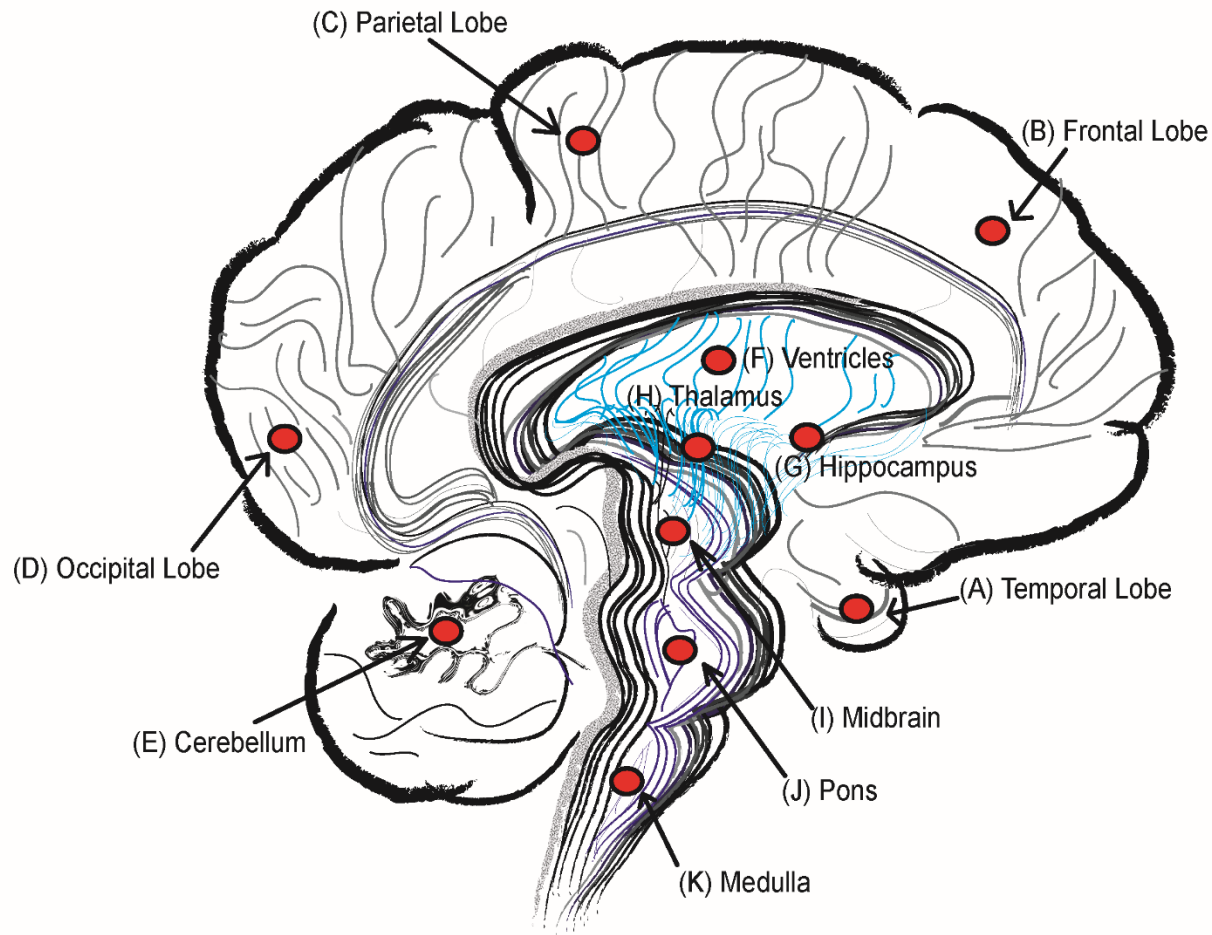


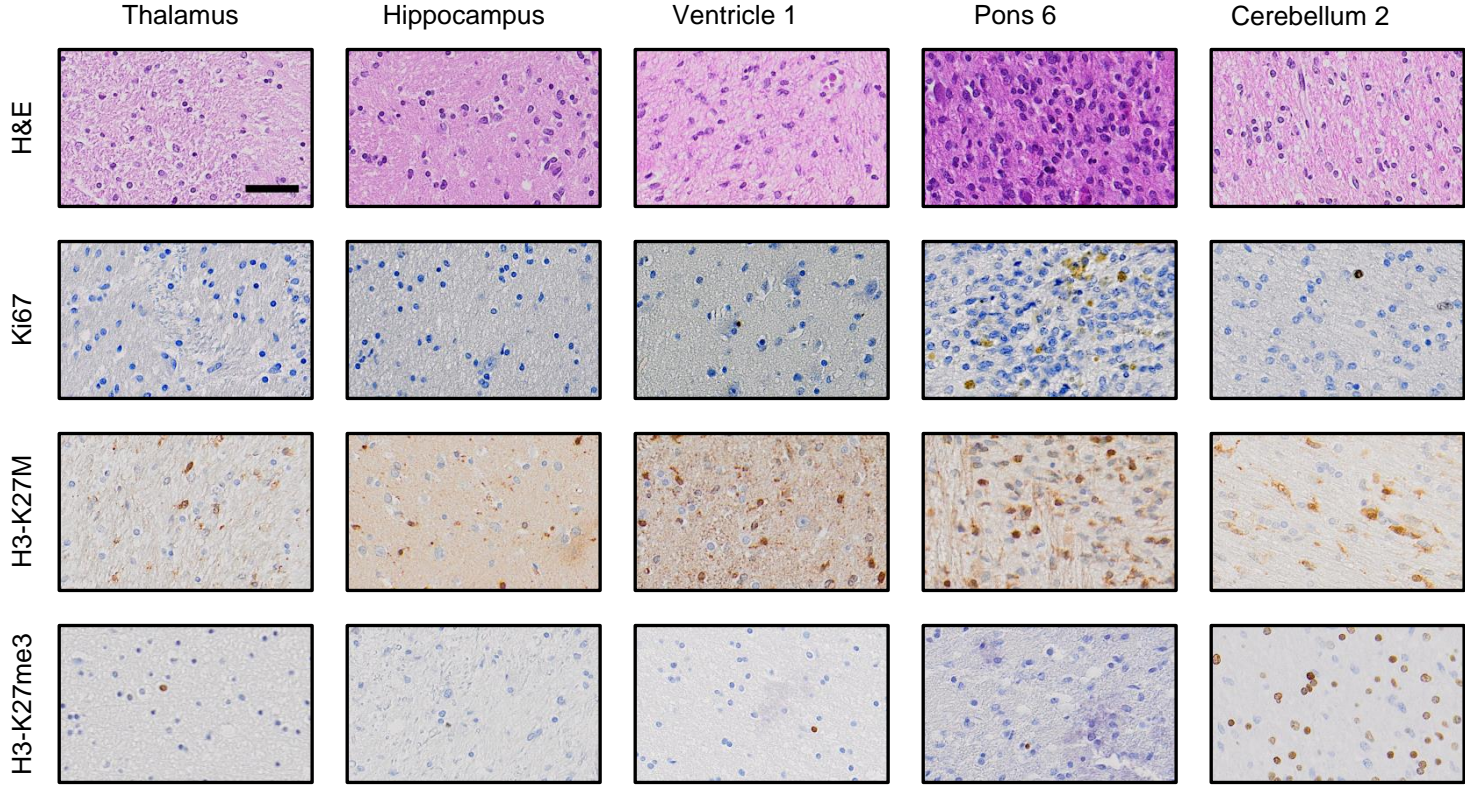
Supplementary Figure 1



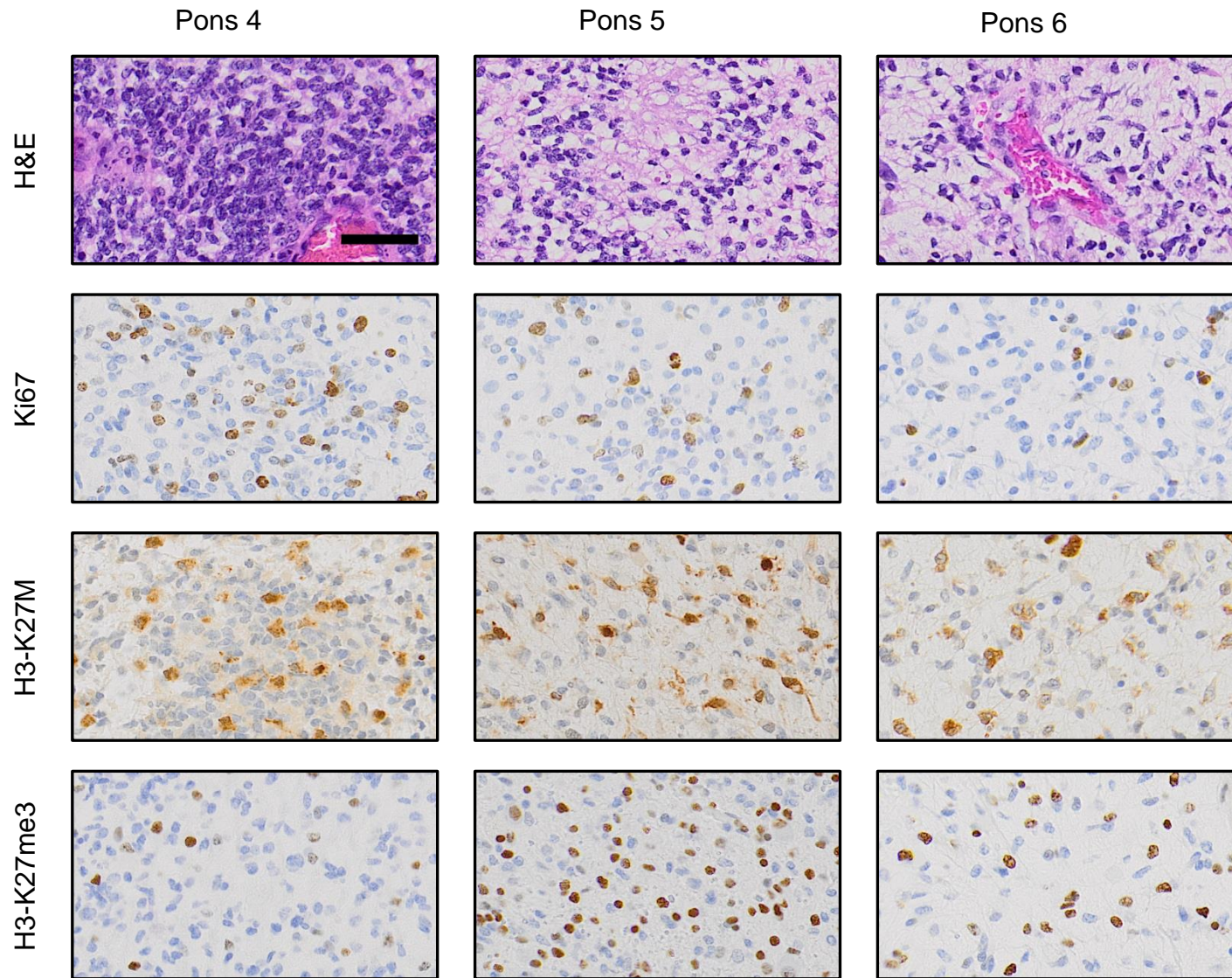
Supplementary Figure 1. Core punch sampling of whole autopsy brains. An average of 6 core punch samples from right and left brainstem (midbrain, pons, medulla) and an average of 10 core punch samples from right and left cerebellum and cerebrum (frontal lobe, parietal lobe, temporal lobe, occipital lobe, lateral ventricles, thalamus, hippocampus) were collected for molecular analyses from frozen and formalin fixed paraffin embedded (FFPE) whole autopsy brains from nine DIPG patients (N=134).

Supplementary Figure 2

(a) DIPG1 (21y 2m; Male; H3.1 K27M)

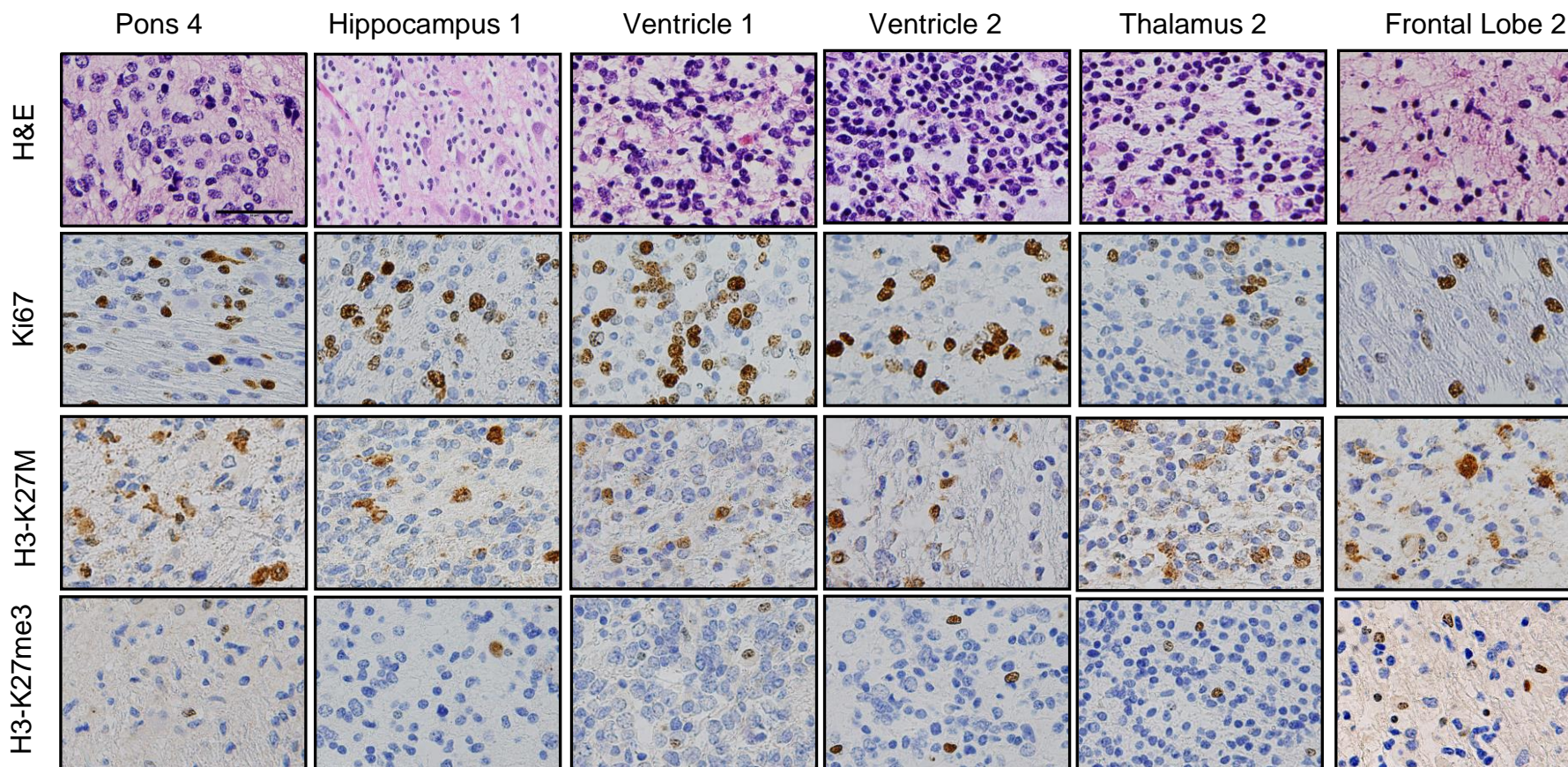


(b) DIPG2 (6y 10m; Male; H3.1 K27M)

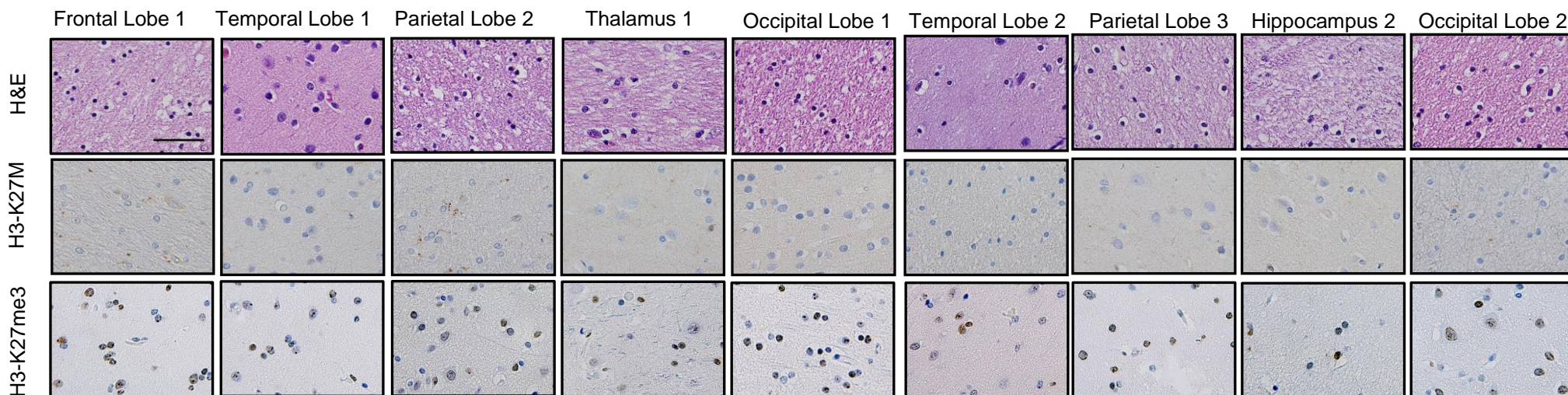


(c) DIPG7; 10y 8m; Female; H3.3 K27M

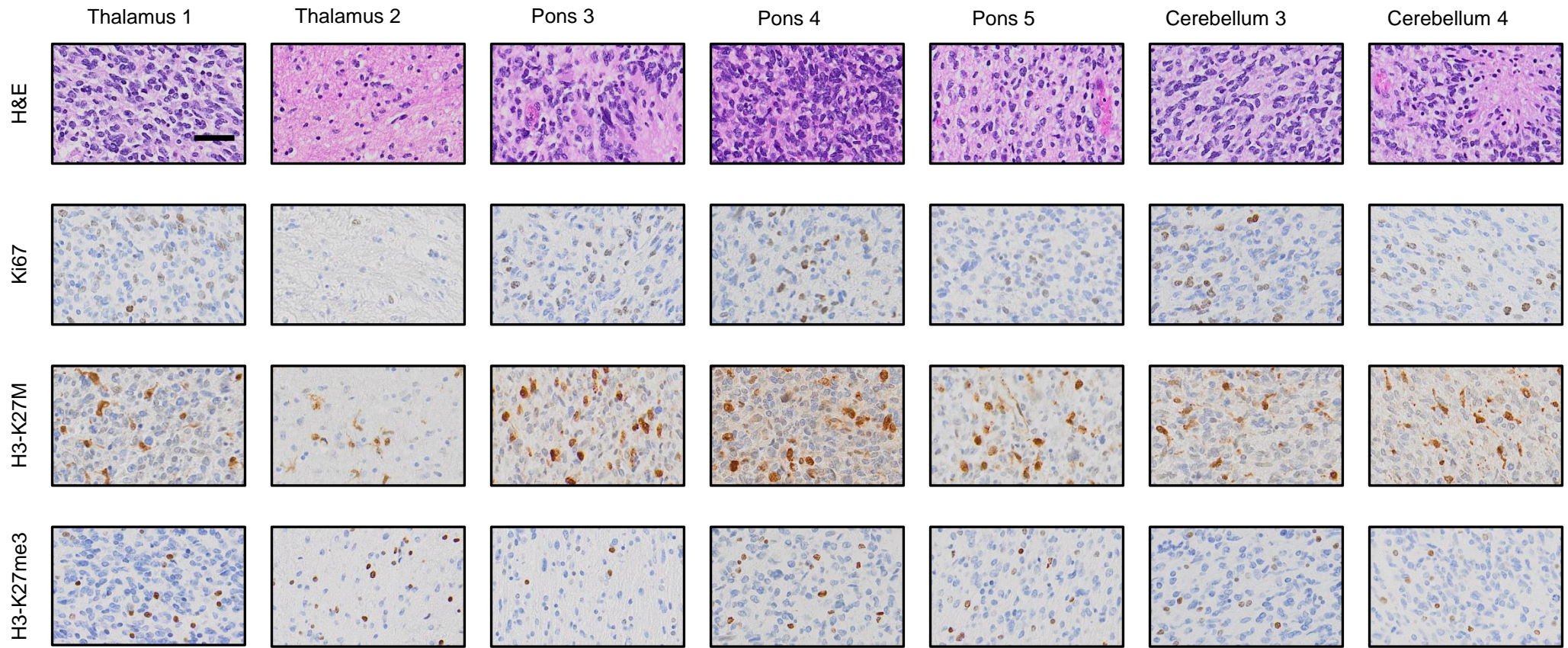
Tumor Locations



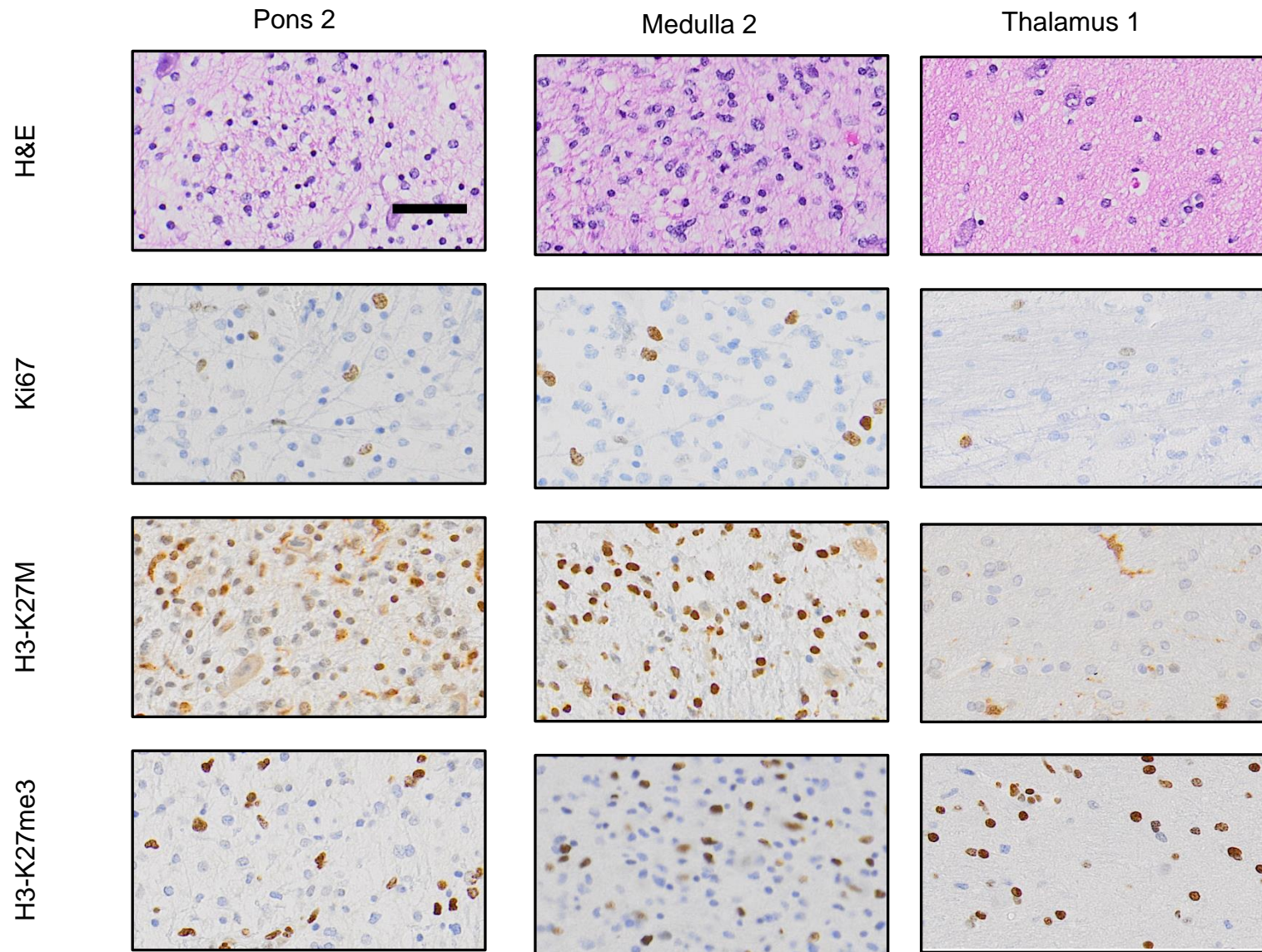
Normal Locations



(d) DIPG8 (5y; Male; H3.3 K27M)

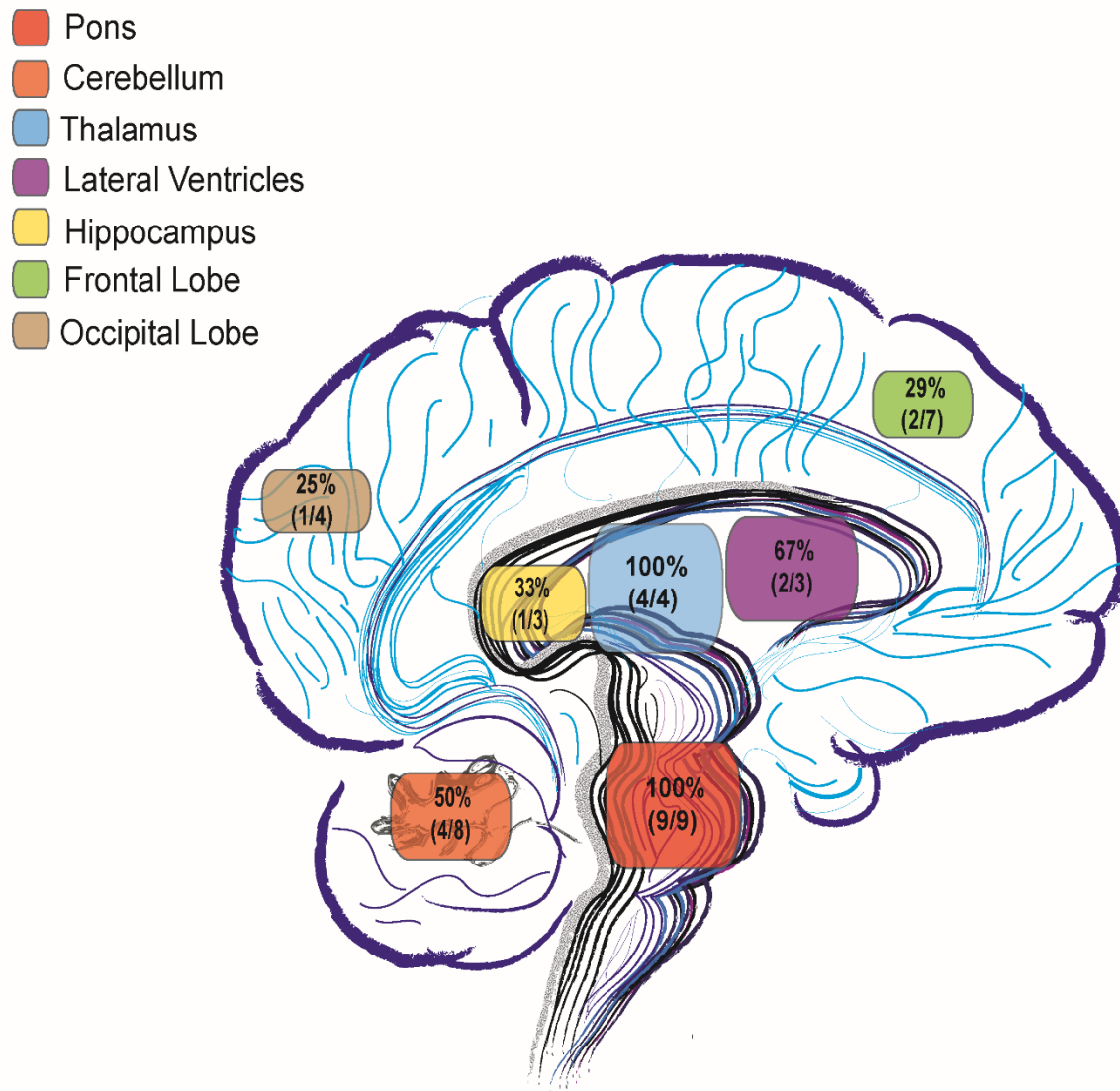


(e) DIPG9 (9y 3m; Male; H3.3 K27M)



Supplementary Figure 2. Immunohistochemical staining of several neuroanatomical locations in DIPG autopsies. Histological sections were probed for Ki67 (proliferation marker), histone 3 K27M mutant (H3-K27M), and histone 3 K27 trimethylation (H3-K27me3) stains in (a) DIPG1 (b) DIPG2 (c) DIPG7 tumor (top) and normal (bottom) (d) DIPG8 and (e) DIPG9. Patient age, gender, and histone 3 mutation status are indicated above each histological staining panel. Tumor spread is seen in proximal (cerebellum and medulla) and distant (thalamus and frontal lobe) brain locations from the primary tumor in the pons. In DIPG9 (e) the area showing normal histology had tumor infiltrate as seen using the specific anti-H3K27M staining or in our molecular studies (Supplementary Table 1). Staining performed on several areas from normal brain DIPG7 (c, bottom) shows no unspecific anti-H3K27M staining in concordance with molecular results (Supplementary Table 1). *Scale bar: 50 μ m.*

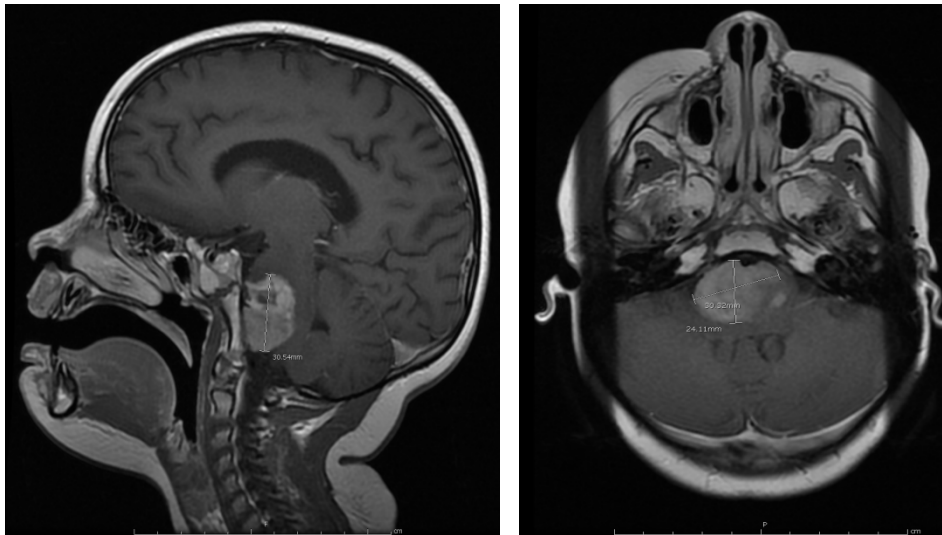
Supplementary Figure 3



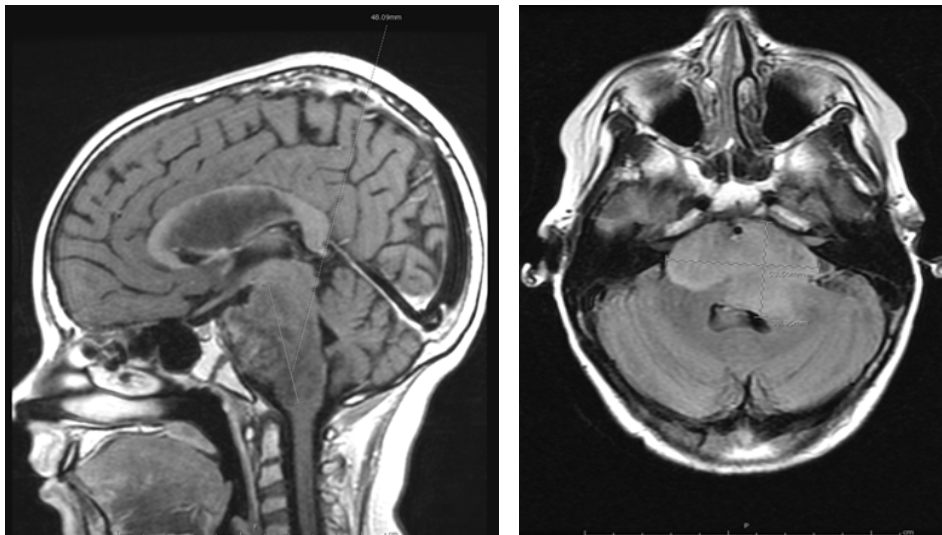
Supplementary Figure 3. Tumor extension in DIPGs. Neuroanatomical location and percent frequency of tumor extension from pons to the cerebellum, thalamus, lateral ventricles, hippocampus, frontal and occipital lobes (color key on left) in brains from nine DIPG patients. Tumor extension was detected by the presence of H3-K27M driver mutation assessed by whole exome sequencing, MiSeq targeted sequencing, and digital droplet PCR molecular analyses. Numbers in parentheses represent the percentage of tumor extension (number of patients with tumor extension/number of patients analyzed).

Supplementary Figure 4

(A)



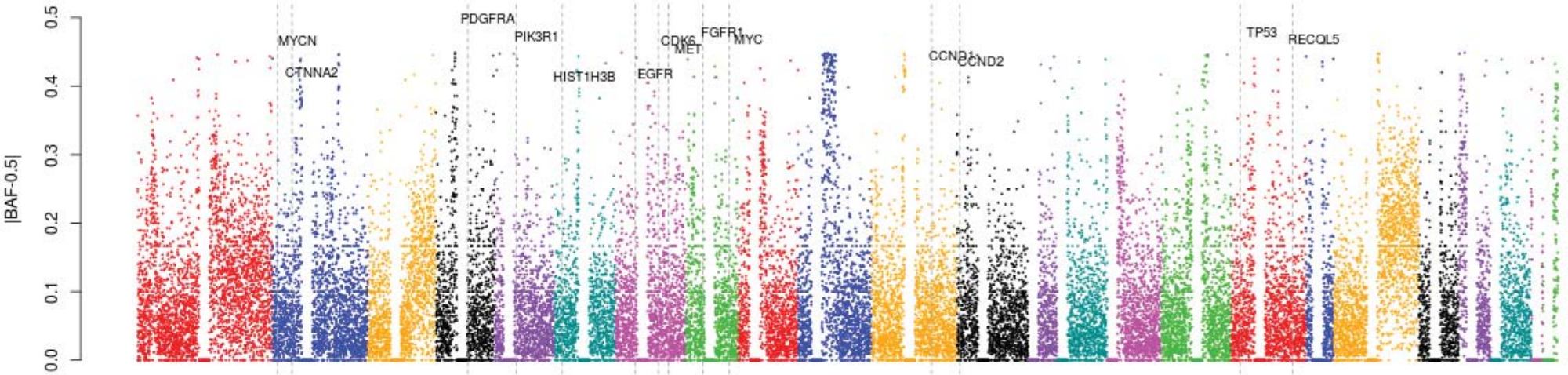
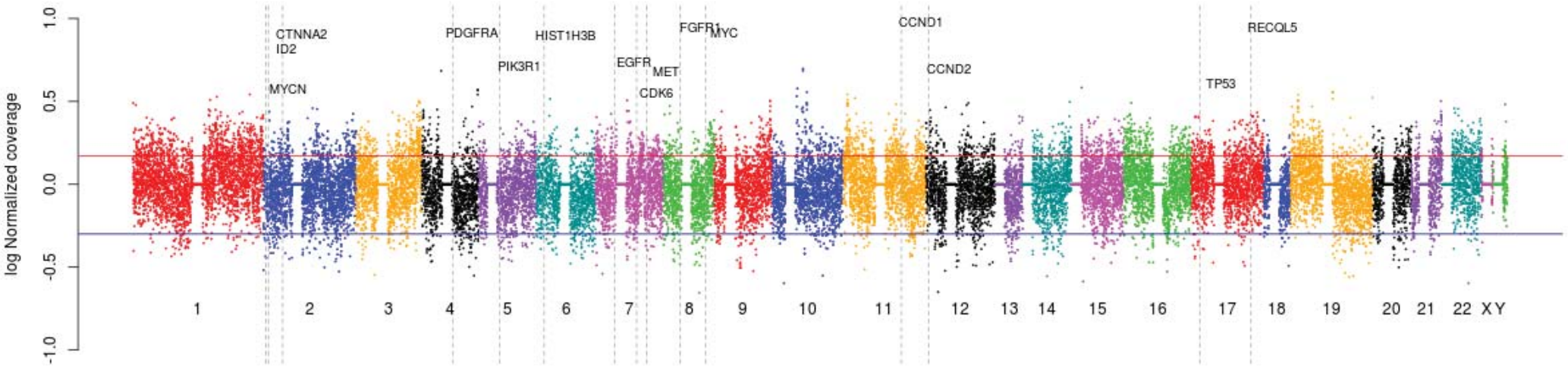
(B)



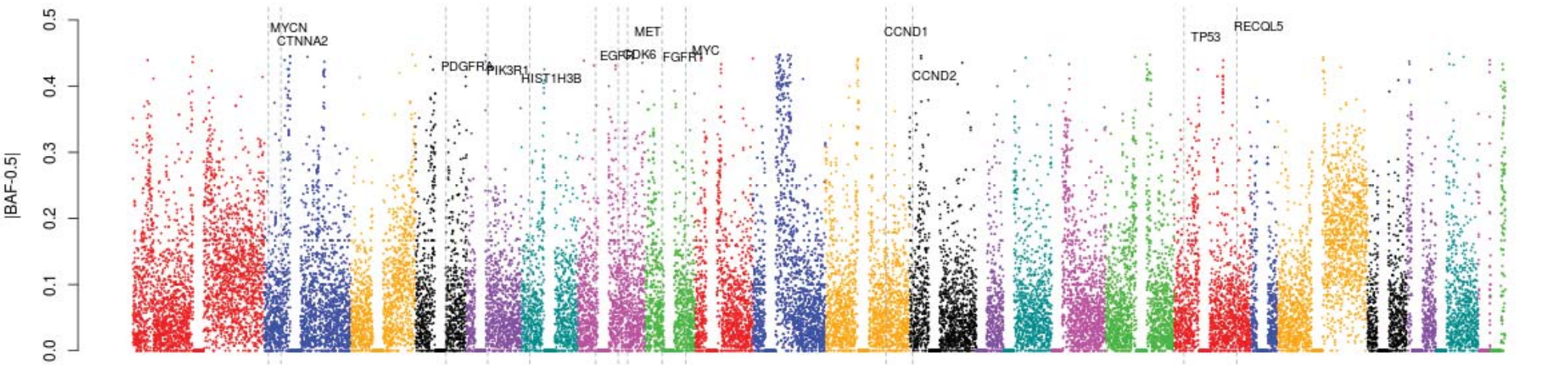
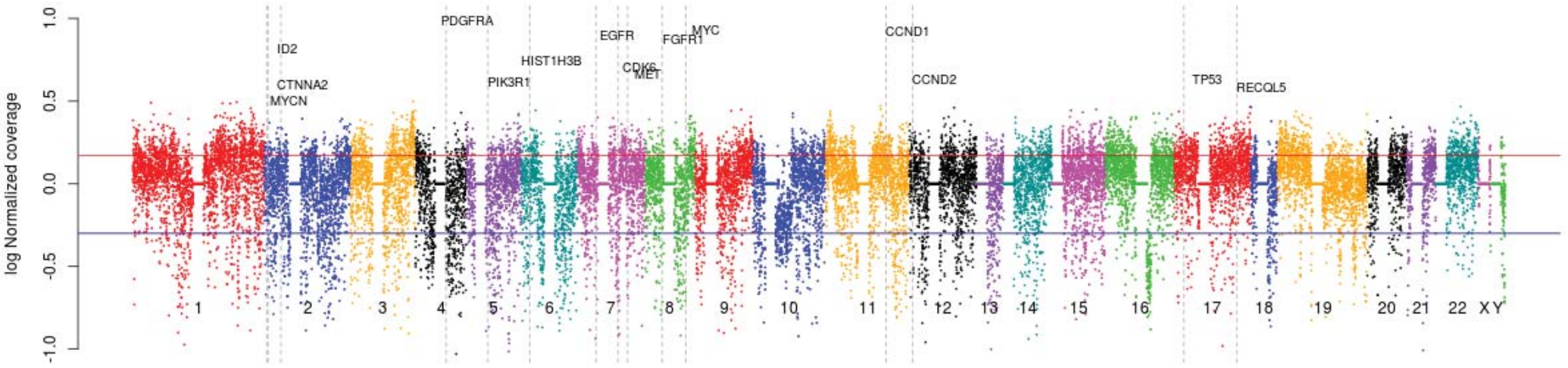
Supplementary Figure 4. Sagittal and axial view MRI of DIPG4 and DIPG7. (A) DIPG patient with non-enhancing lesion in the pons and absence of tumor extension to the cerebral cortex (DIPG4). (B) DIPG patient with non-enhancing pontine lesion and tumor extension to the lateral ventricles (DIPG7). These MRIs were taken from patients 2 weeks (DIPG7) and 3 months (DIPG4) before death, respectively.

Supplementary Figure 5

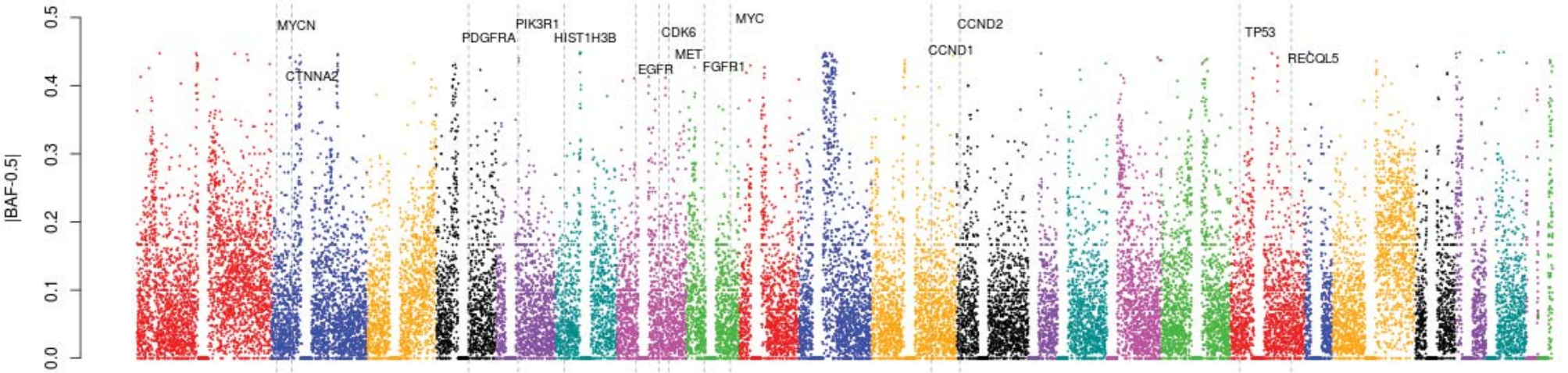
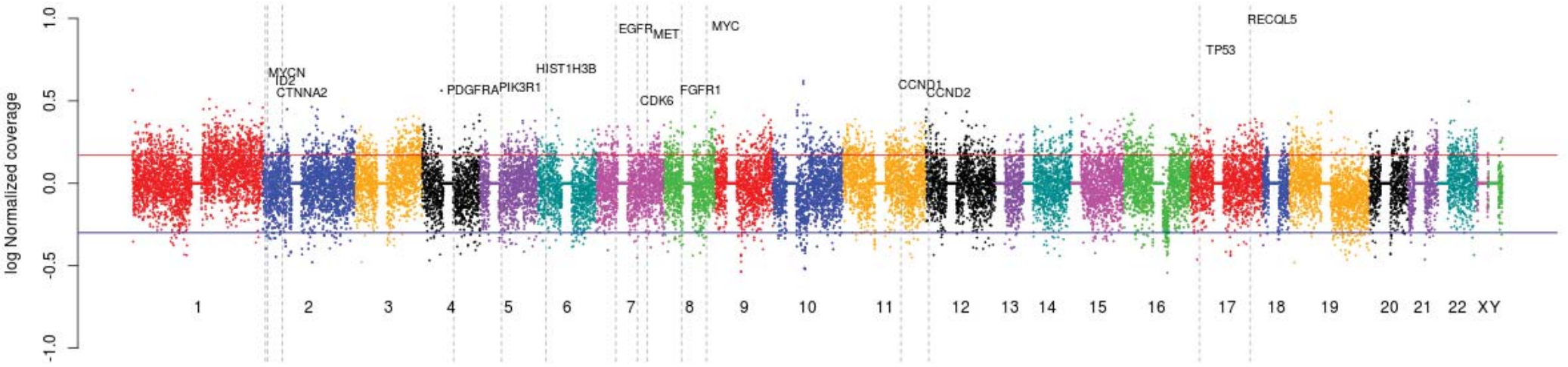
DIPG1-Medulla



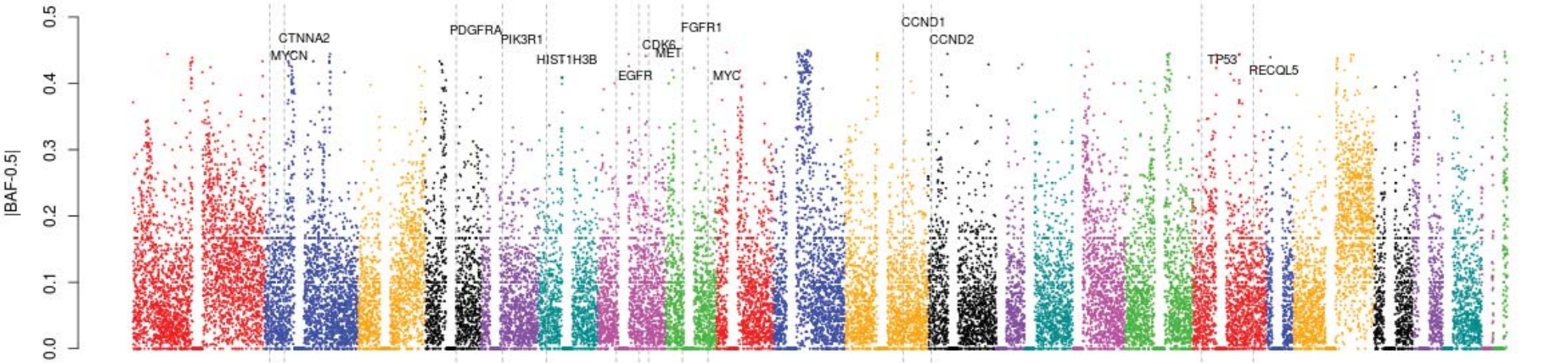
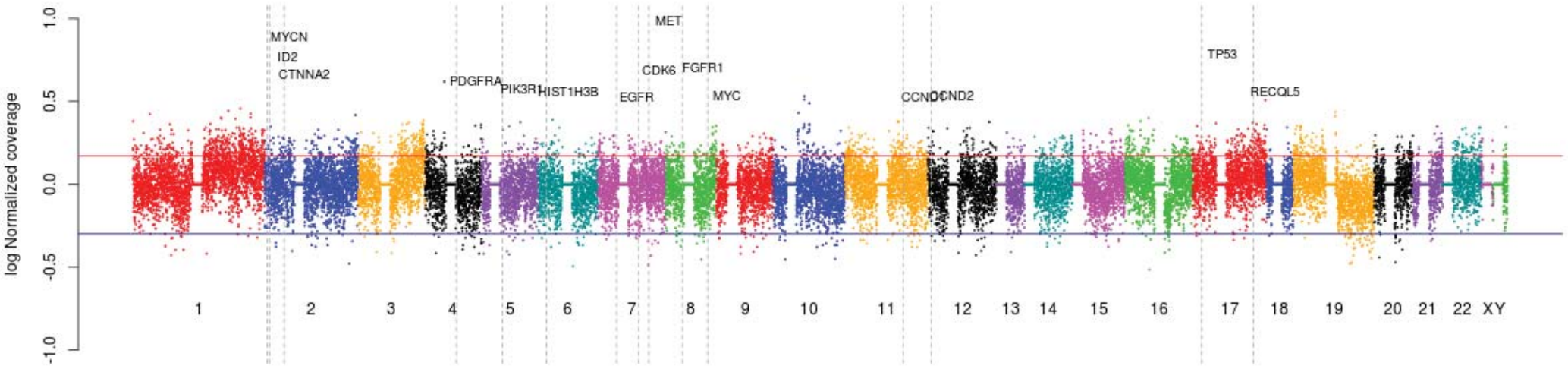
DIPG1-Midbrain



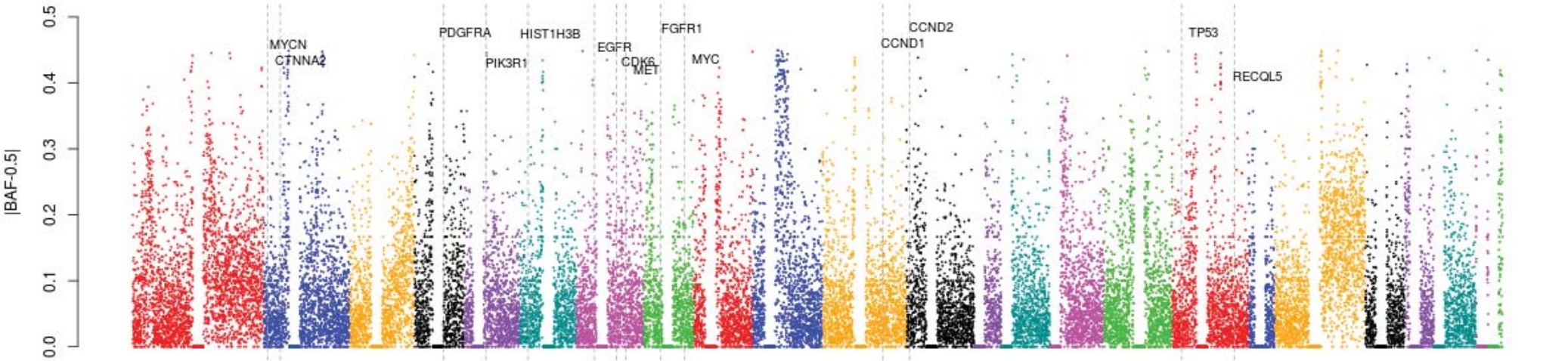
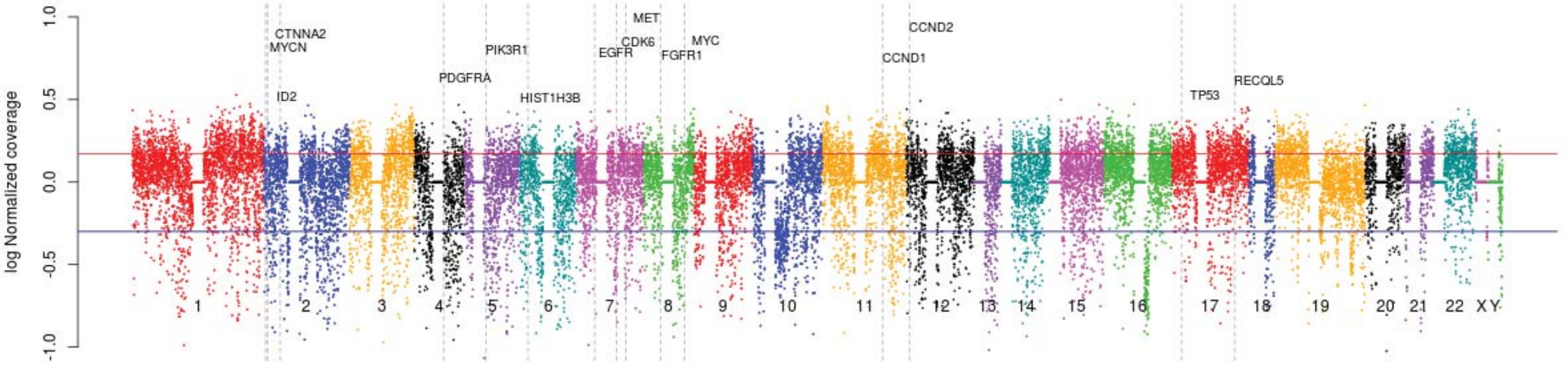
DIPG1-Pons 1



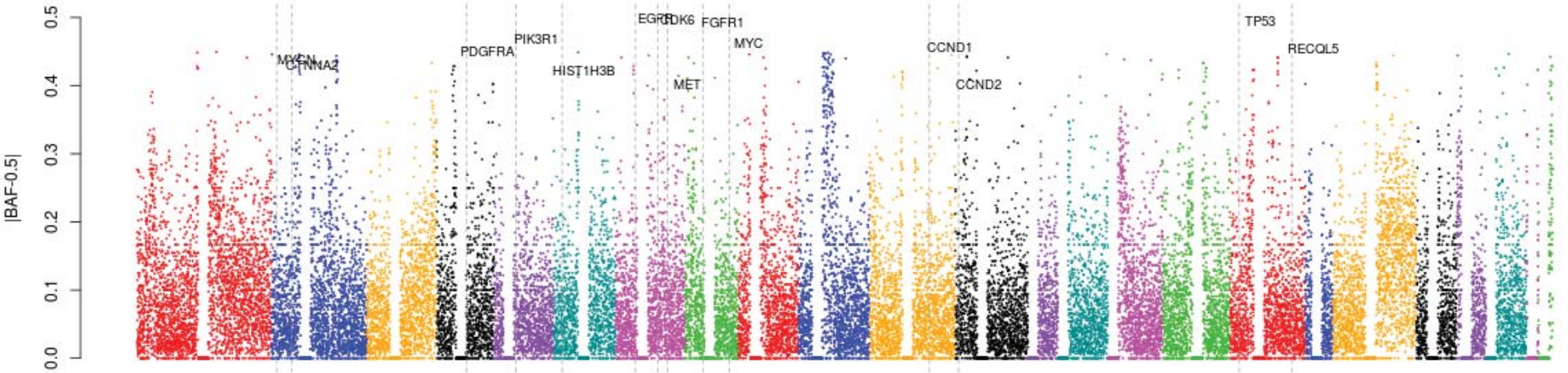
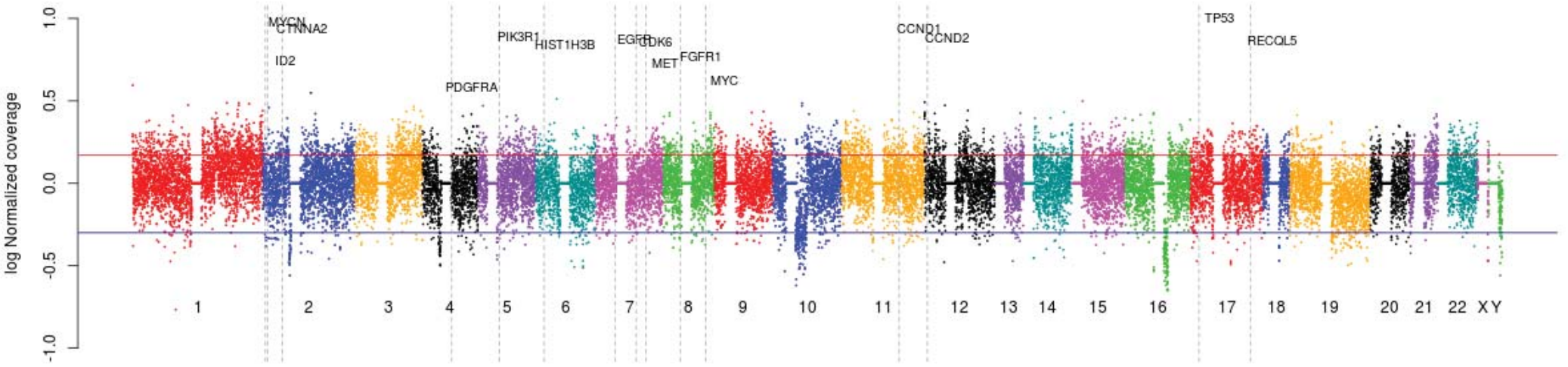
DIPG1-Pons 2



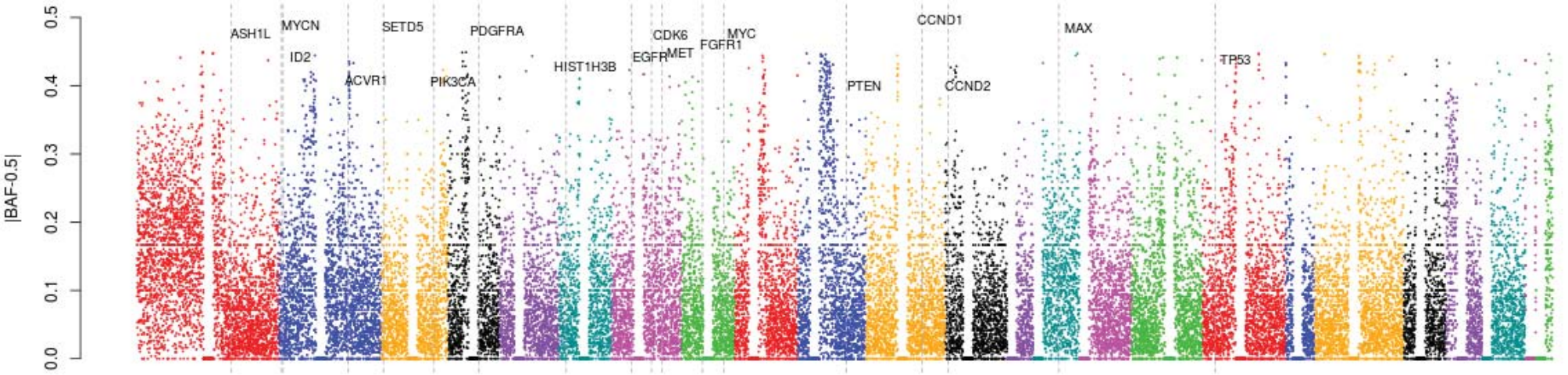
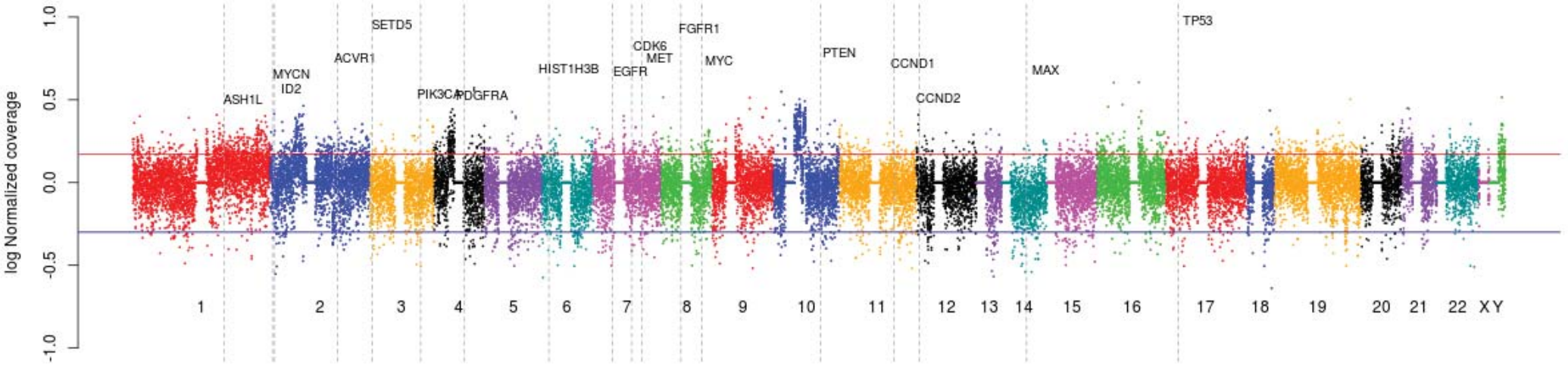
DIPG1-Pons 3



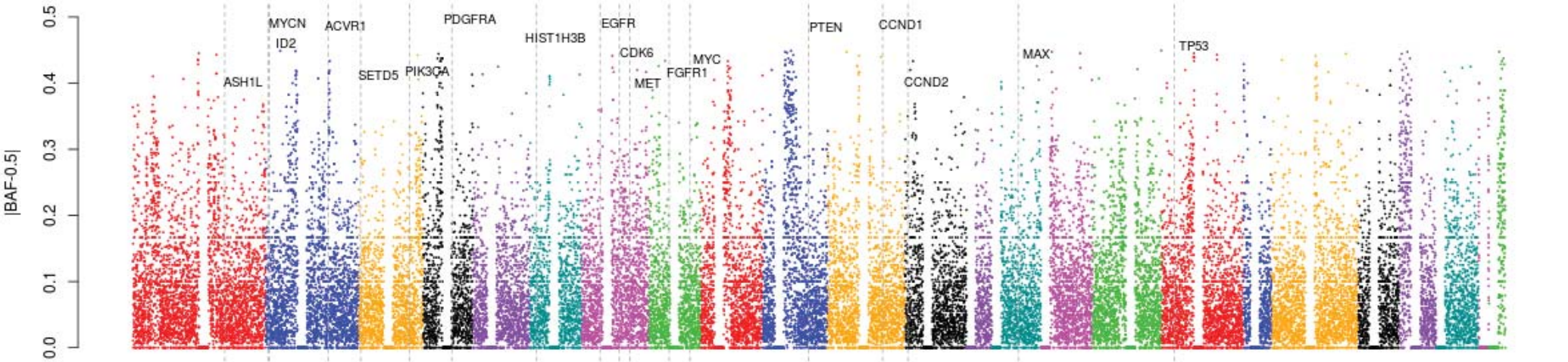
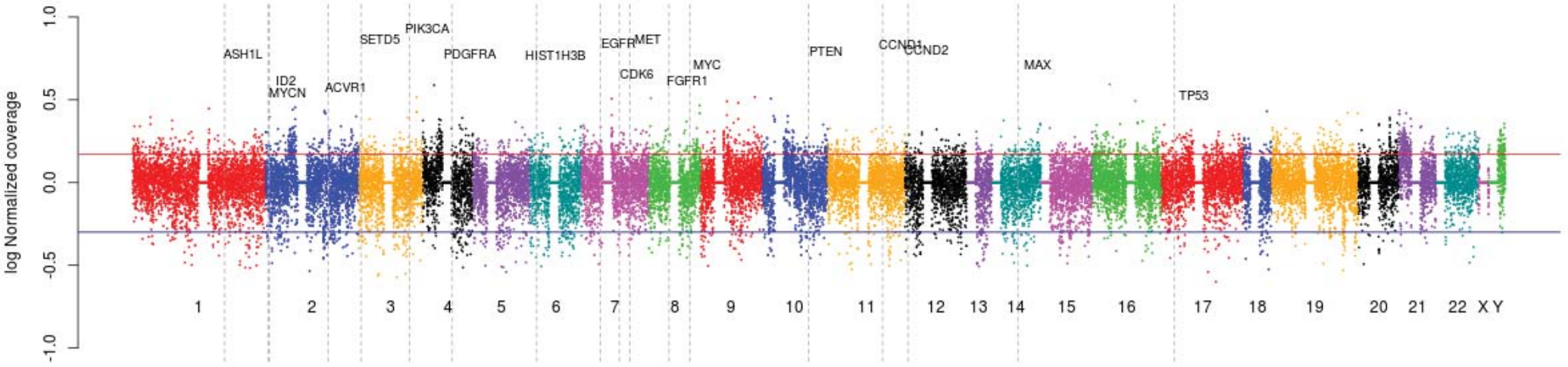
DIPG1-Pons 4



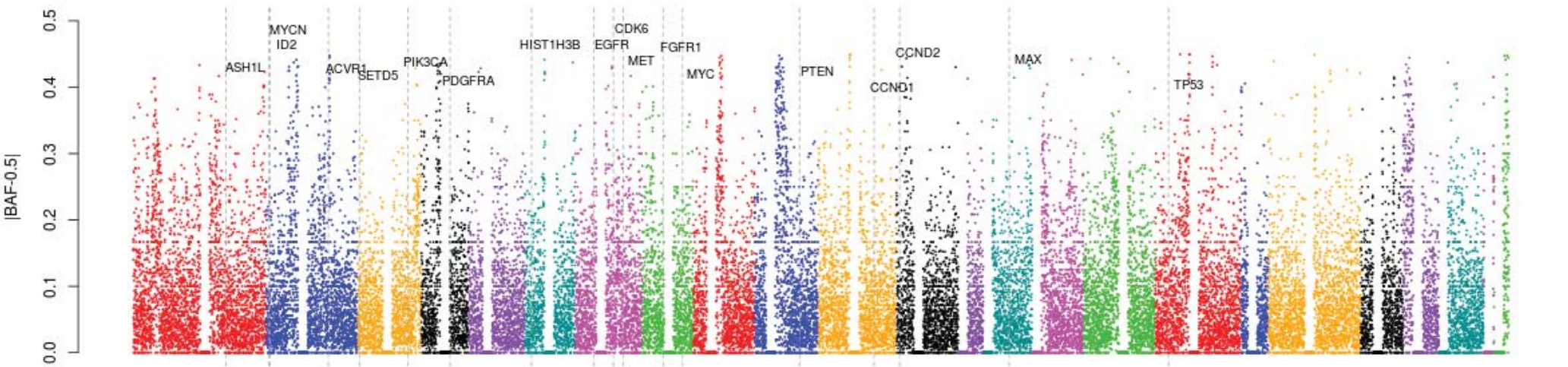
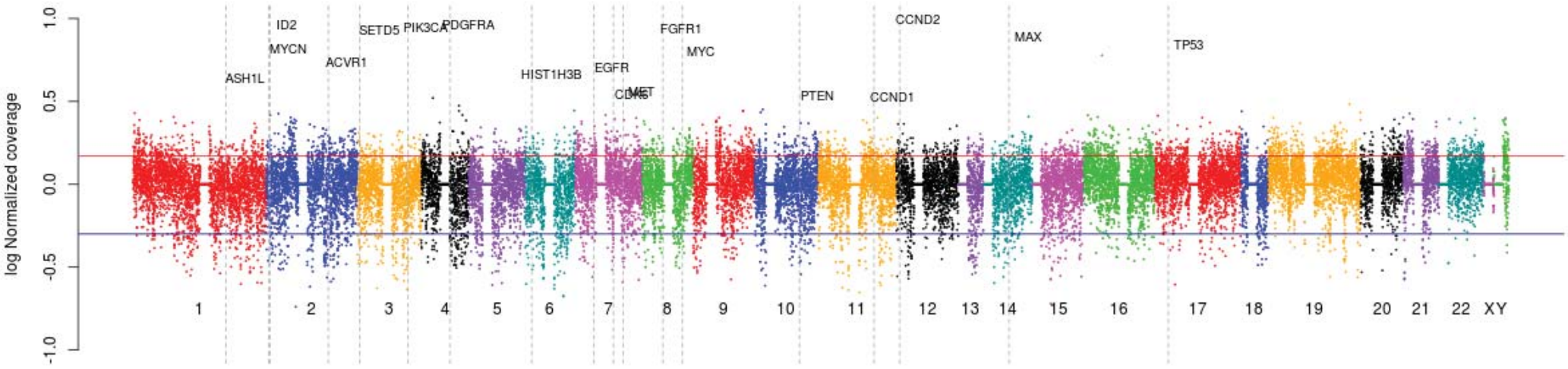
DIPG2-Cerebellum 1



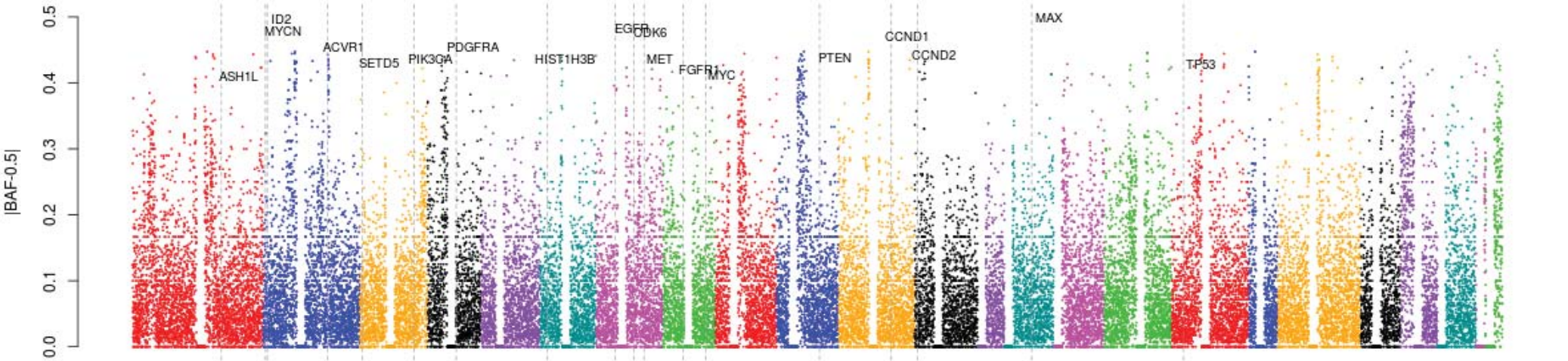
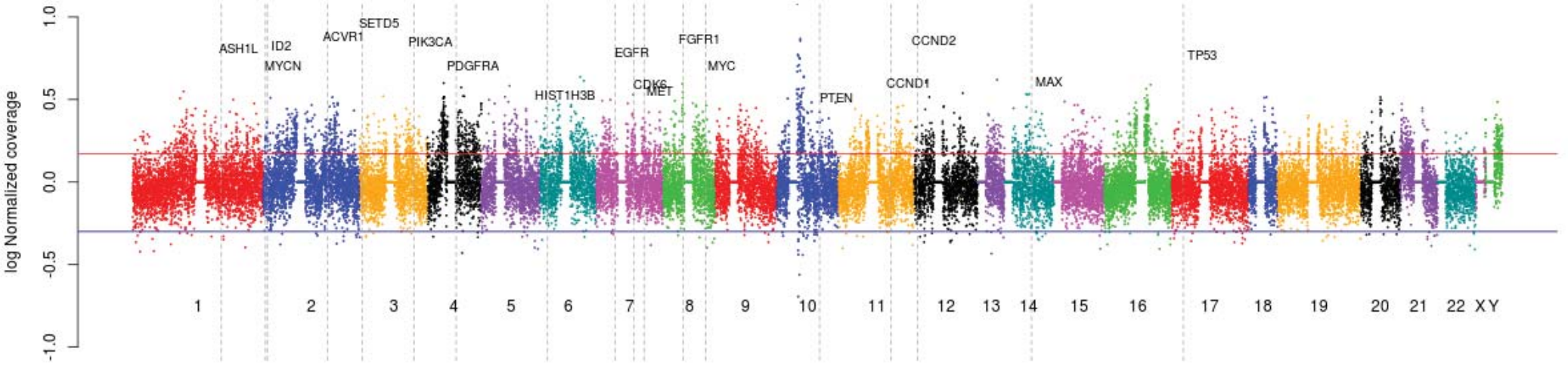
DIPG2-Cerebellum 2



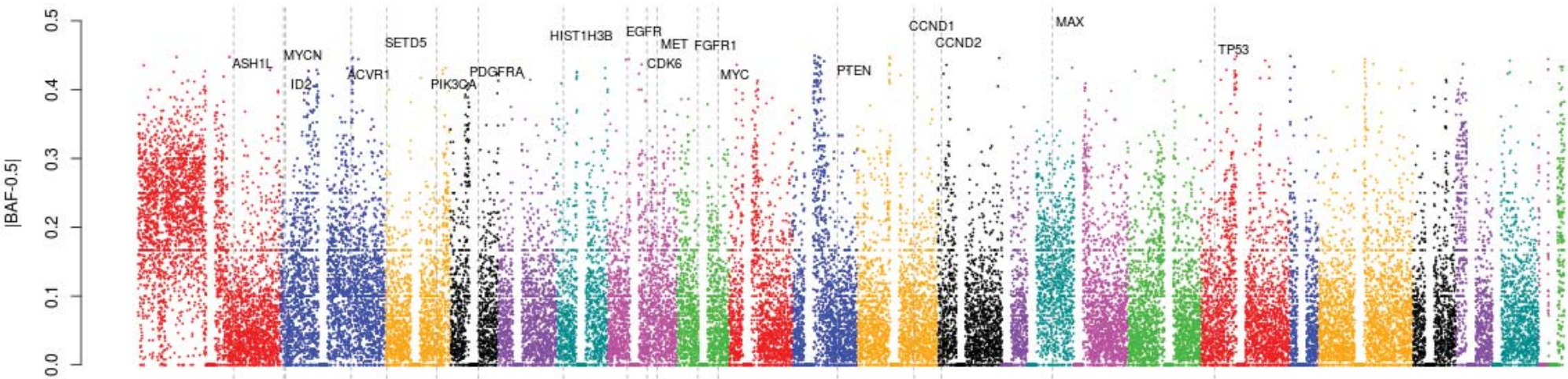
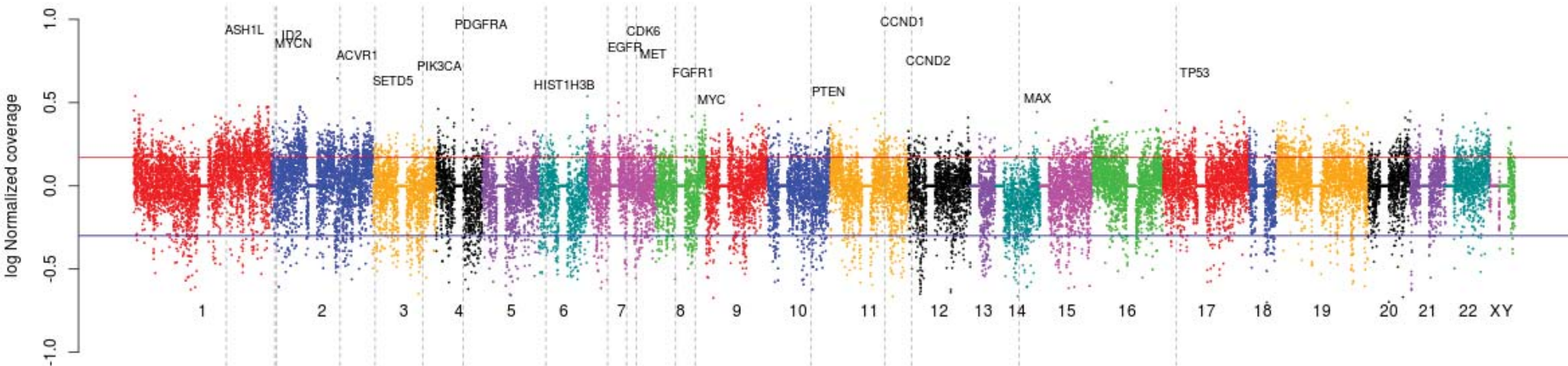
DIPG2-Cerebellum 3



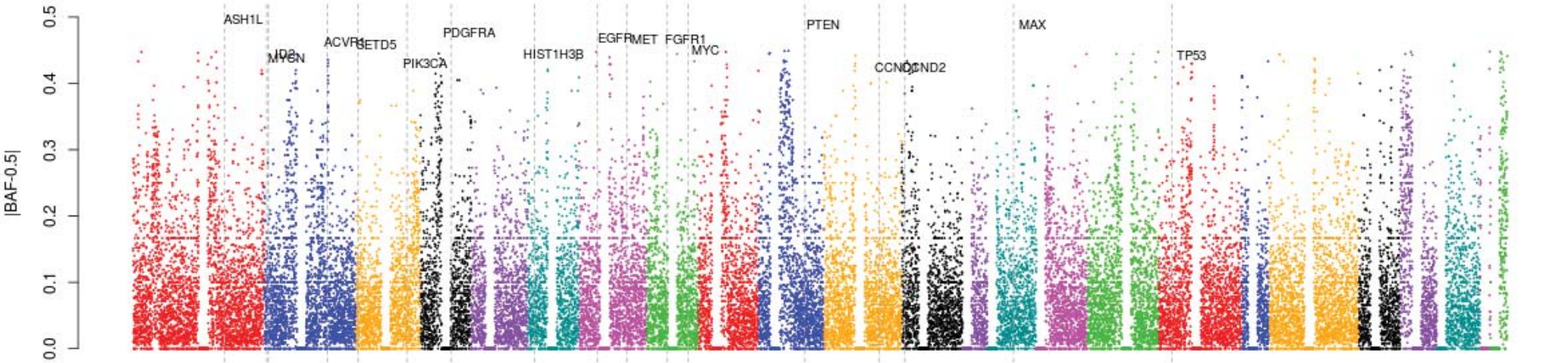
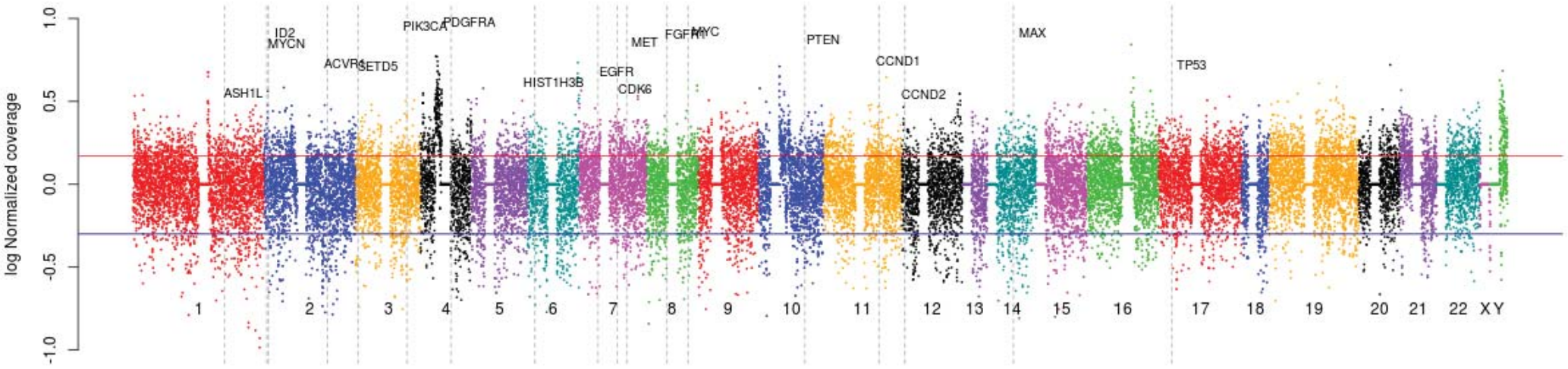
DIPG2-Frontal Lobe 3



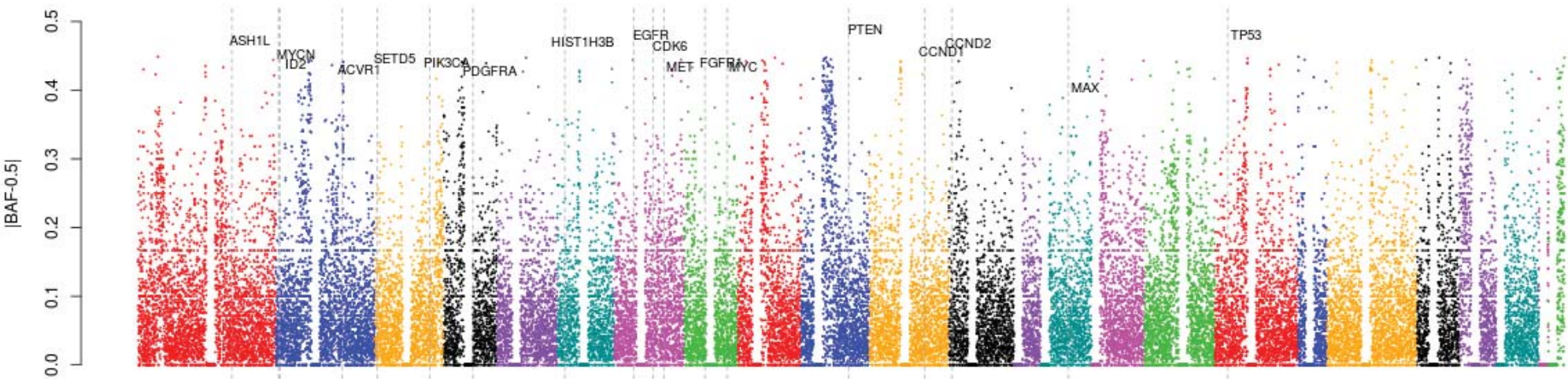
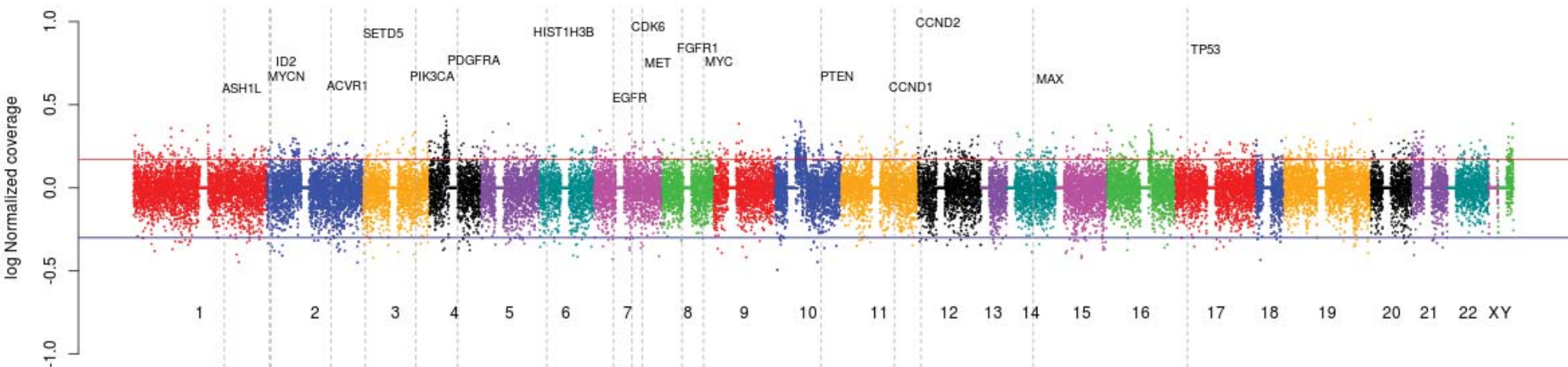
DIPG2-Medulla



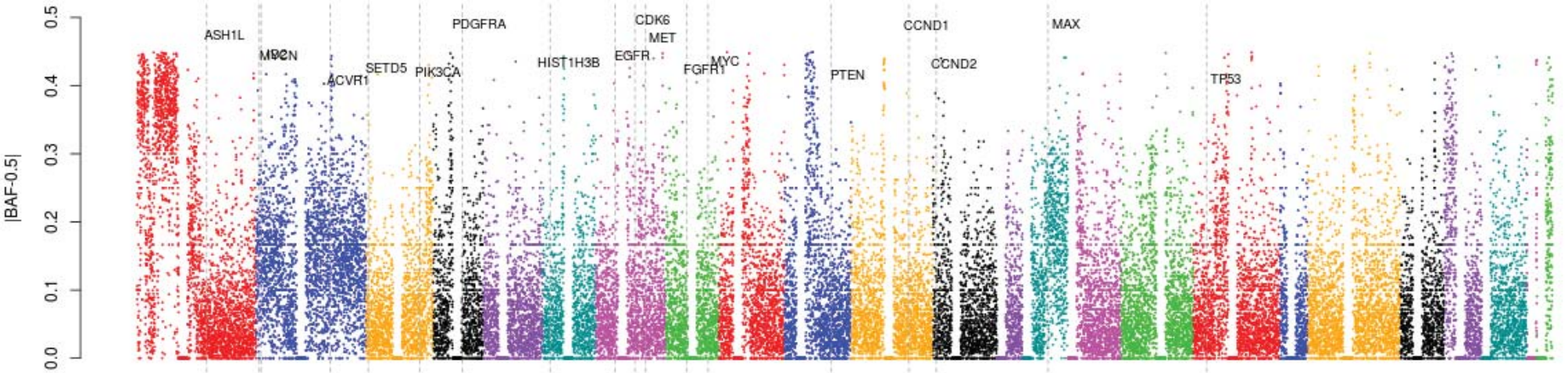
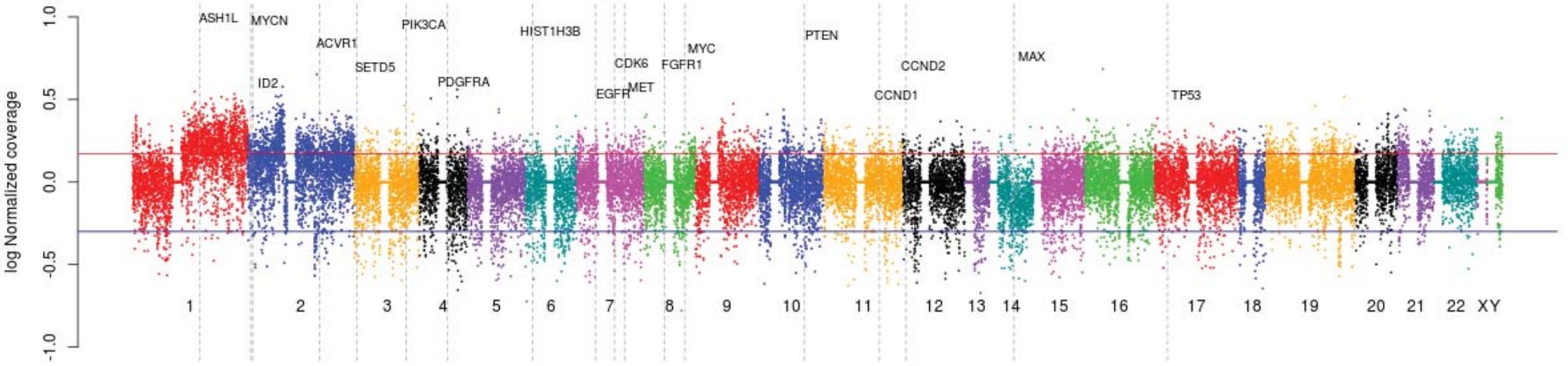
DIPG2-Occipital Lobe 2



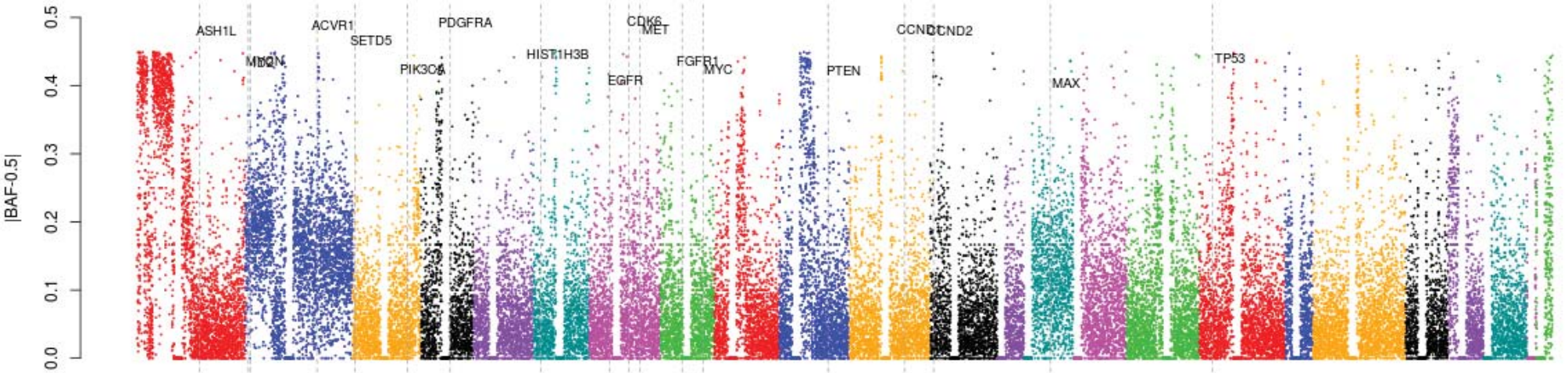
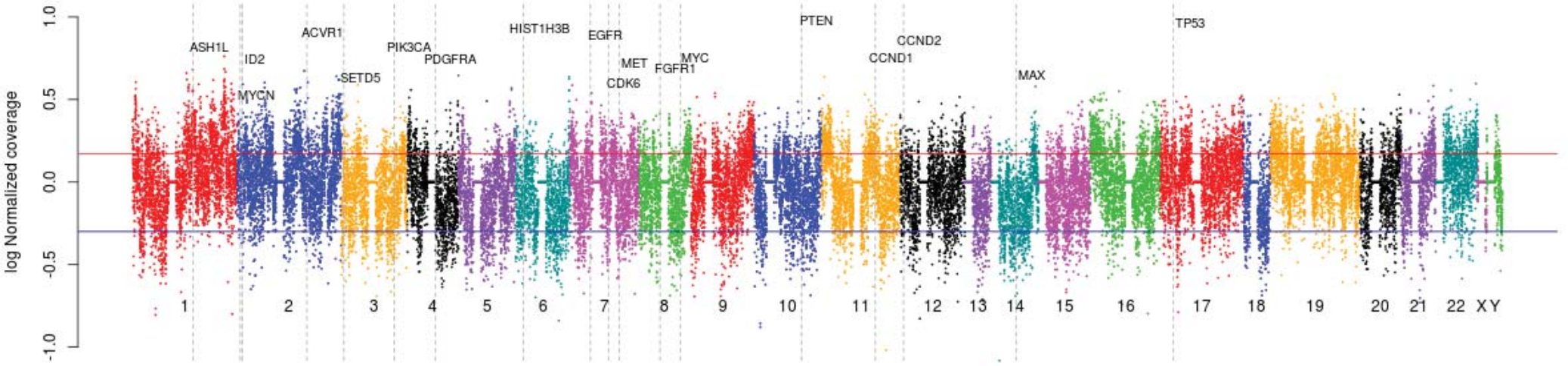
DIPG2-Parietal Lobe 4_lobe



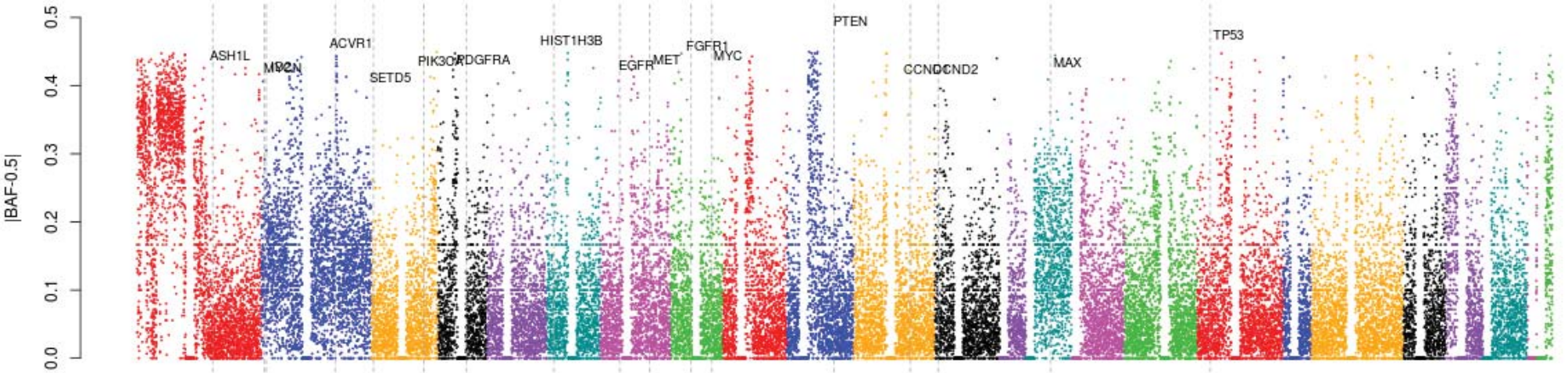
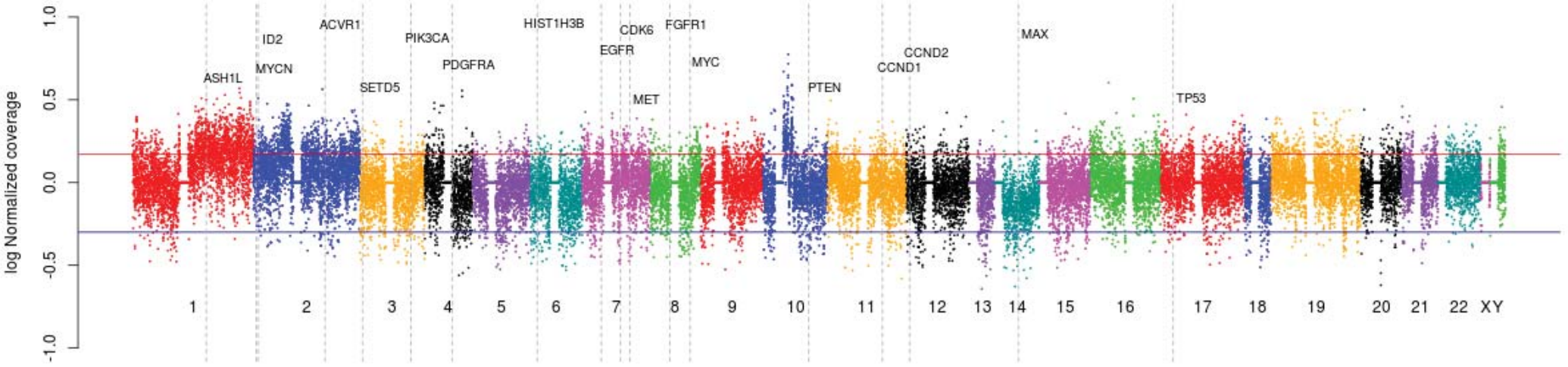
DIPG2-Pons 1



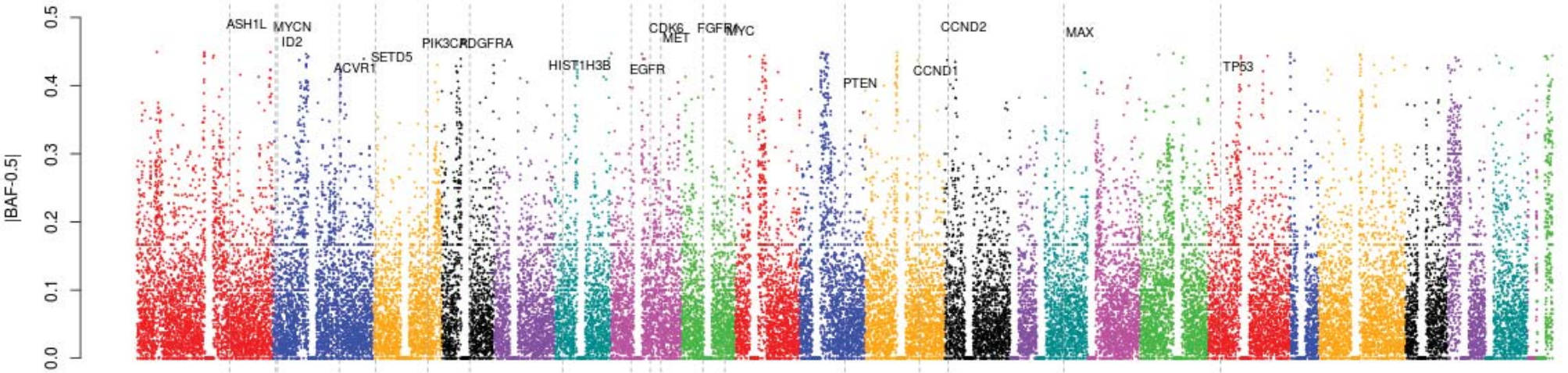
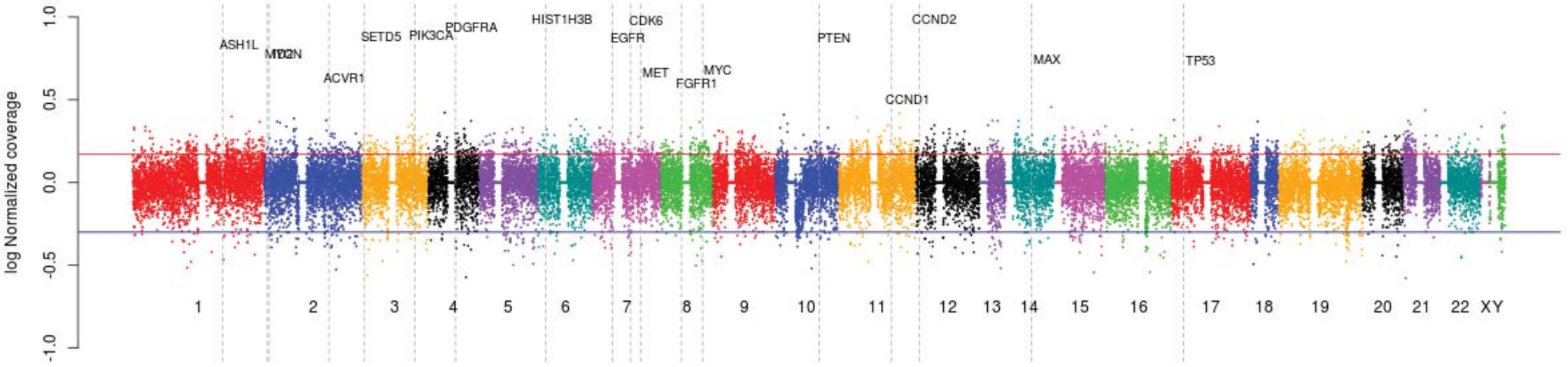
DIPG2-Pons 2



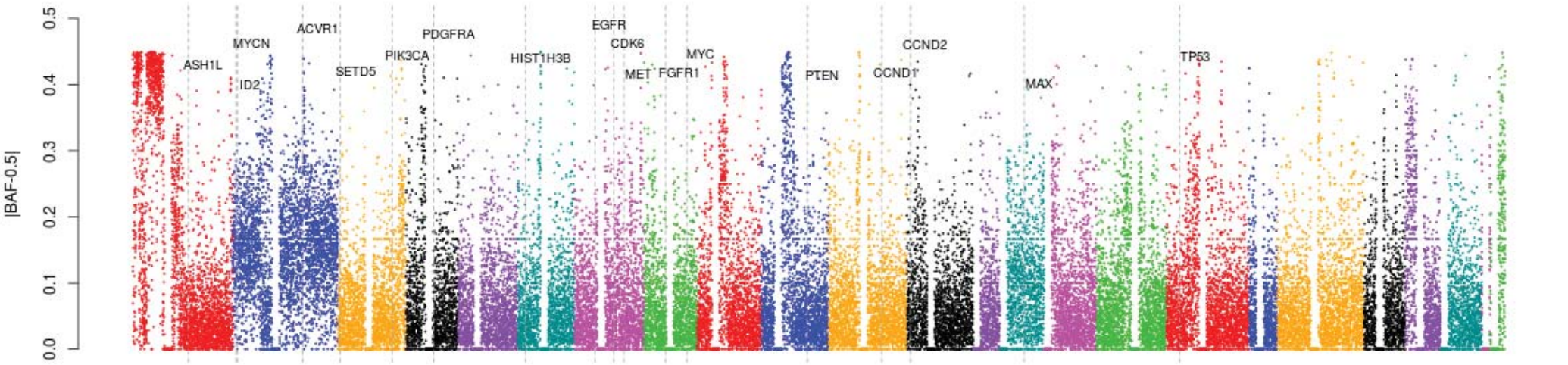
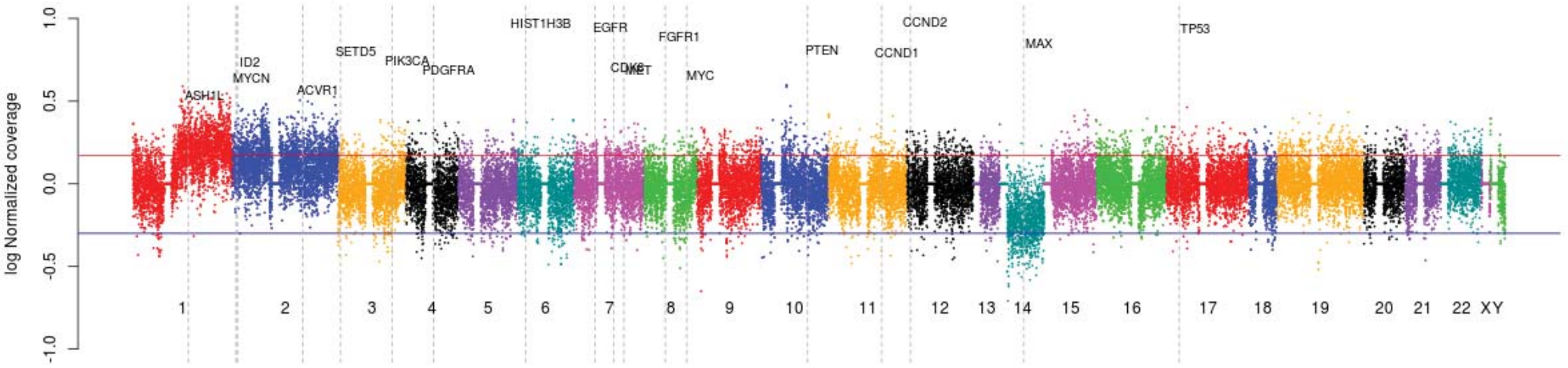
DIPG2-Pons 3



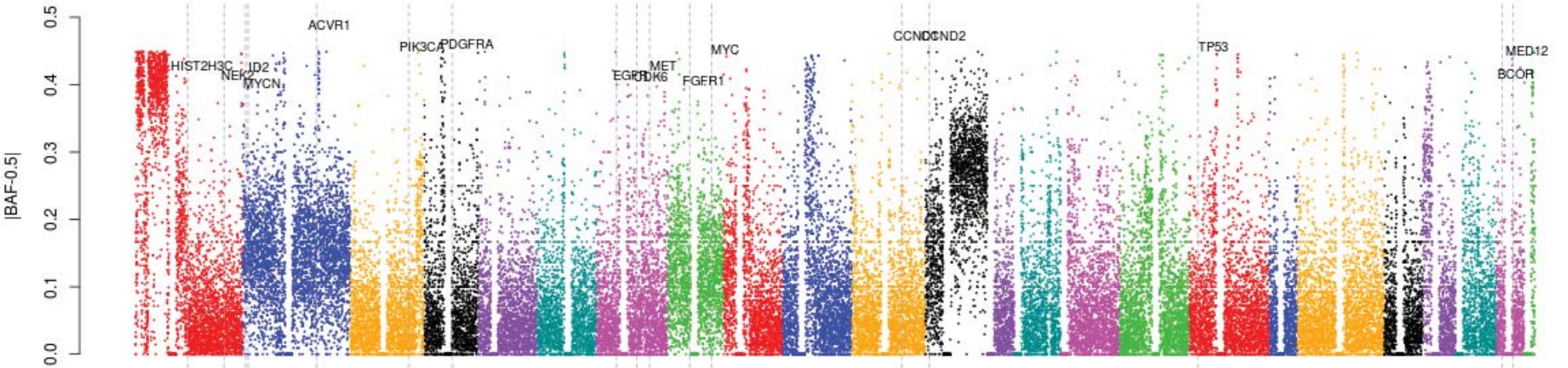
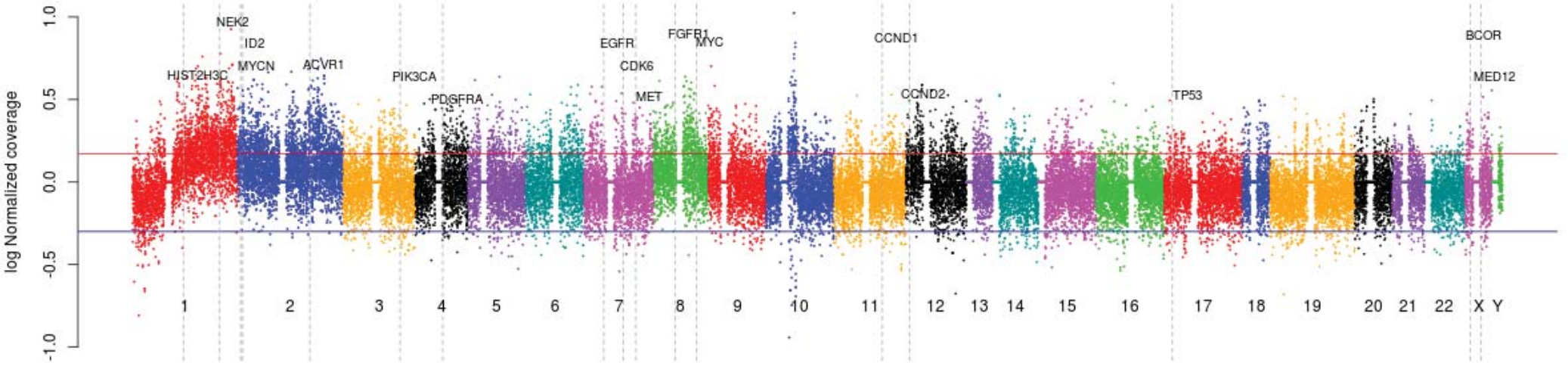
DIPG2-Temporal Lobe 3



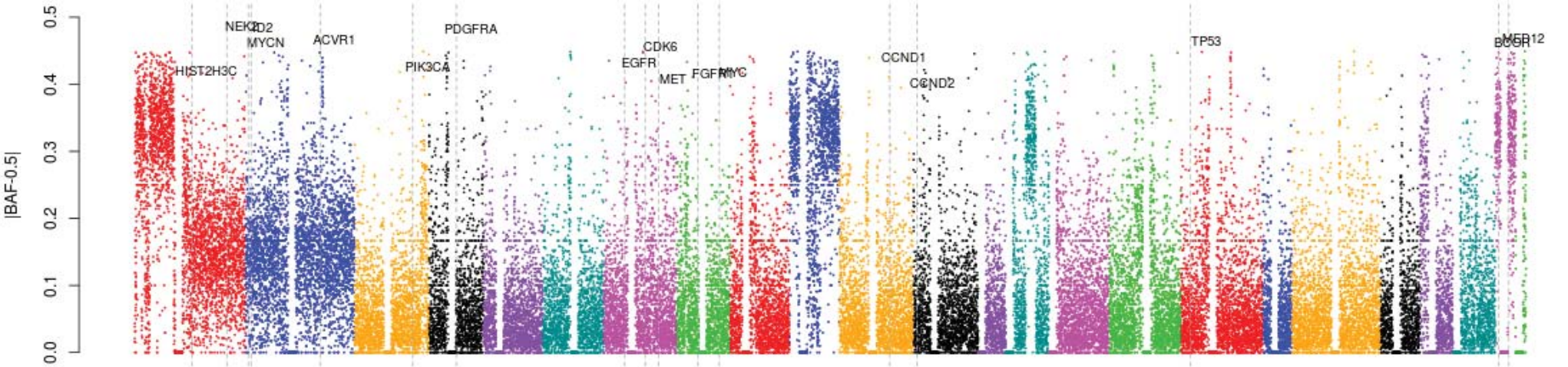
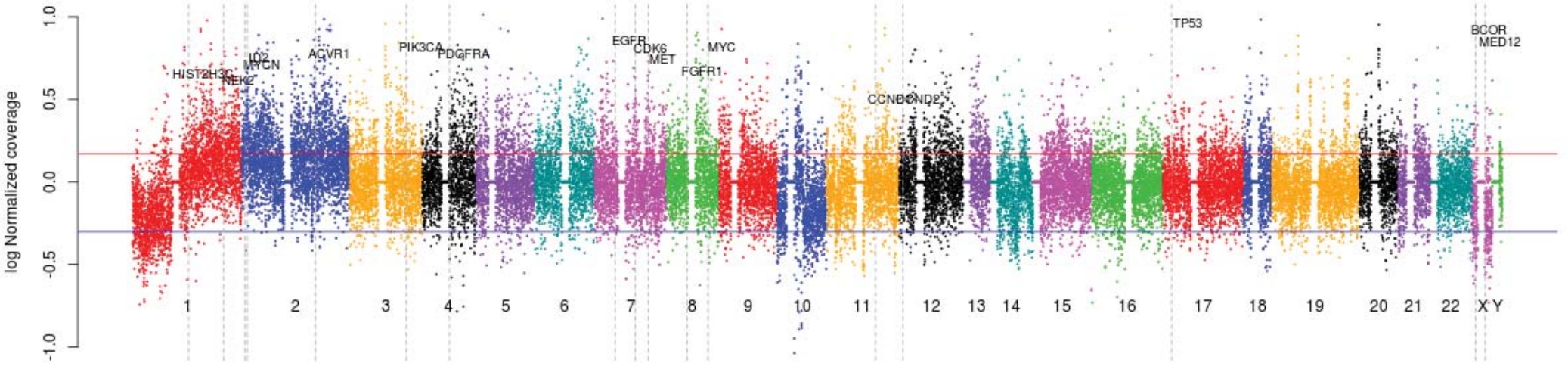
DIPG2-Thalamus



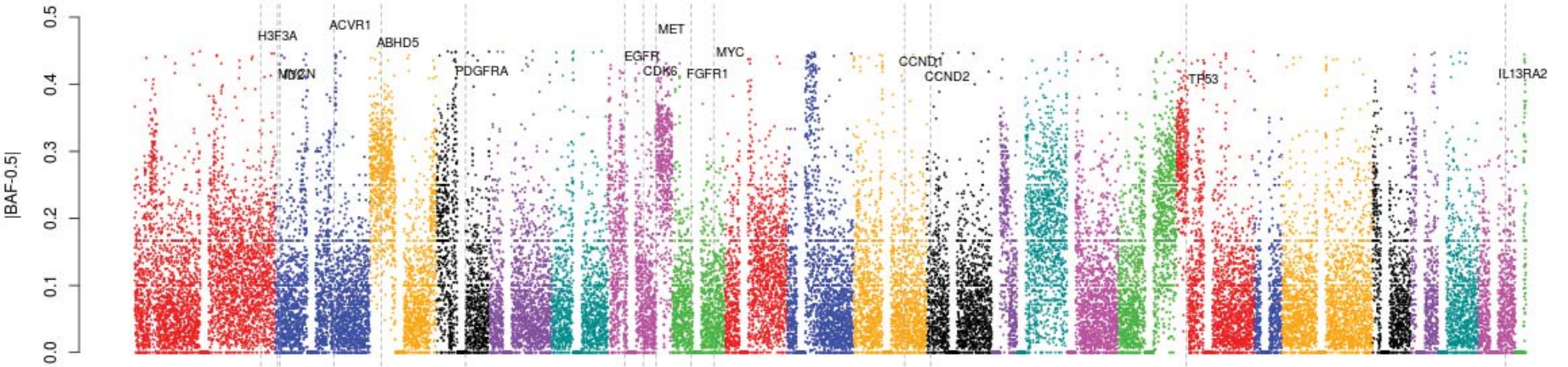
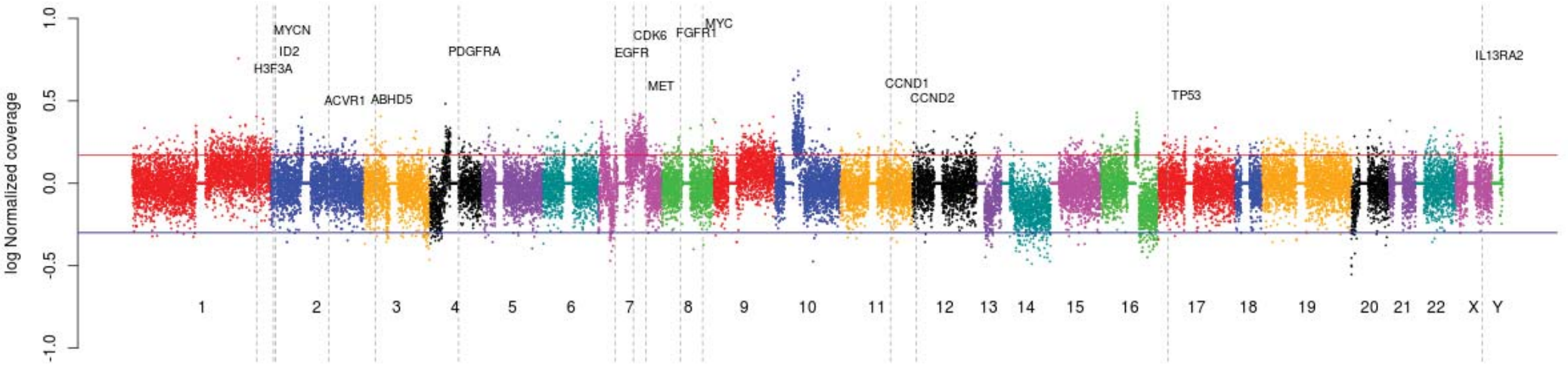
DIPG3-Pons 1



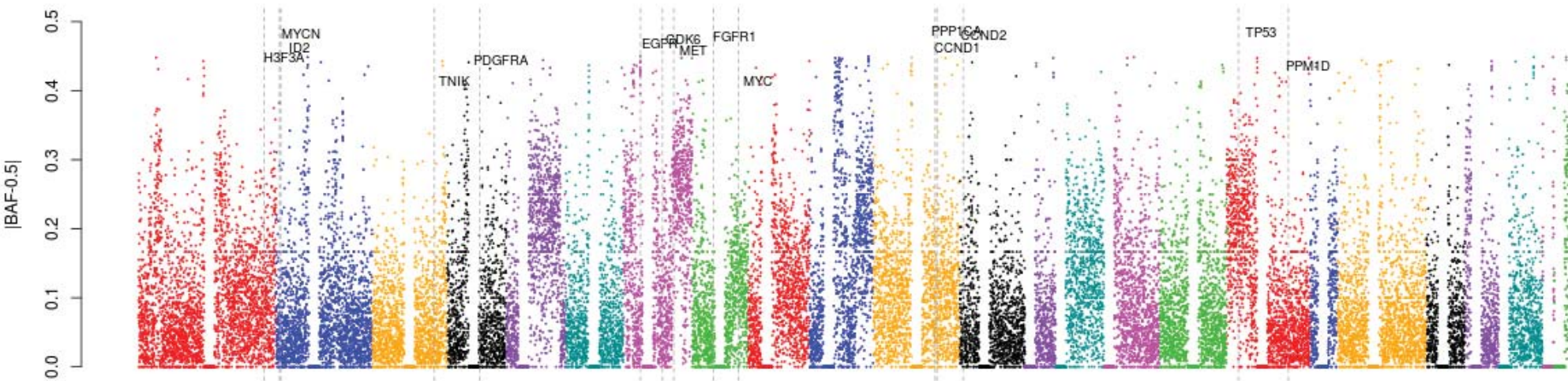
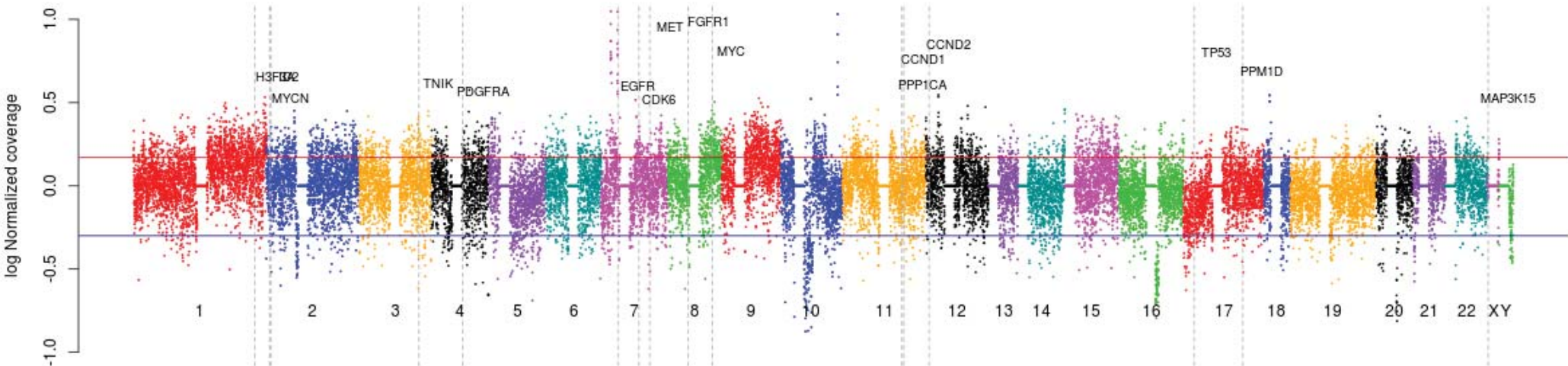
DIPG3-Pons 2



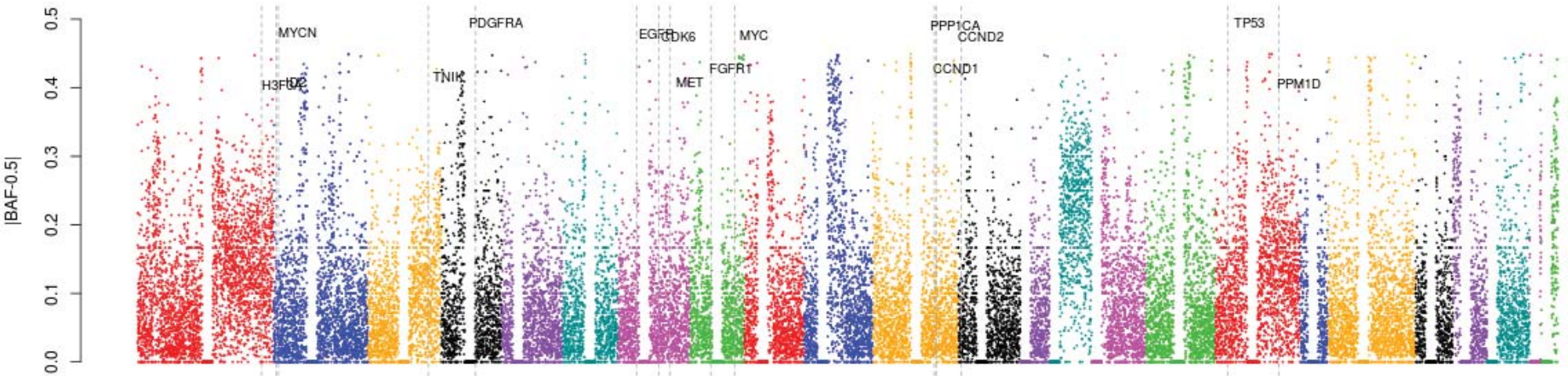
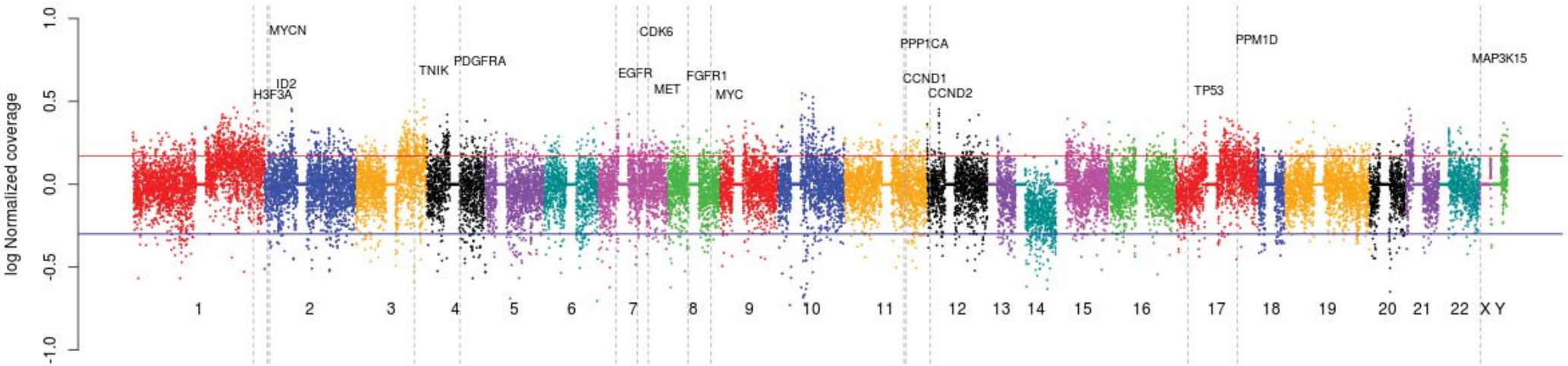
DIPG4-Pons 1



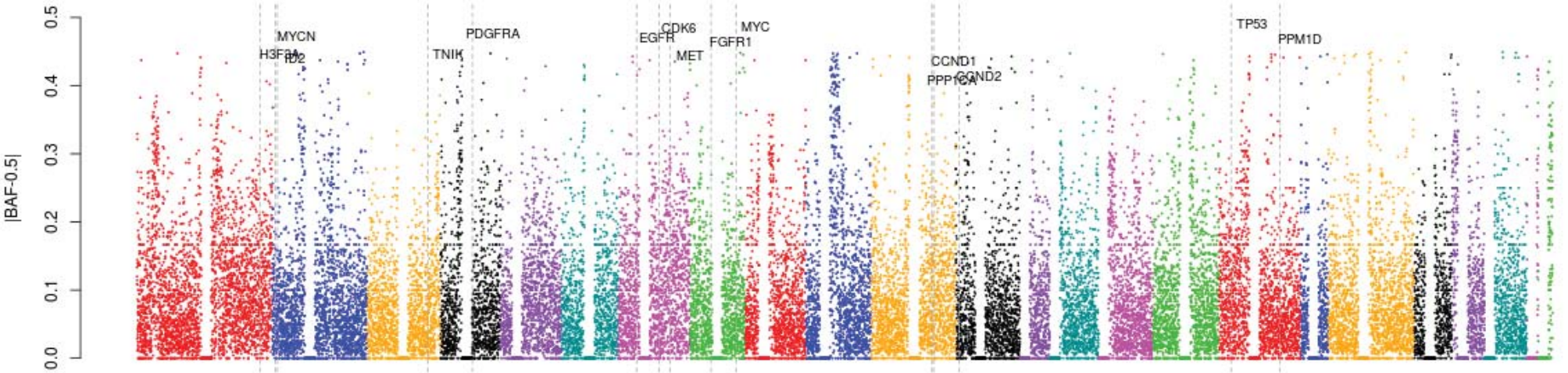
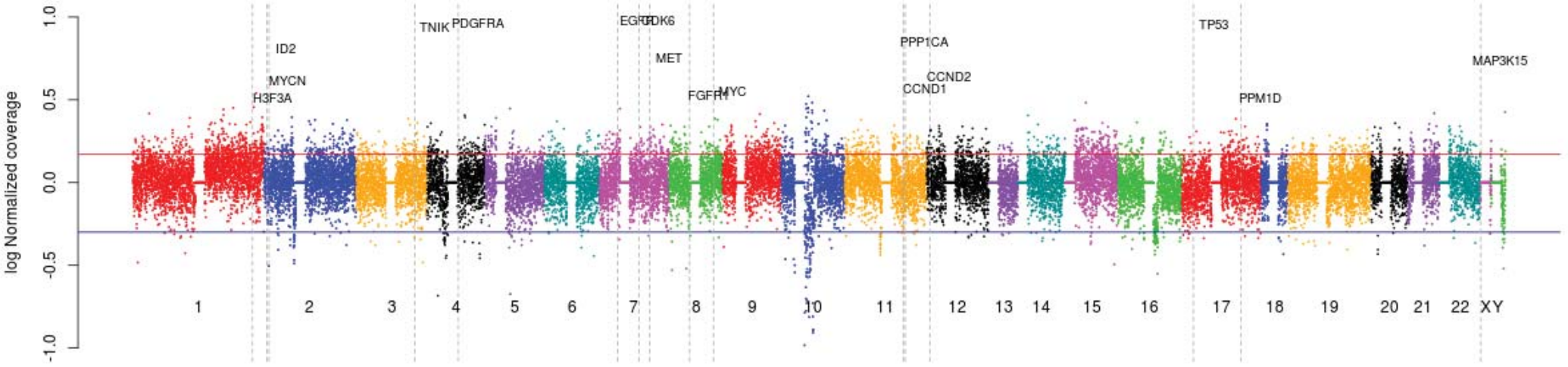
DIPG5-Medulla



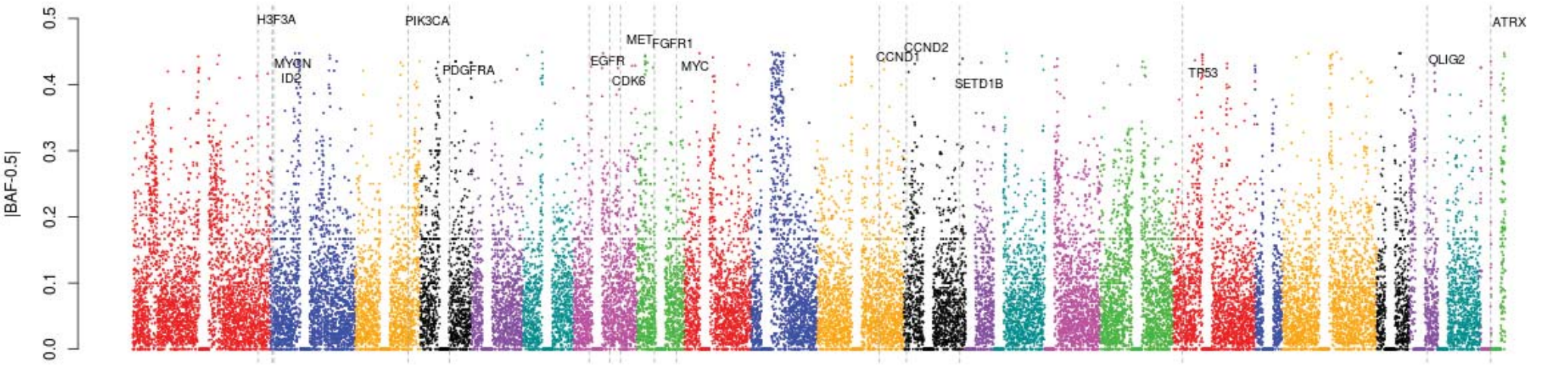
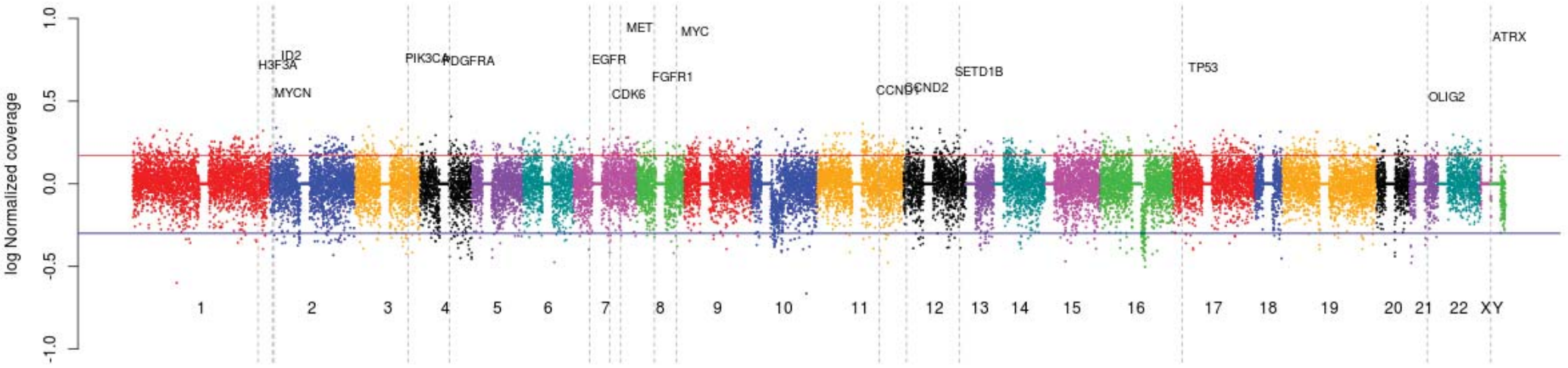
DIPG5-Midbrain 1



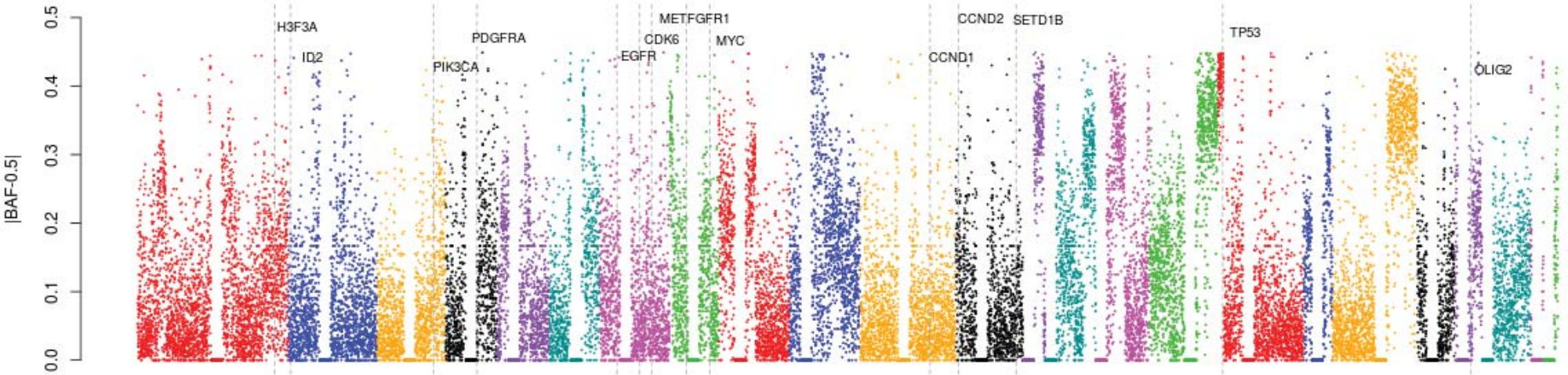
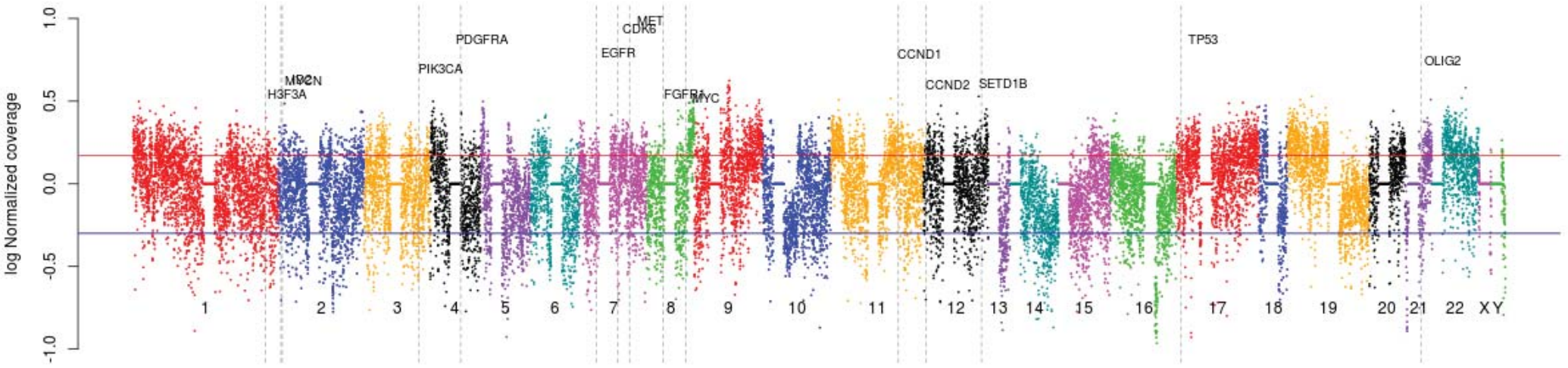
DIPG5-Pons 1



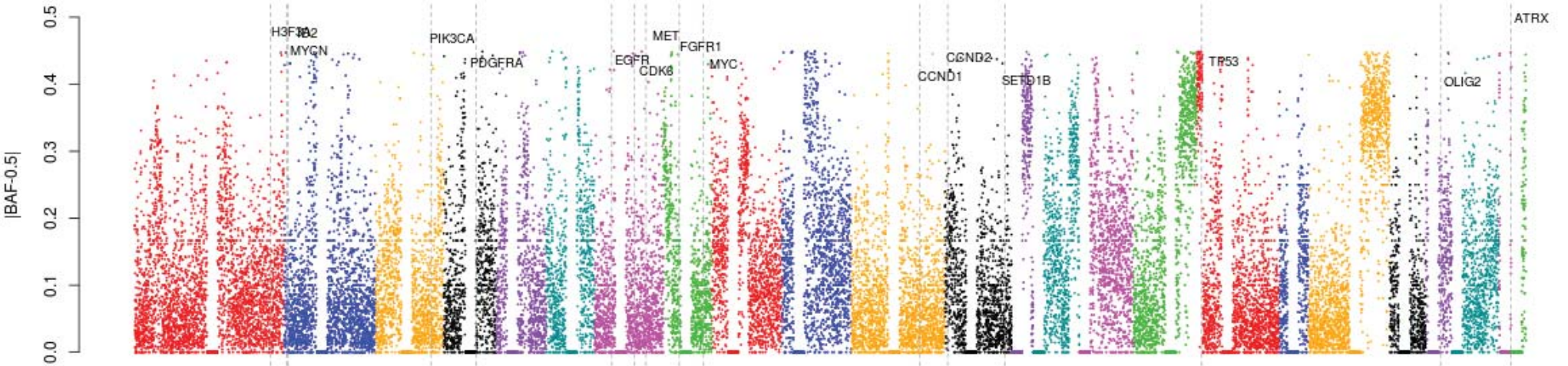
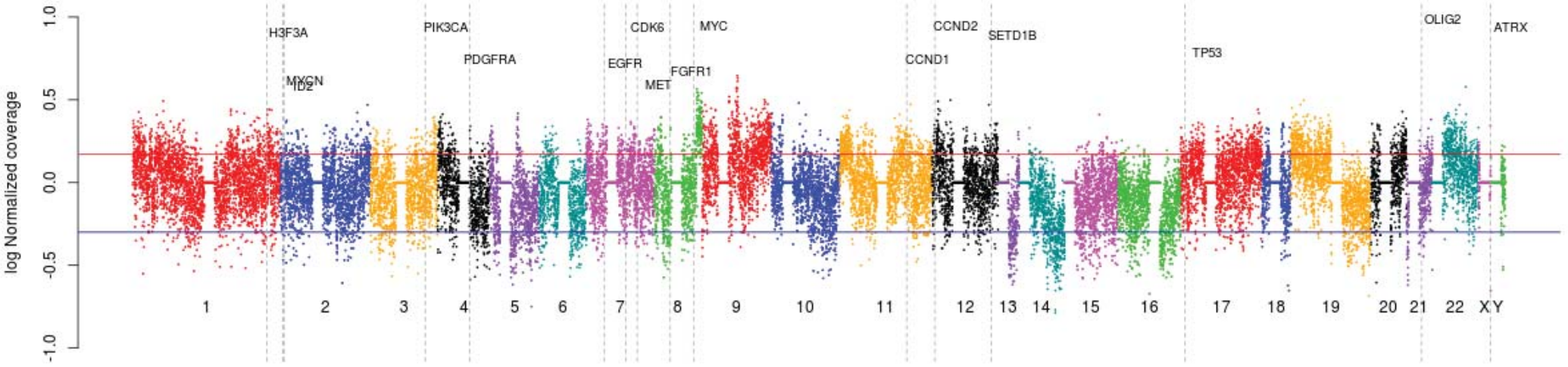
DIPG6-Midbrain



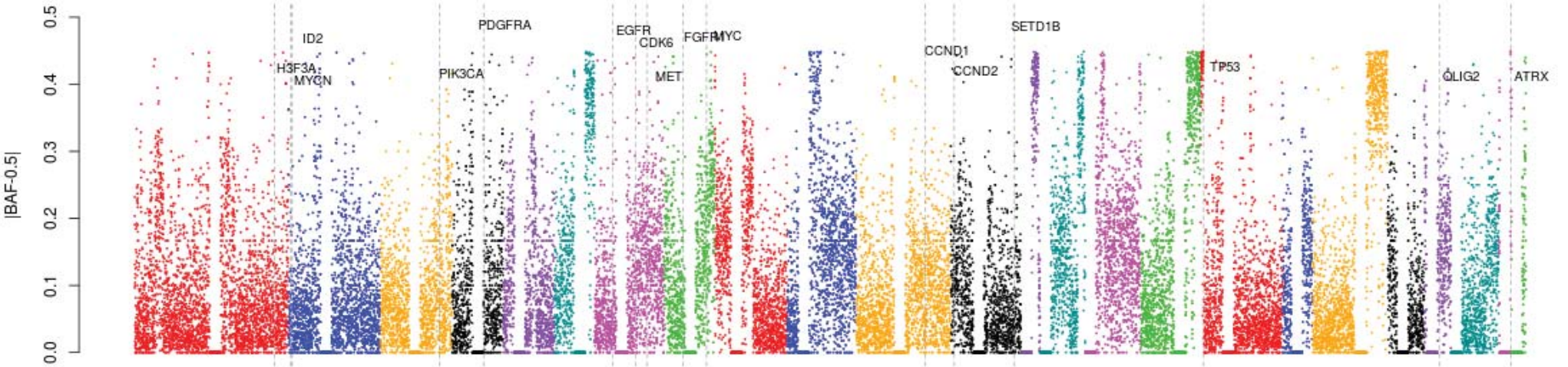
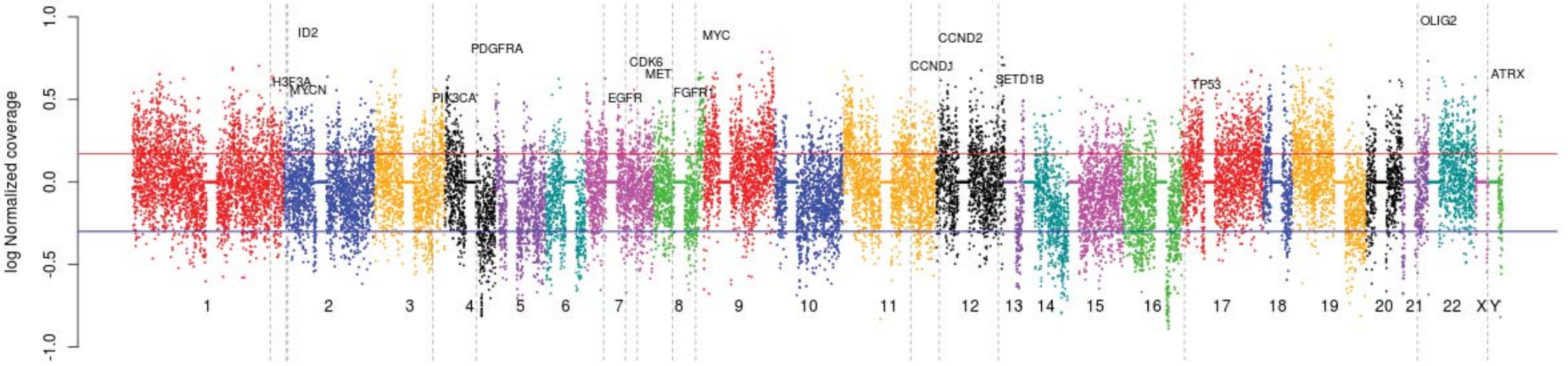
DIPG6-Pons 1



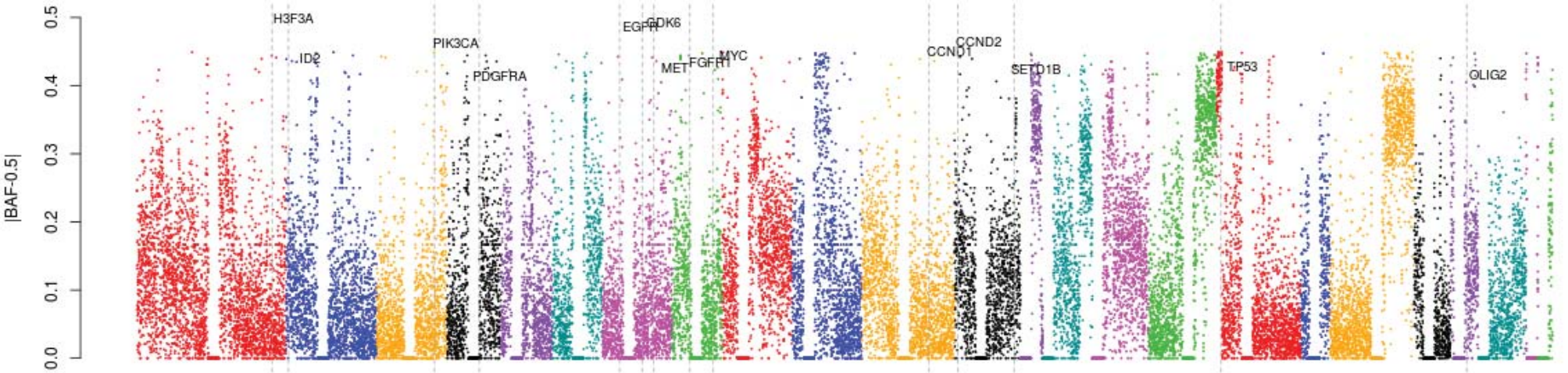
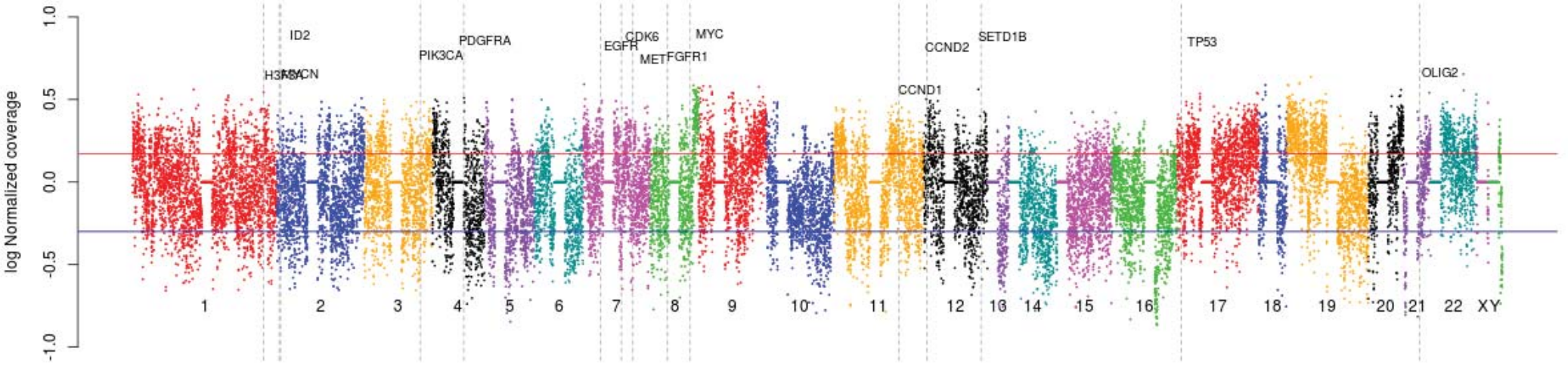
DIPG6-Pons 2



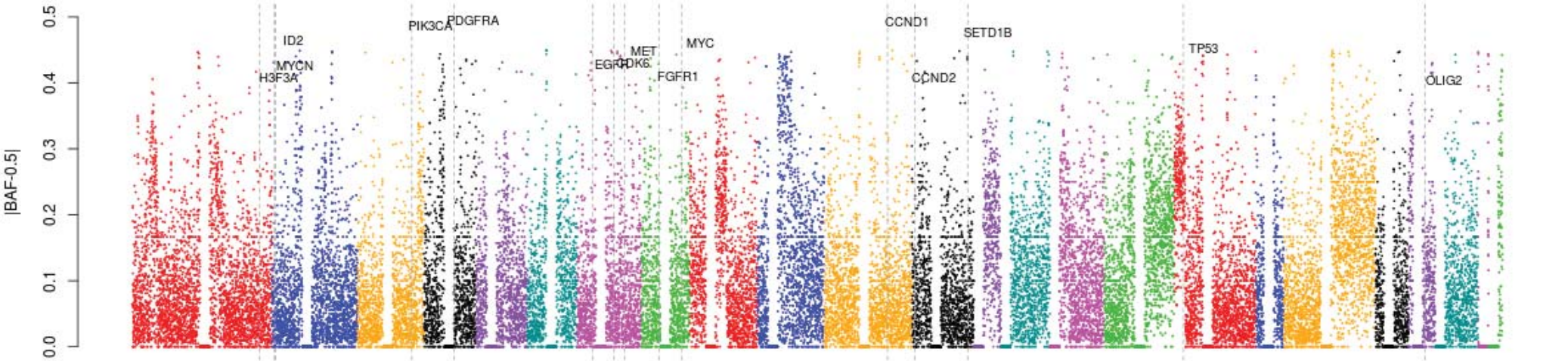
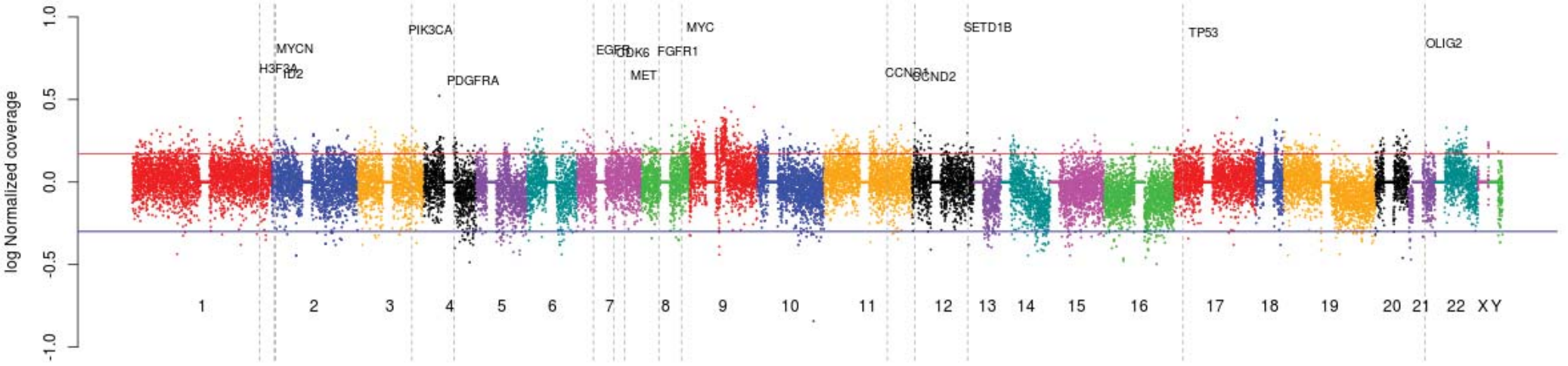
DIPG6-Pons 3



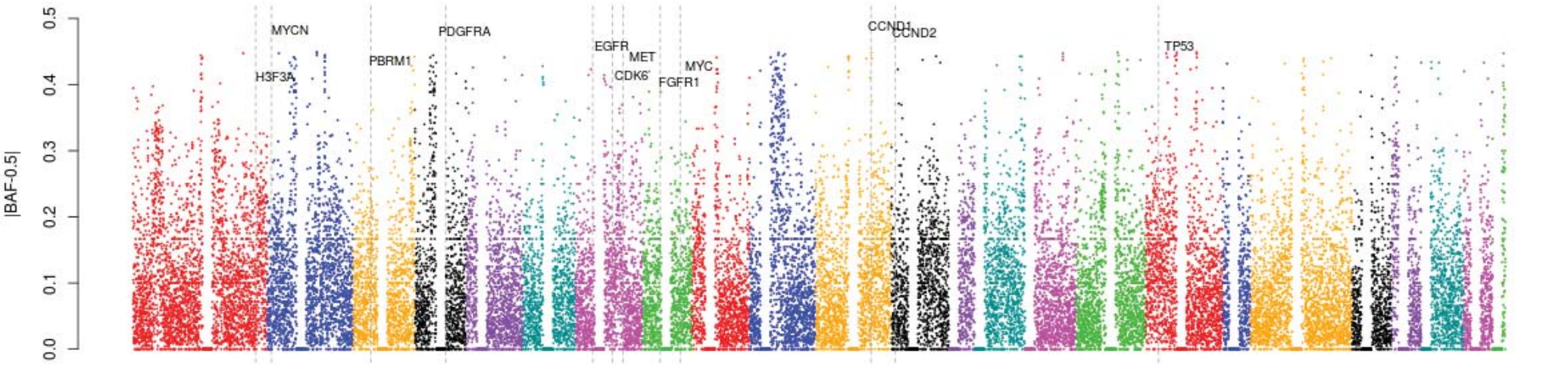
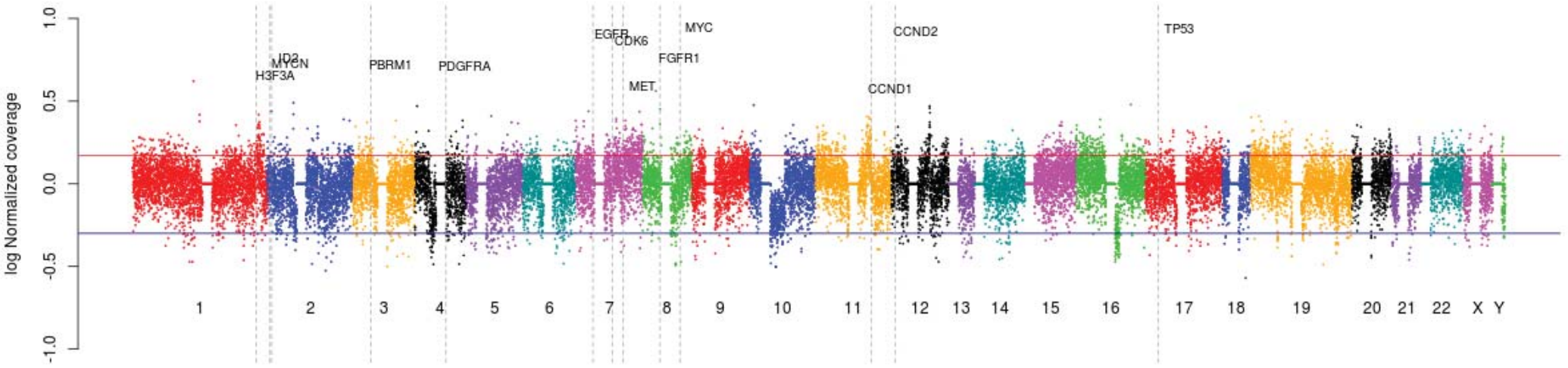
DIPG6-Pons 4



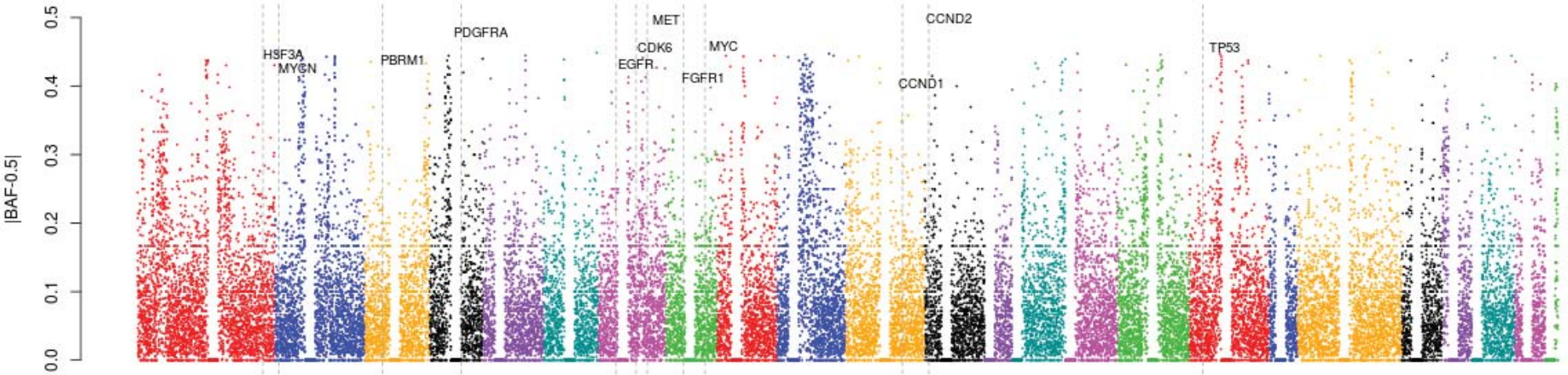
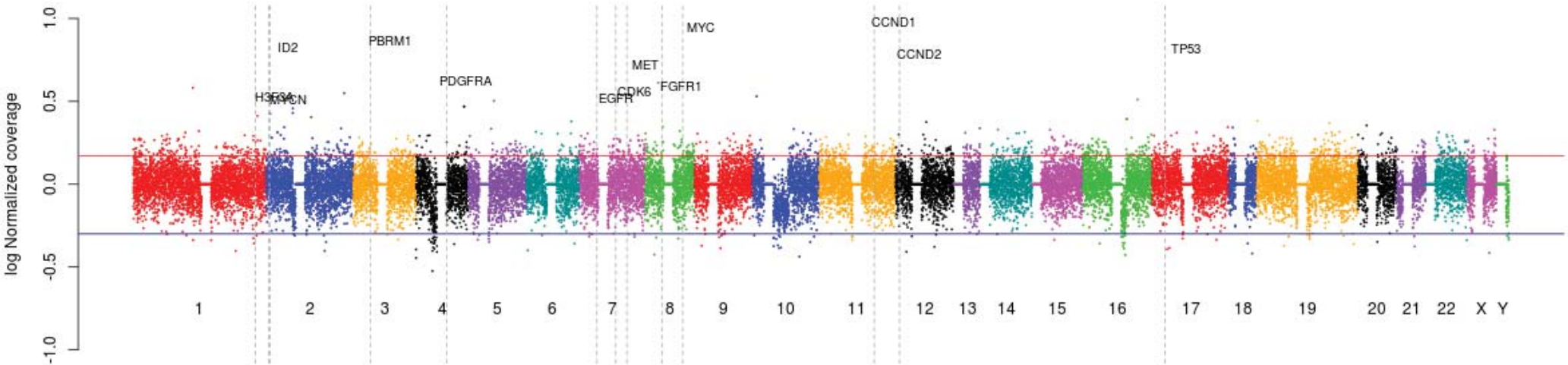
DIPG6-Pons 5



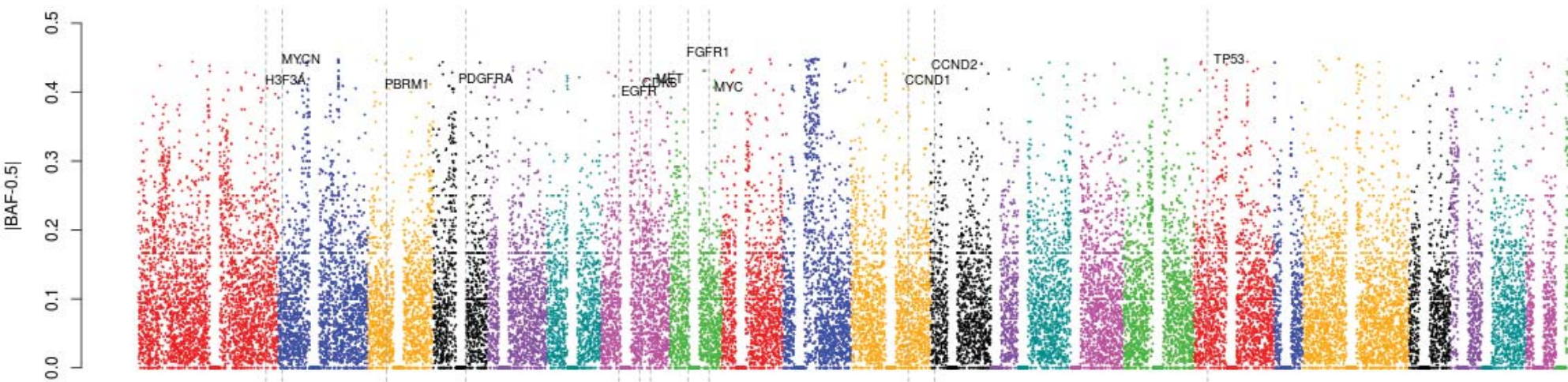
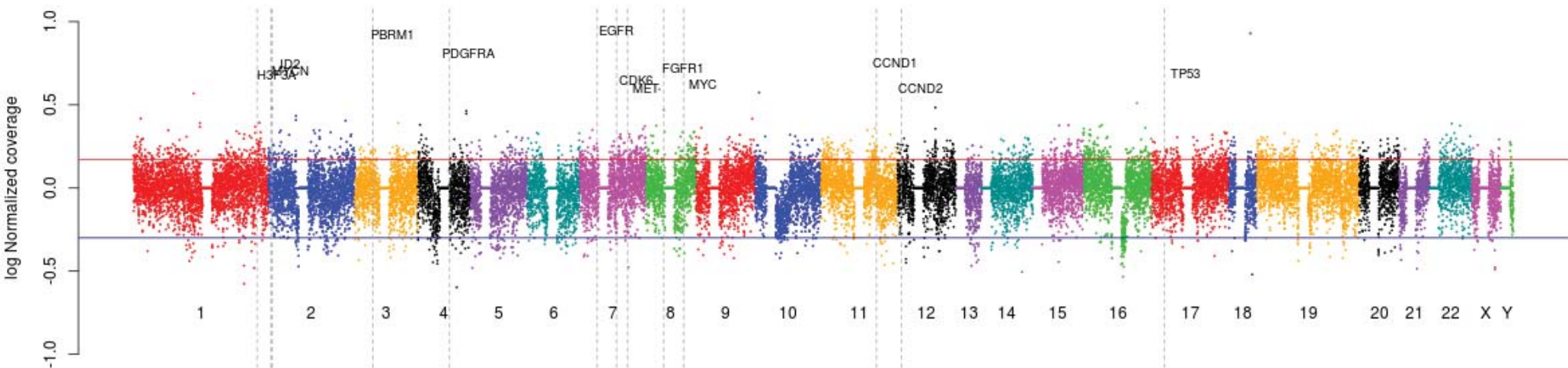
DIPG7-Cerebellum 1



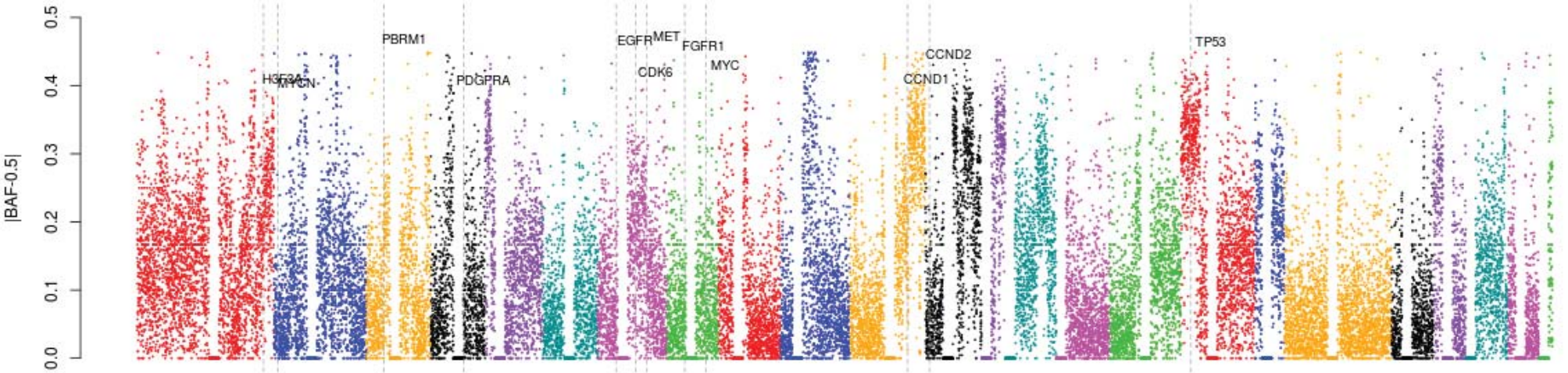
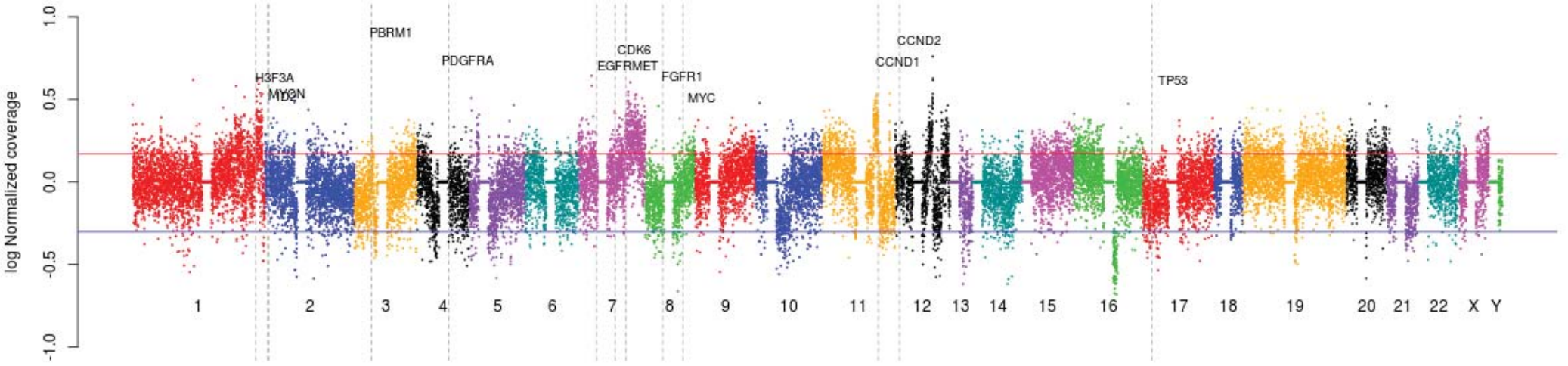
DIPG7-Cerebellum 2



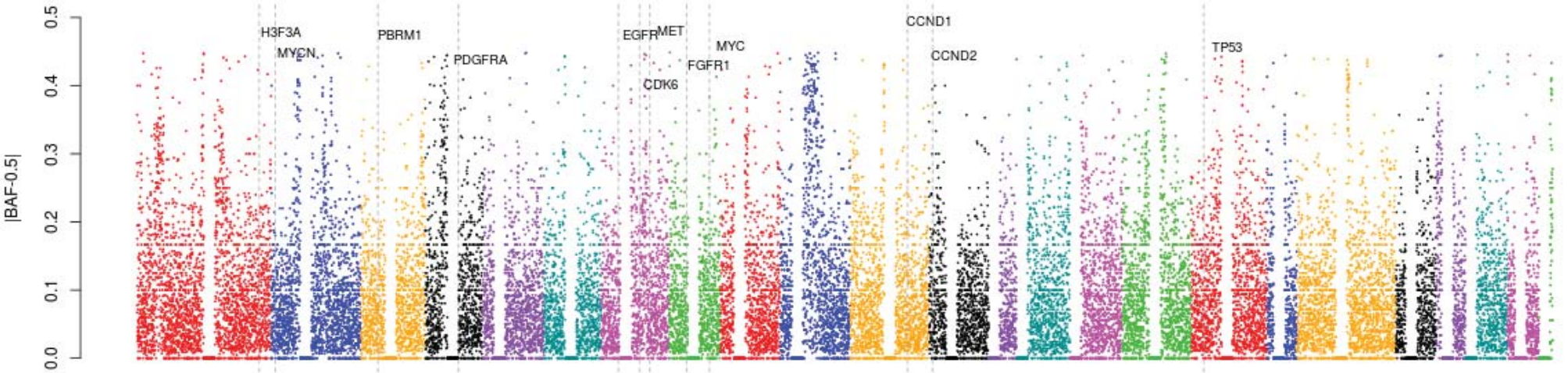
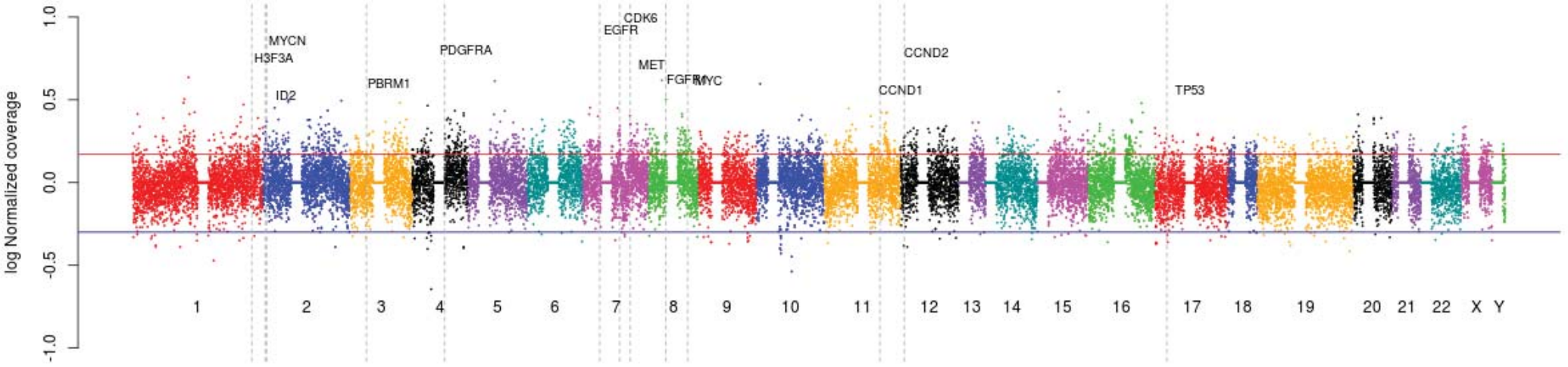
DIPG7-Medulla



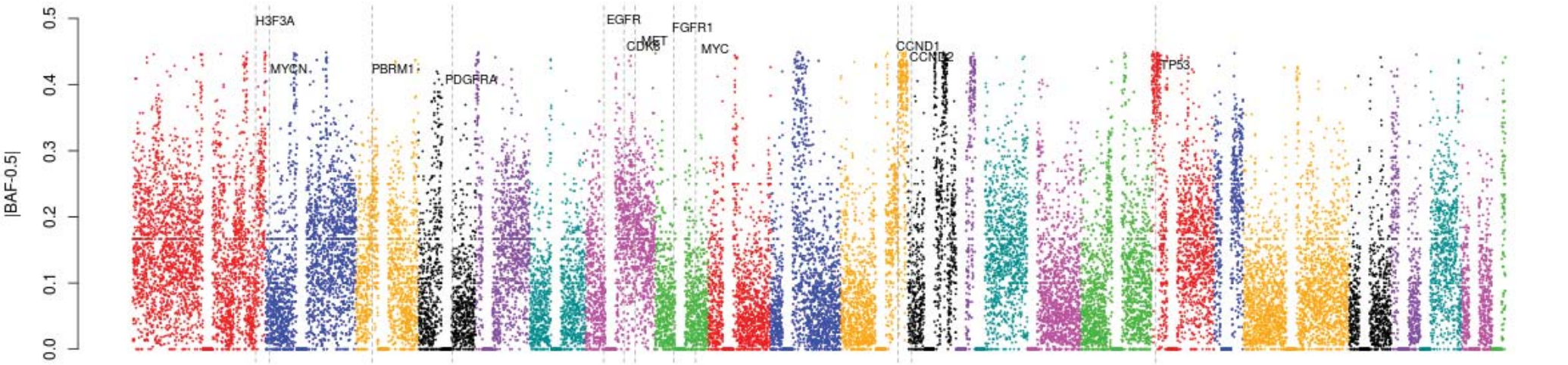
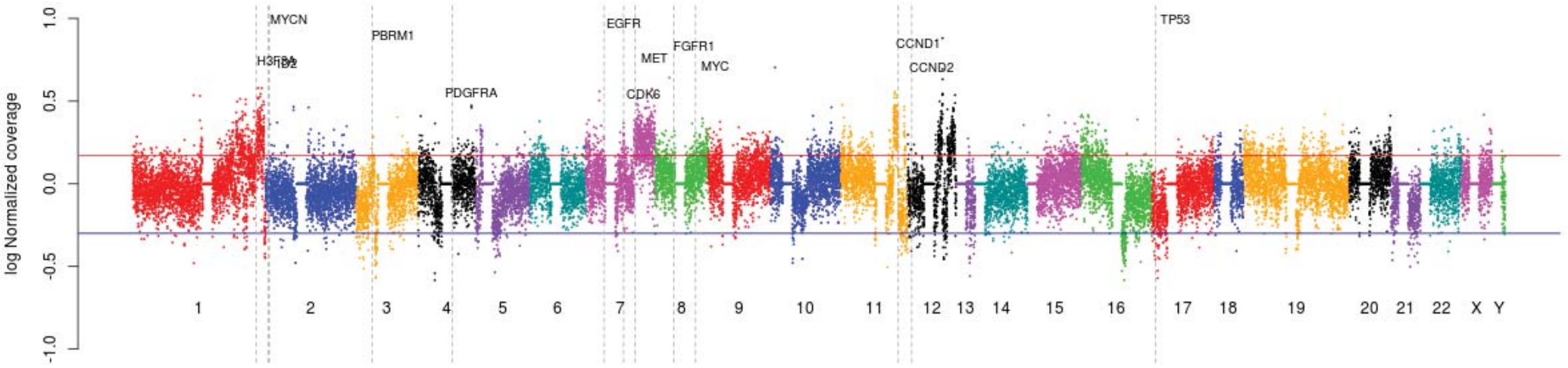
DIPG7-Pons 1



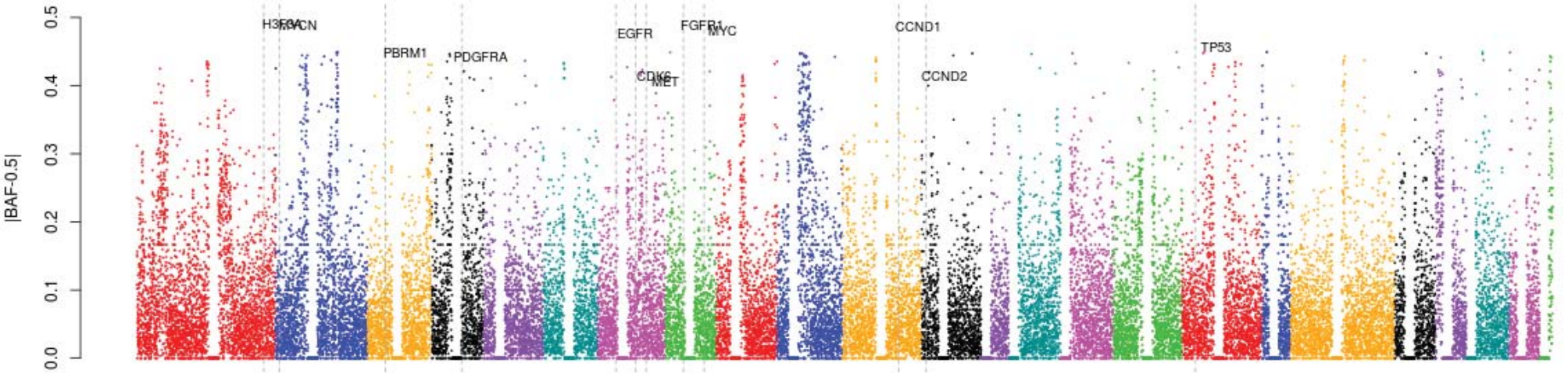
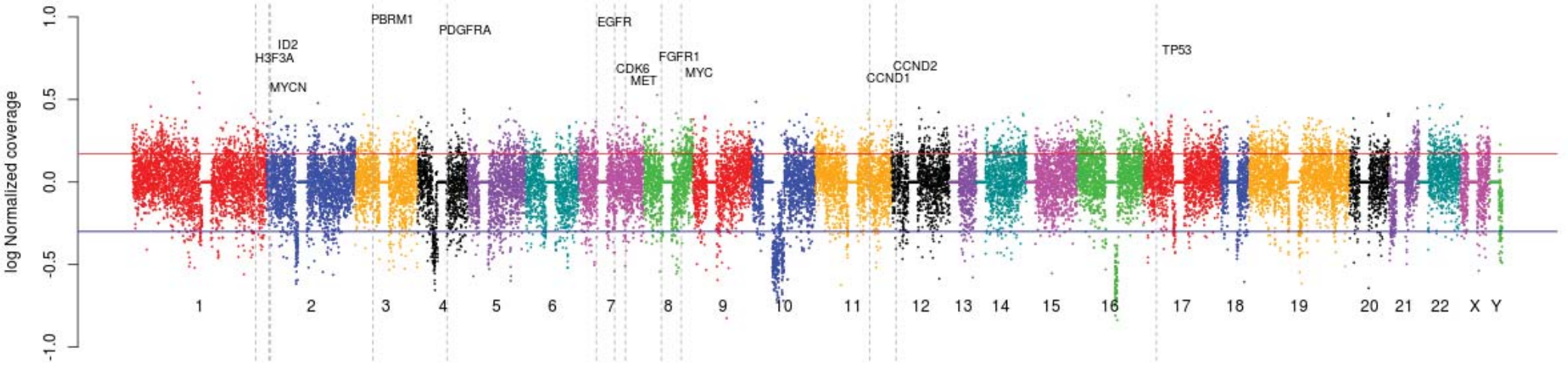
DIPG7-Pons 2



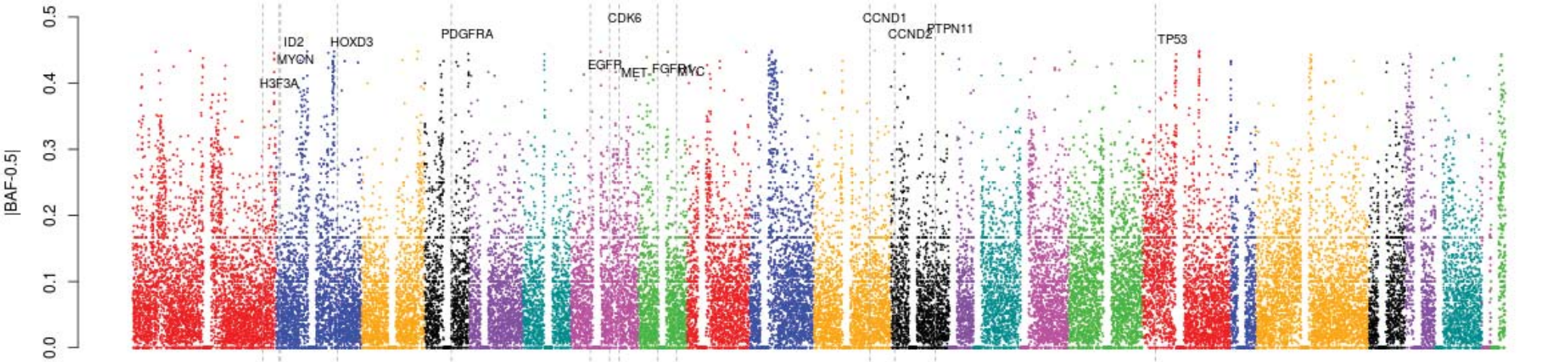
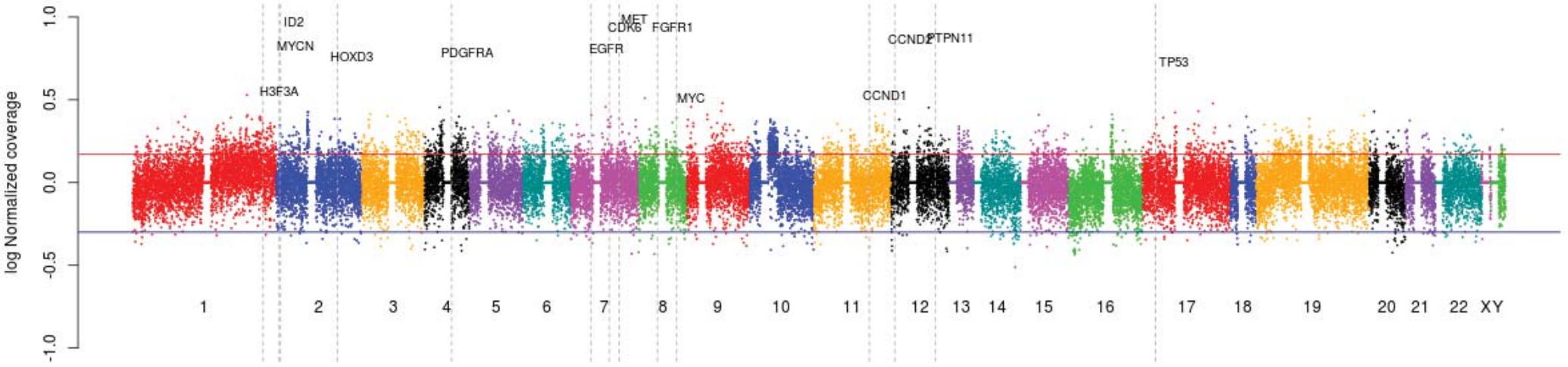
DIPG7-Pons 3



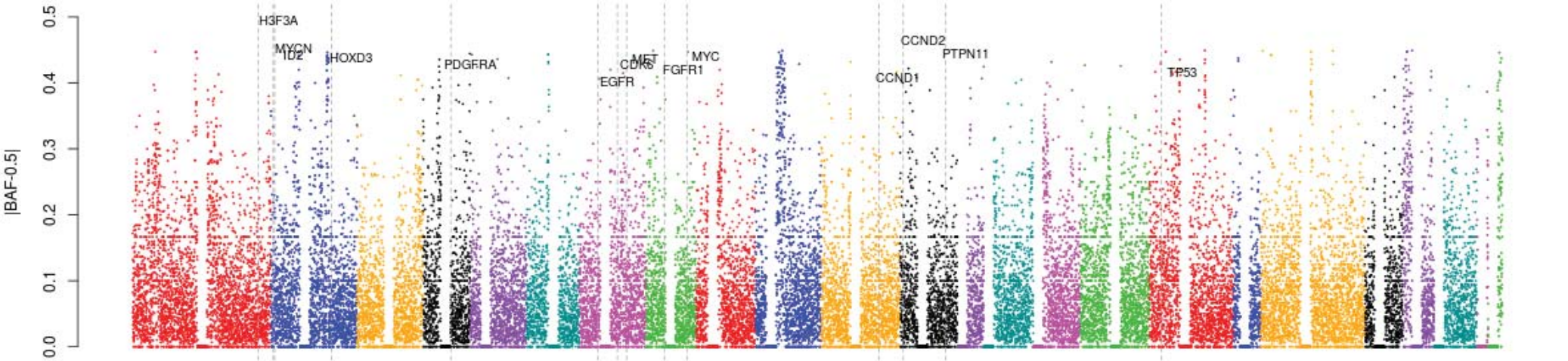
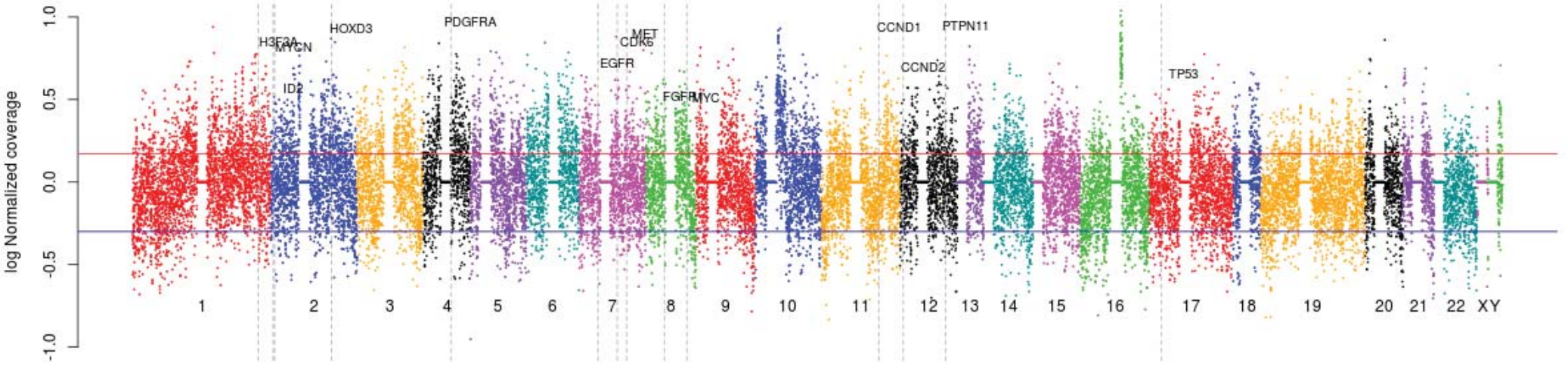
DIPG7-Pons 5



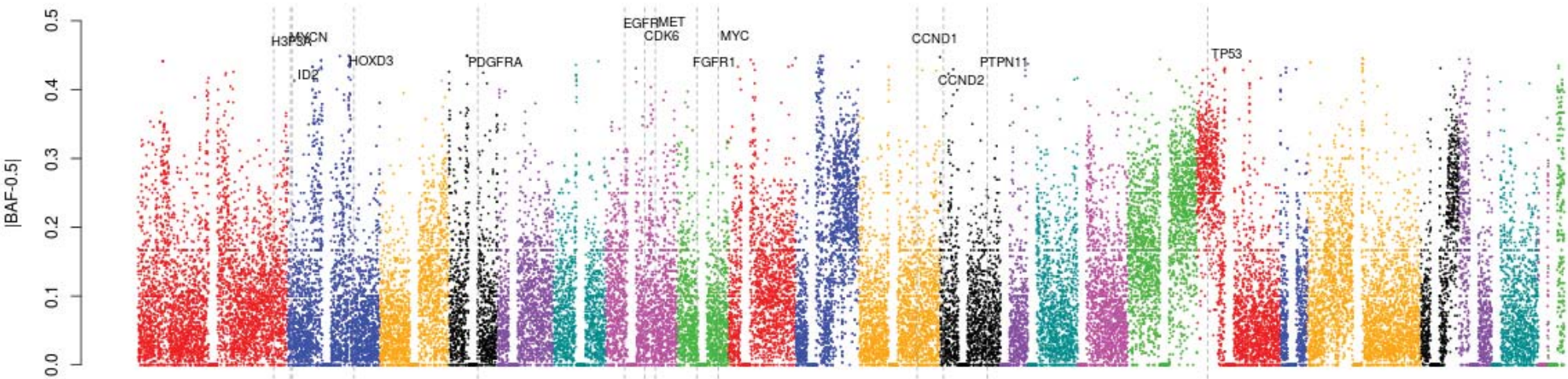
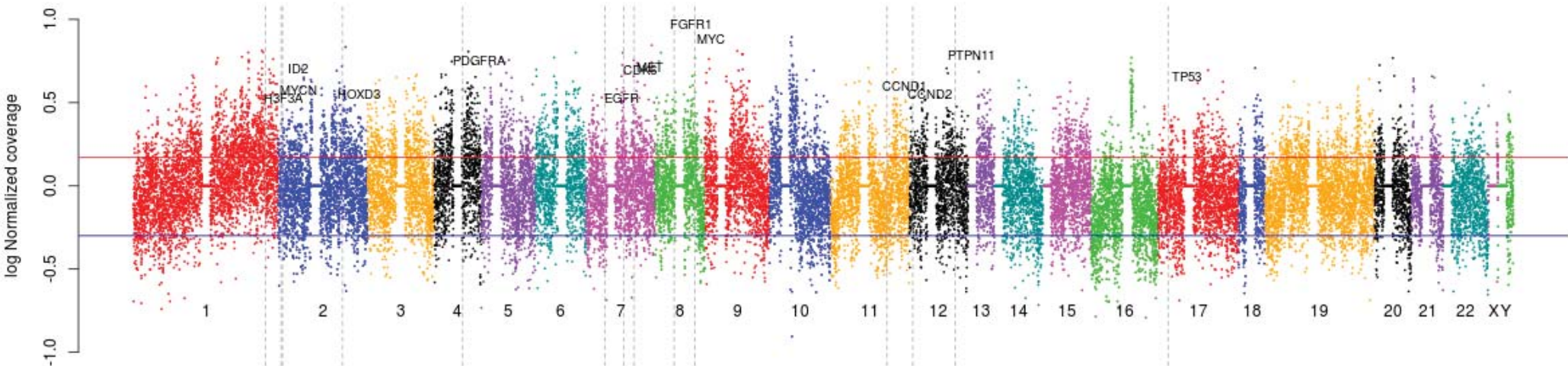
DIPG8-Cerebellum 1



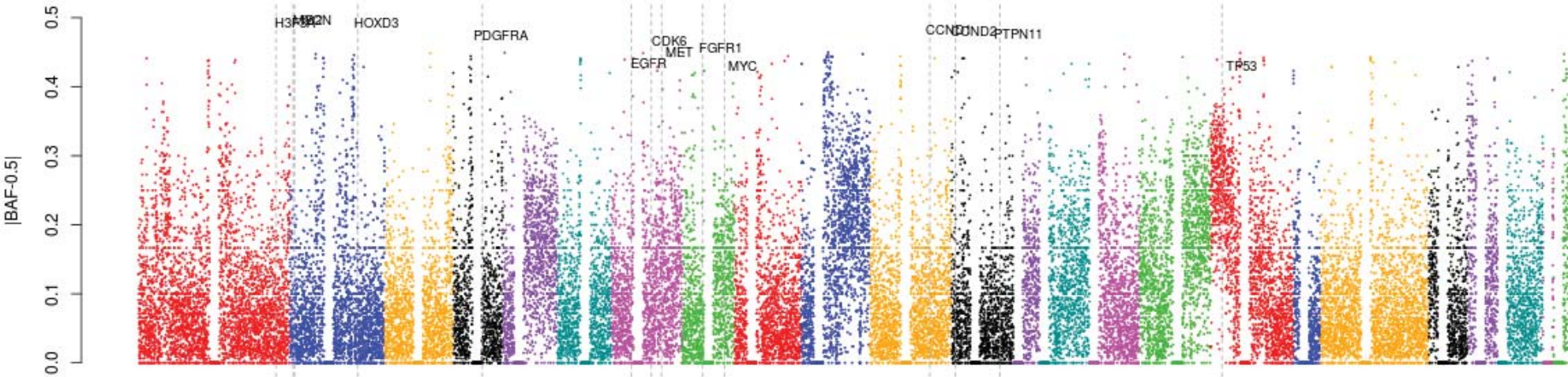
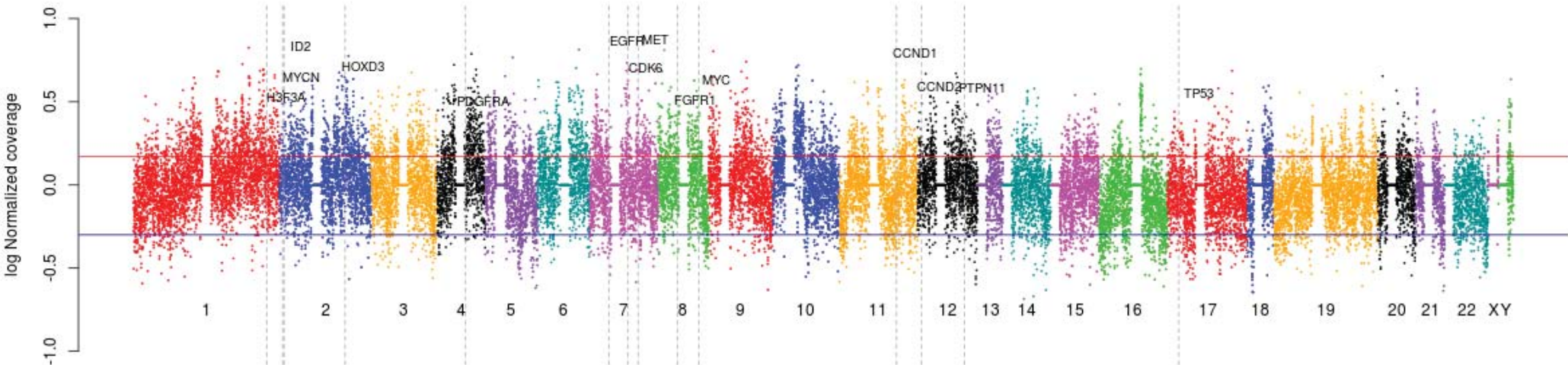
DIPG8- Frontal Lobe 1



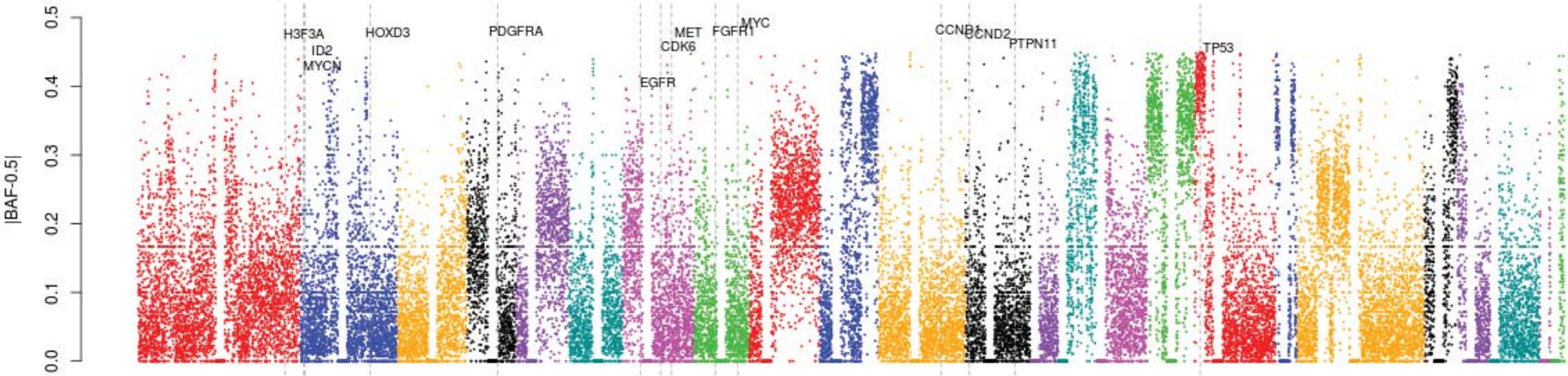
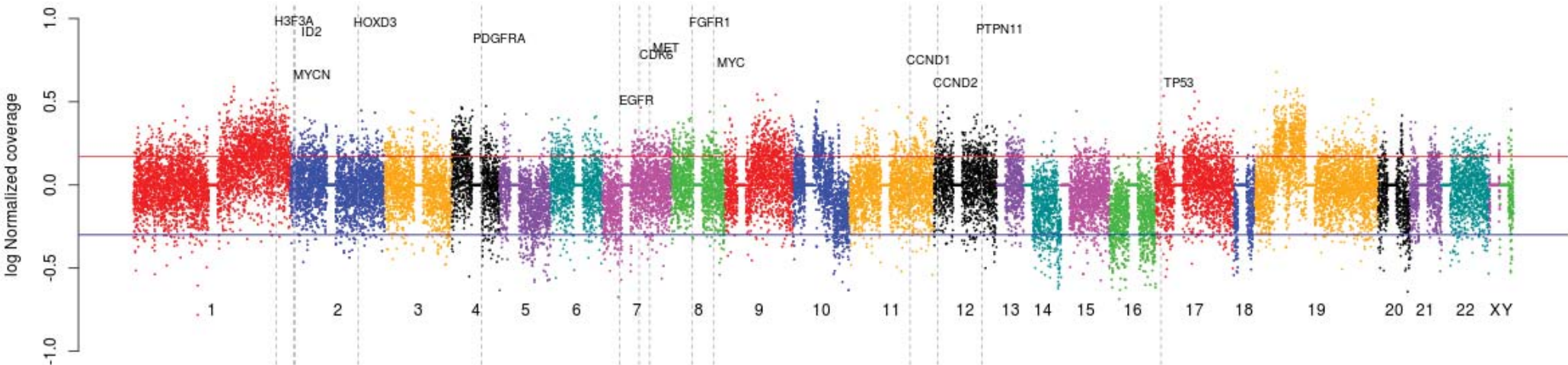
DIPG8-Medulla 1



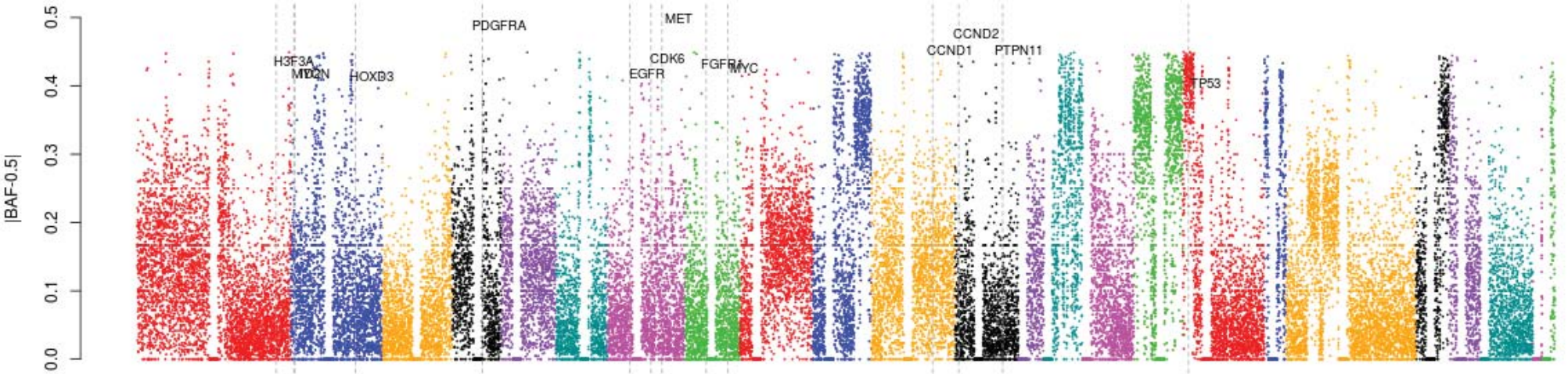
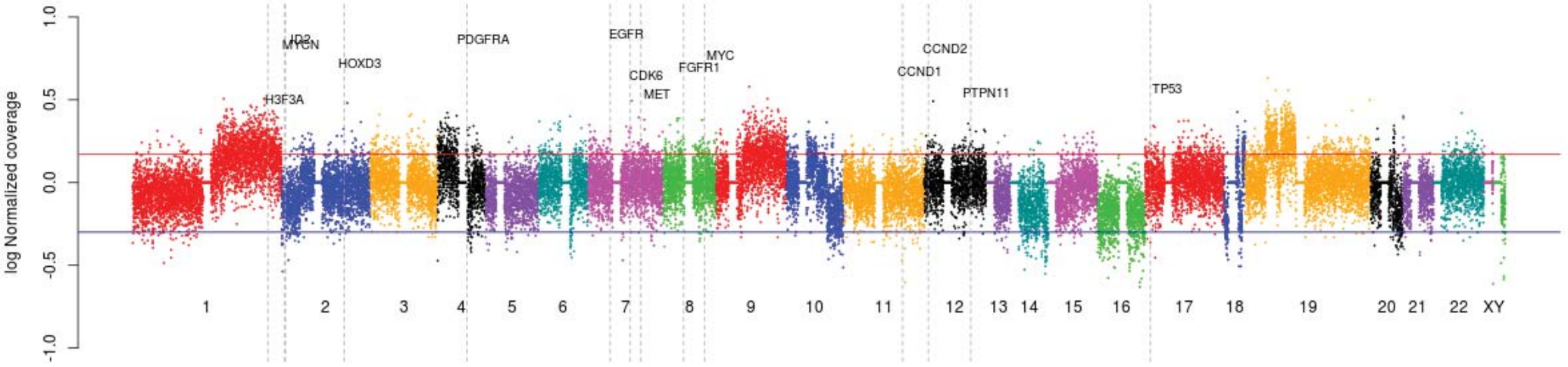
DIPG8-Medulla 2



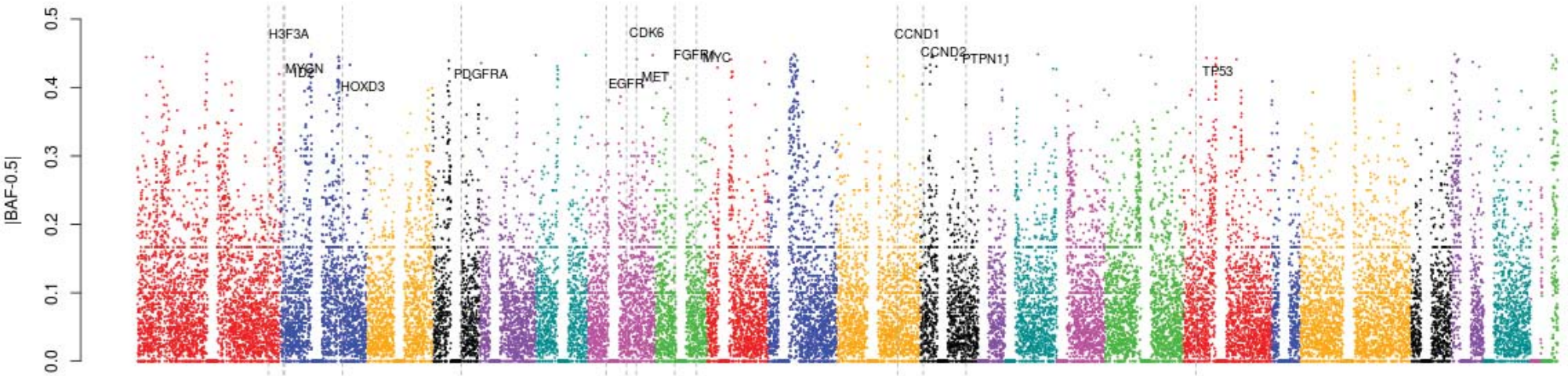
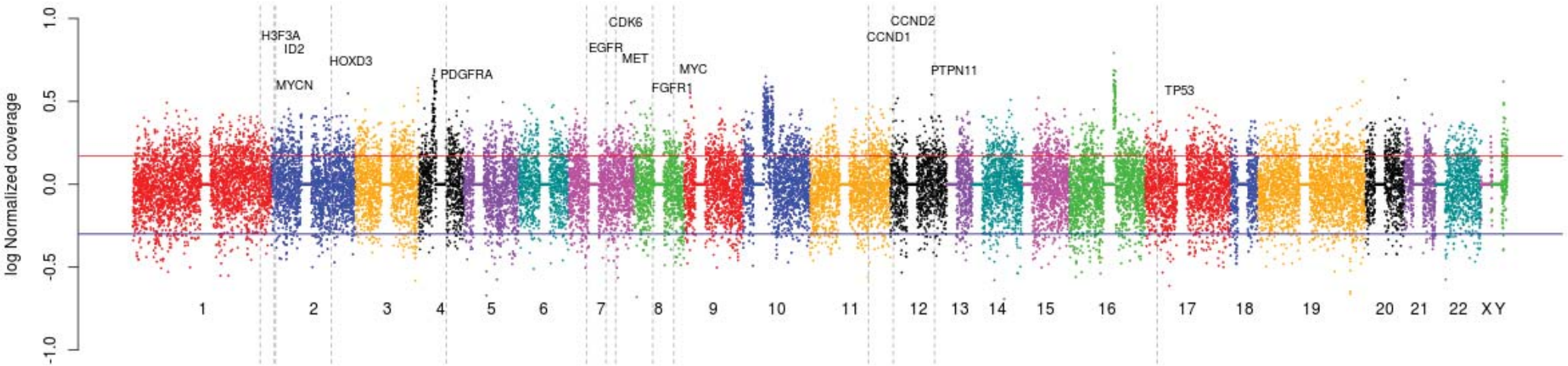
DIPG8-Medulla 3



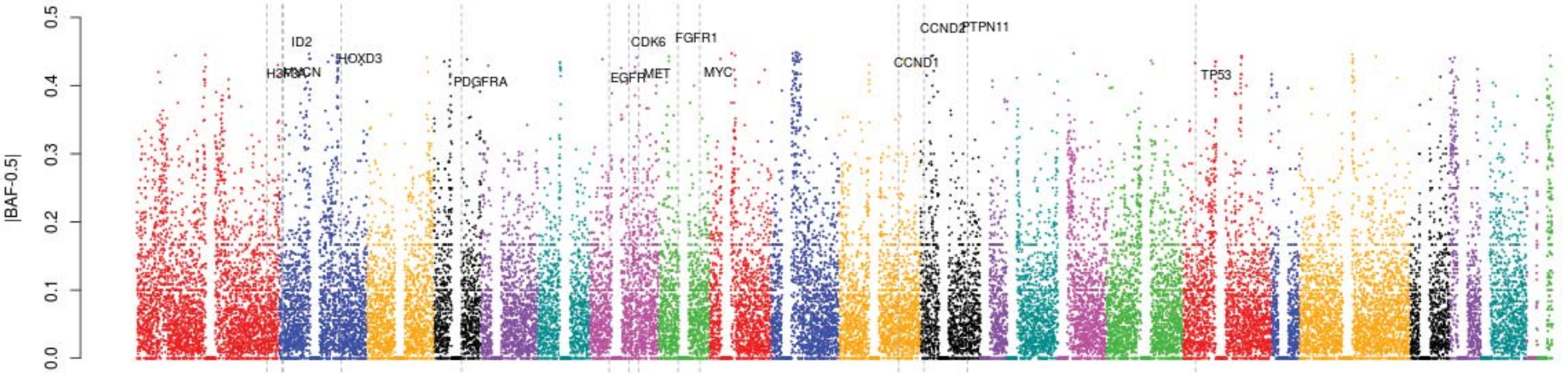
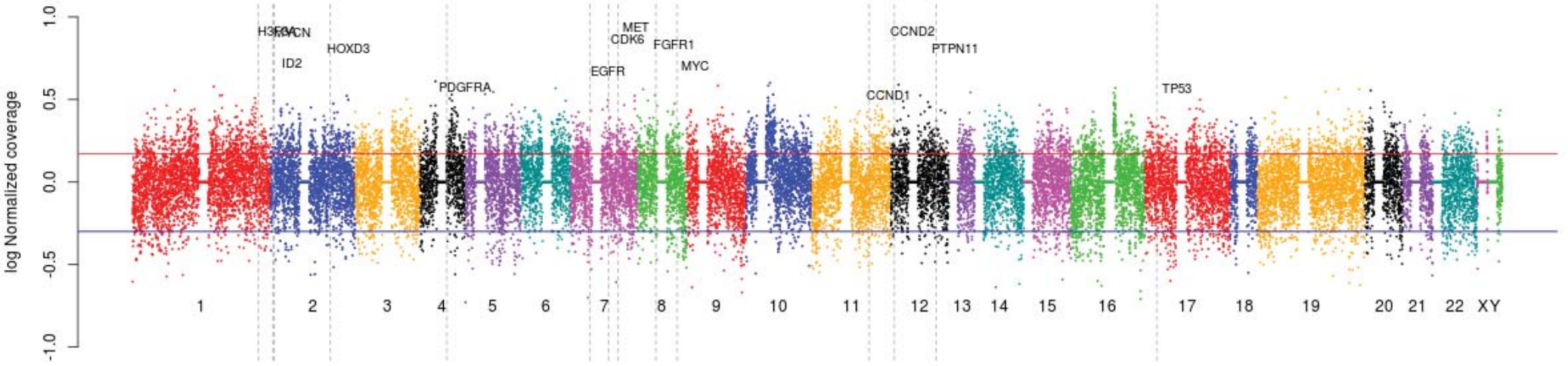
DIPG8-Midbrain



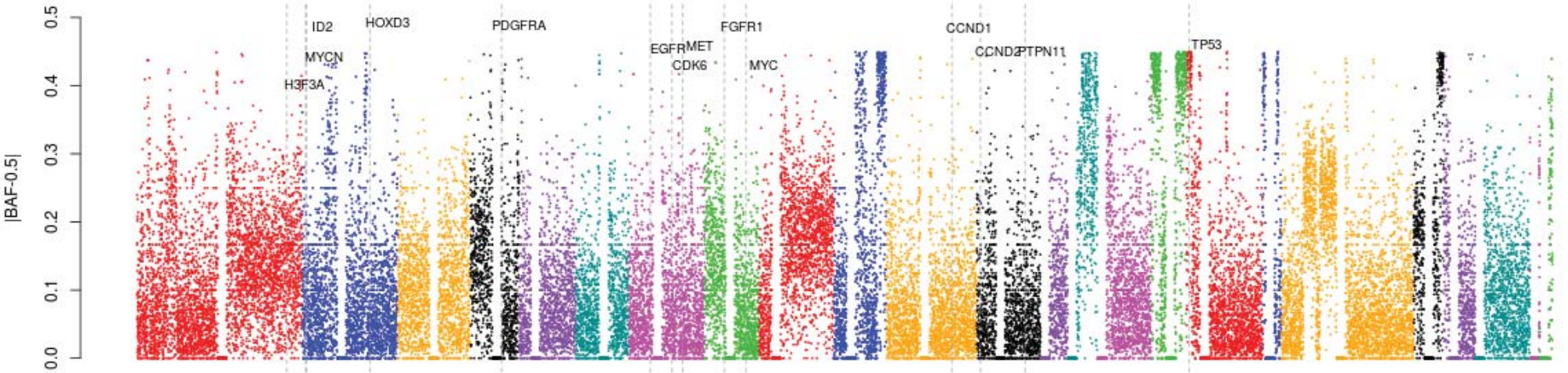
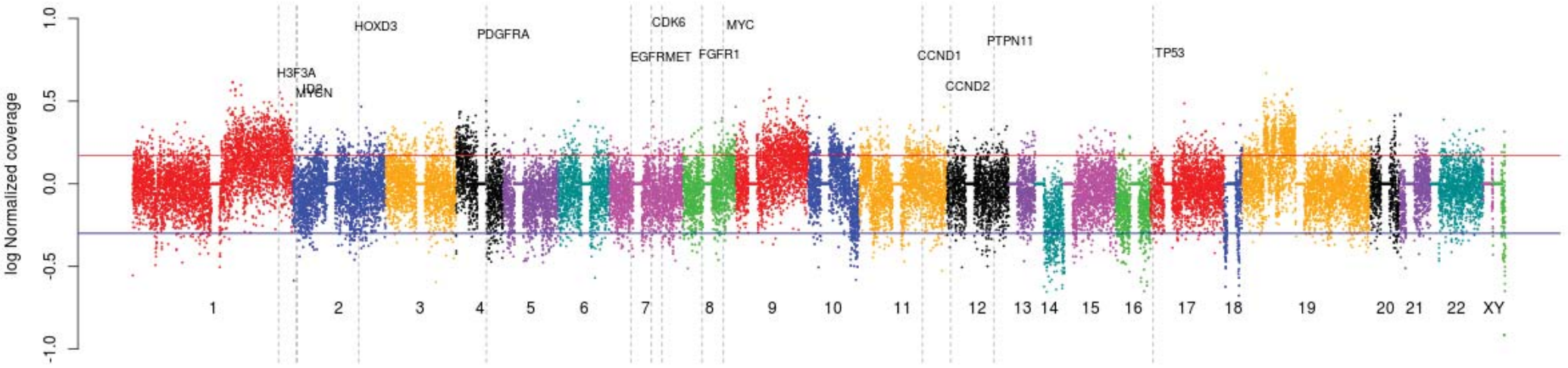
DIPG8-Occipital Lobe



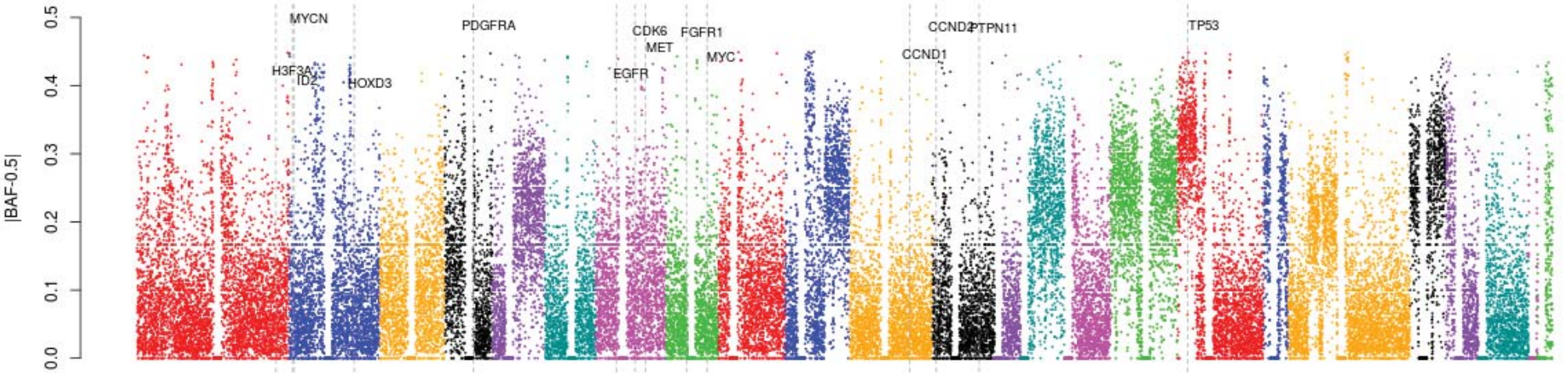
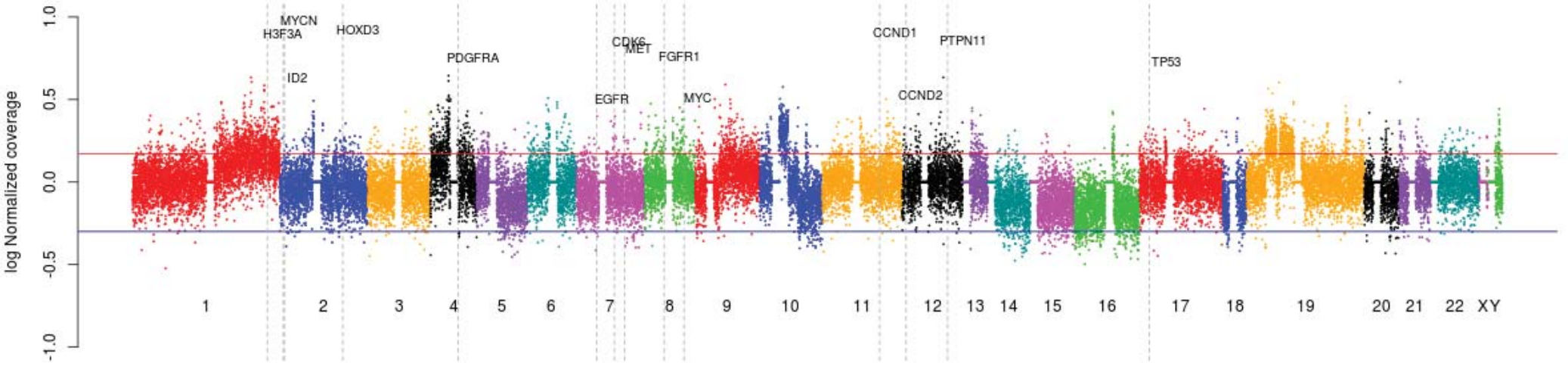
DIPG8-Parietal Lobe 2



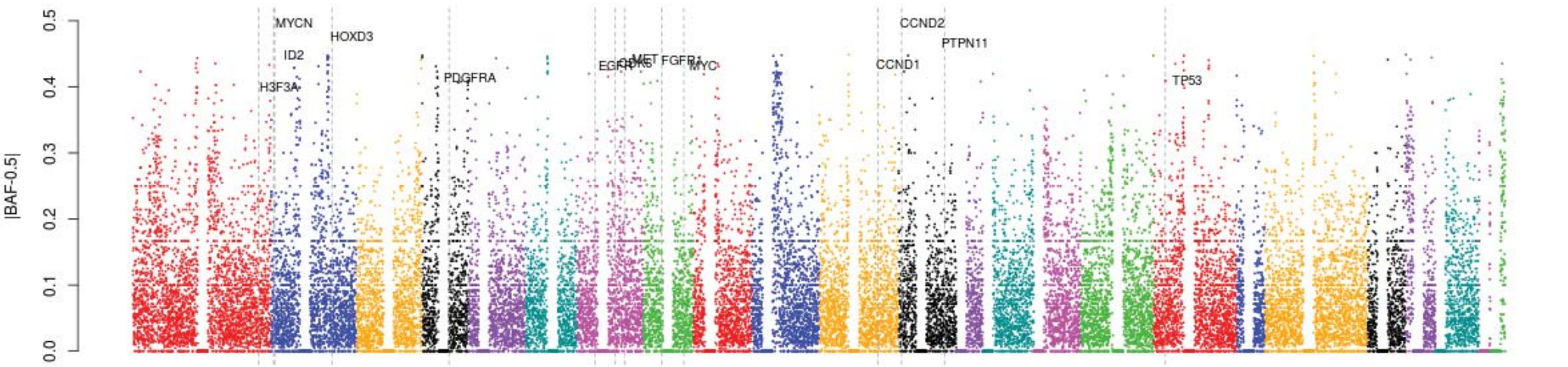
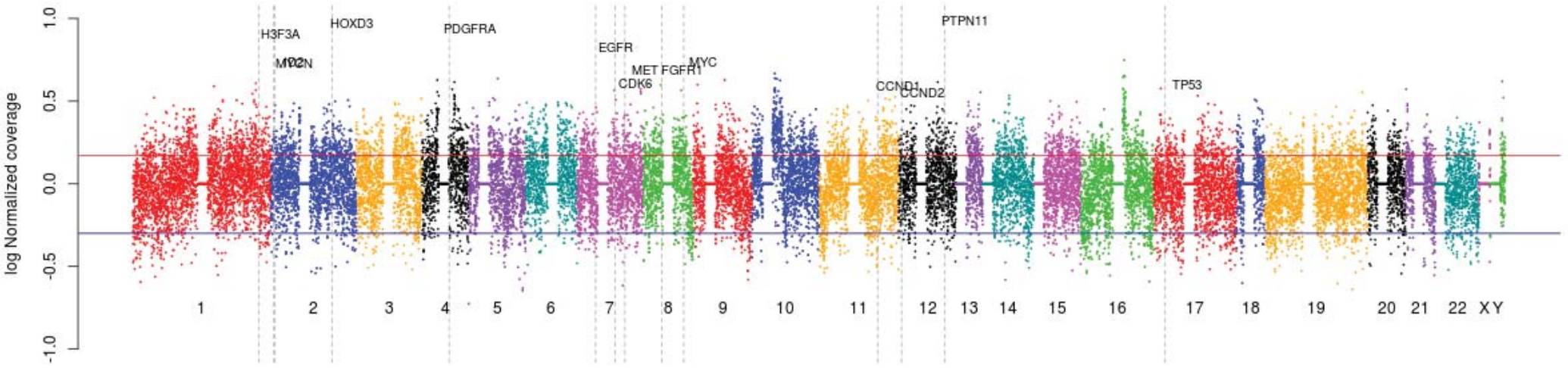
DIPG8-Pons 1



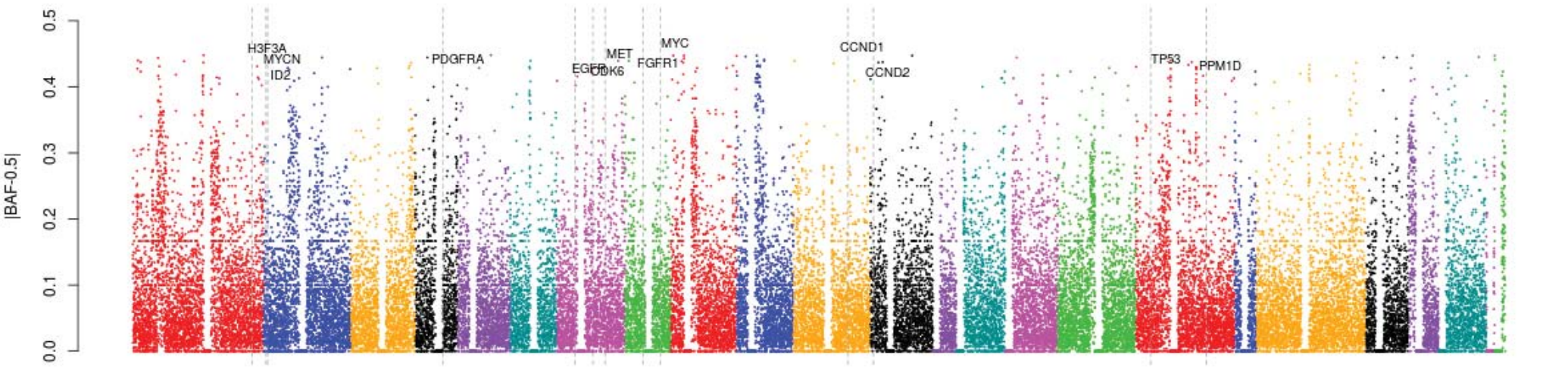
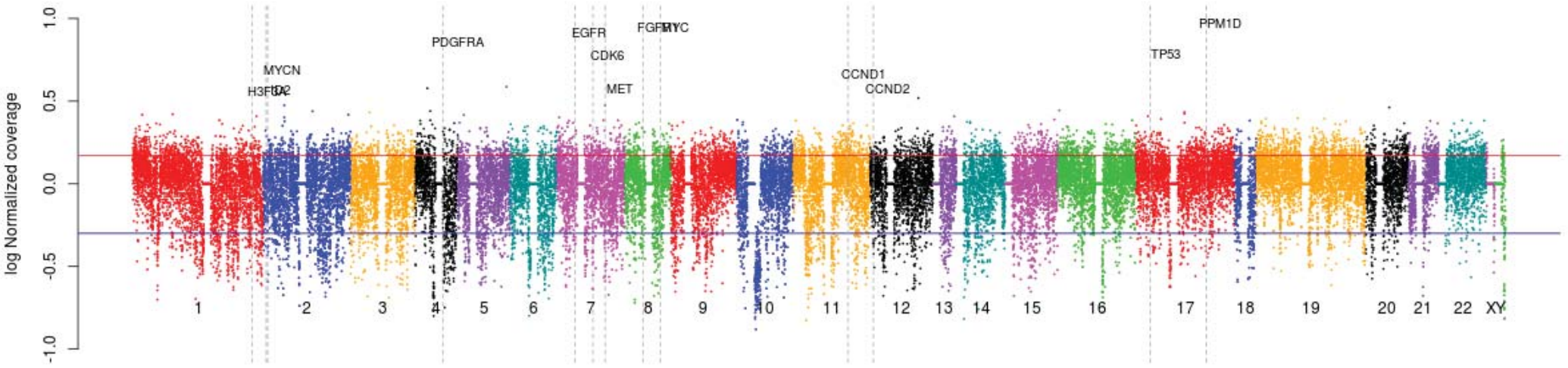
DIPG8-Pons 2



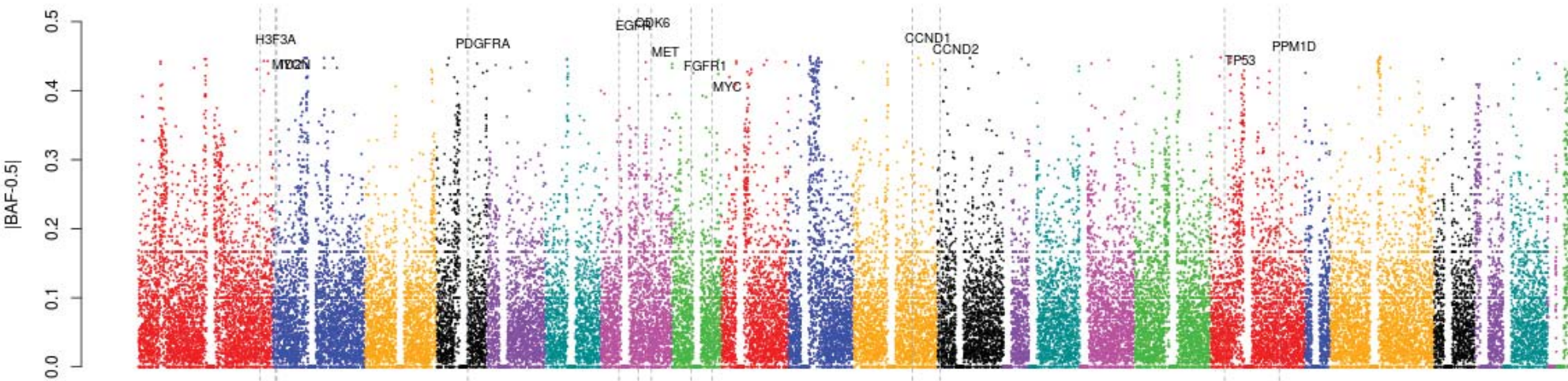
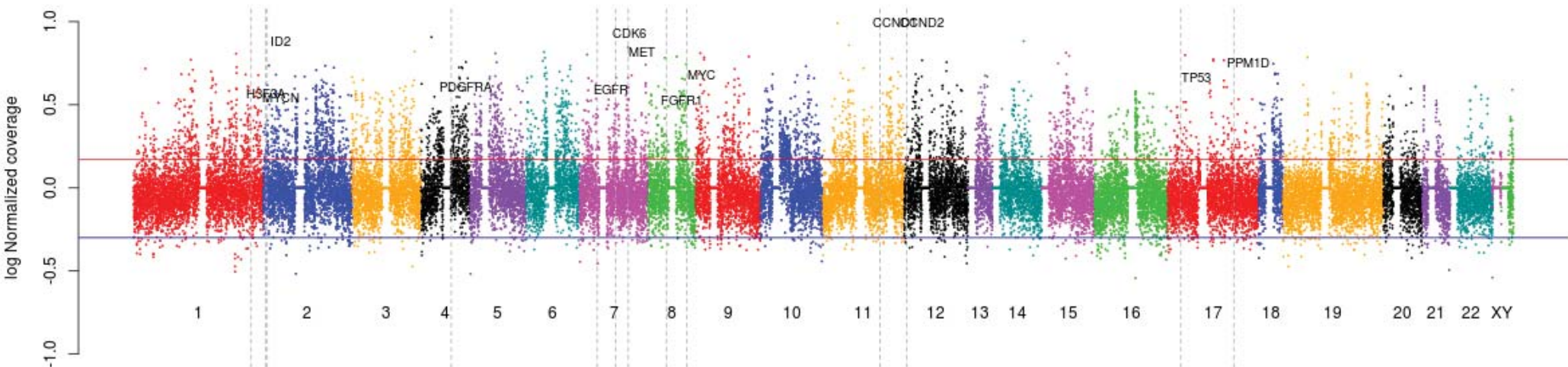
DIPG8-Ventricle



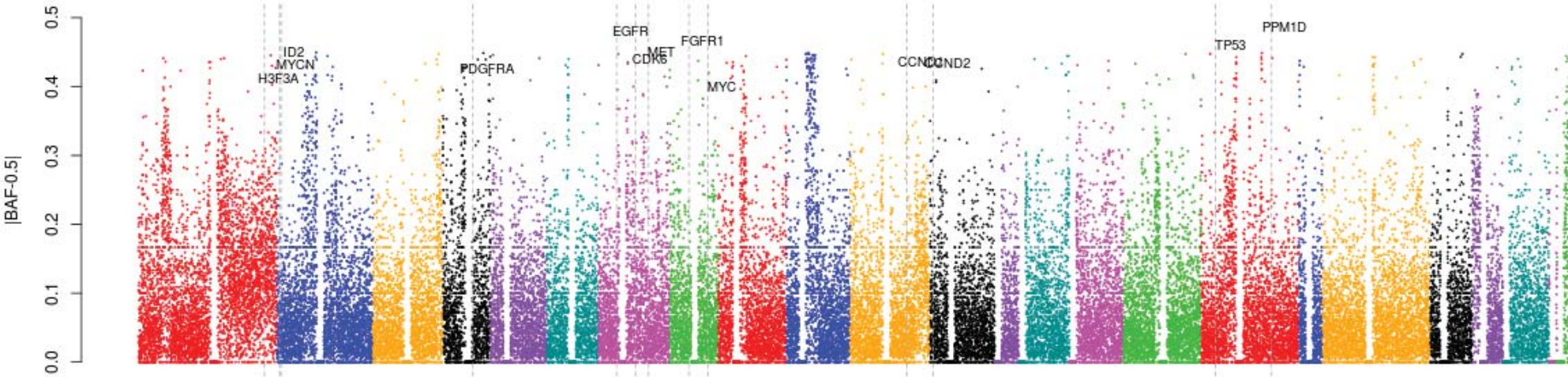
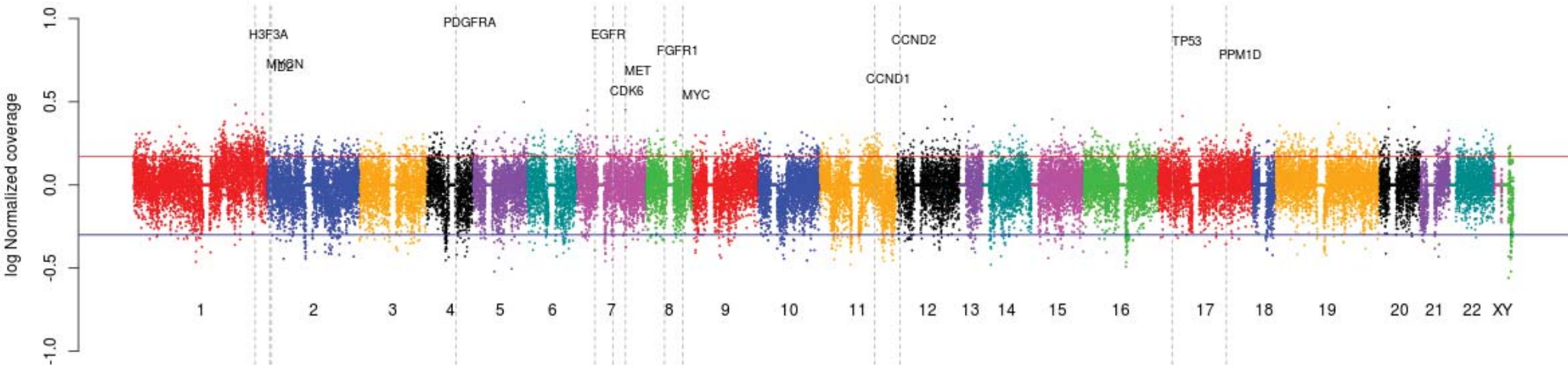
DIPG9-Cerebellum 3



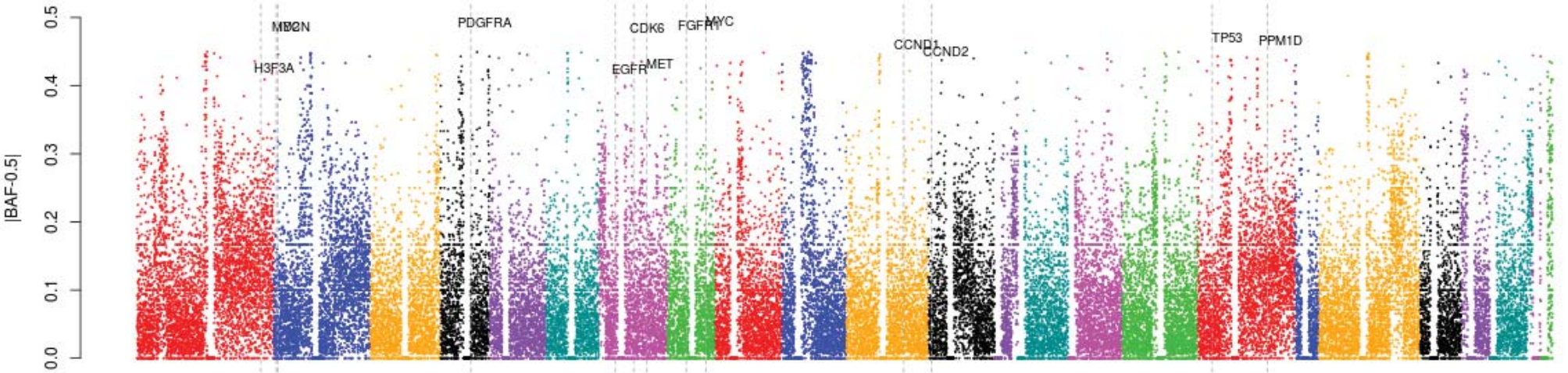
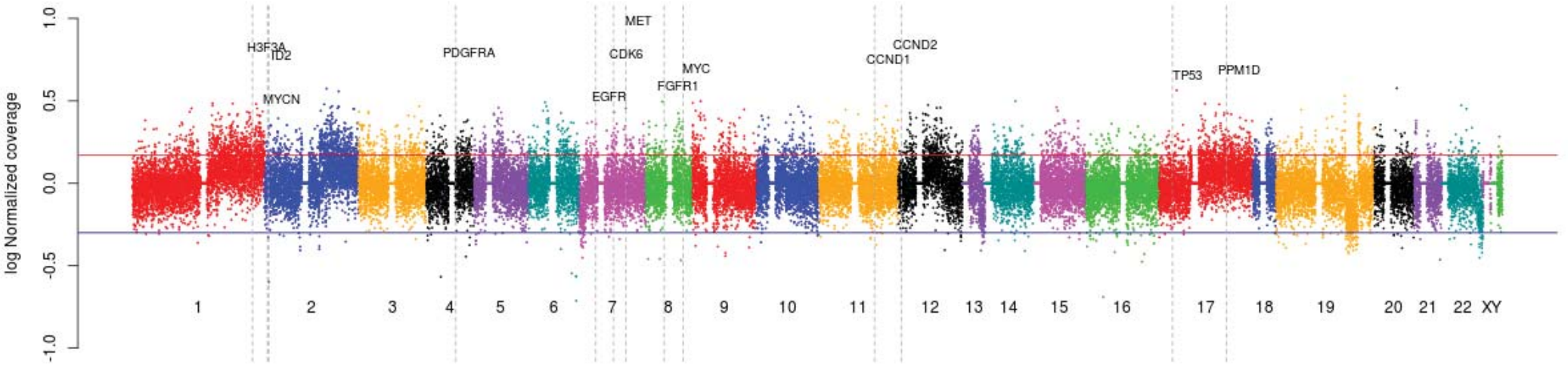
DIPG9-Frontal Lobe 3



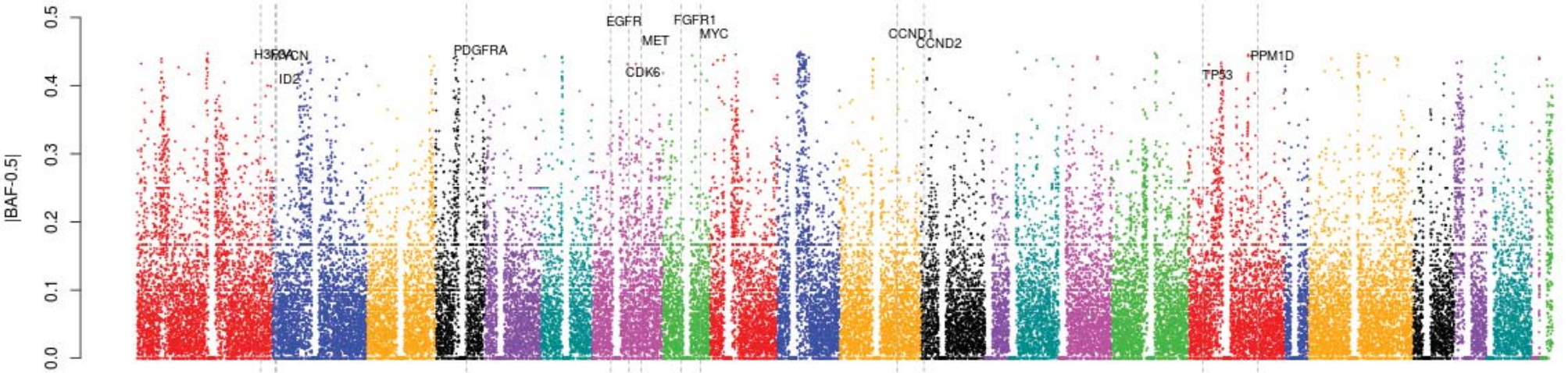
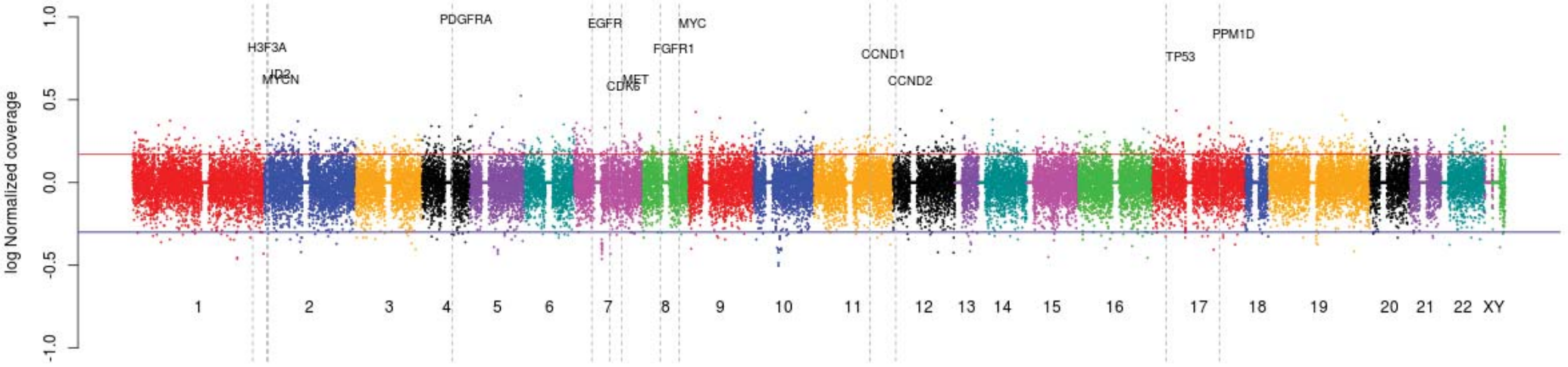
DIPG9-Medulla 1



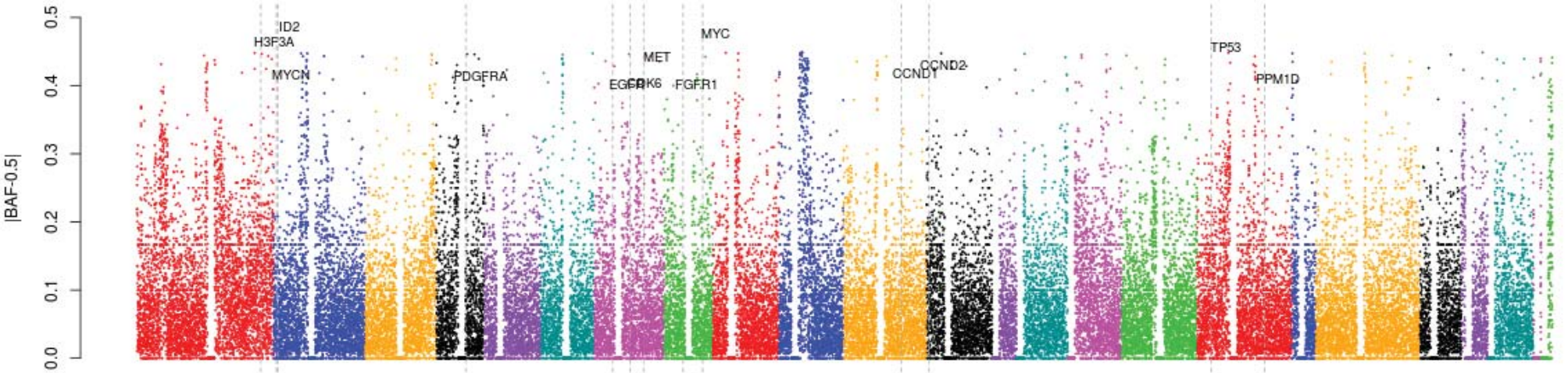
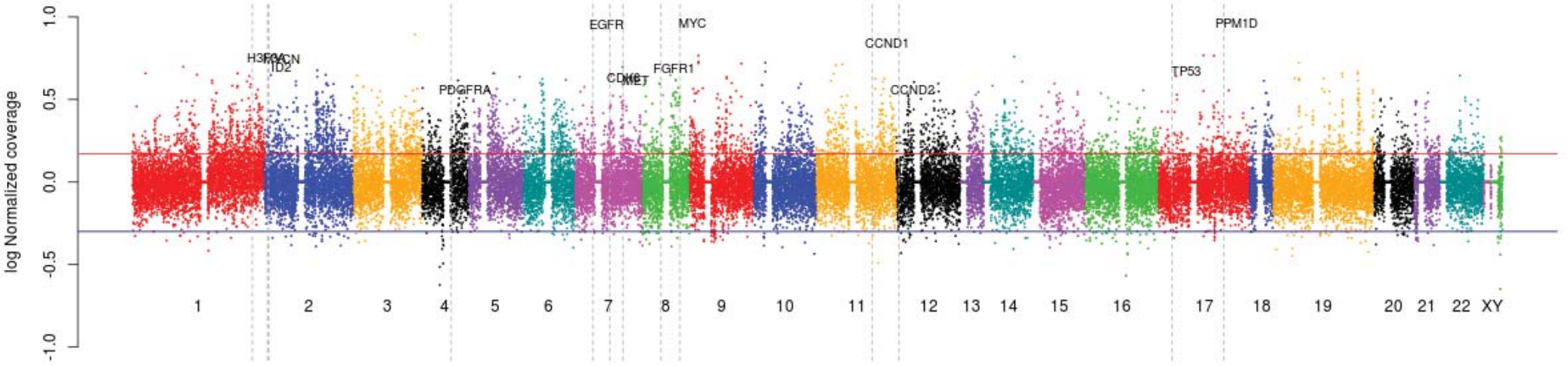
DIPG9-Midbrain



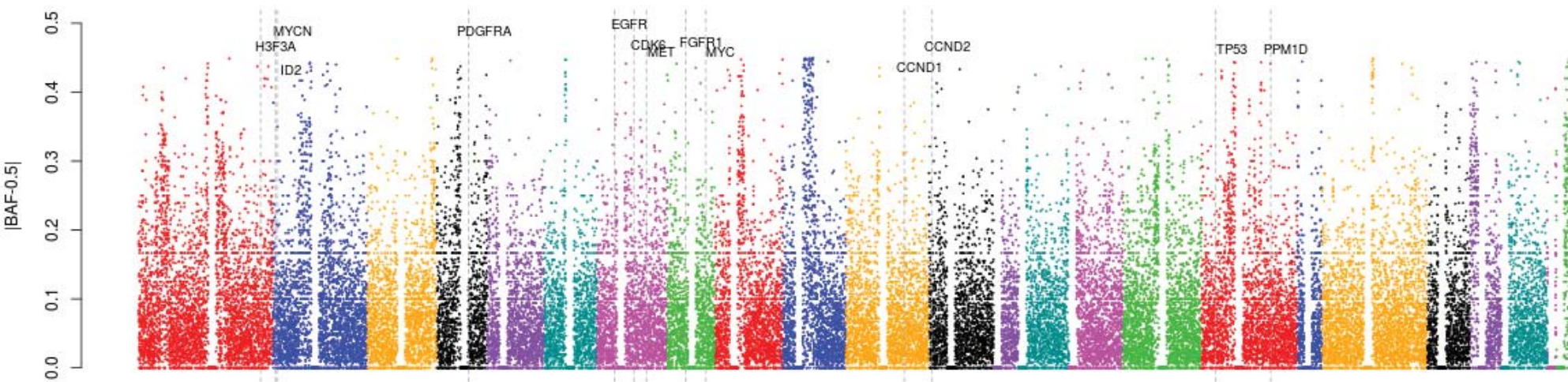
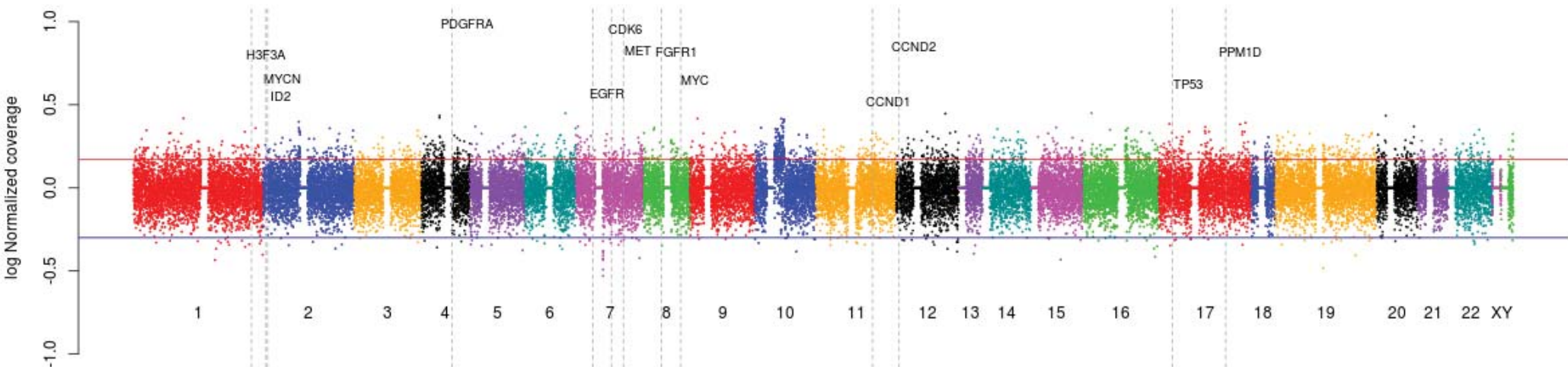
DIPG9-Parietal Lobe 3



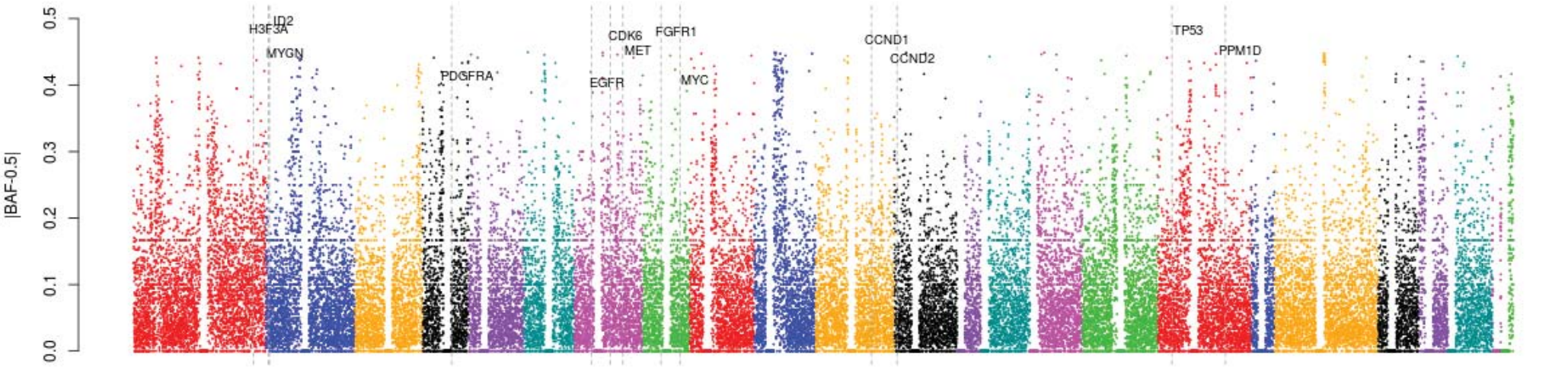
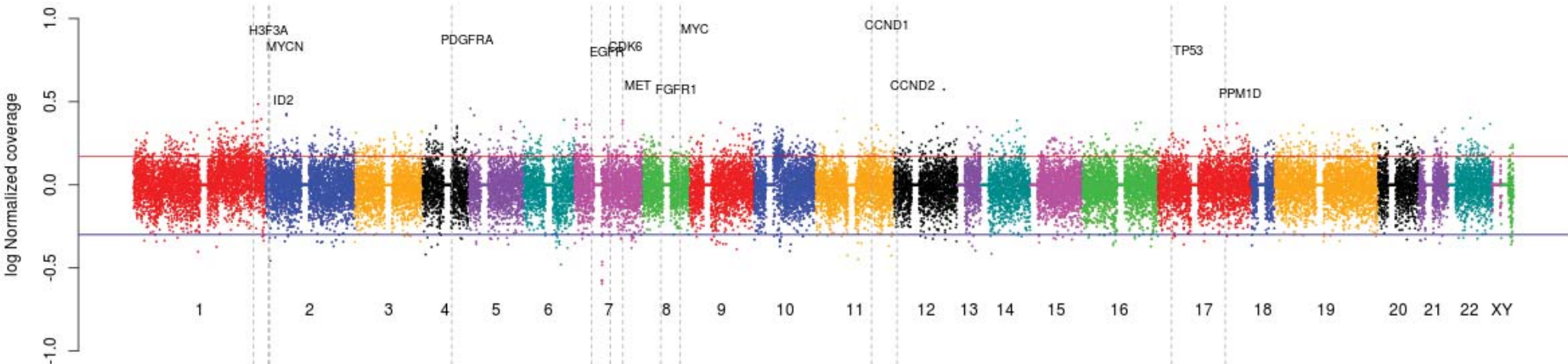
DIPG9-Pons 1



DIPG9-Temporal Lobe 3

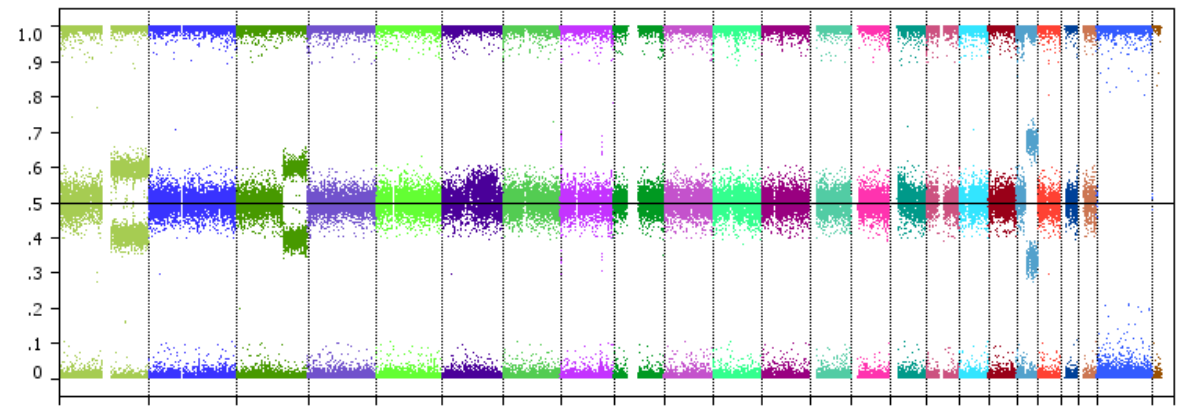
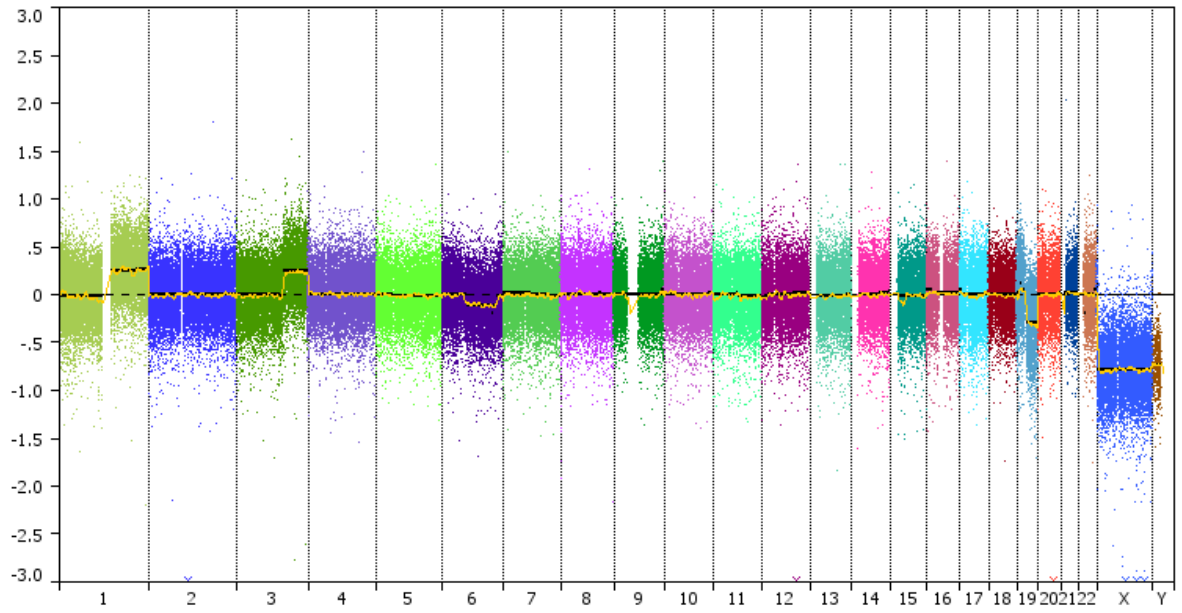


DIPG9-Ventricle

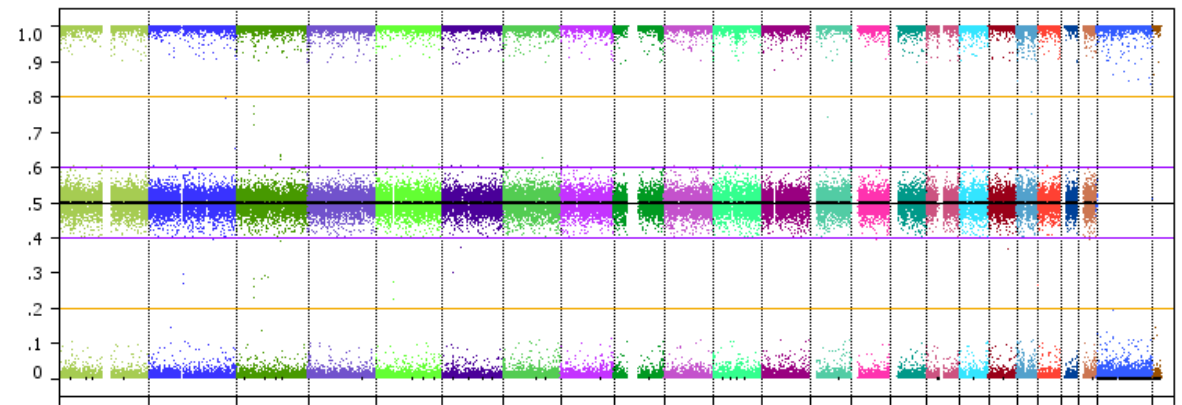
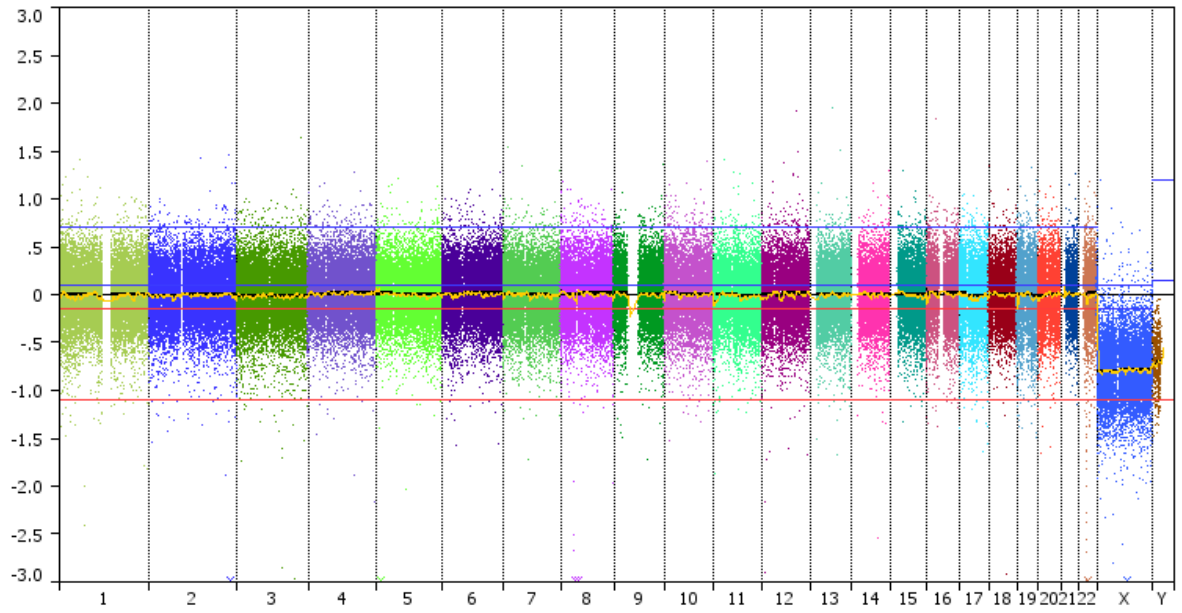


Supplementary Figure 5: CNV calling using whole exome sequencing data. Whole exome sequencing data was used to estimate copy number variation events in DIPG samples using deviation of B allele frequency from 50% as well as normalized coverage data (See online methods).

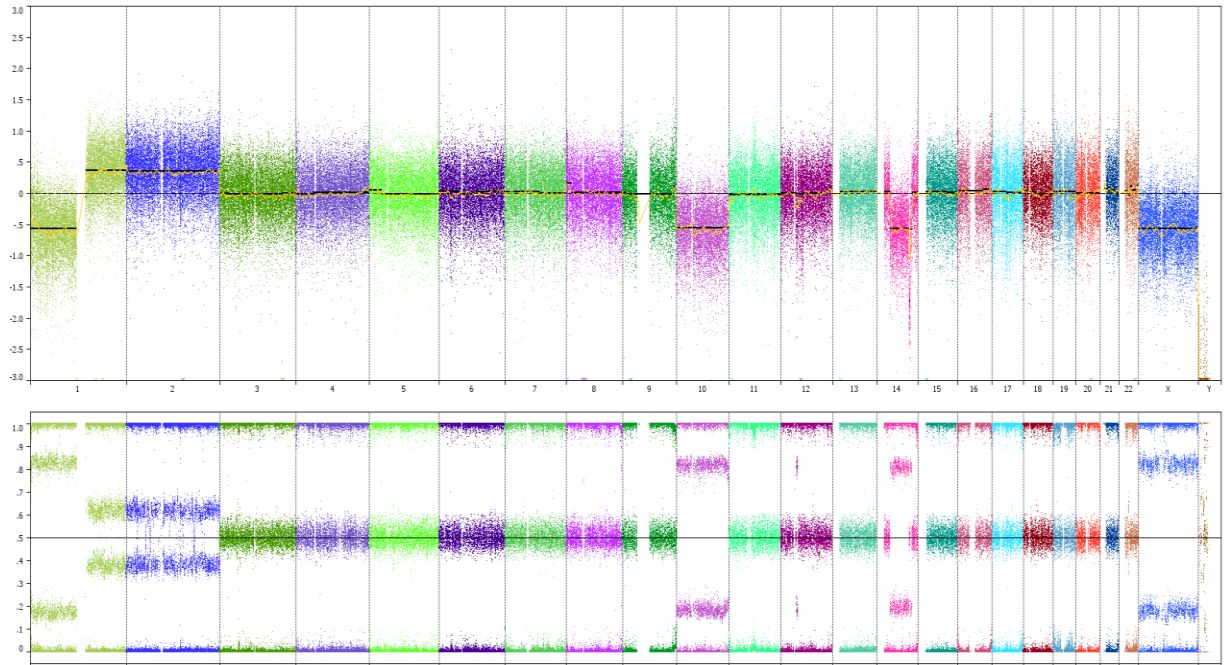
Supplementary Figure 6



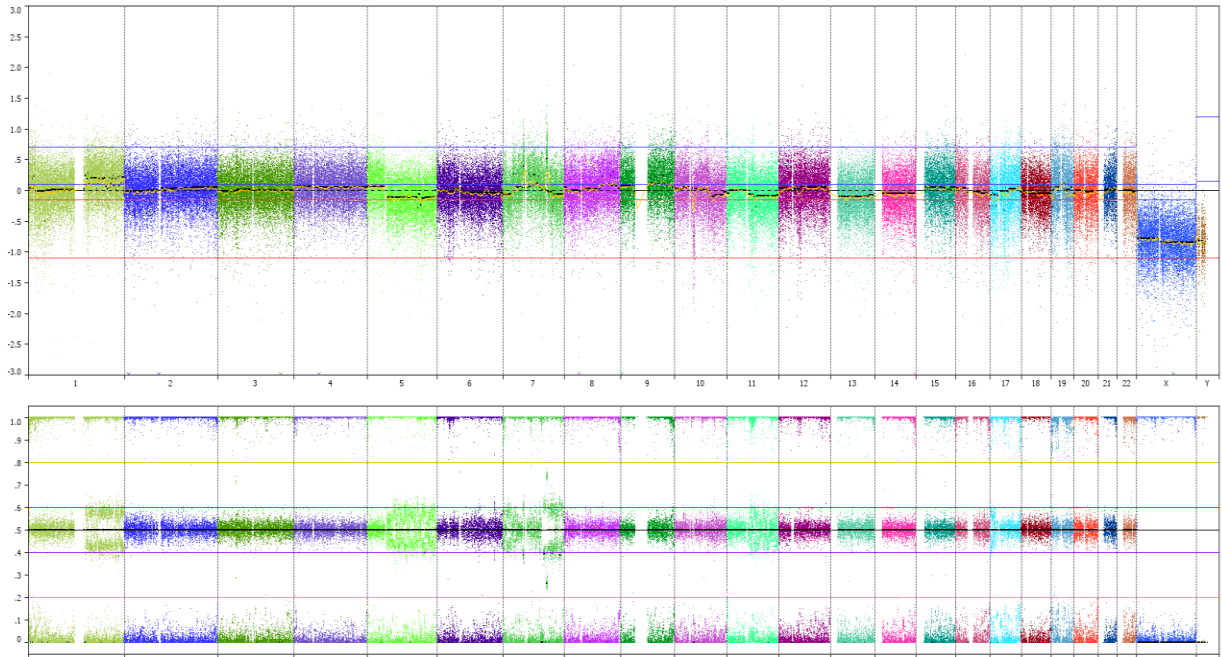
DIPG1-Pons1



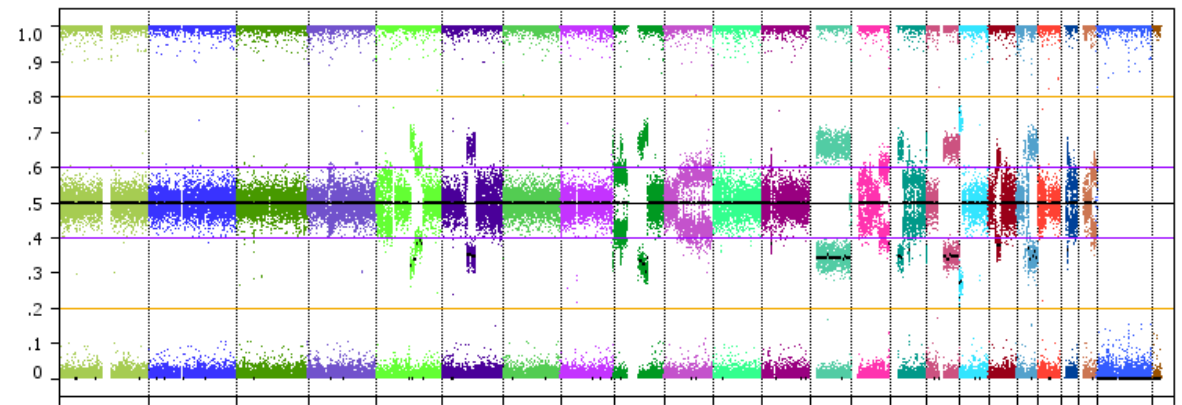
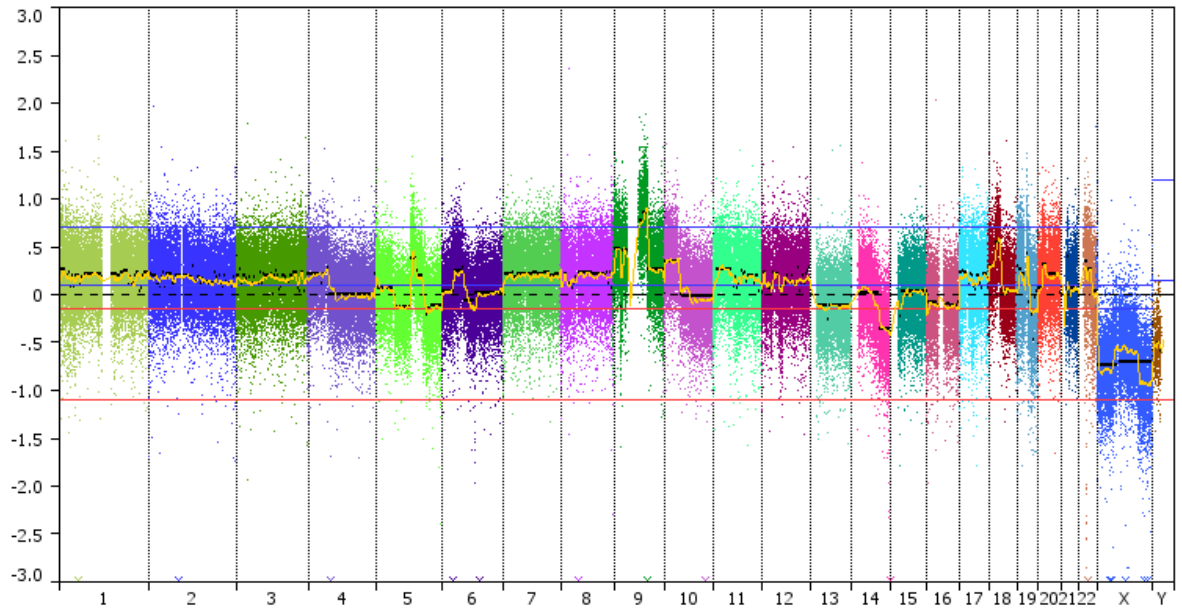
DIPG2-Cerebellum3



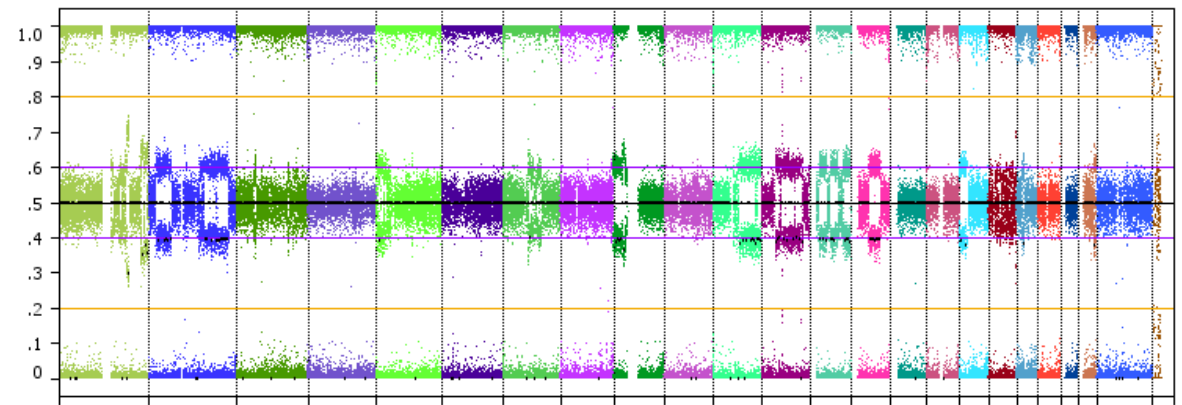
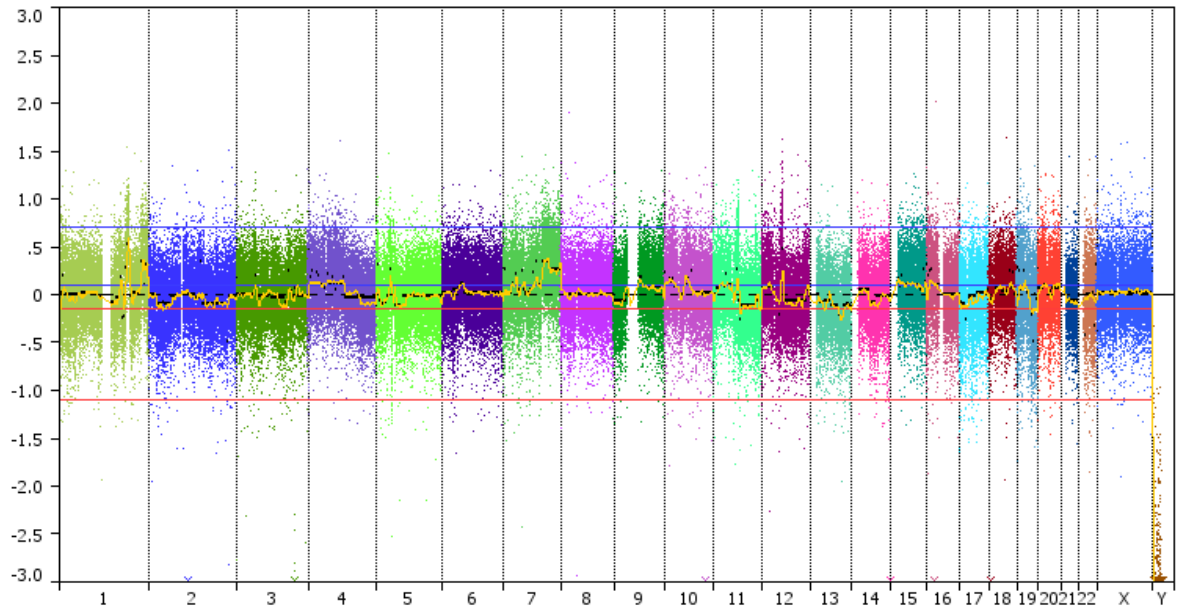
DIPG3-Pons2



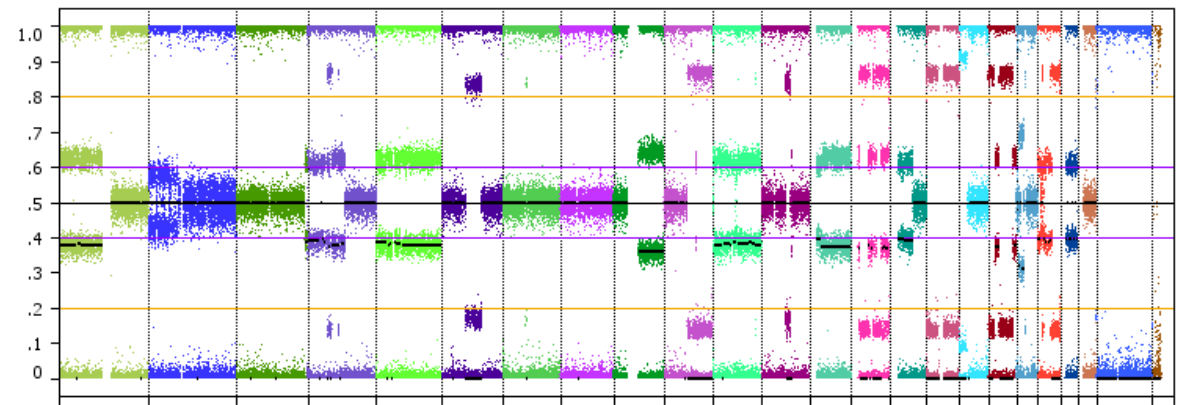
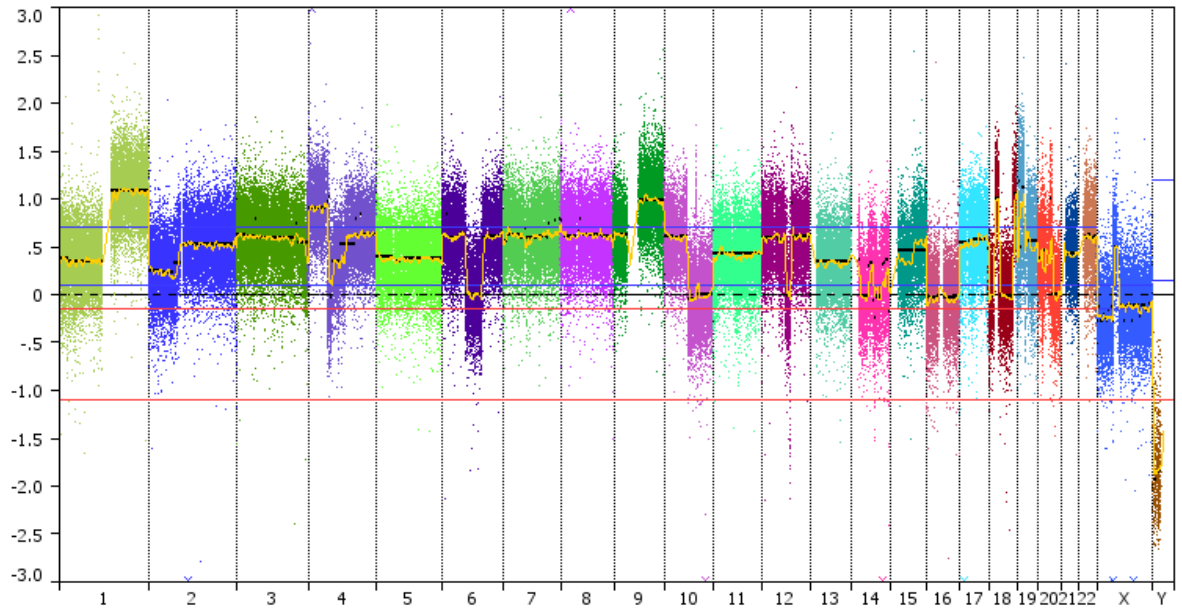
DIPG5-Pons1



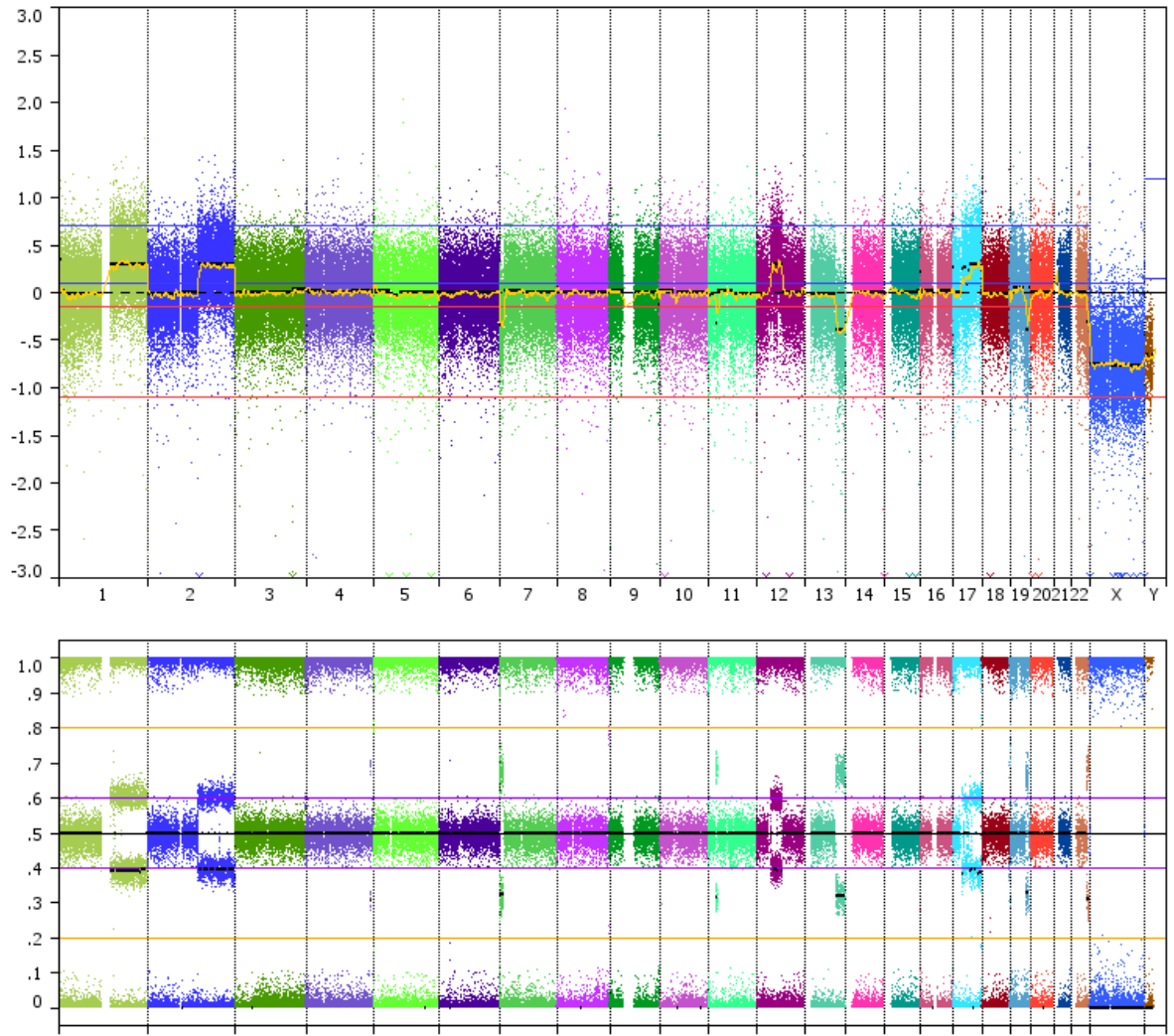
DIPG6-Pons5



DIPG7-Cerebellum1



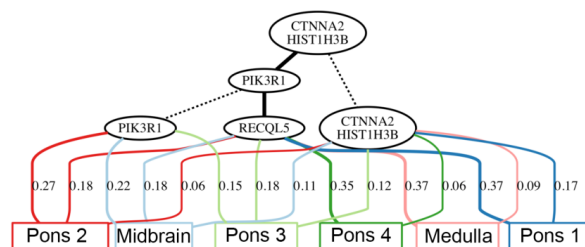
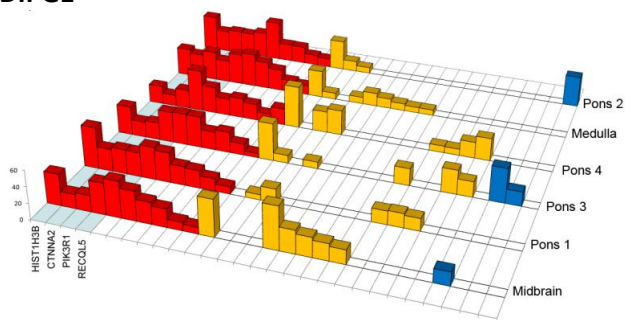
DIPG8-Midbrain



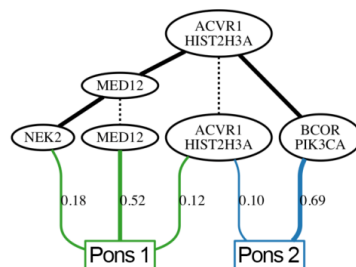
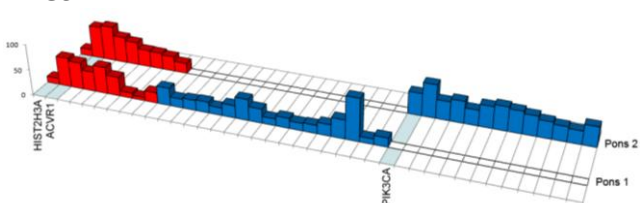
DIPG9-Midbrain

Supplementary Figure 6: OncoScan CNV arrays were used on selected samples to validate whole exome sequencing based CNV calling. Using the same DNA used for whole exome sequencing, one sample from each patient was selected to assess copy number variation events (except for patient DIPG4 for which no material was available). OncoScan results shown here correlated with whole exome based CNV calling.

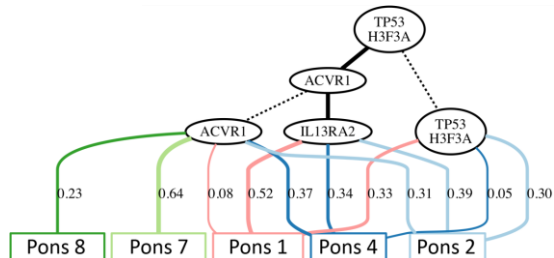
DIPG1



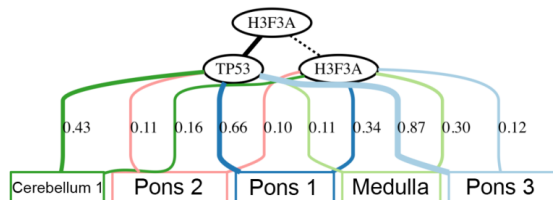
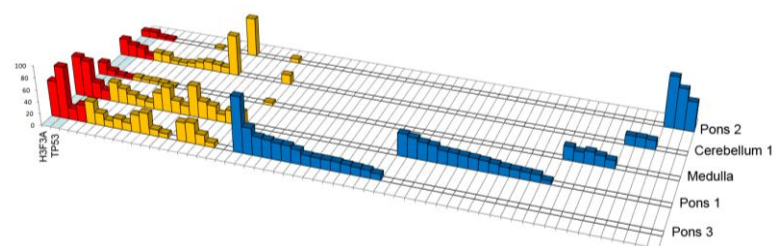
DIPG3



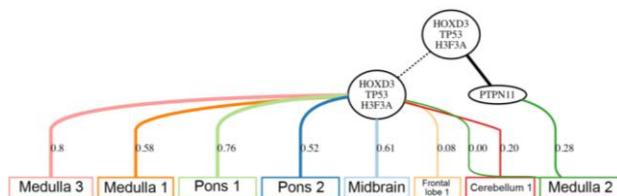
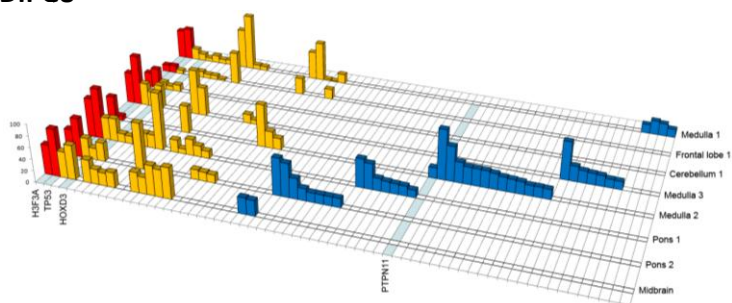
DIPG4*



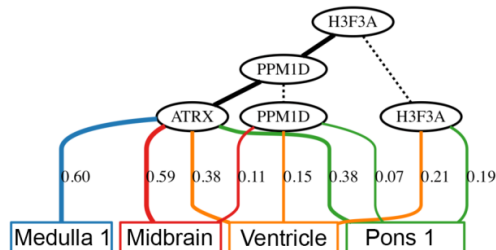
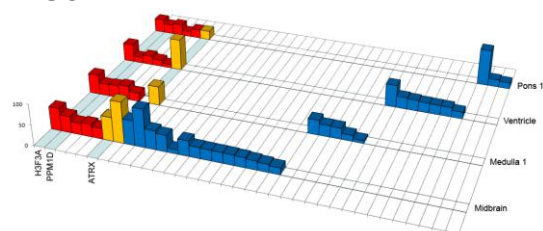
DIPG7



DIPG8



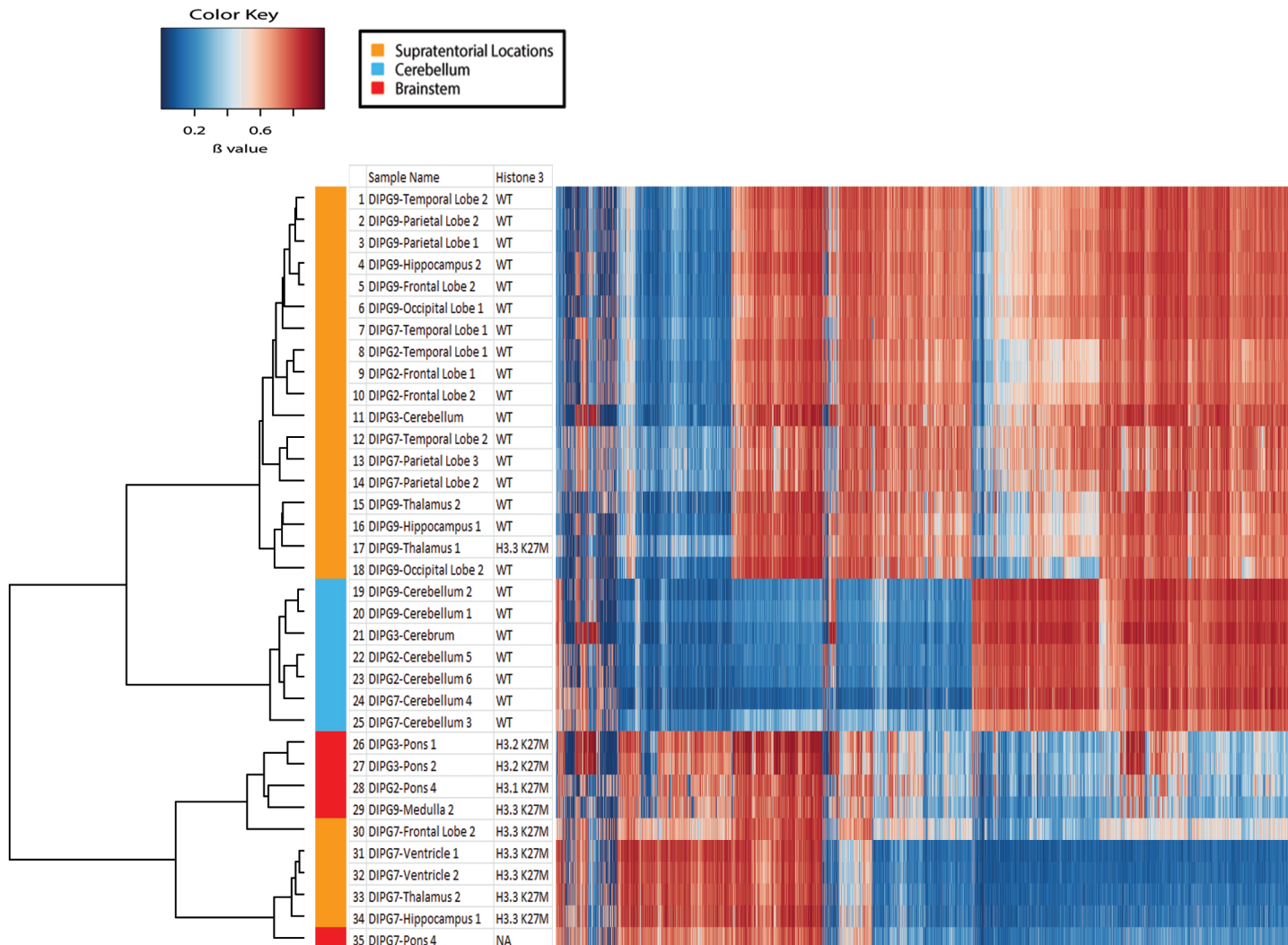
DIPG 9



Supplementary Figure 7: Evolutionary reconstruction and non-synonymous somatic mutations allele frequency values for six patients (not included in Figure 2). Left: Histograms represent the raw allele frequency values from whole exome sequencing data (Supplementary table 2.) Right: Evolutionary trees reconstructed using CNV corrected allele frequencies from the deep amplicon sequencing data targeting the candidate genes found in whole exome sequencing.

*We had access to whole exome sequencing data from only one sample for patient DIPG4. The evolutionary tree for this patient was reconstructed based on deep amplicon sequencing data, except the frequency values were not corrected for CNV events (due to lack of whole exome sequencing data on the rest of the samples from this patient.) Lacking a robust correction for copy number, we consider this phylogeny less reliable and hence do not discuss it in the main text.

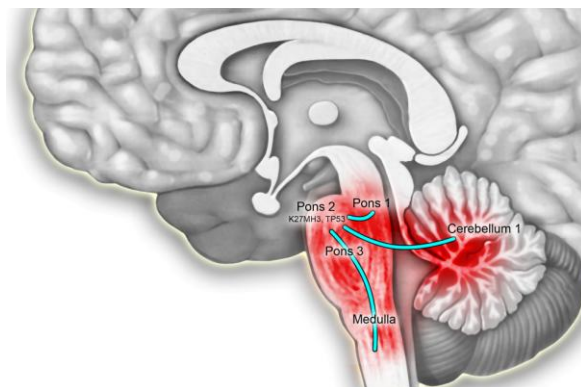
Supplementary Figure 8



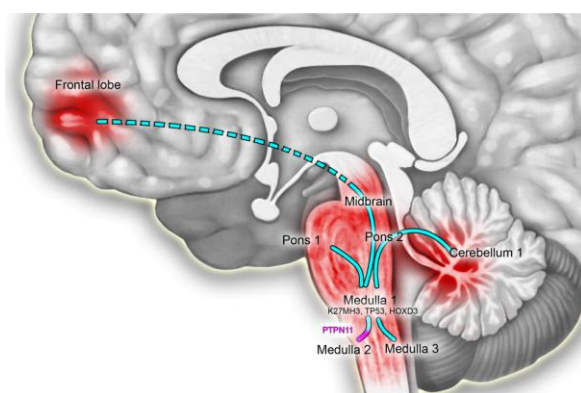
Supplementary Figure 8. Clustering analysis of global DNA methylation from various neuroanatomical locations for four DIPGs. Samples representing histone 3 K27M and wild type neuroanatomical locations for DIPG2 (H3.1-K27M), DIPG3 (H3.2-K27M), DIPG7 and DIPG9 (H3.3-K27M) were analyzed using unsupervised hierarchical clustering for the 5000 most variable probes. Global DNA methylation clustering demonstrates clustering patterns based on histone 3 mutation status, rather than neuroanatomical location. DIPG9 thalamus 1 is a H3-K27M sample clustering with the H3-WT group possible due to the presence of only few tumor cells only detectable by sensitive genomic screening and which based on their low content do not alter the global methylation profile.

Supplementary Figure 9

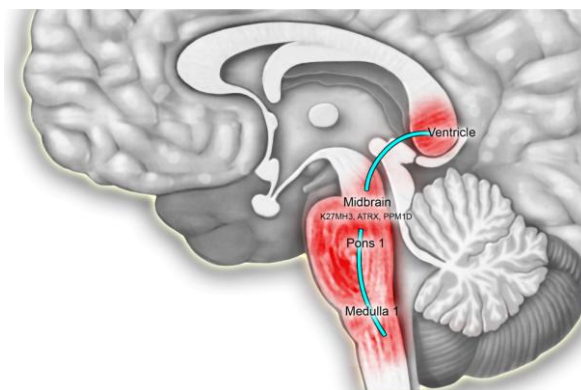
DIPG7



DIPG8

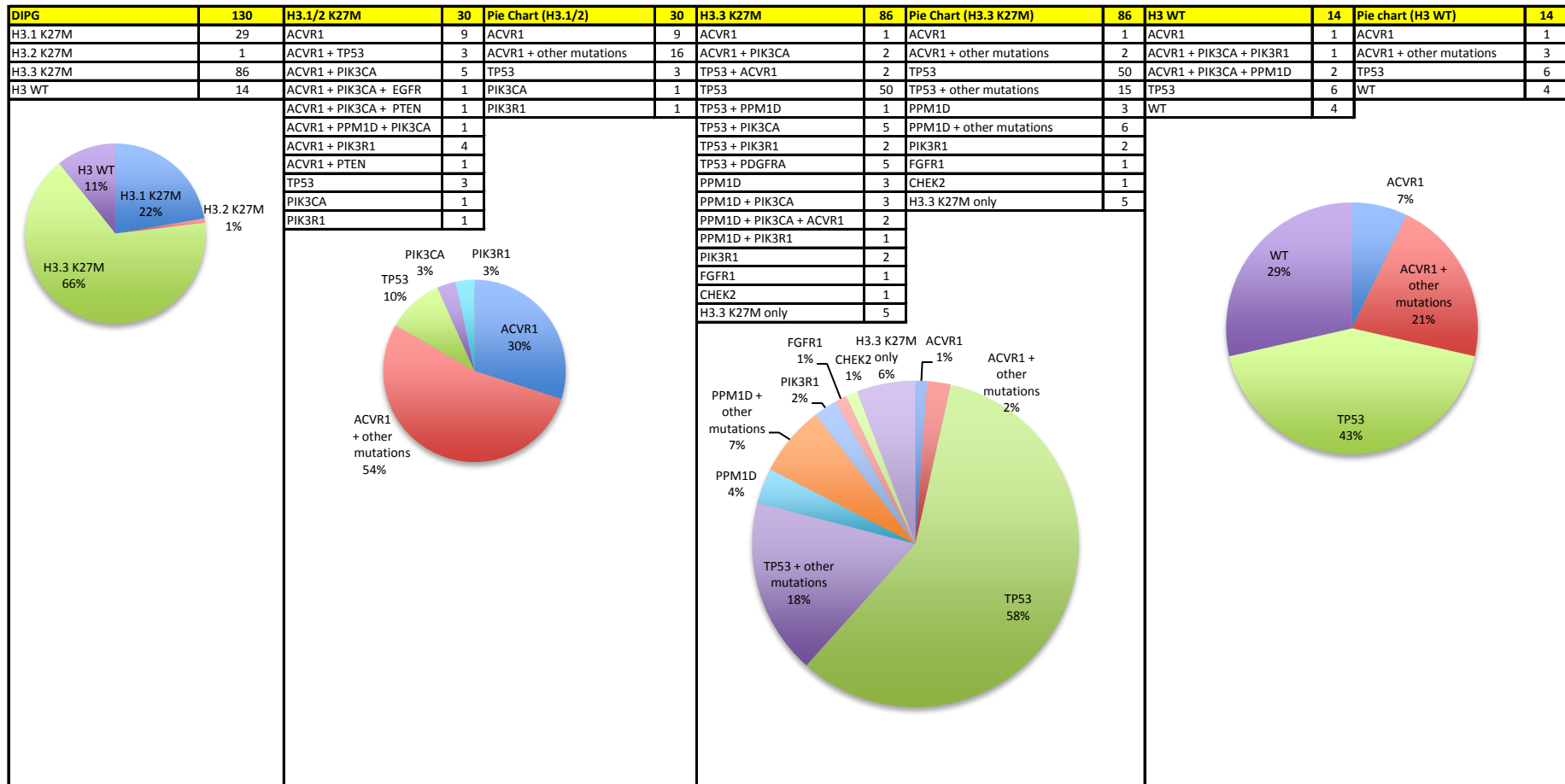


DIPG9



Supplementary Figure 9: Tumor spread outside the brain stem in patients DIPG7, DIPG8, and DIPG9. In DIPG7, spread to cerebellum and across the brain stem is consistent with early event in tumor evolution as it does not harbor additional mutations. In DIPG8, the anatomical spread path towards the frontal lobe is not known and hence is represented as a dashed line (--). Spread to “Medulla 2” in this patient is presumed to happen relatively later than the spread towards the cerebellum since it harbors an additional PTPN1 mutation. The spread towards the medulla and ventricle in DIPG9 is suggested to also happen relatively late in tumor evolution since both samples in these areas share the subclonal ATRX and PPM1D mutations seen in the primary tumor.

Supplementary Figure 10



Supplementary Figure 10: Summary of integrated dataset of Diffuse Intrinsic Pontine Glioma samples from four published studies (n=121) and present work (n=9) showing distribution of histone 3 mutation status as well as mutation combinations of oncohistone partners for all patients. For a detailed list, please refer to Supplementary Table 15.

Supplementary Table 1: Clinical and Molecular Data from 134 Punch Cores Taken From Autopsy Brains of 9 DIPG patients

PATIENT INFORMATION						MOLECULAR ANALYSES											IMMUNOHISTOCHEMISTRY													
Patient ID	Diagnosis	Gender	Age	Treatment		MISEQ RUN 1 (2014)			MISEQ RUN 2 (2015)				Whole Exome Sequencing				Stains				Comments									
				Chemotherapy	Radiation	Sample Name	Neuroanatomical Location	Tissue Type	ddPCR Run1	Histone 3	Other Genes	Histone 3	Other Genes	Methylation	RNA-Seq	ddPCR Run2	Histone 3	Other Genes	H&E	Ki67		H3 K27M	H3 K27me3							
DIPG1 (H3.1)	DIPG (GBM)	Male	21y 2m	SAHA, Avastin, and Irinotecan	Yes	DIPG1-Frontal Lobe 1	B	Frozen			WT						1 core punch	WT	WT	no mutations					cerebral cortex & cerebellum is FFPE					
						DIPG1-Frontal Lobe 2	B	FFPE																√			√	√		
						DIPG1-Cerebellum 1	E	FFPE																		√	√	√	√	
						DIPG1-Cerebellum 2	E	FFPE						WT				H3.1B K27M (4.6%)	failed	failed						√	√	√	√	
						DIPG1-Ventricle 1	F	FFPE																		√	√	√	√	
						DIPG1-Ventricle 2	F	FFPE																		√	√	√	√	
						DIPG1-Hippocampus	G	FFPE																		√	√	√	√	
						DIPG1-Thalamus	H	FFPE																		√	√	√	√	
						DIPG1-Midbrain	I	Frozen						H3.1B K27M (19.32%)	PIK3R1 splicing (24.85%), CTNNA2 A260S (25.33%), RECQL5 splicing (17.24%)		1 Core Punch	H3.1B K27M (25%)	H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
						DIPG1-Pons 1	J	Frozen					H3.1 K27M (52%)	H3.1B K27M (24.4%)	PIK3R1 splicing (23.75%), CTNNA2 A260S (27.39%), RECQL5 splicing (21.99%)		1 Core Punch	H3.1B K27M (34%)	H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
						DIPG1-Pons 2	J	Frozen						H3.1B K27M (29%)	PIK3R1 splicing (29.33%), CTNNA2 A260S (33.04%), RECQL5 splicing (18.02%)				H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
						DIPG1-Pons 3	J	Frozen						H3.1B K27M (23.1%)	PIK3R1 splicing (21.02%), CTNNA2 A260S (22.55%), RECQL5 splicing (16.85%)		1 Core Punch	H3.1B K27M (23%)	H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
						DIPG1-Pons 4	J	Frozen					H3.1 K27M (19%)	H3.1B K27M (18.56%)	PIK3R1 splicing (21.97%), CTNNA2 A260S (20.6%), RECQL5 splicing (21.14%)		1 Core Punch	H3.1B K27M (23%)	H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
						DIPG1-Pons 5	J	FFPE																			√	√	√	√
						DIPG1-Pons 6	J	FFPE																			√	√	√	√
						DIPG1-Medulla	K	Frozen					H3.1 K27M (24%)	H3.1B K27M (25.43%)	PIK3R1 splicing (25.60%), CTNNA2 A260S (24.21%), RECQL5 splicing (18.95%)		1 Core Punch	H3.1B K27M (30%)	H3.1 K27M	CTNNA2 A260S; PIK3R1 splicing										
DIPG2-Temporal Lobe 1	A	FFPE							WT						Does not cluster with K27M group					√	√	√	√							
DIPG2-Temporal Lobe 2	A	FFPE						WT																						
DIPG2-Temporal Lobe 3	A	Frozen								WT				WT	WT	no mutations														
DIPG2-Temporal Lobe 4	A	FFPE						WT																						
DIPG2-Frontal Lobe 1	B	FFPE						WT						Does not cluster with K27M group																
DIPG2-Frontal Lobe 2	B	FFPE						WT						Does not cluster with K27M group						√	√	√	√							
DIPG2-Frontal Lobe 3	B	Frozen						WT		WT				WT	WT	no mutations														
DIPG2-Parietal Lobe 1	C	FFPE						WT																						
DIPG2-Parietal Lobe 2	C	FFPE						WT												√	√	√	√							
DIPG2-Parietal Lobe 3	C	FFPE						WT																						
DIPG2-Parietal Lobe 4	C	Frozen								WT					WT	WT	no mutations													

DIPG2 (H3.1)	DIPG (GBM)	Male	6y 10m	N/A	Yes	DIPG2- Parietal Lobe 5	C	Frozen			WT			WT	WT	no mutations													
						DIPG2- Occipital Lobe 1	D	FFPE	WT																				
						DIPG2- Occipital Lobe 2	D	Frozen			WT							WT	WT	no mutations									
						DIPG2- Cerebellum 1	E	Frozen			H3.1B K27M (15.49%)								H3.1B K27M (16%)	H3.1 K27M	ACVR1 G328V								
						DIPG2- Cerebellum 2	E	Frozen											failed	WT	no mutations								
						DIPG2- Cerebellum 3	E	Frozen											WT	WT	no mutations								
						DIPG2- Cerebellum 4	E	Frozen											WT	WT	no mutations								
						DIPG2- Cerebellum 5	E	FFPE			WT													✓	✓	✓	✓		
						DIPG2- Cerebellum 6	E	FFPE			WT													✓	✓	✓	✓		
						DIPG2- Thalamus	H	Frozen			H3.1 K27M (49%)	ACVR1 G328V (70%)	H3.1B K27M (45.03%)							H3.1B K27M (45%)	H3.1 K27M	ACVR1 G328V; PIK3CA H1047R; MAX R51Q; PTEN A126S							
						DIPG2-Pons 1	J	Frozen			H3.1 K27M (37%)	ACVR1 G328V (57%)	H3.1B K27M (33.44%)							H3.1B K27M (41%)	H3.1 K27M	ACVR1 G328V; MAX R51Q							
						DIPG2-Pons 2	J	Frozen					H3.1B K27M (37.14%)							H3.1B K27M (30%)	H3.1 K27M	ACVR1 G328V; PIK3CA H1047R							
						DIPG2-Pons 3	J	Frozen					H3.1B K27M (33.22%)							1 Core Punch	H3.1B K27M (39%)	H3.1 K27M	ACVR1 G328V; MAX R51Q;						
						DIPG2-Pons 4	J	FFPE			H3.1 K27M (49%)	ACVR1 G328V (70%)													✓	✓	✓	✓	
						DIPG2-Pons 5	J	FFPE																	✓	✓	✓	✓	
						DIPG2-Pons 6	J	FFPE																	✓	✓	✓	✓	
						DIPG2- Medulla	K	Frozen					H3.1B K27M (20.05%)							1 Core Punch	H3.1B K27M (26%)	H3.1 K27M	ACVR1 G328V; MAX R51Q;						
						DIPG3 (H3.2)	DIPG (GBM)	Female	7y 5m	No	Yes	DIPG3- Cerebrum	N/A	Frozen			WT				1 Core Punch		WT	no mutations					
												DIPG3- Cerebellum	E	Frozen			WT								1 Core Punch		WT	no mutations	
DIPG3-Pons 1	J	Frozen			H3.2 K27M (8%)							ACVR1 G328V (47%)							1 Core Punch		H3.2 K27M	ACVR1 G328V							
DIPG3-Pons 2	J	Frozen			H3.2 K27M (5%)	ACVR1 G328V (57%)								1 Core Punch		H3.2 K27M	ACVR1 G328V; PIK3CA H1047R												
DIPG4 (H3.3)	DIPG (GBM)	Female	5y 6m	Capecitabine	Yes	DIPG4- Frontal Lobe 1	B	Frozen			WT						WT	WT	no mutations										
						DIPG4- Frontal Lobe 2	B	FFPE																✓	✓	✓	✓		
						DIPG4- Cerebellum	E	FFPE																	✓	✓	✓	✓	
						DIPG4-Pons 1	J	Frozen					H3.3 K27M (46.47%)							H3.3 K27M (44%)	H3.3 K27M	TP53 C3F; ACVR1 R206H							
						DIPG4-Pons 2	J	Frozen			H3.3 K27M (45%)	ACVR1 R206H (30%)	H3.3 K27M (50.29%)																
DIPG4-Pons 3	J	FFPE																		✓	✓	✓	✓						

cerebral cortex & cerebellum

DIPG7 (H3.3)

DIPG (GBM)

Female

10y 8m

SAHA

Yes

DIPG7- Frontal Lobe 2	B	FFPE		H3.3 K27M (26%)					Clusters with K27M group						√	√	√	√
DIPG7- Parietal Lobe 1	C	Frozen		WT		WT				1 core punch	WT	WT	no mutations					
DIPG7- Parietal Lobe 2	C	FFPE	WT						Does not cluster with K27M group						√	√	√	√
DIPG7- Parietal Lobe 3	C	FFPE	WT						Does not cluster with K27M group						√	√	√	√
DIPG7- Occipital Lobe 1	D	FFPE		WT											√	√	√	√
DIPG7- Occipital Lobe 2	D	FFPE		WT											√	√	√	√
DIPG7- Cerebellum 1	E	Frozen				H3.3 K27M (28.5%)	TP53 R43H (20.12%), PBRM1 splicing (0.35%)				H3.3 K27M (26%)	H3.3 K27M	TP53 R43H					
DIPG7- Cerebellum 2	E	Frozen									WT	WT	no mutations					
DIPG7- Cerebellum 3	E	FFPE		WT					Does not cluster with K27M group						√	√	√	√
DIPG7- Cerebellum 4	E	FFPE				WT			Does not cluster with K27M group						√	√	√	√
DIPG7- Ventricle 1	F	FFPE		H3.3 K27M (64%)					Clusters with K27M group						√	√	√	√
DIPG7- Ventricle 2	F	FFPE		H3.3 K27M (48%)					Clusters with K27M group						√	√	√	√
DIPG7- Hippocampu s 1	G	FFPE		H3.3 K27M (59%)					Clusters with K27M group						√	√	√	√
DIPG7- Hippocampu s 2	G	FFPE		WT											√	√	√	√
DIPG7- Thalamus 1	H	FFPE		WT											√	√	√	√
DIPG7- Thalamus 2	H	FFPE		H3.3 K27M (52%)					Clusters with K27M group						√	√	√	√
DIPG7-Pons 1	J	Frozen		H3.3 K27M (26%)		H3.3 K27M (66.75%)	TP53 R43H (67.55%), PBRM1 splicing (0.23%)				H3.3 K27M (64%)	H3.3 K27M	TP53 R43H					
DIPG7-Pons 2	J	Frozen				H3.3 K27M (9.74%)	TP53 R43H (6.88%), PBRM1 splicing (0.22%)				H3.3 K27M (10%)	H3.3 K27M	TP53 R43H					
DIPG7-Pons 3	J	Frozen				H3.3 K27M (56.87%)	TP53 R43H (87.04%), PBRM1 splicing (0.51%)			2 Core Punches	H3.3 K27M (58%)	H3.3 K27M	TP53 R43H					
DIPG7-Pons 4	J	FFPE		N/A					Clusters with K27M group						√	√	√	√
DIPG7-Pons 5	J	Frozen				WT				2 Core Punches	WT	WT	no mutations					
DIPG7-Pons 6	J	FFPE													√	√	√	√
DIPG7- Medulla	K	Frozen				H3.3 K27M (20.87%)	TP53 R43H (6.57%), PBRM1 splicing (0.36%)				H3.3 K27M (21%)	H3.3 K27M	TP53 R43H					
DIPG8- Temporal Lobe 1	A	FFPE	WT															
DIPG8- Temporal Lobe 2	A	FFPE	WT															
DIPG8- Frontal Lobe 1	B	Frozen		WT		H3.3 K27M (5.11%)	TP53 E162X (11.6%), HOXD3 H431Q (4.48%), PTPN11 A72V (0.09%)				H3.3 K27M (9%)	H3.3 K27M	TP53 E162X					
DIPG8- Frontal Lobe 2	B	FFPE	WT												√	√	√	√
DIPG8- Frontal Lobe 3	B	FFPE	WT												√	√	√	√
DIPG8- Parietal Lobe 1	C	FFPE	WT															
DIPG8- Parietal Lobe 2	C	Frozen				WT					WT	WT	no mutations					
DIPG8- Occipital Lobe	D	Frozen				WT					H3.3 K27M (21%)	WT	TP53 (3.13%)					

cerebral
cortex &
cerebellum
is FFPE

DIPG8 (H3.3)

DIPG (GBM)

Male

5y

Temozolomide, Avastin, SAHA

Yes

DIPG8-Cerebellum 1	E	Frozen				H3.3 K27M (17.96%)	TP53 E162X (22.17%), HOXD3 H431Q (4.27%), PTPN11 A72V (0.08%)			H3.3 K27M (24%)	H3.3 K27M	TP53 E162X					
DIPG8-Cerebellum 2	E	Frozen				WT				WT	WT	TP53 (2.03%)					
DIPG8-Cerebellum 3	E	FFPE	H3.3 K27M (6%)										✓	✓	✓	✓	
DIPG8-Cerebellum 4	E	FFPE	H3.3 K27M (8%)										✓	✓	✓	✓	
DIPG8-Ventricle	F	Frozen				WT				WT	WT	no mutations					
DIPG8-Hippocampus 1	G	FFPE	WT			WT				WT	failed	failed					
DIPG8-Hippocampus 2	G	FFPE	WT														
DIPG8-Thalamus 1	H	FFPE	WT										✓	✓	✓	✓	
DIPG8-Thalamus 2	H	FFPE	H3.3 K27M (4%)										✓	✓	✓	✓	
DIPG8-Midbrain	I	Frozen				H3.3 K27M (41.66%)	TP53 E162X (77.78%), HOXD3 H431Q (46.62%), PTPN11 A72V (0.09%)			H3.3 K27M (16%)	H3.3 K27M	TP53 E162X					
DIPG8-Pons 1	J	Frozen				H3.3 K27M (62.5%)	TP53 E162X (91.58%), HOXD3 H431Q (49.78%), PTPN11 A72V (0.01%)			H3.3 K27M (71%)	H3.3 K27M	TP53 E162X					
DIPG8-Pons 2	J	Frozen				H3.3 K27M (44.42%)	TP53 E162X (69.52%), HOXD3 H431Q (25.18%), PTPN11 A72V (0.14%)			H3.3 K27M (21%)	H3.3 K27M	TP53 E162X					
DIPG8-Pons 3	J	FFPE	H3.3 K27M (12%)										✓	✓	✓	✓	
DIPG8-Pons 4	J	FFPE	H3.3 K27M (17%)										✓	✓	✓	✓	
DIPG8-Pons 5	J	FFPE	H3.3 K27M (13%)										✓	✓	✓	✓	
DIPG8-Medulla 1	K	Frozen		H3.3 K27M (48%)		H3.3 K27M (47.33%)	TP53 E162X (58.07%), HOXD3 H431Q (28.14%), PTPN11 A72V (0.33%)			H3.3 K27M (48%)	H3.3 K27M	TP53 E162X					
DIPG8-Medulla 2	K	Frozen				H3.3 K27M (30.03%)	TP53 E162X (10.27%), HOXD3 H431Q (7.25%), PTPN11 A72V (17.68%)			H3.3 K27M (34%)	H3.3 K27M	TP53 E162X					
DIPG8-Medulla 3	K	Frozen				H3.3 K27M (55.43%)	TP53 E162X (85.62%), HOXD3 H431Q (39.88%), PTPN11 A72V (0.15%)			H3.3 K27M (57%)	H3.3 K27M	TP53 E162X					
DIPG9-Temporal Lobe 1	A	FFPE	WT														
DIPG9-Temporal Lobe 2	A	FFPE	WT					Does not cluster with K27M group									
DIPG9-Temporal Lobe 3	A	Frozen								WT	WT	no mutations					
DIPG9-Frontal Lobe 1	B	FFPE	WT										✓	✓	✓	✓	
DIPG9-Frontal Lobe 2	B	FFPE	WT					Does not cluster with K27M group									
DIPG9-Frontal Lobe 3	B	Frozen		WT		WT				WT	WT	no mutations					
DIPG9-Parietal Lobe 1	C	FFPE	WT					Does not cluster with K27M group									
DIPG9-Parietal Lobe 2	C	FFPE	WT					Does not cluster with K27M group									
DIPG9-Parietal Lobe 3	C	Frozen				WT				WT	WT	no mutations					
DIPG9-Occipital Lobe 1	D	FFPE	WT					Does not cluster with K27M group									
DIPG9-Occipital Lobe 2	D	FFPE	WT					Does not cluster with K27M group									

ABT888
(Naloxone)

DIPG9 (H3.3)	DIPG (GBM)	Male	9y 3m	1 venperay Temozolomi de in maintenanc e phase	Yes	DIPG9-Occipital Lobe 3	D	Frozen				WT			WT	WT	no mutations						
						DIPG9-Cerebellum 1	E	FFPE		WT				Does not cluster with K27M group					√	√	√	√	
						DIPG9-Cerebellum 2	E	FFPE		WT				Does not cluster with K27M group					√	√	√	√	
						DIPG9-Cerebellum 3	E	Frozen				WT			1 Core Punch	failed	WT	no mutations					
						DIPG9-Ventricle	F	Frozen			H3.3 K27M (36.49%)		ATRX V1514D (37.84%), PPM1D E525X (24.68%)			H3.3 K27M (32%)	H3.3 K27M	PPM1D E525X; ATRX V1514D					
						DIPG9-Hippocampus 1	G	FFPE	WT					Does not cluster with K27M group									
						DIPG9-Hippocampus 2	G	FFPE	WT					Does not cluster with K27M group									
						DIPG9-Thalamus 1	H	FFPE	H3.3 K27M (11.5%)					Does not cluster with K27M group						√	√	√	√
						DIPG9-Thalamus 2	H	FFPE	WT					Does not cluster with K27M group									
						DIPG9-Midbrain	I	Frozen			H3.3 K27M (48.57%)		ATRX V1514D (62.76%), PPM1D E525X (48.87%)		1 Core Punch	H3.3 K27M (49%)	H3.3 K27M	PPM1D E525X; ATRX V1514D					
						DIPG9-Pons 1	J	Frozen		H3.3 K27M (26%)	H3.3 K27M (31.42%)		ATRX V1514D (39.03%), PPM1D E525X (22.06%)			H3.3 K27M (24%)	H3.3 K27M	PPM1D E525X; ATRX V1514D					
						DIPG9-Pons 2	J	FFPE												√	√	√	√
						DIPG9-Medulla 1	K	Frozen			H3.3 K27M (46.6%)		ATRX V1514D (64.19%), PPM1D E525X (30.9%)			H3.3 K27M (4%)	H3.3 K27M	PPM1D E525X					
						DIPG9-Medulla 2	K	FFPE	H3.3 K27M (59%)					Clusters with K27M group						√	√	√	√

Supplementary Table 1: Clinical and molecular data from 134 punch cores taken from autopsy brains of 9 Diffuse Intrinsic Pontine Glioma (DIPG) patients analyzed in this study. Neuroanatomical locations: A = temporal lobe, B = frontal lobe, C = parietal lobe, D = occipital lobe, E = cerebellum, F = lateral ventricles, G = hippocampus, H = thalamus, I = midbrain, J = pons, K = medulla

Supplementary Table 2: Sample distribution from 9 DIPG patient samples used in this study

Patient ID	Total # of samples analyzed (IHC and molecular)	MOLECULAR ANALYSES #			IHC only #	IHC and molecular	
		Total #	Brainstem #	Non brainstem #		Tumor #	Normal #
DIPG1	16	8	6	2	8	13	3
DIPG2	28	26	5	21	2	9	19
DIPG3	4	4	2	2	0	2	2
DIPG4	13	6	5	1	7	8	5
DIPG5	9	6	4	2	3	6	3
DIPG6	11	9	7	2	2	8	3
DIPG7	26	25	6	19	1	12	14
DIPG8	26	26	9	17	0	16	10
DIPG9	25	24	4	20	1	7	18
TOTAL	158*	134**	48	86	24***	81	77
AVERAGE	18	15	5	10	3	9	9

Supplementary Table 2: Sample distribution and neuroanatomical locations from 9 DIPG patients (DIPG1-DIPG9) for all methods used in this study. *Number includes all samples used in this study. **Number excludes samples with only IHC data. ***Number of samples with only IHC data.

Supplementary Table 3: Tumor extension assessed by molecular analysis

Patient ID	Pons	Medulla	Midbrain	Cerebellum	Thalamus	Lateral Ventricles	Hippo-campus	Frontal Lobe	Occipital Lobe	Other Cerebral location
DIPG1	(4/4)	(1/1)	(1/1)	(1/1)				(0/1)		
DIPG2	(4/4)	(1/1)		(1/6)	(1/1)			(0/3)	(0/2)	
DIPG3	(2/2)			(0/1)					(0/1)	(0/1)
DIPG4	(5/5)							(0/1)		
DIPG5	(1/1)	(1/1)	(1/2)	(0/2)						
DIPG6	(5/6)		(1/1)	(0/1)				(0/1)		
DIPG7	(3/4)	(1/1)		(1/4)	(1/2)	(2/2)	(1/2)	(1/2)		
DIPG8	(5/5)	(3/3)	(1/1)	(3/4)	(1/2)	(0/1)	(0/2)	(1/3)	(1/1)	
DIPG9	(1/1)	(2/2)	(1/1)	(0/3)	(1/2)	(1/1)	(0/2)	(0/3)	(0/3)	
# patients with tumor extension	9	6	5	4	4	2	1	2	1	
Total # patients analyzed	9	6	5	8	4	3	3	7	4	
Percentage	100%	100%	100%	50%	100%	67%	33%	29%	25%	
# samples with tumor extension	31	9	5	6	4	3	1	2	1	
Total # samples analyzed	33	9	6	22	6	4	5	13	6	
Percentage	94%	100%	83%	27%	67%	75%	20%	15%	17%	

Supplementary Table 3: Tumor extension in different neuroanatomical locations tested by different molecular analyses (WES, MiSeq, ddPCR) for 9 Diffuse Intrinsic Pontine Glioma brain samples (DIPG1-DIPG9) analyzed in this study. Numbers in parentheses:(number of samples with H3-K27M/number of samples analyzed by molecular methods). **Cells in bold indicate samples with extension detected by molecular analysis.**

Supplementary Table 4: Whole Exome Sequencing data for 67 samples (9 Patients; DIPG1-DIPG9)

	K27M Marker	Accompanying (potential) driver mutations			
DIPG1	HIST1H3B_p.K27M	PIK3R1_splicing	CTNNA2_p.A260S		
DIPG2	HIST1H3B_p.K27M	ACVR1_p.G328V	PIK3CA_p.H1047R	MAX_p.R51Q	PTEN_p.A126S
DIPG3	HIST2H3C_p.K27M	ACVR1_p.G328V	PIK3CA_p.H1047R		
DIPG4	H3F3A_p.K27M	ACVR1_p.R206H	TP53p.C3F		
DIPG5	H3F3A_p.K27M	TP53p.A175fs	PPM1D_p.W427X	ATRXp.D94fs	
DIPG6	H3F3A_p.K27M	TP53_p.G113D	PIK3CA_p.H1047R	ATRXsplicing	OLIG2p.P215fs
DIPG7	H3F3A_p.K27M	TP53_p.R43H			
DIPG8	H3F3A_p.K27M	TP53_p.E162X	TP53p.P58fs		
DIPG9	H3F3A_p.K27M	PPM1D_p.E525X	ATRX_p.V1514D		

Supplementary Table 4: Summary of Whole Exome Sequencing data for 67 samples from 9 Diffuse Intrinsic Pontine Gliomas patients (DIPG1-DIPG9) demonstrating the main driver mutations as well as the accessory mutations.

Supplementary Table 5 Whole Exome Sequencing-based CNV calling

	Patient ID		H3.1												H3.2			H3.3																								
			DIPG 1				DIPG 2				DIPG 3	DIPG 4	DIPG 5	DIPG 6			DIPG 7			DIPG 8			DIPG 9																			
	Age	Gender	21y 2m		10m				7y 5m	4y 6m	5y	10y 8m			5y			3y 3m																								
	Male	Female	Male	Male	Female	Female	Female	Male	Female	Male	Male	Female	Female	Female	Male	Male	Female	Male	Male	Male																						
		Midbrain	Pons 1	Pons 2	Pons 3	Pons 4	Medulla	Frontal Lobe 1	Pons 1	Pons 2	Pons 3	Medulla	Cerebellum 1	Cerebellum 2	Cerebellum 3	Cerebellum 4	Cerebellum 5	Occipital Lobe 1	Occipital Lobe 2	Occipital Lobe 3	Thalamus	Frontal Lobe 3	Pons 2	Pons 1	Pons 2	Cerebellum	Cerebellum	Frontal Lobe 1	Pons 1	Medulla	Midbrain	Midbrain	Midbrain	Medulla	Cerebellum 3	Occipital Lobe 3	Occipital Lobe 3	Occipital Lobe 3	Frontal Lobe 3	Venticle	Pons 1	
CNV (Complete Arm Gain/Loss)	1p																																									
	1q																																									
	2p																																									
	2q																																									
	3p																																									
	4p																																									
	4q																																									
	5p																																									
	5q																																									
	6p																																									
	6q																																									
	9p																																									
	9q																																									
	10p																																									
	10q																																									
	11q																																									
	12p																																									
	12q																																									
	13q																																									
	14q																																									
	15q																																									
16p																																										
16q																																										
17p																																										
17q																																										
18p																																										
18q																																										
19p																																										
19q																																										
20p																																										
20q																																										
21p																																										
21q																																										
CCND1																																										
CCND2																																										
MYC																																										
MYCN																																										
TP53																																										
EGFR																																										
SOX2																																										
PDGFRA																																										
CDK6																																										
FGFR1																																										

CNV
■ Complete arm loss
■ Complete arm gain
■ Copy Neutral LOH

Supplementary Table 5: Copy Number Variation (CNV) calling based on Whole Exome Sequencing data for 67 samples from 9 Diffuse Intrinsic Pontine Glioma patients (DIPG1-DIPG9).

Supplementary Table 6: Integrated dataset of DIPG samples from four published studies (n=121) and present work (n=9)

Sample	Tumor Grade	Age	Gender	H3 mutation	ACVR1	TP53	PPM1D	PIK3CA	PIK3R1	PTEN	EGFR	FGFR1	PDGFRA	Other mutations	References
SJHGG106_A	IV	3.72	F	H3.1 K27M	R258G			H1047R							Wu et al. 2014
SJHGG074_D	IV	4.63	F	H3.1 K27M	R258G			I102_V105del							Wu et al. 2014
SJHGG056_A	IV	5.16	F	H3.1 K27M	R258G				L380del						Wu et al. 2014
SJHGG118_A	IV	3.40	M	H3.1 K27M	G328V				K567_I571>						Wu et al. 2014
SJHGG008_A	IV	3.66	F	H3.1 K27M	G356D	R273C									Wu et al. 2014
SJHGG070_A	N/A	4.42	F	H3.1 K27M	G328E	G245C									Wu et al. 2014
SJHGG069_A	III	8.40	F	H3.1 K27M	G328E								Amp (Microarray CNV)		Wu et al. 2014
SJHGG047_A	IV	4.27	M	H3.1 K27M	G328E										Wu et al. 2014
SJHGG065_A	IV	6.58	M	H3.1 K27M	R206H										Wu et al. 2014
SJHGG077_A	IV	4.20	F	H3.1 K27M	R206H										Wu et al. 2014
SJHGG079_A	N/A	5.68	M	H3.1 K27M	R258G										Wu et al. 2014
SJHGG117_A	IV	2.58	F	H3.1 K27M	G328E			H1047R			R108K				Wu et al. 2014
SJHGG061_A	IV	6.90	F	H3.3 K27M	G328V			P449_L456>L					Amp (Microarray CNV)	ATRX gC335>FR	Wu et al. 2014
SJHGG058_A	IV	5.20	F	H3.3 K27M	G356D	E180fs									Wu et al. 2014
SJHGG002_A	IV	14.95	F	H3.3 K27M		R213X			E545K						Wu et al. 2014
SJHGG048_A	IV	6.12	M	H3.3 K27M		R156_R158>R (LOH)			E545K				Amp (Microarray CNV)		Wu et al. 2014
SJHGG057_A	IV	12.30	F	H3.3 K27M			W427X		E110del						Wu et al. 2014
SJHGG050_A	IV	5.70	F	H3.3 K27M			I466_P467fs		K575_R577del						Wu et al. 2014
SJHGG006_A	IV	3.66	F	H3.3 K27M		gE224A (LOH)									Wu et al. 2014
SJHGG102_D	IV	4.75	F	H3.3 K27M		R273C, Q144X									Wu et al. 2014
SJHGG105_A	IV	3.35	F	H3.3 K27M		E339X, R248W									Wu et al. 2014
SJHGG051_A	IV	15.70	M	H3.3 K27M	S241F								Amp (Microarray CNV)	NF1 I719fs (LOH)	Wu et al. 2014
SJHGG010325_A1	IV	10.23	M	H3.3 K27M	R273H								Amp (Microarray CNV)		Wu et al. 2014
SJHGG062_A	IV	12.50	F	H3.3 K27M	S241Y								N659K		Wu et al. 2014
SJHGG001_A	IV	5.33	F	H3.3 K27M	R273C (LOH)								A529>15aa		Wu et al. 2014
SJHGG045_A	N/A	11.78	M	H3.3 K27M		M237I (LOH)									Wu et al. 2014
SJHGG060_A	IV	13.27	M	H3.3 K27M		R248Q (LOH)									Wu et al. 2014
SJHGG073_A	N/A	11.10	M	H3.3 K27M		R273C, K120M									Wu et al. 2014
SJHGG007_A	IV	10.90	M	H3.3 K27M	S241F									NF1 R816X, SV (NF1 CNTN5)	Wu et al. 2014
SJHGG053_A	IV	7.31	F	H3.3 K27M	V173A										Wu et al. 2014
SJHGG055_A	IV	6.43	F	H3.3 K27M	H193Y										Wu et al. 2014
SJHGG101_A	IV	17.24	F	H3.3 K27M	C176Y										Wu et al. 2014
SJHGG004_D	IV	5.47	M	H3.3 K27M		L130_N131>L									Wu et al. 2014
SJHGG010324_A1	N/A	9.56	M	H3.3 K27M		R174_E180>R									Wu et al. 2014
SJHGG009_A	IV	6.01	M	H3.3 K27M	R303X										Wu et al. 2014
SJHGG066_A	IV	9.00	F	H3.3 K27M	R234C								Amp (Microarray CNV)		Wu et al. 2014
SJHGG049_A	IV	13.77	F	H3.3 K27M										CCND2 (Amp)	Wu et al. 2014
SJHGG054_A	IV	3.96	F	H3.3 K27M		G66V									Wu et al. 2014
SJHGG067_A	IV	5.60	F	H3.3 K27M											Wu et al. 2014
SJHGG109_A	IV	8.64	F	H3.3 K27M											Wu et al. 2014
SJHGG068_A	IV	2.97	M	H3.3 K27M		Y163C									Wu et al. 2014
SJHGG059_A	IV	8.90	M	WT	G328V			E545K	T576_R577>R						Wu et al. 2014
SJHGG064_A	IV	10.20	M	WT	G328V		Q404X	H1047R							Wu et al. 2014
SJHGG071_A	N/A	6.87	F	H3.3 K27M	G328E		S516X	E545K						ATRX T1610R	Wu et al. 2014
SJHGG005_A	IV	5.35	F	WT	G328V										Wu et al. 2014
SJHGG003_A	IV	15.58	M	WT		P47_F54fs (LOH)							Amp (Microarray CNV), SV (DIP2C_PDGFRA)		Wu et al. 2014
SJHGG052_A	IV	5.94	F	WT		R158G (LOH)									Wu et al. 2014
SJHGG063_A	IV	5.40	F	WT											Wu et al. 2014
SJHGG076_A	IV	10.25	F	H3.3 K27M			E405X								Wu et al. 2014
SJHGG078_A	IV	6.46	M	WT											Wu et al. 2014
SJHGG075_A	IV	1.82	M	WT		A88fs							Amp (Microarray CNV)		Wu et al. 2014
mHGA1	IV	10	M	H3.3 K27M										CHEK2 splicing	Fontebasso et al. 2014
mHGA2	IV	3.5	F	H3.3 K27M	R206H			E545K							Fontebasso et al. 2014
mHGA3	IV	5	M	H3.3 K27M		R273H									Fontebasso et al. 2014
mHGA4	IV	3	F	H3.3 K27M		IHC POSITIVE									Fontebasso et al. 2014
mHGA14	IV	10	F	H3.3 K27M		R282W						Amp (450K CNV), I543_V544	NF1 LOSS (450K CNV)		Fontebasso et al. 2014
mHGA15	IV	8	F	H3.3 K27M		R273P									Fontebasso et al. 2014
mHGA16	IV	12	M	H3.3 K27M		R248W									Fontebasso et al. 2014
mHGA17	IV	6	F	H3.3 K27M		R156Pfs*17)									Fontebasso et al. 2014
mHGA18	IV	11	F	H3.3 K27M		R248W		V344G		LOSS (450K CNV)		Amp (450K CNV)			Fontebasso et al. 2014
mHGA19	IV	7	M	H3.3 K27M		A159Pfs*11, LOSS (450K CNV)									Fontebasso et al. 2014
mHGA20	IV	7	M	H3.3 K27M		V173L									Fontebasso et al. 2014
mHGA21	IV	4.5	F	H3.3 K27M		R175H									Fontebasso et al. 2014
mHGA22	IV	4	M	H3.3 K27M		S241P									Fontebasso et al. 2014
mHGA23	IV	7	F	H3.3 K27M		V157F									Fontebasso et al. 2014
mHGA24	IV	13	F	H3.3 K27M		P152Rfs*27			K447_Y452del					ATRX Q1034fs	Fontebasso et al. 2014
mHGA26	III	14	F	H3.3 K27M								N546K, 655_655del		ATRX C1122Vfs*8)	Fontebasso et al. 2014
mHGA30	IV	6	M	H3.1 K27M		E258G									Fontebasso et al. 2014
mHGA31	IV	4	M	H3.1 K27M	R258G				L449_H450delinsY						Fontebasso et al. 2014
mHGA32	IV	3	M	H3.1 K27M	G328E			C420R							Fontebasso et al. 2014
mHGA33	III	3	F	H3.1 K27M	G328V		C478X	Y1038F, H1047R							Fontebasso et al. 2014
mHGA34	IV	9	F	H3.1C K27M	G356D										Fontebasso et al. 2014
mHGA35	IV	17	M	WT		V173L, LOSS (450K CNV)							Amp(450K CNV)		Fontebasso et al. 2014

mHGA38	IV	6	M	H3.3 K27M		R158G										Fontebasso et al. 2014
mHGA39	IV	7	F	H3.3 K27M		S241F										Fontebasso et al. 2014
mHGA40	IV	8	M	H3.3 K27M		Y220C										Fontebasso et al. 2014
HSJD_DIPG002		6	F	H3.1 K27M	R258G			H1047R								Taylor et al. 2014
NCHP_DIPG011		4.8	F	H3.1 K27M	G328E				573-576 del					IGF2R K162R		Taylor et al. 2014
NCHP_DIPG108		7.5	M	H3.1 K27M	G328V					R130X						Taylor et al. 2014
HSJD_DIPG004		10	F	H3.1 K27M	G328E											Taylor et al. 2014
NCHP_DIPG052		4.6	M	H3.1 K27M	G328V											Taylor et al. 2014
NCHP_DIPG113		4.6	M	H3.1 K27M	G356D											Taylor et al. 2014
NCHP_DIPG114		8.6	M	H3.1 K27M				Q546K								Taylor et al. 2014
NCHP_DIPG103		5.8	F	H3.1 K27M		E221X										Taylor et al. 2014
NCHP_DIPG081		6.7	M	H3.3 K27M		R273H										Taylor et al. 2014
NCHP_DIPG105		6.6	F	H3.3 K27M		R175H										Taylor et al. 2014
NCHP_DIPG061		11.9	F	H3.3 K27M		V157F										Taylor et al. 2014
NCHP_DIPG102		10.3	M	H3.3 K27M		C135Y										Taylor et al. 2014
HSJD_DIPG001		6	F	H3.3 K27M		G244S										Taylor et al. 2014
NCHP_DIPG112		5.6	F	H3.3 K27M		R273C										Taylor et al. 2014
NCHP_DIPG111		10.6	F	H3.3 K27M		156fs										Taylor et al. 2014
NCHP_DIPG109		6.2	M	H3.3 K27M		R273C										Taylor et al. 2014
NCHP_DIPG065		10.2	M	H3.3 K27M		A159V				K567E						Taylor et al. 2014
NCHP_DIPG006		6.3	M	H3.3 K27M			L513X	E542K								Taylor et al. 2014
HSJD_DIPG003		6	M	H3.3 K27M			S468X	H1047R								Taylor et al. 2014
HSJD_DIPG008		6.5	M	H3.3 K27M			429fs									Taylor et al. 2014
NCHP_DIPG107		8.8	M	H3.3 K27M						K379E						Taylor et al. 2014
NCHP_DIPG101		3.9	F	H3.3 K27M										ATM 1190-1191fs		Taylor et al. 2014
HSJD_DIPG007		9.9	M	H3.3 K27M	R206H		428fs	C420R								Taylor et al. 2014
HSJD_DIPG106		12.1	M	WT			483fs	H1047R								Taylor et al. 2014
HSJD_DIPG104		4.4	M	WT										IGF2R D1830R		Taylor et al. 2014
HSJD_DIPG110		5.7	F	WT										NF1 R192X		Taylor et al. 2014
DIPG03	GBM	NA	M	H3.1C K27M		1										Buczkowicz et al. 2014
DIPG62	AA	NA	M	H3.1 K27M	G328E	R43H										Buczkowicz et al. 2014
DIPG26	LGA	NA	F	H3.3 K27M		S339P		E545G								Buczkowicz et al. 2014
DIPG07	GBM	NA	M	H3.3 K27M		R43H										Buczkowicz et al. 2014
DIPG08	GBM	NA	F	H3.3 K27M		R26fs										Buczkowicz et al. 2014
DIPG24	GBM	NA	F	H3.3 K27M		R141C										Buczkowicz et al. 2014
DIPG27	GBM	NA	M	H3.3 K27M		1										Buczkowicz et al. 2014
DIPG57	GBM	NA	F	H3.3 K27M		1										Buczkowicz et al. 2014
DIPG60	GBM	NA	M	H3.3 K27M		1										Buczkowicz et al. 2014
DIPG06	GBM	NA	M	H3.3 K27M		S109P										Buczkowicz et al. 2014
DIPG19	GBM	NA	F	H3.3 K27M		V41L										Buczkowicz et al. 2014
DIPG28	AA	NA	M	H3.3 K27M		1							T281P			Buczkowicz et al. 2014
DIPG30	GBM	NA	M	H3.3 K27M	1											Buczkowicz et al. 2014
DIPG04	GBM	NA	M	H3.3 K27M		V25F										Buczkowicz et al. 2014
DIPG61	GBM	NA	M	H3.3 K27M		1										Buczkowicz et al. 2014
DIPG58	GBM	NA	F	H3.3 K27M					Q325X, N344del							Buczkowicz et al. 2014
DIPG25	AA	NA	F	H3.3 K27M												Buczkowicz et al. 2014
DIPG29	GBM	NA	F	WT												Buczkowicz et al. 2014
DIPG18	GBM	NA	M	WT		1										Buczkowicz et al. 2014
DIPG1	GBM	21.2	M	H3.1 K27M					Splicing						CTNNA2 A260S	Present study
DIPG2	GBM	6.10	M	H3.1 K27M	G328V			H1047R			A126S				MAX R51Q	Present study
DIPG3	GBM	7.5	F	H3.2C K27M	G328V			H1047R								Present study
DIPG4	GBM	5.6	F	H3.3 K27M	R206H	C3F										Present study
DIPG5	GBM	8	M	H3.3 K27M		A175fs	W427X								ATRX D94fs	Present study
DIPG6	Astrocytoma	6	M	H3.3 K27M		G113D		H1047R							ATRX splicing, OLIG2 P215fs	Present study
DIPG7	GBM	10.8	F	H3.3 K27M		R43H										Present study
DIPG8	GBM	5	M	H3.3 K27M		P58fs, E162X										Present study
DIPG9	GBM	9.3	M	H3.3 K27M			E525X								ATRX V1514D	Present study

Supplementary Table 6: Integrated dataset of Diffuse Intrinsic Pontine Glioma samples from four published studies (n=121) and present work (n=9) showing ditribtuion of histone mutations.

Supplementary Table 7: Mutation analysis of biopsy (obtained at diagnosis) and autopsy pairs from the same patients

Sample Name	Type	H3 Mutation	Other Accompanying Mutations	
SJHGG001	Autopsy	H3.3 K27M	TP53_p.R273C (LOH)	PDGFRA_p.A529>15aa
	Biopsy	H3.3 K27M	TP53_p.R273C (LOH)	PDGFRA_p.A529>15aa
SJHGG002	Autopsy	H3.3 K27M	TP53_p.R213*	PIK3CA_p.E545K
	Biopsy	H3.3 K27M	TP53_p.R213*	TP53_p.L252P
mHGA18	Autopsy	H3.3 K27M	TP53_p.R248W	PIK3CA_p.V344G
	Biopsy	H3.3 K27M	TP53_p.R248W	PIK3CA_p.V344G

Supplementary Table 7: Mutation status of the main histone drivers and accompanying genes obtained from analysis of samples acquired at biopsy and autopsy from three Diffuse Intrinsic Pontine Glioma (DIPG) patients.

1 **Discussion**

2 Genomic analysis of autopsy DIPG brains provides a fascinating opportunity to reconstruct the
3 evolution of this deadly tumor, and has important implications on potential therapeutic
4 interventions. Despite the sample number limitation based on disease incidence, this study
5 provides insight into the spatial and temporal evolution of the tumor genome in DIPG helping
6 derive a general model for this disease.

7

8 Our analysis suggests that H3K27M mutations are the initial oncogenic event in DIPGs.
9 However, these mutations are not solely sufficient for tumor formation, as they require, nearly
10 universally, obligate partners for effective tumorigenesis. There is some disagreement in the
11 cancer field as to the definition of “driver” versus “passenger” mutations. The designation
12 passenger has been used for any variant that is not absolutely essential to the tumorigenesis.
13 However, for the purpose of this study – and more generally – we postulate three categories of
14 mutations: 1) Main driver mutations are essential to initiate and continue tumorigenesis; 2)
15 Accessory driver mutations can further promote and accelerate tumor growth but are not
16 absolutely essential; 3) Passenger mutations are neutral and do not affect the tumor. H3K27M
17 and their obligate partners appear thus to be the main driver mutations, since they always co-
18 occur together throughout each tumor.

19 These K27MH3.3 partner mutations are mostly genetic alterations affecting the TP53 pathway,
20 mainly *TP53* and less frequently *PPM1D* genes. H3K27M mutants also favor association with
21 *ACVR1* gain of function mutations and in rare cases (1 in our dataset and 2 in other reported
22 cases) *PIK3R1* mutations. Interestingly, even if more frequent, *PIK3CA* mutations, including
23 *PIK3CAH1047R*, clearly fall into the accessory driver category, since they were invariably sub-

1 clonal in our dataset and in some other reported cases. Similarly, a PTEN mutation was also
2 found in some but not all tumor areas in DIPG2 (Fig. 3). We can only speculate at this time as to
3 why mutations that ultimately activate the PI3K/AKT pathway appear clonal for some of the
4 component genes and subclonal for others. This may be related to the dosage effect of pathway
5 activation, as the loss of the regulatory unit (*PIK3R1*) leads to increased baseline activity of the
6 wild-type kinase whereas the PIK3CAH1047R mutation has been associated with higher kinase
7 activity and oncogenic potential through cellular reprogramming and induction of stemness
8 properties²⁴. Further studies are needed to help ascertain the specific role of these accessory
9 driver mutations in oncohistone tumorigenesis.

10

11 Once acquired, the main driver partnership is maintained throughout the course of the disease
12 (diagnosis to autopsy), in all cells across the primary tumor site, and in tumor spread throughout
13 the brain. The obligate partner is not chosen randomly. One speculation about the phenotypic
14 advantage of the gain of additional mutations may be that the partners are selected as those best
15 suited to increase levels of mutant H3 in cells. Indeed, H3K27M mutants show a dose-dependent
16 effect in inducing cellular proliferation^{26,27}. Canonical H3 variants are the most abundant
17 histones in cells and need cell-cycle division (S-phase) for synthesis, whereas non-canonical
18 H3.3 variants represent ~5% of all H3 in cells and are present throughout the cell cycle.
19 Accordingly, H3K27M mutations in cell cycle-dependent canonical histones may co-occur with
20 alterations affecting ACVR1, a growth factor that induces cell division, resulting in the synthesis
21 of wildtype and mutant H3.1 proteins. H3K27M mutations in the non-canonical H3.3 histone are
22 mainly associated with mutations affecting the TP53 pathway, as these possibly offer evasion
23 from cell death and senescence and provide the needed opportunity for the mutant H3.3K27M to

1 exert its effect over a longer period, effectively reshaping the epigenome and drive tumor
2 formation.

3

4 The previously unsuspected homogeneity for main driver mutations across the course of the
5 disease we uncover in this study indicates that efforts to cure DIPG should be directed at the
6 oncohistone partnership, as other genetic alterations are generally sub-clonal. Our findings
7 further indicate that needle biopsies recommended to orient care are representative of the main
8 drivers in DIPG even if the regional heterogeneity of other secondary targetable alterations, such
9 as PIK3CA mutations, may not be fully captured. Based on early tumor spread, efforts to cure
10 DIPG should aim for early systemic tumor control as opposed to regimens focused on the pons.

1 **Methods**

2 **Patient samples**

3 All patient samples were collected with informed consent in accordance with the respective
4 Ethics Review Boards of the institutions that provided them. DIPG post-mortem specimen
5 procurement was performed as previously described¹⁹. Briefly, brainstem and cerebellum were
6 removed *en bloc* from the whole brain, and dissected into ~9 transverse sections. The cerebral
7 cortex was dissected into ~11 coronal sections. The brainstem, cerebellum, and cerebral cortex
8 sections were alternatively frozen or fixed in formalin. A total of 158 samples were studied by
9 immunohistochemistry and molecular analyses, representing various neuroanatomical locations
10 such as frontal, parietal, temporal, occipital lobes, thalamus, lateral ventricles, hippocampus,
11 midbrain, pons, medulla, and cerebellum (Supplementary Tables 1-3; Supplementary Fig. 1). For
12 molecular studies (RNA, DNA) 134 core punches were obtained using a biopsy punch and
13 plunger (2mm, #33-31, Integra Miltex, York, PA). All histological sections were reviewed by
14 neuropathologists (C-Y.H.; J.K.), according to the World Health Organization (WHO)
15 classification of tumors. Demographical, clinical and histopathological characteristics of all
16 specimens are presented in Supplementary Table 1.

17

18 **Immunohistochemistry**

19 Immunohistochemistry was performed on 5µm thick FFPE slides. Briefly, slides were de-
20 paraffinized, processed for epitope retrieval, DAB detected using reagents customized for the
21 Leica BOND-MAX automated stainer (Leica Biosystems, Buffalo Grove, IL). Processed slides
22 were probed by immunohistochemical assay for hematoxylin and eosin (H&E), Ki67, H3-K27M,
23 and H3-K27me3 as previously described¹² (Supplementary Fig. 2).

1

2 **Antibodies**

3 Rabbit monoclonal anti-Ki67 (Biocare Medical, Concord, CA) were pre-diluted and ready to use.
4 Rabbit polyclonal anti-H3K27M (#ABE419 Millipore, Billerica, MA, 1:500), rabbit monoclonal
5 anti-H3K27me3 (C36B11, #9733 Cell Signaling, Beverly, MA 1:75) were diluted in Bond
6 primary antibody diluent (#AR9352 Leica Biosystems, Buffalo Grove, IL). Secondary detection
7 was conducted using the Bond polymer refined detection kit (Leica Biosystems, Buffalo Grove,
8 IL). Slides were counterstained for hematoxylin nuclear stain.

9

10 **RNA and DNA extractions**

11 Frozen tissue samples were homogenized with Trizol, and nucleic acids were phase separated
12 using chloroform. Total RNA was extracted according to the PicoPure RNA Isolation kit
13 (Arctus Bioscience Inc. Mountain view, CA). Genomic DNA was extracted from frozen tissue
14 using the Genra Puregene DNA extraction kit, or from FFPE tissue using the QiaAmp DNA
15 mini kit according to the manufacturer's instructions (Qiagen, Valencia, CA). All DNA
16 quantifications were conducted using the Quant-iT Picogreen dsDNA assay kit (Life
17 Technologies, Carlsbad, CA).

18

19 **Droplet Digital PCR**

20 Digital droplet PCR (ddPCR) assays were performed according to standard methods. Briefly,
21 each 20 ul reaction contained 1X ddPCR Supermix for Probes (Bio-Rad), 900 nM gene specific
22 HPLC-purified forward and reverse primer, 250 nM gene-specific mutant or wild-type LNA
23 probe and 12.5-25.0 ng genomic DNA. Each reaction was mixed with 60 ul Droplet Generation

1 Oil (Bio-Rad), partitioned into ~12,000 -16,000 droplets in QX100 Droplet Generator (Bio-Rad),
2 transferred to a 96-well plate and sealed. The primers and probes were designed by Integrated
3 DNA Technologies (IDT) as follows: forward primer for *H3F3A*:
4 5'-GTACAAAGCAGACTGCCCCGCAAAT-3', reverse primer
5 5'-GTGGATACATACAAGAGAGACTTTGTCCC-3'. Forward primer for *HIST1H3B*: 5'-
6 ACAGACGTCTCTGCAGGCAAGC-3', reverse primer 5'-GGCGGTAACGGTGAGGCTTT-
7 3'. *H3F3A* K27M wild-type probe (HEX) CA+C+T+C+T+T+GC and mutant probe (FAM)
8 CA+CT+C+A+T+GCG. *HIST1H3B* K27M wild-type probe (HEX) T+CGC+A+A+GAG+CG
9 and mutant probe (FAM) TCGC+A+T+G+AGCG. The PCRs were performed in a T100
10 Thermal Cycler (Bio-Rad) with the following cycling conditions: 1× (95°C for 10 min), 50×
11 (95°C for 30 s, 53°C or 57°C for 60 s, with 2°C /s ramp rate and 1× (98°C for 10 min).
12 Following end-point amplification, the fluorescence intensity of individual droplets was
13 measured with the QX100 Droplet Reader (Bio-Rad) and data analysis was performed with
14 QuantaSoft droplet reader software (Bio-Rad). Positive and negative droplet populations were
15 detected either automatically or manually on two-dimensional graphs and target DNA
16 concentrations were calculated using the Poisson statistics²⁸. The absolute transcript levels were
17 initially computed as copies/μl PCR for both mutant and wild-type and then presented as percent
18 of total gene copy.

19

20 **Whole-exome sequencing**

21 Genomic DNA was extracted from multiple post-mortem samples patient using the standard
22 extraction methods as described by Qiagen. Nextera Rapid Capture Exome kit was used to
23 prepare the paired-end libraries according to the manufacturer's instructions using on average 36

1 ng of total starting genomic DNA. Sequencing was performed on Illumina HiSeq 2000 using
2 rapid-run mode with 100 bp paired-end reads. Next, adaptor sequences were removed; reads
3 were trimmed for quality using the FASTX-Toolkit. An in-house program was used to ensure the
4 presence of exclusively paired-reads to be used in further steps of the analysis. We next aligned
5 the reads using Burrows-Wheeler Aligner (BWA) 0.7.7 to hg19 as reference genome. Indel
6 realignment was performed using the Genome Analysis Toolkit (GATK)²⁹ We next marked the
7 duplicate reads using Picard and excluded them from further analyses as previously described⁶.
8 The coverage of consensus coding sequence (CCDS) bases was assessed using GATK. The
9 average coverage over all the samples was 70x. The majority of samples had >90% of CCDS
10 bases covered by at least 10 reads and > 83% of CCDS bases covered by at least 20 reads.
11 We called SNVs and short indels using SAMtools mpileup with the extended base alignment
12 quality (BAQ) adjustment (-E)^{30,31}. Next we filtered them for quality so that at least 10% of
13 reads supporting each variant call. We used both ANNOVAR³² and in-house tools to annotate
14 the variants and to identify whether these variants affect protein-coding sequence and if they had
15 previously been observed in datasets including the 1000 Genomes Project data set (November
16 2011), the National Heart, Lung, and Blood Institute (NHLBI) Grand Opportunity (GO) exomes
17 or in approximately 3,000 exomes previously sequenced at our center. Results of whole exome
18 sequencing are summarized in Supplementary Table 4 and presented for individual patients as
19 follows: DIPG1 in Supplementary Data 1; DIPG2 in Supplementary Data 2; DIPG3 in
20 Supplementary Data 3; DIPG4 in Supplementary Data 4; DIPG5 in Supplementary Data 5;
21 DIPG6 in Supplementary Data 6; DIPG7 in Supplementary Data 7; DIPG8 in Supplementary
22 Data 8; DIPG9 in Supplementary Data 9.

23

1 **Exploratory targeted high-depth DNA sequencing of hotspot mutations**

2 Genomic DNA from DIPG samples were used for high-depth sequencing using Illumina MiSeq
3 platform. The MiSeq panel covers exons in 16 Histone H3 isoforms (10 H3.1, 2 H3.2 and 3 H3.3
4 genes) and covering hotspot mutations such as *IDH1* mutation (codon 132), *IDH2* (codons 140
5 and 172), *ACVRI* (exons 6-9) and *BRAF* (exons 11 and 15) and *PPM1D* (exon 6). Genomic
6 DNA samples were sequenced using the MiSeq sequencing platform (Illumina) as previously
7 described⁶ with an average coverage of > 20,000X of the analogous K27M base change across
8 the three histone variants.

9 To estimate allele frequency of mutations identified using whole exome sequencing genomic
10 DNA from the same samples was also used for high-depth sequencing on the MiSeq platform
11 with an average coverage of > 4,000 x. Reads were mapped to the reference genome (human
12 hg19) using the BWA genome alignment^{30,31}. Alignment files were fed to an in-house program
13 to calculate different variations' allele frequencies at the desired positions.

14

15 **RNA Sequencing**

16 We used Qiagen RNeasy Lipid Tissue Mini kit to extract RNA from tumor DIPG3 (Pons 1 and
17 Pons 2) according to manufacturer's instructions. Library was prepared using rRNA depletion
18 methods according to instruction from Epicentre (manufacturer) to achieve greater coverage of
19 mRNA. Paired-end sequencing was performed on the Illumina HiSeq 2000 platform.

20

21 **DNA Methylation Analysis**

22 Methylation profiling data was analyzed as previously described⁶. The raw data were subject to
23 quality control and preprocessing utilizing the R package minfi, and normalized for technical

1 variation between the Infinium I and II probes using the SWAN method. We removed probes on
2 sex chromosomes (chrX, Y), those containing SNPs (dbSNP:
3 <http://www.ncbi.nlm.nih.gov/SNP/>) as well as non-specific probes that bind to multiple genomic
4 locations. Unsupervised hierarchical clustering was performed using average linkage and
5 Pearson rank correlation distance on the top 5,000 most variable probes selected based on
6 standard deviation of beta values (β -values).

7

8 **Copy Number Variation analysis**

9 To study copy number variations in our samples we developed an in-house program to calculate
10 the deviation of B allele frequency from 50% as well as normalized coverage from whole exome
11 sequencing data (adapted from methods used in FishingCNV³³ and ExomeAI³⁴). Different CNV
12 events (duplication, deletion, copy neutral LOH) were called based upon the B allelic imbalance
13 and the status of the normalized coverage as follow: Deviation from 50% B allele frequency and
14 an increase in normalized coverage was considered as amplification, Deviation from 50% B
15 allele frequency and decrease in normalized coverage as deletion, and Deviation from 50% B
16 allele frequency and no change in the normalized coverage was considered as potential copy
17 neutral loss of heterozygosity. We mainly assessed the CNV events at the chromosomal arm
18 level. The results of our CNV detection are presented in Supplementary Fig. 5 and
19 Supplementary Table 14.

20

21 **OncoScan verification of the CNV events**

22 OncoScan® FFPE Assay Kit, provided by Affymetrix, is a platform based on Molecular
23 Inversion Probe (MIP) technology, used to asses copy number and loss of heterozygosity using

1 small amounts of DNA from FFPE samples. We performed this assay on several samples in
2 order to verify our WES based CNV detection method. Genomic DNA was quantified using
3 Picogreen protocol (Quant-iT™ PicoGreen® dsDNA Products, Invitrogen, P-7589) and read on
4 SpectraMAX GeminiXS Spectrophotometer. The OncoScan® FFPE Assay Kit was used
5 according to the manufacturer's instructions (Affymetrix). A GeneAmp PCR system 9700
6 Thermal Cycler (Life Technologies) was used from the Anneal stage to the Denaturation stage.
7 QC gels of the PCR and HaeIII digest products were performed on E-Gel® 48 4% Agarose Gels
8 using Mother E-Base™ Device (Life Technologies) and imaged with SYNGENE GeneGenius
9 Bio Imaging System (Syngene). The digest DNA target was hybridized on OncoScan® Array,
10 (Affymetrix) and incubated at 49°C in the Genechip® Hybridization oven 640 (Affymetrix) for
11 17 hours at 60 rpm. OncoScan® Arrays were then washed in a GeneChips® Fluidics Station 450
12 (Affymetrix) using OncoScan® Stain and Wash Reagents according to the manufacturer's
13 instructions (Affymetrix). The microarrays were finally scanned on a GeneChip® scanner 3000
14 (Affymetrix). Data QC analysis was performed with the OncoScan Consol 1.2.0.50 software
15 (Affymetrix) using OncoScan Analysis Library files r1.1. OncoScan® Positive and negative
16 Control supplied in the OncoScan® FFPE Assay Kit were used for internal controls to assess the
17 performance of each run. CNV events were called using the normalized data using Nexus
18 Express for OncoScan 3.1 (Affymetrix). The OncoScan plots are represented in Supplementary
19 Fig. 6.

20

21 **Constructing evolutionary trees**

22 We used PhyloWGS³⁵ to reconstruct the tumor phylogeny, which uses a Bayesian approach to
23 infer cellular frequencies from mutation allele frequencies. It applies Dirichlet process to cluster

1 mutations with similar cellular frequencies and the tree-structured stick-breaking process to
2 model the clonal evolutionary tree. For multi-region samples from the same patient, we
3 normalized the read counts used for phylogenetic tree construction by copy number counts from
4 CNV analysis. Read counts were corrected for the CNV events (Supplementary Fig. 5; Table 3)
5 for each sample as follow: in case of duplication $Ref' = Ref$, $Alt' = Alt/2$; in case of deletion Ref'
6 $= Ref + Alt$, $Alt' = Alt$; in case of a gene on chromosome X $Ref' = Ref *2 + Alt$, $Alt' = Alt$; in
7 case of Copy Neutral LOH $Ref' = Ref+Alt/2$, and $Alt' = Alt/2$. The reconstructed trees were
8 redrawn in Graphviz by using in-house scripts adapted from AncesTree³⁶ to show the trajectories
9 of mutations with contribution greater than 0.05.

10

11 **Differential Expression and Gene Set Enrichment Analysis**

12 We aligned RNASeq data from DIPG3-Pons 1 and DIPG3-Pons 2 (RNA extraction and
13 sequencing protocol described before) using STAR 2.3.0³⁷, and analyzed differentially expressed
14 genes using DeSEQ2³⁶. The top 200 over expressed and top 200 under expressed genes in
15 DIPG3-Pons 2 (the sub-clone with PIK3CA activating mutation) were analyzed for geneset
16 enrichment using both AmiGo 2 tool³⁸ provided by Gene Ontology and DAVID³⁹. We used
17 PANTHER Overrepresentation Test (release 20150430) for analysis type by Amigo 2
18 (annotation version and release date: GO Ontology database Released 2015-08-06.) The top
19 pathways found by AmiGo 2 to be enriched with enrichment folds higher than 5 and (Bonferroni
20 < 0.05) were retained. We used functional annotation clustering and set the stringency to the
21 highest in DAVID and filtered the results for enrichment folds higher than 5 (Bonferroni
22 corrected p -value < 0.05 .) We used the set of genes with at least 50 RNASeq reads in both
23 DIPG3 Pons1 and 2 combined as background gene set in this analysis (Supplementary Data 10).

1
2
3
4
5
6
7
8
9
10
11
12

URLs

FASTX-Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/; Genome Analysis Toolkit (GATK), <http://www.broadinstitute.org/gsa/wiki/>; Picard, <http://picard.sourceforge.net/>; SAMtools, <http://samtools.sourceforge.net/>; dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/>.
Amigo 2, <http://amigo.geneontology.org/amigo>; DAVID, <https://david.ncifcrf.gov/>.

Accession Numbers

Whole-exome sequencing data for all tumors (along with RNA sequencing data for DIPG3), and also DNA methylation profiles can be accessed through the European Genome Archive (EGA) at the following accession EGAS00001001654 and Gene Expression Omnibus (GEO) under accession GSE77353 respectively.

1 **References**

- 2
- 3 1. Buczkowicz, P., Bartels, U., Bouffet, E., Becher, O. & Hawkins, C. Histopathological
- 4 spectrum of paediatric diffuse intrinsic pontine glioma: diagnostic and therapeutic
- 5 implications. *Acta Neuropathol* **128**, 573-581 (2014).
- 6 2. Khuong-Quang, D.A. *et al.* K27M mutation in histone H3.3 defines clinically and
- 7 biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. *Acta*
- 8 *Neuropathol* **124**, 439-447 (2012).
- 9 3. Schwartzenuber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling
- 10 genes in paediatric glioblastoma. *Nature* **482**, 226-231 (2012).
- 11 4. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas
- 12 and non-brainstem glioblastomas. *Nat Genet* **44**, 251-253 (2012).
- 13 5. Buczkowicz, P. *et al.* Genomic analysis of diffuse intrinsic pontine gliomas identifies
- 14 three molecular subgroups and recurrent activating ACVR1 mutations. *Nat Genet* **46**,
- 15 451-456 (2014).
- 16 6. Fontebasso, A.M. *et al.* Recurrent somatic mutations in ACVR1 in pediatric midline
- 17 high-grade astrocytoma. *Nat Genet* **46**, 462-466 (2014).
- 18 7. Taylor, K.R. *et al.* Recurrent activating ACVR1 mutations in diffuse intrinsic pontine
- 19 glioma. *Nat Genet* **46**, 457-461 (2014).
- 20 8. Wu, G. *et al.* The genomic landscape of diffuse intrinsic pontine glioma and pediatric
- 21 non-brainstem high-grade glioma. *Nat Genet* **46**, 444-450 (2014).
- 22 9. Kumar, A. *et al.* Deep sequencing of multiple regions of glial tumors reveals spatial
- 23 heterogeneity for mutations in clinically relevant genes. *Genome Biol* **15**, 530 (2014).

- 1 10. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer
2 evolutionary dynamics. *Proc Natl Acad Sci U S A* **110**, 4009-4014 (2013).
- 3 11. Szerlip, N.J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and
4 PDGFRA amplification in glioblastoma defines subpopulations with distinct growth
5 factor response. *Proc Natl Acad Sci U S A* **109**, 3041-3046 (2012).
- 6 12. Bechet, D. *et al.* Specific detection of methionine 27 mutation in histone 3 variants
7 (H3K27M) in fixed tissue from high-grade astrocytomas. *Acta Neuropathol* **128**, 733-741
8 (2014).
- 9 13. Caretti, V. *et al.* Subventricular spread of diffuse intrinsic pontine glioma. *Acta*
10 *Neuropathol* **128**, 605-607 (2014).
- 11 14. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by
12 multiregion sequencing. *N Engl J Med* **366**, 883-892 (2012).
- 13 15. Eleveld, T.F. *et al.* Relapsed neuroblastomas show frequent RAS-MAPK pathway
14 mutations. *Nat Genet* **47**, 864-871 (2015).
- 15 16. Wu, X. *et al.* Clonal selection drives genetic divergence of metastatic medulloblastoma.
16 *Nature* **482**, 529-533 (2012).
- 17 17. Johnson, B.E. *et al.* Mutational analysis reveals the origin and therapy-driven evolution
18 of recurrent glioma. *Science* **343**, 189-193 (2014).
- 19 18. Walker, D.A. *et al.* A multi-disciplinary consensus statement concerning surgical
20 approaches to low-grade, high-grade astrocytomas and diffuse intrinsic pontine gliomas
21 in childhood (CPN Paris 2011) using the Delphi method. *Neuro Oncol* **15**, 462-468
22 (2013).

- 1 19. Kambhampati, M. *et al.* A standardized autopsy procurement allows for the
2 comprehensive study of DIPG biology. *Oncotarget* **6**, 12740-12747 (2015).
- 3 20. Castel, D. *et al.* Histone H3F3A and HIST1H3B K27M mutations define two subgroups
4 of diffuse intrinsic pontine gliomas with different prognosis and phenotypes. *Acta*
5 *Neuropathol* **130**, 815-827 (2015).
- 6 21. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell*
7 **149**, 979-993 (2012).
- 8 22. Popic, V. *et al.* Fast and scalable inference of multi-sample cancer lineages. *Genome Biol*
9 **16**, 91 (2015).
- 10 23. Zhang, L. *et al.* Exome sequencing identifies somatic gain-of-function PPM1D mutations
11 in brainstem gliomas. *Nat Genet* **46**, 726-730 (2014).
- 12 24. Koren, S. *et al.* PIK3CA(H1047R) induces multipotency and multi-lineage mammary
13 tumours. *Nature* **525**, 114-118 (2015).
- 14 25. Karar, J. & Maity, A. PI3K/AKT/mTOR Pathway in Angiogenesis. *Front Mol Neurosci*
15 **4**, 51 (2011).
- 16 26. Lewis, P.W. *et al.* Inhibition of PRC2 activity by a gain-of-function H3 mutation found in
17 pediatric glioblastoma. *Science* **340**, 857-861 (2013).
- 18 27. Funato, K., Major, T., Lewis, P.W., Allis, C.D. & Tabar, V. Use of human embryonic
19 stem cells to model pediatric gliomas with H3.3K27M histone mutation. *Science* **346**,
20 1529-1533 (2014).
- 21 28. Hindson, C.M. *et al.* Absolute quantification by droplet digital PCR versus analog real-
22 time PCR. *Nat Methods* **10**, 1003-1005 (2013).

- 1 29. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for
2 analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).
- 3 30. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler
4 transform. *Bioinformatics* **25**, 1754-1760 (2009).
- 5 31. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
6 2078-2079 (2009).
- 7 32. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic
8 variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
- 9 33. Shi, Y. & Majewski, J. FishingCNV: a graphical software package for detecting rare
10 copy number variations in exome-sequencing data. *Bioinformatics* **29**, 1461-1462 (2013).
- 11 34. Nadaf, J., Majewski, J. & Fahiminiya, S. ExomeAI: detection of recurrent allelic
12 imbalance in tumors using whole-exome sequencing data. *Bioinformatics* **31**, 429-431
13 (2015).
- 14 35. Deshwar, A.G. *et al.* PhyloWGS: reconstructing subclonal composition and evolution
15 from whole-genome sequencing of tumors. *Genome Biol* **16**, 35 (2015).
- 16 36. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion
17 for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
- 18 37. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21
19 (2013).
- 20 38. Carbon, S. *et al.* AmiGO: online access to ontology and annotation data. *Bioinformatics*
21 **25**, 288-289 (2009).
- 22 39. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of
23 large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).

1 **Acknowledgements**

2 The authors would like to express their sincere gratitude toward all the patients' families, as well
3 as the staff at the McGill University and Genome Quebec Innovation Centre for excellent
4 technical expertise, library preparation and sequencing. This work was performed within the
5 context of the I-CHANGE consortium (International Childhood Astrocytoma inTegrated
6 Genomics and Epigenomics consortium) and supported by funding from Genome Canada,
7 Genome Quebec, The Institute for Cancer Research of the Canadian Institutes for Health
8 Research (CIHR) McGill University and the Montreal Children's Hospital Foundation. It was
9 also supported by UL1TR000075 and KL2TR000076 from the NIH National Center for
10 Advancing Translational Sciences, The Smashing Walnuts Foundation (Middleburg, VA), The
11 Zickler Family Foundation (Chevy Chase, MD), Goldwin Foundation (St. Lincoln, NE), The
12 Piedmont Community Foundation (Middleburg, VA), The Musella Foundation (Hewlett, NY),
13 The Mathew Larson Foundation (Franklin Lake, NJ), and Brain Tumor Foundation for Children
14 (Atlanta, GA). N. Jabado is a member of the Penny Cole lab and the recipient of a Chair de
15 Recherche from le Fond de Recherche en Sante au Quebec. J. Majewski holds a Canada
16 Research Chair (tier 2). T. Gayden is supported by a studentship from the CIHR. D. Bechet is
17 supported by a studentship from the T.D trust/Montreal Children's Hospital Foundation. E.
18 Panditharatna is a predoctoral student in the Molecular Medicine Program of the Institute for
19 Biomedical Sciences at the George Washington University. This work is from a dissertation to
20 be presented to the above program in partial fulfillment of the requirements for the Ph.D. degree.

21

22

23

1

2 **Author contributions**

3 E.P., L.M., D.F., D.B. performed experiments, E.P., L.M., H.N., R.L., T.G., M.K., E.I.H., M.O.
4 S.P.C., A.S., C.Y.H., J.K., N.J., J.N., J.M. analyzed the data and produced figures and tables,
5 K.L.L., B.E., W.J.I., A.S.M., C.S., K.E.W., R.J.P. provided tissue samples. N.J., J.N. and J.M.
6 provided project leadership and designed the study.

7

8 **Conflict of Interest**

9 The authors declare that they have no competing financial interests.

10

11

1 **Figure Legends**

2 **Figure 1.** Oncogenic alterations in 41 sub-regions from nine DIPG patients from whole exome
3 sequencing data. Samples representing different anatomical locations within each patient are
4 represented in columns. The mutations (in rows) were selected based on published datasets in
5 pediatric glioblastoma and specifically DIPG. Mutations were divided into two subgroups; driver
6 mutations which are essential for tumor initiation/maintenance and accessory driver mutations,
7 which can further promote and accelerate tumor growth, but are not absolutely essential for
8 tumor initiation or maintenance.

9
10 **Figure 2. Selected examples of clonal evolution within DIPG tumors.** Left: histograms show
11 the raw allele frequencies (whole exome sequencing data) for each somatic mutation in different
12 autopsy regions within each tumor. Red: ubiquitous mutations across regions; yellow: mutations
13 shared in at least two regions; blue: mutations seen in only one region. Right: Phylogenetic trees
14 constructed from the mutation allele frequencies of deep amplicon sequencing data showing the
15 order of evolution along with support probabilities (upper portions of graphs) and clonal mixing
16 proportions within samples (lower portions). For clarity, only mutations selected to be likely
17 oncogenic are shown. A) DIPG5: a rare case harboring both TP53 and PPM1D mutations, which
18 are generally found to be mutually exclusive. PPM1D and TP53 mutations occur in distinct
19 clones and are both secondary to H3K27M. ATRX is also secondary and subclonal. B) DIPG6:
20 while it is impossible to resolve the order of H3/TP53/ATRX mutations' appearance, PIK3CA is
21 clearly sub-clonal and appears in the later stages of evolution within this tumor. C) DIPG2: the
22 H3.1 K27M and ACVR1 main driver mutations are ubiquitous, occur at similar frequencies
23 across all samples, and their mutations order cannot be resolved. Conversely, other accessory

1 driver mutations are clearly secondary in order of appearance, and are present only in distinct
2 subclones.

3

4 **Figure 3. Tumor spread in DIPG.** A) Tumor spread in DIPG2 in the thalamus, cerebellum and
5 brain stem. Tumor extension in thalamus harbors secondary mutation PIK3CA, MAX, and
6 PTEN which indicates late spread from both Pons 1 and Pons 2. Extension towards cerebellum is
7 relatively early in the tumor evolution as it lacks secondary mutations found in the primary
8 tumor and other brainstem spread. B) Evolution of tumor in patient DIPG3. Autopsy revealed
9 two morphologically and histologically distinct regions of the tumor, indicated as DIPG3 Pons 1
10 (low-grade) and DIPG3 Pons 2 (high-grade). Exome sequencing identified 11 SNVs and several
11 large scale CNAs common to both regions. Shared alterations included H3.2 K27M and ACVR1
12 G328V mutations that are likely the main driver mutations in this patient. The analysis also
13 indicated a clear clonal substructure of the two regions, with 18 SNVs and 1 CNA found only in
14 DIPG3 Pons 1, and 11 SNVs and 3 CNAs unique to DIPG3 Pons 2. Intriguingly, DIPG3 Pons 2
15 carries the activating PIK3CA H1047R mutation, which occurs early in the evolution of this sub-
16 clone judging by its high allelic frequency. PIK3CA H1047R is associated with multi-potency²⁴
17 and PI3K activation with angiogenesis and growth and this mutation likely contributes to tumor
18 aggressiveness and high-grade features of Pons 2 compared to Pons 1 in DIPG3. *Scale bar 500*
19 *μm.*