

# A Pressure Test to Make 10 Molecules in 90 Days: External Evaluation of Methods to Engineer Biology

Arturo Casini,<sup>†,‡,□</sup> Fang-Yuan Chang,<sup>†,§,□</sup> Raissa Eluere,<sup>†,‡,□</sup> Andrew M. King,<sup>†,§,□</sup> Eric M. Young,<sup>†,§,□</sup> Quentin M. Dudley,<sup>†,||</sup> Ashty Karim,<sup>†,||</sup> Katelin Pratt,<sup>†,‡</sup> Cassandra Bristol,<sup>†,‡</sup> Anthony Forget,<sup>†,‡,§</sup> Amar Ghodasara,<sup>§</sup> Robert Warden-Rothman,<sup>†,§</sup> Rui Gan,<sup>†,||</sup> Alexander Cristofaro,<sup>†,§</sup> Amin Espah Borujeni,<sup>†,§</sup> Min-Hyung Ryu,<sup>§</sup> Jian Li,<sup>||</sup> Yong-Chan Kwon,<sup>||</sup> He Wang,<sup>||</sup> Evangelos Tatsis,<sup>⊥</sup> Carlos Rodriguez-Lopez,<sup>⊥</sup> Sarah O'Connor,<sup>⊥</sup> Marnix H. Medema,<sup>#</sup> Michael A. Fischbach,<sup>†,||</sup> Michael C. Jewett,<sup>†,||</sup> Christopher Voigt,<sup>\*,†,‡,§,||</sup> and D. Benjamin Gordon<sup>\*,†,‡,§</sup>

<sup>†</sup>The Foundry, 75 Ames Street, Cambridge, Massachusetts 02142, United States

<sup>‡</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States

<sup>§</sup>Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States,

<sup>||</sup>Department of Chemical and Biological Engineering, Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, United States

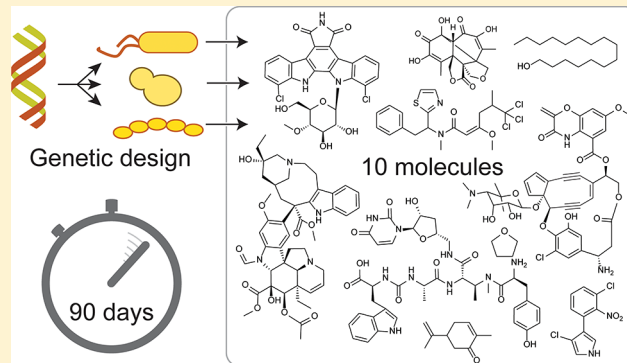
<sup>⊥</sup>Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, United Kingdom

<sup>#</sup>Bioinformatics Group, Wageningen University, Wageningen 6708 PB, The Netherlands

<sup>||</sup>Department of Bioengineering and Chemistry, Engineering & Medicine for Human Health, Stanford University, Stanford, California 94305, United States

## Supporting Information

**ABSTRACT:** Centralized facilities for genetic engineering, or “biofoundries”, offer the potential to design organisms to address emerging needs in medicine, agriculture, industry, and defense. The field has seen rapid advances in technology, but it is difficult to gauge current capabilities or identify gaps across projects. To this end, our foundry was assessed via a timed “pressure test”, in which 3 months were given to build organisms to produce 10 molecules unknown to us in advance. By applying a diversity of new approaches, we produced the desired molecule or a closely related one for six out of 10 targets during the performance period and made advances toward production of the others as well. Specifically, we increased the titers of 1-hexadecanol, pyrolnitrin, and pacidamycin D, found novel routes to the enediyne warhead underlying powerful antimicrobials, established a cell-free system for monoterpene production, produced an intermediate toward vincristine biosynthesis, and encoded 7802 individually retrievable pathways to 540 bisindoles in a DNA pool. Pathways to tetrahydrofuran and barbamide were designed and constructed, but toxicity or analytical tools inhibited further progress. In sum, we constructed 1.2 Mb DNA, built 215 strains spanning five species (*Saccharomyces cerevisiae*, *Escherichia coli*, *Streptomyces albidoflavus*, *Streptomyces coelicolor*, and *Streptomyces albobinaceus*), established two cell-free systems, and performed 690 assays developed in-house for the molecules.



## INTRODUCTION

Periodically, critical-needs crises threaten infrastructure, manufacturing, defense, and human health. Emerging disease prompts the need for rapid routes to complex pharmaceuticals,<sup>1–4</sup> and the U.S. FDA maintains a roster of drugs whose supply will soon fall short of demand.<sup>5</sup> In 2003, shortages of para-aramid fibers for Kevlar jeopardized the supply of protective vests for American soldiers, leading the U.S. Defense Logistics Agency to buy out entire production lines and identify alternative sources.<sup>6</sup> During 1999–2000, the semiconductor industry was threatened by

consecutive explosions at two of the world’s three hydroxylamine plants.<sup>7</sup> The same year, pharmaceutical production was inhibited by acetonitrile shortages.<sup>8</sup> As the bioeconomy emerges,<sup>9,10</sup> it will be increasingly considered as a potential avenue to address such needs. There was an early incident in 2008 where bioderived propanediol was used as an alternative deicer when a worldwide potassium acetate shortage threatened to shut down commercial

Received: December 19, 2017

Published: February 26, 2018

and Air Force runways.<sup>11,12</sup> Sourcing a chemical from an existing industrial bioprocess is relatively straightforward, but what if the need requires building a new organism? Historically, the costs and development time in the biotechnology industry would be too prohibitive to even consider. However, they are dropping rapidly, and biology will soon be at the table when such crises hit.

Centralized genetic engineering facilities consolidate foundational technology around the process of designing and constructing an organism. These facilities integrate genetic tools, software, automation, and manufacturing processes to streamline the engineering of biological systems. Globally, roughly a dozen such facilities have been established in government, academic, and commercial institutions,<sup>13–25</sup> and each emphasizes different foundational technologies, including computer-aided design and artificial intelligence, DNA synthesis, genome construction, automated strain development, and high-throughput screening. Some specialize in particular organisms (e.g., synthesis of complete *S. cerevisiae* chromosomes), core technologies (e.g., protein engineering), or application spaces (e.g., bioenergy).

During periods of rapid technology development, it can be challenging to compare disparate approaches, particularly when failures are unpublished and work at some facilities is proprietary. To this end, third-party time-limited assessments can provide valuable perspectives on technology readiness and reveal bottlenecks. In biology, a preeminent example is CASP (Critical Assessment of Protein Structure Prediction), where amino acid sequences are provided to computational groups prior to the release of 3D structural data, and they must submit predictions within three months.<sup>26</sup> Over the last 20 years, CASP has been critical in guiding future technology development. Similar assessments exist for protein–protein interaction prediction and gene annotation.<sup>27,28</sup> Beyond biology, third-party technical evaluations are common, including aeronautics,<sup>29</sup> robotics,<sup>30</sup> and finance.<sup>31</sup>

To assess our capabilities, we were subjected to a pressure test administered by the U.S. Defense Advanced Research Projects Agency (DARPA) from August 11 to November 11, 2016. We were provided the names of 10 target molecules and allowed three months to research, design, and develop strains to make as many of the molecules as possible. Neither the Foundry nor any of our academic partners had done prior work with these molecules, and we would be not told in advance of their identity or even that they would be small molecules. Contractually, the design challenges could have been anything from materials to living cells that can sense and respond. We were also told neither when the test would start nor its length. The end of the test had a hard stop, at which time we had to report on our results, which formed the basis for this article. For the purposes of scientific publication, some additional assays were repeated after this period to obtain error bars or for additional chemical confirmation of the product (noted in the figure captions and [Methods](#)).

The molecule list spanned simple chemicals that had already been produced by recombinant organisms (e.g., 1-hexadecanol), complex natural products from plants for which there is no enzyme information (e.g., epicolactone), and chemicals with no known biological route (e.g., tetrahydrofuran). [Table 1](#) provides the full list, our and DARPA's initial estimates of difficulty, the state-of-the-art to date, and a summary of our results. Upon receiving the list, we launched into a design phase. Some of this effort was low-tech, including searching the literature to define strategies, followed by a period that resembled a global “scavenger hunt” to identify laboratories with expertise in the molecules, who we approached either for materials or for guidance (occasionally

preceded by laughter, since some of these targets had been worked on for decades). In some cases, assistance was hindered by unwillingness or inability to share materials due to intellectual property. Once a strategy was set, a suite of computer-aided design tools were used to design DNA libraries for each strain. Surprisingly, DNA synthesis or acquiring strains often required the bulk of the time. On average, only 4 weeks remained to construct DNA libraries via automated assembly, transform these into diverse species, and screen and confirm the results ([Table 1](#)). Assay development and extraction/chemical confirmation of the products, particularly the complex natural products, were often limiting in the remaining time.

The article is organized by the different approaches taken. First, we describe targets where the key challenge was enzyme identification. This included mining enzymes for a single step to increase titer (1-hexadecanol) and to fill in a seven-enzyme gap in a large biosynthetic pathway (vincristine). Second, when a natural route could not be taken, a retrosynthetic approach was implemented, where enzymes were artificially combined to build a desired target molecule from a central metabolite. This was used to design a six-enzyme pathway from acetyl-CoA to epicolactone and a five-enzyme pathway from tryptophan to rebeccamycin aglycon (and 540 other bisindoles). Third, methods to rewire or rebuild regulatory networks were applied when a pathway was known but not active or optimal under laboratory conditions. This spanned complete refactoring of the pathway to eliminate native regulation (pyrrolnitrin), the use of inserted T7 RNAP promoters to attempt to “wake up” transcription (barbamide), and the overexpression of an activator to facilitate host transfer (pacadimicin D). Fourth, we were unable to obtain the published organism that produces C-1027, so we identified and confirmed two new producing strains using bioinformatics and then cloned and transferred the pathway. Finally, we developed cell-free systems when the product or key intermediates were expected to be toxic. This was used to demonstrate the production of limonene from mevalonate (toward carvone and other monoterpenoids) and was attempted for tetrahydrofuran (THF) (enzyme expression demonstrated, but no product observed). Collectively, the range of approaches applied demonstrates the broad technology needs required when operating under time constraints.

## ■ MINING ENZYMES AND BALANCING THEIR EXPRESSION

For 1-hexadecanol and vincristine, the core challenge was enzyme identification. Finding genes has been facilitated by “part mining”, where bioinformatics tools are used to identify enzymes from the database, their genes are codon-optimized and synthesized, and then the library of enzymes is screened.<sup>32,33</sup> Improvements can arise from better kinetics, fewer side products, disrupted feedback inhibition, or superior expression and folding. The expression level of the enzyme has to be optimized as well. This gets more complicated for pathways where the balance between enzyme levels can be critical. Finding the right balance requires the assembly of libraries of pathways where different genetic parts (e.g., promoters, ribosome binding sites, terminators) are used to control each gene, followed by screening of all variants.<sup>34–36</sup> There are various computational tools to guide this process, from design of experiments to mechanistic metabolic models.<sup>37–40</sup>

**1-Hexadecanol.** Fatty alcohols are used as fuels, pharmaceutical emulsifiers, lubricants, detergents, surfactants, and paint additives and are derived from palm oil and petroleum.<sup>41</sup>

Table 1. Pressure Test Molecules

Target <sup>a</sup>	Structure	Applications	Predicted Difficulty <sup>b</sup>		Prior State of Art	Outcome		Time (Days) <sup>f</sup>		
			DARPA	Foundry		Titer	Organism	Design	Materials <sup>g</sup>	Build & Test
1-hexadecanol		Industrial lubricant, detergent, surfactant	1	2	160 mg/L in <i>S. cerevisiae</i> <sup>c,44</sup>	238 mg/L	<i>S. cerevisiae</i>	30	60	90
Pyrrolnitrin		Human topical antifungal	2	1	Produced in <i>Pseudomonas</i> <sup>103</sup> Titers not reported.	11 mg/L	<i>E. coli</i>	30	60	90
Rebeccamycin		Antitumor agent	3	4	Produced in <i>E. coli</i> <sup>75</sup> Titers not reported.	0.7 mg/L rebeccamycin aglycon <sup>e</sup>	<i>E. coli</i>	30	60	90
Carvone		Mosquito repellent, food storage pesticide	4	3	None	23 mg/L limonene intermediate	Cell-free <i>E. coli</i>	30	60	90
Pacidamycin D		Antibiotic against wound pathogens	5	5	Produced in <i>S. lividans</i> <sup>118</sup> Titers not reported.	10-50 mg/L <sup>e</sup>	<i>S. albidoflavus</i> J1074	30	60	90
Barbamide		Marine anti-fouling agent	6	7	None	None	<i>E. coli</i>	30	60	90
Vincristine		Potent chemotherapeutic, toxic	7	10	Extracted from Madagascar periwinkle <sup>182</sup>	None; delivered 1.5Mb DNA library	<i>S. cerevisiae</i>	30	60	90
C-1027		Cytotoxic enediyne antibiotic	8	6	56 mg/L in <i>Streptomyces</i> sp. CB02366 (chromophore) <sup>d,183</sup>	2 mg/L	<i>S. albobovineus</i> ATCC 33031	30	60	90
Epicolactone		Antimicrobial and antifungal	9	8	None	None; designed retrosynthetic pathway	<i>S. cerevisiae</i>	30	60	90
THF		Solvent, adhesive, metal degreaser	10	9	None	None; confirmed in-vitro enzyme expression	Cell-free <i>E. coli</i>	30	60	90

<sup>a</sup>Targets previously produced by recombinant organisms: 1-hexadecanol, pyrrolnitrin, rebeccamycin, and pacidamycin D. Known pathways or enzymes, but not previously produced recombinantly: carvone, barbamide, and C-1027. Targets with limited or no enzyme information: vincristine and epicolactone. Targets with no known biological route: THF. <sup>b</sup>In the first 2 days of the test, chemical targets were ranked in terms of estimated difficulty, based on estimated complexity of the pathway and on availability of prior evidence of production as noted above. DARPA and the Foundry (our group) ranked targets independently. <sup>c</sup>Titer of strain from the Zhao lab (U Illinois) when measured under equivalent growth conditions (Methods). The Zhao lab reports a titer of 1.1 g/L under fermentation conditions.<sup>44</sup> <sup>d</sup>Chromophore yield is estimated from the reported chromoprotein titer of 750 mg/L (Methods). <sup>e</sup>Titers for rebeccamycin aglycon and pacidamycin D are estimates. Yields were approximated from indirect evidence (Methods). <sup>f</sup>To accommodate delays in third-party gene synthesis, work was extended for 2 weeks for rebeccamycin and THF. <sup>g</sup>Time required to obtain strain from outside source or DNA from a vendor.

Biosynthetic routes to various fatty alcohols have been reported.<sup>42</sup> 1-Hexadecanol is used as a fastener lubricant in the Army, Navy, and Air Force.<sup>43</sup> The highest reported titer for 1-hexadecanol was achieved in *S. cerevisiae* by expressing a barn owl fatty acyl-CoA reductase (FAR) in a strain whose metabolism had been engineered to increase the availability of malonyl-CoA acid and acetyl-CoA precursors by adding three genes (ACC1, ACL1, ACL2) and deleting one (RPD3).<sup>44,45</sup> This strain was obtained from the Zhao lab (U Illinois) and used as a starting point.

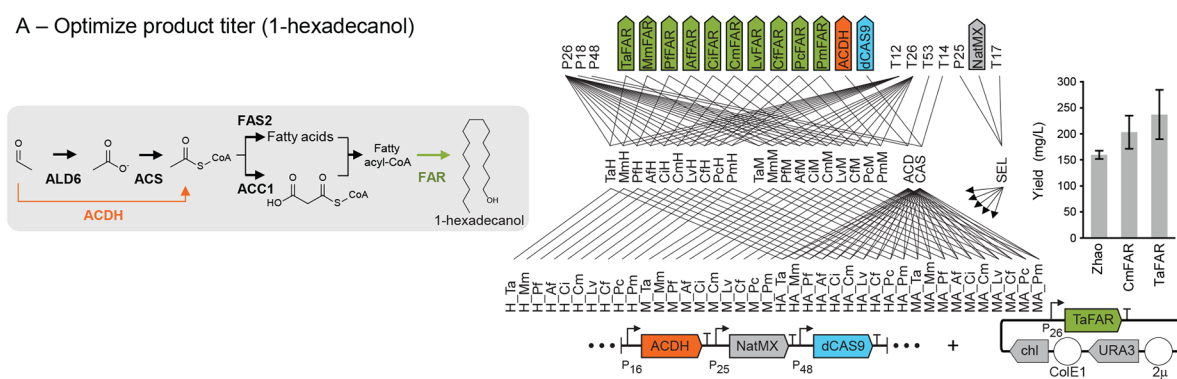
We hypothesized that better titers could be achieved by mining alternative FAR enzymes and increasing the precursor acetyl-CoA. Candidate FAR genes were identified from the NCBI sequence database (Methods), and eight new FARs at varying evolutionary distances were selected. The set included genes from birds (peregrine falcon, crested ibis, and emperor penguin), mammals (Bactrian camel, Yangtze river dolphin, and sperm whale), and reptiles (sea turtle and pit viper), all of which were codon optimized for yeast expression. To improve precursor availability, we simultaneously tested the acetylating acetaldehyde dehydrogenase (ACDH) enzyme EutE from *Listeria innocua*,<sup>46</sup> which produces acetyl-CoA from acetaldehyde with lower ATP cost than native metabolism. Preserving the RPD3 deletion identified by Zhao and co-workers,<sup>44</sup> a full-factorial design-of-experiments library was designed to simultaneously screen for the optimal FAR enzyme, FAR expression level, and impact of the ACDH enzyme (Figure 1a). FAR variants

were assembled onto individual *E. coli*-yeast 2  $\mu$  shuttle vectors (pY124) within cassettes encoding expression at different strengths, which were achieved by using combinations of constitutive promoters and terminators developed in-house. The plasmids were then tested in strains containing ACDH integrated into chromosome XV, as well as strains without it. We successfully constructed 60 strains comprising a total of 621kb of sequence-verified heterologous DNA and tested 28 of them. We found that when combined with ACDH and the RPD3 deletion, two FARs enabled production at levels higher than the published strain (160 mg/L under our growth conditions): the published FAR from barn owl FAR, which was 238 mg/L, and a new FAR from sea turtle, which was 204 mg/L (Figure 1a).

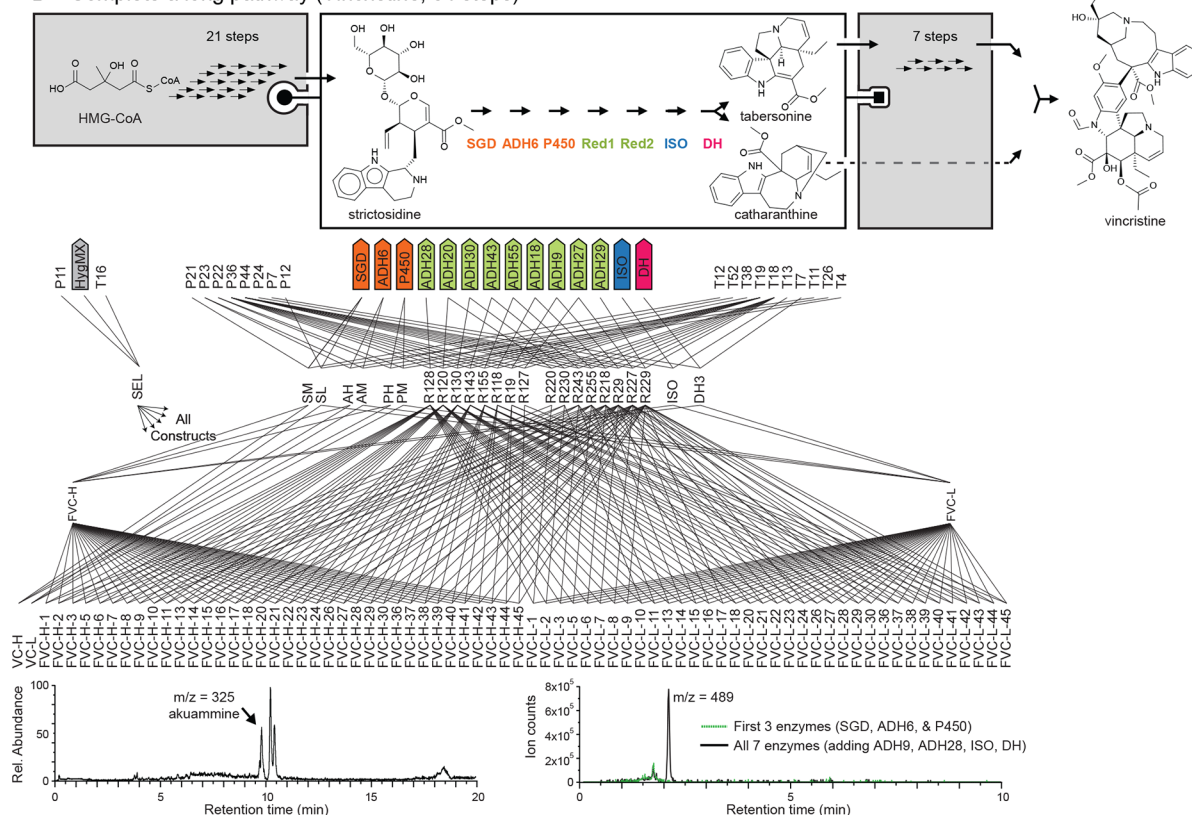
**Vincristine.** Monoterpene indole alkaloids (MIAs) encompass a large class of plant-derived natural products with strong physiological activity, including quinine, many of which have medical applications.<sup>47,48</sup> Vincristine is a potent anticancer chemotherapeutic that interferes with cell division and has toxic side effects including neuropathic pain and low white blood cells. It is derived from the Madagascar periwinkle, one ton of which is required to produce one ounce (costing \$140 000), thus motivating researchers to find an alternative route. The complete pathway from HMG-CoA to vincristine is thought to require over 35 enzymes<sup>49</sup> (Figure 1b). Two research laboratories previously constructed different portions of the pathway, but they cannot be connected because of a multienzyme gap.



## A – Optimize product titer (1-hexadecanol)



## B – Complete a long pathway (Vincristine, 34 steps)



**Figure 1.** Mining enzymes (1-hexadecanol and vincristine). (A) 1-Hexadecanol production in yeast. Left: The designed pathway for 1-hexadecanol production in *S. cerevisiae* introduces ACDH and FAR enzymes into a  $\Delta$ RPD3 background. Center: Assembly of a library of 1-hexadecanol strains. Individual genetic parts indicated at top include genes (FAR variants, ACDH, dCAS9, and NatMX as a selection marker), promoters (“P”), and terminators (“T”), for which sequences are provided in [Supplementary Table 1](#). Lines indicate how parts are combined into different constructs. Intermediate constructs are identified by the abbreviation of the FAR gene they contain and whether flanking promoters and terminators favor medium (“M”) or high (“H”) gene expression levels. ACD, CAS, and SEL refer to intermediate constructs harboring ACDH, dCAS9, and the selection marker. dCAS9 was included to facilitate future optimization, but was not utilized in this work. Completed strains at the bottom are annotated with an additional “A” if ACDH and dCAS9 are included in the design. Right: Titters of the best performing strains from this work compared to a reference strain provided by Zhao and co-workers<sup>44</sup> grown under the same conditions. Best performers contained CmFAR (sea turtle) or TaFAR (barn owl) expressed at high levels in strains also expressing ACDH. Experiments were repeated in triplicate after the testing period. Bottom: Constructs in best performing strain, comprising three genes integrated in the genome and TaFAR with a strong constitutive promoter on a 2  $\mu$  plasmid. (B) Proposed pathway to bridge between previously reported pathways<sup>50,51</sup> for partial biosynthesis for vincristine. The proposed pathway converts strictosidine (left) to tabersonine and catharanthine, which are further modified and combined to produce vincristine via known steps.<sup>51</sup> Small arrows indicate individual enzymatic steps previously reported. Middle: Assembly of a library of pathway variants for the bridge. Gene, promoter (“P”), and terminator (“T”) identities and sequences are provided in [Supplementary Table 8](#). Intermediate constructs are represented according to the designed expression level of each gene as determined by flanking promoters and terminators (“H,” “M,” and “L” for high, medium, and low). Complete constructs at the bottom are assigned unique design identifiers. Pairs of reductase variants (genes beginning “ADH”) were introduced on plasmids, and all other genes were integrated into the genome. Bottom: Mass spectrometry results from strictosidine feeding assays performed after the evaluation period that show production of the akuammine intermediate ([Supplementary Figure 9e](#)) and other unidentified products as detected by MRM triple-quadrupole (left). IT-ToF comparison of products from feeding assays on a partial pathway (3 enzymes) and a complete pathway (7 enzymes) reveal an unidentified product ( $m/z = 489$ ) produced only by the complete pathway (right).

The O'Connor group (John Innes Center) built a 21-gene pathway in *S. cerevisiae* from HMG-CoA to strictosidine.<sup>50</sup> Separately, the De Luca group built the seven-gene pathway from tabersonine to vindoline,<sup>51</sup> which is two steps from vincristine. Little is known regarding the enzymes that convert strictosidine to tabersonine, which would be required to complete the pathway (Figure 1b, top). We collaborated with the O'Connor lab (John Innes Center) to design and construct a library to complete this pathway.

To fill the pathway gap, a library of pathway variants of three to six genes each was assembled, which tested different combinations of candidate genes identified based on recent, coexpression analysis of *Catharanthus roseus*.<sup>52</sup> The transcriptomics data identified an alcohol dehydrogenase (geissoschizine synthase) and a cytochrome P450 (geissoschizine oxidase) that are involved in the biosynthesis of preakumacine, a key intermediate. We rationalized that conversion of preakumacine to catharanthine and tabersonine proceeds via one or more reduction steps to stemmadenine, followed by dehydration and cyclization. Gene candidates were therefore selected based on having similar expression profiles and homology to genes that might be expected to carry out these predicted reactions. Because so little was known about reductases, we included nine candidates. All genes were codon-optimized for yeast expression, synthesized, and assembled into a library of yeast strains employing them in different combinations and at different levels of expression. Genes were integrated into the genome to produce four parent strains: two strains with all non-reductase genes (high and low expression levels) and two strains with only the first three genes of the pathway (also high and low). The reductases were cloned into plasmids, enabling us to access 36 unique reductase pairs in any parent strain via transformation. In total 74 strains were constructed, comprising 1.5 Mb of heterologous DNA in 2 weeks. The pathways were delivered to the JIC because of the complexity in screening. In feeding experiments performed at JIC after the evaluation period, production of akuammine, a deformed degradation product of an MIA biosynthetic intermediate, was observed ( $m/z = 325$ ). Unidentified compounds were also observed in strains that were not present in the parent strains, for example at  $m/z = 489$ , which may represent MIA biosynthetic intermediates or derivatized intermediates (Figure 1b).

## ■ RETROSYNTHETIC DESIGN

When a natural route to a target molecule is not available, enzymes from unrelated pathways have to be combined to build it stepwise from a starting metabolite.<sup>53,54</sup> This is referred to as "retrosynthesis," a term borrowed from the organic chemistry community, where complex molecules are built from simple chemical precursors.<sup>55</sup> Retrosynthetic approaches that use enzymes as opposed to chemical transformations constitute a new field, and although computational algorithms have been developed, access to this software is limited.<sup>56,57</sup> Retrosynthetic design was required to identify a pathway to epicolactone because the producing organism and genome sequence were unavailable. A similar approach had to be taken for THF (described in a later section) because a natural biological route to this molecule has not been described. Retrosynthesis was also applied to enable a user to design a bisindole structure and then be able to rapidly retrieve and assemble the enzymes required.

**Epicolactone.** The fungus *Epicoccum nigrum* is used by the Brazilian sugar cane industry to promote root growth and inhibit pathogens.<sup>58</sup> It is a member of the fascinating family of "black yeasts", well studied for their tolerance of extreme stress

conditions, such the high radiation environments of nuclear reactors and outer space.<sup>59,60</sup> *E. nigrum* produces the multicyclic tropolone epicolactone, which has antimicrobial and antifungal activity.<sup>61</sup> Enzyme mining was inhibited because the genome sequence<sup>62</sup> was not publicly available at the time of the pressure test, and we were unable to obtain a prepublication draft, as the authors did not respond to a request. Moreover, due to epicolactone's unusual chemical structure, we could not find related pathways by searching biochemical databases for similar molecules.<sup>63</sup>

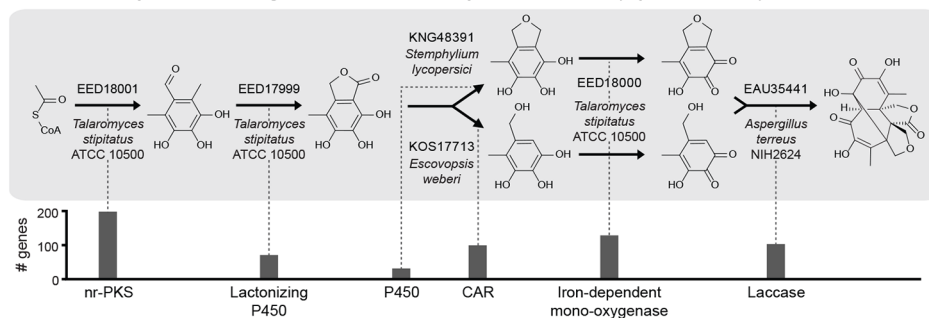
Therefore, we adopted a retrosynthetic approach in which we computationally designed a complete pathway for the production of epicolactone. An eight-step chemical synthesis route had been previously developed,<sup>64</sup> and this was used as a guide to identify enzymes capable of each of these conversions. The assignment of an enzyme class to each step was done manually, guided by a literature search as well as pathway databases and tools such as MIBiG<sup>65</sup> and antiSMASH<sup>63</sup> (Methods). The steps were as follows: (1) A PK-NRP to generate the aryl aldehyde precursor, (2) a P450 for methyl oxidation, (3) a flavin-dependent reductase for conversion of pendant aldehyde to carboxylate, (4) a nonheme-iron-dependent dioxygenase for pendant carboxylate elimination, (5) a flavin-dependent reductase for lactone reduction, (6) a flavin-dependent monooxygenase for vicinal phenol oxidation, and (7) a laccase for monomer coupling, followed by (8) spontaneous ring rearrangements to produce the desired compound.<sup>66–72</sup> Multiple enzymes were identified for each step via a sequence homology search<sup>73</sup> to known members of each enzyme class. To improve the likelihood of success, hits were limited to tropolone-like biosynthetic gene clusters (identified from MIBiG and antiSMASH) (Figure 2a). No pathways could be tested within the project period due to time constraints.

**Rebeccamycin.** The core structure of bisindoles is a tryptophan dimer, the chemical decoration of which leads to diverse pharmaceutical functions, including antibiotics, antivirals, and antitumor agents.<sup>74</sup> Rebeccamycin, produced by *Lechevalieria aerocolonigenes*, is an antitumor agent that inhibits DNA topoisomerase I.<sup>74</sup> Heterologous production of bisindoles has been reported in actinomycetes and *E. coli*.<sup>75,76</sup> We first developed a pathway in *E. coli* to the precursor rebeccamycin aglycon (dichloroarcyriaflavin A) and then expanded this to encompass a larger class of bisindoles. This approach could be generalized to form the basis of a system to rapidly build pathways to chemicals to test against an emerging threat.

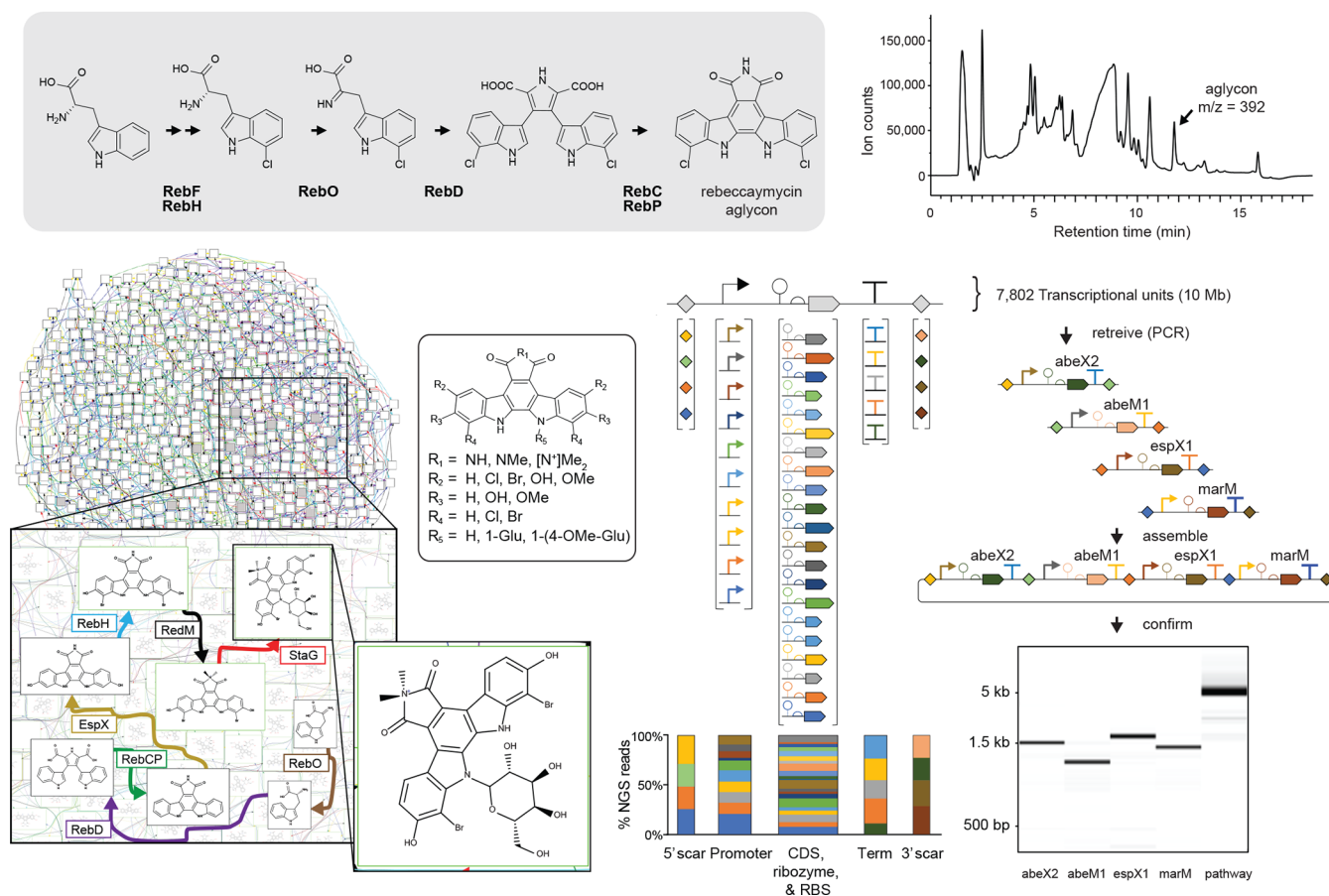
To produce rebeccamycin aglycon, we assembled an *E. coli* strain containing all necessary genes for biosynthesis and export (*rebCDFHOPT*). Genes were codon optimized for heterologous expression in *E. coli*, and individual ribosome binding sites were computed for each gene to maximize expression.<sup>77</sup> Optimized genes were then organized into synthetic operons under the control of T7 promoters and insulated with ribozymes.<sup>78</sup> Production of rebeccamycin aglycon in *E. coli* was confirmed via LCMS (Figure 2b).

We then sought to create a system where a user could access any desired bisindole within a large class of possibilities by retrieving and assembling the associated pathway from a preconstructed pool of uniquely barcoded DNA constructs.<sup>79</sup> Several groups have demonstrated that it is possible to generate up to 32 bisindole derivatives by expressing different combinations of heterologous modifying enzymes.<sup>80–82</sup> However, these efforts only combine two pathways at a time, limiting them to only a fraction of all possible pathways. To establish a systematic, on-demand system, we identified 21 additional genes known to make chemical modifications

## A – Retrosynthetic design - no known biosynthetic route (Epicolactone)



## B – Assembly that encompasses a complete molecule space (Rebeccamycin; 540 bisindoles)



**Figure 2.** Retrosynthetic design (epicolactone and rebeccamycin). (A) Retrosynthetic pathway design for complete biosynthesis of epicolactone. Counts of candidate enzymes found via BLAST search of MIBiG (Supplementary Table 10) are indicated (Methods). (B) Top: Biosynthetic pathway for rebeccamycin aglycon. Genes were codon optimized and expressed in *E. coli*, and the product was detected via LC/MS at  $m/z$  325 and further confirmed based on UV absorbance and chlorination signature in the mass spectrum (Methods). Bottom left: System to access 540 different bisindoles. The cloud represents the collection of in-silico predictions of pathways to all chemical products that can be accessed by employing subsets of the genes in the library, which were selected based on their ability to modify a bisindole core (Methods). A high-resolution version of the cloud diagram is provided in Supplementary Figure 11. Edges indicate operations by specific enzymes, and boxes contain unique chemical products. A pathway to any desired product can be designed by tracing a path from tryptophan. The diversity of the chemical space is also represented as a diagram of chemical substitutions on acryiaflavin-A. Bottom right: Composition of pool of transcriptional units built via one-pot assembly.<sup>184</sup> Each transcriptional unit contains one of the available cloning scars (diamonds), promoters, ribozyme insulators, RBSs, genes, and terminators (part names and sequences are provided in Supplementary Table 6). Below, the relative level of representation of each part in the pool is indicated as determined by next-gen sequencing (Methods). For validation of the assembly scheme, a four-gene pathway (*abeX2*, *abeM1*, *espX1*, *marM*) was assembled after the evaluation period by retrieving unique transcriptional units from the pool by PCR and joining them together via type-IIs assembly (Methods).

to the bisindole core.<sup>80,83–88</sup> We then predicted all chemical products accessible via different selections of these genes and found that the set could access up to 540 unique bisindole products when coexpressed with enzymes to assemble the

core structure (Methods). Of these predicted products, 98% are not annotated in PubChem. Of the known compounds, about half exhibit antitumor properties. The genes were then codon optimized for heterologous expression in *E. coli*, synthesized,



and assembled into a pool of transcriptional units using a one-pot assembly approach<sup>79</sup> which linked each of them to assorted regulatory, cloning, and DNA barcode sequences. This yielded thousands of modular transcriptional units ready for assembly into arbitrary bisindole-modification pathways up to four genes long. By sequencing the pool, 7802 different uniquely barcoded transcriptional units were identified (Methods). The barcodes enable us to retrieve individual constructs from the pool via PCR, which we can then assemble into a pathway designed to produce any bisindole in the set.

## ■ ELIMINATING NATIVE REGULATION

Some targets were associated with a known biosynthetic gene cluster. Clusters are controlled by internal regulation and embedded within the greater global regulatory network of the cell, such that they are active only under defined environmental conditions.<sup>89,90</sup> This can result in suboptimal, or completely silent, expression under laboratory conditions.<sup>91</sup> Regulation can also hinder the transfer of a cluster to a new production host.

The process of “refactoring” seeks to redesign the DNA sequence of a cluster to remove native regulation, replacing it with synthetic parts and control.<sup>92–95</sup> This can be exhaustive, where all the genes are codon-optimized, organized into artificial operons, and placed under the control of synthetic parts. Pyrrolnitrin production in *E. coli* was optimized in this way by refactoring the cluster and deleting a native repressor. Less invasive approaches are often used, for example, placing the operons under the control of an inducible promoter.<sup>96–98</sup> We attempted this with barbamide by inserting T7 RNAP promoters into the cluster, but this failed due to genetic instability. Finally, expression can be activated by deleting an internal repressor or overexpressing an activator.<sup>91</sup> This was applied to the pacidamycin D cluster so that it could be expressed in four *Streptomyces* hosts to identify the optimal producer.

**Pyrrolnitrin.** Phenylpyrroles are chemical derivatives of pyrrolnitrin that are used in agriculture as a prophylactic against pathogenic fungi.<sup>99</sup> It is also used as a human topical antifungal against *Trichophyton*, which causes athlete's foot and jock itch, major problems for combat forces.<sup>100,101</sup> Pyrrolnitrin is produced by a number of *Pseudomonas* species and is the active compound in living biocontrol agents.<sup>102</sup> Successful transfer of the four-gene operon to *E. coli* has been reported,<sup>103</sup> but production levels have not been reported for either the native producer or *E. coli*. We took two approaches to increase pyrrolnitrin production in *E. coli*. First, the biosynthetic gene cluster was refactored so that the genes could be tuned to optimize titer. Second, a native repressor was knocked out to increase the metabolic supply of the tryptophan precursor.

The four native genes for pyrrolnitrin production are encoded on a single operon<sup>104</sup> under the control of a regulated promoter. This promoter responds to different signals across pseudomonads, often requiring a quorum sensing system and the *rpoS* stress response sigma factor.<sup>105–107</sup> The promoter is induced at high cell density during the transition to stationary phase, but turns off at late stages of growth. To achieve production in *E. coli*, we codon-optimized the genes from the PRN operon of *Pseudomonas chlororaphis* for expression in *E. coli* and split them into four individual transcriptional units, each comprising a T7 promoter, an optimized RBS,<sup>77</sup> a ribozyme-based insulator, and a terminator. Transcriptional units were built by gene synthesis and type-II assembly and then assembled into a complete pathway in a second round. The pathway was tested in strains engineered for IPTG-inducible control of a T7 RNA polymerase,

thereby enabling independent control of the pathway expression (Figure 3a).<sup>108</sup> Initially after observing no production, we optimized the growth media and conditions to achieve a baseline production of 3.8 mg/L. By moderating expression levels of the pyrrolnitrin genes (IPTG induction at 0.1 mM rather than 1.0 mM), we increased production to 4.8 mg/mL. Finally, deletion of the TrpR repressor has been shown to boost production in systems requiring high levels of tryptophan synthesis.<sup>109,110</sup> Indeed, we found that moving the refactored cluster to a *trpR* knockout strain increased production to 11 mg/L (Figure 3a).

**Barbamide.** Biofouling on marine vessels is a major problem, costing \$260 M annually for removal and increased fuel costs.<sup>111</sup> One management approach is to incorporate antifouling agents into marine paints.<sup>112</sup> Barbamide is a potent molluscicide extracted from the marine cyanobacterium *Moorea producens* (formerly *Lyngbya majuscula*).<sup>113</sup> Because a related NRPS/PKS compound (lyngbyatoxin) from the same organism had been successfully produced in *E. coli*,<sup>114</sup> we attempted the heterologous transfer of the 26 kb gene cluster<sup>115</sup> to this organism.

To transfer the strain, homologous recombination was used to assemble amplified fragments from the original gene cluster. Since lyngbyatoxin production in *E. coli* required addition of promoter sequences, and it was the only known demonstration of heterologous expression of a cyanobacterial gene cluster in *E. coli*, T7 promoters were added to the barbamide cluster, resulting in four synthetic operons (*barABCD*, *barEF*, *barGH*, and *barKJI*). The engineered cluster was cloned into a shuttle vector to facilitate downstream testing in multiple hosts. After whole-plasmid sequence verification, we tested the construct in *E. coli*, but detected no production. DNA sequencing of the plasmid after transfer to an expression strain revealed that four of the enzymes (*barEFGH*) were selectively deleted. Examination of the junction at the deletion site revealed that we likely facilitated recombination by introducing 25 bp promoter sequences that were identical (Figure 3b).

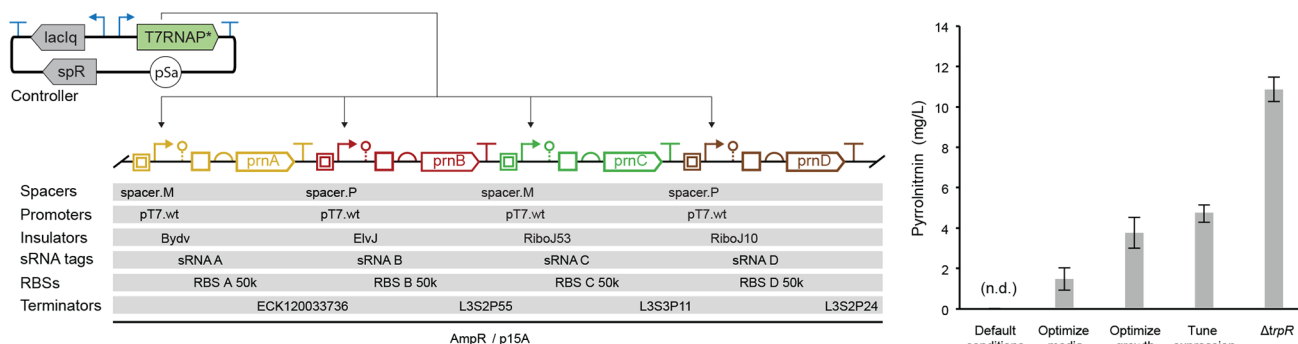
**Pacidamycin D.** The rise of antibiotic resistance is a global problem that is expected to worsen, as there is a dearth of candidates in pharmaceutical pipelines.<sup>3</sup> Uridyl peptides, of which pacidamycin D is a member, are potent via a novel mode of action against pseudomonads, which are associated with high mortality for burn and wound infections and for which multiple drug resistance is a problem.<sup>116,117</sup> The 30 kb pacidamycin D biosynthetic gene cluster had been identified in *S. coeruleo-ubidus*<sup>118,119</sup> and heterologously expressed in *S. lividans*<sup>118</sup> but not in any other hosts. Titters have not been reported.

First, we attempted conjugal transfers of the gene cluster into multiple *Streptomyces* strains (*S. albidoflavus* J1074,<sup>120</sup> *S. coelicolor* M1146,<sup>121</sup> *S. albobovineus* ATCC 33021), but we observed no production. Our attention focused on two genes in the cluster whose functions were previously reported as unknown due to a lack of homology to other genes at the time.<sup>122</sup> We found that one of the genes (*pacI*) shared 89% sequence identity with *ssaA* in the sansanmycin gene cluster, which had recently been shown to be a novel activator.<sup>123</sup> Thus, we placed it under the control of the constitutive  $P_{\text{ermE}}$  promoter, which is commonly used in *Streptomyces* engineering.<sup>124</sup> Transfer of the modified gene cluster successfully enabled pacidamycin D production in all streptomycetes we tested (Figure 3c). In *S. albidoflavus* J1074, we observed production at 13-fold the abundance of the native producer.

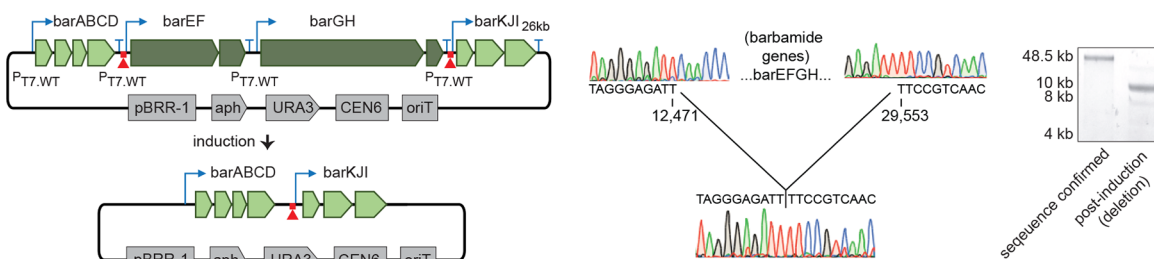
## ■ IDENTIFICATION OF A NOVEL NATURAL PRODUCER

Even if there is a known natural producer of a needed chemical that has been reported in the literature, it may be unobtainable

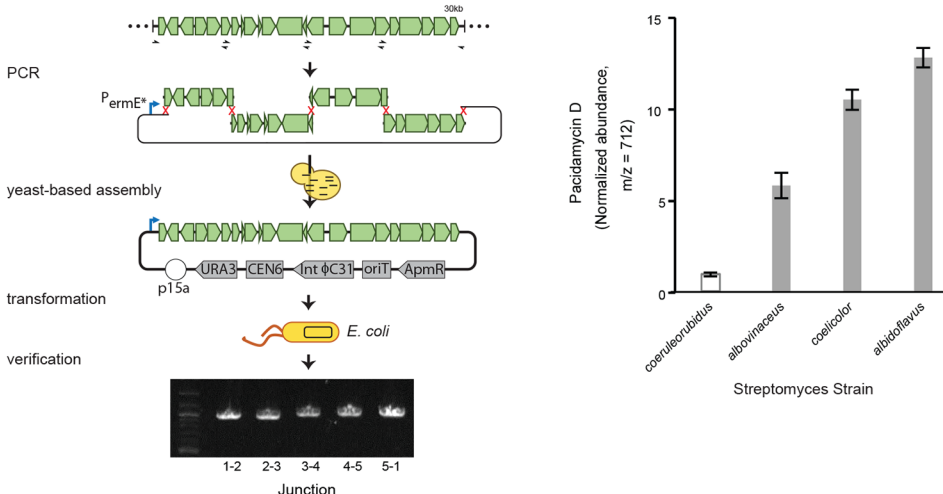
## A - Refactoring and optimizing expression (Pyrrolnitrin)



## B - Loss of gene cluster upon transfer to expression host (Barbamide)



## C - Adding expression control to boost production in heterologous hosts (Pacidamycin D)

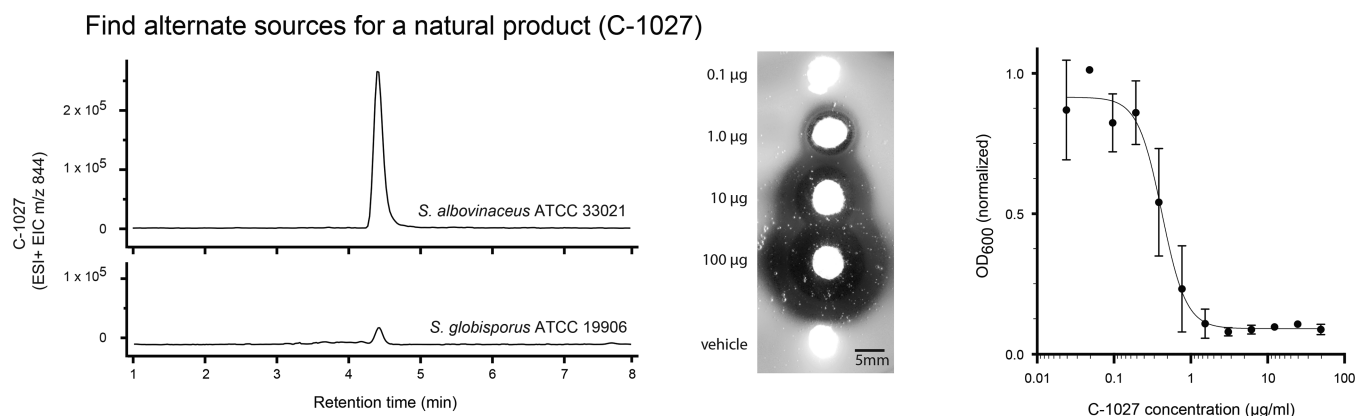


**Figure 3.** Eliminating native regulation (pyrrolnitrin, barbamide, and pacidamycin D). (A) Left: Refactored expression cassette for pyrrolnitrin production. Each refactored transcription unit comprises a CDS codon-optimized for heterologous expression in *E. coli*, a ribozyme-based insulator, a T7 promoter, a terminator, a unique target sequence to allow for sRNA-based knockdown<sup>18</sup> (not used), and a randomized spacer sequence to separate promoters from upstream regulatory elements. Part sequences are provided in [Supplementary Table 5](#). Expression levels were controlled via IPTG-inducible expression of T7RNAP included on a second plasmid. Right: Successive improvements in pyrrolnitrin production in *E. coli*, resulting from testing three different growth media, three different growth conditions, and two different T7RNAP induction levels and boosting tryptophan production by using a  $\Delta trpR$  strain (Methods). (B) Sequence-verified barbamide gene cluster cloned into a shuttle vector for heterologous expression in Gram-negative bacteria. T7 promoters were inserted as indicated. Sequencing (center) and gel (right) reveal partial loss of cluster after transfer to production strain. The sequences flanking the deletion region correspond to a pair of identical T7 promoters (red triangles). (C) Scheme for transfer of the native pacidamycin D gene cluster to heterologous hosts. The native cluster (top left) was PCR amplified and reconstructed into a plasmid via yeast-based assembly. The constitutive promoter  $P_{ermE^*}$  was added in this step. The plasmid was then transferred to *E. coli* for verification, propagation, and eventual conjugation with destination *Streptomyces* strains. Right: Heterologous pacidamycin D production in *Streptomyces* strains. *S. coeruleorubidus* is the native producer.

due to lost or inaccessible stocks, export controls, restricted use due to intellectual property constraints, or a high BSL requirement. If the genome is available, it is possible to synthesize the gene cluster DNA, but this is difficult if it is large, and it

assumes both that no genes are required outside the cluster and that it will be functional after transfer to a heterologous host. An alternative approach is to identify a different strain that has the ability to produce the same chemical. This would have been





**Figure 4.** Identification of a novel natural producer (C-1027). Production of the active form of the C-1027 chromophore in previously unreported *Streptomyces* strains. Left: LC/MS shows chromophore production ( $m/z = 884$ ) in both strains. Right: Antibacterial activity was confirmed via inhibition assays of crude extracts from *S. albobinaceus* ATCC 33021 against *Bacillus subtilis* PY7. C-1027 chromophore levels within extracts were measured via LC/MS for a diffusion assay (center) and for dose–response quantitation ( $IC_{50} = 0.43 \pm 0.04 \mu\text{g/mL}$ ) (Methods).

a daunting task even a decade ago, but the number of sequenced strains in databases is growing rapidly (the Joint Genome Institute had 47 516 microbial genomes as of October 2016<sup>125</sup>). We chose to apply this approach to produce C-1027 because we could not obtain the published strains and the gene cluster is large and complex.

**C-1027.** Enediyne antibiotics are among the most cytotoxic chemicals known, with a potent ability to destroy DNA.<sup>126</sup> Their chemical structure contains a “delivery system” that docks DNA, a “trigger device” that responds to nucleophilic attack from a nearby nucleotide, and a “warhead” that when triggered induces a double-strand break and cell death.<sup>126,127</sup> The enediyne C-1027 is used as an antitumor agent.<sup>128,129</sup> *Streptomyces globisporus* strain C-1027 was reported to produce this compound in 1998,<sup>130</sup> and the strain was subsequently optimized by overexpressing native activators.<sup>131</sup> We requested this strain from the researchers, and, unfortunately, there was no route to obtaining it or derivatives, even by MTA. Although the complete sequence of the biosynthetic gene cluster has been published,<sup>128</sup> complete synthesis during the pressure test was impractical due to its size and complexity (76 kb, 56 genes). We therefore focused on searching for other native producers. After the pressure test, the same group published another *Streptomyces* strain that produces a 10-fold higher titer (750 mg/L of chromoprotein, corresponding to 56 mg/L of the C-1027 chromophore), identified by screening the Natural Products Library Initiative at the Scripps Research Institute using PCR primers targeting the gene cluster.<sup>132</sup> It is noteworthy that it was also not possible to obtain this strain.

A simple BLAST search found nearly identical (93% and 98% identity) gene clusters in two strains available in public repositories: *S. globisporus* ATCC 19906 and *S. albobinaceus* ATCC 33021.<sup>133</sup> By testing both strains under various growth conditions, we confirmed that they both produce C-1027 via LC-MS and a growth-inhibition assay (Figure 4). We observed that *S. albobinaceus* ATCC 33021 produced the C-1027 chromophore at 2 mg/L under unoptimized growth conditions, about 20 times the level of *S. globisporus* ATCC 19906. The cluster was also cloned into a plasmid system using yeast-based assembly and transferred to heterologous strains (*S. albidoflavus* J1074, *S. coelicolor* M1146, and *E. coli*). Although C-1027 production was not detected, we did detect production in *S. albidoflavus* of a known intermediate, 3-enolpyruvoylanthrinate,<sup>134</sup> which

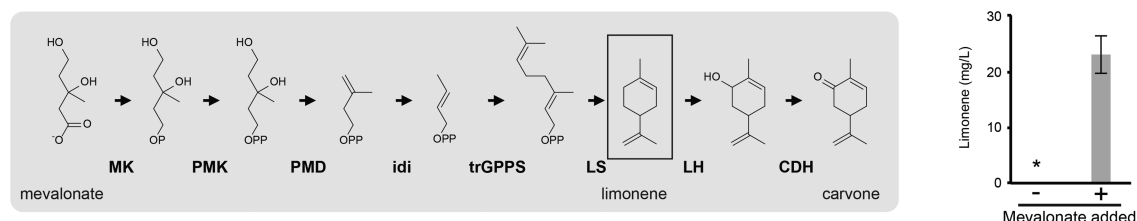
forms the intercalative DNA binding aspect of the chromophore (Methods; Supplementary Figure 3).

## CELL-FREE SYSTEMS

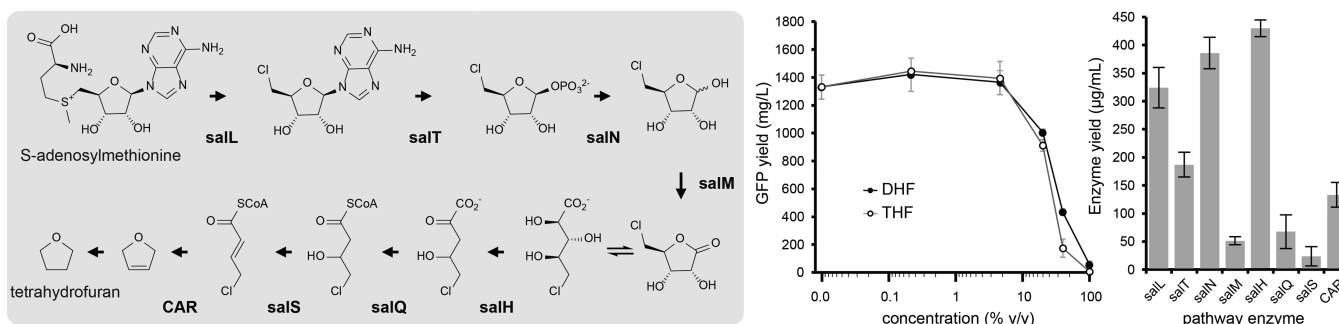
Metabolic pathways can be reconstituted *in vitro* using a cell-free system that contains all the necessary components for transcription/translation, precursors, redox, and energy.<sup>135–137</sup> This approach is valuable when the product or an intermediate is toxic. Toxicity can be difficult to predict *a priori*, but for two of the pressure test molecules we anticipated problems early. THF is thought to be toxic because it inhibits enzymatic reactions, and as a solvent it can indiscriminately disrupt the structure and activities of biological macromolecules.<sup>138,139</sup> Pathway intermediates for carvone production (e.g., limonene) are used as antimicrobials and are toxic because they spontaneously produce hydroperoxides that cause oxidative damage to cellular machinery.<sup>140,141</sup> THF and carvone are not made by bacteria or yeasts, so it was not possible to include binding proteins, transporters, or work with a resistant host, as was possible for the other antimicrobial molecules described in earlier sections. Specifically, we employed a new cell-free platform that enables systematic optimization and debugging of biosynthetic pathways.<sup>136,142–144</sup> In this platform, cell-free cocktails for synthesizing target small molecules are assembled by combining crude cell lysates containing one or more overexpressed pathway enzymes. Exploration of pathway variants is therefore reduced to a pipetting exercise, without need for genetic manipulation. Such systems can aid in selecting optimal enzyme candidates and expression levels, which can inform subsequent optimization in living cells. Cell-free expression systems were developed for both THF and carvone, and enzyme expression was validated. Toward carvone, we were able to produce the precursor limonene, but no product could be obtained for THF.

**Carvone.** Carvone is a monoterpene with many potential applications, including as a mosquito repellent<sup>145,146</sup> and consumption-safe pesticide for food storage.<sup>147</sup> It is produced in plant oils, notably mint, and can be produced by the bacterium *Rhodococcus erythropolis* when grown on carveol/limonene.<sup>148,149</sup> It can be produced via a two-enzyme recombinant pathway in *E. coli* at low levels when limonene is fed as a precursor.<sup>150</sup> Thus, we decided to focus on the first half of the pathway from the metabolite geranyl pyrophosphate to limonene. However, limonene is very toxic in *E. coli*<sup>140</sup> (potentially

## A - Prototyping monoterpene biosynthesis in vitro (Carvone / limonene)



## B - Cell-free prototyping of a novel pathway for a toxic molecule (Tetrahydrofuran)



**Figure 5.** Cell-free systems (carvone and tetrahydrofuran). (A) Biosynthetic pathway for carvone, with limonene intermediate indicated (rectangle). Gene sequences are provided in [Supplementary Table 3](#). The partial pathway (up to limonene biosynthesis) was reconstructed in a cell-free production system previously engineered to produce mevalonate,<sup>186</sup> and production of limonene was observed via GC/MS only upon feeding mevalonate (15 mM) to the system. (B) Novel pathway proposed for THF, based on cyclization of an intermediate of salinosporamide-A biosynthesis. Center: The toxicity of THF and DHF to the host system was first confirmed in the presence of up to 5% THF or DHF by measuring production of GFP (center). Right: *In vitro* yields of all cell-free synthesized enzymes were quantified by monitoring <sup>14</sup>C-leucine incorporation (right) (Methods).

limiting the published carvone titer<sup>150</sup>), so we decided to construct this pathway first using a cell-free system based on *E. coli* lysates. At the time of the pressure test, cell-free systems had been used for monoterpenoid conversions from menthone feedstocks,<sup>151</sup> but not from central metabolism. After the pressure test, Bowie and co-workers published a highly optimized cell-free system to produce limonene from glucose.<sup>152</sup>

To prototype monoterpene pathways, the platform described above was used.<sup>136</sup> We successfully built and tested pathways to convert mevalonate to limonene by using the set of enzymes described by Lee and co-workers from *S. cerevisiae*, *E. coli*, *Abies grandis* (grand fir), and *M. spicata* (spearmint)<sup>153</sup> and mixing them together along with substrate (mevalonate), salts, and cofactors. Unoptimized, our system produced 23 mg/L limonene over 12 h (Figure 5a) (Methods). Converting this to carvone would require addition of a limonene hydroxylase,<sup>154</sup> such as from *Mentha spicata* (spearmint), and a carveol dehydrogenase,<sup>155</sup> such as from *Mentha piperita* (peppermint), which we did not have time to complete during the evaluation period.

**Tetrahydrofuran.** THF is used as an industrial solvent, PVC adhesive, metal degreaser, and precursor to polymers, including Spandex.<sup>156</sup> THF presented a dual challenge because of its known toxicity<sup>138,157</sup> and because there is no known biosynthetic route. Pathways are known for compounds containing functionalized furans within larger molecules (e.g., THF ligands),<sup>158</sup> but not the isolated monomer.

We designed a novel pathway based on cyclization of a four-carbon compound in the salinosporamide-A pathway<sup>159</sup> and tested it in a cell-free production system.<sup>136</sup> Our approach was to find a compound that could be cyclized into 2,3-dihydrofuran

(DHF), which could then be chemically modified to produce THF. Key was finding a biologically accessible four-carbon intermediate with a leaving group that would favor this novel cyclization. We identified a candidate in the salinosporamide-A pathway,<sup>159</sup> 4-chloro-crotonyl-CoA, which we theorized would cyclize via a chlorine leaving group upon enzymatic removal of coenzyme-A with a carboxylic acid reductase.

To test this scheme, seven enzymes were codon-optimized for expression in *E. coli* from the salinosporamide-A pathway, as well as a carboxylic acid reductase known to act on a similar intermediate,<sup>160</sup> and were cloned into individual plasmids. After verifying that all the enzymes are produced via *E. coli* cell-free protein synthesis (CFPS) and that the system functions correctly in the presence of DHF/THF (Figure 3b), the CFPS reactions were combined together and substrate, salts, and cofactors were added. Product was not detected, which could indicate that not all enzymes were active or expressed at sufficiently high levels or that starting substrates were diverted to other molecules in the lysate, among others.

## DISCUSSION

The pressure test evaluated how quickly, when given a random organic molecule, that we could get an initial production system up and running. Different molecules have different constraints and require different paths, and the assessment did a good job of surveying this space. This is reflected in the diversity of approaches we had to deploy: there was no single automated cookie-cutter strategy that could be applied across the board by fully automated systems. Nor was there a pre-established decision scheme to govern which approaches to attempt for a

given compound. However, the test did require a centralized facility to bring together focused designers from different backgrounds and access to computer-aided design (CAD) tools and a high-throughput pipeline for DNA construction, screening, and analysis.

The test begins to evaluate the preparedness of centralized facilities to address a rapid response need for a molecule, but the present work covers only the *first phase* of such a pipeline, spanning from target identification to the initial measurement of product. The titers we were able to achieve ranged from 3.8 to 238 mg/L. A *second phase* would involve strain optimization, where the host is modified to optimize titer and yield by modifying core metabolism, redox balancing, and reducing product toxicity.<sup>161–164</sup> This would be a significant effort; for example, the optimization of farnesene production required modifying 1.2% of the yeast genome.<sup>165</sup> Breakthrough technologies are needed to design, implement, and analyze genome-scale engineering efforts.<sup>166–170</sup> Generalizable approaches to screen development would also accelerate this phase. Scaling-up the optimized strain to pilot production would encompass a *third phase*. This would benefit from miniaturized bioreactors that accurately mimic large-scale production, the demystification of the media/growth optimization, and synthetic regulation to implement process control and pathway switching. Finally, a *fourth stage* would comprise the development of a full fermentation process. This would benefit from pilot scale and national facilities that can be rapidly reconfigured for different strains, feedstocks, and products.<sup>171,172</sup> One can imagine establishing a network of facilities that specialize in different phases and then systematically applying pressure tests designed for each to spur technology development until there is a full development pipeline from desired molecule to full bioprocess. Fully implementing this would require standards to facilitate the hand-offs between phases.

The pressure test exposed gaps and associated needs particular to the first phase. First, the design process was hindered by gaps in CAD tools. Literature and database searches were inadequately linked to actionable design tools, and routine steps became noticeably time-consuming when rushed, including codon optimization, oligo design, plasmid editors, sequence confirmation, and biophysical methods and required constantly switching between software.<sup>173</sup> Retrosynthesis was largely performed manually, and, while software has been developed, it is not publically accessible. Sourcing DNA was another major bottleneck, requiring about half of the allotted time. We ordered 198 genes from four vendors, which required 3–8 weeks for delivery. Had this been reduced to 3 weeks, it would have enabled a second design iteration during the pressure test, and 3 or 4 cycles could have been performed if it were further reduced to a few days. Finally, analytical chemical methods (NMR, mass-spec, etc.) for nontargeted metabolomics to identify a product and verify its structure were woefully inadequate in the context of a timed test. This was particularly true for complex natural products and pathway intermediates and was confounded by difficulties in obtaining standards. Across the process, the gaps were largely mundane and practical, and redirected our attention away from “fancier” research areas, such as artificial intelligence, droplet microfluidics, and robotic automation.

Small molecules only scratch the surface of what is possible with engineering biology. The potential beyond is staggering. The synthesis of whole bacterial genomes has been demonstrated, and the field is on the cusp of building a synthetic yeast

genome<sup>174</sup> and booting up technologies toward a human genome.<sup>175</sup> Design lags, but is accelerating quickly, and the convergence of CAD tools will lead to new levels of genetic engineering. Our ambition is to be able to design genetic systems at the scale of genomes consisting of thousands of genetic parts. These capacities could be extended to harness pathways to non-organic materials with atomic precision, including metals and silica. Many rapid response scenarios could require the design of entire living organisms that target emerging human<sup>176</sup> or agricultural disease,<sup>177</sup> probiotics for resilience against theater pathogens, “smart” diagnostics,<sup>178</sup> treating pollutants,<sup>179</sup> deployed field sensors,<sup>180</sup> or scavenging new sources of minerals,<sup>181</sup> just to name a few. How long before foundries can be expected to design, from the bottom up, 10 genomes to create synthetic living organisms in 90 days?

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b13292.

Supplementary Methods, Supplementary Figures 1–10, and Supplementary Tables 1–9 (PDF)

Excel file (XLSX)

Supplementary Figure 11 (TIF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*cavoigt@gmail.com

\*ben@synbiofoundry.org

### ORCID

Christopher Voigt: 0000-0003-0844-4776

### Author Contributions

□ A. Casini, F.-Y. Chang, R. Eluere, A. King, and E. M. Young contributed equally.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We are grateful to Huimin Zhao (UIUC) for supplying the yeast strain used as a positive control for 1-hexadecanol, Yeojoon Yoon (Ewha Woman's University) for the pYJ1614 plasmid containing the barbamide gene cluster, Taek Soon Lee (JBEI) for providing limonene strains and advice, Toyonobu Usuki (Sophia University) for providing C-1027 standards, and Gerry Wright (McMaster University) for providing strains (*S. albidoflavus* J1074, *S. coelicolor* M1146, *E. coli* ET12567/pUZ8002) and vectors (pIJ10257, pSET152). We also thank Bradley Moore at UCSD and Wenjun Zhang at UC Berkeley for their insightful pathway design discussions. K.P., A.E.B., J.Z., A.C., Q.D., A.C., A.F., C.B., A.G., R.E., R.W.R., R.G., J.K., H.E., M.A.F., M.J., C.A.V., and D.B.G. are funded by the U.S. Defense Advanced Research Projects Agency's Living Foundries program award HR0011-15-C-0084. A.Ka. is funded by the NSF GRFP, E.Y. is funded by a research award to the Foundry from DSM, F.Y.C. is funded by a research award to the Foundry from Novartis, and A.Ki. is sponsored by a Banting postdoctoral fellowship.

## ■ REFERENCES

(1) Prado, N. D.; Pereira, S. S.; da Silva, M. P.; Morais, M. S.; Kayano, A. M.; Moreira-Dill, L. S.; Luiz, M. B.; Zanchi, F. B.; Fuly, A.



- L.; Huacca, M. E.; Fernandes, C. F.; Calderon, L. A.; Zuliani, J. P.; Pereira da Silva, L. H.; Soares, A. M.; Stabeli, R. G.; Fernandes, C. F. *PLoS One* **2016**, *11*, e0151363.
- (2) Qiu, X.; Wong, G.; Audet, J.; Bello, A.; Fernando, L.; Alimonti, J. B.; Fausther-Bovendo, H.; Wei, H.; Aviles, J.; Hiatt, E.; Johnson, A.; Morton, J.; Swope, K.; Bohorov, O.; Bohorova, N.; Goodman, C.; Kim, D.; Pauly, M. H.; Velasco, J.; Pettitt, J.; Olinger, G. G.; Whaley, K.; Xu, B.; Strong, J. E.; Zeitlin, L.; Kobinger, G. P. *Nature* **2014**, *514*, 47–53.
- (3) Brown, E. D.; Wright, G. D. *Nature* **2016**, *529*, 336–343.
- (4) Fischbach, M. A.; Walsh, C. T. *Science* **2009**, *325*, 1089–1093.
- (5) <https://www.accessdata.fda.gov/scripts/drugshortages/default.cfm>.
- (6) United States Government Accountability Office. United States Congress. Senate Committee on Armed Services. United States House Committee on Armed Services. *Defense logistics: actions needed to improve the availability of critical items during current and future operations: report to Congressional committees*. U.S. Government Accountability Office: Washington, D.C., 2005.
- (7) [http://articles.mcall.com/2000-06-15/news/3303465\\_1\\_explosion-chemical-hydroxylamine](http://articles.mcall.com/2000-06-15/news/3303465_1_explosion-chemical-hydroxylamine).
- (8) Majors, R. E. *Lc Gc North America* **2009**, *27*, 458.
- (9) Carlson, R. *Nat. Biotechnol.* **2016**, *34*, 247–255.
- (10) [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/national\\_bioeconomy\\_blueprint\\_april\\_2012.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/national_bioeconomy_blueprint_april_2012.pdf).
- (11) [http://www.aci-na.org/static/entransit/faa\\_deicing\\_alts\\_9-15-08.pdf](http://www.aci-na.org/static/entransit/faa_deicing_alts_9-15-08.pdf).
- (12) <http://www.cryotech.com/defense-logistics-agency-awards-outstanding-readiness-support-award>.
- (13) Chao, R.; Mishra, S.; Si, T.; Zhao, H. *Metab. Eng.* **2017**, *42*, 98–108.
- (14) Yu, A. Q.; Juwono, N. K.; Foo, J. L.; Leong, S. S.; Chang, M. W. *Metab. Eng.* **2016**, *34*, 36–43.
- (15) Blanch, H. W.; Adams, P. D.; Andrews-Cramer, K. M.; Frommer, W. B.; Simmons, B. A.; Keasling, J. D. *ACS Chem. Biol.* **2008**, *3*, 17–20.
- (16) Richardson, S. M.; Mitchell, L. A.; Stracquadanio, G.; Yang, K.; Dymond, J. S.; DiCarlo, J. E.; Lee, D.; Huang, C. L.; Chandrasegaran, S.; Cai, Y.; Boeke, J. D.; Bader, J. S. *Science* **2017**, *355*, 1040–1044.
- (17) Chambers, S.; Kitney, R.; Freemont, P. *Biochem. Soc. Trans.* **2016**, *44*, 687–688.
- (18) Fletcher, L.; Rosser, S.; Elflick, A. *Biochem. Soc. Trans.* **2016**, *44*, 692–695.
- (19) <https://twistbioscience.com/>.
- (20) <http://www.ginkgobioworks.com/>.
- (21) <https://amyris.com/>.
- (22) <https://www.zymergen.com/>.
- (23) <https://www.syntheticgenomics.com/>.
- (24) <https://syncti.org/>.
- (25) <https://www.musebio.com/>.
- (26) Moul, J.; Fidelis, K.; Kryshtafovich, A.; Schwede, T.; Tramontano, A. *Proteins: Struct., Funct., Genet.* **2016**, *84*, 4–14.
- (27) Wodak, S. J.; Janin, J. *Proteins: Struct., Funct., Genet.* **2017**, *85*, 357–358.
- (28) Radivojac, P.; Clark, W. T.; Oron, T. R.; Schnoes, A. M.; Wittkop, T.; Sokolov, A.; Graim, K.; Funk, C.; Verspoor, K.; Ben-Hur, A.; Pandey, G.; Yunes, J. M.; Talwalkar, A. S.; Repo, S.; Souza, M. L.; Piovesan, D.; Casadio, R.; Wang, Z.; Cheng, J.; Fang, H.; Gough, J.; Koskinen, P.; Toronen, P.; Nokso-Koivisto, J.; Holm, L.; Cozzetto, D.; Buchan, D. W.; Bryson, K.; Jones, D. T.; Limaye, B.; Inamdar, H.; Datta, A.; Manjari, S. K.; Joshi, R.; Chitale, M.; Kihara, D.; Lisewski, A. M.; Erdin, S.; Venner, E.; Lichtarge, O.; Rentzsch, R.; Yang, H.; Romero, A. E.; Bhat, P.; Pacanaro, A.; Hamp, T.; Kassner, R.; Seemayer, S.; Vicedo, E.; Schaefer, C.; Achten, D.; Auer, F.; Boehm, A.; Braun, T.; Hecht, M.; Heron, M.; Honigschmid, P.; Hopf, T. A.; Kaufmann, S.; Kiening, M.; Krompass, D.; Landerer, C.; Mahlich, Y.; Roos, M.; Bjorne, J.; Salakoski, T.; Wong, A.; Shatkay, H.; Gatzmann, F.; Sommer, I.; Wass, M. N.; Sternberg, M. J.; Skunca, N.; Supek, F.; Bosnjak, M.; Panov, P.; Dzeroski, S.; Smuc, T.; Kourmpetis, Y. A.; van Dijk, A. D.; ter Braak, C. J.; Zhou, Y.; Gong, Q.; Dong, X.; Tian, W.; Falda, M.; Fontana, P.; Lavezzo, E.; Di Camillo, B.; Toppo, S.; Lan, L.; Djuric, N.; Guo, Y.; Vucetic, S.; Bairoch, A.; Linial, M.; Babbitt, P. C.; Brenner, S. E.; Orenge, C.; Rost, B.; Mooney, S. D.; Friedberg, I. *Nat. Methods* **2013**, *10*, 221–227.
- (29) Hossain, M. *Journal of Organization Design* **2014**, *3*, 46–52.
- (30) Krotkov, E.; Hackett, D.; Jackel, L.; Perschbacher, M.; Pippine, J.; Strauss, J.; Pratt, G.; Orłowski, C. The DARPA Robotics Challenge Finals Results and Perspectives. *J. Field Robotics* **2017**, *34*, 229–240.1002/rob.21683.
- (31) <https://www.federalreserve.gov/bankinfo/stress-tests/CCAR/201503-comprehensive-capital-analysis-review-preface.htm>.
- (32) Bayer, T. S.; Widmaier, D. M.; Temme, K.; Mirsky, E. A.; Santi, D. V.; Voigt, C. A. *J. Am. Chem. Soc.* **2009**, *131*, 6508–6515.
- (33) Bokinsky, G.; Peralta-Yahya, P. P.; George, A.; Holmes, B. M.; Steen, E. J.; Dietrich, J.; Lee, T. S.; Tullman-Ercek, D.; Voigt, C. A.; Simmons, B. A.; Keasling, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 19949–19954.
- (34) Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. *Science* **2010**, *330*, 70–74.
- (35) Xu, P.; Gu, Q.; Wang, W.; Wong, L.; Bower, A. G.; Collins, C. H.; Koffas, M. A. *Nat. Commun.* **2013**, *4*, 1409.
- (36) Smanski, M. J.; Bhatia, S.; Zhao, D.; Park, Y.; L, B. A. W.; Giannoukos, G.; Ciulla, D.; Busby, M.; Calderon, J.; Nicol, R.; Gordon, D. B.; Densmore, D.; Voigt, C. A. *Nat. Biotechnol.* **2014**, *32*, 1241–1249.
- (37) Roehner, N.; Young, E. M.; Voigt, C. A.; Gordon, D. B.; Densmore, D. *ACS Synth. Biol.* **2016**, *5*, 507–517.
- (38) Oberortner, E.; Cheng, J. F.; Hillson, N. J.; Deutsch, S. *ACS Synth. Biol.* **2017**, *6*, 485–496.
- (39) Hillson, N. J.; Rosengarten, R. D.; Keasling, J. D. *ACS Synth. Biol.* **2012**, *1*, 14–21.
- (40) King, Z. A.; Lloyd, C. J.; Feist, A. M.; Palsson, B. O. *Curr. Opin. Biotechnol.* **2015**, *35*, 23–29.
- (41) Mudge, S. M.; Belanger, S. E.; Nielsen, A. M. *Fatty Alcohols: Anthropogenic and Natural Occurrence in the Environment*; Royal Society of Chemistry: Cambridge, UK, 2008.
- (42) Akhtar, M. K.; Turner, N. J.; Jones, P. R. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 87–92.
- (43) [http://everyspec.com/MIL-SPECS/MIL-SPECS-MIL-L/MIL-L-87132B\\_8347/](http://everyspec.com/MIL-SPECS/MIL-SPECS-MIL-L/MIL-L-87132B_8347/).
- (44) Feng, X.; Lian, J.; Zhao, H. *Metab. Eng.* **2015**, *27*, 10–19.
- (45) Hellenbrand, J.; Biester, E. M.; Gruber, J.; Hamberg, M.; Frentzen, M. *BMC Biochem* **2011**, *12*, 64.
- (46) Mueller, U.; Wu, L.; Raamsdonk, L. M.; Winkler, A. A. Patent WO2009013159A2, 2014.
- (47) De Luca, V.; Salim, V.; Levac, D.; Atsumi, S. M.; Yu, F. *Methods Enzymol.* **2012**, *515*, 207–229.
- (48) O'Connor, S. E.; Maresh, J. J. *Nat. Prod. Rep.* **2006**, *23*, 532–547.
- (49) Zhu, J.; Wang, M.; Wen, W.; Yu, R. *Pharmacogn. Rev.* **2015**, *9*, 24–28.
- (50) Brown, S.; Clastre, M.; Courdavault, V.; O'Connor, S. E. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 3205–3210.
- (51) Qu, Y.; Easson, M. L.; Froese, J.; Simionescu, R.; Hudlicky, T.; De Luca, V. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 6224–6229.
- (52) Tatsis, E. C.; Carqueijeiro, I.; de Bernonville, T. D.; Franke, J.; Dang, T. T. T.; Oudin, A.; Lanoue, A.; Lafontaine, F.; Stavrinides, A. K.; Clastre, M.; Courdavault, V.; O'Connor, S. E. *Nat. Commun.* **2017**, *8*, 8.
- (53) Moon, T. S.; Yoon, S. H.; Lanza, A. M.; Roy-Mayhew, J. D.; Prather, K. L. *Appl. Environ. Microbiol.* **2009**, *75*, 589–595.
- (54) Zhou, H.; Vonk, B.; Roubos, J. A.; Bovenberg, R. A.; Voigt, C. A. *Nucleic Acids Res.* **2015**, *43*, 10560–10570.
- (55) Nicolaou, K. C.; E, J. S. *Classics in Total Synthesis: Targets, Strategies*; Wiley-VCH: Weinheim, 1996.
- (56) <http://20n.com/>; <http://arzeda.com>.

- (57) Hadadi, N.; Hatzimanikatis, V. *Curr. Opin. Chem. Biol.* **2015**, *28*, 99–104.
- (58) Favaro, L. C.; Sebastianes, F. L.; Araujo, W. L. *PLoS One* **2012**, *7*, e36826.
- (59) Robertson, K. L.; Mostaghim, A.; Cuomo, C. A.; Soto, C. M.; Lebedev, N.; Bailey, R. F.; Wang, Z. *PLoS One* **2012**, *7*, e48674.
- (60) Onofri, S.; Barreca, D.; Selbmann, L.; Isola, D.; Rabbow, E.; Horneck, G.; de Vera, J. P.; Hatton, J.; Zucconi, L. *Stud. Mycol.* **2008**, *61*, 99–109.
- (61) Araujo, F. D. D.; Favaro, L. C. D.; Araujo, W. L.; de Oliveira, F. L.; Aparicio, R.; Marsaioli, A. J. *Eur. J. Org. Chem.* **2012**, *2012*, 5225–5230.
- (62) Fokin, M.; Fleetwood, D.; Weir, B. S.; Villas-Boas, S. *Genome Announc* **2017**, *5*, e00557-17.10.1128/genomeA.00557-17.
- (63) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Brucoleri, R.; Lee, S. Y.; Fischbach, M. A.; Muller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. *Nucleic Acids Res.* **2015**, *43*, W237–243.
- (64) Ellerbrock, P.; Armanino, N.; Ilg, M. K.; Webster, R.; Trauner, D. *Nat. Chem.* **2015**, *7*, 879–882.
- (65) Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; de Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; Cruz-Morales, P.; Duddela, S.; Dusterhus, S.; Edwards, D. J.; Fewer, D. P.; Garg, N.; Geiger, C.; Gomez-Escribano, J. P.; Greule, A.; Hadjithomas, M.; Haines, A. S.; Helfrich, E. J.; Hillwig, M. L.; Ishida, K.; Jones, A. C.; Jones, C. S.; Jungmann, K.; Kegler, C.; Kim, H. U.; Kotter, P.; Krug, D.; Masschelein, J.; Melnik, A. V.; Mantovani, S. M.; Monroe, E. A.; Moore, M.; Moss, N.; Nuttmann, H. W.; Pan, G.; Pati, A.; Petras, D.; Reen, F. J.; Rosconi, F.; Rui, Z.; Tian, Z.; Tobias, N. J.; Tsunematsu, Y.; Wiemann, P.; Wyckoff, E.; Yan, X.; Yim, G.; Yu, F.; Xie, Y.; Aigle, B.; Apel, A. K.; Balibar, C. J.; Balskus, E. P.; Barona-Gomez, F.; Bechthold, A.; Bode, H. B.; Borriess, R.; Brady, S. F.; Brakhage, A. A.; Caffrey, P.; Cheng, Y. Q.; Clardy, J.; Cox, R. J.; De Mot, R.; Donadio, S.; Donia, M. S.; van der Donk, W. A.; Dorrestein, P. C.; Doyle, S.; Driessen, A. J.; Ehling-Schulz, M.; Entian, K. D.; Fischbach, M. A.; Gerwick, L.; Gerwick, W. H.; Gross, H.; Gust, B.; Hertweck, C.; Hofte, M.; Jensen, S. E.; Ju, J.; Katz, L.; Kaysser, L.; Klassen, J. L.; Keller, N. P.; Kormanec, J.; Kuipers, O. P.; Kuzuyama, T.; Kypripides, N. C.; Kwon, H. J.; Lautru, S.; Lavigne, R.; Lee, C. Y.; Linquan, B.; Liu, X.; Liu, W.; Luzhetskyy, A.; Mahmud, T.; Mast, Y.; Mendez, C.; Metsa-Ketela, M.; Micklefield, J.; Mitchell, D. A.; Moore, B. S.; Moreira, L. M.; Muller, R.; Neilan, B. A.; Nett, M.; Nielsen, J.; O’Gara, F.; Oikawa, H.; Osbourn, A.; Osburne, M. S.; Ostash, B.; Payne, S. M.; Pernodet, J. L.; Petricek, M.; Piel, J.; Ploux, O.; Raaijmakers, J. M.; Salas, J. A.; Schmitt, E. K.; Scott, B.; Seipke, R. F.; Shen, B.; Sherman, D. H.; Sivonen, K.; Smanski, M. J.; Sosio, M.; Stegmann, E.; Sussmuth, R. D.; Tahlan, K.; Thomas, C. M.; Tang, Y.; Truman, A. W.; Viaud, M.; Walton, J. D.; Walsh, C. T.; Weber, T.; van Wezel, G. P.; Wilkinson, B.; Willey, J. M.; Wohlleben, W.; Wright, G. D.; Ziemert, N.; Zhang, C.; Zotchev, S. B.; Breitling, R.; Takano, E.; Glockner, F. O. *Nat. Chem. Biol.* **2015**, *11*, 625–631.
- (66) Mayer, A. M.; Staples, R. C. *Phytochemistry* **2002**, *60*, 551–565.
- (67) Walsh, C. T.; Wencewicz, T. A. *Nat. Prod. Rep.* **2013**, *30*, 175–200.
- (68) Huijbers, M. M.; Montersino, S.; Westphal, A. H.; Tischler, D.; van Berkel, W. J. *Arch. Biochem. Biophys.* **2014**, *544*, 2–17.
- (69) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253–2277.
- (70) Bugg, T. D.; Ramaswamy, S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 134–140.
- (71) Cox, R. J.; Al-Fahad, A. *Curr. Opin. Chem. Biol.* **2013**, *17*, 532–536.
- (72) Davison, J.; al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7642–7647.
- (73) Medema, M. H.; Takano, E.; Breitling, R. *Mol. Biol. Evol.* **2013**, *30*, 1218–1223.
- (74) Sanchez, C.; Mendez, C.; Salas, J. A. *Nat. Prod. Rep.* **2006**, *23*, 1007–1045.
- (75) Hyun, C. G.; Bililign, T.; Liao, J.; Thorson, J. S. *ChemBioChem* **2003**, *4*, 114–117.
- (76) Alkhalaf, L. M.; Du, Y. L.; Ryan, K. S. *Methods Enzymol.* **2016**, *575*, 21–37.
- (77) Espah Borujeni, A.; Salis, H. M. *J. Am. Chem. Soc.* **2016**, *138*, 7016–7023.
- (78) Lou, C.; Stanton, B.; Chen, Y. J.; Munsky, B.; Voigt, C. A. *Nat. Biotechnol.* **2012**, *30*, 1137–1142.
- (79) Woodruff, L. B. A.; Gorochowski, T. E.; Roehner, N.; Mikkelsen, T. S.; Densmore, D.; Gordon, D. B.; Nicol, R.; Voigt, C. A. *Nucleic Acids Res.* **2017**, *45*, 1553–1565.
- (80) Sanchez, C.; Zhu, L.; Brana, A. F.; Salas, A. P.; Rohr, J.; Mendez, C.; Salas, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 461–466.
- (81) Sanchez, C.; Mendez, C.; Salas, J. A. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, S60–S68.
- (82) Du, Y. L.; Ryan, K. S. *ACS Synth. Biol.* **2015**, *4*, 682–688.
- (83) Du, Y. L.; Williams, D. E.; Patrick, B. O.; Andersen, R. J.; Ryan, K. S. *ACS Chem. Biol.* **2014**, *9*, 2748–2754.
- (84) Chang, F. Y.; Brady, S. F. *J. Am. Chem. Soc.* **2011**, *133*, 9996–9999.
- (85) Chang, F. Y.; Ternei, M. A.; Calle, P. Y.; Brady, S. F. *J. Am. Chem. Soc.* **2013**, *135*, 17906–17912.
- (86) Chang, F. Y.; Ternei, M. A.; Calle, P. Y.; Brady, S. F. *J. Am. Chem. Soc.* **2015**, *137*, 6044–6052.
- (87) Sanchez, C.; Butovich, I. A.; Brana, A. F.; Rohr, J.; Mendez, C.; Salas, J. A. *Chem. Biol.* **2002**, *9*, 519–531.
- (88) Brady, S. F.; Chao, C. J.; Handelsman, J.; Clardy, J. *Org. Lett.* **2001**, *3*, 1981–1984.
- (89) Fischbach, M.; Voigt, C. A. *Biotechnol. J.* **2010**, *5*, 1277–1296.
- (90) Song, M.; Sukovich, D. J.; Ciccarelli, L.; Mayr, J.; Fernandez-Rodriguez, J.; Mirsky, E. A.; Tucker, A. C.; Gordon, D. B.; Marlovits, T. C.; Voigt, C. A. *Nat. Commun.* **2017**, *8*, 14737.
- (91) Scherlach, K.; Hertweck, C. *Org. Biomol. Chem.* **2009**, *7*, 1753–1760.
- (92) Chan, L. Y.; Kosuri, S.; Endy, D. *Mol. Syst. Biol.* **2005**, *1*, 0018.
- (93) Temme, K.; Zhao, D.; Voigt, C. A. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7085–7090.
- (94) Shao, Z.; Rao, G.; Li, C.; Abil, Z.; Luo, Y.; Zhao, H. *ACS Synth. Biol.* **2013**, *2*, 662–669.
- (95) Guo, C. J.; Chang, F. Y.; Wyche, T. P.; Backus, K. M.; Acker, T. M.; Funabashi, M.; Taketani, M.; Donia, M. S.; Nayfach, S.; Pollard, K. S.; Craik, C. S.; Cravatt, B. F.; Clardy, J.; Voigt, C. A.; Fischbach, M. A. *Cell* **2017**, *168*, 517–526.
- (96) Smanski, M. J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M. A.; Voigt, C. A. *Nat. Rev. Microbiol.* **2016**, *14*, 135–149.
- (97) Biggins, J. B.; Liu, X.; Feng, Z.; Brady, S. F. *J. Am. Chem. Soc.* **2011**, *133*, 1638–1641.
- (98) Wenzel, S. C.; Gross, F.; Zhang, Y.; Fu, J.; Stewart, A. F.; Muller, R. *Chem. Biol.* **2005**, *12*, 349–356.
- (99) Kilani, J.; Fillinger, S. *Front. Microbiol.* **2016**, *7*, 2014.
- (100) [http://www.caplock.com/Anti\\_Fungal\\_Treatment.htm](http://www.caplock.com/Anti_Fungal_Treatment.htm).
- (101) Tripathi, R. K.; Gottlieb, D. J. *Bacteriol.* **1969**, *100*, 310–318.
- (102) Cartwright, D. K.; Chilton, W.; Benson, D. *Appl. Microbiol. Biotechnol.* **1995**, *43*, 211–216.
- (103) Hammer, P. E.; Hill, D. S.; Lam, S. T.; Van Pée, K.-H.; Ligon, J. M. *Appl. Environ. Microbiol.* **1997**, *63*, 2147–2154.
- (104) Hammer, P. E.; Hill, D. S.; Lam, S. T.; Van Pée, K. H.; Ligon, J. M. *Appl. Environ. Microbiol.* **1997**, *63*, 2147–2154.
- (105) Schmidt, S.; Blom, J. F.; Pernthaler, J.; Berg, G.; Baldwin, A.; Mahenthiralingam, E.; Eberl, L. *Environ. Microbiol.* **2009**, *11*, 1422–1437.
- (106) Selin, C.; Fernando, W. D.; de Kievit, T. *Microbiology* **2012**, *158*, 896–907.
- (107) Sarniguet, A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1225–1259.
- (108) Temme, K.; Hill, R.; Segall-Shapiro, T. H.; Moser, F.; Voigt, C. A. *Nucleic Acids Res.* **2012**, *40*, 8773–8781.
- (109) Zhao, Z.-J.; Zou, C.; Zhu, Y.-X.; Dai, J.; Chen, S.; Wu, D.; Wu, J.; Chen, J. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1921–1929.

- (110) Rodrigues, A. L.; Trachtmann, N.; Becker, J.; Lohanatha, A. F.; Blotenberg, J.; Bolten, C. J.; Korneli, C.; de Souza Lima, A. O.; Porto, L. M.; Sprenger, G. A.; Wittmann, C. *Metab. Eng.* **2013**, *20*, 29–41.
- (111) Kirschner, C. M.; Brennan, A. B. *Annu. Rev. Mater. Res.* **2012**, *42*, 211–229.
- (112) Schultz, M. P.; Bendick, J. A.; Holm, E. R.; Hertel, W. M. *Biofouling* **2011**, *27*, 87–98.
- (113) Orjala, J.; Gerwick, W. H. *J. Nat. Prod.* **1996**, *59*, 427–430.
- (114) Ongley, S. E.; Bian, X.; Zhang, Y.; Chau, R.; Gerwick, W. H.; Müller, R.; Neilan, B. A. *ACS Chem. Biol.* **2013**, *8*, 1888–1893.
- (115) Chang, Z.; Flatt, P.; Gerwick, W. H.; Nguyen, V.-A.; Willis, C. L.; Sherman, D. H. *Gene* **2002**, *296*, 235–247.
- (116) Fernandes, P. B.; Swanson, R. N.; Hardy, D. J.; Hanson, C. W.; Coen, L.; Rasmussen, R. R.; Chen, R. H. *J. Antibiot.* **1989**, *42*, 521–526.
- (117) Okamoto, K.; Sakagami, M.; Feng, F.; Togame, H.; Takemoto, H.; Ichikawa, S.; Matsuda, A. *Org. Lett.* **2011**, *13*, 5240–5243.
- (118) Rackham, E. J.; Gruschow, S.; Ragab, A. E.; Dickens, S.; Goss, R. J. *ChemBioChem* **2010**, *11*, 1700–1709.
- (119) Zhang, W.; Ostash, B.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16828–16833.
- (120) Zaburanyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. *BMC Genomics* **2014**, *15*, 97.
- (121) Gomez-Escribano, J. P.; Bibb, M. J. *Microb. Biotechnol.* **2011**, *4*, 207–215.
- (122) Chen, W.; Qi, J.; Wu, P.; Wan, D.; Liu, J.; Feng, X.; Deng, Z. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 401–417.
- (123) Li, Q.; Wang, L.; Xie, Y.; Wang, S.; Chen, R.; Hong, B. *J. Bacteriol.* **2013**, *195*, 2232–2243.
- (124) Bibb, M. J.; White, J.; Ward, J. M.; Janssen, G. R. *Mol. Microbiol.* **1994**, *14*, 533–545.
- (125) Chen, I. A.; Markowitz, V. M.; Chu, K.; Palaniappan, K.; Szeto, E.; Pillay, M.; Ratner, A.; Huang, J.; Andersen, E.; Huntemann, M.; Varghese, N.; Hadjithomas, M.; Tennessen, K.; Nielsen, T.; Ivanova, N. N.; Kyrpides, N. C. *Nucleic Acids Res.* **2017**, *45*, D507–D516.
- (126) Nicolaou, K. C.; Smith, A. L.; Yue, E. W. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5881–5888.
- (127) Kraka, E.; Cremer, D. *J. Am. Chem. Soc.* **2000**, *122*, 8245–8264.
- (128) Liu, W.; Christenson, S. D.; Standage, S.; Shen, B. *Science* **2002**, *297*, 1170–1173.
- (129) Li, W.; Li, X.; Huang, T.; Teng, Q.; Crnovcic, I.; Rader, C.; Shen, B. *Bioorg. Med. Chem.* **2016**, *24*, 3887–3892.
- (130) Hu, J.; Xue, Y.-C.; Xie, M.-Y.; Zhang, R.; OTANI, T.; MINAMI, Y.; YAMADA, Y.; MARUNAKA, T. *J. Antibiot.* **1988**, *41*, 1575–1579.
- (131) Li, X.; Yu, T.; He, Q.; McDowall, K. J.; Jiang, B.; Jiang, Z.; Wu, L.; Li, G.; Li, Q.; Wang, S.; Shi, Y.; Wang, L.; Hong, B. *Mol. Microbiol.* **2015**, *96*, 1257–1271.
- (132) Yan, X.; Ge, H.; Huang, T.; Yang, D.; Teng, Q.; Crnovčić, I.; Li, X.; Rudolf, J. D.; Lohman, J. R.; Gansemans, Y. *mBio* **2016**, *7*, e02104–02116.
- (133) Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K. S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W. *Nat. Chem. Biol.* **2014**, *10*, 963–968.
- (134) Van Lanen, S. G.; Lin, S.; Shen, B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 494–499.
- (135) Karim, A. S.; Dudley, Q. M.; Jewett, M. C. *Industrial Biotechnology: Microorganisms* **2016**, 125–148.
- (136) Karim, A. S.; Jewett, M. C. *Metab. Eng.* **2016**, *36*, 116–126.
- (137) Bujara, M.; Schümperli, M.; Pellaux, R.; Heinemann, M.; Panke, S. *Nat. Chem. Biol.* **2011**, *7*, 271–277.
- (138) Moody, D. E. *Drug Chem. Toxicol.* **1991**, *14*, 319–342.
- (139) Sikkema, J.; de Bont, J. A.; Poolman, B. *Microbiol. Rev.* **1995**, *59*, 201–222.
- (140) Chubukov, V.; Mingardon, F.; Schackwitz, W.; Baidoo, E. E.; Alonso-Gutierrez, J.; Hu, Q.; Lee, T. S.; Keasling, J. D.; Mukhopadhyay, A. *Appl. Environ. Microbiol.* **2015**, *81*, 4690–4696.
- (141) Sikkema, J.; De Bont, J.; Poolman, B. *Microbiological reviews* **1995**, *59*, 201–222.
- (142) Dudley, Q. M.; Karim, A. S.; Jewett, M. C. *Biotechnol. J.* **2015**, *10*, 69–82.
- (143) Dudley, Q. M.; Anderson, K. C.; Jewett, M. C. *ACS Synth. Biol.* **2016**, *5*, 157810.1021/acssynbio.6b00154.
- (144) Kay, J. E.; Jewett, M. C. *Metab. Eng.* **2015**, *32*, 133–142.
- (145) Koul, O.; Walia, S.; Dhaliwal, G. *Bioprocess Int.* **2008**, *4*, 63–84.
- (146) Tripathi, A. K.; Prajapati, V.; Kumar, S. *J. Econ. Entomol.* **2003**, *96*, 1594–1601.
- (147) de Carvalho, C. C. C. R.; da Fonseca, M. M. R. *Food Chem.* **2006**, *95*, 413–422.
- (148) van der Werf, M. J.; Boot, A. M. *Microbiology* **2000**, *146*, 1129–1141.
- (149) van der Werf, M. J.; Swarts, H. J.; de Bont, J. A. *Appl. Environ. Microbiol.* **1999**, *65*, 2092–2102.
- (150) Carter, O. A.; Peters, R. J.; Croteau, R. *Phytochemistry* **2003**, *64*, 425–433.
- (151) Toogood, H. S.; Ni Cheallaigh, A.; Tait, S.; Mansell, D. J.; Jervis, A.; Lygidakis, A.; Humphreys, L.; Takano, E.; Gardiner, J. M.; Scrutton, N. S. *ACS Synth. Biol.* **2015**, *4*, 1112–1123.
- (152) Korman, T. P.; Opgenorth, P. H.; Bowie, J. U. *Nat. Commun.* **2017**, *8*, 15526.
- (153) Alonso-Gutierrez, J.; Chan, R.; Batth, T. S.; Adams, P. D.; Keasling, J. D.; Petzold, C. J.; Lee, T. S. *Metab. Eng.* **2013**, *19*, 33–41.
- (154) Lupien, S.; Karp, F.; Wildung, M.; Croteau, R. *Arch. Biochem. Biophys.* **1999**, *368*, 181–192.
- (155) Ringer, K. L.; Davis, E. M.; Croteau, R. *Plant Physiol.* **2005**, *137*, 863–872.
- (156) Pruckmayr, G.; Dreyfuss, P.; Dreyfuss, M. *Kirk-Othmer Encyclopedia of Chemical Technology* **2000**, *1*, 1–30.
- (157) Yao, Y.; Guan, J.; Tang, P.; Jiao, H.; Lin, C.; Wang, J.; Lu, Z.; Min, H.; Gao, H. *Bioresour. Technol.* **2010**, *101*, S213–S221.
- (158) Albertson, A. K.; Lumb, J. P. *Angew. Chem., Int. Ed.* **2015**, *54*, 2204–2208.
- (159) Eustaquio, A. S.; McGlinchey, R. P.; Liu, Y.; Hazzard, C.; Beer, L. L.; Florova, G.; Alhamadsheh, M. M.; Lechner, A.; Kale, A. J.; Kobayashi, Y.; Reynolds, K. A.; Moore, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 12295–12300.
- (160) Yim, H.; Haselbeck, R.; Niu, W.; Pujol-Baxley, C.; Burgard, A.; Boldt, J.; Khandurina, J.; Trawick, J. D.; Osterhout, R. E.; Stephen, R.; Estadilla, J.; Teisan, S.; Schreyer, H. B.; Andrae, S.; Yang, T. H.; Lee, S. Y.; Burk, M. J.; Van Dien, S. *Nat. Chem. Biol.* **2011**, *7*, 445–452.
- (161) Federowicz, S.; Kim, D.; Ebrahim, A.; Lerman, J.; Nagarajan, H.; Cho, B.-k.; Zengler, K.; Palsson, B. *PLoS Genet.* **2014**, *10*, e1004264.
- (162) Dunlop, M. J.; Dossani, Z. Y.; Szmids, H. L.; Chu, H. C.; Lee, T. S.; Keasling, J. D.; Hadi, M. Z.; Mukhopadhyay, A. *Mol. Syst. Biol.* **2011**, *7*, 487.
- (163) Park, J. H.; Lee, S. Y. *Curr. Opin. Biotechnol.* **2008**, *19*, 454–460.
- (164) Alper, H.; Moxley, J.; Nevoigt, E.; Fink, G. R.; Stephanopoulos, G. *Science* **2006**, *314*, 1565–1568.
- (165) Meadows, A. L.; Hawkins, K. M.; Tsegaye, Y.; Antipov, E.; Kim, Y.; Raetz, L.; Dahl, R. H.; Tai, A.; Mahatdejkul-Meadows, T.; Xu, L.; Zhao, L.; Dasika, M. S.; Murarka, A.; Lenihan, J.; Eng, D.; Leng, J. S.; Liu, C. L.; Wenger, J. W.; Jiang, H.; Chao, L.; Westfall, P.; Lai, J.; Ganesan, S.; Jackson, P.; Mans, R.; Platt, D.; Reeves, C. D.; Saija, P. R.; Wichmann, G.; Holmes, V. F.; Benjamin, K.; Hill, P. W.; Gardner, T. S.; Tsong, A. E. *Nature* **2016**, *537*, 694–697.
- (166) Schellenberger, J.; Que, R.; Fleming, R. M.; Thiele, I.; Orth, J. D.; Feist, A. M.; Zielinski, D. C.; Bordbar, A.; Lewis, N. E.; Rahmadian, S. *Nat. Protoc.* **2011**, *6*, 1290.
- (167) Burgard, A. P.; Pharkya, P.; Maranas, C. D. *Biotechnol. Bioeng.* **2003**, *84*, 647–657.
- (168) Warner, J. R.; Reeder, P. J.; Karimpour-Fard, A.; Woodruff, L. B.; Gill, R. T. *Nat. Biotechnol.* **2010**, *28*, 856–862.
- (169) Horwitz, A. A.; Walter, J. M.; Schubert, M. G.; Kung, S. H.; Hawkins, K.; Platt, D. M.; Hernday, A. D.; Mahatdejkul-Meadows, T.; Szeto, W.; Chandran, S. S. *Cell Systems* **2015**, *1*, 88–96.



- (170) Rogers, J. K.; Church, G. M. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2388–2393.
- (171) <http://www.mbi.org/>.
- (172) <http://abpdu.lbl.gov/>.
- (173) Appleton, E.; Densmore, D.; Madsen, C.; Roehner, N. *Curr. Opin. Chem. Biol.* **2017**, *40*, 111–118.
- (174) Annaluru, N.; Muller, H.; Mitchell, L. A.; Ramalingam, S.; Stracquadanio, G.; Richardson, S. M.; Dymond, J. S.; Kuang, Z.; Scheifele, L. Z.; Cooper, E. M.; Cai, Y.; Zeller, K.; Agmon, N.; Han, J. S.; Hadjithomas, M.; Tullman, J.; Caravelli, K.; Cirelli, K.; Guo, Z.; London, V.; Yeluru, A.; Murugan, S.; Kandavelou, K.; Agier, N.; Fischer, G.; Yang, K.; Martin, J. A.; Bilgel, M.; Bohutski, P.; Boulter, K. M.; Capaldo, B. J.; Chang, J.; Charoen, K.; Choi, W. J.; Deng, P.; DiCarlo, J. E.; Doong, J.; Dunn, J.; Feinberg, J. I.; Fernandez, C.; Floria, C. E.; Gladowski, D.; Hadidi, P.; Ishizuka, I.; Jabbari, J.; Lau, C. Y.; Lee, P. A.; Li, S.; Lin, D.; Linder, M. E.; Ling, J.; Liu, J.; Liu, J.; London, M.; Ma, H.; Mao, J.; McDade, J. E.; McMillan, A.; Moore, A. M.; Oh, W. C.; Ouyang, Y.; Patel, R.; Paul, M.; Paulsen, L. C.; Qiu, J.; Rhee, A.; Rubashkin, M. G.; Soh, I. Y.; Sotuyo, N. E.; Srinivas, V.; Suarez, A.; Wong, A.; Wong, R.; Xie, W. R.; Xu, Y.; Yu, A. T.; Koszul, R.; Bader, J. S.; Boeke, J. D.; Chandrasegaran, S. *Science* **2014**, *344*, 55–58.
- (175) Boeke, J. D.; Church, G.; Hessel, A.; Kelley, N. J.; Arkin, A.; Cai, Y.; Carlson, R.; Chakravarti, A.; Cornish, V. W.; Holt, L.; Isaacs, F. J.; Kuiken, T.; Lajoie, M.; Lessor, T.; Lunshof, J.; Maurano, M. T.; Mitchell, L. A.; Rine, J.; Rosser, S.; Sanjana, N. E.; Silver, P. A.; Valle, D.; Wang, H.; Way, J. C.; Yang, L. *Science* **2016**, *353*, 126–127.
- (176) Hwang, I. Y.; Koh, E.; Kim, H. R.; Yew, W. S.; Chang, M. W. *Drug Resist. Updates* **2016**, *27*, 59–71.
- (177) Fox, J. L. *Nat. Biotechnol.* **2015**, *33*, 122.
- (178) Riglar, D. T.; Giessen, T. W.; Baym, M.; Kerns, S. J.; Niederhuber, M. J.; Bronson, R. T.; Kotula, J. W.; Gerber, G. K.; Way, J. C.; Silver, P. A. *Nat. Biotechnol.* **2017**, *35*, 653–658.
- (179) Brune, K. D.; Bayer, T. S. *Front. Microbiol.* **2012**, *3*, 3.
- (180) Stratis-Cullum, D. N. *J. Anal. Bioanal. Tech.* **2013**, *S7*, 1–4.
- (181) <http://universalbiomining.com/>.
- (182) Bernhardt, P.; McCoy, E.; O'Connor, S. E. *Chem. Biol.* **2007**, *14*, 888–897.
- (183) Yan, X.; Ge, H.; Huang, T.; Hindra; Yang, D.; Teng, Q.; Crnovcic, I.; Li, X.; Rudolf, J. D.; Lohman, J. R.; Gansemans, Y.; Zhu, X.; Huang, Y.; Zhao, L. X.; Jiang, Y.; Van Nieuwerburgh, F.; Rader, C.; Duan, Y.; Shen, B. *mBio* **2016**, *7*, 7.
- (184) Woodruff, L. B.; Gorochofski, T. E.; Roehner, N.; Mikkelsen, T. S.; Densmore, D.; Gordon, D. B.; Nicol, R.; Voigt, C. A. *Nucleic Acids Res.* **2017**, *45*, 1553–1565.
- (185) Ghodasara, A.; Voigt, C. A. *Nucleic Acids Res.* **2017**, *45*, 8116–8127.
- (186) Dudley, Q. M.; Anderson, K. C.; Jewett, M. C. *ACS Synth. Biol.* **2016**, *5*, 1578–1588.