Materials and Methods

Animals

All experimental protocols were reviewed and approved by the Institutional Animal Care and Use Committee of Tokyo Women's Medical University (approval ID: 13-99-2-B). Mice were purchased from Sankyo Labo Service.

Cell collection

$CD34^-$KSL ($c\text{-}kit^+Sca1^+Lin^-$) LT-HSCs or $CD34^+$KSL ST-HSCs were sorted, as described previously [36]. In brief, we isolated bone marrow cells from 8- to 10-week-old C57BL/6 mice and stained them with antibodies for CD34 (RAM34, eBiosciences, San Diego, CA), Sca-1 (E13-161.7, BD Biosciences Pharmingen, San Jose, CA), c-kit (2B8, BD Biosciences Pharmingen), and a lineage marker (Lineage Detection Kit, Miltenyi Biotec Inc., Bergisch Gladbach, Germany). Subsequently, we analyzed the stained cells using a MoFlo XDP cell sorter system (Beckman Coulter, Fullerton, CA).

RNA sequencing and real-time PCR

After obtaining total RNA extracts from 5000 LT- or ST-HSCs using Isogen (Nippon Gene, Tokyo, Japan) in triplicate, we synthesized cDNA using a SMARTer Pico cDNA amplification kit (Clonetech, Mountain View, CA) and amplified them with 20 cycles of PCR. Using the standard protocols for the SOLiD system, we sequenced the amplified cDNA using a SOLiD sequencer (Life Technologies, Carlsbad, CA), as described previously [36]. In the RT-PCR assay, total RNA was obtained from the sorted cells and cDNA was synthesized as described above. We performed RT-PCR using a TaqMan Gene Expression Assay (Life Technologies) for the genes indicated with the BioMark HD system (Fludigm, South San Francisco, CA).

Read mapping and quantification

We used the TopHat (v1.4.1)/Cufflinks (v.2.0.2) pipeline [33] with the sequenced reads (quality score, >15). The pipeline was coupled to Bowtie (v.0.12.7) [62]. We employed the recursive read mapping method, as described previously [32]. In brief, we applied TopHat by truncating the $3'$ ends of unmapped reads and by realigning the reads using more stringent parameters. We set the parameters empirically, which were used sequentially, as the read length, "-initial-read-mismatches", "-segment-mismatches", and "-segment-length": (50, 3, 2, 25), (46, 3, 2, 23), (42, 3, 2, 21), (38, 2, 0, 19), and (34, 2, 0, 17).

The pipeline, which quantifies RNA abundance as fragments per kilobase of exon per million mapped reads (FPKM), mapped sequenced reads to the mouse genome (mm9), and then assembled transcripts with uniquely mapped reads (uni-reads) for each replicate. We used Cuffcompare to merge all the transcript assemblies; 14,728 and 14,128 RefSeq-annotated genes in LT- and ST-HSCs, respectively. Using the merged transcript assembly, we performed Cuffdiff, which calculates FPKMs across all replicates and detects DEGs via two-group t-tests coupled to a Benjamini–Hochberg false discovery rate (FDR) procedure. We further used transcripts that satisfied the following conditions: successful deconvolution, FDR of <0.05, complete match of intron chain, and FPKM of >0.001. The mouse genome and RefSeq annotation were downloaded from http://genome.ucsc.edu/.

Long-term competitive reconstitution assay

We cultured $CD34^-$KSL HSCs derived from C57BL/6-Ly5.1 congenic mice for 5 days with or without 20μM GW1929 (Sigma-Aldrich, St. Louis, MO) in S-Clone SF-03 medium (Sanko-Junyaku Co., Tokyo, Japan) supplemented with 0.5% bovine serum albumin (Sigma, St. Louis, MO) and 50 ng/ml mouse stem cell factor and 50 ng/ml mouse TPO (all from R&D systems, Minneapolis, MN). Next, we performed a long-term competitive reconstitution assay by transplanting cultured cells with 5×10^5 whole bone marrow competitor cells derived from C57BL/6-Ly5.2 Wt mice into lethally irradiated (9.5 Gy) C57BL/6-Ly5.2 Wt mice.

Log-linear model (LLM)

Suppose that we consider binary-stated (absence or presence) TFs $\{A, B, C\}$. The observed counts fall into 2^3-dimensional contingency table by cross-classifying the TF states. The full model (FM), which contains all the possible interactions, gives the logarithms of probabilities as follows:

$$\log p_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}, \quad (1)$$

where i, j and k are the state indices of $\{A, B, C\}$, λs are unknown parameters, λ_{ij}^{AB}, λ_{ik}^{AC} and λ_{jk}^{BC} represent the interaction effects among the indexed variables. If an instance of A is independent of B, FM can be reduced to a reduced model (RM) with respect to the hierarchy [31], which is given as follows:

$$\log p_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}. \quad (2)$$

This model can be reformulated as

$$p_{ijk} = (p_{i+k} \cdot p_{+jk})/p_{++k}, \quad (3)$$

where "+" denotes the summation over the corresponding index. This formula is equivalent to $\Pr(A=i,B=j|C=k) = \Pr(A=i|C=k.)\Pr(B=j|C=k)$, which means that A and B are independent in the conditional distribution given C ($A \perp\!\!\!\perp B|C$).

To find the most parsimonious RM, we remove an interaction term from the current model and measure two p-values for the asymptotic χ^2 test of a likelihood ratio G^2 statistic [31]. The p-values comprise p_FM, which is the difference between FM and RM, and p_RM, which is the difference between the current model and RM. We accept a removal if it yields the largest p_RM (≥ 0.01), and we terminate if any removal test yields <0.01 for either p_RM or p_FM.

Iterative random sampling for LLM

A large number of TFs can easily yield a vast dimensional contingency table. To find a near optimal parsimonious model even in such higher-dimensional space, we designed an iterative sampling scheme that allowed us to calculate interaction probability Pr as follows.

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ is an undirected graph, where $\mathcal{V}$ is a finite set of vertices (TFs) and $\mathcal{E}$ is a set of edges, which represent the interactions between vertex pairs. The scheme is as follows.

1. 1. $\mathcal{S} = \{s_1, \ldots, s_k\}$, a nonredundant combination of TFs, is selected randomly from all TFs ($k=10$ in the present study).
2. For all possible vertex pairs (s_i, s_j), the trial number $ntry_{ij}$ of an edge between s_i and s_j is counted (i.e., FM of k variables).

3. LLM infers the best model $\mathcal{G}' = (\mathcal{S}, \mathcal{E}')$, where $\mathcal{E}'$ is a set of edges that represents TF–TF interactions.
4. For all possible vertex pairs (s_i, s_j), if an edge in $\mathcal{E}'$ links a pair (s_i, s_j), the observed edge frequency $nobs_{ij}$ for this pair is counted.
5. For all possible vertex pairs (s_i, s_j), the interaction probability Pr for a pair (s_i, s_j) is updated using $nobs_{ij}/ntry_{ij}$.
6. If $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{E}$ is a set of edges ($Pr = 1.0$), is not changed with a large number of samplings ($= 100{,}000$); therefore, this procedure is terminated. Otherwise, steps 1–5 are repeated.

Linear regression model

We used a multivariate regression model

$$\log Y_i = \sum_j w_j X_{ij} + e_i, \tag{4}$$

$$X_{ij} = \sum_k x_k, \tag{5}$$

where Y_i is the expression of gene i, X_{ij} is TGAS of the jth TFBS in the promoter region of gene i, w_j is RC of the jth TFBS, and e_i is the error term. TGAS is the sum of scores x_k, where k represents the position of the jth TFBS in promoter i. We tested the following forms of x_k.

- I: matrix similarity s of TFBS j scored using MATCH[43] ($x_k = s_k$).
- II: TGAS I modified by a location-dependent weight L,

$$x_k = s_k \times L_k. \tag{6}$$

- III: TGAS II weighted by the expression fold change (F) of TFs,

$$x_k = s_k \times L_k \times \sum_{k'} F_{k'}, \tag{7}$$

where k' is the index of TFs binding to TFBS j. If FPKM for TF is ≤ 3, we use $F = 1$.

- IV: the same as TGAS III, but we removed TFBSs where none of the TFs had FPKM of > 3.
- V: TGAS III weighted using both Fs of interactive TFs and the interaction probability Pr estimated by LLM,

$$x_k = s_k \times L_k \times \left(\sum_{k'} F_{k'} + I_{k'}\right) \tag{8}$$

$$I_{k'} = \sum_{l'=1}^{k'} \sum_{j'>l'}^{k'} F_{l'} F_{j'} Pr_{l'j'}. \tag{9}$$

We used a published method to calculate L [40]. First, we calculated the distribution of TFBS j in bins ($= 500$ bp) of promoter regions and created a histogram H_{real}. Next, we randomized the positions of TFBS j and created a histogram H_{rand}. L for the kth TFBS j is given by the following:

$$L_k = \begin{cases} 0, & \text{if } H_{real}(m) < H_{rand}(m) \\ \dfrac{H_{real}(m) - H_{rand}(m)}{H_{real}(m)}, & \text{if } H_{real}(m) \geq H_{rand}(m), \end{cases} \tag{10}$$

where m represents the index of bin that corresponds to the position of the kth TFBS j. This location-dependent weight takes a value between 0 and 1, where a higher weight implies nonrandom occurrence.

Stepwise selection of the regression model

We built a regression model with the explanatory variable X and then reduced the model using AIC. Let the reduced model be Y' with X'. $X - X' = \{x_1, x_2, \ldots\}$ is the variables removed on the basis of AIC. V is the set of all pairwise terms of $x_i x_j$ ($i \neq j$). We searched any elements of V that improve Pearson's correlation coefficient r of 5-fold CV on testing datasets.

1. Randomly select v_i ($\in V$) and add it to X', which yields X''.
2. Perform 5-fold CV with X'' and calculate the averaged r on testing datasets.
3. If the r has been improved, update X'' to X'.
4. Repeat step 1–3 until all v_i have been tested.
5. Calculate Pearson's correlation coefficient R between observed and predicted FPKMs of all genes by using the final model.

We run this procedure 100 times using different random seeds. The final R is referred to as a model quality in this study.

Bioinformatics analysis

We obtained array-based gene expression profiles [8,9] from BloodExpress [63], RNA-seq data for megakaryocyte/erythroid precursors and megakaryocytes from http://genome.ucsc.edu/encode/, and RNA-seq data for MII oocytes and two-cell embryos from DDBJ DRA001066. The public RNA-seq datasets were analyzed using the pipeline mentioned above. To search putative TFBSs and TFs in TRANSFAC professional (released in January 2013) [39], we prepared $\pm$5kb DNA sequences from transcription start sites (TSSs) annotated in RefSeq (http://www.ncbi.nlm.nih.gov/refseq/), and applied the MATCH tool in the minimize false-positive mode [43].

To analyze the enriched GO terms, we used the DAVID Bioinformatics Resources [35]. Significant terms detected by DAVID (EASE score, a modified Fisher's exact p-value, < 0.01) were grouped into representative ancestor terms in the dataset GO Slim2 using CateGOrizer [64]. We used the R programming language (http://www.r-project.org/) for regression modeling and to perform statistical tests. Although all p-values were adjusted by Bonferroni correction (Tables S6 and S8–S11), we used uncorrected p-values throughout this study to avoid too conservative interpretation that would reduce biologically meaningful findings.

Data access

The RNA-seq data generated in this study have been deposited in the DDBJ (DNA Data Bank of Japan) Sequence Read Archive (DRA) under accession number DRA001213. The online version of LLM is available at http://dbtmee.hgc.jp/tools/.

Supporting Information

Figure S1 Correlation analysis of gene expression levels measured using RNA-seq assays. (A) Reproducibility based on triplicate analyses of LT- and ST-HSCs. (B) Comparison of the gene expression correlations in the present study to those reported by Karlsson et al. [15], who purified HSCs using $CD48^-$, $CD150^+$, $CD34^-$, $CD9^{high}$ KSL for LT-HSCs and $CD48^-$, $CD150^+$, $CD9^{low}$ KSL for ST-HSCs.
(EPS)

Figure S2 Contribution of higher-order TF interaction scores estimated by LLM. (A) Statistical differences of 2 regression coefficient (RC) ensembles of a TFBS found commonly by TGAS III and V (two-sample *t*-test). (B) Distribution of the TF interaction score I_k in Equation 9.
(EPS)

Figure S3 Box plots of RCs estimated by 100 iterations of regression modeling with TGAS V. Pos and Neg represent the positive (red) and negative (blue) mean values of RCs (red line), respectively.
(EPS)

Figure S4 Subnetworks involved in ST-HSC regulation. Although the majority of TF-coding genes found in ST-HSCs (Figure 4A) were not differentially expressed, 26 differentially expressed TFs that putatively bind to 21 TFBSs were present among DEGs (Class A and Class B).
(EPS)

Figure S5 Propensity of the TFBS activities inferred from public RNA-seq datasets. We applied our method to public RNA-seq datasets related to sequential cell development (A) and lineage commitment (C). Our procedure evaluates the averaged R of 5-fold CV on testing datasets (blue line). If a model improved R in testing, the model was accepted and its R value between the observed and predicted gene expression of all genes was measured (red line). (B) Of 147 TFBSs ($p<0.05$), 67 TFBSs (Class A; upregulated in Oo) and 80 TFBSs (Class B; upregulated in 2C) exhibited significant gains and losses of activity ($p<0.001$). In addition, 73% (49/67) of Class A and 52.5% (42/80) of Class B genes exhibited no changes in the effects of their TFBS activities between cells, i.e., positive (negative) in Oo was still positive (negative) in 2C. We found that 16% (8/49) of Class A and 83% (35/42) of Class B genes had increased activities in 2C compared with Oo. (D) Among 150 TFBSs ($p<0.05$), 98 TFBSs (Class A, upregulated in MEP) and 114 TFBSs (Class B, upregulated in Mk) exhibited significant gains and losses of activity ($p<0.001$). We also found that 83% (81/98) of Class A and 76% (87/114) of Class B genes exhibited no changes in the effects of their TFBS activities. All of the TFBSs in both classes exhibited increases in the strengths of their activities in Mk compared with MEP. R, Pearson's correlation coefficient; Oo, MII oocytes; 2C, 2-cell embryo; MEP, megakaryocyte/erythroid precursor; Mk, megakaryocyte.
(EPS)

Table S1 RNA-seq mapping statistics.
(XLSX)

Table S2 Differentially expressed cell-surface molecules.
(XLSX)

Table S3 Differentially expressed transcription factors.
(XLSX)

Table S4 Transcription factors categorized into Class C.
(XLSX)

Table S5 Low expressed transcription factors (Class D).
(XLSX)

Table S6 Average regression coefficient of 142 TFBSs.
(XLSX)

Table S7 Classification of MkE, GM, and Lymphoid-associated genes.
(XLSX)

Table S8 TFBSs significantly different in the regression coefficient between LT- and ST-HSCs (Class A).
(XLSX)

Table S9 TFBSs significantly different in the regression coefficient between LT- and ST-HSCs (Class B).
(XLSX)

Table S10 Enriched GO terms in Class A.
(XLSX)

Table S11 Enriched GO terms in Class B.
(XLSX)

Table S12 Result of log-linear model in Class A.
(XLSX)

Table S13 Result of log-linear model in Class B.
(XLSX)

Acknowledgments

We thank Drs S. Mitani and T. Furukawa for their great help with RNA sequencing. Computational resources were provided by the supercomputer system at Human Genome Center, the Institute of Medical Science, the University of Tokyo.

Author Contributions

Conceived and designed the experiments: SJP KN. Performed the experiments: TU YS MY. Analyzed the data: SJP MSA. Contributed reagents/materials/analysis tools: TU SJP. Wrote the paper: SJP.

References

1. Hoang T (2004) The origin of hematopoietic cell type diversity. Oncogene 23: 7188–98.
2. Forsberg EC, Bhattacharya D, Weissman IL (2006) Hematopoietic stem cells: expression profiling and beyond. Stem Cell Rev 2: 23–30.
3. Sanjuan-Pla A, Macaulay IC, Jensen CT, Woll PS, Luis TC, et al. (2013) Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature 502: 232–6.
4. Yamamoto R, Morita Y, Oochara J, Hamanaka S, Onodera M, et al. (2013) Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. Cell 154: 1112–26.
5. Passegue E, Wagers AJ, Giuriato S, Anderson WC, Weissman IL (2005) Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. J Exp Med 202: 1599–611.
6. Forsberg EC, Prohaska SS, Katzman S, Heffner GC, Stuart JM, et al. (2005) Differential expression of novel potential regulators in hematopoietic stem cells. PLoS Genet 1: e28.
7. Zhong JF, Zhao Y, Sutton S, Su A, Zhan Y, et al. (2005) Gene expression profile of murine longterm reconstituting vs. short-term reconstituting hematopoietic stem cells. Proc Natl Acad Sci U S A 102: 2448–53.
8. Mansson R, Hultquist A, Luc S, Yang L, Anderson K, et al. (2007) Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity 26: 407–19.
9. Ficara F, Murphy MJ, Lin M, Cleary ML (2008) Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. Cell Stem Cell 2: 484–96.

10. Kent DG, Copley MR, Benz C, Wohrer S, Dykstra BJ, et al. (2009) Prospective isolation and molecular characterization of hematopoietic stem cells with durable self-renewal potential. Blood 113: 6342–50.
11. Chotinantakul K, Leeanansaksiri W (2012) Hematopoietic stem cell development, niches, and signaling pathways. Bone Marrow Res 2012: 270425.
12. Kunisaki Y, Bruns I, Scheiermann C, Ahmed J, Pinho S, et al. (2013) Arteriolar niches maintain haematopoietic stem cell quiescence. Nature 502: 637–43.
13. Liu P, Barb J, Woodhouse K, Taylor JG, Munson PJ, et al. (2011) Transcriptome profiling and sequencing of differentiated human hematopoietic stem cells reveal lineage-specific expression and alternative splicing of genes. Physiol Genomics 43: 1117–34.
14. Lu R, Neff NF, Quake SR, Weissman IL (2011) Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. Nat Biotechnol 29: 928–33.
15. Karlsson G, Rorby E, Pina C, Soneji S, Reckzeh K, et al. (2013) The tetraspanin cd9 affords high-purity capture of all murine hematopoietic stem cells. Cell Rep 4: 642–8.
16. Weishaupt H, Sigvardsson M, Attema JL (2010) Epigenetic chromatin states uniquely define the developmental plasticity of murine hematopoietic stem cells. Blood 115: 247–56.
17. Wilson NK, Foster SD, Wang X, Knezevic K, Schutte J, et al. (2010) Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell 7: 532–44.
18. Bissels U, Bosio A, Wagner W (2012) Micrornas are shaping the hematopoietic landscape. Haematologica 97: 160–7.
19. Whichard ZL, Sarkar CA, Kimmel M, Corey SJ (2010) Hematopoiesis and its disorders: a systems biology approach. Blood 115: 2339–47.
20. Bonzanni N, Garg A, Feenstra KA, Schutte J, Kinston S, et al. (2013) Hard-wired heterogeneity in blood stem cells revealed using a dynamic regulatory network model. Bioinformatics 29: i80–8.
21. Hannah R, Joshi A, Wilson NK, Kinston S, Gottgens B (2011) A compendium of genome-wide hematopoietic transcription factor maps supports the identification of gene regulatory control mechanisms. Exp Hematol 39: 531–41.
22. Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, et al. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell 144: 296–309.
23. Will B, Vogler TO, Bartholdy B, Garrett-Bakelman F, Mayer J, et al. (2013) Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. Nat Immunol 14: 437–45.
24. Mirshekar-Syahkal B, Haak E, Kimber GM, van Leusden K, Harvey K, et al. (2013) Dlk1 is a negative regulator of emerging hematopoietic stem and progenitor cells. Haematologica 98: 163–71.
25. Gazit R, Garrison B, Rao T, Shay T, Costello J, et al. (2013) Transcriptome analysis identifies regulators of hematopoietic stem and progenitor cells. Stem Cell Reports 1: 266–280.
26. Osawa M, Hanada K, Hamada H, Nakauchi H (1996) Long-term lymphohematopoietic reconstitution by a single cd34-low/negative hematopoietic stem cell. Science 273: 242–5.
27. Ema H, Morita Y, Yamazaki S, Matsubara A, Seita J, et al. (2006) Adult mouse hematopoietic stem cells: purification and single-cell assays. Nat Protoc 1: 2979–87.
28. Bussemaker HJ, Foat BC, Ward LD (2007) Predictive modeling of genome-wide mrna expression: from modules to molecules. Annu Rev Biophys Biomol Struct 36: 329–47.
29. Park SJ, Nakai K (2011) A regression analysis of gene expression in es cells reveals two gene classes that are significantly different in epigenetic patterns. BMC Bioinformatics 12 Suppl 1: S50.
30. Irie T, Park SJ, Yamashita R, Seki M, Yada T, et al. (2011) Predicting promoter activities of primary human dna sequences. Nucleic Acids Res 39: e75.
31. Lauritzen S (1996) Graphical Models. New York: Oxford University Press.
32. Park SJ, Komata M, Inoue F, Yamada K, Nakai K, et al. (2013) Inferring the choreography of parental genomes during fertilization from ultralarge-scale whole-transcriptome analysis. Genes Dev 27: 2736–48.
33. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of rna-seq experiments with tophat and cufflinks. Nat Protoc 7: 562–78.
34. Benz C, Copley MR, Kent DG, Wohrer S, Cortes A, et al. (2012) Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. Cell Stem Cell 10: 273–83.
35. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using david bioinformatics resources. Nat Protoc 4: 44–57.
36. Umemoto T, Yamato M, Ishihara J, Shiratsuchi Y, Utsumi M, et al. (2012) Integrin-alphavbeta3 regulates thrombopoietin-mediated maintenance of hematopoietic stem cells. Blood 119: 83–94.
37. Domen J, Cheshier SH,Weissman IL (2000) The role of apoptosis in the regulation of hematopoietic stem cells: Overexpression of bcl-2 increases both their number and repopulation potential. J Exp Med 191: 253–64.
38. Peng C, Chen Y, Shan Y, Zhang H, Guo Z, et al. (2012) Lsk derived lsk- cells have a high apoptotic rate related to survival regulation of hematopoietic and leukemic stem cells. PLoS One 7: e38614.
39. 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: 316–9.
40. Chen X, Xu H, Yuan P, Fang F, Huss M, et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133: 1106–17.
41. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, et al. (2008) Genome-scale dna methylation maps of pluripotent and differentiated cells. Nature 454: 766–70.
42. Moignard V, Macaulay IC, Swiers G, Buettner F, Schutte J, et al. (2013) Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat Cell Biol 15: 363–72.
43. Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, et al. (2003) Match: A tool for searching transcription factor binding sites in dna sequences. Nucleic Acids Res 31: 3576–9.
44. Boyle P, Despres C (2010) Dual-function transcription factors and their entourage: unique and unifying themes governing two pathogenesis-related genes. Plant Signal Behav 5: 629–34.
45. Whitfield TW, Wang J, Collins PJ, Partridge EC, Aldred SF, et al. (2012) Functional analysis of transcription factor binding sites in human promoters. Genome Biol 13: R50.
46. Diffner E, Beck D, Gudgin E, Thoms JA, Knezevic K, et al. (2013) Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia. Blood 121: 2289–300.
47. Chute JP, Ross JR, McDonnell DP (2010) Minireview: Nuclear receptors, hematopoiesis, and stem cells. Mol Endocrinol 24: 1–10.
48. Henke BR, Blanchard SG, Brackeen MF, Brown KK, Cobb JE, et al. (1998) N-(2-benzoylphenyl)-l-tyrosine ppargamma agonists. 1. discovery of a novel series of potent antihyperglycemic and antihyperlipidemic agents. J Med Chem 41: 5020–36.
49. Brown KK, Henke BR, Blanchard SG, Cobb JE, Mook R, et al. (1999) A novel n-aryl tyrosine activator of peroxisome proliferator-activated receptor-gamma reverses the diabetic phenotype of the zucker diabetic fatty rat. Diabetes 48: 1415–24.
50. Liebermann DA, Gregory B, Hoffman B (1998) Ap-1 (fos/jun) transcription factors in hematopoietic differentiation and apoptosis. Int J Oncol 12: 685–700.
51. Shen LJ, Chen FY, Zhang Y, Cao LF, Kuang Y, et al. (2013) Mycn transgenic zebrafish model with the characterization of acute myeloid leukemia and altered hematopoiesis. PLoS One 8: e59070.
52. Yu S, Jing X, Colgan JD, Zhao DM, Xue HH (2012) Targeting tetramer-forming gabpbeta isoforms impairs self-renewal of hematopoietic and leukemic stem cells. Cell Stem Cell 11: 207–19.
53. Ghiaur G, Yegnasubramanian S, Perkins B, Gucwa JL, Gerber JM, et al. (2013) Regulation of human hematopoietic stem cell self-renewal by the microenvironment's control of retinoic acid signaling. Proc Natl Acad Sci U S A.
54. Okada S, Fukuda T, Inada K, Tokuhisa T (1999) Prolonged expression of c-fos suppresses cell cycle entry of dormant hematopoietic stem cells. Blood 93: 816–25.
55. Wozniak RJ, Keles S, Lugus JJ, Young KH, Boyer ME, et al. (2008) Molecular hallmarks of endogenous chromatin complexes containing master regulators of hematopoiesis. Mol Cell Biol 28: 6681–94.
56. Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, et al. (2013) Charting a dynamic dna methylation landscape of the human genome. Nature 500: 477–81.
57. 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: 4225–31.
58. Zhang P, Behre G, Pan J, Iwama A, Wara-Aswapati N, et al. (1999) Negative cross-talk between hematopoietic regulators: Gata proteins repress pu.1. Proc Natl Acad Sci U S A 96: 8705–10.
59. Peixoto A, Monteiro M, Rocha B, Veiga-Fernandes H (2004) Quantification of multiple gene expression in individual cells. Genome Res 14: 1938–47.
60. Warren L, Bryder D, Weissman IL, Quake SR (2006) Transcription factor profiling in individual hematopoietic progenitors by digital rt-pcr. Proc Natl Acad Sci U S A 103: 17807–12.
61. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, et al. (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498: 236–40.
62. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short dna sequences to the human genome. Genome Biol 10: R25.
63. Miranda-Saavedra D, De S, Trotter MW, Teichmann SA, Gottgens B (2009) Bloodexpress: a database of gene expression in mouse haematopoiesis. Nucleic Acids Res 37: D873–9.
64. Hu ZL, Bao J, Reecy J (2008) Categorizer: A web-based program to batch analyze gene ontology classification categories. Online J Bioinform 9: 108–12.

STEM CELLS

TISSUE-SPECIFIC STEM CELLS

Mesenchymal Stem Cells Cancel Azoxymethane-Induced Tumor Initiation

MASANAO NASUNO,[a] YOSHIAKI ARIMURA,[a] KANNA NAGAISHI,[b] HIROYUKI ISSHIKI,[a] KEI ONODERA,[a] SUGURU NAKAGAKI,[a] SHUHEI WATANABE,[a] MASASHI IDOGAWA,[c] KENTARO YAMASHITA,[a] YASUYOSHI NAISHIRO,[d] YASUSHI ADACHI,[a] HIROMU SUZUKI,[e] MINEKO FUJIMIYA,[c] KOHZOH IMAI,[f] YASUHISA SHINOMURA[a]

Key Words. Mesenchymal stem cells • Azoxymethane • Tumor initiation • Colorectal cancer • Chemoprevention

[a]Department of Gastroenterology, Rheumatology, and Clinical Immunology, [b]Department of Anatomy, [c]Department of Medical Genome Sciences, Research Institute for Frontier Medicine, [d]Department of Educational Development, and [e]Department of Molecular Biology, Sapporo Medical University, Chuo-ku, Sapporo, Japan; [f]Center for Antibody and Vaccine Therapy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Correspondence: Yoshiaki Arimura, M.D., Ph.D., Department of Gastroenterology, Rheumatology, and Clinical Immunology, Sapporo Medical University, S-1, W-16, Chuo-ku, Sapporo 060-8543, Japan. Telephone: +81-11-611-2111, Ext. 3211, 3213; Fax: +81-11-611-2282; e-mail: arimura@sapmed.ac.jp

Received August 8, 2013; accepted for publication October 12, 2013; first published online in STEM CELLS *EXPRESS* November 8, 2013.

1066-5099/2014/$30.00/0

http://dx.doi.org/10.1002/stem.1594

ABSTRACT

The role of mesenchymal stem cells (MSCs) in tumorigenesis remains controversial. Therefore, our goal was to determine whether exogenous MSCs possess intrinsic antineoplastic or proneoplastic properties in azoxymethane (AOM)-induced carcinogenesis. Three in vivo models were studied: an AOM/dextran sulfate sodium colitis-associated carcinoma model, an aberrant crypt foci model, and a model to assess the acute apoptotic response of a genotoxic carcinogen (AARGC). We also performed in vitro coculture experiments. As a result, we found that MSCs partially canceled AOM-induced tumor initiation but not tumor promotion. Moreover, MSCs inhibited the AARGC in colonic epithelial cells because of the removal of O^6-methylguanine (O^6MeG) adducts through O^6MeG-DNA methyltransferase activation. Furthermore, MSCs broadly affected the cell-cycle machinery, potentially leading to G1 arrest in vivo. Coculture of IEC-6 rat intestinal cells with MSCs not only arrested the cell cycle at the G1 phase, but also induced apoptosis. The anti-carcinogenetic properties of MSCs in vitro required transforming growth factor (TGF)-β signaling because such properties were completely abrogated by absorption of TGF-β under indirect coculture conditions. MSCs inhibited AOM-induced tumor initiation by preventing the initiating cells from sustaining DNA insults and subsequently inducing G1 arrest in the initiated cells that escaped from the AARGC. Furthermore, tumor initiation perturbed by MSCs might potentially dysregulate WNT and TGF-β-Smad signaling pathways in subsequent tumorigenesis. Obtaining a better understanding of MSC functions in colon carcinogenesis is essential before commencing the broader clinical application of promising MSC-based therapies for cancer-prone patients with inflammatory bowel disease. STEM CELLS *2014;32:913–925*

INTRODUCTION

Stem and progenitor cells are well-known direct cellular targets of genetic alterations in human carcinogenesis [1–3]. Previous studies have altered our perception of stromal cells from being innocent bystanders to active promoters in the neoplastic process [4–6]. Carcinoma formation accompanied by well-orchestrated desmoplastic reactions [7] closely resembles wound healing and scar formation, and entails the constant availability of growth factors, cytokines, and matrix-remodeling proteins that render the tumor site as a "wound that never heals" [8]. Recent studies have shown that bone marrow-derived mesenchymal stem cells (MSCs) are recruited in large numbers to the stroma of developing tumors [9, 10].

However, the role of MSCs in tumorigenesis remains an intensely debated topic. Khakoo et al. demonstrated that intravenously injected human MSCs possess intrinsic antineoplastic properties in an in vivo model of Kaposi's sarcoma by inhibition of Akt activity in a cell–cell contact-dependent manner [11]. In contrast, Karnoub et al. demonstrated that MSCs within the stroma of the tumor microenvironment facilitate metastatic spread via paracrine signals of C—C motif chemokine 5 that is secreted de novo by MSCs [12].

A meta-analysis of chemoprevention studies has suggested that azoxymethane (AOM)-based rodent models of carcinogenesis are valuable for prediction of chemopreventive efficacy in humans, which is better than that of other models [13, 14]. The prominent advantages of the AOM/dextran sulfate sodium (DSS) colitis-associated carcinogenesis model are that factors influencing tumor initiation [15, 16] should result in changes of the average tumor number per animal, whereas differences of the average tumor size typically provide evidence for factors involved in tumor

progression [17, 18]. AOM is a genotoxic agent that initiates cancer by alkylation of DNA, in which O6-methylguanine (O^6MeG) is a highly cytotoxic, apoptotic, mutagenic, recombinogenic, and clastogenic DNA adduct [19]. Conversely, dextran sulfate sodium (DSS) is not genotoxic, but rather a proinflammatory tumor promoter [20]. Studies in rodents have revealed that AOM-induced tumors resemble human colorectal cancer at the molecular level, which displays dysregulation of the canonical WNT signaling pathway, similar target genes [21–24], and mutation of *K-ras* [25].

Our goal was to determine whether MSCs possess intrinsic antineoplastic or proneoplastic properties in an AOM-induced tumorigenesis model. Because MSCs are prime candidates for use in cell- and gene-based therapies [26, 27], this essential information must be obtained before implementing the broader clinical application of MSC therapies.

Materials and Methods

For detailed Materials and Methods, refer to Supporting Information.

Experimental Animals

Animal studies were performed under the supervision of the Committee for Animal Research of Sapporo Medical University in accordance with protocols approved by the Institutional Animal Care and Use Committee. All animals were maintained according to the guidelines of the Committee for Animal Research of Sapporo Medical University. Lewis rats were purchased from Charles River Laboratories Japan (Yokohama, Japan; http://www.crj.co.jp), and SD-TG (CAG-EGFP) rats were purchased from Japan SLC, Inc. (Hamamatsu, Japan; http://www.jslc.co.jp/) [28]. All rats were aged 6 weeks and were female unless indicated otherwise and were housed under pathogen-free conditions and received autoclaved food and water ad libitum.

Cell Lines and Culture Conditions

Bone marrow cells were harvested by inserting a needle into the shaft of the femur or tibia and flushing it with 30 ml α-modified Eagle's medium (αMEM) containing 20% fetal bovine serum (FBS). To harvest rat MSCs [29], the cell suspensions were passed through a 70-μm nylon filter (Becton Dickinson, Franklin Lakes, NJ; http://www.bd.com/us/) and plated in 75-cm^2 flasks. Cells were grown in alpha-modified Eagle's medium (MEM) containing 20% FBS at 37°C and 5% CO_2. After 3 days, the medium was replaced with fresh αMEM containing 10% FBS, and the adherent cells were grown to 80% confluence to obtain passage 0. In accordance with the International Society for Cellular Therapy criteria [30], cells between passages 3 and 5 were used for subsequent experiments [26]. To harvest rat hematopoietic stem cells (HSCs) [31], CD90.1 $(Thy1.1)^+$ cells were magnetically labeled with CD90.1 MicroBeads (Miltenyi Biotec GmbH, Gladbach, Germany; https://www.miltenyibiotec.com/en/) for 15 minutes. Then, the cell suspension was loaded onto a MACS column that was placed in the magnetic field of a MACS separator. The magnetically labeled $CD90.1^+$ cells were retained and then eluted as the positively selected cell fraction. Detailed protocols and data sheets are available at http://www.miltenyibiotec.com. To prepare conditioned medium from rat MSCs (MSC-CM), MSCs (4×10^5 cells per 150-mm culture dish) were seeded and cultured to confluency. Then, the medium was changed to serum-free Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA; lifetechnologies.com), and the rat MSCs were cultured for a further 48 hours. The conditioned medium was collected, centrifuged at 300*g* for 5 minutes, filtered using a 0.22-μm syringe filter, and then stored at −80°C until use.

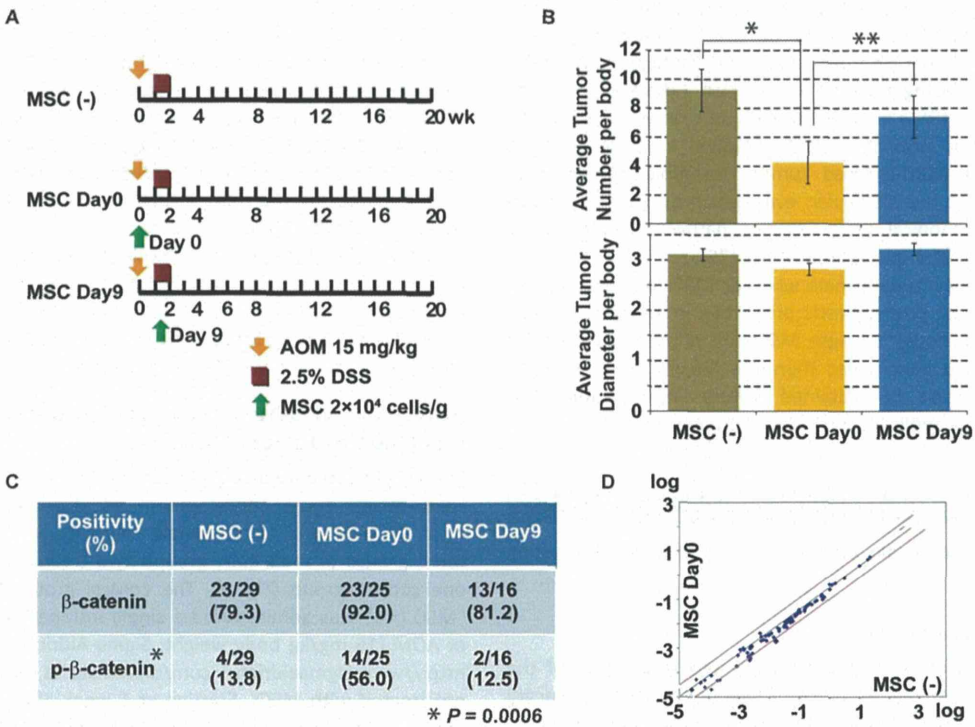
IEC-6 cells obtained from the American Type Culture Collection (Manassas, VA; http://www.atcc.org/) and 3Y1 rat fibroblasts (3Y1-B Clone 1–6) [32] obtained from JCRB Cell Bank (Saito, Japan; http://cellbank.nibio.go.jp/) were maintained as recommended by the depositors.

AOM/DSS Colitis-Associated Carcinoma Model and Evaluation of Tumor Growth

We adopted the two-stage colon tumor model that mimics colitis-driven tumor development as described by Tanaka et al. [33]. A total of 39 female Lew rats were divided into three groups ($n = 13$ each group): two treatment groups and one control group (Fig. 1). The control group designated as "MSC (−)" was administered a single intraperitoneal injection of AOM (15 mg/kg body weight; Sigma-Aldrich, St. Louis, MO; http://www.sigmaaldrich.com/united-states.html) and was not treated with MSCs. Starting at 1 week after injection, the animals received 2.5% DSS (molecular weight 9,000–20,000; Sigma-Aldrich) in drinking water for 7 days, and then received no further treatment for 18 weeks. The group designated as "MSC Day0" was intravenously administered 2×10^4 MSCs/g body weight on day 0 when AOM was injected, and the group designated as "MSC Day9" was administered MSCs on day 9 following the AOM administration corresponding to day 2 after receiving DSS in drinking water. Evaluation of tumor growth in the AOM/DSS colitis-associated carcinoma model is described in Supporting Information.

Analysis of Aberrant Crypts

Fifteen female Lew rats were divided into three groups ($n = 5$ each group): two treatment groups and one control group (Fig. 3). The control group designated as MSC (−) was administered two separate intraperitoneal injections of AOM (15 mg/kg body weight) at 7 days apart (Days 0 and 7) and was not treated with MSCs. No further treatment was performed for 3 weeks. The group designated as "MSC Day1" was intravenously administered 2×10^4 MSCs/g body weight on day 1, which was 1-day before AOM was first injected, and the group designated as "MSC Day8" was administered MSCs on day 8 (1-day after the second AOM administration). The rats were killed at the end of the study (week 4) by transcardial perfusion with 4% paraformaldehyde in phosphate buffered saline (PBS). The colon was removed and cut open from the anus to the cecum along the longitudinal axis. We defined the rectum as the segment at 2 cm proximal to the anus, and divided the entire colon into three segments each measuring approximately 7 cm in length: the distal colon including the rectum and the middle and proximal colon. The colon was spread flat between sheets of filter paper and fixed in 10% buffered formalin. Then, the colon tissues were stained with 0.2% methylene blue in saline according to the procedure described by Bird [34] to observe aberrant crypts (ACs).



Positivity (%)	MSC (-)	MSC Day0	MSC Day9
β-catenin	23/29 (79.3)	23/25 (92.0)	13/16 (81.2)
p-β-catenin*	4/29 (13.8)	14/25 (56.0)	2/16 (12.5)

* $P = 0.0006$

Figure 1. Effects of mesenchymal stem cells (MSCs) in the azoxymethane (AOM)/dextran sulfate sodium (DSS) colitis-associated carcinoma model. AOM/DSS model rats ($n = 39$) were classified into three groups ($n = 13$ each group) according to the timing of MSC administration in carcinogenetic phases: MSC-untreated control (MSC [−]), tumor initiation (day 0; MSC Day0), and tumor promotion phases (day 9; MSC Day9), as shown in panel A. Brown, orange, and blue bars represent data obtained from MSC (−), MSC Day0, and MSC Day9 groups, respectively. The upper panel of B shows the average tumor number per rat (a total of 271 tumors developed), and the lower panel indicates the average tumor size. *, $p = .023$; **, $p = .008$. The upper row of the panel C shows that β-catenin protein expression ($n = 70$ tumor tissues) was regarded as positive when the band intensity was stronger than that in the normal colon in Western blot analyses. In all, 23 of 29 (79.3%), 23 of 25 (92.0%), and 13 of 16 (81.2%) tumors examined were positive for β-catenin expression in MSC (−), MSC Day0, and MSC Day9 groups, respectively. The lower row indicates positivity of phospho-β-catenin at Ser33/37/Thr41 in Western blot analyses in the three groups. The contingency table analysis reached statistical significance (*, $p = .0006$). The WNT signal pathway polymerase chain reaction (PCR) array of representative samples pairs obtained from MSC (−) and MSC Day0 groups was depicted in panel D. PCR and Western blot analyses were performed in triplicate unless specified otherwise. Abbreviations: AOM, azoxymethane; DSS, dextran sulfate sodium; MSC, mesenchymal stem cell.

Based on the McLellan and Bird [35] definition, aberrant crypts (ACs) were defined as those that (i) were larger than normal crypts, (ii) had an increased pericryptal space that separated them from normal crypts, (iii) had a thicker layer of epithelial cells that often stained darkly, and (iv) generally had oval rather than circular openings. The number of aberrant crypt foci (ACF) per colon, the number of ACs in each focus, and the location of each focus were determined by stereomicroscopy (Olympus, Tokyo, Japan; http://www.olympus.co.jp/jp/) at 40× magnification. The mucosa of the distal segments was scraped off and subjected to Western blot analysis.

Effects of MSCs on the AARGC

The five experimental groups included the MSC-untreated control group administered PBS, the group designated as "MSC" administered MSCs, groups designated as "hematopoietic stem cells (HSCs)" and "3Y1" administered HSCs or 3Y1 rat fibroblasts, respectively, and the group designated as "MSC-conditioned medium (MSC-CM)" treated with MSC-CM at 24 hours before AOM administration. Each group consisted of five rats and received a single subcutaneous injection of AOM (15 mg/kg body weight) at 09:00 hours. The rats were then killed by CO_2-induced narcosis at the indicated intervals from 8 to 48 hours. The entire colon was removed immediately, cut open, and flushed with ice-cold saline. Segments measuring 2 cm were taken from the rectal end of the distal portion. These segments were immediately fixed in 10% paraformaldehyde overnight at room temperature and then embedded in paraffin. The mucosa on the remaining segments was scraped off and subjected to subsequent analyses.

β-Catenin Nucleotide Sequence

Sequencing was performed by the classical Sanger method [36].

WNT Signaling Pathway PCR Array Analysis

A rat WNT signaling pathway RT^2 profiler polymerase chain reaction (PCR) array (SuperArray Bioscience, Frederick, MD; http://www.sabiosciences.com/) was performed according to the manufacturer's instructions.

Analysis of the Cell Cycle and Apoptosis

The cell cycle was assessed by flow cytometry and Ki67 immunohistochemistry. The apoptotic cell fraction was

determined by terminal deoxynucleotidyl transferase-mediated dUTP nick-end labeling (TUNEL) reactions.

Immunofluorescence of the DNA Adduct of O^6MeG

The level of DNA alkylation was analyzed by immunofluorescence of distal colon sections using an antibody specific for the DNA adduct of O^6MeG. Frozen sections (4 μm) were prepared, rehydrated, and incubated with 3% hydrogen peroxide in 50% ethanol for 15 minutes at room temperature. Antigen retrieval was carried out with a Retrieval Solution (DAKO, Carpinteria, CA; http://www.dako.com/) for 10 minutes at 105°C in an autoclave. RNase treatment (20 μl RNase A at 10 mg/ml and 5 μl RNase T at 10 U/ml in 1,000 μl PBS, pH 7.5) was carried out for 1 hour at 37°C, and then stopped by treatment with a 140-mmol/l sodium chloride (NaCl) solution for 5 minutes at 4°C. DNA unwinding was achieved by alkali treatment (1,500 μl of 70 mmol/l NaOH/140 mmol/l NaCl and 1,000 μl of absolute methanol) before applying Protein Block (DAKO) for 10 minutes at room temperature. The sections were then incubated at room temperature overnight with an anti-O^6MeG monoclonal antibody (clone EM 2–3; Squarix Biotechnology, Marl, Germany; http://www.squarix.de/) diluted at 1:1,000 in PBS. The next day, the sections were washed in PBS three times for 5 minutes each before applying an Alexa Fluor 594-labeled secondary anti-mouse IgG. Sections were counterstained with DAPI, dehydrated, and cover slipped for observation under a LSM 510 META. The primary antibody was omitted for the negative control.

MTT Assay

Cell proliferation was measured by a MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide] dye reduction assay [37].

RNA Isolation and qPCR Analysis

Quantitative Real-Time PCR (qPCR) was performed using TaqMan Universal PCR Master Mix (Applied Biosystems, Carlsbad, CA; lifetechnologies.com) for 40 cycles at 95°C for 15 seconds and 60°C for 1 minutes by standard methods.

Western Blot Analysis

Western blot analysis was performed according to standard methods.

MSC and IEC-6 Cell Coculture Experiments

Green fluorescent protein (GFP)-labeled rat MSCs were cocultured with the rat small intestinal cell line IEC-6. MSCs and IEC-6 cells cultured separately were included as controls. The cells were cocultured in RPMI 1,640 (Sigma-Aldrich) supplemented with 10% FBS, penicillin (100 U/ml), and streptomycin (100 μg/ml) at 37°C with 5% CO_2. A total of 1×10^6 MSCs were seeded per 100-mm dish for the MSC control and harvesting MSC-CM, and 1×10^6 IEC-6 cells per dish for the cell line control. For direct coculture, 1×10^6 MSCs per dish were preseeded for 2–3 hours. Then, 1×10^6 IEC-6 cells per dish were added and cultured for up to 72 hours with/without AOM treatment. For indirect coculture, 1,250 IEC-6 cells per well were preseeded in the lower chamber of a Transwell (0.4-μm pores, 48 wells; Corning, Tewksbury, MA; http://www.corning.com) for 2–3 hours. Then, the same number of MSCs per well were added to the upper chamber of the Transwell and cultured for up to 72 hours with/without AOM, methylazoxymethanol (MAM; Wako Pure Chemical Industries, Tokyo, Japan; http://www.wako-chem. co.jp/), or O^6-benzylguanine (O^6BG; Sigma-Aldrich), which binds irreversibly to and inhibits the DNA repair enzyme O^6MeG-DNA methyltransferase (Mgmt). Because AOM is metabolized into the active metabolite methylazoxymethanol (MAM) by Cyp2e1, we confirmed whether *Cyp2e1* was expressed in IEC-6 cells [38]. For absorption of transforming growth factor (TGF)-β, 1.0 μg/ml anti-transforming growth factor (TGF)-β neutralizing antibody (Clone # 9016; R&D Systems, Minneapolis, MN; http://www.rndsystems.com/) was added to the direct coculture condition.

Statistical Analysis

To compare means between two groups, parametric and nonparametric analyses were performed using the unpaired Student's *t*-test and the Mann–Whitney *U*-test, respectively. Categorical variables were compared using the chi-square test, exact *p* value based on Pearson's statistic, or the Monte Carlo method. For multiple comparisons, we applied analysis of variance (ANOVA), especially in serial assessments, and two-way repeated measures (mixed between-within subjects) analysis of variance (ANOVA) followed by the Bonferroni test [39]. A difference was considered significant at $p < .05$ in all two-tailed tests. The SPSS Statistics 17.0 software package (SPSS Inc., Chicago, IL; http://www.spss.com/) was used for all statistical analyses.

Results

MSCs Reduce the Tumor Number but Not the Tumor Size in AOM/DSS Colitis-Associated Tumorigenesis

We explored whether MSCs affected tumor initiation or promotion in the AOM/DSS model and the associated mechanism (Fig. 1A). The average tumor number per rat was significantly decreased by up to half of the expected level when MSCs were simultaneously injected with AOM (MSC Day0 group; $p = .008$ compared with the untreated control and $p = .023$ compared with the MSC Day9 group; upper panel in Fig. 1B and Supporting Information Fig. S1A). In contrast, the average tumor diameter was not significantly different among the groups as shown in the lower panel of Figure 1B. In this model, factors that influence tumor initiation should result in changes of the average tumor number per animal, whereas differences in average tumor sizes typically provide evidence of factors involved in tumor progression [18]. Therefore, these results suggest that MSC partially cancel AOM/DSS-induced tumor initiation.

MSCs Profoundly Affect the Mutational Spectra During the Tumor Initiation Phase

As shown in Figure 1C, the tumor β-catenin expression analyzed by Western blotting was not significantly different among MSC (−), MSC Day0, and MSC Day9 groups. However, β-catenin was more frequently phosphorylated (56%, 14 of 25 tumors) in the MSC Day0 group than that in the MSC (−)

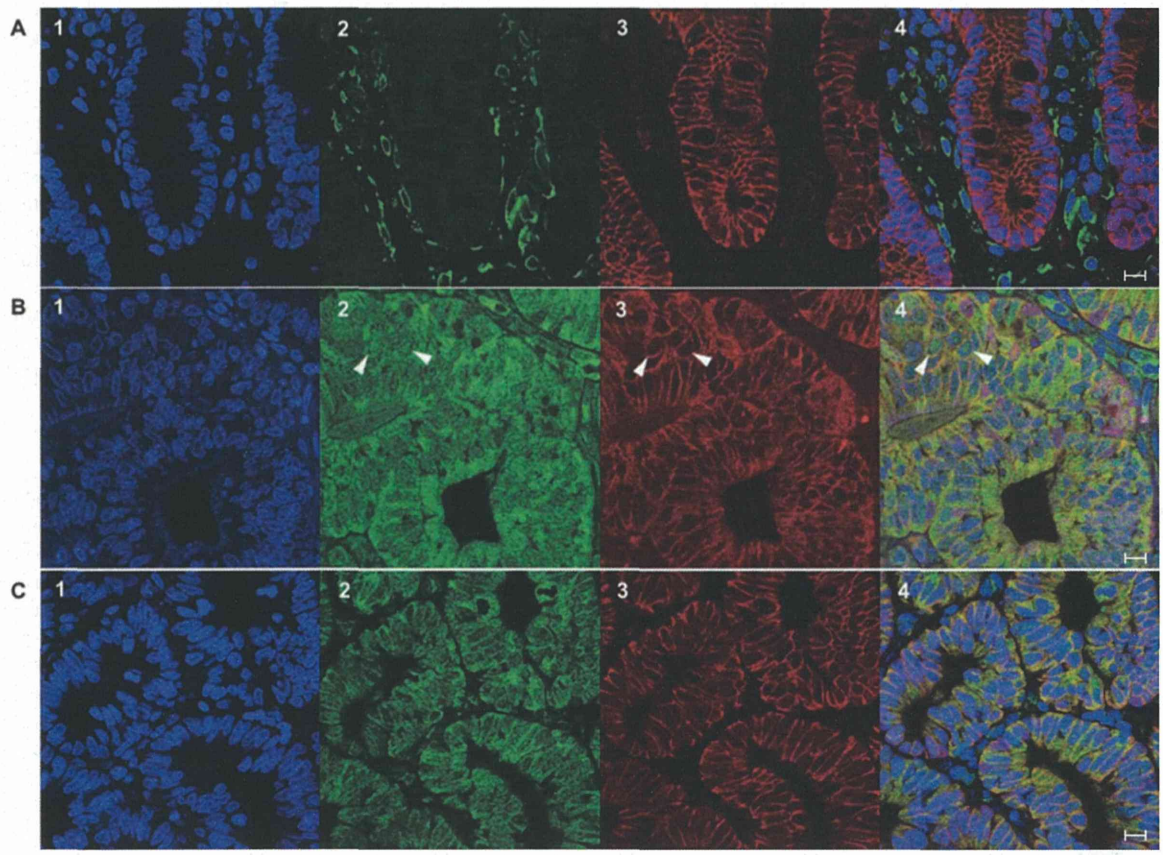


Figure 2. Double immunofluorescence staining of Smad2 and β-catenin. From the left to right, panels show nuclear counterstaining with DAPI (1), visualization of Smad2 (2) and β-catenin (3) using Alexa Fluor 488- and 594-labeled secondary antibodies, respectively, and merged images (4). Panel A shows normal colonic tissues of azoxymethane-untreated healthy rats, panel B shows tumor tissues from the mesenchymal stem cell (MSC) (−) group, and panel C shows tumor tissues from the MSC Day0 group. White arrowheads in panel B indicate nuclear staining of Smad2. Bar scales at the right lower corner in each panel indicated 10 μm.

group (13.8%, 4 of 29 tumors) and the MSC Day9 group (12.5%, 2 of 16, $p = .0006$). Furthermore, the mutation spectrum of *β-catenin* was quite different between MSC (−) and MSC Day0 groups. The codon 34 missense mutation (GGA–GAA) was the most frequent (11 of 25, 44%) in the MSC (−) group. In addition to the above mutation (5 of 15, 33.3%), the codon 32 missense mutation (GAT–AAT) was also frequently mutated (5 of 15, 33.3%) in the MSC Day0 group. Four (66.7%) of the six mutated regions detected in the MSC Day0 group and 12 (85.7%) of the 14 regions detected in the MSC (−) group appeared to be unique (Supporting Information Table S3). Of the 89 genes, 79 (88.8%) genes of the WNT signal pathway examined in the WNT PCR array were downregulated in MSC Day0 tumors compared with that in MSC (−) tumors (Fig. 1D).

Receptor-regulated Smad representing canonical TGF-β-Smad signaling was confined to the cell membrane of the laminar propria stromal cells, and β-catenin was expressed only on the cell membranes of crypt epithelial cells in the normal colon (Fig. 2A). Phospho-Smad2 expression representing activated TGF-β signaling was not significantly different between MSC (−) and MSC Day0 groups in Western blot analyses (data not shown). The total Smad2 protein level was upregulated locally in the cytoplasm and partially in the nuclei (white arrowheads in Fig. 2B). β-Catenin was slightly upregulated in membranous and cytoplasmic staining of the colon carcinomas in MSC (−) group rats (Fig. 2B). In contrast, both Smad2 and β-catenin were localized only on the membrane of colon carcinoma cells in MSC Day0 group rats (Fig. 2C). MSC engraftment was observed in tumors established at 20 weeks after AOM administration (data not shown). Therefore, these results suggest that MSCs profoundly affect the mutational spectra during the tumor initiation phase, leading to distinct WNT and canonical TGF-β-Smad signaling in subsequent tumorigenesis and even in the established tumors.

MSCs Reduce the Formation of ACF

Next, we determined whether MSCs affect aberrant crypt foci (ACF) formation and the timing of MSC administration during tumor initiation induced by AOM for the most efficacious chemoprevention (Fig. 3A). The average ACF density was significantly lower in both pre-AOM (MSC Day−1; $p = 4.7E\text{-}4$) and post-AOM (MSC Day8; $p = .001$) treatment groups than that in the MSC (−) control group (Fig. 3B; Supporting Information Fig. S1B). As depicted in Figure 3C, ACF were formed more frequently in the distal colon than in the proximal colon as reported previously [34]. ACF formation was suppressed in both the distal and middle colons of both treatment groups (MSC Day1 and MSC Day8) with no significant differences between the two treatment groups. The multiplicity of ACs

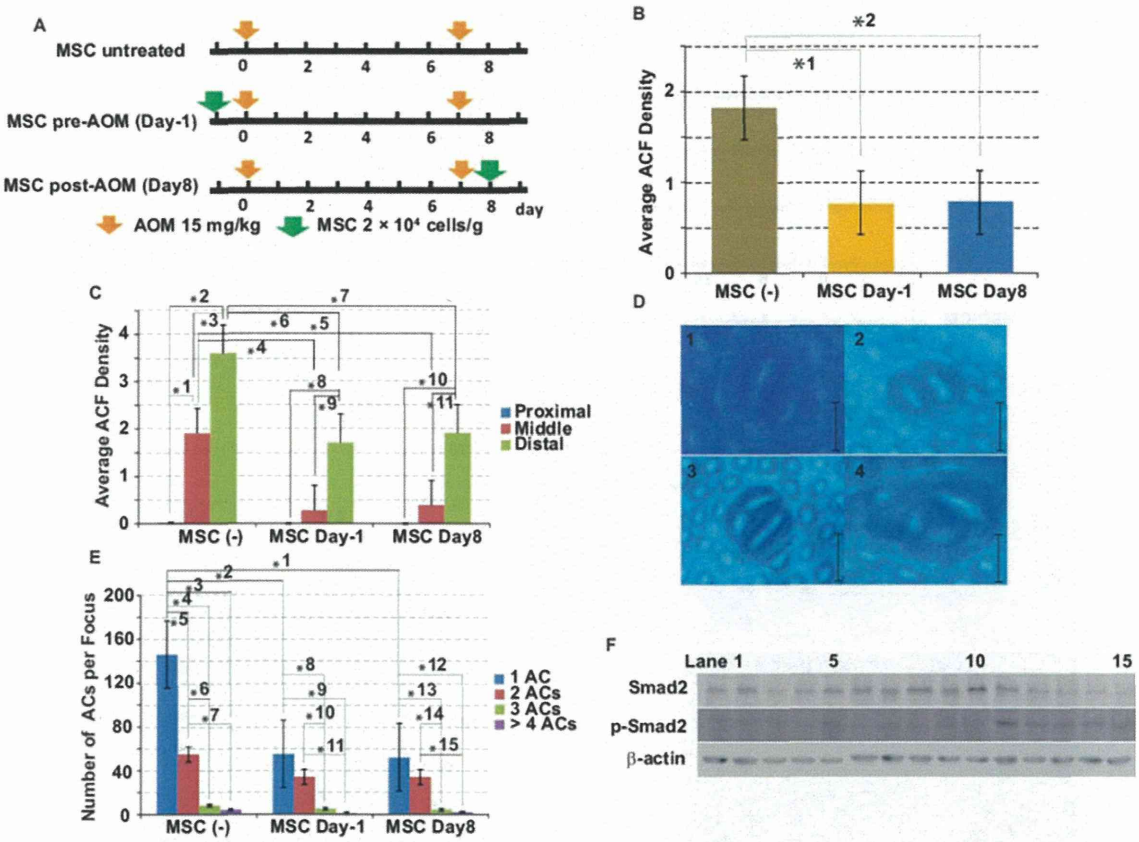


Figure 3. Analysis of the aberrant crypt foci (ACF) model. ACF model rats ($n = 15$) were classified into three groups ($n = 5$ each group) according to the timing of mesenchymal stem cell (MSC) administration either before or after administration of two separate of azoxymethane (AOM) on Days 0 and 7: MSC (−), MSC pre-AOM (Day–1), and MSC post-AOM (day 8) groups (A). A total of 400 ACF developed: 213, 95, and 92 in MSC (−), MSC Day–1, and MSC Day8 groups, respectively. The average ACF density is shown in panel B. *1, $p = 4.7E\text{-}4$; *2, $p = .001$. The average ACF density in the proximal, middle, and distal colon is shown in panel C. *1, $p = .02$; *2, $p = 1.6E\text{-}4$; *3, $p = .037$; *4, $p = .017$; *5, $p = .022$; *6, $p = .002$; *7, $p = .004$; *8, $p = 4.0E\text{-}5$; *9, $p = 2.2E\text{-}4$; *10, $p = 2.1E\text{-}7$; *11, $p = 1.9E\text{-}6$. Representative ACs, one to more than four ACs per focus, are shown in panels D1–4, respectively. Scale bars: 50 μm. The average density of ACs per focus, one to more than four ACs, is shown in panel E. *1, $p = .009$; *2, $p = .011$; *3, $p = 5.2E\text{-}5$; *4, $p = 7.3E\text{-}5$; *5, $p = .005$; *6, $p = 7.3E\text{-}5$; *7, $p = 5.3E\text{-}5$; *8, $p = 1.7E\text{-}4$; *9, $p = 7.9E\text{-}5$; *10, $p = .021$; *11, $p = .005$; *12, $p = 4.0E\text{-}6$; *13, $p = 6.4E\text{-}6$; *14, $p = .001$; *15, $p = .001$. Panel F shows the analysis of transforming growth factor-β signaling by Western blotting of Smad2 and phospho-Smad2. Lanes 1–5, 6–10, and 11–15 show data for MSC (−) control MSC Day–1, and MSC Day8 groups, respectively. Abbreviations: AC, aberrant crypt; ACF, aberrant crypt foci; AOM, azoxymethane; MSC, mesenchymal stem cell.

per focus, as shown in Figure 3D1 (one AC) to D4 (>four ACs), was reciprocally related to the frequency of the ACs. Among these ACs, one AC/focus was significantly reduced by MSC treatment, although there was no significant difference between the treatment groups (Fig. 3E). Canonical TGF-β-Smad signaling represented by phospho-Smad2 was activated in all colonic epithelia (5 of 5) of the MSC Day8 group and not in the colonic epithelia of MSC (−) or MSC Day–1 groups (Fig. 3F). Surprisingly, these results suggest that MSCs elicit a chemopreventive effect on formation of the prototype ACF (one AC/focus), both as a preventive measure in the preinitiation phase (MSC Day1) and a treatment measure in the post-initiation phase (MSC Day8). However, it is unknown why the canonical TGF-β-Smad signals were distinctly activated by the two measures.

MSCs Suppress the AARGC

To obtain a further mechanistic insight into the antineoplastic properties of MSCs in AOM-induced carcinogenesis, we examined whether MSCs affect the acute apoptotic response of a genotoxic carcinogen (AARGC) in vivo (Fig. 4A) [40]. The acute apoptotic response of a genotoxic carcinogen (AARGC) peaked at 8 hours after AOM administration, which was significantly suppressed only in the MSC-treated group compared with that in the MSC-untreated control, HSC, 3Y1, and MSC-CM groups (Fig. 4B, 4C). The Ki-67 labeling index of the colonic epithelia was significantly decreased at 24 and 48 hours only in the MSC group compared with that in the other groups (Fig. 4D, 4E). Western blot analyses revealed suppression of Akt in the AARGC (8 hour) observed in control groups was significantly activated in MSC groups, whereas activation of p38 in the AARGC observed in control groups was slightly suppressed in MSC groups (Fig. 4F). Consequently, these results suggest that AARGC suppression is a specific property of MSCs, which does not involve other cell types or humoral factors produced by MSCs. Because the AARGC is accepted as one of the in vivo mechanisms that suppress tumorigenicity, further experiments are necessity to explain why MSCs possess chemopreventive and AARGC suppression effects.