



## Evolution of Apoptosis Proteins Domain Rearrangement to Create New Protein Architectures

### Introduction

The human genome has about 30,000 different genes. This number is similar to that found in other organisms. However, humans and other vertebrates have a larger diversity of proteins due to the larger number of ways in which different protein domains are used together within proteins; i.e. a larger number of protein architectures.

**Note: Please see the activity “Vertebrates, flies and worms: Protein domain usage compared” for definitions and detailed descriptions of protein domains and protein architectures.**

In this activity we will look at the evolution of apoptotic proteins by comparing the apoptotic proteins from different groups of organisms. This activity will demonstrate how a large number of diverse apoptotic proteins evolved from a small number of ancient domains. This ability to create large numbers of diverse proteins from a small number of domains is characteristic of the vertebrate genome, and explains in part why genome size and the number of genes is not representative of the complexity of the organism.

This activity also shows an example of the type of work that can be done in the field of genomics now that the genomes of different organisms have been sequenced. Research like this answers questions about individual proteins, protein families and bigger picture questions such as evolution.

This activity is based on the work presented in the following reference

Aravind, L. Dixit, VM, Koonin, EV. 2001. Apoptotic Molecular Machinery: Vastly Increased Complexity in Vertebrates Revealed by Genome Comparisons. *Science* (291):1279-1284

### Apoptosis

Apoptosis is also known as programmed cell death. In this process, a cell receives a signal to die from surrounding cells, the environment, or sometimes from the lack of a specific signal to continue living. Once the cell receives the signal to die, it commits suicide in an organized fashion that is not toxic to surrounding cells.

Apoptosis plays a critical role in many processes.

- Development (embryonic in particular, but also after birth) – cells that have served a function in development and are no longer needed are removed by apoptosis. Also, cells that do not form proper connections with surrounding cells apoptose (particularly relevant to nervous system development)
- Immune system (we eliminate immune cells that respond to antigens present on our own cells so that we don't mount an immune response against ourselves).
- Elimination of cancerous or infected cells
- Aging
- Stress response – cells that have been damaged by some stress are signaled to die by apoptosis so that they do not become necrotic and release toxic components onto surrounding cells and tissues.

## Proteins Domains Involved in Apoptosis

There are several different types of protein domains involved in apoptosis.

- Receptor domains – bind to a protein that is the signal for the cell to begin the apoptotic process. The receptor domain is usually but not always extracellular.
- Adaptor domains – transmit the signal from the receptor domains to the enzyme domains. The adaptor domains may transmit the signal under some circumstances but not others, allowing more complex regulation of the process. The adaptor domains are usually but not always intracellular.
- Enzyme domains – have an enzymatic function, usually as a protease or kinase. These functions usually begin the apoptotic process.

Proteins involved in apoptosis contain one or more of these types of domains. Some proteins contain all three of the above types of domains that are required to initiate apoptosis. More frequently the process is carried out by two or more proteins, whose interaction is required to initiate the apoptotic process. Complex webs of interaction are possible in higher eukaryotes with receptor proteins interacting with a variety of different adaptor domain containing proteins under a variety of circumstances. This complicated web of activation allows for finer control of the process.

The activity today will include the domains listed below.

### Example Receptor Domain



TIR - Toll – Interleukin 1 Receptor Domain

### Example Adaptor Domains



D - Death Domain



C - CARD (Capase Recruitment Domain)



P - Pyrin

### Example Enzyme Domains



Capase



Kinase



Ap-ATPase



Ap-GTPase

### Other Domains



p53F – p53 fold domain



ank – ankyrin repeats



Irr - leucine rich repeat



tpr – tetratricopeptide repeat



Ig – Immunoglobulin domain



Tig – Transcription factor immunoglobulin domain



Zu5 – zona pellicuda Unc-5 domain

## The Tree of Life

Scientists now use a 3 domain system, instead of the older 5 kingdom system, to classify life on earth. In this system life is divided into three domains, bacteria, archaea, and eukaryotes. The root of the big tree, or the common ancestor for all life on earth today is thought to be an ancient bacterium.

### Activity

The beads have been arranged on strings representing the types of domain arrangements found in various groups of organisms. The different types of proteins found in different groups of organisms have been placed in bags. Your goal is to identify which bag of proteins corresponds to which spot on the evolutionary tree. Remember that evolutionarily more simple organisms (and protein combinations) are found towards the bottom of the tree, and complexity increases over time (towards the top of the tree).

Begin by examining the proteins in bags 1 – 6. **Do not look at the bacteria or plant bags at this time.**

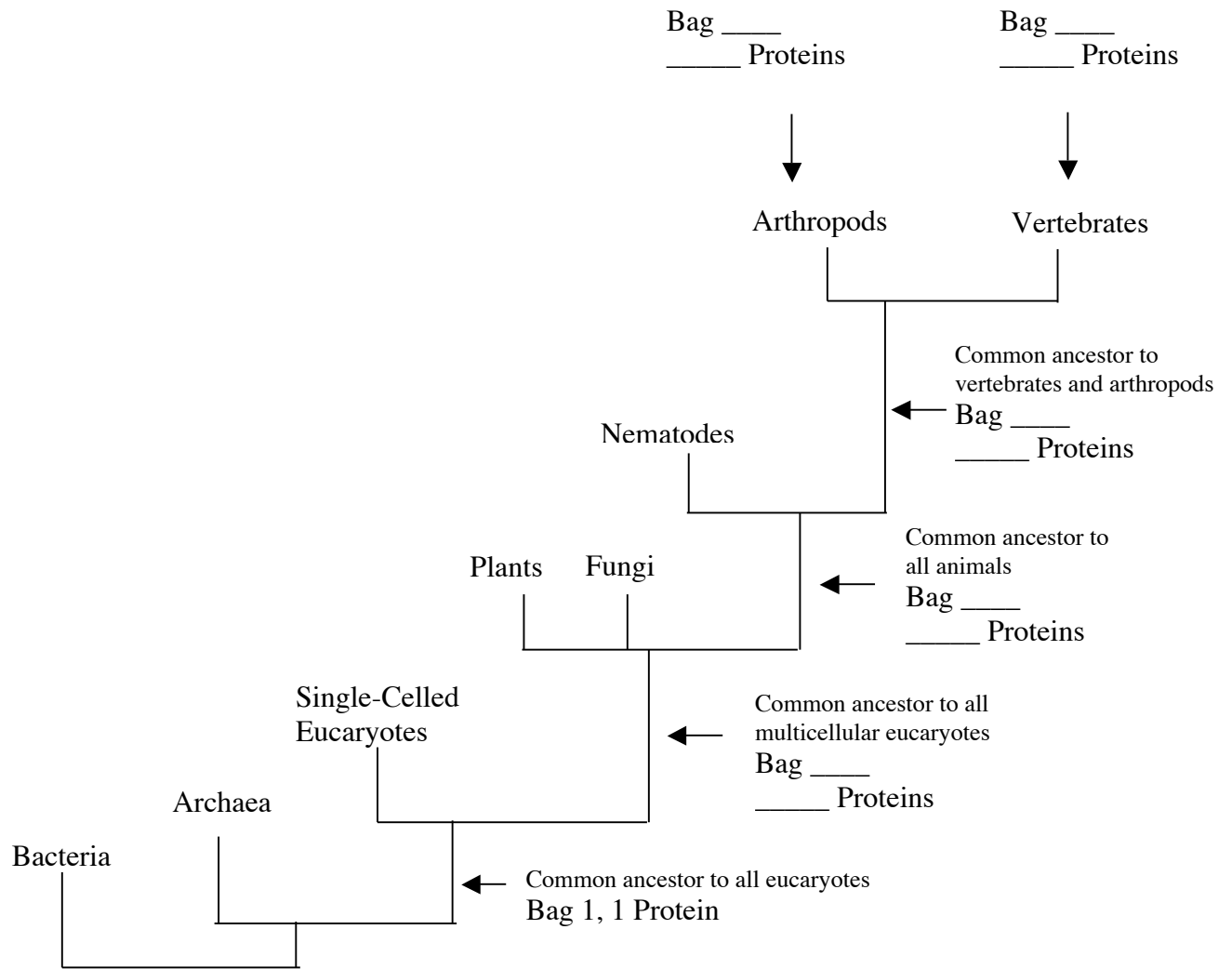
Remove the proteins from the bags (while keeping track of which came from which bag). Arrange the proteins so that you can easily compare the proteins from the different bags. Look at the tree of life on the next page. The contents of bag 1 are found in all eukaryotes.

At each spot on the evolutionary tree where you see the following:

Bag \_\_\_\_\_  
\_\_\_\_\_ Proteins

fill in the bag number (1 – 6) that corresponds to that point on the evolutionary tree and the number of proteins found in that bag.

Then on the following page, sketch the different proteins that are new to each category of proteins.



Sketch the different proteins unique to each group of organisms

All eucaryotes

Higher eucaryotes (plants, fungi, animals)

Animals

Arthropods and Vertebrates

Arthropods

Vertebrates

Answer the following questions

1. How many different proteins are found with a C domain

in vertebrates?

in all animals?

One of the characteristics of vertebrates is a large expansion in the number of apoptotic proteins involving C domains. The NACHT NTPase domain (not covered in this activity) is also very largely expanded in its use in vertebrates.

2. Where do the following domains first appear

Ap-GTPase?

D?

C?

Tir?

P?

### **Summary**

Throughout evolution two different processes occurred

- New domains evolved
- Existing domains were reshuffled to create a larger number of proteins with more highly refined functions.
- Although not shown in this activity, domains are also occasionally lost

Sequencing the genome of several different organisms has allowed us to answer questions about the evolution of organisms, proteins and protein domains.

Work in this area is just beginning. There are many questions still to be answered.

Advanced extension questions

Look at the proteins in the bag labeled “Bacteria” These proteins are found in some (Actinomycetes and Cyanobacteria) but not all bacteria. Also look at the proteins in the bag labeled “Plants.”

First, fill out the left hand column of the table below with the names of the different protein domains found in bacteria. Then, in the right hand column, not which groups of eucaryotes contain each of these domains. Note – one domain has been included in the table as an example.

<b>Bacterial Protein Domain</b>	<b>Eucaryotes in which this domain is found</b>
Kinase	Arthropods and vertebrates

What does the lack of these domains in the common ancestor for all eucaryotes suggest about the origin of these domains in more advanced eucaryotes or bacteria (list all possibilities you can think of). While considering this, take into account that these domains are found in some but not all bacteria.

The domains present in some bacteria and the animal lineage but not their common ancestors are thought to have transferred horizontally from bacteria to the animal lineage. These domains include Tir and kinase. Similarly, Ap-ATPase is thought to have been transferred horizontally from bacteria to plants.