ABSTRACTS POSTERS

Laura Adam
Automatic translation of high-level bio-designs into wet-lab protocols through learning and formalizing expert's practices

Pol Arranz-Gibert
BBB shuttles: From design to application

M. Boone
Screening entire proteomes for yeast-expressable fragments

Costas Bouyioukos
GREAT: Genome REgulatory Architecture Tools

Baudoin Delépine & Vincent Libis
Expanding biosensing abilities through computer-assisted design of metabolic pathways

Jean Marie François
Ex nihilo design and in vivo implementation of Synthetic Metabolic Pathways for the Production of 2,4-Dihydroxybutyric Acid

Raik Grünberg
A field guide to automated cloning

Brian Jester1,
Context Sensitivity: The Blackbox of Synthetic Biology

Rodrigo Ledesma-Amar
Metabolic engineering strategy to produce oleochemicals from raw starch

Guillaume Le Treut
The different structures of the DNA

Ákos Nyerges
pMAGE: a highly precise, portable method for microbial genome-scale engineering

Marleen Renders
A method for the in vitro evolution of XNA ligases

Jorge Rodríguez Grande
Bacterial computation using Plasmids as Wires

Michael G. Sadovsky
Evolutionary orchestra: taxonomy and function in chloroplasts and mitochondria

Neil Swainston
Integrated informatics to support the SYNBIOCHEM Design-Build-Test cycle for production of fine and specialty chemicals

Antoine Vigouroux
Taking control on E. coli's genes expression using CRISPR interference

Torsten Waldminghaus
Synthetic secondary chromosomes to study chromosome maintenance in Escherichia coli

Automatic translation of high-level bio-designs into wet-lab protocols through learning and formalizing expert's practices

Laura Adam and Eric Klavins
Department of Electrical Engineering
University of Washington, Seattle WA
ladam@uw.edu

Lab automation services[1] are emerging around the world, making it possible for synthetic biologists to focus on designing, modeling and characterizing their synthetic biology systems. To facilitate the design step, several computer assisted bio-design tools[2], databases of genetic parts and data standards[3] have been developed. However, a formalization gap between the part-level bio-designs and the fabrication of corresponding organisms remains. The successful construction of a synthetic yeast strain requires mastering a variety of wet-lab procedures for which intermediate DNA-level designs are required.

Here, we introduce Apprentyeast, a Python library that handles the wet lab realization of high-level synthetic designs on behalf of the organism designers while learning the best practices from experts. The basic input is a synthetic yeast object defined as a parent synthetic yeast from which DNA fragments can be inserted or removed. Apprentyeast takes advantage of genomic databases[4,5], Coral[6] --a molecular biology DNA-level design Python package-- and Aquarium[7] --a human-in-the-loop lab automation system that includes a Laboratory Information Management System-- to design, or re-use, the necessary primers, fragments, plasmids and strains, and finally submit the corresponding jobs for automated construction and quality control through a web API. It also includes higher-level design functions, such as gene knockouts or tagging proteins. As an architect's apprentice, Apprentyeast learns the most commonly used construction strategies in the lab and preferentially uses them. We mined all the Gibson assemblies performed through Aquarium in order to infer rules, which are then used to automatically design and choose fragments for the plasmid backbones.

The successful use of Apprentyeast to create synthetic yeast strains in our wet-lab illustrates it is possible to translate high-level bio-designs into protocols. By formalizing and mining expert's local knowledge, synthetic biologists have access to the best practices on how to obtain a yeast strain implementing a synthetic biology design. Such a tool will lower the entry barrier to synthetic biology for individuals with little to no molecular biology experience by abstracting away the wet-lab realization aspects.

References:

1. Check Hayden, E. The automated lab. *Nature* **516,** 5–6 (2014).
2. Lux, M. W. M., Bramlett, B. W. B., Ball, D. a & Peccoud, J. Genetic Design Automation: engineering fantasy or scientific renewal? *Trends Biotechnol.* **30,** 120–6 (2012).
3. Galdzicki, M. *et al.* The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nat. Biotechnol.* **32,** 545–550 (2014).
4. Cherry, J., Adler, C., Ball, C. & Chervitz, S. SGD: Saccharomyces genome database. *Nucleic acids* **26,** (1998).
5. Chang, D. T.-H., Huang, C.-Y., Wu, C.-Y. & Wu, W.-S. YPA: an integrated repository of promoter features in Saccharomyces cerevisiae. *Nucleic Acids Res.* **39,** D647–D652 (2011).
6. Bolten, N. & Klavins, E. Towards a sequence-level DNA design specification language. in *7th INTERNATIONAL WORKSHOP ON BIO-DESIGN AUTOMATION* 82 (2015).
7. Klavins, E. *The Aquarium Project*. (2015). at <http://klavinslab.org/aquarium.html>

BBB shuttles: From design to application

Pol Arranz-Gibert (1), Meritxell Teixidó (1) and Ernest Giralt (1,2),

1-Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology ; 2-Department of Organic Chemistry, University of Barcelona

In recent decades, the efforts channelled into the development of peptides as promising molecular tools to study the chemistry and biology behind cellular metabolism have been greatly facilitated by the well-established chemical synthesis of these molecules. In addition, their natural abundance, as peptides or parts of proteins, has generated interest as therapeutics. Peptide blood-brain barrier (BBB) shuttles are a recent approach to overcome the BBB, the most important gate to the CNS. Transport to and from the brain is controlled by this selective barrier, which blocks the uptake of more than 98% of small molecules and almost 100% of the large-molecule drugs. These shuttles, a novel class of therapeutics, are short peptide sequences able to cross the BBB by diverse mechanisms while carrying other molecular entities. Those designed to cross by passive diffusion are suitable for small molecular weight therapeutics whereas active transport (endocytic mechanisms) allows the transport of supramolecular entities. Here we report on the design and study of peptide BBB shuttles that use passive and active transport. We also present a study case of the applicability of active transport shuttles for gene therapy to treat Friedreich's Ataxia, in which HSV-1 viral particles of around 200 nm are required to cross the BBB.

Screening entire proteomes for yeast-expressable fragments

Boone, M.(1,2), Callewaert N. (1,2)

1. Medical Biotechnology Center, VIB, Ghent, Belgium
2. Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

Despite major strides in heterologous protein expression in the last few decades, it is at present still largely unpredictable whether a given protein can be produced in a particular host. This unpredictability denotes that our knowledge of how the cellular machinerie exactly functions is incomplete. We decided to tackle this problem for the yeast secretory system by determining the full range of domain-sized protein fragments from the human proteome that can be secreted without substantial degradation. Our method integrates the construction of high-complexity cDNA fragment libraries, high-throughput sorting of secreting yeast clones, and deep sequencing as a readout. After significant optimization of several aspects of the platform, we here present data from our proof-of-concept screens in both Pichia pastoris as well as Saccharomyces cerevisiae, showing the feasibility of the approach and its potential role in improving characterization of secretory system components.

GREAT: Genome REgulatory Architecture Tools

Costas Bouyioukos[1], François Bucchini[1], Mohamed Elati[1] and François Képès[1,2]

[1]*institute of Systems and Synthetic Biology, CNRS, Genopole, UEVE, Evry, France.*
[2]*Department of Bioengineering, Imperial College, London, UK.*

Recent advances in genomics, transcriptomics and genome structural biology have revealed significant insights on the non-random arrangement of genes on one hand (Képès 2004) and on the interplay between transcription, gene position and genome structure on the other (Dekker et al., 2013; Dorman, 2013; Weng & Xiao, 2014).
Here we present an online software suite designed to perform a systematic and integrated analysis of regular patterns along genomes. The software is based on an algorithm to detect periodicities. It provides an easy to use interface to execute complicated analyses of regular patterns, rich visualisation of results and output open for further analyses and development.
The suite comprises two software tools. GREAT:SCAN:patterns, a pipeline for systematic study of periodic patterns, clustering and visualisation and GREAT:SCAN:precision, a novel transcription factor binding sites (TFBS) prediction machine learning tool.

GREAT:SCAN:patterns. The algorithm begins by exhaustively analysing all predicted periods and calculating weighted (exact) p-values. The first step outputs a rank of periods based on the exact p-value. On the second step a clustering algorithm detects clusters of genes that are in-phase on the modulo period coordinates and provides an insight of possible local spatial proximity of genes. On the third and last step a more fine-tuned search for regularities is taking place based on a variable size sliding window which detects regularities on specific domains of the chromosome.
In this work, we present a complete analysis of the 7 major TFs of *E. coli* and report initial evidences that regions of periodic arrangement are associated with the macro-domain organisation of the genome of this bacterium.
The software has been initially developed to detect periods on co-regulated genes however it can work with any gene set of interest as well as with any set of genomic positions of interest, including but not limited to chip-seq data.

GREAT:SCAN:precision is a novel implementation of a machine learning tool for TFBS prediction (Elati et al. 2013) which incorporates two different data source (views) in a classifier: a) direct DNA sequence motif readout and b) genome layout readout from the genomic coordinate. The underlying rationale is based on the emerging observation that co-regulated genes are positioned at periodic intervals along the chro-mosome. The combined classifier is then obtained with an iterative weight update scheme, using a modified version of the AdaBoost algorithm. We will report on the novel prediction of *E. coli* TFBS as well as insights on the interplay between sequence motif and position by studying their associations with spectral correlation methods.

Képès, F. *Periodic transcriptional organization of the E. coli genome*. Journal Molecular Biology 340, 957-964 **2004**

Dekker, J.; Marti-Renom, M. A. & Mirny, L. A. *Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data*. Nature Reviews. Genetics, 14, 390-403, **2013**

Dorman, C. J. *Genome architecture and global gene regulation in bacteria: making progress towards a unified model*. Nature Reviews. Microbiology, 11, 349-355, **2013**

Junier, I.; Hérisson, J. & Képès, F. *Periodic pattern detection in sparse boolean sequences*. Algorithms Molecular Biology, 5, 31, **2010**

Weng, X. & Xiao, J. *Spatial organization of transcription in bacterial cells*. Trends in Genetics, **2014**

Elati, M., Nicolle, R., Junier, I., Fernandez, D., Fekih, R., Font, J. and Képès, F. *PreCisIon: PREdiction of CIS-regulatory elements improved by gene's positION*. Nucleic Acids Research. 41, 1406-15 **2013**

Expanding biosensing abilities through computer-assisted design of metabolic pathways

Vincent Libis*, Baudoin Delépine*, Jean-Loup Faulon

Detection of chemical signals is critical for cells in nature as well as in synthetic biology where they serve as inputs for designer circuits. Important progress has been made in the design of signal processing circuits triggering complex biological behaviors but the range of small molecules recognized by sensors as inputs is limited. The ability to detect new molecules will increase the number of synthetic biology applications but direct engineering of tailor-made sensors takes time. Here we describe a way to immediately expand the range of biologically detectable molecules by systematically designing metabolic pathways that transform non-detectable molecules into molecules for which sensors already exist. We leveraged computer-assisted design to predict such sensing-enabling metabolic pathways and we built several new whole-cell biosensors for molecules such as cocaine, parathion, hippuric acid and nitroglycerin.

Ex nihilo design and in vivo implementation of Synthetic Metabolic Pathways for the Production of 2,4-Dihydroxybutyric Acid

Thomas Walther[1-4], Clément Auriol[1-4], Romain Irague[1-4], Christopher M. Topham[1-4], Clémentine Dressaire[1-4], Audrey Baylac[1-4], Florence Calvayrac[1-4], Hélène Serrano-Bataille[1-4], Julie Fredonnet[1-4], Hélène Cordier[1-4], Isabelle André[1-4], Marc Maestracci[5], Robert Huet[5], Magali Remaud-Simeon[1-4], Jean Marie François[1-4]

[1]Université de Toulouse; INSA, UPS, INP; LISBP, 135 Avenue de Rangueil, 31077 Toulouse, France ; [2]INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, Toulouse, France ; [3]CNRS, UMR5504, Toulouse, France ; [4]TWB, 3 rue des Satellites, Canal Biotech Building 2, 31400 Toulouse, France; [5]Adisseo SA, Antony Parc II, 10 place General de Gaulle, 92160 Antony, France.

2,4-Dihydroxybutyric acid (DHB) is a molecule with considerable potential as a versatile chemical synthon. Notably, it may serve as a precursor for chemical synthesis of the methionine analogue 2-hydroxy-4-(methylthio) butyrate, thus, targeting a considerable market in animal nutrition. However, petrochemical synthesis of DHB is not economically viable, and no natural metabolic pathways exist for the biochemical production of DHB. We have therefore conceived three synthetic metabolic pathways for the synthesis of DHB starting from the natural metabolites malate or homoserine. Two of them depart from malate via activation of malate by phosphorylation or acylation to yield malylphosphate or malyl-CoA, respectively. These compounds are then subject to two successive rounds of reduction to yield DHB via the common intermediate malate semialdehyde. The production of DHB from homoserine is the third pathway which requires deamination of homoserine to yield 2-oxo-4-hydroxybutyrate followed by the reduction of the latter to obtain DHB.

These three pathways bear enzymatic steps whose natural enzymatic activities do not exist. We obtained these “missing” activities by rational design based on structural and mechanistic knowledge, and/or by screening of candidate enzymes acting on sterically similar cognate substrates. Each of the individual pathways had its particular thermodynamic and physiological requirements which were met by rational strain design. The pathways were expressed in genetically optimised *Escherichia coli* strains for glucose to malate, and the production of a fairly good amount of DHB from glucose was obtained.

A field guide to automated cloning

Xingjian Xu, Almer van der Sloot, Raik Grünberg

IRIC, University of Montreal

The assembly of larger DNA constructs, often from a mix of gene synthesized and in-house fragments, is the starting point for most synthetic biology projects. Despite major technical advances, DNA assembly remains a bottleneck in many laboratories. Fully automated robotic cloning has been achieved within selected companies but has traditionally been considered too expensive and complex for academic settings.

We have implemented a robotic synthetic biology workstation that automates the complete multifragment DNA assembly work flow. This includes fragment PCR setup, cleanup, Gibson assembly, transformation, spreading on standard-format microplates, robust colony picking, colony PCR, DNA miniprep and auxiliary steps. Usability and user-adoption is facilitated by a strictly modular design as well as by convenient configuration through MS-Excel tables.

We discuss costs, practical issues and experiences with automated assembly design (j5) as well as an in-house sample tracking solution.

Context Sensitivity: The Blackbox of Synthetic Biology

Brian Jester1, Noémie Brisemeur1, Costas Bouyioukos1, Hafez el Sayyed1,2, François Képès1

1-iSSB - CNRS FRE3561 University of Evry-Val-d'Essonne 5 rue Henri Desbruères Genopole Campus 1, Bât. 6 F-91030 ÉVRY CEDEX, FRANCE,
2-Current address Collège de France 11, place Marcelin Berthelot 75231 Paris Cedex 05 France

The bacterial chromosome is embedded within a highly complex environment where it is constantly subjected to a huge range of forces. These interactions impact both the three- dimensional architecture and function of the chromosome in a dynamic way. Even though huge volumes of work have been dedicated to define specific aspects of chromosomal structure and how chromosomes perform their essential tasks, much is still unknown. One of the genomic characteristics that remains elusive is context sensitivity. This can be defined as the variable performance of a genetic circuit when embedded in different genomic locations.
In this work we have begun to unravel and characterize the underlying mechanisms that contribute towards context sensitivity. The construction of a suite of new genomic engineering tools and techniques has enabled us to quantitate the extent of context sensitivity throughout the genome. Building upon these technologies we have constructed several libraries of strains where we have been able to systematically isolate and quantitate the impact that several specific genetic variables have upon context sensitivity. Using computational tools developed within our lab to analyze the different uniquely derived biological datasets we have been able to identify some common features among them. These results suggest that the organization of the genome has evolved to take advantage of a periodic gene layout that promotes the 3D co-localization of co-regulated genes within the cell. Our growing understanding of these rules governing genome layout has laid the foundation for future rational de novo genome design projects.

Metabolic engineering strategy to produce oleochemicals from raw starch

Rodrigo Ledesma-Amaro1,2, Thierry Dulermo1,2, Jean-Marc Nicaud1,2

1INRA, UMR1319, MICALIS, Domaine de Vilvert, F-78352 Jouy-en-Josas, France
2 AgroParisTech, UMR Micalis, Jouy-en-Josas, France

rodrigo.ledesma@grignon.inra.fr

Microbial oils are promising alternatives to petroleum for the sustainable production of biofuels and chemicals. Among the possible producers, oleaginous yeasts are able to naturally accumulate high amounts of lipids (20-90% of the DCW). By far, the most studied and engineered oleaginous yeast is *Yarrowia lipolytica*, which has been previously modified to accumulate up to 90% of its DCW as lipids from glucose. Despite this high accumulation of biolipids, the process to produce them at large scale is not economically feasible, which has limited their use to high value lipids such as omega 3 fatty acids. Therefore, our aim is to reduce the production cost of biolipids in *Y. lipolytica* by 1) expanding the range of waste products and cheap materials that can be used as substrates and 2) the production of modified fatty acids with higher market value. Here we present the engineering of *Yarrowia* to produce lipids from raw starch. Starch is one of the most abundant sugar on earth usually found in high amounts in many industrial waste effluents. It is a glucose polymer that requires at least two enzymatic activities to be degraded, alpha amylase and glucoamylase, none of them found in *Y. lipolytica*. After the heterologous expression of two synthetic genes coding for the mentioned activities, the modified strain was able to grow and produce a few lipids from raw starch. We therefore transferred this strategy to a metabolically engineered strain with more than 9 genes modified that allow it to accumulate high amount of lipids. As a result, the strain produces high amount of neutral lipids from raw starch and from a waste effluent, thus reducing the production cost of bio-oils. Finally, this multiple engineered yeast can serve as a starting point to further modification to produce high value lipids from cheap substrates.

The different structures of the DNA

Guillaume Le Treut, Henri Orland and François Képès

Institute of Systems and Synthetic Biology, Genopole, Évry, France
Institut de Physique Théorique, CEA, Saclay, France

DNA is a biopolymer, well known for embedding the genetic program of one living organ- ism in its sequence of nucleotides, which is transmitted through generations. Noting that the ultimate goal of one living cell is to divide and transmit its genome with as few mutations as possible, the sequence of the nucleotides appears immutable. To the contrary, the topology of the DNA molecule in the nucleus changes through the life cycle, and it can adopt very different types of structures, which entail a variety of biological responses. The resulting states of the DNA have in common to be induced by DNA-binding proteins. In the case of transcription, it has become clear that proteins called transcription factors (TF), can bind and bring into prox- imity regions possibly far apart along the sequence of nucleotides. Namely, HiC experiments have demonstrated that co-regulated genes tend to colocalize, and presumably, are transcribed in transcription factories.

From a physical standpoint, it remains to be understood what is the essential mechanism in- ducing the different structures observed. In order to address this question, we adopt a simplified model in which DNA chromosomes are assimilated to polymer chains. We first characterize the thermodynamical equilibrium of DNA chains interacting with a solution of non-specific bind- ing proteins using a Flory-Huggins free energy model. Namely, we explored the dependence on DNA and protein concentrations of the DNA collapse. For physiologically relevant values of the DNA-protein affinity, this collapse gives rise to a biphasic regime with a dense and a dilute phase; the corresponding phase diagram was computed. In a second approach based on Brownian Dynamics simulations, we show that the dense phase has either a molten globule or a crystalline structure, depending on the DNA bending rigidity, which is influenced by the ionic strength. These results are valid at the thermodynamical equilibrium and should therefore be consistent with many biological processes, whose characteristic timescales range typically from 1 ms to 10 s.

pORTMAGE: a highly precise, portable method for microbial genome-scale engineering

Ákos Nyerges, Bálint Csörgő, István Nagy, Balázs Bálint, Péter Bihari, Viktória Lázár, Gábor Apjok, Kinga Umenhoffer, Balázs Bogos, György Pósfai, Csaba Pál

Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, H-6726, Hungary

Multiplex bacterial genome editing opened a new avenue of research in both biotechnology and basic research. However, currently available tools for bacterial genome manipulation suffer from three major limitations. They i) have been optimized for a few laboratory model strains, ii) demand either extensive plasmid construction or the modification of the host genome prior to large-scale genome engineering, and iii) lead to the accumulation of numerous off-target modifications. These issues have serious implications on wide-spread applicability of genome-scale engineering.

To overcome these roadblocks of genome-scale engineering, we present the plasmid-based, portable Multiplex Automated Genome Engineering (pORTMAGE), a system that allows highly efficient modification on a wide variety of bacteria without producing any undesired, off-target effects. We characterized a dominant mutation in the methyl-directed mismatch repair (MMR) system and integrated its power to precisely disrupt mismatch-repair in target cells. With the integration of this advance, we developed a new workflow for genome engineering and demonstrated its applicability for high-throughput genome editing by efficient modification of multiple loci without any observable off-target mutagenesis. Due to the highly conserved nature of the bacterial MMR system, pORTMAGE simultaneously expands the range of bacterial species for genome-scale engineering. By placing the entire synthetic operon into a broad-host vector, we demonstrated the feasibility of genome editing on enterobacterial species. For the first time, pORTMAGE allowed us to generate mutant libraries in biotechnologically and clinically relevant bacteria with unprecedented ease and to address the extent of conservation of the molecular mechanisms underlying antibiotic resistance in clinical pathogens.

These advances allow the maturation of multiplex genome engineering into a more versatile genome editing tool and paves the way towards the efficient production of highly valuable bio-products, more precise therapeutic applications and the investigation of central issues in functional genomics and evolution.

.

A method for the in vitro evolution of XNA ligases

Marleen Renders
Rega Institute for Medical Research - KU Leuven

Synthetic biologists envisage the creation of new biological systems for different purposes (e.g. finding the origin of life, the production of compounds/materials/energy, maintaining and enhancing human health and environment). In addition to its scientific and technological challenges, synthetic biology has raised world-wide concerns about ethics, biosecurity and biosafety since it may lead to the design of harmful organisms (e.g. the recent creation of an H5N1 avian flu virus mutant that is transmissible between mammals has raised fears to trigger a human pandemic by lab escape). What renders Nature extremely vulnerable for genetically modified organisms (GMOs) is their common system for nucleic acid (NA) and protein biosynthesis. Ideally, no genetic cross-talk should be possible between GMOs and natural systems. Additionally, the proliferation of GMOs should be dependent on the exogenous supply of an unnatural molecule. In theory, both can be achieved in GMOs by the introduction of an 'orthogonal genetic system' that is unable to cross-talk with natural nucleic acids and is built of unnatural precursors that have to be provided as nutrients in the medium.
We develop XNA ligases that link XNA stretches to each other *in vitro.* The development of such ligases is critical to the further development of the field of synthetic biology since it can provide gene-sized XNA fragments for transliteration *in vivo*. Directed evolution methods are used in order to obtain this milestone.

Synthetic biologists envisage the creation of new biological systems for different purposes (e.g. finding the origin of life, the production of compounds/materials/energy, maintaining and enhancing human health and environment). In addition to its scientific and technological challenges, synthetic biology has raised world-wide concerns about ethics, biosecurity and biosafety since it may lead to the design of harmful organisms (e.g. the recent creation of an H5N1 avian flu virus mutant that is transmissible between mammals has raised fears to trigger a human pandemic by lab escape). What renders Nature extremely vulnerable for genetically modified organisms (GMOs) is their common system for nucleic acid (NA) and protein biosynthesis. Ideally, no genetic cross-talk should be possible between GMOs and natural systems. Additionally, the proliferation of GMOs should be dependent on the exogenous supply of an unnatural molecule. In theory, both can be achieved in GMOs by the introduction of an 'orthogonal genetic system' that is unable to cross-talk with natural nucleic acids and is built of unnatural precursors that have to be provided as nutrients in the medium.

We develop XNA ligases that link XNA stretches to each other in vitro. The development of such ligases is critical to the further development of the field of synthetic biology since it can provide gene-sized XNA fragments for transliteration in vivo. Directed evolution methods are used in order to obtain this milestone.

Bacterial computation using Plasmids as Wires

Jorge Rodríguez Grande, María del Pilar Garcillán Barcia, Fernando de la Cruz.

Institute of Biomedicine and Biotechnology of Cantabria (IBBTEC), Santander

A fundamental aim of synthetic biology is to build biological devices with computational abilities. These devices are usually based on transcriptional regulators and other diffusion-dependent systems. This approach poses some limitations on scalability, because of gene expression noise, and the scarcity of characterized, reliable transcriptional regulators.

Trying to solve these issues, we present a novel approach for the construction of logic gates in bacteria. We employ bacterial conjugation as a robust, modular communication system. Our approach is based on the intrinsic modularity of conjugative systems. The wide variety of orthogonal conjugative described in nature permits scalability, and the digital nature of conjugation enables binary logic. In this work we present our progress towards achieving synthetic devices using bacterial plasmids as wires.

Evolutionary orchestra: taxonomy and function in chloroplasts and mitochondria

Michael G. Sadovsky, *ICM SB RAS*, *msad@icm.krasn.ru*
Yulia A. Putintseva[†], [†]*SFU*, *kinomanka85@mail.ru*

We studied the relations in *structure* ÷ *function* ÷ *taxonomy* triad. *Structure* everywhere below is triplet frequency dictionary; *taxonomy* is defined morphologically (by somatic genome, ultimately), and the *function* is the worst thing to define, due to the unobservable diversity of that latter. Here we stipulate that *function* is determined by the whole genome of an organelle (these are mitochondria and chloroplasts).

The key idea of the study is to figure out a clusterization (if any), in the space of triplet frequencies, and then check, whether the revealed clusters comprise taxonomically close genomes; same is true for the functionally close genomes. To do that, a set of 2883 mitochondria genomes (247 chloroplast entities, respectively) have been converted into the frequency dictionaries W_3 thus representing each genome as a point in 63-dimensional space; the triplets with the least standard deviation observed over the dataset have been eliminated. These are GCG and GAC, for mitochondria and chloroplasts, respectively.

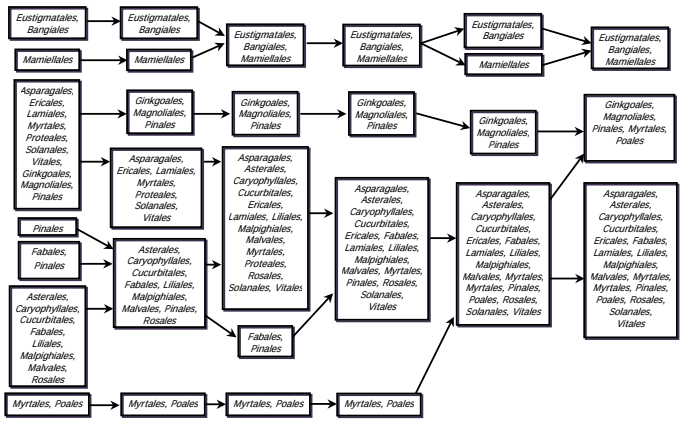

Figure 1: *Upward classification of chloroplasts genomes, case of stable clusterization.*

We used K-means to develop unsupervised upward classification of genomes (mitochondria and chloroplasts separately). A series of classifications for $2 \div 8$ classes has been developed. Then the species composition of each class (in the series) has been studied. Next, a redistribution pattern of the cluster members was studied: as the number of clusters obtained due to K-means was changed from L classes for $L-1$ classes, then the composition of the "higher" classes was a matter of analysis, in the terms of the composition of the "lower" classes. The observed pattern is a layer graph; that latter was far from a fully connected one, both for chloroplasts (see Fig. 1) and mitochondria.

Another approach to figure out the relation between structure and taxonomy (as well, as the structure and function) is based on elastic map clusterization. We used *ViDaExpert* software by A. Zynoviev to do that; the map parameters are shown in the legend (see Fig. 2). Unlike K-means, the elastic map technique is much more stable, in terms of the distribution of the data points over the clusters. Fig. 2 shows the distribution of mitochondria genomes of chordata species only. The figure shows the distribution of clades over the clusters; evidently, it is extremely non-random, but exhibits a high correlation between the cluster developed in frequencies, and the taxonomy.

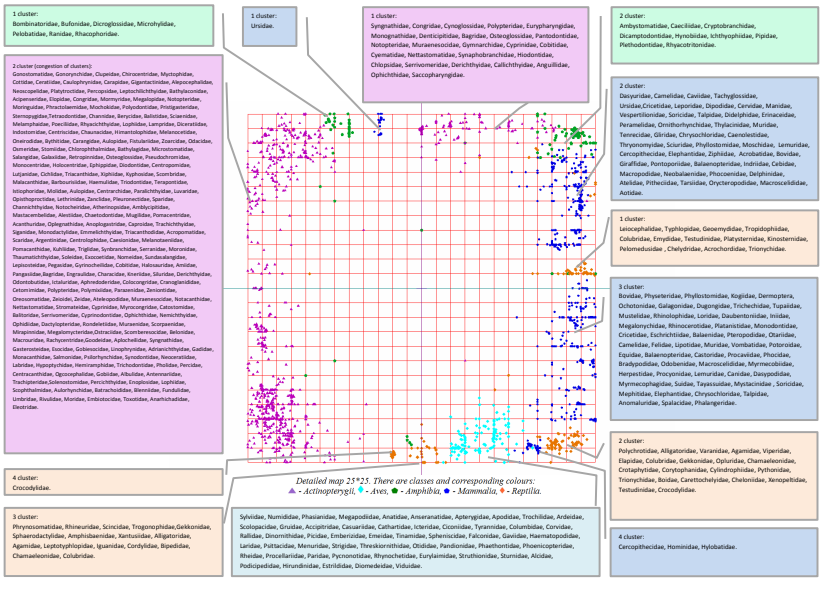

Figure 2: *Clusterization of mitochondria genomes by elastic map, soft,* 25×25 *size.*

We explored the same approach to figure out the relation between the structure and function. To do that, the data set comprising the genomes of chloroplasts and mitochondria of the same species has been developed. Since all the entities belong the same species, then one could expect very minor impact of taxonomy on the clusterization, but the function goes ahead. The results obtained here completely prove the strong relation between structure and function.

Conclusion: extremely high correlation between clusterization provided by triplet composition of organelle genomes and taxonomy (determined by the somatic genome) proves unambiguously the high synchrony in evolution of these two (physically independent) genetic systems. Same is true for *structure* ÷ *function* relation observed over the ensemble of chloroplasts and mitochondria of the same species.

Integrated informatics to support the SYNBIOCHEM Design-Build-Test cycle for production of fine and specialty chemicals

Neil Swainston[1], Pablo Carbonell[1], Andrew Currin[1], Mark Dunstan[1], Donal Fellows[1], Adrian Jervis[1], Rosalind Le Feuvre[1], Nicolas Rattray[1], Christopher Robinson[1], Reynard Speiss[1], Sandra Taylor[1], Alan Williams[1], Cunyu Yan[1], Pedro Mendes[1,2], Eriko Takano[1], Nicolas Turner[1], Nigel Scrutton[1]

[1]Manchester Centre for Synthetic Biology of Fine and Specialty Chemicals (SYNBIOCHEM), Manchester Institute of Biotechnology, University of Manchester, Manchester M1 7DN, United Kingdom

[2]Center for Quantitative Medicine and Department of Cell Biology, UConn Health, Farmington, Connecticut, USA.

The Manchester Centre for Synthetic Biology of Fine and Specialty Chemicals (SYNBIOCHEM) applies a **Design-Build-Test** cycle to the metabolic engineering of high value target compounds. Fundamental to this endeavour is the application and development of an integrated informatics infrastructure to support this work.

From a **Design** perspective, existing tools such as RetroPath [1] and XTMS [2] are being further developed in order to select candidate pathways and heterologous enzymes through application of cheminformatics and machine learning approaches. Novel graph databases are being developed to aid this work, warehousing chemical, enzyme and pathway data from a number of resources including ChEBI [3], Uniprot and MNXref allowing for multiple resources to be queries quickly and intuitively.

In supporting the **Build** aspect, novel DNA design algorithms [4] and assembly methods [5] are being developed to encode controlled enzyme and pathway libraries from which to apply intelligent directed evolution approaches [6].

Testing such variant libraries will be undertaken through targeted and untargeted metabolomics experiments, for which existing analysis pipelines [7] will be further developed. Other screening approaches will include high-throughput colorimetric and fluorometric assays coupled with next-generation sequencing. Consequently, a **Data** management system is in development to organise this work, based on existing e-lab notebooks and software such as SEEK [8] and JBEI ICE.

The Centre is committed to the development of software to support the metabolic engineering lifecycle, support appropriate standardisation efforts, and release code as open source.

[1] Retropath: automated pipeline for embedded metabolic circuits. Carbonell P, et al. *ACS Synth Biol*. 2014, **3**:565-77.
[2] XTMS: pathway design in an eXTended metabolic space. Carbonell P, et al. *Nucleic Acids Res*. 2014, **42**:W389-94.
[3] ChEBI in 2016: Improved services and an expanding collection of metabolites. Hastings J, et al. *Nucleic Acids Res*. 2015, pii: gkv1031.
[4] GeneGenie: optimized oligomer design for directed evolution. Swainston N, et al. *Nucleic Acids Res*. 2014, **42**:W395-400.
[5] SpeedyGenes: an improved gene synthesis method for the efficient production of error-corrected, synthetic protein libraries for directed evolution. Currin A, et al. *Protein Eng Des Sel*. 2014, **27**:273-80.
[6] Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Currin A, et al. *Chem Soc Rev*. 2015, **44**:1172-239.
[7] Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Brown M, et al. *Analyst*. 2009, **134**:1322-32.
[8] SEEK: a systems biology data and model management platform. Wolstencroft K, et al. *BMC Syst Biol*. 2015, **9**:33.

Taking control on *E. coli*'s genes expression using CRISPR interference

Antoine Vigouroux
Synthetic Biology Lab,
Morphogenesis and Microbial Growth,
Institut Pasteur

Abstract

The description of an adaptive immune system in bacteria, known as CRISPR, was one of the recent years' most groundbreaking discoveries in microbiology. CRISPR systems contain RNA-guided nucleases which have been used for numerous biotechnological applications, notably genome editing. In this project, which is a collaboration between the Synthetic Biology and the Morphogenesis and Microbial Growth groups of Institut Pasteur, we will make use of a CRISPR-based system allowing to take full control on the expression levels of arbitrary bacterial genes. In a dialogue between this technique and single-cell fluorescence microscopy, we will study for the first time the effects of CRISPR interference at the single cell level and its response heterogeneity across a cell population. By combining this approach with a morphogenetic study of cells which have undergone such a CRISPR treatment, we will be able to investigate how the genetic circuits of E. coli are wired to control the bacterium's shape and size throughout its replication cycle.

Synthetic secondary chromosomes to study chromosome maintenance in *Escherichia coli*

Daniel Schindler, Sonja Messerschmidt, Franziska Kemter and Torsten Waldminghaus

LOEWE Center for Synthetic Microbiology, SYNMIKRO, Philipps-University Marburg, Hans-Meerwein-Str. 6, D-35043 Marburg, Germany

Advances in *de novo* synthesis of DNA and assembly methodology make construction of synthetic chromosomes a reachable goal [1]. Considering a potential design leads almost certainly to the question what the essential parts of a chromosome are. Investigations on this question have mainly focused on the minimal set of genes needed to allow cells to live. However, chromosomes are more than arrays of genes. Chromosomes need systems to replicate, segregate and organize the encoded genetic information. We explore such chromosome maintenance systems by application of synthetic biology approaches and have designed and constructed the small synthetic secondary chromosome synVicII [2]. The replication origin stems from the secondary chromosome of the human pathogen *Vibrio cholerae* which carries a natural secondary chromosome. Application of a new assay for the assessment of replicon stability based on flow-cytometric analysis of an unstable GFP variant revealed an improved stability of synVicII compared to a secondary replicon based on *E. coli oriC*. Cell cycle analysis and determination of cellular copy numbers of synVicII indicate that replication timing of synVicII in *E. coli* is comparable to the natural chromosome II in *V. cholerae*. Now, synVicII is used as backbone to construct bigger replicons (~100kbp) with variations in the distribution of DNA motifs relevant for chromosome maintenance. Downstream functional characterization should allow deeper understanding of chromosome maintenance. Our long term goal is to establish synthetic secondary chromosomes as experimental system to study chromosome maintenance and to provide chromosome construction rules for biotechnology applications.

[1] Schindler D., Waldminghaus T. 2015. Synthetic chromosomes. *FEMS Microbiology Reviews*. pii: fuv030. [Epub ahead of print]

[2] Messerschmidt S. J., Kemter F. S. , Schindler D., Waldminghaus T. 2015. Synthetic secondary chromosomes in *Escherichia coli* based on the replication origin of *Vibrio cholerae* chromosome II. *Biotechnology Journal*, 10(2):302-14.

Deinococci genome scale structure study to design and engineer heterologous metabolic pathways

Julie Zaworski (1, 2), Jean-Paul Leonetti (2) and François Képès (1)

1 institute of Systems and Synthetic Biology, Genopole, CNRS, University of Evry, France
2 Deinove, Cap Sigma/ZAC Euromédecine II, Grabels, France

These last 20 years, bio-manufacturing companies have been coming across three major pitfalls. These are the toxicity of metabolic intermediates produced during the fermentation process, the instability of genetic constructions and the multiplicity of stimuli required for triggering the synthesis of target compounds. The aim of this thesis is to overcome these obstacles using DEINOVE team's expertise on Deinococci combined with the MEGA team's expertise on genome structure and its expression regulation. The analysis of the genome architecture will be carried out with two in-house softwares of the GREAT:SCAN suite, "patterns" will allow to describe the patterns of relative position of genes regulated by relevant transcription factors. "PreCisIon" will permit to predict further gene targets for the same transcription factor. The most promising ones will be tested at the bench for validation and characterization. Thereafter, a Deinococcus geothermalis library of strains with characterized insertion sites will be created. It will allow accelerating genomic insertion of genes encoding an interesting metabolic pathway in the near future. This project will allow revising and calibrating models and algorithms for the future genome designs, for the benefit of all biotechnological vectors and particularly Deinococci ones.