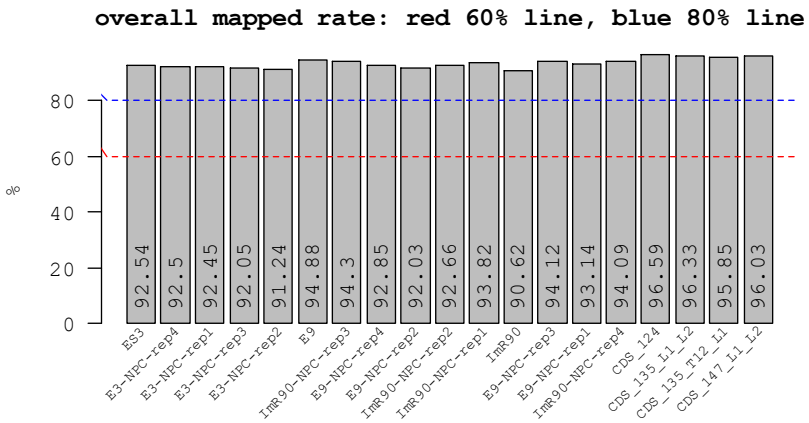
Supplementary Figures

Contents

Figure S13 Data quality for the 19 RNA-seq profiles. **A**, Mapping rates vary 90%~96%. **B**, the 19 profiles have ~ 80% of their reads uniquely map to the human genome. Samples prefixed with 'CDS' are *in vivo* human NP.

A



B

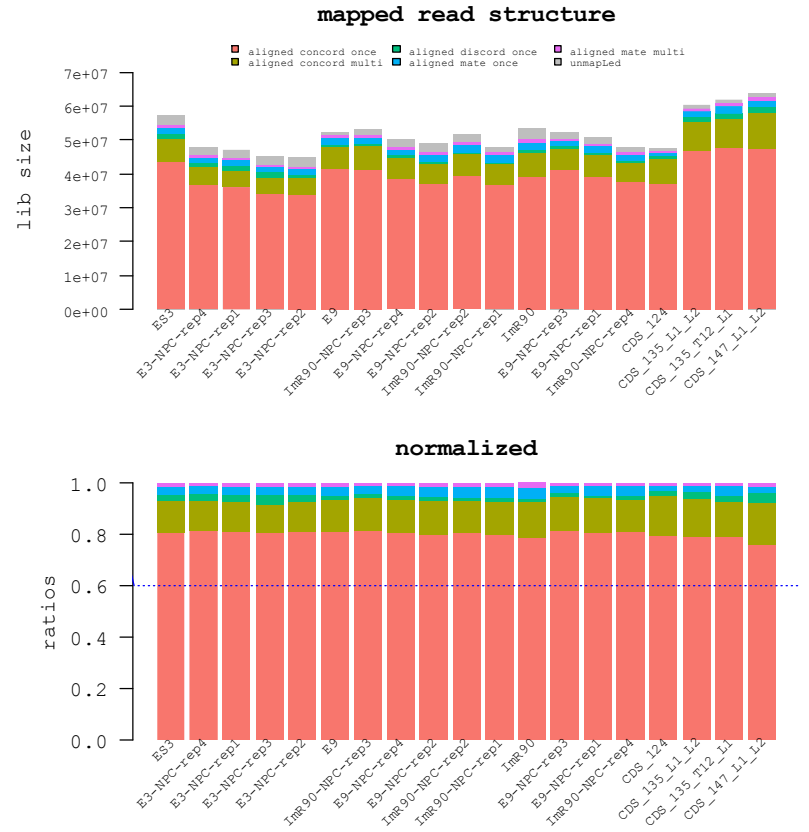


Figure S14 About three thousand differentially expressed genes (DEGs) were detected between the before and after differentiated samples, as shown in this volcano plot. Log2(FC)=log2(after-differentiation/before-differentiation). NPC: *in vitro* NPC differentiations; hESCs: the three stem cell lines.

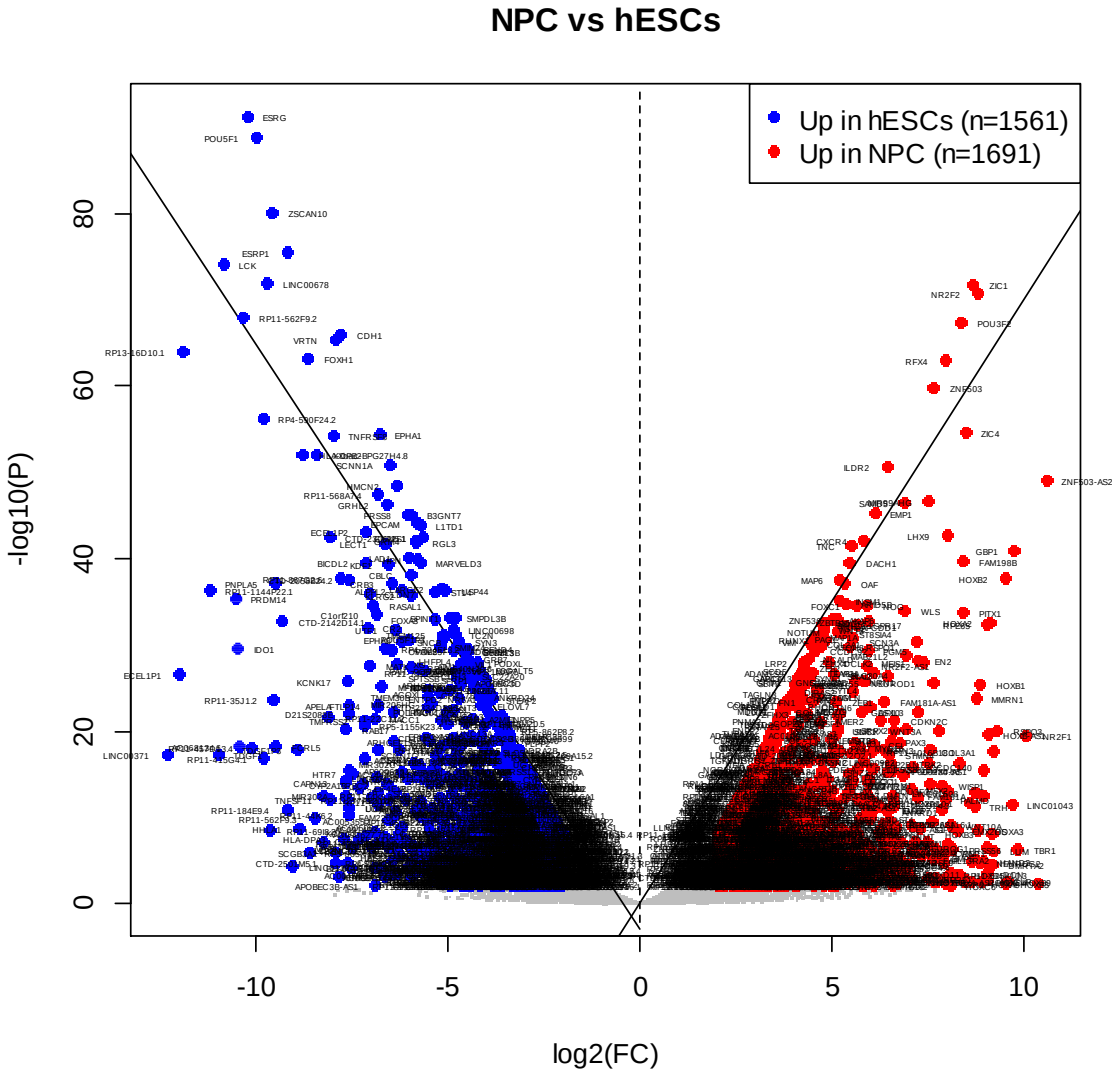


Figure S10 Go term and pathway enrichment analyses show matrisome (the genes of ECMs), and its regulators among the top most significant terms, for the 1691 genes up-regulated in our NPC differentiation. Size of circle indicates –log10(p). Complete list of significant terms and respective overlapping genes can be found in supplementary file g.

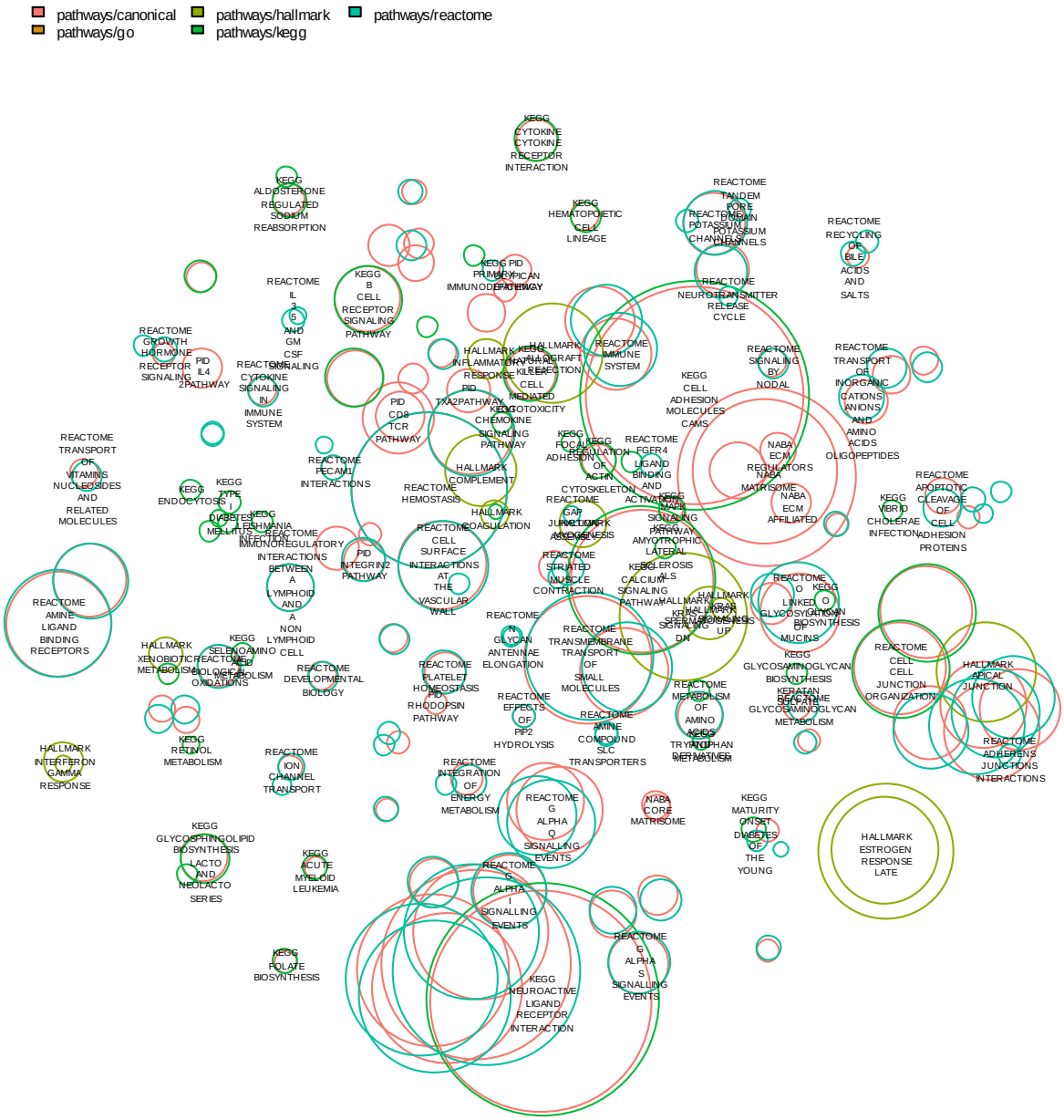


Figure S15, differentiated products of the three stem cell lines have similar responses and up-regulated genes. A, this Venn-diagram shows that 87.8% (1486 out of 1691) of the DEGs up in the NPC differentiations are not varying among the different lines. The lines do have a few hundred DEGs, but B-D show that these genes are only differences in the before differentiation lines.

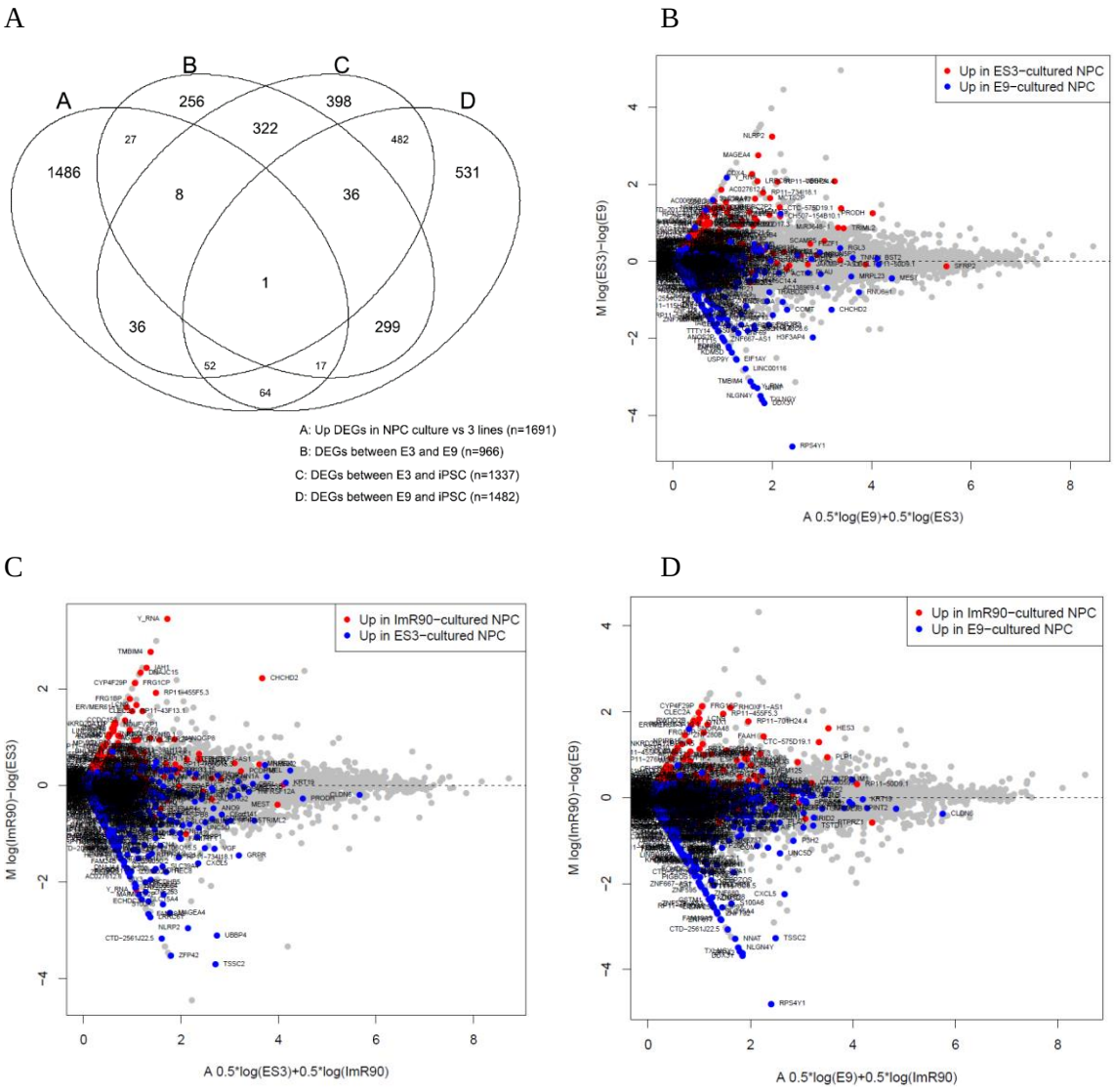


Figure S6, Core matrisome genes that are strongly (log2(fold-change)>2) and significantly (FDR<0.05) highly expressed in our NPC differentiations. Blue: lowly expressed, white: middle, red: highly expressed. Gene-wise standardization performed as in **Figure 3A,G**.

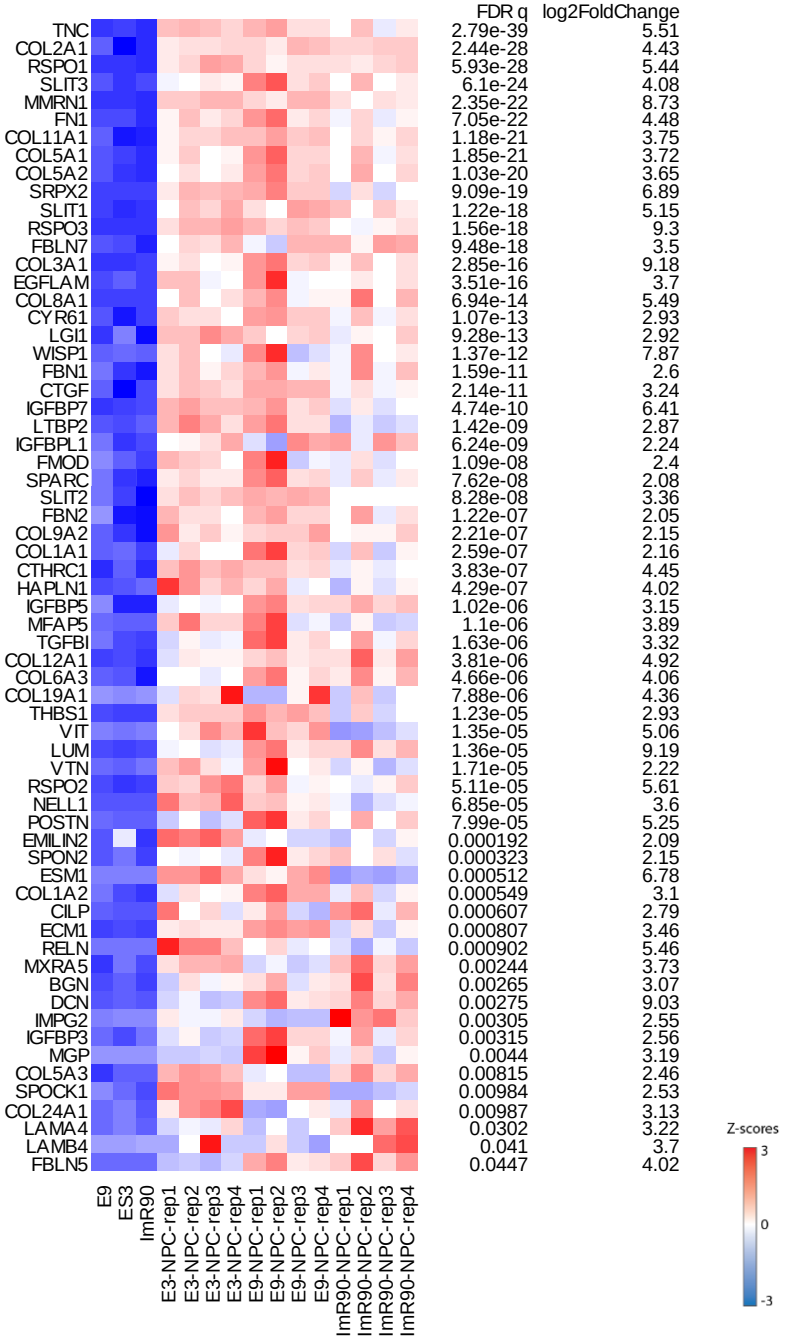


Figure S7, Core matrisome genes that are significantly highly (FDR<0.05) but less strongly (log2(fold-change)>0 and <2) expressed in our NPC differentiations.

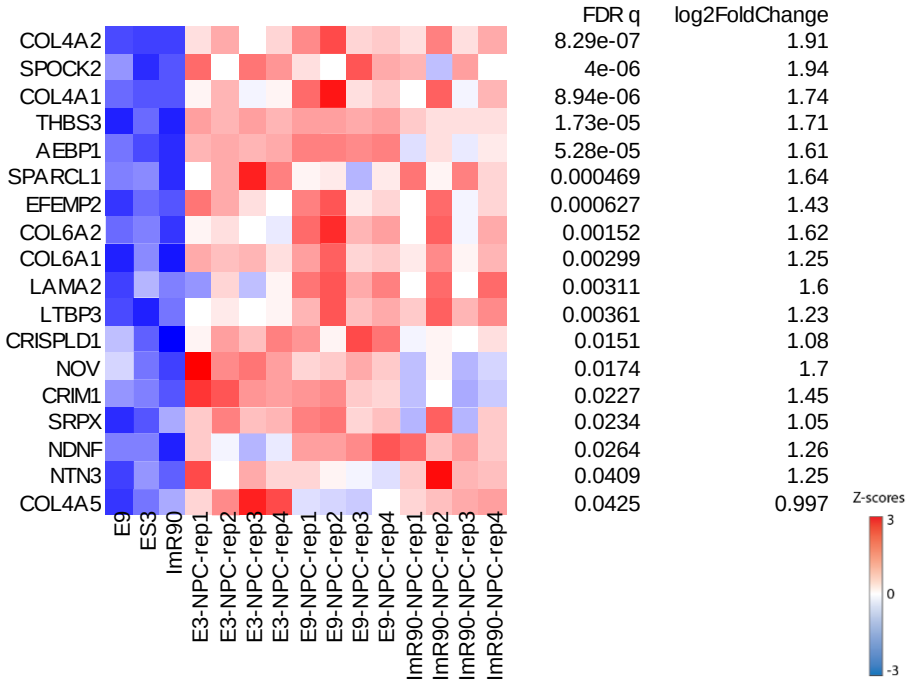


Figure S8, non-Core matrisome genes that are strongly (log2(fold-change)>2) and significantly (FDR<0.05) highly expressed in our NPC differentiations. Color-codes same as **Figure S7**.

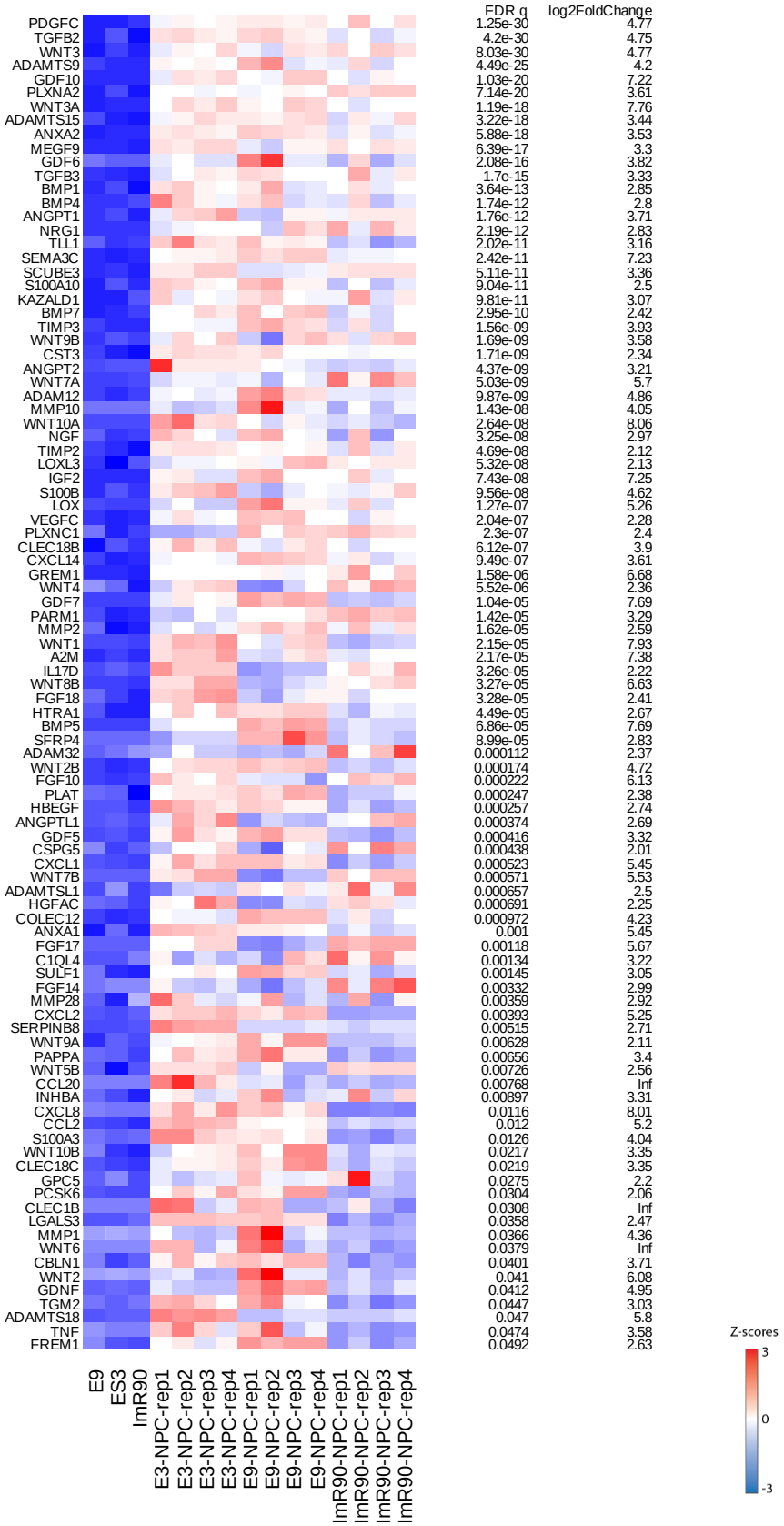


Figure S9, non-Core matrisome genes that are significantly highly (FDR<0.05) but less strongly (log2(fold-change)>0 and <2) expressed in our NPC differentiations.

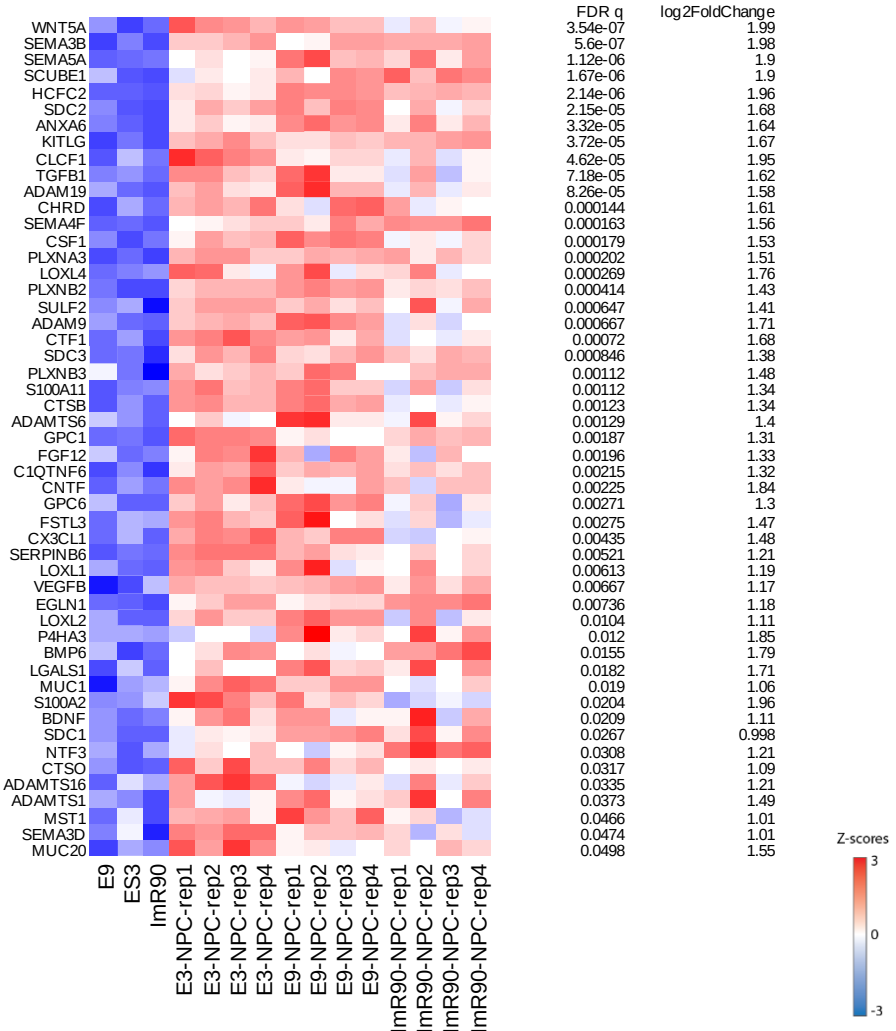


Figure S11, cross-comparisons of *in vitro* NPCs with an in-house set of four adolescent or young *in vivo* NP. Similarity was measure based on genome-wide Pearson correlation coefficients.

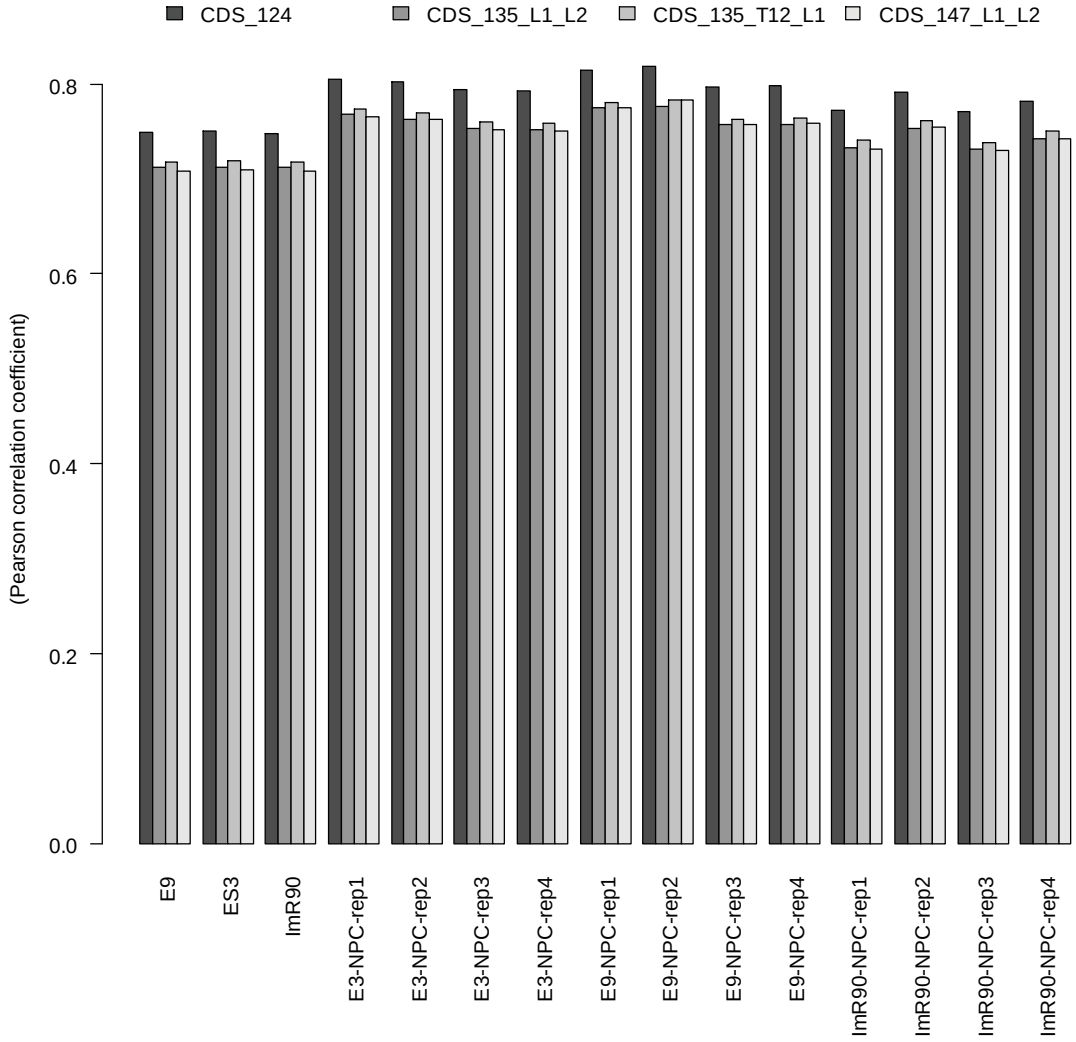


Figure S12, A, cross-comparisons of *in vitro* NPC differentiations with a set of three 'healthy' NP samples in a published data (microarray, see Methods). The three individuals are 20~30 years older than the first *in vivo* data used for cross-comparisons (Figure 3B-F; Methods). Similarity was measure based on genome-wide Pearson correlation coefficients. The Student's *t*-test *p*-value is 8.4×10^{-13}.

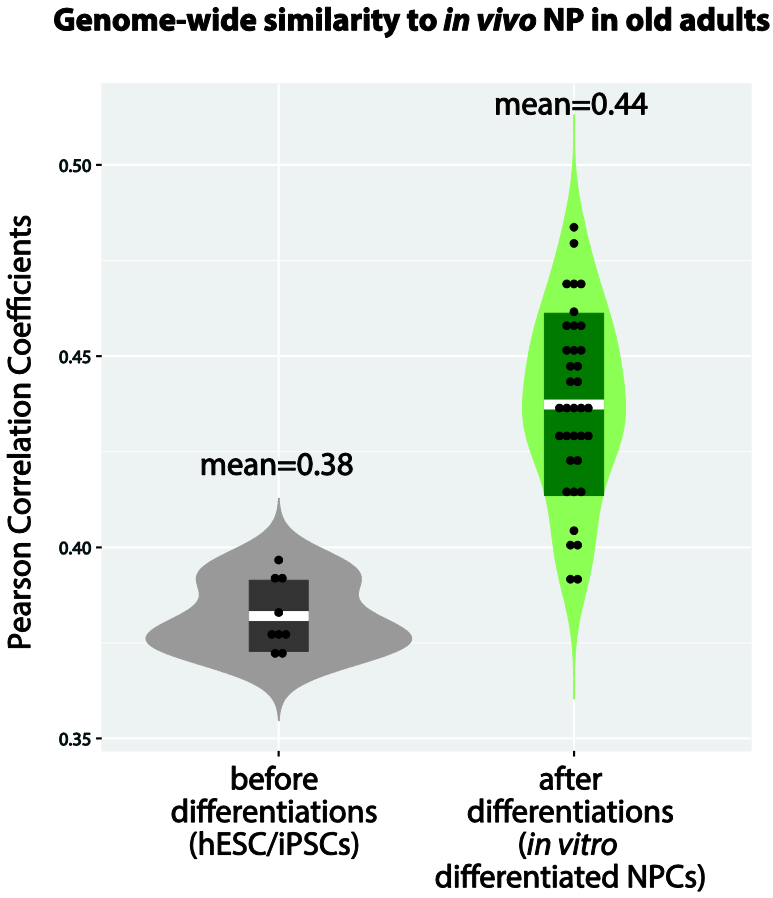


Figure S4, Stem cell markers POU5F1(Oct-4), SOX2 and NANOG were significantly down-regulated in the NPC differentiations. Samples prefixed with 'CDS' are *in vivo* human NP. Log2FC refers to $\log_2$(*in vitro* NPC)-$\log_2$(hESC or iPSC).

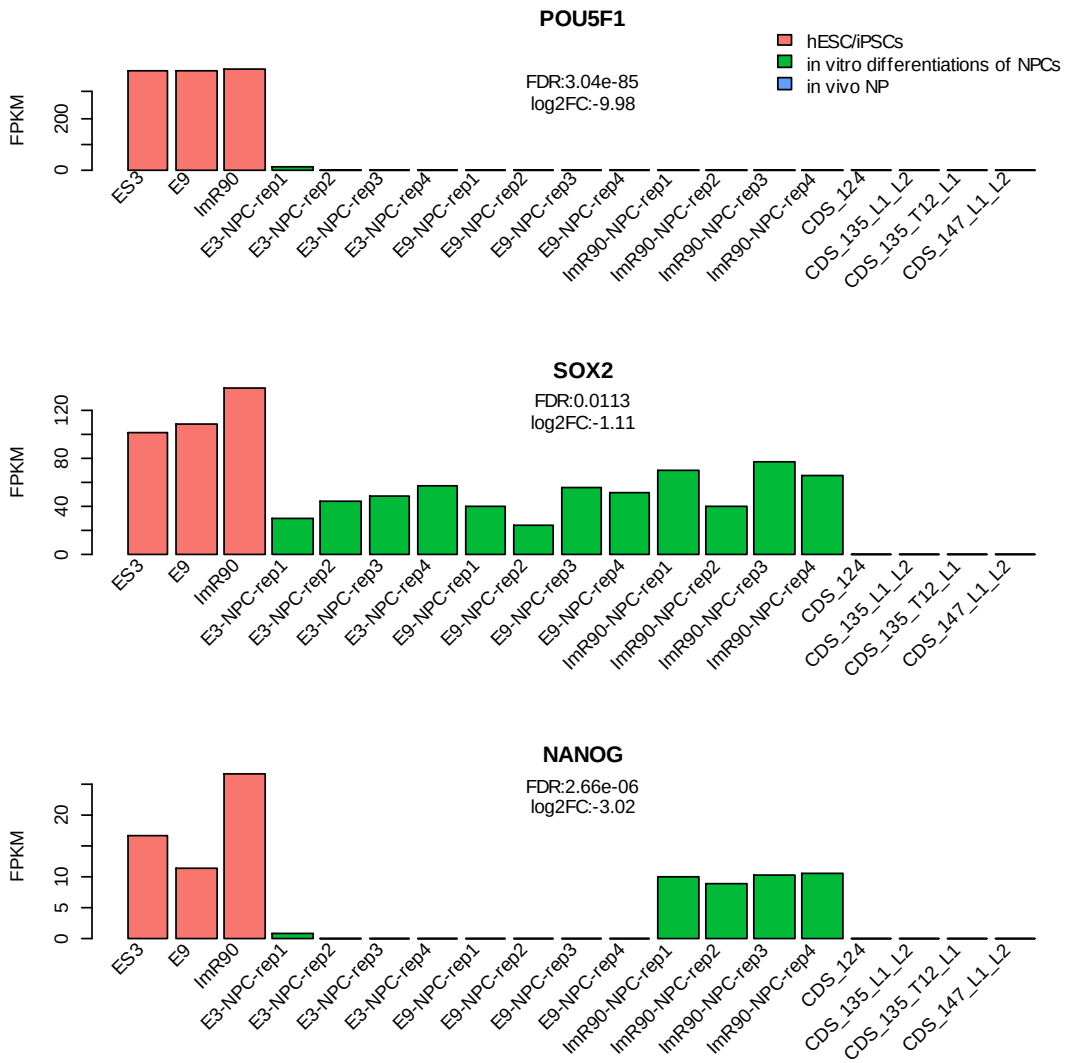


Figure S5, Down-stream targets of Tgf-β pathway, ID1/2/3/4, PITX2, CDKN2B, significantly up-regulated in the NPC differentiations. Log2FC refers to $\log_2$(*in vitro* NPC)-$\log_2$(hESC or iPSC). FDR represents the statistical significance between the first three hESC/iPSC samples, and the remaining *in vitro* differentiations.

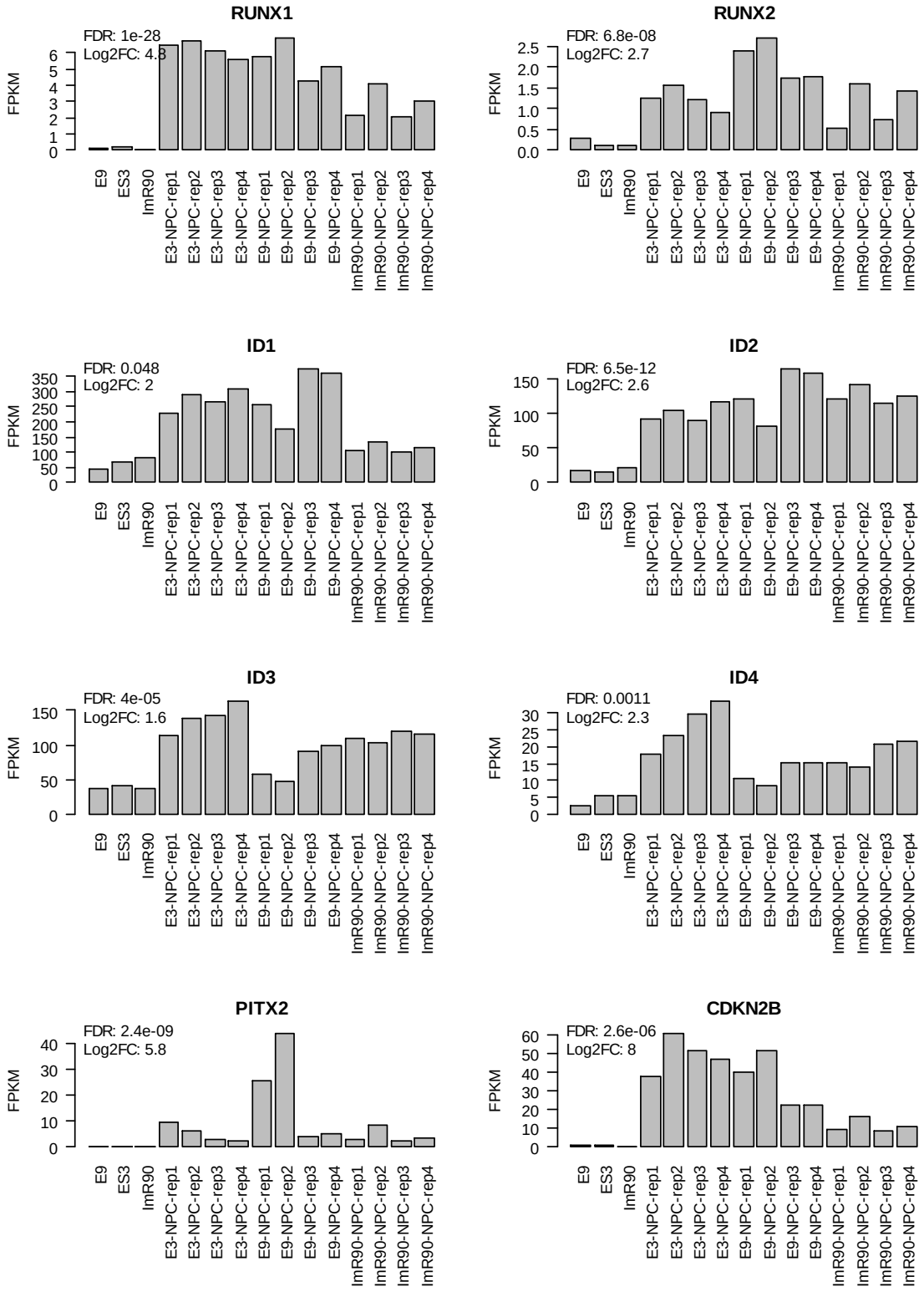


Figure S1, A panel of reported markers were up-regulated in NP differentiations. Log2FC refers to $\log_2$(*in vitro* NPC)-$\log_2$(hESC or iPSC). FDR represents the statistical significance between the first three hESC/iPSC samples, and the remaining *in vitro* differentiations.

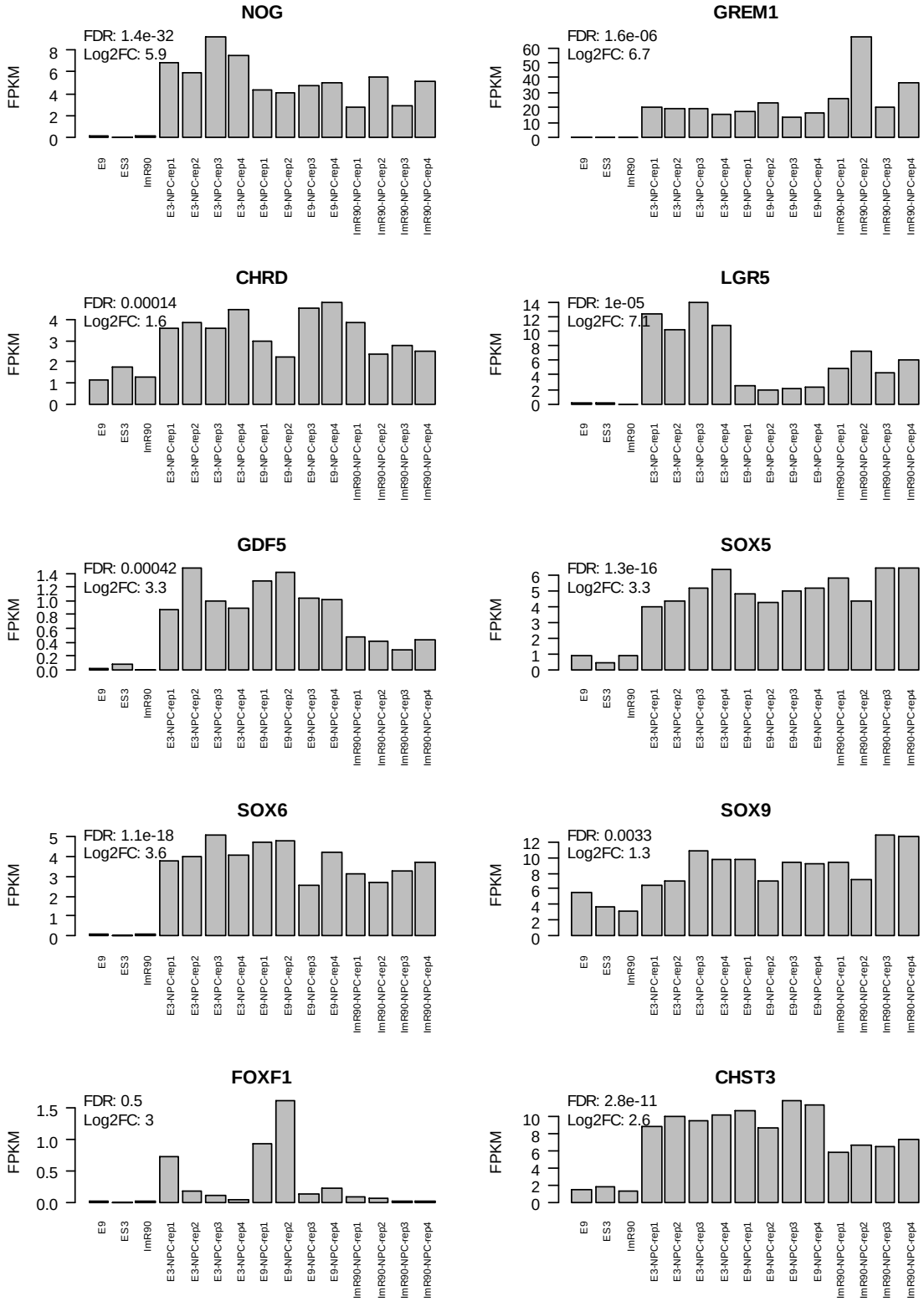


Figure S2, Profiles of additional reported markers or key genes in NP differentiations and hESC/iPSCs. Log2FC refers to $\log_2$(*in vitro* NPC)-$\log_2$(hESC or iPSC). FDR represents the statistical significance between the first three hESC/iPSC samples, and the remaining *in vitro* differentiations.

.

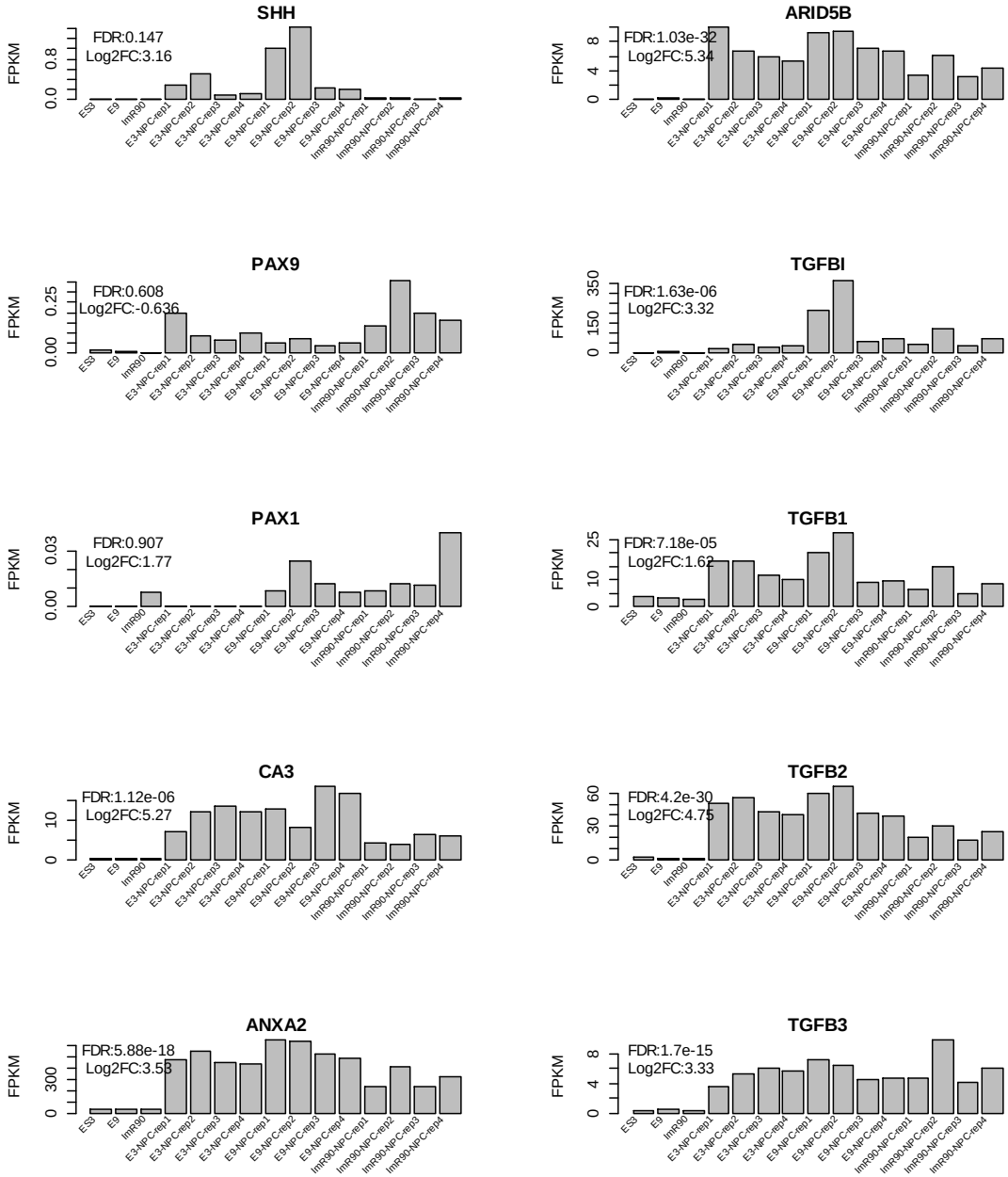


Figure S3, Profiles of additional reported markers or key genes in NP differentiations and hESC/iPSCs. Log2FC refers to $\log_2$(*in vitro* NPC)-$\log_2$(hESC or iPSC). FDR represents the statistical significance between the first three hESC/iPSC samples, and the remaining *in vitro* differentiations.