Beyond HOX: the role of ParaHox genes in normal and malignant hematopoiesis

Vijay P. S. Rawat¹, R Keith Humphries², Christian Buske¹

¹The Institute of Experimental Cancer Research, Comprehensive Cancer Center and University Hospital Ulm, 89081 Ulm, Germany; ²The Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, Canada

Short title: ParaHox genes in normal and malignant hematopoiesis

Corresponding author:

Christian Buske, M.D.

Institute of Experimental Cancer Research

University of Ulm

Albert-Einstein-Allee 11

89081 Ulm

Phone: +49-731 50065800

FAX: +49-731 50065822

E-Mail: christian.buske@uni-ulm.de
Abstract

During the past decade it was recognized that homeobox gene families such as the clustered Hox genes play pivotal roles both in normal and malignant hematopoiesis. More recently, similar roles have also become apparent for members of the ParaHox gene cluster, evolutionarily closely related to the Hox gene cluster. This is in particular demonstrated for the caudal – type homeobox genes (Cdx) genes, known to act as upstream regulators of Hox genes. The CDX gene family member CDX2 belongs to the most frequent aberrantly expressed proto-oncogenes in human acute leukemias and is highly leukemogenic in experimental models. Correlative studies indicate that CDX2 functions as master regulator of perturbed HOX gene expression in human acute myeloid leukemia, locating this ParaHox gene at a central position for initiating and maintaining HOX gene dysregulation as a driving leukemogenic force. There are still few data about potential upstream regulators initiating aberrant CDX2 expression in human leukemias or about critical downstream targets of CDX2 in leukemic cells. Characterizing this network will hopefully open the way to therapeutic approaches targeting deregulated ParaHox genes in human leukemia.
Introduction

The last decade has seen large advances in our understanding of the regulatory network in prenatal and adult hematopoiesis. As one pivotal class of regulatory factors the highly conserved family of homeobox genes have been identified as playing major roles not only in normal hematopoiesis but also in leukemogenesis. The homeobox was first identified in the 1980s as a sequence motif shared among Drosophila homeotic genes (the HOM-C complex) which play key roles in embryonic differentiation along the anterior–posterior axis. The existence of homeotic genes was already proposed in the early 1900s based on observations of Drosophila mutants with changes in the body structure. The homeobox is now known to be present in many genes in virtually all eukaryotic species. It is estimated that the human genome contains at least 200 homeobox genes. Unlike the HOM-C/HOX genes, which are organized in gene clusters, most homeobox genes are dispersed throughout the genome. Homeobox-containing genes constitute a gene family characterized by a highly conserved 183-nucleotide sequence encoding a 61-amino acid domain, the homeodomain (HD). These homeodomains are structurally related to the helix-turn-helix motif of prokaryotic DNA-binding proteins and have sequence-specific DNA binding activity. Homeobox genes are divided into two classes. Class I includes clustered genes (HOX) that comprises 39 members, while the Class II divergent homeobox genes are dispersed through the genome and include smaller families such as the MSX, ParaHox (CDX), PAX, DLX and Engrailed (EN) groups. Although individual members of homeobox families often share little sequence similarity other than the homeobox, some families have additional conserved sequence motifs that contribute to their distinct functional properties.

A key role of homeobox genes for normal and malignant stem – and progenitor cell
behavior was first shown for Hox genes. Initial hints that these genes are essential for regulating hematopoietic stem cells (HSCs) and progenitor cells came from data showing that Hox genes of the A and B cluster are highly expressed in normal murine and human hematopoietic stem (HSCs) and progenitor cells, but silenced in their more differentiated progeny. Studies in experimental models have convincingly proven the functional relevance of Hox genes for normal hematopoietic behavior in mouse and man: e.g. Hoxa9-/- bone marrow cells manifest a severe defect at the level of HSCs and committed progenitors; on the other hand aberrant expression of Hox/HOX genes such as Hoxa9 or Hoxa10 can induce acute myeloid leukemia (AML) in mice, whereas constitutive expression of Hoxb4 increases the self-renewal capacity of murine HSCs without initiating leukemic transformation. Similar data were obtained for the human system using retroviral gene transfer techniques and the NOD/SCID xenograft model. Early data revealed aberrant expression of HOX genes in patients with acute leukemia and this was confirmed later by microarray analyses in large cohorts of patients with AML thereby associating aberrant expression of HOX genes such as HOXA9 with clinical outcome. The rich body of data on the role of Hox genes in normal and malignant hematopoiesis has been reviewed in detail before. In the last years, however, it was shown that other homeobox genes, not belonging to the clustered Hox genes, have a crucial role in regulating early hematopoietic cells and facilitating leukemic transformation. One example is the so called ‘Three-amino-acid loop extension’ (TALE) superfamily, forming an ‘atypical class’ of homeobox genes which show very low sequence identity with other classes of homeobox genes. Most prominent members of this family are Meis1 and Pbx1. Both can directly interact with Hox genes and with each other, thereby forming trimeric complexes, which alter Hox –
DNA binding properties and in the case of Meis1 accelerate Hox induced leukemogenesis. This review will focus on another homeobox gene family, the ParaHox cluster, which was shown to be involved in normal as well as malignant hematopoiesis.

The ParaHox cluster genes: emerging key players in normal and malignant hematopoiesis

**Genomic organization and molecular characteristics**

The ParaHox gene cluster was discovered in 1998 by Brooke et al. who reported that Gsx (genomic screened homeobox), Xlox (Xenopus laevis homeobox) and Cdx (Caudal-type homeobox) genes form their own cluster in the cephalochordate amphioxus. This cluster was termed the ParaHox cluster as it showed homology to the different paralogue groups of the Hox cluster. In mice and man homologous genes cluster on chromosome 5 and 13q12, respectively: these comprise the genes Gsh1/GSH1, Xlox/XLOX (also known as Pdx1/PDX1) and CDX2. In contrast to chordates, mammals have additional ParaHox genes, namely Gsh2/GSH2, Cdx1/CDX1 and Cdx4/CDX4, which are spread over three different chromosomes in the human genome.

The ParaHox and Hox clusters are thought to originate from an ancient hypothetical ProtoHox cluster by duplication before the split of protostomes and deuterostomes, characterizing the ParaHox cluster as one of the evolutionary early homeobox gene clusters. This also explains the homologies of the ParaHox genes to paralogue
groups of the Hox cluster. The paralogue groups of the Hox cluster subdivide the Hox genes into four groups: the Anterior Group (Paralogue Group (PGs) 1–2), Group 3, Central Group (PGs 4–8) and Posterior Group (PGs 9–13) \(^6,30\). The Gsx gene is more closely related to Hox ‘Anterior Group’ genes than to other ParaHox genes; Xlox is more similar to Hox ‘Group 3’, and Cdx is more similar to the Hox ‘Posterior Group’ genes.

Structure – functional analyses have characterized the molecular characteristics of ParaHox proteins in detail (table 1)(Fig.1b). As for homeobox genes, all the ParaHox genes are characterized by a DNA binding homeodomain and in part by a N-terminal transactivation domain. However, homology between the different family members of the ParaHox genes are relatively low and proteins display individual features. Thus, the ParaHox gene PDX1 contains an antennapedia-type hexapeptide motif (PFPWMK), which preferentially binds the DNA motif 5’-[CT]TAAT[TG]-3’. PDX1 is phosphorylated by HIPK2 on Ser-268 upon glucose accumulation, which mediates subnuclear localization shifting from nuclear periphery to nucleoplasm \(^31\). GSH1 contains an additional Snail/Gfi (SNAG) domain essential for transcriptional repression. Based on bioinformatics analysis it is postulated that GSH genes bind to DNA as monomers or as homo- and/or heterodimers, in a sequence-specific manner. CDX1, in contrast to other CDX members, does not contain the highly conserved pentapeptide domain for binding of the TALE cofactor Pbx1.

**Lessons from knock-out mice: initial clues to function of ParaHox genes**

Knock-out mice have been an important tool for our understanding of ParaHox gene function and their essential role in body development (table 1). *Gsh1* homozygous mutant mice are phenotypically normal at birth, but fail to grow properly after 7-10 days.
Homozygous mutants exhibit extreme dwarfism, sexual infantilism and significant perinatal mortality with total white blood cell count (WBC) 5-fold decreased in mutant mice. Both lymphoid and myeloid cells were similarly reduced while macrophages were not affected. \textit{Gsh2} knockout mice display severe structural alterations in the brain accompanied by apnea and reduced blood oxygen levels, leading to death within the first day after birth. \textit{Pdx-1} was shown to be crucial for the development of the pancreas and anterior duodenum. \textit{Cdx1-null} mutant mice are viable and fertile, but display anterior homeotic transformations of vertebrae accompanied by posterior shifts of Hox gene expression in the somatic mesoderm. \textit{Cdx2-null} mutant embryos die between 3.5 and 5.5 days post coitum (dpc), and heterozygotes have tail abnormalities and exhibit anterior homeotic transformations involving the cervical and upper thoracic vertebrae, ribs, and midgut endoderm, underlining the crucial role of Cdx genes for anterior-posterior regional identity. 90% of Cdx2 heterozygote mutant mice develop multiple intestinal adenomatous polyps, particularly in the proximal colon, suggesting that Cdx2 mutations might be the primary event in the genesis of some intestinal tumours. In contrast to \textit{Cdx2}, \textit{Cdx4-null} murine embryos are born healthy and do not show any major morphological or histological abnormalities.

\textit{Cdx2} and \textit{Cdx4} are key regulators of embryonic hematopoiesis

Over the past several years it has become evident, that ParaHox Cdx family members are critical for ordered proliferation and differentiation of embryonic hematopoietic cells. This was particularly demonstrated for the Cdx genes Cdx4 and Cdx2. Evidence for a key role of \textit{Cdx} genes in embryonic hematopoiesis initially came from studies in zebrafish: homozygous mutations in \textit{cdx4} caused a severe defect in embryonic
hematopoiesis with a profound reduction of hemoglobin positive erythroid cells resulting in severe anemia \(^{41}\). These blood defects were accompanied by deregulated or reduced expression of \(hoxb4, hoxb5a, hoxb6b, hoxb7a, hoxb8a, hoxb8b\) and \(hoxa9a\). Of note, the blood deficiency in zebrafish embryos could be rescued by overexpression of \(hoxb7a\) or \(hoxa9a\) but not by \(scl\), indicating that the hematopoietic defect is at least partly caused by perturbation of distinct Hox genes \(^{41}\). The same group demonstrated that morpholino-mediated knockdown of \(cdx1a\) in a \(cdx4\) mutant background caused a severe perturbation of Hox gene expression and a complete failure to specify blood in zebrafish. Furthermore, these Cdx-deficient embryos displayed a significant reduction in expression of key hematopoietic genes such as \(scl\), \(runx1\), and \(gata1\) \(^{42}\). This hematopoietic defect could be efficiently rescued by hoxa9 overexpression, but not by hoxb7 \(^{42}\).

In mouse, ectopic expression of \(Cdx4\) enhances hematopoietic mesoderm formation and further promotes blood progenitor specification and hematopoietic engraftment of murine embryonic stem cells (ESCs) in adult mice \(^{43}\). In addition, \(Cdx4\) overexpression rescues defective blood progenitor formation in mouse ESCs deficient in \(Mll\), a Hox regulator involved in definitive hematopoiesis \(^{44}\). Surprisingly, \(Cdx4\) knock-out mice do not show any major hematopoietic defects \(^{45}\). Indeed, in both germline and conditional \(Cdx4\) knockout models, there is no apparent defect in HSC and progenitor cells as well as mature blood cells. This indicates that Cdx4 as a single factor is not essential for normal adult hematopoietic stem cell functions in mice and that Cdx4 function can be compensated at least partly by other Cdx genes \(^{45}\). One candidate for this is \(Cdx2\): Cdx2-deficient ESCs show impaired production of multipotential blood progenitor colonies \emph{in vitro} consistent with an essential role of \(Cdx2\) in early embryonic hematopoiesis \(^{46}\). By
injecting $Cdx2^{-/-}$ or $Cdx2^{+/+}$ ESCs into normal lacZ$^+$ blastocysts Wang et al. could demonstrate an intrinsic requirement for $Cdx2$ during primitive embryonic hematopoiesis in vivo $^{46}$. Based on genome-wide expression analysis they reported that $Cdx2$ deficiency dramatically perturbs expression profiles of $Hox$ genes and genes involved in ‘signal transduction’, ‘cell growth and proliferation’ and ‘hematopoiesis’. Expression of several components of the canonical signal transduction pathways such as the Jak/Stat and sonic hedgehog pathways were altered in $Cdx2^{-/-}$ EBs. Expression of several hematopoiesis-specific genes including $Scl$, $Gata1$, $Runx1$, and $\beta$-$H1$ globin were markedly decreased in $Cdx2$-deficient EBs $^{46}$, supporting the notion that $Cdx2$ is a key player during hematopoietic development.

In contrast to $Cdx4$ and $Cdx2$, loss of $Cdx1$ gene function does not affect hematopoietic development as illustrated by homozygous $Cdx1$ mutant mice $^{35}$, but also by studies in zebrafish, where $cdx1a$ had no effect on rostral blood island (RBI) hematopoiesis and the intermediate cell mass (ICM) development $^{42}$. The distinct effects of the different members of the $Cdx$ gene family was elegantly shown in a study, which ectopically expressed $Cdx1$, $Cdx2$ and $Cdx4$ in murine ESCs: whereas $Cdx1$ and $Cdx4$ both promoted the hematopoietic potential of the EB - derived CD41$^+$c-kit$^+$ population, $Cdx2$ grossly reduced EB - derived hematopoietic potential. Expression of all the different $Cdx$ genes increased Hox gene expression with the notable difference that $Cdx2$ induced changes in both the anterior and posterior HoxA and HoxB genes, whereas $Cdx1$ and $Cdx4$ only altered posterior Hox gene expression $^{47}$.

Taken together, these data demonstrate that $Cdx$ genes play essential and both distinct and overlapping roles in embryonic hematopoiesis.
ParaHox genes in malignant hematopoiesis

ParaHox genes have emerged as pivotal players in leukemogenesis with a rich body of evidence linking aberrant expression of the ParaHox gene Cdx2 and other ParaHox genes to acute leukemia (table1)(Fig.2).

GSH2

The Gsh2 gene was previously characterized as a brain-specific ParaHox gene important in forebrain and hindbrain formation of the mouse. Peter Marynen's group has shown that GSH2 is not expressed in adult human normal hematopoietic cells. However the group has reported ectopic expression of GSH2 in four patients with AML carrying the translocation t(4;12)(q11-q12;p13), which itself generated no detectable fusion gene. GSH2 is located at 4q11-q12 and thus located closely to the respective breakpoint, which implies that the translocation induces ectopic expression of GSH2. Indeed the authors could demonstrate that aberrant expression of GSH2 caused focus formation in the NIH3T3 assay as a surrogate marker for the oncogenic potential of this ParaHox gene 48.

Cdx1

In normal adult hematopoietic tissue, expression of this homeobox gene was not reported so far and there are no data suggesting involvement of Cdx1 in malignant hematopoiesis. However, in the adult Cdx1 is known to be involved in the intestinal differentiation and the development of intestinal metaplasias 49-51.
**CDX2**

Aberrant expression of CDX2 is a frequent event in patients with AML

Several studies have reported that CDX2 is normally not expressed in adult murine and human normal hematopoiesis (Fig. 2). Ectopic expression of CDX2 in AML was reported by Chase et al. in 1999 in a patient with AML M1 carrying the chromosomal translocation t(12;13)(p13;q12), creating a novel fusion of ETV6 and CDX2. In addition, aberrant expression of CDX2 was found in a case of blast crisis of chronic myeloid leukemia (CML-BC) with no detectable chromosomal abnormality. The fusion partner ETV6 is an important regulator of hematopoietic stem cell survival and frequently affected by translocations. Later we and others have reported that CDX2 is aberrantly expressed in more than 80% of human AML patients, characterizing CDX2 as one of the most frequent aberrantly expressed proto-oncogenes in human leukemia (Fig. 2). Aberrant expression of CDX2 was found in all types of cytogenetic subgroups. Scholl et al. reported the highest expression of CDX2 in patients with t(9;11) followed by normal karyotype AML (CN - AML) and t(15;17) positive patients, similar to our results with highest expression of CDX2 in CN-AML followed by t(9;11) and t(15;17) positive cases. Aberrant expression of CDX2 was not only more common in patients with CN-AML compared to patients with abnormal karyotype (89% vs. 64%, respectively), but the expression level of CDX2 was significantly and more than 15-fold higher in patients with CN-AML as well. Furthermore, the level of CDX2 expression correlated with aberrant HOX gene expression in human AML. The analyses did not show any correlation of the NPM1 mutational status with the frequency and level of aberrant CDX2 expression. Aberrant CDX2 expression is also detectable in most of the human AML cell lines. The mechanisms responsible for disruption of the normal CDX2 silencing in human AML are
unknown and are not due to promoter mutation, gene amplification or abnormal promoter methylation. In contrast to CDX2, all CN-AML cases were negative for expression of the other members of the CDX gene family CDX1 and CDX4.

**CDX2** is aberrantly expressed in the majority of patients with acute lymphoblastic leukemia (ALL)

Ectopic expression of **CDX2** is not restricted to AML patients: we showed that **CDX2** is aberrantly expressed in the majority of adult ALL patients, with 81% overall positivity and a higher overall expression level than detectable in CN-AML patients. The expression of **CDX2** showed high variation between the different ALL subgroups (classified according to the criteria of the ‘European Group for the immunological characterization of leukemias’ (EGIL)) with highest median expression in pre-T-ALL followed by c-ALL and pro-B ALL. In contrast, patients with B-ALL/ Burkitt lymphoma and thymic T-ALL exhibited only low **CDX2** expression. Also the number of patients aberrantly expressing **CDX2** differed considerably between ALL subtypes, with 100% positivity in patients with pro-B ALL, c-ALL and Ph⁺ ALL versus only 40% positivity in B-ALL/ Burkitt lymphoma and around 70% in thymic and pre-T ALL. The expression level of **CDX2** in ALL patients significantly correlated with treatment outcome (high expression levels were associated with inferior overall survival) of the analyzed patients. Although this study was limited by a low number of patients, **CDX2** remained an independent risk factor even after adjusting for the risk factors age and presence of molecular markers by bivariate Cox regression. These data were confirmed for pediatric ALL where the vast majority of patients also showed aberrant expression of **CDX2**. Also in this study,
aberrant expression of CDX2 significantly correlated to an inferior disease outcome. Another recent independent study reported CDX2 expression in 61% of de novo adult ALL patients and decrease of CDX2 expression in patients with complete remission and increase in relapse. Similar to AML the mechanism of aberrant CDX2 expression in ALL is not understood. We have not identified any difference in the methylation level of the CDX2 promoter in CDX2-positive and -negative ALL patients. In contrast to AML, where aberrant CDX2 expression correlated with HOX gene deregulation, there was only a low correlation between the expression of HOX genes with the expression of CDX2, in line with reports that HOX gene deregulation is less common in ALL than in AML. Taken together, these data link aberrant expression of CDX2 to myeloid as well as lymphoid leukemia. So far it is unknown, whether the stage of differentiation of the cell, which is initially hit by aberrant CDX2 expression, determines the phenotype of the leukemia. Experiments testing whether aberrant expression of CDX2 in lymphoid progenitor cells generates ALL in mice are lacking so far.

Cdx2 is highly leukemogenic in experimental models

In a murine BM transplantation model it was shown that ectopic expression of the Parahox gene CDX2, but not expression of the fusion gene ETV6/CDX2, which was described in a patient with translocation t(12;13)(p13;q12) positive AML, was highly leukemogenic. This was the first direct evidence that the ParaHox gene Cdx2 is highly leukemogenic and induces aggressive AML when aberrantly expressed in hematopoietic progenitor cells. Retrovirally enforced expression of Cdx2 in murine hematopoietic progenitors resulted in increased expression of leukemogenic Hox genes such as
Hoxa5, Hoxa7, Hoxa9, Hoxa10 as well as Hoxb3, Hoxb6 and Hoxb8\textsuperscript{55}. These data are in agreement with previous findings that Cdx2 can deregulate Hox gene expression in embryonic as well as adult non-hematopoietic tissues\textsuperscript{37,64}. The ability to upregulate Hox gene expression and to induce AML was clearly dependent on the N-terminal transactivation domain of Cdx2, known to harbor serine phosphorylation sites, which affect the transcriptional activity of the gene\textsuperscript{55,65}. Interestingly, the ETV6-CDX2 fusion gene also lacks the N-terminal transactivation domain of CDX2, which would explain the obvious discrepancy of the oncogenic potential between Cdx2 and the ETV6-CDX2 fusion.

Aberrant Cdx2 expression also led to an altered expression of a number of genes involved in lymphopoiesis such as Lef1, Tcf3 or Id3. These genes are frequently deregulated in lymphoid leukemia\textsuperscript{58,66-67} and are able in the case of Lef1 to induce ALL in transplanted mice\textsuperscript{55}.

The vast majority of data on the leukemogenicity of Cdx2 comes from murine models. However, at least at the level of human leukemia cell lines it could be demonstrated, that leukemic growth depends on aberrant CDX2 expression: thus, shRNA mediated depletion of CDX2 in the human pre-B cell line Nalm-6 and several AML cell lines significantly impaired leukemic cell proliferation and clonogenicity\textsuperscript{56,58}.

CDX4 is less leukemogenic than Cdx2

It has been shown that overexpression of Cdx4 confers enhanced replating and proliferative potential to transduced murine hematopoietic cells. Mice transplanted with Cdx4 overexpressing bone marrow developed a myeloproliferative disorder, which
progressed to AML over time \(^\text{68}\). In this model aberrant expression of Cdx4 had to collaborate with the homeobox gene Meis1 to induce AML in all experimental animals \(^\text{68}\), thus clearly differing from Cdx2, which caused rapid disease in all transplanted mice\(^\text{63}\). In line with this the long latency of leukemia development in primary transplanted animals suggested that aberrant expression of Cdx4 alone is not sufficient to induce leukemia but rather induces a pre-leukemic stage. Loss of Cdx4 significantly prolonged the latency of disease onset in a mouse model of MLL-AF9-induced AML and altered the phenotype of the resultant disease with increased expression of B- and T-lymphoid markers \(^\text{45}\). The mechanism of Cdx4 induced leukemogenesis is not completely understood. However, recently it has been shown that CDX4 activates transcription of the Hox gene HOXA10 in myeloid cells and binds to a cis element of its promoter. Vice versa HOXA10 is able to activate CDX4 expression, pointing to a positive feedback loop between CDX4 and a HOX gene, which itself is linked to myeloid leukemogenesis \(^\text{17,69-70}\). It was also demonstrated that Cdx4 collaborates with Menin to activate Hoxa9 expression by binding with Menin at the same regulatory region at the Hoxa9 locus \(^\text{71}\). In summary, these data categorize CDX4 as an important collaborative partner in Hox associated leukemogenesis.

In AML patients \(CDX4\) is expressed aberrantly in approximately 25% of all cases. \(CDX4\) expression, however, did not correlate to a certain leukemia subtype and its expression did not correlate to HOX gene deregulation \(^\text{68}\). This finding suggests that the ability of CDX4 to deregulate HOX gene expression is of minor relevance for its role in human leukemia.
Mechanisms of ParaHox induced leukemogenesis – speculations and future directions

Gene expression profiling and functional data clearly support a critical involvement of CDX2 in human leukemogenesis. However, molecular mechanisms of Cdx2 induced leukemogenesis are poorly understood. One obvious possibility is that CDX2 exerts its leukemogenic potential by deregulating expression of leukemogenic genes of the HOXA and HOXB cluster, known to be dysregulated in 65% of all AML cases, in particular in normal karyotype (CN-AML) and MLL rearranged AML \(^{55,56}\) (Fig. 2). This would classify the CDX-HOX axis as one of the major driving forces in human AML. However, it could be clearly demonstrated that aberrant Cdx2 expression, besides Hox deregulation, also led to an altered expression of a number of genes involved in embryonic development, lymphopoiesis, Wnt pathway and genes involved in stem cell function. These data support the hypothesis that Cdx2 might deregulate multiple pathways in leukemia, which would also explain the differences between Cdx2 and Cdx4 with regard to their leukemogenic potential.

So far strategies to target Cdx2-induced transformation remain elusive. One possibility would be to identify druggable signaling pathways on which Cdx2 activity are dependent. Interestingly, it was shown in a colon cancer cell line that the transcriptional activity of Cdx2 depends on phosphorylation of the N-terminal S60 site by ERK kinases. We have shown that the N-terminal activation domain of Cdx2, at which the S60 site is localized, is essential for Cdx2 induced leukemogenesis and perturbation of Hox gene expression \(^{55}\). Recent studies documented that Erk1/2 and their upstream effectors Mek1/2 are constitutively activated in primary human AML cells \(^{72,73,74}\). Elevated p-ERK levels were found in 83.3% of the AML patients \(^{75}\). As aberrant expression of CDX2 is detected in
80% of all AML cases regulation of Cdx2 could be indirectly or directly linked with MEK/ERK signaling \(^{72,76}\) (Fig. 3a,b). Therefore, it will be interesting to test whether posttranscriptional modification of CDX2 is crucial for its leukemogenic properties and whether MAPK inhibitors are able to block Cdx2 induced leukemogenesis.

Another possibility to block Cdx2-induced leukemogenesis would be to inhibit mechanisms, which induce and maintain aberrant Cdx2 expression. However, mechanisms that lead to aberrant expression of CDX2 in leukemogenesis are still a mystery. As demonstrated by Scholl et al., the methylation levels of the promoter region of CDX2 in AML patients did not correspond to its expression levels \(^{56}\). In addition, treatment of CDX2-negative cell lines with demethylating agents did not lead to an induction of CDX2 expression \(^{77}\), suggesting that CDX2 expression in leukemia is independent of promoter hypomethylation as well as of the methylation level of cis-regulatory elements. Amplification of the CDX2 locus was detected only in a small percentage of patients with complex karyotype, however not in AML patients with normal karyotype, which are characterized by high CDX2 expression levels \(^{56,78}\). Another possibility, leading to aberrant transcription of CDX2 could be deregulation of trans-acting pathways, which target cis-regulatory elements in the 5' region of the CDX2 gene as suggested by Hinoi et al. \(^{77}\). However, it was reported that CDX2 expression is monoallelic in most of the analyzed leukemia patients \(^{56}\), which indicates that the altered CDX2 expression is not primarily due to differences in the expression of trans-acting factors but rather due to altered cis-regulatory elements, which affect only one allele. Mutations were not detected in the promoter region of CDX2. However, this does not exclude still unknown mutations in cis-regulatory elements of the CDX2 gene, which
cause aberrant expression. It would be interesting, therefore, to expand the sequence analysis of the CDX2 gene beyond its coding and promoter region and to conduct a more comprehensive sequence analysis. One target of genetic alterations for example could be negative regulators, which Wang and Shashikant postulated to be present in an 11.4 kb sequence flanking of Cdx2. Disruption of these elements in hematopoietic cells could lead to an aberrant expression of the normally silenced CDX2 as well as mutations within two enhancer regions in the first intron of CDX2, described to induce ectopic expression of Cdx2 in the murine embryo. Interestingly, several binding sites for FGF and Wnt signaling were identified in these regions. Mutations therefore could possibly affect the binding of factors such as FGF4, which previously was shown to target Cdx2 (Fig. 3). Furthermore, it is probable that Cdx2 not only regulates but also is targeted by Wnt signaling as already demonstrated for other ParaHox genes such as Cdx1 and Cdx4. Another pathway which was shown to regulate Cdx2 expression is Ras/MAPK signaling, a pathway which links extracellular stimuli amongst others to proliferation, differentiation and cell survival. As both the Wnt and Ras/MAPK pathways are frequently deregulated in leukemia, this deregulation in combination with genetic alterations in the cis-regulatory elements of Cdx2 could possibly affect the transcription of the gene (Fig. 3).

Conclusion

In summary, there is a great body of evidence that ParaHox genes play a pivotal role in normal hematopoiesis, but also in leukemogenesis. These data add another homeobox gene family to the growing number of homeobox genes known to be critically involved in
orchestrating normal hematopoiesis and promoting leukemogenesis. Thus, the story ‘homeobox genes in normal and malignant hematopoiesis’ is not at its end but is still growing rapidly since the last review on this topic in this journal nearly 10 years ago. In these 10 years dysregulation of homeobox genes has been recognized as one of the molecular hallmarks of human AML. We know today that homeobox genes are key players for determining ‘stemness’ and are essential for initiating and maintaining self-renewal programs on which leukemic stem cell function depends. ParaHox genes such as Cdx2 are particularly powerful to transform hematopoietic progenitors into leukemic stem cells, at least partly because it stands upstream of Hox genes and is able to perturb expression of several Hox genes at the same time. It is intriguing to speculate that targeting dysfunctional ParaHox genes such as CDX2 in human leukemia might reverse HOX gene dysfunction on a larger scale thereby depriving the leukemic stem cell the molecular ‘environment’ it needs to propagate leukemic growth. It is indeed unfortunate that we know so little about the mechanisms which induce aberrant CDX2 expression in AML patients as this would open possibilities to develop strategies to re-silence CDX2 expression in leukemic cells. Another hope is to identify signaling cascades which are essential for CDX2 induced transformation. Many questions are still open and clarification of the role of other members of the ParaHox gene cluster such as GSH2 for leukemogenesis was just initiated. We can be confident that the next years will generate exciting data about the role of ParaHox genes for normal tissue homeostasis and will improve our understanding how to correct ParaHox dysfunction in human leukemias.
Acknowledgements:

We apologize to authors whose work could not be cited due to space constraints. We acknowledge all members of our institutes for vivid discussions on this topic. We like to thank Silvia Thoene, Shiva Bamezai and Tamoghna Mandal for assisting in the preparation of figures. Our work cited is supported by grants of the Else Kröner-Fresenius-Foundation, the DFG and the BMBF (all to V.P.S. Rawat & C. Buske).

Author Contributions:

V.P.S.R., C.B. and R.K.H all contributed to the writing of the manuscript. The authors have no conflicts of interest to disclose.
References:


<table>
<thead>
<tr>
<th>Human Gene</th>
<th>Protein size (kDa)</th>
<th>Homology between man and mouse*</th>
<th>Chromosomal translocation</th>
<th>Aberrant expression in leukemia</th>
<th>Transforming potential</th>
<th>Phenotype of knockout mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSH1</td>
<td>27</td>
<td>98.01%</td>
<td>------</td>
<td>unknown</td>
<td>Not tested</td>
<td>Phenotypically normal at birth, but 75% of mice die at 4 weeks of age. Marked growth retardation, infertile, low WBC count.</td>
</tr>
<tr>
<td>GSH2</td>
<td>32</td>
<td>89.14%</td>
<td>CHIC2-ETV6&lt;sup&gt;48&lt;/sup&gt;</td>
<td>AML (case reports)&lt;sup&gt;48&lt;/sup&gt;</td>
<td>transforms NIH 3T3 cells in vitro&lt;sup&gt;48&lt;/sup&gt;</td>
<td>Born healthy, but death within 24h after birth. Homozygous mutants suffer from apnea and severe brain alterations.&lt;sup&gt;33&lt;/sup&gt;</td>
</tr>
<tr>
<td>PDX1</td>
<td>30</td>
<td>87.99</td>
<td>------</td>
<td>unknown</td>
<td>Not tested</td>
<td>Death at 1 dpp; apancreatic.&lt;sup&gt;34&lt;/sup&gt;</td>
</tr>
<tr>
<td>CDX1</td>
<td>28</td>
<td>85.28</td>
<td>------</td>
<td>Not expressed&lt;sup&gt;55&lt;/sup&gt;</td>
<td>Not tested</td>
<td>Viable and fertile, anterior homeotic transformations; posterior shifts of Hox gene expression in the somitic mesoderm.&lt;sup&gt;35&lt;/sup&gt;</td>
</tr>
<tr>
<td>CDX2</td>
<td>33</td>
<td>93.89</td>
<td>ETV6-CDX2&lt;sup&gt;52&lt;/sup&gt;</td>
<td>AML, ALL, CML in blast crises&lt;sup&gt;32,55-56&lt;/sup&gt;</td>
<td>causes AML in mice&lt;sup&gt;63&lt;/sup&gt;</td>
<td>peri - implantation embryonic lethality; heterozygous mutants show growth retardation and anterior transformation.&lt;sup&gt;36&lt;/sup&gt;</td>
</tr>
<tr>
<td>CDX4</td>
<td>30</td>
<td>82.98</td>
<td>------</td>
<td>AML&lt;sup&gt;68&lt;/sup&gt;</td>
<td>Overexpression induces AML in mice&lt;sup&gt;45,68&lt;/sup&gt;</td>
<td>Viable, only a mild anterior transformation with very low penetrance.&lt;sup&gt;39&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

* = at amino acid level
Figure Legends

Figure 1: Characteristics of ParaHox genes

a) Chromosomal location of human Parahox genes. Chromosomal locations of murine ParaHox genes are given in brackets. The direction of transcription is indicated by an arrow. b) Protein structure of Parahox genes: shown is the protein structure depicting the N-terminal transactivation domain (blue), the Pbx interacting motif (PIM)(brown), the homeodomain (green) and validated phosphorylation sites (orange)

Figure 2. Cdx expression in normal and malignant hematopoiesis

a) Normal hematopoiesis –\textit{Cdx4} and \textit{HOX A} and B cluster genes are expressed in normal stem and progenitor cells while \textit{Cdx2} and \textit{Cdx1} remain silenced. b) Malignant hematopoiesis – ectopic expression of \textit{CDX2} at the level of stem or progenitor cells induces increased expression of Hox A and B cluster genes and leukemic transformation of hematopoietic cells. The proportion of AML patients, which ectopically express \textit{CDX2} and overexpress HOX genes is indicated. \textit{CDX4} is expressed only in a minority of patients and \textit{CDX1} expression is not detectable in leukemia patients.

Figure 3. Hypothetical model of Cdx gene regulation in normal and malignant hematopoiesis

a) Normal hematopoiesis: based on results from studies in embryogenesis the canonical Wnt pathway regulates the expression of \textit{CDX4} by the Lef1/Tcf complex in hematopoiesis. The FGF - MAPK pathway is not constitutively active in stem and progenitors cells. b) Malignant hematopoiesis: either deregulation of the WNT pathway or FGF signaling induces aberrant CDX2 expression in human leukemia.
Fig. 1

a)

![Diagram of chromosome regions showing GSH, PDX, and CDX with their respective chromosome locations: Ch: 13q12, Ch: 5q31-33, Ch: 4q12-13, and Ch: X13.2.]

b)

![Diagram showing the amino acid positions of PDX1, GSH1, GSH2, CDX1, CDX2, and CDX4, highlighting the transactivation domain, PIM domain, homeodomain, and phosphorylation sites.]
Fig. 2

(a) Normal Hematopoiesis

Normal Hematopoiesis

Cdx/CDX

CDX2 No expression
CDX1 No expression

HOX expression

CDX4 expression

Stem cell → Progenitor → Oligo-lineage precursor → Mature cells

(b) Malignant Hematopoiesis

Malignant Hematopoiesis

Cdx/CDX

HOX expression

Acute Myeloid Leukemia

<table>
<thead>
<tr>
<th>Expression</th>
<th>% of positive patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOX</td>
<td>100</td>
</tr>
<tr>
<td>CDX4</td>
<td>80</td>
</tr>
<tr>
<td>CDX2</td>
<td>90</td>
</tr>
<tr>
<td>CDX1</td>
<td>0</td>
</tr>
</tbody>
</table>

Stem cell → Transformed stem cells → Leukemia initiating cells
Fig. 3

a) Normal hematopoiesis

b) Malignant hematopoiesis (hypothetical model)
Beyond HOX: the role of ParaHox genes in normal and malignant hematopoiesis

Vijay P.S. Rawat, R. Keith Humphries and Christian Buske