

SUPPLEMENTARY DATA

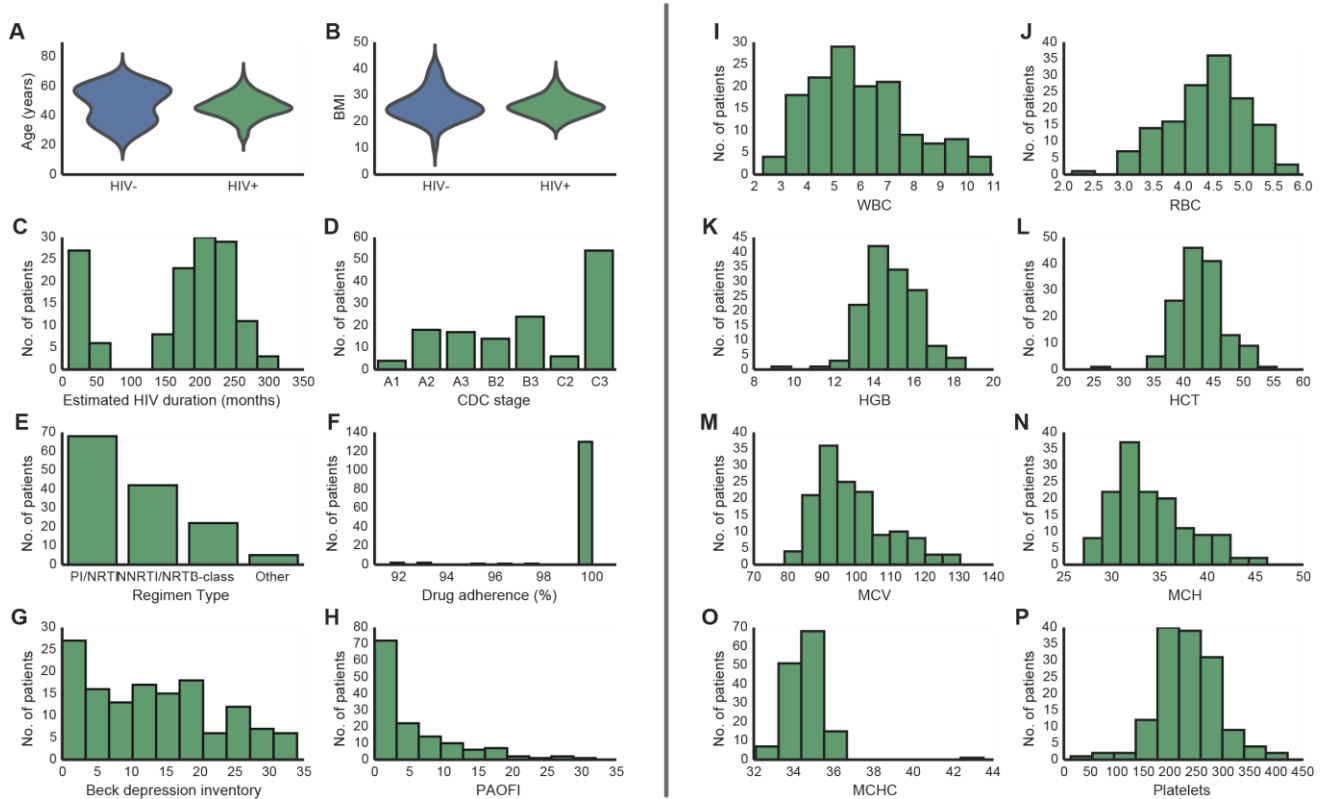


Figure S1. Clinical variables and blood-based biomarkers collected for the primary cohort, related to Experimental Procedures. **A-B**, Comparison of age and body mass index (BMI) across 137 cases and 42 control patients. **C-H**, Clinical variables collected from HIV-infected patients at the time of sample collection. HIV staging was based on standards of the Centers for Disease Control (CDC). Regime types indicate 2-class (PI, protease inhibitor; NRTI, nucleoside reverse-transcriptase inhibitor; NNRTI, non-nucleoside reverse-transcriptase inhibitor); 3-class, (PI/NRTI/NNRTI); PAOFI, patient's assessment of own functioning inventory. **I-P**, Complete blood counts from HIV-infected patients. WBC, white blood cell count; RBC, red blood cell count; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration. For further characterization of HIV patients see **Table S1**.

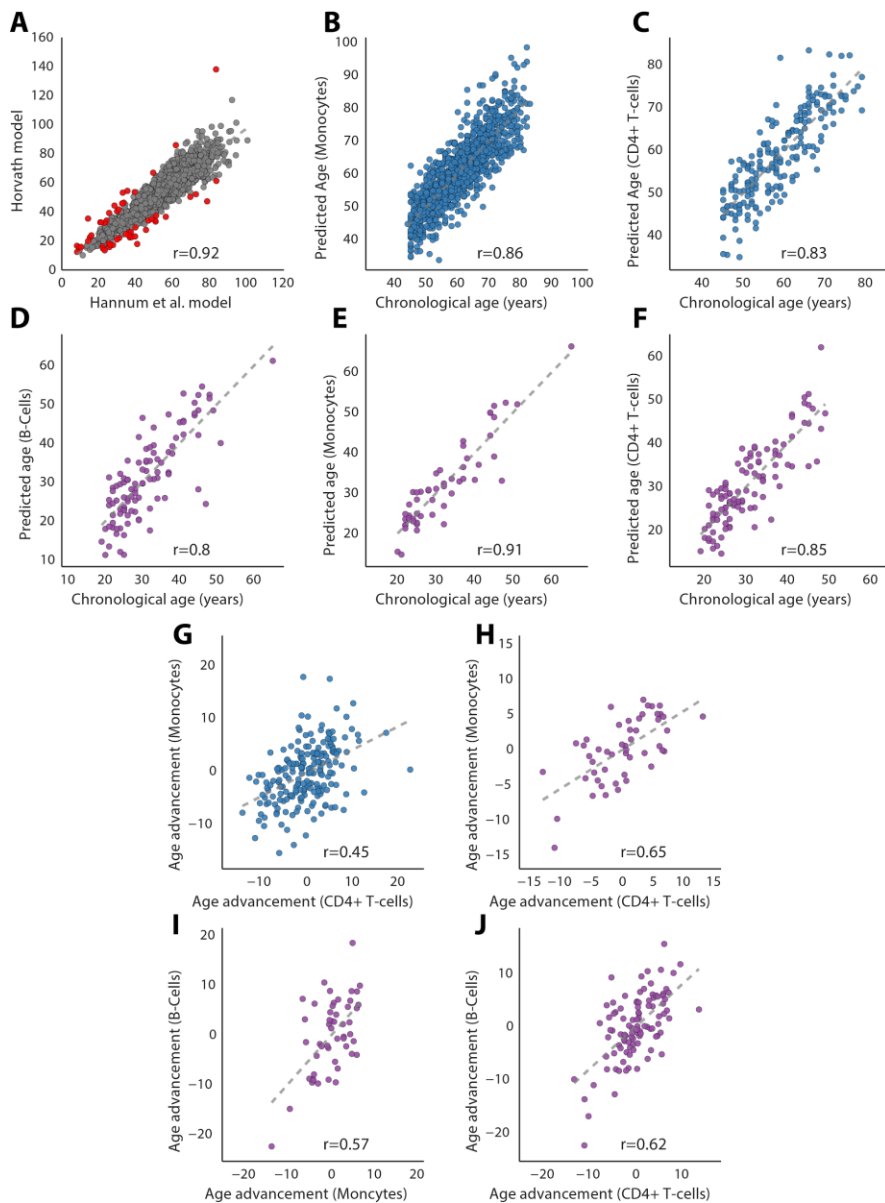


Figure S2. Evaluation of epigenetic age predictions in sorted cell datasets, related to Figure 3. **A**, Scatter plot comparing the ages predicted using the Hannum *et al.* and Horvath models on 1688 samples obtained from sorted cell datasets. Red points indicate patients that were discarded due to disagreement between the two aging models ($n=54$). **B-C** Accuracy of the consensus model (y-axis) to predict true chronological age (x-axis) in sorted cell datasets from Reynolds *et al.* ($n=1130$ and 201 for monocytes and CD4+ T-cells, respectively). **D-F**, Accuracy of the consensus model (y-axis) to predict true chronological age (x-axis) in sorted cell datasets from Absher *et al.* ($n=54$, 103 , and 102 for monocytes, B cells, and CD4+ T-cells, respectively). **G-H**, Scatter plots comparing age advancement in patient matched samples for Reynolds *et al.* (**G**) and Absher *et al.* (**H-J**).

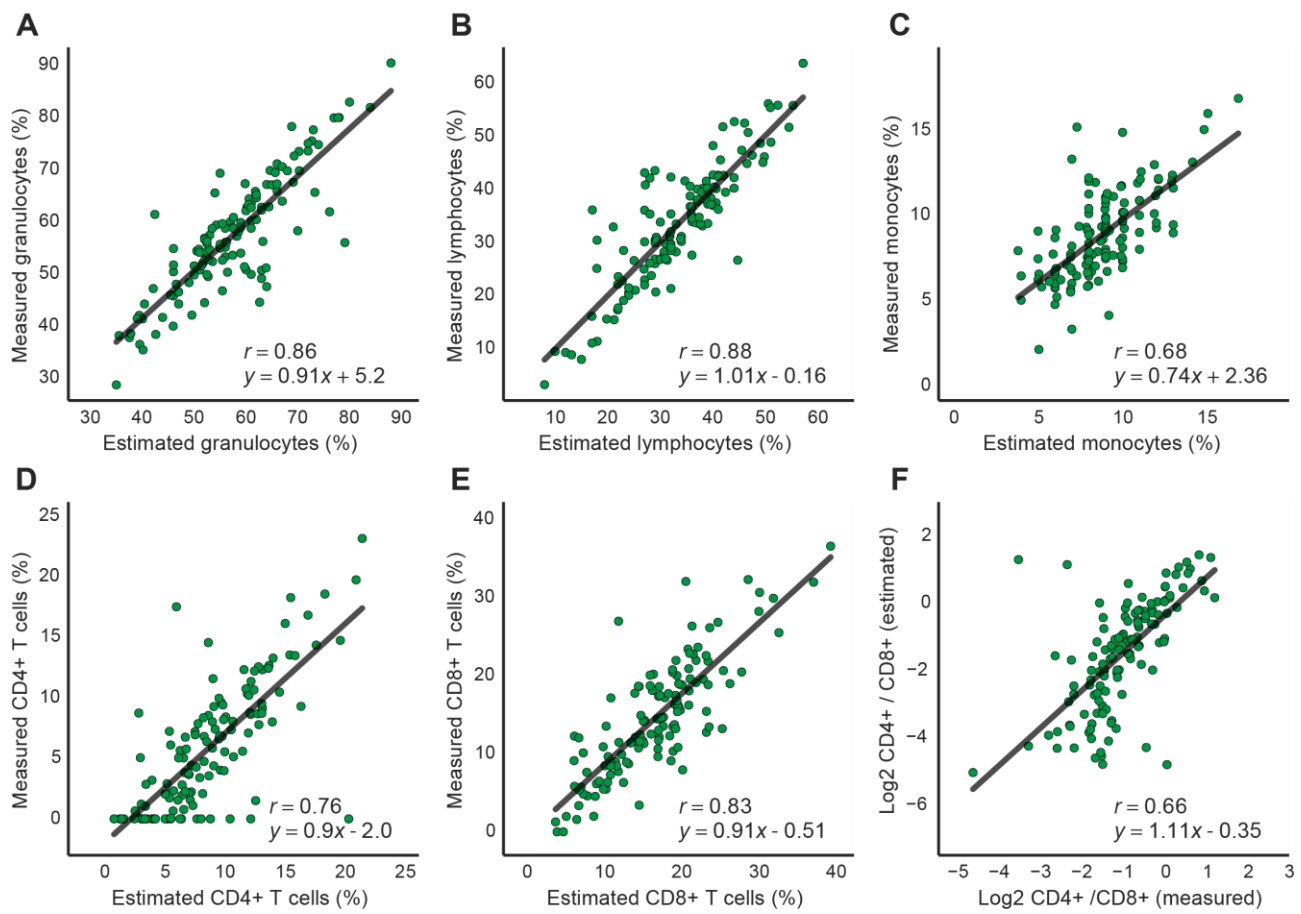


Figure S3. Concordance of estimated cell counts with lab measured values in HIV-infected patients, related to Experimental Procedures. Percentages of each cell type (A-F) were estimated using the Jaffe and Irizarry method. Black lines indicate regression.

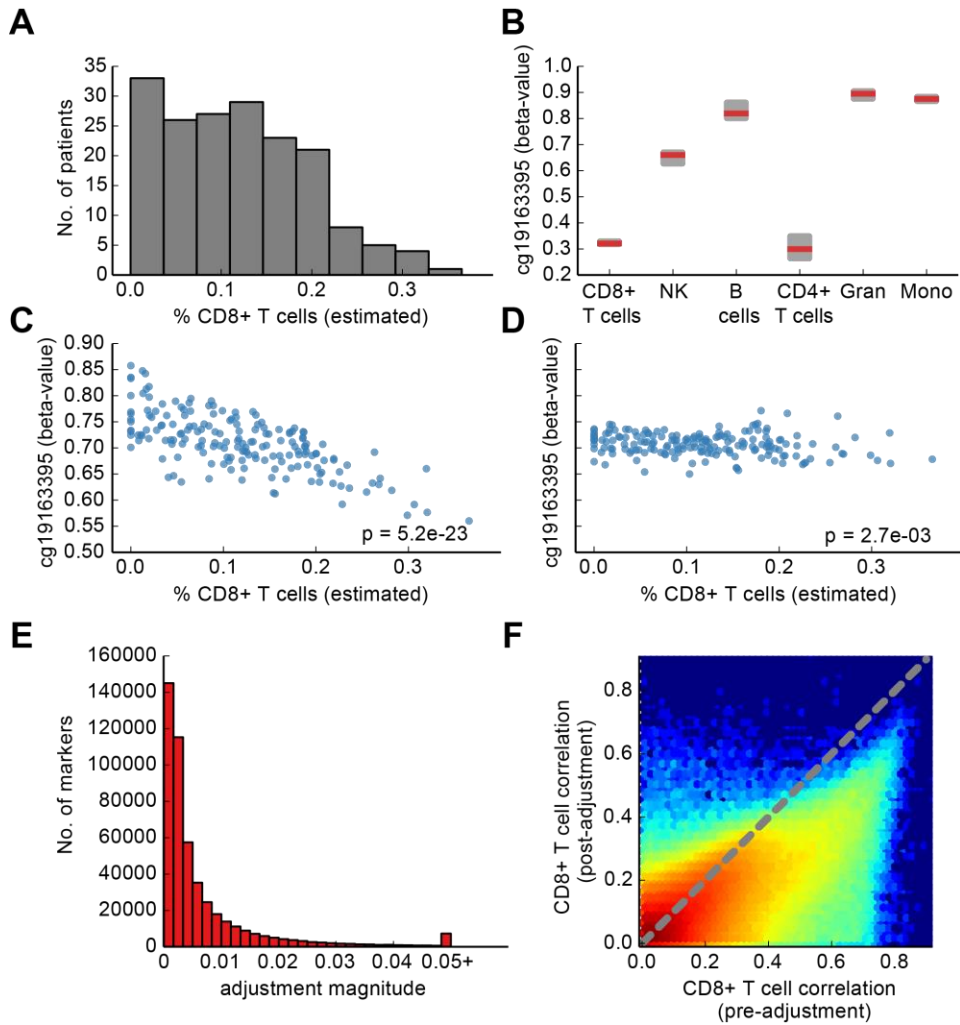


Figure S4. Summary of cell composition adjustment procedure, related to Experimental Procedures. **A**, Distribution of estimated CD8+ T cell composition across the primary patient cohort (Methods). **B**, Distribution of CpG methylation fraction for marker cg19163395 across flow sorted blood populations for six individuals obtained from the Houseman *et al.* dataset. NK, Natural Killer cells; Gran, Granulocytes; Mono, Monocytes. **C-D**, Association of cg19163395 methylation fraction *versus* CD8+ T cell abundance in unadjusted (**C**) and adjusted (**D**) data. **E**, Distribution of the standard deviation of magnitude adjustment across the primary cohort for each marker measured. **F**, Hex-bin plot showing correlation of marker methylation levels (logit-adjusted beta values) with CD8+ T cell abundance before and after cell composition adjustment.

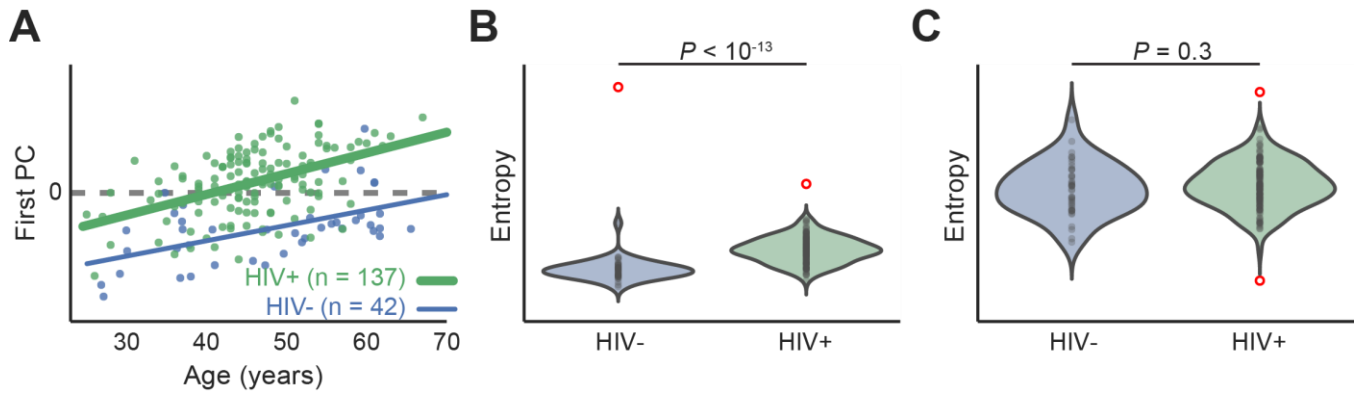


Figure S5. Observed disorder (entropy) in different sets of CpG markers, related to Experimental Procedures. **A**, First principal component (PC) of 7967 age associated markers that trend away from disorder across HIV-infected patients (green) and healthy controls (blue, **Experimental Methods**). **B**, Relative entropy comparing HIV+ to HIV- individuals across the 473,044 markers passing quality control. **C**, Relative entropy comparing HIV+ to HIV- individuals across 436 markers used at least one of the epigenetic models of aging. Significance assessed by Mann-Whitney U test. Note that red circles indicate outliers ± 3 standard deviations away from the mean. These are not used to fit the violin profile, but are used in the statistical assessment.

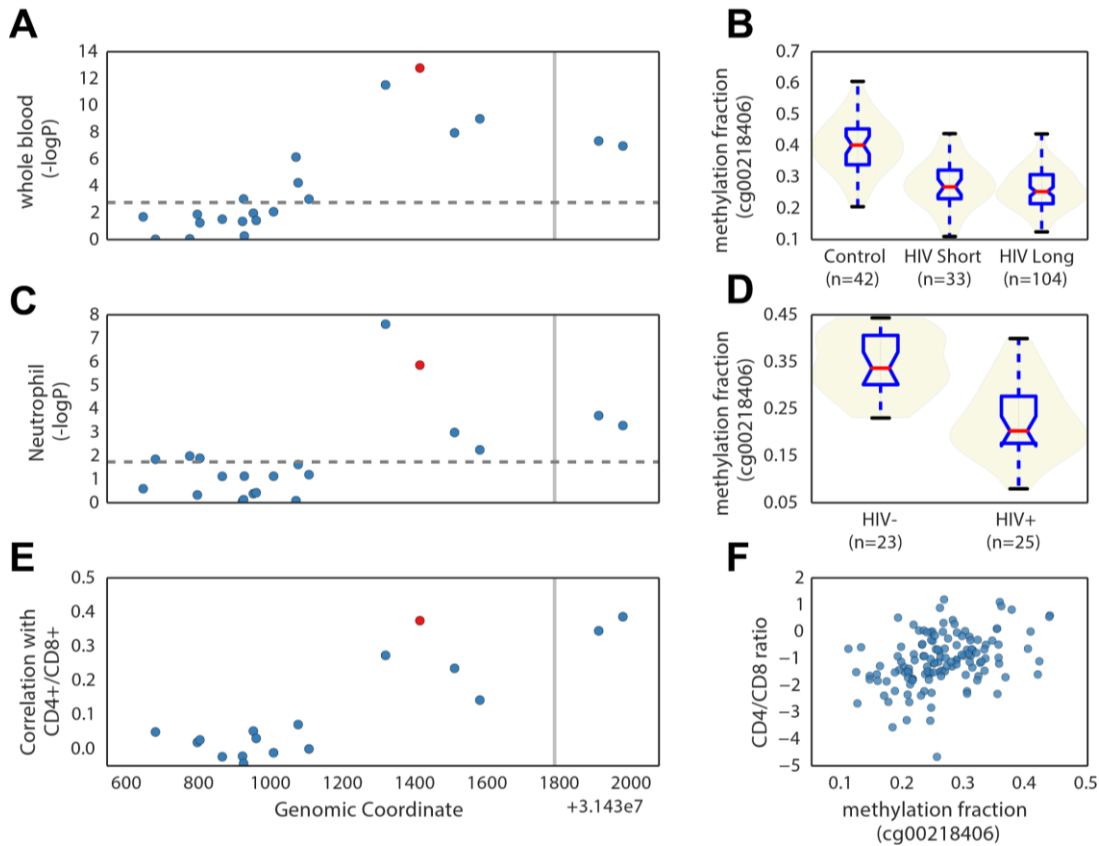


Figure S6. Exploration of methylation markers annotated to HCP5, related to Figure 5. A, Log p-values for association of methylation with HIV infection in primary whole blood samples. **B,** Violin plots showing the distribution of methylation values for cg0028406, the most HIV-associated marker in this region. **C,** Log p-values for association of methylation with HIV infection in the purified neutrophil samples. **D,** Violin plots showing the distribution of methylation values for cg0028406. **E,** Correlation of methylation values with CD4+/CD8+ T-cell ratio among HIV+ subjects. **F,** Scatter plot of CD4+/CD8+ T-cell ratio versus cg0028406 methylation levels. Vertical line corresponds to rs2395029, a SNP having major association with HIV host control. Red points in left-hand plots represent cg0028406, the probe profiled in the right-hand plots.

Table S1. Clinical information and meta-data for HIV+ patients and HIV– controls; related to Experimental Procedures. (separate Excel file).

Table S2. Markers associated with age and HIV status; related to Figure 1, Figure 4, and Figure 5 (separate Excel file).

Table S3. Multivariate linear model comparing age and HIV with the first principal component of age-associated methylation probes, related to Figure 1.

	coef	std err	t	P> t	95.0% Conf. Int.
Intercept	0.2209	0.026	8.482	8.78e-15	0.169 0.272
age	-0.0032	0.001	-6.275	2.64e-09	-0.004 -0.002
HIV	-0.0948	0.010	-9.075	2.25e-16	-0.115 -0.074

R-squared: 0.389

Adjusted R-squared: 0.382

Table S4. Performance of different aging models in healthy subjects used in this study, related to Figure 2.

Dataset	EPIC				Hannum <i>et al.</i>				Primary Cohort			
	n	r	Error		n	r	Error		n	r	Error	
Yrs			%	Yrs			%	Yrs			%	
Hannum <i>et al.</i>	662	0.82	3.69	6.88	538	0.87	4.19	8.33	42	0.92	3.77	8.79
Horvath	662	0.75	4.64	8.79	538	0.82	5.14	10.07	42	0.91	3.74	8.76
Consensus	662	0.82	3.69	6.89	538	0.87	4.12	8.06	42	0.93	3.49	8.27
Consensus (Data filtered)	637	0.82	3.65	6.79	497	0.86	4.06	7.81	40	0.94	3.30	7.60

n, number of samples; *r*, Pearson’s correlation coefficient; Error (Yrs): Mean absolute prediction error in years; Error (%): Average prediction error scaled by age magnitude.

Table S5. Clinical information and meta-data for HIV+ patients and HIV– controls in sorted cell validation datasets; related to Experimental Procedures (separate Excel file).

Table S6. Overlap of age and HIV-associated markers within specific genomic contexts, related to Figure 4.

	Odds Ratio	p-value
All Markers	1.29	3.88e-59
Promoter	1.22	2.50e-08
Gene Body	1.24	4.04e-13
TSS	1.33	2.57e-26
DHS Site	1.46	8.45e-29
Enhancer	1.46	1.73e-26
PRC2	1.68	3.32e-81
CpG Island	1.70	9.32e-110