**ORIGINAL ARTICLE**

**Introduction**

Stem cells have attracted much attention recently because of their unique biological behaviors and attractive clinical usages. CD34⁺ hematopoietic stem cells (HSCs), for example, have been used for clinical stem cell transplantation [1]. Clinically, HSCs can be obtained from three different anatomic locations: bone marrow (BM), peripheral blood (PB), and umbilical cord blood (CB). To improve our ability to utilize various HSCs, it is necessary to understand the differences between these stem cells and to decipher their underlying mechanisms.

In humans, most HSCs and precursors express CD34⁺ antigen while being negative for the CD38 hematopoietic lineage marker [2,3]. The CD34⁺ CD38⁻ population can be divided into long-term and short-term functional HSCs with surface makers CD34⁺ CD33⁻ CD38⁻ Rholow and CD34⁺ CD33⁻ CD38⁻ Rohl"}, 356–369

**SUMMARY**

**Objective:** Somatic CD34⁺ CD38⁻ stem cells can differentiate into cells of hematopoietic and endothelial lineages and have been clinically used to treat diseases. These stem cells can be obtained from cord blood (CB), bone marrow or granulocyte-macrophage colony-stimulating factor–mobilized peripheral blood. Unmasking genes differentially expressed in hematopoietic stem cells (HSCs) from different anatomic locations can improve our understanding of their basic biological features and help in clinical decision making when applying different HSCs.

**Materials and Methods:** We performed microarray analysis on human CD34⁺ CD38⁻ HSCs isolated from CB, bone marrow and peripheral blood. Systems biology and advanced bioinformatics tools were used to better understand the biological modules and genetic networks accompanying each HSC subtype.

**Results:** We identified HSC genes differentially expressed in various HSCs and found them to be involved in critical biological processes such as cell cycle regulation, cell motility, and endogenous antigen presentation. Among these three HSC types, HSCs from CB expressed the fewest rejection and immune response-associated genes, thereby showing the best potential as a transplantation source. Analysis of HSC-enriched genes using systems biology tools revealed a complex genetic network functioning in different CD34⁺ CD38⁻ cells, in which several genes act as hubs, such as MYC in CB HSCs and hepatic growth factor in bone marrow HSCs, to maintain the stability or connectivity of the whole network.

**Conclusion:** This study provides the foundation for a more detailed understanding of CD34⁺ CD38⁻ HSCS from different sources, and reveals the potentials of different HSCs for different clinical applications. [Taiwan J Obstet Gynecol 2009;48(4):356–369]

**Key Words:** CD34, hematopoietic stem cells, systems biology, transplantation
Properties of Various CD34+ CD38− Stem Cells

respectively [4]. Recently, the CD133 antigen has also been used for HSC isolation [5,6]. The gene expression patterns of different human CD34+ HSCs isolated from various anatomic locations have been characterized by genomic analyses, which revealed genes involved in self-renewal, differentiation and lineage choice [2,7]. However, the assignment of biological significance to the filtered genes or even to each HSC type remains difficult. In this post-genomics era, modern bioinformatics and systems biology tools can help to predict biological behaviors or interactions of gene products. Deductions of de novo gene relationships by exploratory computational tools and systems biology algorithms can provide a systematic approach to discovering novel molecular events and relationships [8–10]. The combination of microarray and computational results enables biologists to speed up their research and to analyze data.

The aims of this study were to identify the genes responsible for stem cell properties in various CD34+ CD38− HSCs and to reveal the biological differences between them using systems biology approaches. We compared the gene expression profiles of HSCs from various anatomic locations. Our analysis revealed that CD34+ CD38− HSCs from different sources possess different biological properties and hence hold different application potentials.

Materials and Methods

CD34+ CD38− cells and primary microvascular endothelial cells (MVECs)
Human CD34+ CD38− cells were isolated from BM, PB or CB of healthy individuals (Poietics; Lonza Group Ltd., Basel, Switzerland). Human MVECs (Clonetics; Lonza Group Ltd.) were cultured in EGM-2 MV BulletKit medium (Lonza Group Ltd.).

RNA isolation and real-time polymerase chain reaction (PCR)
Total messenger RNA was extracted using an RNeasy mini kit (catalog no. 74106; Qiagen, Hilden, Germany) and 100 ng to 1 μg of total RNA was reverse transcribed using a First cDNA Synthesis Kit (catalog no. K1612; Fermentas, Glen Burnie, MD, USA). For quantitative real-time PCR analysis, the human pre-messenger RNA sequence was obtained from the National Center for Biotechnology Information AceView program (http://www.ncbi.nlm.nih.gov/AceView/). All primers were designed across introns in the Primer3 Web site (http://frodo.wi.mit.edu/primer3/) or Primer Express software (Applied Biosystems, Foster City, CA, USA).

Thermodynamics and primer specificity analyses were performed using the Vector NTI suite (Invitrogen, Carlsbad, CA, USA) and the National Center for Biotechnology Information reverse e-PCR program (http://www.ncbi.nlm.nih.gov/sutils/e-PCR/reverse.cgi/). Real-time PCR reactions were performed using Maxima SYBR Green qPCR Master Mix (catalog no. K0222; Fermentas), and specific products were detected and analyzed using a StepOne sequence detector (Applied Biosystems). The expression level of each gene was normalized to the expression level of glyceraldehyde 3-phosphate dehydrogenase.

Array probe preparation, data analysis, and functional network analyses
Total RNA collection, complementary RNA probe preparation, array hybridization and data analysis were performed as previously described [10,11]. Expression profiles of CD34+ CD38− HSCs, MVECs and PB mononuclear cells were implemented using the Affymetrix HG-U133A chip (Affymetrix, Santa Clara, CA, USA). Gene annotation enrichment analysis was performed using the DAVID 2008 tool (http://david.abcc.ncifcrf.gov/) [12]. EASE score, a Fisher’s exact test-based scoring system, was used to calculate p values to determine the probability of the number of genes being specifically associated (enriched) with a given gene ontology (GO) term or occurring by random chance. The Ingenuity Pathway Analysis (IPA) Web tool developed by the Ingenuity Co. (http://www.ingenuity.com/) was used to construct functional regulatory networks of gene profiles. IPA uses the Ingenuity Pathways Knowledge Base to identify known interactions between focus genes as well as other genes that are not present in the gene list. IPA then determines a statistical score for each network according to the fit of the network to the set of focus genes. The score is the negative log of p and denotes the likelihood of the focus genes in the network being found together by chance.

Results

Molecular signatures of various CD34+ CD38− HSCs
To access molecular signature genes for HSCs, we obtained human CD34+ CD38− HSCs from umbilical CB. Abundant expression of known “stemness” or precursor genes (including GATA2, GATA1 and RUNX1) in CD34+ CD38− cells were verified by quantitative PCR (Figure 1A).

To compare the gene expression profiles and biological variations between HSCs from different anatomic
locations, we also isolated CD34\(^+\) CD38\(^−\) HSCs from BM and granulocyte-macrophage colony-stimulating factor–mobilized PB from healthy individuals. The gene expression profiles of various HSCs were determined and compared. The gene expression profiles of these three cell types were implemented at least in triplicate using the Affymetrix HG-U133A chip. According to the statistical pipeline used [10,11,13], genes differentially

![Figure 1. Transcriptome analysis of hematopoietic stem cells (HSCs) isolated from different anatomic locations. (A) Enriched expression of stemness and precursor genes in isolated HSCs. CD34\(^+\) CD38\(^−\) HSCs isolated by magnetic beads were subjected to total RNA isolation and quantitative polymerase chain reaction analysis. Primary microvascular endothelial cells were used as a mature progeny control. Mean expression levels of target genes were compared with that of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) control. (B) Multidimensional scaling (MDS) plots using the whole transcriptome (upper) or filtered genes (lower) show the discrimination ability of the obtained molecular signature for cell groups. Each spot represents a single array sample. BM = bone marrow; CB = cord blood; PB = peripheral blood. (C) A heat map shows genes differentially expressed in different CD34\(^+\) CD38\(^−\) HSCs. Genes in red, increased expression; in blue, decreased.](image-url)
expressed between each cell type (the molecular signature) were identified. A total of 910 probe sets were uniquely upregulated in CB HSCs, 450 in BM HSCs, and 912 in PB HSCs (with a positive false discovery rate threshold, \( q < 0.005 \)). In contrast, 220 probe sets were uniquely downregulated in CB HSCs, 851 in BM HSCs, and 1287 in PB HSCs (\( q < 0.005 \)). The discrimination ability of these genes was assessed using multidimensional scaling (Figure 1B). A gene expression heat map for these genes indicated their unique expression in different CD34\(^+\)CD38\(^−\) HSCs (Figure 1C).

The top 50 genes most strongly expressed or repressed in CD34\(^+\)CD38\(^−\) CB HSCs are listed in Tables 1 and 2, respectively. The top 50 genes in BM HSCs and PB HSCs are listed in Tables 3 and 4, respectively. In CB HSCs, genes involved in transplant rejection (such as \( HLA-A \), \( HLA-B \), \( HLA-C \), \( HLA-DPA1 \), \( HLA-DQA1 \) and \( HLA-DQB1 \)) or immune response (such as interferon-induced proteins and CXCR4) were strongly downregulated (labeled with asterisks, Table 2).

In BM HSCs, genes involved in cell survival (such as the baculoviral IAP repeat-containing 5, also known as survivin [14]) and cell cycle progression (such as cyclin A2) were most active (Table 3). In contrast, in PB HSCs, cell cycle inhibitors such as cyclin-dependent kinase inhibitor 1C (\( CDKN1C \); also known as \( p57KIP2 \)) were overexpressed (Table 4), suggesting a less active cell cycle or cell proliferating condition in HSCs from PB. PB HSCs expressed the highest levels of \( PROM1 \) (\( CD133 \)) gene, a marker of HSCs and other stem cell types, such as colon epithelial stem cells [13], indicating a higher stemness status of this HSC subtype. PB HSCs also expressed more genes involved in endothelial (vascular endothelial growth factor A) and megakaryocytic (\( RUNX1 \) [15]) differentiation, indicating their multipotency. Ezrin, a cell membrane-skeleton cross-linker that can increase cellular motility [16,17] was also strongly expressed in PB HSCs (Table 4).

**Functional grouping reveals unique biological properties in each HSC subtype**

Gene signatures provided indications of the functional differences among these three HSC subtypes. To provide quantitative evidence and to gain more insights into the functional consequences of differential gene expression patterns, probe sets enriched or repressed in each HSC subtype were subjected to the GO database search [18] to find statistically overrepresented functional groups within each gene list. The DAVID 2008 Web-based tool, a graph theory evidence-based method to agglomerate gene/protein identifiers [12], was used for this task.

The GO categories of biological processes statistically overrepresented (\( p < 0.05 \)) are shown in Figure 2. The predominant processes in CB HSCs included those pertaining to DNA repair, nuclear messenger RNA splicing and export, and ribosomal RNA or transfer RNA processing (Figure 2A). In contrast, response to stress and endogenous antigen processing and presentation pathways were less active in CB-derived HSCs (Figure 2A). Of the three HSC subtypes, HSCs from BM possessed the largest number of genes involved in cell cycle progression (especially those involved in mitosis and spindle biogenesis) and transplant rejection (i.e. major histocompatibility complex class I endogenous antigen processing and presentation). BM HSCs also possessed the fewest genes involved in cell cycle arrest and cellular migration, indicating an active proliferation yet stationary property of HSCs inside BM (Figure 2B). HSCs from PB expressed the most genes involved in cell motility, reflecting the fact that they were collected after granulocyte-macrophage colony-stimulating factor mobilization (Figure 2C). PB HSCs may also replicate less and participate more in immune response or angiogenesis, because the genes involved in these biological processes were the most differentially expressed in PB HSCs (Figure 2C).

**Coordinated changes in functional networks of CB HSC upregulated genes**

The above genomics data suggest different biological potentials of HSC subtypes isolated from various anatomic locations. Increasing evidence shows that genes do not act individually but collaborate in genetic networks. To better understand how genes enriched in each population of CD34\(^+\)CD38\(^−\) cells are related and how they relate to cellular function, we performed genetic network analysis for signature genes using the IPA Web tool and the Ingenuity knowledge database. The knowledge base behind IPA summarizes known molecular interactions evidenced in the published literature. The term network in IPA does not refer to a biological or canonical pathway with a distinct function (i.e. angiogenesis) but is, rather, a reflection of all the interactions of a given protein as defined in the literature (see “Materials and Methods”).

In CB HSCs, a major network consisting of 258 genes was identified (Figure 3A). This network revealed genes whose roles may be more crucial than others in CB HSCs. \( GATA1 \), \( MYC \), \( NMYC \), \( JUN \) and \( YYHAZ \) are the “hub” genes (genes involved in the greatest number of interactions with other components) that either preserve the stability of the whole network or link submodules in the same network to each other (Figure 3A). These hub genes are all involved in the regulation
Table 1. Top 50 genes enriched in cord blood CD34+ CD38− hematopoietic stem cells

<table>
<thead>
<tr>
<th>Probe set ID</th>
<th>UniGene ID</th>
<th>Gene title</th>
<th>Gene symbol</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>203196_at</td>
<td>Hs.508423</td>
<td>ATP-binding cassette, sub-family C (CFTR/MRP, member 4)</td>
<td>ABCC4</td>
<td>chr13q32</td>
</tr>
<tr>
<td>212224_at</td>
<td>Hs.76392</td>
<td>Aldehyde dehydrogenase 1 family, member A1</td>
<td>ALDH1A1</td>
<td>chr9q21.13</td>
</tr>
<tr>
<td>204446_s_at</td>
<td>Hs.89499</td>
<td>Arachidonate 5-lipoxygenase</td>
<td>ALOX5</td>
<td>chr10q11.2</td>
</tr>
<tr>
<td>204174_s_at</td>
<td>Hs.507658</td>
<td>Arachidonate 5-lipoxygenase-activating protein</td>
<td>ALOX5AP</td>
<td>chr13q12</td>
</tr>
<tr>
<td>203304_at</td>
<td>Hs.533336</td>
<td>BMP and activin membrane-bound inhibitor homolog (Xenopus laevis)</td>
<td>Bambi</td>
<td>chr10p12.3-11.2</td>
</tr>
<tr>
<td>202391_at</td>
<td>Hs.201641</td>
<td>Brain abundant, membrane attached signal protein 1</td>
<td>BASP1</td>
<td>chr5p15.1-p14</td>
</tr>
<tr>
<td>215440_s_at</td>
<td>Hs.184736</td>
<td>BEX family member 4</td>
<td>BEX4</td>
<td>chrXq22.1-q22.3</td>
</tr>
<tr>
<td>218829_s_at</td>
<td>Hs.709651</td>
<td>Chromodomain helicase DNA binding protein 7</td>
<td>CHD7</td>
<td>chr8q12.2</td>
</tr>
<tr>
<td>201735_s_at</td>
<td>Hs.481186</td>
<td>Chloride channel 3</td>
<td>CLCN3</td>
<td>chr4q33</td>
</tr>
<tr>
<td>205624_at</td>
<td>Hs.646</td>
<td>Carboxypeptidase A3 (mast cell)</td>
<td>CPA3</td>
<td>chr3q21-q25</td>
</tr>
<tr>
<td>205159_at</td>
<td>Hs.592192</td>
<td>Colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)</td>
<td>CSF2RB</td>
<td>chr2q13.1</td>
</tr>
<tr>
<td>218858_at</td>
<td>Hs.112981</td>
<td>DEP domain containing 6</td>
<td>DEPDC6</td>
<td>chr8q24.12</td>
</tr>
<tr>
<td>219799_s_at</td>
<td>Hs.179608</td>
<td>Deoxyribo-nucleic acid reduced (SDR family) member 9</td>
<td>DHR59</td>
<td>chr2q31.1</td>
</tr>
<tr>
<td>209560_at</td>
<td>Hs.533717</td>
<td>Delta-like 1 homolog (Drosophila)</td>
<td>DLK1</td>
<td>chr14q32</td>
</tr>
<tr>
<td>204720_s_at</td>
<td>Hs.647643</td>
<td>DnaJ (Hsp40) homolog, subfamily C, member 6</td>
<td>DNAJC6</td>
<td>chr1pter-q31.3</td>
</tr>
<tr>
<td>201123_s_at</td>
<td>Hs.534314</td>
<td>Eukaryotic translation initiation factor 5A</td>
<td>EIF5A</td>
<td>chr17p13-p12</td>
</tr>
<tr>
<td>201794_s_at</td>
<td>Hs.706750</td>
<td>Fas apoptotic inhibitory molecule</td>
<td>FAIM</td>
<td>chr3q22.3</td>
</tr>
<tr>
<td>218170_at</td>
<td>Hs.483296</td>
<td>Guanine nucleotide binding protein (G protein), gamma 11</td>
<td>GNG11</td>
<td>chr7q21</td>
</tr>
<tr>
<td>215780_s_at</td>
<td>Hs.436687</td>
<td>Histidine decarboxylase</td>
<td>HDC</td>
<td>chr8q24.12</td>
</tr>
<tr>
<td>203394_s_at</td>
<td>Hs.250666</td>
<td>High mobility group AT-hook 2</td>
<td>HMGA2</td>
<td>chr12q12</td>
</tr>
<tr>
<td>208025_s_at</td>
<td>Hs.505924</td>
<td>High mobility group box 1</td>
<td>HMGB1</td>
<td>chr12q12</td>
</tr>
<tr>
<td>209204_at</td>
<td>Hs.436792</td>
<td>Intercellular adhesion molecule 4 (Landsteiner-Wiener blood group)</td>
<td>ICAM4</td>
<td>chr19p13.2-cen</td>
</tr>
<tr>
<td>211953_s_at</td>
<td>Hs.643743</td>
<td>Importin 5</td>
<td>IPOS</td>
<td>chr13q32.2</td>
</tr>
<tr>
<td>218170_at</td>
<td>Hs.483296</td>
<td>Isoch ros matsina domain containing 1</td>
<td>ISO1</td>
<td>chr5q21-q33.3</td>
</tr>
<tr>
<td>205885_s_at</td>
<td>Hs.694732</td>
<td>Integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)</td>
<td>ITGA4</td>
<td>chr2q31.1</td>
</tr>
<tr>
<td>221558_s_at</td>
<td>Hs.555947</td>
<td>Lymphoid enhancer-binding factor 1</td>
<td>LEF1</td>
<td>chr4q23-q25</td>
</tr>
<tr>
<td>209204_at</td>
<td>Hs.436792</td>
<td>Lim domain only 4</td>
<td>LMO4</td>
<td>chr1p22.3</td>
</tr>
<tr>
<td>201669_s_at</td>
<td>Hs.519009</td>
<td>Myristoylated alanine-rich protein kinase C substrate</td>
<td>MARCKS</td>
<td>chr6q22.2</td>
</tr>
<tr>
<td>212473_s_at</td>
<td>Hs.706750</td>
<td>Mitochrondial ribosomal protein S31</td>
<td>MRPS31</td>
<td>chr13q44.11</td>
</tr>
<tr>
<td>212604_at</td>
<td>Hs.154655</td>
<td>Mitochondrial ribosomal protein S31</td>
<td>MYH10</td>
<td>chr7p13</td>
</tr>
<tr>
<td>212372_at</td>
<td>Hs.16355</td>
<td>Myosin, heavy chain 10, non-muscle</td>
<td>NMT2</td>
<td>chr10p13</td>
</tr>
<tr>
<td>205006_s_at</td>
<td>Hs.60339</td>
<td>N-myristoyltransferase 2</td>
<td>NMT2</td>
<td>chr10p13</td>
</tr>
<tr>
<td>208857_s_at</td>
<td>Hs.279257</td>
<td>Protein-l-isoaaspartate (l-aspartate) O-methyltransferase</td>
<td>PCMT1</td>
<td>chr4q24-q25</td>
</tr>
<tr>
<td>218676_s_at</td>
<td>Hs.285218</td>
<td>Phosphatidylinositol transfer protein</td>
<td>PCTP</td>
<td>chr17q21-q24</td>
</tr>
<tr>
<td>206726_at</td>
<td>Hs.128433</td>
<td>Prostaglandin D2 synthase, hematopoietic</td>
<td>PGDS</td>
<td>chr4q22.3</td>
</tr>
<tr>
<td>209318_s_at</td>
<td>Hs.444975</td>
<td>Pleiomorphic adenoma gene-like 1</td>
<td>PLAG1</td>
<td>chr6q24-q25</td>
</tr>
<tr>
<td>213241_at</td>
<td>Hs.584845</td>
<td>Psoriasin 1</td>
<td>PSN1C1</td>
<td>chr12q23.3</td>
</tr>
<tr>
<td>219485_s_at</td>
<td>Hs.522752</td>
<td>Proteasome (prosome, macropain) 265 subunit</td>
<td>PSMD10</td>
<td>chrXq22.3</td>
</tr>
<tr>
<td>206157_at</td>
<td>Hs.591286</td>
<td>Pentraxin-related gene, rapidly induced by IL-1 beta</td>
<td>PTX3</td>
<td>chr3q25</td>
</tr>
<tr>
<td>212012_s_at</td>
<td>Hs.332197</td>
<td>Peroxidase homolog (Drosophila)</td>
<td>PXDN</td>
<td>chr2p25</td>
</tr>
<tr>
<td>218668_s_at</td>
<td>Hs.119889</td>
<td>Rap2C, member of RAS oncogene family</td>
<td>RAP2C</td>
<td>chrXq25</td>
</tr>
<tr>
<td>202988_s_at</td>
<td>Hs.75256</td>
<td>Regulator of G-protein signaling 1</td>
<td>RG5</td>
<td>chr1q31</td>
</tr>
<tr>
<td>205352_s_at</td>
<td>Hs.478153</td>
<td>Serpin peptidase inhibitor, clade I (neuroserpin), member 1</td>
<td>SERPIN1</td>
<td>chr3q26.1</td>
</tr>
<tr>
<td>203775_at</td>
<td>Hs.489190</td>
<td>Solute carrier family 25, member 13 (citrin)</td>
<td>SLC25A13</td>
<td>chr7q21.3</td>
</tr>
<tr>
<td>220059_at</td>
<td>Hs.435579</td>
<td>Signal transducing adaptor family member 1</td>
<td>STAP1</td>
<td>chr4q13.2</td>
</tr>
<tr>
<td>201147_s_at</td>
<td>Hs.701968</td>
<td>TIMP metallopeptidase inhibitor 3</td>
<td>TIMP1</td>
<td>chr2q12.1</td>
</tr>
<tr>
<td>219478_at</td>
<td>Hs.36688</td>
<td>WAP four-disulfide core domain 1</td>
<td>WFDC1</td>
<td>chr6q24.3</td>
</tr>
</tbody>
</table>
## Table 2. Top 50 genes downregulated in cord blood CD34<sup>+</sup> CD38<sup>-</sup> hematopoietic stem cells

<table>
<thead>
<tr>
<th>Probe set ID</th>
<th>UniGene ID</th>
<th>Gene title</th>
<th>Gene symbol</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>209160_at</td>
<td>Hs.78183</td>
<td>Aldo-keto reductase family 1, member C3</td>
<td>AKR1C3</td>
<td>chr10p15-p14</td>
</tr>
<tr>
<td>207848_at</td>
<td>Hs.89648</td>
<td>Arginine vasopressin</td>
<td>AVP</td>
<td>chr20p13</td>
</tr>
<tr>
<td>205681_at</td>
<td>Hs.227817</td>
<td>BCL2-related protein A1</td>
<td>BCL2A1</td>
<td>chr15q24.3</td>
</tr>
<tr>
<td>201641_at</td>
<td>Hs.118110</td>
<td>Bone marrow stromal cell antigen 2</td>
<td>BST2</td>
<td>chr19p13.2</td>
</tr>
<tr>
<td>208683_at</td>
<td>Hs.350899</td>
<td>Calpain 2, (m/II) large subunit</td>
<td>CAPN2</td>
<td>chr1q41-q42</td>
</tr>
<tr>
<td>209732_at</td>
<td>Hs.85201</td>
<td>C-type lectin domain family 2, member B</td>
<td>CLEC2B</td>
<td>chr12p13-p12</td>
</tr>
<tr>
<td>213415_at</td>
<td>Hs.655445</td>
<td>Chloride intracellular channel 2</td>
<td>CLIC2</td>
<td>chrXq28</td>
</tr>
<tr>
<td>217028_at</td>
<td>Hs.593413</td>
<td>Chemokine (C-X-C motif) receptor 4</td>
<td>CXCR4*</td>
<td>chr2q21</td>
</tr>
<tr>
<td>218094_s_at</td>
<td>Hs.650505</td>
<td>Dysbindin domain containing 2</td>
<td>DBND2</td>
<td>chr20q13.12</td>
</tr>
<tr>
<td>209383_at</td>
<td>Hs.505777</td>
<td>DNA-damage-inducible transcript 3</td>
<td>DDIT3</td>
<td>chr12q13.1-q13.2</td>
</tr>
<tr>
<td>201324_at</td>
<td>Hs.436298</td>
<td>Epithelial membrane protein 1</td>
<td>EMP1</td>
<td>chr12p12.3</td>
</tr>
<tr>
<td>205767_at</td>
<td>Hs.115263</td>
<td>Eregulin</td>
<td>EREG</td>
<td>chr4q13.3</td>
</tr>
<tr>
<td>212998_x_at</td>
<td>Hs.728</td>
<td>Major histocompatibility complex, class II, DQ beta 1</td>
<td>HLA-DQB1*</td>
<td>chr14q24-q31</td>
</tr>
<tr>
<td>206082_at</td>
<td>Hs.654480</td>
<td>HLA complex P5</td>
<td>HCPS*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>213932_x_at</td>
<td>Hs.181244</td>
<td>Major histocompatibility complex, class I, A</td>
<td>HLA-A*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>209141_at</td>
<td>Hs.549053</td>
<td>Major histocompatibility complex, class I, B</td>
<td>HLA-B*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>208812_x_at</td>
<td>Hs.654404</td>
<td>Major histocompatibility complex, class I, C</td>
<td>HLA-C*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>213537_at</td>
<td>Hs.347270</td>
<td>Major histocompatibility complex, class II, DP alpha 1</td>
<td>HLA-DPA1*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>213831_at</td>
<td>Hs.387679</td>
<td>Major histocompatibility complex, class II, DQ alpha 1</td>
<td>HLA-DQA1*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>211654_x_at</td>
<td>Hs.409934</td>
<td>Major histocompatibility complex, class II, DQ beta 1</td>
<td>HLA-DQB1*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>200599_s_at</td>
<td>Hs.192374</td>
<td>Heat shock protein 90kDa beta (Grp94), member 1</td>
<td>HSP90B1*</td>
<td>chr12q24.2-q24.3</td>
</tr>
<tr>
<td>214453_at</td>
<td>Hs.82316</td>
<td>Interferon-induced protein 44</td>
<td>IFI44*</td>
<td>chr1p31.1</td>
</tr>
<tr>
<td>204439_at</td>
<td>Hs.389724</td>
<td>Interferon-induced protein 44-like</td>
<td>IFI44L*</td>
<td>chr1p31.1</td>
</tr>
<tr>
<td>204415_at</td>
<td>Hs.511731</td>
<td>Interferon, alpha-inducible protein 6</td>
<td>IFI6*</td>
<td>chr1p35</td>
</tr>
<tr>
<td>201601_x_at</td>
<td>Hs.458414</td>
<td>Interferon induced transmembrane protein 1 (9-27)</td>
<td>IFITM1*</td>
<td>chr11p15.5</td>
</tr>
<tr>
<td>205376_at</td>
<td>Hs.658245</td>
<td>Inositol polyphosphate-4-phosphatase, type II, 105kDa</td>
<td>INPP4B</td>
<td>chr4q31.21</td>
</tr>
<tr>
<td>204698_at</td>
<td>Hs.459265</td>
<td>Interferon stimulated exonuclease gene 20kDa</td>
<td>ISG20*</td>
<td>chr15q26</td>
</tr>
<tr>
<td>213281_at</td>
<td>Hs.525704</td>
<td>Jun oncogene</td>
<td>JUN</td>
<td>chr1p32-p31</td>
</tr>
<tr>
<td>203542_s_at</td>
<td>Hs.150557</td>
<td>Kruppel-like factor 9</td>
<td>KLF9</td>
<td>chr9q13</td>
</tr>
<tr>
<td>214056_at</td>
<td>Hs.632486</td>
<td>Myeloid cell leukemia sequence 1 (BCL2-related)</td>
<td>MCL1</td>
<td>chr1q21</td>
</tr>
<tr>
<td>212509_s_at</td>
<td>Hs.250723</td>
<td>Matrix-remodelling associated 7</td>
<td>MTRA7</td>
<td>chr17q25.1-q25.2</td>
</tr>
<tr>
<td>218966_at</td>
<td>Hs.487036</td>
<td>Mucin 18 kDa</td>
<td>MUC18</td>
<td>chr15q21</td>
</tr>
<tr>
<td>202074_s_at</td>
<td>Hs.332706</td>
<td>Optineurin</td>
<td>OPTN</td>
<td>chr10p13</td>
</tr>
<tr>
<td>202464_s_at</td>
<td>Hs.195471</td>
<td>6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3</td>
<td>PFKFB3</td>
<td>chr10p14-p15</td>
</tr>
<tr>
<td>204279_at</td>
<td>Hs.654585</td>
<td>Proteasome (prosome, macropain) subunit, beta type, 9</td>
<td>PSMB9</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>204897_at</td>
<td>Hs.199248</td>
<td>Prostaglandin E receptor 4 (subtype EP4)</td>
<td>PTGER4</td>
<td>chr5p13.1</td>
</tr>
<tr>
<td>204070_at</td>
<td>Hs.17466</td>
<td>Retinoic acid receptor responder (tazarotene induced) 3</td>
<td>RARRES3</td>
<td>chr11q23</td>
</tr>
<tr>
<td>212044_s_at</td>
<td>Hs.523463</td>
<td>Ribosomal protein L27a</td>
<td>RPL27A</td>
<td>chr11p15</td>
</tr>
<tr>
<td>214041_x_at</td>
<td>Hs.433701</td>
<td>Ribosomal protein L37a</td>
<td>RPL37A</td>
<td>chr2q35</td>
</tr>
<tr>
<td>202028_s_at</td>
<td>Hs.380953</td>
<td>Ribosomal protein L38</td>
<td>RPL38</td>
<td>chr17q23-q25</td>
</tr>
<tr>
<td>221943_x_at</td>
<td>Hs.380953</td>
<td>Ribosomal protein L38</td>
<td>RPL38</td>
<td>chr17q23-q25</td>
</tr>
<tr>
<td>200908_s_at</td>
<td>Hs.437594</td>
<td>Ribosomal protein, large, P2</td>
<td>RPLP2</td>
<td>chr11p15.5-p15.4</td>
</tr>
<tr>
<td>204030_at</td>
<td>Hs.134665</td>
<td>Schwannomin interacting protein 1</td>
<td>SCHIP1</td>
<td>chr3q25.33</td>
</tr>
<tr>
<td>200665_s_at</td>
<td>Hs.111779</td>
<td>Secreted protein, acidic, cysteine-rich (osteonectin)</td>
<td>SPARC</td>
<td>chr5q31.3-q32</td>
</tr>
<tr>
<td>217995_at</td>
<td>Hs.511251</td>
<td>Sulfide quinone reductase-like (yeast)</td>
<td>SQORL</td>
<td>chr15q15</td>
</tr>
<tr>
<td>208763_s_at</td>
<td>Hs.522074</td>
<td>TSC22 domain family, member 3</td>
<td>TSC22D3</td>
<td>chrXq22.3</td>
</tr>
<tr>
<td>202096_s_at</td>
<td>Hs.202</td>
<td>Translocator protein (18 kDa)</td>
<td>TSPO</td>
<td>chr22q13.31</td>
</tr>
<tr>
<td>202112_at</td>
<td>Hs.440848</td>
<td>von Willebrand factor</td>
<td>VWF*</td>
<td>chr12p13.3</td>
</tr>
<tr>
<td>201368_at</td>
<td>Hs.503093</td>
<td>Zinc finger protein 36, C3H type-like 2</td>
<td>ZFPM3L2</td>
<td>chr2p22.3-p21</td>
</tr>
<tr>
<td>201369_s_at</td>
<td>Hs.503093</td>
<td>Zinc finger protein 36, C3H type-like 2</td>
<td>ZFPM3L2</td>
<td>chr2p22.3-p21</td>
</tr>
</tbody>
</table>

*Discussed in the text.*
### Table 3. Top 50 genes enriched in bone marrow CD34⁺ CD38⁻ hematopoietic stem cells

<table>
<thead>
<tr>
<th>Probe set ID</th>
<th>UniGene ID</th>
<th>Gene title</th>
<th>Gene symbol</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>205239_at</td>
<td>Hs.645475</td>
<td>Amphiregulin (schwannoma-derived growth factor)</td>
<td>AREG</td>
<td>chr4q13-q21</td>
</tr>
<tr>
<td>214575_s_at</td>
<td>Hs.72885</td>
<td>Azurocidin 1 (cationic antimicrobial protein 37)</td>
<td>AZU1</td>
<td>chr19p13.3</td>
</tr>
<tr>
<td>218332_at</td>
<td>Hs.334370</td>
<td>Brain expressed, X-linked 1</td>
<td>BEX1</td>
<td>chrXq21-q23</td>
</tr>
<tr>
<td>202095_s_at</td>
<td>Hs.514527</td>
<td>Baculoviral IAP repeat-containing 5 (survivin)</td>
<td>BIRC5*</td>
<td>chr17q25</td>
</tr>
<tr>
<td>205950_s_at</td>
<td>Hs.23118</td>
<td>Carbonic anhydrase I</td>
<td>CA1</td>
<td>chr8q13-q22.1</td>
</tr>
<tr>
<td>214315_x_at</td>
<td>Hs.515162</td>
<td>Calreticulin</td>
<td>CALR</td>
<td>chr19p13.3-p13.2</td>
</tr>
<tr>
<td>213226_at</td>
<td>Hs.58974</td>
<td>Cyclin A2</td>
<td>CCNA2</td>
<td>chr4q12</td>
</tr>
<tr>
<td>206488_s_at</td>
<td>Hs.120949</td>
<td>CD36 molecule (thrombospondin receptor)</td>
<td>CD36</td>
<td>chr7q11.2</td>
</tr>
<tr>
<td>218157_x_at</td>
<td>Hs.22065</td>
<td>CDC42 small effector 1</td>
<td>CDC42SE1</td>
<td>chr1q21.2</td>
</tr>
<tr>
<td>209057_s_at</td>
<td>Hs.485471</td>
<td>CDCS cell division cycle 5-like (S. pombe)</td>
<td>CDC5</td>
<td>chr1p21.3</td>
</tr>
<tr>
<td>201131_s_at</td>
<td>Hs.461086</td>
<td>Cadherin 1, type 1, E-cadherin (epithelial)</td>
<td>CDH1*</td>
<td>chr16q22.1</td>
</tr>
<tr>
<td>209714_s_at</td>
<td>Hs.84113</td>
<td>Cyclin-dependent kinase inhibitor 3</td>
<td>CDKN3</td>
<td>chr14q22</td>
</tr>
<tr>
<td>219555_s_at</td>
<td>Hs.55028</td>
<td>Centromere protein N</td>
<td>CENPN</td>
<td>chr16q23.2</td>
</tr>
<tr>
<td>205382_s_at</td>
<td>Hs.155597</td>
<td>Complement factor D (adipsin)</td>
<td>CFD</td>
<td>chr19p13.3</td>
</tr>
<tr>
<td>214665_s_at</td>
<td>Hs.406234</td>
<td>Calcium binding protein P22</td>
<td>CHP</td>
<td>chr15q13.3</td>
</tr>
<tr>
<td>210140_at</td>
<td>Hs.143212</td>
<td>Catenin (cadherin-associated protein), beta 1, 88kDa</td>
<td>CTNNB1*</td>
<td>chr3p21.3</td>
</tr>
<tr>
<td>204971_at</td>
<td>Hs.518198</td>
<td>Cystatin A (stefin A)</td>
<td>CSTA</td>
<td>chr3q21</td>
</tr>
<tr>
<td>201533_s_at</td>
<td>Hs.476018</td>
<td>Catenin (cadherin-associated protein), beta 1, 88kDa</td>
<td>CTNNB1*</td>
<td>chr3p21.3</td>
</tr>
<tr>
<td>205653_s_at</td>
<td>Hs.421724</td>
<td>Cathepsin G</td>
<td>CTSG</td>
<td>chr14q11.2</td>
</tr>
<tr>
<td>202079_at</td>
<td>Hs.302082</td>
<td>Dipeptidase 3</td>
<td>DPEP3</td>
<td>chr16q22.1</td>
</tr>
<tr>
<td>214313_s_at</td>
<td>Hs.158688</td>
<td>Eukaryotic translation initiation factor 5 B</td>
<td>EIF5B</td>
<td>chr2q11.2</td>
</tr>
<tr>
<td>206871_at</td>
<td>Hs.99863</td>
<td>Elastase 2, neutrophil</td>
<td>ELA2</td>
<td>chr19p13.3</td>
</tr>
<tr>
<td>219672_at</td>
<td>Hs.274309</td>
<td>Erythroid associated factor</td>
<td>ERAF</td>
<td>chr16p11.2</td>
</tr>
<tr>
<td>209480_at</td>
<td>Hs.409934</td>
<td>Major histocompatibility complex, class II, DQ beta 1</td>
<td>HLA-DQB1*</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>207165_at</td>
<td>Hs.72550</td>
<td>Hyaluronan-mediated motility receptor (RHAMM)</td>
<td>HMMR</td>
<td>chr5q33.2-qter</td>
</tr>
<tr>
<td>201163_s_at</td>
<td>Hs.47908</td>
<td>Insulin-like growth factor binding protein 7</td>
<td>IGFBP7</td>
<td>chr4q12</td>
</tr>
<tr>
<td>213792_s_at</td>
<td>Hs.465744</td>
<td>Insulin receptor</td>
<td>INSR</td>
<td>chr19p13.3-p13.2</td>
</tr>
<tr>
<td>219306_at</td>
<td>Hs.658939</td>
<td>Kinesin family member 15</td>
<td>KIF15</td>
<td>chr3p21.31</td>
</tr>
<tr>
<td>202209_at</td>
<td>Hs.111632</td>
<td>LSM3 homolog, U6 small nuclear RNA associated</td>
<td>LSM3</td>
<td>chr3p25.1</td>
</tr>
<tr>
<td>213975_s_at</td>
<td>Hs.706744</td>
<td>Lysozyme (renal amyloidosis)</td>
<td>LYZ</td>
<td>chr12q15</td>
</tr>
<tr>
<td>217692_at</td>
<td>Hs.674373</td>
<td>Mago-nashi homolog 2, proliferation-associated</td>
<td>MAGOH2</td>
<td>chr17p13.1</td>
</tr>
<tr>
<td>203948_s_at</td>
<td>Hs.458272</td>
<td>Myeloperoxidase</td>
<td>MPO*</td>
<td>chr17q23.1</td>
</tr>
<tr>
<td>206023_at</td>
<td>Hs.418367</td>
<td>Neuromedin U</td>
<td>NMU</td>
<td>chr4q12</td>
</tr>
<tr>
<td>218039_at</td>
<td>Hs.615092</td>
<td>Nucleolar and spindle associated protein 1</td>
<td>NUSAP1</td>
<td>chr15q15.1</td>
</tr>
<tr>
<td>212239_at</td>
<td>Hs.132225</td>
<td>Phosphonoside-3-kinase, regulatory subunit 1 (alpha)</td>
<td>PIK3R1</td>
<td>chr5q13.1</td>
</tr>
<tr>
<td>201407_s_at</td>
<td>Hs.705383</td>
<td>Protein phosphatase 1, catalytic subunit, beta isoform</td>
<td>PPP1CB</td>
<td>chr2p23</td>
</tr>
<tr>
<td>203401_at</td>
<td>Hs.654581</td>
<td>Phosphotyrosyl phosphatase synthetase 2</td>
<td>PRPS2</td>
<td>chrXp22.3-p22.2</td>
</tr>
<tr>
<td>220051_at</td>
<td>Hs.72026</td>
<td>Protease, serine, 21 (testisin)</td>
<td>PRSS21</td>
<td>chr16p13.3</td>
</tr>
<tr>
<td>207341_at</td>
<td>Hs.928</td>
<td>Proteinase 3</td>
<td>PRTN3</td>
<td>chr19p13.3</td>
</tr>
<tr>
<td>222077_s_at</td>
<td>Hs.708122</td>
<td>Rac GTPase activating protein 1</td>
<td>RACGAP1</td>
<td>chr12q13.13</td>
</tr>
<tr>
<td>206851_at</td>
<td>Hs.73389</td>
<td>Ribonuclease, RNase A family, 3</td>
<td>RNASE3</td>
<td>chr14q24-q31</td>
</tr>
<tr>
<td>214097_at</td>
<td>Hs.190968</td>
<td>Ribosomal protein S21</td>
<td>RPS21</td>
<td>chr20q13.3</td>
</tr>
<tr>
<td>202917_s_at</td>
<td>Hs.416073</td>
<td>S100 calcium binding protein A8</td>
<td>S100A8</td>
<td>chr1q21</td>
</tr>
<tr>
<td>220232_at</td>
<td>Hs.379191</td>
<td>Stearoyl-CoA desaturase 5</td>
<td>SCAD5</td>
<td>chr4q21.22</td>
</tr>
<tr>
<td>202083_s_at</td>
<td>Hs.464184</td>
<td>SEC14-like 1 (S. cerevisiae)</td>
<td>SEC14L1</td>
<td>chr17q25.1-q25.2</td>
</tr>
<tr>
<td>204496_at</td>
<td>Hs.401843</td>
<td>Sterol, calmodulin binding protein 3</td>
<td>STRN3</td>
<td>chr14q13-q21</td>
</tr>
<tr>
<td>206438_s_at</td>
<td>Hs.167165</td>
<td>Tectonic family member 2</td>
<td>TCTN2</td>
<td>chr12q4.31</td>
</tr>
<tr>
<td>202704_at</td>
<td>Hs.653129</td>
<td>Transducer of ERBB2, 1</td>
<td>TOB1</td>
<td>chr17q21</td>
</tr>
<tr>
<td>201690_s_at</td>
<td>Hs.368433</td>
<td>Tumor protein DS2</td>
<td>TPDS2</td>
<td>chr8q21</td>
</tr>
<tr>
<td>202954_at</td>
<td>Hs.93002</td>
<td>Ubiquitin-conjugating enzyme E2C</td>
<td>UBE2C</td>
<td>chr20q13.12</td>
</tr>
</tbody>
</table>

*Discussed in the text.
<table>
<thead>
<tr>
<th>Probe set ID</th>
<th>UniGene ID</th>
<th>Gene title</th>
<th>Gene symbol</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>216356_x_at</td>
<td>Hs.458427</td>
<td>BA1-associated protein 3</td>
<td>BAIAP3</td>
<td>chr16p13.3</td>
</tr>
<tr>
<td>212757_s_at</td>
<td>Hs.699281</td>
<td>Calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma</td>
<td>CAMK2G</td>
<td>chr10q22</td>
</tr>
<tr>
<td>213133_s_at</td>
<td>Hs.106070</td>
<td>Cyclin-dependent kinase inhibitor 1C (p57Kip2)</td>
<td>CDKN1C</td>
<td>chr11p15.5</td>
</tr>
<tr>
<td>219400_at</td>
<td>Hs.408730</td>
<td>Contactin associated protein 1</td>
<td>CNTNAP1</td>
<td>chr2q11-qter</td>
</tr>
<tr>
<td>209432_s_at</td>
<td>Hs.522110</td>
<td>cAMP responsive element binding protein 3</td>
<td>CREB3</td>
<td>chr9pter-p22.1</td>
</tr>
<tr>
<td>217200_s_at</td>
<td>Hs.355264</td>
<td>Cytochrome b-561</td>
<td>CYB561</td>
<td>chr17q21</td>
</tr>
<tr>
<td>209178_at</td>
<td>Hs.570079</td>
<td>DEAH (Asp-Glu-Ala-His) box polypeptide 38</td>
<td>DHX38</td>
<td>chr16q21-q22.3</td>
</tr>
<tr>
<td>208810_at</td>
<td>Hs.490745</td>
<td>DnaJ (Hsp40) homolog, subfamily B, member 6</td>
<td>DNAJB6</td>
<td>chr7q36.3</td>
</tr>
<tr>
<td>219759_s_at</td>
<td>Hs.591249</td>
<td>Endoplasmic reticulum aminopeptidase 2</td>
<td>ERAP2</td>
<td>chr5q15</td>
</tr>
<tr>
<td>205021_s_at</td>
<td>Hs.434286</td>
<td>Ezrin</td>
<td>EZR</td>
<td>chr6q25.2-q26</td>
</tr>
<tr>
<td>211742_s_at</td>
<td>Hs.5509</td>
<td>Ecotropic viral integration site 2B</td>
<td>EVI2B</td>
<td>chr17q11.2</td>
</tr>
<tr>
<td>208621_s_at</td>
<td>Hs.712548</td>
<td>Forkhead box N3</td>
<td>FOXN3</td>
<td>chr14q31.3</td>
</tr>
<tr>
<td>204128_s_at</td>
<td>Hs.514590</td>
<td>Hepatocyte growth factor-regulated tyrosine kinase substrate</td>
<td>HGS</td>
<td>chr7q25</td>
</tr>
<tr>
<td>208569_at</td>
<td>Hs.248174</td>
<td>Histone cluster 1, H2ab</td>
<td>HIST1H2AB</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>221435_x_at</td>
<td>Hs.709864</td>
<td>Hydroxypyruvate isomerase homolog (E. coli)</td>
<td>HYI</td>
<td>chr17q25</td>
</tr>
<tr>
<td>209342_s_at</td>
<td>Hs.106070</td>
<td>Iduronate 2-sulfatase (Hunter syndrome)</td>
<td>IDS</td>
<td>chrXq28</td>
</tr>
<tr>
<td>208569_at</td>
<td>Hs.248174</td>
<td>Histone cluster 1, H2ab</td>
<td>HIST1H2AB</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>221435_x_at</td>
<td>Hs.709864</td>
<td>Hydroxypyruvate isomerase homolog (E. coli)</td>
<td>HYI</td>
<td>chr17p13.13</td>
</tr>
<tr>
<td>203753_at</td>
<td>Hs.644653</td>
<td>Pecanex homolog (Drosophila)</td>
<td>PCNX</td>
<td>chr1p34.2</td>
</tr>
<tr>
<td>211668_s_at</td>
<td>Hs.77274</td>
<td>Plasminogen activator, urokinase</td>
<td>PLAU</td>
<td>chr10q24</td>
</tr>
<tr>
<td>41577_at</td>
<td>Hs.45719</td>
<td>Protein phosphatase 1, regulatory (inhibitor) subunit 16B</td>
<td>PP1R16B</td>
<td>chr20q11.23</td>
</tr>
<tr>
<td>203859_s_at</td>
<td>Hs.631841</td>
<td>Parallemmin</td>
<td>PALM</td>
<td>chr1q21.3</td>
</tr>
<tr>
<td>220777_s_at</td>
<td>Hs.189915</td>
<td>Kinesin family member 13A</td>
<td>KIF13A</td>
<td>chr6p21.3</td>
</tr>
<tr>
<td>210626_at</td>
<td>Hs.520819</td>
<td>Insulin induced gene 1</td>
<td>INSIG1</td>
<td>chr7q25</td>
</tr>
<tr>
<td>209099_x_at</td>
<td>Hs.224012</td>
<td>Jagged 1 (Alagille syndrome)</td>
<td>JAG1</td>
<td>chr10p12.1</td>
</tr>
<tr>
<td>203149_s_at</td>
<td>Hs.655455</td>
<td>Poliovirus receptor-related 2 (herpesvirus entry mediator B)</td>
<td>PVRL2</td>
<td>chr1q21.3</td>
</tr>
<tr>
<td>202963_at</td>
<td>Hs.709229</td>
<td>Regulatory factor X, 5 (influences HLA class II expression)</td>
<td>RFX5</td>
<td>chr1q21.3</td>
</tr>
<tr>
<td>211181_x_at</td>
<td>Hs.149261</td>
<td>Runt-related transcription factor 1</td>
<td>RUNX1</td>
<td>chr21q22.3</td>
</tr>
<tr>
<td>201250_s_at</td>
<td>Hs.706748</td>
<td>Solute carrier family 2 (facilitated glucose transporter), member 1</td>
<td>SLC2A1</td>
<td>chr1p35-p31.3</td>
</tr>
<tr>
<td>210993_s_at</td>
<td>Hs.656534</td>
<td>SMAD family member 1</td>
<td>SMAD1</td>
<td>chr4q31</td>
</tr>
<tr>
<td>212625_at</td>
<td>Hs.43812</td>
<td>Syntaxin 10</td>
<td>STX10</td>
<td>chr19p13.13</td>
</tr>
<tr>
<td>203753_at</td>
<td>Hs.644653</td>
<td>Transcription factor 4</td>
<td>TCF4</td>
<td>chr1p34.2</td>
</tr>
<tr>
<td>202085_s_at</td>
<td>Hs.50382</td>
<td>Tight junction protein 2 (zona occludens 2)</td>
<td>TJP2</td>
<td>chr9q31-q33</td>
</tr>
<tr>
<td>206472_s_at</td>
<td>Hs.709205</td>
<td>Transducin-like enhancer of split 3 [E(sp1) homolog, Drosophila]</td>
<td>TLE3</td>
<td>chr3p11.22</td>
</tr>
<tr>
<td>221702_s_at</td>
<td>Hs.288912</td>
<td>TM2 domain containing 3</td>
<td>TM2D3</td>
<td>chr15q22.3</td>
</tr>
<tr>
<td>219253_at</td>
<td>Hs.376722</td>
<td>Transmembrane protein 185B (pseudogene)</td>
<td>TMMEM185B</td>
<td>chr4q31</td>
</tr>
<tr>
<td>202510_s_at</td>
<td>Hs.525607</td>
<td>Tumor necrosis factor, alpha-induced protein 2</td>
<td>TNFAIP2</td>
<td>chr2q14.2</td>
</tr>
<tr>
<td>209295_at</td>
<td>Hs.521456</td>
<td>Tumor necrosis factor receptor superfamily, member 10b</td>
<td>TNFRSF10B</td>
<td>chr4q31.3</td>
</tr>
<tr>
<td>221962_x_at</td>
<td>Hs.700368</td>
<td>Ubiquitin-conjugating enzyme E2H (UBC8 homolog, yeast)</td>
<td>UBE2H</td>
<td>chr7q32</td>
</tr>
<tr>
<td>214792_x_at</td>
<td>Hs.25348</td>
<td>Vesicle-associated membrane protein 2 (synaptobrevin 2)</td>
<td>VAMP2</td>
<td>chr7p13.1</td>
</tr>
<tr>
<td>212171_x_at</td>
<td>Hs.73793</td>
<td>Vascular endothelial growth factor A</td>
<td>VEGFA</td>
<td>chr6p12</td>
</tr>
<tr>
<td>202932_at</td>
<td>Hs.194148</td>
<td>V-yes-1 Yamaguchi sarcoma viral oncogene homolog 1</td>
<td>YES1</td>
<td>chr15q31.3-33.1</td>
</tr>
<tr>
<td>219571_s_at</td>
<td>Hs.431471</td>
<td>Zinc finger protein 12</td>
<td>ZNF12</td>
<td>chr7p22.1</td>
</tr>
<tr>
<td>219540_at</td>
<td>Hs.460645</td>
<td>Zinc finger protein 267</td>
<td>ZNF267</td>
<td>chr16p11.2</td>
</tr>
</tbody>
</table>

*Discussed in the text.*
Figure 2. Unique biological processes in hematopoietic stem cells from different anatomic locations. Selected biological processes overrepresented by genes enriched or downregulated in CD34+ CD38− cells isolated from (A) cord blood, (B) bone marrow or (C) peripheral blood. Probe sets differentially expressed (either up or down) in each hematopoietic stem cell subtype were subjected to gene set enrichment analysis using the DAVID 2008 Web tool. These categories were selected from the Biological Process organizing principle in the Gene Ontology project. Categories in red indicate increased expression; those in green indicate decreased expression. The number of genes, their percentage in cell type-specific probe sets, and p values for each category significantly (p < 0.05) overrepresented are listed. NLS = nuclear localization sequence; Ag = antigen.
of a specific sub-network in CB HSCs (Figures 3B–D), suggesting their critical roles in maintaining the basic properties and biological function of CB HSCs.

GATA1 regulates a specific sub-network in CB HSCs (Figure 3D). This transcription factor was identified as an erythroid-specific gene essential for erythrocytic differentiation at relatively early stages [19]. KLF1 is also found in this GATA1 sub-network. GATA1 and KLF1 (EKLF) are expressed in precursors and are definitive differentiation-associated genes [20], while GATA1 also cooperates with RUNX1 in megakaryocytic differentiation [15]. The presence of GATA1 indicates that HSCs from CB may contain relatively more lineage-specific precursors than those isolated from BM or PB.

**Genetic networks of BM HSCs**
Hepatic growth factor, \( \text{KRAS} \), \( \text{CTNNB1} \) (\( \beta \)-catenin), \( \text{CDH1} \) (also known as E-cadherin), \( \text{PAX3} \) and \( \text{CDC2} \) are hub genes in BM HSCs (Figure 4). Hepatic growth factor and myeloperoxidase, which can both support self-renewal and keep HSCs in an undifferentiated state [13], are highly expressed in this subtype of HSC (Table 3 and Figure 4).

**Discussion**
Somatic and embryonic stem cells maintain the ability to self-renew and differentiate and, therefore, possess great potentials in regenerative medicine for healing tissues that are unable to repair themselves. In this study, we performed an extensive comparative transcriptome and gene network analyses of CD34\(^+\) HSCs from different sources. The results of these analyses will help to unravel the riddle of HSCs and contribute to developments in cell-based therapy.

In addition to identifying specific genes, we also applied refined computational methods to highlight key functional networks. There is increasing recognition that a systematic approach is necessary to view the overall molecular events responsible for a given biological process [21,22]. In this study, we applied systems biology tools to reveal the functional influences of differential gene expression. Knowledge of the enriched or repressed GO categories in each HSC subtype will improve understanding of their basic biological properties and aid their development as potential clinical applications.
Figure 3. Genetic networks as a framework for the interpretation of cord blood CD34+ CD38− stem cell biology. (A) Functional networks composed of multiple genes, some of which (such as MYC, NMYC, GATA1, YWHAZ and JUN) are hub genes with the
Cell cycle regulators were upregulated in all three HSC types, indicating a restricted cell cycle regulation. However, in BM HSCs, genes involved in mitosis, chromosome condensation and mitotic sister chromatid segregation were more abundant, indicating a relatively active cell cycle and replication. The downregulation of cell cycle arrest genes in BM HSCs was consistent with this theme. Genes involved in cell motility and small GTPase (Rho) signal transduction were downregulated, indicating an inert and immobile property of HSCs inside BM.

In PB HSCs, the regulation of transcription from RNA polymerase II promoter and the NF-κB and MAPKKK cascades were active, indicating the physiological activities of these stem cells. For example, angiogenesis and vascular development-related genes were highly expressed in PB HSCs, which are more motile, suggesting that they should aid vascular morphogenesis and could be useful for the treatment of vascular diseases.

Among the three HSC subtypes analyzed, HSCs from CB appear to provide an excellent source for transplantation. Genes involved in immunogenic responses, such as major histocompatibility complex class I-mediated endogenous antigen processing/presentation and cell defense response, are least active in this subtype, suggesting a lower risk of transplant rejection or graft-versus-host disease. HSCs from PB or BM, especially those from PB, expressed more genes associated with active antigen presentation, response to biotic stimuli, and lymphocyte (B cells) activation, and would, therefore, be less suitable than CB HSCs in clinical applications.

Our analyses had certain limitations. The functional networks were mapped based on predetermined knowledge databases (GO and IPA), in which molecular interactions were established under various physiological or pathologic conditions. It is known that molecular functions vary according to the cellular and tissue contexts, and the stemness genes may, therefore, not interact in a similar way in human CD34+ cells. Functional analyses are needed for the rigorous evaluation of individual gene interactions deduced by in silico modeling.

Acknowledgments

The authors acknowledge the technical services provided by the Microarray and Gene Expression Analysis Core Facility of the National Yang-Ming University.
Figure 4. The major genetic network in bone marrow hematopoietic stem cells. Hub genes are labeled.

VGH Genome Research Center (VYMGC). This work was supported by the Mackay Memorial Hospital (MMH-HB-97-05) and the National Science Council (NSC97-3111-B-010-004).

References


