A SYNTHETIC BIOLOGY BASED CELL LINE ENGINEERING PIPELINE

Christopher S. Stach, University of Minnesota, Minneapolis, USA
Sofie O'Brien, Tung Le, Arpan Bandyopadhyay, Meghan McCann, Conor O'Brien, Zion Lee, Prodromos Daoutidis, Wei-Shou Hu, Nikunj Somia, Michael J. Smanski, University of Minnesota, Minneapolis, USA
Xinning Chen, East China University of Science and Technology, China
Liang Zhao, East China University of Science and Technology, China

Keywords: Synthetic biology, glyco-engineeering, combinatorial design, genome sequencing, ATAC-seq

An ideal host cell line for deriving cell lines of high recombinant protein production should be stable, predictable, and amenable to rapid cell engineering or other forms of phenotypical manipulation. In the past few years we have employed genomic information to identify "safe harbors" for exogenous gene integration in CHO cells, deployed systems modeling and optimization to design pathways and control strategies to modify important aspects of recombinant protein productivity, and established a synthetic biology approach to implement genetic changes, all with the goal of creating a pipeline to produce "designer" cell lines.

Chinese hamster ovary (CHO) cells are the preferred platform for protein production. However, the Chinese hamster genome is unstable in its ploidy, is subject to long and short deletions, duplications, and translocations. In addition, gene expression is subject to epigenetic changes including DNA methylation, histone modification and heterochromatin invasion, thus further complicating transgene expression for protein production in cell lines. With these issues in mind, we set out to engineer a CHO cell line highly amenable to stable protein production using a synthetic biology approach. We compiled karyotyping and chromosome number data of several CHO cell lines and sublines, identified genomic regions with high a frequency of gain and loss of copy number using comparative genome hybridization (CGH), and verified structural variants using sequencing data. We further used ATAC (Assay for Transposase-Accessible Chromatin) sequencing to study chromatin accessibility and epigenetic stability within the CHO genome. RNA-seq data from multiple cell lines were also used to identify regions with high transcriptional activity. Analysis of these data allowed the identification of several "safe harbor" loci that could be used for cell engineering.

Based on results of the data analysis and identification of "safe harbors", we engineered an IgG producing cell line with a single copy of the product transgene as a template cell line. This product gene site is flanked by sequences for recombinase mediated cassette exchange, therefore allowing easy substitution of the IgG producing gene for an alternative product gene. Furthermore, a "landing pad" for multi-gene cassette insertion was integrated into the genome at an additional site. Together, these sites allowed engineering of new cell lines producing a fusion protein and Erythropoietin to be generated from the template cell line. To enable rapid assembly of product transgenes and genetic elements for engineering cell attributes into multi-gene cassettes, we adopted a golden-gate based synthetic biology approach. The assembly of genetic parts into multi-gene cassettes in a LEGO-like fashion allowed different combinations of genes under the control of various promoters to be generated quickly for introduction into the template cell line.

Using this engineered CHO cell line, we set out to study metabolism and product protein glycosylation for cell engineering. To guide the selection of genetic elements for cell engineering, we developed a multi-compartment kinetic model, as well as a flux model of energy metabolism and glycosylation. The transcriptome meta-data was used extensively to identify genes and isoforms expressed in the cell line and to estimate the enzyme levels in the model. The flux model was used to identify and the LEGO-like platform was used to implement the genetic changes that can alter the glycosylation pattern of the IgG produced by the template cell line. Concurrently we employed a systems optimization approach to identify the genetic alterations in the metabolic pathway to guide cell metabolism toward a favorable state. The model prediction is being implemented experimentally using the synthetic biology approach.

In conclusion, we have illustrated a pipeline of rational cell line engineering that integrates genomic science, systems engineering and synthetic biology approaches. The promise, the technical challenges and possible limitations will be discussed in this presentation.