



REGULAR ARTICLE

The *HOX* Code as a “biological fingerprint” to distinguish functionally distinct stem cell populations derived from cord blood

Stefanie Liedtke^{a,1,2}, Anja Buchheiser^{a,2,3}, Julia Bosch^{a,2,4},
Frank Bosse^{b,5,6}, Fabian Kruse^{b,5,6}, Xiaoyi Zhao^{a,2,7},
Simeon Santourlidis^{a,2,7}, Gesine Kögler^{a,*,8}

^a Institute for Transplantation Diagnostics and Cell Therapeutics, Heinrich-Heine-University Medical Center, Düsseldorf, Germany

^b Division for Molecular Neurobiology, Department of Neurology, Heinrich-Heine-University Medical Center, Düsseldorf, Germany

Received 19 November 2009; received in revised form 3 March 2010; accepted 18 March 2010

Abstract Mesenchymal stem cells (MSC) have been isolated from almost every adult tissue. In cord blood (CB), different non-hematopoietic CD45⁻, CD34⁻ adherent cell populations can be generated: the cord blood derived MSC (CB-MSC), that behave almost like MSC from bone marrow (BM-MSC), and unrestricted somatic stem cells (USSC) which show a distinct differentiation potential into all three germ layers. However, distinguishing these populations easily by molecular markers is still a concern. In this study we were able to present the HOX expression pattern of USSC, CB-MSC and BM-MSC, which in fact allows a discrimination of these populations.

Briefly, RT-PCR analysis of the HOX code revealed a high similarity between BM-MSC and CB-MSC, which are both HOX-positive, whereas USSC resembled H9 embryonic stem cells HOX-negative. Especially HOXA9, HOXB7, HOXC10 and HOXD8 are good candidate markers to discriminate MSC from USSC. Thus, our data suggest that the “biological fingerprint” based on the HOX code can be used to distinguish functionally distinct MSC populations derived from bone marrow and cord blood.

© 2010 Elsevier B.V. All rights reserved.

* Corresponding author. Institute for Transplantation Diagnostics and Cell Therapeutics, Heinrich-Heine-University Medical Center, Moorenstrasse 5, D-40225 Düsseldorf, Germany. Fax: +49 211 8104340.

E-mail addresses: liedtke@itz.uni-duesseldorf.de (S. Liedtke), buchheiser@itz.uni-duesseldorf.de (A. Buchheiser), bosch@itz.uni-duesseldorf.de (J. Bosch), bosse@uni-duesseldorf.de (F. Bosse), fabian.kruse@uni-duesseldorf.de (F. Kruse), zhao@itz.uni-duesseldorf.de (X. Zhao), santourlidis@itz.uni-duesseldorf.de (S. Santourlidis), koegler@itz.uni-duesseldorf.de (G. Kögler).

¹ Conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing.

² Fax: +49 211 8104340.

³ Data analysis and interpretation.

⁴ Real time PCR analysis.

⁵ Statistical analysis of Affymetrix data.

⁶ Fax: +49 211 8118411.

⁷ Epigenetic analysis.

⁸ Conception and design, administrative support, final approval of manuscript.

Introduction

The presence of primitive non-hematopoietic stem/progenitor cells in cord blood was reported by our group (Kogler et al., 2005) and confirmed by others (Chan et al., 2007; Chang et al., 2006; Kern et al., 2006; Kim et al., 2005). In our lab, characterization of unrestricted somatic stem cells (USSC) from CB that have unique proliferation capacities and can be differentiated *in vitro* into mesodermal, endodermal and ectodermal lineages was performed (Greschat et al., 2008; Sensken et al., 2007; Kogler et al., 2004; Trapp et al., 2008). Although USSC possess several overlapping features with MSC derived from CB or from bone marrow (BM), such as immunophenotype, osteogenic and chondrogenic *in vitro* and *in vivo* differentiation potential, USSC differ from BM-MSCs with regard to their immunological behavior (van den Berk et al., 2009; Winter et al., 2008) and their neural differentiation potential (Greschat et al., 2008; Kogler et al., 2004). MSC can be differentiated into osteoblasts, adipocytes or chondrocytes in culture or *in vivo* (Prockop, 1997; Prockop et al., 2003). Over the last years a characterization of a large number of CB-derived cell lines in terms of their adipogenic, neural and endodermal differentiation potential was achieved. Based on these results, a classification of CB-derived cell lines regarding their adipogenic differentiation potential was suggested (Kogler et al., 2009). In an actual work by Jansen et al., functional differences between USSC, BM-MSCs and AdAS were analyzed on global transcriptome level and several differentially expressed genes were defined (Jansen et al., 2009), but markers capable of distinguishing between USSC and CB-MSCs are still lacking. Therefore, several questions were addressed here: Which possible molecular markers can distinguish between USSC and CB-MSCs derived from cord blood, and what kind of impact could these markers have on biological functions or transplantation.

Homeobox genes encode homeodomain-containing transcription factors determining the positional identity along the anterior-posterior body axis of animal embryos (Krumlauf, 1994). In humans, the 39 known *HOX* genes are distributed among four clusters *HOXA* to *HOXD*, located in chromosomes 7, 17, 12 and 2, respectively. *HOX* genes are expressed sequentially 3' to 5' along the anterior-posterior axis during embryogenesis, termed "temporal and spacial colinearity" (Kmita and Duboule, 2003). The typical *HOX* code of a cell describes the specific expression of functionally active *HOX* genes in distinct tissues (Kessel and Gruss, 1991). Recent findings revealed that this intrinsic *HOX* code of a cell reflects a continuation of embryonic patterning (Morgan, 2006), and several studies have reported on specific *HOX* gene expression in adult human tissues (Yamamoto et al., 2003; Takahashi et al., 2004). Ackema et al. recently described in mice that mesenchymal stroma cells from different organs are characterized by distinct topographic *HOX* codes (Ackema and Charite, 2008). Ackema et al. reported that even if there is a broad similarity of all MSC tested, these can be subdivided by their specific topographic *HOX* code depending on the tissue of origin. Another study by Chang et al. revealed that fibroblasts from different anatomic sites across the human body express distinct *HOX* patterns (Chang et al., 2002). In addition Chang et al. presented data that more than 1000 genes are differentially

expressed due to the anatomical origin of the cell. Others were able to confirm in mice that the typical *HOX* code can be sufficient to indicate the positional identity of a cell and that the position-specific *HOX* code is independent of the age of the donor (Rinn et al., 2006, 2008). In an actual work by Hwang et al., *HOXC10* was defined as a potential marker for discriminating between human amnion- and decidua-derived mesenchymal stem cells (Hwang et al., 2009).

However, analyzing *HOX* expression patterns in different cell types seems to be useful to clarify the cell origin and subdivide similar cell types, like mesenchymal stem cells, cells regarding their tissue of origin. Thus an expression analysis of all 39 known human *HOX* genes in functionally distinct adherent non-hematopoietic cell populations derived from cord blood was performed and possible markers to distinguish between them were defined in this study.

Results

Analysis of Affymetrix chips revealed *HOX* genes as potential molecular markers to distinguish between USSC and CB-MSCs

The primary basis of this work was a DNA-array (Affymetrix) including in total 5 CB-derived cell lines (USSC $n=3$; CB-MSCs $n=2$). To first assess relatedness between the CB-derived stem cell populations, a principal component analysis (PCA) was performed. As depicted in Fig. 1, the two independent samples of CB-MSCs (red) grouped together, while the three USSC samples (blue) are more divergent but in one plane. The divergency between the cell populations reflects the biological heterogeneity of the samples. These already preliminary data indicate that two different populations exist in cord blood, which can be distinguished by their differentiation potential (Kogler et al., 2009). Further analysis revealed 271 probesets, which are significantly differentially expressed between USSCs and CB-MSCs. Of these, 158 probesets were upregulated and 113 probesets were downregulated in USSC. Subsequently, the Affymetrix Ids, according to the 271 differentially expressed probesets, were subjected to a Functional Annotation Chart analysis offered by the DAVID Bioinformatics Resources 2008 homepage (<http://www.david.abcc.ncifcrf.gov/>) (Dennis et al., 2003; Sherman et al., 2007; Huang da et al., 2009). The results of this analysis are shown in Table 1. Chart Report is an annotation-term-focused view, which lists annotation terms and their associated genes under study. This tool is useful to discover enriched functional-related gene groups. Analysis outcome revealed that the detected GO terms are mainly functionally related to transcription factors and developmental processes. Checking the detailed gene lists of the top 10 GO terms, it was found that the first 6 functional groups comprise 17 *HOX* genes, a list of which is presented in the supplement Table S2. Consequently, an analysis of the expression data of all 39 *HOX* genes in USSC and CB-MSCs followed and all hits belonging to *HOX* genes from the Affymetrix chips were extracted. Fig. 2 depicts the relative expression values taken from the Affymetrix chip of all 39 *HOX* genes. The values from USSC ($n=3$) and CB-MSCs ($n=2$) were compared to each other. Again, USSC revealed almost no expression of *HOX* genes, whereas CB-MSCs cell

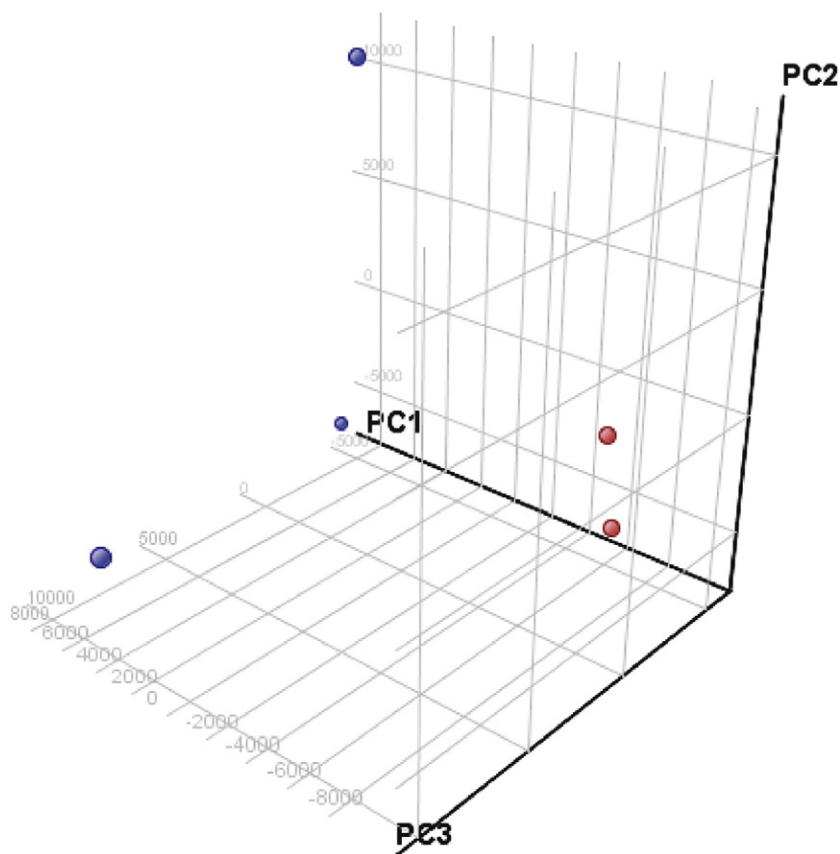


Figure 1 Principal Component Analysis (PCA) of global gene expression profile of USSC and CB-MSc. Red: CB-MSc, blue: USSC. Principal Component Analysis (PCA) calculates the PCA scores and visually represents them in a 3D scatter plot. The first, second and third principal component (PC) capture the sample populations' variability respectively. The first PC separates maximally between CB-MSc and USSC. The second and third PC separates the biological variabilities within the two different populations.

lines showed high expression of numerous *HOX* genes of each cluster. The most prominent ones in CB-MSc are *HOXA9*, *HOXA10*, *HOXB2*, *HOXB7*, *HOXC6*, *HOXC10* and *HOXD8*. *HOXC6* and *HOXC10* reached the highest expression with relative expression values at 7000 (*HOXC10*) and 10000

(*HOXC6*). These preliminary results ascertain that the "biological fingerprint" of a specific cell type is reflected by a typical *HOX* code, and, as a consequence thereof, the *HOX* code can be used as a molecular marker to discriminate USSC from CB-MSc.

Table 1 Functional Annotation Chart

Term	Count	%	PValue	Benjamini
GO:0043565 ~ sequence-specific DNA binding	24	10,48%	6,27E-09	1,80 E-05
GO:00 03700 ~ transcription factor activity	31	13,54%	1,12E-07	1,61 E-04
GO:0007275 ~ multicellular organismal development	54	23,58%	6,16E-07	0,003227888
GO:0032502 ~ developmental process	64	27,95%	8,57E-06	0,022260891
GO:0030528 ~ transcription regulator activity	33	14,41%	6,09E-05	0,056751315
GO:0032501 ~ multicellular organismal process	64	27,95%	3,95E-05	0,066772989
GO:0009653 ~ anatomical structure morphogenesis	29	12,66%	1,09E-04	0,133539358
GO:0048856 ~ anatomical structure development	44	19,21%	2,33E-04	0,217104075
GO:0009887 ~ organ morphogenesis	15	6,55%	4,63E-04	0,293219656
GO:0001501 ~ skeletal development	11	4,80%	4,14E-04	0,304147747

271 probeset IDs that were found to be significantly differentially expressed by USSC and CB-MSc were further analyzed by the DAVID Functional Annotation Chart Tool to identify GO terms with unbalanced distribution of this query list compared to a background list (here: Affymetrix HG-U133_Plus_2 probeset list). The top 10 GO terms are shown.

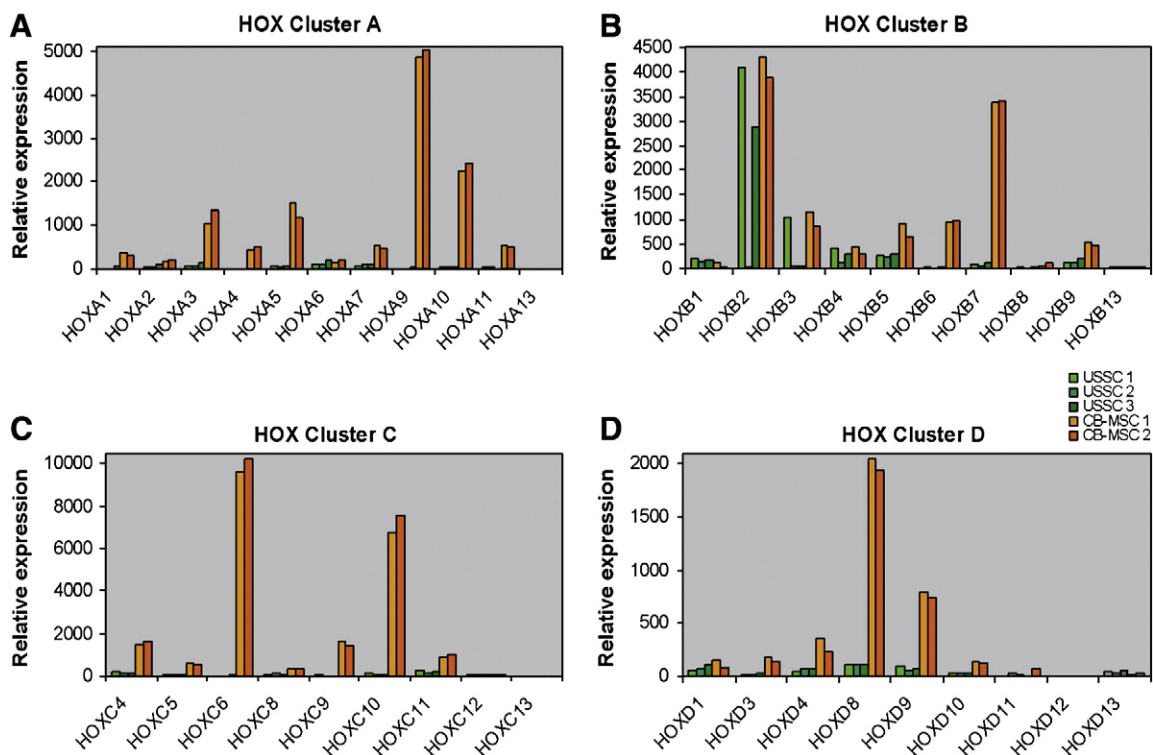


Figure 2 Comparison of relative mRNA expression levels of 39 *HOX* genes in USSC and CB-MSC determined by careful data analysis of our Affymetrix HG-U133_Plus_2 expression arrays. Relative expression values are shown for *HOX* Cluster A (A), B (B), C (C) and D (D). If more than one probeset was used in the array, mean values are depicted in the graph. Raw data is available in Table S1. The mRNA steady state levels are represented for each sample of USSC ($n=3$) and CB-MSC ($n=2$).

Epigenetic status of *HOX* genes in USSC in comparison to ESC

In addition to the array data, the epigenetic status of the *HOX* genes in USSC, CB-MSC and H9 embryonic stem cells was determined. Especially USSC in comparison to H9 cells, which preserve their pluripotency by repressing their *HOX* genes, was assessed (Soshnikova and Duboule, 2008). Based on the Affymetrix data it would be expected to see methylation ranging over all 4 clusters, which would explain as a result the lack of *HOX* gene expression in USSC and H9 cells. Roche NimbleGen provides sensitive and specific DNA methylation microarrays that allow a precise identification of methylated DNA regions across whole genomes or within biologically focused regions including promoters and CpG islands (Mohn et al., 2008; Weber et al., 2007; Zilberman et al., 2007; Rauch et al., 2007; Yasui et al., 2007). The DNA methylation analysis of all four *HOX* clusters is depicted in Fig. 3. The methylation status of H9 embryonic stem cells in comparison to USSC is very similar and corresponds to the expression data deriving from affymetrix chips (Table S3). Except for some cases, the peaks reflecting the methylation status of the genomic region is comparable between H9 embryonic stem cells and USSC. The *HOXA* cluster is highly methylated in all populations. The CB-MSC seem to be higher methylated as compared to USSC, which would not be expected by the affymetrix data, because many of the *HOXA* genes are detectable. The methylation of the *HOXA* cluster is significantly lower in USSC as compared to H9. Nevertheless,

based on affymetrix data, transcription of the *HOXA* genes is rarely detectable in USSC and H9. The clusters B-D are less methylated in all three populations and would allow, in some cases, an expression of several *HOX* genes. This might explain why CB-MSC seem to express several *HOX* genes of all four clusters based on the affymetrix data whereas USSC and H9 do not. Still, non-methylated parts within the *HOX* gene cluster do not directly refer to transcription with respect to other regulative mechanisms of transcription, like posttranscriptional regulation by splicing, or posttranslational by modification and ubiquitination. Due to the fact that the epigenetic analysis can only serve as a prediction for the possible expression of genes, RT-PCR analysis is mandatory to define putative markers in order to distinguish USSC from CB-MSC.

Determination of specific *HOX* expression pattern by RT-PCR

To validate the preliminary array data, primers specific for each *HOX* gene (Table S1) were designed. The expression of each known *HOX* gene was detected by RT-PCR (Table 2). This analysis included USSC cell lines ($n=7$), CB-MSC cell lines ($n=7$) and BM-MSC cell lines ($n=7$) to show the differences between these populations. In addition, several common cell lines like HEK, Hela, NHDF, H9 and nTERA-2, as well as adipose tissue-derived MSC and some distinct tissues like femoral muscle, brain and liver, were analyzed. After

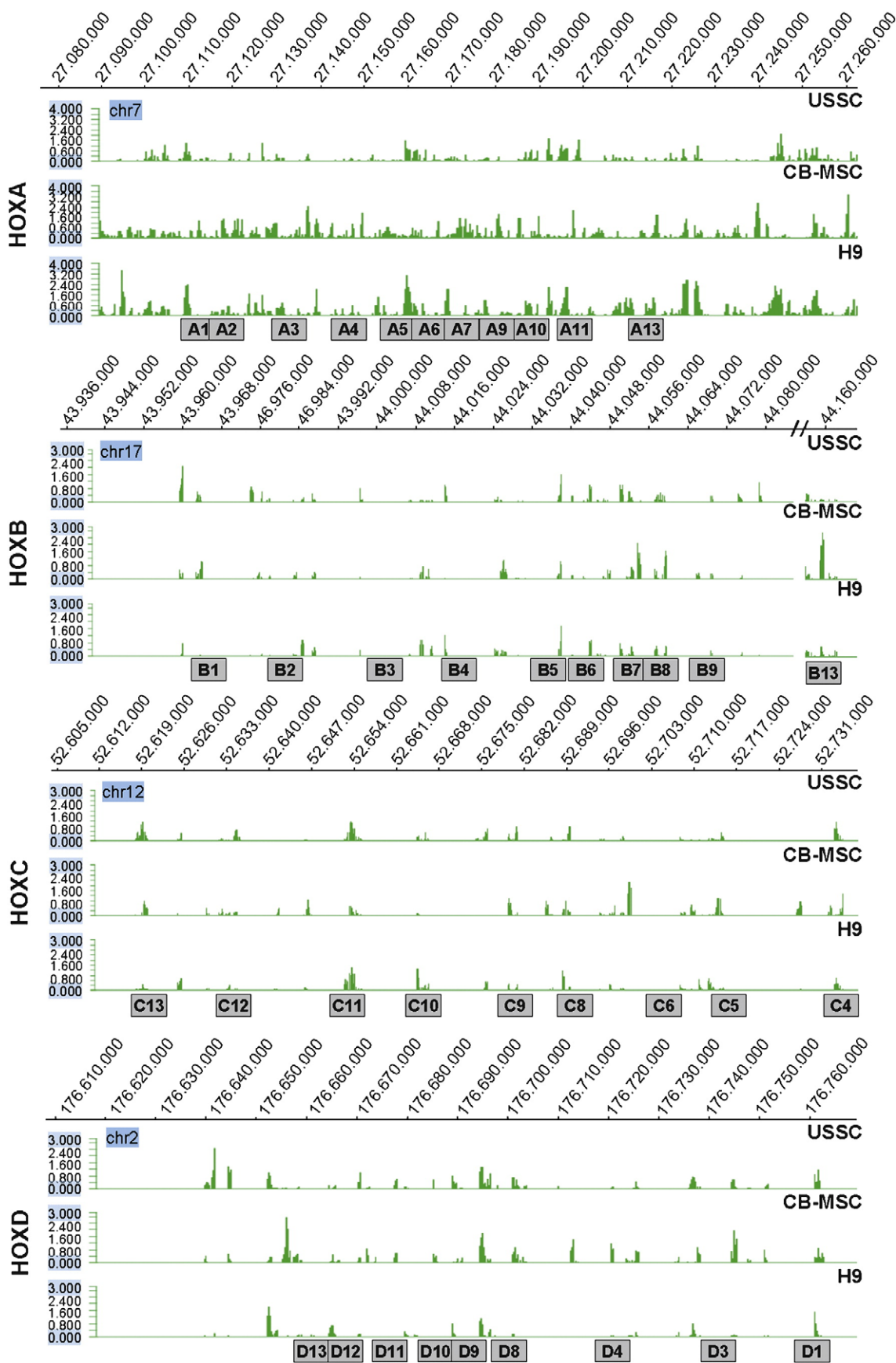


Figure 3 DNA methylation analysis of the 4 *HOX* loci by NimbleGen. A global overview of the DNA methylation status of the *HOX* gene clusters A, B, C and D of USSC, CB-MSK and H9 ES cells is given. Grey boxes display the localization of the *HOX* genes within the cluster.

visualization on agarose gels, the band intensity was defined either as negative -, weakly expressed (-) or as one of three levels positively expressed (+, ++ and +++). The RT-PCR results revealed that BM-MSC samples as well as CB-MSC samples share a high homology regarding their *HOX* code. With the exception of the CB-MSC samples 4, 5 and 7, which express fewer *HOX* genes in comparison to the other CB-MSC samples. This can reflect biological variations, or one can hypothesize that other MSC-like population are present in cord blood. The most highly expressed *HOX* genes in BM-MSC as well as in CB-MSC are *HOXA5*, 9 and 10, *HOXB6* and 7, *HOXC4-10* and *HOXD3-4*, and 8. In both MSC populations, the 5' positioned *HOX* genes are barely or not expressed in all four *HOX* clusters. The 3' *HOX* genes are mainly expressed in Cluster A and C, while in Cluster D they show only partial expression. In the *HOXB* cluster the expression of 3' *HOX* genes is almost absent. Further detailed examination of each single *HOX* gene revealed *HOXD9* as a putative distinguishing feature between CB-MSC (*HOXD9*⁺) and BM-MSC (*HOXD9*⁻). However, both MSC populations show quite a similar expression pattern. Compared to both MSC populations, USSC show almost no expression of *HOX* genes was detected. Only in some cases a very low expression of different *HOX* genes detected. The only gene we expected to be highly expressed was *HOXB2* based on the Affymetrix data (Fig. 2). USSC line 2 revealed no expression of *HOXB2* (Table S3). Since only one probe was used in affymetrix (205453_at), and taking into account the possible heterogeneity of this special gene within the USSC populations, *HOXB2* expression might be not representative in USSC. However, after testing two different primer pairs, RT-PCR analysis revealed expression of *HOXB2* only in adipose tissue derived MSC and HEK and never in USSC showing that the primer pair used in this analysis worked. The USSC share the *HOX*-negative expression pattern with the embryonal cell line H9 and the embryonal teratocarcinoma cell line nTera-2. It is well documented that embryonic stem cells preserve their pluripotency by repressing their *HOX* genes (Soshnikova and Duboule, 2008). The *HOX*-negative status of USSC therefore might reflect the higher immaturity of this cell type in comparison to MSC and could explain the distinct differentiation potential. Additionally, further tissue samples were tested from femoral muscle, brain and liver where muscle possessed a *HOX*-positive profile and brain and liver a *HOX*-negative status.

In analogy to the Affymetrix data the most prominent genes in CB-MSC are *HOXA9*, *HOXA10*, *HOXB7*, *HOXC6*, *HOXC10* and *HOXD8*.

Definition of *HOXA9*, *HOXB7*, *HOXC10* and *HOXD8* as molecular markers to distinguish USSC and CB-MSC by quantitative RT-PCR

Although *HOXA10* and *HOXC6* revealed the highest expression based on the Affymetrix data and the RT-PCR data, these genes were excluded from further analysis due to alternative splice variants. The primers used in this experiment detect both splice variants. The function of these splice variants has not been elucidated yet in detail (Shimeld et al., 1993; Lawrence et al., 1995; Benson et al., 1995), therefore *HOXA9*, *HOXB7*, *HOXC10* and *HOXD8* were defined as good potential molecular markers to discriminate between USSC and CB-MSC.

In a simple RT-PCR approach, 20 different cell lines derived from cord blood and 3 different BM-MSC lines were tested (data not shown). Of the 20 CB-derived cell lines tested, 3 were classified as CB-MSC and 17 as USSC cell lines based on their adipogenic differentiation potential, which is tested routinely in our lab as quality control (Kogler et al., 2009) (Fig. S1). Based on the 4 *HOX* genes tested, it was possible to confirm that USSC as well as embryonic stem cells are negative for the 4 *HOX* markers tested, whereas BM-MSC, CB-MSC and highly differentiated adipose tissue-derived MSC are positive for these markers. Hence the differentiation potential of cell lines derived from cord blood can be linked to the expression of the *HOX* genes. Moreover, testing the 4 molecular *HOX* markers facilitates the characterization of cord blood derived cell lines in comparison to the time consuming determination of the differentiation potential (Sensken et al., 2007; Kogler et al., 2004, 2009).

In a next step the data was confirmed in a quantitative real time PCR approach. As a negative control the embryonal teratocarcinoma cell line nTera-2 was employed, which was negative for *HOX* gene expression in the RT-PCR approach. The positive control was the HEK cell line, known to be positive for the majority of *HOX* genes. All samples applied here were normalized to the HEK cell line. Fig. 4 depicts the results of the quantitative PCR. Comparison of the expression values of each cell type tested showed that USSC cell lines are *HOX*-negative on quantitative level, confirming the preliminary RT-PCR results. By comparison, CB-MSC and BM-MSC expressed many *HOX* genes. *HOXC10* reached the highest expression of all markers tested, in fact expressing 7-30 fold more as compared to the HEK cell line. By contrast, *HOXA9*, *HOXB7* and *HOXD8* achieved moderate levels of gene expression, comparable to the positive control. Nevertheless, evaluating the gene expression of the four markers tested on quantitative level definitively confirmed the potential of the *HOX* markers to distinguish CB-MSC from USSC.

Discussion

MSC populations can be isolated from numerous human tissues (Kogler et al., 2004; Zuk et al., 2001; Goodwin et al., 2001), but the relatedness of these cells remains largely unknown. Cord blood is currently used as an alternative to bone marrow as a source of stem cells for hematopoietic reconstitution after ablation. It is also under intense preclinical investigation for a variety of indications ranging from stroke, to limb ischemia, to myocardial regeneration (Riordan et al., 2007; Hu et al., 2006; Brzoska et al., 2006; Leor et al., 2006; Newcomb et al., 2006). Key questions concerning the differences between cord blood-derived cell populations regarding their origin, differentiation potential, tumorigenicity, and availability still need to be answered (Buchheiser et al., 2009). Over the last years, numerous CB-derived cell lines were generated and characterized based on their differentiation potential in our lab. Beyond the biological small variations between the cell lines, it was detected that adipogenic differentiation is present in CB-MSC, but absent in USSC (Kogler et al., 2009). In order to distinguish easily between these two populations derived from cord blood, it is mandatory to define cell type-specific markers. Here, we were able to present the *HOX* code as a

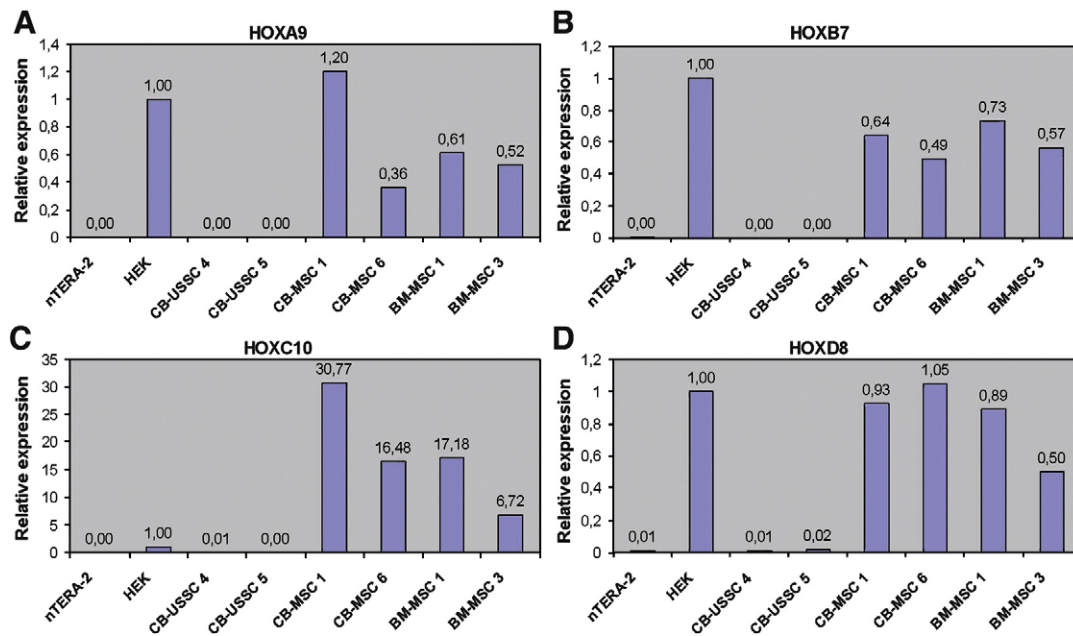


Figure 4 Representative expression patterns of the new putative marker genes *HOXA9*, *HOXB7*, *HOXC10* and *HOXD8*. Differential expression of the four *HOX* genes were verified by means of real-time PCR. Relative changes in gene expression were calculated using the $\Delta\Delta C_t$ -method with GAPDH as internal standard and normalized to human embryonic kidney (HEK) cells.

“biological fingerprint” to distinguish functionally distinct MSC populations derived from cord blood. *HOX* gene expression especially in BM-MSC, CB-MSC and USSC was evaluated by RT-PCR. Several common cell lines like HEK, HeLa, NHDF, H9 and nTera-2, as well as adipose tissue-derived MSC and some distinct tissues like femoral muscle, brain and liver were included in this study to complete the analysis. Furthermore, *HOXA9*, *HOXB7*, *HOXC10* and *HOXD8* were defined as potential molecular markers, which are highly differentially expressed in USSC and CB-MSC. In an actual work by Jansen et al., functional differences between USSC, BM-MSC and AdAS were analyzed on global transcriptome level (Jansen et al., 2009). Within the top 25 genes that are upregulated in BM-MSC compared to USSC, they found *HOXC10* four fold differentially expressed, which could also be confirmed by the data provided here. In a recent work by Hwang et al., *HOXC10* was defined as a potential marker to distinguish amnion- and decidua-derived mesenchymal stem cells (Hwang et al., 2009). Our data are in agreement with the findings of Hwang et al., therefore, determining the function of *HOXC10* in these cells is an interesting task for future prospects. In addition, the determination of the specific *HOX* codes revealed a high similarity between BM-MSC and CB-MSC. These two cell populations could be distinguished by only some *HOX* genes, namely *HOXD9* and *HOXD10* as the most prominent ones in RT-PCR experiments. Taken together with the results of the affymetrix chips (Table S3) only *HOXD9* is expressed in CB-MSC and is absent in BM-MSC. Expression of many *HOX* genes was also present in adipose tissue-derived MSC confirming the data by Lee et al. (Lee et al., 2004), who described that the genetic expression

profiles of BM-MSC and adipose tissue-derived MSC is similar. Whether this correlation reflects the ability of MSC to differentiate into adipocytes remains elusive. The expression of *HOX* genes was mainly absent in USSC, H9 and nTera-2 cells. In addition, brain and liver tissues were *HOX*-negative. Taken together, we were able to document here that CB-MSC resemble BM-MSC and that USSC are more similar to embryonic stem cells based on their *HOX* code without expressing the specific ES-cell markers Oct4, Sox2 and Nanog (Kogler et al., 2009; Liedtke et al., 2007, 2008; Buchheiser et al., 2008).

The aspect of the *HOX*-negative status is also an important factor relevant for transplantation:

In a recent publication, interesting biological and functional aspects of *HOX* genes were highlighted (Wang et al., 2009). One is the influence of the *HOX* code in the process of adult bone regeneration (Leucht et al., 2008) and wound healing (Creuzet et al., 2002). In the work of Leucht et al. it was documented that *HOX*-negative mandibular skeletal progenitor cells adopt a *HOX*-positive profile when transplanted into a *HOX*-positive tibial defect. Conversely, *HOX*-positive tibial skeletal progenitor cells maintain the *HOX* status even when transplanted into a *HOX*-negative mandibular defect (Leucht et al., 2008). In this context the *HOX*-negative status of USSC in comparison to CB-MSC is important. In a no injury *in utero* sheep model it was shown that USSC have the potential to differentiate into parenchymal liver cells (Kogler et al., 2004). Our analysis revealed that USSC as well as liver were *HOX*-negative. Taking into account that *HOX*-negative stem or precursor cells can adopt the *HOX*-positive status if they are transplanted in *HOX*-positive

Notes to Table 2:

Band intensities were depicted with +++ for highly expressed, ++ for strongly expressed, + for expressed, (-) for weakly expressed and – for not expressed. Abbreviations: a. (adult), f. (fetal), HEK (human embryonic kidney), NHDF (normal human dermal fibroblasts), BM-MSC (bone marrow mesenchymal stem cells), CB-MSC (cord blood mesenchymal stem cells), USSC (unrestricted somatic stem cells).

tissues but not vice versa (Leucht et al., 2008; Creuzet et al., 2002), it can be hypothesized that USSC would have a higher regenerative potential in comparison to CB-MSC or other MSC derived from bone marrow. As reflected by their typical *HOX* code, it can be speculated that matching the *HOX* code of the transplanted stem cell with the host tissue is mandatory to achieve engraftment and regenerative healing.

Currently, *HOX* gene expression has been mainly focused on embryonic patterning in drosophila (Akam, 1998). In the murine and human system, the literature comprises a lot of information about the association of *HOX* gene expression and cancer development, but the identification of critical *HOX* subsets and their functional role in cancer onset and maintenance requires further investigation (McGonigle et al., 2008). In some recent publications, the *HOX* code was mainly used as a “biological fingerprint” of different cell types (Ackema and Charite, 2008; Rinn et al., 2008; Hwang et al., 2009). In 2004 Takahashi et al. (Takahashi et al., 2004) presented expression profiles of *HOX* genes in human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR. Their results showed that *HOX* genes are organ-specifically expressed. However, as presented here, the “biological fingerprint” can be applied to define functionally distinct MSC populations derived from cord blood and other tissues. Regarding the fact that cord blood stem cells are a valuable source of neonatal cells, the definition of the *HOX* status might be important for transplantations.

The 4 markers defined in this study are now used routinely in our lab to prospectively define the cell lines generated. We hope that our findings will support the idea that the *HOX* code is the “biological fingerprint” of a cell useful to determine and distinguish different cell types, and in the case of adult stem cells as demonstrated here can provide additional information about a putative regenerative potential important for transplantations.

Materials & Methods

Generation and Expansion of CB-derived cells

USSC and CB-MSC were generated by the same method. Classification of the adherent cells into USSC and CB-MSC was only possible after generation by determining the adipogenic differentiation potential (Fig. S1). CB was collected from umbilical cord vein with informed consent of the mother. MNC were obtained by ficoll (Biochrom, density 1.077 g/cm³) gradient separation followed by ammonium chloride lysis of RBCs. 5–7 10⁶ CB MNC /ml were cultured in T75 culture flasks (Corning) in DMEM low glucose (Cambrex) with 30% FCS (Perbio), 10⁻⁷ M dexamethasone (Sigma-Aldrich), penicillin / streptomycin and L-glutamine (PSG;Cambrex). When colonies were detected, cells were expanded without dexamethasone in a closed system applying cell stacks (Corning). Cord blood derived stem cells (USSC and CB-MSC) were incubated at 37 °C in 5% CO₂ in a humidified atmosphere. Reaching 80% confluence, cells were detached with 0.25% trypsin (Cambrex) and replated 1:3. Each cell line generated was obtained from an individual cord blood sample since the frequency of the cells is very low (Kogler et al., 2006).

Total RNA extraction and Reverse Transcription

Total RNA was extracted from cell lines and cell samples in a 40 µl volume applying the Rneasy Kit (Qiagen) according to the manufacturer’s instructions. Determination of RNA concentrations was carried out by applying a Nanodrop device (NanoDrop Technologies). Reverse transcription was performed for 1 h at 50 °C using the First-strand cDNA Synthesis Kit (Invitrogen) and the enclosed oligo(dT)₂₀ Primer. About 500 ng total RNA was converted into first-strand cDNA in a 20 µl reaction. All control reactions provided with this system were carried out to monitor the efficiency of cDNA-synthesis. Prior to PCR, the completed first-strand reaction was heat-inactivated at 85 °C for at least 10 min. Finally, cDNA was treated with RNaseH according to the manufacturer’s protocol.

RT-PCR and real time PCR

RT-PCR was carried out by designing intron-spanning primers specific for each *HOX* gene (Thermo Scientific). GAPDH was used as reference gene for normalization in all experiments. Approximately 15 ng of cDNA was used for subsequent RT-PCR-analysis in a total volume of 25 µl containing 1x PCR-buffer, 0.2 µM of each primer, 1.5 mM MgCl₂, 0.2 mM each dNTP and 1 U *Taq* DNA Polymerase (Invitrogen) at the following conditions: (1) 2 min at 95 °C for initial Denaturation and *Taq* Polymerase activation, (2) 30 sec at 95 °C, 30 sec at 56 °C, (3) 30 sec at 72 °C for 35 cycles, 5 min at 72 °C for final extension of PCR products. PCR was performed on a Mastercycler ep gradient S (Eppendorf). Subsequently, aliquots of the RT-PCR products and related controls were analyzed on 2% agarose gel electrophoresis.

Real time PCR was carried out with SYBR® Green PCR Mastermix (Applied Biosystems) using 50 ng template cDNA. All reactions were run in duplicates/triplicates, respectively, on an ABI 7700 Detection System (Applied Biosystems). The sequences for the primers were carefully examined and checked for their specificity (a list of primers used is shown in the supplementary data (Table S1)). Evaluation of *Taq* Man Gene Expression Assays (Applied Biosystems) was performed with the SDS 2.3 software. Relative changes in gene expression were calculated following the $\Delta\Delta C_t$ -method with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as internal standard and normalized to human embryonic kidney (HEK) cells. Relative gene expression was illustrated as mean values.

Analysis of microarrays

Cell lines used for affymetrix chips were cultured for 4 days. On day 4 RNA was extracted according to the Rneasy Kit protocol (Qiagen). Approx. 5 µg of total RNA of each preparation were converted into labeled cRNA according to the manufacturer’s Expression manual Version 2 (Affymetrix, Santa Clara CA, USA). Aliquots of the labelled and subsequently fragmented cRNA were hybridized to GeneChip® HG-U133_Plus_2 microarrays (Affymetrix, Santa Clara CA, USA). Following several washing steps the hybridized microarrays were scanned on a GC Scanner 3000 with G7 update. Digitized signal intensities were determined and raw data quality was independently evaluated. The data were further analyzed using Genespring 10.1 Software (Agilent).

Probesets were filtered for fold-changes of ≥ 2.0 between groups (USSC vs. CB-MSC) and significant regulations were identified by an unpaired T-test with FDR-correction for multiple testing (Benjamini-Hochberg, $\alpha = 10\%$).

Annotation

Lists containing the differentially expressed probeset IDs were subjected to the “*Functional Annotation Chart Tool*” provided by DAVID Bioinformatics Resources (<http://www.david.abcc.ncifcrf.gov/>) (Dennis et al., 2003; Sherman et al., 2007; Huang da et al., 2009).

DNA methylation analysis by NimbleGen

1 μg genomic DNA from each cell line, the USSC, CB-MSC and the ES cell line were sonicated to 300-1000 fragment size by the Vibra Cell 75022 Ultrasonic Processor. These DNA samples then underwent immunoprecipitation of methylated DNA employing the Diagenode’s MeDIP kit in accordance to manufacturer’s instructions. Amplification of input and output samples occurred applying the Genome Plex® Complete WGA Kit (Sigma Aldrich) as described in the user’s guide. Hybridization of 1 μg of each amplified DNA sample was performed on NimbleGen 385 K RefSeq Promoter Arrays HG18 containing all known RefSeq genes (Roche). The promoter regions on these arrays are covered by 50-mer probes with approximately 100 bp spacing. The hybridization procedure was applied as suggested by the manufacturer. The hybridized arrays were scanned on an Axon 4000B microarray scanner (Molecular Devices, Sunnyvale CA), and the images were analyzed with Axon GenePix software version 4.1. Image and data analyses were done by NimbleScan version 2.5 and SignalMap version 1.9 software.

Acknowledgments

First of all, we would like to thank Maria Kluth, Stefanie Geyh and Foued Ghanjati for their excellent technical support. We are grateful to Oliver Brüstle for generously providing H9 RNA, and to Paolo Bianco and Hans Werner Müller for helpful discussions. The work was supported by the Deutsche Forschungsgemeinschaft (DFG) project Ko2119/6-1 and the José Carreras Leukemia Foundation grant DJCLS-R07/05v.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.scr.2010.03.004.

References

- Ackema, K.B., Charite, J., 2008. Mesenchymal stem cells from different organs are characterized by distinct topographic Hox codes. *Stem Cells Dev.* 17, 979–991.
- Akam, M., 1998. Hox genes: from master genes to micromanagers. *Curr. Biol.* 8, R676–R678.
- Benson, G.V., Nguyen, T.H., Maas, R.L., 1995. The expression pattern of the murine Hoxa-10 gene and the sequence recognition of its homeodomain reveal specific properties of Abdominal B-like genes. *Mol. Cell. Biol.* 15, 1591–1601.
- Brzoska, E., Grabowska, I., Hoser, G., Streminska, W., Wasilewska, D., Machaj, E.K., Pojda, Z., Moraczewski, J., Kawiak, J., 2006. Participation of stem cells from human cord blood in skeletal muscle regeneration of SCID mice. *Exp. Hematol.* 34, 1262–1270.
- Buchheiser, A., Liedtke, S., Houben, A.P., Waclawczyk, S., Stephan, M., Radke, T.F., Wernet, P., Kogler, G., 2008. Cord Blood as a Very Valuable Source of Neonatal Cells but Embryonic-Like Nature Reevaluated. *Blood* 112, 997 abstract.
- Buchheiser, A., Liedtke, S., Looijenga, L.H., Kogler, G., 2009. Cord blood for tissue regeneration. *J. Cell. Biochem.* 108, 762–768.
- Chan, S.L., Choi, M., Wnendt, S., Kraus, M., Teng, E., Leong, H.F., Merchav, S., 2007. Enhanced in vivo homing of uncultured and selectively amplified cord blood CD34+ cells by cotransplantation with cord blood-derived unrestricted somatic stem cells. *Stem Cells* 25, 529–536.
- Chang, H.Y., Chi, J.T., Dudoit, S., Bondre, C., van de Rijn, M., Botstein, D., Brown, P.O., 2002. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc. Natl. Acad. Sci. U. S. A.* 99, 12877–12882.
- Chang, Y.J., Shih, D.T., Tseng, C.P., Hsieh, T.B., Lee, D.C., Hwang, S.M., 2006. Disparate mesenchyme-lineage tendencies in mesenchymal stem cells from human bone marrow and umbilical cord blood. *Stem Cells* 24, 679–685.
- Cruzet, S., Couly, G., Vincent, C., Le Douarin, N.M., 2002. Negative effect of Hox gene expression on the development of the neural crest-derived facial skeleton. *Development* 129, 4301–4313.
- Dennis Jr., G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4, P3.
- Goodwin, H.S., Bicknese, A.R., Chien, S.N., Bogucki, B.D., Quinn, C.O., Wall, D.A., 2001. Multilineage differentiation activity by cells isolated from umbilical cord blood: expression of bone, fat, and neural markers. *Biol. Blood Marrow Transplant.* 7, 581–588.
- Greschat, S., Schira, J., Kury, P., Rosenbaum, C., de Souza Silva, M.A., Kogler, G., Wernet, P., Muller, H.W., 2008. Unrestricted somatic stem cells from human umbilical cord blood can be differentiated into neurons with a dopaminergic phenotype. *Stem Cells Dev.* 17, 221–232.
- Hu, C.H., Wu, G.F., Wang, X.Q., Yang, Y.H., Du, Z.M., He, X.H., Xiang, P., 2006. Transplanted human umbilical cord blood mononuclear cells improve left ventricular function through angiogenesis in myocardial infarction. *Chin. Med. J. (Engl.)* 119, 1499–1506.
- Huang da, W., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Hwang, J.H., Seok, O.S., Song, H.R., Jo, J.Y., Lee, J.K., 2009. HOXC10 as a Potential Marker for Discriminating between Amnion- and Decidua-Derived Mesenchymal Stem Cells. *Cloning Stem Cells* 11, 269–279.
- Jansen, B.J., Gilissen, C., Roelofs, H., Schaap-Oziemlak, A., Veltman, J., Raymakers, R.A., Jansen, J., Kogler, G., Figdor, C.G., Torensma, R., Adema, G.J., 2009. Functional Differences between Mesenchymal Stem Cell Populations Are Reflected by Their Transcriptome. *Stem Cells Dev.*
- Kern, S., Eichler, H., Stoeve, J., Kluter, H., Bieback, K., 2006. Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood, or adipose tissue. *Stem Cells* 24, 1294–1301.
- Kessel, M., Gruss, P., 1991. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67, 89–104.
- Kim, B.O., Tian, H., Prasongsukarn, K., Wu, J., Angoulvant, D., Wnendt, S., Muhs, A., Spitkovsky, D., Li, R.K., 2005. Cell transplantation improves ventricular function after a myocardial infarction: a preclinical study of human unrestricted somatic stem cells in a porcine model. *Circulation* 112, 196–1104.
- Kmita, M., Duboule, D., 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science* 301, 331–333.

- Kogler, G., Sensken, S., Airey, J.A., Trapp, T., Muschen, M., Feldhahn, N., Liedtke, S., Sorg, R.V., Fischer, J., Rosenbaum, C., Greschat, S., Knipper, A., Bender, J., Degistirici, O., Gao, J., Caplan, A.I., Colletti, E.J., Almeida-Porada, G., Muller, H.W., Zanjani, E., Wernet, P., 2004. A new human somatic stem cell from placental cord blood with intrinsic pluripotent differentiation potential. *J. Exp. Med.* 200, 123–135.
- Kogler, G., Radke, T.F., Lefort, A., Sensken, S., Fischer, J., Sorg, R.V., Wernet, P., 2005. Cytokine production and hematopoiesis supporting activity of cord blood-derived unrestricted somatic stem cells. *Exp. Hematol.* 33, 573–583.
- Kogler, G., Sensken, S., Wernet, P., 2006. Comparative generation and characterization of pluripotent unrestricted somatic stem cells with mesenchymal stem cells from human cord blood. *Exp. Hematol.* 34, 1589–1595.
- Kogler, G., Critser, P., Trapp, T., Yoder, M., 2009. Future of cord blood for non-oncology uses. *Bone Marrow Transplant.* 44, 683–697.
- Krumlauf, R., 1994. Hox genes in vertebrate development. *Cell* 78, 191–201.
- Lawrence, H.J., Sauvageau, G., Ahmadi, N., Lopez, A.R., LeBeau, M.M., Link, M., Humphries, K., Largman, C., 1995. Stage- and lineage-specific expression of the HOXA10 homeobox gene in normal and leukemic hematopoietic cells. *Exp. Hematol.* 23, 1160–1166.
- Lee, R.H., Kim, B., Choi, I., Kim, H., Choi, H.S., Suh, K., Bae, Y.C., Jung, J.S., 2004. Characterization and expression analysis of mesenchymal stem cells from human bone marrow and adipose tissue. *Cell. Physiol. Biochem.* 14, 311–324.
- Leor, J., Guetta, E., Feinberg, M.S., Galski, H., Bar, I., Holbova, R., Miller, L., Zarin, P., Castel, D., Barbash, I.M., Nagler, A., 2006. Human umbilical cord blood-derived CD133+ cells enhance function and repair of the infarcted myocardium. *Stem Cells* 24, 772–780.
- Leucht, P., Kim, J.B., Amasha, R., James, A.W., Girod, S., Helms, J.A., 2008. Embryonic origin and Hox status determine progenitor cell fate during adult bone regeneration. *Development* 135, 2845–2854.
- Liedtke, S., Enczmann, J., Waclawczyk, S., Wernet, P., Kogler, G., 2007. Oct4 and its pseudogenes confuse stem cell research. *Cell Stem Cell* 1, 364–366.
- Liedtke, S., Stephan, M., Kogler, G., 2008. Oct4 expression revisited: potential pitfalls for data misinterpretation in stem cell research. *Biol. Chem.* 389, 845–850.
- McGonigle, G.J., Lappin, T.R., Thompson, A., 2008. Grappling with the HOX network in hematopoiesis and leukemia. *Front Biosci.* 13, 4297–4308.
- Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., Schubeler, D., 2008. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* 30, 755–766.
- Morgan, R., 2006. Hox genes: a continuation of embryonic patterning? *Trends Genet.* 22, 67–69.
- Newcomb, J.D., Ajmo Jr., C.T., Sanberg, C.D., Sanberg, P.R., Pennypacker, K.R., Willing, A.E., 2006. Timing of cord blood treatment after experimental stroke determines therapeutic efficacy. *Cell Transplant.* 15, 213–223.
- Prockop, D.J., 1997. Marrow stromal cells as stem cells for nonhematopoietic tissues. *Science* 276, 71–74.
- Prockop, D.J., Gregory, C.A., Spees, J.L., 2003. One strategy for cell and gene therapy: harnessing the power of adult stem cells to repair tissues. *Proc. Natl. Acad. Sci. U. S. A.* 100 (Suppl 1), 11917–11923.
- Rauch, T., Wang, Z., Zhang, X., Zhong, X., Wu, X., Lau, S.K., Kernstine, K.H., Riggs, A.D., Pfeifer, G.P., 2007. Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay. *Proc. Natl. Acad. Sci. U. S. A.* 104, 5527–5532.
- Rinn, J.L., Bondre, C., Gladstone, H.B., Brown, P.O., Chang, H.Y., 2006. Anatomic demarcation by positional variation in fibroblast gene expression programs. *PLoS Genet.* 2, e119.
- Rinn, J.L., Wang, J.K., Allen, N., Brugmann, S.A., Mikels, A.J., Liu, H., Ridky, T.W., Stadler, H.S., Nusse, R., Helms, J.A., Chang, H.Y., 2008. A dermal HOX transcriptional program regulates site-specific epidermal fate. *Genes Dev.* 22, 303–307.
- Riordan, N.H., Chan, K., Marleau, A.M., Ichim, T.E., 2007. Cord blood in regenerative medicine: do we need immune suppression? *J. Transl. Med.* 5, 8.
- Sensken, S., Waclawczyk, S., Knaupp, A.S., Trapp, T., Enczmann, J., Wernet, P., Kogler, G., 2007. In vitro differentiation of human cord blood-derived unrestricted somatic stem cells towards an endodermal pathway. *Cytotherapy* 9, 362–378.
- Sherman, B.T., Huang da, W., Tan, Q., Guo, Y., Bour, S., Liu, D., Stephens, R., Baseler, M.W., Lane, H.C., Lempicki, R.A., 2007. DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. *BMC Bioinform.* 8, 426.
- Shimeld, S.M., Gaunt, S.J., Coletta, P.L., Geada, A.M., Sharpe, P.T., 1993. Spatial localisation of transcripts of the Hox-C6 gene. *J. Anat.* 183 (Pt 3), 515–523.
- Soshnikova, N., Duboule, D., 2008. Epigenetic regulation of Hox gene activation: the waltz of methyls. *Bioessays* 30, 199–202.
- Takahashi, Y., Hamada, J., Murakawa, K., Takada, M., Tada, M., Nogami, I., Hayashi, N., Nakamori, S., Monden, M., Miyamoto, M., Katoh, H., Moriuchi, T., 2004. Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Exp. Cell Res.* 293, 144–153.
- Trapp, T., Kogler, G., El-Khattouti, A., Sorg, R.V., Besselmann, M., Focking, M., Buhrle, C.P., Trompeter, I., Fischer, J.C., Wernet, P., 2008. Hepatocyte growth factor/c-MET axis-mediated tropism of cord blood-derived unrestricted somatic stem cells for neuronal injury. *J. Biol. Chem.* 283, 32244–32253.
- van den Berk, L.C., Jansen, B.J., Siebers-Vermeulen, K.G., Netea, M.G., Latuhihin, T., Bergevoet, S., Raymakers, R.A., Kogler, G., Figdor, C.C., Adema, G.J., Torensma, R., 2009. Toll-like receptor triggering in cord blood mesenchymal stem cells. *J. Cell. Mol. Med.*
- Wang, K.C., Helms, J.A., Chang, H.Y., 2009. Regeneration, repair and remembering identity: the three Rs of Hox gene expression. *Trends Cell Biol.* 19, 268–275.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., Schubeler, D., 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* 39, 457–466.
- Winter, M., Wang, X.N., Daubener, W., Eyking, A., Rae, M., Dickinson, A.M., Wernet, P., Kogler, G., Sorg, R.V., 2008. Suppression of Cellular Immunity by Cord Blood-Derived Unrestricted Somatic Stem Cells is Cytokine Dependent. *J. Cell. Mol. Med.*
- Yamamoto, M., Takai, D., Yamamoto, F., 2003. Comprehensive expression profiling of highly homologous 39 hox genes in 26 different human adult tissues by the modified systematic multiplex RT-pCR method reveals tissue-specific expression pattern that suggests an important role of chromosomal structure in the regulation of hox gene expression in adult tissues. *Gene Expr.* 11, 199–210.
- Yasui, D.H., Peddada, S., Bieda, M.C., Vallerio, R.O., Hogart, A., Nagarajan, R.P., Thatcher, K.N., Farnham, P.J., Lasalle, J.M., 2007. Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19416–19421.
- Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., Henikoff, S., 2007. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* 39, 61–69.
- Zuk, P.A., Zhu, M., Mizuno, H., Huang, J., Futrell, J.W., Katz, A.J., Benhaim, P., Lorenz, H.P., Hedrick, M.H., 2001. Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng.* 7, 211–228.