# Transcriptional alterations in glioma result primarily from DNA-methylation independent mechanisms

Franck Court, Elisa Le Boiteux, Anne Fogli, Mélanie Müller-Barthélémy, Catherine Vaurs-Barrière, Emmanuel Chautard, Bruno Pereira, Julian Biau, Jean-Louis Kemeny, Toufic Khalil, Lucie Karayan-Tapon, Pierre Verrelle and Philippe Arnaud

## Supplemental Figures

- **Figure S1:** DNA methylation alteration in IDHwt and IDHmut glioma samples
- **Figure S2:** Transcriptional alterations in IDHwt and IDHmut glioma samples.
- **Figure S3**: Different molecular pathways can lead to the same aberrant gene expression pattern in IDHwt and IDHmut samples.
- **Figure S4:** Most of transcriptional affected genes do no show methylation alteration in their CGI/promoter in IDHwt and IDHmut samples.
- **Figure S5:** Gene Ontology
- **Figure S6:** Expression of genes with a methylated CGI/promoter
- **Figure S7:** Affected genes are not enriched in 5-hmC
- **Figure S8:** Gene repression is associated with H3K27me3 gain and affect mainly genes with brain-specific expression profile
- **Figure S9:** Four classes of expression defects in IDHmut glioma samples
- **Figure S10:** Genes with bivalent chromatin signature in ES cells are more prone to be deregulated in IDHmut glioma.
- **Figure S11:** Gene repression is associated with H3K27m3 gain in IDHmut glioma.
- **Figure S12**: Summary table of data obtained by treating HM450K data with conditions described in material and method section (initial analysis) or by using filter described in Zhou *et al.*, 2017 associated to a Wilcoxon-rank sum test (« wilcox test & Zhou *et al.*, 2017 »).

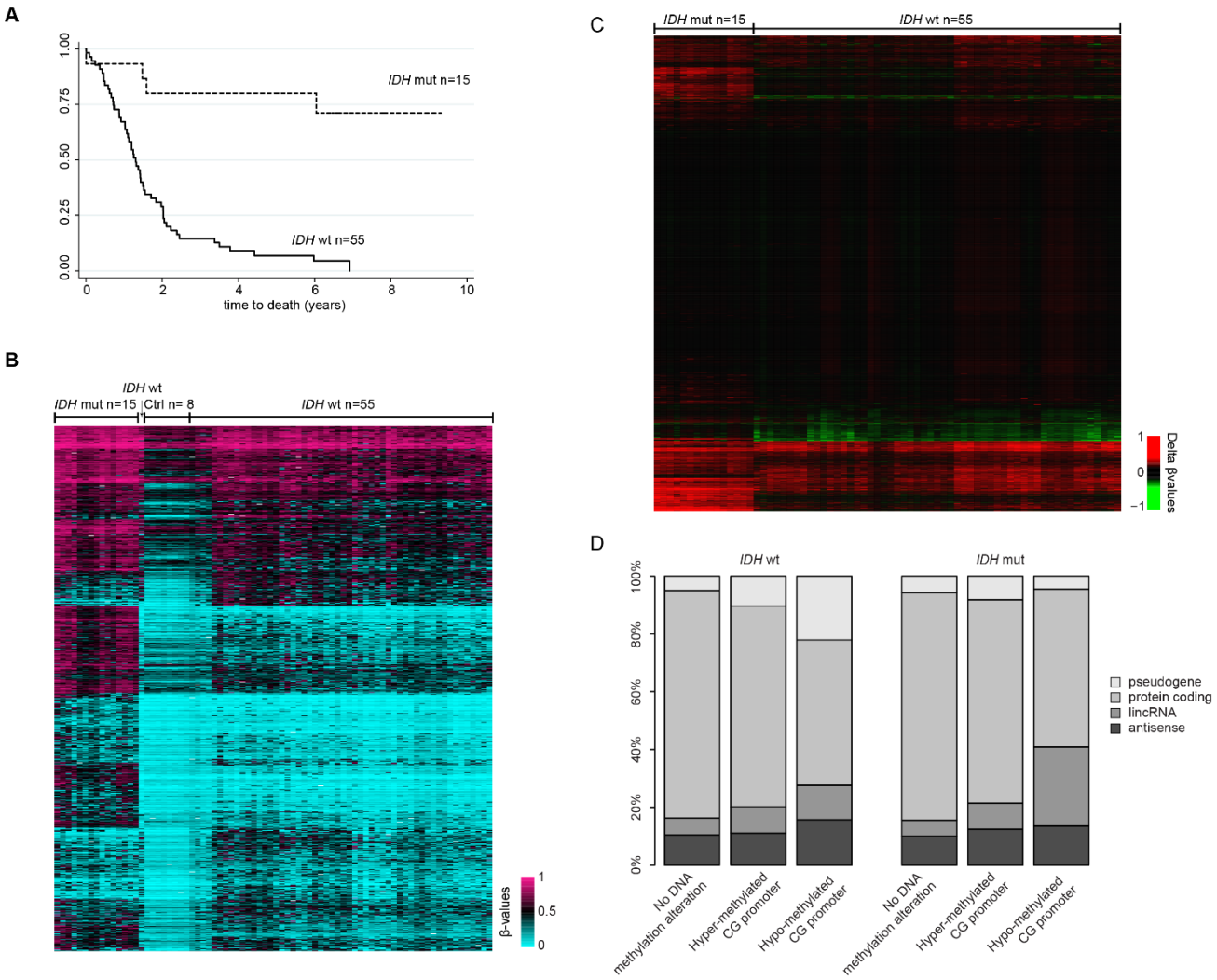
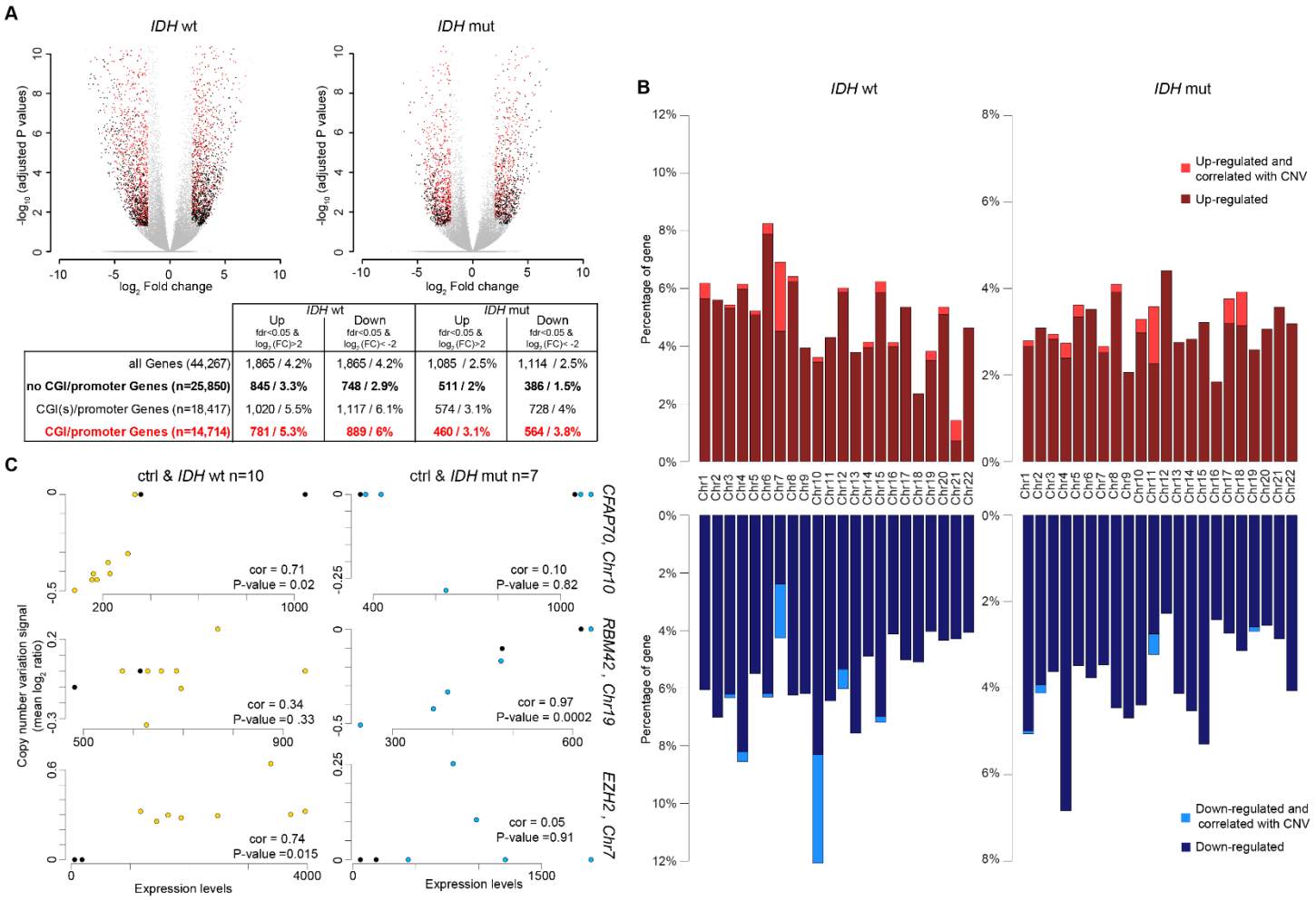



**Figure S1: DNA methylation alteration in *IDH*wt and *IDH*mut glioma samples**

**A)** Kaplan–Meier survival curves for patients with *IDH*wt (n=55) and *IDH*mut (n=15) glioma confirmed that the *IDH* status is a significant prognostic marker: 2-year survival rate of 27% in patients with *IDH*wt (median survival of 1.2 years/14.4 months) and 83% in patients with *IDH*mut glioma (median survival not reached). **B)** Visualization of the DNA methylation level (β-values) at the 86,157 probes (row) located in the 11,795 CGI/promoters analyzed in this study in *IDH*mut, controls and *IDH*wt glioma samples (column) identified a G-CIMP profile in *IDH*mut glioma samples. **C)** Differential DNA methylation level *vs* controls (delta of the means of β-values) of the 11,795 CGIs (row) analyzed in this study in *IDH*mut and *IDH*wt glioma samples (columns). **D)** Classification of the genes associated with hyper-, hypo-methylated or unaffected CGI/promoters, respectively, in *IDH*wt (upper panel) and *IDH*mut (lower panel) glioma samples.




| | IDH wt Up fdr<0.05 & log2(FC)>2 | IDH wt Down fdr<0.05 & log2(FC)< -2 | IDH mut Up fdr<0.05 & log2(FC)>2 | IDH mut Down fdr<0.05 & log2(FC)< -2 |
|---|---|---|---|---|
| all Genes (44,267) | 1,865 / 4.2% | 1,865 / 4.2% | 1,085 / 2.5% | 1,114 / 2.5% |
| **no CGI/promoter Genes (n=25,850)** | **845 / 3.3%** | **748 / 2.9%** | **511 / 2%** | **386 / 1.5%** |
| CGI(s)/promoter Genes (n=18,417) | 1,020 / 5.5% | 1,117 / 6.1% | 574 / 3.1% | 728 / 4% |
| **CGI/promoter Genes (n=14,714)** | **781 / 5.3%** | **889 / 6%** | **460 / 3.1%** | **564 / 3.8%** |

**Figure S2: Transcriptional alterations in *IDH*wt and *IDH*mut glioma samples.**
**A)** Volcano plot analysis with DESeq2 of the differential expression of all annotated genes (n= 44,267) in *IDH*wt (left) and *IDH*mut (right) glioma samples compared with healthy brain controls. Genes that were significantly up- or down-regulated are symbolized by a red dot for CGI/promoter-associated genes, and by a black dot for the others. The numerical values are given in the table. Data are given for all CGI(s)/promoter associated genes (n=18,417) and for those retained in this analysis following filtering, as described in the Materials and Methods section (n=14,714). **B)** Distribution per chromosome of transcriptionally up- (upper panels) and down-regulated (lower panels) CGI/promoter-associated genes in *IDH*wt and *IDH*mut glioma samples. The relative proportion of gene in which transcription alteration correlated with CNV is indicated. **C)** Details of the correlation analyses between CNV and expression for the *CFAP70*, *EZH2* and *RBM42* genes in *IDH*wt (yellow dots in the left panels) and *IDH*mut (blue dots in the right panels) glioma samples. Black dots indicate the expression value in healthy brain. *CFAP70*downregulation and *EZH2* overexpression correlated with CNV in *IDH*wt glioma samples. *RBM42* downregulation correlated with CNV in *IDH*mut glioma samples.

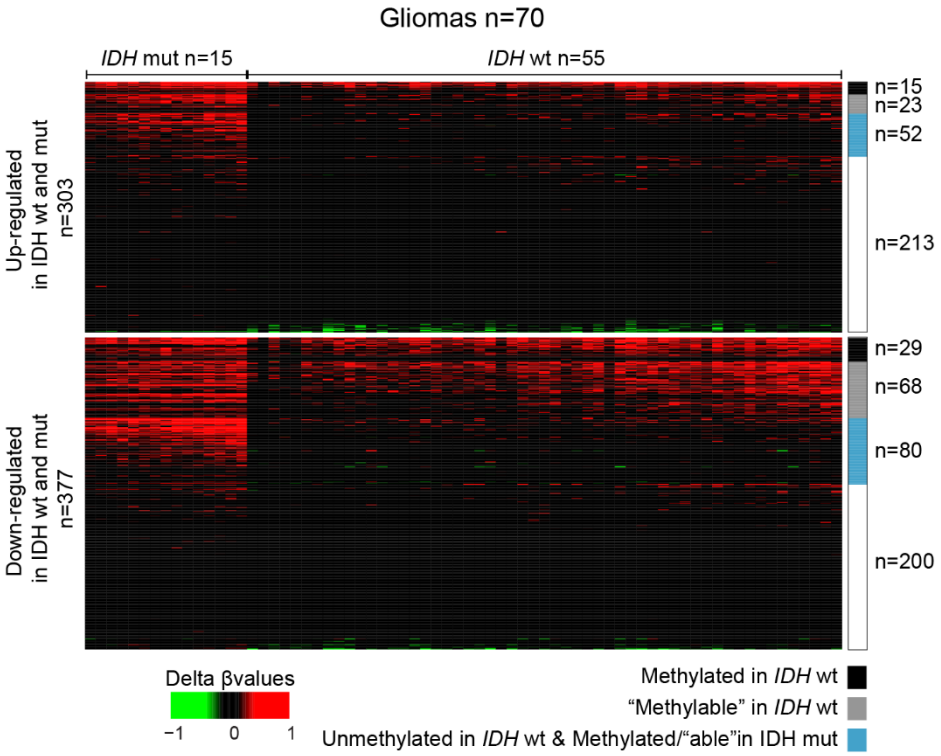



**Figure S3: Different molecular pathways can lead to the same aberrant gene expression pattern in *IDH*wt and *IDH*mut samples.**

Differential methylation pattern, *vs* healthy controls, of genes affected both in *IDH*wt and *IDH*mut glioma samples. Genes identified as methylated and methylable in *IDH*wt glioma samples are shown in black and grey, respectively, on the right column. Genes unmethylated in *IDH*wt glioma samples and methylated/"methylable" in *IDH*mut glioma samples are shown in blue.

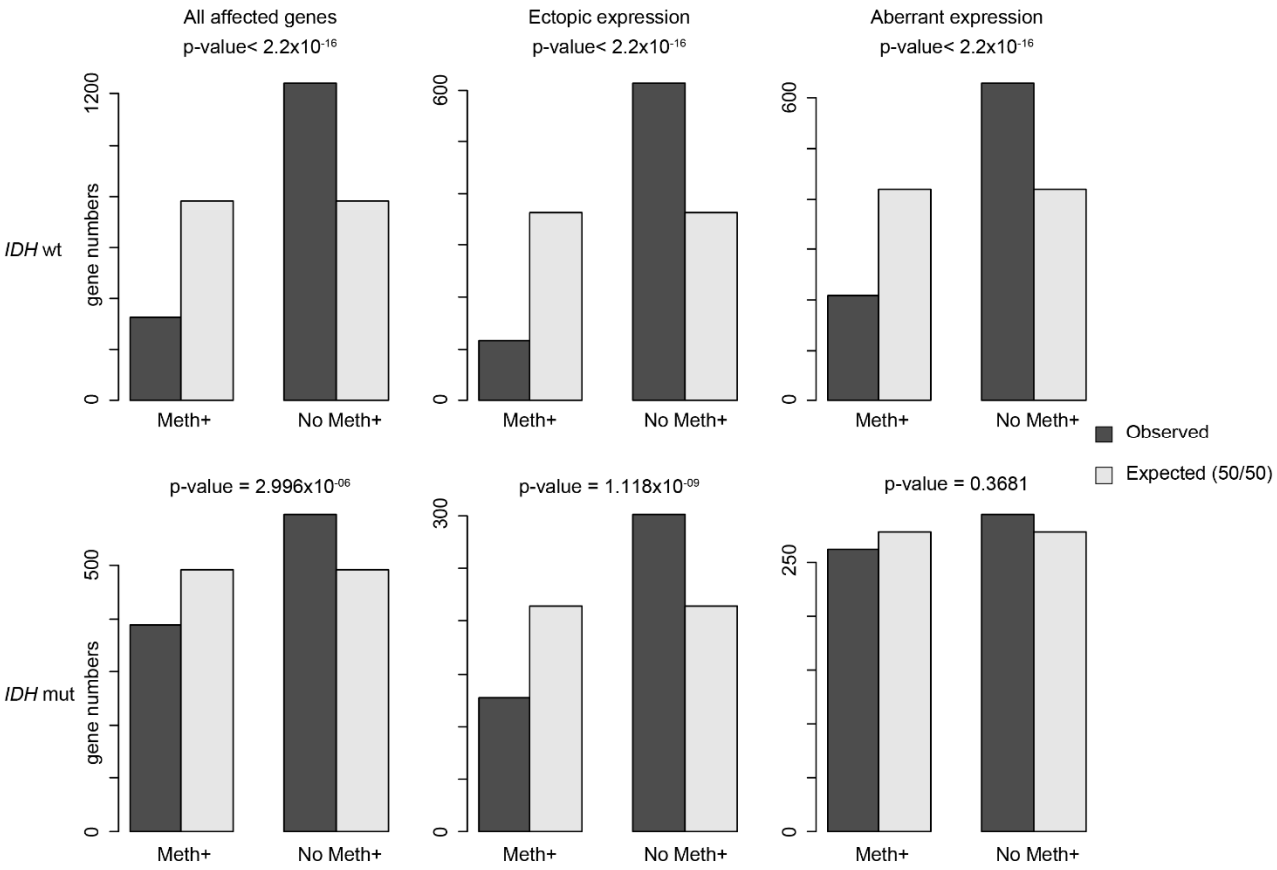



**Figure S4: Most of transcriptional affected genes do no show methylation alteration in their CGI/promoter in *IDH*wt and *IDH*mut samples.**

Fisher's exact test assuming the null hypothesis that gene associated (Meth+) and not associated (No Meth) to DNA methylation alteration in their CGI/promoter are evenly distributed among affected genes.

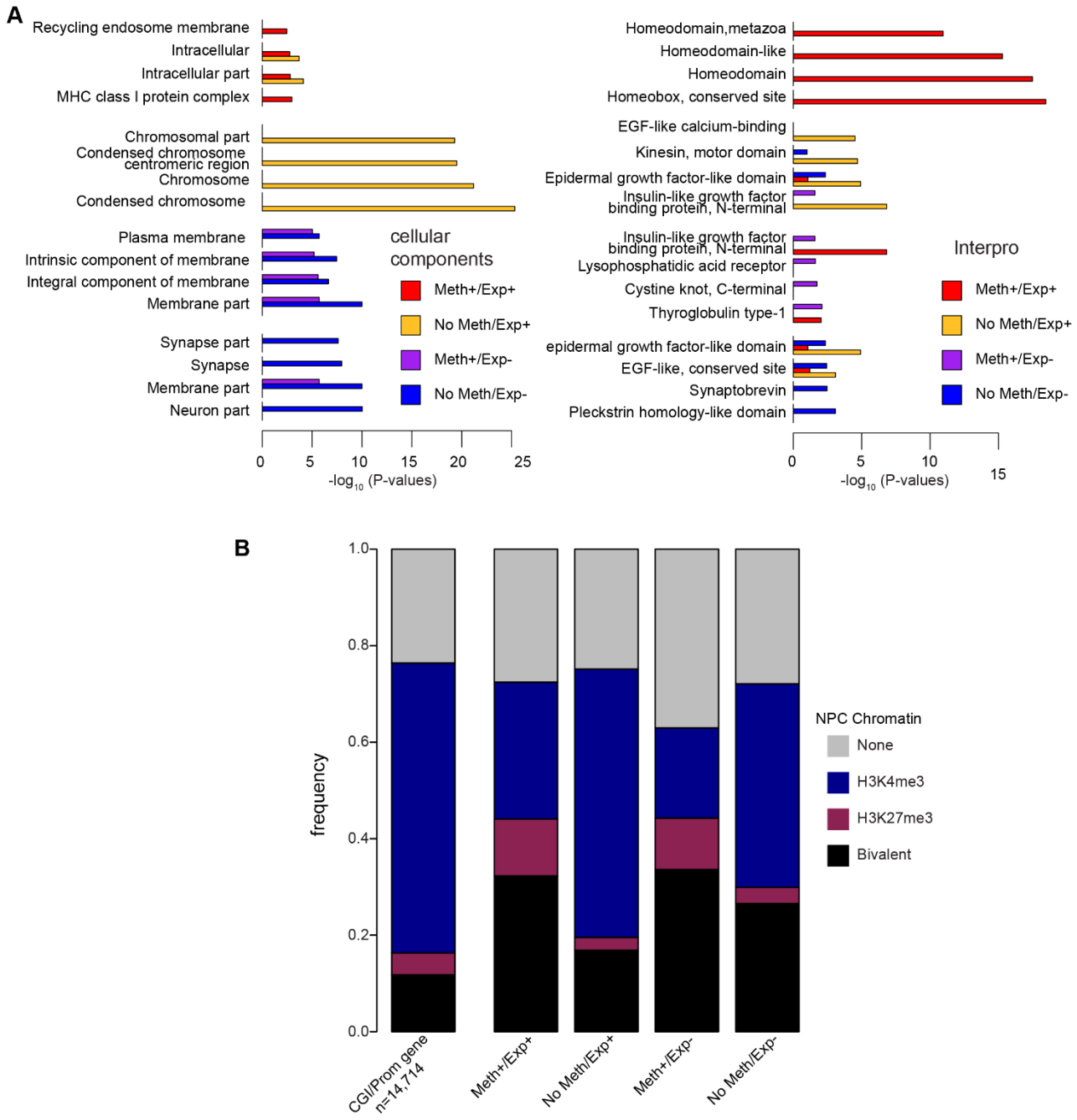



**Figure S5: Gene Ontology**

**A**) Gene Ontology terms enriched in genes belonging to each defect category identified in *IDH*wt glioma samples. For each category, the four highest terms are shown. **B**) Distribution of genes belonging to each defect category according to their chromatin signature in human NPC (none: gray; bivalent: black; H3K4me3-only: blue; H3K27me3-only: red). As reference, the distribution of the 14,714 genes analyzed in this study according to their chromatin signatures in human ES cells is shown in the left panel.

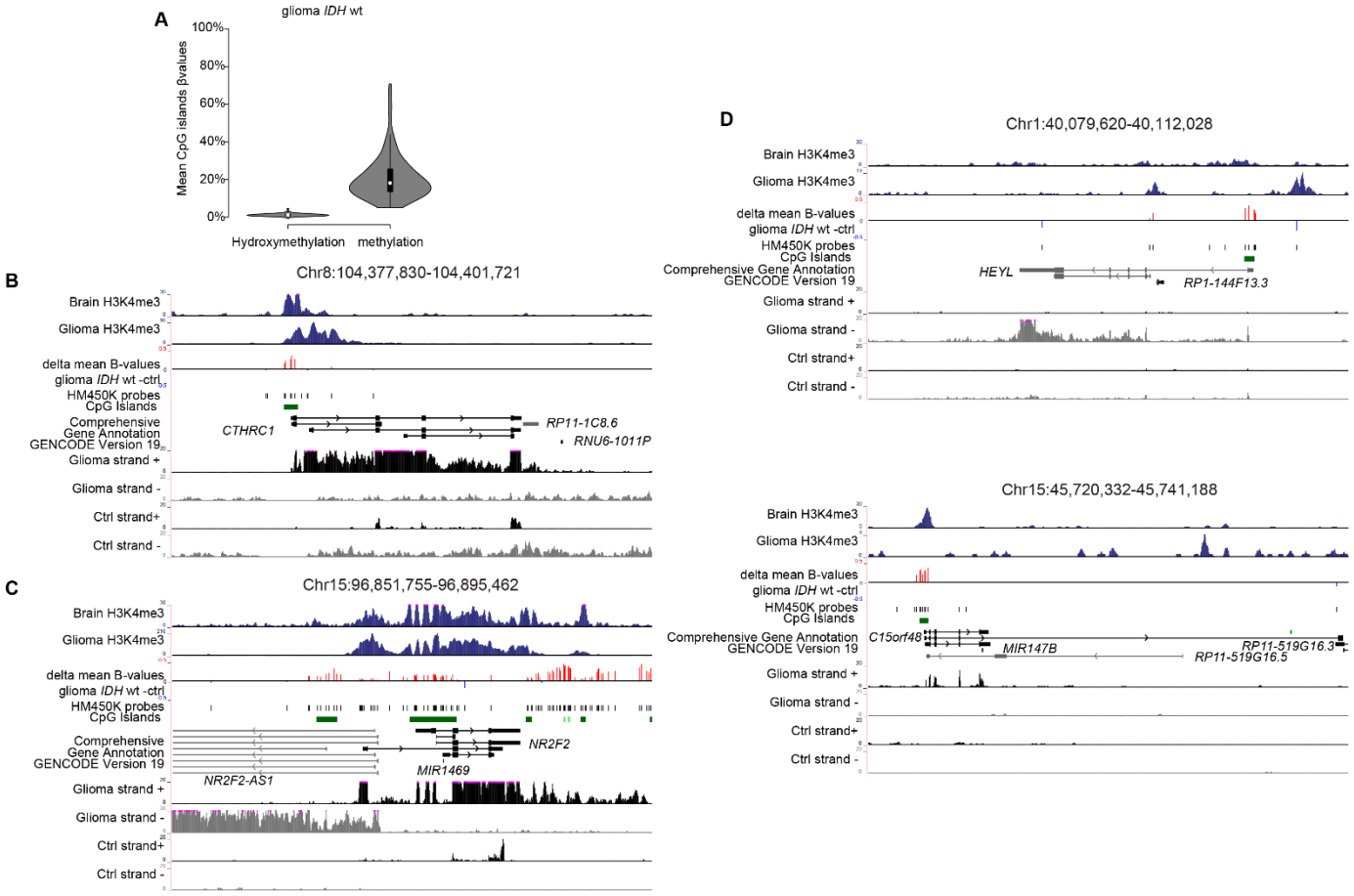



**Figure S6: Expression of genes with a methylated CGI/promoter**

**A)** Levels of hydroxymethylation and methylation, by oxidative bisulfite sequencing, at the CGI/promoter of Meth+/Exp+ genes in an independent cohort of 30 *IDH*wt samples. Data were retrieved from Johnson *et al.*, 2015. **B-C)** Genome browser view showing H3K4me3 enrichment, differential DNA methylation, and oriented RNA-seq signal at **B**) *CTHRC1* (representative example of genes in which transcription initiates from an H3K4me3-marked TSS embedded in a methylated CGI/promoter, **C**) *NR2F2* (example of genes in which transcription initiates from an alternative TSS, and **D**) *HEYL* and *C15orf44*, where transcription initiates from a methylated CGI/promoter.

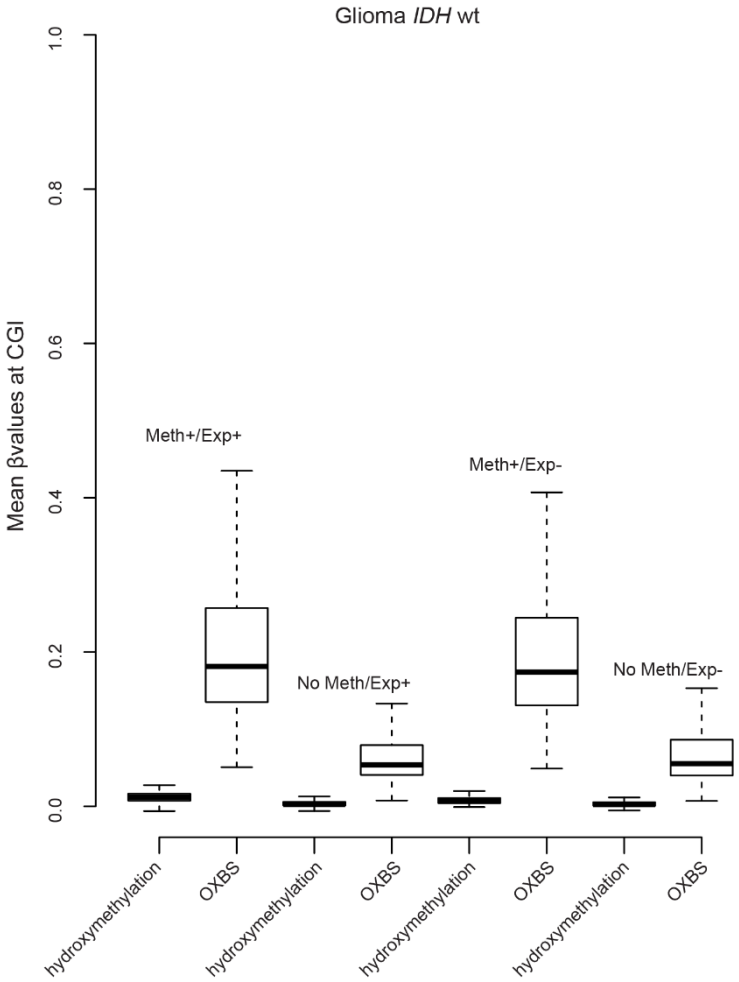



**Figure S7: Affected genes are not enriched in 5-hmC**
Hydroxymethylation and methylation, derived from oxydative bisulfite (OXBS), levels at CGI/promoter of genes for each defect group in an independent cohort of 30 *IDH*wt samples. Data were retrieved from Johnson *et al*., 2015.

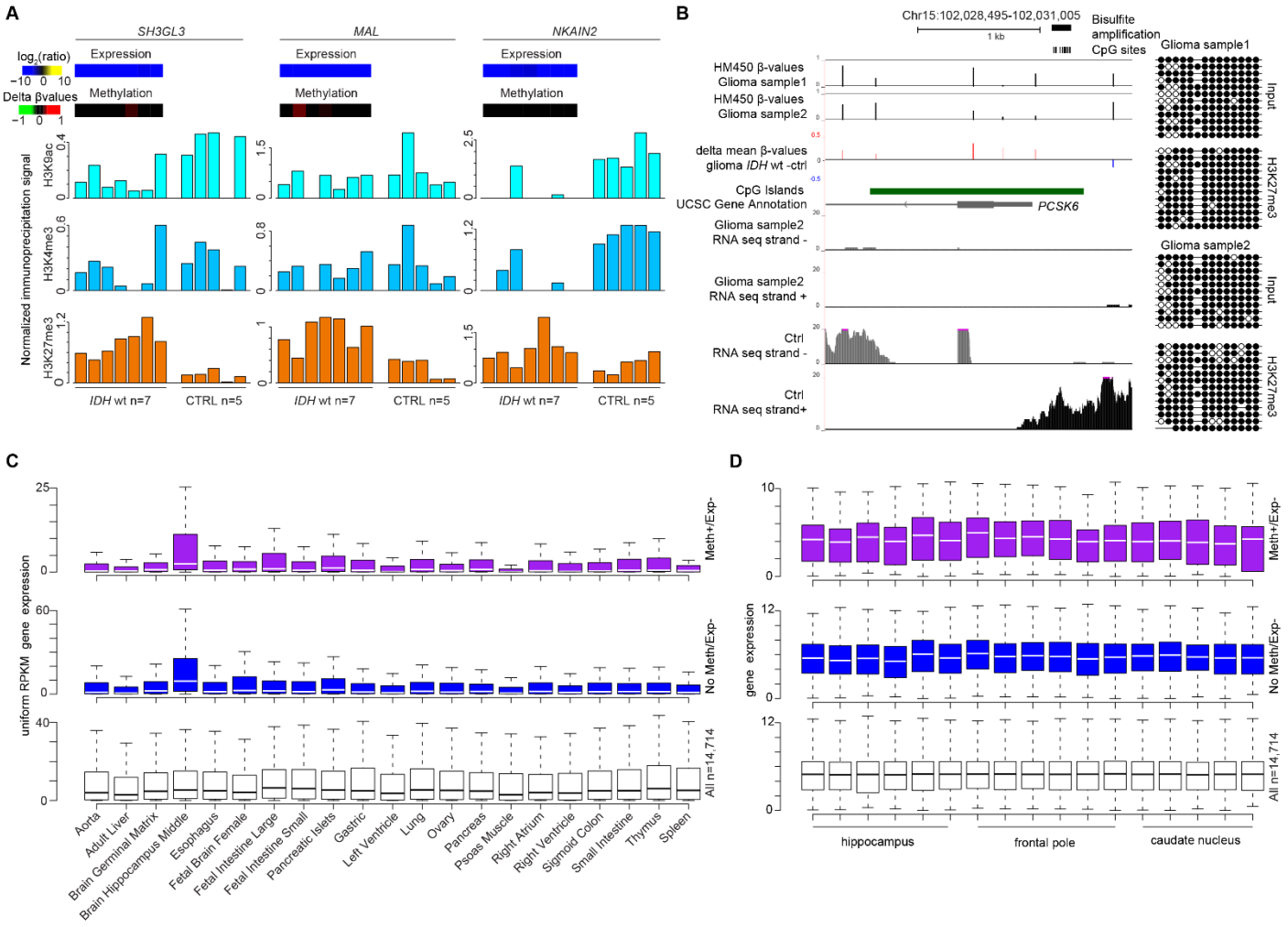



**Figure S8: Gene repression is associated with H3K27me3 gain and affects mainly genes with brain-specific expression profile**

**A)** ChIP analysis of H3K9ac, H3K4me3 and H3K27me3 at selected genes in *IDH*wt glioma (n=7) and control (n=5) samples. The precipitation level was normalized to that obtained at the *TBP* promoter (for H3K4me3 and H3K9ac) and at the *SP6* promoter (for H3K27me3). **B)** Bisulfite-based sequencing data for the *PCSK6* CGI/promoter using input and H3K27me3-immunoprocipated ChIP fractions showed that DNA methylation and H3K27me3 can coexist in glioma cells. Each horizontal row of circles represents the CpG dinucleotides on an individual chromosome. Solid circles, methylated CpG dinucleotides; open circles, unmethylated CpG dinucleotides. The relative position of the bisulfite amplicon is showed on the *PCSK6* locus browser view (right panel). **C)** Median expression of "Meth+/Exp-" (purple column) and "No Meth/Exp-" (blue column) genes as well as of all CGI/promoter genes (white column) in 21 tissues (publicly available normalized RNA-seq data). Genes in the "No Meth/Exp-" group are strongly expressed specifically in adult hippocampus **D**) The expression level in hippocampus is representative of the expression level in other brain parts, as shown for the frontal pole and caudal nucleus.

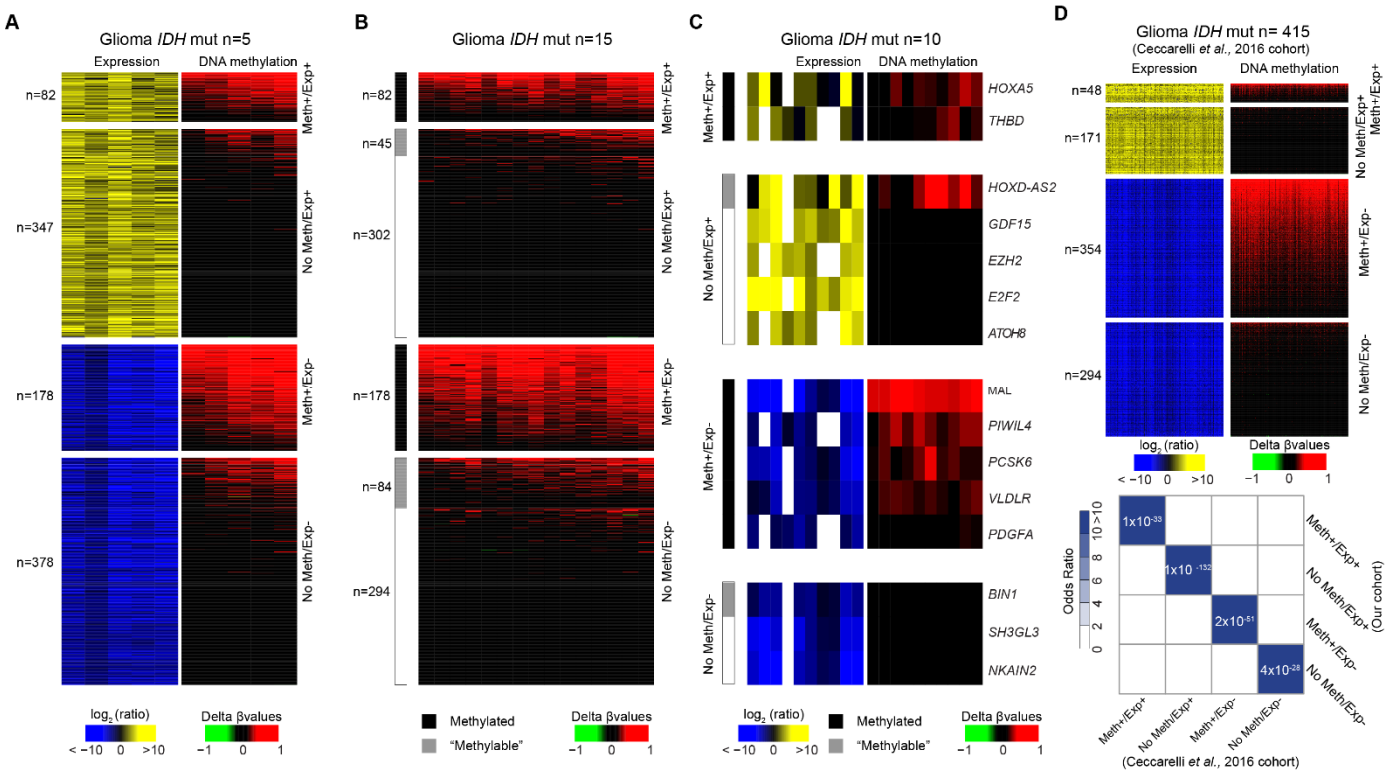



**Figure S9: Four classes of expression defects in IDHmut glioma samples**

**A)** Integrative analysis of gene expression and methylation assessed in 5 *IDH*mut glioma samples identified four main defect classes. **B)** Differential DNA methylation analysis of all the *IDH*mut glioma samples (n=15) classified according the defect classes defined in **A)**. The Methylated and "Methylable" status of the genes is indicated in the left column. **C)** Integrative analysis of the differential expression and methylation (*vs* control) for selected genes from each class defect in 10 *IDH*mut glioma samples. **D)** Integrative analysis of differential gene expression and methylation in an independent cohort of 415 *IDH*mut glioma samples (validation cohort) identified again the four main defect categories. Odds ratio and significance of the overlap (Fisher's exact test) between the data obtained for this validation cohort and our cohort, considered for each defect category, are shown in the panel at the bottom.

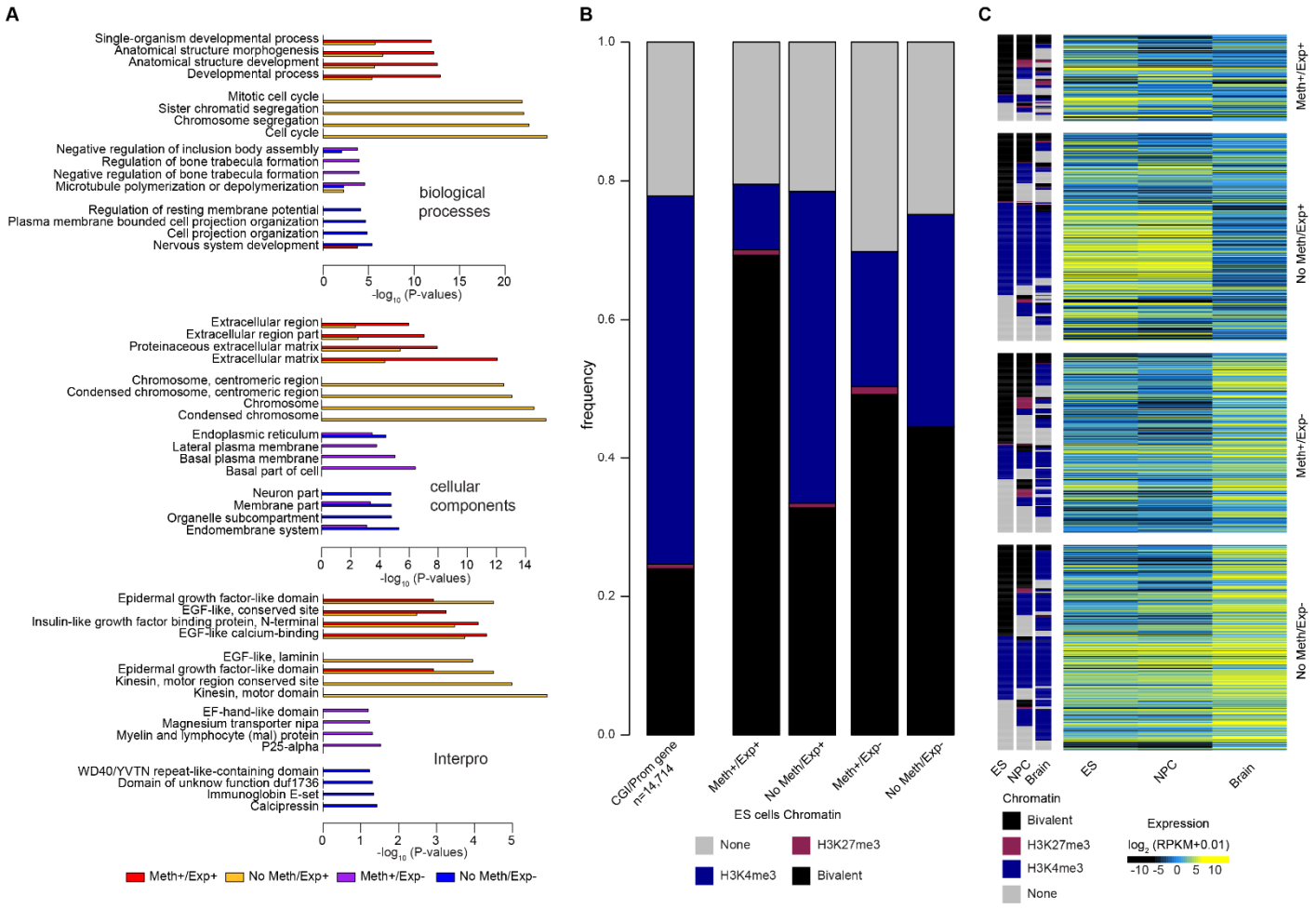



**Figure S10: Genes with bivalent chromatin signature in ES cells are more prone to be deregulated in *IDH*mut glioma.**

**A)** Gene Ontology terms enriched in genes belonging to each defect category identified in *IDH*mut glioma samples. For each category, the four highest terms are shown. **B)** Distribution of genes belonging to each defect category according to their chromatin signatures in human ES cells (none: gray; bivalent: black; H3K4me3-only: blue; H3K27me3-only: red). As a reference, the distribution according to their chromatin signatures in human ES cells of all the 14,714 genes analyzed in this study is shown in the left panel. **c)** Expression level and associated chromatin signatures of genes belonging to each defect category in human ES cells, neural progenitors cells (NPC) and healthy brains.

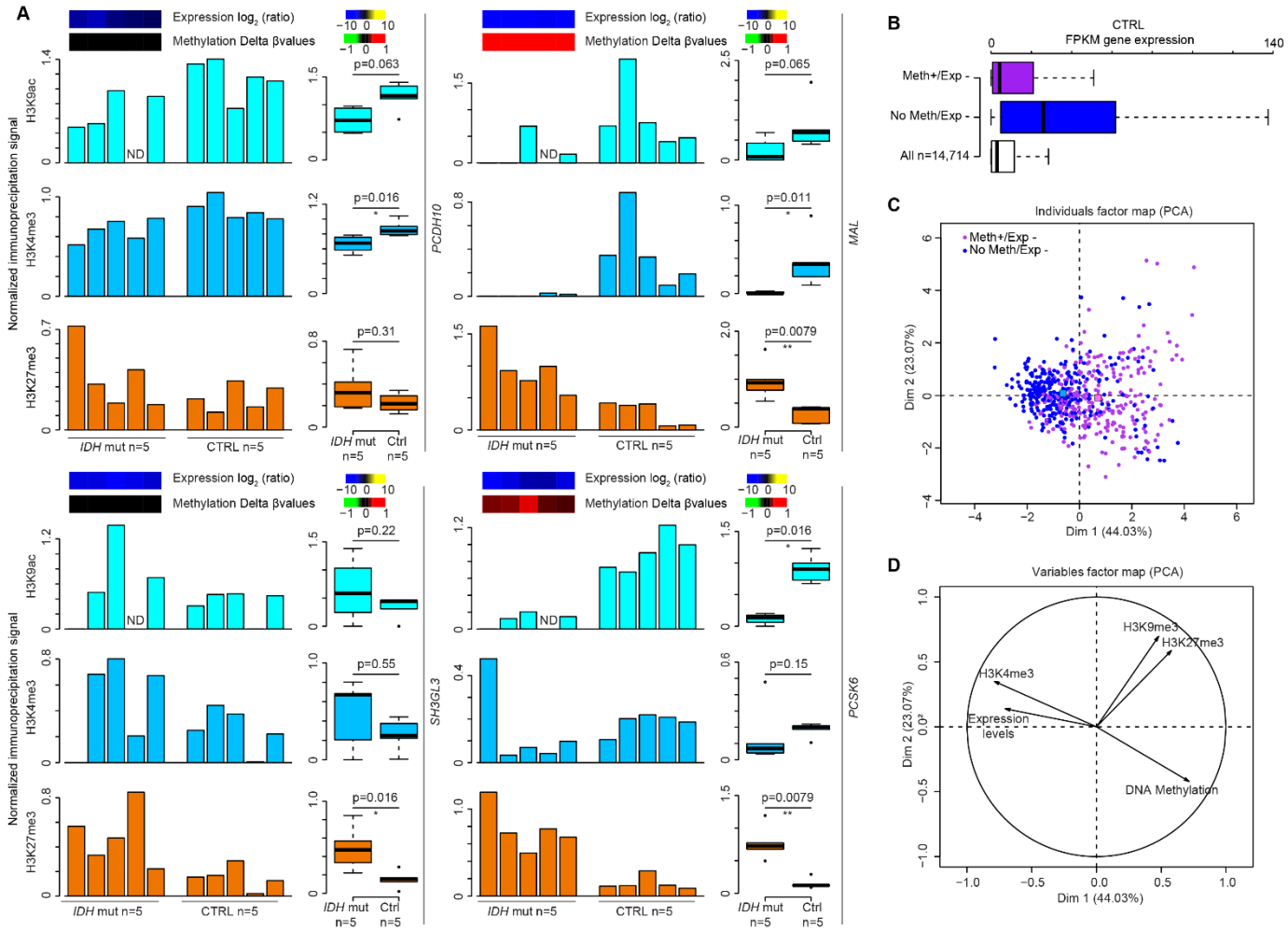



**Figure S11: Gene repression is associated with H3K27m3 gain in *IDH*mut glioma.**
**A)** ChIP analysis of H3K9ac, H3K4me3 and H3K27me3 at selected Meth+/Exp- and No Meth/Exp- genes in IDHmut (n=5) and control (n=5) samples. The precipitation level was normalized to that obtained at the *TBP* promoter (for H3K4me3 and H3K9ac) and at the *SP6* promoter (for H3K27me3) (p-values calculated with the Mann-Whitney *U* test). Data are given individually for each sample (left) or merged (right). **B**) Expression level of "Meth+/Exp-" (purple column) and "No Meth/Exp-" (blue column) genes and of all analyzed genes (white column) in healthy controls. **C-D)** Principal component analysis. Two-dimensional scatter plot of the values of each "Meth+/Exp-" (purple dots) and "No Meth/Exp-" (blue dots) gene along the 1st (Dim 1) and 2nd (Dim 2) principal component C). For each defect category, the centroids are shown by colored squares. **D)** H3K4me3 and gene expression levels in healthy brain are the variables that most contributed to and were significantly correlated with the first principal component.

| | *IDH* wt | | *IDH* mut | |
|---|---|---|---|---|
| | initial analysis | wilcox test & Zhou *et al*., 2016 | initial analysis | wilcox test & Zhou *et al*., 2016 |
| | total 14,714 | total 13,942 | total 14,714 | total 13,942 |
| Meth+/Exp+ | 0.8 | 0.8 | 0.9 | 0.9 |
| No Meth/Exp+ | 4.1 | 4.2 | 2 | 2.1 |
| Meth+/Exp- | 1.4 | 1.6 | 1.7 | 1.9 |
| No Meth/Exp- | 4.2 | 4.3 | 2 | 2.0 |

**Figure S12**: Summary table of data obtained by treating HM450K data with conditions described in material and method section (initial analysis) or by using filter described in Zhou *et al*., 2017 associated to a Wilcoxon-rank sum test (« wilcox test & Zhou *et al*., 2017 »).