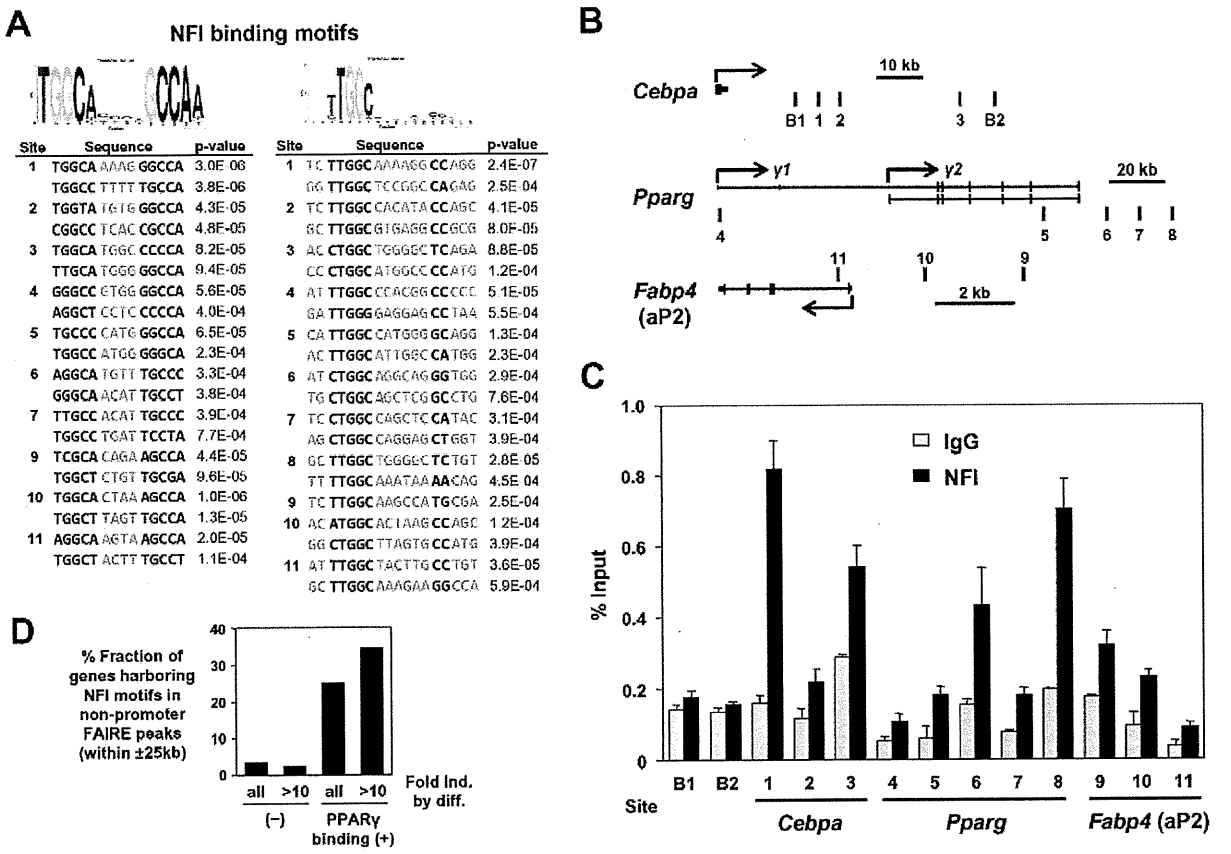


Figure 8. NFI occupy the adipocyte-specific FAIRE peaks and/or the PPARγ binding sites near PPARγ, C/EBPα, and aP2 genes. (A) The NFI binding motifs identified in the adipocyte-specific FAIRE peaks and/or the PPARγ binding sites in the vicinity of PPARγ, C/EBPα and aP2. For site numbers, see (B). (B) Genomic location of the regions examined. B1 and B2 are unrelated genomic regions used as background negative controls. (C) ChIP-qPCR analysis using an anti-NFI antibody (H-300). (D) Percent fraction of genes harboring NFI motifs in non-promoter FAIRE peaks (within ±25 kb) were higher when the genes were bound by PPARγ (within ±25 kb) and induced during differentiation.
doi:10.1371/journal.pgen.1002311.g008

expression appears mediated through multiple regulatory elements distal to transcription start sites (TSSs): greater induction of gene expression by differentiation means greater likelihood that more adipocyte-specific FAIRE peaks are associated with the gene (Figure 2D and 2E). This implies that optimal gene transcriptional regulation may require coordinated actions of multiple regulatory elements. Therefore, although valuable and informative, the proximal promoter assay may not always be sufficient (e.g., AdipoR2, see Figure S2B and Figure 3). Nevertheless, the importance of proximal promoter regions is obvious given the fact that many proximal promoter regions are successfully used to generate tissue-specific transgenic lines. Recently, Mikkelsen et al. demonstrated in adipocytes that many *cis*-regulatory elements are often not conserved between human and murine adipocytes even though the expression pattern of genes is conserved [28]. They observed that such motifs were located within linage-specific transposon insertions. Existence of multiple regulatory elements around biologically important genes *could* be a mechanism by which cells maintain key gene regulations against genomic changes during evolution. Clustering of regulatory elements could also result from an accumulative effect of such evolutional genomic changes.

Computational motif analysis is used to discover new transcription-factor binding motifs in sequences inferred from genome-wide studies such as ChIP-seq [61]. In genome-wide ChIP analysis of transcription factors, motif analysis is used to obtain their accurate binding motifs *and* discover unknown DNA binding factors that co-localize with the transcription factors of interest, for example, see [27,62,63]. The analyses, however, relied on prior knowledge about transcription factors and the regions to be analyzed are limited to their biding sites. In contrast, the combination of motif analyses and mapping of regulatory elements by FAIRE-seq does not require such prior knowledge, hence offers a distinct advantage in unbiased screening for novel transcription factors important in given biological processes. In our study, we retrieved the motifs for PPARγ and C/EBPs and for known regulators that top the list of the motifs identified in the adipocyte- or preadipocyte-specific FAIRE peaks (Figure 5, Figures S6 and S7). Furthermore, we demonstrated that NFIA and NFIB were functionally required for proper adipocyte differentiation (Figure 6). These results demonstrated that motif analyses of cell type–specific FAIRE peaks are useful in identifying regulators of a biological process in an unbiased manner.

To our knowledge, few studies have employed motif analysis and our unbiased approaches in investigating enhancer-like DNA regions. Mikkelsen et al. recently employed ChIP-seq for H3K27ac to define enhancer regions specific for adipocyte differentiation. Both studies similarly detected the motifs for PPAR, C/EBPs and

AP-1 in the most enriched motifs. There are, however, differences. Mikkelsen discovered PLZF and SRF as novel negative regulators [28] and we found NFIA and NFIB as regulators of adipocyte differentiation—perhaps due to differences in methods. First, we directly compared FAIRE peaks and H3K27ac peaks detected in the Mikkelsen study and found considerable, but not complete, overlap especially in the non-promoter regions: 94% of 10,461 promoter FAIRE peaks and 45% of 27,320 non-promoter FAIRE peaks overlapped H3K27ac in 3T3-L1 on day 0. There may be different classes of enhancer elements that prefer either H3K27ac or open chromatin. Also, we used two parameters to sort motifs: the statistical significance of enrichment (p-value) in either kind of cell type–specific FAIRE peaks; and the motif enrichment ratio between the adipocyte- and preadipocyte-specific FAIRE peaks (see [28]). The combination guarantees significant enrichment of the peaks' motifs and the difference in their number depending on whether they are adipocyte- or preadiocyle-specific. The motifs for PLZF and SRF were not on the top of our list since the p-values were not significant—probably due to relatively lower occurrence, although we also found a significant enrichment ratio of 0.37 and 0.50, respectively. We calculated p-values and the enrichment ratios of the top motifs in the Mikkelsen's study by using our adipocyte- and preadipocyte-specific FAIRE peaks and found general similarity (Figure S10). Overall, both studies notably demonstrate the utility of the combining computational motif analysis and unbiased mapping of regulatory elements in identifying new regulators of adipocyte differentiation.

Siersbæk et al. recently employed DNase-seq to investigate genome-wide change in open chromatin structure at various time points during 3T3-L1 differentiation [6]. They reported dramatic increase in the number of open chromatin sites in the first hours of differentiation. Such regions included what they termed "hot spots" that were bound by multiple adipogenic regulators, facilitating binding of PPARγ and C/EBPα during the late stage of differentiation. We found that the DNaseI hypersensitive sites in 3T3-L1 cells on day 0 or day 6 in the Siersbaek study [6] significantly overlapped the FAIRE peaks on day 0 or day 8 in our study (78.8% and 80.9%, respectively) (Figure S11), suggesting that both methods detect similar open chromatin regions. Although limited amount of motif analyses of the DNase I sites was conducted in their study, we think a combination of motif analysis and DNase-seq should work in a similar way.

The NFI family was identified as site-specific DNA-binding protein that bound to the adenovirus origin of replication [56,57]. Although defects in development of organs such as brain, lung, tooth, bone and skeletal muscle in *Nfia*, *Nfib*, *Nifc* and *Nfix*-deficient mice were documented [64,65,66,67,68,69], no publication has reported direct evidence that NFI family transcription factors are involved in adipogenesis, but it is a reasonable supposition since bone, muscle and adipocytes have a common mesenchymal precursor [70]. Interestingly, Graves et al. demonstrated that NFI was bound to the adipogenic −5.4 kb enhancer region in the aP2 promoter [71], which is the original adipogenic enhancer region where the PPARγ/RXR heterodimer was found to bind and act [72]. The NFI binding motif they examined by gel shift assay [72] was close to the best-characterized PPARγ binding sites in the region, and was also in site 9 (Figure 8A, right panel, site 9), which was indeed bound by NFI in ChIP assay (Figure 8C). Forced overexpression of NFIA in 3T3-L1 cells dramatically induced expression of PPARγ, C/EBPα and aP2 and caused lipid droplet formation before initiation of differentiation. Our ChIP data suggest that activation of these genes by NFIA is through direct binding of NFI to regulatory elements near these genes. In overexpression experiments, NFIB did not activate the adipogenic genes (Figure 7). NFI factors are known to undergo extensive alternative splicing [57]. We speculate that this could be due to truncation of the C-terminus caused by lack of exons 10 and 11 in the NFIB cDNA that we cloned (NM_001113209.1) while the NFIA clone completely matched NM_010905.3. NFI was also implicated in functions of other nuclear receptors such as the androgen receptor (AR), estrogen receptor (ER) and glucocorticoid receptor [4,73,74]. Further studies are necessary to elucidate the mode of action of NFIs and positioning of NFIs in the adipogenic regulatory network.

Materials and Methods

Cell Culture

3T3-L1, NIH-3T3, 3T3-F442A and HEK293T cells were maintained in DMEM, supplemented with 10% FBS. For adipocyte differentiation, two days after confluence, 3T3-L1 cells were treated with dexamethasone (1 μM), IBMX (0.5 mM), and insulin (5 μg/ml) (DMI) for 48 hours, followed by treatment with insulin alone, with medium replacement every two days thereafter. For differentiation of 3T3-F442A, cells were treated with insulin (5 μg/ml) after confluence, with medium replacement every two days.

Animal Studies

All animal works have been conducted according to the institutional guidelines.

Antibodies

Generation of characterization of antibodies for human PPARγ and human RXRα was described previously [24]. Rabbit polyclonal anti-histone H3 trimethyl K4 (ab8580) was from Abcam. Antibodies against CTCF were from Upstate (#07–729). Anti-NFI antibody (H-300) was from Santa Cruz (sc-5567).

FAIRE

FAIRE experiments were performed based on a protocol published by Giresi et al. [7]. Briefly, cells were fixed with 1% formaldehyde for five minutes at room temperature, the fixation stopped by adding 2.5 M glycine (final 125 mM). Fixed cells were scraped and collected in 15 ml tubes (4×10^6 cells/tube) and washed twice with cold PBS, then 8×10^6 cells were re-suspended in 800 μl of MC lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% NP-40) and incubated on ice for ten minutes. After spinning for four minutes at 8000 rpm, the pellet was re-suspended in 400 μl SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0, proteinase inhibitor cocktail) and incubated on ice for ten minutes. Glass beads (size, 200 mg) (Polysciences Inc. #05483-250) were added and the DNA was sheared by sonicator. Next, we added 200 μl cold ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH 8.0, 167 mM NaCl), and after spinning for one minute at 8,000 rpm, supernatant was transferred to a new 1.5 ml tube. Aliquote was taken, de-crosslinked, purified by phenol/chloroform extraction, and run on a gel to ensure average fragment sizes of 300 bp. Remaining samples were processed three times by phenol/chloroform extraction to recover DNA not bound by nucleosome in the water phase. The samples were de-crosslinked by overnight incubation at 65°C and purified by ethanol precipitation. They were subsequently treated with RNase A (final 50 ug/ml), purified by QIAquick PCR purification kit (Qiagen) and used for subsequent analyses.

Chromatin Immunoprecipitation (ChIP)

ChIP was performed as descried previously [24,75]. For ChIP using anti-PPARγ, RXRα and CTCF antibodies, 3T3-L1 cells

were cross-linked with 1% formaldehyde for ten minutes at room temperature and were prepared for ChIP as described previously. For ChIP using anti-H3K4me3 antibody, the nuclei of 3T3-L1 cells were prepared by centrifugation through a sucrose gradient and were digested with MNase (TaKaRa). After centrifugation, the supernatant was used for ChIP. Sequences of primers used for qPCR were listed in Table S1.

High-Throughput Sequencing and Peak Detection

High-throughput sequencing was performed by using the Genome Analyzer System (GA II) (Illumina) as described elsewhere [76]. In short, we repaired ends of DNA samples, created 3′-dA overhang, ligated Illumina adaptors, size-fractioned the samples by gel extraction and amplified them with 8 cycles of PCR according to the manufacturer's instructions. We then purified the DNA and performed cluster generation and 36 cycles of sequencing on an Illumina cluster station and 1G analyzer following the manufacturer's instructions. Sequences were mapped to the reference murine genome, NCBI build 37 (mm9). Peak detection was performed using Findpeaks 3.1.9.2 [77] with a false discovery rate (FDR) cut-off of 1×10^{-4}. Operations such as intersections, unions, and subtractions of genome regions were performed with a web-based GALAXY genome analysis tool [78,79].

Average Signal Profiling

Average profiling of FAIRE and histone modifications near transcription start sites or FAIRE peaks were generated using "sitepro" in the CEAS package [80].

Adipocyte- and Preadipocyte-Specific FAIRE Peaks

For definition, we first ranked peaks based on signal intensity, which were detected in 3T3-L1 on either day 0 or day 8 with a FDR of 10^{-4}. We then classified each peak into tertiles (high, mid, low) for either day the peak that had the higher percentile (see also the 4-by-4 table in Figure 3B).

Gene Ontology Analysis

Gene ontology annotation analysis was performed using DAVID (ver. 6.7) [38]. The top 2,000 genes were used, sorted by the number and maximum height of the adipocyte-specific FAIRE peaks within a region ±25 kb from TSS. For genes bound by PPARγ, we used the top 931 genes with more than three PPARγ binding sites within a region ±25 kb from TSS. To detect enrichment of specific—rather than general—terms, following the instructions of DAVID's developer, we used GOTERM_BP_4 and GOTERM_BP_5, and sorted result lists by using both fold enrichment and Benjaini p-value [38,81].

Clustering Analysis

Statistical clustering analyses of the PPARγ binding sites and the adipocyte-specific FAIRE peaks were performed as described in references [47,48].

Enriched Motif Analysis

Enrichment analyses of known motifs were performed with AME ver. 4.6.0 in the MEME suite [82]. After removing repeat regions with RepeatMasker [83], DNA sequences from the center 150 bp regions of the top 2,000 cell type–specific FAIRE peaks were analyzed with a fixing partition of 2,000, dinucleotide randomization and p-value threshold of 10^{-4} and p-value report threshold of 0.05. We used the licensed version of TRANSFAC database (Release 2010.4) [51] and the JASPAR CORE database [52].

Motif enrichment ratios (adipocyte-/preadipocytes-specific FAIRE) for motifs in the TRANSFAC or JASPAR CORE database were determined by a method described in reference [28]. Instances of motifs were enumerated in the adipocyte- or preadipocytes-specific FAIRE peaks by using FIMO ver. 4.6.0 in the MEME suite, with a p-value threshold of 10^{-4}, normalized by total nucleotide length. Motif enrichment ratios were determined by dividing the normalized adipocyte enrichment values by preadipocyte values.

MEME ver. 4.3.0 [40] was used to identify de novo motifs over-represented in the adipocyte- or preadipocyte-specific FAIRE peaks and the PPARγ binding sites. After removing repeat regions with RepeatMasker [83], DNA sequences from the center 150 bp regions of the top 800 cell type–specific FAIRE peaks with higher signals were used for the analyses. Identified enriched de novo motifs were next analyzed by TOMTOM in the MEME suite for comparison against a database of known motifs.

Gel Shift Assay and Reporter Assay

The Gel shift assay and luciferase reporter assay were performed as previously described [84,85]. For the luciferase assay, putative PPRE motifs were cloned in tandem (3× or 6×) into pGL3 basic reporter plasmid (Promega) together with the tk minimal promoter. The −5.4 kb aP2 promoter luciferase construct is described in reference [84].

Knockdown of NFIA and NFIB by siRNA in 3T3-L1 Cell Differentiation

The 3T3-L1 cells were transfected with either control siRNA or siRNA for murine NFIA and NFIB (Santa Cruz Biotechnology, sc-37007, sc-36045 and sc-43566, Sigma MISSION siRNA, SASI_Mm02_00309629, 00309630, 00307243, 00307244) by using Lipofectamine RNAiMAX (Invitrogen) just before they reached confluence. Induction of differentiation (the DMI treatment) was started two days after confluence, as described in a method for differentiation of 3T3-L1 cells.

Oil-Red-O Staining

The 3T3-L1 adipocytes were washed with PBS, fixed with formalin for 30 minutes at room temperature, rinsed with 60% isopropanol and stained with oil red O solution—freshly made by mixing 0.5% oil red O in isopropyl alcohol and water (3:2)—and left to sit for one hour; the cells were then washed with water and dried.

mRNA Expression Analysis

Total RNA was isolated using TRIzol reagent (Invitrogen), then 0.5 μg of the total RNA was reverse transcribed using high-capacity cDNA reverse transcription kits (Applied Biosystems #4375222) and random hexamers. Real-time quantitative PCR (SYBR green) analysis was performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Primer sequences are listed in Table S1. Expression was normalized to 36B4.

Microarray Analysis

Transcriptome analysis of 3T3-L1 during differentiation by using a GeneChip Mouse Genome 430 2.0 array (Affimetrix) was described previously [24]. Heat maps were generated by using GENOMICA, developed by Yaniv Lubling and Eran Segal at the Weizmann Institute of Science. Microarray data of 3T3-L1 and NIH-3T3 cells used in Figure S11 was obtained from GEO (accession number GSE10246).

Retroviral Expression System

We amplified NFIA and NFIB coding sequences from cDNA prepared from adipocytes using primers listed in Table S1, and cloned them into retroviral pMXs-puro vectors. We also made a dominant negative NFIA that lacks the C-terminal transactivation/repression domain (NFIA-DN) [58]. Plat E cells were transfected with pMXs-puro plasmids using Lipofectamine 2000 (Invitrogen). Culture medium containing viruses after two day incubation was centrifuged at 2,000 rpm for 5 min and supernatant was collected and supplemented with 10 μg/ml polybrene. Conditioned medium with viruses was used to infect 3T3-L1 cells and then selection was started by adding 2 μg/ml puromycin and incubated for 2 days.

Accession Numbers

FAIRE-seq and ChIP-seq raw data are deposited into the DNA data bank of Japan (DDBJ accession number: DRA000378).

Supporting Information

Figure S1 Genomic distribution and characterization of promoter and non-promoter FAIRE peaks in 3T3-L1. (A) Location analysis of FAIRE peaks relative to RefSeq genes in 3T3-L1 (day 0). Promoter FAIRE peaks were defined as those located within +/−500 bp of RefSeq transcription start sites (TSSs). Notably, only 8% of the non-promoter FAIRE peaks were located in the −5 kb proximal promoter region, and the vast majority of them were located in distal regions such as introns and intergenic regions. (B) Average profiles of FAIRE and H3K4me3 signals around the TSSs of genes with high, moderate and low expression levels. Signal intensity from microarray data was used for classification by the signal's expression levels. The X-axis indicates distance from the TSS. (C) Percent fractions of the FAIRE peaks (promoter and non-promoter) that overlapped CTCF binding sites as well as H3K4me1 and H3K4me3 positive regions. (D) Average profiles of FAIRE, H3K4me1, H3K4me3, and H3K27ac signals around the FAIRE peaks in promoter and non-promoter regions. The X-axis shows distance from the center of the FAIRE peaks. The FAIRE peaks located within +/−100 bp from RefSeq TSSs were analyzed for promoter FAIRE peaks. The promoter FAIRE peaks showed H3K4me3(+)/H3K4me1(−) modification whereas the non-promoter FAIRE peaks showed H3K4me3(−)/H3K4me1(+) modification.
(TIF)

Figure S2 Clustering of multiple adipocyte-specific non-promoter FAIRE peaks and PPARγ binding sites near *Mgll*, *Adipor2* and *Slc2a4*. Clusters of multiple adipocyte-specific FAIRE peaks and/or PPARγ binding sites were located in genomic regions near *Mgll* (A), *Adipor2* (B) and *Slc2a4*(Glut4) (C) in 3T3-L1 adipocytes. In some cases—e.g., *Slc2ar4* (Glut4) and *Ybx2* in (C)—multiple genes were located in such regions. Bars below the FAIRE signal represent statistically significant FAIRE positive peaks (FDR<10^{-4}). Red asterisks indicate the adipocyte-specific FAIRE peaks on day 8 (see Figure 2B for definition). Blue arrow heads in (B) indicate the PPARγ binding regions in the intron 1 of *Adipor2* tested in Figure 3.
(TIF)

Figure S3 Clustering of multiple adipocyte-specific non-promoter FAIRE peaks and PPARγ binding sites near *Cebpa*, *Srebf1*and *Cidec*. Clusters of multiple adipocyte-specific FAIRE peaks and/or PPARγ binding sites were located in genomic regions near *Cebpa* (A), *Srebf1* (B) and *Cidec* (C).
(TIF)

Figure S4 Binding sites for PPARγ and RXRα in 3T3-L1 cells. (A) De novo motif analysis (MEME) of the center 150 bp of the PPARγ/RXRα binding regions (top 400) in 3T3-L1, day 8. Of note, there is a 5′ extension AGT, which corresponds to the interaction between the PPARγ hinge region and DNA identified by crystal structure analysis [41]. (B) A heat map showing enrichment of PPARγ in the vicinity of genes up-regulated during differentiation. The horizontal bars in the right panel indicate each gene bound by PPARγ (+/−25 kb from TSS, day 8) (C) Ontology analysis with DAVID of genes bound by PPARγ [13].
(TIF)

Figure S5 Co-regulation of neighboring genes during adipocyte differentiation. (A, B) Genomic loci near (A) co-regulated *Mrpl12*, *Slc25a10* and *Gcgr* and (B) co-regulated *Hsd11b1*, *G0s2* and *Lamb3*. Note, there are clusters of the adipocyte-specific FAIRE peaks (asterisks) and the PPARγ binding sites encompassing the co-regulated genes. (C) Microarray analysis showing co-regulation of *Mrpl12*, *Slc25a10* and *Gcgr*, and co-regulation of *Hsd11b1*, *G0s2* and *Lamb3* included in the clusters of multiple adipocyte-specific FAIRE peak and PPARγ binding sites.
(TIF)

Figure S6 Known motif enrichment analysis of the adipocyte- or preadipocyte-specific FAIRE peaks (JASPAR CORE motifs). Enrichment analysis of the adipocyte- (left) and the preadipocyte-specific (right) FAIRE peaks for known motifs in the JASPAR CORE database performed with AME in the MEME suite by the same methods used in Figure 5.
(TIF)

Figure S7 De novo motif analysis of the adipocyte-specific FAIRE peaks. MEME ver. 4.3.0 was used to identify de novo motifs over-represented in the adipocyte- and preadipocyte-specific FAIRE peaks and PPARγ binding sites. After removing repeat regions, DNA sequences from the center 150 bp regions of top 800 cell type–specific FAIRE peaks with higher signals were used for the analyses. Identified enriched de novo motifs were analyzed by TOMTOM in the MEME suite for comparison against a database of known motifs.
(TIF)

Figure S8 Suppression of adipocyte differentiation by knockdown of NFIA and NFIB by using different siRNAs.
(TIF)

Figure S9 Comparison of FAIRE Peaks between undifferentiated 3T3-L1 and NIH-3T3 cells. (A) A heat map showing enrichment of the 3T3-L1- and NIH-3T3-specific FAIRE peaks in the vicinity (+/−25 kb from TSS) of genes sorted by using the ratio of expression levels in 3T3-L1 or NIH-3T3. The FAIRE peaks specific to 3T3-L1 or NIH-3T3 were enriched in the vicinity of genes whose expression levels were higher in 3T3-L1 or NIH-3T3, respectively. (B) Known motif analysis of the 3T3-L1-specific FAIRE peaks (vs NIH-3T3). The binding motif for EBF and PPARγ/RXR were among the top scored motifs.
(TIF)

Figure S10 The enrichment ratios of the top motifs in Mikkelsen's study [28] by using the adipocyte- and preadipocyte-specific FAIRE peaks.
(TIF)

Figure S11 Comparison of DNase-seq in Siersbæk's study [6] and FAIRE-seq peaks near *Klf5*, *Pparg* and *Cebpa* gene. DHS stands for DNase I hypersensitive sites.
(TIF)

Table S1 Sequences of primers.
(DOC)

Acknowledgments

The authors would like to thank: Jason Lieb for providing the FAIRE protocol; the DNA data bank of Japan and Database Center for Life Science (DBCLS) for providing computer resources; all the staff and members of Takashi Kadowaki's lab for technical help and discussion; Kaori Shiina, Shogo Yamamoto, Genta Nagae, Yasuharu Kanki, Atsushi Okabe, Yoichiro Wada, Seitaro Nomura, Kenta Magoori, Takeshi Inagaki, Toshiya Tanaka, and other members of the Laboratory of Systems Biology and Medicine, Research Center, Daizo-Koinuma in the Department of Molecular Pathology; Kenichi Takayama and Satoshi Inoue in the Department of Geriatric Medicine, Yumiko Oishi-Tanaka and Ichiro Manabe in the Department of Cardiovascular Medicine, Kousuke Watanabe and Daiya Takai in the Department of Respiratory Medicine, Graduate School of Medicine, the University of Tokyo—all, for technical help and suggestions.

Author Contributions

Conceived and designed the experiments: H Waki, M Nakamura, T Yamauchi, K Wakabayashi. Performed the experiments: H Waki, M Nakamura, K Wakabayashi, J Yu, L Hirose-Yotsuya, K Take, W Sun, T Aoyama. Analyzed the data: H Waki, M Nakamura, K Wakabayashi, T Fujita, S Tsutsumi, T Yamauchi, M Iwabu, M Okada-Iwabu. Contributed reagents/materials/analysis tools: H Waki, M Nakamura, K Wakabayashi, T Fujita, S Tsutsumi. Wrote the paper: H Waki, M Nakamura. Supervised the design of the experiments: K Ueki, T Kodama, T Yamauchi, S Tsutsumi, J Sakai, H Aburatani, T Kadowaki.

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.
2. Wu C (1980) The 5′ ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. Nature 286: 854–860.
3. Song L, Crawford GE (2010) DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. Cold Spring Harb Protoc 2010: pdb prot5384.
4. John S, Sabo PJ, Thurman RE, Sung MH, Biddie SC, et al. (2011) Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat Genet.
5. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, et al. (2009) Histone modifications at human enhancers reflect global cell type–specific gene expression. Nature 459: 108–112.
6. Siersbaek R, Nielsen R, John S, Sung MH, Baek S, et al. (2011) Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. Embo J 30: 1459–1472.
7. Giresi PG, Lieb JD (2009) Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). Methods 48: 233–239.
8. Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome Res 17: 877–885.
9. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447: 799–816.
10. Gaulton KJ, Nammo T, Pasquali L, Simon JM, Giresi PG, et al. (2010) A map of open chromatin in human pancreatic islets. Nat Genet 42: 255–259.
11. Rosen E, Eguchi J, Xu Z (2009) Transcriptional targets in adipocyte biology. Expert Opin Ther Targets 13: 975–986.
12. Waki H, Tontonoz P (2007) Endocrine functions of adipose tissue. Annu Rev Pathol 2: 31–56.
13. Barak Y, Nelson MC, Ong ES, Jones YZ, Ruiz-Lozano P, et al. (1999) PPAR gamma is required for placental, cardiac, and adipose tissue development. Mol Cell 4: 585–595.
14. Kubota N, Terauchi Y, Miki H, Tamemoto H, Yamauchi T, et al. (1999) PPAR gamma mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. Mol Cell 4: 597–609.
15. Rosen ED, Sarraf P, Troy AE, Bradwin G, Moore K, et al. (1999) PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. Mol Cell 4: 611–617.
16. Tontonoz P, Hu E, Spiegelman BM (1994) Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. Cell 79: 1147–1156.
17. Imai T, Takakuwa R, Marchand S, Dentz E, Bornert JM, et al. (2004) Peroxisome proliferator-activated receptor gamma is required in mature white and brown adipocytes for their survival in the mouse. Proc Natl Acad Sci U S A 101: 4543–4547.
18. Lehmann JM, Moore LB, Smith-Oliver TA, Wilkison WO, Willson TM, et al. (1995) An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). J Biol Chem 270: 12953–12956.
19. Wu Z, Bucher NL, Farmer SR (1996) Induction of peroxisome proliferator-activated receptor gamma during the conversion of 3T3 fibroblasts into adipocytes is mediated by C/EBPbeta, C/EBPdelta, and glucocorticoids. Mol Cell Biol 16: 4128–4136.
20. Tontonoz P, Spiegelman BM (2008) Fat and beyond: the diverse biology of PPARgamma. Annu Rev Biochem 77: 289–312.
21. Nielsen R, Pedersen TA, Hagenbeek D, Moulos P, Siersbaek R, et al. (2008) Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. Genes Dev 22: 2953–2967.
22. Lefterova MI, Zhang Y, Steger DJ, Schupp M, Schug J, et al. (2008) PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. Genes Dev 22: 2941–2952.
23. Nakachi Y, Yagi K, Nikaido I, Bono H, Tonouchi M, et al. (2008) Identification of novel PPARgamma target genes by integrated analysis of ChIP-on-chip and microarray expression data during adipocyte differentiation. Biochem Biophys Res Commun 372: 362–366.
24. Wakabayashi K, Okamura M, Tsutsumi S, Nishikawa NS, Tanaka T, et al. (2009) The peroxisome proliferator-activated receptor gamma/retinoid X receptor alpha heterodimer targets the histone modification enzyme PR-Set7/Setd8 gene and regulates adipogenesis through a positive feedback loop. Mol Cell Biol 29: 3544–3555.
25. Hamza MS, Pott S, Vega VB, Thomsen JS, Kandhadayar GS, et al. (2009) De-novo identification of PPARgamma/RXR binding sites and direct targets during adipogenesis. PLoS ONE 4: e4907. doi:10.1371/journal.pone.0004907.
26. Steger DJ, Grant GR, Schupp M, Tomaru T, Lefterova MI, et al. (2010) Propagation of adipogenic signals through an epigenomic transition state. Genes Dev 24: 1035–1044.
27. Lefterova MI, Steger DJ, Zhuo D, Qatanani M, Mullican SE, et al. (2010) Cell-specific determinants of peroxisome proliferator-activated receptor gamma function in adipocytes and macrophages. Mol Cell Biol 30: 2078–2089.
28. Mikkelsen TS, Xu Z, Zhang X, Wang L, Gimble JM, et al. (2010) Comparative epigenomic analysis of murine and human adipogenesis. Cell 143: 156–169.
29. Okamura M, Kudo H, Wakabayashi K, Tanaka T, Nonaka A, et al. (2009) COUP-TFII acts downstream of Wnt/beta-catenin signal to silence PPARgamma gene expression and repress adipogenesis. Proc Natl Acad Sci U S A 106: 5819–5824.
30. Sakabe NJ, Nobrega MA (2010) Genome-wide maps of transcription regulatory elements. Wiley Interdiscip Rev Syst Biol Med. pp 422–437.
31. Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. Cell 137: 1194–1211.
32. Mori T, Sakaue H, Iguchi H, Gomi H, Okada Y, et al. (2005) Role of Kruppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. J Biol Chem 280: 12867–12875.
33. Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35: D61–65.
34. Robertson AG, Bilenky M, Tam A, Zhao Y, Zeng T, et al. (2008) Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. Genome Res 18: 1906–1917.
35. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, et al. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39: 311–318.
36. Karlsson M, Contreras JA, Hellman U, Tornqvist H, Holm C (1997) cDNA cloning, tissue distribution, and identification of the catalytic triad of monoglyceride lipase. Evolutionary relationship to esterases, lysophospholipases, and haloperoxidases. J Biol Chem 272: 27218–27223.
37. Nishino N, Tamori Y, Tateya S, Kawaguchi T, Shibakusa T, et al. (2008) FSP27 contributes to efficient energy storage in murine white adipocytes by promoting the formation of unilocular lipid droplets. J Clin Invest 118: 2808–2821.
38. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44–57.
39. Kliewer SA, Umesono K, Noonan DJ, Heyman RA, Evans RM (1992) Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. Nature 358: 771–774.
40. Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2: 28–36.
41. Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, et al. (2008) Structure of the intact PPAR-gamma-RXR-alpha nuclear receptor complex on DNA. Nature: 350–356.

42. Schmidt SF, Jorgensen M, Chen Y, Nielsen R, Sandelin A, et al. (2011) Cross-species comparison of C/EBPa and PPARg profiles in mouse and human adipocytes reveals interdependent retention of binding sites. NCBI GEO (Gene Experssion Omnibus): GSE27450.
43. Yamauchi T, Kamon J, Ito Y, Tsuchida A, Yokomizo T, et al. (2003) Cloning of adiponectin receptors that mediate antidiabetic metabolic effects. Nature 423: 762–769.
44. Kadowaki T, Yamauchi T, Kubota N, Hara K, Ueki K, et al. (2006) Adiponectin and adiponectin receptors in insulin resistance, diabetes, and the metabolic syndrome. J Clin Invest 116: 1784–1792.
45. Tsuchida A, Yamauchi T, Takekawa S, Hada Y, Ito Y, et al. (2005) Peroxisome proliferator-activated receptor (PPAR)alpha activation increases adiponectin receptors and reduces obesity-related inflammation in adipose tissue: comparison of activation of PPARalpha, PPARgamma, and their combination. Diabetes 54: 3358–3370.
46. Sun X, Han R, Wang Z, Chen Y (2006) Regulation of adiponectin receptors in hepatocytes by the peroxisome proliferator-activated receptor-gamma agonist rosiglitazone. Diabetologia 49: 1303–1310.
47. Crawford GE, Holt IE, Whittle J, Webb BD, Tai D, et al. (2006) Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). Genome Res 16: 123–131.
48. Stitzel ML, Sethupathy P, Pearson DS, Chines PS, Song L, et al. (2010) Global epigenomic analysis of primary human pancreatic islets provides insights into type 2 diabetes susceptibility loci. Cell Metab 12: 443–455.
49. Ji H, Vokes SA, Wong WH (2006) A comparative analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. Nucleic Acids Res 34: e146.
50. Ebisuya M, Yamamoto T, Nakajima M, Nishida E (2008) Ripples from neighbouring transcription. Nat Cell Biol 10: 1106–1113.
51. Wingender E, Chen X, Hehl R, Karas H, Liebich I, et al. (2000) TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 28: 316–319.
52. Bryne JC, Valen E, Tang MH, Marstrand T, Winther O, et al. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. Nucleic Acids Res 36: D102–106.
53. Gupta RK, Arany Z, Seale P, Mepani RJ, Ye L, et al. (2010) Transcriptional control of preadipocyte determination by Zfp423. Nature 464: 619–623.
54. Tominaga S, Yamaguchi T, Takahashi S, Hirose F, Osumi T (2005) Negative regulation of adipogenesis from human mesenchymal stem cells by Jun N-terminal kinase. Biochem Biophys Res Commun 326: 499–504.
55. Hu E, Kim JB, Sarraf P, Spiegelman BM (1996) Inhibition of adipogenesis through MAP kinase-mediated phosphorylation of PPARgamma. Science 274: 2100–2103.
56. Nagata K, Guggenheimer RA, Hurwitz J (1983) Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc Natl Acad Sci U S A 80: 6177–6181.
57. Gronostajski RM (2000) Roles of the NFI/CTF gene family in transcription and development. Gene 249: 31–45.
58. Namihira M, Kohyama J, Semi K, Sanosaka T, Deneen B, et al. (2009) Committed neuronal precursors confer astrocytic potential on residual neural precursor cells. Dev Cell 16: 245–255.
59. Green H, Kehinde O (1974) Sublines of mouse 3T3 cells that accumulate lipid. Cell 1: 113–116.
60. Jimenez MA, Akerblad P, Sigvardsson M, Rosen ED (2007) Critical role for Ebf1 and Ebf2 in the adipogenic transcriptional cascade. Mol Cell Biol 27: 743–757.
61. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 10: 669–680.
62. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, et al. (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell 122: 33–43.
63. Koinuma D, Tsutsumi S, Kamimura N, Taniguchi H, Miyazawa K, et al. (2009) Chromatin immunoprecipitation on microarray analysis of Smad2/3 binding sites reveals roles of ETS1 and TFAP2A in transforming growth factor beta signaling. Mol Cell Biol 29: 172–186.
64. das Neves L, Duchala CS, Tolentino-Silva F, Haxhiu MA, Colmenares C, et al. (1999) Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. Proc Natl Acad Sci U S A 96: 11946–11951.
65. Steele-Perkins G, Plachez C, Butz KG, Yang G, Bachurski CJ, et al. (2005) The transcription factor gene Nfib is essential for both lung maturation and brain development. Mol Cell Biol 25: 685–698.
66. Steele-Perkins G, Butz KG, Lyons GE, Zeichner-David M, Kim HJ, et al. (2003) Essential role for NFI-C/CTF transcription-replication factor in tooth root development. Mol Cell Biol 23: 1075–1084.
67. Messina G, Biressi S, Monteverde S, Magli A, Cassano M, et al. (2010) Nfix regulates fetal-specific transcription in developing skeletal muscle. Cell 140: 554–566.
68. Plachez C, Lindwall C, Sunn N, Piper M, Moldrich RX, et al. (2008) Nuclear factor I gene expression in the developing forebrain. J Comp Neurol 508: 385–401.
69. Driller K, Pagenstecher A, Uhl M, Omran H, Berlis A, et al. (2007) Nuclear factor I X deficiency causes brain malformation and severe skeletal defects. Mol Cell Biol 27: 3855–3867.
70. Park KW, Halperin DS, Tontonoz P (2008) Before they were fat: adipocyte progenitors. Cell Metab 8: 454–457.
71. Graves RA, Tontonoz P, Ross SR, Spiegelman BM (1991) Identification of a potent adipocyte-specific enhancer: involvement of an NF-1-like factor. Genes Dev 5: 428–437.
72. Tontonoz P, Graves RA, Budavari AI, Erdjument-Bromage H, Lui M, et al. (1994) Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of two nuclear hormone receptors, PPAR gamma and RXR alpha. Nucleic Acids Res 22: 5628–5634.
73. Eeckhoute J, Carroll JS, Geistlinger TR, Torres-Arzayus MI, Brown M (2006) A cell type–specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. Genes Dev 20: 2513–2526.
74. Jia L, Berman BP, Jariwala U, Yan X, Cogan JP, et al. (2008) Genomic androgen receptor-occupied regions with different functions, defined by histone acetylation, coregulators and transcriptional capacity. PLoS ONE 3: e3645. doi:10.1371/journal.pone.0003645.
75. Kaneshiro K, Tsutsumi S, Tsuji S, Shirahige K, Aburatani H (2007) An integrated map of p53-binding sites and histone modification in the human ENCODE regions. Genomics 89: 178–188.
76. Kawase T, Ohki R, Shibata T, Tsutsumi S, Kamimura N, et al. (2009) PH domain-only protein PHLDA3 is a p53-regulated repressor of Akt. Cell 136: 535–550.
77. Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, et al. (2008) FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics 24: 1729–1730.
78. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, et al. (2010) Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol Chapter 19: Unit 19 10 11–21.
79. Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol 11: R86.
80. Shin H, Liu T, Manrai AK, Liu XS (2009) CEAS: cis-regulatory element annotation system. Bioinformatics 25: 2605–2606.
81. http://david.abcc.ncifcrf.gov/forum/cgi-bin/ikonboard.cgi?act = ST;f = 3;t = 1311.
82. McLeay RC, Bailey TL (2009) Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinformatics 11: 165.
83. Chen N (2004) Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics Chapter 4: Unit 4 10.
84. Waki H, Park KW, Mitro N, Pei L, Damoiseaux R, et al. (2007) The small molecule harmine is an antidiabetic cell type–specific regulator of PPARgamma expression. Cell Metab 5: 357–370.
85. Davies BS, Waki H, Beigneux AP, Farber E, Weinstein MM, et al. (2008) The expression of GPIHBP1, an endothelial cell binding site for lipoprotein lipase and chylomicrons, is induced by peroxisome proliferator-activated receptor-gamma. Mol Endocrinol 22: 2496–2504.
86. Schmidt SF, Jorgensen M, Chen Y, Nielsen R, Sandelin A, et al. (2011) Cross species comparison of C/EBPalpha and PPARgamma profiles in mouse and human adipocytes reveals interdependent retention of binding sites. BMC Genomics 12: 152.
87. Graves RA, Tontonoz P, Spiegelman BM (1992) Analysis of a tissue-specific enhancer: ARF6 regulates adipogenic gene expression. Mol Cell Biol 12: 1202–1208.

The EMBO Journal (2011) 30, 2582–2595 | © 2011 European Molecular Biology Organization | All Rights Reserved 0261-4189/11
www.embojournal.org

THE EMBO JOURNAL

Epigenetically coordinated GATA2 binding is necessary for endothelium-specific *endomucin* expression

Yasuharu Kanki[1], Takahide Kohro[1], Shuying Jiang[2], Shuichi Tsutsumi[1], Imari Mimura[1], Jun-ichi Suehiro[1], Youichiro Wada[1], Yoshihiro Ohta[1], Sigeo Ihara[1], Hiroko Iwanari[1], Makoto Naito[2], Takao Hamakubo[1], Hiroyuki Aburatani[1], Tatsuhiko Kodama[1] and Takashi Minami[1,*]

[1]LSBM, The Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan and [2]Department of cellular function, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan

GATA2 is well recognized as a key transcription factor and regulator of cell-type specificity and differentiation. Here, we carried out comparative chromatin immunoprecipitation with comprehensive sequencing (ChIP-seq) to determine genome-wide occupancy of GATA2 in endothelial cells and erythroids, and compared the occupancy to the respective gene expression profile in each cell type. Although GATA2 was commonly expressed in both cell types, different GATA2 bindings and distinct cell-specific gene expressions were observed. By using the ChIP-seq with epigenetic histone modifications and chromatin conformation capture assays; we elucidated the mechanistic regulation of endothelial-specific GATA2-mediated *endomucin* gene expression, that was regulated by the endothelial-specific chromatin loop with a GATA2-associated distal enhancer and core promoter. Knockdown of *endomucin* markedly attenuated endothelial cell growth, migration and tube formation. Moreover, abrogation of GATA2 in endothelium demonstrated not only a reduction of endothelial-specific markers, but also induction of mesenchymal transition promoting gene expression. Our findings provide new insights into the correlation of endothelial-expressed GATA2 binding, epigenetic modification, and the determination of endothelial cell specificity.

The EMBO Journal (2011) **30,** 2582–2595. doi:10.1038/emboj.2011.173; Published online 10 June 2011
Subject Categories: chromatin & transcription; differentiation & death
Keywords: ChIP-seq; *endomucin*; epigenetics; GATA; transcription

*Corresponding author. LSBM, The Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 153-8904, Japan. Tel.: +81 3 5452 5403; Fax: +81 3 5452 5232; E-mail: minami@med.rcast.u-tokyo.ac.jp

Received: 1 September 2010; accepted: 1 May 2011; published online: 10 June 2011

Introduction

Vascular homoeostasis is a critical component of the embryonic and adult stages. Vascular dysfunction is strongly correlated with various diseases; cancer, atherosclerosis, diabetic retinopathy, and sepsis (Minami and Aird, 2005). The endothelium is not only the inner cell layer of vessels, but also the central player for communication with the outer microenvironment. In many cases, endothelial cell activation responding to extracellular signals leads to the spatial and temporal regulation of gene expression. Thus, investigating the molecular mechanisms behind endothelial-cell-specific gene regulation might uncover the key regulators of vascular homoeostasis and onset of vascular diseases.

Currently, six GATA family members have been identified, GATA1 to 6, as belonging to a class of evolutionally conserved C2H2 zinc-finger transcription factors (Patient and McGhee, 2002). In endothelial cells, GATA2, 3, and 6 are selectively expressed (Minami *et al*, 2004). However, their expression levels are not equal. GATA3 is predominantly expressed and functions in endothelial cells derived from large vessels, whereas, GATA2 is uniformly expressed in all endothelial cells. Microvascular endothelial cells preferentially express only GATA2 (Song *et al*, 2009).

GATA2 was initially identified as an activator of endothelin 1 expression in endothelial cells (Wilson *et al*, 1990). Targeted disruption of the *gata2* gene in mice resulted in embryonic lethality between embryonic days 9.5–11.5, due to defects in primitive haematopoiesis and haemogenesis (Tsai *et al*, 1994). In addition to the developmental stage, GATA2 is also recognized as an important regulator of endothelial selective gene expression, including platelet/endothelial cell adhesion molecule (PECAM)1, endothelial-nitric oxide synthase, von Willebrand factor (vWF), Down syndrome critical region-1, vascular endothelial cell adhesion molecule-1, KDR, and GATA2 itself (Jahroudi and Lynch, 1994; Zhang *et al*, 1995; Gumina *et al*, 1997; Kappel *et al*, 2000; Minami and Aird, 2001; Minami *et al*, 2001, 2009; Wozniak *et al*, 2007). Although these findings have suggested that GATA2 has a crucial role in the gene expression profile in vascular endothelial cells, the molecular mechanism by which GATA2 controls many endothelial specifically expressed genes remains largely unknown.

Recently, a hierarchical approach to investigating transcriptional regulation has shown to be the most informative, including consideration of transcription factors, histone modifications, and chromatin conformational changes proceeding gene expression. Modifications associated with trimethylated histone 3 lysine 4 (H3K4me3) in the promoter and monomethylated histone 3 lysine 4 (H3K4me1) in the enhancer, are tightly regulated with lineage specificity (Barski *et al*, 2007; Kim *et al*, 2010). More recently, it has been reported that cell-type specific gene activation and silencing can accompany

dynamic chromatin conformational changes, resulting in different accessibility of the transcriptional machinery (Fullwood *et al*, 2009; Visel *et al*, 2009).

Endothelial cells exhibit a wide range of phenotypic variability depending on the local environment throughout the vascular tree (Aird, 2007). Moreover, recent studies have demonstrated that endothelial to mesenchymal transition (EndMT) can occur in various pathological states in cancer and cardiac fibrosis (Zeisberg *et al*, 2007a, b). During the EndMT, endothelial cells are characterized by their loss of endothelial markers, such as a PECAM1 and gain of the mesenchymal markers, such as an α-smooth muscle (SM)-actin, resulting in the loss of cell–cell junctions and acquisition of invasive and migratory properties (Potenta *et al*, 2008).

In this report, we performed chromatin immunoprecipitation with deep sequencing (ChIP-seq) to determine genome-wide occupancy of GATA2 in endothelial cells and compared it with the respective gene expression profiles to understand GATA2 function on a genome-wide scale. Comparison of epigenetic markers and chromatin conformation between endothelial cells and non-endothelial cells, we identified a new GATA2 responsive endothelial-specific marker gene, *endomucin*. Moreover, abrogation of the GATA2 function in endothelium demonstrated not only reduction of the endothelial-specific marker, but also induction of the EndMT promoting gene expression. These results provide new insights into the cooperation between endothelial-expressed GATA2 binding and epigenetic modification, resulting in the determination of endothelial-cell-specific gene expression.

Results

GATA2 is uniformly expressed in the vascular endothelium

GATA factor has been considered as a key transcription factor regulated with cell-type specificity and differentiation. In endothelial cells, GATA2 has mainly been reported to be involved in the regulation of gene expression. We have previously reported both GATA3 mRNA and protein are predominantly expressed in large vessel endothelium, whereas, GATA2 mRNA is commonly expressed in endothelial cells derived from both microvascular- and larger-endothelium (Song *et al*, 2009). To identify whether GATA2 protein is indeed expressed in the vascular endothelium *in vivo*, we generated a monoclonal antibody against the antigen (residues 192–245) of human GATA2 (Figure 1A). The antibody specifically reacts with GATA2, but not GATA3, illustrating its selectivity (Figure 1B). Immunohistochemical staining revealed that GATA2 was specifically expressed in nuclei of dermal microvascular endothelium and artery endothelial cells of the human dermis (Figure 1C). Moreover, GATA2-derived signals were detected in the small vessels of the lung, heart, kidney, and tumour-derived microvascular endothelial cells (data not shown). In cultured endothelial cells, GATA2-positive staining was uniformly found in nuclei (Figure 1D).

Genome-wide and cell-type specific GATA2 binding was detected in microvascular endothelial cells

Next, we wished to determine how endothelial cell expressed GATA2 regulates the cell-type specific gene expression

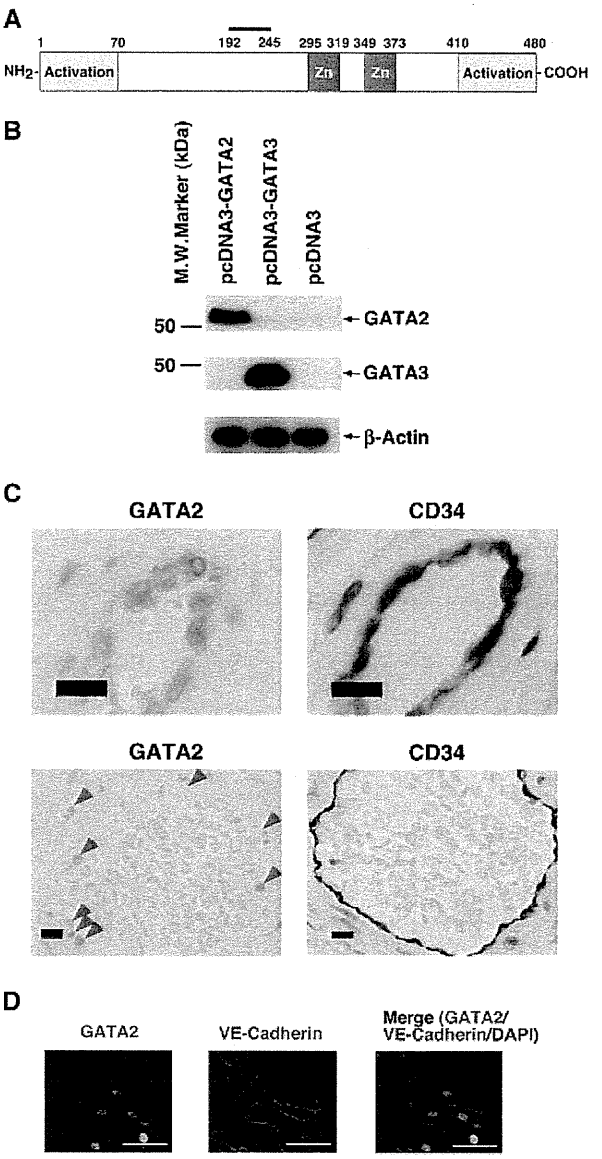


Figure 1 GATA2 expression in endothelium. (**A**) The schematic representation of human GATA2. Zn, zinc-finger domain. The bar indicates the region of the antigen for the monoclonal antibody. (**B**) Human GATA2 (pcDNA3-GATA2), GATA3 (pcDNA3-GATA3), or mock control (pcDNA3) were transfected in COS7 cells. Whole-cell lysates were prepared and blotted with an anti-GATA2, and anti-GATA3 antibody. Anti-β-actin antibody was used as a loading control. (**C**) Subcutaneous tissues were sectioned and subjected to immunohistochemical staining. GATA2 was shown with light brown in the nucleus (left). CD34 was with dark brown (right). Indicative of the microvascular and artery endothelium, respectively. Bar: 10 μm. The results are representative of five independent experiments. (**D**) Immunofluorescent staining of GATA2 (left) and VE-Cadherin (middle) in HMVEC. Merged images with DAPI are shown in right. Bar: 50 μm.

genome wide. Thus, at first, we performed duplicate ChIP with GATA2 antibody, and then the precipitated genome was sequenced comprehensively (duplicate ChIP-seqs). To test the accuracy of these ChIP assays with GATA2 antibody, we choose the endothelial cell specifically expressed vWF promoter region as a positive control; and non-expressed MyoD1 promoter region as a negative control (Supplementary Figure

SI). The GATA2-mediated immunoprecipitated DNA and non-immunoprecipitated whole genome control (input DNA) were used to prepare libraries for deep sequencing and analysed using massively parallel sequencing. The genome-wide GATA2-binding regions were calculated by two independent methods, QuEST (Valouev *et al*, 2008) and Sole-search (Blahnik *et al*, 2010) (see Supplementary data in detail). We found a total of 5737 and 5805 regions, which were identified as GATA2-enriched areas from the first and second ChIP-seqs, respectively (Supplementary Table SI). To reconstruct the GATA2-associated binding regions, we clustered the regions into four sections based on the distance from the transcription start site (TSS) in the respective genes. As shown in Figure 2A, up to 57% GATA2-binding regions were positioned around 10 kbp from each transcript. Among them, 9% were in the proximal promoter of the gene, 42% were located in the intron. The remaining 43% were located at the intergene. Collectively, these data suggest that the defined GATA2-binding regions derived from the comprehensive ChIP-seq data were not selectively located within the proximal promoter of each gene, rather, scattered widely in the whole genome. These findings were consistent with other transcription factors, oestrogen receptor, and FOXA1 as performed in tumour cells (Lupien *et al*, 2008).

Next, we identified the commonly recognized motif from the whole GATA2-mediated ChIP genome sequences. As shown in Figure 2B, the GATA2-recognized sequence (A/T)GATA(A/G) was indeed determined to be the highest enriched binding element, with *E*-value $1.3 \times e^{-1848}$. In addition to GATA2, Ets- and AP-1-recognized sequences were also found as the second and third enriched elements.

Recently, GATA2 ChIP-seq results were reported in the erythroid lineage cells, K562, using a similar strategy (Fujiwara *et al*, 2009). Interestingly, microvascular endothelial cells and K562 both commonly express GATA2, but the GATA2-regulated genes were not similar. Therefore, we compared GATA2 associations among the whole genome between HMVEC and K562 cells. In K562 cells, duplicate ChIP-seq with GATA2 resulted in the GATA2-recognized motif being selected as the highest enriched sequence. However, in contrast to HMVEC, the second and third most enriched motifs were Gfi-1b and TAL1, respectively (Figure 2B). As shown in Figure 2C, in total 896 and 334 genes were found to be GATA2 enriched in HMVEC and K562 cells, respectively. Among them, a total of 118 genes overlapped in both cells (Figure 2C), suggesting that distinct GATA2 associations with recognized elements in a cell-type specific manner could be observed.

To further investigate the correlation between GATA2 binding and resulting expression in HMVEC and K562 cells,

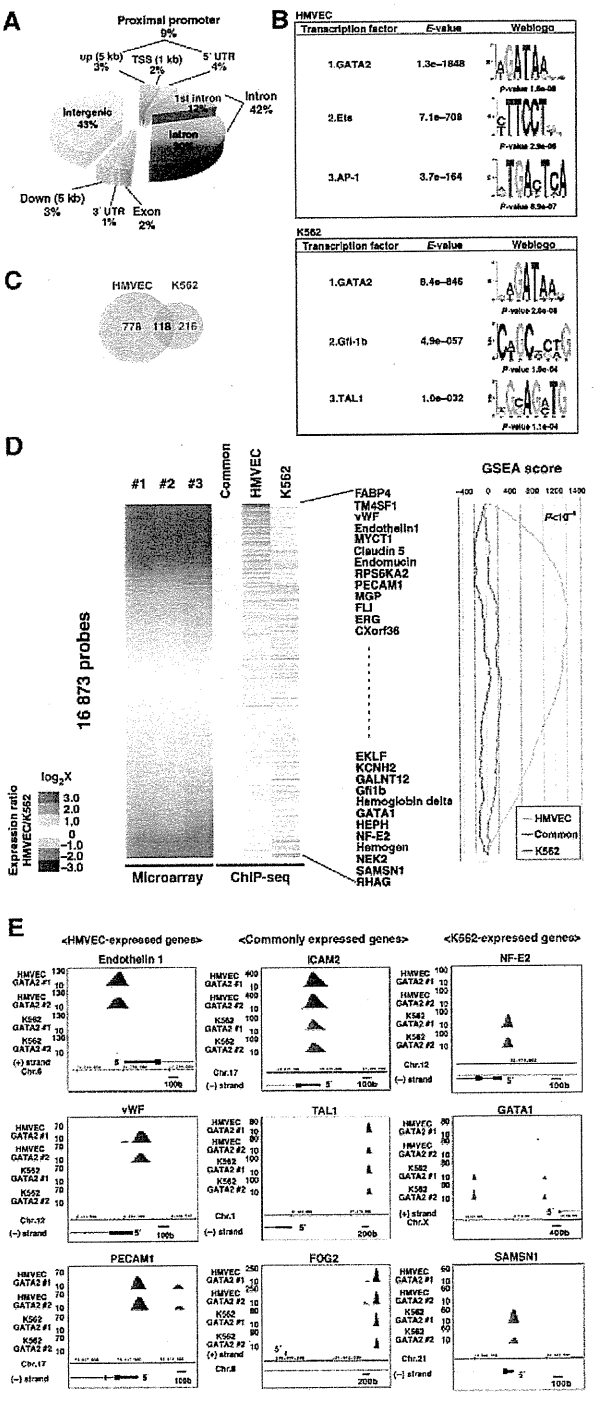


Figure 2 ChIP-seq results in HMVEC and K562. (**A**) Distribution of GATA2-binding regions in HMVEC. GATA2-binding regions (QuEST score ⩾30) were classified according to the distance from the transcription start sites of known genes. (**B**) Determination of the sequence recognized by GATA2 in HMVEC and K562. MEME method (Bailey and Elkan, 1994) was used for identification of enriched sequences and displayed with correlation of the size of character and the rate of enrichment. *E*-value and *P*-value mean the probability *de novo* enriched sequences obtained from ChIP-seq are matched to the displayed Weblogo, and known consensus motifs by chance, respectively. (**C**) Venn diagrams depicting the overlap of genes with GATA2 binding within the up 5 kbp, TSS 1 kbp, 5′UTR, and first intron, in between HMVEC and K562 cells. (**D**) Comparison of the GATA2 binding and the expression value in HMVEC and K562. The heat map from the expression ratio (HMVEC versus K562) was shown in left. Each gene obtaining the GATA2-binding sites from the ChIP-seq was shown with grey bar in middle. Common indicates the finding the GATA2 association into the gene commonly in HMVEC and K562. *HMVEC* and *K562* design the unique GATA2 association separately in each cell. GSEA analysis showed the correlation between the expression ratio (HMVEC versus K562) and the GATA2-binding gene profiles in right. (**E**) Each three representative ChIP-seq with GATA2 result derived from HMVEC (left), common (middle), or K562 (right) expressed genes. #1; first ChIP-seq, and #2; second ChIP-seq.

we performed comprehensive mRNA expression analysis by triplicate microarrays in both cell types. We selected the gene set probes that exhibited significant expression, as a >100 with average difference in at least either HMVEC or K562 cells (Supplementary Table SII). Subsequently, a total of 16 873 probes were sorted by the ratio of the expression levels in HMVEC versus K562 cells as illustrated with a heat map (Figure 2D, left column). To integrate the ChIP-seq data from HMVEC and K562, a grey bar was shown for each probe corresponding to the positive within the up 5 kbp, TSS 1 kbp, 5′UTR and first intron of the ChIP-seq results (Figure 2D, middle column). Interestingly, roughly one third of genes expressed were categorized as highly HMVEC-expressed genes (means red in Figure 2D), which were also relatively dense in ChIP-seq in HMVEC, but not in K562 cells. In contrast, one third of genes were highly expressed in K562 cells (green in Figure 2D), which were also dense in ChIP-seq in K562 cells, yet were sparse in HMVEC (Figure 2C, middle column). Moreover, based on a ranking list of the expression ratio, HMVEC versus K562, a highly significant ($P<10^{-6}$) GSEA score (Subramanian *et al*, 2005) was observed with GATA2 binding only in HMVEC, whereas a minus score was found with GATA2 binding specifically in K562 cells. Significant correlation was not observed within the common GATA2 bindings in both cells (Figure 2D, right column). Interestingly, the HMVEC highly expressed group included well-known endothelial-cell-specific genes, such as vWF, endothelin 1, PECAM1, while K562 cells highly expressed many important genes for erythroid cell differentiation, such as erythroid krüppel-like factor, GATA1, Gfi-1b, and NF-E2 (Figure 2D). The representative GATA2 binding at the proximal promoter derived from duplicate ChIP-seqs was shown in Figure 2E. Consistent with the expression, HMVEC specifically expressed endothelin 1, vWF, and PECAM1 showed GATA2 binding only in HMVEC (Figure 2E, left). In contrast, erythroid-specific NF-E2, GATA1, and SAM domain-containing protein (SAMSN1) revealed specific GATA2 binding in K562 cells alone (Figure 2E, right). Both stable GATA2 binding in HMVEC and K562 cells were found with commonly expressed genes, ICAM2; TAL1; and Friend of GATA (FOG) 2 (Figure 2E, middle). Taken together, these findings suggest that the presence of the GATA sequence alone in the genome cannot explain the association of GATA2. Rather, distinct cell-type specific regulation in endothelial cells and erythroid cells should exist for regulating GATA2 association.

GATA2 regulates endothelial-cell-specific gene expression and cell specificity

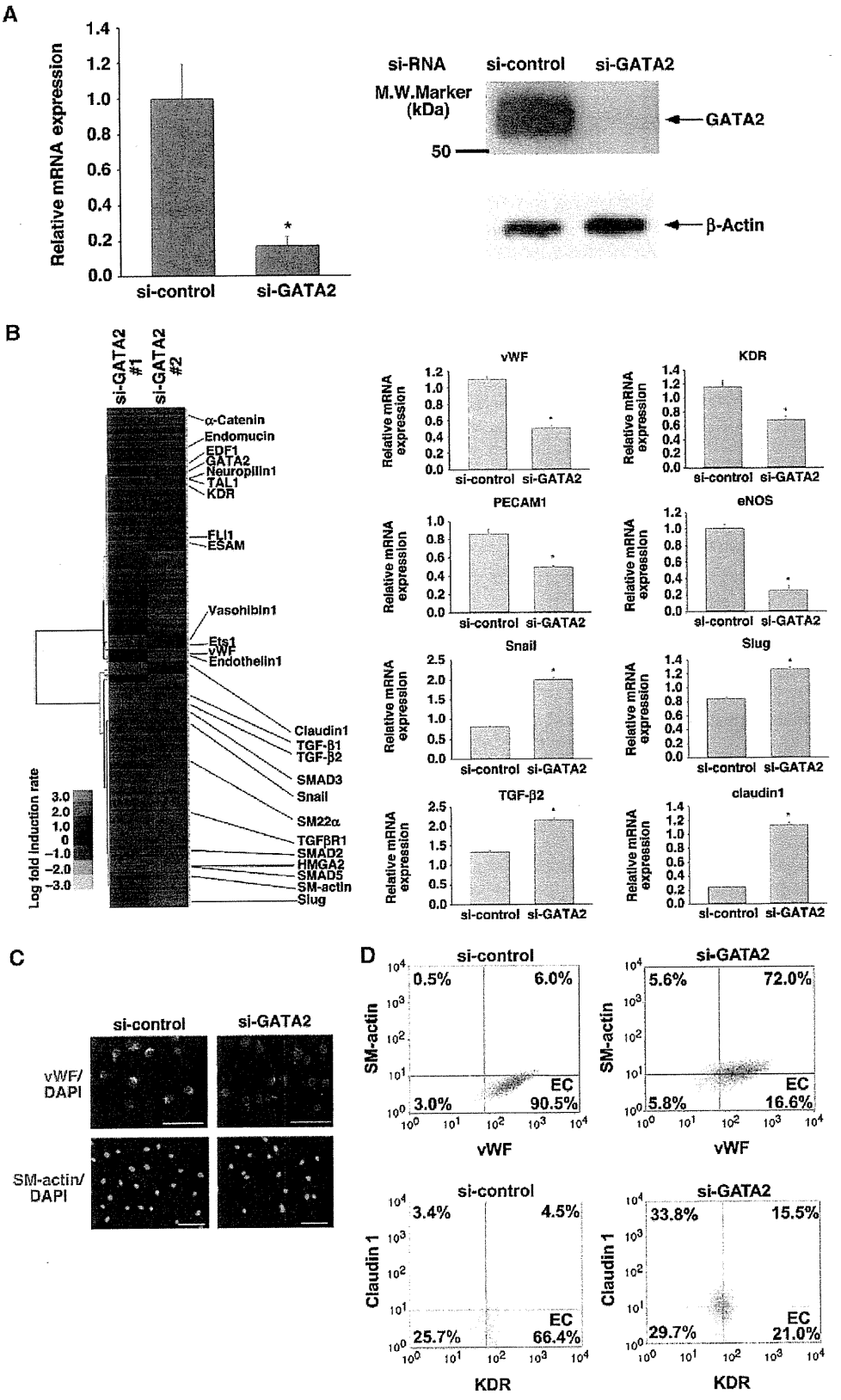
Next, to survey those genes regulated by GATA2 in microvascular endothelial cells, we used siRNA to knockdown GATA2 expression in HMVEC. As shown in Figure 3A, transfection of HMVEC with siRNA to GATA2 (si-GATA2), but not control (si-control), resulted in a 92% reduction in GATA2 mRNA. Moreover, virtually all GATA2 protein expression was blocked by the si-GATA2 treatment (Figure 3A). To comprehensively identify possible GATA2-regulated genes, we employed duplicate microarrays performed in parallel with HMVEC transfected with either si-control or si-GATA2. To identify genes regulated by GATA2, the selection criteria was set at >0.75 log fold upregulated or downregulated genes in the presence of si-GATA2, compared with si-control. A total of 788 (upregulated) and 818 (downregulated) genes were identified and clustered (Figure 3B, whole gene list shown in Supplementary Table SIII). Among them, GATA2 knockdown resulted in the reduction of the endothelial-specific marker, endothelin 1, vWF, KDR, endothelial-specific adhesion molecule, and endothelial differentiation factor-1. In contrast, GATA2 reduction led to the upregulation of the non-endothelial markers, SM-actin, SM22α, and claudin 1. Surprisingly, epithelial mesenchymal transition (EMT)-directed transcription factor, Snail, Slug, and HMGA2, were upregulated (Figure 3B). EndMT-directed TGF-β and the mediator SMAD were also increased (Figure 3B). The representative microarray-based findings were subsequently validated by quantitative real-time PCR (Figure 3B, right graphs). Significant GATA2 binding to the proximal promoter was also detected in TGF-β2 and Slug from duplicate ChIP-seqs (Supplementary Figure SII). To test whether GATA2 reduction leads to the loss of the endothelial specificity, we examined immunofluorescent staining with anti-vWF and SM-actin antibodies as endothelial and mesenchymal markers, respectively. As shown in Figure 3C, si-GATA2-treated cells exhibited a reduction of vWF-positive staining and a concomitant induction of SM-actin-positive staining. To quantify this ratio, we performed FACS analysis. Control siRNA-treated HMVEC resulted in 96.5% of cells staining gated-positive for vWF and 6.5% positive for SM-actin expression. In contrast, si-GATA2-treated HMVEC revealed a decrease in gated-positive cells to 88.6%, with a dramatic increase in cells gated-positive for SM-actin (77.6%) (Figure 3D). Moreover, GATA2 knockdown caused a roughly 50% reduction in the endothelial marker KDR, and up to a 63.5% induction of the mesenchymal marker; claudin 1 (Figure 3D). Collectively, these findings suggested that GATA2 is not only a key regulator of the expression of several endothelial-cell-specific genes, but also necessary for the maintenance of endothelial specificity.

GATA2 regulates endothelial-cell-specific endomucin gene expression

We decided to further focus our analysis on *endomucin* gene expression, since it showed the greatest reduction under si-GATA2 treatment. Human *endomucin* gene structure has been studied by two independent groups (Kinoshita *et al*, 2001; Samulowitz *et al*, 2002). ChIP-seq data suggested that GATA2 was recruited to the proximal promoter in the *endomucin* gene. The peak extended from 500 bp upstream to 200 bp downstream of the TSS (Figure 4A). This segment includes three consensus GATA-binding motifs (Figure 4C). To confirm whether these regions were in an active chromatin state, we performed ChIP and quantitative PCR (ChIP–qPCR) analysis using antibodies against acetylated histone 3 (H3Ac), acetylated histone 4 (H4Ac), trimethylated lysine 4 of histone 3 (H3K4me3), and monomethylated lysine 4 of histone 3 (H3K4me1), as well as GATA2. In addition, we employed ChIP–qPCR analysis with antibodies against serine 5 phosphorylated active polymerase II (polII), and the transcriptional cofactor, p300. As a negative control, we selected the promoter region from zinc-finger protein (znf) 649, which is known to be completely silent in endothelial cells. As shown in Figure 4B, the *endomucin* promoter indicated a profound enrichment with active PolII, H4Ac, H3Ac, H3K4me3, H3K4me1, p300, and GATA2 itself. In contrast, control IgG failed to enrich the *endomucin* promoter in HMVEC.

To determine whether GATA2 is required for *endomucin* gene expression, we employed luciferase reporter analysis.

We isolated a fragment from the human *endomucin* promoter from −1968 to +108 relative to the TSS, and subcloned it into the luciferase vector pGL3 (Figure 4C). The plasmid, endomucin-luc, was transiently transfected into primary cultured human umbilical vein endothelial cells (HUVEC), HMVEC, human skin fibroblast, and K562 cells. As shown



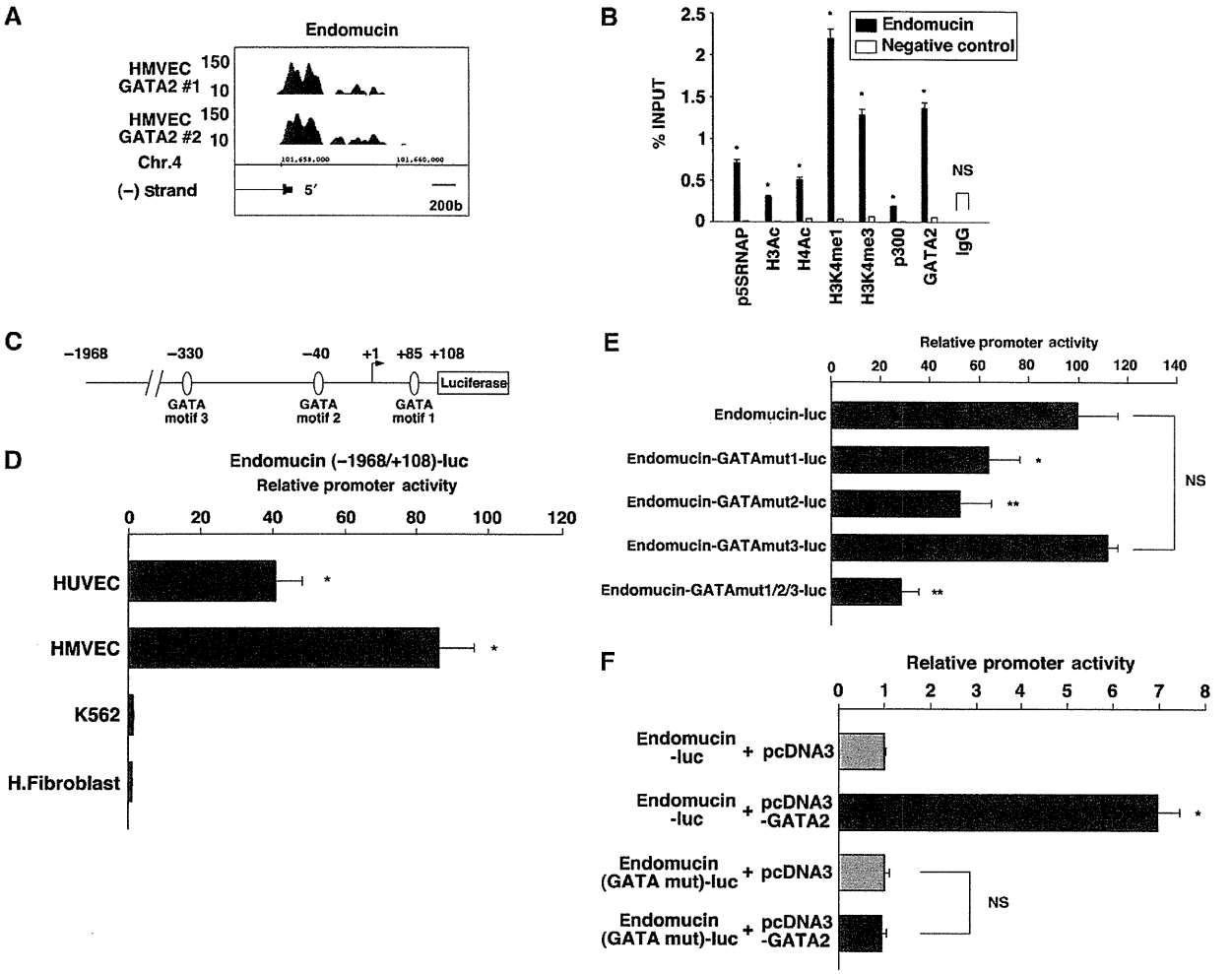


Figure 4 GATA2-mediated *endomucin* gene expression. (**A**) Signals of duplicate ChIP-seqs data. Displayed enrich scores were calculated by QuEST, and visualized in the IGB software (Affymetrix). Significant GATA2-binding signals are shown in grey. - denotes Refseq genes located on the minus strand. (**B**) Formalin-fixed chromatin was immunoprecipitated with antibodies as shown. Immunoprecipitated DNA was quantified by the qPCR using the specific primers for *endomucin* and the znf 649, as a negative control. Fold enrichment were compared with non-immunoprecipitated chromatin, %INPUT. Data are expressed as mean and standard deviations obtained from three independent experiments, $n=3$. $^*P<0.001$ compared with negative control, znf 649. NS, non-significant. (**C**) Schematic representation of the endomucin (−1968/+108)-luc and three consensus GATA motifs. (**D**) Endomucin (−1968/+108)-luc and TK-luc were transiently transfected into HUVEC, HMVEC, K562, or fibroblasts, and assayed for the luciferase activities. Each cell was cotransfected with pRL-CMV to normalize the transfection efficiency. The results show the mean and the standard deviations from the ratio of the endomucin-luc versus TK-luc, with three independent experiments, $n=3$. $^*P<0.001$ compared with the activity from fibroblasts. (**E**) HMVEC were transiently transfected with either wild-type or a GATA-point mutated-luc. The results show the mean and the standard deviations of luciferase light units relative to the wild-type endomucin-luc, obtained in triplicate, $n=4$. $^*P<0.02$ and $^{**}P<0.001$, compared with the activity from the wild type. NS, non-significant. (**F**) COS7 cells were transiently cotransfected with the combination both endomucin-luc or GATA-point mutant, and expression plasmids for GATA2 (pcDNA3-GATA2) or mock control (pcDNA3). The results show the mean and the standard deviations of luciferase light units relative to the pcDNA3 alone, obtained in triplicate from three independent experiments, $n=3$. $^*P<0.01$ compared with the activity from the pcDNA3 alone. NS, non-significant.

Figure 3 Essential role for GATA2 in microvascular endothelial cells. (**A**) (Left) Quantitative real-time PCR analysis of GATA2 mRNA in HMVEC transfected with si-control or si-GATA2. Data are expressed as the mean and the standard deviations of expression levels relative to cyclophilin A obtained in triplicate from five independent experiments, $n=3$. $^*P<0.001$ compared with si-control. (Right) Western blot analysis of HMVEC transfected with si-control or si-GATA2. The membrane was immunoblotted with anti-GATA2 antibody. Anti-β-actin antibody was used as a loading control. The results are representative from three independent experiments. (**B**) (Left) Heat map of the HMVEC transfected with si-control or si-GATA2 derived from microarrays. Colour intensity: green—lower and red—higher, shown relative to median (black). (Right graphs) Quantitative real-time PCR analysis of the representative genes. Data are expressed as the mean and the standard deviations of expression levels relative to cyclophilin A obtained from three independent experiments, $n=3$. $^*P<0.05$ compared with si-control. (**C**) Immunofluorescent stainings of HMVEC treated with si-control or si-GATA2, by using the anti-vWF (green) or SM-actin (red) antibodies. Nuclei was stained with DAPI (blue). The merged images show the representative from four independent experiments. Bar: 100 μm. (**D**) FACS analysis of HMVEC treated with si-control or si-GATA2. The results show the representative from three independent experiments. EC, window from the endothelial fraction.

in Figure 4D, the *endomucin* promoter was greatly upregulated in HUVEC and HMVEC, but not in K562 nor fibroblasts. To determine whether the three GATA elements at +85 (motif 1), −40 (motif 2), and −330 (motif 3) contribute equally to *endomucin* gene expression in HMVEC, we generated point mutations of the GATA motif to TTTA in each region, and systematically measured luciferase activities. As shown in Figure 4E, compared with the wild-type promoter, mutation of motif 3 failed to reduce promoter activity. However, mutation of motif 1 and 2 resulted in a 37 and 48% reduction of promoter activity, respectively. Mutation of all three sites in combination further reduced (72%) promoter activity (Figure 4E), with mutation in motif 1 and 2 resulted in a similar reduction (~70%) of promoter activity (data not shown). These results suggest that the two proximal GATA motifs are necessary for *endomucin* promoter activity in endothelial cells.

Subsequently, to test whether GATA2 transactivates the *endomucin* gene expression, COS7 cells were cotransfected with either an endomucin-luc or a GATA mutant-luc, along with a GATA2 expression vector or a mock control vector, pcDNA3. As shown in Figure 4F, GATA2 overexpression resulted in 6.9-fold increase in *endomucin* promoter activity, while GATA2 overexpression failed to transactivate the GATA-mutated *endomucin* promoter. In large vessel endothelium, GATA2, 3, and 6 were expressed (Song *et al*, 2009). Thus, we next verified whether *endomucin* expression was commonly regulated by the GATA family, or selectively regulated by GATA2 alone. In cotransfection assays with GATA2, 3, or 6, comparable GATA-mediated transactivation activities were detected with core-GATA element-containing thymidine kinase (TK)-luc (Supplementary Figure SIIIA). Importantly, under these conditions, GATA2 potentiated the highest transactivation activity in *endomucin*-luc (Supplementary Figure SIIIB). Compared with GATA2, GATA3 and 6 indicated weaker, but statistically significant upregulation of *endomucin* promoter activity (Supplementary Figure SIIIB). Moreover, knockdown experiments in HUVEC revealed that *endomucin* was reduced ~90% via si-GATA2 treatment. Knockdown of GATA3 or 6 also showed more than half of the reduction. Combined siRNAs against GATA2, 3, and 6 resulted in a far more profound reduction of *endomucin* expression (Supplementary Figure SIIIC). Collectively, these findings suggest that among the GATA family, GATA2 is the most prominent factor for *endomucin* gene expression. The *endomucin* promoter specifically functions in endothelial cells, with the +85 and −40 GATA elements important to confer promoter activity.

Identification of the GATA-mediated distal enhancer regulating endomucin expression

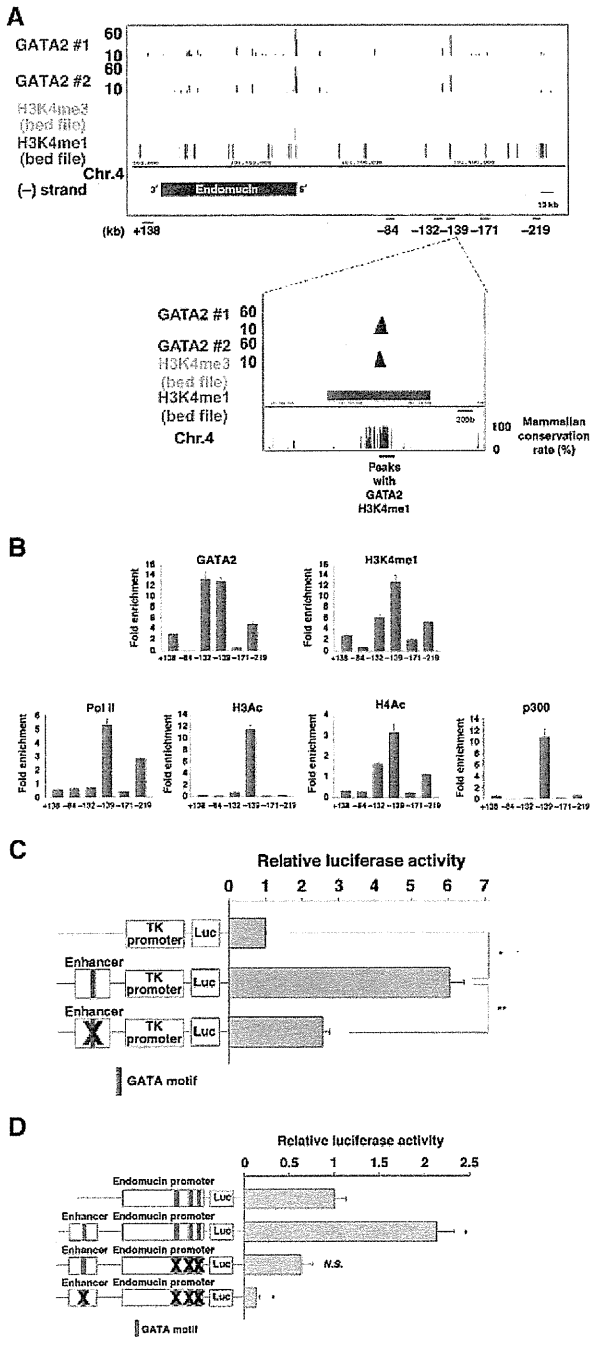
From the above GATA2-mediated ChIP-seq findings, defined GATA2-binding regions were not concentrated in the proximal promoter regions, as many were found in the introns and intergenic regions. It has previously been reported that β-globin gene expression is regulated through a distal enhancer region bound by GATA factor (Kim *et al*, 2007). Thus, we wished to evaluate whether similar GATA2-mediated distal enhancers also existed in the genome of endothelial cells, and whether they regulated *endomucin* gene expression in an endothelial-cell-specific manner. To that end, we first evaluated ChIP-seq information with H3K4me1 and H3K4me3 bindings in HMVEC. Regions that showed specifically higher H3K4me1 signals compared with H3K4me3 were considered enhancer regions (Robertson *et al*, 2008). Thus, we searched the H3K4me1 signals surrounding the *endomucin* locus. As shown in Figure 5A, H3K4me3 signals (means the active core promoter definition) were detected only at the TSS. More importantly, H3K4me1-specific signals were detected at several points in the locus. In all, −219, −171, −139, −132, and +138 kbp regions (Figure 5A, bars) were chosen and validated with ChIP–qPCR analysis with H3K4me1 (Figure 5B). Compared with the negative control region at −84 kbp, H3K4me1 showed statistically significant binding to the positive region. Among them, −139 kbp region showed the highest and −171 kbp showed the lowest binding (Figure 5B). Subsequently, we searched the GATA2 binding within the locus. From the duplicate ChIP-seqs, GATA2 associations were commonly detected at −139 kbp and at a weaker level at −132 kbp (Figure 5A). ChIP–qPCR validation resulted in significant bindings of GATA2 to −132 and −139 kbp regions (Figure 5B). To verify which region from the −132 or −139 kbp is epigenetically designed as an enhancer, we performed ChIP–qPCR analysis with antibodies against active PolII, H3Ac, H4Ac, and p300. As shown in Figure 5B, compared with the negative control region at −84 kbp, only the −139 kbp region showed strong merged binding by p300, H3Ac, H4Ac, and active PolII. Moreover, the comparative genome analysis among mammalian species resulted in sharp homologous signals at −139 kbp (Figure 5A, bottom). Collectively, these findings suggest that the −139 kbp region is the main GATA2-binding enhancer within the *endomucin* locus.

To further investigate whether the −139 kbp region enhanced promoter activity, we employed TK promoter-based luciferase assays. The core area (822 bp) at the −139 kbp region contains the consensus GATA element (Supplementary Figure SIV). Thus, we cloned and ligated the core 822 bp in upstream of the TK promoter. As shown in Figure 5C, addition of the GATA enhancer resulted in more than a six-fold increase in TK promoter activity over the TK basal promoter alone. In contrast, a GATA-mutated enhancer led to a >80% reduction of enhancer activity. Subsequently, to test whether the core area at the −139 kbp region works as an enhancer connected with the intact proximal *endomucin* promoter, we generated constructs and performed reporter assays. The *endomucin* promoter (2.2 kbp) contained high luciferase activity. Ligated with the core enhancer resulted in a further 2.2-fold increase in promoter activity (Figure 5D). While a GATA-mutated enhancer failed to transactivate the *endomucin* promoter. Moreover, an all GATA motifs mutated construct revealed a profound reduction in the basal *endomucin* promoter activity (Figure 5D). Taken together, these findings suggest that the −139 kbp region is in an active chromatin state and the core-GATA-binding element works as an enhancer in endothelial cells.

GATA2 regulates the chromatin structure of the endomucin locus in endothelial cells

GATA factor binds and activates not only a proximal promoter but also a distal enhancer region. It was also reported that GATA factor binding regulates the chromatin conformation (Jing *et al*, 2008). Therefore, to evaluate whether GATA binding at the proximal promoter and distal enhancer of the

endomucin locus also induce chromatin remodelling, we performed chromatin conformation capture (3C) assays (Dekker *et al*, 2002). *Endomucin* was specifically expressed in endothelial cells but not K562 cells (Figure 6A). To compare the chromatin structure between HMVEC and K562 cells by 3C assay, we picked nine different points within the *endomucin* locus (Figure 6B). As shown in Figure 6C, 3C-mediated PCR amplification was detected not only in the TSS region, but also at the GATA associated −132 and −139 kbp regions (see Figure 5A and B) of the *endomucin* locus in HMVEC. In contrast, the chromatin derived from K562 cells exhibited 3C products only detected in the TSS region (Figure 6C).



To verify whether GATA2 association directly mediates chromatin conformation at the *endomucin* locus, we treated cells with siRNA against GATA2 and again performed 3C assays. As shown in Figure 6D, si-GATA2 treatment led to more than a 90% reduction in GATA2 association with GATA sites in the *endomucin* proximal promoter (TSS) and distal enhancer (−139 kbp). Importantly, GATA2 knockdown reduced chromatin-loop formation at −139 kbp and the TSS of the *endomucin* locus (Figure 6E).

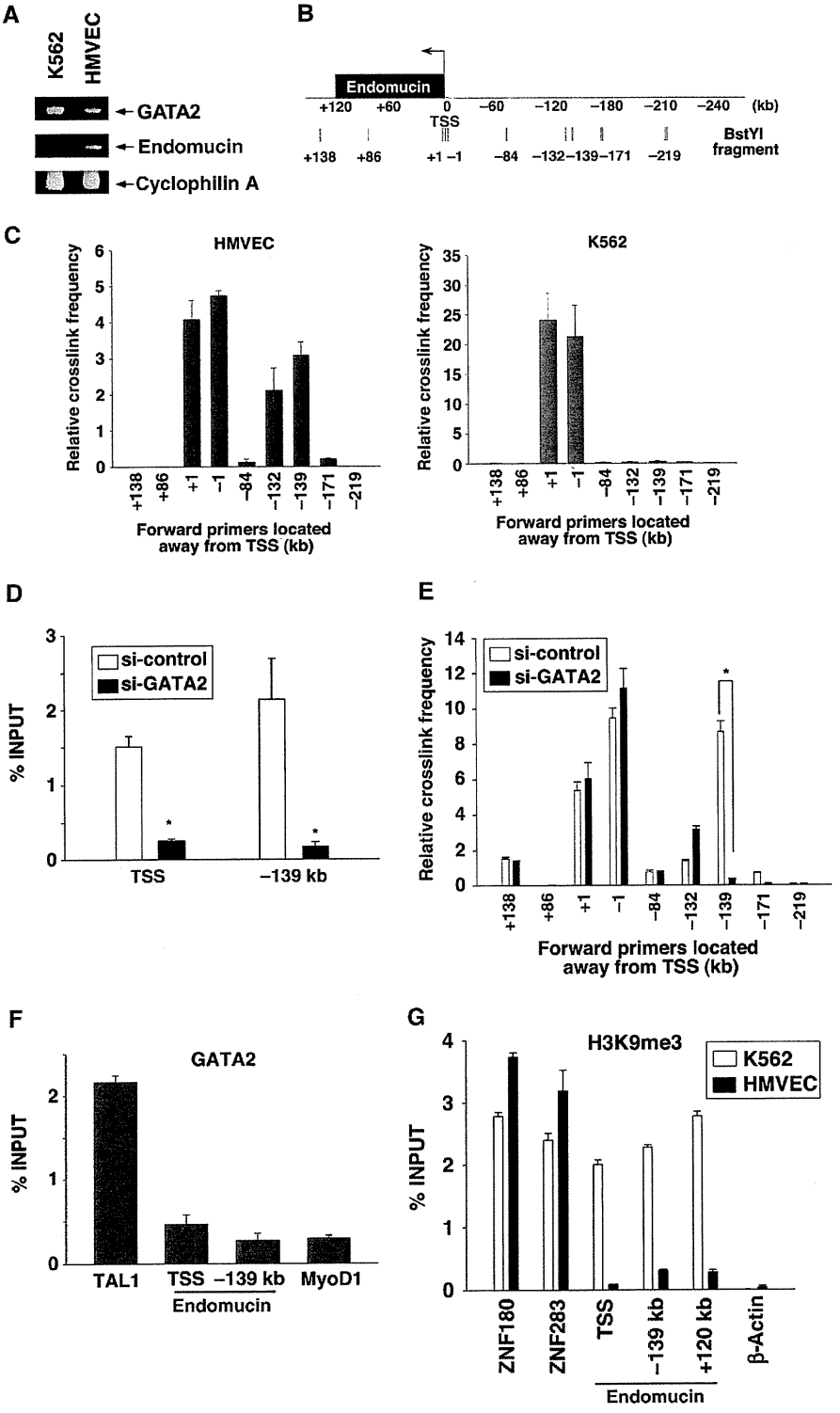
Next, we investigated why the *endomucin* gene was silent in K562 cells. GATA2 was highly expressed in both K562 and HMVEC cells (Figure 6A). However, GATA2-mediated ChIP-seq in K562 cells revealed non-GATA2 binding to the *endomucin* locus of the TSS and −139 kbp. The ratio was comparable to the non-expressed MyoD1 gene promoter (Figure 6F). In contrast, significant GATA2 binding was found in the well-expressed TAL1 gene promoter in K562 cells (Figure 6F). Epigenetically, silent gene loci are often determined by the binding of histon 3 lysine 9 trimethylation (H3K9me3) (Kouzarides, 2007). As shown in Figure 6G, high H3K9me3 enrichments were detected in the *endomucin* locus in K562 cells but not in HMVEC. Moreover, significant H3K9me3 enrichments were found in non-expressive genes, znf 180 and 283, in both K562 and HMVEC, while the ubiquitously expressed β-actin gene indicated neglectable H3K9me3 signals in both cells. Compared with HMVEC, the entire *endomucin* locus showed no active histone marks in K562 cells (Raney *et al*, 2011). Taken together, these findings suggest that GATA2 expressed in endothelial cells binds to the H3K9me3 free regulatory element on several lineage specific genes, including *endomucin*, and mediates endothelial-cell-specific chromatin conformation and determines endothelial cell specificity.

GATA2-regulated endomucin is essential for cell growth, migration, and tube formation in microvascular endothelial cells

Having established a critical role for GATA2 in *endomucin* gene expression in microvascular endothelial cells, we next

Figure 5 Identification of the GATA2-binding enhancer for the *endomucin* locus. (**A**) In all, 390 kbp UCSC hg18 genome browser view around the *endomucin* gene. The gene is transcribed from right to left. In HMVEC, GATA2-binding regions (blue) and H3K4me1 (red), H3K4me3 (pink) enriched regions were shown. Each black bar below the frame indicates the ChIP–qPCR amplification region, and the number show the distance (kbp) from TSS. At lower panel, magnified image around the −139 kbp region in which GATA2 signals and H3K4me1-positive signals are coexisted. Green bar shows the conservation rate (%) between mammalian species. (**B**) ChIP analysis with specific antibodies. Immunoprecipitated DNA were amplified with specific primers described in (**A**) and Supplementary Table SIV. Fold enrichment were calculated with the ratio from the znf 649 signals as a negative control. Data are expressed as the mean and the standard deviations obtained from three independent experiments, $n=3$. (**C**) HMVEC were transfected with TK-luc, GATA enhancer-TK-luc, or GATA enhancer (GATA mut)-TK-luc. HMVEC were cotransfected with pRL-CMV to normalize for the transfection efficiency. The results show the mean and the standard deviations of luciferase light units obtained in triplicate from three independent experiments, $n=3$. $^{*}P<0.001$ compared with TK-luc, $^{**}P<0.01$ compared with GATA enhancer-TK-luc. (**D**) HMVEC were transfected with *endomucin* (−1968/+108) promoter containing either −139 kbp core enhancer or GATA-mutated enhancer, $n=3$. $^{*}P<0.01$ compared with endomucin (−1968/+108)-luc. NS, non-significant.

wished to evaluate the functional relevance of *endomucin* in endothelial cells. To that end, we used siRNA to knockdown expression of *endomucin*. Transfection of HMVEC with two independent siRNAs against *endomucin* (si-endomucin oligo 1 and 2) resulted in >90% inhibition of the mRNA expression (Supplementary Figure SV). Of these, we first analysed the cell growth level. Consistent with previous findings (Mammoto *et al*, 2009), GATA2 knockdown impaired growth factor-mediated cell growth. Importantly, *endomucin* knockdown also attenuated the growth rate to 5.6 and 25.7% by siRNA oligo 1 and 2, respectively (Figure 7A). There was no evidence of increased apoptosis in siRNA-treated HMVEC



(data not shown). Next, we performed cell migration assays using a modified Boyden chamber. Fluorescently labelled HMVEC transfected with control siRNA or siRNA against GATA2 or *endomucin* were plated in the fibronectin-coated upper chamber. The addition of endothelial culture media with growth factors resulted in significant increases in migration of control siRNA-treated cells (Figure 7B). In contrast, GATA2 knockdown abolished, and *endomucin* knockdown impaired by 80.5 and 88.9% (oligo 1 and 2, respectively), the growth factor-mediated cell migration (Figure 7B). Coordinate function from endothelial cell growth and migration are important for angiogenesis activities. Thus, we finally performed tube formation assays of HMVEC. HMVEC were fluorescently labelled and incubated in a collagen gel. As shown in Figure 7C, control siRNA-treated HMVEC promoted growth and migration towards each cell, leading to the formation of capillary-like cords or tubes. Quantitation revealed a 68.4 and 81.6% decrease in tube

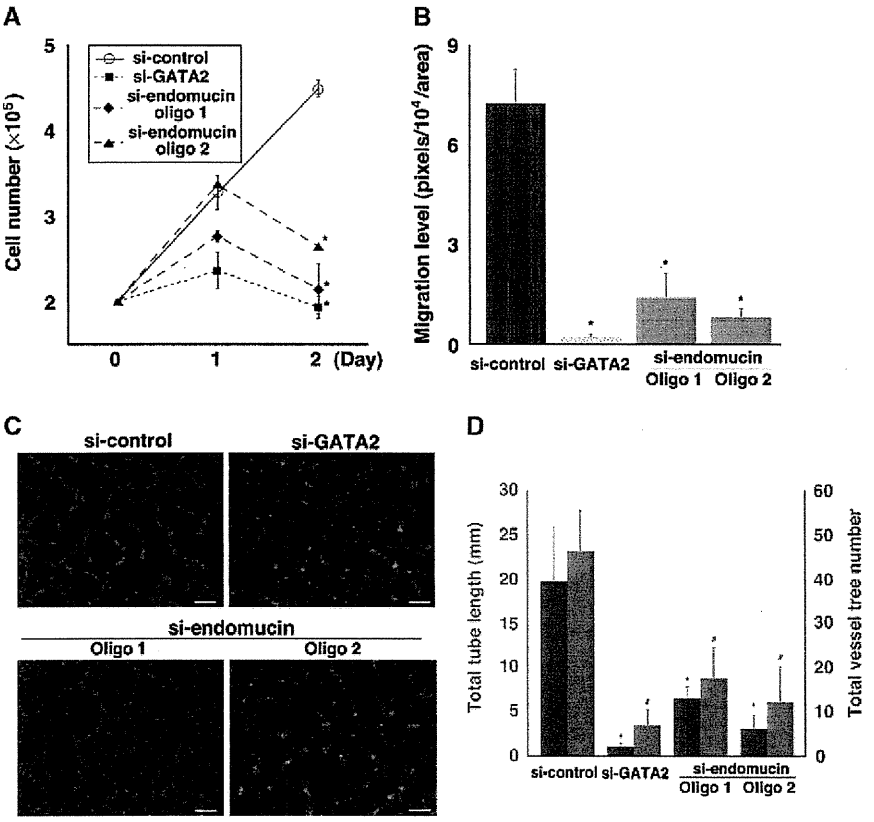


Figure 7 Knockdown of *endomucin* impair endothelial cell growth, migration, and matrix tube formation. (**A**) HMVEC were treated with siRNA against control, GATA2, or *endomucin* (oligo 1 and 2). After 6 h, HMVEC were re-seeded the concentration at 2×10^5. After 1 and 2 days, each cell number was counted. Data are shown with the mean and the standard deviation from three independent experiments, $n=4$. $^*P<0.001$ compared with si-control. (**B**) Migrated cells were quantified by cell image analyser in four independent experiments, $n=3$. $^*P<0.001$ compared with si-control. (**C**) Tube formation assays of si-control, si-GATA2, or si-endomucin-treated HMVEC plated on collagen gel. Capillary-like morphology was observed under the fluorescent microscopy. Bar: 100 μm. (**D**) Quantification of the total tube length (red) and total vessel tree number (blue) was calculated using the cell image analyser from three arbitral optical images per two independent experiments. Data are expressed as mean and standard deviation from six independent results, $n=3$. * and $^{\#}P<0.01$ compared si-control in tube length and vessel tree number, respectively.

Figure 6 Endothelial-cell-specific epigenetical regulation in the *endomucin* locus. (**A**) RT–PCR was carried out with specific primers for GATA2, *endomucin*, or cyclophilin A by using the total RNA from HMVEC and K562 cells. The data show the representative from three independent experiments. (**B**) Long range interactions between the *endomucin* promoter and enhancer measured with 3C-qPCR. The organization of the *endomucin* locus is displayed. A grey bar represents each *Bst*YI fragment. The grey bar indicates the *bait* region of the *endomucin* promoter. The numbers depict the distance (kbp) from the TSS. (**C**) 3C assay in HMVEC (left) and K562 (right). The *x* axis in each graph represents the location of the primers relative to the TSS. The *y* axis represents the cross-link frequency relative to it from the BAC DNA encompassing the *endomucin* region. Data are shown as the mean and the standard error from data in triplicate from two independent 3C experiments. (**D**) ChIP analysis with GATA2 antibody. Immunoprecipitated DNA were quantified with primer pairs described in Supplementary Table SIV. Fold enrichment was calculated with the INPUT. Data are expressed as the mean and the standard deviations obtained from three independent experiments, $n=3$. (**E**) HMVEC were transfected with si-control or si-GATA2. After 48 h, 3C assays were employed. Data are shown as the mean and the standard error from data in triplicate from two independent 3C experiments. (**F**) ChIP analysis with GATA2 antibody in K562. Primer pairs in TAL1 and MyoD1 promoters were used as a positive and a negative control, respectively. Data are expressed as the mean and the standard deviations obtained from triplicate samples, $n=3$. (**G**) ChIP analysis with H3K9me3 antibody in HMVEC and K562. Primer pairs in znf 180, 283, and β-actin were used as a positive and a negative control, respectively. Data are expressed as the mean and the standard deviations obtained from triplicate samples, $n=3$.

length in control versus *endomucin* siRNA (oligo 1 and 2, respectively)-treated cells (Figure 7D, red). Similarly, total vessel tree number was greatly attenuated by siRNAs against GATA2 or *endomucin* (Figure 7D, blue). Taken together, these findings suggest that *endomucin* is critical for cell growth, migration, and angiogenesis activities. Cell-specific expression regulating transcription factor; GATA2, and the downstream key target; *endomucin*, are indispensable for endothelial cell maintenance and physiological function.

Discussion

Recent technological advancements now allow the defining of transcription factor binding regions and unveiling of the histone code on a genome-wide scale. In this report, we identified GATA2 binding in human primary cultured endothelial cells, and then comprehensively mapped and compared these signals with epigenetic histone marks. Duplicate ChIP-seqs with GATA2 demonstrated that only a small number of the GATA2-binding elements were located within the proximal promoter, the majority of which were consistent with a previous promoter study of GATA2 in endothelial cells. Most GATA2 associations were defined in the intergenic or distal regions, far from the core promoter of the gene (see Figure 2A). This GATA2-binding pattern is not unique, but consistent with other reports of ChIP-seq analysis on the whole genome (Bieda *et al*, 2006; Chong *et al*, 2010; Schmidt *et al*, 2010). Such a finding suggests that not only transcription factor binding to the core promoter, but also coordinated transcriptional regulations involved in distal enhancer, silencer, and insulator, might occur in cells. For example, the *gata2* gene itself, GATA2 was reported to bind to the 9.5-kbp downstream of the mouse *gata2*, which enhances endothelial-cell-specific GATA2 gene expression (Khandekar *et al*, 2007). In another group, GATA2 has been shown to bind to the 77-kbp upstream of the mouse *gata2* (Grass *et al*, 2006). In our ChIP-seq data with HMVEC, GATA2 associations were observed at the +9.8 and −80.6 kbp of the human *gata2* gene (Supplementary Figure SVI), which exactly corresponds to the +9.5 and −77 kbp of the mouse *gata2* gene, suggesting that sparse GATA2 binding at the genome-wide level could be functional. Endothelial-cell-specific human *gata2* expression might be regulated with GATA2 binding itself.

Our microarray analysis with siRNA against GATA2 revealed that GATA2 is indispensable for many endothelial-cell-specific genes consistent with previous reports for vWF, KDR, and endothelin 1 (Wilson *et al*, 1990; Minami *et al*, 2001; Liu *et al*, 2009). Unexpectedly, GATA2 knockdown also indicated a role in the induction of cell conversion markers; Snail, Slug, HMGA2, and TGF-β2. EMT and EndMT were observed in pathological stages of cancer metastasis and cardiac fibrosis (Zeisberg *et al*, 2007a, b). Moreover, we observed that GATA2 and 3 expression significantly declined in endothelium under inflammatory and tumour-invaded conditions (Song *et al*, 2009 and data not shown). While interesting, our cell staining studies showed that abrogation of GATA2 alone resulted in only a limited conversion. When lacking GATA2, the mesenchymal markers; SM-actin and claudin 1 were greatly induced, in turn, endothelial marker; KDR was markedly reduced. However, vWF was weakly reduced. A possible reason for only a partial reduction of vWF may be due to protein stability; another is based on the compensation from other endothelial-cell-specific transcription factors. Indeed, vWF is stored in the Weibel-Palade body (Ewenstein *et al*, 1987), which might affect the protein stability. Alternatively, we have recently shown the vWF gene is specifically transactivated by not only GATA but also Ets family member, ERG in endothelium *in vivo* (Liu *et al*, 2009, 2011). Correlating these findings, the combination of transcription factors, FOXC2 and Ets family; Etv2, was shown to have pivotal roles in endothelial specificity and differentiation in zebrafish (De Val *et al*, 2008). Thus, loss of a combination of GATA2 and another transcription factor might result in complete EndMT even in fully differentiated endothelial cells. Collectively, these findings strongly support that GATA2 is a key transcription factor for the maintenance of endothelial cells.

In investigating GATA2-mediated regulation, we chose to focus on the *endomucin* gene, as it exhibited the largest reduction in expression upon GATA2 knockdown. *Endomucin* has been reported to be an endothelial-cell-specific mucin involved in cell–cell or cell–extracellular matrix interactions (Ueno *et al*, 2001). Angiogenic stimulus increases *endomucin* gene expression in HUVEC (Liu *et al*, 2001). Moreover, here we show that knockdown of the *endomucin* by specific siRNA treatment resulted in the suppression of growth, migration, and matrix tube formation in HMVEC (see Figure 7). Recently, it has been shown that reduction in GATA2 levels leads to the abrogation of tube formation (Mammoto *et al*, 2009). Therefore, it appears that the GATA2-endomucin axis has an important role in vascular stability and angiogenesis. Further study with GATA2-ChIP-seq under detailed various angiogenic and inflammatory stimulus, or the generation of *endomucin* null mutations could serve to uncover the detailed mechanisms.

Duplicate ChIP-seqs with GATA2 in HMVEC indeed enriched for the sequence (A/T)GATA(A/G), a well-recognized consensus binding sequence. Such an element is never randomly enriched, since the associated *E*-value is $1.3 \times e^{-1848}$ (see Figure 2B). In our data, GATA2 was commonly expressed within HMVEC and K562 cells, but different GATA2-binding patterns indicated a cell-type specific manner. Moreover, GATA2 occupancy was positioned with H3K4me1, p300, and active PolII-positive enhancer region in the *endomucin* locus in endothelial cells. Therefore, it is interesting to consider the role of GATA2 in regulating endothelial cell specificity. Indeed, it is not well understood how GATA factor binds the appropriate sites with cell-type specificity. One possibility is that GATA factor interacts with other transcription factors, which select the binding region and define the tissue-specific expression. As shown in Figure 2B, the second and third enriched sequences via ChIP-seqs with GATA2 were different between endothelial cells and erythroids. K562 cells coenriched with Gfi-1b and TAL1, as reported with essential factors for erythroid differentiation (Saleque *et al*, 2002; Hall *et al*, 2003). In contrast, HMVEC coenriched with Ets family and AP-1. Our newly finding −139 kbp *endomucin* enhancer also involved AP-1 and Ets binding region (Supplementary Figure SIV). Functional interaction between GATA and AP-1 was reported in endothelial-specific endothelin 1 expression (Kawana *et al*, 1995). The importance of Ets family in endothelium specificity was discussed above. Moreover, we

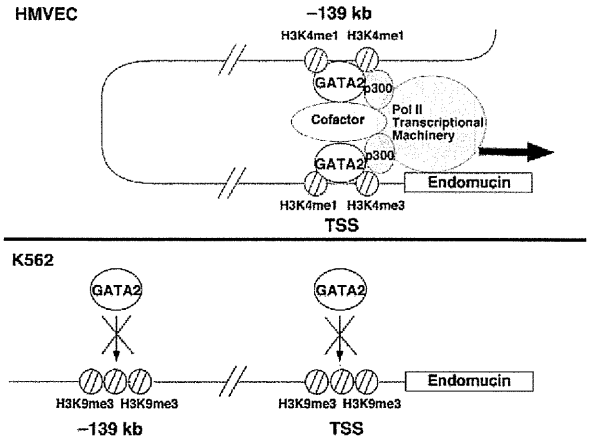


Figure 8 Schematic model for the endothelial-cell-specific *endomucin* gene expression.

recently reported that GATA interacts with Ets factor, ELF1, which cooperatively induces endothelial-specific Tie2 gene expression (Song *et al*, 2009).

Another possibility is that GATA factor interacts with an epigenetically modified enzyme, which regulates chromatin conformation and drives tissue specificity. For example, GATA2 interacts with histone deacetylase (HDAC)3 and HDAC5 in haematopoietic stem cells (Ozawa *et al*, 2001). GATA1 interacts with EZH2 and Suz12 in erythroids that regulates the erythroid cell differentiation (Yu *et al*, 2009). However, it is not well understood whether pre-setting epigenetically modified enzymes allow the association of tissue-specific transcription factors, or the transcription factors positively regulate the epigenetic circumstance. Further analysis would be needed to define the possibility for endothelial cell differentiation.

Chromatin conformation undergoes dynamic changes in various differentiated cells. In erythroid cells, 3C assays were performed to show the communication between the promoter and locus control enhancer region in β-globin gene. GATA1 and FOG1 interaction regulates chromatin-loop formation (Kim *et al*, 2007). Here, we first show the endothelial-cell-specific GATA2-mediated chromatin conformation. Our 3C assays revealed a chromatin loop in the *endomucin* locus. GATA2 was indispensable for such loop formation. Similar to the β-globin locus in erythroids, a region over 100 kbp apart from the TSS region communicated with the core promoter, mediated by GATA2, p300, and active PolII. Such interactions might be associated with the active histone code in an endothelial-cell-specific manner, since H3K4me1 and H3K4me3 bindings were also detected in distal enhancer and core promoter, respectively, in HMVEC, while K562 cells did not show H3K4me1-positive or H3K4me3-positive, but rather H3K9me3-positive heterochromatin formation. In addition, our reporter assays revealed that the −139 kbp distal regulatory region also functioned as an enhancer for the heterologous (TK) promoter and the intact proximal *endomucin* promoter. Especially in *endomucin*, both of the GATA motifs in the distal enhancer and proximal promoter were essential for the expression in endothelial cells. Collectively, the −139 kbp distal regulatory region could function for not only the endothelial-specific chromatin-loop conformation, but also downstream *endomucin* gene expression.

In summary, our findings reveal a mechanism by which GATA2 binds and regulates the tissue-specific *endomucin* gene expression (Figure 8). Moreover, loss of GATA2 initiates cell-type conversion in several populations occurring even in epigenetically defined endothelial cells. Cooperation with epigenetically regulatory enzymes and tissue-specific transcription factors would define the endothelial cell commitment and differentiation.

Materials and methods

Additional information is available in the Supplementary data.

Cell culture

HUVEC and HMVEC were purchased from LONZA. All endothelial cells were cultured in EGM-2MV complete medium (LONZA). Human skin fibroblast (LONZA) and COS7 cells (ATCC CRL-1651) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% heat-inactivated fetal bovine serum (FBS). K562 cells (ATCC CCL-243) were grown in RPMI 1640 medium plus 10% FBS.

Immunohistochemistry

HMVEC or frozen tissue sections obtained from human dermis were fixed and incubated with an anti-VE-Cadherin antibody (R&D systems), anti-vWF antibody (Abcam), anti-SM-actin (Sigma), anti-CD34 (NOVOCASTRA), or an anti-GATA2 antibody. Sections were washed and incubated with secondary antibody labelled with Alexa-Fluor 488 or Alexa-Fluor 594 (Invitrogen) for 1 h. The slides were then washed and mounted in ProLong Gold antifade reagent with DAPI (Invitrogen), and examined by fluorescence microscopy.

DNA microarray

HMVEC were transfected with either si-control or si-GATA2 for 48 h. RNA was harvested and purified with Trizol (Invitrogen). Preparation of cRNA and hybridization of probe arrays were performed according to the manufacturer's instructions (Affymetrix). Annotation of the probe numbers and targeted sequences are shown on the Affymetrix web page.

ChIP-seq analysis

All protocols for Illumina/Solexa sequence preparation, sequencing, and quality control are provided by Illumina. A brief summary of the technique and minor protocol modifications are described in Supplementary data.

Chromosome conformation capture (3C) assay

HMVEC or K562 cells were cross-linked with 1% formaldehyde for 10 min. The reaction was stopped by the addition of 125 mM glycine for 5 min. Nuclei were re-suspended with restriction enzyme buffer, treated with 7.5 μl of 20% SDS at 37°C for 1 h, and then incubated with 50 μl of 20% Triton X-100 for 1 h. Samples were treated with 400 U of *Bst*YI at 37°C for 16 h. After enzyme digestion, 40 μl of 20% SDS was added and incubated at 65°C for 10 min. The samples were diluted with ligation buffer and treated with T4 DNA ligase (2000 U) at 16°C for 16 h. Samples were finally reverse cross-linked and purified with spin columns (Qiagen). All primers were shown in Supplementary Table SIV. Three BAC clones (RP11-1041D18, RP11-760F14, RP11-891J2) were used to validate 3C assay.

FACS analysis

si-control-or si-GATA2-treated HMVEC were harvested with 0.05% trypsin plus 0.6 mM EDTA. Following cells were diluted with PBS plus 1% FBS, and fixed with ethanol. After washing by PBS twice, samples were re-suspended with PBS plus 0.2% Triton X-100, and incubated with antibodies for α-SM-actin (Sigma), vWF (Abcam), claudin 1 (Abnova), or KDR (Cell Signaling). PE-labelled $F(ab')_2$ fragments of donkey anti-mouse IgG (for α-SM-actin or claudin 1) or goat anti-rabbit IgG labelled with Alexa-Fluor 488 (for vWF or KDR) were used as secondary antibodies . Finally, samples were re-suspended $\sim 6 \times 10^5$ cells/ml with PBS plus 1% FBS, and immediately analysed by flow cytometry (Becton-Dickinson).

Data access
The data indicated in this publication are accessible through National Center for Biotechnology Information; Gene Expression Omnibus (GSE 28304) for microarrays, and Sequence Read Archive (SRA 030934.1) for ChIP-seqs.

Supplementary data
Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

This study was supported by the Leading-edge Research Promotion fund from Japan Society for the Promotion of Science (to TM), and in part supported by a Grant-in-Aid for Scientific Research on Innovative Areas from Ministry of Education, Culture, Sports, Science, and Technology in Japan (to TM), and in part by Mochida Memorial and Sankyo Science Foundation in Japan (to TM). We are grateful to Dr Shogo Yamamoto and Ms Akashi Izumi (The University of Tokyo) for technical assistance on ChIP-seqs and microarrays. We thank Dr Patrick C Reid (Peptidream Inc., Tokyo) for critical reading of the manuscript.

Author contributions: YK, SJ, J-iS, IM, TatK., and TM performed and designed the experiment; TK, ST, YO, SI, and HA analysed the data; HI, YW, MN, and TH contributed new reagents; and TM wrote the paper

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aird WC (2007) Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circ Res* **100:** 174–190

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2:** 28–36

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823–837

Bieda M, Xu X, Singer MA, Green R, Farnham PJ (2006) Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. *Genome Res* **16:** 595–605

Blahnik KR, Dou L, O'Geen H, McPhillips T, Xu X, Cao AR, Iyengar S, Nicolet CM, Ludascher B, Korf I, Farnham PJ (2010) Sole-Search: an integrated analysis program for peak detection and functional annotation using ChIP-seq data. *Nucleic Acids Res* **38:** e13

Chong HK, Infante AM, Seo YK, Jeon TI, Zhang Y, Edwards PA, Xie X, Osborne TF (2010) Genome-wide interrogation of hepatic FXR reveals an asymmetric IR-1 motif and synergy with LRH-1. *Nucleic Acids Res* **38:** 6007–6017

De Val S, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, Ehlers ML, Agarwal P, Visel A, Xu SM, Pennacchio LA, Dubchak I, Krieg PA, Stainier DY, Black BL (2008) Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell* **135:** 1053–1064

Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* **295:** 1306–1311

Ewenstein BM, Warhol MJ, Handin RI, Pober JS (1987) Composition of the von Willebrand factor storage organelle (Weibel-Palade body) isolated from cultured human umbilical vein endothelial cells. *J Cell Biol* **104:** 1423–1433

Fujiwara T, O'Geen H, Keles S, Blahnik K, Linnemann AK, Kang YA, Choi K, Farnham PJ, Bresnick EH (2009) Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* **36:** 667–681

Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY *et al* (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* **462:** 58–64

Grass JA, Jing H, Kim SI, Martowicz ML, Pal S, Blobel GA, Bresnick EH (2006) Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol Cell Biol* **26:** 7056–7067

Gumina RJ, Kirschbaum NE, Piotrowski K, Newman PJ (1997) Characterization of the human platelet/endothelial cell adhesion molecule-1 promoter: identification of a GATA-2 binding element required for optimal transcriptional activity. *Blood* **89:** 1260–1269

Hall MA, Curtis DJ, Metcalf D, Elefanty AG, Sourris K, Robb L, Gothert JR, Jane SM, Begley CG (2003) The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12. *Proc Natl Acad Sci USA* **100:** 992–997

Jahroudi N, Lynch DC (1994) Endothelial-cell-specific regulation of von Willebrand factor gene expression. *Mol Cell Biol* **14:** 999–1008

Jing H, Vakoc CR, Ying L, Mandat S, Wang H, Zheng X, Blobel GA (2008) Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell* **29:** 232–242

Kappel A, Schlaeger TM, Flamme I, Orkin SH, Risau W, Breier G (2000) Role of SCL/Tal-1, GATA, and ets transcription factor binding sites for the regulation of flk-1 expression during murine vascular development. *Blood* **96:** 3078–3085

Kawana M, Lee ME, Quertermous EE, Quertermous T (1995) Cooperative interaction of GATA-2 and AP1 regulates transcription of the endothelin-1 gene. *Mol Cell Biol* **15:** 4225–4231

Khandekar M, Brandt W, Zhou Y, Dagenais S, Glover TW, Suzuki N, Shimizu R, Yamamoto M, Lim KC, Engel JD (2007) A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development* **134:** 1703–1712

Kim SI, Bultman SJ, Jing H, Blobel GA, Bresnick EH (2007) Dissecting molecular steps in chromatin domain activation during hematopoietic differentiation. *Mol Cell Biol* **27:** 4551–4565

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465:** 182–187

Kinoshita M, Nakamura T, Ihara M, Haraguchi T, Hiraoka Y, Tashiro K, Noda M (2001) Identification of human endomucin-1 and -2 as membrane-bound O-sialoglycoproteins with anti-adhesive activity. *FEBS Lett* **499:** 121–126

Kouzarides T (2007) Chromatin modifications and their function. *Cell* **128:** 693–705

Liu C, Shao ZM, Zhang L, Beatty P, Sartippour M, Lane T, Livingston E, Nguyen M (2001) Human endomucin is an endothelial marker. *Biochem Biophys Res Commun* **288:** 129–136

Liu J, Kanki Y, Okada Y, Jin E, Yano K, Shih SC, Minami T, Aird WC (2009) A +220 GATA motif mediates basal but not endotoxin-repressible expression of the von Willebrand factor promoter in Hprt-targeted transgenic mice. *J Thromb Haemost* **7:** 1384–1392

Liu J, Yuan L, Molema G, Regan E, Janes L, Beeler D, Spokes KC, Okada Y, Minami T, Oettgen P, Aird WC (2011) Vascular bed-specific regulation of the von Willebrand factor promoter in the heart and skeletal muscle. *Blood* **117:** 342–351

Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M (2008) FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132:** 958–970

Mammoto A, Connor KM, Mammoto T, Yung CW, Huh D, Aderman CM, Mostoslavsky G, Smith LE, Ingber DE (2009) A mechanosensitive transcriptional mechanism that controls angiogenesis. *Nature* **457:** 1103–1108

Minami T, Aird WC (2001) Thrombin stimulation of the vascular cell adhesion molecule-1 promoter in endothelial cells is mediated by tandem nuclear factor-kappa B and GATA motifs. *J Biol Chem* **276:** 47632–47641

Minami T, Aird WC (2005) Endothelial cell gene regulation. *Trends Cardiovasc Med* **15:** 174–184

Minami T, Murakami T, Horiuchi K, Miura M, Noguchi T, Miyazaki J, Hamakubo T, Aird WC, Kodama T (2004) Interaction between

hex and GATA transcription factors in vascular endothelial cells inhibits flk-1/KDR-mediated vascular endothelial growth factor signaling. *J Biol Chem* **279:** 20626–20635

Minami T, Rosenberg RD, Aird WC (2001) Transforming growth factor-beta 1-mediated inhibition of the flk-1/KDR gene is mediated by a 5′-untranslated region palindromic GATA site. *J Biol Chem* **276:** 5395–5402

Minami T, Yano K, Miura M, Kobayashi M, Suehiro J, Reid PC, Hamakubo T, Ryeom S, Aird WC, Kodama T (2009) The Down syndrome critical region gene 1 short variant promoters direct vascular bed-specific gene expression during inflammation in mice. *J Clin Invest* **119:** 2257–2270

Ozawa Y, Towatari M, Tsuzuki S, Hayakawa F, Maeda T, Miyata Y, Tanimoto M, Saito H (2001) Histone deacetylase 3 associates with and represses the transcription factor GATA-2. *Blood* **98:** 2116–2123

Patient RK, McGhee JD (2002) The GATA family (vertebrates and invertebrates). *Curr Opin Genet Dev* **12:** 416–422

Potenta S, Zeisberg E, Kalluri R (2008) The role of endothelial-to-mesenchymal transition in cancer progression. *Br J Cancer* **99:** 1375–1379

Raney BJ, Cline MS, Rosenbloom KR, Dreszer TR, Learned K, Barber GP, Meyer LR, Sloan CA, Malladi VS, Roskin KM, Suh BB, Hinrichs AS, Clawson H, Zweig AS, Kirkup V, Fujita PA, Rhead B, Smith KE, Pohl A, Kuhn RM *et al* (2011) ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic Acids Res* **39** (Database issue): D871–D875

Robertson AG, Bilenky M, Tam A, Zhao Y, Zeng T, Thiessen N, Cezard T, Fejes AP, Wederell ED, Cullum R, Euskirchen G, Krzywinski M, Birol I, Snyder M, Hoodless PA, Hirst M, Marra MA, Jones SJ (2008) Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res* **18:** 1906–1917

Saleque S, Cameron S, Orkin SH (2002) The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* **16:** 301–306

Samulowitz U, Kuhn A, Brachtendorf G, Nawroth R, Braun A, Bankfalvi A, Bocker W, Vestweber D (2002) Human endomucin: distribution pattern, expression on high endothelial venules, and decoration with the MECA-79 epitope. *Am J Pathol* **160:** 1669–1681

Schmidt D, Wilson MD, Ballester B, Schwalie PC, Brown GD, Marshall A, Kutter C, Watt S, Martinez-Jimenez CP, Mackay S, Talianidis I, Flicek P, Odom DT (2010) Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328:** 1036–1040

Song H, Suehiro J, Kanki Y, Kawai Y, Inoue K, Daida H, Yano K, Ohhashi T, Oettgen P, Aird WC, Kodama T, Minami T (2009) Critical role for GATA3 in mediating Tie2 expression and function in large vessel endothelial cells. *J Biol Chem* **284:** 29109–29124

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102:** 15545–15550

Tsai FY, Keller G, Kuo FC, Weiss M, Chen J, Rosenblatt M, Alt FW, Orkin SH (1994) An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* **371:** 221–226

Ueno M, Igarashi K, Kimura N, Okita K, Takizawa M, Nobuhisa I, Kojima T, Kitamura T, Samulowitz U, Vestweber D, Shimomura T, Suda T, Nakashima K, Taga T (2001) Endomucin is expressed in embryonic dorsal aorta and is able to inhibit cell adhesion. *Biochem Biophys Res Commun* **287:** 501–506

Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglou S, Myers RM, Sidow A (2008) Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* **5:** 829–834

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457:** 854–858

Wilson DB, Dorfman DM, Orkin SH (1990) A nonerythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. *Mol Cell Biol* **10:** 4854–4862

Wozniak RJ, Boyer ME, Grass JA, Lee Y, Bresnick EH (2007) Context-dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J Biol Chem* **282:** 14665–14674

Yu M, Riva L, Xie H, Schindler Y, Moran TB, Cheng Y, Yu D, Hardison R, Weiss MJ, Orkin SH, Bernstein BE, Fraenkel E, Cantor AB (2009) Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* **36:** 682–695

Zeisberg EM, Potenta S, Xie L, Zeisberg M, Kalluri R (2007a) Discovery of endothelial to mesenchymal transition as a source for carcinoma-associated fibroblasts. *Cancer Res* **67:** 10123–10128

Zeisberg EM, Tarnavski O, Zeisberg M, Dorfman AL, McMullen JR, Gustafsson E, Chandraker A, Yuan X, Pu WT, Roberts AB, Neilson EG, Sayegh MH, Izumo S, Kalluri R (2007b) Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat Med* **13:** 952–961

Zhang R, Min W, Sessa WC (1995) Functional analysis of the human endothelial nitric oxide synthase promoter. Sp1 and GATA factors are necessary for basal transcription in endothelial cells. *J Biol Chem* **270:** 15320–15326