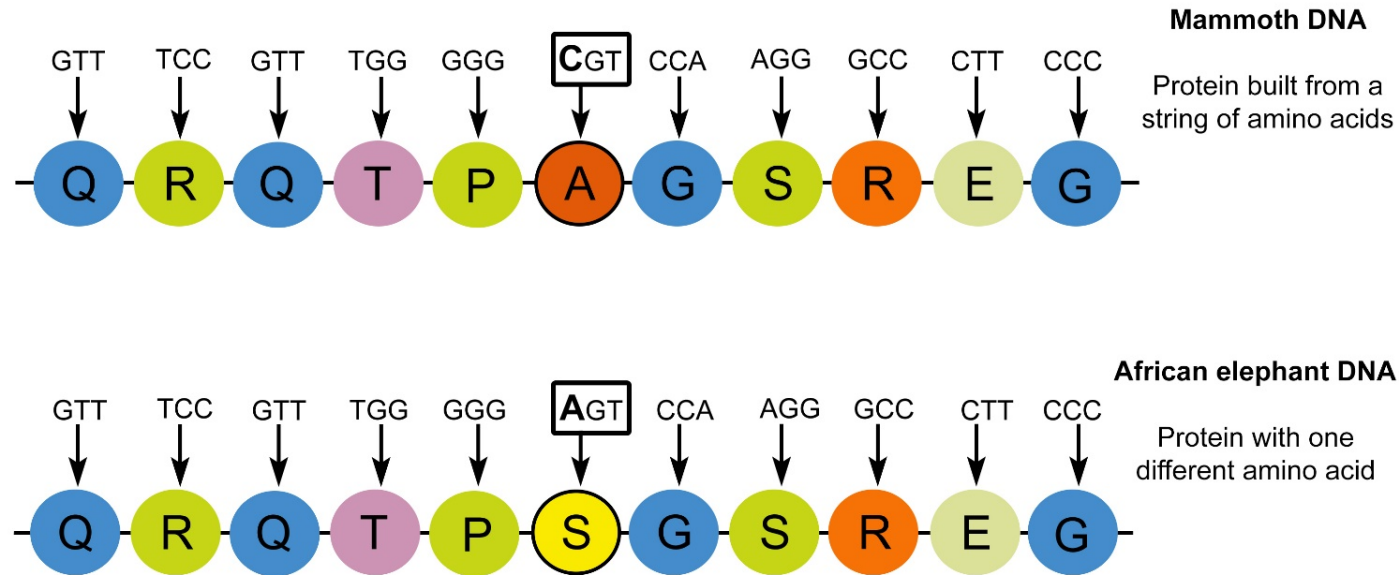


# Protein Sequences Provide Evidence for Evolution

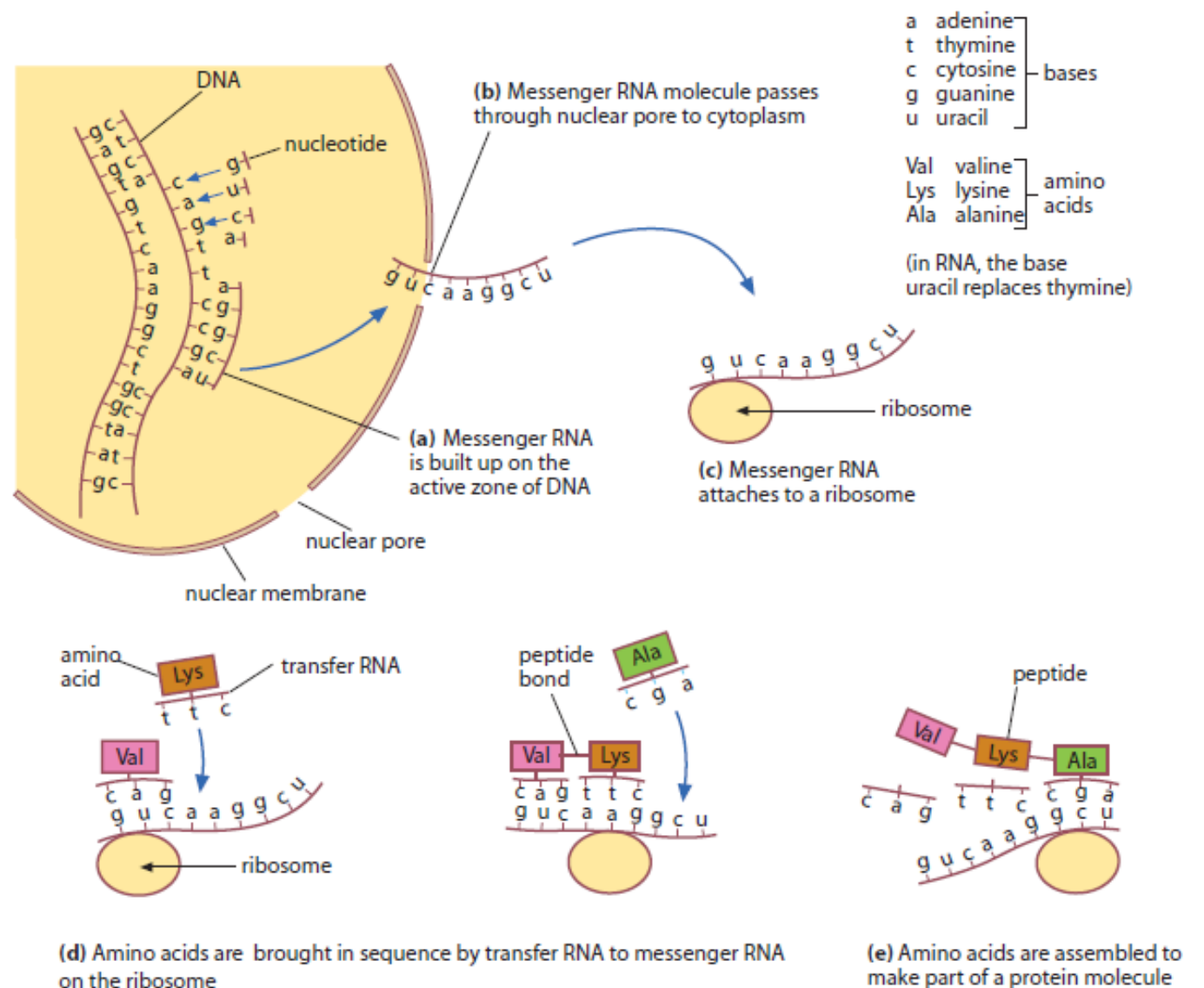


Protein Sequences  
Bioinformatics  
Comparative Genomics

Date:	Human Biology Year 12 ATAR
<p><b>Do Now</b></p> <p>Get out equipment</p> <p><b>Lesson Agenda</b></p> <p>1: Do Now</p> <p>2: DNA evidence for evolution – protein sequences</p> <p>3: Bioinformatics</p> <p>3: Lesson summary</p> <p><b>Suggested Study</b></p> <ul style="list-style-type: none"> <li>Read through today’s notes and textbook section</li> </ul> <p><b>NEXT LESSON</b></p> <p>Fossil evidence for evolution</p>	<p><b>Learning Aims</b></p> <ul style="list-style-type: none"> <li>Review protein synthesis and protein structure</li> <li>Define the term “ubiquitous protein”</li> <li>Describe Cytochrome C</li> <li>Explain why Cytochrome C has changed so little over evolutionary time</li> <li>Describe how Cytochrome C is compared and how similarity is measured to determine common ancestry</li> <li>Describe how other proteins (eg haemoglobin) can be used to provide evidence of common ancestry</li> <li>Define bioinformatics and describe its function</li> <li>Discuss annotation as an aspect of bioinformatics</li> <li>Explore some ways that evolutionary relationships between living things can be represented.</li> </ul>
	<p><b>Key Vocabulary</b></p> <p>Ubiquitous</p> <p>Cytochrome C</p> <p>Amino Acid</p> <p>Protein</p>

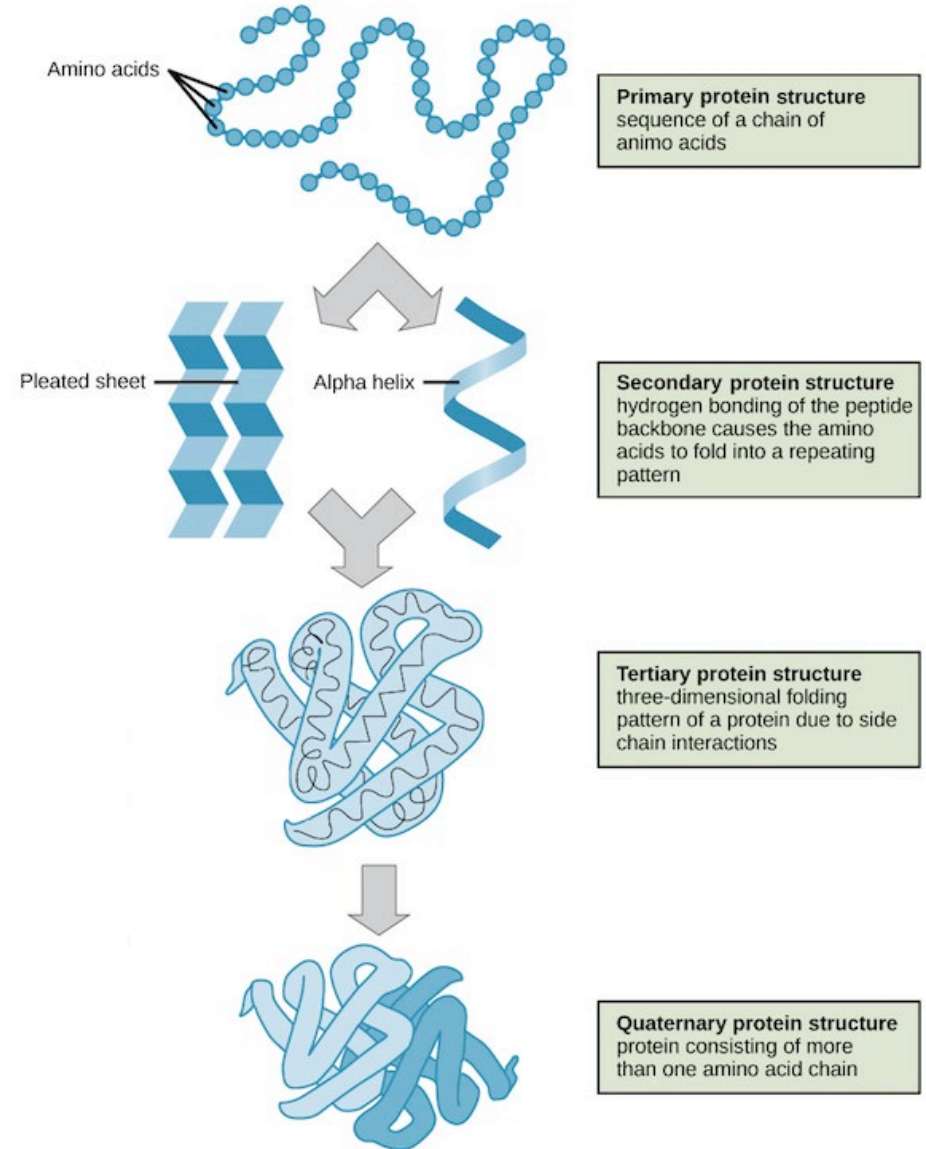
# Genes in DNA code for protein production

- Genes are long sequences of DNA base triplets that code for amino acids to be assembled (see picture) into long strands.
- These strands fold, and bind, and form specifically structured proteins.
- Each human gene codes for between 1 and 3 proteins.
- Proteins perform a variety of functions:
  - Structural material for cells
  - Control of chemical reactions (enzymes are proteins)
  - Control of some aspects of membrane transport (as protein channels and/or receptors)
- Genes can be switched on or off depending on the cell's purpose (remember cell differentiation?)



# Protein Sequences Provide Evidence for Evolution

- Some protein sequences can be compared to assess relatedness and common ancestry among living things, providing evidence for evolution
- Proteins
  - Long chains of amino acids in precise sequences
  - Amino acids are coded for by base triplets
  - There are 20 amino acids that form different proteins by being sequenced in different combinations
- We can compare specific proteins between species to assess degree of relatedness.
  - Can look at degree of similarity
  - Same species almost identical
  - Degree of difference allows an estimate of amount of evolution since two organisms diverged from common ancestor



# Protein Sequences Provide Evidence for Evolution

- Amino Acids and Protein Sequences
  - Amino acids in proteins are usually represented by 3 letter code
  - Eg ALA = Alanine
  - Sometimes given a single letter code as easier to read and compare – see table
- Ubiquitous Proteins
  - Ubiquitous means “everywhere”.
  - These are proteins that occur in all living things from bacteria to humans
  - They perform the same function in all organisms
  - One example is Cytochrome C – essential for cellular respiration

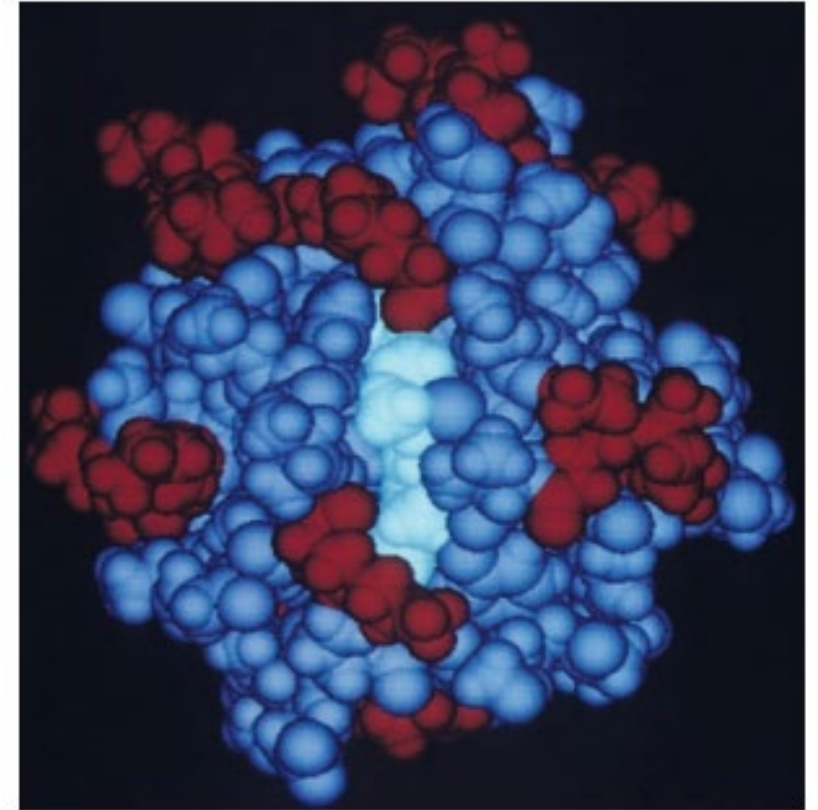
**Table 15.3** The three-letter and single-letter codes used for amino acids

Amino acid	Three-letter code	Single-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I

Amino acid	Three-letter code	Single-letter code
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

# Cytochrome C

- Is an example of a ubiquitous protein
  - Human Cytochrome C has 104 amino acids
  - 37 of the amino acids in Cytochrome C are found in every molecule sequenced, from bacteria to humans
  - Performs essential step in cellular energy production
  - Has changed little over millions of years of evolution (why?)
- Strong evidence that the genes for Cytochrome C production originated from a primitive microbe ancestor, more than 2000 million years ago



**Figure 15.3** A model of cytochrome C

# Comparing Cytochrome C

- Comparing Cytochrome C
  - Aligned so maximum number of positions with same aa can be determined.
  - More similarity = more recent common ancestor
  - Eg: Cytochrome C of chimpanzees, gorillas and humans is identical
    - Recent divergence from common ancestor
  - Cytochrome C of Rhesus monkeys differs from humans by one aa.
    - Still recent divergence from common ancestor, but not as recent as for great apes.
- Other proteins provide similar evidence
  - Eg: haemoglobin is not ubiquitous but is found in all mammals. In humans and chimps, the aa sequence for haemoglobin is identical. It differs by one aa from gorillas, and from gibbons by 2.

*Learning Aim: Describe how Cytochrome C is compared and how similarity is measured to determine common ancestry.*

*Learning Aim: Describe how haemoglobin can provide similar evidence.*



**Table 15.4** Differences in amino acids in cytochrome C between humans and other species

Species compared with humans	Number of differences from human cytochrome C
Chimpanzee	0
Gorilla	0
Rhesus monkey	1
Patas monkey	1
Rabbit	9
Cow	10
Pigeon	12
Bullfrog	18
Tuna	21
Fruit fly	24
Yeast	44

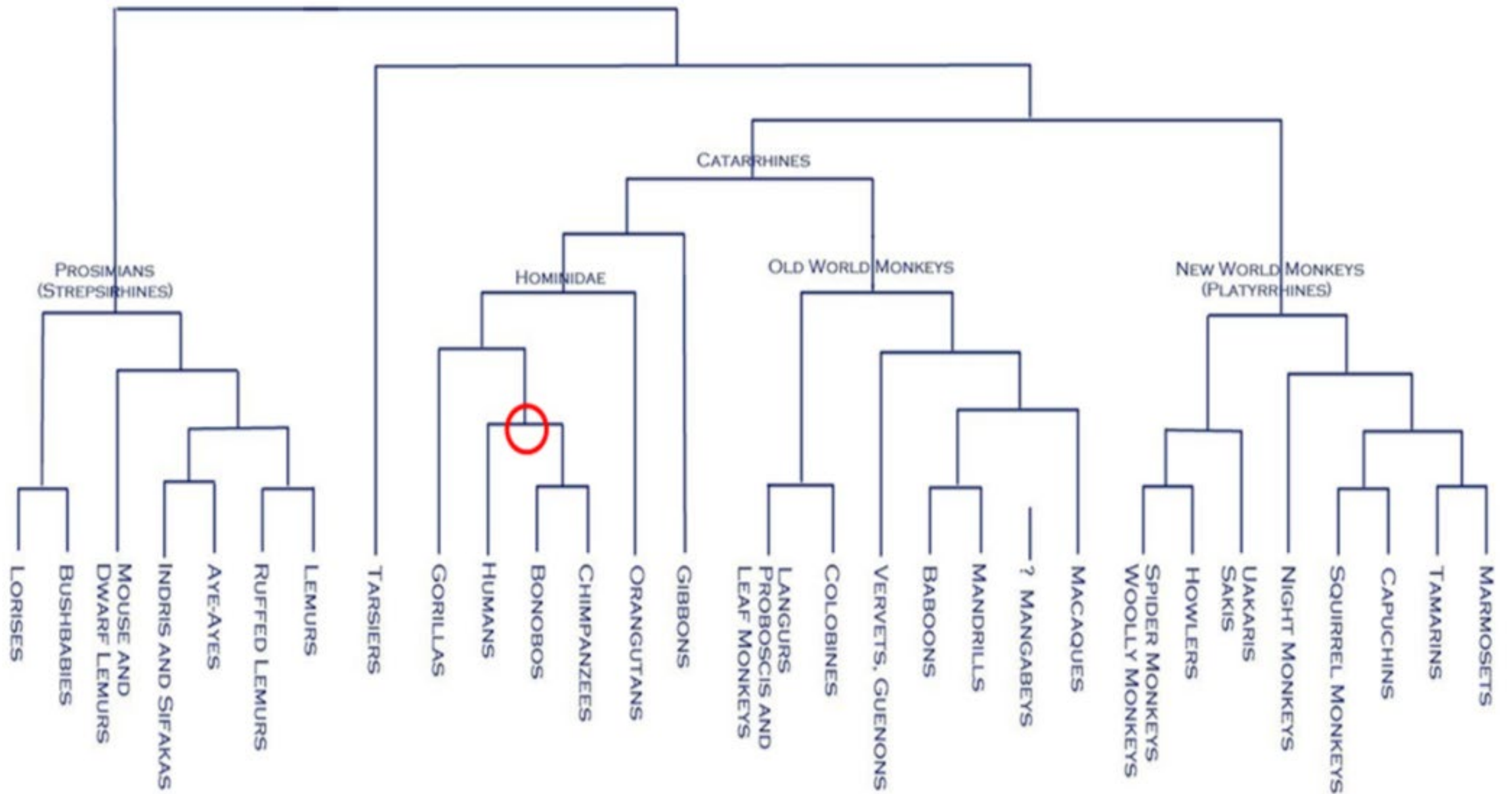


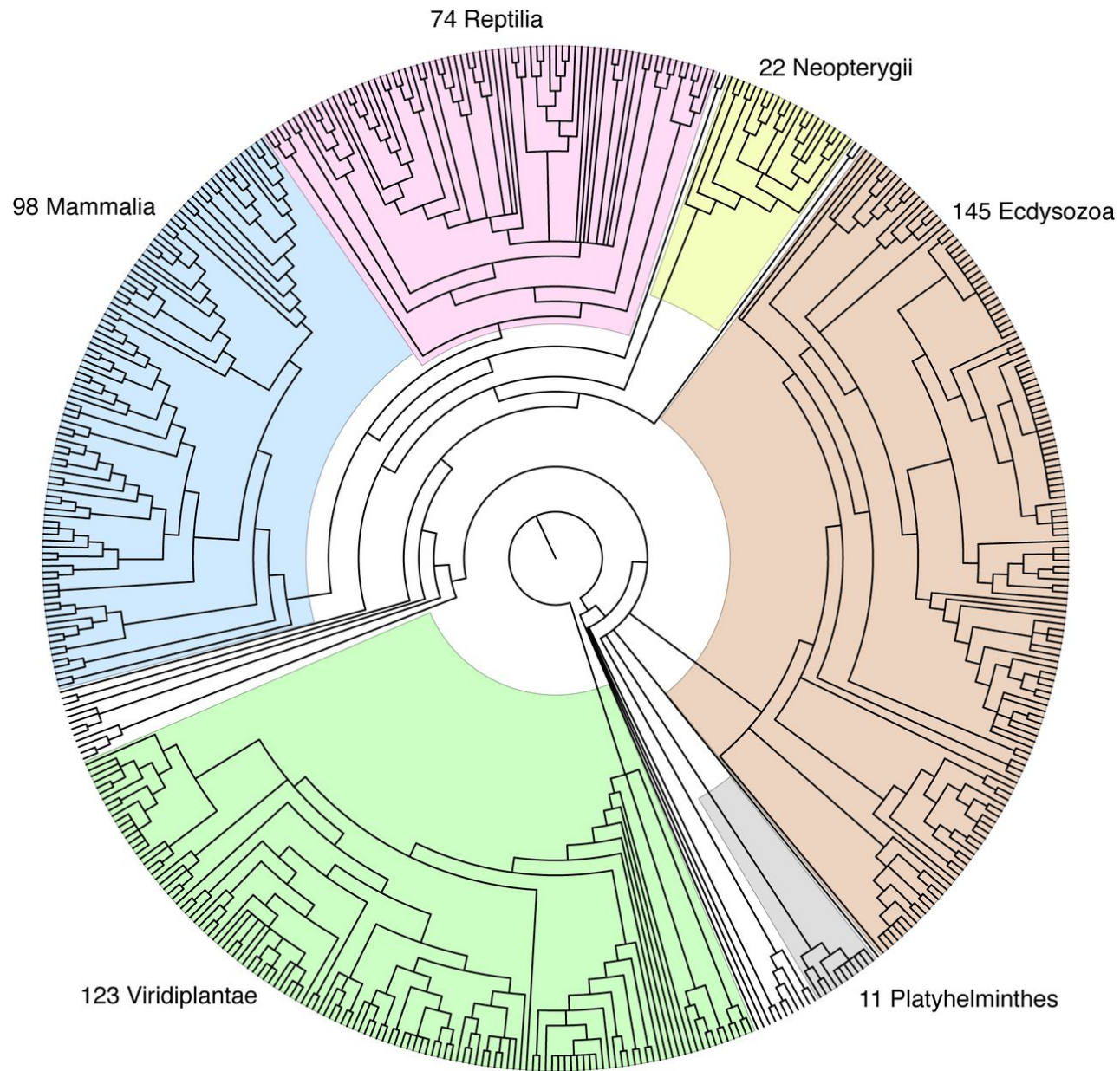
# Bioinformatics

- Bioinformatics
  - Combines biological science, computer science, stats and applied maths to understand and model biological processes
  - Use of computers to describe and analyse living things at a molecular level
  - Has assisted evolutionary biologists to trace evolution by measuring changes in DNA
- Recent developments allow comparison of entire genomes
  - “annotation” – identification of genes and other biological features – an aspect of bioinformatics
  - Identifies base sequences of individual genes within the genome.
  - Does this by identifying Start and Stop codons where sequences begin and end.
  - Does not discover what these genes do, just labels them for later analysis.
  - Done by computer once genome has been sequenced – most genomes are too large to be annotated by hand

# Comparative Genomics Summary

- Genome sequences of different species are compared.
  - Eg ERVs, mtDNA, known gene locations and sequences, ubiquitous proteins etc.
  - Identifies regions of similarity and difference.
  - Can ID genes that are preserved, and those that provide individual characteristics and differences
  - Has observed a high degree of similarity between closely related organisms, and diversity in different evolutionary lineages.
  - Has resulted in reclassification of some species, as the evidence shows a slightly different relationship than was originally hypothesised.





*Learning Aim: Explore some ways that evolutionary relationships between living things can be represented*

Tree of Life Explorer Website:

<https://www.onezoom.org/>