Normalization strategies for real-time expression data in *Chlamydia trachomatis*

Vítor Borges [a], Rita Ferreira [a], Alexandra Nunes [a], Paulo Nogueira [b], Maria J Borrego [a], João P Gomes [a]

[a] Department of Infectious Diseases, National Institute of Health, Av. Padre Cruz, 1649-016 - Lisbon, Portugal
[b] Department of Epidemiology, National Institute of Health, Av. Padre Cruz, 1649-016 - Lisbon, Portugal

Introduction

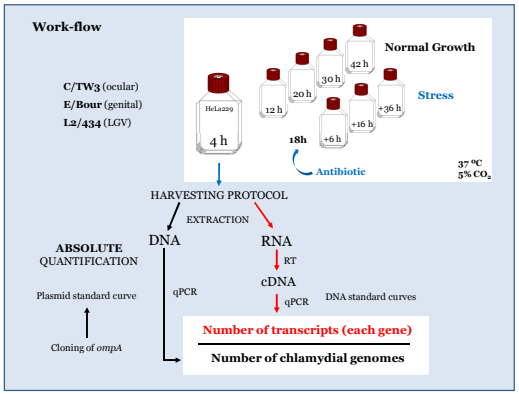
Transcriptomics has been a fundamental approach to get insights into the biology of the genetically non-tractable *C. trachomatis* and gene expression has been mostly evaluated by real-time quantitative reverse transcription PCR (RT-qPCR). A proper normalization strategy is probably the major requirement when performing relative expression assays, where inappropriate methodologies can lead to inaccurate data and incorrect conclusions. In fact, it is mandatory that the mRNA levels of the endogenous control gene (the most commonly employed strategy) do not vary significantly throughout the biological process under study. In the case of *C. trachomatis*, **the validation of the suitability of endogenous controls** for normalization **is still lacking**. The present study constitutes the first evaluation of putative endogenous control genes for real-time expression assays in *C. trachomatis*.

Objectives

• **To assess the expression stability of 10 genes** for their potential use as endogenous controls in RT-qPCR assays at both normal and stress (antibiotic treatment) growth conditions throughout the developmental cycle of three *C. trachomatis* strains with different tissue tropism.

• **To test the applicability of geNorm** and **Normfinder** to our data. These are two widely employed softwares used for the assessment of the most stably expressed genes.

Methodology



• Selected genes: *yraL*, *fer*, CT147, *oppA_2*, *radA*, *yaeI*, *hemN_2*, *tyrP_2*, *map* and *16SrRNA* (based on a previous microarrays study [1]). Information about these genes is described in Borges *et al*, 2010 [2].

• T_{25} cm² flasks of confluent HeLa 229 cell monolayer were inoculated with *C. trachomatis* prototype strains representing distinct disease groups: C/TW3 (ocular), E/Bour (genital) and L2/434 (LGV).

• The developmental cycle under normal conditions was interrupted at time-points 4, 12, 20, 30 and 42 h post-infection (pi). For cultures under D-cycloserine treatment [30µg/ml at 18 h (pi)] the disruption of the developmental cycle was performed at 6, 16 and 36 h post-treatment (pt).

• All flasks were subjected to the same cell harvesting protocol described in Borges *et al*, 2010 [2]. The final supernatant was collected, homogenized and rigorously divided into **two identical aliquots for independent DNA and RNA extractions**.

• Random hexamers were used for all reverse transcription (RT) reactions. cDNA obtained at each time-point was used for transcripts quantification, whereas the extracted DNA was used for cloning-based absolute quantification of bacterial genomes.

• **The number of transcripts** of each gene **was normalized against the number of chlamydial genomes** determined for each time-point.

• The evaluation of the expression stability was assessed by the calculation of the absolute fold-difference between the normalized expression value at each time-point and the mean expression value determined for the entire infectious cycle under normal growth conditions.

• Gene expression stability was also evaluated using two MS Excel applications, geNorm (version 3.5) [3] and Normfinder (version 0.953) [4].

• The final results were based on three independent experiments for all strains.

References

1. Belland, R. J., Zhong, G., Crane, D. D., Hogan, D., Sturdevant, D., Sharma, J., Beatty, W. L., Caldwell, H. D., 2003. Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. Proc Natl Acad Sci USA. 100, 8478-8483.
2. Borges V., Ferreira R., Nunes A., Nogueira P., Borrego M.J., Gomes J.P., 2010. Normalization strategies for real-time expression data in *Chlamydia trachomatis*. J Microbiol Methods. 82, 256-264.
3. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3, research0034.1-0034.11.
4. Andersen, C. L., Jensen, J. L., Orntoft, T. F., 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64, 5245-5250.

Results

Part 1 - Chlamydial growth rate

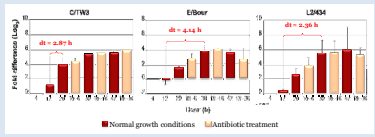


Fig. 1. Growth profile of C/TW3, E/Bour and L2/434 prototype strains. For each strain, the results at each time-point are represented as a mean ± SD of the logarithmic fold-difference of the genome copy number (in each T_{25} cm² flask) compared to 4 h (pi). Antibiotic was added at 18 h (pi). **dt – doubling time during the exponential phase**.

• Overall increase in chlamydial genomes (from 4 h to 42 h pi): **43-fold** (C/TW3), **22-fold** (E/Bour) and **188-fold** (L2/434).

• Doubling times (dt): **2.87 h** (SD ±0.25) (C/TW3), **4.14 h** (SD ±1.05) (E/Bour) and **2.36 h** (SD ±1.65) (L2/434).

• The antibiotic treatment did not seem to affect the normal increase of genome copy number.

Part 2 - Gene expression stability evaluation

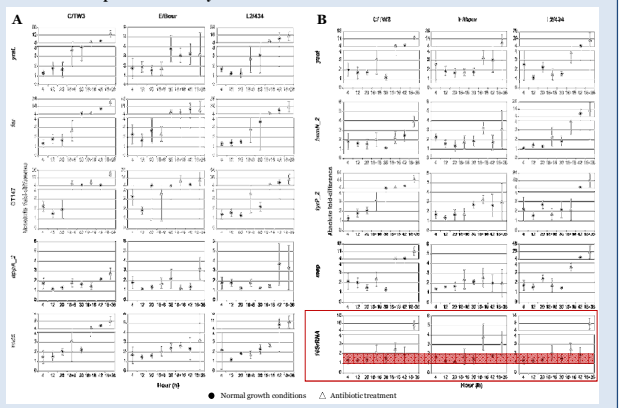


Fig. 2. Evaluation of expression stability of the 10 genes for the three prototype strains. Each graph represents the mean (± SD) absolute fold-difference of expression values of each time-point compared to the mean expression value calculated for the entire developmental cycle under normal growth conditions. Expression values were normalized against the number of bacterial genomes. All RT reactions were performed by using random hexamers.

Part 3 – Top-ranked genes

Table 1. Top-ranked genes experimentally determined.

C/TW3				E/Bour				L2/434			
Normal Conditions		Antibiotic Treatment		Normal Conditions		Antibiotic Treatment		Normal Conditions		Antibiotic Treatment	
Best-ranked Genes	Σ FD*	Best-ranked Genes	Σ FD*	Best-ranked Genes	Σ FD*	Best-ranked Genes	Σ FD*	Best-ranked Genes	Σ FD*	Best-ranked Genes	Σ FD*
16SrRNA	6.27	oppA_2	6.06	**16SrRNA**	7.46	oppA_2	6.34	**16SrRNA**	7.46	oppA_2	6.64
oppA_2	8.13	hemN_2	8.27	oppA_2	7.92	map	6.96	oppA_2	9.85	16SrRNA	13.39
hemN_2	8.81	16SrRNA	11.1	map	8.77	tyrP_2	8.63	hemN_2	14.17	radA	15.05

* sum of the absolute fold-differences (FD) between each time-point and the mean calculated for the whole cycle

Table 2. Top-ranked genes according to **geNorm** and **Normfinder** applications

C/TW3		E/Bour		L2/434	
geNorm	NormFinder	geNorm	NormFinder	geNorm	NormFinder
yraL	*radA*	*radA*	*radA*	*yraL*	*yaeI*
fer	*yraL*	*tyrP_2*	*map*	CT147	*radA*
tyrP_2	*yaeI/map*	*map*	*yraL*	*fer*	*hemN_2*

The softwares were only applied to the expression data obtained during the normal growth conditions.

• ***16SrRNA*** was clearly the most stable gene under **normal growth conditions**.

• ***oppA_2*** was the least variable gene under **antibiotic treatment**.

• All other genes showed an incoherent rank position among strains.

• For the top-ranked genes, **no agreement** was found to exist between the results of **both softwares** for each strain.

• Furthermore, the experimentally best-ranked genes, *16SrRNA* and *oppA_2*, were never placed among the most stable genes according to geNorm and Normfinder.

Normal conditions

• **Most genes** showed **variation above four-fold** to the respective mean.

• ***16SrRNA*** was the only gene for which expression variation **was always below two-fold**.

• *oppA_2* and *hemN_2* also exhibited variations below two-fold except for one time-point for each strain.

Antibiotic treatment

• ***oppA_2*** and ***hemN_2*** were **the least variable genes** but the expression variation was up to four-fold to the mean value (12.7-fold for *hemN_2* for L2/434).

Part 4 - "Fine-tune" of *16SrRNA* by RT target-specific priming

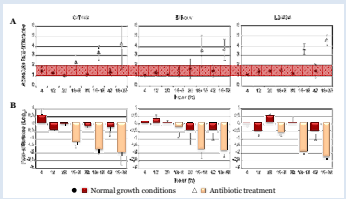


Fig. 3. RT-qPCR results of *16SrRNA* by using RT target-specific priming. (A) represent the mean (± SD) absolute fold-difference of expression values of each time-point compared to the mean expression value calculated for the entire development cycle under normal growth conditions. (B) *16SrRNA* expression profile throughout development. It depicts the same data as in (A) but the logarithm transformation enables the representation of up- or down-regulation phenomena for each time-point.

• Expression results obtained with RT target-specific priming mirrored the ones obtained with random hexamers.

• Although ***16SrRNA*** was the most stable gene under normal growth conditions, its expression was clearly **down-regulated in the antibiotic-treated cultures.**

Conclusions

• ***16SrRNA* was the most stably expressed gene throughout the chlamydial normal developmental cycle**, which consolidates its use as an endogenous control in relative expression assays. However, it was very unstable under antibiotic treatment, suggesting prudence when using ribosomal genes in expression experiments involving stress conditions.

• ***oppA_2* was the least affected under antibiotic treatment**, and may be a good endogenous control candidate to be used in upcoming gene expression studies when other stress conditions are implicated.

• **Genomic DNA likely constitutes a good option for data normalization** since it is less influenced by experimental constraints that are especially relevant for intracellular organisms and it also seems to be less subject to variation than expression of endogenous control genes when working under stress conditions. However, a drawback of this strategy is that the genomic DNA (denominator) is independently extracted from the RNA (numerator) which requires high reproducibility of both extraction methods.

• The **geNorm** and **Normfinder** algorithms revealed **contrasting results** and seem **inappropriate for the selected pool of genes** and they should be used with caution for any other organism for which no "guiding" transcriptome is available yet, because these softwares require a previous raw knowledge of gene stability.

Acknowledgements

This work was supported by a grant from Fundação para a Ciência e a Tecnologia (FCT) (PTDC/BIA-BCM/71117/2006) (to JPG) which funds the fellowships of VB and RF. AN was the recipient of a Ph.D. grant (SFRH/BD/25651/2005) from FCT.

This work was published at the Journal of Microbiological Methods, 2010, 82, 256-264.