CORRESPONDENCE

Robust molecular subgrouping and copy-number profiling of medulloblastoma from small amounts of archival tumour material using high-density DNA methylation arrays

Volker Hovestadt · Marc Remke · Marcel Kool · Torsten Pietsch · Paul A. Northcott · Roger Fischer · Florence M. G. Cavalli · Vijay Ramaswamy · Marc Zapatka · Guido Reifenberger · Stefan Rutkowski · Matthias Schick · Melanie Bewerunge-Hudler · Andrey Korshunov · Peter Lichter · Michael D. Taylor · Stefan M. Pfister · David T. W. Jones

It is now clear that medulloblastoma (MB), one of the most clinically challenging paediatric brain tumours, is not a single disease entity. Rather, it comprises four distinct molecular subgroups (Wnt pathway activated (WNT), Sonic hedgehog pathway activated (SHH), and the less well-characterised Group 3 and Group 4) [7, 15], which are highly divergent in terms of their patient demographics, underlying biology, and survival outcomes [4, 6]. These subgroups

are becoming increasingly important, not only for refining the discovery of prognostic markers or therapeutic targets, but also for the design of prospective clinical trials. Patients with WNT subgroup tumours, for example, generally have a favourable prognosis, and may benefit from a reduction or omission of radiotherapy or chemotherapy to spare neurological side-effects or other toxicities, as is now being prospectively tested in upcoming trials both in North America and Europe. In contrast, patients with poor prognosis Group 3 tumours may benefit from intensification of up-front therapy. Furthermore, many new targeted therapeutics are likely to be efficacious in only one subgroup, such as smoothened

V. Hovestadt · M. Zapatka · P. Lichter
Division of Molecular Genetics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

M. Remke · F. M. G. Cavalli · V. Ramaswamy · M. D. Taylor
Program in Developmental and Stem Cell Biology and
The Arthur and Sonja Labatt Brain Tumour Research Centre,
The Hospital for Sick Children, 555 University Avenue,
Toronto, ON, Canada

M. Remke · V. Ramaswamy · M. D. Taylor
Department of Laboratory Medicine and Pathobiology,
University of Toronto, Toronto, ON, Canada

M. Remke · V. Ramaswamy · M. D. Taylor
Division of Neurosurgery, The Hospital for Sick Children,
555 University Avenue, Toronto, ON, Canada

M. Kool · P. A. Northcott · S. M. Pfister · D. T. W. Jones (✉)
Division of Pediatric Neurooncology, German Cancer Research
Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg,
Germany
e-mail: davidjones@cantab.net

T. Pietsch
Department of Neuropathology, University of Bonn Medical
Centre, Sigmund-Freud-Strasse 25, 53105 Bonn, Germany

R. Fischer · M. Schick · M. Bewerunge-Hudler
Genomics and Proteomics Core Facility, German Cancer
Research Center (DKFZ), Im Neuenheimer Feld 280,
69120 Heidelberg, Germany

G. Reifenberger
Department of Neuropathology, Heinrich-Heine-University
Düsseldorf, Moorenstrasse 5, 40225 Düsseldorf, Germany

S. Rutkowski
Department of Pediatric Hematology and Oncology, University
Medical Center Hamburg-Eppendorf, Martinistrasse 52,
20246 Hamburg, Germany

A. Korshunov
Department of Neuropathology, University of Heidelberg,
Im Neuenheimer Feld 220, 69120 Heidelberg, Germany

A. Korshunov
Clinical Cooperation Unit Neuropathology, German Cancer
Research Center (DKFZ), Im Neuenheimer Feld 280,
69120 Heidelberg, Germany

S. M. Pfister
Department of Pediatric Oncology, Hematology and
Immunology, Heidelberg University Hospital, Im Neuenheimer
Feld 430, 69120 Heidelberg, Germany

inhibitors for SHH pathway-driven MB [1, 2]. A phase III clinical trial randomising SMO inhibition against standard of care in relapsed SHH-MB patients will start recruiting in mid-late 2013. A method for accurate and robust classification into tumour subgroups that is applicable to standard pathology specimens is therefore of key clinical relevance.

The MB subgroups were originally defined based on gene expression profiling from fresh-frozen tumour material [7]. Whilst there are methods to apply such an RNA-based analysis to formalin-fixed paraffin-embedded (FFPE) material, classification accuracy is inferior to that obtained with frozen tissue, particularly when analysing older samples [9]. Furthermore, the use of immunohistochemistry as an alternative subgrouping method [7] has proved difficult to standardise across multiple neuropathology laboratories. The use of a DNA-based platform for subgrouping has clear advantages due to the superior stability of DNA compared with RNA. Methylation profiling has recently been applied for the subgrouping of large series of, for example, glioblastoma and chronic lymphocytic leukaemia samples [5, 10, 14]. It has also been proposed as being suitable for medulloblastoma subclassification, although the older Illumina Golden-Gate platform assessed only a limited subset of genes, and a proportion of samples remained unclassifiable [12]. Also, whilst the concordance between methylation and expression reported by Schwalbe et al. was fairly good (81.5 %), some WNT and SHH-subgroup tumours were misclassified—a clinically important distinction for forthcoming trials. We therefore applied the Illumina Infinium HumanMethylation450 BeadChip array (450k array) to generate genome-wide methylation profiles of a large series of medulloblastoma samples (see Supplementary Methods). The first cohort comprised 107 frozen MB samples collected within the ICGC PedBrain Tumor Project (Heidelberg cohort) [3]. Of these, 86 had matching Affymetrix U133 plus 2.0 expression array data, allowing for a direct comparison between the subgroup classifications of the two methods. Unsupervised k-means consensus clustering on all CpG probes with a standard deviation >0.25 ($n = 21,092$) clearly indicated the presence of four subgroups (Fig. 1a). These methylation subgroups very closely recapitulated the gene expression subgroups of the matching tumours (95.3 % concordance, Rand index = 0.86, $p < 0.0001$; Fig. 1b, c). As expected from previous subgrouping studies, the discordant cases involved switches between Group 3 and Group 4, while the WNT and SHH groups were clearly distinct (Fig. 1c, d).
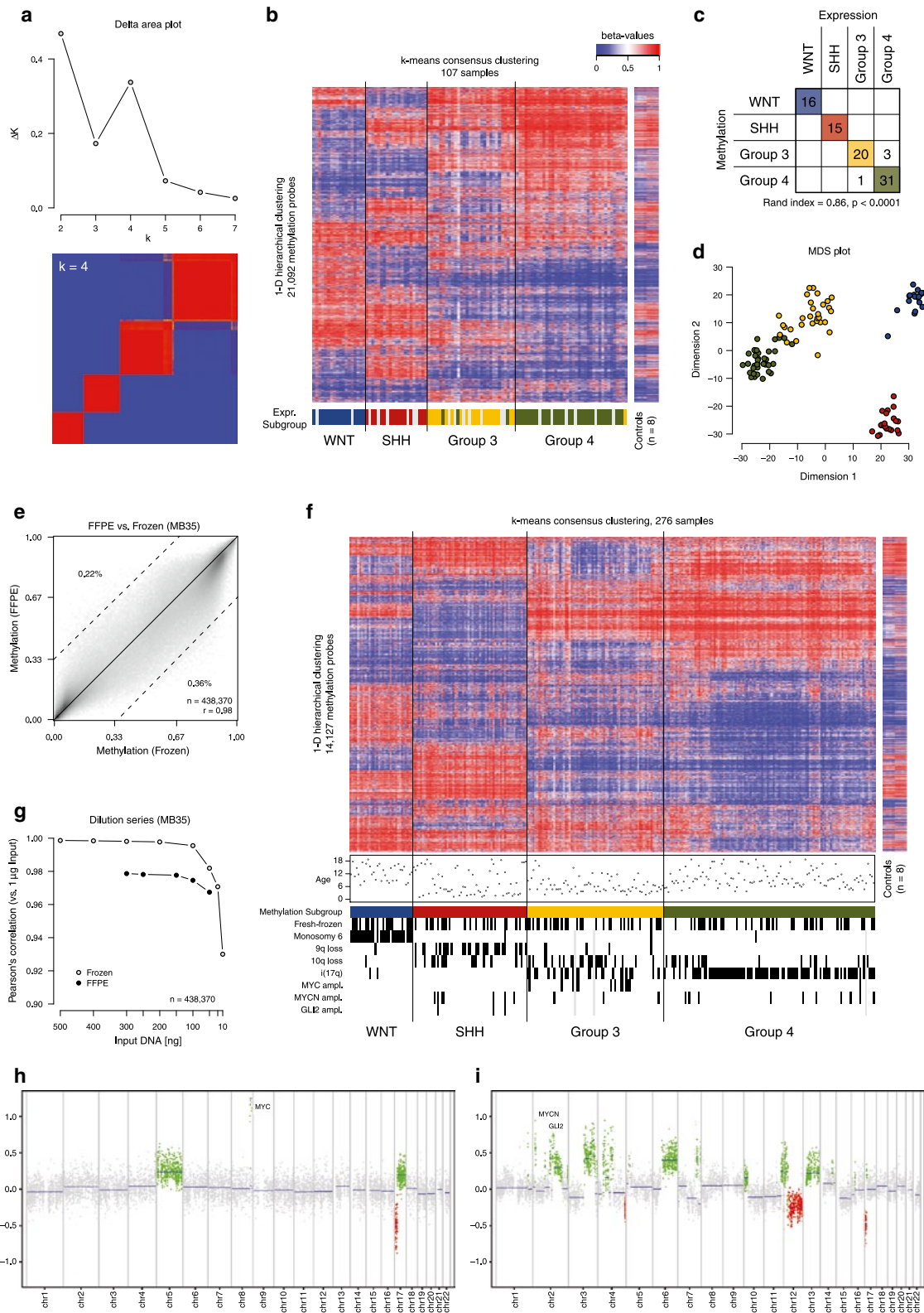
The 450k array is also suitable for analysis of DNA from FFPE material. Profiling of the same tumour from both frozen and FFPE material ($n = 3$) showed a higher correlation than the maximum correlation between different donors, indicating the robustness of this assay for archival tissue (mean Pearson's correlation for paired samples, $r = 0.987$, maximum correlation for any two unpaired samples,

Fig. 1 a k-means consensus clustering of the Heidelberg fresh-frozen cohort ($n = 107$) using the 21,092 most variable CpG probes (SD > 0.25) indicates the presence of four major subgroups in the DNA methylation data. **b** Heatmap of DNA methylation values within the four subgroups derived from the consensus clustering. The gene expression subgroup of the matched samples is indicated below the heatmap. Eight normal cerebellum controls are also shown for comparison. **c** Concordance chart of the gene expression versus DNA methylation-derived subgroups for each sample. **d** Multidimensional scaling (MDS) analysis of the same samples and same CpG probes used for the consensus clustering. **e** Correlation of DNA methylation values derived from fresh-frozen and FFPE material from a single tumour sample. **f** Heatmap of DNA methylation values across a combined set of the Heidelberg fresh-frozen cohort and the FFPE tumour cohort ($n = 276$). Patient age and copy-number events derived from the 450k array data are indicated below the heatmap. **g** Correlation of DNA methylation values from a dilution series of fresh-frozen and FFPE tumour DNA from a single sample. **h** Copy-number plot of a Group 3 medulloblastoma from the FFPE series, showing stereotypic *MYC* amplification and i(17q). **i** Copy-number plot of an SHH medulloblastoma from the FFPE series displaying evidence of dramatic structural changes, reminiscent of chromothripsis

$r = 0.975$; Fig. 1e). We therefore profiled an independent set of 169 FFPE MB samples using the methylation array. These were confirmed as unique samples using SNP genotyping probes from the array platform, which allows for testing to detect duplicate samples from the same patient. Consensus clustering of these samples together with the fresh-frozen cohort did not show any grouping by type of material, and all samples could be assigned a subgroup annotation (Fig. 1f). Using a reduced 48–CpG signature to train a support vector machine (SVM) classifier on the frozen tissue cohort, we were able to predict a tumour subgroup for the FFPE samples with an extremely close match to the clustering subgroup (97.6 % concordance, Rand index = 0.93, $p < 0.0001$; Supplementary Figure 1a, b). This signature allows for simple classification of single clinical samples without the need for comparison against a larger reference dataset. The data for these signature probes for the Heidelberg frozen tumour cohort are given in Supplementary Table 1.

We also investigated the impact of input DNA quantity on resulting data quality, using a dilution series of a single sample down to as little as 10 ng input material (the manufacturer's recommended input is 500 ng for fresh-frozen material or 250 ng for FFPE DNA). For both fresh-frozen and FFPE DNA, there was a very high correlation between profiles at all input quantities down to 50 ng (Fig. 1g). The frozen sample was also tested with 25 and 10 ng of input, resulting in a slightly lower correlation. However, even at 10 ng, the sample would still have been accurately classified as an SHH tumour (Supplementary Figure 1c). Thus, this platform may be suitable for molecular subgrouping even when DNA quantities are limiting.

To further validate the broad applicability of this technique, we examined an additional independent tumour cohort with matching expression subgrouping data (derived as previously described [8]), for which the 450k arrays

a Delta area plot

b k-means consensus clustering, 107 samples

c Expression

d MDS plot

e FFPE vs. Frozen (MB35)

f k-means consensus clustering, 276 samples

g Dilution series (MB35)

h

i

were run in an entirely separate facility (Toronto cohort, $n = 60$). SNP genotyping from the array indicated that four samples were also part of the Heidelberg frozen cohort. As with the frozen versus FFPE comparison, the correlation

between these paired samples run in different facilities was higher than any other pairwise comparison (mean Pearson's correlation for paired samples, $r = 0.988$). The derived methylation subgroups of the remaining 56 unique samples

again gave a very close overlap with the expression-defined subgroups (94.6 % concordance, Rand index = 0.86, $p < 0.0001$), showing the robustness of this platform for the classification of MB, independent of where the data are generated (Supplementary Figure 2a, b).

A further major benefit of using this comprehensive array platform rather than a targeted gene panel is the ability to generate genome-wide copy-number profiles using the intensity measures of the methylation probes, with a good concordance to other copy-number platforms such as CGH or SNP arrays, as we have recently described [14]. This allowed us to detect clinically relevant copy-number aberrations, such as MYC/MYCN/GLI2 gene amplifications, from the FFPE as well as the frozen tumour samples (Fig. 1f, h). Stereotypic MB copy-number changes showed the expected subgroup distribution (e.g. monosomy 6 in WNT tumours, 9q/10q loss in SHH, MYC amplification in Group 3, i(17q) in Group 3/Group 4; Fig. 1f). For 66 samples from the Heidelberg cohort, copy-number data from whole-genome sequencing (WGS) were also available, and were assessed for the alterations indicated in Fig. 1f. All scoring was consistent between WGS and 450k array profiles. Furthermore, 10/60 SHH-MBs showed patterns of dramatic copy-number change, reminiscent of chromothripsis [13] (Fig. 1i). We have previously linked this phenomenon to TP53 mutations (typically germline) in SHH-MB [11]. This tool may therefore aid in identifying medulloblastoma patients with a particularly high risk of having underlying Li Fraumeni syndrome.

In summary, we demonstrate here a method for reliable classification of medulloblastoma into molecular subgroups, and tumour copy-number profiling, using a commercially available DNA methylation array platform that performs well on either frozen or FFPE tumour material. We also show that this technology can be reproducibly applied with low amounts of starting material, at different institutes, and with the benefit of easier handling compared with FFPE-derived RNA. We therefore believe that this platform holds great potential for refining the information obtainable from large, archival tumour series. Most importantly, we also expect that this will become one of the key technologies for risk stratification and patient cohort selection in the next generation of large, biology-led, multi-centre clinical trials.

References

1. Gajjar A, Stewart CF, Ellison DW et al (2010) A phase I pharmacokinetic trial of sonic hedgehog (SHH) antagonist GDC-0449 in pediatric patients with recurrent or refractory medulloblastoma: a pediatric brain tumor consortium study (PBTC 25). J Clin Oncol 28(18s):Abstr. CRA9501

2. Geoerger B, Aerts I, Casanova M et al (2012) A phase I/II study of LDE225, a smoothened (Smo) antagonist, in pediatric patients with recurrent medulloblastoma (MB) or other solid tumors. J Clin Oncol 30(Suppl.):Abstr. 9519

3. Jones DTW, Jager N, Kool M et al (2012) Dissecting the genomic complexity underlying medulloblastoma. Nature 488(7409):100–105. doi:10.1038/nature11284

4. Kool M, Korshunov A, Remke M et al (2012) Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. Acta Neuropathol 123(4):473–484. doi:10.1007/s00401-012-0958-8

5. Kulis M, Heath S, Bibikova M et al (2012) Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet 44(11):1236–1242. doi:10.1038/ng.2443

6. Northcott PA, Jones DTW, Kool M et al (2012) Medulloblastomics: the end of the beginning. Nat Rev Cancer 12(12):818–834. doi:10.1038/nrc3410

7. Northcott PA, Korshunov A, Witt H et al (2011) Medulloblastoma comprises four distinct molecular variants. J Clin Oncol 29(11):1408–1414. doi:10.1200/JCO.2009.27.4324

8. Northcott PA, Shih DJ, Peacock J et al (2012) Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature 488(7409):49–56. doi:10.1038/nature11327

9. Northcott PA, Shih DJ, Remke M et al (2012) Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. Acta Neuropathol 123(4):615–626. doi:10.1007/s00401-011-0899-7

10. Noushmehr H, Weisenberger DJ, Diefes K et al (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 17(5):510–522. doi:10.1016/j.ccr.2010.03.017

11. Rausch T, Jones DTW, Zapatka M et al (2012) Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell 148(1–2):59–71. doi:10.1016/j.cell.2011.12.013

12. Schwalbe EC, Williamson D, Lindsey JC et al (2013) DNA methylation profiling of medulloblastoma allows robust subclassification and improved outcome prediction using formalin-fixed biopsies. Acta Neuropathol 125(3):359–371. doi:10.1007/s00401-012-1077-2

13. Stephens PJ, Greenman CD, Fu B et al (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144(1):27–40. doi:10.1016/j.cell.2010.11.055

14. Sturm D, Witt H, Hovestadt V et al (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22(4):425–437. doi:10.1016/j.ccr.2012.08.024

15. Taylor MD, Northcott PA, Korshunov A et al (2012) Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol 123(4):465–472. doi:10.1007/s00401-011-0922-z