Spring 5-12-2016

Mammalian systems biotechnology: An integrative framework for combining in silico modeling and multi-Omics datasets in different CHO parental cell lines

Dong Yup Lee
National University of Singapore, cheld@nus.edu.sg

Sarantos Kyriakopoulos
Biopolis

Meiyappan Lakshmanan
Biopolis

Kok Siong Ang
Biopolis

Alison Lee
Biopolis

See next page for additional authors

Authors
Dong Yup Lee, Sarantos Kyriakopoulos, Meiyappan Lakshmanan, Kok Siong Ang, Alison Lee, Xuezhi Bi, Andy Tan, Ying Swan Ho, Peiqing Zhang, Yuansheng Yang, and Say Kong Ng

Mammalian systems biotechnology: an integrative framework for combining in silico modeling and multi-omics datasets in different CHO parental cell lines

Dong-Yup Lee, Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117585, Singapore & Bioprocessing Technology Institute (BTI), Agency for Science, Technology and Research (A*STAR), 20 Biopolis Way, #06-01 Centros, Singapore 138668, Singapore
cheld@nus.edu.sg
Sarantos Kyriakopoulos, Meiyappan Lakshmanan, Kok Siong Ang, Alison Lee, Xuezhi Bi, Andy Tan, Ying Swan Ho, Peiqing Zhang, Yuansheng Yang, Say Kong Ng, BTI, A*STAR, 20 Biopolis Way, #06-01 Centros, Singapore 138668, Singapore

Key Words: multi-omics analysis, CHO parental cell lines, genome-scale model, systems biotechnology

The increasing availability of multi-omics data from Chinese hamster ovary (CHO) cell cultures entails both opportunity and challenges toward next generation cell culture engineering. Herein, we present a comprehensive and integrative framework to systematically combine trancriptome, proteome, metabolome and glycome datasets in conjunction with a genome-scale metabolic model of CHO cells. We then apply the framework to compare and contrast the metabolic characteristics of the three commonly used parental cell lines (CHO-K1, CHO-DUKXB11 and CHO-DG44) so that "global" attributes of the parental hosts (e.g. growth related characteristics, glycosylation patterns, etc.) could be highlighted (Figure 1).

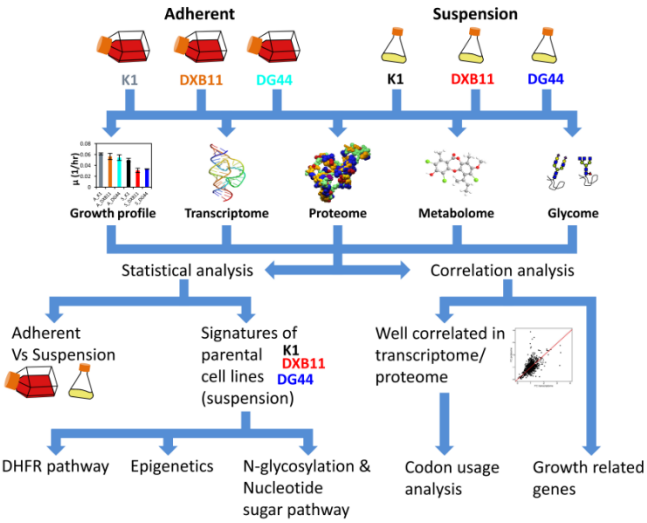

Figure 1 – Integrative framework combining the multi-omics profiling and systems analysis of CHO cell lines

The unique characteristics of the adherent against the suspension cell lines reveal that the latter are in an oxidative stress and that they differentially express genes/proteins associated with the lipid biosynthetic process. The unique transcriptomic and proteomic signatures of the different suspension cell lines, more relevant in an industrial context than the adherent, reflect the known historic divergence of the cell lines, *i.e.* the very different nature of the -DG44 cell line than the other two. Genes/proteins related with the purine nucleotide biosynthetic process (as expected, due to the *Dhfr* gene copy number differences), epigenetic regulation and programmed cell death present the major expression differences between the three parental cell lines. As far as the host N-glycome for each of the cell lines is concerned, it reveals similar profiles. Nevertheless, the cell lines present several differences in the expression of N-glycosylation related genes (*e.g. Man2a1* and *Fut8* are differentially expressed for -DG44 and *Mgat4a* for the -DXB11 cell line) and the pools of nucleotide sugar donors (-K1 presents higher UDP-Glc / UDP-Gal and CMP-sialic acid pools than -DG44; while -DG44 higher GDP-Fuc pools). Growth profiles of the various cell lines were also assessed and our results demonstrate that -K1 cells present significantly higher growth rate than the other two cell lines in suspension culture. Interestingly, adherent cells present a significantly faster growth profile than suspension cells that we attribute to the different media used for the two culture formats, i.e. to the presence of serum for adherent cells.

The integrative framework also involves the use of the genome-scale metabolic model as a scaffold to map the multiomics datasets. Such an analysis allows us to readily pinpoint the heterogeneity in cellular metabolism between the multiple conditions and/or cell lines tested, as well as their correlations. Moreover, the correlation analysis of transcriptome and proteome for a given cell line revealed the plausible regulatory intracellular events that can be targeted for genetic engineering to achieve the enhanced productivity and quality of recombinant proteins in the context of bioprocessing. Interestingly, we identified many differences in the reactions associated with the N-glycan processing pathways for the various parental cell lines analyzed, which may be associated with different glycosylation capacity. Further investigation at the glycomics level may validate our hypothesis that choice of CHO hosts should be product-specific. It is expected that our results can serve as the golden standard for the comprehensive comparison of the various CHO cell lines used worldwide.