *et al.* (2014) Deregulation of feedback inhibition of phosphoenolpyruvate carboxylase for improved lysine production in *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* 80, 1388–1393
194. Brunk, E. *et al.* (2018) Characterizing posttranslational modifications in prokaryotic metabolism using a multiscale workflow. *Proc. Natl. Acad. Sci. U. S. A.* 115, 11096–11101
195. Wang, H.H. *et al.* (2012) Genome-scale promoter engineering by coselection MAGE. *Nat. Methods* 9, 591–593
196. Lee, M.J. *et al.* (2018) Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm. *Nat. Chem. Biol.* 14, 142–147
197. Nielsen, A.A.K. *et al.* (2016) Genetic circuit design automation. *Science* 352, aac7341
198. Gao, X.J. *et al.* (2018) Programmable protein circuits in living cells. *Science* 361, 1252–1258