Microarray-based detection of genomic signatures related with the tumor recurrence in Glioblastoma patients

Álvaro Abella-Bascarán[*,1], **Eloi Casals-Puig**[*] **and Samuel Miravet-Verde**[*]
[*]Pompeu Fabra University, Barcelona (Spain)

ABSTRACT

Glioblastoma tumors, in addition to be the most frequent and aggressive type of brain tumor in humans, are notorious for their resistance to therapy. The aim of this work was to identify which molecular profiles are related with the resistance to treatment, by means of the analysis of microarray data from 80 tumor samples (proceeding from the work of (Murat *et al.* 2008)). The results show ...

KEYWORDS Microarray; tumour recurrence; glioblastoma.

Introduction

Glioblastoma multiforme is the most frequent and aggressive brain tumor in humans, involving glial cells, with an incidence of 2–3 cases per 100,000 person life-years in Europe and North America(Bleeker *et al.* 2012). Its treatment can involve chemotherapy, radiation and surgery. Median survival with standard-of-care radiation and chemotherapy with the alkylating agent temozolomide is only 15 months(Johnson *et al.* 2012) while the median survival without treatment is 4 and a half months.

Regretfully glioblastomas are notorious for resistance to therapy, which has been attributed to DNA-repair proficiency, a multitude of deregulated molecular pathways, and, more recently, to the particular biologic behavior of tumor stem-like cells, as it is exposed in the work of Anastasia Murat(Murat *et al.* 2008). In that case the HOX and EGFR related pathways were identified as differentially expressed using several cluster procedures. However, a deeper analysis based on more general techniques can be able to determine the molecular profiles specific for treatment resistance. To achieve that goal, the same set of gene expression profiles of 80 patients has been used.

Materials and Methods

Tumor Samples and Patient Characteristics

We analyzed data from 80 frozen glioblastoma samples. The data comprised 70 tumors from initial surgery and 10 samples resected at recurrence. All patients were treated within a phase II or a randomized phase III trial(Stupp *et al.* 2002, 2005) The study includes 21 females and 55 males, with a a median age of 52 (range, 26 to 70 years). Out of the 76 patients, 28 received radiotherapy treatment only, and 48 received TMZ/radiotherapy treatment.

doi: 10.1534/genetics.XXX.XXXXXX
Manuscript compiled: Saturday 6th June, 2015%
[1]Affiliation correspondence email: alvaro.abella01@estudiant.upf.edu

Gene Expression Profiling

The microarray data with gene expression profiling was obtained from the Gene Expression Omnibus (GEO) database at http://www.ncbi.nlm.nih.gov/geo/ (accession-number GSE7696). The data was created by the study [XXX] from probes prepared with the Enzo BioArray-High Yield Kit (Enzo Life Sciences, Farmingdale, NY) for double amplification and were hybridized to Affymetrix HG-133Plus2.0 GeneChips (Affymetrix, Santa Clara, CA). The data used had been normalized to the expression of the EIF2C3, DNAJA4, and B2M genes that exhibited little variation in the data set.

Data Analysis and Statistical Methods

Analyses were carried out in R, a free software environment available at http://www.r-project.org/. The quality of the microarray data was assessed by a variety of quality checks. Raw chip images were visually inspected to ensure the absence of artifacts. The intensity distributions of the samples showed a similar exponential distribution, without any artifactual distribution. We used the linear probe level model (PLM) (Bolstad 2004; Brettschneider *et al.* 2007) to verify the absence of artifacts in the chip pseudoimages created from the weights and

residuals of the sample PLM's. We used Normalized Unscaled Standard Errors (NUSE) (Bolstad 2004) to evalute the deviation of the chip probsets. Samples with NUSE median value higher than 1.05 were removed. Using the same models, we also evaluated the Relative Log Expression (RLE) values (Bolstad 2004; Brettschneider *et al.* 2007) to determine technical biases on particular chips. No particular deviation of the median or the interquartile range was found on any sample. The expression intensities for all probe sets from Affymetrix CEL-files were estimated using robust multiarray average with probe-level quantile normalization followed by median polish summarization (Irizarry *et al.* 2003) as implemented in the BioConductor software (http://www.bioconductor.org/). The inspection of MA plots ensured the absence of fluorescent intensity dependent biases (Bolstad 2004). After the quality assurance process, NNN out of the MMM samples were discarded from the analysis.

Batch effect, blablabla

The Benjamini-Hochberg procedure was applied for multiple testing correction (false-discovery rates) (Benjamini and Hochberg 1995)

blablabla

Results and Discussion

The results and discussion should not be repetitive. The results section should give a factual presentation of the data and all tables and figures should be referenced; the discussion should not summarize the results but provide an interpretation of the results, and should clearly delineate between the findings of the particular study and the possible impact of those findings in a larger context. Authors are encouraged to cite recent work relevant to their interpretations. Present and discuss results only once, not in both the Results and Discussion sections. It is sometimes acceptable to combine results and discussion. The text should be as succinct as possible. Heed Strunk and White's dictum: "Omit needless words!"

Additional guidelines

Numbers

In the text, write out numbers nine or less except as part of a date, a fraction or decimal, a percentage, or a unit of measurement. Use Arabic numbers for those larger than nine, except as the first word of a sentence; however, try to avoid starting a sentence with such a number.

Units

Use abbreviations of the customary units of measurement only when they are preceded by a number: "3 min" but "several minutes". Write "percent" as one word, except when used with a number: "several percent" but "75%." To indicate temperature in centigrade, use ° (for example, 37°); include a letter after the degree symbol only when some other scale is intended (for example, 45°K).

Nomenclature and Italicization

Italicize names of organisms even when when the species is not indicated. Italicize the first three letters of the names of restriction enzyme cleavage sites, as in HindIII. Write the names of strains in roman except when incorporating specific genotypic designations. Italicize genotype names and symbols, including all components of alleles, but not when the name of a gene is the same as the name of an enzyme. Do not use "+" to indicate wild type. Carefully distinguish between genotype (italicized) and phenotype (not italicized) in both the writing and the symbolism.

Examples of Article Components

The sections below show examples of different header levels, which you can use in the primary sections of the manuscript (Results, Discussion, etc.) to organize your content.

First level section header

Use this level to group two or more closely related headings in a long article.

Second level section header

Second level section text.

Third level section header: Third level section text. These headings may be numbered, but only when the numbers must be cited in the text.

Figures and Tables

Figures and Tables should be labelled and referenced in the standard way using the `\label{}` and `\ref{}` commands.

Sample Figure

Figure 1 shows an example figure.

Sample Video

Figure 2 shows how to include a video in your manuscript.

Sample Table

Table 1 shows an example table. Avoid shading, color type, line drawings, graphics, or other illustrations within tables. Use tables for data only; present drawings, graphics, and illustrations as separate figures. Histograms should not be used to present data that can be captured easily in text or small tables, as they take up much more space.

Tables numbers are given in Arabic numerals. Tables should not be numbered 1A, 1B, etc., but if necessary, interior parts of the table can be labeled A, B, etc. for easy reference in the text.

Sample Equation

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with $\mathrm{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i}^{n} X_i \qquad (1)$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

Literature Cited

Benjamini, Y. and Y. Hochberg, 1995 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) **57**: 289 – 300.

Bleeker, F. E., R. J. Molenaar, and S. Leenstra, 2012 Recent advances in the molecular understanding of glioblastoma **108**: 11–27.

Table 1 Students and their grades

Student	Grade[a]	Rank	Notes
Alice	82%	1	Performed very well.
Bob	65%	3	Not up to his usual standard.
Charlie	73%	2	A good attempt.

[a] This is an example of a footnote in a table. Lowercase, superscript italic letters (a, b, c, etc.) are used by default. You can also use *, **, and *** to indicate conventional levels of statistical significance, explained below the table.

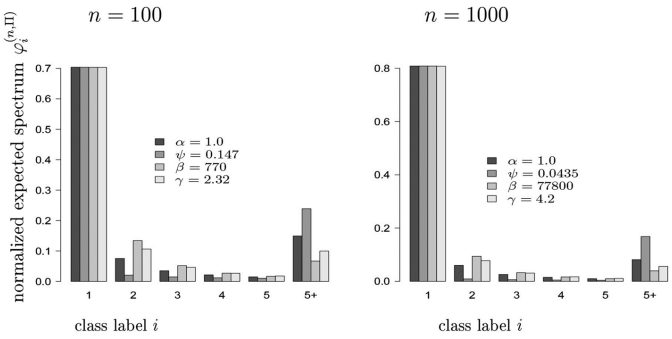

Figure 1 Example figure from 10.1534/genetics.114.173807. Please include your figures in the manuscript for the review process. You can upload figures to Overleaf via the Project menu. Upon acceptance, we'll ask for your figure files to be uploaded in any of the following formats: TIFF (.tiff), JPEG (.jpg), Microsoft PowerPoint (.ppt), EPS (.eps), or Adobe Illustrator (.ai). Images should be a minimum of 300 dpi in resolution and 500 dpi minimum if line art images. RGB, CMYK, and Grayscale are all acceptable. Halftones should be high contrast with sharp detail, because some loss of detail and contrast is inevitable in the production process. Figures should be 10-20 cm in width and 1-25 cm in height. Graph axes must be exactly perpendicular and all lines of equal density. Label multiple figure parts with A, B, etc. in bolded type, and use Arrows and numbers to draw attention to areas you want to highlight. Legends should start with a brief title and should be a self-contained description of the content of the figure that provides enough detail to fully understand the data presented. All conventional symbols used to indicate figure data points are available for typesetting; unconventional symbols should not be used. Italicize all mathematical variables (both in the figure legend and figure) , genotypes, and additional symbols that are normally italicized.

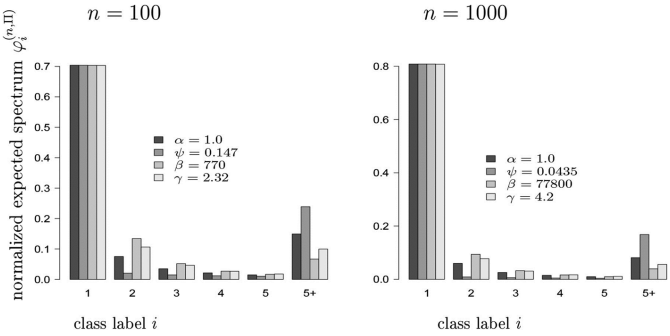

Figure 2 Example movie (the figure file above is used as a placeholder for this example). *GENETICS* supports video and movie files that can be linked from any portion of the article - including the abstract. Acceptable formats include .asf, avi, .wav, and all types of Windows Media files.

Bolstad, B. M., 2004 *Low-level Analysis of High-density Oligonucleotide Array Data: Background, Normalization and Summarization*. Ph.D. thesis.

Brettschneider, J., F. Collin, B. M. Bolstad, and T. P. Speed, 2007 Quality assessment for short oligonucleotide microarray data. Access p. 32.

Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, 2003 Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) **4**: 249–264.

Johnson, D. R., A. M. Sawyer, C. A. Meyers, B. P. O'Neill, and J. S. Wefel, 2012 Early measures of cognitive function predict survival in patients with newly diagnosed glioblastoma. Neuro-Oncology **14**: 808–816.

Murat, A., E. Migliavacca, T. Gorlia, W. L. Lambiv, T. Shay, M. F. Hamou, N. De Tribolet, L. Regli, W. Wick, M. C. M. Kouwenhoven, J. A. Hainfellner, F. L. Heppner, P. Y. Dietrich, Y. Zimmer, J. G. Cairncross, R. C. Janzer, E. Domany, M. Delorenzi, R. Stupp, and M. E. Hegi, 2008 Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. Journal of Clinical Oncology **26**: 3015–3024.

Stupp, R., P. Y. Dietrich, S. O. Kraljevic, A. Pica, I. Maillard, P. Maeder, R. Meuli, R. Janzer, G. Pizzolato, R. Miralbell, F. Porchet, L. Regli, N. de Tribolet, R. O. Mirimanoff, and S. Leyvraz, 2002 Promising survival for patients with newly diagnosed glioblastoma multiforme treated with concomitant radiation plus temozolomide followed by adjuvant temozolomide. Journal of Clinical Oncology **20**: 1375–1382.

Stupp, R., W. P. Mason, M. J. van den Bent, M. Weller, B. Fisher,

M. J. B. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn, J. Curschmann, R. C. Janzer, S. K. Ludwin, T. Gorlia, A. Allgeier, D. Lacombe, J. G. Cairncross, E. Eisenhauer, and R. O. Mirimanoff, 2005 Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. New England Journal of Medicine **352**: 987–996.