HAL open science

The TALE never ends: a comprehensive overview of the role of PBX1, a TALE transcription factor, in human developmental defects

Laura Mary, Delphine Leclerc, David Gilot, Marc-Antoine Belaud-Rotureau, Sylvie Jaillard

To cite this version:

Laura Mary, Delphine Leclerc, David Gilot, Marc-Antoine Belaud-Rotureau, Sylvie Jaillard. The TALE never ends: a comprehensive overview of the role of PBX1, a TALE transcription factor, in human developmental defects. Human Mutation, 2022, 43 (9), pp.1125-1148. 10.1002/humu.24388 . hal-03660222

HAL Id: hal-03660222

https://hal.science/hal-03660222

Submitted on 12 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Accepted Article

Laura MARY ORCID iD: 0000-0002-4981-3999

Sylvie Jaillard ORCID iD: 0000-0002-0593-0261

The TALE never ends: a comprehensive overview of the role of PBX1, a TALE transcription factor, in human developmental defects

Laura Mary[1,2], Delphine Leclerc[3], David Gilot[1,3], Marc-Antoine Belaud-Rotureau[1,2], Sylvie Jaillard[1,2]

[1] Service de Cytogénétique et Biologie Cellulaire, CHU Rennes, Rennes, France.

[2] INSERM, EHESP, IRSET - UMR_S 1085, Université Rennes 1, Rennes, France.

[3] Inserm U1242, Université de Rennes, France, Centre de lutte contre le cancer Eugène Marquis, Rennes, France

Corresponding author:

Laura Mary

Service de Cytogénétique et Biologie Cellulaire, CHU Pontchaillou

2 rue Henri Le Guilloux, F-35000 Rennes, France.

Phone: +33 2 99 28 83 38

laura.mary@chu-rennes.fr

Funding

This research did not receive any specific grant from commercial, or not-for-profit sectors. Authors would like to thank AVIESAN (Plan Cancer).

Abstract:

PBX1 is a highly conserved atypical homeodomain transcription factor (TF) belonging to the TALE (Three Amino acid Loop Extension) family. Dimerized with other TALE proteins, it can interact with numerous partners and reach dozens of regulating sequences, suggesting its role as a pioneer factor. PBX1 is expressed throughout the embryonic stages (as early as the blastula stage) in vertebrates. In human, *PBX1* germline variations are linked to syndromic renal anomalies (CAKUTHED). In this review, we summarized available data on PBX1 functions, PBX1-deficient animal models and *PBX1* germline variations in humans. Two types of genetic alterations were identified in *PBX1* gene. *PBX1* missense variations generate a severe phenotype including lung hypoplasia, cardiac malformations and sexual development defects (DSDs). Conversely, truncating variants generate milder phenotypes (mainly cryptorchidism and deafness). We suggest that defects in PBX1 interactions with

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/humu.24388.

.

various partners, including proteins from the HOX (HOXA7, HOXA10 etc.), WNT (WNT9B, WNT3) and Polycomb (BMI1, EED) families are responsible for abnormal proliferation and differentiation of the embryonic mesenchyme. These alterations could explain most of the defects observed in humans. However, some phenotype variability (especially DSDs) remains poorly understood. Further studies are needed to explore the TALE family in greater depth.

Keywords: Transcription factor – development – PBX1 – DSD – phenotype-genotype

1) Introduction

1.1) The PBC family of transcription factors (TFs)

Numerous classes of homeodomain (HD) transcription factors (TFs) have been described in the animal kingdom. One of these highly conserved super-classes of TFs is characterized by the insertion of a three aminoacid loop in the HD which has been named the TALE TF (for Three Amino acid Loop Extension, (Bertolino, Reimund, Wildt-Perinic, & Clerc, 1995)). Two TALE families, the MEINOX and PBC TFs, are able to interact with HOX and non-HOX proteins (reviews by (Jozwik & Carroll, 2012; Laurent, Bihan, Omilli, Deschamps, & Pellerin, 2008; Merabet & Mann, 2016; Penkov *et al.* 2013; Pi *et al*. 2020)). They also have the ability to recruit regulators of chromatin accessibility and transcriptional co-repressors (Asahara, Dutta, Kao, Evans, & Montminy, 1999; Laurent *et al.* 2008; Longobardi *et al.* 2014). These highly versatile functions make MEINOX and PBC proteins key factors in development, cell maintenance and cancer (see reviews by (Blasi, Bruckmann, Penkov, & Dardaei, 2017) for cancer and (Licia Selleri, Zappavigna, & Ferretti, 2019) for development).

The PBC family is composed of four proteins, namely PBX 1 to 4, with highly conserved sequences among vertebrates. As TALE TFs, their structure is composed of an HD and two protein-protein-interacting domains named PBC-A and B. The HD enables PBX proteins to interact with other TFs and DNA, but the PBC domains are

Accepted Article

.

critical for interactions between PBX and MEINOX (PKNOX (also named PREP) and MEIS) proteins (Longobardi *et al.* 2014). Hetero-dimerization of these proteins is essential to allow PBX DNA-binding capacities, inducing: 1) the import of PBX proteins into the nucleus by hiding their nuclear export signals (NES) (Abu-Shaar, Ryoo, & Mann, 1999; Jaw *et al.* 2000), 2) conformational changes of both partners stabilizing their interaction with DNA (Mathiasen *et al.* 2016), 3) the selection of targeted DNA with PBX protein binding properties relying on the DNA consensus sequence of their MEINOX partner (Blasi *et al.* 2017). Since the phenomena of PKNOX or MEIS binding to PBX proteins are mutually exclusive (Blasi *et al.* 2017), the properties of PBX are linked to its partner. Briefly, PKNOX and MEIS proteins have more or less opposite effects on cell fate. PKNOX proteins favour transcription initiation and preferentially bind to the promoter regions of housekeeping genes. PKNOX1 has been extensively studied in the setting of both development and cancer (Laurent *et al.* 2008; Penkov *et al.* 2013); it is ubiquitously expressed during development and is also considered as a tumour-suppressing gene. PKNOX2 is less well known, and shares structure and functions that are close to PKNOX1. PKNOX2 is mainly expressed in the heart, the brain, muscles, limb buds and the ovaries, and regulates gene expression, with less effective binding properties to PBX proteins than PKNOX1 (Fognani *et al.* 2002; Haller, Rambaldi, Kovács, Daniels, & Featherstone, 2002; Zhou *et al.* 2013). MEIS proteins target promoter-remote regions of developmental genes and favour cellular survival and proliferation. MEIS1, expressed early and ubiquitously, is involved in cancer, hematopoiesis and embryonic development, especially for eye, limb, heart and vasculature formation (review by (Paul, Zhang, & He, 2020)). MEIS2 shares common functions with MEIS1 and is expressed during development mainly in neural structures, neural crest-derived cells,

Accepted Article

.

the head and the gut, with overlapping patterns with MEIS1 (Berenguer & Duester, 2021; Fujita *et al.* 2016). Finally, MEIS3 has been far less studied in mammals. Its role in cancer varies with the tumour type (Meng *et al.* 2021). Meis3 has been shown to regulate neural crest cell migration, neural plate organization and pancreatic formation during mouse development (J. Liu, Wang, Birnbaum, & Stoffers, 2010).

1.2) PBX1 expression and role

Most of the PBX proteins are expressed throughout the embryonic stages (as early as the blastula stage in vertebrates (Choe, Lu, Nakamura, Lee, & Sagerström, 2009)) and maintained in mature tissues. During development, PBX1 is able to bind to promoters and promote histone modifications and chromatin accessibility, promoting gene transcription (Choe, Ladam, & Sagerström, 2014; Britta Moyo Grebbin *et al.* 2016). This binding can precede efficient gene transcription by hours or days (Choe *et al.* 2014). In addition, PBX1 (and TALE TFs in general) can use distinct DNA-binding motifs and protein partners depending on the embryonic stage. This suggests key roles for PBC proteins, from early development and throughout life (for an extensive review on TALE TFs and development, see (Licia Selleri *et al.* 2019)).

In the PBC family, PBX1 and 2 are the most widely expressed both during development and in mature organs (Roberts, van Dijk, & Murre, 1995; L. Selleri *et al.* 2001), with overlapping patterns (Capellini *et al.* 2006). Intriguingly, only PBX1 seems to be mandatory for embryonic development, since *Pbx2*-KO mice are viable and healthy (Licia Selleri et al., 2004). Multiple hypotheses may explain such a discrepancy, but a quantitative model, relying on global cell concentration of Pbx proteins, seemed the most accurate (see (Licia Selleri et al., 2004) and paragraph 3.3 below). Numerous studies have investigated ubiquitous patterns of PBX1 expression

.

Accepted Article

during development, mainly in mouse embryos (Capellini *et al.* 2010, p. 20, 2008; DiMartino *et al.* 2001; Golonzhka *et al.* 2015; Britta Moyo Grebbin *et al*. 2016; Lampe *et al.* 2008; Lescroart & Zaffran, 2018; Manley, Selleri, Brendolan, Gordon, & Cleary, 2004; Catherine A. Schnabel, Godin, & Cleary, 2003; L. Selleri *et al.* 2001; Sonnet, Rezsöhazy, & Donnay, 2012; Vitobello *et al.* 2011). In human adults, PBX1 is ubiquitously expressed, predominantly in glandular tissues, female tissues and the bladder (Fagerberg *et al.* 2014). It remains difficult to define the role of PBX1 proteins because alternative splicing can produce up to seven isoforms for PBX1 in humans, only three with known functions (PBX1a, PBX1b and PBX1d (Asahara *et al.* 1999; Sengupta, Liang, Potula, Chang, & Morel, 2012)). All proteins exhibit a degree of specificity in their patterns of expression (Laurent *et al.* 2008; C. A. Schnabel, Selleri, Jacobs, Warnke, & Cleary, 2001; Sengupta *et al.* 2012). PBX1a is mainly expressed in the brain and during adulthood. Conversely, PBX1b is expressed ubiquitously and preferentially during the embryonic period (Catherine A. Schnabel, Godin, *et al.* 2003). The PBX1d isoform seems specific to CD4+ T cells and is considered as a susceptibility factor for lupus (Cuda *et al.* 2012; Niu, Sengupta, Titov, Choi, & Morel, 2017). PBX1 has numerous partners leading to hetero-dimerization or hetero-trimerization (PKNOX1 and 2 proteins, MEIS1 and 2, FOXC1, HOX proteins etc. (Asahara *et al.* 1999; Berry, O'Neill, Coca-Prados, & Walter, 2005; Dard *et al.* 2019; Laurent *et al.* 2008)). In these larger complexes, PBX1 is able to recruit chromatin-remodelling factors with either repressing or enhancing activity (Asahara *et al.* 1999; Laurent *et al.* 2008; Penkov *et al.* 2013). PBX1 target genes include ubiquitous developmental genes (*PAX1/PAX9*, *WNT9B/WNT3*, *PITX3*, etc. (Capellini *et al.* 2008; Ma *et al.* 2015; Villaescusa *et al.* 2016) and organ-specific genes (*HIF1A*,

.

Accepted Article

DCX, *TH*, *GNRH* genes, etc. (Britta M. Grebbin & Schulte, 2017; Kocabas *et al.* 2015; Rave-Harel *et al.* 2004)), driving cell proliferation, pluripotency and differentiation (Fukuda *et al.* 2015; Gemel, Jacobsen, & MacArthur, 1999; Kocabas *et al.* 2015; Wei, Yu, & Li, 2018). Table 1 sums-up the most relevant *PBX1* cofactors and target genes with implication in development demonstrated in functional studies. For extensive gene modifications induced by PBX1 alone or by the chimeric E2A-PBX1 protein, see (Diakos *et al.* 2014; Thiaville *et al.* 2012).

PBX1 structure is similar to what is observed in other PBC family members, with three functional domains: an HD and two protein-protein-interacting domains, PBC-A and B (Figure 1). The heptad 77-84 in PBC-A and two regions (around Lys153 and heptad 197-211) in PBC-B are required for PBX1-MEINOX interactions (Bruckmann *et al.* 2020). Between the PBC-A and PBC-B domains, a poly-alanine stretch (108-150) confers flexibility on PBX1 and thus enhances its DNA-binding ability by favouring a conformational change upon binding (Blasi *et al.* 2017). DNA binding of PBX1 HD relies strongly on an alpha-helix (located in position 271-288) (Farber & Mittermaier, 2011). PBX1 HD is able to recognize a generic octameric motif (TGATNNAT), where N represents nucleotides that are specified by the HOX partner of PBX1 alone or PBX1-PKNOX/MEIS dimers (Longobardi *et al.* 2014). However, DNA motif recognition of PBX1 varies through development and depends on its partners (Ladam *et al.* 2018). Initially, PBX1 was thought to interact only with HOX proteins carrying a specific hexapeptide motif in their homeodomain called a W motif (Neuteboom, Peltenburg, van Dijk, & Murre, 1995; L. Selleri *et al.* 2001). However, recent studies have demonstrated that PBX1 is also able to interact with other amino acid motifs of HOX HD in place of, or in addition to, the W motif (Dard *et al.* 2019; Merabet & Mann, 2016). The use of these alternative motifs allows HOX

Accepted Article

.

Accepted Article

proteins to adopt different conformations, increasing the range of binding sites that are recognized. This increase in DNA recognition and binding gives the HOX-PBX1 complex the ability to recognize and bind low-affinity, non-consensus binding sites (Merabet & Mann, 2016) and thus increases the number of genes regulated by PBX1-HOX partners (Dard *et al.* 2019).

Localization of PBX1 is under the control of two nuclear localization signals (NLS) located in the HD (Berthelsen, Kilstrup-Nielsen, Blasi, Mavilio, & Zappavigna, 1999), and two nuclear export signals (NES) located close to and in the PBC-A domain (Saleh, Huang, Green, & Featherstone, 2000). As a monomer, PBX1 is folded in such a way that the NES (N-terminus) comes into close contact with the HD, hiding the NLS and thus retaining PBX1 in the cytosol (Saleh *et al.* 2000). PBX1 hetero-dimerization changes the conformation of PBX1, exposing the NLSs (Longobardi *et al.* 2014). The nuclear import of PBX1 can also be based on serine phosphorylation, depending on the cell type (Kilstrup-Nielsen, Alessio, & Zappavigna, 2003).

Even if the target genes of PBXs are not fully discovered yet, these proteins, and especially PBX1, have long been linked to cell fate, embryonic development and cancer in humans. The first human phenotype linked to *PBX1* dysfunction was acute leukaemia by way of the formation of an *E2A-PBX1* fusion transcript (Kamps, Look, & Baltimore, 1991; C.-H. Lin *et al.* 2019). Further studies have also highlighted the role of *PBX1* in promoting the proliferation of various solid tumours (Blasi *et al.* 2017; He *et al.* 2017; Jung *et al.* 2016; Magnani *et al.* 2015; Wei *et al.* 2018). More recently, *PBX1* constitutional heterozygous deletions or point variations have been linked to developmental defects, especially to congenital anomalies of the kidney and urinary tract (CAKUTHED; MIM# 617641)(Le Tanno *et al.* 2017). Despite growing knowledge on the role of *PBX1*, human developmental phenotypes induced by *PBX1*

.

variations are not yet fully understood. The increasing number of cases and the pleiotropy of the symptoms are presented in this review focused on the human developmental phenotype induced by *PBX1* variations.

2) *PBX1* variations in humans

The *PBX1a* transcript (NM_002585.4) will be used as a reference in the present manuscript. Point variations and structural variants with an allele count of 1 were omitted, since we could not exclude the presence of very mildly affected individuals in the databases used (gnomADv2, DGV database and (Alsmadi *et al.* 2014; Audano *et al.* 2019; Thareja *et al.* 2015; Uddin *et al.* 2015; Wong *et al.* 2013)). Two hundred and fifty-eight point variations or small indels and 24 structural variations (15 deletions, 6 duplications and 3 insertions) have been reported to date in *PBX1*. Only 2 of these late deletions have an impact on the *PBX1* coding region (exon 1, (Alsmadi *et al.* 2014; Thareja *et al.* 2015)). The exact count for each variation type is reported in Supp. Table S1.

Thirty patients with heterozygous pathogenic mutations, and partial or complete heterozygous deletions of *PBX1* have been reported in the literature (Alankarage *et al.* 2020; Bertoli-Avella *et al.* 2021; Eozenou *et al.* 2019; Fitzgerald, Powell-Hamilton, Shillingford, Robinson, & Gripp, 2021; Fromer *et al.* 2014; Heidet *et al.* 2017; Le Tanno *et al.* 2017; Riedhammer *et al.* 2017; Slavotinek *et al.* 2017), and 10 additional patients are reported in the ClinVar, DECIPHER and HGMD databases. Their clinical data is summarized in Table 2. The PBX1 structure and its related variants are summarized in Figure 1 and Supp. Figure S1. None of these variations was reported in gnomAD, except one (NM_001204961.1:c.862C>T, p.(Arg288Ter) in a heterozygous state and at a frequency of $0.3.10^{-4}$).

2.1) *PBX1* loss-of-function (LOF)

.

Accepted Article

Accepted Article

Whole or partial gene deletions account for 43% of all the *PBX1*-mutated patients, and nonsense variations (stop gains) or frame-shift mutations (including indels) amount to 23% of the patients. Nonsense or frame-shift variations mainly affect the polyalanine stretch between PBC-A and PBC-B (56%). Only one nonsense variation was located outside a functional domain (Figure 1).

To date, seventeen patients have been reported with whole or partial deletions of *PBX1* (Fitzgerald *et al.* 2021; Heidet *et al.* 2017; Le Tanno *et al.* 2017) of whom 3 were siblings (Fitzgerald *et al.* 2021). The size of these deletions ranged from 37 kb (Fitzgerald *et al.* 2021) to 18 Mb (patient 398083 from DECIPHER). Nine patients were described as presenting LOF variations (nonsense, splice or frame-shift variations, including indels) and 6 had documented phenotypes. The clinical data of these patients is summarized in Figure 2 panel A and extensively reported in Supp. Table S2. Phenotypes are highly similar between variants resulting in LOF and *PBX1* gene deletions; we therefore considered them as the same entity. Most of these patients (83%) displayed various combinations of renal and/or urinary tract anomalies, the most common defect being renal hypoplasia. The only genital anomaly reported was cryptorchidism. Dysplastic ears were frequently reported (43%), sometimes associated with deafness. Cardiac malformations were mainly conotruncal defects and patent ductus arteriosus. Rare cases of arrhythmia were also documented. Interestingly, these cardiac anomalies were only reported in patients with *PBX1* deletions. The hypothesis of the contiguous gene syndrome to explain certain extra-renal phenotypes has been proposed (Le Tanno *et al.*,- 2017) but the cardiac phenotype could not be linked to another gene in close proximity to *PBX1*. The implication of *PBX1* in the cardiac phenotype is discussed below.

2.2) *PBX1* missense variations

.

Accepted Article

Missense variations were diagnosed for thirteen (33%) of the *PBX1*-mutated patients. Most of these missense variations identified in the HD of PBX1 lie in the NLS (Figure 1), with two recurrent positions affected (Arg234 and Arg235), which were studied using functional analyses [63, 64](see below).

Of these 13 patients, 9 had detailed clinical features (Arts *et al.* 2020; Eozenou *et al.* 2019; Kia, Sarafoglou, Mooganayakanakote Siddappa, & Roberts, 2019; Slavotinek *et al.* 2017)(Figure 2 panel B and Supp. Table S2). Renal or urinary anomalies were less prevalent in this group than in the *PBX1* LOF group and were milder (hypoplastic or horseshoe kidneys, ureteral dilatation or hyperechogenic kidneys). Dysplastic ears were rarer. One third of the patients had mild skeletal anomalies (brachydactyly, clinodactyly, radioulnar synostosis). Unlike *PBX1* LOF patients, no central nervous system malformations were reported. Cardiac anomalies were frequent (55%) and similar to the cardiac defects found among patients carrying *PBX1* deletions.

In contrast with PBX1 LOF patients, genital anomalies were frequent and severe. Only one 46,XX individual was described in this cohort as having a bicornuate uterus. Six of the 8 46,XY individuals had various severe phenotypes ranging from cryptorchidism and micropenis to complete sex reversal. These last phenotypes included a nearly normal female phenotype for both internal and external genitalia and gonads (1 patient), various gonadal anomalies (ovotestis, streak gonads) and internal genitalia malformations. Lung hypoplasia was observed in 4 patients. For 3 patients, this lung hypoplasia was associated with diaphragmatic thinning or defect (different from diaphragmatic hernias).

Functional assessment of PBX1 HD variations

.

Accepted Article

Slavotinek *et al.* (Slavotinek *et al.* 2017) studied the functional roles of *PBX1* mutations, notably generating the p.(Arg234Pro) and the p.(Arg235Gln) mutants in HEK293T transfected cells with plasmids encoding these variants. They demonstrated that these variations affected neither levels of protein expression nor nuclear localizations in their conditions. Both mutant proteins showed reduced transactivation properties of a luciferase gene preceded by a *PBX1* promoter, but in a *Pbx1*-KO murine cell line, only p.(Arg234Pro) had reduced transactivation properties. The authors suggested that the interactions between endogenous PBX1 protein in HEK293T cells and mutated transfected PBX1 could adversely affect these results. Eozenou *et al.* (Eozenou *et al.* 2019) studied the p.(Arg235Gln) mutant in HEK293T cells and demonstrated that, despite a few mutant proteins localized in the nucleus, the variation favoured a cytoplasmic retention of mutated PBX1. Moreover, they demonstrated that the mutant protein was unable to form heterodimers with HOXB8. Although the binding of the mutant protein to PKNOX1 was maintained, the binding of mutant PBX1 to EMX2 was reduced and abolished with CBX2. The discrepancy in nuclear localization between these two publications that studied the same variation in the same cellular model raises interesting questions, such as the impact of *in vitro* approaches on PBX1 function. Eozenou *et al.* used myc-tagged PBX1 expression vectors, unlike Slavotinek *et al.*, who transfected HEK293T cells with the cDNA of interest (plus *PKNOX1* cDNA) and analysed the protein content of the cell compartments by Western-blotting experiments. Variations might also result to the expression levels of PBX1 reached after plasmid transfection. However, mechanisms other than the use of NLS enable PBX1 to reach the nucleus: phosphorylation of serine residues in the PBC-B domain also has a role in PBX1 nuclear import (Figure 1, (Kilstrup-Nielsen *et al.* 2003)). In contrast, interaction with proteins of the

.

Accepted Article

cytoskeleton, namely NMHCB or FLNA, retains PBX1 in the cytosol (Berry *et al.* 2005; Huang, Paliouras, Rambaldi, Lasko, & Featherstone, 2003). Both proteins compete with MEIS1 to bind PBX1, but it is the interaction between MEIS1 (or PKNOX1) and PBX1 that enables PBX1 conformational change and exposure of its NLS (Longobardi *et al.* 2014; Saleh *et al.* 2000). It is tempting to postulate that PKNOX1 overexpression could partially overcome mutant PBX1 cytoplasmic retention. On the contrary, a myc-tag may worsen an abnormal conformational change induced by a morbid variation in the PBX1 NLS domain.

Thus, these studies suggest that variations in the NLS lead to a partial loss of PBX1 nuclear localization, but that they also prevent, at least partly, PBX1 interaction with some of its cofactors, two defects that could explain the phenotype induced in humans which recalls, but does not fully phenocopy, the *Pbx1*-KO mouse phenotype.

Two other variations in the HD have been reported to date but not extensively studied: p.(Cys273Tyr) and p.(Gln279Arg) (ClinVar VCV000598775 and (Bertoli-Avella *et al.* 2021)). Both native amino acids are highly conserved across species (Supp. Figure S1). Both residues are positioned on a segment undergoing conformational changes (Ser271-Arg288, (Blasi *et al.* 2017)), which are important in increasing DNA binding ability. We can suggest that these variants alter DNA binding rather than PBX1 localization (as demonstrated with other variants by (Slavotinek *et al.* 2017)) or they disrupt interaction with non-HOX partners. Functional studies are required to explore these hypotheses.

Functional assessment of PBX1 non-HD variations

Only one variation in the PBC domains was analysed in a functional study by Slavotinek et al. (Slavotinek *et al.* 2017), with the protocol previously mentioned. The authors demonstrated that, as for variations in the HD, expression levels of the mutant

.

PBX1 Arg227Pro were comparable to the WT protein. Mutant proteins also reduced transactivation properties in HEK293T cells but not in a *Pbx1*-KO murine cell line. The authors observed no change in nuclear levels in cells expressing the mutant protein compared to WT transfected cells.

The four other variations have not been extensively studied. One patient from a cohort of schizophrenic individuals was described with a substitution of the threonine in position 88 by an isoleucine (Fromer *et al.* 2014). This variation is located in PBX1 NES. PBX1 NES prevents PBX1 nuclear localization by interacting with its NLS (Saleh *et al.* 2000) and it is masked after PBX1 dimerization with PKNOX1 (Berthelsen *et al.* 1999). The substitution of a threonine by a more hydrophobic amino acid could disrupt NES/NLS interactions. In fact, the NLS has an arginine-rich structure that could be destabilized by hydrogen bond modifications, but this hypothesis remains to be explored.

Two siblings were described with a variation substituting the arginine in position 107 for a tryptophan (Arts *et al.* 2020). This Arg107 is located just before one poly-alanine segment (Leu108-Tyr150) that plays an important role in PBX1 flexibility and DNA binding (Blasi *et al.* 2017). Deletion of this helix region (positions 105-113) mildly altered the ability of PBX1 to dimerize with PKNOX1, but did not influence PBX1 dimerization with MEIS1 (Bruckmann *et al.* 2020). However, *in silico* prediction tools (HOPE tool (Dunlavy, O'Leary, Klimov, & Thirumalai, 2005; Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010)) suggest that this amino acid substitution would be damaging for the correct folding of the alpha-helical region (positions 103-116, (Bruckmann et al., 2020)). Arg107 is conserved in all species (Supp. Figure S1).

Accepted Article

.

Accepted Article

One patient was reported with a substitution of the arginine in position 184 by a proline (Alankarage *et al.* 2020; Slavotinek *et al.* 2017). The Arg184 was conserved up to Drosophila. This substitution could disrupt the alpha helix structure implicated in interaction with MEINOX proteins ((Bruckmann *et al.* 2020), HOPE results). Moreover, this arginine is located in a MAP-kinase recognition motif (Calvo, Knoepfler, McGrath, & Kamps, 1999) and phosphorylation is a way for PBX1 to reach the nucleus ((Kilstrup-Nielsen *et al.* 2003), see above).

Finally, another patient was reported with a substitution of the methionine in position 224 by a lysine (Slavotinek *et al.* 2017). Met224 is conserved in all species (including *C. elegans*). This methionine is located in a potential "minimal inhibitory helix", a domain that could play a role in preventing PBX1 monomers from binding DNA, an effect suppressed when PBX1 dimerizes (Calvo *et al.* 1999). However, the function of this domain has not been assessed in more recent studies.

Thus, missense variations in the PBC domains can lead to various deleterious effects: absence of interaction with MEINOX proteins, modification of DNA binding stringency or alterations in nuclear import. However, more robust functional studies are needed to assess the role of these variants.

3) Shedding light on human *PBX1*-induced phenotypes using animal models

Animal models of heterozygous *PBX1* LOF are not helpful in explaining these developmental anomalies, since heterozygous mice for a *Pbx1* null allele merely exhibit smaller size than their wild-type littermates, even if they express only half the normal level of the protein (L. Selleri *et al.* 2001). One notable exception concerns mice heterozygous for the p.(Arg184Pro) mutation that sometimes present with a phenotype close to that of their KO littermates, although it tends to be less severe

.

(Alankarage *et al.* 2020). However, PBX1-null mouse malformations clearly recall the developmental phenotype reported in humans (Figure 2 panel C)(L. Selleri *et al.* 2001).

3.1) Head, face and neck morphogenesis anomalies

The main defect observed in *Pbx1*-KO mice is a homeotic transformation of structures derived from the second branchial arch and the first pharyngeal groove into first branchial arch derivatives (L. Selleri *et al.* 2001). PBX proteins are regulators of HOXA2 expression in rhombomere 4 (Lampe *et al.* 2008), which gives rise to the ectomesenchyme of the second branchial arch (Couly, Grapin-Botton, Coltey, & Le Douarin, 1996). However, the mouse phenotype is not completely reminiscent of *Hoxa2*-KO mice and *Pbx1*-deficient animals also present with delayed third pouch development, including thymus and parathyroid (Manley *et al.* 2004). *Pbx1*-KO mice also display external and inner ear anomalies that are not fully explained by *Hox* gene dysregulation. Hmx1, a homeodomain transcription factor, plays a role in lateral facial morphogenesis, notably in pinna formation, and possesses a downstream regulatory sequence that can be bound by the Hox-PBX-MEIS complex (Rosin *et al.* 2016). Deletions of this highly conserved regulatory region are responsible for the *dumbo* phenotype in mice and rats (Quina *et al.* 2012). Mutations in another homeobox gene, *Emx2*, which is known to be regulated by PBX1 in limb development (see below), also have a role in middle and inner ear formation. Mice mutated on this gene present ossicular chain malformations and abnormal hair cell development (Holley *et al.* 2010). Abnormal patterns or expression of these two genes, combined with *HOXA2* dysregulation, could be involved in the cranial and hearing phenotypes observed in patients, but this remains to be further explored. Finally, it has been demonstrated that PBX proteins control facial morphogenesis by promoting epithelial apoptosis through the Pbx-Wnt-p63-Irf6 network, and a compound loss of Pbx proteins in mice leads to

.

Accepted Article

Accepted Article

cleft lip and palate (Ferretti *et al.* 2011). Interestingly, none of the *PBX1*-mutated patients previously reported presented oral clefts, but SNP in *PBX1*, *WNT9B* and *TP63* was significantly associated with non-syndromic cleft lip and palate in humans (Maili *et al.* 2020).

3.2) Skeletal and limb deformities or hypoplasia

Mice KO for *Pbx1* develop constant and severe skeletal anomalies compared to what is observed in humans. These anomalies mainly involve the cervical and thoracic vertebrae, the ribs and sternum, and proximal structures of the limbs (L. Selleri *et al.* 2001). The scapula, clavicle and humerus in particular are hypoplastic and misshapen, resembling some malformations described in patients. In mice, Pbx1 and Pbx2 control both Polycomb and Hox TFs (namely Bmi, Eed, Hoxa4, Hoxb5, Hoxb8, Hoxc6, and Hoxd3) to establish hind limb positioning, (Capellini *et al.* 2008). Pbx1, in association with Emx2, also controls the expression of scapula-specific genes such as *Alx1* (Capellini *et al.* 2010). More recent studies have also suggested that PBX1 and PBX2 work together in regulating the pattern of the proximal limbs in a dose-dependent manner (Eyal *et al.* 2019). PBX1/PBX2 also interact with GLI3, one of the main controllers of the global limb pattern regulatory network (Eyal *et al.* 2019) and an organizer of neural tube patterns. At cellular level, mice lacking Pbx1 display diminished chondrocyte proliferation with precocious ossification (L. Selleri *et al.* 2001). Pbx1 normally repress the ability of Hoxa10 to activate osteoblast-related gene programs (Gordon *et al.* 2010). This property could explain why *Pbx1*-KO mice display early abnormal osseous formation. The fact that heterozygous mice for a *Pbx1*-null allele are smaller than WT animals also suggests early halting of growth as a result of premature ossification. This late phenotype has not been reported in

.

humans, but most of the patients reported to date were children and we cannot fully assess the impact of their *PBX1* variant on their adult stature.

3.3) Abnormal cardiac outflow tract development

Pbx-deficient mice present heart defects that mimic the malformations observed in humans, arising from abnormal cardiac outflow tract development (Stankunas *et al.* 2008). The anomalies observed in mice partly depend on which Pbx protein is deficient, but *Pbx1* seems to be the main gene implicated in these malformations, with relatively minor roles for *Pbx2* and *Pbx3*. These effects on the cardiac outflow tract are at least partially explained by the PBX1 requirement for correct expression of *PAX3* in pre-migratory neural crest cells, a cell population highly implicated in heart development (review by (Lescroart & Zaffran, 2018)). We can also postulate a role of MEIS2 in this expression, since MEIS2 mutations in humans lead to similar defects in neural crest-derived structures (Douglas et al., 2018; Fujita *et al.* 2016). At cellular level, PBX1 controls both cardiac precursor proliferation in synergy with GATA4, and cell cycle arrest in mature cardiomyocytes by transcriptional activation of CDKN2a, with a suspected role of MEIS1 as a cofactor (Hatzistergos *et al.* 2019). Thus, PBX1 is a key factor in heart formation, and it is worth noting that cardiac anomalies are not very frequent in patients with *PBX1* LOF, unlike patients carrying missense variants (26% vs. 55%). In fact, some authors (Stankunas *et al.* 2008) demonstrated that the severity of heart defects in various *Pbx*-deficient mice correlated with the total concentration level of PBX proteins. We can postulate that a decrease in PBX1 concentrations through nonsense-mediated decay can be counterbalanced, at least partially, by other PBX proteins (Stankunas *et al.* 2008) in patients carrying *PBX1* LOF variants. In fact, PBX proteins share the same HD sequence and are able to target the same binding sites (Longobardi *et al.* 2014; Penkov

Accepted Article

et al. 2013). *Pbx1*-KO mice have a more severe phenotype while they also carry a second hit in *Pbx2*, despite the fact that *Pbx2*-KO mice are viable and healthy (L. Selleri *et al.* 2001). This finding highlights the redundancy of PBX proteins (Capellini *et al.* 2006). In contrast, *PBX1* missense variants can lead to defective PBX1 proteins retaining PBX1 cofactors (dominant-negative effect, see paragraph 4 and Figure 3), preventing partial rescue by other PBX proteins and leading to severe cardiac phenotypes.

3.4) Lung hypoplasia and delayed maturation

PBX1 (and especially the PBX1b isoform) is expressed early in the pulmonary mesenchyme during the initial stages of budding and outgrowth of the bronchial tree. Pbx1b immuno-reactivity persists in the pulmonary interstitium until adulthood in mice (C. A. Schnabel *et al.* 2001). *Pbx1*-KO mouse embryos show severe lung hypoplasia (L. Selleri *et al.* 2001) and conditional KO in mouse lungs has demonstrated abnormal lung maturation with compact terminal saccules and severely reduced expression of surfactant genes (Li *et al.* 2014). The loss of *Pbx1* in the lung mesenchyme also results in continuous vascular constriction and elastin matrix anomalies (McCulley *et al.* 2018). The role of Pbx1 in the developing lung resides, at least partly, in its ability to control the expression of Fgf10, a key factor for alveolar type-II cells and differentiation of lung epithelial progenitors, in collaboration with HoxB4 and Meis1 (Li *et al.* 2014). PBX1 interactions with *HOXA5*, *HOXB5* and, more recently, the repression of *TBX2*, are also of importance in tracheal budding, mesenchyme induction and lung maturation (Jeannotte, Gotti, & Landry-Truchon, 2016; Lüdtke *et al.* 2021).

3.5) Diaphragmatic thinning and defects

Accepted Article

.

Accepted Article

A few patients carrying *PBX1* missense variations presented diaphragmatic thinning and defects (Arts *et al.* 2020; Slavotinek *et al.* 2017). By studying the embryonic transcriptome in the search for genes implicated in diaphragmatic hernia, Russel et al. (Russell *et al.* 2012) identified *PBX1* as a candidate gene. In the *Pbx1*-KO mice they generated, the authors demonstrated that the homozygous mutant mice displayed diaphragmatic and muscle-pattern defects, with regions of the diaphragm reduced to a thin membrane with poor or absent musculature and abnormal muscle marker expression. The authors postulated that these defects originated from a dysregulation of *Hoxa5* and *Hoxb5* through dysregulation of retinoic acid (RA) synthesis, since previous studies had demonstrated that PBX1 was able to control RA synthesis (Vitobello *et al.* 2011). Other signalling pathways, such as SHH (PBX1 interacts with GLI3 during limb formation, see above) could be implicated in the diaphragmatic defects highlighted in *Pbx1*-deficient mice and *PBX1*-mutated patients. The absence of PBX1 could lead to defects in the mesenchymal proliferation of the pleuroperitoneal folds, a structure from which the diaphragm originates. A recent study reported a decrease in *Pbx1*, *Meis1*, and *Runx1* expression in rats with diaphragmatic defects and lung hypoplasia, suggesting reduced mesenchymal cell proliferation as a common mechanism for both diaphragmatic defects and abnormal bronchial budding in *Pbx1*-deficient rodents (Takahashi, Friedmacher, Zimmer, & Puri, 2020). Impaired mesenchymal cell proliferation has been previously demonstrated as a mechanism leading to renal anomalies (see below). Finally, PBX1 is also a cofactor for proteins implicated in myoblast differentiation, such as M-cadherin, (Y.-J. Lin, Kao, Hsiao, & Chen, 2021).

3.6) Pancreas hypoplasia and splenic anomalies

.

Accepted Article

Pbx1-KO mice are also characterized by hypoplastic pancreas and asplenia (L. Selleri *et al.* 2001). Intriguingly, only two patients (siblings) presented asplenia or small spleen associated with accessory spleens. No pancreatic phenotype (pancreas hypoplasia, maldigestion, diabetes…) was reported among *PBX1* patients. In *Pbx1* mouse mutants, Tlx1 and Nkx2.5, early markers for splenic progenitor cells, are not expressed (Brendolan *et al.* 2005). PBX1 is able to bind to TLX1 promoter, alone and in cooperation with TLX1 itself, underlining the fact that PBX1 is a key element in spleen development. During embryogenesis, both *PBX1* and *PBX2* are expressed in the developing pancreas. *PBX1* is mainly expressed in the pancreas mesenchyme, unlike *PBX2*, which is more broadly expressed, including in the pancreatic epithelium (Zhang *et al.* 2006). Both PBX proteins, in association with PKNOX1, bind to the pancreatic *PAX6* enhancer (Zhang *et al.* 2006). Thus, in *Pbx1*-deficient mouse embryos, Pax6 expression is strongly reduced in the pancreas, which is not the case for Pdx1, although this TF is also a cofactor of Pbx1 (Asahara *et al.* 1999). Finally, as demonstrated in the limbs (Capellini *et al.* 2006), the pancreatic defects observed in *Pbx1*-KO mouse embryos are more severe when a second *Pbx2* hit is present (Zhang *et al.* 2006). In the adult pancreas, the expression of *Pbx1* isoforms differs in exocrine and endocrine tissues, Pbx1b being expressed in acinar cells and Pbx1a preferentially in Langherans islets (Asahara *et al.* 1999). The absence of pancreatic phenotypes among patients could be explained by a partial rescue of PBX1 function by PBX2, as suggested for heart development (see above). Finally, mild pancreatic features could be underreported by patients or clinicians, or appear later in life (since most of the PBX1-mutated patients described were young children).

3.7) CAKUTHED

.

Accepted Article

In KO mice, renal anomalies are similar to the human CAKUTHED phenotype, with hypoplastic caudally positioned and ventrally rotated kidneys. Renal hypoplasia is linked to oligonephronia, due to a reduced nephrogenic zone and nephronic differentiation (Catherine A. Schnabel, Godin*, et al.* 2003*)*. Schnabel *et al.* (Catherine A. Schnabel, Godin, *et al.* 2003) demonstrated that Pbx1 was expressed in the nephrogenic mesenchyme, mainly in the stroma. Renal defects in KO mice rely on both deficient ureteric branching and an abnormal extension of the induced nephrogenic mesenchyme. These anomalies could be the consequence of abnormal interactions between Pbx1 and other TFs, such as *Hoxd* genes (Patterson & Potter, 2004). The *Hoxd* gene cluster has been shown to be implicated in the regulation of ureteric branching response and the maintenance of the structural integrity of the tubular epithelia (Di-Poï, Zákány, & Duboule, 2007). However, the *Hoxd* genes are known to interact with Pbx1 and Meis1 during limb bud pattern development (Capdevila, Tsukui, Rodríquez Esteban, Zappavigna, & Izpisúa Belmonte, 1999). Similar interactions could explain the renal phenotype observed in *Pbx1*-KO mice and *PBX1*-mutated patients, but this hypothesis remains to be explored. Moreover, PBX1 also has a prominent role in the establishment of renal vascular patterns: conditional *Pbx1*-KO in mouse renal vascular mural cells (VMC) leads to a loss of *Pdgfrb* down-regulation, premature VMC differentiation, vascular pattern defects and thus early kidney dysfunction (Hurtado *et al.* 2015). These findings suggest a prominent role of PBX1 in the induction and differentiation of the renal mesenchyme and especially of vasculature progenitors, essential for organ development and function.

3.8) Genital organs and genitalia anomalies

Analyses of mice KO for *Pbx1* showed that the anomalies observed for kidneys were part of a larger spectrum of defects of the urogenital ridge mesodermal

.

Accepted Article

derivatives (Catherine A. Schnabel, Godin, *et al.* 2003). PBX1 is expressed early in the coelomic epithelium, nephrogenic cord and adrenogenital precursors, it is maintained in the somatic cells of the undifferentiated gonad and finally in the interstitium of the mature gonad (Catherine A. Schnabel, Godin, *et al.* 2003). Gonads of *Pbx1*-KO animals initiate sexual differentiation but remain rudimentary due to reduced cell proliferation rate, and the expression of the steroid hydroxylase is barely detectable in gonads of male mutant mice. Interestingly, the authors also demonstrated that, in addition to gonadal anomalies, the mice completely lacked adrenal formations, due to the absence of *Nr5a1* expression (SF-1, a key transcription factor implicated in steroidogenesis and sexual differentiation, review by (Tremblay & Viger, 2003)). However, SF-1 expression appeared to be retained in gonadal cells, suggesting a differential regulation of *Nr5a1* expression depending on the tissue (Catherine A. Schnabel, Selleri, & Cleary, 2003). These results demonstrate that *Pbx1* does not interfere in the establishment of the urogenital ridge intermediate mesoderm but plays a prominent role in the establishment of the bipotential gonad (Catherine A. Schnabel, Selleri, *et al.* 2003) and is active upstream from *Nr5a1* only in adrenal development. Moreover, *Pbx1*-KO mice lack Müllerian ducts, but not Wollfian ducts which are normally formed and joined. These findings have led some authors to explore the role of *PBX1* in Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome. Burel et al. (Burel *et al.* 2006) did not find morbid variations in the *HOX* genes or *PBX1* in MRKH patients, but Ma et al. (Ma *et al.* 2015) found an increased risk of MRKH syndrome in association with rare SNP in *PBX1* and *WNT9B*, which are known to interact together in facial morphogenesis (see above and (Ferretti *et al.* 2011; Maili *et al.* 2020)). More recently, Chen *et al*. (Chen *et al.* 2021) also found MRKH candidate variants in *WNT9B*, *HOXA10* and *EMX2*. These last two genes are known to interact with or be

.

Accepted Article

recruited by PBX1 during limb formation and osteoblast activation (Capellini *et al.* 2010; Gordon *et al.* 2010). It is also worth noting that *Hoxa10*-KO male mice present bilateral cryptorchidism (Satokata, Benson, & Maas, 1995), a phenotype also observed in some 46,XY patient carriers of *PBX1* truncating variants. Thus, we can hypothesize that a partial disruption of PBX1 interaction with these effectors, especially *HOXA10*, could explain certain mild phenotypes observed in *PBX1*-mutated patients. However, the discrepancies between mouse models and *PBX1*-mutated patients, especially for adrenal agenesis and sex-dependent gonadal/genital malformations, are not well known. These differences recall the variations observed between *NR5A1*-mutated patients and *Nr5a1*-KO mice: XY mice are sex-reversed and present adrenal agenesis, unlike patients who can either have isolated adrenal deficiency or 46,XY DSD (Biason-Lauber 2000). The phenotypes of *Nr5a1*-KO mice are known to vary depending on the strain studied, suggesting a prominent role of other genetic modifiers. This phenomenon could also play a partial role in the phenotypic variability observed in *PBX1*-mutated patients, where DSDs are not fully penetrant. These results demonstrate that the precise role of PBX1 in human gonadal and genital development has not yet been fully understood.

4) Genotype-phenotype variations in *PBX1*-mutated patients

Two distinct phenotypes can be identified for *PBX1*-mutated patients (as suggested by (Slavotinek *et al.* 2017), Figure 2 A and B). First, a phenotype of syndromic CAKUTHED with skeletal anomalies, deafness, ear dysplasia, neurodevelopmental delay and cryptorchidism is observed for patients carrying a LOF mutation or a gene deletion. A more severe phenotype is observed for patients with missense variations, with gonadal and genital anomalies that can lead to complete sex reversal in 46,XY individuals and lung hypoplasia sometimes associated with

.

Accepted Article

diaphragmatic thinning. This discrepancy has already been discussed by Slavotinek et al. (Slavotinek *et al.* 2017). Genotype-phenotype correlations are of importance to further assess protein function but above all to provide appropriate genetic counselling and suggest the most suitable management for the patients.

Phenotypic variation depending on the mutation type is a mechanism previously described for other TFs. For example, *GLI3* mutations cause Greig syndrome (MIM# 175700) in case of inactivating variations or haplo-insufficiency (through gene deletion), unlike truncating variants in the middle third of the gene, which are responsible for Pallister-Hall syndrome (MIM# 146510). Similarly, we can postulate different mechanisms to explain the phenotypic variability observed in *PBX1*.

4.1) Nonsense-mediated decay (NMD)

Stop or frame-shift mutations in *PBX1* are thought to lead to mRNA NMD, as described for other TFs such as *PAX6* (Lima Cunha, Arno, Corton, & Moosajee, 2019). The NMD of *PBX1* mRNA could lead to a decrease in PBX1 concentration, but the WT allele and other PBX proteins (Figure 3) could counterbalance this loss, at least partially. There is partial functional redundancy between PBX proteins (Stankunas *et al.* 2008). However, total concentration levels of all PBX proteins are also of importance for normal cellular functions, as explained above for cardiac development. Dose-dependency has been previously reported for other TALE TFs, such as *MEIS1* (Marcos *et al.* 2015). Haplo-insufficiency has been described in numerous diseases caused by mutations in TFs (Seidman & Seidman, 2002). Haplo-insufficiency could be the only pathogenic mechanism reported for the gene (for example, *MYRF*, *HAND2* or *NFIB*, (Cohen *et al.* 2020; Rossetti *et al.* 2019), or it could be associated with various gene function disruptions caused by missense mutations (for example *PAX6*). However, these later mutations are often responsible for phenotypes that are milder

.

than haplo-insufficiency (Lima Cunha *et al.* 2019), unlike what is observed in *PBX1*-mutated patients.

4.2) mRNA instability and post-translational protein modifications

Mechanisms other than NMD can alter mRNA stability or translation, leading to a drastic decrease in the total protein levels. mRNA secondary structures are known to alter protein expression through changes in mRNA half-life (Mauger *et al.* 2019). Any nucleotide alteration of the secondary structure of mRNA can lead to premature mRNA decay and haplo-insufficiency-like phenotypes in patients carrying *PBX1* missense variations. Micro-ARNs (miRNA) are also of importance in regulating mRNA levels: 3'UTR variations affecting miRNA binding sites are implicated in human diseases such as Tourette's syndrome (review by (Kawahara, 2014)). Very recently, a SNP in the 3'UTR sequence of *PBX1* was linked to poor prognosis in breast and gastric cancers, by alteration of the regulatory affinity of miR-522-3p (Mohammadi, Salehzadeh, Talesh Sasani, & Tarang, 2021). Thus, the occurrence of aberrant miRNA binding sites in *PBX1* could also play a role in the milder phenotypes caused by missense variants. It is worth noting that *PBX1* 3'UTR variants are rare in the general population (64 variants reported in the gnomAD database, of which only 11 had an allele count > 10). However, 3'UTR variants are not fully explored by exome sequencing, a routine first-tier technique, and they may be under-diagnosed in CAKUTHED patients.

Post-translational changes such as ubiquitinylation or phosphorylation are of importance for correct protein function or degradation. At least 14 phosophorylation sites are reported in PBX1 (https://www.phosphosite.org/). Only the sites located in the PBC-B domain have been studied *in vivo*, with proven effect on PBX1 nuclear

Accepted Article

.

localization (Kilstrup-Nielsen *et al.* 2003), and no germline variation affecting, or creating, a phosphorylation site has been reported to date.

4.3) Dominant-negative or gain-of-function (GoF) effects

Missense variations can also lead to dominant-negative or GoF effects, especially when proteins are engaged in multimeric complexes. This type of mechanism is well known in TP53-linked pathologies for example, where TP53 tetramers formed by both normal and mutant proteins (Gencel-Augusto 2020) affect the stability and localization of the complex. These mechanisms were also observed in complexes composed of different subunits, such as the chromatin-remodelling SWI/SNF complex. Mutations in genes encoding subunits of this complex are known to cause Coffin-Siris syndrome (MIM# 135900): depending on the protein affected, both dominant-negative effects and GoF have been described (Kosho, Okamoto, & Coffin-Siris Syndrome International Collaborators, 2014). Since *PBX1* forms hetero-dimers or trimers with various cofactors, we can postulate that *PBX1* missense variations can generate the expression of abnormal PBX1 proteins that can compete with normal PBX1 TF, either for cofactor dimerization or DNA binding (Figure 3).

4.4) Second genetic hit and epigenetic alterations

However, since PBX1 can interact with various cofactors with pleiotropic modalities, a second genetic hit (such as a rare SNP) in one of these cofactors, in another TALE protein, or in one of PBX1 DNA-binding sites, could play a role in this phenotypic variability. The potential role of a second variation in another TALE protein has been partially discussed above: Pbx1 KO mouse models have a more severe phenotype when they carry a heterozygous variant in another Pbx protein (Capellini et al., 2008). We can postulate that partial disruption of a MEIS or PKNOX partner could worsen the phenotype. The hypothesis of a second genetic hit has

Accepted Article

.

Accepted Article

already been reported for other diseases linked to TF mutations, such as *SHH* for example: some of the intra-familial variability observed for *SHH*-induced holoprosencephaly has been linked to a second pathogenic variation in another causative gene (Nanni et al., 1999). Animal models support this hypothesis of a second hit, with a more severe phenotype when there are additive variants in partner or target genes (for example, otocephaly and *otx2* in zebrafish, (Chassaing et al., 2012) or double mutations in *MECP2* and *CDKL5* in Rett-like syndrome (Jdila et al., 2020)). In contrast, TF binding variations induced by rare SNPs in TFs binding sites have been recently described, especially in common diseases such as thalassemia (Cheng et al., 2020; Savinkova et al., 2013). Despite growing knowledge in this field, *in silico* predictions of variations in TF binding sites remain challenging and require experimental validation (Moradifard, Saghiri, Ehsani, Mirkhani, & Ebrahimi-Rad, 2020). Finally, variations in sequence or level of expression of miRNAs targeting PBX1 could partly explain both intra- and inter-individual variability by modulating the impact of *PBX1* variations from one tissue (or individual) to another. Recent studies have linked the roles of miRNAs to PBX1 function and degradation in both development processes and cancer (Cao et al., 2022; N. Liu et al., 2020).

Shedding light on these mechanisms would be of help to better understand PBX1 function and regulation, and also to refine genetic counselling.

5) Concluding remarks

First discovered in cancer, PBX proteins, and especially PBX1, have proved to be essential in organism development, from the early gastrulation stages to adulthood. As a transcription cofactor, PBX1 interacts with numerous proteins, including those of the HOX family, and reaches dozens of regulating sequences, even in the heterochromatin state, suggesting that PBX proteins could act as pioneer factors. This

.

pleiotropy explains the polymalformative phenotypes observed in both *Pbx1*-deficient mice and patients carrying heterozygous *PBX1* variations. By altering PBX1 cellular localization and its interaction with its partners or DNA binding, these variations impact the wide network under control of PBX1. However, despite a growing number of studies about this particular family of PBC/MEINOX proteins, questions are still emerging: the mechanisms behind the phenotypic variability among *PBX1*-mutated patients, the discrepancies between human phenotypes and animal models (such as sexual development anomalies), the whole landscape of PBX protein redundancy, or the key elements developed by organisms to compensate for PBX1 deficiency. The end of the TALE has not yet been written.

Acknowledgements

Figures 2 and 3 were created with the help of BioRender.com tool.

Competing interests

There are no competing interests for any author.

Ethics approval statement

Ethics approval was not sought for the present study because it only includes previously published data from human and animal subjects.

Authorship

Laura Mary: Investigation, Writing - Original Draft, Delphine Leclerc: Review & Editing, David Gilot, Marc-Antoine Belaud-Rotureau, Sylvie Jaillard: Supervision, Writing - Review & Editing.

Data availability Statement

The data that support the findings of this study is openly available on ClinVar at https://www.ncbi.nlm.nih.gov/clinvar/ or DECIPHER at

.

Accepted Article

https://www.deciphergenomics.org/ (reference numbers are listed in Table 2) or can be found in the publications listed in the Reference section.

REFERENCES

Abu-Shaar, M., Ryoo, H. D., & Mann, R. S. (1999). Control of the nuclear localization of Extradenticle by competing nuclear import and export signals. *Genes & Development*, *13*(8), 935-945. doi: 10.1101/gad.13.8.935

Alankarage, D., Szot, J. O., Pachter, N., Slavotinek, A., Selleri, L., Shieh, J. T., … Dunwoodie, S. L. (2020). Functional characterization of a novel PBX1 de novo missense variant identified in a patient with syndromic congenital heart disease. *Human Molecular Genetics*, *29*(7), 1068-1082. doi: 10.1093/hmg/ddz231

Alsmadi, O., John, S. E., Thareja, G., Hebbar, P., Antony, D., Behbehani, K., & Thanaraj, T. A. (2014). Genome at Juncture of Early Human Migration: A Systematic Analysis of Two Whole Genomes and Thirteen Exomes from Kuwaiti Population Subgroup of Inferred Saudi Arabian Tribe Ancestry. *PLoS ONE*, *9*(6). doi: 10.1371/journal.pone.0099069

Arts, P., Garland, J., Byrne, A. B., Hardy, T. S. E., Babic, M., Feng, J., … Scott, H. S. (2020). Paternal mosaicism for a novel PBX1 mutation associated with recurrent perinatal death: Phenotypic expansion of the PBX1-related syndrome. *American Journal of Medical Genetics. Part A*, *182*(5), 1273-1277. doi: 10.1002/ajmg.a.61541

Asahara, H., Dutta, S., Kao, H. Y., Evans, R. M., & Montminy, M. (1999). Pbx-Hox heterodimers recruit coactivator-corepressor complexes in an isoform-specific manner. *Molecular and Cellular Biology*, *19*(12), 8219-8225. doi: 10.1128/mcb.19.12.8219

Audano, P. A., Sulovari, A., Graves-Lindsay, T. A., Cantsilieris, S., Sorensen, M., Welch, A. E., … Eichler, E. E. (2019). Characterizing the Major Structural Variant Alleles of the Human Genome. *Cell*, *176*(3), 663-675.e19. doi: 10.1016/j.cell.2018.12.019

Accepted Article

.

Accepted Article

Berenguer, M., & Duester, G. (2021). Role of Retinoic Acid Signaling, FGF Signaling and Meis Genes in Control of Limb Development. *Biomolecules*, *11*(1), 80. doi: 10.3390/biom11010080

Berry, F. B., O'Neill, M. A., Coca-Prados, M., & Walter, M. A. (2005). FOXC1 Transcriptional Regulatory Activity Is Impaired by PBX1 in a Filamin A-Mediated Manner. *Molecular and Cellular Biology*, *25*(4), 1415-1424. doi: 10.1128/MCB.25.4.1415-1424.2005

Berthelsen, J., Kilstrup-Nielsen, C., Blasi, F., Mavilio, F., & Zappavigna, V. (1999). The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes & Development*, *13*(8), 946-953. doi: 10.1101/gad.13.8.946

Bertoli-Avella, A. M., Beetz, C., Ameziane, N., Rocha, M. E., Guatibonza, P., Pereira, C., … Bauer, P. (2021). Successful application of genome sequencing in a diagnostic setting: 1007 index cases from a clinically heterogeneous cohort. *European Journal of Human Genetics: EJHG*, *29*(1), 141-153. doi: 10.1038/s41431-020-00713-9

Bertolino, E., Reimund, B., Wildt-Perinic, D., & Clerc, R. G. (1995). A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *The Journal of Biological Chemistry*, *270*(52), 31178-31188. doi: 10.1074/jbc.270.52.31178

Blasi, F., Bruckmann, C., Penkov, D., & Dardaei, L. (2017). A tale of TALE, PREP1, PBX1, and MEIS1: Interconnections and competition in cancer. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *39*(5). doi: 10.1002/bies.201600245

Brendolan, A., Ferretti, E., Salsi, V., Moses, K., Quaggin, S., Blasi, F., … Selleri, L. (2005). A Pbx1-dependent genetic and transcriptional network regulates spleen ontogeny. *Development (Cambridge, England)*, *132*(13), 3113-3126. doi: 10.1242/dev.01884

Bruckmann, C., Tamburri, S., De Lorenzi, V., Doti, N., Monti, A., Mathiasen, L., … Blasi, F. (2020). Mapping the native interaction surfaces of PREP1 with PBX1 by cross-

.

Accepted Article

linking mass-spectrometry and mutagenesis. *Scientific Reports*, *10*(1), 16809. doi: 10.1038/s41598-020-74032-w

Burel, A., Mouchel, T., Odent, S., Tiker, F., Knebelmann, B., Pellerin, I., & Guerrier, D. (2006). Role of HOXA7 to HOXA13 and PBX1 genes in various forms of MRKH syndrome (congenital absence of uterus and vagina). *Journal of Negative Results in Biomedicine*, *5*, 4. doi: 10.1186/1477-5751-5-4

Calvo, K. R., Knoepfler, P., McGrath, S., & Kamps, M. P. (1999). An inhibitory switch derepressed by pbx, hox, and Meis/Prep1 partners regulates DNA-binding by pbx1 and E2a-pbx1 and is dispensable for myeloid immortalization by E2a-pbx1. *Oncogene*, *18*(56), 8033-8043. doi: 10.1038/sj.onc.1203377

Cao, M., Tian, K., Sun, W., Xu, J., Tang, Y., & Wu, S. (2022). MicroRNA-141-3p inhibits the progression of oral squamous cell carcinoma via targeting PBX1 through the JAK2/STAT3 pathway. *Experimental and Therapeutic Medicine*, *23*(1), 97. doi: 10.3892/etm.2021.11020

Capdevila, J., Tsukui, T., Rodríquez Esteban, C., Zappavigna, V., & Izpisúa Belmonte, J. C. (1999). Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. *Molecular Cell*, *4*(5), 839-849. doi: 10.1016/s1097-2765(00)80393-7

Capellini, T. D., Di Giacomo, G., Salsi, V., Brendolan, A., Ferretti, E., Srivastava, D., … Selleri, L. (2006). Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. *Development (Cambridge, England)*, *133*(11), 2263-2273. doi: 10.1242/dev.02395

Capellini, T. D., Vaccari, G., Ferretti, E., Fantini, S., He, M., Pellegrini, M., … Zappavigna, V. (2010). Scapula development is governed by genetic interactions of Pbx1 with its family members and with Emx2 via their cooperative control of Alx1. *Development (Cambridge, England)*, *137*(15), 2559-2569. doi: 10.1242/dev.048819

Capellini, T. D., Zewdu, R., Di Giacomo, G., Asciutti, S., Kugler, J. E., Di Gregorio, A., & Selleri, L. (2008). Pbx1/Pbx2 govern axial skeletal development by controlling

.

Accepted Article

Polycomb and Hox in mesoderm and Pax1/Pax9 in sclerotome. *Developmental Biology*, *321*(2), 500-514. doi: 10.1016/j.ydbio.2008.04.005

Chassaing, N., Sorrentino, S., Davis, E. E., Martin-Coignard, D., Iacovelli, A., Paznekas, W., … Jabs, E. W. (2012). OTX2 mutations contribute to the otocephaly-dysgnathia complex. *Journal of Medical Genetics*, *49*(6), 373-379. doi: 10.1136/jmedgenet-2012-100892

Chen, N., Zhao, S., Jolly, A., Wang, L., Pan, H., Yuan, J., … Zhu, L. (2021). Perturbations of genes essential for Müllerian duct and Wölffian duct development in Mayer-Rokitansky-Küster-Hauser syndrome. *American Journal of Human Genetics*, *108*(2), 337-345. doi: 10.1016/j.ajhg.2020.12.014

Cheng, Y., Sun, Y., Ji, Y., Jiang, D., Teng, G., Zhou, X., … Xu, C. (2020). Novel compound variants of the AR and MAP3K1 genes are related to the clinical heterogeneity of androgen insensitivity syndrome. *Bioscience Reports*, *40*(5), BSR20200616. doi: 10.1042/BSR20200616

Choe, S.-K., Ladam, F., & Sagerström, C. G. (2014). TALE factors poise promoters for activation by Hox proteins. *Developmental Cell*, *28*(2), 203-211. doi: 10.1016/j.devcel.2013.12.011

Choe, S.-K., Lu, P., Nakamura, M., Lee, J., & Sagerström, C. G. (2009). Meis cofactors control HDAC and CBP accessibility at Hox-regulated promoters during zebrafish embryogenesis. *Developmental cell*, *17*(4), 561-567. doi: 10.1016/j.devcel.2009.08.007

Cohen, A. S. A., Simotas, C., Webb, B. D., Shi, H., Khan, W. A., Edelmann, L., … Singh, R. (2020). Haploinsufficiency of the basic helix-loop-helix transcription factor HAND2 causes congenital heart defects. *American Journal of Medical Genetics. Part A*, *182*(5), 1263-1267. doi: 10.1002/ajmg.a.61537

Couly, G., Grapin-Botton, A., Coltey, P., & Le Douarin, N. M. (1996). The regeneration of the cephalic neural crest, a problem revisited: The regenerating cells originate from

.

Accepted Article

the contralateral or from the anterior and posterior neural fold. *Development (Cambridge, England)*, *122*(11), 3393-3407.

Cuda, C. M., Li, S., Liang, S., Yin, Y., Potula, H. H. S. K., Xu, Z., … Morel, L. (2012). Pre-B cell leukemia homeobox 1 is associated with lupus susceptibility in mice and humans. *Journal of Immunology (Baltimore, Md.: 1950)*, *188*(2), 604-614. doi: 10.4049/jimmunol.1002362

Dard, A., Jia, Y., Reboulet, J., Bleicher, F., Lavau, C., & Merabet, S. (2019). The human HOXA9 protein uses paralog-specific residues of the homeodomain to interact with TALE-class cofactors. *Scientific Reports*, *9*(1), 5664. doi: 10.1038/s41598-019-42096-y

Diakos, C., Xiao, Y., Zheng, S., Kager, L., Dworzak, M., & Wiemels, J. L. (2014). Direct and Indirect Targets of the E2A-PBX1 Leukemia-Specific Fusion Protein. *PLoS ONE*, *9*(2), e87602. doi: 10.1371/journal.pone.0087602

DiMartino, J. F., Selleri, L., Traver, D., Firpo, M. T., Rhee, J., Warnke, R., … Cleary, M. L. (2001). The Hox cofactor and proto-oncogene Pbx1 is required for maintenance of definitive hematopoiesis in the fetal liver. *Blood*, *98*(3), 618-626. doi: 10.1182/blood.v98.3.618

Di-Poï, N., Zákány, J., & Duboule, D. (2007). Distinct Roles and Regulations for Hoxd Genes in Metanephric Kidney Development. *PLoS Genetics*, *3*(12). doi: 10.1371/journal.pgen.0030232

Douglas, G., Cho, M. T., Telegrafi, A., Winter, S., Carmichael, J., Zackai, E. H., … Juusola, J. (2018). De novo missense variants in MEIS2 recapitulate the microdeletion phenotype of cardiac and palate abnormalities, developmental delay, intellectual disability and dysmorphic features. *American Journal of Medical Genetics. Part A*, *176*(9), 1845-1851. doi: 10.1002/ajmg.a.40368

Dunlavy, D. M., O'Leary, D. P., Klimov, D., & Thirumalai, D. (2005). HOPE: A homotopy optimization method for protein structure prediction. *Journal of Computational*

.

Accepted Article

Biology: A Journal of Computational Molecular Cell Biology, *12*(10), 1275-1288. doi: 10.1089/cmb.2005.12.1275

Eozenou, C., Bashamboo, A., Bignon-Topalovic, J., Merel, T., Zwermann, O., Lourenco, D., … Brauner, R. (2019). The TALE homeodomain of PBX1 is involved in human primary testis-determination. *Human Mutation*, *40*(8), 1071-1076. doi: 10.1002/humu.23780

Eyal, S., Kult, S., Rubin, S., Krief, S., Felsenthal, N., Pineault, K. M., … Zelzer, E. (2019). Bone morphology is regulated modularly by global and regional genetic programs. *Development (Cambridge, England)*, *146*(14). doi: 10.1242/dev.167882

Fagerberg, L., Hallström, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., … Uhlén, M. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular & Cellular Proteomics: MCP*, *13*(2), 397-406. doi: 10.1074/mcp.M113.035600

Farber, P. J., & Mittermaier, A. (2011). Concerted dynamics link allosteric sites in the PBX homeodomain. *Journal of Molecular Biology*, *405*(3), 819-830. doi: 10.1016/j.jmb.2010.11.016

Ferretti, E., Li, B., Zewdu, R., Wells, V., Hebert, J. M., Karner, C., … Selleri, L. (2011). A conserved Pbx-Wnt-p63-Irf6 regulatory module controls face morphogenesis by promoting epithelial apoptosis. *Developmental Cell*, *21*(4), 627-641. doi: 10.1016/j.devcel.2011.08.005

Fitzgerald, K. K., Powell-Hamilton, N., Shillingford, A. J., Robinson, B., & Gripp, K. W. (2021). Inherited intragenic PBX1 deletion: Expanding the phenotype. *American Journal of Medical Genetics. Part A*, *185*(1), 234-237. doi: 10.1002/ajmg.a.61932

Fognani, C., Kilstrup-Nielsen, C., Berthelsen, J., Ferretti, E., Zappavigna, V., & Blasi, F. (2002). Characterization of PREP2, a paralog of PREP1, which defines a novel sub-family of the MEINOX TALE homeodomain transcription factors. *Nucleic Acids Research*, *30*(9), 2043-2051.

.

Accepted Article

Fromer, M., Pocklington, A. J., Kavanagh, D. H., Williams, H. J., Dwyer, S., Gormley, P., … O'Donovan, M. C. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature, 506*(7487), 179-184. doi: 10.1038/nature12929

Fujita, A., Isidor, B., Piloquet, H., Corre, P., Okamoto, N., Nakashima, M., … Matsumoto, N. (2016). De novo MEIS2 mutation causes syndromic developmental delay with persistent gastro-esophageal reflux. *Journal of Human Genetics, 61*(9), 835-838. doi: 10.1038/jhg.2016.54

Fukuda, S., Hoggatt, J., Singh, P., Abe, M., Speth, J. M., Hu, P., … Pelus, L. M. (2015). Survivin modulates genes with divergent molecular functions and regulates proliferation of hematopoietic stem cells through Evi-1. *Leukemia, 29*(2), 433-440. doi: 10.1038/leu.2014.183

Gemel, J., Jacobsen, C., & MacArthur, C. A. (1999). Fibroblast growth factor-8 expression is regulated by intronic engrailed and Pbx1-binding sites. *The Journal of Biological Chemistry, 274*(9), 6020-6026. doi: 10.1074/jbc.274.9.6020

Golonzhka, O., Nord, A., Tang, P. L., Lindtner, S., Ypsilanti, A. R., Ferretti, E., … Rubenstein, J. L. R. (2015). Pbx regulates patterning of the cerebral cortex in progenitors and postmitotic neurons. *Neuron, 88*(6), 1192-1207. doi: 10.1016/j.neuron.2015.10.045

Gordon, J. A. R., Hassan, M. Q., Saini, S., Montecino, M., van Wijnen, A. J., Stein, G. S., … Lian, J. B. (2010). Pbx1 represses osteoblastogenesis by blocking Hoxa10-mediated recruitment of chromatin remodeling factors. *Molecular and Cellular Biology, 30*(14), 3531-3541. doi: 10.1128/MCB.00889-09

Grebbin, Britta M., & Schulte, D. (2017). PBX1 as Pioneer Factor: A Case Still Open. *Frontiers in Cell and Developmental Biology, 5*, 9. doi: 10.3389/fcell.2017.00009

Grebbin, Britta Moyo, Hau, A.-C., Groß, A., Anders-Maurer, M., Schramm, J., Koss, M., … Schulte, D. (2016). Pbx1 is required for adult subventricular zone neurogenesis. *Development (Cambridge, England), 143*(13), 2281-2291. doi: 10.1242/dev.128033

.

Accepted Article

Haller, K., Rambaldi, I., Kovács, E. N., Daniels, E., & Featherstone, M. (2002). Prep2: Cloning and expression of a new prep family member. *Developmental Dynamics*, *225*(3), 358-364. doi: 10.1002/dvdy.10167

Hatzistergos, K. E., Williams, A. R., Dykxhoorn, D., Bellio, M. A., Yu, W., & Hare, J. M. (2019). Tumor Suppressors RB1 and CDKN2a Cooperatively Regulate Cell-Cycle Progression and Differentiation During Cardiomyocyte Development and Repair. *Circulation Research*, *124*(8), 1184-1197. doi: 10.1161/CIRCRESAHA.118.314063

He, C., Wang, Z., Zhang, L., Yang, L., Li, J., Chen, X., … Zhu, Z. (2017). A hydrophobic residue in the TALE homeodomain of PBX1 promotes epithelial-to-mesenchymal transition of gastric carcinoma. *Oncotarget*, *8*(29), 46818-46833. doi: 10.18632/oncotarget.17473

Heidet, L., Morinière, V., Henry, C., De Tomasi, L., Reilly, M. L., Humbert, C., … Jeanpierre, C. (2017). Targeted Exome Sequencing Identifies PBX1 as Involved in Monogenic Congenital Anomalies of the Kidney and Urinary Tract. *Journal of the American Society of Nephrology: JASN*, *28*(10), 2901-2914. doi: 10.1681/ASN.2017010043

Holley, M., Rhodes, C., Kneebone, A., Herde, M. K., Fleming, M., & Steel, K. P. (2010). Emx2 and early hair cell development in the mouse inner ear. *Developmental Biology*, *340*(2), 547-556. doi: 10.1016/j.ydbio.2010.02.004

Huang, H., Paliouras, M., Rambaldi, I., Lasko, P., & Featherstone, M. (2003). Nonmuscle Myosin Promotes Cytoplasmic Localization of PBX. *Molecular and Cellular Biology*, *23*(10), 3636-3645. doi: 10.1128/MCB.23.10.3636-3645.2003

Hurtado, R., Zewdu, R., Mtui, J., Liang, C., Aho, R., Kurylo, C., … Herzlinger, D. (2015). Pbx1-dependent control of VMC differentiation kinetics underlies gross renal vascular patterning. *Development (Cambridge, England)*, *142*(15), 2653-2664. doi: 10.1242/dev.124776

Jaw, T. J., You, L. R., Knoepfler, P. S., Yao, L. C., Pai, C. Y., Tang, C. Y., … Sun, Y. H. (2000). Direct interaction of two homeoproteins, homothorax and extradenticle, is

.

Accepted Article

essential for EXD nuclear localization and function. *Mechanisms of Development*, *91*(1-2), 279-291. doi: 10.1016/s0925-4773(99)00316-0

Jdila, M. B., Triki, C. C., Ghorbel, R., Bouchalla, W., Ncir, S. B., Kamoun, F., & Fakhfakh, F. (2020). Unusual double mutation in MECP2 and CDKL5 genes in Rett-like syndrome: Correlation with phenotype and genes expression. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, *508*, 287-294. doi: 10.1016/j.cca.2020.05.037

Jeannotte, L., Gotti, F., & Landry-Truchon, K. (2016). Hoxa5: A Key Player in Development and Disease. *Journal of Developmental Biology*, *4*(2), E13. doi: 10.3390/jdb4020013

Jozwik, K. M., & Carroll, J. S. (2012). Pioneer factors in hormone-dependent cancers. *Nature Reviews. Cancer*, *12*(6), 381-385. doi: 10.1038/nrc3263

Jung, J.-G., Shih, I.-M., Park, J. T., Gerry, E., Kim, T. H., Ayhan, A., … Wang, T.-L. (2016). Ovarian Cancer Chemoresistance Relies on the Stem Cell Reprogramming Factor PBX1. *Cancer Research*, *76*(21), 6351-6361. doi: 10.1158/0008-5472.CAN-16-0980

Kamps, M. P., Look, A. T., & Baltimore, D. (1991). The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials. *Genes & Development*, *5*(3), 358-368. doi: 10.1101/gad.5.3.358

Kawahara, Y. (2014). Human diseases caused by germline and somatic abnormalities in microRNA and microRNA-related genes. *Congenital Anomalies*, *54*(1), 12-21. doi: 10.1111/cga.12043

Kia, F., Sarafoglou, K., Mooganayakanakote Siddappa, A., & Roberts, K. D. (2019). Partial gonadal dysgenesis associated with a pathogenic variant of PBX1 transcription factor. *BMJ Case Reports*, *12*(7). doi: 10.1136/bcr-2018-227986

Kilstrup-Nielsen, C., Alessio, M., & Zappavigna, V. (2003). PBX1 nuclear export is regulated independently of PBX-MEINOX interaction by PKA phosphorylation of the PBC-B domain. *The EMBO Journal*, *22*(1), 89-99. doi: 10.1093/emboj/cdg010

.

Accepted Article

Kocabas, F., Xie, L., Xie, J., Yu, Z., DeBerardinis, R. J., Kimura, W., … Zheng, J. (2015). Hypoxic metabolism in human hematopoietic stem cells. *Cell & Bioscience, 5*, 39. doi: 10.1186/s13578-015-0020-3

Kosho, T., Okamoto, N., & Coffin-Siris Syndrome International Collaborators. (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics, 166C*(3), 262-275. doi: 10.1002/ajmg.c.31407

Ladam, F., Stanney, W., Donaldson, I. J., Yildiz, O., Bobola, N., & Sagerström, C. G. (2018). TALE factors use two distinct functional modes to control an essential zebrafish gene expression program. *ELife, 7*. doi: 10.7554/eLife.36144

Lampe, X., Samad, O. A., Guiguen, A., Matis, C., Remacle, S., Picard, J. J., … Rezsohazy, R. (2008). An ultraconserved Hox-Pbx responsive element resides in the coding sequence of Hoxa2 and is active in rhombomere 4. *Nucleic Acids Research, 36*(10), 3214-3225. doi: 10.1093/nar/gkn148

Laurent, A., Bihan, R., Omilli, F., Deschamps, S., & Pellerin, I. (2008). PBX proteins: Much more than Hox cofactors. *The International Journal of Developmental Biology, 52*(1), 9-20. doi: 10.1387/ijdb.072304al

Le Tanno, P., Breton, J., Bidart, M., Satre, V., Harbuz, R., Ray, P. F., … Coutton, C. (2017). PBX1 haploinsufficiency leads to syndromic congenital anomalies of the kidney and urinary tract (CAKUTHED) in humans. *Journal of Medical Genetics, 54*(7), 502-510. doi: 10.1136/jmedgenet-2016-104435

Lescroart, F., & Zaffran, S. (2018). Hox and Tale transcription factors in heart development and disease. *The International Journal of Developmental Biology, 62*(11-12), 837-846. doi: 10.1387/ijdb.180192sz

Li, W., Lin, C.-Y., Shang, C., Han, P., Xiong, Y., Lin, C.-J., … Chang, C.-P. (2014). Pbx1 activates Fgf10 in the mesenchyme of developing lungs. *Genesis (New York, N.Y.: 2000), 52*(5), 399-407. doi: 10.1002/dvg.22764

.

Accepted Article

Lima Cunha, D., Arno, G., Corton, M., & Moosajee, M. (2019). The Spectrum of PAX6 Mutations and Genotype-Phenotype Correlations in the Eye. *Genes*, *10*(12), E1050. doi: 10.3390/genes10121050

Lin, C.-H., Wang, Z., Duque-Afonso, J., Wong, S. H.-K., Demeter, J., Loktev, A. V., … Cleary, M. L. (2019). Oligomeric self-association contributes to E2A-PBX1-mediated oncogenesis. *Scientific Reports*, *9*(1), 4915. doi: 10.1038/s41598-019-41393-w

Lin, Y.-J., Kao, C.-H., Hsiao, S.-P., & Chen, S.-L. (2021). The cooperation of cis-elements during M-cadherin promoter activation. *The Biochemical Journal*, *478*(4), 911-926. doi: 10.1042/BCJ20200535

Liu, J., Wang, Y., Birnbaum, M. J., & Stoffers, D. A. (2010). Three-amino-acid-loop-extension homeodomain factor Meis3 regulates cell survival via PDK1. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(47), 20494-20499. doi: 10.1073/pnas.1007001107

Liu, N., Zhang, Z., Li, L., Shen, X., Sun, B., Wang, R., … Yuan, W. (2020). MicroRNA-181 regulates the development of Ossification of Posterior longitudinal ligament via Epigenetic Modulation by targeting PBX1. *Theranostics*, *10*(17), 7492-7509. doi: 10.7150/thno.44309

Longobardi, E., Penkov, D., Mateos, D., De Florian, G., Torres, M., & Blasi, F. (2014). Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *243*(1), 59-75. doi: 10.1002/dvdy.24016

Lüdtke, T. H., Wojahn, I., Kleppa, M.-J., Schierstaedt, J., Christoffels, V. M., Künzler, P., & Kispert, A. (2021). Combined genomic and proteomic approaches reveal DNA binding sites and interaction partners of TBX2 in the developing lung. *Respiratory Research*, *22*(1), 85. doi: 10.1186/s12931-021-01679-y

Ma, W., Li, Y., Wang, M., Li, H., Su, T., Li, Y., & Wang, S. (2015). Associations of Polymorphisms in WNT9B and PBX1 with Mayer-Rokitansky-Küster-Hauser

.

Accepted Article

Syndrome in Chinese Han. *PloS One*, *10*(6), e0130202. doi: 10.1371/journal.pone.0130202

Magnani, L., Patten, D. K., Nguyen, V. T. M., Hong, S.-P., Steel, J. H., Patel, N., … Coombes, R. C. (2015). The pioneer factor PBX1 is a novel driver of metastatic progression in ERα-positive breast cancer. *Oncotarget*, *6*(26), 21878-21891. doi: 10.18632/oncotarget.4243

Maili, L., Letra, A., Silva, R., Buchanan, E. P., Mulliken, J. B., Greives, M. R., … Hecht, J. T. (2020). PBX-WNT-P63-IRF6 pathway in nonsyndromic cleft lip and palate. *Birth Defects Research*, *112*(3), 234-244. doi: 10.1002/bdr2.1630

Manley, N. R., Selleri, L., Brendolan, A., Gordon, J., & Cleary, M. L. (2004). Abnormalities of caudal pharyngeal pouch development in Pbx1 knockout mice mimic loss of Hox3 paralogs. *Developmental Biology*, *276*(2), 301-312. doi: 10.1016/j.ydbio.2004.08.030

Marcos, S., González-Lázaro, M., Beccari, L., Carramolino, L., Martin-Bermejo, M. J., Amarie, O., … Bovolenta, P. (2015). Meis1 coordinates a network of genes implicated in eye development and microphthalmia. *Development (Cambridge, England)*, *142*(17), 3009-3020. doi: 10.1242/dev.122176

Mathiasen, L., Valentini, E., Boivin, S., Cattaneo, A., Blasi, F., Svergun, D. I., & Bruckmann, C. (2016). The flexibility of a homeodomain transcription factor heterodimer and its allosteric regulation by DNA binding. *The FEBS Journal*, *283*(16), 3134-3154. doi: 10.1111/febs.13801

Mauger, D. M., Cabral, B. J., Presnyak, V., Su, S. V., Reid, D. W., Goodman, B., … McFadyen, I. J. (2019). MRNA structure regulates protein expression through changes in functional half-life. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(48), 24075-24083. doi: 10.1073/pnas.1908052116

McCulley, D. J., Wienhold, M. D., Hines, E. A., Hacker, T. A., Rogers, A., Pewowaruk, R. J., … Sun, X. (2018). PBX transcription factors drive pulmonary vascular adaptation to birth. *The Journal of Clinical Investigation*, *128*(2), 655-667. doi: 10.1172/JCI93395

.

Accepted Article

Meng, L., Tian, Z., Wang, J., Liu, X., Zhang, W., Hu, M., … Zhang, Y. (2021). Effect of myeloid ecotropic viral integration site (MEIS) family genes on tumor microenvironment remodeling and its potential therapeutic effect. *Translational Andrology and Urology*, *10*(2), 594-608. doi: 10.21037/tau-20-1163

Merabet, S., & Mann, R. S. (2016). To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. *Trends in Genetics: TIG*, *32*(6), 334-347. doi: 10.1016/j.tig.2016.03.004

Mohammadi, M., Salehzadeh, A., Talesh Sasani, S., & Tarang, A. (2021). Rs6426881 in the 3'-UTR of PBX1 is involved in breast and gastric cancers via altering the binding potential of miR-522-3p. *Molecular Biology Reports*, *48*(11), 7405-7414. doi: 10.1007/s11033-021-06756-5

Moradifard, S., Saghiri, R., Ehsani, P., Mirkhani, F., & Ebrahimi-Rad, M. (2020). A preliminary computational outputs versus experimental results: Application of sTRAP, a biophysical tool for the analysis of SNPs of transcription factor-binding sites. *Molecular Genetics & Genomic Medicine*, *8*(5), e1219. doi: 10.1002/mgg3.1219

Nanni, L., Ming, J. E., Bocian, M., Steinhaus, K., Bianchi, D. W., Die-Smulders, C., … Muenke, M. (1999). The mutational spectrum of the sonic hedgehog gene in holoprosencephaly: SHH mutations cause a significant proportion of autosomal dominant holoprosencephaly. *Human Molecular Genetics*, *8*(13), 2479-2488. doi: 10.1093/hmg/8.13.2479

Neuteboom, S. T., Peltenburg, L. T., van Dijk, M. A., & Murre, C. (1995). The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(20), 9166-9170. doi: 10.1073/pnas.92.20.9166

Niu, Y., Sengupta, M., Titov, A. A., Choi, S.-C., & Morel, L. (2017). The PBX1 lupus susceptibility gene regulates CD44 expression. *Molecular Immunology*, *85*, 148-154. doi: 10.1016/j.molimm.2017.02.016

.

Accepted Article

Patterson, L. T., & Potter, S. S. (2004). Atlas of Hox gene expression in the developing kidney. *Developmental Dynamics*, *229*(4), 771-779. doi: https://doi.org/10.1002/dvdy.10474

Paul, S., Zhang, X., & He, J.-Q. (2020). Homeobox gene Meis1 modulates cardiovascular regeneration. *Seminars in Cell & Developmental Biology*, *100*, 52-61. doi: 10.1016/j.semcdb.2019.10.003

Penkov, D., Mateos San Martín, D., Fernandez-Díaz, L. C., Rosselló, C. A., Torroja, C., Sánchez-Cabo, F., … Torres, M. (2013). Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. *Cell Reports*, *3*(4), 1321-1333. doi: 10.1016/j.celrep.2013.03.029

Pi, W.-C., Wang, J., Shimada, M., Lin, J.-W., Geng, H., Lee, Y.-L., … Chen, W.-Y. (2020). E2A-PBX1 functions as a coactivator for RUNX1 in acute lymphoblastic leukemia. *Blood*, *136*(1), 11-23. doi: 10.1182/blood.2019003312

Quina, L. A., Kuramoto, T., Luquetti, D. V., Cox, T. C., Serikawa, T., & Turner, E. E. (2012). Deletion of a conserved regulatory element required for Hmx1 expression in craniofacial mesenchyme in the dumbo rat: A newly identified cause of congenital ear malformation. *Disease Models & Mechanisms*, *5*(6), 812-822. doi: 10.1242/dmm.009910

Rave-Harel, N., Givens, M. L., Nelson, S. B., Duong, H. A., Coss, D., Clark, M. E., … Mellon, P. L. (2004). TALE Homeodomain Proteins Regulate Gonadotropin-releasing Hormone Gene Expression Independently and via Interactions with Oct-1. *The Journal of biological chemistry*, *279*(29), 30287-30297. doi: 10.1074/jbc.M402960200

Riedhammer, K. M., Siegel, C., Alhaddad, B., Montoya, C., Kovacs-Nagy, R., Wagner, M., … Hoefele, J. (2017). Identification of a Novel Heterozygous De Novo 7-bp Frameshift Deletion in PBX1 by Whole-Exome Sequencing Causing a Multi-Organ

.

Accepted Article

Syndrome Including Bilateral Dysplastic Kidneys and Hypoplastic Clavicles. *Frontiers in Pediatrics*, *5*, 251. doi: 10.3389/fped.2017.00251

Roberts, V. J., van Dijk, M. A., & Murre, C. (1995). Localization of Pbx1 transcripts in developing rat embryos. *Mechanisms of Development*, *51*(2-3), 193-198. doi: 10.1016/0925-4773(95)00364-9

Rosin, J. M., Li, W., Cox, L. L., Rolfe, S. M., Latorre, V., Akiyama, J. A., … Cox, T. C. (2016). A distal 594 bp ECR specifies Hmx1 expression in pinna and lateral facial morphogenesis and is regulated by the Hox-Pbx-Meis complex. *Development (Cambridge, England)*, *143*(14), 2582-2592. doi: 10.1242/dev.133736

Rossetti, L. Z., Glinton, K., Yuan, B., Liu, P., Pillai, N., Mizerik, E., … Scott, D. A. (2019). Review of the phenotypic spectrum associated with haploinsufficiency of MYRF. *American Journal of Medical Genetics. Part A*, *179*(7), 1376-1382. doi: 10.1002/ajmg.a.61182

Russell, M. K., Longoni, M., Wells, J., Maalouf, F. I., Tracy, A. A., Loscertales, M., … Donahoe, P. K. (2012). Congenital diaphragmatic hernia candidate genes derived from embryonic transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(8), 2978-2983. doi: 10.1073/pnas.1121621109

Saleh, M., Huang, H., Green, N. C., & Featherstone, M. S. (2000). A conformational change in PBX1A is necessary for its nuclear localization. *Experimental Cell Research*, *260*(1), 105-115. doi: 10.1006/excr.2000.5010

Satokata, I., Benson, G., & Maas, R. (1995). Sexually dimorphic sterility phenotypes in HoxalO-deficient mice. *Nature*, *374*(6521), 460-463. doi: 10.1038/374460a0

Savinkova, L., Drachkova, I., Arshinova, T., Ponomarenko, P., Ponomarenko, M., & Kolchanov, N. (2013). An Experimental Verification of the Predicted Effects of Promoter TATA-Box Polymorphisms Associated with Human Diseases on Interactions between the TATA Boxes and TATA-Binding Protein. *PLoS ONE*, *8*(2), e54626. doi: 10.1371/journal.pone.0054626

.

Accepted Article

Schnabel, C. A., Selleri, L., Jacobs, Y., Warnke, R., & Cleary, M. L. (2001). Expression of Pbx1b during mammalian organogenesis. *Mechanisms of Development*, *100*(1), 131-135. doi: 10.1016/s0925-4773(00)00516-5

Schnabel, Catherine A., Godin, R. E., & Cleary, M. L. (2003). Pbx1 regulates nephrogenesis and ureteric branching in the developing kidney. *Developmental Biology*, *254*(2), 262-276. doi: 10.1016/s0012-1606(02)00038-6

Schnabel, Catherine A., Selleri, L., & Cleary, M. L. (2003). Pbx1 is essential for adrenal development and urogenital differentiation. *Genesis (New York, N.Y.: 2000)*, *37*(3), 123-130. doi: 10.1002/gene.10235

Seidman, J. G., & Seidman, C. (2002). Transcription factor haploinsufficiency: When half a loaf is not enough. *The Journal of Clinical Investigation*, *109*(4), 451-455. doi: 10.1172/JCI15043

Selleri, L., Depew, M. J., Jacobs, Y., Chanda, S. K., Tsang, K. Y., Cheah, K. S., … Cleary, M. L. (2001). Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. *Development (Cambridge, England)*, *128*(18), 3543-3557.

Selleri, Licia, DiMartino, J., van Deursen, J., Brendolan, A., Sanyal, M., Boon, E., … Cleary, M. L. (2004). The TALE Homeodomain Protein Pbx2 Is Not Essential for Development and Long-Term Survival. *Molecular and Cellular Biology*, *24*(12), 5324-5331. doi: 10.1128/MCB.24.12.5324-5331.2004

Selleri, Licia, Zappavigna, V., & Ferretti, E. (2019).'Building a perfect body': Control of vertebrate organogenesis by PBX-dependent regulatory networks. *Genes & Development*, *33*(5-6), 258-275. doi: 10.1101/gad.318774.118

Sengupta, M., Liang, S., Potula, H.-H., Chang, L.-J., & Morel, L. (2012). The SLE-associated Pbx1-d isoform acts as a dominant-negative transcriptional regulator. *Genes and immunity*, *13*(8), 653-657. doi: 10.1038/gene.2012.43

Slavotinek, A., Risolino, M., Losa, M., Cho, M. T., Monaghan, K. G., Schneidman-Duhovny, D., … Shieh, J. (2017). De novo, deleterious sequence variants that alter the

.

Accepted Article

transcriptional activity of the homeoprotein PBX1 are associated with intellectual disability and pleiotropic developmental defects. *Human Molecular Genetics*, *26*(24), 4849-4860. doi: 10.1093/hmg/ddx363

Sonnet, W., Rezsöhazy, R., & Donnay, I. (2012). Characterization of TALE genes expression during the first lineage segregation in mammalian embryos. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, *241*(11), 1827-1839. doi: 10.1002/dvdy.23873

Stankunas, K., Shang, C., Twu, K. Y., Kao, S.-C., Jenkins, N. A., Copeland, N. G., … Chang, C.-P. (2008). Pbx/Meis deficiencies demonstrate multigenetic origins of congenital heart disease. *Circulation Research*, *103*(7), 702-709. doi: 10.1161/CIRCRESAHA.108.175489

Takahashi, T., Friedmacher, F., Zimmer, J., & Puri, P. (2020). Pbx1, Meis1, and Runx1 Expression Is Decreased in the Diaphragmatic and Pulmonary Mesenchyme of Rats with Nitrofen-Induced Congenital Diaphragmatic Hernia. *European Journal of Pediatric Surgery: Official Journal of Austrian Association of Pediatric Surgery... [et Al] = Zeitschrift Fur Kinderchirurgie*. doi: 10.1055/s-0040-1714736

Thareja, G., John, S. E., Hebbar, P., Behbehani, K., Thanaraj, T. A., & Alsmadi, O. (2015). Sequence and analysis of a whole genome from Kuwaiti population subgroup of Persian ancestry. *BMC Genomics*, *16*(1). doi: 10.1186/s12864-015-1233-x

Thiaville, M. M., Stoeck, A., Chen, L., Wu, R.-C., Magnani, L., Oidtman, J., … Wang, T.-L. (2012). Identification of PBX1 Target Genes in Cancer Cells by Global Mapping of PBX1 Binding Sites. *PLoS ONE*, *7*(5). doi: 10.1371/journal.pone.0036054

Tremblay, J. J., & Viger, R. S. (2003). A mutated form of steroidogenic factor 1 (SF-1 G35E) that causes sex reversal in humans fails to synergize with transcription factor GATA-4. *The Journal of Biological Chemistry*, *278*(43), 42637-42642. doi: 10.1074/jbc.M305485200

Uddin, M., Thiruvahindrapuram, B., Walker, S., Wang, Z., Hu, P., Lamoureux, S., … Scherer, S. W. (2015). A High-Resolution Copy Number Variation Resource for Clinical and

.

Accepted Article

Population Genetics. *Genetics in medicine: official journal of the American College of Medical Genetics*, *17*(9), 747-752. doi: 10.1038/gim.2014.178

Venselaar, H., Te Beek, T. A. H., Kuipers, R. K. P., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*, *11*, 548. doi: 10.1186/1471-2105-11-548

Villaescusa, J. C., Li, B., Toledo, E. M., Rivetti di Val Cervo, P., Yang, S., Stott, S. R., … Arenas, E. (2016). A PBX1 transcriptional network controls dopaminergic neuron development and is impaired in Parkinson's disease. *The EMBO Journal*, *35*(18), 1963-1978. doi: 10.15252/embj.201593725

Vitobello, A., Ferretti, E., Lampe, X., Vilain, N., Ducret, S., Ori, M., … Rijli, F. M. (2011). Hox and Pbx factors control retinoic acid synthesis during hindbrain segmentation. *Developmental Cell*, *20*(4), 469-482. doi: 10.1016/j.devcel.2011.03.011

Wei, X., Yu, L., & Li, Y. (2018). PBX1 promotes the cell proliferation via JAK2/STAT3 signaling in clear cell renal carcinoma. *Biochemical and Biophysical Research Communications*, *500*(3), 650-657. doi: 10.1016/j.bbrc.2018.04.127

Wong, L.-P., Ong, R. T.-H., Poh, W.-T., Liu, X., Chen, P., Li, R., … Teo, Y.-Y. (2013). Deep Whole-Genome Sequencing of 100 Southeast Asian Malays. *American Journal of Human Genetics*, *92*(1), 52-66. doi: 10.1016/j.ajhg.2012.12.005

Zhang, X., Rowan, S., Yue, Y., Heaney, S., Pan, Y., Brendolan, A., … Maas, R. L. (2006). Pax6 is regulated by Meis and Pbx homeoproteins during pancreatic development. *Developmental Biology*, *300*(2), 748-757. doi: 10.1016/j.ydbio.2006.06.030

Zhou, W., Zhu, H., Zhao, J., Li, H., Wan, Y., Cao, J., … Guo, X. (2013). Misexpression of Pknox2 in mouse limb bud mesenchyme perturbs zeugopod development and deltoid crest formation. *PloS One*, *8*(5), e64237. doi: 10.1371/journal.pone.0064237

.

Accepted Article

Figure Legends

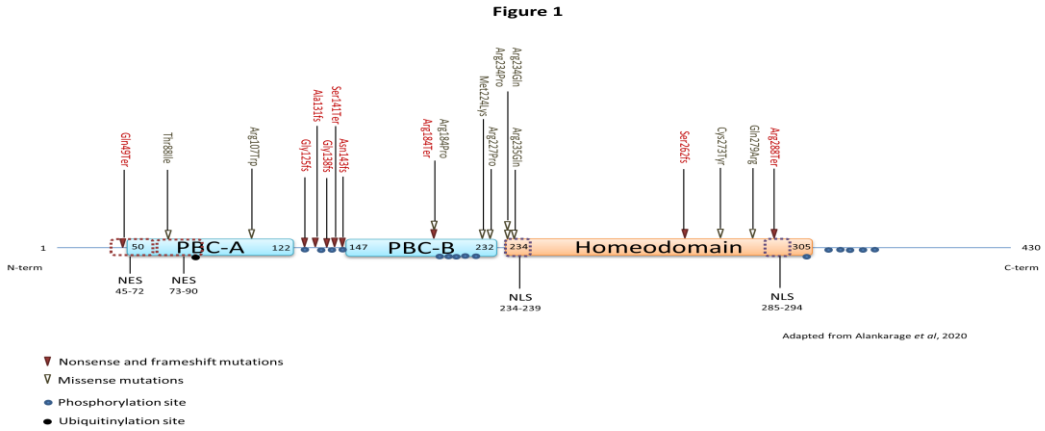

Figure 1: Diagram of the linear structure of PBX1 longest isoform (NP_002576.1). The numbers indicate the amino-acid position by which each functional domain begins and ends. The plain red triangles represent the localization of nonsense variations and the brown triangles represent missense variations. The blue and black dots represent phosphorylation and ubiquitinylation sites respectively. N-term, C-term: N- and C-termini of the protein. NES: Nuclear Export Signal, NLS: Nuclear Localization Signal.

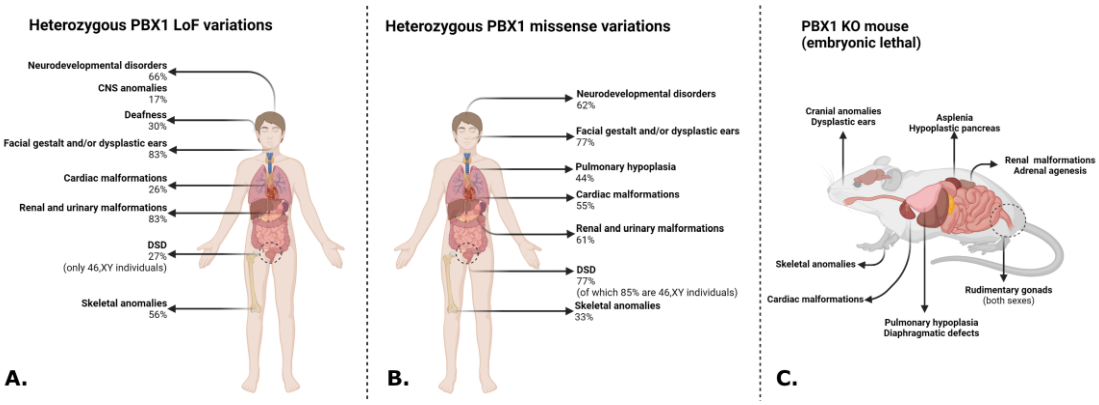

Figure 2: Comparison of the human phenotypes generated by PBX1 heterozygous LoF (A) or missense (B) variants with the phenotype observed in PBX1-null mice (C). DSD: Disorders of Sexual Development.

.

Accepted Article

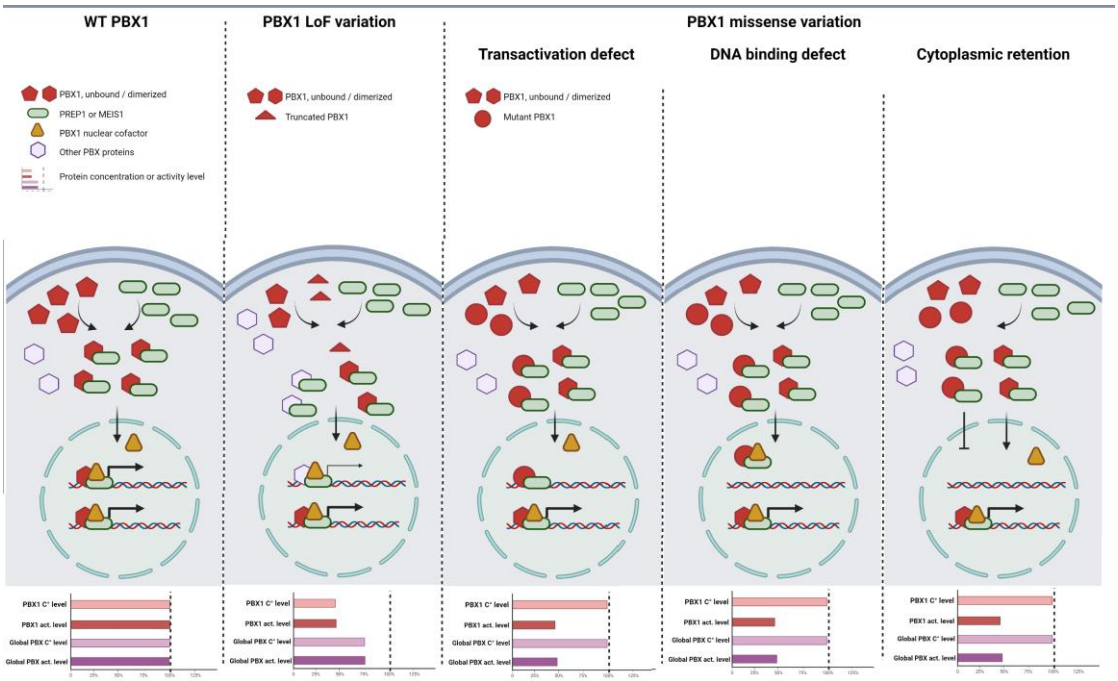

Figure 3: Diagrammatic representation of the hypothesized molecular consequences of PBX1 variations. C°: concentration.

Table 1: Summary of known PBX1 cofactors and target genes in development.

Protein	Cofactor or Target	Interaction with PBX1	Interaction consequence	Cell or tissue type studied	Source
PBX1 cofactor / enhancer					
N CoR/SMRT	Cofactor	Recruited by PBX1	Repression of target. Sensitive to PBX1 isoform (PBX1a)	(HEK) 293T cells	(Asahara, Dutta, Kao, Evans, & Montminy, 1999)
PDX1	Cofactor	Forms heterodimers with PBX1	Activation / repression of target		
CREB binding protein	Cofactor	Recruited by PBX1	Activation of target. Sensitive to PBX1 isoform (PBX1b)		
FLNA	Cofactor	Involved in proper subnuclear localization of PBX1 through phosphorylation	Inhibition of PBX1 export to the nucleus	A7 melanoma cells	(Berry, O'Neill, Coca-Prados, & Walter, 2005)
FoxC1	Cofactor	Recruited by PBX1	Cooperative binding of DNA		
TLX1	Cofactor	Forms heterodimers with PBX1 in spleen	Splenic formation and progenitor proliferation	Splenic progenitors	(Brendolan et al., 2005)
Emx2	Cofactor	Recruited by PBX1	Scapula blade patterning	Dermomyotome	(Capellini et al., 2010)
HoxA7	Cofactor	Binds to PBX1-Meis complex	Trimer formation under control of DNA topology and proteic environment	HEK, HeLa and MCF7 cells	(Dard et al., 2019)
HoxB3	Cofactor	Binds to PBX1-Meis complex			
HoxC8	Cofactor	Binds to PBX1-Meis complex			
HoxA9	Cofactor	Binds to PBX1-Meis complex	Trimer formation insensitive to DNA topology and proteic environment		
HoxA10	Cofactor	Forms heterodimers with PBX1	Repress HOXA10 target genes activation	Osteoblasts and C3H10T1/2 and MC3T3-E1 cells	(Gordon et al., 2010)
NMHCB	Cofactor	Downregulation of PBX1	Compete with Meis1 for PBX1 dimerization, retains PBX1 in the	HEK293T, Cos-7 cells and	(Huang, Paliouras, Rambaldi, Lasko, &

.

Accepted Article

			cytoplasm	Drosophila S2 cells	Featherstone, 2003)
AP2γ	Cofactor	Putative PBX1 cofactor in ER-dependant cancers	Allows the accessibility to chromatin of ER	ER-dependant cancer cells	(Jozwik & Carroll, 2012)
FOXA1	Cofactor	Putative PBX1 cofactor in ER-dependant cancers		MCF-7 cells	
ZFPIP	Cofactor	Binds to PBX1-HOXA9 complex	Inhibit PBX1-HoxA9 DNA binding	Embryonic female genital tract, head and limb bud	(Laurent, Bihan, Omilli, Deschamps, & Pellerin, 2008)
HDAC, HAT	Cofactor	Recruited by PBX1	Repression / activation of target		(Laurent et al., 2008; Penkov et al., 2013)
MEIS1	Cofactor	Needed for PBX1 function (binding mutually exclusive with PKNOX)	Targets promoter-remote regions and developmental genes, favors transcriptionnal regulation and tumorigenesis	MEF cells, mouse embryos	(Blasi, Bruckmann, Penkov, & Dardaei, 2017; Laurent et al., 2008; Penkov et al., 2013)
MEIS2	Cofactor	Needed for PBX1 function (binding mutually exclusive with PKNOX)	Tumorigenesis	HepG2 cells	(Bjerke, Hyman-Walsh, & Wotton, 2011)
KLF4	Cofactor	Binds to PBX1-Meis2 complex			
SIN3B	Cofactor	Recruited by PBX1	Part of a repressor complex including HDAC1, HDAC3, mSIN3B and N-CoR/SMRT		(Laurent et al., 2008)
PKNOX 1	Cofactor	Needed for PBX1 function (binding mutually exclusive with Meis)	Promoters and house-keeping genes, transcription initiation, tumor suppression	MEF cells, mouse embryos	(Blasi et al., 2017; Laurent et al., 2008; Penkov et al., 2013)
PKNOX 2	Cofactor	Needed for PBX1 function (binding mutually exclusive with Meis)	Binds TALE and HoxA1 promoters in early development and removed after retinoic acid-induced differenciation	Mouse embryos	(De Kumar et al., 2017)
CDX2	Cofactor	Forms heterodimers with PBX1	Dimer enhances the activation of proglucagon	Pancreatic islets	(Liu, Branch, & Jin, 2006)
HoxA5	Cofactor	Forms heterodimers with PBX1	Hox target gene activation	Bl21 transfected cells	(Lu & Kamps, 1996)
HoxB5	Cofactor	Forms heterodimers with PBX1			
HoxB8	Cofactor	Forms heterodimers with PBX1			
RUNX1	Cofactor	Regulated by chimeric E2A-PBX1	Promote cell proliferation and tumorigenesis	pre-B ALL cells	(Pi et al., 2020)
Sp2	Cofactor	Recruited by PBX1-Prep1 + Nf-y	Enhances PBX1-Prep DNA binding on promoters	Mouse embryonic fibroblasts	(Völkel et al., 2018)
PBXIP1	Cofactor	Repress PBX1-Hox DNA binding	Erythroid differentiation, various implications in cancer	Hematopoietic Stem cells (HSC)	(Khumukcham & Manavathi, 2021)
MyoD	Cofactor	Recruited by PBX1	Myogenin expression during somitogenesis	Mouse embryo	(Cho, Mallappa, Hernández-Hernández, Rivera-Pérez, & Imbalzano, 2015)
Evi-1	Enhancer	Enhancer of PBX1	Evi-1 promotes PBX1 expression	HSC	(Fukuda et al., 2015)
TET	Enhancer	Favors TALE binding on DNA	Histone 5mC oxidation	P19 mouse embryonal carcinoma cells	(Mahé et al., 2017)
PBX1 target					
MEOX1	Target	Regulated by PBX1	Mediates the cellular growth signal of PBX1 in ovarian cancers	Ovarian cancer cells (OVCAR3)	(Thiaville et al., 2012)
Jun	Target	Regulated by PBX1/PKNOX1 or Pbx1/MEIS1 + Oct-1		Non-small cell lung cancer	(Blasi et al., 2017)
HoxA4	Target	Indirectly regulated by PBX1 through regulation of BMI1 and EED	Partial control of mesoderm patterning	Mouse embryo limb mesoderm	(Capellini et al., 2008)
HoxB5	Target				
HoxB8	Target				
HoxC6	Target				
HoxD3	Target				
Pax1	Target	Regulated by PBX1	Partial control of Polycomb gene mesoderm patterning		
BMI1	Target				
EED	Target				
Polycomb	Target				

.

Smad	Target	Recruited by PBX1-PKNOX1	Mediators of activin signaling	LβT2 cells	(Bailey, Rave-Harel, McGillivray, Coss, & Mellon, 2004)(Bailey et al., 2004)
FSHB	Target	Regulated by PBX1	Regulated by PBX1		
Alx1	Target	Regulated by PBX1-Emx2	Scapula blade patterning	Mouse embryo dermomyotome	(Capellini et al., 2010)
DBX1	Target	Repressed by PBX1	Control of dorsoventral patterning of the spinal cord	Pbx mutant cortical progenitors	(Golonzhka et al., 2015)
DMRTA1	Target	Repressed by PBX1	Control of dorsoventral patterning of ventral zone cortical progenitors		
Dcx	Target	Regulated by PBX1	Promoter early bookmarked by pioneer factor PBX1	Progenitor cells and newborn neuroblasts	(Britta M. Grebbin & Schulte, 2017; Britta Moyo Grebbin et al., 2016)
Th	Target	Regulated by PBX1			
STAT3	Target	Activated by PBX1	Promote STAT3/JAK2 pathway and chemoresistance	Ovarian cancer	(Jung et al., 2016)
Hif1a	Target	Activated by PBX1-Meis-HoxA9	Favors hypoxic metabolism in hematopoietic stem cells (HSC)	HSC	(Kocabas et al., 2015)
HoxA2	Target	Regulated by PBX1-Hox	Hox–Pbx responsive cis-regulatory sequence = Hoxa2 regulation in rhombomere 4	P19 and COS7 cells	(Lampe et al., 2008)
Fgf10	Target	Activated by PBX1-Meis1-HoxB4	Promote lung development and surfactant secretion	Mouse embryonic lungs	(Li et al., 2014)
Wnt9b / Wnt3	Target	Regulated by PBX1	Mouse face morphogenesis	Mouse embryos	(Ferretti et al., 2011; Maili et al., 2020)
CD44	Target	Regulated by PBX1	Activation of T cells, risk factor of lupus disease (isoform PBX1-d)	Jurkat T cells	(Niu, Sengupta, Titov, Choi, & Morel, 2017)
BMP4	Target	Regulated by PBX1 in presence of retinoic acid	Promote differenciation of P19 cells under retinoic acid stimulation	P19 mouse embryonal carcinoma cells	(Qin, Haberbusch, Soprano, & Soprano, 2004)
DCN	Target	Regulated by PBX1 in presence of retinoic acid	Promote differenciation of P19 cells under retinoic acid stimulation		
GNRH1	Target	Regulated by PBX1/Prep1 or Pbx1/Meis1 + Oct-1	Promote restricted expression of GnRH in hypothalamic neurons	GT1-7 cells	(Rave-Harel et al., 2004)
NFE2L1	Target	Activated by PBX1	Ventral midbrain specification of dopaminergic neurons	Ventral midbrain E12.5 tissues	(Villaescusa et al., 2016)
Pitx3	Target	Activated by PBX1			
GBF1	Target	Activated by PBX1			
TPCN1	Target	Activated by PBX1			
TMEM218	Target	Activated by PBX1			
B3GNTL1	Target	Activated by PBX1			
CANX	Target	Activated by PBX1			
IQCD	Target	Activated by PBX1			
LRRC48	Target	Activated by PBX1			
ABHD2	Target	Activated by PBX1			
CDK5R2	Target	Repressed by PBX1			
ANKRD54	Target	Repressed by PBX1			
GPR108	Target	Repressed by PBX1			
PRR13	Target	Repressed by PBX1			
RAB40C	Target	Repressed by PBX1			
ACYP1	Target	Repressed by PBX1			
COQ7	Target	Repressed by PBX1			
NFATC2IP	Target	Repressed by PBX1			
GTPBP1	Target	Repressed by PBX1			
ZFP810	Target	Repressed by PBX1			
RNF5	Target	Repressed by PBX1			
RWDD4A	Target	Repressed by PBX1			
MED30	Target	Repressed by PBX1			
IER5L	Target	Repressed by PBX1			
SIX4IP	Target	Repressed by PBX1			

Accepted Article

ZSWIM4	Target	Repressed by PBX1			
ONECUT2	Target	Repressed by PBX1			
ALDH1A2	Target	Targeted by PBX1-HoxA1	Retinoic acid synthesis in hindbrain	Mouse embryos	(Vitobello et al., 2011)
CCND1	Target	Targeted by STAT3, STAT3 phosphorylation enhanced by PBX1	Promote STAT3/JAK2 pathway and chemoresistance	ccRCC cells	(Wei, Yu, & Li, 2018)
Pax6	Target	Regulated by PBX1	Pancreatic development	Mouse model	(Zhang et al., 2006)

"Interaction consequence" means either the direct molecular effects induced by PBX1 interaction with a cofactor, or the cellular (or developmental) consequence of PBX1 binding on a target gene's promoter.

Table 2: Molecular and clinical summary of previously published *PBX1*-mutated patients.

Heterozygous variant (NM_002585.4)	**Protein change (NP_002576.1)**	**Expected effect**	**Functional studies**	**Chromosomal sex**	**Face and head anlies**	**Renal and urinary anlies**	**External genitalia anlies**	**Gonadal/uterine anlies**	**Skeletal anlies**	**Cardiac malformations**	**Pulmonary hypoplasia**	**Deafness**	**NCS anlies**	**ID**	**Other**	**Article or accession number**
c.145C>T	p.(Gln49Ter)	Truncated protein or NMD (LoF)	No	XY	NC	Yes	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar VCV000666556
c.263C>T	p.(Thr88Ile)	Impaired protein function	No	XY	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	Yes	(Fromer et al., 2014)
c.319C>T	p.(Arg107Trp)	Impaired protein function	No	XX	Yes	Yes	No	Yes	No	No	Yes	NC	No	NC	Yes	(Arts et al., 2020)
c.319C>T	p.(Arg107Trp)	Impaired protein function	No	XY	Yes	Yes	Yes	Yes	No	Yes	Yes	NC	No	NC	Yes	(Arts et al., 2020)
c.370_371dup	p.(Gly125fs)	Truncated protein or NMD (LoF)	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar VCV000817312.1
c.392del	p.(Ala131fs)	Truncated protein or NMD (LoF)	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar RCV001008581.1
c.413_419del	p.(Gly138fs)	Truncated protein or NMD (LoF)	No	XY	Yes	Yes	Yes	No	Yes	No	No	No	No	Yes	No	(Riedhammer et al., 2017)
c.422C>G	p.(Ser141Ter)	Truncated protein or NMD (LoF)	No	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	(Heidet et al., 2017)
c.428del	p.(Asn143fs)	Truncated protein or NMD (LoF)	No	XX	No	Yes	No	No	Yes	No	No	Yes	No	No	No	
NG_028246.2(PBX1):c.511-2A>G	Splice defect	Truncated protein or NMD (LoF)	No	XY	No	Yes	No	No	No	No	No	No	No	No	No	
c.550C>T	p.(Arg184Ter)	Truncated protein or NMD (LoF)	No	XX	Yes	Yes	No	No	No	No	No	No	No	Yes	No	

.

Accepted Article

Accepted Article

NC	p.(Arg184Pro)	Impaired protein function	No	XY	Yes	Yes	No	No	Yes	Yes	No	No	No	Yes	No	(Alankarage et al., 2020)
c.671T>A	p.(Met224Lys)	Impaired protein function	No	XY	Yes	No	Yes	No	No	Yes	Yes	No	No	Yes	Yes	(Slavotinek et al., 2017)
c.680G>C	p.(Arg227Pro)	Impaired protein function	Yes	XY	Yes	No	Yes	No	No	No	No	No	No	Yes	No	
c.701G>A	p.(Arg234Gln)	Impaired protein function	No	XX	NC	Yes	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar VCV000559855
c.701G>C	p.(Arg234Pro)	Impaired protein function	Yes	NC	No	Yes	No	No	No	Yes	No	No	No	Yes	No	(Slavotinek et al., 2017)
c.704G>A	p.(Arg235Gln)	Impaired protein function	Yes	XY	Yes	No	Yes	Yes	No	No	No	No	No	Yes	No	(Kia, Sarafoglou, Mooganayakanakote Siddappa, & Roberts, 2019; Slavotinek et al., 2017)
c.704G>A	p.(Arg235Gln)	Impaired protein function	Yes	XY	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	No	No	(Slavotinek et al., 2017)
c.704G>T	p.(Arg235Gln)	Impaired protein function	Yes	XY	No	No	Yes	Yes	Yes	No	No	No	No	No	No	(Eozenou et al., 2019)
c.783dup	p.(Ser262fs)	Truncated protein or NMD (LoF)	Yes	NC	Yes	No	No	No	No	No	No	Yes	No	Yes	No	(Slavotinek et al., 2017)
c.818G>A	p.(Cys273Tyr)	Impaired protein function	No	NC	NC	Yes	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar VCV000598775
c.836A>G	p.(Gln279Arg)	Impaired protein function	No	NC	NC	Yes	NC	NC	NC	NC	NC	NC	NC	NC	NC	ClinVar RCV001251145.1
c.862C>T	p.(Arg288Ter)	Truncated protein or NMD (LoF)	Yes	NC	Yes	Yes	No	No	Yes	No	No	No	No	Yes	No	(Eozenou et al., 2019)
arr[hg19] 1q23.3(164749027_164786500)x1	p.?Deletion of exon 3-6	Truncated protein or NMD (LoF)	No	XY	Yes	Yes	Yes	No	Yes	Yes	No	No	No	Yes	Yes	(Fitzgerald, Powell-Hamilton, Shillingford, Robinson, & Gripp, 2021)
arr[hg19] 1q23.3(164749027_164786500)x1	p.?Deletion of exon 3-6	Truncated protein or NMD (LoF)	No	XX	Yes	Yes	No	No	Yes	No	No	No	Yes	No	Yes	
arr[hg19] 1q23.3(164749027_164786500)x1	p.?Deletion of exon 3-	Truncated protein or NMD (LoF)	No	XX	Yes	Yes	No	No	Yes	No	No	No	No	No	Yes	

	6															
Whole gene deletion		Haploinsufficiency	No	XY	Yes	Yes	No	No	No	No	No	No	Yes	Yes	Yes	(Le Tanno et al., 2017)
Whole gene deletion		Haploinsufficiency	No	XX	Yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes	No	
Whole gene deletion		Haploinsufficiency	No	XX	Yes	Yes	No	No	No	Yes	No	No	No	Yes	No	
Whole gene deletion		Haploinsufficiency	No	XY	Yes	Yes	Yes	No	Yes	No	No	No	No	No	No	
Whole gene deletion		Haploinsufficiency	No	XX	Yes	Yes	No	No	Yes	No	No	Yes	Yes	Yes	Yes	
Whole gene deletion		Haploinsufficiency	No	XY	No	Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	No	
Whole gene deletion		Haploinsufficiency	No	XX	Yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes	Yes	
Whole gene deletion		Haploinsufficiency	No	XY	Yes	Yes	Yes	No	No	No	No	Yes	No	Yes	No	
Whole gene deletion		Haploinsufficiency	No	XX	No	Yes	No	No	No	No	No	Yes	No	No	No	(Heidet et al., 2017)
Whole gene deletion		Haploinsufficiency	No	XY	Yes	Yes	No	No	No	No	No	No	No	Yes	No	
Whole gene deletion		Haploinsufficiency	No	XY	Yes	NC	NC	NC	Yes	Yes	No	No	No	Yes	Yes	DECIPHER 398083
Whole gene deletion		Haploinsufficiency	No	XY	Yes	Yes	No	Yes	No	No	No	No	No	Yes	No	DECIPHER 387270
Whole gene deletion		Haploinsufficiency	No	XX	Yes	No	No	No	No	No	No	No	No	No	No	DECIPHER 276516
Whole gene deletion		Haploinsufficiency	No	XY	Yes	No	Yes	No	Yes	No	No	No	No	Yes	Yes	DECIPHER 262739

NC: not communicated, NMD: Nonsense-Mediated Decay, LoF: Loss of Function, Anlies: anomalies, CNS: Central Nervous System, ID: Intellectual Disability.

Accepted Article