Supporting Information

Figure S1. Correlation of transcript levels to modified domains: positional effect

For each of the 15 domain codes, expression levels of transcripts that were assigned to domains of this code are plotted against the absolute distance of the domain to the TSS. “CC” is the Spearman Rank Correlation Coefficient between absolute distance and expression level.

Figure S2. Verification of gene expression array data

Shown are average $\log_2$ values over two independent biological replicates. (A) Comparison of $\log_2$ intensities, the scale of the real-time PCR data was adjusted to the array intensities. (B) The $\log_2$ fold changes as determined from expression arrays in comparison to real-time PCR fold changes normalized to Hprt1.

Figure S3. Validation of ChIP enrichment

From each cell type two biological replicates (set1 and set2) were analyzed for each histone modification. Fold changes were calculated relative to input. Primers named “all_mr” were designed for sites showing enrichment on arrays (positive controls), primers named “no_mr” were designed for non-enriched sites (negative controls). (A) Validation of ChIP enrichment before linear amplification. Normal rabbit IgG ChIP was analyzed as control. ChIP with normal rabbit IgG showed no enrichment over input at any tested site. (B) Validation of modified sites identified by ChIP-chip using amplified material.

Figure S1

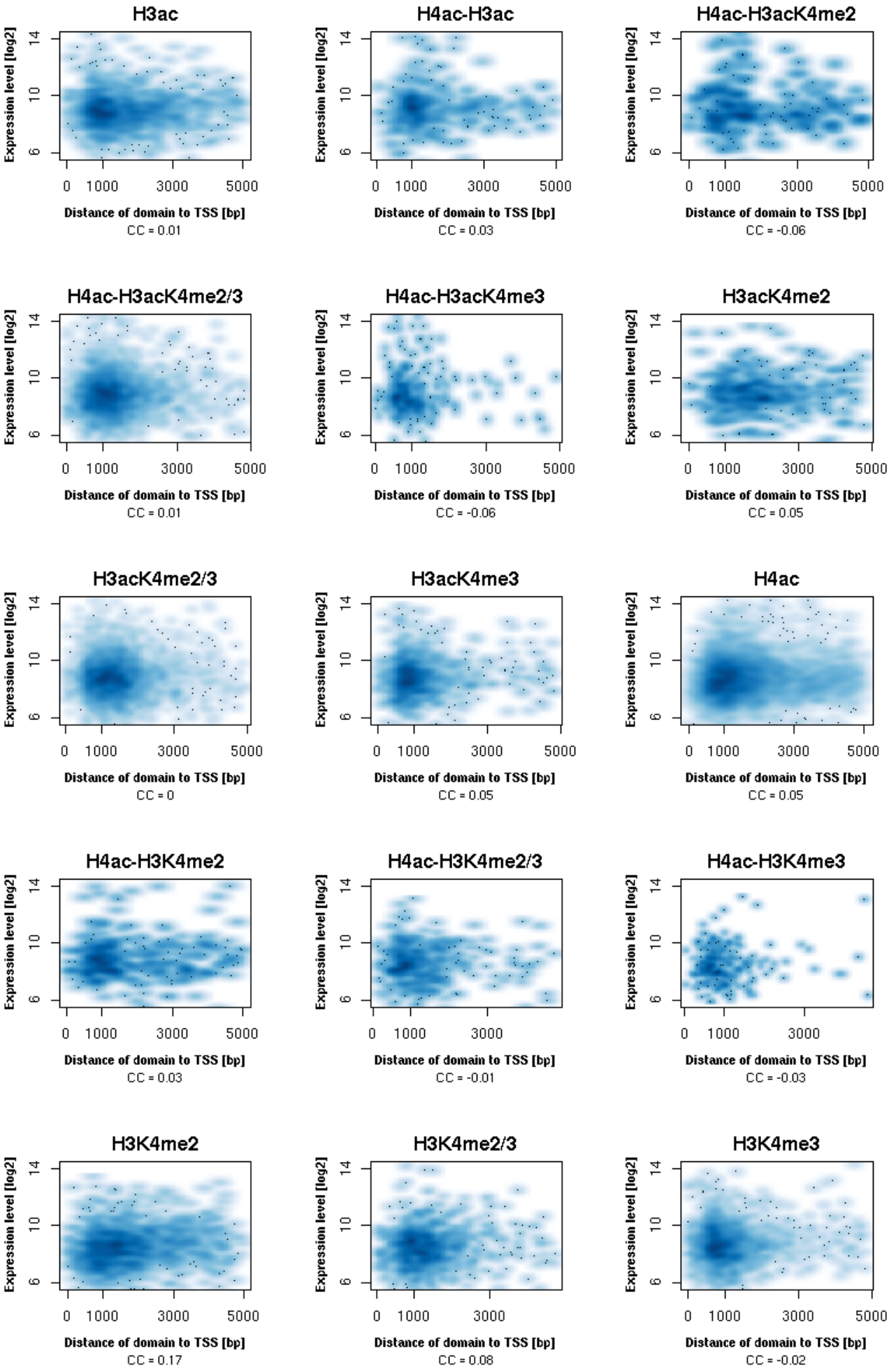

H3ac
Expression level [log2]
Distance of domain to TSS [bp]
CC = 0.01
H4ac-H3ac
CC = 0.03
H4ac-H3acK4me2
CC = -0.06
H4ac-H3acK4me2/3
CC = 0.01
H4ac-H3acK4me3
CC = -0.06
H3acK4me2
CC = 0.05
H3acK4me2/3
CC = 0
H3acK4me3
CC = 0.05
H4ac
CC = 0.05
H4ac-H3K4me2
CC = 0.03
H4ac-H3K4me2/3
CC = -0.01
H4ac-H3K4me3
CC = -0.03
H3K4me2
CC = 0.17
H3K4me2/3
CC = 0.08
H3K4me3
CC = -0.02

Figure S2A

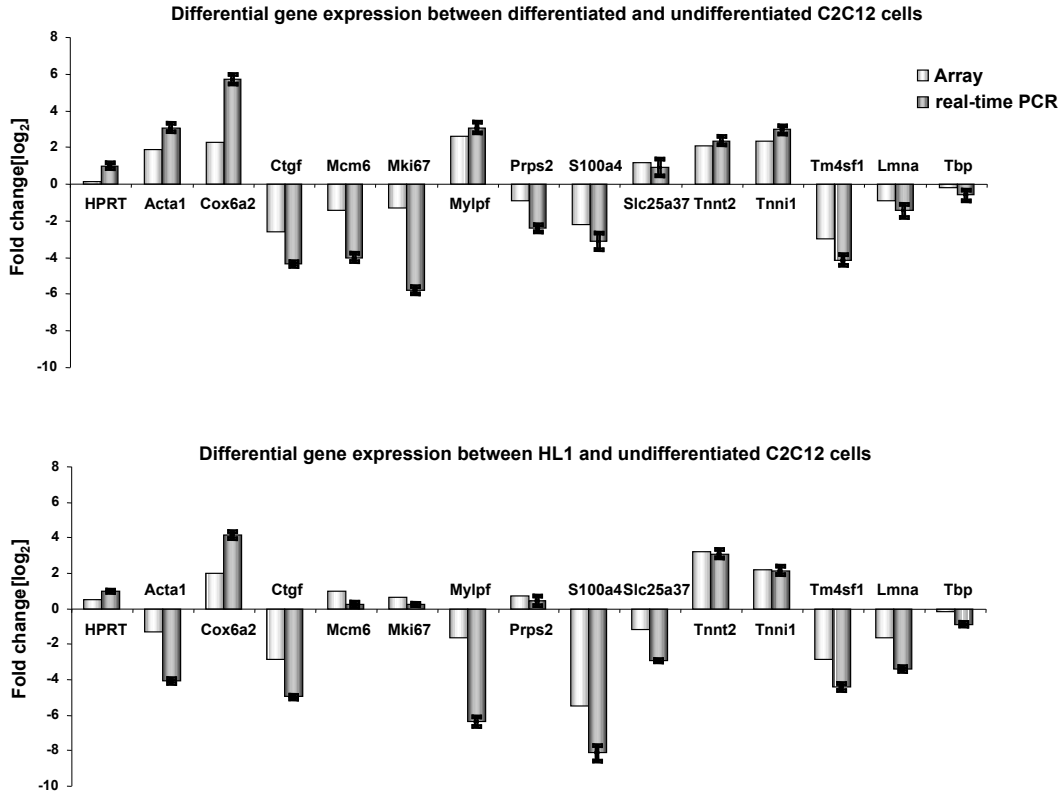

Differential gene expression between differentiated and undifferentiated C2C12 cells
Array
real-time PCR
Fold change[log2]
8
6
4
2
0
-2
-4
-6
-8
-10
HPRT
Acta1
Cox6a2
Ctgf
Mcm6
Mki67
Mylpf
Prps2
S100a4
Slc25a37
Tnnt2
Tnni1
Tm4sf1
Lmna
Tbp
Differential gene expression between HL1 and undifferentiated C2C12 cells
Fold change[log2]
8
6
4
2
0
-2
-4
-6
-8
-10
HPRT
Acta1
Cox6a2
Ctgf
Mcm6
Mki67
Mylpf
Prps2
S100a4
Slc25a37
Tnnt2
Tnni1
Tm4sf1
Lmna
Tbp

Figure S2B

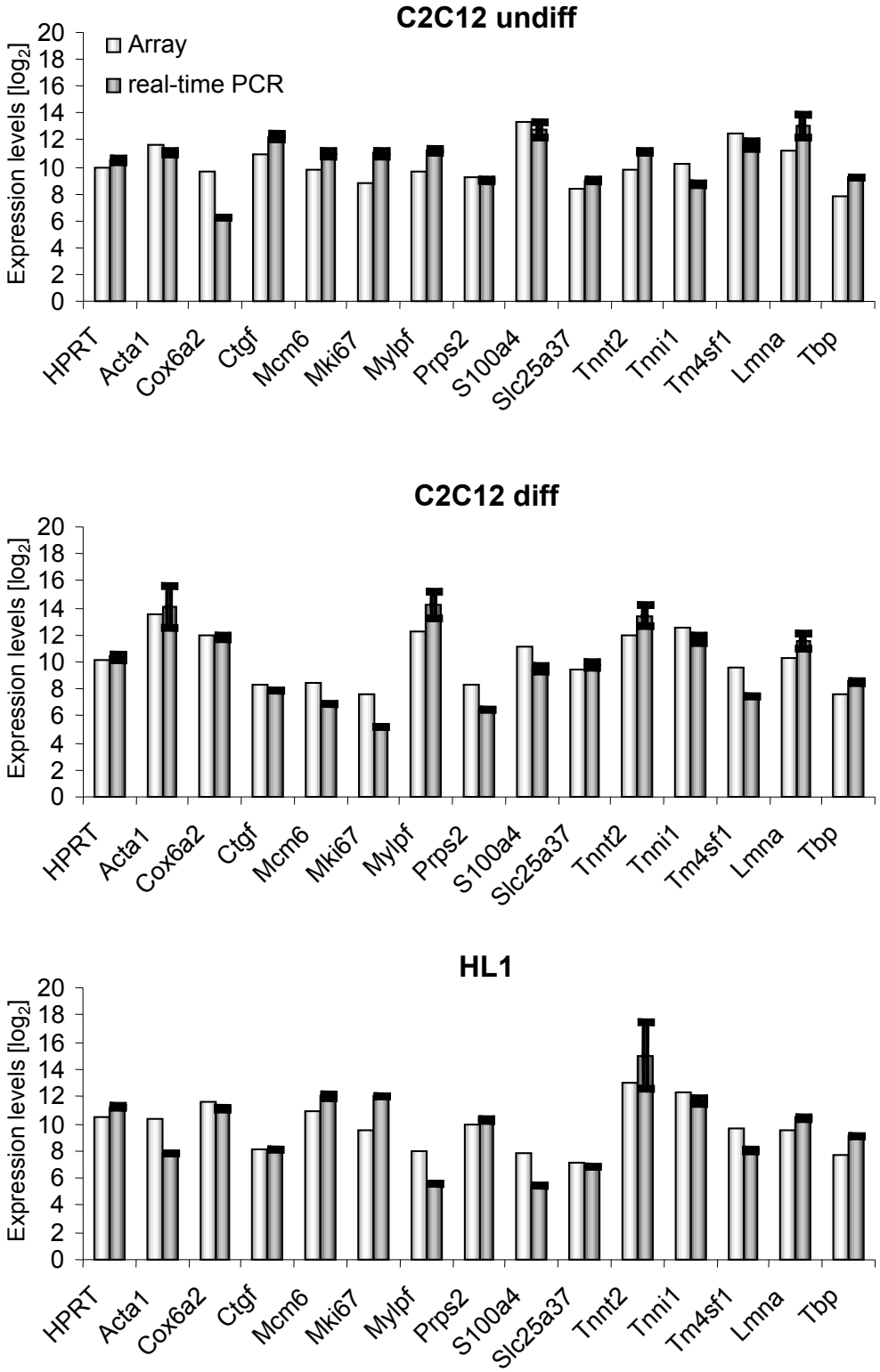
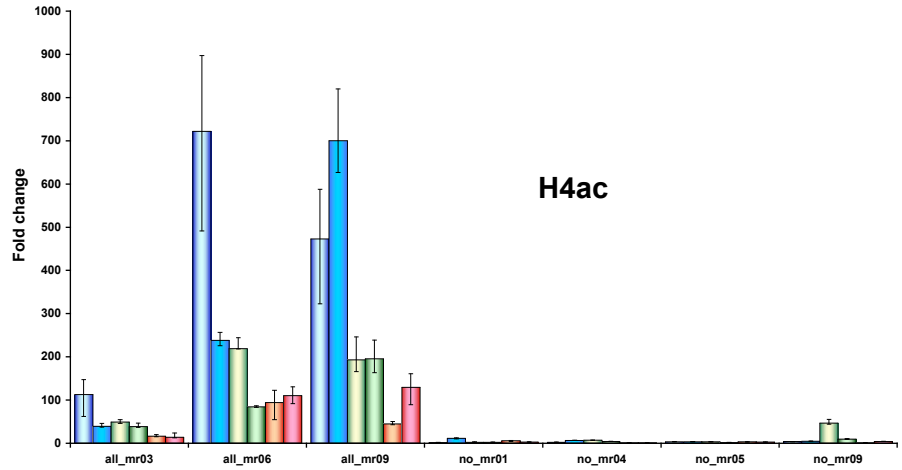
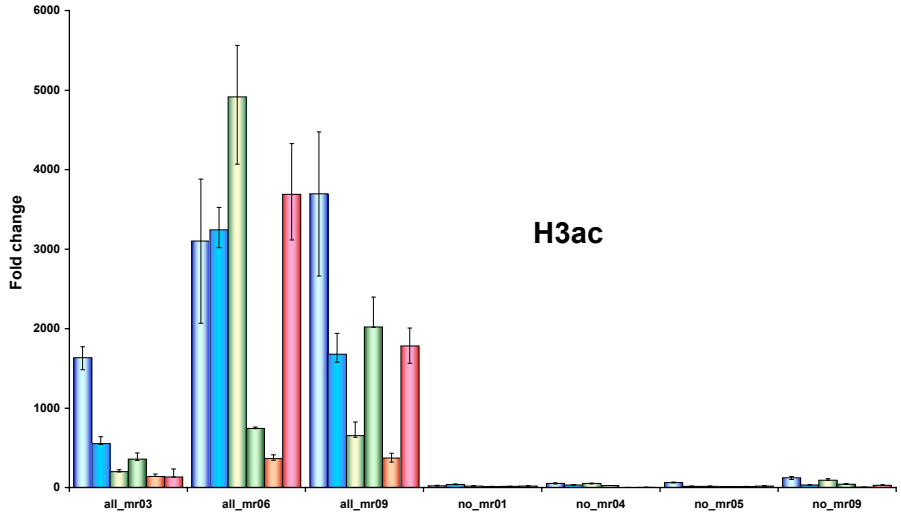
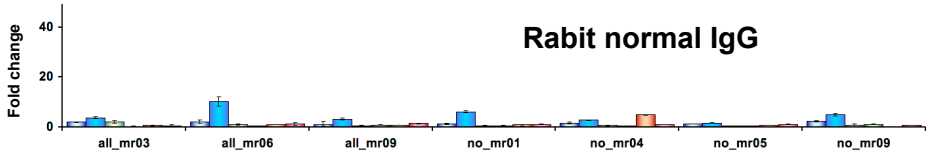
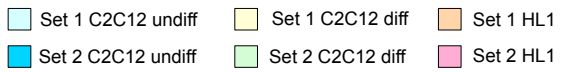

C2C12 undiff
Array
real-time PCR
Expression levels [log2]
0
2
4
6
8
10
12
14
16
18
20
HPRT
Acta1
Cox6a2
Ctgf
Mcm6
Mki67
Mylpf
Prps2
S100a4
Slc25a37
Tnnt2
Tnni1
Tm4sf1
Lmna
Tbp
C2C12 diff
Expression levels [log2]
HPRT
Acta1
Cox6a2
Ctgf
Mcm6
Mki67
Mylpf
Prps2
S100a4
Slc25a37
Tnnt2
Tnni1
Tm4sf1
Lmna
Tbp
HL1
Expression levels [log2]
HPRT
Acta1
Cox6a2
Ctgf
Mcm6
Mki67
Mylpf
Prps2
S100a4
Slc25a37
Tnnt2
Tnni1
Tm4sf1
Lmna
Tbp

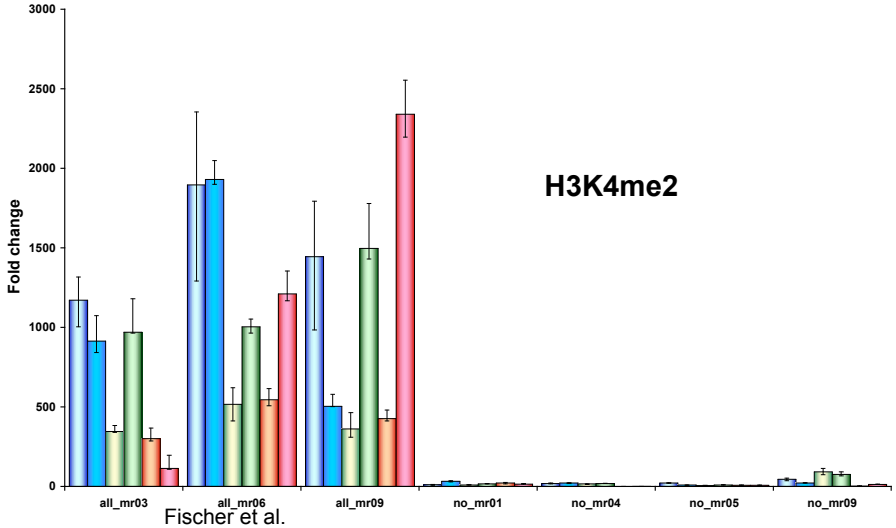
Figure S3A


H4ac
Fold change
1000
900
800
700
600
500
400
300
200
100
0
all_mr03
all_mr06
all_mr09
no_mr01
no_mr04
no_mr05
no_mr09


H3ac
Fold change
6000
5000
4000
3000
2000
1000
0
all_mr03
all_mr06
all_mr09
no_mr01
no_mr04
no_mr05
no_mr09


Rabit normal IgG
Fold change
40
20
0
all_mr03
all_mr06
all_mr09
no_mr01
no_mr04
no_mr05
no_mr09


Set 1 C2C12 undiff
Set 1 C2C12 diff
Set 1 HL1
Set 2 C2C12 undiff
Set 2 C2C12 diff
Set 2 HL1


H3K4me2
Fold change
3000
2500
2000
1500
1000
500
0
all_mr03
all_mr06
all_mr09
no_mr01
no_mr04
no_mr05
no_mr09

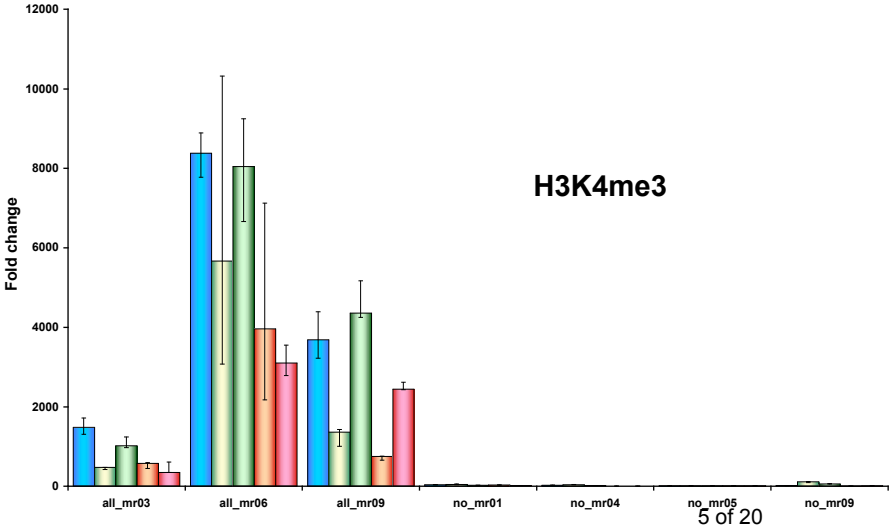

H3K4me3
Fold change
12000
10000
8000
6000
4000
2000
0
all_mr03
all_mr06
all_mr09
no_mr01
no_mr04
no_mr05
no_mr09

Figure S3B

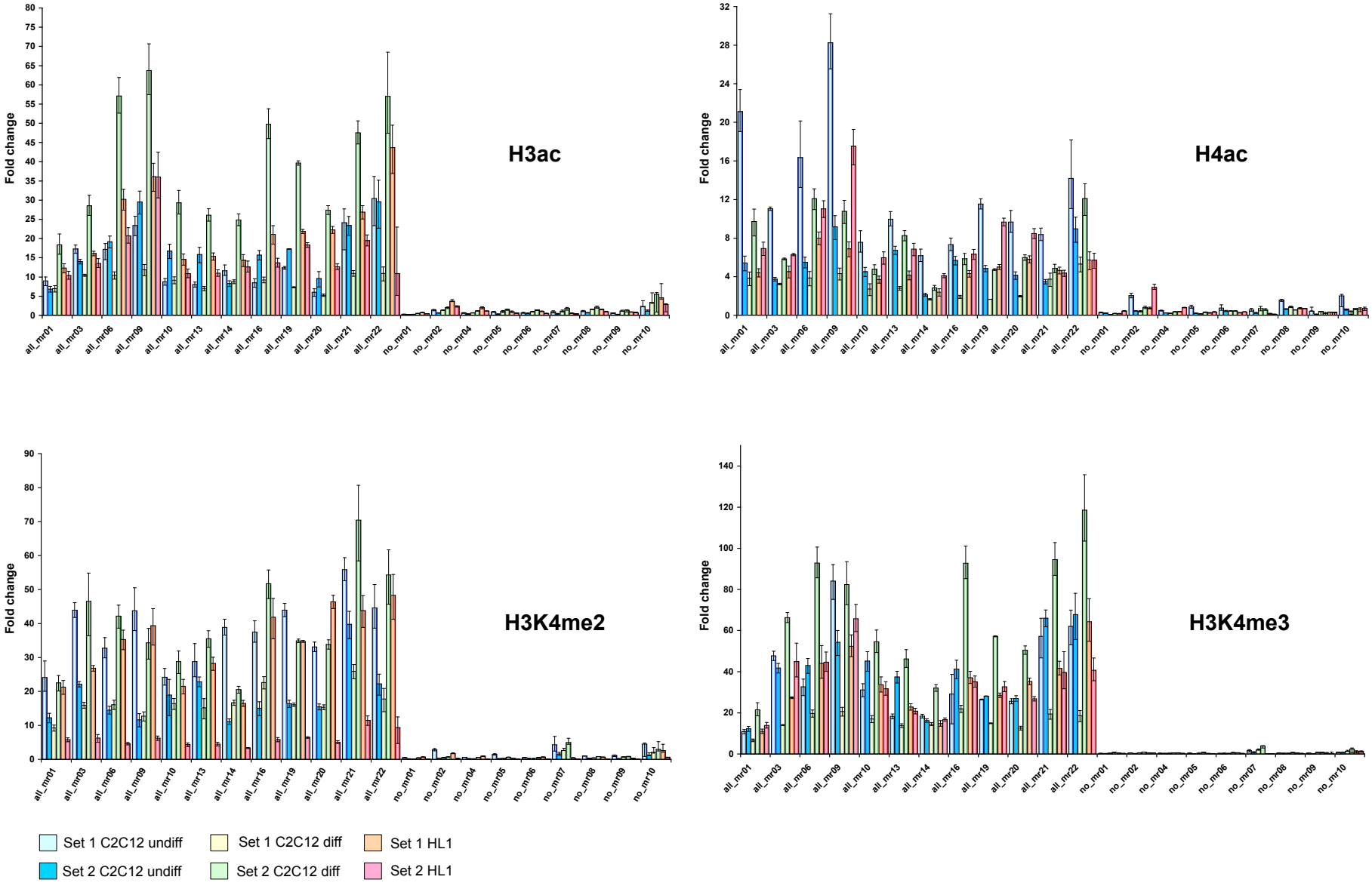

H3ac
H4ac
H3K4me2
H3K4me3
Fold change
all_mr01
all_mr03
all_mr06
all_mr09
all_mr10
all_mr13
all_mr14
all_mr16
all_mr19
all_mr20
all_mr21
all_mr22
no_mr01
no_mr02
no_mr04
no_mr05
no_mr06
no_mr07
no_mr08
no_mr09
no_mr10
Set 1 C2C12 undiff
Set 1 C2C12 diff
Set 1 HL1
Set 2 C2C12 undiff
Set 2 C2C12 diff
Set 2 HL1

Table S1. Sites enriched for histone modifications per cell type and comparison to published data

Table lists the number and median size of identified modified sites of each modification and cell type. The median site sizes and the average number of sites per gene are compared to published data (1) .

Modifications	Number of Modified Sites			Median Site Size [bp]				Average Number of Sites per Gene			
	C2C12u	C2C12d	HL1	C2C12u	C2C12d	HL1	HepG2 [1]	C2C12u	C2C12d	HL1	HepG2 [1] *
H3ac	3,059	3,248	3,210	637	645	609	703	0.34	0.36	0.35	0.50
H4ac	2,925	3,026	2,940	561	561	543	n.a.	0.32	0.33	0.32	n.a.
H3K4me2	3,297	3,205	3,378	621	613	608	605	0.36	0.35	0.37	0.31
H3K4me3	3,493	3,538	3,357	647	647	607	659	0.39	0.39	0.37	0.36

u, undifferentiated; d, differentiated ; n.a., not available

*These values were calculated from the number of modified sites as listed in Table 1 of the publication by Bernstein *et al*. [1], assuming a total of 1,397 genes on Chr 21 and Chr 22 (Ensembl v36).

1. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, et al. (2005) Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse. Cell 120: 169-181.

Table S2. Transcripts categorized by associated modified sites: *p* values for pair wise comparison of transcript categories as shown in Figure 5A

Table gives *p* values for pair wise comparison of categories as shown in Figure 5A. Transcripts were categorized according to the modification status of the respective gene into five groups corresponding to the rows and columns of the table. Expression levels of the categories were compared by two-sided two-sample Wilcoxon tests, and *p* values were adjusted for multiple testing using the Bonferroni procedure. For each comparison, the table gives sign of the difference and the *p* value: +, row category has higher levels than column category; -, row category has lower levels than column category; o, no rejection.

Transcript Category	No modification	H3K4me2	H3K4me3	H4ac
H3ac	+ ($<10^{-30}$)	+ (4×10^{-8})	+ (8×10^{-7})	+ (2×10^{-3})
H4ac	+ ($<10^{-30}$)	+ (0.04)	o	
H3K4me3	+ ($<10^{-30}$)	o		
H3K4me2	+ ($<10^{-30}$)			

Table S3. Transcripts categorized by associated modified domains: *p* values for pair wise comparison of categories as shown in Figure 5B

Table gives *p* values for pair wise comparison of categories as shown in Figure 5B. Transcripts were categorized according to presence of modified domains in the respective gene. Only the eight most frequent domain types are listed. Expression levels of the categories were compared by two-sided two-sample Wilcoxon tests, and *p* values were adjusted for multiple testing using the Bonferroni procedure. For each comparison, the table gives sign of the difference and the *p* value: +, row category has higher levels than column category; -, row category has lower levels than column category; o, no rejection.

Transcript Category	No modification	H3K4me2	H3K4me2/3	H3K4me3	H3ac-K4me3	H3ac-K4me2/3	H4acH3ac-K4me2/3	H4ac
H3ac	+ ($<10^{-30}$)	+ (1×10^{-16})	+ (2×10^{-19})	+ (6×10^{-8})	+ (3×10^{-2})	+ (4×10^{-5})	+ (2×10^{-2})	+ (9×10^{-6})
H4ac	+ ($<10^{-30}$)	+ (3×10^{-6})	+ (5×10^{-8})	o	o	o	o	
H4ac-H3ac-K4me2/3	+ ($<10^{-30}$)	+ (7×10^{-11})	+ (1×10^{-13})	+ (7×10^{-3})	o	o		
H3ac-K4me2/3	+ ($<10^{-30}$)	+ (8×10^{-9})	+ (2×10^{-11})	o	o			
H3ac-K4me3	+ (2×10^{-16})	+ (2×10^{-6})	+ (8×10^{-8})	o				
H3K4me3	+ (6×10^{-12})	+ (2×10^{-4})	+ (4×10^{-5})					
H3K4me2/3	o	o						
H3K4me2	o							

Table S4. Gene Ontology (GO) association for differentially expressed genes

GO association for genes differentially expressed between C2C12 undifferentiated and differentiated cells (A and B) and between C2C12 undifferentiated cells and HL-1 cells (C and D).

Table S4A. GO association of genes upregulated in C2C12 differentiated cells compared to undifferentiated cells

ID	*p* value	Odds Ratio	Expected Count	Observed Count	Size	Term	Ontology
GO:0006937	2.1E-06	31.74	1	6	10	regulation of muscle contraction	Biological Process
GO:0006986	7.3E-04	9.39	1	5	16	response to unfolded protein	Biological Process
GO:0005861	1.2E-08	Inf	0	6	6	troponin complex	Cellular Component
GO:0030017	9.7E-06	10.75	1	8	24	sarcomere	Cellular Component
GO:0005783	1.2E-05	4.52	5	16	100	endoplasmic reticulum	Cellular Component
GO:0043292	2.6E-05	9.03	1	8	27	contractile fiber	Cellular Component
GO:0015629	2.9E-05	5.83	3	11	53	actin cytoskeleton	Cellular Component

Table S4B. GO association of genes downregulated in C2C12 differentiated cells compared to undifferentiated cells

ID	*p* value	Odds Ratio	Expected Count	Observed Count	Size	Term	Ontology
GO:0000279	9.0E-12	18.05	2	16	29	M phase	Biological Process
GO:0000278	6.6E-11	14.63	3	16	32	mitotic cell cycle	Biological Process
GO:0006260	1.6E-09	12.01	3	15	33	DNA replication	Biological Process
GO:0051301	2.7E-09	15.31	2	13	31	cell division	Biological Process
GO:0043283	7.9E-07	3.09	17	40	264	biopolymer metabolism	Biological Process
GO:0007067	1.5E-06	13.51	1	9	21	mitosis	Biological Process
GO:0007059	2.5E-05	19.65	1	6	10	chromosome segregation	Biological Process
GO:0051241	3.0E-05	Inf	0	4	4	negative regulation of organismal physiological process	Biological Process
GO:0006334	9.8E-05	13.08	1	6	12	nucleosome assembly	Biological Process
GO:0007160	1.0E-04	21.61	1	5	8	cell-matrix adhesion	Biological Process
GO:0006333	1.1E-04	9.23	1	7	17	chromatin assembly or disassembly	Biological Process
GO:0006268	1.4E-04	51.38	0	4	5	DNA unwinding during replication	Biological Process

GO:0051726	2.2E-04	4.20	4	12	51	regulation of cell cycle	Biological Process
GO:0006270	4.0E-04	25.67	0	4	6	DNA replication initiation	Biological Process
GO:0000910	4.0E-04	25.67	0	4	6	cytokinesis	Biological Process
GO:0007126	4.0E-04	12.94	1	5	10	meiosis	Biological Process
GO:0051321	4.0E-04	12.94	1	5	10	meiotic cell cycle	Biological Process
GO:0007088	4.1E-04	Inf	0	3	3	regulation of mitosis	Biological Process
GO:0007001	4.4E-04	5.00	3	9	33	chromosome organization and biogenesis (sensu Eukaryota)	Biological Process
GO:0006323	5.7E-04	4.79	3	9	34	DNA packaging	Biological Process
GO:0006461	6.6E-04	7.82	1	6	28	protein complex assembly	Biological Process
GO:0006139	8.2E-04	2.10	20	37	310	nucleobase, nucleoside, nucleotide and nucleic acid metabolism	Biological Process
GO:0005694	4.9E-08	7.63	4	16	54	chromosome	Cellular Component
GO:0043228	2.4E-07	3.47	14	35	208	non-membrane-bound organelle	Cellular Component
GO:0005634	5.3E-07	2.98	23	50	378	nucleus	Cellular Component
GO:0000793	9.1E-07	80.55	1	6	7	condensed chromosome	Cellular Component
GO:0000776	1.3E-04	52.49	0	4	5	kinetochore	Cellular Component
GO:0000786	1.9E-04	16.55	1	5	9	nucleosome	Cellular Component
GO:0005876	3.9E-04	Inf	0	3	3	spindle microtubule	Cellular Component
GO:0005657	8.2E-04	17.47	1	4	7	replication fork	Cellular Component

Table S4C. GO association of genes upregulated in HL-1 cells compared to undifferentiated C2C12 cells

ID	*p* value	Odds Ratio	Expected Count	Observed Count	Size	Term	Ontology
GO:0006941	2.0E-04	34.15	1	5	6	striated muscle contraction	Biological Process
GO:0015980	2.0E-04	3.96	6	14	39	energy derivation by oxidation of organic compounds	Biological Process
GO:0044275	2.7E-04	6.25	3	9	19	cellular carbohydrate catabolism	Biological Process

GO:0006096	3.8E-04	6.92	2	8	16	glycolysis	Biological Process
GO:0005975	6.8E-04	2.95	8	17	58	carbohydrate metabolism	Biological Process

Table S4D. GO association of genes downregulated in HL-1 cells compared to undifferentiated C2C12 cells

ID	*p* value	Odds Ratio	Expected Count	Observed Count	Size	Term	Ontology
GO:0007275	2.5E-07	2.66	29	56	235	development	Biological Process
GO:0007155	4.0E-07	4.41	9	24	65	cell adhesion	Biological Process
GO:0000902	3.5E-06	4.88	6	18	45	cellular morphogenesis	Biological Process
GO:0007160	4.1E-06	48.78	1	7	8	cell-matrix adhesion	Biological Process
GO:0042060	5.1E-05	11.18	2	8	13	wound healing	Biological Process
GO:0006817	1.2E-04	12.16	2	7	11	phosphate transport	Biological Process
GO:0007399	1.5E-04	3.34	8	18	57	nervous system development	Biological Process
GO:0001525	1.7E-04	4.72	4	12	30	angiogenesis	Biological Process
GO:0009887	1.9E-04	3.13	9	19	63	organ morphogenesis	Biological Process
GO:0030216	2.8E-04	Inf	1	4	4	keratinocyte differentiation	Biological Process
GO:0001568	3.5E-04	3.69	6	14	41	blood vessel development	Biological Process
GO:0048637	8.7E-04	6.93	2	7	14	skeletal muscle development	Biological Process
GO:0031012	2.5E-08	6.11	7	22	47	extracellular matrix	Cellular Component
GO:0005576	2.7E-06	2.42	30	54	223	extracellular region	Cellular Component
GO:0005581	1.4E-04	19.26	1	6	8	collagen	Cellular Component
GO:0005604	3.8E-04	9.02	2	7	12	basement membrane	Cellular Component
GO:0005615	4.5E-04	1.99	27	44	201	extracellular space	Cellular Component

Table S5. Linear model and obtained coefficients

Given is a linear model relating absolute expression level to cell type, presence of modified sites, median probe GC content and interactions. The table specifies for each predictor variable the coefficient estimate, its standard error and *p* value for the null hypothesis that the coefficient is equal to 0.

The model (in S-plus/R formula notation)

$$y \sim H3ac + H4ac + H3K4me2 + H3K4me3 + GC + H3ac{:}H4ac + H4ac{:}H3K4me2 + H4ac{:}H3K4me3 + H4ac{:}H3K4me2{:}H3K4me3 + H3ac{:}H4ac{:}H3K4me2 + H3ac{:}H4ac{:}H3K4me3 + H3ac{:}H4ac{:}H3K4me2{:}H3K4me3 + H3ac{:}H3K4me2{:}H3K4me3 + H3ac{:}H3K4me2 + H3ac{:}H3K4me3 + H3K4me2{:}H3K4me3$$

where

y: expression level of transcript in cell line
H3ac: indicator variable for transcript's associated modification H3ac; it is 1 if at least one H3ac is associated to the transcript, 0 otherwise.
H4ac, H3K4me2, H3K4me3: analogous to H3ac
GC : median percent GC content of all expression microarray probes mapped to transcript
cell.type : one of "C2C12U", "C2C12D" or "HL1"

The expression "A:B" denotes the interaction term between predictors A and B, and the function "lm" of R version 2.4 was used to fit the model.

Coefficients of model

	Estimate	Std.Error	*t* value	*p* value
Intercept	4.26	0.06	76.24	$< 2*10^{-16}$
H3ac	**0.58**	**0.05**	**11.07**	**$< 2*10^{-16}$**
H4ac	**0.38**	**0.03**	**13.28**	**$< 2*10^{-16}$**
H3K4me2	0.06	0.05	1.19	1
H3K4me3	**0.39**	**0.04**	**9.01**	**$< 2*10^{-16}$**
GC	8.22	0.02	75.22	$< 2*10^{-16}$
cell.type.C2C12U	-0.02	0.02	-0.99	1
cell.type.C2C12D	0	0.02	-0.25	1
H3ac:H4ac	-0.23	0.1	-2.34	0.37
H4ac:H3K4me2	0.08	0.09	0.93	0.35
H4ac:H3K4me3	**-0.33**	**0.09**	**-4.44**	**$3.1*10^{-3}$**
H3ac:H3K4me2	-0.23	0.11	-2.09	0.7
H3ac:H3K4me3	**-0.37**	**0.11**	**-3.11**	**$1.7*10^{-4}$**
H3K4me2:H3K4me3	**-0.33**	**0.08**	**-4.1**	**$7.9*10^{-4}$**
H4ac:H3K4me2:H3K4me3	0.05	0.15	0.31	1
H3ac:H4ac:H3K4me2	-0.11	0.18	-0.63	1
H3ac:H4ac:H3K4me3	0.24	0.16	2.33	0.38
H3ac:H3K4me2:H3K4me3	**0.46**	**0.14**	**3.4**	**$1.3*10^{-2}$**
H3ac:H4ac:H3K4me2:H3K4me3	-0.08	0.23	-0.33	1

p values have been corrected for multiple testing using the Bonferroni procedure.

Certain interaction terms between modifications are significantly different from zero. This confirms the non-additivity of the modification effects.

Tables S6. Contingency Tables for Differential Upregulation versus Modification Gains

For each of the four histone modifications, a contingency table relates the proportion of transcripts that are differentially upregulated to the proportion of transcripts that gain the respective modification during differentiation.

The following tables contrast differential upregulation of transcripts in differentiated C2C12 (rows) against whether these transcripts gain histone modification marks during differentiation (columns).

Table S6A. H4ac

	H4ac no change	H4ac gain
No diff. expression	10025	825
Significant up	97	**29**

Table S6B. H3ac

	H3ac no change	H3ac gain
No diff. expression	10267	583
Significant up	114	12

Table S6C. H3K4me2

	H3K4me2 no change	H3K4me2 gain
No diff. expression	10259	591
Significant up	109	17

Table S6D: H3K4m3

	H3K4me3 no change	H3K4me3 gain
No diff. expression	10281	569
Significant up	112	14

Table S7. Logistic regression model and obtained coefficients
Associating modification gains during differentiation to differentially up-regulated genes.

Logistic regression analysis of binary indicator variable for upregulation between undifferentiated and differentiated C2C12 cells against gain of modified sites. The table gives for each predictor variable the coefficient estimate, its standard error of the *p* value for the null hypothesis that the coefficient is equal to 0.

The model (in S-plus/R formula notation)

dy ~ dH3ac + dH4ac + dH3K4me2 + dH3K4me3

where

dy: indicator variable: 1 if the transcript is found at significantly higher level in differentiated C2C12 cells than in undifferentiated cells, 0 otherwise
dH3ac: factor variable for transcript's H3ac modification change with two levels: gain or no change. For gain, the transcript had no H3ac modification in undifferentiated cells but in differentiated cells.
dH4ac, dH3K4me2, dH3K4me3: analogous to *dH3ac*

Because individual observations are actually matched pairs of transcripts before and after differentiation and the median probe GC content stays constant during differentiation, we did not include it as a predictor in this model.

The function "glm" of R version 2.4 with option family="binomial" was used to fit the model.

Coefficients of model

	Estimate	Std.Error	z value	*p* value
Intercept	-4.73	0.11	-43.72	$<2*10^{-16}$
H3ac.gain	0.12	0.33	0.35	1
H4ac.gain	**1.18**	**0.22**	**5.55**	$\mathbf{1.4*10^{-7}}$
H3K4me2.gain	0.73	0.3	2.41	0.08
H3K4me3.gain	0.35	0.33	1.08	1

p values have been corrected for multiple testing using the Bonferroni procedure.

Gain of H4ac is significantly associated with differentially upregulated genes. 62 out of 126 up-regulated transcripts, however, do not show any modification gains (significant intercept).

Table S8. Logistic regression model and obtained coefficients

Associating modification losses to differentially down-regulated genes.

Logistic regression analysis of binary indicator variable for downregulation between undifferentiated and differentiated C2C12 cells against loss of modified sites. The table gives for each predictor variable the coefficient estimate, its standard error of the *p* value for the null hypothesis that the coefficient is equal to 0.

The model (in S-plus/R formula notation)

dy ~ dH3ac + dH4ac + dH3K4me2 + dH3K4me3

where

dy: indicator variable: 1 if the transcript is found at significantly lower level in differentiated C2C12 cells than in undifferentiated cells, 0 otherwise
dH3ac: factor variable for transcript's H3ac modification change with two levels: loss or no change. For loss, the transcript had a H3ac modification in undifferentiated cells but not in differentiated cells.
dH4ac, dH3K4me2, dH3K4me3: analogous to *dH3ac*

The function "glm" of R version 2.4 with option family="binomial" was used to fit the model.

Coefficients of model

	Estimate	Std.Error	z value	*p* value
Intercept	-4.23	0.09	-49.31	$<2*10^{-16}$
H3ac.loss	0.49	0.3	1.67	0.48
H4ac.loss	0.18	0.27	0.69	1
H3K4me2.loss	-0.04	0.31	-0.14	1
H3K4me3.loss	0.64	0.28	2.28	0.11

p values have been corrected for multiple testing using the Bonferroni procedure

Table S9. List of primers used for verification of microarray expression analysis

All primers are exon spanning. f - forward primer; r – reverse primer

MGI Symbol	ENSEMBL Transcript ID	Primer Name	Sequence	Partner Primers	Orientation
Acta1	ENSMUST00000034453	acta1_rt_m_f	TTGTGTGTGACAACGGCTCTG	acta1_rt_m_r	f
Acta1	ENSMUST00000034453	acta1_rt_m_r	ACCCACGTAGGAGTCCTTCTGA	acta1_rt_m_f	r
Cox6a2	ENSMUST00000033049	Cox6a2-exp-f.1	CCCAGAGTTCATCCCGTATCA	Cox6a2-exp-r.1	f
Cox6a2	ENSMUST00000033049	Cox6a2-exp-r.1	TGGAAAAGCGTGTGGTTGC	Cox6a2-exp-f.1	r
Ctgf	ENSMUST00000020171	Ctgf-exp-f.1	CATCTCCACCCGAGTTACCAA	Ctgf-exp-r.1	f
Ctgf	ENSMUST00000020171	Ctgf-exp-r.1	TGTCCGGATGCACTTTTTGC	Ctgf-exp-f.1	r
Hprt1	ENSMUST00000026723	hprt_m_f	AAACAATGCAAACTTTGCTTTCC	hprt_m_r	f
Hprt1	ENSMUST00000026723	hprt_m_r	GGTCCTTTTCACCAGCAAGCT	hprt_m_f	r
Lmna	ENSMUST00000029699 ENSMUST00000036252	lmna_m_f1	CGCAACAAGTCCAACGAGG	lmna_m_r1	f
Lmna	ENSMUST00000029699 ENSMUST00000036252	lmna_m_r1	TGGGAAGCGATAGGTCATCA	lmna_m_f1	r
Mcm6	ENSMUST00000027601	Mcm6-exp-f.1	TCTTCCTTGCCT GCCATGT	Mcm6-exp-r.1	f
Mcm6	ENSMUST00000027601	Mcm6-exp-r.1	TCTCAGCGGTCTGTTCCTCATC	Mcm6-exp-f.1	r
Mki67	ENSMUST00000033310	Mki67-exp-f.1	CTGTGAGGCTGAGACATGGAGA	Mki67-exp-r.1	f
Mki67	ENSMUST00000033310	Mki67-exp-r.1	TGGCTTGCTTCCATCCTCAT	Mki67-exp-f.1	r
Mylfp	ENSMUST00000032910	Mylpf-exp-f.1	AGCTCCAACGTCTTCTCCATGT	Mylpf-exp-r.1	f
Mylfp	ENSMUST00000032910	Mylpf-exp-r.1	TCGATAATGCCATCCCTGTTC	Mylpf-exp-f.1	r
Prps2	ENSMUST00000026839	Prps2-exp-f.1	AGATGCTGGAGGAGCCAAAA	Prps2-exp-r.1	f
Prps2	ENSMUST00000026839	Prps2-exp-r.1	CCATCCGGTCCACTTCATTT	Prps2-exp-f.1	r
S100a4	ENSMUST00000001046	S100a4-exp-f.1	GCTCAAGGAGCTACTGACCAGG	S100a4-exp-r.1	f
S100a4	ENSMUST00000001046	S100a4-exp-r.1	CCAAGTTGCTCATCACCTTCTG	S100a4-exp-f.1	r
Slc25a37	ENSMUST00000037064	Slc25a37-exp-f.1	CCACCCTACTCCACGATGCA	Slc25a37-exp-r.1	f
Slc25a37	ENSMUST00000037064	Slc25a37-exp-r.1	CGCCACACTGTCCGGATACA	Slc25a37-exp-f.1	r
Tbp	ENSMUST00000014911 ENSMUST00000039079 ENSMUST00000080441	tbp_m_f1	TGCCACACCAGCTTCTGAGA	tbp_m_r1	f
Tbp	ENSMUST00000014911 ENSMUST00000039079 ENSMUST00000080441	tbp_m_r1	GATGACTGCAGCAAATCGCTT	tbp_m_f1	r
Tm4sf1	ENSMUST00000029376	Tm4sf1-exp-f.1	TACGAAAACTACGGCAAGCG	Tm4sf1-exp-r.1	f
Tm4sf1	ENSMUST00000029376	Tm4sf1-exp-r.1	CACAGTAAGCAGATCCCACGAT	Tm4sf1-exp-f.1	r
Tnni1	ENSMUST00000027674	tnni1_rt_m_f	GCTCTAAGCACAAGGTGTCCAT	tnni1_rt_m_r	f
Tnni1	ENSMUST00000027674	tnni1_rt_m_r	TTCCTCCAGTCTCCTACCTCGA	tnni1_rt_m_f	r
Tnnt2	ENSMUST00000027671	Tnnt2-exp-f.1	CAGACTCTGATCGAGGCTCACT	Tnnt2-exp-r.1	f
Tnnt2	ENSMUST00000027671	Tnnt2-exp-r.1	GACGCTTTTCGATCCTGTCTT	Tnnt2-exp-f.1	r

Table S10. List of primers used for ChIP-chip verification

f - forward primer; r – reverse primer

MGI Gene Symbol	ENSMBL Closest Transcript ID	GeneBank Sequence Accession Number	Primer Name	5' to 3' Sequence	Genomic Sequence Region	Partner Primer	Orientation
Mpz	ENSMUST00000070758	AC163497	all_mr01_f1	CCACGGTTTTGAGGATTCCA	chr1:172987900-172988350	all_mr01_r1	f
Mpz	ENSMUST00000070758	AC163497	all_mr01_r1	TTCTCCCTTTGCCTTGCCA	chr1:172987900-172988350	all_mr01_f1	r
Tm9sf4	ENSMUST00000089027	AC078911	all_mr03_f1	TTAAAAACACCTCTGGCCCTG	chr2:152853750-152854300	all_mr03_r1	f
Tm9sf4	ENSMUST00000089027	AC078911	all_mr03_r1	CCTCCACTCTCATCCACAAAGA	chr2:152853750-152854300	all_mr03_f1	r
Rragc	ENSMUST00000030399 ENSMUST00000072215	AL606962	all_mr06_f1	CAGCGATCTGCTTACGGAATTA	chr4:123418800-123419200	all_mr06_r1	f
Rragc	ENSMUST00000030399 ENSMUST00000072215	AL606962	all_mr06_r1	CACGTGCGAAAGGCAATTAG	chr4:123418800-123419200	all_mr06_f1	r
Rps11 ; Rpl13a	ENSMUST00000051978 ENSMUST00000003518 ENSMUST00000083285	AC126256	all_mr09_f1	AGCTAAATCCCGTCTCAGGCAT	chr7:44995900-44996400	all_mr09_r1	f
Rps11 ; Rpl13a	ENSMUST00000051978 ENSMUST00000003518 ENSMUST00000083285	AC126256	all_mr09_r1	AGTTCCGGAGACCCTCCAGTAA	chr7:44995900-44996400	all_mr09_f1	r
Dctn5	ENSMUST00000033156	AC122232	all_mr10_f1	ACATATGTAAACTGCCCCCGTT	chr7:121924700-121925000	all_mr10_r1	f
Dctn5	ENSMUST00000033156	AC122232	all_mr10_r1	TGTGCATTCAGACCCCACTTC	chr7:121924700-121925000	all_mr10_f1	r
Polr3b	ENSMUST00000077175	AC140333	all_mr13_f1	TAATTGCTTCACGGTGAACTGC	chr10:84052890-84053060	all_mr13_r1	r
Polr3b	ENSMUST00000077175	AC140333	all_mr13_r1	TGCCAAAGATGTCAAGGTTCAG	chr10:84052890-84053060	all_mr13_f1	r
Aldh3a2	ENSMUST00000066277 ENSMUST00000074127 ENSMUST00000071413	AL672172	all_mr14_f1	CACAGCCCCTCTTTACCAGAA	chr11:61082000-61082400	all_mr14_r1	f
Aldh3a2	ENSMUST00000066277 ENSMUST00000074127 ENSMUST00000071413	AL672172	all_mr14_r1	TCCAGGCATGGTAAGACCTCTA	chr11:61082000-61082400	all_mr14_f1	r
Txndc5	ENSMUST00000035988	AC154747	all_mr16_f1	TTGGATTCCACAGGCACATT	chr13:38534600-38535000	all_mr16_r1	f
Txndc5	ENSMUST00000035988	AC154747	all_mr16_r1	TGGCTGTGTTTATTGCTGAGC	chr13:38534600-38535000	all_mr16_f1	r
Sfrs2ip	ENSMUST00000047835	AC158769	all_mr19_f1	CCGCTTAGGAATGCAATGAA	chr15:96289800-96290200	all_mr19_r1	f
Sfrs2ip	ENSMUST00000047835	AC158769	all_mr19_r1	GCGAAATACTTGCACACAGGA	chr15:96289800-96290200	all_mr19_f1	r
Nhlrc2	ENSMUST00000071423	AC116849	all_mr20_f1	TTTCGGACCCTTTTGCACTC	chr19:56602500-56603200	all_mr20_r1	f
Nhlrc2	ENSMUST00000071423	AC116849	all_mr20_r1	CCTCCATGCAGCCAATTCTT	chr19:56602500-56603200	all_mr20_f1	r

1810074P20Rik	ENSMUST00000038705	AL833775	all_mr21_f1	GTTCCTCCCAAACTTGATGTGA	chr4:41040964-41041134	all_mr21_r1	f
1810074P20Rik	ENSMUST00000038705	AL833775	all_mr21_r1	GCAGCGTGCTAATAGCTCTGTC	chr4:41040964-41041134	all_mr21_f1	r
Rpl21	ENSMUST00000035983	AC124828	all_mr22_f1	GGTGGTCTTCAAGTTACCCTGG	chr5:147143984-147144446	all_mr22_r1	f
Rpl21	ENSMUST00000035983	AC124828	all_mr22_r1	CCTCTTAGCAAAAGAGGCCAAA	chr5:147143984-147144446	all_mr22_f1	r
Cdh8	ENSMUST00000067860 ENSMUST00000067839 ENSMUST00000093249	AC162867	no_mr01_f1	AGGTTCCAGAGATAGGAACCCA	chr8:102300000-102303000	no_mr01_r1	f
Cdh8	ENSMUST00000067860 ENSMUST00000067839 ENSMUST00000093249	AC162867	no_mr01_r1	GGCCACCATCTGATTTAGCA	chr8:102300000-102303000	no_mr01_f1	r
Slc26a9	ENSMUST00000049027	AC161805	no_mr02_f1	CCGCTGAATGTGACCTATTGTC	chr1:133571000-133572000	no_mr02_r1	f
Slc26a9	ENSMUST00000049027	AC161805	no_mr02_r1	AAGGTCCCAAATGAACAGCC	chr1:133571000-133572000	no_mr02_f1	r
Rp2h	ENSMUST00000067979 ENSMUST00000033372	BX294384	no_mr04_f1	TCCCAGCAGCTCTTACCACATG	chrX:19531200-19532000	no_mr04_r1	f
Rp2h	ENSMUST00000067979 ENSMUST00000033372	BX294384	no_mr04_r1	TCAACCAACACTTGGATACCCA	chrX:19531200-19532000	no_mr04_f1	r
XP_916743.1	ENSMUST00000062483	AC122287	no_mr05_f1	GCACCCAGGCATTTTCTTTCA	chr1:120718500-120719500	no_mr05_r1	f
XP_916743.1	ENSMUST00000062483	AC122287	no_mr05_r1	TGTGTGTCAGTTCGGAGCTGAG	chr1:120718500-120719500	no_mr05_f1	r
Dock7	ENSMUST00000030282 ENSMUST00000097962 ENSMUST00000075836	AL935325	no_mr06_f1	TCTCCTGCCAACCTTGTGTGT	chr4:98522600-98523600	no_mr06_r1	f
Dock7	ENSMUST00000030282 ENSMUST00000097962 ENSMUST00000075836	AL935325	no_mr06_r1	AATTTGGAACTTCTCCCTCCTG	chr4:98522600-98523600	no_mr06_f1	r
Syt6	ENSMUST00000090697 ENSMUST00000098785	AC123057	no_mr07_f1	GCTGCTAAAGGCAGAAATGTGG	chr3:103704000-103706000	no_mr07_r1	f
Syt6	ENSMUST00000090697 ENSMUST00000098785	AC123057	no_mr07_r1	AATGGAAAAGGCGCTCTGG	chr3:103704000-103706000	no_mr07_f1	r
Zswim6	ENSMUST00000052377	CT572986	no_mr08_f1	GTTTCTGGCTCCGGTTGTATTG	chr13:108908700-108910100	no_mr08_r1	f
Zswim6	ENSMUST00000052377	CT572986	no_mr08_r1	TGTGTGCAGAAGCTGACCTCT	chr13:108908700-108910100	no_mr08_f1	r
Calb2	ENSMUST00000003754	AC163615	no_mr09_f1	CATCTGATGCAATCCGCCA	chr8:113051000-113051600	no_mr09_r1	f
Calb2	ENSMUST00000003754	AC163615	no_mr09_r1	AATCTTCCCCAATTCCCACA	chr8:113051000-113051600	no_mr09_f1	r
Foxj1	ENSMUST00000036215 ENSMUST00000078514	AL645861	no_mr10_f1	AATCTCCTCTTCCCACCCAAAC	chr11:116149000-116151000	no_mr10_r1	f
Foxj1	ENSMUST00000036215 ENSMUST00000078514	AL645861	no_mr10_r1	CTCCTTATTCAATGCCTTTGCC	chr11:116149000-116151000	no_mr10_f1	r

Table S11. Sources considered for array design

Human or mouse transcripts expressed in heart, skeletal or smooth muscle were selected from several sources as listed in the first column. The number of transcripts from each source is listed in the second column. The unified list of transcripts from these sources was represented on ChIP and expression arrays.

Source	Number of Transcripts
Key genes of cardiac development	55
Human chromosome 21 transcipts in Ensembl v26	211
Manually selected controls	204
Transcripts expressed in human heart Kaynak *et al.* [1]	2,546
Symatlas human atrioventricularnode – A/B [2]	2,399 / 2,399
Symatlas human cardiac myocytes – A/B [2]	4,786 / 3,981
Symatlas human heart – A/B[2]	3,391 / 3,978
Symatlas human skeletal muscle – A/B [2]	1,889 / 1,761
Symatlas human smoothmuscle – A/B [2]	5,296 / 5,237
Symatlas mouse heart	1,665
Symatlas mouse skeletal muscle	1,793
Transcripts expressed in mouse hearts Tabibiazar *et al.* [3]	132
All transcription factors listed in Transfac [4] as of Jan 2005	2,236

1. Kaynak B, von Heydebreck A, Mebus S, Seelow D, Hennig S, et al. (2003) Genome-wide array analysis of normal and malformed human hearts. Circulation. pp. 2467-2474.
2. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101: 6062-6067.
3. Tabibiazar R, Wagner RA, Liao A, Quertermous T (2003) Transcriptional profiling of the heart reveals chamber-specific gene expression patterns. Circ Res 93: 1193-1201.
4. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, et al. (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 34: D108-110.