VOLUME 28 · NUMBER 14 · MAY 10 2010

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Polymorphisms of *LIG4, BTBD2, HMGA2,* and *RTEL1* Genes Involved in the Double-Strand Break Repair Pathway Predict Glioblastoma Survival

Yanhong Liu, Sanjay Shete, Carol J. Etzel, Michael Scheurer, George Alexiou, Georgina Armstrong, Spyros Tsavachidis, Fu-Wen Liang, Mark Gilbert, Ken Aldape, Terri Armstrong, Richard Houlston, Fay Hosking, Lindsay Robertson, Yuanyuan Xiao, John Wiencke, Margaret Wrensch, Ulrika Andersson, Beatrice S. Melin, and Melissa Bondy

From the Departments of Epidemiology, Neuro-Oncology, and Pathology, The University of Texas M. D. Anderson Cancer Center; Department of Pediatrics and Dan L. Duncan Cancer Center, Baylor College of Medicine; Department of Integrative Nursing Care, University of Texas Health Science Center School of Nursing, Houston, TX; Department of Neurosurgery, University Hospital of Ioannina, Ioannina, Greece; Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, United Kingdom; Department of Epidemiology and Biostatistics, Neurological Surgery, University of California, San Francisco, San Francisco, CA; and the Department of Radiation Sciences, Oncology, Umeå University Hospital, Umeå, Sweden.

Submitted October 19, 2009; accepted January 28, 2010; published online ahead of print at www.jco.org on April 5, 2010.

Supported by Grants No. 5R01CA119215, 5R01CA070917, and K07CA093592 from the National Institutes of Health (to M. D. Anderson); the American Brain Tumor Association and the National Brain Tumor Society; The Wellcome Trust; Grant No. R01CA52689 from the National Institutes of Health; University of California, San Francisco (UCSF) Brain Tumor SPORE P50CA097257 (to University of California, San Francisco); the Swedish Research Council; Umeå University Hospital excellence grant; the Cancer Foundation of Northern Sweden; the Swedish Cancer Society; Grant LSHC-CT-2004-502943 Mol Cancer Med from the European Union; and by acta oncologica through the Royal Swedish Academy of Science (to the Swedish centre).

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Melissa L. Bondy, PhD, Department of Epidemiology, Unit 1340, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030; e-mail: mbondy@mdanderson.org.

0732-183X/10/2814-2467/$20.00

DOI: 10.1200/JCO.2009.26.6213

ABSTRACT

Purpose
Glioblastoma (GBM) is the most common and aggressive type of glioma and has the poorest survival. However, a small percentage of patients with GBM survive well beyond the established median. Therefore, identifying the genetic variants that influence this small number of unusually long-term survivors may provide important insight into tumor biology and treatment.

Patients and Methods
Among 590 patients with primary GBM, we evaluated associations of survival with the 100 top-ranking glioma susceptibility single nucleotide polymorphisms from our previous genome-wide association study using Cox regression models. We also compared differences in genetic variation between short-term survivors (STS; ≤ 12 months) and long-term survivors (LTS; ≥ 36 months), and explored classification and regression tree analysis for survival data. We tested results using two independent series totaling 543 GBMs.

Results
We identified *LIG4* rs7325927 and *BTBD2* rs11670188 as predictors of STS in GBM and *CCDC26* rs10464870 and rs891835, *HMGA2* rs1563834, and *RTEL1* rs2297440 as predictors of LTS. Further survival tree analysis revealed that patients ≥ 50 years old with *LIG4* rs7325927 (V) had the worst survival (median survival time, 1.2 years) and exhibited the highest risk of death (hazard ratio, 17.53; 95% CI, 4.27 to 71.97) compared with younger patients with combined *RTEL1* rs2297440 (V) and *HMGA2* rs1563834 (V) genotypes (median survival time, 7.8 years).

Conclusion
Polymorphisms in the *LIG4, BTBD2, HMGA2,* and *RTEL1* genes, which are involved in the double-strand break repair pathway, are associated with GBM survival.

J Clin Oncol 28:2467-2474. © 2010 by American Society of Clinical Oncology

INTRODUCTION

Glioblastoma (GBM) is the most common and most malignant primary brain tumor in US and European countries, with an annual incidence of approximately three in 100,000 people newly diagnosed each year.[1] Despite recent advances in treatment including surgical resection followed by concurrent chemotherapy with radiation, the median survival remains approximately 9 to 15 months.[2] Nevertheless, a subset of patients survived for longer than 3 years. Although certain clinical features, such as younger age, good Karnofsky performance status at the time of diagnosis, and extent of resection, are well-known prognostic parameters.[3-5] However, it is likely that other, as-yet-unknown genetic factors may help predict which patients are more likely to have this prolonged survival. Therefore, it is important to determine the genetic factors that influence survival for this rapidly fatal disease and, by doing so, perhaps uncover the molecular signatures of long-term survivorship.

Subtypes of GBM exist, despite indistinguishable features by pathologic evaluation with differing survival durations and responses to treatment.[6] Some genetic aberrations in GBM have been known for years, such as *MGMT*, *ERBB2*, *EGFR*, *TP53*, and

PTEN.[7-9] Recent studies demonstrated that mutations in *IDH1*[10] and alterations in multiple networking genes (*POLD2, CYCS, MYC, AKR1C3, YME1L1, ANXA7,* and *PDCD4*)[11] were associated with GBM survival. Polymorphisms in *TERT*,[12] *IL4R*,[13] *EGF*,[14,15] and *CX3CR1*[16] genes and GBM survival have also been described. However, to date, no prognostic biomarkers have been found to be sufficiently reliable or useful for clinical practice.

Recently, our group[17] and others[18] found, using GWAS (genome-wide association study) methods, several single nucleotide polymorphisms (SNPs) that were strongly associated with glioma susceptibility. However, both of those studies focused on cancer susceptibility rather than outcomes. For this study, we applied a survival analysis to the 100 top-ranking glioma susceptibility SNPs from the GWAS to determine whether they were also associated with GBM survival. We used two analytic strategies to fulfill this goal. First, we examined overall survival in 590 patients with GBM. Second, to determine whether there are specific genetic variations that distinguish long-term survivors (LTS; survival $\geq$ 36 months) from short-term survivors (STS; survival $\leq$ 12 months), we further compared the differences in gene variations in 97 LTS and 202 STS. We then tested our findings in two additional independent data series.

PATIENTS AND METHODS

Study Population

The population for this study was a subset of GBMs from a prospective study of patients with malignant glioma (N = 1,304) consecutively diagnosed and treated at The University of Texas M. D. Anderson Cancer Center (MDACC), Houston, TX.[17,19] The patients were participants in an ongoing glioma epidemiology study from 1992 to 2008, with follow-up until May 2009. We selected primary patients with GBM (International Classification of Diseases code 94403) who were $\geq$ 18 years old at the time of diagnosis. We only included white patients because there are very few non-white patients—11 Hispanic, 12 African American, and three Asian for any meaningful analyses.

Treatment and survival data (dates of death or last contact) were collected retrospectively from medical record review for all patients. Dates of death were confirmed by querying the Social Security Death Index. From the obtained data, we observed patients from the earliest date of registration at MDACC in May 1992 to the date of last contact or death through May 2009. The follow-up periods ranged from 1 month to 17 years (ie, 204 months). The study was approved by M. D. Anderson institutional review board, and written informed consent was obtained from each participant.

Subjects for the Replication Phase

As a validation group, we obtained data for patients from two other institutions (Appendix A1, online only). The first group included 352 newly diagnosed patients with GBM who were seen at the University of California, San Francisco (UCSF) between 2001 and 2006.[18] The second group included 191 patients with GBM ascertained in an outcome study at Umeå University Hospital accrued 2002 to 2007.

Selection of the SNPs and Genotyping Assays

The top-ranking 100 glioma susceptibility SNPs with the smallest *P* values in our GWAS were selected and examined in this study (Appendix Table A2, online only). We contracted with Illumina to conduct the genotyping using the Human 610-Quad Bead Chips (Illumina, San Diego, CA). Subsequent genotyping of SNPs was conducted using either the Illumina 317k chip by decode Genetics (UCSF samples) or single-base primer extension chemistry matrix assisted laser desorption and ionization time of flight mass spectrometry detection by Sequenom (Swedish samples).

Statistical Methods

Survival time was defined as the time between the date of diagnosis and date of death for deceased patients or the last contact date for living patients. The overall survival time was estimated using Kaplan-Meier methods, and log-rank analysis was performed to compare survival curves between groups. Hazard ratios (HRs) and their corresponding 95% CIs were estimated using Cox regression, with adjustment for age, sex, and extent of resection. Genotype frequencies of the LTS and the STS were compared using χ^2 tests. We calculated the odds ratios (ORs) and 95% CIs by unconditional logistic regression analysis with adjustment for diagnosis age, sex, and extent of resection.

To evaluate the chance of obtaining a false-positive association in our data set, we used the false-positive report probability (FPRP) test[20] and the Bayesian false-discovery probability (BFDP) test.[21] For our analyses, we used the moderate range of prior probabilities .08 and .05; the FPRP and BFDP cutoff value of .2 and .8, respectively, as suggested by the authors for summary analyses.[20,21]

Finally, we produced a classification and regression tree (CART) for survival data to identify higher-order interactions between clinical factors and genetic variants using the RPART package[22] in S-PLUS Version 8.0.4 (TIBCO, Palo Alto, CA). CART is a prognostic system with a hierarchical structure, based on recursive portioning that build a decision tree to identify subgroups at higher risk of death. Specifically, the recursive procedure starts at the root node and uses a log-rank statistic to determine the first optimal split and each subsequent split of the data set. This process continues until the terminal nodes have no subsequent statistically significant splits or the terminal nodes reach a prespecified minimum size ($n > 10$). Each end node contains the numbers of total and censored observations falling into the current category, as well as a HR and 95% CI adjusted by age, sex, and extent of resection. The reference group is the one with the smallest percentage of censored observations. Kaplan-Meier estimation of the cumulative survival for each subgroup was also performed; log-rank *P* value was set to indicate significance.

RESULTS

Patient Characteristics

The final population of the MDACC study consisted of 590 patients with primary GBM (Table 1). At the end of our study (May 2009), 544 (92.2%) of 590 patients with GBM had died. The median survival time (MST) was 22.63 months. Among these 590 patients with GBM, 202 were STS, and 97 were LTS (52 survived $\geq$ 5 years). The MST was 7.8 and 76.1 months in the STS and LTS, respectively. No patients survived in the STS, whereas 48 (49.5%) were still alive in the LTS.

Examinations of the demographic data for the LTS and STS shows that patients in the LTS (mean age, 48.1 years) were significantly younger than patients in the STS (mean age, 55.9 years; $P < .001$). Also, patients in the LTS were much more likely to have undergone a gross total resection than were patients in the STS (57.6% *v* 21.6%; $P < .001$).

Association Between SNPs and Overall Survival

Of the 100 SNPs evaluated, six showed a statistically significantly correlation with overall survival according to the log-rank test and Cox regression analysis (Table 2). *LIG4* rs7325927, *BTBD2* rs11670188, *RGS22* rs4734443, and chromosome 3 rs13099725 were associated with STS, whereas *RTEL1* rs2297440 and rs6010620 were associated with LTS. A strong gene-dosage effect on survival was

Table 1. Demographics and Treatment Information for the Study Population

Variable	All Patients (N = 590)							By Survival Time (n = 299)				
	Patients*		No. of Deaths*	MST (months)	Log-Rank *P*	Cox Regression		LTS (n = 97)*		STS (n = 202)*		
	No.	%				HR	95% CI	No.	%	No.	%	χ^2 *P*
Age at diagnosis, years					< .0001							< .0001
Mean	52.31							48.11		55.97		
SD	11.54							12.93		10.53		
Range	18-86							18-78		20-86		
≤ 50	232	39.3	196	31.61		1	Reference	55	56.7	55	27.2	
> 50	358	60.7	348	18.35		1.63	1.37-1.95	42	43.3	147	72.8	
Sex					.2587							.412
Male	365	61.9	339	21.87		1	Reference	53	54.6	122	60.4	
Female	225	38.1	205	25.49		0.91	0.76-1.08	44	45.4	80	39.6	
Radiation therapy					NA							NA
No	2	0.3	2	1		1	Reference	0	0	3	1.7	
Yes	517	87.6	477	24.67		NA		88	100.0	174	98.3	
Chemotherapy					< .0001							NA
No	12	2.0	12	6.42		1	Reference	0	0	11	6.9	
Yes	478	81.0	437	25.99		0.19	0.11-0.34	89	100.0	148	93.1	
Extent of resection					< .0001							< .0001
Biopsy	104	17.3	101	16.80		1	Reference	9	12.3	56	33.5	
Subtotal	225	38.1	214	20.21		0.77	0.61-0.97	22	30.1	75	44.9	
Gross	183	31.0	158	31.23		0.48	0.37-0.62	42	57.6	36	21.6	

Abbreviations: HR, hazard ratio; MST, median survival time (months); LTS, long-term survivors; STS, short-term survivors; NA, not available.
*Numbers do not add up to the column totals due to missing values.

observed when the six SNPs were analyzed in combination (Table 2), with highly significant differences for overall survival (log-rank $P < .001$). The MST decreased progressively as the number of at-risk SNPs increased.

Association of SNPs With Long-Term Survival Between LTS and STS

Eight SNPs were statistically different between LTS and STS. As presented in Table 3, *LIG4* rs7325927 and *BTBD2* rs11670188 were associated with STS, whereas *CCDC26* rs10464870 and rs891835 (in strong linkage disequilibrium [LD]), *VPS8* rs6765837, *RTEL1* rs2297440, and rs6010620 (in strong LD), *HMGA2* rs1563834 were associated with LTS. The joint effect of these combined eight SNPs showed that the risk of early death increased progressively as the number of at-risk SNPs increased ($P_{trend} < .001$).

To evaluate the robustness of these findings, we calculated FPRP and BFDP values at two levels of prior probabilities (.08 and .05) for the eight SNPs (Table 4). At a prior probability level of .08, six of these eight SNPs (*LIG4* rs7325927, *BTBD2* rs11670188, *CCDC26* rs10464870 and rs891835, *VPS8* rs6765837, and *RTEL1* rs2297440) remained noteworthy (FPRP ≤ 0.2, or BFDP ≤ 0.8). At a very low prior probability of .05, only *LIG4* rs7325927 remained noteworthy.

Survival Tree Analysis

We built a survival tree using the clinical variables (age and sex) and the six SNPs that were found to be noteworthy at prior probability of 0.08. Figure 1A depicts the survival tree structure generated by CART analysis. The first split on the decision tree was age, confirming that age was the most important risk factor among those considered. Further inspection of the tree structure suggested distinct patterns of survival, resulting seven terminal nodes belonging to two sides of the tree. In particular, young patients (≤ 50 years) with combined variant type (V) of *RTEL1* rs2297440 and *HMGA2* rs1563834 had the best prognosis, MST 93.9 months (7.8 years; node 1). This subgroup has 10 patients, all were LTS, and eight of them were still alive. Using node 1 as reference group, the young individuals with wild type (W) of *RTEL1* rs2297440, *BTBD2* rs11670188 (V), and *CCDC26* rs10464870 (W) exhibited worse survival (MST, 13.8 months; node 5) and a 16-fold risk of death (HR, 16.84; 95% CI, 3.88 to 73.04). Node 5 has 19 patients, all of them were deceased. Likewise, the older patients (> 50 years) with *LIG4* rs7325927 (V) also showed a very short survival time (MST, 14.8 months; node 7) and exhibited the highest risk of death (HR, 17.53, 95% CI, 4.27 to 71.97), supporting the protective effect of *RTEL1, HMGA2,* and *CCDC26,* and risk effect of old age, *BTBD2* and *LIG4* found in the main effect analysis. Figure 1B shows the Kaplan-Meier estimates of cumulative survival for each end node.

Replication Results

We further looked at the six noteworthy SNPs (*LIG4* rs7325927, *BTBD2* rs11670188, *CCDC26* rs10464870 and rs891835, *VPS8* rs6765837, and *RTEL1* rs2297440) using two independent series. In the UCSF study, except for *VPS8* rs6765837, all of other five SNPs were available for confirmation. In the Swedish study, *VPS8* rs6765837 and *BTBD2* rs11670188 were unavailable for analysis. The confirmation analysis was adjusted for age and sex. Only the heterozygote of *BTBD2* rs11670188 in UCSF group showed significant association with overall

Table 2. Association Between Genotype and Overall Survival (N = 590)

Gene and SNP	No.		MST (months)	Log-Rank *P*	Cox Regression*	
	Patients	Deaths			HR	95% CI
Risk SNPs						
Unknown rs13099725						
AA	434	402	23.31	.007	1	Reference
AG	140	127	25.1		0.95	0.77 to 1.22
GG	16	15	9.18		2.06	1.22 to 3.61
AA + AG	574	529	23.84	.002	1	Reference
GG	16	15	9.18		1.91	1.18 to 3.53
RGS22 rs4734443						
CC	230	214	22.35	.011	1	Reference
TC	269	244	25.93		0.92	0.74 to 1.20
TT	91	86	16.39		1.36	1.07 to 1.76
TC + CC	499	458	24.64	.003	1	Reference
TT	91	86	16.39		1.33	1.01 to 1.82
BTBD2 rs11670188						
AA	354	320	25.81	.032	1	Reference
AG	198	188	19.77		1.29	1.03 to 1.58
GG	38	36	20.52		1.33	0.94 to 1.81
AG + GG	236	224	19.89	.009	1.25	1.06 to 1.65
LIG4 rs7325927						
CC	309	282	25.26	.058	1	Reference
TC	217	200	21.87		1.19	0.97 to 1.45
TT	64	61	15.6		1.63	1.24 to 2.29
TC + TT	281	261	16.81	.008	1.54	1.06 to 2.12
No. of risk alleles						
0-1	454	415	24.94	.005	1	Reference
2-3	121	115	18.98		1.25	1.00 to 1.53
4	15	14	9.86		1.69	1.01 to 2.95
Protective SNPs						
RTEL1 rs2297440						
CC	391	366	21.76	.021	1	Reference
TC	179	162	25.82		0.83	0.60 to 1.02
TT	19	15	27.95		0.55	0.26 to 0.94
TC + TT	198	177	26.78	.02	0.81	0.62 to 0.99
RTEL1 rs6010620						
GG	387	362	21.78	.025	1	Reference
AG	183	166	25.71		0.81	0.63 to 1.04
AA	20	16	27.1		0.57	0.27 to 0.95
AG + AA	203	182	26.58	.021	0.80	0.61 to 0.98
No. of protective alleles						
0	387	362	21.77	.023	1	Reference
1	182	165	25.78		0.84	0.71 to 1.02
2	20	16	27.02		0.53	0.28 to 0.92
No. of at-risk alleles for the 6 SNPs†						
0-2	494	452	25.28	< .001	1	Reference
3-4	83	80	14.37		1.75	1.25 to 2.19
5-6	12	11	9.41		1.63	0.88 to 2.86

Abbreviations: SNP, single nucleotide polymorphism; HR, hazard ratio. MST, median survival time (months).
*Adjusted for diagnosis age, sex and extent of resection.
†At-risk alleles were defined as the minor allele of the risk SNPs and the common allele of the protective SNPs.

survival (HR, 1.31; 95% CI, 1.03 to 1.68; Appendix Table A3, online only). None of the SNPs were significantly different between STS and LTS in both replication data sets (data not shown).

DISCUSSION

In this study, we identified *LIG4* rs7325927 and *BTBD2* rs11670188 as predictors of STS in GBM and *CCDC26* rs10464870 and rs891835, *HMGA2* rs1563834, and *RTEL1* rs2297440 as predictors of LTS. Consistent with most other studies,[23] we found age at diagnosis to be a very important predictive factor in GBM survival. In the Surveillance, Epidemiology and End Results data, 5-year survival rates are approximately 13% for 15 to 45 year olds and only 1% for those ≥ 75 years old.[24] Our LTS (mean age, 48.1 years) was significantly younger than the STS (mean age, 55.9 years). Moreover, age is the initial split on the survival tree, confirming that age is the most important risk factor.

Table 3. Association Between the STS and LTS

Gene and SNP	LTS (n = 97)		STS (n = 202)			Logistic Regression*	
	No.	%	No.	%	*P*	OR	95% CI
Risk SNPs							
BTBD2 rs11670188							
AA	69	71.1	117	57.9	.067	1	Reference
AG	23	23.7	71	35.1		2.12	1.09 to 4.35
GG	5	5.2	14	6.9		1.79	0.43 to 5.26
AG + GG	28	28.9	85	42.1	.02	2.03	1.09 to 3.81
LIG4 rs7325927							
CC	62	64.6	104	51.5	.001	1	Reference
TC	28	29.2	72	35.6		2.33	1.13 to 4.55
TT	6	6.2	26	12.9		3.70	1.09 to 11.11
TC + TT	34	35.4	98	48.5	.004	2.45	1.32 to 4.56
No. of risk alleles							
0	43	44.8	62	30.7	.046	1	Reference
1	49	51.0	129	63.9		1.92	1.12 to 3.23
2	4	4.2	11	4.4		2.78	0.66 to 9.89
Protective SNPs							
RTEL1 rs2297440							
CC	57	58.8	142	70.3	.048	1	Reference
TC	36	37.1	58	28.7		0.52	0.28 to 0.97
TT	4	4.1	2	1.0		0.16	0.02 to 1.37
TC + TT	40	41.2	60	29.7	.027	0.50	0.27 to 0.93
RTEL1 rs6010620							
GG	57	58.8	141	69.8	.082	1	Reference
AG	36	37.1	58	28.7		0.53	0.28 to 0.97
AA	4	4.1	3	1.5		0.41	0.06 to 2.86
AG + AA	40	41.2	61	30.2	.041	0.53	0.29 to 0.98
HMGA2 rs1563834							
GG	63	66	153	75.7	.021	1	Reference
AG	29	27.8	48	23.8		0.63	0.32 to 1.20
AA	4	4.1	1	0.5		0.06	0.01 to 0.78
AG + AA	33	34	49	24.3	.045	0.56	0.27 to 0.99
CCDC26 rs10464870							
TT	49	50.5	133	65.8	< .001	1	Reference
TC	38	39.2	66	32.7		0.62	0.38 to 1.01
CC	10	10.3	3	1.5		0.11	0.04 to 0.49
TC + CC	48	49.5	69	34.2	.016	0.53	0.30 to 0.89
CCDC26 rs891835							
TT	48	49.5	128	63.4	.01	1	Reference
TG	41	42.3	70	34.6		0.56	0.30 to 1.01
GG	8	8.2	4	2.0		0.19	0.05 to 0.65
TG + GG	49	50.5	74	36.6	.026	0.51	0.28 to 0.93
VPS8 rs6765837							
TT	18	18.6	64	31.7	.056	1	Reference
TC	58	59.8	97	48		0.46	0.23 to 0.91
CC	21	21.6	41	20.3		0.54	0.24 to 1.22
TC + CC	79	81.4	138	68.3	.028	0.49	0.26 to 0.92
No. of protective alleles							
0-1	8	8.2	53	26.2	< .001	1	Reference
2-3	70	72.2	126	62.4		0.26	0.12 to 0.61
4-6	19	19.6	23	11.4		0.17	0.08 to 0.54
No. of at-risk alleles for the 8 SNPs†							
0-2	8	8.2	56	27.7	< .001	1	Reference
3-4	62	63.9	124	61.4		3.57	1.54 to 8.43
5-8	27	27.8	22	10.9		8.53	3.23 to 22.98

Abbreviations: STS, short-term survivors; LTS, long-term survivors; SNP, single nucleotide polymorphism; OR, odds ratio.
*Adjusted for age, sex, and extent of resection.
†At-risk alleles were defined as the minor allele of the risk SNPs and the common allele of the protective SNPs.

Further, different genes play roles in different age group. The older patients (> 50 years) with variant type of *LIG4* rs7325927 showed the worst prognosis; whereas young patients (≤ 50 years) with combined variant type of *RTEL1* rs2297440 and *HMGA2* rs1563834 had the best prognosis.

A major finding in this study was the consistent association of *LIG4* rs7325927 with STS. *LIG4* rs7325927 was the only noteworthy SNP at the FPRP low prior probability of .05. Survival tree analysis showed that older patients with *LIG4* rs7325927 (V) exhibited the highest risk of death, which strongly suggests that *LIG4* rs7325927

Table 4. FPRP and BFDP Values for the Significant SNPs on Associations of Between the STS and LTS

Gene	SNP	OR*	95% CI*	FPRP Prior		BFDP Prior	
				0.08	0.05	0.08	0.05
LIG4	Rs7325927	2.45	1.32 to 4.56	**0.09**	**0.15**	**0.53**	**0.65**
BTBD2	Rs11670188	2.03	1.09 to 3.81	0.38	0.50	**0.79**	0.86
CCDC26	Rs10464870	0.53	0.30 to 0.89	0.27	0.38	**0.72**	**0.80**
CCDC26	Rs891835	0.51	0.28 to 0.93	0.39	0.52	**0.79**	0.86
VPS8	Rs6765837	0.49	0.26 to 0.92	0.38	0.50	**0.78**	0.85
RTEL1	Rs2297440	0.50	0.27 to 0.93	0.38	0.50	**0.79**	0.86
RTEL1	Rs6010620	0.53	0.29 to 0.98	0.49	0.62	0.83	0.89
HMGA2	Rs1563834	0.56	0.27 to 0.99	0.51	0.64	0.84	0.89

NOTE. Bold data have a noteworthy association at the 0.2 FPRP or 0.8 BFDP level suggesting a true association.
Abbreviations: FPRP, false-positive report probability; BFDP, Bayesian false-discovery probability; SNP, single nucleotide polymorphism; STS, short-term survivors; LTS, long-term survivors; OR, odds ratio.
*OR were calculated for the heterozygote and mutant homozygous *v* the common homozygous types, as reported in Table 3.

or a genetic variant in LD with this SNP is associated with STS. *LIG4* (DNA ligase IV; OMIM 601837) is essential for V(D)J recombination and DNA double-strand break repair (DSBR) through nonhomologous end joining (NHEJ). As a critical protein involved in NHEJ, *LIG4* forms a heterodimer with *XRCC4* to execute the final rejoining step of NHEJ. Polymorphisms of this gene have been related to risk of glioma[25] and multiple myeloma,[26] as well as to the survival of breast cancer.[27]

Another promising finding is the association of *BTBD2* rs11670188 with STS. Coincidentally, *BTBD2* (blood-tumor barrier

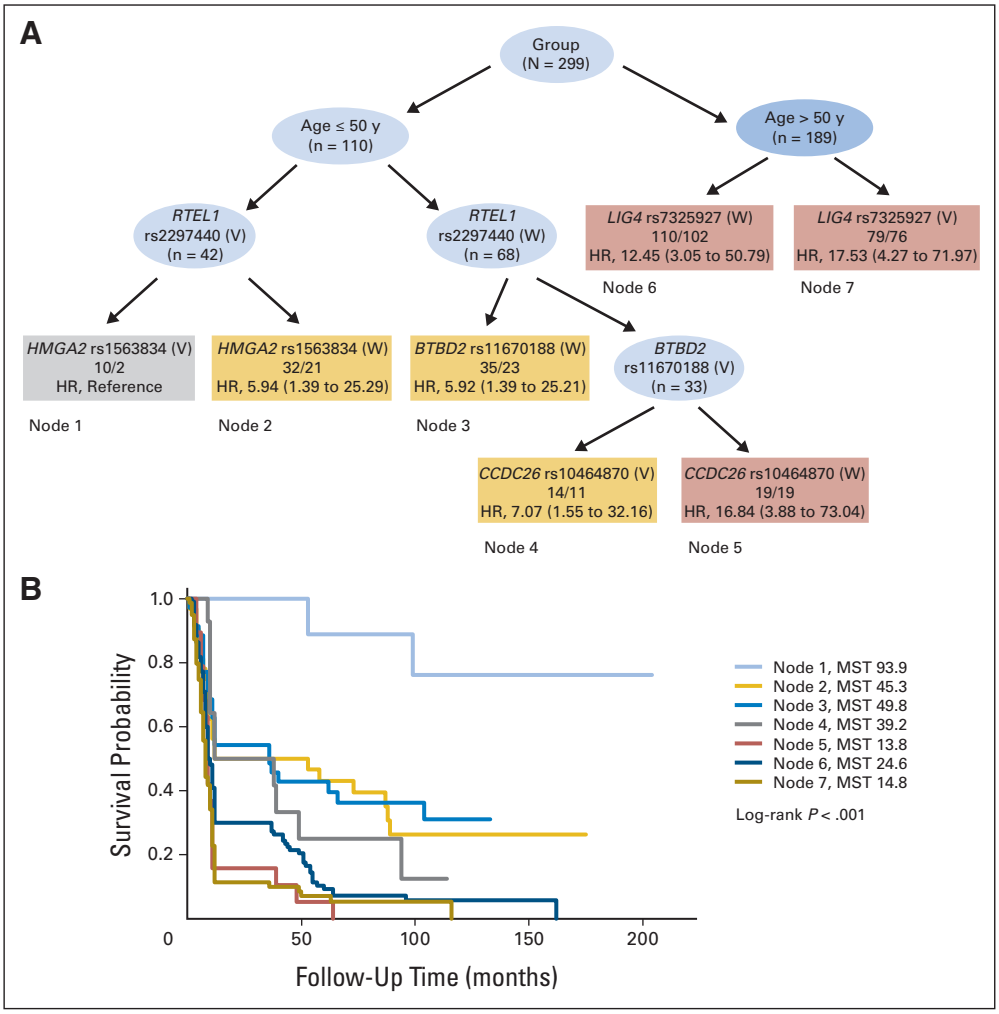

Fig 1. Survival tree analysis. (A) Survival tree. The number of patients and events/deceased are indicated in the nodes. (B) Kaplan-Meier curves of survival times in patients for seven nodes. Node1 had a significantly improved prognosis, with a median survival time (MST; in months) of 7.8 years, compared with 1.2 years for node 5. y, years; V, variant type; W, wild type; HR, hazard ratio.

modification; OMIM 608531) has also been implicated in DSBR pathway because of its tight interaction with *Top1*. *Top1* can cause double-stranded breaks[28] and plays a central role in regulating cell survival.[29,30] Bredel et al[31] have shown that high expression of DNA Top II alpha is associated with prolonged survival in patients with GBM. The *CCDC26*, *HMGA2,* and *RTEL1* genes are also of interest. *CCDC26* modulates retinoic acid, which in turn increases programmed cell death in GBM cells and reduces telomerase activity[32-34]; whereas both *HMGA2* (high-mobility group protein family, member 2; OMIM 600698)[35,36] and *RTEL1* (regulator of telomere elongation helicase 1; OMIM 608833)[37] were recently proved to be directly involved in DSBR pathway.

It is interesting to note that of the five important genes noted in this study, four (*LIG4*, *BTBD2*, *HMGA2,* and *RTEL1*) are directly or indirectly involved in the DSBR pathway, suggesting a very strong genetic interaction network. This is particularly intriguing because DSBR plays a prominent role in cell survival, maintenance of genomic integrity, and prevention of tumorigenesis. DSBs can be generated by endogenous reactive oxygen species or destabilization of stalled replication forks as well as by exposure to a variety of exogenous agents, including ionizing radiation (IR) and chemotherapeutic agents. Furthermore, IR is the only established risk factor for glioma,[38-40] and IR is also used in cancer radiation therapy. Germline variation in DSBR capabilities may affect survival because of altered response to radiation or chemotherapy. Therefore, understanding how these DSBR genes and SNPs function epistatically in the same biologic pathway that influences survival of patients with GBM, whether as enhancers or inhibitors of response to radiation or chemotherapy would be very interesting in future studies.

Although our study demonstrates a strong association of certain SNPs with outcome, and there is biologic plausibility for the associations observed in our study. However, there are limitations of our study. The main one is that most of our findings did not reach statistically significant level in the replication series. However, replication failure should not be surprising or be interpreted as necessarily refuting the initial findings because of the potential problems such as population stratification. Our study and the UCSF study are clinic based while the Swedish study was mainly population based. When comparing the characteristics of the three groups, we observed large discrepancies in age at diagnosis (median 52 years in ours *v* 56 in Swedish and UCSF cases), and survival time (MST 22.6 months in our data set *v* 16.3 in the UCSF, 13.9 months in Swedish set).

Genetic heterogeneity—this heterogeneity may exist even in populations of the same ethnic group such as whites with European ancestry.[41,42] We compared the distribution of four SNPs in the three groups, and found one (*LIG4* rs7325927) was significant in our data set versus Swedish and UCSF, one (*RTEL1* rs2297440) was marginal significant in ours and Swedish versus UCSF (Appendix Table A4, online only), indicating differences in the case populations. Different environmental exposures and different patterns of care would also lead to greater difficulties in replication. As raised by Morgan et al,[43] replication studies face the risk of nonvalidation because of the lack of generalizability, especially in a heterogeneous treatment population for a pathologically and clinically heterogeneous tumor such as GBMs (GBM, the glioblastoma "multiforme," the latter term now taken away, was a very useful reminder of this fact). Gorroochurn et al[44] point out, "even if these problems were to be remedied, trying to replicate many initial findings, even if they are quite significant, maybe be predisposed to failure and should not be interpret as necessarily contradiction the initial association." The same feeling has been echoed elsewhere,[45] and Liu et al[46] have provided a theoretical justification. Thus, on a second look, replication, while important and valuable, is difficult to achieve and may not be sufficient or necessary. Additional information from other lines of evidence, such as detailed molecular mechanistic studies, may be useful for validating and illuminating the functional relevance of genes identified in our study.

Another limitation is the possibility of effect change in treatment on survival because first-line GBM treatment has been changed from involved-field cranial irradiation to chemoirradiation with daily temozolomide in 2005. However, among 590 patients with GBM, 93.6% (552) were diagnosed before 2005; only 6.4%[38] were diagnosed after 2005, and none of them were LTS. Therefore, we do not think that there will be any bias toward genes that are important for DNA repairs. The third limitation is the potential limited applicability of the findings to other ethnic groups, such as African Americans, Asians, and Hispanics.

In conclusion, we have provided evidence to implicate that *LIG4* rs7325927 and *BTBD2* rs11670188 are predictors of GBM STS and that *CCDC26* rs10464870 and rs891835, *RTEL1* rs2297440 and rs6010620, and *HMGA2* rs1563834 are predictors of LTS. It is highly likely that the polymorphisms of these genes, which are directly or indirectly involved in the DSBR pathway, will be novel and potentially prognostic biomarkers for GBM survival, and this could be the beginning of a risk assessment model that could predict which patients would be LTS or STS.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** None **Research Funding:** Terri Armstrong, Schering-Plough, Genentech **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Yanhong Liu, Sanjay Shete, Richard Houlston, Melissa Bondy
Financial support: Sanjay Shete, Carol J. Etzel, Margaret Wrensch, Beatrice S. Melin, Melissa Bondy
Administrative support: Melissa Bondy
Provision of study materials or patients: Georgina Armstrong, Mark Gilbert, Ken Aldape, Terri Armstrong, Yuanyuan Xiao, John Wiencke, Margaret Wrensch, Beatrice S. Melin, Melissa Bondy

Collection and assembly of data: Georgina Armstrong, Spyros Tsavachidis, Mark Gilbert, Ken Aldape, Terri Armstrong, John Wiencke, Ulrika Andersson
Data analysis and interpretation: Yanhong Liu, Sanjay Shete, Carol J. Etzel, Michael Scheurer, Fu-Wen Liang, Richard Houlston, Fay Hosking, Lindsay Robertson, Yuanyuan Xiao, Margaret Wrensch, Ulrika Andersson, Beatrice S. Melin, Melissa Bondy
Manuscript writing: Yanhong Liu
Final approval of manuscript: Yanhong Liu, Sanjay Shete, Carol J. Etzel, Michael Scheurer, George Alexiou, Georgina Armstrong, Spyros Tsavachidis, Fu-Wen Liang, Mark Gilbert, Ken Aldape, Terri Armstrong, Richard Houlston, Fay Hosking, Lindsay Robertson, Yuanyuan Xiao, John Wiencke, Margaret Wrensch, Ulrika Andersson, Beatrice S. Melin, Melissa Bondy

REFERENCES

1. Holland EC: Glioblastoma multiforme: The terminator. Proc Natl Acad Sci U S A 97:6242-6244, 2000

2. Stupp R, Mason WP, van den Bent MJ, et al: Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352:987-996, 2005

3. Scott JN, Rewcastle NB, Brasher PM, et al: Which glioblastoma multiforme patient will become a long-term survivor? A population-based study. Ann Neurol 46:183-188, 1999

4. Curran WJ Jr, Scott CB, Horton J, et al: Recursive partitioning analysis of prognostic factors in three Radiation Therapy Oncology Group malignant glioma trials. J Natl Cancer Inst 85:704-710, 1993

5. Lacroix M, Abi-Said D, Fourney DR, et al: A multivariate analysis of 416 patients with glioblastoma multiforme: Prognosis, extent of resection, and survival. J Neurosurg 95:190-198, 2001

6. Colman H, Aldape K: Molecular predictors in glioblastoma: Toward personalized therapy. Arch Neurol 65:877-883, 2008

7. Martinez R, Schackert G, Yaya-Tur R, et al: Frequent hypermethylation of the DNA repair gene MGMT in long-term survivors of glioblastoma multiforme. J Neurooncol 83:91-93, 2007

8. Krex D, Klink B, Hartmann C, et al: Long-term survival with glioblastoma multiforme. Brain 130: 2596-2606, 2007

9. Cancer Genome Atlas Research Network: Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455:1061-1068, 2008

10. Parsons DW, Jones S, Zhang X, et al: An integrated genomic analysis of human glioblastoma multiforme. Science 321:1807-1812, 2008

11. Bredel M, Scholtens DM, Harsh GR, et al: A network model of a cooperative genetic landscape in brain tumors. JAMA 302:261-275, 2009

12. Wang L, Wei Q, Wang LE, et al: Survival prediction in patients with glioblastoma multiforme by human telomerase genetic variation. J Clin Oncol 24:1627-1632, 2006

13. Scheurer ME, Amirian E, Cao Y, et al: Polymorphisms in the interleukin-4 receptor gene are associated with better survival in patients with glioblastoma. Clin Cancer Res 14:6640-6646, 2008

14. Bhowmick DA, Zhuang Z, Wait SD, et al: A functional polymorphism in the EGF gene is found with increased frequency in glioblastoma multiforme patients and is associated with more aggressive disease. Cancer Res 64:1220-1223, 2004

15. Vauleon E, Auger N, Benouaich-Amiel A, et al: The 61 A/G EGF polymorphism is functional but is neither a prognostic marker nor a risk factor for glioblastoma. Cancer Genet Cytogenet 172:33-37, 2007

16. Rodero M, Marie Y, Coudert M, et al: Polymorphism in the microglial cell-mobilizing CX3CR1 gene is associated with survival in patients with glioblastoma. J Clin Oncol 26:5957-5964, 2008

17. Shete S, Hosking FJ, Robertson LB, et al: Genome-wide association study identifies five susceptibility loci for glioma. Nat Genet 41:899-904, 2009

18. Wrensch M, Jenkins RB, Chang JS, et al: Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 41:905-908, 2009

19. Liu Y, Scheurer ME, El-Zein R, et al: Association and interactions between DNA repair gene polymorphisms and adult glioma. Cancer Epidemiol Biomarkers Prev 18:204-214, 2009

20. Wacholder S, Chanock S, Garcia-Closas M, et al: Assessing the probability that a positive report is false: An approach for molecular epidemiology studies. J Natl Cancer Inst 96:434-442, 2004

21. Wakefield J: A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet 81:208-227, 2007

22. Therneau T, Atkinson B: Technical Report Series No. 61, An Introduction to Recursive Partitioning Using the RPART Routines. Rochester, MN, Department of Health Science Research, Mayo Clinic, 1997

23. Chang SM, Parney IF, Huang W, et al: Patterns of care for adults with newly diagnosed malignant glioma. JAMA 293:557-564, 2005

24. Central Brain Tumor Registry of the United States: SEER: Statistical report: Primary brain tumors in the United States, 2000-2004. Chicago, IL, Central Brain Tumor Registry of the United States, 2008

25. Liu Y, Zhou K, Zhang H, et al: Polymorphisms of LIG4 and XRCC4 involved in the NHEJ pathway interact to modify risk of glioma. Hum Mutat 29:381-389, 2008

26. Roddam PL, Rollinson S, O'Driscoll M, et al: Genetic variants of NHEJ DNA ligase IV can affect the risk of developing multiple myeloma, a tumour characterised by aberrant class switch recombination. J Med Genet 39:900-905, 2002

27. Goode EL, Dunning AM, Kuschel B, et al: Effect of germ-line genetic variation on breast cancer survival in a population-based study. Cancer Res 62:3052-3057, 2002

28. Xu L, Yang L, Hashimoto K, et al: Characterization of BTBD1 and BTBD2, two similar BTB-domain-containing Kelch-like proteins that interact with Topoisomerase I. BMC Genomics 3:1, 2002

29. Champoux JJ: DNA topoisomerases: Structure, function, and mechanism. Annu Rev Biochem 70:369-413, 2001

30. Wang JC: Cellular roles of DNA topoisomerases: A molecular perspective. Nat Rev Mol Cell Biol 3:430-440, 2002

31. Bredel M, Piribauer M, Marosi C, et al: High expression of DNA topoisomerase IIalpha and Ki-67 antigen is associated with prolonged survival in glioblastoma patients. Eur J Cancer 38:1343-1347, 2002

32. Yin W, Rossin A, Clifford JL, et al: Co-resistance to retinoic acid and TRAIL by insertion mutagenesis into RAM. Oncogene 25:3735-3744, 2006

33. Jiang M, Zhu K, Grenet J, et al: Retinoic acid induces caspase-8 transcription via phospho-CREB and increases apoptotic responses to death stimuli in neuroblastoma cells. Biochim Biophys Acta 1783: 1055-1067, 2008

34. Das A, Banik NL, Ray SK: Differentiation decreased telomerase activity in rat glioblastoma C6 cells and increased sensitivity to IFN-gamma and taxol for apoptosis. Neurochem Res 32:2167-2183, 2007

35. Li AY, Boo LM, Wang SY, et al: Suppression of nonhomologous end joining repair by overexpression of HMGA2. Cancer Res 69:5699-5706, 2009

36. Cleynen I, Van de Ven WJ: The HMGA proteins: A myriad of functions (Review). Int J Oncol 32:289-305, 2008

37. Barber LJ, Youds JL, Ward JD, et al: RTEL1 maintains genomic stability by suppressing homologous recombination. Cell 135:261-271, 2008

38. Little MP, De Vathaire F, Shamsaldin A, et al: Risks of brain tumour following treatment for cancer in childhood: Modification by genetic factors, radiotherapy and chemotherapy. Inter J Cancer 78:269-275, 1998

39. Bondy ML, Wang LE, El-Zein R, et al: Gamma-radiation sensitivity and risk of glioma. J Natl Cancer Inst 93:1553-1557, 2001

40. Neglia JP, Robison LL, Stovall M, et al: New primary neoplasms of the central nervous system in survivors of childhood cancer: A report from the childhood cancer survivor study. J Natl Cancer Inst 98:1528-1537, 2006

41. Mueller JC, Lohmussaar E, Magi R, et al: Linkage disequilibrium patterns and tagSNP transferability among European populations. Am J Hum Genet 76:387-398, 2005

42. Lundmark PE, Liljedahl U, Boomsma DI, et al: Evaluation of HapMap data in six populations of European descent. Eur J Hum Genet 16:1142-1150, 2008

43. Morgan TM, Krumholz HM, Lifton RP, et al: Nonvalidation of reported genetic risk factors for acute coronary syndrome in a large-scale replication study. JAMA 297:1551-1561, 2007

44. Gorroochurn P, Hodge SE, Heiman GA, et al: Non-replication of association studies: "pseudo-failures" to replicate? Genet Med 9:325-331, 2007

45. Clarke GM, Carter KW, Palmer LJ, et al: Fine mapping versus replication in whole-genome association studies. Am J Hum Genet 81:995-1005, 2007

46. Liu YJ, Papasian CJ, Liu JF, et al: Is replication the gold standard for validating genome-wide association findings? PLoS One 3:e4037, 2008

