Lecture 9: Protein Architecture V: Evolution, Function & Classification

Margaret A. Daugherty
Fall 2003

Definitions

Biochemical evolution: study of how proteins (and other molecules) and biochemical pathways have changed through time.

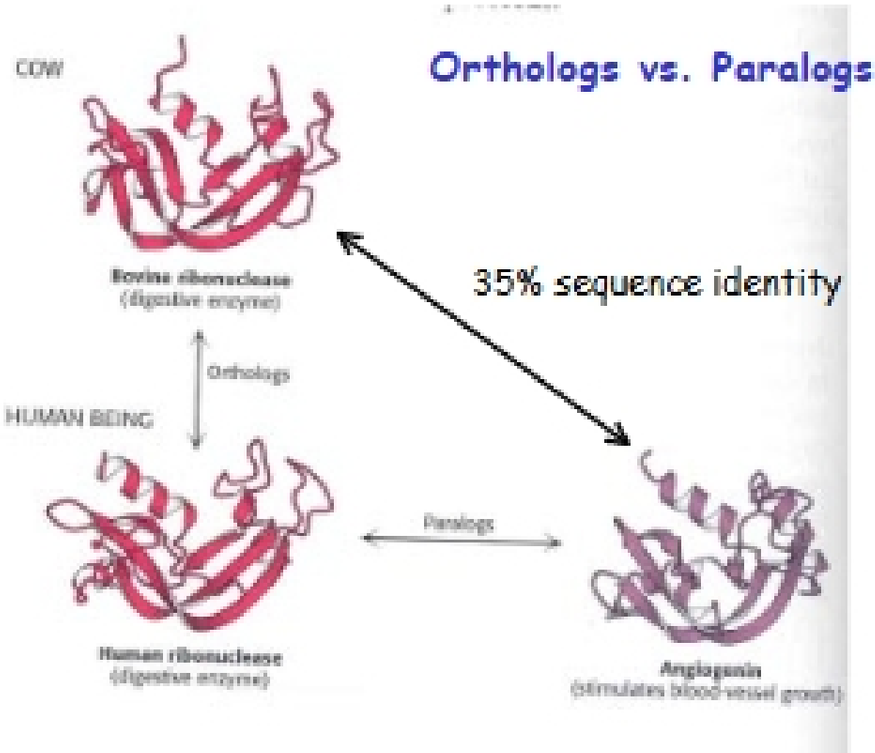
Homologous proteins: those derived from a common ancestor; share a significant degree of sequence similarity

i). **paralogs** - homologs present within a species;

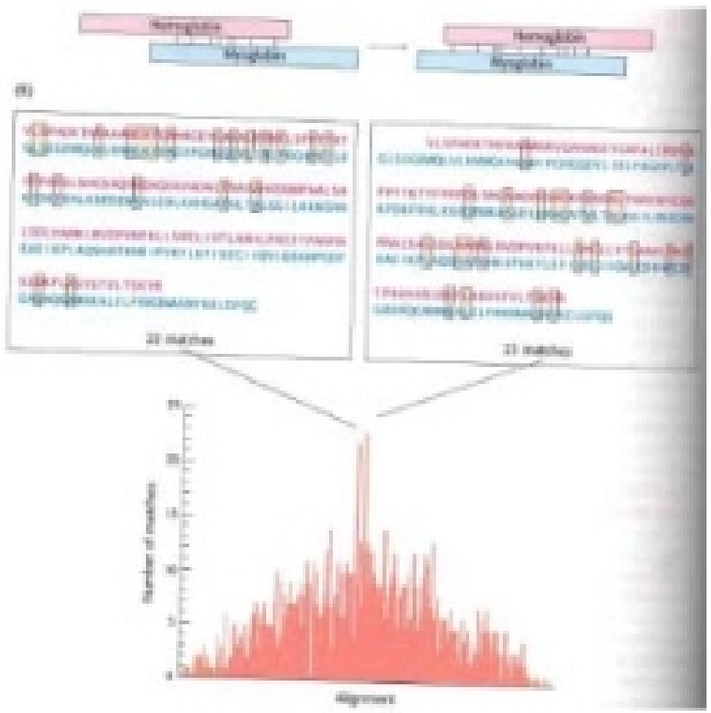
ii). **orthologs**: -homologs that are present in different species

Divergent evolution: Proteins derived from common ancestors

Orthologs vs. Paralogs



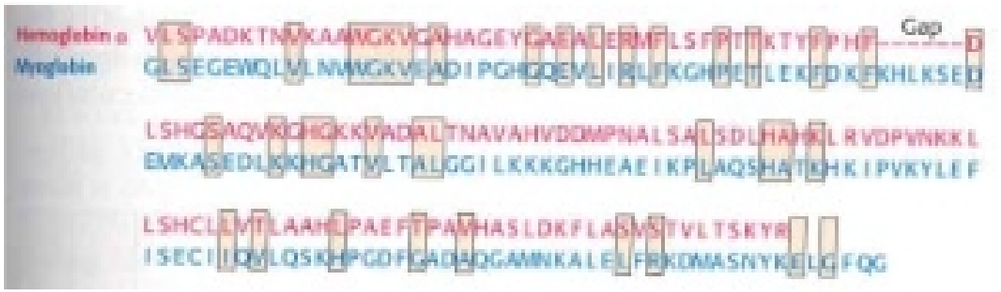
Sequence alignments: statistical survey of sequences



In this example we are looking at identical matches;

Slide one sequence along other and count up matches

Insertion of gaps helps alignments

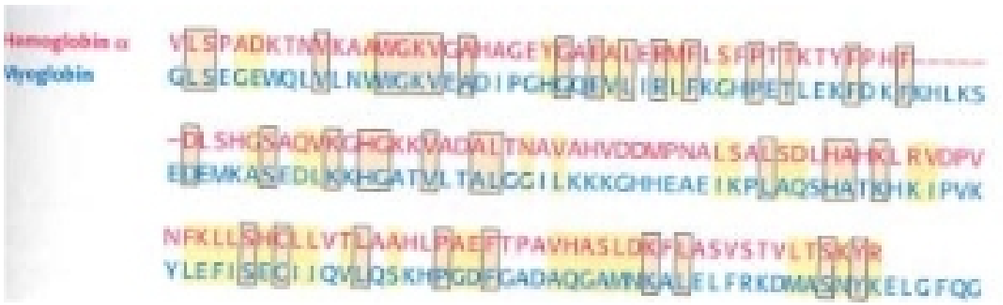


36 matches

Conservative substitutions

Replacing one amino acid by another similar in size & charge

e.g., asp -> glu *can you name other conservative changes?*



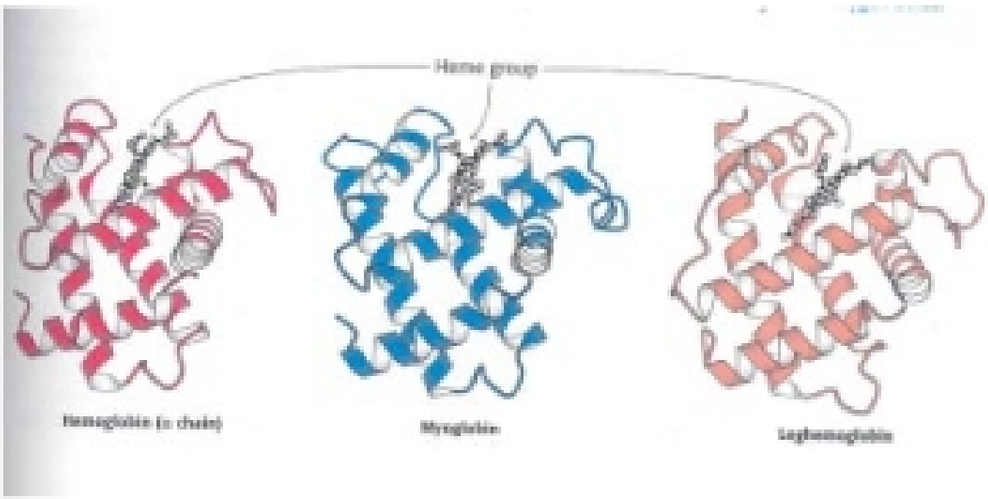
Rule of thumb: (assuming a 100 residue protein)

>25% identity : homologous sequence

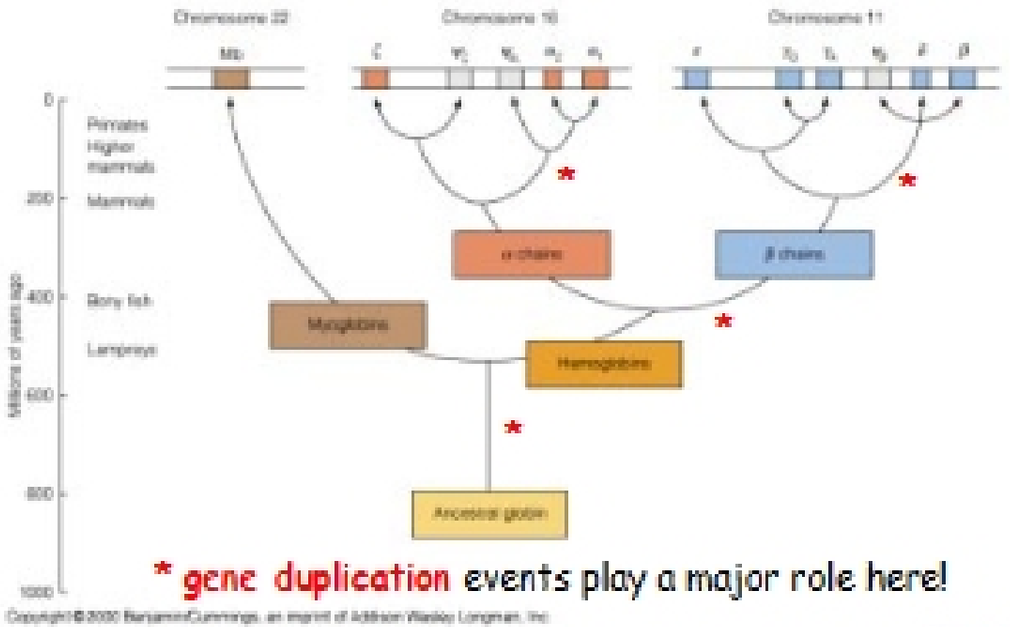
<15% identity: probably not homologous

Paralogs: the oxygen binding heme proteins

Tertiary structure is more conserved than primary structure

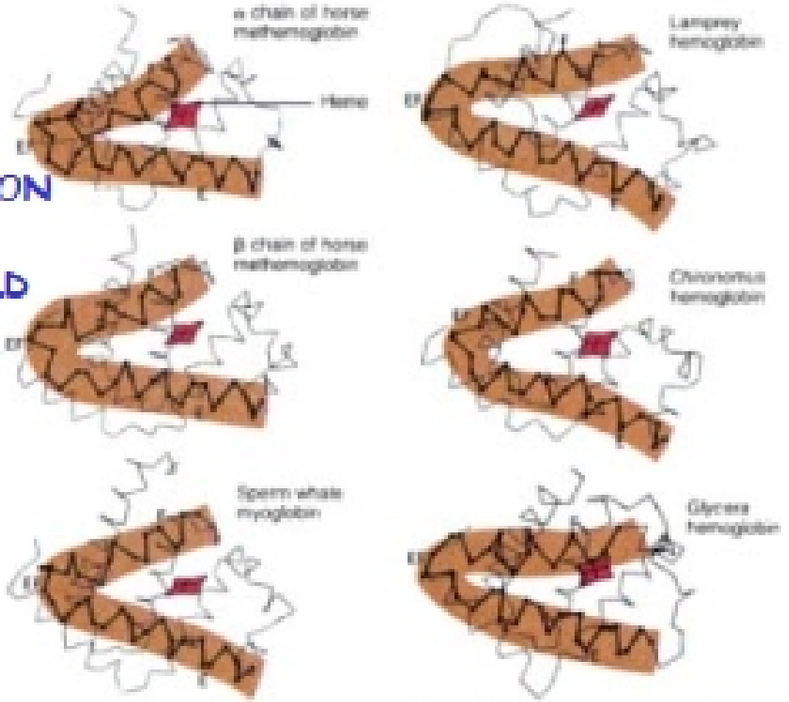


PHYLOGENETIC TREE FOR THE GLOBINS



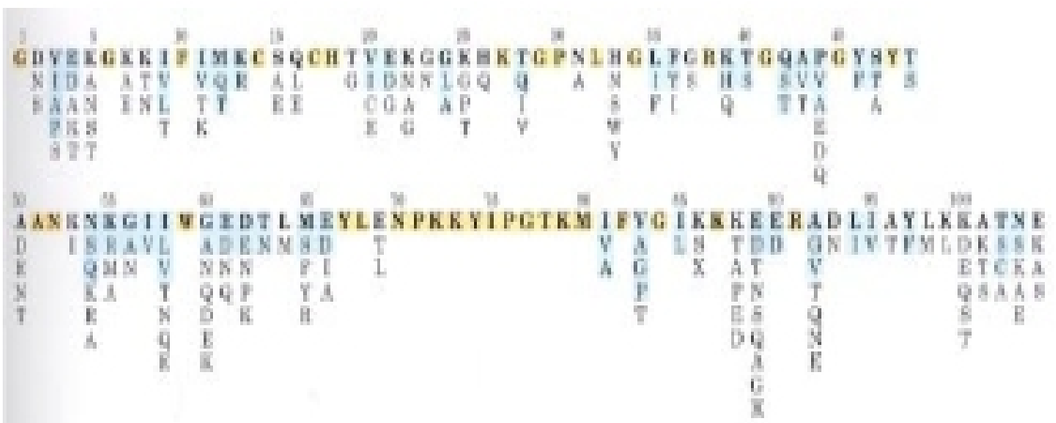
* gene duplication events play a major role here!

CONSERVATION IN THE GLOBIN FOLD



Sequence Alignment: Identification of homologous proteins

cytochrome c: a favorite example



Yellow - invariant; blue - conservative; unshaded; non-conserved

Comparison of sequence differences

	Chimpanzee	Sheep	Rattlesnake	Carp	Snail	Moth	Yeast	Cauliflower	Parsnip
Human	0	10	14	18	20	31	44	44	43
Chimpanzee		10	14	18	20	31	44	44	43
Sheep			20	11	24	27	44	46	46
Rattlesnake				26	28	33	47	45	43
Carp					26	26	44	47	46
Garden snail						28	48	51	50
Tobacco hornworm moth							44	44	41
Baker's yeast (iso-1)								47	47
Cauliflower									13

# of differences is proportional to the phylogenetic difference