Advances in *de novo* strain design using integrated systems and synthetic biology tools

Chiam Yu Ng¹, Ali Khodayari¹, Anupam Chowdhury¹ and Costas D. Maranas

Address: Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, USA

Corresponding author: Costas D. Maranas (costas@psu.edu)

¹Joint first authors

Abstract

Recent efforts in expanding the range of biofuel and biorenewable molecules using microbial production hosts have focused on the introduction of non-native pathways in model organisms and the bio-prospecting of non-model organisms with desirable features. Current challenges lie in the assembly and coordinated expression of the (non-)native pathways and the elimination of competing pathways and undesirable regulation. Several systems and synthetic biology approaches providing contrasting top-down and bottom-up strategies, respectively, have been developed. In this review, we discuss recent advances in both *in silico* and experimental approaches for metabolic pathway design and engineering, with a critical assessment of their merits and remaining challenges.

Introduction

Microbial production has the unique advantage over chemical catalysis in that it can co-opt thousands of enzymes finely tuned by nature and leverage the host’s biological processes for cofactor regeneration, catalytic machinery assembly/disassembly and housekeeping functions. Advancement in metabolic engineering has increased the range of bio-based chemical products in microbial hosts, including therapeutics such as artemisinin [1], bioplastic precursors such as 1,4-butanediol [2], and biodiesel fatty esters and fatty acids [3]. Despite several success stories, only few metabolic engineering products achieve performance metrics that currently merit commercialization [4].

Increasing demands on maximizing production potential have highlighted the importance of developing tools that can identify more efficient pathways, both native and heterologous to a given host, from a (often given) substrate to the target chemical. Subsequently, metabolic intervention strategies are drawn to reconfigure the host metabolism for channeling additional flux towards the selected pathway and eliminating carbon and redox losses towards undesirable products, followed by the construction and evaluation of the strains. However, there are several challenges to the successful implementation of this design-build-test-learn loop (Figure 1). Enzymes are sensitive to temperature and pH and cannot be universally expressed in all hosts with a controllable rate of expression [5]. Another challenge lies in the successful implementation of computational predictions due to incomplete/erroneous modeling descriptions as well as the inability to precisely modulate gene expression to match model predictions. Furthermore, the current capacity to generate combinatorial variants far exceeds the throughput of screening.

© 2015. This manuscript version is made available under the Elsevier user license http://www.elsevier.com/open-access/userlicense/1.0/
In this review, we focus on recent advances in systems and synthetic biology for synthetic metabolic pathway design and optimization. We first describe the recently developed computational tools for identifying \textit{de novo} biosynthetic pathways. Next, we discuss computational stoichiometry-based and kinetic-based approaches for strain optimization. Finally, we discuss the recently developed synthetic biology and genome engineering techniques for synthetic pathway and network engineering.

Pathway prospecting for synthetic routes

Synthetic pathways from a source metabolite to a target chemical must satisfy a number of performance criteria such as (i) maximal use of native reactions [6] (Figure 1b, in orange), (ii) minimal number of reaction steps or equivalently total enzymatic load [7] (Figure 1b, in blue), (iii) maximization of product yield [8] (Figure 1b, in grey), (iv) cofactor balance in the overall pathway [6], and (v) thermodynamic feasibility of the overall pathway and individual steps [9]. \textit{A priori} assessment of whether these criteria are met requires knowledge of the metabolism of the host organism (e.g., codified as a genome-scale network (Figure 1a) and other database resources. Optimizing native pathways generally requires less effort as both catalytic components and regulatory structures are already in place [5]. In contrast, expression of non-native reactions is more complex but the upside is that they can in many cases significantly improve the yield of target products [10].

A number of pathway-prospecting tools have been developed recently (Table 1a) leveraging advances in computational power and availability of well-curated databases of metabolic reactions. Elementary Flux Mode (EFM) derived approaches relying on Linear Programming (LP) formulations can now be extended to genome-scale models and comprehensive reaction databases to search for \textit{de novo} pathways. Variations of this approach have been implemented for reconfiguring novel amino acid synthesis pathways in \textit{E. coli} [7], engineering hosts for biomass-coupled chemical production (e.g., SSDesign [11]) and designing novel pathways for CO\textsubscript{2}-fixation [12]. Alternatively, the computational intractability of EFMs in searching from thousands of reaction candidates [7] can be ameliorated by using graph-based tools for pathway design [13]. These tools can now rank pathways based on product yield [14] and cost of transcription and translation (e.g., DESHARKY [15]), as well as prevent selection of thermodynamically infeasible intermediate reactions (e.g., Metabolic Tinker [9]). In addition, atom-mapping information have also been incorporated to trace the fate of individual atoms (especially carbon) that filters out non-carbon transferring paths (e.g., Carbon Flux Path [16]).

Recent tools in pathway design use a set of “reaction rules” instead of lists of reactions to predict novel pathways without being restricted to previously catalogued reactions in nature. A retrosynthetic algorithm selects the intermediate metabolites between source and target chemical by satisfying the rules of chemical transformation defined by the set of reaction operators (e.g., BNICE [17], XTMS [18]). For example, in the recent GEM-Path approach [6], Biochemical Reaction Operators (BROs) serve as reaction templates for conversion of metabolites. Using an iterative algorithm to trace back from the target molecule one reaction at a time, metabolites are assigned to the BROs using a scoring mechanism based on how similar a metabolite is to the existing host metabolome. A reaction is accepted if it is present in a curated database, or is similar enough to an existing enzyme in the database to catalyze the putative reaction. The algorithm proceeds to identify the previous reaction in the linear pathway terminating when a metabolite present in the host metabolome is identified.

Despite enormous progress over the past few years, available pathway design procedures are generally restricted to only (near) linear routes from the source to the target metabolite. Linear pathway designs generally miss cyclic networks with potential for higher efficiency (both carbon and energy) of production. In addition, by restricting the degrees of freedom to just the source and target metabolite, the
identification of alternative co-reactants/co-products combinations are ignored. While post-processing efforts restore stoichiometry-balance of pathways [6], this may lead to designs with suboptimal carbon and energy efficiencies. Compatibility of a heterologous pathway with the metabolic host of interest is also often not adequately addressed at the design stage [19]. While some of the procedures minimize the number of heterologous enzymes [6,20], or choose enzymes phylogenetically closest to the host [21], there is no guarantee that the synthetic pathway would be host compatible. In addition, existing computational procedures do not directly assess the toxicity potential of intermediate metabolites. As more toxicity data for model organisms is collected (e.g., PanDaTox [22]), toxicity prediction tools (e.g., EDGE [23]) would increasingly become more commonplace in scoring synthetic pathways. Likewise, kinetic properties of the enzymes in the pathway would increasingly be queried to find the most active routes to the target chemical (e.g., DESHARKY [15]).

Modeling-driven pathway engineering
Once the designed pathway is introduced in the host strain, the metabolic fluxes need to be re-apportioned towards the target product (Figure 1c). Several optimization-based computational techniques have been developed to achieve this aim [24] (Table 1b). The scope of these approaches has been expanded through the use of synthetic biology tools. These predictive tools comprised of stoichiometry-only approaches [25], kinetic models [26] or hybrid combinations thereof [27] make quantitative predictions on metabolism upon metabolic interventions.

A number of efforts have integrated non-native pathways into the production host metabolic model by simply expanding the stoichiometric matrix using Flux Balance Analysis (FBA) techniques. For example, Proportional Flux Forcing (PFF) [28] was developed to explore the effect of substrate competition upon insertion of non-native genes using the GDLS algorithm. This is achieved by forcing a fixed fraction of the flux passing through the substrate into the alternative heterologous pathways formed by the introduced genes. This procedure was used for enhancing free fatty acid production in *E. coli*. In another effort, Yim *et al* [29] integrated a biopathway prediction algorithm with computational strain design protocols to identify synthetic pathways producing non-native products from common metabolic intermediates in *E. coli*. They first constructed an ensemble of 10,000 pathways producing 1,4-butanediol (14BDO) from mixed sugar streams. The best engineering strategies were then identified using the OptKnock algorithm [30] improving the yield of 14BDO production for the two best-ranked synthetic pathways.

Stoichiometry-based approaches are limited to steady-state conditions and are generally unable to describe the rate of reaction in terms of the underlying pool of metabolite concentrations and enzyme abundances. Therefore, the identified metabolic engineering strategies may not be implementable. For example, for a suggested up-regulation the corresponding enzymatic activity and metabolite concentrations may not be reachable and/or physiologically allowable. These shortcomings can potentially be addressed by kinetic models that directly track both enzyme levels and concentrations [31].

The application of kinetic-based models in synthetic biology, however, is still hampered by a number of challenges, chief among which are the paucity, *in vivo* applicability and universality of kinetic parameter data. In an effort to alleviate this problem, Farasat *et al* proposed SEAMAPs to build a kinetic model for a given modular synthetic pathway [32]. They used RBS Library Calculator to design minimal number of experiments, which varied expression of each enzyme in the pathway over 10,000-fold, to parameterize the kinetic model of neurosporene production pathway in *E. coli*. In another effort, a kinetic model was developed to identify the rate-limiting step of an *in vitro* ATP-free synthetic pathway for production of hydrogen from pretreated biomass sugars [33].
Successful implementation of kinetic expressions to guide metabolic interventions requires that regulatory interactions at the substrate, transcriptional, translational and post-translational levels are adequately captured [34]. Using metabolite concentration and enzyme activity as model variables, significant progress in the integration of substrate level regulatory interactions in kinetic models has been achieved [35]. Implementation of transcription level regulatory interactions in stoichiometry-based models is limited to Boolean representations or introduction of ad hoc constraints to shrink the flux ranges in concert with transcriptomic and proteomic data. This posture generally assumes that a positive correlation exists between metabolic flux and gene expression levels though there exists ample counter-examples [36]. Generally, the predictive accuracy of these approaches is highly condition dependent [37]. The scope of kinetic models can be further expanded to integrate transcription-level regulatory events. This can potentially be achieved using phenomenological Hill equations or partition functions to describe the rate of mRNA synthesis from a given promoter in terms of transcription factor (TF) activities [38]. This could ultimately enable the integration of transcriptional regulatory events with models of metabolism.

Kinetic-based modeling approaches show promise in capturing the dynamic behavior of metabolic pathways and regulatory interactions. Integration of such models with system-level omics data and computational metabolic engineering tools provides an avenue for understanding and subsequently optimizing the function of synthetic pathways. Robust model parameterization in response to genetic/environmental perturbations remains difficult. Many efforts are currently underway towards resolving this challenge by proposing more efficient optimization approaches [39], reducing parameters search space by structural analysis [40] and efficient sampling approaches [41].

Synthetic biology and genome engineering tools for implementation of pathway predictions

Ideally, the engineered strain should match as close as possible the desired flux distribution predicted through metabolic modeling. This requires among other considerations precise and reliable control of gene transcription and mRNA translation. Various genetic parts including promoters, RBSs, terminators, TFs and small regulatory RNA (sRNA) have been extensively characterized, offering an unprecedented range of parts for engineering gene expression. However, the performance of genetic parts is often non-conserved across different contexts (e.g., host, genetic, media) thus requiring performance re-assessment in the desirable conditions [42]. Alternatively, several computational techniques have been developed for designing context-specific sequences (e.g., RBS Calculator [43]) and selecting the optimum combination (e.g., OptCircuit [44]) of genetic parts (Table 1c).

Optimization of gene expression involves a vast design sequence space (i.e., $4^n$, where $n$ is the length of a transcription unit). In addition, several parameters can affect gene expression such as codon usage, tRNA availability, secondary structures, presence of RNases binding sites, internal Shine-Dalgarno sequences and repeats [45]. Some of these parameters are contradictory thus confounding the task of codon selection. For example, rare codons instead of common codons at the N-terminal tend to reduce secondary structures thus increasing the rate of protein translation [46]. A number of gene design tools attempt to integrate all these sometimes conflicting requirements [45], albeit with limited experimental verification. An alternative approach is to rely on high-throughput gene synthesis technology to generate large codon-usage variants set for a particular protein and then screen for high-expression variants, thus circumventing the need-to-know all design rules [47].

In addition to ensuring proper expression of all genes, the expression level of the entire synthetic pathway must also be carefully tuned to prevent imbalance in cellular resources (e.g., biomass precursors, proteins and redox equivalents [48]) and accumulation of toxic metabolites [49]. Combinatorial approaches geared towards optimizing the expression of all enzymes use various techniques ranging from targeting a rational
selection of genes to random editing at a genome-scale. Combinatorial DNA assembly techniques such as Gibson Assembly [50] and DNA Assembler [51] are commonly used to fuse genes or operons with libraries of regulatory parts (e.g., promoter [52-54], RBS [55] and copy number [55]). When the expression of a large number of genes must be manipulated, they are often partitioned into separate modules based on their functions to reduce the search space. Notably, Ajikumar et al varied the expression of methylerythritol-phosphate (MEP) pathway and taxadiene synthesis modules by changing their promoter and copy number [56]. Despite exploring just a small fraction of the entire combinatorial space, their combinatorial variants achieved up to 15,000-fold change in taxadiene production [56]. On a larger scale, the *Klebsiella oxytoca* nitrogen fixation gene cluster (103 parts) was refactored by employing a combination of combinatorial design and assembly approach [5].

Advances in high-throughput genome engineering techniques have accelerated the construction of large strain libraries (Figure 1e). Methods such as MAGE [57] and synthetic sRNAs [58] directly target tens of pre-selected genes with high specificity, whereas other techniques such as gTME [59], SCALEs [60] and TRMR [61] first generate libraries of strains with randomized genome-wide mutations and subsequently perform selection to identify the genotype conferring the desired traits. Both approaches complement one another, as the latter can be used for identifying subsets of genes required for wider range of expression tuning by the former. Recently, CRISPR-Cas9 (or dCas9) system has emerged as a versatile tool for multiplex genome engineering [62]. This system requires the design of highly orthogonal guide RNA(s) with minimal off-target activity. Collectively, these approaches are already capable of rapidly generating large combinatorial libraries, however, lack of high-throughput assays currently limits their applications to phenotypes with colorimetric (e.g., carotenoids production) or growth-based assays (e.g., metabolite tolerance).

As for other difficult to detect metabolites, intracellular biosensors have led to a number of success stories [63]. TF-based biosensors couple the expression of reporter proteins (typically fluorescent protein or antibiotic resistance marker) to the level of a metabolite of interest, enabling isolation of desired mutants via fluorescence-activated cell sorting (FACS) or positive growth selection [64,65]. Raman et al recently fine-tuned their biosensors so as only cells that produce target chemicals above a certain threshold would survive [66]. They performed FBA to identify target genes for MAGE genome engineering and then employed their biosensors to select for high producers, resulting in 36-fold and 22-fold improvement of naringenin and glucaric acid production, respectively.

While it is important to map desired traits to their genotype, sequencing and characterizing the entire combinatorial libraries remains cost-prohibitive. Microarray technology has been previously employed for parallel genotype-phenotype mapping of large gene knockout and overexpression libraries [60,61]. Recently, the tracking combinatorial engineered libraries (TRACE) method combines DNA assembly and next-generation sequencing for simultaneous genotype mapping of individual cells within a large combinatorial population [67]. TRACE was then employed to efficiently track both combinatorial diversity and evolution trajectory of a MAGE population [67].

All of these techniques have significantly improved pathway engineering and the strain construction process. In addition, they provide useful datasets for validation and refinement of computational tools and stoichiometric and kinetic metabolic models. For example, multiplex gene-knockout techniques can rapidly generate multiple knockout mutants to validate synthetic lethality predictions [68]. The refined model can then be used for predicting genetic manipulation strategies with better accuracy.

**Future perspectives**
With an ever expanding ability to construct, screen and characterize large mutant strain libraries, emphasis will likely shift on the ability to analyze large heterogeneous datasets for guiding the discovery of improved variants. Both machine learning inspired approaches that look for patterns in “big data” and predictive frameworks that seamlessly integrate different layers of biological processes would be needed. Moving beyond the scope of biological functions catalogued in nature, prospecting tools such as BNICE [17] and GEM-Path [6] could pro-actively be used to pinpoint desirable enzyme substrate and/or cofactor activity changes by harnessing enzyme plasticity. This has the potential for the discovery of more direct routes to target chemicals bypassing enzymatic or regulatory bottlenecks. Furthermore, the systematic discovery and pro-active elimination of undesirable secondary enzymatic functions that could drain carbon flux away from the main product could help shorten the strain design cycle [69]. Both tasks require the reliable re-design of enzymes either using de novo [70] or evolutionary techniques [71]. Several computational design techniques have been proposed for improving enzyme turnover number, substrate specificity, reduced allosteric inhibition, etc. [72-74], but reliable protein design remains elusive [75]. Improvements in our ability to model and predict the outcome of biological processes and networks coupled with falling DNA synthesis costs and efficient DNA assembly tools are bringing closer to fruition the dream of the design and assembly of synthetic production hosts uniquely tailored for bio-product biosynthesis.

Acknowledgements

This study was supported by the United States Department of Energy (DOE) grant DE-SC0008091.
### Table 1. Computational tools for pathway prediction, strain design and genetic circuit redesign

<table>
<thead>
<tr>
<th>Method</th>
<th>Merits</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNICE [17]</td>
<td>• The reaction rules are constructed by generalizing the bond breakage/formation at the active site using bond-electron matrix (BEM) information</td>
<td>These procedures define biochemical reaction Operators (BRO) or reaction rules to suggest conversion of a metabolite to a product. Starting from the target product, they use an iterative algorithm to identify existing and de novo reactions to retrace back to a source metabolite (or a host metabolome).</td>
</tr>
<tr>
<td></td>
<td>• Additional modules to curate the pathways based on thermodynamic analysis of the reactions</td>
<td></td>
</tr>
<tr>
<td>XTMS [18]</td>
<td>• Improves on an earlier RetroPath procedure by ranking pathways on a combined score based on pathway heterogeneity, enzyme promiscuity, and metabolite toxicity information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Uses molecular signature information to establish a common core of reaction rules</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The generality of the reaction-rule can be modified based on requirements of computational tractability</td>
<td></td>
</tr>
<tr>
<td>GEM-Path [6]</td>
<td>• The reaction rules are manually curated based on first three entries of EC classification</td>
<td>These procedures use Elementary Flux Modes (EFM) derived approach to convert a source metabolite to the target chemical. Using a Linear programming optimization formulation, they identify a stoichiometry-balanced path that minimizes the sum of metabolic flux required to produce a non-zero amount of target product.</td>
</tr>
<tr>
<td></td>
<td>• Additional module identifies reaction removals in the host to rank the pathways on their growth-coupled yield</td>
<td></td>
</tr>
<tr>
<td>Bar Evan et al [12]</td>
<td>• Explores from a database of reactions (KEGG) to construct minimal paths of conversion of a substrate to a product</td>
<td>These procedures use a bilevel optimization algorithm to identify a minimum set of heterologous reactions that induces the production of a non-native desired chemical in a host organism.</td>
</tr>
<tr>
<td></td>
<td>• Require additional manual curation to cofactor-balance the pathways</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Not automated for identification of alternate paths</td>
<td></td>
</tr>
<tr>
<td>Bordbar et al [7]</td>
<td>• EFMs in genome-scale networks</td>
<td>These methods use graph-search techniques to find a path between a source metabolite (or host metabolome) and a target metabolite by searching from a database of metabolic reactions. These are fast techniques that generally cannot ensure stoichiometry-balanced paths with no yield information of the products.</td>
</tr>
<tr>
<td></td>
<td>• Pathways are not ranked based on any metric</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Computationally challenging for implementation on database of reactions</td>
<td></td>
</tr>
<tr>
<td>SSdesign [11]</td>
<td>• Restricts the metabolic network to EFMs that ensure both growth-coupled and non-growth coupled modes of production of target chemical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Computationally intractable for large networks as the procedure explores through all the EFMs</td>
<td></td>
</tr>
<tr>
<td>OptStrain [8]</td>
<td>• All identified solutions ensure maximum theoretical yield of production of the target chemical</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Subsequent OptKnock step [30] step identifies reaction deletions in the existing network to couple target chemical with biomass production</td>
<td></td>
</tr>
<tr>
<td>SimOptStrain [20]</td>
<td>• Simultaneous identification of reaction deletion in host metabolism and addition of heterologous reactions combines both the OptStrain steps</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Engineering strategies with higher product yield overlooked by OptStrain can be identified</td>
<td></td>
</tr>
<tr>
<td>DESHARKY [15]</td>
<td>• Pathways scored based on transcriptional and translational cost of expressing heterologous genes in the host network</td>
<td></td>
</tr>
<tr>
<td>Metabolic tinker [9]</td>
<td>• Additional thermodynamic information of reactions is used to restrict infeasible pathway designs between a source and a target metabolite</td>
<td></td>
</tr>
<tr>
<td>Carbon Flux Path [16]</td>
<td>• Use atom-transition information for reactions to limit pathways involving carbon transfer</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Identifies KO strategies</td>
<td>Requires stoichiometry representation of the metabolism</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>OptKnock [30]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDLs [76]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OptORF [77]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OptForce [78]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emilio [79]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BiMOMA [20]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RobustKnock [80]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k-OptForce [81]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kamp and Klamt [82]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OptCircuit [44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh et al [83]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zomorrodi and Maranas [84]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBS Calculator [43]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure Legend

Figure 1. Pictorial overview of computational and experimental techniques for strain development and pathway engineering.

References


This work introduces the Gem-Path procedure, a retrosynthetic algorithm of designing de novo paths from the metabolome of a host organism to a non-native target product using pre-formulated reaction operators. Combined with stoichiometry balance and thermodynamics feasibility modules, this approach has a post-processing step for engineering the host for growth-coupled product.


This work describes a Mixed Integer Linear Programming (MILP) optimization formulation that can successfully identify all the shortest, linearly independent pathways in a genome-scale network. Along with discovering hidden transcriptional regulation controlling pathway flux, this method also identified novel minimal routes for amino acid synthesis in E. coli and S. cerevisiae.


This work extends an earlier Carbon Flux path approach to introduce atom-transition details between metabolites in reactions to identify paths that only involve carbon transfer between a source and a target metabolite.


This work develops an Expression-Dependent gene Effects (EDGE) procedure in which a MILP optimization problem predicts whether the overexpression of an in vivo (or non-native) gene would be deleterious/advantageous to the growth potential of a host organism. This approach was used to predict and verify “toxic genes” in several organisms and suggest a procedure of distinguishing cancerous cells from healthy one due to repressed expression of these “toxic genes”.


This study uses k-OptForce procedure to identify the minimal interventions that improve the yield of succinate production using a hybrid kinetic-stoichiometry model of *E. coli* metabolism.


This work presents a thermodynamic model to quantify the level of gene expression as a function of regulatory tuning variables using partition functions.


In this study, the Ensemble Modeling (EM) paradigm is integrated with continuation method for robustness analysis (EMRA) of the (non-)native pathways.


This work employs rationally designed synthetic small regulatory RNAs (sRNAs) for precise control of target gene expression levels. Using synthetic sRNAs, they improved the production of tyrosine by combinatorial knock-down of four genes across fourteen different E. coli strains. They further targeted over 130 genes, including essential genes, to optimize cadaverine production.


This work employed FBA-guided multiplex genome engineering and growth-coupled transcription factor-based biosensors to optimize narigenin and glucaric acid biosynthetic pathway.

This study presented TRACE, an approach that performs genome mutation mapping of $>10^5$ individual cells from combinatorially engineered populations.


