MOLECULAR STUDIES ON *HOXD* GENES: NEW INSIGHTS INTO THE MECHANISMS OF LIMB DEVELOPMENT AND PATHOGENESIS.

Dissertation zur Erlangung des akademischen Grades des Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie der Freien Universität Berlin

vorgelegt von

BARBARA DŁUGASZEWSKA
aus Poznań, Polen

September 2005

1. Gutachter: Prof. Dr. Stefan Mundlos

2. Gutachter: Prof. Dr. Horst Kreß

Disputation am 5. Dezember 2005

TABLE OF CONTENTS

ABBREVIATIONS

Ac	Acetate
acc. no.	Accession number
AD	Activation domain
AER	Apical ectodermal ridge
Ala	Alanine
AP	Anteroposterior
3-AT	3-amino-1,2,4-triazole
BAC	Bacterial artificial chromosome
BD	DNA binding domain
BMP	Bone morphogenetic protein
DAPI	4',6-Diamidino-2-phenylindole
DBCR	Disease-associated balanced chromosomal rearrangements
DEPC	Diethylpyrocarbonate
DIG	Digoxygenin
DMSO	Dimethyl sulfoxide
DO	Dropout supplement
DV	Dorsoventral
ELCR	Early limb control region
FGF	Fibroblast growth factor
FISH	Fluorescence *in situ* hybridisation
GCR	Global control region
Gli3	Transcription factor, GLI-Kruppel family member
Gli3R	Repressor from of Gli3
HISS	Heat inactivated sheep serum
HOX	Homeodomain-containing genes
Hoxd13-HD	Hoxd13 protein lacking homeodomain
Hoxd13+14Ala	Hoxd13 protein with an expanded Ala stretch for 14 Ala residues
Hoxd13_2Ala	Hoxd13 protein with the reduced Ala stretch to 2 Ala residues
Ig	Immunoglobulin
LPM	Lateral plate mesoderm
Meis	Myeloid ecotropic viral integration site 1 homologues

NCBI	National Center for Biotechnology Information
NLS	Nuclear localisation signal
OD_{600}	Optical density at the wave length 600 nm
ORF	Open reading frame
PBS	Phosphate Buffered Saline
Pbx	Pre B-cell leukemia transcription factors
PD	Proximodistal
Peg10	Paternally expressed 10
PFA	Paraformaldehyde
RA	Retinoic acid
RARE	Retinoic acid responsive element
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SD –T	SD medium lacking tryptophane
SD –THULL	SD medium lacking tryptophane, histidine, uracil, lysine and leucine
SD –TL	D medium lacking tryptophane and leucine
SD –L	SD medium lacking leucine
Shh	Sonic hedgehog
SPD	Synpolydactyly
TGF-β	Transforming growth factor β
UTR	Untranslated region
Wnt	Family of highly conserved secreted signalling molecules
w/v	Weight per volume
YAC	Yeast artificial chromosome

1 INTRODUCTION

1.1 Origin of tetrapod limbs

Tetrapods (four-footed, from Greek: *tetra* four + *pod-, pous* foot), i.e. vertebrates with two pairs of limbs, have a very long history, starting in the Devonian period, meaning at least 360 million years ago. At that time, the transition between fishes and tetrapods occurred, often referred to as the fin-to-limb transition.

It is believed that limbs evolved to facilitate exploiting a shallow-water environment and they were originally used horizontally as props or/and paddles (Lebedev 1997). Their proximal parts share homology with fish fins, however more distal limb structures, including digits, are present in tetrapods only. The first known digited tetrapods living in the late Devonian, *Ichthyostega*, *Acanthostega* and *Tulerpeton*, were polydactylous and probably aquatic (Coates and Clack 1990; Lebedev 1997; Clack 2002). Spreading over the land, achieved during tetrapod evolution, was associated with profound limb changes including reduction in the digit number (to maximal 5) and increase in the number of carpus and tarsus bones. This process was followed by other morphological, anatomical and physiological changes allowing adaptation to the terrestrial lifestyle.

1.2 Limb development in embryogenesis

1.2.1 Anatomical view

Limb development follows the same sequence of events in all higher vertebrates. Limb buds are derived from specific regions of the so-called lateral plate mesoderm (LPM), which comprises two stripes of tissue that run along the length of the main body axis. Positioning of the limb field (group of cells which give rise to the limb bud) is dependent on complex signalling and can be initiated only if the prospective limb-forming region responds properly to the signals from more medial tissues. After the limb field has been specified, the limb bud induction and its outgrowth along three major axes, proximodistal (PD), anteroposterior (AP) and the dorsoventral (DV) starts. In respect to the limb structure, proximal-to-distal patterning refers to the formation of the stylopod (upper arm or thigh), through the zeugopod (forearm or shank) up to the autopod (wrist and hands or ankle and feet), respectively. The

anterior part of the limb is defined by the location of the digit number 1 (thumb), whereas the posterior part corresponds to the location of the digit number 5 (little finger) in pentadactylous tetrapods. Due to thorough studies in the past decades, many factors and pathways responsible for limb patterning processes along each of the axes have been determined. Most studies were done in chicken or mouse models. Especially, transplantation experiments, injection of signalling molecules into animal embryos, transgene introduction or knockout technology were useful in these complex analyses.

1.2.2 Molecular basis of limb patterning

1.2.2.1 Establishment of the proximodistal axis in the limb

Limb outgrowth along the PD axis is dependent on the apical ectodermal ridge (AER), a layer of tissue that covers the rim of the distal tip of the limb bud. It has been observed that the AER function is mediated by different members of the fibroblast growth factor (FGF) superfamily (Niswander and Martin 1992; Niswander et al. 1993) and that FGF signalling seems to be essential for proper limb patterning both at the early as well as at the late stage of the limb bud outgrowth (Summerbell 1974; Rowe and Fallon 1982; Saunders 1998; Lewandoski et al. 2000).

AER initiation is a complex event, depending on synthesis and transport of many different molecules. At the beginning, a very important role is ascribed to FGF and Wnt proteins which are synthesised in the mesoderm underlying the prospective limb bud and which activate target genes (for instance other *FGF*s) in the AER (Capdevila and Izpisua Belmonte 2001; Tickle and Munsterberg 2001; Barrow et al. 2003). Later, the limb bud outgrowth is regulated by FGF or retinoc acid (RA) signalling, depending on the distance to the AER. The distal part of the bud located close to the AER is under the influence of different fibroblast growth factors, which are capable to repress retinoic acid production. Lack of RA, together with the expression of “distal” bone morphogenetic proteins (BMPs) and homeobox (Hox) proteins, blocks activation of the proximising genes *Meis1* and *Meis2* in the distal limb bud. The proximal part of the future limb is further away from the AER, therefore it lacks FGF molecules. Instead, RA signalling is activated and can induce expression of *Meis1* and *Meis2* (Capdevila et al. 1999; Mercader et al. 1999; Mercader et al. 2000; Capdevila and Izpisua Belmonte 2001).

The maintenance of the AER occurs via two positive feedback loops (Laufer et al. 1994; Niswander et al. 1994; Zuniga et al. 1999; Capdevila and Izpisua Belmonte 2001; Panman and Zeller 2003). The first one is established between Fgf10, expressed in the limb bud mesoderm, and Fgf8 from the AER. In the second loop, FGF from the AER activates *Sonic hedgehog* (*Shh*) expression in the posterior distal mesenchyme. Shh, acting via *Formin* and *Gremlin* genes, switches off BMP signalling, which in turn enables FGF activation in the AER.

Up to now two models explaining the mechanism of proximal-to-distal patterning in the developing limb bud have been proposed. The older one, called progress zone model, assumes that an internal clock controlled by the AER determines the fate of the cells lying underneath, in the so-called progress zone. While proliferation proceeds, older cells leave the progress zone and escape from the influence of the AER. Cells, which left the progress zone earlier, give rise to the more proximal parts of the limb, whereas cells, which stayed longer under the AER control, contribute to the distal parts of the limb (Summerbell et al. 1973). Recently, the progress zone model has been questioned and a new hypothesis has been proposed (Dudley et al. 2002; Sun et al. 2002). According to it, cells are "labelled" as proximal or distal very early, and the limb development corresponds to the outgrowth of the pre-specified domains.

1.2.2.2 Establishment of the AP axis

Establishment of the anterioposterior axis in the limb bud is tightly connected with the AER. Fgf4 and other FGFs expressed in the AER are able to activate *Shh* in the posterior part of the limb bud, called zone of polarising activity (ZPA). However, the induction of *Shh* is only possible in the presence of posteriorising factors like RA and Hox proteins (Johnson et al. 1994; Niswander et al. 1994; Knezevic et al. 1997; Mackem and Knezevic 1999; Catala 2000; Capdevila and Izpisua Belmonte 2001; Panman and Zeller 2003). AP asymmetry in the limb bud is subsequently established by the negative feedback loop between Shh and the repressor form of Gli3 (Gli3R). The interplay between these two proteins results in formation of the Shh gradient along the AP axis. Lack of Shh signalling in the anterior part of the limb bud results in expression of specific genes, which give anterior identity to the mesenchyme. Posterior mesenchyme is specified by the positive feedback loop between 5' *Hox* genes, *Shh* and *dHand* (te Welscher et al. 2002a; te Welscher et al. 2002b; Panman and Zeller 2003; Zakany et al. 2004). A schematic representation of these pathways can be seen in Fig. 1.

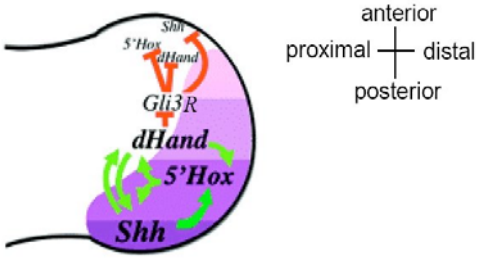


Fig. 1 Establishment of the anterioposterior (AP) asymmetry in the limb bud. Only the most important players are shown. 5' *Hox* genes are expressed in the posterior part of the limb bud and activate Sonic hedgehog (Shh) and the transcription factor dHand, which subsequently activate each other. All these proteins inhibit accumulation of transcriptional repressor Gli3R in the posterior part of the limb bud. Therefore, Gli3R is present only in the anterior structures, where it suppresses transcription of 5' *Hox, dHAND* and *Shh.* The resulting Shh gradient (purple zones) along the AP axis drives expression of different genes in the anterior and the posterior parts of the limb bud. Adapted from Zakany *et al.* 2004.

1.2.2.3 Limb patterning along the dorsoventral axis

Establishment of the DV axis of the limb bud is strictly dependent on the formation of the DV boundary at the mid-point of the AER. Wnt/β-catenin signalling from the ectodermal ridge activates BMPs in the ventral ectoderm, which subsequently induce expression of the *Engrailed 1* gene (*En-1*), coding for a homeobox-containing transcription factor. Presence of En-1 proteins specifies ventral ectoderm and blocks expression of *Wnt7a*, which is active only in the En-1-free dorsal cells of the distal limb bud. Wnt7a signalling from the dorsal ectoderm induces expression of the LIM-homeodomain factor Lmx1b in the same tissue. Thus, both Wnt7a and Lmx1b are responsible for the establishment of the dorsal pattern (Capdevila and Izpisua Belmonte 2001). In addition, it is known that *Radical fringe* (*Rfng*), expressed in the dorsal ectoderm and in the whole AER of chicken limb buds, might be also involved in the DV patterning (Tickle and Munsterberg 2001).

1.3 *HOX* genes in limb development

1.3.1 Chromosomal clustering of *Hox* genes is linked to a specific expression pattern during embryogenesis

Hox genes code for a conserved family of homeobox-containing transcription factors. They are usually clustered and can be found in genomes of different organisms, for instance cnidarians, nematodes, arthropods, echinoderms, cephalochordates and vertebrates (Martinez et al. 1999; Aboobaker and Blaxter 2003; Hill et al. 2003; Wagner et al. 2003). In mammals *Hox* genes are organised in four clusters named A, B, C and D located on different chromosomes. Each cluster consists of 9–11 genes, which are expressed according to the spatio-temporal collinearity rule along the primary body axis. This means that the anterior genes, located at the 3' extremities of the complexes, are expressed earlier and more anterior in the embryo than the posterior genes, located at the 5' ends of the clusters (Duboule and Dollé 1989; Izpisúa-Belmonte et al. 1991). The pattern of *Hox* gene expression is evolutionary conserved and can be observed not only in vertebrates but also in invertebrate species, for example in *Drosophila*, which contains a single homeobox cluster (HOM-C complex) (Fig. 2).

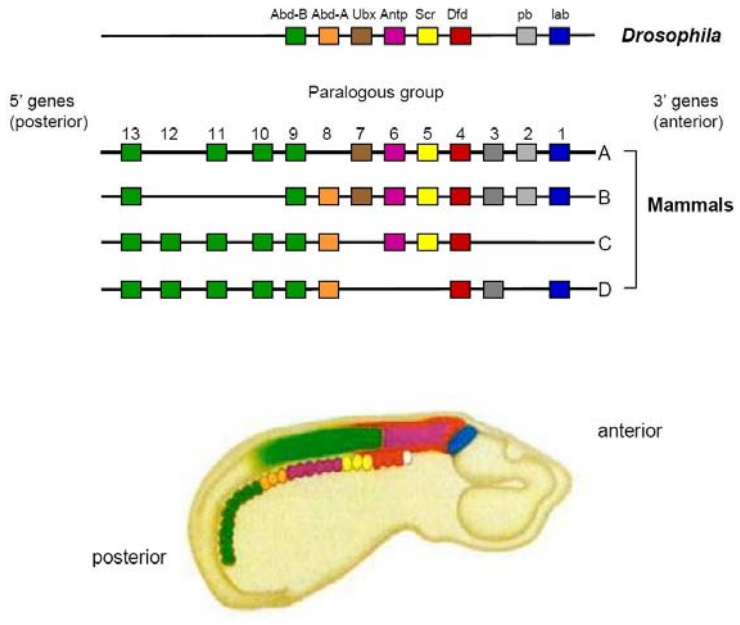


Fig. 2 Organisation of the *Drosophila* HOM-C complex and of the corresponding four *Hox* complexes in mammals. Genes marked with the same colour exhibit the highest homology, thus correspond to the same paralogous groups. Below, expression pattern of *Hox* genes along the main body axis in the mouse embryo. Different colours correspond to expression domains of various *Hox* genes, as shown in the upper panel. Adapted from Carroll 1995.

In addition, in certain tissues *Hox* genes are expressed in a quantitative order. For instance, in limbs of higher vertebrates *Hoxd* genes follow the so called third collinearity rule in such a way that the most posterior gene *Hoxd13* is expressed very strongly, whereas genes located towards the 3' end of the cluster, have a progressively reduced expression level (Kmita et al. 2002a).

1.3.2 *Hox* genes and pre-patterning of the embryo

Nested *Hox* gene expression along the primary body axis forms a pre-pattern, which can define prospective organ regions. For example, anterior expression boundaries of *Hoxc6*, *Hoxc8* and *Hoxb5* in the lateral plate mesoderm of vertebrate embryos fit exactly to the the regions where the forelimb fields are specified, thus suggesting a role of these *Hox* genes in the determination of these particular regions in the embryo (Nelson et al. 1996). In addition, it has been shown that *Hoxb5* knockout mice develop the shoulder girdle shifted, which corresponds to the shift in *Hox* expression domains compared to the wildtype mice (Burke et al. 1995; Rancourt et al. 1995; Gaunt 2000). It is possible that also other *Hox* genes influence the pre-specification of the limb fields. For example, it has been observed that ectopic expression of the *Hoxb8* gene in the anterior part of the limb bud induces an additional ZPA. Furthermore, the lack of limbs in snakes correlates with specific changes in *Hox* expression domains (Cohn and Tickle 1999).

1.3.3 Posterior *Hoxa* and *Hoxd* genes and their role in limb patterning

As previously mentioned, *Hox* genes seem to play a very important role in specification of the limb field, but it is also known that they are essential at later stages of limb development, namely for the establishment of the PD and the AP limb axes.

Expression profiles of the *Hoxd9-13* genes in limb buds differ depending on the stage of development. In the first phase, the posterior *Hox* genes are expressed in the entire limb bud. According to the collinearity rules, more anterior genes such as *Hoxd9* and *Hoxd10* are expressed first, followed by expression of the more posterior genes. A similar profile can be observed for the *Hoxa9-13* genes. In phase II, a clear change in the expression domains is visible, namely the *Hoxd* genes are activated in the posteriorly nested order. It means that the more anterior genes, for instance *Hoxd10*, are expressed in the anterior part of the limb bud, whereas expression of the posterior genes like *Hoxd13* is restricted to the posterior

mesoderm. In the last phase, the *Hoxd* genes are expressed only in the distal part of the limb bud, and there is a switch in the AP expression domains, so that the anterior genes are expressed in the posterior part of the limb bud and *vice versa* (Fig. 3) (Izpisua-Belmonte and Duboule 1992; Duboule 1994; Nelson et al. 1996).

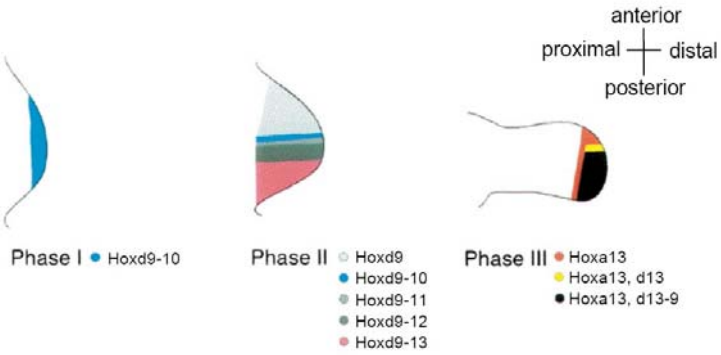


Fig. 3 Three phases of *Hoxd* genes expression in developing limb buds. See text for further explanations. Adapted from Shubin *et al.* 1997.

Overlapping expression domains of *Hox* genes create a dynamic pattern for Hox proteins activity. However, it is known that the region in which few *Hox* genes are expressed, is dominated by the most 5' gene (i.e. the most posterior one). This phenomenon is called posterior prevalence and leads to the situation that in various limb regions, different *Hox* genes at different timepoints play a dominant role. Thus, during phase I, expression of *Hoxd9* and *Hoxd10* specifies the stylopod. Zeugopod patterning is accomplished during both phases I and II, whereas digit formation is dependent on expression of the most posterior *Hox* genes during phases II and III (Johnson and Tabin 1997).

This model has been confirmed by the observation of skeletal defects and *Hox* expression domains in different *Hox* mutants. Thus, *Hoxd9* and double *Hoxa9*/*Hoxd9* knockout mice show forelimb defects, affecting the humerus (Fromental-Ramain et al. 1996), whereas *Hoxd9*/*Hoxd10* double mutants show alterations in the hindlimb skeleton, visible on the border between the stylopod and the zeugopod, which is similar to the defects observed in the single *Hoxd10* mutant mice. In addition, in a small percentage of *Hoxd9*/*Hoxd10* $^{-/-}$ mice the humerus is also deformed (Carpenter et al. 1997; de la Cruz et al. 1999). Inactivation of both *Hoxa10* and *Hoxd10* affects the femur, knee joint and tibia/fibula in mice (Wahba et al. 2001). Moreover, improper development of the thigh and shank has been described in mice lacking either *Hoxd11* or *Hoxa11* genes and in the double mutants (Small and Potter 1993;

Davis and Capecchi 1994; Davis et al. 1995; Favier et al. 1995; Boulet and Capecchi 2004). Additionally, in *Hoxd11* knock-out mice metacarpals, phalanges and wrist bones are also affected (Davis and Capecchi 1994; Favier et al. 1995). Interestingly, compound mutants for *Hoxa10* and *Hoxd11* show a more severe phenotype, giving the evidence that zeugopod development is dependent on the proper expression of several posterior *Hox* genes from both *Hoxa* and *Hoxd* paralogous groups (Favier et al. 1996; Wahba et al. 2001). Finally, production of various single and compound mutant mice indicated that four posterior genes, *Hoxa13, Hoxd11*, *Hoxd12* and *Hoxd13* regulate digit development in a dose-dependent manner (Dolle et al. 1993; Davis and Capecchi 1996; Kondo et al. 1996; Zakany et al. 1997; Kondo et al. 1998). Moreover, *Hoxa11* and *Evx2*, the latter one being located at the proximal end of the *Hoxd* complex, were also shown to contribute to digit morphogenesis, however to a lesser extend (Zakany and Duboule 1999).

1.3.4 A special role of the posterior *Hoxd* genes in the development of digited limbs and the contribution of different pathways to digit formation

It has been proposed that during evolution *Hox* genes acquired new functions, which enabled development of new structures, for instance digits, which appeared only after the fin-to-limb transition. This hypothesis has been supported by the recognition that the specific expression domains of the 5' *Hoxd* genes known for tetrapods, is not established during development of teleost pectoral fins (Sordino et al. 1995). For a long time it was not clear, what exactly happens on the molecular level and how Hoxd proteins regulate digit formation. However, recent data revealed links between *Hox* genes and pathways which play an established role in digit development.

Until now it was known that digit patterning is dependent on *Sonic hedgehog*, as concluded from the analysis of *Shh* mutant mice (Kraus et al. 2001). Moreover, Shh-signalling was shown to induce various genes, among them *BMPs*. Their expression in the interdigital necrotic zone suggested their contribution to apoptotic events separating prospective digits (Yokouchi et al. 1996; Zou and Niswander 1996; Chen and Zhao 1998; Drossopoulou et al. 2000; Guha et al. 2002). Recently, it has been shown that *BMPs* can also be directly activated by posterior *Hox* genes (Suzuki et al. 2003; Knosp et al. 2004). Furthermore, other experiments indicated that the early posterior repression of 5' *Hoxd* genes is required for the localised expression of *Shh*, which in turn promotes late activation of *Hoxd* genes leading to digit asymmetry (Zakany et al. 2004). Moreover, it has been lately shown that Gli3, the

intercellular mediator of Shh, directly interacts with 5' Hoxd proteins and it was suggested that the varying Gli3:Hoxd ratio across the limb bud is responsible for differential activation of target genes (Chen et al. 2004). Therefore, according to the current knowledge, the interplay between Hoxd, Gli3 and the Shh- and the BMP-signalling is thought to pattern the prospective digit area.

1.3.5 Mechanisms controlling expression of *Hox* genes

The identification of regulatory sequences responsible for gene expression is fundamental to obtain the full knowledge about gene function, its connection to the cellular network and its possible implication in diseases. Thus, a very important role of *Hox* genes for the development and patterning of the embryo led many scientists to investigate regulation of these genes in more detail. Especially, a lot of effort has been put to find promoters and enhancers, as well as to explain the collinear expression of *Hox* genes. Moreover, special attention has been directed on the regulation of the 5' *Hox* genes in developing limb buds.

1.3.5.1 Regulation of single Hox genes

Different studies revealed that several anterior *Hox* genes respond to RA treatment, so it was not surprising that retinoic acid responsive elements (RAREs) have been found in enhancer regions of different *Hoxa*, *Hoxb* and *Hoxd* genes (Maconochie et al. 1996; Morrison et al. 1996; Gould et al. 1998; Packer et al. 1998; Zhang et al. 2000; Oosterveen et al. 2003). In addition to the induction by RA, some of these genes are controlled by other mechanisms as well, for instance autoregulation (Popperl and Featherstone 1992; Packer et al. 1998; Manzanares et al. 2001; Yau et al. 2002), cross-regulatory interactions (Gould et al. 1997; Maconochie et al. 1997; Manzanares et al. 2001; Yau et al. 2002) or activation by other transcription factors (Sham et al. 1993; Manzanares et al. 1997; Manzanares et al. 1999; Manzanares et al. 2002).

In many cases enhancers specific for single *Hox* genes, flanking these from the 3' or the 5' side, were found (Whiting et al. 1991; Eid et al. 1993; Gerard et al. 1993; Knittel et al. 1995; Shashikant et al. 1995; Becker et al. 1996; Morrison et al. 1997; Kwan et al. 2001). Interestingly, these regulatory elements can also be shared between neighbouring genes, as shown for the *Hoxa* and *Hoxb* clusters (Gould et al. 1997; Sharpe et al. 1998; Oosterveen et al. 2003).

Recently, several genes for microRNAs have been proposed to lie within *Hox* clusters and to downregulate expression of single genes, as shown *in vivo* for *Hoxb8* and *in vitro* for *Hoxb8*, *Hoxc8*, *Hoxd8*, and *Hoxa7* (Calin et al. 2004; Mansfield et al. 2004; Yekta et al. 2004).
All these data, although valuable for understanding the regulation of *Hox* genes, do not explain their specific nested expression domains thought to result from the clustered organisation on chromosomes. Therefore, it has been proposed that in addition to the regulation of single genes driven by their promoters and local regulatory elements, other regions controlling and regulating the expression of the whole cluster have to be present as well.

1.3.5.2 Global regulation of the Hox clusters

Molecular mechanisms responsible for driving the collinear expression of *Hox* genes have been proposed by several authors (Deschamps et al. 1999; Kmita and Duboule 2003). The first hypothesis suggests that the mechanism of the collinearity is dependent on the progressive accessibility of *Hox* transcriptional units from one end of the cluster to the other. This might involve the process of opening the chromatin structure by transcription of one gene, which would be spread out on to the neighbouring regions. According to this model, expression of the most anterior *Hox* genes is initiated by retinoic acid (Roelen et al. 2002), and upon induction sequential activation of the more 5' genes occurs. The second hypothesis proposes that the collinear activation of the *Hox* genes is driven by local cis-acting elements, which show increasing or decreasing affinity to certain signalling molecules. The existing gradient of these molecules could be "read" along the cluster, allowing expression of the *Hox* genes in the proper way. The third mechanism assumes that a global control region (GCR), located outside the clusters, can regulate several genes in a relatively promoter-unspecific manner. These three mechanisms are not exclusive; on the contrary, they could work in combination with each other, depending on the site and the stage of *Hox* gene expression.

1.3.5.3 Regulation of the posterior Hox genes during limb development

It has been observed that the posterior *Hoxd* genes, *Hoxd10*, *Hoxd11, Hoxd12* and *Hoxd13* show very similar expression domains in presumptive digits (Sordino and Duboule 1996), therefore it has been proposed that these four transcription units are under the control of the same enhancer, which could regulate their spatial and temporal expression in developing limbs (van der Hoeven et al. 1996; Herault et al. 1999). Moreover, it has been suggested that this element (called digit enhancer) is located centromeric to the *Hoxd* complex (Kondo and Duboule 1999; Spitz et al. 2001).

Recently, a conserved region (called region XII) located at the 5' end of the *Hoxd* cluster has been described and it has been shown to be required for the quantitative collinearity of the *Hoxd* genes in limbs (Kmita et al. 2002a; Kmita et al. 2002b). In addition, an approximately 40 kb large segment of human DNA located further 5' to the *HOXD* cluster, has been found to contain the digit enhancer and to control the expression of both *Hoxd* and *Evx2* genes (Spitz et al. 2003). Moreover, a region regulating *Hoxd* gene expression before Shh signalling (**e**arly **l**imb **c**ontrol **r**egion – ELCR) has been lately proposed to be located 3' to the whole complex (Zakany et al. 2004).

Thus, the current model proposes that at the early stage of limb development, the ELCR controls phase II of *Hoxd* gene expression (more 5' genes become activated progressively in more posterior domains). The Hoxd-dependent Shh production probably causes a switch into the later phase of *Hoxd* gene regulation, which is controlled by the global elements located 5' to the complex. The AP expression domains of the *Hoxd* genes change, and at the same timepoint the quantitative collinearity is established by interactions between the remote digit enhancer and local regulatory elements (for instance region XII). The strongest effect, resulting in the highest expression level, is exerted on the most 5' gene (*Hoxd13*) and progressively weaker effects on more anterior genes (Fig. 4) (Deschamps 2004).

Regulation of *Hoxa* expression has not been studied so extensively as that of the *Hoxd* genes. However, recent analyses indicated that over 900 kb long regions upstream of the *Hoxa* and the *Hoxd* clusters are higly conserved. This gave rise to the hypothesis that also for the posterior *Hoxa* genes, limb-specific expression might be dependent on global regulatory elements present upstream of the cluster. Further functional tests partially confirmed this theory (Lehoczky et al. 2004).

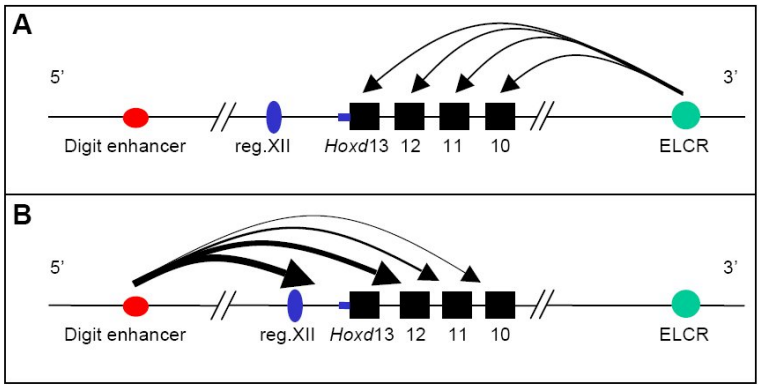


Fig. 4 Schematic representation of 5' *Hoxd* regulation in limb buds. *A*: The hypothetical early limb control region (ELCR) located 3' to the *Hoxd* cluster controls the early phase of *Hoxd* gene expression. *B*: In the later phase, *Hoxd* expression is regulated by the digit enhancer located 5' to the cluster. Local regulatory elements, for instance region XII or other sequences located within the cluster (marked in blue), can co-operate with the globally acting digit enhancer, leading to establishment of the quantitative collinearity. Thickness of the arrows indicates the strength of the enhancement and corresponds to the expression level of a particular gene. Adapted from Zeller and Deschamps 2002, and Deschamps 2004.

1.3.6 *HOX* gene mutations and limb malformations in humans

To date, mutations in four human *HOX* genes, namely *HOXA11*, *HOXA13, HOXD10* and *HOXD13*, have been found. All these mutations, as expected, are associated with limb malformations.

In two families a single nucleotide deletion within the second exon of *HOXA11*, resulting in a frameshift and a premature stop codon, has been found to co-segregate with the proximal radial-ulnar synostosis (Thompson and Nguyen 2000). Different changes in *HOXA13*, including missense and nonsense mutations, polyalanine expansions or small deletions within the promoter region, cause hand-foot-uterus syndrome, a rare dominantly inherited condition affecting distal limbs and genitourinary tract (HFUS, OMIM #140000) (Mortlock and Innis 1997; Goodman et al. 2000), or Guttmacher syndrome (OMIM #176305) (Innis et al. 2002). Recently, a missense mutation in the *HOXD10* gene has been described to be the cause of isolated congenital vertical talus, also known as rocker-bottom feet (CVT, OMIM # 192950), and Charcot-Marie-Tooth disease (CMT, OMIM# 118220) in a big American family of Italian descent (Shrimpton et al. 2004).

The first described mutation within the *HOXD13* gene was an imperfect alanine-coding trinucleotide expansion in the first exon of the gene. This insertion has been linked to

synpolydactyly (SPD, OMIM #186000), a dominant inherited limb disorder affecting exclusively autopods (Muragaki et al. 1996). SPD is characterised by syndactyly of the third and fourth fingers and the fourth and fifth toes, both associated with polydactyly. Subsequent studies revealed that in more than 20 families published up to now, SPD is caused by the same pathological polyalanine tract expansions in HOXD13 protein and that the size of these expansions correlates with the severity of the phenotype (Akarsu et al. 1996; Goodman et al. 1997; Kjaer et al. 2002). Other mutations found in the *HOXD13* gene such as intragenic frameshift deletions, predicted to result in truncated proteins, an acceptor splice site mutation and a missense mutation in exon 2, cause an atypical form of SPD (Goodman et al. 1998; Calabrese et al. 2000; Debeer et al. 2002; Kan et al. 2003). Interestingly, a different missense mutation in the same exon 2 of *HOXD13* has been found in a family with a dominantly inherited combination of brachydactyly and polydactyly (Caronia et al. 2003). The SPD phenotype was also observed in 2 related patients with a microdeletion at the 5' end of the *HOXD* cluster, which removes *HOXD9* to *HOXD13* and extends 85 kb upstream of *HOXD13* (Goodman et al. 2002). In contrast, larger deletions involving chromosome 2q31.1, where the *HOXD* complex is located, have been associated with minor digital anomalies (Nixon et al. 1997; Slavotinek et al. 1999), or with major limb defects (Boles et al. 1995; Nixon et al. 1997; Goodman 2002), or with a combination of severe limb and genital abnormalities (Del Campo et al. 1999).

1.3.7 Hox proteins in complexes

The main role of Hox transcription factors is to regulate the pattern of chondrogenic differentiation in limbs, probably by activation a variety of target genes. However, different experiments *in vitro* revealed a poor affinity of single Hox proteins to the DNA and a low specificity of this binding (Gehring et al. 1994; Pellerin et al. 1994; Lu et al. 1995). Thus, it has been suggested that in order to increase the affinity and to generate binding specificity, Hox proteins form multiprotein-DNA complexes. Known Hox-interaction partners are homeodomain-containing proteins Pbx1 and Pbx2 (Lu et al. 1995; Chang et al. 1996; Knoepfler et al. 1996; Shen et al. 1996; Knoepfler and Kamps 1997; Lu and Kamps 1997). Binding of both Pbx molecules is dependent on the YPWM motif in Hox proteins from paralogous groups 1 – 8 and on the specific tryptophane residues in paralogues 9 –10 (Chang et al. 1995; Knoepfler and Kamps 1995; Chang et al. 1996; Shen et al. 1996; Shen et al. 1997b). Furthermore, Pbx proteins have been shown to dimerise with Meis (Chang et al.

1997; Knoepfler et al. 1997), and since Hox paralogues 9-13 bind Meis as well (Shen et al. 1997a), trimeric complexes of Meis, Pbx and Hox can be formed (Shanmugam et al. 1999). It is known that all three classes of proteins are present in the proximal part of a limb bud, therefore it is believed that Meis/Pbx/Hox complexes control development of the proximal limb structures by regulating transcription of downstream targets (Capdevila and Izpisua Belmonte 2001). Furthermore, trimeric complexes between Prep, Pbx and 3' Hox proteins have also been described (Berthelsen et al. 1998). However, in this case Prep proteins do not interact directly with Hox, and the binding occurs via Pbx (Ferretti et al. 1999; Fognani et al. 2002).

Little is known about interaction partners of the most posterior Hox paralogues. As already mentioned, Meis proteins can bind them *in vitro*, but it is rather unlikely that this binding occurs *in vivo* as well, since *Hox* paralogues 11-13 are expressed only in the distal part of the limb, whereas *Meis* expression is inhibited in this region. Recently, it has been shown that the zinc finger transcription factor Gli3, which plays a role in the AP limb patterning, directly binds the homeodomains of Hoxd11, Hoxd12 and Hoxd13 proteins (Chen et al. 2004). Furthermore, it is known that Hoxa13 and Hoxd13 functionally cooperate with Sp1, a GC-box binding transcription factor (Suzuki et al. 2003). However, up to date no other interaction partners for Hoxd13 are known. Thus, it would be interesting to find and define more factors, which co-operate with this and other Hox proteins in order to regulate distal limb development.

1.4 Outline of the project

Human disorders are currently investigated very thoroughly at the molecular level, and many disease-causing genes have been identified so far. However, it became clear that knowledge about the defective gene or even the mutated nucleotide is often not sufficient for prediction of the clinical phenotype or the severity and course of the disease. This is because genes, and most of the proteins which they encode, do not act alone, but they are parts of different pathways and can be regulated or modified by the action of other genes in various ways. To investigate the factors and mechanisms that play a role in the clinical variability of Mendelian disorders, the Collaborative Research Centre in Berlin has been founded. As a part of this research the molecular pathology of *HOXD*-related limb malformations is being studied.

The goal of my project was to investigate at the cytogenetic and molecular level the autosomal translocation t(2;10)(q31.1;q26.3) carried by a male patient presented with synpolydactyly and mental retardation. Systematic analysis revealed that the breakpoint on chromosome 2 is located in the vicinity of the *HOXD* cluster and that it does not disrupt any known gene. The knowledge about the complexity of *HOXD* regulation mechanisms allows us to hypothesise that the translocation might have disturbed these subtle mechanisms by position effect, thus being causative for the limb phenotype in the patient. The second part of the project focused on the search for Hoxd13 interaction partners. A yeast two-hybrid screen has been performed, and afterwards candidate genes were studied using RNA *in situ* hybridisation, immunofluorescence and coimmunoprecipitation methods. The preliminary results presented here give new insights into the molecular mechanisms of limb development and pathogenesis.

2 MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals and reagents

Standard chemicals and reagents are listed in Table 1 in the alphabetical order.

Table 1 Standard chemicals and reagents

Name	Company
Acetic anhydride	Sigma
Acrylamide mix [Rotiphorese® Gel 40 (29:1)]	Roth
L-Adenine hemisulfate salt	Sigma
Agar	Difco
Agarose	Invitrogen
Ammmonium chloride	Merck
Ammonium persulfate (APS)	Invitrogen
Ampicillin	Sigma
L-Arginine HCl	Sigma
Bacto peptone	Difco
Bacto yeast extract	Difco
Blocking reagent	Boehringer
Bovine serum albumin (BSA)	Sigma
Bradford reagent	Sigma
Bromophenol blue	Fluka
Calcium chloride	Merck
Chloroform	Merck
Citric acid	Merck
Complete mini protease inhibitor cocktail tablets	Roche
4’6-diamino-2-phenyl-indole (DAPI)	Serva
[α-^{32}P]-Dctp	Amersham
Dextran blue	Fluka
Dextran sulfate	Sigma
Diethylpyrocarbonate (DEPC)	Sigma

Name	Company
Dimethyl sulfoxide (DMSO)	Merck
Disodiumhydrogen phosphate	Merck
Dithiothreitol (DTT)	Sigma
dNTPs	MBI Fermentas
DO supplement	BD Biosciences
DPBS (PBS for cell culture)	BioWhittacker
Dulbecco's Modified Eagle's Medium (DMEM) with 4,5 g/l glucose	BioWhittacker
Ethanol	Merck
Ethidium bromide	Serva
Ethylenediaminetetraacetic acid (EDTA)	Merck
Fetal Bovine Serum (FBS)	Biochrom AG
Fluoromount-G	Science Services
Formamide	Merck
Formaldehyde	Roth
GeneRuler 1 kb and 100 bp DNA ladders	MBI Fermentas
Glacial acetic acid	Merck
Glass beads (425-600 μm).	Sigma
Glucose	Merck
L-Glutamine solution for cell culture	BioWhittacker
Glutaraldehyde	Sigma
Glycerol	Merck
Glycine	Sigma
Heat inactivated sheep serum (HISS)	Gibco
Heparin sodium salt	Sigma
Herring sperm DNA	Roche
Hydrogen peroxide	Merck
L-Histidine HCl monohydrate	Sigma
Hybridime human placental DNA	HT
4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid sodium (HEPES)	Sigma
L-Isoleucine	Sigma
Isopropanol	Merck
Kanamycin	Invitrogen
L-Leucine	Sigma

Name	Company
Lithium acetate	Sigma
Lithium chloride	Merck
L-Lysine HCl	Sigma
Magnesium chloride	Merck
Magnesium sulfate	Merck
Maleic acid	Merck
Manganese chloride	Sigma
β-Mercaptoethanol	Merck
Methanol	Merck
L-Methionine	Sigma
3-[N-Morpholino]propanesuflonic acid (MOPS)	Sigma
N,N,N ,N –Tetramethylethylenediamine (TEMED)	Invitrogen
Nonidet P40 (NP40)	Fluka
OPTIMEM with GLUTAMAX	Gibco
Paraformaldehyde (PFA)	Sigma
Pd(N)6 random hexamers	Pharmacia
Penicillin/streptomycin antibiotic solution for cell culture	Invitrogen
Phenol	Roth
Phenol red	Fluka
L-Phenylalanine	Sigma
Polyethylene glycol 3350 (PEG 3350)	Sigma
Polyethylene glycol 6000 (PEG 6000)	Merck
Polyfectamine transfection reagent	Qiagen
Potassium acetate	Merck
Potassium chloride	Merck
Potassium dihydrogen phosphate	Merck
Precision Plus Protein Kaleidoscope Standards	Bio-Rad
L-Proline	Sigma
Protector RNase Inhibitor	Roche
Ribonucleic acid from Baker’s yeast type III (tRNA)	Sigma
Rubidium chloride	Sigma
Salmon sperm DNA	Sigma
Sephadex G-50	Pharmacia

Name	Company
Sodium acetate	Merck
Sodium azide	Sigma
Sodium chloride	Merck
Sodium citrate	Merck
Sodium deoxycholate	Sigma
Sodium dihydrogen phosphate	Merck
Sodium dodecyl sulfate (SDS)	Serva
Sodium hydroxide	Merck
Tetramisole hydrochloride (levamisole)	Sigma
Thiamine hydrochloride	Sigma
L-Threonine	Sigma
Tris (hydroxymethyl)-aminomethane	Merck
Triton X-100	Roth
TRIzol reagent	Invitrogen
L-Tryptophane	Sigma
Trypton	Difco
Tween 20	Sigma
L-Tyrosine	Sigma
L-Uracil	Sigma
L-Valine	Sigma
Yeast nitrogen base without amino acids	Difco

2.1.2 Buffers and solutions

Aqueous solutions were prepared using autoclaved Millipore water. For sterilisation solutions were autoclaved or passed through a 0.45 µm filter (Millipore). Solutions used for Southern blot hybridisation are listed in Table 2. Buffers and solutions used for denaturing RNA gel and Northern hybridisation can be found in Table 3. Solutions used for yeast lysis and β-galactosidase assay are listed in Table 4. Solutions required for SDS polyacrylamide gel electrophoresis (SDS-PAGE), Western blot analysis and immunofluorescence are listed in Table 5. Solutions used for whole mount and section *in situ* hybridisation can be found in Table 6. Buffers used for preparing chemical competent *E.coli* cells are listed in Table 7.

Table 2 Solutions used for Southern blot hybridisation

Solution	**Components**
Nucleus buffer (pH 7.5)	10 mM Tris-HCl 25 mM EDTA 75 mM NaCl
20 × SSC (pH 7.0)	300 mM Sodium citrate 3 M NaCl
Denaturing solution	0.5 M NaOH 1.5 M NaCl
0.5 M Phosphate buffer (pH 7.4)	77.4 mM Na_2HPO_4 22.6 mM NaH_2PO_4
PEG hybridisation buffer	250 mM NaCl 125 mM Na_2HPO_4 1 mM EDTA 7% (w/v) SDS 10% (w/v) PEG 6000
5 × OLB (-dCTP)	0.1 mM each dATP, dGTP, dTTP 1 M HEPES 0.425 mM pd(N)6 25 mM $MgCl_2$ 250 mM Tris-HCl 0.36% (v/v) β-Mercaptoethanol
Stop solution (pH 7.5)	10 mM Tris-HCl 5 mM EDTA 2% SDS 0.1% Dextran blue 0.1% Phenol red
TES (pH 7.5)	10 mM Tris-HCl 5 mM EDTA 0.2% SDS
Washing buffer	40 mM Phosphate buffer 0.5% (w/v) SDS

Table 3 Solutions for Northern blot hybridisation

Solution	**Components/Company**
25 × MOPS solution	1 M MOPS 0.25 M NaAc 50 mM EDTA (pH 7.0)
Electrophoresis buffer	1 × MOPS solution 0.66 M Formaldehyde
Northern blot loading buffer	50% Formamide 2.2 M Formaldehyde 1 × MOPS solution 40 µg/ml Ethidium bromide
QuikHyb Hybridization Solution	Stratagene
Washing solution	2 × SSC 0.1% (w/v) SDS
High stringency washing solution	0.1 × SSC 0.1% (w/v) SDS

Table 4 Solutions used for yeast lysis and β-galactosidase assay

Solution	**Components**
Yeast lysis buffer	2% Triton X-100 1% SDS 100 mM NaCl 10 mM Tris-HCl (pH 8.0) 1 mM EDTA (pH 8.0)
Z-buffer (pH 7.0)	60 mM $Na_2HPO_4 \cdot 2H_2O$ 40 mM $NaH_2PO_4 \cdot H_2O$ 10 mM KCl $MgSO_4 \cdot 7H_2O$
Substrate solution	100 ml Z- buffer 0.27 ml β-Mercaptoethanol 33.4 mg X-Gal

Table 5 Solutions used for SDS-PAGE, Western blot analysis and immunocytochemistry

Solution	**Components**
10 × Protease inhibitor solution	1 tablet of complete mini protease inhibitor cocktail per 1 ml H_2O
Cell lysis buffer	150 mM NaCl 50 mM Tris-HCl (pH 7.5) 1% NP40 1 × Protease inhibitor solution
2 × Sample buffer	125 mM Tris-HCl (pH 6.8) 4% SDS 10% Glycerol 0.006% Bromophenol blue 2% β-Mercaptoethanol
Protein loading buffer	350 mM Tris-HCl (pH 6.8) 10% SDS 30% Glycerol 9,3% DTT 0.012% Bromophenol
1 × Electrophoresis buffer	25 mM Tris-HCl 250 mM Glycine 0.1% SDS
Western blot transferring buffer	192 mM Glycine 25 mM Tris-HCl 20% Methanol
1 × PBS (pH 7.3)	137 mM NaCl 2.7 mM KCl 10.1 mM Na_2HPO4 1.8 mM KH_2PO4
PBST	0.1% Tween 20 in 1 × PBS
Blocking buffer for Western blot	5% (w/v) nonfat dry milk in 1 × PBST
Blocking solution for immunocytochemistry	1 × PBS 10% FBS 0.05% NaN_3

Table 6 Compound solutions used for whole mount and section *in situ* hybridisation

Solution	**Components**
DEPC-H_2O	0.1% (v/v) DEPC in H_2O
DEPC-PBS	0.1% (v/v) DEPC in 1 × PBS
DEPC-PBST	0.1% (v/v) Tween 20 and 0.1% DEPC in 1 × PBS
PBST/Glycine	0.2% (w/v) Glycine in DEPC-PBST
PBST/Tetramisole	0.05 % (w/v) Tetramisole in DEPC-PBST
RIPA buffer	0.05% SDS 150 mM NaCl 1% (v/v) NP40 0.5% (w/v) Sodium deoxycholate 1 mM EDTA 50 mM Tris-HCl in DEPC-H_2O
PFA/Glutaraldehyde	4% PFA/PBS (pH 7.4) 0.2% Glutaraldehyde
20 × DEPC-SSC (pH 7.0)	300 mM Sodium citrate 3 M NaCl 0.1 % DEPC
Heparin	100 mg/ml Heparin in 4 × DEPC-SSC
Hybe buffer	50 % Formamide 0.1 % Tween 20 50 µg/ml Heparin 5 × DEPC-SSC (pH 4.5, adjusted with citric acid) diluted in DEPC-H_2O
SSC/FA/T	2 × SSC 50% Formamide 0.1 % Tween 20
5 × MABT	0.5 M Maleic acid (pH 7.5) 0.75 M NaCl 5 % Tween 20
RNase solution	0.5 M NaCl 10 mM Tris-HCl (pH 7.5) 0.1 % Tween 20
Alkaline phosphatase buffer	100 mM NaCl 50 mM $MgCl_2$ 100 mM Tris-HCl (pH 9.5) 0.05 % (w/v) Tetramisole 0.1 % Tween 20

Solution	Components
10 × Wash buffer	4 M NaCl 1 M Tris-HCl (pH 7.5) 0.5 M EDTA
TNE	10 mM Tris-HCl (pH 7.5) 500 mM NaCl 1 mM EDTA
NTMT (pH 9.5)	100 mM NaCl 100 mM Tris-HCl (pH 9.5) 50 mM $MgCl_2$ 0.1% Tween 20
Hybridisation solution (for paraffin sections)	10 mM Tris-HCl (pH 7.5) 600 mM NaCl 1 mM EDTA 0.25% SDS 10% Dextran sulfate 1 × Denhardt solution 200 µg/ml Ribonucleic acid from Baker's yeast type III 50% Formamide

Table 7 Solutions used for preparing chemical competent *E. coli* cells

Solution	Components
Buffer A	30 mM KAc 10 mM $CaCl_2$ 100 mM RbCl 50 mM $MnCl_2$ 15% Glycerol
Buffer B	10 mM MOPS/NaOH (pH 7.0) 75 mM $CaCl_2$ 10 mM RbCl 15% Glycerol

2.1.3 Media

2.1.3.1 Yeast media

Yeast media used for propagation the wild type strains L40 and AH109, as well as selective media required for the yeast two-hybrid assay are listed in Table 8.

Table 8 Yeast media and dropouts (DO)

Solution	**Components**
YPD medium (pH 5.8)	2% Bacto peptone 1% Bacto yeast extract 1% Glucose
YPD agar (pH 5.8)	20 g Agar per 1 l YPD medium
YPDA medium	2% Bacto peptone 1% Bacto yeast extract 2% Glucose 0.003% Adenine
YPDA agar	20 g Agar per 1 l YPDA medium
10 × DO supplement	0.03% (w/v) L-Isoleucine 0.15% (w/v) L-Valine 0.02% (w/v) L-Adenine hemisulfate salt 0.02% (w/v) L-Arginine HCl 0.02% (w/v) L- Histidine HCl monohydrate 0.1% (w/v) L-Leucine 0.03% (w/v) L-Lysine HCl 0.02% (w/v) L-Methionine 0.05% (w/v) L-Phenylalanine 0.2% (w/v) L-Threonine 0.02% (w/v) L-Tryptophane 0.03% (w/v) L-Tyrosine 0.02% (w/v) L-Uracil
10 × DO –T	10 × DO supplement lacking tryptophane
10 × DO –L	10 × DO supplement lacking leucine
10 × DO –TL	10 × DO supplement lacking tryptophane and leucine
10 × DO –THULL	10 × DO supplement lacking tryptophane, histidine, uracil, lysine and leucine
SD medium	0.67% Yeast nitrogen base without amino acids 2% Glucose 1 × appropriate DO supplement
SD agar	20 g Agar per 1 l SD medium

2.1.3.2 Bacterial media

E.coli strains XL1-Blue, DH5α, GM2163 and STBL4 were cultured in LB medium implemented with appropriate antibiotics, the strain HB101 was cultured in M9 minimal medium. Media with their components are listed in Table 9. Dropout solution (DO) used for preparing M9 minimal medium can be found in the previous section (Table 8).

Table 9 Bacterial media

Medium	**Components**
LB medium	10 g/l Trypton 5 g/l Bacto yeast extract 10 g/l NaCl
LB agar	20 g Agar per 1 l LB medium
M9 minimal medium	6 g/l $Na_2HPO_4 \cdot 2H_2O$ 3 g/l KH_2PO_4 1 g/l NH_4Cl 0.5 g/l NaCl 2 mmol/l $MgSO_4$ 0.1 mmol/l $CaCl_2$ 1 mmol/l Thiamine 1× DO –L 8 g/l Glucose 40 mg/l Proline 50 mg/l Ampicillin
M9 agar	16 g Agar per 1 l M9 minimal medium

2.1.3.3 Cell culture media

List of components for COS1 medium can be found in Table 10.

Table 10 Cell culture media

Solution	**Components**
COS1 medium	DMEM medium with 4,5 g/l glucose 5% FBS 1% L-Glutamine 1% Penicillin/streptomycin antibiotic solution

2.1.4 Enzymes

All restriction endonucleases for cloning and genomic DNA digestion were purchased from New England Biolabs, Invitrogen or MBI Fermentas. Reactions were performed in supplied reaction buffers following the manufacturers' instructions. Additional enzymes used are listed in Table 11.

Table 11 Additional enzymes

Enzyme	**Concentration**	**Company**
Taq DNA polymerase	5 U/µl	Promega
AmpliTaq DNA polymerase	5 U/µl	Perkin Elmer
Pfu DNA polymerase	2.5 U/µl	Stratagene
TaKaRa LA *Taq* polymerase	5 U/µl	TaKaRa
Expand Long Template Enzyme mix	5 U/µl	Roche
Klenow fragment	2 U/µl	Roche
Superscript reverse transcriptase	200 U/µl	Invitrogen
DNase I RNase-free	10 U/µl	Roche
Proteinase K	10 mg/ml	Invitrogen
T3 RNA polymerase	20 U/µl	Roche
T7 RNA polymerase	20 U/µl	Roche
RNase A	10 mg/ml	Roche
T4 DNA ligase	400 U/µl	Promega
Shrimp alkaline phosphatase (SAP)	20 U/µl	Fermentas
E. coli exonuclease I	1 U/µl	Fermentas
Calf intestine alkaline phosphatase	1 U/µl	Fermentas

2.1.5 Kits

All kits used are listed in Table 12.

Table 12 List of commercial kits

Kit	Company
BigDye terminator cycle sequencing ready reaction kit	PE Biosystems
10 × DIG RNA labelling mix	Roche
BM Purple AP Substrate	Roche
QIAprep spin miniprep kit	Qiagen
QIAgen plasmid midi and maxi kits	Qiagen
QIAquick gel extraction kit	Qiagen
Roti™-Lumin Chemiluminescence substrate	Roth

2.1.6 Vectors and universal primers

All vectors used for expression studies as well as those used for intermediate cloning steps are listed in Table 13. Multiple cloning sites and/or maps of modified vectors as well as of the vectors obtained from other labs (marked in the last column of the Table 13) can be found in the appendices 11.1, 11.2, 11.3 and 11.4. Vector-specific primers used for colony PCR and/or sequencing of inserts can be found in Table 14.

Table 13 Vectors

Vector	Resistance gene	Company/Remarks
pGEM®-T Easy Vector System I	Ampicillin	Promega
pCRII®-TOPO®	Ampicillin/ Kanamycin	Invitrogen
pVP16	Ampicillin/ Leucine	constructed by Stan Hollenberg*
pBMT116	Ampicillin/ Tryptophane	constructed by Paul Bartel and Stan Fields*
pGBKT7	Kanamycin/ Tryptophane	BD Biosciences
pcDNA-Flag	Ampicillin	modified pcDNA3.1 vector from Invitrogen*
pTL1-HA2	Ampicillin	modified pSG5 vector from Stratagene**

* A kind gift from the group of Prof. W. Birchmeier, Max-Delbrueck-Center, Berlin.
** A kind gift from the group of Prof. E. Wanker, Max-Delbrueck-Center, Berlin.

Table 14 Vector-specific primers

Vector	Primers pairs	Primer sequences	T_A
pGEM®-T Easy and pCRII®-TOPO®	M13for	GTAAAACGACGGCCAGTG	52°C
	M13rev	GGAAACAGCTATGACCATG	
pVP16	VP16F1	GGATTTACCCCCCACGACTCC	62°C
	VP16R1	AGGGTTTTCCCAGTCACGACGTT	
pBMT116	BTM116F1	TCAGCAGAGCTTCACCATTG	55°C
	BTM116R1	GAGTCACTTTAAAATTTGTATAC	
pGBKT7	Y2H-T7	TAATACGACTCACTATAGGGC	54°C
	Y2H-BD	TTTTCGTTTTAAAACCTAAGAGTC	
pcDNA-Flag	T7	AATACGACTCACTATAGGGAA	54°C
	BGHrev	TAGAAGGCACAGTCGAGG	
pTL1-HA2	pSG fw.	TCTGCTAACCATGTTCATGCC	58°C
	pSG rev.	GGACAAACCACAACTAGAATG	

2.1.7 Antibodies

All primary and secondary antibodies used in this study are listed in Table 15.

Table 15 Antibodies

Antibody	Company
Monoclonal anti-c-myc antibody	Invitrogen
Anti-HA antibody produced in rabbit	Sigma
Monoclonal anti-HA agarose conjugate	Sigma
Monoclonal LexA antibody	Clontech
Anti-Flag® M2 Affinity Gel	Sigma
Monoclonal anti-Flag antibody	Sigma
Polyclonal anti-Flag antibody	Sigma
Alexa Fluor® 546 goat anti-rabbit IgG	Molecular Probes
Alexa Fluor® 488 goat anti-mouse IgG	Molecular Probes
Anti-rabbit IgG (Goat), peroxidase conjugate	Oncogene
Anti-Mouse IgG (Goat), peroxidase conjugate	Oncogene
Anti-digoxigenin-AP	Roche

2.1.8 Primers

All primers used in this study were synthesised by MWG Biotech or by Metabion, Germany.

2.1.9 Genomic clones, ESTs and genomic and cDNA libraries

All human YAC, BAC and cosmid clones, as well as EST clones and the spotted human chromosome 2-specific cosmid library (Livermore) were obtained from the Resource Centre of the German Human Genome Project (RZPD).
Mouse embryonic cDNA library (from pooled stages E9.5 – E10.5) cloned in pVP16 vector was a kind gift from the group of Prof. W. Birchmeier, Max-Delbrueck-Center, Berlin.

2.1.10 Bacteria

For all recombinant DNA techniques, competent *Escherichia coli* strain DH5α or XL1-Blue cells (Stratagene) were used. Isolation of prey plasmids from the yeast two-hybrid screen was performed with the help of the *E. coli* strain HB101 (Promega). Non-methylated plasmids were isolated from the dam^-/dcm^- *E.coli* strain GM2163. Cloning of repetitive-rich sequences was performed with the help of the *E.coli* strain STBL4 (Invitrogen).

2.1.11 Yeast

Yeast strain L40 was obtained from Invitrogen, whereas AH109 strain was purchased from Clontech.

2.1.12 Mammalian cell lines

Transient transfection experiments were performed in COS1 cells (African green monkey kidney cells).

2.1.13 Mouse embryos

Mouse embryos were derived from crosses of wildtype C57Bl and Bl10 mice.

2.2 Methods

2.2.1 DNA isolation

Genomic DNAs from lymphoblastoid or fibroblast cell lines were extracted according to standard protocols (Sambrook et al. 1989). Briefly, approximately 1×10^8 cells were suspended in 10 ml of nucleus buffer (see Table 2), and lysed by addition of SDS to a final concentration of 0.5%. Cell lysates were subjected to overnight proteinase K digestion. DNA was extracted by phenol/chloroform, precipitated with 96% ethanol and washed twice with 70% ethanol.

Plasmid DNAs were isolated using QIAprep miniprep kit or QIAgen plasmid midi and maxi kits according to the manufacturer's instruction.

2.2.2 Fluorescence *in situ* hybridisation (FISH)

Metaphase chromosomes were prepared from the Epstein-Barr virus transformed lymphoblastoid cell line, which was derived from the patient with the translocation t(2;10)(q31.1;q26.3). YAC probes were selected from the Whitehead Institute database, and BAC clones were found with the use of the NIX programme available on the UK HGMP Resource Centre webpage. To select cosmid clones, screening of the spotted human chromosome 2-specific cosmid library (Livermore) was performed. A pool of PCR products selected from the breakpoint region of the patient was labelled with [α^{32}P]-dCTP by random priming and was preannealed with Hybridime at 65°C for 1.5 h. Hybridisation to the library filters was performed at the same temperature overnight. Afterwards, the filters were washed in 40mM phosphate / 0.5% SDS solution. Signals were detected by autoradiography and positive clones were selected for FISH.

DNAs from all clones were isolated according to standard protocols (Sambrook et al. 1989). Purified DNAs were labelled with either digoxigenin-11-dUTP or biotin-16-dUTP by nick translation and were hybridised to metaphase chromosomes of the patients. Signals were detected either by anti-digoxigenin or by FITC- or Cy3- conjugated avidin and were visualised by fluorescence microscopy, as described previously (Wirth et al. 1999).

2.2.3 Amplification of DNA probes for Southern blot hybridisation

DNA probes for Southern blot hybridisation were amplified by PCR with the *Taq* DNA polymerase (Promega) in the supplied buffer (1.5 mM $MgCl_2$ final concentration). PCR reactions were carried out in 50 µl reaction volumes and contained 0.2 µM of each primer pair, 0.2 mM dNTPs, 1 U *Taq* DNA polymerase and approximately 100 ng genomic DNA as template. Cycling conditions included an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at the specific annealing temperature, 1 minute at 72°C, and a final extension step at 72°C for 7 minutes. The PCR products were separated on a 1% agarose gel and purified using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's instruction. Selected probes used for Southern blot hybridisation, and the respective primer sequences, together with the annealing temperatures (T_A) for each PCR reaction are given in Table 16.

Table 16 Probes for Southern blot hybridisation

Probe	**Primer name**	**Primer sequence [5' → 3']**	**T_A**
538A12_49750 (694 bp)	538A12_49750for 538A12_49750rev	GCTTCCCATTGCAGGTGTAAA ATTACTGGTCATCAATATCTAGC	56°C
538A12_79400 (558 bp)	538A12_79400for 538A12_79400rev	AACTCAACATAAACTTTTCCAAAG GAATGTAAAATATAGACATTTGACATTG	56°C
300B2_80000 (611 bp)	300B2_80000for 300B2_80000rev	TCCTTTGCACACAGGTCCTTTC TGGTGACCATGGCAGTCATAC	58°C
300B2_84000 (766 bp)	300B2_84000_for 300B2_84000_rev	GAGAAGGACTAGAGAGGATG CACAGGTATTTGATATGTTGTCAGC	58°C
300B2_90000 (991 bp)	300B2_90000_for 300B2_90000_rev	GGAAGATGTTGAACAGGTGAGAG CACAGAAAGCACGTGGCTGC	60°C

2.2.4 Isotope-labelling of probes for Southern blot hybridisation

Gel purified DNA probes were denatured for 10 min at 95°C, chilled on ice and radiolabelled using random hexanucleotide primers in 1 × OLB buffer. Essentially, the reaction was carried out at 37°C for at least 1 hour in 25 µl volume containing 25 ng of DNA probe, 20 µCi [α-^{32}P]-dCTP and 2 U Klenow fragment. The reaction was stopped by adding 50 µl of stop solution. To remove the excess of non-incorporated dNTPs and random hexamers, the

reaction mixture was separated by passing through a Sephadex G-50 column and the labelled probe was eluted with TES buffer. Solutions used in this section are listed in Table 2.

2.2.5 Southern blot hybridisation

Genomic DNAs from the patient with the translocation t(2;10)(q31.1;q26.3) and from controls were digested with appropriate restriction enzymes and separated on 0.7% agarose gels. Subsequently, gels were pre-incubated in denaturing solution twice for 30 minutes each, and the DNAs were capillary transferred to nylon membranes (Roth) using 10 × SSC. After an overnight transfer, membranes were incubated for 10 minutes in 50 µM phosphate buffer and the DNAs were fixed by UV crosslinking at 254 nm for 2 min. For hybridisation, membranes were pre-hybridised in PEG hybridisation buffer supplemented with 0.1 mg/ml denatured herring sperm DNA as blocking reagent for at least 2 hours at 65°C. Isotope-labelled DNA probes were denatured for 10 min at 95°C, chilled on ice and blocked with human placental DNA Hybridime in PEG hybridisation buffer for 1 hour at 65°C before added to the pre-hybridisation solution. Membranes were hybridised with isotope-labelled probes at 65°C overnight, washed with washing buffer and exposed to Kodak X-Omat AR film at –80°C for 1-7 days. Solutions used in this section are listed in Table 2.

2.2.6 Genomic walking

For breakpoint cloning, genomic walking was performed, as described elsewhere (Siebert et al. 1995). Genomic DNAs from the patient and a control were digested with appropriate restriction enzymes, phenol/chloroform extracted, and ethanol precipitated. Approximately 1 µg of each digested DNA was ligated to pre-annealed adaptor oligos using T4 DNA ligase (Promega) in supplied buffer. After overnight ligation at 15°C, the reaction was stopped by heating at 70°C for 10 min. 1 µl of each ligation mixture was used as a template in the first round of nested PCR. All PCR reactions were carried out with TakaRa LA *Taq* polymerase (TaKaRa) in 50 µl volume of 1× PCR buffer provided by the manufacturer. The first round PCR contained 30 µM of adaptor primer AP1 and one of the sequence-specific primers: 84364for [for der(10)] or 85206rev1 [for der(2)]. Cycling condition comprised an initial denaturation step at 94°C for 3 minutes, followed by 30 cycles of 40 seconds at 94°C, 40 seconds at 60°C, 3 minutes at 72°C, and a final extension at 72°C for 15 minutes. 1 µl of

each first round PCR product was used as a template in the second round of PCR reaction, which contained 30 µM each of nested adaptor primer AP2 and one of sequence-specific primers 84715for [for der(10)] or 85163rev2 [for der(2)]. Cycling condition for the second round PCR was essentially the same as the first round, except for the annealing temperature, which was adjusted to 58°C. Primers and adaptor sequences are listed in Table 17.

Table 17 Primer/Adaptor sequences for genomic walking

Name	**Sequence [5' → 3']**
Adaptor-long	CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT
Adaptor-short*	AATTACCTGCCCGG
AP1	GGATCCTAATACGACTCACTATAGGGC
AP2	TATAGGGCTCGAGCGGC
84364for	CAGATTGTGATTAGATCAGGAG
84715for	GACTTAAAATTGCAGCGTGTGTTTC
85206rev1	GTGTATCTATCTGAGCTCCATG
85163rev2	TTCAGCCTTAAGTCAAAATGTTGG

* 5' phosphate modification

Amplified fragments were isolated from 1% agarose gels, purified with Qiagen columns, subcloned into pGEM-T Easy vector and sequenced using M13 universal primers.

2.2.7 Screening for mutations in the *HOXD13* gene in the translocation patient

In order to screen the *HOXD13* gene in the translocation patient, several PCR reactions were performed using patient genomic DNA as template. First, a 172 bp long fragment encoding the alanine stretch was amplified using HsHOXD13for_1 and HsHOXD13rev_1 primers. Amplification reaction contained 1× PCR Buffer II (Perkin Elmer), 0.75 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 µM of each primer and 2.0 units of AmpliTaq DNA polymerase (Perkin Elmer) in a final volume of 50 µl. Initial denaturation was at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for 1 minute, and extension at 72°C for 2 minutes. The product was purified with the QIAQuick Gel Extraction Kit (Qiagen) and sequenced with HsHOXD13for_1 and HsHOXD13rev_1 primers. Later,

additional PCR reactions were performed in order to screen the whole coding sequence and the splice site acceptor of the *HOXD13* gene. The very 5' end of the coding sequence was amplified with HsHOXD13for_b and HsHOXD13_rev_a primers using patient genomic DNA as template. The reaction contained 10 µl FailSafe PCR Premix J (Epicentre), 2.5 mM of each primer and 0.3 µl Expand Long Template Enzyme mix (Roche) in a final volume of 20 µl. Cycling conditions included an initial denaturation step at 94°C for 2 min, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 61°C, 2 minutes at 68°C, and a final extension step at 68°C for 10 minutes. The 3' end of exon 1, the part of the intron containing the splice site acceptor, and the whole exon 2 of *HOXD13* gene were amplified in three PCR reactions using the following primer pairs: HsHOXD13_for_a and HsHOXD13rev, HsHOXD13for2 and HsHOXD13rev2, HsHOXD13for3 and HsHOXD13rev3, respectively. PCR reactions contained 1 × PCR buffer (Promega), 0.4 µM of each primer, 0.2 mM dNTP, 8% DMSO and 1 unit *Taq* DNA polymerase (Promega) in a total volume of 50 µl. Initial denaturation was at 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at the appropriate temperature (T_A) for 30 seconds, and extension at 72°C for 45 seconds.

All PCR products were separated on the 1% agarose gel, purified using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's instruction, subcloned into pGEM-T Easy vector and sequenced using M13 universal primers. Primer sequences used for *HOXD13* amplification are listed in Table 18.

Table 18 Primers used for amplification the *HOXD13* gene

Primer name	**Primer sequence [5' → 3']**	**PCR product***	**T_A**
HsHOXD13for_b HsHOXD13_rev_a	GGAGCTGGGACATGGACG AATGCGTCCC GGCGAACAC	8881-9027 bp (147 bp long)	61°C
HsHOXD13for_1 HsHOXD13rev_1	CAGTGCCGCGGCTTTCTCTC CTACAACGGCAGAAGAGGAC	8982-9153 (172 bp long)	60°C
HOXD13_for_a HsHOXD13rev	CTCGTCGTCGTCCTCTTCTG GACATACGGCAGCTGTAGTAGC	9125-9289 bp (165 bp long)	62°C
HsHOXD13for2 HsHOXD13rev2**	CTACTACAGCTGCCGTATGTCG TCGGTCCCTTTTCTCCATC	9269-9675 bp (407 bp long)	58°C
HsHOXD13for3** HsHOXD13rev3	AGCTAGGTGCTCCGAATATCC AAGCTGTCTGTGGCCAACC	10405-10739 bp (335 bp long)	58°C

* Positions corresponding to the *HOXD13* gene sequence, GenBank entry AB032481.

** Primers are lying in the intron of the *HOXD13* gene.

2.2.8 RNA isolation

Total RNAs from lymphoblastoid or fibroblast cell lines, as well as the RNA from mouse embryos were isolated using TRIzol reagent (Invitrogen) according to the manufacturer's protocol.

2.2.9 Analysis of *MGMT* expression by Northern blot

In order to analyse *MGMT* expression in lymphoblastoid cell lines total RNAs from the translocation patient and from control cell lines were isolated. Approximately 10 µg of each RNA was mixed with 2.7 volumes of Northern blot loading buffer and denatured for 10 minutes at 65°C. Afterwards, samples were chilled on ice, loaded onto a denaturing 1% agarose gel containing 2.2 M formaldehyde and 1 × MOPS and separated at 28 V for 16–24 hours. After electrophoresis the gel was washed 3 times with DEPC-H_2O in order to remove formaldehyde, and the RNA was capillary transferred to the nylon membrane Hybond™-XL (Amersham Pharmacia Biotech) using 10 × SSC. After an overnight transfer, the RNA was fixed in UV Stratalinker 1800 (Stratagene) for 2 minutes. In order to remove the rests of formaldehyde the membrane was subsequently baked for 2 hours at 80°C.

The *MGMT* and *G3PDH* probes were amplified from total RNA isolated from control cell lines. About 2.5 µg of the RNA was used for the first strand synthesis with Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's protocol. Probes were PCR-amplified using specific primers (see table Table 19) and 1µl of the prepared cDNA as a template. Cycling conditions and purification of PCR products were basically the same as described in section 2.2.3. Labelling and purification of radioactive probes were performed as described in section 2.2.4.

Table 19 Probes used for Northern hybridisation

Probe	Primer name	Primer sequence [5' → 3']	T_A
MGMT (403 bp)	MGMT_ex.2_for MGMT_ex.4_rev	GGACA AGGATTGTGA AATGAAACG TCTCATTGCT CCTCCCACTG	58°C
G3PDH (454 bp)	G3PDH_for G3PDH_rev	GACCACAGTCCATGCCATCACT GTCCACCACCCTGTTGCTGTAG	58°C

Prehybridisation and hybridisation of radioactive probes was performed in QuikHyb® Hybridization Solution (Stratagene) at 68°C for 30 minutes and 3 hours, respectively. Washing was performed according to the manufacturer's protocol. Signals were detected on the phosphoimager. Intensity of the *MGMT* signals was compared to the intensity of the *G3PDH* (housekeeping gene) signals using ImageQuant software (Amersham Biosciences).
MGMT expression in different human tissues was examined by hybridisation of the *MGMT* probe to the human multiple tissue Northern blot (Clontech) and human fetal multiple tissue Northern blot (Clontech). Hybridisation and washing were performed as described above.

2.2.10 RT-PCR analysis

Total RNAs from cell lines, human brain or mouse embryos (E16.5) were digested with the RNase-free DNase I (Roche) in the supplied buffer, according to the manufacturer's protocol. Reverse transcription (RT) reactions were performed according to the manufacturer's instruction using the Superscript reverse transcriptase (Invitrogen) or water (for RT negative control). Efficiency of the first strand synthesis was subsequently controlled by amplification of the housekeeping gene *G3PDH* on cDNA templates. High quality cDNAs were used as templates for expression studies of homologous regions identified in the vicinity of the *HOXD* cluster, as well as for analysis of human ESTs found close to the chromosome 2 breakpoint region in the translocation patient. All RT-PCR reactions were carried out in 50 µl reaction volumes and contained 1 × reaction buffer (1.5 mM $MgCl_2$ final concentration), 0.2 µM of each primer pair, 0.2 mM dNTPs, 1 U *Taq* DNA polymerase and 2 µl of proper cDNA as template. Cycling conditions included an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at the specific annealing temperature, 1 minute at 72°C, and a final extension step at 72°C for 7 minutes. After the first PCR reaction, 2 µl of each product was used for the second round of PCR with the same primers and reaction conditions. In parallel, reactions with the RT negative control were performed in order to make sure that PCR products did not derive from genomic contamination of the cDNA. PCR products were analysed on 1 % agarose gels.
In general, analyses of conserved elements on chromosome 2 were performed using RNA from mouse embryos, whereas human ESTs were amplified either on reported source tissue or on cell lines, in case when the source tissue was not accessible. Normally, primers were designed according to both ends of tested sequences. In case amplification with such primer

pairs did not result in a specific product, inner primers (which did not allow amplification of the whole sequence of interest) were ordered and used for RT-PCR experiments.

Primers used for human EST expression analysis can be found in Table 20. Sequences of primers used for analysis of conserved regions together with the expected length of PCR products and annealing temperatures are listed in Table 21.

Table 20 Primers used for expression analysis of EST sequences on chromosome 2q31

EST	**Primer name**	**Primer sequence [5' → 3']**	**T_A**
AI075926	Hs_EST8_for Hs_EST8_rev1	GTAGTGATACTCAGAAACTGACAC AGAGGGATCATTTACAGCAGG	56°C
AW850653	Hs_EST10_for1 Hs_EST10_rev1	ATTGCTAACCTAATGTGTGAGC GTGGTATGGACCAAAAACCAG	58°C
BG980989	Hs_EST5_for1 Hs_EST5_rev1	AAGTCACATTTAATGGGAGGATC CAGGACATTCTGATGCTAAACG	58°C
BG980131	Hs_EST3_for Hs_EST3_rev	TCGTTGGGAAGATCAAATGAG GGTTCCCATGGTCTGTTTC	56°C
BG979719 and BG979037	Hs_EST6/16_for Hs_EST6/16_rev	AATGCTCCATGCATACCCAAC CAGGGCCACTATAATGTCTC	58°C
BF815673	Hs_EST4_for1 Hs_EST4_rev1	CTTCTATATTCCAGTTATGATTTCAAGG ACACATCTGGAATGAAAACATAAACC	54°C
AW858552 and AW858470	Hs_EST2_for Hs_EST15_rev	CAGGACAATTAGTTCTAGAAGG CTCCTATTTAAATTGCTCACACTC	56°C
AW167235	Hs_EST17_for Hs_EST17_rev	TGACTATCTCTGTTGGGTAGAG GTGGGTTTTCAGAAATCTGAGC	
BE064736 and BE064727 and BE065063 and BE064976	4ESTfor 4ESTrev	TAACCCTATGTAGCTGGTGC CACACAATGCTCTCTCATTGG	58°C
BG952464	Hs_78_for Hs_78_rev2	CACATTCCTCCTCCTTCATTC GATGTGCACTGTCCATTTTAGATC	56°C
AW937867 and AW937872	Hs_EST9/14for Hs_EST9/14rev	GCAGAAATTCTTTGTGAAAGGAG ACTGTACACACAGGATTGTGC	58°C

Table 21 Primers used for expression analysis of high homology regions (HHR)

HHR	Primer name	Primer sequence [5' → 3']	PCR product	T_A
1.	Mm_46_for Mm_46_rev	GTACTTGGTAGCCCTTCAAG CTGAAAAATGAACTGCTTCTAGC	185 bp	56°C
2.	Mm_47_for rej57rev	CTCTCCCTAGCCCACTTAG CTATTACCCTGGCGAAACC	390 bp	56°C
3.	Mm_48_for1 Mm_48_rev1	CATGCTAACAACGCCCCTAG AGCCCCACTTTGCCTTCC	224 bp	56°C
4.	Mm_48_for2_new Mm_48_rev2_new	GGAATTGCTCATTAAATTGCCTAC GAATTTGACTTGGGGTGGACT	98 bp	56°C
5.	Mm_361_for1 Mm_361_rev1	TTTGGAAGATGTATAATGCAATATAAAC GAGCAATTGAAAGTAATATGGGCA	244 bp	56°C
6.	Mm_50_for1 Mm_50_rev1	CAATATATTTTTCAATCCTGACTGTTGG TGAGGCAGTGGCACTAAATG	85 bp	60°C
7.	Mm_50_for2 Mm_50_rev2_new	ATAGAGAAAATGCGTAGATGTGCC TTCAGCTACAGAGAGCTCCCCCA	94 bp	60°C
8.	Mm_51_for1 Mm_51_rev1	GTTTTACACTCTAAATGAAAGCCAC TACCTTACTTTAGCAGCGTGG	92 bp	56°C
9.	Mm_365_for1 Mm_365_rev1	CCATTGTCATGCAAGCACAG GGGATTTGTCCTTTTTATCTAGTC	341 bp	60°C
10.	Mm_54_for rej67rev	CATGGCGAATTCAGTATGAAGG GCAGCTATTTAGCTCGAATTGG	275 bp	58°C
11.	Mm_65_for1_new1 Mm_65_rev1_new1	ATGTGTCTAGGAAGGACATGC AGAGTTCAGCGACATTTGCCTC	299 bp	58°C
12.	Mm_65_for2 Mm_65_rev2	GACAACAATGCCTCGGAAG GACTTCTGAGCTTTCAGAGTG	133 bp	56°C
13.	Mm_68_for1 Mm_68_rev1	CAGCTGTACCCATGAGCATC GGGAGACAATGAGAACGTTTG	130 bp	56°C
14.	Mm_68_for2 Mm_68_rev2	TCATGCACTTCGTACACCTG TTCCTTGACCTAGAAATACGATAC	106 bp	56°C
15.	Mm_70_for Mm_70_rev	GAGTCAGAAAATTGCGATTCATTCC CTTGGCTCCAACAGAGTAGC	196 bp	56°C
16.	Mm_73_for Mm_73_rev	GGGACATGCCATATATTAGCAG GATACAGGAATCGAAGAAAACAGG	446 bp	56°C
17.	Mm_77_for1 Mm_77_rev1	CCTTGTTTTCTTCTGGCCATTTC TTACCAGTGTGACAGTATTAGAAAG	95 bp	56°C

HHR	Primer name	Primer sequence [5' → 3']	PCR product	T_A
18.	Mm_77_for2 Mm_77_rev2	CCTGTCTCCAGAGATGGATC GTCTCTGATGTGTAAGATCAAGC	143 bp	56°C
19.	Mm_78_for Mm_78_rev	GAAGGCACCTCTCACATAAG GTGCTCTATAATTCTACGTGAAG	532 bp	56°C
20.	Mm_79_for Mm_79_rev	CAGAGATCACCCTCTTTCAG TACATAGAATTGTCTTCTGGACC	106 bp	56°C
21.	Mm_80_for Mm_80_rev	GTAAAAGCAGCACACAGTAGTC CTTAGATCTTGATTCTACTCCAACC	221 bp	56°C
22.	Mm_81_for_new Mm_81_rev	GTAGTTTAATGCCCAGCGG GTGTAAAGCAGTTGCTAGAAATC	178 bp	58°C
23.	Mm_86_for Mm_86_rev	TTCTCCGTGAAAGGAGGAG GTGCTGTCACTGAATTCCTTG	158 bp	56°C
24.	Mm_89_for Mm_89_rev	TCTGTCCCTTCTCAAATGGAAG GTGGAGTTAAAGTAGACATATGAGC	86 bp	56°C
25.	Mm_90_for_new2 Mm_90_rev	ATCCATTCTCTGCCCCACTC ATCAGAGGTATTATGGGTGAGC	89 bp	58°C
26.	Mm_95_for Mm_95_rev	CTTGTGAGGCATTAAGATGTTCTC GGTGCTGCCGTGTTAGTATG	116 bp	56°C
27.	Mm_97_for Mm_97_rev	CCCAGTGGCCTTTCTAGTC GGTATGTAGGGCAGGAATATG	118 bp	56°C
28.	Mm_98_for Mm_98_rev	CCTTATTCTAGAATGGCCCAG CTTTGTAAGAGCCCAGAAATGG	122 bp	56°C
29.	Mm_100_for1 Mm_100_rev1	GAGCTAAATTTCCTAGATGGTTATG CTTGGAAAGGTTCTAGGTGTC	127 bp	56°C
30.	Mm_100_for2 Mm_100_rev2	CATAAAACCGGGGCTCCCA CTTATGGGGACTAATGACTAATTCC	226 bp	56°C
31.	Mm_105_for Mm_105_rev_new1	CCCTGCAAATTATAAGCAGCTC CTAATGAAAAGCAGAGGCAAATGAG	126 bp	58°C
32.	Mm_105_for_new1 Mm_105_rev	CACATTTCCTATCAGCCCCT TGCAGTTTGTGACTCCCAAAG	470 bp	56°C
33.	Mm_106_for1 Mm_106_rev1	CTGATCTTTTCCTCTAGCCAG CCTGAGCCCAAGTATTCAC	190 bp	56°C
34.	Mm_106_for2 Mm_106_rev2	TCTTCCCAGGAATGCATCTG CCCTTTTCCTGTTGCTATTTCC	356 bp	56°C
35.	Mm_106_for3 Mm_106_rev3	ACGGACGGTTGTTACACTAG TTTCTGAATCATGCTGACGACG	394 bp	56°C

HHR	Primer name	Primer sequence [5' → 3']	PCR product	T_A
36.	Mm_107_for Mm_107_rev	TGATCAACTTTACTCCTGTTGCTT CTTTATGATATCAGTCACACAG	125 bp	52°C
37.	Mm_108_for Mm_108_rev	TGTCAGGAGTACTAGGAAATGG AATGGATCCTCTCTAGGGGTGT	147 bp	58°C
38.	Mm_111_for Mm_111_rev_new	CATGTTTTGAGAGGTCAACAATGTC CTTCACAAGGAGCCTCAGATG	110 bp	58°C
39.	Mm_113_for2 Mm_113_rev	TCTAAGTAAGAGAACAGATGTGG GTACTTCCTTAATGTAAAGCTCG	92 bp	56°C

2.2.11 Construction of yeast and mammalian expression vectors

DNA manipulations were carried out using standard techniques (Sambrook et al. 1989). In order to facilitate cloning, appropriate restriction sites were incorporated into gene-specific primers used for PCR amplification. Amplified DNA fragments were cloned via A overhangs into the pCRII-TOPO vector as intermediate products or digested with relevant restriction enzymes and cloned directly into proper vectors. In all cases a *Taq*/*Pfu* polymerase mix (*Taq*:*Pfu* in proportion 36:1) was used for amplification in order to improve the fidelity. Sequence identity of each construct was verified by sequencing.

2.2.11.1 Hoxd13 cloning

The yeast two-hybrid bait construct LexA_Hoxd13-HD was cloned in two steps. First, the Hoxd13 fragment lacking the homeodomain (*Hoxd13-HD*) was amplified using primers listed in Table 22 on the cDNA template derived from mouse embryonic stage E14.5. Afterwards, the PCR product and the pBTM16 vector carrying the LexA DNA binding domain were digested with *Eco*RI and *Sal*I restriction nucleases, ligated using T4 DNA ligase (Promega) according to the manufacturer's protocol and transformed into *E. coli*. Similarly, for construction of the vector GAL_Hoxd13-HD, the *Hoxd13-HD* fragment was amplified using specific primers (see Table 22) and cloned into the pGBKT7 vector carrying the GAL DNA binding domain via *Eco*RI and *Sal*I sites.

In order to clone *Hoxd13-HD* into the pTL1-HA2 vector, an *Eco*RI-site-containing 5' primer and a stop codon- and the *Xho*I-site-containing 3' primer were used for amplification of the insert (Table 22). Following restriction digestion with *Eco*RI and *Xho*I enzymes, the PCR product was ligated with the vector, giving rise to the Hoxd13-HD-pTL1-HA2 construct.

All PCR reactions were performed in 1 × PCR buffer (Fermentas) and contained 1.5 mM $MgCl_2$, 1 µM each primer, 8% DMSO and 0.1 mM dNTPs. Cycling conditions included an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at the specific annealing temperature, 1 minute at 72°C, and a final extension step at 72°C for 10 minutes.

pTL1-HA2 constructs carrying the wild type *Hoxd13* gene (wtHoxd13-pTL1-HA2), as well as the *Hoxd13* gene with the expanded stretch for additional 14 Ala residues (Hoxd13+14Ala-pTL1-HA2) or with the alanine stretch reduced to two residues only (Hoxd13_2Ala-pTL1-HA2) were kindly provided by A. Albrecht (MPI for Molecular Genetics, FG Mundlos).

Table 22 Primers for amplification and cloning of *Hoxd13*

Cloned fragment	**Primer**	**Sequence [5' → 3']***	**PCR product****	**Cloning vector**
Hoxd13-HD	EcoRI-HoxD13-For	GAGAATTCGGAATGAGCCGCTCGGGGACT	Hoxd13 (1-816 bp)	pBTM116
	SalI-HoxD13-Rev	CTGTCGACTCCCCGTCGGTAGACGCA		
Hoxd13-HD	EcoRI-HoxD13-For	GAGAATTCGGAATGAGCCGCTCGGGGACT	Hoxd13 (1-816 bp)	pGBKT7
	Hoxd13-SalI-rev-sto	CTGTCGACTTATCCCCGTCGGTAGACGCA		
Hoxd13-HD	EcoRI-HoxD13-For	GAGAATTCGGAATGAGCCGCTCGGGGACT	Hoxd13 (1-816 bp)	pTL1-HA2
	Hoxd13-XhoI-rev-sto	CTCTCGAGTTATCCCCGTCGGTAGACGCA		

* Recognition sites for restriction endonucleases are underlined within primer sequences.
** Positions correspond to the *Hoxd13* mRNA sequence NM_008275.

2.2.11.2 Peg10 cloning

At the beginning, *Peg10* fragments found in the yeast two-hybrid screen were cloned into the mammalian expression vector pcDNA-Flag, later however the full ORF2 of *Peg10* was amplified and cloned into the same vector.

The yeast two-hybrid inserts were cut out of the respective prey plasmids using *Not*I restriction endonuclease and the inserts were cloned directly into the *Not*I-digested and dephosphorylated pcDNA-Flag vector.

For cloning ORF2 of *Peg10*, inserts were amplified on mouse cDNAs from embryonic stages E9.5, E11.5, E12.5 and E13.5 using specific primers with incorporated recognition sites for restriction nucleases *Not*I or *Xba*I. Following digestion with *Not*I and *Xba*I, PCR products were ligated into the pcDNA-Flag vector. Primer sequences, annealing temperature (T_A) and length of amplicon can be found in Table 23.

Table 23 Primers used for *Peg10* amplification and cloning

Primer	**Sequence [5' → 3']***	**T_A**	**PCR product****
Peg10_Not_ORF2_for	TTGCGGCCGCTTATGCTCCA GATTCATATGCCGG	60°C	Peg10-ORF2 (1502-3172 bp)
Peg10_Xba_ORF2_R	GGTCTAGACTACGCAGCAC TGCAGGATG		

* Recognition sites for restriction endonucleases used for cloning the PCR products are underlined within primer sequences.
** Positions correspond to the *Peg10* mRNA sequence AB091827.

PCR products as well as plasmids carrying ORF2 inserts were sequenced with gene-specific primers (see Table 24). Additionally, recombined plasmids were sequenced with vector-specific primers (see Table 14) in order to confirm the cloning boundaries.

Table 24 *Peg10*-specific primers used for sequencing

Primer	**Sequence [5' → 3']**	**Primer binding site***	**T_A**
Peg10_seq4_for	CAATTGCCTCGGGCCCAATC	1665–1684 bp	60°C
Peg10_seq5_for	CCGCATCAGTATCCGCATCC	2021-2040 bp	60°C
Peg10_seq6_for	CATATGAATCCGGATCCACATCAC	2276–2299 bp	60°C
Peg10_seq7_for	TTGACCCTAACATTGAGATGATTCC	2619–2643 bp	60°C
Peg10_seq8_rev	GTCCACGAAATTTCGCAGAGC	2702–2722 bp	60°C
Peg10_seq9_rev	GACAAATTCACCATAGCTTGCCAG	2900–2923 bp	60°C

* Positions correspond to the *Peg10* mRNA sequence AB091827.

2.2.11.3 Cloning of other candidate genes identified in the yeast two-hybrid screen

Fragments of *Dlxin-1*, *Wtip*, *Limk1*, *Limd1* and *Cnot3* found in the yeast two-hybrid screen, were subcloned into the pcDNA-Flag vector basically as described for *Peg10* yeast two-hybrid fragments. The inserts were cut out of the respective prey plasmids using *Not*I restriction endonuclease and cloned directly into the *Not*I-digested and dephosphorylated pcDNA-Flag vector. All fragments were cloned in frame and corresponded to the following positions, for *Dlxin-1*: 1252–1600 bp (GenBank entry AB029448), for *Wtip*: 706–1177 bp (GenBank entry NM_207212), for *Limk1*: 380–803 bp (GenBank entry NM010717), for *Limd1*: 1974–2287 bp (GenBank entry NM_013860), for *Cnot3*: 639–941 bp (GenBank entry NM_146176).

The full length *Wtip* sequence cloned into the pcDNA-Flag vector (Wtip-pcDNA-Flag) was a kind gift of N. Verhey van Wijk (MPI for Molecular Genetics, FG Mundlos).

2.2.12 Small scale yeast transformation

LexA_Hoxd13-HD and the empty pBTM116 vector were transformed to the yeast cells L40 based on the protocol from the Yeast Protocols Handbook (Clontech). Briefly, a single colony of the wild type L40 yeast strain was used to inoculate 40 ml of YPD liquid medium supplemented with 50µg/ml ampicillin. Flasks were incubated overnight at 30°C with vigorous shaking. On the next day approximately 10-20 ml of the overnight culture was diluted in 300 ml of the fresh YPD medium, so that the OD_{600} of the new culture was between 0.2 and 0.25. Yeast cells were cultivated at 30°C for 3-5 hours until the OD_{600} reached 0.5-0.7. Furthermore, the cells were pelleted by centrifugation for 5 minutes at 1000 g at room temperature, washed with 40 ml sterile water and centrifuged again under the same conditions. The pellet was resuspended in 1.5 ml sterile 100 mM LiAc. Competent cells prepared in this way were used for heat shock-based transformation. Shortly, 50µl of competent yeast cells were mixed together with 0.2 µg of the appropriate plasmid, 50 µg of the denatured carrier DNA (salmon sperm DNA) and 300 µl sterile 40% PEG/100mM LiAc solution. All components were vortexed and incubated for 30 minutes at 30°C. Afterwards 35 µl DMSO were added and the tubes were incubated for 15 minutes at 42°C, followed by a short incubation on ice (heat shock). Yeast cells were shortly centrifuged, the pellet was resuspended in 100 µl sterile water and the suspension was streaked on SD agar plates lacking tryptophane (SD –T).

The construct GAL_Hoxd13-HD and the empty vector pGBKT7 were transformed into *ADE2*-deficient AH109 yeast. In order to culture the AH109 strain, YPDA medium was used. Preparation of yeast competent cells as well as transformation were performed according to the protocol described above.

2.2.13 Preparation of yeast protein extracts

A 10 ml liquid culture was prepared from each transformed yeast strain and was incubated in appropriate medium at 30°C overnight. On the next day, OD_{600} was measured and the amount of yeast culture corresponding to $OD_{600} = 1$ was centrifuged for 30 seconds at 8000 rpm in a table centrifuge. The supernatant was discarded and the pellet was resuspended in 100 µl 2 × Sample buffer (see Table 5). Samples were boiled for 5 minutes, afterwards cooled on ice and stored at –20°C.

2.2.14 Yeast two-hybrid screen

The yeast two-hybrid library screen was carried out using the LexA system. The yeast L40 strain carrying the vector LexA_Hoxd13-HD was transformed with the mouse cDNA library from pooled embryonic stages E9.5 – E10.5 constructed in the pVP16 vector carrying the VP16 activation domain. The screening protocol resembles the small-scale yeast transformation protocol. In short, one colony of the LexA_Hoxd13-HD positive yeast was cultured overnight in 5 ml of the SD –T medium at 30°C. The primary culture was subsequently transferred into a bigger flask containing 100 ml SD –T medium and was incubated overnight under the same conditions. On the third day, the culture was diluted in 300 ml YPD so that the OD_{600} reached 0.2 – 0.25. Yeast cells were grown approximately 5 hours until the culture reached $OD_{600} = 0.5$. Afterwards, cells were centrifuged for 5 minutes at 1000 g, washed with 10 ml 100 mM LiAc, pelleted again and resuspended in 2 ml of 100 mM LiAc. 50 µl of yeast competent cells were used for a mini transformation with the empty pVP16 vector (negative control) according to the protocol described in section 2.2.12. After transformation yeast suspension was in parallel streaked on SD plates lacking tryptophane and leucine (SD –TL), and on SD agar plates lacking tryptophane, histidine, uracil, lysine and leucine (SD –THULL).

At the same time 2 ml of yeast competent cells were mixed with 200 µg of the mouse cDNA library and with 3 mg of the denatured carrier DNA (salmon sperm DNA), and kept at the

room temperature for 3 minutes. After adding 20 ml of 40% PEG/100 mM LiAc solution yeast cells were incubated for 30 minutes at 30°C with shaking (200 rpm). Subsequently 2 ml DMSO was added and cells were delicately shaked for 15 minutes at 42°C in order to keep the whole volume equally warm. Later, cells were cooled for 1 minute on ice (heat shock) and 400 ml of warm YPD medium supplemented with 50 µg/ml ampicillin was added. Following incubation for 1h at 30°C, cells were pelleted at 1000 g for 5 minutes, washed in 40 ml sterile H_2O, pelleted once more and resuspended in 5 ml H_2O. 5 µl of transformed cells were used to prepare serial solutions 1:10, 1:100, 1:1000, 1:10000 and 1:100000 which were plated on 12 × 12 cm plates containing SD medium lacking tryptophane and leucine (SD –TL) (control of transformation efficiency). The rest of the cells was plated on SD medium lacking tryptophane, histidine, uracil, lysine and leucine (SD –THULL) in order to identify protein-protein interactions. The clones surviving from the nutrition selection (HIS3-positive colonies) were re-streaked three times on the same SD –THULL medium in order to reduce the number of false positives, and later they were tested for expression of the second reporter gene (see section 2.2.15). Different media used during the yeast two-hybrid screen are listed in Table 8.

2.2.15 LacZ colony filter assay

In order to screen HIS3-positive yeast colonies for the expression of *LacZ* (the second marker gene indicating interaction between proteins) a colony filter assay has been performed. Briefly, 12 × 12 cm filters were cut out of Whatman paper 3MM and were put onto SD –TL plates on which positive yeast colonies and a negative control were growing (yeast transformed with the LexA_Hoxd13-HD construct and the empty pVP16 vector). Yeast cells that attached to the filters were permeabilised by repeated freeze/thaw procedure (three times freezing in liquid nitrogen for ~10 seconds each, and then thawing at room temperature). Fresh Whatman filters were put into clean 12 × 12 cm plates and were soaked with 7 ml of the substrate solution containig X-Gal. Filters with yeast colonies were placed on the pre-soaked filters in the plates and incubated at 37°C. Filters were monitored for the appearance of blue colour once per hour. The reaction was stopped when the negative control (yeast clone transformed with the bait and the empty prey vector) started to get the blue staining. Only colonies showing very intensive blue colour were considered as true positives. Solutions used in this section can be found in Table 4.

2.2.16 Analysis of prey inserts in double positive yeast clones

For all positive yeast clones that survived from nutrition selection and turned blue in the presence of the X-Gal, inserts of prey plasmids were amplified by colony PCR. In short, fresh yeast colonies were picked and dispersed in 50 µl H_2O by vigorous vortexing. Cell suspensions were subjected to three times freeze/thaw treatment (frozen for ~10 seconds in liquid nitrogen, and thawed at room temperature). Five µl of the supernatant containing the released plasmids were used as template for PCR. PCR reactions were carried out with VP16F1 and VP16R1 primers. Cycling conditions included an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 62°C, 1 minute at 72°C, and a final extension step at 72°C for 7 minutes. Positive PCR products were enzymatically purified using shrimp alkaline phosphatase (SAP) and *E. coli* exonuclease I in the supplied buffer according to the manufacturer's protocol. After purification, PCR products were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Samples were analysed on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems), according to the manufacturer's instructions. Sequences were analysed by BLAST searches against the non-redundant nucleotide database at the National Center for Biotechnology Information (NCBI).

2.2.17 Preparation of chemical competent *E. coli* cells

A single colony of *E. coli* HB101, DH5α, XL1-Blue, GM2163 or STBL4 cells was used for inoculation of 5ml LB medium without antibiotics. Cells were grown overnight at 37°C with vigorous shaking. On the next day the overnight culture was used to inoculate 500 ml of the fresh LB medium and afterwards bacteria were cultured under the same conditions until they reached OD_{600} = 0.9 – 1.0. Subsequently the culture was centrifuged for 15 minutes in a GSA rotor at 8000 rpm at 4°C. The medium was discarded and the bacterial pellet was resuspended in 125 ml of buffer A and incubated 30-60 minutes on ice. Next, the centrifugation step was repeated and the pellet was resuspended in 20 ml of buffer B. Bacteria were aliquoted, frozen in liquid nitrogen and stored at -80°C. Media and solution used for preparing competent cells can be found in Table 7 and Table 9.

2.2.18 Isolation of prey plasmids from double positive yeast colonies

Isolation of prey plasmids from HIS3- and LacZ-positive yeast clones was performed in three steps, which included isolation of total yeast DNA, separation of the plasmid DNA from yeast genomic contamination in HB101 *E. coli* strain, and finally plasmid purification from bacterial cells.

Each HIS3/LacZ double positive yeast colony was cultured overnight at 30°C in SD medium lacking leucine. The next day, cells were harvested by centrifugation for 10 seconds at 13000 rpm. The supernatant was discarded and the cell pellet was resuspended in 200 µl of yeast lysis buffer. Subsequently, 200 µl of phenol/chloroform and approximately 0.3 g of glass beads were added to the buffer. The yeast cell suspension was vortexed for 2 minutes in order to destroy cell walls, and centrifuged at room temperature for 5 minutes at 13000 rpm. Yeast total DNA was precipitated by adding 1/10 volume of 3 M NaAc and 2.5 volumes of 96% ethanol to the supernatant, and by centrifugation for 15 minutes at 13000 rpm at 4°C. The DNA pellet was washed twice with 70% ethanol, dried and suspended in 15 µl sterile water.

Approximately 10–15 µl of yeast total DNA was used to transform chemical competent HB101 *E. coli* cells with the "heat shock" method according to standard procedures (Sambrook et al. 1989). Before plating, cells were harvested at 2500 rpm for 5 minutes. The pellet was washed twice with M9 minimal medium, resuspended in 80 µl of M9 minimal medium, plated on M9 agar and incubated at 37°C. Colony growth was observed 36-48 hours after transfection. Plasmid DNAs were isolated from HB101 colonies using the QIAprep spin miniprep kit according to the manufacturer's protocol.

All solutions, kits and media used in this section are listed in Table 4, Table 8 and Table 9.

2.2.19 Confirmation of protein-protein interactions in yeast

Each positive prey plasmid was separately transformed into the L40 yeast strain carrying the LexA_Hoxd13 construct according to the small scale yeast transformation protocol described in section 2.2.12. After transformation, cells were plated on SD –THULL agar. Interactions between GAL_Hoxd13-HD construct and the candidate genes were confirmed in yeast strain AH109 basically in the same way, however in order to detect protein-protein interactions yeast cells were plated on a slightly different medium (SD agar lacking adenine, histidine, leucine and tryptophane).

2.2.20 Cell culture and DNA transfection

COS1 cells were cultured in the appropriate medium at 37°C in 5% CO_2. For immunocytochemistry studies, cells were grown on glass coverslips in 12-well plates. Transient transfection was performed using Polyfect transfection reagent (Qiagen), according to the manufacturer's protocol. Typically, 4×10^4 cells per well were seeded 24 hours before the experiment, and transfected with 0.6 µg plasmid DNA and 4 µl Polyfect Transfection Reagent.

For Western blot and coimmunoprecipitation analysis transfection was performed in 6-well plates and the amount of plasmid DNA and transfection reagent was adjusted proportionally to the surface of wells.

2.2.21 Immunocytochemistry

Immunofluorescence experiments were performed 48 hours after transfection. All steps were performed at room temperature. First, medium was sucked off and the cells on coverslips were rinsed with PBS and fixed in 4% paraformaldehyde/PBS for 15 minutes. Later, cells were washed one time with PBS to remove excess of paraformaldehyde, and permeabilised in PBS buffer containing 0.2% Triton X-100 for 15 minutes. After washing three times in PBS, cells were blocked for 45 minutes in blocking solution and incubated for 1 hour with primary antibodies diluted 1:250 in blocking solution. After washing three times with PBS, cells were incubated for another hour with the appropriate Alexa Fluor-conjugated secondary antibodies diluted 1:500 in blocking solution. Subsequently, cells were washed 3 times with PBS, incubated in DAPI/PBS (1:2000) solution for 10 minutes in order to stain nuclei, and washed twice in PBS to remove excess of DAPI. Coverslips were fixed on slides with Fluoromount-G slide mounting medium and cells were visualised using 63 × oil-immersion lens on an Axiovert 200 M fluorescence microscope (Zeiss) equipped with filters for excitation of green, red and blue. Digital images were captured using an AxioCam MRm camera and the AxioVision 4.2 fluorescence image analysis software.

For the detection of different Hoxd13 proteins anti-HA antibodies produced in rabbit and Alexa Fluor® 546 goat anti-rabbit immunoglobulins were used. Detection of Peg10, Dlxin-1, Wtip, Limk1, Limd1 and Cnot3 proteins was performed with the use of monoclonal anti-Flag antibodies and Alexa Fluor® 488 goat anti-mouse immunoglobulins.

Antibodies and solutions used in this section are listed in Table 5 and Table 15.

2.2.22 Coimmunoprecipitation

Coimmunoprecipitation experiments were performed 48 hours after transfection. All steps were done at 4°C in order to avoid protein degradation. COS1 cells were shortly washed with ice cold 1 × PBS and cell-containing plates were immediately transferred onto ice. 400 µl of cold lysis buffer containing protease inhibitors was added to every well and cells were scraped off the bottom of the wells with plastic scrapers (Biochrom). Afterwards, COS1 cells were incubated for 1 hour with very delicate shaking. Cell lysates were collected in eppendorf tubes and centrifuged for 15 minutes at 14000 rpm in a table centrifuge. Lysates were transferred into fresh tubes and pellets were discarded. Total protein concentration in lysates was measured according to the standard Bradford method (Bradford 1976) using Anthos Reader 2020 (Anthos Labtec Instruments, Austria).

For one immunoprecipitation reaction 30 µl of anti-Flag gel or anti-HA agarose conjugate (both called beads for short) were used. Appropriate beads were washed 3 times with 800 µl of the lysis buffer, which was followed by addition of cell lysates containing 0.8 –1.0 mg of total protein. Bead suspensions were incubated for 2 hours with delicate shaking, pelleted and washed 4 times with 1 ml of lysis buffer. In the end, 45µl of protein loading buffer were added, samples were denatured for 5 minutes at 95°C and stored at –20°C.

2.2.23 SDS-PAGE and Western blot analysis

Polyacrylamide gels were prepared according to standard protocols (Sambrook et al. 1989). Protein samples were heated at 95°C in protein loading buffer for 5 min and separated by 10-15% SDS-PAGE. For Western blot analysis, proteins were transferred from gels to microporous polyvinylidene difluoride (PVDF) membrane Immobilon-P (Millipore) using a mini tank transfer unit TE22 (Amersham Bioscience) according to the manufacturer's instructions. After transfer, blots were incubated with Western blot blocking buffer for 1 hour at room temperature, and then incubated overnight at 4°C with primary antibodies diluted in blocking buffer. Blots were washed three times in PBST, and incubated for 1 hour at room temperature with peroxidase-conjugated anti- rabbit IgG or anti-mouse IgG. Following 5 times washes in PBST, blots were incubated with Roti-Lumin chemiluminescence substrate (Roth) according to the manufacturer's protocol and exposed to Fuji X-ray film. Anti-HA and anti-Flag antibodies were used at 1:2000 dilution. Anti- c-myc monoclonal antibodies were used at 1:500 dilutions. Peroxidase-conjugated anti-mouse and anti-rabbit antibodies were

used at 1:2000 dilution. Solutions and antibodies used in this section are listed in Table 5 and Table 15.

2.2.24 Synthesis of RNA probes for *in situ* hybridisation

The part of the *Peg10* coding sequence corresponding to 2075–2348 bp according to GenBank entry AB091827 (yeast two-hybrid positive prey number 37) was cloned into the pcDNA-Flag vector, as described in section 2.2.11.2. One of the positive plasmids containing the insert cloned in the antisense orientation was used to prepare the probe for RNA *in situ* hybridisation experiments. For negative control, the vector pcDNA-Flag_37, which carried the same *Peg10* insert in a sense orientation, was used. Similarly, the *Limk1*, *Limd1* and *Cnot3* yeast two-hybrid clones were subcloned into pcDNA-Flag vector (described in the section 2.2.11.3). Plasmids carrying inserts in frame were used for synthesis of negative controls, whereas vectors with inserts cloned in the antisense orientation were used to synthesise *in situ* probes. In short, 30 µg of each plasmid was linearised with 3 units of *Bcl*I restriction endonuclease in a total reaction volume of 100 µl. After 3 hours of digestion at 37°C, 20 µg proteinase K was added to the reaction in order to remove ribonucleases. Starting from this step, all used solutions were treated with DEPC or they were taken from fresh aliquots in order to ensure that they were RNase-free. After 1 hour incubation at 37°C the reaction was stopped and 400 µl DEPC-H_2O was added, followed by phenol/chloroform extraction of the plasmid DNA, according to standard procedures (Sambrook et al. 1989). Precipitated vector was resuspended in 25 µl DEPC-H_2O and used as a template for *in vitro* transcription. The reaction was performed in a total volume of 20 µl and contained 1 µg of vector, 1 × DIG labelling mix, 40 units of RNase inhibitor, 1 × transcription buffer (Roche), and 40 units of T7 RNA polymerase. All components were incubated for 2 hours at 37°C, afterwards 20 units of RNase-free DNase I were added in order to remove the plasmid DNA template and incubation was proceeded. After 15 minutes, the reaction was stopped by adding 2 µl of 0.2 M EDTA, synthesised RNA was precipitated using 1/10 volume of 4 M LiCl and 2.5 volumes of ethanol, and diluted in 100 µl DEPC-H_2O. The quality of the synthesised RNA probe was checked on 1% agarose gel.

2.2.25 Whole mount *in situ* hybridisation

Isolation of mouse embryos was performed in DEPC-PBS. Clean embryos were fixed overnight in 4% PFA/PBS at 4°C. On the next day embryos were washed at room temperature two times for 5 minutes in DEPC-PBST, two times for 5 minutes in 50% methanol/DEPC-PBST, and finally one time for 5 minutes in 100% methanol. In the end, embryos were soaked in fresh 100% methanol and stored at –20°C.

Hybridisation with DIG-labelled probes was performed essentially as described elsewhere (Albrecht et al. 2002). In short, embryos were rehydrated at 4°C in 75%, 50%, 25% methanol/DEPC-PBST (10 minutes per each solution) and washed 2 times in ice-cold DEPC-PBST. Afterwards, embryos were bleached with 6% hydrogen peroxide for 1 hour at 4°C, followed by 3 washes in DEPC-PBST (10 minutes each). Younger embryos (E10.5 and E11.5) were digested for 3 minutes with 10 µg/µl proteinase K in proteinase K buffer at room temperature. For older embryos (E12.5), proteinase K concentration was raised to 20 µg/µl and the reaction time was increased up to 5 minutes. After digestion, embryos where washed 2 times in PBST/Glycine, 2 times in DEPC-PBST, 3 times in RIPA buffer, again 3 times in DEPC-PBST and fixed for 20 minutes in PFA/Glutaraldehyde. Subsequently, embryos were washed 3 times in DEPC-PBST and incubated in hybe buffer:DEPC-PBST (1:1 dilution) for 10 minutes, followed by a single washing step in hybe buffer at room temperature. Prehybridisation of embryos was performed in hybe buffer at 65°C. Prior to hybridisation, DIG-labelled RNA probes were diluted 1:100 and denatured for 5 minutes at 80°C, and added together with tRNA (100 µg/ml) to the prehybridisation solution. Hybridisation was performed at 65°C overnight. In order to remove unbound probe, embryos were washed 2 times for 30 minutes in hybe buffer at 65°C, followed by a single washing step in 50% RNase solution/50% hybe buffer at room temperature and by digestion with RNase A (100 µg/ml in RNase solution) for 1 hour at 37°C. Subsequently, embryos were incubated for 5 minutes in 50% RNase solution/50% SSC/FA/T at room temperature, and washed with SSC/FA/T at 65°C (2 times for 5 minutes, 3 times for 10 minutes and 6 times for 30 minutes). After cooling down to room temperature embryos were washed for 10 minutes with (1:1) SSC/FA/T/1 × MABT and subsequently 2 times for 10 minutes with 1 × MABT. Prior to the incubation with antibodies, embryos were blocked with 10% Blocking reagent (Boehringer) diluted in 1 × MABT for 1 hour at the room temperature. At the same time 1 × MABT solution containing 1% Blocking reagent and alkaline phosphatase-conjugated anti-DIG antibodies (diluted 1:5000) was prepared and incubated for 1 hour at 4°C. After the

blocking step, embryos were incubated overnight at 4°C in anti-DIG-antibody-containing solution and subsequently transferred to room temperature and washed 8 times for 1 hour in fresh PBST/Tetramisole solution in order to remove unbound antibodies. Staining of the embryos was based on the enzymatic reaction performed by alkaline phosphatase (AP) in AP buffer using the BM Purple AP Substrate (Roche) according to the manufacturer's protocol. After the reaction, embryos were washed in alkaline phosphatase buffer, fixed in PFA/Glutaraldehyde and stored at 4°C. Images of the stained embryos were captured using the Leica MZ 12.5 stereomicroscope (Leica) coupled to the AxioCam HRc camera and the AxioVision 4.2 image analysis software.
Solutions and antibodies used in this section are listed in Table 6 and Table 15.

2.2.26 Section *in situ* hybridisation

Mouse embryos were isolated and fixed as described in the previous section. On the second day the embryos were washed twice for 15 minutes in fresh DEPC-PBS, followed by incubation for 1 hour in 70% ethanol. After changing the ethanol, embryos were incubated in the tissue processor Leica TP 1020 (Leica) according to the manufacturer's protocol, embedded in paraffin and sectioned at 7 µm with the Reichert Jung 2050 microtome (Reichert Jung). Sections were attached to glass slides, baked for 1 hour at 60°C, dewaxed in xylene and rehydrated using 100%, 75%, 50% and 25% ethanol concentrations. At the end, slides were washed 2 times in DEPC-PBS, followed by a fixation step in 4% PFA/PBS for 10 minutes at room temperature. After 2 washes in fresh DEPC-PBST, mouse sections were digested with proteinase K (1.5 µg/ml) for 10 minutes and washed again 2 times in DEPC-PBST, followed by a second fixation step in 4% PFA/PBS for 5 minutes. Afterwards, sections were acetylated for 10 minutes with 0.25% acetic anhydride, washed 2 times in DEPC-PBST and prehybridised for 1–4 hours in hybridisation solution (for paraffin sections) in a humified slide box. Two µl of the specific probe was denatured in 100 µl of hybridisation solution (for paraffin sections) and hybridised to the slides at 65°C overnight. On the next day slides were rinsed with 5 × SSC, washed with 1 × SSC/50% formamide for 30 minutes at 65°C and in TNE for 10 minutes at 37°C, followed by RNase digestion (20 µg/ml diluted in wash buffer). Later, slides were washed in TNE for 10 minutes at 37°C, followed by a single wash in 2 × SSC for 20 minutes at 65°C and 2 washes in 0.2 × SSC in the same conditions. For detection of DIG-labelled probes, slides were washed 2 times with 1 × MABT at room

temperature and blocked in 1 × MABT containing 20% heat inactivated sheep serum (HISS). 1:2500 diluted alkaline phosphatase-conjugated anti-digoxygenin antibodies were pre-incubated for 2 hours at 4°C in 5% HISS/1 × MABT and pipetted onto the slides. After an overnight incubation at 4°C, slides were washed 3 times in 1 × MABT, incubated for 10 minutes in NTMT and developed with BM Purple AP Substrate (Roche), similarly as described for whole mount *in situ* hybridisation. After the reaction, slides were rinsed with NTMT, washed 2 times for 5 minutes in PBS, fixed in 4% PFA/PBS and embedded with the help of Tissue-TEK (A. Hartenstein).

For cryo-sections, freshly collected embryos were placed into a chamber filled with OCT cryomount medium (Sakura) and frozen in dry ice/ethanol. The resulting frozen blocks were stored at –80°C. Prior to sectioning, blocks were equilibrated for 24–72 hours at –20°C and sectioned at 10 µm with the HM 560 Cryo-Star cryostat (MICROM International GmbH). Hybridisation with DIG-labelled probes and signal detection was performed with the use of the Genesis RSP 150 automation system supplied with the Gemini pipetting software (Tecan Group Ltd.) as described elsewhere (Carson et al. 2005).

Images were captured using the Leica DMR light microscope (Leica) coupled to the AxioCam HRc camera and the AxioVision 4.2 image analysis software.

3 RESULTS

3.1 Analysis of a balanced translocation t(2;10)(q31.1;q26.3) in a male patient with SPD

Disease-associated balanced chromosomal rearrangements (DBCR) form a unique resource in linking genotypes and phenotypes. Therefore mapping of chromosomal breakpoints in patients with DBCR is a very powerful approach to identify disease genes (Bugge et al. 2000). Moreover, the same method can be a good starting point to investigate regulatory elements controlling gene expression (Fantes et al. 1995; Lauderdale et al. 2000; Kleinjan et al. 2001; Griffin et al. 2002). In order to obtain more data about molecular mechanisms leading to limb development in humans, a detailed cytogenetic and molecular study on a balanced translocation t(2;10)(q31.1;q26.3) in a male patient with limb defects was performed.

3.1.1 Clinical description of the male patient with a de novo balanced translocation t(2;10)(q31.1;q26.3)

The patient is a thirteen-year-old boy carrying an apparently balanced chromosome rearrangement t(2;10)(q31.1;q26.3) (schematically depicted in Fig. 5*A*). He is the second child of healthy unrelated parents, born after normal pregnancy without fetal distress. Clinical examination after birth revealed severe malformations of hands and feet as well as a dysmorphic face. The hands were short, had 6 fingers, and absence of the distal phalanges including the nails was noted. The feet showed complete absence of toes II to V and a rudimentary first toe with a missing distal phalanx. X-rays of the hands showed 5 short metacarpals and 6 digits, each consisting only of a single phalanx (Fig. 5*B*-*C*). Metacarpal III was bifurcated at its end giving rise to two digits. The patient was not able to fully extend his elbows and knees due to contractions. Bilateral inguinal hernias were noted and surgically corrected. X-rays of the thorax showed hypoplasia of the medial ends of both clavicles. Ultrasound of the skull revealed hypoplasia of the cerebellum. Further clinical examinations demonstrated developmental delay, deep set eyes (left > right) a progressive scoliosis, narrow shoulders, ataxia, coxa valga, and short stature (110 cm at age 6, < p3). The patient received

several surgical corrections of the hands including removal of polydactylous digits. He has developed well and is currently attending a special school.

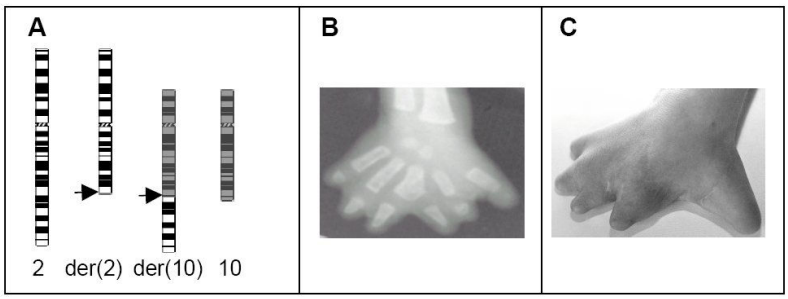


Fig. 5 *A*: Schematic representation of chromosomes in the patient with the translocation t(2;10)(q31.1;q26.3). Breakpoints at 2q31.1 and 10q26.3 are marked with arrows. *B,C*: An X-ray and a photo of the right hand of the patient after surgical correction.

3.1.2 Cytogenetic investigation of the chromosomal breakpoints in the patient

Breakpoint mapping was performed using the fluorescence *in situ* hybridisation (FISH) technique. Several large genomic clones were selected from the regions of interest and were hybridised to the patient's metaphase chromosomes. Number and localisation of detected signals helped to narrow down the breakpoint regions (see Fig. 6*A*).

For the chromosome 2 breakpoint, YAC clone 751E12 (with marker D2S2257 located at 188 cM) gave signals on chromosome 2 and der(2), which indicates that this clone lies proximal to the breakpoint. Another YAC clone, 785G8 (with markers D2S148 and D2S2173 located at 190 cM) showed signals on chromosome 2 and der(10), thus being distal to the breakpoint (data not shown). Thereby, the breakpoint region on chromosome 2 was narrowed down to a 2 cM long interval. This region contains several genes, among others the *HOXD* genes, which were considered as interesting candidates for the patient's phenotype. In order to investigate, whether the translocation disrupts the *HOXD* cluster, BAC clone RP11-514D19 (GenBank acc. no. AC016915), containing the *EVX2* and *HOXD8-13* genes, were hybridised to the patient's chromosomes. The results indicated that this clone lies distal to the breakpoint (Fig. 6*B*). Additional hybridisation experiments with a series of BAC clones selected from the region proximal to *EVX2* led to the identification of the breakpoint-spanning clone RP11-538A12 (GenBank acc. no. AC016761) (Fig. 6*C*).

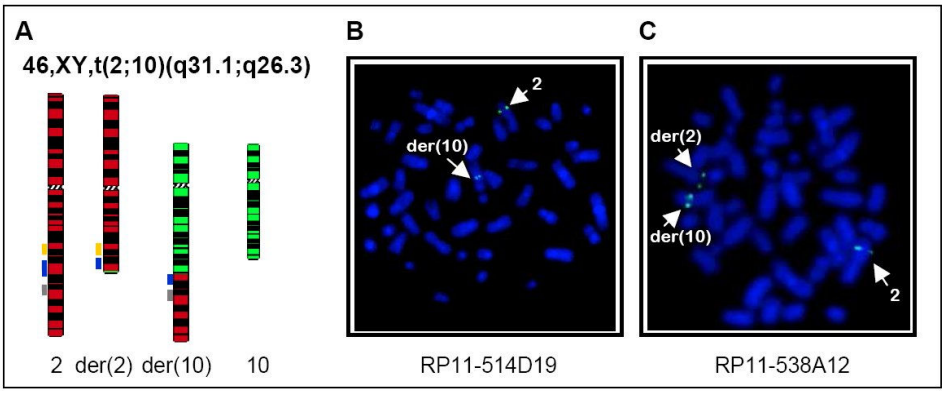


Fig. 6 *A*: Principle of FISH. Genomic clones from the breakpoint region on chromosome 2 (yellow, blue and red bars) are used as hybridisation probes on patient metaphase chromosomes. The breakpoint-spanning clone (blue) shows signals on the normal chromosome 2, the derivative 2 [der(2)] and the derivative 10 [der(10)]. For a proximal (yellow) or a distal (red) clone, only two signals are detected in each case: on chromosome 2 and der(2) or chromosome 2 and der(10), respectively. *B*: FISH results showing hybridisation of BAC clone RP11-514D19 to patient's chromosomes. Signals on chromosomes 2 and der(10) indicate that the clone lies distal to the breakpoint. *C*: FISH results showing hybridisation of the breakpoint-spanning clone RP11-538A12 to patient's chromosomes; signals were detected on chromosome 2, der(2) and der(10) (indicated by arrows).

To map the breakpoint more precisely, PCR products generated with primers selected from the clone RP11-538A12 were used for screening the chromosome 2-specific cosmid library. Positive clones were hybridised to the patient chromosomes, and as a result the proximal and distal cosmids, LLNLc128A0237 and LLNLc128F0946 respectively, were identified. Sequencing of the cosmid ends, together with database searches, revealed that these clones slightly overlap by approximately 0.9 kb. The overlapping sequence lies in an approximately 4.5 kb long LINE repeat, thus the region covering this repeat and extending at least 3 kb in both directions was considered as the region of interest, where the breakpoint must have occurred.

Breakpoint mapping on chromosome 10 was performed in a similar way. YAC 743G11, one of six clones that have been tested, gave signals on der(10), which placed the breakpoint distal to the marker D10S1201. Further mapping showed that BAC RP11-300B2 (GenBank acc. no. AL355531) was breakpoint spanning (not shown). Moreover, BACs RP11-267K7 (GenBank acc. no. AL359508) and RP11-290N15 (GenBank acc. no. AC022688) were found to be proximal and clones 218C11 (GenBank acc. no. AL353725) and 355P15 (GenBank acc. no. AC011849) to be distal to the breakpoint (Fig. 7 and data not shown). BLAST searches

combined with information from the UCSC Genome Browser Gateway database allowed establishing a contig, which indicated the overlapping regions between the proximal and the breakpoint-spanning BACs, as well as between the breakpoint-spanning clone and the distal ones. The common regions have been excluded from further investigation, and an approximately 60 kb region found solely on the breakpoint-spanning BAC has been considered as the candidate region (see Fig. 7).

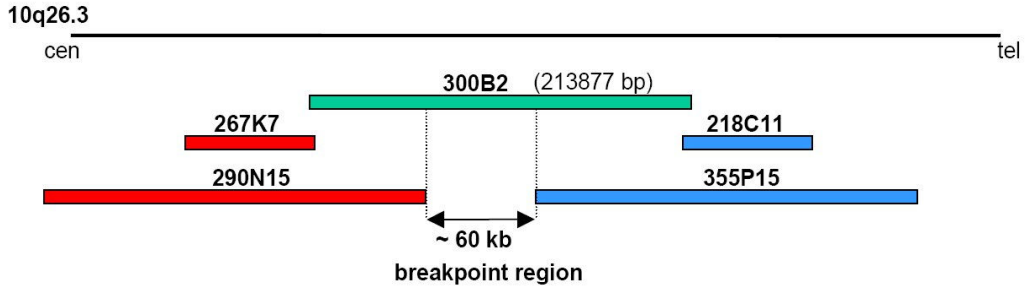


Fig. 7 BAC contig on chromosome 10q26.3. The breakpoint-spanning BAC 300B2 is shown in green, proximal BACs are shown in red and distal BACs in blue. The breakpoint region is located between the distal and the proximal BACs.

3.1.3 Southern blot experiments and cloning of the breakpoints

In order to localise precisely the breakpoint on chromosome 2, patient and control DNAs were digested with several restriction enzymes and transferred onto nylon membranes. Probes derived from the candidate region were radioactively labelled and hybridised to the blots. The probe 538A12_49750 gave aberrant restriction fragments in patient DNA digested with *Bam*HI or *Hin*dIII, whereas the more centromeric probe 538A12_79400 showed aberrant restriction fragments in patient DNA cut with *Bgl*II or *Pst*I. Analysis of the restriction site positions narrowed down the breakpoint region to approximately 1 kb (Fig. 8*A*).

Mapping of the chromosome 10 breakpoint was performed in a similar way, although the candidate region was much larger. Initially, eight probes from the 60 kb long breakpoint region were amplified and hybridised to digested patient and control DNAs. Two probes, 300B2_80000 and 300B2_90000 gave additional bands seen in patient DNA cleaved with *Hpa*I and not in the control. Moreover, different lengths of the aberrant restriction fragments observed with these two probes indicated that the breakpoint maps between these probes, thus is located between 80778 bp and 90997 bp on the breakpoint-spanning BAC (data not

shown). In order to refine the breakpoint region, several new probes were designed and used for hybridisation. As a result, aberrant bands were obtained with probe 300B2_84000 in the *Mbo*I or *Ssp*I digest. This allowed narrowing down the breakpoint region to approximately 700 bp (Fig. 8*B*).

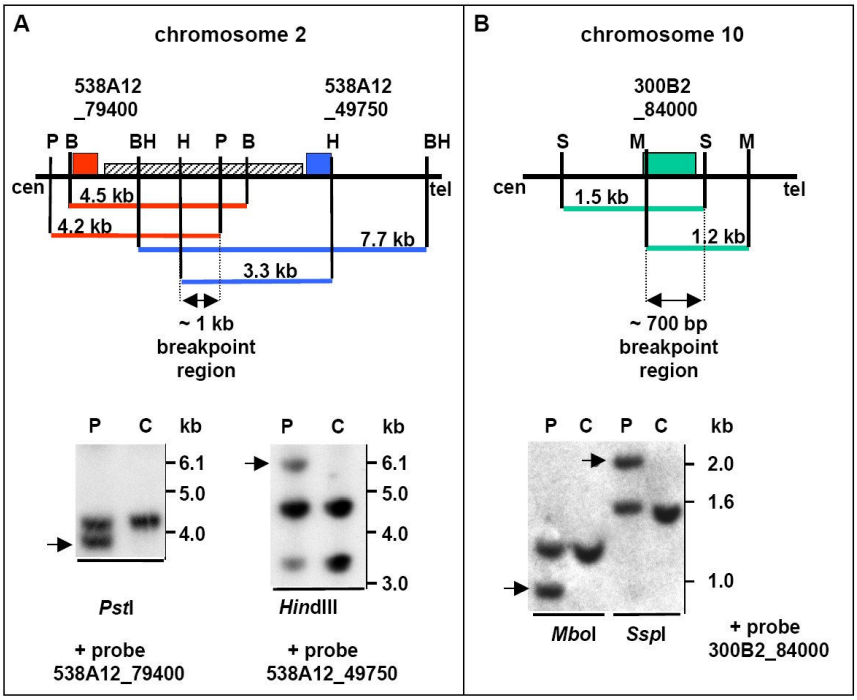


Fig. 8 *A*: Schematic representation of the breakpoint region on BAC RP11-538A12 and the localisation of probes 538A12_79400 and 538A12_49750 (shown as red and blue boxes, respectively) used for Southern blot hybridisation. Positions and sizes of normal restriction fragments are indicated and enzymes are marked as follows: B = *Bgl*II, BH = *Bam*HI, H = *Hin*dIII, P = *Pst*I. The striped box between the two probes indicates the LINE repeat. Lower panels: Southern blot analysis with probes 538A12_49750 and 538A12_79400. DNAs from the patient (P) and a control (C) were digested with *Pst*I or *Hin*dIII restriction enzymes. Aberrant bands present only in the patient and not in the control (arrows), indicate that the breakpoint on chromosome 2 is located within the *Pst*I and *Hin*dIII restriction fragments (approximately 1 kb long). Probe 538A12_49750 recognises in addition to the normal 3.3 kb long *Hin*dIII restriction fragment also another fragment of approximately 4.6 kb in both patient and control DNAs.

B: Schematic representation of the breakpoint region on BAC RP11-300B2. Probe 300B2_84000 is shown as the green box. Positions and sizes of normal restriction fragments are indicated and enzymes are marked as follows: S = *Ssp*I, M = *Mbo*I. Lower panel: Southern blot hybridisation of the probe 300B2_84000 to patient (P) and control (C) DNAs digested with *Ssp*I or *Mbo*I restriction nucleases. Additional bands (shown with arrows) present only in patient and not in control DNAs indicate that the breakpoint on chromosome 10 is located within the *Ssp*I and *Mbo*I restriction fragments (approximately 700 bp long as shown in the upper panel).

In order to clone the breakpoints on both chromosomes, adaptors were ligated to the patient DNA digested with *Eco*RI enzyme. Subsequent PCR reactions with adaptor and sequence specific primers gave rise to two products, approximately 500 bp long for the der(10) and approximately 1.1 kb long for the der(2), which were subcloned and sequenced. The results indicated that the breakpoint on chromosome 2 is located between positions 107910 bp and 107912 bp of clone RP11-538A12. The nucleotide G at position 107911 bp is missing, moreover on der(2) a 9 bp long insertion is present between the original chromosome 2 and chromosome 10 sequences. The breakpoint on chromosome 10 occured between positions 85078 bp and 85079 bp of BAC RP11-300B2 (Fig. 9).

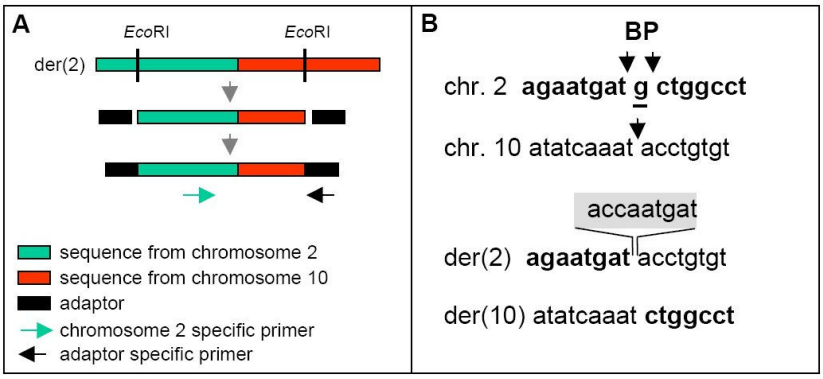


Fig. 9 *A*: Principle of breakpoint cloning. For simplicity only one derivative chromosome [der(2)] is shown. Genomic DNA from the patient was digested with *Eco*RI restriction enzyme. The restriction fragments were then ligated to adaptors on both ends. The junction fragments were amplified by PCR using primers specific for the sequence on chromosome 2 and for the sequence of the adaptor. Sequencing of the PCR products showed the exact sites of the chromosome 2 and chromosome 10 breakpoints.

B: Chromosome 2, 10, der(2) and der(10) sequences in the translocation patient. Chromosome 2-derived sequences are shown in bold, the underlined guanine residue from chromosome 2 is missing in both derivative chromosomes. The 9-bp insertion of unknown origin within der(2) (shaded box) separates chromosome 2- and chromosome 10-derived sequences.

3.1.4 The *MGMT* gene is disrupted by the breakpoint but it is still expressed from the intact chromosome 10 in the translocation patient

Computational analysis of the breakpoint region on chromosome 10 indicated that the breakpoint-spanning BAC 300B2 contains the methylguanine-DNA methyltransferase gene (*MGMT*) encoding a DNA-repair enzyme. The breakpoint occurred between exons 1 and 2 of

MGMT, upstream of the ATG codon (Fig. 10). Therefore one allele of this gene must be inactive in the patient.

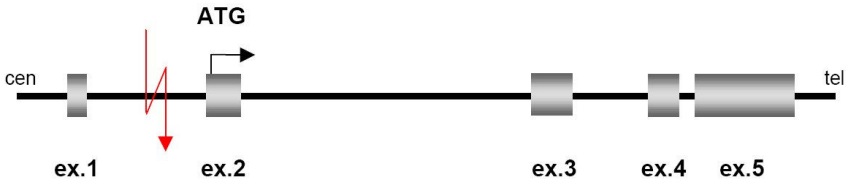


Fig. 10 Organisation of the *MGMT* gene. Five exons span over 300 kb of genomic DNA (not drawn to scale). The breakpoint in the translocation patient (marked in red) disrupted *MGMT* between exons 1 and 2.

In order to examine the expression of *MGMT* in the patient, RT-PCR and Northern blot experiments have been performed. Both approaches indicated that the expression of *MGMT* is maintained in the patient lymphoblastoid cell line. Moreover, Northern blot results suggested that the amount of the *MGMT* mRNA compared to the total mRNA (represented by the housekeeping gene *G3PDH*) is lower in the patient than in control cell lines (Fig. 11*A*). Quantification analysis performed using the ImageQuant software confirmed that the intensity of the *MGMT* signals in the patient is approximately two times lower than in controls. This result is in agreement with the previous data which indicated that in the patient one of the *MGMT* alleles has been disrupted by the breakpoint.

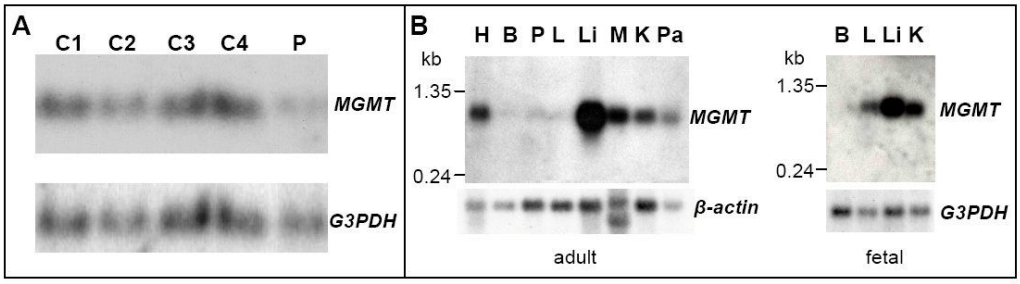


Fig. 11 *A*: Northern blot hybridisation of the *MGMT* probe (exons 2-4) to four different control (C1, C2, C3 and C4) and patient (P) total RNAs. *G3PDH* served as a control for RNA loading. The *MGMT* signal is lower in the patient compared to the controls. *B*: Hybridisation of the same *MGMT* probe to human multiple tissue Northern (MTN) blots shows ubiquitous expression in adult and foetal tissues. H = heart, B = brain, P = placenta, L = lung, Li = liver, M = skeletal muscle, K = kidney, Pa = pancreas. The length of the detected transcript corresponds to the literature data (Nakatsu et al. 1993). *β-actin* or *G3PDH* served as a control for RNA loading.

Moreover, the *MGMT* probe was hybridised to commercial human multiple tissue Northern blots from adult and foetal tissues. The results showed ubiquitous expression of this gene in different tissues, with the highest level of expression in liver and the lowest in brain (Fig. 11*B*).

3.1.5 Analysis of the breakpoint region on chromosome 2

3.1.5.1 The breakpoint on chromosome 2 does not disrupt any known gene

Sequencing of the breakpoint region revealed the exact position of the chromosome 2 break, which appeared to lie approximately 390 kb away from the *HOXD* genes (see Fig. 12). To clearify whether there is any gene disrupted on chromosome 2 or not, the region around the breakpoint has been analysed.

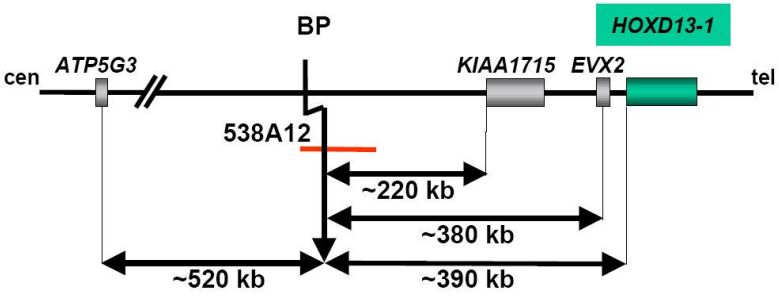


Fig. 12 Schematic map of the breakpoint region on chromosome 2q31.1. The breakpoint spanning BAC clone RP11-538A12 is shown in red. Known genes neighbouring the breakpoint (BP) are schematically depicted (boxes). The *HOXD* cluster containing 9 genes is represented by the green box. Approximate distances between genes are indicated (not drawn to scale). Cen and tel indicates the orientation of the centromere and the telomere, respectively.

First, the NIX program was used to find known genes and ESTs. The complete sequence of the breakpoint-spanning BAC 538A12 was analysed but there were no obvious genes found on this clone. However, there were good matches to 14 EST sequences derived from colon, breast, uterus or testis. All ESTs were collinear with the genomic sequence and corresponded to nine different positions in the genome. Since it was not clear whether the ESTs were parts of genes or just genomic contaminations, RT-PCR experiments were performed. All primers were first tested on human genomic DNA and only working primer pairs were used for

further experiments. Lack of appropriate human tissues was the reason for using human cell lines as the RNA source for RT-PCR. None of the ESTs could be amplified using 35 PCR cycles, but few of them gave products after 70 cycles. However, such a high number of cycles required for amplification of the putative ESTs suggests that they are probably genomic contaminations rather than expressed sequences. Therefore, it is very likely that there are no genes located on the breakpoint-spanning clone.

The list of putative ESTs and the RT-PCR results can be found in appendix (section 11.5).

3.1.5.2 The HOXD13 gene is not mutated in the patient

HOXD genes located on chromosome 2 were shown to cause limb abnormalities in humans and in mice. Especially, different mutations in the *HOXD13* gene cause synpolydactyly (SPD) (Goodman et al. 1997; Goodman et al. 1998; Calabrese et al. 2000; Debeer et al. 2002; Kan et al. 2003). Since SPD can be observed in the patient described in this study, it was theoretically possible that his phenotype is caused by a mutation in *HOXD13*, rather than by the translocation. In order to test this hypothesis, the whole coding sequence and parts of the intron of *HOXD13* in the patient were amplified (see Fig. 13) and sequenced. The results did not show any mutation within the entire *HOXD13* gene, hence the patient's phenotype is most probably caused by the translocation.

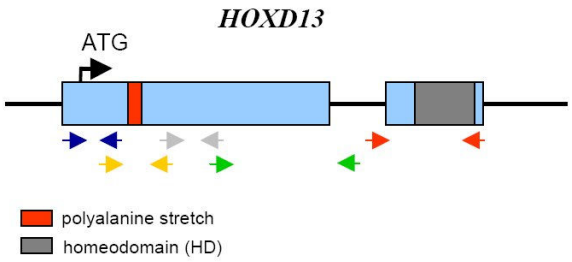


Fig. 13 Schematic representation of the human *HOXD13* gene. Two exons (blue) encode the nuclear homeoprotein. Regions coding for conserved domains are marked with red and grey. For *HOXD13* amplification, five different primer pairs (coloured arrows) were used. Not drawn to scale.

3.1.5.3 The human sequence upstream to the breakpoint shows high homology to the corresponding region in mouse

Abnormal limb development in the patient with a translocation t(2;10)(q31.1;q26.3) suggests involvement of *HOXD* genes in the origin of the phenotype. However, *HOXD13* in the proband is neither mutated nor directly truncated by the translocation. These data gave rise to

the hypothesis that the translocation might have changed expression of *HOXD* genes by disturbing their normal regulation. This in turn might have resulted in the limb phenotype of the patient. I focused my interest on the region centromeric to the breakpoint, since this part of chromosome 2 was separated from the *HOXD* cluster in the translocation patient (see also Fig. 12), and might contain important regulatory elements necessary for proper *HOXD* expression.

In theory, highly conserved non-coding sequences are good candidates for regulatory elements, thus sequence comparison between different species is generally considered to be useful in the identification of gene enhancers or silencers. Since this approach has already been successfully implemented (Loots et al. 2000), I decided to use it in my study. Approximately 130 kb of human sequence lying centromeric to the breakpoint were compared to the corresponding mouse region (approximately 70 kb long) using the PipMaker programme. As a result, 45 elements equal or longer than 100 bp and showing at least 70% identity between human and mouse were found. Five of these elements are associated with human EST sequences BG952464, AW937867 and BE064736, and their expression (except for EST BE064736) was confirmed by RT-PCR in human cell lines or in brain (data not shown). One sequence corresponds to the mouse EST AK015352 expressed in testis. It was not clear whether the other 39 elements are coding or not. In order to test this, RT-PCR experiments with primers specific for these elements were performed on RNA from mouse stage E16.5. Because of the high similarity between human and mouse genes, it was justified to use mouse instead of human RNA for RT-PCR analysis. Moreover, mouse tissues were much more easily accessible. All tested sequences but one could not be amplified using less than 70 PCR cycles, suggesting that they might be non-coding. In conclusion, with this approach it was not possible to reduce significantly the number of putative regulatory elements and to create a good basis for further tests. For this reason functional analyses of the candidate regions have not been pursued.

The list of putative candidates for regulatory sequences with their characteristics is presented in appendix (section 11.6).

3.2 Screening for Hoxd13 interaction partners

As the sequencing of the human genome is completed, the focus of research is shifting to the functional analysis of gene products. Most proteins do not act alone, instead they interact with each other or with other components of a cell. Thus, a very complex structural and functional network is formed. Therefore, one of the very important tasks in molecular biology is not only to study the function of an isolated protein, but also to identify its interacting partners. This knowledge provides new clues about the normal function of the protein in the cell. Moreover, this knowledge is indispensable in understanding the molecular events leading to proper or abnormal development of our body. Since my interest was focused on molecular bases of limb development, in the second part of my project I have searched for Hoxd13 interaction partners.

3.2.1 Yeast two-hybrid screen as a commonly used method to identify protein-protein interactions

One of the most commonly used methods allowing identification of protein-protein interactions is the yeast two-hybrid system. This technique takes advantage of the fact that transcription factors require two functional domains in order to activate gene expression. One of these domains is responsible for DNA binding (BD - **b**inding **d**omain), whereas the second one, called activation domain (AD), induces transcription. In the yeast two-hybrid system these two domains are physically separated and therefore they cannot activate genes, however bringing both of them into close proximity allows reconstitution of a fully active transcription factor.

In this study, the mouse Hoxd13 protein (bait) has been fused to the DNA-binding domain derived from the prokaryotic LexA protein. Proteins from a mouse embryonic library (preys) were fused to the activation domain originating from the viral protein VP16. The bait and prey fusion proteins were expressed in L40 yeast strain, which contains LacZ and HIS3 marker genes under the control of the LexA operon. Induction of the reporter genes could occur only when a prey protein from the library interacts with Hoxd13, resulting in spatial association of the LexA DNA binding domain and the VP16 activation domain (Fig. 14). Expression of HIS3 is normally assayed by growth on media lacking histidine and *LacZ*

expression can be monitored by colorimetry or other methods detecting β-galactosidase activity.

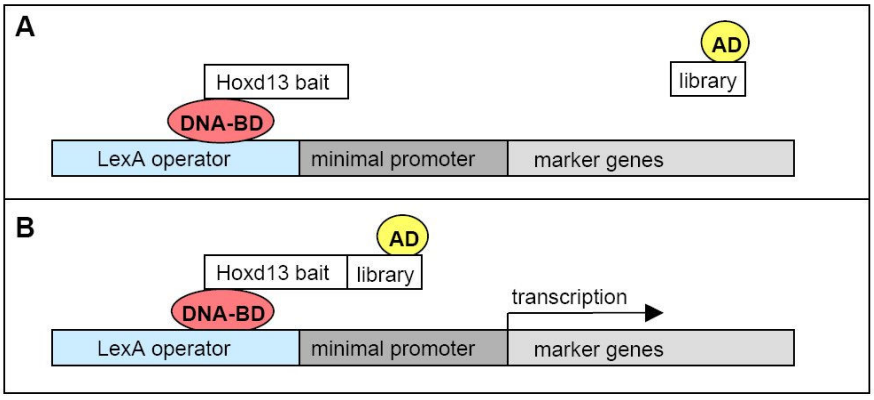


Fig. 14 Yeast two-hybrid principle. *A*: Hoxd13 fused to the LexA DNA-binding domain (DNA-BD) can bind the LexA operator but is not able to activate transcription of the marker genes (*HIS3* and *LacZ*) by itself. Library proteins fused to the activation domain (AD) cannot bind DNA and therefore marker genes are not expressed. *B*: Interaction between Hoxd13 and a specific library protein brings both the DNA binding and activation domains together which allows transcription of the *HIS3* and *LacZ* marker genes (adapted from Matchmaker LexA Two-Hybrid System, User Manual, BD Biosciences).

3.2.2 Construction of the bait vector

The yeast two-hybrid method is based on creation of a novel transcriptional activator. Therefore, in order to obtain reasonable results using this technique, it is essential to use a bait lacking transcriptional activity. This prerequisite seems to exclude transcription factors as possible baits, especially if it is not clear which domain of the protein possesses transcriptional activity. However, the structure of Hoxd13 is known quite well. Similarly to its human homologue (for the gene structure see: Fig. 13), mouse Hoxd13 contains a polyalanine (poly-Ala) repeat in the N-terminal region and a homeodomain (HD) located at the C-terminus. It has been proposed that the N-terminus is responsible for protein-protein interactions (Shen et al. 1996), whereas the homeodomain is known to bind DNA and activate transcription. Therefore, to generate the construct without transcriptional activity, it was necessary to use Hoxd13 lacking the homeodomain.

Construction of the bait vector for the yeast two-hybrid screen included RT-PCR amplification and cloning of the N-terminal part of mouse *Hoxd13* (*Hoxd13-HD*) into the pBMT116 vector containing the bacterial LexA DNA-binding domain and carrying the *Trp1*

marker gene (the pBMT116 map can be found in appendix, section 11.2). The resulting construct, called LexA_Hoxd13-HD (Fig. 15*A*), should generate a fusion protein consisting of the LexA DNA-binding domain at the N terminus followed by the mouse Hoxd13 sequence without homeodomain.

Before the yeast two-hybrid analysis was carried out, expression of this protein had been confirmed in yeast. Shortly, cell lysates were prepared from yeast transformed with the LexA_Hoxd13-HD construct and from the non-transformed control, and loaded on a polyacrylamide gel. Western blot analysis using anti-LexA specific antibodies allowed detection of a band of approximately 50 kDa, corresponding to the expected size of the LexA_Hoxd13-HD fusion protein, only in the transformed cells (Fig. 15*B*).

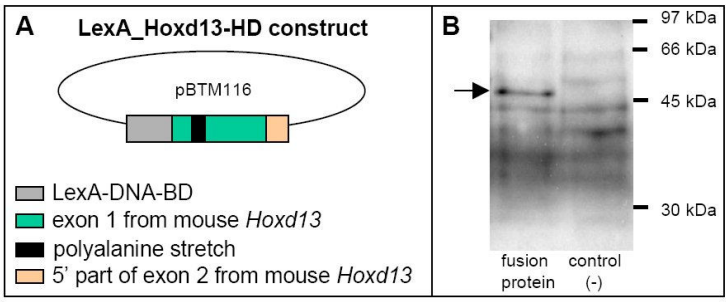


Fig. 15 *A*: Schematic representation of the bait construct used for the yeast two-hybrid screen. Exon 1 and the 5' end of exon 2 from the mouse *Hoxd13* gene have been cloned into the pBTM116 vector, downstream to the LexA DNA-binding-domain-encoding gene (LexA-DNA-BD). The part of *Hoxd13* coding for the homeodomain (3' end of exon 2) has not been included in the bait vector. *B*: Western blot showing expression of the LexA_Hoxd13-HD fusion protein in yeast. Anti-LexA antibodies detect a band of approximately 50 kDa only in cells transformed with the LexA_Hoxd13-HD construct (arrow), whereas no band of this size can be detected in the non-transformed control.

Furthermore, potential auto-activation activity of the LexA_Hoxd13-HD construct has been examined. As the first step, yeast cells carrying the LexA_Hoxd13-HD vector were streaked on SD plates lacking tryptophane and histidine (SD –TH). Single colonies which appeared after 4-5 days on SD –TH plates suggested that the bait alone could be able to activate the *HIS3* marker gene (not shown). However, it is also known that the *HIS3* promoter in the L40 yeast strain possesses basal activity even in the absence of bait-prey interactions. To overcome this problem, 3-amino-1,2,4-triazole (3-AT) which is a competitive inhibitor of the

HIS3 gene product, is usually added to the yeast medium. Testing of several different concentrations of 3-AT indicated that upon addition of 10 mM 3-AT no yeast colonies could grow on SD -TH plates (data not shown). Thus, under these conditions the LexA_Hoxd13-HD fusion protein was not capable to activate expression of the marker gene, hence it was suitable for the yeast two-hybrid screen.

3.2.3 Yeast two-hybrid screen

To co-express bait and prey proteins, yeast cells carrying the LexA_Hoxd13-HD bait construct were transformed with the mouse embryonic cDNA library cloned in the pVP16 vector. For a negative control, the bait was co-transformed with the empty pVP16 vector. After transformation, yeast cells were plated in parallel on SD plates lacking tryptophane and leucine (SD –TL) to check for transformation efficiency, and on SD plates lacking tryptophane, histidine, uracil, leucine and lysine (SD –THULL) supplemented with 10 mM 3-AT to detect interactions.

Colonies on the SD –TL plates were observed already 2 days after transformation and their number indicated that approximately 1.6×10^8 clones were screened. This means that the library, which has approximately 5×10^6 clones, was covered more than 30 times in the screen. The first positive colonies on the SD –THULL plates were seen 3 days after transformation, and they have been collected every 24 hours starting at day 4 and finishing at day 7 after transformation. At the same time, no colony growth was observed on the control plate, indicating that there was no unspecific interaction between the bait and the prey vector (Fig. 16).

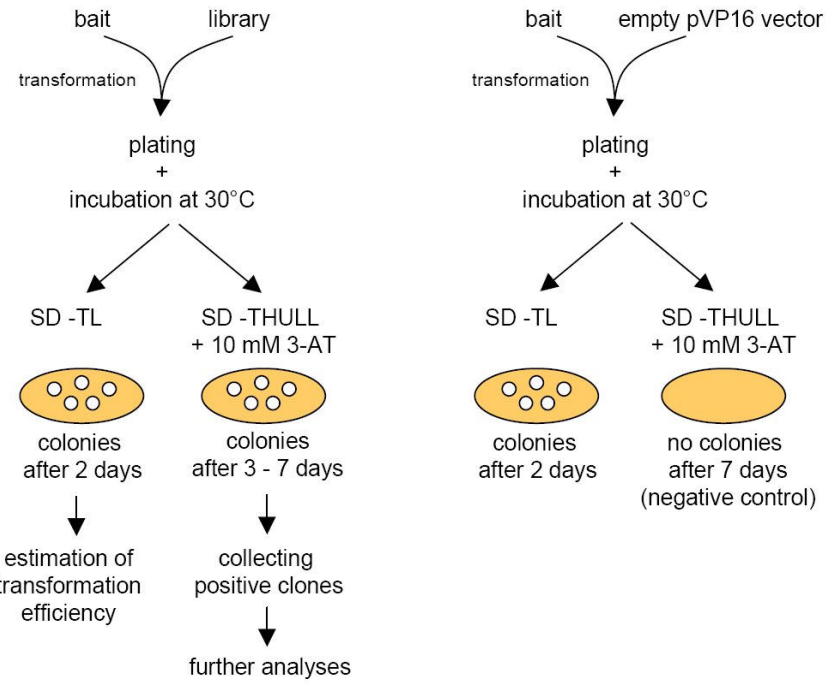


Fig. 16 Scheme of the yeast two-hybrid screen. Bait and prey constructs were co-expressed in yeast cells. After incubation at 30°C on selective plates (SD -THULL), positive yeast colonies were collected and the isolated prey clones were further analysed. For the negative control, the bait was co-transformed with the empty prey vector into yeast cells. Lack of colonies on SD -THULL medium indicated that the bait itself cannot activate marker genes. Colony growth on SD -TL medium served as a positive control of transformation.

However, due to technical limitations, a yeast two–hybrid screen usually generates a large number of false-positive interactions. In order to facilitate the identification of true positives, different approaches were implemented. First, collected clones were re-streaked 3 times on fresh SD –THULL plates. Afterwards, the expression of the second marker gene (*LacZ*) was examined in all HIS3-positive clones. In a colony-lift filter assay only 138 clones showed a strong blue staining, and these double positives ($HIS3^+/LacZ^+$) were used for further analysis. The rest of the clones showed either much weaker or no detectable expression of β-galactosidase (see Fig. 17). A summary showing the number of positive colonies picked after the screen is shown in Table 25.

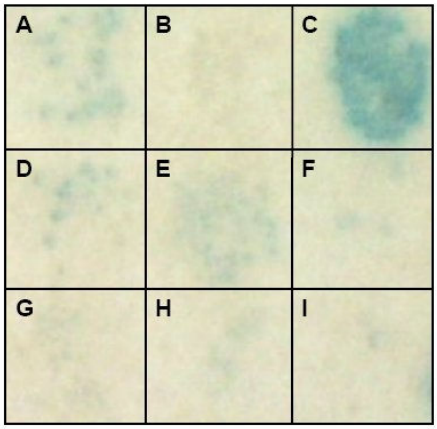


Fig. 17 Colony-lift filter assay. HIS3-positive clones were transferred to Whatman paper filters. The enzymatic reaction was performed using X-Gal as substrate. Blue staining indicates that the *LacZ* marker gene coding for β-galactosidase is expressed in yeast. Results are shown only for eight HIS3-positive clones (panels A and C-I; each panel corresponds to a single yeast clone). Panel B represents the negative control (yeast clone transformed with the LexA_Hoxd13-HD construct and the empty pVP16 vector). Only the clone seen in panel C showed intensive blue staining and therefore it was considered as a potential Hoxd13 interaction partner.

Table 25 Results of the screen showing the number of positive colonies

Day after transformation	No. of $HIS3^+$ colonies collected	No. of $HIS3^+$ colonies with intensive blue staining
4	128	22
5	252	38
6	182	26
7	401	52
Total	963	138

3.2.4 Analysis of positive colonies

Double positive clones ($HIS3^+$/$LacZ^+$) collected after the screen were re-streaked on fresh SD –THULL plates, and used for colony PCR in order to amplify the prey inserts. PCR products were sequenced with pVP16-specific primers and obtained sequences were subjected to BLAST against public databases at NCBI. The results of this analysis revealed that the positive preys represent 53 different genes. Among them, 2 hits were within 3' UTRs, whereas 51 hits matched to ORFs. Out of these 51 ORF-specific hits, 37 genes were cloned in frame, 3 were not in frame and 11 were cloned in the antisense orientation. The list of 37 genes represented by preys cloned in frame can be found in Table 26.

Table 26 List of genes found in the yeast two-hybrid screen

Gene name	GenBank accession no.	Process/function
Dlxin-1	AB029448	Limb development
Peg10	AB091827	Chromatin / RNA binding
Cnot3	NM_146176	
Dazap2	NM_011873	
Pax3	NM_008781	
Tead2	NM_011565	
Snrp70	AK077425	
Snrp116	NM_011431	
Pold4	AF515709	Proliferation / apoptosis
Nedd9	NM_017464	
Bat3	NM_057171	
Pins	AY081187	
Limk1	NM_010717	Protein modification / signal transduction
Ppp2ca	NM_019411	
Ddr2	NM_022563	
Map4k4	NM_008696	
Pkm2	NM_011099	Cellular metabolism and movement
Bckdha	NM_007533	
Pla2g4b	BC016255	
Mical3	NM_153396	
Kifc5a	NM_053173	
Itch	AF037454	
Ubce7ip3	AF124663	
Angptl2	NM_011923	Secreted / extracellular matrix proteins
Fbln2	NM_007992	
Fn1	NM_010233	
Col18a1	NM_009929	
mKIAA1046	AK122429	Unknown function
mKIAA0863	AB093268	
mKIAA0054	AB093209	
mKIAA0222	NM_183033	
Drpla	NM_007881	
Bcap37	NM_007531	
Odz3	NM_011857	
Rnf38	NM_175201	
Wtip	NM_207212	
Limd1	NM_013860	

3.2.5 Peg10 as a putative Hoxd13 interaction partner

Among the candidate genes found in the yeast two-hybrid screen *Peg10* was represented by 8 different clones. *Peg10* (GenBank accession no. AB091827) is transcribed into an approximately 6.4 kb long transcript with two long open reading frames (ORFs). ORF1 encodes a putative zinc finger domain of the CCHC class commonly found in the Gag protein of retroviruses. ORF2 has the potential to encode a polypeptide containing a consensus motif for an aspartyl protease catalytic site, which is characteristic for retroviral proteases. Similarly as in retroviruses, these two ORFs overlap and the ribosomal frameshift can lead to production of a long polyprotein (Fig. 18) (Shigemoto et al. 2001; Manktelow et al. 2005).

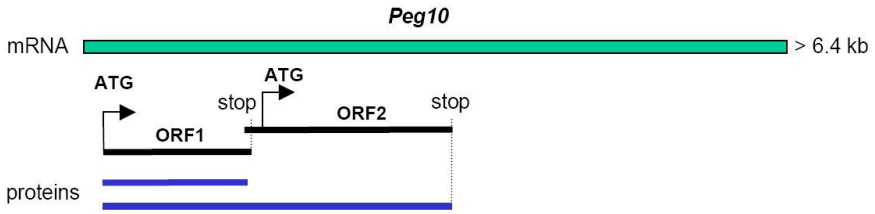


Fig. 18 *Peg10* and ribosomal slippage mechanism. *Peg10* mRNA encodes two open reading frames, ORF1 and ORF2. Translation of the transcript gives rise to two proteins. The first protein corresponds to ORF1, whereas the second one to both ORF1 and ORF2. The longer protein is generated by ribosomal slippage which leads to a frameshift and abolishment of the first stop codon. The next recognised termination codon is present at the end of ORF2.

All *Peg10* clones found in the yeast two-hybrid screen correspond to ORF2 and are cloned in frame. Their localisation within the gene can be seen in Fig. 19. Since some clones contain overlapping sequences, only five of them were used for further analysis.

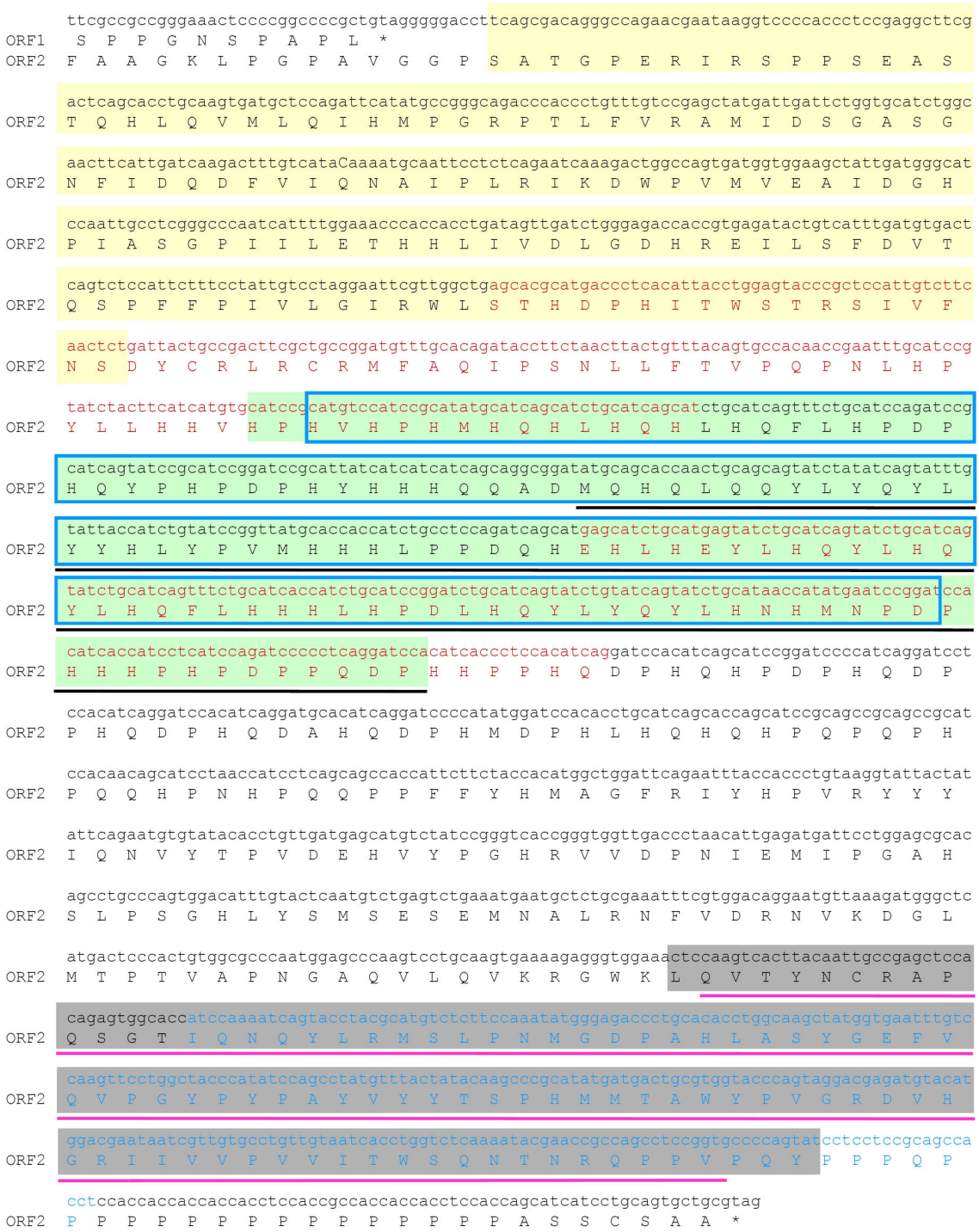


Fig. 19 Part of the *Peg10* gene sequence corresponding to the end of ORF1 and the complete ORF2 (according to the GenBank entry AB091827). Clones found in the yeast two-hybrid screen are marked as follows: clone no. 649 - light yellow box; clone no. 318 - red letters; clone no. 40 - light green box; clone no. 81 - blue frame; clone no. 37 - black underline; clone no. 22 - blue letters; clone no. 173 - grey box; clone no. 143 - pink underline.

3.2.5.1 Confirmation of the interaction with Hoxd13-HD in LexA and GAL4 systems

Prey plasmids corresponding to the *Peg10* gene were isolated from five positive clones numbered 37, 40, 143, 318 and 649 (Fig. 19 shows their inserts sequences). To confirm the interaction between preys and Hoxd13-HD in yeast, wild type L40 cells were co-transformed separately with the bait (LexA_Hoxd13-HD construct) and every single prey plasmid. Moreover, preys were co-transformed with the empty pBTM116 vector (bait vector) in order to test them for autoactivation. Growth on SD –THULL medium observed only for colonies transformed with the bait and a prey indicates interaction between Hoxd13-HD and parts of Peg10 fragments in the LexA system (Fig. 20).

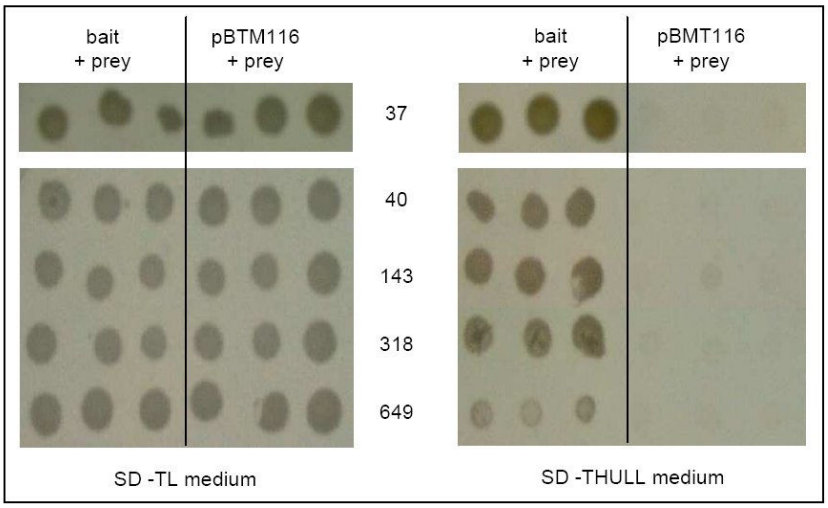


Fig. 20 Yeast cells co-transformed with the LexA_Hoxd13-HD bait or pBTM116 empty vector and one of the *Peg10* prey plasmids (numbers 37, 40, 143, 318 or 649) were spotted in triplicates on selective media. Colony growth on SD -TL plates indicates that both plasmids entered the cell (positive control of transformation), whereas presence of colonies on SD –THULL medium confirms the interaction between Hoxd13 lacking the homeodomain and different Peg10 fragments. Specificity of these interactions is proven by absence of colonies which carry a prey and the empty bait plasmid (pBTM116) on SD -THULL medium.

In addition, the cDNA for Hoxd13 lacking the homeodomain (Hoxd13-HD) was cloned into the pGBKT7 bait vector containing the GAL4 DNA-binding domain. Sequencing of the construct and expression analysis of the fusion protein performed by Western blot indicated that the GAL_Hoxd13-HD bait could be used in yeast two-hybrid experiments. Similarly as for the LexA_Hoxd13-HD construct, the GAL_Hoxd13-HD bait was co-transformed with the isolated prey plasmids. Expression of nutrition marker genes visualised by the growth of

yeast colonies on selective media indicated that the interactions between different parts of Peg10 and the Hoxd13 protein take also place in the GAL4 yeast system (data not shown).

3.2.5.2 Cloning of Peg10

Since the yeast two-hybrid method usually generates a lot of false positives, any interaction discovered in yeast has to be confirmed in the mammalian system as well. The prerequisite for this includes cloning the gene of interest into a mammalian expression vector. There are two identical GenBank entries representing the full length *Peg10* gene (AB091827 and NM_130877), moreover their alignment with the *Edr* gene (GenBank ID: AJ006464) suggests that all three sequences correspond to the same gene (sequence alignment can be found in appendix, section 11.7). Two significant differences between AB091827 and AJ006464 are present in the 5' region of both sequences (resulting in various N-termini of the predicted open reading frames) and in the length of the repeat which lies within ORF2.

As all *Peg10* clones found in the yeast two-hybrid screen corresponded to ORF2, my interest was focused on this particular part of the gene. To determine its real sequence, ORF2 was amplified in eight independent RT-PCR reactions performed on mouse cDNAs derived from four different embryonic stages. All RT-PCR products appeared to be identical, suggesting that the true sequence of the *Peg10* ORF2 was obtained. However, comparison of the RT-PCR product with the GenBank entries AB091827 and AJ006464 showed that all three sequences differ in the length of imperfect repeats present in the gene (Fig. 21). Interestingly, the various lengths of the QDPH-encoding repeats do not cause frameshifts, therefore all putative proteins have the same C-terminus.

Repetitive sequences are usually difficult to clone, since they are prone to various rearrangements during bacterial replication. However, special *E.coli* strains increasing insert stability have been designed. One of them, the recombinase-deficient STBL4 *E.coli* strain, was used for cloning *Peg10* ORF2 into the mammalian expression vector pcDNA-Flag. In spite of much effort (isolation and analysis of 35 transformants), cloning of the full length RT-PCR product corresponding to the ORF2 was not successful. All analysed clones carried mutations, mostly deletions, within the repetitive region of *Peg10*. In three analysed plasmids, the inserts appeared to be identical to the *Peg10* sequence represented by the GenBank entry AB091827. One of these clones, named Peg10-ORF2_pcDNA-Flag was used for further experiments.

```
AJ006464-ORF2  AATCCGGATCCACATCACCACTCTCATCCAGATCCCCCTCAGGATCCACATCACCCTCCA
RT-PCR_product AATCCGGATCCACATCACCATCCTCATCCAGATCCCCCTCAGGATCCACATCACCCTCCA
AB091827-ORF2  AATCCGGATCCACATCACCATCCTCATCCAGATCCCCCTCAGGATCCACATCACCCTCCA

AJ006464-ORF2  CATCAGGATCCACATCAGCATCCGGATCCCCATCAGGAT--------------------
RT-PCR_product CATCAGGATCCACATCAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCACAT
AB091827-ORF2  CATCAGGATCCACATCAGCATCCGGATCCCCATCAGGAT--------------------

AJ006464-ORF2  ------------------------------------------------------------
RT-PCR_product CAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCACATCAGCATCCGGATCCC
AB091827-ORF2  ------------------------------------------------------------

AJ006464-ORF2  ------------------------------------------------------------
RT-PCR_product CATCAGGATCCTCCACATCAGGATCCACATCAGGATGCACATCAGCATCAGGATCCCCAT
AB091827-ORF2  ------------------------------------------------------------

AJ006464-ORF2  ------GCACATCAGGATCCCCATCAGGATCCCCATCAGGATGCACATCAGGATCCACAT
RT-PCR_product CAGGATGCACATCAGGATCCCCATCAGGATCCCCATCAGGATGCACATCAGGATCCACAT
AB091827-ORF2  ------------------------------------------------------------

AJ006464-ORF2  CAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCACATCAGGATGCACATCAG
RT-PCR_product CAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCACATCAGGATGCACATCAG
AB091827-ORF2  ------------------------CCTCCACATCAGGATCCACATCAGGATGCACATCAG
```

Fig. 21 Sequence comparison between the RT-PCR product and GenBank entries AB091827 and AJ006464 corresponding to the ORF2 of *Peg10*. For simplicity, only the region containing imperfect repeats is shown. Differences in the nucleotide sequences are marked in red.

3.2.5.3 Immunocytochemistry studies

For cellular localisation studies, the Peg10-ORF2_pcDNA-Flag construct was transfected into COS1 cells and the protein was detected with anti-Flag specific antibodies. The results indicated that the protein is present in the cytosol (Fig. 22).

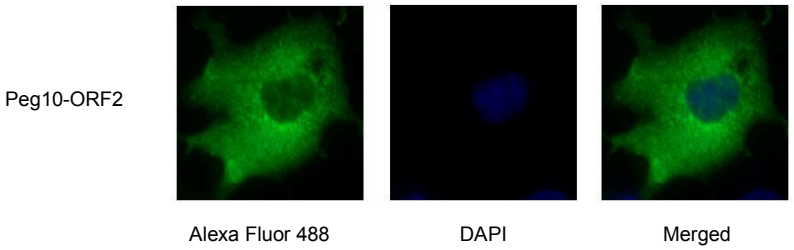


Fig. 22 Cytosolic localisation of Peg10-ORF2 (green) in transiently transfected COS1 cells. The nucleus is visualised by the DAPI-staining (blue).

A similar experiment was performed with different Hoxd13 expression constructs, namely wtHoxd13-pTL1-HA2 (carrying the full length Hoxd13), Hoxd13-HD-pTL1-HA2 (containing Hoxd13 lacking the homeodomain), Hoxd13+14Ala-pTL1-HA2 (containing Hoxd13 with the Ala stretch expanded for fourteen additional alanine residues) and Hoxd13_2Ala-pTL1-HA2 (carrying Hoxd13 with the Ala stretch reduced to two alanines). The results showed that wild type and Hoxd13_2Ala proteins localise to the nucleus, whereas the mutant lacking the homeodomain as well as the Hoxd13+14Ala are present in the cytosol, building aggregate-like structures (Fig. 23). All observed signals were specific, since the transfection of COS1 cells with empty pcDNA-Flag or pTL1-HA2 vectors resulted in no fluorescently labelled cells (data not shown).

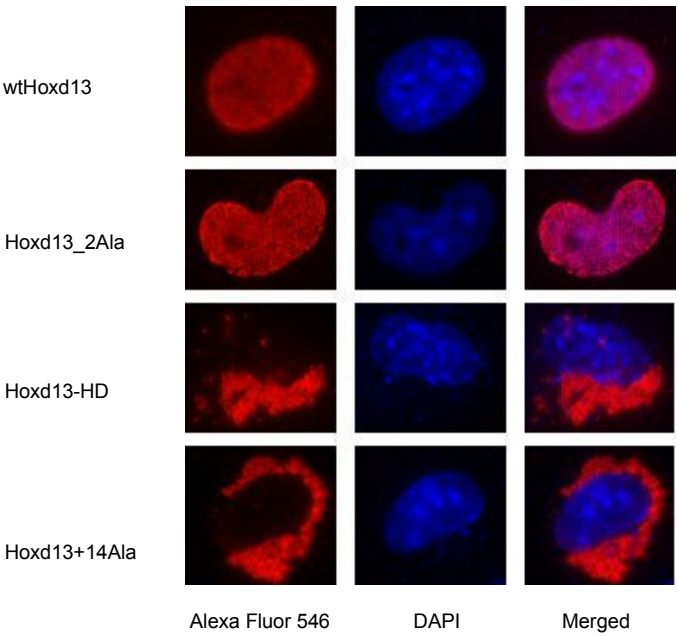


Fig. 23 Single transfection of COS1 cells with the constructs carrying wild type or mutant *Hoxd13*. Red signals indicate cellular localisation of wild type and mutant Hox proteins. Nuclei are labelled with DAPI (blue).

In the next step, different Hoxd13 constructs were co-expressed with Peg10-ORF2 in COS1 cells, in order to examine the putative interaction between the proteins. In all double-transfected cells, Peg10-ORF2 perfectly co-localised with cytosolic aggregates formed by Hoxd13-HD or Hoxd13+14Ala proteins (Fig. 24). Moreover, although Peg10 is detected only in the cytosol of single transfected COS1 cells, it can be seen in the nucleus of COS1 cells co-expressing either the full length Hoxd13 or Hoxd13_2Ala (Fig. 24). However, co-localisation in the nucleus was seen only in approximately 20% of double-transfected cells.

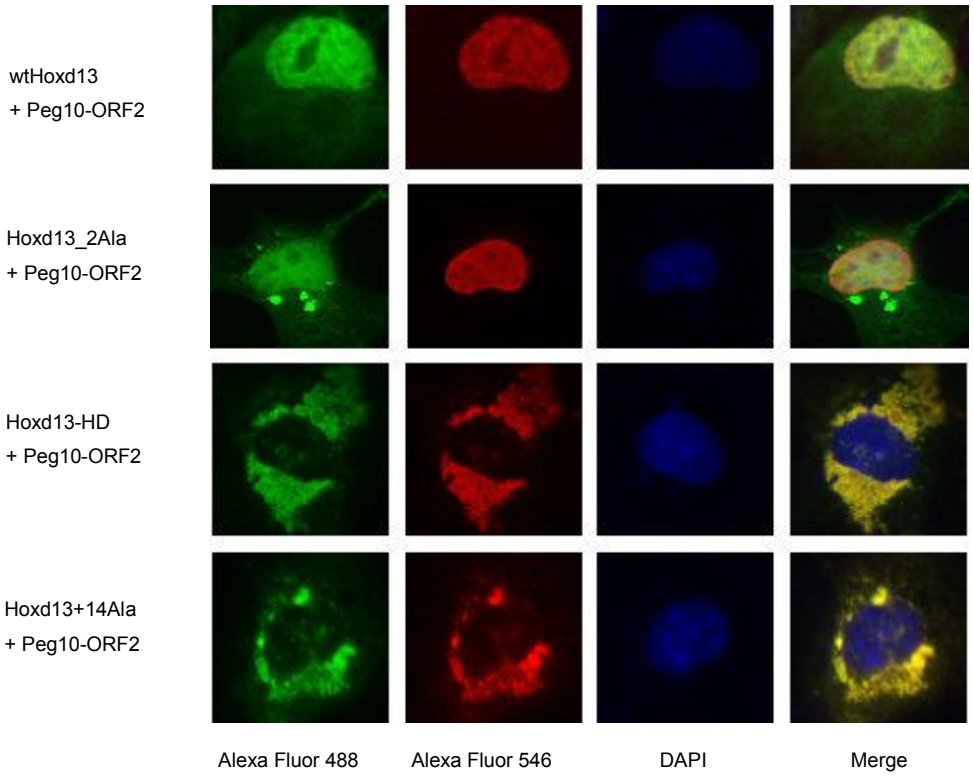


Fig. 24 Co-localisation of Peg10-ORF2 (green) with wild type and mutant Hoxd13 proteins (red) in COS1 cells. Peg10-ORF2 expression can be observed in the nucleus (stained with DAPI in blue) of cells co-transfected with wtHoxd13 or Hoxd13_2Ala. Cells co-transfected with the Peg10-ORF2_pcDNA and Hoxd13-HD-pTL1-HA2 or Hoxd13+14Ala-pTL1-HA2 constructs show co-localisation of both proteins in the cytosol.

3.2.5.4 Coimmunoprecipitation assay

The co-localisation results suggested that the Hoxd13 protein could interact with Peg10. In order to confirm this, coimmunoprecipitation studies have been performed. In preliminary experiments, COS1 cells were transfected separately with different HA-tagged Hoxd13 constructs or with Peg10-ORF2_pcDNA-Flag. Western blot analysis of the cell lysates indicated that all proteins were easily detectable with polyclonal anti-HA or anti-Flag antibodies, respectively (data not shown).

The immunoprecipitation was performed with anti-Flag antibodies, followed by detection of the Hoxd13 proteins with anti-HA antibodies. The results indicated that all Hoxd13 proteins could bind Peg10-ORF2 (Fig. 25*A*). Later, these results were confirmed by additional

experiments, in which the immunoprecipitation with anti-HA antibodies was followed by Peg10 detection with anti-Flag antibodies (Fig. 25*B*).

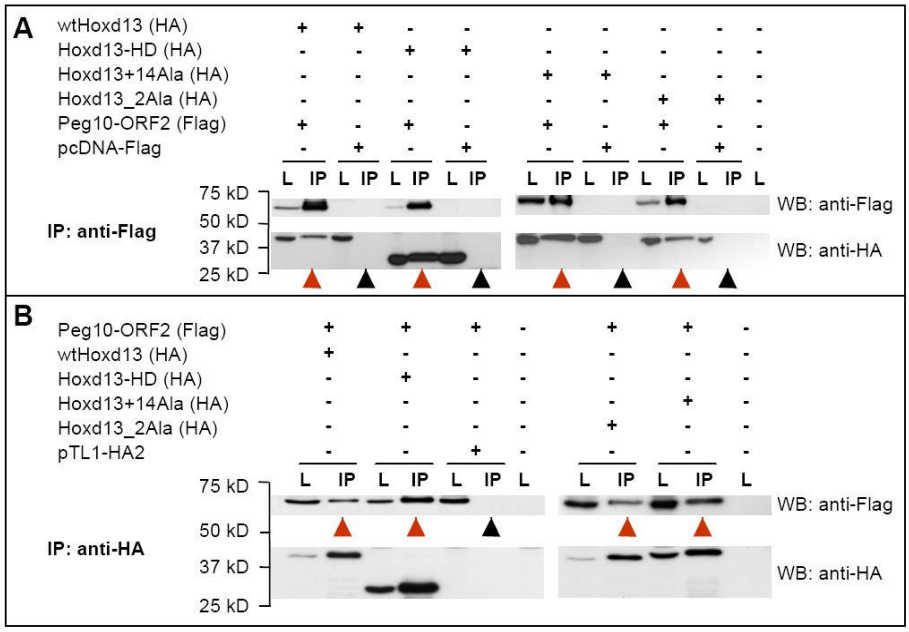


Fig. 25 Coimmunoprecipitation results. *A*: COS1 cells were co-transfected with different HA-tagged Hoxd13 constructs and with the Flag-tagged Peg10-ORF2 or the empty pcDNA-Flag vector (for negative controls). Cell lysates were used for precipitation with anti-Flag beads. In each case, lysates (L) and precipitates (IP) were run on polyacrylamide gels, followed by protein detection with anti-tag antibodies. Precipitation efficiency can be seen on anti-Flag blots (positive controls), whereas protein-protein interactions are detected using anti-HA antibodies. Hoxd13 proteins co-precipitate with Peg10-ORF2 (red arrowheads) but they cannot be pulled down by the empty pcDNA-Flag vector (black arrowheads), indicating specificity of interaction. *B*: Confirmation of Hox-Peg10 binding after antibody switch. Hoxd13 proteins were precipitated with anti-HA beads, shown on anti-HA blots. Peg10-ORF2 was pulled down together with Hox proteins (red arrowhead) but not with the empty pTL1-HA2 vector (black arrowheads), as shown on anti-Flag blots.

3.2.5.5 Whole-mount and section in situ hybridisation

The results obtained in the immunocytochemistry and coimmunoprecipitation assays strongly suggest that Hoxd13 wild type and mutant proteins could bind Peg10. However, these experiments were performed in an artificial system and it is not clear, whether these two proteins could interact *in vivo*. In order to address this question, whole mount *in situ* hybridisation was performed. An antisense probe for the *Peg10* gene was hybridised to mouse embryos from stages E10.5, E11.5 and E12.5. The results indicated that at the earlier stages

(E10.5) *Peg10* is present in most distal parts of the limb buds, however this domain seems to expand more posteriorly in E11.5 old limb buds. At a later stage (E12.5) *Peg10* is expressed in digits and the proximal limb region (Fig. 26). Section *in situ* hybridisation performed on E13.5 and E15.5 limbs showed *Peg10* signals in digits, metacarpals, carpals and in limb muscles. This pattern differs from *Hoxd13* expression domains (Fig. 27).

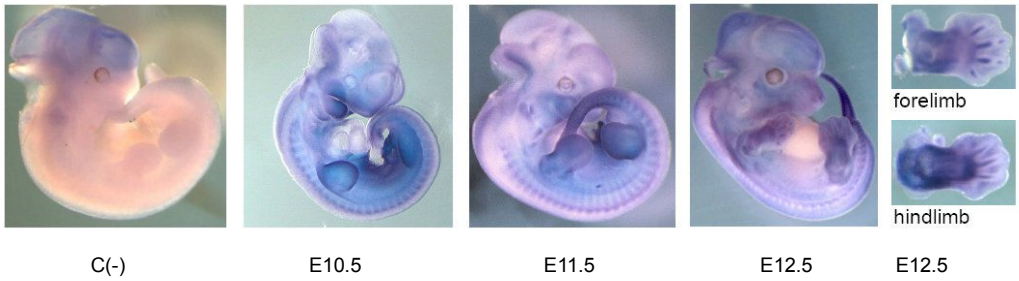


Fig. 26 Whole mount *in situ* hybridisation showing expression of *Peg10* during mouse embryonic development. *Peg10* is clearly upregulated in limb buds and in the tail tip (blue signals). Negative control [C(-)] shows no staining in both structures, indicating that the *Peg10* signal is specific. For the E12.5 stage, in addition to the whole embryo, magnified images of a fore- and a hindlimb are shown.

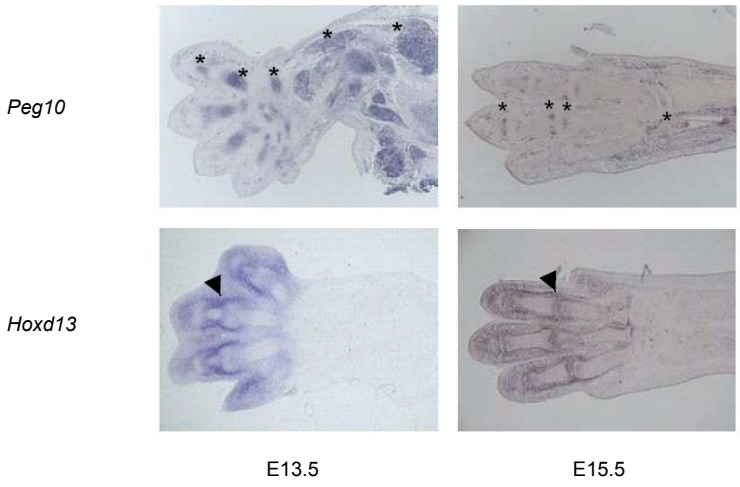


Fig. 27 Section *in situ* hybridisation showing expression of *Peg10* and *Hoxd13* in mouse limbs. At stage E13.5, *Peg10* transcripts are present in the condensations of phalanges, metacarpals and carpals, as well as in muscles (asterisks). *Hoxd13* shows a different expression pattern, being present in the perichondrium (arrowhead). At stage E15.5, both genes seem to maintain their expression pattern, however *Peg10* signals in the distal limb are very weak. Hybridisation of gene-specific probes was performed on paraffin sections (stage E13.5) or cryo-sections (E15.5).

3.2.6 Investigation of other putative Hoxd13 binding proteins

In addition to *Peg10*, five other genes identified in the yeast two-hybrid screen were considered as promising candidate Hoxd13 binding partners. Co-expression of each single clone with LexA_Hoxd13-HD or with GAL_Hoxd13-HD in yeast, confirmed interactions between the candidates and Hoxd13 in both LexA and GAL4 two-hybrid systems (not shown). Further experiments including cellular localisation studies in COS1 cells as well as *in situ* hybridisation are presented in the following sub-sections.

3.2.6.1 Studies on Dlxin-1

The first interesting gene, for which two slightly different clones have been found in the yeast two-hybrid screen, is called *Dlxin-1* or *Maged1* (**m**elanoma **a**nti**ge**n, family **D**, **1**) and it shows a limb-specific expression pattern (Matsuda et al. 2003). The longer prey plasmid isolated in the screen contains a 349 bp long insert (1252–1600 bp according to the GenBank entry AB029448) which covers the region encoding the WQXPXX repeat of the Dlxin-1 protein. This domain is known to interact with Dlx5, a homeobox-containing transcription factor important for limb development (Masuda et al. 2001). The same WQXPXX repeat together with the second distinctive domain of Dlxin-1, the necdin homology domain (NHD), have been shown to be necessary for binding of Msx2, another homeodomain-containing transcription factor which plays a role in osteoblast differentiation. The NHD alone binds Ror2, a receptor tyrosine kinase, mutations of which are causative for limb pathologies: brachydactyly type B (OMIM #113000) or Robinow syndrome (OMIM #268310) (Matsuda et al. 2003). All these data made Dlxin-1 an interesting candidate for a Hoxd13-interaction partner.

Transfection of COS1 cells with pcDNA-Flag vector carrying the 349 bp long *Dlxin-1* fragment isolated in the yeast two-hybrid screen, and subsequent immunocytochemistry studies revealed that the truncated Dlxin-1 protein localises to the cytosol (Fig. 28). Co-transfection with the wtHoxd13-pTL1-HA2 construct did not influence the cellular localisation of the truncated Dlxin-1 protein (Fig. 28), suggesting that Hoxd13 and Dlxin-1 do not bind each other in COS1 cells.

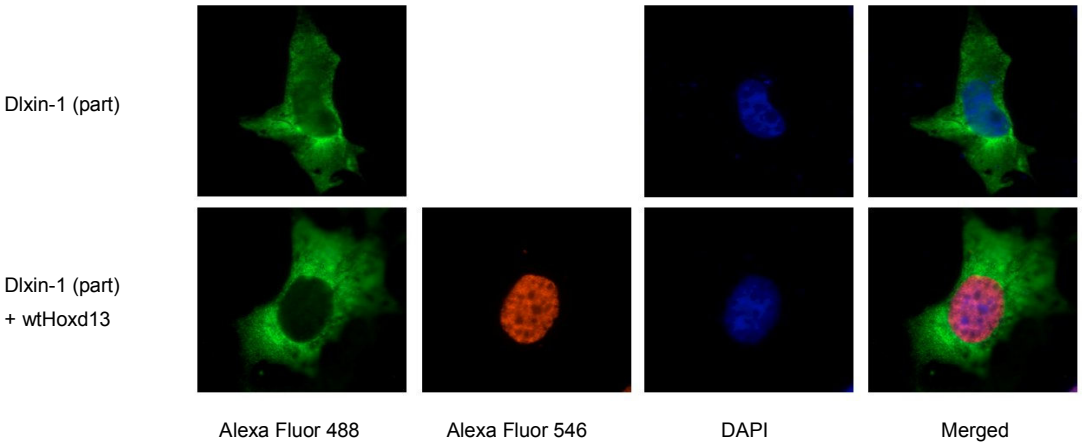


Fig. 28 Overexpression of the Dlxin-1 construct in COS1 cells. Truncated Dlxin-1 (green) localises to the cytosol and does not change its cellular localisation in cells co-transfected with wild type Hoxd13 (red). Cell nuclei are stained with DAPI in blue.

3.2.6.2 LIM domain-containing genes

Among putative Hoxd13 binding partners, three LIM domain-containing proteins were identified in the yeast two-hybrid screen, Wtip (GenBank accession number NP_997095), Limk1 (GenBank accession number NP_034847) and Limd1 (GenBank accession number NP_038888). LIM domains are cysteine- and histidine-rich domains which bind two zinc ions. They appear to mediate protein-protein interactions and are found in many key regulators of developmental pathways (Dawid et al. 1998).

Wtip gene (WT1-interacting protein also known as similar to LIM domains containing 1) has been represented by a single prey clone with the insert covering the region encoding two LIM domains. Whole mount *in situ* hybridisation of the *Wtip* probe to mouse embryos (E10.5 and E11.5) showed a limb specific expression pattern of this gene. Moreover, it has been shown that the Wtip protein binds Ror2, the receptor tyrosine kinase indispensable for normal skeletal development (Verhey van Wijk, not published). Overexpression of full length Wtip in COS1 cells indicated that the protein is localised in the cytosol (Fig. 29) and does not co-localise with Hoxd13 in nuclei of cells co-transfected with the wtHoxd13-pTL1-HA2 construct (Fig. 29).

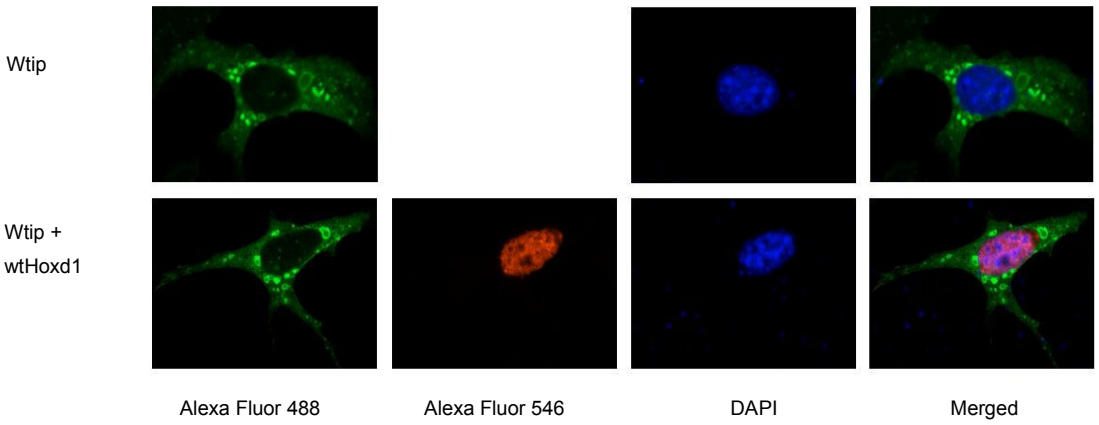


Fig. 29 Overexpression of Wtip in COS1 cells. Wtip (green) localises to the cytosol in single-transfected cells. Co-transfection with the wtHoxd13_pTL1-HA2 construct (red signal) does not influence Wtip cellular localisation. Blue staining (DAPI) marks cell nuclei.

Limk1 (LIM-domain containing protein kinase) is the second LIM protein identified as a potential Hoxd13 interaction partner. The protein contains two LIM domains, a PDZ- and a kinase domain, and it has not been connected to limb development yet. Interestingly, it has been shown that the human homologue, LIMK1, can shuttle between the cytosol and the nucleus (Yang and Mizuno 1999) and is involved in Golgi dynamics, membrane traffic and cytoskeletal organisation (Stanyon and Bernard 1999; Rosso et al. 2004). Two prey clones encoding the Limk1 protein were identified in the yeast two-hybrid screen, the shorter one covers 380-830 bp and the longer one 317-794 bp within the *Limk1* sequence NM_010717. The LIM domain-encoding region is included in both inserts. The shorter *Limk1* fragment was subcloned into the mammalian expression vector pcDNA-Flag and used for cellular localisation studies. The results indicated that the partial Limk1 protein shows both cytosolic and nuclear localisation when overexpressed in COS1 cells, suggesting a possible interaction with the wild type Hoxd13 in the nucleus. However, the same truncated Limk1 protein does not seem to co-localise with the aggregates formed by Hoxd13 lacking the homeodomain (Fig. 30). Whole mount *in situ* hybridisation performed with the *Limk1*-specific probe showed ubiquitous expression of this gene in mouse embryos (E10.5 – E12.5) (Fig. 31).

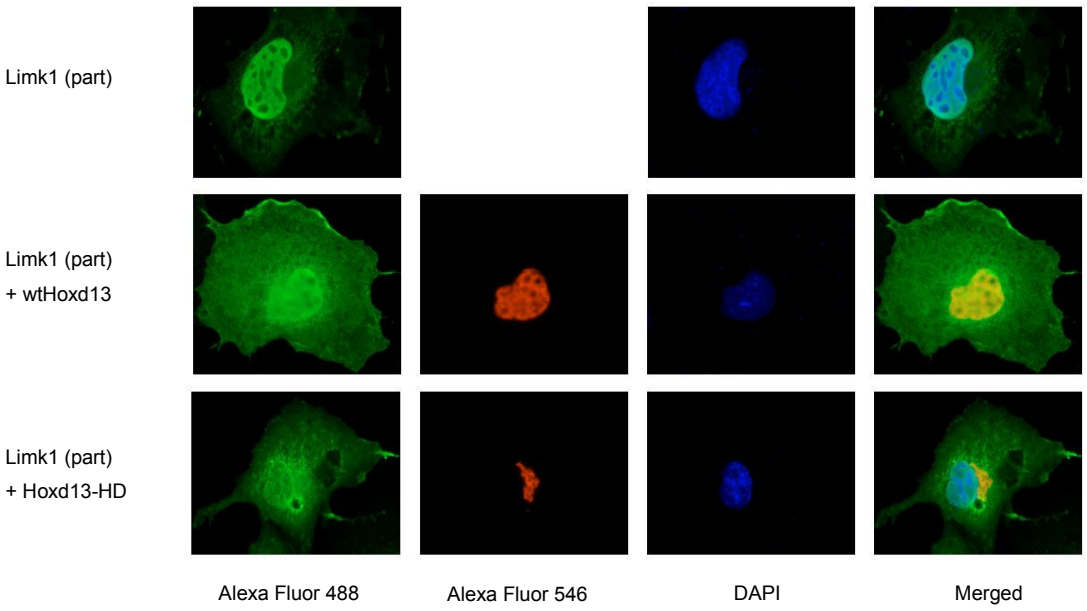


Fig. 30 Partial Limk1 protein (green) is present in both the cytosol and the nucleus (blue) of single-transfected COS1 cells. Upon co-transfection with the wild type Hoxd13 or Hoxd13-HD (red signals) there is no detectable change in the localisation of the partial Limk1 protein.

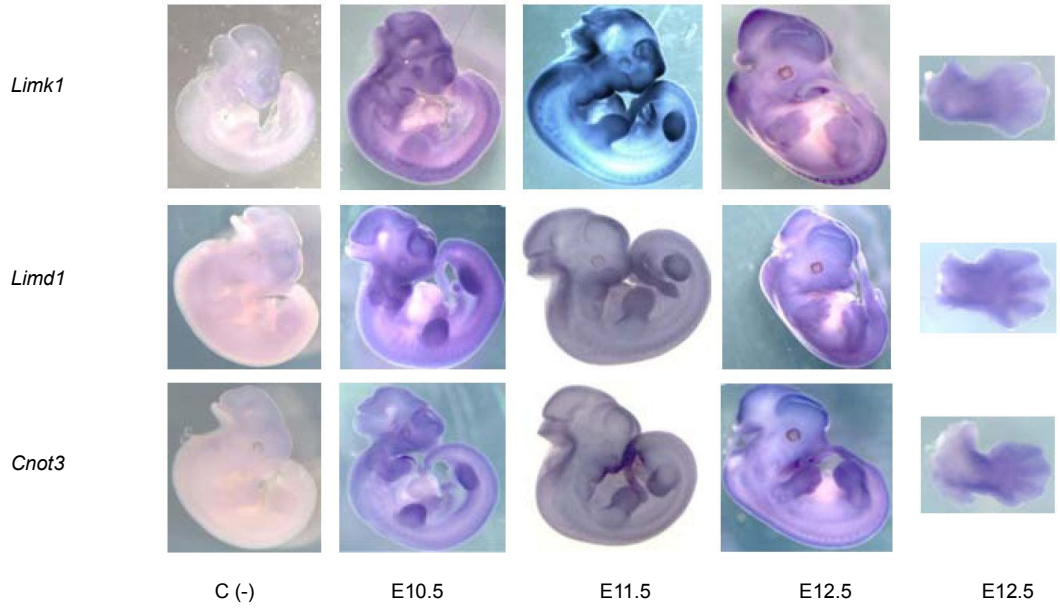


Fig. 31 Whole mount *in situ* hybridisation showing expression of *Limk1*, *Limd1* and *Cnot3* (blue) during mouse embryonic development. All three genes are expressed in limb buds, however *Limk1* expression seems to be more ubiquitous. Negative controls [C(-)] show only very slight staining in the head, indicating specificity of detected signals. For stage E12.5, in addition to the whole embryo, magnified images of forelimbs are shown.

Limd1 (LIM domains containing 1) is a novel gene, being very poorly characterised so far. The human LIMD1 protein is a tumor suppressor. Its ability to bind the retinoblastoma protein (pRB), allows shuttling of LIMD1 between the cytosol and the nucleus (Sharp et al. 2004). Sequence analysis of mouse Limd1 revealed that the protein contains three LIM domains, which according to the yeast two-hybrid data, might be responsible for Hoxd13 binding. In order to examine the expression pattern of *Limd1* during mouse embryonic development, whole mount *in situ* hybridisation experiments were performed. The results showed that *Limd1* is expressed in mice limb buds (Fig. 31). For further analysis, the *Limd1* fragment identified in the yeast two-hybrid screen (1974–2287 bp within the GenBank entry NM_013860), was cloned into the pcDNA vector and overexpressed in COS1 cells. Single transfection experiments indicated that the partial Limd1 protein is present throughout the whole cell. Co-expression with the wild type Hoxd13 or with Hoxd13-HD showed co-localisation of the partial Limd1 protein with both Hoxd13 proteins (Fig. 32).

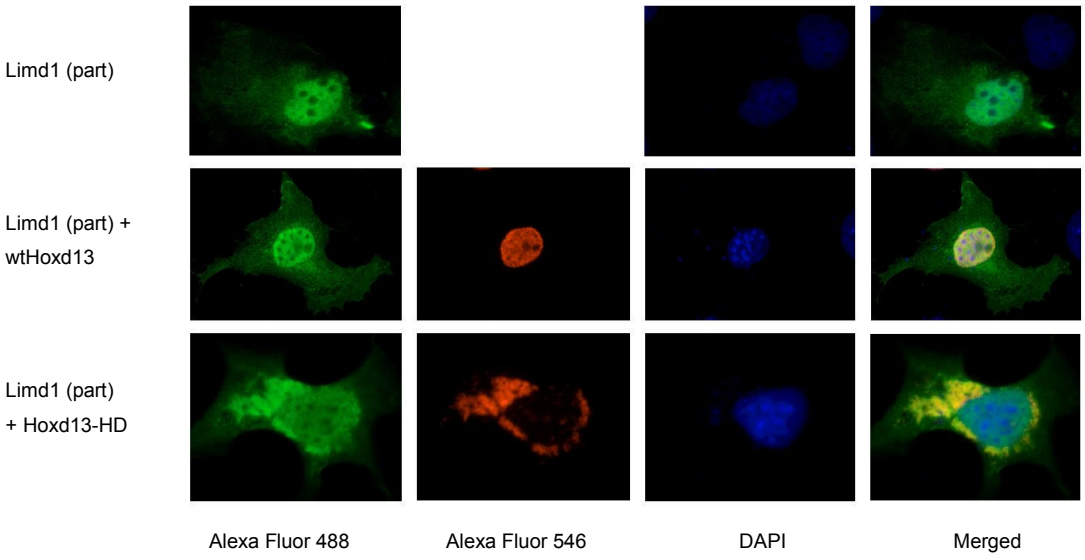


Fig. 32 Overexpression of the truncated Limd1 protein in COS1 cells. Single transfection shows that Limd1 (green) is localised in the nucleus (labelled with DAPI in blue) and spreads in the whole cytosol. Double transfections with wtHoxd13 proteins (red) show co-localisation of both proteins in the nucleus. Furthermore, co-expression with Hoxd13 lacking the homeodomain (red), changes Limd1 pattern in the cytosol, suggesting a possible interaction between both proteins.

3.2.6.3 Cnot3 as a putative Hoxd13 binding partner

Cnot3 (GenBank accesion number NM_146176) encodes a subunit of the CCR4-NOT complex which is known to control global gene expression. Furthermore, it has been shown that mouse and human Cnot3 orthologues can bind other transcriptional regulators (Aoki et al. 2002; Yin et al. 2005). Cnot3 protein has a proline-serine rich region at the C-terminus and two coiled-coil domains at the N-terminus; the function of the latter ones is not clear. The only prey clone corresponding to *Cnot3* which was found in the yeast two-hybrid screen, covers the region encoding the second coiled-coil domain (639–941 bp within the GenBank entry NM_146176).

In order to perform cellular localisation studies, the short prey insert was overexpressed in COS1 cells, either alone or together with *Hoxd13* constructs. The results indicated that the partial Cnot3 protein co-localises with both the wild type Hoxd13 protein and its N-terminal part (Fig. 33). Whole mount *in situ* hybridisation showed that *Cnot3* is expressed in embryonic limb buds (Fig. 31) and therefore it is possible that it could interact with Hoxd13 in order to regulate expression of target genes.

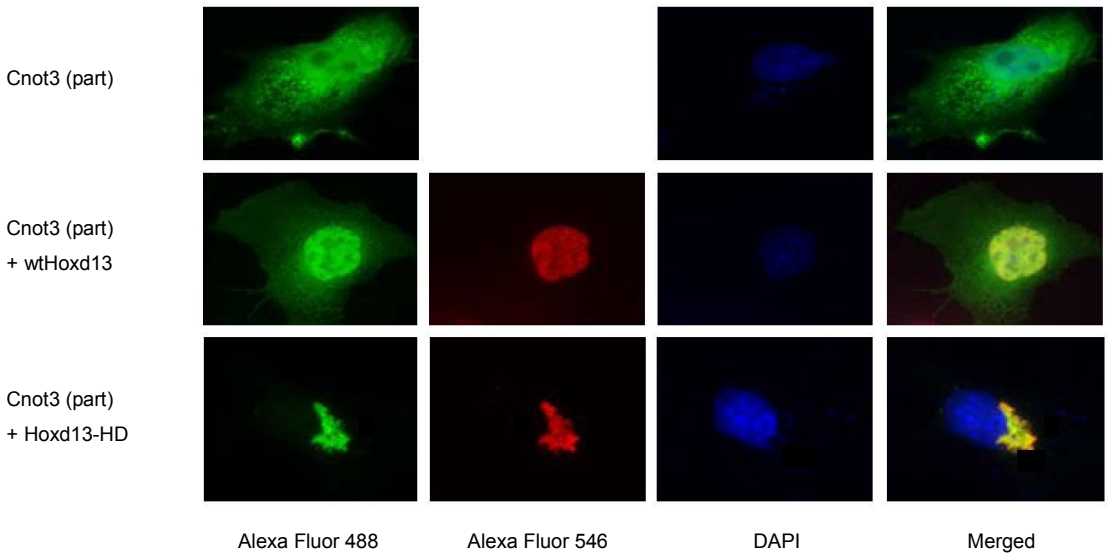


Fig. 33 Overexpression of the partial Cnot3 protein in COS1 cells. Single transfection shows that Cnot3 (green) is localised in the cytosol and the nucleus (labelled with DAPI in blue). Double transfections with Hoxd13 proteins (red signals) indicate that the Cnot3 fragment co-localises in the nucleus with the wild type Hoxd13 and in the cytosol with Hoxd13 lacking the homeodomain.

4 DISCUSSION

4.1 Genotype-phenotype correlation in the patient with a translocation t(2;10)(q31.1;q26.3)

In this study, the translocation t(2;10)(q31.1;q26.3) present in a male patient with skeletal abnormalities and mental retardation has been investigated. Cytogenetic and molecular analysis indicated that the breakpoint on chromosome 10 disrupts the *MGMT* gene, whereas on chromosome 2 the break is localised approximately 390 kb centromeric to the *HOXD* cluster. Moreover, there was no indication that any gene might be disrupted on chromosome 2. The relevance of both breakpoints for the patient's phenotype will be discussed in the following sections.

4.1.1 The *MGMT* gene on chromosome 10 disrupted by the breakpoint does not seem to be responsible for the limb phenotype of the patient

The methylguanine-DNA methyltransferase (*MGMT*) gene disrupted in the patient encodes an enzyme involved in the repair of O^6-alkylguanine-containing DNA. The O^6-alkylguanine-DNA adducts are potent pre-mutagenic lesions, since the modified guanine preferentially mispairs with thymine, instead of building a pair with cytosine. As a result, the G:C → A:T mutations may appear in the DNA. MGMT removes the alkyl group from the DNA by transferring it to its own cysteine residue. This process irreversibly inactivates the protein and the alkylated form of MGMT accumulates as a dead-end product. Therefore, the capacity to repair the O^6-alkylguanine residues is limited by the number of MGMT molecules present in the cell.

In the 1990s two groups created *MGMT* knockout mice in order to study the gene function. Tsuzuki *et al*. showed that the *MGMT* $^{-/-}$ mice were normal, except for a slight growth retardation (Tsuzuki et al. 1996), whereas Glassner *et al.* did not observe any pathologies and abnormalities during development of the knockout mice (Glassner et al. 1999). Therefore, it is rather unlikely that *MGMT* contributes to limb formation. On the contrary, recent data suggest a link between MGMT and cancer. Downregulation of the *MGMT* gene might participate in tumour formation, whereas its upregulation seems to prevent this process

(Sakumi et al. 1997; Nakamura et al. 2001; Oue et al. 2001; Reese et al. 2001; Zhou et al. 2001; Smith-Sorensen et al. 2002). Thus, all these data suggest that disruption of the *MGMT* gene probably did not cause the limb malformations present in the patient.

4.1.2 Could the disruption of *MGMT* have influenced the mental status of the patient?

MGMT is ubiquitously expressed in adult human tissues, as shown in this study in Northern blot experiments. However, different tissues express this gene at various levels, with the highest expression in liver and a very low expression in brain. A similar situation can be seen in foetal tissues. The amount of mRNA of the *MGMT* gene in different cell types correlates with the enzymatic activity of the protein measured by several groups (Grafstrom et al. 1984; Wiestler et al. 1984; Pegg et al. 1985). However, during embryonic development a critical factor can be the rate of DNA repair relative to that of DNA replication. Therefore, some authors compared the activity of MGMT to the activity of DNA polymerase in extracts from different tissues (Krokan et al. 1983). Measured in this way, the lowest activity of MGMT was found in brain of most foetuses. These data suggest that the brain might be more exposed to the risk of DNA mutations than other organs.

Quantitative analysis of *MGMT* expression revealed that in lymphoblastoid cell lines derived from the translocation patient, the level of *MGMT* mRNA was reduced by approximately 50% compared to the control. This is in accordance with the finding that the translocation in the patient caused the disruption of one *MGMT* allele. Decrease of *MGMT* expression in patient's brain during embryogenesis might have led to accumulation of DNA mutations in neuronal tissue, which might have influenced the development of cognitive functions in the patient.

4.1.3 There is no evidence for any gene disrupted on chromosome 2 in the patient with the translocation t(2;10)(q31.1;q26.3)

The results of RT-PCR experiments suggest that no gene is disrupted on the chromosome 2 in the translocation patient. However, recent sequence data available from the UCSC Genome Browser Gateway (May 2004 assembly) indicate that much more ESTs are present on the breakpoint-spanning BAC RP11-538A12 and on the neighbouring BACs than thought before. Since these ESTs are either unspliced or repetitive, it is rather unlikely that any of them might represent a gene.

4.1.4 *HOXD* genes located close to the breakpoint on chromosome 2 are good candidates for the limb phenotype in the patient

The translocation patient presented in this study shows a phenotype similar to that of SPD patients with mutations in the *HOXD13* gene. However, sequence analysis of *HOXD13* in the patient showed no abnormalities within this gene, indicating that the SPD phenotype must have been caused by other defects, most likely by effects of the translocation.

Experiments performed in mice indicate that four posterior *Hox* genes, *Hoxa13, Hoxd11*, *Hoxd12* and *Hoxd13* can control digit development (Dolle et al. 1993; Davis and Capecchi 1996; Kondo et al. 1996; Zakany et al. 1997; Kondo et al. 1998). Progressive reduction in the dose of these genes results in adactylous limbs in mutant mice. Interestingly, the intermediate stage in the pentadactyl to adactyl transition is characterised by polydactyly. The latter phenotype is present in a triple *Hoxd11*/*Hoxd12*/*Hoxd13* knockout mouse, suggesting that SPD is caused by a loss of function of the *HOXD11-13* genes, rather than inactivation of *HOXD13* alone. Therefore, it is very likely that in the translocation patient described in this study, the breakpoint located in the vicinity of the *HOXD* genes affected their proper expression leading to the SPD phenotype. Various hypotheses for the putative deregulation of the *HOXD* genes are discussed in the following sections.

4.1.5 Chromosomal rearrangements can cause disorders in humans and mice via position effect

Balanced chromosomal rearrangements may be associated with pathological phenotypes in humans. One possible mechanism for this is disruption of a gene caused by a breakpoint. However, during the past few years several patients with balanced chromosomal rearrangements and a disease phenotype, who do not carry a disrupted gene, were reported (Kleinjan and van Heyningen 1998; Marlin et al. 1999; Di Paola et al. 2004; Kleinjan and van Heyningen 2004; Lower et al. 2004; Muncke et al. 2004; Tadin-Strapps et al. 2004). This phenomenon can be explained by a so called position effect, which is defined as a deleterious change in the level of gene expression caused by a change in the normal chromosomal environment of the gene (Kleinjan and van Heyningen 1998).

There are several known cases in humans, where breakpoints involving chromosomal band 2q31 are thought to cause improper development of skeletal structures via position effect. One example is a family with mesomelic dysplasia and vertebral defects, carrying a balanced

translocation t(2;8) (Spitz et al. 2002). In addition, a female patient reported with a pericentric inversion inv(2)(p15q31.1) showed bilateral aplasia of radial, ulnar and fibular bones, hypoplasia and dislocation of both tibiae and defects in metacarpals and phalanges (Dlugaszewska et al. 2005). Another known case is a female patient with a balanced translocation t(2;10)(q31.1;q23.3) showing shortening and aplasia of upper limb structures, affecting zeugo- and autopods, and with a slight dextro-convex scoliosis (Dlugaszewska et al. 2005). In all these cases the breakpoints on the long arm of chromosome 2 have been mapped to regions telomeric to the *HOXD* cluster. It is believed that these rearrangements cause position effects, resulting in misregulation of *HOXD* gene expression.

An additional support comes from the analysis of the X-ray induced mouse mutant *ulnaless* (*Ul*), which shows a very severe phenotype, with an affected zeugopod and almost complete absence of ulnae (Herault et al. 1997; Peichel et al. 1997). Recently, it has been shown that an inversion occurred on chromosome 2 in the *ulnaless* mouse, with the breakpoints surrounding the *Hoxd* cluster (Spitz et al. 2003). It is very likely that this rearrangement is the reason for misregulation of posterior *Hoxd* expression observed in limb buds of the *Ul* mouse (Peichel et al. 1997) and therefore the cause for the limb phenotype.

4.1.5.1 Mechanisms leading to a position effect

In general, different mechanisms can lead to a position effect. First, a chromosomal rearrangement might separate regulatory elements from the gene, thus resulting in its misexpression. Secondly, a gene and an enhancer element from another transcriptional unit could be juxtaposed by a rearrangement. Thirdly, one gene together with its regulatory elements might be placed next to a second gene, and the competition for the same regulatory elements between both of them could change the expression level of the first gene. Lastly, the rearrangement could lead to position effect variegation. Assuming that the translocation in the patient presented in this study led to a change in *HOXD* gene expression, it would be interesting to find out which of these mechanisms contributed to the disease phenotype.

It has been suggested that expression of the posterior *Hoxd* genes in limb buds is dependent on both local regulatory elements lying in the direct vicinity of these genes as well as on enhancers lying outside of the cluster. Especially, two global elements, a putative early limb control region and the digit enhancer, lying respectively 3' and 5' of the *Hoxd* genes are involved in the regulation of the whole complex (Deschamps 2004). Since the breakpoint in the translocation patient occurred 5' to the *HOXD* complex, the important question to answer

is, whether 5' regulatory regions are affected by the rearrangement. Recently, the position of the human digit enhancer has been narrowed down to a 40 kb long sequence present on BAC clone 504O20 (Spitz et al. 2003). In the translocation patient, this region is neither disrupted nor separated from the *HOXD* cluster by the rearrangement. However, it is still possible that other regulatory elements responsible for *HOXD* expression may also be present further upstream to this digit enhancer. This hypothesis is supported by the fact that the region 5' to the 2q31.1 breakpoint in the translocation patient was shown to be highly similar to the corresponding region in mouse. It is broadly accepted that sequence conservation maintained during evolution, reflects an important role of the conserved elements. Since there is no evidence for a gene in this region, the presence of regulatory elements is one possible explanation for the sequence similarity close to the breakpoint. If this theory was true, it would be possible that these, till now unknown, regions could be affected by the translocation.

Another mechanism leading to the position effect suggests that the rearrangement, which brings together two different chromosomes, could place the *HOXD* cluster under the influence of another regulatory element located on chromosome 10 or that the impact of the digit enhancer on the *HOX* genes could be reduced by another transcriptional unit lying on the centromeric side of the breakpoint. The first suggestion is possible, however at the present state of knowledge it has a very speculative character. Nothing is currently known about regulatory elements of *MGMT* or other genes in the vicinity of the breakpoint at 10q26.3. The second mechanism is rather unlikely, since *MKI67*, the closest gene located approximately 1.4 Mb centromeric to the chromosome 10 breakpoint, lies probably too far away to be able to compete with *HOXD* genes for the digit enhancer.

The last mechanism, the classical position effect variegation causes silencing of a gene by inserting it into or nearby a heterochromatic region. Alternatively, a long-range insulator or another boundary element may be removed by the rearrangement, which results in spreading the heterochromatin and inactivating of the whole locus. However, it is rather unlikely that one of these mechanisms caused the putative misregulation of the *HOXD* genes in the patient, since the whole *HOXD* cluster has been placed into the middle of the transcriptionally active region at 10q26.

4.1.5.2 HOXD gene expression may be influenced by the accessibility of the entire cluster for transcription factors

Regulation of posterior *Hoxd* genes in limb buds is most likely regulated by the interplay between *cis*-acting elements and *trans*-acting factors. Till now two groups of genes, *Polycomb* (*PcG*) and *trithorax* (*trxG*) have been implicated in maintenance of the active or the silent state of *Hox* genes in *Drosophila* (Simon 1995). Later, *PcG* and *trxG* homologues were also found in mammals (Schumacher and Magnuson 1997). Different *Polycomb* or *trithorax* mouse mutants showed homeotic transformations in the vertebral skeleton which corresponded to the shift in *Hox* gene expression domains (Alkema et al. 1995; Akasaka et al. 1996). These observations suggested that the role for *Polycomb* and *trithorax* genes in the regulation of *Hox* gene expression along the main body axis is conserved between *Drosophila* and mammals.

Results published during the last few years provided evidence that PcG genes may also play a role in proper expression of posterior *Hoxd* genes in mouse limb buds (Barna et al. 2000; Barna et al. 2002). According to these studies, Bmi-1 belonging to the Polycomb group proteins binds Plzf, a nuclear zinc finger protein. Plzf can recognise and bind different regulatory sequences within the *Hoxd* locus, and it can remodel the chromatin by histone deacetylation, which results in *Hoxd* repression. In addition, Plzf can mediate long-distance interactions between *cis* regulatory elements within the *Hoxd* locus. Therefore, Plzf together with its interacting partner Bmi-1 are excellent candidates for factors which could integrate both local and global regulatory mechanisms in order to mediate the correct expression of posterior *Hoxd* genes. It is very likely that they are not the only players in this complex system and that other regulatory proteins will be discovered soon.

Binding of transcriptional regulators to DNA might be dependent on proper chromatin architecture (Kornberg and Lorch 1992; Nourani et al. 2004). Since it has been proposed that the chromatin structure of any locus can be determined by the combination of *cis*-acting elements and by the wider chromosomal and nuclear environment (Kleinjan and van Heyningen 2004), it is plausible that chromosome rearrangements could alter the chromatin architecture. In fact, changes in chromatin structure have been proposed following insertion of some transgenes (de Graaff et al. 2003) or in case of small deletions (Jiang et al. 2003). Hence, it is possible that the translocation, which has occurred in the patient presented here, changed the chromatin structure around the *HOXD* locus. This event might have modified the access to the chromatin for transcription factors, disturbed the interplay between *cis*- and

trans-acting regulatory elements and resulted in deregulation of *HOXD* gene expression and the limb phenotype in the patient.

4.1.6 Are others genes possibly involved in the patient's phenotype?

Interestingly, changes in the global chromatin structure might also influence expression of other genes near the chromosome 2 breakpoint. The closest gene, located only 220 kb telomeric to the breakpoint in the translocation patient, is *KIAA1715*. It spans over 75 kb of the genomic sequence at 2q31 and it is transcribed from the opposite strand compared to the *HOXD* cluster. Another candidate, *EXV2*, is lying 380 kb telomeric to the breakpoint and has the same orientation as *KIAA1715*. Mouse homologues of both genes, *Lnp* and *Evx2*, show the same expression pattern in limb buds and external genitalia as *Hoxd13*, which suggests that all three transcription units are under the control of the same regulatory sequences. In addition, *Lnp* is also expressed in the developing central nervous system in a highly similar pattern to that of *Evx2*, and it has a specific expression domain in the eyes, the heart and the forebrain (Spitz et al. 2003). The neural enhancer that may activate *KIAA1715* and *EVX2* is located in part within the same 40 kb region as the digit enhancer mentioned earlier. Therefore, it is possible that both the limb and the neuronal expression domains of *KIAA1715* and *EVX2* have been affected by the translocation via position effect. Since the translocation patient has cognitive deficits in addition to limb abnormalities, it is tempting to link the central nervous system phenotype with disturbed expression of *KIAA1715* or *EVX2*.

4.2 Search for interaction partners of Hoxd13 protein

The second part of this study focused on the search for Hoxd13 interaction partners in order to shed more light on the molecular basis of limb development. It has been suggested that Hox proteins act in complexes (see also section 1.3.7), however little is known about Hox cofactors playing a role in the development of distal limbs. To address this question, a yeast two-hybrid screen was performed, and in this approach many putative Hoxd13 interaction partners were identified. Several candidates were analysed in more detail, and the results of these studies will be discussed in the following sections.

4.2.1 Peg10 is a putative Hoxd13 binding protein

Peg10, the paternally expressed gene 10, has been mapped to an approximately 1 Mb long cluster of imprinted genes on mouse chromosome 6 (Ono et al. 2003). It is highly conserved between species, with homologous sequences in humans, cow, rat, mink, pig, rhesus and pufferfish. The presence of two long overlapping open reading frames (called ORF1 and OFR2) and the similarity of their predicted amino acid sequences to retroviral proteins Gag and Pol suggested that *Peg10* is a retrotransposon fossil in the mammalian genome (Ono et al. 2001; Shigemoto et al. 2001). Similarly as for other viral genes, a single *Peg10* transcript gives rise to two partially different proteins (Shigemoto et al. 2001; Lux et al. 2005).

4.2.1.1 Parts of the Peg10 protein bind Hoxd13-HD in yeast

In the yeast two-hybrid screen performed with the LexA_Hoxd13-HD bait, eight different *Peg10* clones were identified. Five of them contained the 5' part of the *Peg10-ORF2*, whereas three other clones were very similar to each other and covered the 3' end of the *Peg10-ORF2*. Binding to Hoxd13 lacking the homeodomain was confirmed in the LexA and GAL4 yeast systems for five positive clones originating from both the 5' and the 3' ends of the ORF2. Therefore, it seemed plausible that the full length Peg10 and Hoxd13 proteins could be real binding partners. Moreover, the results suggested that the N-terminus of Hoxd13 is sufficient for the binding. For Peg10, the putative interacting regions are located at the N- and the C-terminus of the Peg10-ORF2, and are separated by approximately 150 amino acids. However, it is possible that these two regions could be brought into proximity by protein folding and in this way they might be both responsible for binding to Hoxd13.

4.2.1.2 Interaction between Peg10 and Hoxd13 in mammalian cells

In spite of intensive attempts it was not possible to clone repetitive sequences present in the *Peg10* gene, therefore for further experiments a partial *Peg10* clone containing the ORF2-encoding sequence identical to GenBank entry AB091827 was used. Subsequent overexpression experiments showed that Peg10-ORF2 perfectly co-localises with the Hoxd13 protein lacking the homeodomain in the cytosol of COS1 cells. Moreover, similar studies have been performed for the wildtype Hoxd13. In the COS1 cells transfected with the Peg10-ORF2 construct, the overexpressed protein showed solely the cytosolic localisation. Co-expression of wildtype Hoxd13 and Peg10-ORF2 induced in some cells a clear change in

Peg10 localisation, from the cytosol to the nucleus. This alteration in cellular localisation suggests that Peg10 might interact with Hoxd13. However, it is not clear, why the co-localisation could not be seen in every double transfected cell. One possible explanation is that Hoxd13 and Peg10 bind each other only in a specific phase of the cell cycle, similarly as it is known for several proteins involved in the DNA-repair or the DNA-replication (Taniguchi et al. 2002; Fan et al. 2004).

The results of the coimmunoprecipitation assay clearly show that Peg10-ORF2 binds to full length Hoxd13. Therefore, it is very likely that the long version of Peg10 containing both ORF1 and ORF2 could also interact with Hoxd13.

Further experiments are needed to answer the questions, whether Hoxd13 binds Peg10 in the direct way and whether this interaction is dependent on DNA binding. In several reported cases, Hox proteins have been shown to interact directly with various proteins, for instance with Pbx or Meis (Shen et al. 1996; Shen et al. 1997), whereas other cofactors like Prep or Sp1 might be bound to Hox in the indirect way (Fognani et al. 2002; Suzuki et al. 2003). Moreover, in the same reports it has been shown that formation of complexes between Hoxd13 and Pbx or Meis is dependent on DNA binding. Interestingly, in all these cases interactions between Hox and other proteins require only the N-terminus of the Hox protein. The same seems to be true for Peg10 binding, since Hoxd13 lacking the homeodomain co-precipitated with Peg10-ORF2. In contrast, binding to Gli3, occurs via the homeodomain of Hoxd13 (Chen et al. 2004). This suggests that Hoxd13 could bind different factors via various domains. Hence, it is possible that Hoxd13, similarly as anterior Hox proteins, could participate in multimeric complexes.

4.2.1.3 Does Peg10 bind Hoxd13 in vivo?

Although it has been shown that Hoxd13 and parts of the Peg10 protein could interact in transformed mammalian cells, the question of much higher biological relevance is whether the binding between Peg10 and Hoxd13 could also occur during mouse embryogenesis. In order to address this question, the expression profile of *Peg10* was examined by whole mount *in situ* hybridisation. Comparison with the known *Hoxd13* expression domains (Albrecht et al. 2002) revealed that at the early stages of mouse development (E10.5) expression of both genes can be detected in similar domains of the distal limb bud. At a slightly later stage (E11.5) expression domain of *Peg10* becomes broader and certainly covers the area expressing *Hoxd13*. At stage E12.5, *Peg10* transcripts can be observed among others in digits

and in the proximal mesoderm, whereas *Hoxd13* is present in the interdigital zones. Different expression domains of both genes were also shown by section *in situ* hybridisation at the later stages of limb development. All these data suggest that Peg10 and Hoxd13 could interact at early stages of development, being expressed in the same regions of limb buds. However, further studies will be needed to confirm the presence of Hoxd13/Peg10 complexes in these tissues. At the later stages of embryogenesis, *Hoxd13* and *Peg10* genes are expressed in different cells and therefore it is rather unlikely that their products can bind each other.

4.2.1.4 A putative role of Peg10 proteins and Hoxd13/Peg10 complexes in limb development

Functional analyses of Peg10 were performed almost exclusively for the protein encoded by ORF1 (Okabe et al. 2003; Tsou et al. 2003). However, the longer version of Peg10 contains both ORF1 and ORF2, suggesting that it could share functional properties with the shorter prtotein. The endogenous *PEG10* has been shown to be upregulated in human hepatocellular carcinoma cells and during liver regeneration in mice. This suggests that the *PEG10* gene product could exert some regulatory function in cell cycle progression. Further experiments supported this hypothesis by showing that overexpression of human PEG10-ORF1 results in an increased rate of G_1 to S transition in 293T cells (Tsou et al. 2003). Moreover, Tsou *et al*. showed that the growth of hepatoma cells is suppressed after their transfection with *PEG10*-specific antisense oligonucleotides. In line with these results, PEG10 overexpression experiments indicated a protective role of this protein in apoptosis (Okabe et al. 2003). All these data are in agreement with the expression pattern of the *Peg10* gene observed in this study in mouse embryonic limbs containing a large number of highly proliferating cells. Interdigital zones of the limb buds, where apoptosis occurs, showed no expression of *Peg10* mRNA.

In vitro studies on the human PEG10-ORF1 protein indicate that it can bind to the activin receptor-like kinase 1 (ALK1) and to other receptors for members of the transforming growth factor beta (TGF-β) superfamily, for instance BMP receptors (Lux et al. 2005). These data provide a link between PEG10 and BMP signalling, which plays an established role in limb development. Binding between diffent receptors and PEG10 assumes that the latter protein must be present in the cytosol, and indeed a few authors report cytosolic localisation of PEG10-ORF1 (Okabe et al. 2003; Tsou et al. 2003; Lux et al. 2005). On the other hand, there are some hints suggesting that PEG10 might be a transcriptional regulator. Sequence comparison revealed that the murine Peg10-ORF1 protein is probably identical to the myelin

expression factor 3 (GenBank acc. number of the nucleotide sequence: AF302691), a brain-derived transcriptional activator containing a predicted nuclear localisation signal (NLS) (Steplewski et al. 1998). Moreover, Okabe *et al.* showed that PEG10-ORF1 can be found throughout the whole cell, thus also in the nucleus (Okabe et al. 2003). Up to now, there are no reports showing cellular localisation of the longer version of PEG10. In this study it was shown that the protein corresponding to the murine Peg10-ORF2 is present in the cytosol, when overexpressed in COS1 cells. However, its localisation can be changed upon co-expression with Hoxd13, and Peg10-ORF2 can co-localise with Hoxd13 in the nucleus. The putative role of Peg10 as a transcription factor can be also supported by the fact that both ORF1 and ORF2 of Peg10 can encode zinc finger domains, which are commonly known DNA-binding motifs.

Considering all these data it is tempting to hypothesise that Peg10 might be involved in BMP and TGF-β signal transduction and it might shuttle between different cell compartments. When present in the nucleus, Peg10 could form complexes with Hoxd13 in order to regulate expression of various target genes during embryogenesis. Since different *Hox* genes, as well as *Peg10*, were described as oncogenes (Okabe et al. 2003; Lawrence et al. 2005), their products could act synergistically by activating other factors responsible for cell proliferation. However, Hoxd13 can also activate genes involved in other processes. For instance, the Hoxd13 protein present in the interdigital zones is thought to induce expression of BMPs which mediate apoptosis, a process necessary for digit separation (Suzuki et al. 2003). Since *Peg10* expression has not been observed in the interdigital zones, it could not co-operate with Hoxd13 in the activation of apoptotic genes. Therefore, various Hoxd13 activities might require numerous interaction partners that would assure the specificity of the DNA binding and would modulate Hoxd13 function.

4.2.1.5 Ala-stretch mutations within Hoxd13 do not influence the binding to Peg10

Human *HOX*-associated pathologies have been extensively investigated for several years. One of these disorders, synpolydactyly (SPD), is caused by extensions of the polyalanine stretch in the HOXD13 protein (Akarsu et al. 1996; Muragaki et al. 1996; Goodman et al. 1997; Kjaer et al. 2002). Studies performed in mice and in transformed cell lines suggested that a similar mutation in murine Hoxd13 results in a misfolded protein which is either degraded or accumulates in the cytosol and therefore cannot fulfil its normal function (Albrecht et al. 2004). In order to see whether mutations in Hoxd13 change its ability to bind

interaction partners, immunocytochemistry and coimmunoprecipitation studies with Peg10 and mutant Hoxd13 proteins were performed.

The results showed that the binding of Peg10-ORF2 to Hoxd13 is not affected by different lengths of alanine expansions in Hoxd13, since both Hoxd13+14Ala and Hoxd13_2Ala co-precipitate with Peg10-ORF2.

Immunocytochemistry studies showed that Peg10-ORF2 co-localises with wild type Hoxd13 or Hox13_2Ala in nuclei of COS1 cells. However, overexpression of Peg10-ORF2 together with the pathogenic Hoxd13+14Ala mutant, changes the cellular localisation of both proteins. Peg10 becomes incorporated into Hox aggregates and cannot enter the nucleus anymore. This suggests that the normal function of Peg10 might be abolished. Similarly, other authors proposed that co-localisation with aggregates might alter functions of various proteins (Boutell et al. 1999; Steffan et al. 2000; Albrecht et al. 2004). Therefore, it is possible that the aggregate sequestration of Peg10 and other Hoxd13 binding partners might contribute to the severity of SPD.

4.2.2 Other potential Hoxd13 interaction partners

Five other putative Hoxd13 interaction partners identified in the yeast two-hybrid screen were examined in this study. Two of them, Dlxin-1 and Wtip, do not co-localise with wildtype Hoxd13 in COS1 cells. Similarly, Limk1 does not show any clear colocalisation with Hoxd13 lacking the homeodomain. These data suggest that there is no binding between Hoxd13 and these candidates in COS1 cells.

Two other genes, *Limd1* and *Cnot3*, seem to be much more interesting. Both are expressed in limb buds during mouse embryonic development. Furthermore, the partial Limd1 and Cnot3 proteins co-localise with both Hoxd13-HD and wildtype Hoxd13 in mammalian cells. Cnot3, which is a member of a transcription regulatory complex, might modify Hoxd13 function and influence expression of different target genes. Limd1 is a novel gene, therefore more studies would be necessary to uncover its function. In general, for both candidates further experiments are needed, involving cloning of full length ORFs and subsequent co-localisation and binding studies in mammalian cells.

4.2.3 Outlook

HOXD-associated human disorders are being currently investigated in detail. Findings from the recent few years allowed researchers to identify several mechanisms on the cellular and

the DNA level that contribute to the ethiology of these disorders. However, an important step in better understanding of these processes is the identification of Hoxd13 binding proteins. Several potential Hoxd13 interaction partners were presented in this thesis, however only one of them, Peg10, was analysed in more detail. Future studies should be performed in order to confirm Hoxd13/Peg10 interaction *in vivo*. For this purpose, Peg10-specific antibodies were generated. Furthermore, functional analysis of Peg10 protein could be performed, including generation of *Peg10* $^{-/-}$ and *Peg10*/*Hoxd13* double knockout mice in order to observe the genetic interaction between both partners. Moreover, Limd1 and Cnot3 should be further analysed to confirm or to exclude their ability to bind Hoxd13. Preliminary experiments might be performed *in vitro*, similarly as it was done for Peg10. In case the results are positive, further studies would be needed to confirm the potential interactions *in vivo* and to elucidate cellular pathways in which Limd1 and Cnot3 proteins take part.

5 SUMMARY

The *HOX* genes encode an evolutionary conserved family of transcription factors playing an important role in embryonic development. In particular, the posterior *HOXD* genes are involved in limb patterning in higher vertebrates. Different *HOXD* mutations are associated with limb anomalies in humans and in mice. One of them, synpolydactyly (SPD), a severe dominant disorder characterised by digit duplications and webbing, has been connected to various mutations within the *HOXD13* gene.

In this study the chromosome translocation t(2;10)(q31.1;q26.3) present in a male patient with mental retardation and SPD was investigated. The breakpoints were mapped and cloned using cytogenetic and molecular approaches. The results indicated that on chromosome 10, *MGMT*, a gene coding for a DNA-repair enzyme has been disrupted. However, up to now there is no link between this gene and limb development, suggesting that *MGMT* is not responsible for the skeletal abnormalities present in the patient. On chromosome 2, the breakpoint occurred approximately 390 kb centromeric to the *HOXD* complex and did not truncate any known gene. However, it has been recently shown that changes in the global chromosomal environment might lead to misregulation of gene expression by so-called position effects. Therefore, it is likely that the translocation affected the precise regulation of *HOXD* expression resulting in their loss of function and leading to limb abnormalities in the patient.

Studies on molecular bases of limb development and pathogenesis performed during the last years led to identification of many pathways and mechanisms responsible for these processes. However, many questions still need to be answered. To find new players involved in distal limb patterning, I have searched for Hoxd13 interaction partners using the yeast two-hybrid technique. Several candidates were identified and further tested in yeast and mammalian systems. One of them, Peg10 has been shown to co-localise with wildtype and mutant Hoxd13 proteins in COS1 cells. Moreover, binding between these proteins was confirmed in the coimmunoprecipitation assays. Whole mount *in situ* hybridisation experiments indicated that *Peg10* is expressed in distal limb buds during mouse embryogenesis. At the early stages of limb development (E10.5 and E11.5) the *Peg10* expression domain overlaps with that of *Hoxd13*, suggesting that both proteins might interact *in vivo*. The results presented in this study together with literature data suggest that Peg10 could modulate Hoxd13 function and

that both interacting proteins might co-operate in order to regulate transcription of various target genes. However, additional experiments will be necessary to confirm this hypothesis and to explain in detail the role of Hoxd13/Peg10 complexes *in vivo*. Moreover, it would be interesting to further analyse two other potential Hoxd13 binding partners, Limd1 and Cnot3. Binding assays and functional studies could give us new data valuable for better understanding limb patterning processes.

6 ZUSAMMENFASSUNG

Die *HOX* Gene kodieren für eine evolutionär konservierte Familie von Transkriptionsfaktoren, die eine wichtige Rolle während der Embryogenese spielen. Insbesondere sind die posterioren *HOXD*-Gene an der Ausbildung der Extremitäten beteiligt. Verschiedene *HOXD*-Mutationen wurden mit Extremitätenfehlbildung bei Mensch und Maus assoziiert. Eine dieser Krankheiten, die Synpolydaktylie (SPD), wird durch verschiedene Mutationen im *HOXD13* Gen verursacht.

In dieser Studie wurde die Translokation t(2;10)(q31.1;q26.3) in einem männlichen Patienten untersucht. Der Junge ist geistig behindert und zeigt verschiedene Knochenanomalien, inklusive SPD. Die chromosomalen Bruchpunkte wurden mit Hilfe von unterschiedlichen zytogenetischen und molekulargenetischen Methoden kartiert und kloniert. Die Ergebnisse zeigten, dass der Bruchpunkt auf dem Chromosom 10 das *MGMT* Gen unterbricht. Da dieses Gen für ein DNA-Reparaturenzym kodiert und da es keine Hinweise zu dessen Rolle in der Extremitätenentwicklung gibt, ist es unwahrscheinlich, dass *MGMT* für die Knochenfehlbildung bei diesem Patienten verantwortlich ist. Der Bruchpunkt auf dem Chromosom 2 wurde ungefähr 390 kb centromerisch vom *HOXD*-Cluster kartiert. Er unterbricht kein bekanntes Gen. In der letzten Zeit wurde bekannt, dass die Fehlregulation der Genexpression mit einer veränderten chromosomalen Umgebung zusammenhängen kann. Dieses Phänomen wird als Positionseffekt bezeichnet. Daher ist es sehr wahrscheinlich, dass die Translokation die präzise Regulation der *HOXD* Gene beeinflusst hat. Dies könnte zum Verlust der *HOXD* Funktion und zum SPD-Phänotyp im Patienten führen.

Studien der letzten Jahre haben zur Identifizierung einer Reihe von Mechanismen geführt, die an Extremitätenentwicklung und –pathogenese beteiligt sind. Trotzdem gibt es dabei noch viele Fragen, die beantwortet werden müssen. Um neue Moleküle zu identifizieren, die in diesen Prozessen eine wichtige Rolle spielen, habe ich nach Hoxd13-Interaktionspartnern mit Hilfe der Hefe-2-Hybrid-Methode gesucht. Mehrere Kandidaten wurden als potenzielle Hoxd13-Bindungsproteine identifiziert und im Hefe- und Affennierenzellsystem weiter analysiert. Einer der Kandidaten, Peg10, kolokalisiert mit den Wildtyp- und den mutanten Hoxd13-Proteinen in COS1-Zellen. Zusätzlich wurde die Bindung zwischen Peg10 und Hoxd13 Proteinen durch Koimmunopräzipitationsstudien nachgewiesen. Die Ergebnisse der Whole mount *in situ* Hybridisierung zeigten, dass *Peg10* während der Mausembryogenese in

den distalen Teilen der Extremitätenknospen exprimiert wird. Die Expressionsdomänen von *Peg10* und *Hoxd13* überlappen sich in den frühen Stadien der Extremitätenentwicklung (E10.5 und E11.5). Das deutet darauf hin, dass die beiden Proteine *in vivo* interagieren könnten. Die Ergebnisse dieser Studie, zusammen mit bereits publizierten Daten, lassen vermuten, dass Peg10 die Funktion von Hoxd13 moduliert und dass die beiden Proteine die Transkription von verschiedenen Zielgenen regulieren. Weitere Experimente sind nötig, um diese Hypothese zu bestätigen und die genaue Rolle der Hoxd13/Peg10 Komplexe *in vivo* zu klären. Zwei andere Kandidaten für Hoxd13-Bindungspartner, Limd1 und Cnot 3, sollten weiter untersucht werden, um ihre möglichen Interaktionen mit Hoxd13 zu bestätigen. Die funktionellen Analysen der Kandidatengene könnten zum besseren Verständnis der molekularen Grundlagen der Extremitätenentwicklung beitragen.

7 ELECTRONIC DATABASE INFORMATION

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/

BLAST, http://www.ncbi.nlm.nih.gov/BLAST/

ClustalW, http://searchlauncher.bcm.tmc.edu/multi-align/multi-align.html

Mouse Genome Informatics (MGI), http://www.informatics.jax.org/

NIX, http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/

OMIM, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM

UCSC Genome Browser Gateway, http://genome.ucsc.edu/cgi-bin/hgGateway/

Whitehead Institute YAC and RH map, http://www.broad.mit.edu/cgi-bin/contig/phys_map

8 REFERENCES

Aboobaker A, Blaxter M (2003) Hox gene evolution in nematodes: novelty conserved. Curr Opin Genet Dev 13:593-8

Akarsu AN, Stoilov I, Yilmaz E, Sayli BS, Sarfarazi M (1996) Genomic structure of HOXD13 gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families. Hum Mol Genet 5:945-52

Akasaka T, Kanno M, Balling R, Mieza MA, Taniguchi M, Koseki H (1996) A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. Development 122:1513-22

Albrecht AN, Kornak U, Boddrich A, Suring K, Robinson PN, Stiege AC, Lurz R, Stricker S, Wanker EE, Mundlos S (2004) A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. Hum Mol Genet 13:2351-9

Albrecht AN, Schwabe GC, Stricker S, Boddrich A, Wanker EE, Mundlos S (2002) The synpolydactyly homolog (spdh) mutation in the mouse -- a defect in patterning and growth of limb cartilage elements. Mech Dev 112:53-67

Alkema MJ, van der Lugt NM, Bobeldijk RC, Berns A, van Lohuizen M (1995) Transformation of axial skeleton due to overexpression of bmi-1 in transgenic mice. Nature 374:724-7

Aoki T, Okada N, Wakamatsu T, Tamura TA (2002) TBP-interacting protein 120B, which is induced in relation to myogenesis, binds to NOT3. Biochem Biophys Res Commun 296:1097-103

Barna M, Hawe N, Niswander L, Pandolfi PP (2000) Plzf regulates limb and axial skeletal patterning. Nat Genet 25:166-72

Barna M, Merghoub T, Costoya JA, Ruggero D, Branford M, Bergia A, Samori B, Pandolfi PP (2002) Plzf mediates transcriptional repression of HoxD gene expression through chromatin remodeling. Dev Cell 3:499-510

Barrow JR, Thomas KR, Boussadia-Zahui O, Moore R, Kemler R, Capecchi MR, McMahon AP (2003) Ectodermal Wnt3/beta-catenin signaling is required for the establishment and maintenance of the apical ectodermal ridge. Genes Dev 17:394-409

Becker D, Jiang Z, Knodler P, Deinard AS, Eid R, Kidd KK, Shashikant CS, Ruddle FH, Schughart K (1996) Conserved regulatory element involved in the early onset of Hoxb6 gene expression. Dev Dyn 205:73-81

Berthelsen J, Zappavigna V, Ferretti E, Mavilio F, Blasi F (1998) The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. Embo J 17:1434-45

Boles RG, Pober BR, Gibson LH, Willis CR, McGrath J, Roberts DJ, Yang-Feng TL (1995) Deletion of chromosome 2q24-q31 causes characteristic digital anomalies: case report and review. Am J Med Genet 55:155-60

Boulet AM, Capecchi MR (2004) Multiple roles of Hoxa11 and Hoxd11 in the formation of the mammalian forelimb zeugopod. Development 131:299-309

Boutell JM, Thomas P, Neal JW, Weston VJ, Duce J, Harper PS, Jones AL (1999) Aberrant interactions of transcriptional repressor proteins with the Huntington's disease gene product, huntingtin. Hum Mol Genet 8:1647-55

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-54

Bugge M, Bruun-Petersen G, Brondum-Nielsen K, Friedrich U, Hansen J, Jensen G, Jensen PK, Kristoffersson U, Lundsteen C, Niebuhr E, Rasmussen KR, Rasmussen K, Tommerup N (2000) Disease associated balanced chromosome rearrangements: a resource for large scale genotype-phenotype delineation in man. J Med Genet 37:858-65

Burke AC, Nelson CE, Morgan BA, Tabin C (1995) Hox genes and the evolution of vertebrate axial morphology. Development 121:333-46

Calabrese O, Bigoni S, Gualandi F, Trabanelli C, Camera G, E. C (2000) A new mutation in HOXD13 homeodomain causes a novel human limb malformation by producing selective loss of function. Europ J Hum Genet 8 (suppl. 1):140

Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 101:2999-3004

Capdevila J, Izpisua Belmonte JC (2001) Patterning mechanisms controlling vertebrate limb development. Annu Rev Cell Dev Biol 17:87-132

Capdevila J, Tsukui T, Rodriquez Esteban C, Zappavigna V, Izpisua Belmonte JC (1999) Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. Mol Cell 4:839-49

Caronia G, Goodman FR, McKeown CM, Scambler PJ, Zappavigna V (2003) An I47L substitution in the HOXD13 homeodomain causes a novel human limb malformation by producing a selective loss of function. Development 130:1701-12

Carpenter EM, Goddard JM, Davis AP, Nguyen TP, Capecchi MR (1997) Targeted disruption of Hoxd-10 affects mouse hindlimb development. Development 124:4505-14

Carroll SB (1995) Homeotic genes and the evolution of arthropods and chordates. Nature 376:479-85

Carson JP, Eichele G, Chiu W (2005) A method for automated detection of gene expression required for the establishment of a digital transcriptome-wide gene expression atlas. J Microsc 217:275-81

Catala M (2000) [Control of the positioning of the vertebrate limb axes during development]. Morphologie 84:17-23

Chang CP, Brocchieri L, Shen WF, Largman C, Cleary ML (1996) Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. Mol Cell Biol 16:1734-45

Chang CP, Jacobs Y, Nakamura T, Jenkins NA, Copeland NG, Cleary ML (1997) Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. Mol Cell Biol 17:5679-87

Chang CP, Shen WF, Rozenfeld S, Lawrence HJ, Largman C, Cleary ML (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. Genes Dev 9:663-74

Chen Y, Knezevic V, Ervin V, Hutson R, Ward Y, Mackem S (2004) Direct interaction with Hoxd proteins reverses Gli3-repressor function to promote digit formation downstream of Shh. Development 131:2339-47

Chen Y, Zhao X (1998) Shaping limbs by apoptosis. J Exp Zool 282:691-702

Clack JA (2002) An early tetrapod from 'Romer's Gap'. Nature 418:72-6

Coates MI, Clack JA (1990) Polydactyly in the earliest known tetrapod limbs. Nature 347:66-69

Cohn MJ, Tickle C (1999) Developmental basis of limblessness and axial patterning in snakes. Nature 399:474-9

Davis AP, Capecchi MR (1994) Axial homeosis and appendicular skeleton defects in mice with a targeted disruption of hoxd-11. Development 120:2187-98

Davis AP, Capecchi MR (1996) A mutational analysis of the 5' HoxD genes: dissection of genetic interactions during limb development in the mouse. Development 122:1175-85

Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR (1995) Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature 375:791-5

Dawid IB, Breen JJ, Toyama R (1998) LIM domains: multiple roles as adapters and functional modifiers in protein interactions. Trends Genet 14:156-62

de Graaff W, Tomotsune D, Oosterveen T, Takihara Y, Koseki H, Deschamps J (2003) Randomly inserted and targeted Hox/reporter fusions transcriptionally silenced in Polycomb mutants. Proc Natl Acad Sci U S A 100:13362-7

de la Cruz CC, Der-Avakian A, Spyropoulos DD, Tieu DD, Carpenter EM (1999) Targeted disruption of Hoxd9 and Hoxd10 alters locomotor behavior, vertebral identity, and peripheral nervous system development. Dev Biol 216:595-610

Debeer P, Bacchelli C, Scambler PJ, De Smet L, Fryns JP, Goodman FR (2002) Severe digital abnormalities in a patient heterozygous for both a novel missense mutation in HOXD13 and a polyalanine tract expansion in HOXA13. J Med Genet 39:852-6

Del Campo M, Jones MC, Veraksa AN, Curry CJ, Jones KL, Mascarello JT, Ali-Kahn-Catts Z, Drumheller T, McGinnis W (1999) Monodactylous limbs and abnormal genitalia are associated with hemizygosity for the human 2q31 region that includes the HOXD cluster. Am J Hum Genet 65:104-10

Deschamps J (2004) Developmental biology. Hox genes in the limb: a play in two acts. Science 304:1610-1

Deschamps J, van den Akker E, Forlani S, De Graaff W, Oosterveen T, Roelen B, Roelfsema J (1999) Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. Int J Dev Biol 43:635-50

Di Paola J, Goldman T, Qian Q, Patil SR, Schute BC (2004) Breakpoint of a balanced translocation (X:14) (q27.1;q32.3) in a girl with severe hemophilia B maps proximal to the factor IX gene. J Thromb Haemost 2:437-40

Dlugaszewska B, Silahtaroglu A, Menzel C, Kubart S, Cohen M, Mundlos S, Tumer Z, Kjaer K, Friedrich U, Ropers HH, Tommerup N, Neitzel H, Kalscheuer VM (2005)

Breakpoints around the HOXD cluster result in various limb malformations. J Med Genet

Dolle P, Dierich A, LeMeur M, Schimmang T, Schuhbaur B, Chambon P, Duboule D (1993) Disruption of the Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs. Cell 75:431-41

Drossopoulou G, Lewis KE, Sanz-Ezquerro JJ, Nikbakht N, McMahon AP, Hofmann C, Tickle C (2000) A model for anteroposterior patterning of the vertebrate limb based on sequential long- and short-range Shh signalling and Bmp signalling. Development 127:1337-48

Duboule D (1994) How to make a limb? Science 266:575-6

Dudley AT, Ros MA, Tabin CJ (2002) A re-examination of proximodistal patterning during vertebrate limb development. Nature 418:539-44

Eid R, Koseki H, Schughart K (1993) Analysis of LacZ reporter genes in transgenic embryos suggests the presence of several cis-acting regulatory elements in the murine Hoxb-6 gene. Dev Dyn 196:205-16

Fan J, Otterlei M, Wong HK, Tomkinson AE, Wilson DM, 3rd (2004) XRCC1 co-localizes and physically interacts with PCNA. Nucleic Acids Res 32:2193-201

Fantes J, Redeker B, Breen M, Boyle S, Brown J, Fletcher J, Jones S, Bickmore W, Fukushima Y, Mannens M (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. Hum Mol Genet 4:415-22

Favier B, Le Meur M, Chambon P, Dolle P (1995) Axial skeleton homeosis and forelimb malformations in Hoxd-11 mutant mice. Proc Natl Acad Sci U S A 92:310-4

Favier B, Rijli FM, Fromental-Ramain C, Fraulob V, Chambon P, Dolle P (1996) Functional cooperation between the non-paralogous genes Hoxa-10 and Hoxd-11 in the developing forelimb and axial skeleton. Development 122:449-60

Ferretti E, Schulz H, Talarico D, Blasi F, Berthelsen J (1999) The PBX-regulating protein PREP1 is present in different PBX-complexed forms in mouse. Mech Dev 83:53-64

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 Res 30:2043-51

Fromental-Ramain C, Warot X, Lakkaraju S, Favier B, Haack H, Birling C, Dierich A, Doll e P, Chambon P (1996) Specific and redundant functions of the paralogous Hoxa-9 and Hoxd-9 genes in forelimb and axial skeleton patterning. Development 122:461-72

Gaunt SJ (2000) Evolutionary shifts of vertebrate structures and Hox expression up and down the axial series of segments: a consideration of possible mechanisms. Int J Dev Biol 44:109-17

Gehring WJ, Affolter M, Burglin T (1994) Homeodomain proteins. Annu Rev Biochem 63:487-526

Gerard M, Duboule D, Zakany J (1993) Structure and activity of regulatory elements involved in the activation of the Hoxd-11 gene during late gastrulation. Embo J 12:3539-50

Glassner BJ, Weeda G, Allan JM, Broekhof JL, Carls NH, Donker I, Engelward BP, Hampson RJ, Hersmus R, Hickman MJ, Roth RB, Warren HB, Wu MM, Hoeijmakers JH, Samson LD (1999) DNA repair methyltransferase (Mgmt) knockout mice are sensitive to the lethal effects of chemotherapeutic alkylating agents. Mutagenesis 14:339-47

Goodman F, Giovannucci-Uzielli ML, Hall C, Reardon W, Winter R, Scambler P (1998) Deletions in HOXD13 segregate with an identical, novel foot malformation in two unrelated families. Am J Hum Genet 63:992-1000

Goodman FR (2002) Limb malformations and the human HOX genes. Am J Med Genet 112:256-65

Goodman FR, Bacchelli C, Brady AF, Brueton LA, Fryns JP, Mortlock DP, Innis JW, Holmes LB, Donnenfeld AE, Feingold M, Beemer FA, Hennekam RC, Scambler PJ (2000) Novel HOXA13 mutations and the phenotypic spectrum of hand-foot-genital syndrome. Am J Hum Genet 67:197-202

Goodman FR, Majewski F, Collins AL, Scambler PJ (2002) A 117-kb microdeletion removing HOXD9-HOXD13 and EVX2 causes synpolydactyly. Am J Hum Genet 70:547-55

Goodman FR, Mundlos S, Muragaki Y, Donnai D, Giovannucci-Uzielli ML, Lapi E, Majewski F, McGaughran J, McKeown C, Reardon W, Upton J, Winter RM, Olsen BR, Scambler PJ (1997) Synpolydactyly phenotypes correlate with size of expansions in HOXD13 polyalanine tract. Proc Natl Acad Sci U S A 94:7458-63

Gould A, Itasaki N, Krumlauf R (1998) Initiation of rhombomeric Hoxb4 expression requires induction by somites and a retinoid pathway. Neuron 21:39-51

Gould A, Morrison A, Sproat G, White RA, Krumlauf R (1997) Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. Genes Dev 11:900-13

Grafstrom RC, Pegg AE, Trump BF, Harris CC (1984) O6-alkylguanine-DNA alkyltransferase activity in normal human tissues and cells. Cancer Res 44:2855-7

Griffin C, Kleinjan DA, Doe B, van Heyningen V (2002) New 3' elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. Mech Dev 112:89-100

Guha U, Gomes WA, Kobayashi T, Pestell RG, Kessler JA (2002) In vivo evidence that BMP signaling is necessary for apoptosis in the mouse limb. Dev Biol 249:108-20

Herault Y, Beckers J, Gerard M, Duboule D (1999) Hox gene expression in limbs: colinearity by opposite regulatory controls. Dev Biol 208:157-65

Herault Y, Fraudeau N, Zakany J, Duboule D (1997) Ulnaless (Ul), a regulatory mutation inducing both loss-of-function and gain-of-function of posterior Hoxd genes. Development 124:3493-500

Hill A, Wagner A, Hill M (2003) Hox and paraHox genes from the anthozoan Parazoanthus parasiticus. Mol Phylogenet Evol 28:529-35

Innis JW, Goodman FR, Bacchelli C, Williams TM, Mortlock DP, Sateesh P, Scambler PJ, McKinnon W, Guttmacher AE (2002) A HOXA13 allele with a missense mutation in the homeobox and a dinucleotide deletion in the promoter underlies Guttmacher syndrome. Hum Mutat 19:573-4

Izpisua-Belmonte JC, Duboule D (1992) Homeobox genes and pattern formation in the vertebrate limb. Dev Biol 152:26-36

Jiang G, Yang F, van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M (2003) Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet 12:2909-21

Johnson RL, Riddle RD, Laufer E, Tabin C (1994) Sonic hedgehog: a key mediator of anterior-posterior patterning of the limb and dorso-ventral patterning of axial embryonic structures. Biochem Soc Trans 22:569-74

Johnson RL, Tabin CJ (1997) Molecular models for vertebrate limb development. Cell 90:979-90

Kan SH, Johnson D, Giele H, Wilkie AO (2003) An acceptor splice site mutation in HOXD13 results in variable hand, but consistent foot malformations. Am J Med Genet 121A:69-74

Kjaer KW, Hedeboe J, Bugge M, Hansen C, Friis-Henriksen K, Vestergaard MB, Tommerup N, Opitz JM (2002) HOXD13 polyalanine tract expansion in classical synpolydactyly type Vordingborg. Am J Med Genet 110:116-21

Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, van Heyningen V (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. Hum Mol Genet 10:2049-59

Kleinjan DA, van Heyningen V (2004) Long-Range Control of Gene Expression: Emerging Mechanisms and Disruption in Disease. Am J Hum Genet 76

Kleinjan DJ, van Heyningen V (1998) Position effect in human genetic disease. Hum Mol Genet 7:1611-8

Kmita M, Duboule D (2003) Organizing axes in time and space; 25 years of colinear tinkering. Science 301:331-3

Kmita M, Fraudeau N, Herault Y, Duboule D (2002a) Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. Nature 420:145-50

Kmita M, Tarchini B, Duboule D, Herault Y (2002b) Evolutionary conserved sequences are required for the insulation of the vertebrate Hoxd complex in neural cells. Development 129:5521-8

Knezevic V, De Santo R, Schughart K, Huffstadt U, Chiang C, Mahon KA, Mackem S (1997) Hoxd-12 differentially affects preaxial and postaxial chondrogenic branches in the limb and regulates Sonic hedgehog in a positive feedback loop. Development 124:4523-36

Knittel T, Kessel M, Kim MH, Gruss P (1995) A conserved enhancer of the human and murine Hoxa-7 gene specifies the anterior boundary of expression during embryonal development. Development 121:1077-88

Knoepfler PS, Calvo KR, Chen H, Antonarakis SE, Kamps MP (1997) Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1. Proc Natl Acad Sci U S A 94:14553-8

Knoepfler PS, Kamps MP (1995) The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. Mol Cell Biol 15:5811-9

Knoepfler PS, Kamps MP (1997) The highest affinity DNA element bound by Pbx complexes in t(1;19) leukemic cells fails to mediate cooperative DNA-binding or cooperative transactivation by E2a-Pbx1 and class I Hox proteins - evidence for selective targetting of E2a-Pbx1 to a subset of Pbx-recognition elements. Oncogene 14:2521-31

Knoepfler PS, Lu Q, Kamps MP (1996) Pbx-1 Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3' to a TAAT motif. Nucleic Acids Res 24:2288-94

Knosp WM, Scott V, Bachinger HP, Stadler HS (2004) HOXA13 regulates the expression of bone morphogenetic proteins 2 and 7 to control distal limb morphogenesis. Development 131:4581-92

Kondo T, Dolle P, Zakany J, Duboule D (1996) Function of posterior HoxD genes in the morphogenesis of the anal sphincter. Development 122:2651-9

Kondo T, Duboule D (1999) Breaking colinearity in the mouse HoxD complex. Cell 97:407-17

Kondo T, Zakany J, Duboule D (1998) Control of colinearity in AbdB genes of the mouse HoxD complex. Mol Cell 1:289-300

Kornberg RD, Lorch Y (1992) Chromatin structure and transcription. Annu Rev Cell Biol 8:563-87

Kraus P, Fraidenraich D, Loomis CA (2001) Some distal limb structures develop in mice lacking Sonic hedgehog signaling. Mech Dev 100:45-58

Krokan H, Haugen A, Myrnes B, Guddal PH (1983) Repair of premutagenic DNA lesions in human fetal tissues: evidence for low levels of O6-methylguanine-DNA methyltransferase and uracil-DNA glycosylase activity in some tissues. Carcinogenesis 4:1559-64

Kwan CT, Tsang SL, Krumlauf R, Sham MH (2001) Regulatory analysis of the mouse Hoxb3 gene: multiple elements work in concert to direct temporal and spatial patterns of expression. Dev Biol 232:176-90

Lauderdale JD, Wilensky JS, Oliver ER, Walton DS, Glaser T (2000) 3' deletions cause aniridia by preventing PAX6 gene expression. Proc Natl Acad Sci U S A 97:13755-9

Laufer E, Nelson CE, Johnson RL, Morgan BA, Tabin C (1994) Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. Cell 79:993-1003

Lawrence HJ, Fischbach NA, Largman C (2005) HOX genes: not just myeloid oncogenes any more. Leukemia 19:1328-30

Lebedev O (1997) Fins made for walking. Nature 390:21-22

Lehoczky JA, Williams ME, Innis JW (2004) Conserved expression domains for genes upstream and within the HoxA and HoxD clusters suggests a long-range enhancer existed before cluster duplication. Evol Dev 6:423-30

Lewandoski M, Sun X, Martin GR (2000) Fgf8 signalling from the AER is essential for normal limb development. Nat Genet 26:460-3

Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, Frazer KA (2000) Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 288:136-40

Lower KM, Kumar R, Woollatt E, Villard L, Gecz J, Sutherland GR, Callen DF (2004) Partial Androgen Insensitivity Syndrome and t(X;5): Are There Upstream Regulatory Elements of the Androgen Receptor Gene? Horm Res 62:208-214

Lu Q, Kamps MP (1997) Heterodimerization of Hox proteins with Pbx1 and oncoprotein E2a-Pbx1 generates unique DNA-binding specifities at nucleotides predicted to contact the N-terminal arm of the Hox homeodomain--demonstration of Hox-dependent targeting of E2a-Pbx1 in vivo. Oncogene 14:75-83

Lu Q, Knoepfler PS, Scheele J, Wright DD, Kamps MP (1995) Both Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine Hox genes, some of which are themselves oncogenes. Mol Cell Biol 15:3786-95

Lux A, Beil C, Majety M, Barron S, Gallione CJ, Kuhn HM, Berg JN, Kioschis P, Marchuk DA, Hafner M (2005) Human retroviral gag- and gag-pol-like proteins interact with the transforming growth factor-beta receptor activin receptor-like kinase 1. J Biol Chem 280:8482-93

Mackem S, Knezevic V (1999) Do 5'Hoxd genes play a role in initiating or maintaining A-P polarizing signals in the limb? Cell Tissue Res 296:27-31

Maconochie M, Nonchev S, Morrison A, Krumlauf R (1996) Paralogous Hox genes: function and regulation. Annu Rev Genet 30:529-56

Maconochie MK, Nonchev S, Studer M, Chan SK, Popperl H, Sham MH, Mann RS, Krumlauf R (1997) Cross-regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. Genes Dev 11:1885-95

Manktelow E, Shigemoto K, Brierley I (2005) Characterization of the frameshift signal of Edr, a mammalian example of programmed -1 ribosomal frameshifting. Nucleic Acids Res 33:1553-63

Mansfield JH, Harfe BD, Nissen R, Obenauer J, Srineel J, Chaudhuri A, Farzan-Kashani R, Zuker M, Pasquinelli AE, Ruvkun G, Sharp PA, Tabin CJ, McManus MT (2004) MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet

Manzanares M, Bel-Vialar S, Ariza-McNaughton L, Ferretti E, Marshall H, Maconochie MM, Blasi F, Krumlauf R (2001) Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. Development 128:3595-607

Manzanares M, Cordes S, Ariza-McNaughton L, Sadl V, Maruthainar K, Barsh G, Krumlauf R (1999) Conserved and distinct roles of kreisler in regulation of the paralogous Hoxa3 and Hoxb3 genes. Development 126:759-69

Manzanares M, Cordes S, Kwan CT, Sham MH, Barsh GS, Krumlauf R (1997) Segmental regulation of Hoxb-3 by kreisler. Nature 387:191-5

Manzanares M, Nardelli J, Gilardi-Hebenstreit P, Marshall H, Giudicelli F, Martinez-Pastor MT, Krumlauf R, Charnay P (2002) Krox20 and kreisler co-operate in the transcriptional control of segmental expression of Hoxb3 in the developing hindbrain. Embo J 21:365-76

Marlin S, Blanchard S, Slim R, Lacombe D, Denoyelle F, Alessandri JL, Calzolari E, Drouin-Garraud V, Ferraz FG, Fourmaintraux A, Philip N, Toublanc JE, Petit C (1999) Townes-Brocks syndrome: detection of a SALL1 mutation hot spot and evidence for a position effect in one patient. Hum Mutat 14:377-86

Martinez P, Rast JP, Arenas-Mena C, Davidson EH (1999) Organization of an echinoderm Hox gene cluster. Proc Natl Acad Sci U S A 96:1469-74

Masuda Y, Sasaki A, Shibuya H, Ueno N, Ikeda K, Watanabe K (2001) Dlxin-1, a novel protein that binds Dlx5 and regulates its transcriptional function. J Biol Chem 276:5331-8

Matsuda T, Suzuki H, Oishi I, Kani S, Kuroda Y, Komori T, Sasaki A, Watanabe K, Minami Y (2003) The receptor tyrosine kinase Ror2 associates with the melanoma-associated antigen (MAGE) family protein Dlxin-1 and regulates its intracellular distribution. J Biol Chem 278:29057-64

Mercader N, Leonardo E, Azpiazu N, Serrano A, Morata G, Martinez C, Torres M (1999) Conserved regulation of proximodistal limb axis development by Meis1/Hth. Nature 402:425-9

Mercader N, Leonardo E, Piedra ME, Martinez AC, Ros MA, Torres M (2000) Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. Development 127:3961-70

Morrison A, Ariza-McNaughton L, Gould A, Featherstone M, Krumlauf R (1997) HOXD4 and regulation of the group 4 paralog genes. Development 124:3135-46

Morrison A, Moroni MC, Ariza-McNaughton L, Krumlauf R, Mavilio F (1996) In vitro and transgenic analysis of a human HOXD4 retinoid-responsive enhancer. Development 122:1895-907

Mortlock DP, Innis JW (1997) Mutation of HOXA13 in hand-foot-genital syndrome. Nat Genet 15:179-80

Muncke N, Wogatzky BS, Breuning M, Sistermans EA, Endris V, Ross M, Vetrie D, Catsman-Berrevoets CE, Rappold G (2004) Position effect on PLP1 may cause a subset of Pelizaeus-Merzbacher disease symptoms. J Med Genet 41:e121

Muragaki Y, Mundlos S, Upton J, Olsen BR (1996) Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. Science 272:548-51

Nakamura M, Watanabe T, Yonekawa Y, Kleihues P, Ohgaki H (2001) Promoter methylation of the DNA repair gene MGMT in astrocytomas is frequently associated with G:C --> A:T mutations of the TP53 tumor suppressor gene. Carcinogenesis 22:1715-9

Nakatsu Y, Hattori K, Hayakawa H, Shimizu K, Sekiguchi M (1993) Organization and expression of the human gene for O6-methylguanine-DNA methyltransferase. Mutat Res 293:119-32

Nelson CE, Morgan BA, Burke AC, Laufer E, DiMambro E, Murtaugh LC, Gonzales E, Tessarollo L, Parada LF, Tabin C (1996) Analysis of Hox gene expression in the chick limb bud. Development 122:1449-66

Niswander L, Jeffrey S, Martin GR, Tickle C (1994) A positive feedback loop coordinates growth and patterning in the vertebrate limb. Nature 371:609-12

Niswander L, Martin GR (1992) Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. Development 114:755-68

Niswander L, Tickle C, Vogel A, Booth I, Martin GR (1993) FGF-4 replaces the apical ectodermal ridge and directs outgrowth and patterning of the limb. Cell 75:579-87

Nixon J, Oldridge M, Wilkie AO, Smith K (1997) Interstitial deletion of 2q associated with craniosynostosis, ocular coloboma, and limb abnormalities: cytogenetic and molecular investigation. Am J Med Genet 70:324-7

Nourani A, Utley RT, Allard S, Cote J (2004) Recruitment of the NuA4 complex poises the PHO5 promoter for chromatin remodeling and activation. Embo J 23:2597-607

Okabe H, Satoh S, Furukawa Y, Kato T, Hasegawa S, Nakajima Y, Yamaoka Y, Nakamura Y (2003) Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. Cancer Res 63:3043-8

Ono R, Kobayashi S, Wagatsuma H, Aisaka K, Kohda T, Kaneko-Ishino T, Ishino F (2001) A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. Genomics 73:232-7

Ono R, Shiura H, Aburatani H, Kohda T, Kaneko-Ishino T, Ishino F (2003) Identification of a large novel imprinted gene cluster on mouse proximal chromosome 6. Genome Res 13:1696-705

Oosterveen T, Niederreither K, Dolle P, Chambon P, Meijlink F, Deschamps J (2003) Retinoids regulate the anterior expression boundaries of 5' Hoxb genes in posterior hindbrain. Embo J 22:262-9

Oue N, Shigeishi H, Kuniyasu H, Yokozaki H, Kuraoka K, Ito R, Yasui W (2001) Promoter hypermethylation of MGMT is associated with protein loss in gastric carcinoma. Int J Cancer 93:805-9

Packer AI, Crotty DA, Elwell VA, Wolgemuth DJ (1998) Expression of the murine Hoxa4 gene requires both autoregulation and a conserved retinoic acid response element. Development 125:1991-8

Panman L, Zeller R (2003) Patterning the limb before and after SHH signalling. J Anat 202:3-12

Pegg AE, Dolan ME, Scicchitano D, Morimoto K (1985) Studies of the repair of O6-alkylguanine and O4-alkylthymine in DNA by alkyltransferases from mammalian cells and bacteria. Environ Health Perspect 62:109-14

Peichel CL, Prabhakaran B, Vogt TF (1997) The mouse Ulnaless mutation deregulates posterior HoxD gene expression and alters appendicular patterning. Development 124:3481-92

Pellerin I, Schnabel C, Catron KM, Abate C (1994) Hox proteins have different affinities for a consensus DNA site that correlate with the positions of their genes on the hox cluster. Mol Cell Biol 14:4532-45

Popperl H, Featherstone MS (1992) An autoregulatory element of the murine Hox-4.2 gene. Embo J 11:3673-80

Rancourt DE, Tsuzuki T, Capecchi MR (1995) Genetic interaction between hoxb-5 and hoxb-6 is revealed by nonallelic noncomplementation. Genes Dev 9:108-22

Reese JS, Allay E, Gerson SL (2001) Overexpression of human O6-alkylguanine DNA alkyltransferase (AGT) prevents MNU induced lymphomas in heterozygous p53 deficient mice. Oncogene 20:5258-63

Roelen BA, de Graaff W, Forlani S, Deschamps J (2002) Hox cluster polarity in early transcriptional availability: a high order regulatory level of clustered Hox genes in the mouse. Mech Dev 119:81-90

Rosso S, Bollati F, Bisbal M, Peretti D, Sumi T, Nakamura T, Quiroga S, Ferreira A, Caceres A (2004) LIMK1 regulates Golgi dynamics, traffic of Golgi-derived vesicles, and process extension in primary cultured neurons. Mol Biol Cell 15:3433-49

Rowe DA, Fallon JF (1982) The proximodistal determination of skeletal parts in the developing chick leg. J Embryol Exp Morphol 68:1-7

Sakumi K, Shiraishi A, Shimizu S, Tsuzuki T, Ishikawa T, Sekiguchi M (1997) Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice. Cancer Res 57:2415-8

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning : A Laboratory Manual. Cold Spring Harbor Laboratory Press

Saunders JW, Jr. (1998) The proximo-distal sequence of origin of the parts of the chick wing and the role of the ectoderm. 1948. J Exp Zool 282:628-68

Schumacher A, Magnuson T (1997) Murine Polycomb- and trithorax-group genes regulate homeotic pathways and beyond. Trends Genet 13:167-70

Sham MH, Vesque C, Nonchev S, Marshall H, Frain M, Gupta RD, Whiting J, Wilkinson D, Charnay P, Krumlauf R (1993) The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. Cell 72:183-96

Shanmugam K, Green NC, Rambaldi I, Saragovi HU, Featherstone MS (1999) PBX and MEIS as non-DNA-binding partners in trimeric complexes with HOX proteins. Mol Cell Biol 19:7577-88

Sharp TV, Munoz F, Bourboulia D, Presneau N, Darai E, Wang HW, Cannon M, Butcher DN, Nicholson AG, Klein G, Imreh S, Boshoff C (2004) LIM domains-containing protein 1 (LIMD1), a tumor suppressor encoded at chromosome 3p21.3, binds pRB and represses E2F-driven transcription. Proc Natl Acad Sci U S A 101:16531-6

Sharpe J, Nonchev S, Gould A, Whiting J, Krumlauf R (1998) Selectivity, sharing and competitive interactions in the regulation of Hoxb genes. Embo J 17:1788-98

Shashikant CS, Bieberich CJ, Belting HG, Wang JC, Borbely MA, Ruddle FH (1995) Regulation of Hoxc-8 during mouse embryonic development: identification and characterization of critical elements involved in early neural tube expression. Development 121:4339-47

Shen WF, Chang CP, Rozenfeld S, Sauvageau G, Humphries RK, Lu M, Lawrence HJ, Cleary ML, Largman C (1996) Hox homeodomain proteins exhibit selective complex stabilities with Pbx and DNA. Nucleic Acids Res 24:898-906

Shen WF, Montgomery JC, Rozenfeld S, Moskow JJ, Lawrence HJ, Buchberg AM, Largman C (1997a) AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. Mol Cell Biol 17:6448-58

Shen WF, Rozenfeld S, Lawrence HJ, Largman C (1997b) The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. J Biol Chem 272:8198-206

Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PW, Reith AD (2001) Identification and characterisation of a developmentally regulated mammalian gene that utilises -1 programmed ribosomal frameshifting. Nucleic Acids Res 29:4079-88

Shrimpton AE, Levinsohn EM, Yozawitz JM, Packard DS, Jr., Cady RB, Middleton FA, Persico AM, Hootnick DR (2004) A HOX gene mutation in a family with isolated congenital vertical talus and Charcot-Marie-Tooth disease. Am J Hum Genet 75:92-6

Shubin N, Tabin C, Carroll S (1997) Fossils, genes and the evolution of animal limbs. Nature 388:639-48

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995) An improved PCR method for walking in uncloned genomic DNA Nucleic Acids Res. Vol. 23, pp 1087-8

Simon J (1995) Locking in stable states of gene expression: transcriptional control during Drosophila development. Curr Opin Cell Biol 7:376-85

Slavotinek A, Schwarz C, Getty JF, Stecko O, Goodman F, Kingston H (1999) Two cases with interstitial deletions of chromosome 2 and sex reversal in one. Am J Med Genet 86:75-81

Small KM, Potter SS (1993) Homeotic transformations and limb defects in Hox A11 mutant mice. Genes Dev 7:2318-28

Smith-Sorensen B, Lind GE, Skotheim RI, Fossa SD, Fodstad O, Stenwig AE, Jakobsen KS, Lothe RA (2002) Frequent promoter hypermethylation of the O6-Methylguanine-DNA Methyltransferase (MGMT) gene in testicular cancer. Oncogene 21:8878-84

Sordino P, Duboule D (1996) A molecular approach to the evolution of vertebrate paired appendages. TREE 11:114-119

Sordino P, van der Hoeven F, Duboule D (1995) Hox gene expression in teleost fins and the origin of vertebrate digits. Nature 375:678-81

Spitz F, Gonzalez F, Duboule D (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell 113:405-17

Spitz F, Gonzalez F, Peichel C, Vogt TF, Duboule D, Zakany J (2001) Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. Genes Dev 15:2209-14

Spitz F, Montavon T, Monso-Hinard C, Morris M, Ventruto ML, Antonarakis S, Ventruto V, Duboule D (2002) A t(2;8) balanced translocation with breakpoints near the human HOXD complex causes mesomelic dysplasia and vertebral defects. Genomics 79:493-8

Stanyon CA, Bernard O (1999) LIM-kinase1. Int J Biochem Cell Biol 31:389-94

Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu YZ, Gohler H, Wanker EE, Bates GP, Housman DE, Thompson LM (2000) The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc Natl Acad Sci U S A 97:6763-8

Steplewski A, Krynska B, Tretiakova A, Haas S, Khalili K, Amini S (1998) MyEF-3, a developmentally controlled brain-derived nuclear protein which specifically interacts with myelin basic protein proximal regulatory sequences. Biochem Biophys Res Commun 243:295-301

Summerbell D (1974) A quantitative analysis of the effect of excision of the AER from the chick limb-bud. J Embryol Exp Morphol 32:651-60

Summerbell D, Lewis JH, Wolpert L (1973) Positional information in chick limb morphogenesis. Nature 244:492-6

Sun X, Mariani FV, Martin GR (2002) Functions of FGF signalling from the apical ectodermal ridge in limb development. Nature 418:501-8

Suzuki M, Ueno N, Kuroiwa A (2003) Hox proteins functionally cooperate with the GC box-binding protein system through distinct domains. J Biol Chem 278:30148-56

Tadin-Strapps M, Warburton D, Baumeister FA, Fischer SG, Yonan J, Gilliam TC, Christiano AM (2004) Cloning of the breakpoints of a de novo inversion of chromosome 8, inv (8)(p11.2q23.1) in a patient with Ambras syndrome. Cytogenet Genome Res 107:68-76

Taniguchi T, Garcia-Higuera I, Andreassen PR, Gregory RC, Grompe M, D'Andrea AD (2002) S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. Blood 100:2414-20

te Welscher P, Fernandez-Teran M, Ros MA, Zeller R (2002a) Mutual genetic antagonism involving GLI3 and dHAND prepatterns the vertebrate limb bud mesenchyme prior to SHH signaling. Genes Dev 16:421-6

te Welscher P, Zuniga A, Kuijper S, Drenth T, Goedemans HJ, Meijlink F, Zeller R (2002b) Progression of vertebrate limb development through SHH-mediated counteraction of GLI3. Science 298:827-30

Thompson AA, Nguyen LT (2000) Amegakaryocytic thrombocytopenia and radio-ulnar synostosis are associated with HOXA11 mutation. Nat Genet 26:397-8

Tickle C, Munsterberg A (2001) Vertebrate limb development--the early stages in chick and mouse. Curr Opin Genet Dev 11:476-81

Tsou AP, Chuang YC, Su JY, Yang CW, Liao YL, Liu WK, Chiu JH, Chou CK (2003) Overexpression of a novel imprinted gene, PEG10, in human hepatocellular carcinoma and in regenerating mouse livers. J Biomed Sci 10:625-35

Tsuzuki T, Sakumi K, Shiraishi A, Kawate H, Igarashi H, Iwakuma T, Tominaga Y, Zhang S, Shimizu S, Ishikawa T, et al. (1996) Targeted disruption of the DNA repair methyltransferase gene renders mice hypersensitive to alkylating agent. Carcinogenesis 17:1215-20

van der Hoeven F, Zakany J, Duboule D (1996) Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. Cell 85:1025-35

Wagner GP, Amemiya C, Ruddle F (2003) Hox cluster duplications and the opportunity for evolutionary novelties. Proc Natl Acad Sci U S A 100:14603-6

Wahba GM, Hostikka SL, Carpenter EM (2001) The paralogous Hox genes Hoxa10 and Hoxd10 interact to pattern the mouse hindlimb peripheral nervous system and skeleton. Dev Biol 231:87-102

Whiting J, Marshall H, Cook M, Krumlauf R, Rigby PW, Stott D, Allemann RK (1991) Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes Dev 5:2048-59

Wiestler O, Kleihues P, Pegg AE (1984) O6-alkylguanine-DNA alkyltransferase activity in human brain and brain tumors. Carcinogenesis 5:121-4

Wirth J, Nothwang HG, van der Maarel S, Menzel C, Borck G, Lopez-Pajares I, Brondum-Nielsen K, Tommerup N, Bugge M, Ropers HH, Haaf T (1999) Systematic characterisation of disease associated balanced chromosome rearrangements by FISH: cytogenetically and genetically anchored YACs identify microdeletions and candidate regions for mental retardation genes. J Med Genet 36:271-8

Yang N, Mizuno K (1999) Nuclear export of LIM-kinase 1, mediated by two leucine-rich nuclear-export signals within the PDZ domain. Biochem J 338 (Pt 3):793-8

Yau TO, Kwan CT, Jakt LM, Stallwood N, Cordes S, Sham MH (2002) Auto/cross-regulation of Hoxb3 expression in posterior hindbrain and spinal cord. Dev Biol 252:287-300

Yekta S, Shih IH, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. Science 304:594-6

Yin X, Warner DR, Roberts EA, Pisano MM, Greene RM (2005) Identification of novel CBP interacting proteins in embryonic orofacial tissue. Biochem Biophys Res Commun 329:1010-7

Yokouchi Y, Sakiyama J, Kameda T, Iba H, Suzuki A, Ueno N, Kuroiwa A (1996) BMP-2/-4 mediate programmed cell death in chicken limb buds. Development 122:3725-34

Zakany J, Duboule D (1999) Hox genes in digit development and evolution. Cell Tissue Res 296:19-25

Zakany J, Fromental-Ramain C, Warot X, Duboule D (1997) Regulation of number and size of digits by posterior Hox genes: a dose-dependent mechanism with potential evolutionary implications. Proc Natl Acad Sci U S A 94:13695-700

Zakany J, Kmita M, Duboule D (2004) A dual role for Hox genes in limb anterior-posterior asymmetry. Science 304:1669-72

Zeller R, Deschamps J (2002) Developmental biology: first come, first served. Nature 420:138-9

Zhang F, Nagy Kovacs E, Featherstone MS (2000) Murine hoxd4 expression in the CNS requires multiple elements including a retinoic acid response element. Mech Dev 96:79-89

Zhou ZQ, Manguino D, Kewitt K, Intano GW, McMahan CA, Herbert DC, Hanes M, Reddick R, Ikeno Y, Walter CA (2001) Spontaneous hepatocellular carcinoma is reduced in transgenic mice overexpressing human O6- methylguanine-DNA methyltransferase. Proc Natl Acad Sci U S A 98:12566-71

Zou H, Niswander L (1996) Requirement for BMP signaling in interdigital apoptosis and scale formation. Science 272:738-41

Zuniga A, Haramis AP, McMahon AP, Zeller R (1999) Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. Nature 401:598-602

9 ACKNOWLEDGEMENTS

I would first like to thank Prof. H.-H. Ropers, Prof. S. Mundlos and Dr. V. Kalscheuer for their support throughout this study. Moreover, I would like to thank Dr. Kalscheuer for her comments on this written work.

I am particularly grateful to Prof. H. Kreß for his permanent support, stimulating discussions and critical reading of this thesis.

I would like to thank Sabine Kübart for fruitful discussions and valuable suggestions, which helped in solving problems in many molecular experiments. Moreover, I thank her for moral support during hard times.

I would like to acknowledge Nicole Verhey van Wijk, a wonderful and patient teacher, for her help in all aspects of yeast and protein work and for her successful attempts to cheer me up.

I would additionally like to thank Corinna Menzel and Petra Viertel for their technical expertise in FISH experiments and Susanne Freier and Hannelore Madle for assistance in cell culture. I am grateful to Jens Ruschmann, Stella-Amrei Kunde, Gundula Leschik and Claudia Mischung for their assistance in sequence analysis, and Dietmar Vogt for assistance in plasmid preparation. I would also like to thank Andrea Albrecht for her cooperativeness and valuable comments on this manuscript and Norbert Brieske, Britta Hoffmann, Florian Witte and Christine Zwingmann for their assistance in RNA *in situ* hybridisation experiments.

I would like to acknowledge all colleagues from the Neurogenetics Group, the Disease and Development Group and all my co-workers from the Student Association at the Max Planck Institute for fruitful discussions and creating a pleasant working atmophere. My special thanks deserved Olivier Hagens, Beatriz Aranda and Joyce So for our numerous discussions during lunch breaks.

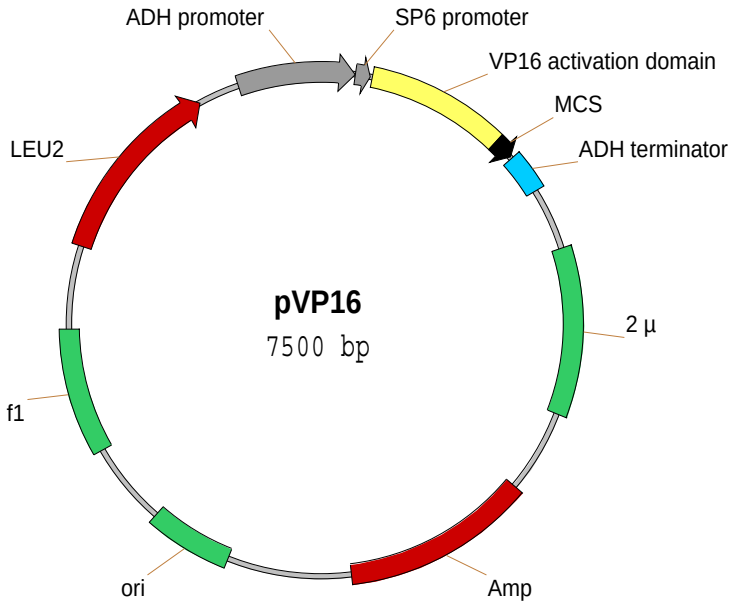
In the end I would like to thank my family and all my friends from Berlin, especially Urs Gruber, Michael Beck, Yoshimi Sugano, Bernd Hartmann, Katharina Schmidt-Brücken and Bettina Reiter for their continous moral support.

10 PUBLICATIONS

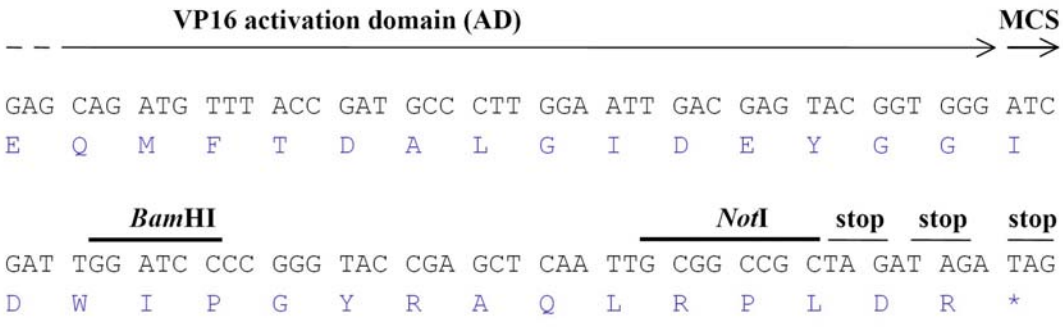
Dlugaszewska B, Silahtaroglu A, Menzel C, Kübart S, Cohen M, Mundlos S, Tümer Z, Kjaer K, Friedrich U, Ropers HH, Tommerup N, Neitzel H, Kalscheuer VM (2005) Breakpoints around the *HOXD* cluster result in various limb malformations. J Med Genet. [Epub ahead of print]

11 APPENDICES

11.1 Vector pVP16

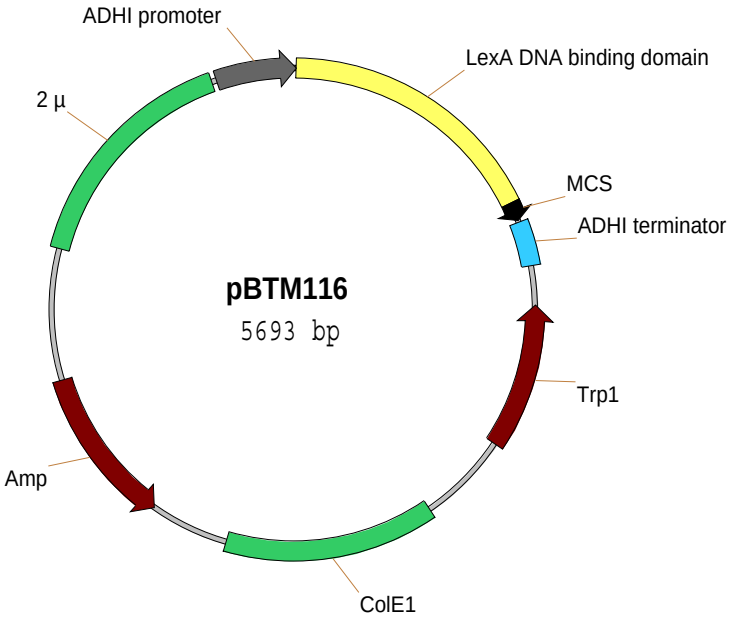


A: Schematic map of the pVP16 vector

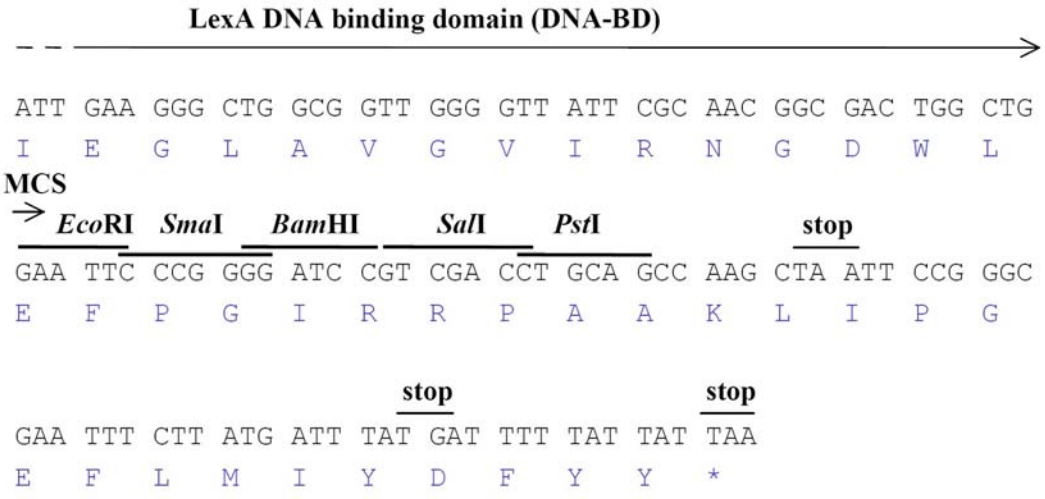


B: Multiple cloning site (MCS) of the pVP16 vector. Only unique restriction sites are marked. Nucleotide triplets are shown in frame together with the corresponding amino acids (blue letters). Stop codons are in all three frames.

11.2 Vector pBTM116

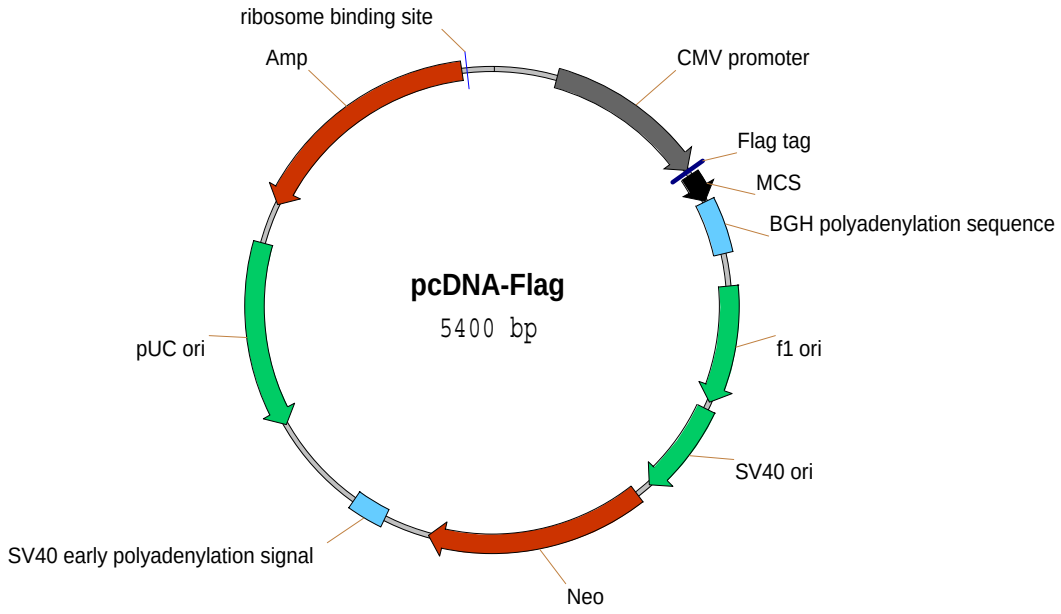


A: Schematic map of the pBTM116 vector

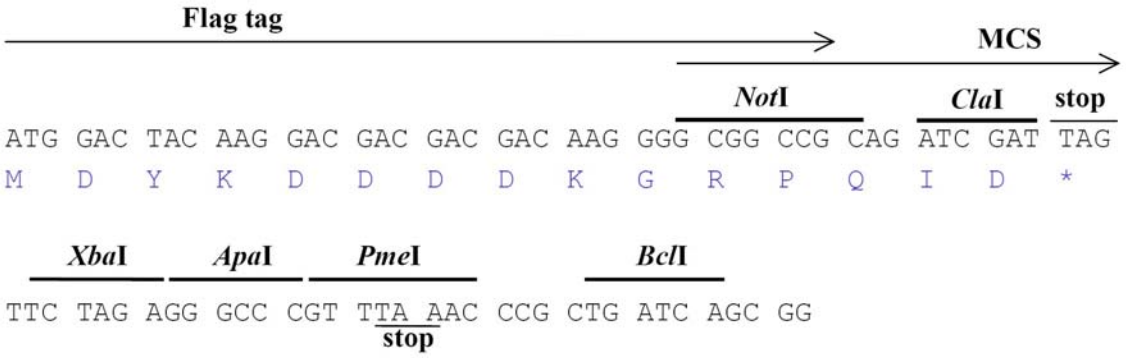


B: Multiple cloning site (MCS) of the pBTM116 vector. Nucleotide triplets are shown in frame together with the corresponding amino acids (blue letters). Stop codons are in all three frames.

11.3 Vector pcDNA-Flag

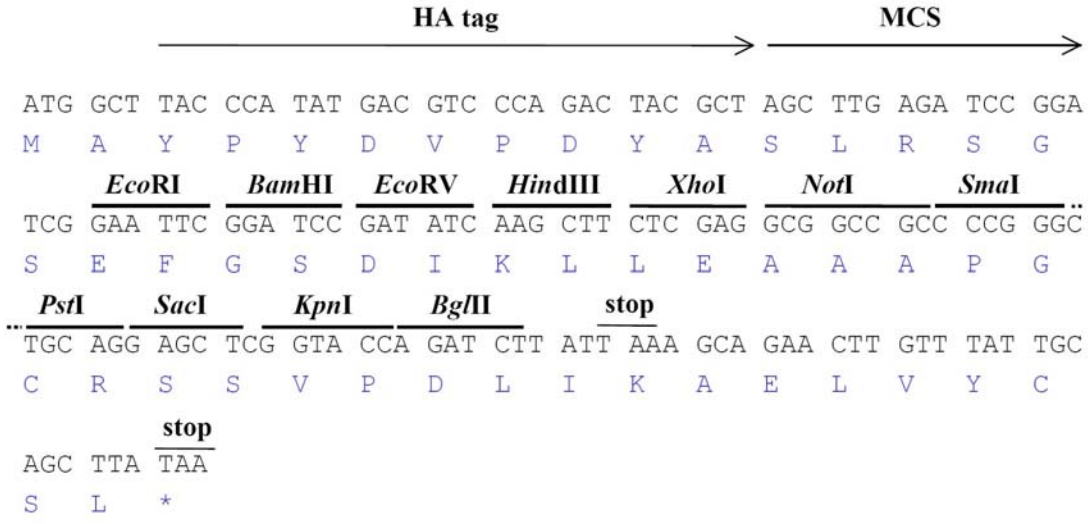


A: Schematic map of the pcDNA-Flag vector



B: Multiple cloning site (MCS) of the pcDNA-Flag vector. Nucleotide triplets are shown in frame together with the corresponding amino acids (blue letters).

11.4 Vector pTL1-HA2



A: Multiple cloning site (MCS) of the pTL1-HA2 vector. Only unique restriction sites are marked. Nucleotide triplets are shown in frame together with the corresponding amino acids (blue letters).

11.5 List of putative ESTs found on BAC RP11-538A12

EST ID (GeneBank)	Location on the BAC RP11-538A12	EST source	RT-PCR results after 70 cycles
AI075926	10898 – 11448 bp	testis	no product
AW850653	58673 – 59165 bp	colon	no product
BG980989	66265 – 66618 bp	colon	product obtained
BG980131	74910 – 75348 bp	colon	product obtained
BG979719	75540 – 75926 bp	colon	product obtained
BG979037	75530 – 75914 bp	colon	
BF815673	105008 – 105211 bp and 111435 – 111638 bp	colon	no product
AW858552	114764 – 115127 bp	colon	product obtained
AW858470	114891 – 115327 bp	colon	
AW167235	128523 – 127998 bp	uterus	no product
BE064736	131134 – 131587 bp	breast	product obtained
BE064727	131107 – 131587 bp	breast	
BE065063	131100 – 131723 bp	breast	
BE064976	131167 – 131664 bp	breast	

11.6 List of putative non-coding high homology regions (HHR) found after human-mouse sequence comparison

HHR no.	HHR location in human	Length of the HHR	Percent of identity	Expression in mouse E16.5
1.	28927 – 29114 bp	188 bp	83%	after 70 cycles
2.	29891 – 30288 bp	398 bp	90%	after 70 cycles
3.	30372 – 30603 bp	232 bp	81%	after 70 cycles
4.	30643– 30748 bp	106 bp	73%	no expression
5.	31141 – 31647 bp	507 bp	85%	after 70 cycles
6.	34556 – 34660 bp	105 bp	81%	no expression
7.	34741 – 34847 bp	107 bp	73%	no expression
8.	34954 – 35053 bp	100 bp	71%	no expression
9.	35681 – 36142 bp	462 bp	92%	after 35 cycles
10.	39320 – 39623 bp	304 bp	82%	after 70 cycles
11.	49103 – 49596 bp	494 bp	91%	after 70 cycles
12.	49684 – 49817 bp	134 bp	93%	after 70 cycles
13.	54194 – 54325 bp	132 bp	70%	after 70 cycles
14.	54693 – 54802 bp	110 bp	75%	after 70 cycles
15.	57843 – 58040 bp	198 bp	70%	after 70 cycles
16.	59793 – 60252 bp	460 bp	82%	after 70 cycles
17.	65075 – 65176 bp	102 bp	90%	after 70 cycles
18.	65190 – 65342 bp	153 bp	82%	after 70 cycles
19.	65918 – 66452 bp	535 bp	86%	after 70 cycles
20.	67322 – 67436 bp	115 bp	74%	after 70 cycles
21.	68498 – 68724 bp	227 bp	93%	after 70 cycles
22.	69476 – 69655 bp	190 bp	82%	after 70 cycles
23.	148938 – 149097 bp	160 bp	84%	after 70 cycles
24.	144670 – 144778 bp	109 bp	80%	after 70 cycles
25.	144541 – 144649 bp	109 bp	77%	no expression
26.	134716 – 143833 bp	118 bp	84%	after 70 cycles
27.	130849 – 130978 bp	130 bp	79%	after 70 cycles
28.	130226 – 130348 bp	123 bp	70%	after 70 cycles
29.	127646 – 127774 bp	129 bp	91%	after 70 cycles
30.	127366 – 127593 bp	228 bp	89%	after 70 cycles
31.	123720 – 123861 bp	142 bp	79%	no expression
32.	123186 – 123719 bp	534 bp	94%	no expression
33.	122697 – 122892 bp	196 bp	78%	after 70 cycles
34.	122291 – 122647 bp	357 bp	84%	after 70 cycles
35.	121787 – 122181 bp	395 bp	94%	after 70 cycles
36.	121468 – 121625 bp	158 bp	73%	no expression
37.	119922 – 120089 bp	168 bp	77%	no expression
38.	114460 – 114591 bp	132 bp	79%	no expression
39.	111467 – 115568 bp	102 bp	88%	no expression

Red letters indicate positions on the human BAC RP11-614N24 (GenBank accesion number AC103916), whereas blue letters correspond to the positions on the breakpoint spanning BAC RP11-538A12 (GenBank acc. no. AC016761).

11.7 Alignment of *Peg10* and *Edr* nucleotide sequences

Comparison of *Peg10* (AB091827) and *Edr* (AJ006464) nucleotide sequences suggests that both of them correspond to the same gene. Identical residues are shown as white letters on the black background. Various translation start codons (according to GenBank data) as well as different length of repetitive sequences found in both entries are marked in red. Stop codons for ORF1 and ORF2 are marked in green.

```
AJ006464     1 ----GAATTCCTGCAGCCGGCCGC--CCTGACCAACTACGACCTGGGGAGAGCAGCCAACCGAGAAGGTCCACCGAGCCT
AB091827     1 ACTTGAGTTGGTGTGTCGAAGAATCCTGACCAACTACGACCTGGGGAGAGCAGCCAACCGAGAAGGTCCACCGAGCCT

AJ006464    75 CGCCTAGGTCTGCTGCGCGGGCTGCGGTCCGGAGCCTTCTCCG-----CGGTCACCCAGTGGACCGGGCCGTCGCGGGAC
AB091827    81 CGCCTAGGTCTGCTGCGCGGGCTGCGGTCCGGAGCCTTCTCCGGACCGCGGTCACCCAGTGGACCGGGCCGTCGCGGGAC

AJ006464   150 CCCTCATCCTTCGTGGCATCGCAGAGGAATCCTCGTGTGGAACAGGCGGGTTTTAAGAACCAAAAGACGCCAACCACGAG
AB091827   161 CCCTCATCCTTCGTGGCATCGCAGAGGAATCCTCGTGTGGAACAGGCGGGTTTTAAGAACCAAAAGACGCCAACCACGAG

AJ006464   230 GGTCCCAGGATCCAGGGCTCCCTCCCCAGGGGAGTGAAGCCCCTCTCACCGCAGCCATGGCTG-TGCGG-TGGTTCCTCC
AB091827   241 GGTCCCAGGATCCAGGGCTCCCTCCCCAGGGGAGTGAAGCCCCTCTCACCGCAGCCATGGCTGCTGCGGGTGGTTCCTCC

AJ006464   308 AACTGCCCGCCCCCTCCCCCTCCCCCTCCTCCCAACAACAACAACAACAACAAG---CCAAAGAGCCCAGGCGTGCCTGA
AB091827   321 AACTGCCCGCCCCCTCCCCCTCCCCCTCCTCCCAACAACAACAACAACAACAACACCCCAAAGAGCCCAGGCGTGCCTGA

AJ006464   385 CGCCGAAGATGATGATGAACGCAGACACGATGAGCTCCCTGAAGACATCAACAACTTTGACGAAGACATGAACAGGCAGT
AB091827   401 CGCCGAAGATGATGATGAACGCAGACACGATGAGCTCCCTGAAGACATCAACAACTTTGACGAAGACATGAACAGGCAGT

AJ006464   465 TTGAGAATATGAACCTGCTGGATCAGGTGGAGTTGCTTGCACAGAGCTACAGTCTGCTGGATCATTTAGATGACTTTGAT
AB091827   481 TTGAGAATATGAACCTGCTGGATCAGGTGGAGTTGCTTGCACAGAGCTACAGTCTGCTGGATCATTTAGATGACTTTGAT

AJ006464   545 GATGATGATGAAGACGATGACTTTGATCCAGAACCTGACCAGGATGAGCTCCCTGAGTACAGTGACGATGATGACCTGGA
AB091827   561 GATGATGATGAAGACGATGACTTTGATCCAGAACCTGACCAGGATGAGCTCCCTGAGTACAGTGACGATGATGACCTGGA

AJ006464   625 GCTTCAGGGTGCTGCAGCAGCCCCTATCCCAAACTTTTTCTCCGATGATGACTGCCTTGAAGACCTTCCTGAGAAGTTCG
AB091827   641 GCTTCAGGGTGCTGCAGCAGCCCCTATCCCAAACTTTTTCTCCGATGATGACTGCCTTGAAGACCTTCCTGAGAAGTTCG

AJ006464   705 ATGGCAACCCTGACATGCTGGGTCCTTTCATGTATCAGTGCCAGCTCTTCATGGAAAAGAGCACCAGAGATTTCTCAGTT
AB091827   721 ATGGCAACCCTGACATGCTGGGTCCTTTCATGTATCAGTGCCAGCTCTTCATGGAAAAGAGCACCAGAGATTTCTCAGTT

AJ006464   785 GACCGCATCCGTGTGTGCTTCGTGACAAGCATGCTGATCGGCCGTGCCGCCCGCTGGGCTACTGCCAAGCTGCAAAGATG
AB091827   801 GACCGCATCCGTGTGTGCTTCGTGACAAGCATGCTGATCGGCCGTGCCGCCCGCTGGGCTACTGCCAAGCTGCAAAGATG

AJ006464   865 TACTTACCTGATGCAACAAACTAACACTGCCTTTATGATGGAGCTGAAGCATGTCTTTGAAGACCCTCAGAGACGTGAAG
AB091827   881 TACTTACCTGATGCA--CAACTA-CACTGCCTTTATGATGGAGCTGAAGCATGTCTTTGAAGACCCTCAGAGACGTGAAG

AJ006464   945 CTGCCAAACGCAAGATCAGACGTCTGCGCCAGGGCCCTGGGCCTGTTGTGGACTACTCCAATGCATTCCAGATGATTGCC
AB091827   958 CTGCCAAACGCAAGATCAGACGTCTGCGCCAGGGCCCTGGGCCTGTTGTGGACTACTCCAATGCATTCCAGATGATTGCC
```

```
AJ006464 1025 CAGGACCTGGATTGGACTGAGCCTGCCCTGATGGATCAGTTCCAGGAAGGTCTCAACCCAGACATTCGCGCAGAGCTGTC
AB091827 1038 CAGGACCTGGATTGGACTGAGCCTGCCCTGATGGATCAGTTCCAGGAAGGTCTCAACCCAGACATTCGCGCAGAGCTGTC

AJ006464 1105 TCGCCAGGAGGCCCCCAAGACCCTGGCTGCTCTGATTACTGCCTGTATTCACATCGAGAGAAGGCTGGCTCGTGACGCTG
AB091827 1118 TCGCCAGGAGGCCCCCAAGACCCTGGCTGCTCTGATTACTGCCTGTATTCACATCGAGAGAAGGCTGGCTCGTGACGCTG

AJ006464 1185 CTGCAAAGCCCGATCCTTCACCCAGAGCCTTGGTGATGCCTCCAAACAGCCAGACCGATCCCACCGAGCCTGTGGGAGGT
AB091827 1198 CTGCAAAGCCCGATCCTTCACCCAGAGCCTTGGTGATGCCTCCAAACAGCCAGACCGATCCCACCGAGCCTGTGGGAGGT

AJ006464 1265 GCCCGCATGCGCCTGTCCAAGGAAGAAAAGGAGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCA
AB091827 1278 GCCCGCATGCGCCTGTCCAAGGAAGAAAAGGAGAGACGCCGCAAAATGAATTTGTGTCTCTACTGTGGCAATGGAGGCCA

AJ006464 1345 TTTCGCCGACACGTGTCCAGCGAAAGCCTCCAAGAATTCGCCGCCGGGAAACTCCCCGGCCCCGCTGTAGGGGGACCTTC
AB091827 1358 TTTCGCCGACACGTGTCCAGCGAAAGCCTCCAAGAATTCGCCGCCGGGAAACTCCCCGGCCCCGCTGTAGGGGACCTTC

AJ006464 1425 AGCGACAGGGCCAGAACGAATAAGGTCCCCACCCTCCGAGGCTTCGACTCAGCACCTGCAAGTGATGCTCCAGATTCATA
AB091827 1438 AGCGACAGGGCCAGAACGAATAAGGTCCCCACCCTCCGAGGCTTCGACTCAGCACCTGCAAGTGATGCTCCAGATTCATA

AJ006464 1505 TGCCGGGCAGACCCACCCTGTTTGTCCGAGCTATGATTGATTCTGGTGCATCTGGCAACTTCATTGATCAAGACTTTGTC
AB091827 1518 TGCCGGGCAGACCCACCCTGTTTGTCCGAGCTATGATTGATTCTGGTGCATCTGGCAACTTCATTGATCAAGACTTTGTC

AJ006464 1585 ATACAAAATGCAATTCCTCTCAGAATCAAAGACTGGCCAGTGATGGTGGAAGCTATTGATGGGCATCCAATTGCCTCGGG
AB091827 1598 ATACAAAATGCAATTCCTCTCAGAATCAAAGACTGGCCAGTGATGGTGGAAGCTATTGATGGGCATCCAATTGCCTCGGG

AJ006464 1665 CCCAATCATTTTGGAAACCCACCACCTGATAGTTGATCTGGGAGACCACCGTGAGATACTGTCATTTGATGTGACTCAGT
AB091827 1678 CCCAATCATTTTGGAAACCCACCACCTGATAGTTGATCTGGGAGACCACCGTGAGATACTGTCATTTGATGTGACTCAGT

AJ006464 1745 CTCCATTCTTTCCTATTGTCCTAGGAATTCGTTGGCTGAGCACGCATGACCCTCACATTACCTGGAGTACCCGCTCCATT
AB091827 1758 CTCCATTCTTTCCTATTGTCCTAGGAATTCGTTGGCTGAGCACGCATGACCCTCACATTACCTGGAGTACCCGCTCCATT

AJ006464 1825 GTCTTCAACTCTGATTACTGCCGACTCTGCTGCCGGATGTTTGCACAGATACCTTCTAACTTACTGTTTACAGCGCCACA
AB091827 1838 GTCTTCAACTCTGATTACTGCCGACTTCGCTGCCGGATGTTTGCACAGATACCTTCTAACTTACTGTTTACAGTGCCACA

AJ006464 1905 ACCGAGTTCGCATCCGTATCTACTTCATCATGTGCATCCGCATGTCCATCCGTATATGCATCAGCATCTGCATCAGCATC
AB091827 1918 ACCGAATTTGCATCCGTATCTACTTCATCATGTGCATCCGCATGTCCATCCGCATATGCATCAGCATCTGCATCAGCATC

AJ006464 1985 TGCATCAGTTTCTGCATCCAGATCCGCATCAGTATCCGCATCCGGATCCGCATTATCATCATCATCAGCAGGCGGATATG
AB091827 1998 TGCATCAGTTTCTGCATCCAGATCCGCATCAGTATCCGCATCCGGATCCGCATTATCATCATCATCAGCAGGCGGATATG

AJ006464 2065 CAGCACCAACTGCAGCAGTATCTATATCAGTATTTGTATTACCACCTGTATCCGGTTATGCACCACCATCTGCCTCCAGA
AB091827 2078 CAGCACCAACTGCAGCAGTATCTATATCAGTATTTGTATTACCATCTGTATCCGGTTATGCACCACCATCTGCCTCCAGA

AJ006464 2145 TCAGCATGAGCATCTGCATGAGTATCTGCATCAGTATCTGCATCAGTATCTGCATCAGTTTCTGCATCACCACCTGCATC
AB091827 2158 TCAGCATGAGCATCTGCATGAGTATCTGCATCAGTATCTGCATCAGTATCTGCATCAGTTTCTGCATCACCATCTGCATC

AJ006464 2225 CGGATCTGCATCAGCATCTGTATCAGTATCTGCATAACCATATGAATCCGGATCCACATCACCACTCTCATCCAGATCCC
AB091827 2238 CGGATCTGCATCAGTATCTGTATCAGTATCTGCATAACCATATGAATCCGGATCCACATCACCATCCTCATCCAGATCCC

AJ006464 2305 CCTCAGGATCCACATCACCCTCCACATCAGGATCCACATCAGCATCCGGATCCCCATCAGGATGCACATCAGGATCCCCA
AB091827 2318 CCTCAGGATCCACATCACCCTCCACATCAGGATCCACAT----------------------------------------
```

```
AJ006464 2385 TCAGGATCCCCATCAGGATGCACATCAGGATCCACATCAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCAC
AB091827 2357 ------------------------------------CAGCATCCGGATCCCCATCAGGATCCTCCACATCAGGATCCAC

AJ006464 2465 ATCAGGATGCACATCAGGATCCCCATATGGATCCACACCTGCATCAGCACCAGCATCCGCAGCCGCAGCCGCATCCACAA
AB091827 2400 ATCAGGATGCACATCAGGATCCCCATATGGATCCACACCTGCATCAGCACCAGCATCCGCAGCCGCAGCCGCATCCACAA

AJ006464 2545 CAGTATCCTAACCATCCTCAGCAGCCACCATTCTTCTACCACATGGCTGGATTCAGAATTTACCACCCTGTAAGGTATTA
AB091827 2480 CAGCATCCTAACCATCCTCAGCAGCCACCATTCTTCTACCACATGGCTGGATTCAGAATTTACCACCCTGTAAGGTATTA

AJ006464 2625 CTATATTCAGAATGTGTATACACCTGTTGATGAGCATGTCTATCCGGGTCACCGGGTGGTTGACCCTAACATTGAGATGA
AB091827 2560 CTATATTCAGAATGTGTATACACCTGTTGATGAGCATGTCTATCCGGGTCACCGGGTGGTTGACCCTAACATTGAGATGA

AJ006464 2705 TTCCTGGAGCGCACAGCCTGCCCAGTGGACATTTGTACTCAATGTCTGAGTCTGAAATGAATGCTCTGCGAAATTTCGTG
AB091827 2640 TTCCTGGAGCGCACAGCCTGCCCAGTGGACATTTGTACTCAATGTCTGAGTCTGAAATGAATGCTCTGCGAAATTTCGTG

AJ006464 2785 GACAGGAATGTTAAAGATGGGCTCATGACTCCCACTGTGGCGCCCAATGGAGCCCAAGTCCTGCAAGTGAAAAGAGGGTG
AB091827 2720 GACAGGAATGTTAAAGATGGGCTCATGACTCCCACTGTGGCGCCCAATGGAGCCCAAGTCCTGCAAGTGAAAAGAGGGTG

AJ006464 2865 GAAACTCCAAGTCACTTACAATTGCCGAGCTCCACAGAGTGGCACCATCCAAAATCAGTACCTACGCATGTCTCTTCCAA
AB091827 2800 GAAACTCCAAGTCACTTACAATTGCCGAGCTCCACAGAGTGGCACCATCCAAAATCAGTACCTACGCATGTCTCTTCCAA

AJ006464 2945 ATATGGGAGACCCTGCACACCTGGCAAGCTATGGTGAATTTGTCCAAGTTCCTGGCTACCCATATCCAGCCTATGTTTAC
AB091827 2880 ATATGGGAGACCCTGCACACCTGGCAAGCTATGGTGAATTTGTCCAAGTTCCTGGCTACCCATATCCAGCCTATGTTTAC

AJ006464 3025 TATACAAGCCCGCATATGATGACTGCGTGGTACCCAGTAGGACGAGATGTACATGGACGAATAATCGTTGTGCCTGTTGT
AB091827 2960 TATACAAGCCCGCATATGATGACTGCGTGGTACCCAGTAGGACGAGATGTACATGGACGAATAATCGTTGTGCCTGTTGT

AJ006464 3105 AATCACCTGGTCTCAAAATACGAACCGCCAGCCTCCGGTGCCCCAGTATCCTCCTCCGCAGCCACCTCCACCACCACCAC
AB091827 3040 AATCACCTGGTCTCAAAATACGAACCGCCAGCCTCCGGTGCCCCAGTATCCTCCTCCGCAGCCACCTCCACCACCACCAC

AJ006464 3185 CACCTCCACCGCCACCACCACCTCCACCAGCATCATCCTGCAGTGCTGCGTACAACCTGTCATGTCCTTTGTAGTCTCTG
AB091827 3120 CACCTCCACCGCCACCACCACCTCCACCAGCATCATCCTGCAGTGCTGCGTACAACCTGTCATGTCCTTTGTAGTCTCTG

AJ006464 3265 CCCTCAACTTGATCCTGTGCAGCTTCTCAATCTATGACTGTGTGGTACTGGACCTTCAGAGGCGCACAGAGCTCAAGTCA
AB091827 3200 CCCTCAACTTGATCCTGTGCAGCTTCTCAATCTATGACTGTGTGGTACTGGACCTTCAGAGGCGCACAGAGCTCAAGTCA

AJ006464 3345 GTTTTCGTCTTGACTGCCACTTTATAAGTTGACAGGCCTGGGTTTTACTTGTTAAAACCTCTCACCATCTCAATCACAGG
AB091827 3280 GTTTTCGTCTTGACTGCCACTTTATAAGTTGACAGGCCTGGGTTTTACTTGTTAAAACCTCTCACCATCTCAATCACAGG

AJ006464 3425 CTGCCAAGTGTCTTTACAAAGAAGCTGATACAAACACAGGCCATGCTGATTTCTTACAGAGGGAGAGAAGAGGAAGAGAA
AB091827 3360 CTGCCAAGTGTCTTTACAAAGAAGCTGATACAAACACAGGCCATGCTGATTTCTTACAGAGGGAGAGAAGAGGAAGAGAA

AJ006464 3505 GAAGAAAGAGGAGGAAGAGGACATGACTTGCCCATATGCTGGGCACCTTATAAAGGAAGCCAGACTTTTCGGTGCAGTAT
AB091827 3440 GAAGAAAGAGGAGGAAGAGGACATGACTTGCCCATATGCTGGGCACCTTATAAAGGA-GCCAGACTTTTCGGTGCAGTAT

AJ006464 3585 GGAAAGGCTTCCGTGATTCTCTTGCTGCACCCCACGAAACTTCACCACCTTCAAACTCCATTTTCACGGTTCCGTTAATT
AB091827 3519 GGAAAGGCTTCCGTTATTCTCTTGCTGCACCCCACGAAACTTCACCACCTTCAAACTCCATTTTCACGGTTCCGTTAATT

AJ006464 3665 TTCAAGGAGCAGCAACTCGACTGGTTCTCTGCTACATGAAACACCTCAGCTTGAAAAGGAAGTGCTCTCTCAGACTGACT
AB091827 3599 TTCAAGGAGCAGCAACTCGACTGGTACTCTGCTACATGAAACACCTCAGCTTGAAAAGGAAGTGCTCTCTCAGACTGACT
```

```
AJ006464 3745 TGTGAGTGTGCCTTCACATTCTGGTGCAAATCATGTGTACCCAAGAACTCTGACATAGCATCTTACCATCATCATGCCAG
AB091827 3679 TGTGAGTGTGCCTTCACATTCTGGTGCAAATCATGTGTACCCAAGAACTCTGACATAGCATCTTACCATCATCATGCCAG

AJ006464 3825 GAATTGGTTTCCTTAAAGTGTGACACTTGGCAACCAGAAGCGTCTGGTGGCTTCAGGTGTGTTAGGGTTTTGTAGAAAGT
AB091827 3759 GAATTGGTTTCCTTAAAGTGTGACACTTGGCAACCAGAAGCGTCTGGTGGCTTCAGGTGTGTTAGGGTTTTGTAGAAAGT

AJ006464 3905 TGTAACCACCTCTCAGTCATGCAAGAGGGTACATCAAAAAGACTTCTAAATTTGGTACCCAGACTTTGACAAGAGCCTGT
AB091827 3839 TGTAACCACCTCTCAGTCATGCAAGAGGGTACATCAAAAAGACTTCTAAATTTGGTACCCAGACTTTGACAAGAGCCTGT

AJ006464 3985 GAGTTAGTTCATTGTAACCAACTTGTCAGGAACTTCAGCAAGAATTTTCATGAATGTTGTTAATTGGCTGGAGAAATACC
AB091827 3919 GAGTTAGTTCATTGTAACCAACTTGTCAGGAACTTCAGCAAGAATTTTCATGAATGTTGTTAATTGGCTGGAGAAATACC

AJ006464 4065 AGCAAAGTTTTGCAATCATCCACCCATTTGTCTGCCATTGTGCATCTGGACATCCATCAACCGCCCACTCAGTGGCATCT
AB091827 3999 AGCAAAGTTTTGCAATCATCCACCCATTTGTCTGCCATTGTGCATCTGGACATCCATCAACCGCCCACTCAGTGGCATCT

AJ006464 4145 ACCATGAAGATGTTTAGAGACGGGAAATGATTTTGCACCCAGATCCACACTCCTCTCATGTCAGATATTTGGAATTAGAG
AB091827 4079 ACCATGAAGATGTTTAGAGACGGGAAATGATTTTGCACCCAGATCCACACTCCTCTCATGTCAGATATTTGGAATTAGAG

AJ006464 4225 CACCGGGAATCGCCTTGAAGATATGAACTTCTAAGAATGGAGTCTCTTGCTACTACCTTGAAACATTTGTTTATCCTCTC
AB091827 4159 CACCGGGAATCGTCTTGAAGATATGAACTTCTAAGAATGGAGTCTCTTGCTACTACCTTGAAACATTTGTTTATCCTCTC

AJ006464 4305 TTTCTGTTTGATTCTACTACTACCATTAACCCTGCTGCAGGATTTGTCACCATTCTGCCTGACTGCTGAGACTCCATTTT
AB091827 4239 TTTCTGTTTGATTCTACTACTACCATTAACCCTGCTGCAGGATTTGTCACCATTCTGCCTGACTGCTGAGACTCCATTTT

AJ006464 4385 GCTGCTATGCAAGGAGATGAAAGAGGCAGGCCTAAAGGAGCAGTAGGACATAACATGGTTTTTATTTCTATCTGTCATGA
AB091827 4319 GCTGCTATGCAAGGAGATGAAAGAGGCAGGCCTAAAGGAGCAGTAGGACATAACATGGTTTTTATTTCTATCTGTCATGA

AJ006464 4465 TCTTAATGGTGAAGTTTCCTTTTTCGTCAGCATTTATCTCTTTGTTATCCTGACATGTTTTAATTAGTTTAGTGGGTTTT
AB091827 4399 TCTTAATGGTGAAGTTTCCTTTTTCGTCAGCATTTATCTCTTTGTTATCCTGACATGTTTTAATTAGTTTAGTGGGTTTT

AJ006464 4545 TTTTTTTCTATTGGTGGTGGTGTTCTTTTGTTGTTGTTTTTTGTGGTCGTCATTTTGATTTTGGATCACTTCGTGTTTTA
AB091827 4479 TTTTTT-CTATTGGTGGTGGTGTTCTTTTGTTGTTGTTTTTTGTGGTCGTCATTTTGATTTTGGATCACTTCGTGTTTTA

AJ006464 4625 CAGTAATTACTTTTAAATGGTGCATTTGCTTCTGATTTTTTTTTTTTTTATGAAGCATCACATCAGTTTACCTCATATCTC
AB091827 4558 CAGTAATTACTTTTAAATGGTGCATTTGCTTCTGATTTTTTTTTTTT-ATGAAGCATCACATCAGTTTACCTCATATCTC

AJ006464 4705 AATTCCATCCTTCATGCATTTTTTTTTTTAACTCATTTGATCTTCAAGCTGCAGAGGGCCTAGCAATGGGTCATCACCTGC
AB091827 4637 AATTCCATCCTTCATGCATTTTTTTTTTTAACTCATTTGATCTTCAAGCTGCAGAGGGCCTAGCAATGGGTCATCACCTGC

AJ006464 4785 AGCCCTGGCATGTACACACGGACATTTGCCACCACTGAAAGCAAACAGTTGGAGAAGTTGGCCACCTGAGTCAAGGAGGT
AB091827 4717 AGCCCTGGCATGTACACACGGACATTTGCCACCACTGAAAGCAAACAGTTGGAGAAGTTGGCCACCTGAGTCAAGGAGGT

AJ006464 4865 GGTGCTGGTGTGAGTTCACAGCTCACAGGGGACGGTGAACGTTGACATTGACTTTGGCAGTGTGTACTATGCTCTACTCC
AB091827 4797 GGTGCTGGTGTGAGTTCACAGCTCACAGGGGACGGTGAACGTTGACATTGACTTTGGCAGTGTGTACTATGCTCTACTCC

AJ006464 4945 TATATATACTCTATAGATGTTAGGCATTAAGGATAAGTGATCTTAAATTTACTGAAATTTTGTTAAGTTGATTAGATTTA
AB091827 4877 TATATATACTCTATAGATGTTAGGCATTAAGGATAAGTGATCTTAAATTTACTGAAATTTTGTTAAGTTGATTAGATTTA

AJ006464 5025 CATTAATTGTTTATATTGTGTTTTTTCTTTTATTTTAGCACTTCCAACAAAAATGCCTTTATTCATTTATTAGGAAGAAA
AB091827 4957 CATTAATTGTTTATATTGTGTTTTTTCTTTTATTTTAGCACTTCCAACAAAAATGCCTTTATTCATTTATTAGGAAGAAA
```

```
AJ006464 5105 TTGGAGTGGCGAACACAAACTAGCAAAATTAATTAATTGGCTGTGGGCCCAAATTTGAATTC----------
AB091827 5037 TTGGAGTGGCGAACACAAACTAGCAAAATTAATTAATTGGCTGTGGGCCCAAATTTGAATTCCTTATTAATTCAAAGTGA

AJ006464      ----------------------------------------------------------------------------
AB091827 5117 AGGAAGTGTGAGAGTTTCTATTGCTGGTCATGAAGTTCACCACGAAGAGTCAGTTTAGTTTGTACCGAGAGGCATTTAGC

AJ006464      ----------------------------------------------------------------------------
AB091827 5197 TGAGAGTGATTTGAGTTGGGCATCTCTATGATGCAGTCCAGACTTGTGTTTAAGTTTACAGGTACCTCTTGGACTCCTGA

AJ006464      ----------------------------------------------------------------------------
AB091827 5277 ATTGATCTCGATGTCACACATCATCGACATCCCGCATCCCACACAGTTCCAAGATTGGCGACAGAGAGCATCCTGTTGGA

AJ006464      ----------------------------------------------------------------------------
AB091827 5357 GAGACCTGACCAGTCACGAAGGACCGCCTGCAGACTCTTCTGCAGGACCATGCTACATATTCAGCTGGAGAGGGAACACG

AJ006464      ----------------------------------------------------------------------------
AB091827 5437 CATCTGATGGGAATGGATTGTTTTTGCTTGTTTCAGAATTTTGCTGTTGAGATGTTGTTAAACTTTGCATTGTTTTTCTT

AJ006464      ----------------------------------------------------------------------------
AB091827 5517 TTTTCTTTTTTTTTCCCTATAGTCAATTAAGAATAAGGGGTAGATAATCATAAGTATTTTGGGCTGGGAGGGATTGTTAA

AJ006464      ----------------------------------------------------------------------------
AB091827 5597 GTAATCTTAGGTGGGTGGGTAATTTAGGAAATTAAGTTAGGATAAGATAGGATAAGATAAGATAAGCTAGGATAAGTTAG

AJ006464      ----------------------------------------------------------------------------
AB091827 5677 ATAAATTAGGATAAGTTAGATAAGTTAGGATAAATTCCATAAGATAGGATAAGTTCCATAAGATAGGATAAGTTAGTTCA

AJ006464      ----------------------------------------------------------------------------
AB091827 5757 CATAGGATAAGCATAGGATAAGTTAGATCACATAGGATAAGTTAGATCACATAGGATAAGTTAGATCAAATACCTCAACA

AJ006464      ----------------------------------------------------------------------------
AB091827 5837 AGTGGACGAGTGTACTTATTGGTCCCTTACCCCTACCAACTTTACCTTTAAGGCCGAGCTCAGAGGGAACTACAGGGAAA

AJ006464      ----------------------------------------------------------------------------
AB091827 5917 TCAGTGTTAGGAGTGAATTGGACATGGATGACATTCTAGTGAAGGTTAGGACGATTAGAGTAATCTAATACAGGCTCTGG

AJ006464      ----------------------------------------------------------------------------
AB091827 5997 GAATTAACTCAAGGGATTTTAGGAATGCCAGATGACAGTCATCATTTGATAAAACTGTTGCTGGAATAAAGTTTAGAATG

AJ006464      ----------------------------------------------------------------------------
AB091827 6077 TTGAATGGTAAATGTTAATGGAAATTGTACGAAAATGAATAAAAGCTCTTTCCATACCCCTCCCCCTCCCTACCCCACTC

AJ006464      ----------------------------------------------------------------------------
AB091827 6157 TTTCTTTCCTCCCCCCCTCCATAACCTACCCCTTCTCCCAAACACACCCCTCCATTTCCTAACTCACCACGTCTCCCCTT

AJ006464      ----------------------------------------------------------------------------
AB091827 6237 CCACTCCCTAACCCACCCCTTCTTCCCTCCCTCCCTCCAGAACCCGCCCCCATCTTCCCCCTCCCTAACACACCCCTTCT

AJ006464      ----------------------------------------------------------------------------
AB091827 6317 TCCTTTCCCCCTCCAACCCCACCCTTCAAACTATATTACTGTATTCTCTTTATTATATTTCTTTGTATTAATAAAACTGT
```

```
AJ006464       ----------
AB091827 6397 ATTGATGTTAC
```