

Low Glucose Concentrations within Typical Industrial Operating Conditions Have Minimal Effect on the Transcriptome of Recombinant CHO Cells

Yogender Kumar Gowtham

Dept. of Bioengineering, Clemson University, Clemson, SC 29634

Christopher A. Sasaki

Inst. of Translational Genomics, Clemson University, Clemson, SC 29634

Dept. of Genetics and Biochemistry, Clemson University, Clemson, SC 29634

Sarah W. Harcum 

Dept. of Bioengineering, Clemson University, Clemson, SC 29634

DOI 10.1002/btpr.2462

Published online April 17, 2017 in Wiley Online Library (wileyonlinelibrary.com)

Typically, mammalian cell culture medium contains high glucose concentrations that are analogous to diabetic levels in humans, suggesting that mammalian cells are cultivated in excessive glucose. Using RNA-Seq, this study characterized the Chinese hamster ovary (CHO) cell transcriptome under two glucose concentrations to assess the genetic effects associated with metabolic pathways, in addition to other global responses. The initial extracellular glucose concentrations used represented high (30 mM) and low (10 mM) glucose conditions, where at the time the transcriptomes were compared, the glucose concentrations were approximately 24 and 4.4 mM for the mid-exponential cultures, where 4.4 mM represents a common target concentration in the biopharmaceutical industry for controlled fed-batch cultures. A recombinant CHO cell line producing a monoclonal antibody was used, such that the impact on glycosylation genes could be evaluated. Relatively few genes were identified as being significantly different ($FDR \leq 0.01$) between the high and low glucose conditions, for example, only 575 genes, and only 40 of these genes had 2-fold or greater differences. Gene expression differences for glycolysis, TCA cycle, and glycosylation-related reactions were minimal and unlikely to have biological significance. This transcriptome study indicates that low glucose concentrations in the culture medium are unlikely to cause any biologically significant or detrimental changes to CHO cells at the transcriptome level. Furthermore, it is well-known that maintaining low glucose concentrations in fed-batch cultures can reduce lactate production, which in turn improves process outcomes. Taken together, the transcriptome data supports the continued development of low glucose-based processes to control lactate. © 2017 American Institute of Chemical Engineers Biotechnol. Prog., 33:771–785, 2017

Keywords: Glycolysis, RNA-Seq, glycosylation, gene expression

Introduction

Chinese hamster ovary (CHO) cells have been used for over 25 years in the biopharmaceutical industry to produce recombinant therapeutic proteins, particularly in the manufacture of several recombinant protein therapeutics including antibodies, interferons, blood factors and growth factors.^{1–5} It is estimated that more than 60% of the recombinant protein therapeutics are manufactured using CHO cells, accounting for more than \$100 billion in global revenues.^{3,5–8} Despite the availability of several mammalian cell line options, the selection of CHO

cells as production host is due to several advantages, such as ease of scale up to large-scale processes, well-developed transfection systems for easy recombinant protein expression and the ability to produce human-compatible proteins with proper post-translational modifications.^{4,9,10}

Over the last few years, the industry has realized an increasing demand for life-saving recombinant protein therapeutics, which has invoked a greater need for improving the yield and productivity of the cell culture processes.^{11–13} Among several factors impacting the yield and productivity, energy metabolism of the recombinant cell lines, such as CHO cells, is one of the major factors that dictate the yield and productivity along with the product quality.^{14–16} For most mammalian cell lines including CHO cells, glucose serves as the primary energy and carbon source required for growth and cell maintenance. Several important metabolic pathways such as glycolysis, pentose pathway and TCA cycle are involved in the catabolism of

Current address of Yogender Kumar Gowtham: KBI Biopharma, Research Triangle Park, 2 Triangle Dr., NC 27709

Additional Supporting Information may be found in the online version of this article.

Correspondence concerning this article should be addressed to S. W. Harcum at harcum@clemson.edu

glucose. Additionally, glucose metabolism is central to several key intermediates for the synthesis of cellular components such as amino acids, fatty acids, and nucleic acids.^{14,17,18} Most mammalian cell culture media, including serum-free media, typically contain glucose concentrations from 25 to 35 mM (~4.5 to 6.5 g/L).^{17–19} Glucose concentrations above 10 mM glucose are analogous to the glucose concentrations observed in humans with uncontrolled diabetes, suggesting that most mammalian cell cultures experience excessive glucose concentrations.^{20–22} Moreover, glucose concentrations have been reported to affect several aspects of cell culture including growth rates, glucose consumption rate, and oxygen uptake rate.^{14,23} CHO cell metabolism under high glucose concentrations is characterized by high glucose uptake rates, followed by incomplete oxidation leading to high lactate secretion even in the presence of plentiful of oxygen; a phenomenon known as the Warburg effect or aerobic glycolysis.^{24–29} Lactate, which is considered a metabolic waste by-product, has long been observed to have adverse effects on mammalian cell growth, therapeutic protein productivity and quality.^{15,24–26,29–33} Lactate production in large-scale culture additionally causes osmolarity changes, especially when on-line base additions are used to control the pH decrease invoked by lactate.^{15,34} In the late stages of fed-batch cultures, a metabolic shift toward lactate consumption can occur, alleviating the inhibitory effect of lactate, which in turn improves cell viability and increases protein titers; however, this metabolic shift cannot always be initiated.^{25,35} The lactate shift primarily occurs at a low glucose levels and high lactate concentrations, but can also be observed under high glucose concentrations when lactate concentrations become inhibitory to glycolysis.^{25,36}

Several approaches have been used to minimize lactate production; however, the interconnectedness between glucose and lactate concentrations is still not fully understood. Approaches to control lactate have included substituting glucose with alternate carbon sources, such as galactose,^{37,38} or glucose feeding mechanisms based on sophisticated pH measurement/timing algorithms.^{34,37,39} Cell engineering approaches have been used to control lactate production, unfortunately, these approaches have not been fully successful.²⁹ Other researchers have focused directly on maintaining low glucose concentrations (~5 mM) via controlled fed-batch conditions interfaced with nonlinear models.^{31,33,40–44} Despite these improvements, there are still biphasic outcomes.^{25,35,36,45}

In the early studies to control the glucose concentration in fed-batch cultures (pre-2000), high glucose concentrations in the batch phase were used, which resulted in high lactate (> 20 mM) at the end of the batch phase and glutamine build-up.⁴⁴ Later, glutamine-limited fed-batch cultures were used to limit lactate, where glucose was severely limited.³³ These studies were groundbreaking in approach, yet the very low set-points used for the glutamine and glucose concentrations (ranging from 0.1 to 0.7 mM), combined with the fact that the cultures were actually often well below these already low set-points, make it difficult to extrapolate these data to more moderately low glucose (~5 mM) and glutamine (~6 mM) levels.³³ Consequently, that research group's subsequent gene expression analysis, which suggested the cultures were undergoing apoptosis, may not be predictive of more moderately low glucose (~5 mM) and glutamine (~6 mM) concentrations.⁴⁶ In a separate study conducted by Lu et al. around the same time with anchorage-dependent CHO cells, it was suggested that extracellular glucose

concentration below 1.2 mM glucose, specifically at 0.25 and 0.54 mM glucose, resulted in significant changes in the intracellular concentrations of ATP, amino acids, and TCA cycle metabolites.²³ Otherwise, Lu et al. observed that these intracellular concentrations were very stable over the extracellular glucose concentrations from 1.2 to 9.4 mM glucose.²³ Thus, evaluating gene expression changes in CHO cells at moderately low glucose concentrations in suspension cultures, more representative of non-starvation and common industrial conditions, are necessary for assessing the feasibility of tight feeding control with respect to metabolic control and subsequent effects on product quality attributes.

In this study, the objective was to evaluate the impact of the extracellular glucose concentrations on the CHO cell transcriptome using the next-generation transcriptomic technique RNA-Seq⁴⁷ for glucose concentrations representative of controlled bioreactors. The recombinant CHO cell line, rCHO DP-12, that produces a monoclonal antibody IgG₁ against interleukin-8, was used. Transcriptome profiles were quantified and compared between the rCHO DP-12 cells grown in a chemical-defined medium with high and low initial glucose concentrations – 30 mM and 10 mM glucose, respectively. At the time of the transcriptome analysis, the extracellular concentrations were approximately 24 and 4.4 mM glucose, respectively, with the same lactate concentrations (approximately 12 mM). Global gene expression differences were characterized where the specific analyses focused on metabolic control genes in glycolysis, the tricarboxylic acid (TCA) cycle and the genes involved in product quality, specifically, glycosylation-related reaction genes.

Materials and Methods

CHO cell line and culture conditions

The rCHO DP-12 cells, clone #1934 (ATCC[®] CRL-12445TM) were purchased from the American Type Culture Collection (ATCC). These cells express a human monoclonal antibody (IgG₁) against Interleukin-8 (IL-8). The CHO cells were grown in a media mixture of 80% DMEM without glucose (Life technologies, NY) supplemented with 10% Cell-Ess[®] serum replacement (Essential Pharmaceuticals LLC, NJ) and 20% CD CHO medium (Life technologies, NY). The 80:20 medium mixture was supplemented with 5 mM L-glutamine (Life Technologies, NY) contains 10 mM glucose and represents the low glucose medium. A glucose stock solution was used to increase the initial glucose concentration to 30 mM glucose representing the high glucose medium. The glucose stock solution was 1.0 M anhydrous dextrose (EMD Millipore, MA) prepared in Milli-Q[®]-purified water (EMD Millipore, MA) which was filter-sterilized through a 0.2 μ m syringe filters (FisherbrandTM Thermo Fischer Scientific, Waltham, MA). It was desired to have a chemically defined, serum-free, animal-component-free, protein-free medium for these studies. The rCHO DP-12 would not adapt to 100% DMEM supplemented with 10% Cell-Ess[®] serum replacement and custom CD CHO medium without glucose was too expensive for this small-scale study. So, in order to achieve 10 mM and 30 mM glucose concentrations, the mixed 80% DMEM medium supplemented with 10% Cell-Ess[®] serum replacement and 20% CD CHO medium was used. The rCHO DP-12 cultures were preconditioned to the high and low glucose media for at least three passages prior to the experiment.

Cells were seeded at approximately 2×10^5 cells/mL in 75 cm² flasks (Corning Inc, NY) and incubated at 37°C in a humidified 5% CO₂ incubator and shaken at 120 rpm. Anti-clumping agent (Life technologies, NY) was added (1:250) to the medium to prevent cell clumping. Cell concentration and viabilities, glucose and lactate concentrations were obtained every 12 hours. Cell concentrations and viabilities were measured using Bio-Rad TC20™ Automated Cell Counter (Bio-Rad Laboratories, CA). The monoclonal antibody concentration was approximated on 4-20% gradient SDS-PAGE gels (BioRad) stained with Coomassie Blue. Glucose and lactate concentrations were obtained using a Yellow Springs Instruments (YSI) 2900 Biochemistry Analyzer (Yellow Springs, OH). All conditions were conducted in duplicates.

Statistical analysis

Statistical analysis for culture parameters were performed using the software JMP® pro 10 (SAS Institute, Cary, NC) or Microsoft® Excel with LINEST and the built in F-statistical tools. T-tests were used to determine if parameters were different between the conditions ($P \leq 0.05$). The standard deviations are reported for growth rates, cell concentrations, glucose and lactate concentrations, and uptake and production rates.

RNA isolation

For transcriptome samples, total RNA was extracted 36 h after inoculating the cells, while the cells were in mid-exponential growth phase. To preserve the integrity of the RNA, RNAlater® cell reagent (Qiagen®, Valencia, CA) was added to the harvested sample immediately upon withdrawal from the flask. Total RNA extraction was performed with the RNeasy® Plus Mini kit (Qiagen®) according to the manufacturer's recommended protocol and quality/integrity was validated on a NanoDrop™ 8000 UV-Vis spectrophotometer (Thermo Scientific, MA) and Agilent 2100 Bioanalyzer System (Agilent, Santa Clara, CA) with RNA 6000 Nano kit (Agilent), respectively. All samples had RNA integrity numbers (RIN) values of greater than nine (9.0).

Library synthesis, sequencing and reference mapping

Each RNA sample was normalized to a standard input concentration (1 µg of total RNA) and an Illumina compatible sequencing library was prepared with the TruSeq RNA sample preparation v2.0 kit (Illumina, San Diego, CA) following the manufacturer's recommended procedures. The resulting cDNA libraries were size validated on an Agilent 2100 Bioanalyzer System to ensure proper fragment distribution (~260 bp) and effective removal of adapter dimers. Sequence libraries were multiplexed and sequence data collected on 2 lanes of a HiSeq® 2500 (Illumina, San Diego, CA) at the Genomics division of David H. Murdock Research Institute (Kannapolis, NC) using a 2×101 bp, paired end read type. Raw sequence reads were demultiplexed on the instrument and reads greater than or equal to 36 bp were cleaned of adapter and low quality bases (less than Phred 15) with the Trimmomatic software package.⁴⁸ The Sequence Read Archive (SRA) accession numbers are SAMN06481667, SAMN06481668, SAMN06481669, SAMN06481670 under BioProject number PRJNA378247.

The trimmed reads obtained were aligned to the Chinese hamster reference genome sequence (GenBank AFTD00000000) with the Subread Aligner (v1.4.5).⁴⁹ Subread performed local alignment of the reads and reported the locations that have the largest overlap among the reads. The FeatureCounts software package was used to assign the reads to the target genes and summarize the data into a read counts table.⁴⁹ This table was subsequently used as input for differential gene expression analysis.

Differential gene expression analysis and gene ontology analysis

Differential gene expression profiling, data filtering, and statistical analysis of the RNA-Seq data was conducted using the edgeR software package.⁵⁰ The edgeR software utilizes a negative binomial distribution model to capture the quadratic mean-variance relationship common to RNA-Seq data.^{51–53} Additionally, edgeR employs an empirical Bayes method to stabilize over-dispersion across genes by sharing information across all observations including replicates^{54,55}; which is crucial for the experimental design in this study. The steps used for RNA-Seq statistical analysis were performed as follows: (1) filter low expressed genes, (2) normalize the library sizes, (3) visualize the multidimensional scaling plots, (4) estimate dispersions across genes, (5) conduct exact tests to identify differentially expressed genes, and (6) perform the overrepresentation analysis. Low expressed genes were removed (filtered) if a gene had less than a specified count-per-million (CPM) value threshold value in at least two samples.⁵⁶ The selected CPM threshold criteria will be discussed in the Results and Discussion section. The R software package segmented (version 0.5-1.1) was used to identify the breakpoint in the genes sorted by CPM.⁵⁷ Each sample library size was normalized using the trimmed mean of M-values (TMM) method in edgeR.⁵⁸ A multidimensional scaling (MDS) algorithm in edgeR was used to visualize the similarities and dissimilarities between the samples.⁵⁸ Tagwise dispersions were estimated for the design matrix using the quantile-adjusted conditional maximum likelihood (qCML) method.⁵⁵ Genes with differential expression due to the glucose concentration were identified using an exact test. This exact test for negative binomial models is highly similar to Fisher's exact test.⁵³ P values were adjusted to control the false discovery rate by a Benjamini-Hochberg procedure and genes with $FDR \leq 0.01$ were deemed as statistically significant, unless otherwise specified.⁵⁹ The overrepresentation analysis was conducted using gene ontology enrichment analysis performed using the goseq software program (version 1.18.0) – a R based Bioconductor package.⁶⁰ The goseq software accounts for selection bias for long and overexpressed gene sequences by calculating the Probability Weighting Function (PWF).⁶¹ For GO analysis, goseq uses two vectors—one containing all the measured genes and the other containing the differentially expressed genes. Additional information containing genes length and the GO categories associated with the genes were supplied. GO terms with $P \leq 0.05$ from the overrepresentation test were considered to be statistically significant. Additionally, manual screening of gene and gene families, that had common stem in its gene symbol, with greater than 1.5-fold expression change was performed to identify genes sensitive to the glucose concentration difference.

Results and Discussion

Glucose is the primary carbon and energy source in most mammalian cell culture medium; however, the rapid consumption of

glucose by proliferating mammalian cells is also characterized by subsequent production of waste metabolites. Lactate is the principal waste metabolite that can cause growth inhibition and decrease recombinant protein productivity.^{24,26,27} It has been well-documented that the metabolic states of CHO cells can be changed to a lower lactate producing state by controlling the glucose concentrations to a lower concentration.^{25,34–36,45,62} RNA-Seq was used to characterize the impact of the glucose concentration on exponentially growing CHO cells, where the glucose concentration was the only major difference at the time of the comparison. To achieve these conditions, rCHO DP-12 cells were seeded into a chemically defined medium with initial low (10 mM) and high (30 mM) glucose concentrations at a low cell concentrations to control the glucose depletion and minimize the lactate build-up. Figure 1 shows the glucose and lactate profiles for these cultures, as well the cell growth profiles. The samples for the transcriptome analysis were taken at 36-h. Up to this point, the specific growth rates for cultures were not significantly different ($P > 0.05$), as shown in Table 1, and represent normal growth rates for CHO cells. Standard errors are reported, unless stated otherwise. Additionally, the cell concentrations were not different for the high and low glucose cultures at harvest, respectively ($P > 0.05$). And, the cell viabilities ($98 \pm 1\%$) were not significantly different at 36-h. The cell viabilities remained above 95% for the entire culture time, while glucose was above 1 mM. Also, shown in Table 1 are the glucose and lactate concentrations at 36-h, as well as the glucose uptake and lactate production rates, and the cell specific glucose and cell specific lactate production rates at 36-h. Only the glucose concentration at 36-h was significantly different between the cultures. Furthermore, both cultures had glucose concentrations well above the limiting concentration of ~ 1 mM at this time.²³ The lactate concentration observed at 36-h is not considered sufficiently high to be inhibitory³⁶ (See Table 1). The monoclonal antibody was detected by SDS-PAGE gels at the end of the cultures. Wipperfman et al. observed less than 45 mg/L monoclonal antibody for this cell line at 3 days (their first sample time point for the monoclonal antibody) with significantly higher cell concentrations.⁶³ Therefore, it was not expected there would be detectable differences in protein productivity or protein quality in 36-h samples. After 36-h, the growth rates diverged due to the depletion of glucose from the low glucose cultures. RNA samples were extracted from cultures in mid-exponential phase at 36-h, when the only major difference between the cultures was the glucose concentration.

Filtering low expression genes

RNA-Seq data obtained from the sequencer was assembled and aligned with the Chinese hamster reference genome.⁶⁴ Overall, 27,629 genes mapped to the Chinese hamster genome. The initial evaluation of the RNA-Seq data for differentially expressed genes used at CPM value of one (1) as per literature standard practice.^{50,56,65,66} A CPM threshold of one indicates that a gene specific sequence with more than one count per million sequences read in at least two or more libraries would pass the threshold. Using the standard practice threshold CPM value of one on the CHO cell glucose data yielded a gene list of 13,603 genes. When the exact test was applied, 690 genes were identified as having differential gene expression between the high and low glucose cultures. Further examination of this 690 gene list revealed that the genes with highest fold differences corresponded to the genes with the lowest counts, that is, lowest CPM. Specifically, for the genes with greater than 2-fold

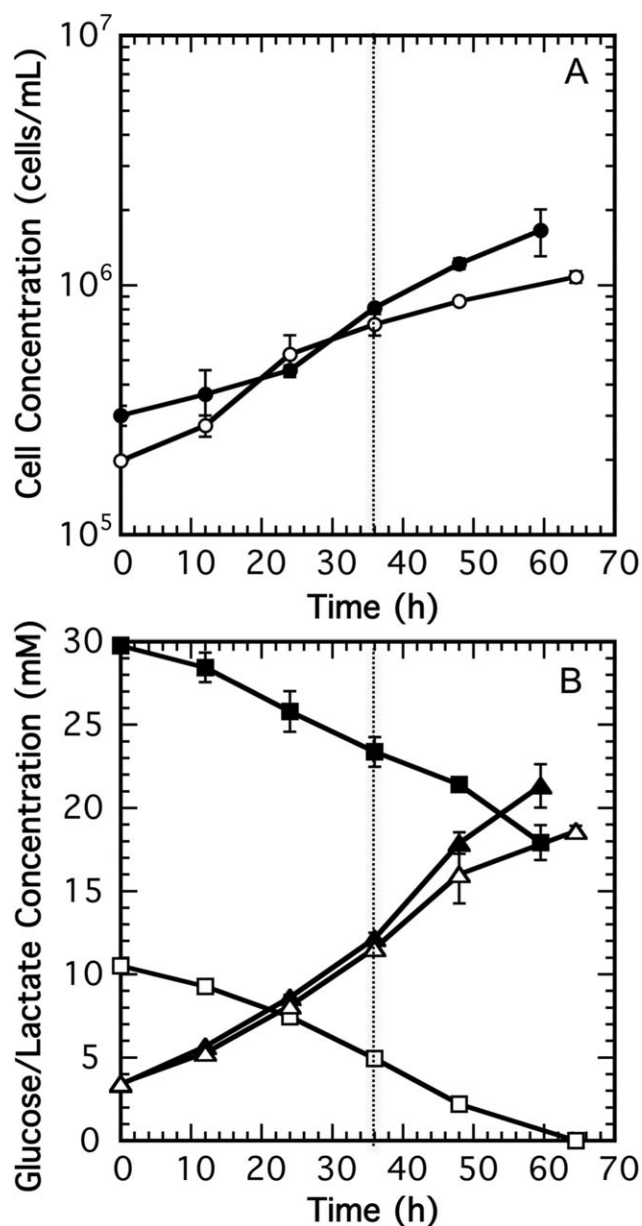


Figure 1. Growth, glucose, and lactate profiles for the high and low glucose cultures.

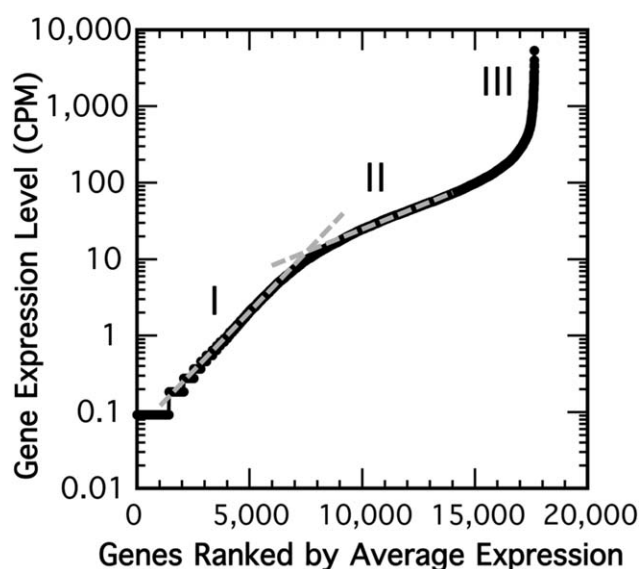
The initial glucose concentrations were 30 mM and 10 mM glucose for the high and low glucose cultures, respectively. (A) Cell concentrations [high glucose (●) and low glucose (○)]; (B) Glucose concentrations [high glucose (■) and low glucose (□)]; and Lactate concentrations [high glucose (▲) and low glucose (△)]. The glucose concentrations were 23.6 ± 0.4 mM and 4.4 ± 0.5 mM glucose for the high and low glucose cultures, respectively, at 36-h. The dashed vertical line indicates the time at which the samples were taken (36-h) for transcriptome analysis. Error bars represent standard deviations.

expression differences (141 genes of the 690 genes), 77 of the 141 genes with 2-fold differences were genes with the absolute lowest CPM values within the 690 gene list. Additionally, the genes identified using the CPM one (1) value contain gene families with inconsistent or contradictory behavior. For example, the gene, *Cacna2d3*, which encodes for the calcium channel alpha 2-delta 3 subunit was found to be 13-fold higher in the high glucose cultures, while the pre-protein gene (*Cacna2d1*) was only 1.3 higher. And, the genes that encode for the other subunits (*Cacna1a*, *Cacnb*,

Table 1. Culture Parameters for the Low and High Glucose Cultures

Culture Parameter	Culture Conditions	
	High Glucose	Low Glucose
Initial glucose (mM)	#29.8 ± 0.4	#10.5 ± 0.5
Glucose at 36-h (mM)	#23.6 ± 0.4	#4.4 ± 0.5
Lactate at 36-h (mM)	12.2 ± 0.3	11.6 ± 0.4
Cell concentration at 36-h (× 10 ⁵ cells/mL)	8.01 ± 0.09	6.97 ± 0.68
Overall growth rate (h ⁻¹)	0.030 ± 0.002	0.030 ± 0.002
Growth rate 0- to 36-h (h ⁻¹)	0.027 ± 0.005	0.035 ± 0.002
Growth rate 12- to 48-h (h ⁻¹)	0.035 ± 0.004	0.032 ± 0.004
Glucose uptake rate at 36-h (mmol/L·h)	-0.181 ± 0.017	-0.156 ± 0.017
Lactate production rate at 36-h (mmol/L·h)	0.244 ± 0.017	0.228 ± 0.021
Cell specific glucose uptake rate at 36-h (nmol/10 ⁶ cell·h)*	-224 ± 77	-224 ± 94
Cell specific lactate production rate at 36-h (nmol/10 ⁶ cell·h)*	301 ± 81	328 ± 104

Only the initial and at 36-h glucose concentrations were different. RNA was harvested for RNA-Seq analysis at 36-h. Measured values are shown with standard errors. #indicates values that are statistically different between the culture conditions ($P \leq 0.05$). *indicates the standard errors were determined using the method described in Hayya et al. for ratios to account for the coefficient of variation. The coefficient of variation for the cell concentrations, glucose uptake and lactate production rates, which were used to obtain the cell specific rates, were approximately 10%, which indicates reasonable control in biological experiments

**Figure 2. Ranked order plot of gene expression levels for the high and low glucose cultures.**

The average gene expression levels of the four samples were used to ranked order the gene expression levels for the high and low glucose cultures. Three distinctly different regions of expression are visible and labeled as regions I, II, and III. The dashed gray lines represent the slopes of regions I and II, where the breakpoint was determined statistically using the software package, segmented.

Cacngl) had been excluded by the CPM threshold of one (1). Since the CPM threshold of one (1) is a convention, within the relatively new field of RNA-Seq analysis, the CPM distribution of the genes for the CHO cell glucose data was further examined to determine a better method for selecting a CPM threshold for the CHO cell data that would have fewer spurious genes called out.

A ranked list of the all the genes was used to decide if there were different types of expression patterns within the gene expression data set. There were 9,994 genes that had no reads in any of the four samples for the two culture conditions, thus the ranked ordered plot contains only 17,635 genes. The plot of the expression levels shows three distinctive regions termed region I, II, and III (Figure 2). The housekeeping genes, which are well-known to have very high expression within an organism,

were located within region III in Figure 2. Region III had a very high slope relative to region II. Housekeeping genes such as β -actin (*Actb*), Glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) and β -2-microglobulin (*B2m*) were located in this region. Region II contained the majority of the genes, while region I contained the genes with the lowest expression. The genes in the middle region corresponded to genes with intermediate CPM values and the slope in this region was the lowest of the three regions. Interestingly, the slope in region I was higher than region II. In reaction kinetics, changes in slope often represent changes in mechanisms.⁶⁷ The observed slope change in Figure 2 between region I and II might not represent a classical reaction mechanism change, but could be due to indexing inaccuracies^{68–70} or some other interpretation variation due the relative newness of the reference sequence status of the CHO genome. While the indexing procedure has several advantages, it also has the potential to introduce false assignment of sequences to the sample libraries.⁶⁸ Moreover, CHO cells have been well-known to have unstable genomes, that is, prone to chromosomal rearrangement, which could create indexing errors.^{71–74} Regardless of the mechanism or cause, it appears in this data, sequences and/or genes below this breakpoint may not have been expressed at biologically meaningful levels and should be removed prior to the statistical analysis.

A statistical software tool was used to identify the breakpoint in the gene expression data shown in Figure 2. This software tool, called segmented, employed a generalized linear model to determine the breakpoint in the regression data. The segmented software tool is a R-based software package.⁵⁷ The data used for the analysis were genes with CPM values from 1.0 to 100, that is, genes in regions I and II. The average CPM value for a gene from all four condition was used. As shown in Figure 2, the breakpoint was determined to be slightly above 10. Since RNA-Seq data is discrete, the final CPM threshold value was rounded down to the nearest integer, which was 10. Using the CPM threshold criteria of 10, 10,090 genes were identified as passing the filter criteria, compared to 13,603 genes that passed the CPM 1 threshold criteria. The differential gene expression analysis by the glucose concentration used these 10,090 genes for further analysis. The multidimensional scaling (MDS) plot showed there was only one dimension that was significant, and even that dimension, the first, had very low log-fold-changes between the conditions.^{75,76} Yet, there was clear separation between the high and low glucose cultures. See

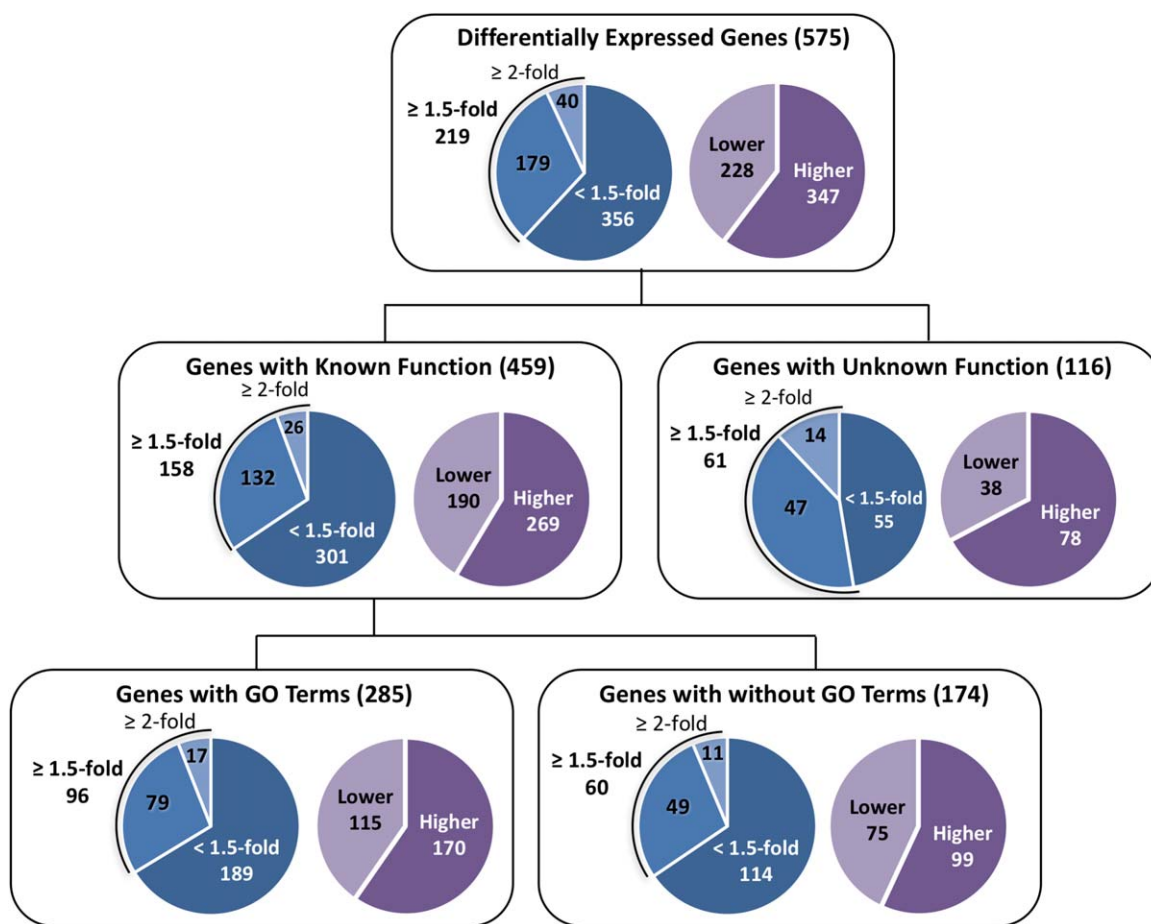


Figure 3. Distribution of the differentially expressed genes between the high and low glucose cultures.

There were 575 genes determined to have significantly different expression between the high and low glucose cultures ($FDR \leq 0.01$). The distribution of the genes with known and unknown function, and genes with and without GO Terms are shown. Each group of genes has been subdivided into genes with higher (darker purple) or lower (lighter purple) expression relative to the high glucose cultures, where higher indicates the gene expression was higher in the high glucose culture. Fold differences between the cultures are also shown. The fold differences were organized into groups as less than 1.5-fold (darkest blue), at least 1.5-fold, but less than 2-fold (middle blue), and at least a 2-fold (lightest blue).

the supplemental materials for the MDS plot (Supporting Information Figure 1).

Differential gene expression analysis

Using the CPM value of 10, an exact test was performed on the 10,090 genes. The exact test analysis yielded 575 genes as identified to be significantly different between the high and low glucose cultures, for example, most likely sensitive to the glucose concentrations ($FDR \leq 0.01$). Figure 3 shows an organizational schematic for these 575 genes based on known and unknown status, as well as expression differences between the culture conditions. For example, for the differential genes with known function (459 genes total), 158 genes had at least a 1.5-fold (≥ 1.5 -fold) difference in expression between the two conditions, while only 26 of these genes at least a 2-fold (≥ 2 -fold) difference. The expression differences were relatively evenly split between higher and lower expression in the high glucose cultures at 269 and 190, respectively. Since the CHO genome (which includes the Chinese hamster and the Chinese hamster ovary cell line derived from *Cricetulus griseus*) only recently received reference Genome status from National Center for Biotechnology Information (NCBI),⁷⁷ the genome is not as

highly annotated as the human and mouse genomes. Consequently, only 285 genes had gene ontology (GO) term annotation.

GO enrichment analysis was performed using the goseq software package on all 285 annotated genes, regardless of fold difference. The goseq analysis identified several biological processes that were over-represented: Protein transport, oxidative-reduction processes, and immune responses. The genes associated with these biological process are listed in Table 2. Additionally, other GO terms associated with a particular gene are included for reference.

From the GO analysis, nine protein transport genes were identified as enriched. Eight of these genes had higher expression in the high glucose cultures (Table 2); however, none of these genes had fold differences greater than 2-fold, and only three genes had fold differences greater than 1.5-fold. Seven of these genes belonged to the RAS oncogene family (*Rab*) or closely related (*Rho*). RAS oncogenes encode proteins that regulate intracellular vesicle formation and transport, mainly between ER and trans-Golgi network.^{78,79} This family and homologs of genes is large, with 85 members in the Chinese hamster genome, where 53 genes passed the CPM 10 threshold criteria. However, only seven of these genes had significantly different expression between

Table 2. Differentially Expressed Genes that were Determined to be Enriched Using the goseq Software Tool ($FDR \leq 0.01$)

Gene Symbol	Gene Description	Fold Difference	Other Associated GO Terms
Protein Transport (GO:0015031)			
Chmp7	charged multivesicular body protein 7, transcript variant X2	1.2	GO:0015031
Gem	GTP binding protein overexpressed in skeletal muscle, transcript variant X2	1.7	GO:0005622, GO:0016020, GO:0003924, GO:0005525, GO:0007165, GO:0007264, GO:0015031
Rab11fp5	RAB11 family interacting protein 5 (class D), transcript variant X2	1.5	GO:0016020, GO:0043231, GO:0008324, GO:0006812, GO:0015031, GO:0055085
Rab14	RAB14, member RAS oncogene family, transcript variant X2	-1.3	GO:0005622, GO:0003924, GO:0005515, GO:0005525, GO:0006886, GO:0006913, GO:0007165, GO:0007264, GO:0015031
Rab34	RAB34, member RAS oncogene family, transcript variant X3	1.3	GO:0005622, GO:0003924, GO:0005515, GO:0005525, GO:0006886, GO:0006913, GO:0007165, GO:0007264, GO:0015031
Rab43	RAB43, member RAS oncogene family, transcript variant X2	1.3	GO:0005622, GO:0003924, GO:0005515, GO:0005525, GO:0006886, GO:0006913, GO:0007165, GO:0007264, GO:0015031
Rab7b	RAB7B, member RAS oncogene family, transcript variant X2	1.4	GO:0005622, GO:0003924, GO:0005515, GO:0005525, GO:0006886, GO:0006913, GO:0007165, GO:0007264, GO:0015031
Rhob	ras homolog family member B, transcript variant X2	1.4	GO:0005622, GO:0005525, GO:0007264, GO:0015031
Rhoq	ras homolog family member Q, transcript variant X2	1.5	GO:0005622, GO:0005525, GO:0007264, GO:0015031
Oxidation-Reduction Process (GO:005114)			
Acad9	Acyl-CoA dehydrogenase family, member 9	-1.3	GO:0003995, GO:0016627, GO:0008152, GO:0055114
Adh7	Alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide, transcript variant X2	2.5	GO:0003824, GO:0005488, GO:0008270, GO:0016491, GO:0008152, GO:0055114
Akrlc2	Aldo-keto reductase family 1, member E2, transcript variant X1	1.5	GO:0016491, GO:0055114
Cbr3	Carbonyl Reductase 3	1.4	GO:0003824, GO:0005488, GO:0016491, GO:0008152, GO:0055114, GO:0004499, GO:0050660, GO:0050661, GO:0055114
Fmo4	Flavin containing monooxygenase 4, transcript variant X1	1.5	GO:0004499, GO:0050660, GO:0050661, GO:0055114
Hmox1	Heme oxygenase (decycling) 1, transcript variant X2	1.3	GO:0004392, GO:0006788, GO:0055114
Idh3b	Isocitrate dehydrogenase 3 (NAD+) beta, transcript variant X1	1.4	GO:000287, GO:0016616, GO:0051287, GO:0055114
Kenab2	Potassium voltage-gated channel, shaker-related subfamily, beta member 2, transcript variant X3	-1.7	GO:0016491, GO:0055114
Leprel1	Leptocan-like 1, transcript variant X2	1.6	GO:0005506, GO:0016491, GO:0016705, GO:0031418, GO:0055114
Ptgr1	Prostaglandin reductase 1, transcript variant X2	1.5	GO:0003824, GO:0005488, GO:0008270, GO:0016491, GO:0008152, GO:0055114
Ptgs2	Prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase), transcript variant X3	1.6	GO:0004601, GO:0020037, GO:0006979, GO:0055114
Xdh	Xanthine dehydrogenase, transcript variant X2	-1.4	GO:0003824, GO:0009055, GO:0016491, GO:0046872, GO:0050660, GO:0051536, GO:0055114
Immune Response (GO:0006955)			
Cxcl3	Chemokine (C-X-C Motif) Ligand 3	-1.3	GO:0005576, GO:0005125, GO:0008009, GO:0006955
Lif	Leukemia inhibitory factor	1.3	GO:0005576, GO:0005125, GO:0006955
Oasl	2'-5'-oligoadenylate synthetase-like	2.6	GO:0003723, GO:0005524, GO:0016740, GO:0006955
Timag1l	Tubulointerstitial nephritis antigen-like 1, transcript variant X2	2.3	GO:0005044, GO:0008234, GO:0030247, GO:0006508, GO:0006955
Tnfrsf9	Tumor necrosis factor (ligand) superfamily, member 9	-1.3	GO:0016020, GO:0005164, GO:0006955

A positive fold difference indicates a gene had higher expression in the culture with the high glucose concentration and conversely a negative fold difference indicates a gene had lower expression in the culture with the high glucose concentration

the conditions. Further examination of the RAS oncogene genes shows that several RAS genes had CPM values between 300 and 900 CPM, where the average RAS family or related genes had CPM values of 117. Yet, five of the seven genes with significant differences only had CPM values between 10 and 53. In contrast, *Rab14* and *Rab34*, which were -1.3 -fold and 1.3 -fold different, had average CPM values of 281 and 163, respectively, counterbalancing each other's change. Hence, the RAS oncogenes with fold differences, only made up a small part of this family, some of the effects were counterbalanced, and most likely would not result in biologically meaningful effects in the cultures.

The GO analysis identified 12 oxidation-reduction process genes as enriched (Table 2). Nine of these genes had higher expression in the high glucose cultures; however, only one gene had a fold difference greater than 2-fold (alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide, transcript variant X2). The oxidative-reduction process genes belonged to several different families involved in oxidation-reduction reactions, such as reductase, oxygenase, and dehydrogenase. Several studies have documented the induction of oxidative stress in response to high glucose exposure; however, in those cases the glucose concentrations were much higher than used in this study.^{80–83} As there was a mixture of reductase and oxidase encoding genes observed to have higher expression in the high glucose culture, all the observed changes, except one was less than 2-fold, and there was a diverse assortment of single gene changes, it is unlikely that these observations resulted in any biologically meaningful effect to the oxidative or reductive responses.

From the GO analysis, five immune response genes were identified as enriched (Table 2). These immune response genes that were significant included two chemokines (*Tnfsf9* and *Cxcl3*) and one cytokines (*Lif*). These genes encoded proteins that induce proliferation and attract neutrophils, and have antiviral activity, respectively. There are numerous chemokine and cytokine encoding genes in the Chinese hamster genome (> 30). Other genes in the tumor necrosis factor receptor superfamily (*Tnfsf*) passed the CPM 10 criteria, but did not have significant differences between the conditions. Only the two *Cxcl3* genes passed the CPM criteria, while the remaining chemokine (C-X-C motif) ligands had no reads or were below even the CPM criteria. *Oas1* and *Tinag11* had greater than 2-fold differences; however, the expression level of the genes was relatively low with average CPM values of 11 and 83, and others related genes were not detected. Other cytokines genes, such as the dedicator of cytokinesis (Dock) genes, were detected, but were not significantly different between the conditions, except for *Dock9*, which was 1.6-fold higher in the high glucose culture with an average CPM of 15. With low observed fold differences, many other chemokines and cytokines without significant difference, it is unclear whether these observations are biologically meaningful. Further evaluation would be required to finalize that assessment.

Since the GO term enrichment analysis could only analyze the genes with GO terms, which represented only approximately half of the differentially expressed genes, a guided screening method was necessary to further examine all the differentially expressed genes. The first analysis identified genes with at least a 2-fold differences between the high and low glucose cultures. Table 3 lists these 48 genes alphabetically. Among the 2-fold genes, 14 genes were classified as

LOC genes, which indicates these genes have uncertain function. Additionally, many of the genes were classified as transcript variants, where transcript variants simply represent alternative mRNA sequences transcribed from the same DNA sequence or gene.⁸⁴ Seven of these genes had average CPM values below 10, which indicated these genes had relatively low expression in CHO cells under these conditions.

The only genes on the 2-fold difference list with a common function were two collagen-associated genes (*Col12a* and *Col6a1*), which had higher expression in the high glucose cultures. *Col12a* and *Col6a1* genes encode for collagen-type VI and XII, respectively. The CPM levels for these genes was relatively low compared to other collagen genes expressed in CHO cells under these conditions. Specifically, there are 44 collagen-associated genes in the Chinese hamster genome, and nine genes passed the CPM 10 threshold. Previous researchers have identified by chemical analysis only collagen-types IV and V in CHO cells at measureable quantities.⁸⁵ The genes that encode for collagen-type IV and V are *Col4a1* and *Col5a2*, respectively. *Col4a1* and *Col5a2* were observed to have CPM values of 417 and 169, respectively. *Col4a1* gene expression was not significantly different between the culture conditions; however, *Col5a2* was observed to be 1.4-fold lower in the high glucose culture ($FDR \leq 0.01$), which was the opposite trend observed for *Col12a* and *Col6a1*. Therefore, it is unlikely that the observed *Col12a* and *Col6a1* fold differences between the high and low glucose cultures resulted in a significant biological difference in collagen composition of the cells.

Another gene that had a high fold difference was the gene that encodes for cytochrome P450 4X1 (LOC100766854). It was observed to have 3.5-fold higher expression in the high glucose cultures. In the Chinese hamster genome, there are 62 annotated cytochrome P450 genes, where five of these genes passed the CPM 10 threshold criteria. Three of the cytochrome P450 genes did not have significant differences in expression between the high and low glucose cultures, including the two cytochrome P450 genes with the highest average CPM values (cytochrome P450 2J3 and cytochrome P450 20A1, transcript variant X2). Additionally, the CPM levels of the cytochrome P450 genes were roughly on average half to one-third of the cytochrome c and b levels. As the exact function of cytochrome P450 4X1 is uncharacterized, it is unclear if it's higher expression in the high glucose culture was of biological significance, as some cytochromes are very specific for a single reaction and other cytochrome have broad and overlapping reaction coverage with other cytochromes.⁸⁶ The other 2-fold genes likewise appear to represent gene expression changes that were not manifested in biological changes, as genes within common families were similarly not observed to have changed.

Glycolysis and TCA cycle genes

Glucose is processed to energy and metabolic precursors via the glycolysis and TCA cycle. There are 83 genes that encode for key enzymes and transport proteins, including isoforms, related to glycolysis and TCA cycle.⁸⁷ Only 55 of these genes were analyzed for differential gene expression, since 20 genes of these genes had no detected reads and eight genes did not meet the CPM threshold criteria. After the exact test ($FDR \leq 0.01$), only two glycolysis genes (*Pfkfb3* and *Pgm1*) and two TCA cycle genes (*Idh3b* and *Suc1g2*) had significantly different expression between the

Table 3. Differentially Expressed Genes that had at Least a 2-Fold Difference Between the High and Low Glucose Cultures (FDR $\leq$ 0.01)

Gene Symbol	Gene Description	Fold Difference	Average CPM
Abcb1b	ATP Binding Cassette Subfamily B Member 1	-2.0	41
Adh7	Alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide, transcript variant X2	2.5	18
Als2cl	ALS2 C-terminal like, transcript variant X7	2.2	9.3
Apobp	Apolipoprotein B receptor	-2.2	36
Cited1	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 1, transcript variant X2	2.4	20
Ckb	Creatine kinase, brain, transcript variant X2	3.8	8.1
Col12a1	Collagen, type XII, alpha 1, transcript variant X3	3.1	19
Col6a1	Collagen, type VI, alpha 1	2.2	84
Cornifin-A	Cornifin alpha	2.2	55
Ctgf	Connective tissue growth factor	5.4	25
Dsp	Desmoplakin, transcript variant X2	3.3	7.4
Dusp1	Dual specificity phosphatase 1	2.1	48
Epha2	EPH receptor A2, transcript variant X2	3.0	11
Fn1	Fibronectin 1, transcript variant X11	2.1	64
Gdpd4	Glycerophosphodiester phosphodiesterase domain containing 4	-6.2	8.2
Gk	Glycerol kinase, transcript variant X1	-2.4	19
Hspb6	Heat shock protein, alpha-crystallin-related, B6, transcript variant X2	2.3	12
Itih5	Inter-alpha-trypsin inhibitor heavy chain family, member 5	2.3	51
Lama2	Laminin, alpha 2, transcript variant X3	2.3	18
LOC100752010	Histone H2B type 1	2.2	18
LOC100753628	Uncharacterized LOC100753628, transcript variant X18	2.0	24
LOC100754646	Small proline-rich protein 2H, transcript variant X2	4.3	29
LOC100766854	Cytochrome P450 4X1	3.5	22
LOC100768367	Retinal dehydrogenase 1	2.9	15
LOC100769145	Liver carboxylesterase 4-like, transcript variant X1	5.1	22
LOC100769432	Liver carboxylesterase 1-like, transcript variant X1	2.3	8.2
LOC100773076	Chromosome unknown open reading frame, human C17orf75, transcript variant X2	2.0	9.1
LOC100773612	Ribosomal protein S6 kinase alpha-2, transcript variant X3	-2.3	17
LOC103159189	Uncharacterized LOC103159189	2.4	15
LOC103159228	Uncharacterized LOC103159228	-2.1	22
LOC103160604	Uncharacterized LOC103160604	-2.0	11
LOC103161223	Uncharacterized LOC103161223, transcript variant X4	-2.2	11
LOC103163538	Uncharacterized LOC103163538, transcript variant X2	3.5	8.4
Mmp12	Matrix metalloproteinase 12 (macrophage elastase), transcript variant X2	3.0	24
Oasl	2'-5'-oligoadenylate synthetase-like	2.6	11
Pcdh7	Protocadherin 7, transcript variant X2	2.6	15
Sema3a	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A, transcript variant X2	3.0	54
Sept3	Septin 3, transcript variant X2	2.2	54
Smad3	SMAD family member 3, transcript variant X2	2.0	41
Tinagl1	Tubulointerstitial nephritis antigen-like 1, transcript variant X2	2.3	83

A positive fold difference indicates a gene had higher expression in the culture with the high glucose concentration and conversely a negative fold difference indicates a gene had lower expression in the culture with the high glucose concentration

high and low glucose cultures. In order to fully explore the effects of extracellular glucose on the transcriptome of the glycolysis and TCA genes, the statistical criterion for the exact test was relaxed to $FDR \leq 0.05$. Using the more relaxed criteria, two additional genes were identified; one in glycolysis (*Tpi1*) and one gene in the TCA cycle (*Pc*). Figures 4 and 5 schematically depict these genes within glycolysis and the TCA cycle, respectively. Also shown in Figures 4 and 5 are the gene expression levels for the gene and subunits, if part of a multimeric enzyme or isoenzymes, with the FDR criterion that each gene met.

Starting with glycolysis, the *Pfkfb3* gene had 1.3-fold higher expression in the high glucose cultures ($FDR \leq 0.01$). *Pfkfb3* encodes for 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3, a bi-functional enzyme that has kinase activity, as well as bisphosphatase activity directed towards the synthesis and degradation of fructose-2,6-bisphosphate. This homodimer enzyme acts as a feedback regulator of glycolysis. Elevated *Pfkfb3* expression has been associated with high glycolytic flux in several tumor cells and rapidly proliferating cells.⁸⁸ Additionally, reduced levels of *Pfkfb3* in a mouse model has been observed to slow the glycolytic

flux.⁸⁹ The gene (*Pgm1*) encoding the enzyme phosphoglucose mutase was observed to have 1.3-fold higher gene expression in the high glucose cultures ($FDR \leq 0.05$). There are several PGM isozymes, which are encoded by different genes and catalyze the transfer of phosphate between the 1 and 6 positions of glucose. The other isoenzymes (*Pgm2* and *Pgm3*) were not observed to have any fold differences between the high and low glucose cultures ($FDR > 0.05$). *Pgm5* was not observed due to no detected reads. The gene (*Tpi1*) encoding the enzyme triosephosphate isomerase 1 was observed to have 1.2-fold higher gene expression in the high glucose cultures ($FDR \leq 0.05$). Triosephosphate isomerase 1 is a homodimer, which reversibly catalyzes the isomerization of glyceraldehyde 3-phosphate to dihydroxyacetone phosphate. It is considered a housekeeping enzyme. Both high and low glucose culture conditions had high expression levels of the gene, with an average CPM value of 319. The pyruvate carboxylase (*Pc*) gene was observed to have a 1.3-fold lower expression in the high glucose cultures ($FDR \leq 0.05$). Pyruvate carboxylase (PC) is a homotetramer that catalyzes the conversion of pyruvate to oxaloacetate in the mitochondria. PC enzyme activity is controlled by the

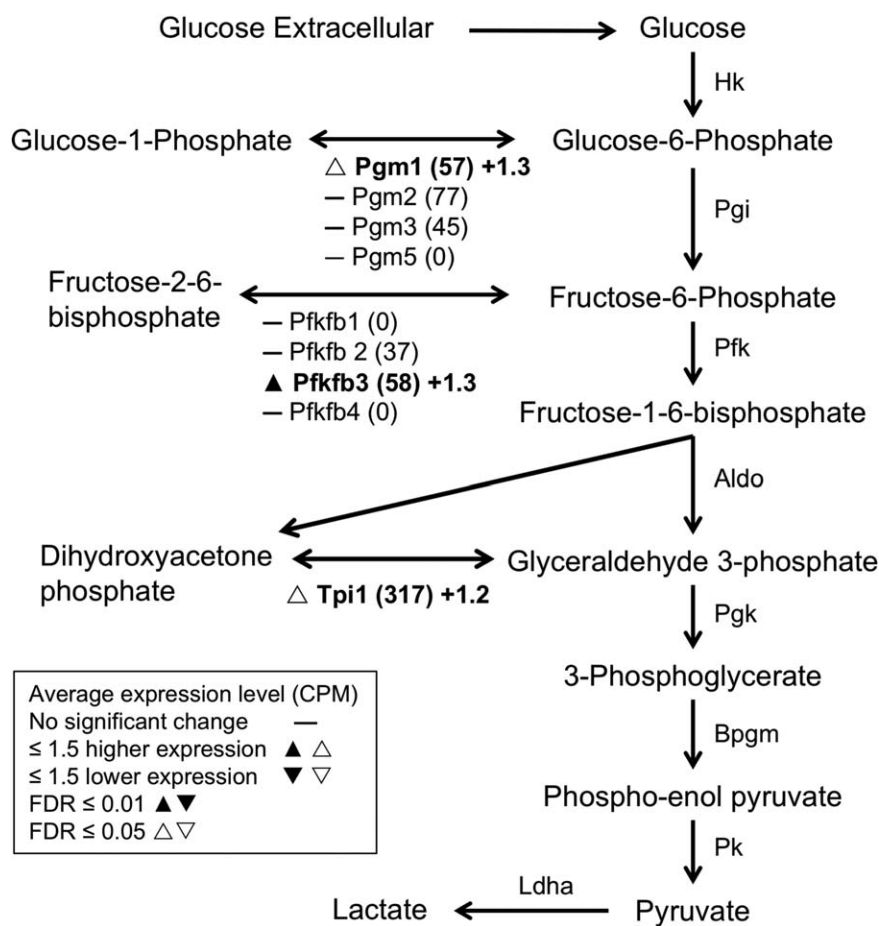


Figure 4. Glycolysis genes with different expression between the high and low glucose cultures.

Phosphoglucosyltransferase 1 (*Pgm1*), 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3 (*Pfkfb3*), and triosephosphate isomerase 1 (*Tpi1*), had significantly different expression between the high and low glucose cultures (FDR ≤ 0.05). Gene expression is shown pictorially by the direction of the triangle: higher expression in the high glucose cultures (▲, △); lower expression in the high glucose cultures (▼, ▽); and no difference (—). Genes that met the more stringent criteria (FDR ≤ 0.01) are shown with fill symbols (▲, ▼) and the less stringent criteria (FDR ≤ 0.05) are shown with open symbols (△, ▽). Standard gene abbreviations are used.

presence of acetyl-CoA and pyruvate in mitochondria; yet, diabetes, which is characterized by elevated glucose, has been shown to increase PC enzyme activity.⁹⁰ In this study, the slightly lower *Pc* gene expression in the high glucose cultures is not consistent with observations in diabetics. The *Idh3b* gene was observed to have 1.4-fold higher gene expression in the high glucose cultures (FDR ≤ 0.01). *Idh3b* encodes the beta subunits of the multimeric enzyme isocitrate dehydrogenases, where this enzyme catalyzes the oxidative decarboxylation of D-threo-isocitrate to 2-oxoglutarate. The other two subunits of the mitochondrial isocitrate dehydrogenase were not observed to have different gene expression levels (FDR > 0.05). The cytosolic isoenzyme subunits (*Idh1* and *Idh2*) were not observed to have any fold differences (FDR > 0.05). The *Suc1g2* gene was observed to have 1.3-fold lower gene expression in the high glucose cultures (FDR ≤ 0.05). *Suc1g2* encodes the alpha subunit of the heterodimer enzyme succinate-CoA ligase, where it converts succinate to succinyl-CoA. *Suc1g2* and *Suc1g1* are both variants of the alpha subunit, while *Suc1a2* encodes the beta subunit. The *Suc1a2* and *Suc1g1* were not observed to have different gene expression levels between the high and low glucose cultures (FDR > 0.05). Overall, the glycolysis and the TCA cycle genes that were observed to have significantly different expression levels all had only slightly fold changes, less than

1.5-fold changes, and many were associated with multi-enzyme complexes where the other subunits did not have significant differences between the culture conditions. Consequently, these gene expression results support the postulate that glucose concentrations between 4.4 and 24 mM do not have a biological meaningful effect on the transcriptome in CHO cells with respect to central metabolism. And, these results are consistent with the observation of Lu et al. that observed that the intracellular metabolite pools associated with glycolysis and the TCA cycle were relatively constants for extracellular glucose concentrations from 1.2 to 9.4 mM.²³

Glycosylation

Protein glycosylation requires precise post-translational modification steps in series to occur where the addition of carbohydrate moieties are constructed and then bound to nascent protein inside the endoplasmic reticulum and Golgi apparatus.^{91,92} Glycosylation of therapeutic proteins is critical in determining the efficacy, immunogenicity and stability of the protein.^{91,93,94} In the biopharmaceutical industry, obtaining consistent glycosylation profiles is necessary for regulatory approval. Fan et al. noticed that the balance of glucose and amino acid metabolism is crucial for glycosylation of recombinant proteins.¹⁸ Therefore,

glycosylation-related gene expression differences were evaluated, in particular, between the high and low glucose cultures.

There are 265 known glycosylation-related genes in CHO cells, previously identified by Xu et al.⁹⁵ Only 156 of these glycosylation genes were analyzed for differential gene expression, since 92 genes had no reads and 17 genes did not meet the CPM 10 threshold criteria. When the glycosylation genes were screened for differential expression, only six glycosylation genes (two copies of one gene) were determined to have significant fold differences ($FDR \leq 0.01$). These significantly different glycosylation genes are listed in Table 4. None of these identified glycosylation genes were

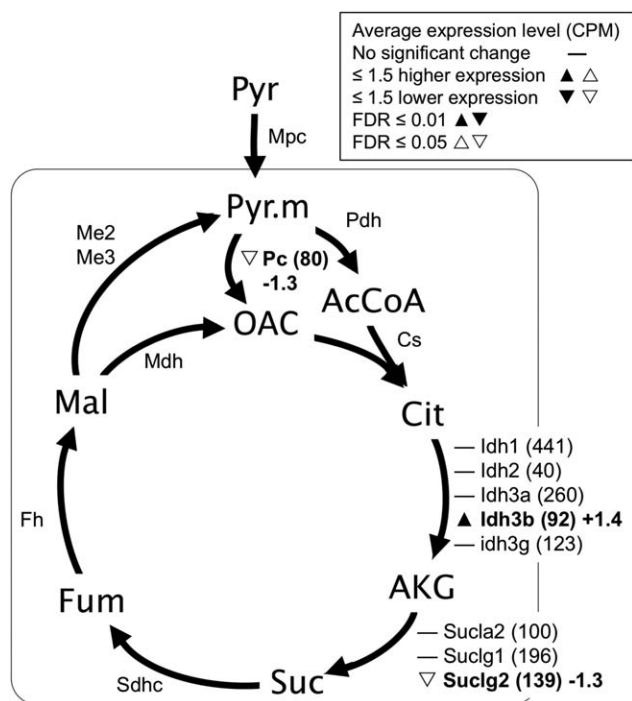


Figure 5. TCA cycles genes with different expression between the high and low glucose cultures.

Pyruvate carboxylase (*Pc*), isocitrate dehydrogenase 3 (NAD(+)) beta (*Idh3b*), and succinate-CoA ligase ADP-forming beta subunit (*Sucla2*) had significantly different expression between the high and low glucose cultures ($FDR \leq 0.05$). Gene expression is shown pictorially by the direction of the triangle: higher expression in the high glucose cultures (▲,▲); lower expression in the high glucose cultures (▼,▼); and no difference (—). Genes that met the more stringent criteria ($FDR \leq 0.01$) are shown with fill symbols (▲,▼) and the less stringent criteria ($FDR \leq 0.05$) are shown with open symbols (▲,▼). Standard gene abbreviations are used.

observed to have greater than or equal to 2-fold differences in expression between the high and low glucose cultures. And, only two genes were observed to have greater than 1.5-fold differences in gene expression—UDP-glucose glycoprotein glucosyltransferase 2 (*Uggt2*) and mannosidase alpha (*Man1c1*). *Uggt2* encodes UDP-glucose glycoprotein glucosyltransferase 2 and was observed to be 1.5-fold lower in the high glucose culture ($FDR \leq 0.01$). UGGT2 recognizes proteins with minor folding defects and re-glycosylates the protein in the endoplasmic reticulum.⁹⁶ The increased expression of *Uggt2* in low glucose cultures could be beneficial, if biologically manifested.

Man1c1 is a mannosidase that trims the alpha-1,2-linked mannose residues from N-linked glycosylation to convert Man(9)GlcNAc(2) toward Man(5)GlcNAc(2) in the Golgi.⁹⁷ *Man1c1* was observed to be 1.7-fold higher in the high glucose cultures ($FDR \leq 0.01$). The lower expression of *Man1c1* low glucose cultures could potentially decrease the trimming reaction rate and slow the processing of protein glycosylation; however, there are many mannosidases that have redundant and higher specificity for trimming alpha-1,2-mannosidases, specifically, *Man1a1* and *Man1a2*. *Man1a1* and *Man1a2* were both observed to not be significantly different between the high and low glucose cultures. Additionally, the average CPM values for *Man1a1* and *Man1a2* was 108, while *Man1c1* only had an average CPM of 9.4. Likewise, the average CPM for *Man2a1*, *Man2a2*, *Man2b1*, *Man2b2*, and *Man2c1* was 109, indicating that *Man1c1* was not a highly expressed mannosidase relative to the other seven mannosidases with similar specificity.⁹⁷ Consequently, the 1.7-fold difference observed between the high and low glucose culture may not have any biological meaningful outcome.

Interestingly, sialyltransferase (*St3gal1*) and a sialidase (*Neu1*) were each observed to have 1.4-fold higher expression in the high glucose cultures, as these genes encode essential opposite functions. Specifically, sialyltransferase enzymes are responsible for adding terminal sialic acid to a glycoprotein in the Golgi, and sialidase enzymes remove the sialic acid from the glycoprotein usually extracellularly.^{91,98} There are six different sialyltransferase genes (*St3gal1-6*). All of the six sialyltransferases were expressed in the CHO cells under the high and low glucose conditions, but only two *St3gal1* transcript variants were differentially expressed (both shown separately in Table 4). The Supporting Information lists all *St3gal1-6* transcript variants with CPM levels. There are four sialidase genes (*Neu1-4*) in the Chinese hamster genome. *Neu4* did not have detectable reads. *Neu2* and *Neu3* did not pass the CPM 10 threshold criteria. *Neu1* was observed to have 1.4-fold higher expression in the high

Table 4. Protein Glycosylation Genes that were Determined to be Significantly Different Between the Low and High Glucose Cultures ($FDR \leq 0.01$)

Gene Symbol	Gene Family	Gene Description	Fold Difference	Average CPM
Fuca1	Fucosidases	fucosidase, alpha-L-1, tissue	-1.4	428
Uggt2	Glucosyltransferases	UDP-glucose glycoprotein glucosyltransferase 2	-1.5	17
Ext1	Glycosyltransferase	exostosin glucosyltransferase 1	1.4	21
Man1c1	Mannosidase	mannosidase alpha class 1C member 1	1.7	9.4
Neu1	Sialidases	neuraminidase 1	1.4	33
St3gal1	Sialyltransferases	ST3 beta-galactoside alpha-2,3-sialyltransferase 1, transcript variant X8	1.4	62
St3gal1	Sialyltransferases	ST3 beta-galactoside alpha-2,3-sialyltransferase 1, transcript variant X8	1.3	13

A positive fold difference indicates a gene had high expression in the culture with the higher glucose concentration and conversely a negative fold difference indicates a gene had lower expression in the culture with the high glucose concentration

glucose cultures. Neu1 is a lysosomal sialidase and the activation of the Neu1 enzyme is strictly dependent on its association with carboxypeptidase protective protein/cathepsin A (PPCA) and β -galactosidase complex.⁹⁹ Neu1 mainly targets the α -2-3 and α -2-6 linkages in oligosaccharides and glycopeptides. The higher expression of *Neu1* in high glucose cultures could be due to its role in glucose feedback mechanisms.¹⁰⁰ In related studies, Ha et al. observed increased Neu3 gene expression relative sialyltransferase gene under ammonium stress.¹⁰¹ They also noted decreased sialylation of the recombinant protein.¹⁰¹ Ha et al. were able to reduce the sialidase gene expression levels with siRNA, and recover much of the sialylation.¹⁰¹ Note, that even though ammonium was not measured in this study for the high and low glucose cultures, it is unlikely that (1) the ammonium concentrations were different between the two conditions and (2) that the ammonium concentrations had reached inhibitory levels,¹⁰² due to the low cell concentrations, short culture duration, and the initial glutamine concentration. Likewise, if any of the observed gene expression changes noted in the present study were to result in a negative biologically meaningful response, siRNA techniques, or permanent silencing could be used to mitigate the effects. However, so far, the sialidase and sialyltransferase gene expression changes appear to be fairly balanced between the high and low glucose cultures, but the sialylation of the glycoprotein should be monitored in further studies with controlled low glucose.

While glycosylation is very important in recombinant protein production, none of the genes observed to have significantly different expression are likely to have an impact on glycosylation, either due to redundancy or counteracting effects. Taken together, these results suggest that between ~24 and 4.4 mM glucose, the CHO cell transcriptome does not undergo significant changes, such that major metabolic pathway enzyme levels were altered. These results; however, do not comprehensively indicate what was occurring metabolically at the substrate-level. Further analysis of the product and metabolism would be required to make these determinations.

Conclusions

For most mammalian cell lines, glucose serves as the primary carbon source required for growth and maintenance. Glucose metabolism is central to several key intermediates for the synthesis of cellular components, such as lipids, sugar chains and nucleic acids. In this study, the impact of two initial glucose concentrations on the CHO cell transcriptome was investigated using next-generation transcriptomic technique, RNA-Seq. For the 27,629 genes annotated in the Chinese hamster genome, 10,090 genes were expressed (CPM ≥ 10). An exact test performed on these genes revealed only 575 genes to be differentially expressed between the high and low glucose cultures. GO enrichment analysis performed on all 285 annotated genes, including genes with very low fold differences, identified several biological processes that were over-represented: Protein transport, oxidative-reduction processes, and immune responses. None of the genes in these GO categories had fold differences ≥ 2 -fold and only, three genes had fold differences greater than 1.5-fold. Guided screening was used to further explore the differentially expressed genes. Forty genes were found to have at least a 2-fold expression

changes; however, none of these genes involved glycolysis, the TCA cycle, or glycosylation-related genes. This gene expression study indicates that controlled low glucose fed-batch cultures are likely to perform well, without any biologically significant or detrimental effects to the CHO cells at the transcriptome level. Additionally, it is well-known that maintaining low glucose concentrations has potential to improve growth and productivity outcomes by reducing waste accumulation, and in turn improving protein glycosylation outcomes. In particular, the transcriptome data support process-related genes will be minimally, if at all, impacted by controlled low glucose control conditions typically used in industrial bioreactors. These data support efforts to reduce the initial glucose concentrations in batch medium and to initiate fed-batch control sooner than is currently practiced.

Acknowledgment

Financial support was provided by the National Science Foundation (NSF) grant number CBET-1218345.

Notation

Aldo = aldolase, fructose-bisphosphate
 Bpgm = 2,3-bisphosphoglycerate mutase
 CHO = Chinese hamster ovary
 CPM = counts per million
 Cs = citrate synthase
 FDR = false discovery rate
 Fh = fumarate hydratase
 GO = gene ontology
 Hk = hexokinase
 Idh = isocitrate dehydrogenase
 IgG₁ = immunoglobulin G subclass 1
 IL-8 = Interleukin-8
 Ldha = lactate dehydrogenase
 Mdh = malate dehydrogenase
 Me = malic enzyme, NADP(+)-dependent (cytosolic and mitochondrial)
 Mpc = mitochondrial pyruvate carrier
 Pc = pyruvate carboxylase
 Pdh = pyruvate dehydrogenase
 Pfk = phosphofructokinase
 Pfkfb = 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase
 Pgi = glucose-6-phosphate isomerase
 Pglk = phosphoglycerate kinase
 Pgm = phosphoglucomutase
 Pk = pyruvate kinase
 siRNA = small interfering RNA
 Sdhc = succinate dehydrogenase
 Sucl = succinate-CoA ligase
 TCA = tricarboxylic acid cycle
 Tpi = triosephosphate isomerase

LITERATURE CITED

- Butler M, Meneses-Acosta A. Recent advances in technology supporting biopharmaceutical production from mammalian cells. *Appl. Microbiol. Biotechnol.* 2012;96:885–894.
- Datta P, Linhardt RJ, Sharfstein ST. An 'omics approach towards CHO cell engineering. *Biotechnol. Bioeng.* 2013;110:1255–1271.
- Kantardjiev A, Zhou W. *Mammalian Cell Cultures for Biologicals Manufacturing*. Springer-Verlag Berlin Heidelberg; 2013.
- Omasa T, Onitsuka M, Kim WD. Cell engineering and cultivation of Chinese hamster ovary (CHO) cells. *Curr. Pharm. Biotechnol.* 2010;11:233–240.

5. Walsh G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* 2014;32:992–1000.
6. Bandaranayake AD, Almo SC. Recent advances in mammalian protein production. *FEBS Lett.* 2014;588:253–260. doi:10.1016/j.febslet.2013.11.035. Recent advances in mammalian protein production. Bandaranayake AD, Almo SC.
7. Jayapal KR, Wlaschin KF, Hu WS, Yap MGS. Recombinant protein therapeutics from CHO cells - 20 years and counting. *Chem. Eng. Prog.* 2007;103:40–47.
8. Glaser V. The biomanufacturing workforce: strategies, statistics, and successes. *Ind. Biotechnol.* 2015;11:199–200.
9. Kim JY, Kim Y-G, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol.* 2012;93:917–930.
10. Wuest DM, Harcum SW, Lee KH. Genomics in mammalian cell culture bioprocessing. *Biotechnol. Adv.* 2012;30:629–638.
11. Craven S, Whelan J. Process analytical technology and quality-by-design for animal cell culture. In: AlRubeai M, editor. *Animal Cell Culture*, Springer International Publishing, Cham, Vol. 9. 2015:647–688.
12. Mehdizadeh H, Lauri D, Karry KM, Moshghbar M, Procopio-Melino R, Drapeau D. Generic Raman-based calibration models enabling real-time monitoring of cell culture bioreactors. *Biotechnol. Prog.* 2015;31:1004–1013.
13. Berry BN, Dobrowsky TM, Timson RC, Kshirsagar R, Ryll T, Wiltberger K. Quick generation of Raman spectroscopy based in-process glucose control to influence biopharmaceutical protein product quality during mammalian cell culture. *Biotechnol. Prog.* 2016;32:224–234.
14. Liu B, Spearman M, Doering J, Lattova E, Perreault H, Butler M. The availability of glucose to CHO cells affects the intracellular lipid-linked oligosaccharide distribution, site occupancy and the N-glycosylation profile of a monoclonal antibody. *J. Biotechnol.* 2014;170:17–27.
15. Templeton N, Dean J, Reddy P, Young JD. Peak antibody production is associated with increased oxidative metabolism in an industrially relevant fed-batch CHO cell culture. *Biotechnol. Bioeng.* 2013;110:2013–2024.
16. Kim DY, Chaudhry MA, Kennard ML, Jardon MA, Braasch K, Dionne B, Butler M, Piret JM. Fed-batch CHO cell t-PA production and feed glutamine replacement to reduce ammonia production. *Biotechnol. Prog.* 2013;29:165–175.
17. Ozturk SS, Thrift JC, Blackie JD, Naveh D. Real-time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor. *Biotechnol. Bioeng.* 1997;53:372–378.
18. Fan Y, Del Val JJ, Muller C, Sen JW, Rasmussen SK, Kontoravdi C, Weilguny D, Andersen MR. Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnol. Bioeng.* 2015;112:521–535.
19. Lee TY, Lin HH, Chen CL, Hwang SM, Tseng CP. Inhibitory effect of excessive glucose on its biochemical pathway and the growth of Chinese hamster ovary (CHO) cells. *J. Carbohydr. Chem.* 2015;34:1–11.
20. Association AD. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2008;31:S55–S60.
21. Taschwer M, Hackl M, Bort JAH, Leitner C, Kumar N, Puc U, Grass J, Papst M, Kunert R, Altmann F, Borth N. Growth, productivity and protein glycosylation in a CHO EpoFc producer cell line adapted to glutamine-free growth. *J. Biotechnol.* 2012;157:295–303.
22. Aehle M, Kuprijanov A, Schaepe S, Simutis R, Lubbert A. Simplified off-gas analyses in animal cell cultures for process monitoring and control purposes. *Biotechnol. Lett.* 2011;33:2103–2110.
23. Lu SB, Sun XM, Zhang YX. Insight into metabolism of CHO cells at low glucose concentration on the basis of the determination of intracellular metabolites. *Process. Biochem.* 2005;40:1917–1921.
24. Lao MS, Toth D. Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. *Biotechnol. Prog.* 1997;13:688–691.
25. Mulukutla BC, Gramer M, Hu WS. On metabolic shift to lactate consumption in fed-batch culture of mammalian cells. *Metab. Eng.* 2012;14:138–149.
26. Mulukutla BC, Khan S, Lange A, Hu WS. Glucose metabolism in mammalian cell culture: new insights for tweaking vintage pathways. *Trends Biotechnol.* 2010;28:476–484.
27. Neermann J, Wagner R. Comparative analysis of glucose and glutamine metabolism in transformed mammalian cell lines, insect and primary liver cells. *J. Cell. Physiol.* 1996;166:152–169.
28. Ozturk S, Hu W-S. *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*. New York: Taylor and Francis; 2005.
29. Seth G, Hossler P, Yee JC, Hu WS. Engineering cells for cell culture bioprocessing - physiological fundamentals. In: *Cell Culture Engineering*, Springer-Verlag, Berlin Heidelberg, Vol.101; 2006:119–164.
30. Young JD. Metabolic flux rewiring in mammalian cell cultures. *Curr. Opin. Biotechnol.* 2013;24:1108–1115.
31. Hu WS, Dodge TC, Frame KK, Himes VB. Effect of glucose on the cultivation of mammalian cells. *Dev. Biol. Standard.* 1987;66:279–290.
32. Omasa T, Higashiyama KI, Shioya S, Suga K. I. Effects of lactate concentration on hybridoma culture in lactate-controlled fed-batch operation. *Biotechnol. Bioeng.* 1992;39:556–564.
33. Wong DCF, Wong KTK, Goh LT, Heng CK, Yap MGS. Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures. *Biotechnol. Bioeng.* 2005;89:164–177.
34. Gagnon M, Hiller G, Luan YT, Kittredge A, DeFelice J, Drapeau D. High-end pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO fed-batch cultures. *Biotechnol. Bioeng.* 2011;108:1328–1337.
35. Le H, Kabbur S, Pollastrini L, Sun ZR, Mills K, Johnson K, Karypis G, Hu WS. Multivariate analysis of cell culture bioprocess data-Lactate consumption as process indicator. *J. Biotechnol.* 2012;162:210–223.
36. Mulukutla BC, Yongky A, Grimm S, Daoutidis P, Hu WS. Multiplicity of steady states in glycolysis and shift of metabolic state in cultured mammalian cells. *Plos One* 2015;10:e0121561. doi:10.1371/journal.pone.0121561
37. Altamirano C, Paredes C, Cairo JJ, Godia F. Improvement of CHO cell culture medium formulation: Simultaneous substitution of glucose and glutamine. *Biotechnol. Prog.* 2000;16:69–75.
38. Altamirano C, Cairo JJ, Godia F. Decoupling cell growth and product formation in Chinese hamster ovary cells through metabolic control. *Biotechnol. Bioeng.* 2001;76:351–360.
39. Altamirano C, Illanes A, Becerra S, Cairo JJ, Godia F. Considerations on the lactate consumption by CHO cells in the presence of galactose. *J. Biotechnol.* 2006;125:547–556.
40. Goudar CT, Joeris K, Konstantinov KB, Piret JM. Logistic equations effectively model mammalian cell batch and fed-batch kinetics by logically constraining the fit. *Biotechnol. Prog.* 2005;21:1109–1118.
41. Craven S, Whelan J, Glennon B. Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller. *J. Process Control.* 2014;24:344–357.
42. Hayter PM, Curling E, Baines AJ, Jenkins N, Salmon I, Strange PG, Tong JM, Bull AT. Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon- γ . *Biotechnol. Bioeng.* 1992;39:327–335.
43. Zhou W, Rehm J, Hu WS. High viable cell concentration fed-batch cultures of hybridoma cells through on-line nutrient feeding. *Biotechnol. Bioeng.* 1995;46:579–587.
44. Zhou WC, Chen CC, Buckland B, Aunins J. Fed-batch culture of recombinant NS0 myeloma cells with high monoclonal antibody production. *Biotechnol. Bioeng.* 1997;55: 783–792.
45. Mulukutla BC, Yongky A, Daoutidis P, Hu W-S. Bistability in glycolysis pathway as a physiological switch in energy metabolism. *Plos One* 2014;9: e98756. doi:10.1371/journal.pone.0098756.
46. Wong DCF, Wong KTK, Lee YY, Morin PN, Heng CK, Yap MGS. Transcriptional profiling of apoptotic pathways in batch and fed-batch CHO cell cultures. *Biotechnol. Bioeng.* 2006;94:373–382.
47. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genetics* 2009;10:57–63.
48. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 2014;30:2114–2120.

49. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 2013;41:e108.
50. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–140.
51. Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* 2007;23:2881–2887.
52. Chen Y, Lun ATL, Smyth GK. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR; Springer International Publishing, Cham, 2014:51–74.
53. Robinson MD, Smyth GK. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* 2008;9:321–332.
54. Fong Y, Wakefield J, De Rosa S, Frahm N. A robust Bayesian random effects model for nonlinear calibration problems. *Biometrics* 2012;68:1103–1112.
55. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012;40:4288–4297.
56. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, Robinson MD. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* 2013;8:1765–1786.
57. Muggeo VMR. Estimating regression models with unknown break-points. *Stat. Med.* 2003;22:3055–3071.
58. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:R25. doi:10.1186/gb-2010-11-3-r25.
59. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 2001;29:1165–1188.
60. Young MD, Wakefield MJ, Smyth GK, Oshlack A. goseq: Gene Ontology testing for RNA-seq datasets; Bioconductor.org, 2012.
61. Oshlack A, Wakefield MJ. Transcript length bias in RNA-seq data confounds systems biology. *Biol. Direct.* 2009;4:
62. Europa AF, Gambhir A, Fu P-C, Hu W-S. Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells. *Biotechnol. Bioeng.* 2000;67:25–34.
63. Wipperfmann A, Klausung S, Rupp O, Albaum SP, Büntemeyer H, Noll T, Hoffrogge R. Establishment of a CpG island microarray for analyses of genome-wide DNA methylation in Chinese hamster ovary cells. *Appl. Microbiol. Biotechnol.* 2014;98:579–589.
64. Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the *Cricetulus griseus* draft genome. *Nat. Biotechnol.* 2013;31:759–765.
65. Educating the Engineer of 2020: Adapting Engineering Education to the New Century (2005) Report. National Academies Press: Washington, DC. DOI:10.17226/11338.
66. Xu G, Strong MJ, Lacey MR, Baribault C, Flemington EK, Taylor CM. RNA CoMPASS: a dual approach for pathogen and host transcriptome analysis of RNA-seq datasets. *PLoS One* 2014;9:e89445.
67. Fogler HS. Elements of Chemical Reaction Engineering; Upper Saddle, NJ: Prentice Hall, 1999.
68. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 2012;40:e3.
69. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protoc.* 2010;2010:prot5448.
70. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 2010;7:111–118.
71. Brinkrolf K, Rupp O, Laux H, Kollin F, Ernst W, Linke B, Kofler R, Romand S, Hesse F, Budach WE. Chinese hamster genome sequenced from sorted chromosomes. *Nat. Biotechnol.* 2013;31:694–695.
72. Cao Y, Kimura S, Itoi T, Honda K, Ohtake H, Omasa T. Construction of BAC-based physical map and analysis of chromosome rearrangement in Chinese hamster ovary cell lines. *Biotechnol. Bioeng.* 2012;109:1357–1367.
73. Derouazi M, Martinet D, Schmutz NB, Flaction R, Wicht M, Bertschinger M, Hacker DL, Beckmann JS, Wurm FM. Genetic characterization of CHO production host DG44 and derivative recombinant cell lines. *Biochem. Biophys. Res. Commun.* 2006;340:1069–1077.
74. Lopez IH, Jadhav V, Baumann M, Auer N, Zotter A, Borth N. Development of a method to quantify genomic rearrangements in Chinese hamster ovary cells. *New Biotechnol.* 2014;31:S187–S187.
75. Gowtham YK, Saski C, Harcum SW. Utilizing RNA-Seq technique to improve molecular understanding of Chinese hamster ovary (CHO) cell bioprocessing. In: *Cell Culture Engineering Conference XV*, La Quinta, CA; 2016.
76. Gowtham YK, Saski C, Harcum SW. Transcriptome analysis of Chinese hamster ovary (CHO) cell lines under low temperature conditions using RNA-seq. In *251th ACS National Meeting & Exposition*, San Diego, CA; 2016.
77. Eukaryotic RefSeq Genome Annotation Status. http://www.ncbi.nlm.nih.gov/genome/annotation_euk/status/ - recent. Accessed on June 12, 2014.
78. Leto D, Saltiel AR. Regulation of glucose transport by insulin: traffic control of GLUT4. *Nat. Rev. Mol. Cell Biol.* 2012;13:383–396.
79. Stenmark H, Olkkonen VM. The rab gtpase family. *Genome Biol.* 2001;2:1.
80. Chang T, Horal M, Jain S, Wang F, Patel R, Loeken M. Oxidant regulation of gene expression and neural tube development: insights gained from diabetic pregnancy on molecular causes of neural tube defects. *Diabetologia* 2003;46:538–545.
81. King GL, Loeken MR. Hyperglycemia-induced oxidative stress in diabetic complications. *Histochem. Cell Biol.* 2004;122:333–338.
82. Rodriguez H, Akman SA, Holmquist GP, Wilson GL, Driggers WJ, LeDoux SP. Mapping oxidative DNA damage using ligation-mediated polymerase chain reaction technology. *Methods-A Companion Methods Enzymol.* 2000;22:148–156.
83. Giugliano D, Ceriello A, Paolisso G. Oxidative stress and diabetic vascular complications. *Diabetes Care* 1996;19:257–267.
84. Pruitt KD, Tatusova T, Klimke W, Maglott DR. NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic Acids Res.* 2008;40:D130–D135. doi:10.1093/nar/gkr1079.
85. Limeback H, Sodek J, Aubin J. Characterization of the collagens synthesized by Chinese hamster ovary cells: effect of colcemid and dibutyl adenosine cyclic monophosphate. *Biochemistry* 1982;21:4720–4729.
86. Nebert DW, Gonzalez FJ. P450 genes: structure, evolution, and regulation. *Annu. Rev. Biochem.* 1987;56:945–993.
87. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* 2014;42:D459–D471.
88. Atsumi T, Chesney J, Metz C, Leng L, Donnelly S, Makita Z, Mitchell R, Bucala R. High expression of inducible 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (iPFK-2; PFKFB3) in human cancers. *Cancer Res.* 2002;62:5881–5887.
89. Mor I, Cheung EC, Vousden KH. Control of glycolysis through regulation of PFK1: old friends and recent additions. *Cold Spring Harbor Symp. Quant. Biol.* 2011;76:211–216.
90. Wallace JC, Jitrapakdee S, Chapman-Smith A. Pyruvate carboxylase. *Int. J. Biochem. Cell Biol.* 1998;30:1–5.
91. Bill, R. M.; Revers, L.; Wilson, I. B. H., *Protein Glycosylation*. Kluwer Academic Publishers: Boston; 1998.
92. Harcum SW, Ozturk S, Hu W. Protein glycosylation. *Biotechnol. Bioprocess. Ser.* 2005;30:113.
93. Jenkins N, Parekh RB, James DC. Getting the glycosylation right: Implications for the biotechnology industry. *Nat. Biotechnol.* 1996;14:975–981.
94. Walsh G. Biopharmaceutical benchmarks – 2003. *Nat. Biotechnol.* 2004;21:865–870.
95. Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol.* 2011;29:735–741.

96. Dejgaard S, Nicolay J, Taheri M, Thomas DY, Bergeron JJM. The ER glycoprotein quality control system. *Curr. Issues Mol. Biol.* 2004;6:29–42.
97. Ben-Ari Fuchs S, Lieder I, Stelzer G, Mazor Y, Buzhor E, Kaplan S, Bogoch Y, Plaschkes I, Shitrit A, Rappaport N, Kohn A, Edgar R, Shenhav L, Safran M, Lancet D, Guan-Golan Y, Warshawsky D, Shtrichman R. GeneAnalytics: an integrative gene set analysis tool for next generation sequencing, RNAseq and microarray data. *Omics* 2016;20:139–151.
98. Gooch CF, Gramer MJ. Glycosidase activities in Chinese hamster ovary cell lysate and cell culture supernatant. *Biotechnol. Prog.* 1993;9:366–373.
99. Miyagi T, Yamaguchi K. Mammalian sialidases: physiological and pathological roles in cellular functions. *Glycobiology* 2012; 22:880–896.
100. Dridi L, Seyrantepe V, Fougerat A, Pan X, Bonneil É, Thibault P, Moreau A, Mitchell GA, Heveker N, Cairo CW. Positive regulation of insulin signaling by neuraminidase 1. *Diabetes* 2013;62:2338–2346.
101. Ha TK, Kim Y-G, Lee GM. Understanding of altered N-glycosylation-related gene expression in recombinant Chinese hamster ovary cells subjected to elevated ammonium concentration by digital mRNA counting. *Biotechnol. Bioeng.* 2015; 112:1583–1593.
102. Chen PF, Harcum SW. Effects of elevated ammonium on glycosylation gene expression in CHO cells. *Metab. Eng.* 2006;8:123–132.

Manuscript received Oct. 10, 2016, and revision received Jan. 7, 2017.