

Basic analyses - Legend to figures (PCa)

UPDATE (February 2022):

- All datasets have the gene information stored with the 2021 [HGNC](#) approved nomenclature.
- Dataset normalization was performed by original authors, and no additional normalization process was applied upon download (their corresponding normalization is described in the corresponding GEO entry).
- Visit the [Datasets](#) tab to access feature information for each of the datasets.
- In microarray datasets, probes interrogating the same gene have been averaged.
- For microarray derived datasets, all probes are included. However, it is important to note that some platforms identify potentially mis-annotated or cross-hybridizing probes. We recommend users to check their genes of interest in the case of the datasets indicated with * in the [Datasets](#) tab.

A. Status in prostate cancer: Violin plots depicting the expression of the gene of interest between non-tumoral (N) and prostate cancer specimens (PCa) in the indicated datasets. The Y-axis represents the Log₂-normalized gene expression (fluorescence intensity values for microarray data or, sequencing reads values obtained after gene quantification with RSEM and normalization using Upper Quartile in case of RNAseq). A Student T-test is performed in order to compare the mean gene expression between two groups.

B. Status by progression: Violin plots depicting the expression of the gene of interest among non-tumoral (N), primary tumor (PT) and metastatic (M) PCa specimens in the indicated datasets. The Y-axis represents the Log₂-normalized gene expression (fluorescence intensity values for microarray data or, sequencing reads values obtained after gene quantification with RSEM and normalization using Upper Quartile in case of RNAseq). An ANOVA test is performed in order to compare the mean gene expression among two groups.

C. Status by Gleason Grade: Violin plots depicting the expression of the gene of interest among PCa specimens of the indicated Gleason grade in the indicated datasets. The Y-axis represents the Log₂-normalized gene expression (fluorescence intensity values for microarray data or, sequencing reads values obtained after gene quantification with RSEM and normalization using Upper Quartile in case of RNAseq). Gleason grade is indicated as GS6, GS7, GS8, GS8+9, GS9, GS10. An ANOVA test is performed in order to compare the mean among groups.

D. Disease Free Survival: Kaplan-Meier curves representing the disease-free survival (DFS) of patient groups selected according to the quartile expression of the gene of interest. Quartiles represent ranges of expression that divide the set of values into quarters. Quartile color code: Q1 (Blue), Q2 + Q3 (Green), Q4 (Red). Each curve represents the percentage (Y-axis) of the population that exhibits recurrence of the disease along time (X-axis, in months) for a given gene expression distribution quartile. Vertical ticks indicate censored patients. Quartile colour code: Q1 (Blue), Q2 plus Q3 (Green), Q4 (Red). A Mantel-Cox test is performed in order to compare the differences between curves, while a Cox proportional hazards regression model is performed to calculate the Hazard Ratio (HR) between the indicated groups.

Note

- All the graphs generated with CANCEERTOOL are produced in PNG (600 ppp) and PDF, which can be edited for example with [*Inkscape*](#) or [*GIMP*](#).
- All the p-values shown on the figures are non-adjusted. Adjusted p-values, calculated using Benjamini & Hochberg (BH) method, are only provided in the Statistical Results table when custom analyses are requested.