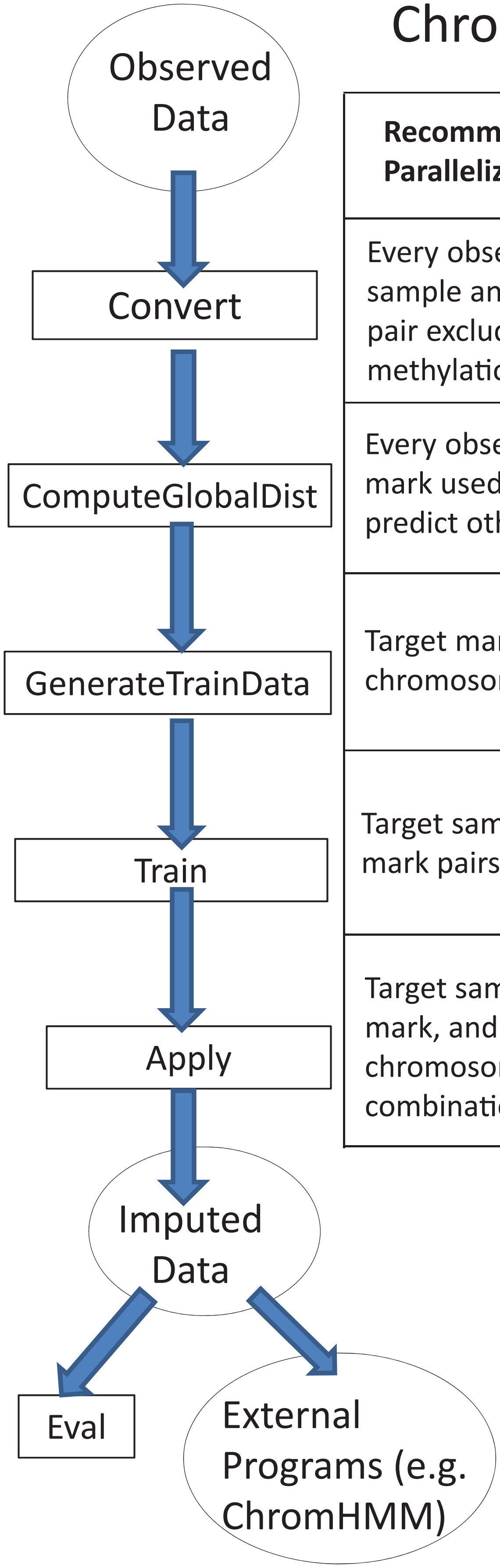
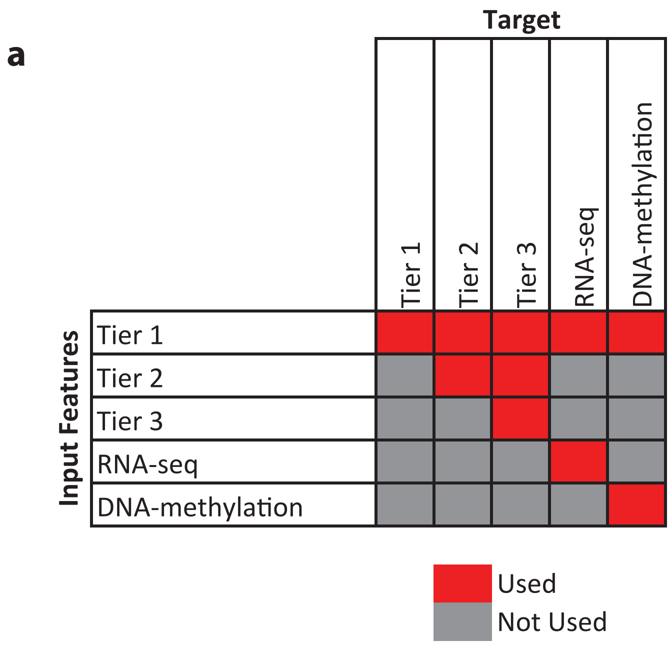
ChromImpute Workflow



Recommended Parallelization	**Number of Jobs**	**Approximate CPU time for hardest jobs**	**Approximate CPU time for average job**
Every observed sample and mark pair excluding DNA-methylation	1088	10min	10min
Every observed mark used to predict other marks	32	2hr	1hr
Target mark and chromosome	782	2hr	1hr
Target sample and mark pairs	4315	4hr	1hr
Target sample, mark, and chromosome combinations	99245	3hr	1hr

Supplementary Figure 1: Workflow of ChromImpute. On left is a flow chart illustrating the various steps to generate the imputed data using the ChromImpute software. First the observed data is converted into the desired target resolution. Second a global distance for each mark between all pairs of samples is computed based on correlation. Third the training data is generated. Fourth the regression predictors are trained. Finally the regression predictors can be applied to generate imputed data for which additional analysis can be conducted. On the right is a recommended parallelization strategy for each command, and then for the imputation application considered here the number of compute jobs that would lead to and the approximate maximum and average CPU time for each job.



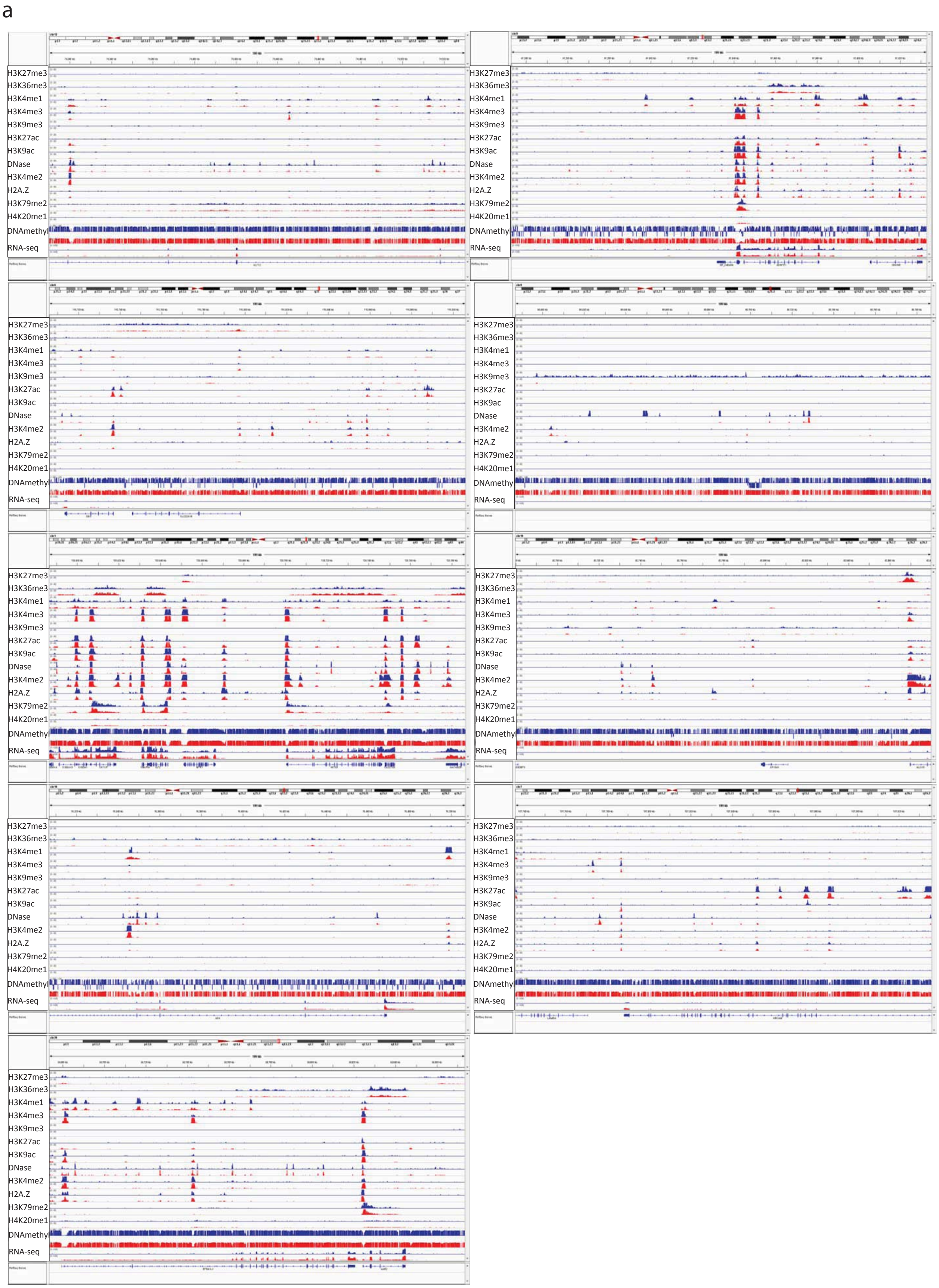
Input Features \ Target	Tier 1	Tier 2	Tier 3	RNA-seq	DNA-methylation
Tier 1	🟥	🟥	🟥	🟥	🟥
Tier 2	⬛	🟥	🟥	⬛	⬛
Tier 3	⬛	⬛	🟥	⬛	⬛
RNA-seq	⬛	⬛	⬛	🟥	⬛
DNA-methylation	⬛	⬛	⬛	⬛	🟥

🟥 Used
⬛ Not Used

b

Tier	Marks	# of Marks	# Samples Represented	Min # Samples Per Mark	Max # of Samples Per Mark	Average # of Samples Per Mark	Total # of Observed Datasets	Total # of Imputed Datasets	Tiers of Other Marks Used in Primary Imputation
1	DNase,H3K27ac,H3K27me3, H3K36me3,H3K4me1,H3K4me3, H3K9ac,H3K9me3	8	127	53	127	106	848	1016	1
2	H2A.Z,H3K4me2,H3K79me2, H4K20me1	4	26	19	23	21.75	87	508	1,2
3	H2AK5ac,H2AK9ac,H2BK120ac, H2BK12ac,H2BK15ac,H2BK20ac, H2BK5ac,H3K14ac,H3K18ac, H3K23ac,H3K23me2,H3K4ac, H3K56ac,H3K79me1,H3K9me1, H3T11ph,H4K5ac,H4K8ac, H4K12ac, H4K91ac	20	10	1	7	4.85	97	2537	1,2,3
	DNA Methylation	1	37	37	56	37	37	127	1
	RNA-seq	1	56	56	56	56	56	127	1

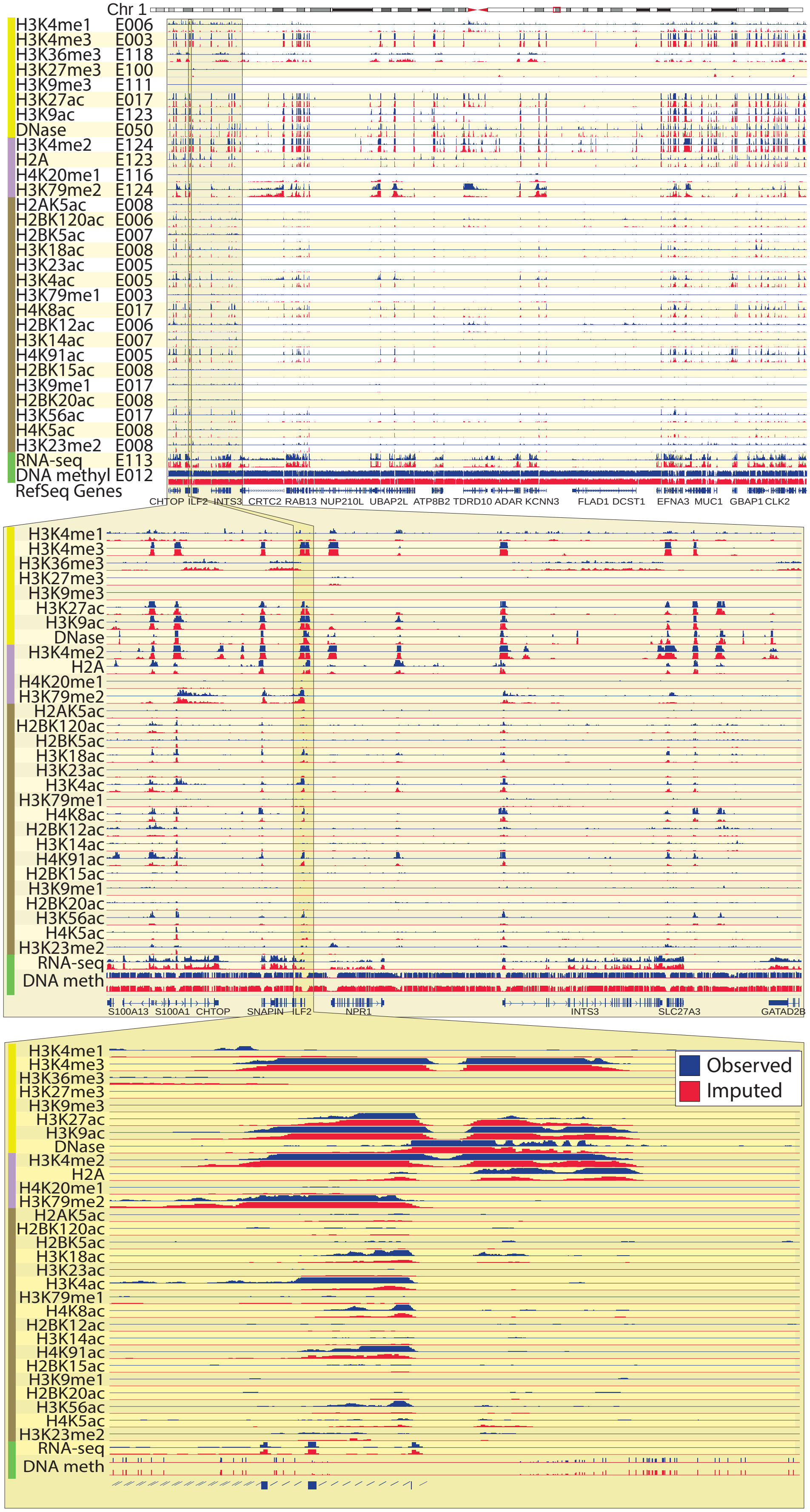
Supplementary Figure 2: Relationship between Mark Tiers for the Main Imputation. **(a)** The rows of the matrix correspond to subsets of marks for input defined in **b**, and the columns the set of target marks. A cell in the matrix is colored red if the corresponding set of marks of the row is used to predict the set of marks of the column for the main imputation. **(b)** The table reports statistics on the different tiers of marks as used for the primary imputation.



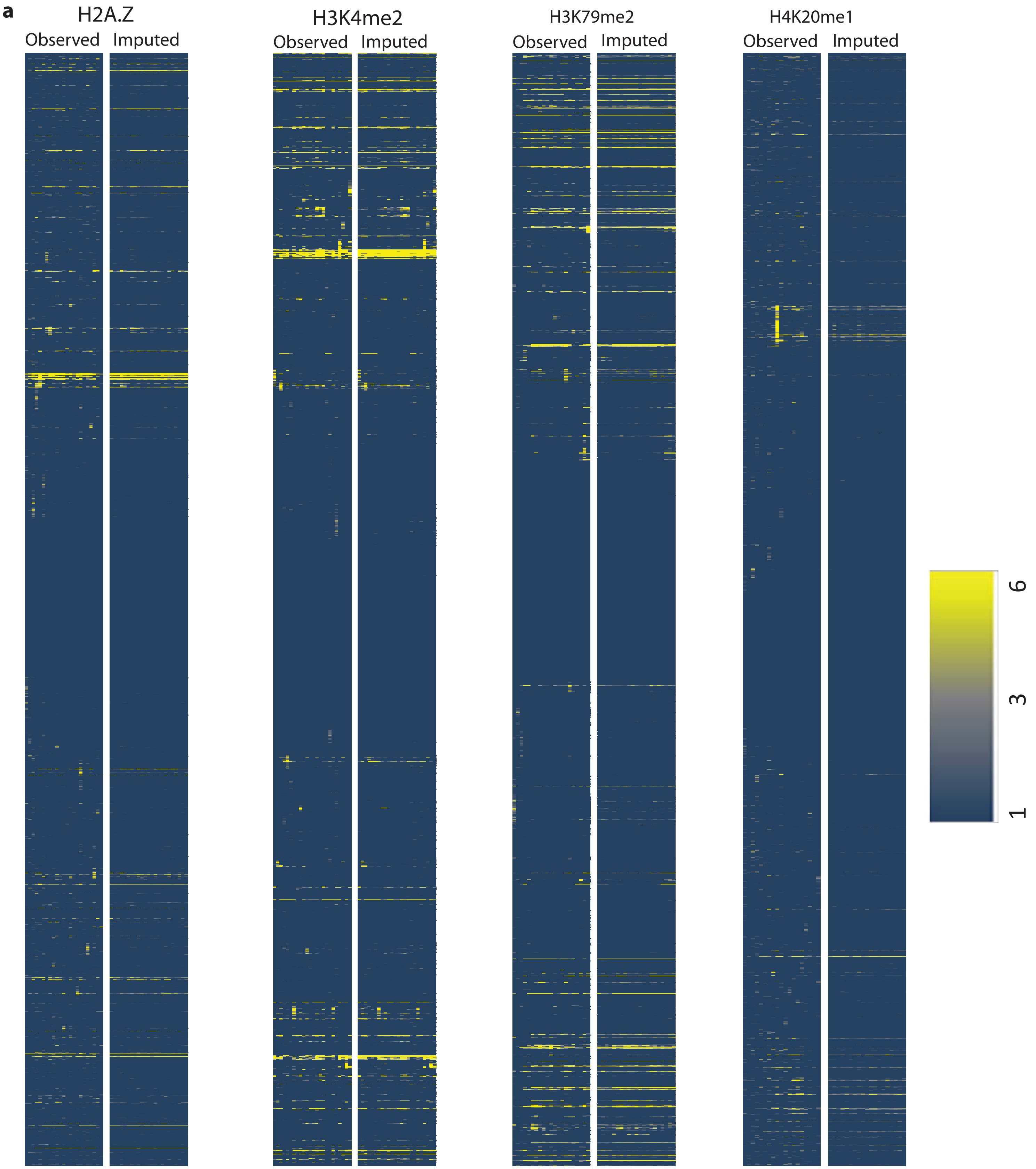
Supplementary Figure 3: Browser Images for Random Loci
(a) The figure shows browser screenshots for all Tier 1,2 marks, DNA-methylation, and RNA-seq at nine randomly selected 200kb loci of which the one with the most signal is also shown in Fig. 2. In blue is the signal for a randomly selected observed track for the mark and below it in red is the corresponding imputed track. DNA methylation values below the horizontal line correspond to missing values. (b) The same as shown in a, but for the Tier-3 marks. (c) Larger 1.5Mb context, and example 5kb close-up also shown for the randomly selected loci with the most signal associated with it. In the bottom set of tracks, note the dip in the nucleosome-free region of the ILF2 gene promoter, and the nucleotide-level concordance in CpG methylation information.


b
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2

c.

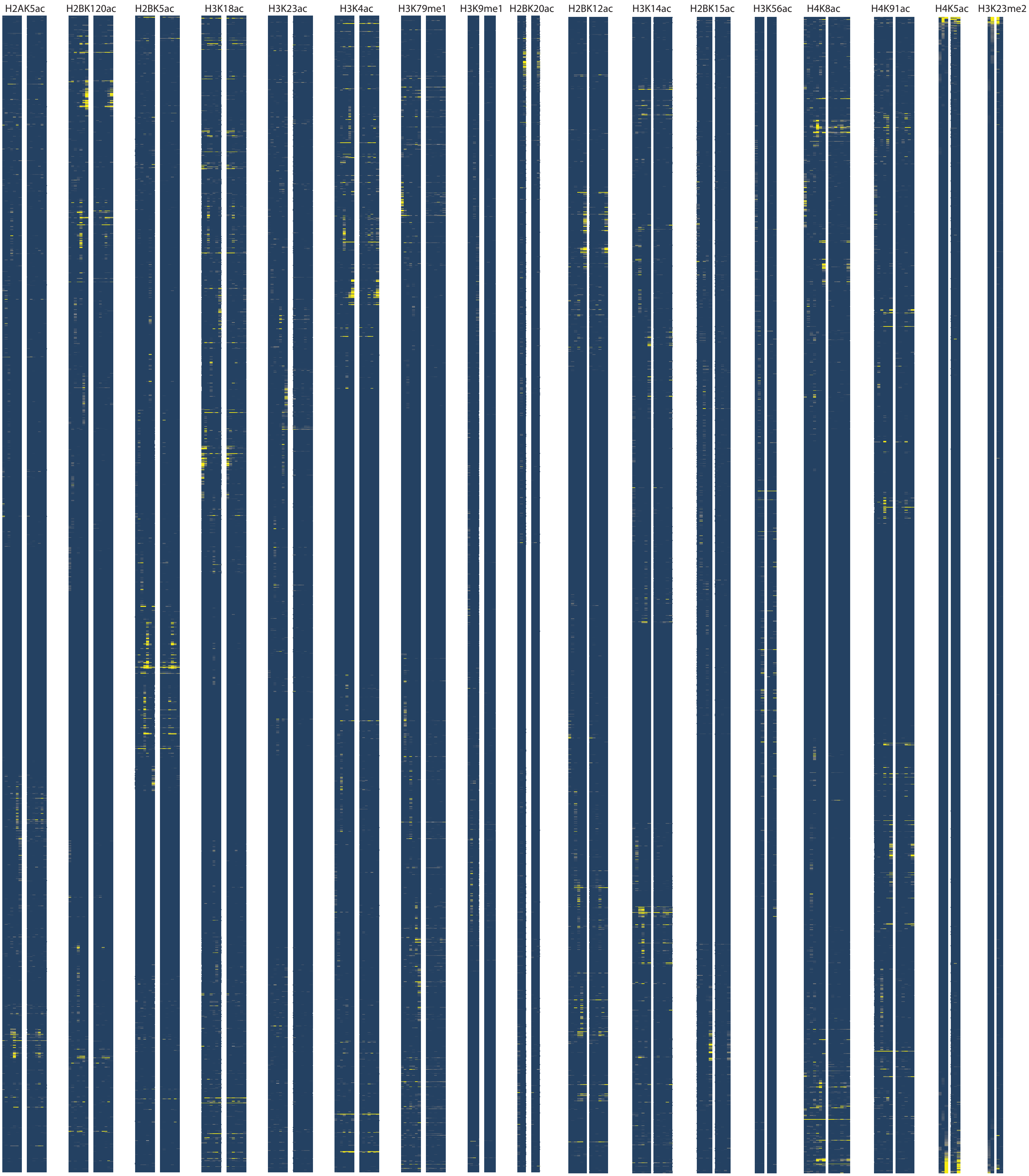


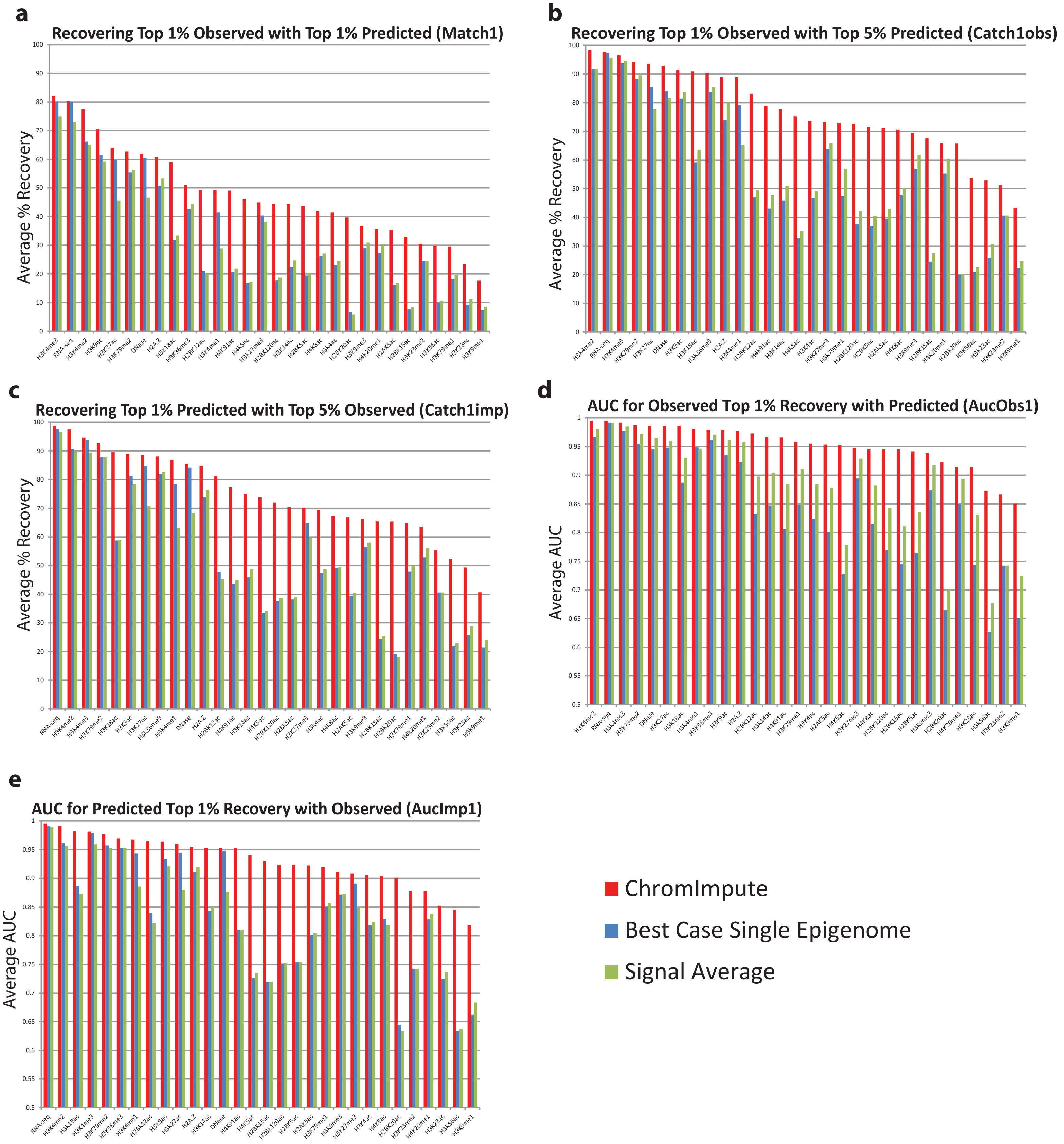
Chr 1
H3K4me1 E006
H3K4me3 E003
H3K36me3 E118
H3K27me3 E100
H3K9me3 E111
H3K27ac E017
H3K9ac E123
DNase E050
H3K4me2 E124
H2A E123
H4K20me1 E116
H3K79me2 E124
H2AK5ac E008
H2BK120ac E006
H2BK5ac E007
H3K18ac E008
H3K23ac E005
H3K4ac E005
H3K79me1 E003
H4K8ac E017
H2BK12ac E006
H3K14ac E007
H4K91ac E005
H2BK15ac E008
H3K9me1 E017
H2BK20ac E008
H3K56ac E017
H4K5ac E008
H3K23me2 E008
RNA-seq E113
DNA methyl E012
RefSeq Genes
CHTOP ILF2 INTS3 CRTC2 RAB13 NUP210L UBAP2L ATP8B2 TDRD10 ADAR KCNN3 FLAD1 DCST1 EFNA3 MUC1 GBAP1 CLK2
H3K4me1
H3K4me3
H3K36me3
H3K27me3
H3K9me3
H3K27ac
H3K9ac
DNase
H3K4me2
H2A
H4K20me1
H3K79me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
RNA-seq
DNA meth
S100A13 S100A1 CHTOP SNAPIN ILF2 NPR1 INTS3 SLC27A3 GATAD2B
H3K4me1
H3K4me3
H3K36me3
H3K27me3
H3K9me3
H3K27ac
H3K9ac
DNase
H3K4me2
H2A
H4K20me1
H3K79me2
H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H4K8ac
H2BK12ac
H3K14ac
H4K91ac
H2BK15ac
H3K9me1
H2BK20ac
H3K56ac
H4K5ac
H3K23me2
RNA-seq
DNA meth
Observed
Imputed



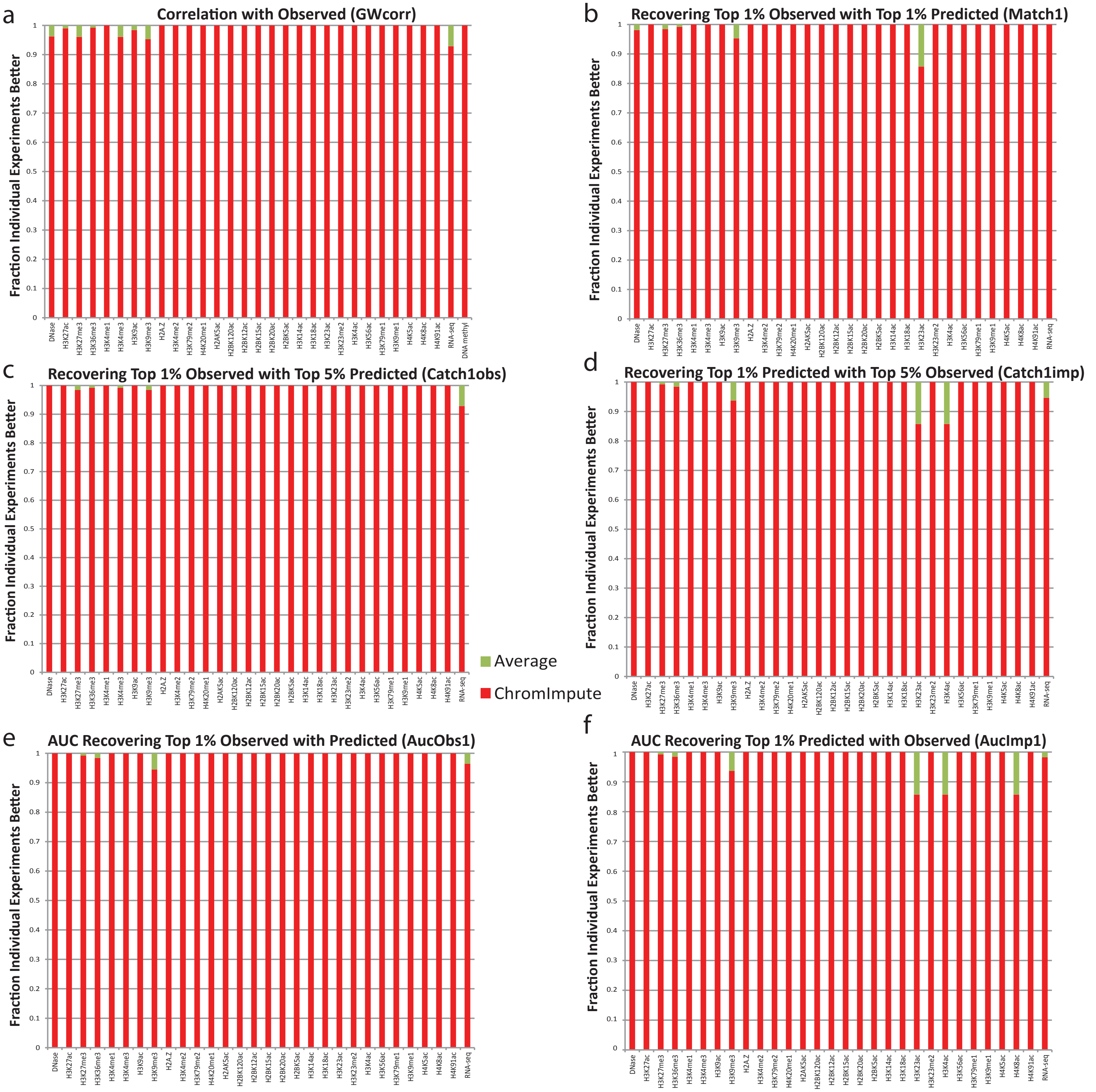
Supplementary Figure 4:
Heatmap of Clustering of Signal at Randomly Selected Positions for Tier-2 and Tier-3 Marks.
Similar heatmaps to those in **Fig. 2b**. The individual columns correspond to samples and the rows correspond to 2,000 randomly selected 25bp intervals that were clustered based on the observed data (left) with the corresponding imputed data also shown (right) for **(a)** Tier-2 marks and **(b)** Tier-3 marks. Only samples for which observed data is available are shown. The coloring corresponds to the signal level as indicated in the legend. Visually the heatmap show an overall agreement though some differences associated with outlier data sets or highly cell type specific behavior can also be seen.

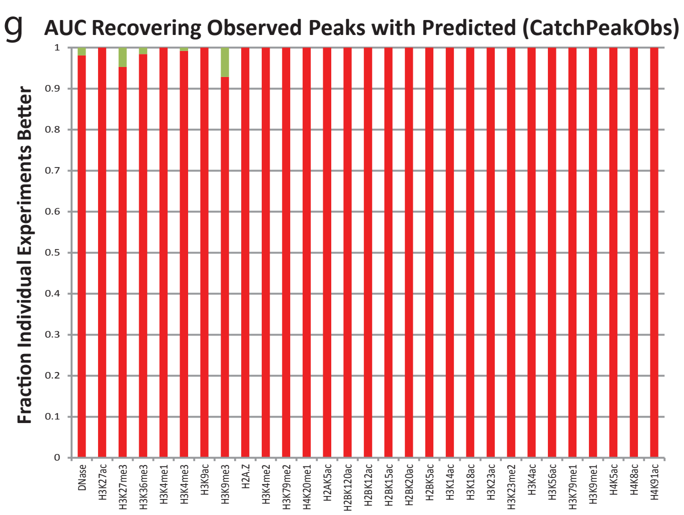
b


H2AK5ac
H2BK120ac
H2BK5ac
H3K18ac
H3K23ac
H3K4ac
H3K79me1
H3K9me1
H2BK20ac
H2BK12ac
H3K14ac
H2BK15ac
H3K56ac
H4K8ac
H4K91ac
H4K5ac
H3K23me2

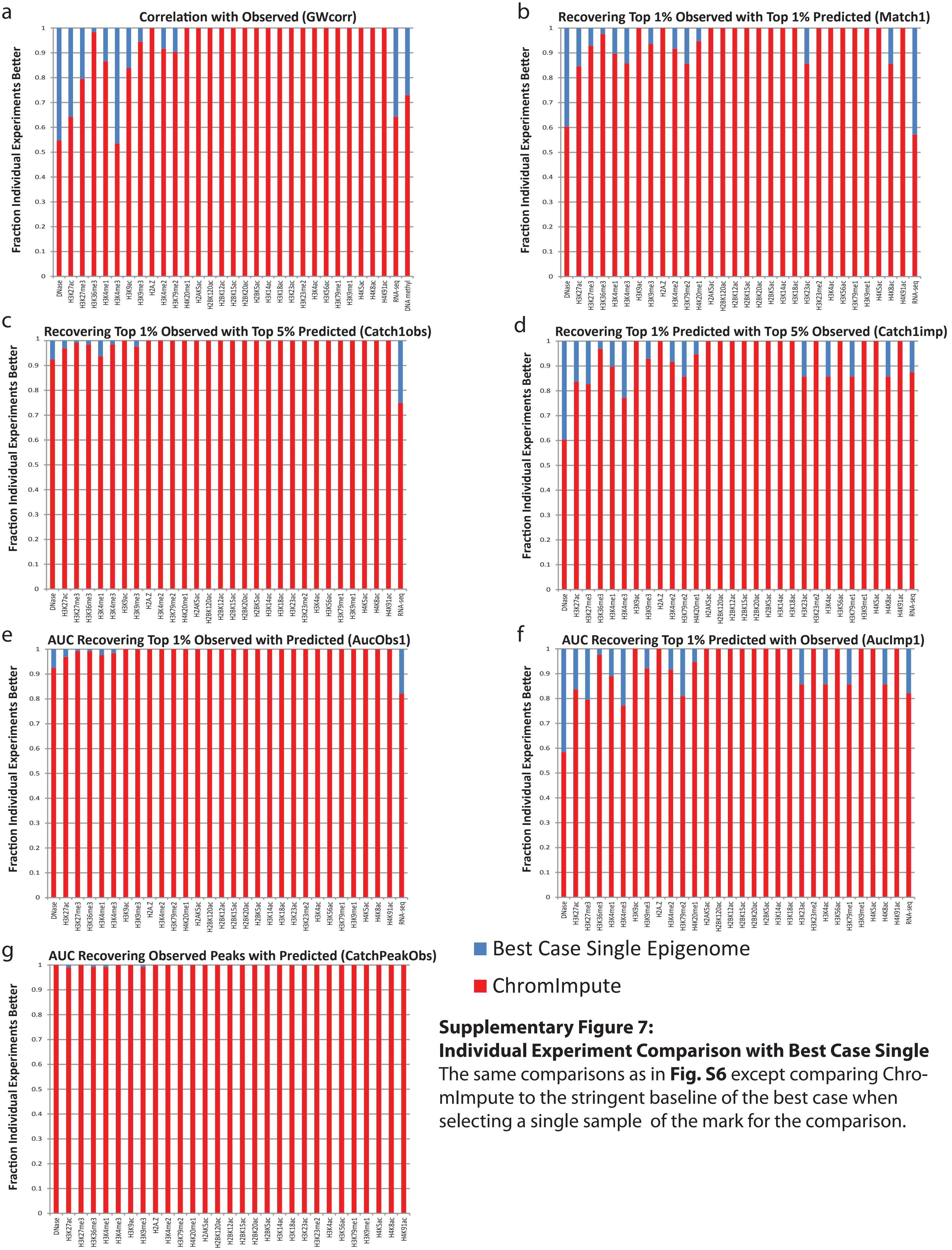


Supplementary Figure 5: Additional Aggregate Comparisons with Stringent Baselines. These comparisons are similar to those shown in **Fig 2c,d** except showing the results based on different evaluation metrics: **(a)** average percent recovery of 25-bp bins in the top 1% observed signal with those in the top 1% of predicted signal **(b)** average percent recovery of top 1% observed signal with the top 5% predicted signal **(c)** average percent recovery of top 1% predicted signal with the top 5% observed signal **(d)** Area under the ROC curve (AUC) for recovering top 1% observed signal when ranking based on predicted signal level **(e)** AUC for recovering top 1% predicted signal when ranking based on observed signal.

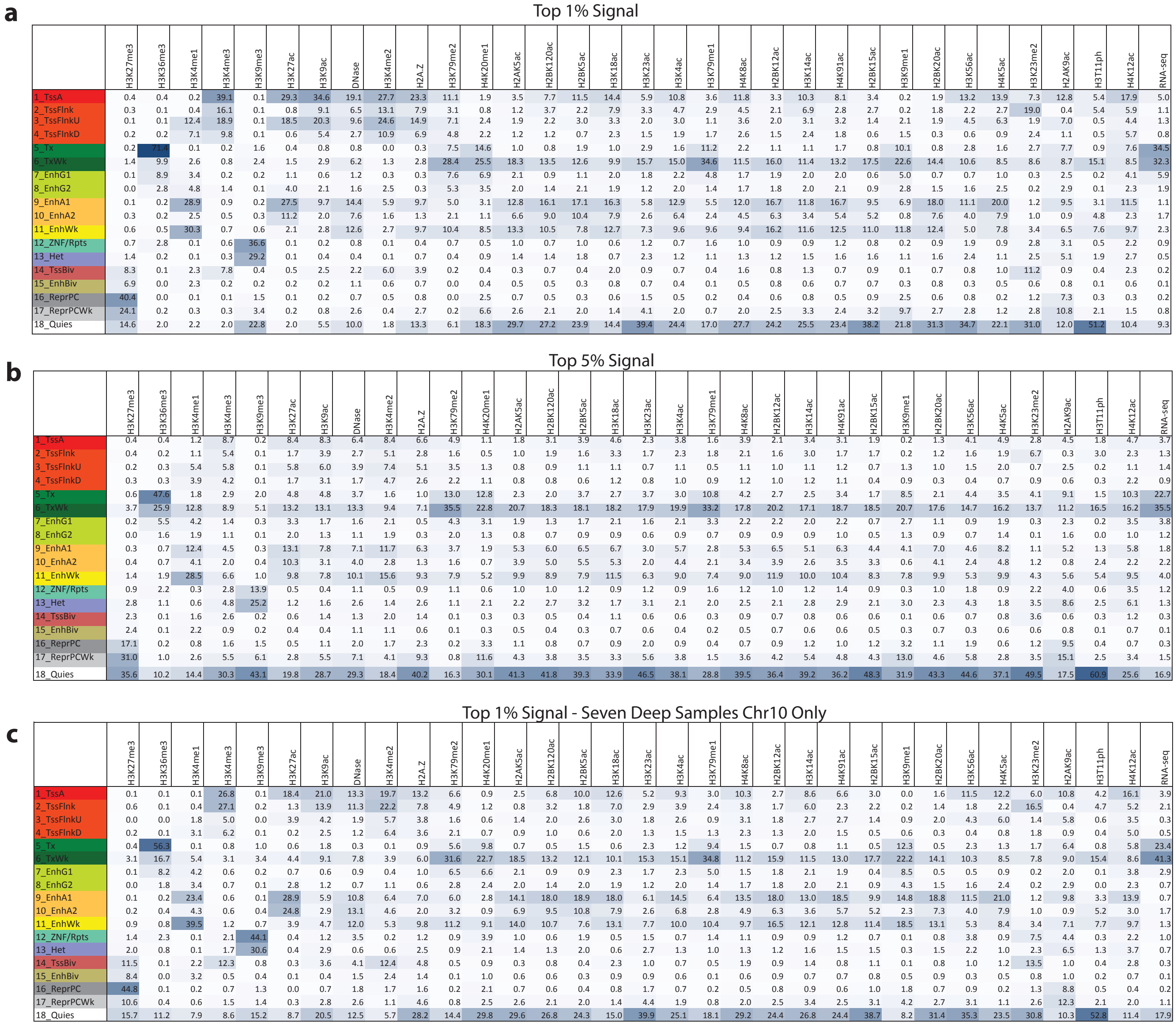




Supplementary Figure 6:
Individual Experiment Comparison with Average Signal.
The figure shows for each mark the fraction of individual experiments for which the ChromImpute predicted data better agrees with the observed data than the stringent baseline of the average signal of the mark in other samples. The comparison of ChromImpute against averaging the signal was done based on **(a)** Pearson correlation, **(b)** agreement for recovering locations of top 1% observed signal with top 1% predicted, **(c)** agreement for recovering top 1% observed signal with top 5% predicted, **(d)** agreement for recovering top 1% predicted with top 5% observed signal **(e)** AUC for recovering top 1% observed with predicted, **(f)** AUC for recovering top 1% predicted with observed, and **(g)** AUC for recovering bases covered by a narrow peak calls on observed data with predicted signal.



Supplementary Figure 7:
Individual Experiment Comparison with Best Case Single
The same comparisons as in **Fig. S6** except comparing ChromImpute to the stringent baseline of the best case when selecting a single sample of the mark for the comparison.



a

Top 1% Signal

	H3K27me3	H3K36me3	H3K4me1	H3K4me3	H3K9me3	H3K27ac	H3K9ac	DNase	H3K4me2	H2A.Z	H3K79me2	H4K20me1	H2AK5ac	H2BK120ac	H2BK5ac	H3K18ac	H3K23ac	H3K4ac	H3K79me1	H4K8ac	H2BK12ac	H3K14ac	H4K91ac	H2BK15ac	H3K9me1	H2BK20ac	H3K56ac	H4K5ac	H3K23me2	H2AK9ac	H3T11ph	H4K12ac	RNA-seq
1_TssA	0.4	0.4	0.2	39.1	0.1	29.3	34.6	19.1	27.7	23.3	11.1	1.9	3.5	7.7	11.5	14.4	5.9	10.8	3.6	11.8	3.3	10.3	8.1	3.4	0.2	1.9	13.2	13.9	7.3	12.8	5.4	17.9	5.0
2_TssFlnk	0.3	0.1	0.4	16.1	0.1	0.8	9.1	6.5	13.1	7.9	3.1	0.8	1.2	3.7	2.2	7.9	3.3	4.7	2.9	4.5	2.1	6.9	2.8	2.7	0.2	1.8	2.2	2.7	19.0	0.4	5.4	5.9	1.1
3_TssFlnkU	0.1	0.1	12.4	18.9	0.1	18.5	20.3	9.6	24.6	14.9	7.1	2.4	1.9	2.2	3.0	3.3	2.0	3.0	1.1	3.6	2.0	3.1	3.2	1.4	2.1	1.9	4.5	6.3	1.9	7.0	0.5	4.4	1.3
4_TssFlnkD	0.2	0.2	7.1	9.8	0.1	0.6	5.4	2.7	10.9	6.9	4.8	2.2	1.2	1.2	0.7	2.3	1.5	1.9	1.7	2.6	1.5	2.4	1.8	0.6	1.5	0.3	0.6	0.9	2.4	1.1	0.5	5.7	0.8
5_Tx	0.2	71.4	0.1	0.2	1.6	0.4	0.8	0.8	0.0	0.3	7.5	14.6	1.0	0.8	1.9	1.0	2.9	1.6	11.2	2.2	1.1	1.1	1.7	0.8	10.1	0.8	2.8	1.6	2.7	7.7	0.9	7.6	34.5
6_TxWk	1.4	9.9	2.6	0.8	2.4	1.5	2.9	6.2	1.3	2.8	28.4	25.5	18.3	13.5	12.6	9.9	15.7	15.0	34.6	11.5	16.0	11.4	13.2	17.5	22.6	14.4	10.6	8.5	8.6	8.7	15.1	8.5	32.3
7_EnhG1	0.1	8.9	3.4	0.2	0.2	1.1	0.6	1.2	0.3	0.3	7.6	6.9	2.1	0.9	1.1	2.0	1.8	2.2	4.8	1.7	1.9	2.0	2.0	0.6	5.0	0.7	0.7	1.0	0.3	2.5	0.2	4.1	5.9
8_EnhG2	0.0	2.8	4.8	1.4	0.1	4.0	2.1	1.6	2.5	0.3	5.3	3.5	2.0	1.4	2.1	1.9	1.2	2.0	1.4	1.7	1.8	2.0	2.1	0.9	2.8	1.5	1.6	2.5	0.2	2.9	0.1	2.3	1.9
9_EnhA1	0.1	0.2	28.9	0.9	0.2	27.5	9.7	14.4	5.9	9.7	5.0	3.1	12.8	16.1	17.1	16.3	5.8	12.9	5.5	12.0	16.7	11.8	16.7	9.5	6.9	18.0	11.1	20.0	1.2	9.5	3.1	11.5	1.1
10_EnhA2	0.3	0.2	2.5	0.5	0.3	11.2	2.0	7.6	1.6	1.3	2.1	1.1	6.6	9.0	10.4	7.9	2.6	6.4	2.4	4.5	6.3	3.4	5.4	5.2	0.8	7.6	4.0	7.9	1.0	0.9	4.8	2.3	1.7
11_EnhWk	0.6	0.5	30.3	0.7	0.6	2.1	2.8	12.6	2.7	9.7	10.4	8.5	13.3	10.5	7.8	12.7	7.3	9.6	9.6	9.4	16.2	11.6	12.5	11.0	11.8	12.4	5.0	7.8	3.4	6.5	7.6	9.7	2.3
12_ZNF/Rpts	0.7	2.8	0.1	0.6	36.6	0.1	0.2	0.8	0.1	0.4	0.7	0.5	1.0	0.7	1.0	0.6	1.2	0.7	1.6	1.0	0.9	0.9	1.2	0.8	0.2	0.9	1.9	0.9	2.8	3.1	0.5	2.2	0.9
13_Het	1.4	0.2	0.1	0.3	29.2	0.1	0.4	0.4	0.1	0.7	0.4	0.9	1.4	1.4	1.8	0.7	2.3	1.2	1.1	1.3	1.2	1.5	1.6	1.6	1.1	1.6	2.4	1.1	2.5	5.1	1.9	2.7	0.5
14_TssBiv	8.3	0.1	2.3	7.8	0.4	0.5	2.5	2.2	6.0	3.9	0.2	0.4	0.3	0.7	0.4	2.0	0.9	0.7	0.4	1.6	0.8	1.3	0.7	0.9	0.1	0.7	0.8	1.0	11.2	0.9	0.4	2.3	0.2
15_EnhBiv	6.9	0.0	2.3	0.2	0.2	0.2	0.2	1.1	0.6	0.7	0.0	0.4	0.5	0.5	0.3	0.8	0.7	0.4	0.1	0.5	0.8	0.6	0.7	0.7	0.3	0.8	0.3	0.5	0.6	0.8	0.1	0.6	0.1
16_ReprPC	40.4	0.0	0.1	0.1	1.5	0.1	0.2	0.7	0.5	0.8	0.0	2.5	0.7	0.5	0.3	0.6	1.5	0.5	0.2	0.4	0.6	0.8	0.5	0.9	2.5	0.6	0.8	0.2	1.2	7.3	0.3	0.3	0.2
17_ReprPCWk	24.1	0.2	0.3	0.3	3.4	0.2	0.8	2.6	0.4	2.7	0.2	6.6	2.6	2.1	2.0	1.4	4.1	2.0	0.7	2.0	2.5	3.3	2.4	3.2	9.7	2.7	2.8	1.2	2.8	10.8	2.1	1.5	0.8
18_Quies	14.6	2.0	2.2	2.0	22.8	2.0	5.5	10.0	1.8	13.3	6.1	18.3	29.7	27.2	23.9	14.4	39.4	24.4	17.0	27.7	24.2	25.5	23.4	38.2	21.8	31.3	34.7	22.1	31.0	12.0	51.2	10.4	9.3

b

Top 5% Signal

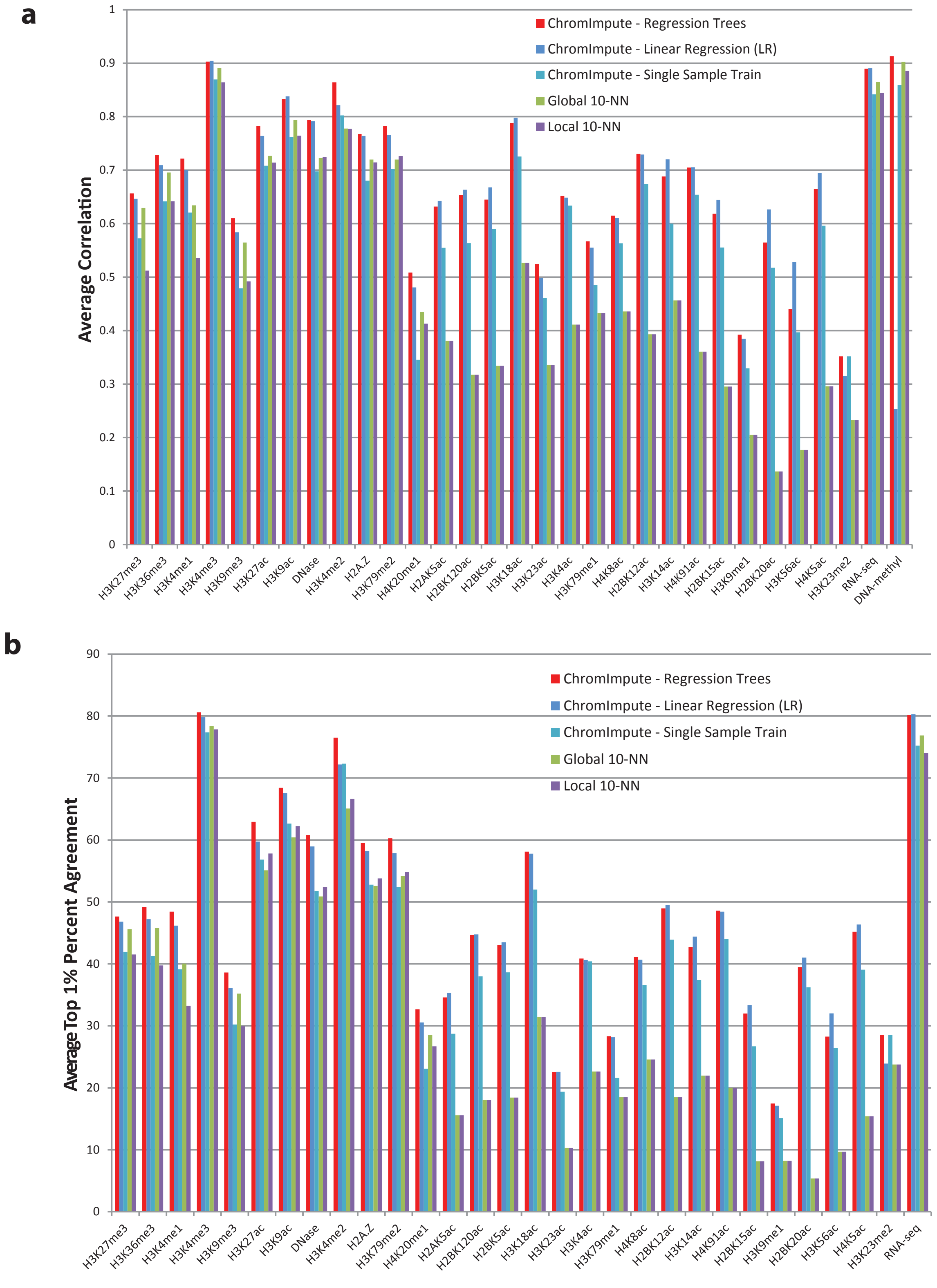
	H3K27me3	H3K36me3	H3K4me1	H3K4me3	H3K9me3	H3K27ac	H3K9ac	DNase	H3K4me2	H2A.Z	H3K79me2	H4K20me1	H2AK5ac	H2BK120ac	H2BK5ac	H3K18ac	H3K23ac	H3K4ac	H3K79me1	H4K8ac	H2BK12ac	H3K14ac	H4K91ac	H2BK15ac	H3K9me1	H2BK20ac	H3K56ac	H4K5ac	H3K23me2	H2AK9ac	H3T11ph	H4K12ac	RNA-seq
1_TssA	0.4	0.4	1.2	8.7	0.2	8.4	8.3	6.4	8.4	6.6	4.9	1.1	1.8	3.1	3.9	4.6	2.3	3.8	1.6	3.9	2.1	3.4	3.1	1.9	0.2	1.3	4.1	4.9	2.8	4.5	1.8	4.7	3.7
2_TssFlnk	0.4	0.2	1.1	5.4	0.1	1.7	3.9	2.7	5.1	2.8	1.6	0.5	1.0	1.9	1.6	3.3	1.7	2.3	1.8	2.1	1.6	3.0	1.7	1.7	0.2	1.2	1.6	1.9	6.7	0.3	3.0	2.3	1.3
3_TssFlnkU	0.2	0.3	5.4	5.8	0.1	5.8	6.0	3.9	7.4	5.1	3.5	1.3	0.8	0.9	1.1	1.1	0.7	1.1	0.5	1.1	1.0	1.1	1.2	0.7	1.3	1.0	1.5	2.0	0.7	2.5	0.2	1.1	1.4
4_TssFlnkD	0.3	0.3	3.9	4.2	0.1	1.7	3.1	1.7	4.7	2.6	2.2	1.1	0.8	0.8	0.6	1.2	0.8	1.0	0.9	1.2	1.0	1.2	1.1	0.4	0.9	0.3	0.4	0.7	0.9	0.6	0.3	2.2	0.9
5_Tx	0.6	47.6	1.8	2.9	2.0	4.8	4.8	3.7	1.6	1.0	13.0	12.8	2.3	2.0	3.7	2.7	3.7	3.0	10.8	4.2	2.7	2.5	3.4	1.7	8.5	2.1	4.4	3.5	4.1	9.1	1.5	10.3	22.7
6_TxWk	3.7	25.9	12.8	8.9	5.1	13.2	13.1	13.3	9.4	7.1	35.5	22.8	20.7	18.3	18.1	18.2	17.9	19.9	33.2	17.8	20.2	17.1	18.7	18.5	20.7	17.6	14.7	16.2	13.7	11.2	16.5	16.2	35.5
7_EnhG1	0.2	5.5	4.2	1.4	0.3	3.3	1.7	1.6	2.1	0.5	4.3	3.1	1.9	1.3	1.7	2.1	1.6	2.1	3.3	2.2	2.2	2.0	2.2	0.7	2.7	1.1	0.9	1.9	0.3	2.3	0.2	3.5	3.8
8_EnhG2	0.0	1.6	1.9	1.1	0.1	2.0	1.3	1.1	1.9	0.3	2.0	1.3	0.8	0.7	0.9	0.9	0.6	0.9	0.7	0.9	0.9	0.9	1.0	0.5	1.3	0.9	0.7	1.4	0.1	1.6	0.0	1.0	1.2
9_EnhA1	0.3	0.7	12.4	4.5	0.3	13.1	7.8	7.1	11.7	6.3	3.7	1.9	5.3	6.0	6.5	6.7	3.0	5.7	2.8	5.3	6.5	5.1	6.3	4.4	4.1	7.0	4.6	8.2	1.1	5.2	1.3	5.8	1.8
10_EnhA2	0.4	0.7	4.1	2.0	0.4	10.3	3.1	4.0	2.8	1.3	1.6	0.7	3.9	5.0	5.5	5.3	2.0	4.4	2.1	3.4	3.9	2.6	3.5	3.3	0.6	4.1	2.4	4.8	1.2	0.8	2.4	2.2	2.2
11_EnhWk	1.4	1.9	28.5	6.6	1.0	9.8	7.8	10.1	15.6	9.3	7.9	5.2	9.9	8.9	7.9	11.5	6.3	9.0	7.4	9.0	11.9	10.0	10.4	8.3	7.8	9.9	5.3	9.9	4.3	5.6	5.4	9.5	4.0
12_ZNF/Rpts	0.9	2.2	0.3	2.8	13.9	0.5	0.4	1.1	0.5	0.9	1.1	0.6	1.0	1.0	1.2	0.9	1.2	0.9	1.6	1.2	1.0	1.2	1.4	0.9	0.3	1.0	1.8	0.9	2.2	4.0	0.6	3.5	1.2
13_Het	2.8	1.1	0.6	4.8	25.2	1.2	1.6	2.6	1.4	2.6	1.1	2.1	2.2	2.7	3.2	1.7	3.1	2.1	2.0	2.5	2.1	2.8	2.9	2.1	3.0	2.3	4.3	1.8	3.5	8.6	2.5	6.1	1.3
14_TssBiv	2.3	0.1	1.6	2.6	0.2	0.6	1.4	1.3	2.0	1.4	0.1	0.3	0.3	0.5	0.4	1.1	0.6	0.6	0.3	0.9	0.7	0.8	0.6	0.6	0.1	0.6	0.7	0.8	3.6	0.6	0.3	1.2	0.3
15_EnhBiv	2.4	0.1	2.2	0.9	0.2	0.4	0.4	1.1	1.1	0.6	0.1	0.3	0.5	0.4	0.3	0.7	0.6	0.4	0.2	0.5	0.7	0.6	0.6	0.5	0.3	0.7	0.3	0.6	0.6	0.8	0.1	0.7	0.1
16_ReprPC	17.1	0.2	0.8	1.6	1.5	0.5	1.1	2.0	1.7	2.3	0.2	3.3	1.1	0.8	0.7	0.9	2.0	0.9	0.4	0.7	0.9	1.2	1.0	1.2	3.2	1.1	1.9	0.6	1.2	9.5	0.4	0.7	0.3
17_ReprPCWk	31.0	1.0	2.6	5.5	6.1	2.8	5.5	7.1	4.1	9.3	0.8	11.6	4.3	3.8	3.5	3.3	5.6	3.8	1.5	3.6	4.2	5.4	4.8	4.3	13.0	4.6	5.8	2.8	3.5	15.1	2.5	3.4	1.5
18_Quies	35.6	10.2	14.4	30.3	43.1	19.8	28.7	29.3	18.4	40.2	16.3	30.1	41.3	41.8	39.3	33.9	46.5	38.1	28.8	39.5	36.4	39.2	36.2	48.3	31.9	43.3	44.6	37.1	49.5	17.5	60.9	25.6	16.9

c

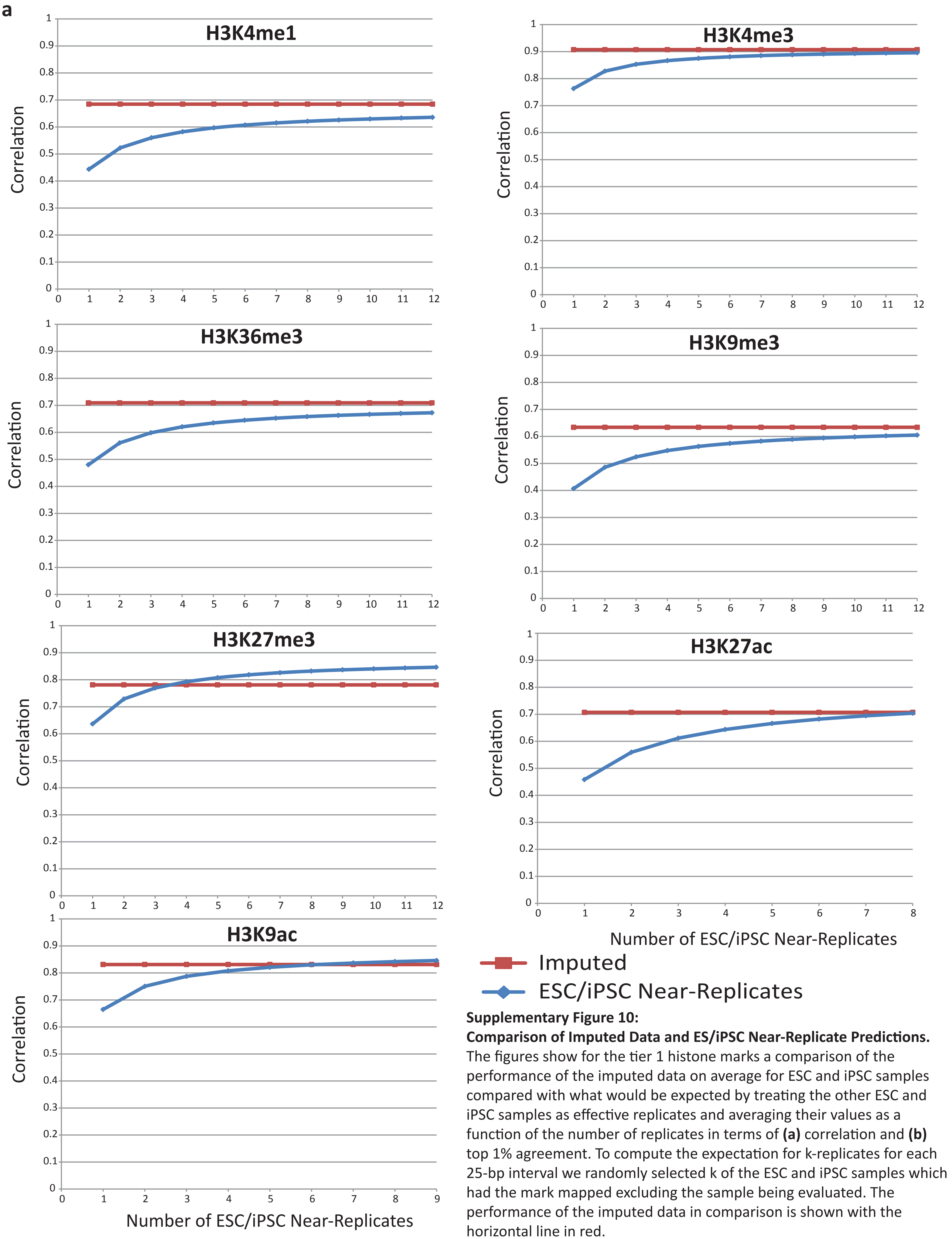
Top 1% Signal - Seven Deep Samples Chr10 Only

	H3K27me3	H3K36me3	H3K4me1	H3K4me3	H3K9me3	H3K27ac	H3K9ac	DNase	H3K4me2	H2A.Z	H3K79me2	H4K20me1	H2AK5ac	H2BK120ac	H2BK5ac	H3K18ac	H3K23ac	H3K4ac	H3K79me1	H4K8ac	H2BK12ac	H3K14ac	H4K91ac	H2BK15ac	H3K9me1	H2BK20ac	H3K56ac	H4K5ac	H3K23me2	H2AK9ac	H3T11ph	H4K12ac	RNA-seq
1_TssA	0.1	0.1	0.1	26.8	0.1	18.4	21.0	13.3	19.7	13.2	6.6	0.9	2.5	6.8	10.0	12.6	5.2	9.3	3.0	10.3	2.7	8.6	6.6	3.0	0.0	1.6	11.5	12.2	6.0	10.8	4.2	16.1	3.9
2_TssFlnk	0.6	0.1	0.4	27.1	0.2	1.3	13.9	11.3	22.2	7.8	4.9	1.2	0.8	3.2	1.8	7.0	2.9	3.9	2.4	3.8	1.7	6.0	2.3	2.2	0.2	1.4	1.8	2.2	16.5	0.4	4.7	5.2	2.1
3_TssFlnkU	0.0	0.0	1.8	5.0	0.0	3.9	4.2	1.9	5.7	3.8	1.6	0.6	1.4	2.0	2.6	3.0	1.8	2.6	0.9	3.1	1.8	2.7	2.7	1.4	0.9	2.0	4.3	6.0	1.4	5.8	0.6	3.5	0.3
4_TssFlnkD	0.2	0.1	3.1	6.2	0.1	0.2	2.5	1.2	6.4	3.6	2.1	0.7	0.9	1.0	0.6	2.0	1.3	1.5	1.3	2.3	1.3	2.0	1.5	0.5	0.6	0.3	0.4	0.8	1.8	0.9	0.4	5.0	0.5
5_Tx	0.4	56.3	0.1	0.8	1.0	0.6	1.8	0.3	0.1	0.9	5.6	9.8	0.7	0.5	1.5	0.6	2.3	1.2	9.4	1.5	0.7	0.8	1.1	0.5	12.3	0.5	2.3	1.3	1.7	6.4	0.8	5.8	23.4
6_TxWk	3.1	16.7	5.4	3.1	3.4	4.4	9.1	7.8	3.9	6.0	31.6	22.7	18.5	13.2	12.1	10.1	15.3	15.1	34.8	11.2	15.9	11.5	13.0	17.7	22.2	14.1	10.3	8.5	7.8	9.0	15.4	8.6	41.3
7_EnhG1	0.1	8.2	4.2	0.6	0.2	0.7	0.6	0.9	0.4	1.0	6.5	6.6	2.1	0.9	0.9	2.3	1.7	2.3	5.0	1.5	1.8	2.1	1.9	0.4	8.5	0.5	0.5	0.9	0.2	2.0	0.1	3.8	2.9
8_EnhG2	0.0	1.8	3.4	0.7	0.1	2.8	1.2	0.7	1.1	0.6	2.8	2.4	2.0	1.4	2.0	1.9	1.2	2.0	1.4	1.7	1.8	2.0	2.1	0.9	4.3	1.5	1.6	2.4	0.2	2.9	0.0	2.3	0.7
9_EnhA1	0.1	0.2	23.4	0.6	0.1	28.9	5.9	10.8	6.4	7.0	6.0	2.8	14.1	18.0	18.9	18.0	6.1	14.5	6.4	13.5	18.0	13.0	18.5	9.9	14.8	18.8	11.5	21.0	1.2	9.8	3.3	13.9	0.7
10_EnhA2	0.2	0.4	4.3	0.6	0.4	24.8	2.9	13.1	4.6	2.0	3.2	0.9	6.9	9.5	10.8	7.9	2.6	6.8	2.8	4.9	6.3	3.6	5.7	5.2	2.3	7.3	4.0	7.9	1.0	0.9	5.2	3.0	1.7
11_EnhWk	0.9	0.8	39.5	1.2	0.7	3.9	4.7	12.0	5.3	9.8	11.2	9.1	14.0	10.7	7.6	13.1	7.7	10.0	10.4	9.7	16.5	12.1	12.8	11.4	18.5	13.1	5.3	8.4	3.4	7.1	7.7	9.7	1.3
12_ZNF/Rpts	1.4	2.3	0.1	2.1	44.1	0.4	1.2	3.5	0.2	1.2	0.9	3.9	1.0	0.6	1.9	0.5	1.5	0.7	1.4	1.1	0.9	0.9	1.2	0.7	0.1	0.8	3.8	0.9	7.5	4.4	0.3	2.2	1.0
13_Het	2.0	0.8	0.1	1.7	30.6	0.4	2.9	0.6	0.6	2.5	0.9	2.1	1.4	1.3	2.0	0.6	2.7	1.3	1.0	1.3	1.2	1.6	1.5	1.5	0.3	1.5	2.2	1.0	2.3	6.5	1.3	3.7	0.7
14_TssBiv	11.5	0.1	2.2	12.3	0.8	0.3	3.6	4.1	12.4	4.8	0.5	0.9	0.3	0.8	0.4	2.3	1.0	0.7	0.5	1.9	0.9	1.4	0.8	1.1	0.1	0.8	1.0	1.2	13.5	1.0	0.4	2.8	0.3
15_EnhBiv	8.4	0.0	3.2	0.5	0.4	0.1	0.4	1.5	2.4	1.4	0.1	1.0	0.6	0.6	0.3	0.9	0.9	0.6	0.1	0.6	0.9	0.7	0.7	0.8	0.5	0.9	0.3	0.5	0.8	1.0	0.2	0.7	0.1
16_ReprPC	44.8	0.1	0.2	0.7	1.3	0.0	0.7	1.8	1.7	1.6	0.2	2.1	0.7	0.6	0.4	0.8	1.7	0.7	0.2	0.5	0.8	0.9	0.6	1.0	1.8	0.7	0.9	0.2	1.3	8.8	0.5	0.4	0.2
17_ReprPCWk	10.6	0.4	0.6	1.5	1.4	0.3	2.8	2.6	1.1	4.6	0.8	2.5	2.6	2.1	2.0	1.4	4.4	1.9	0.8	2.0	2.5	3.4	2.5	3.1	4.2	2.7	3.1	1.1	2.6	12.3	2.1	2.0	1.1
18_Quies	15.7	11.2	7.9	8.6	15.2	8.7	20.5	12.5	5.7	28.2	14.4	29.8	29.6	26.8	24.3	15.0	39.9	25.1	18.1	29.2	24.4	26.8	24.4	38.7	8.2	31.4	35.3	23.5	30.8	10.3	52.8	11.4	17.9

Supplementary Figure 8: Observed Chromatin State Distribution of Top 1 and 5% Observed Signal. The figure shows for each indicated mark (columns) the average chromatin-state distribution of the 25-bp intervals falling within the **(a)** top 1% signal **(b)** top 5% of signal. The chromatin states were taken from (Roadmap Epigenomics Consortium et al, 2015)[10] and defined based on observed data for six marks H3K4me1, H3K4me3, H3K36me3, H3K27me3, H3K9me3 and H3K27ac (also see **Fig. S33** for emission parameters). The average is computed based on only those samples in which both the mark and chromatin-states for this 18-state model were defined (all samples that had H3K27ac). In **(c)** the distribution of top 1% signal is shown, but limited only to chr10 and the seven samples with deep mark coverage.

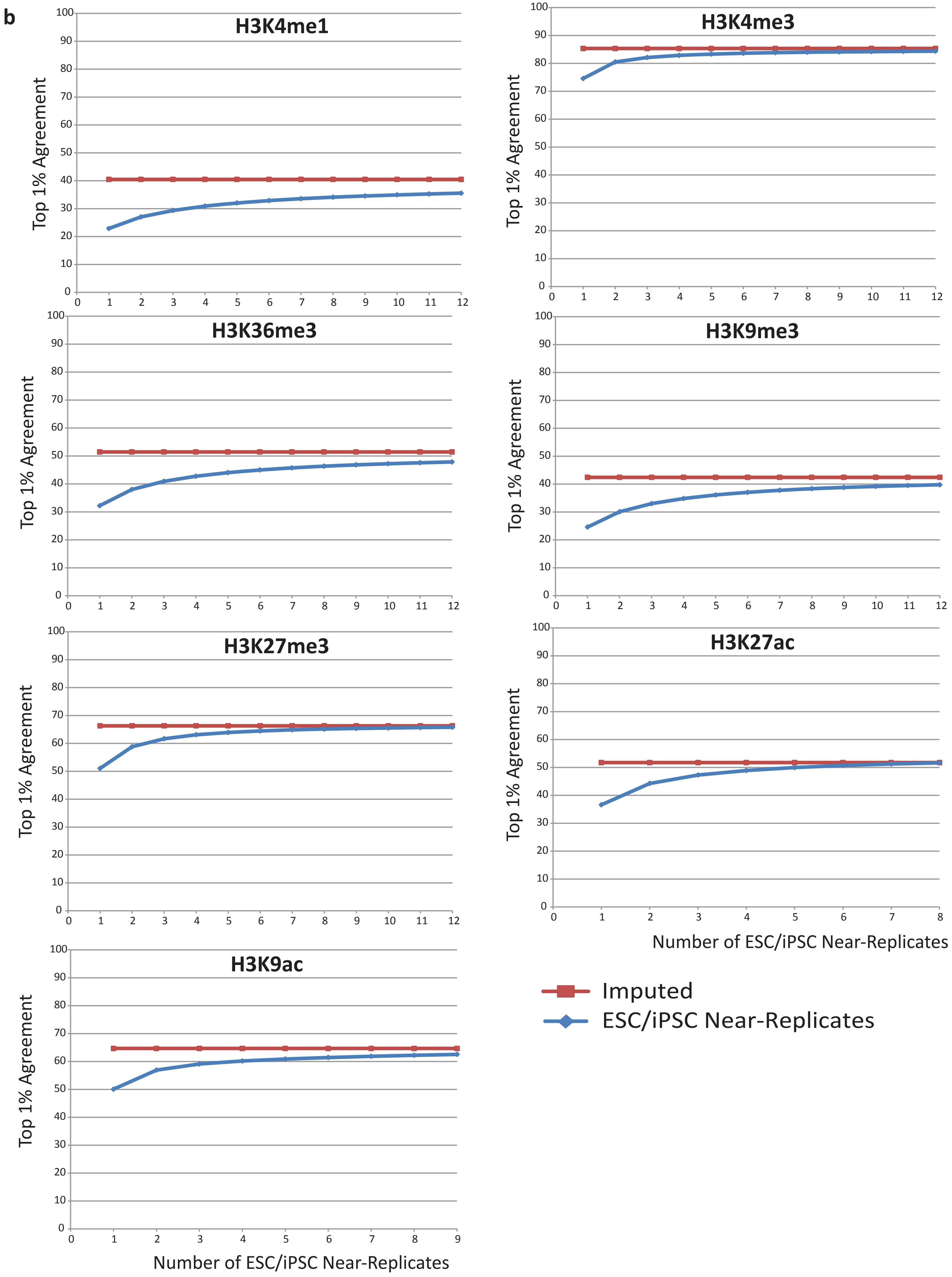


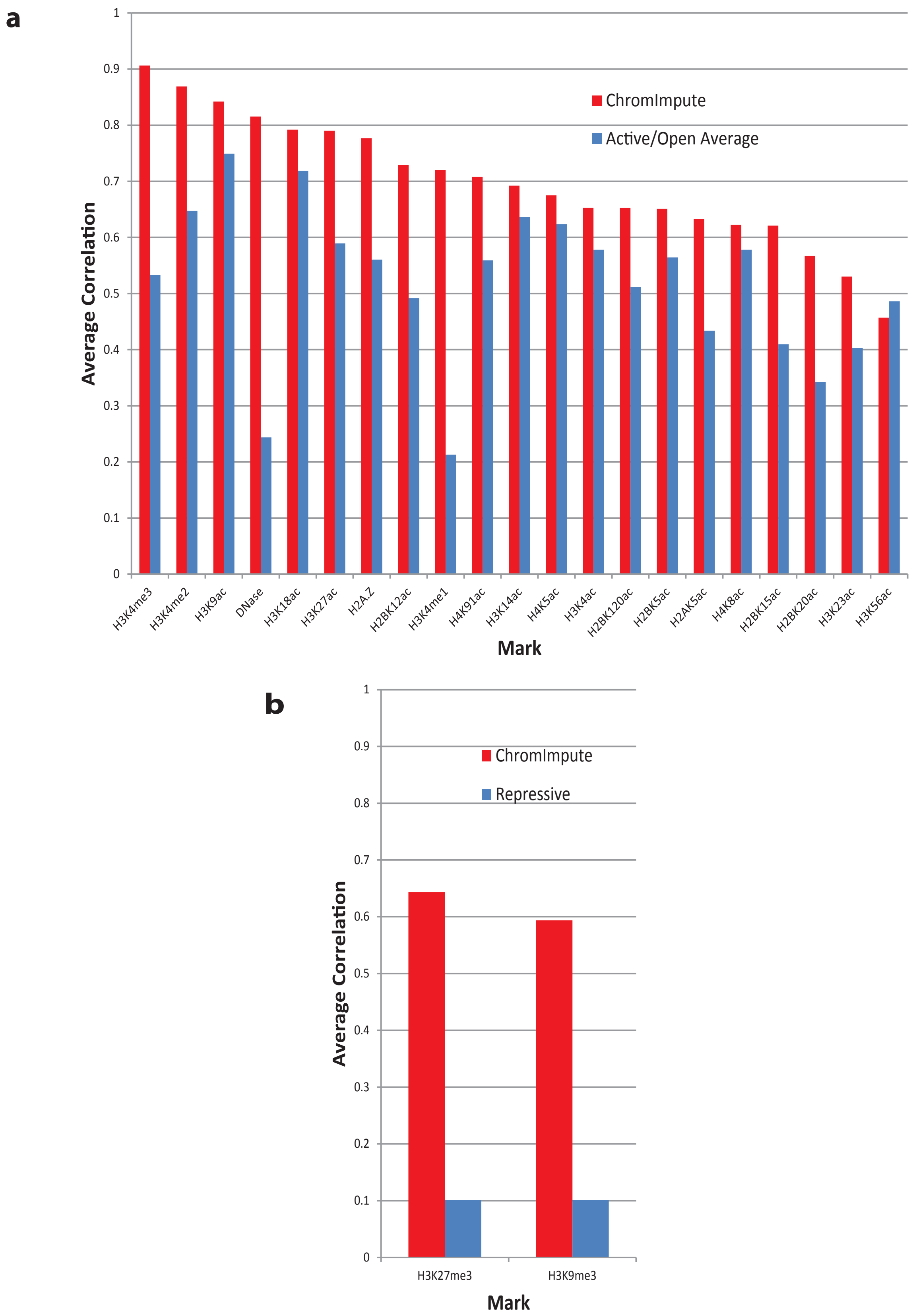
Supplementary Figure 9: Methodological Comparisons.
The graphs show a comparison based on **(a)** Correlation and **(b)** Top 1% agreement metrics for: (1) the standard ChromImpute which is based on an ensemble of regression trees, (2) ChromImpute with the same features and ensemble training strategy except using linear regression opposed to regression trees, (3) ChromImpute with regression trees trained on only on a single sample which was chosen to be the globally most correlated based on H3K4me1, except using H3K4me3 when training for H3K4me1. (4) Predictions based on averaging the target mark in up to the 10 nearest-neighboring samples having the target mark where the distance is determined based on the global correlation distance measure with H3K4me1, except using H3K4me3 when trying to predict H3K4me1. (5) The same as in (4) except using the local Euclidean distance opposed to the global correlation. Evaluation was limited to chr10.



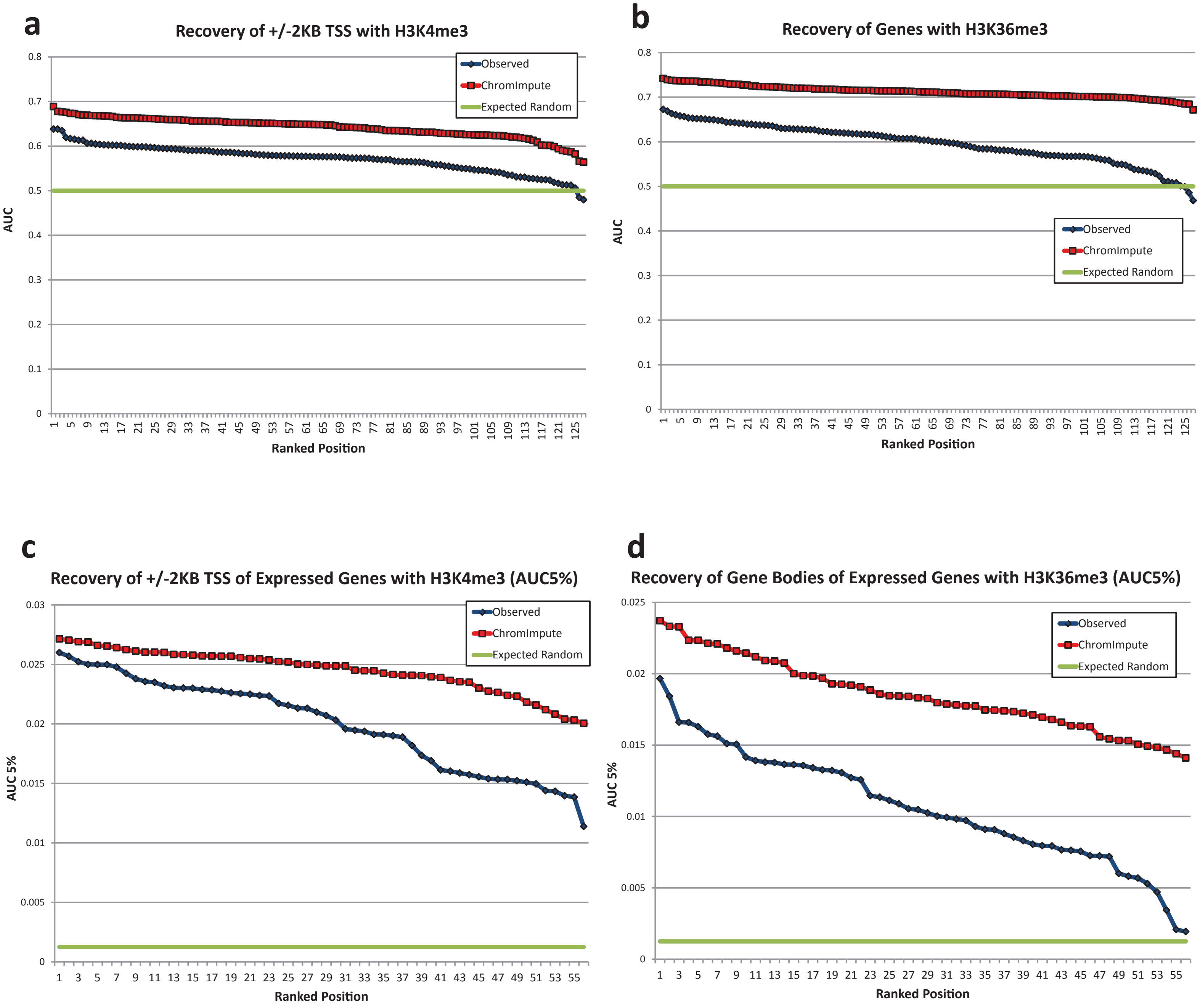
Supplementary Figure 10:

Comparison of Imputed Data and ES/iPSC Near-Replicate Predictions. The figures show for the tier 1 histone marks a comparison of the performance of the imputed data on average for ESC and iPSC samples compared with what would be expected by treating the other ESC and iPSC samples as effective replicates and averaging their values as a function of the number of replicates in terms of **(a)** correlation and **(b)** top 1% agreement. To compute the expectation for k-replicates for each 25-bp interval we randomly selected k of the ESC and iPSC samples which had the mark mapped excluding the sample being evaluated. The performance of the imputed data in comparison is shown with the horizontal line in red.

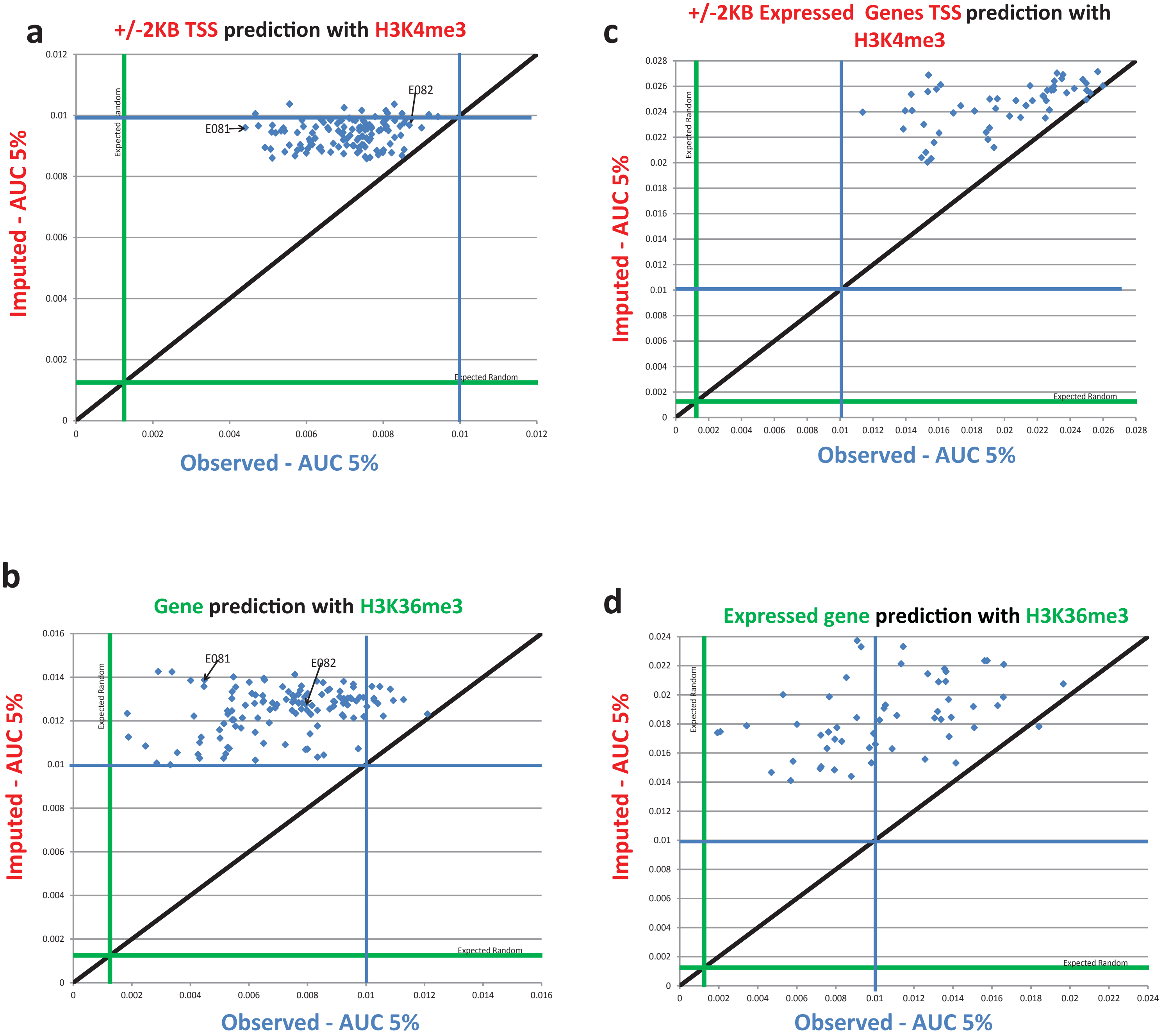

b
H3K4me1
H3K4me3
H3K36me3
H3K9me3
H3K27me3
H3K27ac
H3K9ac
Top 1% Agreement
0
10
20
30
40
50
60
70
80
90
100
0
1
2
3
4
5
6
7
8
9
10
11
12
Number of ESC/iPSC Near-Replicates
Imputed
ESC/iPSC Near-Replicates



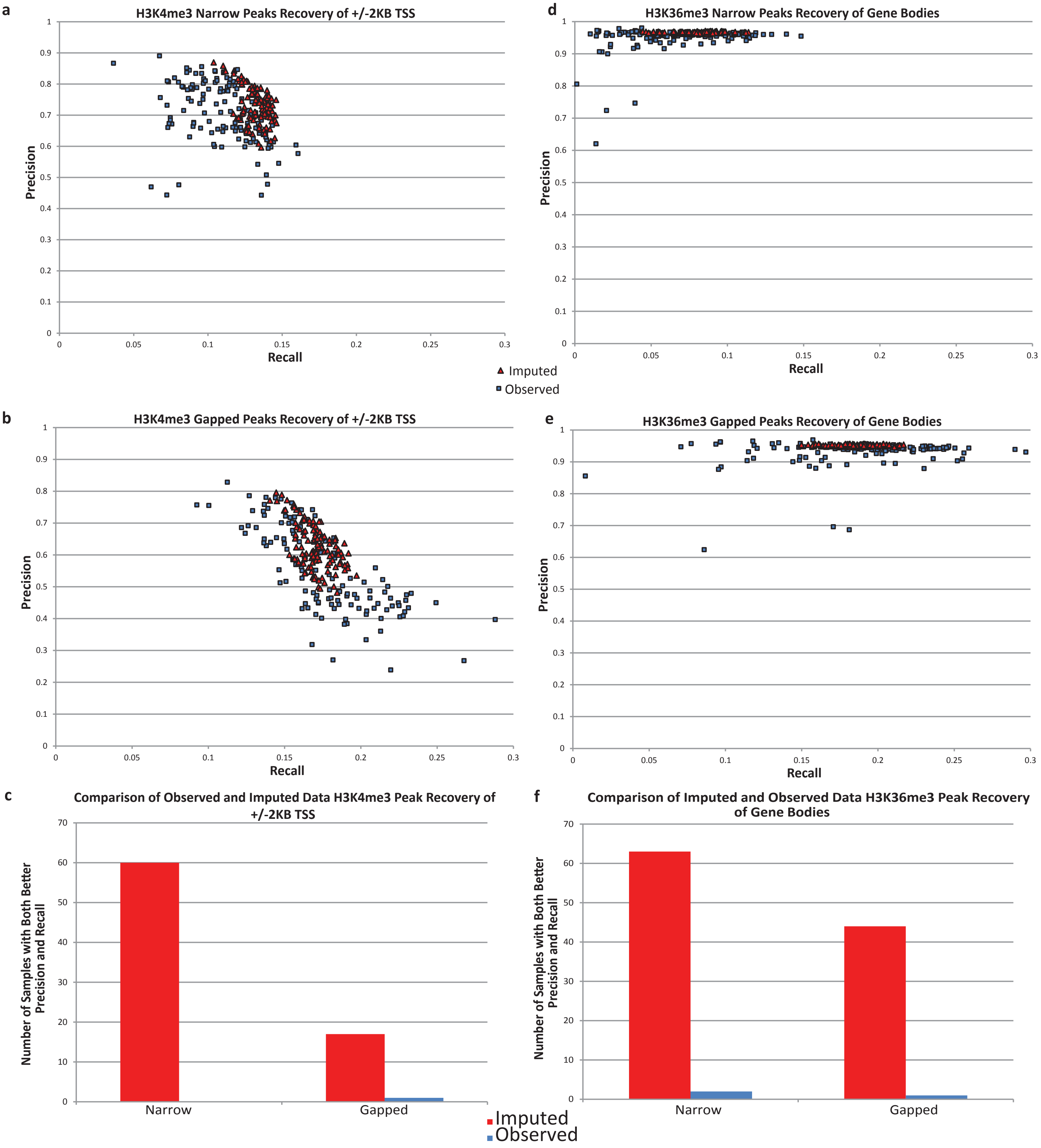
Supplementary Figure 11: Comparison with Averaging Subset of Other Marks in the Same Sample. (a) A comparison of predicting a set of putative active/open marks (H3K4me1/2/3, H2A.Z, DNase, and acetylations) by averaging all other such marks from the same sample compared with ChromImpute predictions evaluated based on average correlation with the observed data. (b) Predicting the repressive marks H3K9me3 and H3K27me3 with the other repressive mark in the same sample compared with ChromImpute predictions evaluated based on average correlation with observed data.



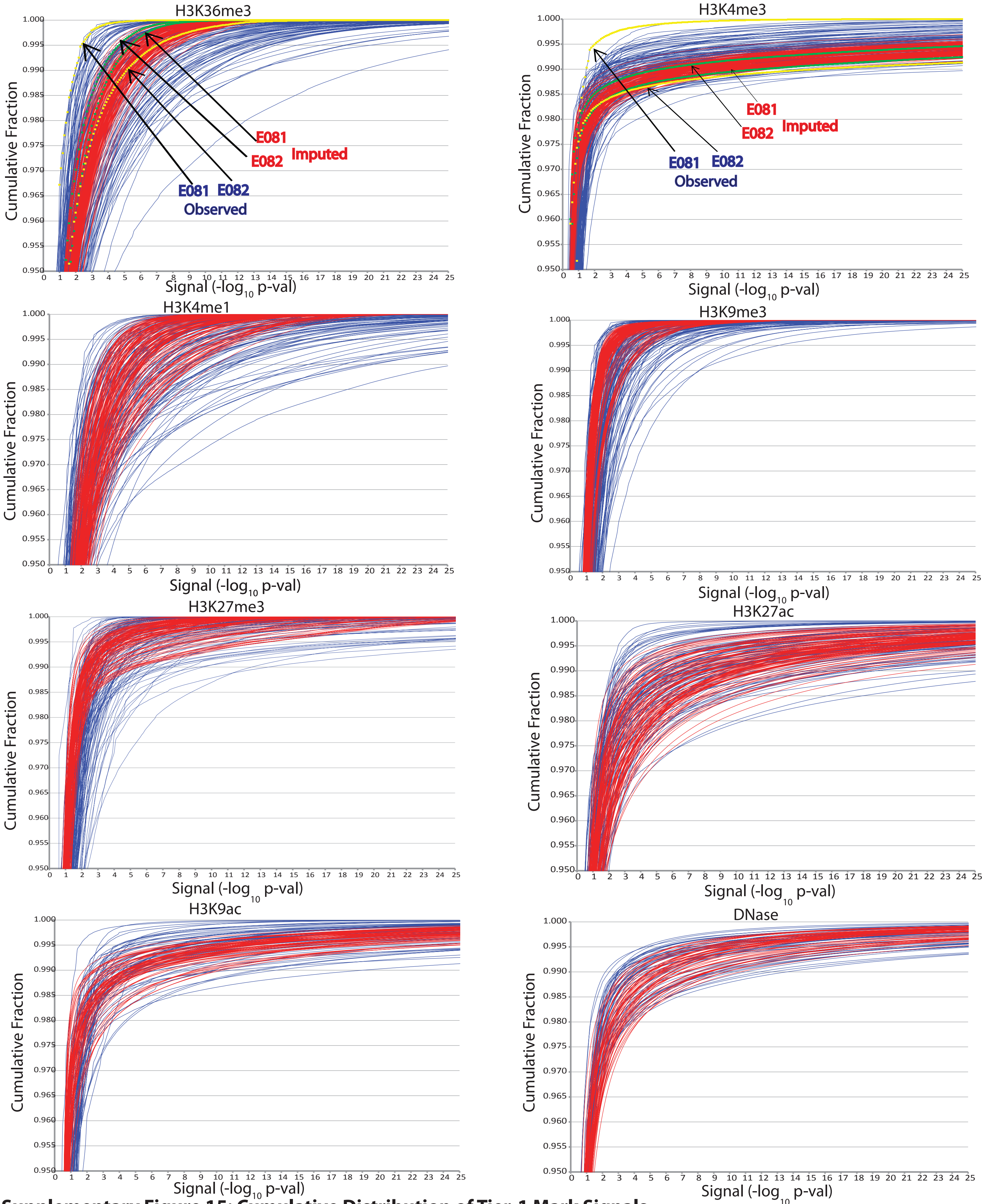
Supplementary Figure 12: Observed and Imputed Mark Recovery of Genomic Features
These are similar plots to main **Fig 3a,b** except **(a,b)** reporting the full AUC for **(a)** H3K4me3 recovery of +/-2KB TSS **(b)** H3K36me3 recovery of gene bodies **(c,d)** reporting the AUC up to a 5% false positive rate based on a set of expressed genes (see **Methods**) for **(c)** the H3K4me3 signal recovering locations within 2kb of the TSS of these genes for the 56 samples with gene expression data available and **(d)** the same as (c) except for gene regions and the H3K36me3 signal.



Supplementary Figure 13:
Scatter Plot Comparison of Observed vs. Imputed Signal at Recovery of Annotated Gene Features. The x-axis of these plots correspond to the area under the ROC curve up to a 5% false positive rate for the observed signal, while the y-axis shows it for the imputed data for **(a)** H3K4me3 recovery of locations within 2kb of annotated transcription start sites **(b)** H3K36me3 recovery of locations within annotated genes **(c)** H3K4me3 recovery of locations within 2kb of annotated transcription start sites for an expressed gene set (see **Methods**) **(d)** H3K36me3 recovery of locations within the expressed gene set (see **Methods**). The black line shows the y=x line, demonstrating in almost all cases the imputed data has better agreement with the annotated gene features. The green lines illustrate what would be expected by random guessing, and the blue lines mark consistently the 0.01 values in each figure.

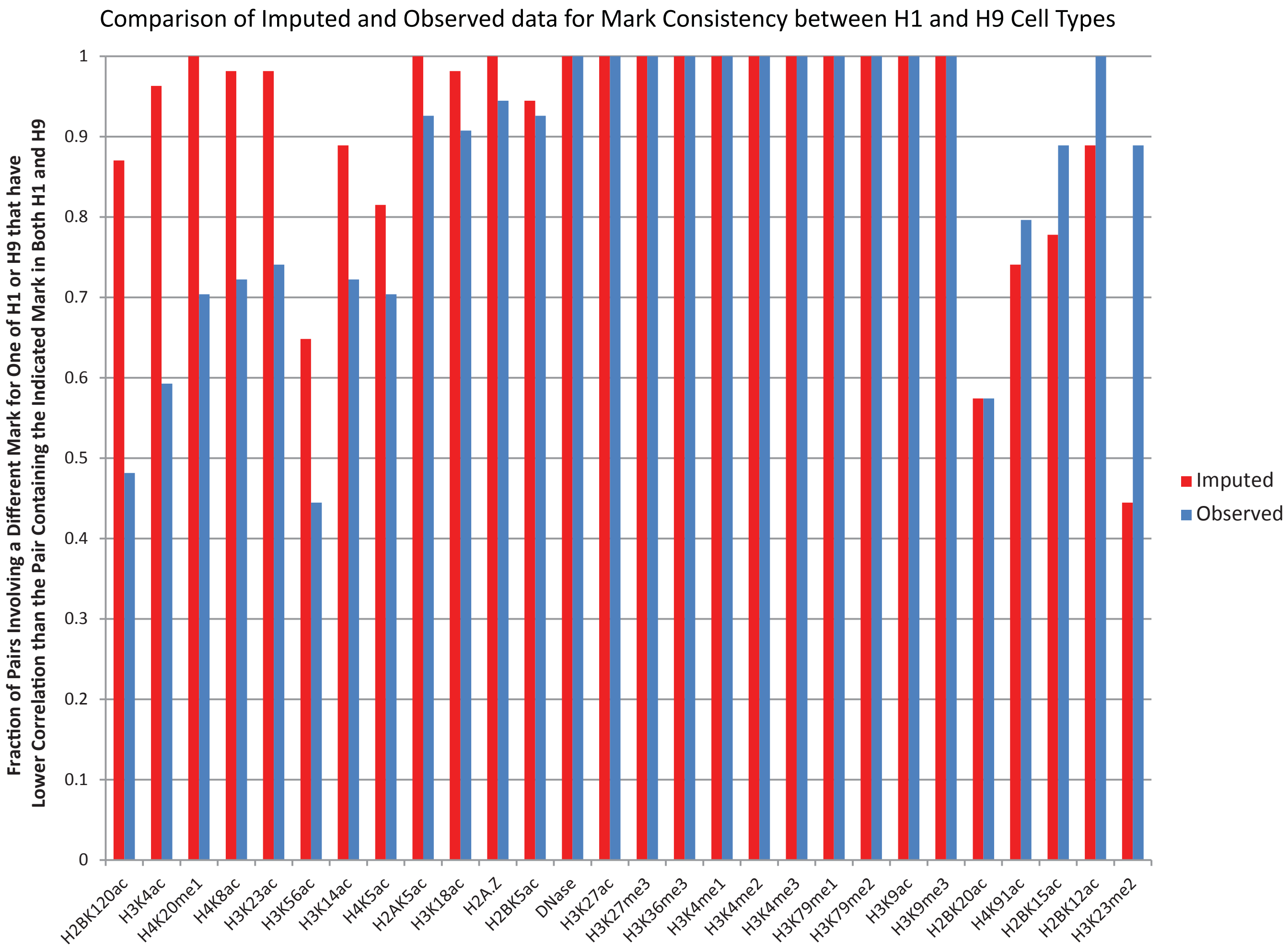


Supplementary Figure 14: Peak Calls Agreement with Annotated Features.
(a) Precision and recall of narrow peak calls for H3K4me3 overlap with locations within 2kb of annotated TSS. Each dot in red corresponds to a peak call set based on the imputed and in blue based on the observed data **(b)** Same as **a** except for gapped peaks. **(c)** The number of samples for which the peak calls based on imputed (observed) H3K4me3 data had both better precision and recall than the corresponding observed (imputed) data shown for separate comparisons based on narrow and gapped peak calls. Peak call sets which were not better in both precision and recall were not counted. **(d-f)** The same as **a-c** except for H3K36me3 overlap with locations within annotated genes.



Supplementary Figure 15: Cumulative Distribution of Tier-1 Mark Signals.

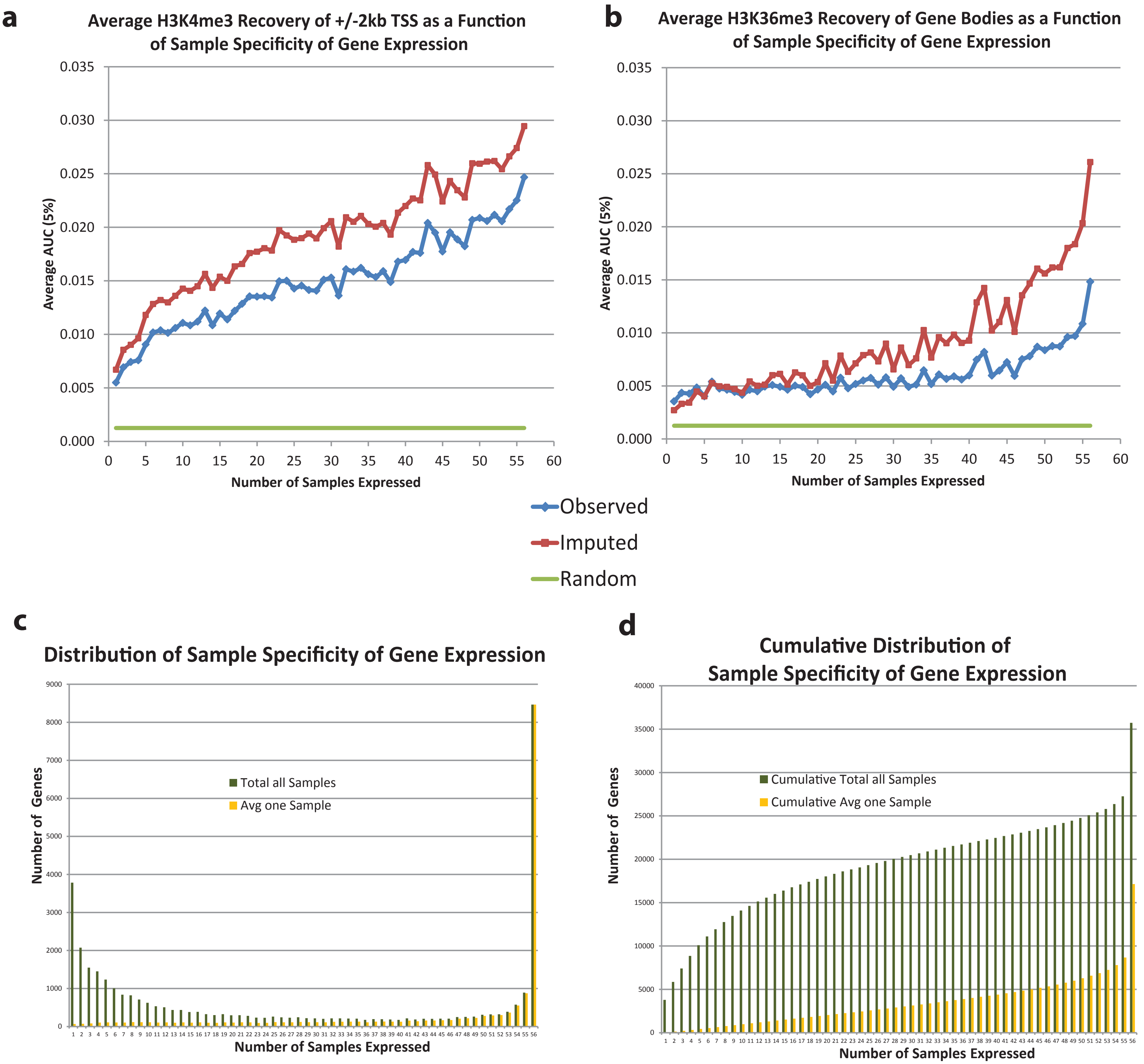
The figure shows the cumulative distribution function plots for the eight Tier-1 marks for each sample based on the observed data in blue and the imputed data in red. These plots show the imputed signal has a more consistent distribution across samples. For H3K4me3 and H3K36me3, in yellow are the cumulative distribution for the observed data of two Fetal brain samples (E081 and E082), while in green for the imputed data, showing even for the same tissue type that the distribution of observed signal can be very different.



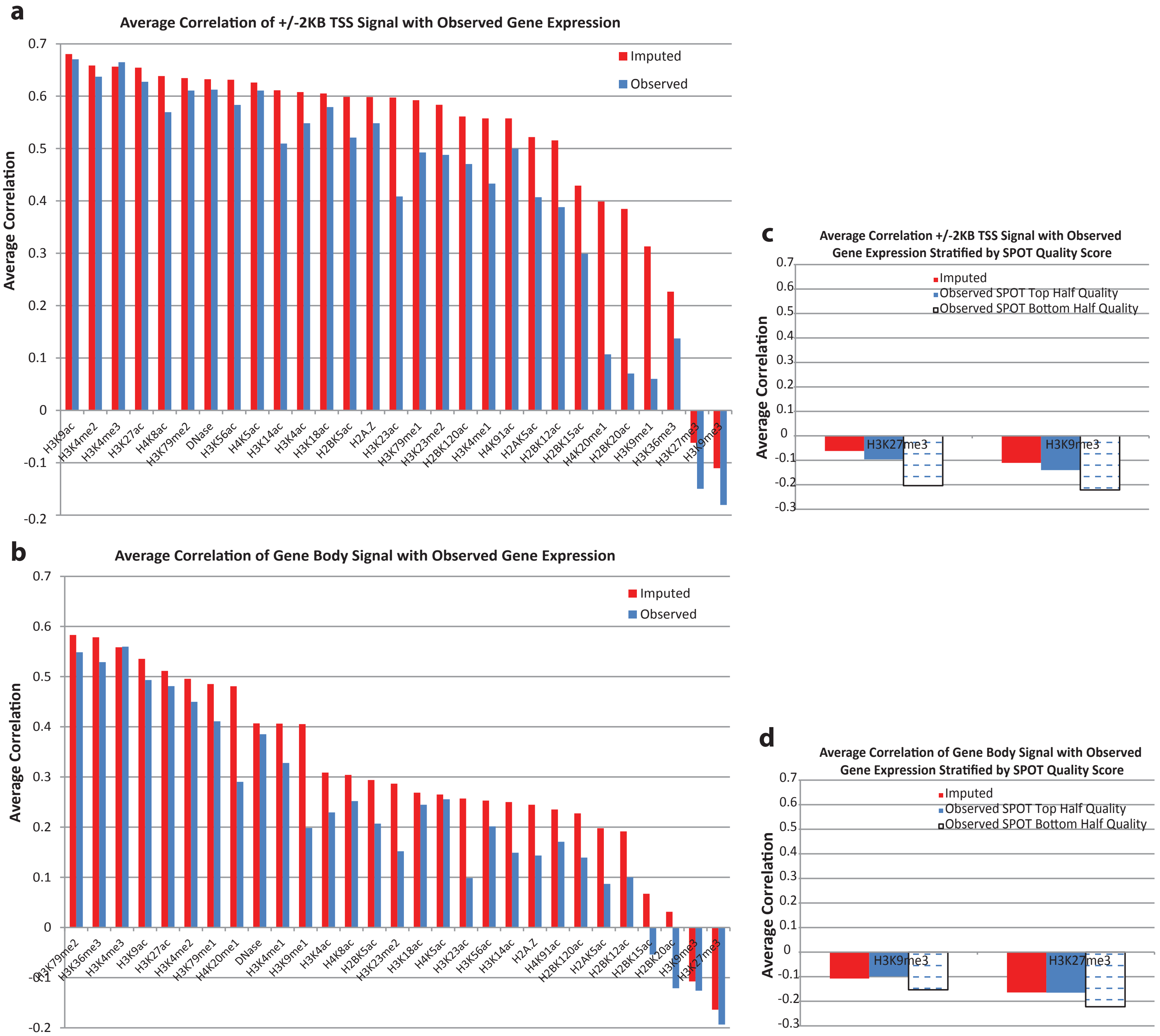
Supplementary Figure 16:

Comparison of Observed and Imputed Relative Mark Agreement Between H1 and H9.

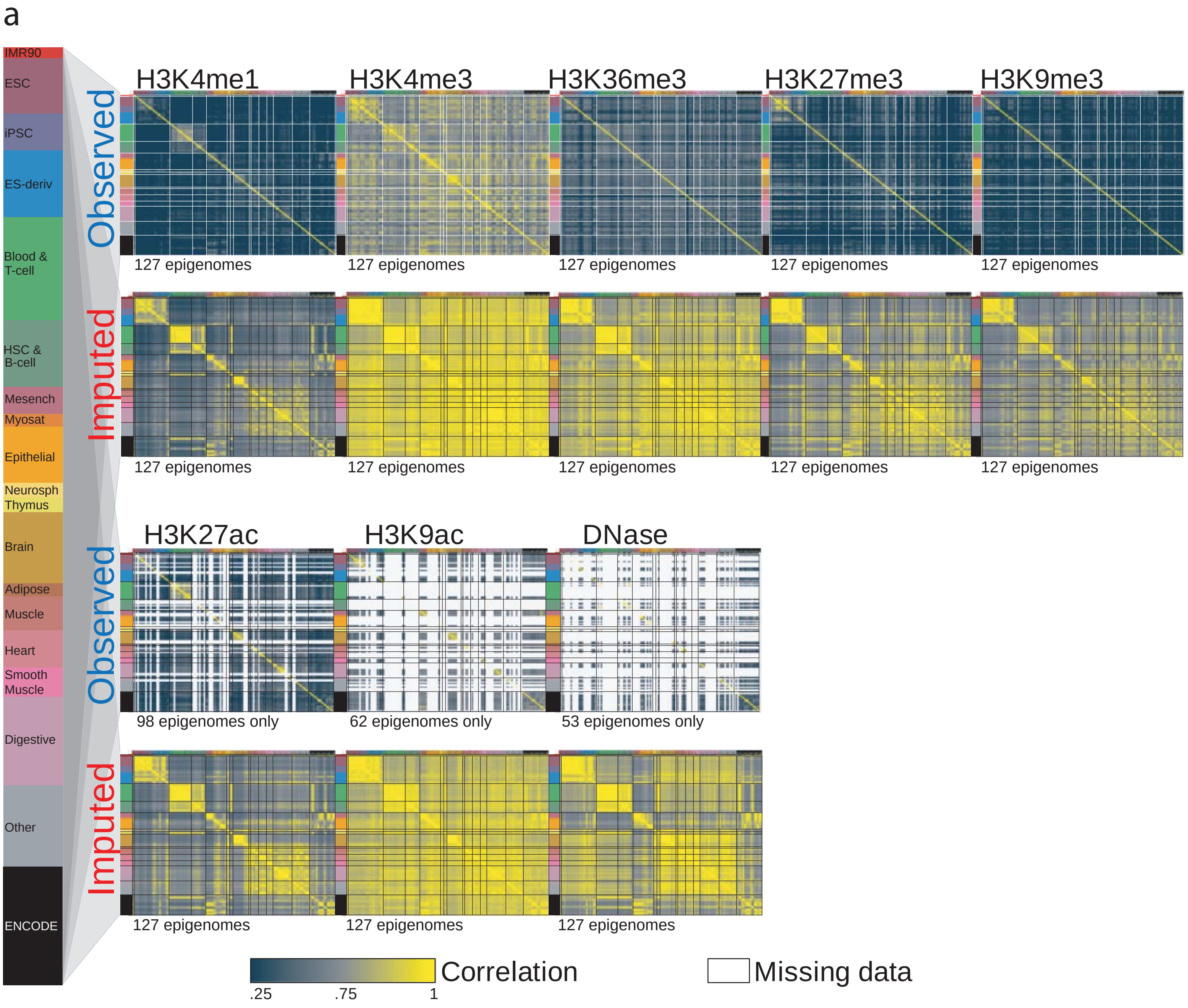
This evaluates for a mark A, how frequently the correlation for the pair (A_{H1}, A_{H9}) is greater than the correlation of any other pair (A_{H1}, B_{H9}) or (B_{H1}, A_{H9}) where B is a mark other than A and the subscript indicates the sample of the experiment which is either of two embryonic stem cell samples, from the H1 or H9 cell lines. This evaluation is done separately for the observed and imputed data. In total the imputed and observed data had different relative agreement on 16 marks, with the imputed data having better relative agreement for 12 of these marks, which is significant ($p<0.04$) based on a binomial test.



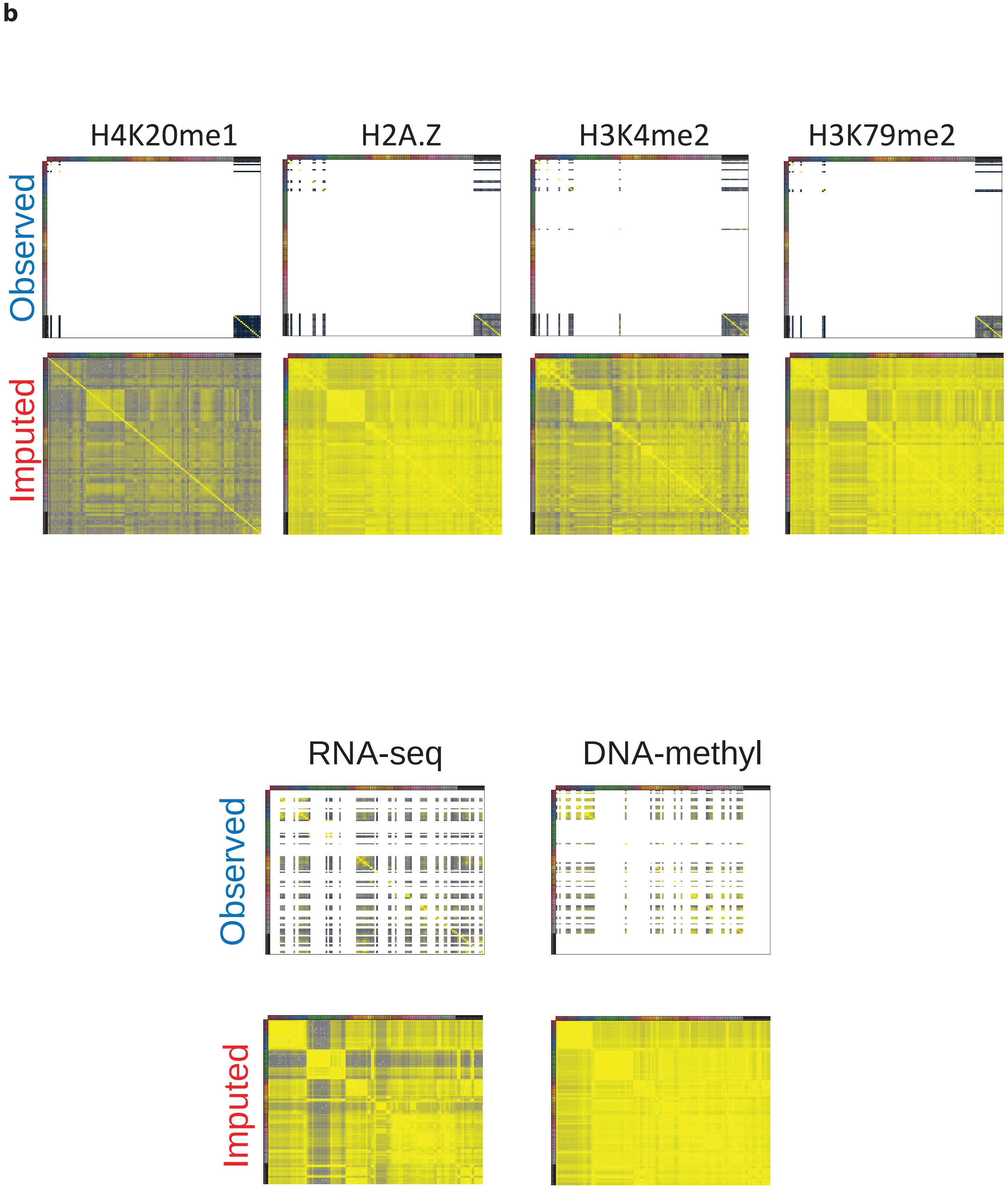
Supplementary Figure 17:
Observed and Imputed Mark Recovery of Genomic Features as a Function of Sample Specificity.
(a) The area under the ROC curve up to a 5% false positive rate for recovering bases +/-2kb of a TSS of an expressed gene in a sample (defined as RPKM ≥ 0.5) conditioned on the number of samples in which the gene is expressed based on the imputed and observed H3K4me3 signal. The negative set is all other bases in the genome except positions corresponding to an expressed gene that are not expressed in the number of samples being conditioned on. The reported AUC is averaged over all the samples with expression data available. **(b)** The same as **a** except for H3K36me3 and gene bodies. **(c)** In green the count of the total number of genes expressed in the specified number of samples. In yellow the average count of the number of genes expressed in the specified number of samples among expressed genes in a given sample. The imputed data has better or similar performance in both evaluations except for recovering the gene bodies with H3K36me3 for a limited number of the most sample specific genes. **(d)** The same as **c** except showing the cumulative totals.



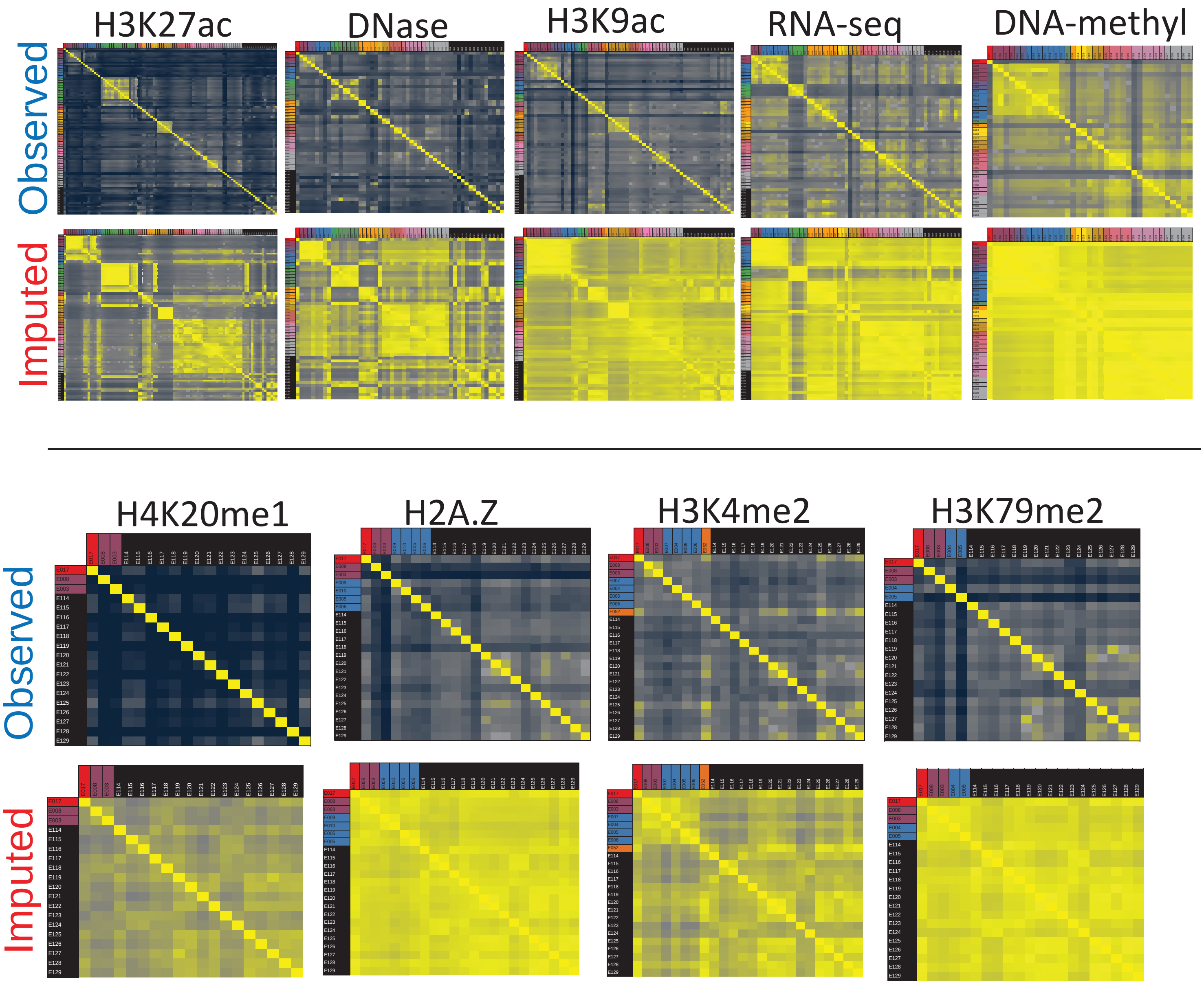
Supplementary Figure 18:
Comparison of Observed and Imputed Chromatin Mark Data Correlation with Observed Gene Expression. **(a)** For each of the tier 1-3 marks mapped in at least two samples for both the imputed and observed data the average correlation of the mark signal within +/-2KB TSS with gene expression. Signal was represented by computing the average of the signal values within the range adding one and then taking a log transformation. The correlations were computed separately for each sample with observed data available for the mark and gene expression, and then averaged. For most marks there was a greater positive correlation with gene expression for the imputed data compared to the observed data. **(b)** The same as **a**, but for annotated gene bodies. **(c)** The correlation of the two repressive marks, H3K27me3 and H3K9me3, for +/-2KB TSS signal with gene expression with the observed data stratified based on whether the sample is in the top half or bottom half of observed datasets considered by the SPOT quality score[10] compared with the imputed data correlation. This shows the stronger negative correlative with gene expression for the observed data for the repressive marks is associated with data sets with lower quality as scored by SPOT. **(d)** The same as **c** but for gene bodies.



Supplementary Figure 19: Sample correlations for imputed vs. observed signal.
The same heatmap of pairwise correlations for observed and imputed data as shown in **Fig. 3e**, but showing **(a)** all the Tier-1 marks **(b)** the Tier-2 marks, RNA-seq, and DNA-methylation **(c)** heatmap of pairwise correlations for observed and imputed data of the Tier-1 marks not mapped in every sample, RNA-seq, DNA-methylation, and Tier-2 marks only showing the subset of samples for which there is observed data available **(d)** heatmap of pairwise correlations for the imputed data for the Tier-3 marks.

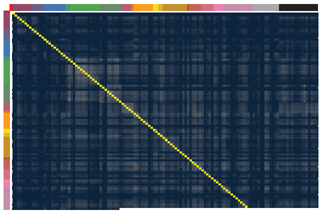

b
H4K20me1
H2A.Z
H3K4me2
H3K79me2
Observed
Imputed
RNA-seq
DNA-methyl
Observed
Imputed

c


H3K27ac
DNase
H3K9ac
RNA-seq
DNA-methyl
Observed
Imputed
H4K20me1
H2A.Z
H3K4me2
H3K79me2
Observed
Imputed
E017
E008
E003
E114
E115
E116
E117
E118
E119
E120
E121
E122
E123
E124
E125
E126
E127
E128
E129

H2AK5ac

H2AK9ac



H2BK5ac

H2BK12ac

H2BK15ac

H2BK20ac

H2BK120ac

H3K4ac

H3K14ac

H3K18ac

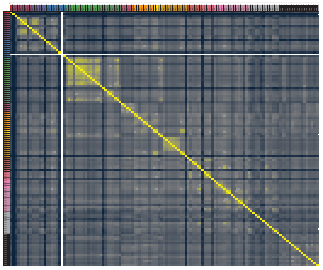
H3K23ac

H3K56ac

H4K5ac

H4K8ac

H4K12ac



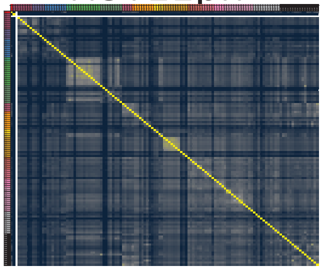
H4K91ac

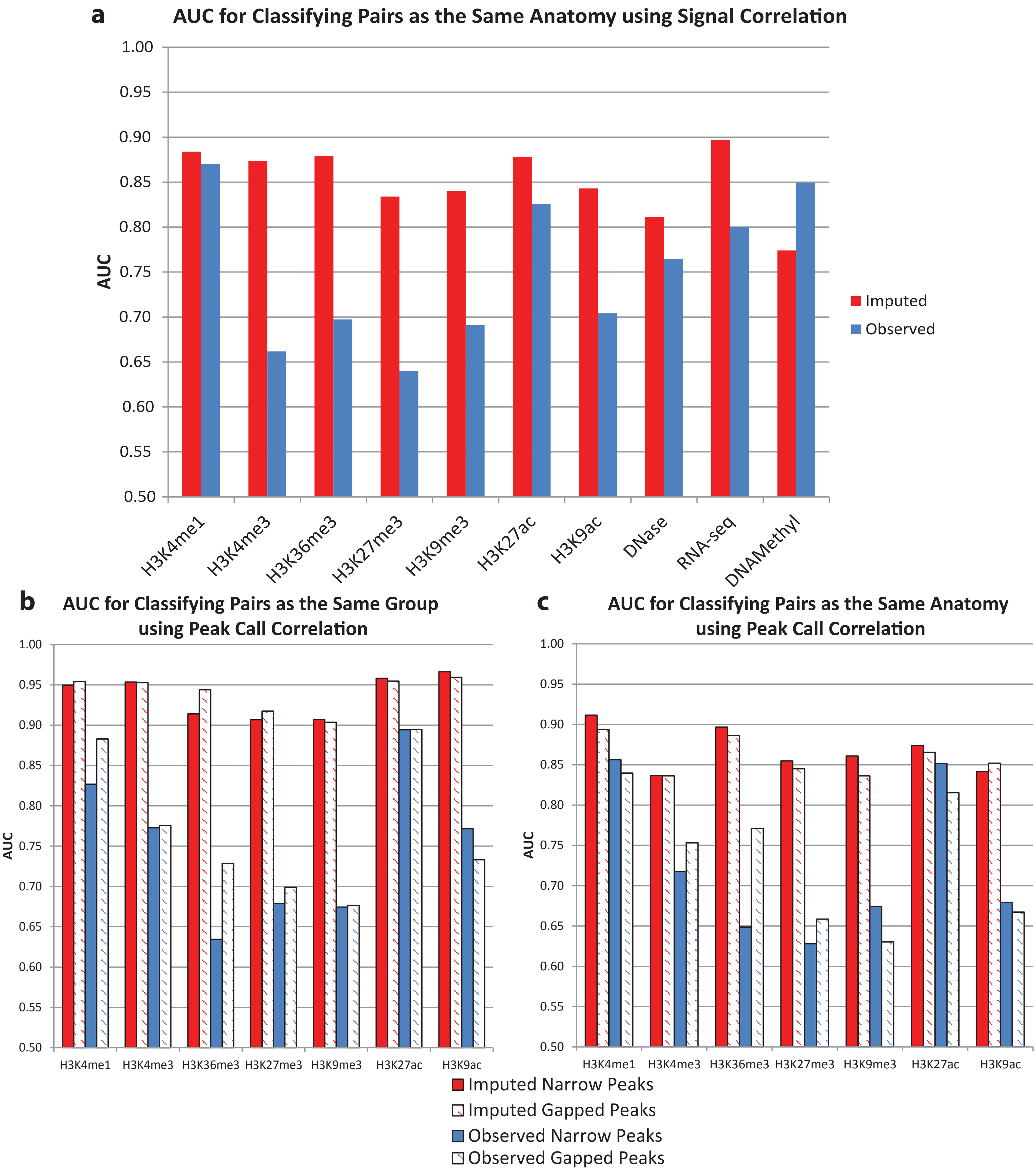
H3K9me1

H3K23me2

H3K79me1

H3T11ph

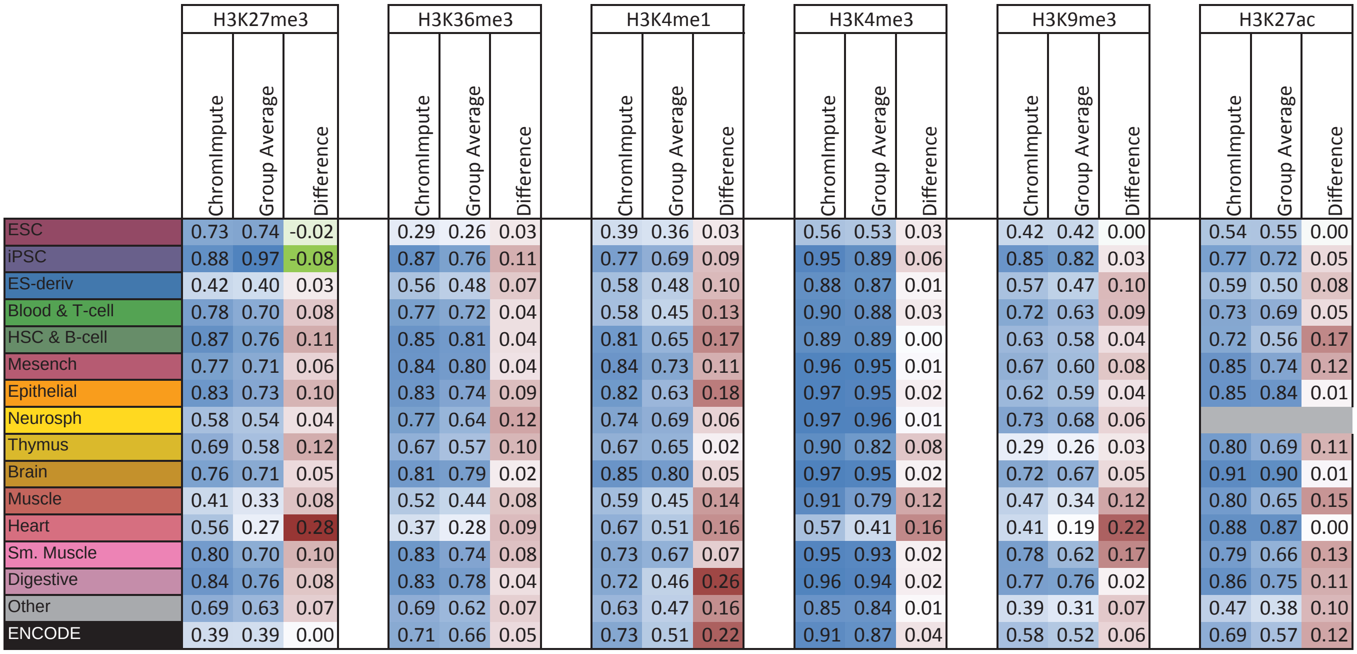



a AUC for Classifying Pairs as the Same Anatomy using Signal Correlation

b AUC for Classifying Pairs as the Same Group using Peak Call Correlation

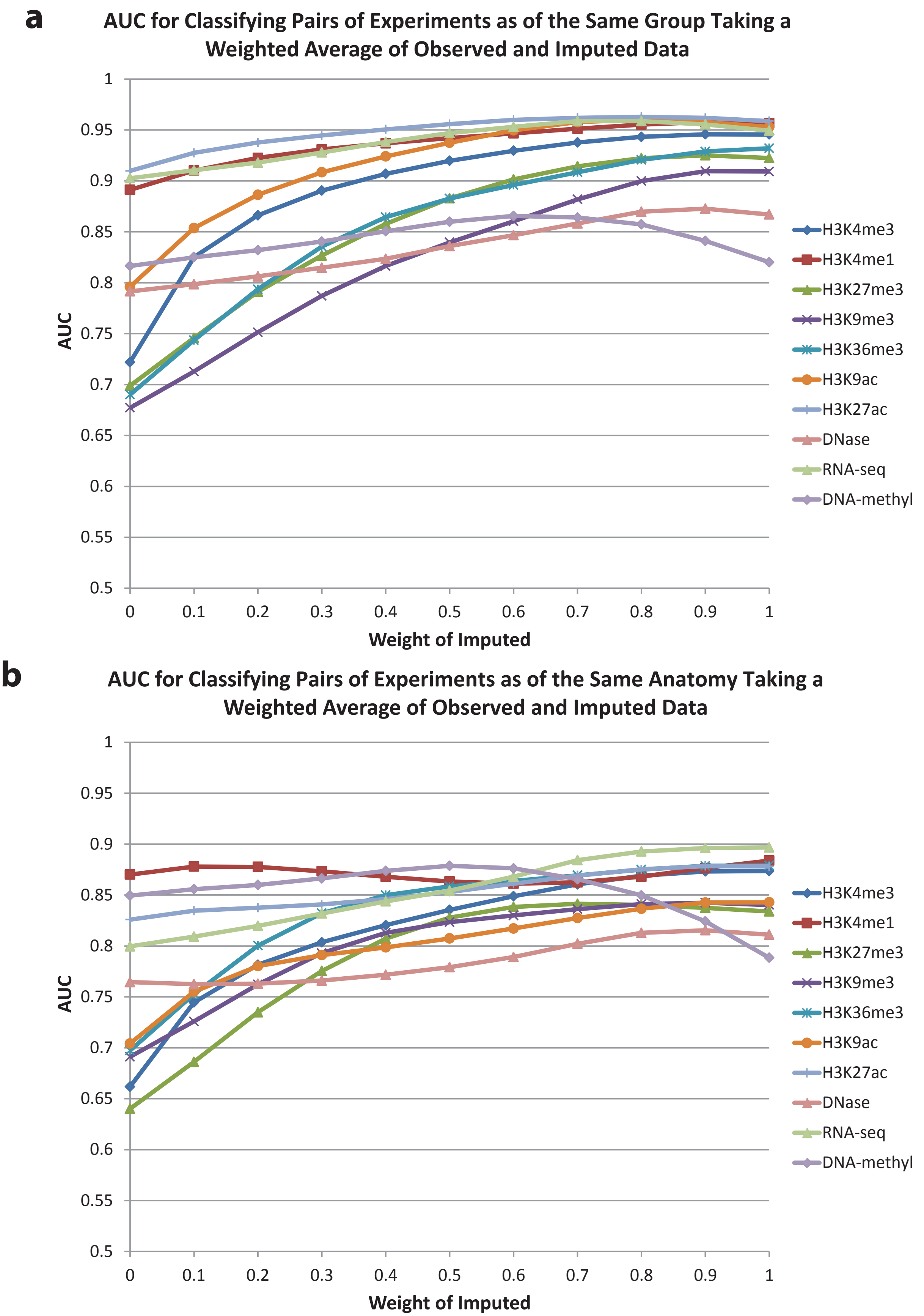
c AUC for Classifying Pairs as the Same Anatomy using Peak Call Correlation

Supplementary Figure 20: Prediction of Same Anatomy and Group Pairs.
(a) This is the same evaluation of mark signal correlation pairs as in **Fig. 3f** except for anatomy annotations[10] (**Table S1**) opposed to sample group annotations. **(b)** The same evaluation as in **Fig. 3f** except based on histone mark peak calls with results shown for both gapped and narrow peak calls. The correlations were computed the same as for the signal except being at the base level and treating a base as having a signal of 1 if it was covered by a peak call and 0 otherwise. **(c)** The same evaluation as in **b**, but for anatomy annotations.



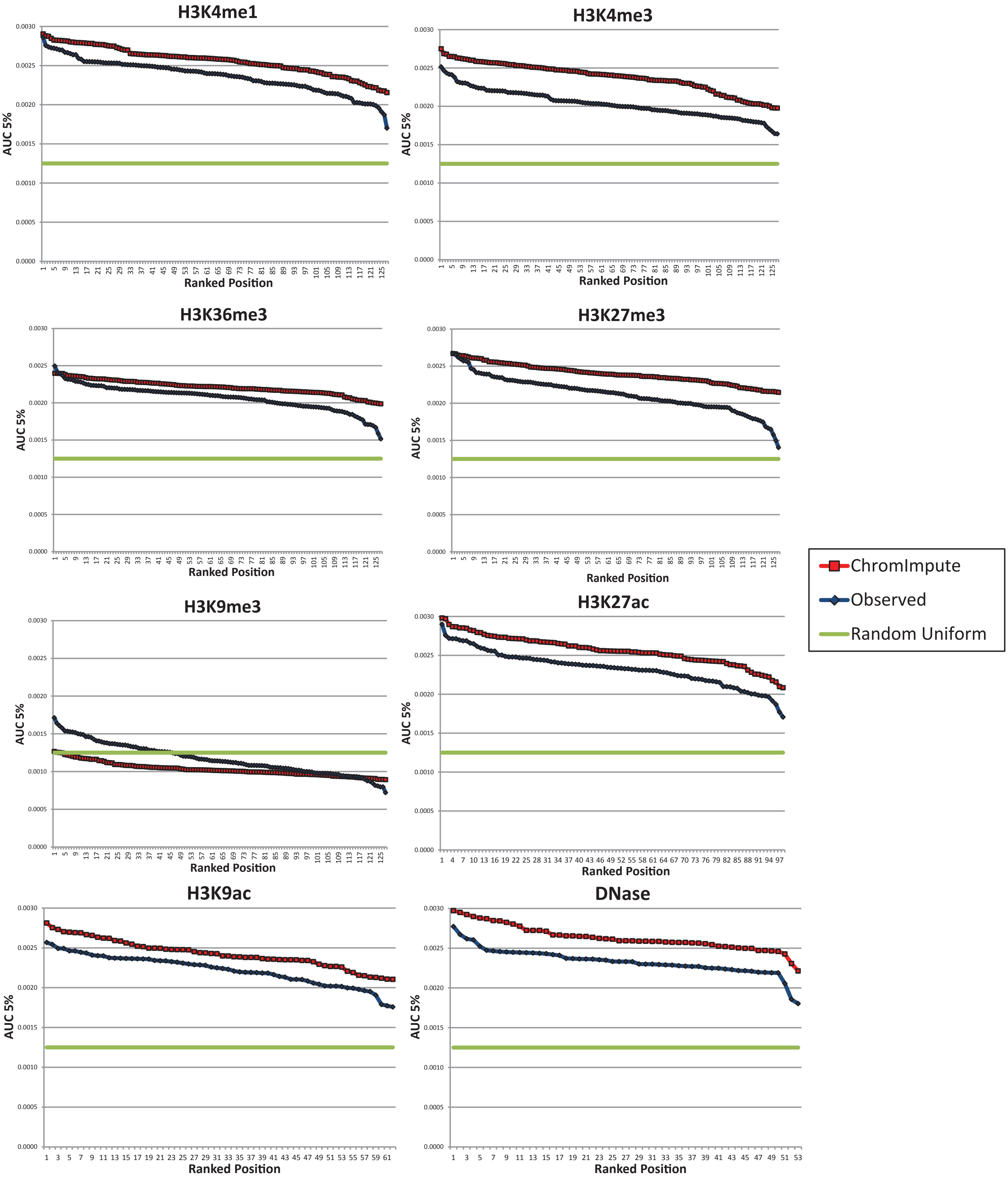
	H3K27me3			H3K36me3			H3K4me1			H3K4me3			H3K9me3			H3K27ac		
	ChromImpute	Group Average	Difference	ChromImpute	Group Average	Difference	ChromImpute	Group Average	Difference	ChromImpute	Group Average	Difference	ChromImpute	Group Average	Difference	ChromImpute	Group Average	Difference
ESC	0.73	0.74	-0.02	0.29	0.26	0.03	0.39	0.36	0.03	0.56	0.53	0.03	0.42	0.42	0.00	0.54	0.55	0.00
iPSC	0.88	0.97	-0.08	0.87	0.76	0.11	0.77	0.69	0.09	0.95	0.89	0.06	0.85	0.82	0.03	0.77	0.72	0.05
ES-deriv	0.42	0.40	0.03	0.56	0.48	0.07	0.58	0.48	0.10	0.88	0.87	0.01	0.57	0.47	0.10	0.59	0.50	0.08
Blood & T-cell	0.78	0.70	0.08	0.77	0.72	0.04	0.58	0.45	0.13	0.90	0.88	0.03	0.72	0.63	0.09	0.73	0.69	0.05
HSC & B-cell	0.87	0.76	0.11	0.85	0.81	0.04	0.81	0.65	0.17	0.89	0.89	0.00	0.63	0.58	0.04	0.72	0.56	0.17
Mesench	0.77	0.71	0.06	0.84	0.80	0.04	0.84	0.73	0.11	0.96	0.95	0.01	0.67	0.60	0.08	0.85	0.74	0.12
Epithelial	0.83	0.73	0.10	0.83	0.74	0.09	0.82	0.63	0.18	0.97	0.95	0.02	0.62	0.59	0.04	0.85	0.84	0.01
Neurosph	0.58	0.54	0.04	0.77	0.64	0.12	0.74	0.69	0.06	0.97	0.96	0.01	0.73	0.68	0.06			
Thymus	0.69	0.58	0.12	0.67	0.57	0.10	0.67	0.65	0.02	0.90	0.82	0.08	0.29	0.26	0.03	0.80	0.69	0.11
Brain	0.76	0.71	0.05	0.81	0.79	0.02	0.85	0.80	0.05	0.97	0.95	0.02	0.72	0.67	0.05	0.91	0.90	0.01
Muscle	0.41	0.33	0.08	0.52	0.44	0.08	0.59	0.45	0.14	0.91	0.79	0.12	0.47	0.34	0.12	0.80	0.65	0.15
Heart	0.56	0.27	0.28	0.37	0.28	0.09	0.67	0.51	0.16	0.57	0.41	0.16	0.41	0.19	0.22	0.88	0.87	0.00
Sm. Muscle	0.80	0.70	0.10	0.83	0.74	0.08	0.73	0.67	0.07	0.95	0.93	0.02	0.78	0.62	0.17	0.79	0.66	0.13
Digestive	0.84	0.76	0.08	0.83	0.78	0.04	0.72	0.46	0.26	0.96	0.94	0.02	0.77	0.76	0.02	0.86	0.75	0.11
Other	0.69	0.63	0.07	0.69	0.62	0.07	0.63	0.47	0.16	0.85	0.84	0.01	0.39	0.31	0.07	0.47	0.38	0.10
ENCODE	0.39	0.39	0.00	0.71	0.66	0.05	0.73	0.51	0.22	0.91	0.87	0.04	0.58	0.52	0.06	0.69	0.57	0.12

Supplementary Figure 21: Comparison of ChromImpute Predictions with Group Average. The table shows for the six marks most deeply profiled and each biological group the average correlation of ChromImpute predictions with the observed data compared to the correlation if the prediction was based on all other observed data sets for the mark within the group. The difference in the correlation between the two corresponding columns is also shown. The evaluation is based on just chr10. These results demonstrate that in almost all cases the ChromImpute predictions show better correspondence to the observed data than predictions based on the group average.

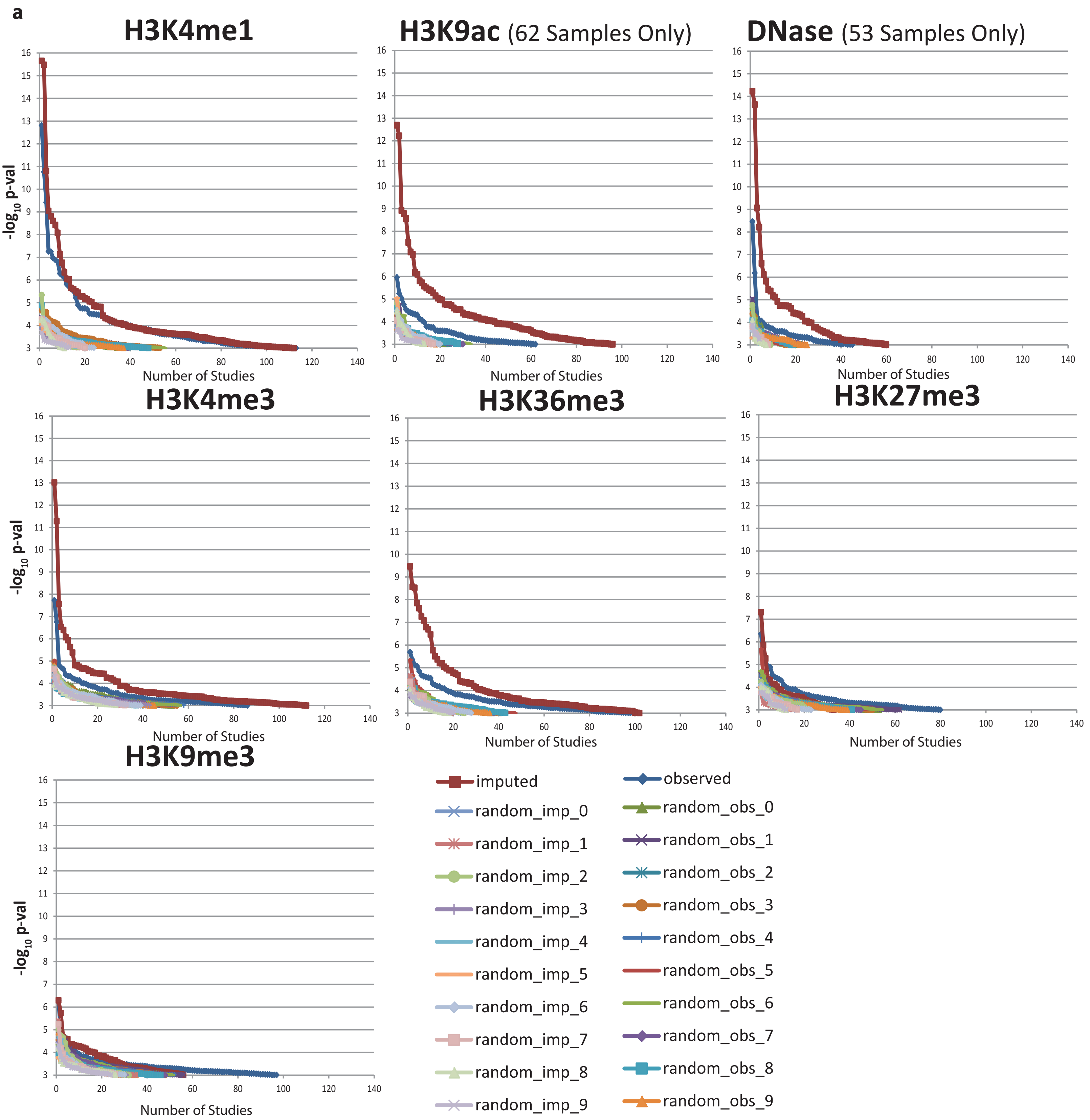

a AUC for Classifying Pairs of Experiments as of the Same Group Taking a Weighted Average of Observed and Imputed Data

b AUC for Classifying Pairs of Experiments as of the Same Anatomy Taking a Weighted Average of Observed and Imputed Data

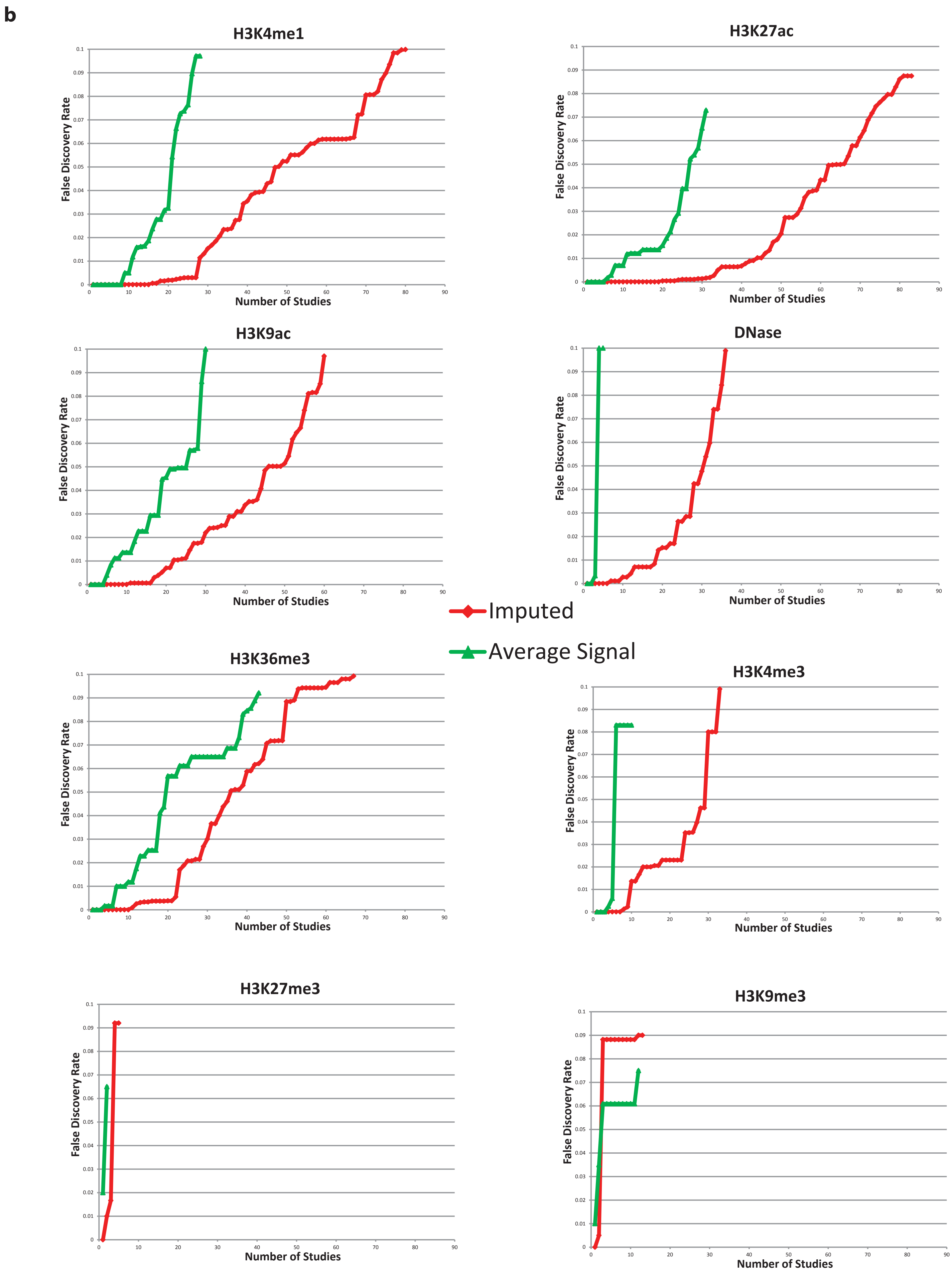
Supplementary Figure 22: Weighted Average Based Predictions of Same Group and Anatomy Pairs. This figure extends the analysis shown in **Fig. 3f** and **Fig. S20a** by showing what the AUC results would be if a weighted average of the observed and imputed data is taken. The weight of the imputed data is shown on the x-axis, with a weight of 0 reducing to the observed data and a weight of 1 reducing to the imputed data. Results for the weights in increments of 0.1 are shown. This shows for all marks except DNA-methylation, the AUC obtained with all weight on the imputed data is either the best or close to the best.

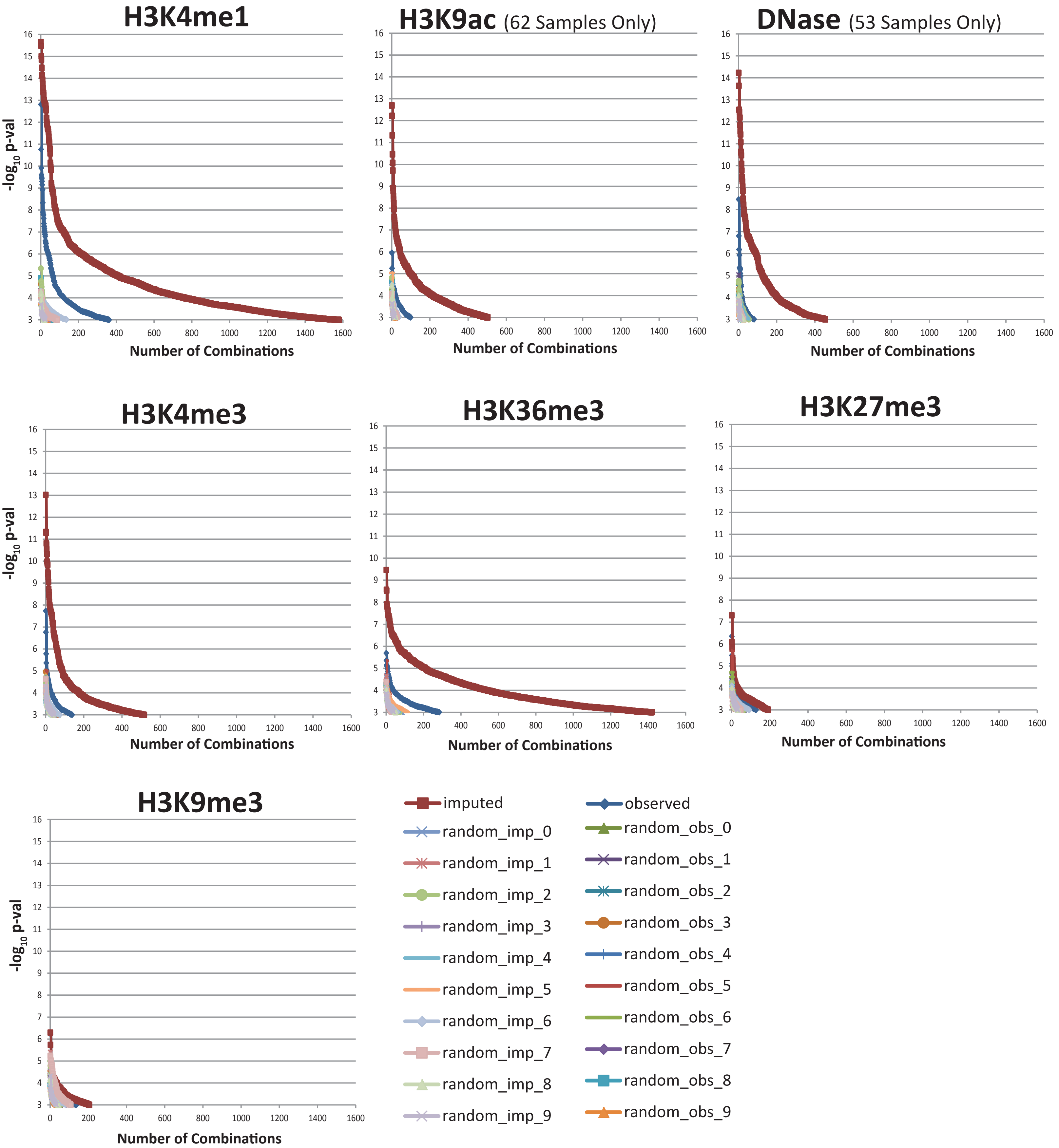


Supplementary Figure 23: Imputed and Observed Data Recovery of NHGRI GWAS Catalog SNPs. The figure compares the observed and imputed data for all Tier-1 marks across each sample where both is available in terms of the extent to which higher signal levels associates with increased recovery of SNPs in the GWAS catalog as measured by the area under the ROC curve up to a 5% false positive rate. The horizontal line is what expected by a uniform random baseline. Other baseline models could be assumed here, but the focus here is comparing the imputed and observed data. For all the marks considered except H3K9me3 the imputed data has better recovery than observed and uniform random baseline, while for H3K9me3 the imputed data shows a stronger depletion than the observed data and uniform baselines.

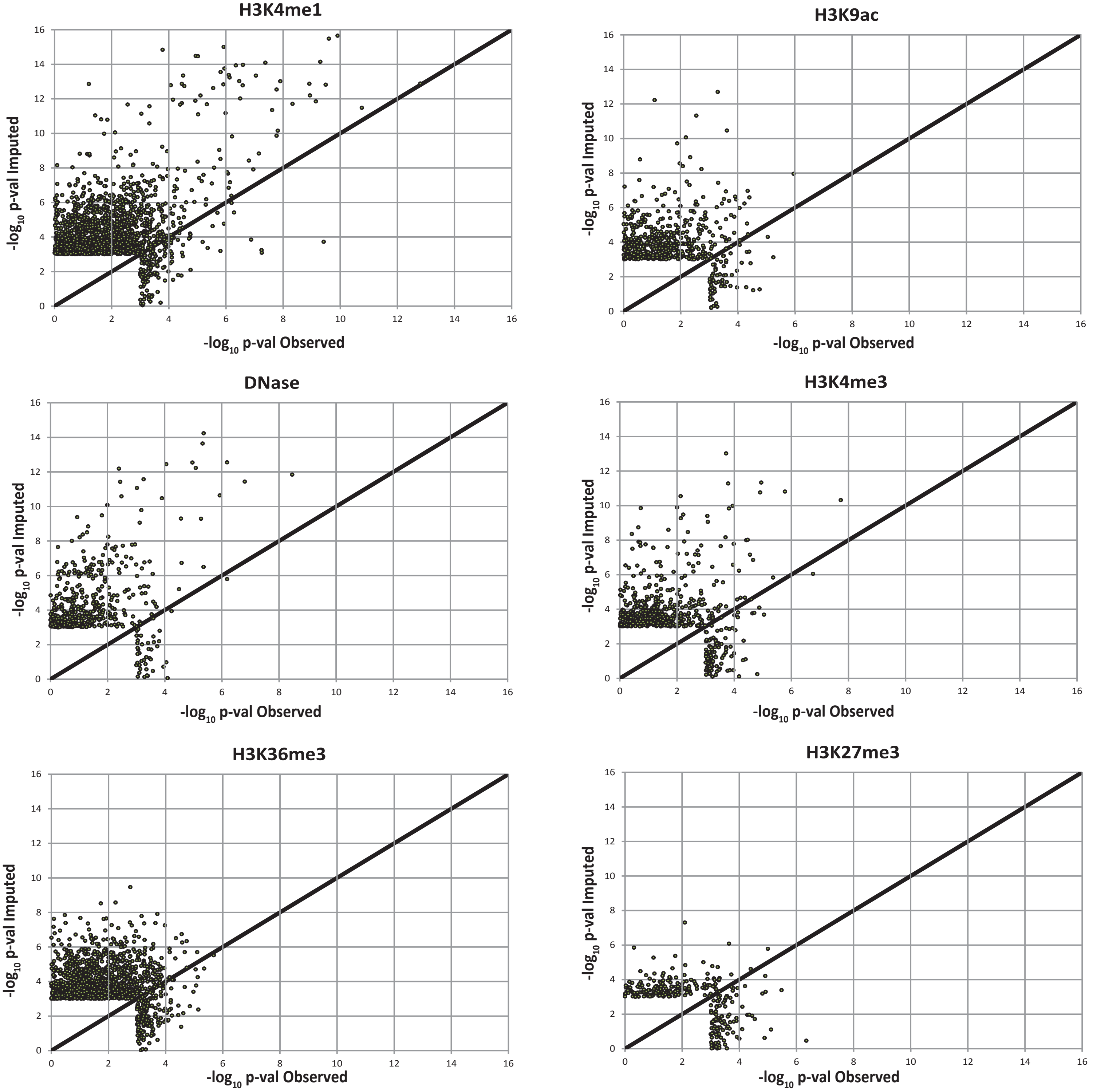


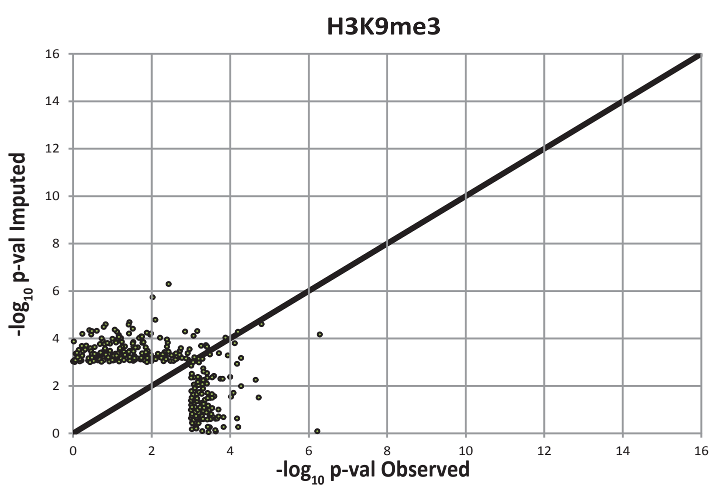
Supplementary Figure 24:
Imputed and Observed Data Correspondence with Genome-wide Association Studies (GWAS) – Max Sample. **(a)** Similar to **Fig. 4a** but for the other Tier-1 marks. The x-axis shows the number of studies for which there was at least one sample for which the indicated mark signal was significantly different for the study identified SNPs compared to a background of all GWAS Catalog SNPs at a significance level indicated on the y-axis based on a Mann-Whitney U test (see **Methods**). This is shown based on the imputed and observed data with the actual GWAS catalog, along with the observed and imputed data based on ten randomizations of the GWAS catalog. For H3K9ac and DNase the comparison was limited to only those samples where both imputed and observed data is available. **(b)** Shows for each of the Tier-1 marks the number of studies that are estimated to be significant in at least one sample based on the imputed data in samples with observed data available at an estimated false discovery rate below 10% and in comparison based on averaging the observed signal across all cell types (see **Methods**).


b
H3K4me1
H3K27ac
H3K9ac
DNase
H3K36me3
H3K4me3
H3K27me3
H3K9me3
False Discovery Rate
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
Number of Studies
0
10
20
30
40
50
60
70
80
90
Imputed
Average Signal

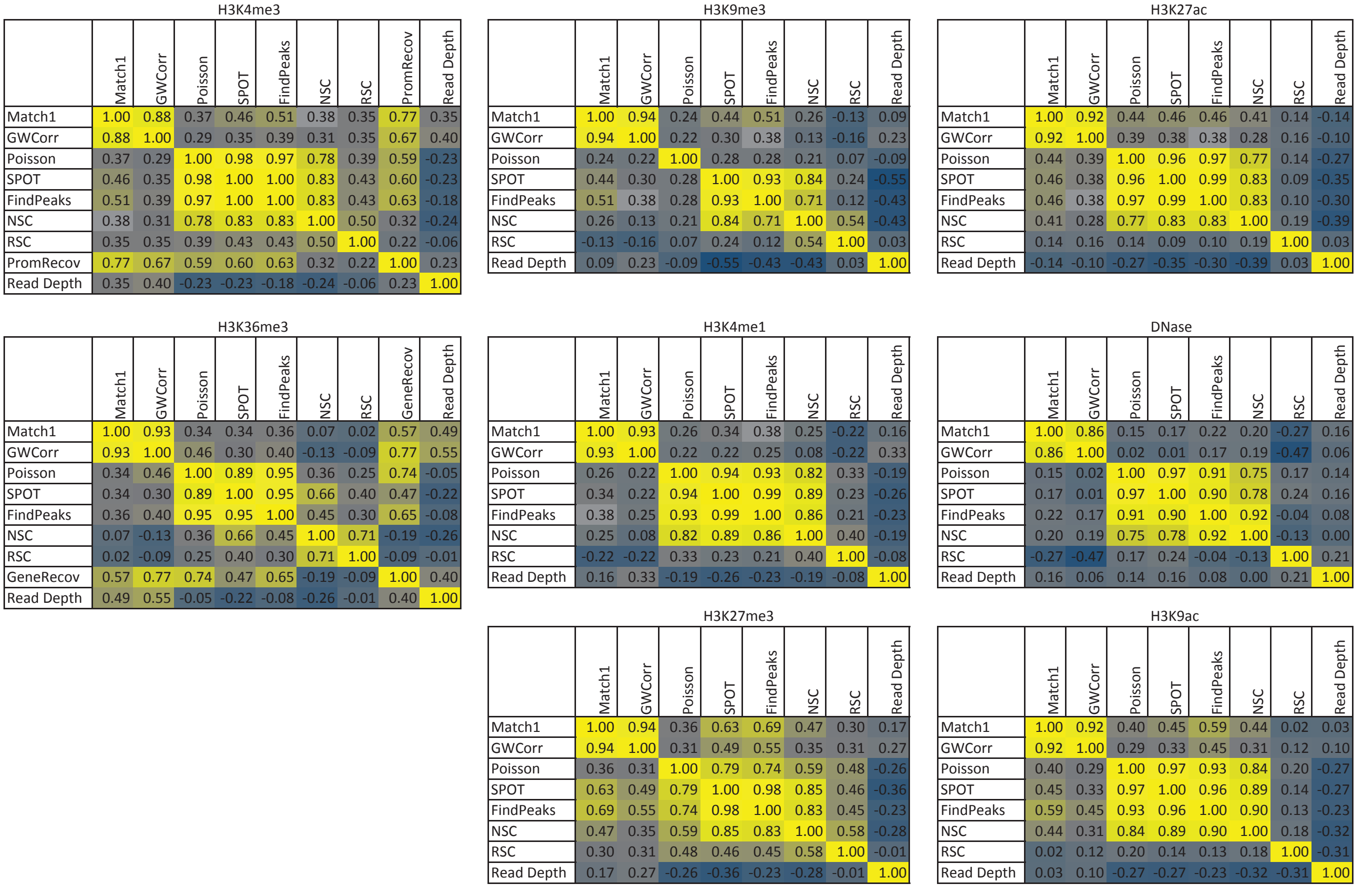


Supplementary Figure 25:
Imputed and Observed Data Correspondence with GWAS – Sample-Study Combinations.
The same as **Fig S24** except based on all combinations of studies and samples, and not just the most significant one per study. The x-axis shows the number of combinations that reached the significance level indicated on the y-axis. This is shown based on the imputed data and observed data with the actual GWAS catalog, along with the observed and imputed data based on ten randomizations of the GWAS catalog.





Supplementary Figure 26:
Imputed-Observed Data GWAS Scatter Plots.
Similar to **Fig. 4b** except for the other Tier 1 marks. The plots are showing the –log$_{10}$ p-value computed for each study-sample combination based on the observed data (x-axis) and imputed data (y-axis) for each combination that had a p-value of 10^{-3} or more significant based on either the imputed or the observed data. The diagonal line is the y=x line showing that most of the more significant study-sample combinations based on either the observed or imputed data are above it and thus have a more significant p-value based on the imputed data.



H3K4me3

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	PromRecov	Read Depth
Match1	1.00	0.88	0.37	0.46	0.51	0.38	0.35	0.77	0.35
GWCorr	0.88	1.00	0.29	0.35	0.39	0.31	0.35	0.67	0.40
Poisson	0.37	0.29	1.00	0.98	0.97	0.78	0.39	0.59	-0.23
SPOT	0.46	0.35	0.98	1.00	1.00	0.83	0.43	0.60	-0.23
FindPeaks	0.51	0.39	0.97	1.00	1.00	0.83	0.43	0.63	-0.18
NSC	0.38	0.31	0.78	0.83	0.83	1.00	0.50	0.32	-0.24
RSC	0.35	0.35	0.39	0.43	0.43	0.50	1.00	0.22	-0.06
PromRecov	0.77	0.67	0.59	0.60	0.63	0.32	0.22	1.00	0.23
Read Depth	0.35	0.40	-0.23	-0.23	-0.18	-0.24	-0.06	0.23	1.00

H3K9me3

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.94	0.24	0.44	0.51	0.26	-0.13	0.09
GWCorr	0.94	1.00	0.22	0.30	0.38	0.13	-0.16	0.23
Poisson	0.24	0.22	1.00	0.28	0.28	0.21	0.07	-0.09
SPOT	0.44	0.30	0.28	1.00	0.93	0.84	0.24	-0.55
FindPeaks	0.51	0.38	0.28	0.93	1.00	0.71	0.12	-0.43
NSC	0.26	0.13	0.21	0.84	0.71	1.00	0.54	-0.43
RSC	-0.13	-0.16	0.07	0.24	0.12	0.54	1.00	0.03
Read Depth	0.09	0.23	-0.09	-0.55	-0.43	-0.43	0.03	1.00

H3K27ac

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.92	0.44	0.46	0.46	0.41	0.14	-0.14
GWCorr	0.92	1.00	0.39	0.38	0.38	0.28	0.16	-0.10
Poisson	0.44	0.39	1.00	0.96	0.97	0.77	0.14	-0.27
SPOT	0.46	0.38	0.96	1.00	0.99	0.83	0.09	-0.35
FindPeaks	0.46	0.38	0.97	0.99	1.00	0.83	0.10	-0.30
NSC	0.41	0.28	0.77	0.83	0.83	1.00	0.19	-0.39
RSC	0.14	0.16	0.14	0.09	0.10	0.19	1.00	0.03
Read Depth	-0.14	-0.10	-0.27	-0.35	-0.30	-0.39	0.03	1.00

H3K36me3

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	GeneRecov	Read Depth
Match1	1.00	0.93	0.34	0.34	0.36	0.07	0.02	0.57	0.49
GWCorr	0.93	1.00	0.46	0.30	0.40	-0.13	-0.09	0.77	0.55
Poisson	0.34	0.46	1.00	0.89	0.95	0.36	0.25	0.74	-0.05
SPOT	0.34	0.30	0.89	1.00	0.95	0.66	0.40	0.47	-0.22
FindPeaks	0.36	0.40	0.95	0.95	1.00	0.45	0.30	0.65	-0.08
NSC	0.07	-0.13	0.36	0.66	0.45	1.00	0.71	-0.19	-0.26
RSC	0.02	-0.09	0.25	0.40	0.30	0.71	1.00	-0.09	-0.01
GeneRecov	0.57	0.77	0.74	0.47	0.65	-0.19	-0.09	1.00	0.40
Read Depth	0.49	0.55	-0.05	-0.22	-0.08	-0.26	-0.01	0.40	1.00

H3K4me1

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.93	0.26	0.34	0.38	0.25	-0.22	0.16
GWCorr	0.93	1.00	0.22	0.22	0.25	0.08	-0.22	0.33
Poisson	0.26	0.22	1.00	0.94	0.93	0.82	0.33	-0.19
SPOT	0.34	0.22	0.94	1.00	0.99	0.89	0.23	-0.26
FindPeaks	0.38	0.25	0.93	0.99	1.00	0.86	0.21	-0.23
NSC	0.25	0.08	0.82	0.89	0.86	1.00	0.40	-0.19
RSC	-0.22	-0.22	0.33	0.23	0.21	0.40	1.00	-0.08
Read Depth	0.16	0.33	-0.19	-0.26	-0.23	-0.19	-0.08	1.00

DNase

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.86	0.15	0.17	0.22	0.20	-0.27	0.16
GWCorr	0.86	1.00	0.02	0.01	0.17	0.19	-0.47	0.06
Poisson	0.15	0.02	1.00	0.97	0.91	0.75	0.17	0.14
SPOT	0.17	0.01	0.97	1.00	0.90	0.78	0.24	0.16
FindPeaks	0.22	0.17	0.91	0.90	1.00	0.92	-0.04	0.08
NSC	0.20	0.19	0.75	0.78	0.92	1.00	-0.13	0.00
RSC	-0.27	-0.47	0.17	0.24	-0.04	-0.13	1.00	0.21
Read Depth	0.16	0.06	0.14	0.16	0.08	0.00	0.21	1.00

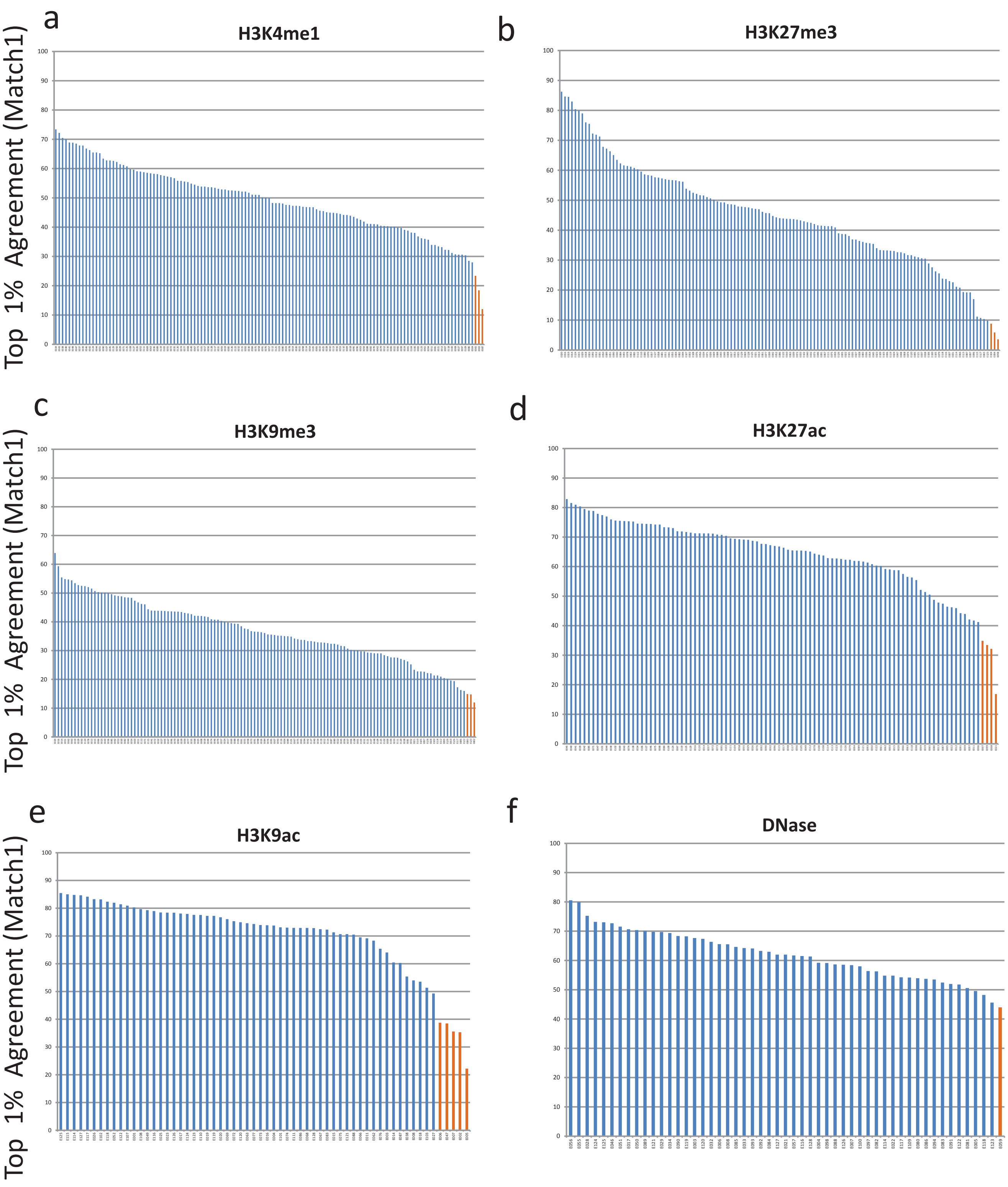
H3K27me3

	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.94	0.36	0.63	0.69	0.47	0.30	0.17
GWCorr	0.94	1.00	0.31	0.49	0.55	0.35	0.31	0.27
Poisson	0.36	0.31	1.00	0.79	0.74	0.59	0.48	-0.26
SPOT	0.63	0.49	0.79	1.00	0.98	0.85	0.46	-0.36
FindPeaks	0.69	0.55	0.74	0.98	1.00	0.83	0.45	-0.23
NSC	0.47	0.35	0.59	0.85	0.83	1.00	0.58	-0.28
RSC	0.30	0.31	0.48	0.46	0.45	0.58	1.00	-0.01
Read Depth	0.17	0.27	-0.26	-0.36	-0.23	-0.28	-0.01	1.00

H3K9ac

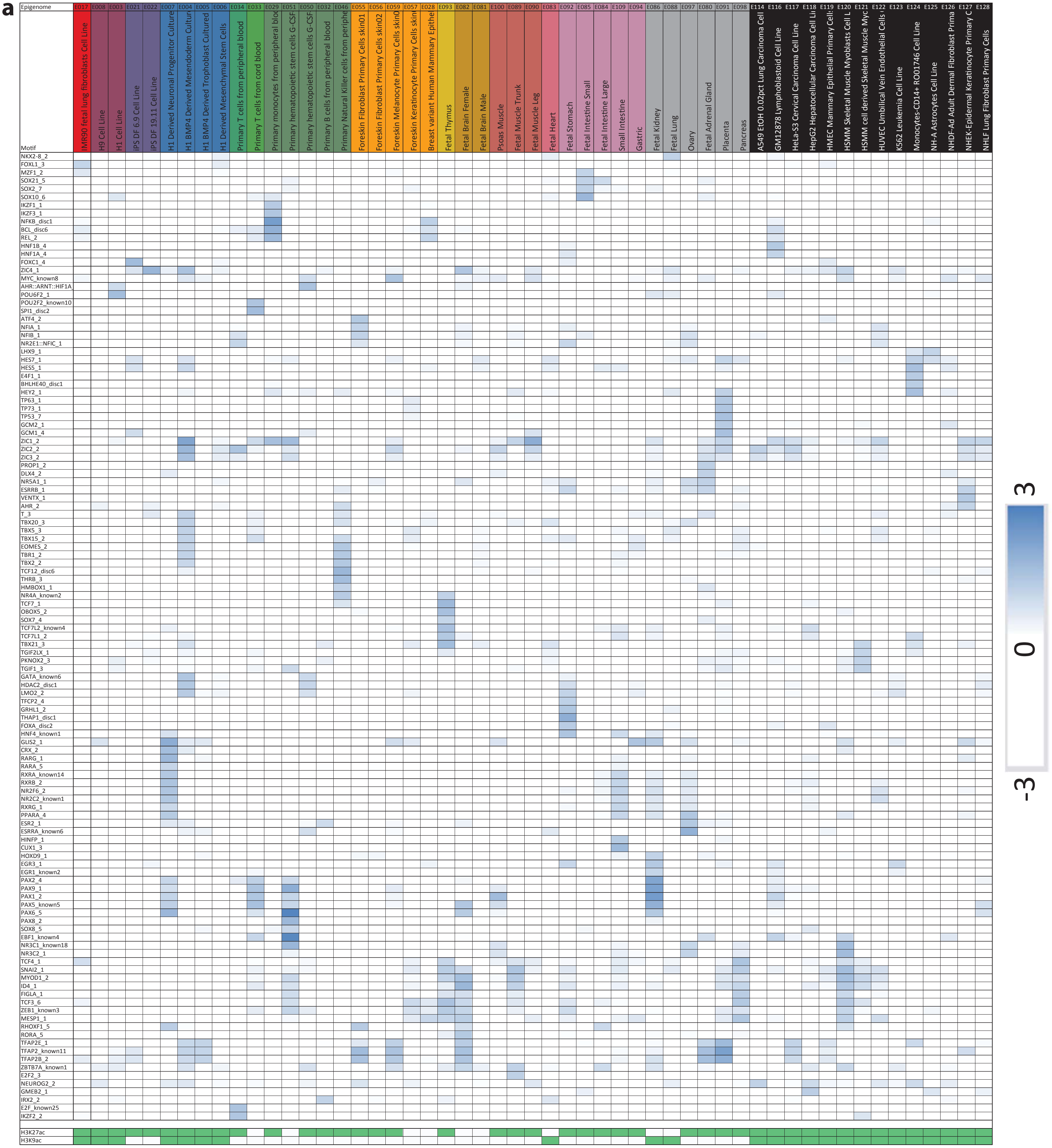
	Match1	GWCorr	Poisson	SPOT	FindPeaks	NSC	RSC	Read Depth
Match1	1.00	0.92	0.40	0.45	0.59	0.44	0.02	0.03
GWCorr	0.92	1.00	0.29	0.33	0.45	0.31	0.12	0.10
Poisson	0.40	0.29	1.00	0.97	0.93	0.84	0.20	-0.27
SPOT	0.45	0.33	0.97	1.00	0.96	0.89	0.14	-0.27
FindPeaks	0.59	0.45	0.93	0.96	1.00	0.90	0.13	-0.23
NSC	0.44	0.31	0.84	0.89	0.90	1.00	0.18	-0.32
RSC	0.02	0.12	0.20	0.14	0.13	0.18	1.00	-0.31
Read Depth	0.03	0.10	-0.27	-0.27	-0.23	-0.32	-0.31	1.00

Supplementary Figure 27: Quality Control (QC) Metric Correlations.
These tables show for each of the Tier-1 marks the Pearson correlation between all the general quality control measures evaluated. The table illustrates the two imputation based quality measures (Match 1, GWCorr) consistently correlate highly with each other, and to a lesser extent with the five non-imputation based general QC measures evaluated (SPOT, Poisson, FindPeaks, NSC, and RSC) along with read depth suggesting the imputation metrics could potentially provide additional unique information for quality control. For H3K36me3 and H3K4me3 the tables also contain correlations with a metric that leveraged relevant gene annotation information, AUC at a 5% false positive rate for recovering gene bodies (GeneRecov) and +/-2KB TSS regions (PromRecov) respectively, for which the imputation QC measures are among the best correlated.



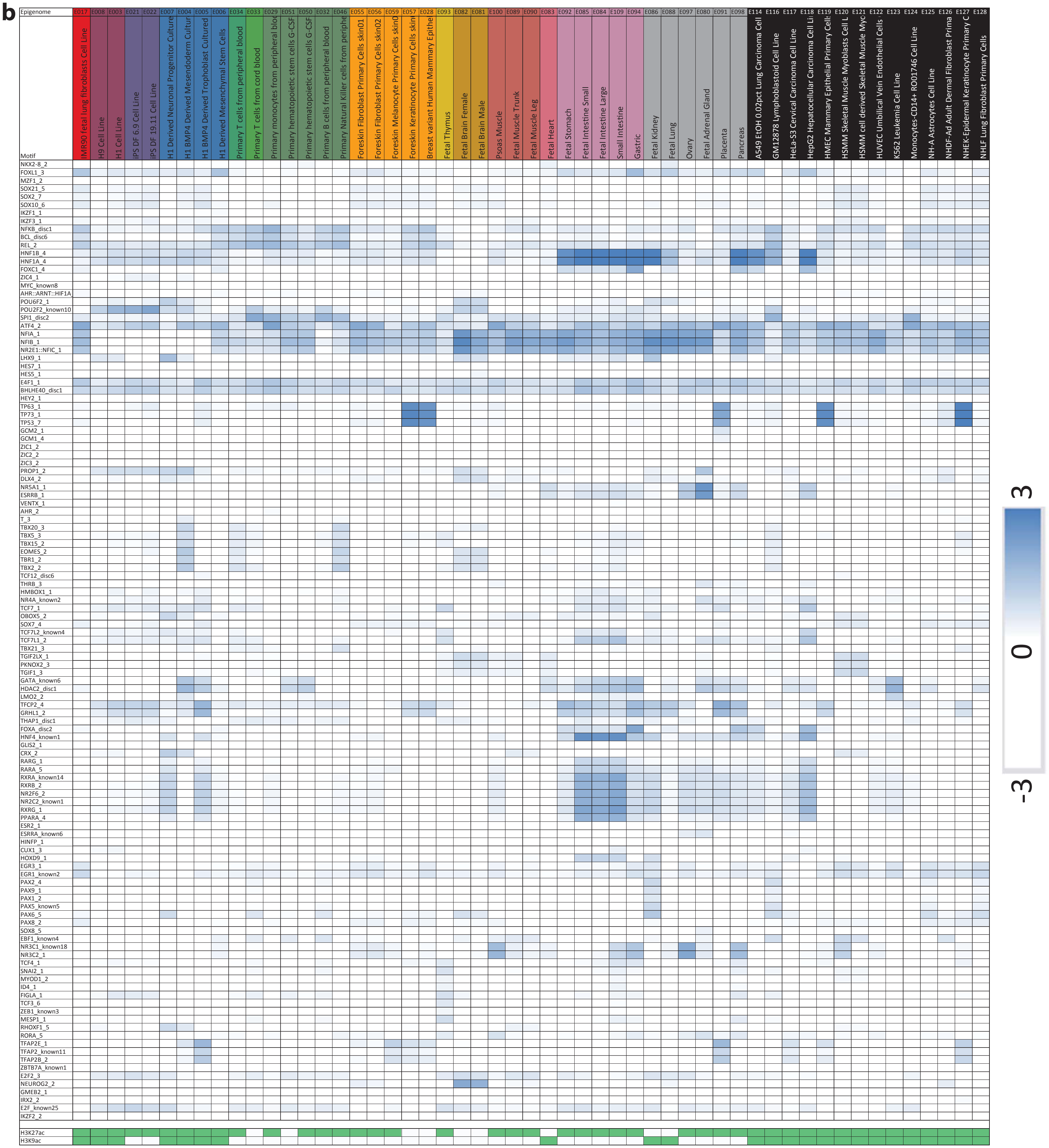
Supplementary Figure 28: Distribution of Imputation Top 1% Agreement Scores.

Similar to **Fig. 5b** except the figure shows the distribution across samples of top 1% signal location agreement percentages between the observed and corresponding imputed data for **(a)** H3K4me1 **(b)** H3K27me3 **(c)** H3K9me3 **(d)** H3K27ac **(e)** H3K9ac and **(f)** DNase. Shown in brown are observed datasets with an imputation agreement score more than two standard deviations below average for the mark. See also **Table S3** for the agreement scores and samples.

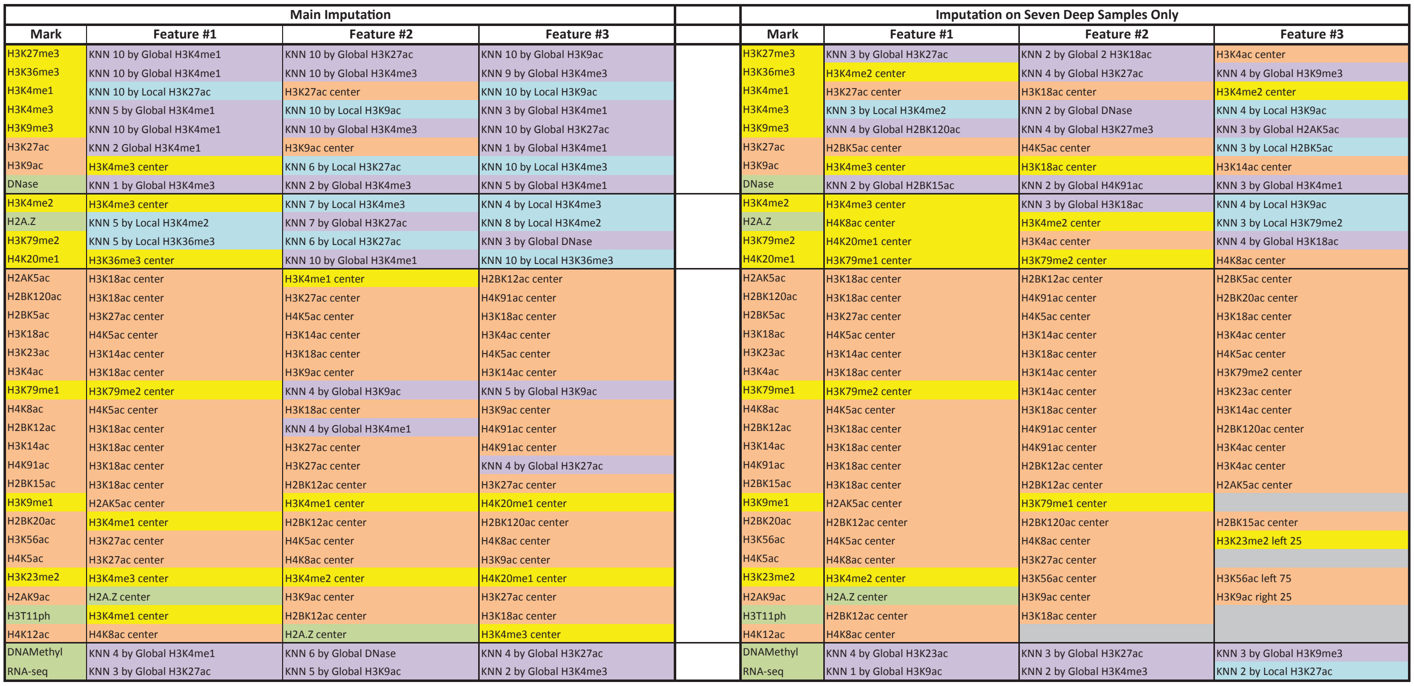


Supplementary Figure 29: Motif Enrichments in Locations of Unexpected DNase signal.

(a) The heatmap shows sequence motif enrichments occurring in locations with unexpected DNase signal, which was defined here as places where the observed DNase signal was above 5, but the imputed signal was below 1. Motif enrichments are shown in $\log_2$ relative to control as computed using a previous described motif enrichment program[57]. The background locations for computing the motif enrichments were locations that had an observed signal above 5. The rows correspond to different motifs, and columns different samples with observed DNase data available. Only motifs which had an enrichment value of at least 1 in at least one sample are shown. If multiple motifs corresponding to the same factor were available only the one with the maximum enrichment for any sample is shown. Along the bottom row is indicated which acetylations marks were available in the sample which could affect the expected signal. **(b)** The same heatmap as in **a**, but showing the motif enrichment for locations that had an observed signal above 5, used in the background for computing the enrichments in **a**, relative to a genomewide background.

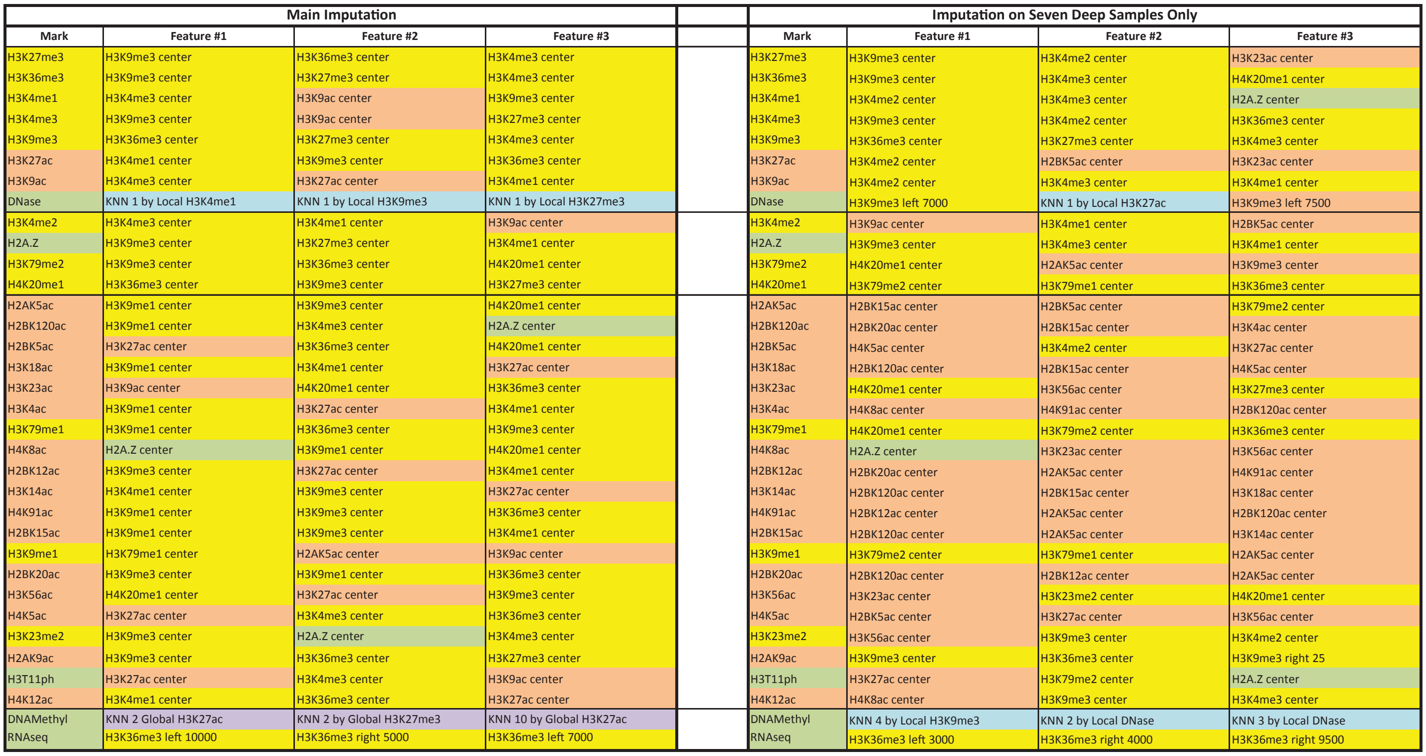

b
Epigenome
Motif
E017
E008
E003
E021
E022
E007
E004
E005
E006
E034
E033
E029
E051
E050
E032
E046
E055
E056
E059
E057
E028
E093
E082
E081
E100
E089
E090
E083
E092
E085
E084
E109
E094
E086
E088
E097
E080
E091
E098
E114
E116
E117
E118
E119
E120
E121
E122
E123
E124
E125
E126
E127
E128
IMR90 fetal lung fibroblasts Cell Line
H9 Cell Line
H1 Cell Line
iPS DF 6.9 Cell Line
iPS DF 19.11 Cell Line
H1 Derived Neuronal Progenitor Culture
H1 BMP4 Derived Mesendoderm Cultur
H1 BMP4 Derived Trophoblast Cultured
H1 Derived Mesenchymal Stem Cells
Primary T cells from peripheral blood
Primary T cells from cord blood
Primary monocytes from peripheral bloc
Primary hematopoietic stem cells G-CSF
Primary hematopoietic stem cells G-CSF
Primary B cells from peripheral blood
Primary Natural Killer cells from periphe
Foreskin Fibroblast Primary Cells skin01
Foreskin Fibroblast Primary Cells skin02
Foreskin Melanocyte Primary Cells skin0
Foreskin Keratinocyte Primary Cells skin
Breast variant Human Mammary Epithe
Fetal Thymus
Fetal Brain Female
Fetal Brain Male
Psoas Muscle
Fetal Muscle Trunk
Fetal Muscle Leg
Fetal Heart
Fetal Stomach
Fetal Intestine Small
Fetal Intestine Large
Small Intestine
Gastric
Fetal Kidney
Fetal Lung
Ovary
Fetal Adrenal Gland
Placenta
Pancreas
A549 EtOH 0.02pct Lung Carcinoma Cell
GM12878 Lymphoblastoid Cell Line
HeLa-S3 Cervical Carcinoma Cell Line
HepG2 Hepatocellular Carcinoma Cell Li
HMEC Mammary Epithelial Primary Cell
HSMM Skeletal Muscle Myoblasts Cell L
HSMM cell derived Skeletal Muscle Myo
HUVEC Umbilical Vein Endothelial Cells
K562 Leukemia Cell Line
Monocytes-CD14+ RO01746 Cell Line
NH-A Astrocytes Cell Line
NHDF-Ad Adult Dermal Fibroblast Prima
NHEK-Epidermal Keratinocyte Primary C
NHLF Lung Fibroblast Primary Cells
NKX2-8_2
FOXL1_3
MZF1_2
SOX21_5
SOX2_7
SOX10_6
IKZF1_1
IKZF3_1
NFKB_disc1
BCL_disc6
REL_2
HNF1B_4
HNF1A_4
FOXC1_4
ZIC4_1
MYC_known8
AHR::ARNT::HIF1A_
POU6F2_1
POU2F2_known10
SPI1_disc2
ATF4_2
NFIA_1
NFIB_1
NR2E1::NFIC_1
LHX9_1
HES7_1
HES5_1
E4F1_1
BHLHE40_disc1
HEY2_1
TP63_1
TP73_1
TP53_7
GCM2_1
GCM1_4
ZIC1_2
ZIC2_2
ZIC3_2
PROP1_2
DLX4_2
NR5A1_1
ESRRB_1
VENTX_1
AHR_2
T_3
TBX20_3
TBX5_3
TBX15_2
EOMES_2
TBR1_2
TBX2_2
TCF12_disc6
THRB_3
HMBOX1_1
NR4A_known2
TCF7_1
OBOX5_2
SOX7_4
TCF7L2_known4
TCF7L1_2
TBX21_3
TGIF2LX_1
PKNOX2_3
TGIF1_3
GATA_known6
HDAC2_disc1
LMO2_2
TFCP2_4
GRHL1_2
THAP1_disc1
FOXA_disc2
HNF4_known1
GLIS2_1
CRX_2
RARG_1
RARA_5
RXRA_known14
RXRB_2
NR2F6_2
NR2C2_known1
RXRG_1
PPARA_4
ESR2_1
ESRRA_known6
HINFP_1
CUX1_3
HOXD9_1
EGR3_1
EGR1_known2
PAX2_4
PAX9_1
PAX1_2
PAX5_known5
PAX6_5
PAX8_2
SOX8_5
EBF1_known4
NR3C1_known18
NR3C2_1
TCF4_1
SNAI2_1
MYOD1_2
ID4_1
FIGLA_1
TCF3_6
ZEB1_known3
MESP1_1
RHOXF1_5
RORA_5
TFAP2E_1
TFAP2_known11
TFAP2B_2
ZBTB7A_known1
E2F2_3
NEUROG2_2
GMEB2_1
IRX2_2
E2F_known25
IKZF2_2
H3K27ac
H3K9ac
3
0
-3

a Root position



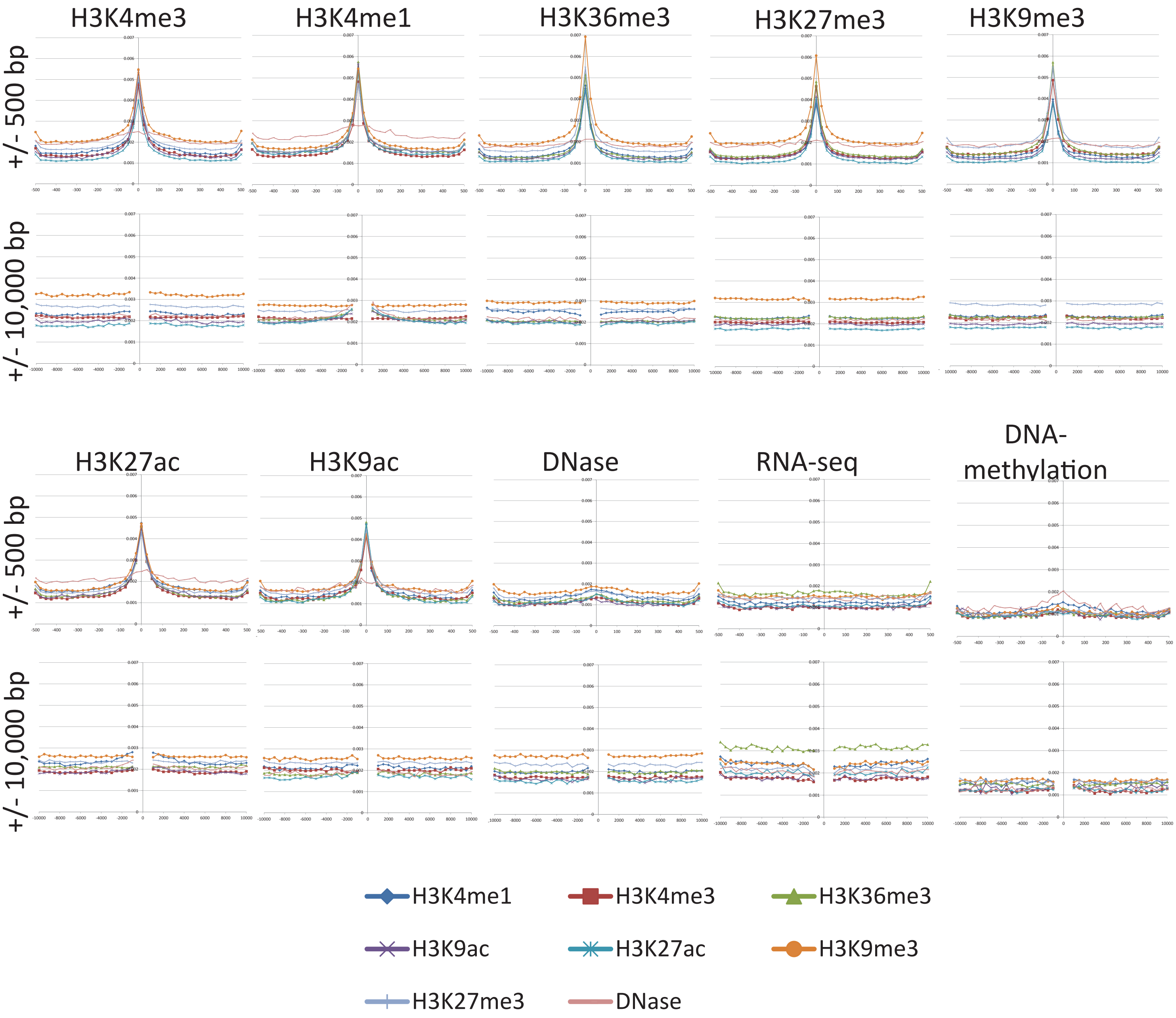
Main Imputation					Imputation on Seven Deep Samples Only			
Mark	**Feature #1**	**Feature #2**	**Feature #3**		**Mark**	**Feature #1**	**Feature #2**	**Feature #3**
H3K27me3	KNN 10 by Global H3K4me1	KNN 10 by Global H3K27ac	KNN 10 by Global H3K9ac		H3K27me3	KNN 3 by Global H3K27ac	KNN 2 by Global 2 H3K18ac	H3K4ac center
H3K36me3	KNN 10 by Global H3K4me1	KNN 10 by Global H3K4me3	KNN 9 by Global H3K4me3		H3K36me3	H3K4me2 center	KNN 4 by Global H3K27ac	KNN 4 by Global H3K9me3
H3K4me1	KNN 10 by Local H3K27ac	H3K27ac center	KNN 10 by Local H3K9ac		H3K4me1	H3K27ac center	H3K18ac center	H3K4me2 center
H3K4me3	KNN 5 by Global H3K4me1	KNN 10 by Local H3K9ac	KNN 3 by Global H3K4me1		H3K4me3	KNN 3 by Local H3K4me2	KNN 2 by Global DNase	KNN 4 by Local H3K9ac
H3K9me3	KNN 10 by Global H3K4me1	KNN 10 by Global H3K4me3	KNN 10 by Global H3K27ac		H3K9me3	KNN 4 by Global H2BK120ac	KNN 4 by Global H3K27me3	KNN 3 by Global H2AK5ac
H3K27ac	KNN 2 Global H3K4me1	H3K9ac center	KNN 1 by Global H3K4me1		H3K27ac	H2BK5ac center	H4K5ac center	KNN 3 by Local H2BK5ac
H3K9ac	H3K4me3 center	KNN 6 by Local H3K27ac	KNN 10 by Local H3K4me3		H3K9ac	H3K4me3 center	H3K18ac center	H3K14ac center
DNase	KNN 1 by Global H3K4me3	KNN 2 by Global H3K4me3	KNN 5 by Global H3K4me1		DNase	KNN 2 by Global H2BK15ac	KNN 2 by Global H4K91ac	KNN 3 by Global H3K4me1
H3K4me2	H3K4me3 center	KNN 7 by Local H3K4me3	KNN 4 by Local H3K4me3		H3K4me2	H3K4me3 center	KNN 3 by Global H3K18ac	KNN 4 by Local H3K9ac
H2A.Z	KNN 5 by Local H3K4me2	KNN 7 by Global H3K27ac	KNN 8 by Local H3K4me2		H2A.Z	H4K8ac center	H3K4me2 center	KNN 3 by Local H3K79me2
H3K79me2	KNN 5 by Local H3K36me3	KNN 6 by Local H3K27ac	KNN 3 by Global DNase		H3K79me2	H4K20me1 center	H3K4ac center	KNN 4 by Global H3K18ac
H4K20me1	H3K36me3 center	KNN 10 by Global H3K4me1	KNN 10 by Local H3K36me3		H4K20me1	H3K79me1 center	H3K79me2 center	H4K8ac center
H2AK5ac	H3K18ac center	H3K4me1 center	H2BK12ac center		H2AK5ac	H3K18ac center	H2BK12ac center	H2BK5ac center
H2BK120ac	H3K18ac center	H3K27ac center	H4K91ac center		H2BK120ac	H3K18ac center	H4K91ac center	H2BK20ac center
H2BK5ac	H3K27ac center	H4K5ac center	H3K18ac center		H2BK5ac	H3K27ac center	H4K5ac center	H3K18ac center
H3K18ac	H4K5ac center	H3K14ac center	H3K4ac center		H3K18ac	H4K5ac center	H3K14ac center	H3K4ac center
H3K23ac	H3K14ac center	H3K18ac center	H4K5ac center		H3K23ac	H3K14ac center	H3K18ac center	H4K5ac center
H3K4ac	H3K18ac center	H3K9ac center	H3K14ac center		H3K4ac	H3K18ac center	H3K14ac center	H3K79me2 center
H3K79me1	H3K79me2 center	KNN 4 by Global H3K9ac	KNN 5 by Global H3K9ac		H3K79me1	H3K79me2 center	H3K14ac center	H3K23ac center
H4K8ac	H4K5ac center	H3K18ac center	H3K9ac center		H4K8ac	H4K5ac center	H3K18ac center	H3K14ac center
H2BK12ac	H3K18ac center	KNN 4 by Global H3K4me1	H4K91ac center		H2BK12ac	H3K18ac center	H4K91ac center	H2BK120ac center
H3K14ac	H3K18ac center	H3K27ac center	H4K91ac center		H3K14ac	H3K18ac center	H4K91ac center	H3K4ac center
H4K91ac	H3K18ac center	H3K27ac center	KNN 4 by Global H3K27ac		H4K91ac	H3K18ac center	H2BK12ac center	H3K4ac center
H2BK15ac	H3K18ac center	H2BK12ac center	H3K27ac center		H2BK15ac	H3K18ac center	H2BK12ac center	H2AK5ac center
H3K9me1	H2AK5ac center	H3K4me1 center	H4K20me1 center		H3K9me1	H2AK5ac center	H3K79me1 center	
H2BK20ac	H3K4me1 center	H2BK12ac center	H2BK120ac center		H2BK20ac	H2BK12ac center	H2BK120ac center	H2BK15ac center
H3K56ac	H3K27ac center	H4K5ac center	H4K8ac center		H3K56ac	H4K5ac center	H4K8ac center	H3K23me2 left 25
H4K5ac	H3K27ac center	H4K8ac center	H3K9ac center		H4K5ac	H4K8ac center	H3K27ac center	
H3K23me2	H3K4me3 center	H3K4me2 center	H4K20me1 center		H3K23me2	H3K4me2 center	H3K56ac center	H3K56ac left 75
H2AK9ac	H2A.Z center	H3K9ac center	H3K27ac center		H2AK9ac	H2A.Z center	H3K9ac center	H3K9ac right 25
H3T11ph	H3K4me1 center	H2BK12ac center	H3K18ac center		H3T11ph	H2BK12ac center	H3K18ac center	
H4K12ac	H4K8ac center	H2A.Z center	H3K4me3 center		H4K12ac	H4K8ac center		
DNAMethyl	KNN 4 by Global H3K4me1	KNN 6 by Global DNase	KNN 4 by Global H3K27ac		DNAMethyl	KNN 4 by Global H3K23ac	KNN 3 by Global H3K27ac	KNN 3 by Global H3K9me3
RNA-seq	KNN 3 by Global H3K27ac	KNN 5 by Global H3K9ac	KNN 2 by Global H3K4me3		RNA-seq	KNN 1 by Global H3K9ac	KNN 2 by Global H3K4me3	KNN 2 by Local H3K27ac

b All positions

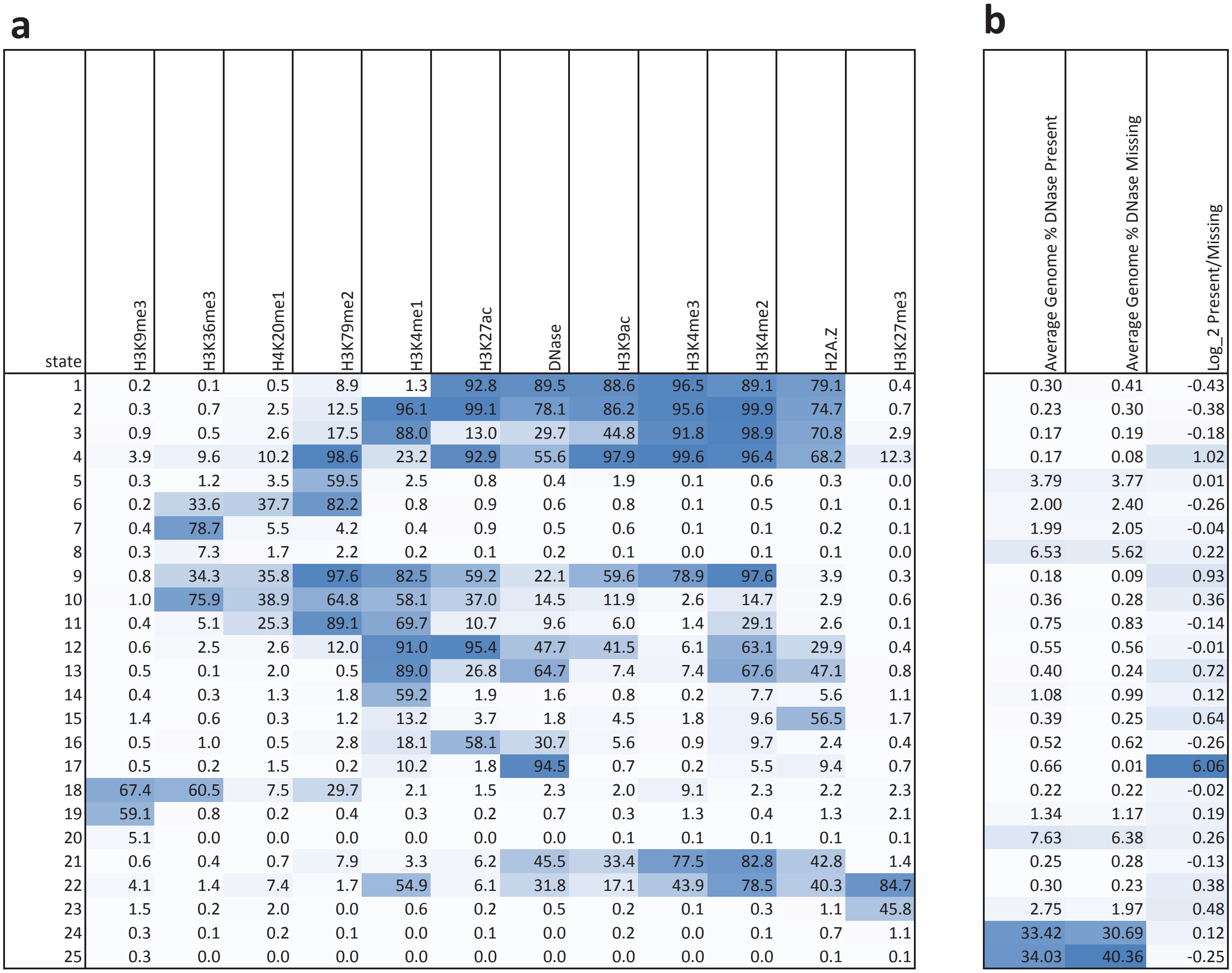


Main Imputation					Imputation on Seven Deep Samples Only			
Mark	**Feature #1**	**Feature #2**	**Feature #3**		**Mark**	**Feature #1**	**Feature #2**	**Feature #3**
H3K27me3	H3K9me3 center	H3K36me3 center	H3K4me3 center		H3K27me3	H3K9me3 center	H3K4me2 center	H3K23ac center
H3K36me3	H3K9me3 center	H3K27me3 center	H3K4me3 center		H3K36me3	H3K9me3 center	H3K4me3 center	H4K20me1 center
H3K4me1	H3K4me3 center	H3K9ac center	H3K9me3 center		H3K4me1	H3K4me2 center	H3K4me3 center	H2A.Z center
H3K4me3	H3K9me3 center	H3K9ac center	H3K27me3 center		H3K4me3	H3K9me3 center	H3K4me2 center	H3K36me3 center
H3K9me3	H3K36me3 center	H3K27me3 center	H3K4me3 center		H3K9me3	H3K36me3 center	H3K27me3 center	H3K4me3 center
H3K27ac	H3K4me1 center	H3K9me3 center	H3K36me3 center		H3K27ac	H3K4me2 center	H2BK5ac center	H3K23ac center
H3K9ac	H3K4me3 center	H3K27ac center	H3K4me1 center		H3K9ac	H3K4me2 center	H3K4me3 center	H3K4me1 center
DNase	KNN 1 by Local H3K4me1	KNN 1 by Local H3K9me3	KNN 1 by Local H3K27me3		DNase	H3K9me3 left 7000	KNN 1 by Local H3K27ac	H3K9me3 left 7500
H3K4me2	H3K4me3 center	H3K4me1 center	H3K9ac center		H3K4me2	H3K9ac center	H3K4me1 center	H2BK5ac center
H2A.Z	H3K9me3 center	H3K27me3 center	H3K4me1 center		H2A.Z	H3K9me3 center	H3K4me3 center	H3K4me1 center
H3K79me2	H3K9me3 center	H3K36me3 center	H4K20me1 center		H3K79me2	H4K20me1 center	H2AK5ac center	H3K9me3 center
H4K20me1	H3K36me3 center	H3K9me3 center	H3K27me3 center		H4K20me1	H3K79me2 center	H3K79me1 center	H3K36me3 center
H2AK5ac	H3K9me1 center	H3K9me3 center	H4K20me1 center		H2AK5ac	H2BK15ac center	H2BK5ac center	H3K79me2 center
H2BK120ac	H3K9me1 center	H3K4me3 center	H2A.Z center		H2BK120ac	H2BK20ac center	H2BK15ac center	H3K4ac center
H2BK5ac	H3K27ac center	H3K36me3 center	H4K20me1 center		H2BK5ac	H4K5ac center	H3K4me2 center	H3K27ac center
H3K18ac	H3K9me1 center	H3K4me1 center	H3K27ac center		H3K18ac	H2BK120ac center	H2BK15ac center	H4K5ac center
H3K23ac	H3K9ac center	H4K20me1 center	H3K36me3 center		H3K23ac	H4K20me1 center	H3K56ac center	H3K27me3 center
H3K4ac	H3K9me1 center	H3K27ac center	H3K4me1 center		H3K4ac	H4K8ac center	H4K91ac center	H2BK120ac center
H3K79me1	H3K9me1 center	H3K36me3 center	H3K9me3 center		H3K79me1	H4K20me1 center	H3K79me2 center	H3K36me3 center
H4K8ac	H2A.Z center	H3K9me1 center	H4K20me1 center		H4K8ac	H2A.Z center	H3K23ac center	H3K56ac center
H2BK12ac	H3K9me3 center	H3K27ac center	H3K4me1 center		H2BK12ac	H2BK20ac center	H2AK5ac center	H4K91ac center
H3K14ac	H3K4me1 center	H3K9me3 center	H3K27ac center		H3K14ac	H2BK120ac center	H2BK15ac center	H3K18ac center
H4K91ac	H3K9me1 center	H3K9me3 center	H3K36me3 center		H4K91ac	H2BK12ac center	H2AK5ac center	H2BK120ac center
H2BK15ac	H3K9me1 center	H3K9me3 center	H3K4me1 center		H2BK15ac	H2BK120ac center	H2AK5ac center	H3K14ac center
H3K9me1	H3K79me1 center	H2AK5ac center	H3K9ac center		H3K9me1	H3K79me2 center	H3K79me1 center	H2AK5ac center
H2BK20ac	H3K9me3 center	H3K9me1 center	H3K36me3 center		H2BK20ac	H2BK120ac center	H2BK12ac center	H2AK5ac center
H3K56ac	H4K20me1 center	H3K27ac center	H3K9me3 center		H3K56ac	H3K23ac center	H3K23me2 center	H4K20me1 center
H4K5ac	H3K27ac center	H3K4me3 center	H3K36me3 center		H4K5ac	H2BK5ac center	H3K27ac center	H3K56ac center
H3K23me2	H3K9me3 center	H2A.Z center	H3K4me3 center		H3K23me2	H3K56ac center	H3K9me3 center	H3K4me2 center
H2AK9ac	H3K9me3 center	H3K36me3 center	H3K27me3 center		H2AK9ac	H3K9me3 center	H3K36me3 center	H3K9me3 right 25
H3T11ph	H3K27ac center	H3K4me3 center	H3K9ac center		H3T11ph	H3K27ac center	H3K79me2 center	H2A.Z center
H4K12ac	H3K4me1 center	H3K36me3 center	H3K27ac center		H4K12ac	H4K8ac center	H3K9me3 center	H3K4me3 center
DNAMethyl	KNN 2 Global H3K27ac	KNN 2 by Global H3K27me3	KNN 10 by Global H3K27ac		DNAMethyl	KNN 4 by Local H3K9me3	KNN 2 by Local DNase	KNN 3 by Local DNase
RNAseq	H3K36me3 left 10000	H3K36me3 right 5000	H3K36me3 left 7000		RNAseq	H3K36me3 left 3000	H3K36me3 right 4000	H3K36me3 right 9500

Supplementary Figure 30: Top Used Features.
(a) (left) The table shows for imputing each mark in the main imputation the top three ranking features in terms of proportional usage as the root feature in the regression trees. Proportional usage in the root was determined as the fraction of times the feature was selected as the root feature for any regression tree built for the mark out of the number of times the feature was eligible for selection. The marks are color coded based on acetylation, methylation, or other. The features are color coded based on if it is the same mark in a sample determined based on local distance, based on global distance, or an acetylation, methylation, or other in the same sample. (right) The same as on the left but for the imputation restricted to the seven samples with deep mark coverage. **(b)** (left) The table shows for imputing each mark in the main imputation the top three ranking features in terms of proportional usage as a feature anywhere in the regression trees. Proportional usage here was determined as the fraction of times the feature was used at any split node in the tree out of the total number of split nodes for which the feature was eligible to be selected. The color coding was the same as in part **a**. (right) The same as on the left but for the imputation restricted to the seven samples with deep mark coverage.



Supplementary Figure 31: Feature Usage Relative to Position.
The figure displays for predicting each of the Tier-1 marks, RNA-seq, and DNA-methylation a plot showing the relative usage of Tier-1 mark signal features at each 25bp position within 500bp of the target position (x-axis) based on the proportion of times it was selected as a split feature in a node anywhere in the tree out of the number of times it was eligible for selection (y-axis), and then directly below that the same plot, but now showing features from 500bp up until 10kb spaced at 500 base pair intervals.



a

state	H3K9me3	H3K36me3	H4K20me1	H3K79me2	H3K4me1	H3K27ac	DNase	H3K9ac	H3K4me3	H3K4me2	H2A.Z	H3K27me3
1	0.2	0.1	0.5	8.9	1.3	92.8	89.5	88.6	96.5	89.1	79.1	0.4
2	0.3	0.7	2.5	12.5	96.1	99.1	78.1	86.2	95.6	99.9	74.7	0.7
3	0.9	0.5	2.6	17.5	88.0	13.0	29.7	44.8	91.8	98.9	70.8	2.9
4	3.9	9.6	10.2	98.6	23.2	92.9	55.6	97.9	99.6	96.4	68.2	12.3
5	0.3	1.2	3.5	59.5	2.5	0.8	0.4	1.9	0.1	0.6	0.3	0.0
6	0.2	33.6	37.7	82.2	0.8	0.9	0.6	0.8	0.1	0.5	0.1	0.1
7	0.4	78.7	5.5	4.2	0.4	0.9	0.5	0.6	0.1	0.1	0.2	0.1
8	0.3	7.3	1.7	2.2	0.2	0.1	0.2	0.1	0.0	0.1	0.1	0.0
9	0.8	34.3	35.8	97.6	82.5	59.2	22.1	59.6	78.9	97.6	3.9	0.3
10	1.0	75.9	38.9	64.8	58.1	37.0	14.5	11.9	2.6	14.7	2.9	0.6
11	0.4	5.1	25.3	89.1	69.7	10.7	9.6	6.0	1.4	29.1	2.6	0.1
12	0.6	2.5	2.6	12.0	91.0	95.4	47.7	41.5	6.1	63.1	29.9	0.4
13	0.5	0.1	2.0	0.5	89.0	26.8	64.7	7.4	7.4	67.6	47.1	0.8
14	0.4	0.3	1.3	1.8	59.2	1.9	1.6	0.8	0.2	7.7	5.6	1.1
15	1.4	0.6	0.3	1.2	13.2	3.7	1.8	4.5	1.8	9.6	56.5	1.7
16	0.5	1.0	0.5	2.8	18.1	58.1	30.7	5.6	0.9	9.7	2.4	0.4
17	0.5	0.2	1.5	0.2	10.2	1.8	94.5	0.7	0.2	5.5	9.4	0.7
18	67.4	60.5	7.5	29.7	2.1	1.5	2.3	2.0	9.1	2.3	2.2	2.3
19	59.1	0.8	0.2	0.4	0.3	0.2	0.7	0.3	1.3	0.4	1.3	2.1
20	5.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1
21	0.6	0.4	0.7	7.9	3.3	6.2	45.5	33.4	77.5	82.8	42.8	1.4
22	4.1	1.4	7.4	1.7	54.9	6.1	31.8	17.1	43.9	78.5	40.3	84.7
23	1.5	0.2	2.0	0.0	0.6	0.2	0.5	0.2	0.1	0.3	1.1	45.8
24	0.3	0.1	0.2	0.1	0.0	0.1	0.0	0.2	0.0	0.1	0.7	1.1
25	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1

b

Average Genome % DNase Present	Average Genome % DNase Missing	Log_2 Present/Missing
0.30	0.41	-0.43
0.23	0.30	-0.38
0.17	0.19	-0.18
0.17	0.08	1.02
3.79	3.77	0.01
2.00	2.40	-0.26
1.99	2.05	-0.04
6.53	5.62	0.22
0.18	0.09	0.93
0.36	0.28	0.36
0.75	0.83	-0.14
0.55	0.56	-0.01
0.40	0.24	0.72
1.08	0.99	0.12
0.39	0.25	0.64
0.52	0.62	-0.26
0.66	0.01	6.06
0.22	0.22	-0.02
1.34	1.17	0.19
7.63	6.38	0.26
0.25	0.28	-0.13
0.30	0.23	0.38
2.75	1.97	0.48
33.42	30.69	0.12
34.03	40.36	-0.25

Supplementary Figure 32: Chromatin State Model with 12-Marks Using Observed Data.
(a) The emission parameters for a chromatin state model learned directly on the Tier-1 and 2 observed data across 127 samples after applying the ChromHMM default binarization and treating as missing cases in which a mark was not available in a sample. **(b)** The first columns show the average percent of the genome assigned to each state for samples where DNase is present, the next column when DNase is absent, and the last column shows the log base two of the ratio between these two columns. This demonstrates that the percentage of the genome assigned to a state associated with DNase (State 17) is highly dependent on whether DNase data was available in the sample.

a

25-state model using 12-marks imputed across 127-samples

	H3K9me3	H3K36me3	H4K20me1	H3K79me2	H3K4me1	H3K27ac	DNase	H3K9ac	H3K4me3	H3K4me2	H2A.Z	H3K27me3	State %
1_TssA	0.4	0.1	0.0	5.0	0.4	89.3	92.3	96.5	99.9	99.1	86.4	3.5	0.18
2_PromU	0.6	0.0	0.5	6.2	99.3	91.5	82.4	96.1	99.6	100.0	95.6	23.4	0.40
3_PromD1	0.2	1.7	7.0	98.4	60.8	99.8	92.2	100.0	100.0	100.0	93.2	6.3	0.41
4_PromD2	0.9	7.3	20.3	94.2	87.5	52.3	8.0	55.1	86.7	98.0	7.2	5.2	0.18
5_Tx5'	0.4	1.1	17.0	76.2	0.3	0.2	0.7	0.0	0.0	0.1	0.0	0.1	2.21
6_Tx	0.7	94.2	45.7	80.9	7.2	1.0	1.0	0.0	0.0	0.4	0.0	0.1	0.69
7_Tx3'	0.1	85.8	1.3	0.8	0.1	0.0	0.5	0.0	0.0	0.0	0.0	0.0	3.46
8_TxWk	0.0	2.4	0.1	1.3	0.1	0.0	0.5	0.0	0.0	0.0	0.0	0.0	5.95
9_TxReg	0.2	27.5	60.2	98.1	98.3	99.9	72.1	92.8	74.0	99.6	6.1	1.3	0.31
10_TxEnh5'	0.2	25.9	49.5	96.2	94.1	94.6	25.7	5.8	2.0	41.6	0.5	0.2	0.38
11_TxEnh3'	0.4	89.6	14.8	11.0	74.3	50.0	20.3	2.9	1.4	11.8	0.5	0.5	0.21
12_TxEnhW	0.1	9.3	48.3	95.5	76.8	3.2	6.6	0.0	0.4	18.9	0.1	0.8	0.51
13_EnhA1	0.2	4.0	1.3	5.9	99.3	99.9	83.7	95.7	38.3	95.7	43.3	0.4	0.24
14_EnhA2	0.2	0.5	0.8	2.6	97.4	97.2	59.1	7.5	9.6	96.8	29.0	0.6	0.33
15_EnhAF	0.3	0.4	0.5	2.1	97.7	94.5	31.0	3.7	1.2	2.3	6.3	0.7	0.49
16_EnhW1	0.1	0.0	0.2	0.5	91.2	16.8	39.1	3.4	15.2	73.3	46.4	0.9	0.27
17_EnhW2	0.1	0.2	0.5	1.0	75.9	0.4	13.8	0.0	0.0	1.3	0.9	0.5	0.90
18_EnhAc	0.3	0.3	0.1	1.1	4.9	64.3	19.4	0.7	0.5	3.3	1.2	0.4	0.26
19_DNase	0.1	0.0	0.1	0.0	3.4	0.3	44.7	0.0	0.0	1.4	6.2	0.1	0.60
20_ZNF/Rpts	88.9	82.0	1.0	15.9	0.5	0.1	0.6	0.2	4.7	1.4	0.0	0.1	0.16
21_Het	69.6	0.2	0.0	0.0	0.1	0.0	1.4	0.0	0.2	0.1	0.0	0.5	0.84
22_PromP	2.6	0.3	0.2	2.0	9.6	11.0	19.6	9.1	34.5	67.6	18.0	1.0	0.20
23_PromBiv	2.2	0.3	2.4	4.0	76.6	15.6	29.5	23.9	63.9	83.1	44.4	96.6	0.24
24_ReprPC	1.1	0.1	0.3	0.3	3.4	0.2	1.2	0.0	0.1	0.3	0.1	72.4	1.27
25_Quies	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	78.61

b

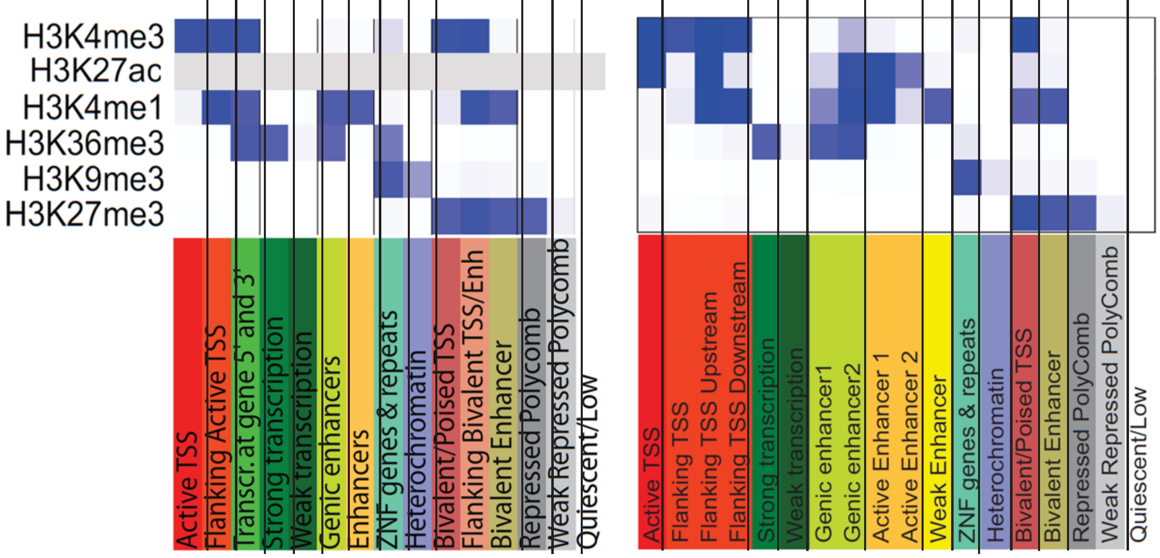
Enrichment for 127-samples 15-state '5-Core-Marks' Observed Model

	1_TssA	2_TssAFlnk	3_TxFlnk	4_Tx	5_TxWk	6_EnhG	7_Enh	8_ZNF/Rpts	9_Het	10_TssBiv	11_BivFlnk	12_EnhBiv	13_ReprPC	14_ReprPCWk	15_Quies
	123.5	2.7	0.3	0.0	0.1	0.0	0.2	0.8	0.0	41.2	2.2	0.5	0.1	0.0	0.0
	44.0	71.3	1.0	0.0	0.1	0.2	4.4	0.5	0.0	64.3	95.3	9.1	0.3	0.1	0.0
	94.2	55.0	12.5	0.0	0.0	0.2	0.8	0.2	0.0	14.6	16.0	1.1	0.1	0.0	0.0
	30.7	47.1	46.0	0.1	0.9	8.0	8.6	2.0	0.0	6.5	4.9	5.7	1.0	0.1	0.0
	0.0	0.0	0.0	1.0	5.4	1.5	1.4	0.7	0.3	0.0	0.0	0.1	0.1	0.2	0.3
	0.0	0.0	0.3	18.6	1.7	14.1	0.2	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	18.1	2.1	1.6	0.1	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	2.5	5.6	0.5	0.5	0.6	0.2	0.0	0.0	0.0	0.1	0.2	0.2
	16.6	55.2	236.4	0.1	0.1	35.5	11.5	0.2	0.0	1.5	2.0	1.1	0.0	0.0	0.0
	0.2	1.6	6.2	1.2	1.2	58.4	20.3	0.5	0.0	0.0	0.0	0.8	0.0	0.0	0.0
	0.1	0.4	4.5	9.5	1.5	74.8	3.8	1.6	0.0	0.0	0.0	0.6	0.1	0.0	0.0
	0.7	0.9	2.7	0.9	3.8	16.9	11.7	0.5	0.0	0.1	0.0	1.0	0.4	0.1	0.1
	6.6	35.8	6.3	0.0	0.2	8.4	22.5	0.2	0.0	0.2	1.1	1.9	0.0	0.0	0.0
	1.1	5.5	0.5	0.0	0.5	3.9	26.9	0.3	0.0	0.2	0.5	4.9	0.1	0.1	0.1
	0.1	0.0	0.0	0.1	1.5	4.7	22.2	0.3	0.1	0.0	0.0	2.8	0.2	0.3	0.2
	4.7	6.9	0.3	0.0	0.9	0.7	17.8	0.2	0.1	2.6	2.4	12.1	0.6	0.8	0.3
	0.1	0.0	0.0	0.2	2.3	1.8	11.9	0.2	0.1	0.0	0.0	3.6	0.5	0.8	0.4
	0.2	0.0	0.0	0.3	2.2	2.5	9.2	0.4	0.3	0.0	0.0	0.9	0.3	0.7	0.5
	0.1	0.0	0.0	0.0	1.0	0.2	4.3	1.1	0.4	0.1	0.1	1.5	0.6	1.5	0.9
	0.0	0.0	0.1	3.9	0.9	0.3	0.0	260.5	6.2	0.0	0.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	0.2	0.0	0.0	22.9	28.4	0.1	0.2	0.1	0.4	0.3	0.4
	22.8	3.8	1.6	0.1	1.1	0.6	4.5	5.4	0.5	18.3	2.5	5.6	2.0	1.4	0.5
	5.5	5.6	0.5	0.0	0.1	0.3	2.2	1.4	0.1	159.8	155.4	183.3	22.2	0.8	0.0
	0.0	0.0	0.0	0.0	0.1	0.0	0.4	0.2	0.2	0.9	0.2	14.2	43.9	5.3	0.1
	0.0	0.0	0.0	0.0	0.5	0.0	0.2	0.1	0.8	0.0	0.0	0.1	0.3	1.1	1.2
base %	0.69	0.45	0.04	3.75	11.71	0.33	2.79	0.19	2.26	0.06	0.06	0.10	0.97	8.29	67.99

c

Enrichment for 98-samples 18-state '5-Core-Marks+H3K27ac' Observed Model

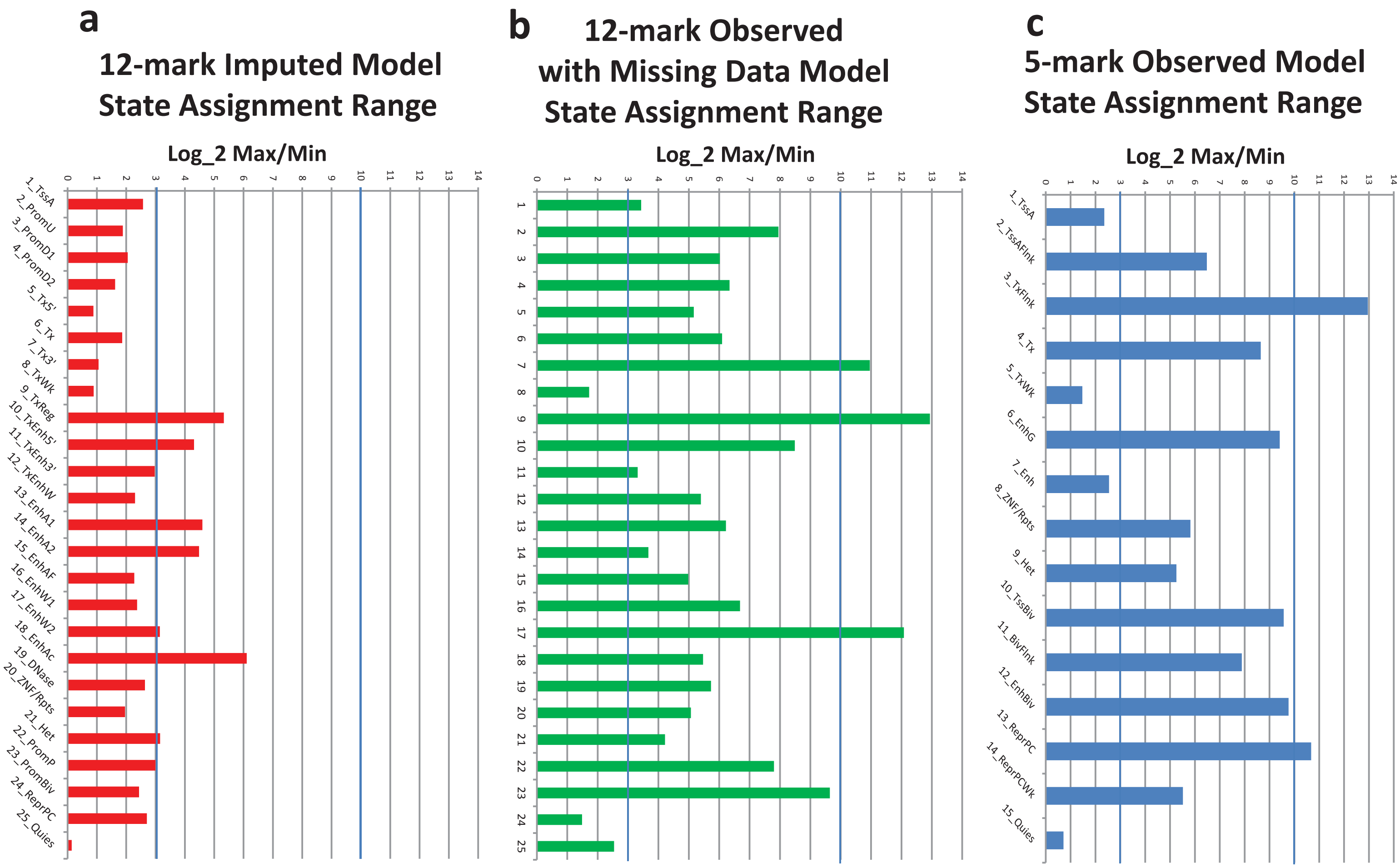
1_TssA	2_TssFlnk	3_TssFlnkU	4_TssFlnkD	5_Tx	6_TxWk	7_EnhG1	8_EnhG2	9_EnhA1	10_EnhA2	11_EnhWk	12_ZNF/Rpts	13_Het	14_TssBiv	15_EnhBiv	16_ReprPC	17_ReprPCWk	18_Quies
126.8	107.5	2.8	3.1	0.0	0.0	0.0	0.0	0.0	1.3	0.1	0.2	0.0	12.0	0.4	0.1	0.0	0.0
41.1	45.8	77.1	65.6	0.0	0.0	0.1	0.5	10.7	3.5	2.0	0.1	0.0	82.4	8.5	0.2	0.0	0.0
124.7	27.0	72.7	33.3	0.0	0.0	0.0	2.6	2.3	0.4	0.3	0.0	0.0	13.1	1.0	0.0	0.0	0.0
9.6	63.3	11.4	86.9	0.1	0.7	8.0	8.2	3.3	4.5	11.2	0.4	0.0	6.1	5.7	1.0	0.1	0.0
0.0	0.1	0.0	0.0	0.8	5.4	2.0	0.0	0.1	1.4	2.1	0.4	0.2	0.0	0.1	0.1	0.2	0.3
0.0	0.0	0.0	0.0	18.1	1.8	23.0	0.9	0.0	0.4	0.4	0.4	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	18.8	2.1	3.9	0.1	0.0	0.2	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	2.5	5.6	0.8	0.0	0.0	1.1	0.7	0.2	0.2	0.0	0.0	0.1	0.2	0.2
23.0	10.9	61.4	42.5	0.0	0.0	8.4	144.9	41.5	3.8	3.0	0.0	0.0	1.3	0.8	0.0	0.0	0.0
0.1	0.4	0.4	4.6	0.5	0.7	42.0	54.9	25.2	14.1	16.7	0.1	0.0	0.0	0.6	0.0	0.0	0.0
0.1	0.4	0.1	1.0	6.9	1.1	87.7	54.4	5.8	7.4	3.9	0.2	0.0	0.0	0.4	0.1	0.0	0.0
0.0	2.5	0.0	3.2	0.6	3.3	17.9	2.2	2.7	9.0	16.0	0.1	0.0	0.1	1.0	0.4	0.1	0.1
7.7	2.8	39.2	12.1	0.0	0.0	1.2	18.9	81.5	11.9	2.7	0.0	0.0	0.4	1.0	0.0	0.0	0.0
0.5	2.7	1.8	12.4	0.0	0.3	1.9	3.5	47.4	17.2	19.2	0.1	0.0	0.2	4.6	0.1	0.1	0.0
0.0	0.2	0.0	0.2	0.1	0.9	3.2	2.8	28.6	28.1	15.5	0.2	0.0	0.0	2.4	0.2	0.2	0.1
0.6	15.5	1.1	20.0	0.0	0.7	0.6	0.2	9.0	11.0	19.0	0.1	0.0	1.8	12.0	0.5	0.7	0.3
0.0	0.4	0.0	0.3	0.2	2.1	2.0	0.3	4.2	15.3	13.9	0.2	0.1	0.0	3.5	0.4	0.7	0.3
0.0	0.5	0.0	0.2	0.3	2.0	2.5	0.3	2.1	18.8	9.0	0.4	0.1	0.0	0.8	0.3	0.6	0.4
0.0	0.4	0.0	0.2	0.0	1.0	0.2	0.0	0.4	3.9	5.5	0.8	0.3	0.5	1.4	0.6	1.5	0.9
0.0	0.1	0.0	0.0	6.5	0.8	1.1	0.2	0.0	0.1	0.0	88.3	1.4	0.1	0.0	0.0	0.0	0.0
0.0	0.1	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.2	0.0	49.5	10.3	1.2	0.1	0.5	0.4	0.2
2.2	65.5	0.6	9.6	0.1	1.0	0.6	0.3	0.6	4.8	5.2	2.2	0.2	7.1	5.8	2.2	1.2	0.5
0.7	20.3	1.2	12.0	0.0	0.1	0.2	0.2	1.0	0.6	2.9	0.6	0.0	163.1	212.8	25.3	0.8	0.0
0.0	0.2	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.2	0.5	0.5	0.1	0.5	14.4	50.3	5.6	0.1
0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.2	0.3	0.3	1.1	0.0	0.1	0.3	1.1	1.2
0.45	0.26	0.27	0.21	3.40	11.80	0.37	0.09	0.63	0.60	1.96	0.57	4.35	0.10	0.09	0.80	8.20	64.83



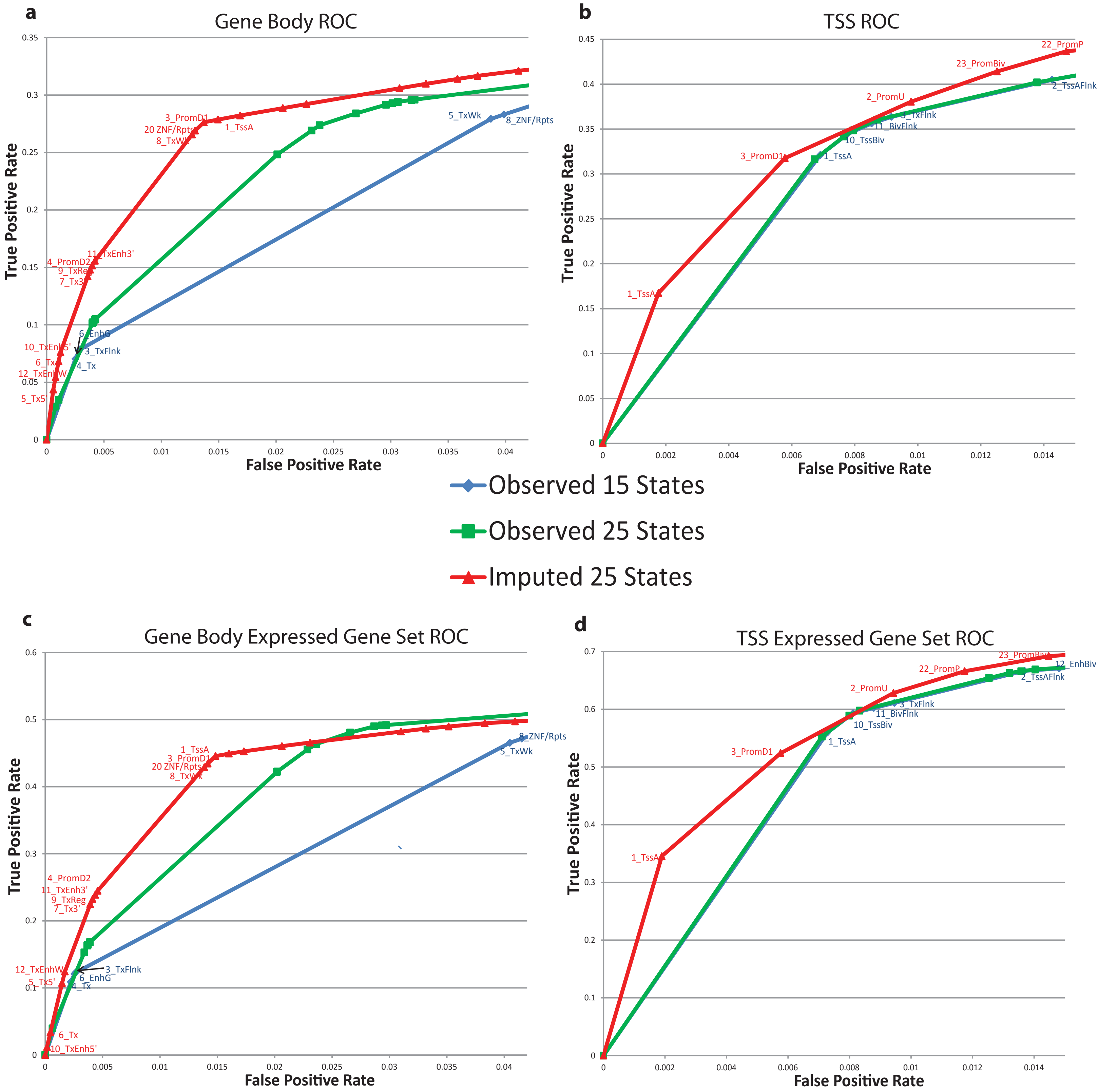
d

	Genome %	CpG.hg19	Exons.Gencodev10.hg19	Genes.Gencodev10.hg19	Introns.Gencodev10.hg19	TES.Gencodev10.hg19	TES_2kbp.Gencodev10.hg19	TSS.Gencodev10.hg19	TSS_2kbp.Gencodev10.hg19	ZNF_genes	Conserved	DNA Methylation Observed	DNA Methylation Imputed	RNA-seq Observed	RNA-seq Imputed
1_TssA	0.18	97.6	10.1	1.3	0.6	3.5	2.5	96.9	9.4	3.6	5.9	0.03	0.13	0.6	0.7
2_PromU	0.41	36.5	4.7	1.1	0.8	2.9	2.4	16.1	7.2	1.9	3.5	0.11	0.20	0.4	0.4
3_PromD1	0.41	55.2	8.5	1.8	1.3	3.6	2.7	36.6	9.7	4.3	4.9	0.04	0.13	1.1	1.2
4_PromD2	0.19	2.3	3.4	1.9	1.8	3.2	2.5	2.5	8.0	4.7	2.2	0.62	0.63	0.7	0.8
5_Tx5'	2.22	0.1	0.6	2.0	2.1	0.7	0.9	0.6	1.2	2.4	0.7	0.90	0.89	0.3	0.4
6_Tx	0.70	1.3	5.2	2.0	1.7	6.0	3.8	2.5	2.8	3.3	2.5	0.91	0.89	1.2	1.3
7_Tx3'	3.48	1.0	5.6	1.9	1.7	5.9	4.3	2.0	2.6	2.8	2.5	0.92	0.91	1.0	1.1
8_TxWk	5.88	0.3	1.9	1.9	1.9	1.7	2.1	0.8	1.5	2.4	1.1	0.90	0.89	0.4	0.5
9_TxReg	0.30	2.8	4.0	1.9	1.8	4.2	2.6	3.9	5.1	1.4	3.3	0.45	0.52	1.2	1.2
10_TxEnh5'	0.38	0.4	1.8	2.0	2.0	2.0	1.5	1.5	2.0	1.4	2.0	0.76	0.76	0.6	0.8
11_TxEnh3'	0.21	1.2	7.2	1.9	1.5	7.3	4.8	2.7	3.0	1.3	3.4	0.80	0.80	1.4	1.5
12_TxEnhW	0.51	0.3	1.2	2.0	2.0	1.2	1.2	1.1	2.3	2.1	1.4	0.82	0.81	0.4	0.6
13_EnhA1	0.22	0.9	2.2	1.3	1.2	1.9	1.8	2.8	2.5	0.8	2.9	0.46	0.56	0.5	0.5
14_EnhA2	0.34	0.4	1.4	1.2	1.2	1.3	1.3	1.8	2.0	0.8	3.0	0.54	0.61	0.2	0.3
15_EnhAF	0.48	0.2	1.3	1.2	1.2	1.2	1.3	1.1	1.6	0.7	2.0	0.68	0.73	0.2	0.3
16_EnhW1	0.28	1.8	1.7	1.0	0.9	1.6	1.4	3.0	3.0	1.0	2.5	0.52	0.54	0.1	0.2
17_EnhW2	0.95	0.3	1.2	1.2	1.3	1.2	1.2	1.0	1.5	0.8	2.1	0.74	0.75	0.2	0.3
18_EnhAc	0.27	0.2	1.1	1.2	1.2	1.0	1.2	1.3	1.6	0.7	1.9	0.75	0.78	0.2	0.3
19_DNase	0.63	0.4	0.9	0.9	0.9	1.0	1.0	1.7	1.2	0.5	2.7	0.57	0.58	0.1	0.1
20_ZNF/Rpts	0.18	0.7	5.0	1.9	1.6	4.0	3.3	1.0	1.5	71.8	1.0	0.93	0.91	0.7	0.9
21_Het	0.91	0.9	1.0	0.7	0.7	0.9	0.8	0.5	0.8	7.7	0.2	0.81	0.79	0.1	0.1
22_PromP	0.20	14.2	3.0	1.2	1.1	2.2	1.8	10.5	5.0	1.5	2.6	0.39	0.43	0.2	0.3
23_PromBiv	0.25	53.5	5.9	1.3	1.0	3.7	2.6	13.0	6.8	0.7	4.5	0.16	0.22	0.2	0.3
24_ReprPC	1.32	4.9	2.0	1.0	0.9	1.7	1.7	1.7	2.8	0.5	1.9	0.63	0.62	0.1	0.1
25_Quies	78.38	0.1	0.5	0.8	0.8	0.6	0.7	0.4	0.6	0.5	0.8	0.85	0.85	0.0	0.1

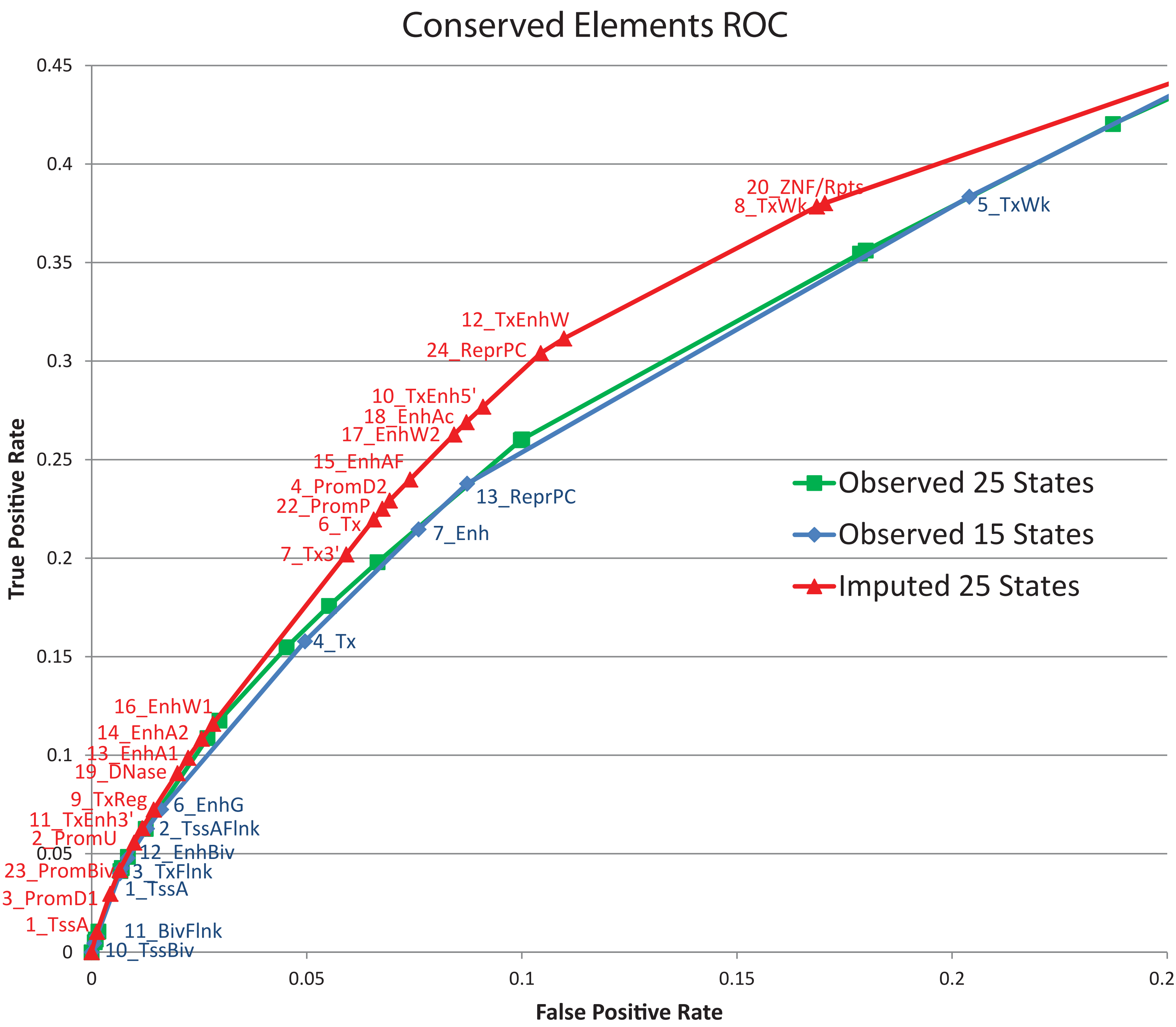
Supplementary Figure 33:
Enrichments of 12-mark Imputation Based Chromatin State Model.
(a) The emission parameters for the 25-state 12-mark imputation based model followed by the median percent of the genome across samples assigned to each state. **(b)** The median enrichments of the state assignments of the imputed data based model for the state assignments of a 15-state model based on observed data for five core marks (H3K4me1, H3K4me3, H3K9me3, H3K27me3, H3K36me3) from (Roadmap Epigenomics Consortium et al, 2015)[10]. The bottom row shows the median state assignment percentages. Below it is a heatmap of the emission parameters and state descriptions from (Roadmap Epigenomics Consortium et al, 2015)[10]. H3K27ac was not used in this model and is grayed out in the heatmap. **(c)** Similar to **b** except comparing state assignments for the imputation model for the state assignments of a 18-state model based on the core marks plus H3K27ac restricted to the 98 samples in which it was defined, also from (Roadmap Epigenomics Consortium et al, 2015)[10]. **(d)** Median state overlaps from **Fig. 6c** here with numeric details. All values are fold enrichments except the column after the state label is genome % and the last four columns report average signal values for observed and imputed DNA-methylation at CGs or RNA-seq data as indicated.



Supplementary Figure 34: Comparison of Chromatin State Model State Assignment Coverage Ranges. The graphs shows for three different models and each state the range of percentage of the genome assigned to each state in terms of log base two ratio of the maximum genome percentage for any sample divided by the minimum genome percentage. The three models are **(a)** 25-state imputation based model based on the 12 Tier-1 and 2 marks. **(b)** 25-state observed data model based on the 12 Tier-1 and 2 marks treating marks as missing data for some samples (**Fig. S32**). **(c)** 15-state observed data model based on 5-core marks mapped in every sample[10] (also see **Fig. S33** for emission parameters). The graphs indicate that the chromatin states inferred based on the imputed data have a more consistent fraction of the genome assigned across samples than based on the observed data.



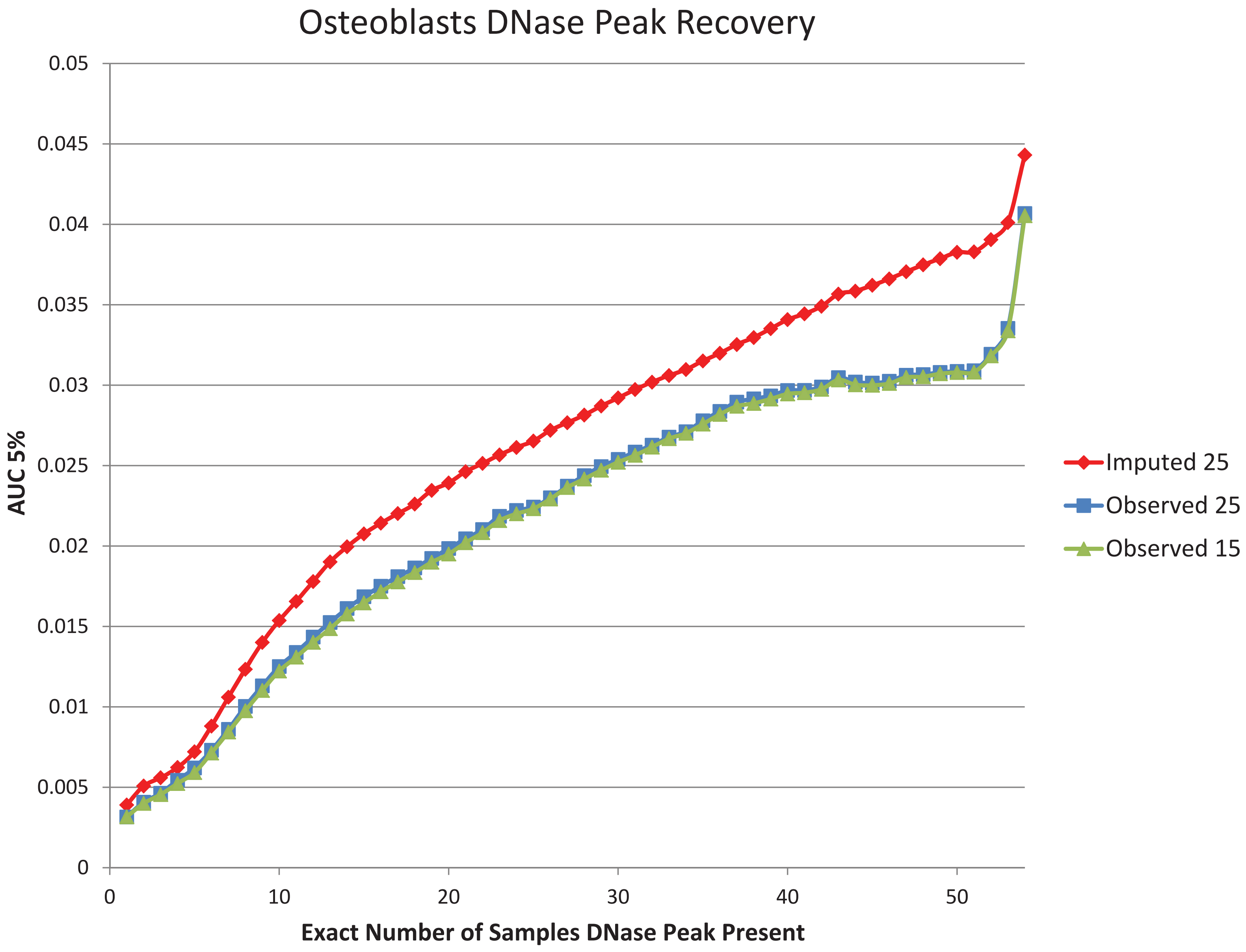
Supplementary Figure 35: Comparison of Chromatin State Models at Recovering Annotated Gene Features. **(a)** The plot compares the chromatin state agreement with annotated genes for the 25-state model based on imputed data for 12-marks, the 15-state model based on observed data for 5-marks[10], and a 25-state model based on observed data for 5-marks learned in the same way as the 15-state model. The plot shows for each model a portion of the best possible ROC curve for recovering bases overlapping annotated genes based on a single ordering of states used across all samples. Predictions are made for each sample and the ROC curve represents the combined results. Labeled on the ROC curve are the top prioritized states for the 25-state imputed and the 15-state observed models at the cumulative true positive and false positive rate after making predictions based on the indicated state is in included. **(b)** The same plot as in **a** but for annotated transcription start sites. **(c), (d)** The same as **a** and **b**, but based on a set of expressed genes (see **Methods**) and only using samples with gene expression data available.



Supplementary Figure 36:

Comparison of Chromatin State Models at Recovering Evolutionarily Conserved Elements.

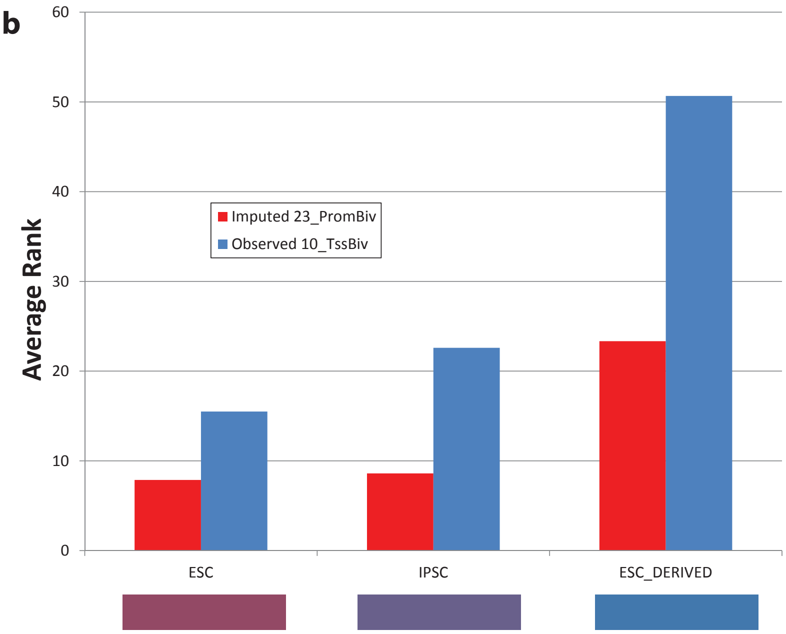
A similar plot to those shown in **Fig. S35**, but for recovering conserved elements based on the PI measure[38,53].



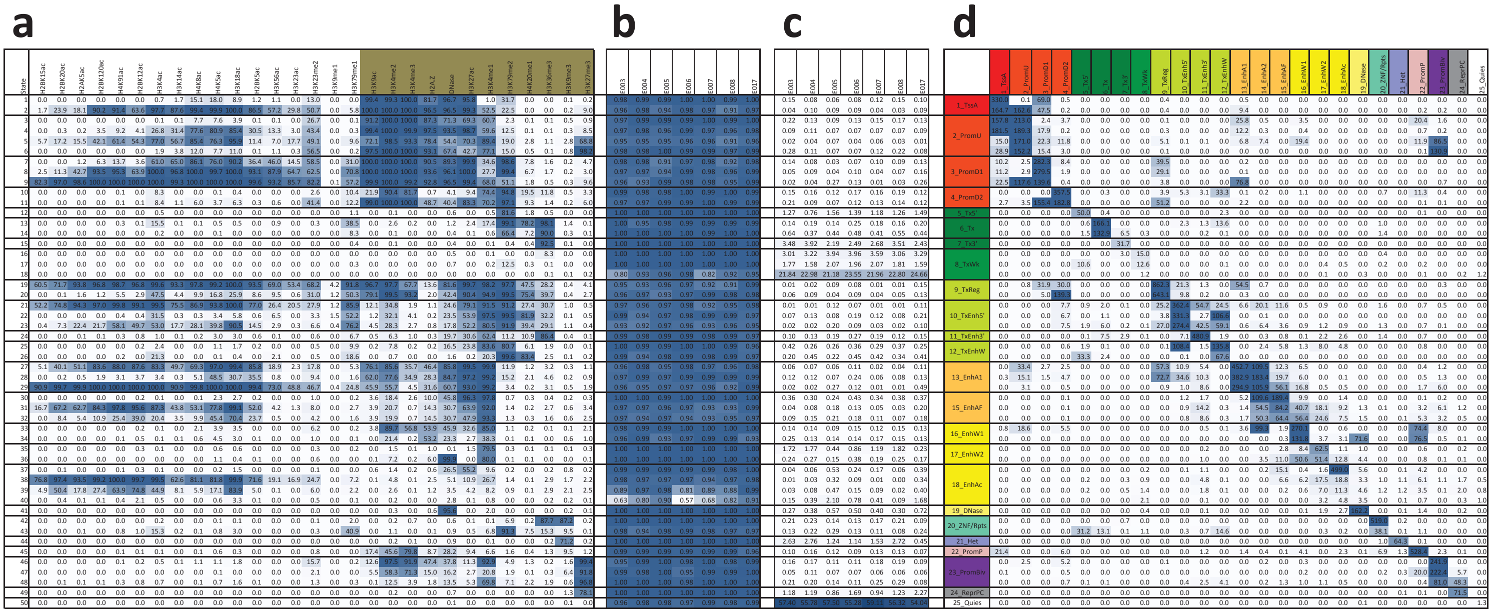
Supplementary Figure 37: Chromatin State Correspondence with Osetoblast DNase Sites. The figure reports the AUC up to a false positive rate of 5% for chromatin state models recovering bases covered by a DNase site in Osteoblast cells corresponding to sample E129. The bases are stratified based on the number of samples that had a DNase peak overlapping it including this dataset in Osteoblast cells. Locations which had a DNase peak, but did not overlap the number of samples being evaluated were excluded for the specific evaluation. This DNase dataset was based on a different DNase protocol[39] than the other datasets and thus was not included in the processed data in (Roadmap Epigenomics Consortium et al, 2015)[10] or used for the imputation here, and E129 otherwise did not have a DNase dataset associated with it. The three chromatin state models being compared are the imputed 25-state model, the observed 15-state model based on the 5-core marks[10] and an observed 25-state model which was generated in the same way as the observed 15-state model except for having a different number of states. In order to generate the ROC curve to compute an AUC 5% value the chromatin states were ordered based on the greatest fold enrichment for the Osteoblast DNase peak bases. This figure demonstrates that the imputed model was better able to recover bases covered by DNase peaks from this data, though this difference generally increased for DNase peaks which were observed in increasing numbers of other samples.

a

Ranking	Imputed Model 23_PromBiv	Observed Model 10_TssBiv
1	E018_iPS-15b Cell Line	E018_iPS-15b Cell Line
2	E014_HUES48 Cell Line	E020_iPS-20b Cell Line
3	E020_iPS-20b Cell Line	E016_HUES64 Cell Line
4	E019_iPS-18 Cell Line	E001_ES-I3 Cell Line
5	E002_ES-WA7 Cell Line	E118_HepG2 Hepatocellular Carcinoma Cell Line
6	E015_HUES6 Cell Line	E015_HUES6 Cell Line
7	E016_HUES64 Cell Line	E019_iPS-18 Cell Line
8	E011_hESC Derived CD184+ Endoderm Cultured Cells	E014_HUES48 Cell Line
9	E008_H9 Cell Line	E012_hESC Derived CD56+ Ectoderm Cultured Cells
10	E003_H1 Cell Line	E070_Brain Germinal Matrix
11	E024_ES-UCSF4 Cell Line	E003_H1 Cell Line
12	E012_hESC Derived CD56+ Ectoderm Cultured Cells	E057_Foreskin Keratinocyte Primary Cells skin02
13	E001_ES-I3 Cell Line	E082_Fetal Brain Female
14	E004_H1 BMP4 Derived Mesendoderm Cultured Cells	E008_H9 Cell Line
15	E007_H1 Derived Neuronal Progenitor Cultured Cells	E011_hESC Derived CD184+ Endoderm Cultured Cells
16	E021_iPS DF 6.9 Cell Line	E092_Fetal Stomach
17	E009_H9 Derived Neuronal Progenitor Cultured Cells	E112_Thymus
18	E089_Fetal Muscle Trunk	E078_Duodenum Smooth Muscle
19	E022_iPS DF 19.11 Cell Line	E002_ES-WA7 Cell Line
20	E091_Placenta	E084_Fetal Intestine Large
21	E092_Fetal Stomach	E056_Foreskin Fibroblast Primary Cells skin02
22	E010_H9 Derived Neuron Cultured Cells	E085_Fetal Intestine Small
23	E056_Foreskin Fibroblast Primary Cells skin02	E021_iPS DF 6.9 Cell Line
24	E013_hESC Derived CD56+ Mesoderm Cultured Cells	E013_hESC Derived CD56+ Mesoderm Cultured Cells
25	E083_Fetal Heart	E103_Rectal Smooth Muscle
26	E090_Fetal Muscle Leg	E004_H1 BMP4 Derived Mesendoderm Cultured Cells
27	E081_Fetal Brain Male	E111_Stomach Smooth Muscle
28	E035_Primary hematopoietic stem cells	E071_Brain Hippocampus Middle
29	E099_Placenta Amnion	E058_Foreskin Keratinocyte Primary Cells skin03
30	E080_Fetal Adrenal Gland	E069_Brain Cingulate Gyrus
31	E093_Fetal Thymus	E055_Foreskin Fibroblast Primary Cells skin01
32	E088_Fetal Lung	E033_Primary T cells from cord blood
33	E086_Fetal Kidney	E025_Adipose Derived Mesenchymal Stem Cell Cultured Cells
34	E059_Foreskin Melanocyte Primary Cells skin01	E101_Rectal Mucosa Donor 29
35	E070_Brain Germinal Matrix	E077_Duodenum Mucosa
36	E036_Primary hematopoietic stem cells short term culture	E086_Fetal Kidney
37	E118_HepG2 Hepatocellular Carcinoma Cell Line	E107_Skeletal Muscle Male
38	E033_Primary T cells from cord blood	E073_Brain Dorsolateral Prefrontal Cortex
39	E030_Primary neutrophils from peripheral blood	E072_Brain Inferior Temporal Lobe
40	E123_K562 Leukemia Cell Line	E100_Psoas Muscle
41	E005_H1 BMP4 Derived Trophoblast Cultured Cells	E075_Colonic Mucosa
42	E027_Breast Myoepithelial Primary Cells	E090_Fetal Muscle Leg
43	E084_Fetal Intestine Large	E127_NHEK-Epidermal Keratinocyte Primary Cells
44	E045_Primary T cells effector/memory enriched from peripheral blood	E108_Skeletal Muscle Female
45	E085_Fetal Intestine Small	E076_Colon Smooth Muscle
46	E113_Spleen	E089_Fetal Muscle Trunk
47	E031_Primary B cells from cord blood	E080_Fetal Adrenal Gland
48	E112_Thymus	E114_A549 EtOH 0.02pct Lung Carcinoma Cell Line
49	E076_Colon Smooth Muscle	E062_Primary mononuclear cells from peripheral blood
50	E110_Stomach Mucosa	E102_Rectal Mucosa Donor 31
51	E039_Primary T helper naive cells from peripheral blood	E009_H9 Derived Neuronal Progenitor Cultured Cells
52	E029_Primary monocytes from peripheral blood	E061_Foreskin Melanocyte Primary Cells skin03
53	E042_Primary T helper 17 cells PMA-I stimulated	E129_Osteoblast Primary Cells
54	E075_Colonic Mucosa	E120_HSMM Skeletal Muscle Myoblasts Cell Line
55	E103_Rectal Smooth Muscle	E068_Brain Anterior Caudate
56	E111_Stomach Smooth Muscle	E059_Foreskin Melanocyte Primary Cells skin01
57	E006_H1 Derived Mesenchymal Stem Cells	E017_IMR90 fetal lung fibroblasts Cell Line
58	E047_Primary T killer naive cells from peripheral blood	E093_Fetal Thymus
59	E057_Foreskin Keratinocyte Primary Cells skin02	E024_ES-UCSF4 Cell Line
60	E109_Small Intestine	E088_Fetal Lung
61	E063_Adipose Nuclei	E026_Bone Marrow Derived Cultured Mesenchymal Stem Cells
62	E046_Primary Natural Killer cells from peripheral blood	E027_Breast Myoepithelial Primary Cells
63	E053_Cortex derived primary cultured neurospheres	E067_Brain Angular Gyrus
64	E044_Primary T regulatory cells from peripheral blood	E128_NHLF Lung Fibroblast Primary Cells
65	E043_Primary T helper cells from peripheral blood	E125_NH-A Astrocytes Cell Line
66	E062_Primary mononuclear cells from peripheral blood	E049_Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells
67	E082_Fetal Brain Female	E121_HSMM cell derived Skeletal Muscle Myotubes Cell Line
68	E054_Ganglion Eminence derived primary cultured neurospheres	E052_Muscle Satellite Cultured Cells
69	E077_Duodenum Mucosa	E023_Mesenchymal Stem Cell Derived Adipocyte Cultured Cells
70	E038_Primary T helper naive cells from peripheral blood	E087_Pancreatic Islets
71	E034_Primary T cells from peripheral blood	E007_H1 Derived Neuronal Progenitor Cultured Cells
72	E078_Duodenum Smooth Muscle	E124_Monocytes-CD14+ RO01746 Cell Line
73	E032_Primary B cells from peripheral blood	E126_NHDF-Ad Adult Dermal Fibroblast Primary Cells
74	E107_Skeletal Muscle Male	E054_Ganglion Eminence derived primary cultured neurospheres
75	E055_Foreskin Fibroblast Primary Cells skin01	E074_Brain Substantia Nigra
76	E037_Primary T helper memory cells from peripheral blood 2	E044_Primary T regulatory cells from peripheral blood
77	E119_HMEC Mammary Epithelial Primary Cells	E010_H9 Derived Neuron Cultured Cells
78	E108_Skeletal Muscle Female	E109_Small Intestine
79	E040_Primary T helper memory cells from peripheral blood 1	E095_Left Ventricle
80	E050_Primary hematopoietic stem cells G-CSF-mobilized Female	E022_iPS DF 19.11 Cell Line
81	E073_Brain Dorsolateral Prefrontal Cortex	E106_Sigmoid Colon
82	E048_Primary T killer memory cells from peripheral blood	E005_H1 BMP4 Derived Trophoblast Cultured Cells
83	E066_Liver	E063_Adipose Nuclei
84	E116_GM12878 Lymphoblastoid Cell Line	E110_Stomach Mucosa
85	E051_Primary hematopoietic stem cells G-CSF-mobilized Male	E079_Esophagus
86	E097_Ovary	E081_Fetal Brain Male
87	E106_Sigmoid Colon	E028_Breast variant Human Mammary Epithelial Cells (vHMEC)
88	E068_Brain Anterior Caudate	E105_Right Ventricle
89	E041_Primary T helper cells PMA-I stimulated	E116_GM12878 Lymphoblastoid Cell Line
90	E017_IMR90 fetal lung fibroblasts Cell Line	E045_Primary T cells effector/memory enriched from peripheral blood
91	E105_Right Ventricle	E053_Cortex derived primary cultured neurospheres
92	E074_Brain Substantia Nigra	E122_HUVEC Umbilical Vein Endothelial Cells Cell Line
93	E117_HeLa-S3 Cervical Carcinoma Cell Line	E091_Placenta
94	E101_Rectal Mucosa Donor 29	E031_Primary B cells from cord blood
95	E087_Pancreatic Islets	E030_Primary neutrophils from peripheral blood
96	E072_Brain Inferior Temporal Lobe	E039_Primary T helper naive cells from peripheral blood
97	E069_Brain Cingulate Gyrus	E035_Primary hematopoietic stem cells
98	E114_A549 EtOH 0.02pct Lung Carcinoma Cell Line	E041_Primary T helper cells PMA-I stimulated
99	E028_Breast variant Human Mammary Epithelial Cells (vHMEC)	E051_Primary hematopoietic stem cells G-CSF-mobilized Male
100	E094_Gastric	E034_Primary T cells from peripheral blood
101	E124_Monocytes-CD14+ RO01746 Cell Line	E006_H1 Derived Mesenchymal Stem Cells
102	E061_Foreskin Melanocyte Primary Cells skin03	E029_Primary monocytes from peripheral blood
103	E115_Dnd41 TCell Leukemia Cell Line	E119_HMEC Mammary Epithelial Primary Cells
104	E102_Rectal Mucosa Donor 31	E097_Ovary
105	E067_Brain Angular Gyrus	E065_Aorta
106	E127_NHEK-Epidermal Keratinocyte Primary Cells	E042_Primary T helper 17 cells PMA-I stimulated
107	E023_Mesenchymal Stem Cell Derived Adipocyte Cultured Cells	E043_Primary T helper cells from peripheral blood
108	E096_Lung	E047_Primary T killer naive cells from peripheral blood
109	E125_NH-A Astrocytes Cell Line	E099_Placenta Amnion
110	E128_NHLF Lung Fibroblast Primary Cells	E048_Primary T killer memory cells from peripheral blood
111	E025_Adipose Derived Mesenchymal Stem Cell Cultured Cells	E050_Primary hematopoietic stem cells G-CSF-mobilized Female
112	E098_Pancreas	E036_Primary hematopoietic stem cells short term culture
113	E049_Mesenchymal Stem Cell Derived Chondrocyte Cultured Cells	E038_Primary T helper naive cells from peripheral blood
114	E071_Brain Hippocampus Middle	E123_K562 Leukemia Cell Line
115	E122_HUVEC Umbilical Vein Endothelial Cells Cell Line	E096_Lung
116	E121_HSMM cell derived Skeletal Muscle Myotubes Cell Line	E032_Primary B cells from peripheral blood
117	E100_Psoas Muscle	E046_Primary Natural Killer cells from peripheral blood
118	E058_Foreskin Keratinocyte Primary Cells skin03	E104_Right Atrium
119	E079_Esophagus	E083_Fetal Heart
120	E120_HSMM Skeletal Muscle Myoblasts Cell Line	E113_Spleen
121	E129_Osteoblast Primary Cells	E117_HeLa-S3 Cervical Carcinoma Cell Line
122	E026_Bone Marrow Derived Cultured Mesenchymal Stem Cells	E115_Dnd41 TCell Leukemia Cell Line
123	E126_NHDF-Ad Adult Dermal Fibroblast Primary Cells	E037_Primary T helper memory cells from peripheral blood 2
124	E052_Muscle Satellite Cultured Cells	E066_Liver
125	E095_Left Ventricle	E040_Primary T helper memory cells from peripheral blood 1
126	E104_Right Atrium	E094_Gastric
127	E065_Aorta	E098_Pancreas



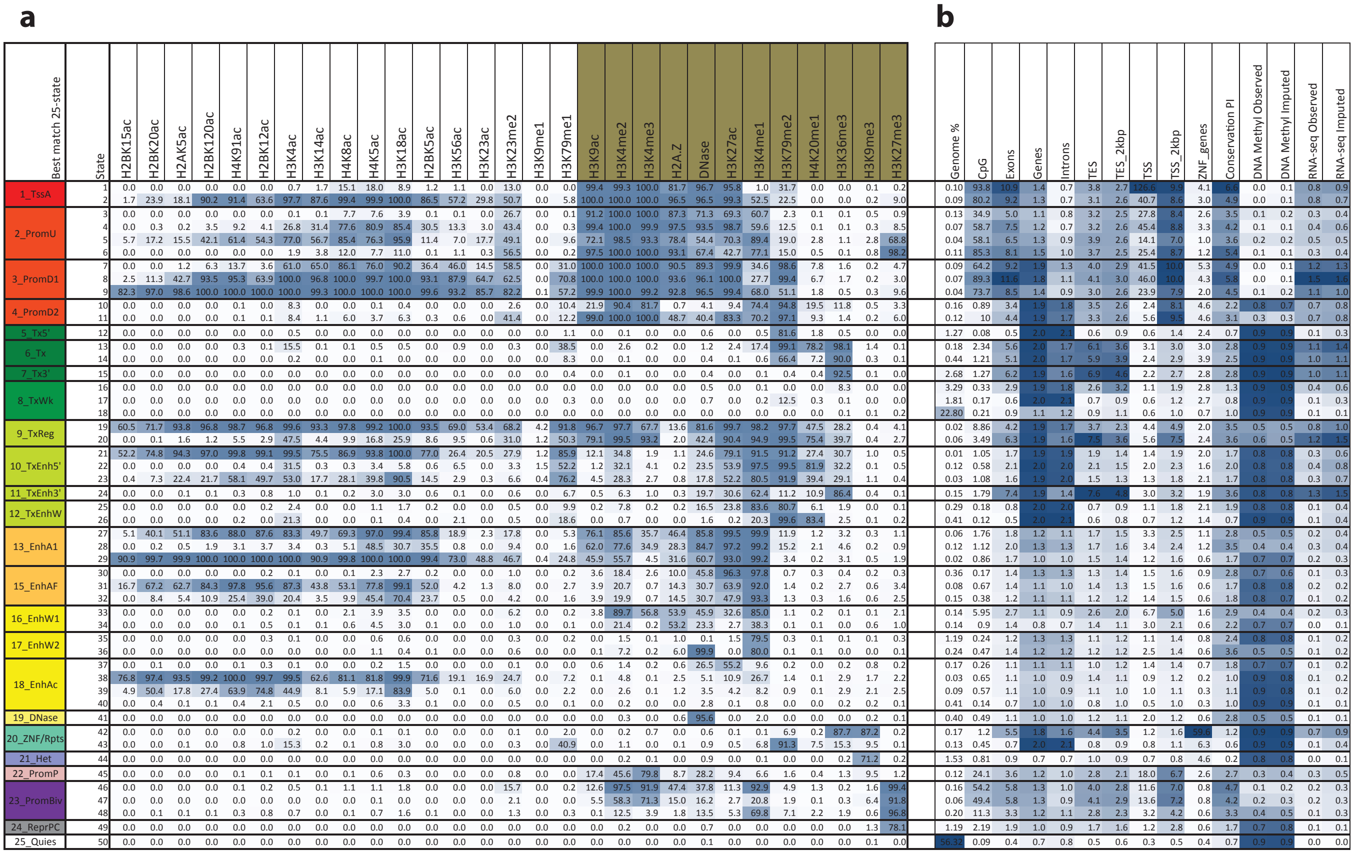
Supplementary Figure 38:
Comparison of Sample Ranking for Bivalent State.
(a) On left the 127 samples are ranked based on the relative amount of genome assigned to state 23_PromBiv a bivalent promoter state from a chromatin state mode learned from the imputed data (**Fig. 6c**). On the right the same samples are ranked, but based on the presence of a bivalent promoter state (10_TssBiv) from a chromatin state model learned based on the observed data for 5-core marks (**Fig. S33**)[10]. The samples are colored based on their biological groups (**Fig. 1**). **(b)** The figure shows that for samples corresponding to Embryonic Stem Cells (ESC), Induced Pluripotent Stem Cells (IPSC), and ESC derived cells (ESC_DERIVED), which could be expected to have a greater presence of the bivalent state[40] relative to other samples, the average of their ranks is lower (corresponding to a greater relative presence of the bivalent state) for the imputed model compared to the observed data model. Colors corresponding to each of these three groups is shown below the graph and the individual samples that comprise them have this color in **a**.



Supplementary Figure 39:

Comparison of 29-mark and 12-mark Imputation Based Chromatin State Models.

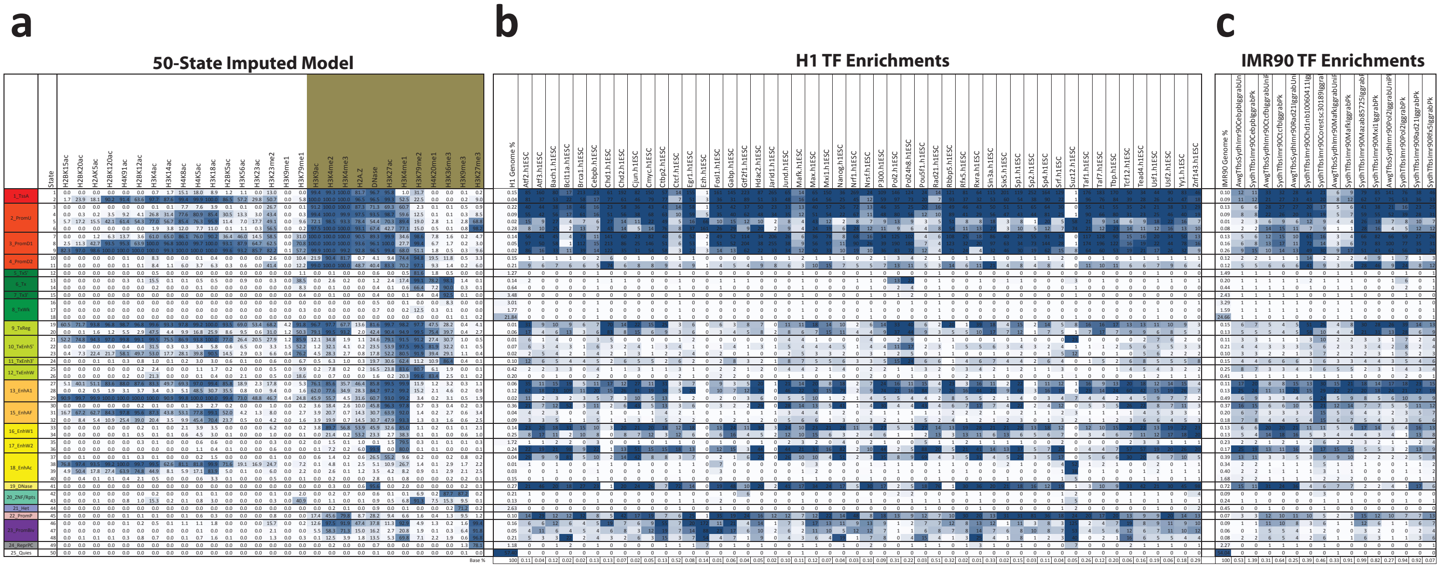
(a) The emission parameters for a 50-state chromatin model learned based on imputed data for 29-marks across the seven samples with deep mark coverage (also see **Fig. 6d**). The marks that are shaded are the Tier-1 and 2 marks. **(b)** The heatmap shows the correlation of the emission parameters with the frequency each mark was found at locations assigned to the state for each of the seven samples. **(c)** The percentage of the genome assigned to each state in the seven samples used to learn the model. **(d)** The median fold enrichment of each state across the seven samples for each state of a 25-state model learned based on the imputed data for the Tier-1 and 2 marks (**Figs. 6c and S33**). The states of the 50-state model have been grouped based on their maximum enrichment for states of this 25-state model.



Best match 25-state	State	H2BK15ac	H2BK20ac	H2AK5ac	H2BK120ac	H4K91ac	H2BK12ac	H3K4ac	H3K14ac	H4K8ac	H4K5ac	H3K18ac	H2BK5ac	H3K56ac	H3K23ac	H3K23me2	H3K9me1	H3K79me1	H3K9ac	H3K4me2	H3K4me3	H2A.Z	DNase	H3K27ac	H3K4me1	H3K79me2	H4K20me1	H3K36me3	H3K9me3	H3K27me3
1_TssA	1	0.0	0.0	0.0	0.0	0.0	0.0	0.7	1.7	15.1	18.0	8.9	1.2	1.1	0.0	13.0	0.0	0.0	99.4	99.3	100.0	81.7	96.7	95.8	1.0	31.7	0.0	0.0	0.1	0.2
	2	1.7	23.9	18.1	90.2	91.4	63.6	97.7	87.6	99.4	99.9	100.0	86.5	57.2	29.8	50.7	0.0	5.8	100.0	100.0	100.0	96.5	96.5	99.3	52.5	22.5	0.0	0.0	0.2	9.0
2_PromU	3	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	7.7	7.6	3.9	0.1	0.1	0.0	26.7	0.0	0.1	91.2	100.0	100.0	87.3	71.3	69.3	60.7	2.3	0.1	0.1	0.5	0.9
	4	0.0	0.3	0.2	3.5	9.2	4.1	26.8	31.4	77.6	80.9	85.4	30.5	13.3	3.0	43.4	0.0	0.3	99.4	100.0	99.9	97.5	93.5	98.7	59.6	12.5	0.1	0.1	0.3	8.5
	5	5.7	17.2	15.5	42.1	61.4	54.3	77.0	56.7	85.4	76.3	95.9	11.4	7.0	17.7	49.1	0.0	9.6	72.1	98.5	93.3	78.4	54.4	70.3	89.4	19.0	2.8	1.1	2.8	68.8
	6	0.0	0.0	0.0	0.0	0.0	0.0	1.9	3.8	12.0	7.7	11.0	0.1	1.1	0.3	56.5	0.0	0.2	97.5	100.0	100.0	93.1	67.4	42.7	77.1	15.0	0.5	0.1	0.8	98.2
3_PromD1	7	0.0	0.0	1.2	6.3	13.7	3.6	61.0	65.0	86.1	76.0	90.2	36.4	46.0	14.5	58.5	0.0	31.0	100.0	100.0	100.0	90.5	89.3	99.9	34.6	98.6	7.8	1.6	0.2	4.7
	8	2.5	11.3	42.7	93.5	95.3	63.9	100.0	96.8	100.0	99.7	100.0	93.1	87.9	64.7	62.5	0.0	70.8	100.0	100.0	100.0	93.6	96.1	100.0	27.7	99.4	6.7	1.7	0.2	3.0
	9	82.3	97.0	98.6	100.0	100.0	100.0	100.0	99.3	100.0	100.0	100.0	99.6	93.2	85.7	82.2	0.1	57.2	99.9	100.0	99.2	92.8	96.5	99.4	68.0	51.1	1.8	0.5	0.3	9.6
4_PromD2	10	0.0	0.0	0.0	0.0	0.1	0.0	8.3	0.0	0.0	0.1	0.4	0.0	0.0	0.0	2.6	0.0	10.4	21.9	90.4	81.7	0.7	4.1	9.4	74.4	94.8	19.5	11.8	0.5	3.3
	11	0.0	0.0	0.0	0.0	0.1	0.1	8.4	1.1	6.0	3.7	6.3	0.3	0.6	0.0	41.4	0.0	12.2	99.0	100.0	100.0	48.7	40.4	83.3	70.2	97.1	9.3	1.4	0.2	6.0
5_Tx5'	12	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.1	0.0	0.0	0.6	0.0	0.5	81.6	1.8	0.5	0.0	0.0
6_Tx	13	0.0	0.0	0.0	0.0	0.3	0.1	15.5	0.1	0.1	0.5	0.5	0.0	0.9	0.0	0.3	0.0	38.5	0.0	2.6	0.2	0.0	1.2	2.4	17.4	99.1	78.2	98.1	1.4	0.1
	14	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	8.3	0.0	0.1	0.0	0.0	0.4	0.1	0.6	66.4	7.2	90.0	0.3	0.1
7_Tx3'	15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.1	0.0	0.0	0.4	0.0	0.1	0.4	0.4	92.5	0.1	0.0
8_TxWk	16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.1	0.0	0.0	8.3	0.0	0.0
	17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.7	0.0	0.2	12.5	0.3	0.1	0.0	0.0
	18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2
9_TxReg	19	60.5	71.7	93.8	96.8	98.7	96.8	99.6	93.3	97.8	99.2	100.0	93.5	69.0	53.4	68.2	4.2	91.8	96.7	97.7	67.7	13.6	81.6	99.7	98.2	97.7	47.5	28.2	0.4	4.1
	20	0.0	0.1	1.6	1.2	5.5	2.9	47.5	4.4	9.9	16.8	25.9	8.6	9.5	0.6	31.0	1.2	50.3	79.1	99.5	93.2	2.0	42.4	90.4	94.9	99.5	75.4	39.7	0.4	2.7
10_TxEnh5'	21	52.2	74.8	94.3	97.0	99.8	99.1	99.5	75.5	86.9	93.8	100.0	77.0	26.4	20.5	27.9	1.2	85.9	12.1	34.8	1.9	1.1	24.6	79.1	91.5	91.2	27.4	30.7	1.0	0.5
	22	0.0	0.0	0.0	0.0	0.4	0.4	31.5	0.3	0.3	3.4	5.8	0.6	6.5	0.0	3.3	1.5	52.2	1.2	32.1	4.1	0.2	23.5	53.9	97.5	99.5	81.9	32.2	0.1	0.5
	23	0.4	7.3	22.4	21.7	58.1	49.7	53.0	17.7	28.1	39.8	90.5	14.5	2.9	0.3	6.6	0.4	76.2	4.5	28.3	2.7	0.8	17.8	52.2	80.5	91.9	39.4	29.1	1.1	0.4
11_TxEnh3'	24	0.0	0.0	0.1	0.1	0.3	0.8	1.0	0.1	0.2	3.0	3.0	0.6	0.1	0.0	0.6	0.0	6.7	0.5	6.3	1.0	0.3	19.7	30.6	62.4	11.2	10.9	86.4	0.4	0.1
12_TxEnhW	25	0.0	0.0	0.0	0.0	0.0	0.2	2.4	0.0	0.0	1.1	1.7	0.2	0.0	0.0	0.5	0.0	9.9	0.2	7.8	0.2	0.2	16.5	23.8	83.6	80.7	6.1	1.9	0.0	0.1
	26	0.0	0.0	0.0	0.0	0.2	0.0	21.3	0.0	0.0	0.1	0.4	0.0	2.1	0.0	0.5	0.0	18.6	0.0	0.7	0.0	0.0	1.6	0.2	20.3	99.6	83.4	2.5	0.1	0.2
13_EnhA1	27	5.1	40.1	51.1	83.6	88.0	87.6	83.3	49.7	69.3	97.0	99.4	85.8	18.9	2.3	17.8	0.0	5.3	76.1	85.6	35.7	46.4	85.8	99.5	99.9	11.9	1.2	3.2	0.3	1.1
	28	0.0	0.2	0.5	1.9	3.1	3.7	3.4	0.3	5.1	48.5	30.7	35.5	0.8	0.0	9.4	0.0	1.6	62.0	77.6	34.9	28.3	84.7	97.2	99.2	15.2	2.1	4.6	0.2	0.9
	29	90.9	99.7	99.9	100.0	100.0	100.0	100.0	90.9	99.8	100.0	100.0	99.4	73.0	48.8	46.7	0.4	24.8	45.9	55.7	4.5	31.6	60.7	93.0	99.2	3.4	0.2	3.1	0.5	1.9
15_EnhAF	30	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.1	2.3	2.7	0.2	0.0	0.0	1.0	0.0	0.2	3.6	18.4	2.6	10.0	45.8	96.3	97.8	0.7	0.3	0.4	0.2	0.3
	31	16.7	67.2	62.7	84.3	97.8	95.6	87.3	43.8	53.1	77.8	99.1	52.0	4.2	1.3	8.0	0.0	2.7	3.9	20.7	0.7	14.3	30.7	63.9	92.0	1.4	0.2	2.7	0.6	3.4
	32	0.0	8.4	5.4	10.9	25.4	39.0	20.4	3.5	9.9	45.4	70.4	23.7	0.5	0.0	4.2	0.0	1.6	3.9	19.9	0.7	14.5	30.7	47.9	93.3	1.3	0.3	1.6	0.6	2.5
16_EnhW1	33	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	2.1	3.9	3.5	0.0	0.0	0.0	6.2	0.0	0.2	3.8	89.7	56.8	53.9	45.9	32.6	85.0	1.1	0.2	0.1	0.1	2.1
	34	0.0	0.0	0.0	0.0	0.1	0.5	0.1	0.1	0.6	4.5	3.0	0.1	0.0	0.0	1.0	0.0	0.1	0.0	21.4	0.2	53.2	23.3	2.7	38.3	0.1	0.1	0.0	0.6	1.0
17_EnhW2	35	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.0	0.6	0.6	0.0	0.0	0.0	0.3	0.0	0.2	0.0	1.5	0.1	1.0	0.1	1.5	79.5	0.3	0.1	0.1	0.1	0.3
	36	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.4	0.1	0.0	0.0	0.6	0.0	0.0	0.1	7.2	0.2	6.0	99.9	0.0	80.0	0.1	0.1	0.0	0.0	0.1
18_EnhAc	37	0.0	0.1	0.0	0.0	0.1	0.3	0.1	0.0	0.0	0.2	1.5	0.0	0.0	0.0	0.1	0.0	0.0	0.6	1.4	0.2	0.6	26.5	55.2	9.6	0.2	0.0	0.2	0.8	0.2
	38	76.8	97.4	93.5	99.2	100.0	99.7	99.5	62.6	81.1	81.8	99.9	71.6	19.1	16.9	24.7	0.0	7.2	0.1	4.8	0.1	2.5	5.1	10.9	26.7	1.4	0.1	2.9	1.7	2.2
	39	4.9	50.4	17.8	27.4	63.9	74.8	44.9	8.1	5.9	17.1	83.9	5.0	0.1	0.0	6.0	0.0	2.2	0.0	2.6	0.1	1.2	3.5	4.2	8.2	0.9	0.1	2.9	2.1	2.5
	40	0.0	0.4	0.1	0.1	0.4	2.1	0.5	0.0	0.0	0.6	3.3	0.1	0.0	0.0	0.5	0.0	0.1	0.0	0.2	0.0	0.0	2.8	0.1	0.8	0.0	0.0	0.2	0.2	0.1
19_DNase	41	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.6	95.6	0.0	2.0	0.0	0.0	0.0	0.2	0.1
20_ZNF/Rpts	42	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.1	0.0	2.0	0.0	0.2	0.7	0.0	0.6	0.1	0.1	6.9	0.2	87.7	87.2	0.2
	43	0.0	0.0	0.1	0.0	0.8	1.0	15.3	0.2	0.1	0.8	3.0	0.0	0.0	0.0	0.3	0.0	40.9	0.0	1.1	0.0	0.1	0.9	0.5	6.8	91.3	7.5	15.3	9.5	0.1
21_Het	44	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.9	0.0	0.1	0.0	0.0	0.2	71.2	0.2
22_PromP	45	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.6	0.3	0.0	0.0	0.0	0.8	0.0	0.0	17.4	45.6	79.8	8.7	28.2	9.4	6.6	1.6	0.4	1.3	9.5	1.2
23_PromBiv	46	0.0	0.0	0.0	0.0	0.1	0.2	0.5	0.1	1.1	1.1	1.8	0.0	0.0	0.0	15.7	0.0	0.2	12.6	97.5	91.9	47.4	37.8	11.3	92.9	4.9	1.3	0.2	1.6	99.4
	47	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	2.1	0.0	0.0	5.5	58.3	71.3	15.0	16.2	2.7	20.8	1.9	0.1	0.3	6.4	91.8
	48	0.0	0.1	0.1	0.1	0.3	0.8	0.7	0.0	0.1	0.7	1.6	0.1	0.0	0.0	1.3	0.0	0.3	0.1	12.5	3.9	1.8	13.5	5.3	69.8	7.1	2.2	1.9	0.6	96.8
24_ReprPC	49	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	1.3	78.1
25_Quies	50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0

Best match 25-state	State	Genome %	CpG	Exons	Genes	Introns	TES	TES_2kbp	TSS	TSS_2kbp	ZNF_genes	Conservation PI	DNA Methyl Observed	DNA Methyl Imputed	RNA-seq Observed	RNA-seq Imputed
1_TssA	1	0.10	93.8	10.9	1.4	0.7	3.8	2.7	126.6	9.9	4.1	6.6	0.0	0.1	0.8	0.9
	2	0.09	80.2	9.2	1.3	0.7	3.1	2.6	40.7	8.6	3.0	4.9	0.0	0.1	0.8	0.7
2_PromU	3	0.13	34.9	5.0	1.1	0.8	3.2	2.5	27.8	8.4	2.6	3.5	0.1	0.2	0.3	0.4
	4	0.07	58.7	7.5	1.2	0.7	3.2	2.6	45.4	8.8	3.3	4.2	0.1	0.1	0.4	0.6
	5	0.04	58.1	6.5	1.3	0.9	3.9	2.6	14.1	7.0	1.0	3.6	0.2	0.2	0.5	0.5
	6	0.11	85.3	8.1	1.5	1.0	3.7	2.5	25.4	8.7	1.2	5.4	0.1	0.1	0.3	0.4
3_PromD1	7	0.09	64.2	9.2	1.9	1.3	4.0	2.9	41.5	10.0	5.3	4.9	0.0	0.1	1.2	1.3
	8	0.07	89.3	11.6	1.8	1.1	4.1	3.0	46.0	10.0	4.3	5.8	0.0	0.1	1.5	1.6
	9	0.04	73.7	8.5	1.4	0.9	3.0	2.6	23.9	7.9	2.0	4.5	0.1	0.2	1.1	1.0
4_PromD2	10	0.16	0.89	3.4	1.9	1.8	3.5	2.6	2.4	8.1	4.6	2.2	0.8	0.7	0.7	0.8
	11	0.12	10	4.4	1.9	1.7	3.3	2.6	5.6	9.5	4.6	3.1	0.3	0.3	0.7	0.8
5_Tx5'	12	1.27	0.08	0.5	2.0	2.1	0.6	0.9	0.6	1.4	2.4	0.7	0.9	0.9	0.1	0.3
6_Tx	13	0.18	2.34	5.6	2.0	1.7	6.1	3.6	3.1	3.0	3.0	2.8	0.9	0.9	1.1	1.4
	14	0.44	1.21	5.1	2.0	1.7	5.9	3.9	2.4	2.9	3.9	2.5	0.9	0.9	1.0	1.1
7_Tx3'	15	2.68	1.27	6.2	1.9	1.6	6.9	4.6	2.2	2.7	2.8	2.8	0.9	0.9	1.0	1.1
8_TxWk	16	3.29	0.33	2.9	1.9	1.8	2.6	3.2	1.1	1.9	2.8	1.3	0.9	0.9	0.4	0.6
	17	1.81	0.17	0.6	2.0	2.1	0.7	0.9	0.6	1.2	2.7	0.8	0.9	0.9	0.1	0.3
	18	22.80	0.21	0.9	1.1	1.2	0.9	1.1	0.6	1.0	0.7	1.0	0.9	0.9	0.1	0.1
9_TxReg	19	0.02	8.86	4.2	1.9	1.7	3.7	2.3	4.4	4.9	2.0	3.5	0.5	0.5	0.8	1.0
	20	0.06	3.49	6.3	1.9	1.6	7.5	3.6	5.6	7.5	2.4	3.6	0.6	0.5	1.2	1.5
10_TxEnh5'	21	0.01	1.05	1.7	1.9	1.9	1.9	1.4	1.4	1.9	2.0	1.7	0.8	0.8	0.3	0.6
	22	0.12	0.58	2.1	2.0	2.0	2.1	1.5	2.0	2.3	1.6	2.0	0.8	0.8	0.4	0.8
	23	0.03	1.08	1.6	1.9	2.0	1.5	1.3	1.3	1.8	2.1	1.7	0.8	0.8	0.3	0.7
11_TxEnh3'	24	0.15	1.79	7.4	1.9	1.4	7.6	4.8	3.0	3.2	1.9	3.6	0.8	0.8	1.3	1.5
12_TxEnhW	25	0.29	0.18	0.8	2.0	2.0	0.7	0.9	1.1	1.8	2.1	1.9	0.8	0.8	0.1	0.4
	26	0.41	0.12	0.5	2.0	2.1	0.6	0.8	0.7	1.2	1.4	0.7	0.9	0.9	0.1	0.4
13_EnhA1	27	0.06	1.76	1.8	1.2	1.1	1.7	1.5	3.3	2.5	1.1	2.8	0.5	0.5	0.2	0.4
	28	0.12	1.12	2.0	1.3	1.3	1.7	1.6	3.4	2.4	1.2	3.5	0.4	0.4	0.3	0.4
	29	0.02	0.86	1.7	1.0	1.0	1.5	1.4	1.2	1.6	0.6	1.7	0.7	0.7	0.2	0.3
15_EnhAF	30	0.36	0.17	1.4	1.3	1.3	1.3	1.4	1.5	1.6	0.9	2.8	0.7	0.6	0.1	0.3
	31	0.08	0.67	1.4	1.1	1.0	1.4	1.3	1.5	1.6	0.6	1.7	0.8	0.8	0.2	0.2
	32	0.15	0.38	1.2	1.1	1.1	1.2	1.2	1.4	1.5	0.6	1.8	0.8	0.7	0.1	0.2
16_EnhW1	33	0.14	5.95	2.7	1.1	0.9	2.6	2.0	6.7	5.0	1.6	2.9	0.4	0.4	0.2	0.3
	34	0.14	0.9	1.4	0.8	0.7	1.4	1.1	3.0	2.1	0.6	2.2	0.7	0.7	0.0	0.1
17_EnhW2	35	1.19	0.24	1.2	1.3	1.3	1.1	1.2	1.1	1.4	0.8	2.4	0.8	0.8	0.1	0.2
	36	0.24	0.47	1.4	1.2	1.2	1.2	1.2	2.5	1.4	0.6	3.6	0.5	0.5	0.1	0.2
18_EnhAc	37	0.17	0.26	1.1	1.1	1.1	1.0	1.2	1.4	1.4	0.7	1.8	0.7	0.7	0.1	0.2
	38	0.03	0.65	1.1	1.0	0.9	1.4	1.1	0.9	1.2	0.2	1.0	0.9	0.8	0.1	0.2
	39	0.09	0.57	1.1	1.0	0.9	1.1	1.0	1.0	1.1	0.3	1.1	0.9	0.8	0.1	0.2
	40	0.41	0.14	0.7	1.0	1.0	0.8	1.0	0.5	1.0	0.3	1.0	0.9	0.9	0.1	0.1
19_DNase	41	0.40	0.49	1.1	1.0	1.0	1.2	1.1	2.0	1.2	0.6	2.8	0.5	0.5	0.1	0.1
20_ZNF/Rpts	42	0.17	1.2	5.5	1.8	1.6	4.4	3.5	1.2	1.6	59.6	1.2	0.9	0.9	0.7	0.9
	43	0.13	0.45	0.7	2.0	2.1	0.8	0.9	0.8	1.1	6.3	0.6	0.9	0.9	0.1	0.4
21_Het	44	1.53	0.81	0.9	0.7	0.7	1.0	0.9	0.7	0.8	4.6	0.2	0.8	0.8	0.0	0.1
22_PromP	45	0.12	24.1	3.6	1.2	1.0	2.8	2.1	18.0	6.7	2.6	2.7	0.3	0.4	0.3	0.5
23_PromBiv	46	0.16	54.2	5.8	1.3	1.0	4.0	2.8	11.6	7.0	0.8	4.7	0.1	0.2	0.2	0.3
	47	0.06	49.4	5.8	1.3	0.9	4.1	2.9	13.6	7.2	0.8	4.2	0.2	0.3	0.2	0.3
	48	0.20	11.3	3.3	1.2	1.1	2.8	2.3	3.2	4.2	0.6	3.3	0.4	0.5	0.1	0.3
24_ReprPC	49	1.19	2.19	1.9	1.0	0.9	1.7	1.6	1.2	2.8	0.6	1.7	0.7	0.8	0.1	0.1
25_Quies	50	56.32	0.09	0.4	0.7	0.8	0.5	0.6	0.3	0.5	0.4	0.7	0.9	0.9	0.0	0.0

Supplementary Figure 40: 29-mark Imputed Data Chromatin State Model General Overlaps. **(a)** The emission parameters of the same expanded mark chromatin state model shown in **Fig. 6d** and **Fig. S39**. **(b)** The median enrichment across the seven samples of the chromatin states for various gene annotations, evolutionarily conserved elements based on the PI measure[38,53], average DNA-methylation levels for CG's based on the observed and imputed data, and average RNA-seq signal levels based on the observed and imputed data also considered in **Figs 6c, S33d**.



Supplementary Figure 41: Chromatin State Transcription Factor Binding Enrichments.
(a) The same emission parameters of the same expanded mark chromatin state model based on imputed data shown in **Fig. 6d** and **Fig. S39-40**. **(b)** The heatmap shows the H1 chromatin state fold enrichment for a collection of H1 ENCODE transcription factor binding datasets based on the uniform processed peak calls[2], which was curated in (Roadmap Epigenomics Consortium et al, 2015)[10]. The first column gives the chromatin state percentages and the last row the genome coverage of the peak call dataset. **(c)** The same heatmap but for the IMR90 chromatin states and a collection of ENCODE transcription factor binding in IMR90 cells[2]. The table shows data for ten different datasets, based on the lab provided peak calls. For five of the data sets there was also available uniformly processed peak calls[2] indicated by the 'Awg' prefix and their enrichments are also shown.

a

Relative Top 1% Agreement and 0.25-concordance for DNA-methylation

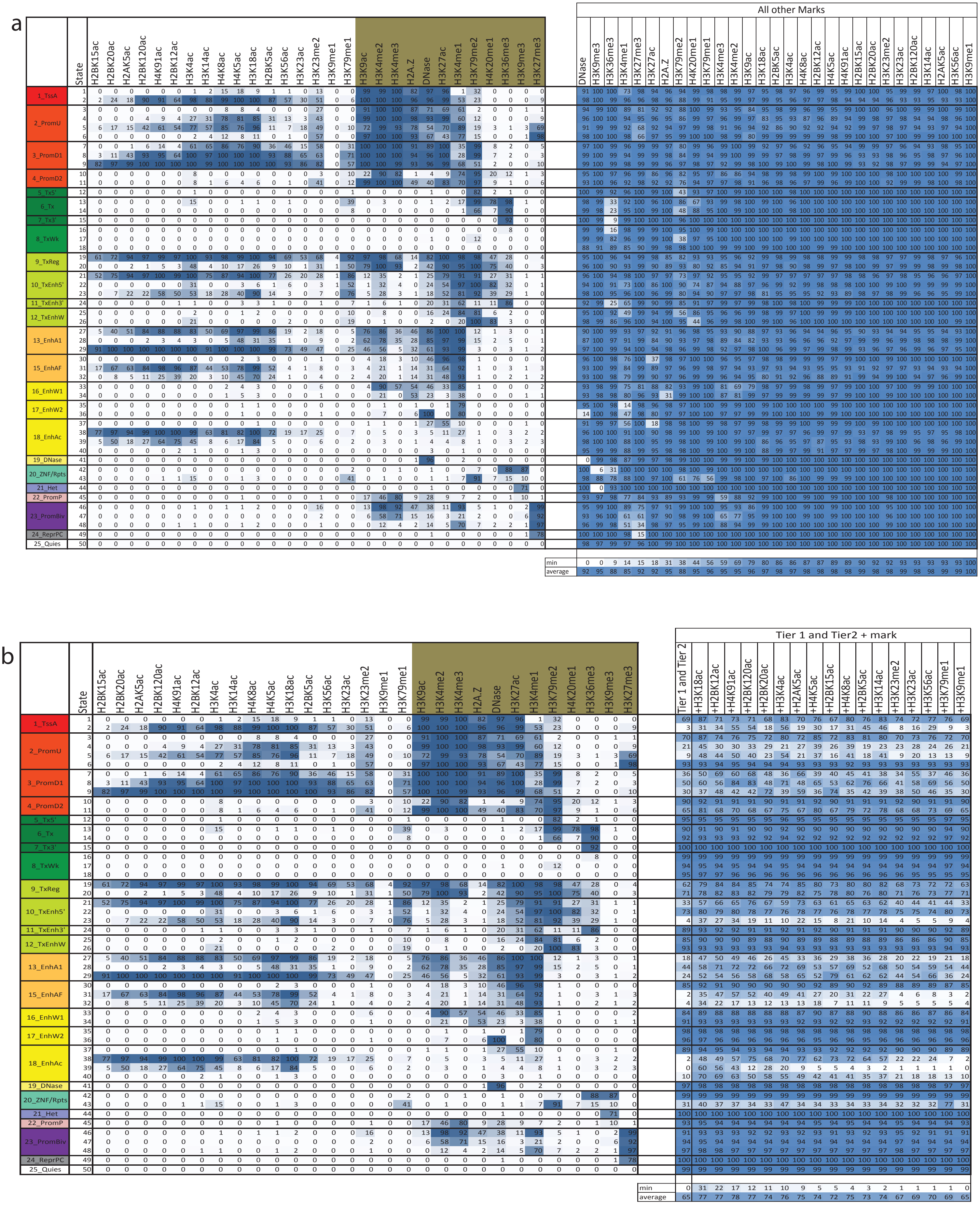
Mark/Feature Set	H3K27me3	H3K36me3	H3K4me1	H3K4me3	H3K9me3	H3K27ac	H3K9ac	DNase	H3K4me2	H2A.Z	H3K79me2	H4K20me1	H2AK5ac	H2BK120ac	H2BK5ac	H3K18ac	H3K23ac	H3K4ac	H3K79me1	H4K8ac	H2BK12ac	H3K14ac	H4K91ac	H2BK15ac	H2BK20ac	H3K56ac	H4K5ac	H3K23me2	RNA-seq	DNA Methylation		All Marks	Acetylations Only
All Tier 1-3 Features	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		1.00	1.00
Same Sample Features Only	0.39	0.26	0.92	0.92	0.24	0.97	0.99	0.70	0.93	0.93	0.82	1.00	0.99	1.00	1.00	0.99	0.99	1.00	0.92	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.20	0.96		0.87	0.99
Same Mark Features Only	0.98	0.97	0.75	0.97	0.97	0.78	0.86	0.98	0.97	0.87	0.98	0.50	0.60	0.60	0.53	0.70	0.47	0.72	0.75	0.67	0.59	0.60	0.58	0.32	0.16	0.32	0.31		1.00	1.00		0.71	0.55
Core + H3K18ac	*1.00*	*0.98*	*0.88*	*0.97*	*1.01*	0.85	0.98	1.00	0.99	1.00	0.97	0.63	0.88	0.90	0.82	*0.71*	0.91	1.04	0.75	0.94	0.87	0.94	0.86	0.85	0.84	0.97	0.96	0.99	1.01	1.00		0.93	0.91
Core + H3K27ac	*1.00*	*1.00*	*0.90*	*0.96*	*1.01*	*0.79*	0.97	1.01	1.00	1.01	0.99	0.63	0.81	0.69	0.90	0.76	0.74	0.76	0.74	0.79	0.64	0.75	0.66	0.51	0.55	0.91	0.90	1.03	1.01	1.00		0.86	0.77
Core + H3K9ac	*0.99*	*1.00*	*0.76*	*0.98*	*1.02*	0.81	*0.95*	0.98	0.99	1.01	0.96	0.62	0.75	0.60	0.68	0.73	0.80	0.72	0.77	0.76	0.60	0.72	0.60	0.53	0.43	0.83	0.75	0.98	1.00	1.00		0.82	0.71
Core + DNase	*0.99*	*0.99*	*0.80*	*0.97*	*1.02*	0.86	0.95	*1.00*	1.00	1.01	0.97	0.61	0.74	0.61	0.69	0.73	0.70	0.72	0.77	0.76	0.61	0.72	0.61	0.47	0.43	0.66	0.68	0.97	1.01	1.00		0.81	0.68
Core (H3K27me3, H3K4me1, H3K4me3, H3K9me3, H3K36me3)	*0.98*	*0.99*	*0.71*	*0.95*	*1.02*	0.79	0.95	1.00	1.00	1.00	0.98	0.62	0.74	0.58	0.67	0.71	0.70	0.70	0.77	0.74	0.59	0.71	0.59	0.45	0.39	0.67	0.66	1.02	1.00	1.00		0.80	0.66
Core + H3K18ac - Same Sample Features Only	*0.33*	*0.12*	*0.70*	*0.57*	*0.23*	0.78	0.95	0.62	0.92	0.93	0.53	0.60	0.85	0.89	0.81	*0.64*	0.90	1.03	0.45	0.92	0.85	0.93	0.85	0.84	0.83	0.97	0.96	0.99	0.21	0.96		0.85	0.90
Core + H3K27ac - Same Sample Features Only	*0.38*	*0.16*	*0.73*	*0.49*	*0.19*	*0.68*	0.95	0.68	0.93	0.91	0.56	0.61	0.74	0.65	0.89	0.71	0.70	0.72	0.49	0.74	0.59	0.73	0.62	0.49	0.55	0.91	0.90	1.03	0.21	0.95		0.78	0.74
Core + H3K9ac - Same Sample Features Only	*0.37*	*0.15*	*0.43*	*0.79*	*0.18*	0.72	*0.92*	0.58	0.93	0.91	0.58	0.60	0.67	0.53	0.63	0.69	0.78	0.68	0.50	0.70	0.54	0.70	0.53	0.50	0.44	0.83	0.74	0.98	0.20	0.95		0.73	0.67
Core + DNase - Same Sample Features Only	*0.37*	*0.15*	*0.61*	*0.57*	*0.23*	0.80	0.93	*0.59*	0.94	0.92	0.55	0.59	0.67	0.55	0.66	0.66	0.65	0.65	0.49	0.70	0.55	0.69	0.56	0.46	0.42	0.67	0.68	0.97	0.20	0.96		0.73	0.64
Core - Same Sample Features Only	*0.35*	*0.15*	*0.31*	*0.17*	*0.20*	0.68	0.92	0.59	0.92	0.90	0.54	0.59	0.65	0.51	0.62	0.64	0.66	0.62	0.51	0.67	0.52	0.69	0.53	0.44	0.39	0.67	0.67	1.02	0.20	0.95		0.70	0.62
Tier 1 and 2 Marks	*1.01*	*1.00*	*0.94*	*1.00*	*1.01*	*0.84*	*0.98*	*1.01*	*0.99*	*1.02*	*0.95*	*1.03*	0.78	0.69	0.89	0.78	0.83	0.68	0.99	0.84	0.66	0.77	0.69	0.56	0.56	0.86	0.91	1.16	1.01	1.00		0.89	0.78
Tier 1 and 2 Marks - Same Sample Features Only	*0.45*	*0.19*	*0.90*	*0.95*	*0.16*	*0.82*	*0.97*	*0.67*	*0.96*	*0.93*	*0.57*	*1.01*	0.73	0.67	0.88	0.75	0.81	0.68	0.96	0.82	0.64	0.76	0.67	0.53	0.56	0.86	0.91	1.16	0.20	0.97		0.85	0.77
H3K27me3		0.92	0.65	0.96	0.98	0.68	0.85	0.98	0.94	0.85	0.94	0.29	0.48	0.45	0.46	0.59	0.37	0.58	0.75	0.59	0.42	0.50	0.43	0.20	0.05	0.21	0.20	0.32	1.00	1.00		0.62	0.44
H3K36me3	0.94		0.65	0.95	0.97	0.61	0.82	0.83	0.93	0.88	0.98	0.46	0.46	0.41	0.43	0.57	0.43	0.54	0.73	0.59	0.40	0.52	0.44	0.22	0.09	0.25	0.28	0.10	1.00	0.99		0.62	0.44
H3K4me1	0.98	1.00		0.96	0.96	0.76	0.86	0.98	0.96	0.92	0.99	0.51	0.72	0.55	0.62	0.67	0.58	0.66	0.75	0.67	0.57	0.62	0.56	0.47	0.44	0.44	0.53	0.15	1.01	0.99		0.73	0.61
H3K4me3	0.95	0.98	0.68		0.97	0.65	0.93	0.86	0.98	0.95	0.94	0.37	0.45	0.45	0.48	0.60	0.61	0.58	0.64	0.63	0.41	0.60	0.45	0.31	0.17	0.63	0.51	1.18	1.00	0.99		0.70	0.53
H3K9me3	0.94	0.94	0.60	0.95		0.62	0.83	0.90	0.93	0.84	0.94	0.23	0.47	0.39	0.43	0.56	0.40	0.54	0.76	0.53	0.41	0.47	0.40	0.21	0.05	0.17	0.24	0.08	0.99	1.00		0.59	0.42
H3K27ac	0.98	0.97	0.87	0.97	0.96		0.93	1.02	0.95	0.96	0.96	0.43	0.73	0.67	0.87	0.72	0.67	0.73	0.63	0.74	0.55	0.67	0.62	0.49	0.49	0.90	0.90	0.24	1.01	1.00		0.79	0.73
H3K9ac	0.98	0.95	0.68	0.98	0.97	0.73		0.92	0.94	0.93	0.87	0.44	0.53	0.50	0.56	0.65	0.77	0.67	0.54	0.67	0.43	0.62	0.48	0.46	0.31	0.83	0.67	0.82	0.99	1.00		0.73	0.62
DNase	1.00	0.94	0.76	0.97	0.96	0.80	0.87		0.96	0.93	0.96	0.38	0.55	0.55	0.58	0.66	0.52	0.68	0.59	0.68	0.50	0.61	0.52	0.37	0.27	0.56	0.52	0.74	1.00	1.00		0.71	0.58
H3K4me2	1.00	0.96	0.79	1.00	0.98	0.73	0.96	0.93		0.97	0.90	0.50	0.56	0.49	0.54	0.64	0.66	0.63	0.65	0.67	0.46	0.62	0.49	0.39	0.33	0.70	0.62	1.01	0.99	1.00		0.74	0.59
H2A.Z	0.86	0.93	0.62	0.94	0.87	0.57	0.84	0.73	0.91		0.73	0.29	0.47	0.45	0.48	0.58	0.49	0.55	0.55	0.69	0.41	0.62	0.47	0.24	0.17	0.41	0.44	0.22	0.97	0.98		0.62	0.49
H3K79me2	0.95	0.92	0.52	0.95	0.90	0.57	0.85	0.82	0.89	0.73		1.03	0.40	0.28	0.35	0.46	0.54	0.41	0.98	0.49	0.25	0.41	0.26	0.21	0.12	0.33	0.28	0.17	0.98	0.97		0.60	0.39
H4K20me1	0.76	0.87	0.35	0.92	0.72	0.42	0.72	0.62	0.85	0.60	0.84		0.29	0.14	0.24	0.30	0.33	0.16	0.66	0.18	0.18	0.24	0.17	0.14	0.09	0.32	0.25	0.34		0.94		0.47	0.26
H2AK5ac	0.99	0.98	0.87	0.97	0.99	0.79	0.87	0.96	0.95	0.90	0.98	0.34		0.71	0.74	0.73	0.68	0.78	0.79	0.73	0.74	0.72	0.69	0.78	0.67	0.60	0.59	0.21	0.99	0.99		0.79	0.74
H2BK120ac	0.96	0.95	0.77	0.97	0.95	0.80	0.88	0.91	0.95	0.91	0.96	0.34	0.83		0.79	0.87	0.74	0.98	0.71	0.86	0.82	0.86	0.87	0.75	0.90	0.75	0.70	0.20	1.00	0.99		0.83	0.84
H2BK5ac	0.97	0.98	0.86	0.97	0.97	0.98	0.90	0.97	0.96	0.93	0.96	0.40	0.86	0.77		0.79	0.75	0.82	0.73	0.80	0.66	0.76	0.73	0.65	0.63	0.93	0.82	0.29	1.00	0.99		0.83	0.80
H3K18ac	0.97	0.96	0.82	0.97	0.97	0.83	0.90	0.96	0.97	0.92	0.87	0.33	0.83	0.89	0.81		0.86	1.04	0.60	0.93	0.83	0.93	0.86	0.83	0.80	0.87	0.91	0.52	0.99	1.00		0.87	0.88
H3K23ac	0.94	0.93	0.67	0.96	0.96	0.66	0.90	0.89	0.93	0.88	0.93	0.49	0.58	0.49	0.52	0.65		0.68	0.78	0.70	0.49	0.61	0.51	0.56	0.45	0.85	0.60	0.28	0.99	0.99		0.73	0.64
H3K4ac	0.94	0.95	0.76	0.96	0.94	0.77	0.89	0.94	0.94	0.94	0.90	0.30	0.77	0.81	0.70	0.84	0.79		0.60	0.84	0.73	0.83	0.80	0.75	0.74	0.76	0.68	0.27	0.99	0.99		0.80	0.79
H3K79me1	0.91	0.97	0.69	0.97	0.97	0.62	0.86	0.89	0.92	0.87	1.04	0.86	0.54	0.40	0.44	0.55	0.52	0.57		0.58	0.43	0.55	0.41	0.29	0.13	0.26	0.27	0.15	0.99	0.99		0.65	0.46
H4K8ac	0.91	0.98	0.77	0.96	0.94	0.73	0.89	0.95	0.93	0.97	0.94	0.30	0.73	0.73	0.69	0.78	0.82	0.88	0.69		0.71	0.80	0.79	0.61	0.48	0.92	0.83	0.25	0.99	1.00		0.80	0.77
H2BK12ac	0.90	0.99	0.86	0.96	0.93	0.73	0.81	1.01	0.95	0.90	0.76	0.30	0.93	0.89	0.73	0.87	0.73	0.88	0.70	0.85		0.85	0.88	0.88	0.94	0.69	0.66	0.26	0.99	0.98		0.83	0.83
H3K14ac	0.84	0.95	0.77	0.96	0.91	0.69	0.85	0.82	0.94	0.93	0.83	0.36	0.80	0.77	0.72	0.81	0.89	0.88	0.78	0.88	0.73		0.80	0.69	0.63	0.76	0.65	0.44	0.98	0.97		0.80	0.79
H4K91ac	0.91	0.98	0.83	0.96	0.91	0.77	0.83	0.91	0.95	0.93	0.77	0.33	0.85	0.90	0.81	0.85	0.77	0.94	0.71	0.92	0.86	0.85		0.76	0.85	0.78	0.82	0.26	0.99	0.98		0.83	0.85
H2BK15ac	0.85	0.92	0.69	0.96	0.93	0.69	0.82	0.93	0.92	0.61	0.99	0.26	0.82	0.72	0.60	0.74	0.68	0.86	0.82	0.68	0.76	0.73	0.70		0.75	0.61	0.51	0.30	1.01	0.99		0.76	0.73
H2BK20ac	0.69	0.85	0.54	0.94	0.79	0.58	0.74	0.71	0.88	0.60	0.76	0.36	0.76	0.91	0.58	0.68	0.58	0.76	0.66	0.49	0.79	0.64	0.76	0.70		0.65	0.59	0.15		0.94		0.69	0.70
H3K56ac	0.73	0.86	0.41	0.91	0.74	0.66	0.81	0.59	0.82	0.70	0.77	0.39	0.52	0.54	0.65	0.55	0.84	0.62	0.61	0.78	0.44	0.63	0.59	0.50	0.47		0.73	0.71		0.91		0.67	0.64
H4K5ac	0.72	0.85	0.61	0.92	0.76	0.93	0.84	0.63	0.86	0.82	0.82	0.31	0.67	0.72	0.81	0.78	0.84	0.78	0.65	0.92	0.57	0.72	0.76	0.58	0.60	0.94		0.27		0.95		0.75	0.78
H3K23me2	0.23	0.05	0.15	0.58	0.09	0.17	0.44	0.34	0.63	0.44	0.10	0.34	0.22	0.17	0.21	0.35	0.49	0.27	0.13	0.28	0.22	0.44	0.21	0.32	0.17	0.63	0.27			0.87		0.34	0.30
H3K18ac +H3K79me2	0.97	0.97	0.80	0.97	0.96	0.83	0.91	0.95	0.97	0.92	*0.87*	1.02	0.84	0.87	0.82	*0.46*	0.92	0.95	0.89	0.94	0.80	0.96	0.85	0.83	0.80	0.89	0.92	0.49	0.99	1.00		0.90	0.88
H3K18ac + H3K79me2 - Same Sample Features Only	0.05	0.10	0.63	0.50	0.30	0.76	0.69	0.52	0.53	0.68	*0.24*	1.00	0.79	0.87	0.81	*0.14*	0.88	0.95	0.91	0.89	0.78	0.95	0.84	0.83	0.80	0.88	0.92	0.49	0.06	0.89		0.71	0.85

b

Relative Coefficient of Determination

Mark/Feature Set	H3K27me3	H3K36me3	H3K4me1	H3K4me3	H3K9me3	H3K27ac	H3K9ac	DNase	H3K4me2	H2A.Z	H3K79me2	H4K20me1	H2AK5ac	H2BK120ac	H2BK5ac	H3K18ac	H3K23ac	H3K4ac	H3K79me1	H4K8ac	H2BK12ac	H3K14ac	H4K91ac	H2BK15ac	H2BK20ac	H3K56ac	H4K5ac	H3K23me2	RNA-seq	DNA Methylation		All Marks	Acetylations Only
All Tier 1-3 Features	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00		1.00	1.00
Same Sample Features Only	0.18	0.31	0.91	0.87	0.22	0.95	0.96	0.45	0.88	0.88	0.71	1.00	0.99	1.00	1.00	0.99	0.99	1.00	0.88	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	0.12	0.76		0.83	0.99
Same Mark Features Only	0.98	0.92	0.64	0.96	0.91	0.67	0.89	0.97	0.92	0.78	0.84	0.07	0.47	0.42	0.39	0.62	0.36	0.58	0.67	0.60	0.47	0.50	0.43	0.24	0.03	0.10	0.16		1.00	0.99		0.61	0.43
Core + H3K18ac	*0.98*	*0.97*	*0.85*	*0.96*	*1.03*	0.78	0.99	0.99	0.99	1.01	0.91	0.56	0.85	0.86	0.72	*0.65*	0.86	1.01	0.74	0.91	0.86	0.95	0.81	0.82	0.78	0.94	0.87	0.97	1.00	1.00		0.91	0.88
Core + H3K27ac	*0.99*	*0.98*	*0.87*	*0.95*	*1.01*	*0.67*	0.95	1.02	1.00	1.01	0.91	0.54	0.77	0.58	0.86	0.69	0.74	0.64	0.76	0.72	0.59	0.65	0.55	0.54	0.46	0.81	0.83	0.99	1.00	0.99		0.82	0.71
Core + H3K9ac	*0.97*	*0.97*	*0.69*	*0.99*	*1.01*	0.70	*0.95*	0.97	0.99	1.02	0.90	0.56	0.72	0.47	0.53	0.67	0.80	0.57	0.78	0.67	0.56	0.63	0.46	0.56	0.36	0.78	0.67	0.90	1.00	0.99		0.78	0.63
Core + DNase	*0.98*	*0.97*	*0.75*	*0.94*	*1.02*	0.75	0.95	*0.99*	0.99	1.01	0.90	0.54	0.71	0.48	0.54	0.66	0.70	0.58	0.77	0.70	0.56	0.63	0.49	0.51	0.32	0.51	0.53	0.90	1.00	1.00		0.76	0.60
Core (H3K27me3, H3K4me1, H3K4me3, H3K9me3, H3K36me3)	*0.96*	*0.97*	*0.65*	*0.94*	*1.01*	0.67	0.95	0.99	0.99	1.01	0.90	0.55	0.71	0.46	0.52	0.65	0.70	0.56	0.77	0.67	0.55	0.63	0.47	0.50	0.31	0.51	0.52	0.94	1.00	0.99		0.75	0.59
Core + H3K18ac - Same Sample Features Only	*0.16*	*0.15*	*0.65*	*0.50*	*0.30*	0.61	0.93	0.36	0.85	0.88	0.42	0.56	0.82	0.84	0.70	*0.52*	0.82	1.00	0.51	0.88	0.84	0.92	0.80	0.82	0.78	0.94	0.87	0.97	0.13	0.70		0.80	0.85
Core + H3K27ac - Same Sample Features Only	*0.19*	*0.17*	*0.65*	*0.40*	*0.27*	*0.40*	0.91	0.42	0.87	0.85	0.46	0.53	0.69	0.54	0.84	0.61	0.67	0.56	0.53	0.61	0.53	0.58	0.50	0.51	0.46	0.80	0.82	0.99	0.13	0.71		0.70	0.66
Core + H3K9ac - Same Sample Features Only	*0.18*	*0.18*	*0.40*	*0.78*	*0.27*	0.48	*0.83*	0.33	0.87	0.85	0.44	0.56	0.63	0.40	0.45	0.58	0.75	0.50	0.52	0.54	0.49	0.56	0.37	0.52	0.36	0.78	0.68	0.90	0.13	0.70		0.65	0.57
Core + DNase - Same Sample Features Only	*0.18*	*0.16*	*0.53*	*0.30*	*0.28*	0.56	0.83	*0.34*	0.88	0.86	0.44	0.52	0.61	0.41	0.49	0.54	0.63	0.48	0.53	0.55	0.48	0.56	0.40	0.48	0.32	0.50	0.51	0.90	0.13	0.71		0.64	0.52
Core - Same Sample Features Only	*0.17*	*0.15*	*0.27*	*0.02*	*0.27*	0.40	0.83	0.34	0.87	0.85	0.43	0.53	0.61	0.38	0.44	0.52	0.63	0.45	0.54	0.53	0.47	0.55	0.38	0.47	0.31	0.49	0.51	0.94	0.13	0.71		0.61	0.50
Tier 1 and 2 Marks	*1.01*	*1.00*	*0.93*	*1.01*	*1.03*	*0.74*	*0.96*	*1.02*	*0.99*	*0.99*	*0.89*	*0.97*	0.75	0.58	0.87	0.71	0.88	0.58	0.99	0.75	0.60	0.62	0.56	0.58	0.48	0.78	0.86	1.17	1.00	1.00		0.84	0.71
Tier 1 and 2 Marks - Same Sample Features Only	*0.18*	*0.19*	*0.86*	*0.91*	*0.22*	*0.63*	*0.88*	*0.43*	*0.90*	*0.85*	*0.46*	*0.97*	0.69	0.56	0.86	0.65	0.83	0.55	0.91	0.68	0.57	0.58	0.53	0.55	0.48	0.78	0.86	1.17	0.13	0.78		0.65	0.67
H3K27me3		0.85	0.54	0.94	0.94	0.56	0.89	0.97	0.88	0.73	0.81	0.17	0.40	0.28	0.30	0.49	0.43	0.42	0.65	0.49	0.34	0.43	0.29	0.25	0.04	0.10	0.09	0.16	0.99	0.99		0.55	0.36
H3K36me3	0.91		0.57	0.97	0.94	0.48	0.84	0.79	0.87	0.84	0.86	0.32	0.42	0.26	0.28	0.49	0.47	0.40	0.73	0.53	0.33	0.44	0.32	0.26	0.04	0.09	0.15	0.07	0.99	0.97		0.55	0.36
H3K4me1	0.97	0.95		0.94	0.90	0.64	0.88	0.98	0.91	0.89	0.89	0.31	0.66	0.43	0.46	0.60	0.58	0.54	0.72	0.62	0.52	0.55	0.45	0.46	0.32	0.24	0.37	0.07	1.00	0.97		0.66	0.52
H3K4me3	0.95	0.95	0.61		0.96	0.52	0.93	0.83	0.95	0.92	0.80	0.12	0.42	0.29	0.33	0.52	0.53	0.42	0.59	0.55	0.34	0.51	0.31	0.30	0.06	0.45	0.35	0.90	1.00	0.97		0.61	0.43
H3K9me3	0.94	0.87	0.51	0.93		0.49	0.87	0.88	0.85	0.71	0.79	0.10	0.40	0.23	0.27	0.46	0.42	0.38	0.63	0.42	0.32	0.41	0.27	0.26	0.04	0.09	0.10	0.08	0.98	0.97		0.52	0.34
H3K27ac	0.97	0.92	0.79	0.95	0.92		0.92	1.04	0.91	0.90	0.84	0.21	0.64	0.52	0.80	0.62	0.56	0.59	0.63	0.63	0.48	0.55	0.48	0.43	0.34	0.71	0.81	0.12	1.01	0.98		0.71	0.63
H3K9ac	0.98	0.89	0.61	1.00	0.93	0.61		0.91	0.91	0.84	0.80	0.32	0.48	0.34	0.42	0.57	0.67	0.52	0.58	0.58	0.37	0.52	0.34	0.43	0.19	0.75	0.55	0.61	0.99	0.98		0.66	0.52
DNase	0.97	0.86	0.66	0.93	0.90	0.68	0.87		0.89	0.83	0.80	0.12	0.43	0.36	0.39	0.55	0.37	0.51	0.55	0.56	0.40	0.46	0.35	0.27	0.12	0.24	0.32	0.47	1.00	0.99		0.59	0.43
H3K4me2	0.97	0.90	0.72	0.96	0.95	0.60	0.89	0.90		0.87	0.71	0.25	0.48	0.34	0.40	0.55	0.58	0.46	0.62	0.56	0.38	0.51	0.35	0.34	0.19	0.45	0.46	0.85	0.99	0.99		0.64	0.47
H2A.Z	0.76	0.89	0.53	0.93	0.83	0.40	0.84	0.67	0.86		0.48	0.11	0.38	0.29	0.32	0.47	0.40	0.41	0.49	0.62	0.34	0.51	0.33	0.13	0.08	0.14	0.24	0.06	0.93	0.92		0.51	0.37
H3K79me2	0.99	0.87	0.44	0.94	0.83	0.41	0.91	0.79	0.84	0.56		0.89	0.36	0.15	0.22	0.32	0.46	0.26	0.94	0.39	0.20	0.30	0.14	0.18	0.05	0.16	0.16	0.12	0.96	0.91		0.53	0.29
H4K20me1	0.61	0.64	0.22	0.91	0.60	0.23	0.83	0.55	0.78	0.34	0.68		0.16	0.05	0.12	0.18	0.22	0.08	0.54	0.10	0.10	0.15	0.07	0.11	0.05	0.15	0.14	0.25		0.74		0.37	0.17
H2AK5ac	1.01	0.94	0.82	0.97	0.95	0.71	0.92	0.95	0.92	0.89	0.88	0.29		0.62	0.65	0.68	0.64	0.67	0.75	0.67	0.70	0.65	0.62	0.73	0.60	0.44	0.44	0.20	0.99	0.97		0.74	0.67
H2BK120ac	0.96	0.89	0.71	0.94	0.90	0.71	0.91	0.91	0.91	0.86	0.82	0.13	0.75		0.67	0.83	0.63	0.88	0.67	0.78	0.76	0.78	0.79	0.64	0.75	0.50	0.52	0.08	1.00	0.98		0.76	0.74
H2BK5ac	0.98	0.93	0.79	0.96	0.92	0.96	0.95	0.97	0.92	0.90	0.84	0.26	0.78	0.66		0.71	0.66	0.69	0.69	0.73	0.60	0.65	0.65	0.60	0.54	0.76	0.71	0.19	0.99	0.98		0.77	0.73
H3K18ac	0.97	0.91	0.77	0.97	0.95	0.78	0.93	0.96	0.94	0.89	0.74	0.19	0.77	0.83	0.71		0.76	1.02	0.59	0.91	0.80	0.93	0.79	0.78	0.73	0.77	0.83	0.47	0.99	1.00		0.82	0.83
H3K23ac	0.91	0.88	0.59	0.95	0.91	0.55	0.95	0.85	0.89	0.82	0.83	0.42	0.54	0.37	0.39	0.60		0.56	0.72	0.62	0.44	0.58	0.40	0.54	0.32	0.69	0.43	0.26	0.98	0.97		0.67	0.56
H3K4ac	0.93	0.91	0.69	0.94	0.87	0.70	0.91	0.94	0.90	0.93	0.75	0.18	0.69	0.71	0.58	0.81	0.66		0.63	0.81	0.68	0.80	0.75	0.68	0.68	0.55	0.52	0.30	0.98	0.97		0.75	0.72
H3K79me1	0.89	0.94	0.60	0.92	0.91	0.48	0.86	0.87	0.87	0.78	0.96	0.72	0.50	0.26	0.31	0.46	0.49	0.42		0.52	0.36	0.48	0.30	0.29	0.06	0.10	0.14	0.12	0.98	0.98		0.59	0.38
H4K8ac	0.91	0.94	0.71	0.95	0.90	0.64	0.91	0.95	0.88	0.94	0.82	0.18	0.67	0.65	0.60	0.73	0.67	0.80	0.70		0.68	0.72	0.75	0.56	0.32	0.69	0.73	0.13	0.99	0.98		0.74	0.70
H2BK12ac	0.84	0.96	0.82	0.95	0.93	0.61	0.83	1.02	0.92	0.90	0.59	0.18	0.86	0.81	0.60	0.82	0.65	0.79	0.65	0.81		0.76	0.81	0.77	0.85	0.51	0.52	0.25	0.98	0.95		0.76	0.75
H3K14ac	0.77	0.91	0.72	0.94	0.87	0.54	0.87	0.77	0.90	0.91	0.69	0.29	0.74	0.69	0.61	0.80	0.80	0.84	0.73	0.86	0.69		0.74	0.65	0.49	0.59	0.51	0.65	0.97	0.92		0.75	0.71
H4K91ac	0.84	0.95	0.78	0.97	0.89	0.69	0.86	0.91	0.92	0.93	0.62	0.22	0.78	0.84	0.75	0.80	0.64	0.87	0.66	0.91	0.82	0.79		0.66	0.78	0.59	0.69	0.27	0.97	0.94		0.78	0.78
H2BK15ac	0.88	0.74	0.61	0.94	0.89	0.58	0.85	0.91	0.86	0.44	0.88	0.17	0.72	0.57	0.46	0.67	0.57	0.72	0.69	0.53	0.66	0.63	0.55		0.67	0.48	0.41	0.36	0.99	0.94		0.68	0.63
H2BK20ac	0.57	0.66	0.43	0.90	0.91	0.36	0.79	0.65	0.84	0.39	0.57	0.15	0.56	0.72	0.36	0.53	0.36	0.59	0.43	0.26	0.63	0.47	0.54	0.56		0.38	0.41	0.11		0.76		0.55	0.53
H3K56ac	0.61	0.70	0.26	0.92	0.69	0.51	0.87	0.56	0.76	0.55	0.61	0.22	0.29	0.31	0.36	0.36	0.47	0.41	0.43	0.52	0.24	0.43	0.33	0.25	0.23		0.55	0.72		0.71		0.51	0.44
H4K5ac	0.60	0.65	0.51	0.92	0.89	0.87	0.87	0.63	0.82	0.69	0.64	0.20	0.52	0.54	0.66	0.69	0.60	0.58	0.50	0.75	0.47	0.50	0.57	0.50	0.49	0.83		0.19		0.77		0.64	0.65
H3K23me2	0.04	0.04	0.10	0.15	0.03	0.04	0.09	0.09	0.33	0.19	0.08	0.28	0.20	0.12	0.13	0.18	0.16	0.23	0.11	0.13	0.17	0.30	0.18	0.29	0.17	0.38	0.16			0.42		0.20	0.18
H3K18ac +H3K79me2	0.97	0.94	0.78	0.98	0.96	0.78	0.97	0.95	0.95	0.90	*0.74*	0.92	0.80	0.83	0.71	*0.32*	0.81	0.92	0.83	0.91	0.81	0.93	0.79	0.81	0.72	0.79	0.84	0.46	0.99	1.00		0.87	0.84
H3K18ac +H3K79me2 - Same Sample Features Only	0.02	0.12	0.58	0.37	0.21	0.63	0.53	0.29	0.32	0.57	0.16	0.94	0.72	0.82	0.71	0.08	0.76	0.92	0.77	0.83	0.73	0.95	0.76	0.77	0.73	0.80	0.83	0.46	0.01	0.38		0.62	0.78

Supplementary Figure 42: Feature Subset Performance for Seven Samples with Deep Mark Coverage.
(a) An extension of the table shown in **Fig. 6a**. The table reports for various feature subsets (rows) the relative average imputation performance for different marks (columns) compared to an imputation conducted using all features. The imputation was done based on only data for the seven samples with deep mark coverage and making no distinctions between Tier-1-3 marks. The feature subsets include using only same sample features as the target, only same mark features from other samples, and all features that can be computed for various subsets of the mark subsets specified, or restricted to same sample features for the mark subset if indicated. The imputation performance is computed based on the agreement in top 1% signal on chr10, except for the DNA methylation which is based on the % of predictions within 0.25 of the observed data on chr10. In italics are those cases in which the target mark is also a mark of the row, but was not used for its imputation. The last two columns show the average performance of the feature subset over all target marks and specifically for acetylations. For the purpose of computing these averages for mark subsets if the target mark is part of the mark subset of the row then a value of 1 is used for the target mark. The mark subset evaluations involving H3K79me2 are limited to five samples where H3K79me2 is present. In gray are values which could not be predicted based on the method. **(b)** The same as **a** except based on the coefficient of determination, which was used here instead of the correlation coefficient so the relative agreement would be meaningful.



a

	State	H2BK15ac	H2BK20ac	H2AK5ac	H2BK120ac	H4K91ac	H2BK12ac	H3K4ac	H3K14ac	H4K8ac	H4K5ac	H3K18ac	H2BK5ac	H3K56ac	H3K23ac	H3K23me2	H3K9me1	H3K79me1	H3K9ac	H3K4me2	H3K4me3	H2A.Z	DNase	H3K27ac	H3K4me1	H3K79me2	H4K20me1	H3K36me3	H3K9me3	H3K27me3
1_TssA	1	0	0	0	0	0	0	1	2	15	18	9	1	1	0	13	0	0	99	99	100	82	97	96	1	32	0	0	0	0
	2	2	24	18	90	91	64	98	88	99	100	100	87	57	30	51	0	6	100	100	100	96	96	99	53	23	0	0	0	9
2_PromU	3	0	0	0	0	0	0	0	0	8	8	4	0	0	0	27	0	0	91	100	100	87	71	69	61	2	0	0	1	1
	4	0	0	0	4	9	4	27	31	78	81	85	31	13	3	43	0	0	99	100	100	98	93	99	60	12	0	0	0	9
	5	6	17	15	42	61	54	77	57	85	76	96	11	7	18	49	0	10	72	99	93	78	54	70	89	19	3	1	3	69
	6	0	0	0	0	0	0	2	4	12	8	11	0	1	0	57	0	0	97	100	100	93	67	43	77	15	0	0	1	98
3_PromD1	7	0	0	1	6	14	4	61	65	86	76	90	36	46	15	58	0	31	100	100	100	91	89	100	35	99	8	2	0	5
	8	3	11	43	93	95	64	100	97	100	100	100	93	88	65	63	0	71	100	100	100	94	96	100	28	99	7	2	0	3
	9	82	97	99	100	100	100	100	99	100	100	100	100	93	86	82	0	57	100	100	99	93	96	99	68	51	2	0	0	10
4_PromD2	10	0	0	0	0	0	0	8	0	0	0	0	0	0	0	3	0	10	22	90	82	1	4	9	74	95	20	12	1	3
	11	0	0	0	0	0	0	8	1	6	4	6	0	1	0	41	0	12	99	100	100	49	40	83	70	97	9	1	0	6
5_Tx5'	12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	82	2	1	0	0
6_Tx	13	0	0	0	0	0	0	15	0	0	1	1	0	1	0	0	0	39	0	3	0	0	1	2	17	99	78	98	1	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	66	7	90	0	0
7_Tx3'	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0
8_TxWk	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9_TxReg	19	61	72	94	97	99	97	100	93	98	99	100	94	69	53	68	4	92	97	98	68	14	82	100	98	98	47	28	0	4
	20	0	0	2	1	5	3	48	4	10	17	26	9	10	1	31	1	50	79	100	93	2	42	90	95	100	75	40	0	3
10_TxEnh5'	21	52	75	94	97	100	99	100	75	87	94	100	77	26	20	28	1	86	12	35	2	1	25	79	91	91	27	31	1	1
	22	0	0	0	0	0	0	31	0	0	3	6	1	6	0	3	1	52	1	32	4	0	24	54	97	100	82	32	0	1
	23	0	7	22	22	58	50	53	18	28	40	90	14	3	0	7	0	76	5	28	3	1	18	52	81	92	39	29	1	0
11_TxEnh3'	24	0	0	0	0	0	1	1	0	0	3	3	1	0	0	1	0	7	1	6	1	0	20	31	62	11	11	86	0	0
12_TxEnhW	25	0	0	0	0	0	0	2	0	0	1	2	0	0	0	0	0	10	0	8	0	0	16	24	84	81	6	2	0	0
	26	0	0	0	0	0	0	21	0	0	0	0	0	2	0	0	0	19	0	1	0	0	2	0	20	100	83	3	0	0
13_EnhA1	27	5	40	51	84	88	88	83	50	69	97	99	86	19	2	18	0	5	76	86	36	46	86	100	100	12	1	3	0	1
	28	0	0	0	2	3	4	3	0	5	48	31	35	1	0	9	0	2	62	78	35	28	85	97	99	15	2	5	0	1
	29	91	100	100	100	100	100	100	91	100	100	100	99	73	49	47	0	25	46	56	5	32	61	93	99	3	0	3	1	2
15_EnhAF	30	0	0	0	0	0	0	0	0	0	2	3	0	0	0	1	0	0	4	18	3	10	46	96	98	1	0	0	0	0
	31	17	67	63	84	98	96	87	44	53	78	99	52	4	1	8	0	3	4	21	1	14	31	64	92	1	0	3	1	3
	32	0	8	5	11	25	39	20	3	10	45	70	24	1	0	4	0	2	4	20	1	14	31	48	93	1	0	2	1	2
16_EnhW1	33	0	0	0	0	0	0	0	0	2	4	3	0	0	0	6	0	0	4	90	57	54	46	33	85	1	0	0	0	2
	34	0	0	0	0	0	0	0	0	1	5	3	0	0	0	1	0	0	0	21	0	53	23	3	38	0	0	0	1	1
17_EnhW2	35	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	79	0	0	0	0	0
	36	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	7	0	6	100	0	80	0	0	0	0	0
18_EnhAc	37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	27	55	10	0	0	0	1	0
	38	77	97	94	99	100	100	99	63	81	82	100	72	19	17	25	0	7	0	5	0	3	5	11	27	1	0	3	2	2
	39	5	50	18	27	64	75	45	8	6	17	84	5	0	0	6	0	2	0	3	0	1	4	4	8	1	0	3	2	3
	40	0	0	0	0	0	2	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0
19_DNase	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	96	0	2	0	0	0	0	0
20_ZNF/Rpts	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	7	0	88	87	0
	43	0	0	0	0	1	1	15	0	0	1	3	0	0	0	0	0	41	0	1	0	0	1	1	7	91	7	15	10	0
21_Het	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	71	0
22_PromP	45	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	17	46	80	9	28	9	7	2	0	1	10	1
23_PromBiv	46	0	0	0	0	0	0	1	0	1	1	2	0	0	0	16	0	0	13	98	92	47	38	11	93	5	1	0	2	99
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	6	58	71	15	16	3	21	2	0	0	6	92
	48	0	0	0	0	0	1	1	0	0	1	2	0	0	0	1	0	0	0	12	4	2	14	5	70	7	2	2	1	97
24_ReprPC	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	78
25_Quies	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

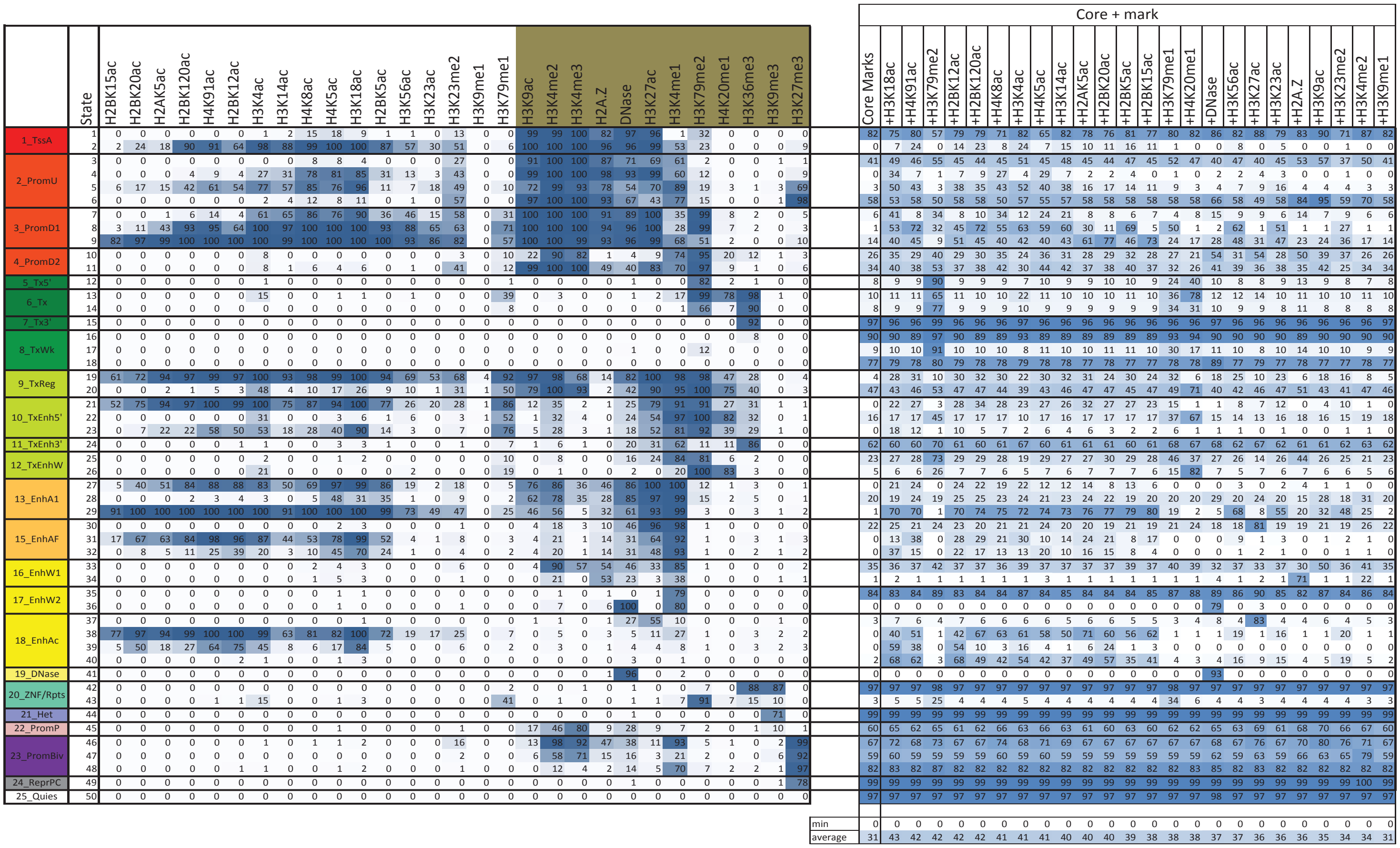
	State	All other Marks																												
		DNase	H3K9me3	H3K36me3	H3K4me1	H3K27me3	H3K27ac	H2A.Z	H3K79me2	H4K20me1	H3K79me1	H3K4me3	H3K4me2	H3K9ac	H3K18ac	H2BK5ac	H3K4ac	H4K8ac	H2BK12ac	H4K5ac	H4K91ac	H2BK15ac	H2BK20ac	H3K23me2	H3K23ac	H2BK120ac	H3K14ac	H2AK5ac	H3K56ac	H3K9me1
1_TssA	1	91	100	100	73	98	94	96	94	99	98	98	99	97	95	98	97	96	99	95	99	100	100	95	99	99	97	100	98	100
	2	98	100	99	96	96	98	96	88	99	91	95	99	97	99	95	96	97	97	98	94	94	94	96	93	93	95	93	93	100
2_PromU	3	94	99	100	89	81	92	92	88	100	99	93	95	84	95	98	99	96	99	95	99	100	100	96	100	99	99	100	99	100
	4	96	100	100	94	95	96	95	86	99	97	96	99	97	83	95	93	87	96	89	94	99	98	97	98	94	94	98	96	100
	5	91	99	99	92	68	92	94	97	99	98	91	96	94	92	86	90	92	92	94	92	99	97	98	97	94	93	97	94	100
	6	98	100	100	98	66	97	95	99	100	100	98	99	80	99	99	99	99	99	98	99	100	100	97	100	99	99	100	100	100
3_PromD1	7	97	100	99	96	99	99	97	80	99	96	99	100	99	92	97	96	92	96	95	94	99	99	97	96	94	93	98	95	100
	8	99	100	99	94	99	99	98	84	99	94	99	100	99	99	97	98	99	97	99	96	93	93	98	96	95	98	97	96	100
	9	99	100	99	98	99	99	96	97	98	97	92	99	98	100	99	100	99	98	100	99	93	92	98	97	99	99	94	97	100
4_PromD2	10	95	100	97	96	99	86	96	82	95	97	86	86	94	98	99	98	99	100	99	100	100	100	98	100	100	100	100	99	100
	11	93	100	96	92	98	92	92	76	94	97	97	98	91	96	98	96	97	99	96	99	100	100	92	99	99	98	100	98	100
5_Tx5'	12	100	99	92	96	100	99	100	43	93	97	100	99	100	100	100	98	100	100	100	100	100	100	100	100	100	100	100	100	100
6_Tx	13	98	99	33	92	100	96	100	86	67	93	99	98	100	99	100	96	100	100	100	100	100	100	100	100	100	100	100	100	100
	14	99	98	23	95	100	99	100	48	88	95	100	99	100	100	100	98	100	100	100	100	100	100	100	100	100	100	100	100	100
7_Tx3'	15	100	99	9	99	100	100	100	96	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
8_TxWk	16	99	99	16	98	99	99	100	95	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	17	99	99	82	96	99	99	100	38	97	95	100	100	100	100	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100
	18	88	91	89	85	90	99	98	98	100	100	99	99	100	99	100	100	100	99	99	100	100	100	100	100	100	100	100	100	100
9_TxReg	19	95	100	96	94	99	98	85	82	93	93	95	96	92	99	98	99	98	98	99	99	98	98	97	97	99	98	97	96	100
	20	96	100	90	93	99	90	89	93	80	92	85	94	91	97	98	90	98	99	97	98	100	99	97	99	99	98	99	98	100
10_TxEnh5'	21	96	100	94	98	100	97	97	73	97	92	95	95	92	99	97	97	97	97	97	98	98	96	97	98	95	96	96	97	100
	22	94	100	91	73	100	86	100	90	74	87	94	88	97	96	99	92	99	98	98	98	100	100	99	100	99	99	99	99	100
	23	98	100	95	96	100	96	99	80	94	90	97	97	98	81	95	95	95	92	93	89	98	97	98	99	96	96	95	99	100
11_TxEnh3'	24	92	99	25	65	99	90	99	85	91	97	99	97	99	98	100	98	100	99	99	99	100	100	100	100	100	100	100	100	100
12_TxEnhW	25	95	100	92	49	99	94	99	56	86	95	96	96	99	98	99	97	100	99	99	99	100	100	100	100	100	100	100	100	100
	26	98	99	86	96	100	94	100	95	44	96	99	98	100	99	100	96	100	100	100	100	100	100	100	100	100	100	100	99	100
13_EnhA1	27	90	100	99	93	97	92	91	96	98	95	93	90	88	97	93	96	94	94	96	95	90	93	94	94	95	95	94	93	100
	28	87	100	97	91	99	84	90	93	97	98	89	84	82	93	93	96	96	96	92	97	99	99	96	99	98	98	98	99	100
	29	97	100	99	94	100	94	98	96	99	98	92	97	98	100	96	99	97	99	98	99	93	96	97	96	99	96	97	95	100
15_EnhAF	30	96	100	98	76	100	37	98	97	100	100	96	95	97	98	98	99	99	99	97	99	100	100	99	100	100	100	100	100	100
	31	93	100	99	84	99	89	97	96	99	98	97	94	93	97	93	94	93	95	95	93	91	92	97	97	93	94	93	94	100
	32	93	100	98	79	99	90	97	96	99	99	96	94	95	80	91	92	96	88	89	91	99	94	99	100	94	97	96	99	100
16_EnhW1	33	93	98	99	75	81	88	82	93	99	100	81	69	79	98	97	99	98	99	97	99	100	100	97	100	99	99	100	100	100
	34	93	98	98	80	96	93	31	99	100	100	87	81	99	97	99	99	99	99	97	99	100	100	99	100	100	100	100	100	100
17_EnhW2	35	95	100	98	14	98	96	98	97	100	100	99	99	100	99	100	100	100	100	99	100	100	100	100	100	100	100	100	100	100
	36	14	100	98	47	98	80	97	97	100	100	97	97	99	99	100	100	100	99	99	100	100	100	100	100	100	100	100	100	100
18_EnhAc	37	91	99	97	56	100	18	98	98	100	99	98	97	99	98	99	99	100	99	99	99	100	99	100	100	100	100	100	100	100
	38	96	100	100	91	100	90	98	99	100	99	100	97	97	99	97	97	96	99	97	99	93	97	98	99	95	97	95	98	100
	39	98	100	100	88	100	95	99	98	100	99	100	99	100	86	96	95	97	87	95	93	98	93	99	100	96	98	96	99	100
	40	95	99	99	88	99	95	98	98	100	99	100	99	100	88	99	97	99	91	97	97	99	97	98	100	99	99	99	100	100
19_DNase	41	0	99	98	87	99	97	98	99	100	100	99	99	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
20_ZNF/Rpts	42	100	6	31	100	100	100	100	98	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	43	98	88	78	88	100	97	100	61	76	56	99	98	100	97	100	86	100	99	99	99	100	100	100	100	100	100	100	100	100
21_Het	44	100	0	93	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
22_PromP	45	93	97	98	77	84	93	89	93	99	99	59	88	92	99	100	100	99	100	99	100	100	100	98	100	100	100	100	100	100
23_PromBiv	46	95	99	100	89	75	97	91	99	100	100	95	93	86	99	100	100	99	100	99	100	100	100	95	100	100	100	100	100	100
	47	93	96	100	61	61	97	90	98	99	100	65	77	95	99	100	100	99	100	99	100	100	100	96	100	100	100	100	100	100
	48	96	99	98	51	34	98	97	95	99	100	88	94	99	99	100	100	100	100	100	100	100	100	99	100	100	100	100	100	100
24_ReprPC	49	100	100	100	98	15	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
25_Quies	50	98	97	99	97	96	100	99	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
min		0	0	9	14	15	18	31	38	44	56	59	69	79	80	86	86	87	87	89	89	90	92	92	93	93	93	93	93	100
average		92	95	88	85	92	92	95	88	95	96	95	95	96	97	98	97	98	98	98	98	99	98	98	99	98	98	99	99	100

b

	State	H2BK15ac	H2BK20ac	H2AK5ac	H2BK120ac	H4K91ac	H2BK12ac	H3K4ac	H3K14ac	H4K8ac	H4K5ac	H3K18ac	H2BK5ac	H3K56ac	H3K23ac	H3K23me2	H3K9me1	H3K79me1	H3K9ac	H3K4me2	H3K4me3	H2A.Z	DNase	H3K27ac	H3K4me1	H3K79me2	H4K20me1	H3K36me3	H3K9me3	H3K27me3
1_TssA	1	0	0	0	0	0	0	1	2	15	18	9	1	1	0	13	0	0	99	99	100	82	97	96	1	32	0	0	0	0
	2	2	24	18	90	91	64	98	88	99	100	100	87	57	30	51	0	6	100	100	100	96	96	99	53	23	0	0	0	9
2_PromU	3	0	0	0	0	0	0	0	0	8	8	4	0	0	0	27	0	0	91	100	100	87	71	69	61	2	0	0	1	1
	4	0	0	0	4	9	4	27	31	78	81	85	31	13	3	43	0	0	99	100	100	98	93	99	60	12	0	0	0	9
	5	6	17	15	42	61	54	77	57	85	76	96	11	7	18	49	0	10	72	99	93	78	54	70	89	19	3	1	3	69
	6	0	0	0	0	0	0	2	4	12	8	11	0	1	0	57	0	0	97	100	100	93	67	43	77	15	0	0	1	98
3_PromD1	7	0	0	1	6	14	4	61	65	86	76	90	36	46	15	58	0	31	100	100	100	91	89	100	35	99	8	2	0	5
	8	3	11	43	93	95	64	100	97	100	100	100	93	88	65	63	0	71	100	100	100	94	96	100	28	99	7	2	0	3
	9	82	97	99	100	100	100	100	99	100	100	100	100	93	86	82	0	57	100	100	99	93	96	99	68	51	2	0	0	10
4_PromD2	10	0	0	0	0	0	0	8	0	0	0	0	0	0	0	3	0	10	22	90	82	1	4	9	74	95	20	12	1	3
	11	0	0	0	0	0	0	8	1	6	4	6	0	1	0	41	0	12	99	100	100	49	40	83	70	97	9	1	0	6
5_Tx5'	12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	82	2	1	0	0
6_Tx	13	0	0	0	0	0	0	15	0	0	1	1	0	1	0	0	0	39	0	3	0	0	1	2	17	99	78	98	1	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	66	7	90	0	0
7_Tx3'	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0
8_TxWk	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9_TxReg	19	61	72	94	97	99	97	100	93	98	99	100	94	69	53	68	4	92	97	98	68	14	82	100	98	98	47	28	0	4
	20	0	0	2	1	5	3	48	4	10	17	26	9	10	1	31	1	50	79	100	93	2	42	90	95	100	75	40	0	3
10_TxEnh5'	21	52	75	94	97	100	99	100	75	87	94	100	77	26	20	28	1	86	12	35	2	1	25	79	91	91	27	31	1	1
	22	0	0	0	0	0	0	31	0	0	3	6	1	6	0	3	1	52	1	32	4	0	24	54	97	100	82	32	0	1
	23	0	7	22	22	58	50	53	18	28	40	90	14	3	0	7	0	76	5	28	3	1	18	52	81	92	39	29	1	0
11_TxEnh3'	24	0	0	0	0	0	1	1	0	0	3	3	1	0	0	1	0	7	1	6	1	0	20	31	62	11	11	86	0	0
12_TxEnhW	25	0	0	0	0	0	0	2	0	0	1	2	0	0	0	0	0	10	0	8	0	0	16	24	84	81	6	2	0	0
	26	0	0	0	0	0	0	21	0	0	0	0	0	2	0	0	0	19	0	1	0	0	2	0	20	100	83	3	0	0
13_EnhA1	27	5	40	51	84	88	88	83	50	69	97	99	86	19	2	18	0	5	76	86	36	46	86	100	100	12	1	3	0	1
	28	0	0	0	2	3	4	3	0	5	48	31	35	1	0	9	0	2	62	78	35	28	85	97	99	15	2	5	0	1
	29	91	100	100	100	100	100	100	91	100	100	100	99	73	49	47	0	25	46	56	5	32	61	93	99	3	0	3	1	2
15_EnhAF	30	0	0	0	0	0	0	0	0	0	2	3	0	0	0	1	0	0	4	18	3	10	46	96	98	1	0	0	0	0
	31	17	67	63	84	98	96	87	44	53	78	99	52	4	1	8	0	3	4	21	1	14	31	64	92	1	0	3	1	3
	32	0	8	5	11	25	39	20	3	10	45	70	24	1	0	4	0	2	4	20	1	14	31	48	93	1	0	2	1	2
16_EnhW1	33	0	0	0	0	0	0	0	0	2	4	3	0	0	0	6	0	0	4	90	57	54	46	33	85	1	0	0	0	2
	34	0	0	0	0	0	0	0	0	1	5	3	0	0	0	1	0	0	0	21	0	53	23	3	38	0	0	0	1	1
17_EnhW2	35	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	79	0	0	0	0	0
	36	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	7	0	6	100	0	80	0	0	0	0	0
18_EnhAc	37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	27	55	10	0	0	0	1	0
	38	77	97	94	99	100	100	99	63	81	82	100	72	19	17	25	0	7	0	5	0	3	5	11	27	1	0	3	2	2
	39	5	50	18	27	64	75	45	8	6	17	84	5	0	0	6	0	2	0	3	0	1	4	4	8	1	0	3	2	3
	40	0	0	0	0	0	2	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0
19_DNase	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	96	0	2	0	0	0	0	0
20_ZNF/Rpts	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	7	0	88	87	0
	43	0	0	0	0	1	1	15	0	0	1	3	0	0	0	0	0	41	0	1	0	0	1	1	7	91	7	15	10	0
21_Het	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	71	0
22_PromP	45	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	17	46	80	9	28	9	7	2	0	1	10	1
23_PromBiv	46	0	0	0	0	0	0	1	0	1	1	2	0	0	0	16	0	0	13	98	92	47	38	11	93	5	1	0	2	99
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	6	58	71	15	16	3	21	2	0	0	6	92
	48	0	0	0	0	0	1	1	0	0	1	2	0	0	0	1	0	0	0	12	4	2	14	5	70	7	2	2	1	97
24_ReprPC	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	78
25_Quies	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	State	Tier 1 and Tier2 + mark																	
		Tier 1 and Tier 2	+H3K18ac	+H2BK12ac	+H4K91ac	+H2BK120ac	+H2BK20ac	+H3K4ac	+H2AK5ac	+H4K5ac	+H2BK15ac	+H4K8ac	+H2BK5ac	+H3K14ac	+H3K23me2	+H3K23ac	+H3K56ac	+H3K79me1	+H3K9me1
1_TssA	1	69	87	71	73	71	68	83	70	76	67	80	76	83	74	72	77	76	69
	2	3	31	34	55	54	18	56	19	30	17	31	45	46	8	16	29	9	3
2_PromU	3	70	87	74	76	75	72	80	72	85	72	83	81	80	70	73	76	72	70
	4	21	45	30	30	33	29	21	27	39	26	39	19	23	23	28	24	26	21
	5	9	48	44	50	40	23	54	21	37	16	41	18	41	9	20	13	13	9
	6	93	93	94	95	94	94	94	93	93	93	92	93	93	93	93	93	93	93
3_PromD1	7	36	50	69	60	68	48	36	66	39	40	45	41	38	34	55	37	46	36
	8	50	60	56	84	83	48	71	48	65	53	62	76	66	41	58	69	56	50
	9	30	37	48	42	42	72	39	59	36	74	35	42	39	38	50	46	35	30
4_PromD2	10	90	92	91	91	91	90	91	91	92	90	91	91	91	92	90	91	91	90
	11	65	81	68	70	68	67	75	67	80	67	79	72	78	68	68	73	69	65
5_Tx5'	12	95	95	95	95	95	95	96	95	95	95	95	95	95	95	95	95	97	95
6_Tx	13	90	91	90	91	90	90	92	90	90	90	90	90	90	90	90	90	94	90
	14	92	93	93	93	92	92	94	92	92	92	92	92	92	92	92	92	97	92
7_Tx3'	15	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
8_TxWk	16	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
	17	94	95	94	95	94	94	95	94	94	94	94	94	94	94	94	94	97	94
	18	95	97	97	96	96	96	96	96	96	96	96	96	96	96	95	95	95	95
9_TxReg	19	62	79	84	84	85	74	74	85	80	73	80	80	82	68	73	72	72	63
	20	71	78	82	83	82	79	79	82	75	78	80	76	80	71	76	73	77	71
10_TxEnh5'	21	33	57	66	65	76	67	59	73	63	61	65	63	62	40	44	41	44	33
	22	73	80	79	80	78	77	76	78	77	76	78	77	78	75	75	74	80	73
	23	4	37	27	34	19	11	10	22	15	8	21	10	14	4	5	5	9	4
11_TxEnh3'	24	89	93	92	92	91	91	92	91	92	90	91	91	91	90	90	90	92	89
12_TxEnhW	25	85	90	90	90	89	88	90	89	89	88	89	88	89	86	86	86	90	85
	26	93	93	93	93	93	93	94	93	93	93	93	93	93	93	93	93	94	93
13_EnhA1	27	18	47	50	49	46	26	45	33	36	29	38	36	28	20	22	19	21	18
	28	44	58	71	72	72	66	72	69	53	57	69	52	68	50	54	59	54	44
	29	24	52	54	56	58	68	58	65	52	79	61	62	62	44	54	66	36	24
15_EnhAF	30	85	92	91	90	90	90	90	90	92	89	90	92	89	88	89	89	87	85
	31	2	35	47	57	52	40	49	41	27	20	31	22	27	4	6	8	3	2
	32	4	34	22	17	13	12	13	13	18	7	11	11	9	5	5	5	5	4
16_EnhW1	33	84	89	88	88	88	88	88	87	90	87	88	90	88	86	86	87	86	84
	34	91	93	93	93	93	93	93	92	93	92	92	93	92	92	92	92	92	91
17_EnhW2	35	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98
	36	96	97	96	96	96	96	96	96	95	96	96	95	96	96	96	96	96	96
18_EnhAc	37	89	94	95	94	93	94	94	93	93	92	92	92	92	90	90	90	89	89
	38	2	48	49	57	75	68	70	77	62	73	72	64	57	22	22	24	7	2
	39	0	60	56	43	12	28	20	9	5	5	4	3	2	1	1	1	1	0
	40	10	70	69	63	50	58	55	49	42	41	41	35	37	21	18	18	13	10
19_DNase	41	97	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98	97	97
20_ZNF/Rpts	42	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
	43	31	40	37	37	34	33	47	34	34	33	34	33	34	32	32	32	77	31
21_Het	44	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
22_PromP	45	93	95	94	94	94	94	94	94	95	94	95	94	94	95	94	94	94	93
23_PromBiv	46	91	93	93	93	92	92	93	92	93	91	93	92	93	95	92	91	91	91
	47	94	95	94	94	94	94	94	94	94	94	94	94	94	97	94	94	94	94
	48	97	98	98	97	97	97	97	97	97	97	97	97	97	98	97	97	97	97
24_ReprPC	49	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
25_Quies	50	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
min		0	31	22	17	12	11	10	9	5	5	4	3	2	1	1	1	1	0
average		65	77	77	78	77	74	76	75	74	72	75	73	74	67	69	70	69	65

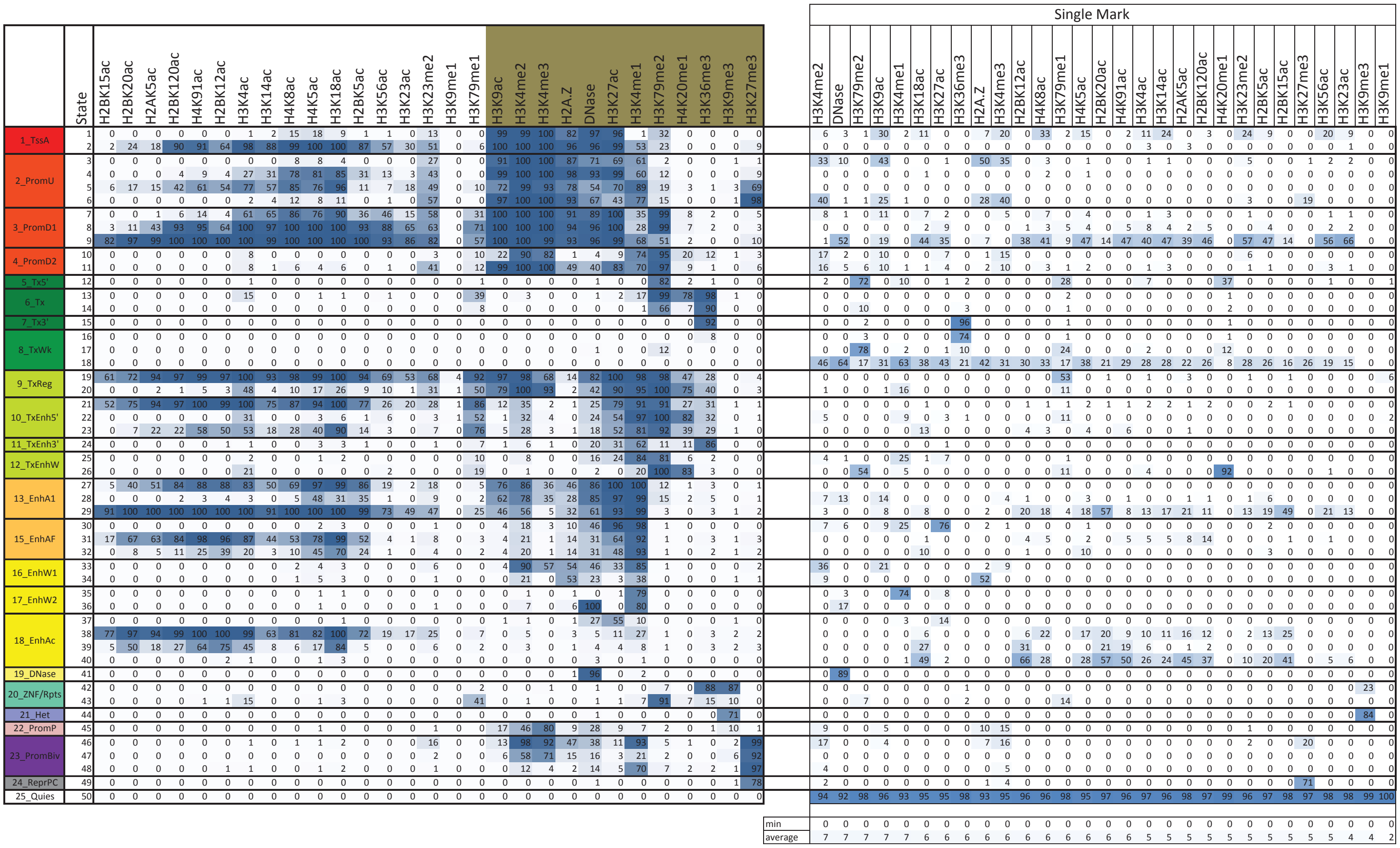
c



	State	H2BK15ac	H2BK20ac	H2AK5ac	H2BK120ac	H4K91ac	H2BK12ac	H3K4ac	H3K14ac	H4K8ac	H4K5ac	H3K18ac	H2BK5ac	H3K56ac	H3K23ac	H3K23me2	H3K9me1	H3K79me1	H3K9ac	H3K4me2	H3K4me3	H2A.Z	DNase	H3K27ac	H3K4me1	H3K79me2	H4K20me1	H3K36me3	H3K9me3	H3K27me3
1_TssA	1	0	0	0	0	0	0	1	2	15	18	9	1	1	0	13	0	0	99	99	100	82	97	96	1	32	0	0	0	0
	2	2	24	18	90	91	64	98	88	99	100	100	87	57	30	51	0	6	100	100	100	96	96	99	53	23	0	0	0	9
2_PromU	3	0	0	0	0	0	0	0	0	8	8	4	0	0	0	27	0	0	91	100	100	87	71	69	61	2	0	0	1	1
	4	0	0	0	4	9	4	27	31	78	81	85	31	13	3	43	0	0	99	100	100	98	93	99	60	12	0	0	0	9
	5	6	17	15	42	61	54	77	57	85	76	96	11	7	18	49	0	10	72	99	93	78	54	70	89	19	3	1	3	69
	6	0	0	0	0	0	0	2	4	12	8	11	0	1	0	57	0	0	97	100	100	93	67	43	77	15	0	0	1	98
3_PromD1	7	0	0	1	6	14	4	61	65	86	76	90	36	46	15	58	0	31	100	100	100	91	89	100	35	99	8	2	0	5
	8	3	11	43	93	95	64	100	97	100	100	100	93	88	65	63	0	71	100	100	100	94	96	100	28	99	7	2	0	3
	9	82	97	99	100	100	100	100	99	100	100	100	100	93	86	82	0	57	100	100	99	93	96	99	68	51	2	0	0	10
4_PromD2	10	0	0	0	0	0	0	8	0	0	0	0	0	0	0	3	0	10	22	90	82	1	4	9	74	95	20	12	1	3
	11	0	0	0	0	0	0	8	1	6	4	6	0	1	0	41	0	12	99	100	100	49	40	83	70	97	9	1	0	6
5_Tx5'	12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	82	2	1	0	0
6_Tx	13	0	0	0	0	0	0	15	0	0	1	1	0	1	0	0	0	39	0	3	0	0	1	2	17	99	78	98	1	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	66	7	90	0	0
7_Tx3'	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0
8_TxWk	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9_TxReg	19	61	72	94	97	99	97	100	93	98	99	100	94	69	53	68	4	92	97	98	68	14	82	100	98	98	47	28	0	4
	20	0	0	2	1	5	3	48	4	10	17	26	9	10	1	31	1	50	79	100	93	2	42	90	95	100	75	40	0	3
10_TxEnh5'	21	52	75	94	97	100	99	100	75	87	94	100	77	26	20	28	1	86	12	35	2	1	25	79	91	91	27	31	1	1
	22	0	0	0	0	0	0	31	0	0	3	6	1	6	0	3	1	52	1	32	4	0	24	54	97	100	82	32	0	1
	23	0	7	22	22	58	50	53	18	28	40	90	14	3	0	7	0	76	5	28	3	1	18	52	81	92	39	29	1	0
11_TxEnh3'	24	0	0	0	0	0	1	1	0	0	3	3	1	0	0	1	0	7	1	6	1	0	20	31	62	11	11	86	0	0
12_TxEnhW	25	0	0	0	0	0	0	2	0	0	1	2	0	0	0	0	0	10	0	8	0	0	16	24	84	81	6	2	0	0
	26	0	0	0	0	0	0	21	0	0	0	0	0	2	0	0	0	19	0	1	0	0	2	0	20	100	83	3	0	0
13_EnhA1	27	5	40	51	84	88	88	83	50	69	97	99	86	19	2	18	0	5	76	86	36	46	86	100	100	12	1	3	0	1
	28	0	0	0	2	3	4	3	0	5	48	31	35	1	0	9	0	2	62	78	35	28	85	97	99	15	2	5	0	1
	29	91	100	100	100	100	100	100	91	100	100	100	99	73	49	47	0	25	46	56	5	32	61	93	99	3	0	3	1	2
15_EnhAF	30	0	0	0	0	0	0	0	0	0	2	3	0	0	0	1	0	0	4	18	3	10	46	96	98	1	0	0	0	0
	31	17	67	63	84	98	96	87	44	53	78	99	52	4	1	8	0	3	4	21	1	14	31	64	92	1	0	3	1	3
	32	0	8	5	11	25	39	20	3	10	45	70	24	1	0	4	0	2	4	20	1	14	31	48	93	1	0	2	1	2
16_EnhW1	33	0	0	0	0	0	0	0	0	2	4	3	0	0	0	6	0	0	4	90	57	54	46	33	85	1	0	0	0	2
	34	0	0	0	0	0	0	0	0	1	5	3	0	0	0	1	0	0	0	21	0	53	23	3	38	0	0	0	1	1
17_EnhW2	35	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	79	0	0	0	0	0
	36	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	7	0	6	100	0	80	0	0	0	0	0
18_EnhAc	37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	27	55	10	0	0	0	1	0
	38	77	97	94	99	100	100	99	63	81	82	100	72	19	17	25	0	7	0	5	0	3	5	11	27	1	0	3	2	2
	39	5	50	18	27	64	75	45	8	6	17	84	5	0	0	6	0	2	0	3	0	1	4	4	8	1	0	3	2	3
	40	0	0	0	0	0	2	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0
19_DNase	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	96	0	2	0	0	0	0	0
20_ZNF/Rpts	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	7	0	88	87	0
	43	0	0	0	0	1	1	15	0	0	1	3	0	0	0	0	0	41	0	1	0	0	1	1	7	91	7	15	10	0
21_Het	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	71	0
22_PromP	45	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	17	46	80	9	28	9	7	2	0	1	10	1
23_PromBiv	46	0	0	0	0	0	0	1	0	1	1	2	0	0	0	16	0	0	13	98	92	47	38	11	93	5	1	0	2	99
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	6	58	71	15	16	3	21	2	0	0	6	92
	48	0	0	0	0	0	1	1	0	0	1	2	0	0	0	1	0	0	0	12	4	2	14	5	70	7	2	2	1	97
24_ReprPC	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	78
25_Quies	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	State	Core + mark																								
		Core Marks	+H3K18ac	+H4K91ac	+H3K79me2	+H2BK12ac	+H2BK120ac	+H4K8ac	+H3K4ac	+H4K5ac	+H3K14ac	+H2AK5ac	+H2BK20ac	+H2BK5ac	+H2BK15ac	+H3K79me1	+H4K20me1	+DNase	+H3K56ac	+H3K27ac	+H3K23ac	+H2A.Z	+H3K9ac	+H3K23me2	+H3K4me2	+H3K9me1
1_TssA	1	82	75	80	57	79	79	71	82	65	82	78	76	81	77	80	82	86	82	88	79	83	90	71	87	82
	2	0	7	24	0	14	23	8	24	7	15	10	11	16	11	1	0	0	8	0	5	0	0	1	0	0
2_PromU	3	41	49	46	55	45	44	45	51	45	48	45	44	47	45	52	47	40	47	40	45	53	57	37	50	41
	4	0	34	7	1	7	9	27	4	29	7	2	2	4	0	1	0	2	2	4	3	0	0	1	0	0
	5	3	50	43	3	38	35	43	52	40	38	16	17	14	11	9	3	4	7	9	16	4	4	4	3	3
	6	58	53	58	50	58	58	50	57	55	57	58	58	58	58	58	58	66	58	49	58	84	95	59	70	58
3_PromD1	7	6	41	8	34	8	10	34	12	24	21	8	8	6	7	4	8	15	9	9	6	14	7	9	6	6
	8	1	53	72	32	45	72	55	63	59	60	30	11	69	5	50	1	2	62	1	51	1	1	27	1	1
	9	14	40	45	9	51	45	40	42	40	43	61	77	46	73	24	17	28	48	31	47	23	24	36	17	14
4_PromD2	10	26	35	29	40	29	30	35	24	36	31	28	29	32	28	27	21	54	31	54	28	50	39	37	26	26
	11	34	40	38	53	37	38	42	30	44	42	37	38	40	37	32	26	41	39	36	38	35	42	25	34	34
5_Tx5'	12	8	9	9	90	9	9	9	7	10	9	9	10	10	9	24	40	10	8	8	9	13	9	8	7	8
6_Tx	13	10	11	11	65	11	10	10	22	11	10	10	10	11	10	36	78	12	12	14	10	11	10	10	11	10
	14	8	9	9	77	9	9	9	10	9	9	9	9	9	9	34	31	10	9	9	8	11	8	8	8	8
7_Tx3'	15	97	96	96	99	96	96	96	97	96	96	96	96	96	96	96	96	97	96	96	96	96	96	96	96	97
8_TxWk	16	90	90	89	97	90	89	89	93	89	89	89	89	89	89	93	94	90	90	90	90	89	90	90	90	90
	17	9	10	10	91	10	10	10	8	11	10	10	11	11	10	30	17	11	10	8	10	14	10	10	9	9
	18	77	79	78	80	79	78	78	79	78	78	77	78	77	77	78	78	89	77	79	77	78	77	77	78	77
9_TxReg	19	4	28	31	10	30	32	30	22	30	32	31	24	30	24	32	6	18	25	10	23	6	18	16	8	5
	20	47	43	46	53	47	47	44	39	43	46	47	47	45	47	49	71	40	42	46	47	51	43	41	47	46
10_TxEnh5'	21	0	22	27	3	28	34	28	23	27	26	32	27	27	23	15	1	1	8	7	12	0	4	10	1	0
	22	16	17	17	45	17	17	17	10	17	16	17	17	17	17	37	67	15	14	13	16	18	16	15	19	18
	23	0	18	12	1	10	5	7	2	6	4	6	3	2	2	6	1	1	1	0	1	0	0	1	1	0
11_TxEnh3'	24	62	60	60	70	61	60	61	67	60	61	61	61	60	61	68	67	68	62	67	62	61	61	62	63	62
12_TxEnhW	25	23	27	28	73	29	29	28	19	29	27	27	30	29	28	46	37	27	26	14	26	44	26	25	21	23
	26	5	6	6	26	7	7	6	5	7	6	7	7	7	6	15	82	7	5	7	6	7	6	6	5	6
13_EnhA1	27	0	21	24	0	24	22	19	22	12	12	14	8	13	6	0	0	0	3	0	2	4	1	1	0	0
	28	20	19	24	19	25	25	23	24	21	23	24	22	19	20	20	20	29	20	24	20	15	28	18	31	20
	29	1	70	70	1	70	74	75	72	74	73	76	77	79	80	19	2	5	68	8	55	20	32	48	25	2
15_EnhAF	30	22	25	21	24	23	20	21	21	24	20	20	19	21	19	21	24	18	18	81	19	19	21	19	26	22
	31	0	13	38	0	28	29	21	30	10	14	24	21	8	17	0	0	0	9	1	3	0	1	2	1	0
	32	0	37	15	0	22	17	13	13	20	10	16	15	8	4	0	0	0	1	2	1	0	0	1	1	0
16_EnhW1	33	35	36	37	42	37	37	36	39	37	37	37	37	39	37	40	39	32	37	33	37	30	50	36	41	35
	34	1	2	1	1	1	1	1	1	3	1	1	1	1	1	1	1	4	1	2	1	71	1	1	22	1
17_EnhW2	35	84	83	84	89	83	84	84	87	84	85	84	84	84	85	87	88	89	86	90	85	82	87	84	86	84
	36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79	0	3	0	0	0	0	0	0
18_EnhAc	37	3	7	6	4	7	6	6	6	6	5	6	6	5	5	3	4	8	4	83	4	4	6	4	5	3
	38	0	40	51	1	42	67	63	61	58	50	71	60	56	62	1	1	1	19	1	16	1	1	20	1	1
	39	0	59	38	0	54	10	3	16	4	1	6	24	1	3	0	0	0	0	0	0	0	0	0	0	0
	40	2	68	62	3	68	49	42	54	42	37	49	57	35	41	4	3	4	16	9	15	4	5	19	5	2
19_DNase	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	93	0	0	0	0	0	0	0	0
20_ZNF/Rpts	42	97	97	97	98	97	97	97	97	97	97	97	97	97	97	98	97	97	97	97	97	97	97	97	97	97
	43	3	5	5	25	4	4	4	5	4	4	4	4	4	4	34	6	4	4	3	4	4	4	4	3	3
21_Het	44	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99
22_PromP	45	60	65	62	65	61	62	66	63	66	63	61	60	63	60	62	62	65	63	69	61	68	70	66	67	60
23_PromBiv	46	67	72	68	73	67	67	74	68	71	69	67	67	67	67	67	67	68	67	76	67	70	80	76	71	67
	47	59	60	59	59	59	59	60	59	60	59	59	59	59	59	59	59	62	59	63	59	66	63	65	79	59
	48	82	83	82	87	82	82	82	82	82	82	82	82	82	82	82	83	85	82	83	82	82	82	82	82	82
24_ReprPC	49	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	100	99
25_Quies	50	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	97	98	97	97	97	97	97	97	97	97
min		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
average		31	43	42	42	42	42	41	41	41	40	40	40	39	38	38	38	37	37	36	36	36	35	34	34	31

d



	State	H2BK15ac	H2BK20ac	H2AK5ac	H2BK120ac	H4K91ac	H2BK12ac	H3K4ac	H3K14ac	H4K8ac	H4K5ac	H3K18ac	H2BK5ac	H3K56ac	H3K23ac	H3K23me2	H3K9me1	H3K79me1	H3K9ac	H3K4me2	H3K4me3	H2A.Z	DNase	H3K27ac	H3K4me1	H3K79me2	H4K20me1	H3K36me3	H3K9me3	H3K27me3
1_TssA	1	0	0	0	0	0	0	1	2	15	18	9	1	1	0	13	0	0	99	99	100	82	97	96	1	32	0	0	0	0
	2	2	24	18	90	91	64	98	88	99	100	100	87	57	30	51	0	6	100	100	100	96	96	99	53	23	0	0	0	9
2_PromU	3	0	0	0	0	0	0	0	0	8	8	4	0	0	0	27	0	0	91	100	100	87	71	69	61	2	0	0	1	1
	4	0	0	0	4	9	4	27	31	78	81	85	31	13	3	43	0	0	99	100	100	98	93	99	60	12	0	0	0	9
	5	6	17	15	42	61	54	77	57	85	76	96	11	7	18	49	0	10	72	99	93	78	54	70	89	19	3	1	3	69
	6	0	0	0	0	0	0	2	4	12	8	11	0	1	0	57	0	0	97	100	100	93	67	43	77	15	0	0	1	98
3_PromD1	7	0	0	1	6	14	4	61	65	86	76	90	36	46	15	58	0	31	100	100	100	91	89	100	35	99	8	2	0	5
	8	3	11	43	93	95	64	100	97	100	100	100	93	88	65	63	0	71	100	100	100	94	96	100	28	99	7	2	0	3
	9	82	97	99	100	100	100	100	99	100	100	100	100	93	86	82	0	57	100	100	99	93	96	99	68	51	2	0	0	10
4_PromD2	10	0	0	0	0	0	0	8	0	0	0	0	0	0	0	3	0	10	22	90	82	1	4	9	74	95	20	12	1	3
	11	0	0	0	0	0	0	8	1	6	4	6	0	1	0	41	0	12	99	100	100	49	40	83	70	97	9	1	0	6
5_Tx5'	12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	82	2	1	0	0
6_Tx	13	0	0	0	0	0	0	15	0	0	1	1	0	1	0	0	0	39	0	3	0	0	1	2	17	99	78	98	1	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	1	66	7	90	0	0
7_Tx3'	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0
8_TxWk	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	12	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9_TxReg	19	61	72	94	97	99	97	100	93	98	99	100	94	69	53	68	4	92	97	98	68	14	82	100	98	98	47	28	0	4
	20	0	0	2	1	5	3	48	4	10	17	26	9	10	1	31	1	50	79	100	93	2	42	90	95	100	75	40	0	3
10_TxEnh5'	21	52	75	94	97	100	99	100	75	87	94	100	77	26	20	28	1	86	12	35	2	1	25	79	91	91	27	31	1	1
	22	0	0	0	0	0	0	31	0	0	3	6	1	6	0	3	1	52	1	32	4	0	24	54	97	100	82	32	0	1
	23	0	7	22	22	58	50	53	18	28	40	90	14	3	0	7	0	76	5	28	3	1	18	52	81	92	39	29	1	0
11_TxEnh3'	24	0	0	0	0	0	1	1	0	0	3	3	1	0	0	1	0	7	1	6	1	0	20	31	62	11	11	86	0	0
12_TxEnhW	25	0	0	0	0	0	0	2	0	0	1	2	0	0	0	0	0	10	0	8	0	0	16	24	84	81	6	2	0	0
	26	0	0	0	0	0	0	21	0	0	0	0	0	2	0	0	0	19	0	1	0	0	2	0	20	100	83	3	0	0
13_EnhA1	27	5	40	51	84	88	88	83	50	69	97	99	86	19	2	18	0	5	76	86	36	46	86	100	100	12	1	3	0	1
	28	0	0	0	2	3	4	3	0	5	48	31	35	1	0	9	0	2	62	78	35	28	85	97	99	15	2	5	0	1
	29	91	100	100	100	100	100	100	91	100	100	100	99	73	49	47	0	25	46	56	5	32	61	93	99	3	0	3	1	2
15_EnhAF	30	0	0	0	0	0	0	0	0	0	2	3	0	0	0	1	0	0	4	18	3	10	46	96	98	1	0	0	0	0
	31	17	67	63	84	98	96	87	44	53	78	99	52	4	1	8	0	3	4	21	1	14	31	64	92	1	0	3	1	3
	32	0	8	5	11	25	39	20	3	10	45	70	24	1	0	4	0	2	4	20	1	14	31	48	93	1	0	2	1	2
16_EnhW1	33	0	0	0	0	0	0	0	0	2	4	3	0	0	0	6	0	0	4	90	57	54	46	33	85	1	0	0	0	2
	34	0	0	0	0	0	0	0	0	1	5	3	0	0	0	1	0	0	0	21	0	53	23	3	38	0	0	0	1	1
17_EnhW2	35	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	79	0	0	0	0	0
	36	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	7	0	6	100	0	80	0	0	0	0	0
18_EnhAc	37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0	1	27	55	10	0	0	0	1	0
	38	77	97	94	99	100	100	99	63	81	82	100	72	19	17	25	0	7	0	5	0	3	5	11	27	1	0	3	2	2
	39	5	50	18	27	64	75	45	8	6	17	84	5	0	0	6	0	2	0	3	0	1	4	4	8	1	0	3	2	3
	40	0	0	0	0	0	2	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0
19_DNase	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	96	0	2	0	0	0	0	0
20_ZNF/Rpts	42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	7	0	88	87	0
	43	0	0	0	0	1	1	15	0	0	1	3	0	0	0	0	0	41	0	1	0	0	1	1	7	91	7	15	10	0
21_Het	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	71	0
22_PromP	45	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	17	46	80	9	28	9	7	2	0	1	10	1
23_PromBiv	46	0	0	0	0	0	0	1	0	1	1	2	0	0	0	16	0	0	13	98	92	47	38	11	93	5	1	0	2	99
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	6	58	71	15	16	3	21	2	0	0	6	92
	48	0	0	0	0	0	1	1	0	0	1	2	0	0	0	1	0	0	0	12	4	2	14	5	70	7	2	2	1	97
24_ReprPC	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	78
25_Quies	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	State	Single Mark																												
		H3K4me2	DNase	H3K79me2	H3K9ac	H3K4me1	H3K18ac	H3K27ac	H3K36me3	H2A.Z	H3K4me3	H2BK12ac	H4K8ac	H3K79me1	H4K5ac	H2BK20ac	H4K91ac	H3K4ac	H3K14ac	H2AK5ac	H2BK120ac	H4K20me1	H3K23me2	H2BK5ac	H2BK15ac	H3K27me3	H3K56ac	H3K23ac	H3K9me3	H3K9me1
1_TssA	1	6	3	1	30	2	11	0	0	7	20	0	33	2	15	0	2	11	24	0	3	0	24	9	0	0	20	9	0	0
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0	0	0	0	0	0	0	1	0	0
2_PromU	3	33	10	0	43	0	0	1	0	50	35	0	3	0	1	0	0	1	1	0	0	0	5	0	0	1	2	2	0	0
	4	0	0	0	0	0	1	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	40	1	1	25	1	0	0	0	28	40	0	0	0	0	0	0	0	0	0	0	0	3	0	0	19	0	0	0	0
3_PromD1	7	8	1	0	11	0	7	2	0	0	5	0	7	0	4	0	0	1	3	0	0	0	1	0	0	0	1	1	0	0
	8	0	0	0	0	0	2	9	0	0	0	1	3	5	4	0	5	8	4	2	5	0	0	4	0	0	2	2	0	0
	9	1	52	0	19	0	44	35	0	7	0	38	41	9	47	14	47	40	47	39	46	0	57	47	14	0	56	66	0	0
4_PromD2	10	17	2	0	10	0	0	7	0	1	15	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0
	11	16	5	6	10	1	1	4	0	2	10	0	3	1	2	0	0	1	3	0	0	0	1	1	0	0	3	1	0	0
5_Tx5'	12	2	0	72	0	10	0	1	2	0	0	0	0	28	0	0	0	7	0	0	0	37	0	0	0	0	1	0	0	1
6_Tx	13	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	14	0	0	10	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
7_Tx3'	15	0	0	2	0	0	0	0	96	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
8_TxWk	16	0	0	3	0	0	0	0	74	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	17	0	0	78	0	2	0	1	10	0	0	0	0	24	0	0	0	2	0	0	0	12	0	0	0	0	0	0	0	0
	18	46	64	17	31	63	38	43	21	42	31	30	33	17	38	21	29	28	28	22	26	8	28	26	16	26	19	15	0	0
9_TxReg	19	0	0	0	0	0	0	0	0	0	0	0	0	53	0	1	0	1	0	3	0	0	1	0	1	0	0	0	0	6
	20	1	0	0	1	16	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10_TxEnh5'	21	0	0	0	0	0	1	0	0	0	0	1	1	1	2	1	1	2	2	1	2	0	0	2	1	0	0	0	0	0
	22	5	0	0	0	9	0	3	1	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	23	0	0	0	0	0	13	0	0	0	0	4	3	0	4	0	6	0	0	1	0	0	0	0	0	0	0	0	0	0
11_TxEnh3'	24	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12_TxEnhW	25	4	1	0	0	25	1	7	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	26	0	0	54	0	5	0	0	0	0	0	0	0	11	0	0	0	4	0	0	0	92	0	0	0	0	1	0	0	0
13_EnhA1	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	28	7	13	0	14	0	0	0	0	0	4	1	0	0	3	0	1	0	0	1	1	0	1	6	0	0	0	0	0	0
	29	3	0	0	8	0	8	0	0	2	0	20	18	4	18	57	8	13	17	21	11	0	13	19	49	0	21	13	0	0
15_EnhAF	30	7	6	0	9	25	0	76	0	2	1	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
	31	0	0	0	0	0	1	0	0	0	0	4	5	0	2	0	5	5	5	8	14	0	0	0	1	0	1	0	0	0
	32	0	0	0	0	0	10	0	0	0	0	1	0	0	10	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0
16_EnhW1	33	36	0	0	21	0	0	0	0	2	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	34	9	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17_EnhW2	35	0	3	0	0	74	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	36	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18_EnhAc	37	0	0	0	0	3	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	38	0	0	0	0	0	6	0	0	0	0	6	22	0	17	20	9	10	11	16	12	0	2	13	25	0	0	0	0	0
	39	0	0	0	0	0	27	0	0	0	0	31	0	0	0	21	19	6	0	1	2	0	0	0	0	0	0	0	0	0
	40	0	0	0	0	1	49	2	0	0	0	66	28	0	28	57	50	26	24	45	37	0	10	20	41	0	5	6	0	0
19_DNase	41	0	89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20_ZNF/Rpts	42	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0
	43	0	0	7	0	0	0	0	2	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21_Het	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	84	0
22_PromP	45	9	0	0	5	0	0	0	0	10	15	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
23_PromBiv	46	17	0	0	4	0	0	0	0	7	16	0	0	0	0	0	0	0	0	0	0	0	2	0	0	20	0	0	0	0
	47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	48	4	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24_ReprPC	49	2	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	71	0	0	0	0
25_Quies	50	94	92	98	96	93	95	95	98	93	95	96	96	98	95	97	96	97	96	98	97	99	96	97	98	97	98	98	99	100
min		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
average		7	7	7	7	7	6	6	6	6	6	6	6	6	6	6	6	5	5	5	5	5	5	5	5	5	5	4	4	2

Supplementary Figure 43: Chromatin State Recovery with Different Mark Subsets.

(a) (left) The emission parameters of the same expanded mark chromatin state model on imputed data also shown in **Fig. 6d** and **Fig. S39-40**. (right) An evaluation of chromatin state recovery of this model using all marks of the model except the indicated mark of the column. Shown for each state is the percentage of locations assigned to it based on a maximum-posterior decoding for the full set of marks that would receive the same state assignment based on performing the posterior decoding with the indicated subset of marks. Along the bottom is the minimum state recovery of any state and the average state recovery. Columns are ordered by increasing minimum state recovery. **(b)** The same as **a** except showing the results for the subset of all Tier-1 and 2 marks along with this set extended by the one additional mark indicated in the column. Columns for the extended set are ordered in decreasing order of minimum state recovery. **(c)** The same as **b** except based on the core mark set of H3K4me3, H3K4me1, H3K36me3, H3K9me3, and H3K27me3 instead of the Tier-1and 2 mark set and the columns are ordered in decreasing average state recovery. **(d)** The same as **c** except showing the chromatin state recovery with only the single mark of the column.