

Supplemental Table 2: Description of the short-read experimental datasets used in this study*ENCODE DATASETS*

Experiment (Assay) type	Sequencing Provider	Cell Line	Link to fasta files	File Name
CTCF	Broad Institute	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeBroadHistone/	wgEncodeBroadHistoneGm12878CtcfStdRawDataRep1
H3K4me3	Broad Institute	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeBroadHistone/	wgEncodeBroadHistoneGm12878H3k04me3StdRawDataRep2V2
H3K27me3	Broad Institute	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeBroadHistone/	wgEncodeBroadHistoneGm12878H3k27me3StdRawDataRep1
H3K9me3	Broad Institute	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeBroadHistone/	wgEncodeBroadHistoneGm12878H3k9me3StdRawDataRep1
Long RNA Seq	Cold Spring Harbor Laboratory	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeCshLongRnaSeq/	wgEncodeCshLongRnaSeqGm12878CellPapFastqRd1Rep1
Methyl Seq	Haib	GM12878	http://hgdownload-test.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeHaibMethylSeq/	wgEncodeHaibMethylSeqGm12878HpaRawDataRep1
GABP	Haib	GM12878	http://hgdownload-test.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeHaibMethylSeq/	wgEncodeHaibTfbsGm12878GbpPcr2xRawDataRep2
DNase-seq	University of Washington	GM12878	http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeUwDnase/	wgEncodeUwDnaseGm12878RawDataRep1
<i>ADDITIONAL GENOMIC DATASETS USED IN STUDY</i>				
Experiment (Assay) type	Sequencing Provider	Cell Line	Link to fasta files	File Name
HuRef Genomic	Duke University	HuRef	http://hgwdev.soe.ucsc.edu/~kehayden/NAR_methods/	HUREF-Genomic
Yoruba Genomic	Duke University	GM129239	http://hgwdev.soe.ucsc.edu/~kehayden/NAR_methods/	GM129239-Genomic