Journal of Biotechnology 189 (2014) 58–69

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

# Gene expression measurements normalized to cell number reveal large scale differences due to cell size changes, transcriptional amplification and transcriptional repression in CHO cells

Dina Fomina-Yadlin, Zhimei Du, Jeffrey T. McGrew*

*Cell Sciences & Technology, Amgen Inc., Seattle, WA 98119, United States*

ARTICLE INFO

*Article history:*
Received 9 July 2014
Received in revised form 18 August 2014
Accepted 25 August 2014
Available online 4 September 2014

*Keywords:*
Chinese Hamster Ovary
Cellular RNA content
Synthetic RNA spike-in controls
Gene expression
Specific productivity

ABSTRACT

Conventional approaches to differential gene expression comparisons assume equal cellular RNA content among experimental conditions. We demonstrate that this assumption should not be universally applied because total RNA yield from a set number of cells varies among experimental treatments of the same Chinese Hamster Ovary (CHO) cell line and among different CHO cell lines expressing recombinant proteins. Conventional normalization strategies mask these differences in cellular RNA content and, consequently, skew biological interpretation of differential expression results. On the contrary, normalization to synthetic spike-in RNA standards added proportional to cell numbers reveals these differences and allows detection of global transcriptional amplification/repression. We apply this normalization method to assess differential gene expression in cell lines of different sizes, as well as cells treated with a cell cycle inhibitor (CCI), an mTOR inhibitor (mTORI), or subjected to high osmolarity conditions. CCI treatment of CHO cells results in a cellular volume increase and global transcriptional amplification, while mTORI treatment causes global transcriptional repression without affecting cellular volume. Similarly to CCI treatment, high osmolarity increases cell size, total RNA content and antibody expression. Furthermore, we show the importance of spike-in normalization for studies involving multiple CHO cell lines and advocate normalization to spike-in controls prior to correlating gene expression to specific productivity ($q_P$). Overall, our data support the need for cell number specific spike-in controls for all gene expression studies where cellular RNA content differs among experimental conditions.



## 1. Introduction

A fundamental assumption of the entire field of comparative transcriptomics has recently been challenged (Loven et al., 2012). Conventional approaches to both global and individual gene expression measurements assume equal cellular RNA content among experimental conditions being compared. This assumption has recently been questioned by the discovery of "global transcriptional amplification" of all actively-transcribed genes that results from c-Myc overexpression in B cells (Lin et al., 2012; Nie et al., 2012). Induction of c-Myc increased cell size and both the mRNA and the rRNA cellular content resulting from augmentation of the entire cellular gene expression program, but these phenotypic observations were masked by traditional global gene expression analysis (Lin et al., 2012; Nie et al., 2012). In order to allow detection of global transcriptional amplification by transcriptomic analysis, Loven et al. (2012) applied the concept of adding spike-in controls proportional to cell number. The use of synthetic spike-in RNA standards for global gene expression analysis using microarrays was previously recommended to allow data comparison across different platforms and protocols used for data collection and analysis (Yang, 2006). Loven et al. (2012) used synthetic RNA spike-in standards (ERCCs) for normalization to cell number, thus, revealing differences in cellular RNA content and enabling cross-platform detection of transcriptional amplification by c-Myc. In addition, spike-in normalization to cell number was equally successful at detecting "global transcriptional repression" in Rett syndrome models driven by MECP2 loss-of-function (Li et al., 2013). Mutant

*Abbreviations:* CCI, cell cycle inhibitor; CHO, Chinese Hamster Ovary; ERCC, External RNA Controls Consortium; HC, heavy chain; IRES, Internal Ribosome Entry Site; LC, light chain; mTOR, mammalian target of rapamycin; mTORI, mTOR inhibitor; $q_P$, specific productivity; qPCR, quantitative real-time PCR; RNA-Seq, next-generation RNA sequencing; RPKM, Reads Per Kilobase of exon model per Million mapped reads; VCD, viable cell density.

* Corresponding author at: 1201 Amgen Court West, MS AW2/D2042, Seattle, WA 98119, United States. Tel.: +1 206 265 7871.
*E-mail address:* mcgrewjt@amgen.com (J.T. McGrew).

http://dx.doi.org/10.1016/j.jbiotec.2014.08.037
0168-1656/

neurons displayed reduced size and cellular RNA content, but conventional global gene expression analysis revealed no change in the majority of genes (Li et al., 2013). In contrast, addition of spike-in standards proportional to cell numbers revealed global transcriptional and translational repression as a consequence of MECP2 deletion (Li et al., 2013), advocating application of this analysis technique to all comparisons involving differences in cellular size.

The onset of the systems biology era in biotechnology has been defined by a wide variety of 'omics-based approaches to characterize the biological basis of desired phenotypic parameters and manipulate them to enhance heterologous protein expression in Chinese Hamster Ovary (CHO) cell lines (Kildegaard et al., 2013). So far, only conventional techniques that assume invariant total RNA yield per cell have been used to evaluate gene expression of CHO cells. For example, many studies have been performed to elucidate the biological basis of specific productivity ($q_P$) and predict $q_P$ from gene expression profiles (Clarke et al., 2011a,b, 2012; Doolan et al., 2012; Kang et al., 2014; Nissom et al., 2006; Yee et al., 2009), but none took into account potential differences in cellular size or total RNA content. Other studies examining effects of small molecule treatments (e.g. sodium butyrate, Kantardjieff et al., 2010) or bioprocess conditions (e.g. temperature, Kantardjieff et al., 2010 and culture osmolarity, Shen et al., 2010) on gene expression in CHO cells also did not take into account changes in cellular RNA content. For this study, we selected two small molecules, one targeting cell cycle progression and the other targeting mTOR signaling, which were hypothesized to have opposite effects on cell size in our biological system. Cell cycle arrest, achieved by either overexpression of an endogenous cell cycle inhibitor ($p21^{CIP1}$) (Bi et al., 2004) or addition of small molecule inhibitors that cause G1 or G1/S arrest (Du et al., 2014; Fingar et al., 2002), has been previously shown to increase mammalian cell size. mTOR signaling through its downstream effectors, ribosomal protein S6 kinase (S6K1) and eukaryotic initiation factor 4E-binding protein (4EBP1), has also been demonstrated to regulate mammalian cell size (Fingar et al., 2002). In fact, mTOR overexpression in CHO-K1 cells stimulated cell cycle progression by promoting G1-to-S phase transition and increased cell size (Dreesen and Fussenegger, 2011). We showed that the assumption of constant cellular size and total RNA yield was inaccurate for several small molecule treatments, multi-cell line comparison studies and changes in culture osmolarity. Therefore, we assessed the use of spike-in standards for evaluation of differential gene expression in CHO cells, focusing on case studies relevant to cell line development and production processes.

## 2. Materials and methods

### 2.1. Cell culture and experimental treatments

Six CHO-derived cell lines (Rasmussen et al., 1998), each expressing a different monoclonal antibody (cell lines A, B, C, D, E and F), were cultured in a proprietary chemically-defined growth media in vented shake-flasks at 36 °C, 5% CO2, 70% relative humidity and shaken at 150–160 rpm in Kuhner incubators. Each cell line was generated using a proprietary expression system. A single mRNA encoded both the antibody light chain (LC) and the LC selectable marker as they were linked by an Internal Ribosome Entry Site (IRES) sequence. Similarly, single mRNA encoded both the HC and the HC selectable marker as they were also linked by an IRES sequence. Viable cell density (VCD), viability and cell diameter were measured with a ViCell automated cell counter (Beckman-Coulter, Inc., Brea, CA). Cellular volume was calculated using the formula for the volume of a sphere: $V = 4/3\pi r^3$.

For the CCI-mTORI treatment study, cells were seeded from day 4 growth cultures at $10 \times 10^6$ viable cells/mL into proprietary chemically-defined production medium in vented 24 deep-well plates (3 mL volume per well) and shaken at 220 rpm in Kuhner incubators. Cultures were treated on day 0 with either CCI (Du et al., 2014) (Amgen proprietary, PCT/US2013/074366, 10 μM final concentration), mTORI (Amgen proprietary, WO/2010/132598, 0.5 μM final concentration) or vehicle control (DMSO, 0.1% final concentration) and daily medium exchanges were performed using fresh media containing an appropriate small molecule (CCI, mTORI or DMSO). Percent DMSO was kept constant among all experimental conditions. Spent medium was used for daily titer measurements and daily specific productivity ($q_P$) calculations. Titer measurements were performed by affinity High Performance Liquid Chromatography (HPLC) using POROS A/20 Protein A column. Daily $q_P$ (pg/cell/day) was calculated according to the simplified formula: $q_P = daily\ titer/daily\ VCD$. On day 3, $3 \times 10^6$ viable cells were collected per condition for gene expression analysis (biological triplicates), snap-frozen and stored at −70 °C for further processing.

For the multi cell line study, ten-day production assays with bolus feeds on days 3, 6 and 8 were performed in chemically-defined production medium in vented shake-flasks as previously described (Fomina-Yadlin et al., 2014). Titer samples were collected on days 3, 6, 8 and 10. For each interval between days [m, n], $q_P$ was calculated according to the formula: $q_P = titer_n / \int_m^n VCDdt/(t_n - t_m)$, where $titer_n$ is the measured cumulative titer at $t_n$ and time ($t$) is expressed in days. On day 6, $3 \times 10^6$ viable cells were collected per condition for gene expression analysis (biological triplicates), snap-frozen and stored at −70 °C.

For osmolarity level study, cells were seeded from day 3 growth cultures by 1:5 split at $\sim 0.5 \times 10^6$ viable cells/mL into proprietary chemically-defined growth media in 24 deep-well plates (3 mL volume per well). Osmolarity was adjusted at 0-h time-point with 5 M NaCl solution in growth medium, as previously described (Shen et al., 2010). Daily VCD, viability and cell diameter measurements were performed. On day 2, $1 \times 10^6$ viable cells were collected per condition for gene expression analysis (biological triplicates), snap-frozen and stored at −70 °C.

### 2.2. ERCC spike-in and RNA extraction

Addition of External RNA Controls Consortium (ERCC) controls was done as previously described (Loven et al., 2012). Specifically, 1 μL of 1:10 diluted ERCC RNA Spike-In Mix 1 (Ambion®, Life Technologies, Grand Island, NY) was added per $1 \times 10^6$ cells. ERCC was added to frozen cell pellets with RLT lysis buffer, and total RNA was isolated with the RNeasy Mini kit (Qiagen, Valencia, CA) according to the manufacturer's protocol, including optional on-column DNAse I digestion, and using 100 μL elution volume. RNA concentration was measured on the Nanodrop 2000 (Thermo Scientific, Wilmington, DE), and RNA quality was assessed using the 2100 Bioanalyzer (Agilent, Santa Clara, CA) with the RNA 6000 Nano Kit (Agilent, Santa Clara, CA) to ensure all samples used for RNA-Seq analysis had RNA Integrity Number (RIN) >9.

### 2.3. RNA-Seq sample processing and analysis

RNA library preparations, sequencing reactions, and initial bioinformatics analysis were conducted at GENEWIZ, Inc. (South Plainfield, NJ). Illumina TruSeq RNA library preparation, clustering, and sequencing reagents were used throughout the process following the manufacturer's recommendations (Illumina, San Diego, CA). Briefly, 1 μg of total RNA was used as starting material for library preparation with the Illumina Truseq RNA preparation Kit V2. Poly-T oligo-attached magnetic beads were used to purify mRNA, which was then fragmented for 8 min. at 94 °C. First strand and second