Data quality control and statistics of qPCR

Data quality control for qPCR

In order to extract meaningful conclusions for the qPCR data, we ensure that the conditions of our PCR reaction are optimized so that the efficiency of the reactions is close to 1 (when efficiency is defined as percentage from 0 to 1) meaning that the amount of template doubles in each cycle. Efficiency of the qPCR for the analysis of the housekeeping gene expression (Gapdh and Gus) was calculated by examining the correlation between Ct number and different concentration of some control tissues using cDNA serial dilutions, and was always close to 1.

To study the relative expression of K^+ channels in BPH versus BPN we used Taqman low density arrays. All the expression assays used were inventoried and validated by the manufacturer, so that under their specified conditions for the amplification, the efficiencies of all the genes studied are also 1 and differ among them in less than 10%, so that the ΔΔCt method could be applied [1].

Statistical analysis of qPCR data

Statistical comparisons of the expression levels of the channel genes under study in BPN and BPH tissues were carried out with the ΔCt values of each sample:

$$Ct_{Channel} - Ct_{housekeeping\ gene}$$

As there is some controversy regarding the election of the most adequate endogenous control for obtaining the ΔCt value, we performed the statistical analysis with two different housekeeping genes, RP18s and Gapdh. Ribosomal protein 18s has the advantages of being an ubiquitous and relatively stable gen that is included in the Taqman low density arrays so that is also an internal control of the reaction. However, its levels of expression are high and therefore it could be less sensitive to variations in expression in the range of the genes of interest. On the contrary, Gapdh, whose expression levels are closer to the levels of the genes under study, serves as external control as its expression is determined from the same samples in an independent reaction. The data obtained with the two sets of ΔCt values showed no important discrepancies, as illustrated in the table II and III (see below), so we used RP18s as the housekeeping gene for representing the data.

For statistical comparisons, the ΔCt values obtained in each sample (control, BPN or experimental, BPH) are subtracted from the average ΔCt of the calibrator (BPN samples) in order to obtain a value of SEM for the control data. Statistical analysis was

performed with the STATGRAPHICS Centurion XVI software (StatPoint Technologies Inc.) using different test for comparison between two samples.

For each set of data, we evaluated first if the data distribution could be adjusted to a normal distribution by using the Saphyro-Wilks test, where p values > 0.05 indicate a normal distribution (null hypothesis). As this was the case for all data set in our study, we applied a student t-test for independent samples to determine if we have statistically significant differences when comparing the means of the two groups (BPN and BPH). In addition, we also use the Mann-Whitney-Wilcoxon test**,** a non-parametric test, that represents a more robust statistical analysis as it compares the median of the two groups to decide whether two populations are equal or not without assuming them to follow the normal distribution.

Although several statistical approaches can be taken to analyze real time data, including multiple regression models, a recent study comparing several statistical methods indicates that more simplified alternatives such as t-test and Wilcoxon two group tests can be used with no differences in the final outcome [2]

Using these tools, we evaluated first the distribution and the possible existence of differences in the expression levels of the two housekeeping genes, RP18s and Gapdh. As shown in Table I, when comparing the data set obtained in BPN and BPH samples, they both show a normal distribution with no statistical differences. After confirming this point, we performed the same statistical analysis on the ΔCt data obtained for each gene under studied in both conditions (BPN and BPH) and using the two housekeeping genes, RP18s (Table II) and Gapdh (Table III).

Data in the tables show the results of this analysis in mesenteric arteries, in which we did not detect expression of Kir1.1 and Kir3.x mRNA in any condition. As indicated in the table, the choice of housekeeping gene did not affect the results and with the exception of SUR2, the statistical analysis of data show no differences between the two tests used. We provide also all p values obtained with both tests (t-test and Mann-Whitney-Wicoxon test), derived from testing the null hypothesis that the differences between ΔCt values in both groups (that is, ΔΔCt) are equal to 0.

References

Livak KJ & Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.

Yuan JS, Reed A, Chen F & Stewart CN, Jr. (2006). Statistical analysis of real-time PCR data. *BMC Bioinformatics* **7**, 85.

Housekeeping genes			Saphyro-Wilks test		Student t-test for independent sample			Mann-Whitney -Wilcoxon test	
	BPN Ct (mean±SD)	BPH Ct (mean±SD)	p BPN	P BPH	t value	p value	Confidence interval	W value	p value
RP18s	11.72±0.62	11.31±0.60	0.083	0.190	1.48671	0.15707	(-0.21592, 1.02792)	31	0.16587
Gapdh	19.78±0.87	19.73±0.86	0.437	0.139	0.04626	0.96361	(-0.79990, 0.83590)	50	0.96984

Table I. Statistical analysis of the Ct values obtained for the two housekeeping genes.
Data are calculated from duplicate samples obtained in five different assays

RP18s relative expression			Saphyro-Wilks test		Student t-test for independent samples			Mann-Whitney -Wilcoxon test	
	BPN ΔCt (mean±SD)	BPH ΔCt (mean±SD)	p BPN	p BPH	t value	p value	Confidence interval 95%	W value	p value
SUR2	17.37±0.88	18.10±0.70	0.270	0.340	-2.18973	**0.04200**	(-1.55771, -0.011029)	65.0	0.11134
Kir6.1	17.99±0.17	19.39±0.60	0.091	0.064	-4.92194	**0.00011**	(-2.00040, -0.803558)	98.0	**0.00033**
Kir6.2	21.57±1.10	23.80±1.70	0.300	0.100	-3.25101	**0.00470**	(-3.67648, -0.782633)	82.5	**0.00251**
Kir2.1	21.35±1.20	22.59±0.92	0.100	0.900	-2.41363	**0.02814**	(-2.32701, -0..50763)	64.0	**0.04225**
Kir2.2	26.61±0.62	26.80±0.49	0.780	0.890	-0.77180	0.45024	(-0.71278, 0.329781)	59.5	0.49580
Kir4.1	22.17±0.79	23.44±0.77	0.490	0.590	-3.64625	**0.00185**	(-2.00649, -0.539512)	88.0	**0.00455**

Table II. Comparison of ΔCt values obtained for all the channel genes showing expression in BPN and BPH mesenteric arteries, using RP18s as endogenous control.
Data are calculated from duplicate samples obtained in five different assays.
Values in bold indicate significant differences between BPN and BPH data.

Gapdh relative expression			Saphyro-Wilks test		Student t-test for independent samples			Mann-Whitney -Wilcoxon test	
	BPN ΔCt (mean±SD)	BPH ΔCt (mean±SD)	p BPN	p BPH	t value	p value	Confidence interval95%	W value	p value
SUR2	9.26±0.30	9.64±0.40	0.65	0.45	-1.77089	0.10195	(-0.86665, 0.08951)	38.0	0.09669
Kir6.1	9.96±0.50	10.98±0.43	0.60	0.24	-4.87243	**0.00012**	(-1.4483, -0.57564)	96.0	**0.00058**
Kir6.2	13.47±0.50	14.82±0.60	0.93	0.2	-4.99773	**0.00013**	(-1.91947, -0.77608)	77.0	**0.00148**
Kir2.1	13.14±0.92	14.29±0.84	0.59	0.89	-2.68162	**0.01707**	(-2.05334, -0.23471)	58.5	**0.03415**
Kir2.2	14.77±0.71	14.22±0.68	0.10	0.58	1.36479	0.20224	(-0.34898, 1.45232)	19.0	0.19939
Kir4.1	14.08±0.30	14.97±0.60	0.37	0.25	-3.96475	**0.00090**	(-1.42890, -0.43907)	90.0	**0.00281**

Table III. Comparison of ΔCt values obtained for all the channel genes showing expression in BPN and BPH mesenteric arteries, using Gapdh as endogenous control.
Data are calculated from duplicate samples obtained in five different assays.
Values in bold indicate significant differences between BPN and BPH data.