# Group4 Assignment-1 Kushagra Agrawal B20296

# Q1

# Seq1

MLLPLLFVACALLAPARAKVYERCELARELMHSHDVPRDQLAVWVCIARHESEYNTTAVGHVNGDGSGDHGIFQISDLYWCSHQGRGKACGMACSELLDDNIADDMQCAKRIHREHQGLSGNGFNAWAVYAQHCKRPERAAKYLEGCFSGTGNEIQQPHEAYGLPPPAITAPPPRARPTAAPVARPVRPAASYGLPPPAAPARRWYAPAAPRPAVVAPVPRFSWQAAPAPPPRPTLRSNPQRVAADPRNFWNWQRGATRRWF

# Mutant 1

MLLZLLFVACALLAPARAMYYERCELARELMHSHDVPRDQLTVWMCISVHESEYNTTAVGHVNPDGZGDHGIFQINDLYWCSHQGRGKACGMACSELLDDNIADDMQCAKRIHRERQPLSGNGFNAWAVYAQHCKRHERAAKYLEGCFSGTGNEIQKQHEAYGLXPPAITAPPPRARPTAAPVARPCRPAASYRLPPPAAPARRWIAPAAPRPAFVAPVQRFSWQAAPIMAPRPSLRSNPQRVAADPRNFWNAQRGATRRWF

# Mutant 2

MLLPITFVACALLAWARAKVYSRCELARELMHSHDVYRDNLAVWVIVAOHNHEYNTTAVGHVNGDRSNAMGIAQISDLYWCSHQGIWKACGDUCSELLDDNIOCDMQCAKRGHRXUQGVSGKGCNAWAVOAQHSKRPERVAKYLEVCFSGEFNEEQTPHEAYGLPPPAITAPPPRAPXTAAPVAXPVPPAASYGLPPPAAPARPWYAPAAPRPAZVSPVPRFSWQAAPAPPPLZTLRSNPNRYAMDXGNMWNWQRGATRYWF

# Mutant 3

RLLPSUYZAPCLLAPARAKVYTYDELARELMHSHUVPRPSLLHWWUIAFWASOXNTTAVZHVNGNGSGPHGIFQISDLYDCSHAGRGKACGMAQSELALDNGYNDMQCAOGIPREHQGMUGNGKLZWLVYAKHZZRNERNASNLFGQFSGTZNEIQFPHOKYGKPPPRIIAPPPRAVPTAAPKRRPVKPAASYGLPPCAAPARRWPUPAFPSXRVVEOXYRFSDQAAPAIPPRPTLRSNPQRMAADPRNFWNWCRGFTRRWF

# Mutant 4

MKLPOZFVKDKLLAOWEQDZYARCEXAEECMHROQVPRDQLXVDVCRARHNSUYNUTXVGHINGDNSGDLFAFKITRLYRCSRTGRGKGCGHYCSRLLSDNXICDMQCFKRQHREHQWLSKFGONAWAVKIVHCKQPERAAKYMEGCMNGEDNEIXFCHEAYTLPPPCITAUPPRSVPTXAPVARPVRPAYQYLTIGMAANAROOGEPHTPRPCVVAPVPIGVWQAAPAGWGRPSHRANQPRVGAOPTRFCNWNRYAORRLF

# Mutant 5

MLLPDVLVATPXLAWARAKGEEKCECDXELXISSVVCRMALASWKXIZRQESEFGSEQVZHZGMDGSYLRGIFQISDLPWRSKOGXGFAFGMWCSPQLDENMGDQZACUKFIDDEHQWLUDOCUYLWIVYADHZYRGENRAKMLLICFKGXPXEIQQPCGMLGCKZPAVVALHQRARPTMCPVARPVRPDWAKGSPHPXVPAURWORPAPPRMAVVRSTERZRCGCAPAPCAAKILQSNOQEVZDPURNFWLXQRGAFHRWF

# Mutant 6

DANXLSFFYCARLUPIRAZGYWRCEVKRELKUOHEPGAMAWFVWNPIAMHCHEYNTLTVHHVUGIQUGYHZUHMISXXPPHSHWYRGPXYGAVAKALLDYNLATEGXFATRISREHZGLUFCDIZAWZVYHQKEUDPVUXAFYLEQCCRGUGUEETQPYEGIGPNOELWTAPPONYHPTAIWGAZEULPADSKGHRDVNEPNLRWYAXRAPRHWVAAAEPVFSTQXPUEPEWRACPERNTSGEGAZMRYNWHYQRPMVHHWF

# Mutant 7

EWKLFFLFACFPLHPCMTDCGEQCWLVIRUMPKHSVPPMQYGLWUIHTTFHSECYTPQQAPOTGNXNTDYKIDQIWEGULMSHCLASKXXCMACSELQWDXPEDWZQPVEPXZFEHHQFXYUGFYZNATLATRULVDHVAAKYGVGXESFKQYRIAQRHEWKIXVKKGIDRCEZMWNKHOAQWARDVDLZAPRELPPFHAPXFRQUAMAUPMCAQXCVLWRFPOQKAUAPVPKTCRZNNIAGSAHZYAORYTWQUGEPRRPN

# Mutant 8

UGUINTPZICANYEAWDVWVOQDHFLYYKPSAOMOZTPRYOCVTTEIRLIOHEYAVGSOCZEAADFMGDRRLGQIVISWWKSMQQRMMYCLTACOZYLOGGAADNPQGANVINRVHFPZIGQGIUAUQDZQGHWVPPOIGTAWXEEXLSGDTQDIQUPTKSYCFRULZIKOXVPOYPXEWADZGRETDDDSXOPNYIPXFGAUMUVLYIPHHHPTCZPICRZWGWATUUHMPOEOPXSGZKEVQMEGWYTCYVSDCYIRVOG

# Mutant 9

VRUPLGANWSAFPSOTEVRWHUZACAWMKYNRTCDAELPMXSZPEXQLUVKVECLPHIZUQVQADWCTRWWNLKMNLNUHEMHZAXKVFCWXQKLEYYNYPUIOVFTFNHTIWIDUQPOFAQGVOWMMZNRQFQDKPCOAAAWPFIOTZQGLVNVMHKVCGHKVLDHZWTWWPKAQRAOXTIADXVCNMGFNLXCQNOLRDDCWWTPQXPPDAUZQWLGROGSCWQCUNGQFRQYZCNWFORLAKDAWYDGEOWSSEEIGPZ

# Seq2

MAHNTEELINKIGSFGRYQVRLFLVASFMVFITTNVLMVMTFSTAEPPWRCTANATSCKLNGTFKPGDKNYEYRCDIPRSDWEFAVDGNFDSIVTEWDLVCKTSTYVSHVNSLAFLFWMLGSIVGSILSDKVGRKKVTFPFFMIVAVSGLISAIAKHYWVFAVFRSLVGIGLGGAGTTAFVLVLEYVGARHRGAVGIGIWFAFVFAMCFLALIAYLLPAWRLLTIVTSAPGLIGFVFWWFTPESLRWLLMKGKSTEAKKTLQKAASVNGKLLVDEDFMLIQEDNDERNIKLGHIGDLFTSRLVAYRTLVSWYCWCVCGMVYYGISLSTPGIGGNMYLNFFISAVCECVGMIIGIFCLDKFGRKKSIAASLWIGGVAMVAAALLSYYDDGSDGYLAGKILMTMVIGKFFITIAFDGIFVHTSELFPTVIRNTALGTSSSSARIGSAASPYIVYSQRVHPLMPFGIMAINALISGILFLSLPETLNRVMPDTVNQNTSACTESPDGGKDVNDDKIGTDL

# Mutant 1

MASNTEELINKIGSFGRYQVRLFLVASNMVFITANVLMVMTFSTAEHPWWHTENATGCKLKGTFKPGDKNYPYRCDIPRIDWEFAVDGNFDSIVTEWDLVCKYSTYVSHVNSLAFLFWMLGSIVGRILSDKVGRKKVTFPNIMIVAVSGLIEAIAKHYWVFAVFRSLVGIGLGGAGTNAICLVLEYVIARHRGAVEIGIWFAVHFAMCFLALIAYQHPAWRLLTIWTCAPGLIGFVFWWFTPESLRVLLMKGKSTEAKKTLQKISSVNGKLLVDEDPMVIQRDNDERNIKLGHIGDLFTSRLVAYRTLVSWYCWCVCGMVYYGISLSTPGIGGNMYLNFFISYZCECXGMIIGISCLDKEGRKKSIAASLWIGGVAMCANALLSYYDDGSDGYLVGKILMTMVIGKFFUTIAFDGIFVITSELFPTVIRNWALGTSSSSARIGSAPSPYIVYSQRVHPLMPFGIMARNALIEGILFLSLPETNNRAAPDTVNQNTXACTESPDGGKDVNDDKIGTDL

# Mutant 2

TAHNTEELINKTGSDGRYGVULFLVASFMVFIYTOVLMVMIFDTAEPPWRCTAUATSCKLNGAFKPGPKNYEYRCDKPRSDWEFAVDGNFISIVTEWDLVCKFSYQQSHVTSLHFLFAMLOSIVGSALSDKVGRKKVTFPFFMIVNVSGLISAIAKHYHVFLVFRSLSGINHGGLGTTAFVLVLEYMLARHRGAVGIWIWGAXVDAMCMLAAIAYLLTRWRLLTIVTVAPGLRGFVVWWFIPESSRWLLMZGKSTEAKKTLQKDASVNGKLLVNEDFNLIQADAAERNIKNMCIGDLFTSRLVAWRTLVSWYCWCVKGMVYYGDSLETIGIGGNMYLNXFINAVCFCVGMIZUIFCLDKFGRQKSIAASLWIGGVAMVAAAFLDSXZPGQGGHLAGKIGMTLVISKFUITVAFDGIFVHTSELTPTVKRNTALGFUYWSARIMSAASPYIHYSURMHPLMPFGIGAINALKQCILFISVPETLNRVMDDTVNQXVSACTESPDGGKDKLGDSIGTDL

# Mutant 3

MUHNREELIPKTGSFGYYQVRLFLQISKMVYITTNOQSVMVKSTXAHPPRCTANGTTZKLENVFKPYNKNYEYRCDHPRSDLEFALAGNFDSIFSEWDLYCKTSDAVVHNOSLTVLLWRLGCUVGSILSHKVGIKYVTFPZZHIGAVSGLISAIAKHYWVFAVFRSLNGIGLGGAMTTEFVYVKEYVLARHRGHDGIGIDEAFVFMMCILVLIAYLLPAWRXVTXETURPNLWADHFWLHTPEKYRWLLMKGKSTEAKMTLQKAAGVNEZHLVDEDFPLIQEDNDERYIXOGHIGDRVHSQLXAYVTKVSWXGWOVCGMVNPGMSLSTAGCGGNMYLNIFISAVCECVDMIIRIMCLDKFORKKSIAASFNIGGDNMVDDALLAYYDDNSDCYLAGSZLMTHVIGHFFITIKDDAUFVHTSEFFPVVIRNTSXGTSSSIARIGSAASPYDVTSQXVHPLMPNGDMAONPLCOAIVFLACPETULRHQPDTVCQNQSDMTDSPDGGIDVNDUKIGHDL

# Mutant 4

MAUNNEEKZUKVCSFGNYQSILFLVASIMVFYTTNVLMQMAFIIAEHQOQTTFNFTSCDLRLDGKQGDKAYEKRWAIPRSDWEUAODGAUDKAVHETDAVYKTZTCTDZVNSLALQVCMLGSEVLIOLFMKVWWOKTTFDSZMSKTVPGFISAIAKHYWVFAIXTSLGGIGLVGAPTOHFVLVLVTOOARHRGAVIWGIUUAFVFPMCDLALIAFLAPALRCLTKUTSVPGLRGFVOWWOTPEKLGWLLMQGKWTEAAKGLQKAASVNXUXOVDHINMLAONVYDVRDIKQGKGQGLWTPRLWLYRTKVSWYTWNECGMVYYPISOSHSGIGKNMYGNFFISCVEGCPGMPIGYVCLDKXGRKTIIADMLWLGGVEQUGIAALSLYDDGSQGYXAGLIUWQMQIMWNFISIAFKGCFVHGSDLWZTVRRNYALGVSSZSARIOAAKSPFRVYNQRVHPLMFFNIZYISKIIHGALVLSLPEILLRAMPDTVQLATSACEFOPDGGKDVNDHKIRQDL

# Mutant 5

MAHNTXZLMFDTNUCGOOQYRDFLGASYMVUWTIVVZYVMRFMTHTEQFRRTAIATVCLLFGROKPKDMNXFCRYINNNSDWPWANDGNFEHGVTEGDNZVATSAAVTRVMPLAQLEWMUFAIVOXDSLDKVQUKLSTFCFFMUVAWGGLISUDVKOKWYSAFDESWVWIGXGKZGSTRYVKVLEYPGMAHRNAVPDGFWRAFUFAYQDOALTAYECPAWPZLFUVTDAPGLIKHVFVEDKMESLRMVGMGDTQFPAKNYMQCAUSVNKKLINDEDFMYIPETNFERNIOMTFIGDHFTQRLVAWMWLVSWYPGKVCGMFYYGOHLSVPNIGQNMAINZFFOAEUXCVGCFIGICHLKKFGDYKUIAUSZWIAKIAMVAAFSLHTQDDCATZYUOGLIZMTMAIGKNYITRAFOXHUCKODOFFPIVIPCYPCGFSSPSARTGSAASPYIXYUQRVPPCHPZEIDRRNWLHSGIORLSIRZTCNRRGFDTVNMNMSACWEXZWGLDDVDNDEFGQDX

# Mutant 6

MAFFYOSQFNKIGSMOFYQPNLFSVAXRMFEIHTNVLMUGEQVMZEOFPMPSYKCESMFYNGZZAPGDKIPQIRTZEGRSDIEFFPDVNYDSIVHVWSWIKKASTYOSIVNLLAGQEELLLSIHGQPASRUVWRGKVTKPFIYIVAXHYHCSOAAYOYWVFKVFCSNVPAZLGGAIGTRUVZLLEYVGFRHRFKNGIVIWCSUOKAMMOYOZMADNRGAMHLDNCSRHPYILIGQCFMXFTGHSLRWDUIPRPSTESILQCPKMQIGDYLQFVDESZLWFERYTDERUOKLOTPIDLRLQTOGQZRTMISZPCEFVPFMMENNOMLOTPMAFESMFSGLFISFTKAZYGMSQLSCCNNAFGWQKWIANELWWGMVIMWEAAWZHZKEDUSFGYZASKTXHQOTRLVFFOUFAGTRFUVHTKOMMPIVERWTVLOPCWSSAZCGSAASSYIVYIQRVVWLIPSOGMQIIFQISYIOAKGCSETANRVMWDNEGQNOSORDPHAUUHKDVYDSAIGTZL

# Mutant 7

TAHVOCELQDLYSCMFKXTHTIGKVACIDVSITTKHPMTYNFWTAEDPFECTZWAMGPSONQHPYPIDHQOERSGMDTRLNWEFWLURTFDSIUFCLOKKTNTFDNSRTVUALAXLFPFLTMIMYSZLYNPVRVVPQTFSSFDFUMQSOXPSHISMAIVFLUVQINYMFZXOMGAPOTANULILZYRRRNTSMCVXCTCONPXQXAFVULEMHSWLXGAWEWLIDALVASPGICMEUWWFDPZSSRIHOXECEIREUOYAWFPZASNVGKUVVHNDRMSSHWDFDHLQCRRERLFDNATKAWDAKXRDAAVYAOTHCNMPSYHQSUPTQXOGAPUQLELFISACCEAOUMSNDLPESDDFGOPHSUWEWWUAIERAMVAFVALSZZXDGVKICLQQKCLMSMVTFSFFIYFAIDSHFLHTGETAIHICVIFIORNSESPAUWGXKFPTYTMIKORNHPMGWFAIAFINOLINGWXFEVLDMTLNSVMWWTFYZFEDSHEASMDWYKDXENCKGITRL

# Mutant 8

MPXUCDEXIUAWGELIMKONRQPHYYXPGVPISOYVLFULPRDWGZLCWIXVPCONQVMLXGSDZYGDKMTIMRUPGPNSDRZFGVADRFDQISSPTTNHGUIAAXMSYHAMLARPLWYWXSUVNEMLWDWAKXYOKUYZFFSUIFKXLPITXZKKHATVHFXKFFGALDGHTVAYFYDFWAVLMWIEUCHHGOSTCXKWHAKKYZGZKHGYYWGXWWAUZLAXSXEPAHKKIXFVISTTOOOQMYVWLGYGQVTEQUKTPGUDIEHHPEGLCOYSFMLZQOLZGMDZGRTLVOFWUAOADIESPTSLVXINIDEONFCVQSEEZYSZXIIQENDWUZGFHIAWCIXCFPGESHVIZDKSXZQUSTZVCLOIXKSGSENCUKQGNKQXGLCGIWLGMYLOFCLLVOYFGTNGGDZHZVHYSEHFKOVHNPSWLTECQWNAFEGHNQLFEIPRVUTCNRHFNAPUKCUHSLKGHOMFIELEATKZRVDGVEFYLNLQXGKXFNPSHNCQQPMVNTIDL

# Mutant 9

AXERAAEAPIDVUZCRLFOMSASOCASAVFMECRRKEXKVIVNOCUQPXDPRSLRMVHFIAACSLPIUDTYMLFAGUMOZPGKZCFMOSFLUQWEUTWAYNWCEHLUWZSCQOMXOFWILXOSEUIXOEAGZMUUWXQPMDUGTWFHAIHNITDFIRCQWSVVSLPVTSIGDVIIOFWQCGLZUIRGGRZRMTHCUVMDTIXYAFARERAOEZFLNHRAKZORMIYTKAPARVIVDMNWIRNFHYRWQSMTOQRXQFEXFSRXUUZNGGKTFKNIZQCGCNGDAFUFNOPPDLPWFEMTZRHLEIRFNDFLTRCCPKOYLIWAQWXTZIHWAOZEYCKTLKLUKYCQPTSAOYENCZCLCOESNYKLFLIOIEATGMSNATNVRCLUGNUSSKETEQATMUYZSEZFOSKIDDFCTISLQOLLCMLPGZUTNCCOSIWDNUOHQPGNIYFHQOUAXHOAMZEEQFXYESXIVLQZLGVEINDSAXECDCCXFUAMXNYTUGOHZAAOTNGYAUETNU

Q2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Protein** | **Organism** | **BLAST Query coverage** | **Percent identity** | **Total score** | **E-value** |
| seq1 | [Lysozyme](https://blast.ncbi.nlm.nih.gov/Blast.cgi) | [Hermetia illucens](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=343691) | 73% | 48.31% | 1012 | 5.00E-52 |
| mutant1\_seq1 | hypothetical protein | [Sarcophaga bullata](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=7385) | 63% | 51.68% | 1045 | 5.00E-43 |
| mutant2\_seq1 | hypothetical protein | [Lucilia cuprina](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=7375) | 65% | 44.44% | 836 | 7.00E-25 |
| mutant3\_seq1 | uncharacterized protein | [Stomoxys calcitrans](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=35570) | 62% | 38.18% | 645 | 9.00E-21 |
| mutant4\_seq1 | uncharacterized protein | [Lucilia cuprina](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=7375) | 64% | 42.61% | 772 | 1.00E-18 |
| mutant5\_seq1 | uncharacterized protein | [Lucilia cuprina](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=7375) | 56% | 33.08% | 400 | 6.00E-08 |
| mutant6\_seq1 | [Lysozyme](https://blast.ncbi.nlm.nih.gov/Blast.cgi) | [Frankliniella occidentalis](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=133901) | 98% | 40.70% | 131 | 1.00E-32 |
| mutant7\_seq1 | - | - |  |  |  |  |
| mutant8\_seq1 | - | - |  | - | - | - |
| mutant9\_seq1 | - | - | - | - | - | - |

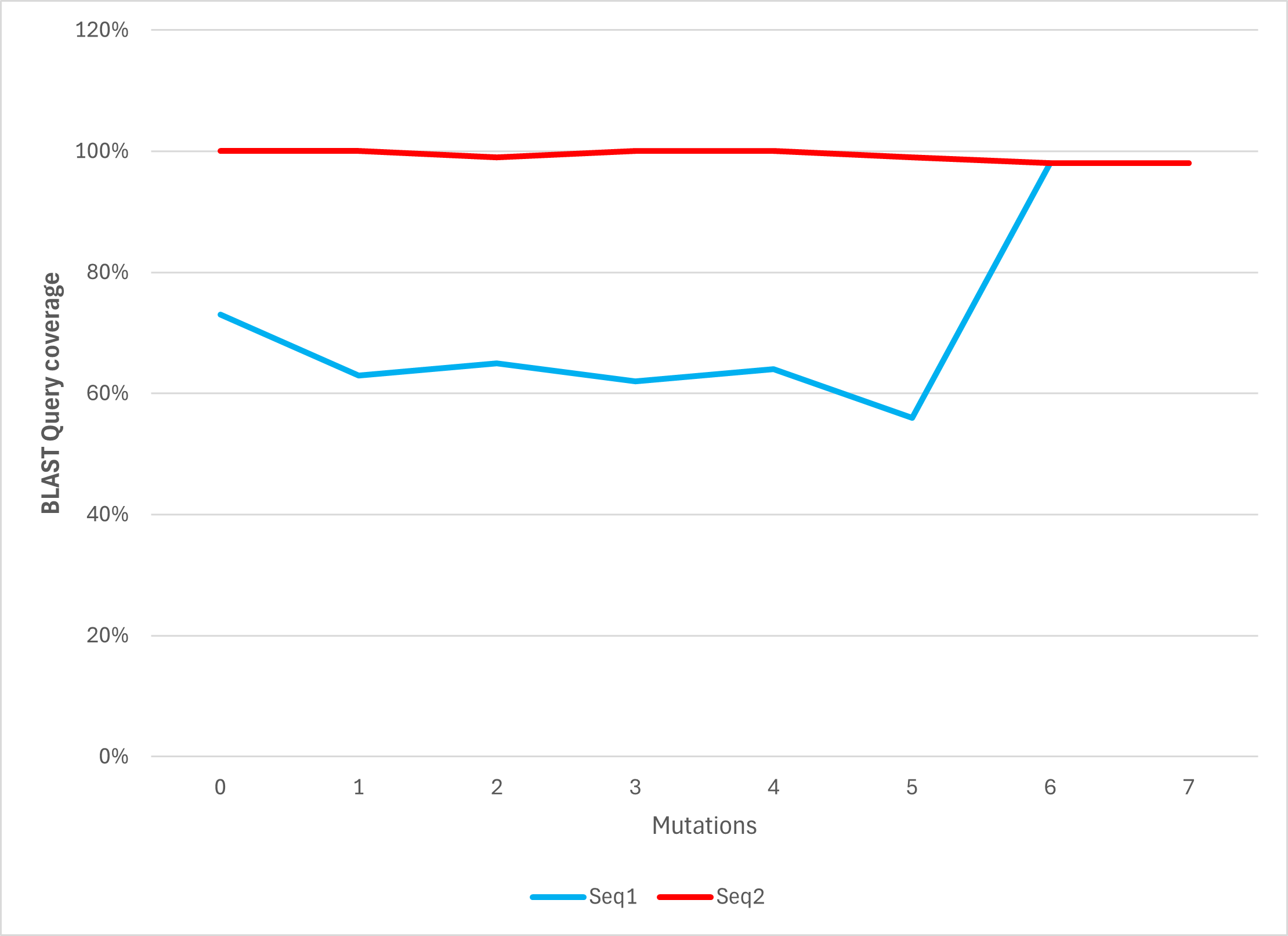
Lysozyme is an enzyme found in tears, saliva, mucus, and other bodily fluids. It plays a key role in the immune system by breaking down the cell walls of bacteria, thereby helping to defend against infections.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Protein** | **Organism** | **BLAST Query coverage** | **Percent identity** | **Total score** | **E-value** |
| seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 100% | 100% | 1057 | 0 |
| mutant1\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 100% | 90.14% | 926 | 0 |
| mutant2\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 99% | 80.23% | 790 | 0 |
| mutant3\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 100% | 70.02% | 679 | 0 |
| mutant4\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 100% | 60.15% | 528 | 0 |
| mutant5\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 99% | 50.19% | 406 | 1.00E-132 |
| mutant6\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 98% | 40.62% | 247 | 3.00E-71 |
| mutant7\_seq2 | organic cation/carnitine transporter 4 | [Exaiptasia diaphana](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2652724) | 98% | 30.74% | 144 | 4.00E-33 |
| mutant8\_seq2 | - | - | - | - | - | - |
| mutant9\_seq2 | - | - | - | - | - | - |

The organic cation/carnitine transporter 4 (OCTN4) is a protein involved in transporting carnitine and other organic cations across cell membranes. It plays a role in various physiological processes, including fatty acid metabolism and energy production.

Q3

0 index defines the original sequence i.e seq1 and seq2



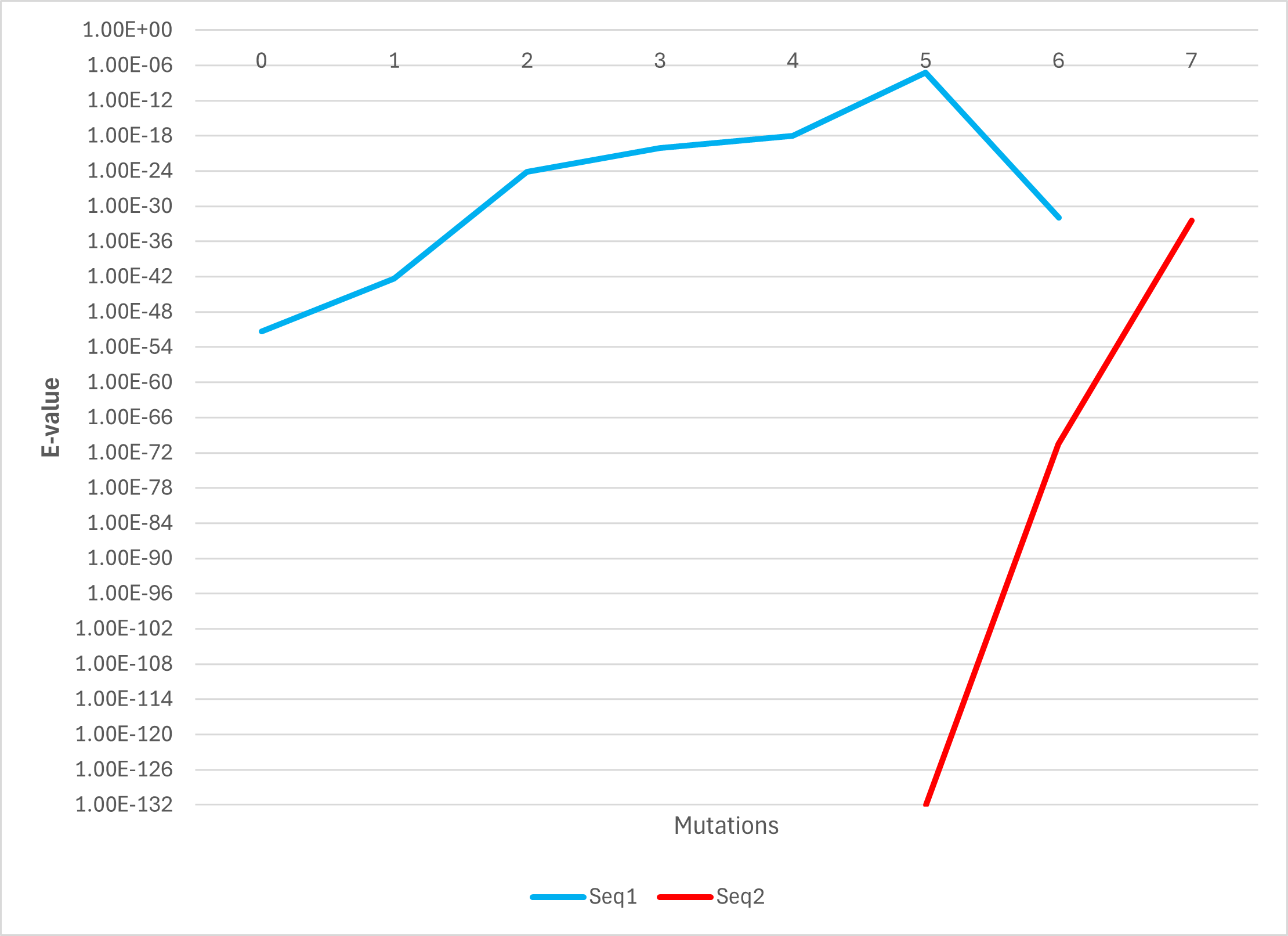
The two sequences do not exhibit the same trend. Seq1 shows a sudden increase in coverage at 4 mutations, while Seq2 maintains a consistently high coverage until a very late stage (after 6 mutations).

These contrasting trends suggest that the underlying sequences or the mutations introduced have different effects on the coverage or the property being measured. Seq1 may require a specific combination of mutations to achieve high coverage, while Seq2 is more robust and maintains high coverage even with multiple mutations.

A graph with red and blue lines

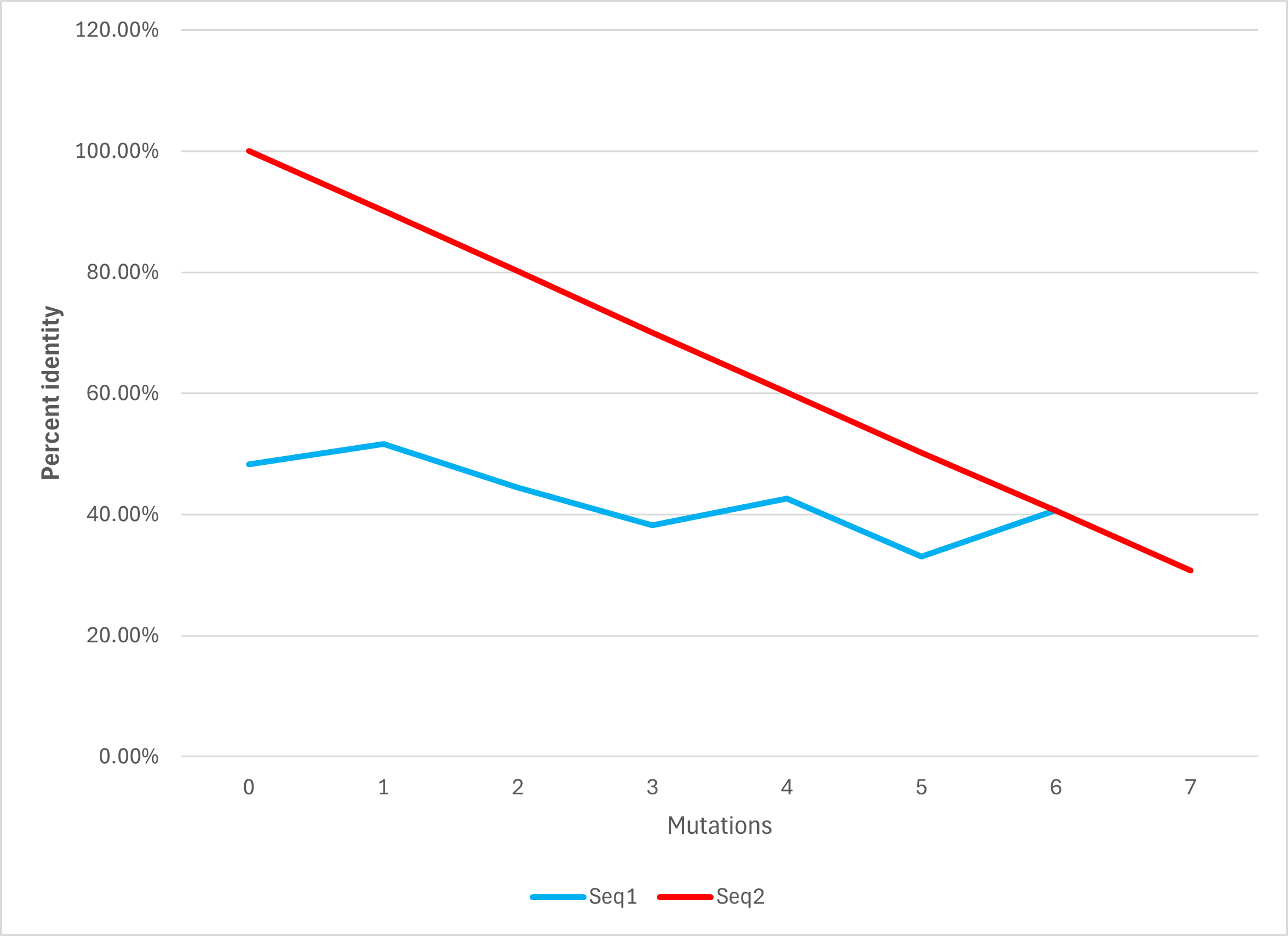
Description automatically generated

As the Hamming distance, which represents the number of differences or mutations between the query sequence and the database sequences, increases, the overall quality or total score of the sequence alignment is expected to decrease. This is because a larger Hamming distance implies fewer identical matches and more mismatches or gaps in the aligned regions. Consequently, the alignments will have a lower score, as they will contain fewer matching positions and more positions with differences or gaps, which contribute negatively to the overall alignment score.



For seq2 E value for mutants 1-5 are 0 that’s why they are not shown in this log function.

The contrasting trends suggest that Seq1 and Seq2 respond very differently to the increasing number of mutations. The contrasting trends suggest that Seq1 and Seq2 respond very differently to the increasing number of mutations. Seq1 exhibits a bell-shaped curve, with an initial increase followed by a decrease, potentially indicating an optimal range of mutations for maximizing the measured value. In contrast, Seq2 maintains a near-constant low value until a specific threshold (6 mutations), after which the value increases exponentially.



As the number of mutations or differences between the query sequence and the sequences in the database increases (i.e., the Hamming distance gets larger), the percentage of identical matches in the aligned regions is expected to decrease. This is because more mutations lead to fewer positions where the bases or residues are identical between the query and the database sequences.

Here seq2 seems more dependent on hamming distance.

Q4

Using Smith Waterman Algorithm:

CALLAPARAKVYERCELAREL

\*\*|\*|\*|\*\*||\*|\*\*\*||\*\*\*

CARLUPIRAZGYWRCEVKREL

Using BLAST:

LLFVACALLAPARAKVYERCELARELMHSHDVPRDQLAVWVCIARHESEYNTTAVGHVNG

L F CA L P RA+ Y RCE+ REL H+ VW IA H EYNT V HV G

LSFFYCARLUPIRAZGYWRCEVKRELKUOHEPGAMAWFVWNPIAMHCHEYNTLTVHHVUG

Two alignments are different but the local alignment we get from smith-waterman can be seen as part of the BLAST alignment