# **Study of self-Induced Mutations on Nail Protein**

#### \* Ms. Sunanda Pandita

School of Computer Science and Engineering, Sandip University Nashik \*sunanda7723@gmail.com **Pravin Gundalwar** 

School of Computer Science and Engineering, Sandip University Nashik gundalwarp@gmail.com

**Abstract-**All primates have nails, which are flat plates that resemble claws and are located at the tips of the fingers and toes. Nails are equivalent to other animals' claws. Alpha-keratin, a polymerized protein with a thick protective coating, makes up fingernails and toenails. Vertebrate hooves, claws, and horns contain alpha-keratin. 144 proteins were found in samples from Caucasian participants during a shotgun proteomic study of the human nail plate. 90% of the mass of the nail plate was made up of 30 identified proteins, the majority of which were keratins and proteins connected to keratins. The majority of the detergent-insoluble fraction was made up of keratins, but several cytoplasmic, membrane, and junctional proteins as well as histones were also found. This suggests that transglutaminases widely utilise readily available proteins as cross-linking substrates.

Keywords: Protein, Amino acid, Keratin, Mutation, Sequence

#### Introduction

A change in the amino acid sequence may affect a protein's structure but not always its function. However, a change at some locations, such as conserved residues, might affect both the protein's structure and function. This paper demonstrates the change in percentage similarity after introducing mutations at different paces in the original aminoacid sequence of a primate nail.

Subjects and samples

- a) Homo Sapiens- Homo sapiens (Latin for "wise man") is the name given to the human species. Homo sapiens is the sole living member of the genus Homo, which contains multiple extinct species.
- b) Gorilla- Gorillas are giant apes that are herbivorous and spend most of their time on the ground in the tropical forests of equatorial Africa. There are two species of gorillas, the eastern and western, and up to five subspecies within the genus gorilla.
- c) Nomascus leucogenya- Nomascus leucogenys (Northern White Cheeked Gibbon) is a species of <u>primates</u> in the family <u>gibbons</u>. They are listed as critically endangered by IUCN and in cites appendix i. They are native to Asia. They are diurnal herbivores. Individuals can grow to 545 mm. Reproduction is viviparous. They have parental care (female provides care and paternal care).

- d) The northern white- Critically endangered cheeked gibbons are found in South and Southeast Asia. The southern white-cheeked gibbon, until thought to be a conspecific, is a close relative. Females of both species look very similar to one another.
- e) Hylobates moloch- The silvery gibbon, or Javan gibbon, is a primate belonging to the genus Hylobatis and subfamily Hylobatidae. It can only be found on the Indonesian island of Java, which is in which it lives in the unspoiled rainforests at elevations of up to 2,450 meters.
- f) Pan troglotytes- Chimpanzees, or chimps as they are more commonly known, are a great ape species found across the tropical forests and savannahs of Africa. There are four recognized subspecies and a possible fifth.
- g) Pan paniscus- A critically endangered great ape along with one of only two species in the genus Pan, the bonobo was formerly known as the pygmy chimpanzee.
- h) Papio Anubis- The anubis, or olive baboon (*P. anubis*), is only slightly smaller than the chacma and olive in colour; the male has a large mane of hair over the head and shoulders. The anubis baboon has a wide range, from the hinterland of Kenya and Ethiopia through the grasslands...
- i) Pongo pygmaeus- Orangutans are great apes native to the rainforests of Indonesia and Malaysia. They are now found only in parts of Borneo and Sumatra, but during the Pleistocene they ranged throughout Southeast Asia and South China. Classified in the genus *Pongo*, orangutans were originally considered to be one species.
- j) Cebus imitator- is a medium-sized New World monkey of the family Cebidae, subfamily Cebinae. Native to the forests of Central America, the white-faced capuchin is important to rainforest ecology for its role in dispersing seeds and pollen.

#### Simulation Tool

Sequence comparison is just one of the many techniques used by computational biologists, such as similarity searching. Other types of biologists frequently use this technique. BLAST (basic local alignment search technique), which compares pairs of sequences to look for local similarities, is the most used tool for this. In the 11 years since its publication, the original study that introduced BLAST has been cited more than 12,000 times, and employing BLAST has turned into a crucial tool in biology. It is crucial to comprehend how it works, what it does, how to utilise it appropriately, and how to analyse outcomes that have already been published. There are numerous contemporary versions of the BLAST algorithm, but NCBI BLAST and WU-BLAST share a common ancestor and are the most popular right now. Washington University in St. Louis offers WU-BLAST, while the National Centre for Biotechnology Information (NCBI) offers NCBI BLAST. The NCBI version of BLAST is described along with its foundations, methods, uses, and possible risks.

- **Maximum Score** The Maximum Score represents the highest possible alignment score (bit-score) that may be achieved between the query sequence & the database segments. It has a relationship with the e-value that is inversely proportional.
- **Total Score** The term "Total Score" refers to the sum of the alignment scores for all of the sequences contained in the shared database.
- **Percent Query Coverage-** The term "Percent Query Coverage" refers to the proportion of the total length of the query that may be located in the segments that have been aligned.
- **E-value-** The e-value provides a measure of the likelihood that the similarity in sequences is not just the result of random chance.
- **Percent Identity-** The percentage of identity reveals how closely the query sequence is related to the sequences that have been aligned.

#### **Experimental Findings**

Original AA sequence of keratin protein of human nail

VTLARTDLEMQIEGLKEELAYLRKNHEEEMLALRGQTGGDVNVEMDAAPGVD LSRILNEMRDQYEQMAEKNRRDAETWFLSKTEELNKEVASNSELVQSSRSEVTE LRRVLQGLEIELQSQLSTKASLENSLEETKGRYCMQLSQIQGLIGSVEEQLAQL RCEMEQQSQEYQILLDVKTRLEHEIATYRRLLXGEDAHLSSQQASGQSYSSREV FTSSSSSSSRQTRPILKEQSSSSFSQGQSS

### ORIGINAL AMINO ACID SEQUENCE OF KERATIN PROTEIN IN HUMAN NAIL

😒 WhatsApp	🗙 🔗 Protein BLAST: search protein: 🗴 😨 keratin - Keratin - Homo sap: 🗙 🕤 Thermosynechococcus vestil: 🗙 🛛 M Inbox (11,003) - sunanda772 🗙 📔 + 🛛 🗠 🗡
< → C (A) b	trichi.nim.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LlINK_LOC=blasthome 🖆 🎓 😇 🛸 🛓 🔲 📵 Error 🔅
NIH	onal Center for Biotechnology Information
BLAST <sup>®</sup> »	Jastp suite Home Recent Results Saved Strategies Help
blastn b	Standard Protein BLAST
Enter Query	BLASTP programs search protein databases using a protein query. more Reset page Bookmark
Enter accession I VTLARTDLEMOIEG DLSRILNEMRBOY EVTELRRVLOGLEI OLAOLRCEMEOOS Or, upload file Job Title	mber(s), gl(s), or FASTA sequence(s)        Clear       Query subrange          CELLYLRKWHEEEMLALRGOTGGDVNVEMDAAPGV       From
Choose Sear	i Set
Databases	Standard databases (nr etc.): Mee O Experimental databases     For more info see What is clustered nr?
Compare	Select to compare standard and experimental database 🕢
Standard Database Organism	Non-redundant protein sequences (nr)       Image: Completions will be suggested       Image: Completions will be suggested
27°C Smoke	📑 🔍 Search 💭 🐗 🤡 🦕 💽 🐻 🗳 🗮 💶 🔷 💦 😵 🕫 🕬 🖬 1155 P203 👁

BLAST Window showing Original AMINO ACID Sequence of KERATIN of Human finger nail pasted in the text box

op: 2849 – 28	65	
🕥 WhatsApp	🗙 😫 NCBI Blast:Protein Sequence 🗙 🧑 keratin - Keratin - Homo sap 🗴	🗙 🛛 🌀 Thermosynechococcus vestit: 🗙 🖌 M Inbox (11,003) - sunanda 772 : 🗙 📔 🕂 🔍 🦳 🗖 🛛 🗙
$\leftrightarrow$ $\rightarrow$ C $\square$ blast.	ncbi.nlm.nih.gov/Blast.cgi	🖻 🖈 😈 🖨 🚺 💽 Error 🚦
An official wet	site of the United States government Here's how you know.~ ional Library of Medicine nal Center for Biotechnology Information	Log in
BLAST <sup>®</sup> » bli	astp suite » results for RID-0PXBHWEV013	Home Recent Results Saved Strategies Help
Job Title	Protein Sequence	How to read this report?     BLAST Help Videos DBack to Traditional Results Page      Filter Results
RID	0PXBHWEV013 Search expires on 03-12 01:22 am Download All	
Program	BLASTP 😧 Citation 🗸	Organism only top 20 will appear
Database	nr <u>See details</u> ✓	lype common name, binomial, taxid or group name
Query ID	Icl Query_10691	+ Add organism
Description	unnamed protein product	Percent Identity E value Query Coverage
Molecule type	amino acid	to to to
Query Length	244	
Other reports	Distance tree of results Multiple alignment MSA viewer	Filter Reset
	Compare these results against the	new Clustered nr database 😧 BLAST
Descriptions	Graphic Summary Alignments Taxonomy	
28°C Smoke	Q Search D 📫	🔮 🧧 😧 🖪 🗘 🧮 🖏 🔷 🛌 💦 ENG 🗇 🕩 🕺 1158 PM 🕢

# BLAST Window showing Result 1 of original AMINO ACID Sequence of Human Nail Keratin protein

🕒 WhatsAp	p 🗙 😣 NCBI Blast:Protein Sequence 🗙 🥺 keratin - Keratin - Homo sap	× G Thermosynechococcus vestit	🗙   📉 Inbox (	11,003) - sunanda7	72: ×   +	~ -	οx
$\leftrightarrow \rightarrow c$	blast.ncbi.nlm.nih.gov/Blast.cgi				🖻 🖈 🛡	* 🕹 🖬 🌘	Error :
2	select all 100 sequences selected	GenPept Graphics	Distance tree	e of results	Multiple alignm	ent MSA Viewe	1 1
	Description	Scientific Name	Max Total Score Score	Query Cover	Per. Acc. Ident Len	Accession	
	keratin [Homo sapiens]	Homo sapiens	482 482	100% 2e-17	1 100.00% 244	AAB30058.2	
	type I keratin 16 [Homo sapiens]	Homo sapiens	473 473	100% 2e-16	4 98.36% 473	AAB35421.1	
	keratin, type I cytoskeletal 16 [Homo sapiens]	Homo sapiens	473 473	100% 2e-16	4 98.36% 473	NP_005548.2	
	keratin 16 [Pan troglodytes verus]	Pan troglodytes verus	471 471	100% 7e-16	4 97.95% 461	BAF62417.1	
	keratin, type I cytoskeletal 16 [Pan troglodytes]	Pan troglodytes	471 471	100% 9e-16	4 97.95% 469	NP_001138310.1	
	keratin, type I cytoskeletal 16 [Pan paniscus]	Pan paniscus	471 471	100% 1e-16	3 97.95% 469	XP_003813891.1	
	keratin, type I cytoskeletal 16 [Gorilla gorilla gorilla]	Gorilla gorilla gorilla	471 471	100% 1e-16	3 97.95% 473	XP_004041720.2	
	keratin, type I cytoskeletal 16 isoform X2 [Nomascus leucogenys]	Nomascus leucogenys	469 469	100% 7e-16	3 97.13% 476	XP_003279481.2	
	keratin, type I cytoskeletal 16 [Hylobates moloch]	Hylobates moloch	467 467	100% 6e-16	2 96.72% 476	XP_032019500.1	
	keratin, type I cytoskeletal 16 isoform X2 [Chlorocebus sabaeus]	Chlorocebus sabaeus	467 467	100% 7e-16	2 96.31% 476	XP_008010896.1	
	keratin, type I cytoskeletal 16 isoform X2 [Papio anubis]	Papio anubis	466 466	100% 1e-16	1 96.31% 476	XP_003913047.1	
	keratin, type I cytoskeletal 16 isoform X2 [Saimiri boliviensis boliviensis]	Saimiri boliviensis boliviensis	466 466	100% 1e-16	1 96.72% 475	XP_003942841.2	
	<u>keratin, type I cytoskeletal 16 [Pongo abelii]</u>	Pongo abelii	466 466	100% 1e-16	1 96.72% 474	XP_002827606.2	
	keratin, type I cytoskeletal 16 [Theropithecus gelada]	Theropithecus gelada	466 466	100% 1e-16	1 96.31% 476	XP_025218966.1	
	<u>keratin, type I cytoskeletal 16 isoform X3 [Sapajus apella]</u>	<u>Sapajus apella</u>	464 464	100% 9e-16	1 96.31% 475	XP_032135572.1	ack
	keratin, type I cytoskeletal 16 isoform X2 [Cebus imitator]	Cebus imitator	464 464	100% 9e-16	1 96.31% 475	XP_017364946.1	eedb
	keratin, type I cytoskeletal 16 [Callithrix jacchus]	Callithrix jacchus	446 446	100% 8e-15	4 94.26% 467	XP_002748663.2	
	PREDICTED: keratin, type I cytoskeletal 16 [Colobus angolensis palliatus]	Colobus angolensis palliatus	437 437	95% 9e-15	0 96.57% 500	XP_011785156.1	
	PREDICTED: LOW QUALITY PROTEIN: keratin, type I cytoskeletal 16 [Rhinopithecus bieti	Rhinopithecus bieti	436 436	95% 9e-15	0 96.57% 480	XP_017735629.1	
28°C Smoke	Q Search 🛄 📫	🔮 📮 C 🖬 🛇	🚊 👊		^ Et	ig ⊜ ¢0) ∎ :	11:59 PM 3/10/2023

# BLAST Window showing Result 2 of original AMINO ACID Sequence of Human Nail Keratin protein

Sr. No	Scientific nome	Max	Total	Query	E-	Percentage
SI. INU.	Scientific fiame	Score	Score	Cover	Value	Identity
					2.00E-	
1	Homo Sapiens	482	482	100%	171	100%
		-			9.00F-	
2	Pan troglodytes	471	471	100%	).00E- 164	97.95%
		-			1.005	
3	Pan paniscus	471	<i>A</i> <b>7</b> 1	100%	1.00E- 163	97 95%
			7/1	10070	105	71.7570
4	Gorilla	471	471	1000/	1.00E-	07.050
		471	471	100%	163	97.95%
5	Nomascus	-			7.00E-	
5	Leucogenya	469	469	100%	163	97.13%
6	II-l-h-tlh	-			6.00E-	
0	Hylobates molocn	467	467	100%	162	96.72%
					7.00E-	
7	chlorocebus sebaeus	467	467	100%	176	96.31%
					1 00F-	
8	papio anubis	466	466	100%	1.00L- 161	96.31%
		1			1.005	
9	Pongo pygmaeus	166	<i>1</i> 66	100%	1.00E- 161	96 72%
		+00	+00	10070	101	JU.1270
10	Cebus imitator			1000/	9.00E-	0.6.010/
		464	464	100%	161	96.31%

# MUTATION 1 at position 12 replace L with Q

VTLARTDLEMQQEGLKEELAYLRKNHEEEMLALRGQTGGDVNVEMDAAPGVDLSR ILNEMRDQYEQMAEKNRRDAETWFLSKTEELNKEVASNSELVQSSRSEVTELRRVL QGLEIELQSQLSTKASLENSLEETKGRYCMQLSQIQGLIGSVEEQLAQLRCEMEQQSQ EYQILLDVKTRLEHEIATYRRLLXGEDAHLSSQQASGQSYSSREVFTSSSSSSSRQTRP ILKEQSSSSFSQGQSS

pp: 2849 – 2	865			
😽 Fw: FIITJEE ::: Student d	letails for 🕆 🖌 📉 Inbox (34) - vaze.career@gmail.c: 🗙 🕴 🔥 My Drive - Google Drive 🛛 🗴 😒 Protein BLAST: search protein dat 🗙 🕂	V	- 0	×
← → C 🔒 blas	st.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome	12 th 13	* 🛛 🕚	:
BLAST <sup>®</sup> » t	Alastp suite Home Recent Results	Saved Strategies	Help	*
	Standard Protein BLAST			
blastn blas	stp blastx tblastn tblastx			-1
Enter Query Se	BLASTP programs search protein databases using a protein query. more	Reset page	Bookmark	)
Enter accession nu	mber(s), gi(s), or FASTA sequence(s) 😮 Clear Query subrange 😯			
VTLARTDLEMQQEGL VDLSRILNEMRDQYE	KEELAYLRKNHEEEMLALRGQTGGDVNVEMDAAPG A From From			
SEVTELRRVLQGLEIE EQLAQLRCEMEQQS	ELQSQLSTKASLENSLEETKGRYCMQLSQIQGLIGSVE T QEYQILLDVKTRLEHEIATYRRLLXGEDAHLSSQQAS To			
Or, upload file	Choose File No file chosen			
Job Title				
	Enter a descriptive title for your BLAST search 😮			
Align two or more	s sequences 🔇			
Choose Search	1 Set			
Databases	Standard databases (nr etc.): Meet O Experimental databases			
	For more info see What is clustered nr?			
Compare	Select to compare standard and experimental database 😵			
Standard				
Organism	Non-redundant protein sequences (nr)			
Optional	Enter organism name or idcompletions will be suggested Add organism			
	Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown 😵			
Ontional	Models (XM/XP) Mon-redundant RefSeq proteins (WP) Uncultured/environmental sample sequences			-
	💾 Q. Search 🛛 🗊 🖬 🧶 🤡 🔳 💷	へ ENG (最 d))	5:23 PM 4/17/2023	ġ

# BLAST WINDOW with search query showing MUTATION 1 at position 12 replace L $\,$

Q

M Fw: FIITJEE ::: Student details for T 🗙 M Inbox (34) - vaze.career@gmail.c 🗙 🛛 🐴 My Drive - Google Driv	e 🗙 😣 NCBI Blast:Protein Seq	uence	× +			×
← → C					e 🖈 🗉 🛊 🛛 🚳	) :
keratin [Homo sapiens]	Homo sapiens	479	479 100	6 4e-170 99.59%	244 AAB30058.2	*
type I keratin 16 [Homo sapiens]	Homo sapiens	470	470 100	6 4e-163 97.95%	473 AAB35421.1	
keratin. type I cytoskeletal 16 [Homo sapiens]	Homo sapiens	470	470 100	6 4e-163 97.95%	473 <u>NP_005548.2</u>	
keratin 16 [Pan troglodytes verus]	Pan troglodytes verus	468	468 100	6 2e-162 97.54%	461 <u>BAF62417.1</u>	
keratin_type I cytoskeletal 16 [Pan paniscus]	Pan paniscus	468	468 100	6 2e-162 97.54%	469 <u>XP_003813891.1</u>	- 1
keratin_type I cytoskeletal 16 [Pan troglodytes]	Pan troglodytes	468	468 100	6 2e-162 97.54%	469 <u>NP_001138310.1</u>	
keratin, type I cytoskeletal 16 [Gorilla gorilla gorilla]	Gorilla gorilla gorilla	468	468 100	6 2e-162 97.54%	473 <u>XP_004041720.2</u>	
keratin_type I cytoskeletal 16 isoform X2 [Nomascus leucogenys]	Nomascus leucogenys	466	466 100	6 2e-161 96.72%	476 XP_003279481.2	- 1
keratin, type I cytoskeletal 16 [Hylobates moloch]	Hylobates moloch	464	464 100	6 1e-160 96.31%	476 <u>XP_032019500.1</u>	
keratin_type I cytoskeletal 16 [Pongo_pygmaeus]	Pongo pygmaeus	463	463 100	6 2e-160 96.31%	474 <u>XP_054313502.1</u>	
keratin. type I cytoskeletal 16 isoform X2 [Saimiri boliviensis boliviensis]	Saimiri boliviensis boliviensis	463	463 100	6 3e-160 96.31%	475 XP_003942841.2	
keratin_type I cytoskeletal 16 [Callithrix jacchus]	Callithrix jacchus	444	444 100	6 7e-153 93.85%	467 <u>XP_002748663.3</u>	
keratin. type I cytoskeletal 16 [Rhinopithecus roxellana]	Rhinopithecus roxellana	434	434 95%	1e-148 96.14%	480 <u>XP_010357842.2</u>	
keratin_type I cytoskeletal 16 [Trachypithecus francoisi]	Trachypithecus francoisi	433	433 959	2e-148 96.14%	472 XP_033038936.1	
PREDICTED: LOW QUALITY PROTEIN: keratin, type I cytoskeletal 16 [Rhinopithecus bieti	Rhinopithecus bieti	433	433 95%	2e-148 96.14%	480 <u>XP_017735629.1</u>	
PREDICTED: keratin, type I cytoskeletal 16 [Colobus angolensis palliatus]	Colobus angolensis palliatus	434	434 959	2e-148 96.14%	500 <u>XP_011785156.1</u>	
keratin, type I cytoskeletal 16 isoform X2 [Piliocolobus tephrosceles]	Piliocolobus tephrosceles	431	431 959	3e-148 95.71%	448 XP_026304479.1	
keratin_type I cytoskeletal 16 [Zalophus californianus]	Zalophus californianus	432	432 100	6 6e-148 90.98%	476 <u>XP_027423980.1</u>	
keratin_type I cytoskeletal 16 isoform X1 [Piliocolobus tephrosceles]	Piliocolobus tephrosceles	431	431 959	7e-148 95.71%	476 <u>XP_023059895.2</u>	
keratin. type I cytoskeletal 16 isoform X2 [Chlorocebus sabaeus]	Chlorocebus sabaeus	431	431 95%	8e-148 96.12%	476 <u>XP_008010896.1</u>	
keratin_type I cytoskeletal 16 isoform X2 [Papio anubis]	Papio anubis	431	431 95%	1e-147 96.12%	476 <u>XP_003913047.1</u>	
keratin, type I cytoskeletal 16 [Theropithecus gelada]	Theropithecus gelada	431	431 959	1e-147 96.12%	476 <u>XP_025218966.1</u>	-
Q Search	🔉 🗰 🐂 💽 🔗 🕻		w	^	ENG (⊕ Φ) ■ 5:23 PM IN (↓) ■ 4/17/202	1 E

**BLAST RESULT of MUTATION 1** 

Sr No	Scientific name	Max	Total	Query	E-	Percentage
51.110.		Score	Score	Cover	Value	Identity
1	Homo Sapiens	170	470	1000/	4.00E-	00.50%
	-	4/9	4/9	100%	1/0	99.39%
2	Pan troglodytes	168	168	1000/	2.00E-	07 5404
		408	400	100%	102	71.34%
3	Pan paniscus	468	468	100%	2.00E- 162	97 54%
			700	10070	102 2.00E	/ T.J.T /0
4	Gorilla	468	468	100%	2.00E- 162	97.54%
	Nomesque				2 00E	
5	Leucogenya	466	466	100%	2.00E- 161	96.72%
		-			1 00F-	
6	Hylobates moloch	464	464	100%	160	96.31%
					8.00E-	
7	chlorocebus sebaeus	431	431	95%	148	96.12%
0					1.00E-	
ð	papio anubis	431	431	95%	147	96.12%
9	Pongo pygmaeus	]			2.00E-	
	rongo pyginaeus	463	463	100%	160	96.31%
10	Cebus imitator	J			3.00E-	
10		430	430	95%	147	95.69%
	J					

MUTATION 2 at position 19 repalace L with E, at position 30 replace M with E

VTLARTDLEMQIEGLKEEEAYLRKNHEEEELALRGQTGGDVNVEMDAAPGVDLSRI LNEMRDQYEQMAEKNRRDAETWFLSKTEELNKEVASNSELVQSSRSEVTELRRVLQ GLEIELQSQLSTKASLENSLEETKGRYCMQLSQIQGLIGSVEEQLAQLRCEMEQQSQE YQILLDVKTRLEHEIATYRRLLXGEDAHLSSQQASGQSYSSREVFTSSSSSSSRQTRPI LKEQSSSSFSQGQSS

pp: 2849 – 2865		
🛛 😽 Fw: FIITJEE ::: Student details for 🖹 🗙 📔 📉 Inbox (39) - vaze.career@gmail.c: 🗙 📔 🐴 My Drive - Google Drive	× 😣 Protein BLAST: search protein dat × +	∨ - 0 X
← → C	DC=blasthome	🖻 🖈 🖻 🛊 🛯 🔕 :
BLAST <sup>®</sup> » blastp suite	Home Recent Results	Saved Strategies Help
blastn blastx tblastn tblastx Standard P	rotein BLAST	
BLASTP programs search protein d	atabases using a protein query, more	Reset page Bookmark
DLSRILNEMROYEOMAEKNRDAETWELSKTEELNKEVASNSELVQSSRS EVTELRRVLQGLEIELQSQLSTKASLENSLEETKGRYCMOLSQIQGLIGSVEE QLQQLRCEIMEQQSQEYQILLDVKTRLEHEINTYRRLLXGEDAHLSQQQASG QSYSSREVFTSSSSSSRQTRPILKEQSSSSFSQGQSS		
Or, upload file Choose File No file chosen		
Job Title		
Enter a descriptive title for your BLAST search V     Align two or more sequences V		
Choose Search Set		
Databases (nr etc.): New C Experimental databases	y experimental clustered nr database Q or more info see What is clustered nr?	
Compare Select to compare standard and experimental database ?		
Standard		
Database         Non-redundant protein sequences (nr)           Image: Contract of the sequences of the sequences (nr)         Image: Contract of the sequences of th		
Organism Optional Enter organism name or idcompletions will be suggested	clude Add organism	
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown 😯		
Exclude Models (XM/XP) Non-redundant RefSeq proteins (WP) Unculture	d/environmental sample sequences	THE COLD14
Q Search	) 🗰 📮 😋 💇 🖪 🚊 🖏	^ ENG ⊕ Φ) ↔ 6:01 PM € IN ↔ 4/17/2023

# BLAST WINDOW with search query showing MUTATION 2 at position 19 replace L with E, at position 30 replace M with E

🛛 🗙 Fw: FIITJEE ::: Student details for T 🗙 📔 🗙 Inbox (39) - vaze.career@gmail.c: 🗙 📔 🐥 My Drive - Google Drive	× 🛿 NCBI Blast:Protein Se	equence	× +			× –	οx
← → C 🌲 blast.ncbi.nlm.nih.gov/Blast.cgi					Ê	x 🗉 🛊	<b>□ ⑤</b> ÷
Select all 100 sequences selected	GenPept Graphics L	Distance	tree of result	s <u>Multiple</u>	<u>ilignme</u>	ent MSA Viewe	1
Description	Scientific Name	Max Score	Total Query Score Cover	E Per. Value	Acc. Len	Accession	
✓ keratin [Homo sapiens]	Homo sapiens	475	475 100%	1e-168 99.18	% 244	AAB30058.2	
V type I keratin 16 [Homo sapiens]	Homo sapiens	467	467 100%	5e-162 97.54	% 473	AAB35421.1	
keratin, type I cytoskeletal 16 [Homo sapiens]	Homo sapiens	467	467 100%	6e-162 97.54	% 473	NP_005548.2	
keratin 16 [Pan troglodytes verus]	Pan troglodytes verus	465	465 100%	3e-161 97.13	% 461	BAF62417.1	
keratin, type I cytoskeletal 16 [Pan paniscus]	Pan paniscus	465	465 100%	4e-161 97.13	% 469	XP_003813891.1	
keratin, type I cytoskeletal 16 [Pan troglodytes]	Pan troglodytes	465	465 100%	4e-161 97.13	% 469	NP_001138310.1	
keratin, type I cytoskeletal 16 [Gorilla gorilla gorilla]	<u>Gorilla gorilla gorilla</u>	465	465 100%	5e-161 97.13	% 473	XP_004041720.2	
keratin, type I cytoskeletal 16 isoform X2 [Nomascus leucogenys]	Nomascus leucogenys	463	463 100%	3e-160 96.31	% 476	XP_003279481.2	
keratin, type I cytoskeletal 16 [Callithrix jacchus]	Callithrix jacchus	440	440 100%	2e-151 93.44	% 467	XP_002748663.3	
keratin, type I cytoskeletal 16 [Rhinopithecus roxellana]	Rhinopithecus roxellana	430	430 95%	3e-147 95.71	% 480	XP_010357842.2	
keratin, type I cytoskeletal 16 [Trachypithecus francoisi]	Trachypithecus francoisi	429	429 95%	3e-147 95.71	% 472	XP_033038936.1	
PREDICTED: keratin, type I cytoskeletal 16 [Colobus angolensis palliatus]	<u>Colobus angolensis palliatus</u>	431	431 95%	4e-147 95.71	% 500	XP_011785156.1	
PREDICTED: LOW QUALITY PROTEIN: keratin, type I cytoskeletal 16 [Rhinopithecus bieti]	Rhinopithecus bieti	430	430 95%	4e-147 95.71	% 480	XP_017735629.1	
keratin, type I cytoskeletal 16 isoform X2 [Piliocolobus tephrosceles]	Piliocolobus tephrosceles	428	428 95%	8e-147 95.28	% 448	XP_026304479.1	
keratin, type I cytoskeletal 16 [Zalophus californianus]	Zalophus californianus	428	428 100%	1e-146 90.57	% 476	XP_027423980.1	S K
keratin, type I cytoskeletal 16 isoform X2 [Saimiri boliviensis boliviensis]	Saimiri boliviensis boliviensis	428	428 95%	1e-146 95.69	% 475	XP_003942841.2	edba
keratin, type I cytoskeletal 16 isoform X1 [Piliocolobus tephrosceles]	Piliocolobus tephrosceles	428	428 95%	1e-146 95.28	% 476	XP_023059895.2	Fe
keratin, type I cytoskeletal 16 isoform X2 [Chlorocebus sabaeus]	Chlorocebus sabaeus	428	428 95%	1e-146 95.69	% 476	XP_008010896.1	
https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=39432	Theropithecus gelada	427	427 95%	3e-146 95.69	% 476	XP_025218966.1	•
Q Search	🍿 📮 🤤 🔮 🖬	ļ	💶 🔀			<sup>6</sup> ⊚ ¢0) ⊡ ,	6:04 PM

**BLAST RESULT of MUTATION 2** 

Sr. No.	Scientific name	Max	Total	Query	E-	Percentage
		Score	Score	Cover	Value	Identity
1	Homo Sapiens	475	475	100%	1.00E- 168	99 18%
			175	10070	4.005	22.1070
2	Pan troglodytes	465	465	100%	4.00E- 161	97.13%
2	Dan naniaana	-			4.00E-	
5	Pan paniscus	465	465	100%	161	97.13%
4	Gorilla			1000	5.00E-	
		465	465	100%	161	97.13%
5	Nomascus	163	163	100%	3.00E-	96 31%
	Leucogenya	+03	-UJ	10070	1.005	20.3170
6	Hylobates moloch	464	464	100%	1.00E- 160	96.31%
		J			8.00E-	
7	chlorocebus sebaeus	431	431	95%	148	96.12%
8					1.00E-	
	papio anubis	431	431	95%	147	96.12%
9	Pongo pygmaeus	642	642	1000/	2.00E-	06 210/
		043	043	100%	100	90.31%
10	Cebus imitator	430	430	95%	3.00E- 147	95.69%
		.20		2010		20.0210

MUTATION 3 at position 7 replace D with L, at position 13 replace E with G

VTLARTLLEMQIGGLKEELAYLRKNHEEEMLALRGQTGGDVNVEMDAAPGVDLSRI LNEMRDQYEQMAEKNRRDAETWFLSKTEELNKEVASNSELVQSSRSEVTELRRVLQ GLEIELQSQLSTKASLENSLEETKGRYCMQLSQIQGLIGSVEEQLAQLRCEMEQQSQE YQILLDVKTRLEHEIATYRRLLXGEDAHLSSQQASGQSYSSREVFTSSSSSSSRQTRPI LKEQSSSSFSQGQSS

2865	
ent details for 🗽 x 🛛 😽 Inbox (40) - vaze.career@gmail.c 🗴 🖌 🔥 My Drive - Google Drive 🛛 x 😵 Protein BLAST: search protein da x 🕇 +	v - 0
blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome	🖻 🖈 🖻 🛊 🖬 🔕
National Library of Medicine National Center for Biotechnology Information	Log in
» blastp suite Home Recent Resulf	ts Saved Strategies Help
Standard Protein BLAST	
blastp blastx tblastn tblastx	
BLASTP programs search protein databases using a protein query. more	Reset page Bookmark
Sequence	
number(s), gi(s), or FASTA sequence(s) 😯 Clear Query subrange 😯	
reQMAEKNRRDAETWFLSKTEELNKEVASNSELVQSSRS ↑ SELQSQLSTKASLENSLETKGRYCMQLSQLQGLGSVEE SSSSSRQTRPILKEQSSSSFSQQQSS ↑ To	
Choose File No file chosen	
Enter a descriptive title for your BLAST search ?	
nore sequences 🕜	
rch Set	
Standard databases (nr etc.):     Image: O Experimental databases     For more info see What is clustered nr?	
Select to compare standard and experimental database 😧	
Non-redundant protein sequences (nr)	
in i	2865

# BLAST WINDOW with search query showing MUTATION 3 at position 7 replace D with L, at position 13 replace E with G

Fw: FIITJEE ::: Student details for T 🗙 📔 🗙 Inbox (40) - vaze.career@gmail.c: 🗙 📔 🙏 My Drive - Google Driv	e 🗙 😒 NCBI Blast:Protein Se	equence	× +			× –	Ð
C 🔒 blast.ncbi.nlm.nih.gov/Blast.cgi					Ŀ	2 🖈 🛡 🛸	•
Sequences producing significant alignments	Downloa	id V	Select c	olumns	Y Show	v <u>100</u> ▼ 🔮	
Select all 100 sequences selected	<u>GenPept Graphics [</u>	Distance	tree of resu	ilts <u>Mu</u>	ltiple aligni	ment MSA Viewe	ſ
Description	Scientific Name	Max Score	Total Query Score Cover	E value	Per. Ac Ident Le	Accession	
keratin [Homo sapiens]	Homo sapiens	475	475 100%	1e-168	99.18% 24	44 <u>AAB30058.2</u>	
✓ type I keratin 16.[Homo sapiens]	Homo sapiens	466	466 100%	2e-161	97.54% 4	73 <u>AAB35421.1</u>	
keratin, type I cytoskeletal 16 [Homo sapiens]	Homo sapiens	466	466 100%	2e-161	97.54% 4	73 <u>NP_005548.2</u>	
keratin 16 [Pan troglodytes verus]	Pan troglodytes verus	464	464 100%	8e-161	97.13% 4	61 <u>BAF62417.1</u>	
keratin, type I cytoskeletal 16 [Pan paniscus]	Pan paniscus	463	463 100%	1e-160	97.13% 4	9 <u>XP_003813891.1</u>	
keratin, type I cytoskeletal 16 [Gorilla gorilla gorilla]	<u>Gorilla gorilla gorilla</u>	464	464 100%	1e-160	97.13% 4	73 <u>XP_004041720.2</u>	
keratin, type I cytoskeletal 16 [Pan troglodytes]	Pan troglodytes	463	463 100%	2e-160	97.13% 4	9 <u>NP_001138310.1</u>	
keratin, type I cytoskeletal 16 isoform X2 [Nomascus leucogenys]	Nomascus leucogenys	461	461 100%	1e-159	96.31% 4	76 <u>XP_003279481.2</u>	
keratin, type I cytoskeletal 16 [Callithrix jacchus]	Callithrix jacchus	440	440 100%	2e-151	93.44% 4	7 <u>XP_002748663.3</u>	
PREDICTED: LOW QUALITY PROTEIN: keratin, type I cytoskeletal 16 [Rhinopithecus bieti]	Rhinopithecus bieti	430	430 95%	4e-147	95.71% 4	30 <u>XP_017735629.1</u>	
keratin, type I cytoskeletal 16 [Rhinopithecus roxellana]	Rhinopithecus roxellana	429	429 95%	5e-147	95.71% 4	30 <u>XP_010357842.2</u>	
keratin, type I cytoskeletal 16 [Trachypithecus francoisi]	Trachypithecus francoisi	429	429 95%	6e-147	95.71% 4	72 <u>XP_033038936.1</u>	
PREDICTED: keratin, type I cytoskeletal 16 [Colobus angolensis palliatus]	Colobus angolensis palliatus	430	430 95%	6e-147	95.71% 5	00 <u>XP_011785156.1</u>	
keratin, type I cytoskeletal 16 isoform X2 [Piliocolobus tephrosceles]	Piliocolobus tephrosceles	427	427 95%	1e-146	95.28% 4	48 <u>XP_026304479.1</u>	
keratin, type I cytoskeletal 16 isoform X2 [Chlorocebus sabaeus]	Chlorocebus sabaeus	428	428 95%	2e-146	95.69% 4	76 <u>XP_008010896.1</u>	
✓ keratin, type I cytoskeletal 16 [Zalophus californianus]	Zalophus californianus	428	428 100%	2e-146	90.57% 4	76 <u>XP_027423980.1</u>	I
keratin, type I cytoskeletal 16 isoform X1 [Piliocolobus tephrosceles]      last.ncbi.nlm.nih.gov/Blast.cri#alnHdr XP 011785156	Piliocolobus tephrosceles	428	428 95%	2e-146	95.28% 4	76 <u>XP_023059895.2</u>	
Q Search	) 🐗 🐂 💽 🤗 🖻			2 440	^ G	ENG (1) 10 4	6:29 PI

**BLAST RESULT of MUTATION 3** 

Sr. No.	Scientific name	Max	Total	Query	E-	Percentage
		Score	Score	Cover	Value	Identity
1	Homo Sapiens	175	175	1000/	1.00E-	00 190/
		475	473	100%	100	99.18%
2	Pan troglodytes	463	463	100%	2.00E- 160	97 13%
			105	10070	1.005	77.1370
3	Pan paniscus	463	463	100%	1.00E- 160	97.13%
					1 00F	
4	Gorilla	464	464	100%	1.00L- 160	97.13%
	Nomascus				1.00E-	
5	Leucogenya	461	461	100%	159	96.31%
6	II. Johataa malaah				4.00E-	
0	Hylobates molocn	427	427	95%	146	95.69%
7		1			2.00E-	
7	chlorocebus sebaeus	428	428	95%	146	95.69%
8					3.00E-	
	papio anubis	427	427	95%	146	95.69%
9	Pongo pygmaeus	407	407	050/	4.00E-	05 (0)
		427	427	93%	140	93.09%
10	Cebus imitator	126	126	95%	1.00E- 145	95 26%
		420	420	<b>7</b> J 70	143	73.2070

### **Culminated Tabular representation of Results**

Max

Score

Sr. No.	Output: Scientific name	Original	Mutation1	Mutation2	Mutation3
1	Homo Sapiens	482	479	475	475
2	Pan troglodytes	471	468	465	463

pp. 2045	2803				
3	Pan paniscus	471	468	465	463
4	Gorilla	471	468	465	464
5	Nomascus				
5	Leucogenya	469	466	463	461
6	Hylobates moloch	467	464	464	427
7	chlorocebus sebaeus	467	431	431	428
8	papio anubis	466	431	431	427
9	Pongo pygmaeus	466	463	643	427
10	Cebus imitator	464	430	430	426

# Total

Score

Sr. No.	Output: Scientific name	Original	Mutation1	Mutation2	Mutation3
1	Homo Sapiens	482	479	475	475
2	Pan troglodytes	471	468	465	463
3	Pan paniscus	471	468	465	463
4	Gorilla	471	468	465	464
5	Nomascus				
5	Leucogenya	469	466	463	461
6	Hylobates moloch	467	464	464	427
7	chlorocebus sebaeus	467	431	431	428
8	papio anubis	466	431	431	427
9	Pongo pygmaeus	466	463	643	427
10	Cebus imitator	464	430	430	426

# Query Cover

Sr.	Output: Scientific name	Original	Mutation1	Mutation2	Mutation3
-----	-------------------------	----------	-----------	-----------	-----------

No.					
1	Homo Sapiens	100%	100%	100%	100%
2	Pan troglodytes	100%	100%	100%	100%
3	Pan paniscus	100%	100%	100%	100%
4	Gorilla	100%	100%	100%	100%
5	Nomascus Leucogenya	100%	100%	100%	100%
6	Hylobates moloch	100%	100%	100%	95%
7	chlorocebus sebaeus	100%	95%	95%	95%
8	papio anubis	100%	95%	95%	95%
9	Pongo pygmaeus	100%	100%	100%	95%
10	Cebus imitator	100%	95%	95%	95%

# E-Value

Sr. No.	Output: Scientific name	Original	Mutation1	Mutation2	Mutation3
1	Homo Sapiens	2.00E- 171	4.00E- 170	1.00E- 168	1.00E- 168
2	Pan troglodytes	9.00E- 164	2.00E- 162	4.00E- 161	2.00E- 160
3	Pan paniscus	1.00E- 163	2.00E- 162	4.00E- 161	1.00E- 160
4	Gorilla	1.00E- 163	2.00E- 162	5.00E- 161	1.00E- 160
5	Nomascus Leucogenya	7.00E- 163	2.00E- 161	3.00E- 160	1.00E- 159
6	Hylobates moloch	6.00E- 162	1.00E- 160	1.00E- 160	4.00E- 146
7	chlorocebus sebaeus	7.00E- 176	8.00E- 148	8.00E- 148	2.00E- 146

8	papio anubis	1.00E- 161	1.00E- 147	1.00E- 147	3.00E- 146
9	Pongo pygmaeus	1.00E- 161	2.00E- 160	2.00E- 160	4.00E- 146
10	Cebus imitator	9.00E- 161	3.00E- 147	3.00E- 147	1.00E- 145

Percentage Identity

Sr. No.	Output: Scientific name	Original	Mutation1	Mutation2	Mutation3
1	Homo Sapiens	100%	99.59%	99.18%	99.18%
2	Pan troglodytes	97.95%	97.54%	97.13%	97.13%
3	Pan paniscus	97.95%	97.54%	97.13%	97.13%
4	Gorilla	97.95%	97.54%	97.13%	97.13%
5	Nomascus Leucogenya	97.13%	96.72%	96.31%	96.31%
6	Hylobates moloch	96.72%	96.31%	96.31%	95.69%
7	chlorocebus sebaeus	96.31%	96.12%	96.12%	95.69%
8	papio anubis	96.31%	96.12%	96.12%	95.69%
9	Pongo abelii	96.72%	96.31%	96.31%	95.69%
10	Cebus imitator	96.31%	95.69%	95.69%	95.26%

#### Conclusion

Because of the simultaneous creation of massive amounts sequencing initiatives & bioinformatics programs such as BLAST, researchers are now able to evaluate the genetic material of a wide diversity of different species. This has helped bioinformatics become a discipline that bridges the gap between computer science and biology, which is essential to the field's further growth. As a result, the use of bioinformatics techniques to the analysis of genome sequence data has developed into a crucial component of contemporary biological research. In the field of bioinformatics, BLAST is one of the most well-known methods, and because it can be accessed by any researcher on the internet, it is frequently utilized to classify sequences according to their functional and taxonomic properties. Its applications

The Ciência & Engenharia - Science & Engineering Journal ISSN: 0103-944X Volume 11 Issue 1, 2023 pp: 2849 – 2865 range from the examination of raw sequence data to the examination of genomic comparisons, and they frequently expand into sequence-based data.

#### Future Scope

It is possible to evaluate evidence of evolution as well as extract valuable patterns in gene expression, as well as classify protein structures, make gene predictions, recognize genes, and diagnose various disorders that affect how genes are expressed, amongst other things. The ability to examine bioinformatics information is provided by data mining, which is also important for the recognition of patterns, categorization of data, for predicting outcomes, or the induction within genetic networks.

#### References

- [1] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search** tool. *J Mol Biol* 1990, 215:403-410.
- [2] NCBI BLASt [http://www.ncbi.nlm.nih.gov/BLAST/]
- [3] WU-BLAST [http://blast.wustl.edu/]
- [4] Baxevanis AD, Ouellette BFF (eds): *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. John Wiley; 1998.
- [5] Durbin R, Eddy S, Krogh A, Mitchison G: *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge: Cam- bridge University Press; 1998.
- [6] Higgins D, Taylor W (eds): *Bioinformatics: Sequence, Structure and Databanks*. New York: Oxford University Press; 2000.
- [7] Kanehisa M: Post-Genome Informatics. New York: Oxford University Press; 2000.
- [8] Gibas L, Jambeck P: *Developing Bioinformatics Computer Skills*. Sebastopol, California: O'Reilly and Associates; 2001.
- [9] Wake DB: Comparative terminology. *Science* 1994, 265:268-269.
- [10] Wake DB: Homoplasy, homology and the problem of 'same- ness' in biology. *Novartis Found Symp* 1999, **222**:24-33.
- [11] Reeck GR, de Haen C, Teller DC, Doolittle RF, Fitch WM, Dicker- son RE, Chambon P, McLachlan AD, Margoliash E, Jukes TH, et al.: "Homology" in proteins and nucleic acids: a terminology muddle and a way out of it. Cell 1987, 50:667
- [12] Pearson WR, Lipman DJ: Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 1988, 85:2444-2448.
- [13] Altschul SF, Boguski MS, Gish W, Wootton JC: Issues in searching molecular sequence databases. Nat Genet 1994, 6:119-129.
- [14] Pearson WR: Searching protein sequence libraries: compari- son of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* 1991, 11:635-650.
- [15] Koski LB, Golding GB: The closest BLAST hit is often not the nearest neighbor. J Mol Evol 2001, 52:540-542.
- [16] Henikoff S, Henikoff JG: Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 1992, 89:10915-10919.

pp: 2849 – 2865

- [17] Dayhoff MO, Schwartz RM, Orcutt BC: A model of evolutionary change in proteins. In: *Atlas of Protein Sequence and Structure*, vol.
- [18] 5. Edited by Dayhoff MO. Washington DC: National Biomedical Research Foundation; 1978:345-352.
- [19] States DJ, Gish W, Altschul SF: Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. *Methods: A Companion to Methods in Enzymology* 1991, 3:66-70.
- [20] Henikoff S, Henikoff JG: Protein family classification based on searching a database of blocks. *Genomics* 1994, **19**:97-107.
- [21] Henikoff S, Henikoff JG: Automated assembly of protein blocks for database searching. *Nucleic Acids Res* 1991, **19:**6565-6572.
- [22] NCBI FTP directory BLAST matrices
- [23] [ftp://ncbi.nlm.nih.gov/blast/matrices]
- [24] Needleman SB, Wunsch CD: A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 1970, **48**:443-453.
- [25] Smith TF, Waterman MS: Identification of common molecular subsequences. J Mol Biol 1981, 147:195-197.
- [26] Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL: GenBank. Nucleic Acids Res 2000, 28:15-18.
- [27] GenBank [http://www.ncbi.nlm.nih.gov/Genbank/]
- [28] Bairoch A, Apweiler R: The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 2000, 28:45-48.
- [29] SWISS-PROT [http://www.expasy.ch/sprot/]
- [30] Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new genera- tion of protein database search programs. *Nucleic Acids Res* 1997, 25:3389-3402.
- [31] Karlin S, Altschul SF: Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc Natl Acad Sci USA* 1993, **90:**5873-5877.
- [32] Lamperti ED, Kittelberger JM, Smith TF, Villa-Komaroff L: Corrup- tion of genomic databases with anomalous sequence. *NucleicAcids Res* 1992, **20**:2741-2747.
- [33] Kristensen T, Lopez R, Prydz H: An estimate of the sequencing error frequency in the DNA sequence databases. *DNA Seq* 1992, **2:**343-346.
- [34] Full list of the BLAST Advanced options
- [35] [http://www.ncbi.nlm.nih.gov/BLAST/full\_options.html]
- [36] States DJ, Botstein D: Molecular sequence accuracy and the analysis of protein coding regions. *Proc Natl Acad Sci USA* 1991,88:5518-5522.
- [37] Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al.: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 1995, 269:496-512.
- [38] Ichikawa T, Suzuki Y, Czaja I, Schommer C, Lessnick A, Schell J, Walden R: Identification and role of adenylyl cyclase in auxin signalling in higher plants. *Nature* 1997, 390:698-701.

- pp: 2849 2865
- [39] Ichikawa T, Suzuki Y, Czaja I, Schommer C, Lessnick A, Schell J, Walden R: Identification and role of adenylyl cyclase in auxinsignalling in higher plants. *Nature* 1998, 396:390.
- [40] Karlin S, Altschul SF: Methods for assessing the statistical sig-nificance of molecular sequence features by using general scoring schemes. Proc Natl Acad Sci USA 1990, 87:2264-2268.
- [41] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3007598/