ACS SyntheticBiology

Editorial

pubs.acs.org/synthbio

Meeting Report for Synthetic Biology for Natural Products 2017: The Interface of (Meta)Genomics, Machine Learning, and Natural Product Discovery

Natural product compounds isolated from plants, animals, and microorganisms have had a society-changing impact in medicine, agriculture, and industry. With new information provided by two decades of genome sequencing efforts and new tools and methods for high-throughput precision genetic engineering, natural product discovery and optimization has been revitalized. Sixty researchers from diverse fields came to Cancun, Mexico over 4 days in early March, 2017 at the Synthetic Biology of Natural Products Conference to talk science, debate, and network all while taking some time to enjoy the sun and sand on Playa Delfines. The presentations covered topics that included new computational and genetic tools for genome mining, detailed biochemical understanding of protein−protein interactions and enzyme mechanisms, and advanced heterologous expression platforms for natural product biosynthesis. This report summarizes the talks from the meeting and highlights the themes that emerged.

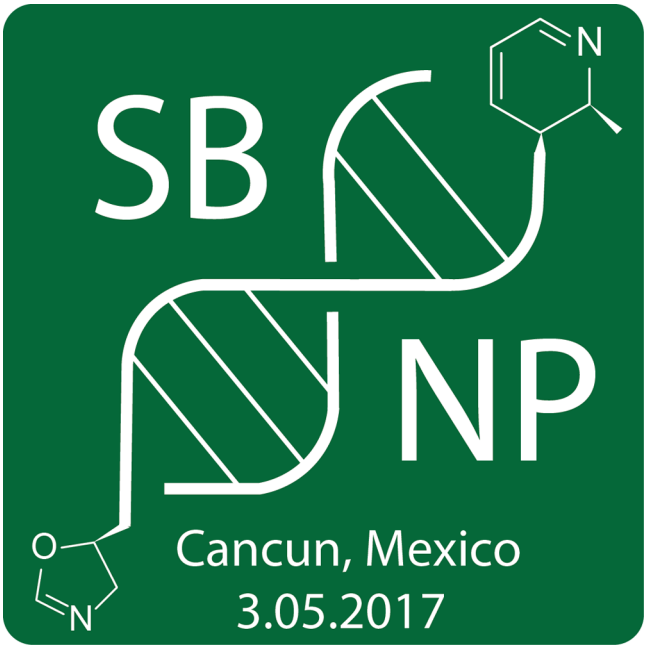

Figure 1. Synthetic Biology for Natural Product Conference, organized by Fusion Conferences Limited, was held at the Fiesta Americana Condesa resort in Cancun, Mexico in early March 2017.

GENE AND PATHWAY DISCOVERY FROM NEW HOSTS OR ENVIRONMENTS

The dramatic cost reduction in DNA sequencing over the past two decades is enabling ambitious sequence-based discovery efforts. Sequencing capabilities have continued to improve to the point where groups have proposed sequencing the 10^{31} basepair metagenome of the entire planet. Sequence information gathered from microbes and plants have dramatically changed the approach we take to discovering natural product biosynthetic genes and pathways, and this was on display during many talks at SBNP17.

Bill Metcalf (University of Illinois at Urbana−Champaign) described collaborative efforts of multiple groups at the University of Illinois and Northwestern to discover new phosphonate natural products by performing sequence-based screens of 10 000 Actinomycetes and full genome sequencing of over 600 strains. They identified biosynthetic gene clusters (BGCs) for the majority of phosphonates produced by the genus *Streptomyces* and were able to isolate and determine the chemical structure of a dozen new compounds, providing a powerful illustration of how genome sequence information can focus discovery efforts to improve success rates compared to the traditional grind-and-find approach. Metcalf posed important questions for the field moving forward. For example, what should the role of Synthetic Biology be in natural products discovery? And what do we need to do to discover every natural product in *Streptomyces*? These ended up being reoccurring themes throughout the conference.

Despite having large and difficult-to-annotate genomes, fungal systems are proving rich sources for natural product discovery and engineering efforts. Colin Harvey (Stanford University) described a collaborative effort to discover new fungal metabolites using engineered *S. cerevisiae* strains as a synthetic biology platform. They found that the design of expression constructs, namely the choice of promoters, routinely made the difference between seeing no natural product at all or obtaining isolable yields of new fungal terpenes and polyketides. A combination of newly characterized promoter elements and strain engineering has given the group an impressive success rate, identifying 17 natural products from their first 41 target BGCs. Claudia Schmidt-Dannert (University of Minnesota Twin Cities) presented work that delves deeper down the fungal branch of the tree of life into the Basidiomycete fungi. Previously under sampled in genome-mining efforts in part due to a lack of genetic tools, Schmidt-Dannert and colleagues are mining many new sesquiterpene BGCs from these mushrooms. With initial work focused around known biologically active scaffolds like protoilludene and hirsutene, their ability to predict terpene structure based on phylogeny of synthase genes is helping to organize the vast genetic potential in this group of organisms. Knowledge gained through their efforts in Basidiomycetes has allowed them to overcome unpredictable splicing events in these fungi to begin applying direct DNA synthesis for functional expression in a model *S. cerevisiae* host. Both the

Received: April 24, 2017
Published: May 19, 2017

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

Harvey and Schmidt-Dannert talks highlighted the importance of taxonomy-guided discovery efforts, with different phyla of fungi showing distinct patterns of natural product biosynthetic potential.

In plants, the growing appreciation of colocalization of on-pathway biosynthetic genes has dramatically accelerated the identification and characterization of full biosynthetic pathways, as highlighted by Anne Osbourn's (John Innes Centre) talk. Her group has taken advantage of this colocalization to discover the genes needed for avenacin A1 production, and are making great progress toward moving this biosynthetic capability from the roots of oats into wheat. Aiding this effort is the discovery that tissue-specific promoters from oat function the same in other plant species. Further, new insights that Osbourn and colleagues have made regarding the epigenetic modifications that demark plant gene cluster boundaries is accelerating the identification of putative BGCs in new genome sequences.

When biosynthetic genes are not colocalized into a BGC, RNA sequencing can be used to focus gene discovery efforts by grouping genes by expression patterns. This technique has been successfully applied by Elizabeth Slattery and her group at Stanford University. Slattery and colleagues filtered thousands of annotated genes by predicted enzyme function as well as by transcription patterns to discover new chemistry from the etoposide, colchicine, and other plant biosynthetic pathways. Importantly, their efforts to produce podophyllotoxin, a key intermediate in the synthesis of etopiside, have improved yields 300-fold and provide an example of how synthetic biology can help ameliorate supply concerns for drugs derived from plant natural products. Joerg Bohlmann (University of British Columbia) and colleagues also relied on RNA sequencing to discover genes involved in plant secondary metabolism, and with surprising results. After surgically removing wood cores from decades old sandalwood trees from a plantation in northern Australia, they discovered transcripts for santalol biosynthesis in the heartwood of mature sandalwood trees, a tissue previously believed to be dead. This provided the last piece of the puzzle, allowing for the stereospecific conversion of santalenes to the commercially relevant Z-santalols.

Several talks at SBNP17 described bypassing cultivation altogether by mining metagenome sequences for new biosynthetic genes and BGCs. Mark Liles (Auburn University) presented his group's efforts to combine BAC cloning of environmental DNA with both next-generation sequencing and direct functional screening. Serendipity led his group to identify and characterize a number of chloramphenicol-modifying enzymes that derivatized the selection agent to kill a Cm^R target strain. This example of "combinatorial metagenomics" could provide a new route toward extending the utility of existing drug scaffolds. Later, Liles demonstrated the advantages of sequencing BAC libraries in identifying diverse biosynthetic genes that would be missed using PCR screening approaches. This talk was supplemented by a poster on the project presented by David Mead, from Varigen Biosciences, who described more details on the extent of sequence diversity captured *via* this approach. Another poster, presented by Dragana Dobrijevic, from University College London, discussed the potential of environmental metagenomics for enzyme discovery, with a focus on transaminases and reductases.

Two speakers provided a perspective on metagenomics-aided discovery from an industrial point of view. David Pompliano (Lodo Therapeutics, Inc.), presented some fascinating data on the biogeographical distribution of biosynthetic potential in soil samples. Certain classes of BGCs are enriched in specific geographic areas, for example nystatins in the desert Southwest or streptolydigins in New England soils. Despite this unequal distribution, deep sampling efforts reinforced the Baas-Becking hypothesis that "everything is everywhere, but the environment selects"; Lodo's founders were able to find BGCs for most major classes of natural product drugs in the parks of New York City. Fortunately, this was not appreciated during the Golden Ages of natural product discovery; mammalian Target of Newyorkcitimycin does not have the same ring. Charles Moore (Novartis), the lone representative from "Big Pharma", described his group's efforts to mine soil metagenome sequences using next-generation sequencing. Their efforts to assemble large contigs from soil metagenome sequencing data underscores the tremendous complexity of soil microbial communities and is pushing the limits of computational resources at Novartis. The high success rates for heterologous production of known compounds and "silent" (or perhaps we can agree to call them "quiet") BGCs achieved by the team at Novartis bodes well for future work to find new molecules directly from environmental DNA.

In other efforts to mine metagenomic sequence data for new natural products, Chris Voigt (Massachusetts Institute of Technology) shared the beginnings of an ambitious effort to produce every specialized metabolite encoded in the human microbiome. In a tour-deforce effort, they have mined a family of small NRPS-containing BGCs to produce several dozen pyrazinones and dihydropyrazinones, many of which are acylated and most of which are new compounds. Included in the new compounds are potent proteasome inhibitors and GPCR ligands, suggesting that these molecules play an important role in modulating human physiology. Alex Wentzel (SINTEF) presented his group's multifaceted approach to finding new chemistry from the Trondheim Fjord, which includes metagenome sequencing, isolation of new marine Actinomycetes, and engineering *S. coelicolor* to produce a "superhost" for heterologous production.

Lastly, when the producing organism is unknown, all is not lost. Michael Smanski from the University of Minnesota Twin Cities presented his group's efforts to develop a biosynthetic route to a potent diterpenoid neuroprotective compound of unknown origin. His group was able to sift through GenBank for enzymes that could be linked together in a synthetic metabolic pathway that yields an advanced intermediate, from which a three-step chemical synthesis affords the target molecule. Smanski presented data to demonstrate the nuanced role that controlling gene expression plays in dictating not only titers but also the fidelity of a biosynthetic pathway. The ability to simultaneously control expression levels across each of the 12 genes in their refactored gene cluster is an example of how tools from synthetic biology are impacting natural product biosynthesis.

MECHANISMS OF NATURAL PRODUCT BIOSYNTHESIS

None of the engineering efforts currently underway would be possible without a solid mechanistic understanding of natural product biosynthesis. Despite several decades of probing natural product biosynthetic pathways with genetic, biochemical, chemical, and structural tools, there remain several unanswered questions in the field. A series of talks provided new insights into mechanisms of biosynthesis and demon-

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

strated that these systems remain a rich source to answer fundamental questions in enzymology.

David Cane (Brown University) presented his group's approach to breaking down the details of α-methyl epimerization during polyketide biosynthesis. Using a creative chemoenzymatic approach to generate β-keto substrates of defined stereochemistry and equilibrium isotope exchange to track the reactions, Cane and colleagues could investigate the roles of various ketoreductase domains from the DEBs gene cluster in directing the stereochemistry of the α-methyl group. They show that "dead" ketoreductase domains, while no longer capable of redox chemistry, can still invert the stereochemistry at the α-carbon. Surprisingly, the amino acid residues responsible for this inversion are in the same active site as those required for the redox chemistry. The nature of the acid/base involved in catalysis remains elusive and provides fodder for future investigations.

The importance of protein−protein interactions in PKSs and NRPSs have thwarted many attempts at reprogramming megasynthases to make designer chemicals. Michael Burkart (University of California San Diego) and his colleagues are dissecting these protein−protein interactions using a combination of chemical probes and protein NMR. Carrier proteins must interact with many different active sites during the catalytic cycles of FAS, PKS, and NRPS enzymes. Producing "crypto carrier proteins" labeled with mechanism-based inhibitors of processing domains allowed the group to cross-link the domains with their interactions preserved. This technology identified surfaces involved in acyl carrier protein:dehydratase domain interactions and gave a snapshot of individual salt-bridges that stabilize the complex. Replacing the mechanism-based inhibitors with substrate mimics provided evidence that the growing polyketide scaffold is sequestered in the ACP during Type II polyketide biosynthesis. These insights will enable improved structure-guided engineering efforts.

Michael Thomas from the University of Wisconsin−Madison provided a detailed account of MbtH-like proteins (MLPs) and their dual roles in promoting NRPS solubility and in activating amino acids by adenylation (A) domains. The complexity in MLP biochemistry is perplexing: a particular MLP may activate one adenylation domain and inhibit another, sometimes improve the affinity of A domains for substrate amino acids, and sometimes improve NRPS solubility, although none of these activities fully explain the overall phenotypes. While the new appreciation of MLP influence on NRPS enzymology is sobering in that it make require decades of measurements to be reexamined, the hope of finding a "universal" MLP akin to the phosphopantetheinyltransferase Sfp would be a boon for NRPS engineering.

In the same theme of upending conventional understanding of biosynthetic mechanisms, Larry Wackett (University of Minnesota Twin Cities) corrected the accepted mechanism of olefin biosynthesis by describing function of a conserved β-lactone synthase, OleC. With the insight that olefin production proceeds through a transient β-lactone intermediate, Wackett and colleagues looked to the BGCs of β-lactone natural products, for example lipstatin and ebilactone, for more examples. More often than not, OleC homologues could be identified, setting the stage for combinatorial biosynthesis of new β-lactone natural products. Additionally, Wackett provided evidence for metabolic compartmentalization into microspheres of enzyme and product during olefin biosynthesis, potentially provided a mechanism for specificity in organisms that coproduce olefins and β-lactone natural products.

Jean-Etienne Bassard, from the University of Copenhagen, presented his work on a more ephemeral example of metabolic compartmentalization. Unlike the microcompartments of olefin production, which are stable enough to view by electron microscopy, the transient association of membrane-bound P450s and a soluble glycosyltransferease required advanced analytical technologies for detection. Bassard described the application of Fluorescent Lifetime Imaging Microscopy to track Förster resonance energy transfer between the cytosolic and membrane-bound components *in planta*. The observation of "metabolons" that continually form and dissociate during the biosynthesis of the Sorghum natural product, dhurrin, points to the intricate and previously unappreciated organization of biosynthesis within cells.

COMBINATORIAL BIOSYNTHESIS AND PATHWAY ENGINEERING

Leveraging the mechanistic understanding of natural product enzymology to redirect biosynthesis promises to enable the engineered production of molecules with desirable properties for commercialization or downstream chemical modification. Several talks at SBNP17 describe approaches and applications for engineering biosynthesis that draw from advances in both our understanding of the underlying biochemistry and from the new tools and approaches from synthetic biology.

Gavin Williams (North Carolina State University) and his pair of poster award-winning graduate students presented their work focused on diversifying complex polyketides. They began by engineering a malonyl-CoA synthetase to generate an enzyme capably of quantitative yields of non-natural extender units. Using a promiscuous *trans*-acyltransferase domain from the kirromycin biosynthetic pathway, they were able to incorporate these non-natural extender units into complex polyketide scaffolds. Incorporation of a propargyl side chain enabled facile derivatization *via* "click-chemistry". By coupling their unique methods for structure diversification with sensitive biosensors also developed by the Williams lab, a coherent directed evolution workflow is envisioned that will enable massively high-throughput screening of permuted genetic constructs. Rebecca Goss (University of St. Andrews) used precursor-directed biosynthesis of the pacidamycin scaffold to produce several aryl halide derivatives. Similar to the propargyl incorporation by Williams, these aryl halides provided a handle for downstream medicinal chemistry. Suzuki−Miyaura cross-coupling reactions could be used with high effiicency to add more aromatic rings to the pacidamycin scaffold.

Taking an alternative approach to polyketide diversification, Jay Keasling from the University of California Berkeley described a platform for directing biosynthesis of small di- and triketides that would be useful compounds in the commodity chemical arena. Such molecules could serve as drop-in oxygenates for gasoline, monomers that will tune the functionality of nylons or other polymers, and more. By exploring the design constraints on domain swapping experiments, Keasling and colleagues better defined the recombination boundaries for functional acyltransferase and ketoreductase domains. By a chance discovery, they learned that simply leaving out NADH from their KR-containing PKS reactions led to spontaneous decarboxylation of diverse diketides to provide a new route toward ketone production.

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

Daisuke Umeno (Chiba University) presented an innovative example that combined protein evolution and pathway engineering to produce long-chain "super"-carotenoid molecules. Umeno and his group used protein engineering strategies to create a farnesyl diphosphate synthase that produces C25 and C30 products as well as a promiscuous carotenoid synthase capable of condensing C15−C25 substrates. By combining these two enzymes in a pathway, the group effectively engineered a biosynthetic band-pass filter to provide specific production of C50 carotenoids. This work provides a roadmap for engineering specificity as an emergent property of pathways built with promiscuous enzymes.

EXPANDING HETEROLOGOUS PRODUCTION SYSTEMS TO DIVERSE HOSTS

A common theme throughout the meeting was the use of heterologous production systems to access new chemistry, facilitate genetic interrogation of pathways, or establish commercially viable production systems. The hurdles to successful heterologous production are numerous and include balancing expression levels, ensuring proper protein folding, providing cofactors, and boosting precursor supplies. Each of these were addressed during the conference. Whether as an academic pursuit or as a practical means to lower cost-of-goods, the heterologous production systems described at SBNP17 show great promise for future natural product discovery and engineering efforts.

Blaine Pfeifer (University at Buffalo) answered the question of whether it was possible to functionally reconstitute a complex Actinomycete megasynthase in *E. coli* by describing his group's progress with the erythromycin type I PKS gene cluster. This required the functional expression and tuning of 23 foreign and 26 total genes together in one cell and was aided by coexpression of chaperones and low temperature growth to ensure proper protein folding. Current work to diversify the sugars has led to the identification of derivatives active against Ery^R reporter strains.

S. cerevisiae has likewise proven itself an effective heterologous production host for a variety of metabolites. As described above, Harvey and Schmidt-Dannert highlighted its utility for new natural product discovery and characterization from the fungal kingdom. Others are using yeast for trans-kingdom expression of biosynthetic pathways. Kristy Hawkins (Antheia) told us about her team's successful effort to port the entire pathway for benzylisoquinoline alkaloids, including clinical drugs like morphine, into a yeast microbial production platform. Metabolic engineering and enzyme improvement are getting them closer toward their goal of reducing the cost of goods 10-fold and adding security into the production of opiates. In the area of flavors and fragrances, Laurent Daviet, from Firmenich, described the importance of terpenoids as specialty chemicals in the production of perfumes, cosmetics, and foods. Mostly derived from plants, the terpenoids of the flavor and fragrance industry are prone to volatile supply chains, fluctuations in price, and often unsustainable harvest. Daviet described his company's efforts to ameliorate these supply issues by moving the terpenoid biosynthetic pathways into a microbial production source. Their success with bringing microbial-produced patchouli oil and ambergris to market underscores the power of this synthetic biology approach.

Instead of transplanting pathways from around the tree of life into advanced model systems like *E. coli* or *S. cerevisiae,* Andriy Luzhetskyy and his team at Saarland University are using modern genome engineering techniques to optimize *S. albidoflavus* J1074 (formerly *S. albus* J1074) as a host for expressing Actinomycete BGCs. They systematically knocked out gene clusters from J1074 to provide a clean background for detecting new chemistry from introduced pathways. Each of the 14 clusters that was knocked out could be empirically connected with compounds *via* LC−MS, raising the question as to whether "silent" gene clusters really exist. A surprising number of natural product gene clusters decreased fitness of the host when knocked out and therefor likely serve important roles in growth or reproduction that are poorly understood. Aside from this "*tabula rasa*" strain of *S. albidoflavus*, Luzhetskyy and colleagues have developed useful tools for *Streptomyces* engineering, including nonleaky induction systems and strains with up to four φC31 insertion sites for increasing the copy number of heterologous pathway genes.

Lastly, nearly all of the plant scientists who presented, including Bassard, Sattely, and Osbourn, noted the versatility of *Nicotiana benthamiana*, a wild tobacco relative, as a heterologous host for investigating plant specialized metabolism. The subcellular localization and multiprotein complexes required for plant natural product biosynthesis are maintained in this host, and a growing number of genetic tools enable rapid pathway characterization. For example, taking advantage of the agrobacterium-mediated HyperTrans expression system, pathway genes could be mixed in a combinatorial manner to produce gram-scale quantities of complex natural products, as Osbourn demonstrated with diverse triterpene pathways. A poster was presented by Sarah D'Adamo (Algenuity) describing her use of a unicellular brown algae, *Phaeodactylum tricornutum*, as a versatile synthetic biology platform for the production of plant triterpene sapogenins.

REDESIGNING THE GENETICS OF NATURAL PRODUCT BIOSYNTHESIS

One of the exciting contributions of synthetic biology in the past decade has been a new suite of DNA synthesis and assembly methods that give genetic engineers freedom to control any aspect of genetic design. We saw several examples on display during the four day conference, from codon optimization to wholesale pathway refactoring. Driven by the dramatic reduction in costs for DNA synthesis over the past decade, the scale of projects that rely on "designer genetic constructs" has skyrocketed, both in industry and academia.

Genetic refactoring is an approach wherein the underlying genetics of a system are completely rewritten toward an objective goal, for example to make a BGC more amenable to high-throughput engineering methods. Huimin Zhao (University of Illinois at Urbana−Champaign), described his group's application of genetic refactoring to discover a new polycyclic tetramate macrolactam from *S. griseus*. Much of the workflow required to design, construct, verify and assay refactored gene clusters has been automated on a standalone robotic platform named iBioFAB, giving graduate students everywhere hope that cloning may soon be a thing of the past. Similarly, Voigt and colleagues at the MIT-Broad Foundry integrate the principles of genetic refactoring into a series of "apps" that have been impedance-matched to allow seamless integration in custom-designed workflows. The speed and agility of the Foundry toward engineering the production of diverse natural product scaffolds was highlighted when Voigt described a recent "pressure test" applied by their funders at DARPA. Within 90 days of receiving a list of target molecules, the group showed

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

substantial progress toward such diverse scaffolds as 1-hexadecanol, vincristine, pacidimicin, rebeccamycin, and carvone, in some cases generating new strains that produce the targets in higher titers than previously reported. In a more focused effort, Smanski demonstrated how the high-throughput DNA assembly enabled in refactored systems can be applied to rationally explore combinations of gene expression in a multigene system as his group optimizes the methylerythritol phosphate pathway in *Streptomyces*.

In a dramatic example of pathway refactoring, Christopher Boddy (University of Ottawa) presented the successful biosynthesis of violacein from a gene cluster redesigned as a single polyprotein, akin to how many viral genomes are organized. Each of the violacein biosynthetic enzymes were expressed as a translational fusion protein with TEV protease tags separating the functional enzymes. Coexpression of the polypeptide with TEV protease in *E. coli* resulted in cleavage into active enzymes and detectable violacein production. The advantage of such a system is that transfer to a new host only requires replacing a single promoter and ribosome binding site upstream of the polyprotein CDS. Moore reported a similar strategy pursued by Novartis that uses self-cleaving 2A peptides in place of TEV protease sites.

Several presenters described methods to mix, match, and optimize biosynthetic pathways without the need to build multigene plasmids. Boddy and Mark Burk from Genomatica each presented complementary approaches to reconstitute biosynthetic pathways using cell-free protein production systems. In such systems, biosynthetic genes can be added to the reaction mixture on discrete plasmids or as linear DNA fragments with gene expression levels controlled by adjusting the concentrations of DNA in the mixture. Cell-free production systems dramatically reduce the time of the standard design−build−test−analyze−learn cycle from days to hours and could be useful in optimizing expression levels or the selection of genes. Combinatorial mixing of discrete biosynthetic plasmids was reported for cellular expression systems as well. Whole pathways can be reconstituted using the aforementioned HyperTrans tobacco expression by coinfiltrating a half-dozen or more different *Agrobacterium* strains. Similarly, Harvey presented an approach wherein a- or α-mating haplotypes of *S. cerevisiae* containing complementary subpathways can be mated in a combinatorial manner to generate unique diploid production strains.

METABOLIC ENGINEERING FOR IMPROVED PRECURSOR SUPPLY

Engineering primary metabolism to channel flux toward precursor molecules that support specialized metabolism is a proven approach for increasing titers. Wolfgang Wohlleben from the University of Tübingen shared three different stories around the central theme of precursor supply, each targeting a different major class of natural compounds. Wohlleben and colleagues improved production titers of the polyketide tacrolimus by boosting lysine (a precursor to the pipecolic acid starter unit) through introduction of a deregulated aspartate kinase from *Corynebacterium*. For the diterpeniod brasilicardin, the group focused on inserting the mevalonate pathway, but interestingly saw the best titers when the first enzyme of this pathway was left out of their constructs. Lastly and most surprisingly, the solution to overproduce glycopeptide balhimycin came through overexpression of an off-pathway gene whose feedback inhibition diverted flux toward the balhimycin precursors. Sonal Ayakar (University of British Columbia) presented a poster describing an alternative approach for increasing flux through precursor pathways by mining metagenome sequences for orthologs of rate-determining enzymatic steps.

Jason Sello (Brown University) presented forward-looking research that would enable coupling natural product biosynthesis with lignocellulose degradation in an attempt to valorize agricultural waste. Sello's team has sequenced the genome of the first reported Actinobacteria capable of degrading plant biomass and have used this information to describe the catabolism of lignin into primary metabolites from a genomic level. This work led to the discovery and characterization of new inducible promoter that responds to aromatic compounds to overexpress target genes, adding to the genetic toolkit for controlling gene expression in *Streptomyces*.

The two previous examples use genomic information to improve precursor levels, but for some natural products it is not obvious which primary metabolic pathways are responsible for supplying precursors. This is particularly relevant when precursor molecules can in theory come from any of several redundant primary metabolic pathways. In these cases, understanding the contribution of each pathway to supporting specialized metabolism is the first step toward rational engineering of titer improvements. Ina Häuslein (Technical University of Munich) shared several examples of an advanced analytical technique called isotopologue profiling that reveals the biosynthetic route through primary metabolism for each atom in the final product. Häuslein and her colleagues are able to generate this data for diverse organisms, from microbes in fermenters to plants growing in the field, by coupling creative methods for administering the isotopically labeled substrate to sensitive mass-spectroscopy or NMR-based analyses.

NEW TOOLS

New tools and methodologies that help with sequence analysis and annotation, physical construction of recombinant DNA plasmids, rational control of gene expression, or analytical chemistry and compound dereplication are being introduced at a dizzying rate in this genomics era of natural products research. We got a chance to hear about a number of new technologies at the SBNP17, both from industry and academia.

A frequent theme throughout the meeting was new computational tools that will be added to the software antiSMASH 4.0 (Medema). It is exciting to see the whole field coming together to support the vision of Marnix Medema to produce a unified community resource to accelerate the identification and characterization of BGCs from sequence information. Some of these tools improve BGC identification in eukaryotic organisms, including plants and fungi. Others aim to more accurately compare BGCs to identify families that produce structurally related natural products. Marc Chevrette, coming from a group of evolutionary biologists at the University of Wisconsin−Madison, described a computational tool that predicts the specificity of adenylation domains from NRPSs by comparing discrete subdomains and identifying monophyletic clades. The computational pipeline has been compiled into a package called SANDPUMA and will be available for predicting structures produced from terpene cyclases as well. Metcalf described an alternative strategy that weights multiple scoring parameters based on active site conservation, gene synteny, and conservation of gene content to generate an overall distance metric for comparing BGCs.

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

When each of these tools are let loose on available genome sequence data, they agree that natural product discovery to date represents only the tip of the iceberg, with many families left uncharacterized. A third computational pipeline described takes the normal genome-centric approach for BGC annotation and turns it upside down. Doug Mitchell (University of Illinois at Urbana−Champaign) presented his group's new tool called RODEO (Rapid ORF Description and Evaluation Online) that probes sequence databases to pull out a conserved BGC family of interest. Especially suited for mining RIPP gene clusters, this tool reannotates target loci by translating each putative reading frame to identify the small precursor peptides often missed during genome annotation. After describing the tool, Mitchell described how his group has applied it to expand the number of known families of lasso peptides. Each of the new software tools described above are implementing aspects of machine learning to improve the accuracy of predictions. Lastly, the Genetic Constructor is a new computer-aided drafting (CAD) tool for genetic engineering that was described by Cornelia Scheitz from Autodesk. The tool is a significant update to sequence manipulation tools that allows genetic engineers to operate at a higher level of abstraction and away from the tedium of looking at A's, T's, C's and G's.

The new tools presented at SBNP17 were not limited to computational biology. A critical, but often rate-limiting, step in heterologous production projects is capturing or assembling whole biosynthetic gene clusters into a mobile genetic element. Chengcang Wu (Intact Genomics) told us about his improved methods for generating large-insert bacterial and fungal artificial chromosomes (BACs and FACs). Routinely obtaining insert sizes greater than 100 kilobases, even large PKS and NRPS clusters can be captured on a single plasmid. A proof of concept experiment in the fungus *Aspergillus terreus* was able to capture three distinct but closely colocalized BGCs on a single insert, leading to the heterologous production of a known benzodiazepine, an ophiobolin-family terpenoid, and a novel lipopeptide from the same construct. Zhen Kang (Jiangnan University) skyped into the conference to tell us about his DATEL method for assembly of large DNA constructs. This system is both scarless and sequence-independent and uses thermostable exonucleases and ligase to join 2−10 DNA fragments with high efficiency. Kang also described his groups RECODE tool for combinatorial mutagenesis, which was used to improve enzymes in an engineered carotenoid pathway. Lastly, Huimin Zhao presented a novel approach for site-specific, endonucleic cleavage of DNA using Argonaute proteins from thermophilic archeon, *Pyrococcus furiosus*. The enzyme *Pf*Ago uses a single-stranded DNA guide to direct it to a complementary single-stranded DNA target generated by heat denaturation. Because *Pf*Ago cuts at a predictable position, two enzymes directed to opposite strands of target locus produce overhanging sticky ends of a user-defined size. These tools combine some of the best attributes of cut-and-paste cloning methods with the versatility of restriction enzyme-independent cloning methods.

We heard many times throughout the conference how precise control of gene expression levels can make or break natural product engineering efforts. Small molecule induction systems, such as Sello's PobR regulator and Luzhetskyy's molecular AND gate mentioned above, provide new options to titrate gene expression in Actinomycetes. Laura Motta-Mena (Optologix, Inc.) presented a new tool for rapid, robust, and dose-dependent gene activation using light-oxygen voltage (LOV) domains from bacteria. With LITE Switch, the first commercially available light-controlled expression kit, Optologix has demonstrated their light-activated promoters in organisms as complex as developing zebrafish. Jose Avalos (Princeton University) has applied the Optologix light induction system to control metabolism. By combining the LOV-domain sensors with genetic inverter circuits, Avalos and his team are able to simultaneously induce and/or repress gene expression in different branches of metabolism to guide flux toward target molecules and reduce the accumulation of toxic shunt products. Further, they have leveraged the light induction system to differentially control metabolism from growth phase (to reach high cell densities) to production phase (to maximize titer).

Brian Pfleger (University of Wisconsin−Madison) provided the lone cyanobacterial talk of the conference. The suite of cis-acting regulatory elements his group has developed from controlling expression in *Synechoccocus* sp. PCC 7002 will accelerate engineering and discovery from cyanobacteria by enabling the type of refactoring approaches mentioned earlier for Actinobacteria, *E. coli*, and yeast. In work that directly transfers to other organisms, Pfleger presented a creative approach toward gaining inducible control over any promoter in the genome. First by adding cleavable TEV-protease sites to the middle of a TALE DNA-binding protein, and later by creating an internal fusion of an allosteric control domain, his group has demonstrated the ability to control the release of TALE proteins that were custom-designed to bind and repress a promoter of interest. TALE proteins have temporarily fallen out of favor because of the excitement around CRISPR/Cas systems. However, coupling new applications such as those presented by Pfleger with improvements in their production, for example the $3 per TALE protein cloning system enabled by Zhao's iBioFAB automation, will ensure their utility for years to come.

PUTTING IT ALL TOGETHER

During his talk, Chris Voigt described synthetic biology as the glue that connects the computational work surrounding BGC discovery from sequence information and the chemistry/biology required to characterize new natural products and determine their biological activities. We caught some inspiring glimpses into the potential of integrated workflows that couples high-throughput and multiplexed genetic engineering experiments with state-of-the-art analytics and data management. Patrick Westfall (Zymergen) and Derek Abbott (Amyris) both described platforms at their respective companies that aim to maximize the type and quality of data collected for every strain engineering project and drive empirical optimization on a scale not previously possible. Westfall shared examples of how Zymergen has applied their "radical empiricism" approach for discovery, optimization, and commercialization of products from microbial fermentation. Whether through developing analytics to predict fermenter performance from microtiter plate data or using combinatorial mutation analysis to make commercial production strains healthier, Westfall demonstrated the importance of a systematic and comprehensive approach to data collection and learning. Abbott gave several examples that prove the axiom, "necessity is the mother of invention". For example, to avoid bottlenecks experienced during LC/MS analysis of engineered strains, Abbott and colleagues developed a dual-column chromatography system to get analysis times down to 1.5 min per strain. Such integrated workflows are not

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743

only the product of industry. Voigt and Zhao described showed us that impedance-matched modular platforms for natural product discovery and engineering are possible in academic settings as well, for example at the MIT-Broad Foundry and iBioFAB, respectively.

CONCLUSIONS

During his talk, Charles Moore shared a poignant image taken from inside an industrial warehouse-sized laboratory at Pfizer during the height of their natural product discovery program in the early 1950's. Long benches covered with thousands of culture dishes were laid out before a cadre of labcoat-dressed technicians ready to pour agar media. The image provided a small glimpse of the monumental investment of resources and effort that drove the "Golden Age" of natural product drug discovery during the 1940's through 1960's. It serves a reminder that Golden Ages do not come about by luck or chance, but instead are the product of unified effort behind a shared vision. They are made, not happened-upon. All signs indicate that the discovery of natural products is primed for a second Golden Age. It is up to us as a community of researchers to make it so. The next Synthetic Biology for Natural Products conference is scheduled for 2019 in Puerto Vallarta, providing an excellent opportunity to see these shared advances.

Michael J. Smanski[*,†]
David Mead[‡]
Claes Gustafsson[§]
Michael G. Thomas[*,∥]

[†]BioTechnology Institute, and Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Twin-Cities, Saint Paul, Minnesota 55108, United States

[‡]Varigen Biosciences, Madison, Wisconsin 53719, United States

[§]ATUM, 37950 Central Court, Newark, California 94560, United States

[∥]Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706, United States

AUTHOR INFORMATION

Corresponding Authors

*E-mail: smanski@umn.edu.
*E-mail: michael.thomas@wisc.edu.

Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

ACKNOWLEDGMENTS

We would like to thank the staff at Fusion Conferences Limited for organizing and operating the conference, and ACS Synthetic Biology for financial support.

DOI: 10.1021/acssynbio.7b00132
ACS Synth. Biol. 2017, 6, 737−743