

## Using Microarray Technology to Select Housekeeping Genes in CHO Cells

Scott M. Bahr, Trissa Borgschulte, Kevin J. Kayser, Matthew V. Caple and Nan Lin

Cell Sciences and Development, SAFC Biosciences 2909 Laclede Avenue, Saint Louis, MO 63103, USA

### Abstract

In the present study, we have identified species-specific housekeeping genes for Chinese Hamster Ovary (CHO) cells using data from gene expression profiling. Genes that are suitable for normalization of quantitative RT-PCR should display relatively stable expression levels across all conditions. We analyzed transcription profiles of several IgG-producing recombinant CHO cell lines under numerous growth conditions using a custom DNA microarray platform. We observed that many of the housekeeping genes commonly used in gene expression profiling in other species showed higher expression variability (typical CV% 10 - 20%). Based on relative expression level variability across over 50 arrays, we selected a novel panel of genes for which we observed stable expression in all cell lines and growth conditions (typically CV% < 10%). Particularly, this panel includes several genes with relatively low expression levels, which may be more appropriate to use in quantifying low abundance transcripts. Selected genes from the panel were used for qRT-PCR normalization as validation. The results reported here are the first in CHO cells and provide a useful tool for gene expression studies for this critical expression platform used in biotherapeutics.

### Materials and Methods

#### Cell Culture and RNA Isolation

The data in this study was obtained from 54 arrays representing 15 different experimental growth conditions that included 2 parental and 3 IgG-producing CHO cell lines. Cells were collected during the mid-logarithmic growth phase for RNA extraction (Day 4-5). RNA isolation and purification was performed with RNeasy MiniKit (Qiagen, Valencia, CA). RNA concentrations were measured using Nanodrop 1100 (Nanodrop Technologies, Wilmington, DE). RNA integrity was measured with Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). All experimental conditions were run with at least biological duplicates.

#### Hybridization

A 4x44K custom microarray platform (Agilent Technologies, Santa Clara, CA) designed with sequences from SAFC Biosciences proprietary CHO Database was used in all microarray studies (Figure 1a). Sample labeling, amplification and hybridization were performed with Agilent's 2-Color Low RNA Input Linear Amplification Kit according to manufacturer's instructions. Chips were washed in Acetonitrile (#50387 Sigma-Aldrich, St Louis, MO) and Stabilization and Drying Solution (Agilent Technologies) to minimize ozone degradation of signal.

Samples are labeled in 2-color technical duplicates and hybridized against a CHO common reference RNA Pool created from an assortment of CHO lines and conditions. This allows direct comparison across experiments using the Log2 values of the sample versus the reference pool.

#### Microarray Quality Control

Agilent's Feature extraction software 9.5 was used to perform dye normalization and QC statistics for overall array quality. Outliers and low quality probes were removed based on the software's recommendations. Expression levels between experiments were normalized using the intensity of the replicated control probes representing 50 control genes in the reference pool on each array.

#### Statistical Analysis

Stably expressed house keeping genes (HKG's) were chosen using the following statistical criteria. Each gene must have a mean expression intensity value above 40, a Log2 ratio within +/- 0.2 of the CHO Reference pool and a coefficient of variation (%CV) below 15% in all individual experiments. Commonly used HKG's such as  $\beta$ -actin (Actb) and  $\beta$ -2-Microglobulin (B2m) were compared to the selected HKG's based on the same criteria.

### Quantitative RT-PCR Validation

Primers were designed against sequences from SAFC Bioscience's CHO Sequence Database using Primer3 software [2] and ordered from Sigma Genosys. The RNA samples from our microarray studies were DnaseI (New England Biolabs, Ipswich, MA) treated followed by Oligo-dT (Sigma Genosys) primed Reverse Transcription. Biological duplicates from the array experiments were pooled for RT reactions. Samples were run in triplicate for each experimental condition and the threshold values (Ct) were averaged. Quantitative RT-PCR (qRT-PCR) was performed on a Stratagene MX3000P (Stratagene, La Jolla, CA). Reactions were run with SYBR® Green Jumpstart™ Taq ReadyMix™ (#54438, Sigma-Aldrich™) mixed with 25ng of cDNA and primers at 500 nM in a final volume of 20  $\mu$ l. Dissociation curve analysis was performed to ensure primer specificity.

### Results and Discussion

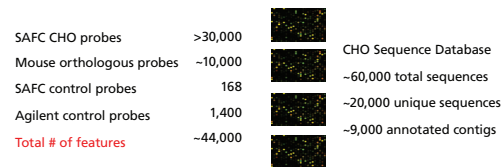


Figure 1a: Probes featured on our custom CHO Microarray (Agilent 4 X 44k)

### CHO Microarray Workflow

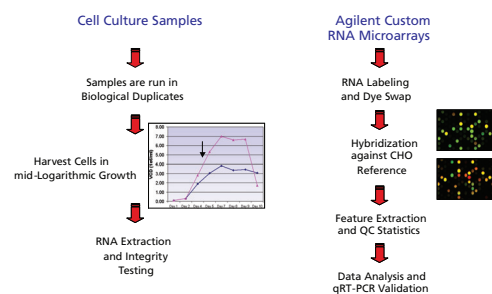


Figure 1b: Microarray analysis work flow

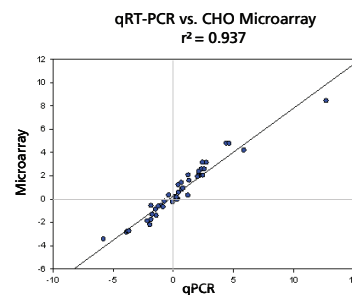


Figure 1c: Correlation of relative expression levels in microarray and qRT-PCR. Results shown are log2 ratios of 38 genes relative to Beta-2 microglobulin (B2m).

Gene Symbol	Gene Name	Location	Mean Expression	Array % CV
Eif3i	eukaryotic translation initiation factor 3, subunit I	Cytoplasm	7540	7.43%
Act5	ARP5 actin-related protein 5 homolog	Unknown	11905	8.58%
Hirip3	HIRA interacting protein 3	Nucleus	5980	8.63%
Pabpn1	poly(A) binding protein, nuclear 1	Nucleus	24600	8.68%
Cog1	component of oligomeric golgi complex 1	Cytoplasm	2560	8.74%
Ap1b1	adaptor-related protein complex 1, beta 1 subunit	Cytoplasm	11600	8.77%
Bsc1L	BCS1-like	Cytoplasm	1270	8.90%
Eftud2	elongation factor Tu GT binding domain containing 2	Nucleus	9535	8.92%
Vezt	vezatin	Plasma Membrane	1155	8.98%
Clta	clathrin, light chain	Plasma Membrane	7230	9.56%
Ube2k	ubiquitin-conjugating enzyme E2K	Cytoplasm	2000	9.57%
Pms2	postmeiotic segregation increased 2	Nucleus	409	12.76%
<b>Actb</b>	<b>beta-actin</b>	<b>Cytoplasm</b>	<b>20244</b>	<b>11.01%</b>
<b>B2M</b>	<b>beta-2-microglobulin</b>	<b>Plasma Membrane</b>	<b>7834</b>	<b>15.40%</b>
<b>Gapdh</b>	<b>glyceraldehyde-3-phosphate dehydrogenase</b>	<b>Cytoplasm</b>	<b>126466</b>	<b>19.06%</b>

Table 1: List of suggested CHO Housekeeping genes sorted by array average %CV. Probe sequences must pass a minimum quality score, have a log2 ratio below +/- 0.2 relative to the CHO RNA Reference sample and have a mean expression value above 40. Shown in the table are commonly used HKG's measured with the same statistical criteria.

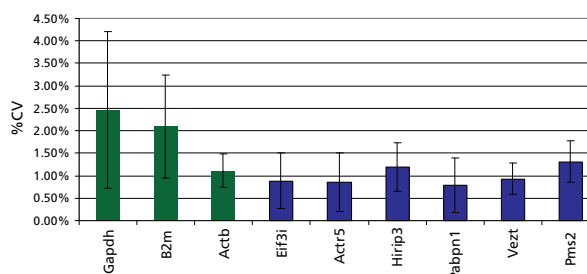


Figure 2: qRT-PCR Validation of HKG's. Graph shows average %CV of Ct values when nine culture conditions were tested in duplicate qRT-PCR reactions.

Gene Symbol	Gene Name	Average Ct Value
Eif3i	eukaryotic translation initiation factor 3, subunit I	21
Pabpn1	poly(A) binding protein, nuclear 1	23
Hirip3	HIRA interacting protein 3	25
Actr5	ARP5 actin-related protein 5 homolog	26
Vezt	Vezatin	27
Pms2	postmeiotic segregation increased 2	27
<b>Actb</b>	<b>beta-actin</b>	<b>17</b>
<b>Gapdh</b>	<b>glyceraldehyde-3-phosphate dehydrogenase</b>	<b>20</b>
<b>B2m</b>	<b>beta-2-microglobulin</b>	<b>22</b>

Table 2: Final selection of HKG's and the average Ct values in the qPCR validation

### Conclusions

- We are the first to report a novel panel of species-specific housekeeping genes (HKG) for CHO cells. These genes were selected based on expression stability from data collected from more than 15 experimental conditions in various CHO cell lines using our custom microarray platform.
- In qRT-PCR validation, the expression stability of our HKG panel is superior or comparable with commonly used HKG's such as ActB, B2m and GAPDH.
- Our HKG panel includes several low expression genes that may be useful for normalization of low abundance transcripts. These HKG's can be used in conjunction with commonly used HKG's for more accurate data normalization.

### Acknowledgements

Christoph Bausch, Research Biotechnology, Sigma-Aldrich  
Yang Liu, Bioinformatics, Sigma-Aldrich

### References

- deJonge et al (2007) Evidence based selection of housekeeping genes. PloS ONE 2(9): e898
- Rozen and Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386
- Hoogewijs et al (2008) Selection and Validation of a set of reliable reference genes for quantitative sod gene expression analysis in C. elegans. BMC Molecular Biology. 9(9).