

Possible roles for the hominoid-specific DSCR4 gene in human cells

Morteza M. Saber^{1,2,3,4}, Marziyeh Karimiavargani^{3,5}, Takanori Uzawa³,
Nilmini Hettiarachchi¹, Michiaki Hamada^{4,6},
Yoshihiro Ito³ and Naruya Saitou^{1,2,7,8*}

¹Population Genetics Laboratory, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan

²Department of Biological Sciences, Graduate School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan

³Nano Medical Engineering Laboratory, RIKEN, Wako, Saitama 351-0198, Japan

⁴Department of Electrical Engineering and Bioscience, Faculty of Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan

⁵Graduate School of Science and Engineering, Saitama University, Saitama, Saitama 388-0825, Japan

⁶Computational Bio Big-Data Open Innovation Laboratory (CBBDOIL), National Institute of Advanced Industrial Science and Technology (AIST), Shinjuku-ku, Tokyo 169-8555, Japan

⁷Department of Genetics, School of Life Science, Graduate University for Advanced Studies, Mishima, Shizuoka 411-8540, Japan

⁸Faculty of Medicine, University of the Ryukyus, Nishihara-cho, Okinawa 903-0215, Japan

(Received 24 February 2020, accepted 28 August 2020; J-STAGE Advance published date: 24 March 2021)

Down syndrome in humans is caused by trisomy of chromosome 21. DSCR4 (Down syndrome critical region 4) is a *de novo*-originated protein-coding gene present only in human chromosome 21 and its homologous chromosomes in apes. Despite being located in a medically critical genomic region and an abundance of evidence indicating its functionality, the roles of DSCR4 in human cells are unknown. We used a bioinformatic approach to infer the biological importance and cellular roles of this gene. Our analysis indicates that DSCR4 is likely involved in the regulation of interconnected biological pathways related to cell migration, coagulation and the immune system. We also showed that these predicted biological functions are consistent with tissue-specific expression of DSCR4 in migratory immune system leukocyte cells and neural crest cells (NCCs) that shape facial morphology in the human embryo. The immune system and NCCs are known to be affected in Down syndrome individuals, who suffer from DSCR4 misregulation, which further supports our findings. Providing evidence for the critical roles of DSCR4 in human cells, our findings establish the basis for further experimental investigations that will be necessary to confirm the roles of DSCR4 in the etiology of Down syndrome.

Key words: Down syndrome, DSCR4, orphan gene, cell migration, human evolution

INTRODUCTION

Down syndrome is a human genetic disease, with the high incidence of 1 out of 700 live births making this disorder the leading genetic cause of mental retarda-

tion and congenital heart disease (Hassold and Jacobs, 1984). The major phenotypes characterizing Down syndrome are variable in age of onset, frequency and severity, and include immune deficiency, heart disease, dysmorphology of facial characters and underlying skeleton, Hirschsprung's disease, alterations of brain structure, early onset of Alzheimer's pathology and increased risk of leukemia. Mental retardation, characterized by certain behavioral and cognitive deficits, is also a com-

Edited by Yoko Satta

* Corresponding author. saitounr@nig.ac.jp

DOI: <https://doi.org/10.1266/ggs.20-00012>

mon feature in Down syndrome individuals (Cohen, 2002; Epstein, 2002; Wiseman et al., 2009; Lana-Elola et al., 2011).

Down syndrome is caused by the inheritance of an extra copy of chromosome 21q. Since the completion of the human genome project, it is known that this region spans ~33.5 million base pairs (Mb) of DNA and contains ~300 genes (Hattori et al., 2000). So far, mouse transgenic models have been the main tool for the analysis of gene–phenotype correlation in Down syndrome (Kahlem et al., 2004; Lyle et al., 2004). Studies of trisomic mouse models containing an extra copy of the mouse genome segment that is orthologous to the q arm of chromosome 21 imply that all trisomic genes are up-regulated by ~50% across multiple tissues (Kahlem et al., 2004; Lyle et al., 2004). Although the underlying mechanism accounting for how the relatively small increment in transcription results in any of the commonly observed phenotypes in Down syndrome is yet unknown, the correspondence of the shortest genomic region shared by Down syndrome individuals with the same Down syndrome characteristics has led to the hypothesis that a critical chromosomal region called Down syndrome critical region (DSCR) contains a dosage-sensitive gene or set of genes whose misregulation is responsible for the emergence of Down syndrome features (Jiang et al., 2015). The previously defined DSCR spans ~5.4 Mb ranging from a proximal boundary between DS21S17 and D21S55 to a distal boundary between MX1 and BCE1 (Korbel et al., 2009). This region harbors about 33 genes that are conserved between human and mouse (Nikolaienko et al., 2005) and has been associated with multiple Down syndrome characteristics such as craniofacial abnormalities, joint hyperlaxity and mental retardation (Lana-Elola et al., 2011). Another study of partial trisomy 21 further defined a 34-kb region as critical to the Down syndrome phenotype (Pelleri et al., 2016).

Down syndrome studies using multiple transgenic mouse models in which the DSCR-orthologous region is overexpressed or underexpressed, in particular chromosome 16 segmental trisomies such as Ts1Cje, Ts65Dn and Ts1Rhr and segmental monosomies including Ms1Rhr, have so far provided valuable information in constructing the phenotypic map of DSCR (Olson et al., 2007). However, investigations on functional annotation of orthologous genes located in the DSCR revealed that, first, the mouse orthologous genes are actually dispersed within the genome and do not have the same synteny as their human counterparts (Gardiner et al., 2003), and, second, not all the genes present in the human DSCR region have orthologs in mouse, which illustrates the inability of mouse models of Down syndrome to accurately simulate the human Down syndrome condition (Saber et al., 2016).

Down syndrome critical region 4 (DSCR4) is a novel *de novo*-originated protein-coding gene in the DSCR region,

and is present only in humans and apes (superfamily Hominoidea), without a known ortholog in non-hominoid organisms (Fig. 1) (Kumar et al., 2017). This gene was first discovered during the human chromosome 21 genome sequencing effort, and was found to have no homolog in the mouse genome (Saber et al., 2016). DSCR4, previously called DSCRB (Nakamura et al., 1997), shares a bidirectional promoter with the DSCR8 gene and is positively and negatively regulated by multiple transcription factors (Toyoda et al., 2002; Dunn et al., 2006; Asai et al., 2008). DSCR4 encodes an experimentally verified protein with 118 amino acids (Uhlén et al., 2015). The amino acid sequence of DSCR4 can potentially form protein secondary structure; however, this protein shows no discernible homology to any known or putative protein domains (Saber et al., 2016). This lack of homology is probably due to the *de novo* origin of DSCR4 and to the fact that two of the three coding exons of this gene are derived from retrotransposons (Saber et al., 2016). A detailed discussion of the transposon-derived origin of DSCR4 exons can be found in Saber et al. (2016). Hypomethylation of the DSCR4 promoter has been shown to be an epigenetic marker for Down syndrome (Du et al., 2011), which is another confounding factor in using mouse models to simulate the Down syndrome condition in humans. However, no study has so far investigated the functionality or subcellular interactions of DSCR4 inside human cells.

In this study, to systematically interrogate the cellular roles and gene regulatory networks involving DSCR4, we overexpressed wild-type DSCR4 in human non-cancerous cells and measured the consequences on the cell transcriptome by differential gene expression analysis. It

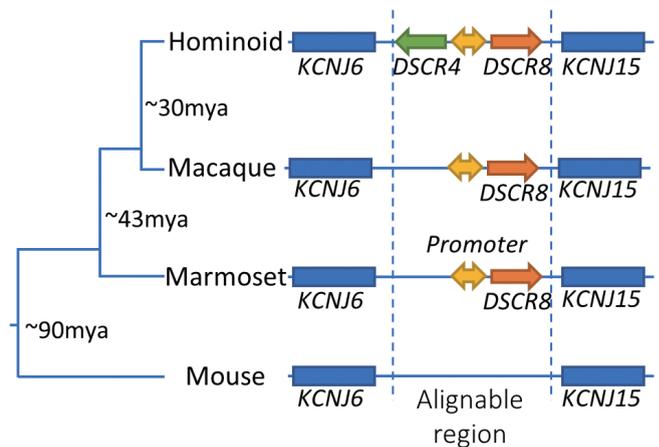


Fig. 1. *De novo* evolution of DSCR4 and its surrounding genomic region. Schematic depiction of the evolution of DSCR4 in common ancestors of hominoids (human, chimpanzee, gorilla, orangutan and gibbons). Evolution of DSCR8 and the bidirectional promoter that drives the expression of both genes are also depicted. Divergence time estimates were taken from Kumar et al. (2017). mya, million years ago.

has been shown that chromosome 21 transcripts are increased in proportion to gene dosage, i.e., 50% more abundant than normal in Down syndrome cells (Kahlem et al., 2004). Therefore, overexpression of DSCR4 is the optimal approach to simulate the Down syndrome condition. Through functional profiling, we investigated the biological processes affected by overexpression of DSCR4 to unravel the likely role of DSCR4 in cellular pathways and its contribution to the etiology of Down syndrome, along with how this gene might contribute to the unique characteristics displayed by the family Hominidae (Saber et al., 2016). Our analysis provides evidence for roles of DSCR4 in the regulation of cell migration and the immune system, which are also biological pathways affected in Down syndrome individuals.

MATERIALS AND METHODS

Cell culture The human papillomavirus 16 E6/E7-transformed cell line HS-27A, an immortalized non-cancerous cell type with epithelial morphology, was purchased from ATCC (CRL-2496) and cultured in RPMI-1640 medium (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Invitrogen) according to ATCC protocols.

Plasmid DNA vector and control vector construction The PTCN-DSCR4 expression vector (BC096162), which contains the full-length cDNA sequence of DSCR4 including upstream and downstream untranslated regions (UTRs), was purchased from transOMIC technologies (Supplementary Fig. S1A). No marker was added to DSCR4 cDNA, to ensure that the function of the extra copy of DSCR4 was not affected. Cytomegalovirus-derived promoter and enhancer sequences placed before DSCR4 cDNA in the vector ensure proper expression in eukaryotic cells. For production of a control plasmid (PTCN-control), we removed the DSCR4 cDNA sequence along with UTR elements from PTCN-DSCR4 using a GeneArt Seamless Cloning and Assembly kit (Supplementary Fig. S1B). With DSCR4 cDNA being the only difference between PTCN-DSCR4 and PTCN-control, we sought to minimize the confounding elements in our differential gene expression analysis.

Kill curve assay The first critical step for generating stably transfected cells is determining the optimal concentration of the selection reagent for selecting stable cell colonies. The purpose of the kill curve assay is to determine the minimum antibiotic concentration needed to kill all the cells over the course of one week. Since the optimal concentration is cell type-dependent, we performed this assay for HS-27A cells using the Cell Counting Kit-8 (Sigma). The HS-27A cells were treated with G418 (Geneticin; Sigma) in a concentration gradient between 0

(negative control) and 1,500 $\mu\text{g/ml}$. The number of cells in each well was counted after a week and the optimum concentration of G418 for treatment of HS-27A cells was determined as 1,400 $\mu\text{g/ml}$.

DNA transfection and stable cell line selection For transfection of HS-27A cells with PTCN-DSCR4 plasmid, first, the optimal concentration of OMNIfect transfection reagent was determined using a concentration gradient of OMNIfect and a pcDNA3 vector containing the GFP marker. In the next step, using the optimal concentration of OMNIfect (2 $\mu\text{l/ml}$), HS-27A cells were transfected with the PTCN-DSCR4 and PTCN-control vectors. Single cell colonies successfully transfected with DSCR4 were selected by incubation with RPMI-1640 containing 10% FBS and 1,400 $\mu\text{g/ml}$ G418 over approximately 21 days.

Gene expression analysis using microarray Total RNA was extracted from three PTCN-DSCR4-transfected and three PTCN-control-transfected HS-27A samples along with one normal HS-27A sample using a PureLink RNA Mini Kit (Invitrogen), and RNA quality was assessed and confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies). The isolated RNA was then carried through the Agilent preparation protocol and each sample was hybridized to one SurePrint G3 Human 8 \times 60K v3 GeneChip (Agilent Technologies). Raw data were processed and analyzed using GeneSpring GX software along with the RobiNA package (Lohse et al., 2012). Quality assessment was conducted by normalized expression value correlation analysis between the three sample groups (Supplementary Fig. S2). Normalized expression values (log base 2) for each chip were calculated using quantile normalization after background correction by the RMA method (Irizarry et al., 2003), which has been proven to measure expression levels reliably. The gene expression matrix was filtered to exclude probe sets that are either not positive or significant, not uniform, not above background or population outliers. Finally, saturated probe sets defined as "Class P" by GeneSpring GX software in PTCN-DSCR4-transfected HS-27A samples with at least a two-fold change in expression compared with PTCN-control-transfected HS-27A samples were determined and used for further gene set enrichment analysis. False discovery rate in all analysis was controlled by applying a Bonferroni adjustment.

Confirmation of DSCR4 gene perturbation using qPCR Quantitative real-time PCR (qPCR) was conducted by synthesizing cDNA from 1 mg DNA-free RNA using random hexamer primers and a PrimeScript First Strand cDNA Synthesis kit (Takara). qPCR was performed using the CFX96 Touch Real-Time PCR Detection System. Expression of DSCR4 and the GAPDH

housekeeping gene was measured using a SYBR Green-based gene expression assay (Applied Biosystems). The primers for qPCR analysis were designed to amplify intron–exon boundaries so as to avoid amplification of any potential DNA contamination. The GAPDH housekeeping gene was used as an endogenous control for normalization. As expected, no significant difference between the cycle threshold (Ct) values was observed for GAPDH across all the samples. Melt curve analysis also confirmed strictly specific amplifications in qPCR assays. The comparative Ct approach was employed with a CFX96 Touch Real-Time PCR machine to quantify DSCR4 expression levels, and confirmed that DSCR4 is overexpressed approximately 9-fold in PTCN-DSCR4-transfected cells compared with PTCN-control-transfected cells (Supplementary Fig. S3).

Functional profiling Differentially expressed probe set IDs in PTCN-DSCR4-transfected HS-27A samples were mapped to ENTREZ IDs and Kyoto Encyclopedia of Genes and Genomes (KEGG) IDs, and successfully mapped IDs were submitted to ClusterProfiler software (Yu et al., 2012). Gene Ontology (GO) terms and Gene Regulatory Networks (GRNs) that were significantly over-represented in the set of differentially expressed genes (DEGs) were identified using ClusterProfiler. Up- and down-regulated gene sets were analyzed independently using SurePrint G3 Human 8×60K v3 gene sets as a reference. GO terms and GRNs with significant enrichment (false discovery rate (FDR) < 0.05) were considered in the analysis. Biological processes enrichment analysis was conducted using data from the Reactome curated pathway database, and pathway enrichment analysis was performed by an enrichment test in the set of manually drawn pathway maps provided by KEGG (Kanehisa et al., 2016). FDR in all analysis was controlled by applying a Bonferroni adjustment.

To investigate the hypothesis that the DSCR4 gene has a unique expression pattern in human tissues and cells, RNA-Seq data of human tissues from the Roadmap Epigenome project (Roadmap Epigenomics Consortium et al., 2015) were retrieved and analyzed. The log-transformed average RPKM (reads per kilobase per million mapped reads) expression scores for the DSCR4 gene and the adjacent DSCR8 gene were calculated across human tissues along with cell lines for which genome-wide gene expression data were available. The YuGene tool was used to scale the gene expression data derived from different platforms for the purpose of performing integrated analyses (Lê Cao et al., 2014). This application also provides a normalized gene expression database using available microarray data across human cells. Using the provided database, YuGene was also employed for further evaluation of tissue-specific expression of the DSCR4 gene.

RESULTS

Computational analysis There is considerable biological evidence indicating functionality of the DSCR4 gene such as its possessing epigenetic marks for an active regulatory region within the promoter, and potential secondary structure in the encoded protein, as well as its harboring a fetal epigenetic marker for Down syndrome and being the binding site for multiple transcription factors. To verify the functionality of DSCR4 from an evolutionary point of view, we first performed derived allele frequency (DAF) analysis using 1000 Genomes Project data (The 1000 Genomes Project Consortium, 2015), which revealed that DSCR4 protein-coding sequences have higher rates of polymorphisms with low-frequency derived alleles compared with DSCR4 intronic sequences (Fig. 2A).

Intrinsic structural disorder (ISD) is the degree to which a given peptide folds as a stable three-dimensional protein, i.e., an ordered protein, versus a flexible and unstructured (disordered) entity. Natural protein sequences are more intrinsically disordered than translated random sequences (Yu et al., 2016; Wilson et al., 2017). Therefore, ISD can be used as a criterion to distinguish functional genes from random ORFs erroneously identified as novel genes. ISD analysis of DSCR4 using IUPred (Dosztányi et al., 2005) revealed that the protein encoded by this gene has a higher degree of ISD compared with its proximate but older protein-coding gene, DSCR8, and with the conserved housekeeping genes GAPDH and ACTB (Fig. 2B). Translated DSCR4 introns, on the other hand, have very low ISD scores.

Differential gene expression analysis We performed experimental gene perturbation analysis for *in vivo* interrogation of DSCR4 functionality. To enhance the reliability of the analysis, data cleaning was performed by discarding DEGs with unsaturated probe signals, which might represent false positives. We found that a total of 253 probes targeting protein-coding genes among 36,427 probes for protein-coding genes (< 0.5%) were differentially expressed between DSCR4-overexpressing HS-27A cells and control samples (Supplementary Fig. S4). Of these 253 DEGs, which were used for further investigations, 131 represent down-regulated and 122 represent up-regulated protein-coding genes (see Supplementary Table S1).

Predicting DSCR4 functions by Gene Ontology analysis After identification of DSCR4 overexpression-mediated DEGs, we performed GO analysis to identify interconnected cellular pathways that were enriched in the DEGs. Functional annotation tools were used to arrange genes in associated categories based on associated GO terms and participation in biological pathways.

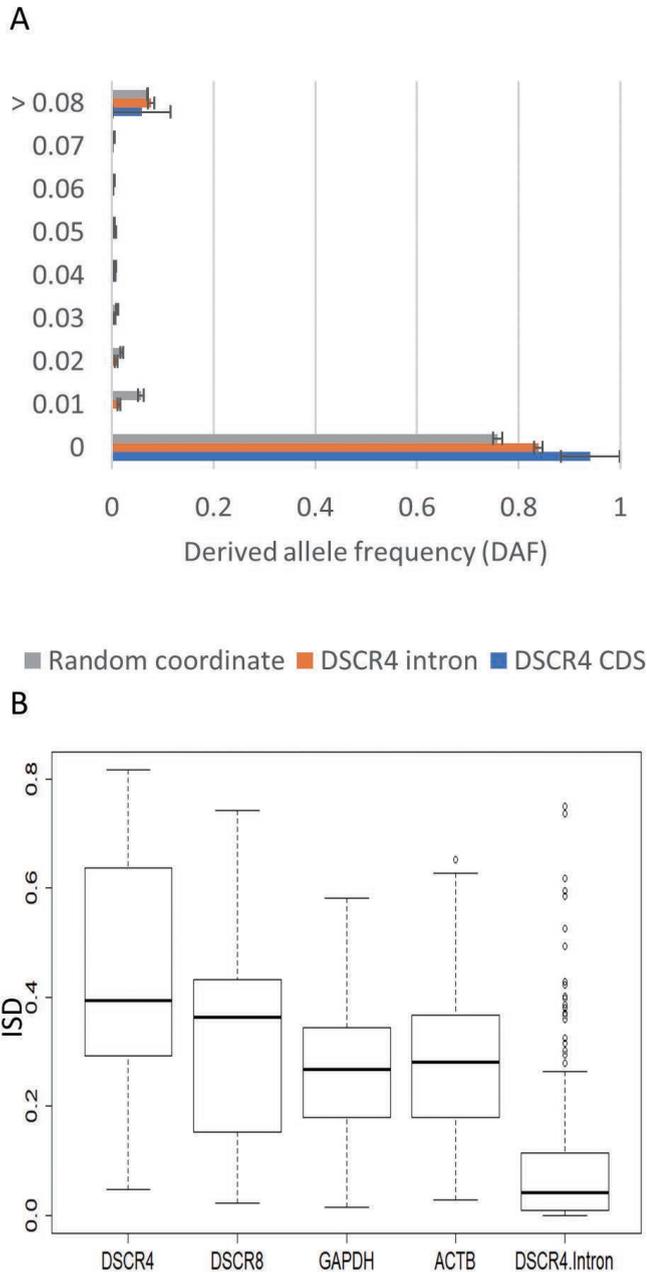


Fig. 2. Functional analysis of DSCR4. (A) Derived allele frequency (DAF) analysis of DSCR4. DSCR4 coding sequences (CDSs) have significantly higher rates of polymorphisms with low-frequency derived alleles compared with DSCR4 intronic sequences (chi-square test, $P = 0.0132$), indicating the action of purifying selection on the CDSs of DSCR4. (B) Intrinsic structural disorder (ISD) characteristics of DSCR4. DSCR4 protein shows a higher degree of intrinsic structural disorder compared with its proximate but older protein-coding gene, DSCR8, and the housekeeping genes GAPDH and ACTB. The ISD score of DSCR4 is also significantly higher than that of DSCR4 translated introns (t -test, $P = 0.002$).

The six GO terms significantly enriched in down-regulated DEGs represent interconnected pathways in the GRNs of human cells (Figs. 3A and 3B). Four of

the six enriched pathways are directly involved in the regulation of movement, migration and motility of cells or cell compartments. On the other hand, the genes up-regulated by DSCR4 overexpression are enriched mainly in the regulation of coagulation and hemostasis processes along with maintenance of apical/basal cell polarity (Figs. 3C and 3D).

If these predictions are correct, we would expect DSCR4 to be expressed mainly in cells with migratory characteristics. We investigated this hypothesis by quantifying the expression of DSCR4 across all human cell transcriptome data available in the Roadmap Epigenomics Project (Roadmap Epigenomics Consortium et al., 2015) and YuGene database (Lê Cao et al., 2014). Only a small portion of highly specialized cells actively migrate within the human body, including leukocytes, stem cells and immune system cells (Entschladen et al., 2005). In line with our predictions, Roadmap data analysis revealed that DSCR4 has significant expression only in K562 cells (Fig. 4). K562 is a human immortalized myelogenous leukemia cell line with erythroleukemia type. Consistent with this finding, analysis of the YuGene database also revealed that monocytes, which are a type of leukocyte, display the highest expression level of DSCR4 (Supplementary Fig. S5).

DSCR4-associated cellular pathways GO and GRN analysis suggested that DSCR4 is involved mainly in the regulation of cell migration, hemostasis and coagulation. What, though, are the cellular pathways through which DSCR4 accomplishes these functions? To answer this question, we reanalyzed the up- and down-regulated DEGs using the Reactome database and KEGG. The Reactome database catalogs a reductionist model which asserts that all biological activity can be represented as events located in subcellular compartments (Fabregat et al., 2016). Consistent with GO and GRN results, these analyses revealed enrichment for several sub-processes involved in cell migration, coagulation and the immune system (Supplementary Fig. S5). Down-regulated DEGs are specifically enriched for the function of semaphorins and integrins, which are important families of proteins involved in the regulation of cell migration (Supplementary Fig. S5A). Strict guidance cues (either repulsive, inhibitory or attractive) normally precede the process of cell migration, and semaphorin receptors are known to govern cell migration mainly by regulating integrin-based cell–substrate adhesion and cytoskeleton dynamics (Huttenlocher and Horwitz, 2011). Semaphorins have also been shown to guide NCC migration in zebrafish (Yu and Moens, 2005). Integrins are also a family of membrane-bound molecules that are important in cell migration, and integrin-based adhesion has served as a model for studying the central role of adhesion in migration (Huttenlocher and Horwitz, 2011). The up-regulated

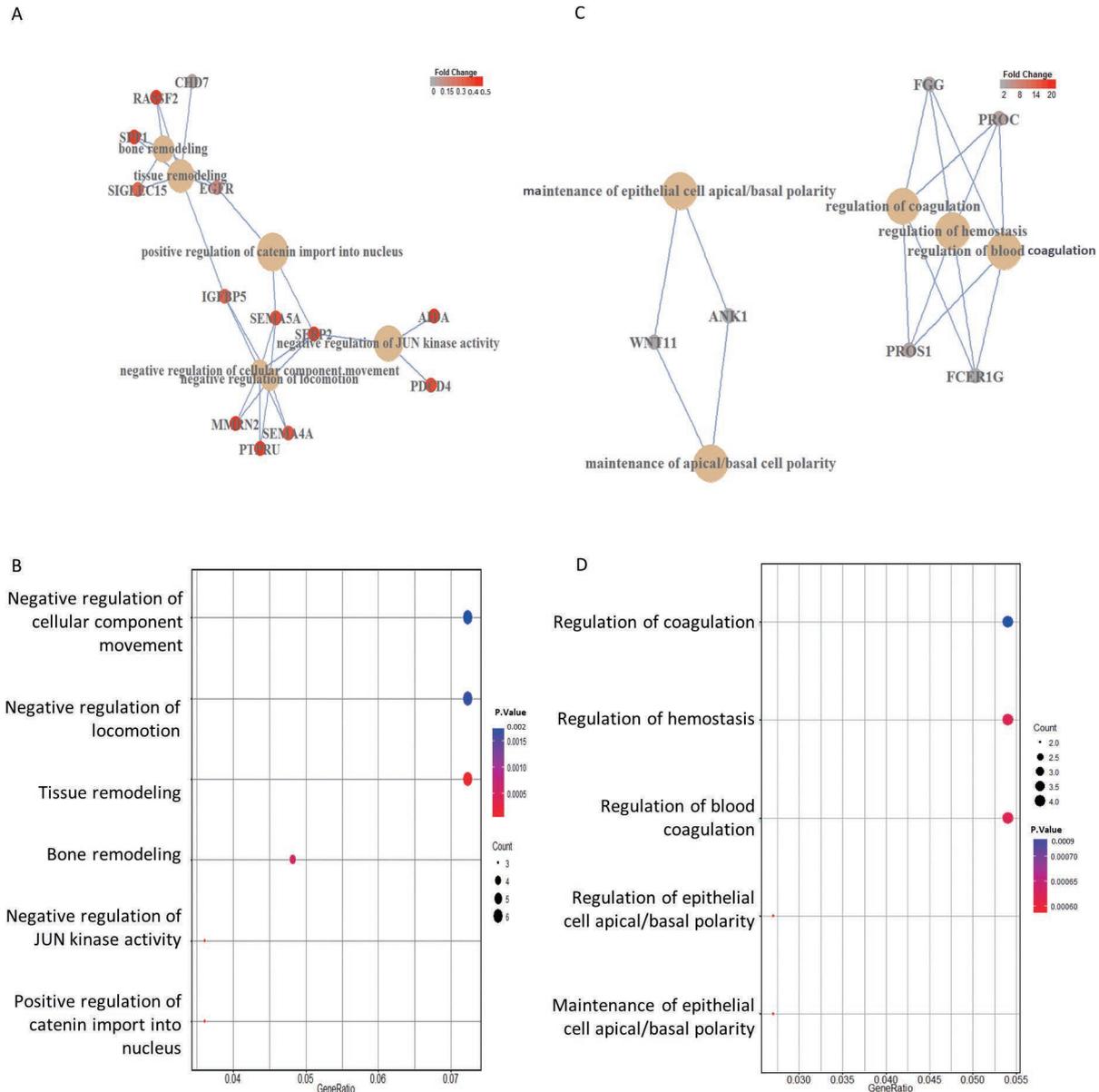


Fig. 3. Gene regulatory network and Gene Ontology analysis of DSCR4 overexpression-mediated DEGs. (A) GRN and (B) GO analysis of DEGs down-regulated by DSCR4 overexpression revealed enrichment of six interconnected pathways involved in the regulation of migration, locomotion and remodeling processes. (C) GRN and (D) GO analysis of up-regulated DEGs, on the other hand, showed enrichment for three interconnected pathways involved in the regulation of blood coagulation and hemostasis along with maintenance of apical/basal cell polarity. Gene Ratio: number of genes within that list which are annotated to the node/size of the list of genes of interest.

DEGs, on the other hand, revealed enrichment for the processes of fibrin clot formation and the complement cascade, which are major components in coagulation and innate immunity, respectively (Supplementary Fig. S5B). From the evolutionary perspective, the two pathways of coagulation and the innate immune system have common ancestry and are highly integrated (Delvaeye and Conway, 2009), which is consistent with our results and supports the hypothesis of DSCR4 being simultaneously involved in functions of the innate immune system

and hemostasis.

The KEGG pathway database is a collection of manually drawn pathway maps representing existing knowledge on molecular interaction, reaction and relation networks (Kanehisa et al., 2016). KEGG analysis further corroborated the results we obtained with the Reactome database by showing that DSCR4 overexpression-mediated DEGs are enriched in coagulation and the complement cascade, and it also revealed that coagulation and the complement cascade are tightly interconnected not only

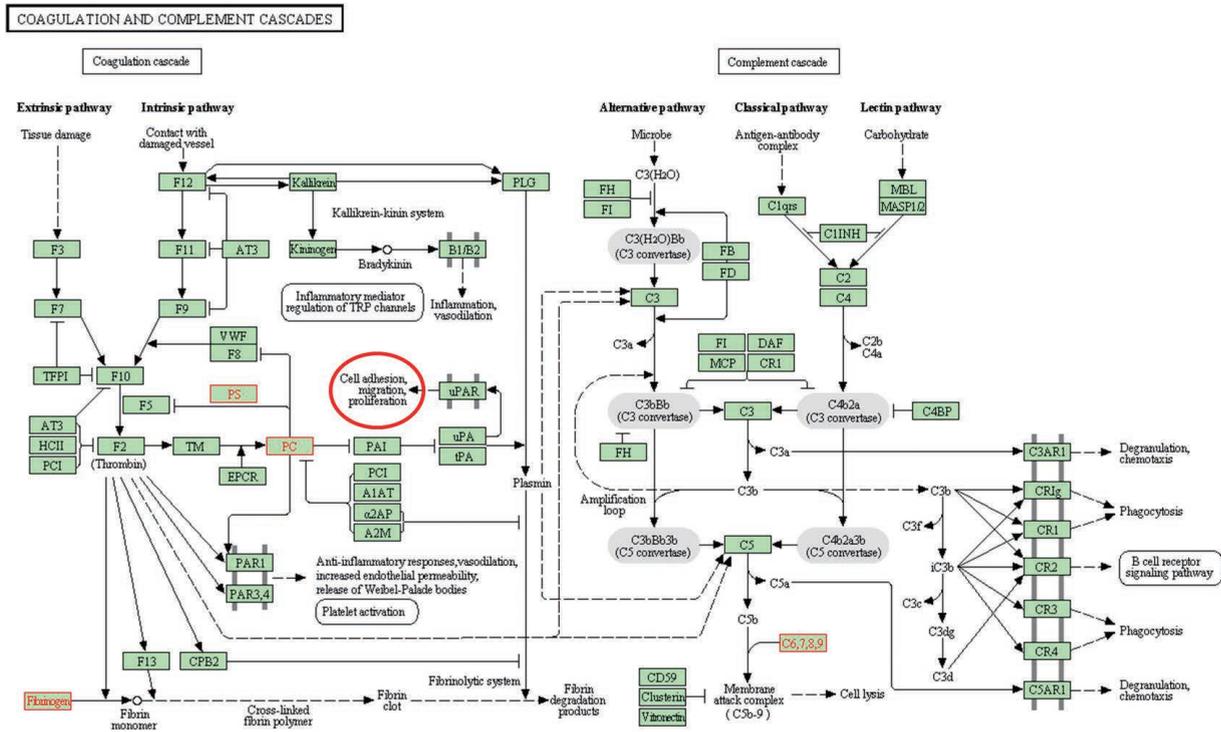


Fig. 4. Expression profile of DSCR4 across human cell lines and tissues. According to Roadmap Epigenomics Project data, DSCR4 and DSCR8, which share a bidirectional promoter, are highly expressed only in K562 cells, a type of leukemia cell. Analysis of transcriptome data provided by Prescott et al. (2015) showed that DSCR4 and DSCR8 also display high expression in human and chimpanzee neural crest cells, which are critical migratory cells involved in facial morphogenesis in the embryo. (1) Data from Prescott et al. (2015). (2) Samples also include esophagus, lung, spleen and fetal large intestine. (3) Samples also include brain germinal matrix, hippocampus, fetal small intestine, stomach, left ventricle, small intestine, sigmoid colon, HEPG2 cells and HMEC cells.

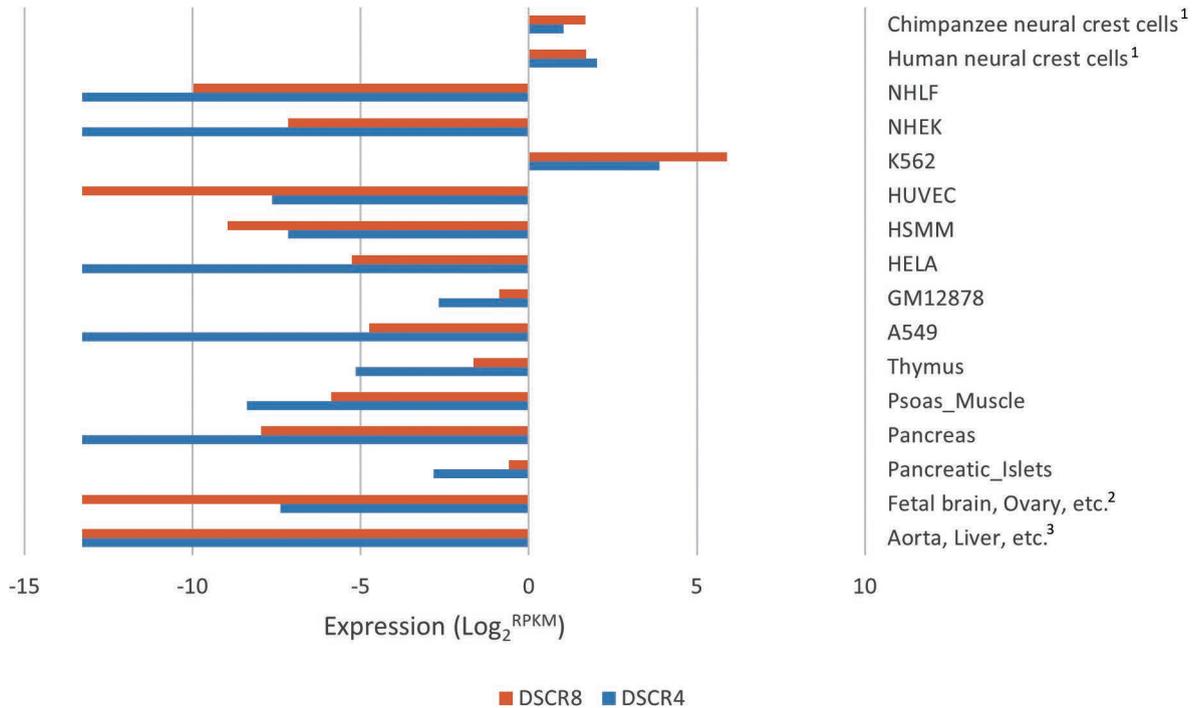


Fig. 5. Interconnection of DSCR4 overexpression-mediated perturbed pathways. KEGG analysis of DSCR4 overexpression-mediated DEGs shows enrichment for the tightly interconnected pathways of the coagulation cascade and the complement cascade (highlighted in red) and further confirm the connection of these cascades with cell adhesion, migration and proliferation (red circle).

with each other but also with processes of cell migration, adhesion and proliferation (Fig. 5). In summary, these results indicate that all the pathways affected by DSCR4 gene perturbation are closely connected and are mainly involved in the processes of cell migration, the immune system and coagulation.

DISCUSSION

DSCR4 is an orphan gene, present only in humans and apes, whose evolution has occurred in multiple steps over the past 100 million years mainly as a result of the action of retrotransposons (Fig. 1). Detailed information on the evolutionary steps by which the three coding exons of DSCR4 arose can be found in Saber et al. (2016). Young *de novo*-originated genes tend to be small in size with relatively low expression levels and weak conservation, which also holds true for DSCR4 and raises doubts about the functionality of this young, experimentally verified, protein-coding gene; recent investigations of ISD, however, support the preadaptation hypothesis of *de novo* gene evolution, indicating that young genes show extreme levels of gene-like traits. These findings also reject the continuum view of *de novo* gene birth, which suggests a series of intermediate stages, or ‘proto-genes’, between non-coding DNA sequences and a fully functional gene (Wilson et al., 2017). The same results were also found in our ISD analysis of DSCR4 (Fig. 2B), which is consistent with the results of DAF analysis (Fig. 2A) and previous findings by Saber et al. (2016) and further supports the functionality of DSCR4.

The present study consistently indicates the likely roles of DSCR4 in cell migration, coagulation and the immune system (Fig. 3). Regulation of cell migration is defined as any process that modulates the frequency, rate or extent of cell migration, and regulation of cellular component movement is defined as any process that modulates the frequency, rate or extent of the movement of a cellular component (The Gene Ontology Consortium, 2000). Tissue and bone remodeling processes are defined as the reorganization or renovation of existing tissues and bones, respectively. Tissue remodeling can either change the characteristics of a tissue, such as in blood vessel remodeling, or result in the dynamic equilibrium of a tissue, such as in bone remodeling (The Gene Ontology Consortium, 2000). Remodeling processes are critical during development, wound repair and metastatic invasion and are driven by coordinated migration of cells through the three-dimensional extracellular matrix (Gjorevski et al., 2015). The other two pathways in which down-regulated DEGs are enriched are also indirectly involved in the regulation of cell migration. The first of these pathways, positive regulation of catenin import into the nucleus, is defined as any process that increases the rate, frequency or extent of the directed

movement of a catenin protein from the cytoplasm into the nucleus (The Gene Ontology Consortium, 2000). The import of β -catenin from cytoplasm to nucleus with the assistance of Wnt protein leads to activation of a signaling pathway named Wnt/ β -catenin signaling (Cai et al., 2013; Jang et al., 2015). It has been shown that the Wnt/ β -catenin signaling pathway is involved in migration of breast cancer cells and metastasis (Cai et al., 2013; Jang et al., 2015). Negative regulation of JUN kinase is the other enriched pathway in down-regulated DEGs and is described as any process that stops, prevents, or reduces the frequency, rate or extent of JUN kinase activity. JUN N-terminal kinases are a group of enzymes that bind and phosphorylate c-JUN proteins. c-jun is a proto-oncogene and is the homolog of the viral oncoprotein v-jun (Zada et al., 2003). In breast cancer cells, c-jun is known to play a key role in migration and invasion of mammary epithelial cells (Jiao et al., 2010). On the other hand, hemostasis and coagulation, which are the main processes enriched in up-regulated DEGs, are critical for the function of the blood and innate immune systems (Esmon et al., 2011; Degen and Palumbo, 2012). Regulation of hemostasis and coagulation are respectively defined as the processes leading to stopping of bleeding (loss of body fluid) or to the arrest of the circulation to an organ and the processes that modulate the frequency, rate or extent of coagulation. Among the other enriched genes involved in the maintenance of apical/basal cell polarity are *wnt11* and *ankyrin1*, which are directly involved in the regulation of cell migration, i.e., a biological function that is also enriched in down-regulated DEGs. Ankyrin1 is a cytoskeleton adaptor protein, and it is affected by p53 and alters cell migration (Hall et al., 2016); and, by interacting with *silberblick*, Wnt11 is involved in controlling cell migration and morphogenesis in zebrafish (Ulrich et al., 2003). In summary, the enriched pathways and processes within the DEGs down-regulated by DSCR4 overexpression consistently indicate a likely role for DSCR4 in the regulation of cell migration. These predictions were further validated by comprehensive analysis of the Roadmap and YuGene expression databases.

T cells and dendritic cells are two other immune system cell types with proven migratory behavior that notably express DSCR4 (Supplementary Fig. S6). Three types of stem cells and their derivatives, including embryonic stem cells (ESCs), ESC-derived neurons and induced pluripotent stem cells (iPSCs), with migratory features are also among the top DSCR4-expressing cells. Blastocysts, the structure formed early in the development of all mammals, contain another type of cell with migratory characteristics that highly expresses DSCR4. NCCs are an embryonic cell population that emerges early in the development of vertebrates and is most relevant to unique human facial traits (Bronner and LeDouarin, 2012). These cells arise from the dorsal part of the neural

tube ectoderm and migrate into the branchial arches and what will later form the embryonic face structure, consequently establishing the central plan of facial morphology (Prescott et al., 2015). Analyses of the YuGene database (Supplementary Fig. S6) and the NCC transcriptome in human and chimpanzee (Prescott et al., 2015) (Fig. 4), indeed revealed high expression of DSCR4 in migratory NCCs as expected. Nakamura et al. (1997), in their original discovery of the DSCR4 gene, also detected abundant expression of this gene in placenta as well as in skeletal and heart muscles. Trophoblasts, which are specialized cells of the placenta, possess migratory behavior (Burrows et al., 1996). Additionally, stem cell migration has been shown to be an essential component of muscle regeneration (Kowalski et al., 2017). These observations further indicate the likely role of DSCR4 in cell migration. In summary, the expression profile of DSCR4 across human cells and tissues supports the predicated functions of this gene revealed by our gene perturbation analysis, and indicates likely roles for DSCR4 in the regulation of cell migration and in the immune system (Fig. 4).

If DSCR4 is involved in the regulation of cell migration, and in coagulation and the immune system, we might expect such features to be observed in human disorders associated with DSCR4 overexpression. Down syndrome is a human disorder characterized by an average 50% overexpression of genes located in this region. Acute leukemia is associated with Down syndrome 10–20-fold more frequently than in the general population, and this disorder is known as the most leukemia-predisposing syndrome (Fong and Brodeur, 1987; Xavier and Taub, 2010). Leukocytes are the human cells that express DSCR4 to the highest level (Fig. 4 and Supplementary Fig. S6); they display active migratory characteristics and have a prominent role in the coagulation process (Bouchard and Tracy, 2003), both of which are biological functions that are enriched in DSCR4 overexpression-mediated DEGs. Another major characteristic of Down syndrome individuals is immune deficiency, especially affecting the innate immune system (Bloemers et al., 2010). This symptom is consistent with our finding that the innate immune system is enriched with genes affected by DSCR4 overexpression (Fig. 5 and Supplementary Fig. S5). In addition, multiple immune system cells including monocytes (leukocytes), T cells and dendritic cells were shown to highly express the DSCR4 gene (Supplementary Fig. S6). Migratory behavior is also essential for the function of immune system cells. Besides highly migratory innate immune cells such as neutrophils, which act as the first line of defense at the site of inflammation, T lymphocytes can also achieve rapid movement (Goldberg et al., 2016). These results indicate an active role for DSCR4 in immune system cells and also in migratory behavior of human cells. Dysmorphology of facial characteristics is another main feature of Down syndrome individuals

that can be explained, at least in part, by our findings that DSCR4 is involved in the regulation of cell migration and that NCCs have notable expression of DSCR4 (Fig. 4 and Supplementary Fig. S6). Cell migration is essential for the function of NCCs, which establish the facial characteristics of the embryo; hence, imbalanced DSCR4 expression may affect the migratory behavior of NCCs, which could in turn affect facial morphology of Down syndrome individuals.

Because there was no previous information on the functions of the DSCR4 gene, we were obliged to use a brute-force approach to identify its functions, which is a statistically challenging objective considering that the cellular effects of short and young genes are usually moderate. The results of our gene perturbation and differential expression investigations, however, are consistent with the transcription profile of human cells and with human diseases in which DSCR4 gene expression is perturbed. This suggests likely roles for DSCR4 in interconnected pathways of cell migration, the immune system and coagulation.

This study is limited by its lack of experimental confirmation of the computational findings. Of great importance is the requirement for *in vitro* analysis to validate the proposed difference in migratory behavior between NCC-differentiated bone marrow cells overexpressing DSCR4 and control cells. Performing this analysis was not feasible for us due to technical limitations with regard to differentiating DSCR4-overexpressing bone marrow cells into NCCs or migratory immune system cells. A second major limitation is the requirement for experimental verification of the possible role of DSCR4 in the formation of the characteristic upward-slanting eyelids and flat nasal bridge in Down syndrome individuals. Investigating this hypothesis was also beyond our reach due to the requirement for *in vivo* experiments in an ape model organism. Notably, this analysis is challenging because of the hominoid-specific occurrence of DSCR4, which precludes the use of common model organisms such as the mouse for such experiments. While acknowledging these limitations, our computational analysis provides a solid hypothesis for future experimental investigation of the hominoid-specific DSCR4 gene. Further experimental research is indeed required to address the above-mentioned limitations and to unravel the roles of this orphan and still-unclassified gene in the etiology of Down syndrome.

DATA AVAILABILITY

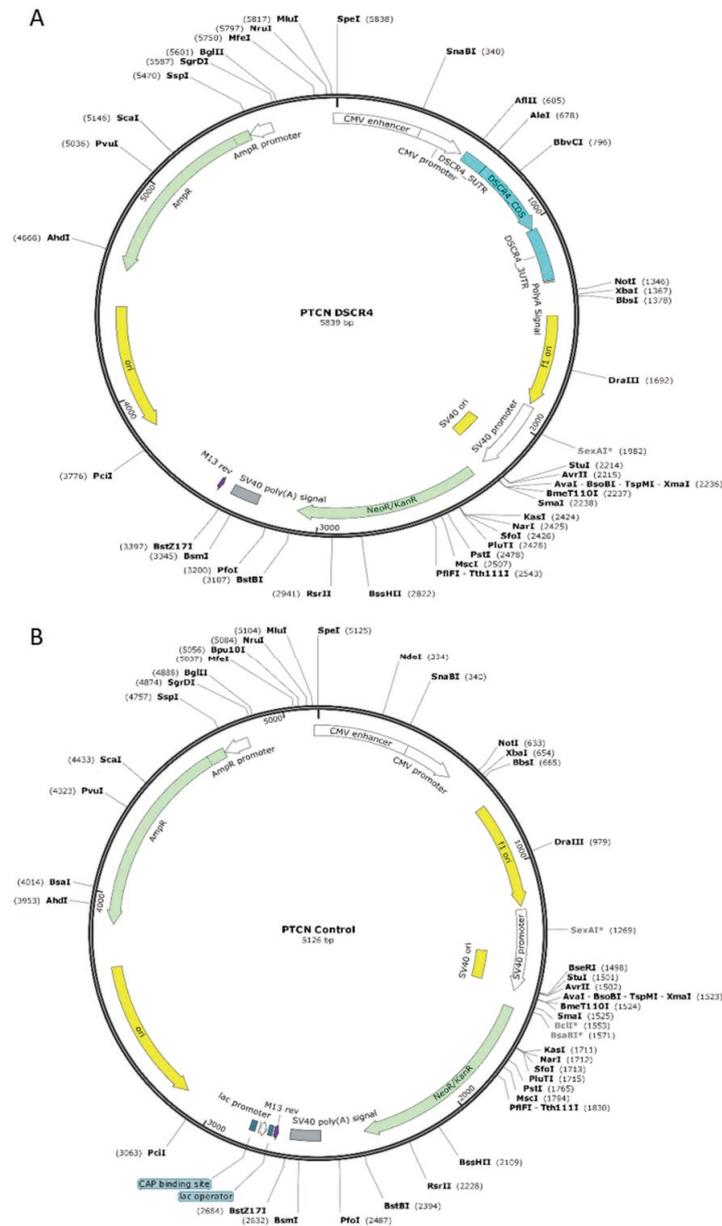
All microarray data have been deposited in the ArrayExpress database with accession number E-MTAB-8779, in accordance with the MIAME guidelines published by the Functional Genomics Data Society (FGED).

This work was supported by a research support grant from the Sasakawa Foundation and a foreign student fellowship from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan to M. M. S. and by a Grant-in-Aid for Scientific Research from MEXT to N. S. Some of the analyses were performed on the National Institute of Genetics of Japan supercomputer.

REFERENCES

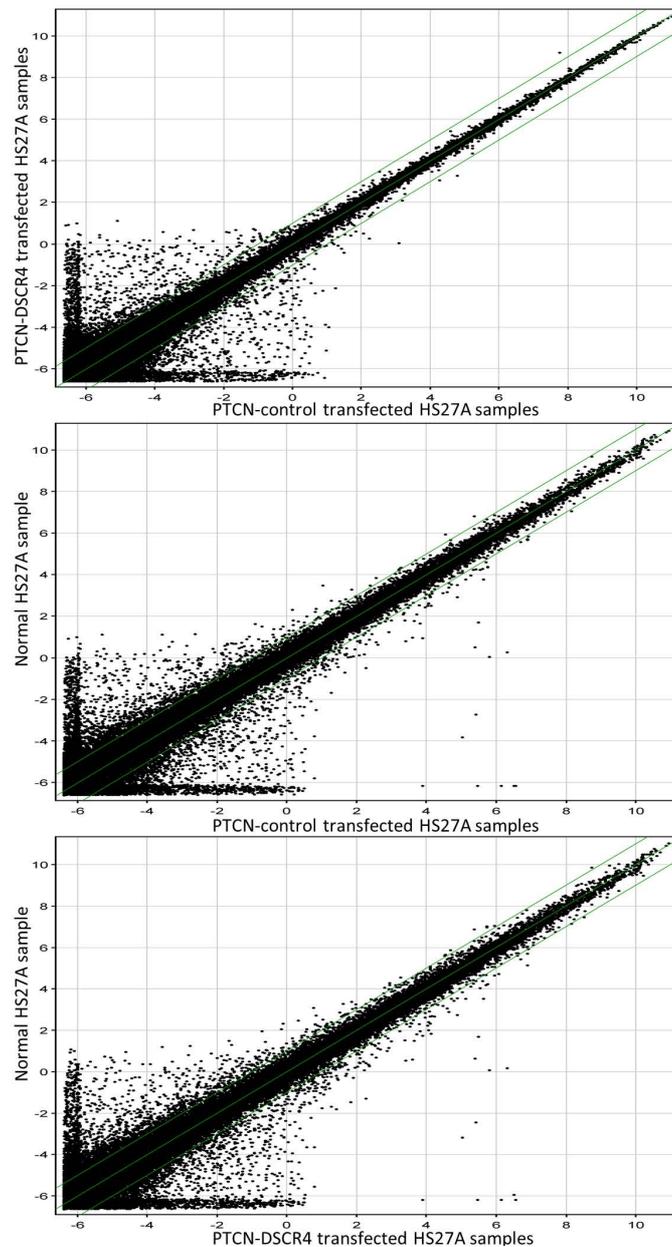
- Asai, S., Yamaki, A., Kudoh, J., Shimizu, N., and Shimizu, Y. (2008) Analysis of the promoter region of human placenta-specific *DSCR4* gene. *Biochim. Biophys. Acta* **1779**, 40–50.
- Bloemers, B. L. P., van Bleek, G. M., Kimpen, J. L. L., and Bont, L. (2010) Distinct abnormalities in the innate immune system of children with Down syndrome. *J. Pediatr.* **156**, 804–809, 809.e1–809.e5.
- Bouchard, B. A., and Tracy, P. B. (2003) The participation of leukocytes in coagulant reactions. *J. Thromb. Haemost.* **1**, 464–469.
- Bronner, M. E., and LeDouarin, N. M. (2012) Development and evolution of the neural crest: an overview. *Dev. Biol.* **366**, 2–9.
- Burrows, T. D., King, A., and Loke, Y. W. (1996) Trophoblast migration during human placental implantation. *Hum. Reprod. Update* **2**, 307–321.
- Cai, J., Guan, H., Fang, L., Yang, Y., Zhu, X., Yuan, J., Wu, J., and Li, M. (2013) MicroRNA-374a activates Wnt/ β -catenin signaling to promote breast cancer metastasis. *J. Clin. Invest.* **123**, 566–579.
- Cohen, W. I. (2002) Health Care Guidelines for Individuals with Down Syndrome—1999 Revision. *In* Down Syndrome: Visions for the 21st Century. (eds.: Cohen, W. I., Nadel, I., and Madnick, M. E.), pp. 237–245. Wiley-Liss, New York.
- Degen, J. L., and Palumbo, J. S. (2012) Hemostatic factors, innate immunity and malignancy. *Thromb. Res.* **129 Suppl 1**, S1–S5.
- Delvaeye, M., and Conway, E. M. (2009) Coagulation and innate immune responses: can we view them separately? *Blood* **114**, 2367–2374.
- Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433–3434.
- Du, Y., Zhang, J., Wang, H., Yan, X., Yang, Y., Yang, L., Luo, X., Chen, Y., Duan, T., and Ma, D. (2011) Hypomethylated *DSCR4* is a placenta-derived epigenetic marker for trisomy 21. *Prenat. Diagn.* **31**, 207–214.
- Dunn, C. A., Romanish, M. T., Gutierrez, L. E., van de Lagemaat, L. N., and Mager, D. L. (2006) Transcription of two human genes from a bidirectional endogenous retrovirus promoter. *Gene* **366**, 335–342.
- Entschladen, F., Drell, T. L., 4th., Lang, K., Masur, K., Palm, D., Bastian, P., Niggemann, B., and Zaenker, K. S. (2005) Analysis methods of human cell migration. *Exp. Cell Res.* **307**, 418–426.
- Epstein, C. J. (2002) 2001 William Allan Award Address. From Down syndrome to the “human” in “human genetics”. *Am. J. Hum. Genet.* **70**, 300–313.
- Esmon, C. T., Xu, J., and Lupu, F. (2011) Innate immunity and coagulation. *J. Thromb. Haemost.* **9 Suppl 1**, 182–188.
- Fabregat, A., Sidiropoulos, K., Garapati, P., Gillespie, M., Hausmann, K., Haw, R., Jassal, B., Jupe, S., Korninger, F., McKay, S., et al. (2016) The Reactome pathway Knowledgebase. *Nucleic Acids Res.* **44**, D481–D487.
- Fong, C.-t., and Brodeur, G. M. (1987) Down’s syndrome and leukemia: epidemiology, genetics, cytogenetics and mechanisms of leukemogenesis. *Cancer Genet. Cytogenet.* **28**, 55–76.
- Gardiner, K., Fortna, A., Bechtel, L., and Davisson, M. T. (2003) Mouse models of Down syndrome: how useful can they be? Comparison of the gene content of human chromosome 21 with orthologous mouse genomic regions. *Gene* **318**, 137–147.
- Gjorevski, N., Piotrowski, A. S., Varner, V. D., and Nelson, C. M. (2015) Dynamic tensile forces drive collective cell migration through three-dimensional extracellular matrices. *Sci. Rep.* **5**, 11458.
- Goldberg, G. L., Cornish, A. L., Murphy, J., Pang, E. S., Lim, L. L., Campbell, I. K., Scalzo-Inguanti, K., Chen, X., McMennamin, P. G., Maraskovsky, E., et al. (2016) G-CSF and neutrophils are nonredundant mediators of murine experimental autoimmune uveoretinitis. *Am. J. Pathol.* **186**, 172–184.
- Hall, A. E., Lu, W.-T., Godfrey, J. D., Antonov, A. V., Paicu, C., Moxon, S., Dalmay, T., Wilczynska, A., Muller, P. A. J., and Bushell, M. (2016) The cytoskeleton adaptor protein ankyrin-1 is upregulated by p53 following DNA damage and alters cell migration. *Cell Death Dis.* **7**, e2184.
- Hassold, T. J., and Jacobs, P. A. (1984) Trisomy in man. *Annu. Rev. Genet.* **18**, 69–97.
- Hattori, M., Fujiyama, A., Taylor, T. D., Watanabe, H., Yada, T., Park, H. S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D. K., et al. (2000) The DNA sequence of human chromosome 21. *Nature* **405**, 311–319.
- Huttenlocher, A., and Horwitz, A. R. (2011) Integrins in cell migration. *Cold Spring Harb. Perspect. Biol.* **3**, a005074.
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264.
- Jang, G.-B., Kim, J.-Y., Cho, S.-D., Park, K.-S., Jung, J.-Y., Lee, H.-Y., Hong, I.-S., and Nam, J.-S. (2015) Blockade of Wnt/ β -catenin signaling suppresses breast cancer metastasis by inhibiting CSC-like phenotype. *Sci. Rep.* **5**, 12465.
- Jiang, X., Liu, C., Yu, T., Zhang, L., Meng, K., Xing, Z., Belichenko, P. V., Kleschevnikov, A. M., Pao, A., Peresie, J., et al. (2015) Genetic dissection of the Down syndrome critical region. *Hum. Mol. Genet.* **24**, 6540–6551.
- Jiao, X., Katiyar, S., Willmarth, N. E., Liu, M., Ma, X., Flomenberg, N., Lisanti, M. P., and Pestell, R. G. (2010) c-Jun induces mammary epithelial cellular invasion and breast cancer stem cell expansion. *J. Biol. Chem.* **285**, 8218–8226.
- Kahlem, P., Sultan, M., Herwig, R., Steinfath, M., Balzereit, D., Eppens, B., Saran, N. G., Pletcher, M. T., South, S. T., Stetten, G., et al. (2004) Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of Down syndrome. *Genome Res.* **14**, 1258–1267.
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462.
- Korbel, J. O., Tirosh-Wagner, T., Urban, A. E., Chen, X.-N., Kasowski, M., Dai, L., Grubert, F., Erdman, C., Gao, M. C., Lange, K., et al. (2009) The genetic architecture of Down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies. *Proc. Natl. Acad. Sci. USA* **106**, 12031–12036.
- Kowalski, K., Kołodziejczyk, A., Sikorska, M., Płaczkiewicz, J., Cichosz, P., Kowalewska, M., Stremińska, W., Jańczyk-Ilach, K., Koblowska, M., Fogtman, A., et al. (2017) Stem cells migration during skeletal muscle regeneration - the

- role of Sdf-1/Cxcr4 and Sdf-1/Cxcr7 axis. *Cell Adh. Migr.* **11**, 384–398.
- Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017) TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819.
- Lana-Elola, E., Watson-Scales, S. D., Fisher, E. M. C., and Tybulewicz, V. L. J. (2011) Down syndrome: searching for the genetic culprits. *Dis. Model. Mech.* **4**, 586–595.
- Lê Cao, K. A., Rohart, F., McHugh, L., Korn, O., and Wells, C. A. (2014) YuGene: a simple approach to scale gene expression data derived from different platforms for integrated analyses. *Genomics* **103**, 239–251.
- Lohse, M., Bolger, A. M., Nagel, A., Fernie, A. R., Lunn, J. E., Stitt, M., and Usadel, B. (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* **40**, W622–W627.
- Lyle, R., Gehrig, C., Neergaard-Henrichsen, C., Deutsch, S., and Antonarakis, S. E. (2004) Gene expression from the aneuploid chromosome in a trisomy mouse model of Down syndrome. *Genome Res.* **14**, 1268–1274.
- Nakamura, A., Hattori, M., and Sakaki, Y. (1997) A novel gene isolated from human placenta located in Down syndrome critical region on chromosome 21. *DNA Res.* **4**, 321–324.
- Nikolaienko, O., Nguyen, C., Crinc, L. S., Cios, K. J., and Gardiner, K. (2005) Human chromosome 21/Down syndrome gene function and pathway database. *Gene* **364**, 90–98.
- Olson, L. E., Roper, R. J., Sengstaken, C. L., Peterson, E. A., Aquino, V., Galdzicki, Z., Siarey, R., Pletnikov, M., Moran, T. H., and Reeves, R. H. (2007) Trisomy for the Down syndrome 'critical region' is necessary but not sufficient for brain phenotypes of trisomic mice. *Hum. Mol. Genet.* **16**, 774–782.
- Pelleri, M. C., Cicchini, E., Locatelli, C., Vitale, L., Caracausi, M., Piovesan, A., Rocca, A., Poletti, G., Seri, M., Strippoli, P., et al. (2016) Systematic reanalysis of partial trisomy 21 cases with or without Down syndrome suggests a small region on 21q22.13 as critical to the phenotype. *Hum. Mol. Genet.* **25**, 2525–2538.
- Prescott, S. L., Srinivasan, R., Marchetto, M. C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F. H., Swigut, T., and Wysocka, J. (2015) Enhancer divergence and *cis*-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83.
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015) Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330.
- Saber, M. M., Adeyemi Babarinde, I., Hettiarachchi, N., and Saitou, N. (2016) Emergence and evolution of Hominidae-specific coding and noncoding genomic sequences. *Genome Biol. Evol.* **8**, 2076–2092.
- The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* **526**, 68–74.
- The Gene Ontology Consortium (2000) Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29.
- Toyoda, A., Noguchi, H., Taylor, T. D., Ito, T., Pletcher, M. T., Sakaki, Y., Reeves, R. H., and Hattori, M. (2002) Comparative genomic sequence analysis of the human chromosome 21 Down syndrome critical region. *Genome Res.* **12**, 1323–1332.
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015) Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419.
- Ulrich, F., Concha, M. L., Heid, P. J., Voss, E., Witzel, S., Roehl, H., Tada, M., Wilson, S. W., Adams, R. J., Soll, D. R., et al. (2003) Slb/Wnt11 controls hypoblast cell migration and morphogenesis at the onset of zebrafish gastrulation. *Development* **130**, 5375–5384.
- Wilson, B. A., Foy, S. G., Neme, R., and Masel, J. (2017) Young genes are highly disordered as predicted by the preadaptation hypothesis of *de novo* gene birth. *Nat. Ecol. Evol.* **1**, 0146.
- Wiseman, F. K., Alford, K. A., Tybulewicz, V. L. J., and Fisher, E. M. C. (2009) Down syndrome—recent progress and future prospects. *Hum. Mol. Genet.* **18**, R75–R83.
- Xavier, A. C., and Taub, J. W. (2010) Acute leukemia in children with Down syndrome. *Haematologica* **95**, 1043–1045.
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287.
- Yu, H.-H., and Moens, C. B. (2005) Semaphorin signaling guides cranial neural crest cell migration in zebrafish. *Dev. Biol.* **280**, 373–385.
- Yu, J.-F., Cao, Z., Yang, Y., Wang, C.-L., Su, Z.-D., Zhao, Y.-W., Wang, J.-H., and Zhou, Y. (2016) Natural protein sequences are more intrinsically disordered than random sequences. *Cell. Mol. Life Sci.* **73**, 2949–2957.
- Zada, A. A., Singh, S. M., Reddy, V. A., Elsässer, A., Meisel, A., Haferlach, T., Tenen, D. G., Hiddemann, W., and Behre, G. (2003) Downregulation of *c-Jun* expression and cell cycle regulatory molecules in acute myeloid leukemia cells upon CD44 ligation. *Oncogene* **22**, 2296–2308.



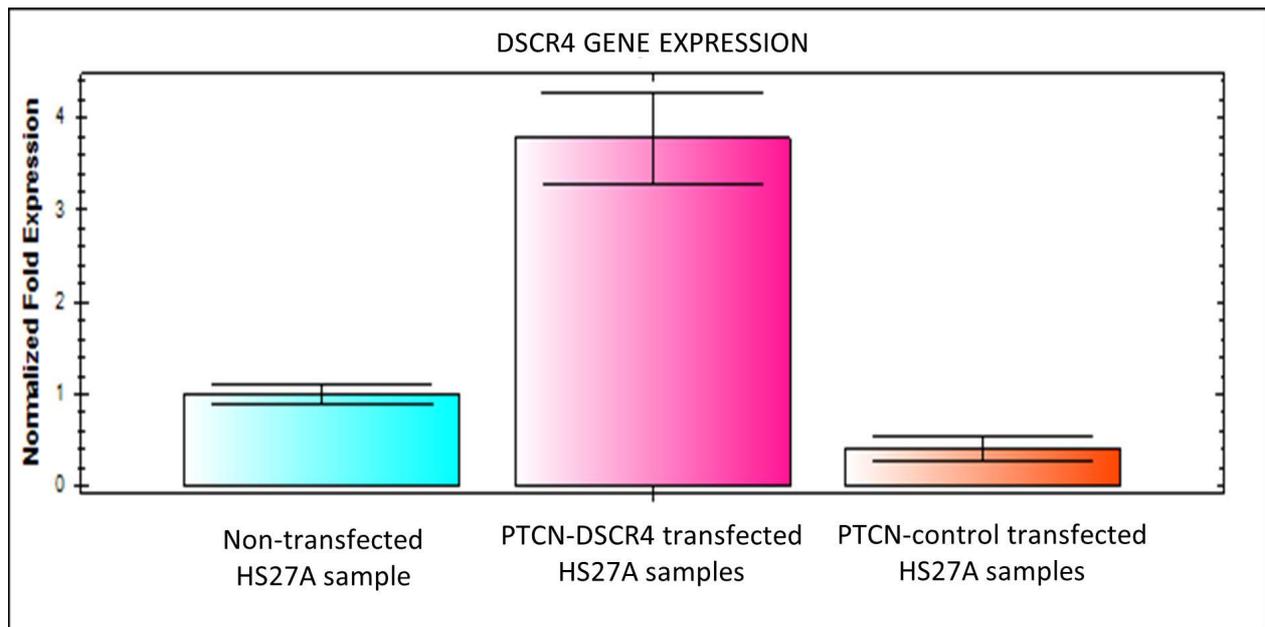
Supplementary Fig. S1. DSCR4 plasmid vector and control vector structure.

(A) PTCN-DSCR4 plasmid carrier includes the full-length cDNA sequence of DSCR4 and its upstream and downstream untranslated regions (UTRs), along with a CMV-derived enhancer and promoter for strong expression in eukaryotic host cells, and eukaryotic and bacterial selection markers. (B) PTCN-control vector was constructed from the PTCN-DSCR4 vector by removing the DSCR4 coding sequence and UTR elements.



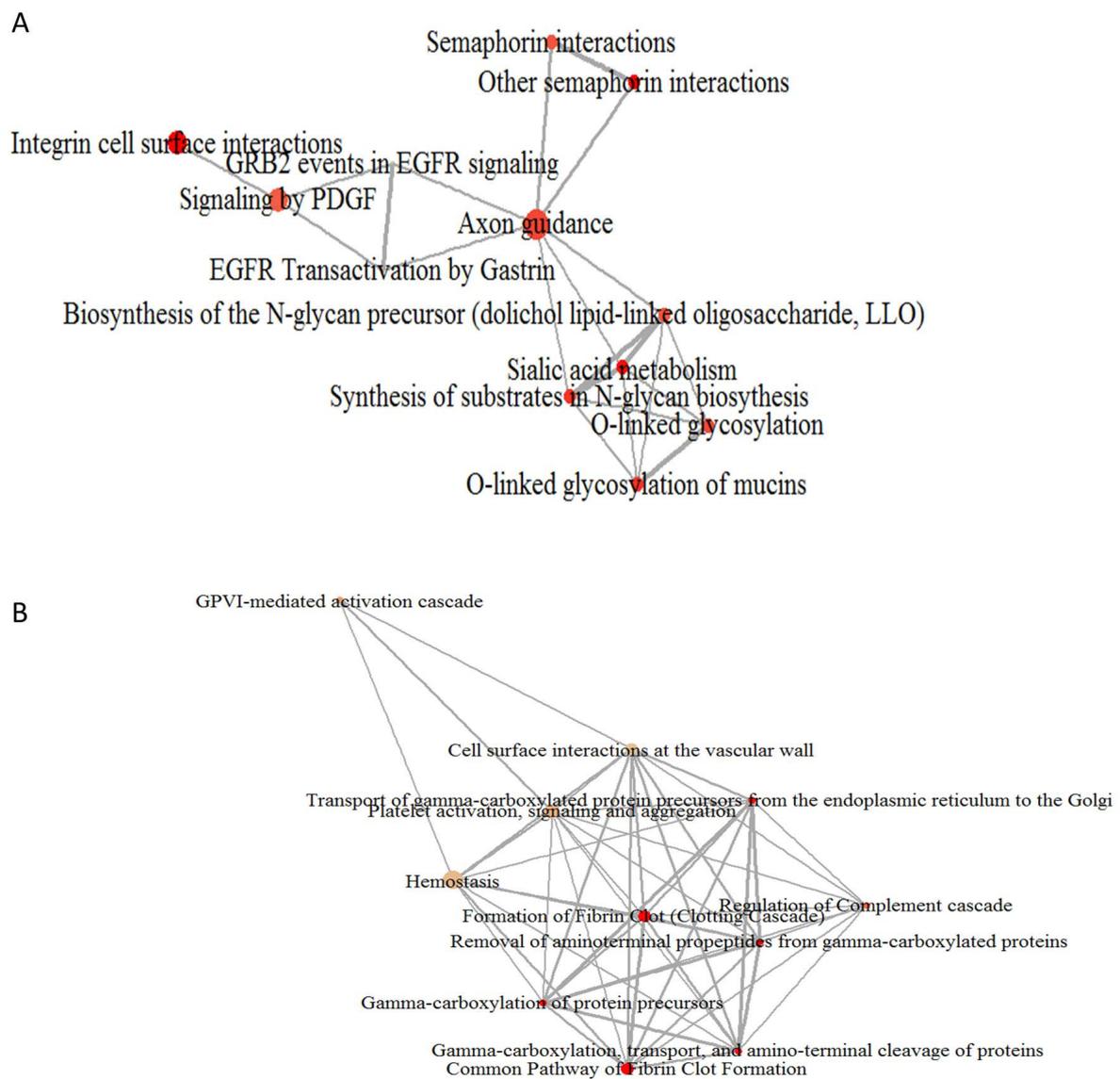
Supplementary Fig. S2. Expression correlation analysis between DSCR4-overexpressing HS-27A cells and control samples.

Scatter plots of expression correlation between DSCR4-overexpressing cells (PTCN-DSCR4-transfected HS-27A cells) and control samples (PTCN-control-transfected HS-27A cells and normal HS-27A cells) confirmed that microarray experiments and the normalization process were properly conducted. PTCN-DSCR4-transfected HS-27A samples also showed more uniform expression values against PTCN-control-transfected samples than against normal HS-27A samples, as expected because PTCN-DSCR4- and PTCN-control-transfected cells underwent the same transfection treatment.



Supplementary Fig. S3. Confirmation of DSCR4 overexpression in DSCR4-transfected HS-27A cells.

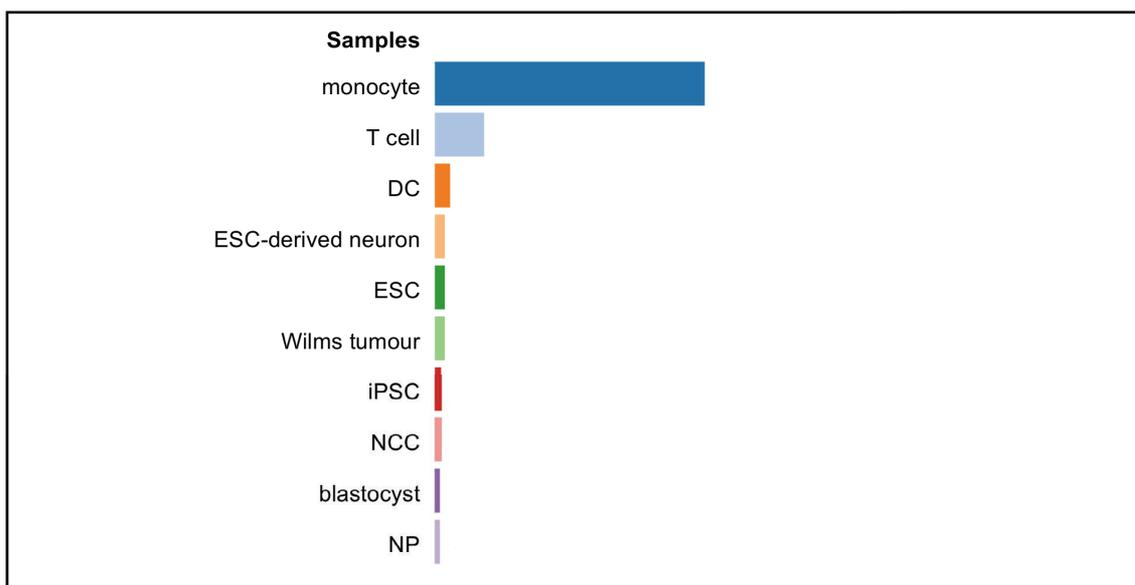
qPCR expression analysis of DSCR4 normalized by GAPDH revealed approximately 9-fold overexpression in PTCN-DSCR4-transfected HS-27A cells compared with control-transfected samples, which is consistent with microarray results and confirms the accuracy of our overexpression procedures. The analysis also showed a 4-fold increase in DSCR4 expression compared with normal (non-transfected) HS-27A cells.



Supplementary Fig. S5. DSCR4-associated cellular pathways.

(A) Cellular pathways enriched in genes down-regulated by DSCR4 overexpression. Semaphorins and integrins are two important classes of membrane proteins that are involved in the regulation of cell migration. (B) Cellular pathways enriched in genes up-regulated by DSCR4 overexpression. Regulation of the complement cascade and the clotting cascade plays critical roles in the innate immune system and hemostasis, respectively.

Sample types highlighted in graph (Top 10 terms)



Supplementary Fig. S6. Expression profiles of DSCR4 in the YuGene database.

DSCR4 has the highest expression value in monocytes, which are a type of leukocytes, followed by two other immune system cells with migratory characteristics. Stem cells, tumor cells, neural crest cells and blastocyst components are the other types of cells with migratory behavior that highly express DSCR4.