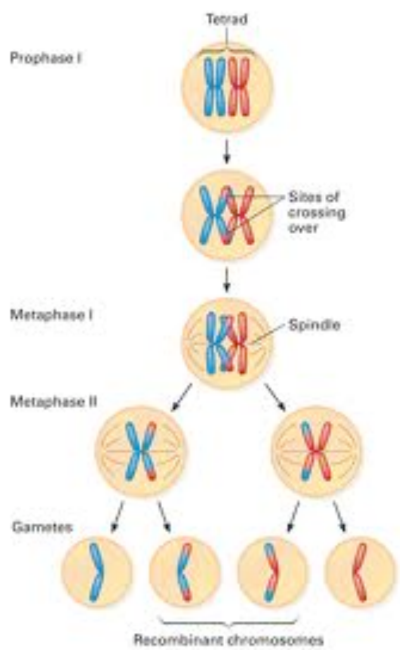
Gene duplication and divergence

Gene **duplication**, followed by **divergence** is a source of new genes in genomes. Divergence literally means to go in different directions. In this context, we mean that *the sequences of the gene copies are becoming different from each other because of the accumulation of mutations*. As you go through the worksheet, you will understand how this process works.

• **What is gene duplication? Write out your explanation of gene duplication:**

• **How do genes get duplicated?**

As you saw in class, genes can get duplicated when there is unequal crossing over of chromosomes. (There are also other known mechanisms, but we did not discuss these.) As you know, during meiosis, homologous chromosomes often exchange equivalent regions. This is called recombination or crossing over (see below). You could also review meiosis using a biology textbook.



Notice that in the figure above, the chromosomes are lined up so that they exchange equivalent regions (as in Figure A below). This means that neither chromosome involved in the crossover gains or loses any information, they simply swap the equivalent sequence from one to the other. Note also, that after meiosis is complete, there are four gametes, each of which has a single copy of each chromosome.

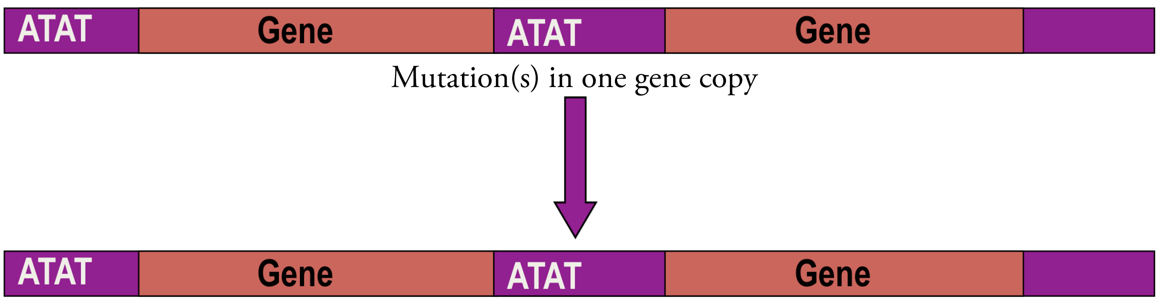
Figure A: shows chromosomes lined up normally, where any crossing over event will result in the exact exchange of equivalent regions between the two chromosomes. For simplicity, only one pair of chromosomes is shown here, rather than the entire tetrad. Notice that the same repeated sequence, represented by ATAT, is present on either side of the gene.

Figure B: shows the same chromosomes, but they are not lined up correctly. However, because they have the same repeated sequences on either side of the gene of interest, it is still possible for them to line up at a different repeat that is offset.
Draw out what happens when these two chromosomes swap sections by crossing over at the offset ATAT repeat shown below. *Each of the resulting chromosomes would go into a gamete (an egg or a sperm). Because only two chromosomes are shown here, you get only 2 gametes, instead of the four that would result from a tetrad in meiosis.*

• **Gene duplication can happen multiple times**, leading to multiple copies of the same gene on a single chromosome. This happens by a repetition of the process above. You may draw out for yourself how further duplication could result in more copies of the gene on the chromosome.

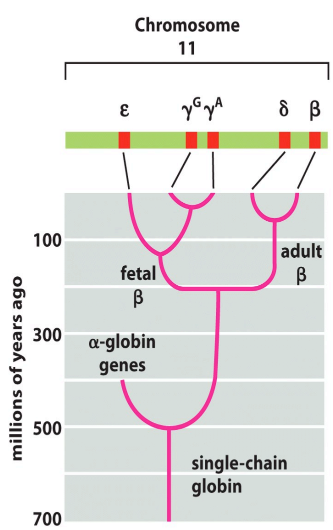
• **How do the copies diverge?**
Once you have more than one copy of a gene on a single chromosome, it is possible for the copies to gradually become different, as the gene copy (or copies) undergoes mutation. Mutations in these copies are not detrimental to the organism, as long as there is at least one "good" original copy. Because of this, the mutant copies can be passed on to the offspring. Remember, mutations are not always harmful, so sometimes, the mutation will result simply

in a slight variation in the properties of the encoded protein. This can turn out to be beneficial to the organism if the variant protein confers an advantage of some kind.
On the figure below with two gene copies, mark one copy a and one copy a'. To start with, both these are identical copies. But as one copy gets mutated, it may become sufficiently different that it encodes a slightly different protein. You would then recognize it as a new, though related, gene and call it b.



You should now have one original copy of the gene, a and one variant, b, *that give rise to two slightly different proteins*. The accumulation of differences (caused by mutation) in the genes is what we call "**divergence**".

The process above is how the original single globin gene is thought to have given rise to two genes that became the α and β globin genes. Subsequent duplication and divergence in these two genes gave rise to the variant forms of α and β that are now seen in humans. In the figure below, you can see that the original single globin gene that was present in early vertebrates was duplicated and gave rise, initially, to the α and β forms. The figure shows the variants that arose in the β gene, but not the α gene. One of the variant forms of the β is the fetal γ form that we discussed in class, that gives rise to the $\alpha2\gamma2$ hemoglobin that binds oxygen with greater affinity than the adult $\alpha2\beta2$ form.



The collection of different globin genes that arose from the common ancestral globin gene are called the globin **gene family**. Like human families, a gene family has related, but recognizably different members. Also like human families, gene families may have dysfunctional members. Such dysfunctional genes are gene copies that have mutations that render them non-functional. Because they look like other members of their gene family, but do not give rise to functional proteins, such non-functional gene family members are called **pseudogenes**.

There are thousands of gene families within the human genome. Like the globin gene family, each of these families is made up of related but slightly different members that arose from an ancestral form. One example is the histone gene family that gives rise to the various different histone proteins that you are familiar with.

How do we know about gene families and how they arise?

Comparison of the genome sequences of different organisms shows us how genomes have changed through time. As we noted, early vertebrates had a single globin gene. More recent descendants of these organisms have more than one globin gene, but these are clearly derived from the original gene, as we can see from the sequence. The figure above shows at what stage in the history of life these different variant forms appeared. Although the globin genes have undergone change over millions of years, there are also recent examples of more rapid changes in genomes that confirm that this mechanism is responsible for the emergence of gene families. For example, scientists have observed at least three independent gene duplication events in the acetylcholine esterase genes of certain mosquitoes over the past 50 years. These duplications have been followed by selection for variant forms of these genes that make the mosquitoes resistant to insecticides. Thus, in a relatively short period of time, mosquitoes have acquired extra copies of the ace-1 gene, some of which have mutations that enable the mosquitoes to survive better in the presence of certain insecticides.