

## BGGN-213: FOUNDATIONS OF BIOINFORMATICS

The find-a-gene project assignment

[amudla@ucsd.edu](mailto:amudla@ucsd.edu)

Anusorn Mudla (A53074020)

[Q1]

Name: STAT1 (Signal Transducer and Activator of Transcription 1)

Accession: NP\_009330.1

Species: Homo Sapiens

[Q2]

Method: TBLASTN2.9.0+

Database: Expressed sequence tags (est)

Organism: Mustela furo (taxid:9669)

**BLAST** » tblastn

Translated BLAST: tblastn

blastn blastp blastx **tblastn** tblastx

TBLASTN search translated nucleotide databases using a protein query. [more...](#)

**Enter Query Sequence**

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

NP\_009330.1

Query subrange [From](#)  [To](#)

Or, upload file  no file selected [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

**Choose Search Set**

Database  [?](#)

Organism [Optional](#)  ☐ exclude [+](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

Exclude [Optional](#) ☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

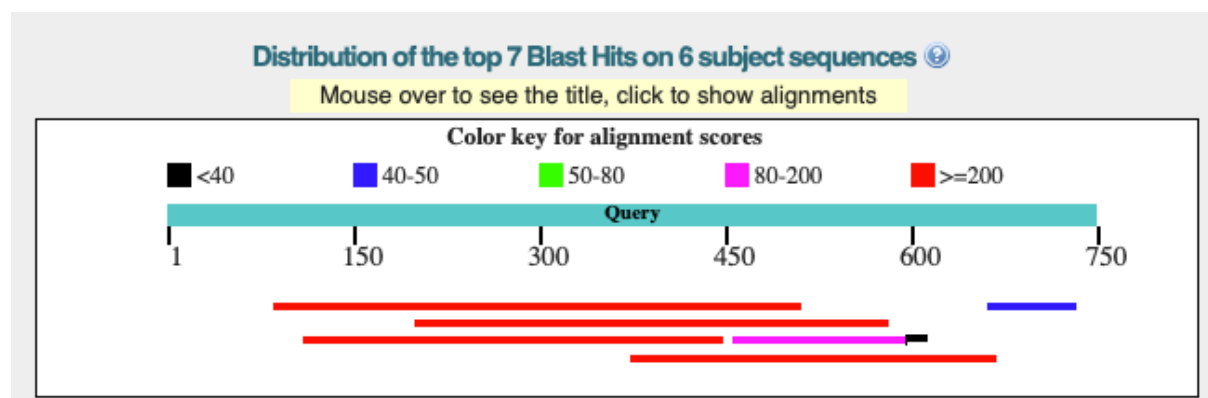
Limit to [Optional](#) ☐ Sequences from type material

Entrez Query [Optional](#)  [YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

**BLAST** Search database **Expressed sequence tags (est)** using **Tblastn** (search translated nucleotide databases using a protein query)

☐ Show results in a new window



## Sequences producing significant alignments:

Select: [All](#) [None](#) Selected:0[Alignments](#) [Download](#) [GenBank](#) [Graphics](#)

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input type="checkbox"/>	<a href="#">Ferret_003_D03 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	679	679	56%	0.0	81.78%	<a href="#">GD181285.1</a>
<input type="checkbox"/>	<a href="#">Ferret_015_J19 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	616	616	50%	0.0	80.93%	<a href="#">GD184919.1</a>
<input type="checkbox"/>	<a href="#">Ferret_013_K21 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	279	279	45%	3e-87	44.96%	<a href="#">GD184235.1</a>
<input type="checkbox"/>	<a href="#">Ferret_010_E11 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	276	276	39%	3e-86	47.65%	<a href="#">GD183067.1</a>
<input type="checkbox"/>	<a href="#">Ferret_015_C03 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	188	219	20%	3e-53	59.57%	<a href="#">GD184746.1</a>
<input type="checkbox"/>	<a href="#">Ferret_004_H24 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence</a>	45.4	45.4	9%	2e-05	43.06%	<a href="#">GD181745.1</a>

[Alignments](#)[Download](#) [GenBank](#) [Graphics](#)

## Ferret\_003\_D03 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence

Sequence ID: [GD181285.1](#) Length: 1280 Number of Matches: 1Range 1: 2 to 1270 [GenBank](#) [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
679 bits(1752)	0.0	Compositional matrix adjust.	362/428(85%)	379/428(88%)	7/428(1%)	+2
Query 87		KRNLQDNFQEDPIQMSMIIYSCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV				146
Sbjct 2		KRNLQDNFQEDP+QMSMII +CLKEER+ILENAQRFNQAQSG+IQSTVMLDKQKELDSKV				181
Query 147		RNVKDKVMCIEHEIKSLEDLQDEYDFCKCTLQNHETNGVAKSDQKQEQLLKKMYLML				206
Sbjct 182		RNVKDKVMCIEHEIK+LEDLQDEYDFCKCTLQNHETNGVAK+DQKQEQLL++KMYLML				361
Query 207		DNKRKEVVHKKIIELLNVTELTQNALINDELVEWKRQSSACIGGPPNACLDQLQNWFTIV				266
Sbjct 362		DNKRKEVV KIIIELLNVTELTQ ALINDELVEWK+RQSSACIGGPPNACLDQLQNWFTIV				541
Query 267		AESLQQVRqgqlkkleeleqkyTYEHDPITKNKQVLWDRFTSLFQQLIQSSFFVVERQPCMP				326
Sbjct 542		AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQ LWDRTSLFQQLIQSSFFVVERQPCMP				721
Query 327		THPQRPLVLKTGVQFTVKLRLLVKLQELNLYNLKVKVLFDKDVNERNTVKGRKFNLGTH				386
Sbjct 722		THPQRPLVLKTGVQFTVKLRLLVKLQELNLYNLKVKVLFDKDVNERNTVKGRKFNLGTH				901
Query 387		TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPLVI				446
Sbjct 902		TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEEL LV LSFET				1075
Query 447		DLETTSLPVVVISNVSQ--LPSGWASILWYNMLVAEPRNLSFFLTPPCARWAQLSEVLSW				504
Sbjct 1076		L T+ P+ ++ LP GWASIL ++ P + F PP + L E SW				1246
Query 505		QFSSVTKR 512				
Sbjct 1247		SFSSSHKR 1270				

>GD181285.1:2-1270 Ferret\_003\_D03 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence

AAGCGTAATCTTCAGGATAATTTTCAGGAAGACCCAGTACAGATGTCTATGATCATCTGTAACGTGTCTGA  
AGGAAGAAAGGAGGATCCTGGAAAATGCTCAGAGATTCAATCAGGCGCAGTCGGGGAGTATCCAGAGCAC  
CGTGATGTTAGACAAACAGAAGGAGCTGGACAGCAAAGTCAGAAATGTGAAGGATAAGGTTATGTGTATA  
GAGCATGAAATAAAGACCTTAGAAGATTTACAAGACGAGTATGACTTCAAGTGCAAAACCTTGCAGAACAA  
GAGAGCATGAAACCAATGGCGTGGCAAAGAATGATCAGAAACAAGAACAGTTGTTAATGCAGAAAGATGA  
TTTAATGCTCGACAATAAGAGAAAGGAAGTAGTTCTCAAATAATAGAGTTGTTGAATGTCACGGAAGCTT  
ACCCAGAAAGCCCTCATGATGAGCTGGTGGAATGGAAGCAGAGACAACAGAGCGCCTGCATCGGGG  
GGCCCCCAACGCTTGTCTTGATCAGCTGCAGAACTGGTTTACCATAGTTGCGGAGAGTCTGCAGCAGGT  
GCGTCAGCAGCTGAAAAAGCTCGAGGAGCTGGAACAGAAATACACCTATGAACACGACCCCTATCACAAAA  
AACAAACAAGGCCTGTGGGACCGCACCTTCAGCCTTTTCCAGCAGCTCATTCAGAGCTCATTTGTGGTGG  
AAAGACAGCCTTGCATGCCAACTCATCCTCAGAGGCCGCTGGTCTTGAAGACCGGGGTCCAGTTTACCCTG  
GAAGTTGAGACTGCTGGTGAAATTGCAAGAGCTGAATTATAATTTGAAAGTCAAAGTCTTATTTGATAAA  
GATGTGAATGAGAGGAATACAGTAAAAGGATTGAGGAAGTTCAACATTTTGGGCACGCACACGAAAGTGA  
TGAACATGGAGGAGTCCACNAATGGGAGTCTGGCGCGGAGTTCCGGCACCTGCAACTGAAAGAACAGAA  
AAATGCTGGCACCAGAACTAACGAGGGCCCTCTCATCGTTACTGAGGAGCTTTCACCTCCTCAGTTTGTAG  
ACCACCTGTTCCAGCCTGGTTTGGTCATTGACCTCAGAACGACCTCTCGGCCCGTTTGGGGATCTCCAAG  
GCCGCCCACTTCCGAGAGGTTGGGCCTCCATCCTGGTGGAACCATCGTGGGAACAAACCCCGAAATTGT  
CCTTCTTTTAAACCCCTGGTGACGAAGGTTGAGCTCTTGGAGAGACTAGCTGGAGTTTCTTCTTCTCA  
CACAAAAGG

[Q3]

Method: Use EMBOSS Transeq to translate DNA sequence from [Q2]. Below are the 6 ORF results.

```
>EMBOSS_001_1
KRNLQDNFQEDPVQMSMIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV
RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQLLIQKMYLML
DNKRKEVVLKIIELLNVTELTQKALINDELVEWKQRQQSACIGGPPNACLDQLQNWFTIV
AESLQQVRQQLKKLEELEQKYTYEHDPIITKNKQGLWDRTFSLFQQLIQSSFFVVERQPCMP
THPQRPLVLKTGVQFTVKLRLLVKLQELNYNLKVVLFDKDVNERNTVKGFRKFNILGTH
TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELSLLSFETTCSSLVWSL
TSERPLGPFQDLQGRPLPRGWASILVETIVGTNPEIVLLFNPPGDEGSALGETSWSFSSS
HKR
>EMBOSS_001_2
SVIFRIIFRKTQYRCL*SSVTV*KKGGGSKMLRDSIRRSRGVSRA*P*CTNRRSWTAKS
EM*RIRLCV*SMK*RP*KIYKTSMTSSAKPCRTESMKPMWQRMIRKNKNC*SRRCI*CS
TIRERK*FSK**SC*MSKNLPRKPSMMMSWNGSRDNAPASGGPPTLVLISCRGTSP*L
RRVCSRCVSS*KSSRSWNRNTPMNTTLSQKTNKACGTAPSAFSSSSFRAHLWWDLSLACQ
LILGRWS*RPSSSP*S*DCW*NCKS*II*I*SKSYLIK*MRGIQ*KDSGSSFTWART
RK**TWRSPXMGVWRRSSGTCN*KNRKMLAPELTRALSLLRSFSSVLRPPVPANFGH*
PQNDLSARLGISKAHFREVGPPSPWPKSWEQTPKLSFFLTPLVTQVQLLERLAGVFLH
TKG
>EMBOSS_001_3
A*SSG*FSGRPSTDVYDHL*LSEGRKEDPGKCSEIQSGAVGEYPEHRDVRQTEGAGQOSQ
KCEG*GYVRA*NKDLRRFTRRV*LQVQNLAEQRA*NQWRGKE*SETRTVNPNEDVNAR
Q*EKSSSQNNRVVECHGTYPESPHQ*AGGHEAETTERLHRGAPQRLS*SAAELVHHSC
GESAAGASAAEKARGAGTEIHL*TRPYHKKQTRPVGPHLQPPFAAHSLEICGGKTALHAN
SSSEAAGLEDRGFVHREVETAGEIARAEL*FESQSLI**RCE*EYSKRQIEVQHFHGAH
ESDEHGGVHXWESGGVPAPATERTEKWHQN*RGPSHRY*GAFT*PQF*DHLFQPGLVID
LRTTSRPVWGSFRPPTSERLGLPHGPNHRGNKPRNCPSF*PPW*RRFSWRD*LEFFFT
QKX
>EMBOSS_001_4
PFV*RRKTPASLSKS*TFVTRGVKKDNFVGVCHDGHQDGGPSTRKWAALIPKRAERS
F*GQ*PNQAGTGGLKTEE*KLLSNDERLALVSSGASIFLFFQLOVPELRRQTPIXGLLHVH
HFRVRAQNVLPESFYCIPLIHIFIK*DFDFQIIQLLQFHQSSQLHGEIDFGLQDQRPL
RMSWHARLSFHHK*ALNELLEKAEGAVPQALFVFCDRVVFVIGVFLFQLLELQLLTHLLQ
TLRNYGEPVLQIKTSVGGPPDAGALLSLPFFHQLIIDEGFLGKFRDIQQLYYFENYFLS
LIVEH*IHLLD*QLFLFLIILCHAIGFMLSVLQGFALVILVL*IF*GLYFMYLTHNLIL
HISDFAVQLLLFV*HHGALDTPRLRLIESLSIFQDPPFFLTQTVTDHRRHLYWVFLKILK
ITL
>EMBOSS_001_5
FCVKKNSS*SLQELNLRHQGG*KEGQFRGLFPRWFFPGWRPNLSEVGGLDGPQTGREVV
LRSMTPGWNRSQN*GVKAPO*R*EGPR*FWCQHSVLSVAGAGTPPFDSSHWTFFCSS
LSCACPKC*TS*ILLYSSSHLYQIRL*LSNYSALAI SPAVTSR*TGPRSSRPAASE
DELACKAVFPQMSSE*AAGKG*RCGPTGLVCFL*GRVHRCISVPAPRAFSAADAPAAD
SPQLW*TSAADQDKRWGAPRCRRSVVSASIPPAHH*GLSG*VP*HSTTLF*ELLPFS
YCRALNTSSGLTTVLVSDHSLPRHWFHALCSARFCT*SHTRLVNLLRSLFHALYTPYPS
HF*LCCPAPSVCILSRCSGYSPTAPD*ISEHFGSSFLPSDSYR*S*TSVLGLPENYPED
YAX
>EMBOSS_001_6
LLCEEELQLVSPRAEPSSPGGLKRRITSGFVPTMVSTRMEAQPLGSGRPWRSPNGPRGR
SEVNDQTRLEQVVSKLRSSESSVTMRGPSLVLPVAFCCSFSCRCRNSAARLPXVDSSMFI
TFVCVPMKLNFLNPFVFLSFTSLSNKTLTKFL*FSSCNFTSSLNFTVNMVTVFKTSGL*
G*VGMQGLCLSTTNEL*MSCWKRLKVRSHRPLCFVIGGSCS*VYFCSSSSSFFSC*RTCCR
LSATHMVQFCS*SRQALGGPPMQALCCLCFHSTSSSLMRAFWVSSVTFNNSIILRTTSFL
LLSSIKYIFWINNCSCF*SFATPLVSCSLFCKVLHLKSYSSCKSSKVFISCSIHITLSF
TFTLLSSSFCLSNITVLWLPDCA*LNL*AFSRILLSFRQLQMIIDICTGSS*KLS*R
LRX
```

Chosen Sequence: The first ORF because it has the longest amino acid before a stop codon.

>Ferret STAT1 (sequence taken from EMBOSS Transeq)

```
KRNLQDNFQEDPVQMSMIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV
RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQLLIQKMYLML
DNKRKEVVLKIIELLNVTELTQKALINDELVEWKQRQQSACIGGPPNACLDQLQNWFTIV
AESLQQVRQQLKKLEELEQKYTYEHDPIITKNKQGLWDRTFSLFQQLIQSSFFVVERQPCMP
THPQRPLVLKTGVQFTVKLRLLVKLQELNYNLKVVLFDKDVNERNTVKGFRKFNILGTH
TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELSLLSFETTCSSLVWSL
TSERPLGPFQDLQGRPLPRGWASILVETIVGTNPEIVLLFNPPGDEGSALGETSWSFSSS
HKR
```

Name: Ferret STAT1

Organism: Mustela putorius furo

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia;  
Mustelidae; Mustelinae; Mustela.

[Q4] : BLASTP search

BLAST®

» blastp suite

Home

Recent Results

Saved Strategies

Help

Standard Protein BLAST

blastnblastpblastxtblastntblastx

BLASTP programs search protein databases using a protein query. [more...](#)

Reset page

Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

>EMBOSS\_001.1  
KRNLDQNFQEDPVQMSMICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV  
RNVKDKVMCIEHIKTLLEDQDEYDFKCKTLQNREHETNGVAKNDQKQELLIQKMYML  
DNKRKEVVLKIELLNVLTETQKALINDELVEWKQRQSQACIGGPPNACLDQLQNWFTIV  
AESLQQVRQLKLEELQKYYEHPITKNKQGLWDRFTSLFQQLIQSSFVVERQPCMP

Query subrange [From](#)   
[To](#)

Or, upload file

Choose File

Job Title

Enter a descriptive title for your BLAST search [Job Title](#)

☐ Align two or more sequences [Align two or more sequences](#)

Choose Search Set

Database

Non-redundant protein sequences (nr) [Database](#)

Organism

Optional

Enter organism name or id-completions will be suggested ☐ exclude [Organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [Organism](#)

Exclude

Optional

☐ Models (XM/XP) ☐ Non-redundant RefSeq proteins (WP) ☐ Uncultured/environmental sample sequences

Program Selection

Algorithm

☐ Quick BLASTP (Accelerated protein-protein BLAST)

☒ blastp (protein-protein BLAST)

☐ PSI-BLAST (Position-Specific Iterated BLAST)

☐ PHI-BLAST (Pattern Hit Initiated BLAST)

☐ DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm [Choose a BLAST algorithm](#)

BLAST

Search database nr using Blastp (protein-protein BLAST)

☐ Show results in a new window

We are beta testing a New Results page

☐ Click here if you would like to see your results in the new format. You can always switch back to the Traditional Results page.

New

The top result is to a STAT1 protein from *Mustela putorius furo*.

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

Alignments

[Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Mustela putorius furo]	732	732	100%	0.0	88.32%	<a href="#">XP_012917317.1</a>
<input type="checkbox"/>	unnamed protein product [Gulo gulo]	731	731	100%	0.0	87.85%	<a href="#">VCW91131.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X1 [Enhydra lutris kenyoni]	729	729	100%	0.0	87.85%	<a href="#">XP_022354714.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X2 [Enhydra lutris kenyoni]	729	729	100%	0.0	87.85%	<a href="#">XP_022355001.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Ursus maritimus]	724	724	100%	0.0	86.92%	<a href="#">XP_008685758.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta [Ursus arctos horribilis]	723	723	100%	0.0	86.92%	<a href="#">XP_026348094.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-like [Leptonyctotes weddellii]	721	721	100%	0.0	86.68%	<a href="#">XP_006727034.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Odobenus rosmarus divergens]	720	720	100%	0.0	86.45%	<a href="#">XP_004392469.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Alluoropoda melanoleuca]	720	720	100%	0.0	86.45%	<a href="#">XP_002918048.2</a>
<input type="checkbox"/>	hypothetical protein PANDA_006420 [Alluoropoda melanoleuca]	720	720	100%	0.0	86.45%	<a href="#">EFB28593.1</a>
<input type="checkbox"/>	signal transducer and activator of [Lynx pardinus]	717	1048	100%	0.0	85.98%	<a href="#">VFV43934.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta [Felis catus]	716	716	100%	0.0	85.98%	<a href="#">XP_006935505.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X2 [Puma concolor]	716	716	100%	0.0	85.98%	<a href="#">XP_025788680.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X1 [Vulpes vulpes]	716	716	100%	0.0	85.75%	<a href="#">XP_025858431.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X2 [Vulpes vulpes]	715	715	100%	0.0	85.75%	<a href="#">XP_025858432.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta [Canis lupus dingo]	715	715	100%	0.0	85.75%	<a href="#">XP_025297339.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Ceratotherium simum simum]	714	714	100%	0.0	85.51%	<a href="#">XP_004426798.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta [Neophocaena asiaeorientalis asiaeorientalis]	714	714	100%	0.0	85.28%	<a href="#">XP_024589435.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta isoform X1 [Camelus bactrianus]	713	713	100%	0.0	85.51%	<a href="#">XP_010947017.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta isoform X3 [Camelus ferus]	713	713	100%	0.0	85.51%	<a href="#">XP_006186875.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1-alpha/beta isoform X1 [Delphinapterus leucas]	712	712	100%	0.0	85.05%	<a href="#">XP_022422976.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta isoform X1 [Vicugna pacos]	712	712	100%	0.0	85.51%	<a href="#">XP_006207347.1</a>
<input type="checkbox"/>	signal transducer and activator of transcription 1 isoform X1 [Sus scrofa]	711	711	100%	0.0	85.05%	<a href="#">XP_020930314.1</a>
<input type="checkbox"/>	PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Manis javanica]	711	711	100%	0.0	85.75%	<a href="#">XP_017520763.1</a>

## Alignments

[Download](#) [GenBank](#) [Graphics](#)

Ferret\_003\_D03 Ferret library from spleen, liver, and thymus Mustela putorius furo cDNA, mRNA sequence

Sequence ID: [GD181285.1](#) Length: 1280 Number of Matches: 1Range 1: 2 to 1270 [GenBank](#) [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
679 bits(1752)	0.0	Compositional matrix adjust.	362/428(85%)	379/428(88%)	7/428(1%)	+2
Query 87		KRNLDQNFQEDPIQMSMIIYSCLEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV				146
Sbjct 2		KRNLDQNFQEDP+QMSMII +CLKEER+ILENAQRFNQAQSG+IQSTVMLDKQKELDSKV				181
Query 147		RNVKDKVMCIEHEIKSLEDLQDEYDFKCKTLQNHETNGVAKSDQKQEQLLKKMYLML				206
Sbjct 182		RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNHETNGVAKNDQKQEQLLIQKMYLML				361
Query 207		DNKRKEVVHKKIIELLNVTELTQNALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIV				266
Sbjct 362		DNKRKEVV KIIIELLNVTELTQ ALINDELVEWK+RQQSACIGGPPNACLDQLQNWFTIV				541
Query 267		AESLQQVRQQLKKLEEELEQKYTYEHDPITKNKQVLWDRTFSLFQQLIQSSFVVERQPCMP				326
Sbjct 542		AESLQQVRQQLKKLEEELEQKYTYEHDPITKNKQGLWDRTFSLFQQLIQSSFVVERQPCMP				721
Query 327		THPQRPLVLKTGVQFTVKRLRLVQLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGTH				386
Sbjct 722		THPQRPLVLKTGVQFTVKRLRLVQLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGTH				901
Query 387		TKVMNMEESTNGSLAAEFRLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI				446
Sbjct 902		TKVMNMEESTNGSLAAEFRLQLKEQKNAGTRTNEGPLIVTEEL LSFET LV				1075
Query 447		DLETTSLPVVVISNVSQ--LPSGWASILWYNMLVAEPRNLSFFLTTPPCARWAQLSEVLSW				504
Sbjct 1076		SL-TSERPLGPFGLQGRPLPRGWASILVETIVGTNPEIVLLF-NPPGDEGSALGET-SW				1246
Query 505		QFSSVTKR 512				
Sbjct 1247		FSS KR				
		SFSSSHKR 1270				

[Download](#) [GenPept](#) [Graphics](#)

unnamed protein product [Gulo gulo]

Sequence ID: [VCW91131.1](#) Length: 750 Number of Matches: 1Range 1: 87 to 512 [GenPept](#) [Graphics](#)[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
731 bits(1886)	0.0	Compositional matrix adjust.	376/428(88%)	385/428(89%)	7/428(1%)
Query 1		KRNLDQNFQEDPVQMSMIIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV			60
Sbjct 87		KRNLDQNFQEDP+QMSMIIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV			146
Query 61		RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNHETNGVAKNDQKQEQLLIQKMYLML			120
Sbjct 147		RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNHETNGVAKNDQKQEQLLIQKMYLML			206
Query 121		DNKRKEVVHKKIIELLNVTELTQKALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIV			180
Sbjct 207		DNKRKEVVHKKIIELLNVTELTQKALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIV			266
Query 181		AESLQQVRQQLKKLEEELEQKYTYEHDPITKNKQGLWDRTFSLFQQLIQSSFVVERQPCMP			240
Sbjct 267		AESLQQVRQQLKKLEEELEQKYTYEHDPITKNKQGLWDRTFSLFQQLIQSSFVVERQPCMP			326
Query 241		THPQRPLVLKTGVQFTVKRLRLVQLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGTH			300
Sbjct 327		THPQRPLVLKTGVQFTVKRLRLVQLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGTH			386
Query 301		TKVMNMEESTNGSLAAEFRLQLKEQKNAGTRTNEGPLIVTEELSLLSFETTC--SSLVW			358
Sbjct 387		TKVMNMEESTNGSLAAEFRLQLKEQKNAGTRTNEGPLIVTEEL LSFET LV			446
Query 359		SL-TSERPLGPFGLQGRPLPRGWASILVETIVGTNPEIVLLF-NPPGDEGSALGET-SW			415
Sbjct 447		L T+ P+ ++ LP GWASIL ++ T P + F NPP S L E SW			504
Query 416		SFSSSHKR 423			
Sbjct 505		FSS KR			
		QFSSVTKR 512			



## [Q5] Relabel sequences for alignment

### >Ferret\_STAT1

KRNLQDNFQEDPVQMSMIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQQLLIQKMYLMLDNKRKEVVLKIIELLNVTTELTKALINDELVEWKQRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQGLWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELSLSFETTCSSSLVWSLTSERPLGPFGLDQGRPLPRGWASILVETIVGTNPEIVLLFNPPGDEGSALGETSWSFSSSHKR

### >Human\_STAT1 |NP\_009330.1 signal transducer and activator of transcription 1-alpha/beta isoform alpha [Homo sapiens]

KRNLQDNFQEDPIQMSMIISCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCKTLQNREHETNGVAKSDQKQEQQLLLKKMYLMLDNKRKEVVHKIIELLNVTTELQNALINDELVEWKRRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQVLWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHLSLFSFETQLCQPGGLVIDLETTSLPVVVISNVSQQLPSGWASILWYNMLVAEPRNLSFFLTTPCARWAQLSEVLSWQFSSVTKR

### >Sea\_otter |XP\_022354714.1:87-512| signal transducer and activator of transcription 1-alpha/beta isoform X1 [Enhydra lutris kenyoni]

KRNLQDNFQEDPIQMSMIICNCLKEERRILENAQRFNQAQSGSIQNTVMLDKQKELDSKVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQQLLIQKMYLMLDNKRKEVVLKIIELLNVTTELTKALINDELVEWKQRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQGLWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHLSLFSFETQLCQPGGLVIDLETTSLPIVVVISNVSQQLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSWQFSSVTKR

### >Cat |XP\_006935505.1:87-512 signal transducer and activator of transcription 1-alpha/beta [Felis catus]

KRNLQDNFQEDPIQMSMIISNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQMLIKMYLMLDSKRKEVVHKIIELLNITELTKALINDELVEWKQRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQGLWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHLSLFSFETQLCQPGGLVIDLETTSLPIVVVISNVSQQLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSWQFSSVTKR

### >Polar\_bear |XP\_008685758.1:146-571 PREDICTED: signal transducer and activator of transcription 1-alpha/beta [Ursus maritimus]

KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHEANGVAKNDQKQEQMLIQKMYLMLDNKRKEVVLKIIELLNVTTELTKALINDELVEWKQRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQGLWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHLSLFSFETQLCQPGGLVIDLETTSLPVVVISNVSQQLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSWQFSSVTKR

### >Finless\_porpoise |XP\_024589435.1:87-512 signal transducer and activator of transcription 1-alpha/beta [Neophocaena asiaeorientalis asiaeorientalis]

KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQAGNIQSTVMLDKQKELDSVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQMLIQKMYLMLDNKRKEVVLKIIELLNVTTELTKALINDELVEWKRRQOSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEEELEQKYTYEHDPIITKNKQALWDRFTSLFQQLIQSSFVVERQPCMPHTHPQRPLVLKTGVQFTVKLRLLVVKLQELNLYNLKVKVLFDDKDVNERNTVKGFRKFNILGHTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTNEGPLIVTEELHLSLFSFETQLCQPGGLVIDLETASLPVVVISNVSQQLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSWQFSSVTKR

**>Beluga\_whale |XP\_022422976.1:87-512 signal transducer and activator of transcription 1-alpha/beta isoform X1 [Delphinapterus leucas]**

KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQAGNIQSTVMLDKQQEELDSRVNRNVKDKVMCI  
EHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQMLLQKMYLMLDNKRKEVVHKIIELLNATEL  
TQKALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEELEQKYTYEHDPI TK  
NKQALWDRFTFSLFQQLIQSSFFVVERQPCMPHPQRPLVLKTGVQFTVKLRLLVVKLQELNYNLKVVLFDK  
DVNERNTVKGFRKFNI LGTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTNEGPLIVTEELHSLSF  
TQLCQPGGLVIDLETASLPVVVISNVSQLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSWQF  
SSVTKR

**>Horse |XP\_001499419.2:87-512 signal transducer and activator of transcription 1-alpha/beta [Equus caballus]**

KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSRVNRNVKDKVMCI  
EQEIKTLEDLQDEYDFKCKTSQNRHEANGVAKNDQKQEQMLLQKMYLMLDNKRKEVVHKIIELLNVTEL  
TQKALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKLEELEQKYTYEHDPI TK  
NKQGLWDRFTFSLFQQLIQNSFFVVERQPCMPHPQRPLVLKTGVQFTVKLRLLVVKLQELNYNLKVVLFDK  
DVNERNTVKGFRKFNI LGTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSF  
TQLCQPGGLVIDLETTSLPVVVISNVSQLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSWQF  
SSVTKR

**>African\_bush\_elephant |XP\_010600974.1 signal transducer and activator of transcription 1-alpha/beta [Loxodonta africana]**

KRNLQDNFQEDPIILMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQK  
ELDSKVRNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNREHETSGAAKNEQKEEQMLLQKMYLMLDNKR  
KEVVHKIIELLNVTELTQNALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQARQQLKKL  
EELEQKYTYEHDPI TKNKQVLWDRFTFSLFQQLIQSSFFVVERQPCMPHPQRPLVLKTGVQFTVKLRLLV  
LQELNYNLTVKVLFDKDVNERNTVKGFRKFNI LGTHKKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTN  
EGPLIVTEELHSLSFETQLCQPGGLVIDLETTSLPVVVISNVSQLPSGWASILWYNMLVAEPRNLSFFLN  
PCAQWAQLSEVLSWQFSSVTKR

**>Bovine |XP\_005202627.2 signal transducer and activator of transcription 1-alpha/beta isoform X1 [Bos taurus]**

KRNLQDNFQEDPIQMSMIICNCLKEERKILDHAQRISQAQSGNIQSTVMLDKQK  
ELDSKVRNVKDKVMSIEHEIKTLEDLQDEYDFKCKTLQNREHETNGVAKNDQKQEQLLLQKMYLMLDNKR  
KEVVLKIEILLNATELTQKALINDELVEWKRRQQSACIGGPPNACLDQLQNWFTIVAESLQQVRQQLKKL  
EELEQKYTYEHDPI TKNKQALWDRFTFSLFQQLIQSSFFVVERQPCMPHPQRPLVLKTGVQFTVKLRLLV  
LQELNYNLKVVLFDKDVNERNTVKGFRKFNI LGTHTKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTN  
EGPLIVTEELHSLSFETQLCQPGGLVIDLETTSLPVVVISNVSQLPSGWASILWYNMLVAEPRNLSFFLN  
PCARWSQLSEVLSWQFSSVTKR

**Alignment:**

Obtained using MUSCLE (version 3.8) at EBI: (see the next page). The sequence is trimmed to yield similar length.

CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```

Ferret_STAT1      KRNLQDNFQEDPVQMSMIICNCLKEERRILENAQRFNQAQSGSIQSTVMLDKQKELDSKV
African_bush_elephant KRNLQDNFQEDPIILMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV
Human_STAT1      KRNLQDNFQEDPIQMSMI IYSCLEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV
Bovine           KRNLQDNFQEDPIQMSMIICNCLKEERKILHAQRISQAQSGNIQSTVMLDKQKELDSKV
Finless_porpoise KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQAGNIQSTVMLDKQKELDSRV
Beluga_whale     KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQAGNIQSTVMLDKQKELDSRV
Horse           KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQSGSIQSTVMLDKQKELDSKV
Sea_otter       KRNLQDNFQEDPIQMSMI IYNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV
Cat             KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV
Polar_bear      KRNLQDNFQEDPIQMSMIICNCLKEERKILENAQRFNQAQSGNIQSTVMLDKQKELDSKV
*****:*****.*****.***:***.***:***.*****:*****.

```

```

Ferret_STAT1      RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQLLIQKMYLML
African_bush_elephant RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTSGAAKNEQKEEQMLLQKMYLML
Human_STAT1      RNVKDKVMCIEHEIKSLEDLQDEYDFKCKTLQNRHEHTNGVAKSDQKQEQLLKMYLML
Bovine           RNVKDKVMSIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQLLQKMYLML
Finless_porpoise RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQMLLQKMYLML
Beluga_whale     RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQMLLQKMYLML
Horse           RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTSQNRHEANGVAKNDQKQEQMLLQKMYLML
Sea_otter       RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQLLIQKMYLML
Cat             RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEHTNGVAKNDQKQEQMLLQKMYLML
Polar_bear      RNVKDKVMCIEHEIKTLEDLQDEYDFKCKTLQNRHEANGVAKNDQKQEQMLLQKMYLML
*****.***:***:*****.*****.***:***:***:***:***:*****

```

```

Ferret_STAT1      DNKRKEVVLKIIELLNVTELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
African_bush_elephant DNKRKEVVHKIIELLNVTELTQNALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Human_STAT1      DNKRKEVVHKIIELLNVTELTQNALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Bovine           DNKRKEVVLKIIELLNATELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Finless_porpoise DNKRKEVVHKIIELLNVTELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Beluga_whale     DNKRKEVVHKIIELLNATELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Horse           DNKRKEVVHKIIELLNVTELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Sea_otter       DNKRKEVVLKIIELLNVTELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Cat             DSKRKEVVHKIIELLNITELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
Polar_bear      DNKRKEVVHKIIELLNVTELTQKALINDELVEWKRRQSSACIGGPPNACLDQLQNWFTIV
*.*****.*****.*****:*****.*****.*****.*****

```

```

Ferret_STAT1      AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQGLWDRFTSLFQQLIQSSSFVVERQPCMP
African_bush_elephant AESLQQARQQLKKLEELEQKYTYEHDPITKNKQVLWDRFTTLFQQLIQSSSFVVERQPCMP
Human_STAT1      AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQVLWDRFTSLFQQLIQSSSFVVERQPCMP
Bovine           AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQALWDRFTSLFQQLIQSSSFVVERQPCMP
Finless_porpoise AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQALWDRFTSLFQQLIQSSSFVVERQPCMP
Beluga_whale     AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQALWDRFTSLFQQLIQSSSFVVERQPCMP
Horse           AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQGLWDRFTSLFQQLIQSSSFVVERQPCMP
Sea_otter       AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQGLWDRFTSLFQQLIQSSSFVVERQPCMP
Cat             AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQGLWDRFTSLFQQLIQSSSFVVERQPCMP
Polar_bear      AESLQQVRQQLKKLEELEQKYTYEHDPITKNKQGLWDRFTSLFQQLIQSSSFVVERQPCMP
*****.*****.*****.*****.*****.*****.*****

```

```

Ferret_STAT1      THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
African_bush_elephant THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLTVKVLFDKDVNERNTVKGFRKFNILGTH
Human_STAT1      THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Bovine           THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Finless_porpoise THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Beluga_whale     THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Horse           THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Sea_otter       THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Cat             THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
Polar_bear      THPQRPLVLKTVGVQFTVKLRLLVLKQELNYNLKVKVLFDKDVNERNTVKGFRKFNILGTH
*****.*****.*****.*****.*****.*****.*****

```

```

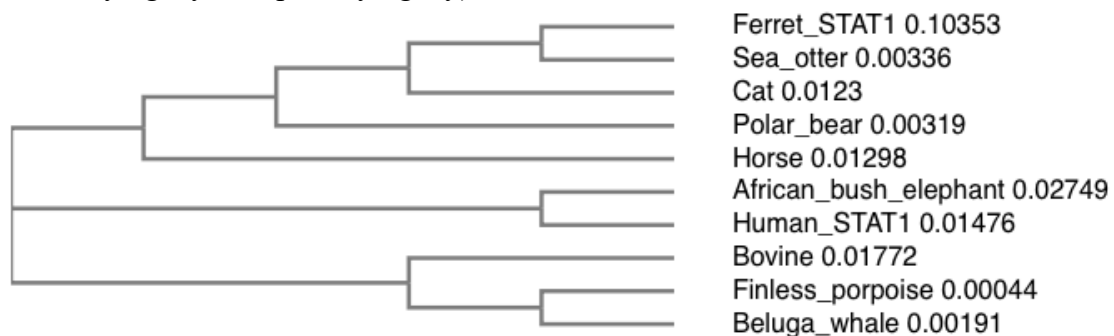
Ferret_STAT1      TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETT-C-SSLVW
African_bush_elephant KVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
Human_STAT1      TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
Bovine           TKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTNEGPLIVTEELHSLSFETQLCQPGGLVI
Finless_porpoise TKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTNEGPLIVTEELHSLSFETQLCQPGGLVI
Beluga_whale     TKVMNMEESTNGSLAAEFRHLQLKEQKNAGARTNEGPLIVTEELHSLSFETQLCQPGGLVI
Horse           TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
Sea_otter       TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
Cat             TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
Polar_bear      TKVMNMEESTNGSLAAEFRHLQLKEQKNAGTRTNEGPLIVTEELHSLSFETQLCQPGGLVI
*****.*****.*****.*****.*****.*****.*****

```

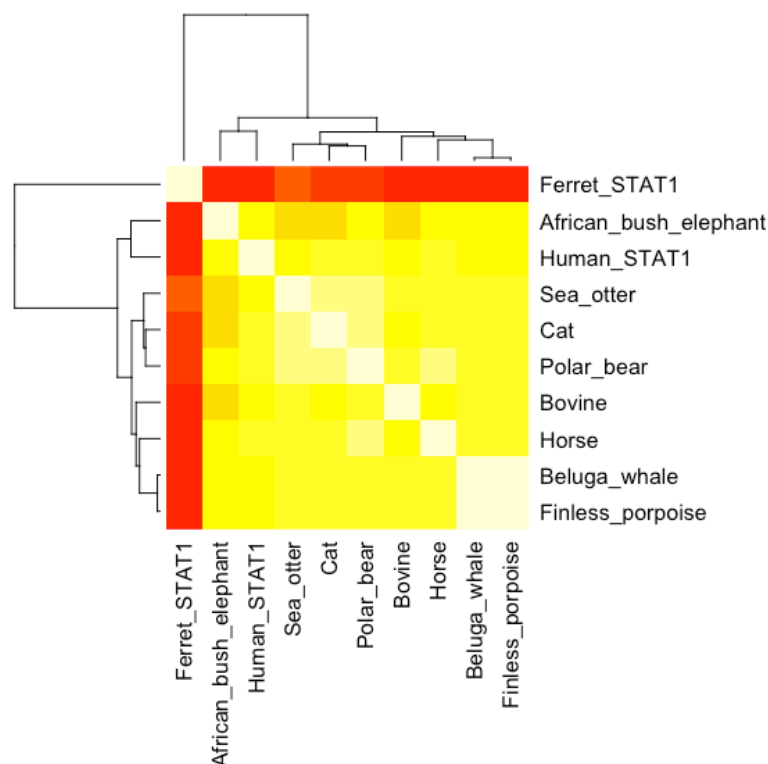


Ferret_STAT1	SL-TSERPLGPFGLDQGRPLPRGWASILVETIVGTNP-EIVLLFNPPGDEGSALGET-SW
African_bush_elephant	DLETTSLPVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLNPPCAQWAQLSEVLSW
Human_STAT1	DLETTSLPVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLTTPPCARWAQLSEVLSW
Bovine	DLETTSLPVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSW
Finless_porpoise	DLETASLPVVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSW
Beluga_whale	DLETASLPVVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSW
Horse	DLETTSLPVVISNV--QLPSGWASILWYNMLVAEPRNLSFFLNPPCARWSQLSEVLSW
Sea_otter	DLETTSLPIVVISNV--QLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSW
Cat	DLETTSLPIVVISNV--QLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSW
Polar_bear	DLETTSLPVVISNV--QLPSGWASILWYNMLVTEPRNLSFFLNPPCARWSQLSEVLSW
	. * * : . * : : : : . ** * * * * . : : : * : : : . * * : * . * *
Ferret_STAT1	SFSSSHKR
African_bush_elephant	QFSSVTKR
Human_STAT1	QFSSVTKR
Bovine	QFSSVTKR
Finless_porpoise	QFSSVTKR
Beluga_whale	QFSSVTKR
Horse	QFSSVTKR
Sea_otter	QFSSVTKR
Cat	QFSSVTKR
Polar_bear	QFSSVTKR
	. * * * * *

**[Q6]** Phylogeny Tree: Use MUSCLE “simple phylogeny” from EBI  
(Tool>Phylogeny>Simple Phylogeny)



**[Q7]** Heatmap

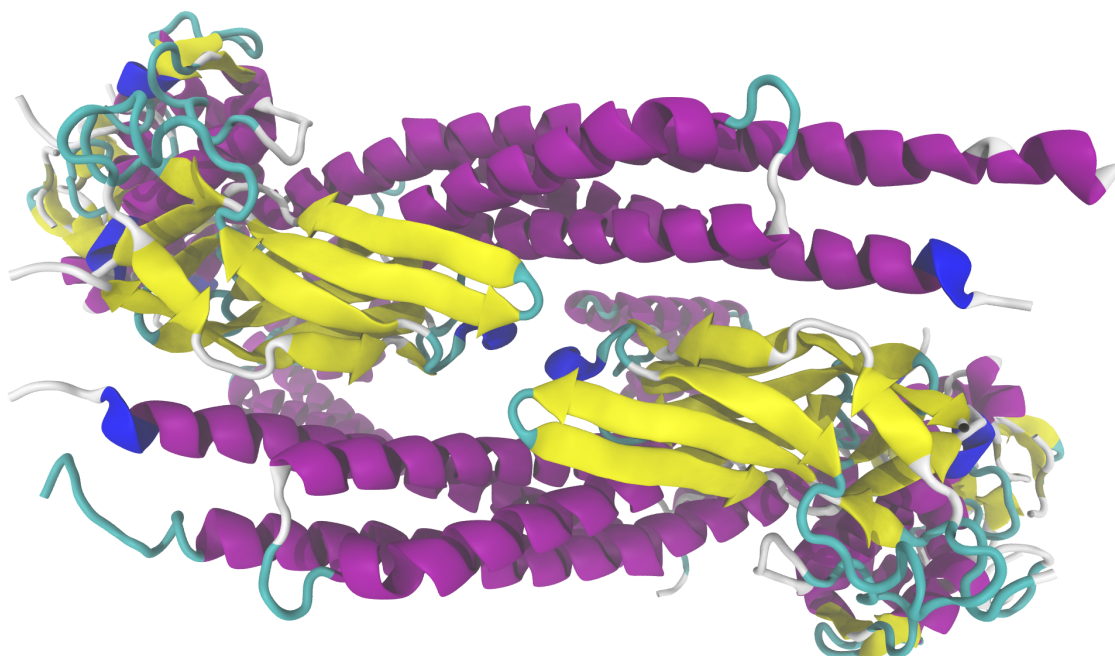


**[Q8]** BLAST the consensus sequence from the alignment using blast.pdb function.

ID	Technique	Resolution	Source	E-value	Identity
1YVL	X-Ray Diffraction	3.0	Homo sapiens	0.00e+00	96.714
1BF5	X-Ray Diffraction	2.9	Homo sapiens	0.00e+00	96.817
3CWG	X-Ray Diffraction	3.05	Mus musculus	2.66e-141	53.827

**[Q9]**

Ferret\_STAT1 is 88% similar to human STAT1 and the consensus sequence is 96.7% similar to the molecule shown below. The molecule is a homodimer of STAT1 which can bind to GAS sequence on the DNA. STAT1 homodimer is formed in response to interferon gamma activation.



**[Q10]: ChEMBL Search**

ChEMBL identifies 4 different targets: CHEMBL6101, CHEMBL1667670, CHEMBL5045, CHEMBL3885638. The first three target is single protein from three different species, while the last target is a protein complex.

CHEMBL6101 shows that there are 754 associated bioactivities to STAT1 protein and 62 associated assays. There are 688 associated compounds. There are 28 ligand efficiency data for targeting this protein. Among 11 inhibitor compounds, none of them has made it to the Phase I clinical trial.

From the binding assays document, a publication shows that thioaryl naphthylmethanone oxime ether can be used an anticancer agent by inhibiting EFG-induced STAT1. DOI: [10.1021/jm500873e](https://doi.org/10.1021/jm500873e)