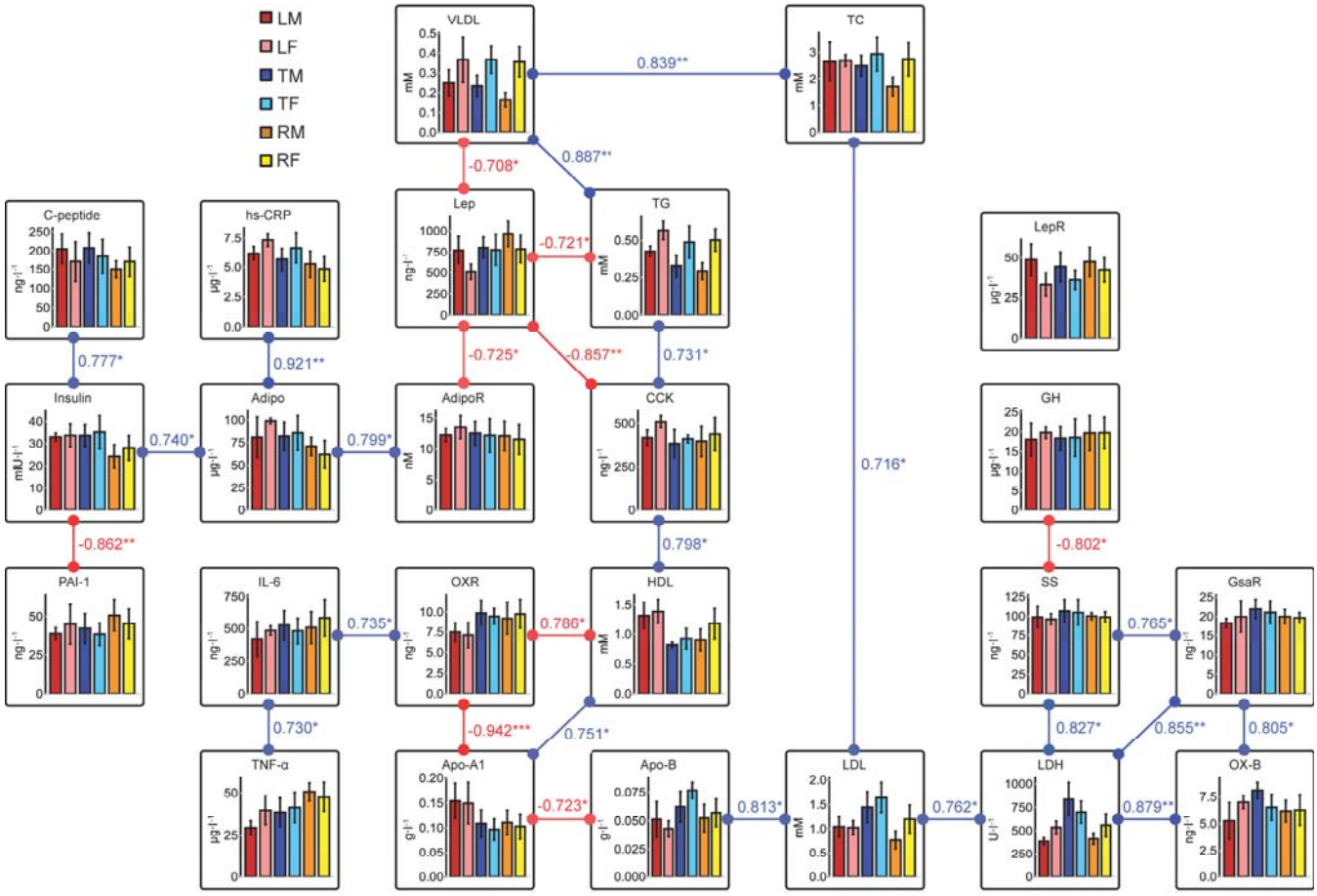
Supplementary Information

An atlas of DNA methylomes in porcine adipose and muscle tissues

Mingzhou Li, Honglong Wu, Zonggang Luo, Yudong Xia, Jiuqiang Guan, Tao Wang, Yiren Gu, Lei Chen, Kai Zhang, Jideng Ma, Yingkai Liu, Zhijun Zhong, Jing Nie, Shuling Zhou, Zhiping Mu, Xiaoyan Wang, Jingjing Qu, Long Jing, Huiyu Wang, Shujia Huang, Na Yi, Zhe Wang, Dongxing Xi, Juan Wang, Guangliang Yin, Li Wang, Ning Li, Zhi Jiang, Qiulei Lang, Huasheng Xiao, Anan Jiang, Li Zhu, Yanzhi Jiang, Guoqing Tang, Miaomiao Mai, Surong Shuai, Ning Li, Kui Li, Jinyong Wang, Xiuqing Zhang, Yingrui Li, Haosi Chen, Xiaolian Gao, Graham S. Plastow, Stephen Beck, Huanming Yang, Jian Wang, Jun Wang, Xuewei Li & Ruiqiang Li

Inventory of Supplementary Information:

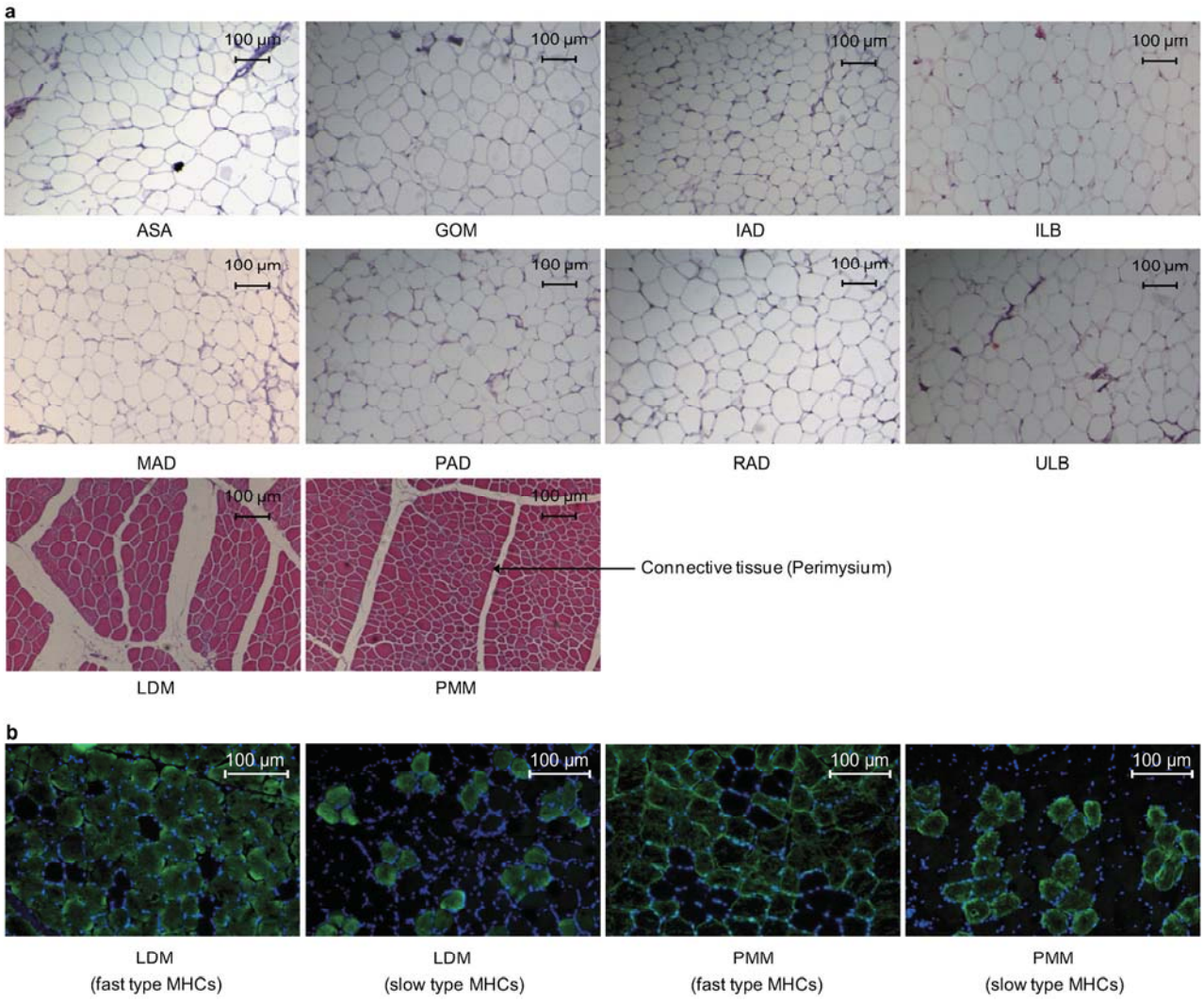
Supplementary Figure S1



Supplementary Figure S1 | Differences of 24 circulating indicators of metabolism in serum. 'L', 'T' and 'R' mean Landrace, Tibetan and Rongchang pig, respectively. 'M' and 'F' mean male and female, respectively. Data normally distributed (Kolmogorov-Smirnov test, $P > 0.05$). The statistical significance was calculated by two-way repeated-measures ANOVA ($n = 9$ per breed per sex, where 'B' and 'S' mean breed and sex, respectively) and is denoted by P_B, P_S, and P_{BxS}. Values are means ± s.d. They are adiponectin (Adipo) ($P_B = 0.0005$, $P_S = 0.648$, $P_{BxS} = 0.515$), adiponectin receptor (AdipoR) ($P_B = 0.490$, $P_S = 0.942$, $P_{BxS} = 0.613$), C-peptide ($P_B = 0.007$, $P_S = 0.215$, $P_{BxS} = 0.076$), cholecystokinin (CCK) ($P_B = 0.536$, $P_S = 0.157$, $P_{BxS} = 0.965$), gastrin receptor (GsaR) ($P_B = 0.019$, $P_S = 0.291$, $P_{BxS} = 0.901$), growth hormone (GH) ($P_B = 0.159$, $P_S = 0.827$, $P_{BxS} = 0.988$), highly sensitive C-reactive protein (hs-CRP) ($P_B = 0.004$, $P_S = 0.855$, $P_{BxS} = 0.077$), insulin ($P_B = 3.525 \times 10^{-5}$, $P_S = 0.418$, $P_{BxS} = 0.457$), interleukin - 6 (IL-6) ($P_B = 0.115$, $P_S = 0.485$, $P_{BxS} = 0.315$), leptin (Lep) ($P_B = 0.04$, $P_S = 0.012$, $P_{BxS} = 0.342$), leptin receptor (LepR) ($P_B = 0.249$, $P_S = 0.0003$, $P_{BxS} = 0.329$), orexin-B (OX-B) ($P_B = 0.001$, $P_S = 0.333$, $P_{BxS} = 0.016$), orexin receptor (OXR) ($P_B = 0.003$, $P_S = 0.841$, $P_{BxS} = 0.684$), plasminogen activator inhibitor-1 (PAI-1) ($P_B = 0.004$, $P_S = 0.224$, $P_{BxS} = 0.693$), tumor necrosis factor-α (TNF-α) ($P_B = 7.757 \times 10^{-5}$, $P_S = 0.842$, $P_{BxS} = 0.515$), somatostatin (SS) ($P_B = 0.024$, $P_S = 0.519$, $P_{BxS} = 0.979$), total cholesterol (TC) ($P_B = 0.0159$, $P_S =$

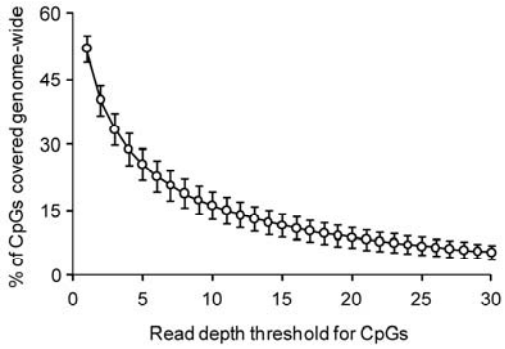
0.0002, P_{BxS} = 0.033), triglycerides (TG) (P_B = 0.017, P_S = 1.936 × 10^{-5}, P_{BxS} = 0.001), high density lipoprotein (HDL) (P_B = 0.0001, P_S = 0.017, P_{BxS} = 0.115), low density lipoprotein (LDL) (P_B = 1.634 × 10^{-8}, P_S = 0.0002, P_{BxS} = 0.261), very-low density lipoprotein (VLDL) (P_B = 0.124, P_S =3.524 × 10^{-9}, P_{BxS} = 0.001), lactate dehydrogenase (LDH) (P_B = 2.068 × 10^{-12}, P_S = 0.821, P_{BxS} = 0.002), apolipoprotein A-1 (Apo-A1) (P_B = 8.633 × 10^{-5}, P_S = 0.081, P_{BxS} = 0.879) and apolipoprotein B (Apo-B) (P_B = 0.0006, P_S = 0.015, P_{BxS} = 0.422). The Pearson's correlation was calculated between each pair of indicators. The bar graph was created using the visualization system VANTED[60], and the blue and red connecting lines showed positive and negative correlations (*P < 0.05, **P < 0.01, ***P < 0.001), respectively.

Supplementary Figure S2



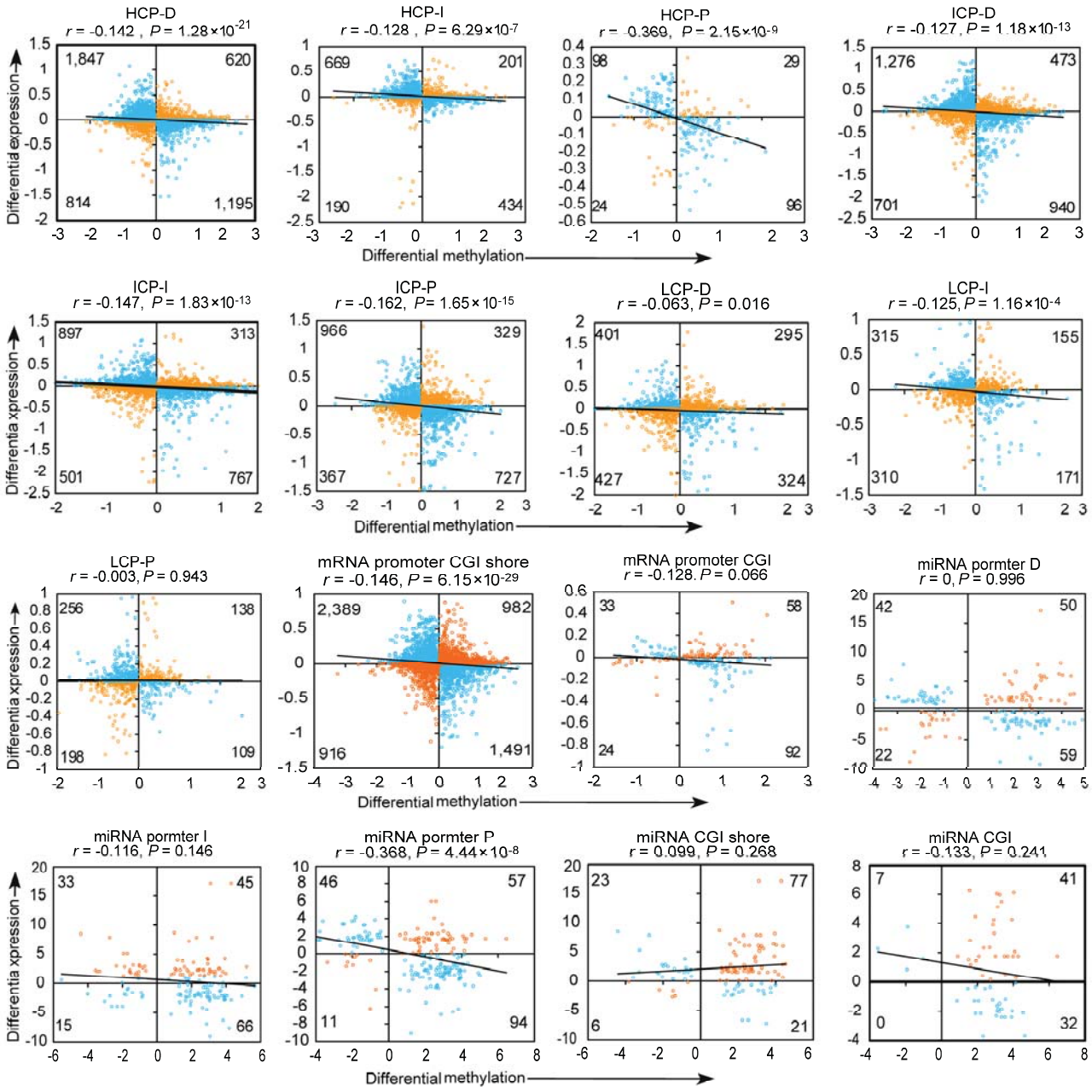
Supplementary Figure S2 | Morphological and immunohistochemistry measurements of adipocytes and myosin heavy chains (MHCs). (**a**) Enlarged photos of the eight adipocytes (ASA: abdominal subcutaneous adipose, GOM: greater omentum, IAD: intermuscular adipose, ILB: inner layer of backfat, MAD: mesenteric adipose, PAD: pericardial adipose, RAD: retroperitoneal adipose, ULB: upper layer of backfat), and two cross-sections of myofibers (LDM: *longissimus dorsi* muscle, and PMM: *psoas* major muscle). Paraffin embedded H&E stain. (**b**) Immunofluorescence against slow-twitch Type I MHCs and fast-twitch Type IIa / IIb MHCs (green) for LDM and PMM cross sections. The nucleuses were stained by DAPI (blue).

Supplementary Figure S3



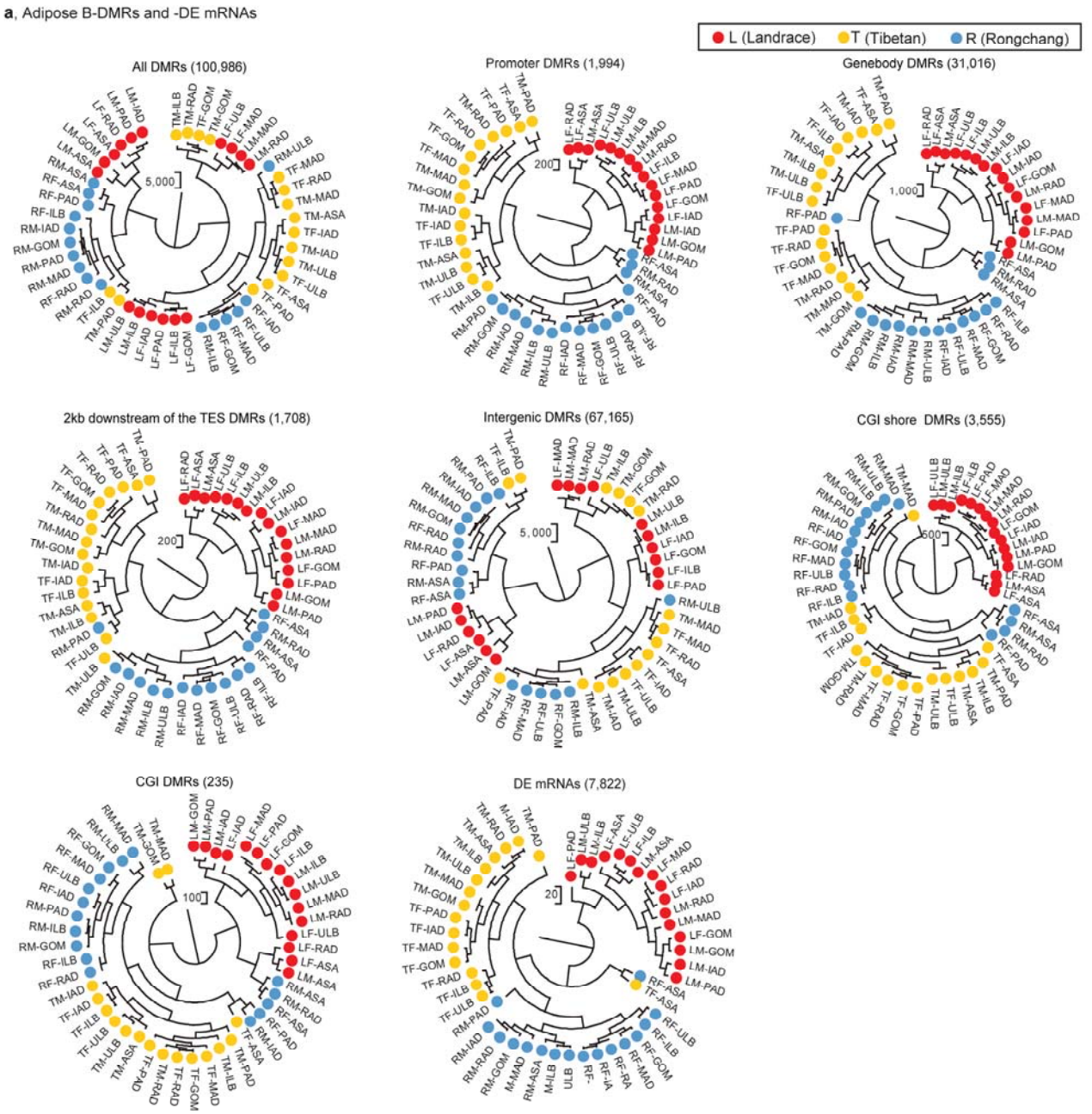
Supplementary Figure. S3 | Percent of CpGs covered by read depth thresholds on average over all samples. Values are means ± s.d (*n* = 180).

Supplementary Figure S4

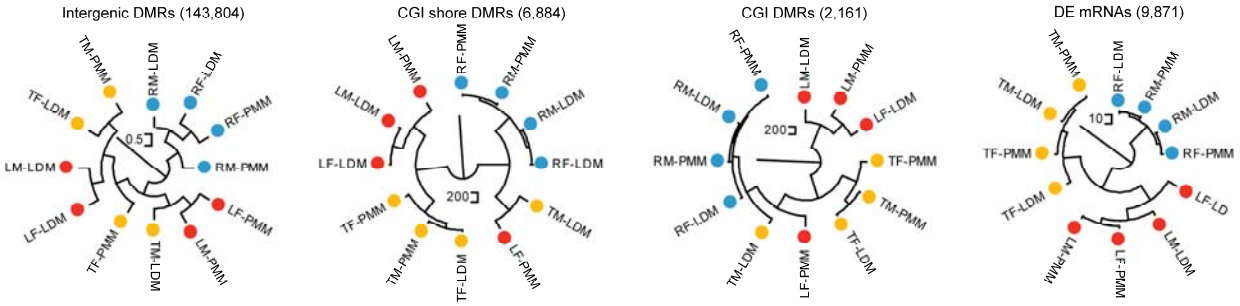


Supplementary Fig. S4 | Correlation between differential methylation and differential mRNA/miRNA expression. A scatter plot and trend line (Pearson correlation) showed correlation between the $\log_2$ ratios of mRNA /miRNA expression difference and the $\log_2$ ratios of the methylation differences. Line represents linear regression. Blue and orange dots represent the mRNA/miRNA– DMR pairs that exhibit the inverse and the same relationships, respectively. The number of mRNA /miRNA–DMR pairs that display a specific relationship are provided in each quadrant (upper left: hypermethylated being downregulated, bottom left: hypomethylated being downregulated, upper right: hypermethylated being upregulated, bottom right: hypermethylated being downregulated).

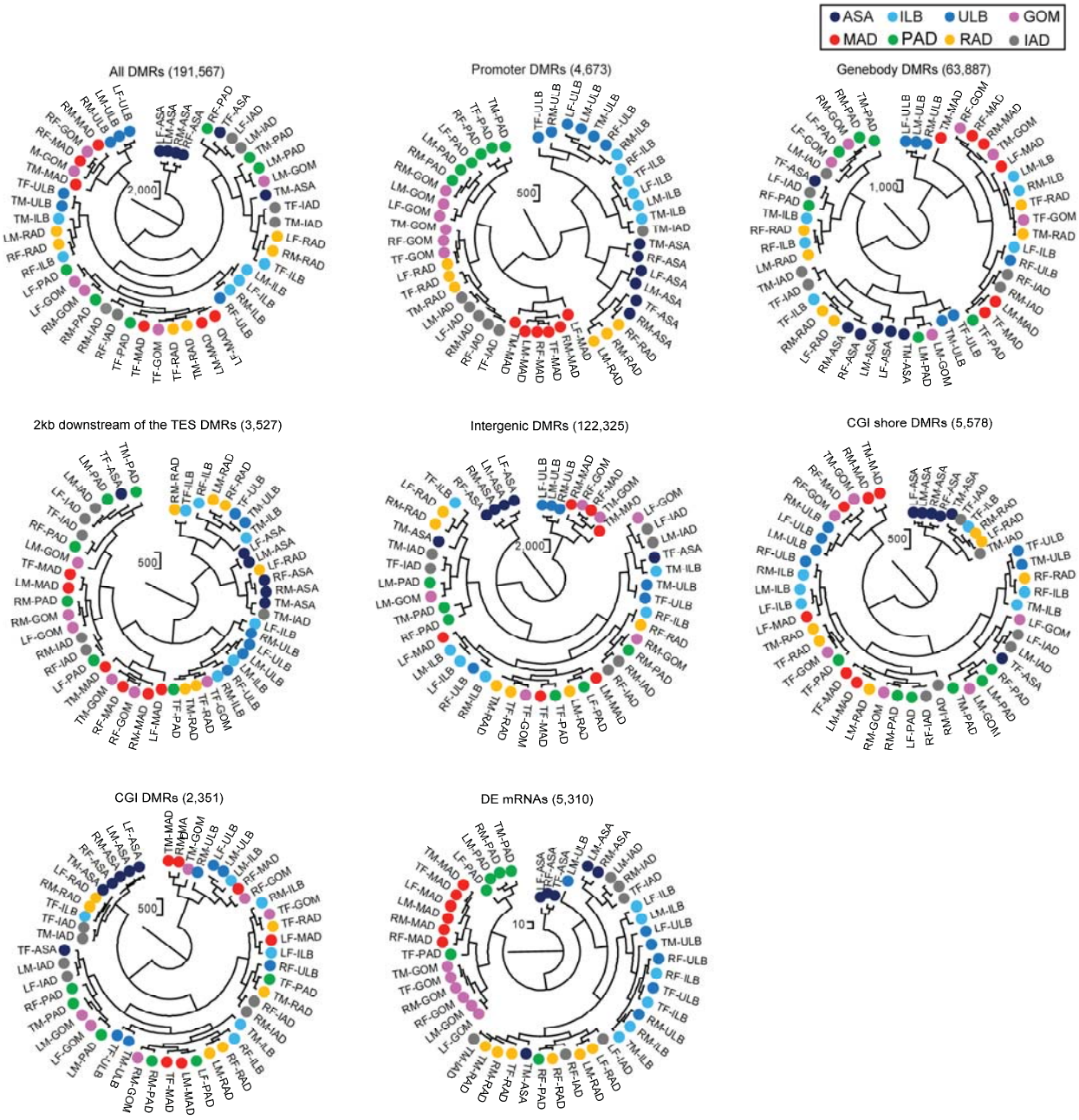
Supplementary Figure S5



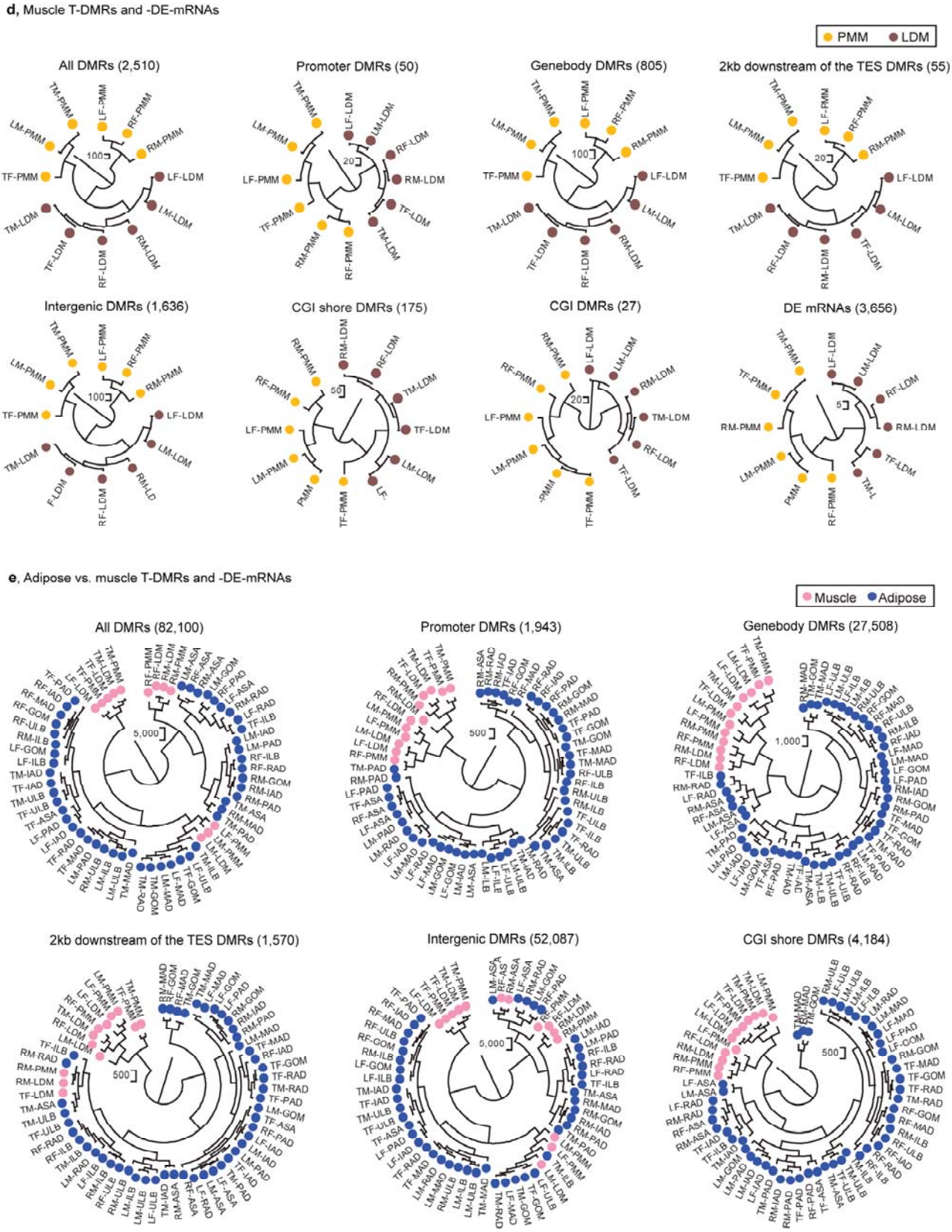
Supplementary Figure S5 cont. 1



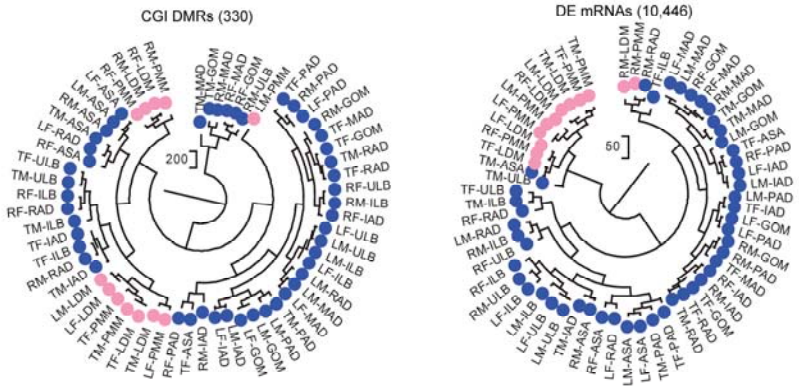
c, Adipose T-DMRs and -DE-mRNAs



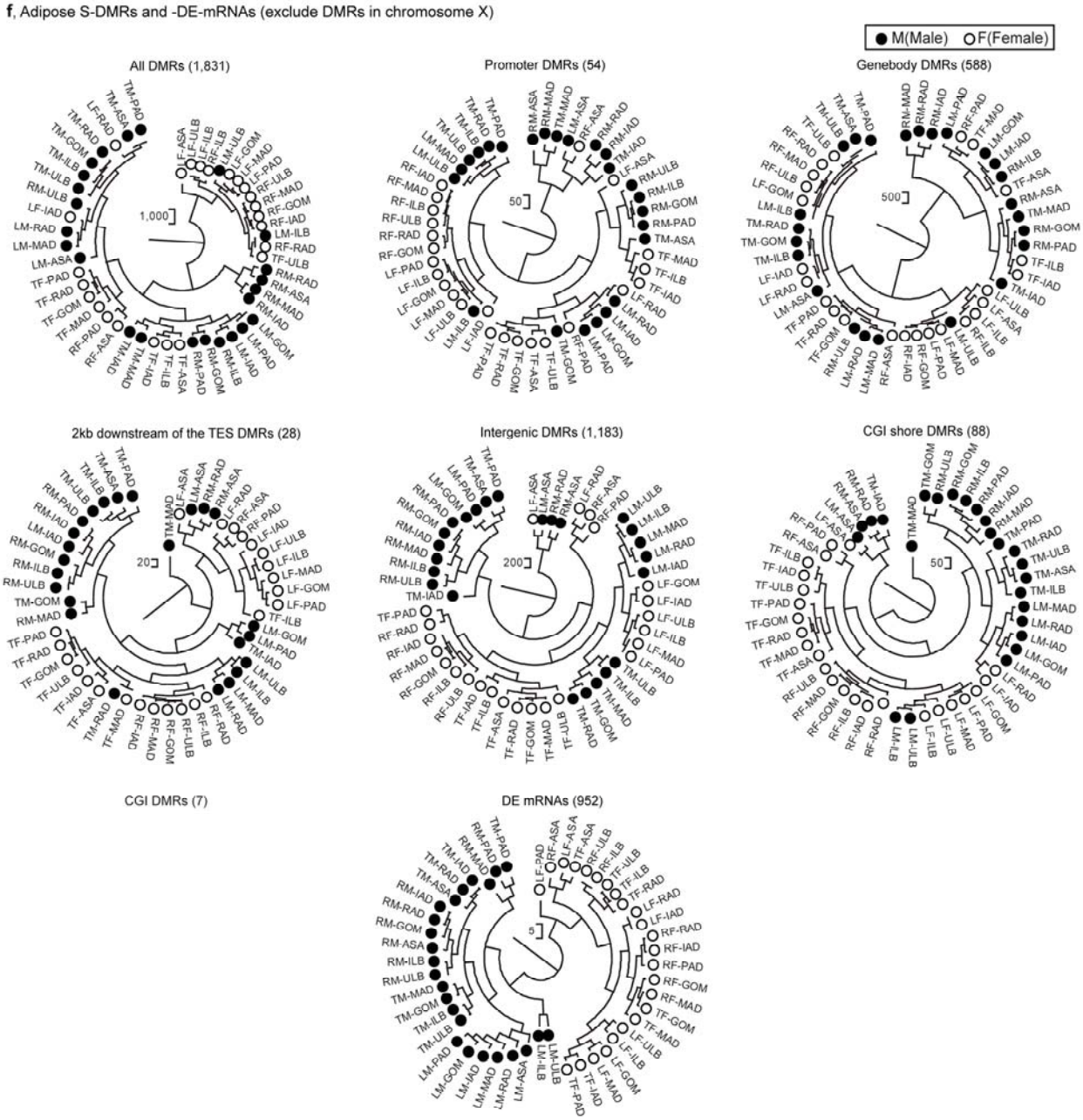
Supplementary Figure S5 cont. 2



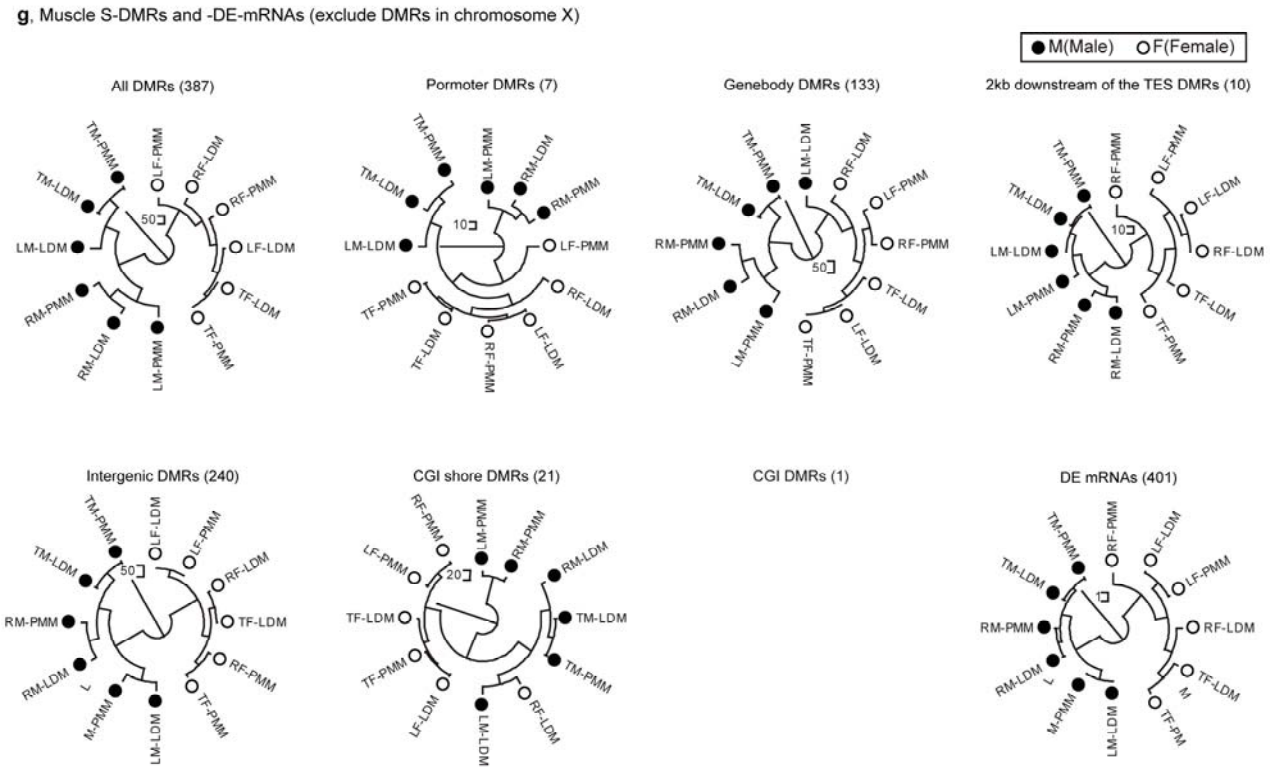
Supplementary Figure S5 cont. 3



f, Adipose S-DMRs and -DE-mRNAs (exclude DMRs in chromosome X)

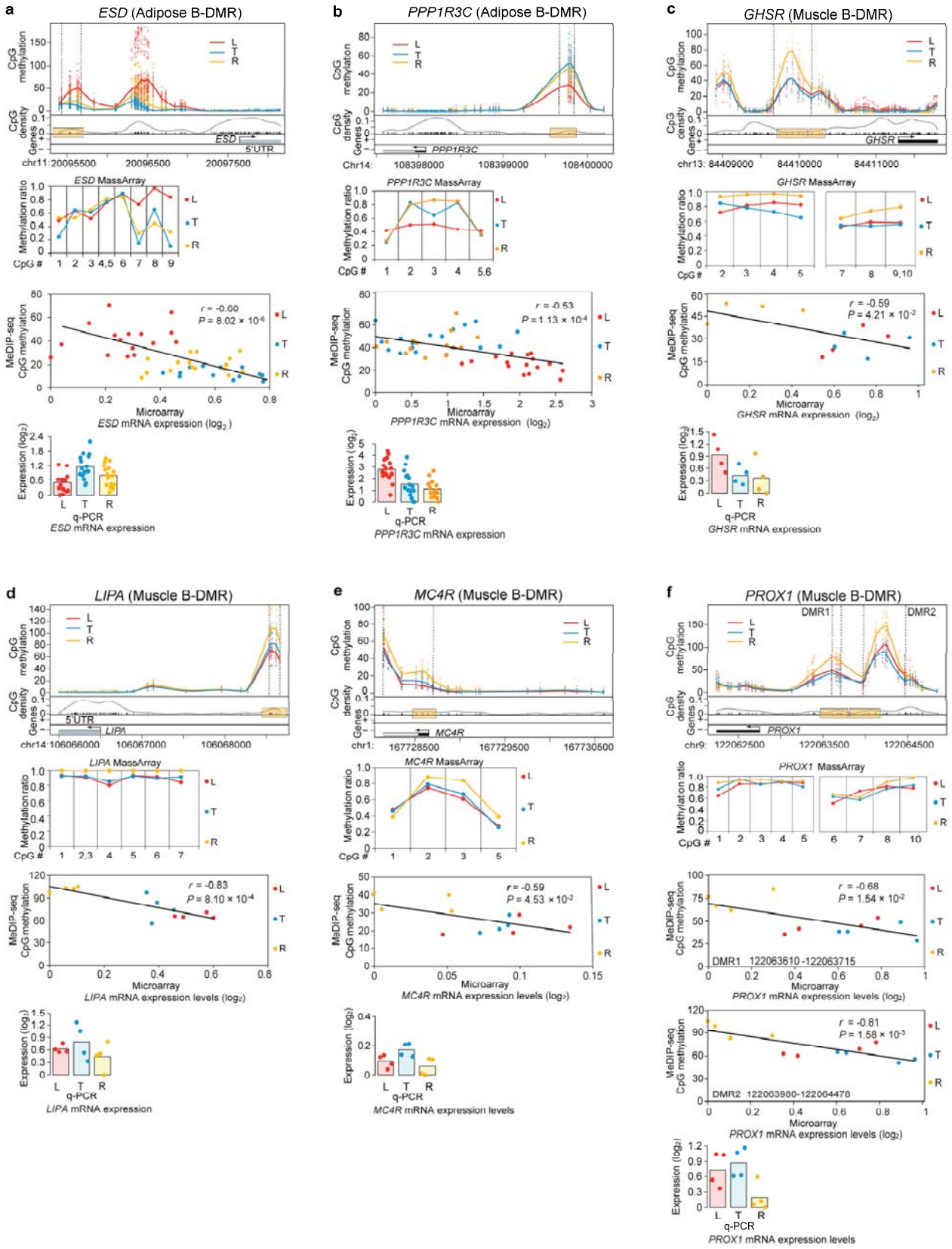


Supplementary Figure S5 cont. 4

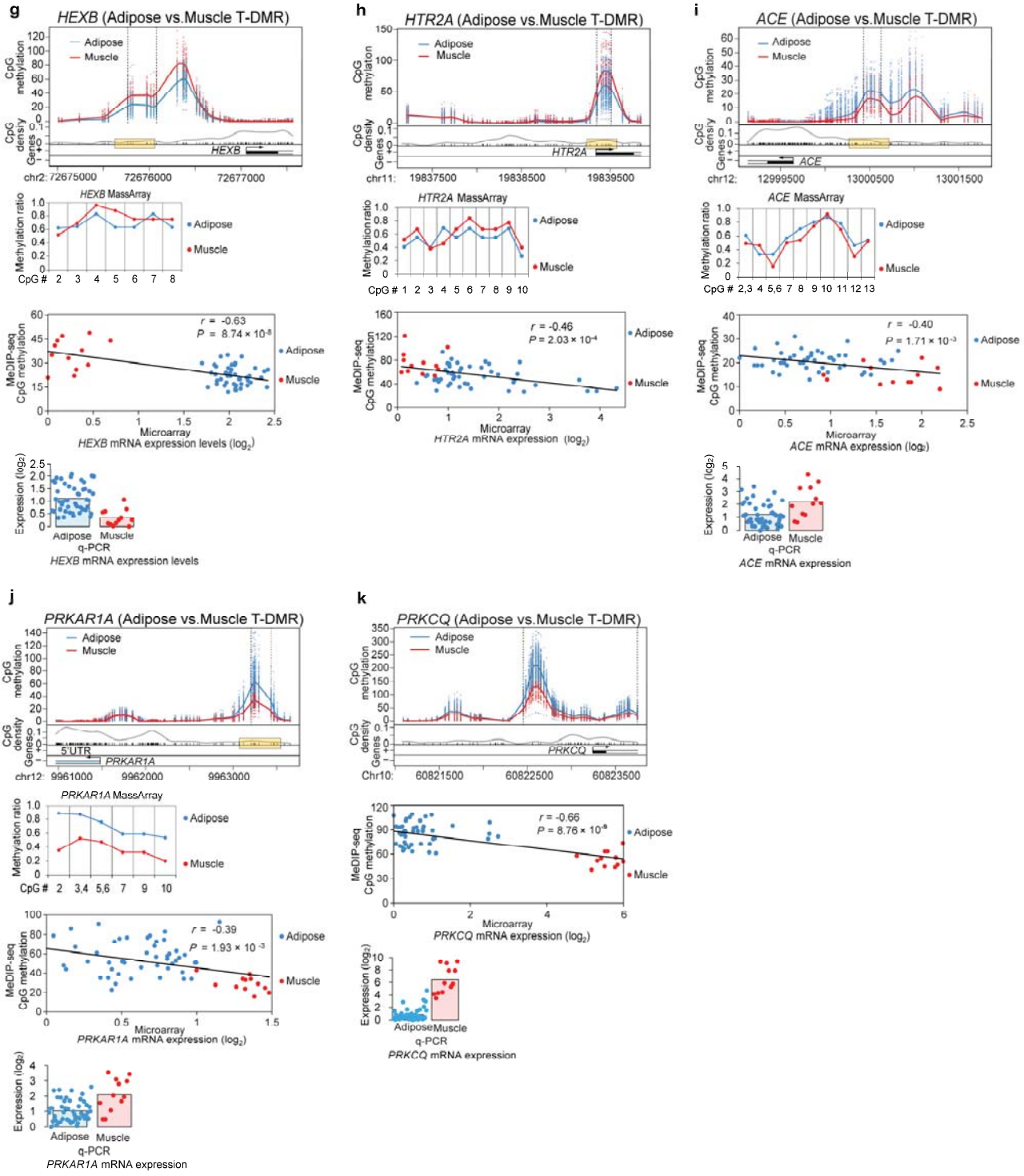


Supplementary Figure S5| Hierarchical clustering of samples using DMRs in different genomic elements and differentially expressed (DE) mRNAs. (**a**) adipose B-DMRs and DE mRNAs, (**b**) muscle B-DMRs and DE mRNAs, (**c**) adipose T-DMRs and DE mRNAs, (**d**) muscle T-DMRs and DE mRNAs, (**e**) adipose vs. muscle T-DMRs, (**f**) adipose S-DMRs and DE mRNAs, (**g**) muscle S-DMRs and DE mRNAs. The number of DMRs or mRNAs used for clustering is provided in brackets. If the number of DMRs are small than 10, the dendrogram has not been generated. The DMRs in promoters include those in miRNA and mRNA.

Supplementary Figure S6



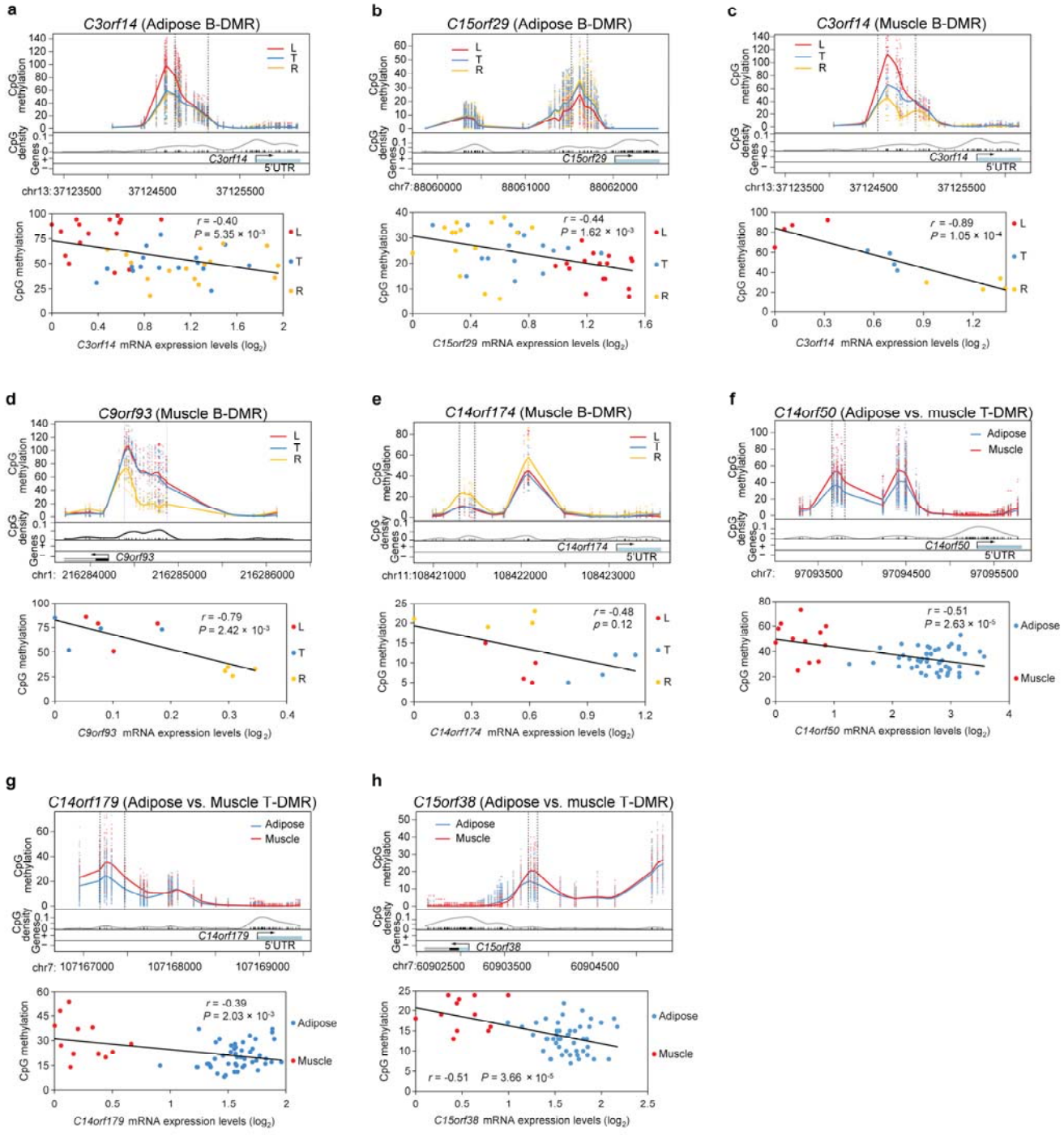
Supplementary Figure S6 cont. 1



Supplementary Figure S6 | Examples of known obesity-related genes showing differential DNA methylation in promoter. (**a** and **b**) The adipose B-DMR in (**a**) *ESD* and (**b**) *PPP1R3C* promoters. (**c-f**) The muscle B-DMR in (**c**) *GHSR*, (**d**) *LIPA*, (**e**) *MC4R* and (**f**) *PROX1* promoters. (**g-k**) The adipose vs. muscle T-DMR in (**g**) *HEXB*, (**h**) *HTR2A*, (**i**) *ACE*, (**j**) *PRKAR1A* and (**k**) *PRKCQ* promoters. Top panels: top half: CpG

methylation. Each point represents methylation level (MeDIP-seq read depth) of a sample at a given CpG site. The curves showed average over the samples. The two vertical dashed lines marked the boundaries of the DMR identified. Lower half: CpG dinucleotides (black tick marks on X axis), CpG density (gray line), TSS (black arrow), exons and introns (filled black and white boxes, respectively). Plus and minus marks denote sense and antisense gene transcription. Second panels: validation of individual CpG methylation by MassArray (mapping to yellow box in upper panel). Third panels: a scatter plot and trend line (Pearson correlation) showing correlation between the $\log_2$ ratios of mRNA expression from microarray and CpG methylation of the DMR from MeDIP-seq. Bottom panels: validation of mRNA expression levels by q-PCR. Bars represent the mean expression level.

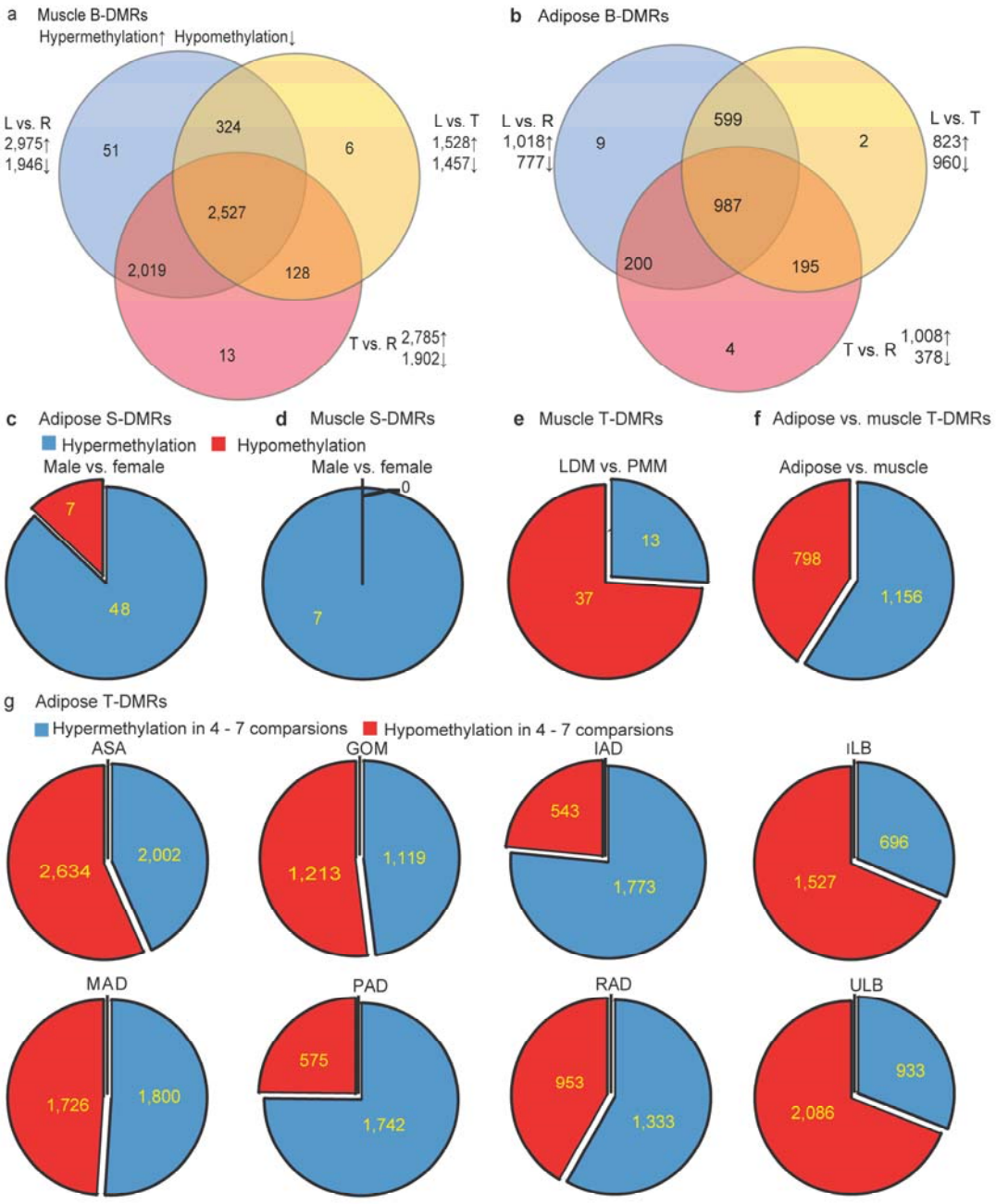
Supplementary Figure S7



Supplementary Figure S7 | Examples of putative genes located in QTLs region affecting fatness and pork quality showing differential DNA methylation in promoter. (**a** and **b**) The adipose B-DMR in (**a**) *C3orf14* and (**b**) *C15orf29* promoters. (**c**-**e**) The muscle B-DMR in (**c**) *C3orf14*, (**d**) *C9orf93* and (**e**) *C14orf174* promoters. (**f**-**h**) The adipose vs. muscle T-DMR in (**f**) *C14orf50*, (**g**) *C14orf179* and (**h**) *C15orf38* promoters. Top panels: top

half: CpG methylation. Each point represents methylation level (MeDIP-seq read depth) of a sample at a given CpG site. The curves showed average over the samples. The two vertical dashed lines marked the boundaries of the DMR identified. Lower half: CpG dinucleotides (black tick marks on X axis), CpG density (gray line), TSS (black arrow), exons and introns (filled black and white boxes, respectively). Plus and minus marks denote sense and antisense gene transcription. Second panels: a scatter plot and trend line (Pearson correlation) showing correlation between the $\log_2$ ratios of mRNA expression from microarray and CpG methylation of the DMR from MeDIP-seq.

Supplementary Figure S8



Supplementary Figure S8 | Statistics of genes with DMRs in promoters. (**a**) Muscle B-DMRs, (**b**) adipose B-DMRs, (**c**) adipose S-DMRs, (**d**) muscle S-DMRs, (**e**) muscle T-DMRs, (**f**) adipose vs. muscle T-DMRs, and (**g**) adipose T-DMRs. To identify specific adipose T-DMRs in a given adipose tissue, we looked for DMRs which hyper- or hypomethylated over half of the target tissues (i.e. a DMR for a given adipose tissue should be hyper- or hypomethylated against more than four of seven target tissues, simultaneously).

Supplementary Table S1 | Summary of MeDIP-seq data production.

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
		Total			**28,185,335,398**	**1,381**	**26,824,959,182**	**1,314**	**81.19**	**74.96**	**95.88**
Landrace	Male	Adipose	ULB	1	170,671,802	8.36	161,873,526	7.93	80.77	74.64	92.22
				2	178,619,948	8.75	168,877,592	8.28	80.84	74.40	91.53
				3	152,207,228	7.46	144,007,480	7.06	80.33	75.65	89.86
			ILB	1	138,296,328	6.78	131,316,922	6.43	79.41	74.39	94.85
				2	151,439,856	7.42	145,699,912	7.14	80.54	74.97	93.82
				3	144,787,884	7.09	140,672,846	6.89	82.57	77.36	90.64
			ASA	1	162,628,568	7.97	155,534,002	7.62	82.89	75.98	97.84
				2	164,251,202	8.05	156,982,468	7.69	83.71	76.33	97.28
				3	142,519,804	6.98	137,021,164	6.71	83.44	78.06	97.13
			GOM	1	150,820,726	7.39	142,813,998	7.00	82.95	76.46	93.67
				2	141,501,820	6.93	137,483,446	6.74	83.27	75.49	91.12
				3	145,506,452	7.13	141,169,458	6.92	82.97	77.58	95.45
			MAD	1	145,851,642	7.15	139,113,920	6.82	81.75	75.46	97.52
				2	238,728,294	11.70	221,793,254	10.87	81.85	75.24	97.16
				3	242,125,282	11.86	225,355,438	11.04	81.83	76.11	97.12
			RAD	1	184,549,414	9.04	177,931,908	8.72	82.03	76.15	96.57
				2	191,816,820	9.40	180,846,106	8.86	81.81	76.10	96.59
				3	181,180,196	8.88	173,067,490	8.48	83.00	77.46	97.89
			PAD	1	188,412,498	9.23	179,213,804	8.78	82.53	75.90	97.19
				2	163,450,970	8.01	156,240,576	7.66	81.67	75.19	96.75
				3	167,865,370	8.23	159,876,762	7.83	81.85	76.33	97.44
			IAD	1	171,113,584	8.38	162,331,310	7.95	82.45	76.72	96.54
				2	164,091,082	8.04	155,737,254	7.63	82.73	77.19	96.07
				3	161,980,238	7.94	154,118,228	7.55	82.34	77.15	97.06
		Muscle	LDM	1	114,983,790	5.63	109,535,498	5.37	80.20	74.96	95.10
				2	157,048,354	7.70	150,063,152	7.35	82.70	76.56	97.35
				3	172,269,232	8.44	163,911,394	8.03	82.35	76.79	97.32
			PMM	1	153,786,950	7.54	147,975,800	7.25	82.59	76.91	96.77
				2	143,767,242	7.04	137,921,732	6.76	82.60	77.23	96.65
				3	134,024,314	6.57	128,612,492	6.30	82.15	77.93	96.55

Supplementary Table S1 cont. 1

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
Landrace	Female	Adipose	ULB	1	157,599,066	7.72	148,762,856	7.29	80.39	74.55	90.22
				2	136,124,998	6.67	129,491,364	6.35	81.01	74.96	91.63
				3	163,622,790	8.02	154,282,780	7.56	80.89	74.53	91.24
			ILB	1	143,011,442	7.01	137,770,430	6.75	80.87	74.50	92.75
				2	131,080,312	6.42	124,543,412	6.10	81.92	75.09	93.74
				3	142,136,092	6.96	132,197,744	6.48	82.29	76.26	92.37
			ASA	1	164,991,988	8.08	157,577,118	7.72	84.00	75.63	97.17
				2	172,631,424	8.46	164,936,716	8.08	84.06	77.51	97.16
				3	162,035,790	7.94	154,273,836	7.56	82.78	74.36	97.38
			GOM	1	152,667,468	7.48	143,607,332	7.04	83.23	75.93	95.48
				2	141,152,880	6.92	134,814,874	6.61	82.98	76.67	95.84
				3	153,694,476	7.53	149,035,830	7.30	83.40	76.50	95.48
			MAD	1	157,592,242	7.72	151,115,210	7.40	81.71	74.32	97.02
				2	175,879,654	8.62	167,407,218	8.20	81.56	75.58	97.03
				3	179,937,280	8.82	173,467,140	8.50	81.40	73.23	96.92
			RAD	1	172,116,398	8.43	164,135,358	8.04	82.73	76.74	97.79
				2	179,267,338	8.78	171,052,790	8.38	83.83	77.52	97.47
				3	177,507,258	8.70	169,186,482	8.29	84.08	76.48	96.37
			PAD	1	184,321,936	9.03	175,354,616	8.59	82.10	75.20	96.95
				2	138,591,374	6.79	134,204,642	6.58	82.08	75.96	98.07
				3	154,426,594	7.57	148,331,456	7.27	82.11	75.58	97.94
			IAD	1	158,852,030	7.78	151,700,886	7.43	82.82	76.36	97.12
				2	152,787,008	7.49	145,242,784	7.12	82.39	76.97	97.28
				3	169,529,462	8.31	160,717,786	7.88	82.35	76.93	97.39
		Muscle	LDM	1	164,211,168	8.05	155,994,150	7.64	82.50	76.37	96.85
				2	154,630,448	7.58	147,049,484	7.21	82.72	77.51	97.74
				3	156,205,226	7.65	146,538,884	7.18	82.16	75.65	96.73
			PMM	1	132,771,734	6.51	127,393,854	6.24	82.60	77.32	96.99
				2	132,486,034	6.49	126,924,070	6.22	82.95	74.14	95.32
				3	156,904,678	7.69	149,390,752	7.32	82.39	77.40	97.16

Supplementary Table S1 cont. 2

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
Tibetan	Male	Adipose	ULB	1	139,627,610	6.84	134,042,084	6.57	82.79	76.81	94.63
				2	163,647,236	8.02	152,385,048	7.47	81.07	76.03	93.78
				3	141,003,626	6.91	134,350,214	6.58	81.06	75.34	94.78
			ILB	1	121,199,910	5.94	116,234,238	5.70	78.60	73.05	93.23
				2	140,715,784	6.90	137,030,556	6.71	82.14	77.35	93.70
				3	152,041,564	7.45	147,524,084	7.23	81.94	76.83	94.33
			ASA	1	171,965,036	8.43	163,845,354	8.03	82.18	75.65	96.68
				2	173,604,582	8.51	163,199,672	8.00	79.77	65.18	94.93
				3	175,633,892	8.61	166,626,484	8.16	80.97	76.08	97.33
			GOM	1	149,452,128	7.32	141,425,334	6.93	77.31	73.68	96.77
				2	190,237,858	9.32	178,601,332	8.75	75.72	72.85	96.60
				3	192,131,238	9.41	180,780,160	8.86	77.57	73.32	97.03
			MAD	1	159,257,508	7.80	153,993,472	7.55	78.47	73.15	95.67
				2	177,962,670	8.72	171,188,318	8.39	73.45	71.10	95.10
				3	188,953,058	9.26	178,142,918	8.73	78.90	74.34	96.21
			RAD	1	177,692,668	8.71	167,567,026	8.21	80.15	73.59	97.81
				2	169,924,006	8.33	161,471,976	7.91	80.74	75.25	97.57
				3	161,722,700	7.92	154,203,474	7.56	78.85	73.70	97.06
			PAD	1	172,016,186	8.43	164,492,072	8.06	81.23	75.65	96.89
				2	129,862,262	6.36	123,649,974	6.06	79.04	65.04	96.50
				3	132,559,270	6.50	126,215,096	6.18	80.99	74.34	97.93
			IAD	1	168,988,868	8.28	159,688,030	7.82	80.35	74.39	97.60
				2	154,469,906	7.57	146,954,462	7.20	81.36	77.07	97.26
				3	168,896,930	8.28	160,867,162	7.88	81.35	75.71	96.81
		Muscle	LDM	1	163,462,556	8.01	157,066,622	7.70	81.77	76.52	95.63
				2	147,379,492	7.22	141,822,334	6.95	81.44	74.39	96.51
				3	161,789,608	7.93	155,600,446	7.62	81.45	75.98	97.08
			PMM	1	157,230,794	7.70	150,220,166	7.36	81.80	77.11	96.62
				2	168,660,446	8.26	158,269,728	7.76	80.78	74.74	97.06
				3	166,384,564	8.15	157,092,740	7.70	80.90	76.04	97.50

Supplementary Table S1 cont. 3

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
Tibetan	Female	Adipose	ULB	1	163,730,720	8.02	158,784,010	7.78	83.30	77.36	92.55
				2	159,319,302	7.81	149,441,768	7.32	81.39	75.63	93.16
				3	159,644,794	7.82	154,664,274	7.58	81.68	74.75	93.46
			ILB	1	156,873,150	7.69	149,362,628	7.32	82.30	77.65	93.02
				2	123,748,300	6.06	117,846,132	5.77	82.28	77.40	93.18
				3	123,625,614	6.06	118,979,702	5.83	82.70	76.76	94.02
			ASA	1	141,589,822	6.94	134,653,820	6.60	81.35	76.74	96.69
				2	171,809,334	8.42	162,671,622	7.97	81.12	76.03	96.67
				3	165,009,576	8.09	154,670,948	7.58	81.29	75.02	97.15
			GOM	1	143,600,730	7.04	135,941,658	6.66	77.57	74.98	96.88
				2	156,072,010	7.65	148,447,264	7.27	80.79	74.76	95.58
				3	143,753,058	7.04	136,858,966	6.71	80.58	73.46	96.88
			MAD	1	146,430,390	7.18	139,955,822	6.86	80.88	75.37	96.59
				2	183,665,138	9.00	173,185,330	8.49	81.24	74.95	96.77
				3	158,642,968	7.77	153,330,346	7.51	80.68	73.92	95.38
			RAD	1	154,513,166	7.57	147,579,272	7.23	80.20	73.52	96.76
				2	109,208,428	5.35	104,793,210	5.13	78.99	71.61	96.95
				3	153,956,896	7.54	146,895,342	7.20	81.19	74.51	97.86
			PAD	1	152,804,834	7.49	146,093,464	7.16	79.59	73.93	95.75
				2	136,208,060	6.67	130,735,532	6.41	80.68	73.35	96.66
				3	162,624,786	7.97	155,185,848	7.60	80.91	73.14	96.34
			IAD	1	168,006,524	8.23	159,956,374	7.84	81.32	76.27	97.66
				2	164,529,350	8.06	156,111,910	7.65	81.82	76.29	97.52
				3	160,974,146	7.89	153,956,596	7.54	81.66	75.14	96.28
		Muscle	LDM	1	162,599,780	7.97	156,210,852	7.65	81.55	74.10	96.65
				2	158,651,926	7.77	152,639,792	7.48	81.78	76.22	96.83
				3	167,419,058	8.20	160,750,332	7.88	81.76	75.54	96.01
			PMM	1	160,800,590	7.88	152,137,686	7.45	80.79	75.63	97.52
				2	165,297,224	8.10	156,547,516	7.67	81.03	75.49	97.36
				3	149,045,916	7.30	141,431,716	6.93	80.98	73.93	97.44

Supplementary Table S1 cont. 4

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
Rongchang	Male	Adipose	ULB	1	161,549,798	7.92	151,911,274	7.44	79.11	72.81	89.66
				2	152,129,462	7.45	143,410,870	7.03	79.08	71.94	90.58
				3	146,106,150	7.16	137,974,052	6.76	78.56	72.75	91.11
			ILB	1	93,864,242	4.60	88,225,832	4.32	74.90	72.66	95.57
				2	137,137,714	6.72	130,116,150	6.38	81.40	74.19	93.89
				3	130,776,638	6.41	123,956,820	6.07	80.89	74.68	94.05
			ASA	1	174,614,830	8.56	166,915,900	8.18	82.89	75.17	96.92
				2	154,872,318	7.59	148,257,530	7.26	83.11	73.62	96.73
				3	167,520,124	8.21	159,940,480	7.84	82.08	75.69	96.25
			GOM	1	136,574,440	6.69	130,712,288	6.40	82.06	75.39	94.26
				2	144,772,284	7.09	135,623,222	6.65	81.63	72.52	95.83
				3	192,220,440	9.42	180,353,348	8.84	77.97	73.18	97.07
			MAD	1	175,814,636	8.61	165,645,462	8.12	77.66	71.41	95.78
				2	174,307,368	8.54	161,491,874	7.91	75.84	69.87	95.52
				3	170,316,442	8.35	164,540,128	8.06	79.66	72.74	96.20
			RAD	1	181,059,094	8.87	172,536,218	8.45	82.37	75.08	97.00
				2	164,383,244	8.05	157,777,182	7.73	82.87	73.70	96.38
				3	181,901,318	8.91	173,537,668	8.50	81.66	75.74	97.16
			PAD	1	162,332,940	7.95	154,863,556	7.59	80.93	73.72	97.59
				2	160,292,240	7.85	152,982,784	7.50	81.13	73.14	97.49
				3	168,265,798	8.25	160,995,312	7.89	81.54	73.97	96.07
			IAD	1	143,232,048	7.02	135,966,436	6.66	79.68	71.64	96.92
				2	136,558,218	6.69	130,138,218	6.38	80.89	72.33	96.66
				3	157,696,740	7.73	149,328,736	7.32	80.83	74.71	97.01
		Muscle	LDM	1	140,495,152	6.88	133,333,724	6.53	81.75	75.88	97.58
				2	139,402,690	6.83	130,362,574	6.39	81.58	75.08	97.05
				3	137,611,004	6.74	131,231,104	6.43	81.20	74.97	94.00
			PMM	1	157,311,320	7.71	151,064,142	7.40	82.39	76.73	94.88
				2	116,535,596	5.71	110,024,052	5.39	82.39	75.17	95.85
				3	118,969,094	5.83	113,499,902	5.56	81.94	76.66	94.87

Supplementary Table S1 cont. 5

Breed	Gender	Tissue type	Tissue symbol	Biological replicate	Number of raw reads	Raw reads (Gb)	Number of clean reads	Clean reads (Gb)	% aligned	% unique	% non-duplicate alignment
Rongchang	Female	Adipose	ULB	1	187,056,228	9.17	177,212,378	8.68	81.26	74.32	94.46
				2	126,060,710	6.18	120,858,500	5.92	80.67	73.19	94.71
				3	146,821,134	7.19	142,710,964	6.99	81.23	74.13	94.00
			ILB	1	114,984,412	5.63	110,322,514	5.41	78.91	73.45	93.07
				2	110,543,488	5.42	106,464,046	5.22	81.14	72.96	93.87
				3	126,839,390	6.22	121,115,662	5.93	82.90	76.25	92.11
			ASA	1	165,090,384	8.09	157,641,250	7.72	82.36	76.06	96.58
				2	164,531,530	8.06	157,312,308	7.71	82.52	74.98	93.89
				3	158,515,502	7.77	151,681,636	7.43	82.61	75.69	95.33
			GOM	1	184,216,102	9.03	172,780,246	8.47	76.80	73.15	96.94
				2	178,100,544	8.73	168,174,020	8.24	78.11	72.33	96.44
				3	188,754,522	9.25	177,296,804	8.69	77.70	72.88	96.57
			MAD	1	183,214,180	8.98	176,517,918	8.65	80.05	73.47	96.57
				2	165,412,160	8.11	157,596,652	7.72	77.27	71.73	96.19
				3	178,759,452	8.76	168,795,332	8.27	78.78	72.32	96.76
			RAD	1	152,691,978	7.48	146,429,082	7.18	82.72	75.77	97.08
				2	161,363,714	7.91	153,700,608	7.53	81.10	73.15	96.92
				3	167,776,238	8.22	159,532,724	7.82	80.16	72.95	97.06
			PAD	1	163,140,210	7.99	156,405,642	7.66	81.93	75.34	96.33
				2	142,961,034	7.01	137,533,656	6.74	81.65	74.01	96.50
				3	152,815,736	7.49	146,850,004	7.20	81.96	74.89	96.43
			IAD	1	156,757,400	7.68	148,254,214	7.26	80.80	73.77	96.96
				2	148,622,582	7.28	140,831,934	6.90	80.69	74.04	97.12
				3	171,408,060	8.40	162,038,504	7.94	80.54	71.20	96.53
		Muscle	LDM	1	138,902,840	6.81	132,228,848	6.48	82.01	76.69	97.18
				2	128,945,146	6.32	123,090,508	6.03	81.65	75.51	96.94
				3	125,590,198	6.15	118,535,974	5.81	81.78	76.72	95.69
			PMM	1	128,746,292	6.31	123,150,474	6.03	80.24	75.83	96.21
				2	92,128,362	4.51	84,013,996	4.12	82.42	76.16	95.89
				3	94,269,464	4.62	81,194,852	3.98	80.62	75.94	95.76

The low-quality reads were filtered from the raw reads and clean reads were used in further analysis. “% aligned” is percent of clean reads aligned on the pig reference genome (version 9.2). “% unique” is percent of reads uniquely aligned over all the aligned reads. The reads with same mapping locations in each sample were taken as potentially duplicated clones created by PCR amplification during sequencing library construction, and hence were removed in the analysis. “% non-duplicate alignment” is percent of uniquely aligned non-duplicate reads over all uniquely aligned reads.

Supplementary Table S2 | Primer sequences used for MassArray.

DMR type	Name	Primer sequence (5' to 3')	Amplicon				
			Chromosome	Start	End	Size (bp)	CpG No.
Adipose B-DMRs	ESD-L	aggaagagagAGATTTGGTGTGTTTTTGGTAAGAG	11	20,095,331	20,095,706	376	9
	ESD-R	cagtaatacgactcactatagggagaaggctAACAAATACTAAATAATCTCAAAAACCA					
	FTO-L	aggaagagagTGTGTGAATTTGGGTAAAGTATTTAAG	6	21,955,684	21,955,938	433	9
	FTO-R	cagtaatacgactcactatagggagaaggctCAAACCAAAAATCAAATCAAAACTC					
	PPP1R3C-L	aggaagagagAGTTGTTTGTTAGGGGTTGGTTAG	14	108,399,563	108,399,874	312	6
	PPP1R3C-R	cagtaatacgactcactatagggagaaggctAAAACCACAACAACACCAAATCT					
Muscle B-DMRs	FTO-L*	aggaagagagTTGGGAGATTATTTATTTTATTTGGAA	6	21,955,789	21,956,269	481	9
	FTO-R*	cagtaatacgactcactatagggagaaggctCACAAACAAAACACTTAAAAAAATACC					
	GHSR-1L*	aggaagagagGATTTAGGGTTTTTGGTTTAAGGTG	13	84,409,711	84,410,134	424	6
	GHSR-1R*	cagtaatacgactcactatagggagaaggctCTTAAAATCCATTCCCAATAATCCT					
	GHSR-2L*	aggaagagagTAATAATGAATAGAGTTTATTATGGATTTG	13	84,410,111	84,410,344	234	4
	GHSR-2R*	cagtaatacgactcactatagggagaaggctACCTTAAACCAAAAACCCTAAATCT					
	LIPA-L*	aggaagagagTTGGAGTTGTAGTTGTTGGTTTATG	14	106,068,456	106,068,753	298	8
	LIPA-R*	cagtaatacgactcactatagggagaaggctCTACAAAACCTTCTATCCCCTATCA					
	MC4R-L*	aggaagagagTTTTTAGATAGATAAAGATTTGGAGAAAA	1	167,728,479	167,728,736	258	5
	MC4R-R*	cagtaatacgactcactatagggagaaggctAAAATCACAAACACCTCAAAAAAAA					
	PROX1-1L	aggaagagagTTTTTGGGGGTGGTTTAAATATTAG	9	122,063,471	122,063,814	344	5
	PROX1-1R	cagtaatacgactcactatagggagaaggctCCCCTAAAAACTTCCATAAAAAAAA					
	PROX1-2L	aggaagagagGGAAGTTTTTAGGGGTTGTTAGAAT	9	122,063,800	122,064,177	378	5
	PROX1-2R	cagtaatacgactcactatagggagaaggctATCCACAAATAAAAATTAACCAAAA					
Adipose T-DMRs	ATP1B1-L*	aggaagagagGGGATAAAGGATGTTAAGGATTGTT	4	85,294,609	85,294,973	365	19
	ATP1B1-R*	cagtaatacgactcactatagggagaaggctCAAACAAAATAATTACCAAACCCAC					
	COL8A2-L*	aggaagagagGTTGGGAGGATTAAGTGAGTTTTTA	6	62,576,166	62,576,652	487	10
	COL8A2-R	cagtaatacgactcactatagggagaaggctATAACTACCAACCTACACCACAACC					
Adipose vs.muscle T-DMRs	ACE-L	aggaagagagAGTTTTTAGTTTATTTGGGGATTGG	12	13,000,272	13,000,725	454	13
	ACE-R	cagtaatacgactcactatagggagaaggctTACTAAAAAACAAATCCTCCCACCT					
	HEXB-L	aggaagagagGGATAGTGGGAAAGAAGTTTTTGAT	2	72,675,699	72,676,132	434	8
	HEXB-R	cagtaatacgactcactatagggagaaggctTAACAACTAACCTACACCACAACCA					
	HTR2A-L*	aggaagagagTTTTGGAGATGAAGTAAGGAGAAGT	11	19,839,235	19,839,576	342	10
	HTR2A-R*	cagtaatacgactcactatagggagaaggctAAAACCCTACCTATAAAACAACCTCA					

For each reverse primer, an additional T7 promoter tag for in vivo transcription was added, as well as a 10-mer tag on the forward primer to adjust for melting temperature differences. L = 10mer - tagged (aggaagagag) + (gene-specific sequence), R = T7 - tagged primers (cagtaatacgactcactatagggagaaggct) + (gene-specific sequence). * : primers were designed for the antisense strand because PCR amplification was not obtained in the forward direction.

Supplementary Table S3| Primer sequences used for q-PCR.

Gene symbol	Primer sequence (5' to 3')	Amplicon size(bp)	Ensembl/Genebank ID
ACE	AAAACCCAGAGCCCAAAC CCTCGGCATACTCGTTCA	103	ENSSSCG00000017296
ATP1B1	ACTGGCAAGCGTGACGAAGATAA CGCCTTACACTCTATGCGGATTT	195	ENSSSCG00000006296
COL8A2	CCGTTAAGTTTGACCGGACTCT CCTTGACGTGTACGTGGTAAGC	117	ENSSSCG00000003635
ESD	AGTAAATGGAAGGCTTATGATGC TTGCGATGAAGTAGTAGCTGTGA	211	ENSSSCG00000009407
FTO	CAGCAGTGGCAGCTGAAATA TGACCAGGTCCCGAAATAAG	133	ENSSSCG00000002832
GHSR	TACGCCACAGTGCTCACCATCA GACCAGCTTTACCCGGCCCTTG	108	ENSSSCG00000011754
HEXB	ATGCCCGGCACAGTAGTTCAGGT CTCCGCCAAGGACACGCTTTCTC	208	ENSSSCG00000014073
LIPA	TATGTGGGTCATTCTCAAGGT CTAGAGGGCTAGTAGCGAACT	133	ENSSSCG00000010450
MC4R	TACCCTGACCATCTTGATTG ATAGCAACAGATGATCTCTTTG	226	ENSSSCG00000004904
PPP1R3C	GAAACAGGAAGCCAAATCACAGA CATTAAGGTCCAAGAGGTCAAAC	170	ENSSSCG00000010464
PRKAR1A	GCCATGTTTCCGGTTTCCTTTAT ACACTGGTTGCCCACTCATTGTT	128	ENSSSCG00000017259
PRKCQ	AGGAAGAACAACGGGAAGACAGA CTCGTGACACTTGACATGGTGGA	195	ENSSSCG00000011131
PROX1	CCGTTTCAGAGTCCGTTAGGT TGGTGGGATGACATCTTGGTC	122	ENSSSCG00000015584
*ACTB**	TCTGGCACCACACCTTCT TGATCTGGGTCATCTTCTCAC	114	ENSSSCG00000007585 / DQ178122
*TBP**	GATGGACGTTCGGTTTAGG AGCAGCACAGTACGAGCAA	124	DQ178129
*TOP2B**	AACTGGATGATGCTAATGATGCT TGGAAAAACTCCGTATCTGTCTC	137	ENSSSCG00000011213 / AF222921

*: *ACTB* (*β* actin), *TBP* (TATA box binding protein) and *TOP2B* (topoisomerase II *β*) are the endogenous control genes.

Supplementary Methods

Body density

Individual body weight was measured before sacrifice. We calculated the pig body volume on the basis of three anthropometric parameters as previously described[61]. The pig body is considered to be like a truncated cone where the base is represented by the abdomen (A), the top by the neck (N) and the length by the body size (BS). Pig body volume was defined (l) as:

$$\text{Volume (l)} = (\pi\,(\text{BS}/3)\,(\text{cm}) \times \{(A)^2\,(\text{cm}) + (N)^2\,(\text{cm}) + (A\,(\text{cm}) \times N\,(\text{cm}))\}) \times 10^{-3}$$

where BS is the body size and A and N are the radius of the abdomen (A) and the neck (N).

It was then possible to determine the density in kg per litter as:

$$\text{Density (kg·l}^{-1}\text{)} = \text{BW (kg)} / \text{V(l)}$$

Circulating indicators of metabolism in serum

At 210 days of age (when the pigs reached peak commercial value), venous blood (50 ml) was collected from each fasted pig immediately before sacrifice. The whole blood was immediately centrifuged at 1,800 g for 10 minutes at RT, and then the resultant sera were stored at -80 °C.

Serum concentrations of total cholesterol (TC), triglycerides (TG), high density lipoprotein (HDL), low density lipoprotein (LDL), very-low density lipoprotein (VLDL), lactate dehydrogenase (LDH), apolipoprotein A-1 (Apo-A1) and apolipoprotein B (Apo-B) were determined in triplicates for each pig by using CL-8000 clinical chemical analyzer (Shimadzu) via standard enzymatic procedures.

Serum levels of 16 hormones / cytokines, which are well known to be associated with ATs deposition and SMTs growth, were measured in triplicates for each pig using a pig-specific ELISA kit (RuiCong). These analyzed molecules included adiponectin (Adipo), adiponectin receptor (AdipoR), C-peptide, cholecystokinin (CCK), gastrin receptor (GsaR), growth hormone (GH), highly sensitive C-reactive protein (hs-CRP), insulin, interleukin - 6 (IL-6), leptin (Lep), leptin receptor (LepR), orexin-B (OX-B), orexin receptor (OXR), plasminogen activator inhibitor-1 (PAI-1), tumor necrosis factor-α (TNF-α) and somatostatin (SS).

Histology of adipocyte and myofiber

After sacrifice, all ATs and SMTs were fixed in 10% neutral buffered formalin solution, embedded in paraffin using TP1020 semi-enclosed tissue processor (Leica), sliced at a thickness of 6 μm using RM2135 rotary microtome (Leica) and stained with hematoxylin and eosin (H&E). The mean diameter of an adipocyte cell was calculated as the geometric average of the maximum and minimum diameter, and 100 cells were measured for each sample in randomly selected fields using a TE2000 fluorescence microscope (Nikon) and Image Pro-Plus 7.0 software

(Media-Cybernetics). The mean adipocyte volume (V) was obtained according to the following formula:

$$V = \pi/ 6 \sum f_i D_i^3 / \sum f_i$$

where D_i is the mean diameter; f_i denotes number of cells with that mean diameter D_i.

The myofiber cross-sectional area was measured as an average of 100 fibers in randomly selected fields.

Myofiber type ratio (fast / slow)

Immunofluorescence staining was performed as previously described[62]. Two MHC monoclonal antibodies, NLC-MHCs and NCL-MHCf (Novocastra), were raised against pig slow-twitch Type I MHCs and fast-twitch Type IIa / IIb MHCs, respectively. The myofiber type ratio was determined by the area rate of fast MHCs versus that of slow MHCs for 100 fields of vision (randomly chosen).

Fatty acid composition

We determined the fatty acid composition of each AT as previously described[63] with some modifications. The fatty acid methyl esters (FAME) were quantified using GC-14C gas chromatograph (Shimadzu), which is equipped with an on-column injector and a flame ionization detector, under the following conditions: 60m × 0.25 mm capillary column size (CP-sil 88 for FAME, Varian - Chrompack), 0.20 μm film thickness, H_2 as carrier gas, flame injector split 20:1 at 260 °C and ionization detector. Response factors were determined by analyzing a standard solution of the relevant pure FAME. Individual compounds were identified by comparing their retention times with those of standards (Sigma). FAMEs were identified by comparison with standards previously run independently or together with samples. Results were expressed as grams of each fatty acid per 100 g of FAMEs detected.

MeDIP–seq

To account for variation of DNA methylome among individuals of a breed, we randomly selected three pigs with a specific sex from each breed as biological replicates. We selected ten tissues from each individual, so in total 180 samples were sequenced separately.

MeDIP DNA libraries were prepared following the protocol as previously described[54]. Each MeDIP library was subjected to paired-end sequencing using Illumina HiSeq 2000 at a 50 bp read length. Five microgram of original DNA was isolated using E.Z.N.A. HP Tissue DNA Midi Kit (Omega) and was sonicated to ~ 100 – 500 bp fragments with a Bioruptor sonicator (Diagenode). Then libraries were constructed by adopting the Illumina Paired-End protocol consisting of end repair, <A> base addition and adaptor ligation steps, which were performed using Illumina's

Paired-End DNA Sample Prep kit following the manufacturer's instructions. Adaptor-ligated DNA was immunoprecipitated by monoclonal anti-methylcytidine antibody. The specificity of the enrichment was confirmed by qPCR using SYBR green mastermix (Applied Biosystems) and primers for positive and negative internal control DNA of non-human samples were supplied in the Magnetic Methylated DNA Immunoprecipitation kit (Diagenode). Cycling of qPCR validation consisted of 95 °C 5 min, followed by 40 cycles 95 °C 15 s and 60 °C 1 min. The enriched fragments with methylation and 10% input DNA were purified on ZYMO DNA Clean & Concentrator-5 columns (ZYMO) following the manufacturer's instructions. DNA was eluted in 30 µl buffer EB and its concentration was measured. Enriched fragments were amplified by adaptor-mediated PCR in a final reaction volume of 50 µl consisting of 23 µl purified DNA, 25 µl Phusion DNA polymerase mix (NEB) and 2 µl PCR primers. Amplification consisted of 94 °C 30 s, 10 cycles of 94 °C 30s, 60 °C 30 s, 72 °C 30 s, followed by prolonged extension for 5 min at 72 °C and the products were held at 4 °C. Amplification quality and quantity were evaluated by Agilent 2100 Analyzer DNA 1000 chips purified by 2% agarose gel and eluted in 15 µl buffer EB. Ultra-high-throughput 50 bp paired-end sequencing was carried out using the Illumina HiSeq 2000 according to manufacturer's instructions. Raw sequencing data was processed by the Illumina base-calling pipeline.

Mapping reads

We filtered low quality MeDIP-seq reads that contained more than 5 'N's or over 50% of the sequence with low quality value (Phred score < 5). The sequencing reads were aligned to the UCSC pig reference genome (Sscrofa9.2, ftp://hgdownload.cse.ucsc.edu/goldenPath/susScr2/), allowing up to four mismatches using SOAP2 (Version 2.21)[55]. Rarely, multiple reads from the same sequencing library mapping to the same genomic location were regarded as potential clonal duplicates due to PCR amplification biases, and therefore were used as one read. To avoid stochastic sampling drift, we filtered out CpG sites that were covered by less than a 10 read depth.

Identification of DMRs

Instead of the classical peak detection method that was controversially suited for DMR identification, we use a novel method which calculates the variation of single CpGs for scanning the genomic regions enriched for methylated CpGs across breeds (B-DMRs), sexes (S-DMRs) and tissues (T-DMRs).

First, the normality and equal variances of read depth at each CpG across different sample groups were tested using Bartlett's test (passed for $P > 0.05$, failed for $P < 0.05$). To compare between two sample groups (i.e. adipose S-DMRs, muscle S-DMRs, muscle T-DMRs and adipose vs. muscle T-DMRs), we used Student's *t*-test (if passed Bartlett's test) or Wilcoxon rank-sum test

(if failed Bartlett's test). To compare among three or more groups (i.e. adipose B-DMRs, muscle B-DMRs and adipose T-DMRs), we used one-way repeated-measures ANOVA followed by Holm-Sidak's post-hoc test (if passed Bartlett's test) or Kruskal-Wallis non-parametric ANOVA followed by Tukey's post-hoc test (if failed Bartlett's test).

If the read depth difference at a single CpG across the sample groups was statistically significant ($P < 0.01$), the CpG was regarded as the seed site of a candidate DMR. Then a 3' downstream adjacent CpG was incorporated with this seed CpG. To highlight the CpG enriched regions, up to 200 bp inter-distance was allowed between the two adjacent CpGs, which resulted in a coverage of ~88.82% of all CpG sites in the pig genome. The average read depth of these two CpGs was under a new round of tests, and would repeatedly go on for the next CpG until encountering a low-variance CpG ($P > 0.01$), which was allowed to be up to 2 kb from the seed CpG. In order to eliminate "trailing smear" (i.e. the difference of a high variance CpG covers up the downstream low variance CpGs), we repeatedly performed the above test procedures from the last CpG in 3' downstream to the first site in 5' upstream (i.e. seed CpG). If five or more CpGs in a genomic region have statistical significance ($P < 0.01$) different read depth across the samples, then this region was taken as a DMR. The resulting *P* values for DMRs were corrected using the Benjamini-Hochberg method (FDR < 0.01, 1,000 permutations).To highlight the specific B-DMRs, S-DMRs and T-DMRs, mutual DMRs (more than 60% overlap) between these three DMR types (excluding the adipose vs. muscle T-DMRs) were filtered out.

Functional enrichment analysis for DMRs in promoters

Functional enrichment analysis of Gene Ontology (GO) and pathway was performed using the DAVID (Database for Annotation, Visualization and Integrated Discovery) web server (http://david.abcc.ncifcrf.gov/)[59]. Genes with DMRs in promoters were mapped to their respective human orthologs, and the lists were submitted to DAVID for enrichment analysis of the significant overrepresentation of GO biological processes (GO-BP), molecular function (GO-MF) terminologies, and KEGG-pathway category. In all tests, the whole known genes were appointed as the background, and *P* values (i.e. EASE score), indicating significance of the overlap between various gene sets, were calculated using Benjamini-corrected modified Fisher's exact test. Only GO-BP, GO-MF or KEGG-pathway terms with a *P* value less than 0.05 were considered as significant and listed.

Gene expression microarray

The total RNA (10 μg) of the 180 samples that corresponded to the samples used for MeDIP-seq sequencing was extracted with TRIzol (Invitrogen) and further purified using an RNeasy column (Qiagen). The total RNA quantity and purity passed the analysis of NanoDrop ND-1000

spectrophotometer (Nano Drop) at 260/280 nm (ratio > 2.0). The integrity of total RNA also passed analysis with the Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent Technologies) with RIN number > 6.0. The labeling procedure was carried out using an RNA Fluorescent Linear Amplification Kit (Agilent Technologies). The sample was labeled with Cy-3 dye only. Fragmentation was carried out by incubation at 60 °C for 30 min in a fragmentation buffer (Agilent Technologies), and the process was stopped by the addition of an equal volume of hybridization buffer (Agilent Technologies). The fragmented target was applied to an Agilent Pig Gene Expression Oligo Microarray (Version 2). Hybridization was carried out at 60 °C for 17 h in a hybridization oven (Robbins Scientific). The hybridized single channel array was scanned with an Agilent microarray scanner. The TIFF image generated was loaded into Feature Extraction Software (Agilent Technologies) for feature data extraction, and data analysis was performed with MultiExperiment Viewer (MeV)[64].

In order to identify differentially expressed (DE) mRNAs ($P < 0.01$) for the clustering analysis, we used one way ANOVA for comparisons of more than two groups or Student's *t*-test for two groups comparisons, with each group corresponding to one of the seven DMR types. Resulting *P* values of above tests were corrected with adjusted Bonferroni method (FDR < 0.01, 1,000 permutations).

In order to obtain high-confidence gene expression data, we mapped 43,603 probes (60 mer in length) to the pig reference genome allowing up to one mismatch, this resulted in 27,955 probes (64.11%) that were uniquely mapped, of which 4,983 (11.43%) probes were uniquely mapped to exons of the Ensembl genes (more than 60% sequence overlap). Multiple probes mapped to the same or different exons of a specific gene were filtered. Only 3,074 probes (7.05%) uniquely representing 3,074 genes were used in subsequent analysis.

Small RNA-seq

Eight ATs and two SMTs of the three female Landrace pigs were used for small RNA-seq. Total RNA was extracted using *mir*Vana™ miRNA isolation kit (Ambion) following manufacturer's instruction and subjected to quality control as above for gene expression microarray.

For each library, equal quantities (5 μg) of small RNA isolated from every three female Landrace pigs were pooled. Approximately 15 μg of small RNA representing each tissue was used for library construction and sequencing. The small RNA fragments (between 10-40 nt) were isolated by polyacrylamide gel electrophoresis (PAGE) and ligated with proprietary adaptors (Illumina). The short RNAs were then converted to cDNA by RT-PCR and each small RNA library was sequenced individually using Illumina Genome Analyzer II according to manufacturer's instructions. The sequencing generated a total of 7.12 Gb of 36-nt single-end reads for ten libraries.

In silico analysis of the small RNA-seq data

Reads were processed using Illumina's Genome Analyzer Pipeline software and then carried out as described by our previous report[57] with some improvements.

The reads were then subjected to a series of additional filters with acceptance criteria derived from the statistics of mammalian miRNAs in miRBase 16.0: (1) not sequencing adapters; (2) containing no more than 80% A, C, G, or T; (3) containing no more than two N (undetermined bases); (4) containing not only A and C or G and T, (5) containing no stretches of A_7, C_8, G_6, or T_7; (6) being longer than 14 nt or shorter than 27 nt; (7) not containing $\geq$ 10 repeats of any dimer, $\geq$ 6 repeats of any trimer, or $\geq$ 5 repeats of any tetramer; (8) been observed more than two times; and (9) not originating from porcine known classes of RNAs (i.e. mRNA in NCBI database[65]; rRNA, tRNA, snRNA and snoRNA in Rfam database[66]; and repetitive sequence elements in Repbase database[67]). All reads were counted and the identical reads were combined into a single kind. The reads that satisfy the acceptance criteria were passed through and called "mappable reads".

The mappable reads were mapped to the UCSC pig reference genome (~2.26 Gb) (Sscrofa9.2, ftp://hgdownload.cse.ucsc.edu/goldenPath/susScr2/) using NCBI Local BLAST. The mapping process included several major steps: (1) map the mappable reads to the known porcine and then to other mammalian pre-miRNAs registered in miRBase 16.0[68]; (2) map the mapped reads in step 1 to pig genome to obtain their genomic locations and annotations; (3) cluster the unmapped sequences in step 2 (that have been mapped to miRBase sequences, but not to the pig genome) as potential novel miRNAs from gaps in the genome; and (4) predict hairpin RNA structures of the mappable reads not-mapped to miRBase in step 1 from the adjacent 60 nt sequences in either direction using UNAFold[69]. To avoid ambiguous reads that have been assigned to multiple positions in pig genome, only reads longer than 18 nt in length were included in step 4.

Mapping and cataloging miRNAs

In ten libraries, miRNAs frequently exhibited extensive sequence-heterogeneity, producing multiple mature variants (named isomiRs) from the miRBase depository as reported in literatures[70-72]. In these cases, as our previous report[57], the most abundant isomiR was chosen as a reference sequence. This provides the most robust approach for evaluation of differential expression. Measuring the abundance of a given miRNA using the count of the most abundant isomiR correlated well with the expression level of the total counts of all isomiRs (average $r = 0.95$, Pearson). In general, the following discussions refer to the most abundant isomiR and its counts, when describing a family of sequences that may vary by length and/or vary by one nucleotide.

The pre-miRNAs and mature miRNA transcripts identified from the mappable reads are divided into five groups with high-to-mid confidence[73] in order (Supplementary Data 9): (1) 301 miRNAs corresponding to 174 known porcine pre-miRNAs which are also mapped to the pig genome. Specifically, 169 miRNAs and 34 miRNA*s are known in miRBase, and 98 were not previously identified (and thus new) porcine miRNA*s; (2) 131 miRNAs corresponding to 87 other known miRBase mammalian pre-miRNAs which are mapped to the pig genome. These miRNAs were labeled PN(a) (porcine novel, “a” type); (3) 79 miRNAs (the most abundant isomiR encoding from the pre-miRNAs in group 1 or 3) corresponding to 79 novel predicted hairpins in the pig genome. These miRNAs were labeled PN(b) (porcine novel, “b” type); (4) 45 miRNAs (mapped to 169 other known miRBase mammalian pre-miRNAs, which do not map to the pig genome) corresponding to 44 novel predicted hairpins in the pig genome. These miRNAs were labeled PN(c) (porcine novel, “c” type); and (5) 458 miRNAs (longer than 18 nt and unmapped to any known miRBase mammalian pre-miRNAs) encompassing 419 candidate pre-miRNAs, which were predicted RNA hairpins derived from the pig genome, and were labeled PC (porcine candidate).

In order to obtain high confidence expression data for the analysis of relationship between miRNA expression and methylation changes, we only used the counts of the predominantly expressed mature product for miRNAs and their respective miRNA*s. In addition, the pre-miRNAs that mapped to multiple genomic loci, or the distinct pre-miRNAs that expressed identical mature miRNAs were also excluded.

Supplemental References

60. Junker, B.H., Klukas, C. & Schreiber, F. VANTED: a system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinformatics* **7**, 109 (2006).

61. Sebert, S.P. *et al*. Childhood obesity and insulin resistance in a Yucatan mini-piglet model: putative roles of *IGF-1* and muscle *PPARs* in adipose tissue activity and development. *Int. J. Obes. (Lond)* **29**, 324-333 (2005).

62. Lefaucheur, L., Ecolan, P., Plantard, L. & Gueguen, N. New insights into muscle fiber types in the pig. *J. Histochem. Cytochem.* **50**, 719-730 (2002).

63. De Campos, R. *et al*. Fatty acid and volatile compounds from salami manufactured with yerba mate (*Ilex paraguariensis*) extract and pork back fat and meat from pigs fed on diets with partial replacement of maize with rice bran. *Food Chem.* **103**, 1159-1167 (2007).

64. Saeed, A.I. *et al*. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**, 374-378 (2003).

65. Pruitt, K.D., Tatusova, T., Klimke, W. & Maglott, D.R. NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic. Acids Res.* **37**, D32-D36 (2009).

66. Gardner, P.P. *et al*. Rfam: updates to the RNA families database. *Nucleic. Acids Res.* **37**, D136-D140 (2009).

67. Kohany, O., Gentles, A.J., Hankus, L. & Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* **7**, 474 (2006).

68. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic. Acids Res.* **39**, D152-D157 (2011).

69. Markham, N.R. & Zuker, M. UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **453**, 3-31 (2008).

70. Glazov, E.A. *et al*. A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res.* **18**, 957-964 (2008).

71. Kuchenbauer, F. *et al*. In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res.* **18**, 1787-1797 (2008).

72. Morin, R.D. *et al*. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* **18**, 610-621 (2008).

73. Ambros, V. *et al*. A uniform system for microRNA annotation. *RNA* **9**, 277-279 (2003).