O que é a Filogenética?

Para que serve a Filogenética?

Classification

Topology

Phylogeography

Molecular Dating

Detecting gene transfer

Molecular evolution

State reconstruction

...

Com que se faz a filogenética?

Dados Morfológicos

Comportamentais

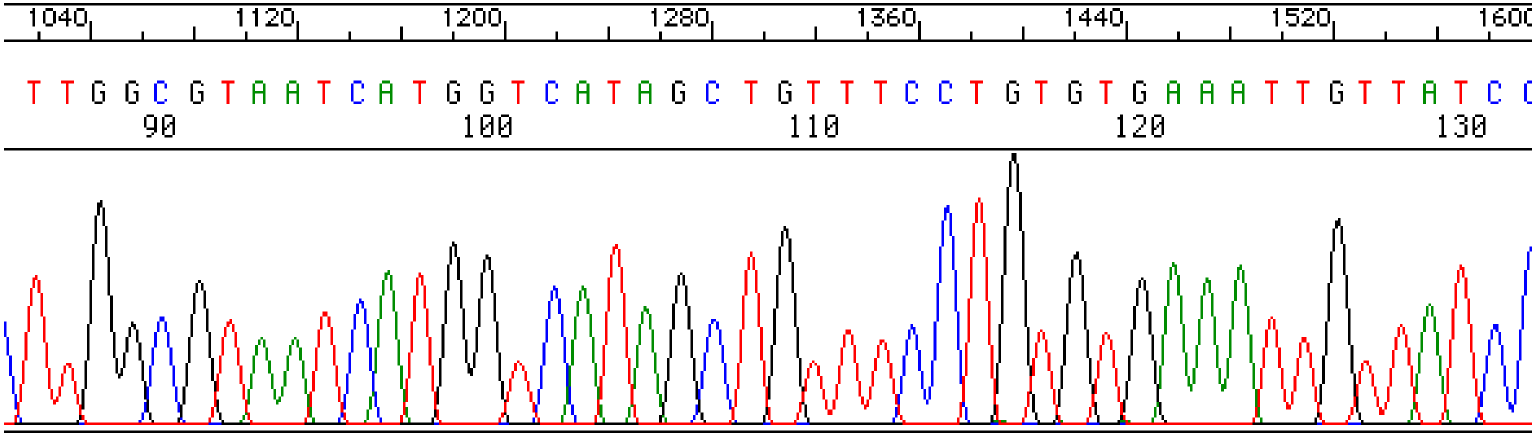
Proteínas

Hybridação

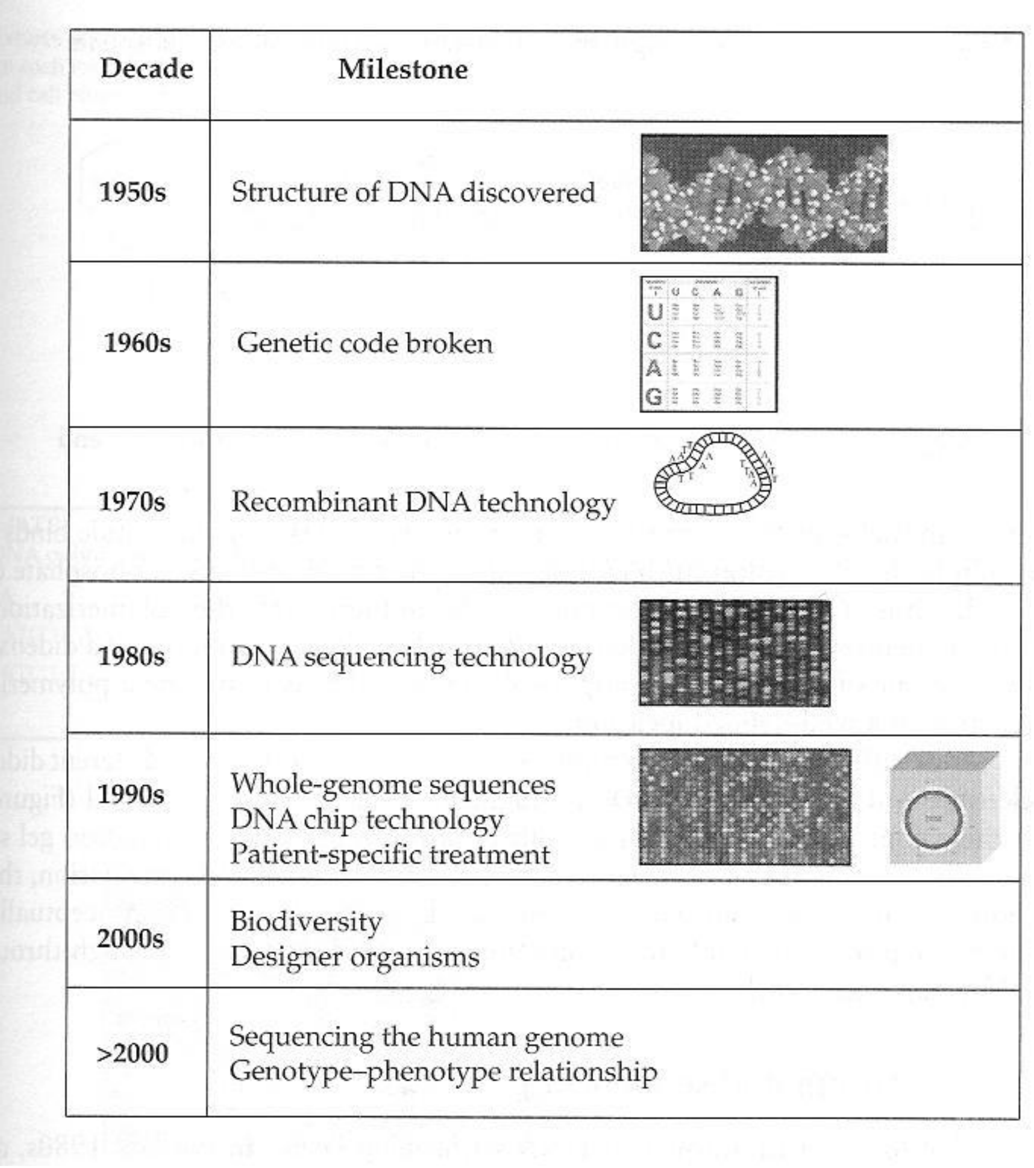
RFLP

Sequenciação

Sequenciação



Sequências



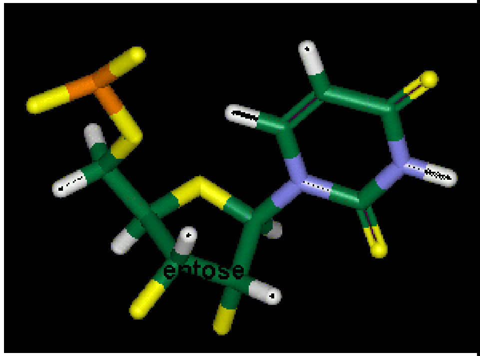
Decade	Milestone
1950s	Structure of DNA discovered
1960s	Genetic code broken
1970s	Recombinant DNA technology
1980s	DNA sequencing technology
1990s	Whole-genome sequences DNA chip technology Patient-specific treatment
2000s	Biodiversity Designer organisms
>2000	Sequencing the human genome Genotype–phenotype relationship

1974 – Sanger Method

Figure 8.3
Some of the key milestones in (1) the discovery of DNA and its content and (2) the development of various DNA technologies.

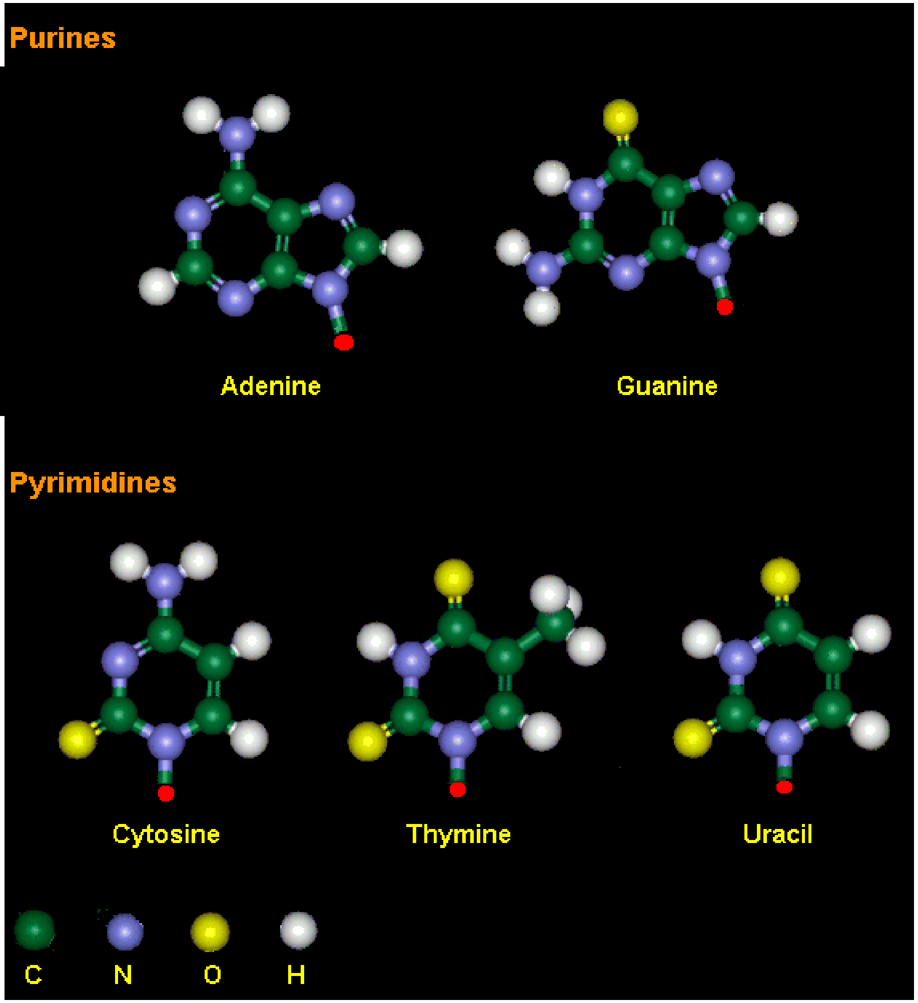
DNA

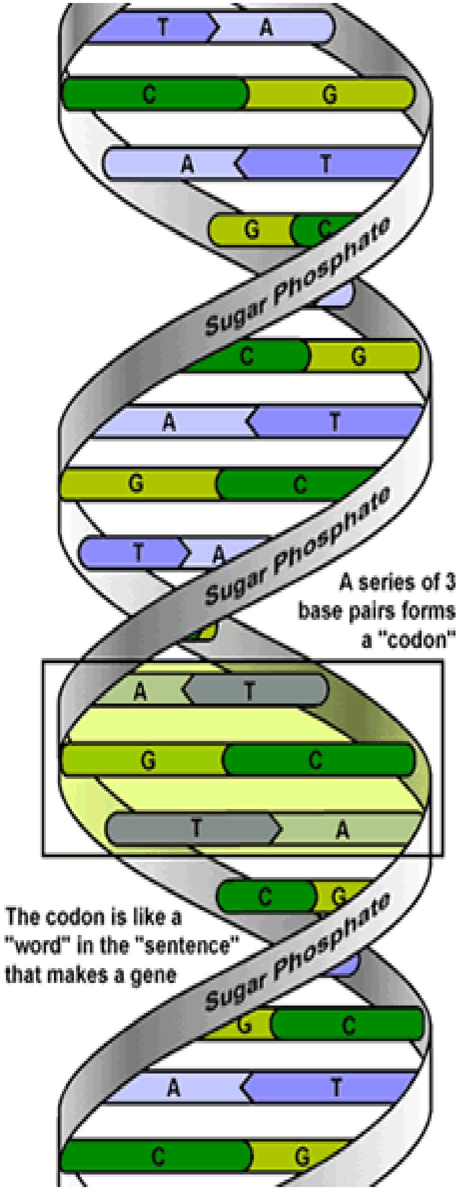
Phosphate Base



Pentose

Nucleotide

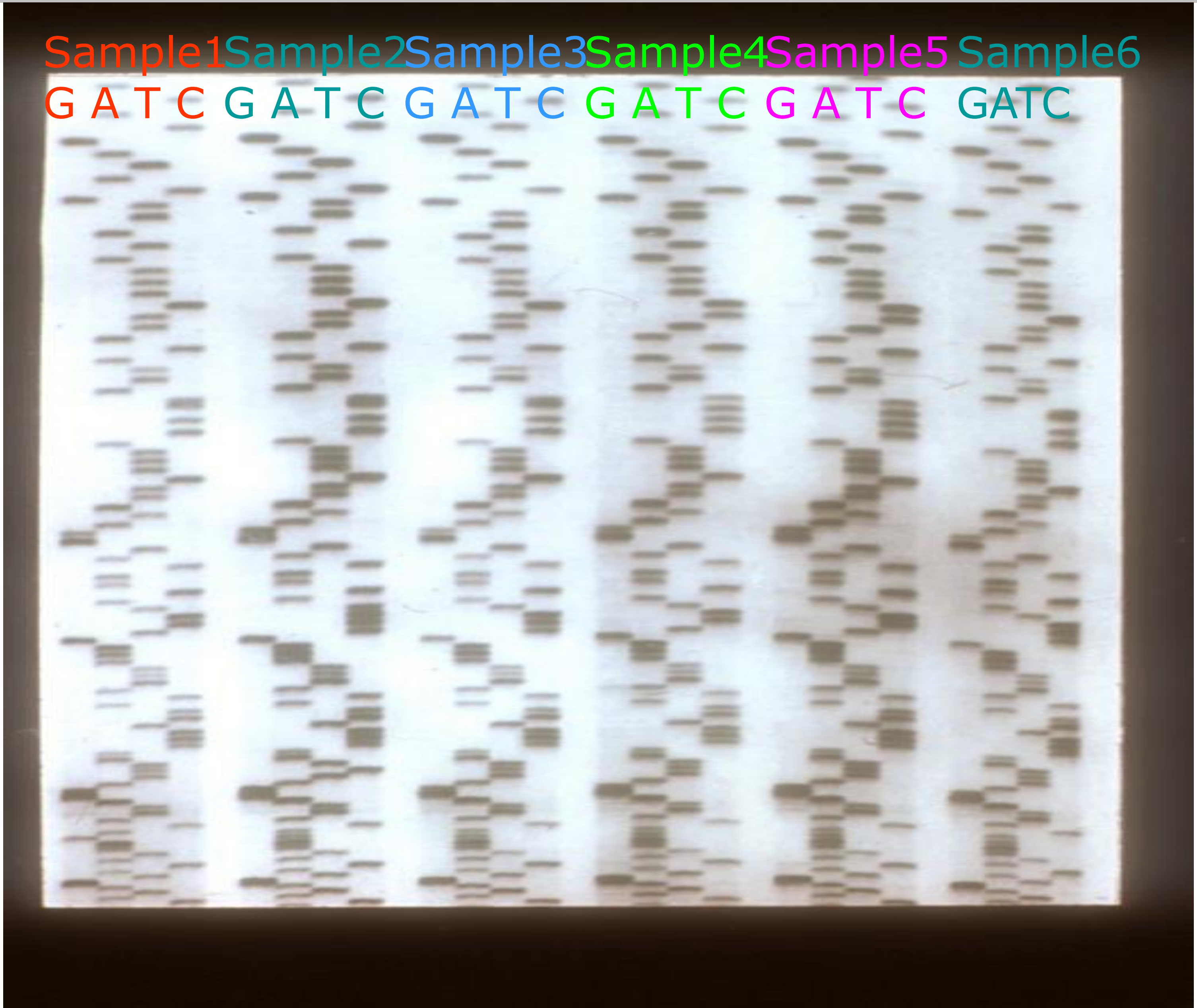




A series of 3 base pairs forms a "codon"

The codon is like a "word" in the "sentence" that makes a gene

Sequenciação Manual



Sequenciação Manual

Sequênciação

Chain termination

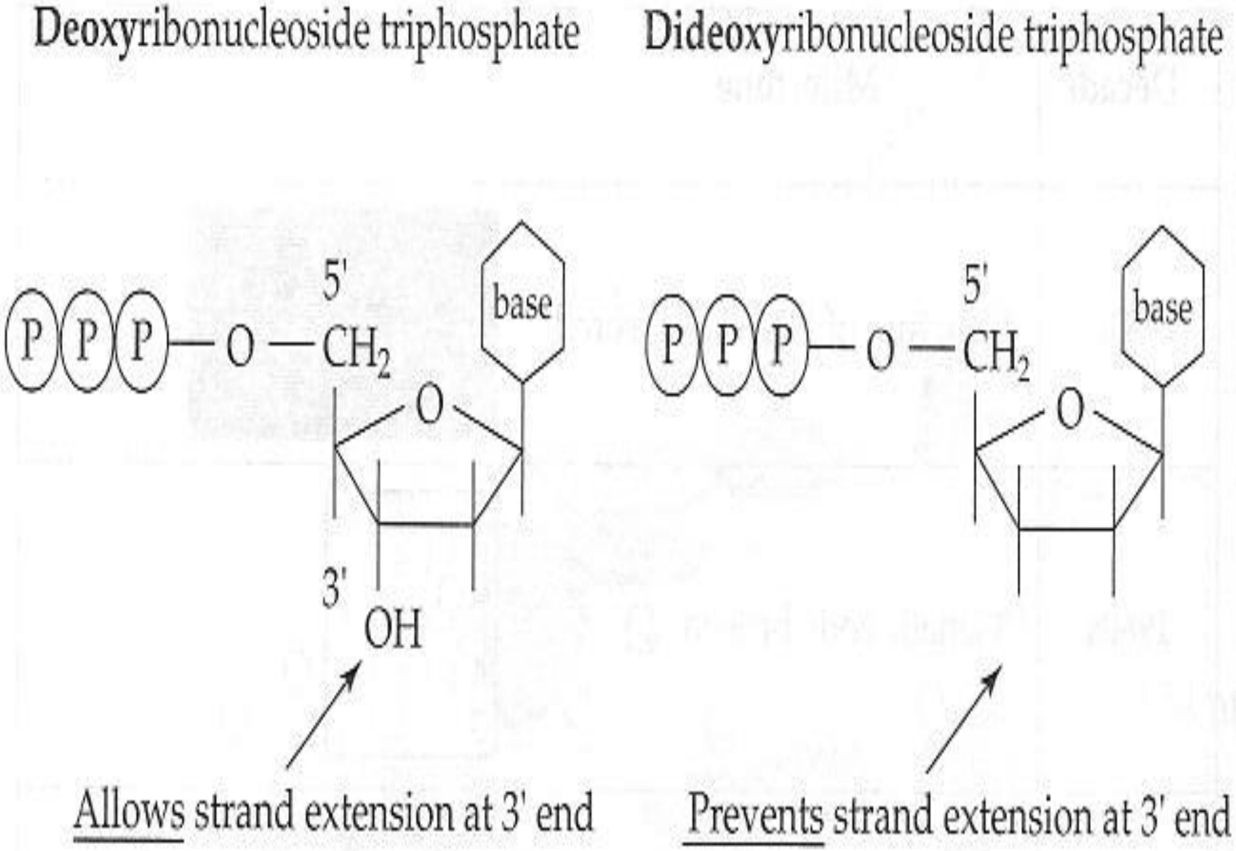
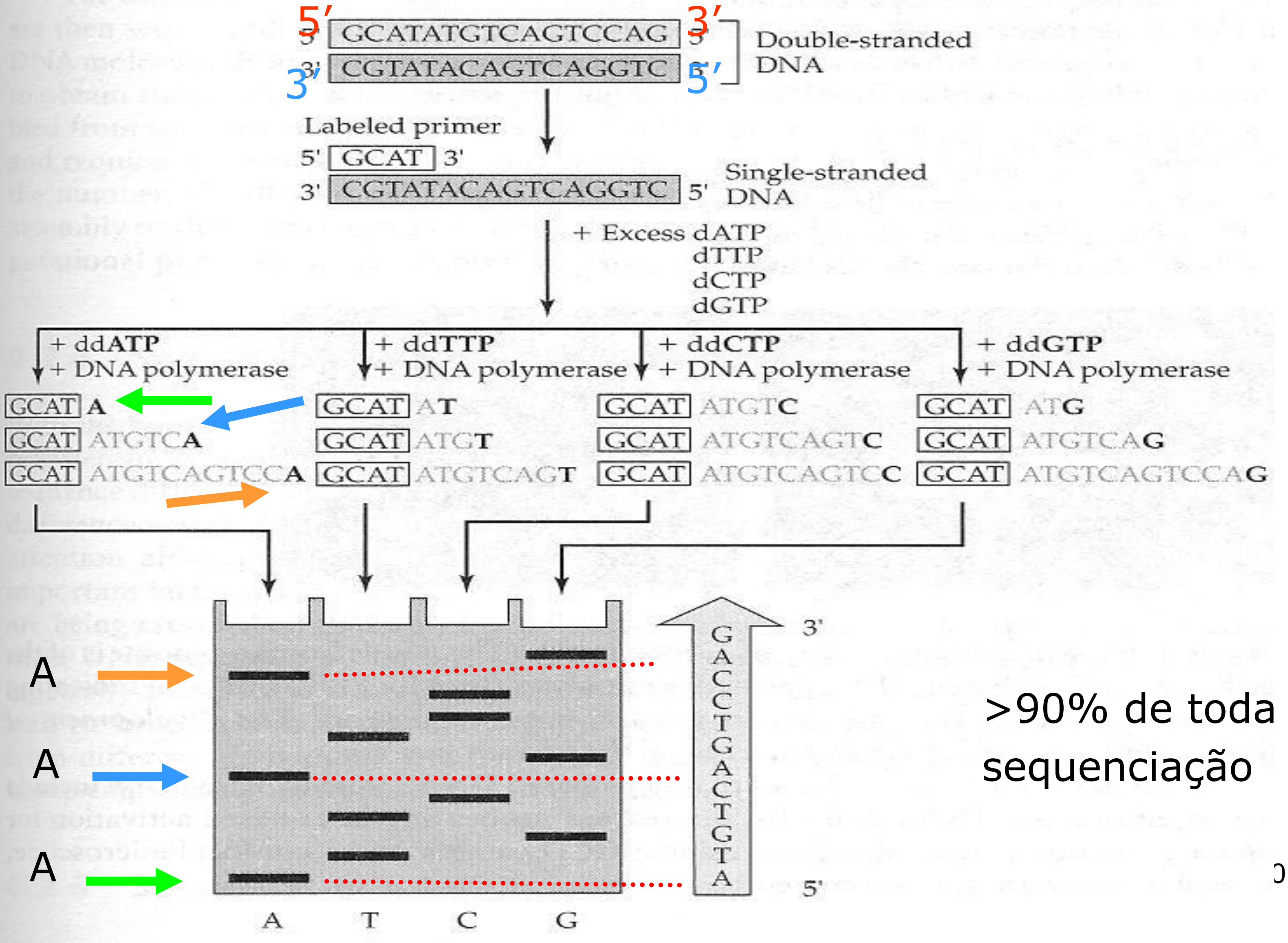


Figure 8.4 The molecular basis for chain termination. The lack of a hydroxyl group in the 3' position of the pentose terminates the polymerization.

“Sanger” Chain termination method



>90% de toda a sequenciação

Sequênciação automática

Began in late 1980s

mid 90s ➔ Applied biosystems automatic sequencer

ABI 377 - gel based - 96 lanes a pop - read length ~500bp - run time ~4-16h => ~40,000 bases/run X 3 runs/day = 120,000

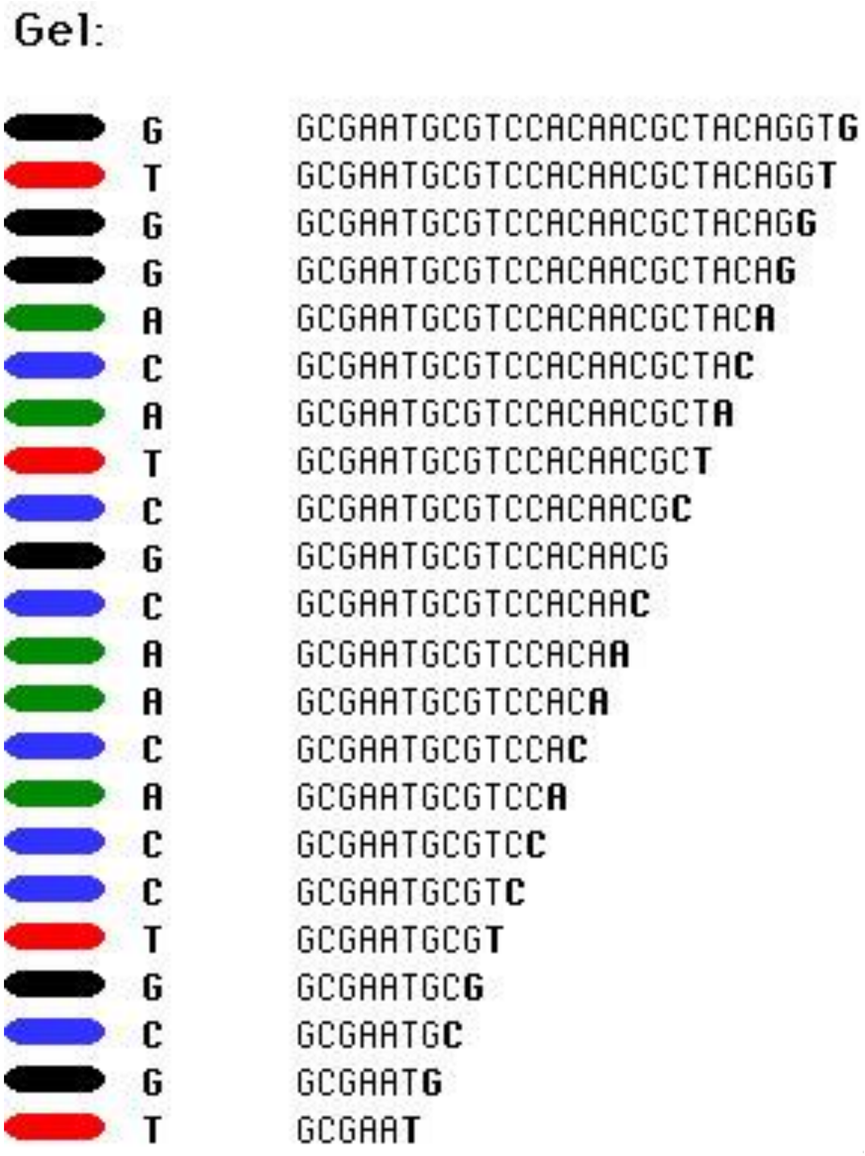
ABI 377XL, 32 templates x 250-400 bases/template x 3 runs/day x 2 base pairs ➔ 75 kilo-base pairs per day

ABI 3700 - 96 templates x 500 bps x 12 runs/day x 2 ➔ 576 kb

ABI 3700 - Capillary based - 48 capillaries - read length ~500bp - run time ~40 minutes => 950,000

3 billion pairs in human genome ➔ 14 years with one machine!

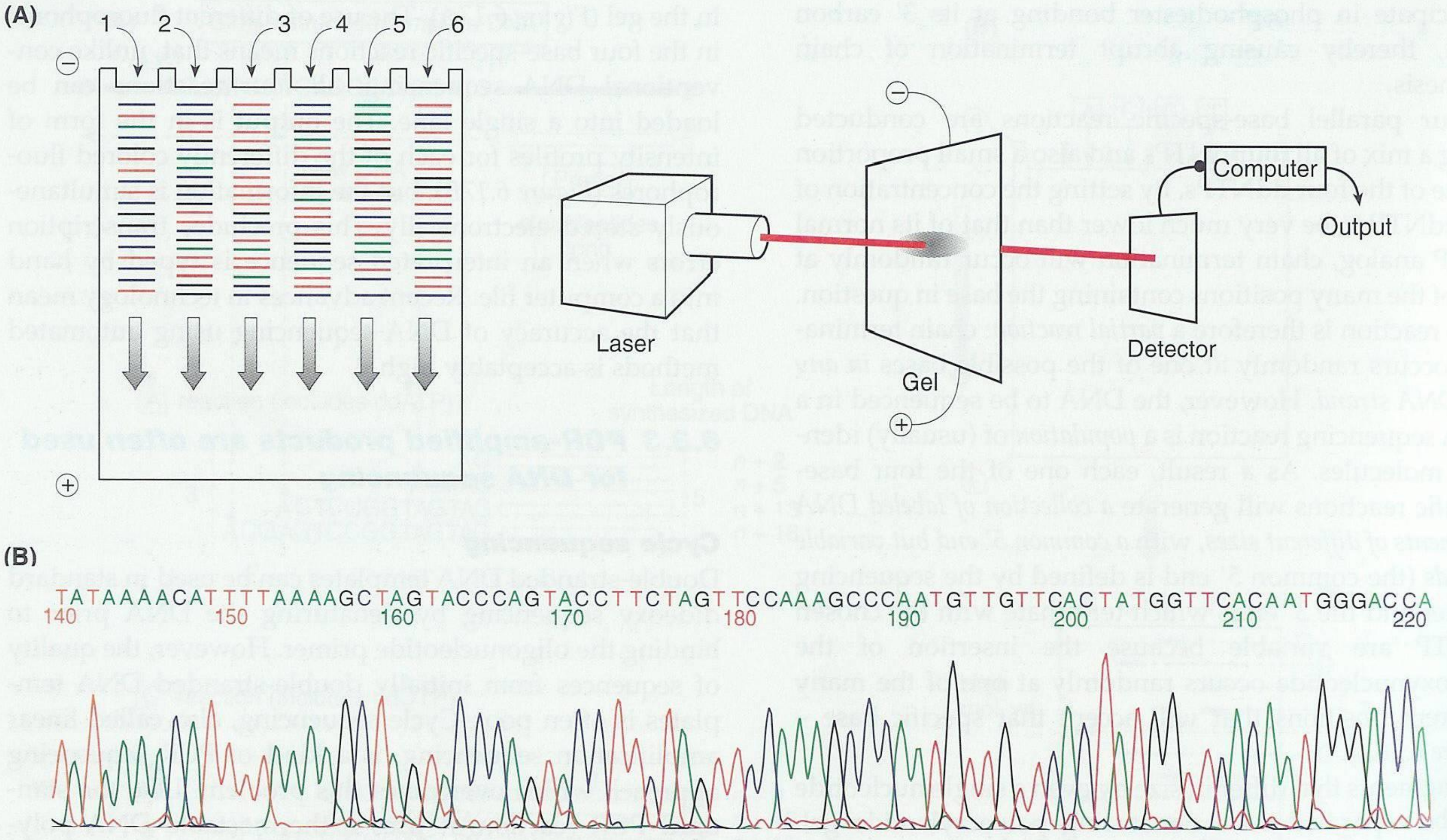
Sequenciadores mais utilizados



Gel:

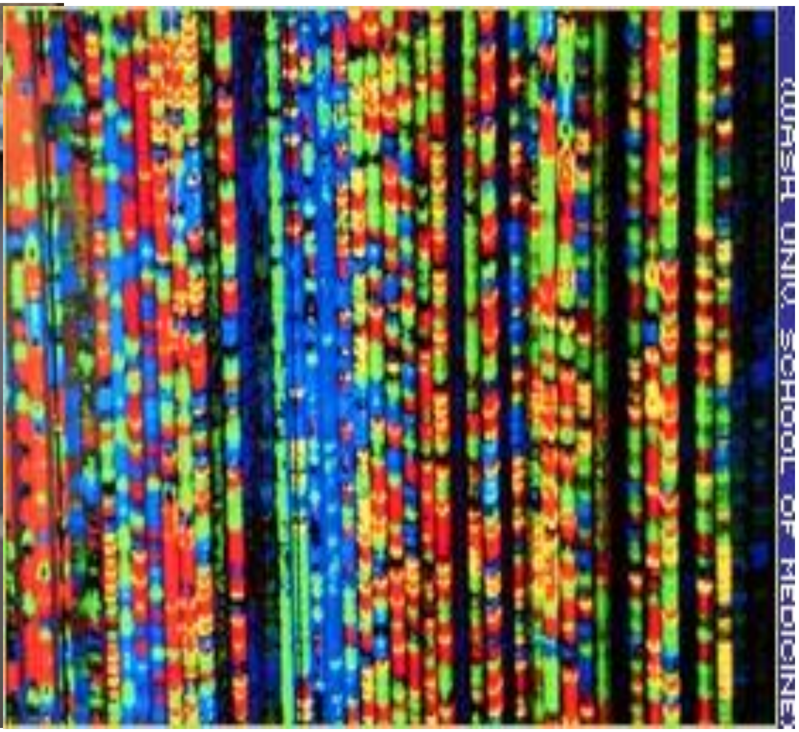
⬛	G	GCGAATGCGTCCACAACGCTACAGGT**G**
🟥	T	GCGAATGCGTCCACAACGCTACAGG**T**
⬛	G	GCGAATGCGTCCACAACGCTACAG**G**
⬛	G	GCGAATGCGTCCACAACGCTACA**G**
🟩	A	GCGAATGCGTCCACAACGCTAC**A**
🟦	C	GCGAATGCGTCCACAACGCTA**C**
🟩	A	GCGAATGCGTCCACAACGCT**A**
🟥	T	GCGAATGCGTCCACAACGC**T**
🟦	C	GCGAATGCGTCCACAACG**C**
⬛	G	GCGAATGCGTCCACAACG
🟦	C	GCGAATGCGTCCACAA**C**
🟩	A	GCGAATGCGTCCACA**A**
🟩	A	GCGAATGCGTCCAC**A**
🟦	C	GCGAATGCGTCCA**C**
🟩	A	GCGAATGCGTCC**A**
🟦	C	GCGAATGCGTC**C**
🟦	C	GCGAATGCGT**C**
🟥	T	GCGAATGCG**T**
⬛	G	GCGAATGC**G**
🟦	C	GCGAATG**C**
⬛	G	GCGAAT**G**
🟥	T	GCGAA**T**

Sequênciação automática

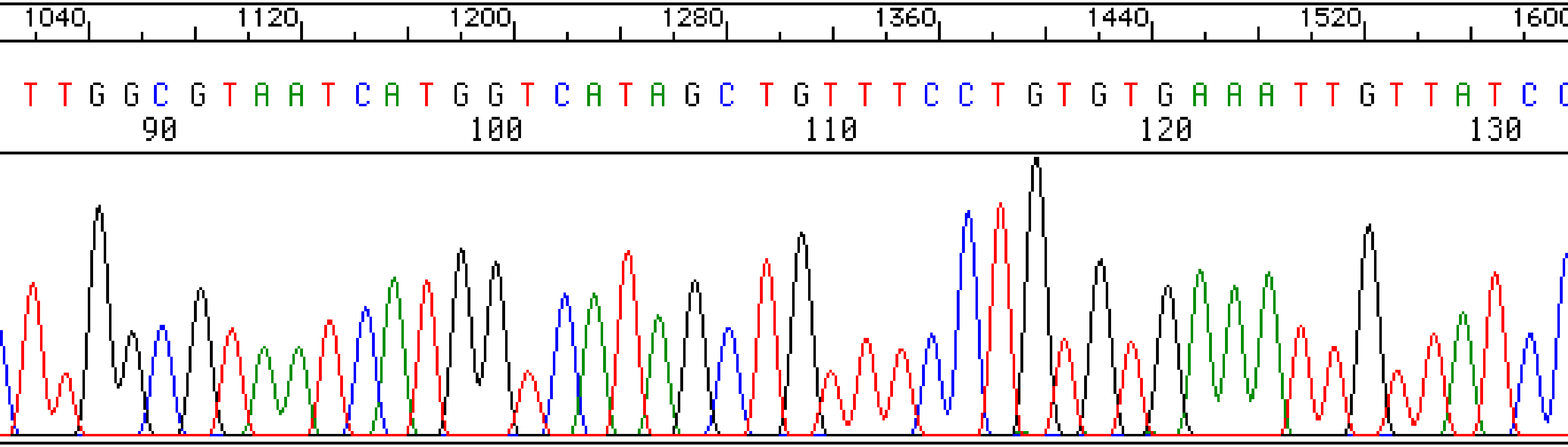


Automated sequencing machines,

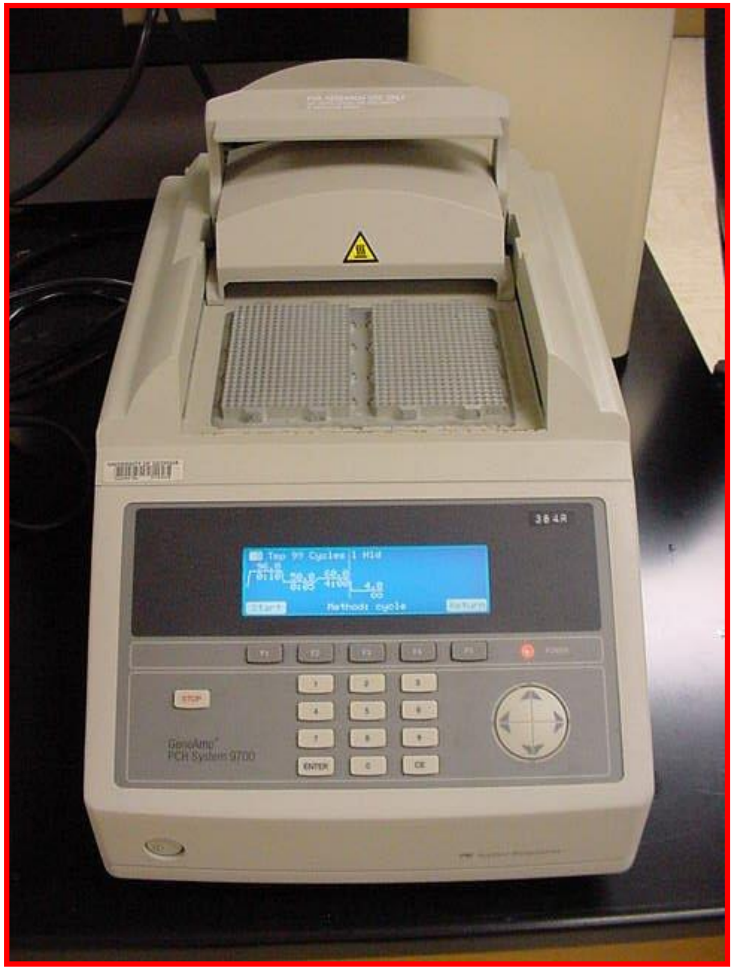
particularly those made by PE Applied Biosystems, use 4 colors of dye, so they can read all 4 bases at once.



Chromatogram



Perkin Elmer 9700



ABI 3700


Applied Biosystems
HITACHI
DT/651
9-14
86%/619
ABI PRISM
3700 DNA Analyzer


EXIT

Outros métodos de sequenciar

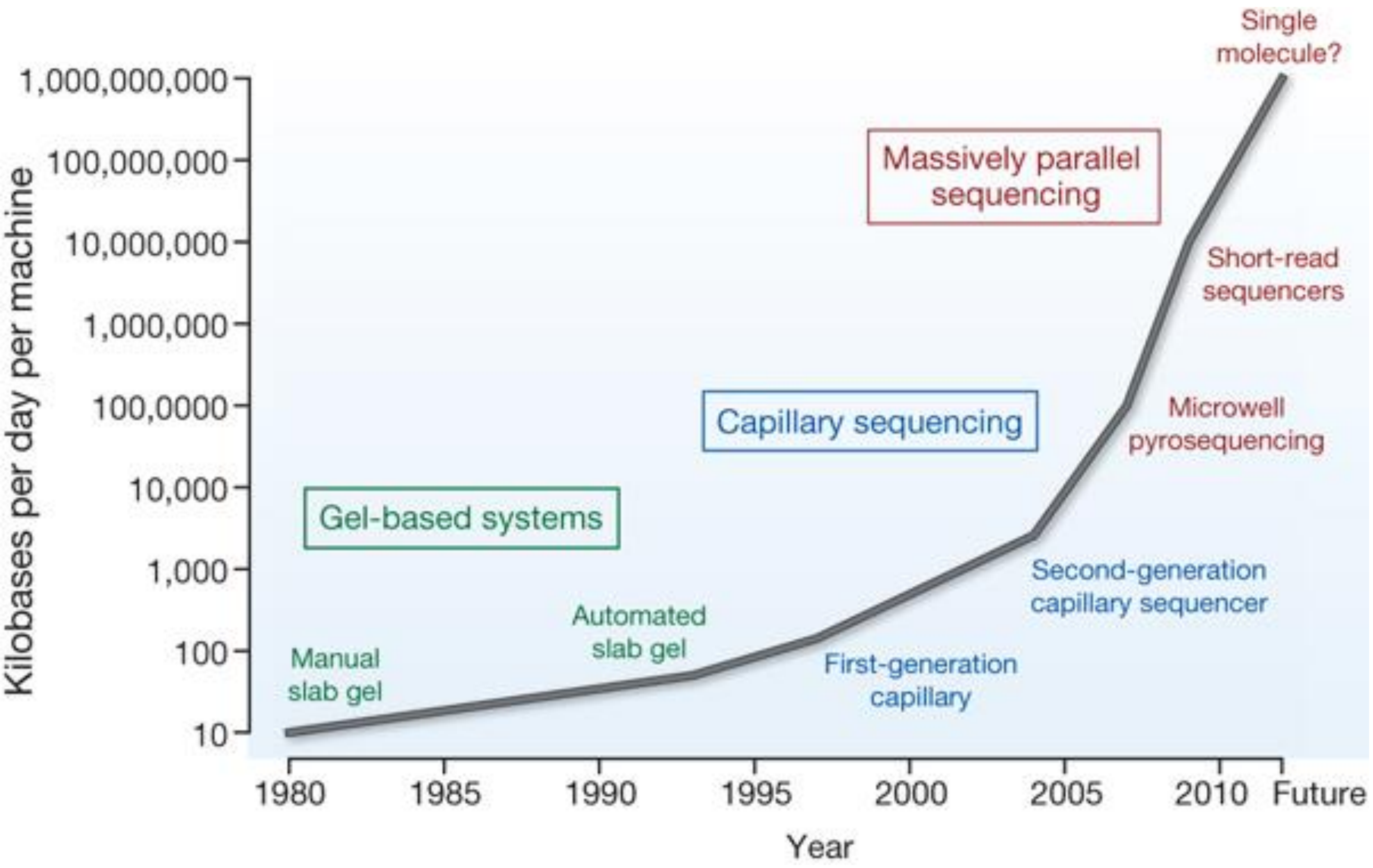
Hybridization method - Hybridize to oligos on a chip

- Affymetrix can do 30K resequence
- Limited by number of features and hybridization specificity

Single molecule methods

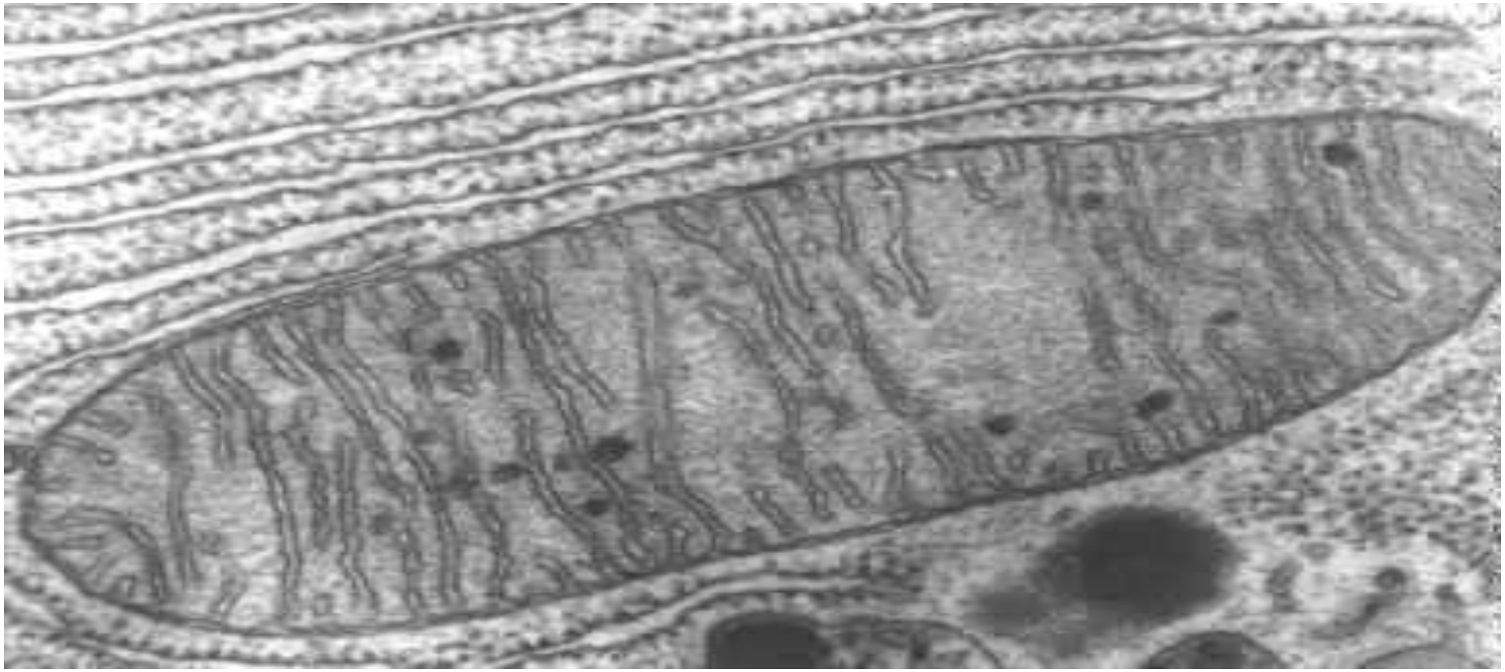
- Pore-base - threads DNA through molecular pore in membrane - bases determined by changes in conductance
- Mass spec - best for small molecules now like SNPs

Improvements in the rate of DNA sequencing over the past 30 years and into the future.



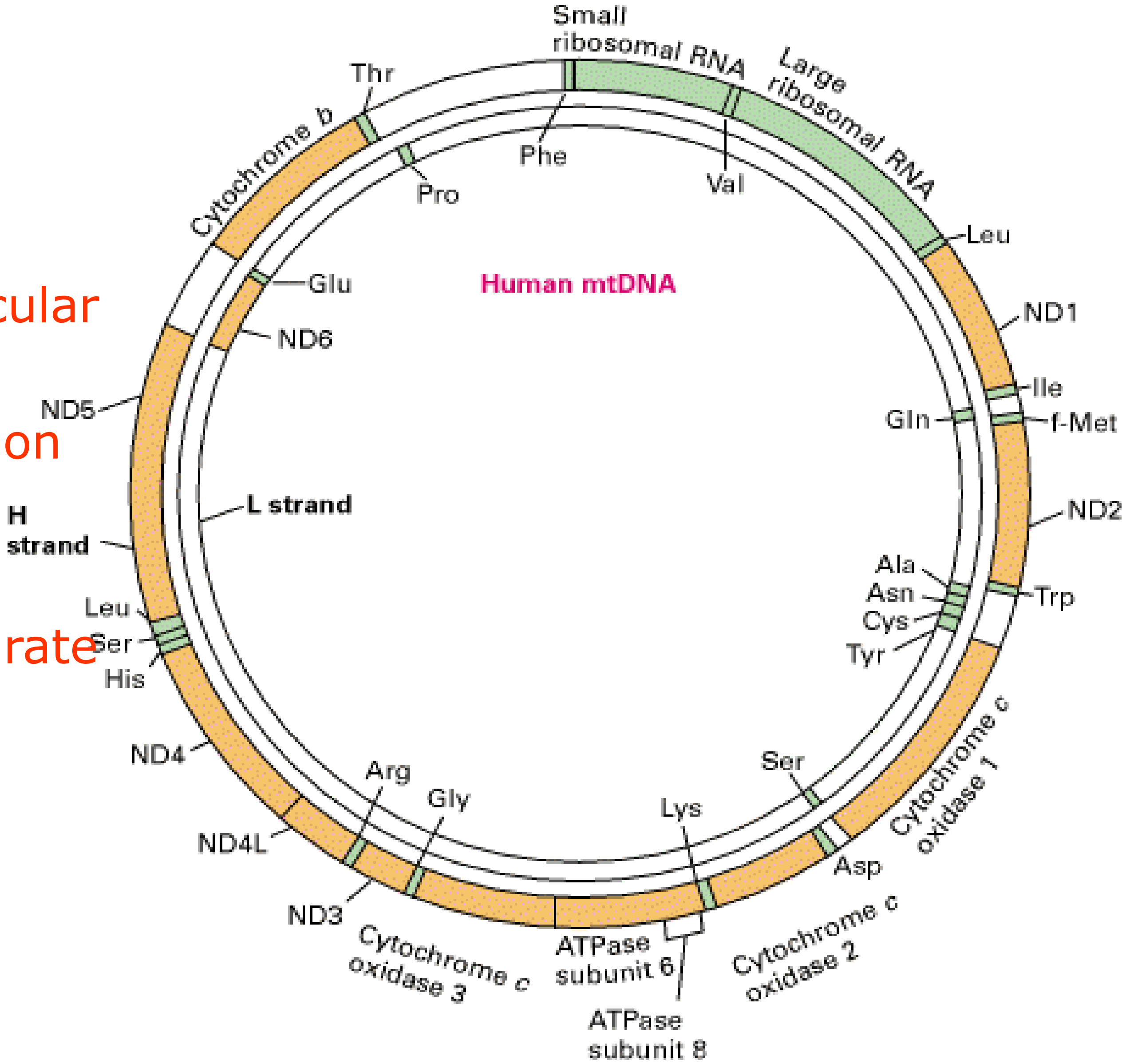
MR Stratton *et al.* *Nature* **458**, 719-724 (2009) doi:10.1038/nature07943

DNA Mitocondrial



DNA Mitocondrial

- Double-stranded, circular
- 16,569 base pairs
- Independent replication
- No recombination
- No DNA repair
- 10X higher mutation rate



DNA Mitocondrial

Oocytes have 100,000 mitochondria

Sperm have < 100 mitochondria

Sperm mitochondria is destroyed after fertilization

Mitochondria is only inherited from the mother

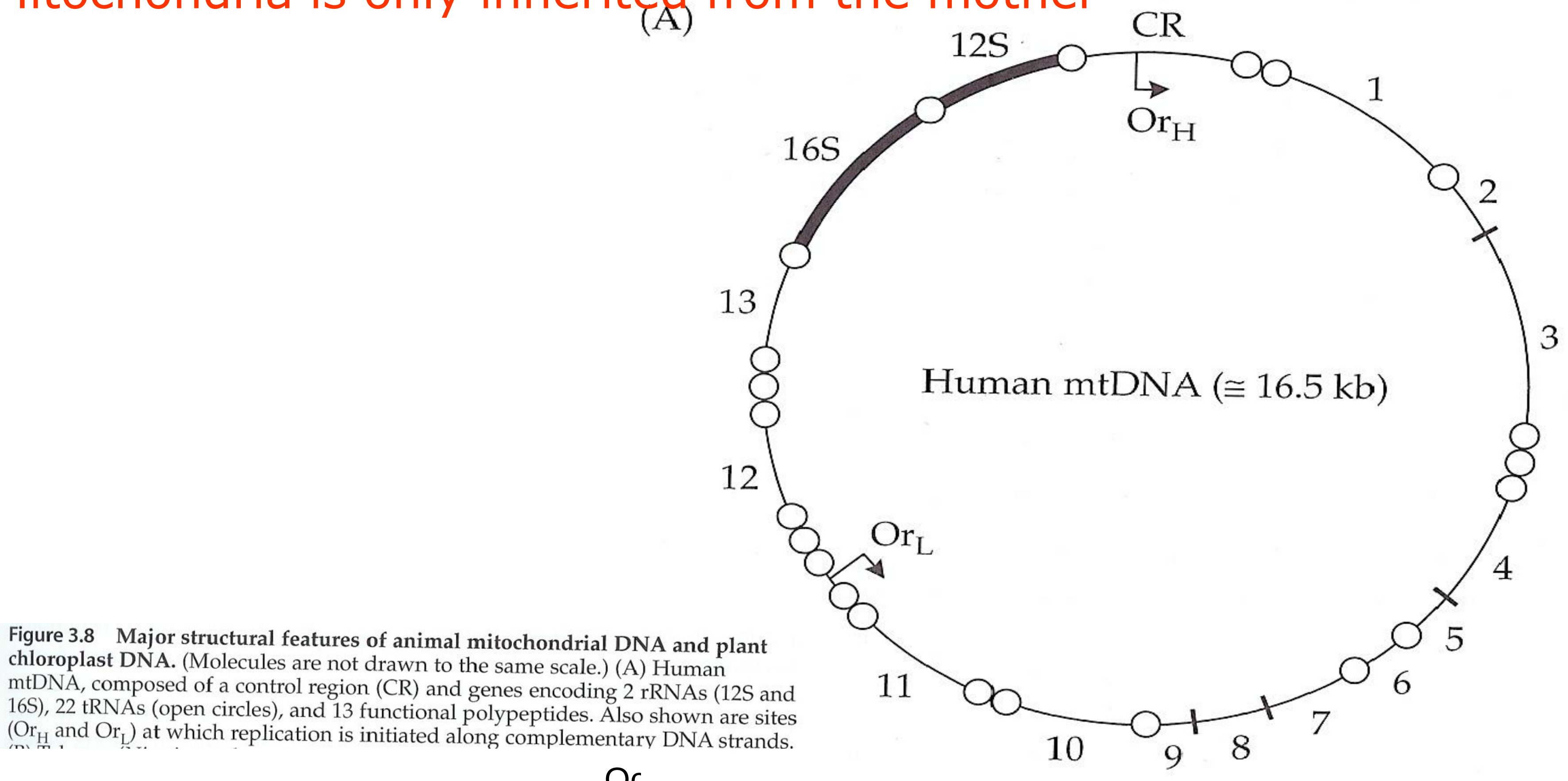
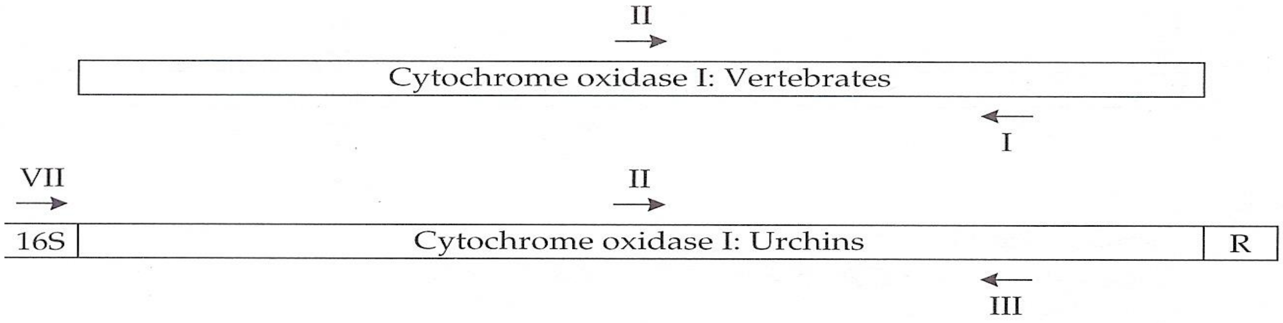


Figure 3.8 Major structural features of animal mitochondrial DNA and plant chloroplast DNA. (Molecules are not drawn to the same scale.) (A) Human mtDNA, composed of a control region (CR) and genes encoding 2 rRNAs (12S and 16S), 22 tRNAs (open circles), and 13 functional polypeptides. Also shown are sites (Or_H and Or_L) at which replication is initiated along complementary DNA strands.

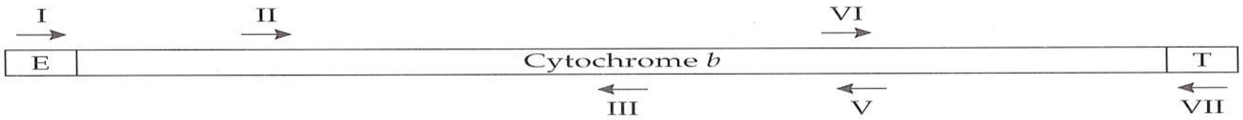
Oc

DNA Mitocondrial - COI



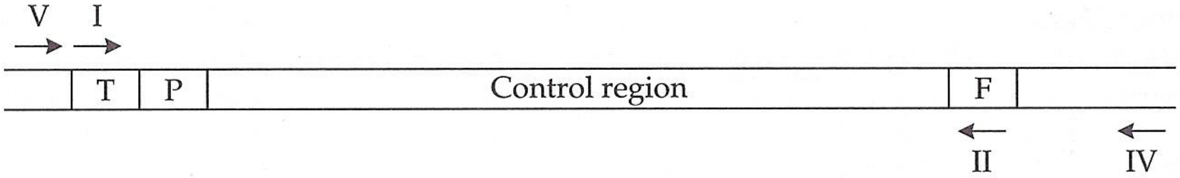
Map position	Primer/Taxa		Sequence position
I.	CO1e-3′	CCA GAG ATT AGA GGG AAT CAG TG	
	Human	. . T . . . . A . . . G . . A	7110
	Frog	. . . . T A . A . . A C	8602
	Urchin	. . . . A G . . G . . A . . C	6992
	Fruit fly	. . . . T A . A . . A T . . . T	2672
II.	CO1f-5′	CCT GCA GGA GGA GGA GAY CC	
	Human	. . C . . C . . . . C . .	6569
	Frog	. . . . . C . . . . T . . C . .	8061
	Urchin	. . . . G . . . . . T . .	6451
	Fruit fly	. . A . . T . . . . T . .	2131
III.	CO1a-3′	AGT ATA AGC GTC TGG GTA GTC	
	Human	G . . G . . T . . A . . G	7227
	Frog	T	8720
	Urchin	T . . . . . G . . . . A	7108
	Fruit fly	. . . G . . . . . A . . A . . . . . A . .	2791
VII.	16SB-5′	ACG TGA TCT GAG TTC AGA CCG G	
	Human	. . . .	3080
	Frog	. . . .	4572
	Urchin	. . . .	5682
	Fruit fly	. . . . . T . . . . T . .	12887

DNA Mitocondrial -cytb



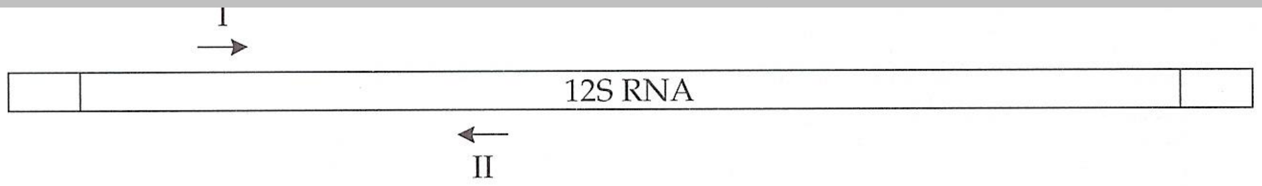
Map position	Primer/Taxa										Sequence position
I.	GLU-5′	TGA	TAT	GAA	AAA	CCA	TCG	TTG			
	Human	...	C..	...	...	T..	...	...			14724
	Frog	...	.C.	...	...	...	C..	...			
	Carp	...	CT.	...	G..	...	C..	...			
	Chicken	C.G	CT.	...	..G	...	.T.	...			
	Shrimp	..C	AT.	...	..T	G..	.T.	.AT			
I.	GLUDG-5′	TGA	CTT	GAA	RAA	CCA	YCG	TTG			14724
II.	CB1-5′	CCA	TCC	AAC	ATC	TCA	GCA	TGA	TGA	AA	
	Human	...	...	...	...	..C	...	...	...	..	14817
	Stingray	...	...	...	...	..C	A.T	...	...	..	
	Frog	...	..A	...	..T	...	T..	.T.	...	..	16321
	Urchin	..C	...	...	C.T	..C	ATT	..G	...	..	14581
III.	CB2-3′	CCC	TCA	GAA	TGA	TAT	TTG	TCC	TCA		
	Human	...	...	...	...	...	...	...	...		15175
	Stingray	...	...	...	...	...	...	...	...		
	Frog	...	...	A..	A..	...	...	...	...		16677
	Urchin	AG.	...	A..	G..	C..	...	C..	C..		14937
V.	CB3-3′	GGC	AAA	TAG	GAA	RTA	TCA	TTC			
	Human	...	G..	...	...	A..	...	...			15560
	Frog	...	G..	...	...	G..	...	...			17065
	Stingray	...	...	...	...	A..	...	...			
	Sturgeon	...	...	G..	A..	G..	...	...			
	Urchin	...	G..	..A	...	G..	C..	...			
	Fruit fly	A..	...	..A	A..	A..	...	...			11325
VI.	CB3R-5′	CAT	ATT	AAA	CCC	GAA	TGA	TAY	TT		
	Rat	..C	..C	...	..A	...	...	..T	..		15560
	Frog	...	...	...	..A	...	...	..C	..		17065
	Stingray	...	...	...	...	...	...	..C	..		
	Urchin	..C	...	C.G	..A	...	..G	..C	..		
	Fruit fly	..C	...	C..	..A	...	...	..T	..		11325
VII.	CB6THR-3′	CTC	CAG	TCT	TCG	RCT	TAC	AAG			
	Human	T.T	..T	...	C..	GT.	...	...			15930
	Frog	...	..T	...	...	G..	...	...			
	Sturgeon	..T	.GA	...	...	GT.	...	...			
	Shark	...	...	...	...	A..	...	...			
	Urchin	.C.	TCT	CTC	CT.	GT.	...	...			
	Fruit fly	A..	TTA	.T.	.T.	AT.	...	...			

DNA Mitocondrial –Control region



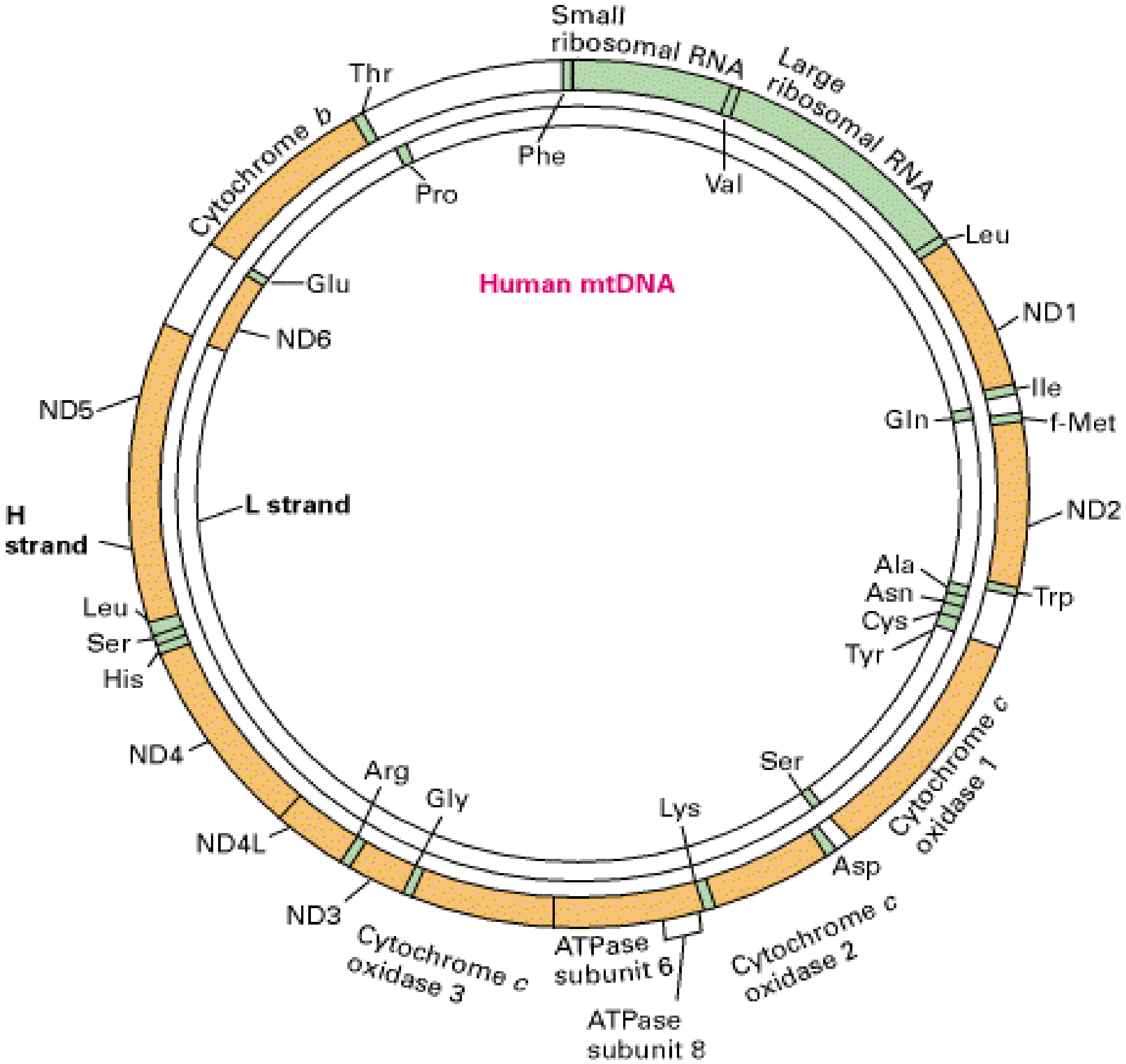
Map position	Primer/Taxa		Sequence position
I.	PRO-5′	CTA CCT CCA ACT CCC AAA GC	
	Human	.C. ..A TT. G.A	15980
	Frog	.C. ..A TTG ..C	
	Sturgeon	TC. ..C TT.	
	Urchin	TAC AT. G..	
II.	PHE-3′	TCT TCT AGG CAT TTT CAG TC	
	Human	C.G ... .AA	625
	Frog	..A ... CA.	
	Urchin	C.. .TG .A.	
V.	CB3R-5′	CAT ATT AAA CCC GAA TGA TAT TT	
	Human	..C ..C ... ..A	15560
	Stingray	... ...A	
	Frog	... ...	
	Urchin	..C ... C.G ..A ... ..G	
	Fruit fly	..C ... C.. ..A	
IV.	12SAR-3′	ATA GTG GGG TAT CTA ATC CCA GTT	
	Human	... ...	1067
	Frog	... ...	
	Fruit fly	... A.A ... ... .T. ...	

DNA Mitocondrial – rRNA 12S

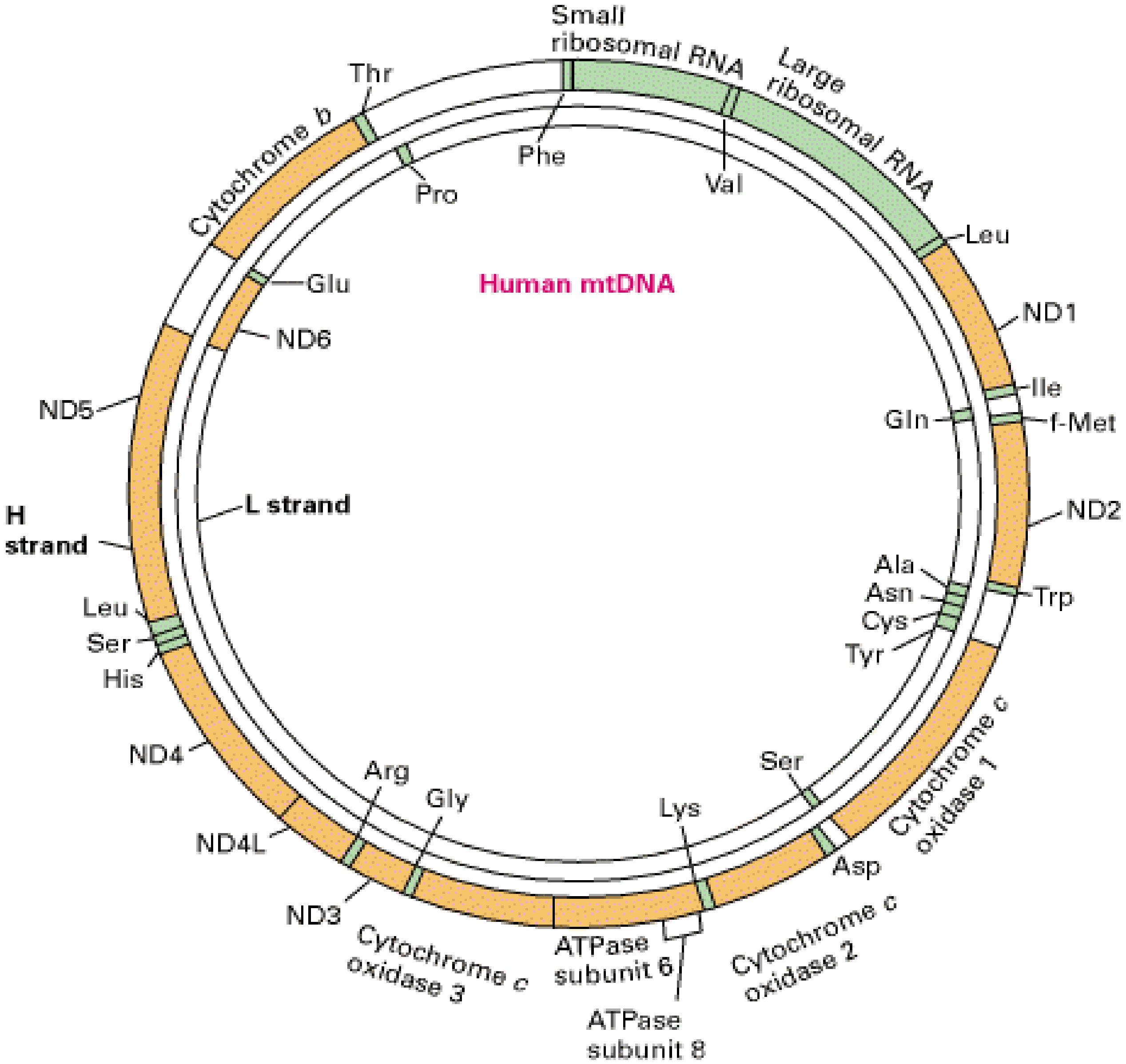


Map position	Primer/Taxa		Sequence position
I.	12SA-5′	AAACTGGGATTAGATACCCCACTAT	
	Human	....	1067
	Frog	....	2486
	Urchin	.....CA............TGT...	491
	Fruit fly	......A............T.T...	14612
	12Sai-5′	AAACTAGGATTAGATACCCTATTAT	
	Human	.....G.............C.C...	1067
	Urchin	....C...............G....	491
	Fruit fly	....	14588
II.	12SB-3′	GAGGGTGACGGGCGGTGTGT	
	Human	....	1478
	Frog	....	2898
	Urchin	...A..........A.....	853
	Fruit fly	A..A.C........A.....	14211
	12Sbi-3′	AAGAGCGACGGGCGATGTGT	
	Human	G..G.T........G.....	1478
	Urchin	G....T..............	855
	Fruit fly	....	14214

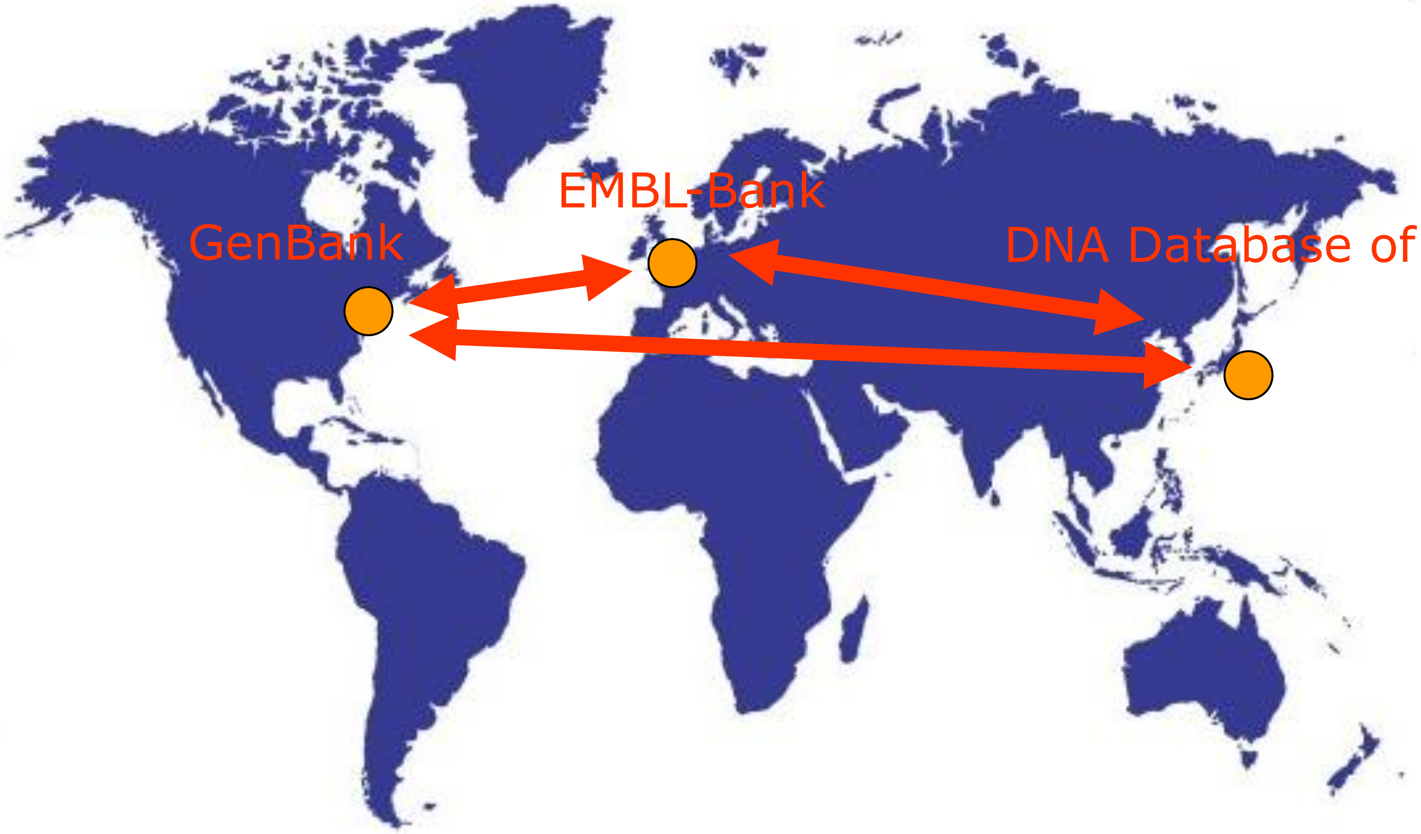
DNA Mitocondrial –velocidades diferentes



DNA Mitocondrial – Genes Nucleares



Bases de Dados



Bases de Dados

GenBank	http://www.ncbi.nlm.nih.gov/
EMBL	http://www3.ebi.ac.uk/Services/DBStats/
UniProt	http://beta.uniprot.org/
EBI	http://www.ebi.ac.uk/
Ensembl	http://www.ensembl.org/index.html

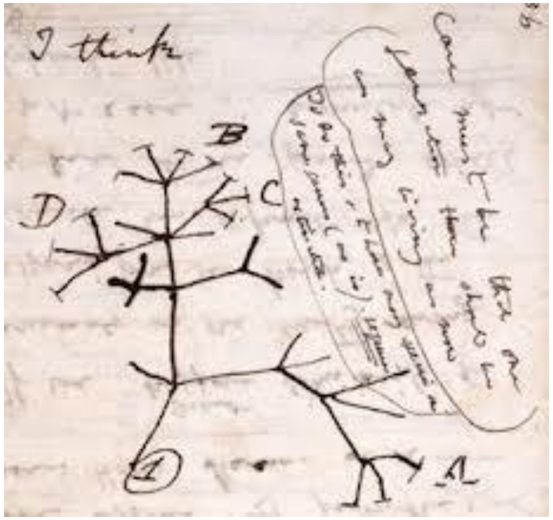
BioEdit	http://www.mbio.ncsu.edu/BioEdit/bioedit.html
Mega	http://www.megasoftware.net/

Referencia para a Filogenética

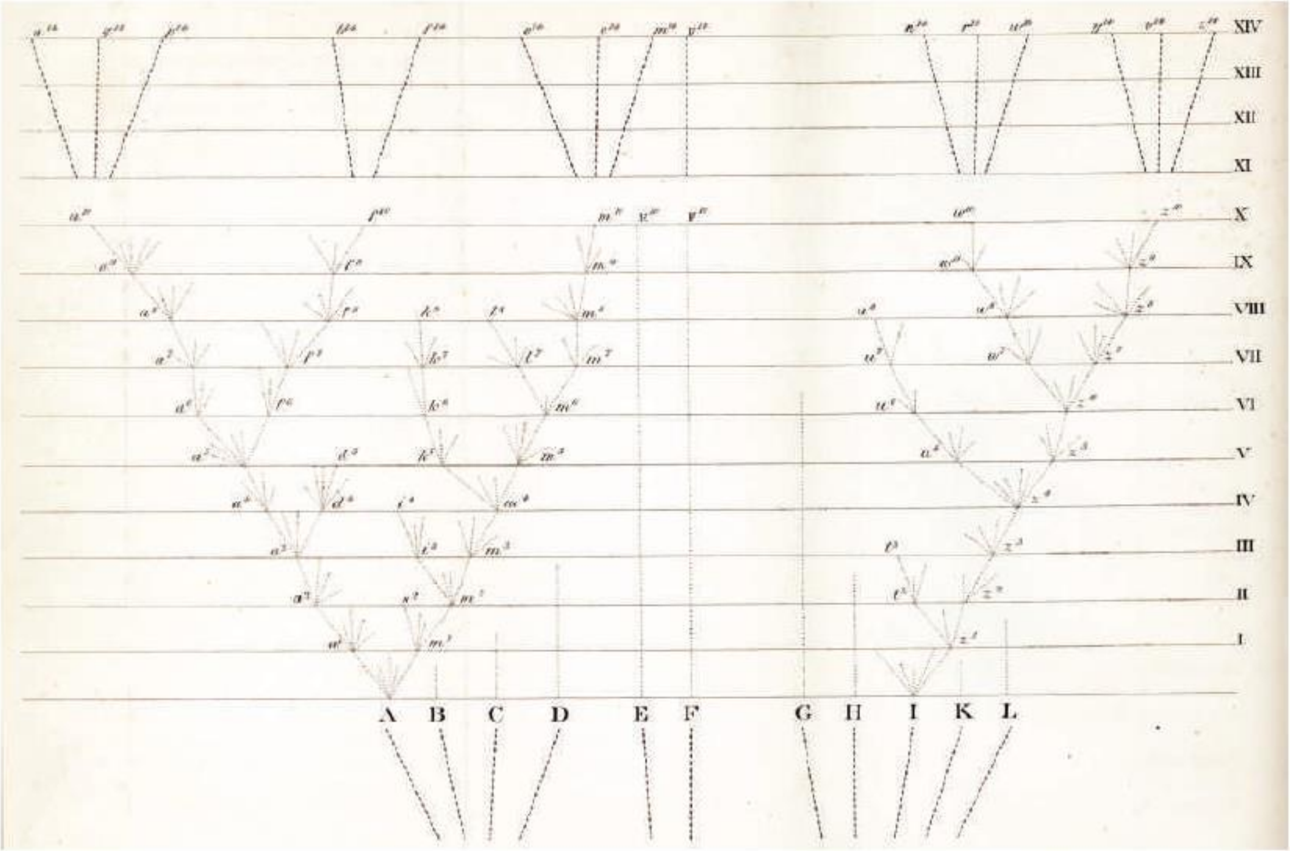
Felsenstein J. 2004. Inferring Phylogenies. Sunderland: Sinauer Associates, Inc.

Lemey P, Salemi, M., and Vandamme, A.-M. 2009. The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing In: Cambridge University Press.

Filogenética

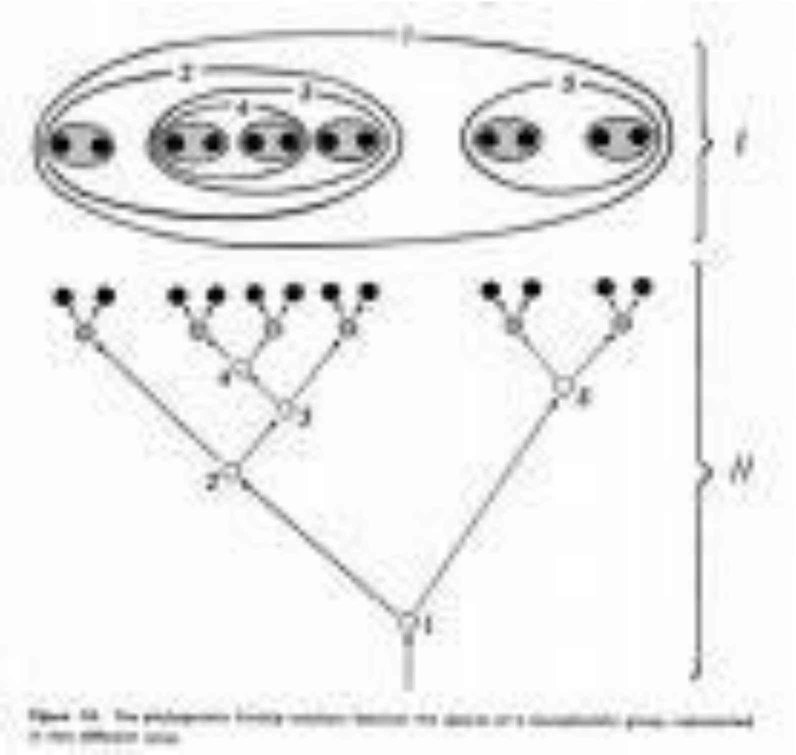


Darwin notebook july 1837



Darwin Origin of Species 1859

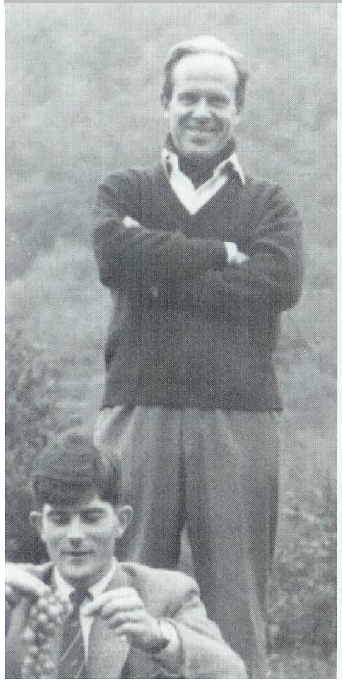
Cladistic analysis - 50s – 60s parsimony perpective

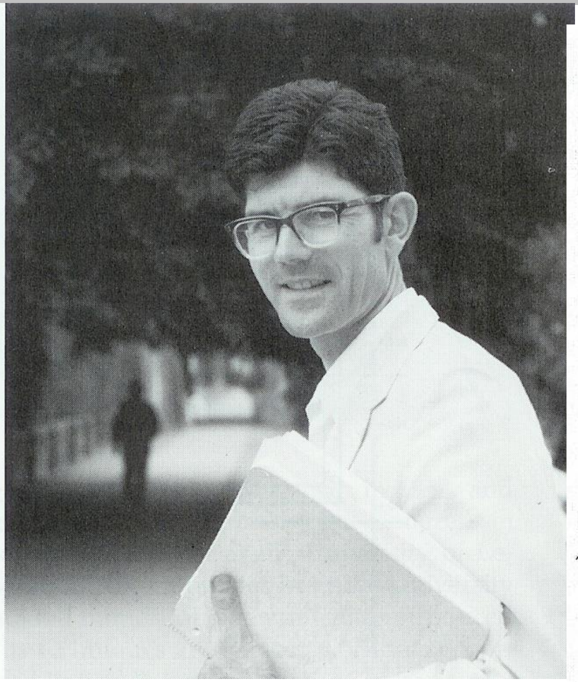


Willi Hennig

Phylogenetic Systematics 1966

Mas terá sido Cavalli Sforza e Edwards 1963…





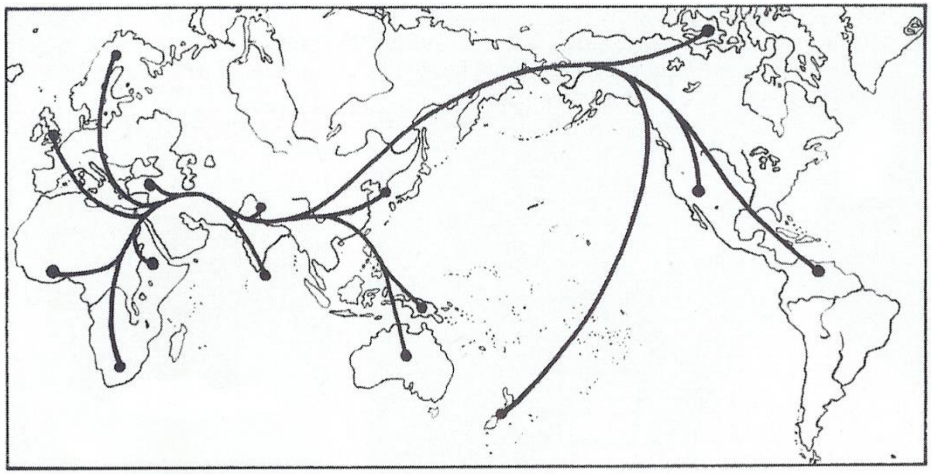


FIG. 1. Topology of the minimum-evolution tree uniting fifteen human populations; constructed on the basis of the frequency of blood-group alleles.

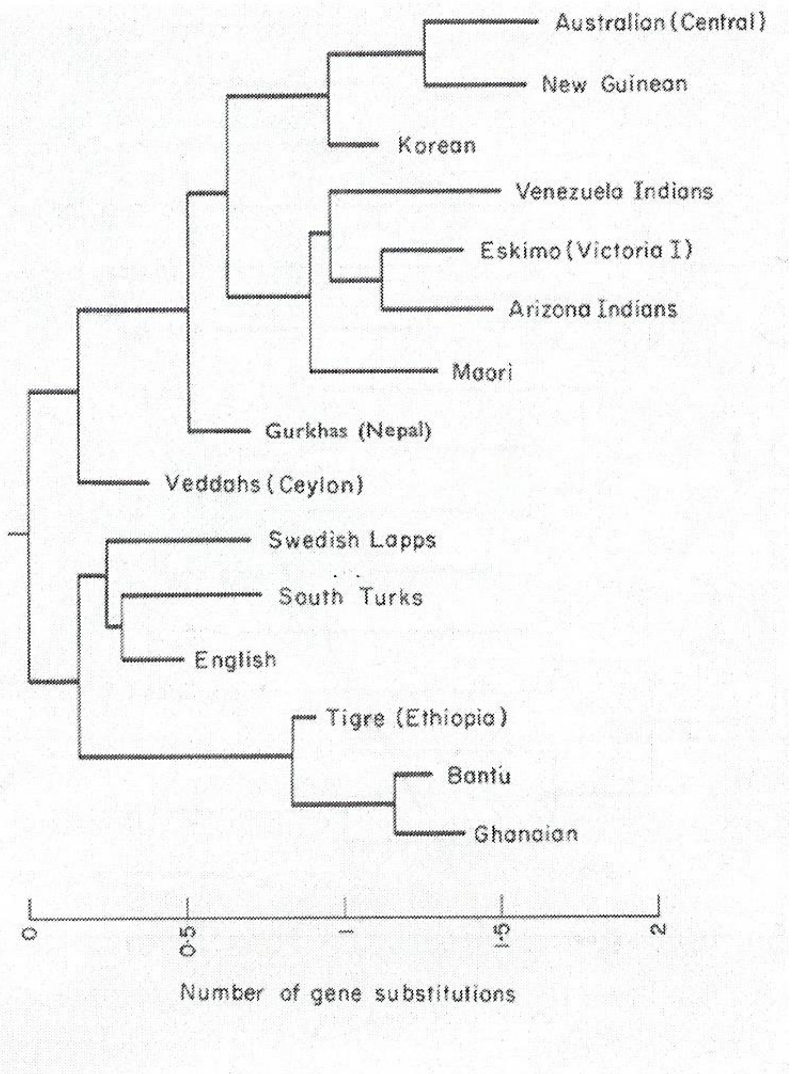
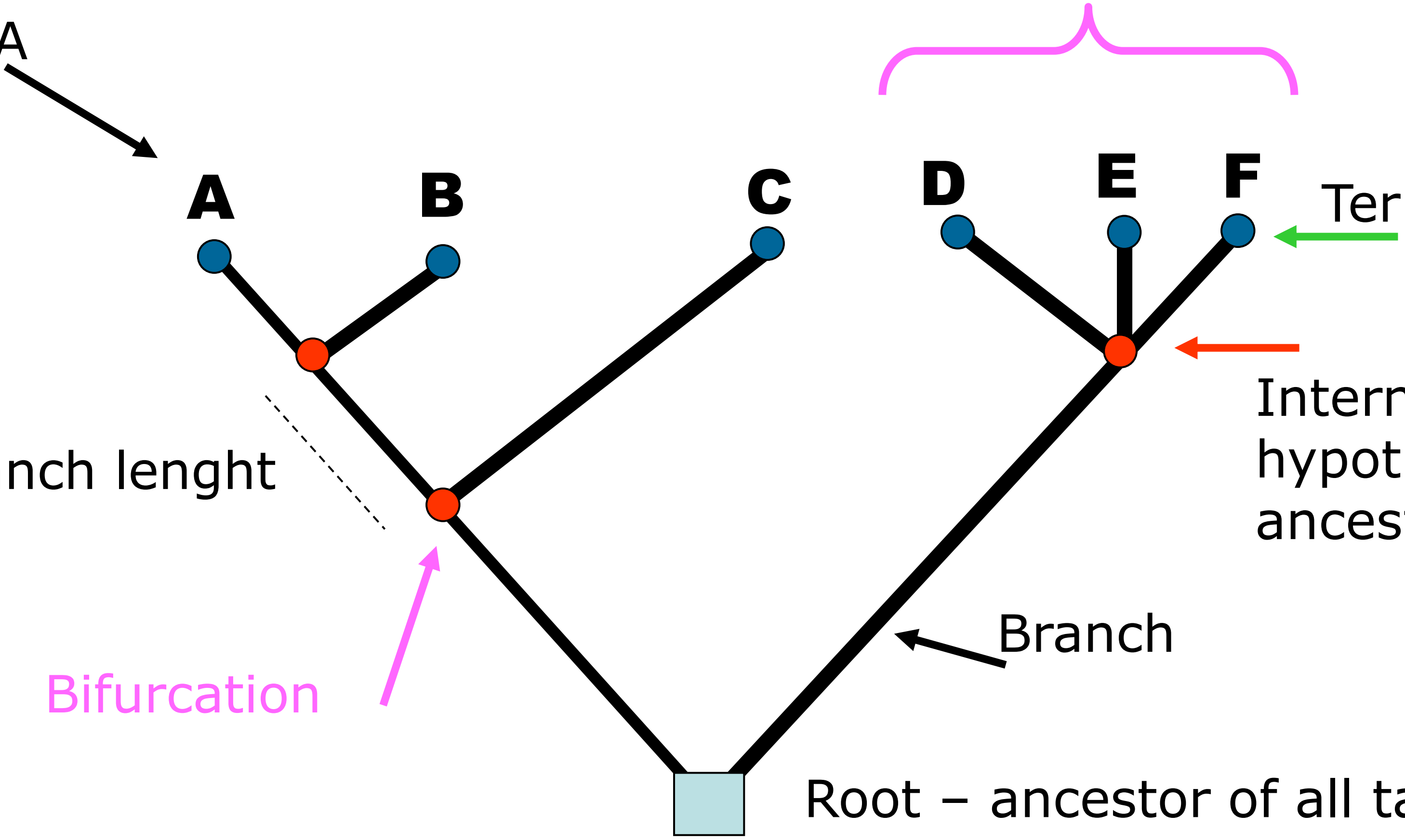


FIG. 5. Tree computed from blood group gene frequencies.

Filogenética – Noções basicas



Operational Taxonomic Unit OTU or TAXA

Polytomy or multifurcation

Terminal node
leaf

Internal node
hypothetical
ancestor

Branch lenght

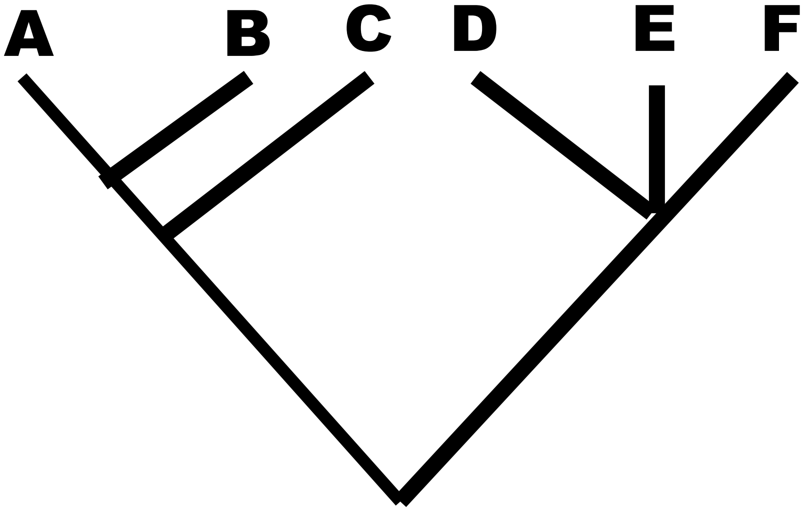
Bifurcation

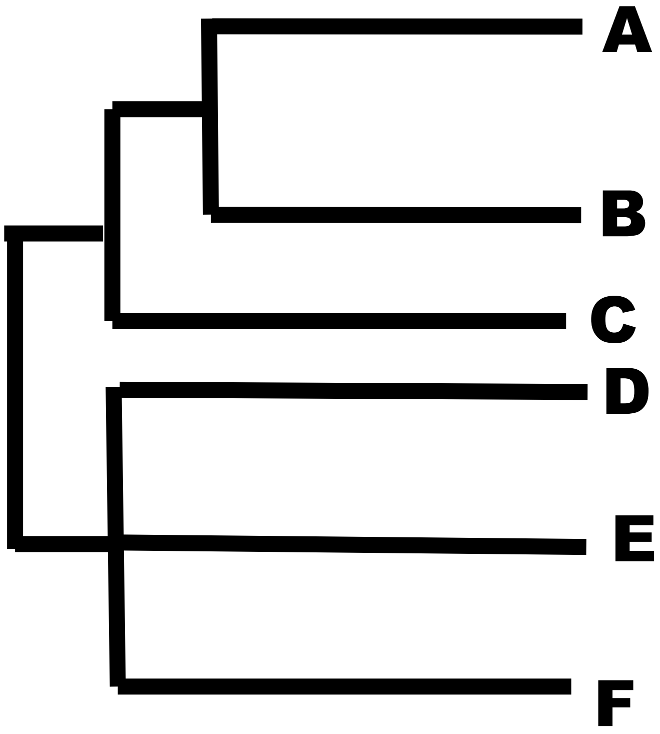
Branch

Root – ancestor of all taxa in tree

Trees as Mobiles

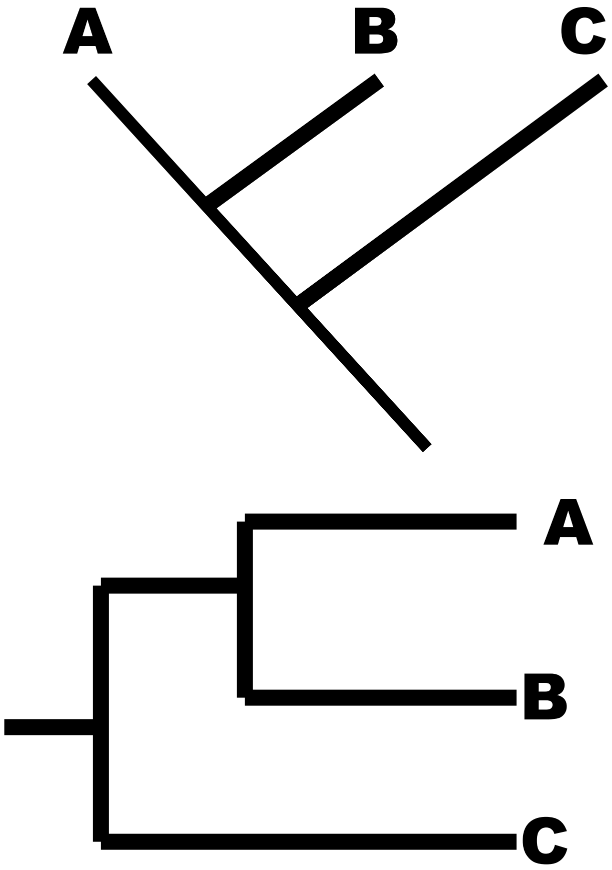
- Unresolved Polytomy (hard and soft)



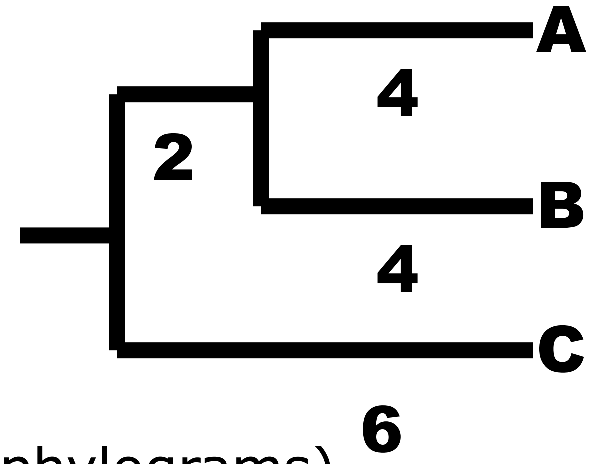


Arvóres filogenéticas

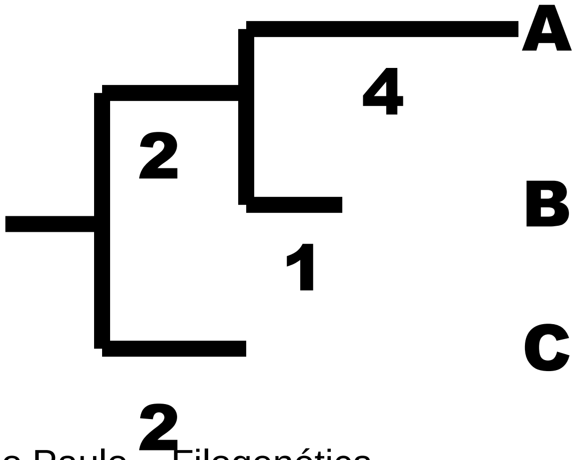
Cladograms



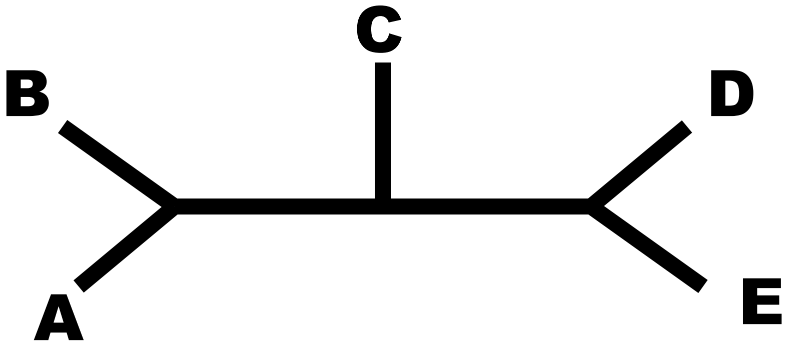
Ultrametric tree (dendograms)

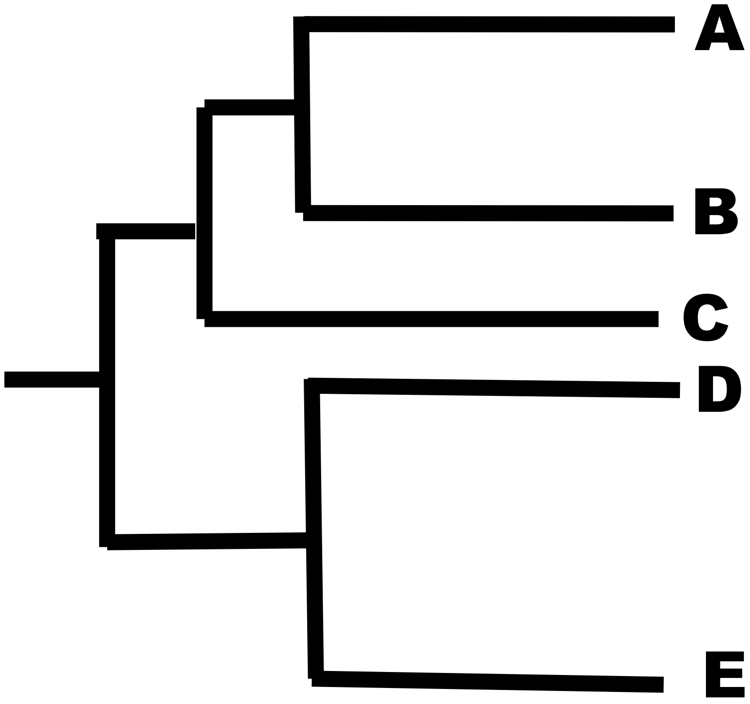


-Additive tree (phylograms)

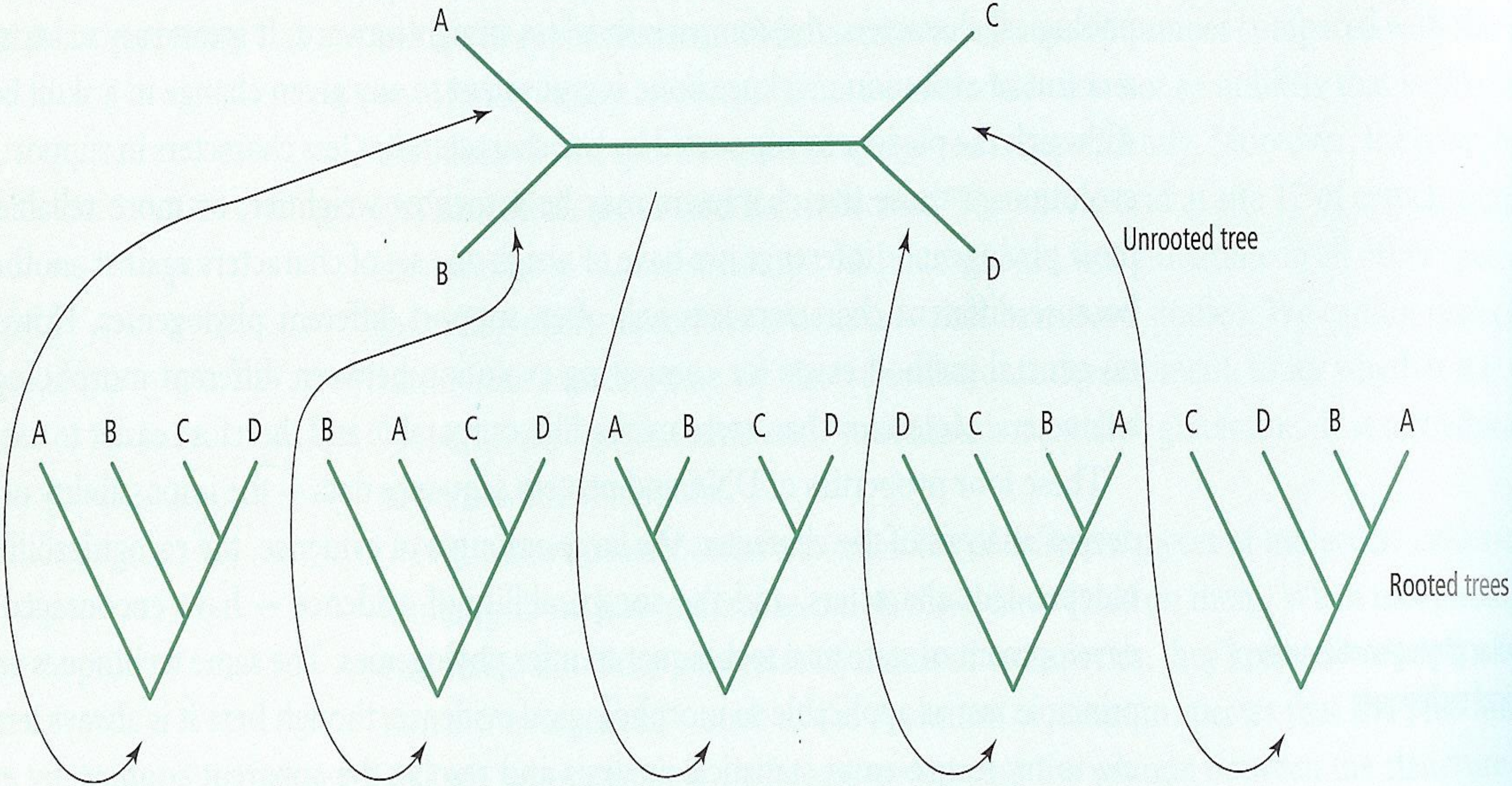


Arvóres filogenéticas – rooted e unrooted





Arvóres filogenéticas – rooted e unrooted



Arvóres filogenéticas – rooted e unrooted

Midpoint

Outgroup

Cladistic terminology and Concepts

-Monophyletic group or clade – an evolutionary assemblage that includes a common and all of its descendents.

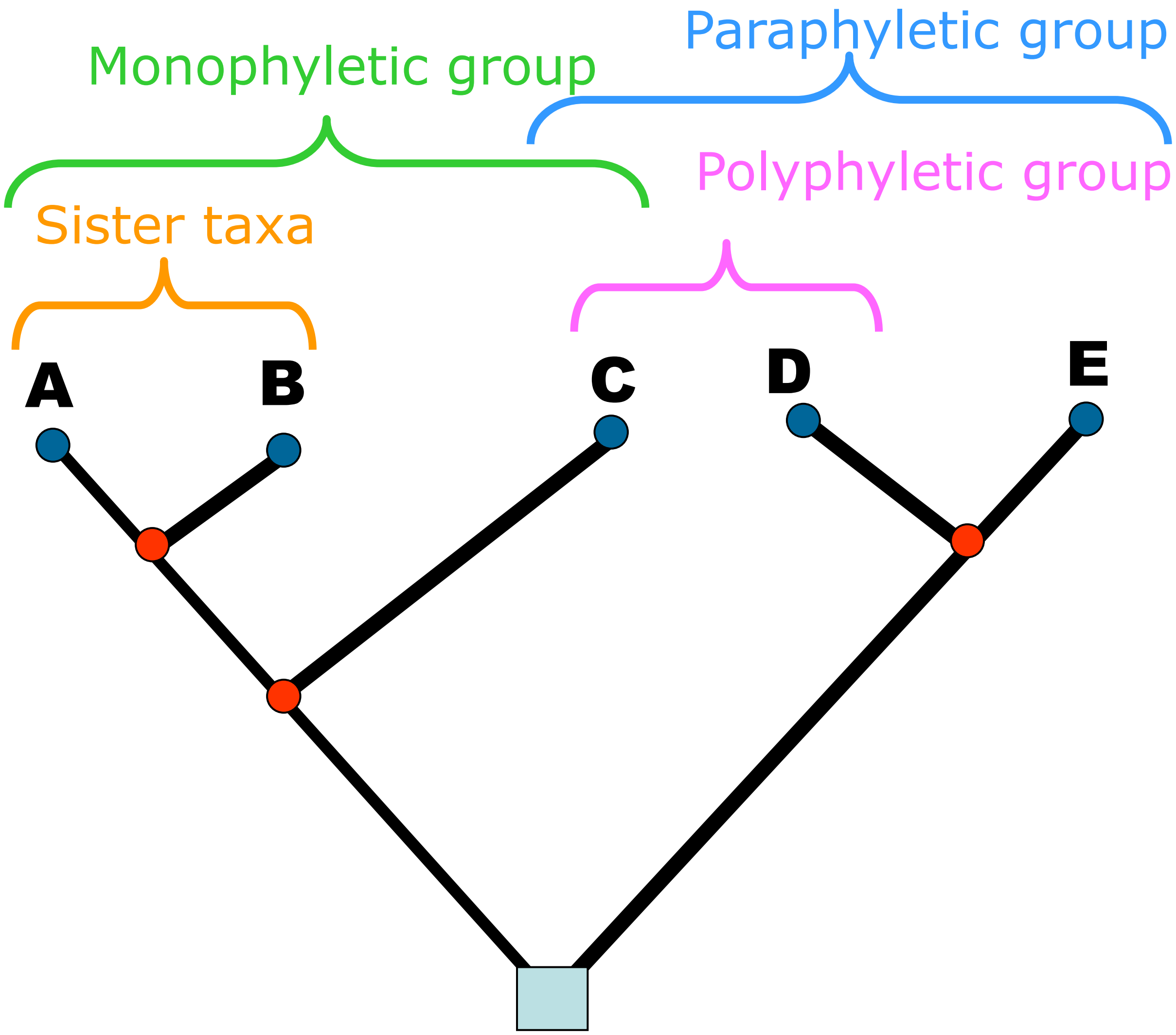
-Polyphyletic group – an artificial assemblage derived form two or more distinct ancestors.

-Paraphyletic group – an artificial assemblage that includes a common ancestor and some but not all of its descendents.

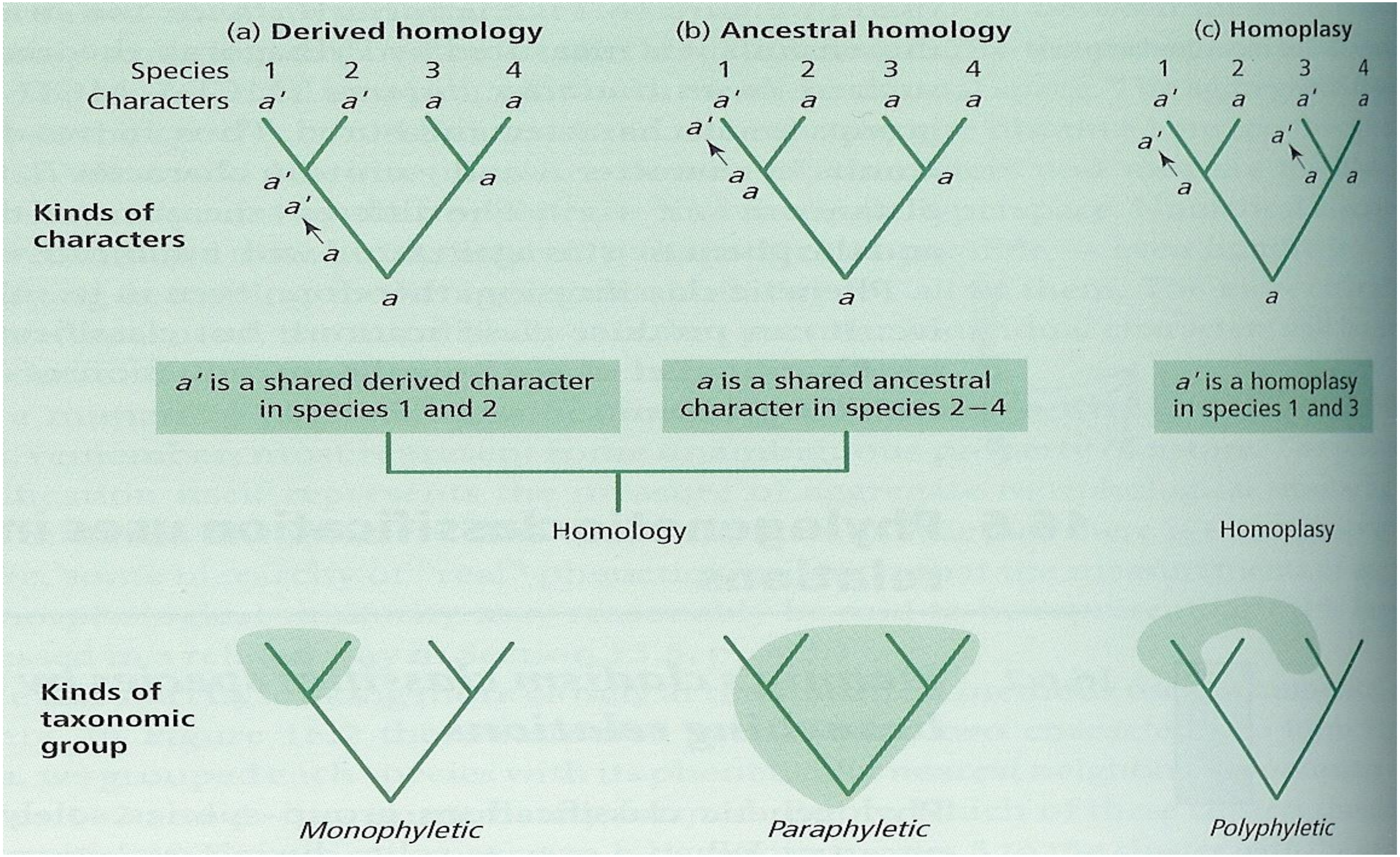
-Outgroup – a taxon phylogenetically outside the clade of interest.

-Sister taxa – a taxa steaming from the same node in a phylogeny

Filogenética – Noções basicas



Cladistic terminology and Concepts



Classes of characters

-PLESIOMORPHY – an ancestral character state (one present in the common ancestor of the taxa under study).

-APOMORPHY – a derived or newly evolved character state.

- SYMPLESIOMORPHY – an ancestral character state shared by two or more descendant taxa.

- SYNAPOMORPHY – a derived character state shared by two or more derived taxa.

- AUTOPOMORPHY – a character state unique to a single taxon.

Character change

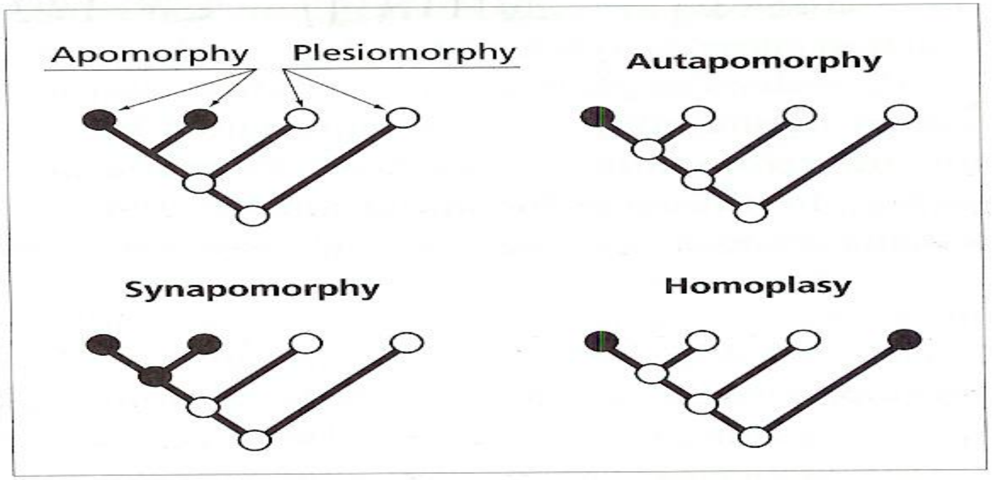
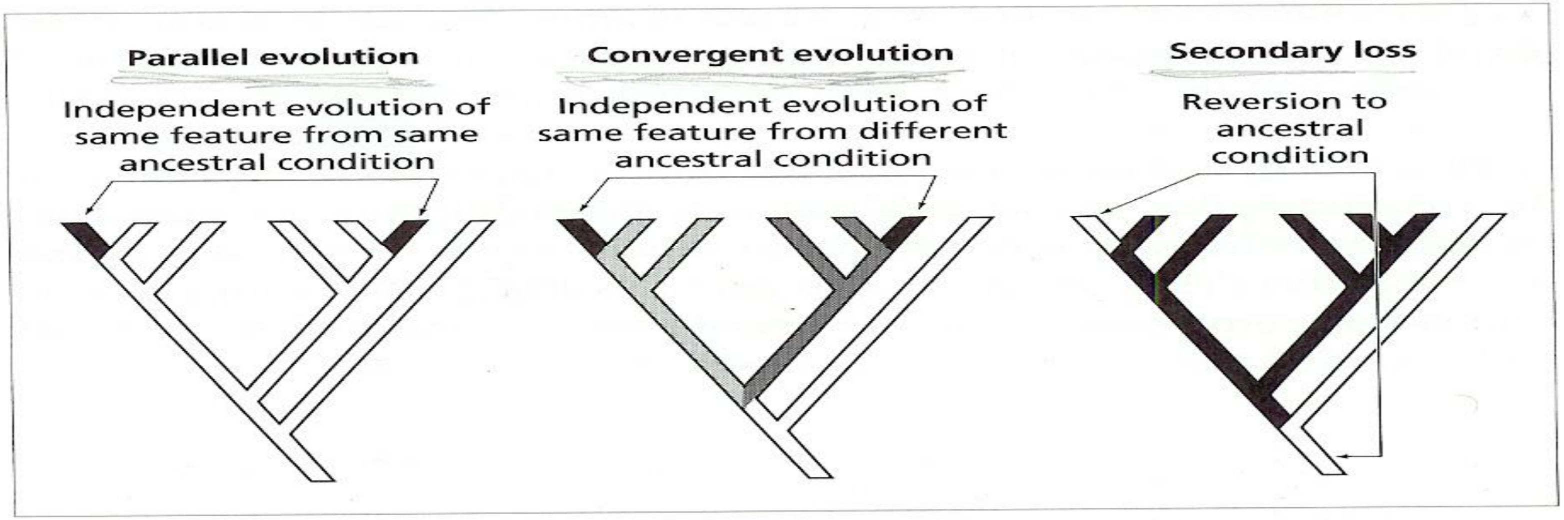


Fig. 2.11 Trees showing the terminology used to describe different patterns of ancestral (○) and derived (●) character states.



Character change

SYMPLESIOMORPHY – an ancestral character state shared by two or more descendant taxa.

SYNAPOMORPHY – a derived character state shared by two or more derived taxa.

AUTOPOMORPHY – a character state unique to a single taxon.

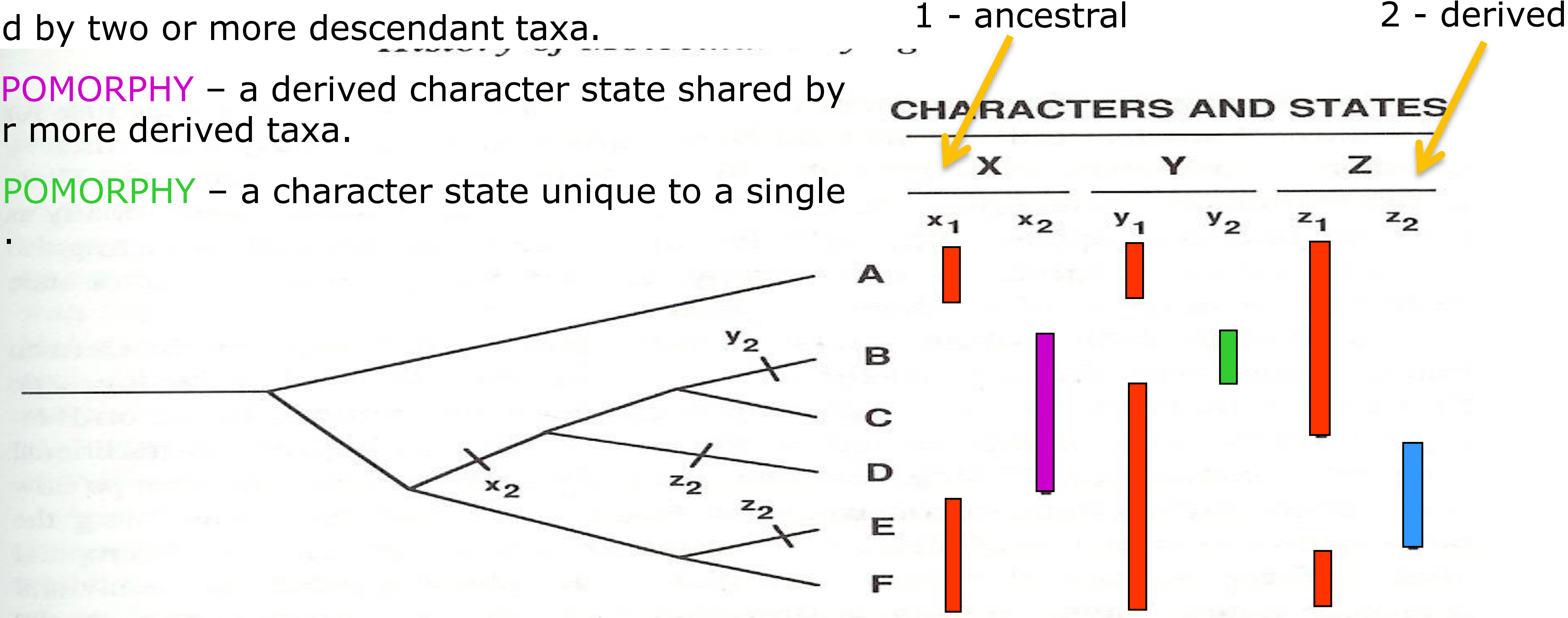
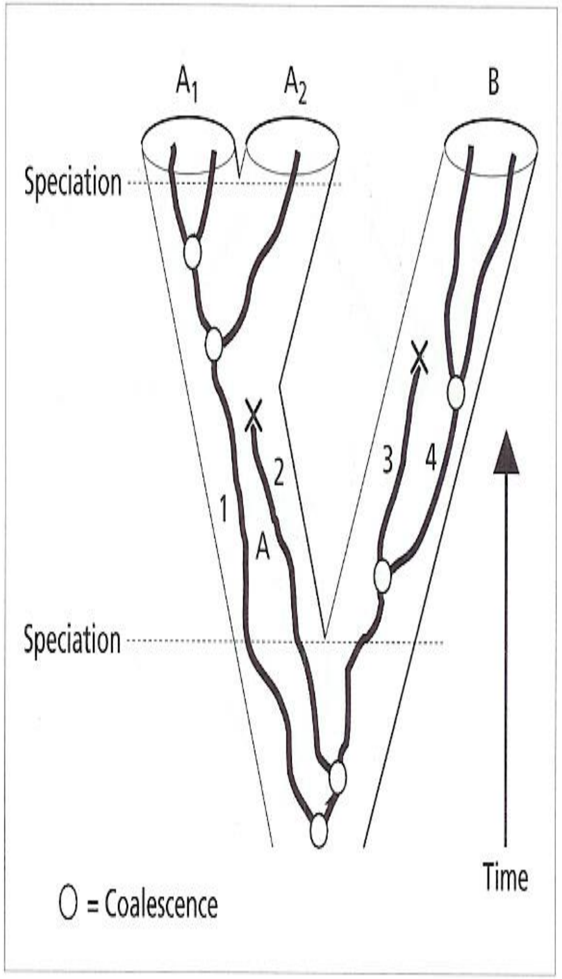
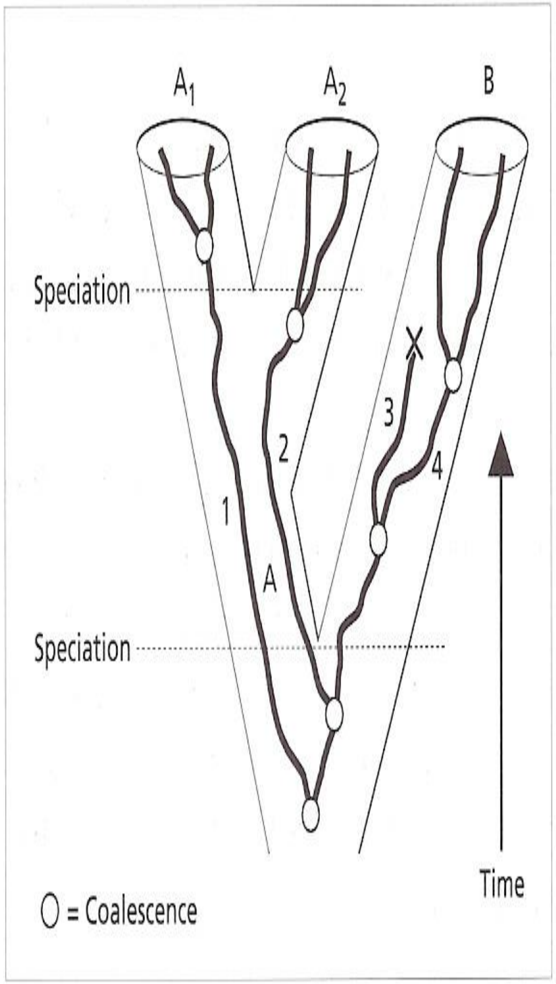


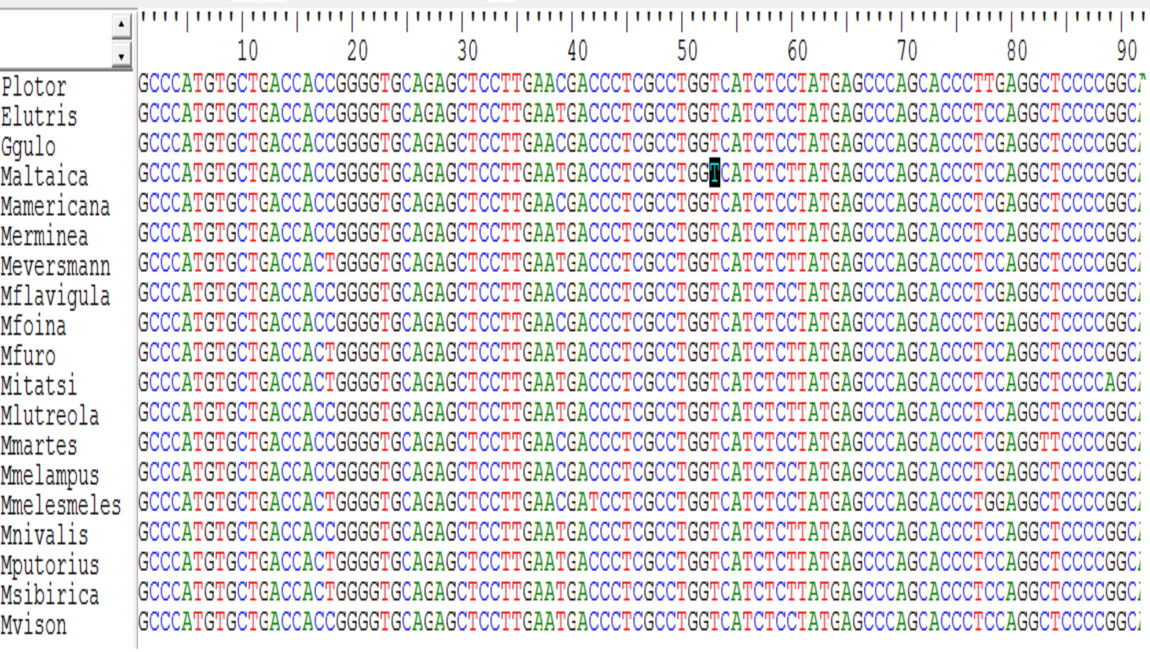
Figure 2.5. Philosophical rationale underlying Hennigian cladistic attempts to distinguish sources of similarity. Shown is the true (but unknown to the researcher) phylogeny for taxa A–F and the distribution of observed binary states for characters X, Y, and Z. Suppose that taxon A is known to be an outgroup for B–F, whose phylogeny is to be reconstructed. Character states x_1, y_1, and z_1, possessed by various taxa and the outgroup, are symplesiomorphs (shared ancestral states), and hence identify no clades. In particular, y_1 and z_1 could be positively misleading in amalgamating ingroup members (C–F and B, C, F, respectively) had their ancestral status gone unrecognized. Character state y_2 defines no multitaxon clade because it is an autapomorph, and z_2 could be misleading as a putative clade marker because it evolved in parallel (convergent) fashion in taxa D and E. Only x_2 is a valid synapomorph in this example, correctly identifying the true clade composed of taxa B–D.

Árvores de genes e de espécies podem contar histórias diferentes





Datasets and phylogenetics trees



```
Plotor      GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTTGAGGCTCCCCGGC
Elutris     GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Ggulo       GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGCTCCCCGGC
Maltaica    GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mamericana  GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGCTCCCCGGC
Merminea    GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Meversmann  GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mflavigula  GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGCTCCCCGGC
Mfoina      GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGCTCCCCGGC
Mfuro       GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mitatsi     GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCAGC
Mlutreola   GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mmartes     GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGTTCCCCGGC
Mmelampus   GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAACGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCGAGGCTCCCCGGC
Mmelesmeles GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAACGATCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTGGAGGCTCCCCGGC
Mnivalis    GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mputorius   GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Msibirica   GCCCATGTGCTGACCACTGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCTTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
Mvison      GCCCATGTGCTGACCACCGGGGTGCAGAGCTCCTTGAATGACCCTCGCCTGGTCATCTCCTATGAGCCCAGCACCCTCCAGGCTCCCCGGC
```