Understanding Fundamental Aspects of Butanol production by *Clostridium beijerinckii*

Caroline B. Milne[1,2], Yi Wang[1], Beth Papanek[1] Yong-Su Jin[1], Nathan D. Price[1,2] and **Hans P. Blaschek[1]**

[1] University of Illinois, Urbana IL; [2] Institute for Systems Biology, Seattle WA

http://bioenergy.illinois.edu

Project Goals: The solventogenic clostridia offer a sustainable approach to petroleum-based production of n-butanol, an important chemical feedstock and potential fuel. With the availability of the genome sequence for *Clostridium beijerinckii* 8052, we can now employ the tools of systems biology in order to gain increased insight into the metabolic and regulatory networks relevant to solvent production. Project goals include examination of: 1) the mutations underlying the *C. beijerinckii* BA101 butanol-overproduction phenotype, 2) the molecular basis for the global shift from acidogenesis to solventogenesis, 3) the genetic basis of butanol tolerance in *C. beijerinckii* and 4) RNA-seq technology for single-nucleotide resolution analysis of the transcriptome of this microorganism.

We have constructed the first genome-scale metabolic model (*i*CM925) for *C. beijerinckii*, containing 925 genes, 938 reactions, and 881 metabolites. To build the model we employed a semi-automated procedure that integrated genome annotation information from KEGG, BioCyc, and The SEED, and utilized computational algorithms with manual curation to improve model completeness. To validate *i*CM925, we conducted fermentation experiments using the NCIMB 8052 strain, and evaluated the ability of the model to simulate measured substrate uptake and product production rates. Experimentally observed fermentation profiles were found to lie within the solution space of the model; however, under an optimal growth objective, additional constraints were needed to reproduce the observed profiles--suggesting the existence of selective pressures other than optimal growth. Notably, a significantly enriched fraction of actively utilized reactions in simulations--constrained to reflect experimental rates--originated from the set of reactions that overlapped between all three annotation databases used ($P = 3.52 \times 10^{-9}$, Fisher's exact test). Inhibition of the hydrogenase reaction was found to have a strong effect on butanol formation--as experimentally observed.

We also conducted a single-nucleotide resolution analysis of the *C. beijerinckii* NCIMB 8052 transcriptome using high-throughput RNA-Seq technology. We identified the transcription start sites and operon structure throughout the genome. We confirmed the structure of important gene operons involved in metabolic pathways for acid and solvent production in *C. beijerinckii* 8052, including *pta-ack*, *ptb-buk*, *hbd-etfA-etfB-crt* (*bcs*) and *ald-ctfA-ctfB-adc* (*sol*) operons; we also defined important operons related to chemotaxis/motility, transcriptional regulation, stress response and fatty acids biosynthesis along with others. We discovered 20 previously non-annotated regions with significant transcriptional activities and 15 genes whose translation start codons were

likely mis-annotated. As a consequence, the accuracy of existing genome annotation was significantly enhanced. Furthermore, we identified 78 putative silent genes and 177 putative housekeeping genes based on normalized transcription measurement with the sequence data. We also observed that more than 30% of pseudogenes had significant transcriptional activities during the fermentation process. Strong correlations exist between the expression values derived from RNA-Seq analysis and microarray data or qRT-PCR results.

References

Milne, C.B, et al. 2011. Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052. BMC Systems Biology. 5:130.

Wang, Yi., et al. 2011. Single-nucleotide resolution analysis of the transcriptome structure of *Clostridium beijerinckii* NCIMB 8052 using RNA-Seq. BMC Genomics, 12:479.

This work was funded by the DOE Office of Science through FOA: DE_FOA-0000368 and an NSF CAREER grant.