



SUPPLEMENTARY FIGURE 3. Prediction of regulatory regions using WGBS data and testing these regions for enrichment in T2D GWAS regions.

A) Different combinations of epigenomic data (top) were combined to generate different sets of refined chromatin states (middle, 11 ChIP-only and 15 ChIP+Meth, ChIP+ATAC and ChIP+ATAC-Meth states, see figure S3B and 3A-B for actual states) using chromHMM. These sets of chromatin states were then tested for enrichment in T2D-related GWAS traits using FGWAS to compare enrichment across states (bottom). B) ChromHMM (I) 11 ChIP-only and 15 state (II) ChIP+ATAC state and (III) ChIP+Meth models. C) Single feature \log_2FE (x-axis) for different enhancer states (grey panels) defined from different combinations of epigenetic marks (y-axis) including ChIP+ATAC+Meth, ChIP+ATAC, ChIP+Meth and ChIP-only. The grey-dashed line indicates the enrichment value of CDS as reference. Enhancers are defined as follows: Strong enhancers are marked by both H3K4me1 and H3K27ac, weak Enhancers are defined by H3K4me1 only, gene enhancers are marked by H3K4me1 and H3K36me3, other enhancers are marked by H3K4me1, H3K4me3 and H3K27ac and are often referred to as TSS upstream regions (only included in the FGWAS T2D model for ChIP-only and ChIP+Meth chromatin states). D) Since chromatin states defined from a different set of epigenomic marks (ChIP-only, ChIP+Meth, ChIP+ATAC and ChIP+ATAC+Meth), as described in S3A-B, are not equivalent and the enrichment can not be easily compared across models, a nested model approach was applied. That is, ChIP-only chromatin states were generated and after evaluating the individual enrichment of each annotation (see Figure 3D), FGWAS maximum likelihood models were defined using ChIP-only, hypomethylated and/or ATAC-seq peak regulatory regions. The combination of all these annotations represented a nested linear model and the changes in maximum likelihood by adding/removing hypomethylated regulatory and ATAC-seq states could be statistically evaluated using a Loglikelihood Ratio Test (LRT) as shown in Figure 3E. E) Maximum likelihood FGWAS nested model combining ChIP-only, ATAC-peaks and LMR states (y-axis) showing \log_2FE enrichment (x-axis) which was used for the LRT in Figure 3E. For all FGWAS enrichment plots the axis has been truncated at -6 to facilitate visualisation and accurate values are provided in the supplementary tables.