Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

New synthetic biology tools for metabolic control

Xueqin Lv[1,3,*], Angeles Hueso-Gil[2,*], Xinyu Bi[1,3], Yaokang Wu[1,3], Yanfeng Liu[1,3], Long Liu[1,3] and Rodrigo Ledesma-Amaro[2]

In industrial bioprocesses, microbial metabolism dictates the product yields, and therefore, our capacity to control it has an enormous potential to help us move towards a bio-based economy. The rapid development of multiomics data has accelerated our systematic understanding of complex metabolic regulatory mechanisms, which allow us to develop tools to manipulate them. In the last few years, machine learning-based metabolic modeling, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) derived synthetic biology tools, and synthetic genetic circuits have been widely used to control the metabolism of microorganisms, manipulate gene expression, and build synthetic pathways for bioproduction. This review describes the latest developments for metabolic control, and focuses on the trends and challenges of metabolic engineering strategies.

Addresses
[1] Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China
[2] Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London, London SW72AZ, UK
[3] Science Center for Future Foods, Jiangnan University, Wuxi 214122, China

Corresponding authors: Long Liu (longliu@jiangnan.edu.cn), Rodrigo Ledesma-Amaro (r.ledesma-amaro@imperial.ac.uk)
[*] These authors contributed equally to this work.

Current Opinion in Biotechnology 2022, **76**:102724

This review comes from a themed issue on **Analytical Biotechnology**

Edited by **Peng Xu** and **Kang Zhou**

https://doi.org/10.1016/j.copbio.2022.102724

Introduction

Compounds of industrial interest can be biomanufactured by those microorganisms that naturally synthetize them, but usually, the production levels of wild type strains are low. Therefore, metabolic engineering is often required to optimize fluxes and ultimately bioprocesses. Although in some cases the deletion, substitution or addition of a gene can be enough to increase the accumulation of a compound, in most of the cases the complexity of metabolism and its regulation require further designs [1,2]. In order to achieve the desired metabolic control, a wide range of sophisticated tools have been developed to date [3]. Within the diversity of methods available, in this work, we review three areas that have proven successful in the last years: machine learning-based models, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) tools, and genetic circuits. These three areas are of particular importance because of their potential in overcoming some of the current challenges in developing efficient cell factories, such as our scarce understanding of metabolic regulation, the limitation of our strain engineering techniques and the lack of capacity to create new regulatory networks with desired behaviors.

With the purpose of predicting the appropriate control of biological functions, whole-cell metabolic models and machine learning have helped us better understand microbial bioprocesses when there is a lack of experimental data, and they have guided engineering strategies [3].

In addition, once the desired genetic targets have been identified *in silico*, microorganisms must be manipulated to optimize their performance. This can be achieved by mechanisms that control synthetic regulatory networks, also known as genetic circuits [4,5]. These circuits present different levels of complexity depending on the application, some of them showing intricate dynamics designed to optimize the synthesis of the compound of interest. Synthetic genetic circuits have the advantage to be programmable and autonomous. Its modular nature gives them stability and predictability. As a result, they have contributed to different areas such as bioproduction, biodiagnosis, microbiome regulation, and biocontainment [6].

In recent years, CRISPR-based tools have become a favorite choice to regulate metabolism, either independently or as part of a genetic circuit. CRISPR was originally discovered as a prokaryotic immune mechanism that targeted potentially harmful exogenous DNA or RNA to silence it and avoid damage [7–9]. The discovery was later applied as a DNA editing technology, and since then it has become one of the most powerful and used tools in biotechnology [10]. Its mechanism of action involves an endonuclease, which is

usually either Cas12 (also known as Cpf1) or Cas9. These nucleases introduce a Double Strand Break (DSB) led by a guide RNA (gRNA) that targets a specific sequence. The following repair can be resolved by endogenous mechanisms. In addition, nuclease-deficient version of Cas12 or Cas9 can be used to perform CRISPRa (activation) and CRISPRi (inhibition), techniques that either upregulate or downregulate gene expression, respectively [11].

In this review, we summarize the current advances in machine learning in relation to metabolic models, the developments in the area of synthetic genetic circuits, and the uses of CRISPR technologies to control metabolism.

Machine learning-based metabolic modeling

With the development of omics technologies, metabolic models are no longer limited to simple gene-protein-reaction interactions. Constraint-based modeling can integrate other factors, such as thermodynamics [12], dynamics [13], gene expression matrices [14], environmental and genetic relationships [5], and metabolic regulation of whole-cell models [3]. The use of machine learning has the potential to facilitate multi-omics data analysis and the building of advanced metabolic network models. In this section of the article, we summarize the applications of machine learning in metabolic modeling.

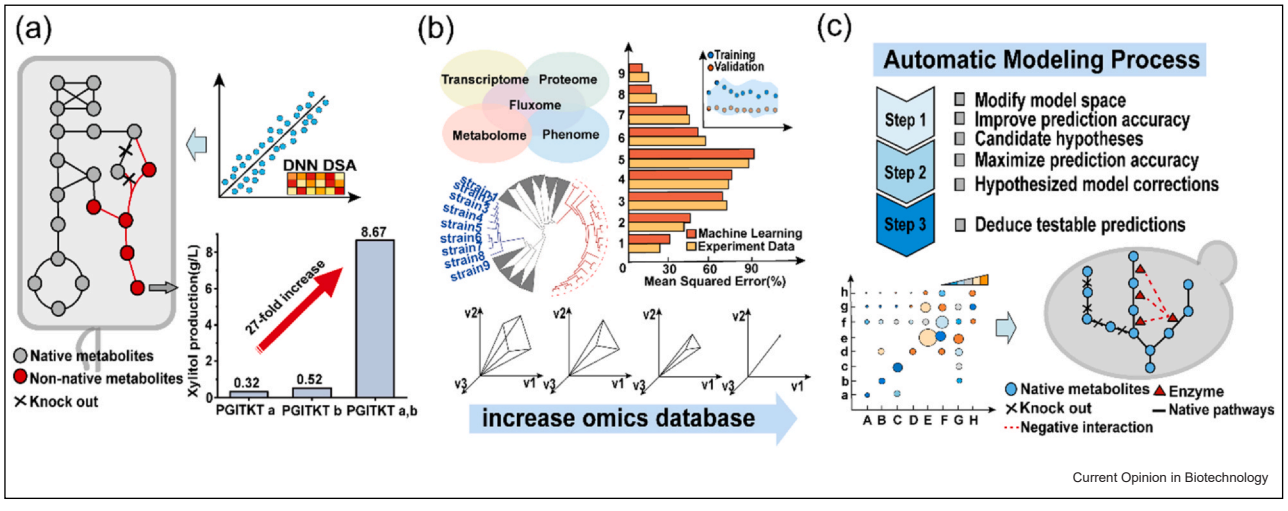
Model-driven production improvement

Metabolic model prediction based on machine learning can help to identify target genes to improve the production of metabolites of interest, as it has been proven in the production of xylitol [15] (Figure 1a and Table 1). In another example, the training sets of the two machine learning algorithms, the automatic recommendation tool (ART) and TeselaGen EVOLVE, were provided by the developed high-throughput biosensor. Then, the data were combined with the genome-scale metabolic model and used to predict the optimal metabolic pathways toward the production of tryptophan. The final results showed that this hybrid approach increased tryptophan titer and productivity by 74% and 43%, respectively [16•]. In a different work, analyzing omics data through the Bayesian factor model in order to design metabolic pathways was also an effective option to increase the biosynthesis of rhamnolipids [17]. Predicting the phenotypes of genetically modified strains under different growth conditions was essential for the adequate development of the metabolic model. As another example of improving the yield of target products, the multiomics model and analytics (MOMA) platform successfully identified key metabolic steps of eight different strains and resulted in an almost twofold increase in the production of isopentenol, with an absolute titer of 920 mg/L [18] (Figure 1c).

Improving model prediction accuracy

Once the metabolic model is established, improving its accuracy becomes a major challenge (Table 1). To this end, an automatic metabolic model integration drive (AMMEDEUS) was developed, which provides the capacity to minimize the errors between model predictions

Figure 1



Representative applications of machine learning-based metabolic mode. **(a)** Model-driven production promotion. In *E.coli* K-12 MG1655, a new method consisting of Deep Neural Network (DNN) and Differential Search Algorithm (DSA) is used to predict the best gene knockout pathway for the model to increase xylitol production. **(b)** Improving model prediction accuracy. Among the 29 bacterial species, using the method of automated metabolic model ensemble-driven elimination of uncertainty with statistical learning (AMMEDEUS) to improve gene importance prediction. **(c)** Exploring the potential metabolic route. In *S. cerevisiae*, machine learning methods are used to coordinate empirical interaction data and model predictions.

Table 1

Machine learning methods currently applied in metabolic models of industrial microorganisms.

Objective	Model type	Machine learning method	Analysis method	Ref.
Model-driven production promotion	Constraint-based model	DNN, DSA	Flux balance analysis, gene knockout	[15]
	Constraint-based model	ART, TeselaGen EVOLVE	Flux balance analysis	[16]
	Kinetic model and multiomics constraint-based model	Bayesian factor modeling	Linlog kinetics	[17]
	Multiomics constraint-based model	GEESE, β-VAE, artificial neural networks	Flux balance analysis	[18]
Improving model prediction accuracy	Constraint-based model	Random forest, Scikit-Learn	pflux balance analysis	[30]
	Constraint-based model	RMLR	Flux balance analysis	[20]
	Dynamic constraint-based model	Decision tree, CART algorithm	TFBA	[21]
	Kinetic model and Multiomics constraint-based model	GA, Scikit-learn	Partial least squares	[22]
	ME-model	DNN, random forest	Flux balance analysis, metabolic flux analysis	[23]
	Multiomics constraint-based model	Principal component analysis	Flux balance analysis, Monte Carlo sampling, gene knockout	[24]
	Transcriptional constraint-based model	Support vector regression, random forest, artificial neural networks, BEMKL, bagged random forest, multimodal artificial neural network, sparse group lasso, NSGA-II, iterative random forests, Shapley additive explanations	Parsimonious flux balance analysis	[25]
Explore the potential metabolic route	Constraint-based model	Support vector machine, k-nearest neighbors, decision tree	MOMA, 13C metabolic flux analysis	[26]
	Constraint-based model	dynEMA, dynEMR-DA, PEMA, principal component analysis, NPLS-DA	N-way, triple cross-validation	[27]
	Dynamic constraint-based model	PMFA, principal component analysis, SPMFA	Flux balance analysis	[28]
	Multiomics constraint-based model	GEESE, β-VAE, artificial neural networks	Flux balance analysis	[18]
	Constraint-based model	Random forest, logistic regression, HC	Flux balance analysis	[29]

ART: Automated Recommendation Tool, BEMKL: Bayesian efficient multiple-kernel learning, dynEMA: dynamic elementary mode analysis, dynEMR-DA: dynamic elementary mode regression discriminant analysis, ME-model: metabolism and gene expression genome-scale metabolic model, NPLS-DA: N-way partial least squares regression discriminant analysis, NSGA-II: non-dominated sorting genetic algorithm II, SPMFA: sparse principal metabolic flux mode analysis, TFBA: thermodynamics-based metabolic flux analysis.

and experimental data within different bacteria [19] (Figure 1b). Machine learning algorithms, such as regularized multinomial logistic regression (RMLR) [20], principal elementary mode analysis (PEMA) [21], and genetic algorithms (GA) [22], can be used to comprehensively explore pathways that affect cell metabolism, therefore with potential to be used to improve strain performance. Importantly, machine learning has actively been deployed to integrate metabolic models with omics data such as genomics, transcriptome, and proteomics, which can be useful for characterizing cell growth and metabolism [23]. In *Escherichia coli* K-12 MG1665, a flux-coupled metabolic subnetwork feature was introduced to improve the prediction of essential genes using a vector machines-based learning strategy. Following this approach, the prediction accuracy of essential response-gene pairs was increased up to 94.28% [24]. In addition, a multiview neural network combining fluxomics and transcriptomics data has been developed, whose prediction accuracy was validated by the use of 27 models together with machine learning methods. In the end, this neural network improved the prediction accuracy of the model while simultaneously analyzing the relationship between gene regulation and metabolic flux [25•].

Exploring potential metabolic biosynthesis pathways

Oftentimes, the metabolic pathway that synthetizes a target product is unknown. The exploration of potential metabolic pathways requires massive omics data and well-established platform analysis. In those cases, metabolic models have proven useful to predict feasible synthesis routes. The MFlux platform was constructed by non-steady-state flux distribution analysis, which can reasonably predict flux groups based on bacterial species, substrate types, growth rates, oxygen conditions, and culture methods, to precisely reveal unknown inter-relationships in metabolic networks [27,28]. Furthermore, principal metabolic flux pattern analysis (PMFA) can resolve differences in gene expression or flux data in undefined metabolic networks by coupling stoichiometric flux analysis and principal component analysis. The experimental results show that PMFA can accurately identify six mitochondrial pathways in response to changes in oxygen by analyzing the culture data of the metabolic network in *Saccharomyces cerevisiae* [26] (Table 1). In order to explore the unknown information between gene expression data and metabolic phenotypes, gene expression latEnt space encoder (GEESE), which is a framework based on deep learning, trains β-Variational Autoencoder (β-VAE) by using deep generative models in order to recognize metabolic data in unknown environments [18]. In order to explore the interaction between genes, machine learning methods (random forest and logistic regression, hierarchical clustering (HC)) have been used to analyze the gene epistasis interaction spectrum. As a result of fusing genetic interaction data with the model, an incorrect annotation in the NAD biosynthesis pathway in *Saccharomyces cerevisiae* was discovered [27] (Figure 1c).

CRISPR derived synthetic biology tools to control metabolism

The applications of CRISPR in the metabolic engineering are vast for both prokaryotes and eukaryotes, including bacteria, yeast, and filamentous fungus [28–30]. Additionally, CRISPR can also help to control the behavior of mammal cell cultures [31] or to improve performance of crops by engineering plant cells [32].

Beyond its utility for creating deletions, substitutions or point mutations, CRISPR technologies have been extended for further applications. For example, due to its ability to introduce DSB in a targeted place in the genome, CRISPR-Cas9 has been used for counter-selection in low-efficiency editing processes. This efficiency could be increased by combining CRISPR with previous protocols for deletion or substitution [33] or with ssDNA recombineering tools such as MAGE [34].

In addition, CRISPR can be multiplexed, enabling the simultaneous manipulation of multiple genes. This is of special interest when more than one gene needs to be edited or regulated at a given time to allow complex rewiring of pathways. Multiplexing has been carried out by controlling the expression of several gRNAs, processing the gRNA arrays using different techniques such as an RNase III, ribozymes or endonucleases like Cys4. These strategies have also been adapted to a wide range of microorganisms [11,29,30].

CRISPRi and CRISPRa

Additional modifications of Cas proteins have expanded the possibilities of CRISPR as a transcription inhibition tool CRISPRi or activation mechanism CRISPRa [11]. Both CRISPRi and CRISPRa are based on the use of deactivated forms of the Cas proteins, achieved by point mutations that remove the nuclease activity.

CRISPRa allows the activation of genes when an activator domain is bound to the Cas protein. For example, in prokaryotes, it can be accomplished when Cas9 is fused to the RNAP omega subunit [35] or AsiA, a phague activator protein [36]. Other possibility to achieve CRISPRa consists of binding an RNA recruiting scaffold to the gRNA to induce the activation of sigma54 and sigma70 promoters. For eukaryotes, there are similar strategies for CRISPRa by adding fusion proteins to Cas9, other proteins joined to the gRNA or using both approaches in parallel [37].

CRISPRi allows the repression of gene expression by targeting deactivated Cas proteins to either the promoter region of the gene or its coding sequence, effect that can

be enhanced by the fusion of repression domains. Their action consists in blocking the RNA polymerase, thus reducing gene expression [38,39]. CRISPRi have offered the possibility to regulate the expression of selected genes, but also have helped in genome imaging, DNA looping or epigenetic modifications [11].

The conditional silencing of genes using CRISPRi has permitted the study of the function of certain metabolic genes that become deleterious when fully deleted. Hence, it is being used as an important tool for metabolic rewiring [33,40]. Furthermore, inducible CRISPRi mechanisms can be used to achieve a dynamic regulation of fluxes. To improve lactate production in *Synechocystis* sp. PCC 6803, the *gltA* gene was targeted with CRISPRi to decouple growth from production, as CO_2 uptake during growth phase was impairing the production [41]. Using this strategy, both stages could be alternated, increasing the lactate accumulation up to 1 g/L. This decoupling strategy using CRISPRi can benefit the scalability of a bioprocess. Dynamic control of the genes *gltA*, *zwf* and *fabI* in *Escherichia coli*, especially during stationary phase, improved the robustness for scalability and allowed the increase of xylitol and citramalate up to 200 g/L and 125 g/L respectively [42]. Another possibility is to convert this conditional induction of CRISPRi in an autonomous switch, letting the microorganism to self-regulate its activity for an optimal performance. For example, using a stationary phase promoter, a protein degradation tag and CRISPRi technology, genes involved in the shikimic acid and glutamic acid production could be repressed depending on the growth phase of *E. coli*, increasing its production titer to 21 and 26 g/L, respectively [43].

Machine learning, OMICS data and models to complement CRISPRi application

The specificity of the gRNA determines how successful a targeting is. This specificity influences its efficiency to recognize the complementary sequence and its ability to produce undesired off-target effects. For balancing both parameters, data analysis and machine learning have helped the design gRNAs, focusing on the energy level (N_{min}) that produces the R loop, a structure formed by the gRNA and its target that is necessary for the recognition of the Cas protein. Therefore, the probability of on-target and off-target bindings of the gRNA can be predicted and improved (Figure 2) [44].

The combination of OMICS data together with CRISPRi has enabled the transitional regulation of pathways in order to increase the production of a certain compound or to investigate fluxes [45]. For example, combining a proteomic analysis with a selective silencing of certain genes, antibiotic production by *Amycolatopsis balhimycina* has been studied. A proteomic assay comparing protein profile during and before antibiotic production shed light on the genes that could be suppressed in order to redirect the metabolism of sugars to the production of balhimycin [46]. In addition to its use for bioproduction optimization, multi-OMICS data in combination with CRISPRi was utilized for the conditional silencing of genes to study the general metabolic landscape of *E. coli* [40]. The study of 30 strains where CRISPRi was induced allowed the characterization of robustness and fitness buffering in the metabolic network.

Furthermore, the use of models can help us select better target genes for the increase of production yields. Using a core kinetic model, certain genes were predicted to influence acetyl-CoA production in *P. putida*. After targeting *accA*, *accC* and *gltA* for conditional silencing using CRISPRi, acetyl-CoA content was increased up to eightfold compared to the wild type. The engineered strains were later used for the synthesis of poly(3-hydroxybutyrate) in bioreactors, increasing by fivefold the production over the *wild type* [47].

Programmable metabolic control enabled by synthetic genetic circuits

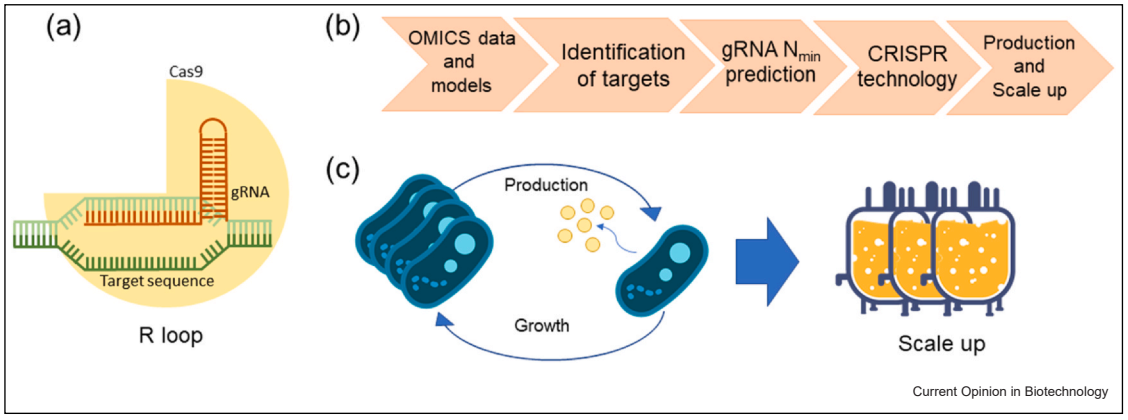
Synthetic genetic circuits, based on concepts taken from the electrical engineering field (like boolean logic gates), can achieve a programable, meticulous, and sophisticated control of cell metabolism [1].

Composition and working principles of synthetic genetic circuits

In synthetic genetic circuits, input signals can be firstly detected by a biosensor, which is followed by information computing and processing. After that, output signals are generated by an actuator to perform desired biological functions (Figure 3a). On the one hand, open-loop circuits, which are induced by environmental factors either manually controlled such as light [48] and temperature [49], or spontaneously controlled such as changes in cell density [50] oxygen [2] and pH [51,52], can be used to control cell metabolism and morphology (Figure 3b). On the other hand, closed-loop circuits equipped with biosensors that respond to intermediate metabolite [53,54] or end-product [55–57] can perform autonomous and continuous feedback control on cell metabolism and growth (Figure 3b).

Before executing desired biological functions, the acquired signals need to be processed and computed by the layered logical operation circuits, which are composed of regulators that act at DNA, transcription, post-transcriptional, or protein level. For example, genetic circuit design automation has been implemented with the help of characterized repressor-based NOT gates [58,59••], thus allowing the construction of sophisticated circuits performed as a digital display by a compiled software [4]. In another example, a set of orthogonal

Figure 2

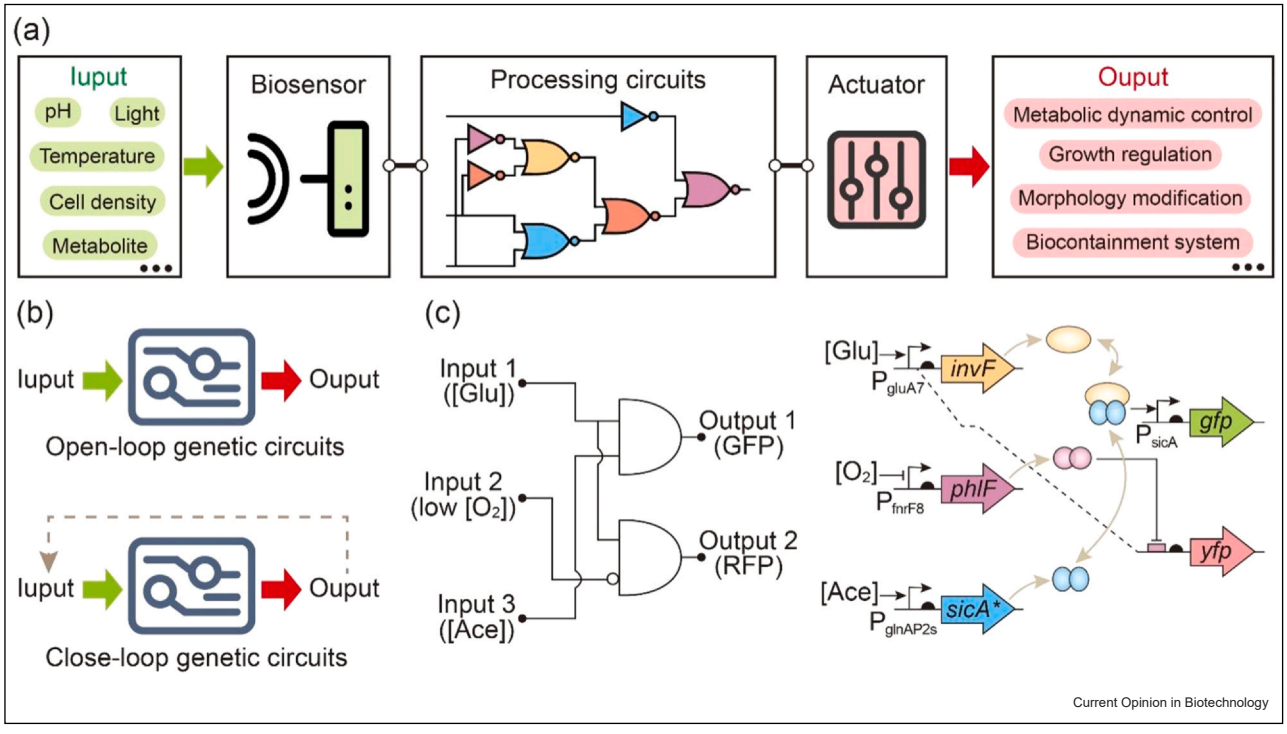


CRISPR technology for Metabolic Engineering. **(a)** Representation of R loop formed between gRNA and the target to be edited or silencing that is recognized by Cas protein. **(b)** Proposed workflow to integrate omics data, models, machine learning, and CRISPR for an efficient control of metabolic functions. **(c)** How CRISPRi can uncouple and alternate growth with bioproduction for a right scale-up of the process.

AND gates have been generated from engineered activator–chaperone pairs [60]. Using these transcriptional logic gates, multiple inputs including feedstock, dissolved oxygen, and byproduct accumulation can be wired together [2] (Figure 3c).

In addition to transcriptional factors, other regulatory mechanisms including trans-encoded small RNAs [56], antisense transcription [61], and CRISPRi/a [53] have also been wired into logic transduction. CRISPRi offers a simple and elegant way to make genetic devices such

Figure 3



Synthetic genetic circuit enabled programmable metabolic control. **(a)** Typical structure of synthetic genetic circuits used for metabolic regulation. **(b)** Comparison between open-loop genetic circuits and close-loop genetic circuits. **(c)** Design and construction of transcription factor-based synthetic genetic circuits.

as toggle switches, oscillators, and incoherent feed-forward loops, which demonstrates its potential to achieve complex applications [62•]. In addition, protein-level regulation such as protease-based controllable protein degradation allows quicker kinetics and ensures a timely response of circuits [63].

Application of synthetic genetic circuits in microbial cell factories

Synthetic genetic circuits can be used to control genes that alter and optimize metabolic fluxes. Improved microbial cell factories can be achieved by using metabolic dynamic pathway regulation, growth and morphology control or population control.

A key challenge in microbial factory design is to balance the distribution of metabolic fluxes to avoid the accumulation of undesired metabolites and growth depletion. To achieve this goal, dynamic pathway regulation strategies have been employed [64]. For instance, bifunctional closed-loop feedback circuits composed of metabolic responsive biosensor and CRISPRi-based NOT gate have been constructed with the purpose of upregulating desired synthetic pathways and down-regulating competitive reactions. Additionally, those circuits were successfully used to achieve the autonomous regulation and dynamic control of intracellular metabolic fluxes, and the N-acetylglucosamine titer increased from 81.7 g/L to 131.6 g/L in a 15-L fed-batch bioreactor [61].

The lifespan of a cell can be also reprogrammed by synthetic genetic circuits designed to regulate parameters like cell size, generation time, and stress tolerance. As an example, in order to enlarge the cell size of *E. coli* for increasing the accumulation capacity of poly lactate-co-3-hydroxybutyrate, the replicative lifespan of *E. coli* was engineered with an open-loop two-output recombinase-based state machine (a system that exists in any of a number of states, in which transitions between states are controlled by inputs) [65]. Furthermore, they have also built an open-loop multioutput recombinase-based state machine to change the chronological lifespan of *E. coli*, leading to the highest titer of butyrate (29.8 g/L) [65].

Stress resistance circuits can also be introduced to improve the stability of engineered strain during the production process. As an example, self-responsive pH adaptable circuits were constructed using base-responsive and acid-responsive promoters [51]. Furthermore, closed-loop addiction circuits able to couple growth and production by linking growth-associated genes and end-product biosensors have also been proposed to avoid population phenotypic heterogeneity and negative mutation accumulation [57,66]. For example, a mevalonic acid biosensor was used to control the expression of two nonconditionally essential genes, and high-yield mevalonic acid production was retained through 95 generations of cultivation, which is equivalent to a 200 m^3 industrial-scale production [67].

Moreover, to prevent the leakage of engineered microorganisms into the environment, biocontainment systems have been designed and introduced with the aid of circuits that could prevent cell growth [52]. In conclusion, reprogrammed metabolism by these genetic circuits not only gives the microbial cell factories efficient production capacity but also ensures their robustness, stability, and safety.

Conclusions and future perspectives

Advances in synthetic biology have contributed to metabolic control strategies and accelerated the development of efficient microbial cell factories. In this review, recent findings for the metabolic regulation of engineered cell are reviewed, focusing on three aspects: metabolic model, CRISPR technology, and genetic circuits. While the methods described above have been used to improve production, additional efforts are needed to widen their applications. We still need to deepen our understanding of biological systems and enhance our capacity to regulate metabolism at will. We expect that the development of whole-cell models of industrial microorganisms will enable the optimization of intracellular resources allocation in microbial cell factories, ultimately improving bioproduction [68]. In addition, it is easy to overlook that the commonly used CRISPR-based metabolic control processes require elaborate debugging due to their complexity. For example, negative feedback on dCas9-based circuits may be implemented to avoid the overload of CRISPRi circuits resulted from the competition between multiple sgRNAs [69•]. Moreover, the burden generated by introducing heterologous synthetic genetic circuits into the cell chassis is not negligible [70], and needs to be studied and overcome.

Author contributions

X.Q.L., A.H.-G., X.Y.B., and Y.K.W. completed the collection and analysis of relevant literatures and the writing of the first draft. Y.F.L., R.L.-A., and L.L. revised the manuscript. R.L.-A. and L.L. designed the manuscript. All authors contributed to the manuscript.

CRediT authorship contribution statement

Xueqin Lv: Writing – original draft, Investigation. Angeles Hueso-Gil: Writing – original draft, Investigation. Xinyu Bi: Writing – original draft. Yaokang Wu: Writing – original draft. Yanfeng Liu: Writing – review & editing. Long Liu: Conceptualization, Writing – review & editing, Funding acquisition. Rodrigo Ledesma-Amaro: Conceptualization, Writing – review & editing.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2020YFA0908300, 2018YFA0900300); the National Natural Science Foundation of China (32070085); the Fundamental Research Funds for the Central Universities (JUSRP52019A); the China National Postdoctoral Program for Innovative Talents (BX2021113). This work also received funding from Biotechnology and Biological Sciences Research Council (BBSRC) (BB/R01602X/1, 19-ERACoBioTech- 33 SyCoLim BB/T011408/1 and BBSRC BB/T013176/1) British Council 527429894, Newton Advanced Fellowship (NAF\R1\201187), Yeast4Bio Cost Action 18229, European Research Council (ERC) (DEUSBIO — 949080), the Bio-based Industries Joint (PERFECOAT — 101022370) under the European Union's Horizon 2020 research and innovation programme, and Fundación Alfonso Martin Escudero Spanish fellowship.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. English MA, Gayet RV, Collins JJ: **Designing biological circuits: synthetic biology within the operon model and beyond**. *Annu Rev Biochem* 2021, **90**:221-244.

2. Moser F, Espah Borujeni A, Ghodasara AN, Cameron E, Park Y, Voigt CA: **Dynamic control of endogenous metabolism with combinatorial logic circuits**. *Mol Syst Biol* 2018, **14**:e8605.

3. Karr Jonathan R, Sanghvi Jayodita C, Macklin Derek N, Gutschow Miriam V, Jacobs Jared M, Bolival B Jr., Assad-Garcia N, Glass John I, Covert Markus W: **A whole-cell computational model predicts phenotype from genotype**. *Cell* 2012, **150**:389-401.

4. Shin J, Zhang S, Der BS, Nielsen AA, Voigt CA: **Programming *Escherichia coli* to function as a digital display**. *Mol Syst Biol* 2020, **16**:e9401.

5. Carrera J, Estrela R, Luo J, Rai N, Tsoukalas A, Tagkopoulos I: **An integrative, multi-scale, genome-wide model reveals the phenotypic landscape of *Escherichia coli***. *Mol Syst Biol* 2014, **10**:735.

6. Xia PF, Ling H, Foo JL, Chang MW: **Synthetic genetic circuits for programmable biological functionalities**. *Biotechnol Adv* 2019, **37**:107393.

7. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E: **Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements**. *J Mol Evol* 2005, **60**:174-182.

8. Mojica FJ, Díez-Villaseñor C, Soria E, Juez G: **Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria**. *Mol Microbiol* 2000, **36**:244-246.

9. Jansen R, Embden JD, Gaastra W, Schouls LM: **Identification of genes that are associated with DNA repeats in prokaryotes**. *Mol Microbiol* 2002, **43**:1565-1575.

10. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity**. *Science* 2012, **337**:816-821.

11. McCarty NS, Graham AE, Studená L, Ledesma-Amaro R: **Multiplexed CRISPR technologies for gene editing and transcriptional regulation**. *Nat Commun* 2020, **11**:1281.

12. Dai Z, Locasale JW: **Thermodynamic constraints on the regulation of metabolic fluxes**. *J Biol Chem* 2018, **293**:19725-19739.

13. Khodayari A, Maranas CD: **A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains**. *Nat Commun* 2016, **7**:13806.

14. Chen K, Gao Y, Mih N, O'Brien EJ, Yang L, Palsson BO: **Thermosensitivity of growth is determined by chaperone-mediated proteome reallocation**. *Proc Natl Acad Sci USA* 2017, **114**:11548-11553.

15. Mohmad Yousoff SN, Baharin A, Abdullah A: **Differential search algorithm in deep neural network for the predictive analysis of xylitol production in *Escherichia coli***. In *Modeling Design and Simulation of Systems*. Edited by Mohamed Ali MS, Wahid H, Mohd Subha NA, Sahlan S, Md, Yunus MA, Wahap AR. Springer; 2017:53-67.

16. • Zhang J, Petersen SD, Radivojevic T, Ramirez A, Pérez-Manríquez A, Abeliuk E, Sánchez BJ, Costello Z, Chen Y, Fero MJ, *et al.*: **Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism**. *Nat Commun* 2020, **11**:4880.
The development of a biosensor can sample multiple sets of data of 250 strains and provide high-quality training sets for machine learning. The training sets of two machine learning algorithms, the ART and TeselaGen EVOLVE, were used to discover the best metabolic pathway of tryptophan in *E. coli*.

17. St John PC, Strutz J, Broadbelt LJ, Tyo KEJ, Bomble YJ: **Bayesian inference of metabolic kinetics from genome-scale multiomics data**. *PLoS Comput Biol* 2019, **15**:e1007424.

18. Barsacchi M, Terre HA, Lió P: **GEESE: metabolically driven latent space learning for gene expression data**. *bioRxiv* 2018,365643.

19. Medlock GL, Papin JA: **Guiding the refinement of biochemical knowledgebases with ensembles of metabolic networks and machine learning**. *Cell Syst* 2020, **10**:109-119 e103.

20. Sridhara V, Meyer AG, Rai P, Barrick JE, Ravikumar P, Segrè D, Wilke CO: **Predicting growth conditions from internal metabolic fluxes in an in-silico model of *E. coli***. *PLoS One* 2014, **9**:e114608.

21. Folch-Fortuny A, Marques R, Isidro IA, Oliveira R, Ferrer A: **Principal elementary mode analysis (PEMA)**. *Mol Biosyst* 2016, **12**:737-746.

22. Dash S, Khodayari A, Zhou J, Holwerda EK, Olson DG, Lynd LR, Maranas CD: **Development of a core *Clostridium thermocellum* kinetic metabolic model consistent with multiple genetic perturbations**. *Biotechnol Biofuels* 2017, **10**:108.

23. Heckmann D, Lloyd CJ, Mih N, Ha Y, Zielinski DC, Haiman ZB, Desouki AA, Lercher MJ, Palsson BO: **Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models**. *Nat Commun* 2018, **9**:5252.

24. Nandi S, Subramanian A, Sarkar RR: **An integrative machine learning strategy for improved prediction of essential genes in *Escherichia coli* metabolism using flux-coupled features**. *Mol Biosyst* 2017, **13**:1584-1596.

25. • Culley C, Vijayakumar S, Zampieri G, Angione C: **A mechanism-aware and multiomic machine-learning pipeline characterizes yeast cell growth**. *Proc Natl Acad Sci USA* 2020, **117**:18869-18879.
A multimodal machine learning framework combining transcriptomics data and metabolic models was developed. The framework used a variety of machine learning algorithms to improve the prediction accuracy of the model and explore the unknown relationship between genes and metabolism in *S. cerevisiae*.

26. Bhadra S, Blomberg P, Castillo S, Rousu J: **Principal metabolic flux mode analysis**. *Bioinformatics* 2018, **34**:2409-2417.

27. Szappanos B, Kovács K, Szamecz B, Honti F, Costanzo M, Baryshnikova A, Gelius-Dietrich G, Lercher MJ, Jelasity M, Myers CL, *et al.*: **An integrated approach to characterize genetic interaction networks in yeast metabolism**. *Nat Genet* 2011, **43**:656-662.

28. Larroude M, Rossignol T, Nicaud JM, Ledesma-Amaro R: **Synthetic biology tools for engineering *Yarrowia lipolytica***. *Biotechnol Adv* 2018, **36**:2150-2164.

29. McCarty NS, Shaw WM, Ellis T, Ledesma-Amaro R: **Rapid assembly of gRNA arrays via modular cloning in yeast**. *ACS Synth Biol* 2019, **8**:906-910.

30. Jiménez A, Muñoz-Fernández G, Ledesma-Amaro R, Buey RM, Revuelta JL: **One-vector CRISPR/Cas9 genome engineering of the industrial fungus *Ashbya gossypii***. *Microb Biotechnol* 2019, **12**:1293-1301.

31. Black JB, Perez-Pinera P, Gersbach CA: **Mammalian synthetic biology: engineering biological systems**. *Annu Rev Biomed Eng* 2017, **19**:249-277.

32. Abdullah, Jiang Z, Hong X, Zhang S, Yao R, Xiao Y: **CRISPR base editing and prime editing: DSB and template-free editing systems for bacteria and plants**. *Synth Syst Biotechnol* 2020, **5**:277-292.

33. Wirth NT, Kozaeva E, Nikel PI: **Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection**. *Microb Biotechnol* 2020, **13**:233-249.

34. Aparicio T, de Lorenzo V, Martínez-García E: **CRISPR/Cas9-enhanced ssDNA recombineering for *Pseudomonas putida***. *Microb Biotechnol* 2019, **12**:1076-1089.

35. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA: **Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system**. *Nucleic Acids Res* 2013, **41**:7429-7437.

36. Ho HI, Fang JR, Cheung J, Wang HH: **Programmable CRISPR-Cas transcriptional activation in bacteria**. *Mol Syst Biol* 2020, **16**:e9427.

37. Santos-Moreno J, Schaerli Y: **CRISPR-based gene expression control for synthetic gene circuits**. *Biochem Soc Trans* 2020, **48**:1979-1993.

38. Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, Olsen KM, Gregg A, Noggle S, Tessier-Lavigne M: **Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9**. *Nature* 2016, **533**:125-129.

39. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression**. *Cell* 2013, **152**:1173-1183.

40. Donati S, Kuntz M, Pahl V, Farke N, Beuter D, Glatter T, Gomes-Filho JV, Randau L, Wang CY, Link H: **Multi-omics analysis of CRISPRi-knockdowns identifies mechanisms that buffer decreases of enzymes in *E. coli* metabolism**. *Cell Syst* 2021, **12**:56-67 e56.

41. Shabestary K, Hernández HP, Miao R, Ljungqvist E, Hallman O, Sporre E, Branco Dos Santos F, Hudson EP: **Cycling between growth and production phases increases cyanobacteria bioproduction of lactate**. *Metab Eng* 2021, **68**:131-141.

42. Ye Z, Li S, Hennigan JN, Lebeau J, Moreb EA, Wolf J, Lynch MD: **Two-stage dynamic deregulation of metabolism improves process robustness & scalability in engineered *E. coli***. *Metab Eng* 2021, **68**:106-118.

43. Gao C, Guo L, Hu G, Liu J, Chen X, Xia X, Liu L: **Engineering a CRISPRi circuit for autonomous control of metabolic flux in *Escherichia coli***. *ACS Synth Biol* 2021, **10**:2661-2671.

44. Wang J, Zhang X, Cheng L, Luo Y: **An overview and metanalysis of machine and deep learning-based CRISPR gRNA design tools**. *RNA Biol* 2020, **17**:13-22.

45. Palazzotto E, Tong Y, Lee SY, Weber T: **Synthetic biology and metabolic engineering of actinomycetes for natural product discovery**. *Biotechnol Adv* 2019, **37**:107366.

46. Gallo G, Renzone G, Alduina R, Stegmann E, Weber T, Lantz AE, Thykaer J, Sangiorgi F, Scaloni A, Puglia AM: **Differential proteomic analysis reveals novel links between primary metabolism and antibiotic production in *Amycolatopsis balhimycina***. *Proteomics* 2010, **10**:1336-1358.

47. Kozaeva E, Volkova S, Matos MRA, Mezzina MP, Wulff T, Volke DC, Nielsen LK, Nikel PI: **Model-guided dynamic control of essential metabolic nodes boosts acetyl-coenzyme A — dependent bioproduction in rewired *Pseudomonas putida***. *Metab Eng* 2021, **67**:373-386.

48. Ding Q, Ma D, Liu GQ, Li Y, Guo L, Gao C, Hu G, Ye C, Liu J, Liu L, *et al.*: **Light-powered *Escherichia coli* cell division for chemical production**. *Nat Commun* 2020, **11**:2262.

49. Wang X, Han JN, Zhang X, Ma YY, Lin Y, Wang H, Li DJ, Zheng TR, Wu FQ, Ye JW, *et al.*: **Reversible thermal regulation for bifunctional dynamic control of gene expression in *Escherichia coli***. *Nat Commun* 2021, **12**:1411.

50. Cui S, Lv X, Wu Y, Li J, Du G, Ledesma-Amaro R, Liu L: **Engineering a bifunctional Phr60-Rap60-Spo0A quorum-sensing molecular switch for dynamic fine-tuning of menaquinone-7 synthesis in *Bacillus subtilis***. *ACS Synth Biol* 2019, **8**:1826-1837.

51. Li C, Gao X, Peng X, Li J, Bai W, Zhong J, He M, Xu K, Wang Y, Li C: **Intelligent microbial cell factory with genetic pH shooting (GPS) for cell self-responsive base/acid regulation**. *Microb Cell Fact* 2020, **19**:202.

52. Stirling F, Naydich A, Bramante J, Barocio R, Certo M, Wellington H, Redfield E, O'Keefe S, Gao S, Cusolito A, *et al.*: **Synthetic cassettes for pH-mediated sensing, counting, and containment**. *Cell Rep* 2020, **30**:3139-3148 e3134.

53. Wu Y, Chen T, Liu Y, Tian R, Lv X, Li J, Du G, Chen J, Ledesma-Amaro R, Liu L: **Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in *Bacillus subtilis***. *Nucleic Acids Res* 2020, **48**:996-1009.

54. Xu P, Li L, Zhang F, Stephanopoulos G, Koffas M: **Improving fatty acids production by engineering dynamic pathway regulation and metabolic control**. *Proc Natl Acad Sci USA* 2014, **111**:11299-11304.

55. Lv Y, Gu Y, Xu J, Zhou J, Xu P: **Coupling metabolic addiction with negative autoregulation to improve strain stability and pathway yield**. *Metab Eng* 2020, **61**:79-88.

56. Yang Y, Lin Y, Wang J, Wu Y, Zhang R, Cheng M, Shen X, Wang J, Chen Z, Li C, *et al.*: **Sensor-regulator and RNAi based bifunctional dynamic control network for engineered microbial synthesis**. *Nat Commun* 2018, **9**:3043.

57. Xiao Y, Bowen CH, Liu D, Zhang F: **Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis**. *Nat Chem Biol* 2016, **12**:339-344.

58. Nielsen AA, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D, Voigt CA: **Genetic circuit design automation**. *Science* 2016, **352**:aac7341.

59. •• Chen Y, Zhang S, Young EM, Jones TS, Densmore D, Voigt CA: **Genetic circuit design automation for yeast**. *Nat Microbiol* 2020, **5**:1349-1360.

Genetic circuit design automation for yeast has been achieved using the software Cello 2.0, which is an update of the previous version developed for *E. coli* before.

60. Moon TS, Lou C, Tamsir A, Stanton BC, Voigt CA: **Genetic programs constructed from layered logic gates in single cells**. *Nature* 2012, **491**:249-253.

61. Xu X, Li X, Liu Y, Zhu Y, Li J, Du G, Chen J, Ledesma-Amaro R, Liu L: **Pyruvate-responsive genetic circuits for dynamic control of central metabolism**. *Nat Chem Biol* 2020, **16**:1261-1268.

62. • Santos-Moreno J, Tasiudi E, Stelling J, Schaerli Y: **Multistable and dynamic CRISPRi-based synthetic circuits**. *Nat Commun* 2020, **11**:2746.

CRISPRi was employed for the remaking of genetic devices including toggle switch, oscillator, and incoherent feed-forward loop, which demonstrates its future applications for processing and computing within complex genetic circuits.

63. Gao C, Hou J, Xu P, Guo L, Chen X, Hu G, Ye C, Edwards H, Chen J, Chen W, *et al.*: **Programmable biomolecular switches for rewiring flux in *Escherichia coli***. *Nat Commun* 2019, **10**:3751.

64. Xu P: **Production of chemicals using dynamic control of metabolic fluxes**. *Curr Opin Biotechnol* 2018, **53**:12-19.

65. Guo L, Diao W, Gao C, Hu G, Ding Q, Ye C, Chen X, Liu J, Liu L: **Engineering *Escherichia coli* lifespan for enhancing chemical production**. *Nat Catal* 2020, **3**:307-318.

66. Lv Y, Qian S, Du G, Chen J, Zhou J, Xu P: **Coupling feedback genetic circuits with growth phenotype for dynamic population control and intelligent bioproduction**. *Metab Eng* 2019, **54**:109-116.

67. Rugbjerg P, Sarup-Lytzen K, Nagy M, Sommer MOA: **Synthetic addiction extends the productive life time of engineered *Escherichia coli* populations**. *Proc Natl Acad Sci USA* 2018, **115**:2347-2352.

68. Goldberg AP, Szigeti B, Chew YH, Sekar JA, Roth YD, Karr JR: **Emerging whole-cell modeling principles and methods**. *Curr Opin Biotechnol* 2018, **51**:97-102.

69. • Huang HH, Bellato M, Qian Y, Cárdenas P, Pasotti L, Magni P, Del Vecchio D: **dCas9 regulator to neutralize competition in CRISPRi circuits**. *Nat Commun* 2021, **12**:1692.
Coupling negative feedback into the dCas9 generator module releasing the competition between multiple sgRNAs, which enabling concurrent and independent regulation of multiple targets by CRISPRi-based genetic circuits.

70. Ceroni F, Boo A, Furini S, Gorochowski TE, Borkowski O, Ladak YN, Awan AR, Gilbert C, Stan GB, Ellis T: **Burden-driven feedback control of gene expression**. *Nat Methods* 2018, **15**:387-393.