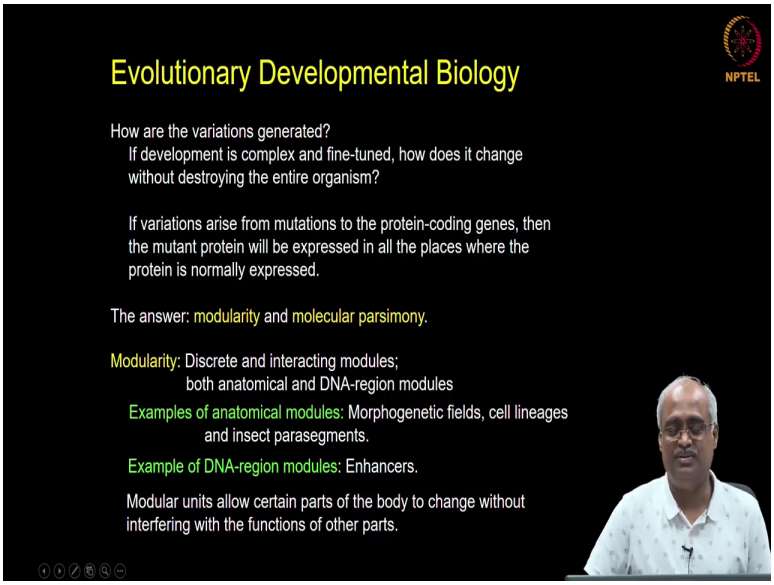
Introduction to Developmental Biology
Prof. Subramaniam
Department of Biotechnology
Indian Institute of Technology- Madras

Lecture No – 26
Evolutionary Developmental Biology (Part 1 of 3)

(Refer Slide Time: 00:15)



Students, welcome back to developmental biology course again. Today we are going to have a new topic and this is going to be what we call as evolutionary developmental biology. So I am sure by now all of you have heard the term evolution. So the theory of evolution is basically the unifying central concept of biology. So as famously you know told by Dobzhansky long time ago, nothing in biology makes sense except in the light of evolution.

So the process of evolution operates by natural selection of the fittest. So this is what you would have learned from your high school and in earlier classes in college or university. So, now for the natural selection to select the fittest you need to have variations in the population in a given population. For example let us say in a human population or if everybody let us say is uniformly 5.5 feet height, then there is no way to select based on height because everyone will be equally competent. If height becomes desirable let us say shorter height is desirable for a given environment.

So variations have to exist in the population then only you can find the shortest person who might fit a given environment. So therefore the basis for natural selection to work is the availability of variations in a

given population. So how are these variations created? Where do they arise from? So they have to arise from variations in the development. So variations in development are what are going to result in a population of human beings. In our example with varying heights so how are these variations permitted in development.

So what is the mechanistic basis of these variations and understanding of that is really at the center of understanding the mechanisms of how evolution operates. And in addition, what are the variations that are permitted by the evolution developmental framework that exists. Like for example if Hox genes that we discussed in the last class, determine the anterior-posterior pattern and segment identity for example in in our vertebral column.

Now what variations are possible within that? What are the limitations and constraints? That is going to determine what is available for evolution to work on and therefore understanding the developmental mechanisms of these features are really critical to explain what is possible in evolution and to predict what kind of structures are going to happen and how that is going to be adapted to a given environment. See in this sense evolution and development are very tightly related and people very affectionately call it as evo-devo.

So that is what we are calling as evolutionary developmental biology. So this is what we are going to discuss in today's class. So as I just said, how are the variations generated? So this is a not an easy question, it is quite complex. So consider the following points and you will understand why this is not a straightforward question. I will take the example of the body patterning in the drosophila or the mammalian axis formation that we discussed in the last class. Everything seems to be very fine tuned.

You know if you want the chicken type of cervical vertebral column, then you have to have Hox gene stopping a little earlier. It should restrict a little bit to the posterior compared to what it is in mouse. So these seem to be very fine tuned and very delicately balanced and tightly controlled. If that is the basis of development, then any variation could end up destroying the whole organism right? The structures will just not fit with the rest of the body plan and as a result that embryo might die.

So, how variations are permitted in such a complex and fine-tuned developmental program? So that is one question. Second, most of us from our molecular biology classes, we would have already learned that our DNA replication is error-prone. The repair mechanism and proofreading notwithstanding and as a result you do get mutant versions of proteins. And if these mutant questions can have differing activities and

that would cause variations then these proteins will be expressed everywhere. We discussed at length about the genome equivalence like all our somatic cells more or less have the same genomic content.

So that means the mutant protein will be there in all parts of the body. Then how that is going to help us create variations? For example, if a protein is involved in height and mutant version of that protein is going to increase or decrease the height, will that protein not have impact in the rest of the embryo development other than height? So even the advent of molecular biology does not provide a straightforward answer to this question of how are variations generated.

So the answers actually come from developmental mechanisms. Remember in the early classes, we learnt about comparative anatomy and even in Hox genes, we saw the comparative development like comparing the expression pattern of Hox genes in chick versus mouse and the number of cervical vertebrae in chick versus mouse.

So we already are familiar with it and that is what provides answer to this question of how variations arise. And the answers are quite simple. In two phrases, it is summarized. One is modularity and another one is molecular parsimony. So these two, we will discuss at length to understand how these two features that the developmental biologists have learnt, provide answer to this question of how variations arise.

We are somewhat familiar with the term, modularity already when we discussed about enhancers. So enhancers are the ones that modulate the transcriptional rate of genes. These enhancers can be upstream in the promoter region or it could be anywhere else, even in the intron or coding sequence and these are cis-acting elements DNA sequences that modulate the rate of transcription. And we saw that these exist in modules like for example in PAX6 gene, you have an enhancer that works in lens, another enhancer that works in pancreas and so on.

So each module is active in one tissue or one time point and not at another tissue or time point. So therefore, they are modules. You have a lens specific module. You have a pancreas specific module and so on. So these are modules of DNA sequences and similarly there are developmental modules. They are discrete and interacting anatomical modules. So when you think of modules do not always think enhancers. Even in this narrow context, there are other modules like tissue modules that exist. Examples are morphogenetic field.

So, you have the retinoic acid posterior to anterior morphogenetic field and you have FGF gradient from posterior to anterior. And these gradients act as modules because they are independent and they function completely independent of the rest of the embryo. For example, if you take our epithelial-mesenchymal interaction discussion, where we saw a mesenchyme from a certain region induces the epithelium to form corresponding structure, be it the body feather of birds or it could be the scales in the thigh region and claws at the tip of the limb.

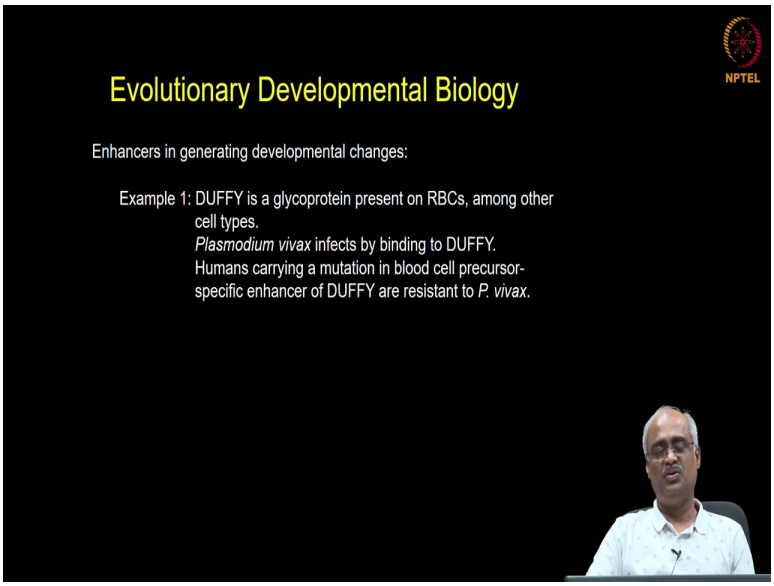
So these are all modifications of the epithelium but they are instructed to make the wing feather, claw or scales by the underlying mesenchyme. Now if you mismatch them, what you find is the epithelium acts as a module. It listens to the underlying mesenchyme and then it behaves that way. So we saw the same in embryonic transplantation experiment between salamanders and frog.

So salamander makes this structure it normally makes based on the modular instruction coming from the frog-underlying tissue that was transplanted. So these are modules. So they function in a way that is independent of what is happening in the rest of the embryo. Some of the cell lineages and the insect parasegments, they all do that. So we saw antennapedia not being expressed in the posterior-most thoracic segment which, meant it did not make wings.

But if you do not have the posterior Hox gene, that ultrabithorax, antennapedia extends into the third segment and now the third segment behaves like the second segment, making a wing. So these are all acting like modules. So these modules, depending on where they are, could actually modify the structures. So we will see many examples, where it will all become very clear, the modular behavior of tissues and the modular behavior of DNA sequences that regulate transcription by stretching them in terms of spatial expression pattern or temporal expression pattern. You can modify the structures, say converting you a normal fingers like digits into webbed structure, making the bat wing and so on.

So the main crux of this is, modular units allows certain parts of the body to change. That is why they are independent without interfering with the functions of the other parts. So this is how variations are made possible.

(Refer Slide Time: 14:06)



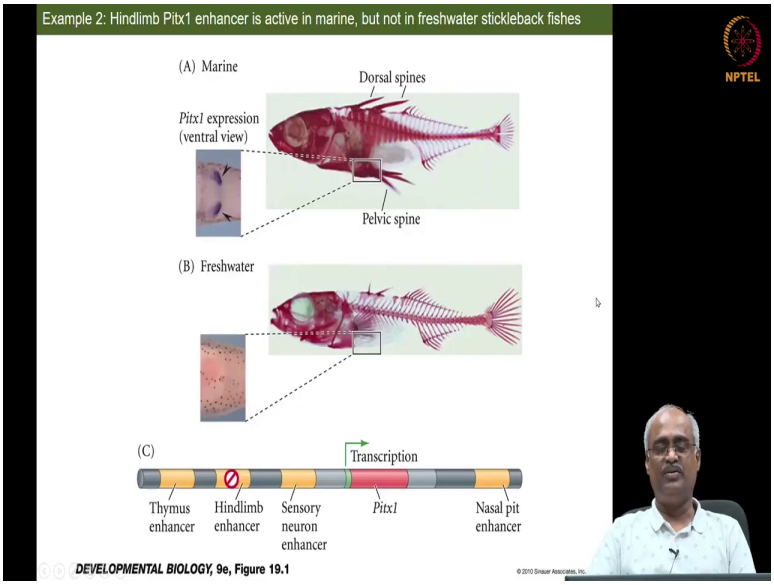
Use of enhancers led to variations which made possible for the natural selection to work and select what fits a given environment. So we have a receptor on our RBCs and also in many other cells called DUFFY, it is a glycoprotein and it is essential for many signaling and it is an essential protein and unfortunately this *plasmodium vivax* (which is a relative of *plasmodium falciparum*) causes malaria and it infects by binding to DUFFY protein present on RBCs.

So many pathogens get into our cells by binding to cell surface receptors. Many viruses also do that. So the *plasmodium vivax* binds to DUFFY on RBCs and that is how it starts the infection. So now it turns out that in certain population of human beings in Africa, they are resistant to *plasmodium vivax* and that is because their blood cells do not express DUFFY. So how come blood cells do not express DUFFY and they are still alive even though DUFFY is an essential protein and is required for survival.

And if these people have a mutation in DUFFY and it is not functional, then how are they surviving. So the answer comes from a specific enhancer in the precursors of the blood cells where the enhancer of DUFFY gene is inactive. As a result, DUFFY is not expressed in the blood cells but it is expressed elsewhere where its function is essential. So as a result, this enhancer mutant that specifically turns off DUFFY in the blood cells has been selected due to the evolutionary pressure of *plasmodium vivax* infection.

So this is how the modularity here in terms of the gene expression helped in generating a variation on which natural selection could work.

(Refer Slide Time: 16:54)



Another example is the fish variety called the sticklebacks because they have these spine-like structures and therefore they are called sticklebacks. So these have thorny spines at the pelvic region in the ventral region of the body.

And these pelvic spines help them avoid predators because these thorns prick them and as a result, the predators do not like to eat these fishes. So this is a survival adaptation that helps it to dodge the predator but when it came to fresh water, there are no such predators for it and therefore this pelvic spine is not required. But then why should it lose the spines? So why did they lose, because there are invertebrate predators here that readily hold these fishes using the spine structures.

So therefore the sticklebacks that lost the ability to make the pelvic spine got selected in the freshwater. So now how come one could have pelvic spine and the other lost it. So the scientists tried to do mutations and finally through those experiments, identified that a mutation in a gene called Pitx1. When they looked at the amino acid sequences, there was absolutely no difference at all between the Pitx1 protein sequence from the marine sticklebacks and the freshwater sticklebacks.

They were identical but then when they looked at the expression pattern, they realized in the ventral part a strong Pitx1 expression is present in the marine water sticklebacks but not in the freshwater sticklebacks.

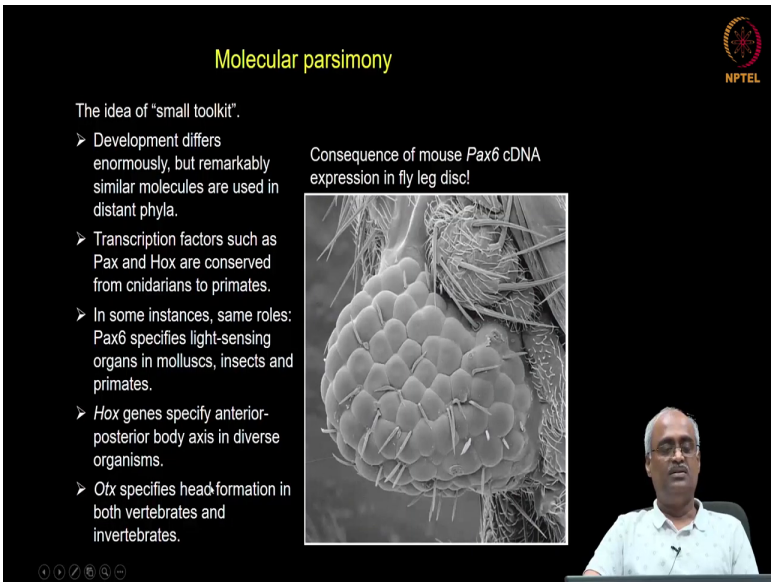
But when they took the amino acid sequences and fused it with the GFP, a reporter protein and introduced into the freshwater fishes, they could see GFP fluorescence in this region of these fishes. And more dramatically when they used the enhancer elements taken from the marine sticklebacks and expressed in

the freshwater Pitx1 gene, fresh water sticklebacks developed pelvic spines. With this experiment, they were able to find out that hind limb enhancer had mutations.

The sequence difference between these two varieties is not in the coding sequence instead it is in the enhancer and due to this mutation, this gene is not expressed in the pelvic region in the freshwater sticklebacks. So this is again example of how turning off a certain enhancer module helps in turning off that gene expression specifically in one part of the body but not in the rest of the places. So these are two really good examples explaining the role of modular enhancers in generating developmental variations.

So, now we move to the next concept. Now we will look at molecular parsimony. What is molecular parsimony and how that helps in generating variations?

(Refer Slide Time: 22:16)



Parsimony means being miserly, using very few things for doing whole lot of things; being frugal. In developmental biology, we call that a small toolkit. So here the developmental mechanism seems to use very few tools to build varieties of structures, morphology and functions.

You can easily imagine this in a machine making world. A screwdriver to tighten a screw in your bicycle can also be used to tighten a screw in an airplane. We are using fewer tools but on very complex machines, a similar thing happens in development as well and that is what we call as molecular persimony. We already have a very good example for this in our recent discussions on Hox genes and anterior-posterior axis.

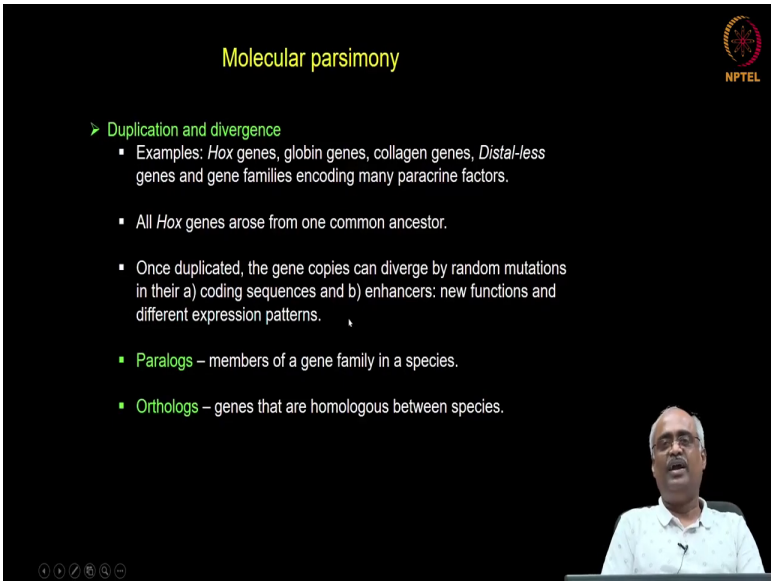
Whichever Hox genes determine segmental identity in drosophila, the same Hox genes determine similar anterior-posterior segment identity in vertebrates. So that explains what is molecular parsimony in a nutshell.

But remarkably similar molecules are used in distant phyla as well. Transcription factors such as Pax6 in eye development and the Hox genes are conserved from cnidarians to primates. Cnidarians are very ancient organisms like sea anemone. They are radially symmetrical organisms that formed even before flatworms, roundworms, insects etc

So starting from cnidarians to primates these genes do similar things; their sequence is conserved and they do very similar functions as well. Pax-6 specifies light sensing organs in molluscs and it does the same in us and in insects. So the use of Pax-6 over a period of about 700 million years has been the same and it is not the sequence alone that is conserved, the very specific organ for which it is required is also conserved. And same is the case with Hox genes.

Pax-6 is required for eye development in all the organisms from molluscs to primates. When Pax-6 is expressed in an insect in the place where its leg will develop, it can instruct the leg imaginal disc to produce eye in the place of leg. So that is the level of conservation. So this is called deep homology which we will discuss in detail a little later. So here we are seeing these things as examples of using small toolkit.

(Refer Slide Time: 28:45)



So starting from ancestral molecules, let us say you have an ancestral Hox which was duplicated and then further duplicated and the two groups of duplicates diverged in different way forming the anterior group that is the antennapedia cluster and then you have the posterior bithorax cluster. So that is how it should have happened and that process is called duplication and divergence. So if a particular gene is essential for early embryonic development, then mutation in that particular gene is going to be lethal.

So such things are not going to be available for the natural selection to work upon but instead during DNA replication , if error happens that might end up duplicating certain proportions of the chromosome, sometimes it is one short stretch of DNA sometimes it can be on a large part of a chromosome.

So once these duplications happen, instead of one copy of the gene, you have two copies. Now the evolutionary pressure is sort of relaxed. As long as one copy remains intact, the embryonic development is going to be all right. And the other copies can be now tinkered with. Its sequences will not have a bigger consequence on the embryonic development and that is how divergence starts.

And this divergence now can lead to new functions, new spatially distinct expression like for example one duplicate expresses in one tissue and the other one expresses in other tissue, where a slight modification in coding sequence could make the tissue develop differently and that could become a selection advantage and so on. So these duplications enable divergence. So many gene families come from such founding members and then you have a group of related genes, that is how you have Hox gene family, globin gene family, collagen gene family, Distal-less family etc.

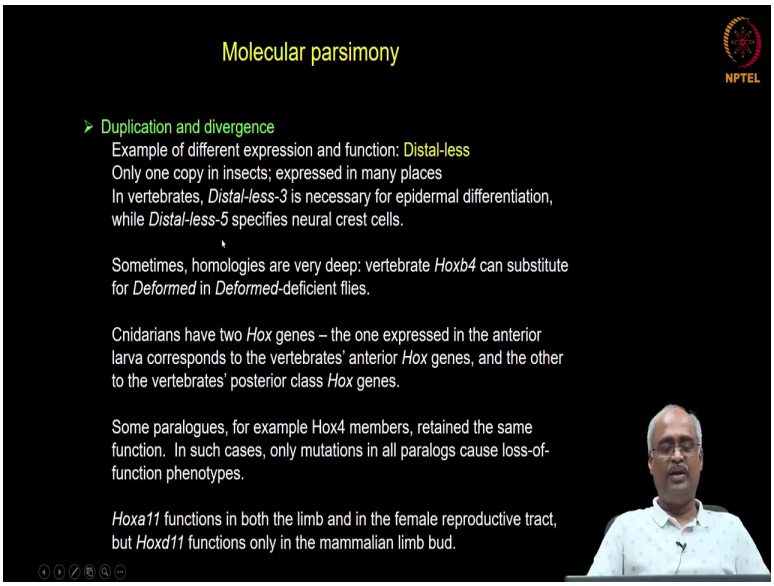
And random mutations can also be in enhancers like in the case of Pitx1 where the hind limb enhancer got mutated in the freshwater sticklebacks. So due to this you could get new functions or different expression pattern. So all this becomes possible when you have multiple copies of genes and therefore divergences in terms of new functionality as well as new domains of expression both become possible.

And such duplicates of the same gene with subtle variations present in a given species are all together called paralogs. So for example, antennapedia is a paralog of Abd-A because both are related proteins. They are homeotic genes and are more closely related to each other than to other gene families but they are all present in drosophila itself and hence, they are paralogs. But if you take Hox4 of vertebrates like humans and look at it in drosophila, you will surely find it to be more related to antennapedia than any other gene family and that you call as orthologs.

So related genes present in different species are called orthologs. These terms apply to proteins as well. So members of a gene family in a species are called paralogs and members of the gene family present in different species, where the homology is across species, are called orthologs. So we continue on this theme to understand some of these dramatic variations that have been accommodated during evolution and they are presented here in bullet forms, each one with a different example.

If you take Distal-less, there is only one copy in insects and it is expressed in most parts of its body. But on the other hand, if you take the case in vertebrates where you have multiple copies, for example Distal-less-3 is required for epidermal differentiation and Distal-less-5 is expressed in neural crest only. So there is specialization here. Instead of one person doing all the jobs, you have employed many more. So that is the advantage of this duplication and divergence.

(Refer Slide Time: 35:41)



If variations in neural crust is going to generate a new body structure that is going to be evolutionarily advantageous, then mutations in Distil-less-5 could enable that without compromising the epidermal differentiation; because a different paralog is taking care of that job.

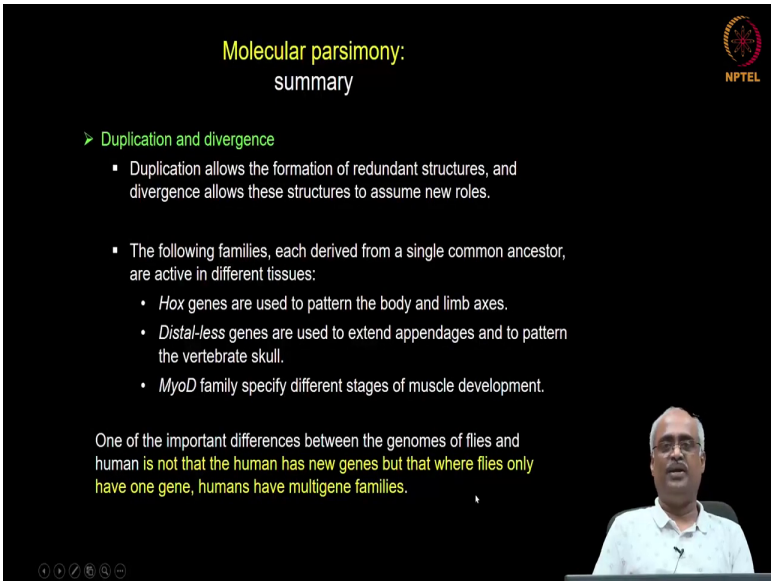
And there are many more such examples. Vertebrate Hox4 can substitute its ortholog in flies that is the Deformed-deficient mutant fly which could be rescued to wild type by expressing the vertebrate Hox4 and where it all starts? It all starts from cnidarians which are like sea anemone kind of organisms, mostly marine although some are there in freshwater as well.

So they have two Hox genes and one of them that is expressed in the anterior part of its larval stage corresponds to the vertebrate's anterior Hox group and its posterior one corresponds to the posterior class of Hox genes. So Hox genes seem to have evolved only once before cnidarians and duplicated in cnidarians. Later, the anterior and posterior seem to have duplicated further and diverged to have 5 anterior and 3 posterior for example in insects.

And then they duplicated again four times to have 4 clusters in vertebrates which we saw in the last class. And there seemed to have been two large such duplications in Hox gene families, one in the vertebrates and another separately in fishes. So we are not going to get into the details; but the main point here is that if you take the example of Hox4, all 4 paralogs of Hox4 do very similar things and therefore, mutating any one of them is not going to create a phenotype and you need to lose all of them.

On the other hand, genes like Hox11 A and D are not expressed in the same place and doing the same functions, which is in contrast to Hox A4, B4, C4, D4 which express in the same place doing the same things. Hox A11 functions both in limb and female reproductive tract while Hox D paralog functions only in the mammalian limb. So you have both variations, one where all retained the same and in another set where you have divergence of the expression pattern and function.

(Refer Slide Time: 39:30)



So in summary what did we learn about this duplication and divergence? So these duplications allow redundant structures like in this case redundant gene sequences and that allows divergence because when you have one wild type copy taking care of the required function, then the other copies can be mutated and experimented with. So therefore you could have new roles coming up and that is how these gene

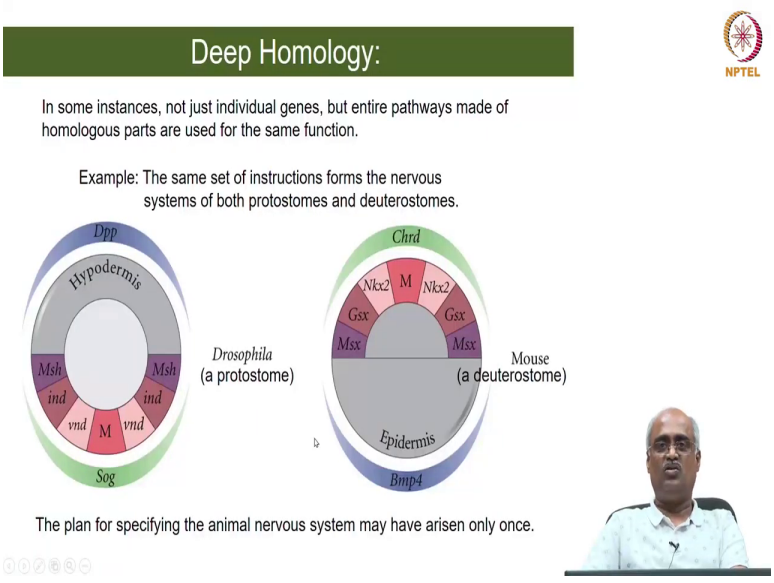
families arose. So you have Hox genes coming from one ancestral founding member generating a family but conserved to pattern the body and limb axis but differently in different organisms.

So let us say hypothetically, immediately after duplication they started functioning in the body plan by anterior-posterior axis and then the limb development and so on. And now the two groups duplicated and made multiple copies and therefore this body anterior-posterior body plan could be altered. So therefore the body plan of a mouse is not the same as drosophila. But still they have the conserved role in the body plan.

But the variation in them, like among their paralogs, creates the variations in the body plan that you see between the two groups of organisms. And similar is true with Distal-less. They are required to extend appendages and to pattern vertebrate skull and; variations in them in the paralogs again creates variations in the appendage and skull structure and similarly for muscle development myoD family members and so on.

The main point that we are realizing here is the difference between flies and humans do not arise by having different genes but by deciding which gene will be expressed, where and when. So it is not that humans have new genes but that humans have multigene families and that is how the difference between the two organisms arises.

(Refer Slide Time: 42:18)



So now let us look at what is deep homology. There are situations where multiple genes functioning in a pathway like signal transduction pathway, are all conserved and do the same function. This is called deep homology.

High degree of conservation is not synonymous with deep homology. It is not that the individual gene is conserved. It is actually a set of genes functioning in a particular process, for example limb development.

They are all conserved and do the same thing and that is what you call deep homology. An example is short gastrulation of drosophila gene and its TGF beta pathway ortholog Decapentaplegic(DPP) function. In fly, DPP determines the dorsal part.

If you recall our discussion on the dorsoventral axis formation the drosophila oocyte, we learned that the Dorsal protein enters the ventral nuclei at the syncytial stage and as a result Dorsal is expressed only in the ventral and Dorsal inhibits Decapentaplegic expression in the ventral side. So DPP is required for countering the activity of the ventralisation signals and as a result, it allows the central nervous system gene expression in a certain order on the dorsal side. So these are the genes that need to be expressed which are induced by the short gastrulation.

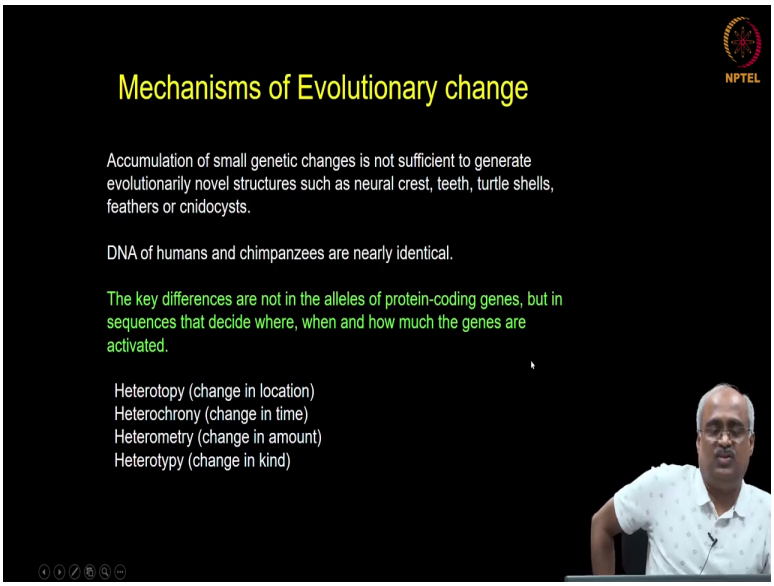
And they are restricted to the ventral side because of the dorsal expression of DPP. So this is how t all central nervous system genes are. The mid body point requires the highest expression of Sog and as it progressively decrease, these genes are activated. And these are countered by the DPP expression. A very similar but inverted structure is what happens in the mouse. So the central nervous system for example in our body as well as in the mouse is on the dorsal, like at our back.

But on the other hand, if you take invertebrates like *C. elegans* or Drosophila, it is on the ventral side of the body and that is because of this expression pattern of these genes. The main point in our current discussion is that central nervous system formation is done by Sog-DPP opposite gradient in drosophila and that is true here again, it is just that the dorsal ventral are inverted but chordin strong expression is what determines this expression pattern; and BMP4, the DPP homolog determines here this pattern.

So it is inverted here. The same genes are doing the same thing. Chordin-BMP4 or Sog-DPP setting up the expression pattern of genes involved in central nervous system specification is the same in both and this is an example of deep homology. And this also tells you that the central nervous system did not evolve multiple times independently because then you will see multiple ways of doing it.

Here what you are seeing, is the same DPP-Sog or its equivalent Chordin-BMP4 is the one doing the CNS specification. In the entire animal kingdom, how to make the central nervous system seem to have arisen only once and then we have diversification from that. And this is what we call as a deep homology.

(Refer Slide Time: 48:20)



Next, we are going to discuss what we call as the mechanisms of how all of this happens. So think of some of these evolutionarily novel structures that did not exist in the organisms, that existed before a certain group. For example, neural crust or teeth are structures that didn't exist in sponges. So how do you make these novel structures by having small changes in protein sequence or even in enhancer modules. How do you make these novel structures? And another major worry comes when you look at the sequences of human and chimpanzees.

So the genome is nearly identical, more than 99% identical and many of the alleles of proteins are very similar but the reality is, chimpanzee is not human. The key difference is not in the alleles of protein-coding genes, instead it is in the sequences that decide where, when and how much the genes are activated; will it come early on and will it be there in the limb bud or will it be in the imaginal disc.

And for how long will it be expressed there. So we saw one example earlier where the duration of FGF expression can determine whether a dog will have a long leg or a shorter leg, i.e how long the proliferation of those chondrocytes will happen before they enter into differentiation. Premature entry means a short leg where they differentiate into bone without producing a lot of precursor cells.

So changing the duration can also alter the structures. So, this is what really makes the difference between chimpanzees and humans, not the exact protein or the sequence itself. And how do you make these changes? They are classified into four groups: heterotopy, heterochrony, heterometry and heterotopy.

So heterotopy means different locations. For example, if a particular gene is expressed in one particular location so far in the course of evolution, now the gene expresses in a different part of the embryo leading to a new structure development. So that you call as heterotopy. And heterochrony is change in time. This change in time could be gene expression; this could also mean the rate of development for example quicker proliferation of cells versus slow proliferation or proliferation for longer time than what it happens in another organism.

So those kind of time dependent variations of developmental process or gene expression, we call as ‘crony’. ‘Tope’ is change in spatial expression and ‘crony’ is change in temporal expression. So in the next class, we will try to understand how these variations are made possible. And how it all ultimately explains how variations in development are generated.