Computational Methods to Assess the Production Potential of Bio-Based Chemicals

Elevated costs and long implementation times of bio-based processes for producing chemicals represent a bottleneck for moving to a bio-based economy. A prospective analysis able to elucidate economically and technically feasible product targets at early research phases is mandatory. Computational tools can be implemented to explore the biological and technical spectrum of feasibility, while constraining the operational space for desired chemicals. In this chapter, two different computational tools for assessing potential for bio-based production of chemicals from different perspectives are described in detail. The first tool is GEM-Path: an algorithm to compute all structurally possible pathways from one target molecule to the host metabolome. The second tool is a framework for Modeling Sustainable Industrial Chemicals production (MuSIC), which integrates modeling approaches for cellular metabolism, bioreactor design, upstream/downstream processes, and economic impact assessment. Integrating GEM-Path and MuSIC will play a vital role in supporting early phases of research efforts and guide the policy makers with decisions, as we progress toward planning a sustainable chemical industry.
Bacterial cells with improved tolerance to isobutyric acid
Bacterial cells genetically modified to improve their tolerance to certain commodity chemicals, such as isobutyric acid and related compounds, and methods of preparing and using such bacterial cells for production of isobutyric acid and related compounds.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, Bacterial Cell Factory Optimization, Global Econometric Modeling, Department of Biotechnology and Biomedicine, Bacterial Synthetic Biology, ALE Technology & Software Development
Authors: Lennen, R. (Intern), Nielsen, A. T. (Intern), Herrgård, M. (Intern), Sommer, M. O. A. (Intern), Feist, A. (Intern), Mohamed, E. T. T. (Intern)
Publication date: 16 Nov 2017

Publication information
IPC: C12P 7/52 A I
Patent number: WO2017194696
Date: 16/11/2017
Priority date: 09/06/2016
Priority number: EP20160173673
Original language: English
Main Research Area: Technical/natural sciences
Source: espacenet
Source-ID: WO2017194696
Publication: Research › Patent – Annual report year: 2017

Bacterial cells with improved tolerance to polyamines
Provided are bacterial cells genetically modified to improve their tolerance to certain commodity chemicals, such as polyamines, and methods of preparing and using such bacterial cells for production of polyamines and other compounds.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, Bacterial Cell Factory Optimization, Global Econometric Modeling, Department of Biotechnology and Biomedicine, Bacterial Synthetic Biology, ALE Technology & Software Development
Authors: Lennen, R. (Intern), Nielsen, A. T. (Intern), Herrgaard, M. (Intern), Sommer, M. O. A. (Intern), Feist, A. (Intern), Tharwat Tolba Mohamed, E. (Intern)
Publication date: 15 Jun 2017

Publication information
IPC: C12P 13/00 A I
Patent number: WO2017097828
Date: 15/06/2017
Priority date: 09/02/2016
Priority number: EP20160154829
Original language: English
Main Research Area: Technical/natural sciences
Source: espacenet
Source-ID: WO2017097828
Publication: Research › Patent – Annual report year: 2017

Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution
There is a need to replace petroleum-derived with sustainable feedstocks for chemical production. Certain biomass feedstocks can meet this need as abundant, diverse, and renewable resources. Specific ionic liquids (ILs) can play a role in this process as promising candidates for chemical pretreatment and deconstruction of plant-based biomass feedstocks as they efficiently release carbohydrates which can be fermented. However, the most efficient pretreatment ILs are highly toxic to biological systems, such as microbial fermentations, and hinder subsequent bioprocessing of fermentative sugars obtained from IL-treated biomass. To generate strains capable of tolerating residual ILs present in treated feedstocks, a tolerance adaptive laboratory evolution (TALE) approach was developed and utilized to improve growth of two different Escherichia coli strains, DH1 and K-12 MG1655, in the presence of two different ionic liquids, 1-ethyl-3-methylimidazolium...
acetate ([C2C1Im][OAc]) and 1-butyl-3-methylimidazolium chloride ([C4C1Im]Cl). For multiple parallel replicate populations of E. coli, cells were repeatedly passed to select for improved fitness over the course of approximately 40 days. Clonal isolates were screened and the best performing isolates were subjected to whole genome sequencing. The most prevalent mutations in tolerant clones occurred in transport processes related to the functions of mdtJI, a multidrug efflux pump, and yhdP, an uncharacterized transporter. Additional mutations were enriched in processes such as transcriptional regulation and nucleotide biosynthesis. Finally, the best-performing strains were compared to previously characterized tolerant strains and showed superior performance in tolerance of different IL and media combinations (i.e., cross tolerance) with robust growth at 8.5% (w/v) and detectable growth up to 11.9% (w/v) [C2C1Im][OAc]. The generated strains thus represent the best performing platform strains available for bioproduction utilizing IL-treated renewable substrates, and the TALE method was highly successful in overcoming the general issue of substrate toxicity and has great promise for use in tolerance engineering.
Adaptive laboratory evolution, Escherichia coli, Ionic liquids, Renewable feedstocks

Increased production of L-serine in Escherichia coli through Adaptive Laboratory Evolution

L-serine is a promising building block biochemical with a high theoretical production yield from glucose. Toxicity of L-serine is however prohibitive for high-titer production in E. coli. Here, E. coli lacking L-serine degradation pathways was evolved for improved tolerance by gradually increasing L-serine concentration from 3 to 100 g/L using adaptive laboratory evolution (ALE). Genome sequencing of isolated clones revealed multiplication of genetic regions, as well as mutations in thrA, thereby showing a potential mechanism of serine inhibition. Other mutations were evaluated by MAGE combined with amplicon sequencing, revealing role of rho, lrp, pykF, eno, and rpoB on tolerance and fitness in minimal medium.

Production using the tolerant strains resulted in 37 g/L of L-serine with a 24% mass yield. The resulting titer is similar to the highest production reported for any organism thereby highlighting the potential of ALE for industrial biotechnology.
Isolation and characterization of the E. coli membrane protein production strain Mutant56(DE3)

Membrane protein production is usually toxic to E. coli. However, using genetic screens strains can be isolated in which the toxicity of membrane protein production is reduced, thereby improving production yields. Best known examples are the C41(DE3) and C43(DE3) strains, which are both derived from the T7 RNA polymerase (P)-based BL21(DE3) protein production strain. In C41(DE3) and C43(DE3) mutations lowering t7rnap expression levels result in strongly reduced T7 RNAP accumulation levels. As a consequence membrane protein production stress is alleviated in the C41(DE3) and C43(DE3) strains, thereby increasing membrane protein yields. Here, we isolated Mutant56(DE3) from BL21(DE3) using a genetic screen designed to isolate BL21(DE3)-derived strains with mutations alleviating membrane protein production stress other than the ones in C41(DE3) and C43(DE3). The defining mutation of Mutant56(DE3) changes one amino acid in its T7 RNAP, which weakens the binding of the T7 RNAP to the T7 promoter governing target gene expression rather than lowering T7 RNAP levels. For most membrane proteins tested yields in Mutant56(DE3) were considerably higher than in C41(DE3) and C43(DE3). Thus, the isolation of Mutant56(DE3) shows that the evolution of BL21(DE3) can be promoted towards further enhanced membrane protein production.

General information

State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, iLoop, Research Groups, Global Econometric Modeling, Microbial Evolution and Synthetic Biology, Stockholm University, Helmholtz Centre for Environmental Research, University of Groningen
Authors: Baumgarten, T. (Ekstern), Schlegel, S. (Ekstern), Wagner, S. (Ekstern), Löw, M. (Ekstern), Eriksson, J. (Ekstern), Bonde, I. (Intern), Herrgard, M. (Intern), Heipieper, H. J. (Ekstern), Nørholm, M. (Intern), Slotboom, D. J. (Ekstern), de Gier, J. (Ekstern)
There is an urgent need to significantly accelerate the development of microbial cell factories to produce fuels and chemicals from renewable feedstocks in order to facilitate the transition to a biobased society. Methods commonly used within the field of systems biology including omics characterization, genome-scale metabolic modeling, and adaptive laboratory evolution can be readily deployed in metabolic engineering projects. However, high performance strains usually carry tens of genetic modifications and need to operate in challenging environmental conditions. This additional complexity compared to basic science research requires pushing systems biology strategies to their limits and often spurs innovative developments that benefit fields outside metabolic engineering. Here we survey recent advanced applications of systems biology methods in engineering microbial production strains for biofuels and -chemicals.
Bio-based chemicals - green, but also sustainable?

For almost two decades, the chemical industry has put great effort into developing bio-chemicals, among others to fight global warming caused by greenhouse gas emissions, one of the biggest threats that are faced by our society today. To facilitate a growing and versatile bio-based chemical production, the US Department of Energy proposed in 2004 a list of 12 building block chemicals which can either be converged through biological or chemical conversions. Moving toward more bio-based chemicals, the chemical industry does not only claim to reduce climate change impacts, but also that they are increasing overall sustainability in chemical production. Whether such claims are justifiable is unclear. When sustainability of bio-based polymer production is assessed, various environmental trade-offs occur that need to be considered. It is not enough to claim that a bio-chemical is sustainable by exclusively looking at reduced global warming impacts related to avoiding oil refining and related greenhouse gas emissions. However, there is big variation of which impacts are assessed and which life cycle stages are included between existing published studies focusing on assessing environmental sustainability of bio-based polymers.

EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in *Saccharomyces cerevisiae*

*Saccharomyces cerevisiae* is widely used in the biotechnology industry for production of ethanol, recombinant proteins, food ingredients and other chemicals. In order to generate highly producing and stable strains, genome integration of genes encoding metabolic pathway enzymes is the preferred option. However, integration of pathway genes in single or few copies, especially those encoding rate-controlling steps, is often not sufficient to sustain high metabolic fluxes. By exploiting the sequence diversity in the long terminal repeats (LTR) of Ty retrotransposons, we developed a new set of integrative vectors, EasyCloneMulti, that enables multiple and simultaneous integration of genes in *S. cerevisiae*. By creating vector backbones that combine consensus sequences that aim at targeting subsets of Ty sequences and a quickly degrading selective marker, integrations at multiple genomic loci and a range of expression levels were obtained, as assessed with the green fluorescent protein (GFP) reporter system. The EasyCloneMulti vector set was applied to balance the expression of the rate-controlling step in the β-alanine pathway for biosynthesis of 3-hydroxypropionic acid (3HP). The best 3HP producing clone, with 5.45 g.L\(^{-1}\) of 3HP, produced 11 times more 3HP than the lowest producing clone, which demonstrates the capability of EasyCloneMulti vectors to impact metabolic pathway enzyme activity.
Engineering and systems level analysis of *Saccharomyces cerevisiae* for production of 3 hydroxypropionic acid via malonyl CoA reductase dependent pathway

In the future, oil- and gas-derived polymers may be replaced with bio-based polymers, produced from renewable feedstocks using engineered cell factories. Acrylic acid and acrylic esters with an estimated world annual production of approximately 6 million tons by 2017 can be derived from 3-hydroxypropionic acid (3HP), which can be produced by microbial fermentation. For an economically viable process 3HP must be produced at high titer, rate and yield and preferably at low pH to minimize downstream processing costs.

**General information**

State: Published

Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, Yeast Metabolic Engineering, iLoop, Applied Metabolic Engineering, Yeast Cell Factories, RWTH Aachen University, Chalmers University of Technology


Number of pages: 13

Publication date: 2016

Main Research Area: Technical/natural sciences

**Publication information**

Journal: Microbial Cell Factories

Volume: 15

Issue number: 53

ISSN (Print): 1475-2859

Ratings:

- BFI (2017): BFI-level 1
- Web of Science (2017): Indexed yes
- BFI (2016): BFI-level 1
- Scopus rating (2016): CiteScore 3.92 SJR 1.446 SNIP 1.228
- Web of Science (2016): Indexed yes
- BFI (2015): BFI-level 1
- Scopus rating (2015): SJR 1.501 SNIP 1.24 CiteScore 4.08
- Web of Science (2015): Indexed yes
- BFI (2014): BFI-level 1
- Scopus rating (2014): SJR 1.672 SNIP 1.471 CiteScore 4.25
- Web of Science (2014): Indexed yes
- BFI (2013): BFI-level 1
- Scopus rating (2013): SJR 1.686 SNIP 1.43 CiteScore 4.22
- ISI indexed (2013): ISI indexed yes
- Web of Science (2013): Indexed yes
- BFI (2012): BFI-level 1
- Scopus rating (2012): SJR 1.392 SNIP 1.312 CiteScore 3.69
- ISI indexed (2012): ISI indexed yes
- Web of Science (2012): Indexed yes
- BFI (2011): BFI-level 1
- Scopus rating (2011): SJR 1.417 SNIP 1.38 CiteScore 3.91
- ISI indexed (2011): ISI indexed yes
- Web of Science (2011): Indexed yes
- BFI (2010): BFI-level 1
- Scopus rating (2010): SJR 1.609 SNIP 1.463
- Web of Science (2010): Indexed yes
- BFI (2009): BFI-level 1
- Scopus rating (2009): SJR 1.276 SNIP 1.206
- Web of Science (2009): Indexed yes
- BFI (2008): BFI-level 1
- Scopus rating (2008): SJR 1.325 SNIP 1.335
Genome-Scale Models

An introduction to genome-scale models, how to build and use them, will be given in this chapter. Genome-scale models have become an important part of systems biology and metabolic engineering, and are increasingly used in research, both in academia and in industry, both for modeling chemical production and for fundamental studies aiming at describing and explaining phenotypic behavior of microbial and mammalian cells.

General information

State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, Applied Metabolic Engineering, University of Minho
Authors: Bergdahl, B. (Intern), Sonnenschein, N. (Intern), Machado, D. (Ekstern), Herrgard, M. (Intern), Förster, J. (Intern)
Number of pages: 40
Pages: 143-182
Publication date: 2016

Host publication information

Title of host publication: Fundamental Bioengineering
Place of publication: Weinheim, Germany
Publisher: Wiley-VCH
Editor: Villadsen, J.
Edition: 1st edition
ISBN (Print): 9783527336746
ISBN (Electronic): 9783527697441
Chapter: 6
Series: Advanced Biotechnology
ISSN: 2365-3035
Main Research Area: Technical/natural sciences
Annotated genome, Eukaryotic models, Genome-scale models, Metabolic networks reconstruction, Metabolome, Model simulations, Proteome, Transcriptome
DOIs: 10.1002/9783527697441.ch06

Publication: Research - peer-review › Book chapter – Annual report year: 2015

Multi-omics Quantification of Species Variation of Escherichia coli Links Molecular Features with Strain Phenotypes

Escherichia coli strains are widely used in academic research and biotechnology. New technologies for quantifying strain-specific differences and their underlying contributing factors promise greater understanding of how these differences

Scopus rating (2007): SJR 1.13 SNIP 1.293
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 0.973 SNIP 0.906
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.99 SNIP 1.056
Scopus rating (2004): SJR 0.615 SNIP 0.478
Scopus rating (2003): SJR 0.528 SNIP 0.229
Web of Science (2003): Indexed yes
Original language: English
3-Hydroxypropionic acid, Saccharomyces cerevisiae, Redox metabolism, Metabolic engineering
Electronic versions:

Bibliographical note
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Source: PublicationPreSubmission
Source-ID: 122774706
Publication: Research - peer-review › Journal article – Annual report year: 2016
significantly impact physiology, synthetic biology, metabolic engineering, and process design. Here, we quantified strain-specific differences in seven widely used strains of *E. coli* (BL21, C, Crooks, DH5a, K-12 MG1655, K-12 W3110, and W) using genomics, phenomics, transcriptomics, and genome-scale modeling. Metabolic physiology and gene expression varied widely with downstream implications for productivity, product yield, and titer. These differences could be linked to differential regulatory structure. Analyzing high-flux reactions and expression of encoding genes resulted in a correlated and quantitative link between these sets, with strain-specific caveats. Integrated modeling revealed that certain strains are better suited to produce given compounds or express desired constructs considering native expression states of pathways that enable high-production phenotypes. This study yields a framework for quantitatively comparing strains in a species with implications for strain selection.

**General information**
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, iLoop, Network Reconstruction in Silico Biology, Big Data 2 Knowledge, Research Groups, University of California
Number of pages: 26
Pages: 238-251
Publication date: 2016
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Cell Systems
Volume: 3
Issue number: 3
ISSN (Print): 2405-4712
Ratings: Web of Science (2017): Indexed yes
Scopus rating (2016): CiteScore 4.31
Original language: English
Electronic versions: BSOG_nihms819693.pdf
DOIs: 10.1016/j.cels.2016.08.013
Source: FindIt
Source-ID: 2345699171
Publication: Research - peer-review › Journal article – Annual report year: 2016

**Predictable tuning of protein expression in bacteria**
We comprehensively assessed the contribution of the Shine-Dalgarno sequence to protein expression and used the data to develop EMOPEC (Empirical Model and Oligos for Protein Expression Changes; http://emopec.biosustain.dtu.dk). EMOPEC is a free tool that makes it possible to modulate the expression level of any *Escherichia coli* gene by changing only a few bases. Measured protein levels for 91% of our designed sequences were within twofold of the desired target level.

**General information**
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Bacterial Synthetic Biology, Bacterial Cell Factory Optimization, iLoop, Research Groups
Number of pages: 8
Pages: 233-236
Publication date: 2016
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Nature Methods
Volume: 13
ISSN (Print): 1548-7091
Ratings: BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Stoichiometric Representation of Gene–Protein–Reaction Associations Leverages Constraint-Based Analysis from Reaction to Gene-Level Phenotype Prediction

Genome-scale metabolic reconstructions are currently available for hundreds of organisms. Constraint-based modeling enables the analysis of the phenotypic landscape of these organisms, predicting the response to genetic and environmental perturbations. However, since constraint-based models can only describe the metabolic phenotype at the reaction level, understanding the mechanistic link between genotype and phenotype is still hampered by the complexity of gene-protein-reaction associations. We implement a model transformation that enables constraint-based methods to be applied at the gene level by explicitly accounting for the individual fluxes of enzymes (and subunits) encoded by each gene. We show how this can be applied to different kinds of constraint-based analysis: flux distribution prediction, gene essentiality analysis, random flux sampling, elementary mode analysis, transcriptomics data integration, and rational strain design. In each case we demonstrate how this approach can lead to improved phenotype predictions and a deeper understanding of the genotype-to-phenotype link. In particular, we show that a large fraction of reaction-based designs obtained by current strain design methods are not actually feasible, and show how our approach allows using the same methods to obtain feasible gene-based designs. We also show, by extensive comparison with experimental 13C-flux data, how simple reformulations of different simulation methods with gene-wise objective functions result in improved prediction accuracy. The model transformation proposed in this work enables existing constraint-based methods to be used at the gene level without modification. This automatically leverages phenotype analysis from reaction to gene level, improving the biological insight that can be obtained from genome-scale models.

General information
Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects

Homologous recombination of single-stranded oligonucleotides is a highly efficient process for introducing precise mutations into the genome of *E. coli* and other organisms when mismatch repair (MMR) is disabled. This can result in the rapid accumulation of off-target mutations that can mask desired phenotypes, especially when selections need to be employed following the generation of combinatorial libraries. While the use of inducible mutator phenotypes or other MMR evasion tactics have proven useful, reported methods either require non-mobile genetic modifications or costly oligonucleotides that also result in reduced efficiencies of replacement. Therefore a new system was developed, Transient Mutator Multiplex Automated Genome Engineering (TM-MAGE), that solves problems encountered in other methods for oligonucleotide-mediated recombination. TM-MAGE enables nearly equivalent efficiencies of allelic replacement to the use of strains with fully disabled MMR and with an approximately 12- to 33-fold lower off-target mutation rate. Furthermore, growth temperatures are not restricted and a version of the plasmid can be readily removed by sucrose counterselection. TM-MAGE was used to combinatorially reconstruct mutations found in evolved salt-tolerant strains, enabling the identification of causative mutations and isolation of strains with up to 75% increases in growth rate and greatly reduced lag times in 0.6 M NaCl.
A Multi-scale, Multi-disciplinary Approach for Assessing the Technological, Economic, and Environmental Performance of Bio-based Chemicals

In recent years, bio-based chemicals have gained interest as a renewable alternative to petrochemicals. However, there is a significant need to assess the technological, biological, economic and environmental feasibility of bio-based chemicals, particularly during the early research phase. Recently, the Multi-scale framework for Sustainable Industrial Chemicals (MuSIC) was introduced to address this issue by integrating modelling approaches at different scales ranging from cellular to ecological scales. This framework can be further extended by incorporating modelling of the petrochemical value chain and the de novo prediction of metabolic pathways connecting existing host metabolism to desirable chemical products. This multiscale, multi-disciplinary framework for quantitative assessment of bio-based chemicals will play a vital role in supporting engineering, strategy and policy decisions as we progress towards a sustainable chemical industry.

A Multi-scale, Multi-disciplinary Approach for Assessing the Technological, Economic, and Environmental Performance of Bio-based Chemicals

In recent years, bio-based chemicals have gained interest as a renewable alternative to petrochemicals. However, there is a significant need to assess the technological, biological, economic and environmental feasibility of bio-based chemicals, particularly during the early research phase. Recently, the Multi-scale framework for Sustainable Industrial Chemicals (MuSIC) was introduced to address this issue by integrating modelling approaches at different scales ranging from cellular to ecological scales. This framework can be further extended by incorporating modelling of the petrochemical value chain and the de novo prediction of metabolic pathways connecting existing host metabolism to desirable chemical products. This multiscale, multi-disciplinary framework for quantitative assessment of bio-based chemicals will play a vital role in supporting engineering, strategy and policy decisions as we progress towards a sustainable chemical industry.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, Global Econometric Modeling
Authors: Herrgard, M. (Intern), Sukumara, S. (Intern), Campodonico Alt, M. A. (Intern), Zhuang, K. (Intern)
Number of pages: 6
Pages: 1151-1156
A Multi-Scale, Multi-Disciplinary Approach for Assessing the Technological, Economic, and Environmental Performance of Bio-Based Chemicals

State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop
Analysis of gene essentiality in Escherichia coli across strains and growth conditions

Different types of knock-out studies have for years been applied in addressing the question of gene essentiality in various organisms. The development within the field of next generation sequencing has paved the way for more extensive studies due to the high throughput. One of these fairly resent methods is transposon insertion sequencing (Tn-Seq), in which a mutant library is constructed by randomly inserting transposons into the genome, the position of which is determined by sequencing. By knowing the number of inserts in each gene in the initial library it is possible to determine if genes are either essential or detrimental for growth in the test condition in question. In this study the TN-Seq method was used to investigate the differences in gene essentiality between four laboratory strains of E.coli subjected to four different growth conditions to investigate the reason for the differences in osmotic and chemical stress tolerance that exists between the strains as well as to assess the commonalities. Based on the sequencing data we identified genes that were essential for growth under the different conditions, some of which are essential in all conditions across strains and others that are specifically essential under certain growth conditions and/or in certain strains. This knowledge is important in the effort to engineer more stress tolerant strains, which are highly relevant for industrial purposes. Here is presented the bioinformatics analysis of the data, which includes one to one comparisons for each strain in each condition to the control condition and a multivariate analysis including all strains across conditions.

Analysis of genetic variation and potential applications in genome-scale metabolic modeling

Genetic variation is the motor of evolution and allows organisms to overcome the environmental challenges they encounter. It can be both beneficial and harmful in the process of engineering cell factories for the production of proteins and chemicals. Throughout the history of biotechnology, there have been efforts to exploit genetic variation in our favor to create strains with favorable phenotypes. Genetic variation can either be present in natural populations or it can be artificially created by mutagenesis and selection or adaptive laboratory evolution. On the other hand, unintended genetic variation during a long term production process may lead to significant economic losses and it is important to understand how to control this type of variation. With the emergence of next-generation sequencing technologies, genetic variation in microbial strains can now be determined on an unprecedented scale and resolution by re-sequencing thousands of strains systematically. In this article, we review challenges in the integration and analysis of large-scale re-sequencing data, present an extensive overview of bioinformatics methods for predicting the effects of genetic variants on protein function, and discuss approaches for interfacing existing bioinformatics approaches with genome-scale models of cellular processes in order to predict effects of sequence variation on cellular phenotypes.
Co-evolution of strain design methods based on flux balance and elementary mode analysis

More than a decade ago, the first genome-scale metabolic models for two of the most relevant microbes for biotechnology applications, Escherichia coli and Saccaromyces cerevisiae, were published. Shortly after followed the publication of OptKnock, the first strain design method using bilevel optimization to couple cellular growth with the production of a target product. This initiated the development of a family of strain design methods based on the concept of flux balance analysis. Another family of strain design methods, based on the concept of elementary mode analysis, has also been growing. Although the computation of elementary modes is hindered by computational complexity, recent breakthroughs have allowed applying elementary mode analysis at the genome scale. Here we review and compare strain design methods and look back at the last 10 years of in silico strain design with constraint-based models. We highlight some features of the different approaches and discuss the utilization of these methods in successful in vivo metabolic engineering applications.
Do genome-scale models need exact solvers or clearer standards?

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, New Bioactive Compounds, Yeast Cell Factories, University of California, Norwegian University of Science and Technology, University of Luxembourg, Sinopia Biosciences, Genomatica Inc, Harvard Medical School, Intrexon, Inc., Virginia Commonwealth University, University of Lausanne, Rose-Hulman Institute of Technology, California Institute of Technology, Babraham Institut, Stanford University, University of Toronto, Biological Research Center, University of Virginia, European Molecular Biology Laboratory, Institute for Systems Biology, University of Wisconsin-Madison, Virginia Tech, Lawrence Livermore National Laboratory, Utah State University, University of Queensland, Pennsylvania State University, Centro de Investigaciones Biologicas
Number of pages: 3
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Molecular Systems Biology
Volume: 11
Issue number: 10
Article number: 831
ISSN (Print): 1744-4292
Ratings:
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
Scopus rating (2016): CiteScore 8.23 SJR 8.366 SNIP 2.15
BFI (2016): BFI-level 2
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 8.731 SNIP 2.395 CiteScore 9.76
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 10.072 SNIP 3.505 CiteScore 11.8
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 9.637 SNIP 2.875 CiteScore 11.84
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 7.904 SNIP 2.417 CiteScore 10.13
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 7.481 SNIP 2.306 CiteScore 8.78
Editorial: Current Challenges in Modeling Cellular Metabolism

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, University of Minho
Authors: Machado, D. (Ekstern), Zhuang, K. (Intern), Sonnenschein, N. (Intern), Herrgard, M. (Intern)
Number of pages: 2
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Bioengineering and Biotechnology
Volume: 3
Article number: 193
Ratings:
Web of Science (2017): Indexed yes
Scopus rating (2016): CiteScore 1 SJR 0.169 SNIP 0.471
Scopus rating (2015): SJR 0.124 SNIP 0
Scopus rating (2014): SJR 0.123 SNIP 0
Original language: English
Metabolism, Modeling formalisms, Metabolic networks, Genome-scale modeling, Kinetic modeling
Electronic versions:
Editorial_Current_Challenges_in_Modeling_Cellular_Metabolism.pdf
DOIs:
10.3389/fbioe.2015.00193

Bibliographical note
© 2015 Machado, Zhuang, Sonnenschein and Herrgård. This is an openaccess article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
Enhanced Protein Production in *Escherichia coli* by Optimization of Cloning Scars at the Vector-Coding Sequence Junction

Protein production in *Escherichia coli* is a fundamental activity for a large fraction of academic, pharmaceutical, and industrial research laboratories. Maximum production is usually sought, as this reduces costs and facilitates downstream purification steps. Frustratingly, many coding sequences are poorly expressed even when they are codon-optimized and expressed from vectors with powerful genetic elements. In this study, we show that poor expression can be caused by certain nucleotide sequences (e.g., cloning scars) at the junction between the vector and the coding sequence. Since these sequences lie between the Shine-Dalgarno sequence and the start codon, they are an integral part of the translation initiation region. To identify the most optimal sequences, we devised a simple and inexpensive PCR-based step that generates sequence variants at the vector-coding sequence junction. These sequence variants modulated expression by up to 1000-fold. FACS-seq analyses indicated that low GC content and relaxed mRNA stability ($\Delta G$) in this region were important, but not the only, determinants for high expression.

Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via \( \beta \)-alanine

Microbial fermentation of renewable feedstocks into plastic monomers can decrease our fossil dependence and reduce global CO2 emissions. 3-Hydroxypropionic acid (3HP) is a potential chemical building block for sustainable production of superabsorbent polymers and acrylic plastics. With the objective of developing *Saccharomyces cerevisiae* as an efficient cell factory for high-level production of 3HP, we identified the \( \beta \)-alanine biosynthetic route as the most economically attractive according to the metabolic modeling. We engineered and optimized a synthetic pathway for *de novo* biosynthesis of \( \beta \)-alanine and its subsequent conversion into 3HP using a novel \( \beta \)-alanine-pyruvate aminotransferase discovered in *Bacillus cereus*. The final strain produced 3HP at a titer of 13.7±0.3 g・L$^{-1}$ with a 0.14±0.0 C-mol・C-mol$^{-1}$ yield on glucose in 80 hours in controlled fed-batch fermentation in mineral medium at pH 5, and this work therefore lays...
the basis for developing a process for biological 3HP production.

**General information**

State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, Yeast Metabolic Engineering, Yeast Cell Factories, University of Copenhagen, Universidade Nova de Lisboa
Pages: 57-64
Publication date: 2015
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Metabolic Engineering
Volume: 27
ISSN (Print): 1096-7176
Ratings:
- BFI (2017): BFI-level 2
- Web of Science (2017): Indexed yes
- BFI (2016): BFI-level 2
- Scopus rating (2016): CiteScore 8.33 SJR 3.54 SNIP 1.864
- Web of Science (2016): Indexed yes
- BFI (2015): BFI-level 2
- Scopus rating (2015): SJR 3.611 SNIP 1.822 CiteScore 8.2
- Web of Science (2015): Indexed yes
- BFI (2014): BFI-level 2
- Scopus rating (2014): SJR 3.381 SNIP 2.034 CiteScore 7.23
- Web of Science (2014): Indexed yes
- BFI (2013): BFI-level 2
- Scopus rating (2013): SJR 4.004 SNIP 2.185 CiteScore 8.43
- ISI indexed (2013): ISI indexed yes
- Web of Science (2013): Indexed yes
- BFI (2012): BFI-level 2
- Scopus rating (2012): SJR 3.032 SNIP 1.858 CiteScore 6.72
- ISI indexed (2012): ISI indexed yes
- Web of Science (2012): Indexed yes
- BFI (2011): BFI-level 1
- Scopus rating (2011): SJR 3.124 SNIP 2.144 CiteScore 6.75
- ISI indexed (2011): ISI indexed yes
- Web of Science (2011): Indexed yes
- BFI (2010): BFI-level 1
- Scopus rating (2010): SJR 2.373 SNIP 1.802
- Web of Science (2010): Indexed yes
- BFI (2009): BFI-level 1
- Scopus rating (2009): SJR 2.575 SNIP 1.421
- Web of Science (2009): Indexed yes
- BFI (2008): BFI-level 1
- Scopus rating (2008): SJR 1.757 SNIP 1.028
- Web of Science (2008): Indexed yes
- Scopus rating (2007): SJR 1.504 SNIP 1.184
- Web of Science (2007): Indexed yes
- Scopus rating (2006): SJR 1.269 SNIP 0.892
- Web of Science (2006): Indexed yes
- Scopus rating (2005): SJR 1.056 SNIP 1.065
- Scopus rating (2004): SJR 1.657 SNIP 1.261
- Scopus rating (2003): SJR 1.168 SNIP 0.884
Methods for modeling Chinese hamster ovary (CHO) cell metabolism.

Embodiments of the present invention generally relate to the computational analysis and characterization of biological networks at the cellular level in Chinese Hamster Ovary (CHO) cells. Based on computational methods utilizing a hamster reference genome, the invention provides methods for identifying a CHO cell line having a desired genetic trait, as well as for generating a desired CHO cell line having a genetic basis for a desired phenotype. Additionally, described herein are methods for constructing and analyzing in silico models of biological networks for CHO cells.

Modeling the Contribution of Allosteric Regulation for Flux Control in the Central Carbon Metabolism of E. coli

Modeling cellular metabolism is fundamental for many biotechnological applications, including drug discovery and rational cell factory design. Central carbon metabolism (CCM) is particularly important as it provides the energy and precursors for other biological processes. However, the complex regulation of CCM pathways has still not been fully unraveled and recent studies have shown that CCM is mostly regulated at post-transcriptional levels. In order to better understand the role of allosteric regulation in controlling the metabolic phenotype, we expand the reconstruction of CCM in Escherichia coli with allosteric interactions obtained from relevant databases. This model is used to integrate multi-omics datasets and analyze the coordinated changes in enzyme, metabolite, and flux levels between multiple experimental conditions. We observe cases where allosteric interactions have a major contribution to the metabolic flux changes. Inspired by these results, we develop a constraint-based method (arFBA) for simulation of metabolic flux distributions that accounts for allosteric interactions. This method can be used for systematic prediction of potential allosteric regulation under the given experimental conditions based on experimental data. We show that arFBA allows predicting coordinated flux changes that
would not be predicted without considering allosteric regulation. The results reveal the importance of key regulatory
metabolites, such as fructose-1,6-bisphosphate, in controlling the metabolic flux. Accounting for allosteric interactions in
metabolic reconstructions reveals a hidden topology in metabolic networks, improving our understanding of cellular
metabolism and fostering the development of novel simulation methods that account for this type of regulation.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, University of Minho
Authors: Machado, D. (Ekstern), Herrgard, M. (Intern), Rocha, I. (Ekstern)
Number of pages: 11
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Bioengineering and Biotechnology
Volume: 3
Article number: 154
Ratings:
Web of Science (2017): Indexed yes
Scopus rating (2016): CiteScore 1 SJR 0.169 SNIP 0.471
Scopus rating (2015): SJR 0.124 SNIP 0
Scopus rating (2014): SJR 0.123 SNIP 0
Original language: English
Escherichia coli, Allosteric regulation, Constraint-based modeling, Metabolism, Systems biology
Electronic versions:
Modeling the contribution of allosteric regulation for flux control in the central carbon metabolism of E. coli.pdf
DOIs:
10.3389/fbioe.2015.00154

Bibliographical note
This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use,
distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the
original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or
reproduction is permitted which does not comply with these terms.
Source: FindIt
Source-ID: 2286805229
Publication: Research - peer-review › Journal article – Annual report year: 2015

Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae
CRISPR/Cas9 is a simple and efficient tool for targeted and marker-free genome engineering. Here, we report the
development and successful application of a multiplex CRISPR/Cas9 system for genome engineering of up to 5 different
genomic loci in one transformation step in baker's yeast Saccharomyces cerevisiae. To assess the specificity of the tool
we employed genome re-sequencing to screen for off-target sites in all single knock-out strains targeted by different
gRNAs. This extensive analysis identified no more genome variants in CRISPR/Cas9 engineered strains compared to
wild-type reference strains. We applied our genome engineering tool for an exploratory analysis of all possible single,
double, triple, quadruple and quintuple gene disruption combinations to search for strains with high mevalonate
production, a key intermediate for the industrially important isoprenoid biosynthesis pathway. Even though we did not
overexpress any genes in the mevalonate pathway, this analysis identified strains with mevalonate titers greater than 41-
fold compared to the wild-type strain. Our findings illustrate the applicability of this highly specific and efficient multiplex
genome engineering approach to accelerate functional genomics and metabolic engineering efforts. (C) 2015 International
Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Synthetic Biology Tools for Yeast
Authors: Jakociunas, T. (Intern), Bonde, I. (Intern), Herrgard, M. (Intern), Harrison, S. J. (Intern), Kristensen, M. (Intern),
Pedersen, L. E. (Intern), Jensen, M. K. (Intern), Keasling, J. (Intern)
Number of pages: 10
Pages: 213-222
Publication date: 2015
Main Research Area: Technical/natural sciences
Publication information
Journal: Metabolic Engineering
CRISPR/Cas9, Multiplex genome editing, Off-target analysis, Mevalonate, Yeast

DOIs:
10.1016/j.ymben.2015.01.008
Multi-scale exploration of the technical, economic, and environmental dimensions of bio-based chemical production

In recent years, bio-based chemicals have gained traction as a sustainable alternative to petrochemicals. However, despite rapid advances in metabolic engineering and synthetic biology, there remain significant economic and environmental challenges. In order to maximize the impact of research investment in a new bio-based chemical industry, there is a need for assessing the technological, economic, and environmental potentials of combinations of biomass feedstocks, biochemical products, bioprocess technologies, and metabolic engineering approaches in the early phase of development of cell factories. To address this issue, we have developed a comprehensive Multi-scale framework for modeling Sustainable Industrial Chemicals production (MuSIC), which integrates modeling approaches for cellular metabolism, bioreactor design, upstream/downstream processes and economic impact assessment. We demonstrate the use of the MuSIC framework in a case study where two major polymer precursors (1,3-propanediol and 3-hydroxypropionic acid) are produced from two biomass feedstocks (corn-based glucose and soy-based glycerol) through 66 proposed biosynthetic pathways in two host organisms (Escherichia coli and Saccharomyces cerevisiae). The MuSIC framework allows exploration of tradeoffs and interactions between economy-scale objectives (e.g. profit maximization, emission minimization), constraints (e.g. land-use constraints) and process- and cell-scale technology choices (e.g. strain design or oxygenation conditions). We demonstrate that economy-scale assessment can be used to guide specific strain design decisions in metabolic engineering, and that these design decisions can be affected by non-intuitive dependencies across multiple scales.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Global Econometric Modeling, Research Groups, iLoop
Authors: Zhuang, K. (Intern), Herrgard, M. (Intern)
Number of pages: 12
Pages: 1-12
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Metabolic Engineering
Volume: 31
ISSN (Print): 1096-7176
Ratings:
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 8.33 SJR 3.54 SNIP 1.864
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 3.611 SNIP 1.822 CiteScore 8.2
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 3.381 SNIP 2.034 CiteScore 7.23
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 4.004 SNIP 2.185 CiteScore 8.43
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 3.032 SNIP 1.858 CiteScore 6.72
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 3.124 SNIP 2.144 CiteScore 6.75
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 2.373 SNIP 1.802
Multi-Scale Technoeconomic Framework for Assessing Viability of Emerging Bio-based Processes

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop
Authors: Zhuang, K. (Intern), Sukumara, S. (Intern), Campodonico Alt, M. (Intern), Herrgard, M. (Intern)
Number of pages: 1
Publication date: 2015

Pharmacogenomic and clinical data link non-pharmacokinetic metabolic dysregulation to drug side effect pathogenesis
Drug side effects cause a significant clinical and economic burden. However, mechanisms of drug action underlying side effect pathogenesis remain largely unknown. Here, we integrate pharmacogenomic and clinical data with a human metabolic network and find that non-pharmacokinetic metabolic pathways dysregulated by drugs are linked to the development of side effects. We show such dysregulated metabolic pathways contain genes with sequence variants affecting side effect incidence, play established roles in pathophysiology, have significantly altered activity in corresponding diseases, are susceptible to metabolic inhibitors and are effective targets for therapeutic nutrient supplementation. Our results indicate that metabolic dysregulation represents a common mechanism underlying side effect pathogenesis that is distinct from the role of metabolism in drug clearance. We suggest that elucidating the relationships between the cellular response to drugs, genetic variation of patients and cell metabolism may help managing side effects by personalizing drug prescriptions and nutritional intervention strategies.
Seven gene deletions in seven days: fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress

Generation of multiple genomic alterations is currently a time consuming process. Here, a method was established that enables highly efficient and simultaneous deletion of multiple genes in *Escherichia coli*. A temperature sensitive plasmid containing arabinose inducible lambda Red recombineering genes and a rhamnose inducible flippase recombinase was constructed to facilitate fast marker-free deletions. To further speed up the procedure, we integrated the arabinose inducible lambda Red recombineering genes and the rhamnose inducible FLP into the genome of *E. coli* K-12 MG1655. This system enables growth at 37 °C, thereby facilitating removal of integrated antibiotic cassettes and deletion of additional genes in the same day. Phosphorothioated primers were demonstrated to enable simultaneous deletions during one round of electroporation. Utilizing these methods, we constructed strains in which four to seven genes were deleted in *E. coli* W and *E. coli* K-12. The growth rate of an *E. coli* K-12 quintuple deletion strain was significantly improved in the presence of high concentrations of acetate and NaCl. In conclusion, we have generated a method that enables efficient and simultaneous deletion of multiple genes in several *E. coli* variants. The method enables deletion of up to seven genes in as little as seven days.
A compendium of genetic variant data

Laboratory strains are genetically unstable if exposed to selective pressure as encountered, for example, during molecular cloning, fermentation, or adaptive laboratory evolution experiments. This genetic variation is the consequence of an adaptation process of the microorganism to stress conditions, e.g., high pressure or temperature, nutrient limitation, or toxic byproduct concentrations. The evolved strains display then new phenotypes: tolerance to a toxic byproduct or higher temperature, improved production rate of a byproduct, or higher uptake rates of nutrients. To understand the effects of those variations, it is necessary to collect and sort this genomic information in an organized fashion, including all relevant physiological data (e.g., growth rate, metabolomics, proteomics, transcriptomics, etc.). We propose a systematic way to collect heterogeneous datasets into a coherent database where the physiological characteristics of mutants can easily be queried. This database contains the experimental information sorted into normalized units. The aim of this repository is to become a golden standard of genetic variation information for microorganisms, providing standardized data obtained from distinct experiments. This compendium of genetic variant is a critical step to develop approaches to automatically and systematically characterize mutated strains in the future.
A kinetic model of thiamine biosynthesis in *Escherichia coli*

Thiamine can only be synthesized by prokaryotes and some eukaryotes, humans for example get it through their diet. Yet, it is key for the correct functioning of the carbohydrate and amino acid metabolism, and thiamine deficiency in humans can cause beriberi, which can result in muscle weakness or cardiovascular problems, among other symptoms. Nowadays it is common to add thiamine to commercial foods. Thus, it is important to produce it in a sustainable and efficient way. One approach to produce thiamine in a sustainable way is to use cell factories, and modeling of the metabolic network can be used to develop strategies for improved process efficiency. Constraint-based modeling methods have been successfully used to increase cell factory productivity. However, these approaches assume that the system is in a steady state, i.e., metabolite concentrations and reaction fluxes are constant along time. Therefore, kinetic models are needed to understand the dynamics of metabolite concentrations and reaction fluxes. We have built a kinetic model for the thiamine biosynthesis pathway in *Escherichia coli*. So far we have used convenience kinetics rate laws to describe the flux rates, but once more data has been collected, we will build enzyme modules where each elementary reaction step is explicitly modeled. This model will be used to understand the pathway dynamics and ultimately suggest genetic manipulation strategies to optimize thiamine production in *E. coli*.

Biotechnology for renewable chemicals

The majority of the industrial organic chemicals are derived from fossil sources. With the oil and gas resources becoming limiting, biotechnology offers a sustainable alternative for production of chemicals from renewable feedstocks. Yeast is an attractive cell factory for sustainable production of chemicals, due to its safe use status, tolerance of low pH and inhibitors, and amenability to large-scale fermentations. There are examples of commercial processes for production of organic acids such as lactic and succinic acids in yeast. We have engineered baker's yeast *Saccharomyces cerevisiae* for the production of non-native 3-hydroxypropionic acid (3HP). 3HP can be chemically dehydrated into acrylic acid and thus can serve as a biosustainable building block for acrylate-based products (diapers, acrylic paints, acrylic polymers, etc.)
Combinatorial Strategies for Improving Multiple-Stress Resistance in Industrially Relevant Escherichia coli Strains.

High-cell-density fermentation for industrial production of chemicals can impose numerous stresses on cells due to high substrate, product, and by-product concentrations; high osmolarity; reactive oxygen species; and elevated temperatures. There is a need to develop platform strains of industrial microorganisms that are more tolerant toward these typical processing conditions. In this study, the growth of six industrially relevant strains of Escherichia coli was characterized under eight stress conditions representative of fed-batch fermentation, and strains W and BL21(DE3) were selected as platforms for transposon (Tn) mutagenesis due to favorable resistance characteristics. Selection experiments, followed by either targeted or genome-wide next-generation-sequencing-based Tn insertion site determination, were performed to identify mutants with improved growth properties under a subset of three stress conditions and two combinations of individual stresses. A subset of the identified loss-of-function mutants were selected for a combinatorial approach, where strains with combinations of two and three gene deletions were systematically constructed and tested for single and multistress resistance. These approaches allowed identification of (i) strain-background-specific stress resistance phenotypes, (ii) novel gene deletion mutants in E. coli that confer single and multistress resistance in a strain-background-dependent manner, and (iii) synergistic effects of multiple gene deletions that confer improved resistance over single deletions. The results of this study underscore the suboptimality and strain-specific variability of the genetic network regulating growth under stressful conditions and suggest that further exploration of the combinatorial gene deletion space in multiple strain backgrounds is needed for optimizing strains for microbial bioprocessing applications.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Bacterial Cell Factories, iLoop, Research Groups
Authors: Lennen, R. (Intern), Herrgard, M. (Intern)
Pages: 6223-6242
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication information
Journal: Applied and Environmental Microbiology
Volume: 80
Issue number: 19
ISSN (Print): 0099-2240
Ratings:
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 4.08
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.891 SNIP 1.308 CiteScore 4.14
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.857 SNIP 1.384 CiteScore 4.02
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.899 SNIP 1.414 CiteScore 4.25
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.975 SNIP 1.429 CiteScore 4.29
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.914 SNIP 1.455 CiteScore 4.12
Evolution of Escherichia coli to 42 °C and Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations.

Adaptive laboratory evolution (ALE) has emerged as a valuable method by which to investigate microbial adaptation to a desired environment. Here, we performed ALE to 42 °C of ten parallel populations of Escherichia coli K-12 MG1655 grown in glucose minimal media. Tightly controlled experimental conditions allowed selection based on exponential-phase growth rate, yielding strains that uniformly converged toward a similar phenotype along distinct genetic paths. Adapted strains possessed as few as 6 and as many as 55 mutations, and of the 144 genes that mutated in total, 14 arose independently across two or more strains. This mutational recurrence pointed to the key genetic targets underlying the evolved fitness increase. Genome engineering was used to introduce the novel ALE-acquired alleles in random combinations into the ancestral strain, and competition between these engineered strains reaffirmed the impact of the key mutations on the growth rate at 42 °C. Interestingly, most of the identified key gene targets differed significantly from those found in similar temperature adaptation studies, highlighting the sensitivity of genetic evolution to experimental conditions and ancestral genotype. Additionally, transcriptomic analysis of the ancestral and evolved strains revealed a general trend for restoration of the global expression state back toward preheat stressed levels. This restorative effect was previously documented following evolution to metabolic perturbations, and thus may represent a general feature of ALE experiments. The widespread evolved expression shifts were enabled by a comparatively scant number of regulatory mutations, providing a net fitness benefit but causing suboptimal expression levels for certain genes, such as those governing flagellar formation, which then became targets for additional ameliorating mutations. Overall, the results of this study provide insight into the adaptation process and yield lessons important for the future implementation of ALE as a tool for scientific research and engineering.
Evolution Reveals A Glutathione-dependent Mechanism Of 3-hydroxypropionic Acid Detoxification
Biologically produced 3-hydroxypropionic acid (3HP) is a potential source for sustainable acrylates and can also find direct use as monomer in the production of biodegradable polymers. For industrial-scale production, high titer, rate and yield are essential; thus there is a need for robust cell factories tolerant to high concentration of 3HP, preferably at low pH. Through adaptive laboratory evolution we selected S. cerevisiae strains with improved tolerance to 3HP at pH 3.5. Genome sequencing of three independent clones identified single-nucleotide changes in the SFA1 gene encoding S- (hydroxymethyl)glutathione dehydrogenase. Introduction of the mutated SFA1 alleles or overexpression of any of the SFA1 alleles in a sfa16 strain enabled growth in the presence of above 40 g/L 3HP. We further found that aldehyde dehydrogenase (ALD6), S-formylglutathione hydrolase (YJL068C) and glutathione play a role in 3HP detoxification. Addition of glutathione relieved growth inhibition by 3HP for several yeast species and for E. coli; but glutathione could not enable growth of a S. cerevisiae sfa16 strain. Based on our findings we propose a 3-hydroxypropionic aldehyde-mediated mechanism underlying 3HP toxicity as well as a glutathione-dependent route for detoxification of 3-hydroxypropionic aldehyde (reuterin). The identified molecular response to 3HP and reuterin may well be a general mechanism for handling resistance to organic acids and aldehydes by living cells.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, CFB - Core Flow, Bacterial Cell Factories, Metagenomics, Fungal Cell Factories, KTH - Royal Institute of Technology, University of Copenhagen
Number of pages: 1
Publication date: 2014
Event: Abstract from Metabolic Engineering X, Vancouver, Canada.
Main Research Area: Technical/natural sciences
Publication: Research - peer-review › Conference abstract for conference – Annual report year: 2015

Evolution reveals a glutathione-dependent mechanism of 3-hydroxypropionic acid tolerance
Biologically produced 3-hydroxypropionic acid (3HP) is a potential source for sustainable acrylates and can also find direct use as monomer in the production of biodegradable polymers. For industrial-scale production, high titer, rate and yield are essential; thus there is a need for robust cell factories tolerant to high concentration of 3HP, preferably at low pH. Through adaptive laboratory evolution we selected S. cerevisiae strains with improved tolerance to 3HP at pH 3.5. Genome sequencing followed by functional analysis identified the causal mutation in SFA1 gene encoding S-(hydroxymethyl)glutathione dehydrogenase. Based on our findings we propose that 3HP toxicity is mediated by 3-hydroxypropionic aldehyde (reuterin) and that glutathione-dependent reactions are used for reuterin detoxification. The identified molecular response to 3HP and reuterin may well be a general mechanism for handling resistance to organic acid and aldehydes by living cells.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop, CFB - Core Flow, Bacterial Cell Factories, Metagenomics, Fungal Cell Factories, KTH - Royal Institute of Technology, University of Copenhagen
Number of pages: 10
Pages: 57-66
Publication date: 2014
Main Research Area: Technical/natural sciences
Publication: Research - peer-review › Journal article – Annual report year: 2014
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 8.33 SJR 3.54 SNIP 1.864
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 3.611 SNIP 1.822 CiteScore 8.2
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 3.381 SNIP 2.034 CiteScore 7.23
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 4.004 SNIP 2.185 CiteScore 8.43
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 3.032 SNIP 1.858 CiteScore 6.72
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 3.124 SNIP 2.144 CiteScore 6.75
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 2.373 SNIP 1.802
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 2.575 SNIP 1.421
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 1.757 SNIP 1.028
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 1.504 SNIP 1.184
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 1.269 SNIP 0.892
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 1.056 SNIP 1.065
Scopus rating (2004): SJR 1.657 SNIP 1.261
Scopus rating (2003): SJR 1.168 SNIP 0.884
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 1.182 SNIP 1.028
Web of Science (2002): Indexed yes
Scopus rating (2001): SJR 0.921 SNIP 0.766
Scopus rating (2000): SJR 0.724 SNIP 0.896
Original language: English
3-hydroxypropionic acid, Tolerance, 3-hydroxypropionic aldehyde (reuterin), Saccharomyces cerevisiae, Adaptive laboratory evolution
Electronic versions:
Evolution_reveals_a_glutathione_dependent_mechanism_of_3_hydroxypropionic_acid_tolerance_Kildegaard2014_MetEng.pdf
DOIs:
10.1016/j.ymben.2014.09.004
Source: PublicationPreSubmission
Source-ID: 100811114
Publication: Research - peer-review › Journal article – Annual report year: 2014
Increased tolerance towards serine obtained by adaptive laboratory evolution

The amino acid serine has previously been identified as one of the top 30 candidates of value added chemicals, making the production of serine from glucose attractive. Production of serine have previously been attempted in E. coli and C. glutamicum, however, titers sufficient for commercial applications have not yet been achieved. This is partly due to the fact that the key serine degradation pathway (serine to glycine), encoded by glyA, has not yet been successfully deleted in E. coli or C. glutamicum. So far, the most successful attempts of serine production have been achieved using a C. glutamicum auxotroph for the cofactor of glyA, however, this requires the use of rich fermentation media or the addition of folic acid. Here, we demonstrate that the two major pathways for degradation of serine can be deleted in E. coli MG1655. In addition to the conversion of serine to glycine (encoded by glyA), the conversion of serine to pyruvate (encoded by sdaA, sdaB and tdcG) was also deleted. As expected, the resulting strain turned out to be susceptible to even low concentrations of serine in the media. In order to improve the tolerance of the strain towards serine, adaptive laboratory evolution was implemented using a state of the art robotics platform. The strain was grown under inhibiting concentrations of serine in minimal media and was periodically transferred to new media during mid log phase. After achieving a desired increase in growth rate, the concentration was serine was gradually increased. During the evolution experiment, the serine tolerance was increased substantially. Genome re-sequencing was subsequently used to analyze the genotype of a number of selected strains. These results reveal insights towards the adaptation process as well as the mechanism of serine tolerance.

Library sequencing strategies for comparative analysis of stress resistance mechanisms in Escherichia coli strains

Transposon insertion sequencing (Tn-Seq) has recently emerged as a powerful next-generation sequencing method that enables querying the contributions of all genes in a bacterial genome toward the fitness of a growing organism. In this method, transposon insertion mutant libraries are constructed and subjected to growth selections. Following selection, the locations of all insertions in the population are counted and can be compared between a control and a target condition, enabling the identification of genes that are both conditionally essential and conditionally detrimental. We have exploited Tn-Seq to probe the basis for the large variations in osmotic and acetate stress tolerance of different laboratory strains of Escherichia coli (K-12 MG1655, BL21(DE3), W, and Crooks). Little is currently known to explain the source of this variation and to enable rational engineering to impart stress tolerance. Tn-Seq revealed many differences and similarities in resistance mechanisms at the genetic level across strains, allowing correlations to be made with growth phenotypes. Cross-strain comparisons of conditionally essential genes and their relative essentiality also suggest a large degree of variation in metabolic flux distributions and regulation of gene expression between strains. A number of direct targets for metabolic engineering of stress resistance via loss-of-function mutations were also discovered, and we show that deletion of a selection of these genes results in improved growth under the original selection condition.
Multi-scale Exploration of the Technical, Economic, and Environmental Dimensions of Bio-based Chemical Production

In recent years, bio-based chemicals have gained traction as a sustainable alternative to petrochemicals. In order to maximize the impacts of researches and investments, there is a need to focus on the most promising combinations of feedstocks, biochemical products, and bioprocesses. To address this issue, we developed a multiscale framework that integrates modeling approaches across scales of cellular metabolism, bioreactor, bioprocess, and economy/ecosystem, and is able to simultaneously assess biological, technological, economic and environmental feasibility of different
production scenarios. Using our framework, we assess the production of two major polymer precursors (1,3-propanediol and 3-hydroxypropionic acid) from two biomass feedstocks (corn-based glucose and soy-based glycerol) using two host organisms (E. coli and S. cerevisiae). We explore the sustainability and economic impacts of a variety of policies and practices (e.g. land-usaeg, energy source mixture, CO₂ emission cap), as well as trade offs between different objectives (e.g. profits for different sectors, emission minimization) for key stakeholders involved in the biochemical value chain (agriculture, energy, and biotechnology sectors).

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, Research Groups, iLoop
Authors: Zhuang, K. (Intern), Herrgard, M. (Intern)
Number of pages: 1
Publication date: 2014

Host publication information
Title of host publication: Abstract Book - DTU Sustain Conference 2014
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Main Research Area: Technical/natural sciences
Conference: DTU Sustain Conference 2014, Lyngby, Denmark, 17/12/2014 - 17/12/2014
Publication: Research - peer-review » Conference abstract in proceedings – Annual report year: 2014

Putting computational modeling at the fingertips of bench biologists
Communication is key in successful collaborations between theoretical and experimental life scientists. In our line of work we integrate physiological and systems-level data of cell factories with constraint-based Modeling approaches to predict suitable targets for metabolic engineering. The most important step in this process is the discourse and prioritization of strategies with the people that actually implement them. Having worked with a particular host organism for many years, experimentalists can often discard strategies based on previous experience or feasibility. Counterintuitive solutions, which are often very interesting from a patenting and biological perspective, require a detailed explanation to convince experimentalists to be worthwhile pursuing. Interactive pathway visualizations have turned out to be tremendously helpful in this context. So far, we have used the high-level programming environment Mathematica and an open source metabolic modeling package (MASS Toolbox) to quickly prototype maps and other interactive widgets. In the long term, we would like to make those publicly accessible using open web technologies.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, iLoop, Research Groups
Authors: Sonnenschein, N. (Intern), Herrgard, M. (Intern)
Publication date: 2014
Event: Poster session presented at 2nd EMBO Conference on Visualizing Biological Data, Heidelberg, Germany.
Main Research Area: Technical/natural sciences
Electronic versions:
VIZBI2014_poster_original_CFB_template.pdf
Links:
http://vizbi.org/2014/
Source: PublicationPreSubmission
Source-ID: 100889190
Publication: Research - peer-review » Poster – Annual report year: 2014

Systematic Evaluation of Methods for Integration of Transcriptomic Data into Constraint-Based Models of Metabolism.
Constraint-based models of metabolism are a widely used framework for predicting flux distributions in genome-scale biochemical networks. The number of published methods for integration of transcriptomic data into constraint-based models has been rapidly increasing. So far the predictive capability of these methods has not been critically evaluated and compared. This work presents a survey of recently published methods that use transcript levels to try to improve metabolic flux predictions either by generating flux distributions or by creating context-specific models. A subset of these methods is then systematically evaluated using published data from three different case studies in E. coli and S. cerevisiae. The flux predictions made by different methods using transcriptomic data are compared against experimentally determined extracellular and intracellular fluxes (from 13C-labeling data). The sensitivity of the results to method-specific parameters is also evaluated, as well as their robustness to noise in the data. The results show that none of the methods outperforms the others for all cases. Also, it is observed that for many conditions, the predictions obtained by simple flux balance analysis using growth maximization and parsimony criteria are as good or better than those obtained using methods that incorporate transcriptomic data. We further discuss the differences in the mathematical formulation of the methods, and their relation to the results we have obtained, as well as the connection to the underlying biological principles of metabolic regulation.
Community structure and function of high-temperature chlorophototrophic microbial mats inhabiting diverse geothermal environments

Six phototrophic microbial mat communities from different geothermal springs (YNP) were studied using metagenome sequencing and geochemical analyses. The primary goals of this work were to determine differences in community
composition of high-temperature phototrophic mats distributed across the Yellowstone geothermal ecosystem, and to identify metabolic attributes of predominant organisms present in these communities that may correlate with environmental attributes important in niche differentiation. Random shotgun metagenome sequences from six phototrophic communities (average 53Mbp/site) were subjected to multiple taxonomic, phylogenetic, and functional analyses. All methods, including G+C content distribution, MEGAN analyses, and oligonucleotide frequency-based clustering, provided strong support for the dominant community members present in each site. Cyanobacteria were only observed in non-sulfidic sites; de novo assemblies were obtained for Synechococcus-like populations at Chocolate Pots (CP_7) and Fischerella-like populations at White Creek (WC_6). Chloroflexi-like sequences (esp. Roseiflexus and/or Chloroflexus spp.) were observed in all six samples and contained genes involved in bacteriochlorophyll biosynthesis and the 3-hydroxypropionate carbon fixation pathway. Other major sequence assemblies were obtained for a Chlorobiales population from CP_7 (proposed family Thermochlorobacteriaceae), and an anoxygenic, sulfur-oxidizing Thermochromatium-like (Gamma-proteobacteria) population from Bath Lake Vista Annex (BLVA_20). Additional sequence coverage is necessary to establish more complete assemblies of other novel bacteria in these sites (e.g., Bacteroidetes and Firmicutes); however, current assemblies suggested that several of these organisms play important roles in heterotrophic and fermentative metabolisms. Definitive linkages were established between several of the dominant phylotypes present in these habitats and important functional processes such as photosynthesis, carbon fixation, sulfur oxidation, and fermentation.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, CFB - Core Flow, Indiana University-Purdue, Montana State University, Joint Genome Institute, Search for Extraterrestrial Intelligence Institute, Western Oregon University, University of Montana
Number of pages: 23
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Microbiology
Volume: 4
Article number: 106
ISSN (Print): 1664-302X
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 4.16 SJR 1.731 SNIP 1.172
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.878 SNIP 1.208 CiteScore 4.15
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.861 SNIP 1.16 CiteScore 3.76
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.751 SNIP 0.951 CiteScore 3.56
ISI indexed (2013): ISI indexed no
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.415 SNIP 0.725 CiteScore 2.78
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.626 SNIP 0.187
Web of Science (2011): Indexed yes
Original language: English
Microbial mats, Microbial interactions, Phototrophic bacteria, Functional genomics, Thermophilic bacteria
Electronic versions:
fmicb_04_00106.pdf
DOIs:
10.3389/fmicb.2013.00106
Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three aquificales lineages.

The Aquificales are thermophilic microorganisms that inhabit hydrothermal systems worldwide and are considered one of the earliest lineages of the domain Bacteria. We analyzed metagenome sequence obtained from six thermal "filamentous streamer" communities (~40 Mbp per site), which targeted three different groups of Aquificales found in Yellowstone National Park (YNP). Unassembled metagenome sequence and PCR-amplified 16S rRNA gene libraries revealed that acidic, sulfidic sites were dominated by Hydrogenobaculum (Aquificaceae) populations, whereas the circum-neutral pH (6.5-7.8) sites containing dissolved sulfide were dominated by Sulfurihydrogenibium spp. (Hydrogenothermaceae). Thermocrinis (Aquificaceae) populations were found primarily in the circum-neutral sites with undetectable sulfide, and to a lesser extent in one sulfidic system at pH 8. Phylogenetic analysis of assembled sequence containing 16S rRNA genes as well as conserved protein-encoding genes revealed that the composition and function of these communities varied across geochemical conditions. Each Aquificales lineage contained genes for CO2 fixation by the reverse-TCA cycle, but only the Sulfurihydrogenibium populations perform citrate cleavage using ATP citrate lyase (Acl). The Aquificaceae populations use an alternative pathway catalyzed by two separate enzymes, citryl-CoA synthetase (Ccs), and citryl-CoA lyase (Ccl). All three Aquificales lineages contained evidence of aerobic respiration, albeit due to completely different types of heme Cu oxidases (subunit I) involved in oxygen reduction. The distribution of Aquificales populations and differences among functional genes involved in energy generation and electron transport is consistent with the hypothesis that geochemical parameters (e.g., pH, sulfide, H2, O2) have resulted in niche specialization among members of the Aquificales.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, CFB - Core Flow, University of New Mexico, Montana State University, Joint Genome Institute, Ehime University, University of Illinois, Portland State University, Los Alamos National Laboratory
Number of pages: 25
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Microbiology
Volume: 4
Article number: 84
ISSN (Print): 1664-302X
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 4.16 SJR 1.731 SNIP 1.172
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.878 SNIP 1.208 CiteScore 4.15
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.861 SNIP 1.16 CiteScore 3.76
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.751 SNIP 0.951 CiteScore 3.56
ISI indexed (2013): ISI indexed no
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.415 SNIP 0.725 CiteScore 2.78
ISI indexed (2012): ISI indexed no
Multi-scale modeling for sustainable chemical production

With recent advances in metabolic engineering, it is now technically possible to produce a wide portfolio of existing petrochemical products from biomass feedstock. In recent years, a number of modeling approaches have been developed to support the engineering and decision-making processes associated with the development and implementation of a sustainable biochemical industry. The temporal and spatial scales of modeling approaches for sustainable chemical production vary greatly, ranging from metabolic models that aid the design of fermentative microbial strains to material and monetary flow models that explore the ecological impacts of all economic activities. Research efforts that attempt to connect the models at different scales have been limited. Here, we review a number of existing modeling approaches and their applications at the scales of metabolism, bioreactor, overall process, chemical industry, economy, and ecosystem. In addition, we propose a multi-scale approach for integrating the existing models into a cohesive framework.

The major benefit of this proposed framework is that the design and decision-making at each scale can be informed, guided, and constrained by simulations and predictions at every other scale. In addition, the development of this multi-scale framework would promote cohesive collaborations across multiple traditionally disconnected modeling disciplines to achieve sustainable chemical production.
Phylogenetic and Functional Analysis of Metagenome Sequence from High-Temperature Archaeal Habitats Demonstrate Linkages between Metabolic Potential and Geochemistry.

Geothermal habitats in Yellowstone National Park (YNP) provide an unparalleled opportunity to understand the environmental factors that control the distribution of archaea in thermal habitats. Here we describe, analyze, and synthesize metagenomic and geochemical data collected from seven high-temperature sites that contain microbial communities dominated by archaea relative to bacteria. The specific objectives of the study were to use metagenome sequencing to determine the structure and functional capacity of thermophilic archaeal-dominated microbial communities across a pH range from 2.5 to 6.4 and to discuss specific examples where the metabolic potential correlated with measured environmental parameters and geochemical processes occurring in situ. Random shotgun metagenome sequence (~40-45 Mb Sanger sequencing per site) was obtained from environmental DNA extracted from high-temperature sediments and/or microbial mats and subjected to numerous phylogenetic and functional analyses. Analysis of individual sequences (e.g., MEGAN and G + C content) and assemblies from each habitat type revealed the presence of dominant archaeal populations in all environments, 10 of whose genomes were largely reconstructed from the sequence data. Analysis of protein family occurrence, particularly of those involved in energy conservation, electron transport, and autotrophic metabolism, revealed significant differences in metabolic strategies across sites consistent with differences in major geochemical attributes (e.g., sulfide, oxygen, pH). These observations provide an ecological basis for understanding the distribution of indigenous archaeal lineages across high-temperature systems of YNP.
The Uses and Future Prospects of Metabolomics and Targeted Metabolite Profiling in Cell Factory Development

The development of cell factories for the production of chemicals has traditionally relied on measurements of product metabolite titers to assess the performance of genetically manipulated strains. With the development of improved metabolomics and targeted metabolite profiling methods, these broader measurements of the cellular metabolic state are now becoming part of the toolbox used to characterize cell factories. In this review we briefly summarize the benefits and challenges of global metabolomics and targeted metabolite profiling methods and discuss the application of these methods in both pathway discovery and cell factory engineering. We focus particularly on exploring the potential of global metabolomics to complement more traditional targeted methods. We conclude the review by discussing emerging trends in metabolomics and how these developments can aid the engineering of better cell factories in the future.
The YNP metagenome project: Environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem

The Yellowstone geothermal complex contains over 10,000 diverse geothermal features that host numerous phylogenetically deeply rooted and poorly understood archaea, bacteria, and viruses. Microbial communities in high-temperature environments are generally less diverse than soil, marine, sediment, or lake habitats and therefore offer a tremendous opportunity for studying the structure and function of different model microbial communities using environmental metagenomics. One of the broader goals of this study was to establish linkages among microbial distribution, metabolic potential, and environmental variables. Twenty geochemically distinct geothermal ecosystems representing a broad spectrum of Yellowstone hot-spring environments were used for metagenomic and geochemical analysis and included approximately equal numbers of: (1) phototrophic mats, (2) “filamentous streamer” communities, and (3) archaeal-dominated sediments. The metagenomes were analyzed using a suite of complementary and integrative bioinformatic tools, including phylogenetic and functional analysis of both individual sequence reads and assemblies of predominant phylotypes. This volume identifies major environmental determinants of a large number of thermophilic microbial lineages, many of which have not been fully described in the literature nor previously cultivated to enable functional and genomic analyses. Moreover, protein family abundance comparisons and in-depth analyses of specific genes and metabolic pathways relevant to these hot-spring environments reveal hallmark signatures of metabolic capabilities that parallel the distribution of phylotypes across specific types of geothermal environments.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, CFB - Core Flow, Montana State University, Joint Genome Institute, Indiana University-Purdue
Number of pages: 15
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Journal: Frontiers in Microbiology
Volume: 4
Article number: 67
ISSN (Print): 1664-302X
Ratings:
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 4.16 SJR 1.731 SNIP 1.172
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.878 SNIP 1.208 CiteScore 4.15
Analyzing the genomic variation of microbial cell factories in the era of "New Biotechnology"

The application of genome-scale technologies, both experimental and in silico, to industrial biotechnology has allowed improving the conversion of biomass-derived feedstocks to chemicals, materials and fuels through microbial fermentation. In particular, due to rapidly decreasing costs and its suitability for identifying the genetic determinants of a phenotypic trait of interest, whole genome sequencing is expected to be one of the major driving forces in industrial biotechnology in the coming years. We present some of the recent studies that have successfully applied high-throughput sequencing technologies for finding the underlying molecular mechanisms for (a) improved carbon source utilization, (b) increased product formation, and (c) stress tolerance. We also discuss the strengths and weaknesses of different strategies for mapping industrially relevant genotype-to-phenotype links including exploiting natural diversity in natural isolates or crosses between isolates, classical mutagenesis and evolutionary engineering.

General information
State: Published
Organisations: Novo Nordisk Foundation Center for Biosustainability, CFB - Core Flow, CFB - Metagenomic Systems Biology, Center for Biological Sequence Analysis
Authors: Herrgard, M. (Intern), Panagiotou, G. (Intern)
Number of pages: 8
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Journal: Computational and Structural Biotechnology Journal
Volume: 3
Issue number: 4
ISSN (Print): 2001-0370
Ratings:
Scopus rating (2016): SJR 1.284 SNIP 1.043 CiteScore 3.16
Scopus rating (2015): SJR 0.784 SNIP 0.543 CiteScore 2.06
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.478 SNIP 0.37 CiteScore 1.03
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.248 SNIP 0.208
Original language: English
Electronic versions:
DOIs:
10.5936/csbj.201210012
Choosing the right platform for the right product: Sustainable production of chemicals in microbial cell factories

The Novo Nordisk Foundation Center for Biosustainability (CFB) is a new non-profit research center focused on sustainable production of biochemicals and therapeutic proteins using microbial and mammalian cell factories. The work at CFB is organized around an iterative loop where cell factories are taken through four stages: 1) In silico design and data analysis, 2) genome engineering, 3) characterization and high throughput screening, and 4) omics based systems analysis. However, before entering this loop, significant two significant decisions need to be made. First, the choice of what products should be made should be made considering biological, technical, economic, and ecological factors. Second, the choice which platform strains will be used to develop production hosts should be made based on characterization of both metabolic and process tolerance traits of strains. I will discuss the role of modeling and systems biology approaches in guiding both the product and strain selection processes.

Bibliographical note
Markus Herrgard
Conference presentation at
Chemical Engineering at the Life Science Interface Conference Sheffield, UK
November 12, 2012
Source: dtu

Metagenomes from High-Temperature Chemotrophic Systems Reveal Geochemical Controls on Microbial Community Structure and Function

The Yellowstone caldera contains the most numerous and diverse geothermal systems on Earth, yielding an extensive array of unique high-temperature environments that host a variety of deeply-rooted and understudied Archaea, Bacteria and Eukarya. The combination of extreme temperature and chemical conditions encountered in geothermal environments often results in considerably less microbial diversity than other terrestrial habitats and offers a tremendous opportunity for studying the structure and function of indigenous microbial communities and for establishing linkages between putative metabolisms and element cycling. Metagenome sequence (14-15,000 Sanger reads per site) was obtained for five high-temperature (>65 degrees C) chemotrophic microbial communities sampled from geothermal springs (or pools) in Yellowstone National Park (YNP) that exhibit a wide range in geochemistry including pH, dissolved sulfide, dissolved oxygen and ferrous iron. Metagenome data revealed significant differences in the predominant phyla associated with each of these geochemical environments. Novel members of the Sulfolobales are dominant in low pH environments, while other Crenarchaeota including distantly-related Thermoproteales and Desulfurococccales populations dominate in suboxic sulfidic sediments. Several novel archaeal groups are well represented in an acidic (pH 3) Fe-oxhydroxide mat, where a higher O-2 influx is accompanied with an increase in archaeal diversity. The presence or absence of genes and pathways important in S oxidation-reduction, H-2-oxidation, and aerobic respiration (terminal oxidation) provide insight regarding the metabolic strategies of indigenous organisms present in geothermal systems. Multiple-pathway and protein-specific functional analysis of metagenome sequence data corroborated results from phylogenetic analyses and clearly demonstrate major differences in metabolic potential across sites. The distribution of functional genes involved in electron transport is consistent with the hypothesis that geochemical parameters (e.g., pH, sulfide, Fe, O-2) control microbial community structure and function in YNP geothermal springs.

General information
State: Published
Organisations: J. Craig Venter Institute, Montana State University, La Jolla Institute for Allergy & Immunology, Ehime University, University of Illinois, Idaho National Laboratory, University of South Alabama
Pages: Article No.: e9773
Connecting extracellular metabolomic measurements to intracellular flux states in yeast

Background: Metabolomics has emerged as a powerful tool in the quantitative identification of physiological and disease-induced biological states. Extracellular metabolome or metabolic profiling data, in particular, can provide an insightful view of intracellular physiological states in a noninvasive manner.

Results: We used an updated genome-scale metabolic network model of Saccharomyces cerevisiae, iMM904, to investigate how changes in the extracellular metabolome can be used to study systemic changes in intracellular metabolic states. The iMM904 metabolic network was reconstructed based
on an existing genome-scale network, iND750, and includes 904 genes and 1,412 reactions. The network model was first validated by comparing 2,888 in silico single-gene deletion strain growth phenotype predictions to published experimental data. Extracellular metabolome data measured in response to environmental and genetic perturbations of ammonium assimilation pathways was then integrated with the iMM904 network in the form of relative overflow secretion constraints and a flux sampling approach was used to characterize candidate flux distributions allowed by these constraints. Predicted intracellular flux changes were consistent with published measurements on intracellular metabolite levels and fluxes. Patterns of predicted intracellular flux changes could also be used to correctly identify the regions of the metabolic network that were perturbed.

Conclusion: Our results indicate that integrating quantitative extracellular metabolomic profiles in a constraint-based framework enables inferring changes in intracellular metabolic flux states. Similar methods could potentially be applied towards analyzing biofluid metabolome variations related to human physiological and disease states.
Decomposing complex reaction networks using random sampling, principal component analysis and basis rotation

ABSTRACT: BACKGROUND: Metabolism and its regulation constitute a large fraction of the molecular activity within cells. The control of cellular metabolic state is mediated by numerous molecular mechanisms, which in effect position the metabolic network flux state at specific locations within a mathematically-definable steady-state flux space. Post-translational regulation constitutes a large class of these mechanisms, and decades of research indicate that achieving a network flux state through post-translational metabolic regulation is both a complex and complicated regulatory problem. No analysis method for the objective, top-down assessment of such regulation problems in large biochemical networks has been presented and demonstrated. RESULTS: We show that the use of Monte Carlo sampling of the steady-state flux space of a cell-scale metabolic system in conjunction with Principal Component Analysis and eigenvector rotation results in a low-dimensional and biochemically interpretable decomposition of the steady flux states of the system. This decomposition comes in the form of a low number of small reaction sets whose flux variability accounts for nearly all of the flux variability in the entire system. This result indicates an underlying simplicity and implies that the regulation of a relatively low number of reaction sets can essentially determine the flux state of the entire network in the given growth environment. CONCLUSION: We demonstrate how our top-down analysis of networks can be used to determine key regulatory requirements independent of specific parameters and mechanisms. Our approach complements the reductionist approach to elucidation of regulatory mechanisms and facilitates the development of our understanding of global regulatory strategies in biological networks.
Reconstruction of biochemical networks in microorganisms.

Systems analysis of metabolic and growth functions in microbial organisms is rapidly developing and maturing. Such studies are enabled by reconstruction, at the genomic scale, of the biochemical reaction networks that underlie cellular processes. The network reconstruction process is organism specific and is based on an annotated genome sequence, high-throughput network-wide data sets and bibliomic data on the detailed properties of individual network components. Here we describe the process that is currently used to achieve comprehensive network reconstructions and discuss how these reconstructions are curated and validated. This review should aid the growing number of researchers who are carrying out reconstructions for particular target organisms.

General information
State: Published
Organisations: University of California, University of Wisconsin-Madison, University of California, San Diego
Authors: Feist, A. (Intern), Herrgard, M. (Intern), Thiele, I. (Ekstern), Reed, J. L. (Ekstern), Palsson, B. Ø. (Ekstern)
Pages: 129-143
Publication date: 2009
Main Research Area: Technical/natural sciences

Publication information
Journal: Nature Reviews. Microbiology
Volume: 7
Issue number: 2
ISSN (Print): 1740-1526
Ratings:
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 15.625 SNIP 5.952 CiteScore 10.91
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 12.689 SNIP 5.263 CiteScore 9.88
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 10.04 SNIP 4.592 CiteScore 10.08
ISI indexed (2013): ISI indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 7.497 SNIP 3.492 CiteScore 10.85
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 6.368 SNIP 2.924 CiteScore 12.23
A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology

Genomic data allow the large-scale manual or semi-automated assembly of metabolic network reconstructions, which provide highly curated organism-specific knowledge bases. Although several genome-scale network reconstructions describe Saccharomyces cerevisiae metabolism, they differ in scope and content, and use different terminologies to describe the same chemical entities. This makes comparisons between them difficult and underscores the desirability of a consolidated metabolic network that collects and formalizes the 'community knowledge' of yeast metabolism. We describe how we have produced a consensus metabolic network reconstruction for S. cerevisiae. In drafting it, we placed special emphasis on referencing molecules to persistent databases or using database-independent forms, such as SMILES or InChI strings, as this permits their chemical structure to be represented unambiguously and in a manner that permits automated reasoning. The reconstruction is readily available via a publicly accessible database and in the Systems Biology Markup Language (http://www.comp-sys-bio.org/yeastnet). It can be maintained as a resource that serves as a common denominator for studying the systems biology of yeast. Similar strategies should benefit communities studying genome-scale metabolic networks of other organisms.
Aerobic fermentation of D-glucose by an evolved cytochrome oxidase-deficient Escherichia coli strain

Fermentation of glucose to D-lactic acid under aerobic growth conditions by an evolved Escherichia coli mutant deficient in three terminal oxidases is reported in this work. Cytochrome oxidases (cydAB, cyoABCD, and cbdAB) were removed from the E. coli K12 MG1655 genome, resulting in the ECOM3 (E. coli cytochrome oxidase mutant) strain. Removal of cytochrome oxidases reduced the oxygen uptake rate of the knockout strain by nearly 85%. Moreover, the knockout strain was initially incapable of growing on M9 minimal medium. After the ECOM3 strain was subjected to adaptive evolution on glucose M9 medium for 60 days, a growth rate equivalent to that of anaerobic wild-type E. coli was achieved. Our findings demonstrate that three independently adaptively evolved ECOM3 populations acquired different phenotypes: one produced lactate as a sole fermentation product, while the other two strains exhibited a mixed-acid fermentation under...
oxic growth conditions with lactate remaining as the major product. The homofermenting strain showed a D-lactate yield of 0.8 g/g from glucose. Gene expression and in silico model-based analyses were employed to identify perturbed pathways and explain phenotypic behavior. Significant upregulation of ygiN and sodAB explains the remaining oxygen uptake that was observed in evolved ECOM3 strains. E. coli strains produced in this study showed the ability to produce lactate as a fermentation product from glucose and to undergo mixed-acid fermentation during aerobic growth. ©American Society for Microbiology. All rights reserved.
Impact of individual mutations on increased fitness in adaptively evolved strains of Escherichia coli

We measured the relative fitness among a set of experimentally evolved Escherichia coli strains differing by a small number of adaptive mutations by directly measuring allelic frequencies in head-to-head competitions using a mass spectrometry-based method. We compared the relative effects of mutations to the same or similar genes acquired in multiple strains when expressed in allele replacement strains. We found that the strongest determinant of fitness among the evolved strains was the impact of beneficial mutations to the RNA polymerase beta and beta' subunit genes. We also identified several examples of epistatic interactions between rpoB/C and glpK mutations and identified two other mutations that are beneficial only in the presence of previously acquired mutations but that have little or no adaptive benefit to the wild-type strain. Allele frequency estimation is shown to be a highly sensitive method for measuring selection rates during competitions between strains differing by as little as a single-nucleotide polymorphism and may be of great use for investigating epistatic interactions.
Network-based prediction of human tissue-specific metabolism.

Direct in vivo investigation of mammalian metabolism is complicated by the distinct metabolic functions of different tissues. We present a computational method that successfully describes the tissue specificity of human metabolism on a large scale. By integrating tissue-specific gene- and protein-expression data with an existing comprehensive reconstruction of the global human metabolic network, we predict tissue-specific metabolic activity in ten human tissues. This reveals a central role for post-transcriptional regulation in shaping tissue-specific metabolic activity profiles. The predicted tissue specificity of genes responsible for metabolic diseases and tissue-specific differences in metabolite exchange with biofluids extend markedly beyond tissue-specific differences manifest in enzyme-expression data, and are validated by large-scale mining of tissue-specificity data. Our results establish a computational basis for the genome-wide study of normal and abnormal human metabolism in a tissue-specific manner.
Microbial regulatory and metabolic networks

Reconstruction of transcriptional regulatory and metabolic networks is the foundation of large-scale microbial systems and synthetic biology. An enormous amount of information including the annotated genomic sequences and the genomic locations of DNA-binding regulatory proteins can be used to define metabolic and regulatory networks in cells. In particular, advances in experimental methods to map regulatory networks in microbial cells have allowed reliable data-driven reconstruction of these networks. Recent work on metabolic engineering and experimental evolution of microbes highlights the key role of global regulatory networks in controlling specific metabolic processes and the need to consider the integrated function of multiple types of networks for both scientific and engineering purposes.

General information
State: Published
Organisations: University of California, San Diego
Authors: Cho, B. (Ekstern), Charusanti, P. (Ekstern), Herrgard, M. (Intern), Palsson, B. ∅. (Ekstern)
Pages: 360-364
Publication date: 2007
Main Research Area: Technical/natural sciences

Publication information
Journal: Current Opinion in Biotechnology
Volume: 18
Issue number: 4
ISSN (Print): 0958-1669
Ratings:
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 8.55 SJR 3.331 SNIP 2.1
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 3.113 SNIP 2.143 CiteScore 7.99
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 3.271 SNIP 2.068 CiteScore 7.45
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 3.322 SNIP 2.198 CiteScore 7.93
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 3.508 SNIP 2.327 CiteScore 7.93
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 3.313 SNIP 2.089 CiteScore 7.76
ISI indexed (2011): ISI indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 3.56 SNIP 2.223
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Scopus rating (2009): SJR 3.772 SNIP 2.085
BFI (2008): BFI-level 2
Scopus rating (2008): SJR 3.324 SNIP 2.009
Scopus rating (2007): SJR 3.058 SNIP 1.959
Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox

The manner in which microorganisms utilize their metabolic processes can be predicted using constraint-based analysis of genome-scale metabolic networks. Herein, we present the constraint-based reconstruction and analysis toolbox, a software package running in the Matlab environment, which allows for quantitative prediction of cellular behavior using a constraint-based approach. Specifically, this software allows predictive computations of both steady-state and dynamic optimal growth behavior, the effects of gene deletions, comprehensive robustness analyses, sampling the range of possible cellular metabolic states and the determination of network modules. Functions enabling these calculations are included in the toolbox, allowing a user to input a genome-scale metabolic model distributed in Systems Biology Markup Language format and perform these calculations with just a few lines of code. The results are predictions of cellular behavior that have been verified as accurate in a growing body of research. After software installation, calculation time is minimal, allowing the user to focus on the interpretation of the computational results.
Systematic condition-dependent annotation of metabolic genes

The task of deriving a functional annotation for genes is complex as their involvement in various processes depends on multiple factors such as environmental conditions and genetic backup mechanisms. This study employs a large-scale model of the metabolism of Saccharomyces cerevisiae to investigate the function of yeast genes and derive a condition-dependent annotation (CDA) for their involvement in major metabolic processes under various genetic and environmental conditions. The resulting CDA is validated on a large scale and is shown to be superior to the corresponding Gene Ontology (GO) annotation, by showing that genes annotated with the same CDA term tend to be more coherently conserved in evolution and display greater expression coherency than those annotated with the same GO term. The CDA gives rise to new kinds of functional condition-dependent metabolic pathways, some of which are described and further examined via substrate auxotrophy measurements of knocked-out strains. The CDA presented is likely to serve as a new reference source for metabolic gene annotation.
Identification of genome-scale metabolic network models using experimentally measured flux profiles

Genome-scale metabolic network models can be reconstructed for well-characterized organisms using genomic annotation and literature information. However, there are many instances in which model predictions of metabolic fluxes are not entirely consistent with experimental data, indicating that the reactions in the model do not match the active reactions in the in vivo system. We introduce a method for determining the active reactions in a genome-scale metabolic network based on a limited number of experimentally measured fluxes. This method, called optimal metabolic network identification (OMNI), allows efficient identification of the set of reactions that results in the best agreement between in silico predicted and experimentally measured flux distributions. We applied the method to intracellular flux data for evolved Escherichia coli mutant strains with lower than predicted growth rates in order to identify reactions that act as flux bottlenecks in these strains. The expression of the genes corresponding to these bottleneck reactions was often found to be downregulated in the evolved strains relative to the wild-type strain. We also demonstrate the ability of the OMNI method to diagnose problems in E. coli strains engineered for metabolite overproduction that have not reached their predicted production potential. The OMNI method applied to flux data for evolved strains can be used to provide insights into mechanisms that limit the ability of microbial strains to evolve towards their predicted optimal growth phenotypes. When applied to industrial production strains, the OMNI method can also be used to suggest metabolic engineering strategies to improve byproduct secretion. In addition to these applications, the method should prove to be useful in general for reconstructing metabolic networks of ill-characterized microbial organisms based on limited amounts of experimental data.

General information
State: Published
Organisations: Technical University of Denmark, University of California
Authors: Herrgard, M. (Intern), Fong, S. S. (Ekstern), Palsson, B. Ø. (Ekstern)
Number of pages: 11
Pages: e72
Publication date: 2006
Main Research Area: Technical/natural sciences

Publication information
Journal: P L o S One
Volume: 2
Issue number: 7
ISSN (Print): 1932-6203
Ratings:
Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae

We describe the use of model-driven analysis of multiple data types relevant to transcriptional regulation of metabolism to discover novel regulatory mechanisms in Saccharomyces cerevisiae. We have reconstructed the nutrient-controlled transcriptional regulatory network controlling metabolism in S. cerevisiae consisting of 55 transcription factors regulating 750 metabolic genes, based on information in the primary literature. This reconstructed regulatory network coupled with an existing genome-scale metabolic network model allows in silico prediction of growth phenotypes of regulatory gene deletions as well as gene expression profiles. We compared model predictions of gene expression changes in response to genetic and environmental perturbations to experimental data to identify potential novel targets for transcription factors. We then identified regulatory cascades connecting transcription factors to the potential targets through a systematic model expansion strategy using published genome-wide chromatin immunoprecipitation and binding-site-motif data sets. Finally, we show the ability of an integrated metabolic and regulatory network model to predict growth phenotypes of transcription factor knockout strains. These studies illustrate the potential of model-driven data integration to systematically discover novel components and interactions in regulatory and metabolic networks in eukaryotic cells.
Bioinformatics Services for Data-Driven Design of Cell Factories and Communities

Omics data is not leveraged effectively in the biotechnology industry due to lack of tools to rapidly access public and private data and to design cellular manipulations or interventions based on the data. With this project we aim to make a broad spectrum of omics data useful to the biotechnology industry covering application areas ranging from industrial biotechnology to human health. We will develop novel approaches for integrative model-based omics data analysis to enable 1) Identification of novel enzymes and pathways by mining metagenomic data, 2) Data-driven design of cell factories for the production of chemicals and proteins, and 3) Analysis and design of microbial communities relevant to human health, industrial biotechnology and agriculture. All research efforts will be integrated in an interactive web-based platform that will be available for the industrial and academic research and development communities, in particular enhancing the competitiveness of biotech SMEs by economizing resources and reducing time-to-market within their respective focus areas. The platform will be composed of standardized and interoperable components that service-oriented bioinformatics SMEs involved in the project can reuse in their own products. An important aspect of the platform will be implementation of different access levels to data and software tools allowing controlling access to proprietary data and analysis tools. Two end-user companies will be involved in practical testing of the platform built within the project using proprietary omics data generated at the companies.

Novo Nordisk Foundation Center for Biosustainability

iLoop

Period: 01/03/2016 → 29/02/2020
Number of participants: 8
cell factories, microbial communities, synthetic biology, systems biology
Acronym: DD-DeCaF
Project participant:
Galkina, Svetlana (Intern)
Redestig, Nils Henning (Intern)
Beber, Moritz Emanuel (Intern)
Dannaher, Danny (Intern)
Project Manager, organisational:
Lohmann, Ricarda (Intern)
Rasmussen, Birte Kastrup (Intern)
Project Coordinator:
Herrgard, Markus (Intern)
Sonnenschein, Nikolaus (Intern)

Financing sources
Source: EU research programme (public)
Name of research programme: Horizon 2020 LEIT BIO
Functional investigations of cell wall alterations in chemical-evolved E. coli strains
One of the largest barriers to achieving economical bio-based production of bulk chemicals such as biofuels and polymer precursors is poor tolerance of microbial production hosts toward high concentrations of excreted product. These concentrations are often in excess of 100 g/L in order to minimize capital and downstream purification costs. However virtually all chemicals at these levels result in stresses and poor growth in the majority of microbial hosts, which can decrease product yields and productivities. To help address this issue, we utilized a robotic platform to evolve parallel populations of Escherichia coli K-12 MG1655 for enhanced growth in the presence of toxic concentrations of 11 chemicals representing diverse functional classes that are of interest as biofuels or their precursors, polymer precursors, and other bulk chemicals and intermediates. Resequencing of over 200 strains and subsequent reconstruction of sets of mutations has provided unparalleled insight on the genomic basis of tolerance. In addition to more specific mechanisms for individual chemicals or classes of chemicals, many broader mechanisms of tolerance have been putatively identified that recur in strains evolved on different chemicals.

One class of common mutations across chemical conditions are coding mutations in genes related to cell wall biogenesis, maintenance, and recycling. It is suspected that many of the strains harboring these mutations feature altered cell morphologies and altered membrane protein and lipid compositions. In order to understand the connection between genotype and phenotype for cell wall mutations, it is proposed to conduct work at EMSL to further characterize the phenotype of a subset of evolved strains with confirmed morphological changes. The proposed tests include using cryogenic transmission electron microscopy and helium ion microscopy to observe cross-sectional and surface modifications of single cells, and performing differential membrane proteomics and lipidomics analyses.
The data obtained from this study will be used to develop further targeted tests on strains with cell wall mutations, and will ultimately be integrated together with other datasets concerning other types of mutations to develop predictive models of chemical and stress tolerance. The direct effect of cell wall mutations on endogenous production and excretion of relevant chemicals will also be tested by employing them directly in engineered production host strains.

Novo Nordisk Foundation Center for Biosustainability

Research Groups

iLoop

Pacific Northwest National Laboratory
Period: 01/10/2015 → 30/09/2016
Number of participants: 2
Project Manager, academic:
Lennen, Rebecca (Intern)
Herrgard, Markus (Intern)

Assessing Life Cycle Impacts of Bioplastics from Dicarboxylic Acids

Department of Management Engineering
Period: 01/10/2015 → 30/09/2018
Number of participants: 3
Phd Student:
Ógmundarson, Ólafur (Intern)
Supervisor:
Herrgard, Markus (Intern)
Main Supervisor:
Fantke, Peter (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Generation of Raw Substrate Utilizing Platform Strains

Novo Nordisk Foundation Center for Biosustainability

Research Groups

iLoop

Network Reconstruction in Silico Biology
Period: 01/06/2015 → 30/04/2019
Number of participants: 4
Phd Student:
Tharwat Tolba Mohamed, Elsayed (Intern)
Supervisor:
Lennen, Rebecca (Intern)
Feist, Adam (Intern)
Main Supervisor:
Herrgard, Markus (Intern)

Generation of Raw Substrate Utilizing Platform Strains

Technical University of Denmark
Period: 01/06/2015 → 31/05/2018
Number of participants: 4
Phd Student:
Tharwat Tolba Mohamed, Elsayed (Intern)
Supervisor:
Feist, Adam (Intern)
Lennen, Rebecca (Intern)
Main Supervisor:
Herrgard, Markus (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Genome-Scale Modeling of Protein Biosynthesis in Methylotrophic Bacteria involved in Animal Feed Production from Natural Gas
Technical University of Denmark
Period: 01/03/2015 → 28/02/2018
Number of participants: 4
Phd Student:
Lieven, Christian (Intern)
Supervisor:
Herrgard, Markus (Intern)
Sonnenschein, Nikolaus (Intern)
Main Supervisor:
Gernaey, Krist V. (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Mass Action Stoichiometric Simulation for Cell Factory Design
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop
Department of Systems Biology
Network Engineering of Eukaryotic Cell Factories
Period: 01/10/2014 → 30/09/2017
Number of participants: 4
Phd Student:
Matos, Marta (Intern)
Supervisor:
Herrgard, Markus (Intern)
Sonnenschein, Nikolaus (Intern)
Main Supervisor:
Andersen, Mikael Rørdam (Intern)
Project

Integration of re-sequencing data into Escherichia coli and Saccharomyces cerevisiae genome-scale metabolic models
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop
Department of Systems Biology
Network Engineering of Eukaryotic Cell Factories
Period: 01/10/2014 → 30/09/2017
Number of participants: 4
Phd Student:
Integration of re-sequencing data into Escherichia coli and Saccharomyces cerevisiae genome-scale metabolic models

Technical University of Denmark
Period: 01/10/2014 → 30/09/2017
Number of participants: 7
Phd Student:
Cardoso, Joao (Intern)
Supervisor:
Herrgard, Markus (Intern)
Sonnenschein, Nikolaus (Intern)
Main Supervisor:
Andersen, Mikael Rørdam (Intern)
Examiner:
Nikel, Pablo Ivan (Intern)
Patil, Kiran Raosaheb (Intern)
Penttilä, Merja Elisa (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Mass Action Stoichiometric Simulation for Cell Factory Design

Technical University of Denmark
Period: 01/10/2014 → 30/09/2017
Number of participants: 7
Phd Student:
Matos, Marta (Intern)
Supervisor:
Herrgard, Markus (Intern)
Sonnenschein, Nikolaus (Intern)
Main Supervisor:
Andersen, Mikael Rørdam (Intern)
Examiner:
Nielsen, Lars Keld (Intern)
Kummer, Ursula (Ekstern)
Ryde, Ulf Sigurd Bror (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Utilize Gasified Biomass and Waste Products (Synthesis Gas) for the Production of Biochemicals

Department of Systems Biology
Period: 01/05/2013 → 29/09/2016
Number of participants: 6
Phd Student:
Redl, Stephanie Maria Anna (Intern)
Supervisor:
Förster, Jochen (Intern)
Main Supervisor:
Nielsen, Alex Toftgaard (Intern)
Examiner:
Herrgard, Markus (Intern)
Simpson, Sean D. (Ekstern)
Soucaille, Philippe (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Marie Curie (EU-stipendium)

Relations
Publications:
Gas Fermentation using Thermophilic Moorella Species for production of Biochemicals
Project: PhD

Investigating of stoichiometric and thermodynamic operating principles in cellular metabolic networks
Department of Systems Biology
Period: 01/11/2009 → 19/03/2013
Number of participants: 6
Phd Student:
Zelezniak, Aleksej (Intern)
Supervisor:
Patil, Kiran Raosaheb (Intern)
Main Supervisor:
Kielland-Brandt, Morten (Intern)
Examiner:
Herrgard, Markus (Intern)
Nielsen, Jens (Intern)
Typas, Athanasios (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Grundforskningsfonden
Project: PhD

Activities:

Data-driven Biotechnology
Period: 09 May 2017 → 10 May 2017
Joao Cardoso (Participant)
Ahmad A. Zeidan (Participant)
Markus Herrgard (Participant)
Nikolaus Sonnenschein (Participant)
Novo Nordisk Foundation Center for Biosustainability
iLoop
Research Groups
Global Econometric Modeling

Description
In silico Identification of metabolite analogues for rational strain Improvement

Related event
Data-driven Biotechnology: Bench, Bioreactor and Bedside
07/05/2017 → 11/05/2017
Hillerød, Denmark
Activity: Attending an event › Participating in or organising a conference

4th Conference in Constraint-based Reconstruction and Analysis
Period: 16 Sep 2015 → 18 Sep 2015
Markus Herrgard (Organizer)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop

Description
COBRA 2015

Related event

4th Conference in Constraint-based Reconstruction and Analysis
16/09/2015 → 18/09/2015
Heidelberg, Germany
Activity: Attending an event › Participating in or organising a conference

Developing an integrated design tool: Recent progress and remaining challenges
Period: 9 Aug 2015
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop

Related event

Advances in Design and Use of Microbial Production Systems: BWC Community Workshop
09/08/2015 → …
Geneva, Switzerland
Activity: Talks and presentations › Conference presentations

Multi-scale modeling of chemical product choices for cell factory development
Period: 8 Jun 2015
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop

Related event

Metabolic Pathway Analysis 2015
08/06/2015 → 12/06/2015
Braga, Portugal
Activity: Talks and presentations › Conference presentations

Using adaptive laboratory evolution to create platform strains for chemical production
Period: 20 Mar 2015
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop
Related event

Seminar
20/03/2015 → …
Leuven, Belgium
Activity: Talks and presentations › Conference presentations

Frontiers in Bioengineering and Biotechnology (Journal)
Period: 1 Jan 2015 → 26 Nov 2015
Markus Herrgard (Editor)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
iLoop

Description
Current Challenges in Modeling Cellular Metabolism

Related journal
Frontiers in Bioengineering and Biotechnology
Scopus rating (2016): CiteScore 1 SJR 0.169 SNIP 0.471, Web of Science (2017): Indexed yes
Indexed in DOAJ
Central database
Activity: Research › Journal editor

Choosing the right product and platform strain for sustainable production of biochemicals
Period: 18 Nov 2014
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability
iLoop

Related event
2013 Congress of The Danish Microbiological Society
18/11/2013 → 18/11/2013
Copenhagen, Denmark
Activity: Talks and presentations › Conference presentations

Genome-scale models of cellular metabolism – from basic science to industrial applications
Period: 11 Sep 2014
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability
iLoop

Related event
KILU Day
11/09/2014 → …
Lund, Sweden
Activity: Talks and presentations › Conference presentations

International Synthetic and Systems Biology Summer School
Period: 15 Jun 2014 → 19 Jun 2014
Markus Herrgard (Organizer)
Novo Nordisk Foundation Center for Biosustainability
Research Groups
Related event

International Synthetic and Systems Biology Summer School
15/06/2014 → 19/06/2014
Taormina, Italy
Activity: Attending an event › Participating in or organising a conference

Systems biology and metabolic modeling for sustainable production of chemicals
Period: 25 Mar 2014
Markus Herrgard (Invited speaker)
Novo Nordisk Foundation Center for Biosustainability

Related event

ChEBI User Workshop: Harnessing ChEBI for systems biology and metabolic modelling
25/03/2014 → 26/03/2014
Hinxton, United Kingdom
Activity: Talks and presentations › Conference presentations