

## Biological Databases

Conducted by Lim Yun Ping

E mail : [yunping@nus.edu.sg](mailto:yunping@nus.edu.sg)

Date : 25 May 2009

Hands on practice

- Compare the difference between using ENTREZ and SRS to search different types of databases and multiple databases at once
- Finding the right sequence more reliably and viewing results in different ways
- Looking at Online Bioinformatics Resources Collection (OBRC)  
<http://www.hsls.pitt.edu/guides/genetics/obrc/>

Please access the following URL for the resources:

**NCBI :** <http://www.ncbi.nlm.nih.gov/Entrez/>

**SRS :** <http://srs.ebi.ac.uk/>

### A. Biology in Databases

#### 1. Introduction to ENTREZ and SRS

a) Use Entrez to search across the databases in NCBI for **WNT4**.

View the number of entries for each category. Click on the Nucleotide Sequence database hyperlink. How many entries are found?

The screenshot shows the NCBI Entrez search engine interface. At the top, there is a navigation bar with links for HOME, SEARCH, SITE MAP, PubMed, Entrez, Human Genome, GenBank, Map Viewer, and BLAST. Below this is a search bar with a red box around it, containing the text "Search across databases". To the right of the search bar are buttons for GO, CLEAR, and Help. Below the search bar, there is a heading "Welcome to the new Entrez cross-database search page". Underneath, there is a grid of database categories, each with an icon and a description. A red box highlights the "Nucleotide" category, which is described as "sequence database (GenBank)". Other categories include PubMed, PubMed Central, Books, OMIM, Site Search, UniGene, CDD, 3D Domains, UniSTS, PopSet, GEO, SNP, and HomoloGene.

Look at the results carefully, you will notice there are sequences from mRNA, genomic and from various organisms. These may be many false positives (hits found which are not related) and there could be false negatives (related information which were not detected) based on a simple, general search.

The screenshot shows the NCBI Entrez search engine interface. The search term "wnt4" is entered in the search box, and the "GO" button is highlighted. The search results are displayed in a grid format, showing the number of hits for each database. The results are as follows:

Database	Count	Description
PubMed	109	biomedical literature citations and abstracts
PubMed Central	258	free, full text journal articles
Site Search	2	NCBI web and FTP sites
Books	20	online books
OMIM	14	online Mendelian Inheritance in Man
OMIA	none	online Mendelian Inheritance in Animals
Nucleotide	167	Core subset of nucleotide sequence records
EST	58	Expressed Sequence Tag records
GSS	3	Genome Survey Sequence records
Protein	137	sequence database
Genome	12	whole genome sequences
Structure	none	three-dimensional macromolecular structures
Taxonomy	none	organisms in GenBank
SNP	454	single nucleotide polymorphism
Gene	58	gene-centered information
HomoloGene	3	eukaryotic homology groups
GENSAT	none	gene expression atlas of mouse central nervous system
Probe	53	sequence-specific reagents
Genome Project	none	genome project information
dbGaP	none	genotype and phenotype
UniGene	10	gene-oriented clusters of transcript sequences
CDD	none	conserved protein domain database
3D Domains	none	domains from Entrez Structure
UniSTS	16	markers and mapping data
PopSet	1	population study data sets
GEO Profiles	1868	expression and molecular abundance profiles
GEO DataSets	5	experimental sets of GEO data
Cancer Chromosomes	none	cytogenetic databases
PubChem BioAssay	none	bioactivity screens of chemical substances
PubChem Compound	none	unique small molecule chemical structures
PubChem Substance	none	deposited chemical substance records
Protein Clusters	none	a collection of related protein sequences

b) To make your search more specific and sensitive, use the "Limits" option. To find only full length mRNA sequences only, limit the search to **exclude ESTs, select molecule : mRNA, location : Genomic RNA/DNA and RefSeq.**

c) View **History**. Did you reduce your results hits considerably?

NCBI Nucleotide search results for 'wnt4'. The search shows 110 results. A red box highlights the 'Top Organisms' list on the right, with a red arrow pointing to 'Homo sapiens (2)'. Below the search results, there are five entries listed with their accession numbers and descriptions. A 'Recent Activity' section is also visible at the bottom right.

d) View the records under “Homo Sapiens”. How many entries are there ?  
 Select the entry with accession number **NM\_030761** and select the display in FASTA format.  
 Copy this sequence to a text file (using Notepad).

d) Click on the link to Homologues of WNT4. In which organisms is the WNT4 gene conserved ?  
 Retrieve the sequences in FASTA format.

>Human

```

MSPRSLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLEVMDSVRRG
AQLAIEECQYQFRNRRWNCSTLDSL PVFGKVVTQGTREAAFVYAISSAGVAFVTRACSSGELEKCGCDR
TVHGVSPQGFQWSGSDNIAYGVAFSQSFVDVRERSKGASSSRALMNLHNNEAGRKA ILTHMRVECKCHG
VSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFKPHTDEDLVYLEPSPDFC
EQDMRSGVLGTRGRTCNKTSKAIDGCELLCCGRGFHTAQVELAERCSCCKFWHCCFVKCRQCQRLVELHTC
R
  
```

>Chimpanzee

```

MSPRSLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLEVMDSVRRG
AQLAIEECQYQFRNRRWNCSTLDSL PVFGKVVTQGTREAAFVYAISSAGVAFVTRACSSGELEKCGCDR
TVHGVSPQGFQWSGSDNIAYGVAFSQSFVDVRERSKGASSSRALMNLHNNEAGRKA ILTHMRVECKCHG
VSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRRVGSSRALVPRTAQFKPHTDET VYWSYPTL
  
```

>Dog

```

MAGTTTLISISAPYYGGDRGVPGLRGTAGTPSPAQSSLTTPATAAPAVPFQPRGGDRPSIRRRHLPNAFC
VPCAVPGTEEATLDGPKALSAPA AVKQGETEFPVTHWLRDLGYLAKLSSVGSISEEETCEKLGKLIQRQ
VQMCKRNLEVMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSL PVFGKVVTQGTREAAFVYAISSAGVAFV
VTRACSSGELEKCGCDRTVHGVSPQGFQWSGSDNIAYGVAFSQSFVDVRERSKGASSSRALMNLHNNEA
GRKA ILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFK
PHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLCCGRGFHTAQVELAERCSCCKFWHC
CFVKCRQCQRLVELHTCR
  
```

>Mouse

```

MSPRSLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLEVMDSVRRG
  
```

AQLAIEECQYQFRNRRWNCSTLDSLVPVFGKVVTQGTREAAAFVYAISSAGVAFVAVTRACSSGELEKCGCDR  
TVHGVSPQGFQWSGCSDNIAYGVAFSQQSFVDVRRERSKGASSSRALMNLHNNEAGRKAILTHMRVECKCHG  
VSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFKPHTDEDLVYLEPSPDFC  
EQDIRSGVLGTRGRTCNKTSKAIDGCELLCCGRGFHTAQVELAERCGRFHWCCFVKCRQCQRLVEMHTC  
R

>Rat

MSPRSLRSLRLLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLEVMDSVRHG  
AQLAIEECQYQFRNRRWNCSTLDSLVPVFGKVVTQGTREAAAFVYAISSAGVAFVAVTRACSSGDLEKCGCDR  
TVHGVSPQGFQWSGCSDNIAYGVAFSQQSFVDVRRERSKGASSSRALMNLHNNEAGRKAILTHMRVECKCHG  
VSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFKPHTDEDLVYLEPSPDFC  
EQDMRSGVLGTRGRTCNKTSKAIDGCELLCCGRGFHTAHVELAERCGRFHWCCFVKCRQCQRLVEMHTC  
R

>Chicken

MSPEYFLRSLLLIILATFSANASNWLYLAKLSSVGSISEEETCEKLGKLIQRQVQMCKRNLEVMDSVRRG  
AQLAIEECQYQFRNRRWNCSTLDTLPVFGKVVTQGTREAAAFVYAISSAGVAFVAVTRACSSGELDKCGCDR  
TVQGGSPQGFQWSGCSDNIAYGVAFSQQSFVDVRRERSKGASSNRALMNLHNNEAGRKAILNMRVECKCHG  
VSGSCEFVKTCWKAMPFRKVGNVLKEKFDGATEVEQSEIGSTKVLVPKNSQFKPHTDEDLVYLDSSPDFC  
DHDLKNGVLGTSGRQCNKTSKAIDGCELMCCGRGFHTDEVEVVERCSCFKHWCCSVKCKPCHRVEIHTC  
R

>Zebrafish

MSSEYLIRSLMLFLALFSANASNWLYLAKLSSVGSISDEETCEKLRGLIQRQVQICKRNVEVMDAVRRG  
AQLAIDECQYQFRNRRWNCSTLESVPVFGKVVTQGTREAAAFVYAISSAASVAFVAVTRACSSGELDKCGCDR  
NVHGVSPQGFQWSGCSDNIAYGVAFSQQSFVDIRERSKGQSSNRALMNLHNNEAGRKAILNMRVECKCHG  
VSGSCEVKTCWKAMPFRKVGNIKEKFDGATEVELRKVGTTKVLVPRNSQFKPHTDEDLVYLDSPDFC  
EHDPRTPGIMGTAGRFNCNKTSKAIDGCELMCCGRGFHTEEVEVVDRCSCFKHWCCYVKCKQCRKMVEMHT  
CR

>Worm

MLKSTQVILIFILLISIVESLSWLALGLAANRFDRDKPGTSCSKLGLTRRQMRFCCKNIDLMEVSRSGS  
LAAHAECQFQFHKRRWNCTLIDPVTHEVIPDVFLYENTRESAFVHAISSAAVAYKVTRDCARGISERCGC  
DYSKNDHSGKSQFYQGCSDNVKFGIGVSKEFVDSAQRRVLMKDDNGTSLLGPSQLSADGMHMINLHNN  
QAGRQVLEKSLRRECKCHGMSGSEMRTCWDSLPNFRHIGMAIKDKFDGAAEVKVKVEDGIEKPRIVMKN  
SQFKRHTNADLVYMTSPDFCESDPLRGILGTKGRQCTLAPNAIDDCSLLCCGRGYEKKVQIVEEKCCK  
FIYCEVRCEPCQKRIEKYLCL

**Homo sapiens wingless-type MMTV integration site family, member 4 (WNT4), mRNA**

[Comment](#) [Features](#) [Sequence](#)

LOCUS	NM_030761	3905 bp	mRNA	linear	PRI 29-MAR-2009
DEFINITION	Homo sapiens wingless-type MMTV integration site family, member 4 (WNT4), mRNA.				
ACCESSION	NM_030761				
VERSION	NM_030761.4 GI:156630997				
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	<a href="#">Homo sapiens</a>				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 3905)				
AUTHORS	Kuulasmaa,T., Jaaskelainen,J., Suppola,S., Pietilainen,T., Heikkilä,P., Aaltomaa,S., Kosma,V.M. and Voutilainen,R.				
TITLE	WNT-4 mRNA expression in human adrenocortical tumors and cultured adrenal cells				
JOURNAL	Horm. Metab. Res. 40 (10), 668-673 (2008)				
PUBMED	<a href="#">18553255</a>				
REMARK	GeneRIF: The abundant WNT-4 mRNA expression in Conn's adenomas and its hormonal regulation in adrenocortical cells suggest a role for WNT-4 in human adrenocortical function.				
REFERENCE	2 (bases 1 to 3905)				
AUTHORS	Yoshida,T., Kitaura,H., Hagio,Y., Sato,T., Iguchi-Ariga,S.H. and Ariga,H.				
TITLE	Negative regulation of the Wnt signal by MM-1 through inhibiting expression of the wnt4 gene				
JOURNAL	Exp. Cell Res. 314 (6), 1217-1228 (2008)				
PUBMED	<a href="#">18281035</a>				
REMARK	GeneRIF: Wnt4 gene, a member of the Wnt-beta-catenin pathway, was identified as a target gene of MM-1.				
REFERENCE	3 (bases 1 to 3905)				
AUTHORS	Phillibert,P., Blason-Lauber,A., Rouzier,R., Pienkowski,C., Paris,F., Konrad,D., Schoenle,E. and Sultan,C.				
TITLE	Identification and functional analysis of a new WNT4 gene mutation among 28 adolescent girls with primary amenorrhea and müllerian duct abnormalities: a French collaborative study				
JOURNAL	J. Clin. Endocrinol. Metab. 93 (3), 895-900 (2008)				
PUBMED	<a href="#">18182450</a>				
REMARK	GeneRIF: WNT4 is involved in the regulation of müllerian duct development and ovarian androgen biosynthesis. WNT4 may also contribute to human follicle development and/or maintenance.				
REFERENCE	4 (bases 1 to 3905)				
AUTHORS	Hande,I.H., Shemer,R., Borochowitz,Z.U., Okopnik,M., Knopf,C., Indelman,H., Drugan,A., Tiosano,D., Gershoni-Baruch,R., Choder,M.				

Change Region Shown  
Customize View

**Sequence Analysis Tools**

[BLAST Sequence](#)  
Find regions of similarity between this sequence and other sequences using BLAST.

[Pick Primers](#)  
Design and test primers for this sequence using Primer-BLAST.

- Articles about WNT4**
- Mutational analysis of the WNT gene family in women with Mayer-Rokitansky-Kuster-Haus [Fertil Steril. 2009]
  - Expression of Wnt4 in human pituitary adenomas regulates activation of the beta-c[Endocr Pathol. 2008]
  - WNT-4 mRNA expression in human adrenocortical tumors and cultured adrenal ce [Horm Metab Res. 2008]
- > See all...

**Reference sequences**

- Protein

**More about the WNT4 gene**

The WNT gene family consists of structurally related genes which encode secreted signaling proteins. These proteins have been implicated in ...

Also Known As: RPI-224A6 7, SERKAL, WN...

**Homologs of WNT4**

The WNT4 gene is conserved in chimpanzee, dog, mouse, rat, chicken, zebrafish, and C.elegans.

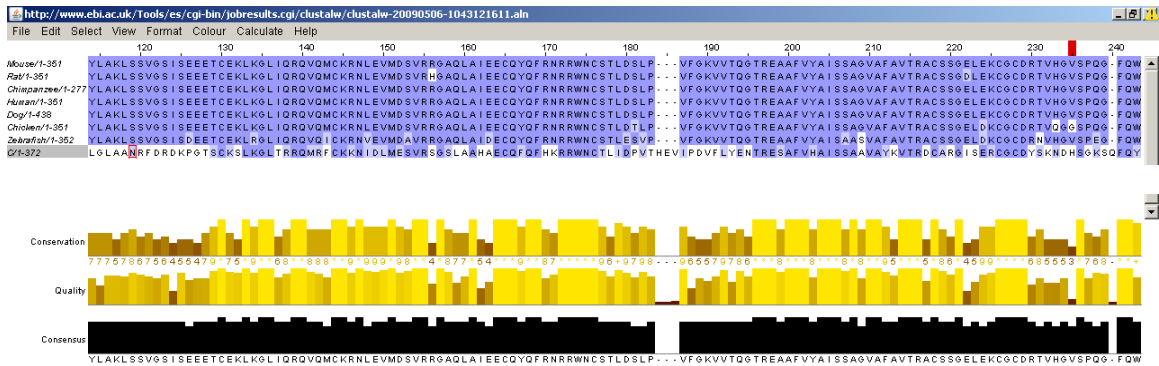
**Order cDNA Clone**

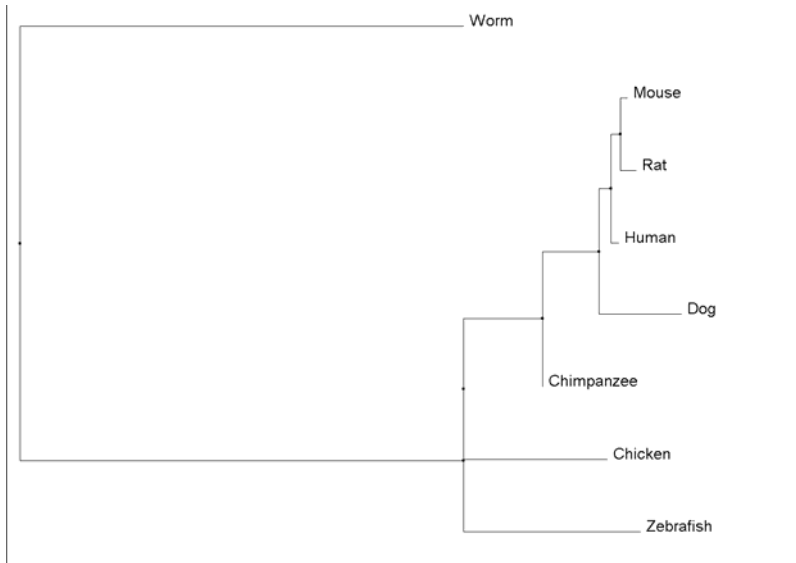
The NIH MGC Collection contains a sequence-verified cDNA clone for WNT4.

**Recent Activity**

Turn Off Clear

e) Use ClustalW: <http://www.ebi.ac.uk/Tools/clustalw/> to generate the multiple sequence alignment and phylogram.  
More information on Jalview is available here: <http://www.jalview.org/documentList.html>





f) Obtain more information from the Link to “Gene” for more information on WNT4:  
[http://www.ncbi.nlm.nih.gov/nuccore/28302128?ordinalpos=1&itool=EntrezSystem2.PEntrez.Sequence.Sequence\\_ResultsPanel.Sequence\\_RVDocSum](http://www.ncbi.nlm.nih.gov/nuccore/28302128?ordinalpos=1&itool=EntrezSystem2.PEntrez.Sequence.Sequence_ResultsPanel.Sequence_RVDocSum)

What additional information do you get? What is the location of this gene ? What are the neighbouring genes upstream and downstream ? What are the other known genes which interact with this gene ?

REMARK GeneRIF: Together with previous observations in animal models, the present data attribute a pivotal role to WNT4 signaling during organogenesis in humans.

REFERENCE 5 (bases 1 to 3905)  
 AUTHORS Miyakoshi,T., Takei,M., Kajiya,H., Egashira,N., Takekoshi,S., Teramoto,A. and Osamura,R.Y.  
 TITLE Expression of Wnt4 in human pituitary adenomas regulates activation of the beta-catenin-independent pathway  
 JOURNAL Endocr. Pathol. 19 (4), 261-273 (2008)  
 PUBMED [19034702](#)

REMARK GeneRIF: These results suggested that activation of Wnt4/Fzd6 signaling through a 'beta-catenin-independent' pathway played a role in proliferation and survival of the pituitary adenoma cells.

REFERENCE 6 (bases 1 to 3905)  
 AUTHORS Tanaka,K., Okabayashi,K., Asashima,M., Perrimon,N. and Kadovaki,T.  
 TITLE The evolutionarily conserved porcupine gene family is involved in the processing of the Wnt family  
 JOURNAL Eur. J. Biochem. 267 (13), 4300-4311 (2000)  
 PUBMED [10866835](#)

REFERENCE 7 (bases 1 to 3905)  
 AUTHORS Vainio,S., Heikkila,M., Kispert,A., Chin,N. and McMahon,A.P.  
 TITLE Female development in mammals is regulated by Wnt-4 signalling  
 JOURNAL Nature 397 (6718), 405-409 (1999)  
 PUBMED [9989404](#)

REFERENCE 8 (bases 1 to 3905)  
 AUTHORS Lescher,B., Haenig,B. and Kispert,A.  
 TITLE sFRP-2 is a target of the Wnt-4 signaling pathway in the developing metanephric kidney  
 JOURNAL Dev. Dyn. 213 (4), 440-451 (1998)  
 PUBMED [9853965](#)

REFERENCE 9 (bases 1 to 3905)  
 AUTHORS Huguet,E.L., McMahon,J.A., McMahon,A.P., Bicknell,R. and Harris,A.L.  
 TITLE Differential expression of human Wnt genes 2, 3, 4, and 7B in human breast cell lines and normal and disease states of human breast tissue  
 JOURNAL Cancer Res. 54 (10), 2615-2621 (1994)  
 PUBMED [8168088](#)

REFERENCE 10 (bases 1 to 3905)  
 AUTHORS Smolich,B.D., McMahon,J.A., McMahon,A.P. and Papkoff,J.  
 TITLE Wnt family proteins are secreted and associated with the cell surface  
 JOURNAL Mol. Biol. Cell 4 (12), 1267-1275 (1993)  
 PUBMED [8167409](#)

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence was derived from [BC057781.1](#), [AF316543.1](#) and [U031281.6](#).  
 On Aug 31, 2007 this sequence version replaced gi:[45439371](#).

Summary: The WNT gene family consists of structurally related genes which encode secreted signaling proteins. These proteins have been implicated in oncogenesis and in several developmental processes,

Q [wnt4 \(167\)](#) Nucleotide  
 Q [wnt4 \(58\)](#) Gene  
 Q [\(wnt4\) AND "Homo sapiens" \(9\)](#) Nucleotide

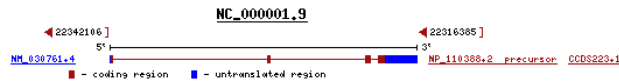
All links from this record

- ▶ Components(Core)
- ▶ Full text in PMC
- ▶ Gene
- ▶ GeneView in dbSNP
- ▶ Genome
- ▶ Genome Project
- ▶ Genome Project
- ▶ HomoloGene
- ▶ Master
- ▶ Probe
- ▶ Protein
- ▶ PubMed
- ▶ PubMed (RefSeq)
- ▶ PubMed (Weighted)
- ▶ Taxonomy
- ▶ Related Sequences
- ▶ Map Viewer
- ▶ OMIM
- ▶ GEO Profiles
- ▶ SNP
- ▶ UniGene
- ▶ UniSTS
- ▶ LinkOut

### Genomic regions, transcripts, and products

(minus strand) Go to [reference sequence details](#)

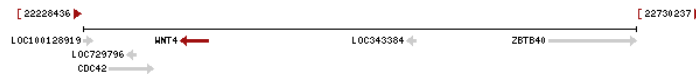
[Try our new Sequence Viewer](#)



### Genomic context

chromosome: 1; Location: 1p36.23-p35.1

[See WNT4 in MapViewer](#)



f) Using the Gene ontology's evidence codes, could you infer some of the possible function and processes of this gene

### Evidence Codes

IC: Inferred by Curator  
IDA: Inferred from Direct Assay  
IEA: Inferred from Electronic Annotation  
IEP: Inferred from Expression Pattern  
IGI: Inferred from Genetic Interaction  
IMP: Inferred from Mutant Phenotype  
IPI: Inferred from Physical Interaction  
ISS: Inferred from Sequence or Structural Similarity  
NAS: Non-traceable Author Statement  
ND: No biological Data available  
RCA: Inferred from Reviewed Computational Analysis  
TAS: Traceable Author Statement  
NR: Not Recorded

### Comments

#### Pathways

KEGG pathway: Basal cell carcinoma  
[05217](#)  
KEGG pathway: Hedgehog signaling pathway  
[04340](#)  
KEGG pathway: Melanogenesis  
[04916](#)  
KEGG pathway: Wnt signaling pathway  
[04310](#)

#### Homology

Mouse, Rat  
[Map Viewer](#)

#### GeneOntology

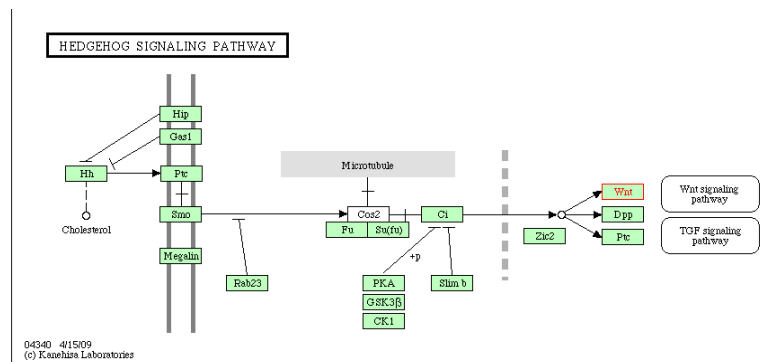
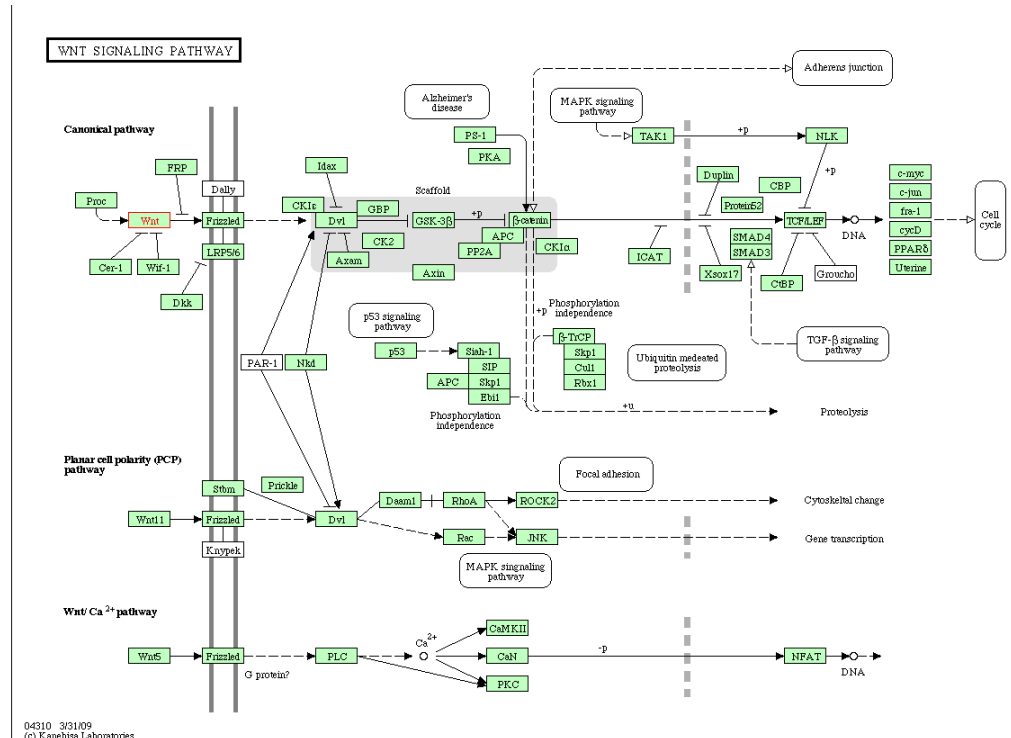
Provided by [GOA](#)

Function	Evidence
<a href="#">extracellular matrix structural constituent</a>	NAS
<a href="#">signal transducer activity</a>	IEA

Process	Evidence
<a href="#">Wnt receptor signaling pathway, calcium modulating pathway</a>	IEA
<a href="#">cell-cell signaling</a>	NAS
<a href="#">multicellular organismal development</a>	NAS

Component	Evidence
<a href="#">extracellular region</a>	NAS
<a href="#">proteinaceous extracellular matrix</a>	IEA

g) Browse through the links to the KEGG databases to find out more about the role played by WNT4 in the WNT signaling and Hedgehog signaling pathways.



Click on the  
 - rectangles representing the genes involved to find out more about the annotations provided.  
 - rounded rectangles to access the links to other pathways affected by the genes.

i) View additional information in UniProt

Protein Accession	Links
P56705.4	<a href="#">GenPept</a> <a href="#">UniProtKB/Swiss-Prot</a>
Q5JYX2	<a href="#">GenPept</a> <a href="#">UniProtKB/TrEMBL</a>
Q8IUM6	<a href="#">GenPept</a> <a href="#">UniProtKB/TrEMBL</a>

Click to view the curated annotations in Uniprot

Additional Links

- MIM [603490](#)
- GeneTests for MIM: [603490](#)
- HPRD [04601](#)
- UCSC [UCSC](#)
- UniGene [Hs.25766](#)

Click to view the additional information in the UCSC Genome browser which displays information in a different style, and contains link to other databases not linked to Entrez gene e.g. Biocarta pathways: <http://www.biocarta.com>

Names and origin		Hide   Top
Protein names	<i>Recommended name:</i> <b>Protein Wnt-4</b>	
Gene names	Name: <b>WNT4</b> ORF Names: UNQ426/PRO864	
Organism	<b>Homo sapiens (Human)</b>	
Taxonomic identifier	9606 [NCBI]	
Taxonomic lineage	Eukaryota › Metazoa › Chordata › Craniata › Vertebrata › Euteleostomi › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorrhini › Catarrhini › Hominidae › Homo	

General annotation (Comments)		Hide   Top
Function	Ligand for members of the frizzled family of seven transmembrane receptors. Probable developmental protein. May be a signaling molecule which affects the development of discrete regions of tissues. Is likely to signal over only few cell diameters ( <a href="#">By similarity</a> ). Overexpression may be associated with abnormal proliferation in human breast tissue.	
Subunit structure	Interacts with PORCN ( <a href="#">By similarity</a> ).	
Subcellular location	<a href="#">Secreted</a> › <a href="#">extracellular space</a> › <a href="#">extracellular matrix</a> .	
Involvement in disease	Defects in WNT4 are a cause of Rokitansky-Kuster-Hauser syndrome (RKH syndrome) [MIM:277000]; also called Mayer-Rokitansky-Kuster-Hauser syndrome (MRKH syndrome or MRKH anomaly). RKH syndrome is characterized by utero-vaginal atresia in otherwise phenotypically normal female with a normal 46,XX karyotype. Anomalies of the genital tract range from upper vaginal atresia to total Muellierian agenesis with urinary tract abnormalities. It has an incidence of approximately 1 in 5'000 newborn girls.  Defects in WNT4 are the cause of female sex reversal with dysgenesis of kidneys, adrenals, and lungs (SERKAL) [MIM:611812]; also known as SERKAL syndrome. ( <a href="#">Ref.12</a> )	
Sequence similarities	Belongs to the <a href="#">Wnt family</a> .	

Ontologies		Hide   Top
<b>Keywords</b>		
Biological process	<a href="#">Wnt signaling pathway</a>	
Cellular component	<a href="#">Extracellular matrix</a> <a href="#">Secreted</a>	
Coding sequence diversity	<a href="#">Polymorphism</a>	
Disease	<a href="#">Disease mutation</a>	
Domain	<a href="#">Signal</a>	
Molecular function	<a href="#">Developmental protein</a>	
PTM	<a href="#">Glycoprotein</a>	
<b>Gene Ontology (GO)</b>		
Biological process	<a href="#">Wnt receptor signaling pathway</a> , <a href="#">calcium modulating pathway</a> Inferred from electronic annotation. Source: InterPro <a href="#">cell-cell signaling</a> Non-traceable author statement. Source: UniProtKB <a href="#">multicellular organismal development</a> Non-traceable author statement. Source: UniProtKB	
Cellular component	<a href="#">proteinaceous extracellular matrix</a> Inferred from electronic annotation. Source: UniProtKB-KW	
Molecular function	<a href="#">extracellular matrix structural constituent</a> Non-traceable author statement. Source: UniProtKB <a href="#">signal transducer activity</a>	

Sequence annotation (Features) Hide | Top

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
<b>Molecule processing</b>					
<input type="checkbox"/> Signal peptide	1 – 22	22	<span style="border: 1px solid orange; border-radius: 50%; padding: 2px;">Potential</span>		
<input type="checkbox"/> Chain	23 – 351	329	Protein Wnt-4		PRO_0000041421
<b>Amino acid modifications</b>					
<input type="checkbox"/> Glycosylation	88	1	N-linked (GlcNAc...) <span style="border: 1px solid orange; border-radius: 50%; padding: 2px;">Potential</span>		
<input type="checkbox"/> Glycosylation	297	1	N-linked (GlcNAc...) <span style="border: 1px solid orange; border-radius: 50%; padding: 2px;">Potential</span>		
<b>Natural variations</b>					
<input type="checkbox"/> Natural variant	12	1	L → P in RKH syndrome; unable to suppress steroidogenesis in an ovarian adenocarcinoma cell line resulting in increased androgen production.		VAR_043497
<input type="checkbox"/> Natural variant	83	1	R → C in RKH syndrome; with androgen excess, normal kidney size and location; unable to suppress expression of steroidogenic enzymes in ovarian; impairs protein		VAR_043498

**Experimental info**

<input type="checkbox"/> Sequence conflict	106	1	T → I in AAG38658. <span style="border: 1px solid orange; border-radius: 50%; padding: 2px;">Ref.1</span>		
<input type="checkbox"/> Sequence conflict	111	1	F → L in AAG38658. <span style="border: 1px solid orange; border-radius: 50%; padding: 2px;">Ref.1</span>		

Sequences Hide | Top

Sequence	Length	Mass (Da)	Tools
<input type="checkbox"/> P56705-1 [UniParc]. Last modified May 10, 2002. Version 4. Checksum: 465D08755C992DA8	FASTA	351	39,052
<pre> 10      20      30      40      50      60 MSPRCLRSL RLLVAVFSA AASNWLYLAK LSSVGSISEE ETCERLKLGI QRVQVMCKRN 70      80      90     100     110     120 LEVMSVRRG AQLAIEECQY QFMRRWNC S TLDLPLVFGK VVTQGTREAA FVYAISSAGV 130     140     150     160     170     180 AFAVTRACSS GELEKCCDR TVHGVSPPGF QWGCSDNIA YGVAFSQSFV DVRESRGAS 190     200     210     220     230     240 SSRALMNLHN NEAGRKAILT HMRVECRCHG VSGSCEVKTC WRAVPPFRQV GHALKERFDG 250     260     270     280     290     300 ATEVEPRRVG SSRALVPRNA QFKFHTDEDL VYLEPSPDFC EQDHRSGVLG TRGRTCNKTS 310     320     330     340     350 KAIDGCELLC CGRGFHTAQV ELAERCSCKF HWCCFVKCRQ CQRLVELHTC R                     </pre>			<input type="text" value="Blast"/> go <ul style="list-style-type: none"> <li><input type="checkbox"/> Blast</li> <li><input type="checkbox"/> ProtParam</li> <li><input type="checkbox"/> Compute pI/MW</li> <li><input type="checkbox"/> ProtScale</li> <li><input type="checkbox"/> PeptideMass</li> <li><input type="checkbox"/> PeptideCutter</li> </ul>
<a href="#">« Hide</a>			

For more information, please refer to Entrez Help:

<http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helpentrez&part=EntrezHelp>

and

[http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helpgene&part=EntrezGene#EntrezGene.How\\_Data\\_Are\\_Display](http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helpgene&part=EntrezGene#EntrezGene.How_Data_Are_Display)

For more information on UniProt, please refer to <http://www.uniprot.org/docs/>

## Database Searching Using SRS

<http://srs.ebi.ac.uk/>

Go to database UR: <http://srs.ebi.ac.uk/>

→ type in keyword in text box → search → limits → select one of the entries from the hits → look at the annotations for more information (accession number, name, length, molecule type, description, literature information, comments/function, sequence → Retrieve sequence in fasta format

e) Use SRS to search **multiple** databases for *Homo sapiens WNT4* mRNA sequence.

Select the **Library Page** tab.

The screenshot shows the EMBL SRS Library Page. At the top, there is a search bar with the text "Enter Text Here" and a "Go" button. Below the search bar, there are several tabs: "Quick Search", "Library Page" (highlighted with a red box), "Query Form", "Results", "Projects", "Views", "Databanks", and "HELP".

On the left side, there are several sections:

- Search Options:** Contains instructions for selecting databanks and entering search terms. A red box highlights the "Standard Query Form" button.
- Tips:** Provides helpful links and information.
- BookMarkLets:** Lists various search filters like Protein Seq, DNA/RNA Seq, and Structures.

On the right side, there is a list of "Available Databanks" with expand/collapse buttons. The "Nucleotide sequence databases" section is expanded, showing a list of databases with checkboxes. A red arrow points to the "EMBL (Release)" checkbox, which is checked. A text box on the right says: "The more databases selected, the longer it will take to obtain the results".

Select the databases you wish to search from the Library Page : **EMBL (Release)**

To submit your query after selecting the databases to search, click on the Standard Query Form button.

The screenshot shows the EMBL search interface. In the 'Search Options' section, 'Combine search terms' is set to '& (AND)'. In the 'Result Display Options' section, 'EMBLSeqSimpleView' is selected. The 'Fields you can search' section shows 'Description' and 'Organism Name' set to 'WNT4' and 'Homo sapiens' respectively. The 'Create a view' section shows 'ID', 'Topology', 'Molecule', 'Data Class', 'Division', 'Sequence Length', and 'Accession Number' selected. The 'Sequence Format' is set to 'embl'.

Select the following fields :  
 Organism Name : Homo sapiens  
 Description : WNT4  
 Sequence format : fasta

The screenshot shows the EMBL search results page. The search query is '[[embirelease-Description:WNT4\*] & [[embirelease-O...]]'. The results table has columns for 'EMBL (Release)', 'Primary Accession (Links to SVA)', 'Accession List', 'Description', and 'Sequence Length'. The 'NCBI BLASTN' tool is selected in the 'Launch analysis tool' section. A red arrow points from a text box to the 'EMBL (Release)' column of the results table.

EMBL (Release)	Primary Accession (Links to SVA)	Accession List	Description	Sequence Length
<input type="checkbox"/> EMBL (Release):DR000161	DR000161	DR000161	TC109791 Human fetal brain, large insert, pCMV expression library Homo sapiens cDNA clone TC109791.5' similar to Homo sapiens wingless-type MMTV integration site family, member 4 (WNT4), mRNA sequence.	545
<input type="checkbox"/> EMBL (Release):AY419208	AY419208	AY419208	Homo sapiens WNT4 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	891
<input checked="" type="checkbox"/> EMBL (Release):AB061675	AB061675	AB061675	Homo sapiens mRNA for WNT4, complete cds.	1061
<input type="checkbox"/> EMBL (Release):AF335591	AF335591	AF335591	Homo sapiens WNT4 (WNT4) gene, partial cds.	6876
<input type="checkbox"/> EMBL (Release):AL031281	AL031281	AL031281	Human DNA sequence from clone RP1-22446 on chromosome 1p35.1-36.23 Contains the 3' UTR of the ELA3A gene for elastase 3A (pancreatic (protease E)), three novel genes, the CDC42 gene for cell division cycle 42 (GTP binding protein, 25kDa), the 3' end of the WNT4 gene for wingless-type MMTV integration site family (member 4) and a CpG island.	129366
<input type="checkbox"/> EMBL (Release):AL445253	AL445253	AL445253	Human DNA sequence from clone RP4-660K3 on chromosome 1 Contains the 5' end of the WNT4 gene for wingless-type MMTV integration site family (member 4) and a CpG island.	17429
<input type="checkbox"/> EMBL (Release):AY009398	AY009398	AY009398	Homo sapiens WNT4 precursor (WNT4) mRNA, complete cds.	1056
<input type="checkbox"/> EMBL (Release):AY033055	AY033055	AY033055	Homo sapiens secreted glycoprotein Wnt4 (WNT4) gene, promoter region.	570
<input type="checkbox"/> EMBL (Release):AY033056	AY033056	AY033056	Homo sapiens secreted glycoprotein Wnt4 (WNT4) gene, promoter region.	6283
<input type="checkbox"/> EMBL (Release):AY033057	AY033057	AY033057	Homo sapiens secreted glycoprotein Wnt4 (WNT4) gene, exon 1 and partial cds.	5334
<input type="checkbox"/> EMBL (Release):AY358947	AY358947	AY358947	Homo sapiens clone DNA48328 WNT4 (UNQ426) mRNA, complete cds.	2049

f) Click to save your results. Select FastaSeqs to set the display to FASTA format.

[Text Entry](#) | [EmblEntry](#) | [EMBLXML](#) | [INSDXML](#) | [Related Data](#)

---

Reset

Previous Entry
Entry 4 of 11 from [Query 1](#)
Next Entry

**Entry Information**

Entry from: [EMBL \(Release\)](#)

**Entry Options**

Launch analysis tool:

Link to related information:

Save entry:

View:

Go to: [General](#) | [Description](#) | [References](#) | [Cross-references](#) | [Sequence](#)

**General Information**

Primary Accession # AB061675  
 Accession # AB061675  
 SRS Entry ID EMBLRELEASE:AB061675  
 Molecule Type linear mRNA  
 Sequence Length 1061  
 Entry Division HUM (*Human*)  
 Entry Data Class STD (*Standard*)  
 Sequence Version AB061675.1  
 Creation Date 12-NOV-2002  
 Modification Date 07-OCT-2008  
 MBL-SVA [AB061675](#)

**Description**

Description Homo sapiens mRNA for WNT4, complete cds.  
 Keywords ;  
 Organism Homo sapiens (human)  
 Organism Classification Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

**References**

1. Katoh,M.; Submitted (11-MAY-2001) to the EMBL/GenBank/DDBJ databases. Contact:Masaru Katoh National Cancer Center Research Institute, Genetics and Cell Biology Section, Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan  
 Position 1-1061
2. Saitoh,T.; Hirai,M.; Katoh,M.; **Molecular cloning and characterization of WNT4 on human chromosome 1p35.1-p36.23** Unpublished.

**Database Cross-references**

ASTD [TRAN0000098771](#)  
 Ensembl-Gn [ENSG00000162552](#); Homo\_sapiens.  
 Ensembl-Tr [ENST00000290167](#); Homo\_sapiens.  
[ENST00000374655](#); Homo\_sapiens.  
 H-InvDB [HIT000059842](#).

**Features**

Key	Location	Qualifier	Value

[Databases](#) | [Tools](#) | [EBI Groups](#) | [Training](#) | [Industry](#) | [About Us](#) | [Help](#) | [Site Index](#)

---

[Quick Search](#) | [Library Page](#) | [Query Form](#) | [Tools](#) | [Results](#) | [Projects](#) | [Views](#) | [Databanks](#) | [HELP](#)

Reset

**Tips**

★ Problems downloading?  
 - see [Downloading Results](#)  
 in the [SRS@EBI FAQ](#)

Saving: Query  1 entries

Output To:  
 Browser Window (HTML)  File (text)

Number of entries to download:

Save As:  
 ASCII text/table  
 Save with view:   
 Column Separator:   
 Generic XML form  
 Using the loader:   
 Specific XML form  
 Using the loader:   
 Using XML Printer:

Save sequence in FASTA format

```

>embl|AB061675|AB061675 Homo sapiens mRNA for WNT4, complete cds.
gaaccatgagtcctccgctcgtgctcgttcgctgcgcctcctcgtcttcgccgctctctcagcccgccgagcaactggctgtacctg
gccaagctgctcgtcgggtgggagcatctcagaggaggagacgtgcgagaaactcaaggcctgatccagaggcaggtgcagatgtgcaa
gcggaacctggaagtcattgactcggctgcgcgcggtgccacgtggccattgaggagtgccagtagcaggtcccggaaccggtcgtgga
actgctccacactcgactccttgccgctcttcggcaagggtgagcgaaggactcgggagggcgcctctcgtgtacgccatctctcgcg
gaagttgatggcttgcaagtgcgcggcgctgcagcagctgggagcctccagggaagtgccgctgtgacagggacagtcagtggggtcacgcc
acagggcttccagtggtcaggatgctctgacaacatcgctcaggtgtggcctctcacagtcgcttctggtgagtgccgggagagaagca
agggggcctcgtccagcagagccctcatgaaacctccacaacaatgaggccggcaggaaggccatcctgacacacatgcgggtggaatgc
aagtgccacggggtgtcaggctcctgtgaggtaaagacgtgctggcgagccgtgccgcctctccgccaggtgggtcacgcactgaagga
gaagttgatggctgcaactgaggtggagccacgcgcgctgggctcctccagggaactggtgccaagcagcagctcaagccgca
cagatgaggacctggtgacttgagcctagccccgactctctgtgagcaggacatgcccagcggcgtctgggacagaggggcccacat
gcaacaagacgtccaagccatcgacgctgtgagctgctgtgctgtggccggtctccacacggcaggtggagctggctgaaagcgc
tcgagctgcaaatccactggtgctgctctgtcaagtgccggcagtgccagcggctcgtggagtgccacagctgcccgatga

```

g) Launch the EMBOSS sequence analysis tool, "Transeq" to translate the sequence.

Reset Previous Entry Entry 4 of 11 from Query 1 Next Entry

Go to: [General](#) [Description](#) [References](#) [Cross-references](#) [Sequence](#)

**Entry Information**  
Entry from: [EMBL](#) ([Release](#))

**Entry Options**  
Launch analysis tool:  
NCBI BLASTN  
Revseq  
SeqmatchallN  
SeqretN  
Showorf  
ShowseqN  
ShuffleseqN  
Silent  
Sixpack  
SplitterN  
StretcherN  
SupermatcherN  
Syco  
Tscan  
**Transeq**  
Trimest  
TrimseqN  
Vectorstrip  
WaterN  
Wtable

**General Information**  
Primary Accession # AB061675  
Accession # AB061675  
SRS Entry ID EMBLRELEASE:AB061675  
Molecule Type linear mRNA  
Sequence Length 1061  
Entry Division HUM (*Human*)  
Entry Data Class STD (*Standard*)  
Sequence Version AB061675.1  
Creation Date 12-NOV-2002  
Modification Date 07-OCT-2008  
EMBL-SVA [AB061675](#)

**Description**  
Description Homo sapiens mRNA for WNT4, complete cds.  
Keywords ;  
Organism Homo sapiens (*human*)  
Organism Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

**References**  
1. Katoh,M.; Submitted (11-MAY-2001) to the EMBL/GenBank/DBJ databases. Contact:Masaru Katoh National Cancer Center Research Institute, Genetics and Cell Biology Section, Genetics Division; Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan  
Position 1-1061  
2. Saitoh,T.; Hirai,M.; Katoh,M.; **Molecular cloning and characterization of WNT4 on human chromosome 1p35.1-p36.23** Unpublished.

**Database Cross-references**  
ASTD [TRAN00000098771](#).  
EMBL-Gen [ENSG00000162552](#); Homo\_sapiens.  
Ensembl Tr [ENST00000290167](#); Homo\_sapiens.  
[ENST00000374655](#); Homo\_sapiens.

Translate the sequence using Transeq.

Quick Search Library Page Query Form Tools Results Projects Views Databanks [HELP](#)

Reset [TRANSEQ--JobName: temp\_job2]

Query found 1 entries

TRANSEQ:temp\_job2\_AB061675\_1

>EMBLRELEASE:AB061675\_1  
AP\*VPARACVRCASSSSPSSQPPRATGCTUPSCRWVGASQRRRRARNRA\*SRGRCRCAS  
GTWKSWTRCAAVPSPWLRSASTSSGTGAGTAPHSFPCSSARW\*RKGLGRRPSTPFLRQ  
VWPLQ\*RGRAAVGSRSAAVTGQCGSRRHRASSGQDALTTSPVWPSHSRLWMCGRARG  
PRPAEP\*TSITTRFP&GRPS\*HTCGWNASATGQAPVR\*RRAGEPCRPSARVVTH\*RRSL  
MVPLRWSHAAMAPPCHWCHATHSSSRTOHRTWCTWSLAFTSVSRTCACACWARGAARHATR  
RFRPFTAVSCCAVAA&STRRRWSWLNAAAANSTGA&SS&G&S&G&S&W&S&CTRADYX

**Apply Options to:**  
 selected results only  
 unselected results only

**Result Options**  
Launch analysis tool:  
Garnier **Launch**  
Extractseq  
FreakP  
Fuzzpro  
**Garnier**  
Helixturnhelix  
Hmoment  
lep  
InfoseqP  
MaskseqP  
MatcherP  
MsbarP  
NeedleP  
Octanol  
PasteseqP  
Patmatdb  
Patmatmotifs  
Pepcoil  
Pepinfo  
Pepnet

Launch another sequence analysis tool to find out the secondary structure of the amino acid sequence obtained in the previous analysis.