【今井　浩三】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Nasuno M, Arimura Y, Nagaishi K, Isshiki H, Onodera K, Nakagaki S, Watanabe S, Idogawa M, Yamashita K, Naishiro Y, Adachi Y, Suzuki H, Fujimiya M, Imai K, Shinomura Y.	Mesenchymal stem cells cancel azoxymethane-induced tumor initiation.	STEM CELLS n/a–n/a DOI	10.1002/stem	1594	2013
Suzuki R, Yamamoto E, Nojima M, Maruyama R, Yamano HO,	Aberrant methylation of microRNA-34b/c is a predictive marker of metachronous gastric cancer risk.	J Gastroenterol	[Epub ahead of print]		2013
Shimizu T, Suzuki H, Nojima M, Kitamura H, Yamamoto E, Maruyama R, Ashida M,	Methylation of a panel of microRNA genes is a novel biomarker for detection of bladder cancer.	Eur Urol	63	1091-1100	2013
Sawada T, Yamamoto E, Suzuki H, Nojima M, Maruyama R, Shioi Y, Akasaka R, Kamimae S,	Association between genomic alterations and metastatic behavior of colorectal cancer identified by array-based comparative genomic hybridization.	Genes Chromosomes Cancer	52	140-149	2013

Kato N, Yamamoto H, Adachi Y, Ohashi H, Taniguchi H, Suzuki H, Nakazawa M, Kaneto H, Sasaki S, Imai K, Shinomura Y.	Cancer detection by ubiquitin carboxyl-terminal esterase L1 methylation in pancreatobiliary fluids.	World J Gastroenterol	19	1718-1727	2013
Adachi Y, Ohashi H, Imsumran A, Yamamoto H, Matsunaga Y, Taniguchi H, Nosho K, Suzuki H, Sasaki Y, Arimura Y, Carbone DP, Imai K, Shinomura Y.	The effect of IGF-I receptor blockade for human esophageal squamous cell carcinoma and adenocarcinoma.	Tumor Biol	[Epub ahead of print		2013
Watanabe S, Arimura Y, Nagaishi K, Isshiki H, Onodera K, Nasuno M, Yamashita K, Idogawa M, Naishiro Y, Murata M, Adachi Y, Fujimiya M, Imai K, Shinomura Y.	Conditioned mesenchymal stem cells produce pleiotropic gut trophic factors.	J Gastroenterol	Nov 12. [Epub ahead of print]		2013

Arimura Y, Isshiki H, Onodera K, Nagaishi K, Yamashita K, Sonoda T, Matsumoto T, Takahashi A, Takazoe M, Yamazaki K, Kubo M, Fujimiya M, Imai K, Shinomura Y.	Characteristics of Japanese inflammatory bowel disease susceptibility loci.	J Gastroenterol	Aug 13. [Epub ahead of print]		2013
Adachi E, Kogayu M, Fjii T, Mae H, Shimizu H, Iwai Y, Shibata H, Suzuki M, Imai K, Koibuchi T.	Tuberculosis examination using whole blood interferon-gamma release assay among health care workers in a Japanese hospital without tuberculosis-specific wards.	SpringerPlus	2:440(05 Sep)		2013
Ito M, Mitsuhashi K, Igarashi H, Nosho K, Naito T, Yoshii S, Takahashi H, Fujita M, Sukawa Y, Yamamoto E, Takahashi T, Adachi Y, Nojima M, Sasaki Y, Tokino T, Baba T, Maruyama R, Suzuki H, Imai K, Yamamoto H, Shinomura Y.	MicroRNA-31 expression in relation to BRAF mutation, CpG island methylation and colorectal continuum in serrated lesions.	J Pathology	In press		2014

安井寛，今井浩三	がん免疫応答の制御：抗がん細胞抗体、その開発のあゆみとがん抗体療法の新たな可能性	実験医学	31(12)	121-127	2013
春日雅人，植木浩二郎，田嶋尚子，野田光彦，大橋健，能登洋，後藤温，小川渉，堺隆一，津金昌一郎，浜島信之，中釜斉，田島和雄，宮園浩平，今井浩三	糖尿病と癌に関する委員会報告	糖尿病	56(6)	374-390	2013
安井寛，今井浩三	バイオ医薬品とＤＤＳ　特集「ＤＤＳ技術の進歩と医療応用」	メディカル・サイエンス・ダイジェスト	40(2)	15-18	2014

【藤渕　航】

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ
藤渕航	iPS細胞からのビッグデータの情報セキュリティと創薬、医療への活用	植田充弘	生命のビッグデータ利用の最前線	CMC出版	日本	2014	

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
山根順子 丸山徹 藤渕航	単細胞技術に基づくiPS細胞の標準化	生体の科学	65巻2号	印刷中	2014

【中畑　龍俊】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Yanagimachi MD., Niwa A., Tanaka T., Honda-Ozaki F., Nishimoto S., Murata Y., Yasumi T., Ito J., Tomida S., Oshima K., Asaka I., Goto H., Heike T., Nakahata T., Saito MK.	Robust and Highly-Efficient Differentiation of Functional Monocytic Cells from Human Pluripotent Stem Cells under Serum- and Feeder Cell-Free Conditions.	PLOS One	8(4)	e59243	2013

Saida S, Watanabe KI, Sato-Otsubo A, Terui K, Yoshida K, Okuno Y, Toki T, Wang R, Shiraishi Y, Miyano S, Kato I, Morishima T, Fujino H, Umeda K, Hiramatsu H, Adachi S, Ito E, Ogawa S, Ito M, Nakahata T, Heike T.	Clonal selection in xenografted TAM recapitulates the evolutionary process of myeloid leukemia in Down syndrome.	Blood	121(21)	4377-87	2013

Kondo T, Asai M, Tsukita K, Kutoku Y, Ohsawa Y, Sunada Y, Imamura K, Egawa N, Yahata N, Okita K, Takahashi K, Asaka I, Aoi T, Watanabe A, Watanabe K, Kadoya C, Nakano R, Watanabe D, Maruyama K, Hori O, Hibino S, Choshi T, Nakahata T, Hioki H, Kaneko T, Naitoh M, Yoshikawa K, Yamawaki S, Suzuki S, Hata R, Ueno S, Seki T, Kobayashi K, Toda T, Murakami K, Irie K, Klein WL, Mori H, Asada T, Takahashi R, Iwata N, Yamanaka S, Inoue H.	Modeling Alzheimer's Disease with iPSCs Reveals Stress Phenotypes Associated with Intracellular Aß and Differential Drug Responsiveness.∴ 2013	Cell Stem Cell.	12(4)	487-96	

Morishima T, Watanabe KI, Niwa A, Hirai H, Saida S, Tanaka T, Kato I, Umeda K, Hiramatsu H, Saito MK, Matsubara K, Adachi S, Kobayashi M, Nakahata T, Heike T.	Genetic correction of HAX1 in induced pluripotent stem cells from a patient with severe congenital neutropenia improves defective granulopoiesis. ;. 2014	Haematologica.	99(1)	19-27	2014
斎藤潤、中畑龍俊	疾患特異的 iPS 細胞	再生医療	12(1)	19-29	2013

【中辻　憲夫】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Miyazak,i T., Nakatsuji, N., Suemori, H.	Optimization of slow cooling cryopreservation for human pluripotent stem cells.	Genesis	Jan;52(1)	49-55	2014
Watatani, K., Xie, Z., Nakatsuji, N. and Sengoku, S.	Global competencies from regional stem cell research: Bibliometrics for investigating and forecasting research trends.	Regen.	Med. 8	659-668	2013

Kumagai, H., Suemori, H., Uesugi, M., Nakatsuji, N., Kawase, E.	Identification of small molecules that promote human embryonic stem cell self-renewal.	Biochem Biophys Res Commun.	434	710-716	2013
Nakatsuji, N., Kawase, E., Miyazaki, T., Minami, I. and Aiba, K.	Multidisciplinary research of human pluripotent stem cells for application to cell therapy and drug discovery.	Tiss. Engin. Regen.	10	1-4	2013
Nakatsuji, N.	Mesoscopic science, where materials become life and life inspires materials. A great opportunity to push back the frontiers of life, materials, and biomaterials sciences.	Biomater.Sci.	1	9-10	2013
Uosaki, H., Magadum, A., Seo, K., Fukushima, H., Takeuchi, A., Nakagawa, Y., Moyes, K.W., Narazaki, G., Kuwahara, K., Laflamme, M., Matsuoka, ., Nakatsuji, N., Nakao, K., Kwon, C., Kass, D.A., Engel, F.B., Yamashita, J.K.	Identification of Chemicals Inducing Cardiomyocyte Proliferation in Developmental Stage-Specific Manner with Pluripotent Stem Cells.	Circ. Cardiovasc. Genet.	6	624-633	2013

【梅澤　明弘】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Fukami M, Tsuchiya T, Vollbach H, Brown KA, Abe S, Ohtsu S, Wabitsch M, Burger H, Simpson ER, Umezawa A, Shihara D, Nakabayashi K, Bulun SE, Shozu M, Ogata T.	Genomic basis of aromatase excess syndrome: recombination- and replication-mediated rearrangements leading to CYP19A1 overexpression.	J Clin Endocrinol Metab.	98(12)	E2013-2021	2013
Higuchi A, Ling QD, Chang Y, Hsu ST, Umezawa A.	Physical cues of biomaterials guide stem cell differentiation fate.	Chem Rev.	113(5)	3297-3328	2013

【西田　幸二】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
西田幸二	角膜上皮の再生医療	日本眼科学会雑誌	117巻1号	69-83	2013
西田幸二、大家義則	イノベーションをマーケットへ シーズ成果報告<再生医療>　角膜上皮幹細胞疲弊症に対する自己培養口腔粘膜上皮細胞シート移植法の多施設共同臨床試験	臨床評価	41巻1号	120-123	2013

【大和　雅之】

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Obokata H, Wakayama T, Sasai Y, Kojima K, Vacanti MP, Niwa H, Yamato M, Vacanti CA.	Stimulus-triggered fate conversion of somatic cells into pluripotency	Nature	505(7485)	641-7.	2014
Suphanantachat S, Iwata T, Ishihara J, Yamato M, Okano T, Izumi Y.	A role for c-Kit in the maintenance of undifferentiated human mesenchymal stromal cells.	Biomaterials	35(11)	3618-26.	2014
Washio K, Kuroda H, Iwata T, Yoshida T, Yamato M, Okano T.	Improved enzymatic treatment for accurate cell counting from extracellular matrix-rich periodontal ligament cell sheets.	Int J Oral Maxillofac Implants.	29(1)	e117-21.	2014
Itoga K, Kobayashi J, Yamato M, Okano T.	Micropatterning with a liquid crystal display (LCD) projector.	Methods Cell Biol.	119	141-58.	2014;
Kikuchi T, Shimizu T, Wada M, Yamato M, Okano T.	Automatic fabrication of 3-dimensional tissues using cell sheet manipulator technique.	Biomaterials.	35(8)	2428-35.	2014

Takaku Y, Murai K, Ukai T, Ito S, Kokubo M, Satoh M, Kobayashi E, Yamato M, Okano T, Takeuchi M, Mochida J, Sato M.	In vivo cell tracking by bioluminescence imaging after transplantation of bioengineered cell sheets to the knee joint.	Biomaterials.	35(7)	2199-206.	2014
Sekine W, Haraguchi Y, Shimizu T, Yamato M, Umezawa A, Okano T.	Chondrocyte differentiation of human endometrial gland-derived MSCs in layered cell sheets.	ScientificWorld Journal	359	109	2013
Iwata T, Yamato M, Ishikawa I, Ando T, Okano T.	Tissue engineering in periodontal tissue.	Anat Rec (Hoboken)	297(1)	16-25	2014
Sakuma M, Kumashiro Y, Nakayama M, Tanaka N, Umemura K, Yamato M, Okano T.	Control of cell adhesion and detachment on Langmuir-Schaefer surface composed of dodecyl-terminated thermo-responsive polymers.	J Biomater Sci Polym Ed.	25(5)	431-43.	2014
Sato M, Yamato M, Hamahashi K, Okano T, Mochida J.	Articular cartilage regeneration using cell sheet technology.	Anat Rec (Hoboken)	297(1)	36-43	2014
Egami M, Haraguchi Y, Shimizu T, Yamato M, Okano T.	Latest status of the clinical and industrial applications of cell sheet engineering and regenerative medicine.	Arch Pharm Res.	37(1)	96-106	2014

Matsumine H, Sasaki R, Takeuchi Y, Miyata M, Yamato M, Okano T, Sakurai H.	Vascularized versus Nonvascularized Island Median Nerve Grafts in the Facial Nerve Regeneration and Functional Recovery of Rats for Facial Nerve Reconstruction Study.	J Reconstr Microsurg.	30(2)	127-36	2014
Yoshimitsu K, Muragaki Y, Maruyama T, Yamato M, Iseki H.	Development and Initial Clinical Testing of "OPECT": An Innovative Device for Fully Intangible Control of the Intraoperative Image-Displaying Monitor by the Surgeon.	Neurosurgery.	10	Suppl 1:46-50.	2014
Sasagawa T, Shimizu T, Yamato M, Okano T.	Expression profiles of angiogenesis-related proteins in prevascular three-dimensional tissues using cell-sheet engineering.	Biomaterials.	35(1)	206-13	2014
Tanaka N, Kondo M, Uchida R, Kaneko M, Sugiyama H, Yamato M, Okano T.	Splitting culture medium by air-jet and rewetting for the assessment of the wettability of cultured epithelial cell surfaces.	Biomaterials.	34(36)	9082-8	2013
Nagase K, Hatakeyama Y, Shimizu T, Matsuura K, Yamato M, Takeda N, Okano T.	Hydrophobized thermoresponsive copolymer brushes for cell separation by multistep temperature change.	Biomacromolecules.	14(10)	3423-33	2013
Kobayashi T, Kan K, Nishida K, Yamato M, Okano T.	Corneal regeneration by transplantation of corneal epithelial cell sheets fabricated with automated cell culture system in rabbit model.	Biomaterials.	34(36)	9010-7	2013

Arisaka Y, Kobayashi J, Yamato M, Akiyama Y, Okano T.	Heparin-functionalized thermoresponsive surface: a versatile cell culture substrate for regulating multivalent affinity binding with heparin-binding proteins by	Organogenesis.	9(3)	125-7	2013
Tadakuma K, Tanaka N, Haraguchi Y, Higashimori M, Kaneko M, Shimizu T, Yamato M, Okano T.	A device for the rapid transfer/transplantation of living cell sheets with the absence of cell damage.	Biomaterials.	34(36)	9018-25.	2013
Matsuzaka N, Nakayama M, Takahashi H, Yamato M, Kikuchi A, Okano T.	Terminal-functionality effect of poly(N-isopropylacrylamide) brush surfaces on temperature-controlled cell adhesion/detachment.	Biomacromolecules.	14(9)	3164-71	2013
Takahashi H, Shimizu T, Nakayama M, Yamato M, Okano T.	The use of anisotropic cell sheets to control orientation during the self-organization of 3D muscle tissue.	Biomaterials.	34(30)	7372-80.	2013
Yoshida T, Iwata T, Umemoto T, Shiratsuchi Y, Kawashima N, Sugiyama T, Yamato M, Okano T.	Promotion of mouse ameloblast proliferation by Lgr5 mediated integrin signaling.	J Cell Biochem.	114(9)	2138-47	2013

Sasagawa T, Shimizu T, Sekiya S, Yamato M, Okano T.	Comparison of angiogenic potential between prevascular and non-prevascular layered adipose-derived stem cell-sheets in early post-transplanted period.	J Biomed Mater Res A.	102(2)	358-65	2014

IV. 研究成果の刊行物・別刷

Computational promoter modeling identifies the modes of transcriptional regulation in hematopoietic stem cells

Sung-Joon Park[1], Terumasa Umemoto[2], Mihoko Saito-Adachi[1], Yoshiko Shiratsuchi[2], Masayuki Yamato[2], Kenta Nakai[1,*]

[1]Human Genome Center, the Institute of Medical Science, the University of Tokyo, Tokyo 108-8639, Japan
[2]Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo 162-8666, Japan
*E-mail: knakai@ims.u-tokyo.ac.jp

Abstract

Extrinsic and intrinsic regulators are responsible for the tight control of hematopoietic stem cells (HSCs), which differentiate into all blood cell lineages. To understand the fundamental basis of HSC biology, we focused on differentially expressed genes (DEGs) in long-term and short-term HSCs, which are closely related in terms of cell development but substantially differ in their stem cell capacity. To analyze the transcriptional regulation of the DEGs identified in the novel transcriptome profiles obtained by our RNA-seq analysis, we developed a computational method to model the linear relationship between gene expression and the features of putative regulatory elements. The transcriptional regulation modes characterized here suggest the importance of transcription factors (TFs) that are expressed at steady state or at low levels. Remarkably, we found that 24 differentially expressed TFs targeting 21 putative TF-binding sites contributed significantly to transcriptional regulation. These TFs tended to be modulated by other nondifferentially expressed TFs, suggesting that HSCs can achieve flexible and rapid responses via the control of nondifferentially expressed TFs through a highly complex regulatory network. Our novel transcriptome profiles and new method are powerful tools for studying the mechanistic basis of cell fate decisions.

Keywords: hematopoiesis, log-linear model, regression model, RNA sequencing, transcription factor

Introduction

Hematopoiesis is a complex and dynamic process, which generates mature blood cells throughout the life of organisms. In the adult bone marrow, long-term hematopoietic stem cells (LT-HSCs) maintain a balanced pool of stem cells, which also differentiates into more mature short-term hematopoietic stem cells (ST-HSCs), multipotent progenitors with a lower self-renewal capacity. It is believed that the blood lineage choice of HSCs is governed by a stepwise cell fate decision [1, 2]. However, recent studies have raised questions about the hierarchical hematopoietic system [3, 4]. Many studies based on genome-wide gene expression profiling [5–9] have demonstrated that specific extrinsic and intrinsic regulators play key roles in hematopoiesis [10–12]. Recently, high-throughput sequencing techniques have been applied widely [13–15], which have provided new insights into *in vivo* transcription factor (TF) binding and epigenetic modifications [16–18]. Systems biology approaches are also enhancing our understanding of the regulatory dynamics of hematopoiesis [19].

Despite the biological importance of the formation of all blood cells via a transition from LT-HSC to ST-HSC, little is known about the mechanism that underlies this early differentiation. A major explanation for this deficiency is a lack of comprehensive genome-wide identification studies and characterizations of the regulatory elements that govern gene expression in HSCs. The profiling of potential key regulators [8,17,20] and the large-scale integration of datasets [21,22] have improved our understanding greatly.

However, these studies are limited to a small number of factors that function in heterogeneous HSCs, which were isolated using different combinations of monoclonal antibodies. Therefore, unconsidered key regulators may exist at this early stage of hematopoiesis. Indeed, novel key factors [23, 24] and new multipotent progenitors [3, 4, 25] have been identified recently.

To address these deficiencies, we developed a computational method on the basis of novel transcriptome data from adult mouse bone marrow HSCs; $CD34^{-}KSL$ ($c\text{-}kit^{+}Sca1^{+}Lin^{-}$) LT-HSCs and $CD34^{+}KSL$ ST-HSCs, a widely used strategy to isolate HSCs at high purity [26, 27]. Our method uses a regression-based approach [28–30] to model the linear relationships between gene expression and the characteristics of regulatory elements compiled from a database. In the present study, we extended this regression modeling-based approach using large-scale log-linear modeling (LLM) [31], which considered the combinatorial nature of TFs. Thus, our method can systematically infer the regulation modes exerted by TFs that are probably necessary for gene expression, as well as suggesting synergistic TF modules. Using our transcriptome profiles and this novel method, we characterized transcriptional regulatory modes related to HSCs, which suggested the functional importance of TFs expressed at steady-state or low levels. Remarkably, we identified 24 differentially expressed TFs that targeted 21 putative TF-binding sites (TFBSs) in LT-HSCs. These TFs might be essential for maintaining the HSC capacity during the early stage of hematopoiesis.

Results

Extensive transcriptome discovery

RNA-seq analysis of HSCs

To establish transcriptional profiles, we extracted total RNA from mouse LT-HSCs ($CD34^{-}KSL$) and ST-HSCs ($CD34^{+}KSL$), and performed SOLiD RNA-seq assays in triplicate. We generated 44–70 million 50 bp short reads, among which 44%–63% were mapped uniquely to the mouse genome (mm9) via our recursive mapping strategy [32]. These uniquely mapped reads (uni-reads) were used for further analysis (Table S1). We used the TopHat/Cufflinks pipeline [33] to quantify the RNA abundance of RefSeq genes as fragments per kilobase of exon per million mapped reads (FPKM). This analysis confirmed the high reproducibility among replicates (Figure S1A). We also assessed the overlap between our profile and public expression profiles [8, 9]. This comparison showed that our RNA-seq assay uniquely identified 8275 and 9220 genes from LT- and ST-HSCs, respectively (Figure 1A). This indicates that our study successfully identified a more detailed transcriptome landscape than previous studies.

The application of different monoclonal antibodies to purify HSC populations may have diverse effects on the resulting expression profiles [2], which are related to issues regarding the functional purification of HSCs [10,26] and the heterogeneous expression in single cells [4,10,34]. In fact, a comparison between our findings and the results of an RNA-seq analysis of HSCs isolated using distinct markers [15] demonstrated that there were great differences, particularly among genes that were expressed at low levels (Figure S1B). In addition, we performed qRT-PCR using 90 genes that were randomly selected from our samples, and confirmed that RNA quantities relative to the housekeeping gene *B2m* were in overall agreement (Figure 1B). However, genes that were expressed at low levels were substantially different. These results suggest the difficulty in detecting and quantifying rare transcripts in HSCs.

Identification of differentially expressed genes (DEGs)

We identified genes with high expression levels (FPKM, > 3) and calculated the fold change (FC) in gene expression. This analysis detected the transcriptionally active state of ST-HSCs (Figure 1C), which supported the results of previous studies [6,7,15]. Our RNA-seq assay detected a higher number of DEGs than those reported previously, which may have been related to our more comprehensive transcriptome

discovery method. We categorized the genes into 4 classes using a change of 2-fold as the threshold [15] (Figure 1D): Class A, 363 genes upregulated in LT-HSC; Class B, 743 genes downregulated in LT-HSC; Class C, 6332 genes with FC ≤ 2 and FPKM > 3; and Class D, 6006 genes with low expression (FPKM, ≤ 3). Thus, Class A and Class B represented DEGs, Class C represented steady-state transcription genes, and Class D represented genes with noisy expression and/or functional low expression genes.

We searched for any gene ontology (GO) terms enriched in DEGs using the DAVID Bioinformatics Resources [35]. Figure 1E shows the representative GO terms (Tables S10 and S11 for complete lists). This analysis showed that DEGs were involved in the immune response, cell–cell communication, and signal transduction. This was not surprising because extrinsic and intrinsic signals and molecules contribute to the biology of HSCs in the bone marrow microenvironment [1, 10, 11, 36]. In addition to these common biological processes, Class A genes were involved particularly in cell death, cell differentiation, and homeostasis, whereas Class B genes were involved in DNA repair, cell cycle progression, and cell organization. These results were consistent with those of previous studies that showed that apoptosis and cell-cycle regulators play critical roles in maintaining a balanced pool of HSCs and in the expansion of progenitor populations [5, 37, 38].

Differentially expressed cell-surface molecules and TFs

DEGs included 77 cell-surface molecules with the "cell surface" (GO:0009986) GO term (Table S2), some of which are known to be associated with hematopoiesis: in Class A, *Vwf*, *Lhcgr*, *Cxcl12*, and *Tgfbr3*; in Class B, *CD244*, *CD33*, and *Clec12a*. *CD34*, which was used to isolate HSCs in this study, exhibited an upregulation of over 12-fold in ST-HSCs compared with LT-HSCs. To obtain high HSC purities, these cell-surface molecules will be useful as alternative or additional markers.

DEGs also included 57 TFs that were annotated in TRANSFAC [39], i.e., 31 in Class A and 26 in Class B (Tables 1 and S3). These differentially expressed TFs included known hematopoietic regulators (e.g., *Gata2*, *Tal1*, and *Satb1*) and previously unconsidered TFs, such as the hepatocyte nuclear factor *Foxa3*, the BTB-domain zinc finger *Zbtb20*, the DNA-binding domain *Arid5a*, and the epigenetic regulator *Uhrf1*. It was noteworthy that a large number of TFs belonged to Class C (303 TFs) and Class D (341 TFs) (Tables S4 and S5). In particular, TFs with synergistic functions in HSCs [17] and that belonged to TF families, such as Fox, Lmo, and Sox (which are required by HSCs), were present in Class C and/or Class D. These results may suggest that, in addition to differentially expressed TFs, TFs with coding genes that are expressed at stable or low levels are functionally important molecules.

Computational modeling of DEG promoters

Workflow overview of promoter modeling

To determine the upstream regulatory elements that are essential for DEG transcription, we used a linear regression model that was used widely for this purpose in previous studies [28, 30]. The underlying assumption of this model is that the expression levels of genes are controlled by the sum of the independent activities of regulators, such as DNA-binding factors or epigenetic marks. These activities can be approximated using high-throughput *in vivo* experiments [40, 41] or knowledge-based computational approaches [25, 30]. As a preliminary test, we applied the linear regression model described in our previous study [29] using ChIP-seq data for 10 major TFs [17]. In this approach, we used genome-wide TF-binding instances that occurred within ± 5 kb regions from transcription start sites (TSSs), and predicted the FPKMs of DEGs by using a simple linear regression model with rigorous statistical tests. However, we were unable to detect any significant effects, and the correlation between the observed and predicted FPKMs was < 0.3. This failure may, in part, reflect the possibility that these TFs exert regulatory functions as distal enhancers, rather than through proximal promoters [17, 42].

To identify regulators from proximal promoter regions comprehensively, we used TRANSFAC [39],

which is a database that curates > 1.5 million ChIP-seq sites, and designed a workflow coupled with intensive computations (Figure 2). First, we prepared the promoter sequences of DEGs and searched for putative TFBSs and mouse TFs that are known to bind to the TFBSs in TRANSFAC using the MATCH tool [43]. This procedure identified 140 and 141 TFBSs for Class A and Class B promoters, respectively. Among these, 70 TFBSs in Class A and 69 TFBSs in Class B were targeted by at least one TF with a highly expressed coding gene (FPKM, > 3). In total, 265 and 267 TFs were involved in Class A and Class B, respectively.

Next, we calculated the TFBS–gene association score (TGAS) using 5 distinct scoring schemes, which were employed as explanatory variables in a linear regression model. These scores considered matrix similarity, positional bias of TFBSs, the expression levels of TFs, and the probability of TF–TF interactions (Materials and Methods). Given a TGAS, we searched exhaustively for the best combination of TFBSs, including pairwise interactions between TFBSs. We performed a 5-fold cross-validation (CV) to avoid the risk of over-fitting. This procedure was repeated 100 times with different random seeds. An ensemble of 100 regression coefficients (RCs) for a TFBS provides statistical information of the estimated regulatory activity of the TFBS. We conducted statistical tests using these ensembles. We applied this workflow to 4 regression models to predict the expression levels of each of the Class A and Class B genes in LT- or ST-HSCs.

We attempted to characterize promoter architectures by testing the different TGASs mentioned above, rather than by comparing our approach with other modeling methods. This was because of the difficulty of implementing existing methods using our inputs and analyzing their results. We also aimed to determine regulatory activities by analyzing 4 models. We characterized the context-dependent function of regulators that activated and repressed the transcription of distinct genes depending on the cellular context [44, 45]. Thus, our approach provided a detailed picture of the regulatory modes involved in context-dependent gene expression.

Inference of higher-order TF interactions

The co-occupancy of a promoter by multiple TFs contributes synergistically to transcriptional regulation. We considered this when calculating TGAS by performing probabilistic LLM [31] coupled with iterative random sampling. The input matrix used for LLM, i.e., n promoters in rows $\times$ m TFs in columns, comprised binary values that represent the existence of TFBSs for m TFs in n promoters. Using this matrix, LLM was employed to infer the conditional (in)dependency of TF occurrences, i.e., TF–TF interactions in higher-order conditional distributions. It should be noted that LLM cannot determine whether an interaction is competitive or cooperative.

The huge number of TFs means that LLM is not adequate to compute them all; therefore, we performed random sampling with 10 arbitrary selected TFs, which means that an inferred TF–TF interaction was observed constantly in the 2^8 state combinations of 8 TFs. This sampling procedure was terminated if an outcome had no effect during 10^5 runs. We calculated the interaction probability Pr for all possible TF pairs using this iterative sampling procedure (Materials and Methods). After repeating the sampling procedure 1,367,639 times for Class A and 1,406,837 times for Class B, we retrieved 50 and 77 interactions ($Pr = 1.0$) from Class A and Class B, respectively (Tables S12 and S13).

Performance of regression models

Overall, Pearson's correlation coefficient Rs in learning and testing of 5-fold CVs showed a slight over-fitting in the range of < 0.2 (Figure S5), which was acceptable in our sense. One of the reasons for this over-fitting was the unbalanced numbers between testing genes and TFBSs; e.g., 72 Class A genes (a subset of 5-fold CV) were tested by a model with over 100 predictors that were trained by the remaining Class A genes. This implies that the constructing of a model to generalize the gene regulation for an