© January 2025 | IJIRT | Volume 11 Issue 8 | ISSN: 2349-6002

Phylogenetic Analysis of Cytochrome b Protein from Different Chameleon Species

ATIF BAIG MIRZA[1], ABHIJIT HIRE[2], SANJAY SHELKE[3], SUVARNA PAWAR[4]

[1, 2, 3]*Student, Department of Biotechnology, Yashwantrao Chavan College, Sillod Dist. Aurangabad (M.H) India.*

[4]*Assistant Professor, Department of Biotechnology, Yashwantrao Chavan College, Sillod Dist. Aurangabad (M.H) India.*

Abstract— The major aim of this project is phylogenetic analysis of Cytochrome b protein from Chameleon species. The phylogenetic tree construction of different chameleon species and to find out the common structural and functional regions among them. We have taken 17 sequences of cytochrome b protein in different chameleon species. The cytochrome b is a common protein in all these 17 species. Cytochrome b is a Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-C1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis. The chameleon is a very distinctive and well-known species of lizard due to the large eyes and curled tail. Some species of chameleon, although not all, are able to change the colour of their skin in order to be camouflaged into their surroundings. Clustal Omega is a new multiple sequence alignment program that used seeded guide trees and HMM profile-profile techniques to generate alignments between 17 sequences. Mega6.0 is an advanced version of the Molecular Evolutionary Genetics Analysis (MEGA) software, which had shown the facilities for building sequence alignments, inferring phylogenetic histories, and conducting molecular evolutionary analysis. Mega6.0 is used to construct the phylogenetic tree. MEME discovered novel, ungapped motifs in the sequences. Prosite described protein domains, families and functional sites as well as associated patterns and profiles to identify them. This analysis will be helpful in drug discovery of particular disease or disorder causes by cytochrome b mutation.

Indexed Terms- Phylogenetic tree, Cytochrome b, Chameleon.

I. INTRODUCTION

Chameleons are reptiles that are part of the iguana suborder. These colorful lizards are known as one of the few animals that can change skin color. However, it is a misconception that chameleons change colors to match their surroundings [1].

Scientific name: Chamaeleonidae

Phylogenetic tree, also known as a phylogeny, is a tree diagram that depicts the lines of evolutionary descent of different species, organisms, or genes from a common ancestor. Phylogenetic analysis is sometimes regarded as being an intimidating, complex process that requires expertise and years of experience. In fact, it is a fairly straightforward process that can be learned quickly and applied effectively. This Protocol describes the several steps required to produce a phylogenetic tree from molecular data for inovices. In the example illustrated here, the program MEGA is used to implement all those steps, thereby eliminating the need to learn several programs, and to deal with multiple file formats from one step to another.[3]

Phylogenies are useful for organizing knowledge of biological diversity, for structural classifications, and for providing insight into events that occurred during evolution. Furthermore, because these trees show descent from a common ancestor, and because much of the strongest evidence for evolution comes in the form of common ancestry, one must understand phylogenies in order to fully appreciate the overwhelming evidence supporting the theory of evolution [4].

The pattern of branching in a phylogenetic tree reflects how species or other groups evolved from a series of common ancestors.

In trees, two species are more related if they have a more recent common ancestor and less related if they have a less recent common ancestor. Phylogenetic trees can be drawn in various equivalent styles. Rotating a tree about its branch points doesn't change the information it carries [5].

Cytochrome b protein plays a key role in structures called mitochondria, which convert the energy from food into a form that cells can use. Cytochrome b is one of 11 components of a group of proteins called as complex III. In mitochondria, complex III performs one step of a process known as oxidative phosphorylation, in which oxygen and simple sugars are used to create adenosine triphosphate (ATP), the cell's main energy source. During oxidative phosphorylation, the protein complexes, including complex III, drive the production of ATP through a step-by-step transfer of negatively charged particles called electrons. Cytochrome b is involved in the transfer of these particles through complex III [6]. Cytochrome b is commonly used as a region of mitochondrial DNA for determining phylogenetic relationships between organisms, due to its sequence variability. It is considered to be most useful in determining relationships within families and genera. Comparative studies involving cytochrome b have resulted in new classification schemes and have been used to assign newly described species to a genus as well as to deep in the understanding of evolutionary relationships [7].

Uniprot:

UniProt is an important collection of protein sequences and their annotations, which has doubled in size to 80 million sequences during the past year. This growth in sequences has prompted an extension of UniProt accession number space from 6 to 10 characters. An increasing fraction of new sequences are identical to a sequence that already exists in the database with the majority of sequences coming from genome sequencing projects.

They created a new proteome identifier that uniquely identifies a particular assembly of a species and strain or subspecies to help users track the provenance of sequences. They present a new website that has been designed using a user-experience design process. They have introduced an annotation score for all entries in UniProt to represent the relative amount of knowledge known about each protein. These scores will be helpful in identifying which proteins are the best characterized and most informative for comparative analysis [8]. All UniProt data is provided freely and is available on the web at (http://www.uniprot.org/). The mission of UniProt is to provide the scientific community with a comprehensive, high quality and freely accessible resource of protein sequence and functional information.

Clustal-Omega:

Clustal-Omega is a general-purpose multiple sequence alignment (MSA) program for proteins. It produces high quality MSAs and is capable of handling data-sets of hundreds of thousands of sequences in reasonable time. In default mode, users give a file of sequences to be aligned and these are clustered to produce a guide tree and this is used to guide a "progressive alignment" of the sequences. There are also facilities for aligning existing alignments to each other, aligning a sequence to an alignment and for using a hidden. This latter procedure is referred to as "external profile alignment" or EPA (computing.bio.cam.ac.uk/local/doc/clustalo.txt.).

Clustal Omega is a completely rewritten and revised version of the widely used Clustal series of programs for multiple sequence alignment. It can deal with very large numbers (many tens of thousands) of DNA/RNA or protein sequences due to its use of the mBED algorithm for calculating guide trees. This algorithm allows very large alignment problems to be tackled very quickly, even on personal computers. The accuracy of the program has been considerably improved over earlier Clustal programs [9].

MEGA6.0:

Molecular Evolutionary Genetics Analysis (MEGA) software, which currently contains facilities for building sequence alignments, inferring phylogenetic histories, and conducting molecular evolutionary analysis. In version 6.0, MEGA now enables the inference of time trees, as it implements the RelTime method for estimating divergence times for all branching points in a phylogeny. A new Time tree Wizard in MEGA6 facilitates this time tree inference by providing a graphical user interface (GUI) to specify the phylogeny and calibration constraints step-by-step. This version also contains enhanced algorithms to search for the optimal trees under evolutionary criteria and implements a more advanced memory management that can double the size of sequence data sets to which MEGA can be applied.

Both GUI and command-line versions of MEGA6 can be downloaded from www.megasoftware net free of charge [10].

An advanced version of the Molecular Evolutionary Genetics Analysis (MEGA) software,which currently contains facilities for building sequence alignments, inferring phylogenetic histories, and conducting molecular evolutionary analysis. In version 6.0, MEGA now enables the inference of timetrees, as it implements the RelTime method for estimating divergence times for all branching points in a phylogeny. A new Time tree Wizard in MEGA6 facilitates this time tree inference by providing a graphical user interface (GUI) to specify the phylogeny and calibration constraints step-by-step. This version also contains enhanced algorithms to search for the optimal trees under evolutionary criteria and implements a more advanced memory management that can double the size of sequence data sets to which MEGA can be applied. Both GUI and command-line versions of MEGA6 can be downloaded from www. megasoftware.net free of charge [11].

MEME:
The MEME Suite is a software toolkit with a unified web server interface that enables users to perform four types of motif analysis: motif discovery, motif–motif database searching, motif-sequence database searching and assignment of function. It offers a significantly expanded set of programs for these tasks compared with the earlier web server. MEME is tool for motif discovery.

MEME (Multiple EM for Motif Elicitation) is one of the most widely used tools for searching for novel ‘signals’ in sets of biological sequences. Applications include the discovery of new transcription factor binding sites and protein domains. MEME works by searching for repeated, ungapped sequence patterns that occur in the DNA or protein sequences provided by the user. Users can perform MEME searches via the web server hosted by the National Biomedical Computation Resource (http://meme.nbcr.net) and several mirror sites. Through the same web server, users can also access the Motif Alignment and Search Tool to search sequence databases for matches to motifs encoded in several popular formats. By clicking on buttons in the MEME output, users can compare the motifs discovered in their input sequences with databases of known motifs, search sequence databases for matches to the motifs and display the motifs in various formats. This article describes the freely accessible web server and its architecture, and discusses ways to use MEME effectively to find new sequence patterns in biological sequences and analyse their significance [15].

Prosite:
PROSITE consists of documentation entries describing protein domains, families and functional sites, as well as associated patterns and profiles to identify them. It is complemented by ProRule, a collection of rules based on profiles and patterns, which increases the discriminatory power of these profiles and patterns by providing additional information about functionally and/or structurally critical amino acids.

PROSITE is largely used for the annotation of domain features of UniProtKB/ Swiss-Prot entries. In order to allow better functional characterization of domains, PROSITE developments focus on subfamily specific profiles and a new profile building method giving more weight to functionally important residues [16].
PROSITE is accessible at: (http://www.expasy.org/prosite*/.)*

II. MATERIAL AND METHODS

I. Retrieval of sequence.

we have taken 17 sequences of cytochrome b protein from different chameleon species.

The sequences of cytochrome b proteins retrieved from uniprot database; it is available on web at *www.uniprot.org.*

II. Multiple Sequence Alignment.

By using clustal omega tool, we have done multiple sequence alignment of different chameleon species of cytochrome b proteins. It helps to find out the sequence similarity in different sequences. We got a tree in statistical format, it saves as *“tree.nwk”* format on notepad.

III. Construction of phylogenetic tree.

Mega6.0 is software for phylogenetic tree construction. It is downloaded from www.megasoftware.net

IV. Motif identification.

Motif scanning means finding all known motifs that occur in a sequence. MEME tool allows to discover novel motifs in collections of unaligned protein sequence. It is available on the web at *meme-suite.org/*

V. Domain identification.

PROSITE consists of documentation entries describing protein domains, families and functional sites as well as associated patterns and profiles to identify them. This tool is available on web at *http://prosite.expasy.org/*

III. RESULT AND DISCUSSION

I. Retrieval Sequences of 17 chameleon species from Uniprot database:

American_chameleon (H9GBQ4)
MGRIEWAMWANEQALASGLILVAGGIVAVSG
QFKRWEFAAYAIAAGVFICLLEYPRGKRKKGS
TMQRCGQQYFTPVVKVFGPLTRNYYVRSILHA
CLAVPAGFLLATILGTVCLAIASGIYLLAATRGE
EWQPIESKPKERPQVGGTIKHPPTNPPPRPPPDA
RKKPPEDGAGGQENPIAVEVEAE

Flapneck Chameleon (B7S688)
MTITRKSNPALKIINHSLIDLPTPTNISTAWNFGSLL
GLMLITQITTGLFLAMHYTADINLAFSSVAHICREV
NFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYL
YKETWNIGIILLFLTMATAFLGYILPWGQMSFWGA
TVITNMLSAIPYIGSNLVNWIWGGFSVNNPTLVRFF
TLHFLTPFLIAGATMLHLLFLHETGSNNPTGLSSN
SDKVSFHPYFPLKDLTTATTTMTLLTLIALFLPTLL
TDPQNFSPANPLSTPPHIMPEWYFLFAYTILRSIPN
KLGGVMALAMSILILVLMPLLHTSKQRTMKFRPTS
QLVFWTMTANIMILTWIGSQPVEHPLTMIGQIATV
MYFLIVLAIMPILASMENKTL

Fischers_Chameleon (B7S675)
MTALRKHYPPLKTINHSLIDLPTPTNISMAWNL
GSLLGLMLMMQILTGLFLTMHYATDTHLAFSS
VSTICREVNYGWLMRNMHANGASMFFICIFLHI
GRGLFYSSYLFKETWIIGVILLLLTMITAFLGYIL
PWGQMSFWGATVITNMLSAIPLIGNDLVNWIW
GGFSVGNPTLTRFFALHFMMPFLIAGTTTLHLL
FLHETGSSNPTGLSSNPDKVLFHPYFTLKDLTT
ALTATTLLMLVALFLPTLLTDPQNFSPANPLMT
PPHIMPEWYFLFAYTILRSIPNKLGGVMALFAA
VSILLLVPMLHTSKQRSMKFRPTSQTIFWIMTA
NILILTWAGSQPVEQPIITIGQVATTAYFLIFIFIT
PIIATLENKYHDMH

Awash_chameleon (B7S6G6)
MTITRKSNPALKIINHSLIDLPTPTNISTAWNFGS
LLGLMLVTQITTGLFLAMHYTADINLAFSSIAHI
SREVNFGWLMRNLHANGASMFFICIYLHIGRG
LYYSSYLYKETWNIGIILLFLTMATAFLGYILPW
GQMSFWGATVITNMLSAIPYIGNNLVNWIWGG
FSVNNPTLIRFFTLHFLTPFLIAGTTTLHLLFLHE
TGSNNPTGLSSNSDKVSFHPYFPLKDLTAATMT
VTLLTLIALFLPTLLTDPQNFSPANPLSTPPHIMP
EWYFLFAYTILRSIPNKLGGVMALLMSILILILM
PLLHTSKQRTMKFRPFSQLMFWTMAANIMILT
WIGSQPVEQPLTTIGQIATALYFLIIVVITPTLAL
MENKTL

Mediterranean_Chameleon (B7S6J2)
MTITRKSNPALKIINHSLIDLPTPTNISTAWNFGS
LLGLMLVTQITTGLFLAMHYTADINLAFSSIAHI
SREVNFGWLMRNLHANGASMFFICIYLHIGRG
LYYSSYLYKETWNIGIILLFLTMATAFLGYILPW
GQMSFWGATVITNMLSAIPYIGNNLVNWIWGG
FSVNNPTLVRFFTLHFLTPFLIAGTTMLHLLFLH
ETGSNNPTGLSSNSDKVSFHPYFPLKDLTTATM
TVTLLTLIALFLPTLLTDPQNFSPANPLSTPPHIM
PEWYFLFAYTILRSIPNKLGGVMALLMSILILIL
MPLLHTSKQRTMKFRPLSQLMFWTMTANIMIL
TWIGSQPVEQPLTTIGQIATTLYFLIIIMITPTLAL
MENKTL

Indian_Chameleon (B7S6B4)
MTITRKSNPALKIINHSLIDLPTPTNISTAWNFGS
LLGLMLVTQITTGLFLAMHYTANIDLAFSSIAHI
SREVNFGWLMRNLHANGASMFFICIYLHIGRG
LYYSSYLHKETWNIGIILLLMTMATAFLGYILP
WGQMSFWGATVITNMLSAIPYIGNNLVNWIW
GGFSVNNPTLIRFFTLHFLTPFLIAGTTMLHLLF
LHETGSNNPTGLSSNSDKVSFHPYFPLKDLTTA
TTTMTLLTLIALFMPTLLTDPQNFSPANPLSTPP
HIMPEWYFLFAYTILRSIPNKLGGVMALFMSIL
VLILMPILHTSKQRTMKFRPLSQLMFWTMTTNI
LILTWVGSQPVEQPLTTIGQISTTLYFLIIIMLMP
TLALMENKTL

African_Chameleon (B8LG21)
MTITRKANPALKIINHSLIDLPTPTNISTAWNFG
SLLGLMLVTQITTGLFLAMHYTADINLAFSSIA
HISREVNFGWLMRNLHANGASMFFICIYLHIGR

GLYYSSYLYKETWNIGIILLFLTMATAFLGYILP
WGQMSFWGATVITNMLSAIPYIGNNLVSWIWG
GFSVNNPTLIRFFTLHFLTPFLIAGTTMLHLLFL
HETGSSNPTGLSSNSDKVPFHPYFPLKDLTTAT
MTLTLLTLIALFLPTLLTDPQNFSPANPLSTPPHI
MPEWYFLFAYTILRSIPNKLGGVMALLMSILILI
LMPLLHTSKQRTMKFRPFSQLMFWTMTANIMI
LTWIGSQPVEQPLTTIGQISTALYFLIIIMITPTLA
LMENKTL

Veild_Chameleon (B7S6C7)
MTIARKSNPALKIINHSLIDLPTPTNISTAWNFGS
LLGLMLVTQITTGLFLAMHYTADINLAFSSIAHI
SREVNFGWLMRNLHANGASMFFICIYLHIGRG
LYYSSYLYKETWNIGIMLLFLTMATAFLGYILP
WGQMSFWGATVITNMLSAIPYIGNNLVNWIW
GGFSVNNPTLVRFFTLHFLTPFLIAGTTMLHLLF
LHETGSNNPTGLSSNSDKVSFHPYFPLKDLTTA
TTTITLLTLIALFLPTLLTDPQNFSPANPLSTPPHI
MPEWYFLFAYTILRSIPNKLGGVMALFMSILILI
LMPLLHTSKQRTMKFRPFSQLMFWTMTANILIL
TWIGSQPVEQPLTTIGQISTTLYFLIIIMFTPTLAL
MENK

Arabian_Chameleon (B7S6E0)
MTIARKSNPALKIINHSLIDLPTPTNISTAWNFGS
LLGLMLVTQITTGLFLAMHYTADINLAFSSIAHI
SREVNFGWLMRNLHANGASMFFICIYLHIGRG
LYYSSYLYKETWNIGIMLLFLTMATAFLGYILP
WGQMSFWGATVITNMLSAIPYIGNNLVNWIW
GGFSVNNPTLIRFFTLHFLTPFLIAGTTMLHLLF
LHETGSNNPTGLSSNSDKVSFHPYFPLKDLTTA
TTTITLLTLIALFLPTLLTDPQNFSPANPLSTPPHI
MPEWYFLFAYTILRSIPNKLGGVMALFMSILILI
LMPLLHTSKQRTMKFRPFSQLMFWTMTANILIL
TWIGSQPVEQPLTTIGQISTTLYFLIIIMLTPTLAL
MENKTL

Panthar_Chameleon (A0G2SPK)
MTITRKHHPMLKIINHSLIDLPTPTNITTAWNFG
SLIGLMLMVQITTGLLLATHYTADVNMAFISVS
HICREVNFGWLTRNLHANGASMFFICIYLHIGR
GLYYSSFLHKETWNIGIILLFMTMATAFLGYILP
WGQMSFWGATVITNMMSAIPYIGNNLVSWIW
GGFSINNPTLIRFFTAHFLLPFIIAGATMLHLLFL
HETGSSNPTGMSSNLDKVTFHPYFPLKDLTTML
MTITLLTLTATLSPTFFMDPENFLPANPLTTPPHI
MPEWYFLFAYAILRSIPNKLGG

Treefrog_Chameleon (Q29YK6)
LLFLVMATAFVGYVLPWGQMSFWGATVITNL
LSAAPYIGTELVQWIWGGFSVDNATLTRFFTFH
FILPFIIAGASMIHLLFLHQTGSSNPTGLNSNPDK
IPFHAYYSYKDAFGFALLLALLAALSTFAPNILG
DPDNFTPANPLVTPPHIKPEWYFLFAYAILRSIP
NKLGGVLALLFSIMILFLMPILHTSNQRTTAFRP
LAKLLFWTLVANTMILTWIGGQP

Spiny_Chameleon (Q29YK6)
LLFLVMATAFVGYVLPWGQMSFWGATVITNL
LSAAPYIGTELVQWIWGGFSVDNATLTRFFTFH
FILPFIIAGASMIHLLFLHQTGSSNPTGLNSNPDK
IPFHAYYSYKDAFGFALLLALLAALSTFAPNILG
DPDNFTPANPLVTPPHIKPEWYFLFAYAILRSIP
NKLGGVLALLFSIMILFLMPILHTSNQRTTAFRP
LAKLLFWTLVANTMILTWIGGQP

Parson_Chameleon (D6RRZ5)
MTITRKSHPVLKTISHSLIDLPTPTNISTAWNFGS
LLGLMLVTQIVTGLFLAMHYTADINLAFNSVSL
ICREVNFGWLMRSLHANGASMFFICIYLHIGRG
LYYSSYLYKETWNVGIILLLLTMATAFLGYVLP
WGQMSFWGATVITNMLSAIPYIGNNLVNWIW
GGFSVNNPTLIRFFTLHFLMPFLIAGTTMIHLLF
LHETGSSNPTGLSSNPDKITFHPYFPLKDLCSAL
TTITLLTLITLFLPTLLTDPQNFSPANPLATPPHI
MPEWYFLFAYTILRSIPNKLGGVVALLMSIFILI
LIPLFHTSKQRTMKFRPTSQLMFWTMTTNILILT
WIGGQPVEQPFTLIGQTSTVIYFLIMTTITPTIAT
LENKAL

Parrotfish_Chameleon (B6CYI0)
LVDLPAPSNISVWWNFGSLLGLCLASQILTGLF
LAMHYTSDIATAFSSVAHICRDVNYGWLIRNM
HANGASFFFICIYLHIGRGLYYGSYLYKETWNT
GVILLLLVMMTAFVGYVLPWGQMSFWGATVI
TNLLSAVPYVGNTLVQWIWGGFSVDNATLT

Brown_Chameleon (B7S662)
MTVMRKSNPILKTINHSLIDLPTPINISTAWNFG
SLLGLLLISQIITGLFLAMHYTADIKLAFNSVSHI
CREVNLGWLLRNLHANGASMFFICIYLHIGRGI
YYASYLYKETWNTGIALLFLTMATAFLGYIL

Gecko_Chameleon (Q9XLH6)
IINNSFIDLPTPPNISAWWNFGSLLGICLIIQITTG
LFLAMHYTADTSLAFSSISHICRNVNYGWLIRN

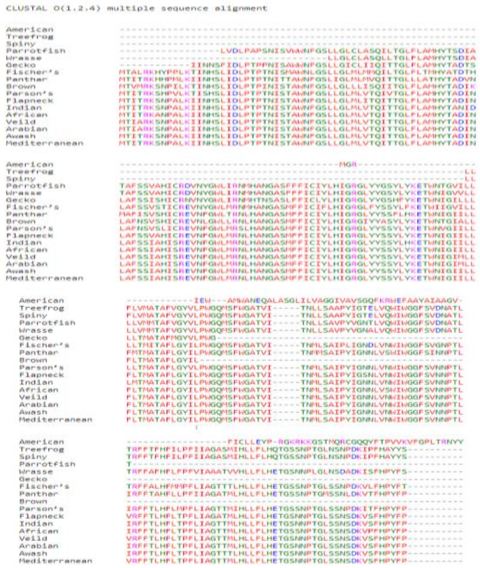
MHTNSASLFFICLYLHIGRGLYYGSHFYKETWN
IGILLLLLTMATAFMGYVLPWG

Wrasse_Chameleon (A0A1W6R997)
LLGLCLASQLLTGLFLAMHYTSDIATAFSSVAH
ICRDVNYGWLIRNMHANGASFFFICIYLHIGRG
LYYGSYLYKETWNIGVVLLLLVMMTAFVGYV
LPWGQMSFWGATVITNLLSAVPYVGNALVQW
IWGGFSVDNATLTRFFAFHFLFPFVIAAATVVH
LLFLHETGSNNPLGLNSDADKISFHPYFSYKDL
LGFAGLLIALTSLALFAPNL

II. Multiple sequence alignment

Multiple sequence alignment is comparison of more than two sequences. By using Clustal-Omega tool we perform multiple sequence alignment of 17 chameleon species on the basis of cytochrome b protein.

Fig.no.1 Multiple sequence alignment result



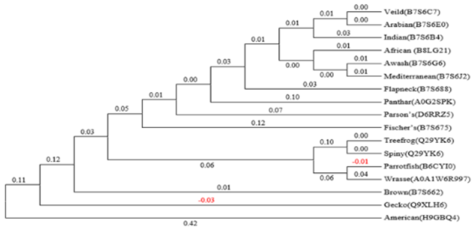
CLUSTAL O(1.2.4) multiple sequence alignment

- In above finding, there are maximum sequence similarities in Parrotfish chameleon, Wrasse chameleon, Gecko chameleon, Fischer's chameleon, Panthar chameleon, Brown chameleon, Parson's chameleon, Flapneck Chameleon, Indian Chameleon, African Chameleon, Veild Chameleon, Arabian Chameleon, Awash Chameleon and Mediterranean Chameleon.
- There are no sequence similarities has been found in American chameleon, Treefrog Chameleon and Spiny chameleon with other chameleons. These are distantly related species with other 14 types of chameleon species, the last 11 species i.e. from Fischer's to Mediterranean are closely related, that shows 99% similarities with each other. Therefore, in them, there are minimum mismatches & No gaps with in them. Regarding to the above conclusion the phylogenetic tree has been constructed by using MEGA 6.0

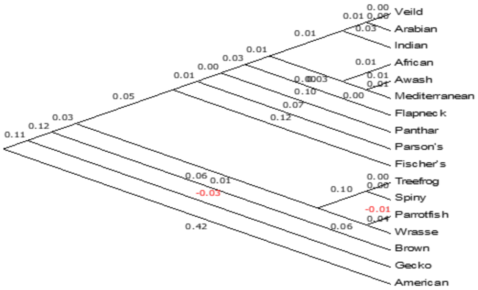
III. Phylogenetic tree construction:

Molecular Evolutionary Genetics Analysis (MEGA) software, which is an integrated tool for automatic and manual sequence alignment, inferring phylogenetic trees, mining web-based databases, estimating rates of molecular evolution, and testing evolutionary hypothesis. By using MEGA6.0. we construct following phylogenetic tree in different forms.

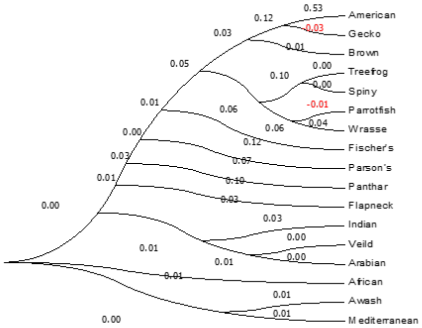
1. Rectangular Form: -



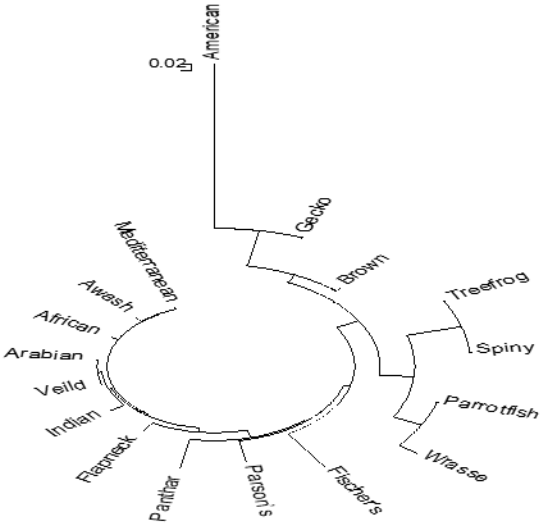
2. Straight Form: -



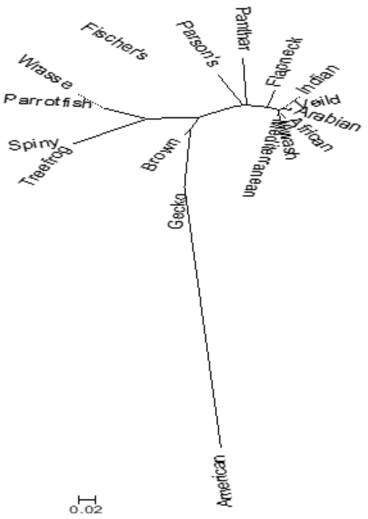
3. Curved Form: -



4. Circular form: -



5. Radiation:-



- There are Five forms of phylogenetic trees are displayed in the result of MEGA6.0 software.
- Veild chameleon & Arabian chameleon, Awash Chameleon & Mediterranean and Treefrog Chameleon & Spiny Chameleons are sharing a more recent common ancestor with each other than other in the tree. Hence, they are closely related with each other i.e. homologous with each other. the evolutionary distance between these species is 0.0 to 0.0, 0.01 to 0.01 and 0.0 to 0.0 respectively.
- The horizontal lines are branches and represent evolutionary lineages changing over time.
- The longer the branch in the horizontal dimension, the larger the amount of change.
- The bar at the bottom of the figure provides a scale for this. In this case the line segment with the number '0.02' shows the length of branch that represents an amount of genetic change of 0.02.
- Flap neck, Panther, Parson's, Fischer's, Brown and American chameleon share a more recent common ancestor with each other than other in the tree hence they are closely related with each other. The evolutionary distance between these species is 0.01 to 0.42 respectively.
- Veild chameleon and Arabian chameleon are closely related with Indian chameleon. The time of divergence among these species are 0.01%
- Awash chameleon and Mediterranean chameleon are closely related with African Chameleon. The time of divergence between these species are 0.0%.
- Parrotfish chameleon and Geko chameleons are distantly related with other 15 species of chameleon i.e. non-homologous with other species.

IV. MOTIF IDENTIFICATION

The MEME Suite allows the biologist to discover novel motifs in collections of unaligned nucleotide or protein sequences, and to perform a wide variety of other motif-based analyses.
The result has shown below:

	Logo	E-value	Sites	Width	More	Submit/Download
1.		7.2e-558	14	50		→
2.		6.2e-510	14	50		→
3.		8.4e-386	13	50		→

Stopped because requested number of motifs (3) found.

MOTIF LOCATIONS

○ Only Motif Sites ○ Motif Sites+Scanned Sites ● All Sequences

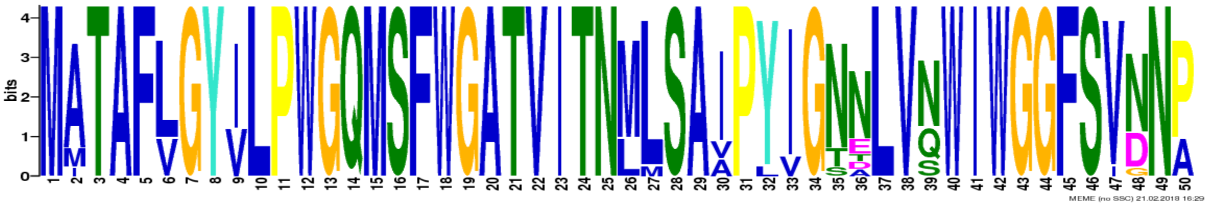
Name	p-value	Motif Location
1. American_chameleon	8.27e-2	
2. Flapneck_Chameleon	1.98e-171	
3. Fischers_Chameleon	5.00e-150	
4. Awash_chameleon	5.07e-167	
5. Mediterranean_Chameleon	7.93e-171	
6. Indian_Chameleon	6.99e-170	
7. African_Chameleon	4.84e-168	
8. Veild_Chameleon	7.79e-171	
9. Arabian_Chameleon	7.93e-171	
10. Panthar_Chameleon	3.67e-158	
11. Treefrog_Chameleon	3.29e-94	
12. Spiny_Chameleon	3.29e-94	
13. Parson_Chameleon	3.24e-163	
14. Parrotfish_Chameleon	9.57e-111	
15. Brown_Chameleon	2.56e-55	
16. Gecko_Chameleon	2.64e-48	
17. Wrasse_Chameleon	6.06e-152	

Log Likelihood Ratio: 1952, Information Content: 200.9, Relative Entropy: 201.2, Bayes Threshold: 9.86763

Name	Start	*p-value*	Sites
2.Flapneck_chameleon	65	1.60e-66	SVAHICREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYLYKETWN
9.Arabian chameleon	65	3.40e-64	SIAHISREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYLYKETWN
8. Veild chameleon	65	3.40e-64	SIAHISREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYLYKETWN
7.African chameleon	65	3.40e-64	SIAHISREVNFGWLMRNLHANGASMFFIC IYLHIGRGLYYSSYLYKETWN
5. Mediterranean chameleon	65	3.40e-64	SIAHISREVNFGWLMRNLHANGASMFFIC IYLHIGRGLYYSSYLYKETWN
4. Awash chameleon	65	3.40e-64	SIAHISREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYLYKETWN
6.Indian chameleon	65	3.03e-63	SIAHISREVNFGWLMRNLHANGASMFFIC IYLHIGRGLYYSSYLHKETWN
17. Wrasse chameleon	30	1.35e-61	SVAHICRDVNYGWLIRNMHANGASFFFICIYLHIGRGLYYGSYLYKETWN
14.Parrotfish Chameleon	48	1.35e-61	SVAHICRDVNYGWLIRNMHANGASFFFICIYLHIGRGLYYGSYLYKETWN
10.Panthar Chameleon	65	1.53e-61	SVSHICREVNFGWLTRNLHANGASMFFICIYLHIGRGLYYSSFLHKETWN
13.Parson's Chameleon	65	2.91e-61	SVSLICREVNFGWLMRSLHANGASMFFICIYLHIGRGLYYSSYLYKETWN

15.Brown Chameleon	65	3.03e-60	SVSHICREVNLGWLLRNLHANGASMFFICIYLHIGRGIYYASYLYKETWN
3. Fischer's Chameleon	65	9.59e-57	SVSTICREVNYGWLMRNMHANGASMFFICIFLHIGRGLFYSSYLFKETWI
16.Gecko Chameleon	53	2.10e-53	SISHICRNVNYGWLIRNMHTNSASLFFICLYLHIGRGLYYGSHFYKETWN

- In above list of chameleon species Arabian, Veild, African, Mediterranean, Awash and Indian chameleon shows complete similarity in their sequence sites with each other.
- And Flap neck, Panthar, Parsons, Brown, Wrasse, Parrotfish and Fishers shows similarity but quite different from above six species and with each other.



Log Likelihood Ratio: 1862, Information Content: 203.6, Relative Entropy: 191.8, Bayes Threshold: 8.03772

Name	Starts	*p-value*	Sites
13.Parson's chameleon	124	2.58e-61	MATAFLGYVLPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
9.Arabian chameleon	124	5.63e-61	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
8.Veild chameleon	124	5.63e-61	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
6.Indian chameleon	124	5.63e-61	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
5. Mediterranean chameleon	124	5.63e-61	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
4. Awash chameleon	124	5.63e-61	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSVNNP
7. African chameleon	124	3.38e-60	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVSWIWGGFSVNNP
2. Flapneck chameleon	124	2.97e-59	MATAFLGYILPWGQMSFWGATVITNMLSAIPYIGSNLVNWIWGGFSVNNP
10. Panthar chameleon	124	1.56e-57	MATAFLGYILPWGQMSFWGATVITNMMSAIPYIGNNLVSWIWGGFSINNP
17.Wrasse chameleon	89	4.86e-57	MMTAFVGYVLPWGQMSFWGATVITNLLSAVPYVGNALVQWIWGGFSVDNA
12. Spiny chameleon	6	5.93e-57	MATAFVGYVLPWGQMSFWGATVITNLLSAAPYIGTELVQWIWGGFSVDNA

11.Treefrog chameleon	6	5.93e-57	MATAFVGYVLPWGQMSFWGATVITNLLSAAPYIGTELV QWIWGGFSVDNA
14.Parrotfish chameleon	107	7.25e-57	MMTAFVGYVLPWGQMSFWGATVITNLLSAVPYVGNTL VQWIWGGFSVDNA
3.Fischer's chameleon	124	1.08e-55	MITAFLGYILPWGQMSFWGATVITNMLSAIPLIGNDLVN WIWGGFSVGNP

- In above list of chameleon species Arabian, Veild, African, Indian, Mediterranean and Awash chameleon shows complete similarity in their sequence sites with each other.

- And Parson's, African, Flap neck, Panther, Wrasse, Spiny, Treefrog, Parrotfish and Fischer's shows similarity but quite different from above six species and with each other.

Log Likelihood Ratio: 1580, Information Content: 192.7, Relative Entropy: 175.3, Bayes Threshold:8.05641

Name	Start	p-value	Sites
9. Arabian chameleon	183	1.44e-59	FLTPFLIAGTTMLHLLFLHETGSNNPTGLSS NSDKVSFHPYFPLKDLTTA
8. Vield chameleon	183	1.44e-59	FLTPFLIAGTTMLHLLFLHETGSNNPTGLSSNSDKVSF HPYFPLKDLTTA
6. Indian chameleon	183	1.44e-59	FLTPFLIAGTTMLHLLFLHETGSNNPTG LSSNSDKVSFHPYFPLKDLTTA
5.Mediterranean chameleon	183	1.44e-59	FLTPFLIAGTTMLHLLFLHETGSNNPTG LSSNSDKVSFHPYFPLKDLTTA
2.Flapneck chameleon	183	1.44e-59	FLTPFLIAGATMLHLLFLHETGSNNPTG LSSNSDKVSFHPYFPLKDLTTA
7. African chameleon	183	1.51e-57	FLTPFLIAGTTMLHLLFLHETGSSNPTG LSSNSDKVPFHPYFPLKDLTTA
4. Awash chameleon	183	9.61e-56	FLTPFLIAGTTTLHLLFLHETGSNNPTG LSSNSDKVSFHPYFPLKDLTAA
13.Parson's chameleon	183	1.64e-54	FLMPFLIAGTTMIHLLFLHETGSSNPTG LSSNPDKITFHPYFPLKDLCSA
10.Panthar chameleon	183	1.56e-52	FLLPFIIAGATMLHLLFLHETGSSNPTGMSSNLDKVT FHPYFPLKDLTTM
3.Fischer's chameleon	183	2.10e-51	FMMPFLIAGTTTLHLLFLHETGSSNPTGLSSNPDKVL FHPYFTLKDLTTA
12.Spiny chameleon	65	9.93e-47	FILPFIIAGASMIHLLFLHQTGSSNPTGLNSNPDKIPFH AYYSYKDAFGF

11.Treefrog chameleon	65	9.93e-47	FILPFIIAGASMIHLLFLHQTGSSNPTGLNSNPDKIPFHAYYSYKDAFGF
17.Wrasse chameleon	148	3.18e-46	FLFPFVIAAATVVHLLFLHETGSNNPLGLNSDADKISFHPYFSYKDLLGF

- In above list of chameleon species Arabian, Veild, Indian, Mediterranean and Flap neck chameleon shows complete similarity in their sequence sites with each other.
- And African, Awash, Flap neck, Parson's, Panthar, Fischer's, Spiny, Treefrog and Wrasse shows similarity but quite different from above five species and with each other.

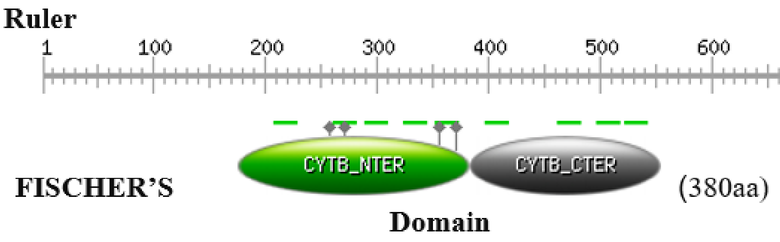
- ➢ Start: The position in the sequence where the motif site starts. If a motif started right at the beginning of a sequence it would be described as starting at position 1.
- ➢ P-value: The probability that an equal or better site would be found in a random sequence of the same length conforming to the background letter frequencies.
- ➢ A motif site with the flanking letters on either side. When the site is not on the given strand then the site and both flanks are reverse complemented so they align.

V. DOMAIN IDENTIFICATION

- Using Prosite online tool, we discovered the domain in given sequences.
- It consists of a database of biologically significant sites and patterns formulated in such a way that with appropriate computational tools it can rapidly and reliably identify which known family of protein (if any) the new sequence belongs to.
- PROSITE's uses include identifying possible functions of newly discovered proteins and analysis of known proteins for previously undetermined activity.
- Prosite results of some species of chameleon has shown below in tubular form:

1. Fischer's:

Fischer's species contains 380 amino acid and domain is observed in between 100-500 region of ruler scale



PS5100321 CYTB_CTER Cytochrome b/b6 C-terminal region profile

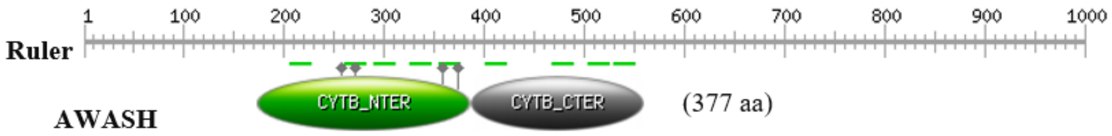
Accession No.	PS510021-109: score= 34.333				
Protein sequence	--------MTItrkhhpmlKIINHSLIDLPTPTNITTAWNFGSLIGLMLMVQITTGLLLA THYTADVNMAFISVSHICREVNFGWLTRNLHANGASMFFICIYLHIGRGLYYSSFL--HK ETWNIGIILLFMTMATAFLGYILPWGQMSFWGATVITNMMSAIPYIGNNLVSWIWGGFSI NNPTLIRFFTAH-FLLPFIIAGATMLHLLFLHETGSSNPT				
Predicted Features	TRANSMEM	33	53	Helical	[condition: none
	METAL	83		Iron (heme2 axial ligand)	[condition: H and not(<AnyFeature:PS51003)] [group:1]
	TRANSMEM	86	106	Helical	[condition: none]
	METAL	97		Iron (heme 1 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	TRANSMEM	115	135	Helical	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	TRANSMEM	150	170	Helical	[condition: none]
	TRANSMEM	178	198	Helical	[condition: none]
	METAL	182		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	METAL	196		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]

PS510032010 CYTB_CTER Cytochrome b/b6 C-terminal region profile:

Accession No.	PS510032010-290: score=25.887				
Protein sequence	GLSSNSDKVSFHPYFPLKDLTTATTITLLTLIALFLPTLLTDPQNFSPANPLSTPPHIM PEWYFLFAYTILRSIPNKLGG-------------------------------------- --				
Predicted Features	TRANSMEM	224	244	Helical	[condition: none]
	TRANSMEM	288	290	Helical	[condition: none]
	TRANSMEM	290	290	Helical	[condition: none]
	TRANSMEM	290	290	Helical	[condition: none]

2. Awash

Awash species contains 377 amino acid and domain is observed in between 100-500 region of ruler scale



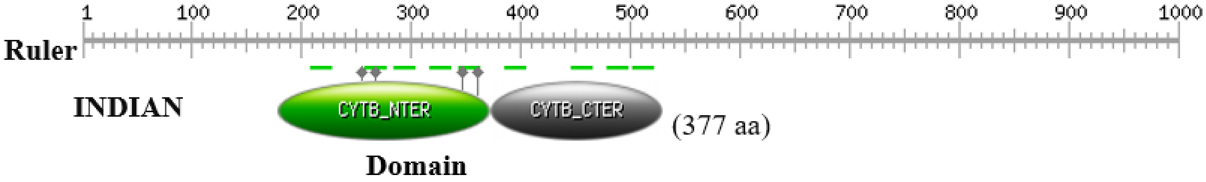
Domain

PS5100321 CYTB_CTER Cytochrome b/b6 C-terminal region profile:

Accession No.	PS510021-109: score= 34.325				
Protein sequence	--------MTItrksnpalKIINHSLIDLPTPTNISTAWNFGSLLGLMLVTQITTGLFLA MHYTADINLAFSSIAHISREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYL--YK ETWNIGIILLFLTMATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSV NNPTLIRFFTLH-FLTPFLIAGTTTLHLLFLHETGSNNPT				
Predicted Features	TRANSMEM	33	53	Helical	[condition: none]
	METAL	83		Iron (heme2 axial ligand)	[condition: H and not(<AnyFeature:PS51003)] [group:1]
	TRANSMEM	86	106	Helical	[condition: none]
	METAL	97		Iron (heme 1 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	TRANSMEM	115	135	Helical	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	TRANSMEM	150	170	Helical	[condition: none]
	TRANSMEM	178	198	Helical	[condition: none]
	METAL	182		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	METAL	196		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]

3. Indian

Indian species contains 377 amino acid and domain is observed in between 100-500 region of ruler scale



PS510021 CYTB_CTER Cytochrome b/b6 C-terminal region profile:

Accession No.	PS510021-209: score= 34.256				
Protein sequence	--------MTItrksnpalKIINHSLIDLPTPTNISTAWNFGSLLGLMLVTQITTGLFLA MHYTANIDLAFSSIAHISREVNFGWLMRNLHANGASMFFICIYLHIGRGLYYSSYL—HK ETWNIGIILLLMTMATAFLGYILPWGQMSFWGATVITNMLSAIPYIGNNLVNWIWGGFSV NNPTLIRFFTLH-FLTPFLIAGTTMLHLLFLHETGSNNPT				
Predicted Features	TRANSMEM	33	53	Helical	[condition: none]
	METAL	83		Iron (heme2 axial ligand)	[condition: H and not(<AnyFeature:PS51003)] [group:1]
	TRANSMEM	86	106	Helical	[condition: none]
	METAL	97		Iron (heme 1 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	TRANSMEM	115	135	Helical	[condition: none]
	TRANSMEM	150	170	Helical	[condition: none]
	TRANSMEM	178	198	Helical	[condition: none]
	METAL	182		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]
	METAL	196		Iron (heme 2 axial ligand)	[condition: H and not(<AnyFeature:PS51003>)] [group:1]

CONCLUSION

- This paper is Analysis of Phylogenetic tree of different Chameleon species on the basis of cytochrome b protein. Cytochrome b is used as a region of mitochondrial DNA. Mitochondrial disease that causes prominent muscular problems are called Mitochondrial myopathy, while mitochondrial disease causes both prominent muscular & neurological problems are called Mitochondrial encephalomyopathy (It refers to the brain).
- The study of this paper may be helpful in future to drug discovery against Chameleon. These species can be protected from the diseases or disorders that causes by mutation in cytochrome b molecule.

REFERENCES

[1] By Alina Bradford, Live Science Contributor | June 3, 2015 12:15am ET

[2] Farias IP[1], Ortí G, Sampaio I, Schneider H, Meyer A: 2001. Aug;53(2):89-103.

[3] Barry G. Hall* Bellingham Research Institute, Bellingham, Washington, from https://academic.oup.com/mbe/article-abstract/30/5/1229/992850.

[4] By David Baum, Ph.D. (Dept. of Botany, University of Wisconsin, 430 Lincoln Ave., Madison, WI)

[5] https://www.khanacademy.org/science/biology/her/tree-of-life/a/phylogenetic-trees

[6] https://ghr.nlm.nih.gov/gene/MT-CYB

[7] Molecular Biology and Evolution. 18 (4): 465–471. doi:10.1093/oxford journals. molbev. a003825. PMID 11264397.

[8] Nucleic Acids Research, 2015, Vol. 43, Database issue Published online 27 October 2014 doi: 10.1093/nar/gku989.

[9] By Sievers F[1], Higgins DG ,2014;1079:105-16. doi: 10.1007/978-1-62703-646-7_6.Clustal Omega, accurate alignment of very large numbers of sequences.

[10] By Timothy L. Bailey,[*]Nadya Williams,[1]Chris Misleh,[1] and Wilfred W. Li[1:]2006 Jul 14. doi: 10.1093/nar/gkl198

[11] By Koichiro Tamura,1,2 Glen Stecher,3 Daniel Peterson,3 Alan Filipski,3 and Sudhir Kumar.

[12] By Research Center for Genomics and Bioinformatics, Tokyo Metropolitan University, Hachioji, Tokyo, Japan

[13] By Center for Evolutionary Medicine and Informatics, Biodesign Institute, Arizona State University

[14] By Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia)

[15] By Timothy L. Bailey Mikael Boden Fabian A. Buske Martin Frith Charles E. GrantLucaClementi JingyuanRen Wilfred W. Li William S. NobleNucleic Acids Research, Volume 37.

[16] By J. A. Sigrist1, Lorenzo Cerutti1 , Edouard de Castro1 , Petra S. Langendijk-Genevaux1 , Virginie Bulliard1 , Amos Bairoch1,2 and Nicolas Hulo1 1 Swiss Institute of Bioinformatics (SIB), Centre Me´dicalUniversitaire and 2 Structural Biology and Bioinformatics Department, University of Geneva, Geneva

[17] By Li W., Jaroszewski,L. and Godzik,A. (2002) Tolerating some redundancy significantly speeds up clustering of large protein databases. Bioinformatics, 18, 77–82.

[18] Hubbard T., Barker,D., Birney,E., Cameron,G., Chen,Y., Clark,L., Cox,T., Cuff,J., Curwen,V., Down,T. *et al*. (2002) The Ensembl genome database project. Nucleic Acids Res., 30, 38–41.

[19] Kersey P., Hermjakob,H. and Apweiler,R. (2000) VARSPLIC: alternatively-spliced protein sequences derived from Swiss-Prot and TrEMBL. Bioinformatics, 11, 1048–1049.

[20] S. B. Needleman and C. D. Wunsch, “A general method applicable to the search for similarities in the amino acid sequence of two proteins,” Journal of Molecular Biology

[21] T. F. Smith and M. S. Waterman, “Identification of common molecular subsequences,” Journal of Molecular Biology, vol. 147

[22] Frith MC, Saunders NFW, Kobe B, Bailey TL. Discovering sequence motifs with arbitrary insertions and deletions. PLoS Comput. Biol. 2008;4 e1000071.

[23] Sanner MF. A component-based software environment for visualizing large macromolecular assemblies. Structure. 2005; 13:447–462.

[24] Sigrist, C.J.A., De Castro, E., Langendijk-Genevaux, P.S., Le Saux, V., Bairoch, A., Hulo, N. 2005

[25] Hedges SB, Kumar S. Precision of molecular time estimates. Trends Genet. 2004; 20:242–247.

[26] Baum, D. (2008) Reading a Phylogenetic Tree: The Meaning of Monophyletic Groups. Nature Education 1(1):190

[27] SouravSingha Roy, RakhiDasgupta, AngshumanBagchi. Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, India.

[28] Arvelakis A, Reczko M, Stamatakis A, Symeonidis A and Tollis IG: Using treemaps to

visualize phylogenetic trees. Proc of ISMBDA2005, in press.

[29] Bader DA, Moret BME and Vawter L: Industrial applications of high-performance computing for phylogeny reconstruction. Proc of SPIE ITCom 4528: 159-168, 2001.

[30] Baker CS and Palumbi SR: Which whales are hunted? A molecular genetic approach to whaling. Science 265: 1538-1539, 1994.