1   An integrated genomic analysis of anaplastic meningioma identifies prognostic molecular

2   signatures

3

4   Grace Collord[1,2,20], Patrick Tarpey[1,20], Natalja Kurbatova[3], Inigo Martincorena[1], Sebastian

5   Moran[4], Manuel Castro[4], Tibor Nagy[1], Graham Bignell[1], Francesco Maura[1,5,6], Jorge

6   Berna[7], Jose M. Tubio[7], Chris E. McMurran[8], Adam M.H. Young[8], Matthew D. Young[1],

7   Imran Noorani[1,8], Stephen J Price[8], Colin Watts[8], Elke Leipnitz[9], Matthias Kirsch[9],

8   Gabriele Schackert[9], Danita Pearson[10], Abel Devadass[10], Zvi Ram[16], V. Peter Collins[10],

9   Kieren Allinson[10], Michael D. Jenkinson[11,19], Rasheed Zakaria[11,12], Khaja Syed[11,12], C.

10   Oliver Hanemann[13], Jemma Dunn[13], Michael W. McDermott[14], Ramez W Kirollos[8],

11   George S. Vassiliou[1,15,], Manel Esteller[4,17,18], Sam Behjati[1,2], Alvis Brazma[3], Thomas

12   Santarius[8*], Ultan McDermott[1*]

13

Affiliations:

14

15   [1] Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, CB10 1SA, UK

16   [2] Department of Paediatrics, University of Cambridge, Cambridge Biomedical Campus,

17   CB2 0QQ, UK

18   [3] European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI,

19   Wellcome Trust Genome Campus, Hinxton, CB10 1SD, UK

20   [4] Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research

21   Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Catalonia, Spain

22   [5] Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy

23   [6] Department of Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan,

24   Italy

25   [7] Phylogenomics Lab, Edificio Torre CACTI, Campus Universitario, Universidad de Vigo,

26   36310 Vigo, Spain

27   [8] Department of Neurosurgery, Department of Clinical Neuroscience, Cambridge

28   University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK

29   [9] Klinik und Poliklink für Neurochirurgie, "Carl Gustav Carus" Universitätsklinikum,

30   Technische Universität Dresden, Fetscherstrasse 74, 01307 Dresden, Germany

31   [10] Department of Pathology, Cambridge University Hospital, CB2 0QQ, Cambridge, UK

32   [11] Department of Neurosurgery, The Walton Centre, Liverpool, L9 7LJ, UK

33   [12] Institute of Integrative Biology, University of Liverpool, Liverpool, L9 7LJ, UK

34   [13] Institute of Translational and Stratified Medicine, Plymouth University Peninsula

35   Schools of Medicine and Dentistry, Plymouth University, Plymouth, Devon PL4 8AA, UK

36   [14] Department of Neurosurgery, UCSF Medical Center, San Francisco, CA 94143-0112,

37   USA

38   [15] Department of Haematology, Cambridge University Hospitals NHS Trust, Cambridge,

39   CB2 0QQ, UK

40   [16] Department of Neurosurgery, Tel-Aviv Medical Center, Tel-Aviv, Israel

41   [17] Physiological Sciences Department, School of Medicine and Health Sciences,

42   University of Barcelona (UB), Catalonia, Spain

43   [18] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia,

44   Spain

45   [19] Institute of Translational Medicine, University of Liverpool, Liverpool, L9 7LJ, UK

46   [20] These authors contributed equally to this work.

47   [*] Corresponding authors

48

49   Correspondence should be addressed to:

50   U.M. (*um1@sanger.ac.uk*), T.S. (*ts381@cam.ac.uk*)

51

52

53 **Abstract**

54 Anaplastic meningioma is a rare and aggressive brain tumor characterised by intractable

55 recurrences and dismal outcomes. Here, we present an integrated analysis of the whole

56 genome, transcriptome and methylation profiles of primary and recurrent anaplastic

57 meningioma. A key finding was the delineation of two distinct molecular subgroups that

58 were associated with diametrically opposed survival outcomes. Relative to lower grade

59 meningiomas, anaplastic tumors harbored frequent driver mutations in SWI/SNF complex

60 genes, which were confined to the poor prognosis subgroup. Our analyses discern two

61 biologically distinct variants of anaplastic meningioma with potential prognostic and

62 therapeutic significance.

63

64 **Introduction**

65 Meningiomas arise from arachnoidal cells of the meninges and are classified as grade I

66 (80% of cases), grade II (10-20%) or grade III (1-3%)[1,2]. Grade III meningiomas comprise

67 papillary, rhabdoid and anaplastic histological subtypes, with anaplastic tumors

68 accounting for the vast majority of grade III diagnoses[2,3]. Nearly half of anaplastic

69 meningiomas represent progression of a previously resected lower grade tumor, whereas

70 the remainder arise *de novo*[2,4,5]. Recurrence rates are 5-20% and 20-40%, respectively,

71 for grade I and 2 tumors[3,6]. By contrast, the majority of anaplastic meningioma patients

72 suffer from inexorable recurrences with progressively diminishing benefit from repeated

73 surgery and radiotherapy and 5-year overall survival of 30-60%[4,7].

74 A recent study of 775 grade I and grade 2 meningiomas identified five molecular

75 subgroups defined by driver mutation profile[8]. In keeping with previous smaller studies,

76 mutually exclusive mutations in *NF2* and *TRAF7* were the most frequent driver events,

77 followed by mutations affecting key mediators of PI3K and Hedgehog signalling[8,9].

78 Recurrent hotspot mutations were also identified in the catalytic unit of RNA polymerase II

79   (*POLR2A*) in 6% of grade I tumors[8]. More recently, a study comparing benign versus *de*

80   *novo* atypical (grade II) meningiomas found the latter to be significantly associated with

81   *NF2* and *SMARCB1* mutations[10]. Atypical meningiomas were further defined by DNA and

82   chromatin methylation patterns consistent with upregulated PRC2 activity, differentially

83   methylated Homeobox domains and transcriptional dysregulation of pathways involved in

84   proliferation and differentiation[10].

85   Despite the high mortality rate of anaplastic meningiomas, efforts to identify therapeutic

86   strategies have been hampered by a limited understanding of the molecular features of

87   this aggressive subtype. Here, we present an analysis of the genomic, transcriptional and

88   DNA methylation patterns defining anaplastic meningioma. Our results reveal two distinct

89   molecular subgroups associated with dramatically different prognoses.

90

91   **Results**

92   **Overview of the genomic landscape of primary and recurrent anaplastic**

93   **meningioma**

94   We performed whole genome sequencing (WGS) on a discovery set of 19 anaplastic

95   meningiomas resected at first presentation ('primary'). A subsequent validation cohort

96   comprised 31 primary tumors characterised by targeted sequencing of 366 cancer genes.

97   We integrated genomic findings with RNA sequencing and methylation array profiling in a

98   subset of samples (Supplementary Table S1). Somatic copy number alterations and

99   rearrangements were derived from whole genome sequencing reads, with RNA

100   sequences providing corroborating evidence for gene fusions. Given the propensity of

101   anaplastic meningioma to recur, we studied by whole genome sequencing 13 recurrences

102   from 7 patients.

103   Excluding a hypermutated tumor (PD23359a, see Supplementary Discussion), the

104   somatic point mutation burden of meningioma was low with a median of 28 somatic

4

105   coding mutations per tumor (range 11 to 71; mean sequencing coverage 66X) in 18

106   tumors interrogated by WGS at first presentation ('primary') (Supplementary Figure S1).

107   Mutational signatures analysis of substitutions identified in whole genome sequences

108   revealed the age-related, ubiquitous processes 1 and 5 as the predominant source of

109   substitutions (Supplementary Figure S2)[11].

110   The rearrangement profile of anaplastic meningioma is relatively quiet, with a median of

111   12 structural rearrangements (range 0–79) in the 18 primary tumor genomes

112   (Supplementary Figure S3, Supplementary Table S3). Somatic retrotransposition events,

113   a significant source of structural variants in over half of human cancers, were scarce

114   (Supplementary Figure S4; Supplementary Table S4)[12]. Analysis of expressed gene

115   fusions did not reveal any recurrent events involving putative cancer genes

116   (Supplementary Table S5).

117   Recurrent large copy number changes were in keeping with known copy number trends in

118   aggressive meningiomas, notably frequent deletions affecting chromosomes 1p, 6q, 14

119   and 22q (Figure 1b, Supplementary Table S6)[2,8,10,13].

120   The genomic landscape of recurrent tumors was largely static both with respect to driver

121   mutations and structural variation. Driver mutations differed between primary and

122   recurrent tumors for two of eleven patients with serial resections available. For seven sets

123   of recurrent tumors studied by whole genome sequencing, only two demonstrated any

124   discrepancies in large copy number variants (PD23344 and PD23346; Supplementary

125   Figure S5). Similarly, matched primary and recurrent samples clustered closely together

126   by PCA of transcriptome data, suggesting minimal phenotypic evolution (Supplementary

127   Figure S6).

128

129

130

Driver genes do not delineate subgroups of anaplastic meningioma

Over 80% of low grade meningiomas segregate into 5 distinct subgroups based on driver mutation profile[8,10]. In anaplastic meningioma, however, we found a more uniform driver landscape dominated by deleterious mutations in *NF2* (Figure 1a). A key feature distinguishing anaplastic meningioma from its lower grade counterparts were driver events in genes of the SWI/SNF chromatin regulatory complex (Figure 1a). The most frequently mutated SWI/SNF component was *ARID1A*, which harbored at least one deleterious somatic change in 12% of our cohort of 50 primary tumors (Supplementary Table S1). *ARID1A* has not been implicated as a driver in grade I or grade II meningiomas[8,10]. Single variants in *SMARCB1*, *SMARCA4* and *PBRM1* were also detected in three tumors (Supplementary Figure S7). In total, 16% of anaplastic meningiomas contained a SWI/SNF gene mutation. By contrast, SWI/SNF genes are mutated in <5% of benign and atypical meningiomas[8,10].

In the combined cohort of 50 primary tumors, we found at least one driver mutation in *NF2* in 70%, similar to the prevalence reported in atypical meningiomas and more than twice that found in grade I tumors[8,10,14]. Interestingly, there was no significant difference in NF2 expression between *NF2* mutant and wild-type tumors (*p*-value 0.960; Supplementary Figure S8). We considered promoter hypermethylation as a source of *NF2* inactivation, but found no evidence of this (Supplementary Table S7). Thus, as observed in other cancer types, non-mutational mechanisms may contribute to *NF2* loss of function in a proportion of anaplastic meningiomas[15-20].

Other driver genes commonly implicated in low grade tumors were not mutated, or very infrequently (Figure 1a). Furthermore, we did not observe an increased frequency of *TERT* promoter mutations, previously associated with progressive or high grade tumors[21,22]. Notably[2,23,24], methylation analysis revealed *CDKN2A* and *PTEN* promoter hypermethylation in 17% and 11% of primary tumors, respectively (Figure 1a). We did not

6

157 find evidence of novel cancer genes in our cohort, applying established methods to

158 search for enrichment of non-synonymous mutations (See Methods and Supplementary

159 Methods). The full driver landscape of anaplastic meningioma, considering point

160 mutations, structural variants with resulting copy number changes and promoter

161 hypermethylation is presented in Supplementary Figure S7.

162

163 **Differential gene expression defines anaplastic meningioma subgroups with**

164 **prognostic and biological significance**

165 We performed messenger RNA (mRNA) sequencing of 31 anaplastic meningioma

166 samples from a total of 28 patients (26 primary tumors and 5 recurrences). Gene

167 expression variability within the cohort did not correlate with clinical parameters including

168 prior radiotherapy, anatomical location or clinical presentation (*de novo* versus

169 progressive tumor) (Supplementary Figure S6). However, multiple unsupervised

170 hierarchical clustering methods and correlation measures consistently delineated two

171 distinct clusters, hereafter referred to as C1 and C2 (Figure 2a-c). The clinical trajectories

172 of patients in these clusters markedly diverged. We retrospectively sought follow-up

173 survival data from the time of first surgery, which was available for 25 of the 28 patients

174 included in the transcriptome analysis (12 patients in C1, 13 in C2; mean follow-up of

175 1,403 days from surgery). We observed a significantly worse overall survival outcome in

176 C1 compared to C2 (p<0.0001, hazard ratio 17.0 (95% CI 5.2-56.0)) (Figure 2g;

177 Supplementary Table S8). The subgroups were well balanced with respect to potential

178 confounding features such as gender, age, radiotherapy and anatomical location

179 (Supplementary Table S9). Paradoxically, a greater proportion of C2 patients had higher

180 Simpson surgical resection scores (indicating more residual tumor after surgery),

181 conventionally a negative prognostic indicator (*P* = 0.081, Fisher's exact test) [4,6].

182

7

Transcriptional programmes segregating anaplastic meningioma

Nineteen hundred genes underpinned the differentiation of anaplastic meningioma into subgroups C1 and C2, which could be reduced to only 6 transcripts selected on the basis of PCA coefficient and differential expression analysis (see Methods; Supplementary Tables S10 and S11; Supplementary Figure S9). Pathway enrichment analysis was most significant for evidence of epithelial-mesenchymal transition (EMT) in the C1 tumors, with concordant loss of E-cadherin (*CDH1*) and upregulation of *CXCL14*, both prognostic biomarkers in diverse other cancers (Supplementary Table S12, Figure 2d-f)[25-32]. The C1 and C2 tumors were further distinguished by significant dysregulation of proliferation, PRC2 activity and embryonic stem cell transcriptional programmes (Supplementary Table S13). Hox genes constituted a notable proportion of the transcripts distinguishing the two anaplastic meningioma subgroups, largely underpinning the significance of pathways involved in tissue morphogenesis. Furthermore, differentially methylated genes were also significantly enriched for Hox genes, with pathway analysis results corroborating the main biological themes emerging from transcriptome analysis (Supplementary Tables S14 and S15).

Comparison of the anaplastic and benign meningioma transcriptome

Previous studies investigating the relationship between meningioma WHO grade and gene expression profiles have included few anaplastic tumors[33,34]. We therefore extended our analysis to include published RNA sequences from 19 benign grade I meningiomas. External data was processed using our in-house pipeline with additional measures taken to minimise batch effects (see Methods and Supplementary Tables S16 and S17). Unsupervised hierarchical clustering and principal component analysis demonstrated clear tumor segregation by histological grade (Figure 3a,b).

8

208   Consistent with there being a coherent biological trend across histological grades and

209   anaplastic meningioma subgroups, we noted significant overlap between genes

210   differentially expressed between grades and between C1 and C2 tumors (hypergeometric

211   distribution P = 5.084E-09). In keeping with this finding, formal pathway analysis identified

212   significant dysregulation of stemness, proliferation, EMT and PRC2 activity

213   (Supplementary Tables S18 and S19). The most significantly dysregulated pathways also

214   included TGF-beta, Wnt and integrin signalling, mediators of invasion and mesenchymal

215   differentiation that are normally in part controlled by NF2 and other Hippo pathway

216   members[15,16,35-37]. Yes-associated protein 1 (Yap1), a cornerstone of oncogenic Hippo

217   signalling, is frequently overexpressed in cancer and synergises with Wnt signalling to

218   induce EMT[15,37-39]. *YAP1* was upregulated in anaplastic tumors (log$_2$ fold change = 2.18,

219   FDR = 0.0062) along with *MYL9* (log$_2$ fold change = 6.25, FDR = 3.76E-18), a key

220   downstream effector essential for Yap1-mediated stromal reprogramming (Figure 3c)[38].

221   Additionally, the anaplastic tumors demonstrated further upregulation of major growth

222   factor receptor and kinase circuits previously implicated in meningioma pathogenesis,

223   notably epidermal growth factor receptor (EGFR), insulin-like growth factor (IGFR),

224   vascular endothelial growth factor receptor (VEGFR) and mTOR complex 1 (mTORC1)

225   kinase complex [40-45].

226

227   **Discussion**

228   Meningiomas constitute a common, yet diverse tumor type with few therapeutic

229   options[1,7,8,10]. Efforts to improve clinical outcomes have been hampered by limited

230   understanding of the molecular determinants of aggressive disease. Here, we explored

231   genomic, epigenetic and transcriptional features of anaplastic meningioma, the most

232   lethal meningioma subtype[4].

233

9

Frequent somatic changes in SWI/SNF complex genes, predominantly *ARID1A,* constitute the main genomic distinction between anaplastic and lower grade meningiomas[8,10]. SWI/SNF acts as a tumor suppressor in many cell types by antagonising the chromatin modifying Polycomb repressive complex 2 (PRC2)[46-48]. Deleterious SWI/SNF mutations unfetter PRC2 activity with profound consequences for primitive developmental pathways frequently co-opted during oncogenesis[48,49]. SWI/SNF inactivation is associated with a stem cell-like phenotype and poor outcomes in diverse cancer types[50-54].

Although anaplastic tumors resist meaningful classification based on driver mutation patterns, unbiased transcriptional profiling revealed two biologically distinct subgroups with dramatically divergent survival outcomes. This finding is emblematic of the limitations of histopathological grading as a risk stratification system for meningioma[1,3,4,13,55]. Interestingly, all SWI/SNF mutations were exclusively identified in the poor prognosis (C1) subgroup (P = 0.016, Fisher's exact test). C1 tumors were further characterised by transcriptional signatures of PRC2 target activation, stemness, proliferation and mesenchymal differentiation. These findings were in part underpinned by differential expression of Hox genes. Acquisition of invasive capacity and stem cell traits are frequently co-ordinately dysregulated in cancer, often through subversion of Hox gene programmes integral to normal tissue morphogenesis[56-58]. Hox genes have a central role in orchestrating vertebrate development and act as highly context-dependent oncogenes and tumor suppressors in cancer[57,59]. Several of the most starkly upregulated Hox genes in the C1 tumors consistently function as oncogenes across a range of solid and haematological malignancies, including *HOTAIR*, *HOXB7*, *HOXA4*, *HOXA-AS2*, *HOXC11,* and *NKX2-2*[57,60-70]. Like many long non-coding RNAs (lncRNA), *HOTAIR* and *HOXA-AS2* modulate gene expression primarily by interacting directly with chromatin

10

260 remodelling complexes, exerting oncogenic activity by recruiting PRC2 to target

261 genes[60,62,69-73]. *HOXA-AS2* has been shown to mediate transcriptional repression of the

262 tumor suppressor gene *CDKN2A* (p16[INK4A]), loss of which is associated with poor

263 meningioma survival[23,24,60,69,70]. Given the antagonistic relationship between the SWI/SNF

264 and PRC2 chromatin regulators, deleterious SWI/SNF mutations and overexpression of

265 lncRNAs known to mediate PRC2 activity emerge as potentially convergent mechanisms

266 underpinning the differences between C1 and C2 tumors[49].

267

268 In the context of recent studies of lower grade meningiomas, our findings raise the

269 possibility that the balance between PRC2 and SWI/SNF activity may have broader

270 relevance to meningioma pathogenesis. Compared to grade I tumors, atypical

271 meningiomas are more likely to harbor *SMARCB1* mutations and large deletions

272 encompassing chromosomes 1q, 6q and 14q. Notably, these genomic regions

273 encompass *ARID1A* and several other SWI/SNF subunit genes. Both *SMARCB1*

274 mutations and the aforementioned copy number changes were associated with

275 epigenetic evidence of increased PRC2 activity, differential Homeobox domain

276 methylation, and upregulation of proliferation and stemness programmes in the atypical

277 tumors.

278

279 The extent to which SWI/SNF depletion plays a role in lower grade meningiomas may be

280 therapeutically relevant. Diverse SWI/SNF mutated cancers exhibit dependence on both

281 catalytic and non-catalytic functions of EZH2, a core subunit of PRC2[74-76]. Several EZH2

282 inhibitors are in development with promising initial clinical results[77]. Other modulators of

283 PRC2 activity, including HOTAIR, may also be relevant therapeutic targets[78,79].

284 Furthermore, growing recognition of the relationship between EMT and resistance to

285 conventional and targeted anti-cancer agents has profound implications for rational

286   integration of treatment approaches[80,81]. Notably, EGFR inhibition has yielded

287   disappointing response rates in meningioma[41,82]. A mesenchymal phenotype is strongly

288   associated with resistance to EGFR inhibitors in lung and colorectal cancer[80,81,83-85].

289   Combining agents that abrogate EMT with other therapies is a promising strategy for

290   addressing cell-autonomous and extrinsic determinants of disease progression and may

291   warrant further investigation in meningioma[80,86].

292

293   This study has revealed prognostically significant anaplastic meningioma subgroups and

294   identified potentially actionable alternations in SWI/SNF genes and other therapeutically

295   tractable targets. However, a substantially larger series of tumors, ideally nested in a

296   prospective multicentre observational study, will be required to expand upon our main

297   findings and explore mechanistic and therapeutic ramifications of meningioma diversity.

298

Methods
Sample selection

DNA was extracted from 70 anaplastic meningiomas; 51 samples at first resection ('primary') and 19 from subsequent recurrences. Matched normal DNA was derived from peripheral blood lymphocytes. Written informed consent was obtained for sample collection and DNA sequencing from all patients in accordance with the Declaration of Helsinki and protocols approved by the NREC/Health Research Authority (REC reference 14/YH/0101) and Ethics Committee at University Hospital Carl Gustav Carus, Technische Universität Dresden, Germany (EK 323122008). Samples underwent independent specialist pathology review (V.P.C and K.A). DNA extracted from fresh-frozen material was submitted for whole genome sequencing whereas that derived from formalin-fixed paraffin-embedded (FFPE) material underwent deep targeted sequencing of 366 cancer genes.

One tumor sample PD23348 (and two subsequent recurrences) separated from the main study samples in a principal components analysis of transcriptomic data (Supplementary Figure S10). Analysis of WGS and RNA sequencing data identified an expressed gene fusion, *NAB2-STAT6*. This fusion is pathognomonic of meningeal hemangiopericytoma, now classified as a separate entity, solitary fibrous tumors[87-89]. We therefore excluded three samples from this tumor from further study. A second sample (PD23354a), diagnosed as an anaplastic meningioma with papillary features, was found to have a strong APOBEC mutational signature as well as an *EML4-ALK* gene fusion (exon 6 EML4, exon 19 ALK) (Supplementary Figure S11)[90]. Therefore this sample was also removed as a likely metastasis from a primary lung adenocarcinoma. The hypermutator sample PD23359a underwent additional pathological review to confirm the diagnosis of

13

324 anaplastic meningioma (K.A., Department of Histopathology, Cambridge University

325 Hospital, Cambridge, UK).

326

327 RNA was extracted from fresh-frozen material from 34 primary and recurrent tumors, 3 of

328 which were from PD23348 and were subsequently excluded from final analyses

329 (Supplementary Table S1).

330

331 **Whole genome sequencing** Short insert 500bp genomic libraries were constructed,

332 flowcells prepared and sequencing clusters generated according to Illumina library

333 protocols[91]. 108 base/100 base (genomic), or 75 base (transcriptomic) paired-end

334 sequencing were performed on Illumina X10 genome analyzers in accordance with the

335 Illumina Genome Analyzer operating manual. The average sequence coverage was

336 65.8X for tumor samples and 33.8X for matched normal samples (Supplementary Table

337 S1).

338

339 **Targeted genomic sequencing**

340 For targeted sequencing we used a custom cRNA bait set (Agilent) to enrich for all coding

341 exons of 366 cancer genes (Supplementary Table S20). Short insert libraries (150bp)

342 were prepared and sequenced on the Illumina HiSeq 2000 using 75 base paired-end

343 sequencing as per Illumina protocol. The average sequence coverage was 469X for the

344 tumor samples.

345

346 **RNA sequencing and data processing**

347 For transcriptome sequencing, 350bp poly-A selected RNA libraries were prepared on the

348 Agilent Bravo platform using the Stranded mRNA library prep kit from KAPA Biosystems.

349 Processing steps were unchanged from those specified in the KAPA manual except for

350   use of an in-house indexing set. Reads were mapped to the GRCh37 reference genome

351   using STAR (v2.5.0c)[92]. Mean sequence coverage was 128X. Read counts per gene,

352   based on the union of all exons from all possible transcripts, were then extracted BAM

353   files using HTseq (v0.6.1)[93]. Transcripts Per kilobase per Million reads (TPM) were

354   generated using an in-house python script

355   (https://github.com/TravisCG/SI_scripts/blob/master/tpm.py)[92,93]. We downloaded

356   archived RNA sequencing FASTQ files for 19 grade I meningioma samples representing

357   the major mutational groups (*NF2*/chr22 loss, *POLR2A*, *KLF4/TRAF7*, *PI3K* mutant)

358   (ArrayExpress: GSE85133)[8]. Reads were then processed using STAR and HTseq as

359   described above. Cancer cell line (n=252) and triple-negative breast cancer (n = 100)

360   RNA sequencing data was generated in-house by the aforementioned sequencing and

361   bioinformatic pipeline.

362   Expressed gene fusions were sought using an in-house pipeline incorporating three

363   algorithms: TopHat-Fusion (v2.1.0), STAR-Fusion (v0.1.1) and deFuse (v0.7.0)

364   (https://github.com/cancerit/cgpRna)[92,94,95]. Fusions identified by one or two algorithms or

365   also detected in the matched normal sample were flagged as likely artefacts. Fusions

366   were further annotated according to whether they involved a kinase or known oncogene

367   and whether they occurred near known fragile sites or rearrangement break points[96]

368   (Supplementary Table S5).

369

370   **Differential gene expression and pathway enrichment analysis**

371   The DESeq2 R package was used for all differential gene expression analyses[97,98].

372   DESeq2 uses shrinkage estimation of dispersion for the sample-specific count

373   normalization and subsequently applies a linear regression method to identify differentially

374   expressed genes (DEGs)[97,98].

375   Preliminary comparison of anaplastic and externally-generated grade I meningioma data

376   revealed evidence of laboratory batch effects, which we mitigated with two

377   batch-correction methods: RUVg and PEER[99,100]. RUVg estimates the factor attributed to

378   spurious variation using control genes that are assumed to have constant expression

379   across samples[101-103]. We selected control genes (*RPL37A, EIF2B1, CASC3, IPO8,*

380   *MRPL19, PGK1* and *POP4*) on the basis of previous studies of suitable control genes for

381   transcript-based assays in meningioma[104]. PEER ('probabilistic estimation of expression

382   residuals') is based on factor analysis methods that infer broad variance components in

383   the measurements. PEER can find hidden factors that are orthogonal to the known

384   covariates. We applied this feature of PEER to remove additional hidden effect biases.

385   The final fitted linear regression model consists of the factor identified by RUVg method

386   that represents the unwanted laboratory batch effect and 13 additional hidden factors

387   found by PEER that are orthogonal to the estimated laboratory batch effect. Using this

388   approach we were able to reduce the number of DEGs from more than 18000 to 8930, of

389   which <4,000 are predicted to be protein-coding.

390   To identify biological pathways differentially expressed between meningioma grades and

391   anaplastic meningioma subgroups we applied a functional class scoring algorithm using a

392   collection of 461 published gene sets mapped to 10 canonical cancer hallmarks

393   (Supplementary Table S21)[56,105-109]. We further corroborated these findings with a more

394   general Gene Ontology (GO) pathway analysis[110].

395

396   **Identification of 6 transcripts recapitulating anaplastic meningioma clusters**

397   Mapped RNA sequencing reads were normalised using the regularised logarithm (rlog)

398   function implemented by the DESeq2 package[97,98]. PCA was performed using the top 500

399   most variably expressed transcripts and the prcomp function (R stats package)[111]. Given

400   that primary component 1 (PC1) was the vector most clearly distinguishing the closely

16

401 clustered C2 subgroup from the more diffusely clustered C1 (Figure 3a), we extracted the

402 top 50 transcripts with the highest absolute PC1 coefficients. We then identified the subset

403 that overlapped with the most significantly differentially expressed genes (absolute $\log_2$

404 fold change > 4 and adjust *p*-value < 0.0001) between i) the C1 and C2 anaplastic

405 meningioma subgroups and ii) the C1 anaplastic meningiomas and the 19 grade I tumors

406 (Supplementary Tables S10 and S17). Iteratively reducing the number of PC1

407 components identified the minimum number of transcripts that recapitulated segregation

408 of C1 and C2 tumors upon unsupervised hierarchical clustering and PCA (Supplementary

409 Table S11; Supplementary Figure S9).

410

411 **Processing of genomic sequencing data**

412 Genomic reads were aligned to the reference human genome (GRCh37) using the

413 Burrows-Wheeler Aligner, BWA (v0.5.9)[112]. CaVEMan (Cancer Variants Through

414 Expectation Maximization: http://cancerit.github.io/CaVEMan/) was used for calling

415 somatic substitutions. Small insertions and deletions (indels) in tumor and normal reads

416 were called using a modified Pindel version 2.0. (http://cancerit.github.io/cgpPindel/) on

417 the NCBI37 genome build[113,114]. Annotation was according to ENSEMBL version 58.

418 Structural variants were called using a bespoke algorithm, BRASS (BReakpoint

419 AnalySiS) (https://github.com/cancerit/BRASS) as previously described[115].

420 The ascatNGS algorithm was used to estimate tumor purity and ploidy and to construct

421 copy number profiles from whole genome data[116].

422

423 **Identification of cancer genes based on the impact of coding mutations**

424 To identify recurrently mutated driver genes, we applied a dN/dS method that considers

425 the mutation spectrum, the sequence of each gene, the impact of coding substitutions

426 (synonymous, missense, nonsense, splice site) and the variation of the mutation rate

427   across genes[115]. To detect genes under significant selective pressure by either point

428   mutations or indels, each gene's *P*-values from dN/dS analysis of substitutions and from

429   the recurrence analysis of indels were combined using Fisher's method. Multiple testing

430   correction (Benjamini-Hochberg FDR) was performed separately for all genes, stratifying

431   the FDR correction to increase sensitivity[117]. To achieve a low false discovery rate a

432   conservative *q*-value cutoff of <0.05 was used for significance and considered significant

433   any gene with qmis_sfdr<0.05 OR qglobal_sfdr<0.05. See Supplementary Methods.

434

435   **Identification of driver mutations in known cancer genes**

436   Non-synonymous coding variants detected by Caveman and Pindel algorithms were

437   flagged as putative driver mutations according to set criteria and further curated following

438   manual inspection in the Jbrowse genome browser[118]. Variants were screened against

439   lists of somatic mutations identified by a recent study of 11,119 human tumors

440   encompassing 41 cancer types and also against a database of validated somatic drivers

441   identified in cancer sequencing studies at the Wellcome Trust Sanger Institute

442   (Supplementary Tables S22 and S23)[119].

443   Copy number data was analysed for homozygous deletions encompassing tumor

444   suppressor genes and for oncogene amplifications exceeding 5 or 9 copies for diploid

445   and tetraploid genomes, respectively. Only focal (<1Mb) copy number variants meeting

446   these criteria were considered potential drivers. Additional truncating events (disruptive

447   rearrangement break points, nonsense point mutations, essential splice site mutations

448   and out of frame indels) in established tumor suppressors were also flagged as potential

449   drivers. Only rearrangements with breakpoints able to be reassembled at base pair

450   resolution are included in this dataset.

451

452

TraFiC pipeline for retrotransposon integration detection

For the identification of putative solo-L1 and L1-transduction integration sites, we used the TraFiC (Transposome Finder in Cancer) algorithm[12]. TraFiC uses paired-end sequencing data for the detection of somatic insertions of transposable elements (TEs) and exogenous viruses. The identification of somatic TEs (solo-L1, Alu, SINE, and ERV) is performed in three steps: (i) selection of candidate reads, (ii) transposable element masking, (iii) clustering and prediction of TE integration sites and (iv) filtering of germline events[12].

Methylation arrays and analysis

We performed quantitative methylation analysis of 850,000 CpG sites in 25 anaplastic meningiomas. Bisulfite-converted DNA (bs-DNA) was hybridized on the Ilumina Infinium HumanMethylationEPIC BeadChip array following the manufacturer's instructions. All patient DNA samples were assessed for integrity, quantity and purity by electrophoresis in a 1.3% agarose gel, picogreen quantification and Nanodrop measurements. Bisulfite conversion of 500 ng of genomic DNA was done using the EZ DNA Methylation Kit (Zymo Research), following the manufacturer's instructions. Resulting raw intensity data (IDATs) were normalized under R statistical environment using the Illumina normalization method developed under the minfi package (v1.19.10). Normalized intensities were then used to calculate DNA methylation levels (beta values). Then, we excluded from the analysis the positions with background signal levels in methylated and unmethylated channels (p>0.01). Finally we removed probes with one or more single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) > 1% in the first 10 bp of the interrogated CpG, as well as the probes related to X and Y chromosomes. From the filtered positions, we selected only CpG sites present both in promoter regions (TSS, 5'UTR and 1st exon) and CpG islands (UCSC database, genome version hg19).

19

479 For the supervised analysis of the probes, CpG sites were selected by applying an

480 ANOVA test to identify statistically significant CpG positions (FDR adjusted p-value <

481 0.01) that were differentially methylated among the compared groups (Δβ > 0.2). Selected

482 CpG sites were later clustered based on the Manhattan distances aggregated by ward's

483 linkage. Finally, the genes corresponding to the selected CpGs were used to perform a

484 Gene Set Enrichment Analysis (GSEA) with curated gene sets in the Molecular

485 Signatures Database[120]. The gene sets used were: H: hallmark gene sets, BP: GO

486 biological process, CC: GO cellular component, MF: GO molecular function and C3: motif

487 gene sets (http://software.broadinstitute.org/gsea/msigdb/collections.jsp). The gene

488 clusters resulting from the hypergeometric test with a FDR adjusted p-value < 0.05 were

489 finally considered. We observed high levels of methylation for *CREBBP* in the majority of

490 tumor samples, however, similar patterns were manifest in normal tissue controls, hence

491 *CREBBP* hypermethyation does not appear to be a feature of oncogenesis in these

492 samples.

493

494 **Mutational signature analysis**

495 Mutational signature extraction was performed using the nonnegative matrix factorization

496 (NNMF) algorithm[11]. Briefly, the algorithm identifies a minimal set of mutational signatures

497 that optimally explains the proportions of mutation types found across a given mutational

498 catalogue and then estimates the contribution of each identified signature to the mutation

499 spectra of each sample.

500

501 **Patient survival analysis**

502 The Kaplan-Meier method was used to analyze survival outcomes by the log-rank

503 Mantel-Cox test, with hazard ratio and two-sided 95% confidence intervals calculated

504 using the Mantel_Haenszel test (GraphPad Prism, ver 7.02). Overall survival data from

505 time of first surgery for each anaplastic meningioma within gene-expression defined

506 subgroups 1 and 2 was collected and used to plot a Kaplan-Meier survival curve.

507

508 **Data availability**

509 All sequencing data that support the findings of this study have been deposited in the

510 European Genome-Phenome Archive and are accessible through the accession numbers

511 EGAS00001000377, EGAS00001000828, EGAS00001000859, EGAS00001001155 and

512 EGAS00001001873. All other relevant data are available from the corresponding author

513 on request.

514

515 **Supplementary Discussion**

516 **A hypermutator anaplastic meningioma with a haploid genome**

517 One primary anaplastic meningioma resected from an 85-year old female (PD23359a)

518 had a hypermutator phenotype, with 27,332 point mutations and LOH across nearly its

519 entire genome (Supplementary Figure S12; Supplementary Table S24). Independent

520 pathological review confirmed the original diagnosis of anaplastic meningioma, and

521 transcriptome analysis confirmed that this tumor clustered appropriately with the rest of

522 the cohort (Figure 3a,b). The majority of the mutations were substitutions, 72% of which

523 were C>T transitions. We identified two deleterious mutations in DNA damage repair

524 mediators: a *TP53* p.R248Q missense mutation and a homozygous truncating variant in

525 the mismatch repair gene *MSH6* (p.L1330Vfs*9). Despite the latter finding, mutational

526 signatures analysis was dominated by signature 1, with no evidence of signatures

527 typically associated with defects in homologous recombination, mismatch repair or *POLE*

528 activity (signatures 3, 6, 10, 15, 20 or 26). The copy number profile is most consistent with

529 this tumor having first undergone haploidization of its genome, with the exception of

530 chromosomes 7, 19 and 20, followed by whole genome duplication (Supplementary

21

531   Figure S12). Of note, several important oncogenes are located on chromosome 7,

532   including *EGFR*, *MET* and *BRAF*. Widespread LOH has been described in a significant

533   proportion of oncocytic follicular thyroid cancers where preservation of chromosome 7

534   heterozygosity has also been observed[121].

535

536

537

References

1   Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* **131**, 803-820, doi:10.1007/s00401-016-1545-1 (2016).

2   Mawrin, C. & Perry, A. Pathological classification and molecular genetics of meningiomas. *J Neurooncol* **99**, 379-391, doi:10.1007/s11060-010-0342-2 (2010).

3   Rogers, C. L. *et al.* Pathology concordance levels for meningioma classification and grading in NRG Oncology RTOG Trial 0539. *Neuro Oncol* **18**, 565-574, doi:10.1093/neuonc/nov247 (2016).

4   Champeaux, C., Wilson, E., Brandner, S., Shieff, C. & Thorne, L. World Health Organization grade III meningiomas. A retrospective study for outcome and prognostic factors assessment. *Br J Neurosurg* **29**, 693-698, doi:10.3109/02688697.2015.1054350 (2015).

5   Moliterno, J. *et al.* Survival in patients treated for anaplastic meningioma. *Journal of neurosurgery* **123**, 23-30, doi:10.3171/2014.10.JNS14502 (2015).

6   Durand, A. *et al.* WHO grade II and III meningiomas: a study of prognostic factors. *J Neurooncol* **95**, 367-375, doi:10.1007/s11060-009-9934-0 (2009).

7   Buttrick, S., Shah, A. H., Komotar, R. J. & Ivan, M. E. Management of Atypical and Anaplastic Meningiomas. *Neurosurg Clin N Am* **27**, 239-247, doi:10.1016/j.nec.2015.11.003 (2016).

8   Clark, V. E. *et al.* Recurrent somatic mutations in POLR2A define a distinct subset of meningiomas. *Nat Genet* **48**, 1253-1259, doi:10.1038/ng.3651 (2016).

9   Clark, V. E. *et al.* Genomic analysis of non-NF2 meningiomas reveals mutations in TRAF7, KLF4, AKT1, and SMO. *Science* **339**, 1077-1080, doi:10.1126/science.1233009 (2013).

10   Harmanci, A. S. *et al.* Integrated genomic analyses of de novo pathways underlying atypical meningiomas. *Nature communications* **8**, 14433, doi:10.1038/ncomms14433 (2017).

11   Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421, doi:10.1038/nature12477 (2013).

12   Tubio, J. M. *et al.* Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* **345**, 1251343, doi:10.1126/science.1251343 (2014).

13   Sahm, F. *et al.* DNA methylation-based classification and grading system for meningioma: a multicentre, retrospective analysis. *Lancet Oncol*, doi:10.1016/S1470-2045(17)30155-9 (2017).

14   Forbes, S. A. *et al.* COSMIC: somatic cancer genetics at high-resolution. *Nucleic acids research* **45**, D777-d783, doi:10.1093/nar/gkw1121 (2017).

15   Harvey, K. F., Zhang, X. & Thomas, D. M. The Hippo pathway and human cancer. *Nat Rev Cancer* **13**, 246-257, (2013).

16   McClatchey, A. I. Merlin and ERM proteins: unappreciated roles in cancer development? *Nat Rev Cancer* **3**, 877-883 (2003).

17   Petrilli, A. M. & Fernandez-Valle, C. Role of Merlin/NF2 inactivation in tumor biology. *Oncogene* **35**, 537-548, doi:10.1038/onc.2015.125 (2016).

18   Morales, F. C., Molina, J. R., Hayashi, Y. & Georgescu, M. M. Overexpression of ezrin inactivates NF2 tumor suppressor in glioblastoma. *Neuro Oncol* **12**, 528-539, doi:10.1093/neuonc/nop060 (2010).

19   Shaw, R. J. *et al.* The Nf2 tumor suppressor, merlin, functions in Rac-dependent signaling. *Dev Cell* **1**, 63-72 (2001).

20   Xiao, G. H., Beeser, A., Chernoff, J. & Testa, J. R. p21-activated kinase links Rac/Cdc42 signaling to merlin. *J Biol Chem* **277**, 883-886, doi:10.1074/jbc.C100553200 (2002).

21   Goutagny, S. *et al.* High incidence of activating TERT promoter mutations in meningiomas undergoing malignant progression. *Brain Pathol* **24**, 184-189, doi:10.1111/bpa.12110 (2014).

22   Koelsche, C. *et al.* Distribution of TERT promoter mutations in pediatric and adult tumors of the nervous system. *Acta Neuropathol* **126**, 907-915, doi:10.1007/s00401-013-1195-5 (2013).

23   Bostrom, J. *et al.* Alterations of the tumor suppressor genes CDKN2A (p16(INK4a)), p14(ARF), CDKN2B (p15(INK4b)), and CDKN2C (p18(INK4c)) in atypical and anaplastic meningiomas. *Am J Pathol* **159**, 661-669, doi:10.1016/S0002-9440(10)61737-3 (2001).

24   Goutagny, S. *et al.* Genomic profiling reveals alternative genetic pathways of meningioma malignant progression dependent on the underlying NF2 status. *Clin Cancer Res* **16**, 4155-4164, doi:10.1158/1078-0432.CCR-10-0891 (2010).

593   25   Paredes, J. *et al.* Epithelial E- and P-cadherins: role and clinical significance in cancer. *Biochimica*
594   *et biophysica acta* **1826**, 297-311, doi:10.1016/j.bbcan.2012.05.002 (2012).

595   26   Berx, G. & van Roy, F. Involvement of members of the cadherin superfamily in cancer. *Cold*
596   *Spring Harbor perspectives in biology* **1**, a003129, doi:10.1101/cshperspect.a003129 (2009).

597   27   Alexiadis, M. *et al.* Transcriptomic analysis of stage 1 versus advanced adult granulosa cell
598   tumors. *Oncotarget* **7**, 14207-14219, doi:10.18632/oncotarget.7422 (2016).

599   28   Augsten, M. *et al.* Cancer-associated fibroblasts expressing CXCL14 rely upon NOS1-derived
600   nitric oxide signaling for their tumor-supporting properties. *Cancer Res* **74**, 2999-3010,
601   doi:10.1158/0008-5472.can-13-2740 (2014).

602   29   Sjoberg, E., Augsten, M., Bergh, J., Jirstrom, K. & Ostman, A. Expression of the chemokine
603   CXCL14 in the tumour stroma is an independent marker of survival in breast cancer. *Br J Cancer*
604   **114**, 1117-1124, doi:10.1038/bjc.2016.104 (2016).

605   30   Zakrzewski, K. *et al.* Transcriptional profiles of pilocytic astrocytoma are related to their three
606   different locations, but not to radiological tumor features. *BMC Cancer* **15**, 778,
607   doi:10.1186/s12885-015-1810-z (2015).

608   31   Zeraati, M. *et al.* Cancer-associated noncoding mutations affect RNA G-quadruplex-mediated
609   regulation of gene expression. *Sci Rep* **7**, 708, doi:10.1038/s41598-017-00739-y (2017).

610   32   Zhao, L. *et al.* Long Noncoding RNA LINC00092 Acts in Cancer-Associated Fibroblasts to Drive
611   Glycolysis and Progression of Ovarian Cancer. *Cancer Res* **77**, 1369-1382,
612   doi:10.1158/0008-5472.can-16-1615 (2017).

613   33   Watson, M. A. *et al.* Molecular characterization of human meningiomas by gene expression
614   profiling using high-density oligonucleotide microarrays. *The American journal of pathology* **161**,
615   665-672, doi:10.1016/s0002-9440(10)64222-8 (2002).

616   34   Wrobel, G. *et al.* Microarray-based gene expression profiling of benign, atypical and anaplastic
617   meningiomas identifies novel genes associated with meningioma progression. *Int J Cancer* **114**,
618   249-256, doi:10.1002/ijc.20733 (2005).

619   35   De Craene, B. & Berx, G. Regulatory networks defining EMT during cancer initiation and
620   progression. *Nature reviews. Cancer* **13**, 97-110, doi:10.1038/nrc3447 (2013).

621   36   Morrow, K. A. *et al.* Loss of tumor suppressor Merlin results in aberrant activation of
622   Wnt/beta-catenin signaling in cancer. *Oncotarget* **7**, 17991-18005, doi:10.18632/oncotarget.7494
623   (2016).

624   37   Kalluri, R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer* **16**, 582-598,
625   doi:10.1038/nrc.2016.73 (2016).

626   38   Calvo, F. *et al.* Mechanotransduction and YAP-dependent matrix remodelling is required for the
627   generation and maintenance of cancer-associated fibroblasts. *Nat Cell Biol* **15**, 637-646,
628   doi:10.1038/ncb2756 (2013).

629   39   Rosenbluh, J. *et al.* beta-Catenin-driven cancers require a YAP1 transcriptional complex for
630   survival and tumorigenesis. *Cell* **151**, 1457-1473, doi:10.1016/j.cell.2012.11.026 (2012).

631   40   Mawrin, C. *et al.* Different activation of mitogen-activated protein kinase and Akt signaling is
632   associated with aggressive phenotype of human meningiomas. *Clin Cancer Res* **11**, 4074-4082,
633   doi:10.1158/1078-0432.ccr-04-2550 (2005).

634   41   Mawrin, C., Chung, C. & Preusser, M. Biology and clinical management challenges in
635   meningioma. *American Society of Clinical Oncology educational book. American Society of*
636   *Clinical Oncology. Meeting*, e106-115, doi:10.14694/EdBook_AM.2015.35.e106 (2015).

637   42   Lopez-Lago, M. A., Okada, T., Murillo, M. M., Socci, N. & Giancotti, F. G. Loss of the tumor
638   suppressor gene NF2, encoding merlin, constitutively activates integrin-dependent mTORC1
639   signaling. *Mol Cell Biol* **29**, 4235-4249, doi:10.1128/mcb.01578-08 (2009).

640   43   James, M. F. *et al.* NF2/merlin is a novel negative regulator of mTOR complex 1, and activation of
641   mTORC1 is associated with meningioma and schwannoma growth. *Molecular and cellular*
642   *biology* **29**, 4250-4261, doi:10.1128/mcb.01581-08 (2009).

643   44   Johnson, M. D., Okedli, E., Woodard, A., Toms, S. A. & Allen, G. S. Evidence for
644   phosphatidylinositol 3-kinase-Akt-p7S6K pathway activation and transduction of mitogenic
645   signals by platelet-derived growth factor in meningioma cells. *Journal of neurosurgery* **97**,
646   668-675, doi:10.3171/jns.2002.97.3.0668 (2002).

647   45   Weisman, A. S., Raguet, S. S. & Kelly, P. A. Characterization of the epidermal growth factor
648   receptor in human meningioma. *Cancer Res* **47**, 2172-2176 (1987).

649 46 Wu, J. I., Lessard, J. & Crabtree, G. R. Understanding the words of chromatin regulation. *Cell* **136**, 200-206, doi:10.1016/j.cell.2009.01.009 (2009).

651 47 Kia, S. K., Gorski, M. M., Giannakopoulos, S. & Verrijzer, C. P. SWI/SNF mediates polycomb eviction and epigenetic reprogramming of the INK4b-ARF-INK4a locus. *Mol Cell Biol* **28**, 3457-3464, doi:10.1128/MCB.02019-07 (2008).

654 48 Wilson, B. G. & Roberts, C. W. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* **11**, 481-492, doi:10.1038/nrc3068 (2011).

656 49 Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci Adv* **1**, e1500447, doi:10.1126/sciadv.1500447 (2015).

659 50 Lichner, Z. *et al.* The chromatin remodeling gene ARID1A is a new prognostic marker in clear cell renal cell carcinoma. *Am J Pathol* **182**, 1163-1170, doi:10.1016/j.ajpath.2013.01.007 (2013).

661 51 Luchini, C. *et al.* Prognostic role and implications of mutation status of tumor suppressor gene ARID1A in cancer: a systematic review and meta-analysis. *Oncotarget* **6**, 39088-39097, doi:10.18632/oncotarget.5142 (2015).

664 52 Le Loarer, F. *et al.* SMARCA4 inactivation defines a group of undifferentiated thoracic malignancies transcriptionally related to BAF-deficient sarcomas. *Nat Genet* **47**, 1200-1205, doi:10.1038/ng.3399 (2015).

667 53 Farshidfar, F. *et al.* Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. *Cell reports* **18**, 2780-2794, doi:10.1016/j.celrep.2017.02.033 (2017).

670 54 Liu, G. *et al.* Prognostic and Clinicopathological significance of ARID1A in endometrium-related gynecological cancers: A meta-analysis. *J Cell Biochem*, doi:10.1002/jcb.26109 (2017).

672 55 Goldbrunner, R. *et al.* EANO guidelines for the diagnosis and treatment of meningiomas. *Lancet Oncol* **17**, e383-391, doi:10.1016/S1470-2045(16)30321-7 (2016).

674 56 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

676 57 Shah, N. & Sukumar, S. The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* **10**, 361-371, doi:10.1038/nrc2826 (2010).

678 58 Xu, Q. *et al.* Long non-coding RNA regulation of epithelial-mesenchymal transition in cancer metastasis. *Cell Death Dis* **7**, e2254, doi:10.1038/cddis.2016.149 (2016).

680 59 Krumlauf, R. Hox genes in vertebrate development. *Cell* **78**, 191-201 (1994).

681 60 Xie, M. *et al.* Long noncoding RNA HOXA-AS2 promotes gastric cancer proliferation by epigenetically silencing P21/PLK3/DDIT3 expression. *Oncotarget* **6**, 33587-33601, doi:10.18632/oncotarget.5599 (2015).

684 61 Bao, X. *et al.* Knockdown of long non-coding RNA HOTAIR increases miR-454-3p by targeting Stat3 and Atg12 to inhibit chondrosarcoma growth. *Cell Death Dis* **8**, e2605, doi:10.1038/cddis.2017.31 (2017).

687 62 Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-1076, doi:10.1038/nature08975 (2010).

689 63 Kim, K. *et al.* HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* **32**, 1616-1625, (2013).

691 64 Li, X. *et al.* Long non-coding RNA HOTAIR, a driver of malignancy, predicts negative prognosis and exhibits oncogenic activity in oesophageal squamous cell carcinoma. *Br J Cancer* **109**, 2266-2278, doi:10.1038/bjc.2013.548 (2013).

694 65 Ozes, A. R. *et al.* NF-kappaB-HOTAIR axis links DNA damage response, chemoresistance and cellular senescence in ovarian cancer. *Oncogene* **35**, 5350-5361, doi:10.1038/onc.2016.75 (2016).

696 66 Shi, J. *et al.* Long non-coding RNA in glioma: signaling pathways. *Oncotarget*, doi:10.18632/oncotarget.15175 (2017).

698 67 Wu, X. *et al.* HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition. *Cancer Res* **66**, 9527-9534, doi:10.1158/0008-5472.CAN-05-4470 (2006).

701 68 Hayashida, T. *et al.* HOXB9, a gene overexpressed in breast cancer, promotes tumorigenicity and lung metastasis. *Proc Natl Acad Sci U S A* **107**, 1100-1105, doi:10.1073/pnas.0912710107 (2010).

25

703 69   Ding, J. *et al.* Long noncoding RNA HOXA-AS2 represses P21 and KLF2 expression transcription
704   by binding with EZH2, LSD1 in colorectal cancer. *Oncogenesis* **6**, e288,
705   doi:10.1038/oncsis.2016.84 (2017).
706 70   Zhao, H., Zhang, X., Frazao, J. B., Condino-Neto, A. & Newburger, P. E. HOX antisense lincRNA
707   HOXA-AS2 is an apoptosis repressor in all trans retinoic acid treated NB4 promyelocytic
708   leukemia cells. *J Cell Biochem* **114**, 2375-2383, doi:10.1002/jcb.24586 (2013).
709 71   Ponting, C. P., Oliver, P. L. & Reik, W. Evolution and functions of long noncoding RNAs. *Cell*
710   **136**, 629-641, doi:10.1016/j.cell.2009.02.006 (2009).
711 72   Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX
712   loci by noncoding RNAs. *Cell* **129**, 1311-1323, doi:10.1016/j.cell.2007.05.022 (2007).
713 73   Khalil, A. M. *et al.* Many human large intergenic noncoding RNAs associate with
714   chromatin-modifying complexes and affect gene expression. *Proceedings of the National Academy
715   of Sciences of the United States of America* **106**, 11667-11672, doi:10.1073/pnas.0904715106
716   (2009).
717 74   Helming, K. C., Wang, X. & Roberts, C. W. Vulnerabilities of mutant SWI/SNF complexes in
718   cancer. *Cancer Cell* **26**, 309-317, doi:10.1016/j.ccr.2014.07.018 (2014).
719 75   Kim, K. H. *et al.* SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of
720   EZH2. *Nat Med* **21**, 1491-1496, doi:10.1038/nm.3968 (2015).
721 76   Bitler, B. G. *et al.* Synthetic lethality by targeting EZH2 methyltransferase activity in
722   ARID1A-mutated cancers. *Nat Med* **21**, 231-238, (2015).
723 77   Kim, K. H. & Roberts, C. W. Targeting EZH2 in cancer. *Nat Med* **22**, 128-134,
724   doi:10.1038/nm.4036 (2016).
725 78   Ozes, A. R. *et al.* Therapeutic targeting using tumor specific peptides inhibits long non-coding
726   RNA HOTAIR activity in ovarian and breast cancer. *Scientific reports* **7**, 894,
727   doi:10.1038/s41598-017-00966-3 (2017).
728 79   Pfister, S. X. & Ashworth, A. Marked for death: targeting epigenetic changes in cancer. *Nature
729   reviews. Drug discovery* **16**, 241-263, doi:10.1038/nrd.2016.256 (2017).
730 80   Gotwals, P. *et al.* Prospects for combining targeted and conventional cancer therapy with
731   immunotherapy. *Nat Rev Cancer* advance online publication, doi:10.1038/nrc.2017.17 (2017).
732 81   Marcucci, F., Stassi, G. & De Maria, R. Epithelial-mesenchymal transition: a new target in
733   anticancer drug discovery. *Nature reviews. Drug discovery* **15**, 311-325, doi:10.1038/nrd.2015.13
734   (2016).
735 82   Norden, A. D. *et al.* Phase II trials of erlotinib or gefitinib in patients with recurrent meningioma. *J
736   Neurooncol* **96**, 211-217, doi:10.1007/s11060-009-9948-7 (2010).
737 83   Byers, L. A. *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to
738   EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR
739   inhibitor resistance. *Clinical cancer research : an official journal of the American Association for
740   Cancer Research* **19**, 279-290, doi:10.1158/1078-0432.ccr-12-1558 (2013).
741 84   Buonato, J. M. & Lazzara, M. J. ERK1/2 blockade prevents epithelial-mesenchymal transition in
742   lung cancer cells and promotes their sensitivity to EGFR inhibition. *Cancer Res* **74**, 309-319,
743   doi:10.1158/0008-5472.can-12-4721 (2014).
744 85   Thomson, S., Petti, F., Sujka-Kwok, I., Epstein, D. & Haley, J. D. Kinase switching in
745   mesenchymal-like non-small cell lung cancer lines contributes to EGFR inhibitor resistance
746   through pathway redundancy. *Clinical & experimental metastasis* **25**, 843-854,
747   doi:10.1007/s10585-008-9200-4 (2008).
748 86   Marcucci, F., Stassi, G. & De Maria, R. Epithelial-mesenchymal transition: a new target in
749   anticancer drug discovery. *Nat Rev Drug Discov* **15**, 311-325, doi:10.1038/nrd.2015.13 (2016).
750 87   Chmielecki, J. *et al.* Whole-exome sequencing identifies a recurrent NAB2-STAT6 fusion in
751   solitary fibrous tumors. *Nat Genet* **45**, 131-132, doi:10.1038/ng.2522 (2013).
752 88   Gao, F. *et al.* Inversion-mediated gene fusions involving NAB2-STAT6 in an unusual malignant
753   meningioma. *Br J Cancer* **109**, 1051-1055, doi:10.1038/bjc.2013.395 (2013).
754 89   Schweizer, L. *et al.* Meningeal hemangiopericytoma and solitary fibrous tumors carry the
755   NAB2-STAT6 fusion and can be diagnosed by nuclear expression of STAT6 protein. *Acta
756   Neuropathol* **125**, 651-658, doi:10.1007/s00401-013-1117-6 (2013).
757 90   Soda, M. *et al.* Identification of the transforming EML4-ALK fusion gene in non-small-cell lung
758   cancer. *Nature* **448**, 561-566, doi:10.1038/nature05945 (2007).

26

759 91   Kozarewa, I. *et al.* Amplification-free Illumina sequencing-library preparation facilitates improved
760   mapping and assembly of (G+C)-biased genomes. *Nature methods* **6**, 291-295,
761   doi:10.1038/nmeth.1311 (2009).

762 92   Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29**,
763   15-21, doi:10.1093/bioinformatics/bts635 (2013).

764 93   Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput
765   sequencing data. *Bioinformatics (Oxford, England)* **31**, 166-169,
766   doi:10.1093/bioinformatics/btu638 (2015).

767 94   McPherson, A. *et al.* deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data.
768   *PLoS computational biology* **7**, e1001138, doi:10.1371/journal.pcbi.1001138 (2011).

769 95   Kim, D. & Salzberg, S. L. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts.
770   *Genome biology* **12**, R72, doi:10.1186/gb-2011-12-8-r72 (2011).

771 96   Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463**,
772   893-898, doi:10.1038/nature08768 (2010).

773 97   Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome biology*
774   **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).

775 98   Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
776   RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

777 99   Peixoto, L. *et al.* How data analysis affects power, reproducibility and biological insight of
778   RNA-seq studies in complex datasets. *Nucleic acids research* **43**, 7664-7674,
779   doi:10.1093/nar/gkv736 (2015).

780 100   Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of
781   expression residuals (PEER) to obtain increased power and interpretability of gene expression
782   analyses. *Nat Protoc* **7**, 500-507, doi:10.1038/nprot.2011.457 (2012).

783 101   Wang, X. *et al.* Analysis of gene expression profiling in meningioma: deregulated signaling
784   pathways associated with meningioma and EGFL6 overexpression in benign meningioma tissue
785   and serum. *PLoS One* **7**, e52707, doi:10.1371/journal.pone.0052707 (2012).

786 102   Savvidis, C. & Koutsilieris, M. Circadian rhythm disruption in cancer biology. *Mol Med* **18**,
787   1249-1260, doi:10.2119/molmed.2012.00077 (2012).

788 103   Sharma, S., Ray, S., Moiyadi, A., Sridhar, E. & Srivastava, S. Quantitative proteomic analysis of
789   meningiomas for the identification of surrogate protein markers. *Sci Rep* **4**, 7140,
790   doi:10.1038/srep07140 (2014).

791 104   Pfister, C., Tatabiga, M. S. & Roser, F. Selection of suitable reference genes for quantitative
792   real-time polymerase chain reaction in human meningiomas and arachnoidea. *BMC Res Notes* **4**,
793   275, doi:10.1186/1756-0500-4-275 (2011).

794 105   Luo, W., Friedman, M. S., Shedden, K., Hankenson, K. D. & Woolf, P. J. GAGE: generally
795   applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161,
796   doi:10.1186/1471-2105-10-161 (2009).

797 106   Iorio, F. *et al.* Population-level characterization of pathway alterations with SLAPenrich dissects
798   heterogeneity of cancer hallmark acquisition. *bioRxiv doi: 10.1101/077701* (2016).

799 107   Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated
800   aggressive human tumors. *Nat Genet* **40**, 499-507, doi:10.1038/ng.127 (2008).

801 108   Sarrio, D. *et al.* Epithelial-mesenchymal transition in breast cancer relates to the basal-like
802   phenotype. *Cancer Res* **68**, 989-997, doi:10.1158/0008-5472.can-07-2017 (2008).

803 109   Wong, D. J. *et al.* Module map of stem cell genes guides creation of epithelial cancer stem cells.
804   *Cell stem cell* **2**, 333-344, doi:10.1016/j.stem.2008.02.009 (2008).

805 110   The Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic acids*
806   *research* **43**, D1049-D1056, doi:10.1093/nar/gku1179 (2015).

807 111   R: A language and environment for statistical computing (R Foundation for Statistical Computing,
808   Vienna, Austria, 2016).

809 112   Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform.
810   *Bioinformatics (Oxford, England)* **25**, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

811 113   Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to
812   detect break points of large deletions and medium sized insertions from paired-end short reads.
813   *Bioinformatics (Oxford, England)* **25**, 2865-2871, doi:10.1093/bioinformatics/btp394 (2009).

814 114 Raine, K. M. *et al.* cgpPindel: Identifying Somatically Acquired Insertion and Deletion Events
815   from Paired End Sequencing. *Current protocols in bioinformatics* **52**, 15.17.11-12,
816   doi:10.1002/0471250953.bi1507s52 (2015).
817 115 Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome
818   sequences. *Nature* **534**, 47-54, doi:10.1038/nature17676 (2016).
819 116 Raine, K. M. *et al.* ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from
820   Whole-Genome Sequencing Data. *Current protocols in bioinformatics* **56**, 15.19.11-15.19.17,
821   doi:10.1002/cpbi.17 (2016).
822 117 Sun, L., Craiu, R. V., Paterson, A. D. & Bull, S. B. Stratified false discovery control for large-scale
823   hypothesis testing with application to genome-wide association studies. *Genet Epidemiol* **30**,
824   519-530, doi:10.1002/gepi.20164 (2006).
825 118 Buels, R. *et al.* JBrowse: a dynamic web platform for genome visualization and analysis. *Genome*
826   *biology* **17**, 66, doi:10.1186/s13059-016-0924-1 (2016).
827 119 Chang, M. T. *et al.* Identifying recurrent mutations in cancer reveals widespread lineage diversity
828   and mutational specificity. *Nat Biotechnol* **34**, 155-163, doi:10.1038/nbt.3391 (2016).
829 120 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting
830   genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550,
831   doi:10.1073/pnas.0506580102 (2005).
832 121 Corver, W. E. *et al.* Genome haploidisation with chromosome 7 retention in oncocytic follicular
833   thyroid carcinoma. *PloS one* **7**, e38287, doi:10.1371/journal.pone.0038287 (2012).
834
835

Author Contributions

G.C. and N.K. performed mRNA expression analysis. P.T. and G.C analysed whole genome and targeted sequencing data. I.M. performed statistical analyses to detect novel driver mutations. S.M. analysed methylation array data. F.M. generated mutational signatures analysis. J.T. and M.C. performed retrotransposon analysis. C.O.H. and J.D. performed protein expression analysis. A.B, S.B. and M.Y. contributed to data analysis strategy. A.Y., T.N., G.R.B and J.T. provided informatic support. T.S., R.W.K., M.K, G.S., D.P., A.D., C.E.M., A.Y., I.N., S.J.P., C.W., Z.R., M.D.J., R.Z., and K. S. provided samples and clinical data. S.B., G.S.V, I.N. and M.W.M. provided conceptual advice. V.P.C and K.A carried out central pathology review. U.M. and T.S. devised and supervised the project. U.M. G.C. and T.S. wrote the manuscript with input from S.B., P.T., and G.S.V. All authors approved the manuscript.

863   **Competing Financial Interests**: All authors declare no competing financial interests.

864

865   Correspondence and requests for materials should be addressed to U.M.

866   (um1@sanger.ac.uk) and T.S. (ts381@cam.ac.uk)
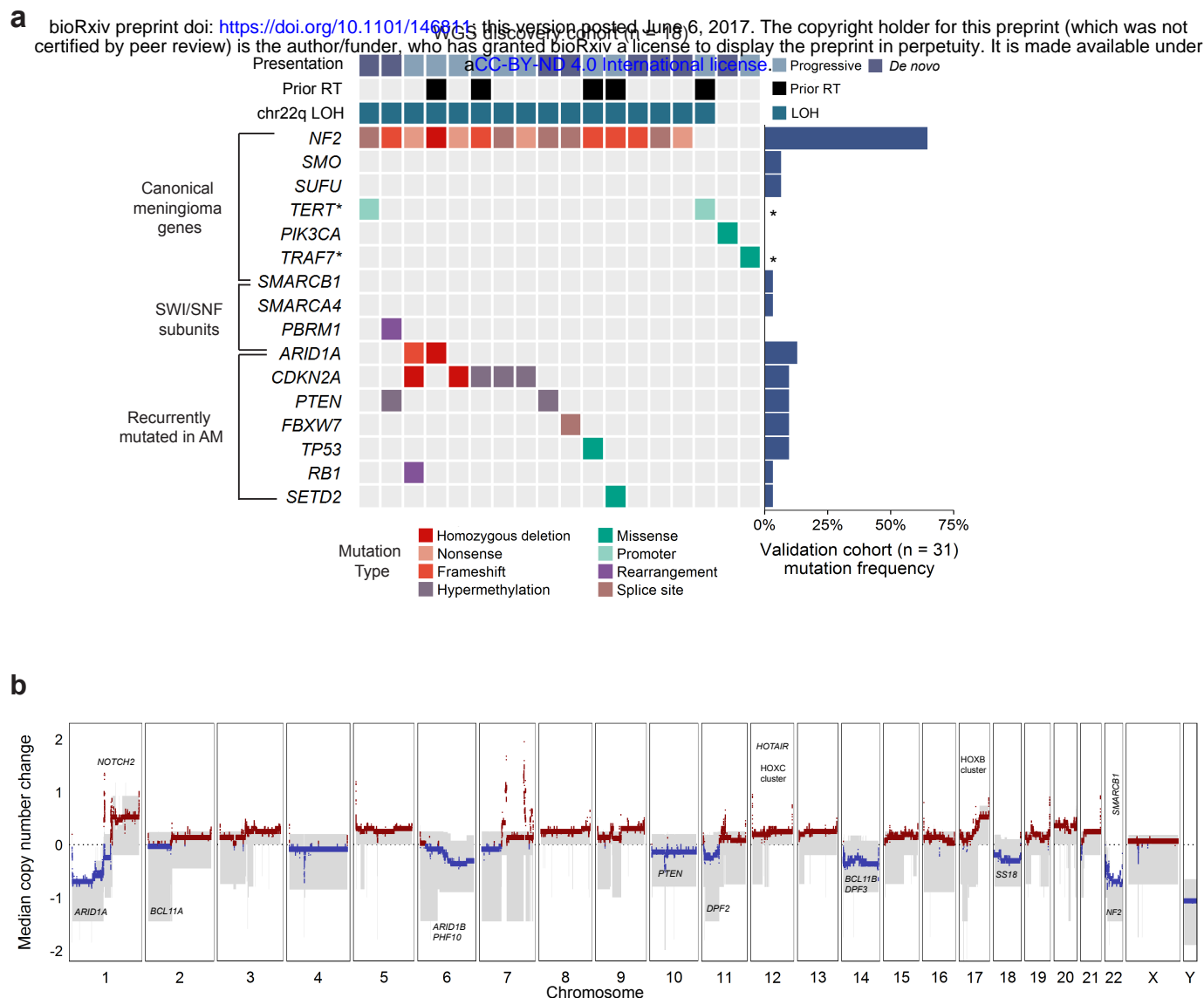
Figure 1

a

b

Figure 1 | The landscape of driver mutations and copy number alterations in anaplastic meningioma. (a) The landscape of somatic driver variants in primary anaplastic meningioma. Somatic mutation and promoter methylation data is shown for a discovery cohort of 18 primary tumours characterised by whole genome sequenceing. Mutations in recurrently altered genes, established meningioma genes and SWI/SNF complex subunits are included. Samples are annotated for chromosome 22q LOH, prior radiotherapy exposure, and clinical presentation (*de novo* verus progression from a lower grade meningioma). The barchart to the right indicates mutation frequency in a validation cohort of 31 primary tumors sequenced with a 366 cancer gene panel. Asterisks indicate genes not included in the targeted sequencing assay. **(b)** Aggregate copy number profile of primary anaplastic meningioma. For the 18 tumors charactersied by whole genome sequencing, the median relative copy number change was calculated across the genome in 10 kilobase segments, adjusting for ploidy. The grey shaded area indicates the first and third quantile of copy number for each genomic segment. The solid red and blue lines represent the median relative copy number gain and loss, respectively, with zero indicating no copy number change. X-axis: Chromosomal position. Y-axis: median relative copy number change. Potential target genes are noted. AM, anaplastic meningioma; LOH, loss of heterozygosity; RT, radiotherapy.
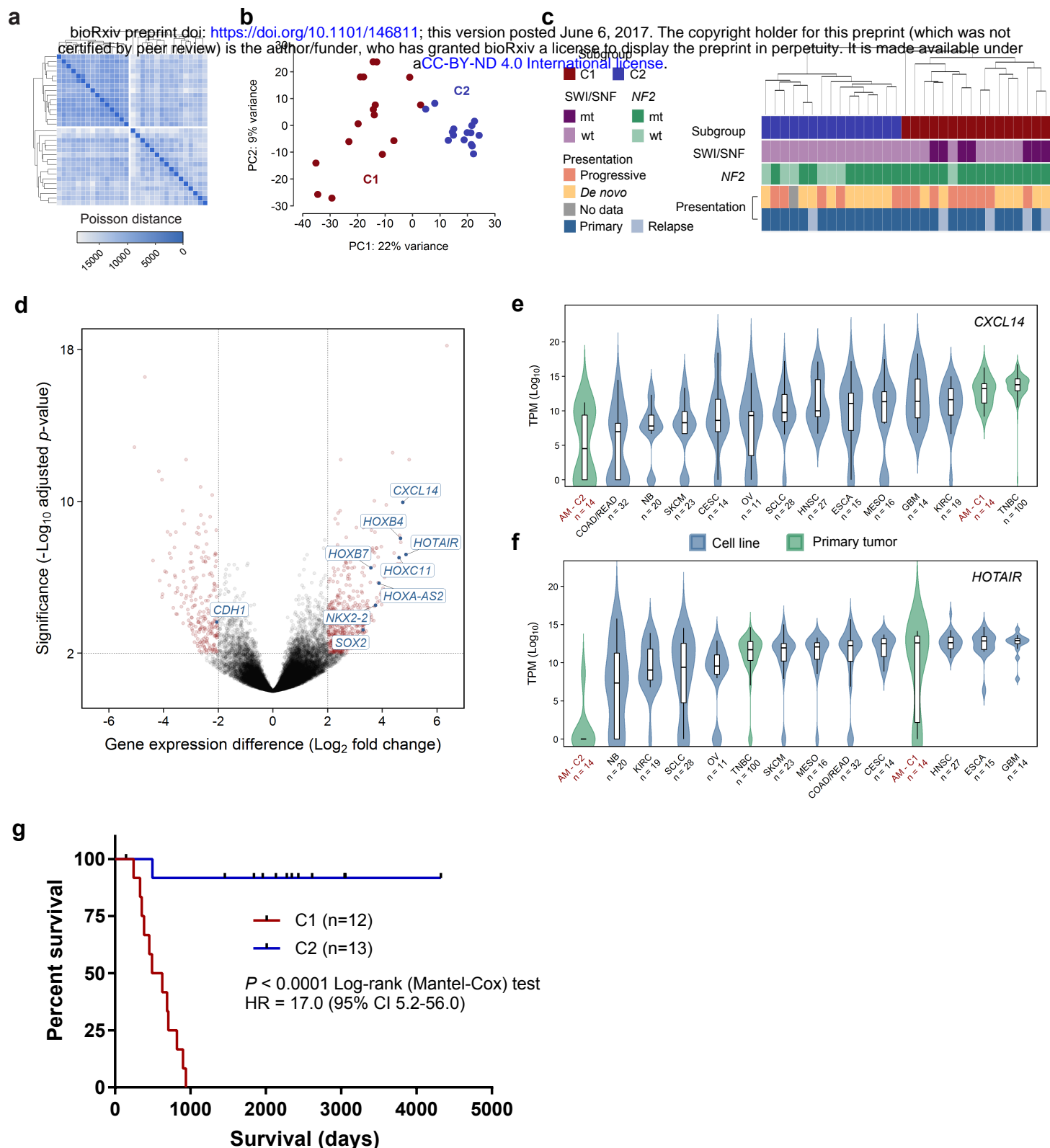
Figure 2

Figure 2 | Transcriptomic classification of anaplastic meningioma. (a) Unsupervised hierarchical clustering and **(b)** Principal component analysis of anaplastic meningioma gene expression revealed two subgroups (denoted C1 and C2). **(c)** Dendrogram obtained by unsupervised clustering annotated with clinical and genomic features. **(d)** Volcano plot depicting genes differentially expressed between C1 versus C2 anaplastic meningioma samples. The horizontal axis shows the $\log_2$ fold change and the vertical axis indicates the -$\log_{10}$ adjusted p-value. Genes with an adjusted p-value < 0.01 and absolute $\log_2$ fold change > 2 are highlighted in red with particular genes of interest indicated. **(e, f)** Box plots of (e) CXLC14 and (f) HOTAIR expression across 31 anaplastic meningomas classified into C1 and C2 subgroups, 100 primary breast tumors, and 219 cancer cell lines from 11 tumor types. Upper and lower box hinges correspond to first and third quartiles, horizontal line and whiskers indicate the median and 1.5-fold the interquarntile range, respectively. Underlying violin plots show data distribution and are color-coded according to specimen source (blue, cell line; green, pimary tumor). X-axis indicates tumor type and number of samples in cohort. Y-axis shows NF2 $\log_{10}$ TPM values. **(g)** Kaplan-Meier curves showing overall survival for 25 (of 28) anaplastic meningioma patients in C1 and C2 subgroups for whom follow-up data was available. Dashes indicate timepoints at which subjects were censored at time of last follow-up. TPM, transcripts per kilobase million; AM, anaplastic meningioma; TNBC, triple negative breast carcinoma; wt, wild-type; mt, mutated; HR, hazard ratio; CI, confidence interval; PC, principal component.
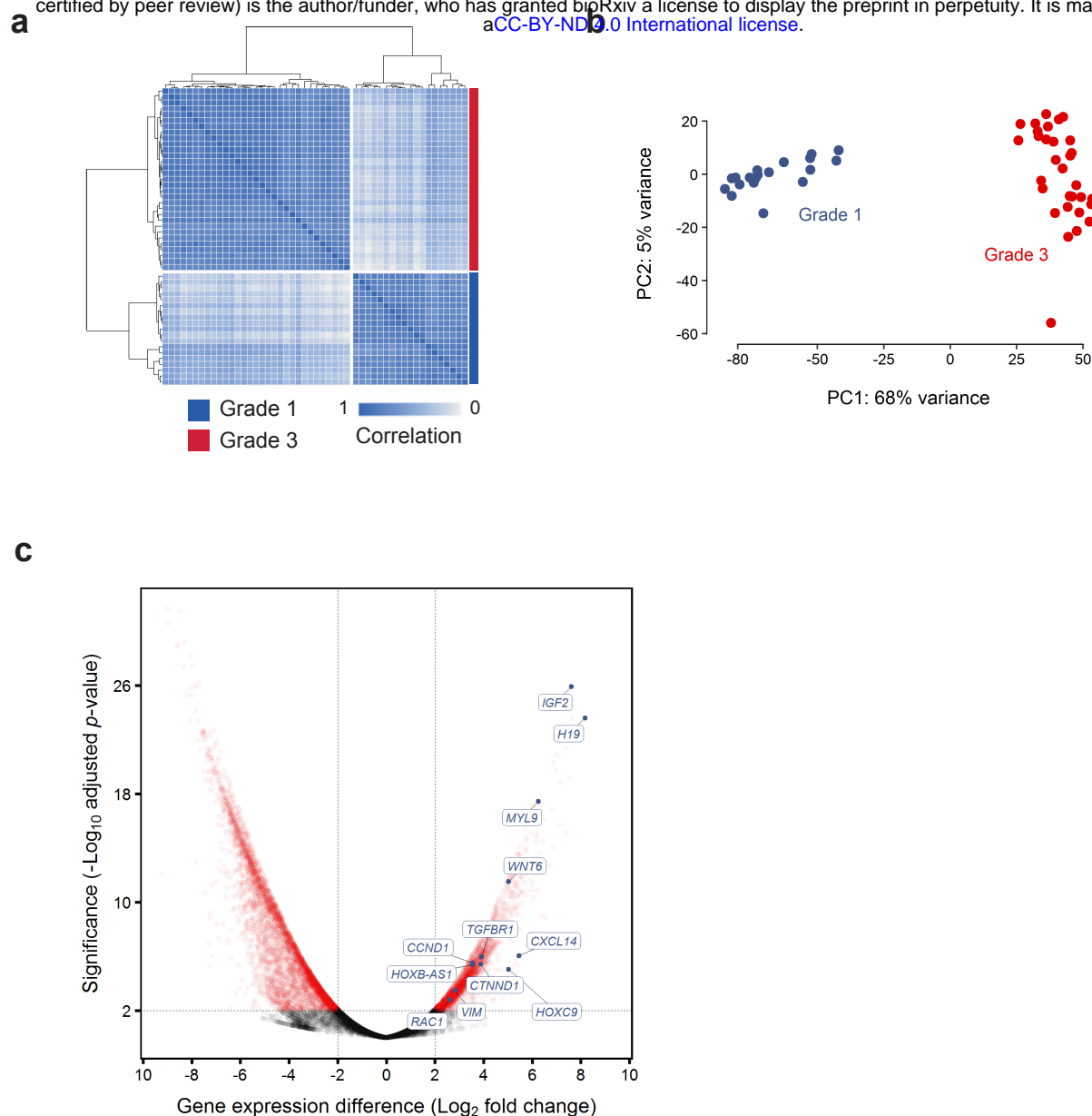
Figure 3

a

b

c

Figure 3 | Differences in gene expression profile between grade 1 and anaplastic meningomas. **(a, b)** Normalised transcript counts from grade 1 and anaplastic meningioma samples clustered by **(a)** Pearson's correlation coefficient and **(b)** principal component analysis. **(c)** Volcano plot illustrating differences in gene expression between anaplastic versus grade 1 meningiomas with selected genes indicated. The horizontal axis shows the $\log_2$ fold change and the vertical axis indicates the $-\log_{10}$ adjusted p-value. Genes with an adjusted p-value < 0.01 and absolute $\log_2$ fold change > 2 are highlighted in red. PC, principal component.