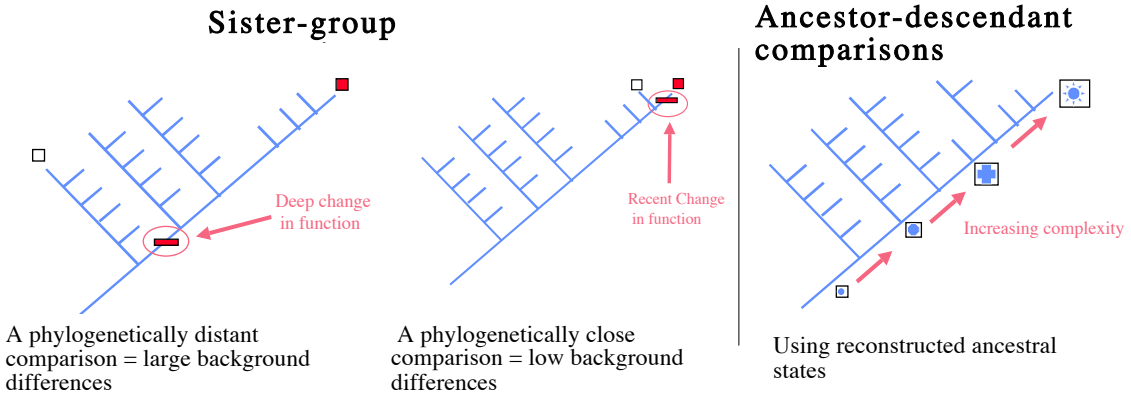
April 20, 2020. **Comparative genomics; evo-devo**

A. Phylogenomics

This is the era of whole-genome sequencing; molecular data are becoming available at a rate unanticipated a few years ago. Sequencing projects in a number of countries have produced a growing number of fully sequenced genomes, providing computational biologists with tremendous opportunities. Something can be learned about the function of genes by examining them in one organism. However, a much richer array of tools is available using a phylogenetic approach (Eisen & Fraser, 2003. Phylogenomics: intersection of evolution and genomics, *Science* 300; 1706-1707) .

As with other traits, there are two types of useful phylogenetic reasoning.. (1) Close sister-group comparisons between lineages differing in a critical phenotype (e.g., desiccation or freeze tolerance) can allow a quick narrowing of the search for genetic causes. (2) Dissecting a complicated, evolutionarily advanced genotype/phenotype complex (e.g., development of the angiosperm flower), by tracing the components back through simpler ancestral reconstructions, can lead to quicker understanding. Hence, phylogenomics allows one to go beyond the use of pairwise sequence similarities, and use phylogenic comparative methods as discussed in this class to confirm and/or to establish gene function and interactions.



B. Evolution and development ("evo-devo")

The last frontier in our understanding of biological forms is an understanding of their developmental origins. Much of the ultimate control of form resides in the genome, yet much also resides in the environment (at levels from the internal cellular environment to the external habitat). The highly interactive and complex nature of developmental processes make it impractical to deduce phenotype from genotype based on first principles.

We need to carefully keep in mind what we mean by "homology" as well. Remember back to the discussion last week about "hemiplasy." For a trait with a complex development involving many interacting loci and several subparts, only portions of its development may be homologous, and other portions may be homoplastic.

The phenotype is an emergent property and its origin can be studied most efficiently by backtracking from the phenotype itself to its structural, physiological, developmental, ecological, and genetic causes.

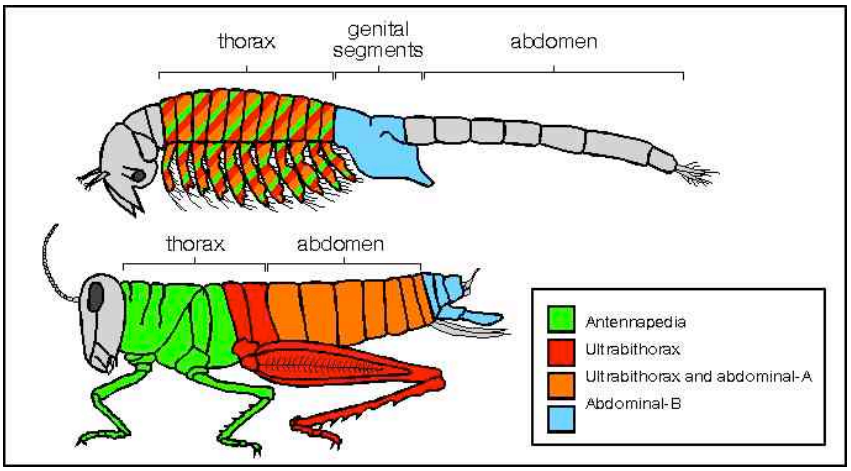
C. Ontogeny and genetics

1) Expression studies
-- use of reporter genes
-- EST studies (cDNAs from target tissues)
-- transcriptomics
-- real time PCR
-- microarray

2) Forward genetics
-- starts with a phenotype and moves towards the gene
-- screen for & isolate relevant mutants
-- map locus through genetic crosses, chromosome mapping, association studies
-- quantitative trait locus (QTL) analyses. Identifying SNPs that correlate with a phenotypic trait.
-- isolate gene & sequence

3) Reverse genetics
-- Starts with a particular gene and assays the effect of its disruption
-- Knockouts of candidate genes by transformation, observe change in phenotypes
-- gene silencing (RNA interference)

4) Gene family evolution
a. Hox genes in animals



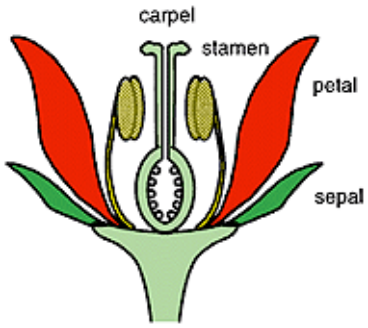
Hox genes are a subset of homeobox genes. Might have arisen by rounds of duplication of an ancestral gene, followed by a quaduplication of the cluster in mammals. Partially overlapping zones of expression which vary in the anterior extent of their expression define distinct regions. Tandem gene duplication can allow retention of gene while new functions are adopted by one copy. Hox gene cluster arose from rounds of tandem duplication. Vertebrates have four Hoxgene complexes. *Amphioxus*, a vertebrate-like chordate, has one Hox cluster which may be close to ancestral Hox complex. (taken from http://www.mun.ca/biology/desmid/brian/BIOL3530/DB_Ch15/BIOL2900_EvoDevo.html)

b. The ABC model in (some) flowering plants

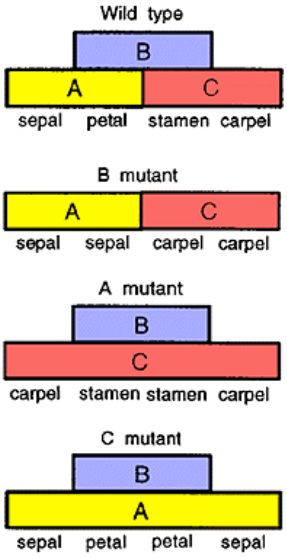
The MADS box is a highly conserved sequence motif found in a family of transcription factors. By now, more than hundred MADS box sequences have been found in species from all eukaryotic kingdoms. The family of MADS domain proteins has been subdivided into several distinct subfamilies. Most MADS domain factors play important roles in developmental

processes. Most prominently, the MADS box genes in flowering plants are the "molecular architects" of flower morphogenesis (source: The MADS-box Gene Home Page; http://www.mpizkoeln.mpg.de/mads/).

MADS-box genes and the ABC model of organ identity determination
Source: Gerbera Lab, Univ Helsinki http://honeybee.helsinki.fi/MMSBL/Gerberalab/abc.html

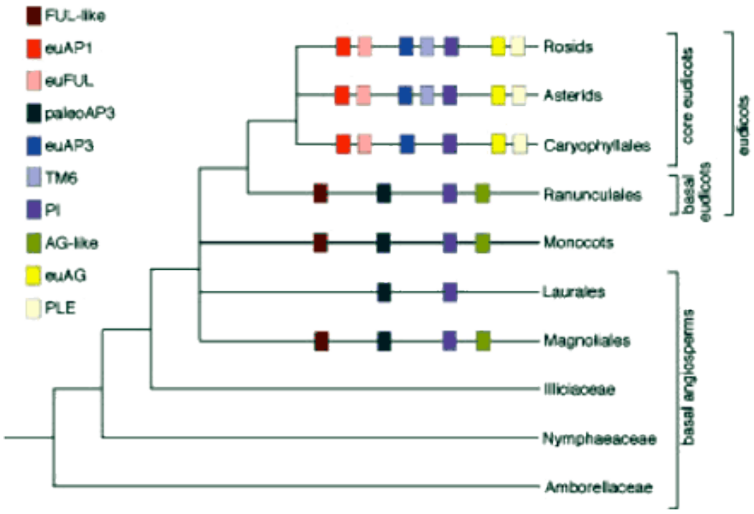


The basic structure of a complete flower consists of four concentric whorls. A simple model has been proposed to predict organ formation in flowers, where three classes of homeotic genes, the so-called ABC-class genes, act alone or together to give rise to sepals (A), petals (A+B), stamens (B+C), and carpels (C).



According to the ABC-model, organ determination in the whorls depends on the combinatorial action of three regulatory functions. A mutation disrupting one of the functions causes a homeotic change in organ identity. Note that the A and C functions are negatively regulating each other: Mutation in one causes expansion in the expression domain of the other. Molecular cloning has indicated that most of these ABC homeotic genes encode a well conserved DNA binding domain, the MADS box, and that this domain has been shown to be capable of binding to specific DNA sequence motifs known as CArG boxes. Because of their essential roles in flower development, and due to the high degree of conservation in the MADS box domain, MADS box genes have been cloned from diverse angiosperm plant species, including petunia, tomato, maize, white campion, sorrel, gerbera, and even one gymnosperm species, spruce. Although the ABC model has been shown to apply in several species other than the model species *Arabidopsis* and *Antirhhinum*, the precise functions of most MADS box genes remain unclear. For general reviews see: Kramer & Hall, 2005. [Evolutionary dynamics of genes controlling floral development, *Current Opinion in Plant Biology* 8: 13-18], and Heijmans, Morel, & Vandenbussche, 2012. [MADS-box genes and floral development: the dark side, *Journal of Experimental Botany* 63:5397–5404].

It seems that, in addition to their essential roles during floral development, MADS box genes act also as regulators for various other aspects of plant development; homologs are found in most Eukaryotes! This is a good example of repurposing of genes, probably though duplication and subfuntionalization.



MADS box gene duplications. Genes corresponding to the different MADS box gene lineages are indicated in the clades where they have been identified *(Irish VF (2003) The evolution of floral homeotic gene function. BioEssays 25:637-646*

.

5) Biochemical pathways

These are great examples of the relationship between ontogeny and phylogeny. By far the easiest way for chemical diversity to evolve is through change in existing biosynthetic pathways. An example from Kip Will's current research:

Bombardier beetles. Geadephaga is the largest clade of organisms that use a single homologous gland system to produce no less than 19 distinct classes of chemical compounds for defense. Four lineages of quinone producing carabid beetles, including four species commonly known as the bombardier beetles, chemically blast their defensive quinones at extremely hot temperatures (up to 100 °C). Transcriptomes for genes involved in quinone production can be examined comparatively to elucidate chemical biosynthetic pathways, and describe the genetic architecture of quinone evolution. The evolutionary history of quinone biosynthesis in carabids can be studied by inferring the phylogenetic history of candidate gene families using the tree topology and branch lengths to test whether genes are ancient and shared among taxa, or if gene diversification is recent and specific to certain lineages. The hypothesis is that the genes up-regulated in secretory cells during quinone synthesis are closely related to those involved in quinone production in the arthropod cuticle.