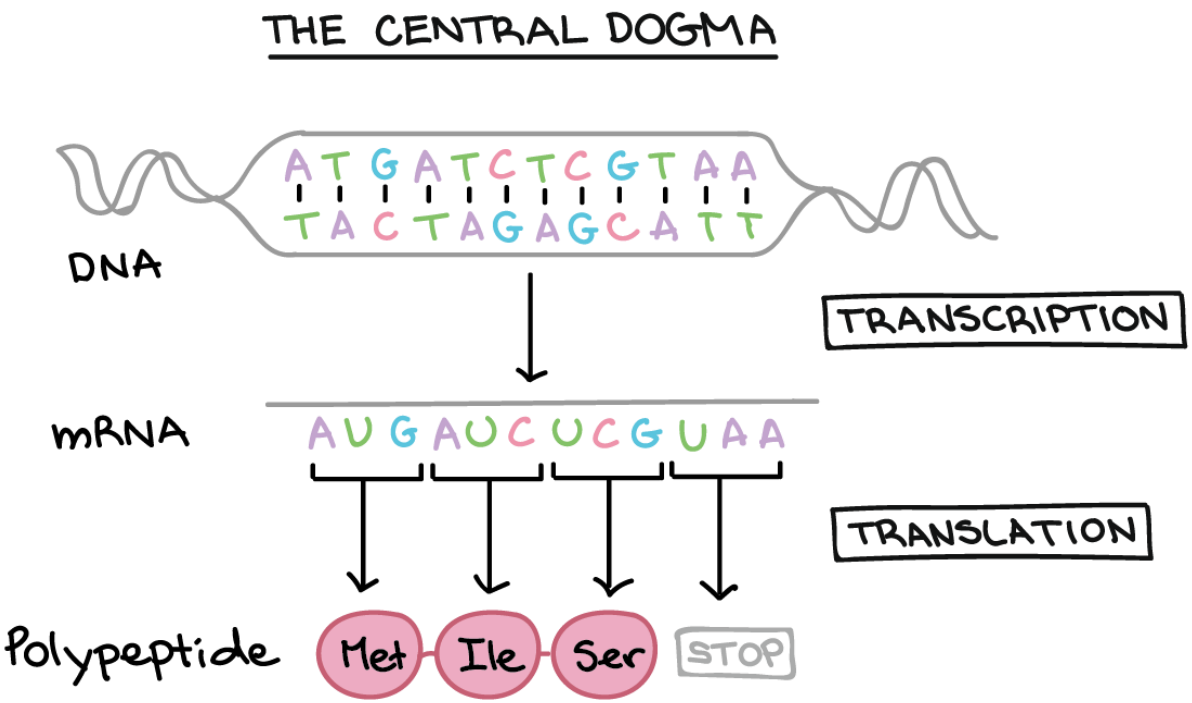
Forensics & Bioinformatics I

Forensics I

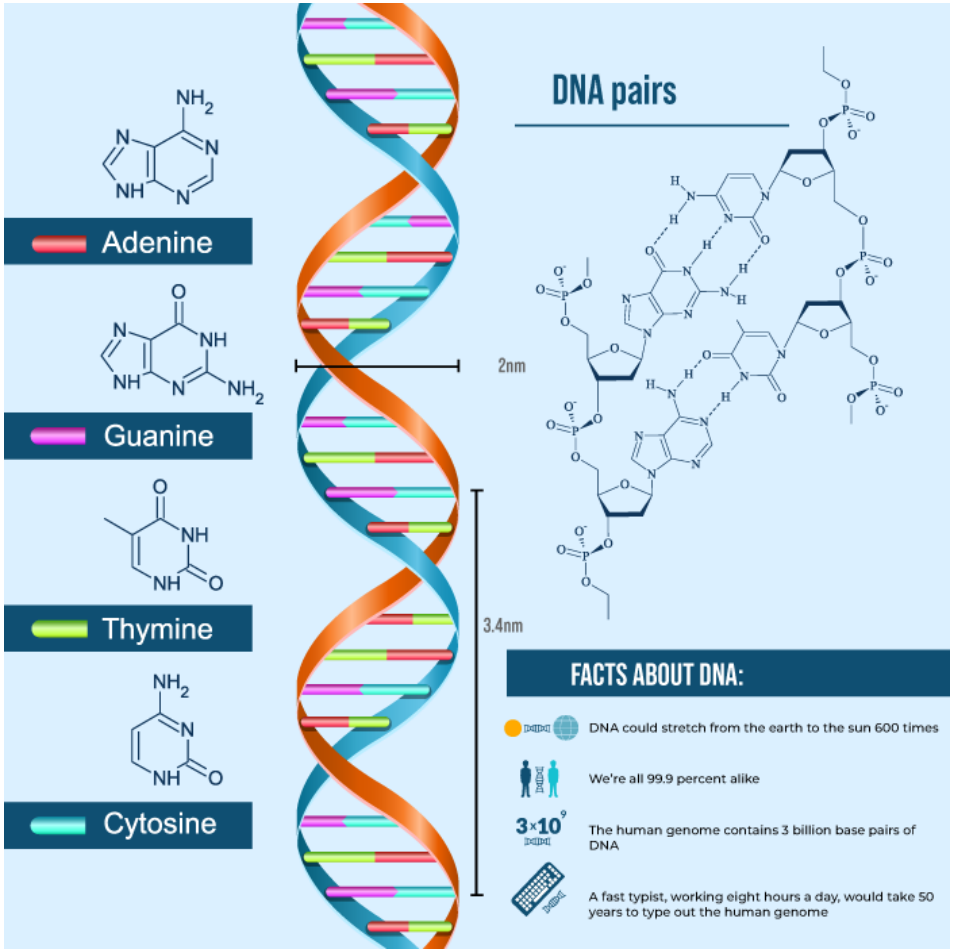
DNA

- **DNA- contains instructions for the development of an organism**
- **Central Dogma: genetic information flows only in one direction, from DNA, to RNA, to protein, or RNA directly to protein.**
- **Double helix structure**
- **Composed of nucleotides**



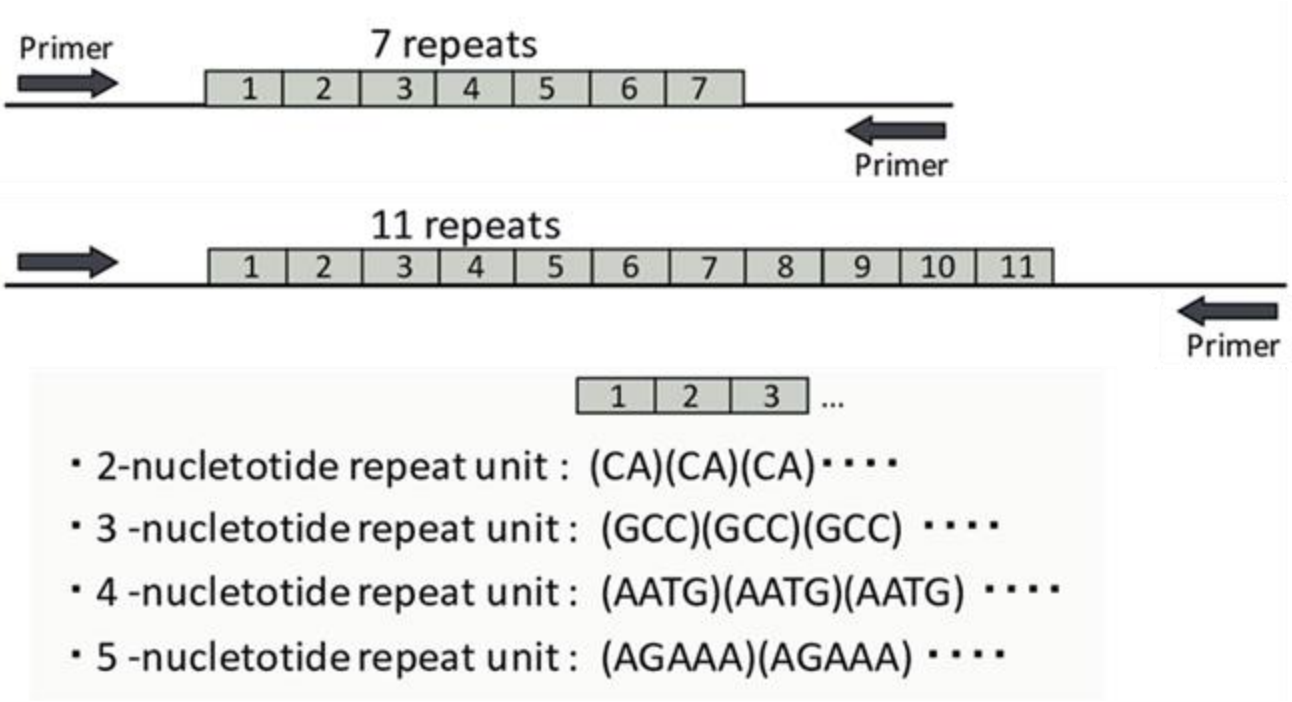
Genetic Code: DNA Structure

- 4 building blocks (nucleotides) of DNA
- Adenine (A), Thymine (T), Cytosine (C), Guanine (G)
- Uracil (U) found in RNA, binds to adenine (takes the place of thymine)
- The two strands in a DNA molecule are:
- **Complementary**
- A pairs with T
- C pairs with G
- **Antiparallel**
- Both strands run in opposite directions to each other



Short-tandem repeats (STRs)

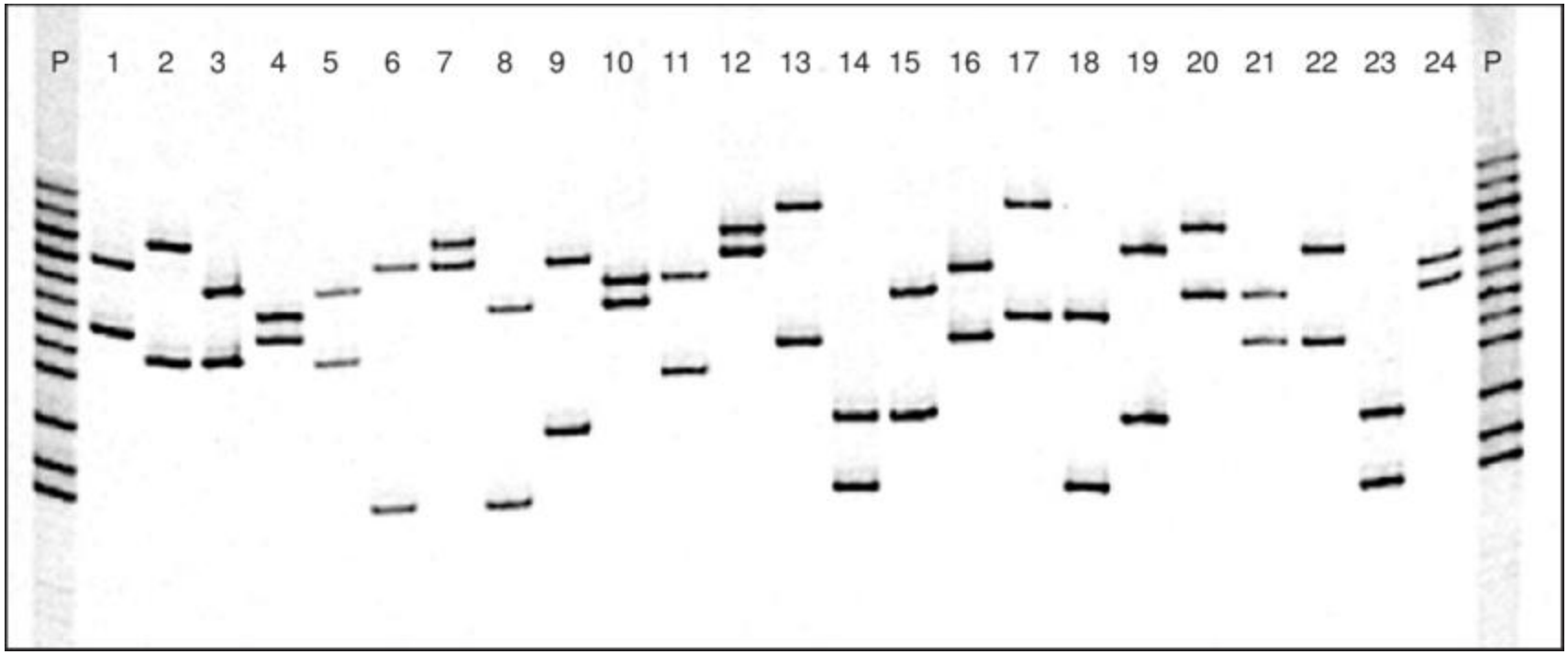
- The **coding region** of the genome encodes proteins that determine physical traits.
- On the other hand, 98-99% of human DNA is considered **non-coding**.
- This region of DNA has unique repeating patterns, known as **short-tandem repeats (STRs)**
  - can be used to differentiate one person from another.
  - present at specific locations, may be repeated , on everyone's chromosomes (loci)
  - 5-20% of people share the same DNA profile at any one STR site.



To match DNA profiles with confidence, scientists analyze many different STR sites concurrently.

Creating a DNA profile

- Multiple STR regions from each DNA sample are amplified by **Polymerase Chain Reaction**.
- These STR regions are then separated by gel electrophoresis based upon STR size.
- DNA-binding fluorescent dyes allow visualization of a banding pattern.



Recognizing and Counting STRs

DNA Sequence		
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Maternal	Person 1
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Paternal	
ACTGACTACCGACCGACCGACCGACCGTTAA	Maternal	Person 2
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Paternal	
ACTGACTACCGACCGACCGACCGACCGACCGTTAA	Maternal	Person 3
ACTGACTACCGACCGACCGACCGTTAA	Paternal	

Chromosome #10 contains STRs that are 4 nucleotides in length

Recognizing and Counting STRs

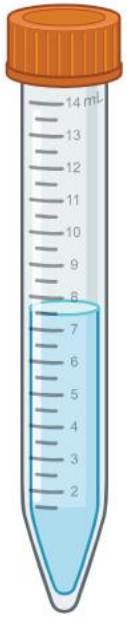
DNA Sequence		
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Maternal	Person 1
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Paternal	
ACTGACTACCGACCGACCGACCGACCGTTAA	Maternal	Person 2
ACTGACTACCGACCGACCGACCGACCGACCGACCGACCGTTAA	Paternal	
ACTGACTACCGACCGACCGACCGACCGACCGACCGTTAA	Maternal	Person 3
ACTGACTACCGACCGACCGACCGTTAA	Paternal	

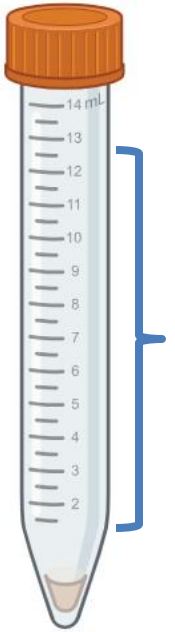
Chromosome #10 contains STRs that are 4 nucleotides in length

Visual Protocol

Before STR regions from each DNA sample are amplified by **PCR** we need to extract DNA.

Saline+
Mouth Rinse





Discard the
Supernatant
(liquid)

What types of cells exist in your mouth?

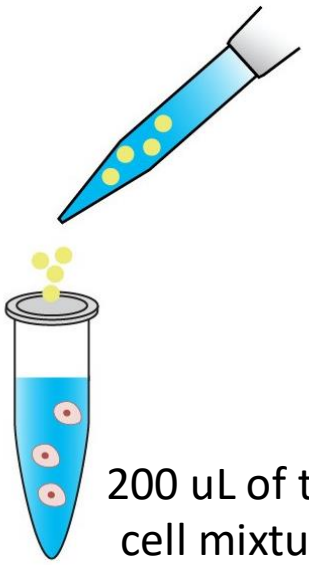
Addition of Chelex resin chelates (absorbs) ions that inhibit the function of Taq polymerase (PCR).

Thermocycler

Fine temperature control:
ability to hold a precisely set temperature with little fluctuation.

99°C for 10 minutes. This will lyse the cells, releasing your DNA.

1 **Add Chelex 100** suspension to sample and mix

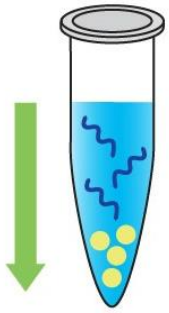


200 uL of the cell mixture

2 **Heat sample**

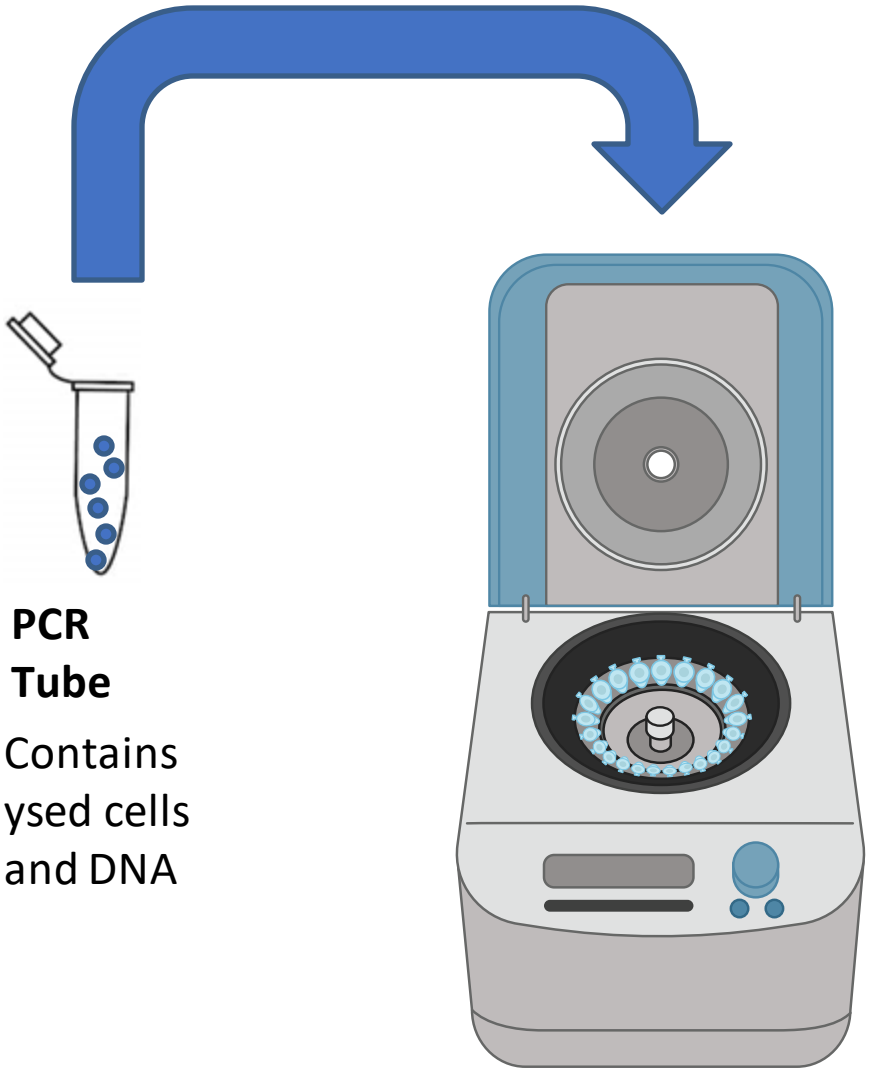
99°C for 10 minutes.

3 **Separate Chelex 100 beads** from sample by centrifugation, settling or filtration



The master mix contains:

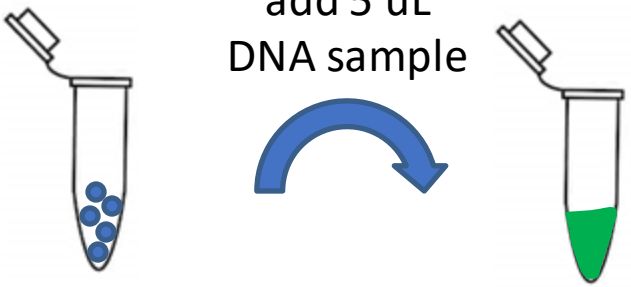
- Buffer
- Loading Dye
- Deoxynucleotides
- Three Pairs Of Primers
- Taq Polymerase
- Water



PCR Tube

Contains lysed cells and DNA

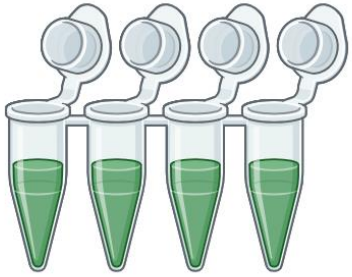
Centrifuge



PCR Tube

25 uL PCR master mix

Thermocycler to amplify the STRs in your DNA



DNA + master mix

Thermocycler

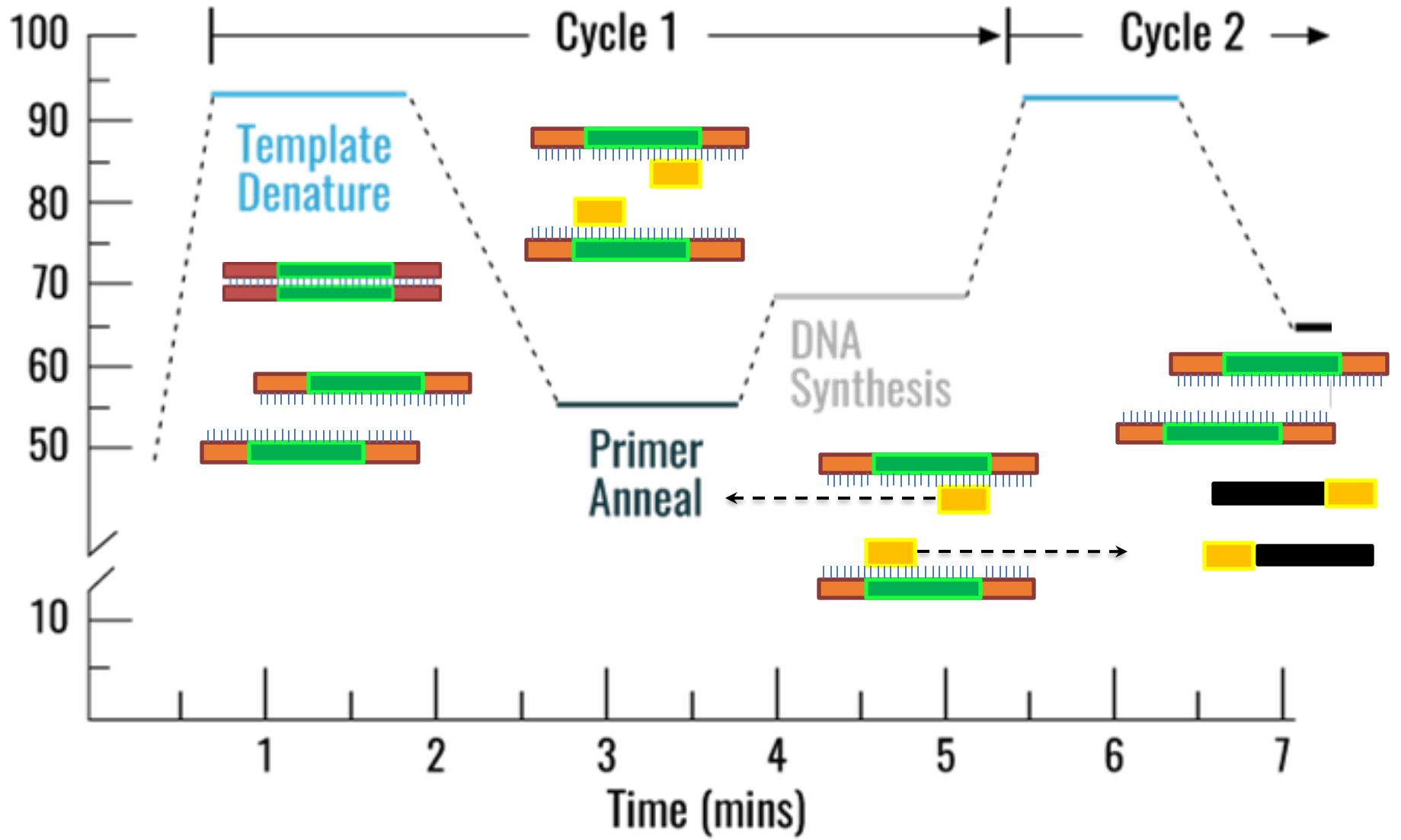
Fine temperature control:

ability to hold a precisely set temperature with little fluctuation.

PCR Program

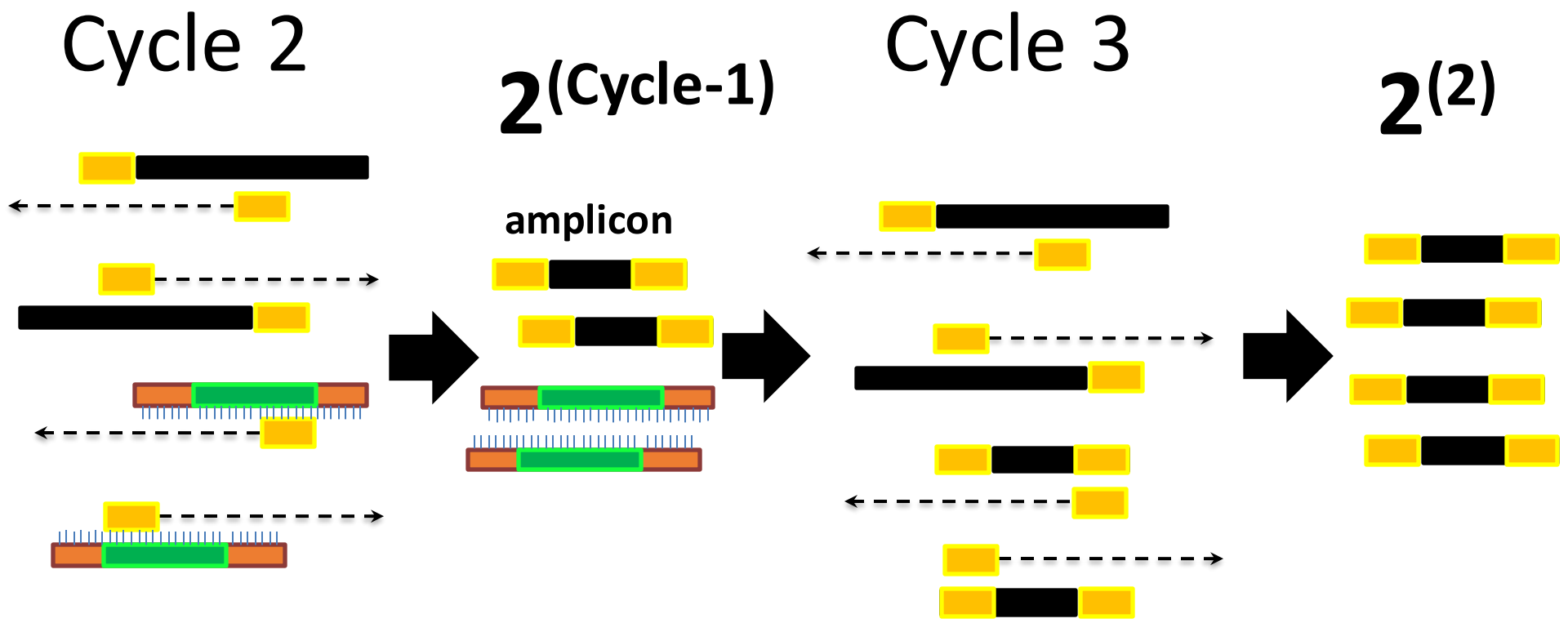
PCR Reaction Components

- DNA template
- Primers
  - short pieces of DNA specific to the sense or antisense strands
  - bind via hydrogen bonds
- DNA polymerase.
  - Taq polymerase is a thermostable enzyme isolated from the bacterium Thermus aquaticus that makes its home in hot springs. Taq polymerase can withstand temperatures greater than 90°C.
- Deoxyribose nucleoside triphosphates or dNTPs
  - will comprise the base pairs in the growing strands
- A reaction buffer, which maintains pH and contains important ions like manganese, magnesium and potassium
  - stabilizes the reaction and provides important cofactors to the polymerase enzyme.
- A solvent PCR grade water
  - free of ions that can inhibit the reaction.

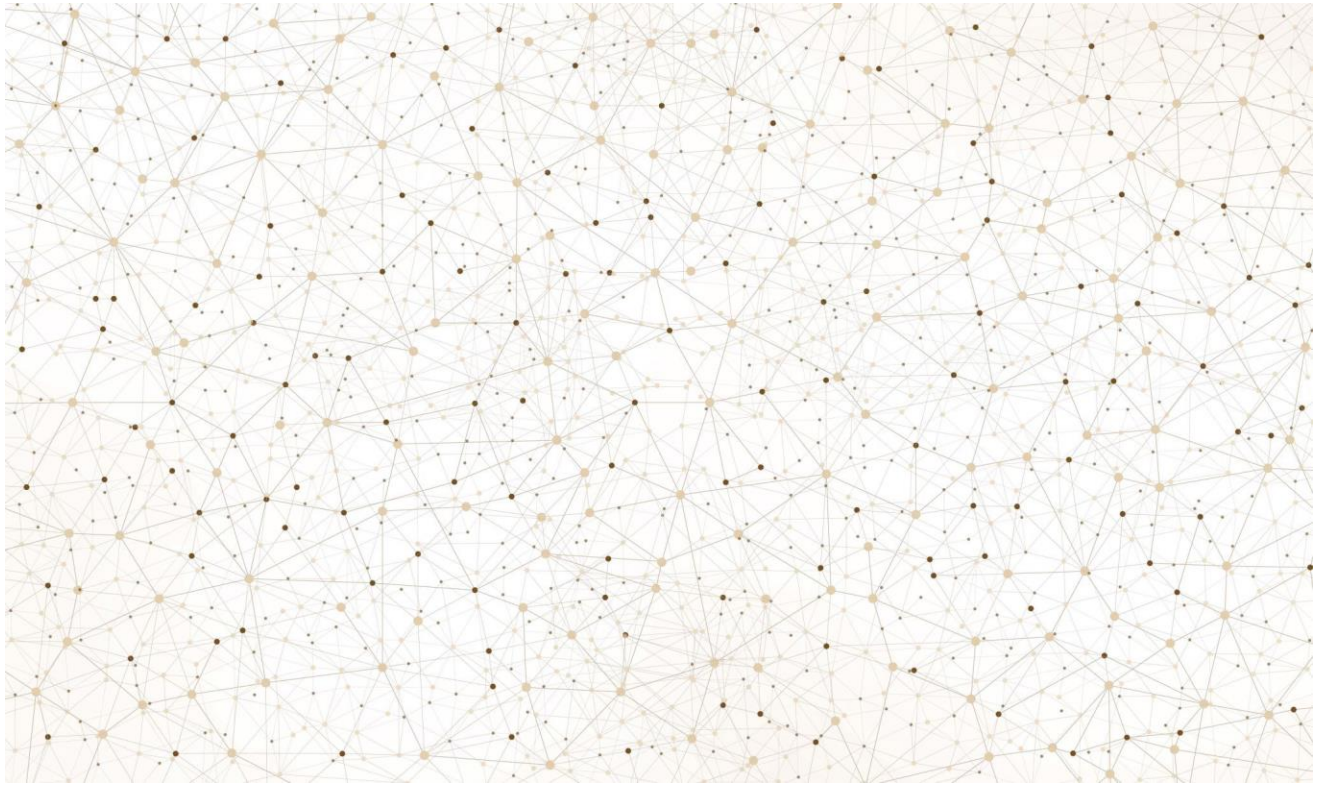


The denaturation of DNA results in the hydrogen bonds between complementary base pairs being broken yielding only single stranded DNAs.	The annealing of short pieces of DNA called primer. Specific to the sense or antisense strands bind via hydrogen bonds.	Elongation using polymerase, begins by adding free dNTPs to the ends of the primer one at a time in the 5’ to 3’ direction to make double stranded DNA.

- In the next cycle primers will bind to single stranded DNA formed via previous extension.
- The short fragment you are trying to amplify, the amplicon, will ultimately form when the polymerase extends from the forward primer on a strand that was generated by amplification from the reverse primer or vice versa.
- Once generated, the amount of amplicon will increase exponentially $(2^{n-1})^*$ in subsequent cycles.



Bioinformatics I



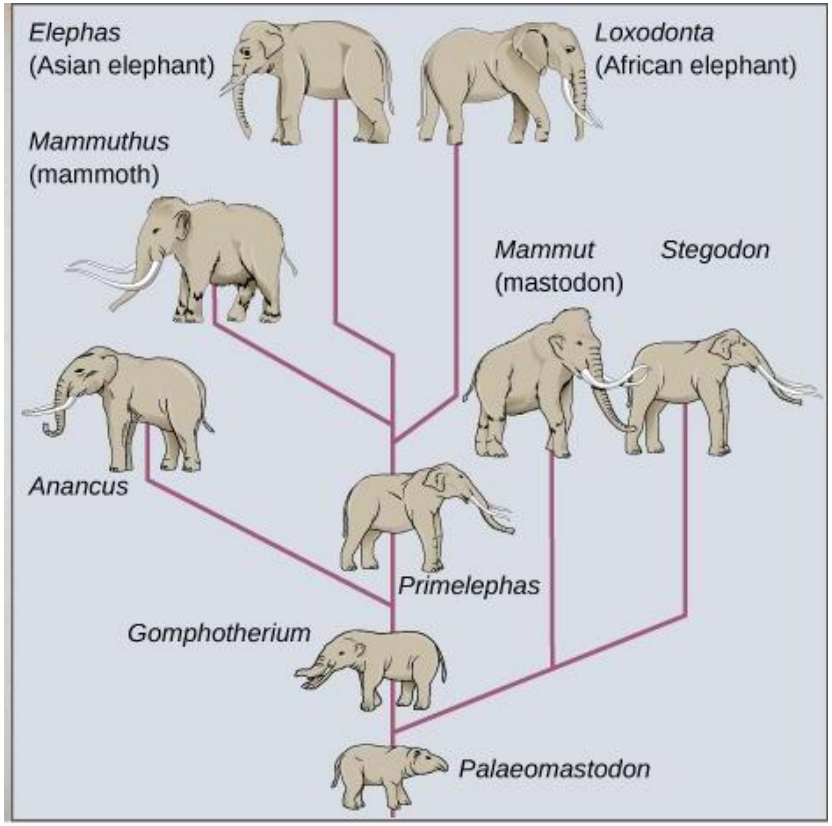
Introduction to Phylogenetic Trees

- A phylogenetic tree can summarize the evolution of various life forms on Earth.

- Scientists consider phylogenetic trees to be a hypothesis of the evolutionary past since one cannot go back to confirm the proposed relationships.

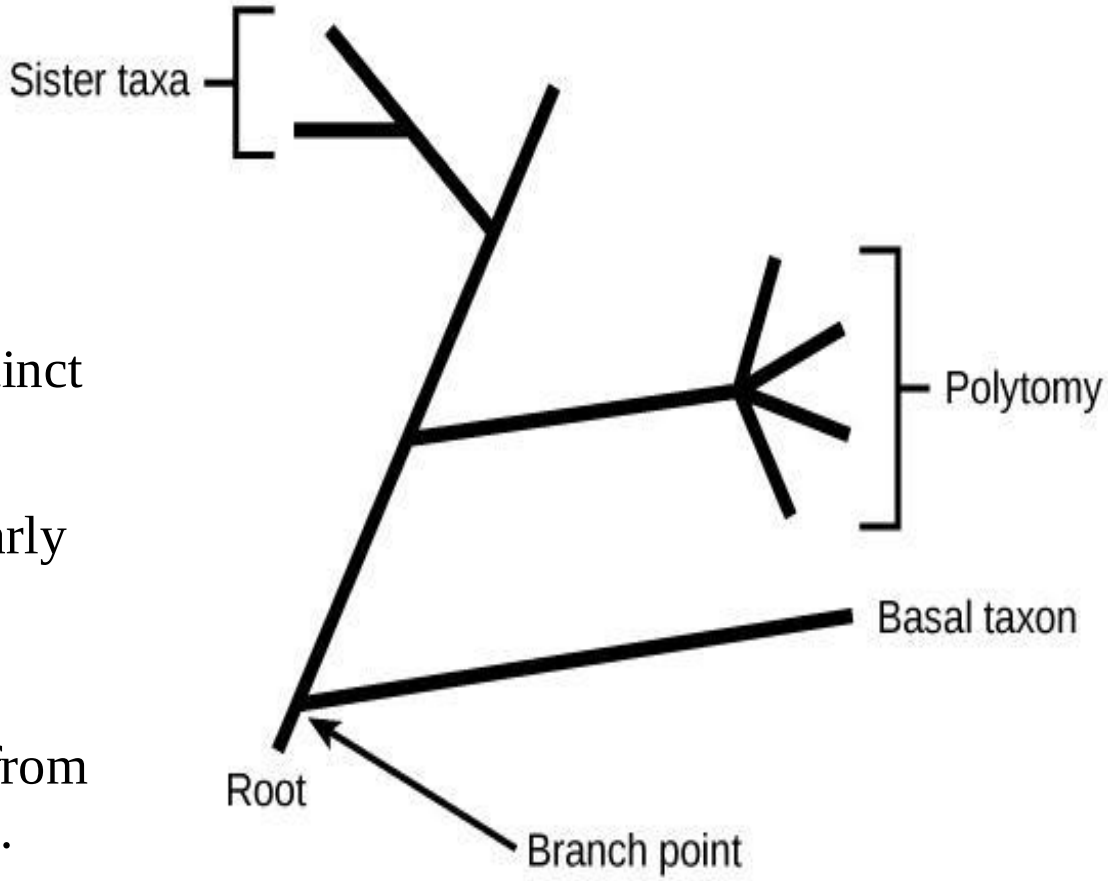
- It is a diagram showing the evolutionary relationships among biological species based on similarities and differences in genetic or physical traits or both.

- Nodes and branches comprise a phylogenetic tree.
  - Internal nodes represent ancestors
  - points in evolution when an ancestor has diverged to form two new species. The length of each branch is proportional to the time elapsed since the split.

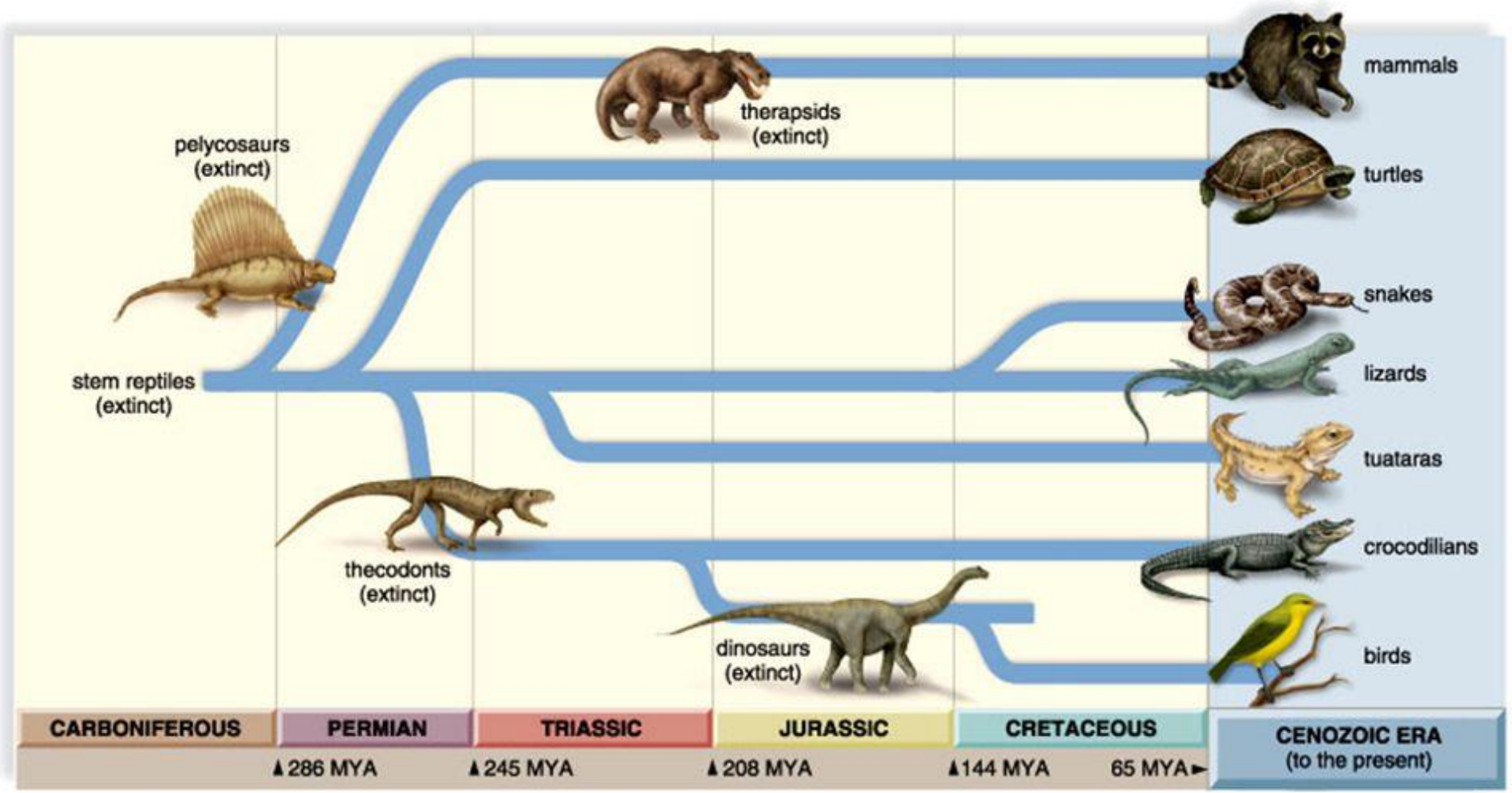


Terminology

- The point where a split occurs, a branch point, represents where a single lineage evolved into a distinct new one.
- We call a lineage that evolved early from the root that remains unbranched a basal taxon.
- We call two lineages stemming from the same branch point sister taxa.

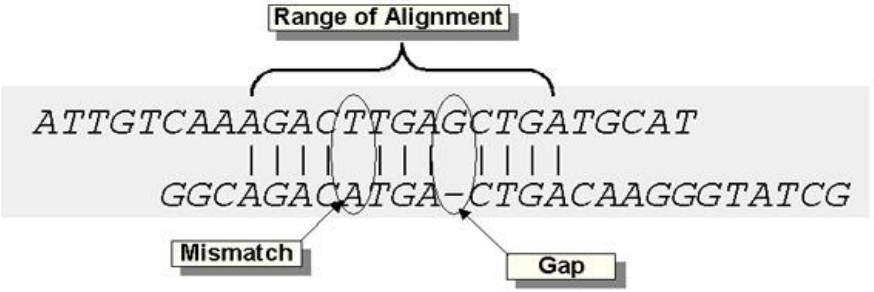


Phylogenetic Tree of Reptiles



Bioinformatics terminology

- Q: What is the Expect (E) value?
- The Expect value (E) is a parameter that describes the number of hits one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the Score (S) of the match increases. Essentially, the E value describes the random background noise. For example, an E value of 1 assigned to a hit can be interpreted as meaning that in a database of the current size one might expect to see 1 match with a similar score simply by chance.
- The lower the E-value, or the closer it is to zero, the more "significant" the match is. However, keep in mind that virtually identical short alignments have relatively high E values. This is because the calculation of the E value takes into account the length of the query sequence. These high E values make sense because shorter sequences have a higher probability of occurring in the database purely by chance.



$S = \sum(\text{identities, mismatches}) - \sum(\text{gap penalties})$

$\text{Score} = \text{Max}(S)$

Other items of interest are:

- **Maximum Score:** the highest alignment score calculated from the sum of the rewards for matched nucleotides or amino acids and penalties for mismatches and gaps.
- **Total Score:** the sum of alignment scores of all segments from the same subject sequence.
- Query Coverage: the percent of the query length that is included in the aligned segments.
- **Identity:** the highest percent identity for a set of aligned segments to the same subject sequence.

- **Biofuel Video Project**

- *Due date for **video—Sunday, October 29 @ 11:59 pm***
- *Double check that video uploaded to YouTube actually is visible to others and plays!*
- *Don't wait until the last minute—uploading takes TIME*
- *Submit link via Peerceptiv (through Canvas)*

- *Due date for **reviews—Sunday, November 5@ 11:59 pm***
- *You **must** submit a video to be eligible to review.*
- *You are graded on your review—don't just give all 5's for everyone—it WILL hurt your grade.*
- *You **must** do three reviews. As you finish one, the next will be assigned.*