

Using Microarray Technology to Select Housekeeping Genes in Chinese Hamster Ovary 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 (HKGs) for Chinese Hamster Ovary (CHO) cells using data from microarray gene expression profiling. HKGs suitable for quantitative RT-PCR normalization should display relatively stable expression levels across experimental conditions. We analyzed transcription profiles of several IgG-producing recombinant CHO cell lines under numerous culture conditions using a custom CHO DNA microarray platform and calculated relative expression variability from 124 microarrays. We selected a novel panel of candidate HKGs based on their low variability in expression from the microarray data. Compared to three traditional HKGs (Gapdh, Actb and B2m), the majority of genes on this panel demonstrated lower or equal variability. Each candidate HKG was then validated using qRT-PCR. Final selection of CHO-specific HKGs include Actr5, Eif3i, Hirip3, Pabpn1, Vezt, Cog1 and Yaf2. The results reported here provide a useful tool for gene expression analyses in CHO cells, a critical expression platform used in biotherapeutics.

Materials and Methods

Cell Culture and RNA Isolation

The cell culture experiments used in the present study involved 25 different experimental growth conditions including two parental and four IgG-producing CHO cell lines (CHO K1 lineage). All cell lines used were cultured in serum-free, animal component free suspension media. Our dataset represented batch and fed-batch cultures in 125-ml shake flasks, TPP bioreactor tubes and 6-well static cultures, with numerous chemically-defined or plant hydrolysate containing media and treatments such as sodium butyrate. All experiments were conducted in biological duplicates. Parental and recombinant cell lines were all included in the analyses. Cells were collected during the early- and mid-logarithmic as well as stationary phases for RNA isolation. Total RNA isolation was performed with RNeasy Mini Kit (Qiagen, Valencia, CA). RNA concentrations were measured using Nanodrop 1100 (Nanodrop Technologies, Wilmington, DE). RNA integrity was analyzed with Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA).

Microarray Design, Hybridization and Expression Data Analysis

An internally developed CHO DNA sequence database containing over 60,000 total sequences and 9,000 annotated contigs was used to design the 4x 44K custom microarray platform (Agilent Technologies, Santa Clara, CA). The average probe length is 60-bp, with a melting temperature (Tm) of 70 °C and a GC content of 45%. Each sequence has two probes designed to hybridize with different regions. In the instance that known genes were not represented in our sequence database, mouse orthologous gene sequences were used for probe design. This array platform was used in all 124 microarray studies.

A common RNA reference pool was created from an assortment of CHO lines and conditions in order to directly compare across experiments. Dye-swap technical replicates were conducted as following: RNA samples were labeled with Cy5 or Cy3 and hybridized against the common reference pool labeled with Cy3 or Cy5, respectively.

Sample labeling, amplification and hybridization were performed with the 2-Color Low RNA Input Linear Amplification Kit (Agilent Technologies). Slides were washed in acetonitrile (Sigma-Aldrich, St. Louis, MO) and treated with Stabilization and Drying Solution (Agilent Technologies) according to manufacturer's protocol. Agilent Feature Extraction software 9.5 (FE 9.5) was used to perform dye normalization and QC statistics based on manufacturer's protocols.

Gene Symbol	RefSeq ID	Microarray Probe	Microarray % CV	Microarray Normalized Intensity*
Actr5	NM_175419	CHO Probe 1	21.51%	64.30
		CHO Probe 2	23.66%	15.17
Ampd2	NM_028779	CHO Probe 1	29.41%	10.96
		CHO Probe 2	22.46%	23.12
		Mouse Probe	19.29%	23.10
Ap3d1	NM_007460	CHO Probe 1	36.14%	9.19
		CHO Probe 2	26.90%	13.89
Cog1	NM_013581	CHO Probe 1	23.74%	10.02
		CHO Probe 2	24.81%	7.15
		Mouse Probe	18.57%	2.02
Eftud2	NM_175317	CHO Probe 1	30.43%	8.21
		CHO Probe 2	27.05%	9.13
		Mouse Probe	25.72%	1.80
Eif3i	NM_018799	CHO Probe 1	18.96%	318.23
		CHO Probe 2	22.52%	77.36
Hirip3	NM_172746	CHO Probe 1	26.72%	64.01
		CHO Probe 2	30.57%	28.01
Nap1L1	NM_015781	CHO Probe 1	25.03%	87.61
		CHO Probe 2	31.05%	19.02
Pabpn1	NM_019402	CHO Probe 1	23.80%	96.24
		CHO Probe 2	26.15%	42.33
Rabep1	NM_019400	CHO Probe 1	18.66%	2.48
		CHO Probe 2	26.35%	6.72
		Mouse Probe	18.59%	2.59
Yaf2	NM_024189	CHO Probe 1	25.48%	8.93
		CHO Probe 2	21.37%	24.36
		CHO Probe 3	28.48%	14.28
		CHO Probe 4	20.07%	5.25
Vezt	NM_172538	CHO Probe 1	19.42%	3.53
		CHO Probe 2	13.53%	4.93
Actb	NM_007393	30 hamster probes	27.13%	109.05
B2m	NM_009735	29 hamster probes	38.23%	62.02
Gapdh	NM_008084	61 hamster probes	44.97%	544.16

* Ratio of arbitrary fluorescent intensity values

Table 1. List of CHO Housekeeping Genes Selected from Microarray Gene Expression Analysis

Statistical Analysis

The relative expression levels of each probe were normalized using the median fluorescence intensity of the reference pool in each dye-swap set. CHO-specific HKG candidates were selected using the following criteria. The coefficient of variation (% CV) was calculated using the expression value of each probe normalized with the median intensity of the reference pool. Each gene must have at least one microarray probe with a % CV below 29.21 % (25% percentile of all probes analyzed, (Figure 1).

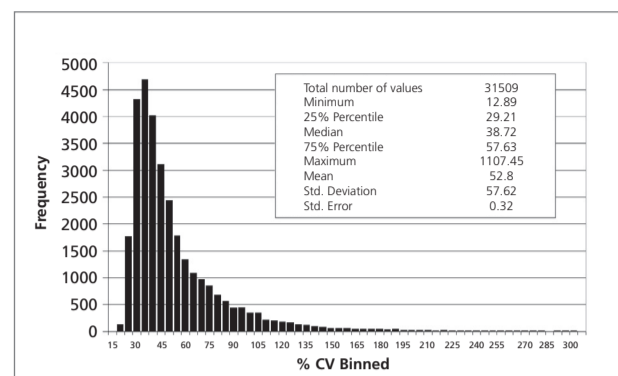


Figure 1. Depicts the distribution of the % CV values from the 31509 probes on 124 microarrays used for the present analysis. Quality control and background estimation probes on the microarray were excluded. The 25% percentile value is 29.21%, which was used for the primary cutoff for CHO HKG selection. The data points (n = 232) with % CV greater than 300% were not shown on this histogram.

Quantitative RT-PCR Validation

Primers were designed against sequences from the internal CHO DNA sequence database using Primer3 software (<http://fokker.wi.mit.edu/primer3/input.htm>). The primers were 18–25 base pairs (bp) with Tm of 64–65 °C. The amplicons were between 156–249 bp (Table 2).

All amplification efficiencies were calculated based on at least 4 dilution points and within a range of 90%–112%. A minimum of ten experimental conditions were randomly selected for qRT-PCR validation, with biological duplicates from the microarray experiments pooled for RT reactions. qRT-PCR was performed on a MX3000P thermocycler (Stratagene, La Jolla, CA) in triplicate for each experimental condition, with threshold cycle values (Ct) averaged from the triplicates. Reactions were run with SYBR Green Jumpstart™ Taq ReadyMix™ (Sigma-Aldrich, S4438) mixed with 1 mL of cDNA prepared from 25 ng/mL RNA and primers at 500 nM in a reaction volume of 20 mL.

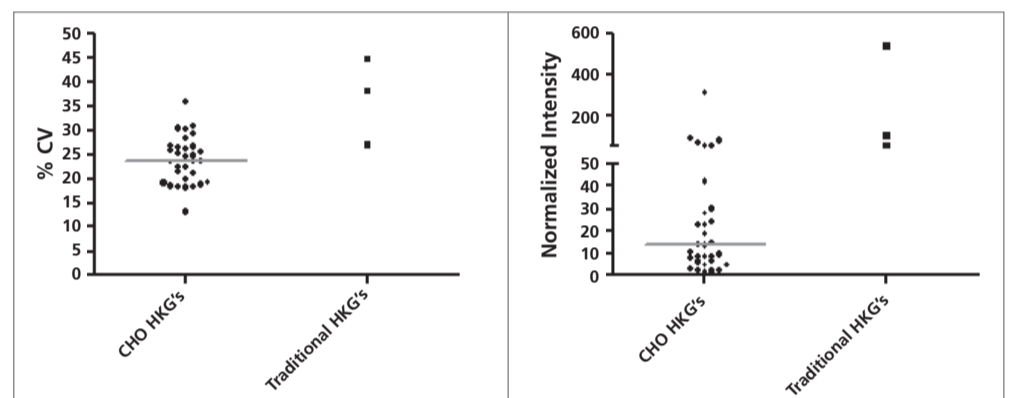


Figure 2. Left panel: % CV distribution of 31 probes of the 12 CHO HKG candidates and three traditional HKGs (Gapdh, Actb and B2m).

Right panel: Relative expression level distribution of the 12 CHO HKG candidates and three traditional HKGs. Each dot in the scatter dot plots represents one microarray probe. Horizontal line represents the median. Average of all the probes of each gene is depicted for the traditional HKGs.

Gene Symbol	Mean Ct	SD	Coefficient of variation	Amplification Efficiency (%)	Description	Location
Eif3i	20.59	0.89	4.48%	92	Eukaryotic translation initiation factor 3, subunit I	Cytoplasm
Pabpn1	22.4	0.6	4.74%	102	Poly(A) binding protein, nuclear 1	Nucleus
Cog1	23.08	1.34	7.22%	112	Component of oligomeric golgi complex 1	Cytoplasm
Actr5	24.75	0.64	5.67%	112	ARP5 actin-related protein 5 homolog (yeast)	Unknown
Hirip3	24.61	0.46	5.21%	106	HIRA interacting protein 3	Nucleus
Yaf2	25.58	1.53	3.30%	90	YY1 associated factor 2	Nucleus
Vezt	25.91	0.95	6.79%	102	Vezatin, adherens junctions transmembrane protein	Plasma Membrane
Actb	16.89	1.54	9.11%	112	Actin, beta	Cytoplasm
B2m	20.26	1.9	9.38%	93	Beta-2-microglobulin	Plasma Membrane
Gapdh	16.74	1.36	8.12%	89	Glyceraldehyde-3-phosphate dehydrogenase	Cytoplasm

*Sub-cellular location is based on Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com).

Table 2. List of final selection CHO HKGs with relative expression levels (Ct) and sub-cellular location*

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 25 experimental conditions in various CHO cell lines using our custom microarray platform.
- In qRT-PCR validation, the expression stability for 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

Carol Kreader, Brian Ward, Christoph Bausch, Yang Liu – Sigma Aldrich Biotech