Developmental Silencing: Actions and Interactions of the Polycomb-Group

by

Tiffani L. Howard

A Dissertation

Presented to the Department of Cell and Developmental Biology
and the Oregon Health Science University School of Medicine

in partial fulfillment for the degree of
Doctor of Philosophy

August 1999

School of Medicine
Oregon Health Science University

Certificate of Approval

This is to certify that the Ph.D. thesis of

Tiffani L. Howard

has been approved.

[redacted]
Stanley Hollenberg
Professor in charge of thesis

[redacted]
Jan Christian
Chairwoman

[redacted]
Matt Thayer
Member

[redacted]
Rich Maurer
Member

Associate Dean for Graduate Studies

TABLE OF CONTENTS

Developmental Silencing: Actions and Interactions of the Polycomb-Group

Table of figures		ii
Acknowledgements		v
Abstract		vi
Chapter 1	Introduction	1
Chapter 2	Material and Methods	13
Chapter 3	Subcellular localization of Th1 is regulated during development	21
Chapter 4	Th1 interacts with the Polycomb-Group	29
Chapter 5	Characterization of additional Th1 partners, CENP-B and TFII-I	38
Chapter 6	Overexpression of PcG proteins causes a specific neural phenoype and affects specific targets in Xenopus	46
Chapter 7	Identification of two new Polycomb-Group proteins, BS69 and PRSM, by the Xenopus assay system	55
Chapter 8	Discussion	59
References		75

List of Figures

Figure	Chapter	Title	Page
Figure 1	1	Th1 subcellular localization is concentration dependent	**21**
Figure 2	1	Cells with ectopic Th1 show exclusively nuclear or cytoplasmic protein expression, or display apoptotic characteristics	22
Figure 3	1	Th1 contains two nuclear localization signals	23
Figure 4	1	Th1 is nuclear in undifferentiated Rcho-1 cells but becomes cytoplasmic upon giant cell differentiation	24
Figure 5	1	Th1 is nuclear in undifferentiated PC-12 cells, but becomes cytoplasmic upon serum deprivation	25
Figure 6	1	In vivo Th1 expression is nuclear at e12.5	26
Figure 7	1	At e12.5 Th1 expression in the heart and sympathetic ganglia is nuclear, but becomes cytoplasmic at e13.5	26
Figure 8	1	Transfected Th1 is trapped in the cytoplasm by co-expression of I-mf-a	27
Figure 9	2	Comparison of mPc12 and mPh2 to related proteins	30
Figure 10	2	Th1 interacts with mPc12, mPh2, and Xpc in the yeast two hybrid system	32
Figure 11	2	Th1-PcG and PcG-PcG interactions in vitro	33
Figure 12	2	Transfected Th1 localizes to subnuclear domains	34
Figure 13	2	Th1 recruits PcG proteins	35
Figure 14	2	Th1 directs PcG proteins to its subnuclear domains	36
Figure 15	3	Strategy for mammalian two-hybrid and transcriptional preference assays	40
Figure 16	3	TFII-I interacts with Th1 in the mammalian two hybrid assay	40
Figure 17	3	Both TFII-I and CENP-B are repressors, each with a defined repression domain	41
Figure 18	3	TFII-I amd CENP-B protein structure	42
Figure 19	3	Ectopic TFII-I and CENP-B colocalize with transfected Th1	43
Figure 20	3	CENP-B colocalizes with and recruits PcG members	44
Figure 21	3	CENP-B null mutant mice do not exhibit axial homeotic mutations	45
Figure 22	4	Sequence comparison of Pcl from Xenopus, mammals and Drosophila	46
Figure 23	4	RT-PCR quantitation of Xpcl1 expression during Xenopus embrogenesis	47
Figure 24	4	Xpcl1 is expressed in the anterior central nervous system	47
Figure 25	4	Strategy for Xenopus assay system	48

Figure 26	4	Phenotypic consequences of Pcl1 overexpression	48
Figure 27	4	Analysis of the Pcl1 overexpression phenotype	50
Figure 28	4	Pcl effects on brain and eye markers	51
Figure 29	4	PcG proteins produce similar anterior neural phenotypes	52
Figure 30	4	PcG proteins produce a similar gross phenotype	52
Figure 31	4	Pcl1 alters expression of En-2 and Krox-20	53
Figure 32	4	PcGs produce similar effects on En-2 expression	55
Figure 33	5	BS69 overexpression in Xenopus embryos causes a PcG overexpression phenotype	56
Figure 34	5	Sections of a BS69 injected embryo reveal increased and disorganized neural tissue	56
Figure 35	5	PRSM shifts and represses En-2	57

This dissertation is dedicated to

Dr. Elver Voth

whose love of science was rooted in his love of God.

Acknowledgements

I am especially grateful to Stan Hollenberg for supporting my efforts both academically and personally. Thank you for the early morning heart-to-hearts. Your instruction has been patient, careful, and generous. In short, thanks, Stan, for choosing a wild card and playing the game fairly to the end.

Thank you, Jan Christian, for generously allowing us to take over your lab, and for your willingness to talk frankly with me.

Thank you, Melinda Walters, Chang Sheng Guo, and Jane Diaz for patiently reading manuscripts (over and over) and listening to practice seminars (over and over).

Thank you, Kent Thornberg, for believing in me.

Thank you, Mom and Dad, for saying, "I knew you could do it," and "I'm so proud of my kids," and "What was a gene again?"

And thank you, Paul, My Love, for constant quiet confidence in me.

Finally, I thank my family and friends for their loyalty. I could not have finished this work without you.

Abstract

During development it is essential that genes are both activated and silenced in a temporally and spatially specific manner. The Polycomb-Group (PcG) comprises a collection of structurally distinct proteins responsible for long-term stable gene silencing. In *Drosophila*, homeotic targets of PcG repression show expanded expression domains in PcG mutant flies, which causes homeotic transformation. Mutation of mammalian PcGs also causes axial transformations, but direct target genes are not yet identified and little is known about how the PcG members choose their targets. This thesis describes a transient assay system which has allowed identification of two new vertebrate target genes, *Engrailed-2* and *Rx2A*, and two new PcG members, BS69 and PRSM. In addition, two hybrid interactions with the tissue specific basic helix-loop-helix transcription factor, Th1, provide clues to the mechanism of PcG selection of targets. We show that Th1, a DNA binding protein, is capable of interacting with multiple PcG members, suggesting that it may bring PcG proteins to a potential repression site. Th1 may also be fundamental in controlling repression versus activation of targets, as it can interact with and recruit a protein at the transcriptionally silent centromere, CENP-B, and a factor associated with the transcriptional machinery, TFII-I. These observations address three fundamental questions: what are the proteins in PcG complexes, what are the vertebrate targets of repression and how are the targets chosen.

Developmental Silencing:

Actions and Interactions of the Polycomb-Group

Chapter 1

Introduction

Helix-loop-helix transcription factors have fundamental roles in development

Basic-helix-loop-helix (bHLH) proteins are transcription factors with crucial roles during development. Skeletal myogenesis, neurogenesis, and hematopoiesis are a subset of the processes that require bHLH expression (Campos-Ortega 1995; Edmondson and Olson 1993; Kageyama *et al.* 1995; Shivdasani and Orkin 1996). Basic-HLH proteins have been grouped into four classes based on protein interactions, DNA binding ability, and transcriptional activity. 1) E proteins, also called Class A, are widely-expressed and act as partner for at least two of the other classes (Murre *et al.* 1989). 2) Tissue-specific bHLH proteins, or Class B, pair with E proteins to produce heterodimers. E protein homodimers and heterodimers with tissue-specific bHLHs are thought generally to function as transcriptional activators (Braun *et al.* 1990; Quong *et al.* 1993; Weintraub 1993). 3) Id proteins lack a basic region and form a non-DNA-binding complex with E proteins and some tissue-specific factors (Benezra *et al.* 1990). 4) hairy and Enhancer-of-split bHLH proteins form a fourth class whose interactions with the other three groups are not well characterized. These proteins act genetically as repressors of bHLH-mediated functions by inhibiting promoter activity (Knust *et al.* 1987).

The DNA binding specificity of Class B bHLH proteins combines with a restricted tissue distribution to select cell-type-specific genes. For example, members of the MyoD family are skeletal muscle specific and regulate muscle differentiation genes, whereas achaete-scute and NeuroD proteins are found in the nervous system and are thought to promote neurogenesis (Kageyama *et al.* 1995). Other Class B proteins, though, have a broader tissue distribution. For example, *SCL/TAL1* RNA is found in cells of the hematopoietic lineage, endothelial cells and in the brain (Kallianpur *et al.* 1994). The

isolation and characterization of new bHLH proteins has helped to define the roles of cell-specific versus more widely-expressed regulators.

Thing1 is tissue specific

The high affinity of tissue-specific bHLH proteins for E proteins provided the strategy for the isolation of a Thing1 (Th1) cDNA (Hollenberg *et al.* 1995), also known as eHAND (HAND1) (Cserjesi *et al.* 1995) and Hxt (Cross *et al.* 1995). Two groups obtained the cDNA by two hybrid protein interaction screens with E protein from *Drosophila* (daughterless) (Hollenberg *et al.* 1995) and mouse (E12) (Cserjesi *et al.* 1995), respectively. Cross and coworkers identified Hxt by screening a phage expression library with the HLH domain of E47. Based on *in situ* hybridization, *Th1* is expressed in both embryonic and extraembryonic components. By embryonic day nine of development (E9.0), *Th1* is expressed in the developing heart, gut, and the pharyngeal arches. The primordial sympathetic ganglia and adrenal medulla begin to express *Th1* RNA soon after. Gut smooth muscle expression is maintained in adult mice and is joined by low level liver and adrenal expression (Cserjesi *et al.* 1995; Hollenberg *et al.* 1995). Extraembryonic *Th1* expression is considerable in trophoblast cells which are present in a number of tissues including the chorion, yolk sac and ectoplacental cone (Cross *et al.* 1995). Th1 expression is more complex than that of most of the known tissue-specific bHLHs suggesting a more general role for this transcription factor. A second cDNA with high homology to Th1 was isolated at the same time and named Thing2 (Th2) (Hollenberg *et al.* 1995). Th2 is also expressed in the heart and a number of other locations where Th1 is not expressed (Cserjesi *et al.* 1995).

The Thing proteins are required for heart development

Th1 and Th2 have fundamental roles in heart development. Chick antisense experiments showed that inhibiting expression of both Th1 and Th2 blocks heart tube

maturation (Srivastava *et al.* 1995), suggesting that the proteins have redundant roles. However, in the mouse heart Th1 and Th2 have only partially overlapping expression patterns. Th1 is found throughout the early heart tube, but later it is restricted to the left ventricle and outflow tract at opposite ends of the heart tube, where it is maintained into adulthood. In contrast, Th2 becomes restricted to the right ventricle, and expression is eventually lost except in the aortic arteries (Biben and Harvey 1997).

Null mutant mice have provided many clues as to the role of the Thing proteins in heart development. Both Th1 and Th2 knockout mice are embryonic lethal at e10.5 in part due to failure of the heart to loop (Srivastava *et al.* 1997). A rightward loop is observed in all vertebrates and is required for heart morphogenesis and chamber formation, but the mechanism for this event is not clear. Differential proliferation or expression of cell adhesion or cell motility (actin) molecules has been postulated but not proven (Olson and Srivastava 1996). In *Drosophila*, two genes are implicated in early heart development; Nkx2.5 (tinman homeodomain protein) is necessary for commitment of heart cells, and Mef2C (MADS box protein) has a role in specifying cardiac muscle (Bour *et al.* 1995; Lilly *et al.* 1995; Ranganayakulu *et al.* 1998). In Nkx2.5 and Mef2C null mutant mice, which also have defective looping, Th1 left handed restriction is lost (Biben and Harvey 1997; Lin *et al.* 1997) placing spatially restricted Th1 cardiac expression under the control of these two important heart specifiers.

Th1 and Th2 null mice have other heart defects that reflect the proteins' natural expression patterns. In the hearts of Th1 mutant mice, the left ventricle is unspecified and trabeculation is absent (Riley *et al.* 1998). Th2 mutant mice lack the right ventricle and have no trabecula or aortic arteries (Srivastava *et al.* 1997; Thomas *et al.* 1998). It has been suggested that the loss of Th2 is partially compensated by upregulation of Th1, allowing Th1 to assume some Th2 functions. These data indicate that Th1 and Th2 have important roles in dictating the onset of looping. In addition, these mutational studies demonstrate that

restricted expression patterns for the two proteins are established to determine the ventricle lineage.

Differentiation of placental trophoblasts is Th1 dependent

Studies of Th1 expression have determined that the highest levels of Th1 mRNA are in the most differentiated trophoblasts within the placenta (Cross *et al.* 1995). The role of Th1 in trophoblast differentiation, both in embryonic stem cells and in cultured Rcho-1 trophoblast cells, has been explored. Blastocyst cells can assume two distinct fates, producing either extraembryonic or embryonic tissues. Cross and coworkers have shown that overexpression of Th1 in blastocysts alters cell fate preference. Th1 causes a disproportionate number of cells of the blastocyst to become trophoblast (extraembryonic) cells. In addition, fewer of each cell type are produced. These researchers suggested that more cells are fated to be trophoblasts, at the expense of the inner cell mass population (embryonic), and that the number of trophoblasts is decreased due to premature differentiation, evidenced by exceptionally large trophoblasts. Additional evidence to support this conclusion was obtained using the trophoblast cell line, Rcho-1, which matures into giant cells when exposed to low serum conditions (Faria and Soares 1991). Th1 expression is endogenous in these cells and increases with differentiation. When the Th1 protein level in undifferentiated cells was elevated by transfection, differentiation was induced in as many as 99% of the transfected cells (Cross *et al.* 1995). Consistent with these observations, Th1 null mutant mice die at e7.5 with trophoblast cell defects leading to an unsuccessful implantation, requiring creation of tetraploid mice for study of heart development (Firulli *et al.* 1998; Riley *et al.* 1998). Together these results lead to the conclusion that Th1 is required for development of the trophoblast lineage.

Downregulation of Thing2 expression causes branchial arch defects

Th2 protein is required for development of the branchial arches, which are transient pockets of cells giving rise to many structures of the head and neck. CATCH-22 syndrome is the term used to describe the numerous diverse human congenital defects caused by abnormalities of the neural crest-derived branchial arches (Lammer and Opitz 1986; Wilson *et al.* 1993). Two distinct mouse strains display characteristics of the CATCH-22 syndrome, endothelin-1 and Th2 null mutant mice (Thomas *et al.* 1998). In the endothelin null mouse both Th1 and Th2 are downregulated, suggesting that loss of Th2 in the endothelin mutant mice may explain the CATCH-22 phenotype observed. The third and fourth branchial arches of endothelin and Th2 mutant mice are greatly reduced. This is not a result of failure of cells to migrate from the neural crest to the arches, or a failure of cells to differentiate into arch cells, as evidenced by the expression of homeobox proteins, Msx1 and 2 at the arch site. Rather, the branchial arches are small as a result of increased apoptosis. These data indicate that Th2 is required for survival of the branchial arch cell population and is therefore critical for later development of the craniofacial structures and aortic arches.

Thing1 is both an activator and repressor of transcription

Analysis of the transcriptional potential of Th1 revealed two activities (Hollenberg *et al.* 1995). As a heterodimer with E protein, Th1 activates transcription, but when tethered to DNA in the absence of E protein, Th1 represses transcription. A set of specific DNA binding sites for the Th1/E protein heterodimer has been identified (Hollenberg *et al.* 1995), but a distinct DNA binding site favoring repression has not. There is a small assortment of suggested targets for activation by Thing proteins, all discovered by identifying genes that are downregulated in the Th1 and Th2 null mutant mouse hearts. *Adenylosuccinate synthase 1*, encoding a cardiac metabolic protein, may be a target of both Th1 and Th2 (Lewis *et al.* 1999). Th1 null mice are devoid of myosin light chain-2 protein (Firulli *et al.* 1998; Riley *et al.* 1998) while Th2 targets include homeobox gene *Msx1* (Thomas *et al.*

1998), *GATA4* (Srivastava *et al.* 1997), and *Ufd1* (Yamagishi *et al.* 1999). Studies show that *Mash2* is ectopically expressed in the Th1 null mouse suggesting the Mash2 might be a target of Th1 repression (Riley *et al.* 1998). To date, the only other report of a possible Thing target for repression is *Th1* in the Th2 knockout. Interestingly, mutation of Th1 does not appear to affect *Th2* expression.

Finding targets of repression, specifying DNA binding sites capable of recruiting Th1 for repression, or identifying partners for Th1 that encourage repression would greatly expedite research in this field. We have screened with Th1 in the yeast two hybrid system and have identified three Polycomb-Group (PcG) partners and two other strong repressors. The biochemical and cellular interactions presented in Chapters Four and Five indicate that Th1 may repress transcription during development by association with a PcG complex. These results suggest a role for tissue specific bHLH proteins in repressing regulatory genes and thus stabilizing developmental decisions.

Epigenetic silencing is fundamental to development

During development cells become committed to distinct lineages through the expression of specific combinations of transcription factors. The signals that determine these combinations are frequently short-lived, for example, Th2 is only expressed embryonically in the heart, yet the cellular lineages it helps define persist for many cell divisions. Elegant genetic studies in *Drosophila* have demonstrated the importance of epigenetic mechanisms for maintaining critical regulatory genes in an active or silent state (Kennison 1995; Orlando and Paro 1995; Simon 1995; Tamkun 1995). One specific silencing mechanism uses the products of the PcG. The PcG has been proposed to generate or stabilize inactive chromatin configurations on specific target genes (for review see (Bienz and Muller 1995; Paro 1995; Pirrotta 1997b).

The majority of the PcG has been identified through analysis of homeotic mutants in *Drosophila*. These mutants show proper restriction of homeotic expression patterns at

the earliest times of expression, but widespread derepression soon thereafter (Celniker *et al.* 1990; Struhl and Akam 1985; Struhl and White 1985; Wedeen *et al.* 1986). The phenotypic consequence is posteriorization of the embryo through ectopic anterior expression of homeotic genes. Several genetic studies indicate that *Drosophila* PcG gene products play a more general role in stabilizing developmental decisions (McKeon *et al.* 1994; Moazed and O'Farrell 1992; Pelegri and Lehmann 1994; Santamaría and Randsholt 1995). Various lines of evidence demonstrate that these PcG activities are direct transcriptional effects. For example, Paro and coworkers have demonstrated that homeotic loci in the *bithorax* complex and the segment polarity gene *engrailed* are physically bound by a PcG protein, Polycomb (Pc) (Strutt and Paro 1997).

The Polycomb-group mechanism for silencing target genes and stabilizing developmental decisions has been conserved between *Drosophila* and vertebrates. Structural homologs of virtually all *Drosophila* PcG proteins have been identified in mammals. Although the structural similarity between species is limited, it is sufficient to mediate similar protein-protein contacts (Alkema *et al.* 1997b; Hashimoto *et al.* 1998; Jones *et al.* 1998; Kyba and Brock 1998; Sewalt *et al.* 1998; van Lohuizen 1998). In fact, the mouse Pc homologue, M33, can partially rescue Pc mutant flies (Müller *et al.* 1995). Function is also conserved, as demonstrated by genetic analysis of mouse PcG homologs. Targeted or spontaneous mutation of five different mammalian PcG genes produces posteriorizing axial transformations of the mouse skeleton and anterior shifts of Hox expression boundaries, effects typical of a homeotic phenotype (Akasaka *et al.* 1996; Core *et al.* 1997; Shumacher *et al.* 1996; Takihara *et al.* 1997; van der Lugt *et al.* 1994). Finally, Bunker and Kingston (Bunker and Kingston 1994) have shown that both *Drosophila* and mammalian PcG proteins have repression activity in a mammalian cell culture system.

PcG complexes are heterogeneous

The PcG has been estimated to consist of 30-40 genes in *Drosophila* (Jurgens 1985; Landecker *et al.* 1994), with fifteen identified at the molecular level, including *Polycomb, polyhomeotic (ph), Polycomblike (Pcl),* and *Posterior Sex Combs (Psc)*. The protein structures deduced from cloned PcG genes are generally unrelated but contain motifs present in other nuclear factors (Pirrotta 1997a). For example, Pc contains a chromodomain motif also found in the heterochromatin-associated protein, HP1 (Paro and Hogness 1991), and both Psc and Supresser of zeste 2 have a ring finger (Brunk *et al.* 1991; van Lohuizen *et al.* 1991)).

Staining polytene chromosomes with specific antibodies to three PcG proteins, Pc, Ph and Pcl, has defined a set of about 100 common binding loci (Franke *et al.* 1992; Lonie *et al.* 1994; Rastelli *et al.* 1993). This co-localization and the co-immunoprecipitation of several PcG proteins (Alkema *et al.* 1997a; Franke *et al.* 1992; Hashimoto *et al.* 1998) strongly support the hypothesis that PcG proteins act in multi-protein complexes (Alberts and Sternglanz 1990; Gaunt and Singh 1990; Paro 1990; Reuter *et al.* 1990). Indirect evidence from genetic analysis (Cheng *et al.* 1994; Soto *et al.* 1995) and immunolabeling *Drosophila* polytene chromosomes (Rastelli *et al.* 1993; Sinclair *et al.* 1998; Stankunas *et al.* 1998) suggests that PcG complexes may be heterogeneous and somewhat gene-specific. For example, the PcG protein Psc is detectable at only half of the cytological loci occupied by Pc, Ph and Pcl. Soto and coworkers have determined, by assaying zygotic and maternal null fly lines, that PcG proteins act at different times and tend to affect different tissues (Soto *et al.* 1995). This suggests that either the binding sites or the PcG complexes formed on the sites are tissue specific. Identifying tissue-specific proteins which directly mediate DNA binding or regulate the composition of PcG complexes would broaden our understanding of the action of PcG proteins. Chapter Four of this thesis provides evidence that the transcription factor, Thing1, is such a bridging protein. Its tissue and subcellular localization are characterized in Chapter Three and additional Th1 partners are discussed in Chapter Five.

Epigenetic gene regulation includes activation as well. In *Drosophila,* Trithorax-Group (TrxG) proteins are responsible for stable activation of target genes. Members of the TrxG have been localized to many of the same polytene chromosome loci as PcG proteins (Chinwalla *et al.* 1995), but direct interactions have not been reported. In addition, some protein motifs are shared between members of both groups, suggesting a similar mechanism of action. These two groups of proteins appear to functionally antagonize each other as evidenced by double mutant flies null for Pc and TRX, where homeotic defects are abolished (Kennison and Tamkun 1988). Brahma, a TrxG protein, was identified in a screen for suppression of Pc and has structural and functional similarity to yeast SWI2 (Tamkun *et al.* 1992). As a member of a large complex, SWI2 modulates chromatin structure to facilitate activation of a variety of target genes (Peterson and Tamkun 1995). Therefore, it is postulated that the TrxG may facilitate activation of target genes by preventing or reversing chromatin packaging put in place by PcG members (Orlando *et al.* 1998).

PcG functions through Polycomb response elements in specific targets

Drosophila binding sites for PcG complexes, termed Polycomb response elements (PREs), have been identified in several genes including *Ultrabithorax (Ubx*), engrailed *(en),* and *Ph* (Chan *et al.* 1994; Chang *et al.* 1995; Chiang *et al.* 1995; Christen and Bienz 1994; Nomura *et al.* 1998; Simon *et al.* 1993; Strutt and Paro 1997). Much has been learned about the structure of complexes on PREs by crosslinking DNA to proteins followed by immunoprecipitating chromatin fragments with PcG-specific antibodies (Orlando and Paro 1993). Chromatin immunoprecipitation indicates that Pc sits near transcriptional start sites and on PREs (Chiang *et al.* 1995; Orlando *et al.* 1998; Orlando and Paro 1993; Strutt and Paro 1997). Suprisingly, the occupancy of a gene by Pc does not necessarily correlate with transcriptional inactivity.

Until recently, moderate DNA binding activity had been identified for only two PcG members, Pcl (Inouye *et al.* 1994) and Mel18 (Kanno *et al.* 1995). Cloning of an additional PcG protein, the product of the *pleiohomeotic locus (pho,)* revealed a YY1 homolog with strong DNA binding ability and specificity (Brown *et al.* 1998). Probable YY1/pho sites have been found near all PcG regulated genes (Mihaly *et al.* 1998). Efficient complex formation may require other PcG binding sites found in association with YY1 sites. Pho may be responsible for initiating DNA binding that is further stabilized by recruitment of other PcG proteins with weaker DNA binding activity. Recently a YY1 and Pc-G binding protein has been identified called RYBP (Garcia *et al.* 1999) which may facilitate the recruitment of PcG proteins to the site of YY1 binding.

Similar to *Drosophila*, mammalian PcG proteins regulate Hox genes, but they also have important roles in cell growth regulation. For example, the mammalian Psc homolog, Bmi-1, has oncogenic properties when overexpressed and down-regulates expression from the tumor suppressor locus *ink4a* (Haupt *et al.* 1991; Jacobs *et al.* 1999; van Lohuizen *et al.* 1991). Similarly, RING1, a constituent of a mammalian PcG complex, transforms Rat1a cells and activates *c-fos* (Satijn and Otte 1999). A growth stimulatory role of the mouse Pc protein, M33, is also inferred from genetic studies (Core *et al.* 1997). In contrast, two studies indicate potential growth inhibitory properties of PcG proteins Mel-18 and hPc2 through repression of *c-myc* (Satijn *et al.* 1997; Tetsu *et al.* 1998). Whether these differences in growth regulation are assay-dependent or PcG protein specific remains to be elucidated.

Two models propose how specific targets are chosen for PcG repression

Many studies have addressed the question of how a specific target is chosen for PcG mediated repression. Two models have been proposed to explain the decision. The first model suggests that a temporally and spatially restricted repressor, such as hunchback (hb), acts as a tether by which the PcG complex is directed to a PRE and is positioned for

complex formation. However, hb binding sites are not sufficient for maintenance of repression; additional PRE sequences must be present. Furthermore, hb is not necessary (Poux *et al.* 1996) and indeed is not present in all sites of PcG repression indicating that other proteins might be capable of tethering the complex. We are proposing that Th1 and other transcription factors are capable of this activity via multiple low affinity interactions with multiple PcG proteins.

An alternative model postulates that PcG proteins are able to sense the state of a gene, active or repressed, and only stabilize the repressed state. This information is likely presented as chromatin modifications, the work of repressors on an inactive promoter. Long after the transcription factors have disengaged from the DNA, modifications like methylation and deacetylation would be read by PcG proteins. Interestingly, one putative PcG member, Mi-2, has been shown to interact with hb which supports the tethering model above (Wade *et al.* 1998; Zhang *et al.* 1998). Mi-2 is also a component of a chromatin remodeling complex, therefore a combination of these two models may be employed.

An additional way of maintaining an established "repressive" state is to sequester inactive genes in compartments devoid of activating factors. Inactive loci have been shown to be in association with heterochromatin. In *Drosophila*, the closer a gene is to the heterochromatic center, the more likely it is to be repressed (Csink and Henikoff 1996). In one mammalian study, six B cell specific genes were assayed for heterochromatin association, marked by the Ikaros protein. Repressed genes were found at the Ikaros site while active genes were not (Brown *et al.* 1997). PcG protein association with heterochromatin has also been established in U2OS cells (Saurin *et al.* 1998). This osteosarcoma cell line expresses high levels of most PcG proteins, which are localized to large subnuclear domains. These domains are adjacent to the centromeric heterochromatin evidenced by immunostaining with PcG and kinetochore antibodies. Together these observations suggest that specific domains with a repression preference are established near

heterochromatin and that placing genes within this PcG-charged environment may help maintain a repressed state.

Focusing on Pcl allows the study of the Polycomb-Group mechanism

To further our understanding of vertebrate PcG action, we have chosen a homolog of *Polycomblike* for further analysis. In *Drosophila*, *Pcl* plays an important, perhaps central role in PcG function, since it genetically interacts with many other PcG loci (Landecker *et al.* 1994). Nevertheless, little about its mechanistic role or its protein interaction properties has been ascertained. For example, Pcl has not been included as a component of either of two biochemically distinct PcG complexes (Alkema *et al.* 1997a; Denisenko *et al.* 1998; Gunster *et al.* 1997; Hashimoto *et al.* 1998; Jones *et al.* 1998; Kyba and Brock 1998; Sewalt *et al.* 1998; van Lohuizen *et al.* 1998), despite the observation that Pcl completely co-localizes with Pc and Ph on polytene chromosomes (Lonie *et al.* 1994). Structurally, Pcl contains two (Cys)4-His-(Cys)3 motifs, known as PHD fingers (Aasland *et al.* 1995). PHD fingers are found in many important transcriptional regulators, but their function remains a mystery. Similar to other mammalian gene families, two distinct *Pcl* genes have been reported (Coulson *et al.* 1998; Kawakami *et al.* 1998). Interestingly, one of the vertebrate Pcl proteins has been demonstrated to have sequence-specific DNA binding activity in vitro (Inouye *et al.* 1994).

We have carried out an analysis of Pcl and other PcG proteins in *Xenopus laevis*. The goals were threefold: 1) to compare the phenotypic consequences of Pcl overexpression relative to other PcG proteins 2) to develop a transient developmental assay for PcG function, and 3) to utilize the assay to identify new vertebrate PcG proteins and targets. Chapter Six details the identification of two vertebrate PcG target genes, *En-2* and *Rx2A*, and provides evidence that vertebrate Pcl functions through a PcG mechanism to generate a response characteristic of PcG overexpression in *Xenopus* embryos. Finally, the data

presented in Chapter Seven confirms the power of the transient Polycomb functional assay with identification of two new PcG proteins.

Chapter 2

Materials and Methods

Plasmid construction

All of the following constructs were kindly created by Stan Hollenberg.

Blunt ended fragments. The following restriction enzyme fragments were converted to blunt ends and used for constructions. Where not indicated as otherwise, each encodes full-length protein, with the exception of the following N-terminal truncations: XPc (8 aa), Ets1 (3 aa), and FosB (1 aa). Th1: SV-Th1 (Hollenberg *et al.* 1995) - EcoRI/XhoI. Th1 (forced dimer): The coding sequence of Th1 was amplified by PCR with primers 5' CGGAATTCAAACATGAACCTCGTGGGC 3' and 5' GAATTCGCCTGGTTTAGCTCCAGCGCCCA 3' - EcoRI. Th1(1-160): Th1 coding sequences were amplified as described previously (Hollenberg *et al.* 1995) - EcoRI/StuI. Th1(30-216): Th1 was PCR-amplified with primers 5' GCGAATTCCAACATGATCGATGGCCCGGCCTCG 3' and 5' GCGATCTAGATCACTGGTTTAGCTCCAG 3' - EcoRI/XbaI. mPCL2: KS-mPCL2 - EcoRI/HindIII. In this plasmid EcoRI was placed upstream of the ATG by PCR with primer 5' CGAATTCCACAATGGCGCAGCTCCCCCGG 3'; HindIII follows the polyA tract. mPh2: KS-TE14-25 RACE (see below) was digested with NcoI (partial)/HindIII, end-filled and inserted into the end-filled BamHI site of pBTM116 in an antisense orientation to create pBTM-αmPh2. Fragments were produced with SalI/AvrII (MBP-mPh2) or EcoRI/AvrII (Tet-AU1-mPh2). XPc: CS-XPolycomb (a gift of A.P. Otte) - AflIII/XbaI. mENX-1: pBS-mEnx-1 (a gift of A. Ullrich) - BssHII/DraI. Mel-18: pSG5-Mel18 (a gift of M. Kanno) was subcloned into Bluescript to place a BamHI site directly upstream of the ATG - BamHI/EcoRI. hEts-1: pCMV-hEts1 (a gift of R. Goodman) -

EagI. FosB: pGEM1-AC113 (Zerial *et al.* 1989) - XmnI/EcoRV. E protein: Two hybrid clone TH20 (mouse E12, a.a. 41-153, Genbank D29919) - NotI.

LexA plasmids. pBTM116 was converted to pLex-A by introduction of the ADE2 gene (Chen *et al.* 1996). pLex2-A and pBTM116-2 were created by digesting pLex-A or pBTM116, respectively, with EcoRI, end-filling and re-ligating. Fusion derivatives were constructed from BamHI/end-filled pBTM116 (pLex-Th1, pLex-Th1(1-160), pBTM116-2 (pLex-E), and pLex-A (pLex-Th1(30-216)). pLex-TE was constructed from pLex-A by sequential insertion of Th1 (forced dimer) into SmaI and the E protein sequence into end-filled BamHI.

pVP16 Derivatives: VP16-fusion derivatives were constructed by insertion into the BamHI site (end-filled) of pVP16 or pVP16+ (Mel-18 only). pVP16+ was kindly provided by A. Zahler and pVP16-CREB was a gift of P. Goldman and R. Goodman.

MBP-Fusion Constructs: MBP fusions were generated by introducing cDNAs into the XmnI site of pmal-c2 (New England Biolabs).

CS derivatives: Coding sequences were inserted into the StuI site of CS2+ (a gift of D. Turner) for in vitro transcription/translation.

Tet-AU1 and Derivatives: Tet-AU1 was constructed from CS2+ by replacing the CMV promoter with the Tet-operators from pUHD10-3 (Gossen *et al.* 1995) and inserting the AU1 coding sequence into the BstBI site of the polylinker. Tet-AU1 fusion derivatives were constructed by introducing cDNAs into the end-filled NcoI site (-mPCL2, -XPc, -, -mENX-1), XhoI site (-Ets1, -FosB), or XbaI site (-mPh2, -Mel-18).

Beta actin-Th1: The CMV promoter of pCS2+ was replaced by the rat β-actin promoter (300 nt) to create pβact. The Th1 coding sequence was introduced into the pβact polylinker.

XPcl1 cDNA: *XPcl1* was isolated by hybridization screening of a *Xenopus* Stage 28-30 head library (a kind gift of R. Harland) using a 890 nt probe from the 5' coding region of mPcl1 (ATG to StyI). The single positive from the screen was sequenced and determined

to lack nt 628-723 and contained a premature stop codon at nt 1137 due to a frameshift. Both defects were corrected by replacement with an RT-PCR product from Stage 36 *Xenopus* RNA using primers 5' TGCACACGCTATGCTGCATATGAAACTGTC 3' and 5' GCACAGGCACTGCTATATACATGCTTTTAG 3'. Analysis of several PCR products showed a region of heterogeneity around nt 1126, including the frameshift seen in the original library isolate. A PCR product without a frameshift was chosen and spliced into the original cDNA by digestion with NdeI/EcoRV.

Sequence comparisons with the Genbank non-redundant database were performed using the BLAST algorithm (Altschul *et al.* 1997). Sequences are available for XPcl1 (AF130453) and mPcl1 (U81490) through Genbank.

Plasmid construction for Xenopus injection and in vitro RNA synthesis: All PcG coding sequences were introduced into the plasmid CS2+ (a gift of D. Turner). CS-XPcl1, CS-mPcl1, CS-Pc(M33), CS-mPh2 (Genbank U81491), and CS-XBmi were constructed by PCR with gene specific primers and contain no gene-specific 5'-untranslated information. CS-mPcl1-stop was constructed by PCR using the sense primer 5' GCCAGAATTCCACAATGGCGTAGCTCCCCCGG 3' and contains a premature termination codon. Capped RNA was synthesized with NotI-linearized templates according to Moon and Christian (Moon and Chrisian 1989) using SP6 RNA Polymerase.

Two hybrid screens and assays

Two hybrid screens and assays were performed together with Stan Hollenberg.

S. cerevisiae strain L40 transformed with pLex-Th1(30-216) or pLex-TE was used to screen an E9.5/10.5 mouse embryo cDNA fusion library as described previously (Hollenberg *et al.* 1995). Bait plasmids contained ADE2 and were eliminated by screening for red colonies (Chen *et al.* 1996). Each library positive was then tested in secondary two hybrid assays by mating with derivatives of AMR70 which express distinct LexA-fusions.

Beta-galactosidase assays were performed in doubly transformed L40 cells after permeabilization (Yocum *et al.* 1984). Activity is expressed as (OD_{405} X 1000)/(OD_{595} X reaction time in minutes).

Hybridization screens and RACE

Race was performed by Stan Hollenberg.

Full-length cDNA for mPCL2 and a partial cDNA for mPh2 (TE14/25) were isolated by hybridization screening of an E16 mouse embryo library (Novagen) with ΔH53 and TE14, respectively. The 5' coding portion of mPh2 was isolated by RACE. Nested antisense primers: SH203 5' CTTCTGCAGCTTGCTTCGGTCTGA 3' and SH204 5' CCACTCGTTTGGTGCATCCCACA 3' were used on E14 mouse embryo total RNA with a Marathon cDNA amplification kit (Clontech) according to the manufacturer's instructions. The 2.1 kb amplification product was partially digested with NcoI, completely with NotI and cloned into NcoI/NotI-digested KS-TE14/25 to generate KS-TE14/25 RACE containing a full-length ORF. RT-PCR with SH203 and SH205 5' GAATTCCTCTTCTGCCACCTGAGCACCCT 3' was performed to confirm the correct apposition of the RACE and cDNA isolates. mPCL2 and mPh2 were sequenced by generating nested deletions. Sequence comparisons were performed with the Genbank non-redundant database using the BLAST algorithm (Altschul *et al.* 1990).

Purification of fusion proteins on maltose affinity resin

The following MBP constructs were constructed and assayed by Angela Ingermann.

Maltose binding protein (MBP) fusions were expressed in E. coli strain BL21(DE3)LysE containing the plasmid pUBS520 (Brinkmann *et al.* 1989). Induction of expression, cell lysis, and binding to amylose resin were according to the manufacturer's instructions (New England Biolabs). Resin-bound fusion proteins were stable at 4°C for several months. The

yield and purity of each fusion protein was determined by Bradford assay and 10% SDS-PAGE.

In vitro binding assays

RNA for translation was produced by in vitro transcription of NotI-digested plasmids with either SP6 (CS-Th1, CS-mPCL2, CS-mPh2, CS2-XPc, CS-Enx1, Tet-AU1-mel18, Tet-AU1-FosB) or T7 RNA Polymerase (CMV-hEts1). RNA was translated in vitro with rabbit reticulocyte lysate (Promega) in the presence of S35-methionine. Full-length unfused proteins are Th1, mPCL2, mPh2, XPc, mENX-1, and hEts-1. Mel-18 and FosB were produced as AU1-fusions to improve translational efficiency. MBP-fusion proteins (3 ug) bound to 5 ul maltose resin were placed in 25 ul NETT buffer (100 mM NaCl, 1 mM EDTA, 20 mM Tris (pH 8.0), and 0.2% Triton X-100) containing 50 ug/ml ethidium bromide. Binding reactions were carried out by adding 2.5 ul in vitro-translated protein for 2 hours at 4°C. Reactions were then washed with 3 X 1 ml NETT and resolved by 10% SDS-PAGE. Signals were visualized by autoradiography and quantitated by PhosphorImager (Molecular Dynamics) scanning.

Antibodies and antibody production

MBP-Th1 fusion protein (Hollenberg *et al.* 1995) was mixed with Freund's incomplete adjuvant and injected into rabbits. The serum of rabbit 46-5 (1:500) detects in vitro translated Th1 as a single band after Western blotting. In transfection controls, α46-5 specifically detects Th1-transfected NIH3T3 cells only and identifies a subcellular distribution of AU1-Th1 fusion protein indistinguishable from AU1 antiserum. GST-115 (TFII-I) fusion protein was injected into chickens and antibody recovered from eggs (Avies Lab). ACA human autoimmune serum was a kind gift of EM Tan. Bmi-1 monoclonal antibody was a kind gift of M. van Lohuizen.

Cell culture and transfection

NIH3T3 mouse embryo fibroblasts were cultured on eight well SuperCell culture slides (Erie Scientific Co.) in Dulbecco's modified eagle medium (DMEM, Gibco BRL) with 10% newborn calf serum, 100 units/ml penicillin G, and 100μg/ml streptomycin. Twenty-four hours after plating the cells, each well was transfected with a calcium phosphate co-precipitate containing 100ng of each plasmid DNA: Tet-AU1-fusion vector, pUHG17-1 (Gossen *et al.* 1995), plus or minus βact-Th1. Five hours post-transfection, cells were washed and fed with media containing doxycycline (Sigma) at 0-1000 ng/ml. Pervanadate was added to cells at a concentration of 500 μM for 30 minutes before fixation.

Immunocytochemistry

Thirty hours after transfection cells were fixed with 4% paraformaldehyde in PBS, washed with PBS and allowed to dry. Cells were incubated 30 minutes in blocking solution (PBS with 10% normal sheep serum, 0.1mg/ml bovine serum albumin, and 0.05% Triton X-100). Primary antibodies and secondary antibodies in blocking solution were then added for one hour each at room temperature. Primary antibodies are rabbit polyclonal anti-Th1 α46-5 (1:100) and mouse monoclonal anti-AU1 antibody (1:300; Babco), mouse monoclonal anti-Bmi-1 (1:300), and human auto-immune serum ACA (1:300). Secondary antibodies are goat anti-rabbit IgG-TRITC (1:200; Sigma) and Cy3 (1:300, Jackson ImmunoReseach Laboratories, Inc.), goat anti-mouse IgG-FITC (1:200; PharMingen) and Cy3 (1:300, Jackson ImmunoReseach Laboratories, Inc.). Hoechst 33258 (Molecular Probes) was included at 1.0 μg/ml. Samples were visualized with a Zeiss Axioplan microscope or by confocal laser scanning microscopy on a BioRad MRC 1000. A number of negative controls were performed to verify colocalization results. Cells visualized by immunofluorescence after treatment with primary or secondary antibodies alone were negative. In addition, each cell scanned by confocal microscopy was first excited with the

wavelength for rhodamine or fluorescein and viewed at the emission wavelength for the opposite fluorophore to rule out bleed-through.

Bone and cartilage staining

Staining was performed on wild type and CENP-B null mutant newborn mice (a kind gift of KH Choo) according to the protocols found in Gruss and Kessel, 1991, and van der Lugt et al., 1994 (van der Lugt *et al.* 1994).

RNA quantitation by RT-PCR

RNA quantitation was performed by Yoshino Yoshitake.

XPcl1 cDNA was synthesized with MLV-Reverse Transcriptase using random hexamer primers and 500 ng total RNA. Primers

5' TGCACACGCTATGCTGCATATGAAACTGTC 3' and

5' CCGTTCAGAACGAGGAGTGTC 3' were used for amplification with Taq DNA Polymerase as described previously (Hollenberg *et al.* 1993). Ornithine decarboxylase primers were used as a control (Nastos *et al.* 1998). Cycle number was determined empirically to be in the linear range for each primer pair.

Xenopus embryo culture, manipulation, and phenotypic analysis

Xenopus embryos were obtained, injected and cultured as described previously by Moon and Christian (1989). Embryonic stages were assigned according to Nieuwkoop and Faber, 1967 (Nieuwkoop and Faber 1967). Capped RNA (0.1-1 ng) was injected at the dorsal marginal zone in one or both of the blastomeres. Phenotypes of injected embryos were scored by visual inspection of embryos between Stages 36 to 42. Severity ranged from minor defects, with eyes decreased in size or organization, to major defects with complete loss of normal head structure. All of the reported PcG proteins are capable of

causing a major defect. Numbers reported are for embryos with either minor or major defects.

RNA in situ hybridization and histological analysis

Injection into embryos on only one side included 200 pg beta-galactosidase RNA. Post fixation embryos were assayed for beta-galactosidase with Red-Gal (Research Organics, Inc.) as a substrate (Flowers and Culp 1995) for 45 minutes. Whole mount in situ hybridization was performed as described (Harland 1991) using BM Purple Substrate (Boehringer Mannheim). *En-2* and *Rx2A* repression was scored if the embryo morphology looked normal but the intensity of staining was visually decreased at least two-fold on the side of injection. A change in shape of the *En-2* or *Rx2A* expression pattern was not scored as repression. A shift in the *En-2* expression domain was scored in embryos with normal morphology if *En-2* expression at the midline was offset. After scoring the phenotype or hybridization signal, some embryos were dehydrated and embedded in paraffin. Ten or twenty μm sections were cut and stained with hematoxylin-eosin or eosin alone (Christian and Moon 1993).

Chapter 3

Subcellular localization of Th1 is regulated during development

Controlled subcellular localization is a frequent form of regulation for nuclear proteins and can yield clues about their function. Analysis of subcellular localization of endogenous and transfected Th1 showed three distinct patterns. In addition to the expected diffuse nuclear pattern, Th1 was also cytoplasmic and in a distinctive nuclear speckle pattern. Characterization of the cytoplasmic and nuclear patterns will be reported, here while the nuclear speckle pattern will be discussed in the next chapter.

Ectopic Th1 protein is either strictly nuclear or cytoplasmic

Th1 localization was first characterized in transfected cells. To determine the specificity of the new Th1 antibody I transfected NIH3T3 mouse fibroblasts with both CS-Th1 (driven by the CMV promoter) and CS-Th2 plasmids. The best rabbit, bleed and concentration were chosen based on highest signal and specificity for Th1 alone, and lowest background.

In these experiments, Th1 was toxic at high levels (brightest fluoresence) but at lower levels presented either a cytoplasmic or nuclear pattern. As shown in Figure 1, strong Th1 expression from the CMV promoter (CS-Th1)

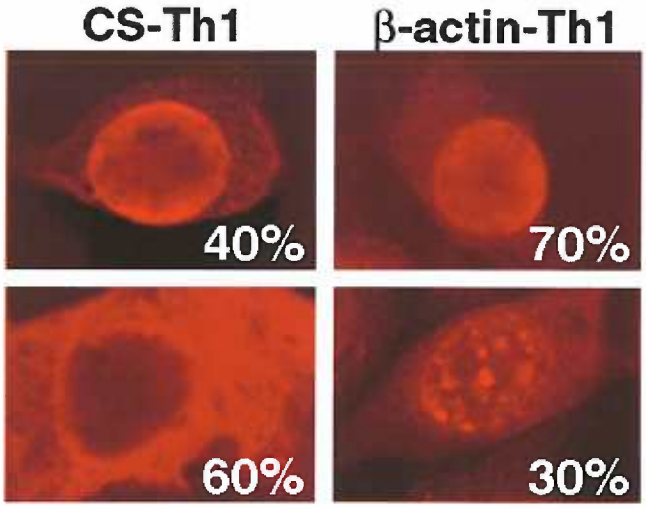

Figure 1 **Th1 subcellular localization is concentration dependent.** NIH3T3 cells were transfected with either CS-Th1 (left panels) or β-actin-Th1 (right panels) followed by immunostaining for Th1 protein. The percent of transfected cells with the pictured localization is reported.

is primarily cytoplasmic (60% of transfected cells). With decreased Th1 expression from the rat ß-actin promoter, Th1 is in the nucleus in 70% of the transfected cells. The remaining 30% of transfectants exhibit a nuclear punctate pattern.

A tetracycline inducible Th1 construct (Tet-Th1) was constructed which allows protein levels to be more closely regulated. When this plasmid was transfected in NIH3T3 cells and the cells were treated with 10 ng/ml doxycycline (a tetracycline analog), a significantly weaker Th1 signal was observed. Figure 2 describes trends in Th1 localization through the first 54 hours of induction. During this time course the concentration of Th1 protein increased steadily based on immunocytochemical inspection. Early in induction (with the lowest levels of Th1) most cells displayed nuclear expression. By 54 hours after induction, levels approximate that of CS-Th1 and expression was split evenly between exclusively nuclear or cytoplasmic localization. A third of the transfected cells were apoptotic. This data suggests that protein levels affect subcellular localization and that Th1 localization is regulated.

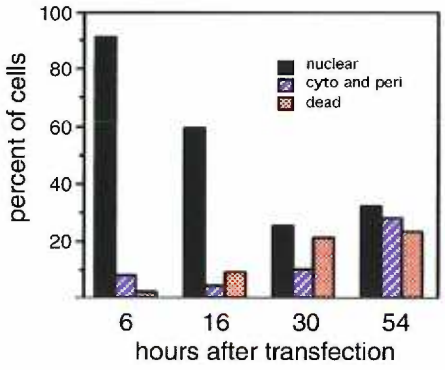

Figure 2 **Cells with ectopic Th1 show exclusively nuclear or cytoplasmic protein expression, or display apoptotic characteristics.** NIH3T3 cells were transfected with tetracycline inducible Th1, treated with tetracycline and assayed for Th1 localization by immunostaining at varying timepoints. The percentage of transfected cells with nuclear or cytoplasmic Th1 or apoptotic characteristics are shown.

A number of experiments follow from this observation: 1) identification of the nuclear localization signal (NLS), 2) determination and manipulation of expression in cell lines with endogenous Th1, and 3) investigation of Th1 expression in sections of developing mouse embryos. This third experiment will provide information about the biologically relevant location of Th1 within the cell. Because we have performed numerous two-hybrid screens with Th1, we have identified protein interactions capable of explaining the divergent subnuclear patterns. Data concerning one such protein will be presented later.

Th1 contains two nuclear localization signals

The NLS of Th1 was identified by tagging Th1 with ß-galactosidase (ß-gal), followed by deletion analysis. Fusing ß-gal to fragments of Th1 provided two benefits. First, this method does not rely on maintaining the Th1 antibody epitope to follow Th1 fragments, and second, location of the fragment can be quickly and easily identified with an enzyme reaction rather than a more lengthy immunocytochemistry procedure.

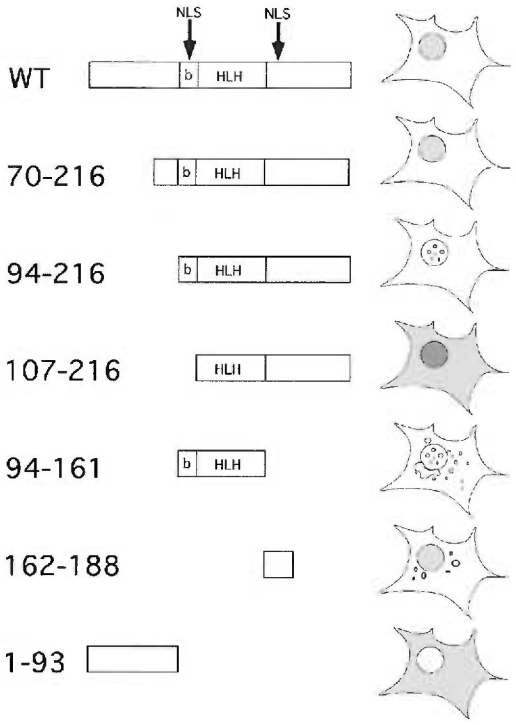

Figure 3 **Th1 contains two nuclear localization signals.** Sequences encoding the indicated amino -acid fragments of Th1 were fused to β-gal, transfected into NIH3T3 cells and assayed for β-gal activity. The subcellular localization of the β-gal-Th1 fusions are represented graphically. The two NLS are indicated by arrows.

Nuclear localization signals are typically stretches of basic residues, therefore we suspected the basic region (also the DNA binding domain) might be the NLS of Th1. In initial experiments Th1 was divided into two pieces: amino acids 1-93 and 94-216 (including the bHLH), and each was fused to ß-gal. ß-gal assays showed that only the bHLH/carboxy-terminal fragment was nuclear and the amino-terminal fragment was not. Further dissection of the bHLH and carboxy-terminal portion of Th1 is reported in Figure 3. Two NLS signals are evident in Th1, neither one being particularly efficient alone, but in combination nuclear transport is complete. One NLS is contained within the basic region (94-161) and the second NLS is just c-terminal to the HLH domain (162-188). The presence of two NLS may allow fine regulation of the extent of nuclear entry of Th1.

Endogenous Th1 localization in Rcho-1 and PC-12 cell lines is regulated

To continue the investigation of Th1 subcellular localization, I chose to use two cell lines derived from in vivo sites of Th1 expression. Rcho-1 cells are a trophoblast cell line isolated from a rat choriocarcinoma tumor, and are capable of undergoing differentiation into the placental cell type, giant cells (Faria and Soares 1991). PC-12 cells are a commonly studied rat pheochromocytoma cell type from an adrenal medulla tumor and undergo a sympathetic neuron like differentiation when exposed to nerve growth factor (NGF) (Tischler and Greene 1975). These two lines are ideal for studying the effects of differentiation on Th1 subcellular localization.

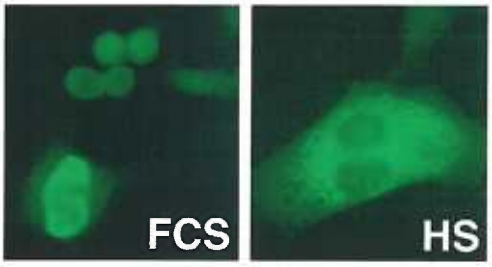

Figure 4 **Th1 is nuclear in undifferentiated Rcho-1 cells but becomes cytoplasmic upon giant cell differentiation.** Rcho-1 cells actively growing in 10% fetal calf serum (FCS) or induced to differentiate in 2.5% horse serum (HS) for four days, were stained with anti-Th1 antibody.

Undifferentiated Rcho-1 cells are small cells with a high ratio of nucleus to cytoplasm. When exposed to low serum conditions (2.5% horse serum) for 48-72 hours, differentiation occurs as evidenced by a dramatic increase in cell size (5-10 fold increase in diameter) and enlargement and fragmentation of the nucleus. I stained these cells with Th1 antibody at different stages of differentiation. Th1 is nuclear in undifferentiated trophoblasts but after exposure to horse serum (48 hrs) a nuclear exclusion is evident (Figure 4 HS) in the giant cells. The nuclear exclusion of Th1 with terminal differentiation suggests that targets may no longer require Th1 in mature giant cells.

PC-12 cells also express Th1 and are easily manipulated to differentiate. After a 24 hour period of serum depletion followed by the addition of 50ng/ml NGF, these small cells begin to put out neurites and express neural markers. Similar to Rcho-1 cells, in PC-12 cells Th1 localization is nuclear in actively dividing cells (Figure 5, serum), but serum removal retains Th1 into the cytoplasm. Addition of NGF decreases Th1 expression, but nuclear expression is reactivated when NGF is removed. NGF removal after

differentiation of PC-12 cells into neurons is analogous to target-derived trophic factor removal in vivo; both events cause apoptosis. Serum removal in NIH3T3 cells transfected with Th1 does not alter Th1 localization. This suggests that additional factors present in cells derived from in vivo sites of Th1 expression are responsible for regulating Th1 localization in response to serum. These data suggest that nuclear exclusion could be a way of regulating transcription of Th1 targets thus allowing differentiation of trophoblasts into giant cells, and PC-12 cells into sympathetic neurons.

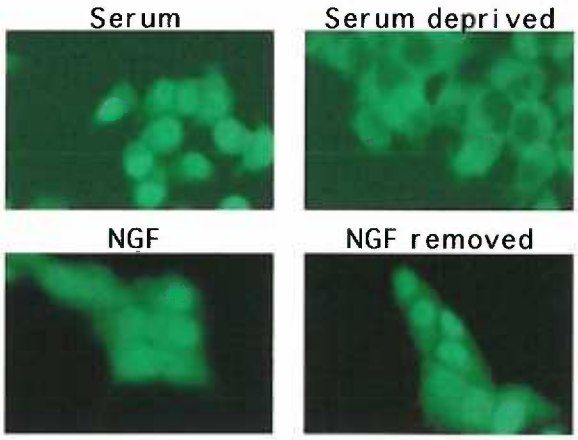

Figure 5 **Th1 is nuclear in undifferentiated PC-12 cells, but becomes cytoplasmic upon serum deprivation. Nuclear Th1 is resumed in apoptotic cells.** PC-12 cells growing under the indicated culture conditions were stained with anti-Th1 antibody. Conditions include: Serum, 10% fetal calf serum; serum deprived, DMEM with 5μg/ml of transferrin and insulin (TI) after 24 hours; NGF, TI plus 50 ng/ml NGF for 48 hours; NGF removed: TI after 24 hours.

Prompted by the discovery that serum deprivation and NGF treatment affected Th1 localization in some cell types, I assayed many other effectors in transfected NIH3T3 cells for this ability. Out of fifteen pharmacological and biological additions only pervanadate affected Th1 localization (data not shown). Pervanadate (500μM) was able to transfer primarily nuclear beta-actin-Th1 (95% nuclear) to a more cytoplasmic pattern (50% nuclear). Pervanadate is a tyrosine phosphatase inhibitor (Kadota *et al.* 1986). This suggests that a decrease in phosphorylation is necessary for the nuclear entry or maintenance of Th1. From these results it is impossible to say whether this is a direct effect on Th1 or an indirect effect on a protein responsible for regulating Th1 entry.

Th1 subcellular localization is regulated in vivo

It was important to verify the cell culture results by looking at Th1 expression in vivo. This was done by staining frozen sections from mouse embryos at different stages of

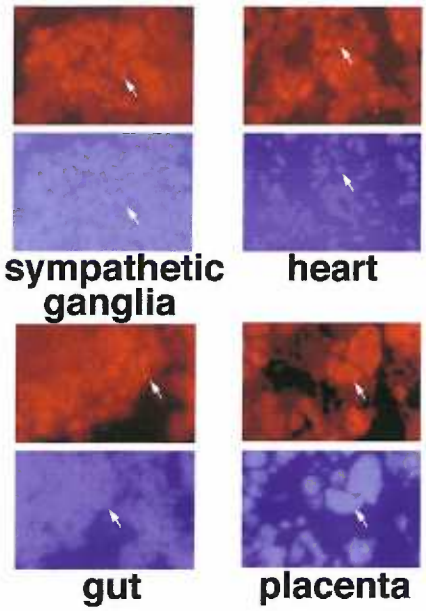

Figure 6 **In vivo Th1 expression is nuclear at e12.5.** Normal mouse embryo frozen sections were stained with anti-Th1 antibody (red) and Hoechst nuclear stain (blue) at e12.5. Representative photographs of sympathetic ganglia, heart, gut and the placenta are shown. Arrows indicate a specific cell that is stained with both anti-Th1 antibody and Hoechst.

development. Although many stages were analyzed, two relevant stages are shown here, embryonic day(e) 12.5 and e13.5. Figure 6 shows representative photographs of Th1 immunostaining at e12.5 within the four sites of Th1 expression, heart, sympathetic ganglia, placenta and gut (Hollenberg *et al.* 1995). Expression of Th1 is exclusively nuclear at this stage suggesting a transcriptional role for Th1 in the development of each of these tissues. However, at e13.5 expression of Th1 in the heart and sympathetic ganglia has become primarily cytoplasmic (Figure 7). This demonstrates that

cytoplasmic Th1 can be measured in both cell culture and in tissue sections. The onset of cytoplasmic Th1 localization in sympathetic ganglia corresponds with

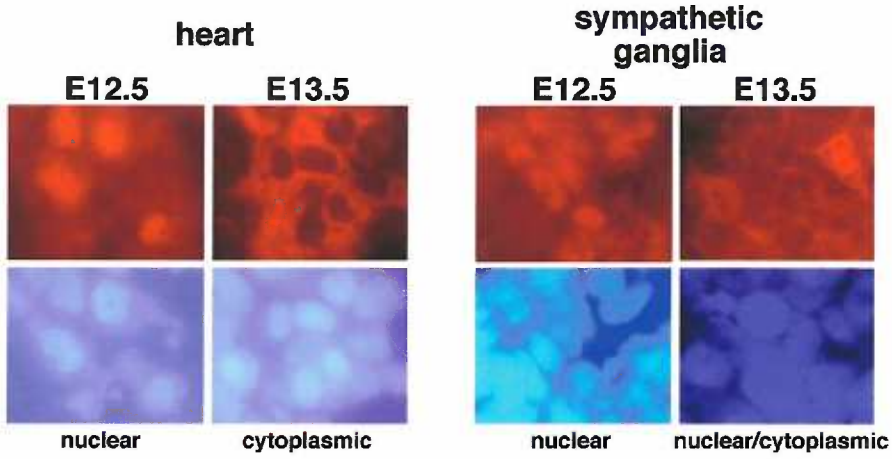

Figure 7 **At embryonic day 12.5, Th1 expression in the heart and sympathetic ganglia is nuclear but becomes cytoplasmic by e13.5.** Mouse embryo frozen sections were stained with anti-Th1 antibody (red) and Hoechst nuclear stain (blue) at e12.5 and e13.5. Representative photographs of sympathetic ganglia and heart are shown.

changes in cell proliferation. Prior to e13.5, sympathetic neurons are actively dividing, followed by a sharp reduction in proliferation. It is possible that nuclear Th1 expression is necessary for proliferation and that its removal from the nucleus is necessary for terminal differentiation of the neurons (as suggested from cytoplasmic transfer of Th1 due to NGF treatment of PC-12 cells (Figure 5). A change in Th1 localization in the heart may signify an important but unidentified event that occurs during heart development.

I-mf-a protein regulates Th1 subcellular localization

One of the successful Th1 two hybrid screens utilized a forced dimer of Th1 and E protein (see Chapter Four). One Th1 partner obtained in this screen was especially interesting, Inhibitor of MyoD family (I-mf-a). I-mf-a has been shown to inhibit nuclear entry of the tissue specific bHLH transcription factors, MyoD, Myf5 and Myogenin in NIH3T3 cells (Chen *et al.* 1996). I-mf-a showed a similar ability to alter Th1 distribution. I-mf-a transfected in conjunction with Th1 in NIH3T3 cells caused a dramatic increase in cytoplasmic Th1, from 5% to 60% (Figure 8). To determine if cytoplasmic transport of Th1 in vivo requires I-mf-a, I obtained I-mf-a null mutant mice (Kraut *et al.* 1998) and looked at Th1 expression in the heart at e13.5. Expression of Th1 in heart is cytoplasmic in both the +/- and -/- mice (data not shown). This suggests that cytoplasmic localization in the developing heart is not absolutely dependent upon expression of I-mf-a. The role of the interaction of these two proteins is still unclear.

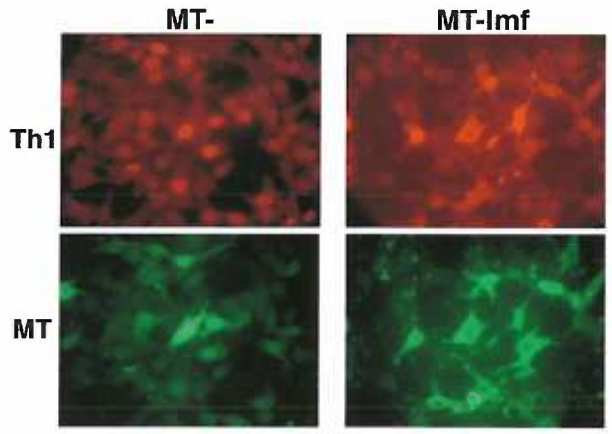

Figure 8 **Transfected Th1 is trapped in the cytoplasm by co-expression of I-mf-a.** NIH3T3 cells were transfected with β-actin-Th1 and either unfused myc-tag (MT) protein, or MT-I-mf-a. Cells were stained with anti-Th1 antibody (red) and anti-MT antibody (green).

Ectopic Th1 induces apoptosis

There is substantial circumstantial evidence that Th1 is involved in apoptosis. 1) High levels of transfected Th1 are toxic in NIH3T3 cells (Figure 1 and 2). Cells stain positive by TUNEL assay (Gavrieli *et al.* 1992), but do not display a DNA ladder (data not shown). 2) I have been unsuccessful in making any stable lines expressing Th1 even though weak and inducible promoters have been employed. 3) Th1 is expressed in the nucleus of PC-12 cells when induced to undergo apoptosis by NGF removal (Figure 5). Deficiencies in apoptosis have not been reported for Th1 null mutant mice (Firulli *et al.* 1998; Riley *et al.* 1998). However, the mutant mice die before apoptosis can be observed in sites of Th1 expression, for example, the sympathetic ganglia.

In summary, the subcellular localization of Th1 is regulated both in ectopically and endogenously expressing cells, and in tissues in vivo. A method for controlling transcriptional events mediated by Th1 may be to export it from the nucleus or capture it in the cytoplasm. As opposed to regulating protein synthesis or degradation, this method could allow rapid entry and exit of Th1 from the nucleus. Regulation may be mediated by protein interactions or by phosphorylation status. Regulated subcellular localization of Th1 during development provides an interesting avenue for further study.

Chapter 4

Th1 interacts with the Polycomb-Group

The yeast two hybrid system provides a method to study the function of a protein by identifying associated proteins (Fields and Song 1989). To apply this strategy to Th1, we have carried out two-hybrid screens with three distinct Th1 baits: the bHLH alone, bHLH plus C-terminus, and full-length, heterodimerized Th1. The isolation of two proteins in the latter two screens is presented here.

The Th1 C-terminus identifies a Polycomblike homologue

Th1 and its close relative, Th2, have a short region of high sequence conservation at their C-termini, with 15 out of 17 residues identical. We reasoned that this sequence conservation is an indicator of an important function. To search for proteins that interact with the C-terminus of Th1, we fused an amino terminal truncation of Th1 to LexA (Lex-Th1(30-216); see Methods). This bait contains both the bHLH and C-terminal region. Lex-Th1(30-216) was used in a yeast two hybrid screen of a mouse embryo cDNA fusion library to identify interacting proteins (Hollenberg *et al.* 1995). Out of 9 x 10^6 clones analyzed, more than 400 tested positive and were further analyzed. The majority of these clones interact with the bHLH domain and encode E proteins. To eliminate these and identify partners which are dependent upon the C-terminal domain, we paired each clone with the C-terminal truncation derivative Lex-Th1(1-160). Only two of the original positives lost the ability to interact. Sequence analysis showed that these clones (ΔH53) contained identical inserts of 360 nt.

Comparison of the encoded peptide with database sequences revealed a similarity to the first of two cysteine- and histidine-rich "PHD" domains (Aasland *et al.* 1995) in the *Drosophila* Polycomblike (dPcl) and trithorax proteins (see Figure 9A). Using the ΔH53 cDNA clone for hybridization we isolated a full-length cDNA encoding a novel protein of 559 amino acids with extensive similarity to dPcl and to a mouse protein, M96 (Inouye *et al.* 1994). This region of similarity extends over 340 amino acids and includes perfect conservation of all cysteines and histidines in both PHD domains (Figure 1A). Based on this extensive structural similarity to dPcl and M96 and interaction with other Polycomb-group homologues (see below), we call this new gene *mPcl1* . The expression of *mPcl1* is widespread, with low, comparable RNA levels in embryo and adult tissues as determined by Northern blot and RT-PCR (data not shown). This relatively ubiquitous expression pattern parallels that seen for *dPcl* (Lonie *et al.* 1994) and suggests that the mPcl1 protein is available for interaction with Th1 in its restricted expression domain during embryogenesis (Cserjesi *et al.* 1995; Hollenberg *et al.* 1995).

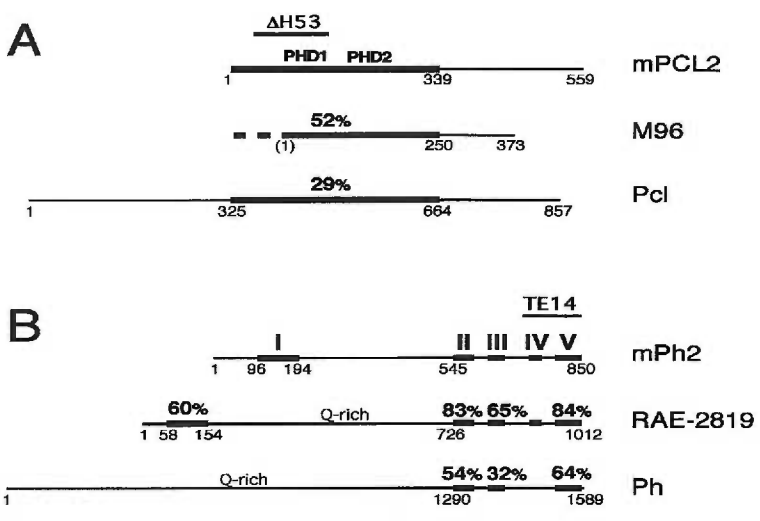

Figure 9. **Comparison of mPcl1 and mPh2 to related proteins.**
A). mPcl1 is compared to mouse M96 (Inouye *et al.* 1994), and *Drosophila* Pcl (Lonie *et al.* 1994). Percentages are sequence identity of the indicated region relative to mPcl1. The conserved Cys4-His-Cys3 PHD1 and PHD2 domains (Aasland *et al.* 1995) and the region encoded by the two hybrid clone ΔH53 are indicated. The dashed region at the amino terminus of M96 is upstream of the proposed initiation codon (Inouye *et al.* 1994).
(B) The deduced amino acid sequence of mPh2 is compared with mouse RAE-2819 (Nomura *et al.* 1994) and *Drosophila* (DeCamillis *et al.* 1992). The amino acids encoded by two hybrid clone TE14 and the highly conserved SPM domain (Bornemann *et al.* 1996) at the C-terminus are indicated.
mPcl1 and mPh2 sequences are available through Genbank Accession numbers U81490 and U81491, respectively.

A Th1-E protein forced dimer identifies a Polyhomeotic gene

A two hybrid screen with an unpaired Th1 bHLH domain is biased. It prevents the identification of partners which require a heterodimer for binding or of under-represented non-bHLH partners that bind to other interfaces. To search for these partners we constructed Lex-TE, a tripartite fusion protein comprised of LexA, full-length Th1, and the bHLH domain from a mouse E protein, E12. In this forced dimer configuration (Neuhold and Wold 1993) the favored interaction should be intramolecular dimerization between the HLH domains of Th1 and E protein. The interaction properties of Lex-TE support this hypothesis: the Lex-TE fusion protein very weakly detects E protein and Id interactions relative to Lex-Th1 and Lex-E, respectively (data not shown).

The Lex-TE protein was used in the two hybrid system to re-screen the mouse embryo cDNA fusion library. Clones were grouped in a secondary analysis according to their preference for Lex-TE, Lex-Th1 and Lex-E. Twenty clones whose interaction with Lex-TE was comparable to or greater than Lex-Th1 were further analyzed. These cDNAs were transferred to a mammalian expression vector and tested for their ability to affect Th1 in two assays: 1) alter the activity of co-transfected Th1 plus E protein on a reporter gene and 2) regulate Th1 sub-cellular localization (see below). One clone, termed TE14, tested strongly positive in both assays (data not shown). Comparison of the 146 amino acid peptide encoded by TE14 with the database reveals high similarity with a mouse polyhomeotic homologue, RAE-2819 (Nomura *et al.* 1994) and *Drosophila* Ph (DeCamillis *et al.* 1992). A combination of hybridization screening with the TE14 cDNA and RACE were used to isolate a full-length open reading frame. As shown in Figure 9B, the encoded protein has extensive similarity with RAE-2819 in the C-terminal 300 amino acids, and lower, but significant identity with *Drosophila* Ph in this region. The region of highest conservation between all three proteins is the SPM domain at the very C-terminus (Bornemann *et al.* 1996), also contained within our two hybrid clone TE14. Since our gene represents the second member of the mammalian polyhomeotic family, we call it

mPh2. The *mPh2* gene is widely expressed in embryos and adult tissues, and in cell lines that express Th1 (data not shown).

Th1 and PcG interact by two hybrid assay

The isolation of two structurally distinct PcG gene products as Th1 binding partners suggested that Th1 might interact with other PcG proteins. To test this hypothesis we constructed activation domain fusions (VP16-) of three vertebrate PcG homologues: a *Xenopus* Polycomb (XPc), (Reijnen *et al.* 1995); ENX-1, a mouse Enhancer of zeste homologue (Hobert *et al.* 1996); and Mel-18 (Tagawa *et al.* 1990). Each was tested for interaction with Lex-Th1 in a yeast two hybrid assay (Figure 10). For comparison, we also assayed unfused LexA, and VP16-fusions of CREB, mPcl1 and mPh2. As expected, VP16-mPcl1 and VP16-mPh2 produce a significant induction in β−gal activity when paired with Lex-Th1, as compared to unfused LexA. VP16-CREB is inactive with either partner. Consistent with a role for Th1 in PcG action, one out of the three additional PcG proteins tested, XPc, also shows an induction of reporter gene activity that is dependent upon Lex-Th1.

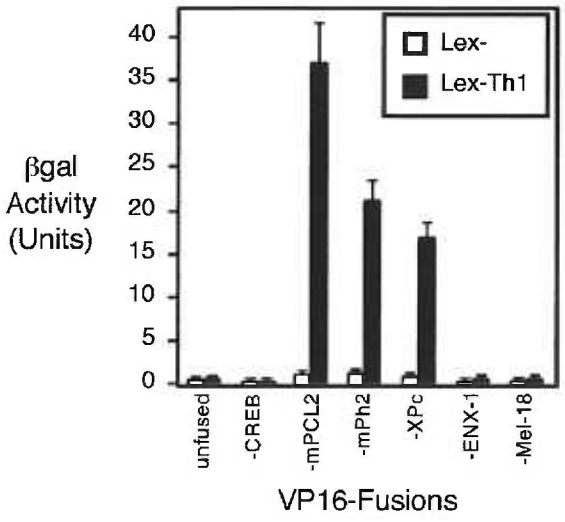

Figure 2-2. **Th1 interacts with mPcl1, mPh2 and XPc in the yeast two hybrid system.** The beta-galactosidase two hybrid signal is shown for pairwise combinations of LexA or Lex-Th1 with VP16-fusions of -CREB, -mPcl1, -XPc, -ENX-1, and Mel-18. Plasmids expressing the indicated pairs of fusion proteins were transformed into yeast strain L40 and assayed for beta-galactosidase activity. VP16-fusions contain full-length sequence with the exception of CREB (aa 1-283), mPh2 (aa 705-850), and XPc (aa 9-521). Values are the mean of three experiments and error bars indicate standard deviation.

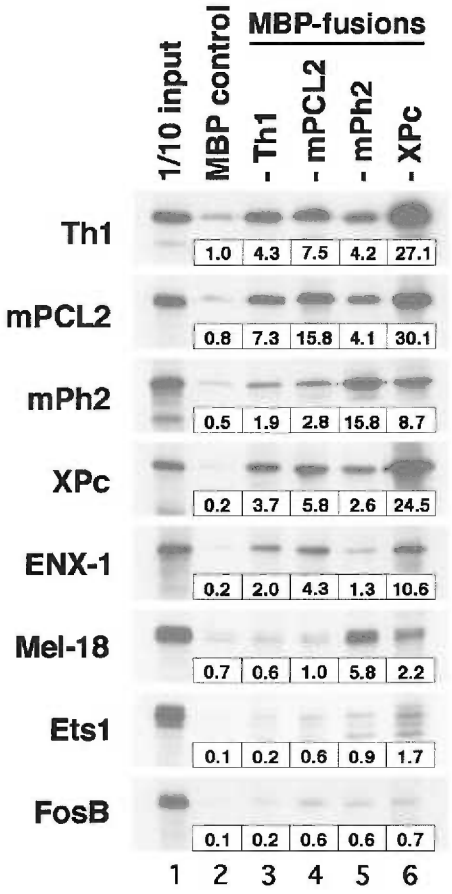

Figure 11. **Th1-PcG and PcG-PcG interactions in vitro.** The proteins indicated at the top were fused in-frame to the malE coding sequence, expressed in E. coli and bound to a maltose affinity resin. Proteins on the vertical axis was translated in vitro with S-35 methionine and tested for binding to each fusion protein. For comparison, lane 1 of each panel shows 10% of total input for each binding reaction. Lane 2 controls for non-specific interactions with MBP or the solid support. The average binding from three experiments (boxes) was determined by Phosphorimager analysis. This data was collected by Angela Ingermann.

Th1-PcG and PcG-PcG interact in vitro

The two-hybrid interactions described above between Th1 and several PcG proteins strongly suggested that Th1 physically binds to PcG members. To provide biochemical evidence for these interactions and to systematically test for other Th1-PcG and PcG-PcG interactions, we have performed a series of in vitro binding assays. Maltose binding protein (MBP) fusions to Th1, mPcl1, mPh2, and XPc were expressed in E. coli, purified and immobilized on maltose affinity resin. Each candidate partner protein was radiolabeled by translation in vitro, and the ability to bind to immobilized fusion protein was quantified. Unfused MBP and in vitro translated Ets1 and FosB serve as negative controls.

As predicted from two hybrid assays, Th1 binds to itself, mPcl1, mPh2, and XPc (Figure 3A). Although not detectable by two hybrid tests, Th1 binding to ENX-1 is also above background. Several PcG proteins also display efficient binding with each other. For example, mPcl1 binds to mPh2 and XPc, and mPh2 binds to XPc and Mel-18. The specificity of these interactions is evident when compared with Ets1 or FosB, which show no binding above background to any of the fusion proteins. This biochemical assay supports our two hybrid interaction

tests and indicates the potential for the formation of multi-protein complexes containing Th1 and PcG proteins.

Th1 is localized to distinct subnuclear domains

We have used indirect immunofluorescence to analyze the subcellular distribution of Th1 in transiently transfected NIH3T3 cells. Three distinct subcellular expression patterns are produced. At high expression levels, when Th1 is controlled by the SV40 or CMV promoter, Th1 protein is predominantly cytoplasmic (Figure 1 and 2). In contrast, when the rat beta-actin promoter directs expression, Th1 protein is detected almost exclusively in the nuclear compartment. Surprisingly, two distinct patterns are evident (Figure 12A). In a typical experiment, 70% of the transfected cells show a uniform Th1 nuclear distribution and the remaining 30% exhibit a punctate expression pattern with concentrated subnuclear localization of Th1. The size and number of detectable Th1 subnuclear domains varies considerably within one transfection, with an average of about 8 and a range from one to greater than 25. Th1 expression domains are phase dark (Figure 12B), with comparable structures not detectable in untransfected cells (data not shown). Thus Th1 is either contributing to and modifying an existing structure or is producing a structure de novo. Hoechst 33258 stains regions of Th1 expression very weakly, thus clearly distinguishing them from the brightly stained chromocenters.

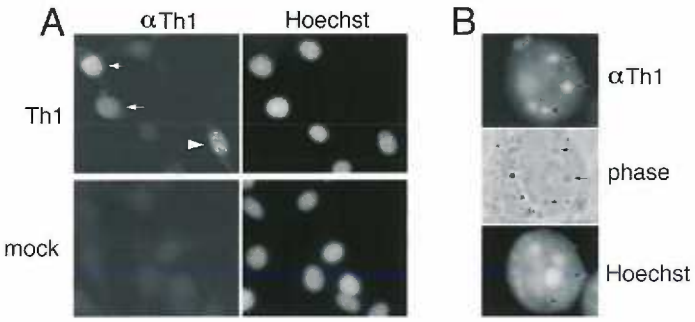

Figure 12. **Transfected Th1 localizes to subnuclear domains.** (A) NIH3T3 cells were transiently transfected with Th1 expression vector (βact-Th1) or were mock-transfected. Th1 protein was detected by α46-5 antibody followed by rhodamine-conjugated secondary antibody. DNA was stained with Hoechst 33258. Note that both punctate (arrowhead) and uniform (small arrows) Th1 nuclear patterns are evident.
(B) A magnified view of a Th1-positive nucleus is shown. The Th1 subnuclear domains (arrows) are phase dark and distinct from the Hoechst-positive regions.

Transiently expressed Th1 and PcG proteins colocalize

The localization of Th1 to discrete subnuclear domains provides a cellular assay for Th1 interaction with potential partners. We have tested the potential of Th1 to recruit co-transfected PcG proteins to subnuclear domains in NIH3T3 cells. Since specific antibodies are not available in most cases, the expression of each candidate partner was monitored through a 9-amino-acid tag (AU1-) fused at the amino terminus. To avoid overexpression artifacts, the level of each fusion protein was controlled by doxycycline regulation of a reverse tetracycline-controlled transactivator (Gossen *et al.* 1995). After co-transfection, the subcellular localization of fusion protein and Th1 was determined by indirect immunofluorescence with a confocal microscope (see Methods).

Table 1. Distribution of Transiently Expressed PcG and Control Proteins (-/+ Th1) in NIH 3T3 Cells

	-Th1	+Th1			
		Nuclear Partner Pattern with Punctate Th1 (%, n = 150 cells)			
Partner	Punctate Nuclear (%, n = 500 cells)	Colocalized Punctate	Neighbor Punctate	Distinct Punctate	Uniform Staining
mPCL2	0	81	2	0	17
mPh2	0	74	5	0	21
XPc	100	34	53	13	0
ENX-1	1.9	49	0	1	50
Mel-18	0	48	2	1	49
Ets1	0	0	0	0	100
FosB	0	0	0	0	100

n values are the sum of three independent experiments.

Figure 13 **Th1 recruits PcG proteins.** NIH3T3 cells were co-transfected with β–actin-Th1 and the PcG proteins and control proteins listed above. Transfected cells containing Th1 in a nuclear punctate pattern were counted and grouped according to distribution of the PcG protein (labeled by immunostaining against an AU1 tag).

Three of the PcG proteins tested (mPcl1, mPh2, and Mel-18) do not produce a punctate pattern when expressed in the absence of Th1, and mENX-1 rarely does (Figure 14A, column a; Figure 13). Instead they produce a relatively uniform nuclear (mPh2, mPcl1, mENX-1) or nuclear plus cytoplasmic distribution (Mel-18). When co-expressed, Th1 has the ability to efficiently alter the subnuclear distribution of these proteins by concentrating them into punctate patterns that perfectly overlap with its own (Figure 14A, columns b-d; Figure 13). The two-hybrid partners mPcl1 and mPh2 give almost 80% co-localization when Th1 distribution is punctate, and ENX-1 and Mel-18 produce this co-localization in about half the cells.

In contrast to the other four PcG proteins tested, XPc expressed alone produces a punctate staining pattern similar to that of Th1 (Figure 14B, e; Figure 13). Every cell with detectable XPc expression shows a punctate pattern, therefore we cannot demonstrate a re-distribution of XPc protein in response to Th1 expression. Nevertheless, co-expression of Th1 and XPc produces co-localization one-third of the time in cells with a Th1 punctate pattern (Figure 14B, panel f; Figure 13), consistent with the results described above. In addition to co-localization with XPc and uniform nuclear expression, Th1 creates two other distinct patterns (Figure 14B and Figure 13): Th1 and XPc are found in similar, but slightly offset patterns (panel g; "neighbor punctate") and unrelated patterns (panel h; "distinct punctate"). All patterns are found in one transfection and are not correlated with protein expression levels. When co-expressed with Th1, three of the other proteins, mPcl1, mPh2, and Mel-18 occasionally produce the neighbor punctate pattern (Figure 13).

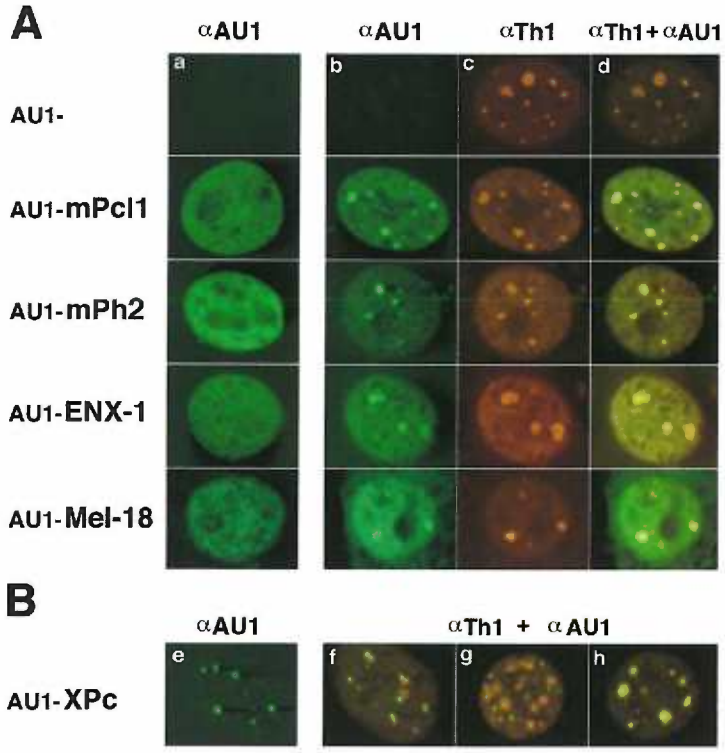

Figure 14 **Th1 directs PcG proteins to its subnuclear domains.** (A) PcG fusions proteins localized by Th1.
(B) XPc produces a subnuclear pattern by itself (e). Co-expression of Th1 produces co-localized (f), neighbor (g) and distinct (h) punctate patterns.
Nuclei of NIH3T3 cells expressing AU1-fusions alone (column a and e) or in the presence of Th1 are shown (columns b-d, f-h). Proteins were detected with the indicated antibodies and fluorescent-conjugated secondaries (see Methods) by scanning laser confocal microscopy. Pseudocolors are: AU1-fusions (red) and Th1 (green). In columns d and panels f-h, the red and green images were overlaid to show regions of co-localization (yellow). NIH3T3 cells were transiently transfected with plasmids encoding AU1-fusions, reverse Tet activator, and Th1. Cells were treated with doxycycline and fixed for immunocytochemistry 30 hours later. Doxycycline concentrations (ng/ml) are: mPh2 (0); m-ENX-1, Ets1, FosB, Mel-18 (10); mPcl1 (100); and XPc and unfused (1000)

Co-localization with Th1 is a very specific phenomenon. We have co-expressed Th1 with seven epitope-tagged transcription factors (Arix, Ets1, FosB, JunB, Xbra, Xgsc, XlBox6; Figure 13 and data not shown). None of these proteins is found in subnuclear domains when expressed alone and none ever co-localizes with Th1 to its subnuclear pattern. These controls underscore the specificity and significance of our observations.

In conclusion, the PcG represents a second class of proteins obtained by two hybrid screen and has become a major focus in our lab. The ability of Th1 to redistribute these proteins within the cell suggests an important interaction, but the nature of the PcG complex makes proving direct interaction difficult. Immunoprecipitation of Th1 and PcG members has been attempted by several means but has been difficult to for technical reasons. Others have demonstrated that PcG interactions are only detectable with endogenous protein levels (Schoorlemmer *et al.* 1997). Further study of PcG interactions requires a functional assay system in an in vivo environment. We have engineered such a system which will be presented in Chapter Six.

Chapter 5

Characterization of additional Th1 partners, CENP-B and TFII-I

In this chapter I characterize a pair of Th1 two-hybrid interaction partners that are strong repressors and colocalize with Th1, but have not been shown to be PcG members. These proteins are CENP-B and TFII-I.

Centromere binding protein-B (CENP-B) was identified as the antigen recognized by autoimmune sera from scleroderma patients (Earnshaw and Rothfield 1985; Muro *et al.* 1990). Staining with CENP-B antiserum reveals a nuclear punctate pattern coincident with the centromere (Sugimoto *et al.* 1994; Yoda *et al.* 1992). Dimers of CENP-B have been shown to bind to a 17-base pair motif called alphoid repeats which flank the centromere (Masumoto *et al.* 1989). CENP-B has been proposed to contribute to condensation of the DNA at the centromere by producing bends through simultaneous binding to distinct alphoid sequences (Muro *et al.* 1992). Despite these observations, CENP-B null mutant mice display no mitotic or meiotic effects suggesting that CENP-B plays no role in chromosome condensation and migration during these events, or its role can be replaced by a protein with redundant function (Hudson *et al.* 1998). However, the CENP-B mutant mice do show a 30% decrease in body weight. Defining the function of CENP-B is still an active area of research.

The second repressor protein obtained by two-hybrid interaction with Th1 is TFII-I. cDNAs for TFII-I have been independently cloned by three individual groups, but it was first described as an initiator binding protein (Roy *et al.* 1991). Ananda Roy generated a large body of work detailing TFII-I dependent transcription from the adenovirus major late core promoter. TFIIA is required for binding of TFIID to the TATA box, initiating transcription. However, TFII-I can allow initiation of transcription in a TATA-less context by binding the initiator (Inr) and forming a complex with TATA binding protein and the

activator, upstream stimulatory factor (USF) (Roy *et al.* 1993; Roy *et al.* 1991). TFII-I and USF are immunologically related and act cooperatively through either the Inr or the USF site (E box) (Roy *et al.* 1991). Roy further showed that when TFII-I was bound by Myc transcription was inhibited (Roy *et al.* 1993), suggesting that TFII-I could provide a switch between on and off transcriptional states. Grueneberg and coworkers identified TFII-I (called SPIN for SRF-Phox1 Interacting protein) through its effects on the *c-fos* promoter (Grueneberg *et al.* 1997). TFII-I stabilizes interactions between serum response factor (SRF) and the homeodomain protein, Phox1, an interaction necessary for *c-fos* activation. Activation of the ras pathway is also required, likely through phosphorylation of TFII-I (Kim *et al.* 1998). A third group cloned TFII-I, termed BAP-135 for Bruton's tyrosine kinase associated protein of 135 kDa, when identifying proteins phosphorylated in response to B cell receptor engagement (Yang and Desiderio 1997). Each of these groups describe the ability of TFII-I to be phosphorylated, and a structure comprised of six novel, but similar 90 amino acid repeats.

Description of strategy

We confirmed the interactions observed in yeast by performing a mammalian interaction assay. By removing the need for specific DNA binding, much can be learned about a protein with suspected transcriptional abilities. This is achieved by fusing the Gal4 DNA binding domain (Gal-) or LexA DNA binding domain (Lex-) to the full length protein of interest and transfecting it in conjunction with a synthetic reporter gene with multimerized Gal4 and LexA DNA binding sites (L8G5-LUC). This strategy was used to test both the interaction strength of two hybrid partners (mammalian two-hybrid) and their transcriptional preference for either activation or repression of a reporter gene.

In the mammalian two hybrid assay, the bait is tethered to the reporter DNA by the Gal4 DNA binding domain and co-transfected with the interacting fragment of the partner which is fused to an efficient transcriptional activation domain from VP16 (-VP16). Interaction strength is measured by the ability of the protein interaction to bring the activator to the reporter gene (see Figure 15). Although this assay is very similar to the system in yeast, surprisingly a large number of partners are excluded because interactions are not maintained in a mammalian cell context.

The transcriptional preference assay can determine whether a protein will activate or repress transcription when tethered to DNA. To measure activation, a Gal4 fusion is expressed in the presence of L8G5-LUC which is silent in the cell line we use (NIH3T3). Activation is measured by an increase in luciferase activity. To measure repression, L8G5-LUC can be activated by co-transfection of Lex-VP16, which can be reversed in the presence of a Gal-fusion protein with repressor activity (see Figure 15).

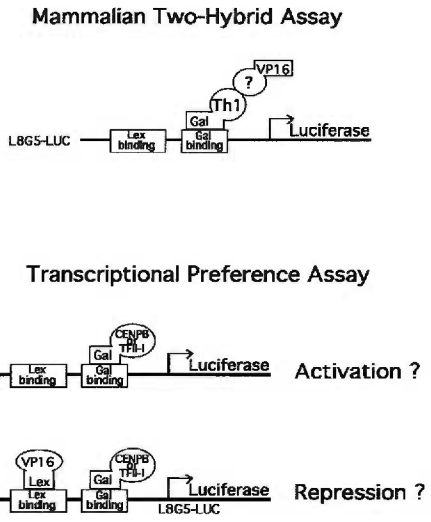

Figure 15 **Strategy for mammalian two-hybrid and transcriptional preference assays.** The L8G5-LUC reporter gene has eight multimerized Lex DNA binding sites (L8) and five Gal4 DNA binding sites (G5) controlling expression of luciferase. This construct is used for both assays. Fusion proteins with either Gal4 or LexA DNA binding domains are co-transfected with the reporter constructs. The reporter can be activated by co-transfection of fusion proteins with the VP16 activation domain.

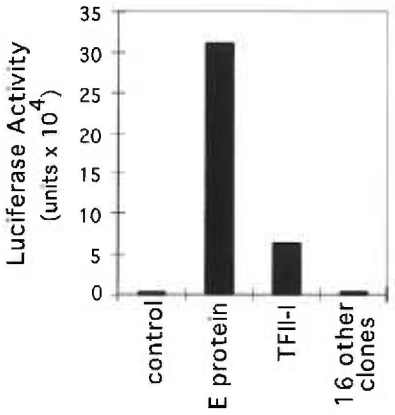

Figure 16. **TFII-I interacts with Th1 in the mammalian two hybrid assay.** The mammalian two hybrid assay was performed in NIH3T3 cells by co-transfection of VP16 fusion proteins of E protein, TFII-I, and CENP-B fragments and Gal-Th1, followed by a luciferase assay. Note that TFII-I interacts with Th1 but 16 other yeast positives do not.

TFII-I interacts in the mammalian two hybrid

To test if TFII-I interacts with Th1 in both yeast and mammalian cells the mammalian two hybrid assay is employed (Figure 16). Gal-Th1 has little or no effect on

the reporter gene, but when co-expressed with TFII-I-VP16, transcription is activated 25 fold. Indeed, out of 16 unique partners obtained in this screen only the interaction with TFII-I is above background. Th1 partner. E protein, is used as a positive control and shows a strong interaction in the mammalian system.

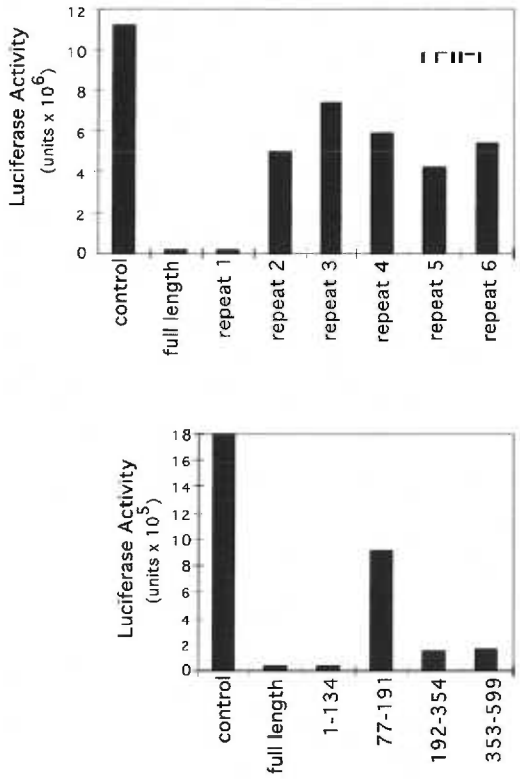

Figure 17 **Both TFII-I and CENP-B are repressors, each with a defined repression domain.** The transcriptional preference assay was performed on NIH3T3 cells transfected with Gal fusion of either full length or fragments of TFII-I (top) or CENP-B (bottom). These constructs were tested in the context of activated transcription of L8G5-LUC for the ability to repress transcription. Repeat one of TFII-I and amino acids 77-191 of CENP-B mediate the strongest repression. Note that the carboxy terminal portions of CENP-B also have some repression activity.

CENP-B and TFII-I are repressors

Full length CENP-B and TFII-I were fused to Gal4, and assayed for transcriptional preference. Both proteins are strong repressors in this assay, capable of repressing Lex-VP16 activated transcription about 50 fold (Figure 17 full lengths). This result requires Gal4 fusion and is not affected by co-transfection of Th1 (data not shown). This level of repression is remarkable, indicating that both transcriptional inhibition and co-transfection of numerous plasmids is very efficient. The ability of CENP-B to repress is not surprising considering its association with the centromere, a transcriptionally silent region. However, as TFII-I has been shown previously to be an activator of transcription (Roy *et al.* 1993; Roy *et al.* 1991), TFII-I mediated repression supports the theory that TFII-I may be a transcriptional switch.

The repression domains of these two proteins were dissected out by fragmenting the cDNA and fusing it to Gal4, much like the ß-gal fusions used in determining the NLS.

The ability to repress is conferred by the first repeat (190/251) of TFII-I (Figure 17). However, CENP-B has multiple repression domains. The strongest is in the amino-terminal fragment (1-134) and is likely within the first 77 amino acids (Figure 17). The carboxy-terminal portion of the protein also has repression activity. The specificity of this activity made it possible to perform very direct two hybrid screens using only the strongest repression domains of each protein in hopes of identifying the mechanism for repression. Surprisingly, screens with the individual domains produced a common partner, NKtarp (Natural Killer tumor antigen related protein (Genbank U82939). NKtarp is most closely related to the herpes immediate early viral gene, ICPO implicated in criticle viral transcriptional regulation and contains a ring finger domain in the amino-terminus, a domain found in three other PcG proteins (Kliewer *et al.* 1989). Therefore, an ICPO-related interaction partner, NKtarp, links CENP-B and TFII-I by a currently unknown repression mechanism.

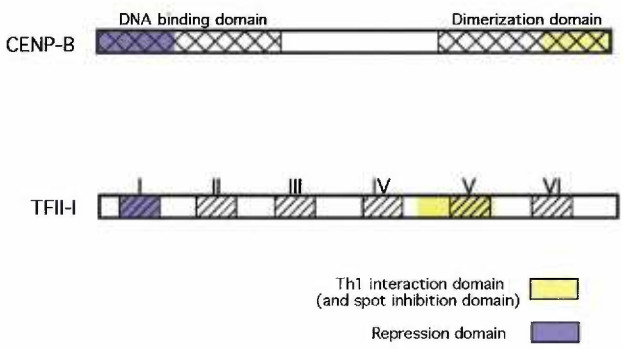

Figure 18 **TFII-I and CENP-B protein structure.** The structure of TFII-I and CENP-B is represented including the Th1 interaction domain and spot inhibition domain(yellow block), and the strongest repression domain(blue block) are indicated. The six repeats of TFII-I are shown as hatched blocks, and the DNA binding domain and dimerization domains of CENP-B are indicated by checkered blocks.

Figure 18 depicts the protein structures of CENP-B and TFII-I. The fragments of CENP-B and TFII-I obtained by two hybrid screen inhibit Th1 spot formation (see Chapter Four) in transfected NIH3T3 cells (indicated by the yellow block). The strongest repression domains are indicated by blue blocks. The CENP-B protein has two known motifs, a DNA binding domain at the amino terminus and a dimerization domain at the carboxy-terminus (Yoda *et al.* 1992). The DNA binding domain is necessary for CENP-B spot formation (data not shown). The six repeats within TFII-I are numbered.

CENP-B and TFII-I colocalize with Th1

TFII-I, endogenous and transfected in NIH3T3 cells, is strictly nuclear in a uniform distribution (data not shown). Transfection of two Th1 constructs change the endogenous and ectopic TFII-I pattern. Cytoplasmic Th1, a product of the CS-Th1 plasmid, pulls TFII-I into the cytoplasm (data not shown) and beta-actin-Th1 transfection recruits TFII-I to the nuclear punctate pattern (Figure 19A). TFII-I is the only protein we have identified that follows Th1 to each of its locations. Together with the exceptionally strong mammalian two-hybrid interaction, this data suggests a high affinity interaction between these two proteins.

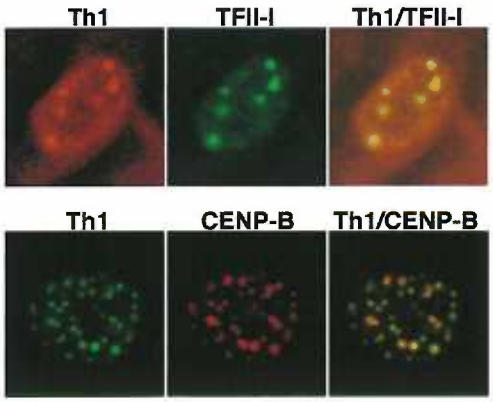

Figure 19 **Ectopic TFII-I and CENP-B colocalize with transfected Th1.** NIH3T3 cells were transiently transfected with β-actin-Th1 and AU1 fusions of either A)TFII-I or B)CENP-B. The cells were fixed and stained with anti-Th1 antibody and anti-AU1 antibody. A rhodamine secondary antibody (red) recognizes Th1 in panel A and AU1-CENP-B in panel B. Likewise, a fluorescein secondary antibody (green) recognizes AU1-TFII-I in panel A and Th1 in panel B. The third frame in each panel is the overlay of the two staining patterns indicating co-localization of the transfected proteins.

When expressed in NIH3T3 cells, CENP-B has a nuclear punctate subcellular localization, much like beta-actin-Th1. In fact, the two proteins efficiently colocalize in co-transfection experiments (Figure 19B). CENP-B colocalization was limited to exogenous levels as a specific antibody was unavailable. It is interesting to note that a Th1 construct with GFP fused to the carboxy-terminus never formed a nuclear punctate pattern until transfected with CENP-B (data not shown), further evidence of a tight interaction. The subnuclear Th1-CENP-B expression domain may be a site for shuttling proteins to functional locations, possibly the centromeric heterochromatin, for long-term repression.

A logical follow-up experiment for each of these studies would be co-immunoprecipitation with Th1. This can be achieved when we obtain specific, high affinity antibodies against TFII-I and CENP-B.

CENP-B has characteristics of a PcG protein

Additional experiments characterizing CENP-B localization link it to the PcG. Co-transfection of CENP-B with the XPc produces overlapping punctate patterns of the two proteins (Figure 20A). I addressed the possibility that CENP-B is capable of recruiting endogenous PcG members by staining for Bmi-1 protein in NIH3T3 cells. This protein has a very fine granular pattern within the nucleus (see bottom cell in figure 3-6B). However, when CENP-B is over-expressed, it concentrates Bmi-1 into a pattern of larger spots staining positive for CENP-B (see top cell in figure 20B). In contrast to the Bmi-1 pattern in NIH3T3 cells, in U2OS cells, PcG proteins, including Bmi-1, are expressed at very high levels in a distictive subnuclear pattern (Alkema *et al.* 1997). In this cell line, transfected CENP-B colocalizes with Bmi-1 (data not shown). Staining with ACA (human auto-immune serum against the centromere) indicates that the endogenous Bmi-1 location is "neighbor punctate" to a subset of the centromeres (figure 20C). The existence of two domains with endogenous protein levels, but a single domain in transfected cells suggests that high levels of CENP-B can drive the association of the separate compartments. Questions concerning CENP-B and PcG colocalization will only be answered when a specific CENP-B antibody becomes available.

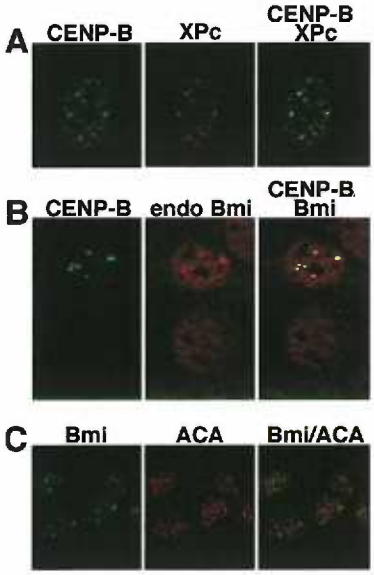

Figure 20 **CENP-B colocalizes with and recruits PcG members.** A) NIH3T3 cells were transiently transfected with GFP-CENP-B and and AU1-XPc. The cells were fixed and stained with anti-AU1 antibody and rhodamine secondary antibody (red). Note that CENP-B colocalizes with XPc in a nuclear punctate pattern. B) NIH3T3 cells were transfected with GFP-CENP-B. The top cell is transfected (evidenced by green GFP) and the bottom cell is not. Anti-Bmi-1 antibody (rhodamine secondary) recognizes endogenous protein. Note that CENP-B recruits Bmi-1 to its compartment. C) Endogenous Bmi-1 and the centromere (ACA antibody) are in adjacent compartments.The third frame in each panel is the overlay of the two staining patterns indicating co-localization of the transfected proteins.

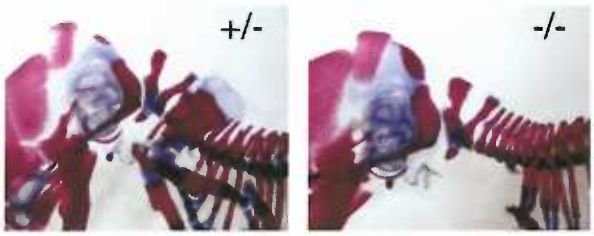

Figure 21 **CENP-B null mutant mice do not exhibit axial homeotic mutations.** Newborn CENP-B null mutant (-/-) and heterozygous +/-) mice were stained for bone (red) and cartilage (blue). A representative photograph of the cervical and cranial region is shown.

Recruitment of exogenous and endogenous PcG members, and a strong ability to repress transcription indicate a possible link between CENP-B and the PcG. When CENP-B null mutant mice (Hudson *et al.* 1998) became available, I assayed littermates for homeotic transformation, a phenotype predicted in PcG mutants. Whole-mount staining of bone and cartilage allows the analysis of cervical vertebra which are easily distinguishable as C1-7. Figure 21 shows +/- and -/- littermates with bone and cartilage stain. The two embryos are indistinguishable indicating that there is no homeotic transformation of the vertebra due to loss of CENP-B protein. This does not completely rule out PcG association. Mutation of some *Drosophila* PcGs have very weak or no effects in *Drosophila* until combined with other PcG mutations, producing a synergistic increase in phenotype.

This class of Th1 interacting partners requires additional study. We have concentrated on CENP-B and TFII-I because of their strong two hybrid interactions with Th1, recruitment by Th1, and Th1 spot inhibition, as well as a strong ability to repress transcription. It is possible that Th1 forms a mechanistic connection with two types of repression, one mediated by the PcG complex and a second mediated by CENP-B and TFII-I effects, conceivably through Nktarp.

To further our analysis of PcG repression, we focused on creating a transient biological assay for PcG function. The experiments were initially performed with Pcl alone, but were verified with three other known PcG proteins. This assay is detailed in the next Chapter.

Chapter 6

Overexpression of PcG proteins causes a specific neural phenotype and affects specific targets in Xenopus

Isolation of a *Xenopus Polycomblike* homolog, *XPcl1*

To initiate an analysis of *polycomblike* in *Xenopus*, we screened a *Xenopus* stage 28-30 head library by hybridization with a mouse *Polycomblike 1* cDNA (Chapter Four). A full-length coding cDNA, termed *XPcl1*, was obtained (see Materials and Methods). Sequence comparison between XPcl1 and other Polycomblike sequences (Figure 22) shows the highest similarity with mPcl1/TcTex3/PHF1 (Coulson *et al.* 1998; Kawakami *et al.* 1998). The N-terminal portions of the proteins, which include the two PHD fingers, are 70% identical. Two additional short blocks of identity, regions II and III, are found in the more divergent C-terminal regions of the proteins. M96A, a second mammalian Polycomblike, is also closely related to XPcl1, but shows less identity in the PHD fingers and region II is not conserved. In contrast, *Drosophila* Pcl is much more divergent, with only 31% identity in the PHD finger regions; regions II and III are absent. Closer analysis of this sequence comparison shows that PHD finger 1 is highly divergent between vertebrates and *Drosophila*, whereas the PHD finger 2 region shows about 50% identity over a region of 100 amino acids (not shown). Others have observed a similar limited sequence conservation between *Drosophila* and vertebrate PcG proteins, such as Polycomb and Polyhomeotic (Nomura *et al.* 1994; Pearce *et al.* 1992).

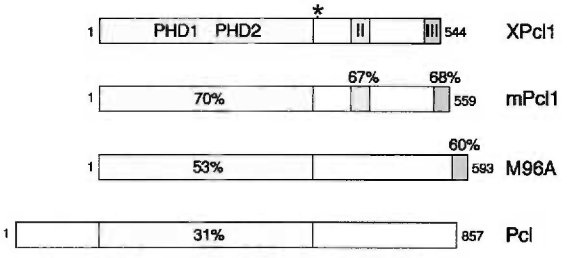

Figure 22 **Sequence comparison of Polycomblike from *Xenopus*, mammals and *Drosophila*.** XPcl1 is compared to mouse Pcl1 (TLH and SMH, unpubl.), M96A (Inouye *et al.* 1994), and *Drosophila* Pcl (Lonie *et al.* 1994). Percentages are amino acid sequence identity of the indicated region relative to XPcl1. The conserved Cys4-His-Cys3 PHD1 and PHD2 domains (Aasland *et al.* 1995), homology regions II and III, and the region of microheterogeneity in XPcl1 (*) are indicated.
Genbank accession numbers are XPcl1 (AF130453) and mPcl1 (U81490). This data was collected by Stan Hollenberg.

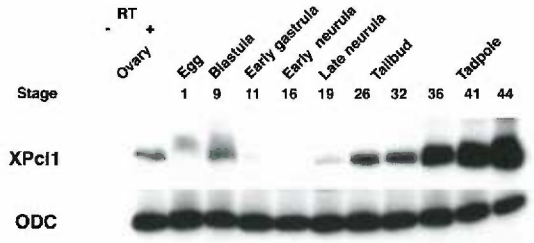

Figure 23 **RT-PCR quantitation of XPcl1 expression during Xenopus embryogenesis.** XPcl1 mRNA levels were quantified using reverse transcription followed by PCR with gene-specific primers. Primers specific for Xenopus Ornithine Decarboxylase (ODC) were used as control to correct for differences in RNA levels and reverse transcription efficiency. Developmental stage is indicated (**Nieuwkoop and Faber 1967**). Cycle number and product sizes are: XPcl1 (27 cycles, 436 bp) and ODC (32 cycles, 234 bp). This data was collected by Yoshino Yoshitake.

XPcl1 is expressed in a spatially-restricted pattern late in embryogenesis

We have quantified the levels of *XPcl1* mRNA throughout embryogenesis by RT-PCR using *XPcl1* gene-specific primers (Figure 23). *XPcl1* RNA is detectable in the zygote as a maternal message. After a decline in levels during gastrulation *XPcl1* RNA levels begin to increase and become significantly higher by tailbud stage. Steady state levels further increase during tadpole stages. This pattern indicates that *XPcl1* RNA has maternal and zygotic components which are temporally similar to *Xenopus* PcG genes *XPolycomb* and *XBmi-1* (Reijnen *et al.* 1995).

Whole mount RNA in situ hybridization was used to determine the spatial distribution of *XPcl1* RNA during embryonic stages. Consistent with our RT-PCR analysis showing low levels early in development, the *XPcl1* RNA signal was very weak at stage 20, but became readily detectable by Stage 26 (Figure 24A). Stage 33 embryos show *XPcl1* RNA expression

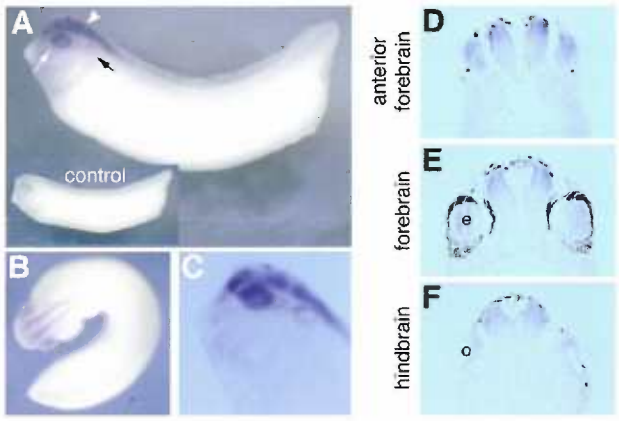

Figure 24 ***XPcl1* is expressed in the anterior central nervous system.** XPcl1 antisense probe was used for whole mount RNA in situ hybridization. XPcl1 expression is restricted to the brain (white arrowhead), eyes (white arrow) and otic vesicles (black arrow) as shown at stage 33, lateral view (A,C), and in a dorsal view (B) at stage 26. Inset of A shows embryo hybridized with sense XPcl1 probe. The embryo in C has been cleared and shows apparent gaps of expression in the brain. Transverse sections of stage 37 embryos after XPcl1 detection are also shown (D-F). Signal is present in the brain, eyes (e), and otic vesicles (o). This data was collected by Yoshino Yoshitake.

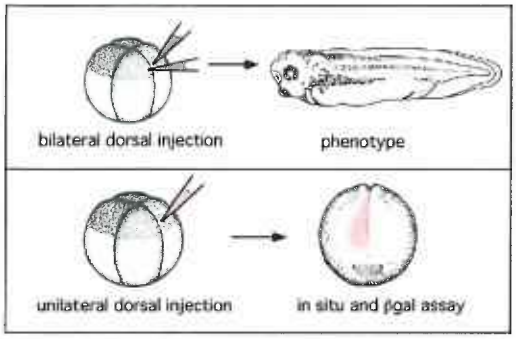

Figure 25 **Strategy for Xenopus assay system.** Upper panel; Bilateral dorsal injection at the four cell stage is followed by visual inspection of stage 36-40 embryos (4 days later) for gross phenotypes. Lower panel: Unilateral dorsal injection at the four cell stage followed by β–gal assay (red product) to determine side of injection and in situ hybridization at stage 19-22 (two days later).

predominantly in the anterior central nervous system including the developing eye. Within the developing brain the *XPcl1* signal is distinctly non-uniform at this stage. As shown in Figure 24C, gaps of significantly weaker signal are visible within the head of a Stage 33 embryo. Transverse sections of *XPcl1*-stained embryos confirm the whole mount signal localization and show signal throughout the brain, eyes and otic vesicles (Figure 24D-F).

Overexpression of XPcl1 in *Xenopus* produces anterior central nervous system defects

Xenopus is commonly used for measuring developmental consequences due to overexpression of an injected RNA (see Figure 25). We used this strategy to measure the phenotypic consequence of early, widespread Pcl1 expression. Preliminary experiments compared the effects of injecting capped RNA transcripts, synthesized from either mouse or *Xenopus Pcl1*. Identical phenotypes were produced regardless of the species origin of *Pcl1*, although mouse *Pcl1* was somewhat more potent (Figure 29). Marginal zone injections into either both dorsal or both ventral blastomeres at the four-cell stage resulted in distinct phenotypes, as expected from dorsal-anterior or ventral-posterior localization, respectively, of injected RNA. Ventrally-injected embryos develop with a partial second axis. In addition some show a tail

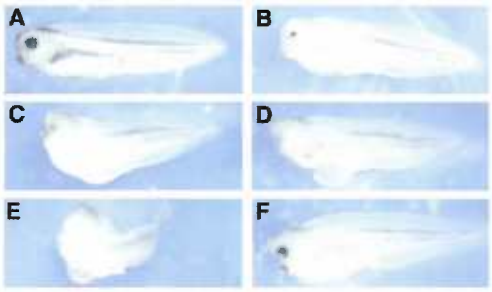

Figure 26 **Phenotypic consequences of Pcl1 overexpression.** Embryos were symmetrically injected in both blastomeres near the dorsal marginal zone at the four-cell stage with 0.4 ng RNA. Control (A) and mPcl1 (B-F) embryos are shown at stage 40. Eyes are small (B,C), absent (D,E) and occasionally fused (F) and embryos show characteristic changes in head shape relative to controls.

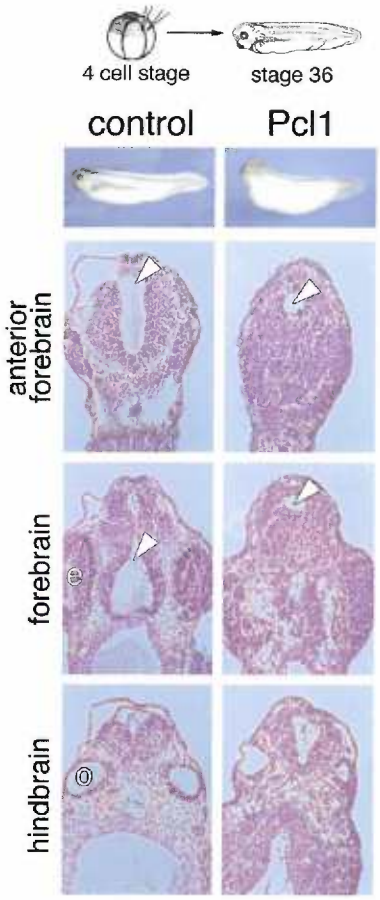

Figure 27 **Analysis of the Pcl1 overexpression phenotype**. Four cell embryos were bilaterally injected with control (beta-galactosidase) or mPcl1 RNA, sectioned at stage 36 (transverse), and stained with hematoxylin-eosin. Note how Pcl1 markedly decreases the size of the ventricle (v) and expands the surrounding neural tissue. The position of the eye (e) and otic vesicle (o) are indicated.

duplication (data not shown). These observations will be published in detail elsewhere. As described below, dorsal injection of Pcl1 also generated a reproducible phenotype and was used for all subsequent experiments.

Dorsal misexpression of Pcl1 had a striking effect on development of the anterior nervous system. Embryos consistently exhibited abnormal head structures, visible in Figure 26 (B-F) as a dorsal bulge and an indentation anterior to the eyes. In addition, the eyes were reduced in size, completely absent, or occasionally fused. The severity of the phenotype was dose-dependent, with 20, 41, and 70% of embryos exhibiting a phenotype at 200, 400 and 1000 pg injected RNA, respectively. To localize the developmental defects, we analyzed Pcl1-injected embryos for correct muscle and nerve formation, using immunostaining. Both muscle and notochord are present and relatively normal (YY data not shown), indicating the effect is localized to the nervous system. To define the phenotype in more detail, Pcl1-injected embryos were sectioned and the gross morphology of the anterior nervous system was examined histologically. These embryos consistently showed dramatic alterations in the structure of the forebrain (Figure 27). Neural tissue is clearly present into the most anterior regions as demonstrated by *N-CAM* (pan-neural) whole mount in situ hybridization (Figure 28), but it appears very disorganized with the normal columnar appearance disrupted (Figure 27). The overall mass of neural tissue is markedly increased at the expense of the ventricle (v), which is virtually absent. The

normally-thin roof and floor plates are greatly thickened and disorganized. The phenotype is graded along the anterior-posterior axis of the brain: the effects are less severe in the midbrain, and the hindbrain appears relatively normal (Figure 27).

The strong effects on forebrain structure seen in Pcl1-injected embryos do not result from gross alterations in anterior-posterior patterning. We have used whole mount RNA in situ hybridization with *N-CAM*, *Rx2A*, *Xotx2*, *X-dll3*, and *Xwnt-8b* antisense probes (Blitz and Cho 1995; Cui *et al.* 1995; Kablar *et al.* 1996; Kintner and Melton 1987; Mathers *et al.* 1997; Pannese *et al.* 1995; Papalopulu and Kintner 1993) to define the changes within the anterior central nervous system (CNS) mediated by Pcl1 overexpression (Figure 28). At stage 19-22, expression of *Rx2A* and *Xotx2* is absent in the developing eye (arrowheads) of embryos made to misexpress Pcl1, although expression of *Xotx2* in the forebrain (arrows) is maintained. Expression of *N-CAM* in the eyes (arrowheads) of Pcl1-misexpressing embryos is completely normal, suggesting that Pcl1 doesn't alter formation of the early eye field, but affects a later step in eye development. At this stage Pcl1 produces no significant alteration of the forebrain pattern of *Xotx2*, *X-dll3*, and *Xwnt-8b*. At stage 25-26, *Xotx2*, *Xwnt-8b*, and *X-dll3* are

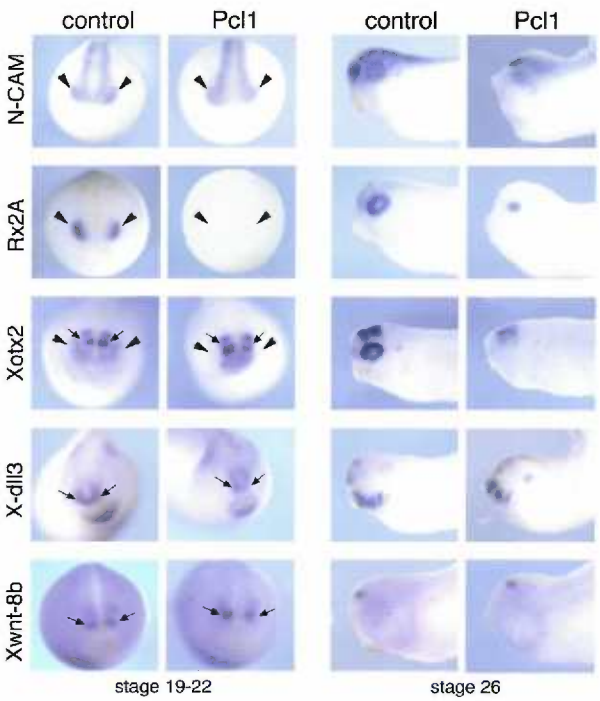

Figure 28 **Pcl1 effects on brain and eye markers.** Control (beta-galactosidase) or mPcl1 RNA was injected bilaterally into the dorsal marginal zone of four-cell stage embryos. Expression of N-CAM (pan-neural), Rx2A (eye), Xotx2 (eye, forebrain, midbrain), X-dll3 (forebrain, otic vesicle, branchial arches, cement gland), and Xwnt8-b (forebrain-midbrain boundary) was assayed by whole mount RNA in situ hybridization. Representative embryos are shown from stage 19-22 (left) and stage 26 (right). Developing eye field is denoted by arrowheads, forebrain-specific gene expression by arrows. Note that neural patterning is unchanged at stage 19-22, even when eye markers are completely absent.

still expressed at their correct positions within the forebrain, but their signals are diminished in some Pcl1-injected embryos (Figure 28, *Xotx2*). These analyses demonstrate that Pcl1 expression does not alter early specification of the anterior nervous system, but clear structural defects become evident at later stages.

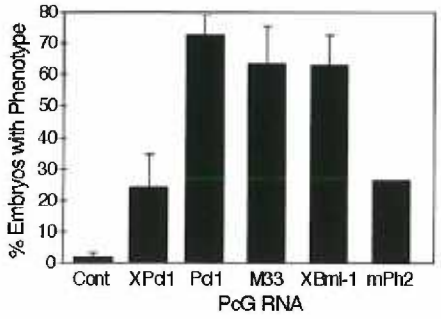

Figure 29 **PcG proteins produce similar anterior neural defects.** Embryos were injected bilaterally at the four-cell stage with 1 ng of each RNA. At stages 36-42 embryos were scored for the distinctive changes in head morphology and eye size shown in Figure 4. At least 30 embryos were analyzed for every condition in each experiment. Values are the average of three or more separate experiments. Error bars show s.e.m. Injected RNAs are: control (beta-galactosidase), Xenopus and mouse Pcl1, M33 (mouse Polycomb1), XBmi-1, and mPh2 (mouse polyhomeotic 2).

PcG proteins generate similar phenotypes in *Xenopus*

We have tested the phenotypic consequences of injecting three other PcG RNAs in the *Xenopus* assay system, focusing on anterior nervous system defects. Four-cell embryos were symmetrically-injected in both dorsal blastomeres with RNA encoding either M33 (mouse Polycomb 1; (Pearce *et al.* 1992), XBmi-1 (Reijnen *et al.* 1995), or mPh2 (mouse Polyhomeotic2; (Chapter Four). All RNAs showed similar defects in nervous system development, with about 60% of the embryos exhibiting changes in anterior brain and eye development (Figure 29 and 30). This level of activity is very similar to that produced by mouse *Pcl1* RNA. We have sectioned embryos from each PcG injection and see changes in anterior neural organization similar to Pcl1

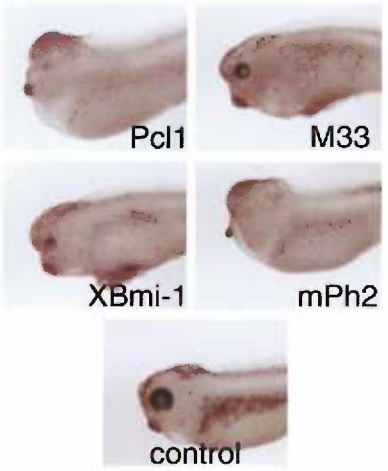

Figure 30 **PcG proteins produce a similar gross phenotype.** Embryos were injected bilaterally at the four-cell stage with 1 ng of each RNA. At stages 36-42 embryos were scored for the distinctive changes in head morphology and eye size shown in Figure 4. Injected RNAs are: control (beta-galactosidase), mouse Pcl1, M33, XBmi-1, and mPh2.

overexpression (data not shown). We conclude that the Pcl1 phenotype is common to several PcG proteins and is therefore a consequence of PcG function.

XPcl1 represses *En-2*

The segment polarity gene *engrailed* is a direct PcG target gene in *Drosophila* (Strutt and Paro 1997). One of its vertebrate homologs, *En-2*, is thus a candidate target gene. Indeed, analysis of the Pcl1 overexpression phenotype by whole mount RNA in situ hybridization showed that *En-2* was repressed in some bilaterally-injected embryos (data not shown). However, genetic studies in mice suggest that *En-2* repression is unlikely to be responsible for the neural phenotype we observe (Joyner *et al.* 1991).

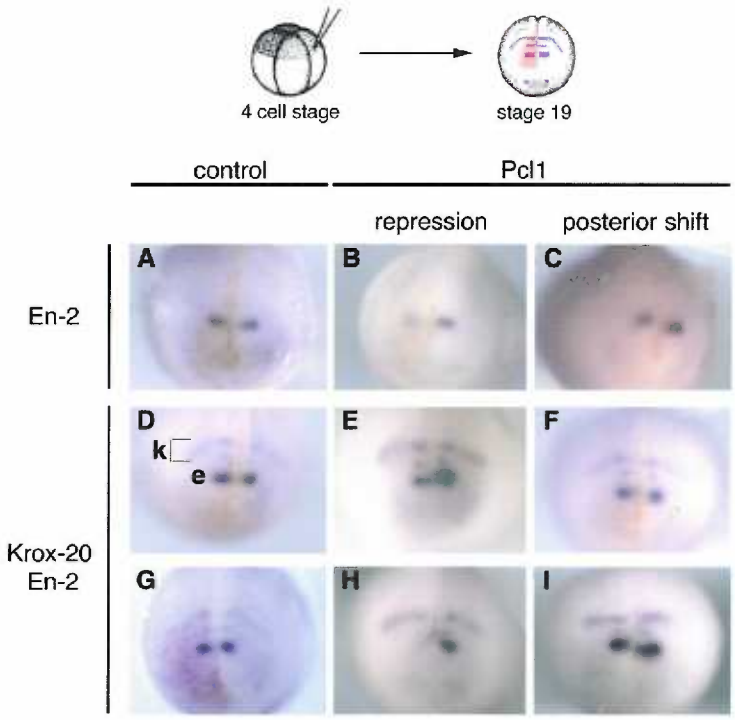

Figure 31 **Pcl1 alters expression of En-2 and Krox-20.** Expression of En-2 (midbrain-hindbrain boundary) and Krox-20 (hindbrain; rhombomeres 3 and 5) were assayed by whole mount RNA in situ hybridization at stage 19. Embryos were injected unilaterally at the four-cell stage with either control (GFP; A,D,G) or mPcl1 RNA, plus RNA encoding N-beta-galactosidase. Beta-galactosidase assay with Red-gal (red product) marks the side of injection (on the left in each panel). Repression of En-2 is dependent upon Pcl1 (B), but uncoupled from repression of Krox-20 (E,H). Posterior shifts of En-2 are also Pcl1-

To provide a more accurate assessment of *En-2* changes, embryos were assayed for *En-2* expression after injection into only one of the two dorsal blastomeres at the four-cell stage. This eliminates the difficulty of comparing control and experimental *En-2* levels and sites of expression in different embryos. The side of injection was marked with RNA encoding a nuclear beta-galactosidase. In situ beta-galactosidase assay with the

chromogenic substrate Red-Gal (see Figure 25 and Materials and Methods) provided an indication of both the side of injection and the targeted region. In control embryos *En-2* is expressed in a characteristic stripe at the midbrain-hindbrain boundary (Davis *et al.* 1991; Hemmati-Brivanlou *et al.* 1991). Injection of control RNA, encoding beta-galactosidase, GFP, or a mutant Pcl1 (mPcl1-stop), produces no consistent change in the *En-2* pattern (Figure 31 A,D,G, Fig 32 and data not shown). In contrast, both mPcl1 and XPcl1 significantly reduce the *En-2* expression level on the injected side in about 15% of the embryos at stage 19-22 (Figure 31B and 32). This repression is gene-specific; double-labeling with probes for both *En-2* and *Krox-20*, which marks rhombomere 3 and 5 in the hindbrain (Bradley *et al.* 1993), shows repression of the *En-2* signal without any change in levels of *Krox-20* (Figure 31E). Figure 31H shows a rare example of complete En-2 repression which locally affects *Krox-20* only at rhombomere 3, possibly an indirect consequence of midbrain-hindbrain alterations (Joyner 1996). The specificity of *En-2* repression is supported by the observation that *Xotx2*, *Xwnt-8b*, *X-dll3*, and *N-CAM* are not repressed by Pcl1 at the same stage (Figure 28).

XPcl1 alters the site of *En-2* expression

In addition to repression, injection of *XPcl1* RNA has the potential to shift the site of *En-2* expression posteriorly (Figure 31 C,F,I). This effect is frequently independent of *En-2* repression, but always occurs on the side of RNA injection. Both mouse and *Xenopus* Pcl1 produce a similar effect, with mPcl1 showing greater efficiency (18% versus 8%, Figure 32). This effect is very rare in control-injected embryos, detectable in less than 0.5 % of embryos.

Changes in the position of expression along the anterior-posterior axis are not specific to *En-2*. For example, *Krox-20* expression shifts in a posterior direction along with *En-2* (Figure 31F). We also detect very localized posterior shifts affecting only *En-2*

or only one of the two *Krox-20* stripes (Figure 31I), suggesting that we are locally rather than globally changing positional identity along the axis.

Other PcG proteins alter *En-2* similar to XPcl1

Based on analyses in *Drosophila*, PcG proteins regulate overlapping sets of target genes. To determine whether *En-2* is a common target of other PcG proteins in this assay, we measured changes in *En-2* expression levels and position mediated by injection of RNA encoding M33, XBmi-1, and mPh2. Figure 32 shows that expression of each of these three PcG proteins produces significant repression of the *En-2* gene, and all show similar efficiencies. In addition, all three PcG proteins generate a posterior shift in *En-2* expression indistinguishable from that caused by Pcl1 (Figure 32). Therefore, four vertebrate PcG proteins generate similar effects on both development of the nervous system and the *En-2* expression pattern.

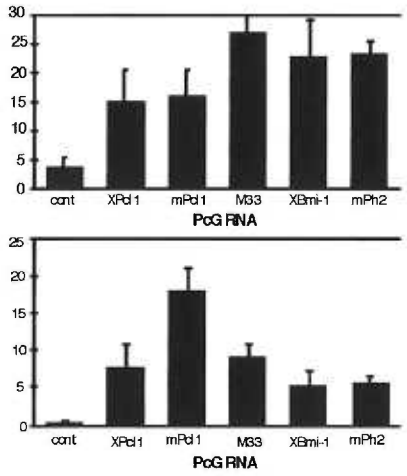

Figure 32 **PcGs produce similar effects on En-2 expression.** Embryos were injected unilaterally and assayed for En-2 repression and position shift as in Figure 8. The percentage of embryos that show the effect and the standard error are given. At least 50 embryos.

Development of a transient vertebrate assay system for PcG biological function allows analysis of potential PcG members obtained by two-hybrid screens. Chapter Seven describes two new PcG proteins identified by this method.

Chapter 7

Identification of two new Polycomb-Group proteins by the Xenopus assay system

Two hybrid screens with mPcl1 fusions yielded two proteins of interest, BS69 and PRSM. The full power of the *Xenopus* assay system for PcG function was demonstrated in the analysis of these proteins. The results of those studies are reported here.

BS69 is a Polycomb-Group protein

Positives from mouse Pcl1 two hybrid screens were subjected to a secondary screen by requiring interaction with *Drosophila* Polycomblike. This identified partners that might have fundamental roles in PcG function as opposed to proteins with functions specific for silencing in mammalian cells. Overall, approximately 30% of the positives from each screen, including most of the strongest positives, corresponded to C-terminal regions of BS69. BS69 was identified previously as a nuclear E1A partner with the ability to block E1A-transactivation (Hateboer *et al.* 1995). Of note, BS69 contains a PHD finger and a bromodomain, domains frequently contained within chromatin regulatory proteins.

The ability of BS69 to repress was tested in the transcriptional preference assay (see Figure 15). Relative to the Gal4 DNA binding domain alone, a Gal4-BS69 fusion reduces luciferase activity, indicating that BS69 has an intrinsic repression activity. Interestingly, Gal4-Pcl1 shows little or no repression activity in this assay. In contrast, wild-type BS69, when co-expressed with Gal4-Pcl1, has the capability to confer repression activity on the fusion protein (Hong Ma, data not shown). This is evidence for in vivo association of BS69 and Pcl1.

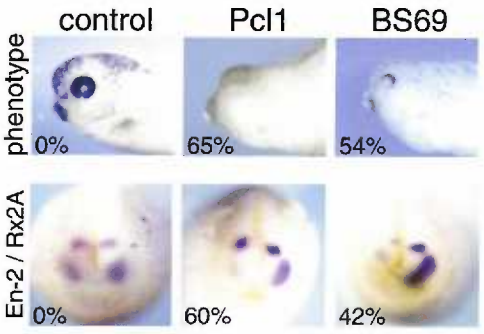

Figure 33 **BS69 overexpression in *Xenopus* embryos causes a PcG overexpression phenotype.** In vitro transcribed capped RNA encoding BS69 or β–gal protein was injected at the dorsal marginal zone of four cell embryos. Embryos were allowed to develop to stage 40 and assayed for head phenotype, or stage 20 and assayed by in situ hybridization for changes in En-2 and Rx2A expression. Percentages of injected embryos showing either a phenotype (top panels) or change (shift or repression) in En-2 or Rx2A message (bottom panels) are indicated in the bottom left of the frame.

The ability of BS69 to bind Pcl1 in vivo and confer repression activity strongly implicates BS69 in the PcG silencing process. To test this hypothesis in a developmental context, I injected BS69 transcripts bilaterally (see Figure 25). As predicted from our in vitro observations, misexpression of BS69 produces effects on *Xenopus* development which are indistinguishable from those produced by Pcl1. The characteristic head morphology produced by PcG overexpression is clearly visible (Figure 33). Sectioning of the mutant embryos revealed disruption and mass increase of the neural tissue as observed in other PcG effected embryos (Figure 34) In addition, the two neural markers repressed by other PcGs, *En-2* and *Rx2A*, are also strongly affected (Figure 33). In the example shown, the posterior shift of *En-2* is also visible. Therefore, BS69 expression alters neural development, represses *En-2* and *Rx2A*, and shifts *En-2* posteriorly, all of which are indicators of PcG action. We conclude that BS69 has the capability to produce PcG function.

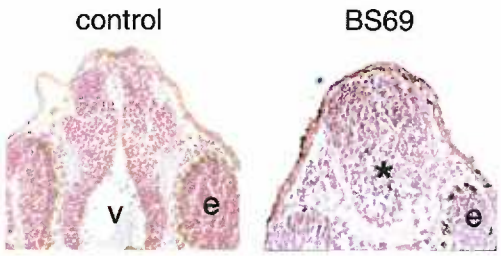

Figure 34 **Sections of a BS69 injected embryo reveal increased and disorganized neural tissue.** In vitro transcribed capped RNA encoding BS69 or β–gal protein was injected at the dorsal marginal zone of four cell *Xenopus* embryos. Embryos were allowed to develop to stage 40. A representative wild type embryo and an embryo with a moderate BS69 phenotype were embedded and cross-sectioned, and stained with Hematoxylin-Eosin.

PRSM is a Polycomb-Group protein

PRSM was the only protein with equivelent interactions with both mouse and *Drosophila* Pcl in the two-hybrid screen. A cDNA containing PRSM has been published (Genbank AAC67541) but represents a

distinct reading frame that completely encompasses our PRSM cDNA. Database searches reveal that our version of PRSM has homologs in *S. cerevisiae, S. pombe, C. elegans* and plants. In *dictyostelium* the PRSM homolog is necessary for stalk differentiatation during development (Genbank AAC67541). This developmental role for PRSM made it an interesting candidate for injection in the *Xenopus* PcG assay system.

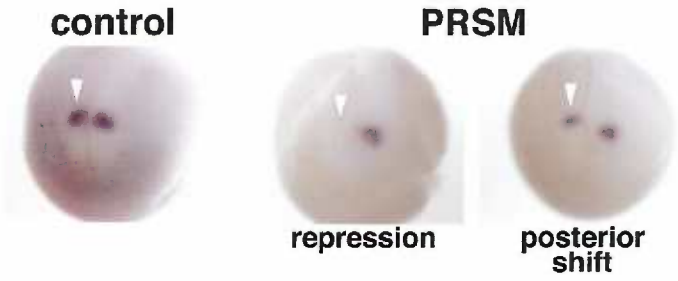

Figure 35 **PRSM overexpression shifts and represses *En-2*.** In vitro transcribed capped RNA encoding PRSM or β–gal protein was injected at the dorsal marginal zone of four cell embryos. Embryos were allowed to develop to stage 20 and assayed by in situ hybridization for changes in En-2 expression. Arrows indicate side of injection.

PRSM overexpression produces an effect on head structure similar to BS69 in developing *Xenopus* embryos, while mediating striking effects on *En-2* expression (Figure 35). A large number of embryos diplayed repressed and shifted *En-2* in the presence of excess PRSM (75% and 19% of injected embryos, respectively). In addition, the degree of posterior shift is often greater than in Pcl embryos, reaching the equivalent of a two rhombomere shift in a posterior direction. This event is rarely seen in Pcl injected embryos.

PRSM repression is apparently not stable. *En-2* expression is reported in Chapter Six at stage 19-22. When later stages (stage 24-26) are assayed in PcG-injected embryos the percentage of affected embryos remains the same. In contrast, PRSM injected embryos assayed at stage 26 no longer exhibit *En-2* effects (data not shown). It is assumed that RNA degradation is a significant factor by this stage due to the progressive loss of ß-gal activity, therefore it is possible that PRSM RNA is no longer available. This result is intriguing because of the proposed role of PcG in maintaining silencing. It suggests that PRSM has a fundamental role in stabilizing the complex.

Both BS69 and PRSM produce a phenotype indistinguishable from the four other PcG proteins assayed (see Chapter Six). We have, therefore, identified two new PcG proteins by the *Xenopus* assay system.

Chapter 8

Discussion

Nuclear localization of Th1 is regulated

Th1 has two nuclear localization signals, suggesting tight regulation of Th1's subcellular localization. Th1 may be excluded from the nucleus by masking of the NLSs. A mask could be applied by interaction with an inhibitory protein, like I-mf-a. Protein interactions or phosphorylation might change the conformation of Th1 to mask the NLS within the folded protein. Finally, masking may occur by a combination of the two former processes.

There are two alternative ways Th1 might be trapped in the cytoplasm other than masking of the NLS: cytoplasmic retention, and active nuclear export. It is not likely that Th1 is retained in the cytoplasm by a cytoplasmic tether. If a tether was employed, increasing levels of Th1 expression would overwhelm the ability of the tether to maintain cytoplasmic retention. Just the opposite is true, with increased Th1 protein levels in culture, Th1 is more likely to be cytoplasmic (Figure 1). This points to another possibility. A nuclear export signal may tag Th1 for active nuclear export. Th1 would appear nuclear when too much Th1 protein swamped the export machinery and backed up in the nucleus. However, there would likely be some cytoplasmic Th1 that had been exported. In general Th1 is localized strictly to the nucleus or strictly to the cytoplasm. Furthermore, inspection of Th1 protein structure did not identify a consensus nuclear export signal. These two observations make the existence of a simple nuclear export signal unlikely. The same theory of masking could be applied to modulate the activity of an export signal. Again, transfection experiments where high levels of Th1 are cytoplasmic discourage this theory. An endogenous masking protein would be the limiting factor in Th1's localization and high Th1 protein levels would swamp out the ability of the mask to inhibit the export signal, thus

trapping Th1 in the nucleus. The exclusion of the processes of cytoplasmic retention by tethering or active nuclear export for controlling Th1 subcellular localization suggests that Th1's movements are controlled via regulation of the nuclear localization signal.

Subcellular localization of Th1 may determine its role

It is possible that Th1 has different roles in different subcellular locations. Nuclear Th1 likely acts as a DNA binding transcription factor with targets including heart specific genes whose expression is necessary for cardiac looping and the specification of the right and left ventricle. These event all occur prior to E10.5. Th1's cytoplasmic role would not be assumed until E13.5, after gross morphological changes in the heart are complete. It is possible that additional subtle remodeling is occurring, or that final differentiation of some cell types within the heart is taking place. Either of these events may require inactivation of Th1 transcriptional targets or initiation of an as yet unidentified cytoplasmic function.

Likewise, in the sympathetic ganglia Th1 may have diverse roles. Acting as a nuclear transcription factor, Th1 targets may be responsible for maintaining the stem cell potential of cells of the sympathetic lineage. At the precise time Th1 becomes cytoplasmic, E13.5, the stem cells within the sympathetic ganglia stop dividing and begin terminal differentiation into sympathetic neurons. Because Th1 expression is maintained in the cytoplasm (at least until E17.5, TLH unpublished), it is likely that there is an additional function for Th1 there.

Th1 may have a role in determining PcG targets

Nuclear Th1 may also play a role in identifying targets for PcG mediated repression. When developmental decisions are being made, Th1 and other transcription factors may be removed from the nucleus to allow a change in the transcriptional state of specific genes. The cytoplasmic localization of Th1, as a result of serum deprivation, supports this model. Serum deprivation is known to induce differentiation in a number of

cell types. This process requires the activation of many tightly regulated, possibly silenced, genes. It is possible that in order to reset transcription of repressed genes it requires that key transcription factors, such as Th1, must be inactivated, in this case by removal from the nucleus. A prominent model for selection of PcG targets suggests that the equilibrium of active and repressive effectors at a promoter determines the final transcriptional state. This equilibrium may be influenced by temporary removal of influential factors from the nucleus.

Ectopic Th1 induces cell death in cultured cells

Ectopic expression of Th1 in culture frequently causes cell death. Other researchers have not identified a role for Th1 in apoptosis, but may not be assaying the right tissue type or developmental timepoint. A decrease in apoptosis has not been mentioned in descriptions of the Th1 or Th2 null mutant mice (Firulli *et al.* 1998; Riley *et al.* 1998). An increase in apoptosis explains the absence of branchial arches three and four in the Th2 mutant mouse (Thomas *et al.* 1998), but this would suggest that Th2 is responsible for cell death prevention as opposed to induction.

Th1 is expressed and localized to the nucleus during both major phases of cell death in the sympathetic ganglia. This is mirrored by nuclear re-entry when PC-12 cells are induced to apoptose. Unfortunately, Thing null mutant mice do not live long enough to investigate cell death in the sympathetic ganglia. Further experiments in PC-12 cells will help define Th1's role in survival of sympathetic neurons.

In summary, the cyptoplasmic versus nuclear localization of Th1 reveals two important points. 1) Th1 function is tightly regulated such that it is necessary to quickly remove it from the nucleus and prevent its role in transcription there. The regulation is controlled by the use of two separate NLSs in Th1 which may be independently manipulated for nuclear entry. 2) Th1 may have both a nuclear function and a cytoplasmic function evidenced by its regulated localization both in culture and the developing mouse embryo. Nuclear functions surely include activation of heart specific genes. Cytoplasmic

functions are as yet unknown. These two characteristics of Th1 make it a transcription factor worth further analysis.

Th1 interacts with multiple PcG proteins

We have used the yeast two hybrid system to identify protein partners for the bHLH protein Th1. Screens using the Th1 C-terminus and a Th1-E protein intramolecular dimer each isolated a novel, structurally distinct member of the PcG. Since PcG proteins are thought to function in multi-protein complexes this suggested Th1 might interact with other PcG proteins. To further investigate this we analyzed the ability of three additional vertebrate PcG proteins to physically interact with Th1 in vitro, in yeast and in mammalian cells. All PcG proteins scored positive in the mammalian cell interaction assay and mPCL2, mPh2 and XPc were positive in all of these assays. This in vitro and cellular evidence suggests that Th1 can act through a PcG mechanism to regulate gene expression.

The ability of Th1 to bind several PcG proteins is very striking and in sum provides a strong indication that Th1 directs formation of a PcG multi-protein complex. Due to technical limitations we are unable to demonstrate simultaneous, direct binding of more than one PcG protein to Th1. Nevertheless, the two hybrid clones TE14 and ΔH53 interact with non-overlapping regions of Th1 in the yeast two hybrid system (SMH unpublished), supporting the hypothesis that Th1 binds at least two PcG proteins at once.

Our in vitro analysis identifies not only a series of PcG interactions with Th1 but also with each other (Figure 11). Several lines of evidence suggest that PcG proteins function as multi-protein complexes (Alberts and Sternglanz 1990; Gaunt and Singh 1990; Paro 1990; Reuter *et al.* 1990). In support of a model for a multi-protein complex, we provide biochemical evidence for many homo- and hetero-PcG interactions. Particularly noteworthy is strong two hybrid and biochemical evidence for mPh2 and Mel-18 interaction. The numerous interactions that are detectable suggest flexibility in how PcG complexes are assembled. A protein that can influence several of these interactions would

be predicted to determine how these proteins form a complex, although further analysis will be required to precisely define the PcG complex or complexes that form with Th1.

This study was initiated based on earlier work which showed the bHLH of Th1 could function as a transferable repression domain (Hollenberg *et al.* 1995). We have found that the bHLH of Th1 is not sufficient for interaction with any of the PcG proteins, indicating that the repression function identified in this earlier study is mediated by protein interaction(s) distinct from those reported here, but instead may be mediated by TFII-I interaction (see Chapter Five).

Subnuclear Th1 Domains and PcG Interactions

The localization of transiently expressed Th1 to a punctate subnuclear pattern is very similar to the pattern observed for *Drosophila* PcG proteins Pc and Ph (Franke *et al.* 1995; Messmer *et al.* 1992) and for many PcG proteins in U2OS (Saurin *et al.* 1998). We have not detected endogenous Th1 in a punctate subnuclear pattern in PC-12 or Rcho-1 cells, or in mouse embryo sections. Either Th1-PcG interactions are cell or stage specific and are not occurring in these cells or they are obscured by a strong uniform nuclear pattern. In addition, Th1 may be associated with PcG proteins even when it adopts a uniform nuclear distribution, an interaction undetectable by our assay. These issues await further analysis.

We have used the subnuclear Th1 pattern produced by transient transfection as a spatial indicator of protein-protein interaction. This method relies on the ability of a PcG partner protein to be "mislocalized" by co-expressed Th1. It does not require all interacting components to be identified and may be capable of detecting indirect interactions. For example, Mel-18 shows no capability to interact directly with Th1 in either the two hybrid assay or biochemical assay, yet shows relatively efficient co-localization in the cellular assay (Figure 13 and 14A). We suggest that endogenous PcG proteins are participating in this process, perhaps to form a bridge between Th1 and Mel-18. This

would help explain the strict dependence on Th1 expression levels for detection of the subnuclear pattern.

We have presented a cellular co-localization assay, in part, due to our inability to demonstrate an interaction between Th1 and PcG proteins by co-immunoprecipitation from cell extracts. Th1 and PcG proteins only co-localize at sub-maximal expression levels, so we do not have the sensitivity to detect the interaction in extracts from transfected cells. Nevertheless, our cellular assay system has a number of advantages over co-immunoprecipitation. For example, the co-localization can be measured in intact, fixed cells and thus is not a result of interaction after lysis, and it can occur at lower expression levels. The ability of Th1 to recruit proteins to its subnuclear expression domain in our assay is difficult to explain without a physical interaction. Formally, Th1 could be acting to target the PcG proteins to its subnuclear domains by an indirect mechanism. We consider this hypothesis very unlikely for two reasons. First, seven randomly selected nuclear proteins were not mislocalized by Th1 expression (TLH unpublished); the effect is specific to the PcG proteins tested. Second, since Th1 and XPc often are frequently found in distinct subnuclear domains, an indirect Th1 effect on PcG protein localization would be expected to direct these proteins to XPc subnuclear sites. Instead, Th1 recruits other PcG proteins to its subnuclear pattern almost exclusively.

The formation of a neighbor punctate pattern by Th1 and XPc and occasionally by other PcG partners is distinct from co-localization (shown for Th1/XPc in Figure 14B, part h) but is further suggestive evidence that co-localization is a regulated biological phenomenon. The neighbor punctate distribution is likely to be formed either as a precursor to co-localization or as a segregation of co-localized proteins. An apparent precedent for the second possibility has been observed with patient autoimmune sera (Ascoli and Maul 1991). Nuclear "dots" formed by unidentified proteins were seen as paired structures and increased in frequency when cells were stimulated to proliferate. We

are currently investigating the relationship between cell growth, cell cycle and co-localization.

A Model for Th1 Function

Current models for PcG repression in *Drosophila* postulate the temporally regulated formation of a multi-protein repression complex on the PREs of a gene. The mechanism by which PcG proteins are recruited to the PRE has not been resolved (Bienz and Muller 1995), and no direct role for tissue-specific factors in this process has been demonstrated, although most work focuses specifically on the *Ubx* gene. Th1 has specific DNA binding properties (Hollenberg *et al.* 1995) and is tissue-specific, thus its role could be to recruit PcG proteins to specific genes. This may be analogous to the action of a GAL4-Pc fusion protein reported by Müller (Muller 1995). In this study transient recruitment of Pc to GAL4 binding sites is sufficient for transient repression, but another PRE(s) is needed to stabilize this activity long term after Pc-GAL4 expression is terminated. Developmental expression of Th1 is transient in some sites, so it could function analogous to a temporary burst of GAL4-Pc, repressing genes by PcG localization. The subsequent stabilization of this activity in the absence of Th1 would be a gene-specific event and dependent upon the presence of other linked PREs.

This model for Th1 action postulates the existence of both a tissue-specific PcG complex (with Th1) and a PRE. Although some evidence suggests that PcG activity is tissue-specific (Soto *et al.* 1995), the mechanism for this effect is not well understood. In our model a Th1-PcG complex recognizes a Th1-specific PRE. The only DNA binding activity that has been measured for Th1 required the presence of E protein as a heterodimer protein. In contrast, we have no evidence that E protein is an obligate partner for Th1-PcG interactions. mPh2 was isolated as a Th1-E partner, but E protein is not required for this interaction. In addition, we have only infrequently detected co-localization of transfected Th1 and E protein in subnuclear domains, and E protein co-expression dramatically reduces

the frequency of the Th1 punctate pattern (TLH unpublished). Finally, Id1 co-expression, which is expected to compete for E protein interaction, has no inhibitory effect on either the formation of Th1 subnuclear domains or the efficiency of Th1-PcG co-localization (TLH unpublished). Therefore, Th1 may have a non-E protein partner interacting with the bHLH domain. We have identified two non-bHLH partners for the Th1 bHLH domain that act as repressors and co-localize to Th1 subnuclear domains, TFII-I and CENP-B (see Chapter Five). This suggests the possibility that Th1 may not be directly binding DNA, but might be mediating DNA binding activity indirectly by the action of the PcG proteins it recruits. Our observation that other bHLH proteins also form domains in the nucleus suggests that this may be a common function for transcription factors. This could provide a link between positively defining a cell fate and silencing the expression of regulators for other lineages.

Th1 protein interactions may allow many functions

Th1 association with PcG members, E protein, CENP-B and TFII-I imply that its functions in transcription may be modulated by its protein interactions. When heterodimerized with E protein, Th1 can activate transcription (Hollenberg *et al.* 1995). However, by association with repressors, its function may be reversed. Th1 repression may be direct via a Th1 DNA binding site in a promoter, or Th1 repression may be mediated by recruitment of the PcG. Th1 interaction with both CENP-B and TFII-I also implies that it may have the ability to associate with both chromatin and the transcriptional machinery. One model for how targets are chosen for PcG repression suggests that the state of a gene is sensed and if already repressed, stabilized by a PcG mechanism. Th1's potential to recognize and influence both states, through diverse protein interactions, and its ability to recruit PcG proteins makes it an ideal candidate for sensing the state of a gene and assembling a PcG complex.

CENP-B may act to sequester genes for PcG repression

Sequences around the centromere, including the region of CENP-B binding, may be a site for PcG repression. As others have suggested, this may be the epicenter of nuclear silencing domain (Brown *et al.* 1997). Interaction of Th1 with CENP-B may facilitate the delivery of Th1 repression targets to this site where PcG proteins would be enlisted to maintain silencing. Evidence for such a site is outlined in studies by Csink and Henikoff which show that repressed genes are in close association with the heterochromatin (Csink and Henikoff 1996). Identification of a PcG localized domain in U20S cells, adjacent to the centromere, also supports the idea that a repressive site may be established at which silencing would be maintained by the pooling of repressor proteins. CENP-B interacts with at least two PcG proteins. With the yeast two hybrid assay we detect an interaction with Pc (SMH unpublished), and CENP-B overexpression can modulate the localization of Bmi-1 in NIH3T3 cells. We postulate that CENP-B helps to create a repression domain in the nucleus by interacting with tissue specific transcription factors and PcG proteins, thus dragging their targets to the centromere.

TFII-I may repress by two different methods

Data suggests that TFII-I acts as both an activator and inhibitor of transcription. TFII-I has two roles in activating transcription of the adenovirus major late core promoter. It is required for binding of TFIID to the TATA box and can mediate TATA-less transcription by binding the initiator element (Roy *et al.* 1993; Roy *et al.* 1991). TFII-I has also been shown to stabilize interactions between serum response factor and transcription factors (Grueneberg *et al.* 1997). Two studies report TFII-I mediated repression: Roy and coworkers showed that TFII-I can inhibit myc activated transcription (Roy *et al.* 1993) and we have described the ability of TFII-I to repress transcription when tethered to DNA through a Gal4 fusion (Figure 17).

PcG interactions with Th1 do not require the bHLH, suggesting that Th1 repression reported in Hollenberg et al [Hollenberg, 1995 #177) is not through a PcG mechanism. TFII-I binds the bHLH of Th1 (SMH unpublished). It is not able to disrupt the E/Th1 interaction, but the interactions are mutually exclusive (TH and SMH unpublished). We suggest that TFII-I may mediate bHLH dependent repression by Th1. In some contexts promoters may be bound by Th1 through TFII-I, possibly forming inactive complexes with the transcriptional machinery.

TFII-I repeat structure allows multiple protein interactions

The repeat structure of TFII-I is ideal for interaction with many transcription factors. Phox-1 has been shown to interact with the second repeat (Grueneberg *et al.* 1997), serum response factor the first repeat (Roy *et al.* 1991), hematopoetic bHLH, SCL-1 binds the first repeat (unpublished communication) and Th1 interacts with the fifth repeat (Figure 18). This propensity for binding a number of transcription factors, possibly in combination, might allow TFII-I to orchestrate complex formation. In this role, TFII-I would stabilize cooperative DNA binding of diverse transcription factors. Through this action it might dictate the relative positions of binding sites on a promoter. Another role for TFII-I might be in sensing the state of a gene by sampling the transcription factors bound at a promoter thus determining whether to recruit TFIID and other components of the transcriptional machinery, or inhibit transcription.

Some non-PcG related functions for TFII-I can be proposed. It is likely that TFII-I interacts with Th1 to stabilize Th1 interactions with other factors or DNA, as it has been shown to do with Phox-1 binding at the c-fos promoter (Grueneberg *et al.* 1997). However, our studies indicate that TFII-I and E protein cannot be bound simultaneously, so what Th1's transcriptional activity would be in this conformation is uncertian. A second possible function for a Th1/TFII-I complex is simply to block Th-mediated transcription by preventing it from binding to DNA.

En-2 as both a direct and indirect target of vertebrate PcG

PcG proteins play fundamental roles in the repression of target genes during development, yet an understanding of their vertebrate roles is just beginning to emerge. Our studies in *Xenopus* have identified two PcG target genes, *En-2* and *Rx2A*, and a characteristic set of phenotypic changes in the nervous system which results from PcG overexpression. These observations demonstrate that *Xenopus* provides a powerful developmental assay system for studying PcG function and identifying new PcG family members and target genes.

Studies in *Drosophila* have identified a number of PcG target genes, including *engrailed* and *Ultrabithorax*. Several lines of evidence demonstrate that these are direct target genes to which a PcG complex is physically bound. Our results suggest that PcG binding to *engrailed* may be conserved between *Drosophila* and vertebrates. Specifically, overexpression of four structurally distinct PcG proteins by RNA injection in *Xenopus* embryos represses *En-2* expression. Consistent with this suggestion, the *En-2* gene is repressed in Rat1a cells stably overexpressing the putative PcG gene *RING1a* (Satijn and Otte 1999). An important test of specificity in our study is provided by expression analysis of five other genes which are not repressed at early stages by Pcl1.

Overexpression of PcG proteins also produced a posterior shift in *En-2* expression. This effect is very likely to be an indirect consequence of PcG action. Numerous genetic studies, in both *Drosophila* and mice, have defined the importance of PcG proteins for maintaining spatial restriction of homeotic gene expression. Removal of PcG function produced posteriorization of the embryos due to an anterior shift of homeotic expression. One overexpression study has demonstrated that the effect can reverse direction: excess Bmi-1 produces an anteriorized embryo (Alkema *et al.* 1995). The posterior shift of *En-2* that we detect is consistent with embryo anteriorization. If we are altering anterior-posterior axis positional information, other markers should also shift posteriorly. We analyzed one

other marker for posterior shift, *Krox-20*. It frequently shifted posteriorly along with *En-2*, approximately one to two rhombomeres. This relatively small change in positional identity observed for *En-2* and *Krox-20* is reminiscent of effects seen in studies of mutation or overexpression of PcG in mice.

Rx2A repression in the eye

The power of studying the PcG in *Xenopus* is demonstrated by our discovery of a second PcG repression target, *Rx2A*. Pcl1 mediates very strong reduction of *Rx2A*, often to undetectable levels, although the eye field is present and expresses *N-CAM* normally. Earlier studies have demonstrated a role for PcG function in eye development: mouse polyhomeotic *(Rae28)* mutant mice exhibit microphthalmy (Takihara *et al.* 1997). Eye-restricted transcription factors may act both hierarchically and through cross-regulation (Loosli *et al.* 1999), thus further studies will be required to determine if *Rx2A* repression by Pcl1 is direct. The changes observed for both *En-2* and *Rx2A* expression underscore the potential for identifying additional PcG targets.

Tissue-specific expression of PcGs in the CNS

XPcl1 shows the strongest expression in the anterior central nervous system, as has been reported for some other PcG genes. For example, expression of the other two identified *Xenopus* PcG genes were reported to be nervous system-specific or enriched (Reijnen *et al.* 1995). In *Drosophila*, many PcG genes show a CNS-enriched pattern of RNA or protein accumulation at later stages of development (Bornemann *et al.* 1998; Buchenau *et al.* 1998; Martin and Adler 1993; Paro and Zink 1993). On the other hand, studies in *Drosophila* and mammals have also reported a relatively uniform distribution of PcG RNA or protein, including *Drosophila* Pcl (Lonie *et al.* 1994; Pearce *et al.* 1992). Our studies with *XPcl1* indicate not only CNS-enrichment, but a distinctive pattern of expression within the nervous system. We have observed discontinuities in the CNS expression

pattern, one of which is overlapping with the *En-2* stripe of expression (YY and SMH, unpublished). The absence of expression of *XPcl1* from these boundaries suggests either that it helps create the boundaries or may be a consequence of them.

In general, models of PcG action don't account for tissue-specific differences in expression level, although tissue-specific effects of PcG activity have been observed in genetic studies (Soto *et al.* 1995). Instead, PcG activity is assumed to be mediated largely by earlier developmental events at a target gene. PcG enrichment and action in the developing brain suggests the importance of tissue-specific expression levels.

PcG Overexpression Phenotype in the CNS

As shown in our overexpression studies and earlier work (Alkema *et al.* 1995), enhanced PcG levels can alter PcG function. The site of greatest sensitivity in our assay, the anterior nervous system, is also the site of highest endogenous expression. This suggests an important role for PcG in helping to fashion the nervous system. Consistent with this hypothesis, mutation of mouse Bmi-1 produces neurological defects (van der Lugt *et al.* 1994). Genetic studies indicate the importance of PcG in maintaining homeotic expression patterns along the anterior-posterior axis, but changes along this axis are unlikely to be responsible for the PcG neural phenotype. Markers for positional identity within the developing anterior nervous system are positioned correctly in embryos at stage 19-22 (Figure 28).

Rather than alterations in positional information, excess cell proliferation could account for the neural defect. PcG proteins directly or indirectly repress genes encoding at least two key components of the cell cycle regulatory mechanism: *ink4a* and *c-myc* (Jacobs *et al.* 1999; Satijn *et al.* 1997; Tetsu *et al.* 1998). PcG overexpression would be predicted to reduce the cell cycle block mediated directly by the CDK inhibitor p16INK4A (for a review see (Sharpless and DePinho 1999). Alternatively, rather than too much cell division, apoptosis could be suppressed in PcG overexpression embryos. c-myc promotes apoptosis

in some cellular contexts, and its repression by PcG proteins could inhibit the cell death events which are readily detectable in the developing *Xenopus* nervous system (Hensey and Gautier 1998). Studies are underway to distinguish between these possibilities and to further define the target genes involved. This approach should led to the discovery of new PcG proteins, as well as new target genes and biological activities.

The Xenopus PcG assay system allows identification of new PcG proteins

The isolation of two new PcG proteins by the *Xenopus* assay is an important achievement. The Polycomb field has been advancing slowly because of the limitation of long and expensive genetic experiments required to identify new PcG members. *Drosophila* genetic analyses have been required for definitive determination of Polycomb function. It is expected that proteins identified in vertebrate assays will be products of *Drosophila* PcG genes that await molecular characterization or that have escaped genetic screens for homeotic mutations. Mammalian PcG members that have been assayed in null mutant mice all have *Drosophila* homologs which have already been shown to have Polycomb function in fly lines. The mouse assay for PcG is useful for verifying functional homology but inappropriate for assaying potential PcG proteins, due to the time and expense required.

Relative to mouse transgenics, the *Xenopus* PcG system has many advantages. The function of potential partners is assayed in a biological system within a developing organism. The assay is fast; PcG function can be determined within two weeks of cDNA isolation. New target genes can be tested easily which allows one to look at affects on many targets. Finally, many embryos can be assayed, which allows detection of small effects. For instance, verification of the observation that PRSM effects are reversible with time required large numbers of embryos collected at specific time points.

The strategy used for identification of BS69 and PRSM as new PcG members demonstrates the potential of this method. The procedure requires performing two hybrid

screens with PcG proteins as bait followed by a secondary screen for species cross-reactivity. Proteins meeting these requirements are assayed in the *Xenopus* system for the ability to repress specific target genes and to cause homeotic transformation. We feel confident this method will allow identification of may new PcG proteins.

BS69 and PRSM are evolutionarily conserved PcG proteins

BS69 was identified by interaction with adenovirus, E1A, an association which inhibits E1A transactivation (Hateboer *et al.* 1995). BS69 contains a PhD domain, also found in Pcl, but Pcl/BS69 interaction is not via this domain (SMH unpublished). Instead, it is likely that other PcG interactions are mediated by association with BS69 through the PHD domain. Pcl interacts with the carboxy-terminus of BS69. This portion of the protein is conserved in *C. elegans*, suggesting that PcG function mediated by BS69 is evolutionarily conserved as well. Finally, the carboxy-terminus is also responsible for interaction with the co-repressor, NCoR, and histone deacetylase (unpublished communication), suggesting that chromatin modification might be one role of BS69 in PcG function.

PRSM is a novel protein, but database searches reveal homologs in other species, including *Drosophila*, allowing further study of this protein by *Drosophila* genetics. A homeotic transformation would be expected in PRSM null mutant fly lines. In addition, because PRSM is the only protein yet assayed in our *Xenopus* system whose activity is reversed with time, it will be interesting to determine if the phenotype can yield clues about this phenomenon.

In conclusion, the *Xenopus* assay will fuel the PcG field through identification of new PcG members. The expansion of this functional group should help to define the mechanism of Polycomb silencing during development.

Collectively, the work described in this thesis represents a significant contribution to the Polycomb-Group field. I have shown data supporting the model that PcG proteins are attracted to specific genes by temporally and spatially regulated transcription factors like Th1, which are able to both bind DNA and interact with and recruit multiple PcG proteins. Spatial regulation of transcription factors may be fine tuned at the subcellular level by control of cytoplasmic versus nuclear localization. Th1 interaction with CENP-B, which binds the transcriptionally silent centromere, may help position Th1 targets in repression domains. Th1 transcriptional activity might be modulated by association with TFII-I, a protein capable of repression, but shown to interact with the transcriptional machinery.

Genetic experiments from other labs have identified many new PcG proteins, but these experiments are lengthy and expensive. I have presented a transient but biological assay system for PcG function which allowed the identification of new vertebrate target genes, *En-2* and *Rx2A,* and two new PcG proteins, BS69 and PRSM. The ability of BS69 to interact with histone deacetylase (unpublished communication) supports the model that PcG repression is maintained through chromatin modification. The characteristics of newly identified PcG members will continue to expand our understanding of PcG repression.

References

Aasland, R., T. J. Gibson, and A. F. Stewart. 1995. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* **20:** 56-59.

Akasaka, T., M. Kanno, R. Balling, M. A. Mieza, M. Taniguchi, and H. Koseki. 1996. A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. *Development* **122:** 1513-1522.

Alberts, B., and R. Sternglanz. 1990. Gene expression. Chromatin contract to silence [news]. *Nature* **344:** 193-194.

Alkema, M. J., M. Bronk, E. Verhoeven, A. Otte, L. J. van't Veer, A. Berns, and M. van Lohuizen. 1997a. Identification of Bmi1-interacting proteins as constituents of a multimeric mammalian polycomb complex. *Genes Dev* **11:** 226-240.

Alkema, M. J., J. Jacobs, J. W. Voncken, N. A. Jenkins, N. G. Copeland, D. P. Satijn, A. P. Otte, A. Berns, and M. van Lohuizen. 1997b. MPc2, a new murine homolog of the Drosophila polycomb protein is a member of the mouse polycomb transcriptional repressor complex. *J Mol Biol* **273:** 993-1003.

Alkema, M. J., N. M. van der Lugt, R. C. Bobeldijk, A. Berns, and M. van Lohuizen. 1995. Transformation of axial skeleton due to overexpression of bmi-1 in transgenic mice. *Nature* **374:** 724-727.

Ascoli, C. A., and G. G. Maul. 1991. Identification of a novel nuclear domain. *J. Cell Biol.* **112:** 785-795.

Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61:** 49-59.

Biben, C., and R. P. Harvey. 1997. Homeodomain factor Nkx2-5 controls left/right asymmetric expression of bHLH gene eHand during murine heart development. *Genes Dev* **11:** 1357-1369.

Bienz, M., and J. Muller. 1995. Transcriptional silencing of homeotic genes in Drosophila. *Bioessays* **17:** 775-784.

Blitz, I. L., and K. W. Cho. 1995. Anterior neurectoderm is progressively induced during gastrulation: the role of the Xenopus homeobox gene orthodenticle. *Development* **121:** 993-1004.

Bornemann, D., E. Miller, and J. Simon. 1996. The Drosophila Polycomb group gene *Sex comb on midleg (Scm)* encodes a zinc finger protein with similarity to polyhomeotic protein. *Development* **122:** 1621-1630.

Bornemann, D., E. Miller, and J. Simon. 1998. Expression and properties of wild-type and mutant forms of the Drosophila sex comb on midleg (SCM) repressor protein. *Genetics* **150:** 675-686.

Bour, B. A., M. A. O'Brien, W. L. Lockwood, E. S. Goldstein, R. Bodmer, P. H. Taghert, S. M. Abmayr, and H. T. Nguyen. 1995. Drosophila MEF2, a transcription factor that is essential for myogenesis. *Genes Dev* **9:** 730-741.

Bradley, L. C., A. Snape, S. Bhatt, and D. G. Wilkinson. 1993. The structure and expression of the Xenopus Krox-20 gene: conserved and divergent patterns of expression in rhombomeres and neural crest. *Mech Dev* **40:** 73-84.

Braun, T., B. Winter, E. Bober, and H. H. Arnold. 1990. Transcriptional activation domain of the muscle-specific gene- regulatory protein myf5. *Nature* **346:** 663-665.

Brown, J. L., D. Mucci, M. Whiteley, M. L. Dirksen, and J. A. Kassis. 1998. The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1 [see comments]. *Mol Cell* **1:** 1057-1064.

Brown, K. E., S. S. Guest, S. T. Smale, K. Hahm, M. Merkenschlager, and A. G. Fisher. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91:** 845-854.

Brunk, B. P., E. C. Martin, and P. N. Adler. 1991. Molecular genetics of the *Posterior sex combs/Suppressor 2 of zeste* region of Drosophila: aberrant expression of the *Suppressor 2 of zeste* gene results in abnormal bristle development. *Genetics* **128:** 119-132.

Buchenau, P., J. Hodgson, H. Strutt, and D. J. Arndt-Jovin. 1998. The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing. *J Cell Biol* **141:** 469-481.

Bunker, C. A., and R. E. Kingston. 1994. Transcriptional repression by Drosophila and mammalian Polycomb group proteins in transfected mammalian cells. *Mol. Cell. Biol.* **14:** 1721-1732.

Campos-Ortega, J. A. 1995. Genetic mechanisms of early neurogenesis in Drosophila melanogaster. *Mol. Neurobiol.* **10:** 75-89.

Celniker, S. E., S. Sharma, D. J. Keelan, and E. B. Lewis. 1990. The molecular genetics of the bithorax complex of Drosophila: cis- regulation in the Abdominal-B domain. *Embo J* **9:** 4277-4286.

Chan, C. S., L. Rastelli, and V. Pirrotta. 1994. A Polycomb response element in the *Ubx* gene that determines an epigenetically inherited state of repression. *EMBO J.* **13:** 2553-2564.

Chang, Y. L., B. O. King, M. O'Connor, A. Mazo, and D. H. Huang. 1995. Functional reconstruction of trans regulation of the *Ultrabithorax* promoter by the products of two antagonistic genes, *trithorax* and *Polycomb*. *Mol. Cell. Biol.* **15:** 6601-6612.

Chen, C. M., N. Kraut, M. Groudine, and H. Weintraub. 1996. I-mf, a novel myogenic repressor, interacts with members of the MyoD family. *Cell* **86:** 731-741.

Cheng, N. N., D. A. Sinclair, R. B. Campbell, and H. W. Brock. 1994. Interactions of polyhomeotic with Polycomb group genes of Drosophila melanogaster. *Genetics* **138:** 1151-1162.

Chiang, A., M. B. O'Connor, R. Paro, J. Simon, and W. Bender. 1995. Discrete Polycomb-binding sites in each parasegmental domain of the bithorax complex. *Development* **121:** 1681-1689.

Chinwalla, V., E. P. Jane, and P. J. Harte. 1995. The Drosophila trithorax protein binds to specific chromosomal sites and is co-localized with Polycomb at many sites. *EMBO J.* **14:** 2056-2065.

Christen, B., and M. Bienz. 1994. Imaginal disc silencers from Ultrabithorax: evidence for Polycomb response elements. *Mech. Dev.* **48:** 255-266.

Core, N., S. Bel, S. J. Gaunt, M. Aurrand-Lions, J. Pearce, A. Fisher, and M. Djabali. 1997. Altered cellular proliferation and mesoderm patterning in Polycomb-M33- deficient mice. *Development* **124:** 721-729.

Coulson, M., S. Robert, H. J. Eyre, and R. Saint. 1998. The identification and localization of a human gene with sequence similarity to Polycomblike of Drosophila melanogaster. *Genomics* **48:** 381-383.

Cross, J. C., M. L. Flannery, M. A. Blanar, E. Steingrimsson, N. A. Jenkins, N. G. Copeland, W. J. Rutter, and Z. Werb. 1995. *Hxt* encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. *Development* **121:** 2513-2523.

Cserjesi, P., D. Brown, G. E. Lyons, and E. N. Olson. 1995. Expression of the novel basic helix-loop-helix gene *eHAND* in neural crest derivatives and extraembryonic membranes during mouse development. *Dev. Biol.* **170:** 664-678.

Csink, A. K., and S. Henikoff. 1996. Genetic modification of heterochromatic association and nuclear organization in Drosophila. *Nature* **381:** 529-531.

Cui, Y., J. D. Brown, R. T. Moon, and J. L. Christian. 1995. Xwnt-8b: a maternally expressed Xenopus Wnt gene with a potential role in establishing the dorsoventral axis. *Development* **121:** 2177-2186.

Davis, C. A., D. P. Holmyard, K. J. Millen, and A. L. Joyner. 1991. Examining pattern formation in mouse, chicken and frog embryos with an En-specific antiserum. *Development* **111:** 287-298.

DeCamillis, M., N. S. Cheng, D. Pierre, and H. W. Brock. 1992. The *polyhomeotic* gene of Drosophila encodes a chromatin protein that shares polytene chromosome-binding sites with Polycomb. *Genes Dev.* **6:** 223-232.

Denisenko, O., M. Shnyreva, H. Suzuki, and K. Bomsztyk. 1998. Point mutations in the WD40 domain of Eed block its interaction with Ezh2. *Mol Cell Biol* **18:** 5634-5642.

Earnshaw, W. C., and N. Rothfield. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91:** 313-321.

Edmondson, D. G., and E. N. Olson. 1993. Helix-loop-helix proteins as regulators of muscle-specific transcription. *J. Biol. Chem.* **268:** 755-758.

Faria, T. N., and M. J. Soares. 1991. Trophoblast cell differentiation: establishment, characterization, and modulation of a rat trophoblast cell line expressing members of the placental prolactin family. *Endocrinology* **129:** 2895-2906.

Fields, S., and O. Song. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340:** 245-246.
Firulli, A. B., D. G. McFadden, Q. Lin, D. Srivastava, and E. N. Olson. 1998. Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat Genet* **18:** 266-270.

Franke, A., M. DeCamillis, D. Zink, N. Cheng, H. W. Brock, and R. Paro. 1992. Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of Drosophila melanogaster. *EMBO J.* **11:** 2941-2950.

Franke, A., S. Messmer, and R. Paro. 1995. Mapping functional domains of the Polycomb protein of Drosophila melanogaster. *Chromosome Res.* **3:** 351-360.

Garcia, E., C. Marcos-Guti#rrez, M. Lorente, J. C. Moreno, and M. Vidal. 1999. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *Embo J* **18:** 3404-3418.

Gaunt, S. J., and P. B. Singh. 1990. Homeogene expression patterns and chromosomal imprinting. *Trends Genet* **6:** 208-212.

Gavrieli, Y., Y. Sherman, and S. A. Ben-Sasson. 1992. Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation. *J Cell Biol* **119:** 493-501.

Gossen, M., S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard. 1995. Transcriptional activation by tetracyclines in mammalian cells. *Science* **268:** 1766-1769.

Grueneberg, D. A., R. W. Henry, A. Brauer, C. D. Novina, V. Cheriyath, A. L. Roy, and M. Gilman. 1997. A multifunctional DNA-binding protein that promotes the formation of serum response factor/homeodomain complexes: identity to TFII-I. *Genes Dev* **11:** 2482-2493.

Gunster, M. J., D. P. Satijn, K. M. Hamer, J. L. den Blaauwen, D. de Bruijn, M. J. Alkema, M. van Lohuizen, R. van Driel, and A. P. Otte. 1997. Identification and characterization of interactions between the vertebrate polycomb-group protein BMI1 and human homologs of polyhomeotic. *Mol Cell Biol* **17:** 2326-2335.

Hashimoto, N., H. W. Brock, M. Nomura, M. Kyba, J. Hodgson, Y. Fujita, Y. Takihara, K. Shimada, and T. Higashinakagawa. 1998. RAE28, BMI1, and M33 are members of heterogeneous multimeric mammalian Polycomb group complexes. *Biochem Biophys Res Commun* **245:** 356-365.

Hateboer, G., A. Gennissen, Y. F. Ramos, R. M. Kerkhoven, V. Sonntag-Buck, H. G. Stunnenberg, and R. Bernards. 1995. BS69, a novel adenovirus E1A-associated protein that inhibits E1A transactivation. *Embo J* **14:** 3159-3169.

Haupt, Y., W. S. Alexander, G. Barri, S. P. Klinken, and J. M. Adams. 1991. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice [see comments]. *Cell* **65:** 753-763.

Hemmati-Brivanlou, A., J. R. de la Torre, C. Holt, and R. M. Harland. 1991. Cephalic expression and molecular characterization of Xenopus En-2. *Development* **111:** 715-724.

Hensey, C., and J. Gautier. 1998. Programmed cell death during Xenopus development: a spatio-temporal analysis. *Dev Biol* **203:** 36-48.

Hobert, O., I. Sures, T. Ciossek, M. Fuchs, and A. Ullrich. 1996. Isolation and developmental expression analysis of Enx-1, a novel mouse Polycomb group gene. *Mech Dev* **55:** 171-184.

Hollenberg, S. M., R. Sternglanz, P. F. Cheng, and H. Weintraub. 1995. Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol. Cell. Biol.* **15:** 3813-3822.

Hudson, D. F., K. J. Fowler, E. Earle, R. Saffery, P. Kalitsis, H. Trowell, J. Hill, N. G. Wreford, D. M. de Kretser, M. R. Cancilla, E. Howman, L. Hii, S. M. Cutts, D. V. Irvine, and K. H. Choo. 1998. Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. *J Cell Biol* **141:** 309-319.

Inouye, C., P. Remondelli, M. Karin, and S. Elledge. 1994. Isolation of a cDNA encoding a metal response element binding protein using a novel expression cloning procedure: the one hybrid system. *DNA Cell Biol.* **13:** 731-742.

Jacobs, J. J., K. Kieboom, S. Marino, R. A. DePinho, and M. van Lohuizen. 1999. The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* **397:** 164-168.

Jones, C. A., J. Ng, A. J. Peterson, K. Morgan, J. Simon, and R. S. Jones. 1998. The Drosophila esc and E(z) proteins are direct partners in polycomb group-mediated repression. *Mol Cell Biol* **18:** 2825-2834.

Joyner, A. L. 1996. Engrailed, Wnt and Pax genes regulate midbrain--hindbrain development. *Trends Genet* **12:** 15-20.

Joyner, A. L., K. Herrup, B. A. Auerbach, C. A. Davis, and J. Rossant. 1991. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. *Science* **251:** 1239-1243.

Jurgens, G. 1985. A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* **316:** 153-155.

Kablar, B., R. Vignali, L. Menotti, M. Pannese, M. Andreazzoli, C. Polo, M. G. Giribaldi, E. Boncinelli, and G. Barsacchi. 1996. Xotx genes in the developing brain of Xenopus laevis. *Mech Dev* **55:** 145-158.

Kadota, S., I. G. Fantus, B. Hersh, and B. I. Posner. 1986. Vanadate stimulation of IGF binding to rat adipocytes. *Biochem Biophys Res Commun* **138:** 174-178.

Kageyama, R., Y. Sasai, C. Akazawa, M. Ishibashi, K. Takebayashi, C. Shimizu, K. Tomita, and S. Nakanishi. 1995. Regulation of mammalian neural development by helix-loop-helix transcription factors. *Crit. Rev. Neurobiol.* **9:** 177-188.

Kallianpur, A. R., J. E. Jordan, and S. J. Brandt. 1994. The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood* **83:** 1200-1208.

Kanno, M., M. Hasegawa, A. Ishida, K. Isono, and M. Taniguchi. 1995. *mel-18*, a Polycomb group-related mammalian gene, encodes a transcriptional negative regulator with tumor suppressive activity. *EMBO J.* **14:** 5672-5678.

Kawakami, S., K. Mitsunaga, Y. Y. Kikuti, A. Ando, H. Inoko, K. Yamamura, and K. Abe. 1998. Tctex3, related to Drosophila polycomblike, is expressed in male germ cells and mapped to the mouse t-complex. *Mamm Genome* **9:** 874-880.

Kennison, J. A. 1995. The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function. *Annu Rev Genet* **29:** 289-303.

Kennison, J. A., and J. W. Tamkun. 1988. Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. *Proc Natl Acad Sci U S A* **85:** 8136-8140.

Kim, D. W., V. Cheriyath, A. L. Roy, and B. H. Cochran. 1998. TFII-I enhances activation of the c-fos promoter through interactions with upstream elements. *Mol Cell Biol* **18:** 3310-3320.

Kintner, C. R., and D. A. Melton. 1987. Expression of Xenopus N-CAM RNA in ectoderm is an early response to neural induction. *Development* **99:** 311-325.

Kliewer, S., J. Garcia, L. Pearson, E. Soultanakis, A. Dasgupta, and R. Gaynor. 1989. Multiple transcriptional regulatory domains in the human immunodeficiency virus type 1 long terminal repeat are involved in basal and E1A/E1B-induced promoter activity. *J Virol* **63:** 4616-4625.

Knust, E., K. A. Bremer, H. Vassin, A. Ziemer, U. Tepass, and J. A. Campos-Ortega. 1987. The enhancer of split locus and neurogenesis in Drosophila melanogaster. *Dev Biol* **122:** 262-273.

Kraut, N., L. Snider, C. M. Chen, S. J. Tapscott, and M. Groudine. 1998. Requirement of the mouse I-mfa gene for placental development and skeletal patterning. *Embo J* **17:** 6276-6288.

Kyba, M., and H. W. Brock. 1998. The Drosophila polycomb group protein Psc contacts ph and Pc through specific conserved domains. *Mol Cell Biol* **18:** 2712-2720.

Lammer, E. J., and J. M. Opitz. 1986. The DiGeorge anomaly as a developmental field defect. *Am J Med Genet Suppl* **2:** 113-127.

Landecker, H. L., D. A. Sinclair, and H. W. Brock. 1994. Screen for enhancers of *Polycomb* and *Polycomblike* in Drosophila melanogaster. *Dev. Genet.* **15:** 425-434.

Lewis, A. L., Y. Xia, S. K. Datta, J. McMillin, and R. E. Kellems. 1999. Combinatorial interactions regulate cardiac expression of the murine adenylosuccinate synthetase 1 gene. *J Biol Chem* **274:** 14188-14197.

Lilly, B., B. Zhao, G. Ranganayakulu, B. M. Paterson, R. A. Schulz, and E. N. Olson. 1995. Requirement of MADS domain transcription factor D-MEF2 for muscle formation in Drosophila. *Science* **267:** 688-693.

Lin, Q., J. Schwarz, C. Bucana, and E. N. Olson. 1997. Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* **276:** 1404-1407.

Lonie, A., R. D'Andrea, R. Paro, and R. Saint. 1994. Molecular characterisation of the *Polycomblike* gene of Drosophila melanogaster, a trans-acting negative regulator of homeotic gene expression. *Development* **120:** 2629-2636.

Loosli, F., S. Winkler, and J. Wittbrodt. 1999. Six3 overexpression initiates the formation of ectopic retina. *Genes Dev* **13:** 649-654.

Martin, E. C., and P. N. Adler. 1993. The Polycomb group gene Posterior Sex Combs encodes a chromosomal protein. *Development* **117:** 641-655.

Masumoto, H., H. Masukata, Y. Muro, N. Nozaki, and T. Okazaki. 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J Cell Biol* **109:** 1963-1973.

Mathers, P. H., A. Grinberg, K. A. Mahon, and M. Jamrich. 1997. The Rx homeobox gene is essential for vertebrate eye development. *Nature* **387:** 603-607.

McKeon, J., E. Slade, D. A. Sinclair, N. Cheng, M. Couling, and H. W. Brock. 1994. Mutations in some Polycomb group genes of Drosophila interfere with regulation of segmentation genes. *Mol. Gen. Genet.* **244:** 474-483.

Messmer, S., A. Franke, and R. Paro. 1992. Analysis of the functional role of the Polycomb chromo domain in Drosophila melanogaster. *Genes Dev.* **6:** 1241-1254.

Mihaly, J., R. K. Mishra, and F. Karch. 1998. A conserved sequence motif in Polycomb-response elements [letter; comment]. *Mol Cell* **1:** 1065-1066.

Moazed, D., and P. H. O'Farrell. 1992. Maintenance of the *engrailed* expression pattern by Polycomb group genes in Drosophila. *Development* **116:** 805-810.

Muller, J. 1995. Transcriptional silencing by the Polycomb protein in Drosophila embryos. *Embo J* **14:** 1209-1220.

Müller, J., S. Gaunt, and P. A. Lawrence. 1995. Function of the Polycomb protein is conserved in mice and flies. *Development* **121:** 2847-2852.

Muro, Y., H. Masumoto, K. Yoda, N. Nozaki, M. Ohashi, and T. Okazaki. 1992. Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. *J Cell Biol* **116:** 585-596.

Muro, Y., K. Sugimoto, T. Okazaki, and M. Ohashi. 1990. The heterogeneity of anticentromere antibodies in immunoblotting analysis. *J Rheumatol* **17:** 1042-1047.

Murre, C., P. S. McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, C. V. Cabrera, J. N. Buskin, S. D. Hauschka, A. B. Lassar, and et al. 1989. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58:** 537-544.

Neuhold, L. A., and B. Wold. 1993. HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74:** 1033-1042.

Nieuwkoop, P. D., and J. Faber. (1967). "The Normal Table of Xenopus Laevis." Garland Publishing, Amsterdam, North Holland.

Nomura, M., Y. Takihara, M. Abdul Motaleb, K. Horie, T. Higashinakagawa, and K. Shimada. 1998. Sequence-specific DNA binding activity in the RAE28 protein, a mouse homologue of the Drosophila polyhomeotic protein. *Biochem Mol Biol Int* **46:** 905-912.

Nomura, M., Y. Takihara, and K. Shimada. 1994. Isolation and characterization of retinoic acid-inducible cDNA clones in F9 cells: one of the early inducible clones encodes a novel protein sharing several highly homologous regions with a Drosophila polyhomeotic protein. *Differentiation* **57:** 39-50.

Olson, E. N., and D. Srivastava. 1996. Molecular pathways controlling heart development. *Science* **272:** 671-676.

Orlando, V., E. P. Jane, V. Chinwalla, P. J. Harte, and R. Paro. 1998. Binding of trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early Drosophila embryogenesis. *Embo J* **17:** 5141-5150.

Orlando, V., and R. Paro. 1993. Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell* **75:** 1187-1198.

Orlando, V., and R. Paro. 1995. Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr. Opin. Genet. Dev.* **5:** 174-179.

Pannese, M., C. Polo, M. Andreazzoli, R. Vignali, B. Kablar, G. Barsacchi, and E. Boncinelli. 1995. The Xenopus homologue of Otx2 is a maternal homeobox gene that demarcates and specifies anterior body regions. *Development* **121:** 707-720.

Papalopulu, N., and C. Kintner. 1993. Xenopus Distal-less related homeobox genes are expressed in the developing forebrain and are induced by planar signals. *Development* **117:** 961-975.

Paro, R. 1990. Imprinting a determined state into the chromatin of Drosophila. *Trends Genet* **6:** 416-421.

Paro, R. 1995. Propagating memory of transcriptional states. *Trends Genet.* **11:** 295-297.

Paro, R., and D. S. Hogness. 1991. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. *Proc. Natl. Acad. Sci. USA* **88:** 263-267.

Paro, R., and B. Zink. 1993. The Polycomb gene is differentially regulated during oogenesis and embryogenesis of Drosophila melanogaster. *Mech. Dev.* **40:** 37-46.

Pearce, J. J., P. B. Singh, and S. J. Gaunt. 1992. The mouse has a Polycomb-like chromobox gene. *Development* **114:** 921-929.

Pelegri, F., and R. Lehmann. 1994. A role of polycomb group genes in the regulation of gap gene expression in Drosophila. *Genetics* **136:** 1341-1353.

Peterson, C. L., and J. W. Tamkun. 1995. The SWI-SNF complex: a chromatin remodeling machine? *Trends Biochem Sci* **20:** 143-146.

Pirrotta, V. 1997a. Chromatin-silencing mechanisms in Drosophila maintain patterns of gene expression. *Trends Genet* **13:** 314-318.

Pirrotta, V. 1997b. PcG complexes and chromatin silencing. *Curr Opin Genet Dev* **7:** 249-258.

Poux, S., C. Kostic, and V. Pirrotta. 1996. Hunchback-independent silencing of late *Ubx* enhancers by a Polycomb Group Response Element. *EMBO J.* **15:** 4713-4722.

Quong, M. W., M. E. Massari, R. Zwart, and C. Murre. 1993. A new transcriptional-activation motif restricted to a class of helix- loop-helix proteins is functionally conserved in both yeast and mammalian cells. *Mol Cell Biol* **13:** 792-800.

Ranganayakulu, G., D. A. Elliott, R. P. Harvey, and E. N. Olson. 1998. Divergent roles for NK-2 class homeobox genes in cardiogenesis in flies and mice. *Development* **125:** 3037-3048.

Rastelli, L., C. S. Chan, and V. Pirrotta. 1993. Related chromosome binding sites for zeste, suppressors of zeste and Polycomb group proteins in Drosophila and their dependence on *Enhancer of zeste* function. *EMBO J.* **12:** 1513-1522.

Reijnen, M. J., K. M. Hamer, J. L. den Blaauwen, C. Lambrechts, I. Schoneveld, R. van Driel, and A. P. Otte. 1995. *Polycomb* and *bmi*-1 homologs are expressed in overlapping patterns in Xenopus embryos and are able to interact with each other. *Mech. Dev.* **53:** 35-46.

Reuter, G., M. Giarre, J. Farah, J. Gausz, A. Spierer, and P. Spierer. 1990. Dependence of position-effect variegation in Drosophila on dose of a gene encoding an unusual zinc-finger protein. *Nature* **344:** 219-223.

Riley, P., L. Anson-Cartwright, and J. C. Cross. 1998. The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. *Nat Genet* **18:** 271-275.

Roy, A. L., C. Carruthers, T. Gutjahr, and R. G. Roeder. 1993. Direct role for Myc in transcription initiation mediated by interactions with TFII-I. *Nature* **365:** 359-361.

Roy, A. L., M. Meisterernst, P. Pognonec, and R. G. Roeder. 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature* **354:** 245-248.

Santamaría, P., and N. B. Randsholt. 1995. Characterization of a region of the X chromosome of Drosophila including *multi sex combs (mxc)*, a Polycomb group gene which also functions as a tumour suppressor. *Mol. Gen. Genet.* **246:** 282-290.

Satijn, D. P., D. J. Olson, J. van der Vlag, K. M. Hamer, C. Lambrechts, H. Masselink, M. J. Gunster, R. G. Sewalt, R. van Driel, and A. P. Otte. 1997. Interference with the expression of a novel human polycomb protein, hPc2, results in cellular transformation and apoptosis. *Mol Cell Biol* **17:** 6076-6086.

Satijn, D. P. E., and A. P. Otte. 1999. RING1 interacts with multiple polycomb-group proteins and displays tumorigenic activity [In Process Citation]. *Mol Cell Biol* **19:** 57-68.

Saurin, A. J., C. Shiels, J. Williamson, D. P. Satijn, A. P. Otte, D. Sheer, and P. S. Freemont. 1998. The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J Cell Biol* **142:** 887-898.

Schoorlemmer, J., C. Marcos-Gutierrez, F. Were, R. Martinez, E. Garcia, D. P. Satijn, A. P. Otte, and M. Vidal. 1997. Ring1A is a transcriptional repressor that interacts with the Polycomb- M33 protein and is expressed at rhombomere boundaries in the mouse hindbrain. *Embo J* **16:** 5930-5942.

Sewalt, R. G., J. van der Vlag, M. J. Gunster, K. M. Hamer, J. L. den Blaauwen, D. P. Satijn, T. Hendrix, R. van Driel, and A. P. Otte. 1998. Characterization of interactions between the mammalian polycomb-group proteins Enx1/EZH2 and EED suggests the existence of different mammalian polycomb-group protein complexes. *Mol Cell Biol* **18:** 3586-3595.

Sharpless, N. E., and R. A. DePinho. 1999. The *INK4A/ARF* locus and its two gene products. *Current Opinion in Genetics & Development* **9:** 22-30.

Shivdasani, R. A., and S. H. Orkin. 1996. The transcriptional control of hematopoiesis. *Blood* **87:** 4025-4039.

Shumacher, A., C. Faust, and T. Magnuson. 1996. Positional cloning of a global regulator of anterior-posterior patterning in mice. *Nature* **383:** 250-253.

Simon, J. 1995. Locking in stable states of gene expression: transcriptional control during Drosophila development. *Curr. Opin. Cell Biol.* **7:** 376-385.

Simon, J., A. Chiang, W. Bender, M. J. Shimell, and M. O'Connor. 1993. Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products. *Dev. Biol.* **158:** 131-144.

Sinclair, D. A., T. A. Milne, J. W. Hodgson, J. Shellard, C. A. Salinas, M. Kyba, F. Randazzo, and H. W. Brock. 1998. The Additional sex combs gene of Drosophila encodes a chromatin protein that binds to shared and unique Polycomb group sites on polytene chromosomes. *Development* **125:** 1207-1216.

Soto, M. C., T. B. Chou, and W. Bender. 1995. Comparison of germline mosaics of genes in the Polycomb group of Drosophila melanogaster. *Genetics* **140:** 231-243.

Srivastava, D., P. Cserjesi, and E. N. Olson. 1995. A subclass of bHLH proteins required for cardiac morphogenesis. *Science* **270:** 1995-1999.

Srivastava, D., T. Thomas, Q. Lin, M. L. Kirby, D. Brown, and E. N. Olson. 1997. Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND [see comments] [published erratum appears in Nat Genet 1997 Aug;16(4):410]. *Nat Genet* **16:** 154-160.

Stankunas, K., J. Berger, C. Ruse, D. A. Sinclair, F. Randazzo, and H. W. Brock. 1998. The enhancer of polycomb gene of Drosophila encodes a chromatin protein conserved in yeast and mammals. *Development* **125:** 4055-4066.

Struhl, G., and M. Akam. 1985. Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of Drosophila. *Embo J* **4:** 3259-3264.

Struhl, G., and R. A. White. 1985. Regulation of the Ultrabithorax gene of Drosophila by other bithorax complex genes. *Cell* **43:** 507-519.

Strutt, H., and R. Paro. 1997. The polycomb group protein complex of Drosophila melanogaster has different compositions at different target genes. *Mol Cell Biol* **17:** 6773-6783.

Sugimoto, K., Y. Hagishita, and M. Himeno. 1994. Functional domain structure of human centromere protein B. Implication of the internal and C-terminal self-association domains in centromeric heterochromatin condensation. *J Biol Chem* **269:** 24271-24276.

Tagawa, M., T. Sakamoto, K. Shigemoto, H. Matsubara, Y. Tamura, T. Ito, I. Nakamura, A. Okitsu, K. Imai, and M. Taniguchi. 1990. Expression of novel DNA-binding protein with zinc finger structure in various tumor cells. *J. Biol. Chem.* **265:** 20021-20026.

Takihara, Y., D. Tomotsune, M. Shirai, Y. Katoh-Fukui, K. Nishii, M. A. Motaleb, M. Nomura, R. Tsuchiya, Y. Fujita, Y. Shibata, T. Higashinakagawa, and K. Shimada. 1997. Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. *Development* **124:** 3673-3682.

Tamkun, J. W. 1995. The role of brahma and related proteins in transcription and development. *Curr Opin Genet Dev* **5:** 473-477.

Tamkun, J. W., R. Deuring, M. P. Scott, M. Kissinger, A. M. Pattatucci, T. C. Kaufman, and J. A. Kennison. 1992. brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68:** 561-572.

Tetsu, O., H. Ishihara, R. Kanno, M. Kamiyasu, H. Inoue, T. Tokuhisa, M. Taniguchi, and M. Kanno. 1998. mel-18 negatively regulates cell cycle progression upon B cell antigen receptor stimulation through a cascade leading to c-myc/cdc25. *Immunity* **9:** 439-448.

Thomas, T., H. Yamagishi, P. A. Overbeek, E. N. Olson, and D. Srivastava. 1998. The bHLH factors, dHAND and eHAND, specify pulmonary and systemic cardiac ventricles independent of left-right sidedness. *Dev Biol* **196:** 228-236.

Tischler, A. S., and L. A. Greene. 1975. Nerve growth factor-induced process formation by cultured rat pheochromocytoma cells. *Nature* **258:** 341-342.

van der Lugt, N. M., J. Domen, K. Linders, M. van Roon, E. Robanus-Maandag, H. te Riele, M. van der Valk, J. Deschamps, M. Sofroniew, M. van Lohuizen, and et al. 1994. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev* **8:** 757-769.

van Lohuizen, M. 1998. Functional analysis of mouse Polycomb group genes. *Cell Mol Life Sci* **54:** 71-79.

van Lohuizen, M., M. Frasch, E. Wientjens, and A. Berns. 1991. Sequence similarity between the mammalian *bmi*-1 proto-oncogene and the Drosophila regulatory genes *Psc* and *Su(z)2*. *Nature* **353:** 353-355.

van Lohuizen, M., M. Tijms, J. W. Voncken, A. Schumacher, T. Magnuson, and E. Wientjens. 1998. Interaction of mouse polycomb-group (Pc-G) proteins Enx1 and Enx2 with Eed: indication for separate Pc-G complexes. *Mol Cell Biol* **18:** 3572-3579.

Wade, P. A., P. L. Jones, D. Vermaak, and A. P. Wolffe. 1998. A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. *Curr Biol* **8:** 843-846.

Wedeen, C., K. Harding, and M. Levine. 1986. Spatial regulation of Antennapedia and bithorax gene expression by the Polycomb locus in Drosophila. *Cell* **44:** 739-748.

Weintraub, H. 1993. Summary: genetic tinkering--local problems, local solutions. *Cold Spring Harb Symp Quant Biol* **58:** 819-836.

Wilson, D. I., J. Burn, P. Scambler, and J. Goodship. 1993. DiGeorge syndrome: part of CATCH 22. *J Med Genet* **30:** 852-856.

Yamagishi, H., V. Garg, R. Matsuoka, T. Thomas, and D. Srivastava. 1999. A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects [see comments]. *Science* **283:** 1158-1161.

Yang, W., and S. Desiderio. 1997. BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. *Proc Natl Acad Sci U S A* **94:** 604-609.

Yoda, K., K. Kitagawa, H. Masumoto, Y. Muro, and T. Okazaki. 1992. A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. *J Cell Biol* **119:** 1413-1427.

Zhang, Y., G. LeRoy, H. P. Seelig, W. S. Lane, and D. Reinberg. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95:** 279-289.