The American Journal of Human Genetics, Volume *102*

Supplemental Data

FUN-LDA: A Latent Dirichlet Allocation Model for Predicting Tissue-Specific Functional Effects of Noncoding Variation: Methods and Applications

Daniel Backenroth, Zihuai He, Krzysztof Kiryluk, Valentina Boeva, Lynn Pethukova, Ekta Khurana, Angela Christiano, Joseph D. Buxbaum, and Iuliana Ionita-Laza

Supplemental Material

Summary of results for six SNPs in the literature, with evidence of regulatory function.

- rs12350739 has been shown to influence human skin color by regulating transcription of nearby *BNC2* pigmentation gene[1]. In Figure S10 we show the predictions for Roadmap tissue E059: Foreskin Melanocyte Primary Cells skin01, the tissue we deemed closest to the one used in the functional study, melanocyte cell lines.

- rs12740374: In[2] the authors show using human-derived hepatocytes that SNP rs12740374 creates a C/EBP (CCAAT/enhancer binding protein) transcription factor binding site and alters the hepatic expression of the SORT1 gene. In Figure S11 we show the predictions for Roadmap tissue E066: Liver, the tissue we deemed closest to the one used in the functional study, human-derived hepatocytes.

- rs356168: In[3], the authors performed allele-specific TaqMan® qRT-PCR analysis in human induced pluripotent stem cells (hIPSC)-derived neurons and show that this SNP regulates the expression of the *SNCA* gene, a gene implicated in the pathogenesis of Parkinson's disease. In Figure S12 we show the predictions for Roadmap tissue E007: H1 Derived Neuronal Progenitor Cultured Cells, the tissue we deemed closest to the one used in the functional study, hIPSC-derived neurons.

- rs2473307: In[4], the authors showed evidence that this SNP, associated with schizophrenia, reduces expression of *CDC42* gene in a human neuronal cell line. In Figure S13 we show the predictions for Roadmap tissue E007, H1 Derived Neuronal Progenitor Cultured Cells.

- rs227727: In[5], the authors show that this SNP, in perfect LD with the most significant GWAS variant, alters the function of an enhancer. In Figure S14, we show the predictions for Roadmap tissue E119, HMEC Mammary Epithelial Primary Cells.

- rs144361550: In[6], the authors show that this SNP, in strong LD with a lead GWAS variant, displays allele-specific transcriptional activity in primary melanocytes. Furthermore, mass spectrometry analyses using melanoma cell line revealed that RECQL is an unequivocal allele-preferential binder of rs144361550. In Figure S15, we show the predictions for Roadmap tissue E059: Foreskin Melanocyte Primary Cells skin01, the tissue we deemed closest to the one used in the functional study, melanocyte cell lines.

Inference and parameter estimation in the variational inference procedure. It can be shown that for a single tissue the lower bound on the log likelihood can be written as

$$\begin{aligned}
L(\boldsymbol{a}, \boldsymbol{w}|\boldsymbol{\alpha}) &= \log\Gamma(\alpha_0+\alpha_1) - \log\Gamma(\alpha_0) - \log\Gamma(\alpha_1) + (\alpha_0-1)(\Psi(a_0)-\Psi(a_0+a_1)) \\
&\quad + (\alpha_1-1)(\Psi(a_1)-\Psi(a_0+a_1)) + \sum_{i=1}^{m} w_i(\Psi(a_1)-\Psi(a_0+a_1)) \\
&\quad + (m-\sum_{i=1}^{m} w_i)(\Psi(a_0)-\Psi(a_0+a_1)) + \sum_{i=1}^{m}(1-w_i)\log f_0(\boldsymbol{Z}_i) + \sum_{i=1}^{m} w_i \log f_1(\boldsymbol{Z}_i) \\
&\quad - \log\Gamma(a_0+a_1) + \log\Gamma(a_0) + \log\Gamma(a_1) - (a_0-1)(\Psi(a_0)-\Psi(a_0+a_1)) \\
&\quad - (a_1-1)(\Psi(a_1)-\Psi(a_0+a_1)) - \sum_{i=1}^{m} w_i \log w_i - \sum_{i=1}^{m}(1-w_i)\log(1-w_i),
\end{aligned}$$

where $\Psi(x) = d\log\Gamma(x)/dx$.

Maximizing $L(\boldsymbol{a}, \boldsymbol{w}|\boldsymbol{\alpha})$ with respect to $\boldsymbol{a}$ and $\boldsymbol{w}$, respectively, we get

$$w_i = \frac{f_1(\boldsymbol{Z}_i) \times \exp(\Psi(a_1))}{f_0(\boldsymbol{Z}_i) \times \exp(\Psi(a_0)) + f_1(\boldsymbol{Z}_i) \times \exp(\Psi(a_1))},$$

and

$$a_1 = \alpha_1 + \sum_{i=1}^{m} w_i \quad \text{and} \quad a_0 = \alpha_0 + \sum_{i=1}^{m}(1-w_i).$$

Given the optimal estimates of $\boldsymbol{a}$ and $\boldsymbol{w}$, we maximize the lower bound $L(\boldsymbol{a}, \boldsymbol{w}|\boldsymbol{\alpha})$ with respect to the hyperparameter $\boldsymbol{\alpha}$ by using the Newton-Raphson method as in[7]. Namely, we update $\boldsymbol{\alpha}$ by iterating:

$$\boldsymbol{\alpha}^{\text{new}} = \boldsymbol{\alpha} - H(\boldsymbol{\alpha})^{-1}\nabla L(\boldsymbol{\alpha}).$$

where the gradient $\nabla L(\boldsymbol{\alpha})$ is:

$$\frac{\partial L(\boldsymbol{\alpha})}{\partial \alpha_r} = \Psi(\alpha_0+\alpha_1) - \Psi(\alpha_r) + \Psi(a_r) - \Psi(a_0+a_1) \qquad \text{for } r = 0, 1,$$

and for the Hessian matrix we have:

$$H(\boldsymbol{\alpha}) = -\text{Diag}\big(\Psi'(\alpha_0), \Psi'(\alpha_1)\big) + \Psi'(\alpha_0+\alpha_1)\mathbf{1}\mathbf{1}'.$$

References

[1] Visser M, Palstra RJ, Kayser M (2014) Human skin color is influenced by an intergenic dna polymorphism regulating transcription of the nearby bnc2 pigmentation gene. *Hum Mol Genet* 23: 5750–5562.

[2] Musunuru K et al. (2010) From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 466: 714–719.

[3] Soldner F et al. (2016) Parkinson-associated risk variant in distal enhancer of ?-synuclein modulates target gene expression. *Nature* 533: 95–99.

[4] Gilks WP, Hill M, Gill M, Donohoe G, Corvin AP, Morris DW (2012) Functional investigation of a schizophrenia gwas signal at the cdc42 gene. *World J Biol Psychiatry* 13: 550–554.

[5] Leslie EJ et al. (2015) Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. *Am J Hum Genet* 96: 397–411.

[6] Choi J et al. (2017) A common intronic variant of PARP1 confers melanoma risk and mediates melanocyte growth via regulation of MITF. *Nat Genet* Epub ahead of print

[7] Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993–1022.

[8] Tewhey R et al. (2016) Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell* 165: 1519–1529.

[9] Kheradpour P, Ernst J, Melnikov A, Rogov P, Wang L, Zhang X, Alston J, Mikkelsen TS, Kellis M (2013) Systematic dissection of regulatory motifs in 2000 predicted human enhancers using a massively parallel reporter assay. *Genome Res* 23: 800–811.

[10] Lu Q, Powles RL, Wang Q, He BJ, Zhao H (2016) Integrative Tissue-Specific Functional Annotations in the Human Genome Provide Novel Insights on Many Complex Traits and Improve Signal Prioritization in Genome Wide Association Studies. *PLoS Genet* 12: e1005947.

[11] Libbrecht MW, Rodriguez O, Weng Z, Hoffman M, Bilmes JA, Noble WS (2017) A unified encyclopedia of human functional DNA elements through fully automated annotation of 164 human cell types. doi: https://doi.org/10.1101/086025

[12] Zhang Y, Hardison RC (2017) Accurate and Reproducible Functional Maps in 127 Human Cell Types via 2D Genome Segmentation. BioRxiv preprint doi: http://dx.doi.org/10.1101/118752.

[13] Pennacchio LA et al. (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444: 499–502.

[14] Degner JF et al. (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* 482: 390–394.

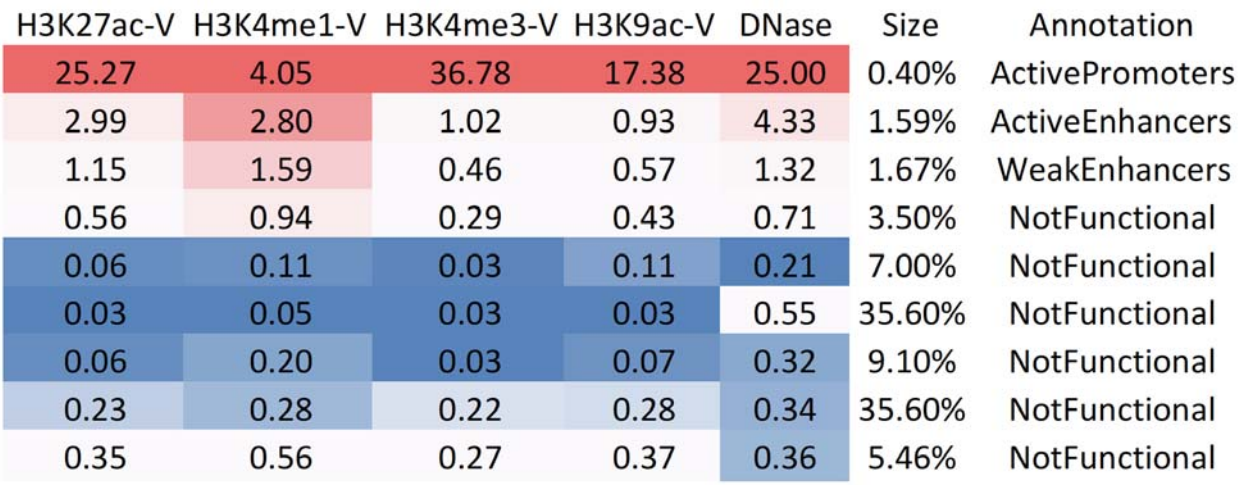


FIGURE S1. Heatmap of epigenetic features vs. class in the FUN-LDA model with nine classes across tissues and cell types in Roadmap.

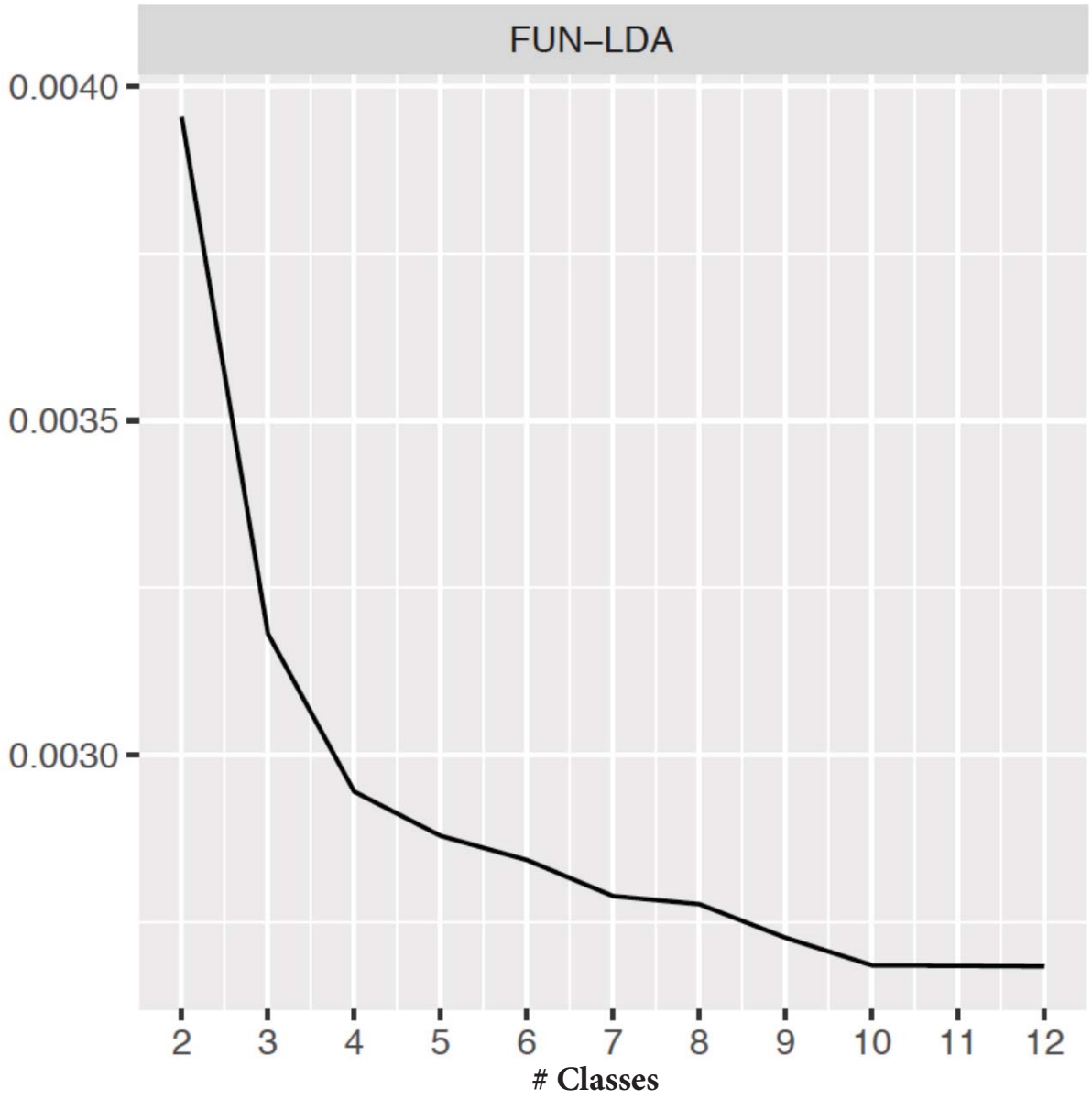


FIGURE S2. Perplexity measure of FUN-LDA models as a function of the number of classes.

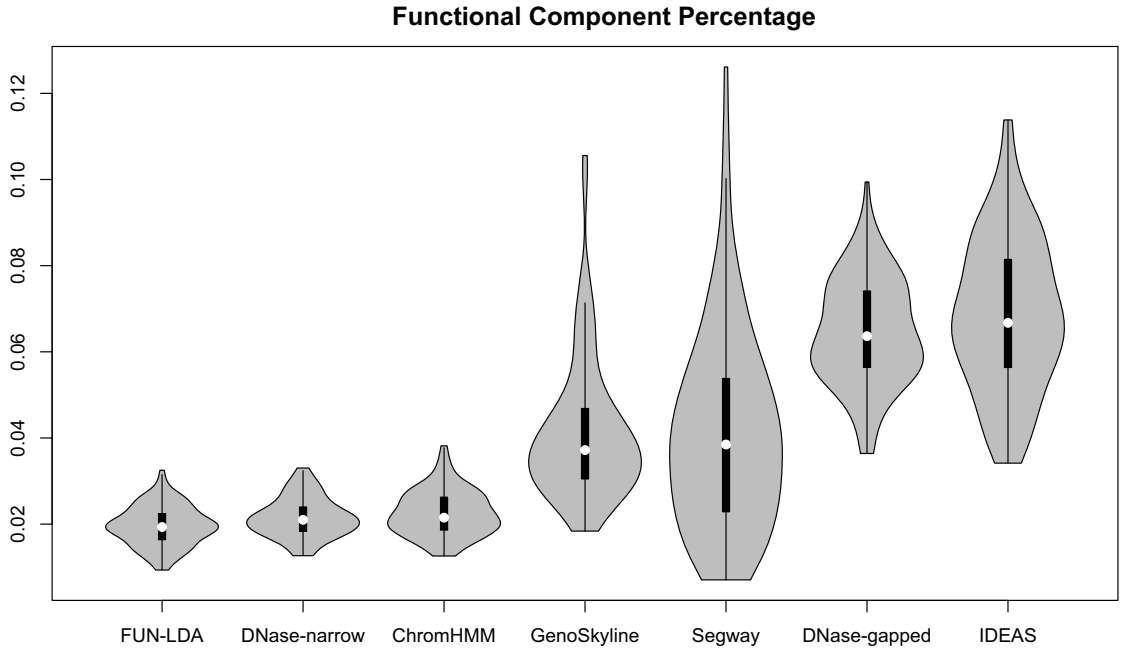


FIGURE S3. Violin plots showing the distribution of proportion of functional variants across tissues in Roadmap for each of several methods.

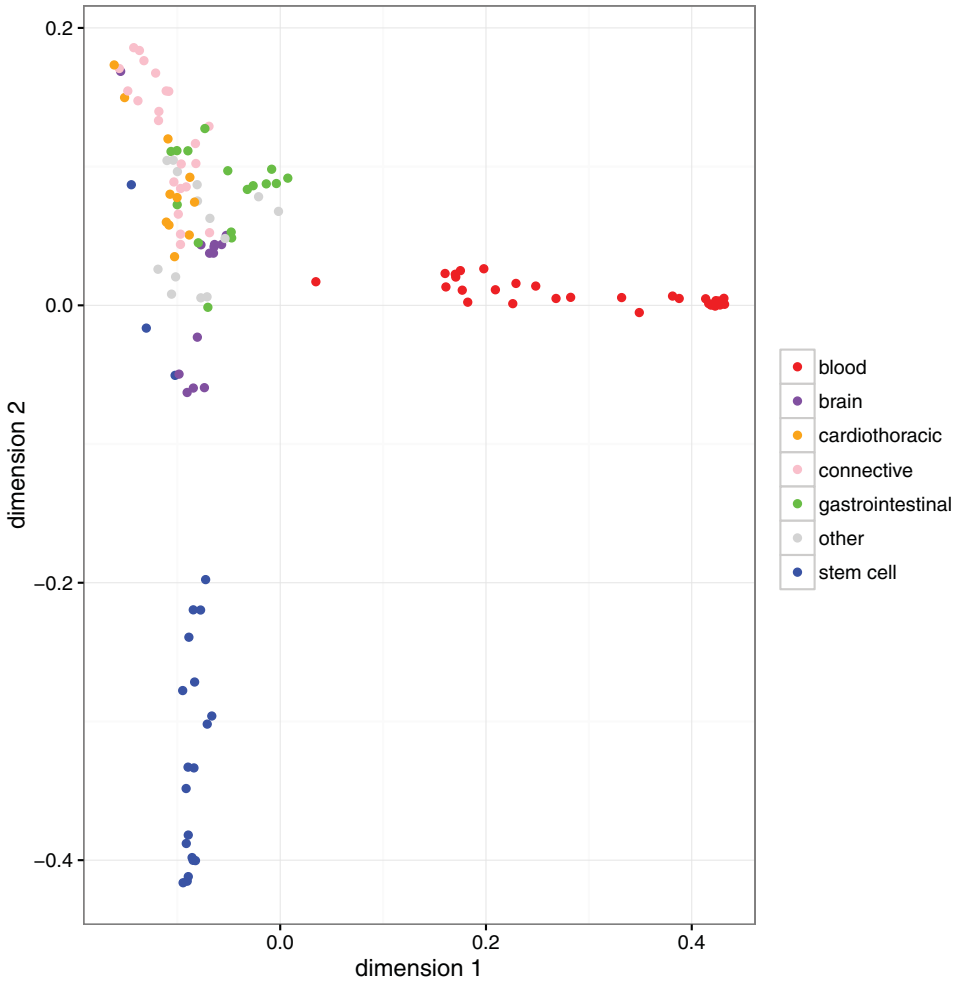
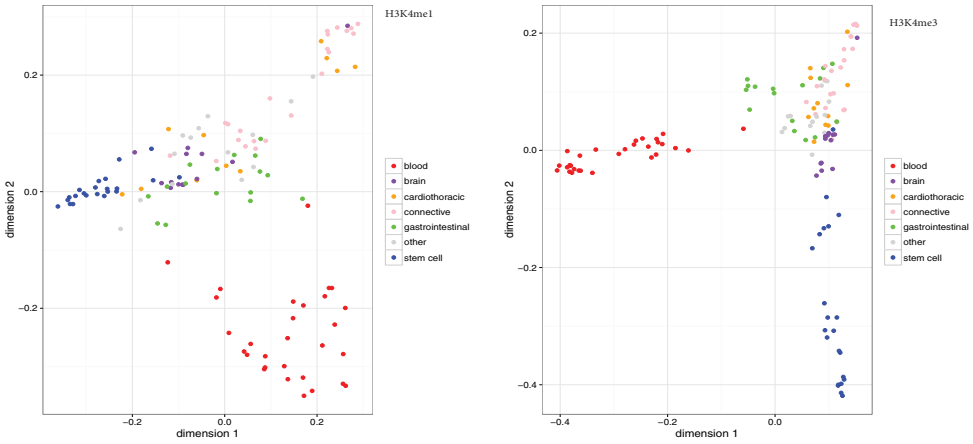
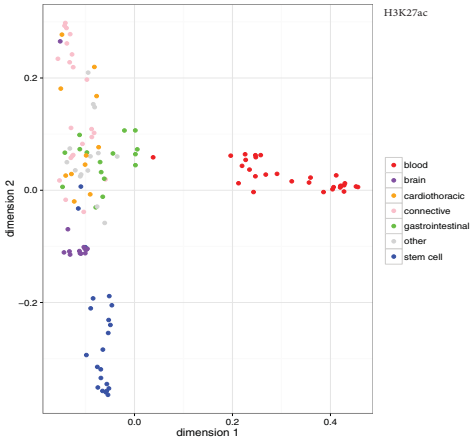


FIGURE S4. Multidimensional scaling plot of the correlations between the functional scores for the different tissues (FUN-LDA).

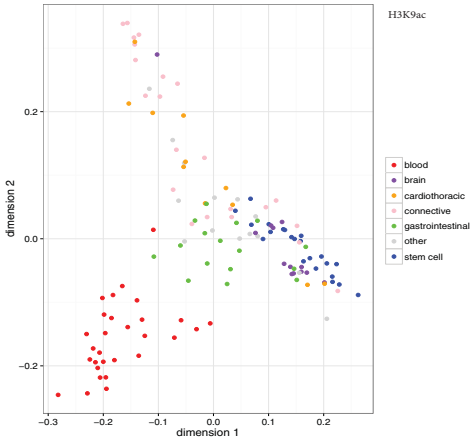


(A) H3K4me1

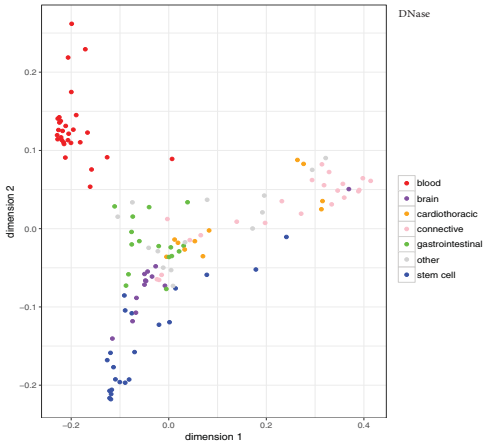
(B) H3K4me3



(C) H3K27ac



(D) H3K9ac



(E) DNase

FIGURE S5. Multidimensional scaling plots of the correlations between the functional scores for the different tissues using individual histone marks and DNase.

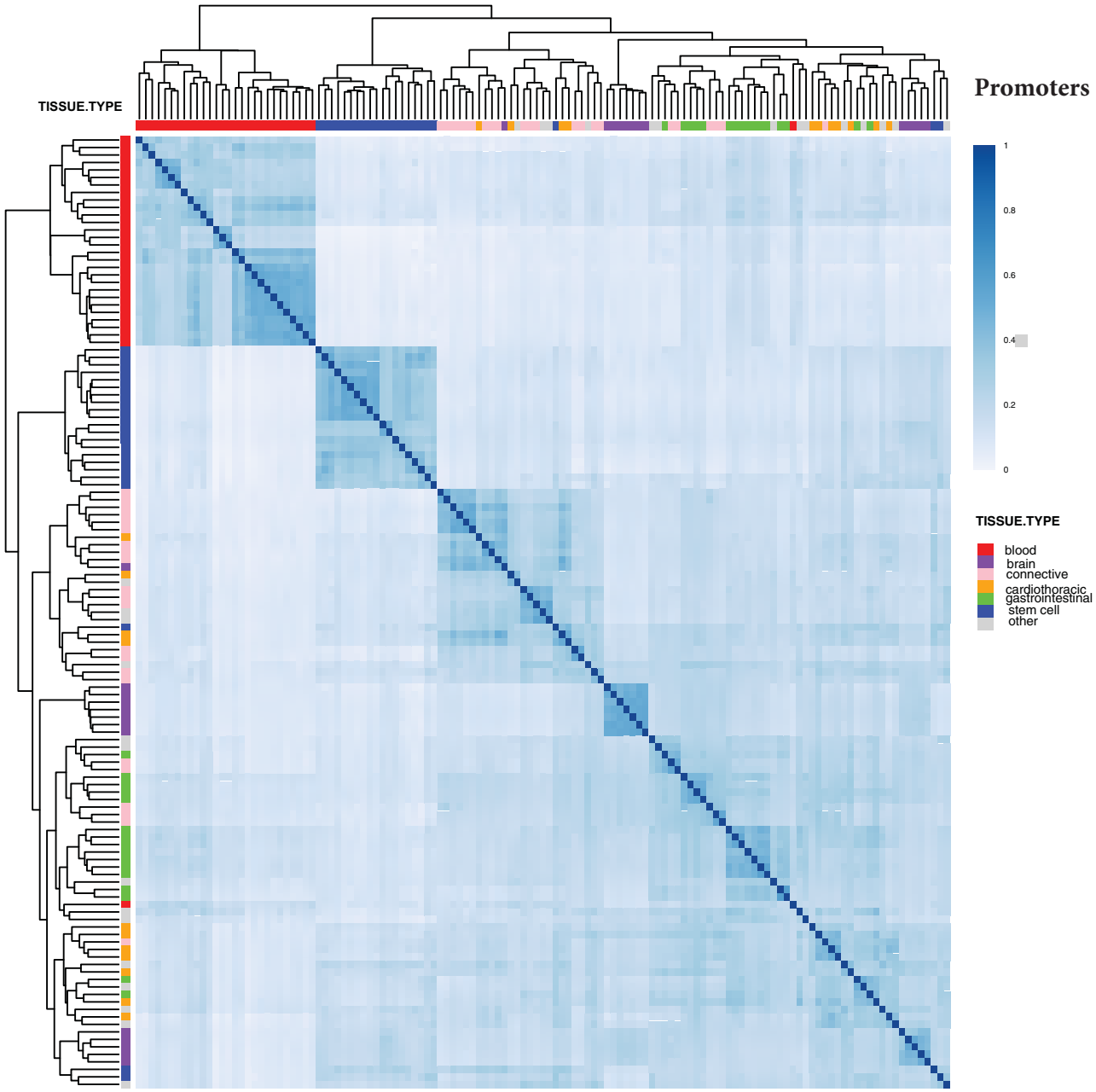


FIGURE S6. Jaccard index of overlap among functional variants falling in *promoter* regions in different cell types and tissues in Roadmap (FUN-LDA). Hierarchical clustering is used to cluster the different cell types and tissues.

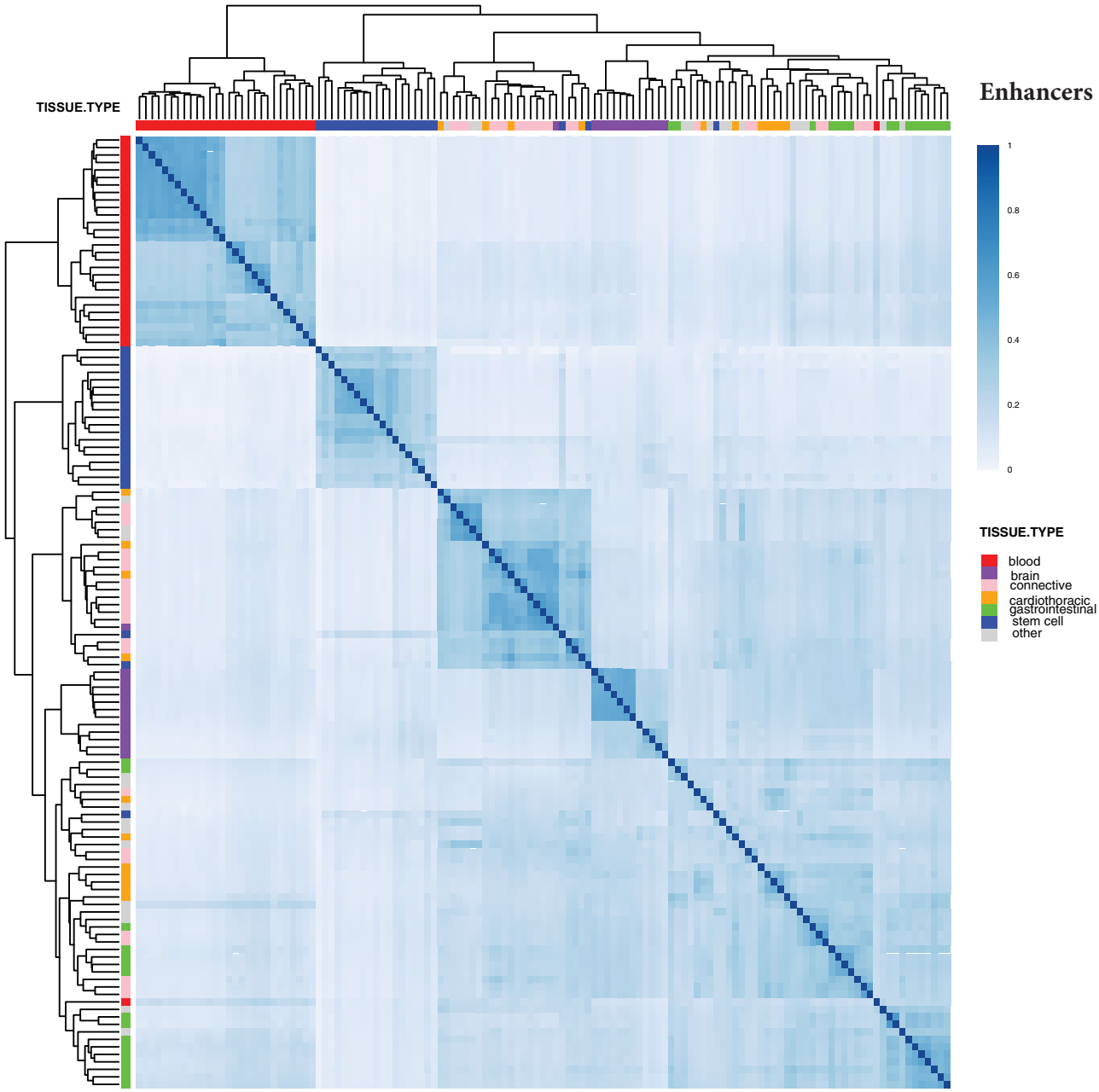


FIGURE S7. Jaccard index of overlap among functional variants falling in *enhancer* regions in different cell types and tissues in Roadmap (FUN-LDA). Hierarchical clustering is used to cluster the different cell types and tissues.

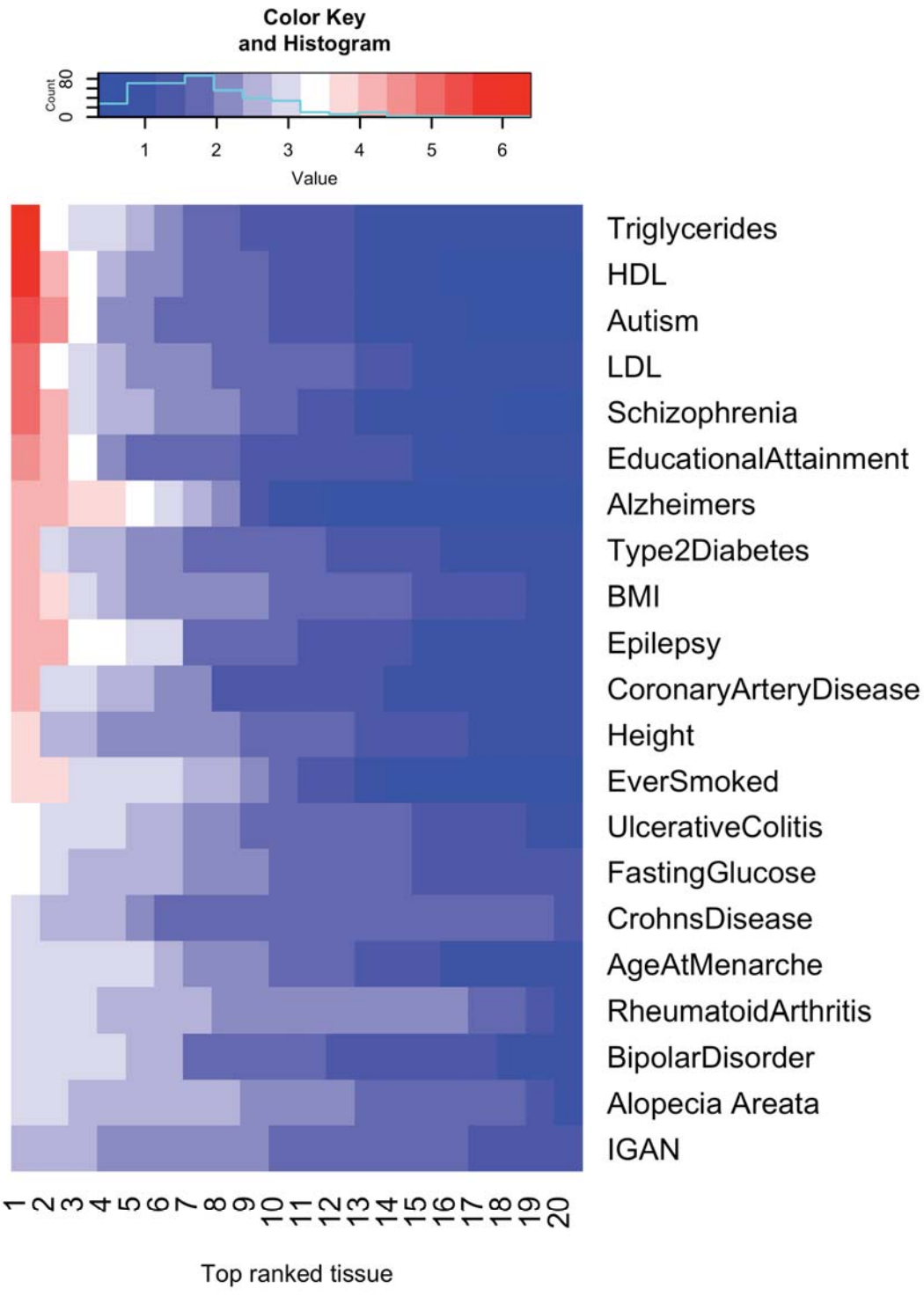
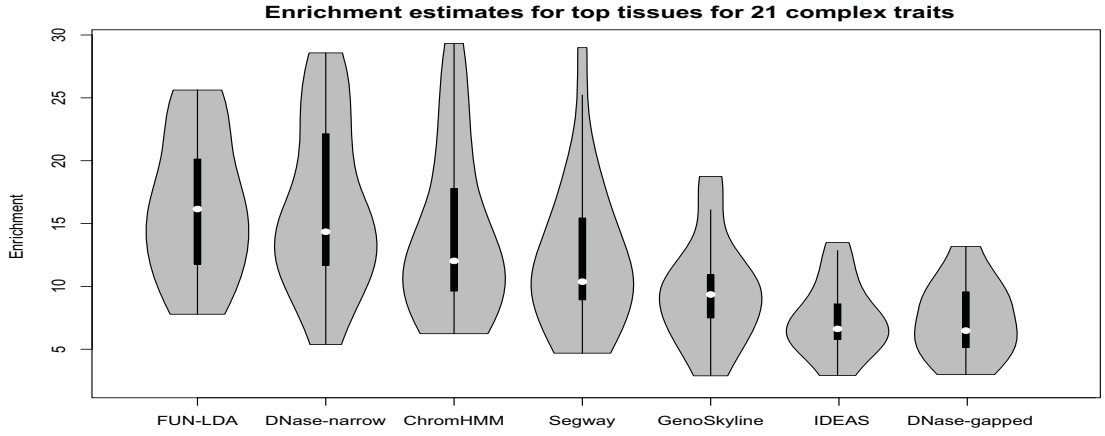
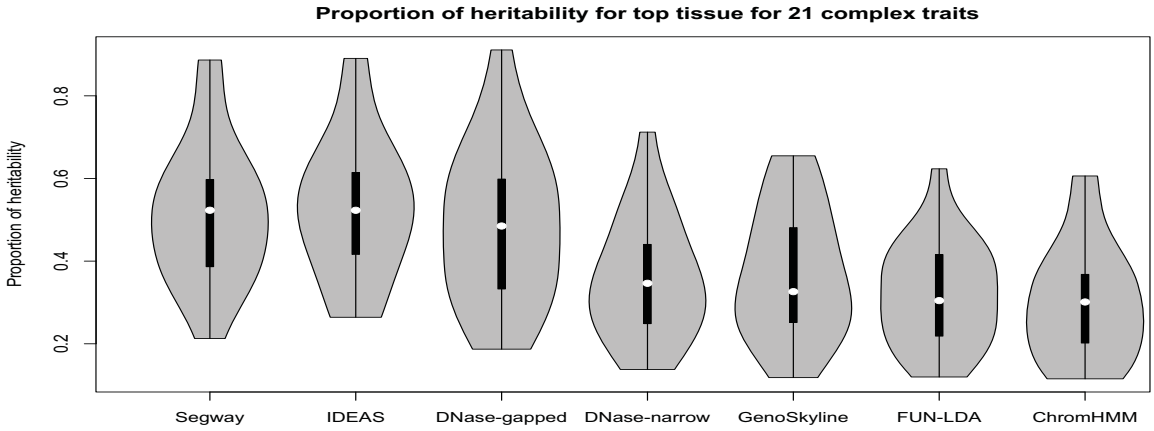


FIGURE S8. Heatmap of LD score regression analysis. Each row presents top 20 tissues for one of the 21 complex traits. Each cell represents the -log10(p value) standardized within trait.



(A) Enrichments estimates.



(B) Proportion of heritability.

FIGURE S9. Enrichment estimates (the proportion of SNP heritability in the functional component divided by the proportion of SNPs in that component) and proportion of heritability for different methods across top tissues for 21 complex traits. Estimates for DNase are omitted since they do not make sense for continuous annotations, such as quantitative DNase.

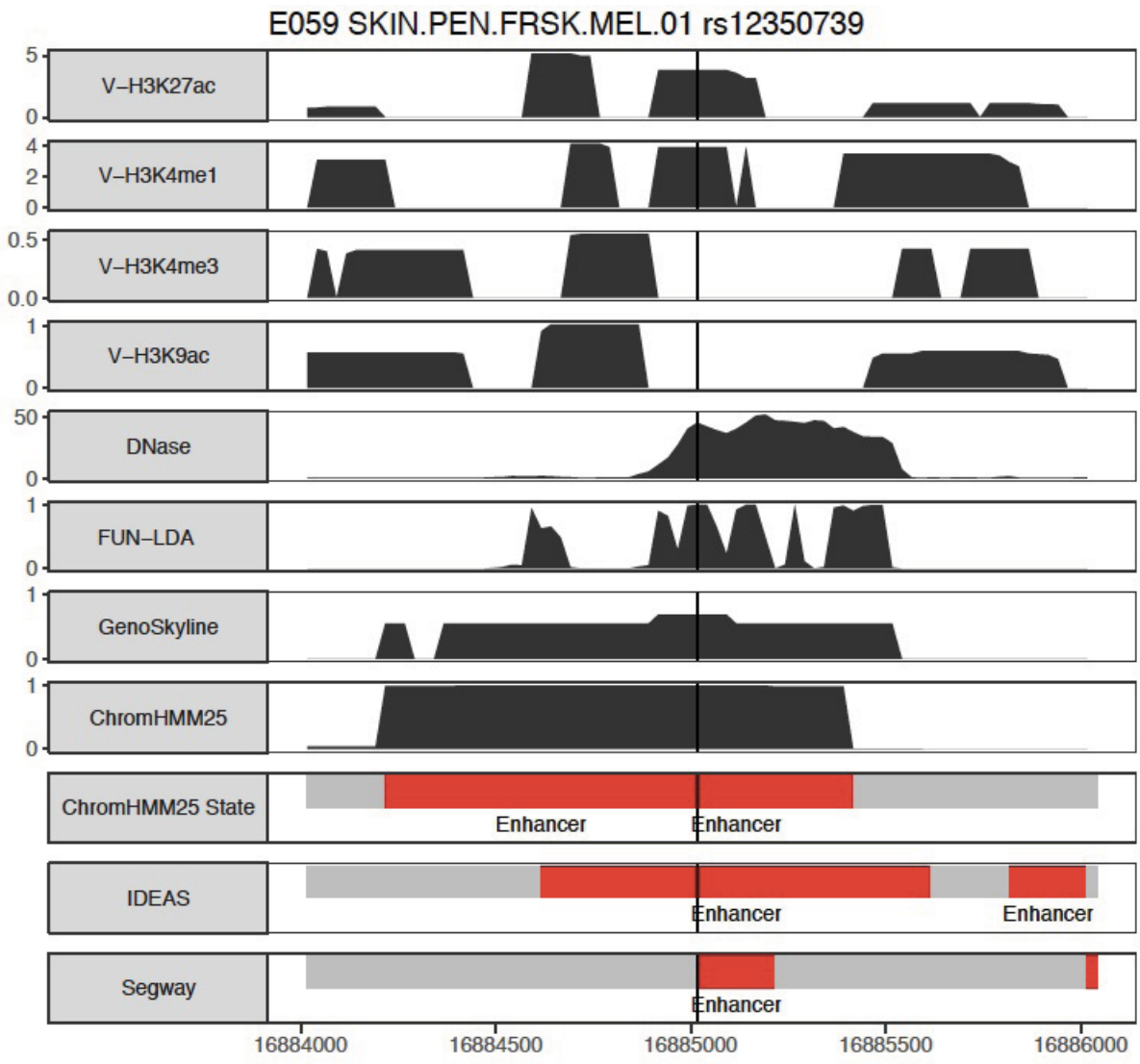


FIGURE S10. rs12350739 in Roadmap tissue E059. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

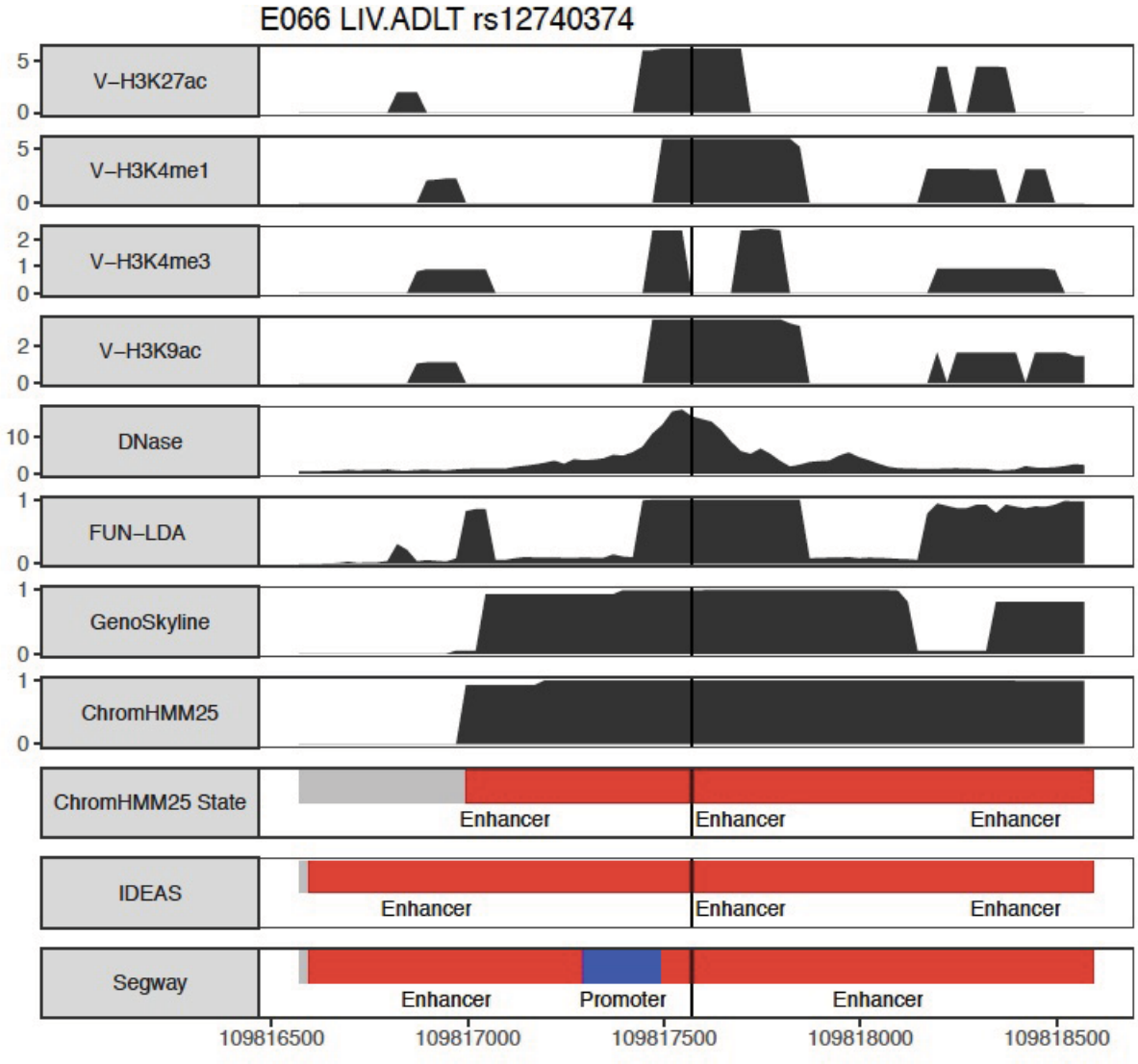


FIGURE S11. rs12740374 in Roadmap tissue E066. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

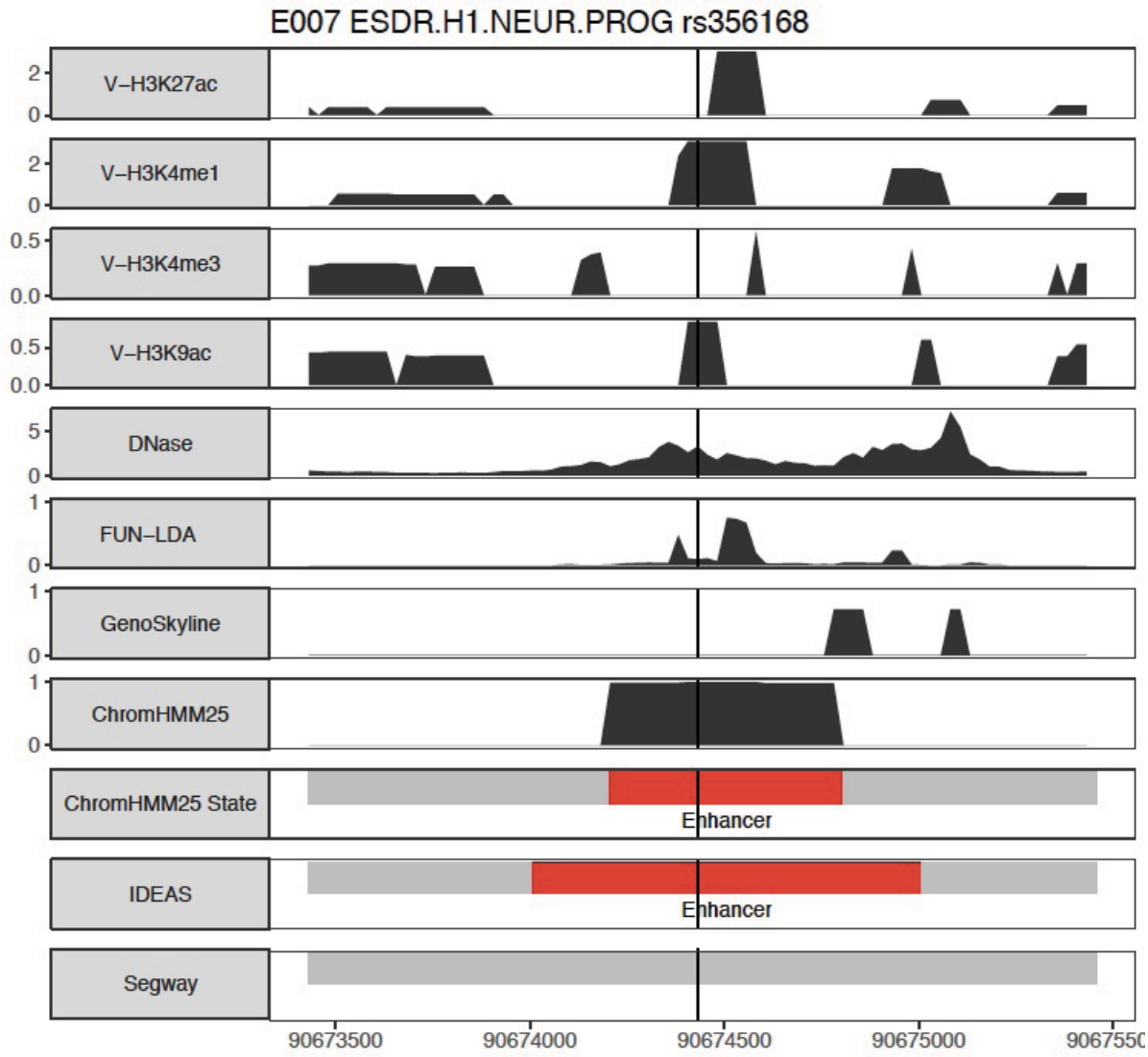


FIGURE S12. rs356168 in Roadmap tissue E007. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

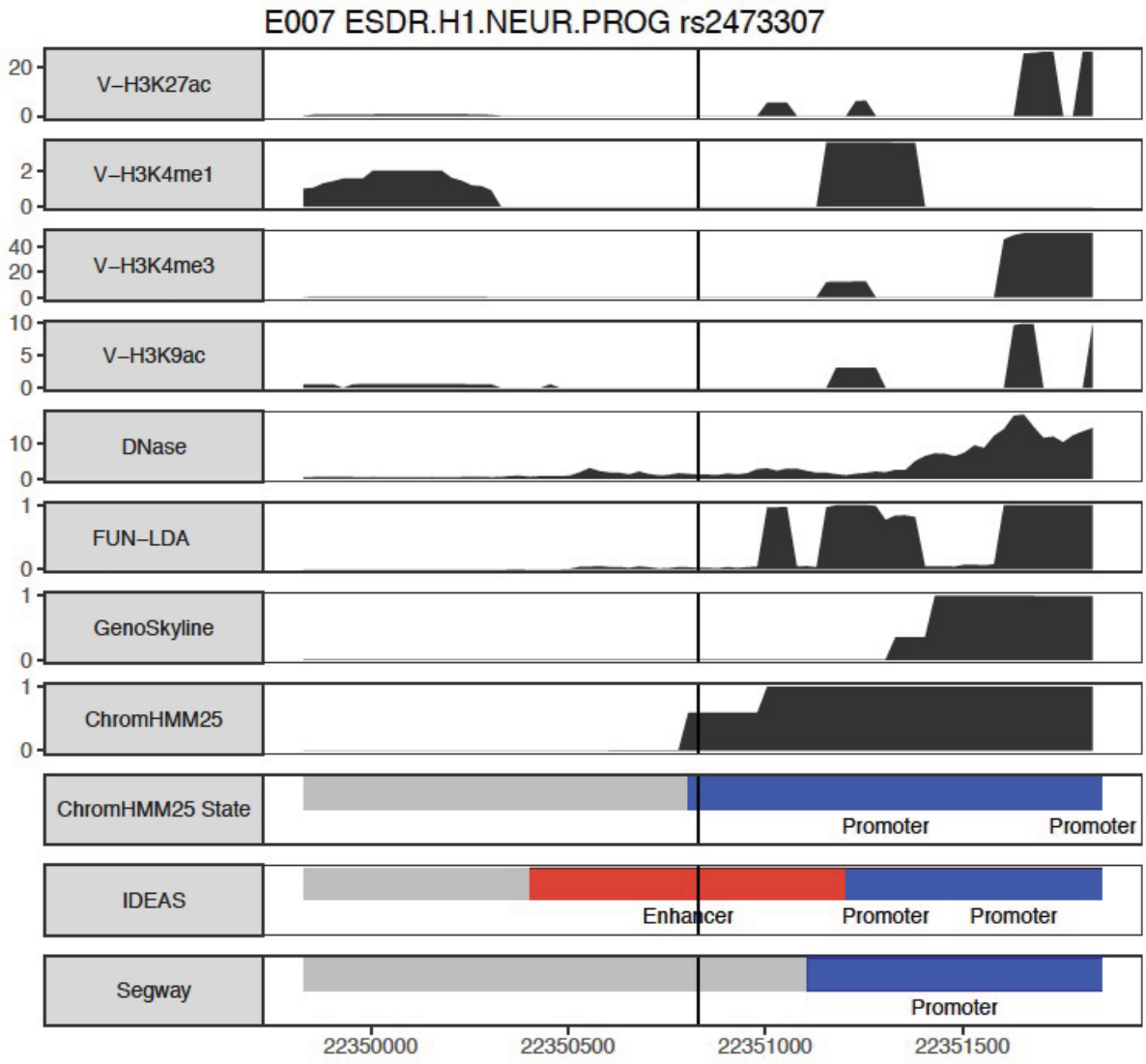


FIGURE S13. rs2473307 in Roadmap tissue E007. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

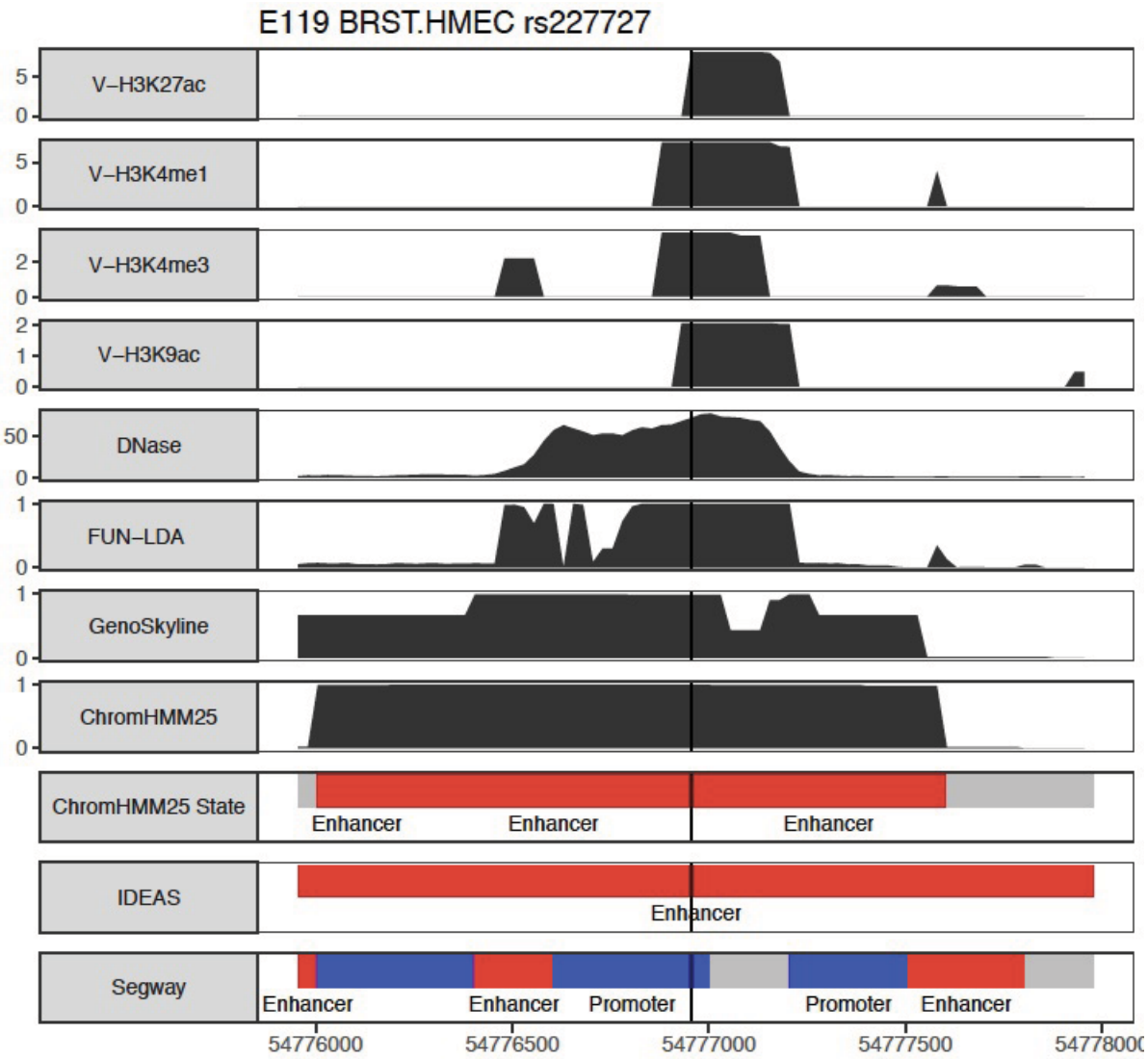


FIGURE S14. rs227727 in Roadmap tissue E119. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

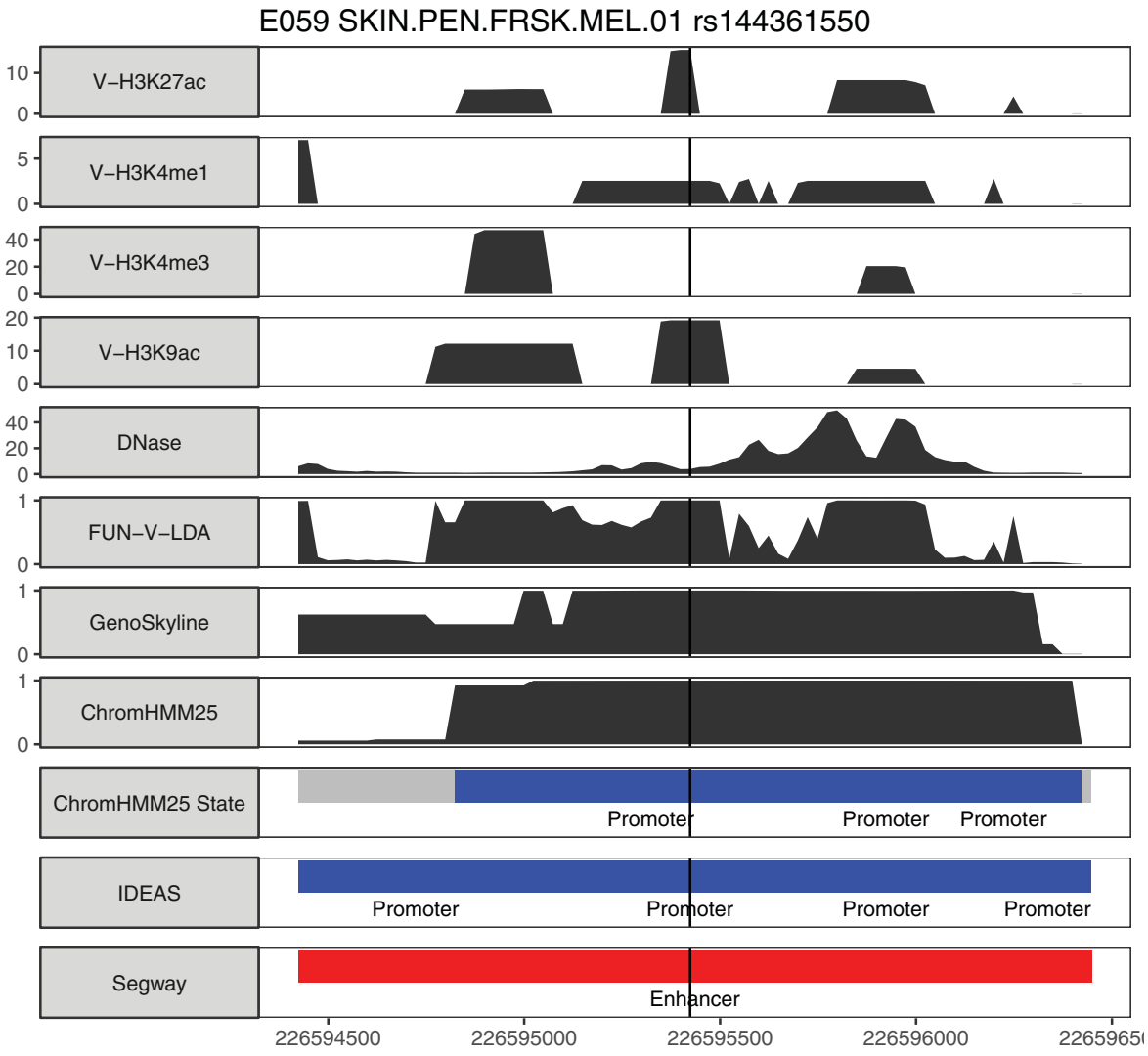


FIGURE S15. rs144361550 in Roadmap tissue E059. Valley scores for four activating histone marks and DNase, posterior probabilities from FUN-LDA, GenoSkyline, and ChromHMM (25 state model), and segmentations from ChromHMM, IDEAS and Segway are shown in 2 kb windows centered around the lead SNPs. For clarity we only highlight in the segmentations the type of states we consider functional (enhancer states in red, promoter states in blue) for the different segmentation approaches.

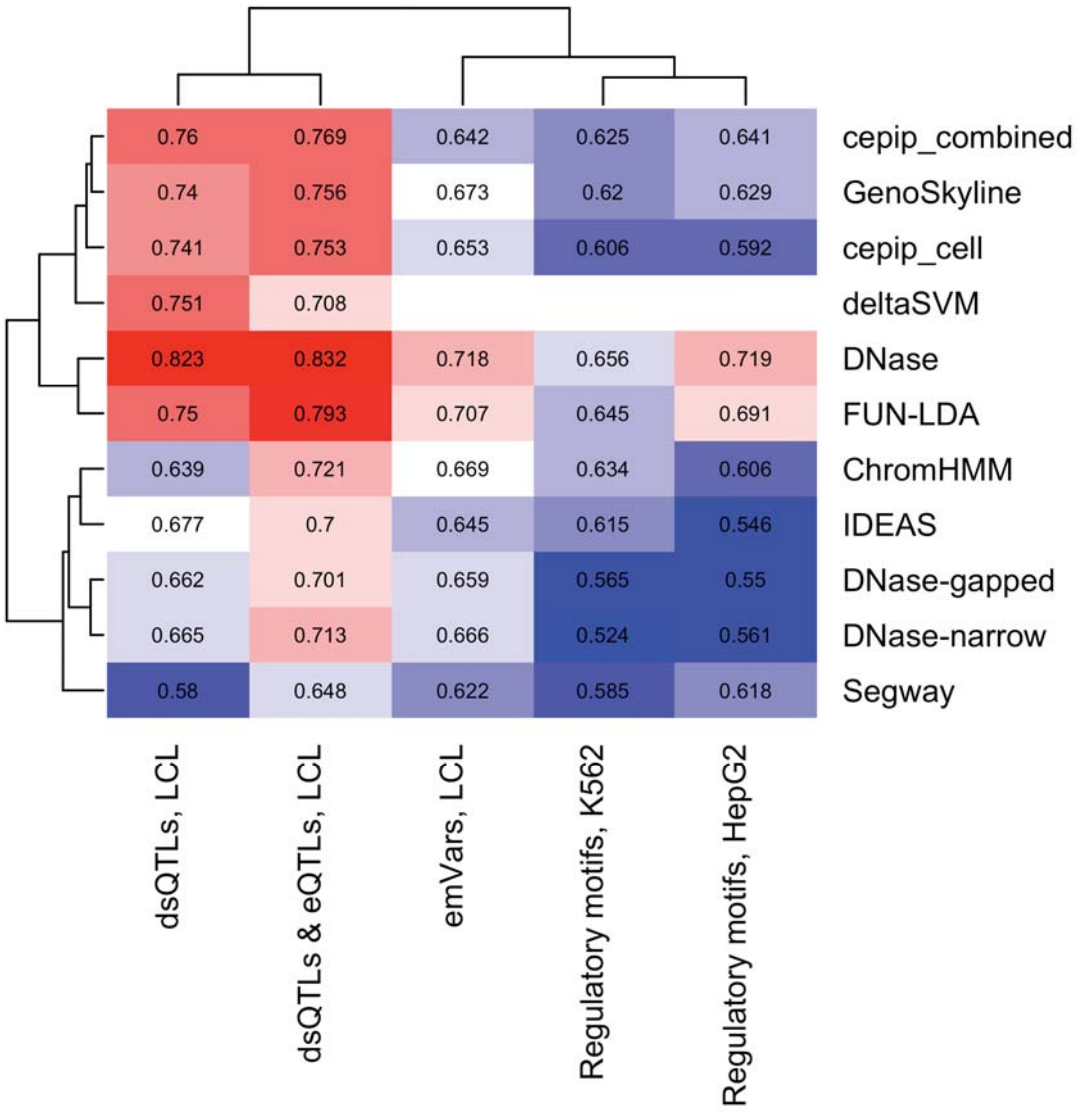


FIGURE S16. Tissue/cell type specific functional predictions. AUROC values for discriminating between variants likely to be functional and control variants. Results are shown for several datasets (three different cell lines) with experimental validation (MPRA) of potential regulatory variants, and one dsQTL dataset (dsQTLs & eQTLs contains a subset of dsQTLs that are also eQTLs). Methods include FUN-LDA, GenoSkyline, ChromHMM (25 state model), Segway, IDEAS, DNase (quantitative, -narrow and -gapped), cepip, and deltaSVM (note that deltaSVM predictions are only available for the dsQTL dataset). Each row represents a method and each column represents a dataset. The methods and datasets are grouped using hierarchical cluster analysis.

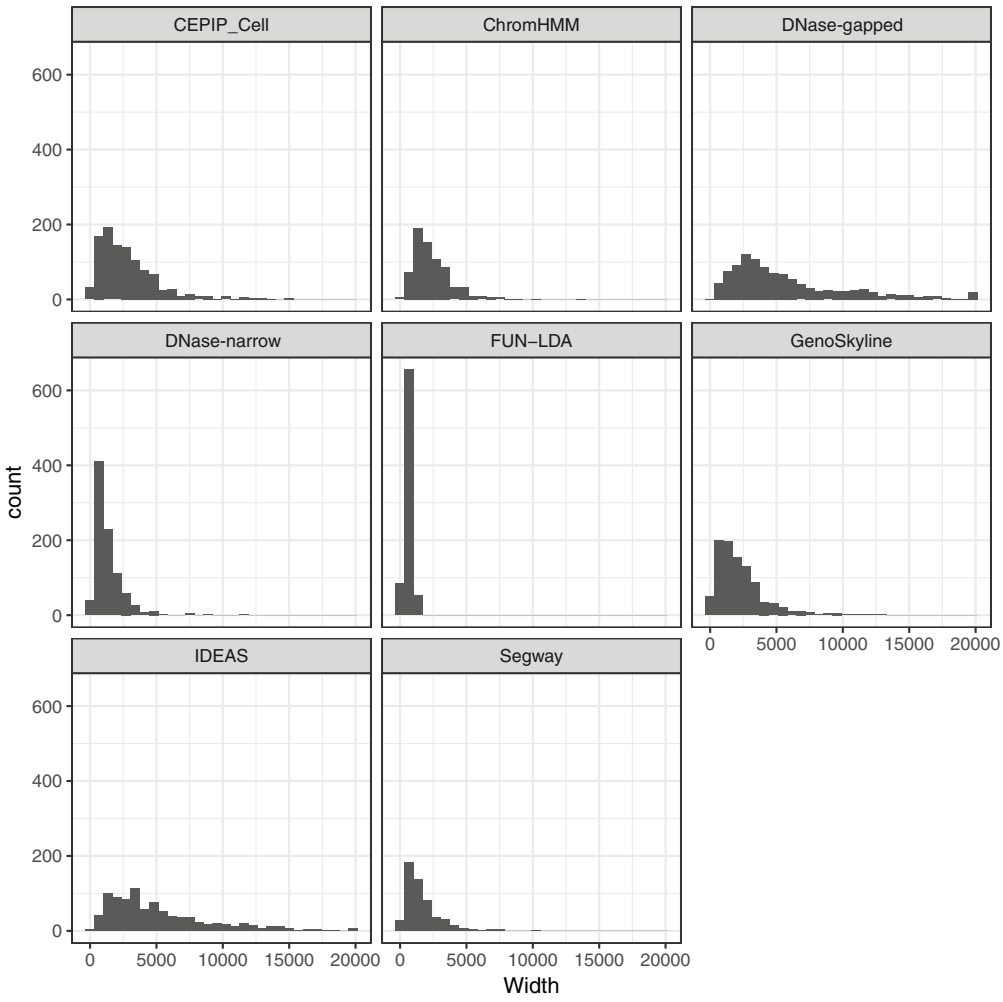


FIGURE S17. Widths of predicted functional regions (in bps) including validated functional variants from[8,9] and the eight confirmed variants in Table S8.

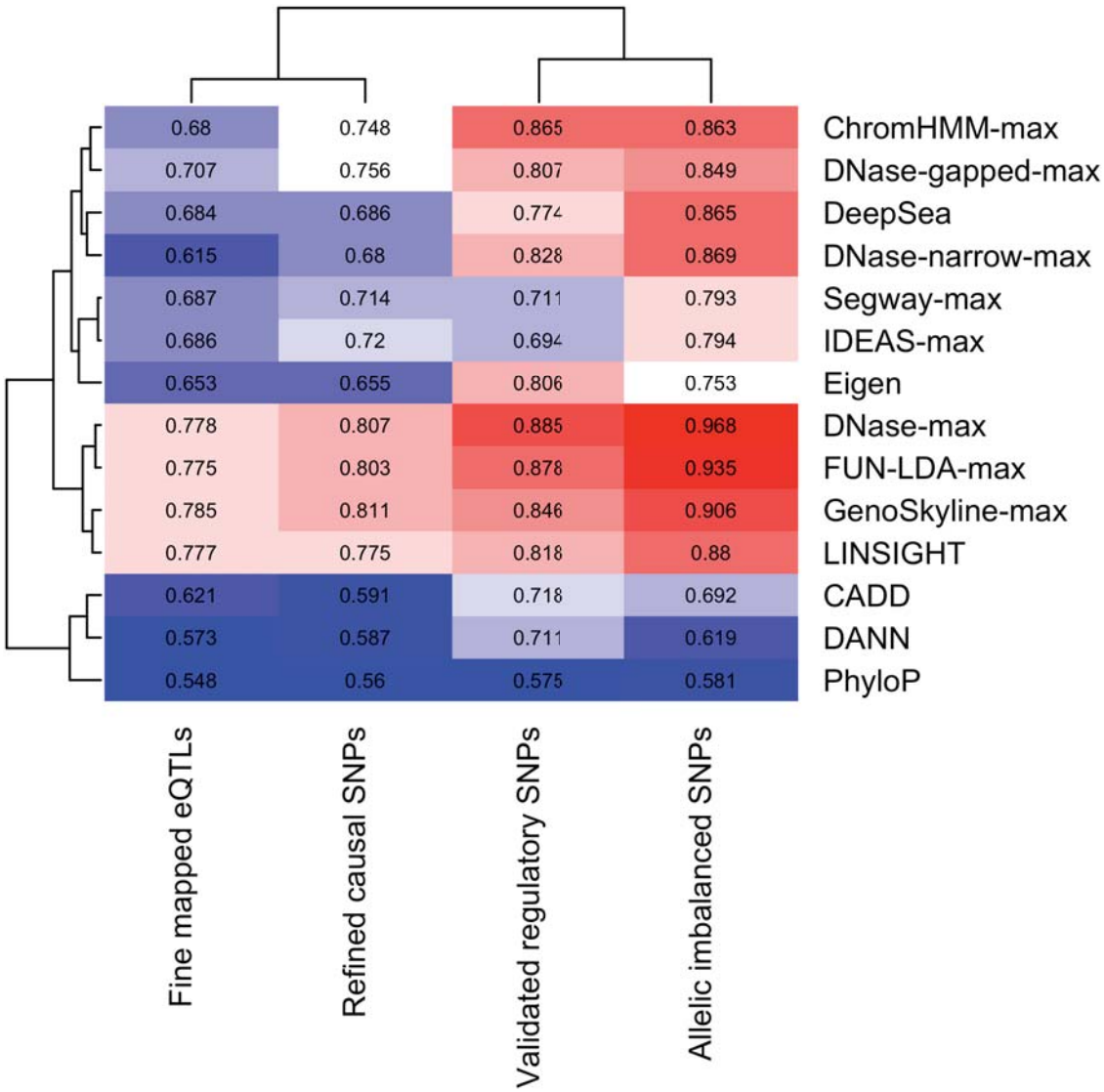


FIGURE S18. Organism level functional prediction. AUROC values for discriminating between variants likely to be functional and control variants for four non-tissue specific datasets. Methods include FUN-LDA-max (maximum across 127 different tissues), GenoSkyline-max, ChromHMM-max (25 state model), Segway-max, IDEAS-max, and DNase-max (quantitative, -narrow and -gapped). In addition, results for several organism level functional prediction methods, including phyloP, Eigen, CADD, DANN, DeepSea and LINSIGHT, are also reported. Each row represents a method and each column represents a dataset. The methods and datasets are grouped using hierarchical cluster analysis.

TABLE S1. Definition of the functional class for the five integrative methods considered.

Method	Functional Class Definition
FUN-LDA	States 1 and 2 (active promoters and enhancers) in Figure S1
GenoSkyline	The functional class as defined in[10]
ChromHMM (25 state model)	1_TssA, 2_PromU, 3_PromD1, 4_PromD2, 13_EnhA1, 14_EnhA2, 15_EnhAF
Segway	Promoters and Enhancers[11]
IDEAS	4_Enh, 6_EnhG, 8_TssAFlnk, 10_TssA, 14_TssWk,17_EnhGA[12]

TABLE S2. GTEx tissues and sample sizes.

Tissue	Sample size
Muscle - Skeletal	361
Whole Blood	338
Skin - Sun Exposed (Lower leg)	302
Adipose - Subcutaneous	298
Artery - Tibial	285
Lung	278
Thyroid	278
Cells - Transformed fibroblasts	272
Nerve - Tibial	256
Esophagus - Mucosa	241
Esophagus - Muscularis	218
Artery - Aorta	197
Skin - Not Sun Exposed (Suprapubic)	196
Heart - Left Ventricle	190
Adipose - Visceral (Omentum)	185
Breast - Mammary Tissue	183
Stomach	170
Colon - Transverse	169
Heart - Atrial Appendage	159
Testis	157
Pancreas	149
Esophagus - Gastroesophageal Junction	127
Adrenal Gland	126
Colon - Sigmoid	124
Artery - Coronary	118
Cells - EBV-transformed lymphocytes	114
Brain - Cerebellum	103
Brain - Caudate (basal ganglia)	100
Liver	97
Brain - Cortex	96
Brain - Nucleus accumbens (basal ganglia)	93
Brain - Frontal Cortex (BA9)	92
Brain - Cerebellar Hemisphere	89
Spleen	89
Pituitary	87
Prostate	87
Ovary	85
Brain - Putamen (basal ganglia)	82
Brain - Hippocampus	81
Brain - Hypothalamus	81
Vagina	79
Small Intestine - Terminal Ileum	77
Brain - Anterior cingulate cortex (BA24)	72
Uterus	70
Brain - Amygdala	62
Brain - Spinal cord (cervical c-1)	59
Brain - Substantia nigra	56
Minor Salivary Gland	51
Kidney - Cortex	26
Bladder	11
Cervix - Ectocervix	6
Fallopian Tube	6
Cervix - Endocervix	5

TABLE S3. Tissues and Cell Types in Roadmap (part 1)

Epigenome.ID	Epigenome.Mnemonic	Standardized.Epigenome.name
E017	LNG.IMR90	IMR90 fetal lung fibroblasts Cell Line
E002	ESC.WA7	ES-WA7 Cells
E008	ESC.H9	H9 Cells
E001	ESC.I3	ES-I3 Cells
E015	ESC.HUES6	HUES6 Cells
E014	ESC.HUES48	HUES48 Cells
E016	ESC.HUES64	HUES64 Cells
E003	ESC.H1	H1 Cells
E024	ESC.4STAR	ES-UCSF4 Cells
E020	IPSC.20B	iPS-20b Cells
E019	IPSC.18	iPS-18 Cells
E018	IPSC.15b	iPS-15b Cells
E021	IPSC.DF.6.9	iPS DF 6.9 Cells
E022	IPSC.DF.19.11	iPS DF 19.11 Cells
E007	ESDR.H1.NEUR.PROG	H1 Derived Neuronal Progenitor Cultured Cells
E009	ESDR.H9.NEUR.PROG	H9 Derived Neuronal Progenitor Cultured Cells
E010	ESDR.H9.NEUR	H9 Derived Neuron Cultured Cells
E013	ESDR.CD56.MESO	hESC Derived CD56+ Mesoderm Cultured Cells
E012	ESDR.CD56.ECTO	hESC Derived CD56+ Ectoderm Cultured Cells
E011	ESDR.CD184.ENDO	hESC Derived CD184+ Endoderm Cultured Cells
E004	ESDR.H1.BMP4.MESO	H1 BMP4 Derived Mesendoderm Cultured Cells
E005	ESDR.H1.BMP4.TROP	H1 BMP4 Derived Trophoblast Cultured Cells
E006	ESDR.H1.MSC	H1 Derived Mesenchymal Stem Cells
E062	BLD.PER.MONUC.PC	Primary mononuclear cells from peripheral blood
E034	BLD.CD3.PPC	Primary T cells fromperipheralblood
E045	BLD.CD4.CD25I.CD127.TMEMPC	Primary T cells effector/memory enriched from peripheral blood
E033	BLD.CD3.CPC	Primary T cells from cord blood
E044	BLD.CD4.CD25.CD127M.TREGPC	Primary T regulatory cells fromperipheralblood
E043	BLD.CD4.CD25M.TPC	Primary T helper cells fromperipheralblood
E039	BLD.CD4.CD25M.CD45RA.NPC	Primary T helper naive cells fromperipheralblood
E041	BLD.CD4.CD25M.IL17M.PL.TPC	Primary T helper cells PMA-I stimulated
E042	BLD.CD4.CD25M.IL17P.PL.TPC	Primary T helper 17 cells PMA-I stimulated
E040	BLD.CD4.CD25M.CD45RO.MPC	Primary T helper memory cells from peripheral blood 1
E037	BLD.CD4.MPC	Primary T helper memory cells from peripheral blood 2
E048	BLD.CD8.MPC	Primary T CD8+ memory cells from peripheral blood
E038	BLD.CD4.NPC	Primary T helper naive cells from peripheral blood
E047	BLD.CD8.NPC	Primary T CD8+ naive cells from peripheral blood
E029	BLD.CD14.PC	Primary monocytes from peripheral blood
E031	BLD.CD19.CPC	Primary B cells from cord blood
E035	BLD.CD34.PC	Primary hematopoietic stem cells
E051	BLD.MOB.CD34.PC.M	Primary hematopoietic stem cells G-CSF-mobilized Male
E050	BLD.MOB.CD34.PC.F	Primary hematopoietic stem cells G-CSF-mobilized Female
E036	BLD.CD34.CC	Primary hematopoietic stem cells short term culture
E032	BLD.CD19.PPC	Primary B cells from peripheral blood
E046	BLD.CD56.PC	Primary Natural Killer cells from peripheral blood
E030	BLD.CD15.PC	Primary neutrophils from peripheral blood
E026	STRM.MRW.MSC	Bone Marrow Derived Cultured Mesenchymal Stem Cells
E049	STRM.CHON.MRW.DR.MSC	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells
E025	FAT.ADIP.DR.MSC	Adipose Derived Mesenchymal Stem Cell Cultured Cells
E023	FAT.MSC.DR.ADIP	Mesenchymal Stem Cell Derived Adipocyte Cultured Cells
E052	MUS.SAT	Muscle Satellite Cultured Cells
E055	SKIN.PEN.FRSK.FIB.01	Foreskin Fibroblast Primary Cells skin01
E056	SKIN.PEN.FRSK.FIB.02	Foreskin Fibroblast Primary Cells skin02
E059	SKIN.PEN.FRSK.MEL.01	Foreskin Melanocyte Primary Cells skin01
E061	SKIN.PEN.FRSK.MEL.03	Foreskin Melanocyte Primary Cells skin03
E057	SKIN.PEN.FRSK.KER.02	Foreskin Keratinocyte Primary Cells skin02
E058	SKIN.PEN.FRSK.KER.03	Foreskin Keratinocyte Primary Cells skin03
E028	BRST.HMEC.35	Breast variant Human Mammary Epithelial Cells (vHMEC)
E027	BRST.MYO	Breast Myoepithelial Primary Cells
E054	BRN.GANGEM.DR.NRSPHR	Ganglion Eminence derived primary cultured neurospheres
E053	BRN.CRTX.DR.NRSPHR	Cortex derived primary cultured neurospheres
E112	THYM	Thymus
E093	THYM.FET	Fetal Thymus

Table S4. Tissues and Cell Types in Roadmap (part 2)

Epigenome.ID	Epigenome.Mnemonic	Standardized.Epigenome.name
E071	BRN.HIPP.MID	Brain Hippocampus Middle
E074	BRN.SUB.NIG	Brain Substantia Nigra
E068	BRN.ANT.CAUD	Brain Anterior Caudate
E069	BRN.CING.GYR	Brain Cingulate Gyrus
E072	BRN.INF.TMP	Brain Inferior Temporal Lobe
E067	BRN.ANG.GYR	Brain Angular Gyrus
E073	BRN.DL.PRFRNTL.CRTX	Brain_Dorsolateral_Prefrontal_Cortex
E070	BRN.GRM.MTRX	Brain Germinal Matrix
E082	BRN.FET.F	Fetal Brain Female
E081	BRN.FET.M	Fetal Brain Male
E063	FAT.ADIP.NUC	Adipose Nuclei
E100	MUS.PSOAS	Psoas Muscle
E108	MUS.SKLT.F	Skeletal Muscle Female
E107	MUS.SKLT.M	Skeletal Muscle Male
E089	MUS.TRNK.FET	Fetal Muscle Trunk
E090	MUS.LEG.FET	Fetal Muscle Leg
E083	HRT.FET	Fetal Heart
E104	HRT.ATR.R	Right Atrium
E095	HRT.VENT.L	Left Ventricle
E105	HRT.VNT.R	Right Ventricle
E065	VAS.AOR	Aorta
E078	GI.DUO.SM.MUS	Duodenum Smooth Muscle
E076	GI.CLN.SM.MUS	Colon Smooth Muscle
E103	GI.RECT.SM.MUS	Rectal Smooth Muscle
E111	GI.STMC.MUS	Stomach Smooth Muscle
E092	GI.STMC.FET	Fetal Stomach
E085	GI.S.INT.FET	Fetal Intestine Small
E084	GI.L.INT.FET	Fetal Intestine Large
E109	GI.S.INT	Small Intestine
E106	GI.CLN.SIG	Sigmoid Colon
E075	GI.CLN.MUC	Colonic Mucosa
E101	GI.RECT.MUC.29	Rectal Mucosa Donor 29
E102	GI.RECT.MUC.31	Rectal Mucosa Donor 31
E110	GI.STMC.MUC	Stomach Mucosa
E077	GI.DUO.MUC	Duodenum Mucosa
E079	GI.ESO	Esophagus
E094	GI.STMC.GAST	Gastric
E099	PLCNT.AMN	Placenta Amnion
E086	KID.FET	Fetal Kidney
E088	LNG.FET	Fetal Lung
E097	OVRY	Ovary
E087	PANC.ISLT	Pancreatic Islets
E080	ADRL.GLND.FET	Fetal Adrenal Gland
E091	PLCNT.FET	Placenta
E066	LIV.ADLT	Liver
E098	PANC	Pancreas
E096	LNG	Lung
E113	SPLN	Spleen
E114	LNG.A549.ETOH002.CNCR	A549 EtOH 0.02pct Lung Carcinoma Cell Line
E115	BLD.DND41.CNCR	Dnd41 TCell Leukemia Cell Line
E116	BLD.GM12878	GM12878 Lymphoblastoid Cells
E117	CRVX.HELAS3.CNCR	HeLa-S3 Cervical Carcinoma Cell Line
E118	LIV.HEPG2.CNCR	HepG2 Hepatocellular Carcinoma Cell Line
E119	BRST.HMEC	HMEC Mammary Epithelial Primary Cells
E120	MUS.HSMM	HSMM Skeletal Muscle Myoblasts Cells
E121	MUS.HSMMT	HSMM cell derived Skeletal Muscle Myotubes Cells
E122	VAS.HUVEC	HUVEC Umbilical Vein Endothelial Primary Cells
E123	BLD.K562.CNCR	K562 Leukemia Cells
E124	BLD.CD14.MONO	Monocytes-CD14+ RO01746 Primary Cells
E125	BRN.NHA	NH-A Astrocytes Primary Cells
E126	SKIN.NHDFAD	NHDF-Ad Adult Dermal Fibroblast Primary Cells
E127	SKIN.NHEK	NHEK-Epidermal Keratinocyte Primary Cells
E128	LNG.NHLF	NHLF Lung Fibroblast Primary Cells
E129	BONE.OSTEO	Osteoblast Primary Cells

TABLE S5. Results from stratified LD score regression for the different methods (part 1).

Trait	Method	Roadmap Epigenome Name	-log10(p)
AgeAtMenarche	ChromHMM	Cortex derived primary cultured neurospheres	4.31
AgeAtMenarche	DNase	hESC Derived CD56+ Ectoderm Cultured Cells	4.76
AgeAtMenarche	DNase-gapped	iPS DF 6.9 Cells	4.16
AgeAtMenarche	DNase-narrow	ES-UCSF4 Cells	7.36
AgeAtMenarche	FUN-LDA	H9 Derived Neuron Cultured Cells	6.15
AgeAtMenarche	GenoSkyline	H1 Derived Neuronal Progenitor Cultured Cells	7.96
AgeAtMenarche	IDEAS	H1 Derived Neuronal Progenitor Cultured Cells	3.47
AgeAtMenarche	Segway	H1 Derived Neuronal Progenitor Cultured Cells	9.91
Alopecia	ChromHMM	Primary T helper cells PMA-I stimulated	3.31
Alopecia	DNase	Primary T helper 17 cells PMA-I stimulated	2.10
Alopecia	DNase-gapped	Primary T helper 17 cells PMA-I stimulated	4.04
Alopecia	DNase-narrow	Primary T helper memory cells from peripheral blood 1	3.81
Alopecia	FUN-LDA	Primary T cells from cord blood	3.90
Alopecia	GenoSkyline	Primary T helper memory cells from peripheral blood 2	3.23
Alopecia	IDEAS	Primary T helper 17 cells PMA-I stimulated	4.48
Alopecia	Segway	Primary T helper 17 cells PMA-I stimulated	5.27
Alzheimers	ChromHMM	Primary hematopoietic stem cells	1.86
Alzheimers	DNase	Monocytes-CD14+ RO01746 Primary Cells	2.05
Alzheimers	DNase-gapped	Primary hematopoietic stem cells G-CSF-mobilized Male	3.96
Alzheimers	DNase-narrow	Primary hematopoietic stem cells G-CSF-mobilized Male	3.59
Alzheimers	FUN-LDA	Primary hematopoietic stem cells G-CSF-mobilized Male	3.78
Alzheimers	GenoSkyline	Monocytes-CD14+ RO01746 Primary Cells	2.91
Alzheimers	IDEAS	Primary hematopoietic stem cells G-CSF-mobilized Male	4.06
Alzheimers	Segway	Primary hematopoietic stem cells G-CSF-mobilized Male	3.79
Autism	ChromHMM	Fetal Brain Female	1.19
Autism	DNase	Primary monocytes from peripheral blood	1.64
Autism	DNase-gapped	Primary monocytes from peripheral blood	2.16
Autism	DNase-narrow	Monocytes-CD14+ RO01746 Primary Cells	1.94
Autism	FUN-LDA	Primary monocytes from peripheral blood	2.41
Autism	GenoSkyline	Brain Dorsolateral Prefrontal Cortex	1.26
Autism	IDEAS	Liver	2.54
Autism	Segway	Monocytes-CD14+ RO01746 Primary Cells	2.34
BipolarDisorder	ChromHMM	Primary monocytes from peripheral blood	2.27
BipolarDisorder	DNase	Monocytes-CD14+ RO01746 Primary Cells	2.23
BipolarDisorder	DNase-gapped	Monocytes-CD14+ RO01746 Primary Cells	3.48
BipolarDisorder	DNase-narrow	Monocytes-CD14+ RO01746 Primary Cells	2.48
BipolarDisorder	FUN-LDA	Fetal Brain Female	3.20
BipolarDisorder	GenoSkyline	Psoas Muscle	3.73
BipolarDisorder	IDEAS	Fetal Brain Male	3.30
BipolarDisorder	Segway	Brain Dorsolateral Prefrontal Cortex	3.70
BMI	ChromHMM	Fetal Brain Female	2.94
BMI	DNase	ES-UCSF4 Cells	1.12
BMI	DNase-gapped	ES-UCSF4 Cells	2.58
BMI	DNase-narrow	ES-UCSF4 Cells	4.29
BMI	FUN-LDA	Brain Germinal Matrix	4.79
BMI	GenoSkyline	Brain Dorsolateral Prefrontal Cortex	6.47
BMI	IDEAS	Brain Angular Gyrus	4.44
BMI	Segway	iPS DF 19.11 Cells	4.49
CoronaryArteryDisease	ChromHMM	Liver	3.38
CoronaryArteryDisease	DNase	Liver	2.62
CoronaryArteryDisease	DNase-gapped	Liver	4.67
CoronaryArteryDisease	DNase-narrow	Lung	3.51
CoronaryArteryDisease	FUN-LDA	Liver	4.61
CoronaryArteryDisease	GenoSkyline	Lung	4.25
CoronaryArteryDisease	IDEAS	Adipose Nuclei	3.65
CoronaryArteryDisease	Segway	Small Intestine	5.70

TABLE S6. Results from stratified LD score regression for the different methods (part 2).

Trait	Method	Roadmap Epigenome Name	-log10(p)
CrohnsDisease	ChromHMM	Primary T helper 17 cells PMA-I stimulated	6.39
CrohnsDisease	DNase	Primary T helper cells PMA-I stimulated	3.84
CrohnsDisease	DNase-gapped	Primary B cells from peripheral blood	6.89
CrohnsDisease	DNase-narrow	Primary T helper 17 cells PMA-I stimulated	6.90
CrohnsDisease	FUN-LDA	Primary B cells from cord blood	6.25
CrohnsDisease	GenoSkyline	Primary Natural Killer cells from peripheral blood	4.95
CrohnsDisease	IDEAS	Primary T helper memory cells from peripheral blood 1	7.60
CrohnsDisease	Segway	Primary T helper 17 cells PMA-I stimulated	7.53
EducationalAttainment	ChromHMM	Fetal Brain Female	4.74
EducationalAttainment	DNase	Fetal Brain Female	3.05
EducationalAttainment	DNase-gapped	Cortex derived primary cultured neurospheres	4.27
EducationalAttainment	DNase-narrow	Fetal Brain Female	3.07
EducationalAttainment	FUN-LDA	Fetal Brain Female	5.84
EducationalAttainment	GenoSkyline	Brain Dorsolateral Prefrontal Cortex	3.61
EducationalAttainment	IDEAS	Fetal Brain Female	7.32
EducationalAttainment	Segway	Fetal Brain Male	5.55
Epilepsy	ChromHMM	Brain Angular Gyrus	2.91
Epilepsy	DNase	Dnd41 TCell Leukemia Cell Line	0.99
Epilepsy	DNase-gapped	Brain Hippocampus Middle	2.36
Epilepsy	DNase-narrow	Fetal Thymus	1.85
Epilepsy	FUN-LDA	Brain Anterior Caudate	4.11
Epilepsy	GenoSkyline	Brain Inferior Temporal Lobe	3.35
Epilepsy	IDEAS	Brain Angular Gyrus	4.40
Epilepsy	Segway	Brain Angular Gyrus	4.51
EverSmoked	ChromHMM	Primary T cells effector/memory enriched from peripheral blood	2.15
EverSmoked	DNase	Brain Inferior Temporal Lobe	0.61
EverSmoked	DNase-gapped	Brain Inferior Temporal Lobe	1.31
EverSmoked	DNase-narrow	Primary hematopoietic stem cells	0.78
EverSmoked	FUN-LDA	Brain Inferior Temporal Lobe	2.68
EverSmoked	GenoSkyline	Brain Inferior Temporal Lobe	2.94
EverSmoked	IDEAS	Brain Angular Gyrus	3.66
EverSmoked	Segway	Brain Inferior Temporal Lobe	4.16
FastingGlucose	ChromHMM	Pancreatic Islets	1.44
FastingGlucose	DNase	Fetal Intestine Small	1.03
FastingGlucose	DNase-gapped	Pancreatic Islets	2.03
FastingGlucose	DNase-narrow	iPS-15b Cells	1.60
FastingGlucose	FUN-LDA	Pancreatic Islets	1.45
FastingGlucose	GenoSkyline	H9 Cells	2.29
FastingGlucose	IDEAS	Pancreatic Islets	3.65
FastingGlucose	Segway	Pancreatic Islets	3.85
HDL	ChromHMM	Primary monocytes from peripheral blood	2.72
HDL	DNase	Liver	3.94
HDL	DNase-gapped	Adipose Nuclei	5.15
HDL	DNase-narrow	Adipose Nuclei	4.37
HDL	FUN-LDA	Liver	4.73
HDL	GenoSkyline	Liver	3.67
HDL	IDEAS	Adipose Nuclei	5.63
HDL	Segway	Liver	4.28
Height	ChromHMM	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	5.55
Height	DNase	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	4.45
Height	DNase-gapped	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	9.99
Height	DNase-narrow	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	10.81
Height	FUN-LDA	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	12.28
Height	GenoSkyline	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	11.31
Height	IDEAS	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	14.59
Height	Segway	Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	13.40

TABLE S7. Results from stratified LD score regression for the different methods (part 3).

Trait	Method	Roadmap Epigenome Name	-log10(p)
IGAN	ChromHMM	Dnd41 TCell Leukemia Cell Line	2.35
IGAN	DNase	Monocytes-CD14+ RO01746 Primary Cells	1.59
IGAN	DNase-gapped	Primary T cells from peripheral blood	3.86
IGAN	DNase-narrow	Primary T helper memory cells from peripheral blood 2	4.13
IGAN	FUN-LDA	Primary Natural Killer cells from peripheral blood	3.28
IGAN	GenoSkyline	Primary mononuclear cells from peripheral blood	3.64
IGAN	IDEAS	Primary T cells from peripheral blood	3.65
IGAN	Segway	Primary Natural Killer cells from peripheral blood	3.23
LDL	ChromHMM	Liver	3.25
LDL	DNase	Liver	1.61
LDL	DNase-gapped	Liver	3.68
LDL	DNase-narrow	Fetal Adrenal Gland	2.64
LDL	FUN-LDA	Liver	4.08
LDL	GenoSkyline	Liver	4.37
LDL	IDEAS	Liver	5.06
LDL	Segway	Liver	4.29
RheumatoidArthritis	ChromHMM	GM12878 Lymphoblastoid Cells	8.25
RheumatoidArthritis	DNase	Primary T helper cells PMA-I stimulated	4.27
RheumatoidArthritis	DNase-gapped	Primary T helper cells PMA-I stimulated	7.60
RheumatoidArthritis	DNase-narrow	Primary T helper cells PMA-I stimulated	7.51
RheumatoidArthritis	FUN-LDA	GM12878 Lymphoblastoid Cells	6.93
RheumatoidArthritis	GenoSkyline	Primary B cells from peripheral blood	5.83
RheumatoidArthritis	IDEAS	GM12878 Lymphoblastoid Cells	8.84
RheumatoidArthritis	Segway	Primary T helper 17 cells PMA-I stimulated	7.93
Schizophrenia	ChromHMM	Fetal Brain Female	11.88
Schizophrenia	DNase	Brain Germinal Matrix	6.64
Schizophrenia	DNase-gapped	Fetal Brain Female	9.01
Schizophrenia	DNase-narrow	Fetal Brain Female	9.12
Schizophrenia	FUN-LDA	Fetal Brain Female	14.70
Schizophrenia	GenoSkyline	Brain Dorsolateral Prefrontal Cortex	8.95
Schizophrenia	IDEAS	Fetal Brain Male	Inf
Schizophrenia	Segway	Fetal Brain Male	Inf
Triglycerides	ChromHMM	Liver	3.49
Triglycerides	DNase	Liver	4.05
Triglycerides	DNase-gapped	Liver	4.89
Triglycerides	DNase-narrow	Liver	4.06
Triglycerides	FUN-LDA	Liver	4.11
Triglycerides	GenoSkyline	Liver	3.63
Triglycerides	IDEAS	Liver	4.30
Triglycerides	Segway	Liver	3.86
Type2Diabetes	ChromHMM	Fetal Kidney	1.79
Type2Diabetes	DNase	Fetal Intestine Small	1.27
Type2Diabetes	DNase-gapped	Pancreatic Islets	3.67
Type2Diabetes	DNase-narrow	HepG2 Hepatocellular Carcinoma Cell Line	2.91
Type2Diabetes	FUN-LDA	Pancreatic Islets	4.21
Type2Diabetes	GenoSkyline	Adipose Nuclei	2.18
Type2Diabetes	IDEAS	Fetal Intestine Small	3.03
Type2Diabetes	Segway	Pancreatic Islets	3.27
UlcerativeColitis	ChromHMM	Primary T helper 17 cells PMA-I stimulated	4.26
UlcerativeColitis	DNase	Primary T helper cells PMA-I stimulated	2.06
UlcerativeColitis	DNase-gapped	Primary T helper 17 cells PMA-I stimulated	3.95
UlcerativeColitis	DNase-narrow	Primary T helper 17 cells PMA-I stimulated	4.84
UlcerativeColitis	FUN-LDA	Primary T helper 17 cells PMA-I stimulated	4.45
UlcerativeColitis	GenoSkyline	Rectal Mucosa Donor 29	3.54
UlcerativeColitis	IDEAS	Primary T helper 17 cells PMA-I stimulated	4.97
UlcerativeColitis	Segway	Primary T helper 17 cells PMA-I stimulated	5.77

TABLE S8. For eight SNPs selected from literature, the tissue or cell type in the original study and the closest tissue in Roadmap that we selected are given.

SNP	Tissue in Functional Study	Selected Roadmap Tissue
rs6801957	murine heart tissue	E104 - Right Atrium
rs12821256	cultured human keratinocytes	E127 - NHEK-Epidermal Keratinocyte Primary Cells
rs12350739	skin epidermal samples/melanocyte cell lines	E059 - Foreskin Melanocyte Primary Cells skin01
rs12740374	primary hepatocytes	E066 - Liver
rs356168	hIPSC-derived neurons	E007 - H1 Derived Neuronal Progenitor Cultured Cells
rs2473307	human neuronal cell line	E007 - H1 Derived Neuronal Progenitor Cultured Cells
rs227727	human embryonic oral epithelial cells	E119 - HMEC Mammary Epithelial Primary Cells
rs144361550	primary melanocytes	E059 - Foreskin Melanocyte Primary Cells skin01

TABLE S9. AUROC for various integrative methods vs. individual epigenetic annotations using MPRA validated variants.

Method	Type	emVars	Regulatory motifs	Regulatory motifs
		E116	E118	E123
FUN-LDA		0.709	0.694	0.646
GenoSkyline		0.674	0.630	0.619
ChromHMM	Integrative	0.668	0.608	0.634
Segway		0.624	0.618	0.585
IDEAS		0.621	0.546	0.615
DNase		0.722	0.719	0.654
DNase-narrow		0.629	0.561	0.524
DNase-gapped		0.653	0.550	0.565
H3K27ac	Single annotation	0.677	0.556	0.597
H3K4me1		0.664	0.545	0.578
H3K4me3		0.692	0.535	0.602
H3K9ac		0.670	0.549	0.615

TABLE S10. AUROC for the segmentation methods ChromHMM, Segway and IDEAS state combinations with maximum AUROC using the MPRA validated variants. Note that the selection of the best state combination is based on combining the variants from all three MPRA datasets in Section 2.4.

Method	TypeState	States in 'functional' group	emVars	Reg. motifs	Reg. motifs
			E116	E118	E123
FUN-LDA	Selected	1_ActiveEnhancers, 2_ActivePromoters	0.709	0.694	0.646
ChromHMM	Best	1_TssA, 2_PromU, 9_TxReg, 13_EnhA1 14_EnhA2, 16_EnhW1, 22_PromP	0.670	0.619	0.661
	Selected	1_TssA, 2_PromU, 3_PromD1, 4_PromD2 13_EnhA1, 14_EnhA2, 15_EnhAF	0.668	0.608	0.634
Segway	Best	Bivalent, RegPermissive, Enhancer, Promoter	0.650	0.591	0.630
	Selected	Enhancer, Promoter	0.624	0.618	0.585
IDEAS	Best	4_Enh, 8_TssAFlnk, 6_EnhG, 10_TssA 19_Enh/ReprPC,11_EnhBiv,15_TssBiv,14_TssWk,17_EnhGA	0.635	0.544	0.614
	Selected	4_Enh, 6_EnhG, 8_TssAFlnk, 10_TssA, 14_TssWk,17_EnhGA	0.621	0.546	0.615

TABLE S11. Tissue/cell type specific functional predictions. AUROC/AUPR values for discriminating between variants likely to be functional and control variants. Results are shown for validated enhancers in ultra conserved sequence elements[13]. Methods include FUN-LDA, GenoSkyline, ChromHMM (25 state model), Segway, IDEAS, and DNase (quantitative, -narrow and -gapped). The tissues with the highest AUROC/AUPR for each method are also shown. Most of the validated enhancers were shown to affect gene expression in embryonic tissue.

Dataset	Method	Top Tissue	AUROC
Ultra conserved	FUN-LDA	hESC Derived CD184+ Endoderm Cultured Cells	0.658
Elements	GenoSkyline	Primary hematopoietic stem cells	0.697
	ChromHMM	hESC Derived CD56+ Ectoderm Cultured Cells	0.604
	Segway	HUES6 Cells	0.588
	IDEAS	hESC Derived CD184+ Endoderm Cultured Cells	0.646
	DNase	hESC Derived CD184+ Endoderm Cultured Cells	0.629
	DNase-narrow	hESC Derived CD184+ Endoderm Cultured Cells	0.568
	DNase-gapped	hESC Derived CD184+ Endoderm Cultured Cells	0.656
Dataset	Method	Top Tissue	AUPR
Ultra conserved	FUN-LDA	hESC Derived CD184+ Endoderm Cultured Cells	0.485
Elements	GenoSkyline	hESC Derived CD184+ Endoderm Cultured Cells	0.490
	ChromHMM	hESC Derived CD56+ Ectoderm Cultured Cells	0.477
	Segway	hESC Derived CD56+ Mesoderm Cultured Cells	0.457
	IDEAS	hESC Derived CD184+ Endoderm Cultured Cells	0.490
	DNase	hESC Derived CD56+ Mesoderm Cultured Cells	0.485
	DNase-narrow	hESC Derived CD56+ Ectoderm Cultured Cells	0.434
	DNase-gapped	hESC Derived CD184+ Endoderm Cultured Cells	0.508

TABLE S12. Comparison of FUN-LDA with organism level functional prediction methods on the tissue/cell type specific datasets. AUROC/AUPR values for discriminating between variants likely to be functional and control variants. Results are shown for several datasets (three different cell lines) with experimental validation (MPRA) of potential regulatory variants, and one dsQTL dataset (dsQTLs & eQTLs contains a subset of dsQTLs that are also eQTLs). Methods include tissue/cell type specific FUN-LDA and organism level methods including phyloP (primate), Eigen, CADD, DANN, DeepSea and LINSIGHT.

Dataset	Method	AUROC	AUPR
emVars in[8], E116	FUN-LDA	0.707	0.468
	Eigen	0.604	0.359
	CADD	0.569	0.241
	DANN	0.532	0.226
	LINSIGHT	0.651	0.325
	phyloP	0.518	0.211
	DeepSea	0.691	0.437
Regulatory motifs in[9], E118/HepG2	FUN-LDA	0.691	0.445
	Eigen	0.636	0.363
	CADD	0.606	0.329
	DANN	0.623	0.381
	LINSIGHT	0.629	0.341
	phyloP	0.596	0.330
	DeepSea	0.646	0.347
Regulatory motifs in[9], E123/K562	FUN-LDA	0.645	0.287
	Eigen	0.573	0.224
	CADD	0.538	0.189
	DANN	0.547	0.196
	LINSIGHT	0.557	0.198
	phyloP	0.549	0.206
	DeepSea	0.606	0.253
dsQTLs in[14], E116	FUN-LDA	0.750	0.374
	Eigen	0.666	0.316
	CADD	0.646	0.296
	DANN	0.571	0.269
	LINSIGHT	0.750	0.386
	phyloP	0.537	0.245
	DeepSea	0.791	0.548
dsQTLs & eQTLs in[14], E116	FUN-LDA	0.793	0.476
	Eigen	0.717	0.425
	CADD	0.703	0.355
	DANN	0.586	0.283
	LINSIGHT	0.736	0.409
	phyloP	0.541	0.231
	DeepSea	0.805	0.564

TABLE S13. Grouping of Roadmap tissues into 10 tissue types.

Epigenome.ID	Type	Epigenome.ID	Type
E022	Stem cell	E117	Connective tissue
E007	Stem cell	E028	Connective tissue
E004	Stem cell	E057	Connective tissue
E002	Stem cell	E058	Connective tissue
E021	Stem cell	E119	Connective tissue
E009	Stem cell	E127	Connective tissue
E010	Stem cell	E071	Brain
E001	Stem cell	E074	Brain
E015	Stem cell	E073	Brain
E018	Stem cell	E068	Brain
E016	Stem cell	E067	Brain
E020	Stem cell	E069	Brain
E014	Stem cell	E072	Brain
E019	Stem cell	E027	Internal organs
E024	Stem cell	E059	Internal organs
E008	Stem cell	E061	Internal organs
E003	Stem cell	E065	Internal organs
E012	Stem cell	E097	Internal organs
E011	Stem cell	E086	Internal organs
E115	Blood	E087	Internal organs
E123	Blood	E100	Internal organs
E030	Blood	E105	Internal organs
E029	Blood	E104	Internal organs
E124	Blood	E095	Internal organs
E035	Blood	E096	Internal organs
E036	Blood	E113	Internal organs
E051	Blood	E079	Internal organs
E050	Blood	E094	Internal organs
E034	Blood	E098	Internal organs
E046	Blood	E081	Fetal brain
E041	Blood	E070	Fetal brain
E047	Blood	E082	Fetal brain
E048	Blood	E054	Fetal brain
E038	Blood	E053	Fetal brain
E045	Blood	E005	Fetal tissue 1
E044	Blood	E099	Fetal tissue 1
E043	Blood	E013	Fetal tissue 1
E039	Blood	E006	Fetal tissue 1
E042	Blood	E083	Fetal tissue 1
E040	Blood	E108	Muscle
E037	Blood	E107	Muscle
E112	Blood	E063	Muscle
E093	Blood	E078	Muscle
E062	Blood	E103	Muscle
E033	Blood	E076	Muscle
E116	Blood	E111	Muscle
E031	Blood	E091	Fetal tissue 2
E032	Blood	E092	Fetal tissue 2
E122	Connective tissue	E089	Fetal tissue 2
E120	Connective tissue	E090	Fetal tissue 2
E121	Connective tissue	E088	Fetal tissue 2
E025	Connective tissue	E080	Fetal tissue 2
E023	Connective tissue	E066	GI
E049	Connective tissue	E110	GI
E026	Connective tissue	E109	GI
E129	Connective tissue	E106	GI
E126	Connective tissue	E075	GI
E052	Connective tissue	E077	GI
E125	Connective tissue	E101	GI
E055	Connective tissue	E102	GI
E056	Connective tissue	E118	GI
E017	Connective tissue	E085	GI
E128	Connective tissue	E084	GI
E114	Connective tissue		