

Title of dataset: Transcriptome Meta Data Compilation for Chinese hamster tissues and CHO cell lines

Authors: Vishwanathan, Nandita; Yongky, Andrew; Johnson, Kathryn C; Fu, Hsu-Yuan; Jacob, Nithya M; Le, Huong; Bandyopadhyay, Arpan

DOI of dataset: <http://dx.doi.org/10.13020/D6JS3F>

Transcriptome Meta Data Compilation

Transcript expression levels for tissues and cell lines have been compiled as an excel spreadsheet to allow for a rapid survey of transcript levels of different genes. This data consists of either microarray or RNA sequencing data for 3 tissues (liver, brain, ovary), 3 parental cell lines (DG44, DXB11, CHO-K1) and 16 recombinant cell lines. Sample description is included as a separate sheet in the workbook. The microarray and sequencing data have been linear and upper quartile normalized respectively. The data is organized as shown in the table below and has been color coded for easy visualization. Gene information contains gene description, symbol and orthologous mouse ENSEMBL identifiers. Columns in pink pertaining to 26 different functional pathways can be sorted for the value 1 to quickly assess expression of genes in that pathway. This is followed by RNA sequencing and microarray expression data for tissues and cell lines. The minimum, mean and maximum expression values of each gene for all the cell lines in both microarray and sequencing datasets have also been included.

Group	Columns
Gene Information	Gene Description, Gene Symbol, Orthologous Ensembl Mouse Gene ID
Pathways	Glycolysis gluconeogenesis, TCA cycle, N glycan biosynthesis, Pentose phosphate pathway, Oxidative phosphorylation, Cell cycle, Insulin signaling pathway, Mapk signaling pathway, Fatty acid metabolism, Ribosome, Biosynthesis of unsaturated fatty acids, p53 signaling pathway, Inositol phosphate metabolism, Erbb signaling pathway, Other glycan degradation, Phosphatidylinositol signaling, O glycan biosynthesis, Amino sugar and nucleotide sugar metabolism, Apoptosis, Pyruvate metabolism, mTOR signaling pathway
RNA-Seq Data	Minimum, Mean and Maximum for cell lines, Brain, Liver
Microarray Data	Minimum, Mean and Maximum for cell lines, Brain, Liver, Ovary

The nomenclature for various samples and the availability of the corresponding transcriptome data is summarized in the table below:

<i>Notation</i>	<i>Description (mRNA Source)</i>	<i>RNA-Seq</i>	<i>Microarray</i>
CHO-K1	Parental CHO-K1		Y
DG44	Parental DG44	Y	Y
DXB11	Parental DXB11		Y
rDG_1	Recombinant DG44 (IgG)		Y
rDG_2	Recombinant DG44 (IgG)		Y
rDG_2a	Recombinant DG44 (IgG)		Y
rDG_2b	Recombinant DG44 (IgG)		Y
rDG_3	Recombinant DG44 (GFP)	Y	
rDG_3a	Recombinant DG44 (GFP)	Y	
rDG_4	Recombinant DG44 (IgG)	Y	
rDG_5	Recombinant DG44 (EPO)	Y	
rDG_6	Recombinant DG44 (EPO)	Y	
rDX_1	Recombinant DXB11(IgG)		Y
rDX_1a	Recombinant DXB11(IgG)		Y
rDX_2	Recombinant DXB11(DHFR)		Y
rDX_2a	Recombinant DXB11(DHFR)		Y
rDX_3	Recombinant DXB11(IgG)		Y
rDX_3a	Recombinant DXB11(IgG)		Y
rDX_4	Recombinant DXB11(IgG)		Y
Brain	Chinese hamster brain tissue	Y	Y
Liver	Chinese hamster liver tissue	Y	Y
Ovary	Chinese hamster ovary tissue		Y