appropriate. Results in graphs are expressed as mean $\pm$ s.d. or mean $\pm$ s.e.m. for the indicated number of observations. All statistical analyses were performed and P values were obtained using the GraphPad Prism software 6.0 or the R software (https://www.r-project.org) and are reported in Source Data.

Code availability. A collection of the R procedures to perform ee-MWW is available at http://github.com/miccec/yaGST. The *RGBM* package is available from CRAN at https://cran.r-project.org/web/packages/RGBM/index.html.

Data availability. Transcriptomic microarray gene expression data have been deposited in ArrayExpress with accession number E-MTAB-6037. Source data for western blot are provided in Supplementary Fig. 1. Source Data for Figs 1–4 and Extended Data Figs 1, 2, 4, 7–10 are included in the online version of the paper.

28. Isserlin, R., Merico, D., Voisin, V. & Bader, G. D. Enrichment Map — a Cytoscape app to visualize and explore OMICs pathway enrichment results. *F1000Res.* **3**, 141 (2014).
29. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
30. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
31. Stransky, N., Cerami, E., Schalm, S., Kim, J. L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **5**, 4846 (2014).
32. Yoshihara, K. *et al.* The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* **34**, 4845–4854 (2015).
33. Iyer, M. K., Chinnaiyan, A. M. & Maher, C. A. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics* **27**, 2903–2904 (2011).
34. Risso, D., Schwartz, K., Sherlock, G. & Dudoit, S. GC-content normalization for RNA-seq data. *BMC Bioinformatics* **12**, 480 (2011).
35. Ceccarelli, M. *et al.* Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
36. Lum, P. Y. *et al.* Extracting insights from the shape of complex data using topology. *Sci. Rep.* **3**, 1236 (2013).
37. Roweis, S. T. & Saul, L. K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**, 2323–2326 (2000).
38. Nielson, J. L. *et al.* Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury. *Nat. Commun.* **6**, 8581 (2015).
39. Singh, G., Memoli, F. & Carlsson, G. E. Topological methods for the analysis of high dimensional data sets and 3D object recognition. In *SPBG* (eds Botsch, M. & Pajarola, R.) 91–100 (Citeseer, 2007).
40. Rizvi, A. H. *et al.* Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat. Biotechnol.* **35**, 551–560 (2017).
41. Rolland, D. *et al.* Global phosphoproteomic profiling reveals distinct signatures in B-cell non-Hodgkin lymphomas. *Am. J. Pathol.* **184**, 1331–1342 (2014).
42. Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* **24**, 2534–2536 (2008).
43. Fenyö, D. & Beavis, R. C. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal. Chem.* **75**, 768–774 (2003).
44. MacLean, B., Eng, J. K., Beavis, R. C. & McIntosh, M. General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* **22**, 2830–2832 (2006).
45. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392 (2002).
46. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003).
47. Pedrioli, P. G. Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol. Biol.* **604**, 213–238 (2010).
48. Fermin, D., Basrur, V., Yocum, A. K. & Nesvizhskii, A. I. Abacus: a computational tool for extracting and pre-processing spectral count data for label-free quantitative proteomic analysis. *Proteomics* **11**, 1340–1345 (2011).
49. Beausoleil, S. A., Villén, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **24**, 1285–1292 (2006).
50. Lee, S. B. *et al.* An ID2-dependent mechanism for VHL inactivation in cancer. *Nature* **529**, 172–177 (2016).
51. Sonoda, Y. *et al.* Formation of intracranial tumors by genetically modified human astrocytes defines four pathways critical in the development of human anaplastic astrocytoma. *Cancer Res.* **61**, 4956–4960 (2001).
52. Rutka, J. T. *et al.* Establishment and characterization of five cell lines derived from human malignant gliomas. *Acta Neuropathol.* **75**, 92–103 (1987).
53. Malina, A. *et al.* Adapting CRISPR/Cas9 for functional genomics screens. *Methods Enzymol.* **546**, 193–213 (2014).
54. Ran, F. A. *et al.* Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
55. Zhao, X. *et al.* The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nat. Cell Biol.* **10**, 643–653 (2008).
56. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25**, 402–408 (2001).
57. Ajaz, S., Czajka, A. & Malik, A. Accurate measurement of circulating mitochondrial DNA content from human blood samples using real-time quantitative PCR. *Methods Mol. Biol.* **1264**, 117–131 (2015).
58. Vazquez, F. *et al.* PGC1α expression defines a subset of human melanoma tumors with increased mitochondrial capacity and resistance to oxidative stress. *Cancer Cell* **23**, 287–301 (2013).
59. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
60. Lasorella, A., Rothschild, G., Yokota, Y., Russell, R. G. & Iavarone, A. Id2 mediates tumor initiation, proliferation, and angiogenesis in Rb mutant mice. *Mol. Cell. Biol.* **25**, 3563–3574 (2005).
61. Zhu, S. *et al.* The bHLH repressor Deadpan regulates the self-renewal and specification of *Drosophila* larval neural stem cells independently of Notch. *PLoS ONE* **7**, e46724 (2012).