

Original Article

Independent validation of a mathematical genomic model for survival of glioma patients

Jason B Nikas

Genomix Inc., Minneapolis, MN 55364, USA

Received May 19, 2016; Accepted May 22, 2016; Epub June 1, 2016; Published June 15, 2016

Abstract: An independent cohort study was conducted to validate a mathematical genomic model for survival of glioma patients that was introduced previously. Of the 102 new subjects that were employed in this study, 40 were long-term survivors (survival ≥ 3 years), and 62 were short-term survivors (survival ≤ 1 year). Utilizing the gene expression of 5 genes as captured by mRNA sequencing of primary tumor tissue, obtained from the initial biopsy during the diagnosis, and prior to the administration of any treatment, the model classified correctly all but three of the 102 subjects. More specifically, of the 62 STS (short-term survivors), 61 were classified correctly (sensitivity = 98.4%); and of the 40 LTS (long-term survivors), 38 were classified correctly (specificity = 95.0%). The 5 gene expression input variables to the model were: *FAM120AOS*, *MXI1*, *OCIAD2*, *PCDH15*, and *PDLIM4*. Of the top 29 most significantly differentially expressed genes between STS and LTS subjects, as identified in the original study, all but one were highly significant. Furthermore, with respect to survival, the model - designed to operate at the molecular level (gene expression of tumor cells) - was also able to statistically differentiate between the two subgroups of the STS group, namely, the STS subjects with lower grade glioma and the STS subjects with glioblastoma; whereas variables either at the tissue level or at the organismal level were not able to do so. Based on these results, and taking into account that accurate clinical prognosis for short-term vs. long-term survival for glioma patients is currently nonexistent, this study provides further, independent evidence for the accuracy and the clinical utility of the model.

Keywords: Glioma, survival, cancer genomics, computational biology, *FAM120AOS*, *MXI1*, *OCIAD2*, *PCDH15*, *PDLIM4*

Introduction

In a previous study [1], I introduced a mathematical genomic model that can identify with a high accuracy those patients with glioma who will experience short-term survival (≤ 1 year), as well as those with long-term survival (≥ 3 years), at the time of diagnosis and prior to surgery and adjuvant chemotherapy. The normalized RNA-Seq gene expression values of the following 5 genes, *FAM120AOS*, *MXI1*, *OCIAD2*, *PCDH15*, and *PDLIM4*, constitute the 5 input variables of the model. Also in the previous study [1], 29 genes were identified as the top most significantly differentially expressed between the two groups, i.e., the STS (short-term survivors) and the LTS (long-term survivors); and that list of 29 genes included the aforementioned 5 input genes to the model. In the original study, 89 subjects (14 STS and 75 LTS) were used for the development and subse-

quent validation of the model. Based on those 89 subjects, the performance of the model was as follows: sensitivity = $(13/14) = 92.9\%$ and specificity = $(72/75) = 96.0\%$. Here in this study, an independent cohort, 102 new subjects (62 STS and 40 LTS) were used for the second, independent validation of the model.

Furthermore, in this study, in order to test and extend the domain of the model, and in order to test its theoretical foundation, the STS group was expanded to include STS subjects with grade IV glioma (glioblastoma multiforme or GBM). All of the STS subjects with GBM were classified correctly by the model. Upon further investigation, there was a statistically significant difference between the STS subjects with LGG (lower-grade glioma) and the STS subjects with GBM. More specifically, according to individual subject scores as calculated by the algorithm of the model, the STS-GBM group had

Independent validation of a genomic model for survival of glioma patients

significantly higher scores than the STS-LGG group. That means that in comparison with the STS-LGG subjects, the STS-GBM subjects were significantly farther away from the cut-off point, which signifies the boundary between short-term and long-term survival. This new experimental evidence lends further support to the theoretical foundation of the model. The model was designed to detect a certain genomic pattern in the tumor cells that represents the early stages of a progressive genomic development, which ultimately leads to patient short-term survival. Both the results of this study in connection with the second and independent validation of the model and the further, independent evidence of its theoretical foundation provide additional support for its clinical utility.

Methods

Data acquisition

The normalized data for 102 (62 STS and 40 LTS) subjects with glioma [G2, G3, or G4 (grade 2, 3, or 4)] generated from mRNA sequencing of primary tumor tissue (using the Illumina HiSeq 2000 sequencer) were downloaded from The Cancer Genome Atlas (TCGA) of the National Cancer Institute under the category LGG and GBM (accessed on 2016-04-09).

Clinical

All STS subjects had a survival ≤ 1 year, and all LTS subjects had a survival ≥ 3 years. All subjects selected for this study had G2, G3, or G4 glioma; all had primary (de novo) glioma of one of the following types: astrocytoma, oligodendroglioma, oligoastrocytoma, or glioblastoma multiforme; and all had surgery, adjuvant chemotherapy (mainly temozolomide), and/or radiation.

None of the subjects selected for this study had a prior diagnosis of brain cancer, and none of them received any neoadjuvant treatment. All tumor tissue samples used in this study were collected via a biopsy that preceded any treatment.

[Table S1](#) contains clinical and demographical information about all 102 subjects employed in this study.

Statistical

General: All analyses performed in this study, including those of, and in connection with, the

algorithm of the model, were based on a bioinformatic methodology that I have developed and presented previously [2-11].

Differential F_1 score analysis: The significance level was set at $\alpha = 0.01$ (two-tailed) to account for the following four comparisons in the differential F_1 score analysis in this study. a) LTS vs. STS subjects. As can be calculated from the F_1 scores of all 102 subjects used in this study ([Table S1](#)), the F_1 scores were normally distributed with respect to both groups (40 LTS vs. 62 STS), but the condition of equality of variance with respect to both groups was not met. Therefore, the Aspin-Welch unequal-variance t-test for normal distributions was used to calculate the probability of significance. b) $LTS_{1,2}$ vs. $STS_{1,2}$. In this analysis, the 89 subjects of the original study [1] (75 LTS and 14 STS) ([Table S7](#) of the original study) were pooled together with the 102 subjects used in this study (40 LTS and 62 STS) ([Table S1](#)). In this case, the F_1 scores of the combined total of 191 subjects were not normally distributed with respect to both groups (115 LTS vs. 76 STS). Therefore, the Mann-Whitney U test for non-parametric distributions was used to calculate the probability of significance, and the approximate probability level with correction was reported. c) STS-LGG vs. STS-GBM. This analysis was performed using all 62 STS subjects of this study. The F_1 scores of the two subgroups of the STS group, namely, STS-LGG and STS-GBM, were parametrically distributed with respect to both of those groups (both the normality and the equality of variance conditions were met) ([Table S1](#)). Therefore, the equal-variance independent t-test was used to calculate the probability of significance. d) $STS-LGG_{1,2}$ vs. $STS-GBM_{1,2}$. In this analysis, all STS subjects from the original study [1], all of which were STS-LGG [subjects # 76-89 ([Table S7](#) of the original study)], were pooled together with all STS subjects of this study. The F_1 scores of the two subgroups of the combined STS group from both studies, namely, $STS-LGG_{1,2}$ and $STS-GBM_{1,2}$, were parametrically distributed with respect to both of those groups (both the normality and the equality of variance conditions were met) ([Table S1](#) and [Table S7](#) of the original study). Therefore, the equal-variance independent t-test was used to calculate the probability of significance.

Differential gene expression analysis: The P-value for each one of the gene variables was calculated using the appropriate test. The

Independent validation of a genomic model for survival of glioma patients

Table 1. Gene expression results of the top 29 genes identified in the original study

No	Gene Name	NCBI Gene ID	DE (STS)	ROC AUC	FC	P	M _L	SD _L	M _S	SD _S
1	ABI1	10006	↓	-0.9427	-1.882	9.09E-18	2452.22	859.41	1302.95	349.07
2	ADO	84890	↓	-0.9048	-1.512	3.40E-14	1140.74	249.96	754.21	211.37
3	AP1S3	130340	↑	0.9306	2.845	1.64E-16	19.10	11.89	54.34	29.52
4	ARNTL2	56938	↑	0.9613	4.264	5.24E-20	24.89	19.77	106.15	47.49
5	ASCC1	51008	↓	-0.9133	-1.580	6.57E-15	964.10	212.08	610.32	185.94
6	CMYA5	202333	↑	0.8827	4.712	1.69E-12	75.46	99.50	355.53	489.87
7	CTBP2	1488	↓	-0.9581	-1.788	1.39E-19	1344.39	293.69	752.00	140.63
8	DIAPH1	1729	↑	0.9165	1.666	3.42E-15	972.26	232.33	1619.89	453.99
9	EIF4EBP2	1979	↓	-0.9480	-1.614	9.05E-16	3649.52	719.00	2260.90	472.38
10	EMP3	2014	↑	0.9996	29.801	1.07E-28	79.11	87.93	2357.65	1452.41
11	ETV7	51513	↑	0.8940	4.613	2.47E-13	11.83	29.23	54.56	61.37
12	FABP5	2171	↑	0.9246	11.846	6.28E-16	23.22	24.82	275.11	304.44
13	FAM120AOS	158293	↑	0.9375	1.676	3.19E-17	608.66	109.96	1019.94	274.71
14	FBXO17	115290	↑	0.9843	7.834	9.51E-24	74.10	69.15	580.50	204.58
15	GJD3	125111	↑	0.9496	10.265	1.52E-18	10.92	11.38	112.08	161.46
16	LOC254559	254559	↓	-0.9754	-6.188	4.14E-22	3141.36	1399.85	507.68	570.74
17	MAP1LC3C	440738	↑	0.9772	21.471	1.87E-22	2.11	2.80	45.34	53.98
18	MARCH5	54708	↓	-0.9222	-1.456	1.06E-15	1049.88	171.21	721.06	153.89
19	MRPL43	84545	↓	-0.8411	-1.366	7.76E-10	1199.77	229.95	878.01	358.98
20	MXI1	4601	↓	-0.9742	-2.530	8.17E-17	2769.10	808.09	1094.39	411.89
21	OCIAD2	132299	↑	0.9649	16.288	1.66E-20	85.37	102.82	1390.48	1458.48
22	PCDH15	65217	↓	-0.9827	-12.035	2.01E-23	1676.88	920.92	139.33	220.01
23	PDLIM4	8572	↑	0.9411	16.783	1.36E-17	77.65	91.82	1303.13	1344.10
24	RAP2A	5911	↓	-0.9722	-2.901	1.38E-21	6042.71	3191.48	2083.20	864.91
25	RBM17	84991	↓	-0.9661	-1.919	1.12E-20	2194.57	516.48	1143.68	322.71
26	SEPHS1	22929	↓	-0.8738	-1.528	7.05E-12	1464.10	407.75	958.03	317.63
27	SLC12A7	10723	↑	0.9044	2.280	3.66E-14	392.13	188.62	893.88	365.23
28	SLC27A3	11000	↑	0.9694	3.966	3.76E-21	189.65	92.57	752.09	859.46
29	TMPRSS3	64699	↑	0.7649	7.528	3.45E-06	6.46	13.21	48.66	84.63
*	TBP	6908	↓	-0.6181	-1.090	4.46E-02	328.23	61.06	301.16	95.87

Gene expression results from this study of the top 29 genes identified in the original study in alphabetical order. The arrows indicate differential expression [over-expression (↑) or under-expression (↓)] of the STS subjects as compared with the LTS subjects. The ROC AUC value, the fold change (FC) value, the P-value, the mean expression value of the LTS subjects (M_L) and their standard deviation (SD_L), the mean expression value of the STS subjects (M_S) and their standard deviation (SD_S) are listed for each gene variable. (*) TBP, a natural control gene, is also listed for comparison purposes.

equal-variance independent t-test (TT) was used for parametric gene variables (both normality and equality of variance conditions were met); the Aspin-Welch unequal-variance t-test (AW) was used for those gene variables that met the normality condition but not the equality of variance condition; and the Mann-Whitney U test (MW) was used for the non-parametric gene variables. The Anderson-Darling test was used to assess normality, and the Levene absolute test for equal variances was used to assess equality of variance throughout this study. The

initial significance level was set at $\alpha_0 = 0.05$ (two-tailed). Since, however, there are 20,531 gene variables (these are all the exome genes that were quantified by the mRNA sequencing of all tumor tissue samples using the Illumina HiSeq 2000 sequencer), the significance level was adjusted using the Bonferroni correction to: $\alpha_B = 2.43 \times 10^{-6}$ (two-tailed). Therefore, in order for any variable to be deemed statistically significant, the following condition must be met: $P < 2.43 \times 10^{-6}$. It should be pointed out here that since this study was designed to be

Independent validation of a genomic model for survival of glioma patients

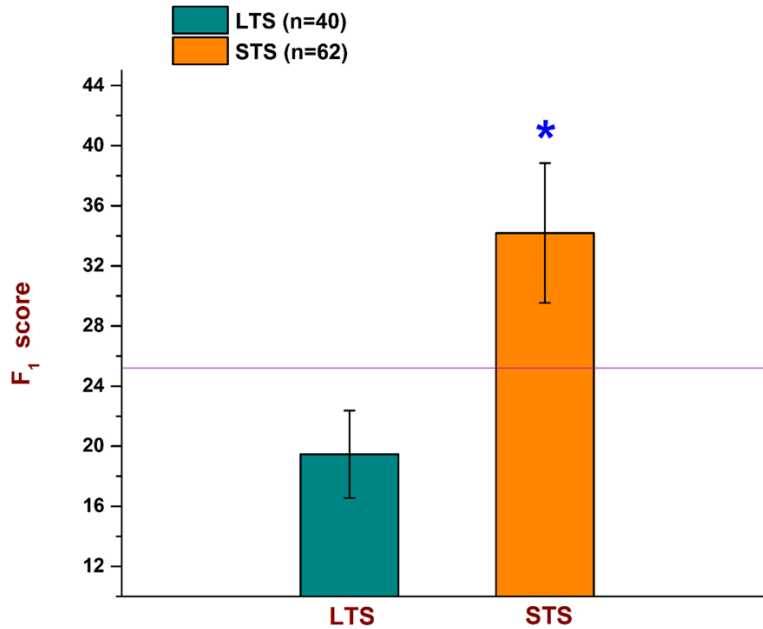


Figure 1. Model performance in the present study. The model classifies a subject as a long-term survivor (LTS) if the F_1 score is < 25.2 or as a short-term survivor (STS) if the F_1 score is ≥ 25.2 . The cut-off score of 25.2 is represented here by the horizontal purple line. In this study, a total of 102 subjects were employed (40 LTS and 62 STS). The mean F_1 score of the LTS subjects was 19.463 (top of the green bar) and their standard deviation (whiskers above or below the top of the green bar) was 2.906. The 95% confidence interval of the mean F_1 score of the LTS subjects was: [18.534, 20.393]. The mean F_1 score of the STS subjects was 34.186 (top of the orange bar) and their standard deviation (whiskers above or below the top of the orange bar) was 4.651. The 95% confidence interval of the mean F_1 score of the STS subjects was: [33.005, 35.368]. The significance level was set at $\alpha = 0.01$ (two-tailed), and the probability of significance for the F_1 was $P = 3.83 \times 10^{-36}$ (Aspen-Welch unequal-variance t-test with a t-statistic = -19.674 and $df = 99.93$). The F_1 scores of all 102 subjects are listed in [Table S1](#).

an independent validation study of the original one, only those 30 gene variables that are listed in [Table 1](#) were examined. Therefore, a much higher threshold of significance ($P < 1.67 \times 10^{-3}$) is warranted. Nevertheless, the far more stringent threshold of significance, namely, $P < 2.43 \times 10^{-6}$ was imposed for the differential gene expression analysis in this study. ROC curve analysis was performed on all 29 gene variables with respect to the two groups (40 LTS vs. 62 STS) in order to calculate their respective ROC AUC value. For all 29 gene variables, the fold change was calculated as the mean expression value of the STS subjects over the mean expression value of the LTS subjects of a particular gene variable.

Cox proportional hazards regression analyses: Two Cox proportional hazards regression analyses were performed: one in connection with all of the STS subjects of this study (STS-LGG vs.

STS-GBM) ([Table S2](#)) and one in connection with all of the STS subjects combined from both studies (STS-LGG_{1,2} vs. STS-GBM_{1,2}) ([Table S3](#)). Each one of those two analyses was performed twice: the first time as Model A, wherein, in addition to the survival and censoring variables, the Group variable (STS-LGG vs. STS-GBM), which represents the tumor histological classification (lower-grade glioma vs. glioblastoma) of the STS subjects, was the only independent variable; and the second time as Model B, wherein, in addition to the survival and censoring variables, five independent variables were used [Group (STS-LGG vs. STS-GBM), Gender, Age, Tumor Histological Type, and Tumor Histological Grade]. The independent variables were inputted as follows: Group (Tumor Histological Classification): STS-LGG = 0 and STS-GBM = 1; Gender: female = 1 and male = 2; Age: as a numerical variable; Tumor Histological

Type: astrocytoma = 1, glioblastoma = 2, oligoastrocytoma = 3, oligodendroglioma = 4; and Tumor Histological Grade: as 2, 3, or 4 corresponding to tumor grades 2, 3, or 4, respectively. The survival cut-off point was set to 1 year since all STS subjects experienced a survival ≤ 1 year. The survival time ([Table S1](#) and [Table S1](#) of the original study) was used as the time-to-event variable. Regarding the censoring variable, since all of the STS subjects experienced a survival ≤ 1 year, i.e. within the survival cut-off point, none of the subjects was censored. The significance level for all Cox analyses was set at $\alpha = 0.05$ (two-tailed).

Computer software

All analyses in this study were carried out with custom software written by JBN in MATLAB R2016a.

Independent validation of a genomic model for survival of glioma patients

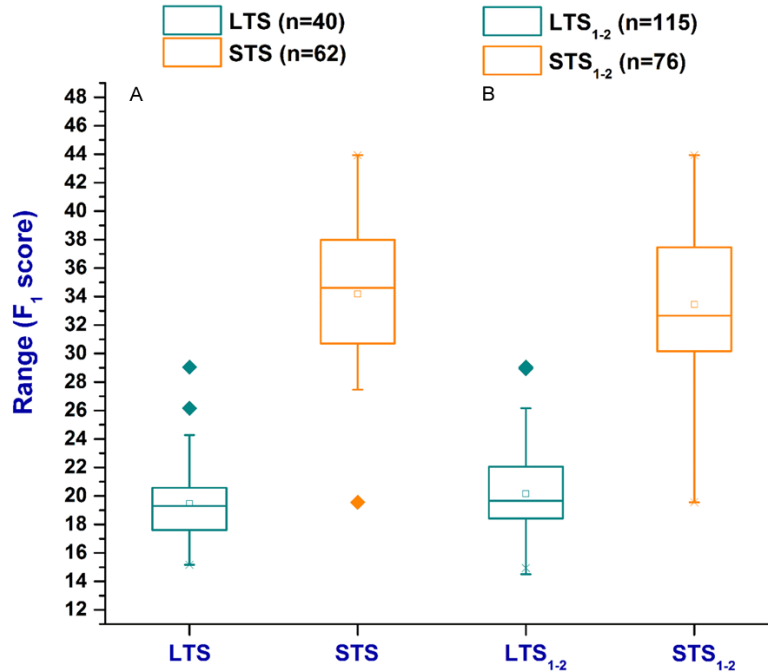


Figure 2. Model performance (present study and overall performance). (A) Box plots of the model performance in the present study. In this study, a total of 102 subjects were employed (40 LTS and 62 STS). The mean F_1 score of the LTS subjects was 19.463, the median was 19.294, and the range was 13.873 [15.175, 29.048]. There were two statistical outliers, represented here by the two green diamonds. The mean F_1 score of the STS subjects was 34.186, the median was 34.606, and the range was 24.370 [19.560, 43.930]. There was one statistical outlier, represented here by the orange diamond. The F_1 scores of all 102 subjects are listed in [Table S1](#). (B) Box plots of the overall model performance. In order to assess the overall performance of the model, thus far, the 89 subjects (75 LTS and 14 STS) used in the original study ([Table S7](#) of the original study) were pooled together with the 102 subjects (40 LTS and 62 STS) used in this study ([Table S1](#)). The mean F_1 score of the 115 LTS_{1-2} subjects was 20.147, the median was 19.653, and the range was 14.547 [14.501, 29.048]. There was one statistical outlier, represented here by the green diamond. The mean F_1 score of the 76 STS_{1-2} subjects was 33.444, the median was 32.666, and the range was 24.370 [19.560, 43.930].

Results

Model performance

In the original study [1], the F_1 algorithm of the model was developed and presented (Equation 1); and the cut-off score of 25.2 was calculated such that if the score of a particular subject was < 25.2 , the subject would be classified as LTS, or if the score was ≥ 25.2 , the subject would be classified as STS.

$$F_1 = \left[\ln \left[\frac{(X_1)^2 + (X_2)^2 + (X_3)^2}{(X_1) \cdot (X_2) \cdot (X_3)} \right] \right]^{\frac{1}{2}} \cdot (10) \quad (\text{Equation 1})$$

In Equation 1 above, $X_1 = FAM120AOS$, $X_2 = PDLIM4$, $X_3 = OCIAD2$, $X_4 = PCDH15$, and $X_5 =$

$MXI1$. The X_1 - X_5 are the normalized RNA-Seq gene expression values of the aforementioned 5 genes. In this second, independent cohort study, utilizing the F_1 algorithm (Equation 1); employing the normalized RNA-Seq gene expression values of the above 5 genes as input variables; and using the same cut-off score of 25.2; the model classified correctly all but three of the 102 new subjects used (40 LTS and 62 STS). More specifically, of the 62 STS subjects, all but one were classified correctly [sensitivity = $(61/62) = 98.4\%$]; and of the 40 LTS subjects, all but two were classified correctly [specificity = $(38/40) = 95.0\%$]. [Table S1](#) lists the F_1 scores of all 102 subjects. Statistical analysis of the scores of all 102 subjects revealed a large significant difference between the two groups (LTS vs. STS) [$P = 3.83 \times 10^{-36}$ (Aspen-Welch unequal-variance t-test with a t-statistic = -19.674 and $df = 99.93$)]. Specifically, the mean F_1 score of the LTS group was 19.463 with a 95% confidence interval of [18.534, 20.393] and SD =

2.906; whereas the mean F_1 score of the STS group was 34.186 with a 95% confidence interval of [33.005, 35.368] and SD = 4.651.

Figures 1 & 2A depict the results of the aforementioned statistical analysis, whereas **Figure 3A** provides a 3D space position of all 102 subjects according to their F_1 scores and shows a clear separation between the two groups (LTS vs. STS).

It should be pointed out here that those statistical values of the two groups were very close to the corresponding values of the two groups in the original study. More specifically, the original study utilized 89 different subjects, and the mean F_1 score of the LTS group was 20.511 (SD

Independent validation of a genomic model for survival of glioma patients

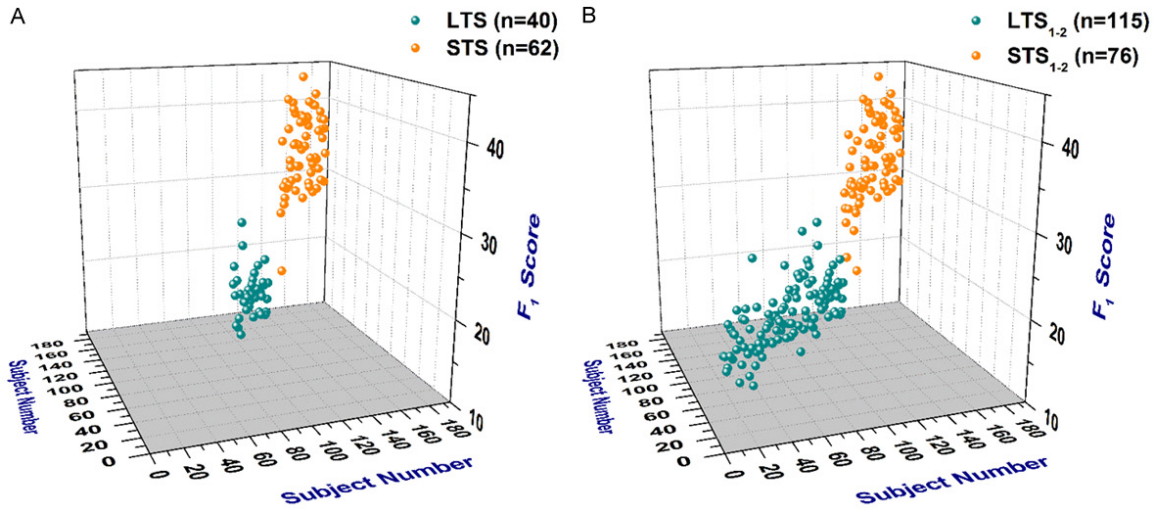


Figure 3. 3D space position of all subjects according to their F_1 scores. (A) 3D space position of all 102 subjects employed in the present study according to their F_1 scores. The F_1 scores of all 102 subjects (40 LTS and 62 STS) are plotted in the z-axis. The subject number is plotted in the x-axis and the y-axis. A plane parallel to the x-y plane that intersects the z-axis at the point 25.2, which is the cut-off score, represents the cut-off plane. Subjects that are classified as STS lie above the cut-off plane, whereas subjects that are classified as LTS lie below the cut-off plane. The two groups are clearly separated. (B) 3D space position of all 191 subjects employed in both studies according to their F_1 scores. The 89 subjects (75 LTS and 14 STS) used in the original study (Table S7 of the original study) were pooled together with the 102 subjects (40 LTS and 62 STS) used in this study (Table S1). The F_1 scores of all 191 subjects (115 LTS_{1-2} and 76 STS_{1-2}) are plotted in the z-axis. The subject number is plotted in the x-axis and the y-axis. A plane parallel to the x-y plane that intersects the z-axis at the point 25.2, which is the cut-off score, represents the cut-off plane. Subjects that are classified as STS lie above the cut-off plane, whereas subjects that are classified as LTS lie below the cut-off plane. The two groups are clearly separated. In order to provide the same perspective as the one in (A), all 102 subjects of this study were assigned the same subject numbers as in (A), and all three axes were scaled the same way as in (A).

= 2.790), while the mean F_1 score of the STS group was 30.157 (SD = 4.068).

In order to assess the overall performance of the model, thus far, the 89 subjects (75 LTS and 14 STS) used in the original study (Table S7 of the original study) were pooled together with the 102 subjects (40 LTS and 62 STS) used in this study (Table S1). Out of $89+102 = 191$ subjects (115 LTS and 76 STS), the model classified correctly all but seven subjects (four in the original study and three in the present one) (Table S1 and Table S7 of the original study). More specifically, out of a total of 76 STS subjects, the model classified correctly all of them except two [overall sensitivity = $(74/76) = 97.4\%$]; and out of a total of 115 LTS subjects, the model classified correctly all of them except five [overall specificity = $(110/115) = 95.7\%$]. Just as was the case in the present study, statistical analysis of the F_1 scores of the combined total of 191 subjects revealed a large significant difference between the two groups (115 LTS vs. 76 STS) [$P = 1.29 \times 10^{-44}$ (Mann-Whitney U test with a z-value = 11.358)].

Specifically, the mean F_1 score of the LTS group was 20.147 with a 95% confidence interval of [19.618, 20.676] and SD = 2.863; whereas the mean F_1 score of the STS group was 33.444 with a 95% confidence interval of [32.350, 34.539] and SD = 4.789. **Figure 2B** depicts the results of the aforementioned statistical analysis; **Figure 3B** provides a 3D space position of all 191 subjects according to their F_1 scores and with respect to the two groups (LTS vs. STS) and shows a clear separation between them; whereas **Figure 4** provides the same 3D space position of all 191 subjects but with respect also to the two STS subgroups ($STS_{LGG_{1-2}}$ and $STS_{GBM_{1-2}}$) and shows also a distinct separation between those two subgroups.

It is both interesting and important to note here that as the subject sample size increased considerably [either from 89 (original study) or from 102 (present study) to 191 (combined total from both studies)], the statistically significant difference between the LTS and the STS subjects also increased considerably.

Independent validation of a genomic model for survival of glioma patients

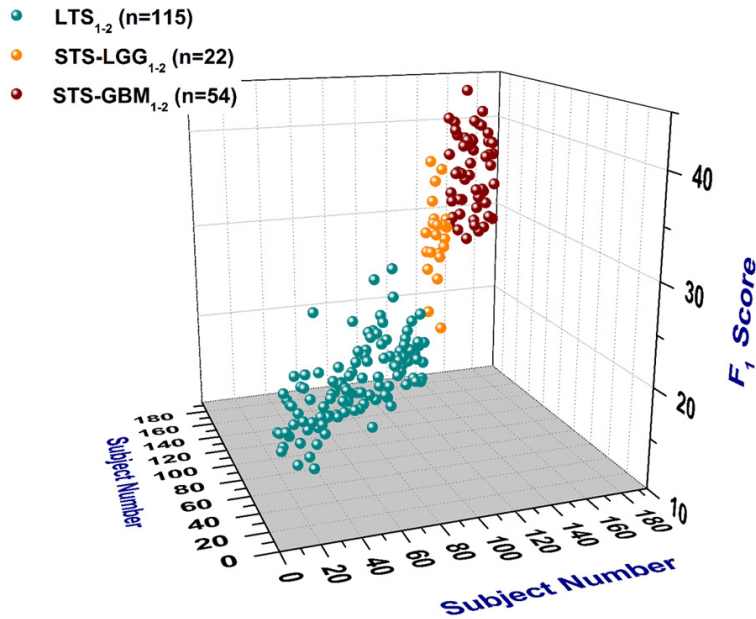


Figure 4. 3D space position of all 191 subjects used in both studies according to their F_1 scores. The F_1 scores of all 191 subjects (115 LTS_{1-2} and 76 STS_{1-2}) from both studies are plotted in the z-axis. The subject number is plotted in the x-axis and the y-axis. A plane parallel to the x-y plane that intersects the z-axis at the point 25.2, which is the cut-off score, represents the cut-off plane. In this 3D scatter plot, the space position of the 115 LTS_{1-2} subjects (green spheres) vs. the 76 STS_{1-2} subjects (orange & dark red spheres) is depicted. There is a clear separation between the LTS_{1-2} subjects and the STS_{1-2} subjects; the former lie below the cut-off plane, whereas the latter lie above it. Furthermore, the two STS_{1-2} subgroups, namely, the 22 $STS-LGG_{1-2}$ [short-term survivors with lower-grade glioma (orange spheres)] vs. the 54 $STS-GBM_{1-2}$ [short-term survivors with glioblastoma (dark red spheres)] are depicted. There is a distinct separation between those two subgroups. The $STS-GBM_{1-2}$ subjects have significantly greater F_1 scores than the $STS-LGG_{1-2}$ subjects do; the former lie much higher above the cut-off plane than the latter do.

Top 29 most significant genes

In the original study [1], the top 29 genes that were the most significant in terms of differential expression between the LTS and STS groups were identified and reported (Table 1 of the original study). Here, using the mRNA-Seq expression of those 29 genes for all 102 subjects, and employing the appropriate tests, as reported in the Methods section, the statistical significance in connection with differential expression between the two groups (LTS vs. STS) was also assessed for each of those 29 genes. All but one of those 29 genes were highly significant according to the very stringent criterion of significance imposed in this study ($P < 2.43 \times 10^{-6}$). It is worth mentioning here, however, that even according to this very stringent criterion of significance, the *TMPRSS3* gene (# 29 in

Table 1) barely missed statistical significance ($P = 3.45 \times 10^{-6}$). **Figure 5** shows the heat map resulted by plotting the expression of the 29 genes for all 102 subjects (40 LTS and 62 STS) and provides a distinct visual separation between the two groups. **Table 1** lists statistical details about the differential expression of those 29 genes between the two groups (LTS vs. STS).

A comparison between the statistical values of the 29 genes in the original study (Table 1 of the original study) with the corresponding statistical values of those 29 genes in this study (**Table 1**) reveals a substantial increase in statistical significance for all of those 29 genes except one (# 29 *TMPRSS3*). It is also important to note here that the 5 genes that are the input variables of the algorithm of the model, namely, *FAM12-OAOS*, *MXI1*, *OCIAD2*, *PCDH15*, and *PDLIM4*, demonstrated increases in the magnitude of statistical significance that were among the highest observed (**Table 1**).

Short-term survival in connection with lower-grade glioma vs. glioblastoma

In this study, in order to be able to test the domain of the model, i.e., whether the model can identify correctly STS subjects (short-term survivors with survival ≤ 1 year) with grade IV glioma (glioblastoma), the STS group was expanded to include 54 STS-GBM subjects. In total, the STS group comprised 62 subjects [8 STS-LGG (short-term survivors with lower-grade glioma) and 54 STS-GBM (short-term survivors with glioblastoma)]. Of the 8 STS-LGG subjects, the model classified correctly all of them except one as STS subjects; and of the 54 STS-GBM subjects, the model classified correctly all of them as STS subjects (**Table S1**).

STS-LGG vs. STS-GBM: Statistical analysis of the F_1 scores of all 62 STS subjects used in this

Independent validation of a genomic model for survival of glioma patients

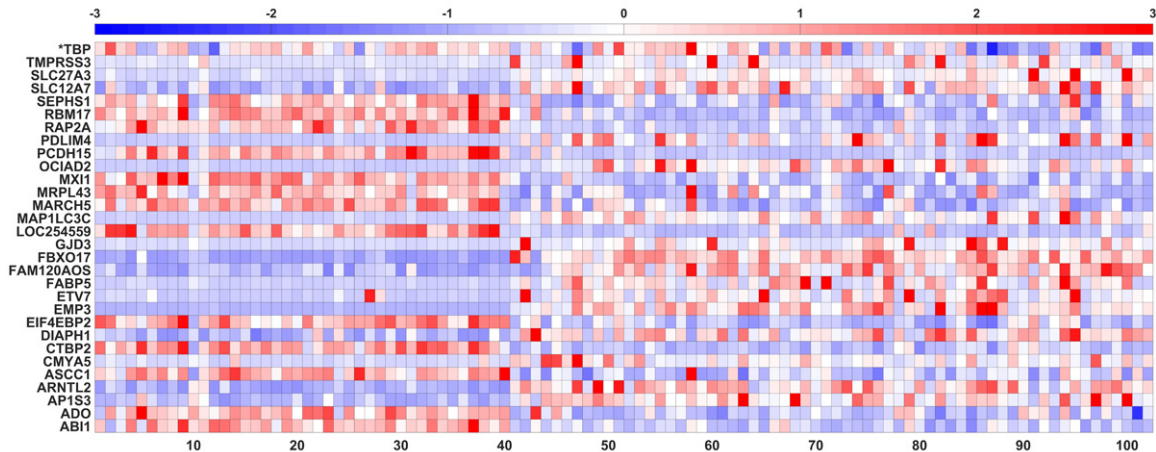


Figure 5. Heat map of the expression of the 29 most significant tumor tissue genes. Heat map of the tumor tissue gene expression, generated from mRNA sequencing, of the 40 LTS subjects (columns # 1-40) (x-axis) and the 62 STS subjects (columns # 41-102) (x-axis) employed in this study with respect to the 29 most significant genes (rows # 1-29) (y-axis), identified in the original study. The order of those 29 genes is alphabetical (the same as the one in [Table 1](#)). The *TBP* gene (*), which is a natural control gene, was also included for comparison purposes. All 30 gene variables were standardized (mean = 0 and SD = 1). The intensity scale of the standardized expression values represents, therefore, the z scores; and it ranges from -3.0 [blue: low expression (3 SD below the mean)] to +3.0 [red: high expression (3 SD above the mean)], with 0 [white (mean = 0)] representing the reference intensity value (mean expression value of all 102 subjects). As can be seen, based on the expression of those 29 most significant genes, there is a distinct overall separation between the LTS and the STS subjects.

study revealed a significant difference between the two STS subgroups (8 STS-LGG vs. 54 STS-GBM) [$P = 6.88 \times 10^{-4}$ (equal-variance independent t-test with a t-statistic = -3.580 and $df = 60$)]. Specifically, the mean F_1 score of the STS-LGG group was 29.158 with a 95% confidence interval of [25.172, 33.143] and SD = 4.767; whereas the mean F_1 score of the STS-GBM group was 34.931 with a 95% confidence interval of [33.789, 36.074] and SD = 4.185. According to the model, therefore, and from the molecular perspective, in comparison with the STS-LGG subjects, the STS-GBM subjects had significantly higher scores and were significantly farther away from the cut-off point of 25.2, which demarcates long-term survival (< 25.2) from short-term survival (≥ 25.2). **Figure 6A** depicts the results of the aforementioned statistical analysis.

Two Cox proportional hazards regression analyses were performed with respect to the two STS subgroups (STS-LGG vs. STS-GBM). The first analysis (Model A) revealed that when examined all by itself, the tumor histological classification (Group) [LGG (lower-grade glioma) vs. GBM (glioblastoma)] had no statistically significant effect on the survival of the STS sub-

jects [Model A: Group variable ($P = 0.2446$)]. To state it differently and equivalently, with respect to survival, the STS subjects with LGG could not be statistically differentiated from the STS subjects with GBM. The second Cox proportional hazards regression analysis (Model B) revealed that when examined together with the gender, age, tumor histological type, and tumor histological grade variables, the tumor histological classification (Group variable) had no statistically significant effect on the survival of the STS subjects either [Model B: Group variable ($P = 0.3475$)]. The second Cox analysis also revealed that gender, age, tumor histological type (astrocytoma, glioblastoma, oligoastrocytoma, or oligodendroglioma), and tumor histological grade - as possible covariates - had no statistically significant effect on the survival of the STS subjects either [Model B: Gender ($P = 0.1546$), Age ($P = 0.2422$), TH Type ($P = 0.7689$), and TH Grade ($P = 0.2621$)]. The log-likelihood of the Model A was $\text{LogL} = -196.951$, whereas the log-likelihood of the Model B was $\text{LogL} = -194.152$ ($P = 0.2312$), indicating that the Model B did not constitute a statistically significant improvement over the Model A. [Table S1](#) contains all of the data used in those analyses, whereas [Table S2](#) lists all of their respective results.

Independent validation of a genomic model for survival of glioma patients

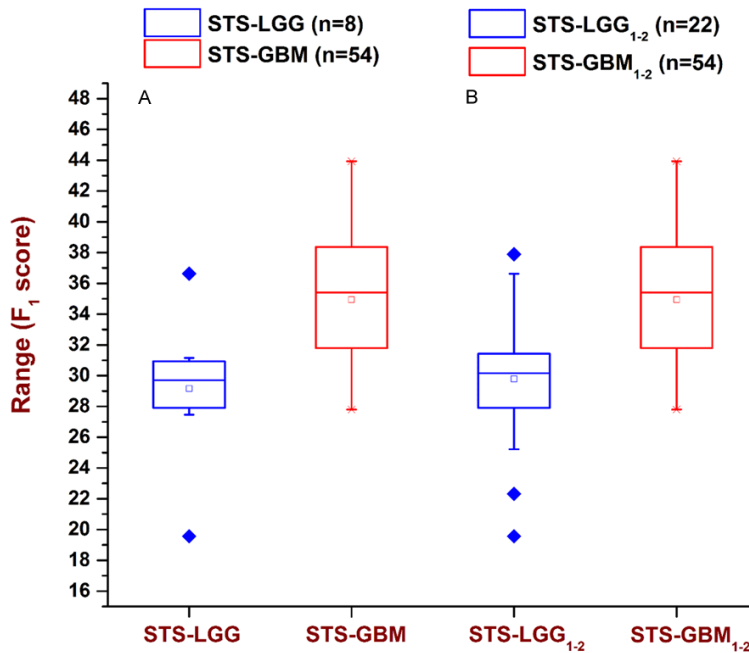


Figure 6. Short-term survivors with lower-grade glioma (STS-LGG) vs. short-term survivors with glioblastoma (STS-GBM). (A) The two box plots depict the results of the analysis between the two STS subgroups [8 STS-LGG subjects (blue) vs. 54 STS-GBM subjects (red)] employed in the present study. The mean F_1 score of the STS-LGG subjects was 29.158, the median was 29.701, and the range was 17.065 [19.560, 36.625]. There were two statistical outliers, represented here by the two blue diamonds. The mean F_1 score of the STS-GBM subjects was 34.931, the median was 35.412, and the range was 16.134 [27.796, 43.930]. The F_1 scores of all 102 subjects used in this study are listed in [Table S1](#). (B) The two box plots depict the results of the analysis between the two STS subgroups [22 STS-LGG_{1,2} subjects (blue) vs. 54 STS-GBM_{1,2} subjects (red)] employed in both the present study and the original one. The mean F_1 score of the STS-LGG_{1,2} subjects was 29.794, the median was 30.163, and the range was 18.336 [19.560, 37.896]. There were three statistical outliers, represented here by the three blue diamonds. The mean F_1 score of the STS-GBM_{1,2} subjects was 34.931, the median was 35.412, and the range was 16.134 [27.796, 43.930].

According to the aforementioned Cox proportional hazards regression analyses, with respect to survival, neither variables at the tissue level [tumor histological classification (LGG vs. GBM), tumor histological type (astrocytoma, glioblastoma, oligoastrocytoma, or oligodendroglioma), or tumor histological grade (glioma grade II, III, or IV)] nor variables at the organismal level (gender or age) were able to statistically discriminate between the two STS subgroups, i.e., STS-LGG vs. STS-GBM. The model, on the other hand, designed to operate at the molecular level (gene expression of tumor cells), was able to statistically discriminate between the two STS subgroups and showed that, with respect to survival, the STS-GBM sub-

jects were significantly worse than the STS-LGG subjects were, that is to say that the former were significantly farther away from the long-term survival area (significantly farther away from the cut-off of 25.2) than the latter were.

STS-LGG_{1,2} vs. STS-GBM_{1,2}: In order to ascertain whether the previous analyses and results about the two STS subgroups (STS-LGG vs. STS-GBM) were skewed in any way by the relative small sample size of one of the two subgroups (8 STS-LGG vs. 54 STS-GBM), all of the STS subjects of the original study (14 STS-LGG) (Table S1 of the original study) were pooled together with all of the STS subjects of this study. The combined STS group, therefore, comprised 22 = 8+14 STS-LGG_{1,2} and 54 STS-GBM_{1,2} subjects. Statistical analysis of the F_1 scores of the combined total of 76 STS subjects (used in both the original and the present studies) revealed a significant difference between the two STS subgroups (22 STS-LGG_{1,2} vs. 54 STS-GBM_{1,2}) [$P = 7.12 \times 10^{-6}$ (equal-variance independent t-test

with a t-statistic = -4.832 and df = 74)]. Specifically, the mean F_1 score of the STS-LGG_{1,2} group was 29.794 with a 95% confidence interval of [27.910, 31.678] and SD = 4.250; whereas the mean F_1 score of the STS-GBM_{1,2} group was 34.931 with a 95% confidence interval of [33.789, 36.074] and SD = 4.185. According to the model, therefore, and from the molecular perspective, in comparison with the STS-LGG_{1,2} subjects, the STS-GBM_{1,2} subjects had significantly higher scores and were significantly farther away from the cut-off point of 25.2, which demarcates long-term survival (< 25.2) from short-term survival (≥ 25.2). **Figure 6B** depicts the results of the aforementioned statistical analysis.

Independent validation of a genomic model for survival of glioma patients

Two Cox proportional hazards regression analyses were performed with respect to the two STS subgroups (22 STS-LGG_{1,2} vs. 54 STS-GBM_{1,2}), using all STS subjects from both the original and the present studies. The first analysis (Model A) revealed that when examined all by itself, the tumor histological classification (Group) [LGG (lower-grade glioma) vs. GBM (glioblastoma)] had no statistically significant effect on the survival of the STS subjects [Model A: Group variable ($P = 0.1387$)]. To state it differently and equivalently, with respect to survival, the STS-LGG_{1,2} subjects could not be statistically differentiated from the STS-GBM_{1,2} subjects. The second Cox proportional hazards regression analysis (Model B) revealed that when examined together with the gender, age, tumor histological type, and tumor histological grade variables, the tumor histological classification (Group variable) had no statistically significant effect on the survival of the STS subjects either [Model B: Group variable ($P = 0.3875$)]. The second Cox analysis also revealed that gender, age, tumor histological type (astrocytoma, glioblastoma, oligoastrocytoma, or oligodendroglioma), and tumor histological grade - as possible covariates - had no statistically significant effect on the survival of the STS subjects either [Model B: Gender ($P = 0.2417$), Age ($P = 0.2835$), TH Type ($P = 0.4171$), and TH Grade ($P = 0.6419$)]. The log-likelihood of the Model A was $\text{LogL} = -255.690$, whereas the log-likelihood of the Model B was $\text{LogL} = 254.147$ ($P = 0.5435$), indicating that the Model B did not constitute a statistically significant improvement over the Model A. [Table S1](#) contains all of the data used in those analyses, whereas [Table S3](#) lists all of their respective results.

It is interesting to point out here the following two observations. 1) By combining together all STS subjects from both studies (the original and the present one), the sample size of the STS-LGG subgroup increased from 8 to 22 (nearly tripled). 2) Given that large sample size increase, the model's ability to discriminate between the two STS subgroups became more statistically significant, whereas the ability of either the variables at the tissue level or the variables at the organismal level to discriminate between the two STS subgroups remained statistically non-significant at best.

Discussion

On the performance of the model

In the original study [1], employing 89 subjects, the model exhibited a sensitivity of 92.9% and a specificity of 96.0%, with a significant difference ($P = 4.05 \times 10^{-18}$) between the scores of the two groups (STS vs. LTS). In this second, independent cohort study, employing 102 new subjects, the model exhibited a sensitivity of 98.4% and a specificity of 95.0%, with a significant difference ($P = 3.83 \times 10^{-36}$) between the scores of the two groups. It is evident, therefore, that the performance of the model in this study improved considerably, and that the statistically significant difference between the two groups increased considerably. If the results of both studies are combined, the overall performance of the model thus far is: overall sensitivity = 97.4% and overall specificity = 95.7%, with a significant difference ($P = 1.29 \times 10^{-44}$) between the scores of the two groups (STS vs. LTS). It should be pointed out here that as the subject sample size increased substantially (either from 89 or from 102 to 191), the statistically significant difference between the scores of the two groups also increased substantially. These overall results provide another perspective of the considerable improvement in the model's performance and demonstrate its overall accuracy.

Given that in this study, 54 STS subjects with GBM were included, and given that the model identified all of them correctly, it stands to reason that the domain of the model, as was theoretically expected in the past, can now be expanded to include subjects with glioma grade II, III, or IV.

On the top 29 most significant genes

In this study, and compared with the original one, all 29 genes that had been identified in the original study as the most important in the process of short-term survival vs. long-term survival in patients with glioma became substantially more significant with the exception of one [# 29 *TMPRSS3* ([Table 1](#))]. Furthermore, and more importantly, and also in terms of statistical significance, the 5 genes that constitute the input variables of the algorithm of the model (*FAM120AOS*, *MXI1*, *OCIAD2*, *PCDH15*, and *PDLIM4*) attained levels that were among

Independent validation of a genomic model for survival of glioma patients

the highest observed in this study (**Table 1** and **Table 1** of the original study). This in itself provides further support for the theoretical foundation of the model.

On short-term survival in connection with lower-grade glioma vs. glioblastoma

The model was designed to operate in two very distinct survival time intervals [short-term survival (≤ 1 year) and long-term survival (≥ 3 years)] and at the molecular level (gene expression of tumor cells). More specifically, the model was designed to quantify a certain genomic state of the tumor cells of a patient and, in relation to the cut-off point, to determine whether that patient will be a short-term survivor or a long-term survivor. Moreover, the farther away a patient's score is from the cut-off point (whether below or above), then 1) the less - or the more - advanced the genomic state of the tumor cells of the patient is, respectively, and 2) the higher the certainty is that the patient belongs to the respective group. The most important point here is that survival - a phenomenon at the organismal level - is determined at the molecular level. Since molecular changes precede cellular changes, which in turn precede tissue changes, which in turn precede organismal changes, it follows that changes at the molecular level may be the earliest accurate prognosticators of developments at the organismal level long before those developments occur. In this study, that was indeed the case. Because the short-term survival time interval is short (≤ 1 year), variables at the tissue level (histology) were not able to detect a significant difference between the STS subjects with lower glioma and STS subjects with glioblastoma in connection with survival. The model, on the other hand, examining genomic changes, i.e., changes at the molecular level, was able to detect a significant difference between those two groups in connection with survival. More specifically, based on the genomic profile of their tumor cells, and in connection with survival, the model calculated scores for the STS subjects with glioblastoma that were significantly higher than those for the STS subjects with lower-grade glioma, which means that, in terms of survival, the former subjects were farther away (higher) from the cut-off point of 25.2, that is to say that they were worse off, than the latter subjects were. Moreover, the ability of the model to detect a significant difference between those two groups increased in

terms of statistical significance as the sample size of the STS subjects increased. In conclusion, the results of this study lend additional, independent support for the model's accuracy, theoretical foundation, and clinical utility.

Acknowledgements

This study was supported by Genomix Inc. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The author would like to thank Emily G. Nikas and Mary & George Headrick for helpful discussions.

Disclosure of conflict of interest

The author is a partner of Genomix Inc.

Authors' contribution

JBN conceived, designed, and carried out all aspects of this study and wrote and edited the manuscript.

Address correspondence to: Dr. Jason B Nikas, Genomix Inc., Minneapolis, MN, USA. E-mail: jbnikas@genomix-inc.com

References

- [1] Nikas JB. A mathematical model for short-term vs. long-term survival in patients with glioma. *Am J Cancer Res* 2014; 4: 862-873.
- [2] Nikas JB, Lee JT, Maring ED, Washechek-Aletto J, Felmlee-Devine D, Johnson RA, Smyrk TC, Tawadros PS, Boardman LA and Steer CJ. A common variant in MTHFR influences response to chemoradiotherapy and recurrence of rectal cancer. *Am J Cancer Res* 2015; 5: 3231-3240.
- [3] Nikas JB. Inflammation and immune system activation in aging: a mathematical approach. *Sci Rep* 2013; 3: 3254.
- [4] Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougale RM, Brown WL, Law EK and Harris RS. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 2013; 494: 366-370.
- [5] Leonard B, Hart SN, Burns MB, Carpenter MA, Temiz NA, Rathore A, Vogel RI, Nikas JB, Law EK, Brown WL, Li Y, Zhang Y, Maurer MJ, Oberg AL, Cunningham JM, Shridhar V, Bell DA, April C, Bentley D, Bibikova M, Cheetham RK, Fan JB, Grocock R, Humphray S, Kingsbury Z, Peden J, Chien J, Swisher EM, Hartmann LC, Kalli KR, Goode EL, Sicotte H, Kaufmann SH and Harris RS. APOBEC3B upregulation and

Independent validation of a genomic model for survival of glioma patients

- genomic mutation patterns in serous ovarian carcinoma. *Cancer Res* 2013; 73: 7222-7231.
- [6] Nikas JB, Low WC and Burgio PA. Prognosis of treatment response (pathological complete response) in breast cancer. *Biomark Insights* 2012; 7: 59-70.
- [7] Nikas JB and Low WC. Linear Discriminant Functions in Connection with the micro-RNA Diagnosis of Colon Cancer. *Cancer Inform* 2012; 11: 1-14.
- [8] Nikas JB, Boylan KL, Skubitz AP and Low WC. Mathematical prognostic biomarker models for treatment response and survival in epithelial ovarian cancer. *Cancer Inform* 2011; 10: 233-247.
- [9] Nikas JB and Low WC. Application of clustering analyses to the diagnosis of Huntington disease in mice and other diseases with well-defined group boundaries. *Comput Methods Programs Biomed* 2011; 104: e133-147.
- [10] Nikas JB and Low WC. ROC-supervised principal component analysis in connection with the diagnosis of diseases. *Am J Transl Res* 2011; 3: 180-196.
- [11] Nikas JB, Keene CD and Low WC. Comparison of analytical mathematical approaches for identifying key nuclear magnetic resonance spectroscopy biomarkers in the diagnosis and assessment of clinical change of diseases. *J Comp Neurol* 2010; 518: 4091-4112.