The molecular evolution of development

Michael D. Purugganan

Summary

Morphological differences between species, from simple single-character differences to large-scale variation in body plans, can be traced to changes in the timing and location of developmental events. This has led to a growing interest in understanding the genetic basis behind the evolution of developmental systems. Molecular evolutionary genetics provides one of several approaches to dissecting the evolution of developmental systems, by allowing us to reconstruct the history of developmental genetic pathways, infer the origin and diversification of developmental gene functions, and assess the relative contributions of various evolutionary forces in shaping regulatory gene evolution. *BioEssays* **20**:700–711, 1998.

How do different species evolve different morphologies? This remains one of the central questions in evolutionary biology, and one whose precise answer has remained elusive. Recent approaches to this question have focused on dissecting the patterns by which developmental mechanisms diversify over evolutionary time.[1,2] Many of the present-day attempts to study the evolution of development are centered at the molecular level and exploit the remarkable progress that has been made at unraveling the molecular mechanisms that control the unfolding of morphology during organismal development.[1–6] It has not escaped the attention of both evolutionists and developmental geneticists that the morphological phenotypes that accompany mutations at some regulatory genes mimic differences observed between species.[7,8] These observations have led to suggestions that variation at these control loci may contribute to interspecies differences in body form and provide the impetus for concerted efforts to test correlations between regulatory gene evolution and morphological diversification.

Recent efforts in comparative developmental genetics have begun to address the evolutionary role of variation in molecular developmental mechanisms by assessing interspecies patterns of conservation and change in developmental regulatory gene expression and attempting to correlate these with morphological differences between taxa.[3,4] This developmental genetic approach is complemented by efforts among evolutionary geneticists who study the molecular evolution of developmental systems, focusing principally on issues of evolutionary history and dynamics.[9,10] The molecular evolutionary genetic approach provides a powerful framework for investigating the origins and history of developmental pathways. It also serves as the basis for assessing the patterns of genetic change that accompany morphological diversification and the evolutionary forces that shape developmental gene structure and function.

The study of the molecular evolution of development revolves around asking two key questions: How do developmental genes evolve? And what are the interconnections between changes at these regulatory genes and the evolution of developmental processes? Despite the central relevance of developmental evolution to the study of morphological diversification, we know very little about the molecular evolution of developmental genetic pathways and the genes that comprise them. The molecular evolutionary approach is central to an emerging evolutionary developmental biology, as it provides detailed outlines of evolutionary histories and mechanisms that are difficult to obtain by other means, and permits in-depth analyses of the dynamics that characterize the evolution of developmental systems. This review explores several aspects of the molecular evolution of developmental systems and discusses some of the insights provided in understanding the diversification of developmental processes.

Funding agency: U.S. Department of Agriculture; Funding agency: National Science Foundation; Funding agency: Alfred P. Sloan Foundation.
*Correspondence to: Michael D. Purugganan, Department of Genetics, North Carolina State University, Box 7614, Raleigh, NC 27695; E-mail: michaelp@unity.ncsu.edu

BioEssays 20:700–711, © 1998 John Wiley & Sons, Inc.

Reconstructing the history of developmental systems

The reconstruction of history through phylogenetic analysis identifies the molecular evolutionary mechanisms that have structured the genetic networks that control developmental processes and the relative timing of evolutionary events.[11] Information on the evolutionary history of developmental systems is imprinted on the sequence of the genes that comprise them. The sequence information within these developmental loci trace evolutionary events that accompany the origin of developmental regulatory genes and gene functions,[9,12–15] the evolutionary processes that lead to the elaboration of developmental pathways,[16] the historical correlations between different, interacting developmental systems,[17,18] and the interconnecting links between molecular and morphological diversity.[19–21] Two series of studies—one on the animal segmentation genes and the other on plant flowering loci—illustrate some of the lessons to be gleaned from a molecular phylogenetic approach.

Early evolution of the HOM/Hox genes

From the phylogenetic perspective, one of the best studied molecular developmental system is the animal homeodomain HOM/*Hox* genes, which encode DNA-binding transcriptional activators involved in anteroposterior (AP) axial patterning in several animal taxa.[13,22] The *Antennapedia*-class homeotic genes form a HOM-C gene cluster of eight loci in *Drosophila melanogaster*. Members of the HOM-C include the genes *proboscidea (pb)*, *labial (lab)*, *Deformed (Dfd)*, *Antennapedia (Antp), Ultrabithorax (Ubx)*, *abdominal-A (abd-A)*, and *Abdominal-B (Abd-B)* and control the specification of segmental identity along the insect body.[22] The genes in the complex are expressed along the AP axis in a sequence that is collinear with their position within the gene cluster. Homologues to these homeotic loci are also found in mammalian species, where they are referred to as *Hox* genes.[13] In humans, a total of 38 genes are organized into four *Hox* clusters (A–D), each of which spans ~100 kb and contains 9–11 loci.

The organizational complexity of the HOM/*Hox* cluster in extant animals has prompted several attempts at phylogenetic reconstruction to analyze the evolutionary events that accompanied the diversification of these developmental genes and infer the organization of the cluster in ancestral metazoan taxa.[13,23–26] Previous analyses have suggested that the ancestral metazoan cluster consisted of three precursor genes that were responsible for head, trunk, and tail patterning in early metazoans,[24] although the lack of support for many of the reconstructed relationships makes it difficult to draw clear conclusions from these early analyses. A recent phylogenetic study, however, has clarified many of the ambiguous relationships among these loci and provides a plausible evolutionary scenario for the diversification of this developmental system.[12]

Based on the reconstructed phylogenetic relationships of the HOM/*Hox* genes in the arthropod *Drosophila melanogaster*, the vertebrates *Homo sapiens* and *Mus musculus*, the cephalochordate *Amphioxus* and the nematode *Caenorhabditis elegans* (Fig. 1), Zhang and Nei[24] suggest that extant members of the HOM/*Hox* gene family were derived not from three precursors, as earlier suggested, but from two related loci that diverged very early in the evolution of the metazoans (Fig. 2A). Given the current expression patterns of various HOM/Hox genes, it is likely that one of these early loci (the A/B/C ancestor) was responsible for anterior or ''head'' patterning, while the other gene (the D/E ancestor) was involved in posterior-like development—two generalized morphogenetic functions that were probably present in very early metazoan species.

One of these ancestral genes, the D/E locus, duplicated at least once in early evolution to form two distinct posteriorly expressed loci—the *Hox9/Hox10* (D group) ancestral locus, and the progenitor of the *Drosophila Abdominal-B* and vertebrate *Hox11* to *Hox 13* loci (the E group genes; Fig. 2). The other ancestral locus underwent a more complex series of duplications, first producing the A group genes which later split to form a *lab*-like and *pb*-like lineage, and the B group/C group ancestor. The latter subsequently split into two genes, one of which resulted in the ancestor of the vertebrate *Hox3* loci, while the other produced the group C genes. Although the phylogenetic analysis indicates a closer relationship between B and C group genes, the bootstrap support is weak (<50%), and it is possible that the ancestral *Hox3* gene was actually derived from A group loci.

All these duplications occurred before the divergence of the Pseudocoelomates (represented by *Caenorhabditis elegans*) and the Coelomates (including the vertebrate and arthropod lines) 750 mya,[27] so that the HOM/*Hox* cluster of this primitive Precambrian ancestral species had 6 loci (Fig. 2).

Reconstruction of the possible developmental functions of this primitive cluster suggests again that one group of genes (the D and E loci) were involved in posterior development; in *Drosophila*, *Abd-B* is expressed in the last abdominal segments that later give rise to genitalia.[28] The other four genes (the B and C group ancestral loci, and the *lab*-like and *pb*-like progenitors) were probably involved in different aspects of anterior or head patterning. The greater number of loci, accompanied by increasing differentiation in expression domains and function, may reflect increased specialization of head and ''tail'' structures in at least the primitive common ancestor of Pseudocoelomates and Coelomates, and possibly an even earlier ancestral metazoan taxa.

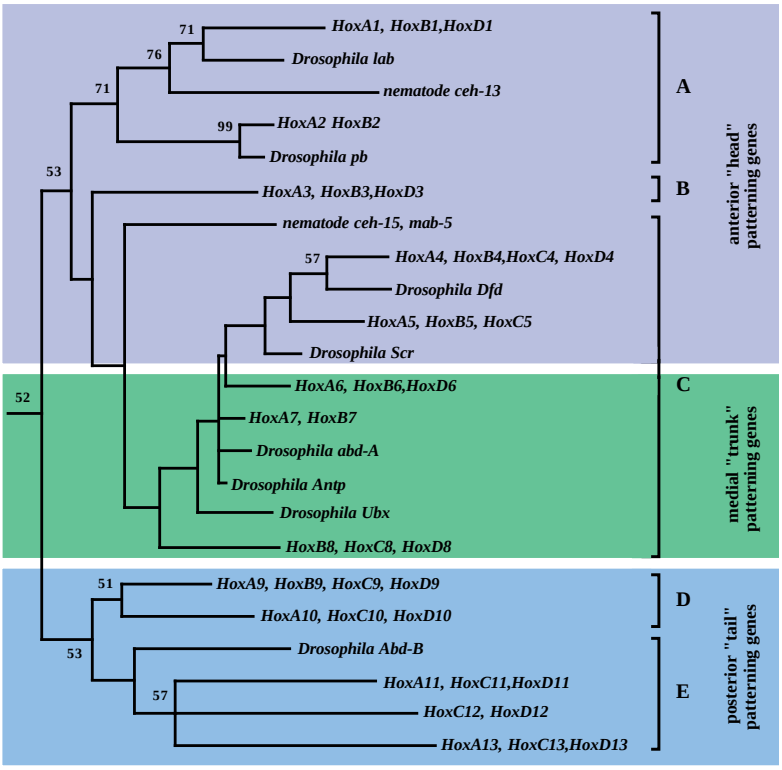

Figure 1. Phylogeny of the *Antennapedia*-class homeobox genes. The phylogeny is a simplified adaptation from Zhang and Nei.[12] The numbers next to the nodes give the bootstrap support, with short nodes of less than 50% support collapsed in the tree. The three major group of genes (anterior, medial, and posterior genes) are indicated. The precise boundary between anterior and medial genes is unclear; the *Drosophila* gene *Scr* is expressed in both the head and thoracic regions.

Subsequent expansion of the HOM/*Hox* cluster proceeded in several directions. Many of these later duplications occurred in the medially expressed C group loci, which underwent a series of diversification events that followed the initial establishment of the *Deformed* lineage (Fig. 1). The phylogenetic relationships within the C group loci are ambiguous, given the low bootstrap supports of most of the groupings, although there is greater support for a separate *Dfd/Hox4* lineage. In *Drosophila melanogaster, Dfd* is expressed in the maxillary and mandibular region during embryonic head development,[29] while mouse *Hox4* genes are expressed in the hindbrain.[30] The duplications that gave rise to the *Drosophila* C group genes—*Sex-combs reduced (Scr), abdominal-A (abd-A), Antennapedia (Antp)* and *Ultrabithorax (Ubx)*—may have occurred separately from those that led to the vertebrate *Hox5* to *Hox8* genes, a possibility supported by the low sequence similarity between these two group of genes outside of the conserved homeodomain[12] (Fig. 2).

Comparative genetic studies indicate that the duplications that gave rise to *Ubx* and *abd-A* are confined to arthropods, myriapods and the *Onychophora*, and that the last common ancestor of these taxa already had the full complement of HOM class genes at the time of their divergence in the later Early Cambrian 530 mya.[31] The expression of both the *Drosophila* and vertebrate loci, however, indicates that genes involved in patterning of the central body trunks were derived later in metazoan evolution from more anteriorly expressed head patterning loci. In *Drosophila, Scr* is expressed in the labial head and parts of the first thoracic segments,[32] but all the other genes are involved primarily in thoracic and abdominal segment differentiation (Fig. 3). Unlike previous analyses that suggest that the ‘‘trunk’’ patterning genes were part of the very early ancestral HOM/*Hox* complex,[24] Zhang and Nei’s work indicates a relatively recent origin in metazoan evolution for genes that pattern the middle of the animal body, which may be correlated with the rise of morphological complexity in the trunk body regions in some animal lineages (Fig. 3).

The relative timing of events in this reconstruction is still ambiguous, given the limited number of metazoan phyla represented in the phylogeny and the low bootstrap support for some gene relationships. Some of the diversification events outlined here may have occurred earlier. Putative identification of planarian (phylum Platyhelminthes) orthologues to the *Drosophila* HOM genes, for example, suggests that the medial group C genes may already have been established with the advent of triploblastic animals, although this is based on only fragments of the homeodomain sequence in flatworm species.[33] Several cnidarian HOM/Hox-like genes (the *Cnox* genes) have recently been isolated, but these loci appear to have diversified separately from those

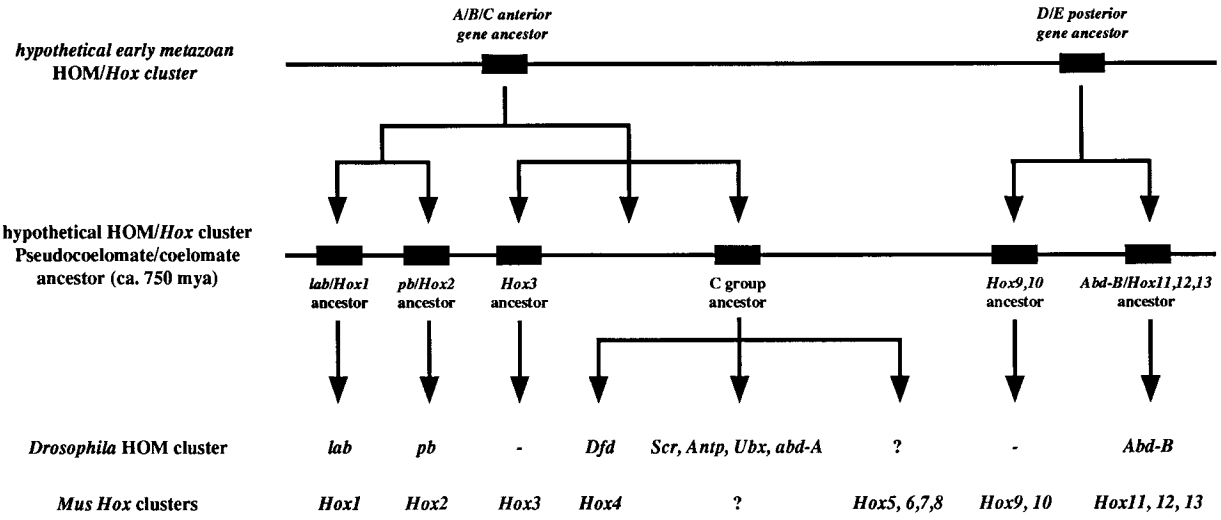

Figure 2. Evolutionary scenario for the diversification of the HOM/*Hox* cluster. The ancestral cluster consisted of anterior and posterior patterning genes that gave rise to the present-day cluster in extant metzoans. Both the *Hox3* and *Hox9/10* lineages were lost in the line leading to *Drosophila*. The sequence of diversification for most of the C group genes is unclear from phylogenetic analysis; in this scenario, we assume an independent origin for the *Scr*, *Ubx*, *Antp*, and *abd-A* genes in arthropods and the *Hox5* to *Hox8* genes in vertebrates. Phylogenetic analysis also shows that the duplication of the *Hox* cluster in vertebrates to four clusters (*HoxA* to *HoxD*) occurred after the split of cephalochordates from vertebrates.[12,13]

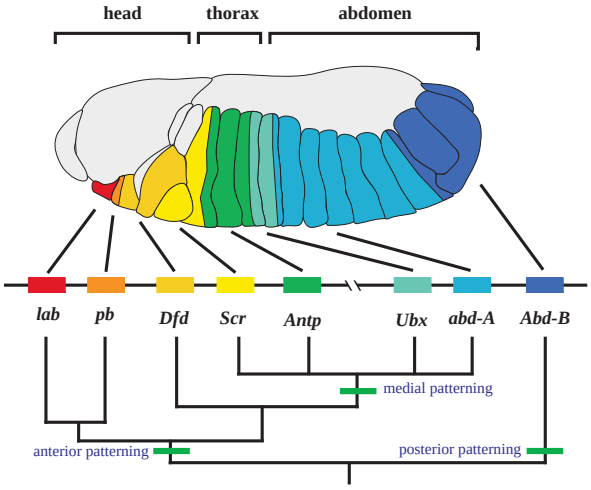

Figure 3. Evolution of developmental function among the *Drosophila* HOM-C genes. The major expression domains of the different *Drosophila* genes are mapped onto the gene phylogeny. The expression patterns strongly suggest that genes in *Drosophila* that are expressed in the thoracic and the anterior abdominal segments are derived from the anterior head segmentation loci. Adapted in part from Carroll.[3]

found in triploblastic species.[34] This suggests that the very early events of HOM/Hox cluster diversification, including the initial separation of anterior (A/B/C) and posterior (D/E) genes may have occurred only in the triploblasts. It will be interesting to carefully reconstruct the HOM/*Hox* phylogeny with a more thorough taxonomic sampling to clarify the relative timing of some gene duplication events.

Diversification of flower developmental genes

The diversification of plant developmental systems has also been explored in a molecular phylogenetic context, including the regulatory loci that interact to direct flower development in the angiosperms. Genetic studies in *Arabidopsis*, *Antirrhinum*, and maize have led to the isolation and characterization of homeotic loci that regulate flower morphogenesis.[5,35] Molecular studies have shown that many of these floral regulatory genes belong to the plant MADS-box gene family, a group of transcriptional activators that, in *Arabidopsis thaliana*, contain approximately 20–25 members.[36] At least five of these genes (*AGAMOUS, APETALA1, CAULIFLOWER, APETALA3*, and *PISTILLATA*) have defined floral homeotic functions. The precise developmental functions of the other members, referred to as *AGLs* (or *AGAMOUS-like* genes) remain poorly understood, although expression studies have shown that most are expressed in developing flowers.[37] A few *AGLs*, however, are preferentially expressed in roots or the early developing plant embryo.[36]

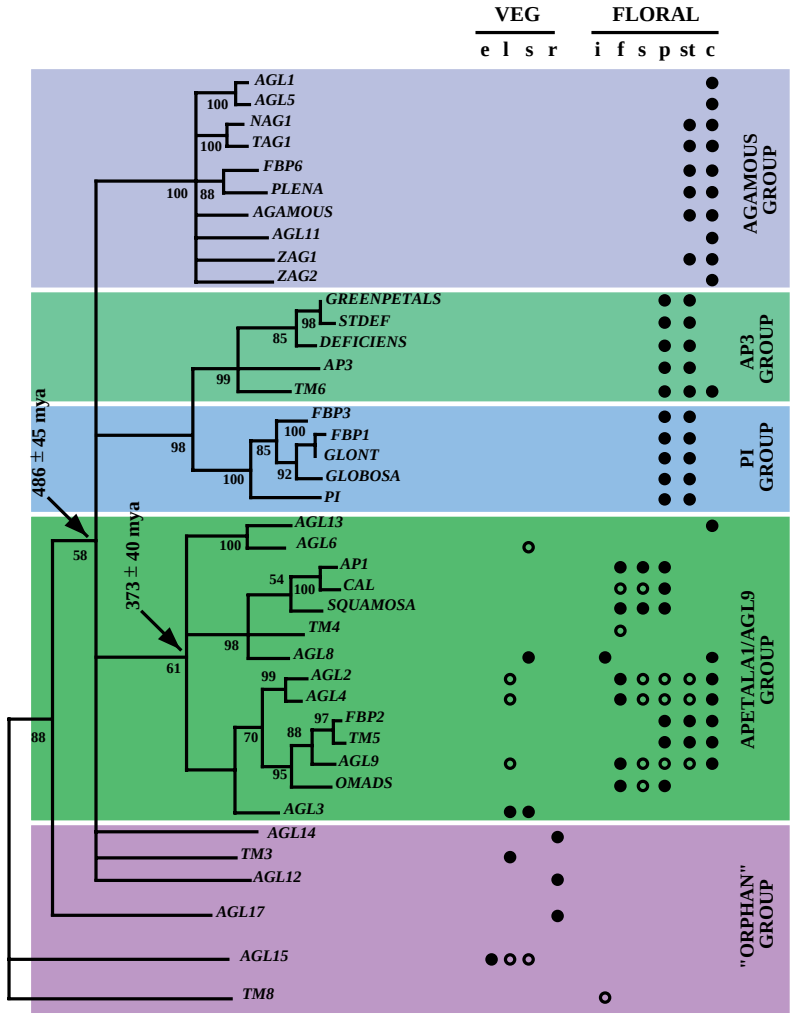

Figure 4. Evolution of the plant MADS-box gene family. The different major plant MADS-box floral homeotic gene groups are indicated. The expression of the different genes in various tissues or organs are indicated, with closed and open circles indicating strong and weak expression detected, respectively (vegetative structures: e, embryo; l, leaves; s, stems; r, roots; reproductive structures: i, inflorescence meristems, f, floral meristems; s, sepals; p, petals; st, stamens; c, carpels). The expression data were compiled from the literature. Molecular clock estimates for the appearance of the floral homeotic gene groups and the diversification of the AP1/AGL9 group genes are indicated.[14,15]

Phylogenetic analysis of the plant MADS-box regulatory gene family[14,15,38,39] shows that most of the various genes within this family are organized into distinct clades, each representing a floral homeotic gene group (the AGAMOUS, APETALA3/PISTILLATA, or AP1/AGL9 groups) whose members share similar developmental functions. The functional diversification within and between gene groups can be assessed by mapping expression patterns onto the gene phylogeny (Fig. 4). On the basis of this analysis, the AGAMOUS and APETALA3/PISTILLATA floral homeotic gene groups have members which show very tight expression profiles. Genes within the AGAMOUS group, for example, are expressed in developing stamens and carpels, which is consistent with genetic analysis—mutations at these loci result in homeotic transformations in these reproductive organs.[40,41] The diverse AP1/AGL9 group, however, includes several distinct genes that are found in flowering plant genomes. In *Arabidopsis*, for example, these include the *APETALA1*, *CAULIFLOWER*, *AGL8, AGL3*, and *AGL9* loci. The expression profiles of the AP1/AGL9-group genes reflect this evolutionary diversity; genes in this clade are expressed in a wider range of tissues and organs in the plant, including several that are expressed in both vegetative and floral structures.[42]

How did these floral homeotic genes originate? The plant MADS-box genes include several loci (the “orphan” genes), which are expressed primarily in vegetative parts of the plant, such as root or embryonic tissues, and are evolutionarily related to the floral developmental loci[36] (Fig. 4). Although resolution at the base of the phylogeny is poor, it suggests that early plant MADS-box genes may have possessed a vegetative developmental function, and subsequent gene duplications led to the current genes, which were evolutionarily recruited to direct diverse aspects of reproductive devel-

TABLE 1. Comparison of Sequence Distances Between Duplicated Structural and Regulatory Genes in the Maize Genome

Duplicated loci	Length (kb)	Nucleotide substitution distance[a] Ks	Ka	Ka/Ks
Regulatory genes				
Ohp1/Ohp2	1.200	0.254 (1.19)	0.0593 (7.25)	0.234
R/B	1.677	0.241 (0.83)	0.0841 (7.84)	0.349
C1/Pl1	0.777	0.159 (1.05)	0.0462 (9.25)	0.291
Ibp1/Ibp2	2.061	0.150 (0.36)	0.0482 (3.62)	0.321
Tbp1/Tbp2	0.606	0.147 (1.20)	0.0066 (1.45)	0.045
Vpl4a/Vpl4b	1.821	0.121 (0.29)	0.0346 (2.69)	0.286
Obf1/Obf2	1.026	0.104 (0.48)	0.0196 (2.95)	0.189
Mean:				0.245
Structural genes				
Orp1/Orp2	1.170	0.298 (1.44)	0.0114 (1.31)	0.038
Ant1/Ant2	1.173	0.227 (1.32)	0.0114 (1.32)	0.050
Cpna/Cpnb	1.734	0.186 (0.55)	0.0126 (0.97)	0.068
Cdc2a/Cdc2b	0.882	0.177 (1.04)	0.0097 (1.46)	0.055
Whp1/C2	1.206	0.169 (0.66)	0.0286 (3.29)	0.169
Fer1/Fer2	0.627	0.168 (1.44)	0.0189 (4.01)	0.113
Pgpa1/Pgpa2	1.170	0.102 (0.39)	0.0494 (5.97)	0.484
Mean:				0.163

Adapted from Gaut and Doebley.[49]
[a]Ks and Ka are synonymous and nonsynonymous distances, respectively. Variances for Ks ($\times 10^3$) and Ka ($\times 10^5$) are shown in parentheses.

opment in the seed plants. The establishment of the distinct floral homeotic gene groups predate the rise of the flowering plants, and it appears that these loci were extant even before the origin of seed plants 285 mya.[43,44] Molecular clock estimates indicate that the gene duplications that gave rise to the floral homeotic gene lineages may have occurred as early as the origin of land plants 450 mya and took place in a relatively short span of evolutionary time (<50–75 myr).[14,15] It remains unclear what the function of the ancestral plant MADS-box genes were, particularly in plant taxa whose reproductive structures are very different from those seen in angiosperms. The rapid early evolution of these genes, however, suggests that specialization of distinct gene lineages may have resulted from selective pressures on ancestral land plant lineages to evolve more complex reproductive morphologies. Although much of this work is at an early stage, these molecular evolutionary analyses provide the historical framework to test hypotheses on the origin and evolution of developmental function of these plant regulatory genes, and can be used to guide experiments to explore the diversification of this gene family.

Tempo of evolution among developmental loci

The rates at which genes evolve reflect the long-term evolutionary forces that have shaped the structure of these loci. It is possible, in some cases, to infer the nature of selective and other evolutionary forces that act on a gene by measuring the nature and patterning of rate variation between domains, loci or species.[16,45,46] The rates of gene evolution have thus become a useful benchmark to contrast the differing evolutionary dynamics between loci.[47]

Rapid mosaic evolution of regulatory genes

How fast do regulatory genes evolve? Pleiotropic effects of mutations at morphogenetic loci would suggest that the evolution of regulatory genes may be highly constrained and therefore evolve at a relatively slow pace.[48] Recent work, however, indicates that while some domains of regulatory proteins (e.g., DNA-binding domains) are indeed evolving fairly slowly, other regions may be among the faster-evolving components of the coding region of eukaryotic genomes.

In the maize genome, one can compare substitution rates for genes that possess duplicate copies as a result of a polyploidization event that occurred at 15–20 mya[49] (Table 1). Since synonymous substitutions may be largely neutral and thus dependent only on the mutation rate, the ratio of nonsynonymous (Ka) to synonymous (Ks) substitutions between gene pairs provides a normalized measure of evolutionary rate for different loci. In a study that involved estimating the substitution rate of 14 maize duplicate gene

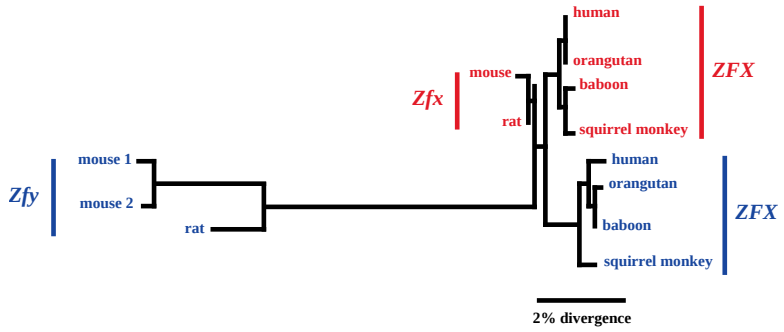

Figure 5. Relationships among primate and rodent *ZFX* and *ZFY* genes. The phylogeny highlights the large evolutionary distance leading to the rodent *Zfy* loci. Adapted from Shimmin et al.[60]

loci, seven regulatory gene pairs had a significantly higher Ka/Ks ratio than seven other structural genes (0.25 vs 0.16), indicating that regulatory proteins evolve faster than structural genes in this plant genome.

This rapid evolution of regulatory loci has been observed in basic helix-loop-helix,[50] MADS-box,[14] HMG-domain,[52,53] and zinc finger proteins.[54] It appears that regulatory proteins evolve in a highly mosaic fashion, with regions of very strong conservation interspersed with regions that evolve at rapid rates. In the plant *R* gene family, for example, which encodes basic-helix-loop-helix transcriptional activators that regulate anthocyanin pigmentation patterning, the conserved regions evolve at about 1.02×10^{-9} nonsynonymous substitutions/site/year, while the rest of the gene evolves at a fourfold faster rate [approximately 4.08×10^{-9} nonsynonymous substitutions/site/year] [50]. The mean Ka/Ks ratio of the rapidly evolving region for these regulatory genes is 0.89, which is significantly greater than the mean value of 0.14 for a set of plant genes and 0.189 for 42 mammalian sequences.

It is unclear why some regions of these regulatory loci evolve so fast. The rapidly evolving regions at some control loci do not appear to be dispensable sequences, since functional and genetic studies demonstrate that, for some genes, alteration of sequence in these domains result in changes in protein function (see, e.g., ref. 51). It has been suggested that the increased levels of nonsynonymous substitutions at these regulatory genes may have adaptive significance; for example, rapid evolution of the *SRY* mammalian sex determination gene, which encodes a DNA-binding transcriptional activator, may be associated with speciation.[52,53] It may be that regulatory diversification may accompany rapid sequence evolution, resulting in phenotypic variations that are observed even between closely related species. This rapid change may also be partly driven by coevolutionary mechanisms between epistatically interacting loci.

Whether this variation in evolutionary rates for regulatory genes are correlated with rates of morphological evolution remains open to question. Adaptive radiations, for example, are characterized by highly diverse species that show strong molecular genetic similarity,[55,56] which led to early suggestions that much of the visible diversity we observe in nature arises from changes in gene regulation rather than structural protein evolution.[55] The issue of correlated rates between molecular and morphological evolution, however, continues to be raised,[57] and recent work does indicate that an association between sequence change and morphological variation may indeed exist.[58] Many of these studies, however, focus on the molecular evolution of structural genes, and it is unclear whether a stronger correlation of rates may be uncovered when comparing the tempo of morphological change with molecular variation in genes that regulate developmental processes. Addressing these issues will require detailed evolutionary and comparative studies to address the precise functions that variable regulatory gene sequences play, and dissect possible causal links between their rapid evolution and organismal diversification.

Regulatory gene evolutionary rates vary between lineages and over time

The rate at which regulatory genes evolve can vary significantly across lineages, and may signal changes in the mutation rates, selective pressures, population dynamics, or even the genetic interactions that these loci experience in different taxa. One interesting example concerns the homologous zinc finger genes *ZFX* and *ZFY* that are encoded in the mammalian X and Y chromosomes.[59,60] These two loci encode transcriptional activators that duplicated prior to the origin of the placental mammals, and are believed to be involved in reproductive development. Evolutionary analysis of these genes in primates (humans, orangutans, baboons, and squirrel monkeys) and rodents (mice and rats), shows that *ZFX* is well conserved across all placental mammals, including both primates and rodents, while *ZFY* is conserved only among primate species but has evolved rapidly in rodent lineages (Fig. 5). Only one nonsynonymous substitution separates rat and mouse *ZFX*, but 38 replacement substitutions differentiate the *ZFY* genes between these two species. Shimmin and co-workers suggest that the evolution of

X-inactivation of *ZFX* in the rodent lineage makes *ZFY* a redundant locus subject to a new selective environment, and this is manifested in an increase in its rate of molecular evolution.[60]

There are also indications that for some regulatory loci most of the sequence evolution has occurred rapidly over short time periods. The pituitary growth hormones, for example, evolve slowly in the vertebrates except for short bursts of evolution in the primates and the artiodactyls. In the latter group, the rate of evolution has increased 25- to 50-fold over the basal rate at specific times in the evolutionary history of these taxonomic lineages.[61,62] In the primates, this rate acceleration appears to have transpired before the divergence of Old World monkeys and apes. It is estimated that a slow rate of evolution predominates 90% of the time for these loci, but 85% of sequence change occurs in short periods of rapid diversification. The pattern suggests that the evolution of these proteins is characterized by punctuated bursts of sequence change, which may be associated with switching of secondary hormone functions during particular periods of their evolution in certain animal taxa.[63] This episodic evolution of regulatory proteins may be a widespread phenomenon; it is believed that rapid sequence evolution follows gene duplication events that characterize the growth of many developmentally important regulatory gene families and may be correlated with functional divergence among paralogous loci.[16]

Microevolution of development

All morphological differences observed between taxa have their ultimate origins as genetic variation within a species. Understanding the evolutionary dynamics of morphological diversification partly involves assessing the levels, patterning, and distribution of diversity in candidate developmental loci to uncover the origin and ascertain the significance of population-level variation. Studies on molecular diversity at the sequence level have proven particularly useful in dissecting the mechanisms that govern the evolution of specific genetic loci.[64,65] Molecular population genetics provides a context for estimating how population history, breeding system, and selection affect molecular variation, and delineating the mechanisms that lead to evolutionary diversification.

Variation at regulatory genes

Levels of genetic diversity at regulatory loci govern the rates of adaptive evolution and limit the degree to which selection at these genes can shape evolutionary change. What is the extent of variation in developmental genes? Moriyama and Powell[66] compiled data on sequence variation in the coding region for 22 nuclear loci in *Drosophila melanogaster* (Table 2). For X-linked loci, the estimate of the mean proportion of nucleotide differences between alleles (π) for four structural genes is 0.0036, while three regulatory loci have a mean of 0.0015. For autosomal loci, the 12 structural genes have a mean π of 0.0049, while the three regulatory genes studied had an average value of 0.0019. Regulatory loci also possess consistently lower levels of nucleotide diversity at noncoding regions.[66] Although they remain limited, the data on regulatory genes in *Drosophila* do document the extent of diversity among some developmental genes and suggest that variation levels may be lower for these regulatory loci, as compared with their structural gene counterparts. By contrast, work on both the Arabidopsis MADS-box floral regulatory gene *CAULIFLOWER*[67] and Zea helix-turn-helix anthocyanin pigmentation control gene *C1*[68] demonstrate appreciable levels of diversity that are closer to species norms. This variation in the levels of intraspecific regulatory gene polymorphisms is the result of the particular demographic and selective forces that each specific locus experiences.[64,65]

What evolutionary forces act on these regulatory genes? Like all other loci in *Drosophila*, variation at specific regulatory genes result from the interplay of various evolutionary forces such as selection and genetic drift, which together will determine the degree to which polymorphisms are available for adaptive changes in organismal morphology.[64] Using several statistical tests to probe the selective history of particular regulatory genes, the levels and patterning of nucleotide diversity within and between loci can be used to infer the nature of these evolutionary forces.[64,65]

Several loci, such as the *Drosophila bride-of-sevenless (boss)*[69] and *Zea mays C1*[68] genes, display a pattern of sequence variation that is consistent with predictions of the neutral model of molecular evolution. For *boss*, which encodes a membrane-bound receptor that triggers photoreceptor cell differentiation in the fly eye, it has been suggested that the position of this gene as part of a developmental network constrains the selective pressures that it may experience.[69] There are cases, however, in which patterns of sequence polymorphism for regulatory genes appear consistent with the possibility of directional selection. The *decapentaplegic (dpp)* gene, which encodes the intercellular signaling molecule transforming growth factor-β (TGF-β), shows evidence for selective divergence in the lineage leading to *Drosophila pseudoobscura*.[70] There is also evidence for recent adaptive fixation of alleles at or near the *Drosophila* sex-determining gene *transformer*[71] and the segment polarity gene *ci*,[72] both of which display low levels of intraspecific polymorphism that may have arisen as a result of selective hitchhiking (Table 2). It is unclear whether the site(s) under selection are at these regulatory loci or are simply tightly linked to them, and what the phenotypic or adaptive correlates of this selection may be.

TABLE 2. Nucleotide Diversity of Regulatory and Structural Gene Coding Regions in *Drosophila melanogaster*

Gene	No. of alleles	Length (kb)	No. of polymorphic sites	Nucleotide diversity[a]	
				π	Θ
X-Linked genes					
Regulatory genes					
ase	6	1.068	6	2.06	2.46
pn	8	1.173	1	0.22	0.33
z	6	0.804	5	2.09	2.74
Mean:				1.46	1.84
Structural genes					
Pgd	13	1.443	4	1.33	0.89
per	6	1.682	20	4.95	5.21
Yp2	6	1.046	9	4.46	3.77
Zw	33	1.558	24	3.84	3.80
Mean:				3.65	3.42
Autosomal genes					
Regulatory genes					
tra	11	0.588	1	0.74	0.58
boss	5	1.566	16	4.86	4.90
ci	10	0.958	0	0.00	0.00
Mean:				1.87	1.83
Structural genes					
Acp26Aa	10	0.792	16	7.35	7.14
Acp26Ab	10	0.270	5	8.81	6.55
Adh	15	0.768	19	8.11	7.61
Adhr	11	0.816	6	1.34	2.51
Lcp1Psi	10	0.384	1	1.22	0.93
Pgi	11	1.674	4	0.78	0.82
Amy-d	8	1.482	37	8.82	9.63
Amy-p	10	1.482	34	9.75	8.11
Sod	11	0.441	7	4.37	5.42
Est-6	13	1.632	45	7.21	8.89
Rh3	5	1.149	2	0.70	0.84
Mlc1	16	0.314	0	0.00	0.00
Mean:				4.87	4.87

Adapted from Moriyama and Powell.[66]
[a]Nucleotide diversities are multiplied by 10^3.

Preliminary surveys document that variation for regulatory genes do exist, and that in certain cases there is some evidence to suggest that selective pressures have patterned the levels of polymorphism and divergence at these loci. The question now is, does the variation we observe at the molecular level translate into phenotypic variation that can lead to between-species evolutionary divergence? In other words, do these polymorphisms matter? Selection experiments indicate that some population-level variation in developmental function can be attributed to polymorphisms at regulatory genes. Work on natural variation in *Drosophila* bristle number are associated with polymorphisms at *achaete-scute* and *scabrous*,[20,73] both of which are neurogenic genes involved in sensory structure development. It has also been shown that ether sensitivity in flies, which results in the dramatic bithorax phenocopy first demonstrated by C.H. Waddington, is correlated with alleles at *Ultrabithorax*.[19] In *Arabidopsis*, the MADS-box floral regulatory locus *CAULIFLOWER* appears to be evolving non-neutrally, and molecular polymorphisms in this gene are associated with the

differential ability of naturally occurring alleles to program early flower formation.[67]

It is clear that molecular polymorphisms observed at some developmental control genes are associated with natural phenotypic variation in a variety of structures, and can serve as the basis for selection to act upon and potentially lead to morphological divergence. What remains to be demonstrated are the precise forces and mechanisms that drive these molecular polymorphisms to fixation between species; recent isolation of morphological trait genes such as *manx* in ascidians,[74] *teosinte-branched1* in maize,[75] or *Bocauli-flower* in *Brassica oleracea*,[51] may allow us to link together intraspecific variation at regulatory loci to between-species morphological differences.

Evolution of regulatory targets

Morphological evolution is correlated with variation in gene expression patterns, which may arise from changes in the *trans*-acting regulatory factors or the *cis*-acting control regions of various loci. The ability of regulatory transcriptional activators to function, for example, depends critically on their recognition and interaction with specific promoter element sequences in various target genes. Other regulatory elements have also been observed in introns and 3′ untranslated regions, and may mediate regulation at a variety of transcriptional and post-transcriptional control points.

We know very little about how *cis*-regulatory regions evolve. Intraspecific variation in promoter sequences has been documented, and it is clear from molecular population genetic analysis that many promoters are constrained in their molecular evolution.[70,76–78] The extent of this constraint is demonstrated in the promoter of the homeodomain pair-rule gene *evenskipped* (*eve*), which is expressed in *Drosophila* embryos in a series of stripes that define parasegmental boundaries. Population genetic analysis of five *D. melanogaster* and six *D. simulans eve* promoter alleles sampled over a broad geographic range estimate that this promoter is 2% diverged between the two species, although *eve* intron and silent coding sites display 6% divergence.[76] The low divergence values of this region indicates that this control sequence is subject to purifying selection, although stabilizing or compensatory selection may also act at these regulatory sequences and partly explain the patterning of variation at this promoter. Evidence for constraint has also been shown for regulatory regions in *dpp*[70] and *Mlc1*,[78] although these regulatory sites are within intron sequences. In the case of *dpp*, this conservation is associated with the presence of transcriptional enhancer sequences, while in *Mlc1* they are found in introns that flank an alternatively spliced exon.

There is evidence that some of these regulatory sites are evolving non-neutrally. This can be demonstrated in the homeodomain-containing pair-rule gene *fushi tarazu* (*ftz*), which is one of the targets of *eve* and which possesses a 1-kb *zebra* element sequence at its 5′ upstream region. Using the patterns of intraspecific polymorphisms, Jenkins and co-workers demonstrate an excess of substitutions in protein-binding sites at the *zebra* element between *D. melanogaster* and *D. simulans*, suggesting adaptive divergence at these target sites between the two species.[79]

Naturally occurring differences in promoter sequences have been correlated with gene expression levels in several genes. In the *Esterase-6* promoter analyzed in 17 *D. melanogaster* lines, polymorphisms in a 325-bp region are found to be in strong linkage disequilibrium and are associated with a peak of variation approaching silent site levels.[77] This region contains elements that control *Est-6* expression in the anterior sperm ejaculatory duct, and heterogeneity in this sequence may arise from yet undefined selective forces; two *Est-6* promoter haplotype groups, designated P1 and P7, are associated with differences in male *Est-6* activity in *D. melanogaster*.

The role of selection in promoter variation and its association with regulatory function has been extensively explored in the *Fundulus heteroclitus* (killifish) *Ldh-B* locus.[80] The *Ldh-B* gene is fixed for alternate alleles in northern and southern coastal Atlantic populations; this differentiation is associated with divergence in developmental rates, hatching times, and swimming performance in selection experiments. The northern population displays twice as much *Ldh-B* activity than the southern population, which is correlated with an increase in gene transcription rates. A molecular population genetic analysis of the *Ldh-B* promoter reveals a peak of variability in a central region of the 5′ flanking regulatory sequence between promoters, and both the Fu and Li and the HKA selection tests reject neutrality as an explanation for the variation. A repressor element in this central region has been revealed in transient reporter construct expression studies using transformed heterologous systems, and may be partly responsible for differences in *Ldh-B* transcriptional levels observed between northern and southern populations.

The molecular population genetic analysis of regulatory regions provides an opportunity to dissect the evolutionary forces that may affect gene transcriptional and post-transcriptional control. Although some of these studies have focused on regulatory genes that directly participate in the construction of organismal morphologies, such as *eve*, *ftz*, and *dpp*, there is still a need for more detailed studies on the microevolution of regulatory sequences among developmental loci. These studies require characterization not only of the *trans*-acting regulatory factors but their *cis*-acting structural gene targets as well, and will help define the evolutionary dynamics of regulated gene expression.

Summary and perspectives

Developmental geneticists have begun a program of comparative studies to systematically assess evolutionary patterns of

regulatory gene expression. Grounded on more detailed investigations in model systems, the hope is that we can correlate regulatory gene evolution with morphological diversification. Evolutionary geneticists are also now in a strong position to begin to address questions on the molecular history of these developmental genetic systems, and on the evolutionary processes that have shaped the structure of these regulatory loci and their functional interactions.

The studies highlighted in this paper are still at too early a stage to draw firm generalizations on how developmental processes evolve at the molecular level, but they do provide directions for possible research into the evolution of developmental genetic systems. Several avenues need to be explored in detail, including the coevolution of interacting pathway members, tests for epistatic selection on regulatory gene alleles, and clear correlations of molecular evolutionary changes with phenotypic variation using modern phylogenetically based comparative methodologies. Some of these studies are already under way, and it is clear that this powerful molecular evolutionary perspective, complemented by developmental genetic studies, should continue to help unravel the evolutionary patterns and processes that underlie the evolution of developmental pathways and organismal morphologies.

Acknowledgments

I thank Jeff Thorne, Jim Mahaffey, Jon Swaffield, Adam Wilkins, members of the Purugganan laboratory, and two anonymous reviewers for their constructive comments. Marianne Barrier assisted in constructing the figures.

References

1 Gould SJ (1977) *Ontogeny and Phylogeny*. Cambridge, MA: Belknap.
2 Raff RA (1996) *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. Chicago: University of Chicago Press.
3 Carroll SB (1995) Homeotic genes and the evolution of arthropods and chordates. *Nature* 376:479–485.
4 Gerhart J, Kirschner M (1997) *Cells, Embryos and Evolution*. Malden, MA: Blackwell Science.
5 Coen E (1991) The role of homeotic genes in flower development and evolution. *Ann Rev Plant Phys Plant Mol Biol* 42:241–279.
6 Palopoli MF, Patel NH (1996) Neo-Darwinian developmental evolution—can we bridge the gap between pattern and process? *Curr Op Genet Dev* 6:502–508.
7 DeSalle R, Carew E (1992) Phyletic phenocopy and the role of developmental genes in morphological evolution in the Drosophilidae. *J Evol Biol* 5:363–374.
8 Coen ES, Nugent JM (1994) Evolution of flowers and inflorescences. *Development (Suppl)*:107–116.
9 Atchley WR, Fitch WM, Bonner-Fraser M (1994) Molecular evolution of the MyoD family of transcription factors. *Proc Natl Acad Sci USA* 91:11522–11526.
10 Sidow A (1992) Diversification of the *Wnt* gene family in the ancestral lineage in vertebrates. *Proc Natl Acad Sci USA* 89:5098–5102.
11 Nei M (1996) Phylogenetic analysis in molecular evolutionary genetics. *Annu Rev Genet* 30:371–403.
12 Zhang JZ, Nei M (1996) Evolution of *Antennapedia*-class homeobox genes. *Genetics* 142:295–303.
13 Ruddle FH et al. (1994) Evolution of *Hox* genes. *Annu Rev Genet* 28:423–442.
14 Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: Diversification of the plant MADS-box gene family. *Genetics* 140:345–356.
15 Purugganan MD (1997) The MADS-box floral homeotic gene lineages predate the origin of seed plants: Phylogenetic and molecular clock estimates. *J Mol Evol* 45:392–396.
16 Ohta T (1989) Role of gene duplication in evolution. *Genome* 31:304–310.
17 Atchley WR, Fitch WM (1995) Myc and Max—Molecular evolution of a family of proto-oncogene products and their dimerization partners. *Proc Natl Acad Sci USA* 92:10217–10221.
18 Kumar S, Balczarek K, Lai CZ (1996) Evolution of the *hedgehog* gene family. *Genetics* 142:965–972.
19 Gibson G, Hogness DS (1996) Effect of polymorphisms in the *Drosophila* regulatory gene *Ubx* on homeotic stability. *Science* 271:200–203.
20 Mackay TFC, Langley CH (1990) Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster*. *Nature* 348:64–66.
21 Templeton AR, Crease TJ, Shah F (1985) The molecular through ecological genetics of *abnormal-abdomen* in *D. mercatorum*. I. Basic genetics. *Genetics* 111:805–818.
22 Kappen C, Schughart K, Ruddle FH (1989) Two steps in the evolution of *Antennapedia*-class vertebrate homeobox genes. *Proc Natl Acad Sci USA* 86:5459–5463.
23 Schughart K, Kappen C, Ruddle FH (1989) Duplication of large genomic regions during the evolution of vertebrate homeobox genes. *Proc Natl Acad Sci USA* 86:7067–7071.
24 Schubert F, Nieselt-Struwe K, Gruss P (1993) The *Antennapedia*-type homeobox genes have evolved from three precursors separated early in metazoan evolution. *Proc Natl Acad Sci USA* 90:143–147.
25 Bailey WJ, Kim J, Wagner G, Ruddle FH (1997) Phylogenetic reconstruction of vertebrate *HOX* cluster duplications. *Mol Biol Evol* 14:843–853.
26 Kappen C, Schughart K, Ruddle FH (1993) Early evolutionary origins of major homeodomain sequence classes. *Genome* 18:54–70.
27 Doolittle RF, et al. (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* 271:470–477.
28 Celniker SE, Keelan DJ, Lewis EB (1989) The molecular genetics of the bithorax complex of *Drosophila*—Characterization of the products of the *Abdominal-B* domain. *Genes Dev* 3:1424–1436.
29 Chadwick R, McGinnis W (1987) Temporal and spatial distribution of trasncripts from the *Deformed* gene of *Drosophila*. *EMBO J* 6:779–789.
30 Horan GSB, Wu K, Wolgemuth R, Behringer RR (1994) Homeotic transformation of cervical vertebrae in *Hoxa-4* mutant mice. *Proc Natl Acad Sci USA* 91:12644–12648.
31 Grenier J, Garber TL, Warren R, Whitington P, Carroll S (1997) Evolution of the entire arthropod *Hox* gene set predated the origin and radiation of the onychophoran/arthropod clade. *Curr Biol* 7:547–553.
32 Pederson J, Kiehart DP, Mahaffey JW (1996) The role of HOM-C genes in segmental transformations—re-examination of the *Drosophila Sex combs reduced* embryonic phenotype. *Dev Biol* 180:131–142.
33 Balavoine G, Telford MJ (1995) Identification of planarian homeobox sequences indicates the antiquity of most *Hox* homeotic gene subclasses. *Proc Natl Acad Sci USA* 92:7227–7231.
34 Kuhn K, Streit B, Schierwater B (1996) Homeobox genes in the cnidarian *Eleutheria dichotoma*: Evolutionary implications for the origin of *Antennapedia*-class (HOM/Hox) genes. *Mol Phyl Evol* 6:30–38.
35 Yanofsky MF (1995) Floral meristems to floral organs—Genes controlling early events in Arabidopsis flower development. *Annu Rev Plant Phys Plant Mol Biol* 46:167–188.
36 Rounsley SD, Ditta G, Yanofsky MF (1995) Diverse roles for MADS-box genes in *Arabidopsis* development. *Plant Cell* 7:1259–1269.
37 Ma H, Yanofsky MF, Meyerowitz EM (1991) *AGL1–AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev* 5:484–495.
38 Doyle JJ (1994) Evolution of a plant multigene family—towards connecting molecular systematics and molecular developmental genetics. *Syst Biol* 43:307–308.
39 Theissen G, Kim JT, Saedler H (1996) Classification and phylogeny of the MADS-box multigene family suggests defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. *J Mol Evol* 43:484–516.
40 Yanofsky MF et al. (1990) The protein encoded by the *Arabidopsis* homeotic gene *AGAMOUS* resembles transcription factors. *Nature* 346:35–39.
41 Mena M et al. (1996) Diversification of C-function activity in maize flower development. *Science* 274:1537–1540.
42 Huang H, et al. (1995) The *Arabidopsis* MADS-box gene *AGL3* is widely expressed and encodes a sequence-specific DNA-binding protein. *Plant Mol Biol* 28:549–567.
43 Munster T, et al. (1997) Floral homeotic genes were recruited from homologous MADS-box genes pre-existing in the common ancestor of ferns and seed plants. *Proc Natl Acad Sci USA* 94:2415–2420.
44 Tandre K, Albert V, Sundas A, Engstrom P (1995) Conifer homologues to genes that control flower development in angiosperms. *Plant Mol Biol* 27:69–78.
45 Hughes AL, Ota T, Nei M (1990) Positive darwinian selection promotes charge profile diversityin the antigen-binding cleft of class I MHC molecules. *Mol Biol Evol* 7:515–524.
46 Easteal S (1990) The patterns of mammalian evolution and the relative rate of molecular evolution. *Genetics* 124:165–173.
47 Li WH, Wu CI, Luo CC (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucelotide and codon changes. *Mol Biol Evol* 2:150–174.
48 McGinnis W, Garber RL, Wirz J, Kuroiwa A, Gehring WJ (1984) A homologous protein coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* 37:403–408.
49 Gaut BS, Doebley JF (1997) DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc Natl Acad Sci USA* 94:6809–6814.

50 Purugganan MD, Wessler SR (1994) Molecular evolution of the plant *R* regulatory gene family. *Genetics* 138:849–854.
51 Kempin SA, Savidge B, Yanofsky MF (1995) Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* 267:522—525.
52 Whitfield LS, Lovell-Badge R, Goodfellow PN (1993) Rapid sequence evolution of the mammalian sex determining gene *SRY*. *Nature* 364:713–715.
53 Tucker PK, Lundrigan B (1993) Rapid evolution of the sex determining locus in Old World mice and rats. *Nature* 364:715–717.
54 DeBono M, Hodgkin J (1996) Evolution of sex determination in *Caenorhabditis*—unusually high divergence of *tra-1* and its functional consequences. *Genetics* 144:587–595.
55 King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.
56 Sturmbauer C, Meyer A (1992) Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* 358:578–581.
57 Vermeij G (1996) Animal origins. *Science* 274:525–526.
58 Omland KE (1997) Correlated rates of molecular and morphological evolution. *Evolution* 51:1381–1393.
59 Shimmin LC, Chang BHJ, Li WH (1993) Male-driven evolution of DNA sequences. *Nature* 362:745–747.
60 Shimmin LC, Chang BHJ, Li WH (1994) Contrasting rates of nucleotide substitution in the X-linked and Y-linked zinc finger genes. *J Mol Evol* 39:569–578.
61 Wallis M (1994) Variable evolutionary rates in the molecular evolution of mammalian growth hormones. *J Mol Evol* 38:619–627.
62 Wallis M (1996) The molecular evolution of vertebrate growth hormones—a pattern of near-stasis interrupted by sustained bursts of rapid change. *J Mol Evol* 43:93–100.
63 Wallis M (1997) Function switching as a basis for bursts of rapid change during the evolution of pituary growth hormone. *J Mol Evol* 44:348–350.
64 Kreitman M, Akashi H (1995) Molecular evidence for natural selection. *Annu Rev Ecol Syst* 26:403–422.
65 Takahata N (1996) Neutral theory of molecular evolution. *Curr Op Genet Dev* 6:767–772.
66 Moriyama EN, Powell JR (1996) Intraspecific nuclear DNA variation in *Drosophila*. *Mol Biol Evol* 13:261–277.
67 Purugganan MD, Suddith JI (1998) Molecular population genetics of the *Arabidopsis CAULIFLOWER* regulatory gene: Nonneutral evolution and naturally occurring variation in floral homeotic function. *Proc. Natl. Acad. Sci. USA* **95:**8130–8134.
68 Hanson MA et al. (1996) Evolution of anthocyanin biosynthesis in maize kernels—the role of regulatory and enzymatic loci. *Genetics* 143:1395–1407.
69 Ayala FJ, Hartl D (1993) Molecular drift of the *bride-of-sevenless* (*boss*) gene in *Drosophila*. *Mol Biol Evol* 10:1030–1040.
70 Richter B, Long MY, Lewontin RC, Nitasaka E (1997) Nucleotide variation and conservation at the *dpp* locus, a gene controlling early development in *Drosophila*. *Genetics* 145:311–323.
71 Walthour CS, Schaffer SW (1994) Molecular population genetics of sex determination genes—the *transformer* gene of *D. melanogaster*. *Genetics* 136:1367–1372.
72 Berry AJ, Ajioka JW, Kreitman, M (1991) Lack of polymorphism on the *Drosophila* fourth chromosome resulting from selection. *Genetics* 129: 1111–1117.
73 Lai CG et al. (1994) Naturally-occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila*. *Science* 266:1697–1702.
74 Swalla BJ, Jefferey W (1996) Requirement of the *manx* gene for expression of chordate features in a tailless ascidian larva. *Science* 274:1205–1208.
75 Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature* 386:485–488.
76 Ludwig MZ, Kreitman M (1995) Evolutionary dynamics of the enhancer region of *even-skipped* in *Drosophila*. *Mol Biol Evol* 12:1002–1011.
77 Odgers WA, Healy MJ,, Oakeshott JG (1995) Nucleotide polymorphism in the 5′ promoter region of *Esterase-6* in *Drosophila melanogaster* and its relationship to enzyme activity variation. *Genetics* 141:215–222.
78 Leicht BG, Muse SV, Hanczye M, Clark AG (1995) Constraints on intron evolution in the gene encoding the mysoin alkali light chain in *Drosophila*. *Genetics* 139:299–308.
79 Jenkins DL, Ortori CA, Brookfield JFY (1995) A test for adaptive change in DNA sequences controlling transcription. *Proc R Soc Lond B* 261:203–207.
80 Schulte PM, Gomez-Chiarri M, Powers DA (1997) Structural and functional differences in the promoter and 5′ flanking region of *Ldh-B* within and between populations of the teleost *Fundulus heteroclitus*. *Mol Biol Evol* 145:759–769.