**PRACTICAL ON DOTS PLOTS AND PAIRWISE SEQUENCE ALIGNMENTS**

**STUDENT: PIERRE WENSEL**

1. Use the Needleman-Wunsch algorithm and the Smith-Waterman algorithm with a Jukes-Cantor model (all mutations present the same probability) and a gap penalty of -5 to find the optimal alignment between these two sequences.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **A** | **C** | **G** | **T** |
| **A** | **5** | **-4** | **-4** | **-4** |
| **C** | **-4** | **5** | **-4** | **-4** |
| **G** | **-4** | **-4** | **5** | **-4** |
| **T** | **-4** | **-4** | **-4** | **5** |

**Compute the score of each alignment. Include the alignment matrix at the end of the document.**

AGTCTTCGTAA

CTTGGTA

**Needleman-Wunsch Algorithm:**

A screenshot of a calendar

Description automatically generated

Global N-W Alignment:

AGTCTTCGTAA

\_ \_ \_CTTGGTA \_

Score= (4 gaps\*-5)+ (6 matches \*5) + (1 mismatch \* -4)=6

**Smith-Waterman Algorithm:**

A grid of numbers and yellow squares

Description automatically generated

Local S-W Alignment:

CTTGGTA

CTTCGTA

Score=(6 matches \*5 ) + (1 mismatch \* -4)= 26

Smith-Waterman (local alignment) gave the higher score of 26 and is therefore the optimal alignment.

1. Perform a pairwise sequence alignment between human fibroblast growth factor receptor 3 (P22607) and mouse proto-oncogene c-ret (P35546). Integrate information from a Dot Plot, PFAM and the Sequence alignment to answer the following questions. **Include the Dot Plot, PFAM output and Sequence Aligment at the end of the document.**
2. Global alignment or Local Aligment? Justify your answer.

**LOCAL ALIGNMENT.** To better justify my answer for this question, I first obtained and examined the FAFSA protein amino acid sequence files for both proteins by running query with their Uniprot IDs in Uniprot database. This was to later have a means of inputting for for dot-plot, to examine relative amino acid lengths, and to verify identity of both eukaryotic-derived (containing exons+introns) (Homo sapiens)human fibroblast growth factor receptor 3 (P22607) and (Mus Musculus) Mouse proto-oncogene c-ret (P35546). Based on these files, it was evident that the sequence P22607 is 806 amino acids and is shorter than the 1115 amino acids of sequence P35546. That a global alignment of these sequences of unequal length may result in high dissmiliarity in longer, unmatched regions compared to those obtained via local alignment methods may be one of various reasons why local alignment is more suitable. Furthermore, I scanned both sequences with InterproSCan by entering their text-based UniProt IDs to examine their domain architecture (quality and quantity). InterproScan revealed that Uniprot ID P22607 (Homo Sapiens) contained four (4) PFAM (curated, reviewed) domains with accession numbers as follows:

PF07714 Protein tyrosine and serine/threonine kinase

PF21165 Fibroblast growth factor receptor 3 transmembrane domain

PF13927 Immunoglobulin domain

PF07679 Immunoglobulin I-set domain

Interproscan also revealed that Uniprot ID P35546 (Mus musculus) contained five (5) PFAM domains with accession numbers as follows:

PF07714 Protein tyrosine and serine/threonine kinase

PF17813 RET Cadherin like domain 4

PF17812 RET Cadherin like domain 3

PF00028 Cadherin domain

PF17756 RET Cadherin like domain 1

Based on InterProScan and PFAM Outputs, our two protein sequences appear to share a highly-conserved, common domain PF07714 (Protein tyrosine and serine/threonine kinase) but no other domains. Preliminarily, high “localized” sequence identity between the two proteins is expected in this shared domain region, and this is additional reason as to why local alignment may be more suitable. Based on the PFAM comparison of relative domain locations within the proteins, it appears that the PF07714 domain spans the amino acids 726-1006 for Mus Musculus P35546 while it spans the amino acids 839-1106 for Homo sapiens P22607. The two protein sequences could therefore be more optimally aligned via local alignment dynamic programmin-based algorithms like Smith-Waterman around the shared common, highly conserved and homologous PF07714 domain around which significantly high % identity is expected to be revealed in Dot Plots. Additionally, it appears that P35546 possesses repeats of highly similiar cadherin domains such that high sequence similarity throughout these paralogous , duplication/mutation-derived cadherin regions would become apparent should P35546 be aligned to itself in a qualitative dotplot or quantitative alignment.

Before then generating a Dot Plot for these 2 proteins, the following FAFSA protein sequence files were obtained from Uniprot using the Uniprot ID P22607 and P35546:

>sp|P22607|FGFR3\_HUMAN Fibroblast growth factor receptor 3 OS=Homo sapiens OX=9606 GN=FGFR3 PE=1 SV=1

MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELS

CPPPGGGPMGPTVWVKDGTGLVPSERVLVGPQRLQVLNASHEDSGAYSCRQRLTQRVLCH

FSVRVTDAPSSGDDEDGEDEAEDTGVDTGAPYWTRPERMDKKLLAVPAANTVRFRCPAAG

NPTPSISWLKNGREFRGEHRIGGIKLRHQQWSLVMESVVPSDRGNYTCVVENKFGSIRQT

YTLDVLERSPHRPILQAGLPANQTAVLGSDVEFHCKVYSDAQPHIQWLKHVEVNGSKVGP

DGTPYVTVLKTAGANTTDKELEVLSLHNVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAE

EELVEADEAGSVYAGILSYGVGFFLFILVVAAVTLCRLRSPPKKGLGSPTVHKISRFPLK

RQVSLESNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGE

GCFGQVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEMEMMKMIGKHKNIIN

LLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLDYSFDTCKPPEEQLTFKDLVSCAYQ

VARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKW

MAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCT

HDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPFEQYSPGGQDTPSS

SSSGDDSVFAHDLLPPAPPSSGGSRT

>sp|P35546|RET\_MOUSE Proto-oncogene tyrosine-protein kinase receptor Ret OS=Mus musculus OX=10090 GN=Ret PE=1 SV=2

MAKATSGAAGLGLKLILLLPLLGEAPLGLYFSRDAYWERLYVDQPAGTPLLYVHALRDAP

GEVPSFRLGQHLYGVYRTRLHENDWIRINETTGLLYLNQSLDHSSWEQLSIRNGGFPLLT

IFLQVFLGSTAQREGECHWPGCTRVYFSFINDTFPNCSSFKAQDLCIPETAVSFRVRENR

PPGTFYHFHMLPVQFLCPNISVKYSLLGGDSLPFRCDPDCLEVSTRWALDRELREKYVLE

ALCIVAGPGANKETVTLSFPVTVYDEDDSAPTFSGGVGTASAVVEFKRKEGTVVATLQVF

DADVVPASGELVRRYTNTLLSGDSWAQQTFRVEHSPIETLVQVNNNSVRATMHNYKLILN

RSLSISESRVLQLAVLVNDSDFQGPGAGGILVLHFNVSVLPVTLNLPRAYSFPVNKRARR

YAQIGKVCVENCQEFSGVSIQYKLQPSSINCTALGVVTSPEDTSGTLFVNDTEALRRPEC

TKLQYTVVATDRQTRRQTQASLVVTVEGTSITEEVGCPKSCAVNKRRPECEECGGLGSPT

GRCEWRQGDGKGITRNFSTCSPSTRTCPDGHCDAVESRDANICPQDCLRADIVGGHERGE

RQGIKAGYGICNCFPDEKKCFCEPEDSQGPLCDALCRTIITAALFSLIISILLSIFCVCH

HHKHGHKPPIASAEMTFCRPAQGFPISYSSSGTRRPSLDSTENQVPVDSFKIPEDPKWEF

PRKNLVLGKTLGEGEFGKVVKATAFRLKGRAGYTTVAVKMLKENASQSELRDLLSEFNLL

KQVNHPHVIKLYGACSQDGPLLLIVEYAKYGSLRGFLRDSRKIGPAYVSGGGSRNSSSLD

HPDERVLTMGDLISFAWQISRGMQYLAEMKLVHRDLAARNILVAEGRKMKISDFGLSRDV

YEEDSYVKKSKGRIPVKWMAIESLFDHIYTTQSDVWSFGVLLWEIVTLGGNPYPGIPPER

LFNLLKTGHRMERPDNCSEEMYRLMLQCWKQEPDKRPVFADISKDLEKMMVKSRDYLDLA

ASTPSDSLLYDDGLSEEETPLVDCNNAPLPRSLPSTWIENKLYGMSDPNWPGESPVPLTR

ADGTSTGFPRYANDSVYANWMVSPSAAKLMDTFDS

The Dot Plot was then generated by visiting the Dotlet app online at <https://dotlet.vital-it.ch> and entering the two FAFSA file amino acid sequences, where Sequence#1 corresponded to P22607 (FGFR3\_HUMAN Fibroblast growth factor receptor 3 (Organism=Homo sapiens)) and Sequence#2 corresponded to P35546 (RET\_MOUSE Proto-oncogene tyrosine-protein kinase receptor Ret (Organism= Mus musculus)). The Dot Plot Window Size was adjusted to 22 to remove noise and threshhold was adjusted also so as to adjust stringency.Based on the dot plot, a distinct, well-defined and relatively continuous diagonal line was observed on the lower left of the plot. This qualitatively signified a relatively high amount of homology and sequence identity and corresponded to the aforementioned shared domain PF07714 (Protein tyrosine and serine/threonine kinase) whose locations were respectively deisgnated by blue rectangles. Because no significant similarity was observed elsewhere in the Dot Plot, local alignment again appears to be more justified for these 2 proteins. The two proteins may be homologous, as they qualitatively suggest and present a high sequence identity and high sequence similarity along the two sequences. The regions with no similarity correspond to the regions that have mutated during evolution.

Despite the aforementioned reasons justifying a local allignment, a global allignment between both sequences was first performed nonetheless, consistent with class video. For this, global allignment using the default BLOSUM62 substuttion matrix initially the FAFSA file amino acid sequences for both proteins were entered at the following website:

<https://www.ebi.ac.uk/Tools/psa/emboss_needle/>

This global alignment resulted in low sequence identity (23.4%), low similarity (31.2%), and high gap (46.8%). In contrast, a local alignment around a smaller subset of 463 amino acids corresponding to the shared domain was performed using the identical default BLOSUM62 substitution matrix at the following website:

<https://www.ebi.ac.uk/Tools/psa/emboss_water/>

This local alignment instead quantitatively resulted in higher sequence identity (43.8%), higher similarity (56.8%), and lower gap (17.1%). With all things considered, local alignment is justified as more suitable for these proteins sharing a relatively small, localized, common, highly conserved domain.

1. Matrix used for the alignment? Jusify your answer.

**BLOSUM85.** The following was considered as substitution matrix selection criteria for performing alignments of the two proteins. BLOSUM matrices usually perform better than PAM matrices

for local alignments of protein sequences in databases, and therefore BLOSUM-based matrixes were selected for our 2 proteins. For conducting local and short alignment (~412 amino acids) of very similar sequences, such as the shared and similar domain PF07714 (Protein tyrosine and serine/threonine kinase)of our two proteins, a lower PAM or, in this case, a higher level **BLOSUM85** matrix was most appropriate for our very similar protein sequences. BLOSUM85 evidently resulted in the highest sequence identity (46.8%), highest similarity (57.5%), and lowest gap (16.5%). For this, the default parameters of gap opening of 10 and gap extension of 0.5 were used, and these represent rarer, less probable changes compared to evolutionary mutations involving mismatches that still conserve protein 3-D structure and fucntionality.

In general, BLOSUM80 would be appropriate for closely related protein families, BLOSUM62 is more effective at detecting potential similitudes, and BLOSUM45 is appropriate for large alignments of divergent sequences. for aligning distantly related sequences: high PAM or low

BLOSUM.When conducting my local alignment, BLOSUM30 proved inappropriate and too low and causes errors. When conducting my global alignment on the two proteins, the % identity and similarity was increased when switching from BLOSUM62 to BLOSUM45.

1. **Identity, similarity, % gap?**

As previously mentioned, when performing, for conducting local and short alignment (~412 amino acids) of very similar sequences, such as the shared and similar domain PF07714 (Protein tyrosine and serine/threonine kinase)of our two proteins, a lower PAM or, in this case, a higher level **BLOSUM85** matrix was most appropriate for our very similar protein sequences. BLOSUM85 evidently resulted in the highest sequence identity (46.8%), highest similarity (57.5%), and lowest gap (16.5%). It is noted that the sequence identity is expected to be higher than sequence % similaity for proteins, unlike for nucleotides. For the local alignment with BLOSUM85, the default parameters of gap opening of 10 and gap extension of 0.5 were used, and these represent rarer, less probable changes compared to evolutionary mutations involving mismatches that still conserve protein 3-D structure and fucntionality.

1. Are these sequences homologous (specify paralogous/orthologous)

**Homologous-Orthologous.** Based on the significantly high % sequence identity of 46.8% (>30%), low E-value (from BLASTp) the alternative hypothesis that the two proteins ((Homo sapiens)human fibroblast growth factor receptor 3 (P22607) and (Mus Musculus) Mouse proto-oncogene c-ret (P35546) are homologous has been proven. These 2 proteins are homologous as they and the shared , conserved domain PF07714 (Protein tyrosine and serine/threonine kinase) are derived from a single common ancestral gene present in the common ancestor of those organisms. The shared domains exhibit similarl 3-D structure and therefore similiar functionality. The 2 proteins are orthologous genes as they are found in different organisms. In contrast, the non-shared, slightly different, cadherin domains of the P35546 protein are likely paralogous to themselves, as they are both present in the same Mus Musculus organism and are related to each other through gene duplication event and subsequent evolutionary mutation events.

**INTERPROSCANS FOR THE TWO PROTEINS:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**PFAM DESCRIPTION OF SHARED DOMAIN:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**PFAM OUTPUT FOR SHARED DOMAIN:**

A screenshot of a computer

Description automatically generated

**PFAM CARTOON OF SHARED DOMAIN:**

A screenshot of a cell phone

Description automatically generated

**PFAM SIGNATURE FOR SHARED DOMAIN:**

A screenshot of a computer

Description automatically generated

**DOT PLOT FOR 2 PROTEINS WITH HIGH NOISE:**

A screenshot of a computer

Description automatically generated

**DOT PLOT FOR 2 PROTEINS WITH LOW NOISE:**

A screenshot of a computer

Description automatically generated

**GLOBAL ALIGNMENT: ALLIGNMENT#1 USING BLOSUM62**

########################################

# Program: needle

# Rundate: Sun 12 Nov 2023 03:32:17

# Commandline: needle

# -auto

# -stdout

# -asequence emboss\_needle-I20231112-033212-0622-86777755-p1m.asequence

# -bsequence emboss\_needle-I20231112-033212-0622-86777755-p1m.bsequence

# -datafile EBLOSUM62

# -gapopen 10.0

# -gapextend 0.5

# -endopen 10.0

# -endextend 0.5

# -aformat3 pair

# -sprotein1

# -sprotein2

# Align\_format: pair

# Report\_file: stdout

########################################

#=======================================

#

# Aligned\_sequences: 2

# 1: FGFR3\_HUMAN

# 2: RET\_MOUSE

# Matrix: EBLOSUM62

# Gap\_penalty: 10.0

# Extend\_penalty: 0.5

#

# Length: 1254

# Identity: 293/1254 (23.4%)

# Similarity: 391/1254 (31.2%)

# Gaps: 587/1254 (46.8%)

# Score: 853.0

#

#

#=======================================

FGFR3\_HUMAN 1 -------------------------------------------------- 0

RET\_MOUSE 1 MAKATSGAAGLGLKLILLLPLLGEAPLGLYFSRDAYWERLYVDQPAGTPL 50

FGFR3\_HUMAN 1 -------------------------------------------------- 0

RET\_MOUSE 51 LYVHALRDAPGEVPSFRLGQHLYGVYRTRLHENDWIRINETTGLLYLNQS 100

FGFR3\_HUMAN 1 -------------------------------------------------- 0

RET\_MOUSE 101 LDHSSWEQLSIRNGGFPLLTIFLQVFLGSTAQREGECHWPGCTRVYFSFI 150

FGFR3\_HUMAN 1 -------------------------------------------------- 0

RET\_MOUSE 151 NDTFPNCSSFKAQDLCIPETAVSFRVRENRPPGTFYHFHMLPVQFLCPNI 200

FGFR3\_HUMAN 1 --------------------------------MGAPACALALCVAVAIVA 18

:.......||| |||

RET\_MOUSE 201 SVKYSLLGGDSLPFRCDPDCLEVSTRWALDRELREKYVLEALC----IVA 246

FGFR3\_HUMAN 19 GASSESLGTEQRVVGRAAEVPGPEPGQQEQLVFGSGDAVELSCP------ 62

| ||..: :.|.||.|

RET\_MOUSE 247 G-----------------------PGANK-------ETVTLSFPVTVYDE 266

FGFR3\_HUMAN 63 ----PPGGGPMGPTVWV-----KDGTGLVPSERVLVGPQRLQVLNASHED 103

|...|.:|....| |:||.:. .|||.:|....

RET\_MOUSE 267 DDSAPTFSGGVGTASAVVEFKRKEGTVVA----------TLQVFDADVVP 306

FGFR3\_HUMAN 104 SGAYSCRQRLTQRVLCHFSVRVTDAPSSGDDEDGEDEAEDTGVDTGAPYW 153

:.....| |.|..:| .|:..|:.|.....:|..

RET\_MOUSE 307 ASGELVR-RYTNTLL-----------------SGDSWAQQTFRVEHSPIE 338

FGFR3\_HUMAN 154 T---------RPERMDKKL---------------LAVPAANTVRFRCPAA 179

| |....:.|| ||| ..|...|:.|.|

RET\_MOUSE 339 TLVQVNNNSVRATMHNYKLILNRSLSISESRVLQLAV-LVNDSDFQGPGA 387

FGFR3\_HUMAN 180 G----------------NPTPSISWLKNGREFRGEHRIGGI--------- 204

| |...:.|:..|.|. |...:||.:

RET\_MOUSE 388 GGILVLHFNVSVLPVTLNLPRAYSFPVNKRA-RRYAQIGKVCVENCQEFS 436

FGFR3\_HUMAN 205 ------KLRHQQWSLVMESVVPSDRGNYTCVVENKFGSIRQ------TYT 242

||:....:.....||.|.......:..|...::|: .||

RET\_MOUSE 437 GVSIQYKLQPSSINCTALGVVTSPEDTSGTLFVNDTEALRRPECTKLQYT 486

FGFR3\_HUMAN 243 LDVLERSPHRPILQAGLPANQTAVLGSDV--EFHCKVYSDAQPHIQWLKH 290

:...:|...|. .||.| ...|.|:.: |..|. |.

RET\_MOUSE 487 VVATDRQTRRQ-TQASL---VVTVEGTSITEEVGCP------------KS 520

FGFR3\_HUMAN 291 VEVN------------GSKVG------PDGTPYVTVLKTAGANT---TDK 319

..|| ||..| .||........|...:| .|.

RET\_MOUSE 521 CAVNKRRPECEECGGLGSPTGRCEWRQGDGKGITRNFSTCSPSTRTCPDG 570

FGFR3\_HUMAN 320 ELEVLSLH--NVTFEDAGEYTCLAGNSIGFSHHSAWLVVLPAEEELVEA- 366

..:.:... |:..:| ||..:.:| .|.. .|.:.::|

RET\_MOUSE 571 HCDAVESRDANICPQD-----CLRADIVG-GHER-------GERQGIKAG 607

FGFR3\_HUMAN 367 ------------------DEAGS---------VYAGILSYGVGFFLFILV 389

|..|. :.|.:.|..:...|.|..

RET\_MOUSE 608 YGICNCFPDEKKCFCEPEDSQGPLCDALCRTIITAALFSLIISILLSIFC 657

FGFR3\_HUMAN 390 V--------------AAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSL 425

| |.:|.|| |.:| ||:

RET\_MOUSE 658 VCHHHKHGHKPPIASAEMTFCR----PAQG-----------FPI------ 686

FGFR3\_HUMAN 426 ESNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLG 475

|.:|..:..| .|.|.|.. ..|...::|.|||||..|..|.||

RET\_MOUSE 687 -SYSSSGTRRP-----SLDSTENQ--VPVDSFKIPEDPKWEFPRKNLVLG 728

FGFR3\_HUMAN 476 KPLGEGCFGQVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSE 525

|.||||.||:||.|.|..: |.||.. .||||||||::|:..:|.||:||

RET\_MOUSE 729 KTLGEGEFGKVVKATAFRL-KGRAGY-TTVAVKMLKENASQSELRDLLSE 776

FGFR3\_HUMAN 526 MEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGL 575

..::|.: .|.::|.|.|||:|.|||.::||||..|:||.|||..|..|.

RET\_MOUSE 777 FNLLKQV-NHPHVIKLYGACSQDGPLLLIVEYAKYGSLRGFLRDSRKIGP 825

FGFR3\_HUMAN 576 DY----------SFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIH 615

.| |.| .|.|..||..||:|.|:|::|||:|||..|.:|

RET\_MOUSE 826 AYVSGGGSRNSSSLD--HPDERVLTMGDLISFAWQISRGMQYLAEMKLVH 873

FGFR3\_HUMAN 616 RDLAARNVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEA 665

|||||||:||.|...|||:||||:|||:..|.|.|.:.||:||||||.|:

RET\_MOUSE 874 RDLAARNILVAEGRKMKISDFGLSRDVYEEDSYVKKSKGRIPVKWMAIES 923

FGFR3\_HUMAN 666 LFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDK 715

|||.:||.||||||||||||||.||||:||||||.|.||.|||.||||::

RET\_MOUSE 924 LFDHIYTTQSDVWSFGVLLWEIVTLGGNPYPGIPPERLFNLLKTGHRMER 973

FGFR3\_HUMAN 716 PANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSA 765

|.||:.::|.:|.:||...|.:||.|..:.:||:::: |.|.| ||||:|

RET\_MOUSE 974 PDNCSEEMYRLMLQCWKQEPDKRPVFADISKDLEKMM-VKSRD-YLDLAA 1021

FGFR3\_HUMAN 766 PFEQYSPGGQDTPSSSSSGDDSVFAHDL---------LPPAPPSSGGSRT 806

.|||.|...||.:...:. ||.:.||:.....

RET\_MOUSE 1022 ----------STPSDSLLYDDGLSEEETPLVDCNNAPLPRSLPSTWIENK 1061

FGFR3\_HUMAN 807 -------------------------------------------------- 806

RET\_MOUSE 1062 LYGMSDPNWPGESPVPLTRADGTSTGFPRYANDSVYANWMVSPSAAKLMD 1111

FGFR3\_HUMAN 807 ---- 806

RET\_MOUSE 1112 TFDS 1115

**GLOBAL ALIGNMENT: ALLIGNMENT#2 USING BLOSUM30 (To increase % identity)**

########################################

# Program: needle

# Rundate: Sun 12 Nov 2023 03:44:00

# Commandline: needle

# -auto

# -stdout

# -asequence emboss\_needle-I20231112-034355-0408-88475795-p1m.asequence

# -bsequence emboss\_needle-I20231112-034355-0408-88475795-p1m.bsequence

# -datafile EBLOSUM30

# -gapopen 10.0

# -gapextend 0.5

# -endopen 10.0

# -endextend 0.5

# -aformat3 pair

# -sprotein1

# -sprotein2

# Align\_format: pair

# Report\_file: stdout

########################################

#=======================================

#

# Aligned\_sequences: 2

# 1: FGFR3\_HUMAN

# 2: RET\_MOUSE

# Matrix: EBLOSUM30

# Gap\_penalty: 10.0

# Extend\_penalty: 0.5

#

# Length: 1206

# Identity: 314/1206 (26.0%)

# Similarity: 483/1206 (40.0%)

# Gaps: 491/1206 (40.7%)

# Score: 1616.0

#

#

#=======================================

FGFR3\_HUMAN 1 MGAPACALALCVAVAIVAGASSESLGTEQRVVGRAAEVPGPEPGQQEQLV 50

.|.|:|.:.| :|....:..|..|:::..:

RET\_MOUSE 1 ----------------MAKATSGAAG-----LGLKLILLLPLLGEAPLGL 29

FGFR3\_HUMAN 51 FGSGDAV--ELSCPPPGGGPMGPTVWVKDGTGLVPSERV---LVGPQRLQ 95

:.|.||: .|....|.|.|:..:..::|..|.|||.|: |:|..|.:

RET\_MOUSE 30 YFSRDAYWERLYVDQPAGTPLLYVHALRDAPGEVPSFRLGQHLYGVYRTR 79

FGFR3\_HUMAN 96 ----------------VLNAS--HEDSGAYSCR-----------QRL--- 113

:||:| |....::|.| |.:

RET\_MOUSE 80 LHENDWIRINETTGLLYLNQSLDHSSWEQLSIRNGGFPLLTIFLQVFLGS 129

FGFR3\_HUMAN 114 -TQR-VLCH--------FSVRVTDA-P--SSGDDEDGEDEAEDTGVDTGA 150

:|| ..|| ||: ::|: | ||...:| .:.::|.|...:

RET\_MOUSE 130 TAQREGECHWPGCTRVYFSF-INDTFPNCSSFKAQD--LCIPETAVSFRV 176

FGFR3\_HUMAN 151 PYWTRPERMDKKLLAVPAANTVRFRCPA--------AGNPTP-------- 184

..:||:.....:.::| |:|.||. .|:..|

RET\_MOUSE 177 -RENRPPGTFYHFHMLP----VQFLCPNISVKYSLLGGDSLPFRCDPDCL 221

FGFR3\_HUMAN 185 --SISWLKNGREFRGEHRI------------------------------- 201

|..|..: ||:|.:..:

RET\_MOUSE 222 EVSTRWALD-RELREKYVLEALCIVAGPGANKETVTLSFPVTVYDEDDSA 270

FGFR3\_HUMAN 202 ----GG-------IKLRHQQWSLVM------ESVVPSDRGNYTCVVENKF 234

|| ::::..::::|: ..|||:. |.::..:.|.:

RET\_MOUSE 271 PTFSGGVGTASAVVEFKRKEGTVVATLQVFDADVVPAS-GELVRRYTNTL 319

FGFR3\_HUMAN 235 GS----IRQTYTLDVLERSPHRPILQAGLPAN------------------ 262

.| .:||: |.:|.|| ::.:.:|

RET\_MOUSE 320 LSGDSWAQQTF------RVEHSPI-ETLVQVNNNSVRATMHNYKLILNRS 362

FGFR3\_HUMAN 263 ---------QTAVLGSDVEF-----------H------------------ 274

|.|||..|.:| |

RET\_MOUSE 363 LSISESRVLQLAVLVNDSDFQGPGAGGILVLHFNVSVLPVTLNLPRAYSF 412

FGFR3\_HUMAN 275 -------------------CKVYSDAQPHIQW-LKHVEVNGSKVG----P 300

|..:|.: .||: |.:..:|.:.:| |

RET\_MOUSE 413 PVNKRARRYAQIGKVCVENCQEFSGV--SIQYKLQPSSINCTALGVVTSP 460

FGFR3\_HUMAN 301 D---GTPYVT--------------------------------VLKTAGAN 315

: ||.:|: |:.:.|:.

RET\_MOUSE 461 EDTSGTLFVNDTEALRRPECTKLQYTVVATDRQTRRQTQASLVVTVEGTS 510

FGFR3\_HUMAN 316 TTDK--------------ELEVL--------------------------- 324

.|:: |.|..

RET\_MOUSE 511 ITEEVGCPKSCAVNKRRPECEECGGLGSPTGRCEWRQGDGKGITRNFSTC 560

FGFR3\_HUMAN 325 --------SLHNVTFE--DAG--EYTCLAGNSIGFSHH-------SAWLV 355

..|..::| ||. :..||..:.:| .|. .|::.

RET\_MOUSE 561 SPSTRTCPDGHCDAVESRDANICPQDCLRADIVG-GHERGERQGIKAGYG 609

FGFR3\_HUMAN 356 V---LPAEEELV--EADEAG---------SVYAGILSYGVGFFLFILVV- 390

: :|.|::.: :.|.:| .:.|.::|:.:..:|.|:.|

RET\_MOUSE 610 ICNCFPDEKKCFCEPEDSQGPLCDALCRTIITAALFSLIISILLSIFCVC 659

FGFR3\_HUMAN 391 -------------AAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLES 427

|..|:|| |..| || :|:.|

RET\_MOUSE 660 HHHKHGHKPPIASAEMTFCR----PAQG-----------FP----ISYSS 690

FGFR3\_HUMAN 428 NASMSSNTPLVRIARLSSGEG--PTLANVSELELPADPKWELSRARLTLG 475

..: |...|.|.|. | |..:::|.|||||:.|..|:||

RET\_MOUSE 691 SGT--------RRPSLDSTENQVP----VDSFKIPEDPKWEFPRKNLVLG 728

FGFR3\_HUMAN 476 KPLGEGCFGQVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSE 525

|.||||:||.||:|.|..: |.||.. :||||||||::|:..:|.||:||

RET\_MOUSE 729 KTLGEGEFGKVVKATAFRL-KGRAGY-TTVAVKMLKENASQSELRDLLSE 776

FGFR3\_HUMAN 526 MEMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGL 575

..::|.: .|:.:|.|:|||:|.|||:::||||..|.||.|||..|:.|.

RET\_MOUSE 777 FNLLKQV-NHPHVIKLYGACSQDGPLLLIVEYAKYGSLRGFLRDSRKIGP 825

FGFR3\_HUMAN 576 DY----------SFDTCKPPEEQ-LTFKDLVSCAYQVARGMEYLASQKCI 614

.| |:| :|:|: ||..||:|.|:|::|||:|||..|.:

RET\_MOUSE 826 AYVSGGGSRNSSSLD---HPDERVLTMGDLISFAWQISRGMQYLAEMKLV 872

FGFR3\_HUMAN 615 HRDLAARNVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPE 664

||||||||:||:|...|||:||||:|||...|.|.|.:.||:||||||.|

RET\_MOUSE 873 HRDLAARNILVAEGRKMKISDFGLSRDVYEEDSYVKKSKGRIPVKWMAIE 922

FGFR3\_HUMAN 665 ALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMD 714

:|||.:||.||||||||||||||:||||.||||||.|.||.|||.||||:

RET\_MOUSE 923 SLFDHIYTTQSDVWSFGVLLWEIVTLGGNPYPGIPPERLFNLLKTGHRME 972

FGFR3\_HUMAN 715 KPANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLS 764

:|.||:.::|.:|.:||.:.|..||:|..:.:||:: :.|.|.| ||||:

RET\_MOUSE 973 RPDNCSEEMYRLMLQCWKQEPDKRPVFADISKDLEK-MMVKSRD-YLDLA 1020

FGFR3\_HUMAN 765 APFEQYSPGGQDTPSSSSSGDDSVFAHDLLP-------PAP---PSSGGS 804

| .|||.|...||. ::.:..| |.| ||::..

RET\_MOUSE 1021 A----------STPSDSLLYDDG-LSEEETPLVDCNNAPLPRSLPSTWIE 1059

FGFR3\_HUMAN 805 RT------------------------------------------------ 806

..

RET\_MOUSE 1060 NKLYGMSDPNWPGESPVPLTRADGTSTGFPRYANDSVYANWMVSPSAAKL 1109

FGFR3\_HUMAN 807 ------ 806

RET\_MOUSE 1110 MDTFDS 1115

#---------------------------------------

#---------------------------------------

**LOCAL ALIGNMENT: ALLIGNMENT#3 USING BLOSUM62**

########################################

# Program: water

# Rundate: Sun 12 Nov 2023 03:48:59

# Commandline: water

# -auto

# -stdout

# -asequence emboss\_water-I20231112-034855-0724-75055289-p1m.asequence

# -bsequence emboss\_water-I20231112-034855-0724-75055289-p1m.bsequence

# -datafile EBLOSUM62

# -gapopen 10.0

# -gapextend 0.5

# -aformat3 pair

# -sprotein1

# -sprotein2

# Align\_format: pair

# Report\_file: stdout

########################################

#=======================================

#

# Aligned\_sequences: 2

# 1: FGFR3\_HUMAN

# 2: RET\_MOUSE

# Matrix: EBLOSUM62

# Gap\_penalty: 10.0

# Extend\_penalty: 0.5

#

# Length: 463

# Identity: 203/463 (43.8%)

# Similarity: 263/463 (56.8%)

# Gaps: 79/463 (17.1%)

# Score: 890.0

#

#

#=======================================

FGFR3\_HUMAN 372 VYAGILSYGVGFFLFILVV--------------AAVTLCRLRSPPKKGLG 407

:.|.:.|..:...|.|..| |.:|.|| |.:|

RET\_MOUSE 640 ITAALFSLIISILLSIFCVCHHHKHGHKPPIASAEMTFCR----PAQG-- 683

FGFR3\_HUMAN 408 SPTVHKISRFPLKRQVSLESNASMSSNTPLVRIARLSSGEGPTLANVSEL 457

||: |.:|..:..| .|.|.|.. ..|...

RET\_MOUSE 684 ---------FPI-------SYSSSGTRRP-----SLDSTENQ--VPVDSF 710

FGFR3\_HUMAN 458 ELPADPKWELSRARLTLGKPLGEGCFGQVVMAEAIGIDKDRAAKPVTVAV 507

::|.|||||..|..|.|||.||||.||:||.|.|..: |.||.. .||||

RET\_MOUSE 711 KIPEDPKWEFPRKNLVLGKTLGEGEFGKVVKATAFRL-KGRAGY-TTVAV 758

FGFR3\_HUMAN 508 KMLKDDATDKDLSDLVSEMEMMKMIGKHKNIINLLGACTQGGPLYVLVEY 557

||||::|:..:|.||:||..::|.: .|.::|.|.|||:|.|||.::|||

RET\_MOUSE 759 KMLKENASQSELRDLLSEFNLLKQV-NHPHVIKLYGACSQDGPLLLIVEY 807

FGFR3\_HUMAN 558 AAKGNLREFLRARRPPGLDY----------SFDTCKPPEEQLTFKDLVSC 597

|..|:||.|||..|..|..| |.| .|.|..||..||:|.

RET\_MOUSE 808 AKYGSLRGFLRDSRKIGPAYVSGGGSRNSSSLD--HPDERVLTMGDLISF 855

FGFR3\_HUMAN 598 AYQVARGMEYLASQKCIHRDLAARNVLVTEDNVMKIADFGLARDVHNLDY 647

|:|::|||:|||..|.:||||||||:||.|...|||:||||:|||:..|.

RET\_MOUSE 856 AWQISRGMQYLAEMKLVHRDLAARNILVAEGRKMKISDFGLSRDVYEEDS 905

FGFR3\_HUMAN 648 YKKTTNGRLPVKWMAPEALFDRVYTHQSDVWSFGVLLWEIFTLGGSPYPG 697

|.|.:.||:||||||.|:|||.:||.||||||||||||||.||||:||||

RET\_MOUSE 906 YVKKSKGRIPVKWMAIESLFDHIYTTQSDVWSFGVLLWEIVTLGGNPYPG 955

FGFR3\_HUMAN 698 IPVEELFKLLKEGHRMDKPANCTHDLYMIMRECWHAAPSQRPTFKQLVED 747

||.|.||.|||.||||::|.||:.::|.:|.:||...|.:||.|..:.:|

RET\_MOUSE 956 IPPERLFNLLKTGHRMERPDNCSEEMYRLMLQCWKQEPDKRPVFADISKD 1005

FGFR3\_HUMAN 748 LDRVLTVTSTDEYLDLSAPFEQYSPGGQDTPSSSSSGDDSVFAHDL---- 793

|:::: |.|.| ||||:| .|||.|...||.:...:.

RET\_MOUSE 1006 LEKMM-VKSRD-YLDLAA----------STPSDSLLYDDGLSEEETPLVD 1043

FGFR3\_HUMAN 794 -----LPPAPPSS 801

||.:.||:

RET\_MOUSE 1044 CNNAPLPRSLPST 1056

#---------------------------------------

#---------------------------------------

**LOCAL ALIGNMENT: ALLIGNMENT#4 USING BLOSUM35 (BLOSUM30 too low/errors)**

########################################

# Program: water

# Rundate: Sun 12 Nov 2023 03:54:32

# Commandline: water

# -auto

# -stdout

# -asequence emboss\_water-I20231112-035429-0034-36437449-p1m.asequence

# -bsequence emboss\_water-I20231112-035429-0034-36437449-p1m.bsequence

# -datafile EBLOSUM35

# -gapopen 10.0

# -gapextend 0.5

# -aformat3 pair

# -sprotein1

# -sprotein2

# Align\_format: pair

# Report\_file: stdout

########################################

#=======================================

#

# Aligned\_sequences: 2

# 1: FGFR3\_HUMAN

# 2: RET\_MOUSE

# Matrix: EBLOSUM35

# Gap\_penalty: 10.0

# Extend\_penalty: 0.5

#

# Length: 994

# Identity: 303/994 (30.5%)

# Similarity: 438/994 (44.1%)

# Gaps: 288/994 (29.0%)

# Score: 1462.5

#

#

#=======================================

FGFR3\_HUMAN 4 PAC----ALALCVAVAIVAGASSESLGTEQRVVGRAAEV--PG------- 40

|.| |..||: .|..|..|..|. ||

RET\_MOUSE 155 PNCSSFKAQDLCI--------------PETAVSFRVRENRPPGTFYHFHM 190

FGFR3\_HUMAN 41 -P----EPGQQEQLVFGSGDAVELSCPPPGGGPMGPTVWVKD----GTGL 81

| .|:......:.:||::.:.|.|...... |.|..| ...:

RET\_MOUSE 191 LPVQFLCPNISVKYSLLGGDSLPFRCDPDCLEVS--TRWALDRELREKYV 238

FGFR3\_HUMAN 82 VPSERVLVGP--QRLQV-----LNASHED------SGAYSCRQRLTQRVL 118

:.:..::.|| ::..| :....|| ||..: |...:

RET\_MOUSE 239 LEALCIVAGPGANKETVTLSFPVTVYDEDDSAPTFSGGVG-----TASAV 283

FGFR3\_HUMAN 119 CHF---------SVRVTDA---PSSGD--------DEDGEDEAEDTGVDT 148

..| :::|.|| |:||: ...|:..|:.|....

RET\_MOUSE 284 VEFKRKEGTVVATLQVFDADVVPASGELVRRYTNTLLSGDSWAQQTFRVE 333

FGFR3\_HUMAN 149 GAPYWT---------RPERMDKKL---------------LAVPAANTVRF 174

.:|..| |...::.|| ||| ..|...|

RET\_MOUSE 334 HSPIETLVQVNNNSVRATMHNYKLILNRSLSISESRVLQLAV-LVNDSDF 382

FGFR3\_HUMAN 175 RCPAAGN-----------PT----PSISWLKNGREFRGEHRIGGI----- 204

:.|.||: |: |....:..::..|...:||.:

RET\_MOUSE 383 QGPGAGGILVLHFNVSVLPVTLNLPRAYSFPVNKRARRYAQIGKVCVENC 432

FGFR3\_HUMAN 205 ----------KLRHQQWSLVMESVVPSDRGNYTCVVENKFGSIRQ----- 239

||:......:..:||.|.......::.|...::|:

RET\_MOUSE 433 QEFSGVSIQYKLQPSSINCTALGVVTSPEDTSGTLFVNDTEALRRPECTK 482

FGFR3\_HUMAN 240 -TYTLDVLERSPHRPILQAGLPANQTAVLGSDV--EFHCKVYSDAQPHIQ 286

.||:...:|...|. .||:| .:.|.|:.: |:.| |

RET\_MOUSE 483 LQYTVVATDRQTRRQ-TQASL---VVTVEGTSITEEVGC-------P--- 518

FGFR3\_HUMAN 287 WLKHVEVN------------GSKVG------PDGTPYVTVLKTAGANT-- 316

|...|| ||.:| .||...:..:.|.:..|

RET\_MOUSE 519 --KSCAVNKRRPECEECGGLGSPTGRCEWRQGDGKGITRNFSTCSPSTRT 566

FGFR3\_HUMAN 317 -TDKELEVLSLHNVTFE--DAG--EYTCLAGNSIGFSHH-------SAWL 354

.|. |...:| ||: ...||..:.:| :|. .|..

RET\_MOUSE 567 CPDG-------HCDAVESRDANICPQDCLRADIVG-GHERGERQGIKAGY 608

FGFR3\_HUMAN 355 VV---LPAEEE-----------LVEADEAGSVYAGILSYGVGFFLFILVV 390

.: :|.|:: |.:|.....:.|.::|..::::|.|:.|

RET\_MOUSE 609 GICNCFPDEKKCFCEPEDSQGPLCDALCRTIITAALFSLIISILLSIFCV 658

FGFR3\_HUMAN 391 --------------AAVTLCRLRSPPKKGLGSPTVHKISRFPLKRQVSLE 426

|.:|:|| |..| ||:...

RET\_MOUSE 659 CHHHKHGHKPPIASAEMTFCR----PAQG-----------FPISYS---- 689

FGFR3\_HUMAN 427 SNASMSSNTPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGK 476

||:| |...|.|.|:.: .|..:::|.|||||:.|..|:|||

RET\_MOUSE 690 -----SSGT---RRPSLDSTENQV--PVDSFKIPEDPKWEFPRKNLVLGK 729

FGFR3\_HUMAN 477 PLGEGCFGQVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEM 526

.||||.||.||.|.|:.: |.||.. :||||||||::|:..:|.||:||.

RET\_MOUSE 730 TLGEGEFGKVVKATAFRL-KGRAGY-TTVAVKMLKENASQSELRDLLSEF 777

FGFR3\_HUMAN 527 EMMKMIGKHKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLD 576

.::|.:: |.::|.|.|||:|.|||.::||||..|.||.|||..|..|..

RET\_MOUSE 778 NLLKQVN-HPHVIKLYGACSQDGPLLLIVEYAKYGSLRGFLRDSRKIGPA 826

FGFR3\_HUMAN 577 Y----------SFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHR 616

| |:| .|.|..||..||:|.|:|::|||:|||..|.:||

RET\_MOUSE 827 YVSGGGSRNSSSLD--HPDERVLTMGDLISFAWQISRGMQYLAEMKLVHR 874

FGFR3\_HUMAN 617 DLAARNVLVTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEAL 666

||||||:||.|...|||:||||:|||...|.|.|.:.||:||||||.|:|

RET\_MOUSE 875 DLAARNILVAEGRKMKISDFGLSRDVYEEDSYVKKSKGRIPVKWMAIESL 924

FGFR3\_HUMAN 667 FDRVYTHQSDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKP 716

||.:||.||||||||||||||:||||.||||||.|.||.|||.||||::|

RET\_MOUSE 925 FDHIYTTQSDVWSFGVLLWEIVTLGGNPYPGIPPERLFNLLKTGHRMERP 974

FGFR3\_HUMAN 717 ANCTHDLYMIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAP 766

.||:.::|.:|.:||...|..||:|..:.:||:::: |.|.| ||||:|

RET\_MOUSE 975 DNCSEEMYRLMLQCWKQEPDKRPVFADISKDLEKMM-VKSRD-YLDLAA- 1021

FGFR3\_HUMAN 767 FEQYSPGGQDTPSSSSSGDDSVFAHDL---------LPPAPPSS 801

.|||.|...||::...:. ||.:.||:

RET\_MOUSE 1022 ---------STPSDSLLYDDGLSEEETPLVDCNNAPLPRSLPST 1056

#---------------------------------------

#---------------------------------------

The identiy decreased. Therefore, I re-ran pairwise alignment with higher BLOSUM

**LOCAL ALIGNMENT: ALLIGNMENT#5 USING BLOSUM85**

########################################

# Program: water

# Rundate: Sun 12 Nov 2023 03:57:45

# Commandline: water

# -auto

# -stdout

# -asequence emboss\_water-I20231112-035741-0250-93371905-p1m.asequence

# -bsequence emboss\_water-I20231112-035741-0250-93371905-p1m.bsequence

# -datafile EBLOSUM85

# -gapopen 10.0

# -gapextend 0.5

# -aformat3 pair

# -sprotein1

# -sprotein2

# Align\_format: pair

# Report\_file: stdout

########################################

#=======================================

#

# Aligned\_sequences: 2

# 1: FGFR3\_HUMAN

# 2: RET\_MOUSE

# Matrix: EBLOSUM85

# Gap\_penalty: 10.0

# Extend\_penalty: 0.5

#

# Length: 412

# Identity: 193/412 (46.8%)

# Similarity: 237/412 (57.5%)

# Gaps: 68/412 (16.5%)

# Score: 936.5

#

#

#=======================================

FGFR3\_HUMAN 391 AAVTLCRLRSPPKKGLGSPTVHKISRFPL------KRQVSLESNASMSSN 434

|..|.|| |.:| ||: .|:.||:|. ...

RET\_MOUSE 673 AEMTFCR----PAQG-----------FPISYSSSGTRRPSLDST---ENQ 704

FGFR3\_HUMAN 435 TPLVRIARLSSGEGPTLANVSELELPADPKWELSRARLTLGKPLGEGCFG 484

.| |....:|.|||||..|..|.|||.||||.||

RET\_MOUSE 705 VP-----------------VDSFKIPEDPKWEFPRKNLVLGKTLGEGEFG 737

FGFR3\_HUMAN 485 QVVMAEAIGIDKDRAAKPVTVAVKMLKDDATDKDLSDLVSEMEMMKMIGK 534

:||.|.|..: |.||.. .||||||||::|:..:|.||.||..::|.: .

RET\_MOUSE 738 KVVKATAFRL-KGRAGY-TTVAVKMLKENASQSELRDLLSEFNLLKQV-N 784

FGFR3\_HUMAN 535 HKNIINLLGACTQGGPLYVLVEYAAKGNLREFLRARRPPGLDY------- 577

|..:|.|.|||:|.|||..:||||..|.||.|||..|..|..|

RET\_MOUSE 785 HPHVIKLYGACSQDGPLLLIVEYAKYGSLRGFLRDSRKIGPAYVSGGGSR 834

FGFR3\_HUMAN 578 ---SFDTCKPPEEQLTFKDLVSCAYQVARGMEYLASQKCIHRDLAARNVL 624

|.| .|.|..||..||:|.|:|::|||:|||..|.:||||||||:|

RET\_MOUSE 835 NSSSLD--HPDERVLTMGDLISFAWQISRGMQYLAEMKLVHRDLAARNIL 882

FGFR3\_HUMAN 625 VTEDNVMKIADFGLARDVHNLDYYKKTTNGRLPVKWMAPEALFDRVYTHQ 674

|.|...|||:||||:|||:..|.|.|.:.||:||||||.|:|||.:||.|

RET\_MOUSE 883 VAEGRKMKISDFGLSRDVYEEDSYVKKSKGRIPVKWMAIESLFDHIYTTQ 932

FGFR3\_HUMAN 675 SDVWSFGVLLWEIFTLGGSPYPGIPVEELFKLLKEGHRMDKPANCTHDLY 724

|||||||||||||.||||.||||||.|.||.|||.||||::|.||:.::|

RET\_MOUSE 933 SDVWSFGVLLWEIVTLGGNPYPGIPPERLFNLLKTGHRMERPDNCSEEMY 982

FGFR3\_HUMAN 725 MIMRECWHAAPSQRPTFKQLVEDLDRVLTVTSTDEYLDLSAPFEQYSPGG 774

.:|.:||...|.:||.|..:..||::.: |.|.| ||||:|

RET\_MOUSE 983 RLMLQCWKQEPDKRPVFADISKDLEKMM-VKSRD-YLDLAA--------- 1021

FGFR3\_HUMAN 775 QDTPSSSSSGDD 786

.|||.|...||

RET\_MOUSE 1022 -STPSDSLLYDD 1032

#---------------------------------------

#---------------------------------------