DNA methylation-based classification of human central nervous system tumours

David Capper*[1,2,98], David T.W. Jones*[3], Martin Sill*[4], Volker Hovestadt*[5+], Daniel Schrimpf[1,2], Dominik Sturm[3,6], Christian Koelsche[1,2], Felix Sahm[1,2], Lukas Chavez[3], David E. Reuß[1,2], Annekathrin Kratz[1,2], Annika K. Wefers[1,2], Kristin Huang[1,2], Kristian W. Pajtler[3,6,7], Leonille Schweizer[1], Damian Stichel[1,2], Adriana Olar[8,9], Nils W. Engel[10,11], Kerstin Lindenberg[2], Patrick N. Harter[12], Anne Braczynski[12], Hildegard Dohmen[29], Boyan K. Garvalov[29], Roland Coras[13], Annett Hölsken[13], Ekkehard Hewer[14], Melanie Bewerunge-Hudler[15], Matthias Schick[15], Roger Fischer[15], Rudi Beschorner[16], Jens Schittenhelm[16], Ori Staszewski[17], Khalida Wani[9], Pascale Varlet[18], Melanie Pages[18], Petra Temming[19], Florian Selt[6,7,20,21], Hendrik Witt[3,6,7], Till Milde[6,7,20,21], Olaf Witt[6,7,20,21], Eleonora Aronica[22], Felice Giangaspero[23], Elisabeth Rushing[24], Wolfram Scheurlen[25], Christoph Geisenberger[26,27], Fausto J. Rodriguez[28], Albert Becker[30], Matthias Preusser[31], Christine Haberler[32], Rolf Bjerkvig[33,101], Jane Cryan[34], Michael Farrell[34], Martina Deckert[35], Jürgen Hench[36], Stephan Frank[36], Jonathan Serrano[37], Kasthuri Kannan[37], Aristotelis Tsirigos[37], Wolfgang Brück[38], Silvia Hofer[39], Stefanie Brehmer[40], Marcel Seiz-Rosenhagen[40], Daniel Hänggi[40], Volkmar Hans[41,42], Stephanie Rozsnoki[43], Jordan R. Hansford[44], Patricia Kohlhof[45], Bjarne W. Kristensen[46], Matthias Lechner[47], Beatriz Lopes[48], Christian Mawrin[49], Ralf Ketter[50], Andreas Kulozik[6], Ziad Khatib[51], Frank Heppner[52,53], Arend Koch[52], Anne Jouvet[54], Catherine Keohane[55], Helmut Mühleisen[56], Wolf Mueller[57], Ute Pohl[58], Marco Prinz[17,59], Axel Benner[4], Marc Zapatka[5], Nicholas G. Gottardo[60,61,62], Pablo Hernáiz Driever[92], Christof M. Kramm[93], Hermann L. Müller[94], Stefan Rutkowski[95], Katja von Hoff[95], Michael C. Frühwald[96], Astrid Gnekow[96], Gudrun Fleischhack[19], Stephan Tippelt[19], Gabriele Calaminus[97], Camelia-Maria Monoranu[63], Arie Perry[64], Chris Jones[65], Thomas S. Jacques[66], Bernhard Radlwimmer[5], Marco Gessi[30], Torsten Pietsch[30], Johannes Schramm[67], Gabriele Schackert[68], Manfred Westphal[69], Guido Reifenberger[70], Pieter Wesseling[71], Michael Weller[72], Vincent Peter Collins[73], Ingmar Blümcke[13], Martin Bendszus[74], Jürgen Debus[75], Annie Huang[76], Nada Jabado[77], Paul A. Northcott[78], Werner Paulus[43], Amar Gajjar[79], Giles Robinson[79], Michael D. Taylor[80], Zane Jaunmuktane[91,99], Marina Ryzhova[81], Michael Platten[82], Andreas Unterberg[26], Wolfgang Wick[83], Matthias Karajannis[84], Michel Mittelbronn[12,100], Till Acker[29], Christian Hartmann[85], Kenneth Aldape[86], Ulrich Schüller[11,87,88], Rolf Buslei[13], Peter Lichter[5], Marcel Kool[3], Christel Herold-Mende[26], David Ellison[89], Martin Hasselblatt[43], Matija Snuderl[90], Sebastian Brandner[91], Andrey Korshunov[1,2], Andreas von Deimling[#1,2] and Stefan M Pfister[#3,6,7]

* co-first authors

# These authors jointly directed this work

Correspondence should be addressed to S.M.P. (s.pfister@dkfz.de) or A.v.D. (Andreas.vonDeimling@med.uni-heidelberg.de)

1 Department of Neuropathology, University Hospital Heidelberg, Heidelberg, Germany

2 Clinical Cooperation Unit Neuropathology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

3 Division of Pediatric Neurooncology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

4 Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg, Germany

5 Division of Molecular Genetics, German Cancer Research Center (DKFZ), Heidelberg, Germany

6 Department of Pediatric Oncology, Hematology and Immunology, University Hospital Heidelberg, Heidelberg, Germany

7 Center for Individualized Pediatric Oncology (ZIPO) and Brain Tumors, Heidelberg, Germany

8 Departments of Pathology and Laboratory Medicine & Neurosurgery, Medical University of South Carolina, Hollings Cancer Center, Charleston, SC, 29425, USA

9 Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

10 Department of Oncology and Hematology with Sections Bone Marrow Transplant and Pneumology, Hubertus Wald Tumorzentrum / University Cancer Center Hamburg, University Medical Center Hamburg , Hamburg, Germany

11 Center for Neuropathology and Prion Research, Ludwig-Maximilians-University, Munich, Germany

12 Institute of Neurology (Edinger Institute), Goethe-University Frankfurt and German Cancer Consortium (DKTK), Partner Site Frankfurt/Mainz, Frankfurt am Main, German Cancer Research Center (DKFZ) Heidelberg, Germany

13 Institute for Neuropathology of the Friedrich-Alexander University of Erlangen-Nürnberg (FAU), Erlangen, Germany

14 Institute of Pathology, University of Bern, Bern, Switzerland

15 Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), Heidelberg, Germany

16 Institute of Pathology and Neuropathology, Department of Neuropathology, University Hospital Tübingen, Tübingen, Germany

17 Institute of Neuropathology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

18 Department of Neuropathology, Centre Hospitalier Sainte Anne, Paris, France

19 Pediatrics III, Pediatric Oncology and Hematology, University Hospital Essen, Essen, Germany

20 CCU Pediatric Oncology (G340) German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), Heidelberg, Germany

21 Translational program, Therapy and Research Center for Pediatric Oncology and Hematology, University Hospital Heidelberg and German Cancer Research Center (DKFZ), Heidelberg, Germany

22 Department of (Neuro)Pathology, Academic Medisch Centrum (AMC) and Swammerdam Institute for Life Sciences, Center for Neuroscience, University of Amsterdam; Stichting Epilepsie Instellingen Nederland (SEIN), The Netherlands

23 Department of Radiological, Oncological and Anatomo-Pathological Sciences, Sapienza University, Rome, Italy and IRCCS Neuromed, Pozzilli (Is), Italy

24 Department of Neuropathology, University Hospital Zurich, Zurich, Switzerland

25 Cnopf'sche Kinderklinik Nuremberg, Nuremberg, Germany

26 Department of Neurosurgery, Heidelberg University Hospital, Heidelberg, Germany

27 Hubrecht Institute-KNAW (Royal Netherlands Academy of Arts and Sciences), Uppsalalaan 8, 3584 CT Utrecht, Netherlands

28 Division of Neuropathology of the Johns Hopkins University School of Medicine, Baltimore, MD, USA

29 Institute of Neuropathology, University of Giessen, Giessen, Germany

30 Department of Neuropathology, University of Bonn, Bonn, Germany

31 Department of Medicine I, Comprehensive Cancer Center Vienna, CNS Unit (CCC-CNS), Medical University of Vienna, Vienna, Austria

32 Institute of Neurology, Medical University of Vienna, Vienna, Austria

33 Department of Biomedicine, University of Bergen, Bergen, Norway

34 Department of Neuropathology, Beaumont Hospital, Dublin, Ireland

35 Department of Neuropathology, University Hospital of Cologne, Cologne, Germany

36 Department of Neuropathology, Institute of Pathology, Basel University Hospital, Basel, Switzerland

37 NYU Langone Medical Center, New York, NY, USA

38 Institute of Neuropathology, University Medical Center Göttingen, Göttingen, Germany

39 Division of Oncology, Luzerner Kantonsspital, Luzern, Switzerland

40 Department of Neurosurgery, University Medical Center Mannheim, University of Heidelberg

41 Institut für Neuropathologie, Evangelisches Krankenhaus Bielefeld gGmbH, Bielefeld, Germany

42 Institut für Neuropathologie, Universitätskinikum Essen, Essen, Germany

43 Institute of Neuropathology, University Hospital Münster, Münster, Germany

44 Children's Cancer Centre, Royal Children's Hospital; Murdoch Children's Research Institute; Department of Pediatrics, University of Melbourne; Melbourne, Australia

45 Institute for Pathology, Katharinenhospital Stuttgart, Stuttgart, Germany

46 Department of Pathology, Odense University Hospital, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

47 University College London Cancer Institute and University College London Hospitals, London, United Kingdom.

48 Department of Pathology, University of Virginia, Charlottesville, Virginia, USA

49 Institute of Neuropathology, Otto-von-Guericke-University, Magdeburg

50 Department of Neurosurgery, University Hospital Saarland, Homburg, Saar, Germany

51 Nicklaus Children's Hospital Brain Institute, Miami, FL, 33155, USA

52 Department of Neuropathology, Charité Universitätsmedizin Berlin, Berlin, Germany

53 Cluster of Excellence, NeuroCure, and Berlin Institute of Health (BIH), Berlin, Germany

54 Département de Pathologie et Neuropathologie, Hôpital Neurologique, Hospices Civils de Lyon, Lyon, France

55 Department of Neuropathology, Cork University Hospital, Cork, Ireland

56 Department of Pathology, Ludwigsburg Hospital, Ludwigsburg, Germany

57 Department of Neuropathology, Leipzig University, Leipzig, Germany

58 Department of Cellular Pathology, Queen's Hospital, Romford, UK

59 BIOSS Centre for Biological Signalling Studies, University of Freiburg, Germany

60 Department of Pediatric Oncology and Haematology, Princess Margaret Hospital for Children, GPO Box D184, Perth, Western Australia 6840, Australia

61 Telethon Kids Institute, University of Western Australia, PO Box 855, Perth, Western Australia 6872, Australia

62 School of Paediatrics and Child Health, University of Western Australia, GPO Box D184, Perth, Western Australia 6840, Australia

63 Department of Neuropathology, Insitute of Pathology, Comprehensive Cancer Center (CCC) Mainfranken, University of Würzburg, Würzburg, Germany

64 Department of Pathology, University of California San Francisco, San Francisco, CA, USA

65 Division of Molecular Pathology, Institute of Cancer Research, London, UK

66 Developmental Biology and Cancer Programme, UCL Great Ormond Street Institute of Child Health and Histopathology Department, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK

67 Medical Faculty, University of Bonn Medical School, Bonn, Germany

68 Department of Neurosurgery, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.

69 Department of Neurosurgery, University Hospital Hamburg-Eppendorf, Hamburg, Germany.

70 Department of Neuropathology, Heinrich Heine University Düsseldorf and German Cancer Consortium (DKTK), partner site Essen/Düsseldorf, German Cancer Research Center (DKFZ) Heidelberg, Germany

71 Department of Pathology, Princess Máxima Center for Pediatric Oncology and University Medical Center Utrecht, Utrecht, the Netherlands, and Department of Pathology, VU University Medical Center, Amsterdam, the Netherlands

72 Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland

73 Department of Pathology, Division of Molecular Histopathology, University of Cambridge, Cambridge, UK.

74 Department of Neuroradiology, Heidelberg University Hospital, Heidelberg, Germany

75 Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany

76 Department of Pediatrics, Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

77 Division of Hematology/Oncology, McGill University, Montreal, QC, Canada

78 Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN, USA

79 Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA

80 Division of Neurosurgery, Arthur and Sonia Labatt Brain Tumor Research Centre, Hospital for Sick Children, University of Toronto, Toronto, Canada

81 NN Burdenko Neurosurgical Institute, Moscow, Russia

82 Department of Neurology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

83 Department of Neurology, Heidelberg University Hospital, Heidelberg, Germany

84 Division of Pediatric Hematology/Oncology, Departments of Pediatrics and Otolaryngology, NYU Langone Medical Center and Perlmutter Cancer Center, New York, NY, USA

85 Department of Neuropathology, Hannover Medical School (MHH), Hannover, Germany

86 Department of Pathology, University of Toronto, Toronto, Ontario, Canada

87 Institute of Neuropathology and Department of Pediatric Hematology and Oncology, University Medical Center, Hamburg-Eppendorf, Germany

88 Research Institute Children's Cancer Center, Hamburg, Germany

89 Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA

90 Division of Neuropathology, Department of Pathology, NYU Langone Medical Center, New York, NY, USA

91 Division of Neuropathology,UCL Hospitlals and Department of Neurodegeneration, Institute of Neurology, University College London, Queen Square, WC1N 3BG, London, UK

92 Department of Pediatric Oncology/Hematology, Charité-Universitätsmedizin Berlin, Berlin, Germany

93 Division of Pediatric Hematology and Oncology, University Medical Center Goettingen, Goettingen, Germany

94 Department of Pediatrics and Pediatric Hematology / Oncology, Klinikum Oldenburg AöR, Medical Campus University Oldenburg, 26133 Oldenburg, Germany

95 Department for Pediatric Hematology and Oncology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

96 Children's Hospital Augsburg, Swabian Children's Cancer Centre, Augsburg, Germany

97 Department of Pediatric Hematology/Oncology, University of Bonn Medical Center, Bonn, Germany

98 Department of Neuropathology, Charité Universitätsmedizin Berlin and German Cancer Consortium (DKTK), Partner Site Berlin, German Cancer Research Center (DKFZ) Heidelberg, Germany
99 Division of Neuropathology, UCL Hospitals and Department of Molecular Neuroscience, Institute of Neurology, University College London, Queen Square, WC1N 3BG, London

100 NORLUX Neuro-Oncology Laboratory, Luxembourg Institute of Health (LIH); Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg; Laboratoire national de santé (LNS), Dudelange; Luxembourg Centre of Neuropathology (LCNP), all Luxembourg.

101 NORLUX Neuro-Oncology Laboratory, Department of Oncology, Luxembourg Institute of Health, Luxembourg.

+ current address: Department of Pathology and Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA, and Broad Institute of MIT and Harvard, Cambridge, MA, USA

Main text

Accurate pathological diagnosis is crucial for optimal clinical management of cancer patients. For the ~100 known brain tumour entities, standardization of the diagnostic process has been shown to be particularly challenging - with substantial inter-observer variability in the histopathological diagnosis and grading of many tumour types [1] [2] [3]. The recent update of the WHO classification of tumours of the central nervous system [4] has added single-gene molecular markers to increase objectivity for some entities, but for the vast majority of classes such markers have not been established. Furthermore, the number of molecularly defined brain tumour classes is currently expanding much more rapidly than the means to reliably diagnose them. A database of the molecular signatures of all known CNS tumours could therefore help to standardise diagnostic procedures. We herein present the development of a comprehensive approach for DNA methylation-based CNS tumour classification across all entities and age groups, and demonstrate its application in a routine diagnostic setting. We show that machine learning algorithms are capable of reliably and robustly differentiating between 82 CNS tumour DNA methylation classes (represented by ~2800 reference samples), including recently defined subgroups not yet covered in official classifications. We further demonstrate that availability of this method may have a substantial impact on diagnostic precision compared with standard methods, resulting in a change of diagnosis in 12% of ~1100 prospective cases analysed at our centre. For broader accessibility we have designed a free and easy to use online classifier tool (www.molecularneuropathology.org) requiring no additional onsite data processing. Our results provide a blueprint for the generation of tumour classifiers across other cancer entities, with the potential to fundamentally transform modern tumour pathology.

The developmental complexity of the brain is reflected in the vast array of distinct brain tumour entities defined in the current WHO classification of CNS tumours [4]. These tumours are clinically and biologically highly diverse, encompassing a wide spectrum from benign neoplasms that can frequently be cured by surgery alone (e.g. pilocytic astrocytoma), to highly malignant tumours responding poorly to any therapy (e.g. glioblastoma). Previous studies reported substantial inter-observer variability in the histopathological diagnosis of many CNS tumours, e.g., in diffuse gliomas [1], ependymomas [2] and supratentorial PNETs [3]. Some molecular subgrouping has been introduced into the update of the WHO classification to partially address this, but only for a few entities, such as medulloblastoma). Furthermore, several single-gene tests based on DNA methylation analysis (e.g., MGMT promoter methylation status), FISH (e.g., 1p/19q, *EGFR, MYC, MYCN, PDGFRA*, 19q13.42, etc.) or immunohistochemistry (CTNNB1, LIN28A, etc.) that are required to cover the most important differential diagnoses have been shown to be difficult to standardize. Such diagnostic discordance and uncertainty confounds decision-making in clinical practice as well as the interpretation and validity of clinical trial results.

The cancer methylome is a combination of both somatically acquired DNA methylation changes and characteristics reflecting the cell of origin [5,6]. The latter property allows, for example, tracing of the primary site of highly dedifferentiated metastases of cancers of unknown origin [7]. It has been convincingly shown that DNA methylation profiling is highly robust and reproducible even from small samples and poor quality material [8], and such profiles have been widely used to subclassify CNS tumours that were previously considered homogeneous diseases [3,9-21]. Based on this preliminary work within single entities, we herein present a comprehensive approach for DNA methylation-based classification of all CNS tumour entities across age groups.

Unsupervised data analysis and definition of 91 DNA methylation classes

To establish a comprehensive CNS tumour reference cohort, we generated genome-wide DNA methylation profiles (minimum of eight cases per group) representing almost all WHO defined neuroectodermal and sellar region tumours [1]. We further profiled mesenchymal tumours, melanoma, diffuse large B-cell lymphoma, plasmacytoma and six types of pituitary adenomas, in total comprising 76 histopathological entities and seven entity variants occurring in the CNS. The DNA methylation cohort was supplemented with data from published series [3,9-24]. All histopathological entities and variants were

analysed by unsupervised clustering both within each entity as well as across histologically similar tumour entities, aiming to identify (i) distinct DNA methylation classes within one histopathological entity and (ii) DNA methylation classes comprising tumours displaying a varied histological phenotype. This iterative process led to the designation of 82 CNS tumour classes characterised by distinct DNA methylation profiles (Figure 1a). Twenty-nine of these were equivalent to a single WHO entity (category 1), 29 represented subclasses within a WHO entity (category 2), in eight the WHO grading was not fully recapitulated (category 3) and in 11 the boundaries of methylation classes were not identical to the entity boundaries of WHO (category 4) (Figure 1a). The remaining five represented DNA methylation classes not defined by the WHO classification (category 5), three of which were recently described [3] as well as the not yet well-defined class of anaplastic pilocytic astrocytoma and one new subclass of infantile hemispheric glioma. There was evidence for several additional classes of rare tumours, with too few cases to be included at present. In consideration of the impact of the tumour microenvironment on the methylation profile, we included 47 tumour samples with a pronounced inflammatory or reactive tumour microenvironment, respectively, both demonstrating distinct methylation profiles. We additionally selected 72 samples representing seven non-neoplastic CNS regions, resulting in a reference cohort of 2,801 samples from 91 classes (Figure 1a) that was visualized using t-SNE dimensionality reduction [25] (Figure 1b) and supported the separation of samples into the defined DNA methylation classes (see also Extended Data Figure 1a, b). Supplementary Information Table 1 gives an overview of methylation class characteristics and Supplementary Information Table 2 shows case-by case information of the reference samples.

To test the stability of separation of methylation classes by t-SNE we performed iterative random down samplings of the reference cohort and observed a high stability of groups irrespective of the exact composition of the cohort (Extended Data Figure 1c, d). To analyse for possible confounding batch effects within our pre-processed reference cohort data set (after adjusting for FFPE versus frozen material) we applied the sva algorithm [26,27]. We found no significant surrogate variables that may have indicated batch effects (data not shown). Assessment of overlayed t-SNE representations of the reference cohort with gender, patient age, material type, array preparation date, tissue source, and tumour purity also did not indicate unexpected associations (Extended Data Figure 2, Extended Data Figure 3a-c, see Extended Data Methods for purity calculation model). For reference astrocytomas, oligodendrogliomas and glioblastomas we additionally performed classification according to the TCGA pan glioma DNA methylation model [28] indicating a strong association of the TCGA classes LGm1-6 with

specific classes defined in our reference cohort (Extended Data Figure 3d, Supplementary Information Table 2).

Classifier development

Application in routine diagnostics requires fast and reproducible classification of samples as well as a measure of confidence for the specific call. To this end, we employed the Random Forest (RF) algorithm [29] to generate 10,000 binary decision trees, incorporating genome-wide information from all 2,801 reference samples and 91 methylation classes (Extended Data Figure 4). Each of these trees assigns a given diagnostic sample to one of the 91 classes, resulting in an aggregate raw score (Figure 2a). To get class probability estimates that can be used to guide diagnostic decision-making, we fitted a multinomial logistic regression calibration model that transforms the raw scores into a probability that measures the confidence in the class assignment ('calibrated score'). The calibration allows comparison of classifier results between classes despite a different raw score distribution from class to class (Extended Data Figure 5a, b). Cross validation of the RF classifier resulted in an estimated error rate of 4.89% for raw and 4.28% for calibrated scores and an area under receiver operating characteristic curve (AUC) of 0.99, indicating a high discriminating power (Figure 2b, c). The vast majority of cross validation misclassifications occurred within eight groups of histologically and biologically closely related tumour classes, distinction of which is currently without clinical impact (with the possible exception of choroid plexus tumours [17]; Figure 2b). We therefore defined eight 'methylation class families' (MCF), for which calibrated scores are summed up to a single score. This reduced the cross-validated error rate for the clinically relevant groupings to 1.14% (Figure 2b, c).

For application to diagnostic tumour samples, a threshold value for the prediction of a matching class is required. Receiver operating characteristic (ROC) analysis on the maximum calibrated scores resulting from cross validation data was performed to estimate possible threshold values (Extended Data Figure 5c, d). A balanced specificity (93.8%) and sensitivity (of 93.4%) was observed at a cut-off of ≥0.836 (Youden index). At a cut-off of ≥0.958, specificity reached 100% but sensitivity was reduced to 82.7%. As a practical compromise, we set our threshold in the middle between both these values (≥0.9). For subclasses within methylation class families, we defined a threshold value of ≥0.5 as sufficient for a valid prediction, as long as all family member scores add up to a total score of ≥0.9. Taking the maximum as prediction threshold and using a multiclass approach [30], overall sensitivity and specificity was 0.989

and 0.999, respectively (Figure 2c). Single class specificity and sensitivity for the ≥0.9 threshold are provided in Supplementary Information Table 3.

Clinical implementation

For evaluation of clinical utility, we prospectively analysed a series of 1,155 diagnostic CNS tumours in parallel with standard histopathological workup (Figure 3a, b). For 51 cases (4%) the material was not suitable for methylation profiling, mostly because of too low tumour cell content or only very limited total material. Methylation profiling was performed for the remaining 1,104 samples and the cases were assigned as either 'matching to a defined DNA methylation class' (calibrated score ≥0.9) or as 'no match' cases (highest score <0.9) (for a case-by-case list see Supplementary Information Table 4). The investigated cases comprised 64 different histopathological entities from both adult (71%) and paediatric patients (29%). The spectrum of entities was enriched for rare and difficult to diagnose cases received for referral, and therefore did not exactly match the distribution seen in daily routine diagnostic practice centres. Histopathological evaluation was performed blinded to DNA methylation profiling results and included standard molecular testing.

In total, 88% of profiled samples (n=977/1,104) matched to an established DNA methylation class with a calibrated classifier score ≥0.9 (Figure 3b). For 838 of these (838/1,104; 76%), results obtained by pathology and DNA methylation profiling were concordant. In 171 of the cases, an unambiguous molecular subgroup could be assigned, which would not have been available based on histopathology evaluation only (e.g., molecular subgroups of medulloblastoma and ependymoma, many of which were included in the latest version of the WHO classification for CNS tumours [4]).

For the remaining 139 samples with a calibrated classifier score ≥0.9, the DNA methylation class was discordant from the pathological diagnosis. These cases were histologically and molecularly re-evaluated, including additional molecular-pathological investigation (DNA copy-number profiling, targeted gene sequencing, gene panel sequencing [31], and gene fusion analysis of a subset of cases, see Supplementary Information Table 5 for case-by-case details). This resulted in a revision of the initial histopathological diagnosis in 129 of the 139 cases (12% of all cases, Figure 3c) in favour of the predicted methylation class. In agreement with a number of recent reports, several of these were IDH wildtype astrocytomas and anaplastic astrocytomas reclassified as IDH wildtype glioblastomas.

Establishing a new diagnosis had a profound clinical impact: A change in WHO grading was observed in 71% of these cases (92/129), with both upgrading (41%, 53/129) and downgrading (30%, 39/129; Figure 3c). Discrepant results could not be resolved in only 10 cases (<1% of profiled cases), and the histopathological diagnosis was retained. To substantiate the impact in clinical practice we contacted five external centres that have started to implement methylation profiling for diagnostic cases using our algorithm. In total, these centres analysed 401 diagnostic cases and in 50 cases (12%) a new diagnosis was established, nearly perfectly recapitulating our rate of reclassification (Extended Data Figure 6a, Supplementary Information Table 6). Analysis of the individual centres indicates that the rate of reclassification varied between 6% and 25% most likely due to differences of the spectrum of investigated cases and more upfront molecular testing by some centres (Extended Data Figure 6b).

Twelve percent of tumours from the prospective cohort (127/1,104) could not be assigned to a DNA methylation class using the rigid calibrated classifier score cut-off of ≥0.9 (Figure 3b). To further clarify the role of these non-classifiable cases we performed an unsupervised t-SNE analysis of the reference cohort together with the diagnostic cohort (Figure 4a). This demonstrated a high overlap of the classifiable cases with the reference cohort, whereas non-classifiable cases frequently fell in the periphery of the reference classes or even completely separate from these and frequently grouped with other non-classifiable cases (Figure 4a). This may indicate that such cases represent rare novel molecular entities that have not been previously recognized. An example for such a likely novel CNS tumour entity is exemplified for one unusual tumour type in Figure 4a, b.

Technical robustness and inter-laboratory comparison

Technical robustness of the RF classifier was investigated by inter-laboratory comparison. Results of two independent laboratories (starting from DNA extraction) were highly correlated, with only two of 51 samples (4%) showing a classifier score slightly lower than 0.9 in one of the centres whereas all other cases were classified identically (Extended Data Figure 7a). Calculation of copy number profiles was also stable across laboratories (Extended Data Figure 7b). To ascertain forward compatibility with developing technologies, we further used the RF classifier to interrogate newer EPIC DNA methylation arrays and high-coverage whole-genome bisulfite sequencing data. For all 27 samples from different CNS tumours profiled on both array platforms, raw scores and calibrated scores were highly correlated

(Extended Data Figure 7c) and running them through the classifying algorithm resulted in the same prediction for every case. Further, for all 20 high-coverage whole-genome bisulfite sequencing samples (11 different CNS tumour entities), the highest prediction score was for the same class as the 450k array, suggesting that our approach is applicable to different DNA methylation profiling techniques with only slight adaptations (Extended Data Figure 7d).

Global dissemination of the platform

To ensure unrestricted community access to our classification system, we created a free web platform for data upload, automatic normalization, Random Forest classification, and PDF report generation (www.molecularneuropathology.org). DNA copy-number profiles [30] and O6-methylguanine-DNA-methyltransferase (*MGMT)* promoter methylation status [31] are additionally provided, since they can be generated from the same data source – thus having the potential of replacing several time- and cost-intensive single-gene tests. A representative website report is shown in Extended Data Figure 8. During upload, the data provider can consent if the data may be used for further classifier development. We expect that this web platform can thereby act as a hub for a worldwide cooperative network to continuously identify and track rare tumour classes so that they can eventually be added to the catalogue of known human cancers. Since the launch of the website in December 2016, over 4500 cases have been uploaded from over 15 participating centres. New biological insights can likely also be gained based on the interrelationships of tumour classes, and by closer examination of how gene methylation affects tumour biology.

In conclusion, we demonstrate that DNA methylation-based CNS tumour classification is a valuable asset for clinical decision making. In contrast to traditional pathology, whereby there is a pressure to assign all tumours to a described entity even for atypical or challenging cases, the objective measure that we provide here allows for 'no match' to a defined class. This information can also be of substantial value in highlighting that a tumor is not a typical example of a given differential diagnosis, and may rather belong to a rarer, yet undefined class. We defined 5 categories of methylation classes that have different clinical implications. Category 1 can be directly translated to WHO entities. Category 2 represent subclasses of WHO entities. For all but ependymal tumor subclasses, currently little clinical implications derive form subclassification and a translation back to the WHO class may be appropriate for clinical

purposes. Category 3 reflects the fact, that WHO grading cannot be fully recapitulated by methylation profiling for several classes. Further data is required to assess if the methylation classes of this category may provide a more robust means of prognostication than histology alone, as has been demonstrated for several other classes [3,11,14]. In category 4, the WHO entity boundaries are not identical to the boundaries of the methylation classes. Until additional data on the exact boundaries become available, this category should be critically discussed in the clinical context and orthogonal testing should be undertaken whenever possible. Category 5 represents putative new entities that are currently not recognized by the WHO and where little data on these cases exists, but where biological rationale for a novel class was considered to be strong.

A study in which reference pathology and molecular diagnostics are blinded for each other´s results is currently ongoing for all childhood brain tumours diagnosed in Germany to objectivise the potential effect of re-classification on patient outcome (http://pediatric-neurooncology.dkfz.de/index.php/en/diagnostics/molecular-neuropathology), with results due over the next few years.

We expect that the principle of using DNA methylation signatures as part of a combined histo-molecular tumor classification will improve diagnostic accuracy not only in neuropathology, but will serve as a blueprint in other fields of tumour pathology. In our experience, adaptation of this technique in diagnostic laboratories is relatively straightforward. Extended Data Figure 9 summarizes a sample workflow for diagnostic implementation.

In particular, the digital nature of methylation data facilitates easy exchange and will allow aggregation of extensive tumor libraries. This will likely result in the detection of exceptionally rare tumor classes and a continued refinement of classifiers. Inclusion of new classes will allow a prompt translation into diagnostic practice, almost certainly resulting in a more dynamic tumor classification. The easy exchange also has great promise for standardization of tumor diagnostics across centres and across clinical trials.

Acknowledgments

We thank Ulrike Lass and Antje Habel from the Department of Neuropathology of the University Hospital Heidelberg for technical support and Iris Oezen for coordination. We thank the Microarray unit of the Genomics and Proteomics Core Facility, German Cancer Research Center (DKFZ), especially Nadja Wermke and Anja Schramm-Glück, for providing excellent methylation services. We thank the German Glioma Network and the Neuroonkologische Arbeitsgemeinschaft (NOA) for data sharing. We thank the DKFZ-Heidelberg Center for Personalized Oncology (DKFZ-HIPO) for technical support and funding through HIPO_036. We also thank the German Childhood Cancer Foundation for funding ("Neuropath 2.0 - Increasing diagnostic accuracy in paediatric neurooncology" (DKS 2015.01)). In other parts, this work was supported by an Illumina Medical Research Grant and the German Cancer Consortium (DKTK) joint funding project 'Next Generation Molecular Diagnostics of Malignant Gliomas'. Additional funding was obtained through the A Kids' Brain Tumour Cure (PLGA) Foundation ("Prospective assessment of predictive/prognostic molecular biomarkers in the SIOP-LGG 201X adaptive phase III clinical trial cohort"), the Friedberg Charitable Foundation, Toronto, Canada (to Matija Snuderl and Matthias Karajannis) and the Stichting Kinderen Kankervrij (KIKA) and Stichting AMC Foundation (to Eleonora Aronica). Dr. Adriana Olar was partially supported by the National Institutes of Health/National Cancer Institute (NIH/NCI) (5T32CA163185). Sebastian Brandner was partially supported by the National Institute of Health Research (NIHR) UCLH/UCL Biomedical Research Centre.

Stefan Pfister, Deimling von Andreas, David Jones, David Capper, Volker Hovestadt, Martin Sill, Melanie Bewerunge-Hudler and Matthias Schick share inventorship of a "DNA-methylation based method for classifying tumor species of the brain". A patent has been applied for this method (EP 3067432 A1). All terms are being managed by the German Cancer Research Center in accordance with its conflict of interest policies. The other authors declare no conflicts of interest.

Author Contributions (CRediT taxonomy)

Conceptualization, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.E., A.v.D., S.M.P.; Methodology, D.C., D.T.W.J., M.Sill, A.Benner, V.Hovestadt, A.v.D., S.M.P.; Software, M.Sill, V.Hovestadt, D.Schrimpf, L.C., D.Stichel, M.Z.; Validation, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, M.Snuderl, A.Korshunov, A.v.D., S.M.P.; Formal Analysis, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, L.C.; Investigation, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Sturm, C.Koelsche, L.C., F.Sahm, D.E.R., A.Kratz, A.K.W., K.H., K.W.P., L.S., K.L., C.G., M.Snuderl, J.Serrano, K.K., A.T.; Resources, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, D.Sturm, L.C., K.W.P., A.O., N.W.E., P.N.H., A.K.B, H.D., B.K.G., R.C., E.H., M.Bewerunge-Hudler, M.Schick, R.F., R.Beschorner, J.Schittenhelm, F.Selt, O.S., K.W., V.P., M.Pages, P.T., H.Witt, E.A., F.G., E.R., W.S., F.J.R., J.H., Z.J., T.A., A.Becker, M.Preusser, C.Haberler, R.Bjerkvig, C.Hartmann, J.C., M.D., S.F., W.B., S.Hofer, S.Brehmer, M.Seiz-Rosenhagen, V.Hans, S.Heim, J.R.H., P.K., B.W.K., M.L., B.L., C.Mawrin, R.K., A.Kulozik, Z.K., F.H., A.Koch, A.Jouvet, C.Keohane, T.M., A.Hölsken, H.Mühleisen, W.M., U.P., M.Prinz, M.F., N.G., C.Monoranu, A.P., C.J., T.S.J., B.R., M.G., T.P., G.R., P.W., M.M., M.Weller, O.W., V.P.C, I.B., M.Bendszus, J.D., A.Huang, N.J., P.A.N., W.P., A.Gajjar, G.W.R., M.D.T., M.R., D.H., M.Platten, W.W., A.U., M.Snuderl, M.Karajannis, K.A., D.E., R.Buslei, P.L., M.Kool, M.H., A.Korshunov, A.v.D., U.S., S.Brandner, C.Herold-Mende, S.M.P., J.Schramm, G.S., M.Westphal, S.R., K.v.H., C.M.K, M.C.F, A.Gnekow, P.H., G.F., S.T., H.L.M, G.C.; Data Curation, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, D.Sturm, L.C., D.Stichel, A.O., M.Kool, A.Korshunov, T.A. S.Brandner, A.v.D., S.M.P.; Writing – Original Draft, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, D.Sturm, C.Koelsche, G.R., P.W., A.v.D., S.M.P.; Visualization, D.C., D.T.W.J., M.Sill, V.Hovestadt, D.Schrimpf, D.Sturm, A.v.D., S.M.P.; Supervision, A.v.D., S.M.P.; Project Administration, D.C., D.T.W.J., S.M.P.; Funding Acquisition, D.C., D.T.W.J., D.Sturm, A.v.D., S.M.P., P.L., C.Herold-Mende, M.Bendszus, J.D., W.W., G.R., M.Snuderl, M.Karajannis, E.A., A.O.

References

1 van den Bent, M. J. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathol.* **120**, 297-304, doi:10.1007/s00401-010-0725-7 (2010).

2 Ellison, D. W. *et al.* Histopathological grading of pediatric ependymoma: reproducibility and clinical relevance in European trial cohorts. *J Negat Results Biomed* **10**, 7, doi:10.1186/1477-5751-10-7 (2011).

3 Sturm, D. *et al.* New Brain Tumor Entities Emerge from Molecular Classification of CNS-PNETs. *Cell* **164**, 1060-1072, doi:10.1016/j.cell.2016.01.015 (2016).

4 Louis, D. N., Ohgaki, H., Wiestler, O. D. & Cavenee, W. K. *WHO Classification of Tumours of the Central Nervous System (revised 4th edition)*. (IARC, 2016).

5 Fernandez, A. F. *et al.* A DNA methylation fingerprint of 1628 human samples. *Genome Res.* **22**, 407-419, doi:10.1101/gr.119867.110 (2012).

6 Hovestadt, V. *et al.* Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510**, 537-541, doi:10.1038/nature13268 (2014).

7 Moran, S. *et al.* Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol.* **17**, 1386-1395, doi:10.1016/S1470-2045(16)30297-2 (2016).

8 Hovestadt, V. *et al.* Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays. *Acta Neuropathol.* **125**, 913-916, doi:10.1007/s00401-013-1126-5 (2013).

9 Korshunov, A. *et al.* Histologically distinct neuroepithelial tumors with histone 3 G34 mutation are molecularly similar and comprise a single nosologic entity. *Acta Neuropathol.* **131**, 137-146, doi:10.1007/s00401-015-1493-1 (2016).

10 Korshunov, A. *et al.* Embryonal tumor with abundant neuropil and true rosettes (ETANTR), ependymoblastoma, and medulloepithelioma share molecular similarity and comprise a single clinicopathological entity. *Acta Neuropathol.* **128**, 279-289, doi:10.1007/s00401-013-1228-0 (2014).

11 Sturm, D. *et al.* Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* **22**, 425-437, doi:10.1016/j.ccr.2012.08.024 (2012).

12 Reuss, D. E. *et al.* Adult IDH wild type astrocytomas biologically and clinically resolve into other tumor entities. *Acta Neuropathol.* **130**, 407-417, doi:10.1007/s00401-015-1454-8 (2015).

13 Holsken, A. *et al.* Adamantinomatous and papillary craniopharyngiomas are characterized by distinct epigenomic as well as mutational and transcriptomic profiles. *Acta Neuropathol Commun* **4**, 20, doi:10.1186/s40478-016-0287-6 (2016).

14 Pajtler, K. W. *et al.* Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell* **27**, 728-743, doi:10.1016/j.ccell.2015.04.002 (2015).

15 Lambert, S. R. *et al.* Differential expression and methylation of brain developmental genes define location-specific subsets of pilocytic astrocytoma. *Acta Neuropathol.* **126**, 291-301, doi:10.1007/s00401-013-1124-7 (2013).

16 Heim, S. *et al.* Papillary Tumor of the Pineal Region: A Distinct Molecular Entity. *Brain Pathol.* **26**, 199-205, doi:10.1111/bpa.12282 (2016).

17 Thomas, C. *et al.* Methylation profiling of choroid plexus tumors reveals 3 clinically distinct subgroups. *Neuro Oncol.* **18**, 790-796, doi:10.1093/neuonc/nov322 (2016).

18 Koelsche, C. *et al.* Melanotic tumors of the nervous system are characterized by distinct mutational, chromosomal and epigenomic profiles. *Brain Pathol.* **25**, 202-208, doi:10.1111/bpa.12228 (2015).

19 Mack, S. C. *et al.* Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature* **506**, 445-450, doi:10.1038/nature13108 (2014).

20 Johann, P. D. *et al.* Atypical Teratoid/Rhabdoid Tumors Are Comprised of Three Epigenetic Subgroups with Distinct Enhancer Landscapes. *Cancer Cell* **29**, 379-393, doi:10.1016/j.ccell.2016.02.001 (2016).
21 Wiestler, B. *et al.* Integrated DNA methylation and copy-number profiling identify three clinically and biologically relevant groups of anaplastic glioma. *Acta Neuropathol.* **128**, 561-571, doi:10.1007/s00401-014-1315-x (2014).
22 Jones, D. T. *et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat. Genet.* **45**, 927-932, doi:10.1038/ng.2682 (2013).
23 Jones, D. T. *et al.* Dissecting the genomic complexity underlying medulloblastoma. *Nature* **488**, 100-105, doi:10.1038/nature11284 (2012).
24 Pietsch, T. *et al.* Prognostic significance of clinical, histopathological, and molecular characteristics of medulloblastomas in the prospective HIT2000 multicenter clinical trial cohort. *Acta Neuropathol.* **128**, 137-149, doi:10.1007/s00401-014-1276-0 (2014).
25 van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *The Journal of Machine Learning Research* **9**, 85 (2008).
26 Leek, J. T. & Storey, J. D. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS genetics* **3**, 1724-1735, doi:10.1371/journal.pgen.0030161 (2007).
27 Leek, J. T. & Storey, J. D. A general framework for multiple testing dependence. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18718-18723, doi:10.1073/pnas.0808709105 (2008).
28 Ceccarelli, M. *et al.* Molecular Profiling Reveals Biologically Discrete Subsets and Pathways of Progression in Diffuse Glioma. *Cell* **164**, 550-563, doi:10.1016/j.cell.2015.12.028 (2016).
29 Breiman, L. Random forests. *Machine learning* **45**, 5-32 (2001).
30 Sokolova, M. & Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manage.* **45**, 427-437, doi:10.1016/j.ipm.2009.03.002 (2009).
31 Sahm, F. *et al.* Next-generation sequencing in routine brain tumor diagnostics enables an integrated diagnosis and identifies actionable targets. *Acta Neuropathol.* **131**, 903-910, doi:10.1007/s00401-015-1519-8 (2016).