

# 시스템생물학 및 합성생물학 기반 미생물 세포 공장 최적화 전략

신종오<sup>1</sup>, 노명현<sup>2</sup>, 백승호<sup>2</sup>, 신종혁<sup>3,4</sup>, 안정호<sup>5,6</sup>, 임성순<sup>7</sup>, 장성호<sup>8,9,10</sup>, 임현규<sup>11,12\*</sup>

## Systems and Synthetic Biology Approaches for Optimizing Microbial Cell Factories

Jongoh Shin<sup>1</sup>, Myung Hyun Noh<sup>2</sup>, Seung-Ho Baek<sup>2</sup>, Jonghyeok Shin<sup>3,4</sup>, Jung Ho Ahn<sup>5,6</sup>, Sung Sun Yim<sup>7</sup>,  
Sungho Jang<sup>8,9,10</sup>, and Hyun Gyu Lim<sup>11,12\*</sup>

Received: 19 August 2025 / Revised: 1 September 2025 / Accepted: 1 September 2025

© 2025 The Korean Society for Biotechnology and Bioengineering

**Abstract:** Microbial cell factories have emerged as pivotal platforms for the sustainable production of biofuels, pharma-

ceuticals, and high-value chemicals. Despite remarkable progress, the rational design and optimization of microbial hosts remain a major challenge due to complex cellular physiology, unknown gene functions, and the limitations of traditional trial-and-error approaches. Integration of systems and synthetic biology approaches can effectively address these barriers. Systems biology enables global understanding and predictive cellular modeling through omics analyses, while synthetic biology provides standardized tools for construction and dynamic regulation of microbial cells. The integration of both fields under the Design-Build-Test-Learn (DBTL) framework has transformed microbial strain development into a data-driven, iterative engineering process. This review comprehensively examines key systems biology tools such as genomics, transcriptomics, proteomics, and metabolomics, as well as synthetic biology approaches including bioparts assembly, genetic circuits, and high-throughput screening. Furthermore, the emerging roles of biofoundries and AI-powered automation platforms are introduced to accelerate DBTL cycles. By unifying systems-level insights with engineering precision, these integrated approaches pave the way for the next generation of high-performance microbial cell factories.

**Keywords:** systems biology, synthetic biology, metabolic engineering

## 1. INTRODUCTION

미생물 세포 공장은 박테리아 및 효모 등 유전적으로 조작된 미생물을 이용해 바이오매스로부터 바이오 연료, 고분자, 의약품 원료, 식품 첨가물 등 다양한 고부가가치 화합물을 생

<sup>1</sup>전남대학교 자연과학대학 생물학과

<sup>1</sup>Department of Biological Sciences, Chonnam National University, Gwangju 61186, Korea

<sup>2</sup>한국화학연구원 바이오화학연구센터

<sup>2</sup>Center for Bio-Based Chemistry, Korea Research Institute of Chemical Technology (KRICT), Ulsan 44429, Korea

<sup>3</sup>한국생명공학연구원 합성생물학연구센터

<sup>3</sup>Synthetic Biology Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea

<sup>4</sup>세종대학교 생명과학대학 식품생명공학과

<sup>4</sup>Department of Food Science and Biotechnology, College of Life Sciences, Sejong University, Seoul 05006, Korea

<sup>5</sup>한국과학기술연구원 청정에너지연구센터

<sup>5</sup>Clean Energy Research Center, Korea Institute of Science and Technology, Seoul 02792, Korea

<sup>6</sup>과학기술연합대학원대학교 한국과학기술연구원 에너지·환경 융합전공

<sup>6</sup>Division of Energy & Environment Technology, KIST School, University of Science and Technology, Seoul 02792, Korea

<sup>7</sup>한국과학기술원 공학생명과학대학원

<sup>7</sup>Graduate School of Engineering Biology, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea

<sup>8</sup>인천대학교 생명·나노바이오공학과

<sup>8</sup>Department of Bioengineering and Nano-Bioengineering, Incheon National University, Incheon 22012, Korea

<sup>9</sup>인천대학교 생명공학부 생명공학전공

<sup>9</sup>Division of Bioengineering, College of Life Sciences, Incheon National University, Incheon 22012, Korea

<sup>10</sup>인천대학교 바이오소재·공정개발 연구소

<sup>10</sup>Research Center for Bio Materials and Process Development, Incheon National University, Incheon 22012, Korea

<sup>11</sup>인하대학교 바이오시스템융합학과

<sup>11</sup>Department of Biological Sciences and Bioengineering, Inha University, Incheon 22212, Korea

<sup>12</sup>인하대학교 인공지능바이오산업기술연구소

<sup>12</sup>Institute of AI-driven Industrial Biotechnology, Inha University, Incheon 22212, Korea

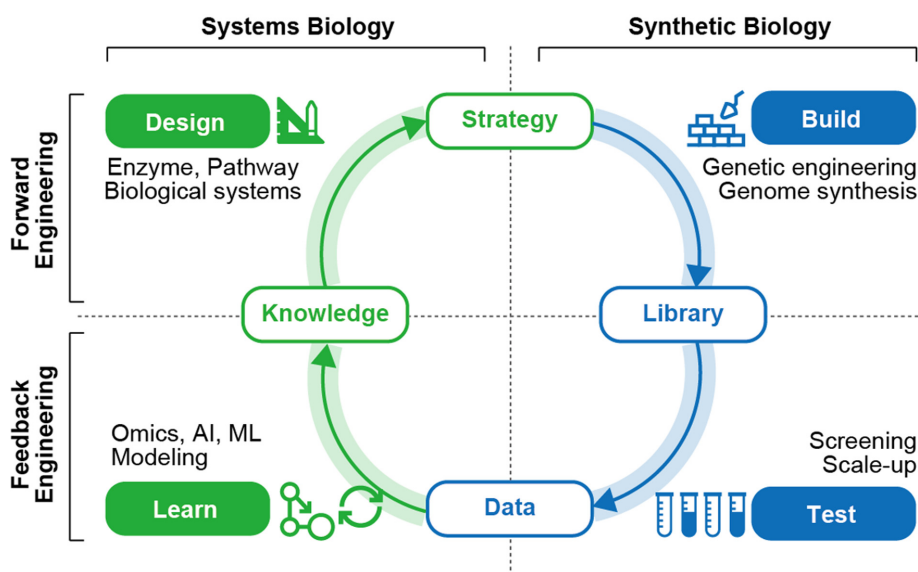
Tel: +82-32-860-7538, Fax: +82-32-860-7726

E-mail: hyungyu.lim@inha.ac.kr

산하는 바이오 공정을 의미한다 [1-3]. 이 기술은 친환경적이고 지속 가능한 생산 방식으로 기존 화학 공정을 대체하거나 보완할 수 있는 미래 산업의 핵심 기술로 주목받고 있으며, 지난 수십 년간 괄목할 만한 성과를 거두어 왔다. 대표적인 사례로, 의약품 분야에서는 Sanofi (Paris, France)가 효모를 통해 말라리아 치료제 아르테미시닌의 전구체인 아르테미신산(artemisinic acid)을 대량 생산하는 공정을 상용화하였으며 [4], 바이오 연료 분야에서는 Amyris (Emeryville, CA, USA)가 효모 유래 파르네센(farnesene)을 기반으로 바이오 제트 연료를 항공기에 실제로 적용하는 데 성공한 바 있다 [5]. 화학 소재 분야에서는 Genomatica (San Diego, CA, USA)가 대장균을 기반으로 1,4-부탄디올(1,4-butanediol, 1,4-BDO) 생산용 세포 공장을 개발하여 연간 3만 톤 규모의 바이오 1,4-BDO 생산 공정을 구축하였고 [6], DSM-firmenich (Maastricht, Netherlands)는 효모를 이용한 비타민 B2 (riboflavin) 생산을 통해 전 세계 시장의 상당 부분을 점유하고 있다. 최근에는 Modern Meadow (Nutley, United States)가 바이오 가죽 생산을 위한 콜라겐 합성에도 성공하여 지속 가능한 소재 개발의 새로운 가능성을 제시하기도 하였다 [7]. 한국에서도 CJ 제일제당(Seoul, Korea) 및 GS 칼텍스(Seoul, Korea)가 각각 폴리하이드록시알카노에이트(polyhydroxyalkanoates) 및 2,3-부탄디올(2,3-butanediol) 상업화에 성공하여, 다양한 상품 제조에 적용하고 있다. 이러한 접근법은 기존 화학 공정 대비 낮은 온도와 압력에서 운용되며 유독성 용매나 촉매를 사용하지 않아 환경적 이점이 크고, 지속 가능한 생산 방식으로 평가받고 있다 [8].

그러나 자연계에서 분리된 야생형 미생물은 특정 화합물을 산업적 규모로 대량 생산하는 데 최적화되어 있지 않으며 [1,9], 목적 산물 생산을 위한 미생물 세포 공장 설계 및 구축 과정에도 다양한 난제가 존재한다 [8]. 생리 측면에서는 첫째로 세포는 생존과 증식에 필수적인 대사 활동에 자원을 우선 분배하므로, 목적 산물의 생산성이 낮다. 둘째로 기질이나 중간 산물, 목적 산물의 축적 또는 공정 조건에서 유래한 세포 스트레스 및 낮은 내성으로 인한 생산성 저하가 발생할 수 있으며, 이를 해결하기 위해 미생물 균주의 내성을 향상시키는 전략이 필요하다 [10]. 대사 경로 측면에서는 첫째로 목적 산물의 생합성 경로가 명확하지 않은 경우, 핵심 경로를 탐색하고 주요 유전자 및 효소를 선별하여 미생물에 도입함으로써 합성 경로 구축이 가능하다 [11,12]. 둘째로, 도입된 이중 생합성 경로 내에서 생산 수율을 제한하는 병목 대사 경로가 발생할 가능성이 높기 때문에, 이러한 병목 구간을 규명하고 대사 흐름을 원활하게 조절해야 목표 생산량을 높일 수 있다. 따라서 산업적 요구 수준을 충족하는 미생물 세포 공장 개발에는 세포 내 대사 네트워크를 인위적으로 재설계하고 최적화하는 대사공학적 접근이 필수적이다 [13]. 무작위 돌연변이를 유발하여 원하는 형질을 보이는 균주를 선별하는 과거의 균주 개발 방식은 정보의 한계로 시행착오에 의존하여 유전자 조작이 이루어졌다. 이러한 접근법은 시간과 비용이 많이 소요되고 성공률이 낮아, 복잡한 생명 시스템을 공학적으로 제어하는 데 명확한 한계를 보였다 [14,15].

이를 극복하기 위한 새로운 패러다임으로 시스템생물학



**Fig. 1.** The DBTL (Design-Build-Test-Learn) cycle integrating systems biology and synthetic biology. The green semicircle represents the systems biology domain, where Learn and Design stages utilize omics data, AI/ML, and modeling to extract knowledge (feedback engineering) and to design enzymes, metabolic pathways, and cellular systems (forward engineering). The blue semicircle represents the synthetic biology domain, where Build and Test stages involve gene editing and genome synthesis (forward engineering), followed by screening and scale-up of constructed variant libraries to generate performance data (feedback engineering). With each cycle, the design rules are refined and fed back into the subsequent Design-Build stages. AI, artificial intelligence; ML, machine learning.

(systems biology)과 합성생물학(synthetic biology)의 융합이 제시되고 있다 (Fig. 1) [16]. 시스템생물학은 유전체, 전사체, 단백질체 등 다양한 수준의 오믹스(omics) 데이터를 통합적으로 분석하여 세포 내 복잡한 상호작용 네트워크를 시스템 수준에서 이해하고 예측하는 ‘분석적 학문’이며 [17], 다중 오믹스 데이터와 네트워크 모델링을 통해 세포를 설계 (design)하고 실험 결과로부터 학습(learn)하는 기반을 제공할 수 있다. 반면, 합성생물학은 표준화된 유전자 부품 (bioparts)을 활용하여 새로운 기능의 유전자 회로와 대사 경로를 합리적으로 설계하고 구축하는 ‘구성적 학문’이다. 표준화된 바이오 부품 조립과 유전자 회로 설계를 통해 세포 공장을 구축(build)하고, 그 기능을 검증(test)하는 기술을 제공하며, 이를 반복적으로 수행하여 미생물 개량 효율을 극대화 한다 [18,19].

앞서 언급한 복합적인 난제들을 체계적으로 해결하기 위한 핵심 전략은 설계-구축-시험-학습(Design-Build-Test-Learn, DBTL) 사이클이다. DBTL 사이클은 미생물 세포 공장 개발을 직관과 경험에 의존하는 방식에서 예측 가능한 공학적 최적화 과정으로 전환시키는 핵심 프레임워크로 작용한다 [20,21]. 이 사이클은 미생물 세포 공장 개발 과정을 체계적이고 반복적인 최적화 과정으로 정의하며, 시스템생물학과

합성생물학 기술들을 유기적으로 연결하는 기능적 다리 역할을 수행한다 [16]. 구체적으로, DBTL 사이클은 시스템 생물학적 분석을 통해 얻은 지식을 바탕으로 합리적 설계를 수행(design)하고, 합성생물학 기술로 정밀하게 균주를 구축(build)하며, 고처리량 기술로 표현형을 신속하게 검증(test)한 후, 그 결과를 다시 시스템 모델에 반영하여 이해를 심화하고 다음 설계에 반영(learn)함으로써 선순환 구조를 완성한다 [22].

DBTL 사이클은 미생물 세포 공장 개발을 표준화, 정량화, 자동화할 수 있는 공학적 접근법으로 자리잡았고, 전 세계 각국은 바이오파운드리(biofoundry)라는 자동화된 설비를 앞다투어 구축하고 있다. 글로벌 바이오파운드리 연맹(Global Biofoundries Alliance)에 가입된 기관의 경우 국제적 표준화와 상호운용성을 확보하고 있으며, 한국 역시 국가 차원의 바이오파운드리 인프라 구축을 본격화하고 있다. 특히 정부는 세계 최초로 합성생물학 진흥법을 제정하고 대통령 직속 국가바이오위원회를 출범시켜, 합성생물학 연구와 산업화를 체계적으로 지원할 수 있는 제도적 기반을 마련하였다. 이러한 노력은 한국이 글로벌 바이오파운드리 네트워크와의 연계를 강화하는 동시에, 차세대 미생물 세포 공장 개발을 위한 연구 허브로 자리매김하는 데 중요한 역할을 할

**Table 1.** Comparison of key omics technologies for microbial cell factory optimization

Technology	Description	Output	Application	Limitation
RNA-Seq	Quantitative analysis of the entire cellular RNA using next-generation sequencing	Transcriptome (gene expression profile)	Identification of transcriptional bottlenecks and regulatory responses; discovery of candidate genes related to metabolic pathways	mRNA abundance does not always correlate with protein/enzyme activity; lacks post-transcriptional regulatory information
Proteomics	Identification and quantification of the entire cellular proteome using mass spectrometry	Proteome (protein expression profile)	Direct quantification of enzymes; identification of bottlenecks in protein secretion/stability; direct link to phenotype	Technically complex and costly; narrower dynamic range compared to transcriptome
Metabolomics	Quantitative analysis of small-molecule metabolites using MS or NMR	Metabolome (metabolite profile)	Measurement of final metabolic flux outputs; direct evidence of metabolic bottlenecks; closest link to phenotype	Analytical challenges due to chemical diversity of metabolites; instability of certain metabolites
ChIP-Seq and ChIP-Exo	Sequencing of DNA fragments bound by specific proteins (mainly TFs) after immunoprecipitation	Genome-wide binding map of specific transcription factors	Identification of master regulators and their target genes; establishment of global regulatory strategies	Requires antibodies for specific TFs; indirect regulatory effects cannot be captured
CRISPRi-Seq	Genome-wide repression of gene expression using dCas9/gRNA libraries followed by fitness analysis	Fitness score of each gene under given conditions	Functional analysis of all genes including essential ones; discovery of unexpected engineering targets	Potential off-target effects; complex library construction and screening
Ribo-Seq	Sequencing of ribosome-protected mRNA fragments to measure translational activity	Ribosome occupancy map; TE	Identification of translational bottlenecks; discovery of hidden regulatory mechanisms (sORFs, ribosome pausing)	Technically demanding with complex data analysis; requires careful removal of rRNA contamination

RNA, ribonucleic acid; Seq, sequencing; mRNA, messenger RNA; MS, mass spectrometry; NMR, nuclear magnetic resonance; CRISPRi, clustered regularly interspaced short palindromic repeats interference; gRNA, guide RNA; TF, transcription factor; TE, translation efficiency; sORFs, small open reading frames; rRNA, ribosomal RNA.

것으로 기대된다.

본 논문에서는 미생물 세포 공장 구축 및 최적화 과정에서 시스템생물학 및 합성생물학 기술들이 DBTL 사이클이라는 프레임워크 안에서 어떻게 기여하는지를 종합적으로 고찰하고자 한다. 먼저, 시스템생물학 기반의 ‘설계’ 및 ‘학습’ 단계 기술들을 살펴보고, 이어서 합성생물학 기반의 ‘구축’ 및 ‘시험’ 단계 기술들을 논의하고자 한다. 마지막으로, 이러한 기술들이 통합되고 자동화된 바이오파운드리 플랫폼을 통해 미래의 미생물 세포 공장 개발이 어떻게 가속화될 것인지 전망하고자 한다.

## 2. 시스템생물학: 합리적 설계를 위한 설계 및 학습 단계

시스템생물학은 생명체를 구성하는 다양한 요소들의 상호작용을 분석하고, 이를 통해 생명현상의 원리를 이해하려고 하는 학문으로 [23], 효율적인 미생물 세포 공장 개발을 위한 기초적 이해를 제공하는 필수 분야이다 (Table 1). 특히 다양한 오믹스 데이터의 생산 및 분석 기술이 발전함에 따라 구성 요소의 개별적 정보 축적뿐만 아니라, 두 개 이상의 구성 요소의 상호작용에 관한 정보도 얻어지고 있다. 이들 데이터를 종합적으로 분석하여 시스템에 대한 이해 및 예측 능력을 높임으로써, 궁극적으로 효율적인 세포 공장을 개발하는 데 기여할 수 있다 [24].

미생물 세포 공장 개발에 있어 DBTL 사이클의 ‘설계’와 ‘학습’ 단계는 시스템생물학 기술을 통해 세포를 하나의 통합된 시스템으로 이해하고, 이를 바탕으로 가장 효과적인 유전공학적인 개입 지점을 예측하는 데 중점을 둔다. 특히, 계층 단위의 대사 모델링부터 전사체, 단백질체, 그리고 여러 오믹스 데이터의 통합 분석에 이르기까지 다계층적 생물학적 데이터를 수집하고 해석함으로써, 경험적 추측을 넘어 데이터 기반의 합리적인 균주 설계를 수행할 수 있다.

### 2.1. 유전체 수준 대사 모델(genome-scale metabolic model, GSMM)

GSMM은 미생물 세포의 전체 대사 네트워크를 수리적으로 모델링한 것으로, 세포가 대사 회로를 통해 탄소원으로부터 바이오매스 및 화합물을 생산하는 과정을 사전 시뮬레이션할 수 있게 한다. 일반적으로 GSMM 구축은 유전자-단백질-반응(gene-protein-reaction) 관계를 규명하여 생화학 반응의 양론 관계를 나타내는 행렬을 구성하는 과정으로 이루어진다. 이러한 GSMM은 유전체 정보에서 얻을 수 있는 모든 대사 반응과 경로를 포함하여, 주어진 조건에서 각 대사 반응의 흐름(flux)을 계산하고 시뮬레이션할 수 있도록 한다 [25].

예시로 플럭스 균형 분석(flux balance analysis, FBA) 방법이 있으며, 구체적으로 이는 질량 보존 법칙과 같은 물리화학적 제약 조건 하에서, 세포 성장과 같은 특정 목표 함수(objective function)를 최적화하는 대사 흐름 분포를 선형 계획법을 통해 예측하는 강력한 방법이다 [26]. 이를 통해, 특

정 물질 생산을 최대화하고자 할 때, 이론적으로 탄소 흐름이 어떤 경로로 어떻게 분배되어야 하는지 예측할 수 있고, 이를 바탕으로 유전자 조작 전략을 수립할 수 있다 [27].

대사공학 분야에서 GSMM의 활용은 혁신적인 변화를 가져왔는데, 가장 중요한 응용 분야는 목표 화합물 생산을 증대시키기 위한 최적의 유전자 결손(knockout), 과발현(over-expression), 또는 발현 저하(knockdown) 대상을 컴퓨터 시뮬레이션을 통해 예측하는 것이다 [26,28]. 예를 들어, FBA 기반의 OptKnock 알고리즘은 목표 화합물의 생산이 세포 성장과 연계되도록 여러 유전자 삭제 조합을 찾아줌으로써, 생산 균주의 대사 경로를 합리적으로 재설계하도록 도움을 주고 있다 [29].

현재 대장균 및 효모 등 산업미생물에 대한 상세한 GSMM이 구축되어 폭넓게 활용되고 있으며, 실제로 다양한 화합물 생산 균주가 개발되어 왔다. 예를 들어, GSMM상에서 *Escherichia coli* 유전체의 유전자 결손 시뮬레이션을 수행함으로써, L-발린 생산성 향상에 기여하는 표적 유전자를 성공적으로 도출한 바 있으며, 이는 무작위 돌연변이 라이브러리 선별에 비해 월등히 효율적으로 적용될 수 있음이 보여진 바 있다 [30]. 또한, GSMM 기반의 대사 플럭스 최적화를 통해 *E. coli*에서 지방산 생산 균주의 수율을 최고 수준으로 높인 사례도 보고된 바 있다 [31]. 또한, GSMM은 경로 예측 알고리즘과 결합하여 세포 내에 존재하지 않는 비고유(non-native) 대사 경로를 포함한 모든 가능한 생합성 경로를 평가하고 최적의 경로를 설계하는 데에도 활용되고 있으며, 더 나아가, 기존에 알려지지 않았던 대사 반응이나 유전자의 기능을 발견하는 등, 생물학적 지식의 공백을 메우는 역할도 수행하고 있다 [26,28].

하지만 초기단계 GSMM은 대사 회로를 구성하는 유전자가 받는 세포 내 조절, 효소의 양 및 실제 반응 속도 등 같은 복잡한 세포 내 동역학을 고려하지 않는다는 한계가 존재한다. 최근 효소의 촉매 속도( $k_{cat}$ ), 조효소(cofactor), 단백질 합성 비용 등을 추가로 고려한 효소 제약 모델(enzyme-constrained model)과 같은 차세대 모델들이 개발되고 있으며, 이는 더욱 정확한 세포 표현형 예측을 가능하게 할 것으로 기대된다 [32,33]. 더불어, 최근에는 GSMM에 다중 오믹스 데이터를 통합하여 모델 예측 정확도를 높이고, 미생물 공동배양 시스템에서의 대사 상호작용을 예측하는 등 응용 범위도 확대되고 있다 [8]. 이러한 GSMM 기반 설계는 대규모 유전자 편집에 따른 세포 대사 재구성을 체계적으로 안내하여, 미생물 세포 공장의 합리적 디자인을 가능케 한다.

### 2.2. 전사체 분석(transcriptome analysis)

전사체 분석은 특정 조건에서 모든 RNA (ribonucleic acid) 전사체의 종류와 양을 정량적으로 측정하는 기술로, 세포 내 모든 유전자들의 전사 상태에 대한 순간적인 정보를 얻을 수 있어, 대규모 유전체 수준에서의 대사공학 연구에 핵심적인 도구로 활용된다 [34]. 초반에는 마이크로어레이(microarray)와 같은 방법이 사용되었으나, 차세대 염기서열 시퀀싱

(Next-generation sequencing) 기술이 등장한 이후 RNA 시퀀싱(RNA-sequencing)으로 대부분 대체되었다.

전사체 분석은 (i) 전사 조절 단백질의 타겟 유전자를 발굴하고, (ii) 스트레스 반응 및 내성 기작(mechanism)을 규명하며, (iii) 대사 경로상의 병목 구간을 식별하는 등 다양한 목적에 활용될 수 있다 [34]. 예를 들어 특정 전사 조절 인자의 표적 유전자를 찾기 위해 해당 단백질을 과발현하거나 유전자 결손 시킨 후 전사체 변화를 비교하면, 조절 인자가 직접 조절하는 유전자군을 도출할 수 있다. 또한, 대사 과부하나 용매 독성, 산화 스트레스 등의 조건에서 전사체 변화를 분석하면, 내성 관련 유전자나 보호 기작(예: efflux pump, chaperone 등)을 밝혀낼 수 있다. 특히 적응진화(adaptive evolution, ALE)로 획득된 내성 균주의 표현형을 해석할 때, 전사체 데이터를 통해 내성에 기여하는 유전자들의 발현 변화를 파악하는 전략이 유용하다. 이 외에도 고생산성 균주와 대조 균주의 전사체 비교를 통해 목표 대사 경로에서 발현이 저하된 효소 유전자(즉, 전사 수준의 병목)나 과도하게 발현된 경쟁 경로 및 스트레스 반응 유전자들을 찾아냄으로써 대사 흐름의 문제점을 진단할 수 있다. 이렇게 밝혀진 문제 유전자들을 억제하거나 추가 발현하여 생산성을 높이는 전략도 가능하다.

그러나 단순히 발현 변화 폭이 큰 유전자를 표적으로 삼는 전통적 접근은 한계가 있으며, 최근에는 전사체 데이터를 대사 네트워크 맥락에서 해석하는 통합 기법들이 도입되어 보다 신뢰성 높은 표적 탐색이 가능해지고 있다. 예를 들어, 차등 발현 유전자(differentially expressed genes)에 의존하지 않고 전사체 변화를 대사 네트워크 기반으로 해석하는 알고리즘이 개발되어, *Saccharomyces cerevisiae*에서 전사체 데이터로부터 신규 대사 경로 설계 전략을 도출한 사례가 보고되었다 [35]. 또한 전사체 기반 스트레인 최적화 도구도 제시되어, 2차 대사 산물 생산 스트렙토마이세스(*Streptomyces*) 균주에서 과발현 유전자를 체계적으로 설계하는 데 성공한 바 있다 [36,37]. 나아가 여러 조건에서 수집된 전사체 데이터를 기계학습 등의 AI (artificial intelligence) 알고리즘으로 통합 분석하여, 동일한 전사 조절을 받는 공동 조절 유전자군(co-regulated gene set)을 찾아내 전장 전사 조절 네트워크를 규명하려는 시도도 확장되고 있다. 실제로, *E. coli*의 방대한 전사체 데이터를 독립 성분 분석(independent component analysis)으로 분해(deconvolution)하여 얻어진 iModulon이라는 유전자군이 실험적으로 규명된 레굴론(regulon)과 상당 부분 겹침이 확인되었다 [38]. 이는 비모델 미생물에서도 응용할 수 있는 접근법으로 [39], 축적된 데이터로부터 유의미한 생물학적 통찰을 도출해내는 강력한 수단이라 할 수 있다. 다만 mRNA (messenger RNA)의 양이 단백질 양이나 효소 활성과 반드시 비례하지 않는다는 한계가 있다. 전사 후 조절, 번역 효율 차이, 단백질 안정성 등 여러 요인이 관여하므로 전사체만으로 세포 기능 상태를 완벽히 파악하기 어렵다. 따라서 보다 정확한 이해를 위해서는 다른 오믹스 데이터와의 통합 분석이 필수적이다 [40].

### 2.3. 단백질체학(proteomics)

단백질체학은 질량 분석기(mass spectrometry)를 기반으로 세포 내 단백질들의 발현량과 변형 상태 등을 전반적으로 분석하는 기법이다 [41]. 단백질은 세포 내에서 실제로 생화학 반응을 수행하고 세포 구조를 형성하는 주체이기 때문에, 단백질체 분석을 통해 전사체 분석보다 실제 효소들의 발현량이나 활성 상태를 더욱 직접적으로 확인할 수 있어 대사 경로의 병목을 보다 정확하게 찾아낼 수 있다. 특히, mRNA의 양과 실제 단백질 양이 완벽하지 않을 수 있다는 점을 고려하면, 단백질체 분석이 세포의 실제 상태를 파악하는 데 더 직접적인 방법이 될 수 있다 [42,43].

미생물 세포 공장 최적화에서 단백질체 분석은 다양한 방식으로 활용된다. 첫째, 목표 대사 경로에 포함된 효소들의 발현 수준을 직접 정량하여, 원하는 효소가 충분히 만들어지지 않거나 혹은 세포 자원이 불필요한 단백질 생산에 소모되고 있는 문제를 직접적으로 확인할 수 있다 [44,45]. 이러한 분석을 통해 전사 수준에서는 정상적으로 발현되지만 번역 수준에서 병목이 발생하는 숨겨진 속도 제한 단계를 발굴할 수 있다. 구체적인 사례로, 대장균에 효모 유래 메발론산 경로를 도입하여 세스퀴테르펜 계열 물질인 아모르파디엔(amorphadiene)을 생산한 연구에서, 표적 단백질 정량 분석(targeted proteomics)을 실시한 결과, 두 개의 효모 효소인 메발론산 키나아제(Mevalonate kinase)와 포스포메발론산 키나아제(Phosphomevalonate Kinase)의 발현이 특히 낮아 전체 경로의 병목으로 작용함을 밝혀낸 바 있다 [45]. 이후 연구진은 해당 효소의 유전자를 코돈 최적화하고 강한 프로모터로 재조합하여 단백질 발현을 높였고, 그 결과 아모르파디엔 생산 수율이 3배 이상 향상되어 500 mg/L 이상의 농도를 달성할 수 있었다.

둘째, 단백질 접힘(folding), 안정성(stability), 분비(secretion) 등 생산된 단백질이 겪는 다양한 과정에서의 병목 현상을 파악할 수 있다. 예를 들어, 단백질 분비 생산 속주로 널리 활용되고 있는 *Bacillus subtilis*의 경우, 단백질체 분석을 통해 분비 경로의 비효율적인 단계를 찾거나, 생산된 목적 단백질을 분해하는 단백질분해효소(protease)를 동정하여 제거하는 전략을 수립할 수 있다. 실제로 *B. subtilis* 활용  $\alpha$ -아밀레이스 생산 연구에서는 단백질체 분석을 통해 여섯 개의 주요 외부 비 프로테아제를 동정하고 이들을 순차적으로 결실시킴으로써 목적 단백질의 안정성을 크게 향상시킨 단백질 세포 공장을 구축한 바 있다 [46]. 또한 *B. subtilis*의 단백질 분비 시스템에 대한 체계적인 단백질체 분석을 통해 Sec-의존적 분비 경로에서 PrsA 지질단백질이 분비 단백질의 접힘에 중요한 역할을 함을 확인하였고, PrsA의 과발현을 통해 여러 이종 단백질의 분비량을 최대 네 배까지 향상시킬 수 있음이 보고되었다 [47].

셋째, 산업 발효 공정에서 존재할 수 있는 생성물 독성 및 삼투압 스트레스 등 다양한 환경에 대한 세포의 반응을 분석하여, 이 때 발현이 증가하는 스트레스 단백질들을 파악하고 이를 조절하여 균주의 강건성(robustness)을 향상시킬 수 있

다 [48,49]. 예를 들어, 고농도 바이오 연료나 유기산 생산 조건에서 열충격 단백질(heat-shock proteins), 샤페론(chaperones), 외막 단백질 등의 발현 변화를 추적함으로써 세포 스트레스의 주요 원인을 파악하고 이를 완화하기 위한 유전공학 전략을 수립할 수 있다.

더 나아가, 단백질체 분석 데이터는 유전체 규모 대사 모델의 정확도를 높이는 데도 활용되고 있다. 효소 제약 모델에 실제 단백질 발현량 데이터를 통합함으로써 단순한 플럭스 균형 분석보다 현실적인 대사 상태 예측이 가능해진다 [1,50]. 이러한 접근은 세포 내 단백질 자원 할당의 한계를 고려한 보다 정교한 대사공학 전략 수립을 가능케 한다. 이처럼 단백질체 수준의 분석은 세포 공장의 실제 작동 메커니즘을 규명하여, 보다 세밀한 최적화 전략 수립에 기여할 수 있으며, 다른 시스템생물학 기반 분석과 상호보완적으로 활용될 때 미생물 세포 공장의 합리적 설계와 성능 향상에 필수적인 정보를 제공한다.

## 2.4. 대사체학(metabolomics)

대사체학은 세포, 조직, 또는 생물체 내에 존재하는 저분자 대사 물질(metabolites) 전체를 정량적으로 분석하여, 세포의 생리적 상태를 파악할 수 있는 도구를 제공한다 [51]. 대사 물질은 유전자 발현과 효소 활동의 최종 산물이자 세포의 생리적 상태를 가장 직접적으로 반영하는 표현형 지표이기 때문에, 대사체 분석을 통해 미생물 세포 공장의 실제 대사 흐름과 생산성을 실시간으로 모니터링할 수 있다 [52].

미생물 세포 공장 최적화에서 대사체 분석은 여러 측면에서 핵심적인 역할을 수행한다. 첫째, 대사 플럭스 분석(metabolic flux analysis, MFA)을 통한 병목 지점 규명이다. 동위원소 표지 기법( $^{13}\text{C}$ -labeling)과 질량분석법을 결합한 MFA는 세포 내 대사 경로의 플럭스를 정량화하여 생산성을 제한하는 속도 결정 단계를 정확히 식별할 수 있다 [53]. 예를 들어, 대장균에서 1,4-BDO 생산 연구에서는 표적 대사체 분석을 통해 1,4-BDO 생합성 경로의 마지막 두 단계가 주요 병목으로 작용함을 확인하였고, 이후 해당 효소들의 유전공학 적 개량을 통해 생산성을 크게 향상시킬 수 있었다 [54,55].

둘째로 중간 대사 물질의 축적 및 독성 모니터링을 통한 균주 및 공정 최적화가 가능하다. 고농도 생산 조건에서는 중간체나 최종 산물의 축적으로 인한 세포 독성이 생산성 저하의 주요 원인이 될 수 있다. 대사체 분석을 통해 이러한 독성 대사 물질의 농도 변화를 실시간으로 추적함으로써, 적절한 배양 조건 조절이나 유전자 회로 설계를 통한 독성 완화 전략을 수립할 수 있다 [56,57]. 또한, 스트레스 대사 물질이나 에너지 상태 지표(adenylate energy charge 등)의 변화를 통해 세포의 상태를 평가하고 배양 환경을 최적화할 수 있다.

끝으로 비표적 대사체 분석(untargeted metabolomics)을 통한 예측하기 어려웠던 대사체 변화를 발견하고, 이를 균주 개량에 반영하는 것이다. 비표적 대사체 분석은 미리 정의된 표적 대사 물질 집합없이 광범위한 질량 스펙트럼 데이터를 수집하여 수천 개의 이온을 탐지할 수 있으며, 이를 통해 예

상치 못한 낮은 수준의 대사 물질 차이를 발견할 수 있는 기회를 제공한다. 예를 들어, 숙신산 생산 *E. coli* 균주의 대사체 프로파일 연구에서는 개량 균주에서 만니톨의 상대적 농도가 예상치 않게 증가함을 발견했으며, 만니톨 탈수소효소 제거를 통해 숙신산 생산을 약 20% 향상시킨 바 있다 [58].

최근에는 공간 대사체학(spatial metabolomics)과 단일 세포 대사체학(single-cell metabolomics) 등 새로운 기법들의 발전으로 세포 내 대사 물질의 공간적 분리와 세포 간 이질성에 대한 이해도 깊어지고 있다 [59,60]. 공간 대사체학은 조직이나 세포 내에서 대사 물질의 위치별 농도 분포를 시각화함으로써 대사 구획화(metabolic compartmentalization)와 국소적 대사 활동을 규명할 수 있게 하며, 단일 세포 대사체학은 세포 집단 내에서 개별 세포의 대사적 다양성과 상태 변화를 실시간으로 추적하여 생산성 향상을 위한 균질한 세포 집단 유지 전략 수립에 중요한 정보를 제공할 수 있다. 따라서, 향후 대사체학 기술은 다른 오믹스 기법들과 함께 미생물 세포 공장의 시스템 수준 이해와 정밀 제어를 실현하는 핵심 플랫폼으로 그 중요성이 더욱 커지고 있다.

## 2.5. 다중 오믹스(multi-omics)

비록 단일 오믹스 데이터는 각기 다른 층위에서 세포를 조망하지만 복잡한 생명 현상의 일부분만을 보여주기 때문에, 개별로는 세포 시스템의 전체상을 설명하는 데 한계가 있다. 유전체 → 전사체 → 단백질체 → 대사체로 이어지는 정보의 흐름을 온전히 추적하고 시스템을 총체적으로 이해하기 위해, 다양한 오믹스 데이터를 동시에 통합 분석하려는 노력이 진행되어 왔다 [61,62]. 다중 오믹스 접근법은 데이터 규모가 방대하여 머신러닝(machine learning, ML) 등의 기법과 결합되고 있으며, 이렇게 얻어진 시스템 수준 지식은 미생물 세포 공장의 설계 및 개선에 강력한 가이드가 된다 [62]. 실제로 다중 오믹스와 AI를 활용해 대사 경로 동태를 예측하고 신규 설계에 반영하는 사례가 늘고 있으며, 이는 유전형-표현형 연결고리를 구명하고 대사 병목의 근본 원인을 밝혀내는 데 큰 도움을 준다.

### 2.5.1 전사 인자결합 분석(ChIP-Seq, ChIP-exo)

ChIP-Seq 혹은 ChIP-exo(크로마틴 면역침강 후 시퀀싱)은 특정 DNA 결합단백질(전사 인자, 억제자 등)이 게놈 어디에 결합하는지를 밝혀주는 방법이다 [63,64]. 이를 통해 세포 내 전사 조절망을 지도화할 수 있으며, 어떤 전사 인자가 목표 대사 경로의 유전자를 억제하거나 활성화하는지 파악할 수 있다. 예를 들어 *E. coli*에는 기능이 완전히 밝혀지지 않은 전사 인자가 많은데, ChIP-exo 기법(ChIP-seq의 해상도 개선)을 활용한 연구에서 40개의 미지 전사 인자의 전장 결합 위치를 확인하였다 [63]. 이렇게 얻은 DNA 결합 정보를 전사체 변화와 연계 분석한 결과, 이들 전사 인자가 조절하는 표적 유전자들이 대사 경로와 막 단백질 등에 관여함을 밝혀내어 미지 인자의 기능을 추정할 수 있었다.

ChIP-Seq 및 ChIP-exo로 구축한 글로벌 전사 조절 네트워크



크 정보는 복잡한 세포 내 유전자 발현 패턴을 사람이 이해할 수 있는 기초를 제공하기에, 세포 공장 디자인에 큰 도움을 줄 수 있다. 개별 유전자를 하나씩 조작하는 대신, 전체 경로를 통제하는 마스터 조절자(master regulator) 전사 인자 하나를 조절함으로써 경로상의 모든 효소 발현을 동시 제어하는 전역적 공학 전략을 구사할 수 있다. 실제로 RegulonDB나 YEASTRACT+ 같은 데이터베이스는 여러 실험을 통해 축적된 TF-타겟 유전자 정보를 제공하여, 특정 대사물 생산 증대에 적합한 TF를 예측하는 데 활용되고 있다 [65-67]. 최근 이를 인공지능과 결합하여 더 정교하게 생산 균주에서 불필요하게 작동하는 억제 인자를 제거하거나, 유용한 활성 인자를 강화하는 등 전사과정 제어를 통한 대사공학이 가능해진다 [68]. 이처럼 여러 데이터를 통합적으로 분석하여 기작이 수반된 지식으로 전환, 이를 바탕으로 신뢰성 있게 유전자 조작을 수행할 수 있게 된다.

### 2.5.2. 고처리량 기능 유전체학(high-throughput functional genomics)

트랜스포존 시퀀싱(transposon-sequencing, Tn-Seq)과 CRISPRi-Seq (clustered regularly interspaced short palindromic repeats inhibition-sequencing) 같은 게놈 레벨 기능 스크리닝(genome-level functional screening) 기법들은 시스템 수준에서 유전자-표현형 관계를 탐색할 수 있게 해준다. 한두 개의 타겟 유전자에 국한된 전통적 대사공학을 넘어 유전체 내 모든 유전자를 특정 표현형(주로 성장 적합성)과 연관 지음으로써 균주 개량 타겟 발굴에 있어 핵심적으로 요구되는 유전체학 기술이다 [69,70].

Tn-Seq은 트랜스포존을 유전체에 무작위 삽입하여 방대한 유전자 파괴 돌연변이 라이브러리를 만들고, 특정 선택압 조건에서 배양한 뒤 생존한 균주들의 트랜스포존 삽입 위치를 시퀀싱하는 기법이다 [69,71]. 어떤 유전자에 트랜스포존이 삽입되면 그 유전자의 기능이 상실되므로, 특정 환경에서 생존에 필수적인 유전자에 트랜스포존이 삽입된 돌연변이체는 군집에서 사라지게 되고, 시퀀싱 신호가 줄어드는 방식으로 탐지된다. 이를 통해 조건 의존적 필수 유전자나, 반대로 제거 시 생장이 유리하거나 불리해지는 유전자 등을 전장 수준에서 선별할 수 있다. 예를 들어 Tn-Seq은 병원성 세균의 숙주 내 생존 필수 유전자를 밝히거나, 리그노셀룰로오스 분해에 필요한 유전자들을 동정하는 데 활용되어 왔다 [72,73]. Tn-Seq 기법의 발전으로 최근에는 미세유체액적 기반 Droplet Tn-Seq도 등장하여, 단일 세포 수준의 복잡한 표현형도 감지할 수 있게 되었다 [74].

CRISPRi-Seq은 불활성화 Cas9(deactivated Cas9, dCas9) 단백질과 가이드 RNA 라이브러리를 이용하여 유전체 내 모든 유전자의 발현을 체계적으로 억제하고, 개별 유전자의 억제 영향을 평가하는 기법이다 [75]. 유전자를 완전히 제거하는 Tn-Seq과 달리 발현량을 조절할 수 있어 필수 유전자도 발현을 부분적으로 억제하여 기능을 연구할 수 있으며, 더욱 미세한 기능 분석을 가능케 한다. 이러한 전장 스크리닝 기술

들은 균주의 강건성이나 생산성 향상을 위한 예상치 못한 유전자 표적을 발굴하거나, 기능 미상 유전자들의 역할을 규명하는 데 유용하다. 실제로 *E. coli*를 대상으로 한 전장 CRISPRi 스크리닝에서는 세포 생리 향상을 유도하는 숨겨진 유전적 표적을 찾아내어, 지질 대사를 개선함으로써 자유 지방산 생산성을 획기적으로 높인 사례가 있다 [76]. 또 다른 예로, 자가영양 이산화탄소 고정 미생물에 CRISPRi-Seq을 적용하여, CO<sub>2</sub> 조건에서 성장에 필수적인 핵심 유전자 수십 개를 동정하고 미지 유전자들의 역할을 밝힘으로써, 해당 유전자들을 엔지니어링하여 CO<sub>2</sub> 고정 효율을 향상시킨 연구도 보고되었다 [77]. 이처럼 Tn-Seq과 CRISPRi-Seq은 전장 기능 유전체학 수단으로서, 특정 생산 조건에서 이로운 혹은 해로운 유전자를 가려내 세포 공장 최적화의 새로운 표적을 발굴하는 데 기여한다. 이렇게 얻은 지식은 DBTL 사이클의 설계 단계로 다시 반영되어, 보다 효율적인 균주 개발을 가능케 한다.

### 2.5.3. 리보솜 프로파일링(ribosome profiling, Ribo-Seq)

Ribo-Seq은 세포 내 리보솜에 결합한 mRNA 조각들을 추출하여 시퀀싱함으로써, 어떤 mRNA가 실시간 번역되고 있는지를 코돈 단위 해상도로 파악하는 기술이다 [40,78]. 이를 통해 단순 mRNA 존재량이 아니라 실제 번역량을 알 수 있어 전사체와 단백질체 사이의 간극을 메워주는 중요한 정보를 제공한다. 대사공학 관점에서 Ribo-Seq의 활용 가치는 크게 두 가지로 나뉜다. 첫째, 주요 유전자들의 번역 효율 평가이다. Ribo-Seq 데이터(번역 중인 리보솜 수)를 RNA-Seq 데이터(전사량)와 비교하면 각 유전자별 번역 효율(translational efficiency)을 계산할 수 있다. 이를 통해 mRNA는 충분히 존재하지만 리보솜이 잘 붙지 못하는 숨은 병목 단계를 찾고, 기존 유전자 주석에 없던 작은 ORF (open reading frame)들을 발굴하며, 특정 코돈에서 리보솜 정체(pause)가 발생하는 지점을 포착해 아미노산 부족이나 mRNA 2차 구조 문제를 진단할 수 있다 [79]. 예를 들어 번역 중 리보솜이 자주 멈추는 코돈을 Ribo-Seq으로 확인함으로써, 해당 코돈에 대응하는 tRNA (transfer RNA)가 고갈되었거나 mRNA 구조가 번역을 방해함을 추론하고 대책을 세울 수 있다.

둘째, 세포 전체 번역 자원 분배도의 파악이다. 고생산 균주 개발 시 생산과 무관한 단백질 합성에 낭비되는 에너지를 줄이는 것도 중요한데, Ribo-Seq은 세포가 어떤 단백질 번역에 자원을 집중하고 있는지 밝혀내 불필요한 경로 억제 등 대사 부담 완화 전략 수립에 기여한다. 실제 *Pseudomonas putida*에서 전사체(RNA-Seq)와 Ribo-Seq을 통합 분석한 결과, 세포가 주요 중심 대사 경로 단백질 번역을 우선시하는 번역 우선순위화 현상이 관찰되었다 [80]. 즉, 세포는 성장에 필요한 경로의 번역에 자원을 집중하며, 상대적으로 부차적인 경로는 번역 효율이 낮게 유지되었다. 이러한 자원 배분 특성은 생산 품목에 따라 번역 자원을 재분배하는 공학적 전략의 필요성을 시사한다. 예컨대 세포 성장기에 불필요한 구조 단백질 번역을 억제하고, 생산기에 생산효소 번역을 극대

화하는 식의 번역 단계 리디자인이 가능하다.

Ribo-Seq으로 얻은 통찰은 이전에 간과되었던 번역 수준의 병목을 해결하거나 세포 자원 재설계를 가능케 함으로써, 미생물 세포 공장의 새로운 최적화 지평을 열어준다. 예를 들어 번역 과정에서 중요한 사페론이나 번역 조절 인자를 제어함으로써 세포 성장과 생산 간 균형을 조절하거나, 불필요한 단백질 합성을 줄여 대사 부담을 완화하는 전략을 구현할 수 있다 [76]. 이는 기존의 전사 수준 또는 대사 수준 개량만으로는 얻기 어려웠던 성과를 이끌어낼 수 있으며, 궁극적으로 균주의 번역 효율 최적화를 통해 생산성을 극대화하는 데 기여한다.

### 3. 합성생물학: 정밀 구축 및 고속 검증

미생물 세포 공장의 성공적인 상용화 사례들은 ‘어떻게 하면 원하는 대사 경로를 빠르고 정밀하게 하나의 세포 안에 구현할 수 있을까?’라는 공통의 과제를 제시한다. 합성생물학은 생물 시스템을 공학적으로 재설계하고 새로운 기능을 부여하는 학문으로서 [81], 시스템생물학이 ‘설계(design)-학습(learn)’ 단계에서 데이터를 해석하고 최적의 개입점을 제안한다면, 합성생물학은 ‘구축(build)-시험(test)’ 단계에서 해당 설계를 표준화·자동화·대량화 실현으로 이어지도록 한다 [82].

#### 3.1. 유전자 부품 및 회로 구축(build: genetic parts and circuit construction)

##### 3.1.1. 바이오 부품 조립 및 모듈형 클로닝

바이오 부품 조립은 프로모터, 리보솜 결합 서열(ribosome binding site, RBS), 코딩 서열(coding sequence), 터미네이터를 원하는 순서로 연결해 기능성 DNA 구조물을 구성하는 과정을 의미한다 [83]. 초기 BioBrick 규격은 접합부 서열을 표준화해 수천 개의 부품을 레고 블록처럼 손쉽게 공유하고

조립할 수 있도록 했으며, 무작위 조합 방식의 BioBrick assembly를 적용한 리코펜 생산 균주에서는 수율이 30% 이상 증가했다 [84]. 이후 번역 효율을 정밀하게 조절할 수 있는 RBS Calculator [85] 및 UTR Designer [86]와 같은 소프트웨어가 개발되었고, RedLibs 알고리즘은 목표 플렉스에 맞춘 RBS 라이브러리를 자동 생성하여 최소한의 변이로 경로를 최적화할 수 있음을 보여주었다 [87].

DNA 조립 기술은 전통적인 제한효소 기반 클로닝에서 시작되었으며, 이는 특정 염기 서열을 인식하는 제한효소로 DNA를 절단하고 T4 DNA 연결효소(T4 DNA ligase)로 연결하는 방식이다 (Table 2). 그러나 제한효소 클로닝은 제한된 인식 서열과 ‘흉터’ 서열 때문에 확장성이 제한되었고, Gibson assembly가 말단의 상동서열을 이용해 흉터 없이 (scarless) 여러 DNA 조각을 동시에 연결하면서 이러한 한계를 극복했다 [88].

모듈식 클로닝은 앞서 언급한 바이오부품의 표준화 개념과 밀접히 연관되어 있으며, 특히 속도와 재현성을 높이기 위한 클로닝 전략을 지칭한다. 합성생물학 연구자들은 다양한 모듈형 벡터와 클로닝 키트를 개발하여 복잡한 다중 유전자 조립을 체계화 해왔다. Golden Gate(Type IIS 효소) 기반 모듈형 클로닝(modular cloning, MoClo) 체계가 레벨 0 (기본적인 유전자 부품)에서 레벨 2 (다중 유전자 회로나 전체 대사 경로)까지 계층적 조립 방식을 확립해, 다유전자 경로를 단 하루 만에 자동으로 조립·검증할 수 있는 길을 열었다 [89,90]. 이러한 모듈성과 계층성은 높은 조합 유연성을 제공하며, 연구자들은 MoClo Yeast Toolkit (YTK)와 같은 특정 생물에 최적화된 표준화된 toolkit와 Addgene과 같은 플라스미드 저장소를 통해 이러한 부품과 시스템을 손쉽게 공유하고 활용할 수 있다 [84,91].

하지만 여전히 여러 유전자를 순차적으로 연결하는 작업은 시간과 노력이 많이 소요되는 병목이 존재하며, 이를 해

**Table 2.** Key DNA assembly methods and their characteristics

Method	Mechanism	Key enzyme	Scar presence	Throughput	Application
Restriction-ligation cloning	Ligation of DNA fragments with compatible sticky or blunt ends using ligase	Restriction enzymes, T4 DNA ligase	Restriction sites remain as scars	Low; sequential assembly	Subcloning of single genes
Gibson assembly	Joining of DNA fragments with overlapping ends by a mix of exonuclease, polymerase, and ligase	5' exonuclease, DNA polymerase, DNA ligase	None	Medium; suitable for assembling 2–15 fragments	Sequence-independent multi-fragment assembly
Golden Gate (MoClo)	Type IIS restriction enzymes cut outside their recognition sites, generating programmable sticky ends for assembly	Type IIS restriction enzymes ( <i>e.g.</i> , BsaI, BpiI), T4 DNA ligase	None	High; assembly of > 10 parts in a single reaction	Synthetic biology standard; combinatorial library construction; hierarchical assembly
Yeast homologous recombination	<i>In vivo</i> homologous recombination within yeast cells	-	None	High; assembly of > 10 parts in a single step	Biofoundry automated workflows

MoClo, modular cloning



결하기 위해 자동화된 대량 조립 기술이 요구된다. YTK와 Standard European Vector Architecture(SEVA)는 숙주 범용성을 확장했으며 [92,93], 효모의 동형 재조합 능력을 이용한 *in vivo* 원스텝 클로닝은 로봇 공정과 결합돼 초고처리량 라이브러리 구축을 가능하게 했다 [18]. 이 방법은 기존의 복잡한 다단계 클로닝 과정을 혁신적으로 단순화하였고, 로봇 플랫폼에서 고처리량으로 수행될 수 있도록 소형화되었다. 모듈식 표준과 자동화의 결합은 대규모 병렬 실험을 가능케 하여, 수많은 변형체 중 최적해를 식별하는 합성생물학의 탐색 능력을 비약적으로 향상시킨다. 현재 K-Biofoundry를 비롯한 주요 바이오파운드리들은 하루 수백 개의 대사 경로 플라스미드를 자동 제작하며, 이는 설계 공간 탐색 속도를 획기적으로 가속화하는 기반이 되고 있다.

### 3.1.2 유전자 회로 설계(genetic circuit design)

유전자 회로 설계는 세포 내에서 유전자들이 상호작용하며 만들어내는 논리 연산(logic)을 의도적으로 설계하는 것을 의미한다. 세포 공장에서 목표 화합물 생산을 위해 관련 유전자를 단순히 과발현하거나 경쟁 경로 유전자를 제거하는 단순한 유전체 조작은 종종 최적의 결과를 내지 못하는 경우가 많았다. 따라서, 고수를 생산을 달성하려면 단순한 정적 과발현으로는 부족하며, 대사 부담과 독성 중간체 축적을 피하기 위해 동적 유전자 회로가 필요하다 [94].

예를 들어, 퀴럼센싱(quorum-sensing) 신호나 세포 내 화학물 농도를 감지하는 감지-응답 회로는 세포 밀도와 내부 대사 상태에 따라 생산 경로를 자동으로 조절하도록 설계될 수 있다 [95]. 또한, 자연계 세포도 감지-응답 시스템이나 대사 조절 회로를 통해 환경 변화에 적응하므로, 합성생물학자는 이러한 원리를 모방해 맞춤형 유전자 회로를 구축하여 세포 공장에 동적 제어 능력을 부여할 수 있다. 예를 들어, 유전자 회로는 바이오센서(biosensor)를 이용하여 특정 대사물질(전구체, 중간체, 최종 산물 등)의 세포 내 농도를 감지하고, 그에 맞춰 관련 유전자의 발현을 자동으로 조절하는 시스템으로 구성될 수 있다. 예를 들어, 알로스테릭 전사 인자(Allosteric transcription factor)를 이용한 바이오센서 회로는 특정 대사 산물을 인지하는 즉시 전사 활성을 전환해 목표 유전자의 발현을 정밀하게 조절할 수 있으며 [16], GlcN6P 센서를 활용한 N-아세틸글루코사민 생산 공정에서는 외부 유도제 없이도  $97\text{ g L}^{-1}$ 의 높은 생산성이 달성되었다 [96]. 성장 단계와 생산 단계를 자동 분리하는 계층형 회로 역시 아이소프로판올 수율을 크게 증대시킨 바 있다 [97]. 따라서, 유전자 회로는 이러한 문제를 해결하기 위한 동적 조절(dynamic regulation) 전략을 제공할 수 있다.

최근엔 CRISPR 시스템을 활용한 프로그래머블(programmable) 회로도 각광받고 있다. dCas9과 sgRNA (single guide RNA)로 구성된 CRISPRi 회로는 임의의 논리로 유전자를 켜고 끌 수 있어, 세포 내 거의 모든 경로를 동적으로 제어하는 할 수 있음을 보여주었다 [98]. 또한, Cello 플랫폼 같은 *in silico* 설계 도구가 복잡한 논리 회로를 자동으로 설계하는

수준까지 발전했다 [99]. 이러한 동적 제어 기술은 세포 자원 배분을 최적화하고, 대사 플럭스의 변동을 완충해 산업 균주의 안정성과 생산성을 동시에 향상시킬 것으로 기대된다.

### 3.1.3. 균주 개발 전략(strain development strategies)

균주 개발 단계에서는 설계한 경로와 회로를 숙주 미생물에 도입하고, 여러 차례의 개량 과정을 거쳐 목표 산물을 고효율로 생산하는 최종 생산 균주를 만들어내는데, 이 과정에서 다양한 기술과 전략이 요구된다. CRISPR/Cas9 유전체 편집 기술은 모델 미생물뿐 아니라 비모델 중에서도 다중 유전자 삽입 및 삭제를 정밀하게 수행할 수 있도록 고도화되어 왔다 [16,100]. 이로 인해 표적 유전자의 정교한 제거(knock-out), 치환, 삽입(knock-in)이 과거보다 훨씬 수월해졌으며, 다중 유전자 편집 또한 단일 단계에서 가능해짐에 따라 수십 개 유전자를 논리적으로 조합 및 변형하는 대사망상 최적화가 이전보다 획기적으로 가속화되었다. 예를 들어, 균주의 미니멀 채시(minimal chassis) 설계는 불필요한 유전자를 제거해 대사 자원을 제품 합성에 집중시켰고 [101], DNA 절단 기능이 제거된 dCas9 단백질을 이용하는 CRISPRi 및 CRISPRa (CRISPR activation) 기술은 유전체를 영구적으로 바꾸지 않으면서 특정 유전자의 발현을 가역적으로 억제하거나 활성화할 수 있는 도구로 자리 잡았다 [102]. 이는 대사 경로의 각 단계의 발현 수준을 미세하게 조율(fine-tuning)하는 데 매우 유용하게 사용된다.

다양한 바이오 부품의 조합을 시험해야 하는 경우에는 병렬 합성 생물학 라이브러리 전략이 활용될 수 있다. 예를 들어, 프로모터 강도와 RBS 서열을 조합한 수백 종 이상의 경로 조합 라이브러리를 구축하고 이를 고처리량으로 테스트 하면, 실험적 진화를 거치지 않고도 최적의 균주를 효과적으로 선별할 수 있다. 예를 들어, 프로모터와 RBS 조합 라이브러리를 대규모로 구축한 뒤, 머신러닝으로 최적 조합을 예측해 경로를 개선하는 전략이 도입되었다 [103].

설계된 균주가 기대한 성능을 보이지 않을 경우에는 ALE가 적용되기도 한다. 이는 미생물을 일정 조건하에서 장기간 배양하여 자연 돌연변이로 조건 적응을 유도하고, 적응된 균주의 유전체를 해독하여 어떤 변이가 적응에 기여했는지 알 수 있으며, 이렇게 얻은 돌연변이는 생산 균주에 다시 도입되어 성능 향상에 활용된다. 즉, 성능이 낮은 균주에 ALE를 적용하여 내성 돌연변이를 확보하고, 유전체 해독 및 재구현을 통해 강건한 생산 균주로 환원하는 접근이 널리 활용되고 있다 [104]. 이와 같이 축적되는 실험-분석 데이터는 DBTL 사이클의 'Learn' 단계로서 다음 'Design'을 돕게 되며, 자동화 플랫폼과의 결합을 통해 최적 균주 개발에 소요되는 시간이 빠르게 단축되고 있다 [22].

## 3.2. 고속 표현형 분석 및 스크리닝(test: high-throughput phenotyping and screening)

Build 단계를 통해 다양한 설계의 균주가 제작되었다면, 다음 단계에서는 이들이 실제로 원하는 기능을 수행하는지를

평가(test)해야 한다. 전통적으로는 생성된 각 균주를 배양하여 high performance liquid chromatography (HPLC) 등의 분석 장비를 이용해 생산물 농도를 측정하는 방식으로 기능을 평가하였다. 그러나 DBTL 사이클에서 ‘구축’ 단계의 발전은 종종 ‘시험’ 단계의 처리 용량을 초과하게 되며, 이로 인해 ‘시험’ 단계가 전체 공정의 병목으로 작용되는 경우가 많았다 [105]. 이로 인해, 합성생물학 시대에는 수백만에서 수십억 개에 달하는 방대한 유전자 라이브러리를 빠르고 정확하게 평가하기 위한 자동화된 고처리량 스크리닝(high-throughput screening, HTS) 기술이 지속적으로 발전해왔다.

### 3.2.1. 무세포 시스템 기반 프로토타이핑(cell-free system prototyping)

무세포 전사-번역(cell-free protein synthesis, CFPS) 시스템은 세포에서 전사 및 번역에 필요한 요소들(RNA 중합효소, 리보솜, tRNA 등)을 추출하여 시험관 내에서 유전자 발현을 재현하는 기술로, 세포를 배양하지 않고도 전사 및 번역 반응을 수행함으로써 유전자 회로와 대사 경로를 수시간 내에 평가할 수 있는 플랫폼이다 [82]. 기존 접근법은 세포를 직접 형질전환하고 배양하는 과정에 시간이 많이 소요되지만, 무세포 시스템은 세포 내부의 복잡한 생리현상을 제거하고 설계된 유전자 부품의 기능만을 분리하여 시험할 수 있다는 장점이 있다 [106].

이러한 특성을 이용하면 합성 유전자 회로를 세포에 넣기 전에 시험관 내에서 테스트함으로써, 유전자 부품(프로모터, RBS 등)이나 간단한 유전자 회로의 순수한 회로 특성(논리 연산, 응답 곡선 등)을 세포 성장에 영향을 받지 않고 파악(prototyping)할 수 있다. 또한, 몇몇 대사 경로는 중간 산물이 세포에 독성을 나타내거나, 산물이 세포막을 통과하지 못하는 등의 문제가 있는데, 이 때 무세포 시스템을 이용하면 기질과 효소, 보조 인자만 존재하는 환경에서 반응을 관찰할 수 있다. 예를 들어, 700종 이상의 효소 조합을 *in vitro*에서 ‘믹스-앤-매치’한 뒤 최적 세트만 *in vivo*로 도입함으로써 개발 사이클을 단축한 사례가 보고되었고 [106], 리포터 센서를 결합한 무세포 바이오센서는 셀룰라아제 라이브러리를 고속으로 평가하는 데 쓰이고 있다 [107]. 무세포 시스템은 이처럼 경로 프로토타이핑에 활용될 뿐만 아니라, 세포 내에서 일어나기 어려운 신규 대사반응망을 구현하거나 비전형적인 기질과 에너지를 활용한 반응을 설계하는 데에도 응용되고 있다 [106].

향후 무세포 시스템은 단순한 스크리닝을 넘어 설계된 합성 회로나 대사 경로를 신속하게 프로토타이핑할 수 있는 표준 도구로 자리잡을 것으로 기대된다. 이는 기능이 없거나 성능이 낮은 설계를 조기에 걸러내는 ‘빠른 실패(quick-fail)’ 전략을 가능하게 하여, 전체 연구 개발 효율을 향상시킬 수 있다. 특히, 음향 액체 취급 로봇(예: Echo 525[Beckman Coulter Life Sciences, Brea, United States])이나 미세유체(microfluidics) 기술과 결합하면 피코리터(picoliter) 수준의 극소량 공간에서 수많은 CFPS 반응을 동시에 수행할 수 있

어, 고가의 시약 소비를 획기적으로 줄이면서도 처리량을 극대화하는 방향으로 기술이 진화하고 있다 [108].

### 3.2.2. 세포 기반 고처리량 스크리닝 기술(cell-based high-throughput screening)

세포 기반 고처리량 스크리닝은 다수의 변이 균주 또는 유전자 라이브러리를 동시에 또는 신속하게 평가하는 기술들을 의미한다. 전통적인 분석 방법인 HPLC나 Gas Chromatography는 시료 전처리가 복잡하고 분석 시간이 길어, 대규모 라이브러리 스크리닝에는 부적합하다 [107]. 대규모 변이체를 제작하는 것은 상대적으로 쉽기에, 변이체들의 성능을 개별적 및 정량적으로 분석할 수 있는 고처리량 스크리닝 기술의 확보가 매우 중요하다.

이를 해결하기 위한 방법으로는 고처리량 배양 방법 혹은 유전자 회로 도입을 통한 스크리닝 개발이 있다. 로봇 핸들러와 마이크로플레이트 리더를 이용하면 96-well, 384-well 플레이트 상에서 많은 균주를 동시 배양하면서 세포 성장, pH, 형광 등을 시간별로 모니터링할 수 있으며, 이를 통해 전통적인 플라스크 실험 수백 개 분의 데이터를 단일 장비로 처리할 수 있어 스크리닝 효율을 획기적으로 향상시킨다. 더 나아가, 특정 대사 산물의 농도를 형광 신호로 변환하는 유전자 센서를 균주에 도입하면, 생산성이 높은 세포일수록 더 강한 형광을 나타내도록 회로를 설계할 수 있다. 이 경우, 수십만에서 수백만 개의 세포를 유세포 형광 활성화 소팅(fluorescence-activated cell sorting, FACS) 장비를 통해 분당 수만 개 이상 분석할 수 있으며, 이 중 형광이 가장 강한 세포만 자동으로 선별 및 수집할 수 있다 [107,109]. NitR 변이 센서를 사용한 muconic acid 스크리닝이나 [110], *cis*-뮤콘산과 리신의 고생산 균주 탐색 [111]과 같은 예제는 이러한 접근의 효율성을 잘 보여준다. 이러한 biosensor-FACS 플랫폼은 자동화와 결합하여 연구자의 개입 없이도 하루에 수십억 개의 세포 중 최고 성능 몇 개를 골라낼 수 있는 수준으로 발전하고 있다.

이 밖에도 세포 기반 고처리량 스크리닝에는 미세유체(microfluidic) 칩 기술을 활용한 액적(droplet) 스크리닝, 대사 산물을 실시간으로 측정하는 질량 분석 기반 플랫폼, ALE를 자동화한 키모스탯(chemostat) 장비 등도 활용되고 있다 [112]. 핵심은 속도와 규모이며, 이들 기술은 DBTL 사이클 중 Test 단계를 병목 없이 수행함으로써 전체 개발 기간을 단축하는 데 기여하고 있다. 고처리량 분석에서 얻어진 방대한 표현형-유전형 데이터는 다시 머신러닝 등을 통해 분석되어 다음 디자인에 반영되고 있으며, 이러한 데이터 주도 설계로의 전환이 미생물 공장 최적화의 성공률을 향상시키고 있다.

## 4. 미래 전망: 자동화와 바이오파운드리

시스템 및 합성생물학의 궁극적인 목표는 수개월에서 수년이 걸리던 DBTL 사이클을 수일 내로 단축하여 생물 시스템

공학의 혁신을 가속화하는 것이며, 이를 실현하기 위한 핵심 전략은 DBTL 사이클의 모든 단계를 통합하고 자동화하는 것이다 [113]. 이를 위해 앞서 언급한 여러 첨단 기술들이 종합적으로 구현된 장비와 시설이 바로 바이오파운드리이다.

바이오파운드리에는 합성생물학의 DBTL 사이클을 높은 처리량과 자동화(automation)로 실행할 수 있는 통합 플랫폼을 가리키며, 구체적으로, 컴퓨터 소프트웨어가 설계한 수많은 유전자 조합을 로봇이 자동으로 합성·조립하고, 로봇 배양기와 분석기가 균주들을 배양 및 측정한다. 그 결과 데이터를 다시 소프트웨어가 학습하여 다음 실험을 디자인하는 폐쇄 루프 운영을 지향한다. 바이오파운드리의 확산과 고도화를 위해 실험 프로토콜과 워크플로우의 표준화, 그리고 계층적 프레임워크가 제안되고 있다 [114]. 이러한 자동화 파이프라인을 통해 인간 연구자 없이 수백에서 수천 회의 DBTL 사이클을 신속하게 반복할 수 있으므로, 최적화에 걸리는 시간을 획기적으로 단축할 수 있다 [115].

전 세계적으로 바이오파운드리 구축이 활발히 진행되고 있으며, 실제로 미국, 영국, EU, 호주, 중국, 일본 등에서는 국가 단위의 대형 바이오파운드리가 운영되어, 의약품 합성부터 소재 생물공학까지 다양한 프로젝트에 활용되고 있다. 대표적으로 미국의 iBioFAB(일리노이대)과 런던 BioFoundry는 Golden Gate 조립, 나노포어 기반 품질 관리 (quality control, QC), 로봇 배양- 분석을 통합해 자일렌 경로 최적화를 77% 앞당긴 사례가 있다 [116]. 영국의 에든버러 유전체 파운드리(Edinburgh Genome Foundry, EGF)는 대표적인 바이오 파운드리로서, Golden Gate 기반의 DNA 조립 파이프라인을 완전히 자동화했으며, 옥스포드 나노포어(Oxford Nanopore) 기술을 이용한 장문 길이 시퀀싱 기반의 고처리량 QC 시스템을 구축하여 대규모 DNA 구축의 신뢰성을 확보한 바 있다 [117]. 우리나라에서도 한국생명공학연구원(Korea Research Institute of Bioscience and Biotechnology, KRIBB)를 중심으로 K-Biofoundry가 구축되어 가동 중이며, 이는 국가 생물경제 핵심 인프라로서 자동화, 로봇공학, AI를 접목하여 연구자들에게 표준화된 대량 실험 서비스를 제공하는 것을 목표로 하고 있다.

더 나아가, 바이오파운드리의 등장은 필연적으로 AI 및 ML 기술의 도입을 촉진하고 있다 [118,119]. 이는 고처리량 자동화 플랫폼은 인간이 직접 분석하기에는 너무 방대하고 복잡한 바이오 빅데이터(bio-big data)를 생성하기 때문이며, ML 모델은 다중 오믹스 데이터, 스크리닝 결과 등 방대한 데이터셋을 학습하여, 인간이 직관적으로 파악하기 어려운 복잡한 생물학적 '설계 규칙'을 발견할 수 있고, 예측 기반 설계를 촉진한다. 예를 들어 독일 Jülich 연구소에서는 Auto BioTech라는 다목적 바이오파운드리를 구축하여, *E. coli*의 플라스미드 조립 및 CRISPR/Cas9 유전체 편집을 완전 자동화하고 나아가 *Corynebacterium glutamicum*와 같은 그람-양성 균주의 형질전환까지 로봇이 자동으로 수행한다 [91].

이와 같이 ML 알고리즘이 다중 오믹스와 HTS 데이터를

학습하면 다음 설계 규칙을 스스로 추천할 수 있고, 이를 자율적으로 수행함으로써 바이오 제조의 속도와 효율이 비약적으로 향상될 것으로 기대된다 [120]. 결국 표준화, 자동화, AI로 무장한 합성생물학 플랫폼은 산업용 미생물 세포 공장의 설계와 제작 패러다임을 근본적으로 변화시키고 있으며, 지속 가능한 바이오 경제 시대의 핵심 인프라로 자리매김하고 있다 [116,121].

## 5. CONCLUSIONS

미생물 세포 공장 개발은 과거 느리고 예측 불가능했던 개발 방식에서 시스템생물학과 합성생물학의 융합, 그리고 이를 통합하는 DBTL 사이클 프레임워크의 확립을 통해 데이터를 기반으로 합리적으로 설계하고, 표준화된 부품으로 정밀하게 구축하며, 고속으로 검증하고, 그 결과를 체계적으로 학습하여 개선하는 공학적 프로세스로 진화하고 있다.

시스템생물학과 합성생물학의 유기적 통합에도 불구하고 미생물 세포 공장 개발 최적화를 위한 많은 과제가 남아있다. 첫째로 유전체 상에는 아직 그 기능이 규명되지 않은 미지의 유전자들이 상당수 존재하여, 이들이 대사 네트워크와 세포에 미치는 영향을 예측하기 어렵게 만든다 [122]. 둘째, DBTL 사이클의 고도화에도 불구하고 각 단계 간의 발전 속도 불균형은 여전히 해결이 필요한 과제이다. 특히 자동화 기술로 '구축' 단계는 비약적으로 빨라졌으나, 구축된 균주의 표현형을 대량으로 정밀하게 분석하는 '시험'과 그 결과를 모델에 반영하는 '학습' 단계가 병목 현상을 보이고 있다 [123]. 또한, 합성생물학 측면에서도 도입된 유전자 부품 간의 예기치 않은 간섭이나 예상치 못한 돌연변이 발생과 균주 진화와 같은 변수도 고려되어야 한다 [124].

그럼에도 불구하고, 고처리량 실험과 시스템적 접근의 발전은 이러한 난제를 점차 극복할 수 있게 한다. 특히 바이오 파운드리에 AI와 ML을 접목하면 '학습' 단계의 가속화와 예측 모델의 정밀도를 높이는 핵심 동력이 될 것이다. 과거 수년이 걸리던 경로 구축이 이제는 수개월 내 가능해지고 있으며, 세포 전체를 하나의 시스템으로 바라보는 총체적 접근은 미생물 세포 공장을 예측 가능하고 자동화된 공학의 영역으로 끌어올리고 있다. 이러한 기술 혁신은 에너지, 환경, 보건 문제 해결에 기여하며, 지속 가능한 바이오 기반 사회 구현을 위한 핵심 역량으로 자리매김할 것이다.

## Acknowledgements

이 논문은 전남대학교 학술연구비(과제번호: 2025-0394-01) 지원에 의하여 연구되었음. 또한 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2024-00403077, RS-2025-02215308).

## REFERENCES

1. Cho, J. S., G. B. Kim, H. Eun, C. W. Moon, and S. Y. Lee (2022) Designing microbial cell factories for the production of chemicals. *JACS Au*. 2: 1781-1799.
2. Lee, S. Y. and H. U. Kim (2015) Systems strategies for developing industrial microbial strains. *Nat Biotechnol*. 33: 1061-1072.
3. Liu, Y. and J. Nielsen (2019) Recent trends in metabolic engineering of microbial chemical factories. *Curr Opin Biotechnol*. 60: 188-197.
4. Ro, D. K., E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*. 440: 940-943.
5. George, K. W., J. Alonso-Gutierrez, J. D. Keasling, and T. S. Lee (2015) Isoprenoid drugs, biofuels, and chemicals--artemisinin, farnesene, and beyond. *Adv Biochem Eng Biotechnol*. 148: 355-389.
6. Burgard, A., M. J. Burk, R. Osterhout, S. Van Dien, and H. Yim (2016) Development of a commercial scale process for production of 1,4-butanediol from sugar. *Curr Opin Biotechnol*. 42: 118-125.
7. Jakab, K., F. Marga, R. Kaesser, T. H. Chuang, H. Varadaraju, et al. (2019) Non-medical applications of tissue engineering: Biofabrication of a leather-like material. *Mater Today Sustain*. 5: 100018.
8. Kim, G. B., H. R. Kim, and S. Y. Lee (2025) Comprehensive evaluation of the capacities of microbial cell factories. *Nat Commun*. 16: 2869.
9. Wehrs, M., D. Tanjore, T. Eng, J. Lievense, T. R. Pray, and A. Mukhopadhyay (2019) Engineering robust production microbes for large-scale cultivation. *Trends Microbiol*. 27: 524-537.
10. Lennen, R. M., H. G. Lim, K. Jensen, E. T. Mohammed, P. V. Phaneuf, et al. (2023) Laboratory evolution reveals general and specific tolerance mechanisms for commodity chemicals. *Metab Eng*. 76: 179-192.
11. Keasling, J. D. (2012) Synthetic biology and the development of tools for metabolic engineering. *Metab Eng*. 14: 189-195.
12. Zhao, Y., X. Jian, J. Wu, W. Huang, C. Huang, et al. (2019) Elucidation of the biosynthesis pathway and heterologous construction of a sustainable route for producing umbelliferone. *J Biol Eng*. 13: 44.
13. Yu, X., J. Guo, J. Sun, and C. Zhang (2025) Navigable genome engineering: Stepwise correlation for precision-guided optimization of microbial cell factory phenotypes. *Synth Biol Eng*. 3: 10003.
14. Wan, S., X. Liu, W. Sun, B. Lv, and C. Li (2023) Current advances for omics-guided process optimization of microbial manufacturing. *Bioresour Bioprocess*. 10: 30.
15. Rowlands, R. T. (1984) Industrial strain improvement: Mutagenesis and random screening procedures. *Enzyme Microb Technol*. 6: 3-10.
16. Ko, Y. S., J. W. Kim, J. A. Lee, T. Han, G. B. Kim, et al. (2020) Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chem Soc Rev*. 49: 4615-4636.
17. Amer, B. and E. E. K. Baidoo (2021) Omics-driven biotechnology for industrial applications. *Front Bioeng Biotechnol*. 9: 613307.
18. Seong, M. J., Y. R. Yoon, K. K. Kwon, H. Kim, S. G. Lee, et al. (2025) Automated construction of a yeast-based multigene library via homologous recombination in a biofoundry workflow. *ACS Synth Biol*. 14: 1549-1556.
19. Kwon, K. K., J. Lee, H. Kim, D. H. Lee, and S. G. Lee (2024) Advancing high-throughput screening systems for synthetic biology and biofoundry. *Curr Opin Syst Biol*. 37: 100487.
20. Kim, Y. M., C. J. Petzold, E. J. Kerkhoven, and S. E. Baker (2021) Editorial: Multi-omics technologies for optimizing synthetic biomanufacturing. *Front Bioeng Biotechnol*. 9: 818010.
21. Jeon, S., Y. J. Sohn, H. Lee, J. Y. Park, D. Kim, et al. (2025) Recent advances in the design-build-test-learn (DBTL) cycle for systems metabolic engineering of *Corynebacterium glutamicum*. *J Microbiol*. 63: e2501021.
22. van Lent, P., J. Schmitz, and T. Abeel (2023) Simulated design-build-test-learn cycles for consistent comparison of machine learning methods in metabolic engineering. *ACS Synth Biol*. 12: 2588-2599.
23. Westerhoff, H. V. and B. O. Palsson (2004) The evolution of molecular biology into systems biology. *Nat Biotechnol*. 22: 1249-1252.
24. Choi, K. R., J. H. Shin, J. S. Cho, D. Yang, S. Y. Lee (2017) Systems metabolic engineering of *Escherichia coli*. *EcoSal Plus*. 7.
25. Orth, J. D., I. Thiele, and B. Ø. Palsson (2010) What is flux balance analysis? *Nat Biotechnol*. 28: 245-248.
26. Oberhardt, M. A., B. Ø. Palsson, and J. A. Papin (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol*. 5: 320.
27. Gu, C., G. B. Kim, W. J. Kim, H. U. Kim, and S. Y. Lee (2019) Current status and applications of genome-scale metabolic models. *Genome Biol*. 20: 121.
28. Kim, B., W. J. Kim, D. I. Kim, and S. Y. Lee (2015) Applications of genome-scale metabolic network model in metabolic engineering. *J Ind Microbiol Biotechnol*. 42: 339-348.
29. Burgard, A. P., P. Pharkya, and C. D. Maranas (2003) Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng*. 84: 647-657.
30. Park, J. H., K. H. Lee, T. Y. Kim, and S. Y. Lee (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.
31. Fang, L., J. Fan, S. Luo, Y. Chen, C. Wang, et al. (2021) Genome-scale target identification in *Escherichia coli* for high-titer production of free fatty acids. *Nat Commun*. 12: 4976.
32. Sánchez, B. J., C. Zhang, A. Nilsson, P. J. Lahtvee, E. J. Kerkhoven, and J. Nielsen (2017) Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol Syst Biol*. 13: 935.
33. Li, F., L. Yuan, H. Lu, G. Li, Y. Chen, et al. (2022) Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. *Nat Catal*. 5: 662-672.
34. Kim, M., B. G. Park, J. Kim, J. Y. Kim, and B. G. Kim (2018) Exploiting transcriptomic data for metabolic engineering: Toward

- a systematic strain design. *Curr Opin Biotechnol.* 54: 26-32.
35. Rodriguez, A., Y. Chen, S. Khoomrung, E. Özdemir, I. Borodina, and J. Nielsen (2017) Comparison of the metabolic response to over-production of p-coumaric acid in two yeast strains. *Metab Eng.* 44: 265-272.
  36. Yi, J. S., M. W. Kim, M. Kim, Y. Jeong, E. J. Kim, et al. (2017) A novel approach for gene expression optimization through native promoter and 5' UTR combinations based on RNA-seq, Ribo-seq, and TSS-seq of *Streptomyces coelicolor*. *ACS Synth Biol.* 6: 555-565.
  37. Kim, M., J. S. Yi, M. Lakshmanan, D. Y. Lee, and B. G. Kim (2016) Transcriptomics-based strain optimization tool for designing secondary metabolite overproducing strains of *Streptomyces coelicolor*. *Biotechnol Bioeng.* 113: 651-660.
  38. Sastry, A. V., Y. Gao, R. Szubin, Y. Hefner, S. Xu, et al. (2019) The *Escherichia coli* transcriptome mostly consists of independently regulated modules. *Nat Commun.* 10: 5536.
  39. Rychel, K., K. Decker, A. V. Sastry, P. V. Phaneuf, S. Poudel, and B. O. Palsson (2021) iModulonDB: A knowledgebase of microbial transcriptional regulation derived from machine learning. *Nucleic Acids Res.* 49: D112-D120.
  40. Calviello, L. and U. Ohler (2017) Beyond read-counts: Ribo-seq data analysis to understand the functions of the transcriptome. *Trends Genet.* 33: 728-744.
  41. Landels, A., C. Evans, J. Noirel, and P. C. Wright (2015) Advances in proteomics for production strain analysis. *Curr Opin Biotechnol.* 35: 111-117.
  42. Rajacharya, G. H., A. Sharma, and S. S. Yazdani (2024) Proteomics and metabolic burden analysis to understand the impact of recombinant protein production in *E. coli*. *Sci Rep.* 14: 12271.
  43. Snoeck, S., C. Guidi, and M. De Mey (2024) "Metabolic burden" explained: Stress symptoms and its related responses induced by (over)expression of (heterologous) proteins in *Escherichia coli*. *Microb Cell Fact.* 23: 96.
  44. Yunus, I. S. and T. S. Lee (2022) Applications of targeted proteomics in metabolic engineering: advances and opportunities. *Curr Opin Biotechnol.* 75: 102709.
  45. Redding-Johanson, A. M., T. S. Bath, R. Chan, R. Krupa, H. L. Szmidi, et al. (2011) Targeted proteomics for metabolic pathway optimization: Application to terpene production. *Metab Eng.* 13: 194-203.
  46. Neef, J., J. M. van Dijk, and G. Buist (2021) Recombinant protein secretion by *Bacillus subtilis* and *Lactococcus lactis*: Pathways, applications, and innovation potential. *Essays Biochem.* 65: 187-195.
  47. Neef, J., C. Bongioni, B. Schmidt, V. J. Goossens, and J. M. van Dijk (2020) Relative contributions of non-essential Sec pathway components and cell envelope-associated proteases to high-level enzyme secretion by *Bacillus subtilis*. *Microb Cell Fact.* 19: 52.
  48. Champomier-Vergès, M. C., E. Maguin, M. Y. Mistou, P. Anglade, and J. F. Chich (2002) Lactic acid bacteria and proteomics: Current knowledge and perspectives. *J Chromatogr B Analyt Technol Biomed Life Sci.* 771: 329-342.
  49. Rosen, R. and E. Z. Ron (2002) Proteome analysis in the study of the bacterial heat-shock response. *Mass Spectrom Rev.* 21: 244-265.
  50. Lee, S. Y., D. Y. Lee, and T. Y. Kim (2005) Systems biotechnology for strain improvement. *Trends Biotechnol.* 23: 349-358.
  51. Holbrook-Smith, D., J. Trouillon, and U. Sauer (2024) Metabolomics and microbial metabolism: Toward a systematic understanding. *Annu Rev Biophys.* 53: 41-64.
  52. Tanaka, K., T. Bamba, A. Kondo, and T. Hasunuma (2024) Metabolomics-based development of bioproduction processes toward industrial-scale production. *Curr Opin Biotechnol.* 85: 103057.
  53. J. A. M. Kaste and Y. Shachar-Hill (2024) Model validation and selection in metabolic flux analysis and flux balance analysis. *Biotechnol Prog.* 40: e3413.
  54. Cortada-Garcia, J., R. Daly, S. A. Arnold, and K. Burgess (2023) Streamlined identification of strain engineering targets for bioprocess improvement using metabolic pathway enrichment analysis. *Sci Rep.* 13: 12990.
  55. Barton, N. R., A. P. Burgard, M. J. Burk, J. S. Crater, R. E. Osterhout, et al. (2015) An integrated biotechnology platform for developing sustainable chemical processes. *J Ind Microbiol Biotechnol.* 42: 349-360.
  56. Cortada-Garcia, J., J. Haggarty, T. Moses, R. Daly, S. A. Arnold, and K. Burgess (2022) On-line untargeted metabolomics monitoring of an *Escherichia coli* succinate fermentation process. *Biotechnol Bioeng.* 119: 2757-2769.
  57. Xu, P., N. Q. Lin, Z. Q. Zhang, and J. Z. Liu (2024) Strategies to increase the robustness of microbial cell factories. *Adv Biotechnol* 2: 9.
  58. Valle, A., Z. Soto, H. Muhamadali, K. A. Hollywood, Y. Xu, et al. (2022) Metabolomics for the design of new metabolic engineering strategies for improving aerobic succinic acid production in *Escherichia coli*. *Metabolomics.* 18: 56.
  59. Hu, T., M. Allam, S. Cai, W. Henderson, B. Yueh, et al. (2023) Single-cell spatial metabolomics with cell-type specific protein profiling for tissue systems biology. *Nat Commun.* 14: 8260.
  60. Alexandrov, T. (2023) Spatial metabolomics: From a niche field towards a driver of innovation. *Nat Metab.* 5: 1443-1445.
  61. St John, P. C. and Y. J. Bomble (2019) Approaches to computational strain design in the multiomics era. *Front Microbiol.* 10: 597.
  62. Costello, Z. and H. G. Martin (2018) A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst Biol Appl.* 4: 19.
  63. Gao, Y., H. G. Lim, H. Verkler, R. Szubin, D. Quach, et al. (2021) Unraveling the functions of uncharacterized transcription factors in *Escherichia coli* using ChIP-exo. *Nucleic Acids Res.* 49: 9696-9710.
  64. Park, P. J. (2009) ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet.* 10: 669-680.
  65. Teixeira, M. C., R. Viana, M. Palma, J. Oliveira, M. Galocha, et al. (2023) YEASTRACT+: A portal for the exploitation of global transcription regulation and metabolic model data in yeast biotechnology and pathogenesis. *Nucleic Acids Res.* 51: D785-D791.
  66. Salgado, H., S. Gama-Castro, P. Lara, C. Mejia-Almonte, G. Alarcón-Carranza, et al. (2024) RegulonDB v12.0: A comprehensive

- sive resource of transcriptional regulation in *E. coli* K-12. *Nucleic Acids Res.* 52: D255-D264.
67. Park, J. Y., M. Jang, E. Choi, S. M. Lee, I. Bang, et al. (2025) ChIP-mini: A low-input ChIP-exo protocol for elucidating DNA-binding protein dynamics in intracellular pathogens. *Nucleic Acids Res.* 53: gkaf009.
  68. Wang, K., X. Zeng, J. Zhou, F. Liu, X. Luan, and X. Wang (2024) BERT-TFBS: A novel BERT-based model for predicting transcription factor binding sites by transfer learning. *Brief Bioinform.* 25: bbae195.
  69. G.C. diCenzo, A. Mengoni, and M. Fondi (2019) Tn-core: A toolbox for integrating Tn-seq gene essentiality data and constraint-based metabolic modeling. *ACS Synth. Biol.* 8: 158-169.
  70. Chen, X., F. Li, X. Li, M. Otto, Y. Chen, V. Siewers (2025) Model-assisted CRISPRi/a library screening reveals central carbon metabolic targets for enhanced recombinant protein production in yeast. *Metab Eng.* 88: 1-13.
  71. Goodall, E. C. A., A. Robinson, I. G. Johnston, S. Jabbari, K. A. Turner, et al. (2018) The essential genome of *Escherichia coli* K-12. *mBio.* 9: e02096-17.
  72. Torres, M., S. Paszti, and L. Eberl (2024) Shedding light on bacteria-host interactions with the aid of TnSeq approaches. *mBio.* 15: e0039024.
  73. Borchert, A. J., A. Bleem, and G. T. Beckham (2023) RB-TnSeq identifies genetic targets for improved tolerance of *Pseudomonas putida* towards compounds relevant to lignin conversion. *Metab Eng.* 77: 208-218.
  74. Nolan, L. M., M. A. Webber, and A. Filloux (2025) Throwing a spotlight on genomic dark matter: The power and potential of transposon-insertion sequencing. *J Biol Chem.* 301: 110231.
  75. de Bakker, V., X. Liu, A. M. Bravo, and J. W. Veening (2022) CRISPRi-seq for genome-wide fitness quantification in bacteria. *Nat Protoc.* 17: 252-281.
  76. Fang, L., X. Hao, J. Fan, X. Liu, Y. Chen, et al. (2025) Genome-scale CRISPRi screen identifies pcnB repression conferring improved physiology for overproduction of free fatty acids in *Escherichia coli*. *Nat Commun.* 16: 3060.
  77. Shin, J., J. Bae, H. Lee, S. Kang, S. Jin, et al. (2023) Genome-wide CRISPRi screen identifies enhanced autolithotrophic phenotypes in acetogenic bacterium *Eubacterium limosum*. *Proc Natl Acad Sci U S A.* 120: e2216244120.
  78. Ingolia, N. T., S. Ghaemmhami, J. R. Newman, and J. S. Weissman (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science.* 324: 218-223.
  79. Arpat, A. B., A. Liechti, M. De Matos, R. Dreos, P. Janich, and D. Gatfield (2020) Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing. *Genome Res.* 30: 985-999.
  80. Tibocho-Bonilla, J. D., V. Gandhi, C. Lieng, O. Moyne, R. Santibáñez-Palominos, and K. Zengler (2025) Model of metabolism and gene expression predicts proteome allocation in *Pseudomonas putida*. *NPJ Syst Biol Appl.* 11: 55.
  81. Endy, D. (2005) Foundations for engineering biology. *Nature.* 438: 449-453.
  82. Laohakunakorn, N. (2020) Cell-free systems: A proving ground for rational biodesign. *Front Bioeng Biotechnol.* 8: 788.
  83. Canton, B., A. Labno, and D. Endy (2008) Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol.* 26: 787-793.
  84. Iverson, S. V., T. L. Haddock, J. Beal, and D. M. Densmore (2016) CIDAR MoClo: Improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. *ACS Synth Biol.* 5: 99-103.
  85. Salis, H. M., E. A. Mirsky, and C. A. Voigt (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol.* 27: 946-950.
  86. Seo, S. W., J. S. Yang, I. Kim, J. Yang, B. E. Min, et al. (2013) Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metab Eng.* 15: 67-74.
  87. Jeschek, M., D. Gerngross, and S. Panke (2016) Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat Commun.* 7: 11163.
  88. Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, and H. O. Smith (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 6: 343-345.
  89. Engler, C., R. Kandzia, and S. Marillonnet (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One.* 3: e3647.
  90. Blázquez, B., D. S. León, J. Torres-Bacete, Á. Gómez-Luengo, R. Kniewel, et al. (2023) Golden standard: A complete standard, portable, and interoperable MoClo tool for model and non-model proteobacteria. *Nucleic Acids Res.* 51: e98.
  91. Rosch, T. M., J. Tenhaef, T. Stoltmann, T. Redeker, D. Kösters, et al. (2024) AutoBioTech—a versatile biofoundry for automated strain engineering. *ACS Synth Biol.* 13: 2227-2237.
  92. Silva-Rocha, R., E. Martínez-García, B. Calles, M. Chavarría, A. Arce-Rodríguez, et al. (2013) The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 41: D666-D675.
  93. Lee, M. E., W. C. DeLoache, B. Cervantes, and J. E. Dueber (2015) A highly characterized yeast toolkit for modular, multipart assembly. *ACS Synth Biol.* 4: 975-986.
  94. Cress, B. F., E. A. Trantas, F. Ververidis, R. J. Linhardt, and M. A. Koffas (2015) Sensitive cells: Enabling tools for static and dynamic control of microbial metabolic pathways. *Curr Opin Biotechnol.* 36: 205-214.
  95. Gupta, A., I. M. Reizman, C. R. Reisch, and K. L. Prather (2017) Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol.* 35: 273-279.
  96. Wu, Y., T. Chen, Y. Liu, R. Tian, X. Lv, et al. (2020) Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in *Bacillus subtilis*. *Nucleic Acids Res.* 48: 996-1009.
  97. Doong, S. J., A. Gupta, and K. L. J. Prather (2018) Layered dynamic regulation for improving metabolic pathway productivity in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 115: 2964-2969.



98. Nielsen, A. A. and C. A. Voigt (2014) Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol.* 10: 763.
99. Nielsen, A. A., B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, et al. (2016) Genetic circuit design automation. *Science.* 352: aac7341.
100. Jakóčiūnas, T., M. K. Jensen, and J. D. Keasling (2016) CRISPR/Cas9 advances engineering of microbial cell factories. *Metab Eng.* 34: 44-59.
101. Chen, X. R., Y. Z., B. Z. Li, and Y. J. Yuan (2024) Genome engineering on size reduction and complexity simplification: A review. *J Adv Res.* 60: 159-171.
102. Choi, K. R. and S. Y. Lee (2016) CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol Adv.* 34: 1180-1209.
103. Naseri, G. and M. A. G. Koffas (2020) Application of combinatorial optimization strategies in synthetic biology. *Nat Commun.* 11: 2446.
104. Sandberg, T. E., M. J. Salazar, L. L. Weng, B. O. Palsson, and A. M. Feist (2019) The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab Eng.* 56: 1-16.
105. Petzold, C. J., L. J. Chan, M. Nhan, and P. D. Adams (2015) Analytics for metabolic engineering. *Front Bioeng Biotechnol.* 3: 135.
106. Rasor, B. J. and T. J. Erb (2025) Cell-free systems to mimic and expand metabolism. *ACS Synth Biol.* 14: 316-322.
107. Kaczmarek, J. A. and K. L. J. Prather (2021) Effective use of biosensors for high-throughput library screening for metabolite production. *J Ind Microbiol Biotechnol.* 48: kuab049.
108. Baranwal, A. K. and S. J. Maerkl (2024) A comprehensive review of microfluidic approaches in cell-free synthetic biology. *Front Synth Biol.* 2.
109. Teng, Y., J. Zhang, T. Jiang, Y. Zou, X. Gong, and Y. Yan (2022) Biosensor-enabled pathway optimization in metabolic engineering. *Curr Opin Biotechnol.* 75: 102696.
110. Wang, G., S. Özmerih, R. Guerreiro, A. C. Meireles, A. Carolas, et al. (2020) Improvement of *cis,cis*-muconic acid production in *Saccharomyces cerevisiae* through biosensor-aided genome engineering. *ACS Synth Biol.* 9: 634-646.
111. Bentley, G. J., N. Narayanan, R. K. Jha, D. Salvachúa, J. R. Elmore, et al. (2020) Engineering glucose metabolism for enhanced muconic acid production in *Pseudomonas putida* KT2440. *Metab Eng.* 59: 64-75.
112. Zhang, Z., Q. Guo, Y. Wang, and H. Huang (2023) High-throughput screening of microbial strains in large-scale microfluidic droplets. *Front Bioeng Biotechnol.* 11: 1105277.
113. Ma, Y., Z. Zhang, B. Jia, and Y. Yuan (2024) Automated high-throughput DNA synthesis and assembly. *Heliyon.* 10: e26967.
114. Kim, H., N. J. Hillson, B. K. Cho, B. H. Sung, D. H. Lee, et al. (2025) Abstraction hierarchy to define biofoundry workflows and operations for interoperable synthetic biology research and applications. *Nat Commun.* 16: 6056.
115. Chao, R., S. Mishra, T. Si, and H. Zhao (2017) Engineering biological systems using automated biofoundries. *Metab Eng.* 42: 98-108.
116. Hamedirad, M., R. Chao, S. Weisberg, J. Lian, S. Sinha, and H. Zhao (2019) Towards a fully automated algorithm driven platform for biosystems design. *Nat Commun.* 10: 5150.
117. Vegh, P., S. Donovan, S. Rosser, G. Stracquadanio, and R. Fragkoudis (2024) Biofoundry-scale DNA assembly validation using cost-effective high-throughput long-read sequencing. *ACS Synth Biol.* 13: 683-686.
118. Cheng, Y., X. Bi, Y. Xu, Y. Liu, J. Li, et al. (2023) Machine learning for metabolic pathway optimization: A review. *Comput Struct Biotechnol J.* 21: 2381-2393.
119. Zhang, Q., W. Chen, M. Qin, Y. Wang, Z. Pu, et al. (2025) Integrating protein language models and automatic biofoundry for enhanced protein evolution. *Nat Commun.* 16: 1553.
120. Gurdo, N., D. C. Volke, and P. I. Nikel (2022) Merging automation and fundamental discovery into the design-build-test-learn cycle of nontraditional microbes. *Trends Biotechnol.* 40: 1148-1159.
121. Kim, G. B., S. Y. Choi, I. J. Cho, D. H. Ahn, and S. Y. Lee (2023) Metabolic engineering for sustainability and health. *Trends Biotechnol.* 41: 425-451.
122. Vanni, C., M. S. Schechter, S. G. Acinas, A. Barberán, P. L. Buttigieg, et al. (2022) Unifying the known and unknown microbial coding sequence space. *Elife.* 11: e67667.
123. Raj, K., N. Venayak, P. Diep, S. A. Golla, A. F. Yakunin, and R. Mahadevan (2021) Automation assisted anaerobic phenotyping for metabolic engineering. *Microb Cell Fact.* 20: 184.
124. Zhang, C., R. Tsoi, and L. You (2016) Addressing biological uncertainties in engineering gene circuits. *Integr Biol.* 8: 456-464.