Genome Editing for the Production of Natural Products in Escherichia coli

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

OPEN ACCESS

1 # Genome Editing for the Production of Natural Products in *Escherichia coli*

2 Ziga Zebec and Nigel S. Scrutton

3 Manchester Institute of Biotechnology and School of Chemistry, The University of Manchester, Manchester,
4 M1 7DN United Kingdom

5 Correspondence to either nigel.scrutton@manchester.ac.uk or zigazebec@gmail.com

6 **Abstract**

7 Natural products such as secondary metabolites (e.g. plant terpenoids) have been found to be a major source
8 of bioactive compounds. These natural products accumulate as complex mixtures with other related
9 compounds and this chemical complexity adds cost to the downstream recovery and purification of natural
10 products from plant biomass. One aim of synthetic biology and metabolic engineering programmes is to
11 produce such compounds from synthetic gene clusters (SGC) in heterologous hosts and thereby achieve more
12 targeted and affordable production. Both fungi and bacteria are common hosts for metabolic engineering in
13 industry. Fungal hosts include *Penicillium chrysogenum, Saccharomyces cerevisiae, Aspergillus niger* and the
14 bacterial hosts *Escherichia coli, Bacillus subtilis, Corynebacterium glutamicum*.[1] *E. coli* is often selected as a
15 host given the ease of its genetic manipulation and the long history [2] [3] of using this organism in laboratory-
16 based bioengineering.[4] [5] The bioengineering of *E. coli* extends also to feedstock pathways to interface and
17 optimise the production of high value compounds from widely available and inexpensive carbon sources.
18 Genome editing is important in these microbial bioengineering programmes and is needed to isolate stable
19 strains and to optimise production. Here we discuss frequently used methods for genome editing in *E. coli* in
20 relation to the production of natural compounds and chemicals.

21

22 **Introduction**

23 Secondary metabolites are high value natural chemicals that are commercially used, for example in the food-
24 flavouring, food-additive, fragrance, pharmaceutical, biochemical and biomedical industries. These natural
25 compounds can be extracted from plant biomass by steam distillation as complex mixtures known as 'Essential
26 Oils'.[6] Many of the natural products, like terpenes and their derivatives (terpenoids), have been tested for
27 potential use as bioactive chemicals and drug candidates for a variety of ailments and many have been shown
28 to have potential benefits in major diseases (e.g. cancer [7] [8] and Alzheimer's[9]) and to possess antimicrobial
29 activity against multi-drug resistant bacterial pathogens,[10] but most compounds still have not been clinically
30 tested. Most secondary metabolites cannot be extracted from plants as pure compounds and are usually
31 isolated as complex mixtures of natural chemicals.[6] The composition of such natural extracts (e.g. Essential
32 Oils) depends on the geographic location and vegetative stage of plants during harvest.[11] An additional
33 problem is the availability of these compounds since crop yields and access to biomass may fluctuate due to
34 environmental effects or other factors (e.g. disease, land repurposing, political barriers).[12] Many compounds

are also present in relatively low quantities and intensive purification is required to recover these products from plant biomass.[6] Alternative routes to sourcing these compounds using synthetic chemistry can be expensive and environmentally costly.[13] These issues have led to an increased emphasis on the production of natural compounds in microbial cells from fermentable and low cost carbon sources.[14] The challenge here is to redirect carbon fluxes from cellular metabolism towards the desired product but is limited by the extent to which natural metabolic pathways – encoded by genetic information contained in the chromosome – can be engineered for this purpose.[15] Production in microbial hosts also presents new downstream processing challenges for recovery of these products from microbial bioreactors. These challenges are often distinct from those associated with purification following production using synthetic chemistry approaches, or from agricultural biomass.

Synthetic biology and metabolic engineering approaches allow one to introduce new 'parts' to the existing genetic frame of a recipient organism enabling extension of existing microbial metabolic frames. These parts are generally organised in synthetic gene clusters (SGC), which comprise the coding sequences for component enzymes and associated regulators (e.g. promoters, ribosome binding sites) required for gene expression. Not only does this enable production of any naturally occurring compound from a known metabolic pathway, but additionally it facilitates production through *de novo* pathway engineering.[1] A crucial requirement for microbial production is a genetically stable production host, which can maintain its genetic and phenotypic form over time to support optimal production.[16] [17] Extracellular DNA in the form of a plasmid is the fastest way to introduce and test a variety of new synthetic pathways in *E. coli*.[18] However, the continuous use of antibiotics in industrial-scale operations is not acceptable for environmental, cost and stability reasons[19] and plasmids are prone to loose productivity after 20 generations due to unequal plasmid segregation.[16] In this review, we focus on strategies that allow chromosomal manipulations in *E. coli* to i) redirect/manipulate the metabolic frame towards the desired product, and ii) insert genetic information for a heterogeneous pathway (i.e. a novel SGC comprising coding sequences and regulatory parts) to produce a stable production strain. We examine the most recent and prominent cases of genetic frame engineering customized for the production of high value natural products through the insertion of heterologous SGCs. Typically this kind of genetic manipulation requires a set of tools to facilitate genetic manipulations *via* recombination-mediated genetic engineering or 'recombineering'.

Evolving the 'genetic frame'

Synthetic biology is a rapidly expanding research field. One aim of this field is to establish routes to stable, pure, natural and environmentally friendly sources of high value natural compounds by production in microbial cell factories. This process starts usually with the selection of a waste carbon source such as cane sugar by-products[20] or crude glycerol[21] to support fermentative growth through primary metabolic pathways.[22] The classical Darwinian approach of adaptation and selection has historically been an important and powerful tool for natural genome engineering, which is limited by the organisms genomic content.[23] This has been explored from various angles, including the isolation of varieties and strains from the environment with acquired and

1 evolved metabolic pathways. For example, a new *E. coli* strain (strain ZH-4) was isolated from bovine rumen

2 and is capable of cellulose utilisation and degradation.[24] As cellulosic wastes are a large, natural and

3 renewable carbon source this shows potential for this metabolic adaption to serve as a genetic framework for

4 further strain modifications.[25] Also, directed evolution in a xylose-based chemostat has enabled the *E. coli*

5 genome to utilize CO_2 and to generate glucose. This was achieved using a heterologous Calvin-Benson-

6 Bassham (CBB) cycle, with RuBisCO, phosphoribulokinase (prk), and carbonic anhydrase expressed from a

7 plasmid.[26] Laboratory-based evolution of *E. coli* has been used to introduce mutations in eight different genes

8 (prs, pgi, serA, glmU, crp, ppsR, xylA, and malT) involved in regulating fluxes between the CBB and tricarboxylic

9 acid (TCA) cycles. However, a genetic reconstruction showed that only five of the genes are sufficient to

10 restore the CO_2 utilising phenotype of *E. coli*.[27] Nature can therefore provide a roadmap for how to tap a

11 carbon source that is not only abundant but is also a major industrial waste product. Laboratory evolution in *E.*

12 *coli* shows that a natural means of genome engineering is important, as it does not require reverse engineering

13 of metabolic pathways to utilise the selected carbon source. [28] Strain evolution using adaptation and selection

14 therefore provides a non-invasive method to engineer the *E. coli* genetic frame to utilise different waste

15 carbon sources and to build up a pool of acetyl-CoA to support metabolic demand. As such, this is an excellent

16 starting point from which to complement this strain by introducing SGCs for the production of complex and

17 valuable natural products and chemicals.

18

19
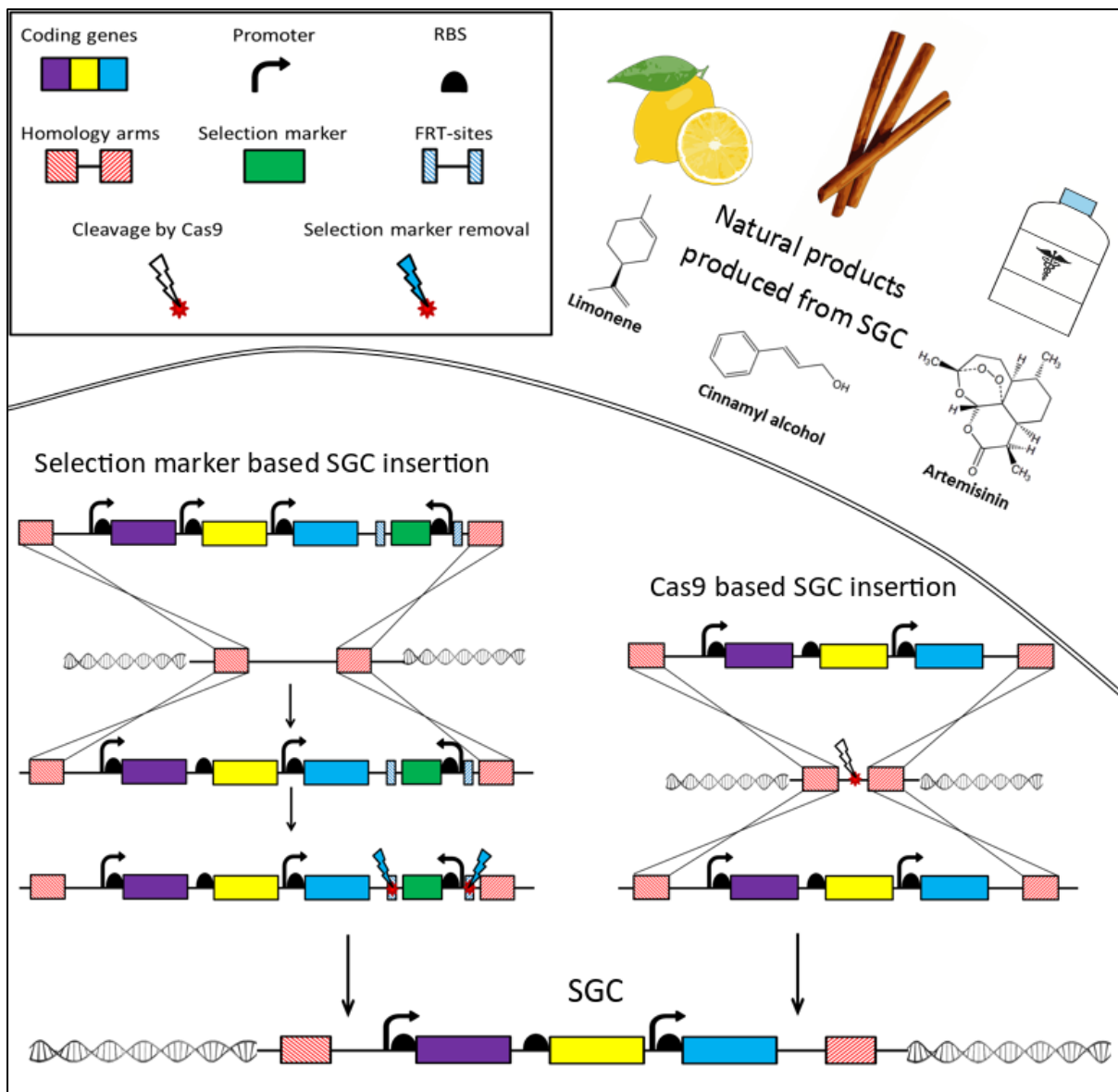
Figure 1. Schematic representation of two of the most common strategies used to insert a synthetic gene cluster (SGC) in *E. coli* with λ-RED. Selection marker based insertion (*, table 1) or Cas9 based insertion (**, table1). Abbreviations are as followed: FRT = flippase recombination target sites; RBS = ribosome binding site

Genome reorganisation and optimisation with the λ-RED system

Genome engineering by adaptation and selection is clearly one route to developing *E. coli* strains with a customised genetic frame for use with a carbon feedstock. However, distribution of the incorporated carbon from the acetyl-CoA pool to downstream production pathways requires more targeted fine-tuning of genes or gene circuits involved in those metabolic pathways.[29] Frequently, this is achieved using the λ-RED system to integrate recombinant and synthetic double stranded (ds) DNA into the *E. coli* chromosome – a method known as 'recombineering' (Figure 1). The λ-RED system was discovered on the linear DNA of the λ bacteriophage and is essential for the virus to transition between lysogenic and the lytic cycles since it enables viral integration into the *E. coli* chromosome.[30] The λ-RED system comprises three independent subunits (Beta, Exo and Gam) where Gam prevents degradation of linear dsDNA by the RecBCD nuclease, Exo (exonuclease)

1   degrades dsDNA in a 5' to 3' manner to produce single-stranded (ss) DNA templates, and Beta binds to ssDNA

2   to facilitate annealing of homology arms (HA) of a minimum 35nt to repair the chromosome during

3   replication.[31] Curing of the remaining plasmids (e.g. helper or donor plasmid) is usually achieved through the

4   use of a thermo-sensitive origin of replication pSC101. This disables plasmid segregation during replication at

5   42˚C and thereby results in plasmid loss.[32] Small changes to the chromosome can be achieved using the λ-RED

6   system by directing ssDNA (oligos) to the lagging strand of the replication fork in a process called 'oligo-

7   mediated allelic replacement'.[3] [33] This approach has been the cornerstone of modern genome engineering

8   for example in the first multiplexed automated genome engineering (MAGE) platform.[34] The study

9   demonstrated that with sophisticated computational design of 90-mer oligos up to 20 genes can be altered

10   simultaneously with an efficiency of 30% (in 2 to 2.5h) by introducing mutations of up to 20 nucleotides or

11   deletions of up to 100nt.[34] Targeted genes located in the 1-deoxy-D-xylulose-5-phosphate pathway were

12   evolved to genetically engineer strains that have up to 5 times higher production levels of the isoprene

13   lycopene compared to the ancestor strain. The natural product lycopene is chemically a tetraterpene

14   assembled from eight isoprene subunits which derive from acetyl-CoA via the mevalonate pathway (MVA) and

15   is most commonly found as a red pigment in fruits like tomatoes. Lycopene is industrially most commonly used

16   as a biological colouring agent. A complementary genome-wide method has also been developed called

17   'traceable multiplex recombineering' (TRMR) that relies on parallel DNA synthesis and recombineering for

18   molecular barcoding.[35] This method changed the expression profile of 95 % of all genes in *E. coli* in 24h and

19   was utilized for mapping gene-traits of 8,000 distinct mutations in seven different environments.[35] Genome

20   wide stop codon replacement (TGA to TAA) has also been achieved using MAGE[36] and was further optimized

21   by manipulating the replication process.[37] Methyl-directed mismatch repair (MMR) systems have also been

22   used for transient suppression of DNA repair systems,[38] which can be applied to generate sets of genetic

23   frames in *E. coli* with wide phenotypical diversity.[39] These strategies can be performed in a high throughput

24   (HTP) manner – generating enormous genetic and phenotypic diversity in *E. coli* – but are limited to maximum

25   insertions of around 20nt.

26

27   **λ-RED recombineering for insertion of SGCs**

28   The λ-RED system also allows insertion of larger synthetic fragments (e.g. genes and even whole SGCs) but this

29   kind of recombineering requires the use of selection markers flanked by 50nt HA or longer.[40] To remove the

30   selective marker, a site-specific recombinase (termed flippase), that requires 'flippase recombination target

31   sites' (FRT) that flank the antibiotic cassette is needed. An interesting way of using the selection marker is

32   through 'chemically induced chromosomal evolution' (CIChE), which explores the principle of tandem

33   duplication of cassettes. The cassettes comprise a SGC and antibiotic marker (5kb) flanked by 1kb homology

34   arms. This needs to be duplicated as a result of adaptation to increased antibiotic concentration, assisted by

35   the λ-RED system, with the whole process being terminated by the deletion of RecA.[16] The forced

36   multiplication of the SGC allows for rapid evolution of the *E. coli* genetic frame improving the production, for

example, of lycopene by 60 % over plasmid borne production levels. Moreover, the strain had stable production levels for more than 80 generations in contrast to the plasmid borne production strain, which was stable for less than 20 generation.[16]

Alternative strategies have allowed insertion of larger SGCs.[41, 42] One strategy has relied on the insertion of landing pads (LP) into specific chromosomal sites via 50nt HA flanked by LP of 50nt random sequences. The LP facilitates recombination site, enabling the insertion of an SGC (up to 7kb) into the previously integrated chromosomal site, while I-SceI endonuclease cleaves in the LP to facilitate recombineering.[41] A different approach is offered by the KIKO (Knock-in/Knock-out) plasmid system, which does not require prior insertion of a LP, but instead requires longer 500nt HA flanked with FRT used to excise the selective marker in an additional step.[42] To date, the largest SGC that has been introduced successfully on the *E. coli* chromosome is 59kb, using recombinase-assisted genome engineering (RAGE).[17] RAGE requires pre-insertion of a LP that contains a *lox* flanked selection marker. This is introduced to the chromosome site by the λ-RED system and is later used to insert the large 59kb SGC via a Cre recombination system derived from the P1 phage.[43] The engineered *E. coli* strain was able to produce 40 % higher titres of ethanol from alginate degradation compared to a comparable plasmid borne system, whilst maintaining the production levels for at least 50 generations.[17] Also, 50kb DNA fragments have been inserted onto the chromosome using only the λ-RED system. This used seven 7kb fragments that had overlapping (minimum 60nt) homology with each other and in this way could be assembled to 50kb inserts in a modular-flexible manner.[44] The λ-RED system is a versatile and the most applicable tool for fast and economical genome engineering in *E. coli*. However, it is potentially limited due to unwanted integrations or genome rearrangements that can happen due to high metabolic burden caused by antibiotic use and spontaneous rearrangements attributed to partial similarities of the SGC and regions of the chromosome.

CRISPR/Cas9 mediated recombineering

The discovery of the bacterial and archaeal viral defence systems called CRISPR (clustered regularly interspaced palindromic repeats)[45] was the basis for developing a new technology for targeted genome engineering, namely CRISPR/Cas9.[46] The technology utilizes the endonuclease activity of Cas9 that is guided by small RNA (crRNA) to its cognate DNA target. DNA cleavage requires not only hybridization between target DNA and crRNA, but also a matching protospacer adjacent motif (PAM e.g. 5'-NGG-3') on the non-targeted DNA strand.[47] The strategy behind the system is to introduce double strand breaks (DSB) with Cas9 in a targeted and programmable fashion. To accomplish the repair of DSB, the λ-RED system needs a DNA template (synthetic oligonucleotides or plasmid borne dsDNA) with HA that are flanking the DSB in the chromosome but does not require any antibiotic marker integration.[48] The first protocols for CRISPR/Cas9 genome editing in *E. coli* were developed using a dual-RNA system, which comprises the guiding crRNA and an antisense tracrRNA needed for association with Cas9. The whole system is based on three components; the helper plasmid that encodes all the genes of enzymes needed for recombineering (Cas9, λ-RED system and tracrRNA), a second plasmid containing a programmed crRNA, and the editing template provided via DNA oligonucleotide.[48] Using

a slightly modified plasmid system up to three target genes can be simultaneously edited. The system is able to integrate up to 4.5kb fragments with efficiency of up to 75 %, accomplished by using 400nt HA and single-guide RNA (sgRNA), which mimics the previously used crRNA-tracrRNA duplex.[49, 50] A variation of this strategy has been used to insert a SGC encoding for the pigment β-carotene, while also altering the genetic frame to increase acetyl-CoA flux availability through the methylerythritol phosphate pathway (MEP) and increasing the production levels of the resulting strain by 3-fold.[51] β-carotene pigment is also a member of the terpene family assembled from isoprene units found in most plants. It possesses a red-orange colour and like lycopene is derived from acetyl-CoA via the MVA pathway and is used as a food colouring agent. Further improvements of the editing strategy were achieved using a linear DNA template with a mutated PAM motif on the 5' site of 100nt long HA to prevent sequential cleavage of the targeted DNA by Cas9 after the initial DSB was repaired by the same template. This enabled a markerless chromosomal integration of a 10kb SGC with 50% efficiency (14kb with 25% efficiency also reported) and a production for isobutanol at levels similar to the plasmid borne production strain.[52] These methods[42, 49] were also recently explored to integrate the melavonate pathway (MVA) and optimize downstream pathways for sucrose utilization (e.g. *cscAKB* operon). Extensive efforts were made for a 5-fold higher production of bisabolene in a genome engineered *E. coli* strain.[53] Bisabolene is a derivative of the MVA pathway which is commonly used as a scent and is a component of many essential oils.[6] This study not only highlights the enormous potential of genome editing, but also points to weaknesses since the 15kb SGC had to be split into 5kb sub-clusters for successful integration. Additional genome editing of the intrinsic regulators (e.g. promoter) inside the SGC was needed because the chromosomally inserted SGC had different expression levels compared to the SGC encoded on the plasmid.[53] A complementary tool in the CRISPR family is the CRISPR optimized multiple automated genome engineering (CRMAGE),[54] which uses the β-subunit of the λ-RED system to introduce 20nt mutations. Here the genotype of a wild type cell is targeted and degraded by the CRISPR/Cas9 system improving the recombineering efficiency to 97 % and allowing genotype rather than phenotype selection.[54] Recently, this tool was used for genome wide excision of mobile genetic elements in *E. coli* to engineer a more stable genetic frame.[55] CRISPR systems are at the centre of a rapidly evolving field with seemingly endless potential for developing tools to manipulate nucleic acids on all levels such as Cpf1[56] or C2c2.[57] These will soon add complementary tools to the existing arsenal of approaches (Table 1) used for genetic engineering in *E. coli*.

Table 1. Overview of most commonly used methods for recombineering.

Method	Size inserted	Components /mode	Homology arms	Product /increase	References
laboratory strain evolution	Not specified	one plasmid /chemostat cultivation	Not utilized	CO_2 utilization	[26]
MAGE	20nt	λ-RED /HTP	35nt	Lycopene /4-fold[35]	[34] [36], [39]
CIChE	5kb	λ-RED /HTP	1kb	Lycopene /60%	[16]
RAGE	59kb	λ-RED & Cre-recombinase /HTP	*lox* site	Ethanol from alginate /40%	[17]
λ-RED high molecular weight [*]	7 x 7kb = 50kb	λ-RED /Single cell	minimum 60nt	Lactose utilization	[44]
KIKO [*]	5.4kb	λ-RED & FRT /Single cell	500nt	Not specified	[42]
Landing pad [*]	7kb	λ-RED & I-SceI & FRT /Single cell	50nt	Not specified	[41]
CRISPR/Cas9 [**]	4.5kb	λ-RED & Cas9 /Single cell	400nt	β-carotene /2.8-fold[52]	[49] [51]
CRISPR/Cas9 PAM-mutation [**]	10kb	λ-RED & Cas9 /Single cell	100nt	Isobutanol /equal level	[52]
CRISPR/Cas9 & KIKO	3x5kb = 15kb	λ-RED & Cas9 /Single cell	500nt	Bisabolene /5-fold	[53]
CRMAGE	20nt	λ-RED /HTP	35nt	Not specified	[54] [55]

1 *Schematically represented in Figure 1; ** Schematically represented in Figure 1

2

3

4 **Conclusions and outlook**

5 Genome engineering can be used to generate *E. coli* strains with a genetic frame that harbours a SGC for the

6 production of high value secondary metabolites. Recent work has demonstrated an ability to engineer genetic

7 frames that utilize abundant and inexpensive carbon sources to produce acetyl-CoA using laboratory strain

8 evolution, MAGE, TRMR or CRMAGE platforms. The λ-RED system is the most routinely used method to

9 introduce SGCs onto the *E. coli* chromosome. However, with this system larger SGCs (up to 8kb) require co-

10 insertion of a selection marker. The use of antibiotic resistance cassettes can be avoided by using a

11 CRISPR/Cas9 system, which introduces a targeted DSB in the chromosome and establishes a selective pressure

12 for *E. coli* to repair the DSB. More specialised methods – for example RAGE – can be used to insert very large

13 fragments (up to 50Kb). Chemically induced chromosomal evolution (CIChE) can also be used to increase the

14 chromosomal SGC copy number to achieve higher production levels. A variety of tools for genetic engineering

15 are therefore available, but these still require exquisite laboratory handling skills for sub-cloning and culturing,

16 detailed investigation of existing tool sets and method optimization. No doubt the future will usher in an

17 explosion in the number of newly identified SGCs[58] as well as new tools for genome engineering.[59] All this

will be timely if we are to reproduce and diversify the natural products provided by nature and manufacture new medicines for societal benefit. The relative ease by which genome editing can now be achieved in *E. coli* should lead to rapid isolation of stable strains for the production of bioactive natural products as well as fine and speciality chemicals, platform chemicals, therapeutics and components for new synthetic biological materials.

References

[1]   J. Nielsen, J. D. Keasling, *Cell* **2016**, 164, 1185.
[2]   R. Simon, U. Priefer, A. Pühler, *Nature biotechnology* **1983**, 1, 784.
[3]   D. Yu, H. M. Ellis, E. C. Lee, N. A. Jenkins, N. G. Copeland, D. L. Court, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 5978.
[4]   A. Schirmer, M. A. Rude, X. Li, E. Popova, S. B. Del Cardayre, *Science* **2010**, 329, 559.
[5]   E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. Del Cardayre, J. D. Keasling, *Nature* **2010**, 463, 559.
[6]   K. H. C. Baser, G. Buchbauer, *Handbook of essential oils: science, technology, and applications*, CRC Press, **2015**.
[7]   V. Bogdanovic, J. Mrdjanovic, I. Borisev, *J. Altern. Complement. Med.* **2017**, 23, 831.
[8]   S. Slezakova, J. Ruda-Kucerova, *Anticancer Res.* **2017**, 37, 5995.
[9]   M. F. Zafeer, F. Firdaus, F. Ahmad, R. Ullah, E. Anis, M. Waseem, A. Ali, M. M. Hossain, *Int. J. Biol. Macromol.* **2018**, 109, 1029.
[10]   R. Subramani, M. Narayanasamy, K. D. Feussner, 3 Biotech **2017**, 7, 172.
[11]   G. Dugo, I. S. d'Alcontres, A. Cotroneo, P. Dugo, *Journal of Essential Oil Research* **1992**, 4, 589.
[12]   A. J. Challinor, J. Watson, D. Lobell, S. Howden, D. Smith, N. Chhetri, *Nature Climate Change* **2014**, 4, 287.
[13]   K. Geoghegan, P. Evans, *Tetrahedron Lett.* **2014**, 55, 1431.
[14]   C. W. Marshall, E. V. LaBelle, H. D. May, *Curr. Opin. Biotechnol.* **2013**, 24, 391.
[15]   A. M. Feist, B. Ø. Palsson, *Nat. Biotechnol.* **2008**, 26, 659.
[16]   K. E. Tyo, P. K. Ajikumar, G. Stephanopoulos, *Nat. Biotechnol.* **2009**, 27, 760.
[17]   C. N. Santos, D. D. Regitsky, Y. Yoshikuni, *Nat. Commun.* **2013**, 4, 2503.
[18]   N. G. Leferink, A. J. Jervis, Z. Zebec, H. S. Toogood, S. Hay, E. Takano, N. S. Scrutton, *ChemistrySelect* **2016**, 1, 1893.
[19]   G. Le Page, L. Gunnarsson, J. Snape, C. R. Tyler, *Environ. Int.* **2017**.
[20]   C. C. Geddes, M. T. Mullinnix, I. U. Nieves, J. J. Peterson, R. W. Hoffman, S. W. York, L. P. Yomano, E. N. Miller, K. T. Shanmugam, L. O. Ingram, *Bioresour. Technol.* **2011**, 102, 2702.
[21]   M. Saini, Z. W. Wang, C. J. Chiang, Y. P. Chao, *Biotechnol. Biofuels* **2017**, 10, 173.
[22]   K. Martinez-Gomez, N. Flores, H. M. Castaneda, G. Martinez-Batallar, G. Hernandez-Chavez, O. T. Ramirez, G. Gosset, S. Encarnacion, F. Bolivar, *Microb. Cell Fact.* **2012**, 11, 46.
[23]   M. Dragosits, D. Mattanovich, *Microb. Cell Fact.* **2013**, 12, 64.
[24]   J. Pang, Z.-Y. Liu, M. Hao, Y.-F. Zhang, Q.-S. Qi, *Biotechnol. Biofuels* **2017**, 10, 165.

[25]   A. Arevalo-Gallegos, Z. Ahmad, M. Asgher, R. Parra-Saldivar, H. M. N. Iqbal, *Int. J. Biol. Macromol.* **2017**, 99, 308.

[26]   N. Antonovsky, S. Gleizer, E. Noor, Y. Zohar, E. Herz, U. Barenholz, L. Zelcbuch, S. Amram, A. Wides, N. Tepper, D. Davidi, Y. Bar-On, T. Bareia, D. G. Wernick, I. Shani, S. Malitsky, G. Jona, A. Bar-Even, R. Milo, *Cell* **2016**, 166, 115.

[27]   E. Herz, N. Antonovsky, Y. Bar-On, D. Davidi, S. Gleizer, N. Prywes, L. Noda-Garcia, K. Lyn Frisch, Y. Zohar, D. G. Wernick, A. Savidor, U. Barenholz, R. Milo, *Nat. Commun.* **2017**, 8, 1705.

[28]   C. P. Long, J. E. Gonzalez, A. M. Feist, B. O. Palsson, M. R. Antoniewicz, *Metab. Eng.* **2017**, 44, 100; C. P. Long, J. E. Gonzalez, A. M. Feist, B. O. Palsson, M. R. Antoniewicz, *Proc. Natl. Acad. Sci. USA* **2018**, 115, 222.

[29]   A. Rodriguez, J. A. Martinez, N. Flores, A. Escalante, G. Gosset, F. Bolivar, *Microb. Cell Fact.* **2014**, 13, 126.

[30]   C. M. Radding, *J. Mol. Biol.* **1970**, 52, 491.

[31]   J. A. Mosberg, M. J. Lajoie, G. M. Church, *Genetics* **2010**, 186, 791.

[32]   M. G. Thompson, N. Sedaghatian, J. F. Barajas, M. Wehrs, C. B. Bailey, N. Kaplan, N. J. Hillson, A. Mukhopadhyay, J. D. Keasling, *Sci. Rep.* **2018**, 8, 1590.

[33]   H. M. Ellis, D. Yu, T. DiTizio, D. L. Court, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 6742.

[34]   H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest, G. M. Church, *Nature* **2009**, 460, 894.

[35]   J. R. Warner, P. J. Reeder, A. Karimpour-Fard, L. B. A. Woodruff, R. T. Gill, *Nat. Biotechnol.* **2010**, 28, 856.

[36]   F. J. Isaacs, P. A. Carr, H. H. Wang, M. J. Lajoie, B. Sterling, L. Kraal, A. C. Tolonen, T. A. Gianoulis, D. B. Goodman, N. B. Reppas, C. J. Emig, D. Bang, S. J. Hwang, M. C. Jewett, J. M. Jacobson, G. M. Church, *Science* **2011**, 333, 348.

[37]   M. J. Lajoie, C. Gregg, J. Mosberg, G. Washington, G. Church, *Nucleic Acids Res.* **2012**, 40, e170.

[38]   A. Nyerges, B. Csorgo, I. Nagy, B. Balint, P. Bihari, V. Lazar, G. Apjok, K. Umenhoffer, B. Bogos, G. Posfai, C. Pal, *Proc. Natl. Acad. Sci. USA* **2016**, 113, 2502.

[39]   G. Kuznetsov, D. B. Goodman, G. T. Filsinger, M. Landon, N. Rohland, J. Aach, M. J. Lajoie, G. M. Church, *Genome Biol.* **2017**, 18, 100.

[40]   S. K. Sharan, L. C. Thomason, S. G. Kuznetsov, D. L. Court, *Nat. Protoc.* **2009**, 4, 206.

[41]   T. E. Kuhlman, E. C. Cox, *Nucleic Acids Res.* **2010**, 38, e92.

[42]   S. Sabri, J. A. Steen, M. Bongers, L. K. Nielsen, C. E. Vickers, *Microb. Cell Fact.* **2013**, 12, 60.

[43]   G. D. van Duyne, in *Mobile DNA III*, *American Society of Microbiology*, **2015**, 119.

[44]   M. Juhas, J. W. Ajioka, *Microb. Cell Fact.* **2016**, 15, 172.

[45]   E. V. Koonin, K. S. Makarova, Y. I. Wolf, *Annu. Rev. Microbiol.* **2017**, 71, 233.

[46]   M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, 337, 816.

[47]   C. Anders, O. Niewoehner, A. Duerst, M. Jinek, *Nature* **2014**, 513, 569.

[48]   W. Jiang, D. Bikard, D. Cox, F. Zhang, L. A. Marraffini, *Nat. Biotechnol.* **2013**, 31, 233.

[49]   Y. Jiang, B. Chen, C. Duan, B. Sun, J. Yang, S. Yang, *Appl. Environ. Microbiol.* **2015**, 81, 2506.

[50]   M. E. Pyne, M. Moo-Young, D. A. Chung, C. P. Chou, *Appl. Environ. Microbiol.* **2015**, 81, 5103.

[51]   Y. Li, Z. Lin, C. Huang, Y. Zhang, Z. Wang, Y. J. Tang, T. Chen, X. Zhao, *Metab. Eng.* **2015**, 31, 13.

[52]   M. C. Bassalo, A. D. Garst, A. L. Halweg-Edwards, W. C. Grau, D. W. Domaille, V. K. Mutalik, A. P. Arkin, R. T. Gill, *ACS Synth. Biol.* **2016**, 5, 561.

[53]   J. Alonso-Gutierrez, D. Koma, Q. Hu, Y. Yang, L. J. G. Chan, C. J. Petzold, P. D. Adams, C. E. Vickers, L. K. Nielsen, J. D. Keasling, T. S. Lee, *Biotechnol. Bioeng.* **2017**.

[54]   C. Ronda, L. E. Pedersen, M. O. Sommer, A. T. Nielsen, *Sci. Rep.* **2016**, 6, 19452.

[55]   K. Umenhoffer, G. Draskovits, A. Nyerges, I. Karcagi, B. Bogos, E. Timar, B. Csorgo, R. Herczeg, I. Nagy, T. Feher, C. Pal, G. Posfai, *ACS Synth. Biol.* **2017**, 6, 1471.

[56]   B. Zetsche, M. Heidenreich, P. Mohanraju, I. Fedorova, J. Kneppers, E. M. DeGennaro, N. Winblad, S. R. Choudhury, O. O. Abudayyeh, J. S. Gootenberg, W. Y. Wu, D. A. Scott, K. Severinov, J. van der Oost, F. Zhang, *Nat. Biotechnol.* **2016**, 35, 31.

[57]   O. O. Abudayyeh, J. S. Gootenberg, S. Konermann, J. Joung, I. M. Slaymaker, D. B. T. Cox, S. Shmakov, K. S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E. S. Lander, E. V. Koonin, F. Zhang, *Science* **2016**, 353.

[58]   S. Mukherjee, R. Seshadri, N. J. Varghese, E. A. Eloe-Fadrosh, J. P. Meier-Kolthoff, M. Goker, R. C. Coates, M. Hadjithomas, G. A. Pavlopoulos, D. Paez-Espino, Y. Yoshikuni, A. Visel, W. B. Whitman, G. M. Garrity, J. A. Eisen, P. Hugenholtz, A. Pati, N. N. Ivanova, T. Woyke, H. P. Klenk, N. C. Kyrpides, *Nat. Biotechnol.* **2017**, 35, 676.

[59]   S. Doron, S. Melamed, G. Ofir, A. Leavitt, A. Lopatina, M. Keren, G. Amitai, R. Sorek, *Science* **2018**.