Published in final edited form as:
Trends Cell Biol. 2009 June ; 19(6): 268–275. doi:10.1016/j.tcb.2009.03.007.

Regeneration, repair and remembering identity: the three Rs of *Hox* gene expression

Kevin C. Wang[1,2], **Jill A. Helms**[3], and **Howard Y. Chang**[1]

[1]Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

[2]Department of Dermatology, University of California San Francisco (UCSF), San Francisco, CA 94115, USA

[3]Department of Surgery, Stanford University School of Medicine, Stanford, CA 94305, USA

Abstract

Hox genes encode transcription factors that specify embryonic positional identity in cells and guide tissue differentiation. Recent advances have greatly increased our understanding of the epigenetic mechanisms that ensure the faithful expression of *Hox* genes in adult cells and which involve the interplay of histone methylation, demethylation and intergenic transcription of long non-coding RNAs. The transcriptional memory of *Hox* genes poses both an opportunity and a challenge for regenerative medicine. Matching the positional identity of transplanted stem cells with that of the host environment, as reflected by their respective *Hox* profiles, is likely to be required to achieve regenerative healing. Strategies to manipulate the plasticity of *Hox* gene expression will probably become a major focus in regenerative medicine.

Introduction

The discovery of genes that control embryonic body segment identity is one of the great triumphs of developmental biology. Recent studies have led to the realization that some of these same genes have ongoing and prominent functions in adult cells. The *Hox* genes in particular, which code for a large family of transcription factors, have key roles in embryonic segmental identity (Box 1, Figure 1; see http://www.youtube.com/watch?v=9k_oKK4Teco for a recent animated tribute), and their expression in adult cells constitutes a form of positional memory – an internal representation by a cell of where it is located within a multicellular organism.

The confluence of two areas of investigation has brought the transcriptional memory of *Hox* genes into focus. On the one hand, substantial progress has recently been made in unraveling mechanisms of epigenetic regulation of *Hox* genes. The fidelity of expression pattern of *Hox* genes is necessary for the normal homeostasis of adult tissues and organs, and mis-expression of *Hox* genes can readily lead to diseases such as cancer [1,2]. On the other hand, the positional memory of cells has important implications for the burgeoning field of

Corresponding author: Chang, H.Y. (howchang@stanford.edu).

regenerative medicine. A proper pattern of *Hox* genes could be programmed to make the desired tissues; conversely, the inability to erase or transcend a fixed pattern of *Hox* genes might be a key factor limiting regeneration in mammals. We describe here newly recognized epigenetic mechanisms that make the transcriptional memory of *Hox* particularly robust, and the implications of the cellular memory of *Hox* expression for tissue homeostasis and developmental plasticity that could prove to be necessary for tissue regeneration.

Persistent expression of *HOX* genes in adulthood

It is increasingly clear that *Hox* genes might have an enduring role in maintaining positional identity throughout the lifetime of an organism. For instance, unbiased global gene expression analysis of adult human fibroblasts, cultured *ex vivo*, showed that such cells maintain large-scale differences – comprising the differential expression of >1000 genes – that reflect the anatomic origin of cells [3]. This scale of differential gene expression between subtypes of fibroblasts is on a par with the level of differential expression seen among currently accepted, distinct cell types, such as in the many types of white blood cell. The most prominent class of genes in the positional memory of fibroblasts is the *HOX* genes, which preserve elements of the embryonic *HOX* expression patterns in these adult cells [3]. Systematic comparison of the gene expression programs of fibroblasts from numerous finely mapped anatomic sites across the human body confirmed that adult fibroblasts consistently expressed distinct patterns of *HOX* genes that were sufficient to indicate the position of the cell along three developmental axes [4]. Differential expression of the *HOXA* and *HOXD* genes reflected the location of the fibroblast on the proximal–distal axis along the upper and lower limbs, whereas differential expression of *HOXC* genes most strongly correlated with anterior–posterior location along the trunk, and expression of *HOXB* genes was associated with origin from internal organs rather than from skin (Figure 1c).

The apparent fidelity of *HOX* expression in adult fibroblasts has been tested by several functional experiments. First, extensive *in vitro* passage showed that fibroblasts maintained their distinct *HOX* expression patterns for >35 cell generations, and the anatomic site-specific *HOX* expression patterns of fibroblasts were not perturbed by soluble factors or direct cell–cell contact from heterotypic cells [5]. The fidelity of *HOX* gene expression in adult fibroblasts indicated that a powerful system of transcriptional memory is probably at work. Indeed, fibroblasts isolated from young versus old human donors – decades apart in age – show little difference in the expected pattern of position-specific *HOX* expression [5].

One function of the ongoing *HOX* expression is tissue homeostasis. Cells in the superficial layer of skin, termed the epidermis, are constantly turned over, and the human epidermis completely replaces itself every 28 days. How do newly generated epidermal cells know where they are located in the body, and whether they should make scalp hairs or thick, hairless skin for the palm? Work in our laboratory showed that the ongoing *HOX* expression in adult skin fibroblasts can provide this positional memory. Adult palmoplantar fibroblasts expressed *HOXA13*, which in turn activated a battery of genes. One of these genes encoded WNT5A, a morphogen that both promoted distal limb extension during development and instructed adult epidermal cells to differentiate toward the palmoplantar fate. In other words,

the same mechanism employed to pattern distal limb outgrowth early in development is used again to specify distal epidermal differentiation in adulthood [6]. These findings are also consistent with the idea that *HOX* genes can be important 'micromanagers' that orchestrate differentiation involving many different cell types and developmental pathways long after embryonic development [7].

The remarkable retention of *HOX* expression and positional identity is not limited just to fibroblasts. Smooth muscle cells [8] and skeletal muscle cells [9,10] are also organized by site-specific patterns of gene expression and, in some cases, also preserve features of the *HOX* code *ex vivo*. Fat deposits also have site-specific differences, as revealed by transcriptional profiling [11]. Mesenchymal stem cells (MSCs) reside in the bone marrow but can give rise to fibroblasts, chrondrocytes and bone. It seems that MSCs derived from different bones can have distinct developmental potentials, and these are correlated with anatomic site-specific differences in the expression of numerous *HOX* genes by the mesenchymal stem cells [12]. In sum, these data suggest that a subset of developmental regulators, particularly the *HOX* genes, not only governs what a segment of an embryo will become (e.g. the forelimb), but they also ensure the development of specific types of skin, muscle, nerve or fat that belong in that body segment.

Despite the strength of these observations, several important caveats remain. First, the position-specific pattern of *HOX* genes (termed the '*HOX* code') in adult cells is notably more sparse and simple compared with the *HOX* code in the embryo, which has been extensively studied in mouse development. For instance, although the developing limb bud shows at least three nested patterns of *HOX* expression along the proximal–distal axis in adult fibroblasts, which demarcate the upper arm, forearm and hand, only two patterns are evident that demarcate the hand from the remainder of the arm [4]. Likewise, whereas each segment of the spine is demarcated by sequential expression of additional *HOX* genes during embyrogenesis, the *HOX* expression in adult fibroblasts shows a biphasic pattern, which switches along the antero–posterior axis at the umbilicus. These discrepancies might have arisen for several reasons. The adult *HOX* code has been mapped in just a few cell types, and the combination of *HOX* expression from multiple cell types might generate a more detailed address code. Alternatively, perhaps maintenance of positional identity in adulthood requires a simpler set of address codes than during initial pattern formation and organogenesis. A second caveat is that, in some cases, the adult HOX expression pattern does not simply reflect a remnant of the embryonic expression pattern. As an example, whereas the *HOX* expression pattern in MSCs approximately follows the 3′–5′ colinearity rule based on the bone of origin, the *HOX* expression patterns in MSCs do not strictly match the *HOX* expression pattern of the individual bones, at least during bone development [12]. This discrepancy raises the possibility that some instances of *HOX* gene expression in adult cells might not reflect positional memory *per se* but, rather, could be examples of the *HOX* transcription factor being deployed for gene regulation in a manner that is independent of embryonic segmental identity.

Epigenetic memory of *Hox* genes

Classical genetics and biochemistry have previously identified powerful epigenetic mechanisms that maintain the appropriate ON and OFF state of *Hox* genes [13,14], and new epigenomic mapping efforts have provided new clues as to how positional identity can be faithfully transmitted from embryogenesis into adulthood and old age. Epigenetics refers to heritable changes in phenotype or gene expression caused by mechanisms other than changes in the underlying DNA sequence; these changes can persist through cell divisions. The trithorax family of factors encodes protein complexes that possess enzymatic activity to methylate Lys4 on the N-terminal tail of histone H3 (H3K4) and trithorax proteins, also known as mixed-lineage leukemia (MLL) proteins, are required for maintaining *Hox* gene activation [14]. Conversely, the Polycomb group proteins promote histone H3 Lys27 methylation, leading to *Hox* gene silencing [15].

Recent epigenomic mapping studies combine chromatin immunoprecipitation, a method that retrieves DNA sequences associated with a particular histone modification, with hybridization to tiling microarrays or high throughput sequencing to read out all of the DNA sequences associated with a particular histone modification or chromatin-binding factor [16]. These studies have provided the picture of the chromatin landscape of a typical gene; the promoter, gene start and gene body are each associated with a canonical pattern of nucleosome occupancy and histone modifications [17,18]. Di- and tri-methylated Lys4 on histone H3 is associated with the promoter and the first several hundred bases of transcribed genes. H3K27me3 and associated Polycomb group proteins occupy the promoters of a subset of transcriptional silent genes. Interestingly, the *HOX* loci emerge as an exception to this chromatin landscape (Figure 2a). In differentiated cells that possess transcriptional memory of *HOX*, such as fibroblasts, the *HOX* loci show extended chromosomal domains of histone modifications and occupancy by their cognate enzymes that encompass multiple adjacent *HOX* genes and intergenic regions [6]. The chromosomal domains of H3K4 and H3K27 methylation are mutually exclusive and can be programmed in a modular manner based on the anatomic origin of the cells [6]. Furthermore, numerous sites of extensive transcription of long non-coding RNAs (lncRNAs) are found within the *HOX* loci that interweave the intergenic and genic regions (Box 2). We recently discovered at least 231 transcribed regions within the four human *HOX* loci [6]. The majority of the *HOX* lncRNAs are expressed in a spatially patterned manner in the tissue and in fact constitute the predominant site-specific output of the *HOX* loci [6]. LncRNAs turn out to have key roles in configuring this epigenetic landscape. Evidence in *Drosophila* and mammalian studies showed that lncRNA transcription can activate or silence *Hox* genes *in cis* or *in trans* [6,19,20]. For instance, a lncRNA in the *HOXC* locus, named HOTAIR, is required *in trans* to silence across 40 kilobases of the *HOXD* locus in adult fibroblasts [6]. HOTAIR accomplishes this task through its interactions with the Polycomb repressive complex 2 (PRC2), targeting H3K27 trimethylation to *HOXD* [6] (Figure 2b).

Several features of the *Hox* chromatin landscape lend substantial robustness to its transcriptional memory. First, the broad domains of histone modifications favor faithful duplication of the chromatin landscape during DNA replication. As chromatin is unraveled during DNA replication, post-translationally modified histones should partition randomly to

both daughter strands, along with new, unmodified histone proteins. Thus, a single modified nucleosome has a probability of being lost during DNA replication, but this probability is greatly reduced for an extended chromatin domain ranging over kilobases and comprising dozens of homogenously modified nucleosomes [21]. Second, the concomitant association of the histone modification enzymes, namely Polycomb or MLL complexes, along with their cognate histone marks on chromatin ensures that any deviation from the desired chromatin pattern can be corrected; for instance, incorporation of an unmodified histone during DNA replication can be modified to the desired state. Furthermore, recent biochemical studies have uncovered families of histone demethylases, enzymes that reverse histone methylation events. Notably, a H3K27 demethylase is physically associated with MLL2 and MLL3 H3K4 methylase complexes [22-24], whereas a H3K4 demethylase is physically associated with the Polycomb H3K27 methylase complex [25]. These associations ensure the univalent nature of H3K4 versus H3K27 methylation in differentiated cells, and prevent any ambiguity in the conveyance of ON versus OFF state of the *Hox* genes. Finally, the extensive transcription of lncRNAs confers an additional element of stability. Because the *Hox* lncRNAs are co-regulated with the *Hox* genes and also control the chromatin states of the *Hox* loci, *Hox* lncRNAs function as a positive feedback loop to continually reinforce the chromatin state. *In cis*, extensive and interweaved transcription of lncRNAs can help maintain an active chromatin state with high RNA polymerase II occupancy across an extended chromosomal region (Figure 2). *In trans*, lncRNAs such as HOTAIR can also functionally demarcate silent chromatin domains on distantly located *Hox* genes and provide cross regulation over multiple *Hox* loci. These unique features can also explain in part why *Hox* genes have stayed together as compact loci during evolution and seem particularly resistant to insertion of repetitive elements [26]. Finally, because Hox lncRNAs are coordinately regulated with their neighboring *Hox* genes, ectopic activation of the Hox loci by *de novo* mechanisms in adult cells could also set the lncRNAs in motion, thereby creating an epigenetic state for their persistent expression.

Positional memory in wound healing and regeneration

Positional identity not only governs what a segment of an embryo will become – for example, the forelimb – but it also ensures the development of specific types of skin, muscle, nerve or fat that belong in that particular body segment. The retention of positional identity in adult differentiated cells might contribute to its faithful homeostasis but limits its plasticity, leading to a loss of regenerative ability in higher vertebrates. The importance of positional memory in regeneration is demonstrated elegantly in the freshwater flatworm planarian [27,28]. Planarians regenerate a head if the head is cut off, and regenerate a tail if the tail is cut off. The polarity changes in the adult organism are possible through modulation of levels of β-catenin, which is the output of the normal developmental Wnt pathway that specifies anterior–posterior polarity (tails have high β-catenin). Silencing of β-catenin after wounding resulted in the inappropriate regeneration of a head instead of a tail [27]. Furthermore, the silencing of this posterior signal can transform some adult heads into tails in uncut animals. These results again illustrate the existence of an ongoing regulatory program that specifies position identity, which maintains dynamic control of tissue homeostasis. Where a wound requires newly regenerated tissue to integrate with the old, this

presumably occurs through the use of sustained instructive cues [27,28]. Earlier studies showed that related planarian species do possess *Hox* genes, and that the *Hox* genes become ectopically expressed in the regenerated tissue [29]. Whether planarian *Hox* genes are required for proper regeneration has not been tested. Nonetheless, a robust system of positional identity and memory must be in place to sense what structures in the animal have been lost and to use the remaining cells to regenerate the missing structures. This regenerative ability is presumed to be dependent on the totipotent, somatic stem cells called neoblasts present in the adult planarian [30].

Do mammalian stem cells possess positional information? Embryonic stem cells seem to repress *Hox* gene expression and, in doing so, preserve their pluripotency [31] but it is simply not known if this is the case in adult stem cells. Tissues that undergo continual remodeling, such as skin and bone, must have a means by which stem cells are informed of the phenotypic identity of their ancestors. Perhaps this information comes from the environment [5]; alternatively, adult stem cells themselves might carry this ancestral information and one mechanism by which this might be accomplished is via retention of a *Hox* code.

Support for this hypothesis comes from the recent observation that adult skeletal stem or progenitor cells retain at least part of the *Hox* code established during fetal development [32]. Limb mesenchymal cells that are destined to form the lower leg bones (tibia and fibula) express several *Hox* genes during development, including *Hoxa11*. These bones undergo constant remodeling and yet, even so, *Hoxa11* expression persists in adult tibial osteoblasts and osteocytes [32]. Not all parts of the skeleton are derived from *Hox*-expressing cells: for example, mesenchymal cells in the first branchial arch that are destined to form the lower jaw (mandible) do not express any *Hox* genes during development [33] and this *Hox*-negative status is maintained into adulthood [32]. Clearly then, *Hox* expression is not a prerequisite for bone formation in either the embryo [33] or in the adult [32]. So what function(s) might this *Hox* code serve?

The role of *Hox* expression in adult stem or progenitor cells was addressed in a series of experiments that exploited two unique features of bone (Figure 3). First, bone has remarkable regenerative potential; and second, skeletal stem or progenitor cells are easy to isolate. Relative to other tissues, bones harbor an abundance of stem cells, which make it possible for our skeletons to undergo continual remodeling (our entire skeletons are replaced every decade or so) and also to repair without leaving behind any scar tissue. This process of bone repair follows the same general program as bone development and bone remodeling: skeletal stem cells give rise to transient amplifying osteo-chondro-progenitor cells that proliferate and then eventually differentiate into osteoblasts or chondrocytes. The only notable differences between bone remodeling and bone repair seem to be the stimulus for new bone formation (i.e. trauma versus hormonal regulation) and the spatial and temporal restriction of osteogenesis to a site of injury versus over the entire skeleton.

Taking advantage of these two attributes, Leucht and colleagues asked whether or not the *Hox* status of a bone is recapitulated during regeneration in mice [32]. *Hox*-expressing tibial bones were injured, and the osteo-chondro-progenitor cells occupying the injury site re-

expressed *Hoxa11*. Likewise, *Hox*-free mandibular bones were injured and, here, the osteo-chondro-progenitor cells in those injury sites seemed to retain their *Hox*-free status. Again, these results emphasize that *Hox* status does not confer differentiation potential onto cells (*Hox*-negative cells can make both bone and cartilage). However, when *Hox*-positive skeletal stem cells from the tibia were grafted into a *Hox*-negative (mandibular) injury, the grafted cells showed a persistence of *Hoxa11* expression in the *Hox*-negative environment of the mandible. Furthermore, the grafted cells failed to differentiate into osteoblasts and instead formed a cartilaginous callus. In the lexicon of bone healing, this could be referred to as 'scarring.' However, *Hox*-negative skeletal stem cells from the mandible readily adopted *Hoxa11* expression when transplanted into a *Hox*-positive tibial injury sites. These grafted cells formed a seamless regenerate of new bone, with no evidence of the cartilage 'scar.'

Together, these data demonstrate that at least some adult stem cells are equipped with a *Hox* code that is retained, even after transplantation. This positional identity is first established during the embryonic period [34,35] and seems to be rigorously maintained throughout the life of an organism and during the regenerative process. Elegant embryonic experiments first carried out using the quail–chick chimera system championed by Nicole le Douarin [36] demonstrate that, at least during early life, the *Hox* status of a cell confers upon it a sense of positional identity and that this identity is unchanged when cells are placed into a new environment. There is a 'flip side' to this: embryonic cells that normally lack *Hox* gene expression integrate seamlessly into their new environment. In accordance with this finding, the '*Hox*-free' condition of cranial neural crest cells has been strongly associated with an extraordinary plasticity that is a key contributor to the evolution of the craniofacial skeleton (for a review, see Ref. [37]). The finding that *Hox*-negative skeletal stem cells begin to express *Hox* genes when placed into a *Hox*-positive environment implies that injury sites have specific *Hox* codes. Furthermore, synchrony between the cells occupying the injury site and the injury environment itself could be a crucial component of normal healing. Future experiments will also have to directly test whether or not a disparity in *Hox* gene expression underlies the ability of any grafted cell to heal wounds more efficiently.

Beyond the *Hox* code, the pattern of *Hox* cofactors in the wound site or adult cells is probably also important. Hox proteins typically rely on transcriptional cofactors that refine and constrain their activities [38] (Box 1). For example, recruitment of one of the cofactors, a member of the forkhead family of transcription factors FoxP1, has recently been demonstrated to have a crucial role in fine-tuning motor-neuron diversification in mice [39,40]. It is conceivable that convergent and/or divergent activities of the Hox proteins and their cofactors in injured tissues and the injury environment contribute to the assembly of local gene regulatory networks that coordinate maintenance or changes in positional identity. Comparisons of the promoters of *Hox* genes and of their cofactors from different organisms should enable us to better understand how genes can be co-opted from one context, such as development, to be used in another, such as regeneration and repair [41].

Concluding remarks

The transcriptional memory of *Hox* genes is a double-edged sword for regenerative medicine. The persistence of positional cues might enable resident lineage-specific stem

cells to repair damaged tissues, but a mismatch of positional identity can prevent distantly located or grafted stem cells to participate in regeneration. A larger implication of these experiments is that the success or failure of grafted cells could be controlled by molecular features that distinguish one type of adult stem or progenitor cell from another. The histologic characteristics of skin grafts are carefully considered before grafting; perhaps the molecular characteristics should be as well. Likewise, physicians carefully consider the blood type of an individual before transfusion. Although once this classification scheme was based on the presence of two blood types, we now recognize 30 distinct characteristics. Future studies are needed to directly test whether adult stem cells actually retain a *Hox* code, and then whether these potential molecular differences are crucial features of successful tissue repair and regeneration. A corollary of this concept would imply that cells erased of their *Hox* expression can somehow regain plasticity found earlier in development, which is in fact observed in induced pluripotent stem cells [42,43]. Our increasing understanding of the epigenetic mechanisms operating on the *Hox* loci should make this goal feasible, perhaps by careful manipulation of appropriate lncRNAs or histone demethylases. A better understanding of the mechanisms of positional identity should be a key goal for regenerative medicine to enable the restoration of tissue function after injury, aging or disease.

Acknowledgments

The authors are supported by grants from the National Institutes of Health. H.Y.C. is supported by the California Institute for Regenerative Medicine.

References

1. McGonigle GJ, et al. Grappling with the HOX network in hematopoiesis and leukemia. Front Biosci. 2008; 13:4297–4308. [PubMed: 18508512]
2. Maroulakou IG, Spyropoulos DD. The study of HOX gene function in hematopoietic, breast and lung carcinogenesis. Anticancer Res. 2003; 23:2101–2110. [PubMed: 12894584]
3. Chang HY, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. Proc Natl Acad Sci U S A. 2002; 99:12877–12882. [PubMed: 12297622]
4. Rinn JL, et al. Anatomic demarcation by positional variation in fibroblast gene expression programs. PLoS Genet. 2006; 2:e119. [PubMed: 16895450]
5. Rinn JL, et al. A dermal HOX transcriptional program regulates site-specific epidermal fate. Genes Dev. 2008; 22:303–307. [PubMed: 18245445]
6. Rinn JL, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007; 129:1311–1323. [PubMed: 17604720]
7. Akam M. Hox genes: from master genes to micromanagers. Curr Biol. 1998; 8:R676–R678. [PubMed: 9768351]
8. Chi JT, et al. Gene expression programs of human smooth muscle cells: tissue-specific differentiation and prognostic significance in breast cancers. PLoS Genet. 2007; 3:1770–1784. [PubMed: 17907811]
9. Donoghue MJ, et al. Mammalian muscle cells bear a cell-autonomous, heritable memory of their rostrocaudal position. Cell. 1992; 69:67–77. [PubMed: 1555243]
10. Grieshammer U, et al. A transgene target for positional regulators marks early rostrocaudal specification of myogenic lineages. Cell. 1992; 69:79–93. [PubMed: 1313337]
11. Gesta S, et al. Evidence for a role of developmental genes in the origin of obesity and body fat distribution. Proc Natl Acad Sci U S A. 2006; 103:6676–6681. [PubMed: 16617105]
12. Ackema KB, Charite J. Mesenchymal stem cells from different organs are characterized by distinct topographic Hox codes. Stem Cells Dev. 2008; 17:979–991. [PubMed: 18533811]

13. Ringrose L, Paro R. Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development. 2007; 134:223–232. [PubMed: 17185323]
14. Schuettengruber B, et al. Genome regulation by polycomb and trithorax proteins. Cell. 2007; 128:735–745. [PubMed: 17320510]
15. Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer. 2006; 6:846–856. [PubMed: 17060944]
16. Collas P, Dahl JA. Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation. Front Biosci. 2008; 13:929–943. [PubMed: 17981601]
17. Barski A, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007; 129:823–837. [PubMed: 17512414]
18. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007; 448:553–560. [PubMed: 17603471]
19. Sanchez-Elsner T, et al. Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to Ultrabithorax. Science. 2006; 311:1118–1123. [PubMed: 16497925]
20. Petruk S, et al. Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx *in cis* by transcriptional interference. Cell. 2006; 127:1209–1221. [PubMed: 17174895]
21. Dodd IB, et al. Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell. 2007; 129:813–822. [PubMed: 17512413]
22. Lan F, et al. A histone H3 lysine 27 demethylase regulates animal posterior development. Nature. 2007; 449:689–694. [PubMed: 17851529]
23. Agger K, et al. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature. 2007; 449:731–734. [PubMed: 17713478]
24. Lee MG, et al. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science. 2007; 318:447–450. [PubMed: 17761849]
25. Pasini D, et al. Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. Genes Dev. 2008; 22:1345–1355. [PubMed: 18483221]
26. Lee AP, et al. Highly conserved syntenic blocks at the vertebrate Hox loci and conserved regulatory elements within and outside Hox gene clusters. Proc Natl Acad Sci U S A. 2006; 103:6994–6999. [PubMed: 16636282]
27. Gurley KA, et al. β-catenin defines head versus tail identity during planarian regeneration and homeostasis. Science. 2008; 319:323–327. [PubMed: 18063757]
28. Petersen CP, Reddien PW. Smed-βcatenin-1 is required for anteroposterior blastema polarity in planarian regeneration. Science. 2008; 319:327–330. [PubMed: 18063755]
29. Orii H, et al. The planarian HOM/HOX homeobox genes (Plox) expressed along the anteroposterior axis. Dev Biol. 1999; 210:456–468. [PubMed: 10357903]
30. Salo E. The power of regeneration and the stem-cell kingdom: freshwater planarians (Platyhelminthes). Bioessays. 2006; 28:546–559. [PubMed: 16615086]
31. Soshnikova N, Duboule D. Epigenetic regulation of Hox gene activation: the waltz of methyls. Bioessays. 2008; 30:199–202. [PubMed: 18293357]
32. Leucht P, et al. Embryonic origin and Hox status determine progenitor cell fate during adult bone regeneration. Development. 2008; 135:2845–2854. [PubMed: 18653558]
33. Creuzet S, et al. Negative effect of *Hox* gene expression on the development of the neural crest-derived facial skeleton. Development. 2002; 129:4301–4313. [PubMed: 12183382]
34. Grapin-Botton A, et al. Plasticity of transposed rhombomeres: *Hox* gene induction is correlated with phenotypic modifications. Development. 1995; 121:2707–2721. [PubMed: 7555700]
35. Couly G, et al. Determination of the identity of the derivatives of the cephalic neural crest: incompatibility between *Hox* gene expression and lower jaw development. Development. 1998; 125:3445–3459. [PubMed: 9693148]
36. Le Douarin NM. The avian embryo as a model to study the development of the neural crest: a long and still ongoing story. Mech Dev. 2004; 121:1089–1102. [PubMed: 15296974]
37. Helms JA, et al. New insights into craniofacial morphogenesis. Development. 2005; 132:851–861. [PubMed: 15705856]

38. Mann RS, Affolter M. Hox proteins meet more partners. Curr Opin Genet Dev. 1998; 8:423–429. [PubMed: 9729718]
39. Dasen JS, et al. Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, FoxP1. Cell. 2008; 134:304–316. [PubMed: 18662545]
40. Rousso DL, et al. Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. Neuron. 2008; 59:226–240. [PubMed: 18667151]
41. True JR, Carroll SB. Gene co-option in physiological and morphological evolution. Annu Rev Cell Dev Biol. 2002; 18:53–80. [PubMed: 12142278]
42. Wernig M, et al. *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature. 2007; 448:318–324. [PubMed: 17554336]
43. Okita K, et al. Generation of germline-competent induced pluripotent stem cells. Nature. 2007; 448:313–317. [PubMed: 17554338]
44. Joshi R, et al. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell. 2007; 131:530–543. [PubMed: 17981120]
45. Kmita M, Duboule D. Organizing axes in time and space; 25 years of colinear tinkering. Science. 2003; 301:331–333. [PubMed: 12869751]
46. McGinnis W, Krumlauf R. Homeobox genes and axial patterning. Cell. 1992; 68:283–302. [PubMed: 1346368]
47. Krumlauf R. Hox genes in vertebrate development. Cell. 1994; 78:191–201. [PubMed: 7913880]
48. Lumsden A, Krumlauf R. Patterning the vertebrate neuraxis. Science. 1996; 274:1109–1115. [PubMed: 8895453]
49. Pearson JC, et al. Modulating *Hox* gene functions during animal body patterning. Nat Rev Genet. 2005; 6:893–904. [PubMed: 16341070]
50. Magli MC. The role of homeobox genes in hematopoiesis. Biotherapy. 1998; 10:279–294. [PubMed: 9592016]
51. Hombria JC, Lovegrove B. Beyond homeosis–HOX function in morphogenesis and organogenesis. Differentiation. 2003; 71:461–476. [PubMed: 14641327]
52. Mann RS, Chan SK. Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. Trends Genet. 1996; 12:258–262. [PubMed: 8763497]
53. Moens CB, Selleri L. Hox cofactors in vertebrate development. Dev Biol. 2006; 291:193–206. [PubMed: 16515781]
54. Amaral PP, et al. The eukaryotic genome as an RNA machine. Science. 2008; 319:1787–1789. [PubMed: 18369136]
55. Zhao J, et al. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008; 322:750–756. [PubMed: 18974356]
56. Nagano T, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science. 2008; 322:1717–1720. [PubMed: 18988810]
57. Pandey RR, et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. Mol Cell. 2008; 32:232–246. [PubMed: 18951091]
58. Guttman M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009
59. Mercer TR, et al. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009; 10:155–159. [PubMed: 19188922]
60. Mattick JS, et al. RNA regulation of epigenetic processes. Bioessays. 2009; 31:51–59. [PubMed: 19154003]

Box 1

Hox genes and homeodomain transcription factors

In mammals, 39 *Hox* transcription factors, each containing a DNA-binding homeodomain, are clustered in four chromosomal loci (see Figure 1 in the main article). The encoded Hox proteins are transcription factors that, in association with accessory factors, bind to specific DNA sequences to activate or repress genes [46-49]. The *Hox* genes, first discovered in fruit flies, are instrumental in regulating body formation during development; they are highly conserved in evolution and control the polarity of the embryo, formation of the anterior–posterior body axis, and establishment of body segmentation during embryogenesis in all species, including humans. In addition, *Hox* genes in vertebrates have been shown to be involved in several other processes including patterning of the limb bud axis, hematopoeisis and organogenesis [50,51]. The *Hox* genes are sequentially expressed in a nested fashion along the anterior–posterior and proximal–distal axes, with the more 5′ genes being expressed more posteriorly and distally. Interestingly, organization of *Hox* genes on the chromosomes reflects their anterior–posterior expression in the body. Mutations of *Hox* genes in early development can transform one body part into another (termed homeosis) and a spate of recent studies have now emphasized the pervasive and enduring roles of positional identity in the control of cell fate decisions.

The exquisite DNA-binding specificity of the different Hox proteins has been demonstrated to be frequently dependent on their interactions with other DNA-binding proteins, which act as Hox cofactors [52]. These include the PBC and MEIS classes of three amino acid loop extension (TALE) homeodomain proteins – the PBC class comprises fly Extradenticle (Exd) and vertebrate Pbx homeoproteins, whereas the MEIS class includes fly Homothorax (Hth) and vertebrate Meis and Prep homeoproteins [51]. These cofactors have pervasive roles as regulators of Hox activity [38,53]. Although some of their functions are clearly *Hox*-dependent, others are less obviously so, suggesting that these proteins might function more broadly to modulate the activities of transcription factor complexes.

Box 2

LncRNAs

Ranging from several hundred bases to dozens of kilobases, lncRNAs are a recently recognized class of RNA-only genes that function by means other than serving as a template for protein synthesis. Recent efforts at genome annotation using multiple approaches, such as expression profiling, chromatin state maps and compilation of expressed sequence tags, have consistently shown that the genome is pervasively transcribed, including many sequences that do not seem to code for proteins (for a review, see Ref. [54]). LncRNAs are typically transcribed by RNA polymerase II, spliced and polyadenylated. Many lncRNAs are interleaved with protein-coding genes in sense and antisense orientations, but thousands of other lncRNAs are clearly demarcated, independent transcriptional units located away from other genes. One emerging functional theme of lncRNAs is their involvement in controlling chromatin states, thereby affecting gene expression. Dosage compensation of female X chromosomes [55], parental-specific expression of imprinted genes [56,57] and epigenetic regulation of *HOX* genes [6,19,20] all involve specific lncRNAs that can act *in cis* or *trans* to regulate the chromatin state of diverse target genes (Figure 2b in the main text). More recently, functional genomics approaches have been used to demonstrate a diverse range of roles for lncRNAs in processes from embryonic stem cell pluripotency to cell proliferation [58]. The characterization of the lncRNAs will no doubt have a considerable impact on our understanding of the genetic programming of complex organisms during evolution, development and disease [59,60].

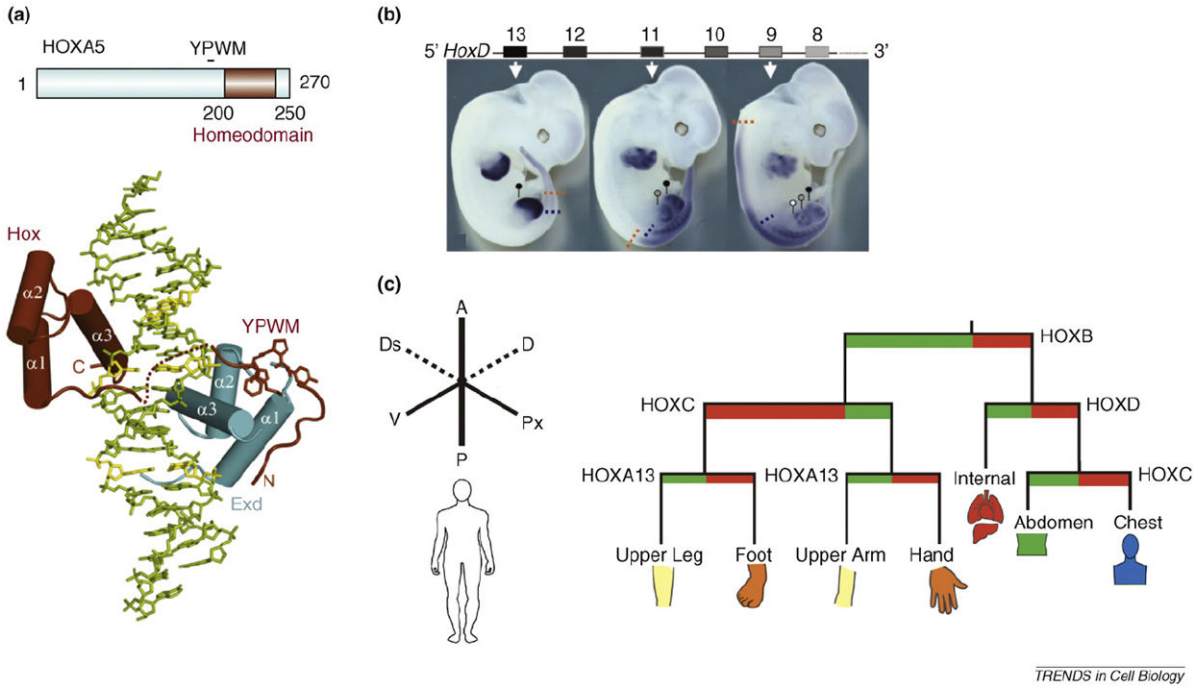

Figure 1.

Hox genes and positional identity. **(a)** *Hox* genes belong in the family of homeodomain-containing transcription factors. The domain organization of *HOXA5* and the crystal structure of the homologous fly Scr *Hox* protein (brown) in complex with Exd (light blue), binding to DNA (green) are shown. Exd is a *Hox* cofactor that acts as an accessory DNA-binding factor. Homeodomains are helix-turn-helix motifs (represented by α1, α2 and α3) that bind specific DNA sequences. Sequences N-terminal of the homeodomain, such as the YPWM motif, are involved in protein–protein interaction with accessory DNA-binding factors such as Exd. Part (a) adapted, with permission, from Ref. [44]. **(b)** Expression of three *HoxD* genes in developing mouse embryos by *in situ* hybridization. The schematic above the *in situ* pictures represents the *HoxD* locus – the individual numbers refer to different *HoxD* genes. The white arrows indicate that HoxD-13, -11 and -9 are shown in the three panels, from left to right. The orange and blue dotted lines represent the neural tube and paraxial mesoderm, respectively – the lines are there to highlight the tissue-specific mechanisms for colinear gene regulation: *Hox* genes tend to have a very distinctive organization in the genome, being arranged in gene clusters in which the order of the genes within the cluster corresponds to (or is 'colinear with') some aspect of the gene expression. In addition to spatial colinearity, in vertebrates there is also a temporal colinearity, where the 3′–5′ arrangement of genes also reflects the temporal order in which they are activated during development. The expression domains of *Hox* genes towards the 3′ UTR seem closer to the head and closer to the trunk (dotted lines), whereas the *Hox* genes towards the 5′ UTR are expressed closer to the tail and closer to fingers or toes on the limbs (pins). For example, the domains of expression of *HoxD*13 in both the neural tube (marker by the orange line) and the paraxial mesoderm (the blue line) are quite similar, whereas in *HoxD*11 and to a greater extent *HoxD*9, one can really appreciate the variation in the transcript domains of the same gene in different tissues. The pins represent nested domains of expression, from the

highest level (black) to lowest (white), again illustrating the colinearity principle. Part (b) reproduced, with permission, from Ref. [45]. **(c)** *HOX* genes as the address code of the human body. As illustrated by the decision tree, differential expression of *HOX* genes reflects the anatomic origin of adult human fibroblasts along the anterior (A)–posterior (P), dorso (D)–ventral (V), and proximal (Px)–distal (Ds) axes, in addition to their origin from cutaneous versus internal organs. Red indicates high expression; green indicates low expression.

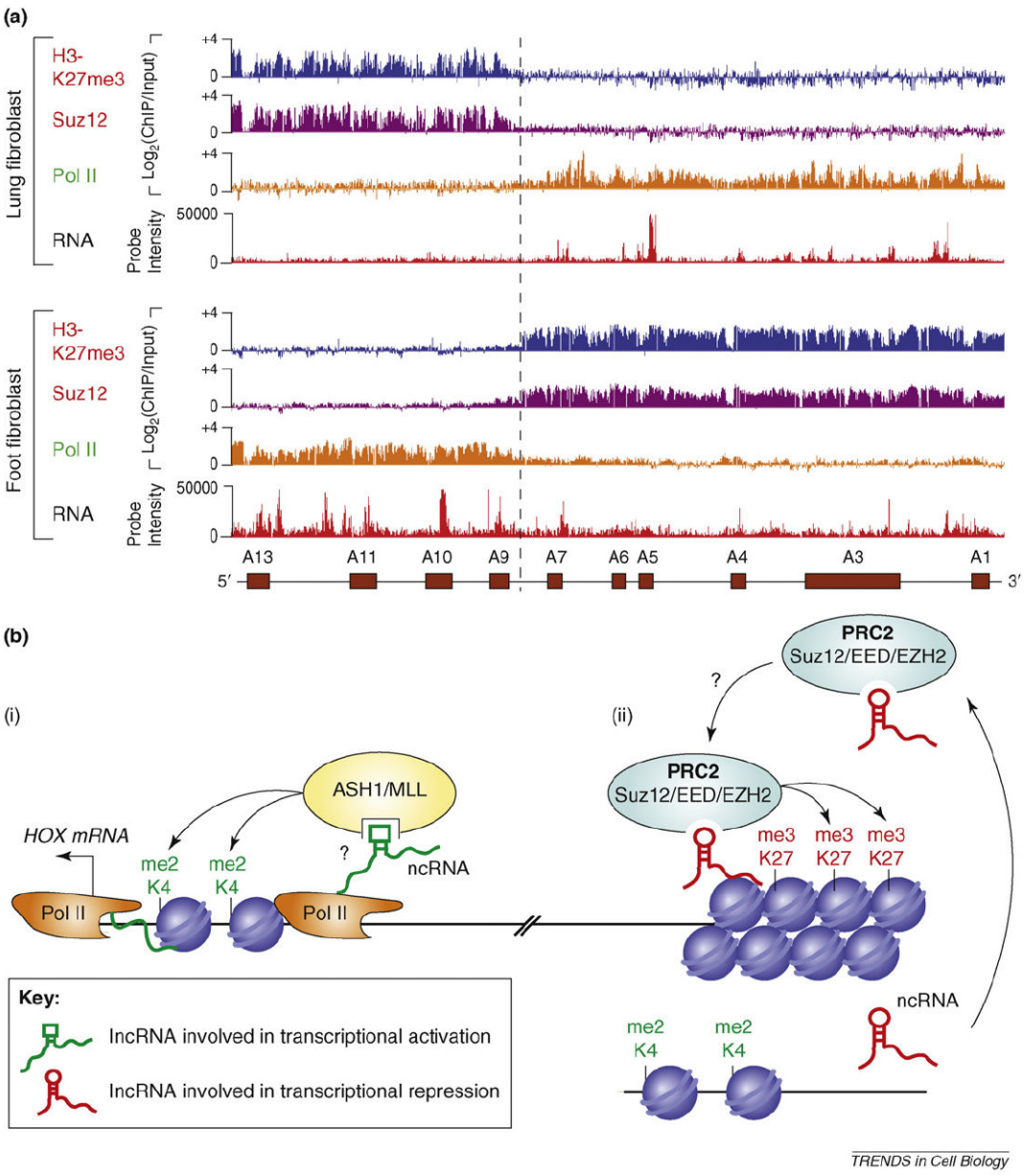

Figure 2.

Epigenetic landscape of *Hox* genes. **(a)** Chromatin state map across~100 kilobases of the human *HOXA* locus (X axis) obtained from chromatin IP on microarray (ChIP-chip) experiments, a technique for isolation and identification of the DNA sequences occupied by specific DNA-binding proteins in cells. Y axes indicate the occupancy of the indicated proteins (represented by the $\log_2$ ratio of ChIP over input) or expression of RNAs (represented by a linear scale of hybridization intensity). Suz12 is a component of the PRC2 that mediates histone H3 Lys27 methylation (H3K27me3). RNA polymerase (Pol) II broadly occupies the regions that are also transcribed and occupied by H3K4me3 (not shown). Note that fibroblasts from two different anatomic origins (lung and foot) can program the *HOXA* locus in diametrically opposite ways but along the same boundary. The

HOXA locus is depicted at the bottom of the map, with *HoxA13* at the distal 5′ end and *HoxA1* at the promixal 3′ end. The dashed line highlights the boundary of opposite configurations of chromatin modification and intergenic transcription. **(b)** Model of chromatin state regulation by lncRNAs. lncRNAs can affect chromatin state *in cis* or *in trans*, and in gene activation or silencing, via histone-modification enzymes. (i) lncRNAs might increase the accessibility of Trithorax group proteins such as ASH1 (absent, small, or homeotic discs 1) or MLL or directly recruit them, leading to trimethylation of histone H3K4 and transcriptional activation of downstream targets such as the *Hox* genes. (ii) By contrast, lncRNAs might also target the PRC2 proteins to trimethylate H3K27 at a distance and render the target genes transcriptionally silent.

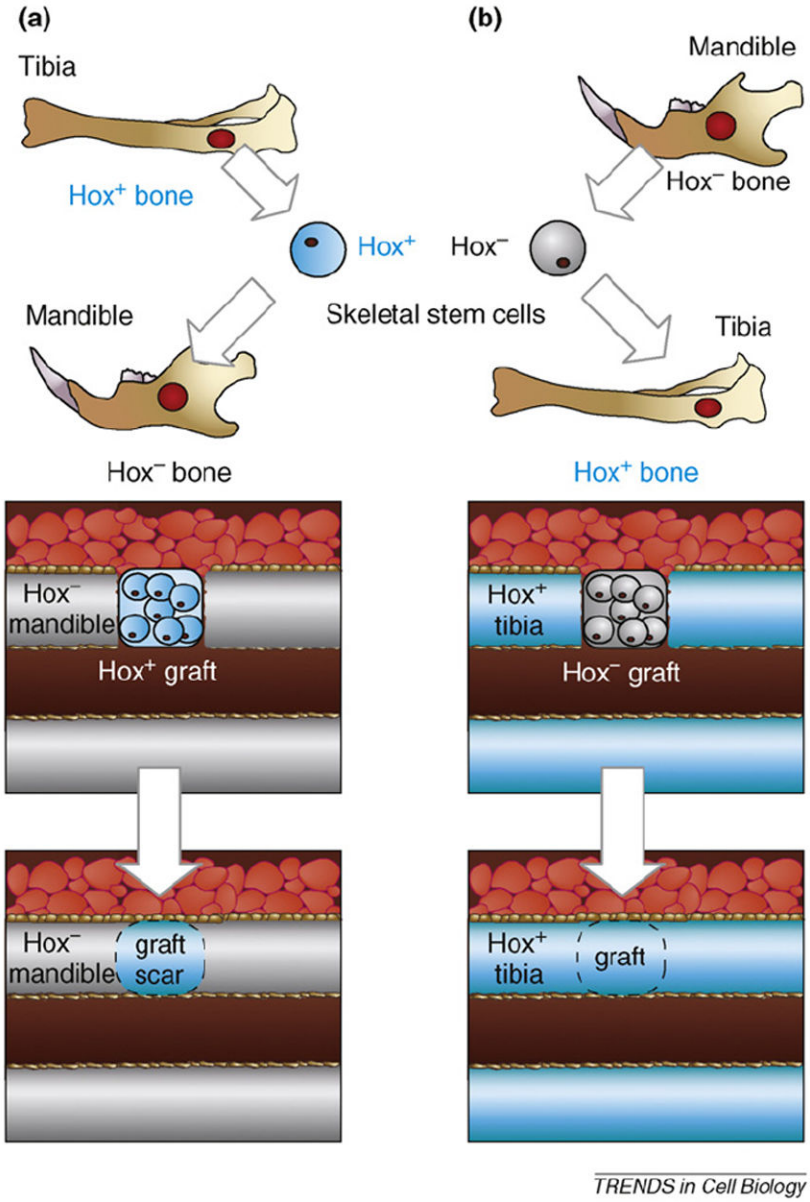

Figure 3.

Hox status and bone regeneration in mouse. Cells expressing *Hox* genes (for example *Hoxa11*) are in light blue and those not expressing *Hox* are indicated in light grey. Hox^+ skeletal stem cells can only heal orthotopic Hox^+ bone injury site (tibia). **(a)** Hox^+ skeletal stem cells cannot repair a Hox^- injury site (mandible). **(b)** Conversely, Hox^- skeletal stem cells will express the ectopic *Hox* gene when transplanted into a Hox^+ injury site and regenerate the ectopic bone.