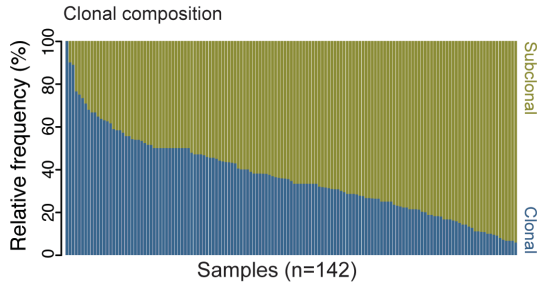
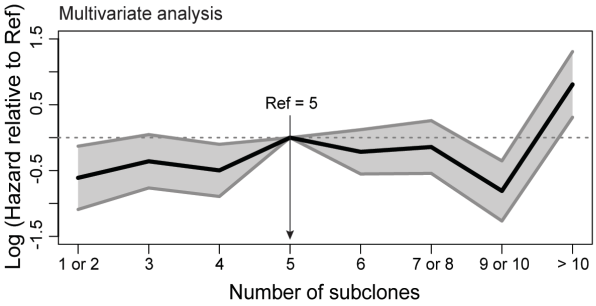
Supplementary Figure S1

A Subclonal architecture



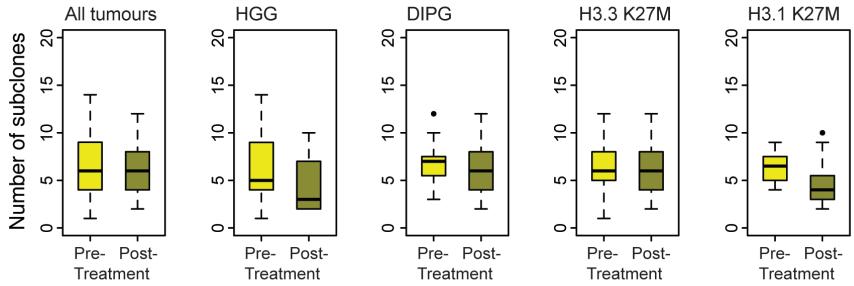
B Overall survival



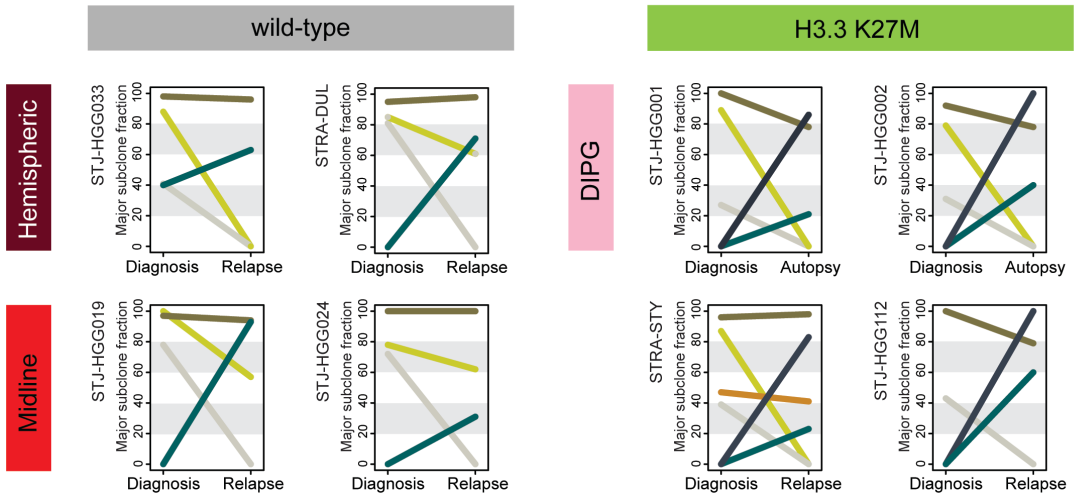
C Cox proportional hazard model

		Hazard ratio	p value
Histone	H3.3 G34R/V	0.7345 (0.5179)	0.8551
	H3.3 K27M	**1.8613 (0.2476)**	**0.0121**
	H3.1 K27M	0.6498 (0.3053	0.1579
	Wild-type	1	
Location	Pons	1.7107 (0.4371)	0.2193
	Midline	1.2274 (0.4394)	0.6410
	Hemispheric	1	
Age		0.9608 (0.0265)	0.1318
No. subclones	>10	**2.9677 (0.4114)**	**0.0082**
	≤10	1	

D Sample timing

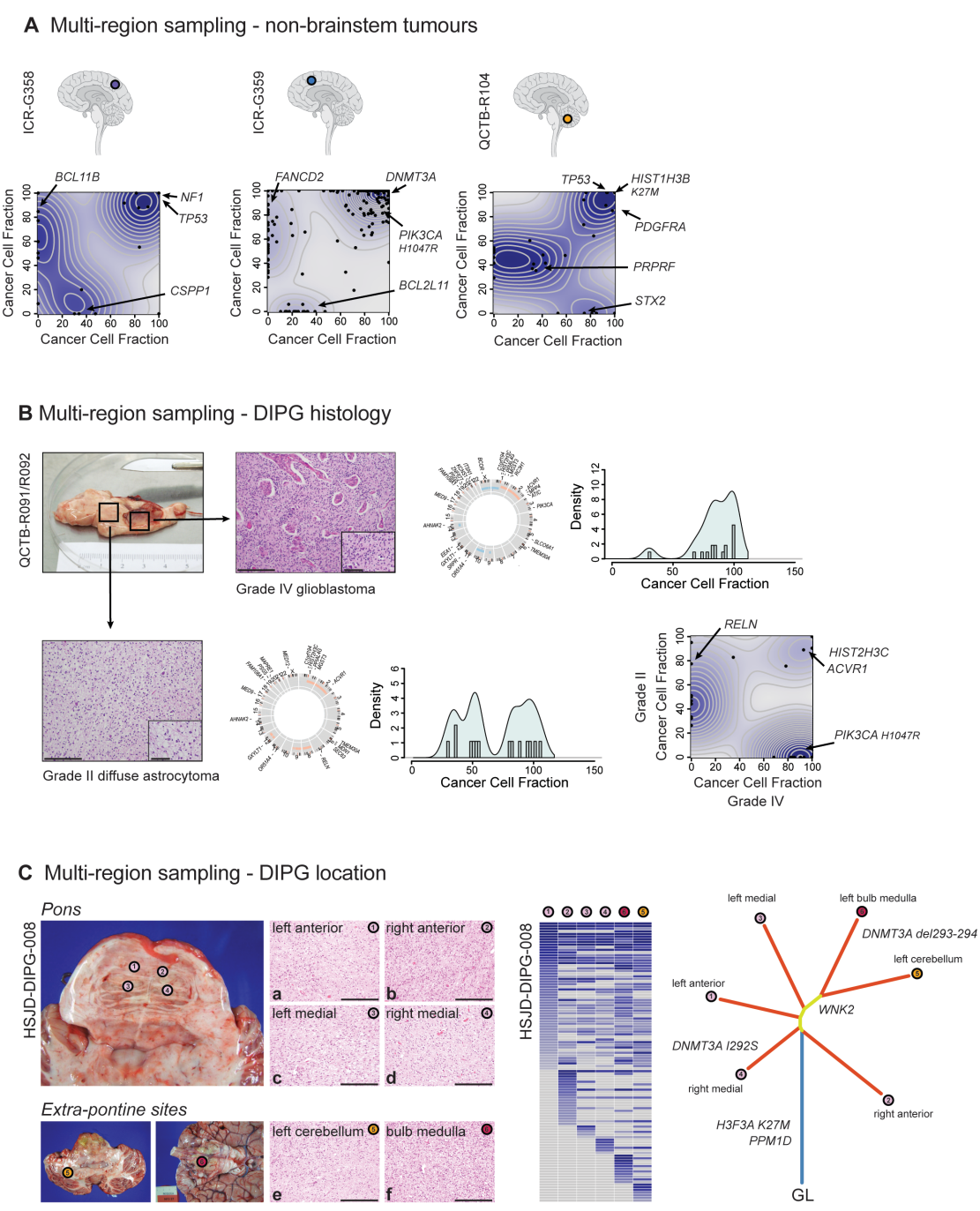


E Longitudinal sampling



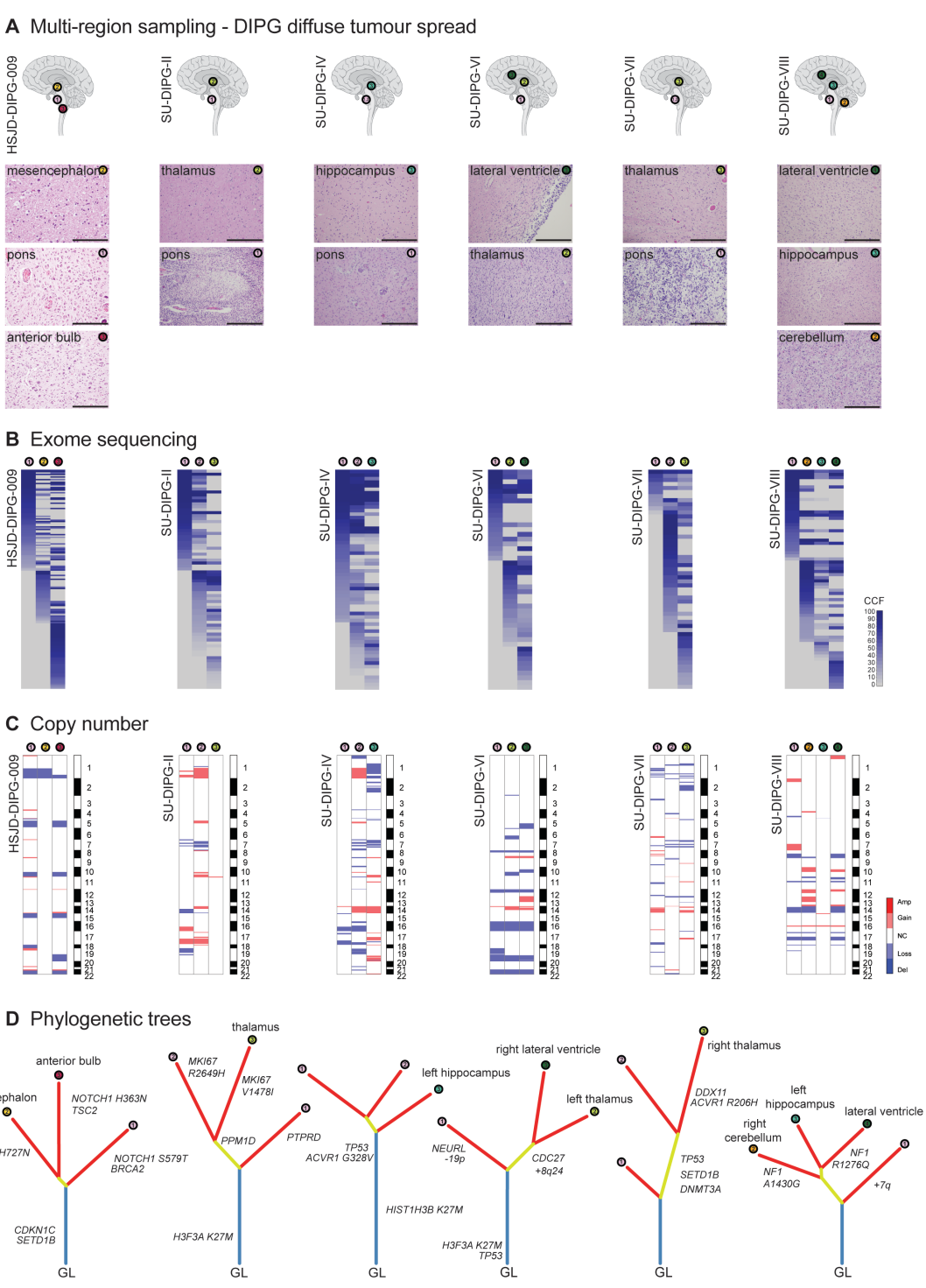
Supplementary Figure S1 – *Prognostic implication of subclonal diversity.* (A) Subclonal architecture. Percentage of somatic variants classified as clonal (blue) or subclonal (dark yellow) are plotted for each of 142 cases, ordered by decreasing clonality. (B) Overall survival. Hazard ratios for the number of subclones (x axis) present in n=142 samples in respect of overall survival are plotted on a log scale (y axis) relative to the reference, in this case n=5 subclones. Shaded areas equals standard error. (C) Cox proportional hazards model. Hazard ratios (with standard errors in brackets) and p values calculate by the Cox proportional hazards model in respect of overall survival for histone H3 mutation (H3.3 G34R/V, n=10; H3.3 K27M, n=61; H3.1 K27M, n=23; wild-type, n=48), anatomical location (pons, n=93; midline, n=20; hemispheric, n=29), age (n=142), and number of subclones (>10, n=8; ≤10, n=134). (D) Sample timing. Boxplots for number of subclones (y axis) in tumours stratified by whether sampling was pre- or post-treatment. No significant differences by ANOVA were observed when all tumours were considered together (n=142), DIPG (n=93) and non-brainstem tumours (n=49) were separated, or when segregated by histone H3 mutation subgroup (H3.3 K27M, n=61; H3.1 K27M, n=23; there were insufficient numbers of post-treatment samples for statistical analysis with H3.3 G34R). The thick line within the box is the median, the lower and upper limits of the boxes represent the first and third quartiles, the whiskers 1.5x the interquartile range, and individual points outliers. (E) Longitudinal paired sampling. Clusters of mutant allele frequencies are aggregated into major subclones. Changes in the tumour fractions of these clusters after treatment are assessed in eight paired samples taken at diagnosis and relapse/autopsy. Coloured lines link related major subclones pre- and post-treatment. Samples are arranged by anatomical location and histone H3 mutation subgroup.

Supplementary Figure S2

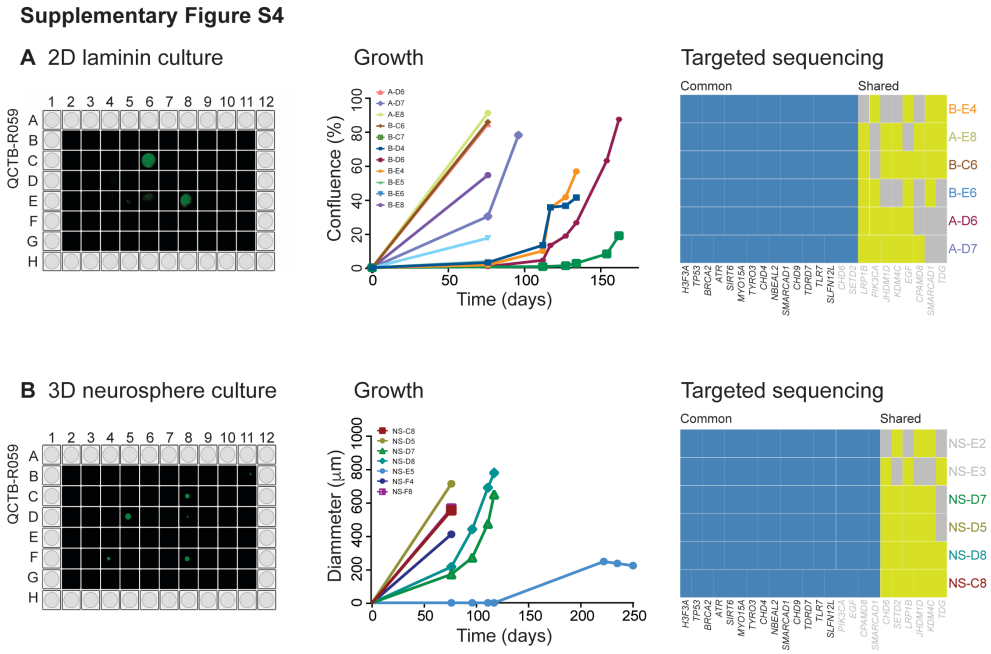


Supplementary Figure S2 - *Multiregion sampling.* (A) Biplots of cancer cell fractions for somatic mutations from distinct samples from the same case for three patients with GBM in differing anatomical locations (labelled). Kernel density heatmap overplotted and specific mutations present either ubiquitously (e.g. *NF1*, *DNMT3A*, *HIST1H3B*) or in one region and not the other (e.g. *BCL11B*, *FANCD2*, *STX2*) are labelled for each case. (B) Multi-region sample based on WHO grade. DIPG sample QCTB-R091/R092 was dissected at autopsy into two regions, one classified as WHO grade IV glioblastoma, the other as WHO grade II diffusely infiltrating glioma, as seen on H&E staining. Scale bar = 100μm (inset scale bar = 25μm). Both samples were exome sequenced, and the CIRCOS plot provided (somatic SNVs and InDels, outer ring; DNA copy number changes (red=gain, blue=loss) and loss of heterozygosity (yellow), inner rings). Cancer cell fraction plots for each component highlight a complex subclonal architecture for both, and the CCF biplot and kernel density heatmap highlight ubiquitous somatic mutations such as in *HIST2H3C* and *ACVR1*, as well as distinct alterations such as *PIK3CA* in the higher grade, and *RELN* in the lower grade area. (C) Multi-region sample based on location. Six different regions of HSJD-DIPG-008 were sampled *post-mortem,* from within and outside the pons. Common, shared and private mutations across all regions were observed, as shown in the contingency heatmap of CCFs and the phylogenetic tree reconstructed using neighbour-joining algorithms based upon the nested subpopulation phylogenies calculated as part of EXPANDS. Ubiquitous clonal mutations such as H3.3 K27M and *PPM1D* could be placed on the trunk of the tree, whilst convergent evolution of distinct mutations in *DNMT3A* were seen in different regions. Extra-pontine sites acquired mutations in *WNK2*. Scale bar = 100μm.

Supplementry Figure S3

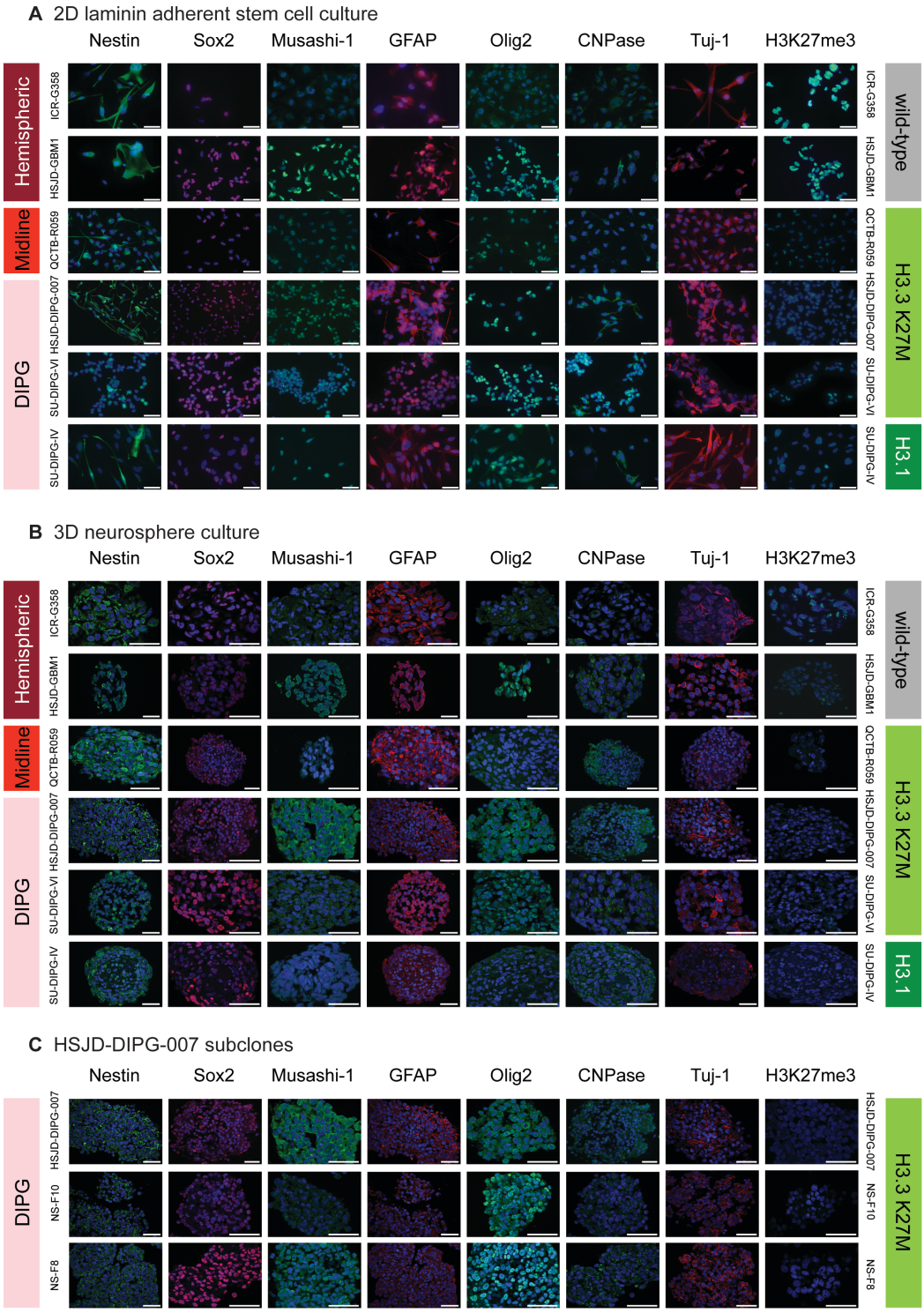


Supplementary Figure S3 – *Multi-region samples of DIPG diffuse spread* (A) Multi-region sampling based on diffuse spread of DIPG cells throughout the brain. Representative H&E images from distinct anatomical locations of six DIPG cases collect at autopsy. Regions are colour-coded and labelled. Scale bar = 100μm. (B) Contingency plots for somatic mutations in topographically distinct regions from six DIPG autopsy specimens. CCFs are plotted for each variant in each sample, with anatomical location highlighted and colour-coded. (C) Copy number profiles for multi-region autopsy samples showing diffuse spread throughout the brain. Red = gain, blue = loss, white = no change. (D) Branching evolution and genotypic convergence. Phylogenetic trees built using maximum parsimony from exome sequencing data. In most cases distant sites appear to escape from pons early in tumour evolution, and evidence of convergent evolution of distinct SNVs in key genes is apparent.



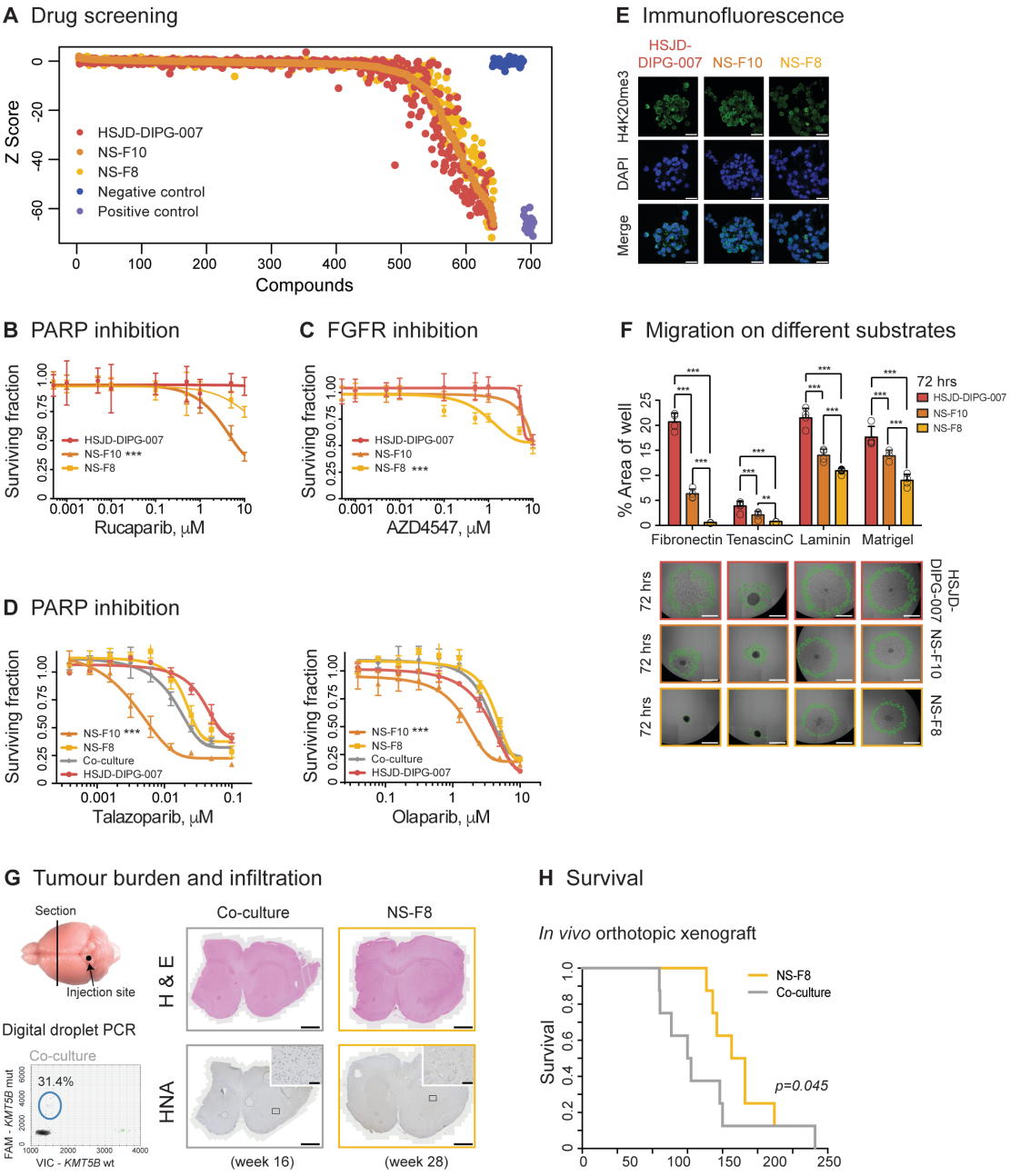
Supplementary Figure S4 – *QCTB-R059 single cell-derived colonies grown on laminin and as neurospheres.* (A) 2D laminin and (B) 3D neurosphere culture of single cell-derived colonies from QCTB-R059 assessed by Celigo S imaging cytometer; time-course of growth of individual colonies, labelled and colour-coded; targeted sequencing contingency plot of somatic mutations identified as common to all subclones (blue), and shared amongst certain subclones (yellow).

Supplementary Figure S5



Supplementary Figure S5 – *Immunophenotyping of in vitro cultures*. (A) 2D laminin cultures of pGBM and DIPG. Immunofluorescence for a range of stem cell and differentiation markers in six pGBM and DIPG primary patient-derived cells established as adherent cultures on laminin. Nestin, Musashi-1 (both stem cell markers), Olig2 (oligodendrocytes), CNPase (neurons) are labelled green; SOX2 (stem cell), GFAP (astrocytes), Tuj-1 (neurons) are labelled red. H3K27me3 is also assessed, labelled green. In all cases, nuclei are stained with DAPI (blue). Cells are labelled as to the anatomical location from which they were derived, and their histone H3 subgroup. (B) 3D neurosphere cultures of pGBM and DIPG. Immunofluorescence for a range of stem cell and differentiation markers in six pGBM and DIPG primary patient derived cells established as 3D neurospheres. Nestin, Musashi-1 (both stem cell markers), Olig2 (oligodendrocytes), CNPase (neurons) are labelled green; SOX2 (stem cell), GFAP (astrocytes), Tuj-1 (neurons) are labelled red. H3K27me3 is also assessed, labelled green. In all cases, nuclei are stained with DAPI (blue). Cells are labelled as to the anatomical location from which they were derived, and their histone H3 subgroup. (C) 3D neurosphere cultures of HSJD-DIPG-007 bulk cells and subclones. Immunofluorescence for a range of stem cell and differentiation markers in heterogeneous HSJD-DIPG-007 bulk cells, subclones NS-F10 and NS-F8. Nestin, Musashi-1 (both stem cell markers), Olig2 (oligodendrocytes), CNPase (neurons) are labelled green; SOX2 (stem cell), GFAP (astrocytes), Tuj-1 (neurons) are labelled red. H3K27me3 is also assessed, labelled green. In all cases, nuclei are stained with DAPI (blue). Cells are labelled as to the anatomical location from which they were derived, and their histone H3 subgroup. All representative images taken from n=3 independent experiments. All scale bars = 50μm.

Supplementary Figure S6



Supplementary Figure S6 – *Functional assessment of HSJD-DIPG-007 subclones*. (A) Drug screening of HSJD-DIPG-007 bulk cells and subclones. An 80 compound drug library was screened against heterogeneous HSJD-DIPG-007 bulk cells and subclones grown as neurospheres, with cell viability as the read-out. Median Z scores across all replicates for each set of cells is plotted, ordered by decreasing value in NS-F10. Negative controls (DMSO) are labelled blue, positive controls (staurosporine) purple. Screens were carried out in n=3 independent experiments. Validation of differential sensitivity of subclones to (B) PARP inhibitor (rucaparib, AG-14699) and (C) FGFR inhibitor (AZD4547) are shown, with effect on cell viability (surviving fraction on y axes) of treatment of heterogeneous bulk cells and subclones with increasing concentrations of drug (x axes, $\log_{10}$ scale). ANOVA was used to test for significance of NS-F10 *versus* NS-F8 and HSJD-DIPG-007 bulk culture for rucaparib, and for NS-F8 *versus* NS-F10 and HSJD-DIPG-007 bulk culture for AZD4547. *** all p values <0.001. Data derived from n=3 independent experiments. (D) Co-culture. In addition to HSJD-DIPG-007 bulk cells and subclones cultured alone, NS-F10 and NS-F8 were cultured mixed in equal ratios and exposed to PARP inhibitors talazoparib and olaparib. Effect on cell viability (surviving fraction on y axes) of treatment of cells with increasing concentrations of drug (x axes, $\log_{10}$ scale). ANOVA was used to test for significance of NS-F10 *versus* NS-F8, co-culture and HSJD-DIPG-007 bulk culture for talazoparib and olaparib. *** all p values <0.001. Data derived from n=3 independent experiments. (E) Immunofluorescence of HSJD-DIPG-007 bulk cells and subclones for H4K20me3. Heterogeneous bulk HSJD-DIPG-007 cells and subclones were stained using an antibody directed against H4K20me3 (green) with nuclei stained with DAPI (blue). Representative images taken from n=3 independent experiments. Scale bar = 50μm. (F) Migration of HSJD-DIPG-007 bulk cells and subclones on different substrates. Migration of HSJD-DIPG-007 bulk cells and subclones NS-F10 and NS-F8 onto fibronectin, tenascin-C, laminin and matrigel at 72 hours, calculated as percentage of total area in the well

covered by migrating cells using the Celigo S cytometer. Representative images are given for each substrate at 72 hours, with extent of tumour cell migration marked in green. Data derived and representative images taken from n=3 independent experiments. All comparisons carried out by ANOVA, **p<0.01. ***p<0.001. Scale bar = 500μm. (G) Tumour burden and infiltration. Co-cultured NS-F10/NS-F8 subclones and NS-F8 alone were implanted directly into the pons of NOD-SCID mice and tumours allowed to form over 8 months. At week 16, co-cultured cells were found in mixed populations by ddPCR and anti-HNA staining to be present diffusely infiltrated throughout the CNS (section indicated by line on mouse brain image, along with injection site), by contrast with F8 cells at 28 weeks. Representative images from a total of n=8 mice per group. Scale bar = 1000μm (inset scale bar = 50μm). (H) Survival. Tumour-bearing animals implanted with co-culture NS-F10/NS-F8 subclones had significantly shorter survival than NS-F8 alone (p=0.045, log-rank test, n=8 mice per group). All graphs represent mean +/- standard deviation.

Legends for supplementary tables

Supplementary Table S1 – *Cancer cell fractions (CCF) for 142 pGBM and DIPG*. All somatic single nucleotide variants (SNVs) and small insertions and deletions (InDels) are provided upon re-analysis of whole genome or exome sequencing data from 142 published pGBM and DIPG, with details of variant and calculated CCF.

Supplementary Table S2 – *Clinical data and number of subclones.* Clinicopathological data for 142 published pGBM and DIPG, along with number of somatic mutations and number of subclones, the latter calculated by the EXPANDS package.

Supplementary Table S3 – *Cancer cell fractions for multiregion sampling*. All somatic single nucleotide variants (SNVs) and small insertions and deletions (InDels), along with CCFs, are provided from exome sequencing of 36 topographically distinct samples from 12 individual pGBM / DIPG patients.

Supplementary Table S4 – *Targeted sequencing of single cell-derived colonies.* Details of somatic single nucleotide variants (SNVs) and small insertions and deletions (InDels) from a custom panel of 435 genes recurrently mutated in pGBM and DIPG (along with additional histone genes), assessed in all single cell-derived colonies from SU-DIPG-VI, QCTB-R059 and HSJD-DIPG-007.

Supplementary Table S5 – *Drug screening of HSJD-DIPG-007.* A library of 80 compounds at varying concentrations were screened against heterogeneous bulk HSJD-DIPG-007 and subclones NS-F10 and NS-F8 grown as neurospheres. Median Z scores for each biological replicate for each well of the screen are provided.

Supplementary Table S6 – *RNA sequencing of HSJD-DIPG-007 subclones.* Raw read counts and fragments per kilobase of exon per million reads mapped (FPKPM) are provided for each transcript in each neurosphere culture assayed. Differential expression analysis between *SUV420H1* mutant subclone NS-F10 and wild-type NS-F8 were assessed by gene set enrichment analysis (gene sets with nominal p values of 0 provided) and DAVID ontology analysis (enrichment scores >1.5 provided). All cell preparations were sequenced n=1.

Legends for supplementary videos

Supplementary Video S1 – *Invasion of co-cultured SU-DIPG-VI subclones.* Individual subclones E6 (green) and D10 (red) of SU-DIPG-VI were differentially labelled and co-cultured in equal ratios. The video shows time-lapse confocal microscopy of invasion onto matrigel over 30 hours, with the otherwise poorly motile E6 cells found to invade further and in greater numbers alongside D10 cells than when cultured alone. Taken from n=3 independent experiments.

Supplementary Video S2 – *Migration of co-cultured HSJD-DIPG-007 subclones.* Individual subclones NS-F8 (green) and NS-F10 (red) of HSJD-DIPG-007 were differentially labelled and co-cultured in equal ratios. The video shows time-lapse confocal microscopy of migration on fibronectin over 24 hours, with the otherwise poorly motile NS-F8 cells found to migrate further and in greater numbers alongside NS-F10 cells than when cultured alone. Taken from n=3 independent experiments.