AAAS WACHTEL PRIZE ESSAY

Moving the needle: Optimizing classification for glioma

Roel G. W. Verhaak[1,2]*

Genomic markers provide unbiased information that is increasingly being used to enhance traditional histopathology approaches for classification of cancer samples.

Diffuse gliomas are the most common type of primary central nervous system (CNS) tumor in adults (*1*), and they have traditionally been classified as oligodendroglioma, astrocytoma, or oligoastrocytoma. World Health Organization (WHO) grade II and grade III oligodendrogliomas vary in mitotic activity and clinical outcome, with 10-year survival rates of 63 and 39%, respectively (*1*). WHO grade II and grade III astrocytomas similarly differ in cellularity and clinical response, with 5-year survival rates of 47 and 27%, respectively (*1*). The oligoastrocytoma category, which traditionally represented 5 to 10% of cases, is increasingly falling out of favor given its low reproducibility. WHO grade IV astrocytoma, known as glioblastoma (GBM), is characterized by necrosis, hypercellularity, and a poor median survival of 15 months (*1*). GBMs represent 55% of the adult diffuse glioma patient population. Median age at diagnosis is 45, 50, and 64 years for oligodendroglioma, astrocytoma, and GBM, respectively (*1*), which means that the number of life years lost because of grade II and grade III gliomas is comparable to grade IV disease. Glioblastomas are standardly treated with surgical resection followed by concomitant radiotherapy and temozolomide. Clinical guidelines for treatment of lower-grade gliomas (LGGs) are less well standardized. Ionizing brain radiation can result in severe cognitive dysfunction, which is a major consideration for patients with relatively favorable prognosis. To overcome the lack of treatment options for glioma, large initiatives such as The Cancer Genome Atlas (TCGA) Consortium were established to perform comprehensive molecular characterization of a patient cohort sufficiently sized to produce statistically meaningful results.

MOLECULAR CHARACTERIZATION OF ADULT DIFFUSE GLIOBLASTOMA

GBM was the first tumor type to be analyzed by TCGA. The first batches of TCGA data were generated around the time when I started a postdoctoral fellowship in the laboratory of M. Meyerson at the Broad Institute/Dana-Farber Cancer Institute. Our group, jointly with other TCGA analysis groups, detected somatic alterations in 206 GBM tumor samples, which showed that most GBMs have defects in three major pathways: (i) p53/stress response, (ii) Rb/cell cycle control, and (iii) receptor tyrosine kinase/Ras/phosphoinositide 3-kinase signal transduction (*2*).

Although our results converged on three established cancer pathways, they also revealed extensive genomic intertumor heterogeneity. During my graduate training, I pioneered the unsupervised analysis of acute myeloid leukemia (AML) transcriptomic profiles. In these studies, we identified AML expression subtypes that coincided with the presence of leukemia gene fusions such as *PML-RAR*, *CFPB-MYH11*, and *AML-ETO* (*3*). Motivated by the results of this work, I applied an iterative consensus clustering method to group 202 GBM samples also included in the genomic analysis and identified four expression-based groups, which we labeled proneural, neural, classical, and mesenchymal. Annotating the subtypes with genomic abnormalities revealed convergence of transcriptional subtype with somatic alterations: *PDGFRA* amplifications and *IDH1* and *TP53* mutations were most frequently found in the proneural group, *EGFR* alterations were found in the classical group, and *NF1* abnormalities were preferentially grouped with mesenchymal GBM (*4*).

After joining the faculty of the University of Texas MD Anderson Cancer Center in 2010, I continued my involvement with TCGA's GBM project. We expanded the molecular profiling data set to 600 GBM samples and used this large cohort to identify frequent mutations in chromatin-organizing genes, find different types of somatic alteration in *EGFR* in nearly 60% of cases, and refine the DNA copy number landscape. Expression classification of the full cohort identified associations between transcriptional subtypes and genomic alterations, including the amplification of *MYC*, *SOX2*, and *CDK4* in proneural tumors and *CCNE1* amplification in the classical subtype. We showed that the temozolomide response marker, *MGMT* methylation, is most accurate in the classical group but bears little sensitivity in the mesenchymal and proneural classes, a finding with potential clinical application. Our results reinforced subtype identity as a fundamental basis for classification of GBM (*5*).

The importance of GBM expression subtyping was shown in a large number of follow-up studies. For example, IDH-mutant GBM samples were universally classified in the proneural group, showed relatively favorable survival and a characteristically distinct DNA hypermethylation profile, and provided evidence that GBMs harboring *IDH1* mutations are secondary tumors, which originated as an LGG (*6*). *Nf1* and *Trp53*/*Nf1* mouse models recapitulating all subtypes and glioma cell line models from all expression classes have been reported, providing evaluation platforms for target discovery and preclinical validation. Expression classes correlate with differences in tumor growth properties as measured by the extent of contrast enhancement and hemodynamic imaging biomarkers, whereas genome markers do not. Response to the U.S. Food and Drug Administration–approved antiangiogenesis inhibitors such as bevacizumab could only be predicted when considering subtype as a covariate (*7*). Since publication in 2010, the expression subtypes we discovered have become a point of reference for the glioma community.

LOWER-GRADE GLIOMA

Motivated by the heterogeneity in pathohistological diagnosis of LGG, the development of a classification method with clinical relevance became top priority for the TCGA LGG working group. The TCGA LGG data set included the multifaceted molecular profiles of 293 grade II and grade III oligodendrogliomas, astrocytomas, and oligoastrocytomas. I directed the working group analysis, which started with clustering of LGG samples using the DNA methylation, DNA copy number mRNA, and microRNA expression platforms. Each of these analyses resulted in subtypes that associated with outcome, histology, and molecular alterations. To identify the common molecular subtypes of LGG, we integrated the results of the

[1]Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [2]Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.
*Corresponding author. Email: rverhaak@mdanderson.org

individual platforms using a clustering of clusters analysis. This approach revealed three subtypes that were consistently found in each data set. Unexpectedly, the three subtypes were summarized by two well-known molecular markers: (i) IDH–wild-type LGG, with a GBM-like median survival of 1.7 years; (ii) IDH-mutant and chromosome 1p/19q co-deleted LGG, with a median survival of 8 years; and (iii) other IDH-mutant LGGs, with a median survival of 6.3 years. The spectrum of genomic abnormalities further separated the three molecular subtypes, with the IDH–wild-type LGG resembling the palette of GBM somatic alterations, the IDH-mutant/1p-19q codeletion group harboring uniform *TERT* promoter mutations and frequent mutations in 1p gene *FUBP1* and 19q gene *CIC*, and the IDH-mutant non-codeletion class, showing no *TERT* promoter mutations but 75% *ATRX* mutations and 95% *TP53* mutations (*8*).

REVISING THE WHO CLASSIFICATION OF ADULT DIFFUSE GLIOMA

Our studies showed that (i) IDH and 1p/19q status improved the classification of grade II/grade III gliomas (*8*), (ii) IDH-mutant GBMs are hypermethylated (*6*), and (iii) IDH–wild-type LGGs are hypomethylated (*8*). Together, our results suggested that the separation of grade II/grade III gliomas from GBM is artificial, and a molecular study across glioma was urgently needed. I, therefore, initiated and supervised the TCGA pan-glioma analysis working group with the goal of analyzing and integrating the molecular and clinical profiles of gliomas across grades and histologies. The final curated data set contained 819 exomes, 932 methylomes, 1084 DNA copy number profiles, and 1045 gene expression profiles. The size of the patient cohort allowed the discovery of previously underappreciated somatic alterations, such as cohesion pathway alterations in 16% of glioma. Analysis of 141 glioma whole-genome sequencing data sets showed increased telomere lengths in *ATRX* mutant cases, which are predominantly IDH-mutant LGG with unaltered 1p/19q.

This data set allowed us to address longstanding questions such as whether IDH-mutant GBMs are molecularly similar to IDH-mutant LGG and whether IDH–wild-type LGG should be classified as GBM. Unsupervised grouping of glioma samples by DNA methylation profiles provided a clear answer: Tumor grade and histology had little effect on DNA methylation pattern. Instead, IDH mutation status and, secondarily, 1p/19q codeletion status were the main drivers of the clustering. Within the group of IDH-mutant gliomas, we identified a 6% subset characterized by a demethylator phenotype, which was associated with poor outcome. However, 5 of 10 recurrent gliomas similarly carried the demethylator pattern, suggesting that demethylation is a marker of glioma progression. Among IDH–wild-type gliomas, we detected a subset with a DNA methylation pattern reminiscent of pilocytic astrocytomas (PAs), which are pediatric grade I gliomas. The somatic abnormality spectrum of the PA subset showed enrichment for *BRAF* and other PA-associated genes, but independent re-review of the hematoxylin and eosin–stained pathology slides confirmed a diagnosis of adult grade II/grade III glioma. Despite their IDH–wild-type status, the PA-like subtype was associated with 5-year

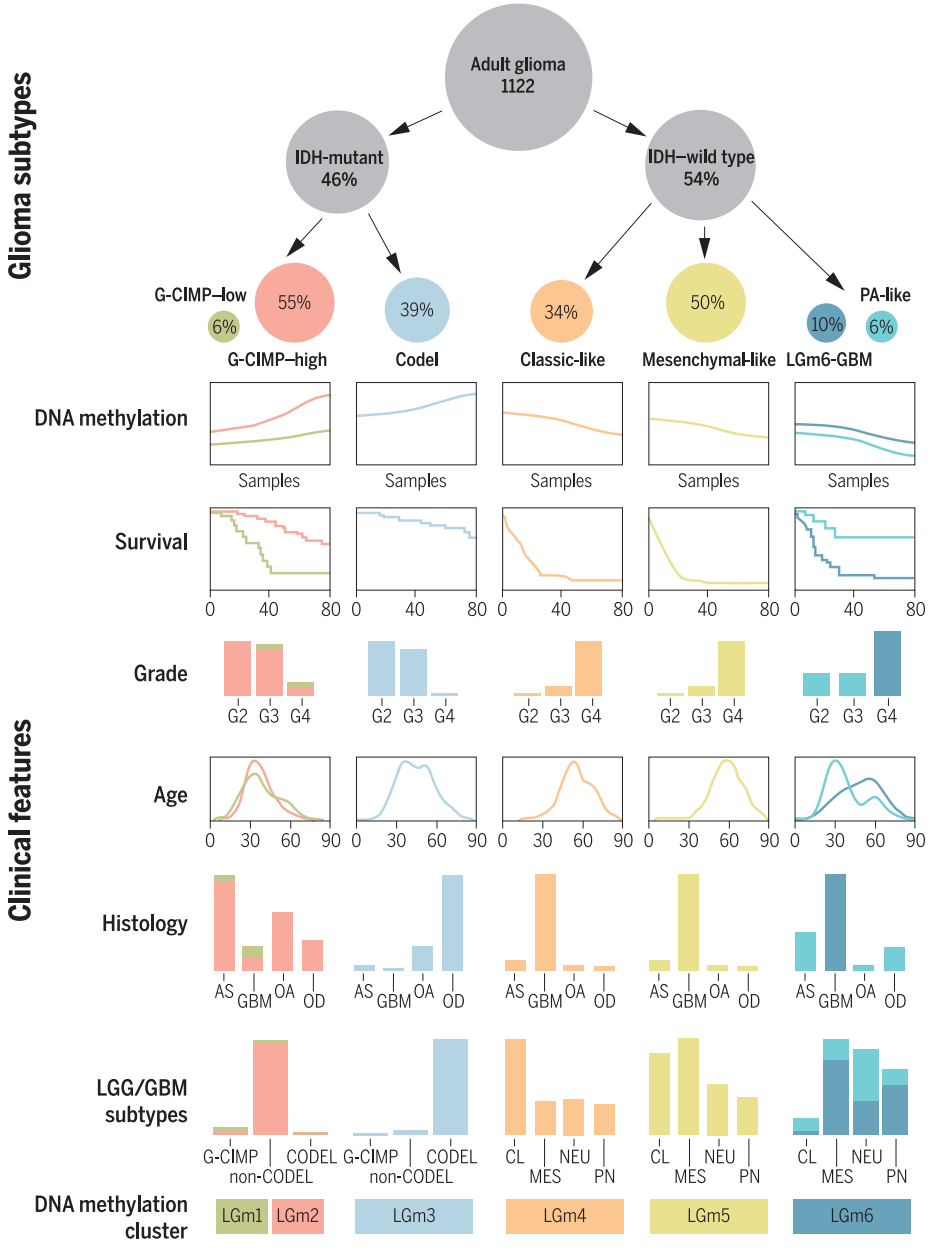

Fig. 1. Subtypes of adult diffuse glioma. Adult gliomas from 1122 patients were grouped into seven subtypes characterized by different patterns of DNA methylation, histology, and clinical characteristics [shown as cartoon representations; adapted from (*9*)]. Survival is shown in months, and age is shown in years. AS, astrocytoma; OA, oligoastrocytoma; OD, oligodendroglioma; CL, classical; MES, mesenchymal; NEU, neural; PN, proneural; CODEL, codeletion; G-CIMP, glioblastoma with a hypermethylation phenotype.

CREDIT: C. BICKEL/*SCIENCE TRANSLATIONAL MEDICINE*

survival rates of more than 50%. Statistical analysis showed that a multivariate model considering grade, age, and DNA methylation subtype provided the optimal predictive model of patient outcome. Overall, our analysis identified seven subtypes of adult diffuse glioma, which differed in DNA methylation profile, somatic alteration spectrum, and patient characteristics including age and survival (Fig. 1) (*9*).

The molecular subtypes we distilled from the seven combinations of histology and grade lay the foundation for a reproducible and clinically relevant classification that incorporates molecular data into the pathological diagnosis. Results from us and others have motivated the 2016 revision of the WHO classification of CNS tumors, which now includes molecular markers in addition to pathohistology (*10*). Optimal clinical classification saves patients from unnecessary side effects and motivates clinicians to provide the most aggressive therapy for those who need it. Our study marked the final chapter of the TCGA glioma group and a personal journey that has forever changed the research landscape of adult glioma.

REFERENCES AND NOTES

1. Q. T. Ostrom, H. Gittleman, J. Fulop, M. Liu, R. Blanda, C. Kromer, Y. Wolinsky, C. Kruchko, J. S. Barnholtz-Sloan, CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2008-2012. *Neuro. Oncol.* **17**, iv1–iv62 (2015).
2. Cancer Genome Atlas Research Network, Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
3. P. J. M. Valk, R. G. W. Verhaak, M. A. Beijen, C. A. J. Erpelinck, S. B. van Waalwijk van Doorn-Khosrovani, J. M. Boer, H. B. Beverloo, M. J. Moorhouse, P. J. van der Spek, B. Löwenberg, R. Delwel, Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* **350**, 1617–1628 (2004).
4. R. G. W. Verhaak, K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkenson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B. A. Weir, S. Gabriel, W. Winckler, S. Gupta, L. Jakkula, H. S. Feiler, J. G. Hodgson, C. D. James, J. N. Sarkaria, C. Brennan, A. Kahn, P. T. Spellman, R. K. Wilson, T. P. Speed, J. W. Gray, M. Meyerson, G. Getz, C. M. Perou, D. N. Hayes; The Cancer Genome Atlas Research Network, Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *PDGFRA*, *IDH1*, *EGFR*, and *NF1*. *Cancer Cell* **17**, 98–110 (2010).
5. C. W. Brennan, R. G. Verhaak, A. McKenna, B. Campos, H. Noushemhr, S. R. Salama, S. Zheng, D. Chakravarty, J. Z. Sanborn, S. H. Berman, R. Beroukhim, B. Bernard, C.-J. Wu, G. Genovese, I. Shmulevich, J. Barnholtz-Sloan, L. Zou, R. Vegesna, S. A. Shukla, G. Ciriello, W. K. Yung, W. Zhang, C. Sougnez, T. Mikkelsen, K. Aldape, D. D. Bigner, E. G. Van Meir, M. Prados, A. Sloan, K. L. Black, J. Eschbacher, G. Finocchiaro, W. Friedman, D. W. Andrews, A. Guha, M. Iacocca, B. P. O'Neil, G. Foltz, J. Myers, D. J. Weisenberger, R. Penny, R. Kucherlapati, C. M. Perou, D. N. Hayes, R. Gibbs, M. Marra, G. B. Mills, E. Lander, P. Spellman, R. Wilson, C. Sander, J. Weinstein, M. Meyerson, S. Gabriel, P. W. Laird, D. Haussler, G. Getz, L. Chin; TCGA Research Network, The somatic genomic landscape of glioblastoma. *Cell* **155**, 462–477 (2013).
6. H. Noushmehr, D. J. Weisenberger, K. Diefes, H. S. Phillips, K. Pujara, B. P. Berman, F. Pan, C. E. Pelloski, E. P. Sulman, K. P. Bhat, R. G. W. Verhaak, K. A. Hoadley, D. N. Hayes, C. M. Perou, H. K. Schmidt, L. Ding, R. K. Wilson, D. Van Den Berg, H. Shen, H. Bengtsson, P. Neuvial, L. M. Cope, J. Buckley, J. G. Herman, S. B. Baylin, P. W. Laird, K. Aldape; The Cancer Atlas Research Network, Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* **17**, 510–522 (2010).
7. E. P. Sulman, M. Won, D. T. Blumenthal, M. A. Vogelbaum, H. Colman, R. B. Jenkins, A. Chakravarti, R. Jeraj, P. D. Brown, K. A. Jaeckle, D. Schiff, J. Dignam, J. N. Atkins, D. Brachman, M. Werner-Wasik, R. Komaki, M. R. Gilbert, M. P. Mehta, K. D. Aldape, Molecular predictors of outcome and response to bevacizumab (BEV) based on analysis of RTOG 0825, a phase III trial comparing chemoradiation (CRT) with and without BEV in patients with newly diagnosed glioblastoma (GBM). *J. Clin. Oncol.* **31**, LBA2010 (2013).
8. The Cancer Genome Atlas Research Network, D. J. Brat, R. G. Verhaak, K. D. Aldape, W. K. Yung, S. R. Salama, L. A. Cooper, E. Rheinbay, C. R. Miller, M. Vitucci, O. Morozova, A. G. Robertson, H. Noushmehr, P. W. Laird, A. D. Cherniack, R. Akbani, J. T. Huse, G. Ciriello, L. M. Poisson, J. S. Barnholtz-Sloan, M. S. Berger, C. Brennan, R. R. Colen, H. Colman, A. E. Flanders, C. Giannini, M. Grifford, A. Iavarone, R. Jain, I. Joseph, J. Kim, K. Kasaian, T. Mikkelsen, B. A. Murray, B. P. O'Neill, L. Pachter, D. W. Parsons, C. Sougnez, E. P. Sulman, S. R. Vandenberg, E. G. Van Meir, A. von Deimling, H. Zhang, D. Crain, K. Lau, D. Mallery, S. Morris, J. Paulauskis, R. Penny, T. Shelton, M. Sherman, P. Yena, A. Black, J. Bowen, K. Dicostanzo, J. Gastier-Foster, K. M. Leraas, T. M. Lichtenberg, C. R. Pierson, N. C. Ramirez, C. Taylor, S. Weaver, L. Wise, E. Zmuda, T. Davidsen, J. A. Demchok, G. Eley, M. L. Ferguson, C. M. Hutter, K. R. Mills Shaw, B. A. Ozenberger, M. Sheth, H. J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J. C. Zenklusen, B. Ayala, J. Baboud, S. Chudamani, M. A. Jensen, J. Liu, T. Pihl, R. Raman, Y. Wan, Y. Wu, A. Ally, J. T. Auman, M. Balasundaram, S. Balu, S. B. Baylin, R. Beroukhim, M. S. Bootwalla, R. Bowlby, C. A. Bristow, D. Brooks, Y. Butterfield, R. Carlsen, S. Carter, L. Chin, A. Chu, E. Chuah, K. Cibulskis, A. Clarke, S. G. Coetzee, N. Dhalla, T. Fennell, S. Fisher, S. Gabriel, G. Getz, R. Gibbs, R. Guin, A. Hadjipanayis, D. N. Hayes, T. Hinoue, K. Hoadley, R. A. Holt, A. P. Hoyle, S. R. Jefferys, S. Jones, C. D. Jones, R. Kucherlapati, P. H. Lai, E. Lander, S. Lee, L. Lichtenstein, Y. Ma, D. T. Maglinte, H. S. Mahadeshwar, M. A. Marra, M. Mayo, S. Meng, M. L. Meyerson, P. A. Mieczkowski, R. A. Moore, L. E. Mose, A. J. Mungall, A. Pantazi, M. Parfenov, P. J. Park, J. S. Parker, C. M. Perou, A. Protopopov, X. Ren, J. Roach, T. S. Sabedot, J. Schein, S. E. Schumacher, J. G. Seidman, S. Seth, H. Shen, J. V. Simons, P. Sipahimalani, M. G. Soloway, X. Song, H. Sun, B. Tabak, A. Tam, D. Tan, J. Tang, N. Thiessen, T. Triche Jr., D. J. Van Den Berg, U. Veluvolu, S. Waring, D. J. Weisenberger, M. D. Wilkerson, T. Wong, J. Wu, L. Xi, A. W. Xu, L. Yang, T. I. Zack, J. Zhang, B. A. Aksoy, H. Arachchi, C. Benz, B. Bernard, D. Carlin, J. Cho, D. DiCara, S. Frazer, G. N. Fuller, J. Gao, N. Gehlenborg, D. Haussler, D. I. Heiman, L. Iype, A. Jacobsen, Z. Ju, S. Katzman, H. Kim, T. Knijnenburg, R. B. Kreisberg, M. S. Lawrence, W. Lee, K. Leinonen, P. Lin, S. Ling, W. Liu, Y. Liu, Y. Liu, Y. Lu, G. Mills, S. Ng, M. S. Noble, E. Paull, A. Rao, S. Reynolds, G. Saksena, Z. Sanborn, C. Sander, N. Schultz, Y. Senbabaoglu, R. Shen, I. Shmulevich, R. Sinha, J. Stuart, S. O. Sumer, Y. Sun, N. Tasman, B. S. Taylor, D. Voet, N. Weinhold, J. N. Weinstein, D. Yang, K. Yoshihara, S. Zheng, W. Zhang, L. Zou, T. Abel, S. Sadeghi, M. L. Cohen, J. Eschbacher, E. M. Hattab, A. Raghunathan, M. J. Schniederjan, D. Aziz, G. Barnett, W. Barrett, D. D. Bigner, L. Boice, C. Brewer, C. Calatozzolo, B. Campos, C. G. Carlotti Jr., T. A. Chan, L. Cuppini, E. Curley, S. Cuzzubbo, K. Devine, F. DiMeco, R. Duell, J. B. Elder, A. Fehrenbach, G. Finocchiaro, W. Friedman, J. Fulop, J. Gardner, B. Hermes, C. Herold-Mende, C. Jungk, A. Kendler, N. L. Lehman, E. Lipp, O. Liu, R. Mandt, M. McGraw, R. Mclendon, C. McPherson, L. Neder, P. Nguyen, A. Noss, R. Nunziata, Q. T. Ostrom, C. Palmer, A. Perin, B. Pollo, A. Potapov, O. Potapova, W. K. Rathmell, D. Rotin, L. Scarpace, C. Schilero, K. Senecal, K. Shimmel, V. Shurkhay, S. Sifri, R. Singh, A. E. Sloan, K. Smolenski, S. M. Staugaitis, R. Steele, L. Thorne, D. P. Tirapelli, A. Unterberg, M. Vallurupalli, Y. Wang, R. Warnick, F. Williams, Y. Wolinsky, S. Bell, M. Rosenberg, C. Stewart, F. Huang, J. L. Grimsby, A. J. Radenbaugh, J. Zhang, Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* **372**, 2481–2498 (2015).
9. M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, A. M. Bradley, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta, S. Anjum, J. Wang, G. Manyam, P. Zoppoli, S. Ling, A. A. Rao, M. Grifford, A. D. Cherniack, H. Zhang, L. Poisson, C. G. Carlotti Jr., D. Pretti da Cunha Tirapelli, A. Rao, T. Mikkelsen, C. C. Lau, W. K. A. Yung, R. Rabadan, J. Huse, D. J. Brat, N. L. Lehman, J. S. Barnholtz-Sloan, S. Zheng, K. Hess, G. Rao, M. Meyerson, R. Beroukhim, L. Cooper, R. Akbani, M. Wrensch, D. Haussler, K. D. Aldape, P. W. Laird, D. H. Gutmann; TCGA Research Network, H. Noushmehr, A. Iavarone, R. G. W. Verhaak, Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
10. D. N. Louis, A. Perry, G. Reifenberger, A. von Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, D. W. Ellison, The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A summary. *Acta Neuropathol.* **131**, 803–820 (2016).

Acknowledgments: I am grateful to R. Wachtel and M. Wachtel and the American Association for the Advancement of Science (AAAS) for awarding me the 2016 AAAS Martin and Rose Wachtel Cancer Research Award. **Funding:** This work was supported by grants to R.G.W.V. from the NIH/National Cancer Institute (R01 CA190121, P01 CA085878, P50 CA127001, P50 CA083639, and P30 CA016672) and the Cancer Prevention and Research Institute of Texas (R140606).

10.1126/scitranslmed.aah4740

Citation: R. G. W. Verhaak, Moving the needle: Optimizing classification for glioma. *Sci. Transl. Med.* **8**, 350fs14 (2016).

Science Translational Medicine

Moving the needle: Optimizing classification for glioma

Roel G. W. Verhaak

Sci Transl Med **8**, 350fs14350fs14.
DOI: 10.1126/scitranslmed.aah4740