

# Analysis of the DNA-Binding Profile and Function of TALE Homeoproteins Reveals Their Specialization and Specific Interactions with Hox Genes/Proteins

Dmitry Penkov,<sup>1,2,7</sup> Daniel Mateos San Martín,<sup>3,7</sup> Luis C. Fernandez-Díaz,<sup>1</sup> Catalina A. Rosselló,<sup>3</sup> Carlos Torroja,<sup>4</sup> Fátima Sánchez-Cabo,<sup>4</sup> H.J. Warnatz,<sup>5</sup> Marc Sultan,<sup>5</sup> Marie L. Yaspo,<sup>5</sup> Arianna Gabrieli,<sup>1</sup> Vsevolod Tkachuk,<sup>2</sup> Andrea Brendolan,<sup>6</sup> Francesco Blasi,<sup>1,\*</sup> and Miguel Torres<sup>3,\*</sup>

<sup>1</sup>IFOM (Foundation FIRC Institute of Molecular Oncology) at the IFOM-IEO Campus, via Adamello 16, 20139 Milan, Italy

<sup>2</sup>Department of Basic Medicine, Lomonosov Moscow State University, Lomonosov Prospect, 31/5, 119192, Moscow, Russia

<sup>3</sup>Cardiovascular Development and Repair Department

<sup>4</sup>Bioinformatics Unit

Centro Nacional de Investigaciones Cardiovasculares (CNIC), Melchor Fernández Almagro 3, 28029 Madrid, Spain

<sup>5</sup>Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, 14195 Berlin, Germany

<sup>6</sup>San Raffaele Scientific Institute, Division of Molecular Oncology, via Olgettina 60, 20123, Milan, Italy

<sup>7</sup>These authors contributed equally to this work

\*Correspondence: francesco.blasi@ifom.eu (F.B.), mtorres@cnic.es (M.T.)

<http://dx.doi.org/10.1016/j.celrep.2013.03.029>

## SUMMARY

The interactions of Meis, Prep, and Pbx1 TALE homeoproteins with Hox proteins are essential for development and disease. Although Meis and Prep behave similarly *in vitro*, their *in vivo* activities remain largely unexplored. We show that Prep and Meis interact with largely independent sets of genomic sites and select different DNA-binding sequences, Prep associating mostly with promoters and house-keeping genes and Meis with promoter-remote regions and developmental genes. Hox target sequences associate strongly with Meis but not with Prep binding sites, while Pbx1 cooperates with both Prep and Meis. Accordingly, *Meis1* shows strong genetic interaction with *Pbx1* but not with *Prep1*. *Meis1* and *Prep1* nonetheless coregulate a subset of genes, predominantly through opposing effects. Notably, the TALE homeoprotein binding profile subdivides Hox clusters into two domains differentially regulated by *Meis1* and *Prep1*. During evolution, Meis and Prep thus specialized their interactions but maintained significant regulatory coordination.

## INTRODUCTION

The specificity of transcription in a crowded eukaryotic chromatin is something of a mystery. Different members of closely related transcription factor families bind near-identical DNA sequences *in vitro*, but their individual function *in vivo* is rarely known. Transcription factors may also bind different cofactors, resulting in differing patterns of DNA recognition and binding. An example is provided by the Hox and TALE (three amino acid loop exten-

sion) families (Moens and Selleri, 2006), which have similar DNA-binding domains. Interaction between the *Drosophila* TALE proteins Extradenticle (Exd) and Homothorax (Hth) targets the two proteins to the nucleus (Chan et al., 1994; Rieckhof et al., 1997) where the complex interacts with Hox proteins, determining their DNA-binding specificity and thereby anteroposterior segmental identity (reviewed in Mann and Affolter, 1998).

The genomes of mammals contain four Exd-related genes (*Pbx*) and two Hth-related subfamilies, *Meis* and *Prep* (the latter also known as *pKnox*), respectively comprising three and two members. The interaction of Exd with Hth or Hox has been retained in all species, and hence in vertebrates Pbx proteins form complexes with Hox, Meis, and Prep. Pbx proteins interact with Prep or Meis through a conserved amino-terminal domain (Berthelsen et al., 1998; Chang et al., 1997; Knoepfler et al., 1997) and with Hox proteins through the homeodomain (Piper et al., 1999). The independent interaction surfaces allow Pbx to form trimers with Prep or Meis and Hox, and this interaction alters the DNA-binding selectivity of the individual Hox proteins (Ferretti et al., 2000; Jacobs et al., 1999; Ryoo et al., 1999). Meis, but not Prep, can also interact directly with posterior Hox proteins (Williams et al., 2005).

The full complexity of the TALE transcriptional regulatory network *in vivo* has not even been estimated. Our knowledge of these factors' DNA sequence specificity is based on *in vitro* selection of target sequences by purified or *in vitro*-translated protein complexes and on the analysis of a limited number of endogenous target sequences. A general observation is that affinity for DNA is low for monomers and increases with heterologous complex formation. Prep and Meis alone preferentially bind the TGACAG hexameric sequence (PM sites) (Berthelsen et al., 1998; Chang et al., 1997; Ferretti et al., 2000; Shen et al., 1997) and Pbx to the TGATTGAT sequence (LeBrun and Cleary, 1994). Prep-Pbx and Meis-Pbx dimers both preferentially bind the decameric sequence TGATTGACAG (Chang et al., 1997; Knoepfler et al., 1997). Pbx-Hox dimers bind octameric motifs

of the type TGATNNAT, in which the variable core determines the Hox paralog group binding (Shen et al., 1997). Studies combining oligonucleotide selection (SELEX) with deep sequencing (SELEX-seq) in *Drosophila* show that the site variants at the variable core can be grouped into three main classes of specificity that obey the colinearity rules and underline the preference of Hox for distinct DNA minor groove topographies (Slattery et al., 2011). X-ray studies showed that in Pbx-Hox binding to the octameric sites, each monomer binds one half-site (LaRonde-LeBlanc and Wolberger, 2003; Piper et al., 1999). Ternary complexes take place through the interaction between Meis/Prep bound to hexameric sites and nearby Pbx-Hox bound to octameric sites through direct Prep/Meis-Pbx interaction (Berthelsen et al., 1998; Ferretti et al., 2000, 2005; Ryoo et al., 1999). Ternary complexes can also form by Meis1 interaction with DNA-bound Pbx-Hox dimers without Meis1 binding DNA (Shanmugam et al., 1999).

Meis and Prep proteins contain two homologous functional domains: the Pbx-interacting domain and the homeodomain. The homeodomains and Pbx-interacting regions of Meis1 and Prep1 are 84% and 63% identical. However, other regions of the two proteins are not conserved, including the C-terminal domain, which is essential for Meis1 oncogenic activity (Bisailon et al., 2011; Wong et al., 2007). Both Prep1 and Meis1 dimerize with Pbx and recognize similar DNA sequences in vitro. Although some specific functions have been identified for Prep and Meis, there is no information about whether their activities are coordinated in vivo.

*Prep1* is ubiquitously expressed from the oocyte to the embryo and the adult (Fernandez-Diaz et al., 2010; Ferretti et al., 1999, 2006). *Meis1* and *Meis2* encode very similar proteins (Moskowitz et al., 1995; Nakamura et al., 1996), and their expression starts around gastrulation and is regionalized (Cecconi et al., 1997; Oulad-Abdelghani et al., 1997). *Pbx1*, *Prep1*, and *Meis1* are developmentally essential genes. *Prep1* null embryos die shortly after implantation, with massive apoptosis and proliferation defects (Fernandez-Diaz et al., 2010). *Pbx1* deletion is embryonically lethal at embryonic day (E) 15.5, and embryos display major homeotic anomalies, organ absence or hypoplasia, hematopoietic defects, and other features (DiMartino et al., 2001; Selleri et al., 2001). *Meis1*-deficient mice die at E14.5 with definitive hematopoietic stem cell failure, megakaryocyte lineage aplasia, lymphatic vasculature defects, heart defects, and eye hypoplasia (Azcoitia et al., 2005; Hisa et al., 2004). While *Prep1* null embryos die early, hypomorphic mutants (*Prep1<sup>hi</sup>*), in which only 3%–7% of the wild-type protein is produced, show variable viability during gestation. *Prep1* hypomorphs show defects in hematopoiesis, including hematopoietic stem cells, eye development, and angiogenesis (Di Rosa et al., 2007; Ferretti et al., 2006). Although impairment of eye development, hematopoiesis, and angiogenesis is common in *Meis1* and *Prep1<sup>i</sup>* mutants, the specific aspects affected are different. In addition, the involvement of these factors in disease is clearly divergent, since *Prep1* acts as a tumor suppressor (Iotti et al., 2011; Longobardi et al., 2010) while *Meis1* is leukemogenic (Moskowitz et al., 1995; Wong et al., 2007) and *Meis1* leukemogenic activity cannot be replaced by *Prep1* (Thorsteinsdottir et al., 2001).

We have undertaken a comprehensive comparative analysis of the genomic interaction and function of Meis, Prep, and

Pbx1 in mouse embryos in vivo. We show that Meis and Prep mostly select distinct genomic sites and DNA motifs and show differential interactions with Hox genes and proteins. Our analysis establishes a framework for understanding the mechanisms of action of TALE proteins in development and disease.

## RESULTS

### Prep and Meis Select, and Drive Pbx1 to, Different Genomic Sites

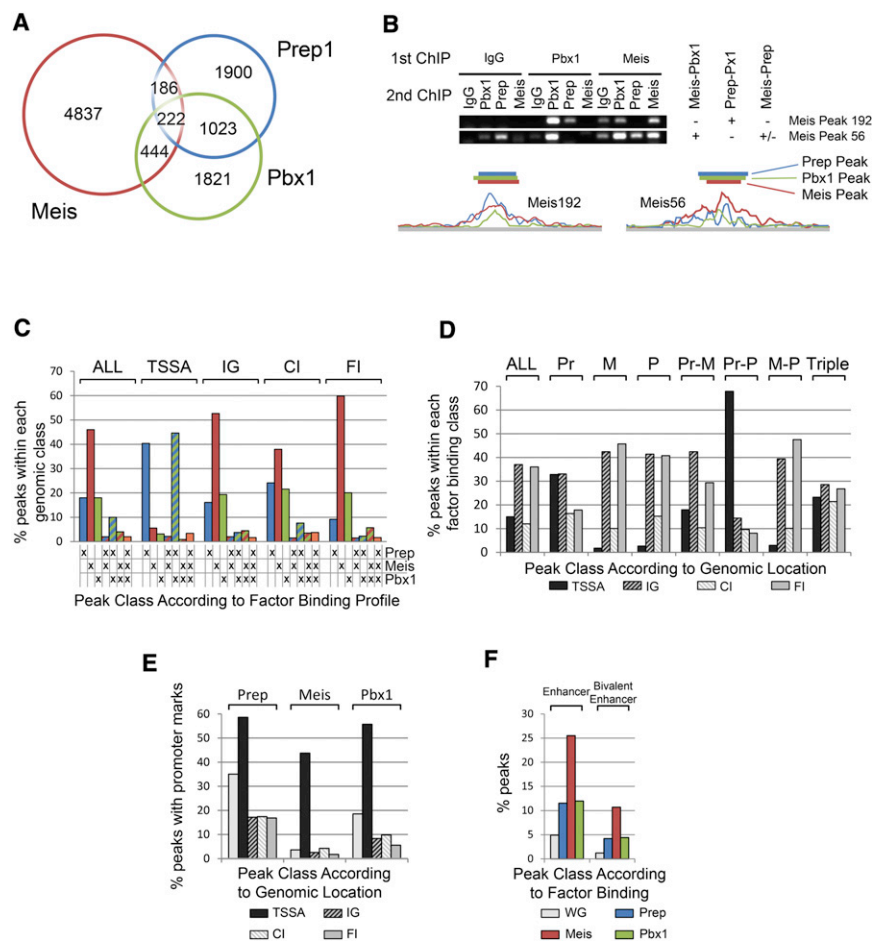
Chromatin immunoprecipitation sequencing (ChIP-seq) on E11.5 embryos with antibodies to Prep1/2, Pbx1, or Meis1/2 (see Experimental Procedures) detected 3,331 peaks for Prep, 5,686 for Meis, and 3,504 for Pbx1 (Tables S1 and S2) (Gene Expression Omnibus [GEO] accession number GSE39609). The nonredundant peak list contains 10,326 genomic regions, of which 82% correspond to single-factor-bound regions, 16% to two-factor-bound regions, and 2% to regions bound by all three (Figure 1A). About half of the Pbx1 and Prep peaks were exclusively bound by these factors (Pbx1<sup>exc</sup>, Prep<sup>exc</sup>), while 85% of Meis peaks were exclusive (Meis<sup>exc</sup>), suggesting more independent activities for Meis compared with Prep and Pbx1 (Figure 1A). Analysis of peak overlaps revealed a lower coincidence of Prep with Meis (Prep-Meis<sup>com</sup>) than with Pbx1 (5.6% versus 30% of Prep peaks; Figure 1A). Almost 30% of the Pbx1 peaks were also bound by Prep, a larger proportion than by Meis (12.6%; Figure 1A). An additional 6% of Pbx1 peaks were simultaneously bound by Prep and Meis (triple peaks).

ChIP-re-ChIP assays of double and triple peaks confirmed that Pbx1 binds simultaneously with either Prep or Meis in a majority of sites (10/17 for Prep and 17/21 for Meis) (Figure S1A). In contrast, in 15/17 Meis-Prep common peaks, these factors do not show simultaneous binding (Figures 1B and S1A). In triple peaks, the most frequent situation was thus alternative binding by either Prep+Pbx1 or Meis+Pbx. The mapping of the relative positions between pairs of the three factors in triple peaks indicates that their binding preferences are cocentered (Figure S1B), indicating that in most cases they bind to the same sequences. Given that in most sites there is no simultaneous binding of Meis and Prep, these factors may compete for binding to the same sequences. The infrequent cases of the simultaneous co-binding of Meis and Prep may thus correspond to the independent binding of the two factors to neighboring target sequences.

In relative terms, Prep-Pbx1 cobinding is therefore predominant with respect to Meis-Pbx1 cobinding. In addition, the analysis of Prep-Pbx2 site occupancy in the thymus, where only Pbx2 is expressed, indicates that the embryonic Pbx1-Prep<sup>com</sup> peaks can be bound by Prep-Pbx2 when Pbx1 is not available (Tables S2 and S3). These data show that Prep-Pbx interactions are predominant in Prep targets and can occur with different Pbx partners.

### Prep Binding Sites Correlate with Transcription Start Sites, while Meis Binding Sites Concentrate in Transcription-Start-Site-Remote Regions

We next analyzed the distribution of the peaks according to their position with respect to RefSeq genes. We classified peaks as transcription start site associated (TSSA) when they appeared



**Figure 1. Meis and Prep Select Different Binding Sites and Gene-Regulatory Regions in Cooperation with Pbx1**

(A) Venn diagram of peak classes containing single, double, and triple binding by Meis, Prep, and Pbx1. Prep-Meis overlap versus Prep-Pbx overlap and Pbx1-Prep versus Pbx1-Meis adjusted p values ( $\text{adjp}$ ) < 0.0001.

(B) ChIP-re-ChIP experiment. Top: PCR amplification of consecutive immunoprecipitations with anti-immunoglobulin G (IgG), anti-Meis, anti-Prep or anti-Pbx1 antibodies. Bottom: Read profile of peaks tested above. Color bands represent peaks as called by PICS, and color lines represent read density. The interpretation of the cobinding is shown to the right of the gel and is based on the comparison of the specific band intensity with that of the control immunoprecipitations (IgG).

(C) Percentage of peaks located in transcription-start site-associated (TSSA), intragenic (IG), close-intergenic (CI), or far-intergenic (FI) regions that belong to each factor binding class.

(D) Percentage distribution of genomic location classes within each factor binding profile category ( $\text{adjp} = 1$  for Meis<sup>exc</sup> versus Meis-Pbx1<sup>com</sup> and  $p = 0.26$  for Pbx1<sup>exc</sup> versus Meis-Pbx1<sup>com</sup>;  $\text{adjp} < 0.0001$  for Prep enrichment in TSSA and deployment in FI classes and for Prep-Pbx1<sup>com</sup> profile versus that of either Prep or Pbx1 alone).

(E) Percentage of Meis, Prep, or Pbx1 peaks containing promoter marks within each genomic location ( $\text{adjp} < 0.0001$  for Prep promoter marks preference in any class and for Meis, only in the TSSA class).

(F) Percentage of Meis, Prep, and Pbx1 peaks containing enhancer marks ( $\text{adjp} < 0.0001$  for Meis association with enhancer marks and  $\text{adjp} < 0.01$  for Prep and Pbx1). WG, whole genome.

See also Figures S1, S2, and Tables S1, S2, and S3.

within  $-500$  to  $+100$  bp from a transcription start site (TSS), intragenic (IG) when they overlapped a transcription unit, close intergenic (CI) when they appeared  $<20$  kb from a TSS, and far intergenic (FI) when they were located  $>20$  kb from the closest TSS. We first studied the abundance of the different peak classes defined by factor binding profile within each of these genomic regions. Within the TSSA class, the most abundant peaks were Prep<sup>exc</sup> and Prep-Pbx1<sup>com</sup> sites, which together represented 85% of all TSSA peaks (Figure 1C). In contrast, in all other genomic regions, single-factor-bound sites were the most abundant peak classes, with Meis<sup>exc</sup> peaks predominating in all classes, but especially in the IG and FI classes. Peaks in the Prep<sup>exc</sup> and Pbx1<sup>exc</sup> lists were moderately represented in non-TSSA classes, where peaks bound by more than one factor were generally of low abundance. Within the IG class, peaks for Prep and Pbx1 show a neutral distribution between exonic and intronic regions; however, Meis peaks show a 4-fold reduction in the expected occurrence in exons ( $p < 0.0001$  for Meis,  $p = 0.1$  for Prep, and  $p = 1$  for Pbx1).

To determine how cobinding modifies the binding preferences of each factor, we studied the distribution of different genomic regions across the peak classes defined by factor binding profile (Figure 1D). Binding distributions for Pbx1<sup>exc</sup>, Meis<sup>exc</sup>, and Meis-

Pbx1<sup>com</sup> were very similar, with low preference for TSSA regions and CI and high preference for IG and FI compared with the distribution shown by all peaks (Figure 1D). These data indicate that Pbx1 and Meis have similar preferences individually and that their cobinding does not change these preferences. Prep alone, in contrast, showed a strong preference for TSSA regions (41-fold enrichment compared to genomic TSSA region content) and a low preference for FI regions. Unlike Meis-Pbx1<sup>com</sup> peaks, the Prep-Pbx1<sup>com</sup> profile diverged sharply from that observed for each factor in isolation, with a marked prevalence of binding to TSSA regions (71.5% for Prep-Pbx1<sup>com</sup> versus 2.6% and 32.8% for Pbx1<sup>exc</sup> and Prep<sup>exc</sup>, respectively) and underrepresentation of all other regions with respect to the Prep<sup>exc</sup> and Pbx1<sup>exc</sup> profiles. These data indicate a strong preference of Prep-Pbx1 dimers for TSSA regions, which is led mainly by Prep since Pbx1 alone does not show any such preference. Common binding of Prep and Meis mostly affected the TSSA and FI classes, appearing at frequencies between those observed for the single factors. Prep-Meis cobinding thus displays mixed properties of the two independent factors and does not generate new binding preferences. The peaks bound by all three factors are predominantly enriched in the TSSA and CI classes in comparison with the whole genome. The

binding preferences of TALE factors in the genome correlate with the global distribution of their occupancy levels (Figure S2).

We next studied the correlations between the identified peaks and known epigenetic marks. Peaks located close to a TSS could coincide with promoters, which are associated with H3K4Me3 and RNAPolII marks (Mikkelsen et al., 2007). Prep peaks are strongly enriched in promoter marks (35% of Prep peaks versus 0.4% in the whole genome), not only in the TSSA category but also within the CI and IG categories and, to a lesser extent, the FI category (Figure 1E). Prep thus appears to have a strong binding preference for promoters and sequences with promoter-like epigenetic marks. This tendency was weaker for Meis peaks, which only correlated significantly with promoter marks in the TSSA and CI peaks, and at a lower proportion than Prep. An intermediate situation was found for Pbx1, which showed a very strong association with promoter marks for the TSSA peaks and a significant association, but weaker than that observed for Prep peaks, in other genomic regions (Figure 1E). Similar analyses of the coincidence of peaks with murine embryonic fibroblast enhancer (H3K4Me1<sup>+</sup>, H3K4Me3<sup>-</sup>) and bivalent enhancer (H3K4Me1<sup>+</sup>, H3K27Ac<sup>+</sup>) marks (Shen et al., 2012) revealed more than a 5-fold enrichment of Meis peaks with both enhancer and bivalent enhancer marks with respect to the whole genome and 2-fold with respect to Prep and Pbx1 peaks (Figure 1F).

Thus, while many Prep and Pbx1 sites are located in promoters, Meis peaks show a preference for enhancers. Interestingly, however, Meis peak sequences are more conserved than those of Prep, Pbx1, and several other developmental transcription factors (Figure 2A). A notable exception is the conservation of HoxC9 binding sites in the embryonic spinal cord (Jung et al., 2010), whose conservation profile is very similar to that of the Meis peaks. The degree of conservation of Meis and HoxC9 peaks is only surpassed by that of p300 peaks in forebrain (Blow et al., 2010).

### Prep and Meis Select Different DNA-Binding Sequences in the Genome, Alone or in Combination with Pbx1

To identify consensus DNA sequences in the identified peaks, we performed an unbiased search using rGADEM software (comparable results were obtained with MEME; data not shown) (Figure S3). For each peak, we searched 300 bp centered on the peak maximum. We obtained two types of motifs: those mapping at a single maximum coinciding with the peak center, which we call core motifs, and those showing a bimodal distribution with maxima symmetrically flanking the peak center or showing a spread distribution, which we call accessory motifs. Within the core motifs, we identified the following known motifs: hexameric sequences resembling or identical to the previously in-vitro-described Meis/Prep consensus (HEXA), octameric sequences similar to Pbx/Hox sites (OCTA), a decameric sequence containing a 5' Pbx1 half-site followed by a Meis/Prep site (DECA), and an extended version of the DECA sequence containing a CCAAT sequence at a fixed distance (DECA<sup>ext</sup>) (Figure S3).

Within the Prep<sup>exc</sup> sites, an unbiased motif search only identified the core motif DECA (Figure 2B). In addition, the DECA motif always appeared in the binding classes in which Prep was present in combination with any other factor/s. In contrast, the

DECA<sup>ext</sup> domain only appeared in the Prep-Pbx1<sup>com</sup> class. In contrast, within the Meis<sup>exc</sup> class, both the HEXA and OCTA motifs were identified but not the DECA motif. HEXA and OCTA motifs also appeared in Meis-Pbx1<sup>com</sup>, while the OCTA motif appeared in all categories in which Meis was present. In the Pbx1<sup>exc</sup> class, a previously undescribed and poorly defined consensus motif was detected. Given the poor definition of this motif, we excluded it from further analysis. The sites identified in the Pbx1 combinations with either Meis or Prep represent, respectively, the binding preferences of Meis or Prep alone, with the previously mentioned exception of DECA<sup>ext</sup> in Prep-Pbx1<sup>com</sup> peaks. The accessory motifs mostly consisted of sequences of low complexity, which nonetheless occurred preferentially in association with specific factors and may enhance binding or allow the binding of cofactors (Figure S3).

We then performed directed searches to determine the abundance of the identified core motifs in the peak sets for each factor and their combinations (Figures 2C and 2D). Overall, 62% of all peaks contained at least one core motif; the HEXA motif was present in about 20% of all peaks, the OCTA in 27%, and the DECA in 29% (Figure 2C). Core sequences were present in 82% of Prep<sup>exc</sup> peaks and 68% of Meis<sup>exc</sup> peaks (Figure 2D).

In Prep<sup>exc</sup> peaks, DECA or DECA<sup>ext</sup> motifs were predominant (48% and 37%, respectively; Figure 2D), while OCTA motifs were not represented over random expectation (4.7% versus 6.4%; *adjp* = 1). In contrast, DECA and especially DECA<sup>ext</sup> motifs were not overrepresented in Meis<sup>exc</sup> peaks (*adjp* = 1), while the HEXA (25%) and the OCTA motifs (42%) were predominant and represented over random expectation (*adjp* < 0.0001; Figure 2D). Pbx1<sup>exc</sup> peaks did not show a strong preference for any of the core motifs, but participation of Pbx1 in the binding increased the presence of the DECA<sup>ext</sup> motif in the Prep profile (74.8% versus 37%; *adjp* < 0.0001) and of the OCTA motif in the Meis profile (55% versus 42%; *adjp* < 0.0001) (Figure 2D).

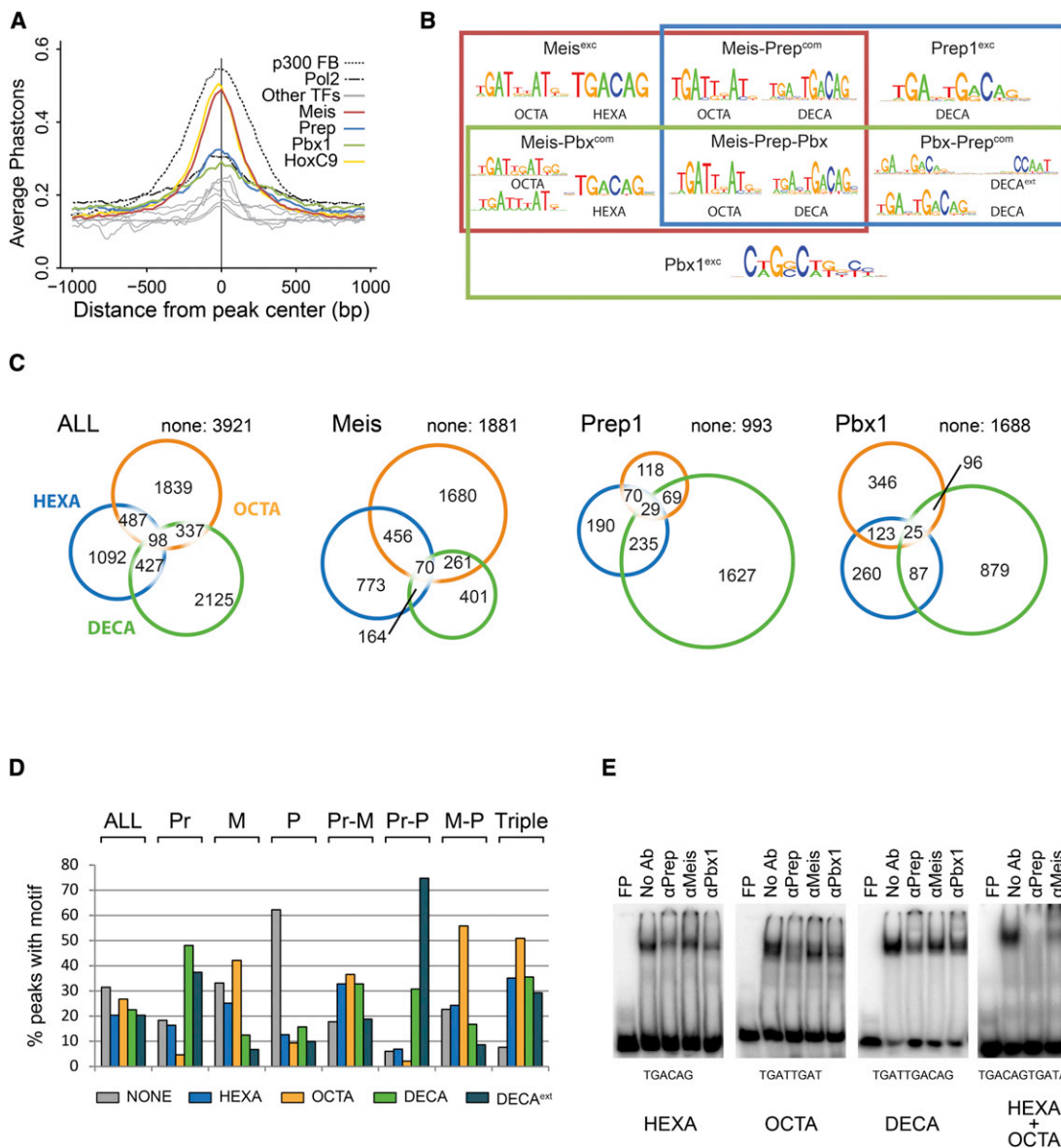
Regarding Meis-Prep<sup>com</sup> peaks, all motifs show an abundance intermediate between that found for each factor independently, with the exception of the HEXA motif, which is more abundant in the common peaks than in the single-factor peaks. The triple-factor peaks have a profile similar to that of the Meis-Prep peaks (chi-square *p* value = 0.37), except for a clear increase in the OCTA sequence (36% versus 50%; *p* = 0.005), again indicating correlation between Pbx1 and the OCTA sequence, provided that Meis is also involved in the binding.

Electrophoretic mobility shift assays (EMSA) of peak sequences from the identified binding motifs showed that while Prep and Pbx can bind any of the core motifs identified, Meis can bind the HEXA and OCTA sequences but can only weakly bind to the DECA sequence (Figure 2E; Extended Results; Figure S4A).

### Meis Binding to the OCTA Motif Corresponds to Pbx-Hox Binding Sites

The abundance of OCTA sites in Meis targets could correspond to a strong association between Meis and Pbx-Hox target sites. In contrast, the low representation of the OCTA motif in the Prep<sup>exc</sup> peaks would then indicate that Prep-Pbx1 mainly selects non-Hox binding sites. Within the OCTA motif, not all base combinations at the variable core of the OCTA motif stimulate





**Figure 2. Meis and Prep Select Different DNA Target Sequences in the Genome**

(A) DNA sequence conservation (vertebrate PhastCons) profile of Meis, Prep, and Pbx1 peaks. For comparison, the plot shows binding sites for HoxC9, HoxA2, p300 forebrain, and other transcription factors (Mahony et al., 2011; Schmidt et al., 2010).

(B) Core sequence motifs identified in exclusive, double, and triple peaks.

(C) co-occurrence of core sequence motifs in each binding class. Boxplots show Pbx1 enrichment factors for peaks cobound by dimers and trimers (Pbx1-Prep, Pbx1-Meis, and Pbx1-Prep-Meis).

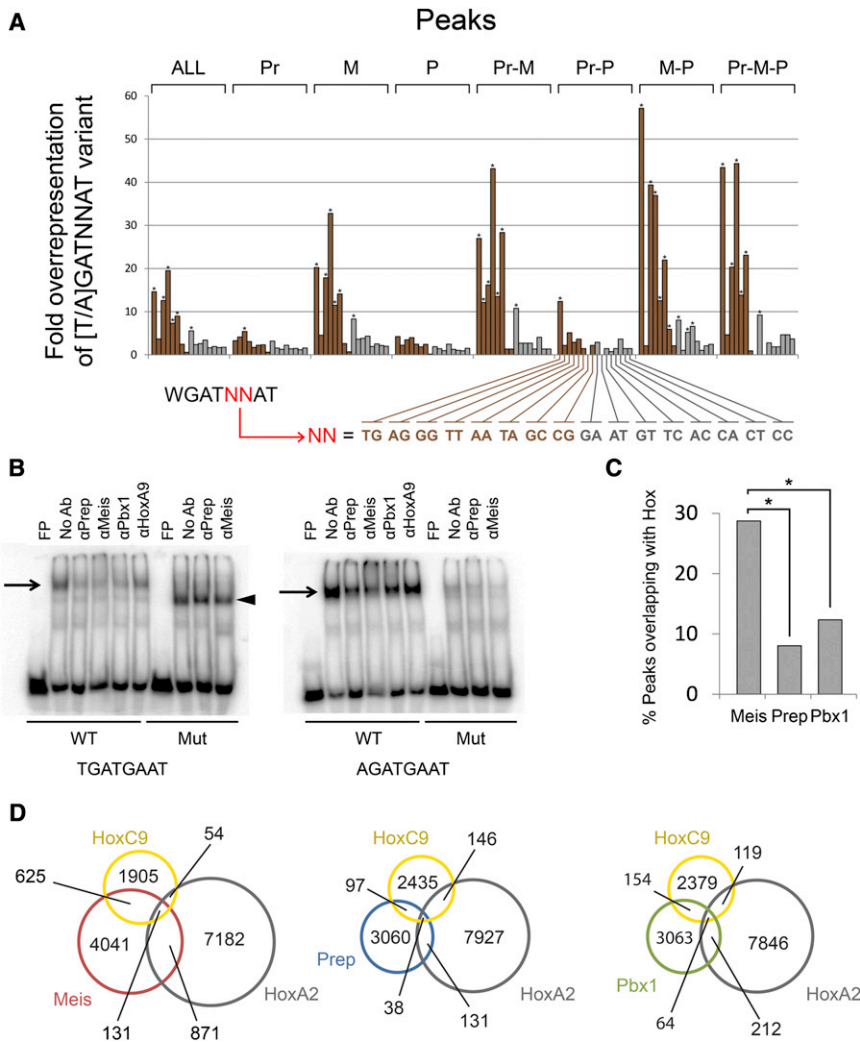
(D) Abundance of core sequence motifs in each factor binding class.

(E) EMSA testing of the in vitro binding ability of the TALE factors. FP, free probe.

See also Figures S3 and S4.

Pbx-Hox dimer binding (Berger et al., 2008; Chan et al., 1994; Chang et al., 1996; Lu and Kamps, 1997; Mann and Chan, 1996; Noyes et al., 2008). We therefore examined the enrichment of each dinucleotide combination at the OCTA variable core (bases 5 and 6) in the peak sets for each factor and their combinations (Figure 3A). Eight two-base combinations have been reported to promote Pbx-Hox binding, while the remaining eight

have not (Slattery et al., 2011; Tümpel et al., 2007). Of the eight that do, five were strongly and significantly overrepresented in all but one of the peak sets (Figure 3A). In contrast, only one of the combinations (GA in the variable core) not previously found to bind Pbx-Hox was overrepresented in various peak sets (Figure 3A). EMSA analyses of sequences from OCTA-containing peaks confirmed various Hox protein binding to the previously



**Figure 3. Hox Binding Motifs and Sites Strongly Correlate with Meis Peaks and Not with Prep Peaks**

(A) Overrepresentation of OCTA variants for each base combination at positions 5 and 6 of the sequence in the peak sets for each factor and their combinations. The two-base combinations that have been previously described to bind Pbx-Hox are shown in brown, while those that have not are shown in gray. Cases in which the screened sequences were found more than five times overrepresented and deviated from random expectation with  $p < 0.001$  are indicated with an asterisk. (B) EMSA testing of in vitro binding ability of the TALE factors and an example Hox protein to the candidate new Hox binding sequence. Mutant probes contain WGATCCAT instead of WGATGAAT. FP, free probe. Arrows indicate the migration of complexes formed between nuclear proteins and DNA. Arrowhead indicates a nonspecific complex. (C) Percentage overlap of  $Prep^{total}$ ,  $Meis^{total}$ , and  $Pbx1^{total}$  peaks with either HoxC9 or HoxA2. Asterisks show  $p < 0.0001$ . (D) Overlap of Prep, Meis, and Pbx1 peaks with HoxC9 and HoxA2 peaks.

analyzed here, show a strong overlap with  $Meis^{total}$  (28% versus 0.42% expected by chance;  $p < 0.0001$ ), a much lower overlap with  $Prep^{total}$ , and moderate overlap with  $Pbx1^{total}$  (Figures 3C and 3D). Common Meis-Pbx1 and Meis-Prep peaks show an increased chance to overlap with Hox (51% of Meis-Pbx1 and 44% of Meis-Prep<sup>com</sup> peaks). Moreover, 68% of Prep-Hox<sup>com</sup> peaks and 79% of Pbx1-Hox<sup>com</sup> peaks are also

Meis peaks (data not shown), again indicating that Meis binding shows the strongest association with Hox binding in these experiments. It is noteworthy, however, that while our analysis is comprehensive for Meis proteins, it is not so for Pbx proteins, so that the lower overlap of Pbx1 with Hox binding sites may be due to the participation of other Pbx family members instead of Pbx1.

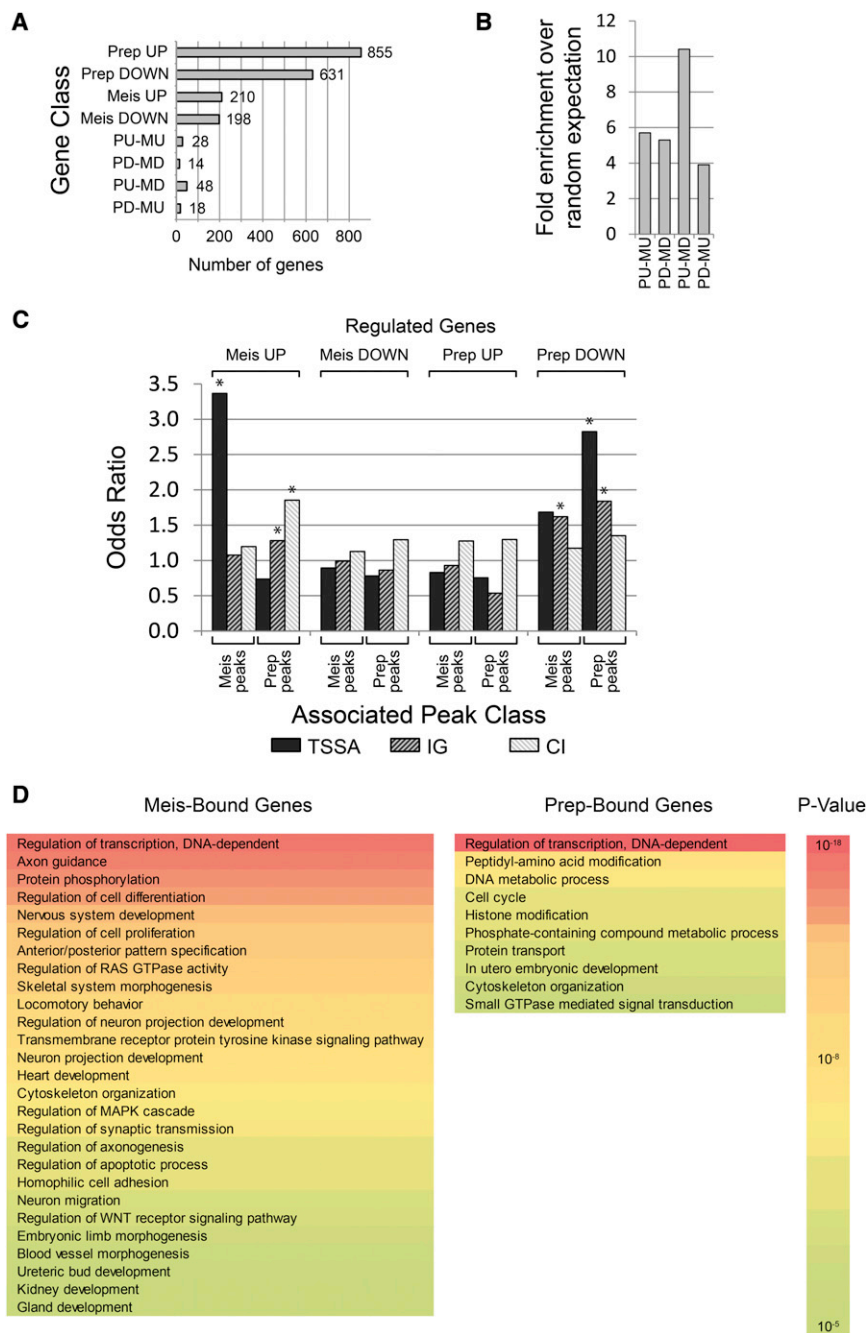
known sequences (Figure S4B) and weak binding of Hoxa9 to the OCTA motif with GA in the variable core (Figure 3B). All peak sets that showed enrichment were Meis-bound, while peak sets in which Meis was not involved showed marginal or no enrichment for Hox-bound base combinations. An exception was the enrichment for the TGATTGAT sequence in the Prep-Pbx1<sup>com</sup> peaks; however, this sequence might be a variant of the DECA sequence. Interestingly, the degree of enrichment in Hox-type sequences increased with cobinding of Pbx1 or Prep with Meis, being maximal in peaks bound by all three factors. These data support the idea that the OCTA sequence represents Pbx-Hox targets and that Meis is the factor most associated with Hox binding sites in the genome. In contrast, Prep does not normally select Hox binding sequences unless the peak is also Meis-bound.

These data suggest that Meis peaks containing an OCTA motif could represent Hox targets. In line with this suggestion, ChIP-seq peaks identified for HoxA2 in E11.5 second branchial arch (Donaldson et al., 2012) and for HoxC9 in E11.5 spinal cord (Jung et al., 2010), despite representing the targets of just 2 of the 39 Hox proteins in embryonic tissues different from those

analyzed here, show a strong overlap with  $Meis^{total}$  (28% versus 0.42% expected by chance;  $p < 0.0001$ ), a much lower overlap with  $Prep^{total}$ , and moderate overlap with  $Pbx1^{total}$  (Figures 3C and 3D). Common Meis-Pbx1 and Meis-Prep peaks show an increased chance to overlap with Hox (51% of Meis-Pbx1 and 44% of Meis-Prep<sup>com</sup> peaks). Moreover, 68% of Prep-Hox<sup>com</sup> peaks and 79% of Pbx1-Hox<sup>com</sup> peaks are also Meis peaks (data not shown), again indicating that Meis binding shows the strongest association with Hox binding in these experiments. It is noteworthy, however, that while our analysis is comprehensive for Meis proteins, it is not so for Pbx proteins, so that the lower overlap of Pbx1 with Hox binding sites may be due to the participation of other Pbx family members instead of Pbx1.

### Prep1 and Meis1 Coordinately Regulate a Subset of Their Target Genes

To investigate functional interactions between Prep1 and Meis1, we compared changes to the transcriptome caused by elimination of either *Meis1* or *Prep1* in mouse embryos. To this end, we performed total RNA sequencing (RNA-seq) in *Meis1*-deficient and *Prep1*<sup>fl/fl</sup> E11.5 embryos. RNA-seq identified 855 upregulated and 631 downregulated transcripts in *Prep1*<sup>fl/fl</sup> embryos and 210 upregulated and 198 downregulated transcripts in *Meis1*-deficient embryos (Figure 4A; Table S4). The affected transcripts in *Meis1* mutants probably represent only a fraction of all *Meis*-regulated genes, since the expression patterns of *Meis1* and *Meis2* overlap considerably in the embryo. To estimate the



**Figure 4. Meis and Prep Target Gene Core-regulation and Functional Annotation**

(A) Numbers of genes up- or downregulated in *Prep1<sup>iii</sup>* mutant and *Meis1*-deficient embryos showing core-regulation. Prep UP (PU) are genes upregulated and Prep DOWN (PD) are genes downregulated in *Prep1<sup>iii</sup>* embryos; Meis UP (MU) are genes upregulated and Meis DOWN (MD) are genes downregulated in *Meis1* loss-of-function embryos.

(B) Extent of core-regulation. The most over-represented set of genes is that composed of genes upregulated in *Prep1<sup>iii</sup>* and downregulated in *Meis1*-deficient mutants. Asterisks show chi-square  $p < 0.0001$ .

(C) Association of regulated genes with peak classes and their genomic location. Graph shows fold enrichment in peak density over whole-genome average (i.e., odds ratio). Asterisk shows  $p < 0.001$ .

(D) Selected Gene Ontology terms associated with bound genes.

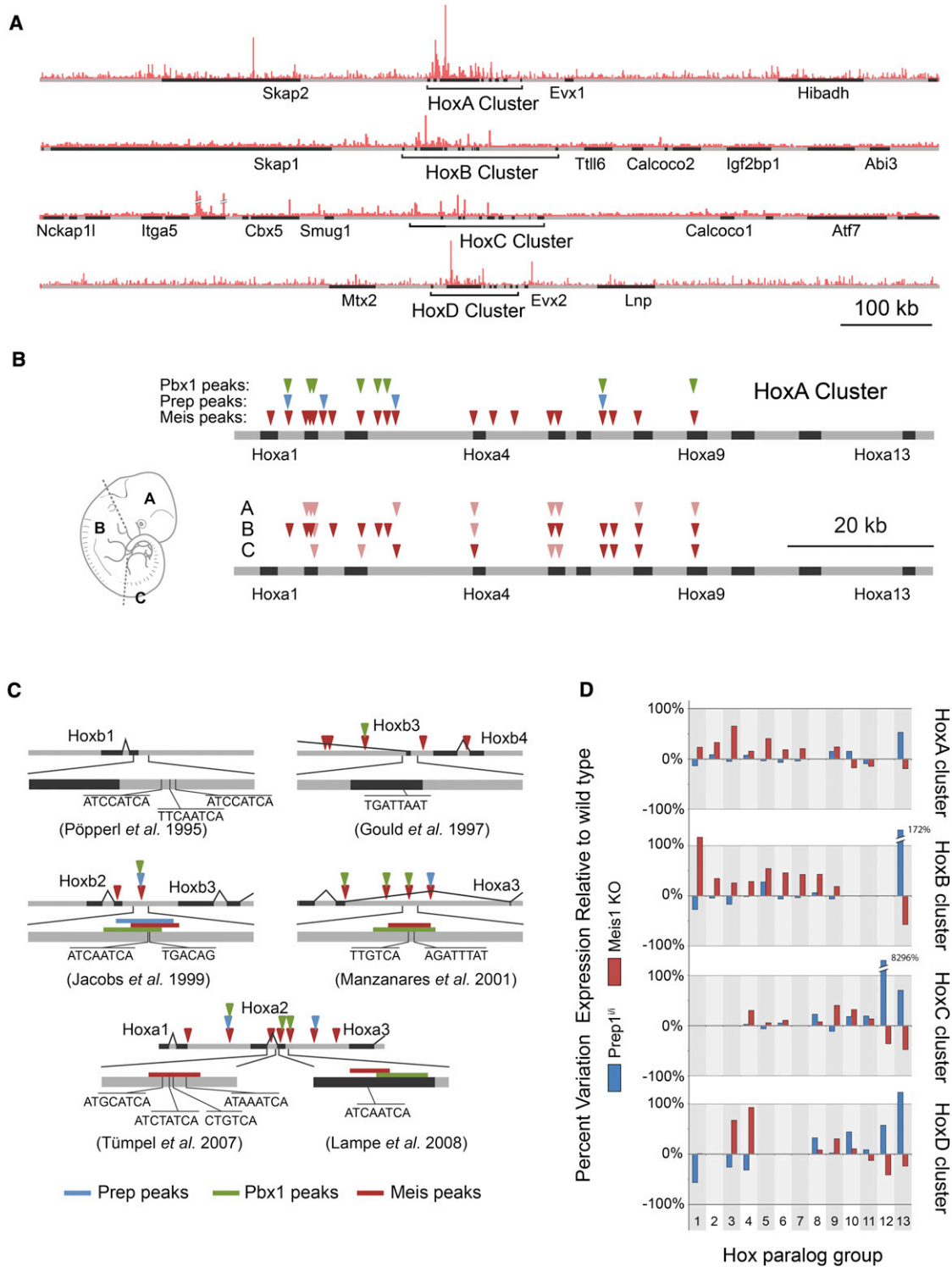
See also Tables S4 and S9.

*Meis1*-deficient and upregulated in the *Prep1*-deficient embryos. This class included 48 genes, or 24% of the genes downregulated in *Meis1*-deficient embryos, and was 10-fold higher than random expectation. These results indicate considerable functional interactions between *Meis1* and *Prep1* in the regulation of gene expression, including cooperative and, more frequently, antagonistic actions.

To identify putative direct transcriptional targets of TALE factors, we examined the correlation between ChIP-seq peaks and the set of genes regulated by *Prep1* and *Meis1* (Figure 4C). Among the genes upregulated in *Meis1*-deficient embryos, we found significant enrichment in TSSA Meis peaks but not in those in other gene regions, suggesting a correlation between TSSA Meis binding and negative regulation of transcription (Figure 4C). Surprisingly, within the genes upregulated in *Meis1*-deficient embryos, there was a significant enrichment in Prep IG

and CI peaks and, conversely, we found some enrichment of IG Meis peaks among the genes upregulated in *Prep1<sup>iii</sup>* embryos, again suggesting a functional interaction between Meis and Prep in gene regulation (Figure 4C). In contrast, we found no enrichment for any factor peaks among genes downregulated in *Meis1*-deficient embryos or upregulated in *Prep1<sup>iii</sup>* embryos. Finally, TSSA and IG Prep peaks were overrepresented among *Prep1<sup>iii</sup>*-downregulated genes and underrepresented among *Prep1<sup>iii</sup>*-upregulated genes, indicating a transcriptional activator function for these Prep sites.

extent of gene core-regulation by *Meis1* and *Prep1*, we determined the frequency of core-regulated genes and the nature of the core-regulation. All classes of core-regulated genes occurred at frequencies 4- to 10-fold higher than expected under the null hypothesis of independence of gene subsets, indicating coordinated actions of these transcription factors in the regulation of specific sets of genes (Figure 4B). We found 108 genes core-regulated by the two factors, corresponding to 26% of *Meis1*-regulated genes and 7% of *Prep1*-regulated genes. Interestingly, the most enriched class was genes downregulated in



**Figure 5. TALE-Factor Binding Subdivides the Hox Clusters in Two Regulatory Domains**

(A) Meis ChIP-seq read profiles (in red) in the Hox cluster environment. Black bars show coding regions.

(B) Top: The HoxA cluster is shown, with Meis, Prep, and Pbx1 ChIP-seq peaks represented in different colors as indicated. Bottom: A similar representation of the HoxA cluster shows Meis ChIP-PCR signals obtained from different embryo portions, as indicated on the scheme on the left. An absent triangle indicates no binding detected, a light-colored triangle indicates positive but not predominant binding compared to other embryo regions, and a dark-colored triangle indicates predominant binding.

(legend continued on next page)



These results indicate an association of Meis TSSA binding with the repression of Meis1 target genes and an association of Prep TSSA binding with the activation of Prep1 target genes. The reported context-dependent repressive activity of Meis/Hth proteins is in agreement with these findings (Elkouby et al., 2012; Huang et al., 2005). In addition, the association of Meis peaks with Prep1-regulated genes and of Prep peaks with Meis1-regulated genes suggests coordinated actions of Meis and Prep on some of their targets.

We next profiled the Gene Ontology annotations of potential Meis and Prep targets (Figure 4D), considering the set of genes with Meis or Prep peaks in their promoters or transcriptional units as potential direct targets. For both factors, target genes encoding transcriptional regulators are strongly overrepresented. Meis-bound genes are strongly enriched for functions involved in several aspects of development, such as AP pattern specification, heart development, nervous system development, and blood vessel morphogenesis. Meis also binds to genes involved in cell processes that potentially mediate its leukemogenic properties, such as cell differentiation and proliferation. In contrast, developmentally associated genes are only weakly overrepresented among Prep targets, which are instead annotated to basal cell functions like DNA and histone modification, protein transport, and signal transduction. These data correlate with the fact that Meis proteins are expressed in a developmentally restricted manner, while Prep is a ubiquitously expressed protein that regulates essential cell functions (Fernandez-Diaz et al., 2010; Iotti et al., 2011).

### The TALE-Factor Binding Landscape Subdivides Hox Clusters into Two Regions with Differential Transcriptional Responses

Examination of the binding sites of Prep, Meis, and Pbx1 in the Hox clusters reveals abundant interaction sites (Figures 5A and 5B; Figure S5A), suggestive of extensive crosstalk and autoregulation within the Hox/TALE network. In the Hox clusters, Meis peaks are the most abundant, occurring mostly in the HoxA and least in the HoxC cluster. Pbx1 binding sites are less abundant, and of the three factors, Prep binding sites are the least abundant. In all cases but one, Pbx1 and Prep sites coincide with Meis peaks. Interestingly, all peaks concentrate in paralog groups 1–9, with no peak present in paralogs 10–13 in any Hox cluster, indicating a subdivision of the Hox clusters into TALE-interactive and TALE-noninteractive regions (Figure 5A; Figure S5A). ChIP-PCR analysis of Meis binding sites in the HoxA cluster indicated that the binding profile was variable in different regions of the embryo and correlated with the expression status of the HoxA cluster in these regions (Figure 5B).

Previous studies had identified six TALE protein binding regions in the Hox clusters that were mostly involved in cooperation with Hox proteins in auto- and cross-regulatory interactions

(Gould et al., 1997; Jacobs et al., 1999; Lampe et al., 2008; Manzanares et al., 2001; Pöpperl et al., 1995; Tümpel et al., 2007). Although those interactions were described at a different developmental stage and only affect a subset of the tissues analyzed here, we found interactions at the precise sites previously described in four out of the six regions (Figure 5C).

We then compared Hox gene expression in Meis1-deficient and Prep1<sup>iii</sup> mutant E11.5 embryos with that in wild-type littermates by RNA-seq. In Meis1-deficient embryos, 22 of the 27 Hox genes from paralog groups 1–9 increased their expression, while expression of the remaining five decreased or was maintained (Figure 5D). In contrast, seven of the eight Hox genes from paralog groups 11–13 decreased their expression and expression of the other was maintained (Figure 5D). Paralog 10 genes showed variable behavior, with Hoxa10 expression being reduced, Hoxc10 increased, and Hoxd10 maintained. Although many of the expression changes are moderate and would not be significant in isolation, the correlation of the expression changes with the position of the genes in the cluster significantly ( $p < 0.05$ ) diverges from the transcriptomic average for all clusters except the 3' part of cluster C (see Experimental Procedures). These results suggest that Meis function moderates the expression of paralog groups 1–9 while enhancing expression of paralog groups 11–13. Paralog group Hox10 seems to be placed in a frontier region, with the influence of Meis activity depending on the specific Hox cluster.

While expression of many Hox cluster genes does not change in Prep1<sup>iii</sup> E11.5 mutants, the 5' genes show changes opposite to those observed in Meis1-deficient embryos (Figure 5D). An opposite regulation to that observed in Meis1-deficient embryos was also observed in all cases for paralog group 1, extending to paralogs 1–4 in the case of the HoxD cluster (Figure 5D).

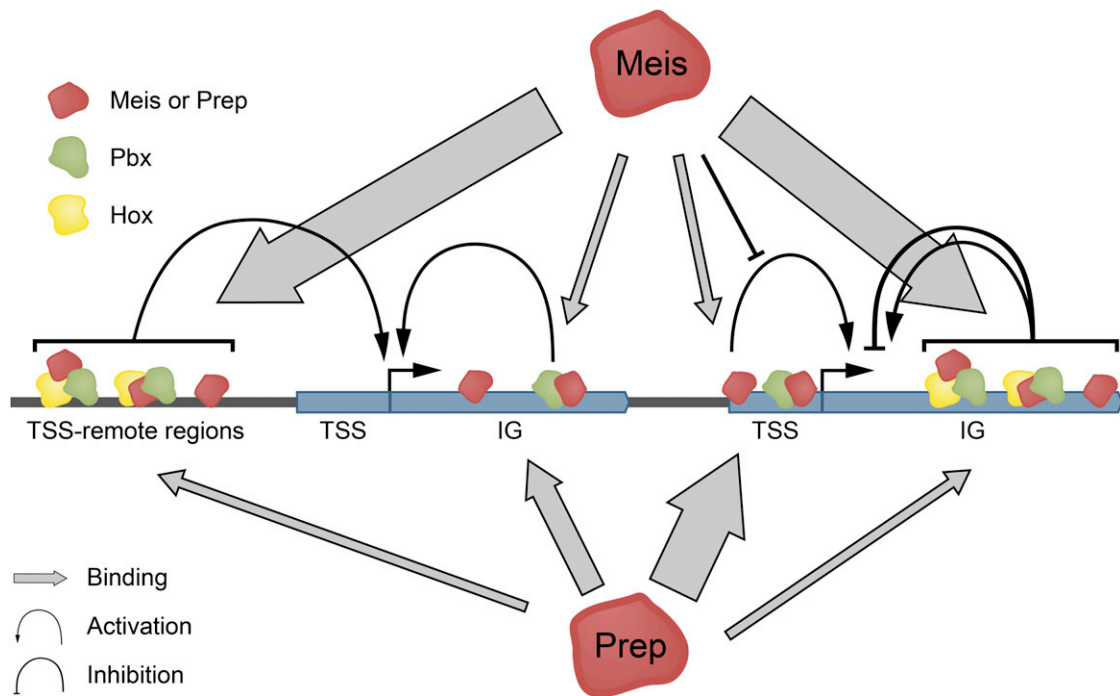
These results show interactions between Prep and Meis in the global modulation of Hox gene expression. From the six previously described regulatory interactions, four are detected in our study, suggesting that the observed regulatory effects involve direct interactions linked to the described binding sites. This view is further supported by the correlation between the Meis binding profile and HoxA cluster expression and by the coincidence between the Meis/Prep/Pbx binding profiles and the transcriptional response landscape in the Hox clusters.

Additional genetic interaction studies showed no interaction between Prep1 and Meis1 loss-of-function alleles (Extended Results; Table S5; Figure S6), suggesting that the critically affected functions in these mutants are independent. This is consistent with the predominantly independent DNA-binding activities observed for each factor. In contrast, the strong genetic interaction between Meis1 and Pbx1 (Table S5) correlates with the predominance of Pbx-Hox binding sites within the Meis ChIP peaks.

(C) Representation of previously described TALE factor binding sites in the Hox clusters. Black bars indicate exons, and gray bars indicate introns or intergenic regions. For each case, a representation of the Hox cluster subregion with the ChIP-seq peaks is shown above and a zoom showing the specific sequences previously described and their position with respect to the peaks here described (color stripes) is shown below.

(D) Differential transcriptional response of the 3' and 5' halves of the Hox clusters to Meis1 and Prep1 deficiency. Graphs show the change in transcript levels (percentage) between controls and mutant Meis1 ko and Prep1<sup>iii</sup> E11.5 embryos.

See also Figure S5.



**Figure 6. A Representation of PREP and MEIS Homeodomain Factor Activity in Regulating Gene Expression**

Meis and Prep proteins cooperate to regulate gene expression. Prep binds mostly to promoters in conjunction with Pbx. Meis binds mainly to non-TSSA regions in cooperation with Hox proteins, often without contacting DNA. They often show opposing activities.

## DISCUSSION

In this study, we have analyzed genomic binding sites for Pbx1, Meis1/2, and Prep1/2. While Pbx3/4 and Meis3 were not studied, the similarities of these proteins to the members of the family studied here suggest that their binding repertoires will be comparable. In addition, the expression patterns of the studied proteins cover the majority of tissues in which their counterparts are expressed. This analysis is thus a near-comprehensive picture of the general binding abilities of TALE factors in the mammalian embryo. The results highlight the specialization of Prep and Meis in binding largely independent genomic elements through selection of different DNA sequences. The contrast with the previously reported *in vitro* activities suggest that Meis and Prep gain additional binding specificity *in vivo* through interaction with cofactors or chromatin landmarks. While Prep1 interacts preferentially with promoters and nearby regions, Meis shows preference for intergenic and intragenic regions away from TSSA regions. Prep could thus directly control promoter activity, while a substantial part of the Meis sites coincides with enhancers. A number of Meis binding sites remain functionally undefined and, despite their high evolutionary conservation, do not correlate with the described marks of known constitutive chromatin factors such as CTCF and others (not shown). These results suggest that a proportion of Meis sites are evolutionarily conserved protein-DNA interaction regions whose function remains to be explored. In addition, Prep mostly participates in dimer formation with Pbx, while Meis is predominant in Pbx-Hox interactions on targets.

These findings suggest that during evolution, Meis and Prep proteins acquired specialized functions enabling them to interact with specific subsets of regulatory regions and target genes (Figure 6). Functions of the MEINOX and PBC proteins extend beyond the regulation of Hox protein activity. The complete loss of function of *exd* or *hth* in *Drosophila* results in the early failure of embryonic development due to defects in cell division at stages when Hox genes are not required (Rauskolb et al., 1993; Salvany et al., 2009). The ontology analysis of the targets bound by Meis and Prep, together with the functions previously described for *Meis1* and *Prep1* in mice, suggest that Prep has specialized in the basic cellular functions, which would be Hox independent, and Meis has specialized in patterning functions more related to Hox activity. Whether the sum of the functions of Meis and Prep corresponds to those exerted by the single MEINOX in flies (Hth) or, alternatively, involves the acquisition by either factor of new functions related to vertebrate evolution remains to be explored.

Regarding the interaction with Pbx, we found that Prep binds DNA preferentially as a dimer with any of the Pbx proteins. The strong overlap between thymic Pbx2 sites and embryonic Pbx1 sites also shows that the two proteins can substitute each other, and hence provides a molecular basis for the concept of Pbx redundancy (Selleri et al., 2004).

The data presented point to Meis factors as major *in vivo* partners of Hox proteins, cooperating with them in target selection with little contribution from Prep. Identifying the genomic binding sites of the 39 mammalian Hox proteins is a major challenge that is still far from being achieved. Given the general requirement of

Hox proteins for cooperation with TALE factors, and the fact that TALE proteins can interact promiscuously with Hox proteins, the putative binding sites presented in this study likely represent the most comprehensive set of in vivo Hox genomic targets yet identified.

Despite the extensive divergence in their genomic binding patterns, *Meis1* and *Prep1* do show coregulation of some downstream genes, with opposing effects predominating. Some of these antagonistic interactions might underlie the opposing roles of *Meis1* and *Prep1* in tumor formation, where *Meis1* function promotes tumor formation while *Prep1* behaves as a tumor suppressor (Iotti et al., 2011; Longobardi et al., 2010; Thorsteinsdottir et al., 2001).

A striking case of coordinated regulation was observed in *Hox* gene regulation, where the TALE protein binding profile and transcriptional regulatory activity subdivide *Hox* clusters into two regions: the paralog 1–9 region and the paralog 10–13 region. These results highlight the important role of TALE factors in globally regulating *Hox* gene expression, in addition to serving as cofactors of Hox proteins. The modulation of *Hox* cluster transcriptional activity may be the result of global conformational changes promoted by TALE factors, since paralog groups 10–13 are not directly bound by TALE factors, yet they are sensitive to their levels.

Our work thus identifies TALE and TALE-Hox binding sites, target genes, and in vivo specificities that increase our understanding of the molecular pathways controlled by this regulatory network in development and disease.

## EXPERIMENTAL PROCEDURES

### ChIP-Seq and ChIP-Re-ChIP

ChIPs were performed using standard methods on E11.5 mice embryo trunks. We used anti-Prep1/2 antibody, anti-Pbx1 antibody, and a mix of anti-Meis antibodies. The same antibodies were used for ChIP-re-ChIP.

### ChIP-Chip

ChIPs of mouse thymocyte lysates were performed as described above with anti-Prep1/2 antibody and anti-Pbx2 antibody on thymuses from 6- to 8-week-old C57B6 mice. The resulting DNA was hybridized to a Nimblegen mouse RefSeq promoter array.

### ChIP-Seq Data Analysis

ChIP DNA was sequenced using an Illumina GAI analyzer. Single-end 36 bp reads were mapped with BWA software against mm9 version of the mouse genome. The alignments were then used for peak calling, de novo motif discovery, and motif identification and validation with MotIV.

### ChIP-PCR

For comparison of AP occupancy of Meis peaks, E11.5 embryos were dissected as shown in the diagram in Figure 5B and ChIP was carried out as described above. DNA was then subjected to PCR for 30 or 35 cycles, depending on primers, to avoid saturation.

### RNA-Seq Data Analysis

Total RNA was purified from whole E11.5 mouse embryos, and a library was prepared and sequenced on the Illumina platform according to the manufacturer's instructions. Approximately 7M reads per sample were aligned to mouse mm9, and transcript expression was estimated with mouse Ensembl 63 genebuild as a reference.

### Exploratory Data Analysis

Peak overlapping, correlation with RNA-seq data, and conservation data aggregation were performed on the Galaxy platform.

Individual instances of the core motifs within all peaks were searched with a local install of the FIMO (Find Individual Motif Occurrences) program from the MEME suite.

Peak profiling was performed using custom Python scripts and the CEAS (Cis-regulatory Element Annotation System) tool.

### EMSA

Nuclear extracts were isolated from cells prepared from E11.5 mouse embryonic body. EMSA reactions were performed following the standard protocol.

### Gene Ontology Analysis

GO term overrepresentation was assessed with GOrilla, comparing the lists of genes with Meis or Prep binding sites against the list of all nuclear genes in Ensembl v63, with a p value cutoff of  $10^{-5}$ . We considered those genes that have a Meis or Prep peak in their promoter (–500 to +100) or within the transcriptional unit.

### Animal Procedures

All animal procedures have been reviewed and approved by the CNIC Animal Experimentation Ethics Committee, according to the National and European regulations.

For further details, see [Extended Experimental Procedures](#).

### ACCESSION NUMBERS

The GEO accession number for the ChIP-seq results reported in this paper is GSE39609.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Results, Extended Experimental Procedures, six figures, and nine tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2013.03.029>.

### LICENSING INFORMATION

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ACKNOWLEDGMENTS

D.P., F.B., and M.T. are grateful to D. Pasini, G. Natoli, A. Roselló, and members of the Blasi and Torres labs for their helpful discussions. We thank L. Modica for Prep1<sup>fl</sup> RNA preparation, J. Wilde and M. Linser for assistance in ChIP-seq library preparation, A. Dopazo for sequencing, V. Amstislavskiy for sequencing data processing, L. Selleri for providing the Pbx1 KO mice, and Jeremy Dasen for antibodies. The collaboration among F.B., M.T., and D.P. was made possible by COST Action BM0805. F.B. was supported by grants from the AIRC (Associazione Italiana Ricerche sul Cancro, 8929), Ministero dell'Università e Ricerca (MERIT; EU FP7 Prepobedia, MIUR-FIRB RBNE08NKH7), the Cariplo Foundation, and the Italian Ministry of Health. M.T. was supported by grants RD06/0010/0008 and BFU2009-08331/BMC from the Spanish Ministerio de Economía y Competitividad (MINECO). D.M.S.M. was supported by a fellowship from the Consejería de Educación de la Comunidad de Madrid and the European Social Fund. D.P. and V.T. were supported by grant 02.740.11.0872 from the Russian Ministry of Education and Science. D.P. was supported by grant 12-04-01659-a from the Russian Foundation for Basic Research. A.B. was supported by Associazione Italiana Ricerche sul Cancro start-up grant #4780. The CNIC is supported by the MINECO and the pro-CNIC Foundation. IFOM is supported by the FIRC

(Fondazione Italiana Ricerche sul Cancro). S. Bartlett (CNIC) provided English editing.

Received: July 18, 2012

Revised: February 19, 2013

Accepted: March 20, 2013

Published: April 18, 2013

## REFERENCES

- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106.
- Azcoitia, V., Aracil, M., Martínez-A, C., and Torres, M. (2005). The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo. *Dev. Biol.* **280**, 307–320.
- Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Peña-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**, 1266–1276.
- Berthelsen, J., Zappavigna, V., Mavilio, F., and Blasi, F. (1998). Prep1, a novel functional partner of Pbx proteins. *EMBO J.* **17**, 1423–1433.
- Bisaillon, R., Wilhelm, B.T., Kros, J., and Sauvageau, G. (2011). C-terminal domain of MEIS1 converts PKNOX1 (PREP1) into a HOXA9-collaborating oncoprotein. *Blood* **118**, 4682–4689.
- Blow, M.J., McCulley, D.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2010). ChIP-Seq identification of weakly conserved heart enhancers. *Nat. Genet.* **42**, 806–810.
- Cecconi, F., Proetzl, G., Alvarez-Bolado, G., Jay, D., and Gruss, P. (1997). Expression of Meis2, a Knotted-related murine homeobox gene, indicates a role in the differentiation of the forebrain and the somitic mesoderm. *Dev. Dyn.* **210**, 184–190.
- Chan, S.K., Jaffe, L., Capovilla, M., Botas, J., and Mann, R.S. (1994). The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein. *Cell* **78**, 603–615.
- Chang, C.P., Brocchieri, L., Shen, W.F., Largman, C., and Cleary, M.L. (1996). Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Mol. Cell. Biol.* **16**, 1734–1745.
- Chang, C.P., Jacobs, Y., Nakamura, T., Jenkins, N.A., Copeland, N.G., and Cleary, M.L. (1997). Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol. Cell. Biol.* **17**, 5679–5687.
- Di Rosa, P., Villaescusa, J.C., Longobardi, E., Iotti, G., Ferretti, E., Diaz, V.M., Miccio, A., Ferrari, G., and Blasi, F. (2007). The homeodomain transcription factor Prep1 (pKnox1) is required for hematopoietic stem and progenitor cell activity. *Dev. Biol.* **311**, 324–334.
- DiMartino, J.F., Selleri, L., Traver, D., Firpo, M.T., Rhee, J., Warnke, R., O'Gorman, S., Weissman, I.L., and Cleary, M.L. (2001). The Hox cofactor and proto-oncogene Pbx1 is required for maintenance of definitive hematopoiesis in the fetal liver. *Blood* **98**, 618–626.
- Donaldson, I.J., Amin, S., Hensman, J.J., Kutejova, E., Rattray, M., Lawrence, N., Hayes, A., Ward, C.M., and Bobola, N. (2012). Genome-wide occupancy links Hoxa2 to Wnt- $\beta$ -catenin signaling in mouse embryonic development. *Nucleic Acids Res.* **40**, 3990–4001.
- Elkouby, Y.M., Polevoy, H., Gutkovich, Y.E., Michaelov, A., and Frank, D. (2012). A hindbrain-repressive Wnt3a/Meis3/Tsh1 circuit promotes neuronal differentiation and coordinates tissue maturation. *Development* **139**, 1487–1497.
- Fernandez-Diaz, L.C., Laurent, A., Girasoli, S., Turco, M., Longobardi, E., Iotti, G., Jenkins, N.A., Fiorenza, M.T., Copeland, N.G., and Blasi, F. (2010). The absence of Prep1 causes p53-dependent apoptosis of mouse pluripotent epiblast cells. *Development* **137**, 3393–3403.
- Ferretti, E., Schulz, H., Talarico, D., Blasi, F., and Berthelsen, J. (1999). The PBX-regulating protein PREP1 is present in different PBX-complexed forms in mouse. *Mech. Dev.* **83**, 53–64.
- Ferretti, E., Marshall, H., Pöpperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. *Development* **127**, 155–166.
- Ferretti, E., Cambroneo, F., Tümpel, S., Longobardi, E., Wiedemann, L.M., Blasi, F., and Krumlauf, R. (2005). Hoxb1 enhancer and control of rhombomere 4 expression: complex interplay between PREP1-PBX1-HOXB1 binding sites. *Mol. Cell. Biol.* **25**, 8541–8552.
- Ferretti, E., Villaescusa, J.C., Di Rosa, P., Fernandez-Diaz, L.C., Longobardi, E., Mazzieri, R., Miccio, A., Micali, N., Selleri, L., Ferrari, G., and Blasi, F. (2006). Hypomorphic mutation of the TALE gene Prep1 (pKnox1) causes a major reduction of Pbx and Meis proteins and a pleiotropic embryonic phenotype. *Mol. Cell. Biol.* **26**, 5650–5662.
- Gould, A., Morrison, A., Sproat, G., White, R.A., and Krumlauf, R. (1997). Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev.* **11**, 900–913.
- Hisa, T., Spence, S.E., Rachel, R.A., Fujita, M., Nakamura, T., Ward, J.M., Devor-Henneman, D.E., Saiki, Y., Kutsuna, H., Tessarollo, L., et al. (2004). Hematopoietic, angiogenic and eye defects in Meis1 mutant animals. *EMBO J.* **23**, 450–459.
- Huang, H., Rastegar, M., Bodner, C., Goh, S.L., Rambaldi, I., and Featherstone, M. (2005). MEIS C termini harbor transcriptional activation domains that respond to cell signaling. *J. Biol. Chem.* **280**, 10119–10127.
- Iotti, G., Longobardi, E., Masella, S., Dardaei, L., De Santis, F., Micali, N., and Blasi, F. (2011). Homeodomain transcription factor and tumor suppressor Prep1 is required to maintain genomic stability. *Proc. Natl. Acad. Sci. USA* **108**, E314–E322.
- Jacobs, Y., Schnabel, C.A., and Cleary, M.L. (1999). Trimeric association of Hox and TALE homeodomain proteins mediates Hoxb2 hindbrain enhancer activity. *Mol. Cell. Biol.* **19**, 5134–5142.
- Jung, H., Lacombe, J., Mazzoni, E.O., Liem, K.F., Jr., Grinstein, J., Mahony, S., Mukhopadhyay, D., Gifford, D.K., Young, R.A., Anderson, K.V., et al. (2010). Global control of motor neuron topography mediated by the repressive actions of a single hox gene. *Neuron* **67**, 781–796.
- Knoepfler, P.S., Calvo, K.R., Chen, H., Antonarakis, S.E., and Kamps, M.P. (1997). Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. *Proc. Natl. Acad. Sci. USA* **94**, 14553–14558.
- Lampe, X., Samad, O.A., Guiguen, A., Matis, C., Remacle, S., Picard, J.J., Rijli, F.M., and Rezsóhazy, R. (2008). An ultraconserved Hox-Pbx responsive element resides in the coding sequence of Hoxa2 and is active in rhombomere 4. *Nucleic Acids Res.* **36**, 3214–3225.
- LaRonde-LeBlanc, N.A., and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev.* **17**, 2060–2072.
- LeBrun, D.P., and Cleary, M.L. (1994). Fusion with E2A alters the transcriptional properties of the homeodomain protein PBX1 in t(1;19) leukemias. *Oncogene* **9**, 1641–1647.
- Longobardi, E., Iotti, G., Di Rosa, P., Mejetta, S., Bianchi, F., Fernandez-Diaz, L.C., Micali, N., Nuciforo, P., Lenti, E., Ponzoni, M., et al. (2010). Prep1 (pKnox1)-deficiency leads to spontaneous tumor development in mice and accelerates EmuMyc lymphomagenesis: a tumor suppressor role for Prep1. *Mol. Oncol.* **4**, 126–134.
- Lu, Q., and Kamps, M.P. (1997). Heterodimerization of Hox proteins with Pbx1 and oncoprotein E2a-Pbx1 generates unique DNA-binding specificities at nucleotides predicted to contact the N-terminal arm of the Hox homeodomain—demonstration of Hox-dependent targeting of E2a-Pbx1 in vivo. *Oncogene* **14**, 75–83.
- Mahony, S., Mazzoni, E.O., McCuine, S., Young, R.A., Wichterle, H., and Gifford, D.K. (2011). Ligand-dependent dynamics of retinoic acid receptor binding during early neurogenesis. *Genome Biol.* **12**, R2.



- Mann, R.S., and Chan, S.K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet.* *12*, 258–262.
- Mann, R.S., and Affolter, M. (1998). Hox proteins meet more partners. *Curr. Opin. Genet. Dev.* *8*, 423–429.
- Manzanares, M., Bel-Vialar, S., Ariza-McNaughton, L., Ferretti, E., Marshall, H., Maconochie, M.M., Blasi, F., and Krumlauf, R. (2001). Independent regulation of initiation and maintenance phases of *Hoxa3* expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. *Development* *128*, 3595–3607.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* *448*, 553–560.
- Moens, C.B., and Selleri, L. (2006). Hox cofactors in vertebrate development. *Dev. Biol.* *297*, 193–206.
- Moskow, J.J., Bullrich, F., Huebner, K., Daar, I.O., and Buchberg, A.M. (1995). Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol. Cell. Biol.* *15*, 5434–5443.
- Nakamura, T., Jenkins, N.A., and Copeland, N.G. (1996). Identification of a new family of Pbx-related homeobox genes. *Oncogene* *13*, 2235–2242.
- Noyes, M.B., Christensen, R.G., Wakabayashi, A., Stormo, G.D., Brodsky, M.H., and Wolfe, S.A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* *133*, 1277–1289.
- Oulad-Abdelghani, M., Chazaud, C., Bouillet, P., Sapin, V., Chambon, P., and Dollé, P. (1997). Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. *Dev. Dyn.* *210*, 173–183.
- Piper, D.E., Batchelor, A.H., Chang, C.P., Cleary, M.L., and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* *96*, 587–597.
- Pöpperl, H., Bienz, M., Studer, M., Chan, S.K., Aparicio, S., Brenner, S., Mann, R.S., and Krumlauf, R. (1995). Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/pbx*. *Cell* *81*, 1031–1042.
- Rauskolb, C., Peifer, M., and Wieschaus, E. (1993). *extradenticle*, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell* *74*, 1101–1112.
- Rieckhof, G.E., Casares, F., Ryoo, H.D., Abu-Shaar, M., and Mann, R.S. (1997). Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* *91*, 171–183.
- Ryoo, H.D., Marty, T., Casares, F., Affolter, M., and Mann, R.S. (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development* *126*, 5137–5148.
- Salvany, L., Aldaz, S., Corsetti, E., and Azpiazu, N. (2009). A new role for hth in the early pre-blastodermic divisions in *Drosophila*. *Cell Cycle* *8*, 2748–2755.
- Schmidt, D., Wilson, M.D., Ballester, B., Schwalie, P.C., Brown, G.D., Marshall, A., Kutter, C., Watt, S., Martinez-Jimenez, C.P., Mackay, S., et al. (2010). Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* *328*, 1036–1040.
- Selleri, L., Depew, M.J., Jacobs, Y., Chanda, S.K., Tsang, K.Y., Cheah, K.S., Rubenstein, J.L., O’Gorman, S., and Cleary, M.L. (2001). Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. *Development* *128*, 3543–3557.
- Selleri, L., DiMartino, J., van Deursen, J., Brendolan, A., Sanyal, M., Boon, E., Capellini, T., Smith, K.S., Rhee, J., Pöpperl, H., et al. (2004). The TALE homeodomain protein Pbx2 is not essential for development and long-term survival. *Mol. Cell. Biol.* *24*, 5324–5331.
- Shanmugam, K., Green, N.C., Rambaldi, I., Saragovi, H.U., and Featherstone, M.S. (1999). PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. *Mol. Cell. Biol.* *19*, 7577–7588.
- Shen, W.F., Montgomery, J.C., Rozenfeld, S., Moskow, J.J., Lawrence, H.J., Buchberg, A.M., and Largman, C. (1997). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol. Cell. Biol.* *17*, 6448–6458.
- Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenko, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* *488*, 116–120.
- Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., and Mann, R.S. (2011). Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. *Cell* *147*, 1270–1282.
- Thorsteinsdottir, U., Kroon, E., Jerome, L., Blasi, F., and Sauvageau, G. (2001). Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Mol. Cell. Biol.* *21*, 224–234.
- Tümpel, S., Cambronero, F., Ferretti, E., Blasi, F., Wiedemann, L.M., and Krumlauf, R. (2007). Expression of *Hoxa2* in rhombomere 4 is regulated by a conserved cross-regulatory mechanism dependent upon *Hoxb1*. *Dev. Biol.* *302*, 646–660.
- Williams, T.M., Williams, M.E., and Innis, J.W. (2005). Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction. *Dev. Biol.* *277*, 457–471.
- Wong, P., Iwasaki, M., Somerville, T.C., So, C.W., and Cleary, M.L. (2007). Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes Dev.* *21*, 2762–2774.