All Figures and Tables in Supplements

Supplementary Table 1: Summary of read statistics and quality parameters
        <u>submitted as separate excel file</u>

Supplementary Table 2: List of GO-terms for tissue-specific H3K4me3 peaks

Supplementary Table 3: Summary of results from QTL analysis

Supplementary Table 4: List of differential TF binding sites in heart and liver

Supplementary Table 5: List of genes located in the QTL hotspot at chr3
        <u>submitted as separate excel file</u>

Supplementary Table 6: Results of the graphical model analysis
        <u>submitted as separate excel file</u>

Supplementary Table 7: Results of ENSEMBL-based analysis of histoneQTL

Supplementary Table 8: Primer sequences for ChIP-qPCR assays

Supplementary Figure 1: Workflow of data analysis

Supplementary Figure 2: Coverage plots for histone marks

Supplementary Figure 3: Effects of SNPs on H3K4me3 quantification

Supplementary Figure 4: Fold change of differential histone marks

Supplementary Figure 5: Validation of differential histone modification marks

Supplementary Figure 6: Quantile-quantile plots for the QTL analyzes of histone modification traits

Supplementary Figure 7: Comparison of differential histone marks between RI and parental strains

Supplementary Figure 8: Representation of competing graphical models

Supplementary Figure 9: Distribution of graphical models

Supplementary Figure 10: Distribution of H3K4me3 peaks according to genomic features

Supplementary Figure 11: Histogram of P-values for the correlation between expression and histone traits

Supplementary Figure 12: Evaluation of H3K27me3 peak calls

Supplementary Table 1   submitted as separate excel file
Summary of read statistics and quality parameters for ChIP-seq and RNA-seq data sets, including sequenced, mapped reads, relative strand cross-correlation (RSC), normalized strand cross-correlation (NSC) and the ENCODE quality flag based on these two metrics.

Supplementary Table 2: List of GO-terms with 20 highest significance levels in tissue-specific H3K4me3 peaks.

Tissue	GO-ID	p-value	GO-Term
Heart	GO:0032501	6.15E-16	multicellular organismal process
	GO:0003012	2.59E-15	muscle system process
	GO:0006936	1.07E-14	muscle contraction
	GO:0003008	3.29E-14	system process
	GO:0007275	1.22E-12	multicellular organismal development
	GO:0061061	2.02E-12	muscle structure development
	GO:0060047	3.10E-12	heart contraction
	GO:0003015	3.10E-12	heart process
	GO:0048731	4.08E-12	system development
	GO:0022803	4.93E-12	passive transmembrane transporter activity
	GO:0015267	4.93E-12	channel activity
	GO:0048856	1.17E-11	anatomical structure development
	GO:0032502	3.42E-11	developmental process
	GO:0014706	4.25E-11	striated muscle tissue development
	GO:0008016	9.64E-11	regulation of heart contraction
	GO:0060537	1.34E-10	muscle tissue development
	GO:0022836	1.36E-10	gated channel activity
	GO:0072358	1.50E-10	cardiovascular system development
	GO:0072359	1.50E-10	circulatory system development
	GO:0006941	1.61E-10	striated muscle contraction
Liver	GO:0006955	2.62E-19	immune response
	GO:0016712	6.07E-15	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen
	GO:0030414	1.57E-14	peptidase inhibitor activity
	GO:0004866	1.96E-14	endopeptidase inhibitor activity
	GO:0004867	3.68E-14	serine-type endopeptidase inhibitor activity
	GO:0002376	7.44E-14	immune system process
	GO:0061135	1.00E-13	endopeptidase regulator activity
	GO:0006952	1.16E-13	defense response
	GO:0004497	3.61E-13	monooxygenase activity
	GO:0020037	9.90E-13	heme binding
	GO:0019752	1.95E-12	carboxylic acid metabolic process
	GO:0043436	1.95E-12	oxoacid metabolic process
	GO:0070330	2.48E-12	aromatase activity
	GO:0006082	2.49E-12	organic acid metabolic process
	GO:0046906	2.62E-12	tetrapyrrole binding
	GO:0042180	2.86E-12	cellular ketone metabolic process

GO:0050776	2.94E-12	regulation of immune response
GO:0061134	1.29E-11	peptidase regulator activity
GO:0050778	1.68E-11	positive regulation of immune response
GO:0002684	3.10E-11	positive regulation of immune system process

Supplementary Table 3: Summary of results from QTL analysis on histone methylation traits and strain distribution patterns in the RI panel.

Tissue	Modification	Marks[a]	Marks with histoneQTL	*cis*	*trans*	Unique traits	Unique marker	Polygenic
heart	H3K4me3	25,064	5,142	4,207	1,059	2,638	2504	252
	H3K27me3	4,214	117	107	80	102	15	3
liver	H3K4me3	31,447	4,409	3,729	1,028	2,945	1,464	199
	H3K27me3	3,776	234	208	145	196	38	2

Regions were obtained from MACS peak calling-based analysis for H3K4me3 and HMM-based peak calling for H3K27me3, respectively.
[a] total number of identified histone methylation peaks

Supplementary Table 4: List of differential TF binding sites in heart and liver tissue.

Tissue	Matrix	p	q	Factor	Description	Ensembl id
heart	V$CETS1P54_01	1.40E-05	0.0040	Ets1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	ENSRNOG00000008941
heart	V$E2F1_Q6_01	1.19E-04	0.0093	E2f1	E2F transcription factor 1	ENSRNOG00000016708
heart	V$MYC_Q2	1.32E-04	0.0093	Myc	myelocytomatosis oncogene	ENSRNOG00000004500
heart	V$MYC_Q2	1.32E-04	0.0093	Max	MYC associated factor X	ENSRNOG00000008049
heart	V$USF_C	4.62E-04	0.0239	Usf1	upstream transcription factor 1	ENSRNOG00000004255
heart	V$TBX5_02	5.06E-04	0.0239	Tbx5	T-box 5	ENSRNOG00000001399
heart	V$MSX1_01	7.10E-04	0.0287	MSX1	msh homeobox 1	ENSG00000163132
heart	V$STAT3_02	8.49E-04	0.0300	Stat3	signal transducer and activator of transcription 3 (acute-phase response factor)	ENSRNOG00000019742
heart	V$USF2_Q6	9.78E-04	0.0307	Usf2	upstream transcription factor 2, c-fos interacting	ENSRNOG00000021030
heart	V$ZF5_01	2.22E-03	0.0605	Zfp161	zinc finger protein 161	ENSRNOG00000016719
heart	V$HAND1E47_01	2.46E-03	0.0605	Hand1	heart and neural crest derivatives expressed 1	ENSRNOG00000002582
heart	V$PAX5_02	2.57E-03	0.0605	PAX5	paired box 5	ENSG00000196092
heart	V$MTF1_Q4	3.64E-03	0.0793	Mtf1	metal-regulatory transcription factor 1	ENSRNOG00000025724
heart	V$DEAF1_02	4.26E-03	0.0861	Deaf1	deformed epidermal autoregulatory factor 1 (Drosophila)	ENSRNOG00000017960
liver	V$ZF5_01	4.68E-04	0.0629	Zfp161	zinc finger protein 161	ENSRNOG00000016719
liver	V$YY1_02	5.20E-04	0.0629	Yy1	YY1 transcription factor	ENSRNOG00000004339
liver	V$HNF4_Q6_03	6.66E-04	0.0629	Hnf4a	hepatic nuclear factor 4, alpha	ENSMUSG00000017950
liver	V$CDPCR3_01	1.37E-03	0.0827	Cux1	cut-like homeobox 1	ENSRNOG00000001424
liver	V$E2F_Q2	1.46E-03	0.0827	E2f1	E2F transcription factor 1	ENSRNOG00000016708
liver	V$PAX5_01	1.76E-03	0.0830	PAX5	paired box 5	ENSG00000196092
liver	V$PAX8_B	2.06E-03	0.0833	Pax8	paired box 8	ENSRNOG00000026203
liver	V$ER_Q6	2.60E-03	0.0919	Esr1	estrogen receptor 1	ENSRNOG00000019358

Supplementary Table 5   <u>**submitted as separate excel file**</u>
List of genes located in the QTL hotspot at chr3 with either a cis-regulating eQTL or a structural variant (SNP/indel) with consequences on the amino acid sequence.

Supplementary Table 6   <u>**submitted as separate excel file**</u>
Integrated genetic analysis of histone modifications H3K4me3, H3K27me3 and gene expression levels using graphical modeling. The graphical models were subdivided into two main classes: models linking the genotype variable only with histone modifications (link.only.modification) or models linking the genotype variable with the gene expression variable (links.expression) which is further subdivided into models with a direct link (direct.link) or indirect link (indirect.link). In the respective columns we report the bootstrap probabilities for each model class.

Supplementary Table 7: Results of ENSEMBL-based analysis of histoneQTL mapping results in heart and liver tissue. Results are used for integration with gene expression data.

	Heart		Liver		Both tissues	
H3K4me3 (traits)	(10,396)		(11,076)			
FDR-cutoff	*cis*	*trans*	*cis*	*trans*	*cis*	*Trans*
0.05	531	1,867	407	144	127	12

Supplementary Table 8: Primer sequences and genomic position for ChIP-qPCR assays.

Histone modification	Tissue	Region	Forward-Primer	Reverse-Primer	Genomic Position
H3K4me3	left ventricle	1	CCCTAGCCTGGTCCATAACCA	TTCCTAGCAGGGTTGGAGGAA	chr14:1,489,583-1,489,652
H3K4me3	left ventricle	2	CTGTCTTTGTTTTGCGATTTCG	GGCGCAGCCTTTTCCTTT	chr1:73,217,126-73,217,192
H3K4me3	left ventricle	3	GCGAGAGGTCGAGCTCATG	CCTACTCGCCTCCTGTCACATC	chr11:32,624,026-32,624,090
H3K4me3	liver	1	CCCTAGCCTGGTCCATAACCA	TTCCTAGCAGGGTTGGAGGAA	chr14:1,489,583-1,489,652
H3K4me3	liver	2	TGCTCAGAATGTGCATGTGAGTT	CCCAATACCACGAGAAAATGGA	chr7:2,110,064-2,110,137
H3K4me3	liver	3	GGCACTTCAGGCGGTAGGT	CCTGCTAGTGGAAAGGCATCA	chr3:5,936,438-5,936,501
H3K4me1	left ventricle	1	AAGGCCACATCTATTCCAATGAG	ACGATACCACTTGAGTTGCTTTGTA	chr19:44,176,907-44,176,981
H3K4me1	left ventricle	2	CAGGAAAGCACATTGCTCATTC	TTTGAACTCCGTGAACCACAAC	chr16:72,285,915-72,285,981
H3K4me1	left ventricle	3	CATTGTGGAAGAAGAGACAGAACAA	TGAGGGAAACATGCGATCAG	chr5:56,813,182-56,813,255
H3K4me1	liver	1	TTGGAGGCTGAAAGCTGGTT	GAAAGGGCGTCTGGTCCAC	chr15:14,191,980-14,192,030
H3K4me1	liver	2	CTGCATCCCAGCCCAGTCT	AGATAGAGCCATCCTGGCCA	chr5:119,061,967-119,062,017
H3K4me1	liver	3	TCCAGCTCACTGAACTGCCC	AGAAAGAATTTGGTGGCTCTGC	chr3:5,299,021-5,299,071
H3K27me3	left ventricle	1	AGCCTCGCTTTGAGCAGTACA	GAGCGTCCAAGCTGAGCAG	chr15:19,393,444-19,393,494
H3K27me3	left ventricle	2	TTGGAGGGATCAGCACTGTG	AAAACCAGTCTCTGGGACTGGA	chr1:2,123,750-2,123,800
H3K27me3	left ventricle	3	TTGGTGCTGTTTCATTTTGATCA	GCTGAAGGGCCTTTGCTACA	chr11:76,487,730-76,487,780
H3K27me3	liver	1	AGCCTCGCTTTGAGCAGTACA	GAGCGTCCAAGCTGAGCAG	chr15:19,393,444-19,393,494
H3K27me3	liver	2	TTGGAGGGATCAGCACTGTG	AAAACCAGTCTCTGGGACTGGA	chr1:2,123,750-2,123,800
H3K27me3	liver	3	GAATTCGGCAACAATCTGGG	TTCATGTGACGCGCTTTCTT	chr13:86,915,949-86,916,000
H4K20me1	left ventricle	1	GGGCACTGCATTTATCTGTGG	TGCTTAGCAAGTGGGCTGTG	chr5:133,093,819-133,093,869
H4K20me1	left ventricle	2	GATGGGATGGTGGCCTACAG	GCAAAGTGTTTGAAGGTCACCAG	chr4:76,814,701-76,814,751
H4K20me1	left ventricle	3	TAACGGACAAACCTGGGTGC	GGTCCCCTGCAGAGTGAACA	chr11:81,474,879-81,474,929
H4K20me1	liver	1	GGGCACTGCATTTATCTGTGG	TGCTTAGCAAGTGGGCTGTG	chr5:133,093,819-133,093,869
H4K20me1	liver	2	GGTCCCTTTCAATGACAGTACACA	TCCTCCCCCCCTTACAGG	chr15:103,559,940-103,560,000
H4K20me1	liver	3	TATCCTCCCTGACCAAGGCA	GGTTGCAAGCTGGATTCCAC	chr6:128,441,298-128,441,348
H3K4me3		Gapdh	TACTTCGGCCACCCTATCCA	CATGCCGGTCTGGCTAAATT	chr4:161,285,658-161,285,725
H3K4me1		Gapdh	CAACAGTCCTGCCTCCTGCT	CCCCAAGCCCTTAATCTGCT	chr4:161,287,876-161,287,926
H3K27me3		Myod1	CGGCTACCCAAGGTGGAGAT	AGAGCCTGCAGACCTTCAATG	chr1:96,910,826-96,910,887
H4K20me1		Gapdh	TACTTCGGCCACCCTATCCA	CATGCCGGTCTGGCTAAATT	chr4:161,285,658-161,285,725

**Supplementary Figure 1: Workflow of data analysis.**

Workflow of data analysis

Replicates / strain / tissue			
ChIP	BN	SHR	30 Ris
H3K4me3	3	3	1
H3K27me3	3	3	1
H3K4me1	3	3	0
H4K20me1	3	3	0

Replicates / strain / tissue			
	BN	SHR	30 Ris
RNA-seq	5	5	1

ChIP-seq read alignment

Quality control and peak calling using
MACS (H3K4me3, H3K4me1) and
HMM (H3K27me3, H4K20me1)

Identified peaks

Filter 1: Selection of peaks called in 3 progenitor
strains and 5 RI strains,
Quantification of reads in each region

Read counts in each region and strain

Quantile normalization of read counts Filter 2:
Selection of regions were read coverage
is > 1X in 3 progenitor strains and 5 RI strains

Normalized reads across strains

Correlation analysis
Peak distribution

Quality assessment of quantitative
ChIP-seq data

DESeq model and p-value correction (FDR)

NB model, permutations and
p-value correction (FDR)

Differential peaks in progenitor strains
(Table 1)

Histone QTLs in RI strains
(Table 2, Supplementary Table 1)

Overlap of histone QTLs with ENSEMBL genes
quantification of reads in annotated promoter
regions and gene bodies

Subset of histoneQTLs with
annotated ENSEMBL gene
(Supplementary Table 2)

RNA-seq reads

Read mapping using TopHat

Mapped reads across known
and predicted slice junctions

Quantification of reads in each gene per kb

Read counts in each gene and strain

Quantile based scaling normalization of read counts,
Filter 1: Selection of expressed genes in three
progenitor strains and 5 RI strains

Normalized RNA reads across strains

NB model, permutations and
p-value correction (FDR)

eQTLs in RI strains
(Supplementary Table 2)

Graphic models of histone and gene
expression QTLs, boot strapping

Integration of histone QTLs
and gene expression data (Figure 3)

Supplementary Figure 2: Coverage plots for histone marks. We plotted the average read coverage of H3K4me3 and H3K4me1 peaks 2,000bp upstream and downstream of the TSS for ENSEMBL-annotated expressed (red) and non-expressed genes (blue) for heart (solid line) and liver tissue (dotted line). For H3K27me3 and H4K20me1 we plotted the average read coverage 2,000bp upstream and downstream of the gene combined with the average coverage per base in 10 bins across the length of the gene body for expressed and non-expressed genes.
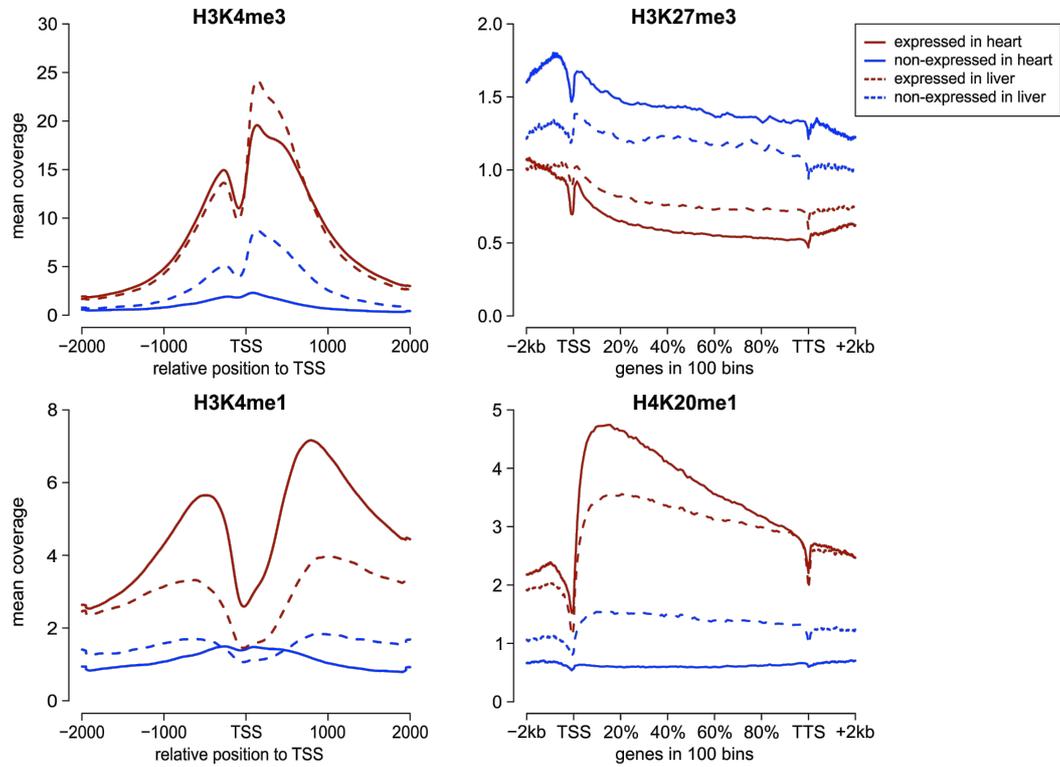
Supplementary Figure 3: Effects of SNPs on H3K4me3 quantification. In order to rule out effects of SNPs on the results of the short read alignment and the quantification of modified regions we have aligned the reads against version of the reference where reference alleles were substituted by SHR alleles. We then quantified H3K4me3 in heart tissue in three biological replicates of SHR animals for all autosomal Macs regions and Ensembl annotated promoter regions. The scatterplot shows the results of the quantification (left: Macs, right: Ensembl). We performed a differential analysis using DESeq and found significant differences only for 50 regions (28 Ensembl) highlighted in red. The majority constitutes regions where the quantification against the reference genome underestimates the true modification levels, therefore these regions can be considered false negatives. Only four of the differential promoter regions gives rise to a differential call in the parental analysis and three to a QTL in the analysis of RI strains (none of the regions were in ensembl annotation based promoter regions). We excluded those regions from further analyses. We concluded that SNPs do not have a major effect on the quanitification of modification levels and we used the quantification against the reference genome for the rest of the analysis.
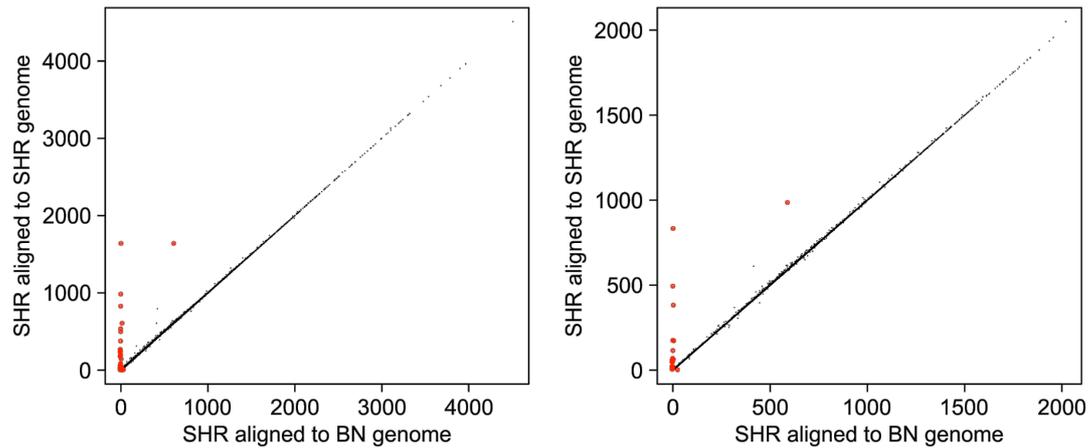
Supplementary Figure 4: Fold change of differential histone marks. The fold change of differential H3K4me3, H3K27me1, H3K4me1 and H4K20me1 marks in heart and liver tissues.
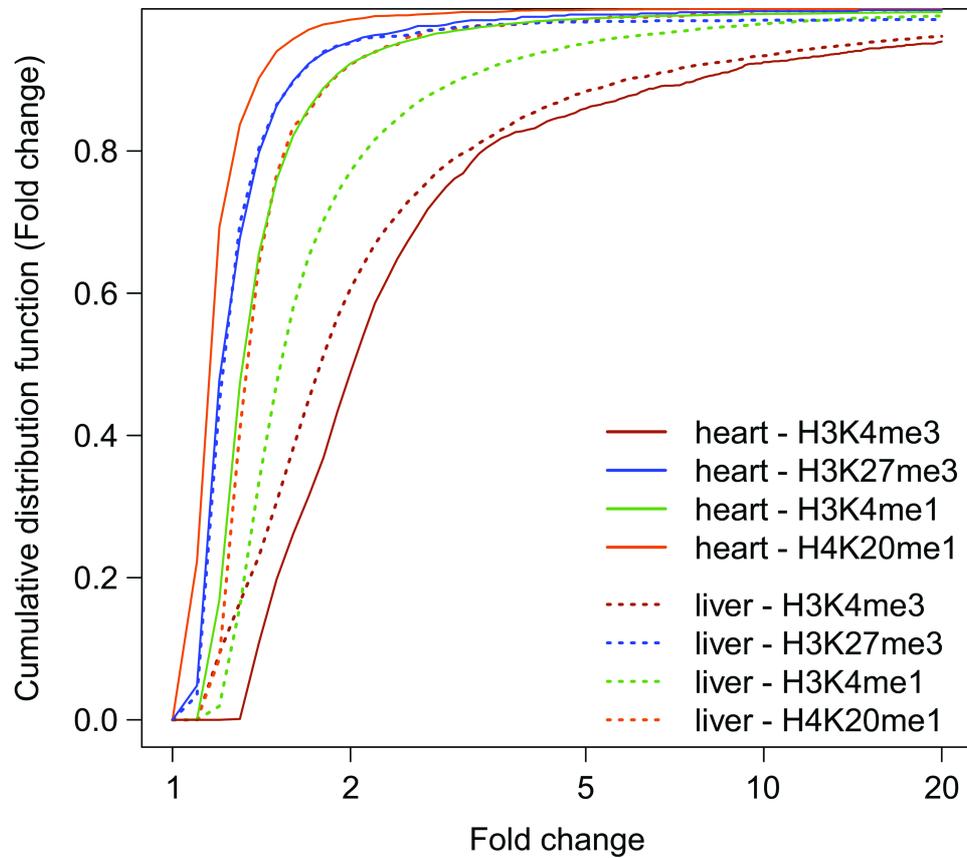
Supplementary Figure 5: Validation of differential histone modification marks. ChIP was performed in heart and liver tissue of three parental BN and three SHR rats. For each histone modification, three randomly chosen differential regions were verified using quantitative PCR (qPCR). Relative fold changes were calculated according to the $2^{-\Delta\Delta Ct}$ method with endogenous control region of *Myod1* (for H3K27me3) or *Gapdh* (for all other marks). All ChIP-qPCR results were significantly different between BN and SHR rats (n = 3 per strain, Mann-Whitney-Test p < 0.05). Data are mean ± s.e.m.
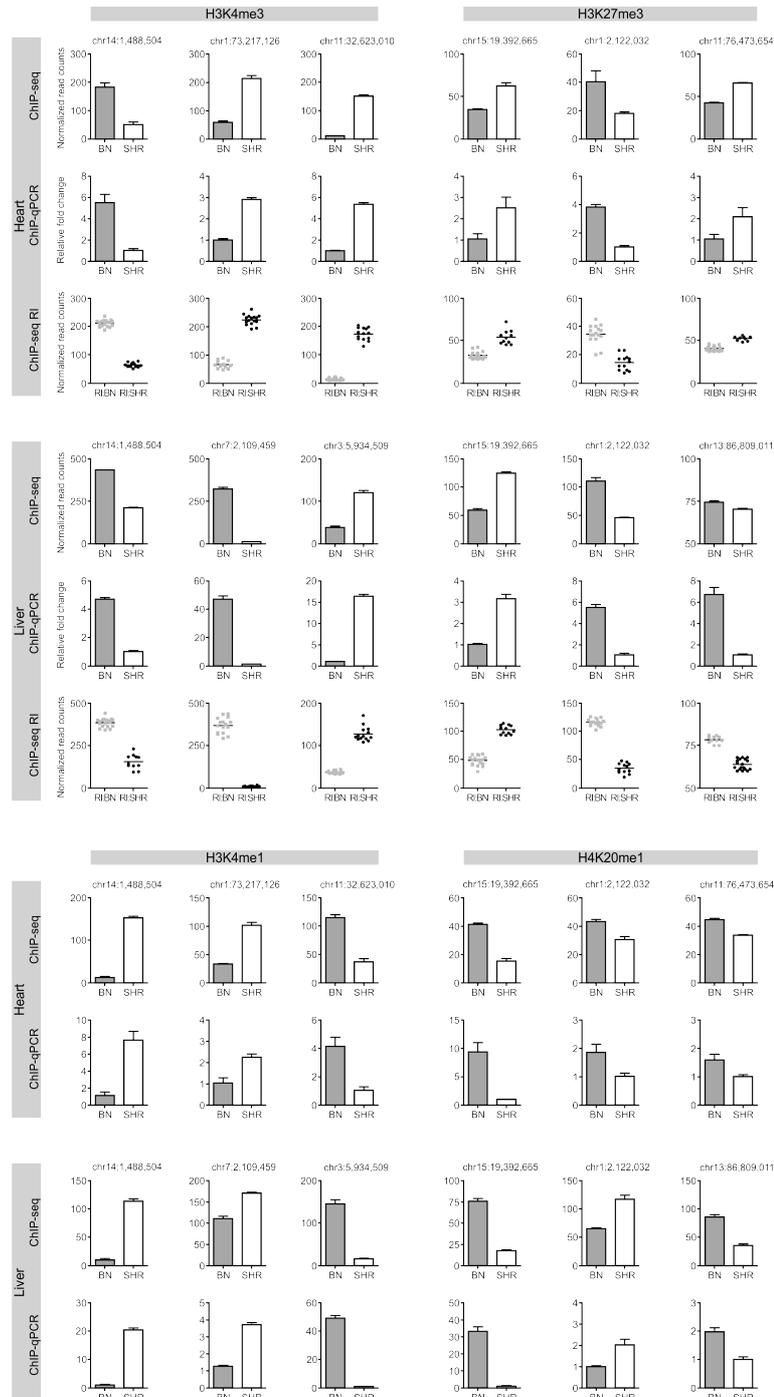
Supplementary Figure 6: Quantile-quantile plots for the QTL analyzes of histone modification traits. For each trait and each tissue we show the observed quantiles of the association statistic plotted against the quantiles of the permutation based null distribution. The traits are occupancy levels of H3k4me3 regions defined by the peak calling analysis (macs), H3K4me3 and H3K27me3 regions defined by annotations of known protein coding genes (Ensembl).
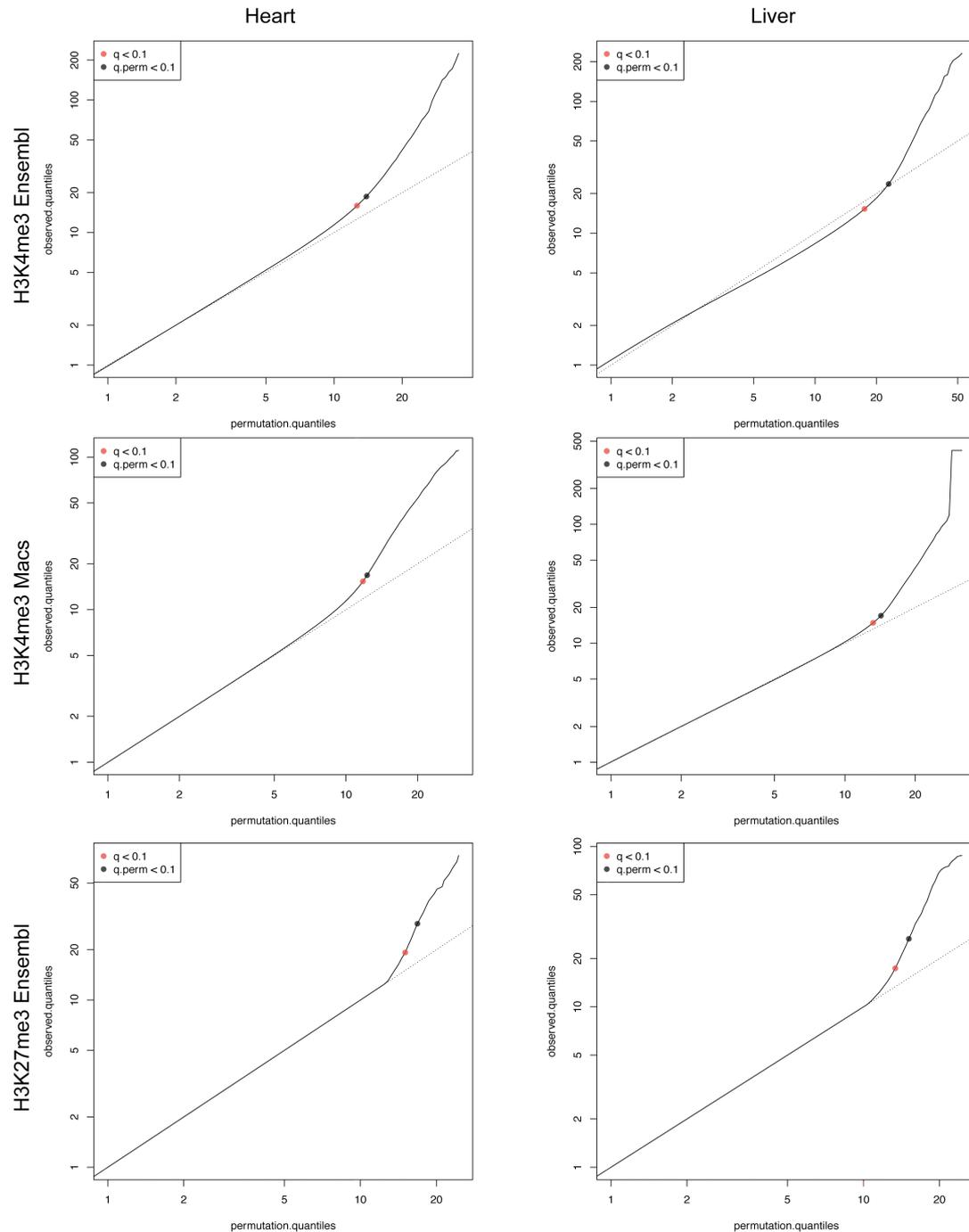
Supplementary Figure 7: Comparison of differential histone marks between RI and parental strains. Four proportional venn diagrams display absolute numbers of differential histone methylation marks in left ventricle and liver tissue of BN and SHR inbred rats which overlay identified *cis* and/or *trans* histoneQTLs in the RI panel.
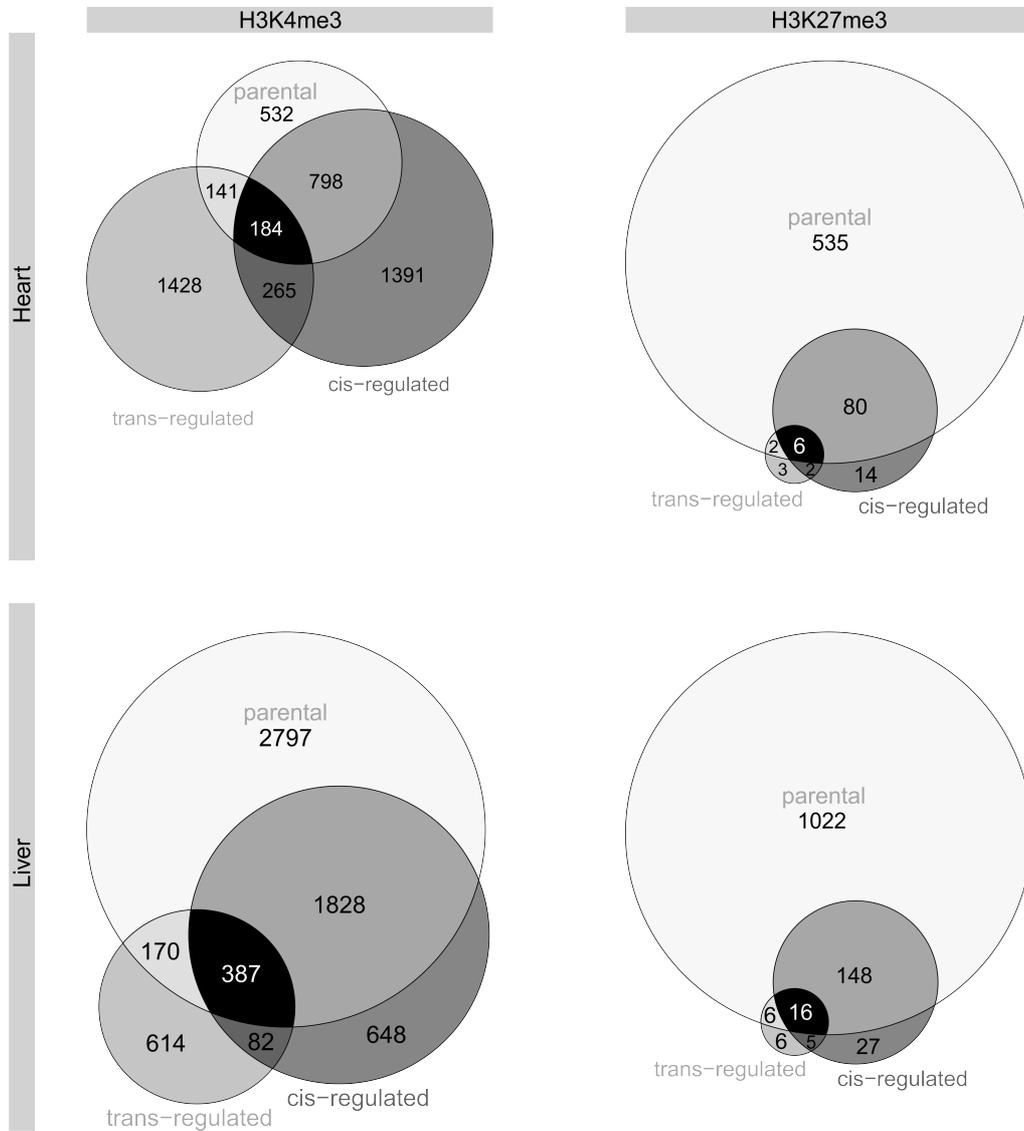
Supplementary Figure 8: Representation of competing graphical models. In addition to the null model where all variables are independent we have compared the Akaike information criterion (AIC) of all models where the SNP can directly influence histone modification and gene expression levels, and histone modification levels can directly influence gene expression.
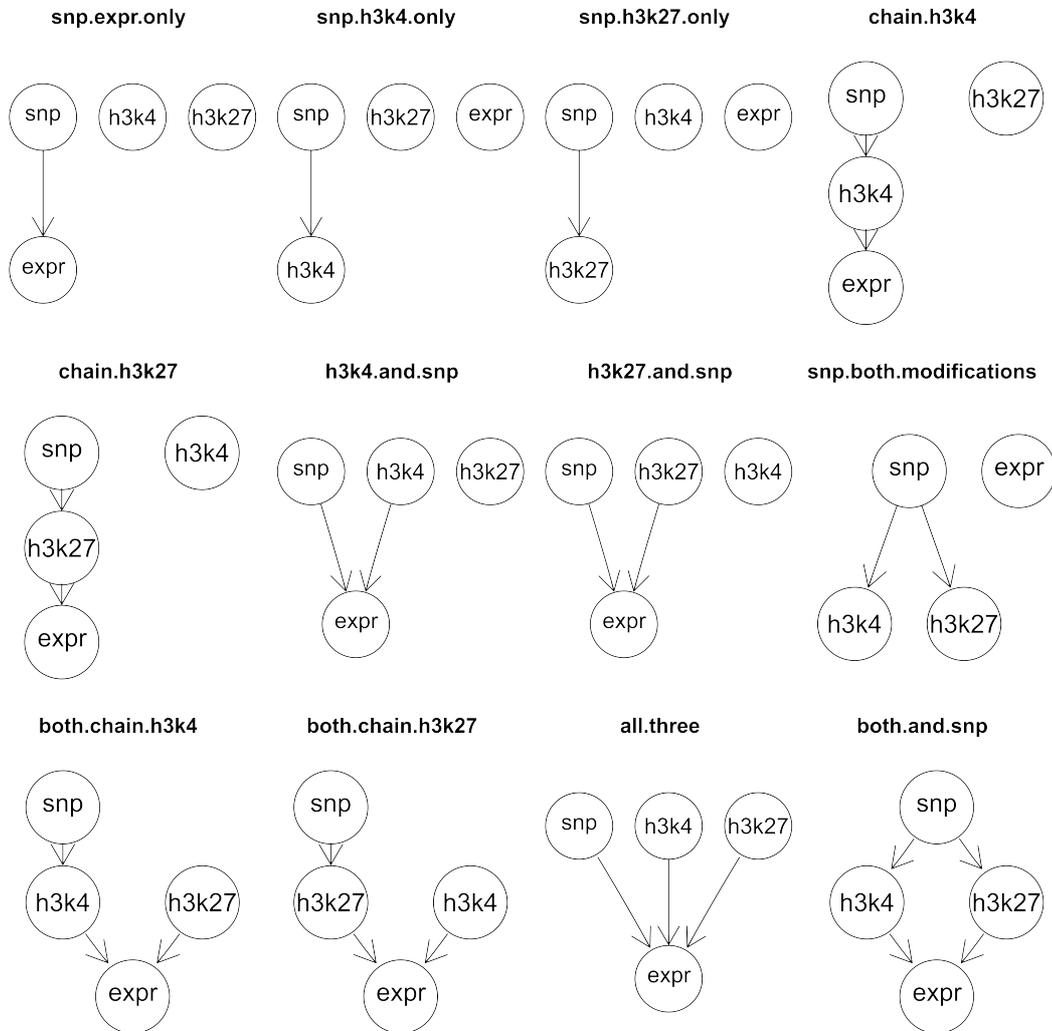
Supplementary Figure 9: Distribution of graphical models. The barplots show the percentage of genes for which each of the models from Figure 4 a-d resulted in the best fit. We estimate these proportions by aggregating the bootstrap results for each model. The proportions are provided separately for each class of genes: "No expression" are genes where histone data is available but gene expression is below the thresholds for analysis, genes of the other classes have expression data and histone data as given by the name of the class, the numbers in brackets represent the number of genes falling in each class.
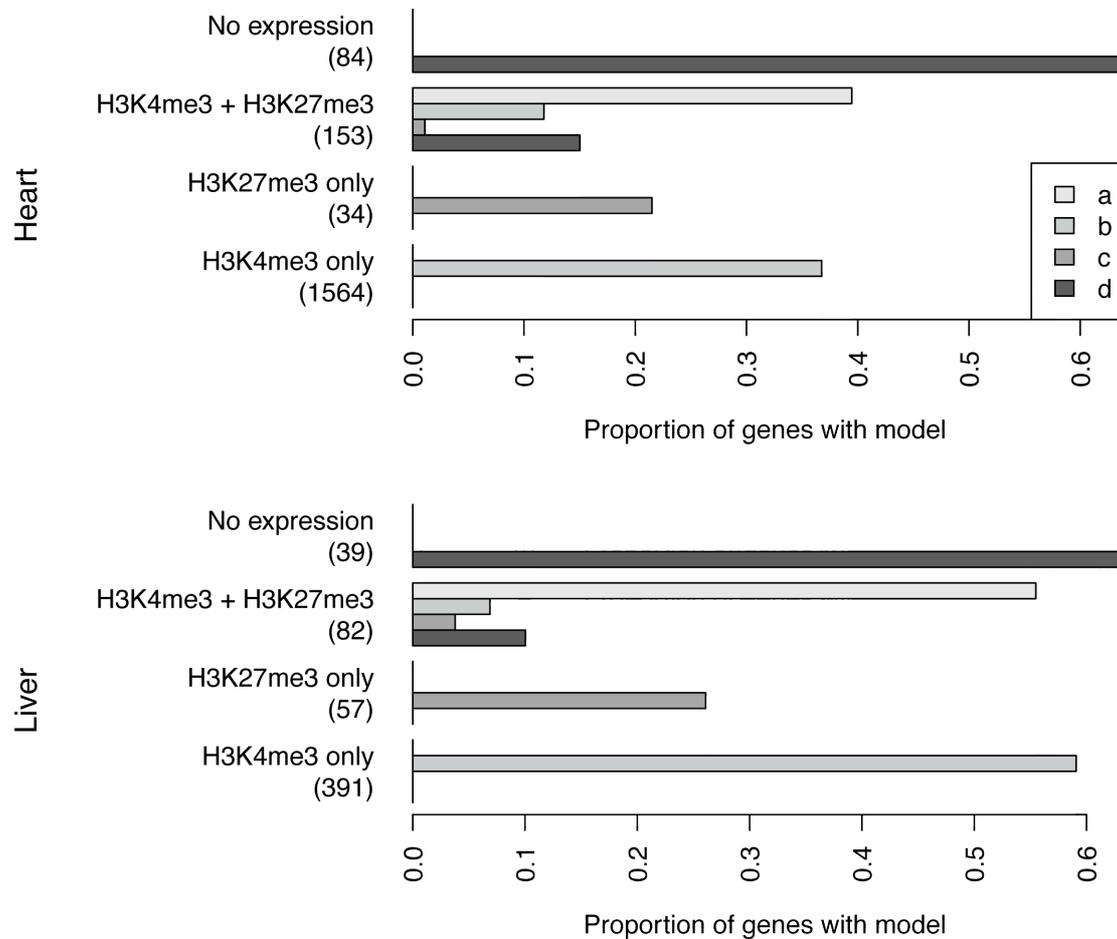
Supplementary Figure 10: Distribution of H3K4me3 peaks according to genomic features.
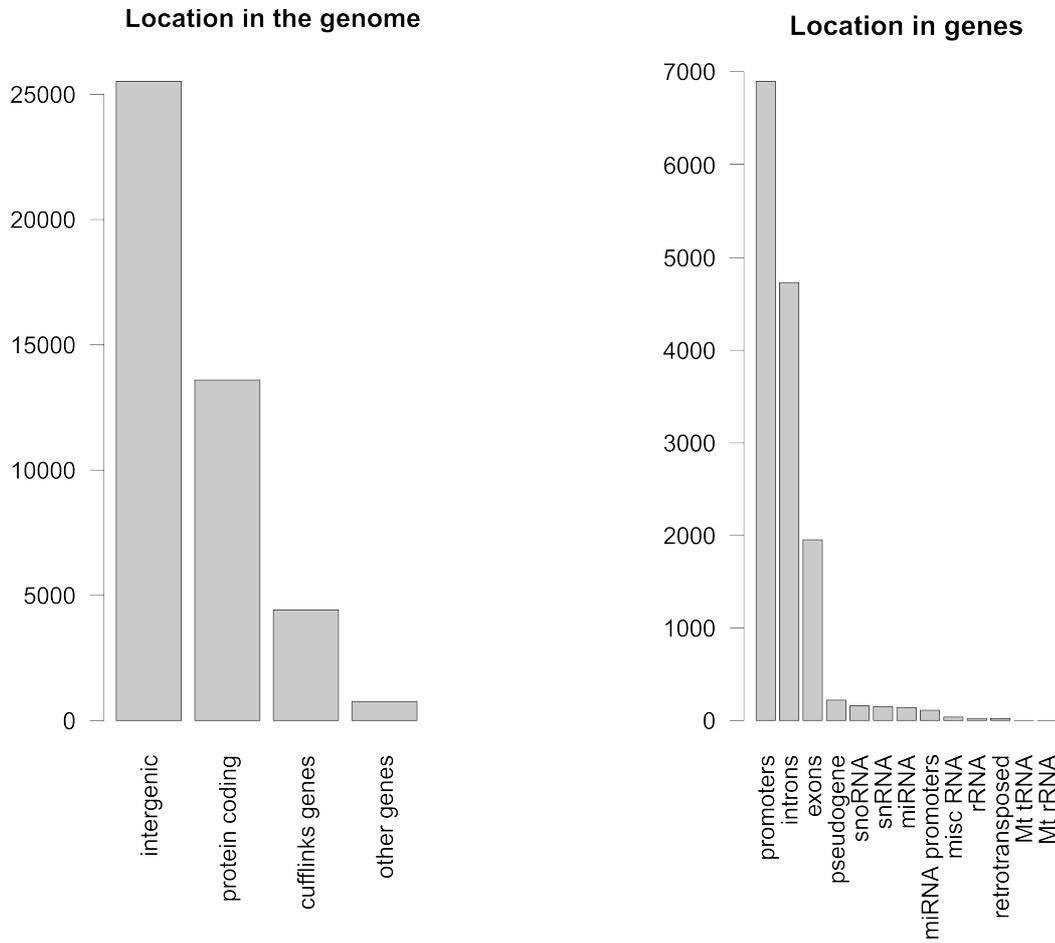
Supplementary Figure 11: Histogram of P-values for the correlation between expression and histone traits. P-values were computed for the Pearson correlation coefficients using the t-distribution. We compared several FDR methods to estimate the proportion of true null hypothesis: fdrtool (Strimmer 2008)(red), locfdr (Efron 2004)(green) and qvalue (Storey 2003)(blue). The qvalue method lead to the smallest estimates and therefore most conservative estimates of the proportion of true null hypothesis.
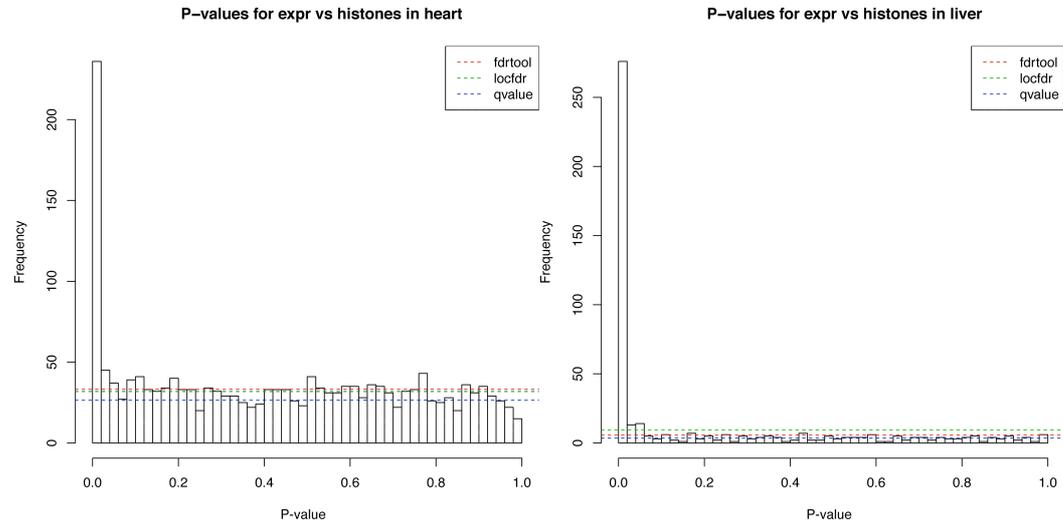
Supplementary Figure 12: Evaluation of H3K27me3 peak calls. We used a data set of qPCR validated positive and negative examples of H3K27me3 regions (Micsinai et al., 2012) for the evaluation of our HMM for peak calling in comparison to macs2 (Zhang et al 2008) and sicer (Zang et al 2009). We treated each region as a single data point and labeled it as zero or one based on the qPCR data. The corresponding ChIP-seq and control data set (GEO accessions GSM721294, GSM721306) was then processed using the standard settings of each peak caller that we compared. Predictions for each data point were obtained by averaging scores over all region calls or bins that were overlapping the validated region. Finally, we evaluated the performance of each method by their ROC curves.