

### Video Tutorial 3.1: The evolution of Hox genes

Homeotic Hox genes were first discovered in fruit flies through the striking phenotypes they can give rise to when the genes are mutated. In these homeotic phenotypes one part of the body along the anterior to posterior (or head to tail) axis is converted to the likeness of another part of the body.

In this example you can see the normal structure of the antenna on the head of a wild-type fruit fly here on the left. But in this fly on the right, that has a homeotic mutation, you can see that this antenna on the head has now been transformed into a leg structure. A structure usually only found on the thorax of the fly. Striking homeotic phenotypes like this are part of why the Hox genes have garnered so much interest.

You can see in this diagram that different regions of the fly express different Hox genes. And it turns out that the normal role of the Hox genes is to provide regional identity to cells along the anterior to posterior axis of the body so that cells have the information needed to form the correct structures in the correct places.

Hox genes achieve this function by regulating other genes. The encoded Hox proteins can bind to DNA and regulate entire sets of other genes that are needed to make the correct structures for a particular region of the body.

When the first Hox genes were sequenced it became clear that they all contained a short stretch, of 180 basepairs, of sequence that showed similarity to all the other Hox genes. This stretch of sequences was called the Homeobox and it encodes the part of the Hox protein that directly binds to DNA. This stretch of the protein is called a Homeodomain. So the DNA sequence is a homeobox and the corresponding encoded protein domain is the homeodomain. The homeodomain generally forms a Helix-Loop-Helix-Turn-Helix structure that allows it to bind to DNA.

There are other genes, not just the Hox genes, that contain Homeoboxes.

Homeobox containing genes in fact form a large and diverse family of genes of which the homeotic Hox genes form just a specific subset.

Excitingly, since their discovery in fruit flies these Hox genes have been shown to play a role in patterning the anterior to posterior axis of most animals, including humans. Hox genes tend to be found in clusters in the genomes of animals. And changes to the exact number, sequence and regulation of the Hox genes have clearly played an important role in the evolution of different body plans in animals as we describe in Chapter 3.

We can therefore make important insights into the evolution of animal bodyplans by tracing the evolution of Hox genes over time.

We can search genome sequences for similarities to the Hox homeobox sequences as a way of identifying Hox genes in the genomes of animals that are around today. But how do we look back across evolutionary time at the history of these genes?

Well, we can trace the evolution of gene families like the Hox genes by mapping the genes onto an evolutionary tree, a phylogenetic tree, of the species of interest.

For example, let's take a look at Hox genes in vertebrates like ourselves. When we look at humans we find Hox genes clustered together in four different chromosomal locations in our genome. Each cluster of Hox genes contains varying numbers of genes indicating that there have been duplications and losses of different genes in different clusters. And based on sequence similarities the Hox genes can be classified into four main types indicated by the different colors here.

But how did we end up with this particular configuration of Hox genes? And, more specifically, what did the Hox genes of the last common ancestor of all vertebrates look like?

Let's take a look at an evolutionary tree of vertebrates (animals with a backbone like ours). When we look at the Hox genes of other mammals, like mice, we find that they have a very similar number of Hox clusters and genes within the clusters to humans. If we look across the tree a little bit further to birds and other tetrapods (those are the four limbed vertebrates) we also see four Hox clusters but occasional loss and gain of some of the individual genes. When when we look at even more distantly related vertebrates like fish we see more variety in Hox composition. The model organism, zebrafish, for example has 7 Hox clusters, many other fish have 8 clusters and salmon have even more clusters. And throughout the fish lineages we also see variation in the composite of genes present in each of the clusters.

So clearly there have been changes, through duplications and deletions of individual genes and likely duplications and deletions of entire clusters of genes, even entire genomes, over the course of vertebrate evolution. And we can see that to look further back in evolutionary time to try to understand what the Hox genes of the last common ancestor of all vertebrates looked like it would be really helpful to reach further afield to more distantly related animals that are close relatives of vertebrates but not themselves vertebrates.

Animals that sit at this key position on the tree of life are non-vertebrate chordates like this animal, Amphioxus.

Chordates are animals that have a structure called a notochord. The notochord runs along the length of the spinal cord and provides structural support to the body. In vertebrates the notochord becomes reduced as the animals develop and as the bones of the vertebral column form; those bones are the vertebrae that we, and all vertebrates, have in our backs.

But in the non-vertebrate chordates like Amphioxus the notochord persists and these animals don't form the bones of a vertebral column. These non-vertebrate chordates, are the closest relatives of the vertebrates. Analysis of their genomes is therefore particularly interesting because of the potential this analysis has to shed light on what the genome of the very first vertebrate ancestors might have looked like.

Researchers have therefore analyzed the genome of animals like Amphioxus to identify the Hox genes present. They have found that Amphioxus contain a single Hox cluster with 15 genes (it seems they have actually duplicated some of the more posterior Hox genes, giving them some extra genes in this group).

From this we can infer that the ancestral precursor to all vertebrates likely had a single Hox cluster of about 14 genes. This cluster then became duplicated twice, probably through two rounds of whole genome duplication, resulting in four Hox clusters in the last common ancestor of all vertebrates.

Individual Hox genes then underwent some additional changes in the lineages giving rise to modern day tetrapods like ourselves. And at the base of the fish lineage it is likely that there was an additional round of whole genome duplication increasing the number of clusters from 4 to 8 Hox clusters in ancestral fish. This was then followed by a variety of additional single gene or single cluster duplications and losses in the various different fish lineages. In the Zebrafish lineage there appears to have been loss of one entire cluster bringing the total number of clusters from 8 to 7, while in the salmon lineage for example there appears to have been a fourth round of whole genome duplication, followed by some gene losses. Overall the fish show quite a lot of individual variations in the exact composition of Hox genes against this general backdrop of whole genome duplication to 8 clusters in the ancestral fish species.

We can extend this same approach of surveying the composition of Hox genes, not just in vertebrates and their nearest relatives but to all multicellular animals and their nearest relatives to ask what the composition of Hox genes might have been in the last common ancestor of all animals.

Here we have a simplified evolutionary tree, a phylogenetic tree, of multicellular animals. There are two main branches to this part of the animal kingdom, the deuterostomes (like us humans) and the protostomes (like fruit flies). These are all considered multicellular bilaterian animals (animals that at some stage in their lifecycle have a bilateral or left-right symmetry to their bodyplan).

Their closest relatives include simpler multicellular animals like Cnidarians (such as sea anemones and jelly fish), Ctenophores (which are comb jellies), Placozoans and sponges. Analysis of Hox genes from all these lineages reveal that the last common ancestor of flies and vertebrates likely had a cluster of about seven distinct Hox genes and these likely functioned in providing regional identity to cells along the antero-posterior axis as we see in modern day flies and vertebrates.

Looking further back gets complicated as it's tricky to discern and interpret sequence similarities over such large evolutionary distances. However it is clear that Cnidarians do also contain a number of Hox genes in their genomes too, so the last common ancestor here on the tree before Cnidarians and the bilaterian animal lineages split likely also had at least three or four different types of Hox genes, though their exact functions are less understood.

By contrast to the Cnidarians, the known genomes of sponges contain a variety of homeobox genes but no clear Hox genes. Even more distantly related fungi and plants also contain homeobox genes but not Hox genes. It is therefore likely that the Hox genes arose early on in the evolution of multicellular animals some time after the sponge lineage had already diverged away from other multicellular animals.

Overall then, we can gain insight into the evolution of gene families like the Hox genes by tracing their ancestry across evolutionary trees that include the species of interest as well as some of their nearest relatives at key positions on the evolutionary tree.

References and further reading:

Ferrier, D.E.K (2012) "Evolution of the Hox Gene Cluster". Chichester: eLS John Wiley & Sons, Ltd.

Holland, P.W., 2013. "Evolution of homeobox genes." Wiley Interdisciplinary Reviews: Developmental Biology, 2(1), pp.31-45.

Pascual-Anaya et al. (2013) "Evolution of Hox gene clusters in deuterostomes" BMC Developmental Biology 13:26.

Image Attributions:

Frog: By Brian Gratwicke [CC BY 2.0 (<http://creativecommons.org/licenses/by/2.0/>)], via Wikimedia Commons

Coelacanth: By No machine-readable author provided. Bogdan assumed (based on copyright claims). [GFDL (<http://www.gnu.org/copyleft/fdl.html>), CC-BY-SA-3.0 (<http://creativecommons.org/licenses/by-sa/3.0/>) or CC BY-SA 2.5-2.0-1.0 (<http://creativecommons.org/licenses/by-sa/2.5-2.0-1.0/>)], via Wikimedia Commons

Fish: By George Chernilevsky (Own work) [Public domain], via Wikimedia Commons

Shark: By Knepp Timothy, U.S. Fish and Wildlife Service [Public domain], via Wikimedia Commons

Hagfish: By J. H. Richard (Contributions to the fauna of Chile Charles Girard) [Public domain], via Wikimedia Commons.

Sturgeon: By Duane Raver/U.S. Fish and Wildlife Service [Public domain], via Wikimedia Commons

Amphioxus: By Jon Houseman (Jon Houseman and Matthew Ford) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0/>)], via Wikimedia Commons and © Hans Hillewaert / via Wikimedia Commons

Dog (Lola), Gecko, Bird, Flies: Photographs by Rachel Hoang, PA, USA.