

Species	Sequence	Position
Homo	-----MCFICGMLKEMSG--LRSECGSSK-----NTFIPGHV	0
Pan	-----MCFICGMLKEMSG--LRSECGSSK-----NTFIPGHV	30
Mus	-----MCFICGMLKEMSG--LRSECGSSK-----NTFIPGHV	0
Rattus	MAGSENTFQPLLLQKVDISILVLRSLVLGEHTGECNNPCTSLKAVFWFISTNQTVIQSRL	60
Homo	-----MVDGTLTLLLS	11
Pan	SFFLDTHDADSVLIPNGCRVSRANQLRTQE-----GLRTQEAG--IMVDGTLTLLLS	81
Mus	-----MLLFAHLLQLLV	12
Rattus	SFLDLA-----GVCRSQSAQKFQSTSQQKHNPSCFSQSGLMRMLPPAYLLQLLA	111
Homo	EALALTQTWAGSHSLKYFHTSVSRPGRGEPFISVGIVDDTQFVRFDNDAAAPRMVPRAP	71
Pan	EALALTQTWAGSHSLKYFHTSVSRPGRGEPFISVGIVDDTQFVRFDNDAAAPRMVPRAP	141
Mus	SAT--VPTQSSPHSLRYFTTAVSRPGLGEPFIIIVGVDDTQFVRFDSDAENPRMEPRAR	70
Rattus	GTM--VLTQTSSHLRYFYTALSRLPGLGEPFIVGVDDTQFVRYSDVENPRMEPRAR	169
Homo	WMEQEGSEYWDRETRSDARTAQIFRVNLRTLRGYYNQSEAGSHTLQWMHGCELGPDGRFL	131
Pan	WMEQEGSEYWDRETRSDARTAQIFRVNLRTLRGYYNQSEAGSHTLQWMHGCDLGPDPGRFL	201
Mus	WIEQEGPEYWERETWKARDMGRNFRVNLRTLLRGYYNQSNDESHTLQWMYGCDDVGPDPGRLL	130
Rattus	WMEQEEPAYWERETRKARDTGRNFKVNLRTLLRYYNQSDDESHTLQWMYGCDDVGPDPGHLL	229
Homo	RGYEQFAYDGKDYLTLNEDLRSWTAVDTAAQISEQKSNDAEAEHQRAYLEDTCVEWLHK	191
Pan	RGYEQFAYDGKDYLTLNEDLRSWTAVDTAAQISERKSNDAEAEHQRAYLEDTCVEWLHK	261
Mus	RGYQCEAYDGDQYISLNEDLRSWTANDIASQISKHKSEAVDEAHQRAYLQGPCVEWLHR	190
Rattus	RGYQCEAYDGRDYISLNEDLRSWTATDMASQASKIKSEEVGEAHQRAYLQGPCVEWLHT	289
Homo	YLEKGKETLLHLEPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQQDGEGHTQDTEL	251
Pan	YLEKGKETLLHLEPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQQDGEGHTQDTEL	321
Mus	YLRLLGNETLQSDPPKAHVTHHPSEDEVTLRCWALGFYPADITLTWQLNGEELTQDMEL	250
Rattus	YLHLGKETLLRSDPPKAHVTLHPRPEGDVTLCWALGFYPADITLTWQLNGEDLTQDMEL	349
Homo	VETRPAGDGTQKWAAVVVPSSGEEQRYTCHVQHEGLPEPVLTRWKPASQP--TIPIVGII	309
Pan	VDTRPAGDGTQKWAAVVVPSSGEEQRYTCHVQHEGLPEPLTRWKPASQP--TIPIVGII	379
Mus	VETRPAGDGTQKWAAVVVPSSGEEQRYTCHVQHEGLPEPLTRWKEPPPTVSNMVIIV	309
Rattus	VETRPAGDGTQKWAAVVVPSSGEEKYTCHEGLPEPLALRWEPSPFTDSSMPVIWV	409
Homo	AGLVLLGSVVS--GA--VVAAVIWRKKSSGGKGGSYS-----KAESDASQGSSEHS	357
Pan	AGLVLLGSVVS--GA--VVAAVMWRKKSSGGKGRSYS-----KAESDSA-----	420
Mus	--LVVLGAVIILGAVV--AFVMKRRRHIGVKGCYAHVLGSKSFQTSWDPQKA-----	357
Rattus	GAVAIIGAVAIIGAVAIIGAVVRRRRKNTGEKGSYAHVLGSKAFQISWDPQKA-----	462
Homo	L 358	
Pan	- 420	
Mus	- 357	
Rattus	- 462	

Przykładowe pozycje konserwatywne:

- VGYVDDTQFVR
- LNEDLRSWTA
- TLRCWALGFYPA
- TRPAGDGTQKWAAVVVP

Muscle

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```
Homo      -----
Pan        ----MCFICGML-----KMSGRLRSECGSSKNTFIPGHV
Mus        -----
Rattus     MAGSENTFQPLLLQKVDISILVLSVLGEHTGECNNPCTSLKAVFWFISTNQTQVIQSRL

Homo      -----MVDGTLTLLLEALALTQTW
Pan        SFFLDTHDADSVLIPNGCRVSRANQLRTQEGRLTQEGAGIMVDGTLTLLLEALALTQTW
Mus        -----MLLFAHLLQLLVSATVPTQS-
Rattus     SFLLDAGVCRSQSAAQKFQSTSQQKHNPHSCFSQSLGMRMLPPAYLLQLLAGTMMVLTQT-
              : . ** ** : . **:

Homo      AGSHSLKYFHTSVSRPGRGEPFISVGYVDDTQFVRFDNDAA SPRMVP RAPWMEQEGSEY
Pan        AGSHSLKYFHTSVSRPGRGEPFISVGYVDDTQFVRFDNDAA SPRMVP RAPWMEQEGSEY
Mus        -SPHSLRYFTTAVSRPGLGEPFIIVGYVDDTQFVRFDSDAENPRMEPRARWIEQEGPEY
Rattus     -SSHSLRYFYTALSRPGLGEPFIIVGYVDDTQFVRYSDEVNPRMEPRARWMEQEEPAY
              ..**.* ** *:*** ***** *****:*. * .*** ** *:*** . *

Homo      WDRETRSARDTAQIFRVNLR LRGYYNQSEAGSHTLQMMHGCELGPDGRFLRGYEQFAYD
Pan        WDRETRSARDTAQIFRVNLR LRGYYNQSEAGSHTLQMMHGCDLGPDGRFLRGYEQFAYD
Mus        WERETWKARDMGRNFRVNLRL LRGYYNQSDNESHTLQMMYGC DVGPDGRLLRGYCEAYD
Rattus     WERETRKARDTGRNFKVNLRL LRGYYNQSDNESHTLQMMYGC DVGPDGHLRLRGYCEAYD
              *:***.*** .. *:***** *****:*****:***:***.*** * **

Homo      GKDYLT LNEDLR SWTA VDTAAQISEQK SNDASEAHQRAYLEDTCVEWLHKYLEKGKETL
Pan        GKDYLT LNEDLR SWTA VDTAAQISERK SNDACEAHQRAYLEDTCVEWLHKYLEKGKETL
Mus        GQDYI SLNEDLR SWTA NDIA SQISKHKSEAVDEAHQRAYLQGPCVEWLHRYLR LRGNETL
Rattus     GRDYI SLNEDLR SWTA TDMASQASKIKSEEVGEAHQRAYLQGPCVEWLHTYLHLGKETL
              *.***:***** * *: * *: * *: * *:*****:..***** ** *:***

Homo      LHLEPPKTHVTHHPISDHEATLR CWALGFYPAEITLTWQQDGEGHTQDTLVETRPAGDG
Pan        LHLEPPKTHVTHHPISDHEATLR CWALGFYPAEITLTWQQDGEGHTQDTLVETRPAGDG
Mus        QRS DPPKAHVTHHPRSEDEVTLRCWALGFYPADITLTWQLNGEELTQDMELVETRPAGDG
Rattus     LRS DPPKAHVTLHPRPEGDVTLRCWALGFYPADITLTWQLNGEDLTQDMELVETRPAGDG
              . :***:** ** .: : *****:***** :* ** ***:*****

Homo      TFQKWA AVVVP SGEEQRYTCHVQHEGLPEPVTLRWKPA--SQPTIPIVGIIAGLVLLGSV
Pan        TFQKWA AVVVP SGEEQRYTCHVQHEGLPEPLTLRWKPA--SQPTIPIVGIIAGLVLLGSV
Mus        TFQKWA AVVVP LKGEQYTYCHVYHEGLPEPLTLRWEPPPSTVSNMVIIVL---VVLGAV
Rattus     TFQKWA AVVVP SGEE LKYTCHVEHEGLPEPLALRWEPSPTDSSMPVIIVLGA VAIIGAV
              ***** *: * ***** *****:***:*. : .: : : : .:***

Homo      -VSGAVV---AAVIWRKKSSGGKG-----GSYSKAEWSDSAQGSSESHL
Pan        -VSGAVV---AAVMWRKKSSGGKG-----RSYSKAEWSDSA-----
Mus        IILGAVV---AFVMKRRRHIGVKGYAHVLGSKSFQTS DWPQKA-----
Rattus     AIIGAVAIIGAVRRRK RNTGKGSYAHVLGSKAFQIS DWPQKA-----
              : ***. * * *. * ** : : :*:***
```

T-coffe

MSA

The multiple sequence alignment result as produced by T-coffee.

T-COFFEE, Version_11.00 (Version_11.00)

Cedric Notredame

SCORE=907

*

BAD **AVG** **GOOD**

*

Homo : 97
Pan : 93
Mus : 97
Rattus : 88
cons : 90

Homo -----
Pan M-----CFICGMLKEMSGLR-SEC-----GSSKNTFIPGHVSFFLDTHDADSV
Mus -----
Rattus MAGSENTFQPLLLQKVDSILVLRSLVLGEHTGECNNPCTSLKAVFWFISTNQTVIQSRLSFLLDAGVCRSQ
cons

Homo -----MVDGTLTLLLLSEALALTQTWAGSHSLKYFHTSVSRPGRGEPFI
Pan LIPNGCRVSRANQLRTQEGLRTQEAGIMVDGTLTLLLLSEALALTQTWAGSHSLKYFHTSVSRPGRGEPFI
Mus -----MLLFAHLLQLLVSATVPTQSSP--HSLRYFTTAVSRPGLGEPFI
Rattus SAAQKFQSTSQQKHNPSCFSQSLGMRMLPPAYLLQLLAGTMVLQTSS--HSLRYFYTALSRPGLGEPFI
cons

Homo SVGYVDDTQFVRFDNDAASPRMVPAPWMEQEGSEYWDRETRSARDTAQIFRVNLRTRLRGYYNQSEAGSHTL
Pan SVGYVDDTQFVRFDNDAASPRMVPAPWMEQEGSEYWDRETRSARDTAQIFRVNLRTRLRGYYNQSEAGSHTL
Mus IVGYVDDTQFVRFDSDAENPRMEPRARWIEQEGPEYWERETWKARDMGRNFRVNLRTLLGYYNQSNDESHTL
Rattus VVGYYVDDTQFVRYDSDVENPRMEPRARWMEQEAPYWERETRKARDTGRNFKVNLRTLLRYYNQSDDESHTL
cons

Homo QWMHGCELGPDGRFLRGYEQFAYDGKDYLTNEDLRSWTAVDTAAQISEQKSNDAEAEHORAYLEDTCVEW
Pan QWMHGCDLGPDPGRFLRGYEQFAYDGKDYLTNEDLRSWTAVDTAAQISERKSNDAEAEHORAYLEDTCVEW
Mus QWMYGC DVGPGRLLRGYCEAYDGDYISLNEDLRSWTANDIASQISKHKSEAVDEAHQRAYLQGPCVEW
Rattus QWMYGC DVGPGRLLRGYCEAYDGRDYISLNEDLRSWTATDMASQASKIKSEEVGEAHQRAYLQGPCVEW
cons

Homo LHKYLEKGKETLLHLEPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQQDGEGHTQDTEL VETRPAGDG
Pan LHKYLEKGKETLLHLEPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQQDGEGHTQDTEL VETRPAGDG
Mus LHRYLRLGNETLQRSDPPKAHVTHHPRSEDEVTLRCWALGFYPADITLTWQLNGEELTQDMEL VETRPAGDG
Rattus LHTYLHLGKETLLRSDPPKAHVTLHPRPEGDVTLRCWALGFYPADITLTWQLNGEDLTQDMEL VETRPAGDG
cons

Homo TFQKWA AVVVP SGEEQRYTCHVQHEGLPEPVTLRWKP--ASOPTIPVIGIAGLVLLGSVVS-GAVVAA---
Pan TFQKWA AVVVP SGEEQRYTCHVQHEGLPEPLTLRWKP--ASOPTIPVIGIAGLVLLGSVVS-GAVVAA---
Mus TFQKWA AVVVP LKGEQYYTCHVYHEGLPEPLTLRWEPPPTVSNMVIIAVL---VVLGAVIILGAVVAF---
Rattus TFQKWA AVVVP SGEEQRYTCHVQHEGLPEPLALRWEPSPFTDSSMPVIVVLGAVAIIGAVAIIGAV
cons

Homo VIWRKKSSGGKGSYSKAEWSDSAQGSES-HSL
Pan VMWRKKSSGGKGRSYSKAEWSDS-----A
Mus VMKRRRHIGVKG-CYAHVLGSKSFQTSQDWPQKA
Rattus VRRRKRTGEGK-SYAHVLGSKAFQISDWPQKA
cons

CLUSTAL format alignment by MAFFT FFT-NS-i (v7.487)

Podsumowanie

Sekwencje są do siebie najbardziej podobne ich środkiem, różnią się natomiast początkiem oraz końcem, tam jest najmniej podobieństw.

c)

Clustal

1: Homo	100.00	98.00	65.20	62.64
2: Pan	98.00	100.00	65.01	57.21
3: Mus	65.20	65.01	100.00	80.11
4: Rattus	62.64	57.21	80.11	100.00

Muscle

1: Homo	100.00	98.00	66.09	63.51
2: Pan	98.00	100.00	66.47	55.02
3: Mus	66.09	66.47	100.00	80.39
4: Rattus	63.51	55.02	80.39	100.00

Mafft

1: Homo	100.00	98.00	66.09	63.51
2: Pan	98.00	100.00	66.47	57.93
3: Mus	66.09	66.47	100.00	80.39
4: Rattus	63.51	57.93	80.39	100.00

Podsumowanie

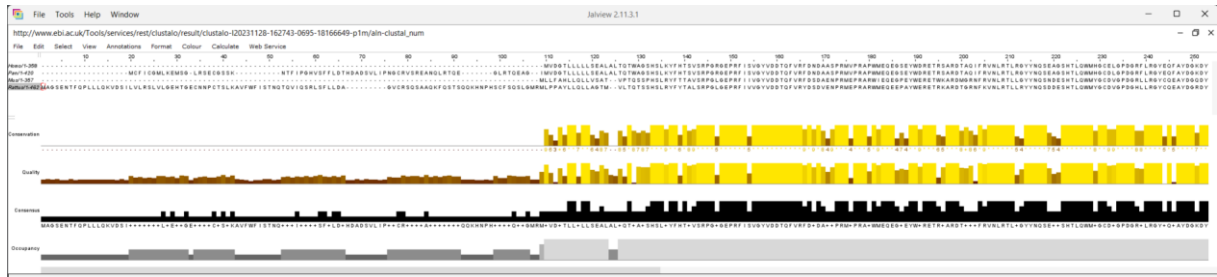
Człowiek – Szympan (98.00).

Mysz – Szczur (80).

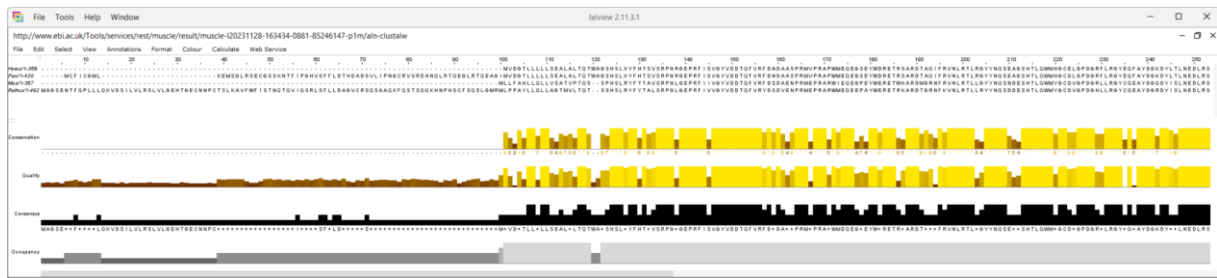
Myślę, że jest to zgodne z oczekiwaniami, ponieważ są to pary ze sobą spokrewnione i nawet po wyglądzie widać wiele podobieństw.

d)

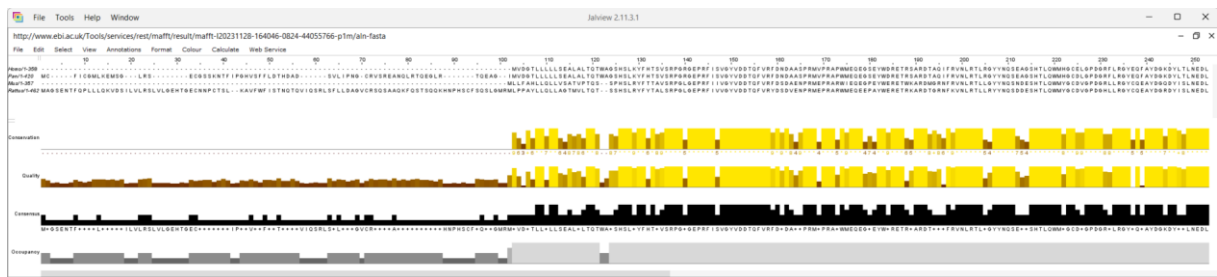
Clustal



Muscle



Mafft



Według dokumentacji + oznacza, że "jeśli wartość modalna jest wspólna dla więcej niż 1 reszty, na wyświetlaczu używany jest symbol „+” z prostego powodu, że nie jest możliwe wyświetlenie wielu znaków w jednym miejscu”.

e)

Clustal

150
|
G R G E P R I
G R G E P R I
G L G E P R I
G L G E P R I

320
|
T H H P I S I
T H H P I S I
T H H P R S E
T L H P R P E

Muscle

140
|
P G R G E P R
P G R G E P R
P G L G E P R
P G L G E P R

310
|
V T H H P I S
V T H H P I S
V T H H P R S
V T L H P R P

Mafft

170
|
R M V P R
R M V P R
R M E P R
R M E P R

300
|
L E K G K E T L
L E K G K E T L
L R L G N E T L
L H L G K E T L

f)

Clustal

Phylogram

Branch length: ☐ Cladogram ☒ Real



Homo 0.0209497
Pan 0.0209497
Mus 0.105042
Rattus 0.105042

Guide Tree

```
(  
(  
Homo:0.0209497  
,  
Pan:0.0209497  
):0.180751  
,  
(  
Mus:0.105042  
,  
Rattus:0.105042  
):0.0966583  
)  
;
```

Mafft

Phylogram

Branch length: ☐ Cladogram ☒ Real



1_Homo_sapiens 0.0405
2_Pan_troglodytes 0.0405
3_Mus_musculus 0.165
4_Rattus_norvegicus 0.165

Guide Tree

```
((  
1_Homo_sapiens  
:0.04050,  
2_Pan_troglodytes  
:0.04050):0.26385,(  
3_Mus_musculus  
:0.16500,  
4_Rattus_norvegicus  
:0.16500):0.13935);
```


Zadanie 2.

Clustal

CLUSTAL O(1.2.4) multiple sequence alignment

```
Gallus      MNGTEGINFYVPMSENKTVGVRSPFEYQYYLAEPWKYRLVCCYIFFLISTGLPINLLTLL      60
Danio       MNGTEGPAFYVPMSENATGVVRSPFEYQYYLVAPWAYGFVAAYMFFLIITGFPVNFLTLY      60
Bos         MNGTEGNFYVPMSENKTVGVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
Canis       MNGTEGNFYVPMSENKTVGVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
Homo        MNGTEGNFYVPMSENATGVVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
Pan         MNGTEGNFYVPMSENATGVVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
Mus         MNGTEGNFYVPMSENATGVVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
Rattus      MNGTEGNFYVPMSENATGVVRSPFEYQYYLAEPWQFSMLAAYMFLIIVLGFPINFLTLY      60
*****

Gallus      VTFKHKKLRQPLNYILLNLAVADLFMACFGFTVTFTAHNGYFVFGPVGCAVEGFFATLG      120
Danio       VTIEHKKLRTPLNLYILLNLAIADLFMVFGGFTTTMYTSLHGYFVFGRLGCNLEGGFATLG      120
Bos         VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGGFATLG      120
Canis       VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGGFATLG      120
Homo        VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTSTLYTSLHGYFVFGPTGCNLEGGFATLG      120
Pan         VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTSTLYTSLHGYFVFGPTGCNLEGGFATLG      120
Mus         VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGGFATLG      120
Rattus      VTVQHKLRTPLNLYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGGFATLG      120
**.*

Gallus      GQVALWSLVLAIERIYVVKPMGNFRFSATHAMGIAFTWMAFSCAAPPLFGWS----      176
Danio       GEMGLKSLVLAIERIMMVCKPVSNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
Bos         GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
Canis       GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      180
Homo        GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
Pan         GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
Mus         GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
Rattus      GEIALWSLVLAIERIYVVKPMGNFRFGENHAIMGVAFWMAACSCAVPLVGS----      176
**.*

Gallus      -----RYMPEGMQCSGPDYYTHNPDYHNESYVLYMFIHFIIPVWVIFFSYGRICKV      230
Danio       -----RYIPEGMQCSGPDYYTPTGVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
Bos         -----RYIPEGMQCSGPDYYTPHEETNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
Canis       HSPVLVRYIPEGMQCSGPDYYTLKPEVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      240
Homo        -----RYIPEGLQCSGPDYYTLKPEVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
Pan         -----RYIPEGLQCSGPDYYTLKPEVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
Mus         -----RYIPEGMQCSGPDYYTLKPEVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
Rattus      -----RYIPEGMQCSGPDYYTLKPEVNNEFVIYMFVHFIIPVWVIFFCYGRLVCTV      230
**.*

Gallus      REAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTL      290
Danio       KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTL      290
Bos         KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTI      290
Canis       KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTL      300
Homo        KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTI      290
Pan         KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTI      290
Mus         KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTL      290
Rattus      KEAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYASVAFYIFTHQGSDFGPIFMTL      290
:***

Gallus      PAFFSKSSSLYNPIIYVLMNKQFRNCMTTICCGKNPLGDDEAST--TVSKTETSQVAPA      350
Danio       PAFFAKTSAVYNPICIYIMNKQFRNCMTTLCCKGNPLGDDEAST--TASKTEASSVSSS      349
Bos         PAFFAKTSAVYNPVIYIMNKQFRNCMTTLCCKGNPLGDDEAST--TVSKTETSQVAPA      348
Canis       PAFFAKSSSIYNPVIYIMNKQFRNCMTTLCCKGNPLGDDEAST--SASKTETSQVAPA      358
Homo        PAFFAKSAAIYNPVIYIMNKQFRNCMTTICCGKNPLGDDEAST--TVSKTETSQVAPA      348
Pan         PAFFAKSAAIYNPVIYIMNKQFRNCMTTICCGKNPLGDDEAST--TVSKTETSQVAPA      348
Mus         PAFFAKSSSIYNPVIYIMNKQFRNCMTTLCCKGNPLGDDEAST--TASKTETSQVAPA      348
Rattus      PAFFAKTASIYNPVIYIMNKQFRNCMTTLCCKGNPLGDDEAST--TASKTETSQVAPA      348
***.*

Gallus      QVSPA      355
Danio       SVSPA      354
Bos         -----      348
Canis       -----      358
Homo        -----      348
Pan         -----      348
Mus         -----      348
Rattus      -----      348
```

- SLVVLAIER
- EVTRMVI
- TGVVRSP

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

Gallus	SQVSPA
Danio	SSVSPA
Bos	SQVAPA
Homo	SQVAPA
Pan	SQVAPA
Canis	SQVAPA
Mus	SQVAPA
Rattus	SQVAPA
	* * *

```

T-COFFEE, Version_11.00 (Version_11.00)
Cedric Notredame
SCORE=990
*
  BAD  AVG  GOOD
*
Homo      : 99
Pan       : 99
Canis     : 99
Bos       : 99
Mus       : 99
Rattus    : 99
Gallus    : 99
Danio     : 99
cons      : 99

Homo      MNGTEGPNFYVPFSNATGVVRSPFEYPQYYLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Pan       MNGTEGPNFYVPFSNATGVVRSPFEYPQYYLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Canis     MNGTEGPNFYVPFSNKTGVVRSPFEYPQYYLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Bos       MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Mus       MNGTEGPNFYVPFSNVTGVVRSPFEQPYLLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Rattus    MNGTEGPNFYVPFSNITGVVRSPFEQPYLLAEPWQFSMLAAYMFLLIIVLGFPINFLTLYVTVQHKKLRTPL
Gallus    MNGTEGINFYVPMNKTGVVRSPFEYPQYYLAEPWKYRLVCCYIFLITGLPINLLTLLYTFKHKKLRLQPL
Danio     MNGTEGPAFYVPMNATGVVRSPFEYPQYYLVAPWAYGFVAAYMFLLIITGFPVNFLLTYVTIEHKKLRTPL

cons      *****  *****,**  *****,*  *****  **  :  :.,*,*,**  *,*,*,***  **,*****

Homo      NYILLNLAVADLFMVLGGFTSTLYTSLHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Pan       NYILLNLAVADLFMVLGGFTSTLYTSLHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Canis     NYILLNLAVADLFMVFGGFTTLLYTSLSHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Bos       NYILLNLAVADLFMVFGGFTTLLYTSLSHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Mus       NYILLNLAVADLFMVFGGFTTLLYTSLSHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Rattus    NYILLNLAVADLFMVFGGFTTLLYTSLSHGYFVFGPTGCNLEGGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Gallus    NYILLNLAVADLFMACFGFTVTFYTAWNGYFVFGPVGCAVEGFFATLGGEIALWSLVVLAIERVYVVCCKPMS
Danio     NYILLNLAIADLFMVFGGFTTMYTSLHGYFVFGRLGCNLEGGFFATLGGEMLKSLVLAIERVMVVCCKPVS

cons      *****,**,* *****,  *** *,** : *****  ** : *****:,*,* *****:,*****:,

Homo      NFRFGENHAIMGVFTWVMALACAAPPLAGWS-----RYIPEGLQCSCGIDYYTLKPEVNNEFSVIY
Pan       NFRFGENHAIMGVFTWVMALACAAPPLAGWS-----RYIPEGLQCSCGIDYYTLKPEVNNEFSVIY
Canis     NFRFGENHAIMGVFTWVMALACAAPPLAGWSLLSHSPLVLRYIPEGMQCSCGIDYYTLKPEVNNEFSVIY
Bos       NFRFGENHAIMGVFTWVMALACAAPPLVGWS-----RYIPEGMQCSCGIDYYTPHEETNNEFSVIY
Mus       NFRFGENHAIMGVFTWVMALACAAPPLVGWS-----RYIPEGMQCSCGIDYYTLKPEVNNEFSVIY
Rattus    NFRFGENHAIMGVFTWVMALACAAPPLVGWS-----RYIPEGMQCSCGIDYYTLKPEVNNEFSVIY
Gallus    NFRFSATHAMMGIAFTWVMAFSCAAPPLFGWS-----RYMPEGMQCSCGPDYYTHNPDYHNEFSVLY
Danio     NFRFGENHAIMGVFTWVMACSCAVPPLVGWS-----RYIPEGMQCSCGVDYYTRTPGVNNEFSVIY

cons      *****,**,* *****,** : *****,** *****,** *****  *****,** : *****,**

Homo      MFVVHFTIPMIIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTH
Pan       MFVVHFTIPMIIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTH
Canis     MFVVHFAIPMIVIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTH
Bos       MFVVHFIPLIVIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWLPYAGVAFYIFTH
Mus       MFVVHFTIPMIVIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIFFLICWLPYASVAFYIFTH
Rattus    MFVVHFTIPMIVIFFCYGQLVFTVKEAAAQQQESATTQKAKEVTRMVIIMVIFFLICWLPYASVAFYIFTH
Gallus    MFVZHFIIPVVVIFFSYGRLLCKVREAAAQQQESATTQKAKEVTRMVLVGLFMAWTPYAVVAFWIFTH
Danio     MFIVHFFIPLIVIFFCYGRLLCTVKEAARQQQESATTQRAEREVTRMVIIMVIAFLICWLPYAGVAFYIFTH

cons      **:*,** **:*,**,* **:*,** : *,** ***** **:*,**,* **:*,** : *,** *:*,** :*,**

Homo      QGSNFGPIFMTIPAFFAKSAAIYNPVIYIMMNKQFRNCMLTTCGKNPLGDDEA-SAT--VSKTETS----
Pan       QGSNFGPIFMTIPAFFAKSAAIYNPVIYIMMNKQFRNCMLTTCGKNPLGDDEA-SAT--VSKTETS----
Canis     QGSDFGPIFMTLPAFFAKSSSIYNPVIYIMMNKQFRNCMLTTCGKNPLGDDEA-SAS--ASKTETS----
Bos       QGSDFGPIFMTIPAFFAKTSAYYNPVIYIMMNKQFRNCMLTTCGKNPLGDDEA-STT--VSKTETS----
Mus       QGSNFGPIFMTLPAFFAKSSSIYNPVIYIMLNKQFRNCMLTTCGKNPLGDDEA-SAT--ASKTETS----
Rattus    QGSNFGPIFMTLPAFFAKTASIYNPVIYIMMNKQFRNCMLTTCGKNPLGDDEA-SAT--ASKTETS----
Gallus    KGADFTATLMAVPAFFSKSSSLYNPVIYVLMNKQFRNCMLTTCGKNPLFGDEDV-SSTVSQSKTEVS----
Danio     QGSEFGPVFMTLPAFFAKTSAYYNPCYICMNKQFRNCMLTTCGKNPFEEEGASTT--ASKTEASSVSS

cons      :*,** : *,**,* **:*,** : *,** **:*,** ***** **:*,**,* **:*,** : *,** *:*,**

Homo      -QVA-----PA
Pan       -QVA-----PA
Canis     -QVA-----PA
Bos       -QVA-----PA
Mus       -QVA-----PA
Rattus    -QVA-----PA
Gallus    -SVSSSQVSPA
Danio     SSVS-----PA

cons      *,**

```

Mafft

CLUSTAL format alignment by MAFFT FFT-NS-i (v7.487)

```
Homo      MNGTEGPNFYVPFSNATGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Pan       MNGTEGPNFYVPFSNATGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Canis     MNGTEGPNFYVPFSNKTGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Mus       MNGTEGPNFYVPFSNVTGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Rattus    MNGTEGPNFYVPFSNITGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Bos       MNGTEGPNFYVPFSNKTGVVRSPFEYQYYLAEPWQFSMLAAYMFLILVLGFPINFLTLY
Danio     MNGTEGPAFYVPMSNATGVVRSPYEQYYLVAPWAYGFVAAYMFFLIITGFPVNFLTLY
Gallus    MNGTEGINFYVPMSNKTGVVRSPFEYQYYLAEPWKYRLVCCYIFFLISTGLPINLLTLL
          *****
          *****:*****:*****:*****:*****:*****:*****:*****:*****
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Pan       VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Canis     VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Mus       VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Rattus    VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Bos       VTVQHKKLRTPLNYILLNLAVADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Danio     VTIHKKLRTPLNYILLNLAIADLFMVVGGFSTLYTSLHGYPVFGPTGCNLEGGFATLG
Gallus    VTFKHKKLRLQPLNYILVNLAVADLFMACFGFTVTFYTAWNGYFVFGPVGCAVEGGFATLG
          **.:*****
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      GEIALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLAGWS----
Pan       GEIALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLAGWS----
Canis     GEIALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLAGWSLLS
Mus       GEIALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVFTWMMALACAAPPLVGS----
Rattus    GEIGLWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLVGS----
Bos       GEIALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLVGS----
Danio     GEMGLKSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLVGS----
Gallus    GQVALWSLVVLAIERVYVVKPMSNFRFGENHAIMGVAFTWMMALACAAPPLVGS----
          **.:*****
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      -----RYIEGLQCSCGIDYYTLKPEVNNEFSVIYMFVHFIPMIIFFCYGQLVFTV
Pan       -----RYIEGLQCSCGIDYYTLKPEVNNEFSVIYMFVHFIPMIIFFCYGQLVFTV
Canis     HSPLVLRYPYIEGMQSCGIDYYTLKPEVNNEFSVIYMFVHFAIPMIVIFFCYGQLVFTV
Mus       -----RYIEGMQSCGIDYYTLKPEVNNEFSVIYMFVHFIPMIVIFFCYGQLVFTV
Rattus    -----RYIEGMQSCGIDYYTLKPEVNNEFSVIYMFVHFIPMIVIFFCYGQLVFTV
Bos       -----RYIEGMQSCGIDYYTPHEETNNEFSVIYMFVHFIPMIVIFFCYGQLVFTV
Danio     -----RYIEGMQSCGVDYYTRPGVNNEFSVIYMFVHFIPMIVIFFCYGRLVCTV
Gallus    -----RYMEGMQSCGPDYYTHNPDYHNESVYLYMFVHFIPMIVIFFSYGRICKV
          **.:*****
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTI
Pan       KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTI
Canis     KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSDFGPIFMTL
Mus       KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTL
Rattus    KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYASVAFYIFTHQGSNFGPIFMTL
Bos       KEAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYAGVAFYIFTHQGSDFGPIFMTI
Danio     KEAARQQQESATTQKAEREVTRMVIIMVIAFLICWVPYAGVAFYIFTHQGSDFGPIFMTL
Gallus    REAAAQQQESATTQKAKEVTRMVIIMVIAFLICWVPYAGVAFYIFTHQGSDFGPIFMTL
          :***
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      PAFFAKSAAIYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASATV--SKTE----T
Pan       PAFFAKSAAIYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASATV--SKTE----T
Canis     PAFFAKSSSIYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASASA--SKTE----T
Mus       PAFFAKSSSIYNPVIYIMLNKQFRNCMLTTICCGKNPLGDDE-ASATA--SKTE----T
Rattus    PAFFAKTASIYNPIIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASATA--SKTE----T
Bos       PAFFAKTSAVYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASTTV--SKTE----T
Danio     PAFFAKTSAVYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASTTV--SKTE----T
Gallus    PAFFAKTSAVYNPVIYIMMNKQFRNCMLTTICCGKNPLGDDE-ASTTV--SKTE----T
          *****:*****:*****:*****:*****:*****:*****:*****:*****

Homo      SQVAPA
Pan       SQVAPA
Canis     SQVAPA
Mus       SQVAPA
Rattus    SQVAPA
Bos       SQVAPA
Danio     SSVSPA
Gallus    SQVSPA
          **.:*****
```


Podsumowanie

Porównując wyniki z tymi zawartymi w poprzednim zadaniu zauważamy, że sekwencje aminokwasów są do siebie dużo bardziej podobne.

c)

Clustal

Percent Identity Matrix - created by Clustal2.1								
1: Gallus	100.00	68.93	70.69	72.41	70.11	70.11	70.11	69.83
2: Danio	68.93	100.00	81.03	79.89	78.74	78.74	79.31	80.75
3: Bos	70.69	81.03	100.00	94.25	93.39	93.39	93.39	93.68
4: Canis	72.41	79.89	94.25	100.00	95.40	95.40	95.69	95.40
5: Homo	70.11	78.74	93.39	95.40	100.00	100.00	94.83	95.11
6: Pan	70.11	78.74	93.39	95.40	100.00	100.00	94.83	95.11
7: Mus	70.11	79.31	93.39	95.69	94.83	94.83	100.00	97.13
8: Rattus	69.83	80.75	93.68	95.40	95.11	95.11	97.13	100.00

Muscle

Percent Identity Matrix - created by Clustal2.1								
1: Gallus	100.00	69.12	72.13	71.55	71.55	72.99	71.26	70.98
2: Danio	69.12	100.00	82.47	79.89	79.89	81.03	80.46	81.90
3: Bos	72.13	82.47	100.00	93.39	93.39	94.25	93.39	93.68
4: Homo	71.55	79.89	93.39	100.00	100.00	95.40	94.83	95.11
5: Pan	71.55	79.89	93.39	100.00	100.00	95.40	94.83	95.11
6: Canis	72.99	81.03	94.25	95.40	95.40	100.00	95.69	95.40
7: Mus	71.26	80.46	93.39	94.83	94.83	95.69	100.00	97.13
8: Rattus	70.98	81.90	93.68	95.11	95.11	95.40	97.13	100.00

Mafft

Percent Identity Matrix - created by Clustal2.1								
1: Homo	100.00	100.00	95.40	94.83	95.11	93.39	79.89	71.55
2: Pan	100.00	100.00	95.40	94.83	95.11	93.39	79.89	71.55
3: Canis	95.40	95.40	100.00	95.69	95.40	94.25	81.03	72.99
4: Mus	94.83	94.83	95.69	100.00	97.13	93.39	80.46	71.26
5: Rattus	95.11	95.11	95.40	97.13	100.00	93.68	81.90	70.98
6: Bos	93.39	93.39	94.25	93.39	93.68	100.00	82.47	72.13
7: Danio	79.89	79.89	81.03	80.46	81.90	82.47	100.00	69.12
8: Gallus	71.55	71.55	72.99	71.26	70.98	72.13	69.12	100.00

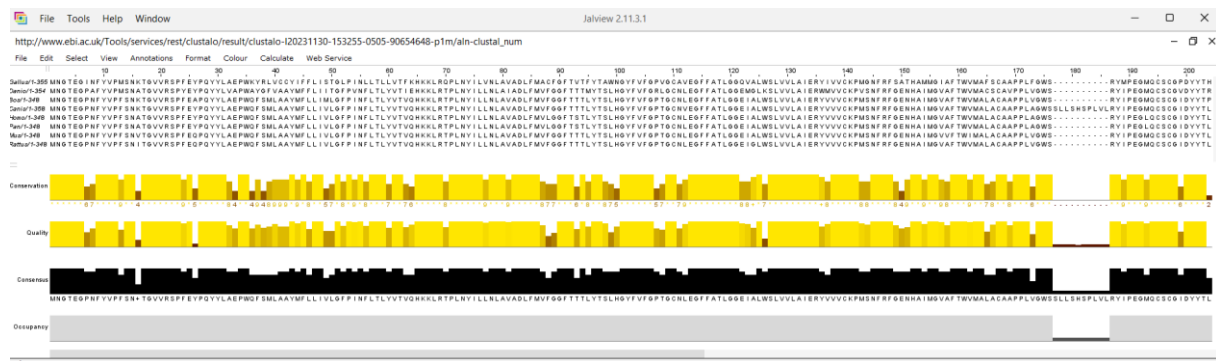
Człowiek – Szympans (100.0)

Mysz – Szczur (97.13)

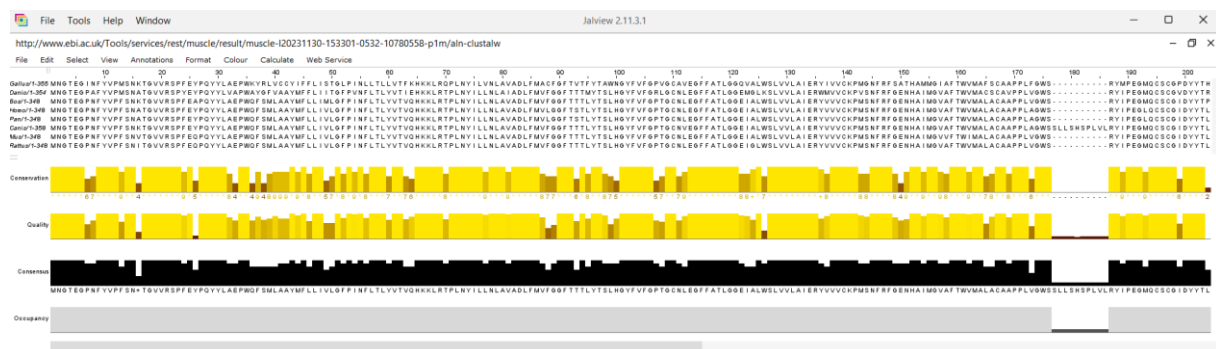
Tak jak w poprzednim zadaniu jest to sytuacja spodziewana ze względu na pokrewieństwo gatunków w parach. Zaskakującym jednak jest, iż wartość ta jest praktycznie bliska 100, czyli maksymalnej wartości. W przypadku człowieka i szympansa jest to właśnie dokładnie 100. Zatem sekwencje rodopsyny zarówno u człowieka jak i szympansa są takie same.

d)

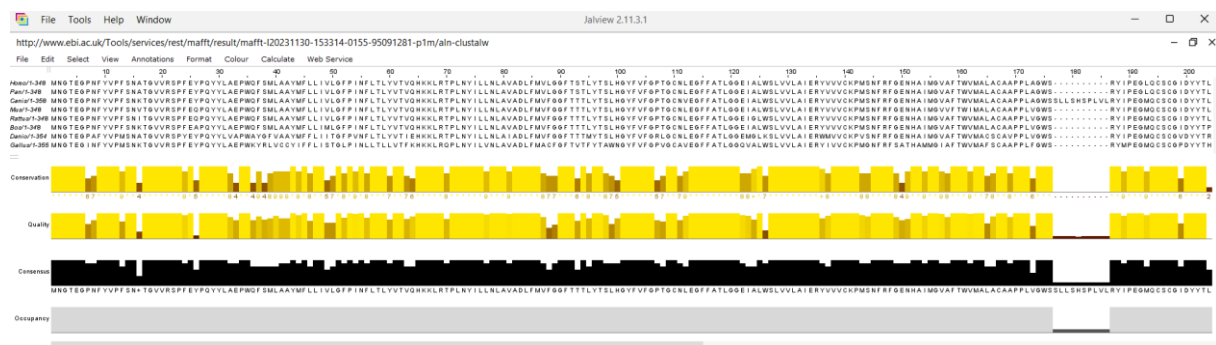
Clustal



Muscle



Mafft



e)

Clustal

210
THNPDYHNE
TRTPGVNNE
TPHEETNNE
TLKPEINNNE
TLKPEVNNE
TLKPEVNNE
TLKPEVNNE
TLKPEVNNE
TLKPEVNNE

SPFEYPC
SPYEYPC
SPFEAPC
SPFEYPC
SPFEYPC
SPFEYPC
SPFEYPC
SPFEQPC
SPFEQPC 24, 26

Muscle

210
HNP DYHNE
RTPGVNNE
PHEETNNE
LKPEVNNE
LKPEVNNE
LKPEINNNE
LKPEVNNE
LKPEVNNE

SPFEYPC
SPYEYPC
SPFEAPC
SPFEYPC
SPFEYPC
SPFEYPC
SPFEYPC
SPFEQPC
SPFEQPC 24, 26

Mafft

FEYPC
FEYPC
FEYPC
FEQPC
FEQPC
FEAPC
YEYPC
FEYPC 24,26

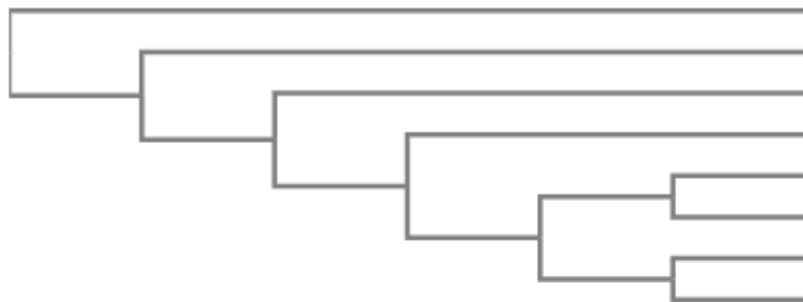
210
TLKPEVNNE
TLKPEVNNE
TLKPEINNNE
TLKPEVNNE
TPHEETNNE
TRTPGVNNE
THNPDYHNE

f)

Clustal

Phylogram

Branch length: ☒ Cladogram ☐ Real



Gallus 0.154533
Danio 0.101794
Bos 0.0328664
Canis 0.0269397
Homo 0
Pan 0
Mus 0.0143678
Rattus 0.0143678

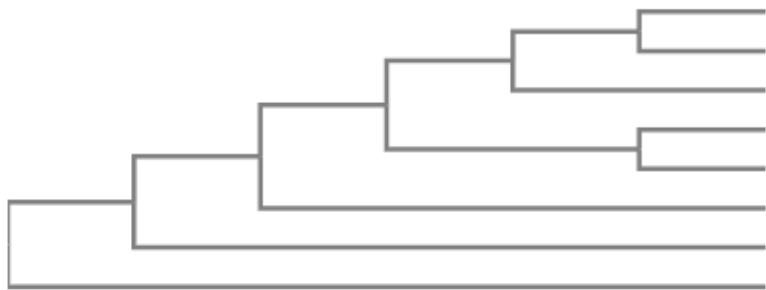
Guide Tree

```
(  
Gallus:0.154533  
,  
(  
Danio:0.101794  
,  
(  
Bos:0.0328664  
,  
(  
Canis:0.0269397  
,  
(  
(  
Homo:0  
,  
Pan:0  
) :0.0251437  
,  
(  
Mus:0.0143678  
,  
Rattus:0.0143678  
) :0.0107759  
) :0.00179598  
) :0.00592672  
) :0.0689275  
) :0.0527392  
)  
;
```

Mafft

Phylogram

Branch length: ☒ Cladogram ☐ Real



1_Homo_sapiens 0
2_Pan_troglodytes 0
3_Canis_familiaris 0.0235
5_Mus_musculus 0.012
6_Rattus_norvegicus 0.012
4_Bos_taurus 0.03772
8_Danio_rerio 0.12262
7_Gallus_gallus 0.18478

Guide Tree

```
(((((  
1_Homo_sapiens  
:0.00000,  
2_Pan_troglodytes  
:0.00000):0.02350,  
3_Canis_familiaris  
:0.02350):0.00213,(  
5_Mus_musculus  
:0.01200,  
6_Rattus_norvegicus  
:0.01200):0.01363):0.01209,  
4_Bos_taurus  
:0.03772):0.08490,  
8_Danio_rerio  
:0.12262):0.06216,  
7_Gallus_gallus  
:0.18478);
```

Zadanie 3

Z obserwacji wynika, że bardziej konserwatywna jest rodopsyna. Myślę, że jest tak dlatego że jak w treści zadania jest napisane, to białko występuje w siatkówce oka, co za tym idzie, jest powiązane ze wzrokiem. Wzrok odgrywa bardzo ważną rolę w przetrwaniu gatunku, dlatego też, to białko znalazło się w tak wielu gatunkach.