Published in final edited form as:

Blood Cells Mol Dis. 2008 ; 40(2): 156–159.

Transcriptional Complexity of the *HOXA9* Locus

Relja Popovic¹, Frank Erfurth², and Nancy Zeleznik-Le^{1,2,3,*}

¹ Molecular Biology Program, Loyola University Chicago, Maywood, Illinois

² Oncology Institute, Loyola University Chicago, Maywood, Illinois

³ Department of Medicine, Loyola University Chicago, Maywood, Illinois

Abstract

Evolutionarily conserved *HOX* genes play an important role during development and hematopoiesis. *HOX* protein products are transcription factors whose precise mechanism of action is still poorly understood. Regulation of *HOX* gene expression has been the topic of various studies. While alternative splicing and alternative promoter usage have been known to increase the number of transcripts across the *HOX* clusters, more recently high-throughput analyses have identified a number of new coding and noncoding RNA molecules whose function is not known. Here we review the transcriptome of the most studied *HOX* locus, *HOXA9*. Strict control of *HOXA9* expression has been shown to play a critical role in hematopoiesis while aberrant expression has been shown important to the development of leukemia. However, it is still unclear how various transcripts from this locus are regulated and what specific role(s) each one of them plays.

Introduction

Drosophila HOM genes were identified in flies with homeotic mutations, in which a body segments, such as antennae, are converted to different segments, such as legs[1]. Subsequently, mammalian homologues of *HOM* genes were identified and named *HOX* genes. Both *HOM* and *HOX* genes play an essential role during embryogenesis by binding to DNA through a highly conserved 183-nt sequence called the homeodomain. Homeodomain containing proteins are transcription factors that play a role in determination of segment identity of the developing embryo along various body axes[2]. During evolution, cluster duplication gave rise to four mammalian *HOX* clusters located on different chromosomes, with a total of 39 *HOX* genes found in 13 different paralogous groups. In each cluster, the spatial and temporal pattern of expression of the *HOX* genes corresponds to their physical position within the cluster. The 3' genes are expressed earlier than the 5' genes, have a more anterior boundary of expression and have a higher sensitivity to transcriptional inducers such as retinoic acid[3]. More recent data has shown that *HOX* genes are also involved in hematopoietic cell development. Certain *HOX* genes are expressed in a lineage specific pattern with 5' genes of the *HOXA* cluster being highly expressed in primitive hematopoietic progenitors and downregulated in more differentiated cells[4].

Analysis of the *HOX* cluster transcriptome shows an extensive repertoire of transcripts whose function is still poorly understood. Extensive splicing and alternative promoter usage adds an

*Corresponding author Nancy Zeleznik-Le, 2160 South 1st Avenue, Bldg 112, Room 337, Maywood, IL 60153, Tel: (708) 327-3368, Fax: (708) 327-3342, Email: nzelezn@lumc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

additional layer of complexity. It is proposed that some *HOX* genes share common cis-regulatory elements which help in the colinear pattern of *HOX* gene expression. Polycistronic messages have also been described for certain *HOX* genes[5]. In most of the knockout and overexpression studies only canonical forms of *HOX* transcripts have been studied, without truly discriminating between similar isoforms. In recent years, a number of noncoding transcripts have also been identified making the *HOX* clusters an even more complex system to decipher. Some of these are antisense transcripts which play a role in regulation of *HOX* gene expression[6]. Of special interest is the presence of microRNAs in the *HOX* cluster which may partially explain the importance of the whole *HOX* cluster in development and disease. Discovery of numerous noncoding and regulatory transcripts in the *HOX* clusters has raised additional questions regarding the role of each one of these transcripts.

HOXA9 Locus

The most studied *HOXA* cluster gene is *HOXA9*. Among 6800 genes tested, *HOXA9* was shown to be the single most correlative marker of poor prognosis in patients with acute myeloid leukemia[7]. This finding confirms previous reports outlining the significant role that *HOXA9* plays not only during normal hematopoiesis but also in leukemia. *Hoxa9* deficient mice have impaired hematopoietic development[8]. Furthermore, the *Hoxa9* locus is a common site of retroviral integrations in leukemia-prone BXH2 mice[9,10]. Resulting leukemias express high levels of *Hoxa7* and *Hoxa9* genes[9]. Overexpression of *Hoxa9* and *Meis1* also leads to immortalization of mouse bone marrow cells and leukemia development in mice[11]. It is not surprising that *Hoxa9*, as a transcription factor, regulates genes involved in stem cell biology[12]. While much work remains to be done on the downstream targets of *HOXA9*, studies have shown that *HOXA9* is able to function both as transcriptional activator and repressor. The *HOXA9* gene itself is a target of leukemic translocation involving chromosome 11(*HOXA9*) and chromosome 7 (*NUP98*) that encodes the NUP98-*HOXA9* oncogenic protein [13,14]. NUP98-*HOXA9* transformation of NIH-3T3 cells depends on *HOXA9*-mediated DNA binding, strengthening the idea that misregulation of *HOXA9* targets is important in disease development[15].

The canonical *HOXA9* transcript contains two exons: exon *CD* and the homeodomain-containing exon II (Fig. 1). This transcript is expressed in variety of embryonic and adult tissues [16]. A portion of exon *CD* can be alternatively spliced giving rise to a different new transcript with a shifted reading frame[17]. The splice site found in the first exon is present not only in humans, but in other species as well suggesting that this alternatively spliced transcript may play an important role in development and/or disease[16]. The frame shift causes a premature stop codon upstream of the homeodomain, leading to production of a non-homeodomain containing protein[17]. Absence of the homeodomain ablates DNA binding activity of the protein. Like canonical *HOXA9*, this alternatively spliced transcript is expressed in numerous tissues and its expression is conserved across species. The truncated protein is found in both the nucleus and cytoplasm which is in contrast to the homeodomain containing *HOXA9* found exclusively in the nucleus. Like canonical *HOXA9*, the truncated protein binds CBP with high efficiency and the two isoforms may compete for CBP as well as for other *HOXA9* interacting proteins[16]. If this model is accurate, the truncated protein may interfere with normal *HOXA9* function as well as have other novel functions that have not been described yet.

Exon II is a particularly interesting region not only because it codes for the DNA binding domain, but because together with the 3' UTR is the most shared region among various *HOXA9* isoforms. Studies have shown existence of a number of transcripts that utilize exon II of the *HOXA9* gene. One of these transcripts starts much more upstream from the canonical *HOXA9*, in the vicinity of the *HOXA11* gene[18]. Other isolated transcripts utilize the previously described splice acceptor site within the *CD* exon and contain all of exon II[17,

Popovic (unpublished)]. The majority of the transcripts presumably share the common 3'UTR as well. Studies that use 3'UTR tags for detecting expression levels of *HOXA9* are not able to differentiate between the various isoforms and their data represent the “total” *HOXA9* pool. More specific probes will need to be designed to dissect the pattern of expression for each one of the transcripts.

Another intriguing region in the *HOXA9* locus is located approximately 4.2kb upstream of the canonical promoter and proximally 3' of the *HOXA10* gene. This region is very highly conserved across multiple species and encodes an additional *HOXA9* exon designated *AB*. The *AB* exon is employed in a number of cloned transcripts as an alternative first exon in lieu of exon *CD*. While some RNAs are spliced to give *AB-exon II* transcripts, others use the internal *CD* splice acceptor site to form an *AB-CD-exon II* transcript[13,19, Popovic(unpublished)]. We have isolated two of these transcripts that differ only in 40bp of sequence used at the end of the *AB* exon. Both of the transcripts are found in bone marrow progenitor cells and sequence analysis predicts that they do not code for protein. It is still not understood the mechanism by which these RNAs are regulated or their precise transcription start site. Recently, Mainguy *et al.* described an antisense transcript transcribed from this region[18]. Interestingly, microRNA 196b is also located in this area and the majority of mir-196 targets tested so far are *HOX* genes themselves[20]. It is possible that the expression of sense and antisense transcripts regulates not only the transcription from this region but also help regulating expression of other *HOX* genes. This mode of regulation would aid in the collinear model of *HOX* gene expression.

Conclusions

There remain many unanswered questions regarding transcriptional regulation of the *HOXA9* locus. It is clear that more detailed studies need to be performed to appreciate the difference in expression patterns of the various *HOXA9* transcripts. Understanding the functional differences between all these transcripts will allow us to better comprehend what goes wrong in the diseases associated with *HOXA9* misexpression. For a long time dogma considered the majority of the DNA sequence to be ‘garbage’, but with the discovery of noncoding RNA molecules, we began to realize the complexity of the transcriptional regulatory machinery. Proper expression of *HOXA9* is evidently very important in normal development and disease. Different *HOXA9* isoforms may possess alternative functions and thus play diverse roles in normal processes as well as in disease progression. Together with other *HOX* genes, *HOXA9* is overexpressed in MLL-associated leukemias, however it is still unclear whether this is also true for all of the transcripts from the locus. MLL may differentially regulate expression levels of the various RNAs and in the presence of MLL fusion proteins the normal transcript balance may be altered. Also, the *HOXA9* locus is a common site for chromosomal rearrangements with the TCR β locus[21]. Patients with those types of translocations have high levels of *HOX* gene expression, including *HOXA9*. The chromosomal breakage occurs in the region between exons *CD* and *AB* and may influence misexpression of multiple transcripts. In efforts to develop successful therapies for treatment of leukemia and other diseases, we must understand the nature and function of the molecules we are trying to target. In the case of *HOXA9*, there is much work that remains to be done.

Acknowledgements

This paper is based on a presentation at the Seventh International Workshop on Molecular Aspects of Myeloid Stem Cell Development and Leukemia in Annapolis, Maryland May 13–16, 2007, sponsored by The Leukemia & Lymphoma Society.

References

1. Bridges, CB.; Morgan, TH. The third-chromosome group of mutant characters of *Drosophila melanogaster*. Carnegie Institution of Washington; Washington: 1923. p. 1-251.
2. Gehring WJ, Affolter M, Burglin T. Homeodomain proteins. Annual Review of Biochemistry 1994;63:487–526.
3. Krumlauf R. Hox genes in vertebrate development. Cell 1994;78:191–201. [PubMed: 7913880]
4. Sauvageau G, Lansdorp PM, Eaves CJ, et al. Differential expression of homeobox genes in functionally distinct CD34+ subpopulations of human bone marrow cells. Proc Natl Acad Sci U S A 1994;91:12223–12227. [PubMed: 7527557]
5. Simeone A, Pannese M, Acampora D, et al. At least three human homeoboxes on chromosome 12 belong to the same transcription unit. Nucleic Acids Res 1988;16:5379–5390. [PubMed: 2898768]
6. Sessa L, Breiling A, Lavorgna G, et al. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. RNA 2007;13:223–239. [PubMed: 17185360]
7. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–537. [PubMed: 10521349]
8. Lawrence HJ, Helgason CD, Sauvageau G, et al. Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid and lymphoid hematopoiesis. Blood 1997;89:1922–1930. [PubMed: 9058712]
9. Nakamura T, Largaespada DA, Shaughnessy JD Jr, et al. Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukaemias. Nat Genet 1996;12:149–153. [PubMed: 8563752]
10. Bijl J, Sauvageau M, Thompson A, Sauvageau G. High incidence of proviral integrations in the Hoxa locus in a new model of E2a-PBX1-induced B-cell leukemia. Genes Dev 2005;19:224–233. [PubMed: 15655112]
11. Kroon E, Kros J, Thorsteinsdottir U, et al. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. EMBO J 1998;17:3714–3725.
12. Dorsam ST, Ferrell CM, Dorsam GP, et al. The transcriptome of the leukemogenic homeoprotein HOXA9 in human hematopoietic cells. Blood 2004;103:1676–1684. [PubMed: 14604967]
13. Borrow J, Shearman AM, Stanton VP Jr, et al. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet 1996;12:159–167. [PubMed: 8563754]
14. Nakamura T, Largaespada DA, Lee MP, et al. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat Genet 1996;12:154–158. [PubMed: 8563753]
15. Kasper LH, Brindle PK, Schnabel CA, et al. CREB binding protein interacts with nucleoporin-specific FG repeats that activate transcription and mediate NUP98-HOXA9 oncogenicity. Mol Cell Biol 1999;19:764–776. [PubMed: 9858599]
16. Dintilhac A, Bihan R, Guerrier D, et al. A conserved non-homeodomain Hoxa9 isoform interacting with CBP is co-expressed with the 'typical' Hoxa9 protein during embryogenesis. Gene Expr Patterns 2004;4:215–222. [PubMed: 15161102]
17. Fujimoto S, Araki K, Chisaka O, et al. Analysis of the murine Hoxa-9 cDNA: an alternatively spliced transcript encodes a truncated protein lacking the homeodomain. Gene 1998;209:77–85. [PubMed: 9524228]
18. Mainguy G, Koster J, Woltering J, et al. Extensive polycistronism and antisense transcription in the Mammalian hox clusters. PLoS ONE 2007;2:e356. [PubMed: 17406680]
19. Kim MH, Chang HH, Shin C, et al. Genomic structure and sequence analysis of human HOXA-9. DNA Cell Biol 1998;17:407–414. [PubMed: 9628584]
20. Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science 2004;304:594–596. [PubMed: 15105502]
21. Soulier J, Clappier E, Cayuela JM, et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood 2005;106:274–286. [PubMed: 15774621]

Figure 1.

Schematic diagram of the *HOXA9* locus and various cloned transcripts. Exons are represented with dark boxes on the top bar. Translocation sites for *TCRβ*, *NUP98* and retroviral integration sites are labeled with arrows. Each transcript is labeled with an associated reference or EST accession number. Dark boxes on the bottom bar represent CpG islands.