

SUPPLEMENTARY INFORMATION

Specific methyl-CpG configurations define cell identity through gene expression regulation.

Teresa Improda ^{1*}, Valentina Morgera ^{1*}, Maria Vitale ¹, Lorenzo Chiariotti ², Fabiana Passaro ², Antonia Feola ¹, Antonio Porcellini ¹, Mariella Cuomo ^{2✉}, Antonio Pezone ^{1✉}

¹Dipartimento di Biologia, Complesso Universitario di Monte Sant'Angelo, Università di Napoli "Federico II", Napoli 80126, Italy;

²Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Napoli 80131, Italy;

*These authors equally contributed to this work.

✉Correspondence: antonio.pezone@unina.it and mariella.cuomo@unina.it

Supplementary Figures

Figure S1. Both-strand methylation status on the BDNF promoter during stem cell differentiation.

(A) Average methylation of the six CpGs of the BDNF during stem cell differentiation and the relative random control at different time points (T0, T2, T4, T8 and T14, color-coded squares on the right). (B) Average methylation of each CpG identified by the nucleotide position. (C) Taxonomy plot of epialleles derived from each population. (D) Shannon entropy of the methylated molecules in the same samples and time points shown in (A). (E) Comparison of the Average methylation of each CpG identified by the nucleotide position shown in (B). (F) Pearson correlation of all analyzed time points. A pairwise comparison between each pair was performed with the student's t-test: * $P < 0.05$ versus T0.

Figure S2. Methylation signatures of BDNF promoters during stem cells differentiation.

(A) PCA of BDNF epialleles during stem cell differentiation and the relative random control at different time points (T0, T2, T4, T8 and T14, color-coded squares on the right). Dendrogram and hierarchical clustering of the frequency (B) and the structure (C) of BDNF epialleles. The similarity of different samples is represented by the horizontal distances on each branch of the dendrogram. 'vegan' R package (version 2.4-1) was used to perform the tests and draw correlograms. (D) Frequency of the methylated cores (MethCores index) in the same population analyzed in Fig. S1A. (E) Frequency of the methylated cores normalized to the average methylation (clonality index) in each population. (F) The structure and the composition of the methylated cores at different time points in the populations of molecules derived from each strand. Each CpG is labeled with a color code on the panel's right side. A pairwise comparison between each pair was performed with the student's t-test: * $P < 0.05$ versus T0.

Figure S3. DDO gene's methylation signatures during stem cell differentiation at region R1.

(A) Structure of the putative mouse *DDO* gene promoter. The blue arrow indicates the transcription start site (+1). Positions of CpG sites are indicated as relative to TSS. The analyzed 3 kb genomic portion is divided in 7 regions (R1-R2-R3-R4-R5-R6-R7) corresponding to different amplicons. (B) Average methylation of the 3 CpGs present in R1 shown in (A) and the relative random control at different time points (P0 and P30, color-coded squares on the right). (C) Average methylation of each CpG identified by the nucleotide position shown in (A), in the populations of molecules shown in (B). (D) Shannon entropy of the methylated molecules in the same samples and time points shown in (B). A pairwise comparison between each pair was performed with the student's t-test: * $P < 0.05$ versus P0. (E) The methylated cores' structure and composition at different times. Each CpG is labeled with a color code on the panel's right side.

Figure S4. DDO gene's methylation signatures during stem cell differentiation at regions R2 and R3.

(A) Average methylation of the 4 CpGs present in R2 shown in (Fig. S3A) and the relative random control at different time points (P0 and P30, color-coded squares on the right). (B) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (A). (C) Shannon entropy of the methylated molecules in the same samples and time points shown in (A). (D) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. (E) Average methylation of the 9 CpGs present in R3 shown in (Fig. S3A) and the relative random control at different time points (P0 and P30, color-coded squares on the right). (F) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (E). (G) Shannon entropy of the methylated molecules in the same samples and time points shown in (E). (H) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. A pairwise comparison between each pair was performed with the student's t-test: * $P < 0.05$ versus P0.

Figure S5. DDO gene's methylation signatures during stem cell differentiation at regions R4 and R5.

(A) Average methylation of the 6 CpGs present in R4 shown in (Fig. S3A) and the relative random control at different time points (E15, P0, P14 and P30, color-coded squares on the right). (B) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (A). (C) Shannon entropy of the methylated molecules in the same samples and time points shown in (A). (D) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. (E) Average methylation of the 6 CpGs present in R5 shown in (Fig. S3A) and the relative random control at different time points (E15, P0 and P30, color-coded squares on the right). (F) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (E). (G) Shannon entropy of the methylated molecules in the same samples and time points shown in (E). (H) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. A pairwise comparison between each pair was performed with the student's t-test: * $P < 0.05$ versus P0.

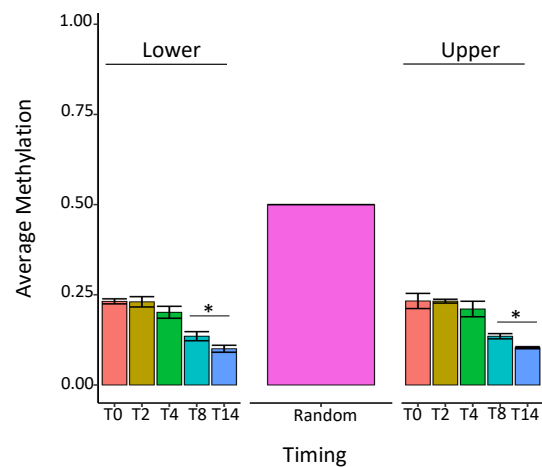
Figure S6. DDO gene's methylation signatures (cores) during stem cell differentiation at regions R6 and R7.

(A) Average methylation of the 6 CpGs present in R6 shown in (Fig. S3A) and the relative random control at different time points (E15, P0 and P30, color-coded squares on the right). (B) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (A). (C) Shannon entropy of the methylated molecules in the same samples and time points shown in (A). (D) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. (E) Average methylation of the 9 CpGs present in R7 shown in (Fig. S3A) and the relative random control at different time points (P0 and P30, color-coded squares on the right). (F) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A), in the populations of molecules shown in (E). (G) Shannon entropy of the methylated molecules in the same samples and time points shown in (E). (H) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. A pairwise comparison between each pair was performed with the student's t-test: *P < 0.05 versus P0.

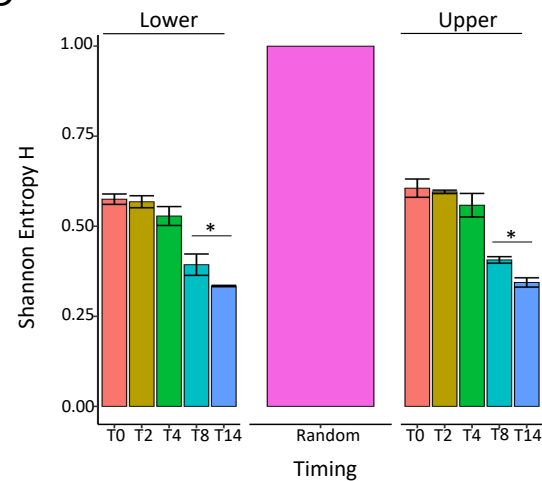
Figure S7. The methylation status of DDO promoter in different tissues and cell lines. (A) The average methylation of the six CpGs shown in (Fig. S3A) of in the DNA extracted from brain areas, several cell types isolated from the same brain areas, immortalized A1 (neuronal) cells, c-myc immortalized neurons at different passages and ESCs upon induction of neural differentiation and the relative random control (B) Average methylation of each CpG identified by the nucleotide position shown in (Fig. S3A) in the populations of molecules shown in (A). (C) Shannon entropy of molecules in the same samples and time points shown in (A). (D) The methylated cores' structure and composition at different time points. Each CpG is labeled with a color code on the panel's right side. A pairwise comparison between each pair was performed with the student's t-test: *P < 0.05 versus T0.

Fig. S1

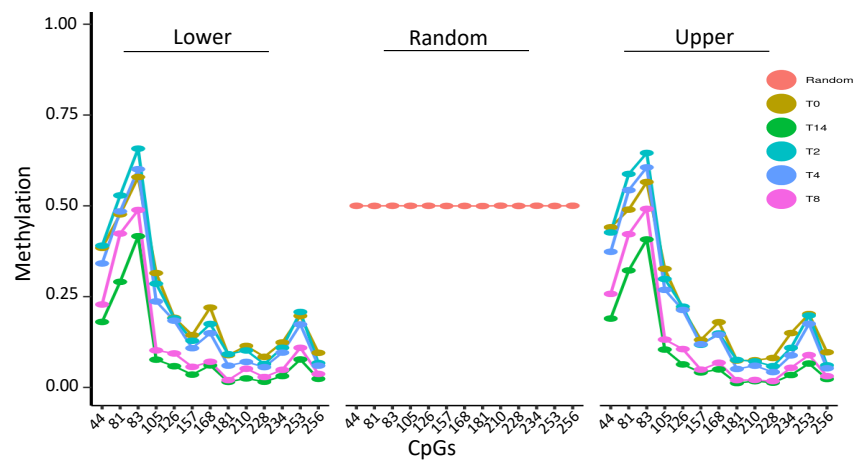
A



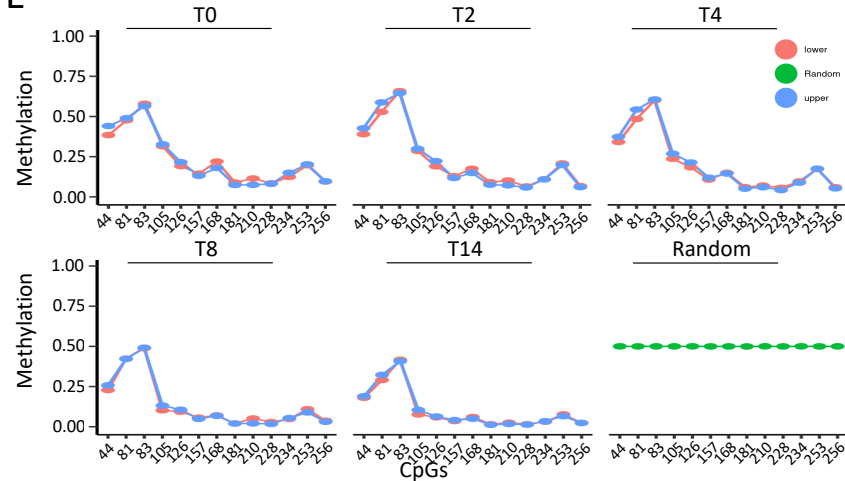
D



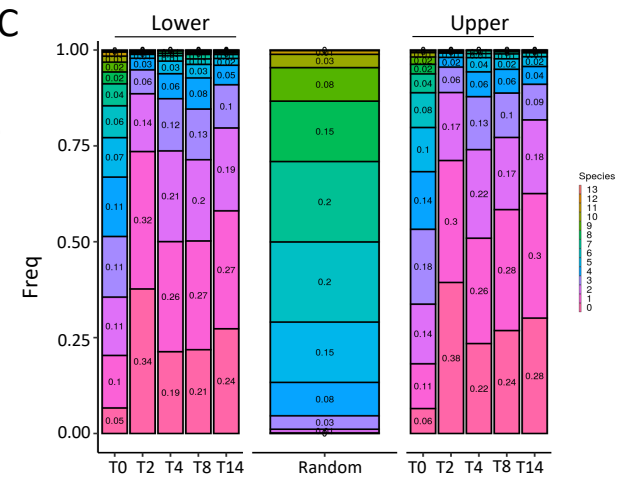
B



E



C



F

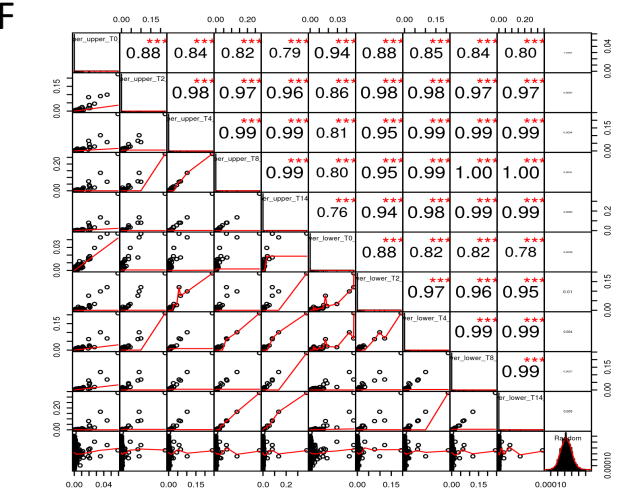


Fig. S2

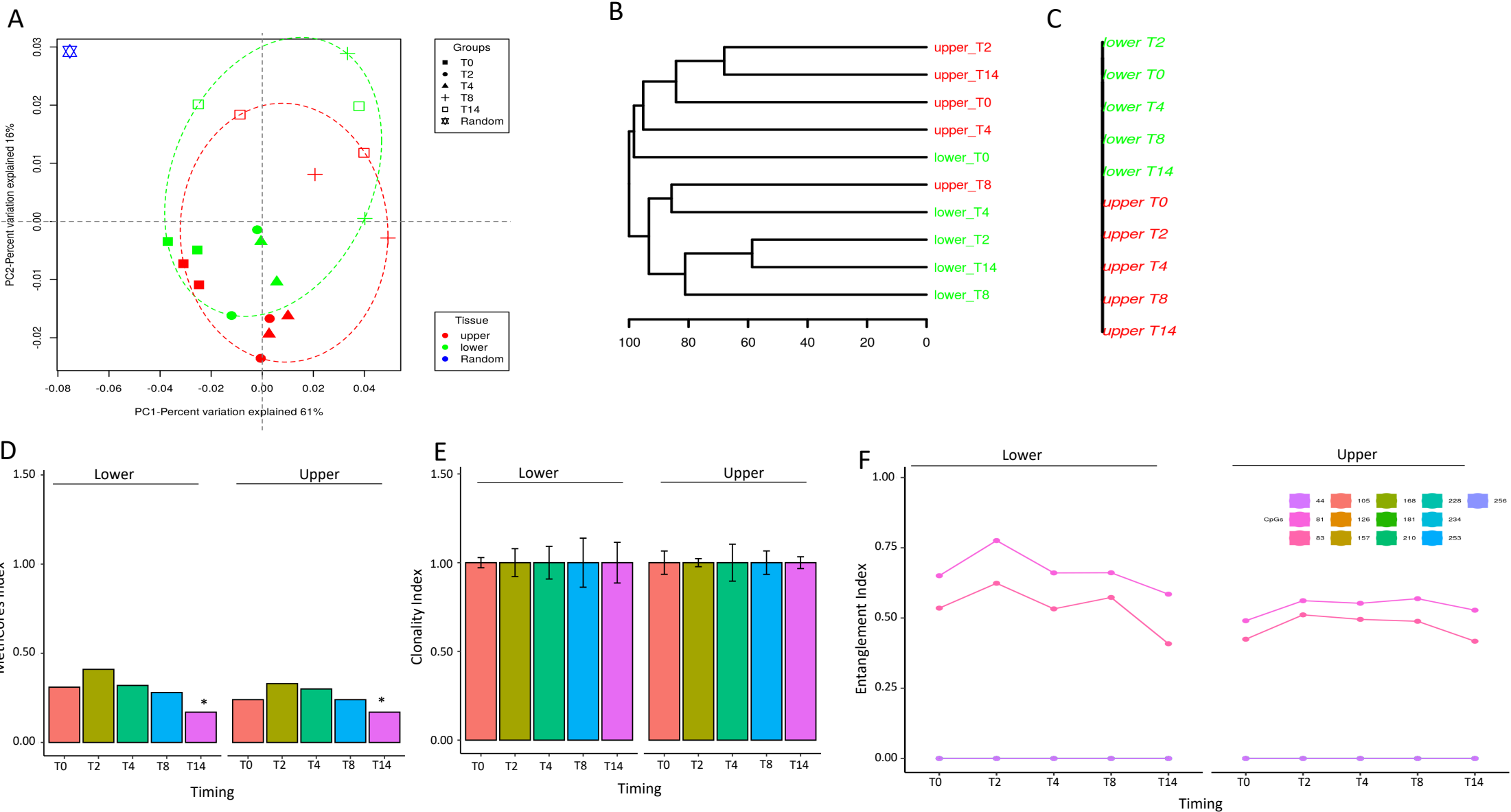
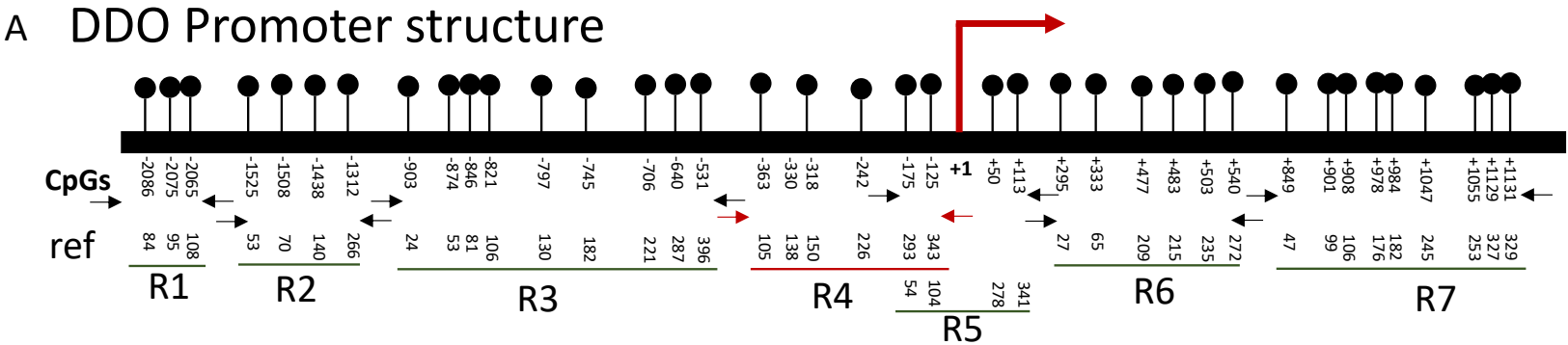


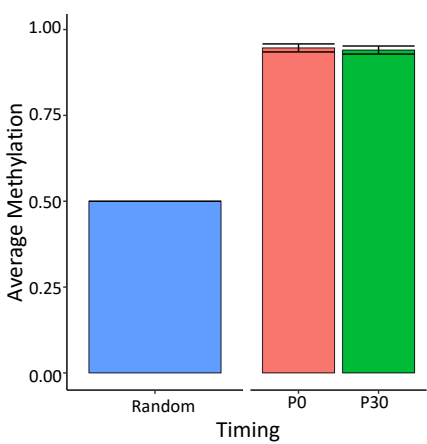
Fig. S3

A DDO Promoter structure

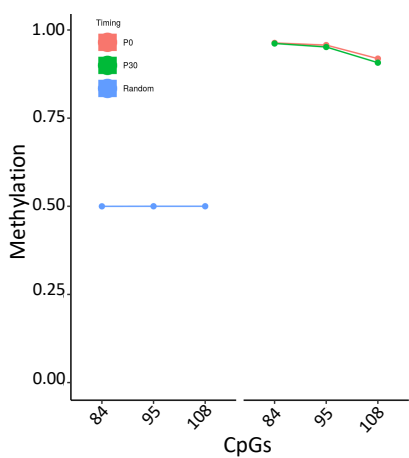


DDO R1

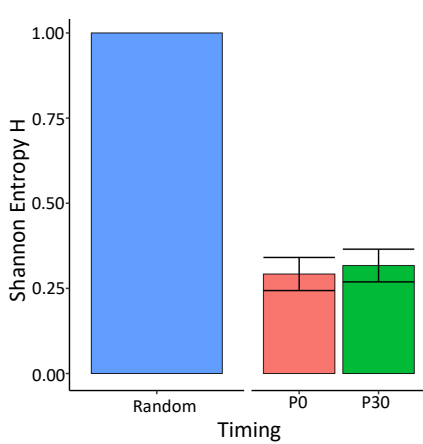
B



C



D



E

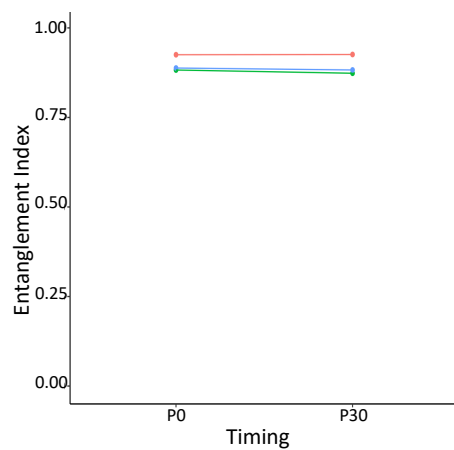


Fig. S4

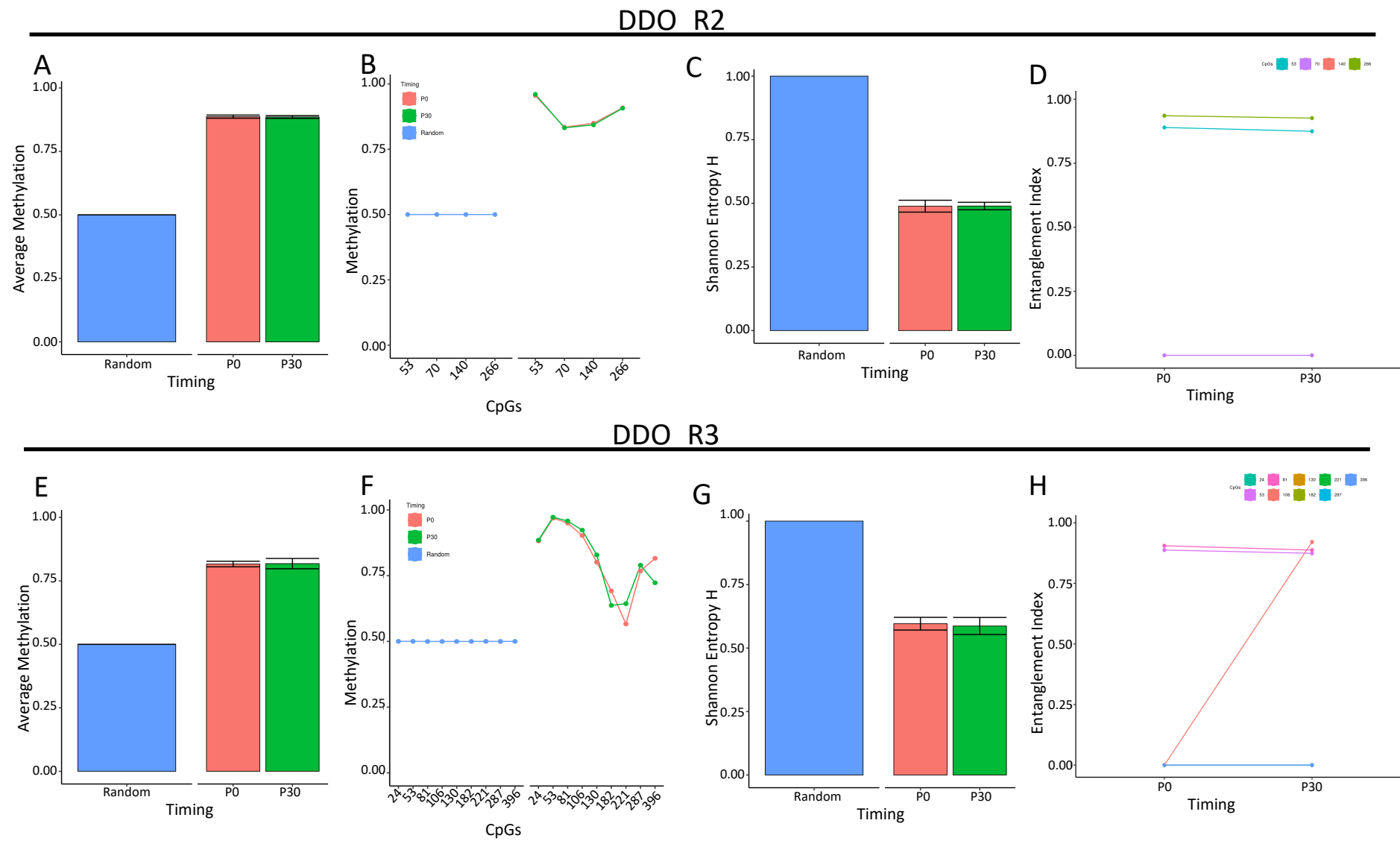


Fig. S5

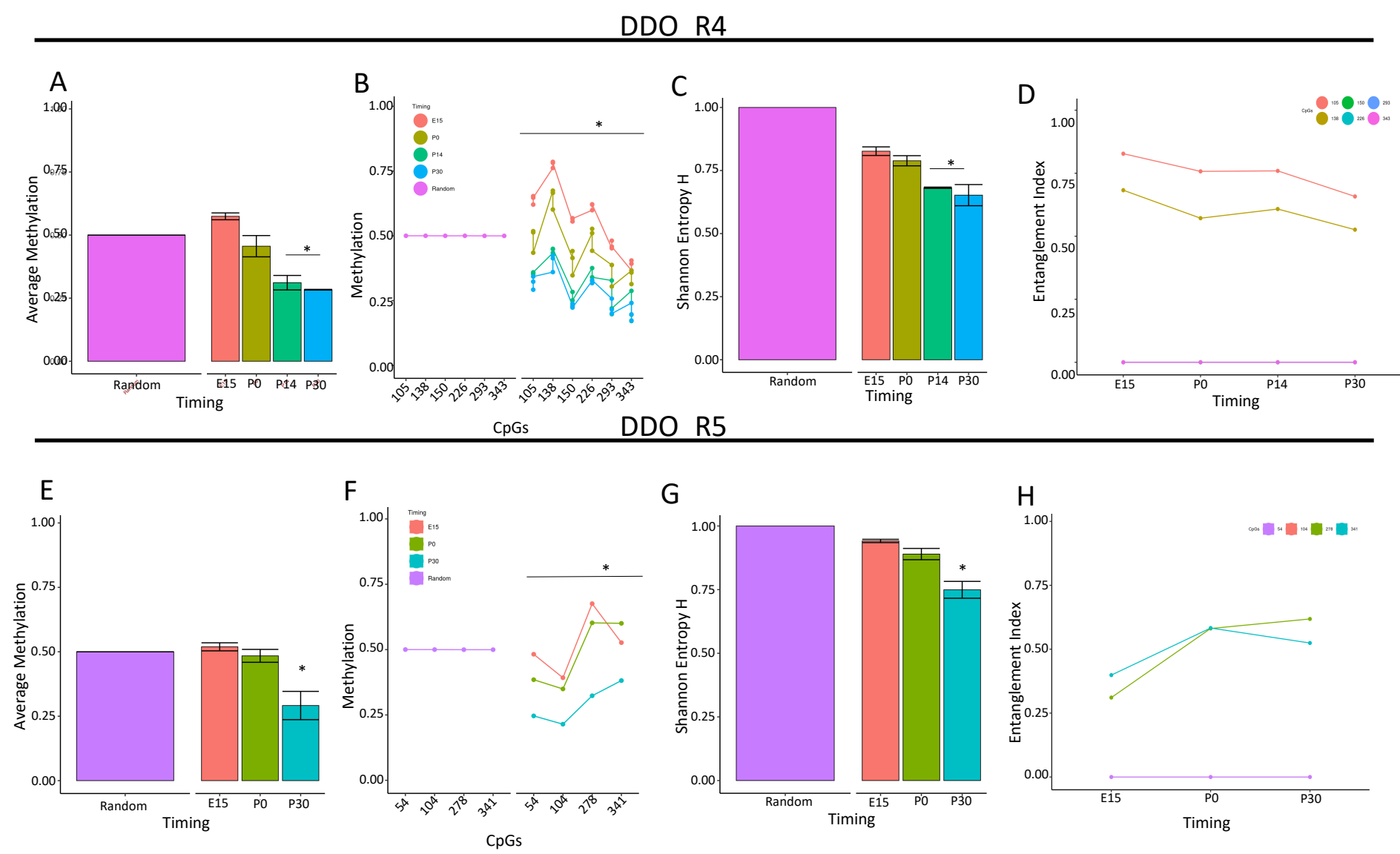


Fig. S6

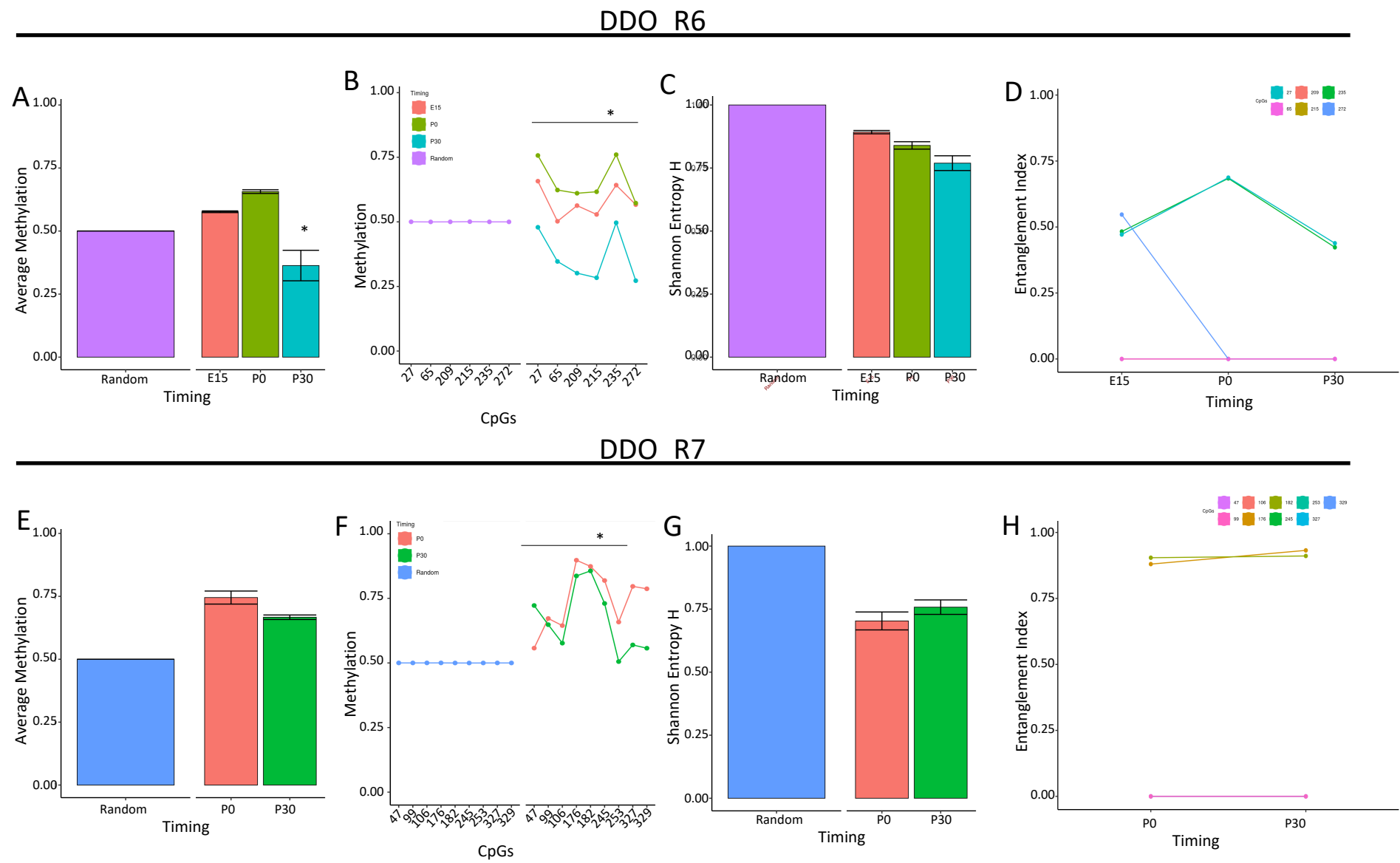


Fig. S7

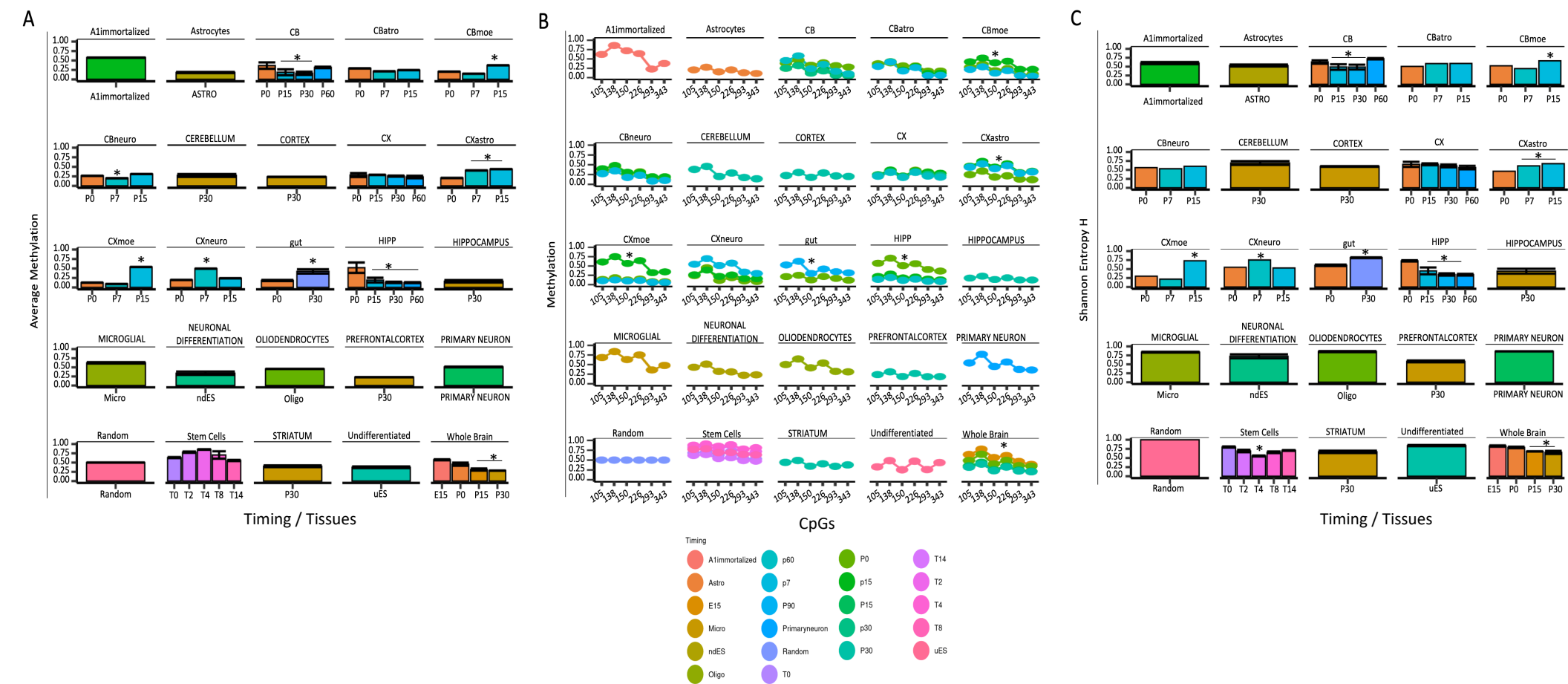


Table S1

Analysis	Foward	Reverse	Gene
Bisulfite	5'-AAAGAATGGAAGAGGAAATTTAGA-3'	5'-TTAGATAGGTTGATTGTTGGTG-3'	NANOG
Bisulfite	5'-AGAAAATAGTGATTTTAGAGTT-3'	5'-GTGTTGATGTTGTAGAGTTTGT-3'	TUBB3
Bisulfite	5'-GATAGGTAAGTAATTTATGGA-3'	5'-GATAATTGGGTATTATGTTA-3'	GFAP
Bisulfite	5'-GGATTGGAAGTGAAAATATTTAT-3'	5'-AGTTTTTGTTAGATTAAATGGA-3'	BDNF plus
Bisulfite	5'-AATAACAACTTAACTTCTATATA-3'	5'-GGTGATAGTTATTAAATTGAAG-3'	BDNF minus
Bisulfite	5'-TAAATATTTATAGAAGGGAATTGG-3'	5'-AAACAAACACTATAACATAAACC-3'	DDO
RNA	5'-CAAGGGTCTGCTACTGAGATGCTCTG-3'	5'-TTTTGTTTGGGACTGGTAGAAGAATCAG-3'	NANOG
RNA	5'-TTGGAACCTGGAACCATGGACAG-3'	5'-CTGTGCCCCCAGCCAGTGA-3'	TUBB3
RNA	5'-GTCAGAAGGCCACCTCAAGAGG-3'	5'-TTGGGCAGCACACAGGCAGAA-3'	GFAP
RNA	5'-ACCCCTCACCTAAATCCTG-3'	5'-TGCACGAATTACCAGAATCAG-3'	BDNF
RNA	5'-ACCACCAGTAATGTAGCGGC-3'	5'-GGTACCGGGGTATCTGCAC-3'	DDO
RNA	5'-TTCACCACCATGGAGAAGGCT-3'	5' – ACAGCCTTGGCAGCGCCAGT3'	GAPDH