Molecular mechanisms of selector gene function and evolution

Richard S Mann* and Sean B Carroll†

Selector proteins regulate the formation and identity of animal body regions, organs, tissues, and cell types. Recent studies have focused on the regulation of the DNA binding and transcriptional regulatory activity of this special class of transcription factors. Elucidation of the architecture of selector-regulated target gene enhancers and gene networks, and comparative studies of selector protein function are providing important insights into the evolution of development and morphology.

Addresses
*Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA; e-mail: rsm10@columbia.edu
†Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin, 53706, USA; e-mail: sbcarrol@facstaff.wisc.edu

Current Opinion in Genetics & Development 2002, **12**:XXX–XXX

0959-437X/02/$ – see front matter
© 2002 Elsevier Science Ltd. All rights reserved.

Abbreviations

Antp	*Antennapedia*
Dll	Distalless
dpp	decapentaplegic
en	engrailed
Exd	Extradenticle
ftz	*fushi tarazu*
hth	*homothorax*
sal	*spalt*
Ubx	Ultrabithorax
Vg	Vestigial

Introduction

The selector gene concept was first articulated by Antonio Garcia-Bellido in 1975 [1] to describe a class of genes that governs the fates of groups of cells. In its original formulation, influenced primarily by studies in *Drosophila* on the behavior of tissues lacking the activity of genes such as engrailed (en) or Ultrabithorax (Ubx), selector gene function was intimately tied to properties associated with lineage compartments [1]. More recently, the term has been used to describe the function of genes that specify cell, tissue, organ, as well as regional identity in animals (Figure 1). It has also been applied more liberally to describe various transcription factors with somewhat less discrete roles. Although we do not wish to split semantic hairs, selector genes do possess special developmental properties and are of particular evolutionary interest, and in order for the concept to be useful, it is important to distinguish selectors from other transcription factors. Here, we include those genes with discrete roles governing the formation and identity of cell types, organs, and other body parts, such as 'classical' selectors and genes more recently understood to possess selector or selector-like roles (Figure 1; Table 1).

It is clear that selector genes perform diverse roles during development. Consequently, removing their activities produces a wide variety of defects that depend in part on which type of selector is affected. Field-specific selectors, such as Eyeless, Vestigial (Vg), or Distalless (Dll), are required for the growth of their respective fields [2–4]. Similarly, in the absence of organ, cell-type, or tissue-specific selectors — such as the mesoderm-specific selector Twist — there is usually a loss of these specific structures (e.g. [5]). In contrast, inactivating region-specific selectors, such as the Hox genes, often results in the transformation of one body part into another. These transformations are sometimes as a result of cross regulation between different region-specific selectors. For example, loss of the Hox gene *Antennapedia* (*Antp*) from the second thoracic leg results in a transformation to antenna and to the derepression of the antennal selector *homothorax* (*hth*) [6,7••]. These observations lead to the idea that there exists an underlying ground plan upon which region-specific selector proteins operate. This ground plan has been examined for the *Drosophila* ventral appendage [8], which, in the absence of regional selector inputs, develops as a leg-like appendage but, unlike normal legs, has only two segments along its proximal–distal axis. Similarly, the forewings of *Drosophila* develop in the absence of Hox input, suggesting that this structure represents the ground plan for the dorsal appendage, which is then shaped into a hindwing (haltere) by the Hox protein Ubx [9]. Interestingly, in both of these examples, in the absence of Hox input, the proximal–distal axes form and field-specific selectors (Dll and Vg, respectively) are expressed normally in the ground plan. Thus, from this morphological point of view, the different types of selectors appear to work independently.

Over the past several years, considerable progress has been made in understanding how individual selector proteins act, both at a molecular level and in broader developmental terms. For the most part, the dramatic effects of selector gene mutations result from either the collapse or perturbation of networks of selector-regulated genes, the architectures of which have only recently begun to come into view. There are two primary components of these networks to consider: the selector proteins themselves and the diverse mechanisms through which their activities are regulated; and the *cis*-regulatory DNA elements that act as developmental integrators of selector protein

function and as units of evolution in the process of morphological diversification.

Here, we focus on recent work that has elucidated some of the molecular mechanisms that underlie selector gene function and selector-regulated *cis*-regulatory elements. We begin by reviewing general features of selector-regulated enhancer architecture. We examine the mechanisms that govern selector protein DNA binding and activity, and highlight both co-factor dependent and independent modes of target gene regulation. Finally, we discuss the molecular evolution of selector proteins and the origin and diversification of selector-regulated networks.

Enhancer architecture

When the selector gene hypothesis was first introduced, it was not clear what types of functions these genes would encode. Using even the most liberal definition of a selector gene, we now know that they all encode proteins that regulate transcription. Consequently, an understanding of how selector genes function requires both the dissection and analysis of directly regulated target gene enhancers. From the work of many labs on various selectors, we now have in hand several elements that are regulated directly by selector proteins (Table 1). In this section, we describe some common themes that have emerged from the analysis of these enhancers that may provide clues to how selector proteins function *in vivo*.

Synergy and combinatorial control

It is well established that enhancers are **[Author: Ok?]** integrators of multiple inputs, and selector-targeted enhancers are no exception. In all cases that have been examined, selector inputs must be combined with other inputs to produce a transcriptional response. The available examples also make it apparent that selector-regulated enhancers often integrate selector inputs with inputs from cell–cell signaling pathways (e.g. [10••–13••,14,15]) (Figure 2). Because signaling pathways are broadly deployed and are not sufficient to impart either tissue- or cell-type specificity, selector proteins are the key collaborators with signaling pathways in the specification of structures or cell types. Selector-regulated enhancers appear to integrate both of these inputs.

Combinatorial control was part of the original concept of selector genes [16]. In that context, combinatorial control was meant to suggest that combinations of selector proteins are required to make a particular structure. For example, the posterior haltere in the fly requires three selector genes: the field selector *vestigial* (*vg*) to specify dorsal appendage identity (wing and haltere), the regional selector *Ubx* to specify haltere instead of wing, and the compartment selector *en* to specify posterior compartment identity. There is now evidence that this type of combinatorial control can be integrated into single enhancers. For example, a wing-specific enhancer from the *spalt* (*sal*) gene is both activated by the Vg–Sd **[Author: Care to define Sd?]** selector complex and also repressed in the haltere by the Ubx Hox selector protein, and this enhancer appears to directly integrate both of these inputs [17].

Different modes of selector protein DNA binding

Some selector proteins have the capacity to bind DNA in multiple ways that impact their transcriptional regulatory properties or DNA-sequence specificities. Pit-1, a POU domain protein that is critical for the generation of specific cell-types in the mammalian pituitary gland, binds to two alternative DNA sequences that differ by 2 base pairs in the spacing between the POU-specific and POU-homeodomain binding sites [18••]. This spacing difference is thought to induce an allosteric change in Pit-1 that is important for it to mediate repression in one pituitary cell type. Sd, the DNA binding component of the Vg–Sd selector complex in *Drosophila*, has also been shown to exhibit a preference for different binding sites that depends on either the presence or absence of Vg [19]. The Hox homeodomain proteins often bind DNA together with two homeodomain protein co-factors which in flies are called Extradenticle (Exd) and Hth (e.g. lab550; Figure 2b). Exd and Hth appear, on some targets, to alter Hox DNA binding specificities (see [1,20] for reviews) whereas, on other targets (see below), Hox proteins may bind without co-factors. Thus, DNA- or co-factor-induced conformational changes apparently give selector proteins added flexibility and selectivity on target gene enhancers.

Although Exd and Hth play a critical role in some Hox-target enhancers, these co-factors are not present in cell nuclei in the distal part of the legs, wings, and halteres [21,22]. Moreover, genetic experiments demonstrate that *exd* has no required function there, even though Hox proteins confer appendage-specific morphologies [22]. Ubx-mediated repression of the *sal* enhancer in the haltere provides an example of an Exd- and Hth-independent mechanism for a Ubx-mediated function. In this enhancer, multiple Hox-binding sites contribute in an additive way to repression [17]. Although it is difficult to rule out that unknown co-factors are involved, no obvious flanking sequences are shared between individual Hox-binding sites, suggesting that multiple DNA-bound Ubx monomers might be sufficient for this Hox input. In fact, the presence of multiple binding sites for selector proteins may be a common strategy among the enhancers that have been characterized to date. For example, there are four Tinman binding sites in the muscle progenitor enhancer from the *even-skipped* gene [10••,12••] (Figure 2) and multiple Sd-binding sites in several wing-specific enhancers [11••]. Similarly, targets of Pha-4, which is required for pharyngeal development in *Caenorhabditis elegans*, usually have multiple Pha-4-binding sites [23••]. In addition, enhancers with higher-affinity Pha-4-binding sites are active earlier than those with lower affinity sites. These data suggest that some selector proteins may function in a cofactor-independent

manner by binding to multiple sites within 0; enhancers [Author: Ok?].

These observations raise an interesting question; why are Hox proteins dependent in some cases on Exd and Hth but in other cases do not require these co-factors? To date, the enhancers that are regulated by Hox proteins in the apparent absence of co-factors are active in the appendages as opposed to the main trunk of the fly. Hox functions in the appendages appear to be less specific than they are in the trunk; Abdominal-A (Abd-A), an abdominal Hox protein, has the capacity to transform wing into haltere (and presumably repress *sal*) just as well as Ubx. In contrast, Ubx and Abd-A act to produce distinct morphogenetic outputs in the trunk [24,25]. Thus, there may be a stricter requirement for selector specificity — and therefore co-factors — in the trunk than in the appendages. The test of this idea must await the identification of additional Hox target enhancers that are active in either the trunk or the appendages.

Repression versus activation by selector proteins

For a selector protein to govern the identity or formation of an organ or body part, it often acts as both a transcriptional repressor as well as an activator. This is especially clear for the Hox selectors. Ubx, for example, activates *dpp* in the visceral mesoderm but represses *Dll* in the ectoderm [26–29]. Recently, two mechanisms have been suggested to account for how *Drosophila* Ubx repression functions have been modulated during evolution: one dependent on phosphorylation and the other on the acquisition of a glutamine/alanine-rich domain [30••,31••]. Whatever the mechanism that allowed *Drosophila* Ubx to become a repressor of *Dll*, this mechanism must still allow Ubx to also act as an activator of other targets. Therefore, a challenge for the future will be to understand how the repressor versus activator functions of Ubx, as well as other selector proteins, are regulated at different target enhancers.

Selector-regulated gene networks are complex

Many selectors regulate genes at multiple steps in the definition of specific organs and cell types. For example, *twist*, a gene that is required for the formation of the mesoderm in flies, appears to target genes at many steps in the definition of mesodermal cell types, including other, more subordinate, selector genes (such as *tinman* and dMef-2) as well as genes required for terminal differentiation of specific mesodermal cell types (such as *myosin61F* and *blow*, a gene required for myoblast fusion) [32–34]. For some target enhancers, such as the *eve* [Author: even-skipped?] muscle enhancer, Twist may even collaborate with one of its own targets, Tinman, which also behaves as a selector protein [10••] (Figure 2a). Similarly, the two Pax-6 genes in *Drosophila*, *eyeless* and *twin of eyeless*, appear to directly regulate genes encoding transcription factors that are high up in the eye development hierarchy, such as *sine oculis*, as well as terminal differentiation genes, such as *rhodopsin1* [35,36••,37]. In fact, Garcia-Bellido [16] anticipated that selector genes would regulate genes required for terminal differentiation (he called them 'realizator genes'); what he did not anticipate, and what we are now beginning to appreciate, was how many additional layers of subordinate selector genes and other transcription factors exist between the tops and bottoms of these hierarchies.

Evolution of selector genes and selector-regulated networks

Comparative analysis of selector-gene distribution and function in arthropods, vertebrates and other taxa has demonstrated that many selectors have a deep evolutionary history, dating to at least the last common ancestor of bilaterians (reviewed in [38]). The discovery of this shared set of genes, while satisfying in terms of identifying some universal features of animal development, presents an evolutionary paradox. Namely, if the selector genes that govern animal form are so conserved, how does animal diversity arise? Several mechanisms have been identified including: first, the duplication and divergence of selector genes; second, changes in the regulation of selector genes; third, changes in selector protein sequence and function; and fourth, evolutionary changes in the sets of target genes regulated by individual selectors. We discuss a few examples of the molecular evolution of selector proteins and some general inferences that can be drawn about the evolution of their target gene networks.

Evolutionary changes in selector protein function: the gain or loss of functional motifs [Author: Please check that hierarchy of headings and subheadings is correct]

The last common ancestor of arthropods possessed a minimum of ten *Hox* selector genes [38]. These included orthologs of the eight classical *Hox* genes known from *Drosophila*, as well as the *Hox3* and *fushi tarazu* (*ftz*) genes. The latter two genes lost their ancestral homeotic roles and acquired novel functions during insect evolution (Figure 3). Recent work has elucidated the evolutionary interval during which these functional changes took place and identified some of the critical protein residues involved in their evolution. In the case of the *Hox3* gene, it appears that a single *Hox3*-derived gene with characteristics of both the *zerknüllt* and *bicoid* genes existed in lower non-Cyclorrhaphan flies. In the stem lineage of Cyclorrhaphan flies, this gene was duplicated and subsequently diversified to give rise to the respective *zerknüllt* and *bicoid* genes known in *Drosophila* [39] (Figure 3).

In *Drosophila*, *ftz* functions as a pair-rule segmentation gene and specifies the identity of a small set of neurons. It appears that the ancestral *ftz* gene also had a *Hox*-type role, as indicated by its regional expression in mites [40], spiders [41], and grasshoppers [42]. Comparative analysis of *ftz* orthologs from a grasshopper [43••] and a beetle [44•] suggest that *ftz* lost its homeotic function or

capacity since the divergence of the beetle and fly lineages, because the beetle *Ftz* protein retains its homeotic potential when expressed in flies. Interestingly, the beetle Ftz protein has also retained the YPWM motif involved in Exd interactions while this motif has degenerated in the fly protein suggesting this motif be particularly important for maintenance of homeotic function [44•]. By contrast, the grasshopper Ftz protein exhibits weak segmentation potential compared with fly Ftz [43••,44•] and the grasshopper protein lacks the LXXLL motif that mediates interactions between Ftz and the critical Ftz-F1 co-factor [45••,46,47]. The beetle protein possesses this motif and shows greater segmentation activity [44•]. These observations suggest that the gain of segmentation activity and the LXXLL motif preceded and/or was accompanied by the loss of homeotic activity and the YPWM motif in the course of insect evolution.

The gain of functional motifs in a *Hox* selector protein, while retaining *Hox* selector function, has recently been described for the arthropod Ubx protein (Figure 3). Functional and phylogenetic analyses indicate that the C terminus of the Ubx protein became extended in length in the crustacean lineage to function as an activity modulating domain, probably via acquisition of serine and threonine **[Author: Ok?]** residues that are dephosphorylated when the Ubx protein acts as a repressor [30••]. In the insect lineage, a sister group to the crustacea, it appears that this modulatory domain was replaced by a glutamine/alanine rich motif that potentiates Ubx repressor function [17]. It is very likely that further functional comparisons of Hox and other selector proteins between lineages will reveal additional lineage-specific gains and losses in functional motifs that play a role in the differential regulation of target genes among taxa and perhaps in morphological diversification.

Evolution of selector-regulated networks

The elucidation of selector-regulated gene networks is in its early stages. Only a few direct targets of any particular selector are known in any model taxon, let alone in other related taxa. However, from the anatomy of selector-regulated enhancers and studies thus far on the general architecture of selector-regulated networks we can draw a few inferences about the evolution of target gene regulation and morphology. With regard to all selectors, it is abundantly clear that most act by regulating potentially very large sets of target genes. Studies of the cell-type regulators Mef-2 [48] and Crx [49••] or the organ-specific selector Pha-4 [23••] suggest that, at a minimum, dozens of genes required for cell-type function are directly and independently regulated by individual selectors. Similarly, it would appear that *Hox* and field-specific selectors regulate substantial sets of genes in any single tissue. It is reasonable to infer that the assembly of these target gene sets, and by extension, the morphology of the structures governed by these selectors, occurred by stepwise addition of individual genes to these sets via acquisition of selector binding sites in their respective *cis*-regulatory elements.

There may be a significant difference, however, in the evolutionary diversification of target gene sets regulated by different classes of selectors. Namely, we expect, and some evidence supports the notion, that cell-type specific gene networks or batteries may be more extensively conserved throughout metazoan evolution whereas Hox-regulated target gene sets may be more divergent. A muscle cell, for example, requires many structural proteins and enzymes to endow this cell type with its characteristic properties and some of these Mef-2-regulated target genes are conserved among flies and vertebrates [48]. By contrast, the story of animal diversification along the primary body axis is largely one of differentiating serially repeated, homologous structures such as segments, appendages, and vertebrae, and so on. Different taxa are characterized by changes in the number, size, shape, and function of serial structures and Hox proteins play a central role in the development and evolution of axial morphology [38]. *Hox*-regulated gene sets may, therefore, be considerably more divergent among taxa than cell-type-specific gene batteries and there is some evidence that morphologically divergent, homologous structures differ in their *Hox*-regulated gene sets [50].

Because of their major influence on axial morphology, there is great interest in understanding, in detail, how *Hox* regulation of a target gene is acquired. From the standpoint of statistical probability, the evolution of large compound or trimeric *Hox* binding sites in one step within *cis*-regulatory elements is an improbable event [17]. It is more appealing to suggest that simpler, perhaps single *Hox* monomer sites, may function as intermediates in the gradual molecular evolution of *Hox*-regulated enhancers (Figure 4). Comparative studies of *Hox*-regulated enhancers, and detailed investigation of potential variation in *Hox*-regulation will shed critical light on the origin, maintenance, and diversification of *Hox*-regulation.

Conclusions and perspectives

The selector gene concept has been very fruitful in formulating ideas about the genetic regulatory logic that governs developmental decisions at the cell, tissue, organ, and body plan level. Although initially stemming from work on *Drosophila*, it is clear that these ideas are useful for thinking about developmental logic throughout the animal kingdom. New tools for elucidating large networks of selector-regulated target genes such as microarrays and bioinformatic algorithms will eclipse the previously standard 'one protein, one target' approaches. However, much work is still required to understand the nuances of selector protein DNA binding and the post-translational modifications and co-factor interactions that modulate their activities. Contrary to long-held assumptions, orthologous proteins are not necessarily functionally equivalent in that taxon-specific changes in regulatory motifs may occur. It is

very likely that subtle enhancer-specific and context-specific interactions have played a role in the diversification of selector protein functions and remain to be elucidated.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Mann RS, Morata G: **The developmental and molecular biology of genes that subdivide the body of *Drosophila*.** *Annu Rev Cell Dev Biol* 2000, **16**:243-271.
2. Quiring R, Walldorf U, Kloter U, Gehring WJ: **Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans.** *Science* 1994, **265**:785-789.
3. Williams JA, Paddock SW, Carroll SB: **Pattern formation in a secondary field: a hierarchy of regulatory genes subdivides the developing Drosophila wing disc into discrete sub-regions.** *Development* 1993, **117**:571-584.
4. Cohen SM, Jürgens G: **Proximal-distal pattern formation in Drosophila: cell autonomous requirement for Distal-less gene activity in limb development.** *EMBO J* 1989, **8**:2045-2055.
5. Baylies M, Bate M: **twist: a myogeneic switch in *Drosophila.*** *Science* 1996, **272**:1481-1484.
6. Struhl G: **A homoeotic mutation transforming leg to antenna in *Drosophila.*** *Nature* 1981, **292**:635-638.

••7. Casares F, Mann RS: **Control of antennal versus leg development in *Drosophila*.** *Nature* 1998, **392**:723-726.

This work examines the ground state for the ventral appendage in *Drosophila* by examining appendages that have no Hox (leg-determining) or *hth* (antennal-determining) inputs. The work also shows that the expression patterns of the Notch ligands Serrate and Delta, which are important for leg segmentation in flies, are regulated by these region-specific selectors.

8. Casares F, Mann R: **The ground state of the ventral appendage in *Drosophila.*** *Science* 2001, **293**:1477-1480.
9. Weatherbee S, Halder G, Hudson A, Kim J, Carroll S: **Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere.** *Genes Dev* 1998, **10**:1474-1482.

••10. Halfon MS, Carmena A, Gisselbrecht S, Sackerson CM, Jiménez F, Baylies MK, Michelson AM: **Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors.** *Cell* 2000, **103**:63-74.

This paper, together with [12••] [Author: Ok?], dissected the direct inputs into the *eve* muscle progenitor enhancer in *Drosophila*. The work elegantly shows how multiple inputs, including two selectors (Twi and Tin) and three signaling systems (Wg, Dpp, and ras/MAPK) collaborate to directly control the activity of this enhancer.

••11. Guss KA, Nelson CE, Hudson A, Kraus ME, Carroll SB: **Control of a genetic regulatory network by a selector gene.** *Science* 2001, **292**:1164-1167.

This paper illustrates that the Vg/Sd field selector system in flies collaborates with multiple signaling systems, including Dpp (Mad/Madea) and Notch (Su[H]), to generate distinct, wing-specific expression patterns. In addition, the activity of synthetic regulatory DNAs show that both inputs are necessary and sufficient to activate transcription.

••12. Knirr S, Frasch M: **Molecular integration of inductive and mesoderm-intrinsic inputs governs *even-skipped* enhancer activity in a subset of pericardial and dorsal muscle progenitors.** *Dev Biol* 2001, **238**:13-26.

This paper, together with [10••] [Author: Ok?], dissected the direct inputs into the *eve* muscle progenitor enhancer in *Drosophila*. The work elegantly shows how multiple inputs, including two selectors (Twi and Tin) and three signaling systems (Wg, Dpp, and ras/MAPK) collaborate to directly control the activity of this enhancer.

••13. Marty T, Vigano A, Ribeiro C, Nussbaumer U, Grieder N: **A HOX complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer.** *Development* 2001, **128**:2833-2845.

This paper dissects the lab550 enhancer, which synergistically integrates a Hox input (via a Lab/Exd/Hth trimeric binding site) with a Dpp input (via Mad and Cre binding sites). The Dpp input is surprisingly complex; a putative repressor attenuates the Dpp response and makes the enhancer Hox-dependent, and a GATA factor, perhaps the tissue-specifying gene *serpent*, is also required for activity.

14. Hepker J, Blackman R, Holmgren R: **Cubitus interruptus is necessary but not sufficient for direct activation of a wing-specific decapentaplegic enhancer.** *Development* 1999, **126**:3669-3677.
15. Affolter M, Mann R: **Development. Legs, eyes, or wings — selectors and signals make the difference.** *Science* 2001, **292**:1080-1081.
16. Garcia-Bellido A: **Genetic control of wing disc development in *Drosophila.*** *Ciba Found Symp* 1975, **29**:161-178.
17. Galant R, Walsh C, Carroll SB: **Hox protein repression of a target gene: Extradenticle-independent, additive action through multiple monomer binding sites.** *Development* 2002, **129**:3115-3126.

••18. Scully K, Jacobson E, Jepsen K, Lunyak V, Viadiu H, Carrière C, Rose D, Hooshmand F, Aggarwal A, Rosenfeld MG: **Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification.** *Science* 2000, **290**:1127-1131.

This work provides a beautiful example of how small differences in a selector binding site can result in distinct transcriptional outputs. In this case, a 2 bp spacing difference in the bipartite Pit-1 binding site is important for this selector to recruit a co-repressor complex in one pituitary cell type.

19. Halder G, Carroll SB: **Binding of the vestigial cofactors switches the DNA-binding specificity of the scalloped selector protein.** *Development* 2001, **128**:3295-3305.
20. Mann RS, Chan SK: **Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins.** *Trends Genet* 1996, **12**:258-262.
21. Rieckhof G, Casares F, Ryoo H, Abu-Shaar M, Mann RS: **Nuclear translocation of Extradenticle requires *homothorax*, which encodes an Extradenticle-related homeodomain protein.** *Cell* 1997, **91**:171-183.
22. González-Crespo S, Morata G: **Control of *Drosophila* adult pattern by *extradenticle*.** *Development* 1995, **121**:2117-2125.

••23. Gaudet J, Mango S: **Regulation of organogenesis by the *Caenorhabditis elegans* FoxA Protein PHA-4.** *Science* 2002, **295**:821-825.

This paper demonstrates that a battery of genes is under the direct control of a single organ-specific selector protein. A combined genetic, microarray, and genomic approach was used to identify genes and *cis*-regulatory elements under the control of the Pha-4 protein in the developing *C. elegans* pharynx. Temporal control of pharynx-specific gene expression was found to correlate with the number and affinity of Pha-4-binding sites in target gene enhancers.

24. Casares F, Calleja M, Sanchez-Herrero E: **Functional similarity in appendage specification by the Ultrabithorax and abdominal-A Drosophila HOX genes.** *EMBO J* 1996, **15**:3934-3942.
25. Morata G, Sanchez-Herrero E: **Patterning mechanisms in the body trunk and the appendages of *Drosophila.*** *Development* 1999, **126**:2823-2828.

26. Capovilla M, Brandt M, Botas F: **Direct regulation of decapentaplegic by Ultrabithorax and its role in Drosophila midgut morphogenesis.** *Cell* 1994, **76**:461-475.

27. Manak J, Mathies L, Scott M: **Regulation of a decapentaplegic midgut enhancer by homeotic proteins.** *Development* 1994, **120**:3605-3619.

28. Sun B, Hursh DA, Jackson D, Beachy PA: **Ultrabithorax protein is necessary but not sufficient for full activation of decapentaplegic expression in the visceral mesoderm.** *EMBO J* 1995, **14**:520-535.

29. Vachon G, Cohen B, Pfeifle C, McGuffin M, Botas J, Cohen S: **Homeotic genes of the Bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene.** *Cell* 1992, **71**:437-450.

••30. Ronshaugen M, McGinnis N, McGinnis W: **Hox protein mutation and macroevolution of the insect body plan.** *Nature* 2002, **415**:914-917.

This paper, together with [Galant, 2002 #119], **[Author: Please clarify which reference this is]** is one of the first reports of functional evolution of a Hox selector protein. It demonstrates that the Ubx protein has functionally diverged in the crustacean and insect lineages. In crustacean Ubx protein, the extended and divergent C terminus is rich in Ser and Thr residues that appear to be phosphorylated *in vivo* to regulate Ubx activity. The *Drosophila* Ubx protein lacks this motif.

••31. Galant R, Carroll SB: **Evolution of a novel transcriptional repression domain in the Ultrabithorax Hox protein in insects.** *Nature* 2002, **415**:848-849.

This paper, together with [Galant, 2002 #3], **[Author: Please clarify which reference this is]** demonstrates that a functional motif arose in the Ubx protein of insects. A C-terminal insect-specific glutamine/alanine rich domain is correlated with greater repression activity on the Dll limb enhancer, suggesting that the evolution of this motif potentiated or provided a new mode of repression activity.

32. Xu X, Yin Z, Hudson J, Ferguson E, Frasch M: **Smad proteins act in combination with synergistic and antagonistic regulators to target Dpp responses to the *Drosophila* mesoderm.** *Genes Dev* 1998, **12**:2354-2370.

33. Yin Z, Xu X, Frasch M: **Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development.** *Development* 1997, **124**:4971-4982.

34. Furlong E, Anderson E, Null B, White K, Scott MP: **Patterns of gene expression during Drosophila mesoderm development.** *Science* 2001, **293**:1629-1633.

35. Sheng G, Thouvenot E, Schmucker D, Wilson D, Desplan C: **Direct regulation of rhodopsin 1 by Pax-6/eyeless in *Drosophila*: evidence for a conserved function in photoreceptors.** *Genes Dev* 1997, **11**:1122-1131.

••36. Punzo C, Seimiya M, Flister S, Gehring W, Plaze S: **Differential interactions of eyeless and twin eyeless with the sine oculis enhancer.** *Development* 2002, **129**:625-634.

Drosophila, *eyeless* and *twin of eyeless*, in the direct regulation of an enhancer from *sine oculis*. They show that these two Pax-6 proteins both bind to and activate the enhancer through different sequences and that these two selector inputs are specifically required to activate *sine oculis* in the two different visual systems in the fly, the ocellus and the compound eye.

37. Niimi T, Seimiya M, Kloter U, Flister S, Gehring W: **Direct regulatory interaction of the eyeless protein with an eye-specific enhancer in the sine oculis gene during eye induction in Drosophila.** *Development* 1999, **126**:2253-2260.

38. Carroll SB, Grenier JK, Weatherbee SD: *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Malden: Blackwell Scientific; 2001.

39. Stauber M, Prell A, Schmidt-Ott U: **A single *Hox3* gene with composite *bicoid* and *zerknüllt expression characteristics in non-Cyclorrhaphan flies.*** *Proc Natl Acad Sci USA* 2002, **99**:274-279.

40. Telford MJ, Thomas R: **Of mites and *zen*: expression studies in a chelicerate arthropod confirm *zen* is a divergent Hox gene.** *Dev Genes Evol* 1998, **208**:591-594.

41. Damen WGM, Tautz D: **A Hox class 3 orthologue from the spider *Cupiennius salei* is expressed in a Hox-gene-like fashion.** *Dev Genes Evol* 1998, **208**:586-590.

42. Dearden P, Grbic M, Falciani F, Akam M: **Maternal expression and early zygotic regulation of the Hox3/zen gene in the grasshopper Schistocerca gregaria.** *Evol Dev* 2000, **2**:261-270.

••43. Alonso C, Maxton-Kuechenmeister J, Akam M: **Evolution of Ftz protein function in insects.** *Curr Biol* 2001, **11**:1473-1478.

This paper presents an example of how a Hox protein acquired a different function in the insect lineage. It demonstrated that the *Schistocerca* Ftz protein still possessed some homeotic activity but very weak segmentation potential, suggesting that the protein had acquired these activities in the course of fly evolution.

•44. Löhr U, Yussa M, Pick L: ***Drosophila fushi tarazu*: a gene on the border of homeotic function.** *Curr Biol* 2001, **11**:1403-1412.

[Author: You hadn't awarded a bullet to this paper, so have added one – please advise (all other papers had been awarded double bullets, i.e. oustanding interest, Ok?)] This paper shows that the Ftz protein lost the YPWM motif and gained the LXXLL Ftz-F1 interaction motif in the course of insect evolution and that these motifs are associated with loss of homeotic and gain of segmentation activity, respectively.

••45. Schwartz C, Sampson H, Hlousek D, Percival-Smith A, Copeland JW, Simmonds A, Kraus H: **FTZ-Factor1 and Fushi tarazu interact via conserved nuclear receptor and coactivator motifs.** *EMBO J* 2001, **20**:510-519.

This paper describes one of the few known cofactors for Hox-type proteins besides Exd and Hth. The Ftz-F1 cofactor interacts with the Ftz protein through a conserved motif that is essential for Ftz segmentation activity.

46. Suzuki T, Kawasaki H, Yu RT, Ueda H, Umesono K: **Segmentation gene product Fushi tarazu is an LXXLL motif-dependent coactivator for orphan receptor FTZ-F1.** *Proc Natl Acad Sci USA* 2001, **98**:12403-12408.

47. Yussa M, Löhr U, Su K, Pick L: **The nuclear receptor Ftz-F1 and homeodomain protein Ftz interact through evolutionarily conserved protein domains.** *Mech Dev* 2001, **107**:39-53.

48. Black B, Olson E: **Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins.** *Annu Rev Cell Dev Biol* 1998, **14**:167-196.

••49. Livesey FJ, Furukawa T, Steffen M, Church G, Cepko C: **Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene *Crx*.** *Curr Biol* 2000, **10**:301-310.

This paper is one of the first descriptions of a cell-type selector target gene battery. The Crx homeoprotein is required for retinal cell development. It is shown that a host of photoreceptor cell-type genes respond to Crx levels and possess Crx binding sites in their putative upstream *cis*-regulatory elements.

50. Weatherbee SD, Nijhout HF, Grunert LW, Halder G, Galant R, Selegue J, Carroll S: ***Ultrabithorax* function in butterfly wings and the evolution of insect wing patterns.** *Curr Biol* 1999, **9**:109-115.

51. Chan S-K, Jaffe L, Capovilla M, Botas J, Mann S: **The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with Extradenticle, another homeoprotein.** *Cell* 1994, **78**:603-615.

52. Zhou B, Bagri A, Beckendorf S: **Salivary gland determination in Drosophila: a salivary-specific, fork head enhancer integrates spatial pattern and allows fork head autoregulation.** *Dev Biol* 2001, **237**:54-67.
53. Ryoo HD, Mann RS: **The control of trunk Hox specificity and activity by Extradenticle.** *Genes Dev* 1999, **13**:1704-1716.
54. Ryoo HD, Marty T, Casares F, Affolter M, Mann RS: **Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex.** *Development* 1999, **126**:5137-5148.
55. Zaffran S, Küchler A, Hsiu-Hsiang L, Frasch M: ***biniou* (*FoxF*), a central component in a regulatory network controlling visceral mesoderm development and midgut morphogenesis in *Drosophila.*** *Genes Dev* 2001, **15**:2900-2915.
56. Popperl H, Bienz M, Studer M, Chan SK, Aparicio S, Brenner S, Mann RS, Krumlauf R: **Segmental expression of Hoxb-1 is controlled by a highly conserved autoregulatory loop dependent upon exd/pbx.** *Cell* 1995, **81**:1031-1042.
57. Di Rocco G, Gavalas A, Popperl H, Krumlauf R, Mavilio F, Zappavigna V: **The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXBA modulate the Hoxb1 auto-regulatory enhancer function.** *J Biochem* 2001, **276**:20506-20515.
58. Si Dong P, Chu J, Panganiban G: **Coexpression of the homeobox genes *Distal-less* and *homothorax* determines *Drosophila* antennal identity.** *Development* 2000, **127**:209-216.
59. Capovilla M, Kambris Z, Botas J: **Direct regulation of the muscle-identity gene *apterous* by a Hox protein in the somatic mesoderm.** *Development* 2001, **128**:1221-1230.
60. Diez del Corral R, Aroca P Gomez-Skarmeta J, Cavodeassi F, Modelell J: **The Iroquois homeodomain proteins are required to specify body wall identity in Drosophila.** *Genes Dev* 1999, **13**:1754-1761.
61. Herranz H, Morata G: **The functions of pannier during Drosophila embryogenesis.** *Development* 2001, **128**:4837-4846.
62. Certel K, Hudson A, Carroll SB, Johnson WA: **Restricted patterning of *vestigial* expression in *Drosophila* wing imaginal discs requires synergistic activation by both Mad and the Drifter POU domain transcription factor.** *Development* 2000, **127**:3173-3183.
63. Panganiban G, Sebring A, Nagy L, Carroll S: **The development of crustacean limbs and the evolution of arthropods.** *Science* 1995, **270**:1363-1366.
64. Wang D, Valdez M, McAnally J, Richardson J, Olsen E: **The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development.** *Development* 2001, **128**:4623-4633.

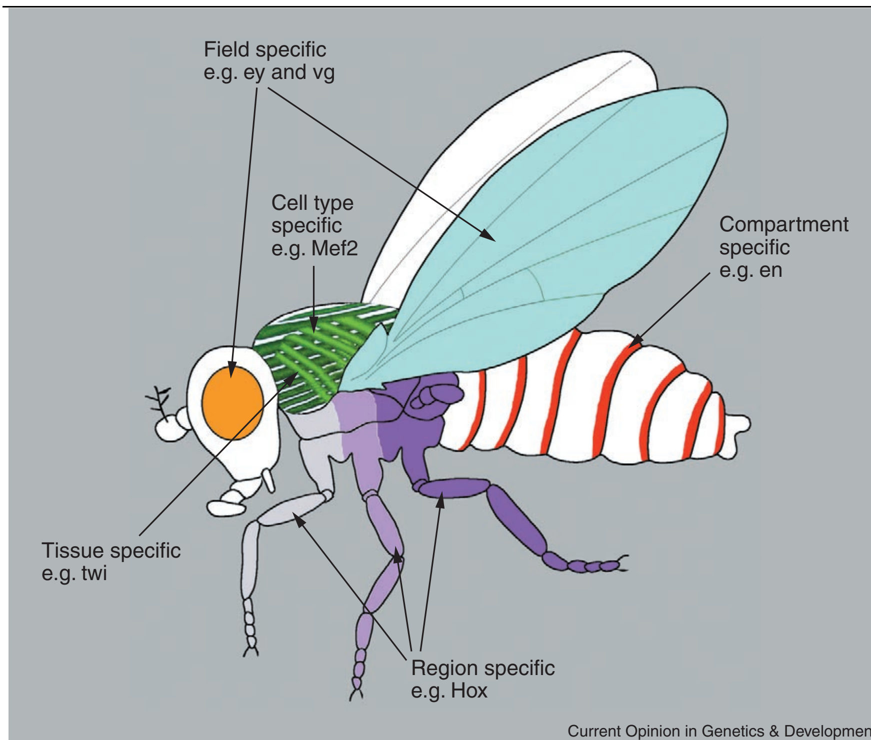

Figure 1

Types of selector proteins and their realm of action. This classification is based primarily on how broadly a selector functions in development. Region-specific selectors, such as Hox proteins, govern the development of a discrete region of the body plan, often a subset of the anterior–posterior axis. Compartment-specific selectors, such as Engrailed (en) and Apterous, govern the development of compartments, such as the posterior or dorsal compartments in *Drosophila*. Field-specific selectors, such as Vg or Dll, govern the development of fields of cells, such as the wing blade or leg fields. Tissue- or organ-specific selectors, such as Twist (twi), govern the development of a specific tissue or organ, such as the mesoderm. Cell-type specific selectors, such as Mef2, govern the development of a specific cell type, such as a muscle cell. For simplicity, only a subset of the expression pattern for each of these selectors is shown.

(a) ***eve* MHE**

(b) ***lab* 550 element**

Current Opinion in Genetics & Development

Figure 2

Combinatorial regulation of gene expression by integration of selector protein and signaling pathway inputs. Selector inputs are shown in aqua squares, co-factors in orange triangles, and signaling or other inputs in red circles. **(a)** The *eve* minimal muscle/heart enhancer (MHE) is regulated by inputs from three signaling pathways (Dpp, Wnt and DER) via three direct DNA-binding signal transducing proteins (Mad [M], dTCF, and Ets [E]), and by two selector proteins Twist (Twi) and Tinman (Tin). Note the multiple binding sites for most inputs distributed over the 312bp element. Data from [10••]; see also [12••]. **(b)**, the *lab*550 element is regulated by inputs from the Dpp signaling pathway (Mad/Media (M)), by the Lab selector protein and the two Hox co-factors Exd and Hth, and by additional regulators (C, GATA, HMG). Data from [13••]; see also [54].

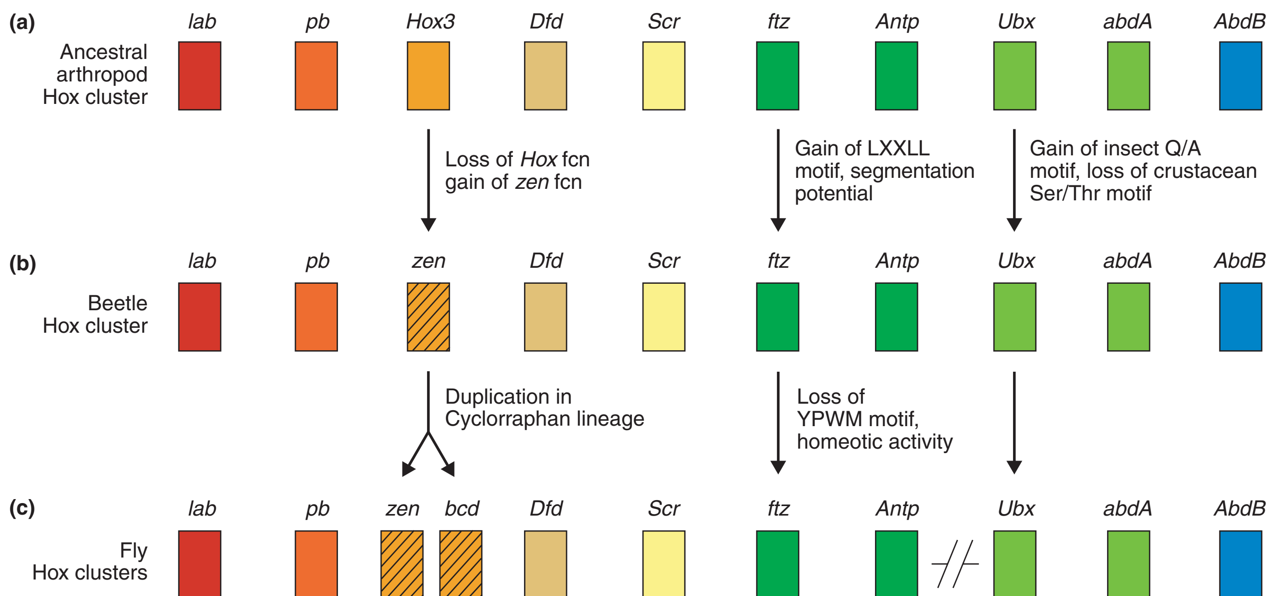

Figure 3

Evolution of Hox selector protein structure and function in the arthropod lineage. **(a)** The last common ancestor of arthropods possessed at least ten distinct *Hox* genes. **(b)** The Hox3, Ftz, and Ubx proteins acquired new roles in the insect lineage associated with the development of the dorsoventral axis (*zerknüllt* [*zen*]), pair-rule segmentation (*ftz*), and abdominal limb repression (*Ubx*). These functional changes are associated with either the acquisition or loss of distinct protein sequence motifs. **(c)** Duplication of a Hox3-type gene in the Cyclorrhaphan fly lineage gave rise to the *zen* and *bicoid* (*bcd*) genes. The fly Ftz protein lost homeotic potential through loss of its YPWM interaction motif, which mediates interactions with Exd.

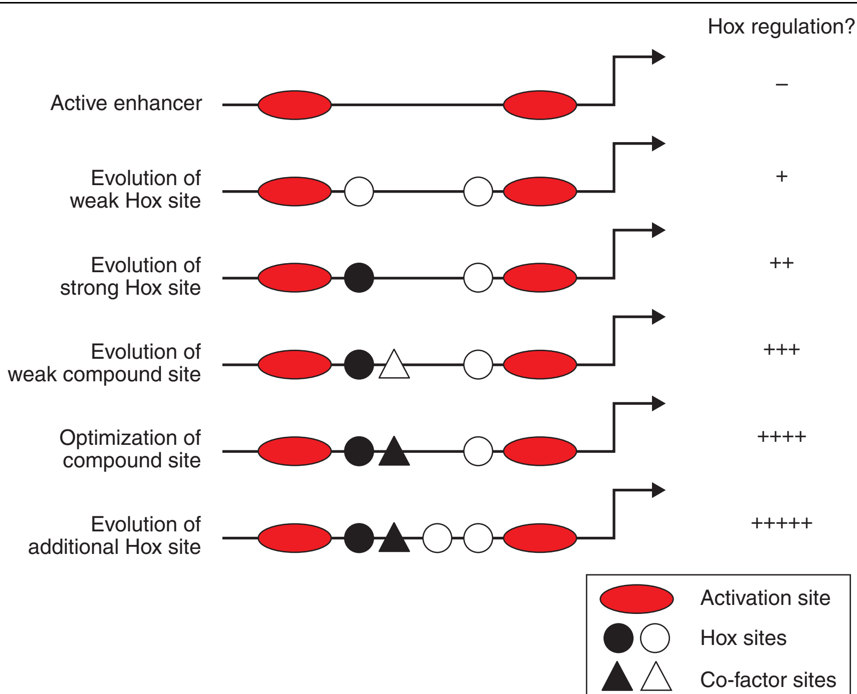

Figure 4

Model for the evolution of Hox protein regulation of a target gene. A typical enhancer with multiple sites for a transcriptional activator (red ovals) is shown. The evolution of weak Hox monomer binding sites (open circles) appears to be sufficient to impart Hox regulation in some enhancers. Directional selection and molecular evolution of the enhancer (depicted as the gain and optimization of binding sites progressing from top to bottom in the figure) may occur through addition of weak co-factor sites (open triangles) that evolve to be higher affinity sites (solid triangles). The critical feature of this model is the gradual, stepwise accumulation of sites that impart quantitatively greater Hox influence over gene regulation (based on [31••]).

Table 1

Examples of selectors and selector-regulated enhancers [Author: Ok as title?].

Selector	Target	Element (bp)	Additional inputs	References
Region specific				
Ubx (Exd/Hth)	*dpp*	dpp674 (674)	Abd-A(–)	[26-28,51]
Scr (Exd/Hth)	*fkh*	fkh (1000)	*wg*(+), *fkh*(+), *dpp*(–), *EGFR*(–)	[52,53]
Lab (Exd/Hth)	*lab*	lab550 (550)	*dpp*(+), Bin(+), ?(–)	[13,54,55]
HoxB1, A1 (Pbx/Meis)	*Hoxb1*	b1ARE (140)	RA(+), Sox2(+), Oct(+)	[56,57]
Hth	Antenna genes	Not known	*Dll*	[7,58]
Antp (Exd/Hth?)	*ap*	ME680 (680)	Not known	[59]
Iro-C	Lateral notum genes	Not known	Not known	[60]
Pnr	Dorsal notum genes	Not known	Not known	[61]
Field specific				
Vg (Sd)	*vg*	vgQ (800)	Mad(+), Dfr(+), *Ubx*(-)	[11,62] [50]*
Vg (Sd)	*sal*	sal (328)	Mad(+), Ubx(–)	[11,31]
Vg (Sd)	*cut*	cut (700)	Mad(+), SuH(+)	[11]
Vg (Sd)	*dpp*	heldout (358)	Ci(+), Pan(+)	[14]
Ey, Toy	*so*	so-10 (428)	Not known	[37]
Ey or Toy	*rh1*	Drh-P3 (319)	Not known	[35]
Dll	Ventral appendage genes	Not known	e.g. Hox, *hth*	[7]
Organ or tissue specific				
Twi	*tin*	tinB (180)	*btd*(–), *eve*(–)	[33]
Tin	*tin*	tinD (349)	Mad(+)	[32]
Tin	*eve*	MHE (312); EMEB (394)	Mad(+), Pan(+,-) Twi(+), Ets(+,-)	[10,12]
Pha-4	many	e.g. TO5E11.3 (196)	not known	[23]
Cell-type specific				
Pit-1	hGH	hGH promoter (320)	TR(+), Sp1(+)	[18]
Mef2	*Mef2c*	SMRR (158)	MyoD family	[64]
Crx	*rhodopsin*	rhodopsin promoter	Not known	[49]

Four classes of selectors are distinguished here: region specific, field specific, organ or tissue specific, and cell-type specific. Not listed are the two compartment-specific selectors in *Drosophila*, Apterous and Engrailed. Column 1: the selector; known or suspected co-factors are in parentheses; column 2: known or suspected target genes; column 3: name and size of element (if identified); column 4: other positive (+) or negative (-) inputs. Genetic inputs are italicized; in those cases where DNA binding sites have been defined, the binding factor is listed; column 5: references. Abbreviations: *dpp*, *decapentaplegic*; *fkh*, *forkhead*; *lab*, *labial*; *ap*, *apterous*; *vg*, *vestigial*; *sal*, *spalt*; *so*, *sine oculis*; *rh1*, *rhodopsin1*; *tin*, *tinman*; *eve*, *even-skipped*; *wg*, *wingless*; hGH, human growth hormone; *Mef2*, *myocyte enhancer factor 2*; Abd-A, Abdominal-A; Ubx, Ultrabithorax; *Dfr*, *Drifter*; *EGFR*, *epidermal growth factor receptor*; Bin, Biniou; RA, retinoic acid; *Dll*, *Distalless*; Ci, Cubitis interruptus [Author: 2 rs, Ok?]; Pan, Pangolin; Twi, Twist; TR, thryoid receptor; Mad, Mothers against Dpp; *btd*, *buttonhead*.

*[Author: Why reference [50] cited independently for Vg(Sd) entry? Shall I make → [11,50,62]?] [Author: Where should reference 63 be cited, please?]