Chapter 1: Introduction

Molecular and Epigenetic Regulation of Development

Neural crest as a stem cell model

A fundamental question in developmental biology is how a pluripotent precursor can generate an amazing diversity of specialized cell types. This involves the process by which stem cells become restricted in their fate potential over time, undergo lineage commitment, and finally differentiate into specific cell types and tissues. Restriction of stem cell potential occurs gradually over time. This necessitates maintenance of a degree of multipotency and plasticity throughout development and even into adulthood, since some tissues contain progenitors with the capacity to de- or trans-differentiate during tissue repair or oncogenesis (Pietersen and van Lohuizen, 2008). In an attempt to understand these important events, much research has been directed at identifying mechanisms that regulate multipotency and understanding the signals that orchestrate lineage-specific differentiation.

The neural crest has been a useful model system to study these processes *in vivo* because of its capability to differentiate into a large number of diverse cell types, its capacity to proliferate, and the persistence of multipotent progenitors within differentiated tissues. In addition, slowly developing model systems like the chicken embryo are amenable to embryological perturbation because of their accessibility and ease of manipulation, thus proving very useful for understanding neural crest development (Le Douarin, 2004). Neural crest precursors are specified during early gastrulation at the border of the presumptive neural plate, and come to reside within the dorsal neural folds by morphological rearrangements during neurulation. Upon neural tube closure, neural crest cells emigrate from its dorsal aspect and undergo one of the most extensive migrations in the vertebrate body, coming to populate almost every

developing tissue. They differentiate into cell types as diverse as cranial bone and cartilage, sensory, parasympathetic, and enteric ganglia, pigment cells, and secretory endocrine cells, among many others (Le Douarin and Kalcheim, 1999). It has been suggested that this highly specialized cell type, which was a major driving force during evolution of the vertebrate predator, should be considered a fourth germ layer (Gans and Northcutt, 1983; Hall, 2000).

Lineage-tracing experiments using the chick embryo have been instrumental in identifying neural crest precursors. These studies have demonstrated that neural crest progenitors are indistinguishable from the rest of the neuroepithelium prior to their emigration from the dorsal neural tube, and that progeny of single-labeled dorsal neuroepithelial cells can contribute to both neural crest or dorsal neural tube derivatives such as roof plate and commissural neurons (Bronner-Fraser and Fraser, 1988, 1989; Selleck and Bronner-Fraser, 1996). Intriguingly, even dorsal neural tube cells expressing canonical "pre-migratory neural crest" markers can contribute to either of these lineages, suggesting that the neural crest is not committed prior to emigration, despite exposure to a number of specification signals (LaBonne and Bronner-Fraser, 1999). Furthermore, clonogenic analyses of migrating neural crest cells have demonstrated that they remain largely multipotent throughout their journey and that fate restriction occurs gradually, at least in a portion of the progenitors. However, a pool of multipotent, if not pluripotent, neural crest stem cells persists even after differentiation in tissues such as the sensory, sympathetic, and enteric ganglia and peripheral nerves, which are able to self-renew in culture. This population is heterogeneous and contains both multipotent neural crest progenitors as well as cells that differentiate into only one or two cell types

(Crane and Trainor, 2006). Nevertheless, even the apparently lineage-restricted progenitors demonstrate a high degree of plasticity and can de-differentiate upon back-transplantation into younger embryo hosts, and trans-differentiate appropriately into alternative lineages in response to novel signals (Le Douarin, 2004). It is likely that this ability may contribute to the oncogenic potential of neural crest cells in neurocristopathies such as neuroblastoma and Schwannoma. However, the highly plastic and multipotent nature of the neural crest also makes it a promising candidate for stem cell therapy, such as for use in peripheral nerve repair (Crane and Trainor, 2006). Not surprisingly, a large amount of research has been dedicated to understanding the timing and mechanisms regulating emergence of this fascinating cell type during early embryogenesis.

Establishment of the neural plate border

It is now widely accepted that neural crest cells are first specified during gastrulation and preceding the emergence of a definitive neural plate, in all vertebrates examined including the chicken embryo. Explants of early chick gastrula-stage medial epiblast generate migratory neural crest cells in culture in the absence of inducing factors, and are able to autonomously differentiate into bona fide crest derivatives such as melanocytes and neurons (Basch et al., 2006). The precise region of the epiblast from which neural crest progenitors arise has been examined by fate mapping and found to coincide with the junction between future neural plate and non-neural ectoderm. This “neural plate border” region is fairly wide and also contains progenitors of neural plate, epidermis, and placodes that are highly intermixed and indistinguishable from each other either

morphologically or molecularly (Garcia-Martinez et al., 1993; Fernandez-Garre et al., 2002; Ezin et al., 2009). In order to understand how progenitors within the neural plate border acquire their distinct cell fates, one must first consider the signaling events that segregate neural tissue from non-neural ectoderm, therefore generating this specialized "in-between" region.

For some time, the process of neural plate induction and specification of the neural lineage was thought to be a relatively simple and "default" process occurring during gastrulation, involving inhibition of bone morphogenetic proteins (BMP) by diffusible factors emerging from a specialized mesodermal signaling center, the "organizer." Accordingly to this scenario, the ectodermal germ layer had an inherent predisposition to a neural fate in the absence of epidermal-derived BMP signals (Hemmati-Brivanlou and Melton, 1997). Moreover, diffusion of BMP inhibitors from the organizer was found to generate a concentration gradient which specified positional information such that epidermal cells were specified at high BMP levels, the neural plate formed where they were absent or low, and intermediate BMP concentrations at the neural plate border specified neural crest fate (LaBonne and Bronner-Fraser, 1999). While experimental support for this "default model" of neural induction came from studies using the *Xenopus* model system and primarily involved data from *in vitro* experiments, investigation of this process in amniotes, as well as more careful reexamination of inductive events in the frog have generated a more complicated picture involving integration of several distinct molecular pathways.

Current data suggest that pre-patterning of the ectoderm and specification of the neural fate occurs prior to gastrulation and that mesoderm induction and

the organizer are dispensable for neural induction in both anamniotes and amniotes (Pera et al., 1999; Kuroda et al., 2004). In chick and mouse, BMP inhibition by organizer-derived inhibitors is neither necessary nor sufficient for neural induction. Rather, FGF signaling appears to play a major role both independently and together with BMP repression (Wilson et al., 2000). Consequently, the only area of the chick epiblast which is affected by direct perturbation of BMP signaling is the border of the prospective neural plate, suggesting that the role of these factors in neural induction may be to maintain the boundary of neural plate formation (Streit and Stern, 1999). Following these reports in the chick, additional studies in *Xenopus* also found a requirement for the FGF pathway in neural induction, as well as for formation of the neural plate border, and consequently, the neural crest (Launay et al., 1996; LaBonne and Bronner-Fraser, 1998). In addition, the canonical Wnt signaling pathway has been implicated in neural induction in chick and *Xenopus*. Two members of the Wnt family are expressed at high levels in lateral epiblast of the chick blastula and have been shown to inhibit neural induction by blocking the ability of FGF to negatively regulate BMP (Wilson et al., 2001). In contrast, Wnt signaling in the *Xenopus* blastula is necessary for BMP inhibition in dorsal ectoderm, prior to diffusion of neuralizing factors from the organizer (Baker et al., 1999). Although the mechanism by which Wnt functions in neural induction varies between *Xenopus* and chick, this pathway is necessary in both organisms for induction of the third ectodermal derivative, the neural crest (LaBonne and Bronner-Fraser, 1998; Garcia-Castro et al., 2002). In summary, specification of the neural plate and neural plate border occurs very early in development in both amniotes and anamniotes, by processes that are generally conserved in other vertebrate model

systems, such as zebrafish and mouse. The neural plate is distinguished from non-neural ectoderm by the integration of signals from three separate pathways: BMP, FGF, and Wnt. In *Xenopus*, neural induction occurs primarily through BMP inhibition in dorsal ectoderm by early Wnt and FGF signals and later signals from the organizer. Cell fates at the neural plate border are specified by cooperative activity of intermediate ectodermal BMP levels and mesodermally derived FGF and Wnt signals. In chick, FGF plays a main role in neural induction by independently promoting neural fate as well as inhibiting BMP in medial epiblast, while high concentrations of ectodermal Wnt in lateral epiblast regulate the lateral extent of the neural plate. Neural plate border fates are specified at the edge of the presumptive neural plate by high levels of BMP and lateral diffusion of ectodermal Wnt, and later maintained by FGF signals emanating from paraxial mesoderm (LaBonne and Bronner-Fraser, 1999; Wilson and Edlund, 2001; Knecht and Bronner-Fraser, 2002). Therefore, by late gastrulation, diffusible growth factor signals have regionalized the ectoderm and the presumptive neural plate border has been established. However, specification of distinct neural plate border fates, including that of the neural crest, requires precise transcriptional readout of these early inductive signals.

Gene regulatory interactions driving early neural crest development

Understanding of transcriptional regulation of neural crest development is largely derived from functional studies in which putative neural crest specifier genes were perturbed by over-expression, dominant-negative inhibition, or antisense oligonucleotide knock-down. Unfortunately, the classical vertebrate model systems used to study neural crest development are not easily amenable

to the kind of genomic *cis*-regulatory analysis that has enabled formulation of detailed gene regulatory circuits for tissue-specific development in other organisms (Davidson et al., 2002). However, a putative gene regulatory network for neural crest development (NC-GRN) has been proposed based on the large collection of data from neural crest perturbation studies and examinations of epistatic relationships between vertebrate neural crest genes (Fig. 1.1, Meulemans and Bronner-Fraser, 2004). The NC-GRN proposes that the inductive events responsible for ectodermal patterning (BMP, FGF, Wnt) activate of group of transcription factors (Msx1, Dlx3/5, Pax3/7, Zic1) at the junction between neural and non-neural ectoderm, specifying this area as the neural plate border. Subsequently, highly coordinated activity of the "neural plate border specifiers" leads to the activation of "neural crest specifier" genes (Snail1/2, FoxD3, SoxE group, Myc, AP-2, Id) specifically in neural crest progenitors residing within the neural plate border (neighboring placode progenitors are specified by an alternative combination of signals). Expression of neural crest specifier genes confers competency to form bona fide neural crest by inducing effector genes which are necessary for delamination from the neural tube, migration along appropriate pathways, and cell type-specific differentiation (Sauka-Spengler and Bronner-Fraser, 2008).

The molecular events leading to specification of the neural plate border are reiterative and highly complex. For example, BMP signals at the edges of the neural plate in combination with Wnt signals from ectoderm induce *Msx1* and *Pax7* in the prospective neural plate border (Tribulo et al., 2003; Monsoro-Burq et al., 2005; Basch et al., 2006). The combination of high concentrations of FGF and low BMP activates the neural specifier *Zic1*, which also functions as a neural

plate border specifier by collaborating with *Msx* and *Pax* to induce downstream neural crest genes (Merzdorf, 2007). In contrast, high levels of ectodermal BMP and Wnt induce the ectoderm specifiers *Dlx5* and *Dlx3*, which function indirectly to position the neural plate border by repressing neuronal fate (Bang et al., 1997; Suzuki et al., 1997; Pera et al., 1999; Streit and Stern, 1999; Luo et al., 2001a; Tribulo et al., 2003; Monsoro-Burq et al., 2005). The fact that many of the neural plate border specifiers do not function uniquely in this region (e.g., Msx1 and Dlx3/5 genes are also ectodermal specifiers, while Zic1 is a neural gene) makes it incredibly difficult to precisely map the neural plate border using gene expression analysis. While we are currently unable to obtain cellular resolution of this process, we do know that the neural plate border region is established by cooperative activity and cross-regulatory interactions between neural plate border specifiers (Meulemans and Bronner-Fraser, 2004).

Some of the regulatory relationships between neural plate border and neural crest specifiers have been described, and attempts at dissection of *cis*-regulatory interactions are currently underway in a number of organisms. For example, one of the earliest neural crest-specific genes activated by Pax3 and Zic1 in *Xenopu*s is *FoxD3*, which promotes neural crest fate by inducing and maintaining expression of other neural crest specifiers such as the *SoxE* genes, and by segregating the neural crest lineage from other cell fates in the dorsal neural tube (Dottori et al., 2001; Kos et al., 2001; Montero-Balaguer et al., 2006; Stewart et al., 2006). Pax3/7 and Zic1 also cooperate with Msx1 to induce the neural crest specifier *Snail2*, which is essential for neural crest migration and also functions as an anti-apoptotic factor and regulator of *SoxE* expression (Nieto et al., 1994; Mayor et al., 1995; LaBonne and Bronner-Fraser, 2000; del Barrio and

Nieto, 2002; Monsoro-Burq et al., 2005; Sato et al., 2005; Taneyhill LA, 2007). Interactions between neural crest specifiers are highly complex, involving extensive auto- and cross-regulation, so that perturbation of one member of this group usually affects expression of all others (Meulemans and Bronner-Fraser, 2004). Interestingly, some neural crest specifiers perform several temporally distinct functions during development. For example, *AP2α* is activated during early development by high levels of BMP in non-neural ectoderm and specifies ectodermal fate by maintaining *Msx1* and *Dlx5* expression (Luo et al., 2002). However, during late neurulation, *AP2α* becomes recruited to the dorsal neural tube and functions in a feedback loop with Slug and Sox9 to maintain neural crest identity (Luo et al., 2003). Other transcription factors that are considered neural crest specifiers, such as c-myc, N-myc, and Id, function mainly as proliferation and survival factors and inhibitors of differentiation (Bellmeyer et al., 2003; Light et al., 2005). The regulatory targets of neural crest specifiers and their function in later stages of neural crest development are reviewed elsewhere (Meulemans and Bronner-Fraser, 2004; Sauka-Spengler and Bronner-Fraser, 2008).

In summary, specification of the neural crest lineage occurs via step-wise and highly coordinated activation of discrete groups of genes during early development. Although a simplistic view of the NC-GRN would suppose that transcriptional events during neural crest development proceed in a hierarchical fashion, we know that interactions between induction factors, neural plate border specifiers, and neural crest specifiers are characterized by a large degree of cross- and auto-regulation and are therefore highly complex. Precise timing of

induction of NC-GRN factors is still largely unknown, and studies in *Xenopus* and lamprey suggest that some neural crest specifiers are expressed as early as gastrulation concomitant with the neural plate border genes (Huang and Saint-Jeannet, 2004; Sauka-Spengler et al., 2007). In addition, expression of inducers such as *BMP* and neural plate border specifiers such as *Msx1* and *Pax7* persists throughout early neural crest development, suggesting the possibility of late roles in maintenance of the neural crest fate and continued regulation of neural crest specifiers. The complexity of the NC-GRN interactions, together with the fact that neural crest progenitors remain multipotent despite continuous exposure to a plethora of specifying signals, suggests the existence of modulatory factors that regulate early neural crest progenitor development.

Epigenetic regulation of embryonic development

The advent of whole-genome analysis by technologies such as ChIP-on-Chip and ChIP-Seq has enabled researchers to obtain a large-scale view of molecular events operating during stem cell development, lineage commitment, and differentiation (Mendenhall and Bernstein, 2008). Data from such studies have demonstrated that a ubiquitous and important mechanism for regulating gene expression during development involves epigenetic modification of chromatin structure. Chromatin state maps have illustrated that the majority of transcription factor families involved in cell type-specific determination are transcriptionally inactive in pluripotent stem cells and are correlated with high levels of trimethylation of histone H3 on lysine 27 ($H3K27me^3$), a mark of compacted heterochromatin. In contrast, in response to differentiation signals, developmental regulator genes become associated with activated polymerase II

and a methylation mark of active transcription, H3K4me^3 (histone H3 trimethylated on lysine 4), resulting in high expression. Concurrently, genes that are involved in maintenance of pluripotency (Oct4, Sox2, Nanog) or in specification of alternative cell lineages become repressed and labeled by H3K27me^3 during differentiation (Mikkelsen et al., 2007).

A recent pivotal study that examined methylation patterns in mouse embryonic stem cells (ESC) demonstrated that surprisingly, a large number of promoters of transcriptionally inactive developmental regulator genes are marked by both repressive and active chromatin marks (H3K27me^3 and H3K4me^3), which have been termed "bivalent" regions. Upon differentiation, genes that were characterized by bivalent domains in stem cells resolve to either one or the other methylation mark, and become preferentially activated or repressed (Bernstein et al., 2006). Based on these findings, it has been suggested that the bivalent domain may function to keep key developmental regulators "poised" to undergo a rapid change in transcriptional activity upon receiving differentiation signals. Therefore, epigenetic chromatin modifications play a vital role during development by regulating transcriptional events, preventing premature activation of lineage specification factors, and modulating inputs of developmental signals by enabling rapid and flexible changes in transcription of target genes (Pietersen and van Lohuizen, 2008). Not surprisingly, dysregulation of epigenetic mechanisms is regularly observed in a large number of human diseases and cancers (Delcuve et al., 2009).

The Polycomb Group of epigenetic repressors

The enzymatic complexes responsible for H3K27 and H3K4 trimethylation, the Polycomb Group (PcG) and Trithorax Group (TxG), have been extensively studied in a number of organisms and developmental processes and are highly conserved throughout evolution in both plants and animals (Schuettengruber et al., 2007; Whitcomb et al., 2007). PcG genes were first identified in *Drosophila* by E. B. Lewis as repressors of the Hox complex, as reflected in their names which describe the homeotic transformations that characterized the mutants (Lewis, 1978). Identification of vertebrate Polycomb orthologs has demonstrated that their role in axial patterning is conserved (Alkema et al., 1995; van der Lugt et al., 1996). Interestingly, PcG proteins play a critical role in maintenance and self-renewal of stem cells by repression of transcriptional regulators (Fig. 1.2A).

Isolation of PcG proteins and their characterization by biochemical assays has demonstrated that this group functions as two separate and sequentially acting complexes: Polycomb Repressive Complex 2 (PRC2) and Polycomb Repressive Complex 1 (PRC1), each of which consists of a set of core components in *Drosophila* and a large number of paralogs in vertebrates (Fig. 1.2B). The PRC2 subunit Enhancer of Zeste (E(z), or vertebrate Ezh1/2) is the key enzymatic partner responsible for trimethylation of histone H3 lysine 27. Three other core components of PRC2, which stimulate its methyltransferase activity, include Extra sex combs (Esc, or vertebrate Eed), Suppressor of Zeste (Su(z)12, or vertebrate Suz12), and Nurf55 (Sparmann and van Lohuizen, 2006; Schuettengruber et al., 2007). The downstream PRC1 complex is thought to recognize the PRC2-catalyzed methylation mark via a chromodomain of the

Polycomb protein (Pc, or vertebrate Cbx2/4/6/7/8). Core PRC1 components also include Polyhomeotic (Ph, or vertebrate Mph1/2 and Phc3), Posterior sex combs (Psc, or vertebrate Bmi-1 and Mel-18), and dRing (vertebrate Ring1B and Ring1A). The Ring1B protein possesses catalytic activity which is used to monoubiquitylate lysine 119 of histone H2A (Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007). This mark is necessary for maintenance of PcG repressive activity, which is lost in Ring1B mutants despite persistence of H3K27me^3 (Wang et al., 2004). Furthermore, presence of all PRC1 and PRC2 core components is necessary for their expression and repressive activity, suggesting extensive auto-regulation as a means of maintaining complex integrity (Boyer et al., 2006; Lee et al., 2006; van der Stoop et al., 2008).

The importance of PcG-mediated gene repression during development has been demonstrated by PcG transgenic mouse models, which exhibit an inability to maintain the embryonic stem cell state, drastic loss of stem cell populations, and embryonic lethality. Analogously, many human cancers are characterized by increased expression of Polycomb genes (Sparmann and van Lohuizen, 2006). In ChIP-on-Chip studies, binding of both PRC2 and PRC1 proteins in mouse and human embryonic stem cells was enriched at the silent promoters of a large number of genes involved in vertebrate development. The PcG target genes included members of such highly conserved transcription factor families as *Hox*, *Dlx*, *Irx*, *Lhx*, *Pou*, *Pax*, *Six*, *Sox*, and *Tbx*, among many others. Upon stimulation with differentiation factors, the Polycomb complexes were removed from chromatin, causing de-repression of target genes and subsequent differentiation. Interestingly, it was the PcG target genes that were preferentially activated upon stem cell differentiation. Similar de-repression of transcription

factor targets was observed in stem cells carrying mutations for one or more PRC members, causing inappropriate differentiation in culture (Boyer et al., 2006; Bracken et al., 2006; Lee et al., 2006).

Furthermore, promoter regions of PcG-target genes are primarily characterized by bivalent domains and often exhibit co-occupancy by PRC1 and PRC2 members, Trithorax group proteins, and RNA polymerase II that is in a "paused" biochemical conformation (Stock et al., 2007; Ku et al., 2008). Therefore, the Polycomb proteins function as critical regulators of development by repressing differentiation-promoting transcription factors in stem cells while maintaining them in a poised state, enabling rapid and highly coordinated activation upon reception of inductive signals. In addition, lineage-appropriate specification requires activation of cell type-specific genetic programs coincident with suppression of signals mediating alternative fates, states which can still be reversed prior to lineage commitment and therefore involve a large degree of plasticity. Finally, terminal differentiation necessitates maintained activation of specialized cell type markers and stable repression of multipotency factors and genes involved in other tissue functions. ChIP-on-Chip analysis of Polycomb binding and histone methylation in a variety of lineage precursors and differentiated cell types have demonstrated that the PcG participates in regulation of all of these processes and aspects of development (Bracken et al., 2006; Pasini et al., 2007; Mohn et al., 2008).

Structure and function of Bmi-1 in stem cell development

The vertebrate PRC1 ortholog of *Drosophila posterior sex combs*, Bmi-1, was one of the first Polycomb genes to be studied as a stem cell and oncogenic factor.

Bmi-1 was first identified in the mouse as a retrovirus-induced cooperator in lymphomagenesis with c-myc, and subsequently named "**B**-cell type specific **M**oloney murine lymphoma retrovirus **i**nsertion site **1** (Haupt et al., 1991). In these transgenic mice, over-expression of Bmi-1 induced lymphomas by inhibiting c-myc-mediated apoptosis in hematopoietic stem cells (Jacobs et al., 1999b). Conversely, Bmi-1 knockout mice exhibited gross defects in the hematopoetic system due to failure of hematopoietic stem cells to proliferate and self-renew (Lessard and Sauvageau, 2003; Park et al., 2003). Mice deficient in Bmi-1 also exhibit defects in self-renewal of other stem cell types, such as CNS (subventricular zone) and PNS (enteric neural crest) progenitors (Molofsky et al., 2003). The targets through which Bmi-1 functions to positively regulate the cell cycle in mouse were identified by double knockout experiments, and involve the $p16^{Ink4a}/p19^{Arf}$ locus of cell cycle repressors (Jacobs et al., 1999a; Molofsky et al., 2005). Subsequent ChIP experiments have demonstrated that Bmi-1 negatively regulates this locus by direct association, and this interaction has been extensively studied due to its role in cell senescence during aging (Bracken et al., 2007). In addition, Bmi-1 functions in stem cell development by regulating a number of differentiation-specific transcription factors in cooperation with other Polycomb members (Bracken et al., 2006). During later development, Bmi-1 is necessary for maintenance of postnatal stem cell populations and regulation of axial patterning by direct repression of homeotic genes (van der Lugt et al., 1994; Molofsky et al., 2003; Cao et al., 2005).

Bmi-1 is a ~40 kDa protein which is characterized by three distinct functional domains (Fig. 1.3A). A highly conserved cysteine-rich RING finger domain located near the N-terminus mediates protein-protein interactions with

the other RING- containing PRC1 members Ring1A and Ring1B (Hemenway et al., 1998; Satijn and Otte, 1999). Since the presence of Bmi-1 in the PRC1 complex has been shown to stimulate ubiquitination activity of Ring1B, it is not surprising that this key interaction domain is also necessary for the oncogenic potential and repressive activity of Bmi-1 (Alkema et al., 1997b; Itahana et al., 2003; Wang et al., 2004). The Bmi-1 protein also contains a conserved helix-turn-helix-turn-helix-turn (HTHTHT) domain which is necessary for interaction with Mph proteins, the mammalian orthologs of *Drosophila polyhomeotic*, and for the ability to repress transcription of Hox genes and other targets (Cohen et al., 1996; Alkema et al., 1997a). A proline, glutamine, serine, threonine-rich, or PEST, domain is localized in the C-terminus of Bmi-1, which may function to target the protein for rapid degradation, although this has not been definitively demonstrated *in vitro* (Alkema et al., 1997b). A putative MAPK-pathway phosphorylation site within this domain may be involved in subcellular translocation of Bmi-1 in response to external signals (Voncken et al., 2005). Based on biochemical protein interaction assays, it has been suggested that Bmi-1 may function as a tethering protein that maintains structural integrity of PRC1 (Fig. 1.3B, Cao et al., 2005).

In addition, the biochemical deletion studies that identified Bmi-1 functional domains also demonstrated that truncated portions of the protein could dimerize with full-length Bmi-1 and other PRC1 proteins, and exhibit dominant-negative effects (Hemenway et al., 1998; Satijn and Otte, 1999; Itahana et al., 2003). Based on these data, an intriguing possibility is that the function of Bmi-1 may be mediated by naturally occurring alternatively spliced isoforms which contain differential combinations of functional domains, therefore modulating the activity of Bmi-1 and the complex. Indeed, there is mounting

evidence that a number of Polycomb proteins and other critical developmental specifiers are regulated by alternative splicing (Alkema et al., 1997a; Yamaki et al., 2002; Tajul-Arifin et al., 2003; Li et al., 2005). Therefore, cell lineage diversification and differentiation during development likely involves several complex layers of regulation: transcriptional specification signals, their modification by epigenetic repressor complexes, and in turn, the modulation of those complexes by alternatively spliced isoforms.

Figure 1.1: Gene regulatory network for neural crest development

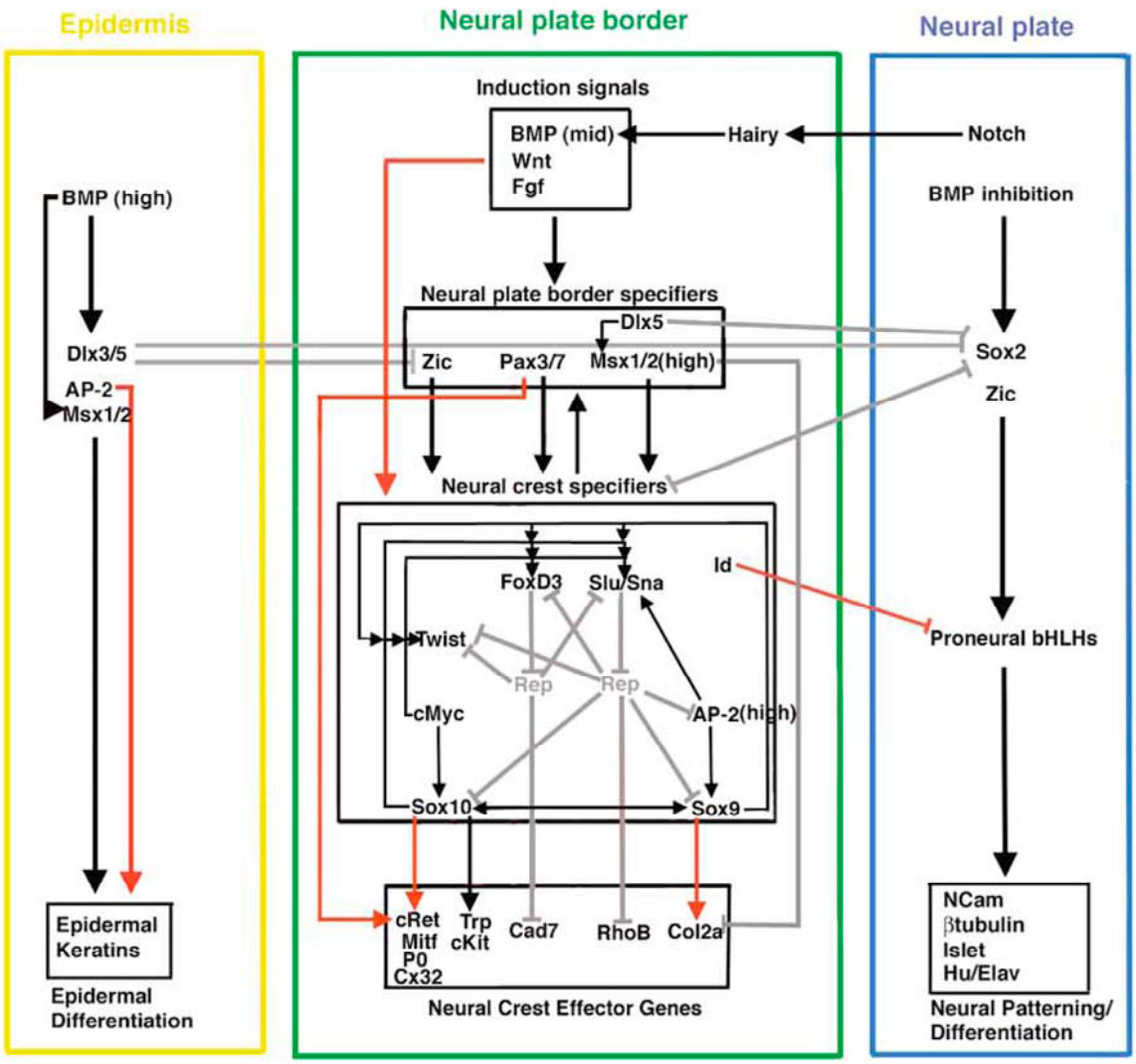

Figure 1.1. Putative gene regulatory network proposed by Meulemans and Bronner-Fraser to describe signaling and transcriptional events at the neural plate border during vertebrate neural crest development (Meulemans and Bronner-Fraser,2004).

Figure 1.2: Biochemical composition of Polycomb Repressive Complexes

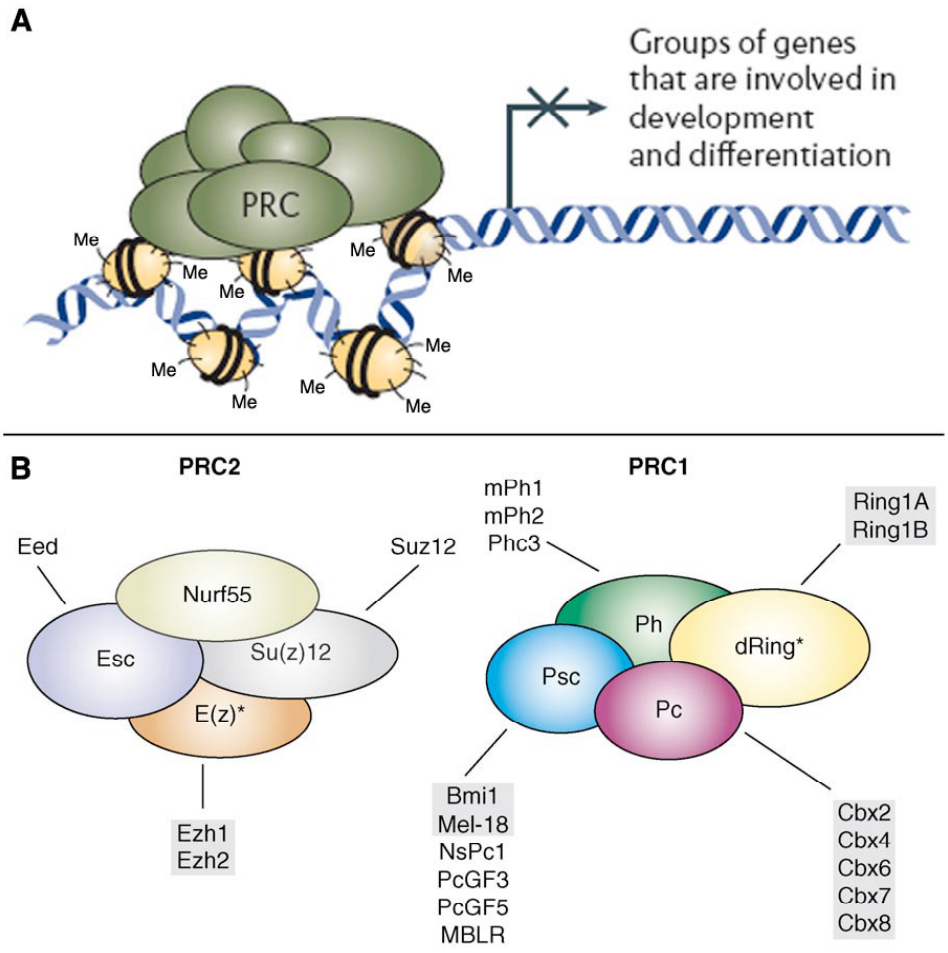

Figure 1.2. The Polycomb Group consists of two discrete complexes with a large number of highly conserved protein partners. **A.** In a simplistic schematic, the Polycomb complex is shown to associate with regulatory regions of genes involved in development and differentiation, causing histone methylation, compaction of heterochromatin, and transcriptional repression. Adapted from Baylin and Ohm, 2006. **B.** Diagram illustrating the core components of *Drosophila* Polycomb Repressive Complex 2 (PRC2) and Polycomb Repressive Complex 1 (PRC1), and their mammalian paralogs. From Whitcomb et al., 2007.

Figure 1.3: Biochemical structure of Bmi-1 protein

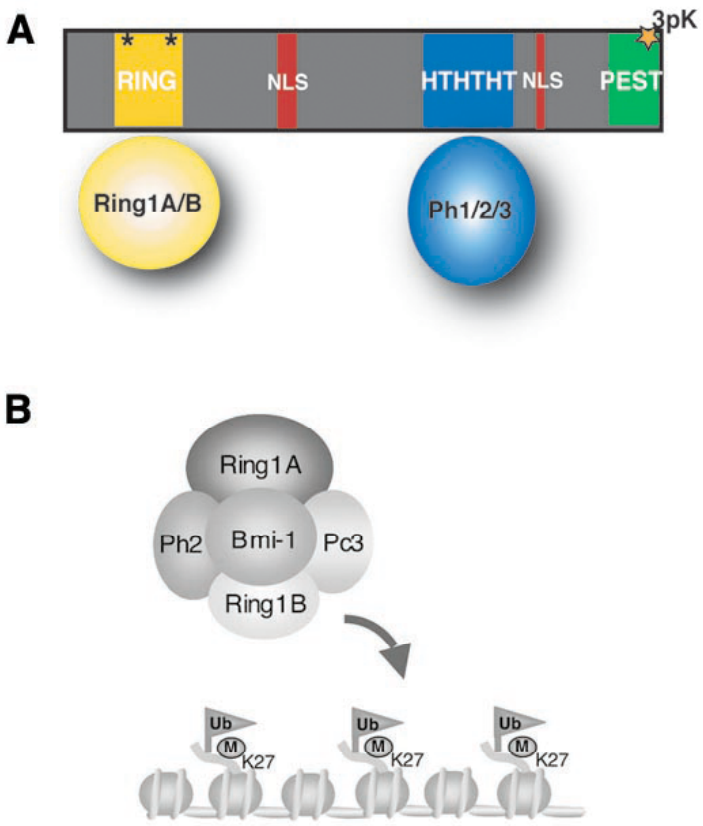

Figure 1.3. A. The PRC1 member Bmi-1 is characterized by several conserved motifs necessary for interaction with other complex members. The RING finger domain (yellow) mediates interaction with RING-containing proteins Ring1A and Ring1B, for which the presence of two conserved cysteine residues (asterisks) is required. The helix-turn-helix-turn-helix-turn (HTHTHT, blue) domain mediates interactions with polyhomeotic proteins Ph1, Ph2, and Phc3. A proline-glutamine-serine-threonine-rich (PEST, green) domain may be involved in protein degradation. A putative downstream MAPK pathway phosphorylation site (orange star) lies within this domain. **B.** Bmi-1 has been biochemically purified as a tethering protein in a complex containing Ring1A,

Ring1B, Ph2, and Pc3 (Cbx8). PRC1 (via Ring1B) ubiquitinates histone H2A on lysine 119 within genomic regions targeted by PRC2. From Cao et al., 2005.