

Supplementary Figure 1. Scatter plot of gene expressions between technical replicates (5ug total RNA as starting material). The left and right plots are about the replicates using THP-1 and HeLa RNA, respectively. The numbers at upper right panels are Pearson's correlation coefficient based on log₂ of TPM.

Supplementary Figure 2. Scatter plot of gene expressions between different starting materials (5ug to 100ng total RNA of THP-1). The numbers at upper right panels shows Pearson's correlation coefficient based on log₂ of TPM and the number of genes detected with both of the compared profiles, with two threshold (0.1 tpm and 1 tpm). The horizontal gray lines indicate 1 tpm. Averaged TPMs in the technical replicates are used.

Supplementary Figure 3. Manual checking of genes called as 'missed' by HeliScopeCAGE and detected by microarray find examples of a) novel promoter within coding exon of *KRT10*, b) novel peaks without EST support for *PLXNB2*, and c) a likely example of upstream locus *AP1TD1* promoter driving intergenic fusion transcripts with *CORT1*.

Supplementary Figure 4. Additional novel differentially expressed regions identified by HeliScopeCAGE using edgeR (A) Novel THP-1 enriched signal chr19:785213..785668+, between the *AZU1* and *PRTN3* loci (AVGTPM HeLa = 1.6, AVGTPM THP-1 = 3182, FDR = 2.73E-146) (B) Novel THP-1 enriched signal chr6:159462337..159462622- (AVGTPM HeLa = 0, AVGTPM THP-1 = 205, FDR = 3.02E-54) (C) Novel HeLa enriched signal chr21:43685411..43685520- (AVGTPM HeLa = 279, AVGTPM THP-1 = 0, FDR = 1.75E-74) (D) HeLa enriched signal chr17:72632120..72632181+ (AVGTPM HeLa = 131, AVGTPM THP-1 = 0, FDR = 1.79E-50). Log scale used, purple = minus strand, green = plus strand.

Supplementary Figure 5. Examples of alternative promoters a) *STARD8* has a THP1 and a HeLa specific promoter b) *PTPN6* has two promoters, one which is specific to HeLa, c) *CXCR1* is only expressed in THP-1 and uses two promoters .

Supplementary Figure 6. logFC of top 2 TCs for genes with more than one highly expressed TC in the 89975 set a) correlation of top TC pairs from the same locus b) correlation of top TC pairs from random locations

Supplementary Figure 7. Length distribution of HeliScopeCAGE tags

Supplementary Figure 8. Improvement of CAGE by HeliScopeCAGE a) Comparison of the technical reproducibility level of 454CAGE and HeliScopeCAGE. Scatter plot of the normalized TSS expression level (in tag-per-million) variation between two selected technical replicate HeliScopeCAGE THP-1 5ug libraries (in blue) is compared to that obtained with two technical replicates 454 CAGE THP-1 from FANTOM4, showing the greater level of measurement consistency obtained with HeliScopeCAGE. (Note that similar results are obtained with any two out of the three HeliScopeCAGE technical replicates produced, see Supp. Fig1 and Supp. Fig2 for HeliScopeCAGE technical replicates self-consistency). b) Log₂ expressions of qRT-PCR (normalized copy number) and the other technologies: HeliScopeCAGE (tpm), 454CAGE (tpm), and microarray (quantile normalized intensities). All of the expressions are averages within replicates. The qRT-PCR and 454 CAGE data are the ones prepared in FANTOM4 (Suzuki et al. 2009; Ravasi et al. 2010) and downloaded from the FANTOM web resource (Kawaji et al. 2011), and only the transcription factors profiled by qRT-PCR are plotted. Gray line represent 2 fold changes in the log₂ scale (4 fold in the original scale) from their regression line.

Supplementary Figure 9. Fraction of TCs falling into sharp (less than 10 bases) or broad (>=10 bases) wide in Helicos, 454 and 454 libraries after PCR bias correction.

Supplementary Figure 10. Fine level TSS preference differences between HeLa and THP-1 in the *B4GALT1* locus are revealed by HeliScopeCAGE.

Supplementary Figure 11. Fine level TSS preference differences between HeLa and THP-1 in the *MTA2* locus are revealed by HeliScopeCAGE.

Supplementary Figure 12. Genomic views of additional genes with clear sense/antisense exon painting – *SRF*, *SPI1*, *IRF8*.

Supplementary Figure 13. MvsA-plot of gene expression profiles starting with 200ng total RNA of THP-1 and 5ug total RNA of HeLa. Differential analysis is performed on this pair, as well as a pair of 5ug THP-1 RNA and 5ug HeLa RNA. All the analysis is based on technical duplicates, and the number of identified differentially expressed genes are indicated in the parentheses. A) 5ug HeLa vs 5ug THP-1, B) 5ug HeLa vs 200ng THP-1.

Supplementary table 1 Number of mapped reads in the HeliScopeCAGE libraries generated

Supplementary table 2 Genes detected and called as differentially expressed between THP-1 and HeLa using Illumina microarrays and HeliscopeCAGE.

Supplementary table 3 Manual checking of Loci called as missed by Refseq and detected by microarray

Supplementary table 4 edgeR analysis of top 89975 TCs and gene annotation – includes novel and alt promoter examples.

Supplementary table 5 Correlations observed for pairs of alternative promoter TCs separated by different distances

Supplementary table 6 Fraction of TCs falling into sharp (less than 10 bases) or broad (≥ 10 bases) wide in Helicos, 454 and 454 libraries after PCR bias correction.

Supplementary table 7 Average CAGE cDNA yields for 5ug totalRNA starting with or without Actinomycin D in reverse transcription reaction.

