

به نام خداوند بخشنده مهربان

روشها و  
پروتکل‌های  
ژنتیکی



# ژنتیکا

[www.genetica.ir](http://www.genetica.ir)

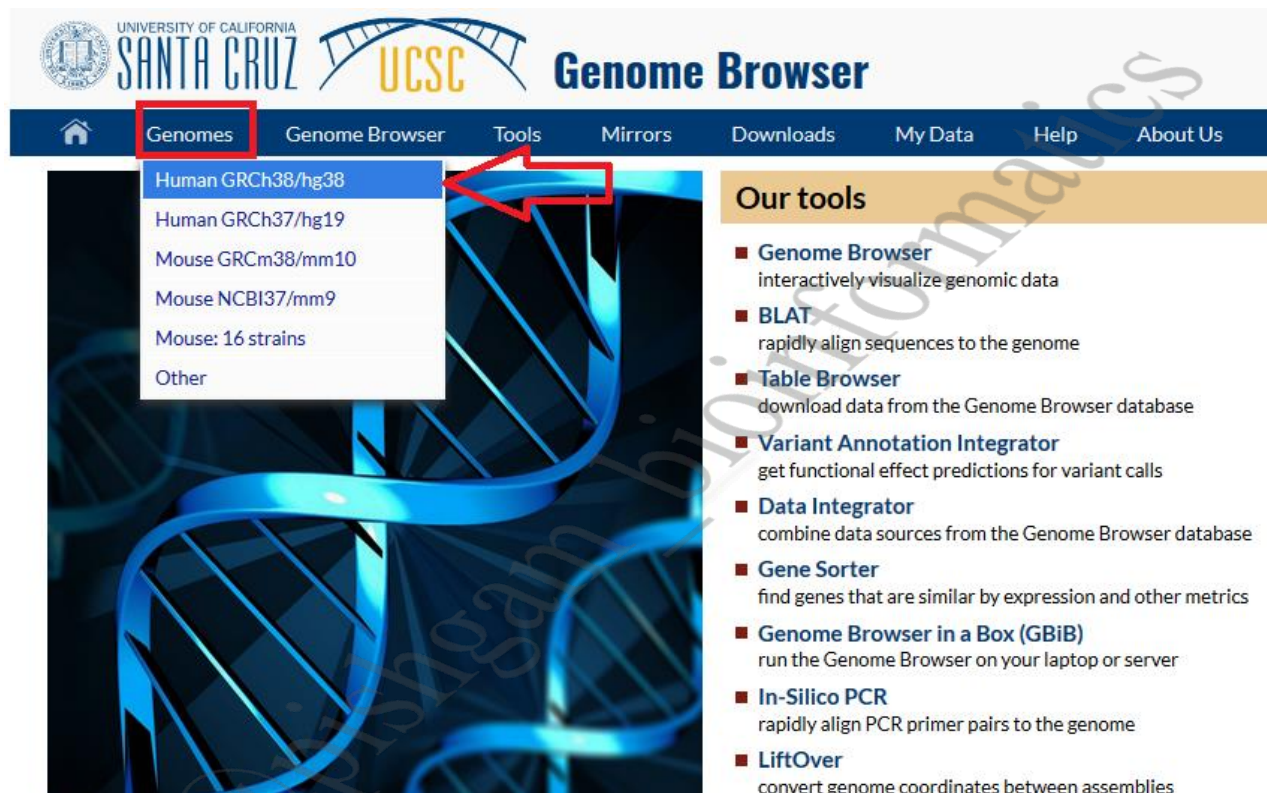
**آموزش سایت UCSC: بررسی بیان ژن ها با RNA-seq و میکرواری، تعیین مسیرهای سیگنالی، عملکرد مولکولی و هم چنین فرایندهای زیستی (gene ontology) ژن ها**

این آموزش شامل موارد زیر می شود:

۱. آموزش نحوه جستجوی ژن ها در سایت UCSC
۲. آموزش بررسی بیان ژن ها در بافت های مختلف با داده های RNA-seq در سایت ucsc
۳. آموزش تعیین عملکرد مولکولی و هم چنین فرایندهای زیستی ژن ها در سایت UCSC
۴. آموزش تعیین مسیرهای سیگنالی مرتبط برای ژن ها در سایت UCSC
۵. آموزش بررسی بیان ژن ها در بافت های مختلف با داده های میکرواری در سایت UCSC
۶. آموزش شناسایی دومین های پروتئین و ساختار سه بعدی پروتئین ها در سایت UCSC

یکی از سایت های بسیار مفید برای بررسی اطلاعات ژن ها مانند بیان آن ها در بافت های مختلف سایت UCSC به آدرس زیر می باشد.

<https://genome.ucsc.edu/>



The image shows the UCSC Genome Browser website. The top navigation bar includes 'Genomes', 'Genome Browser', 'Tools', 'Mirrors', 'Downloads', 'My Data', 'Help', and 'About Us'. The 'Genomes' menu is expanded, showing a list of genome assemblies: 'Human GRCh38/hg38', 'Human GRCh37/hg19', 'Mouse GRCh37/mm9', 'Mouse NCBI37/mm9', 'Mouse: 16 strains', and 'Other'. A red box highlights the 'Genomes' menu, and a red arrow points to the 'Human GRCh38/hg38' option. On the right side, the 'Our tools' section lists various tools: 'Genome Browser' (interactively visualize genomic data), 'BLAT' (rapidly align sequences to the genome), 'Table Browser' (download data from the Genome Browser database), 'Variant Annotation Integrator' (get functional effect predictions for variant calls), 'Data Integrator' (combine data sources from the Genome Browser database), 'Gene Sorter' (find genes that are similar by expression and other metrics), 'Genome Browser in a Box (GBiB)' (run the Genome Browser on your laptop or server), 'In-Silico PCR' (rapidly align PCR primer pairs to the genome), and 'LiftOver' (convert genome coordinates between assemblies).

برای جستجوی ژن ها در سایت UCSC در صفحه اصلی این سایت همانند شکل فوق از منوی Genomes گزینه Human GRCh38/hg38 را انتخاب نمایید تا به صفحه زیر منتقل شوید.

در صفحه باز شده همانند شکل فوق در کادر ۱ اسم ژن موردنظر خود را وارد نمایید. در صورتی که اسم وارد شده برای سایت UCSC شناخته شده باشد چند گزینه با پیشنهاد می دهد، بعد از انتخاب ژن موردنظر بر روی گزینه go کلیک نمایید.

نکته: در صورتی که سایت گزینه ای را برای اسم وارد شده پیشنهاد نمود نیز بر روی گزینه go کلیک نمایید.

بعد از کلیک بر روی گزینه go به صفحه زیر منتقل خواهید شد.

## Known Genes

GDNF (uc063dca.1) at chr5:37815651-37839686 - Neurotrophic factor that enhances survival and morphological differentiation of dopaminergic neurons (GDNF)

GDNF-AS1 (uc021xxo.2) at chr5:37840438-37874894 - Homo sapiens GDNF antisense RNA 1 (head to head) (GDNF-AS1), transcript variant GDNF-AS1 (uc063dcd.1) at chr5:37869907-37874894 - Homo sapiens GDNF antisense RNA 1 (head to head) (GDNF-AS1), transcript variant GDNF-AS1 (uc063dcc.1) at chr5:37816026-37835491 - glial cell derived neurotrophic factor (from HGNC GDNF)

GDNF (uc063dcb.1) at chr5:37816026-37834908 - glial cell derived neurotrophic factor (from HGNC GDNF)

GDNF-AS1 (uc031sjd.2) at chr5:37873425-37875799 - GDNF antisense RNA 1 (head to head) (from HGNC GDNF-AS1)

GDNF (uc011cpi.2) at chr5:37812677-37839680 - Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 1, mRNA.

GDNF (uc011cph.3) at chr5:37815651-37835827 - Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 4, mRNA.

GDNF (uc011cpg.3) at chr5:37815651-37835827 - Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 3, mRNA.

GDNF (uc011cpf.2) at chr5:37812677-37835491 - Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 2, mRNA.

GDNF (uc011cpd.1) at chr5:37812677-37834822 - Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 6, mRNA.

GFRA2 (uc064kvc.1) at chr8:21779550-21788159 - GDNF family receptor alpha 2 (from HGNC GFRA2)

GZF1 (uc061vvi.1) at chr20:23368871-23370421 - GDNF-inducible zinc finger protein 1 (from HGNC GZF1)

GZF1 (uc061vvh.1) at chr20:23365219-23368907 - GDNF-inducible zinc finger protein 1 (from HGNC GZF1)

GFRA1 (uc057wds.1) at chr10:116272004-116272643 - GDNF family receptor alpha 1 (from HGNC GFRA1)

GFRA1 (uc009xyr.4) at chr10:116056933-116273284 - Homo sapiens GDNF family receptor alpha 1 (GFRA1), transcript variant 3, mRNA. (from RefSeq NM\_001162732.4)

GFRA1 (uc003pcm.1) at chr6:55327469-55402493 - Homo sapiens GDNF family receptor alpha like (GFRA1), mRNA. (from RefSeq NM\_207410)

GFRA3 (uc003lcn.4) at chr5:138252379-138274671 - Homo sapiens GDNF family receptor alpha 3 (GFRA3), mRNA. (from RefSeq NM\_001496)

GFRA4 (uc002wio.4) at chr20:3659909-3663399 - Homo sapiens GDNF family receptor alpha 4 (GFRA4), transcript variant 2, mRNA. (from RefSeq NM\_001162732.4)

GFRA1 (uc057wdr.1) at chr10:116064401-116273033 - Receptor for GDNF. Mediates the GDNF-induced autophosphorylation and activation of GDNF family receptor signaling pathway. (from RefSeq NM\_001162732.4)

GFRA2 (uc003wzw.2) at chr8:21692935-21788323 - Homo sapiens GDNF family receptor alpha 2 (GFRA2), transcript variant 3, mRNA. (from RefSeq NM\_001162732.4)

GFRA2 (uc003wzv.2) at chr8:21693182-21788459 - Homo sapiens GDNF family receptor alpha 2 (GFRA2), transcript variant 2, mRNA. (from RefSeq NM\_001162732.4)

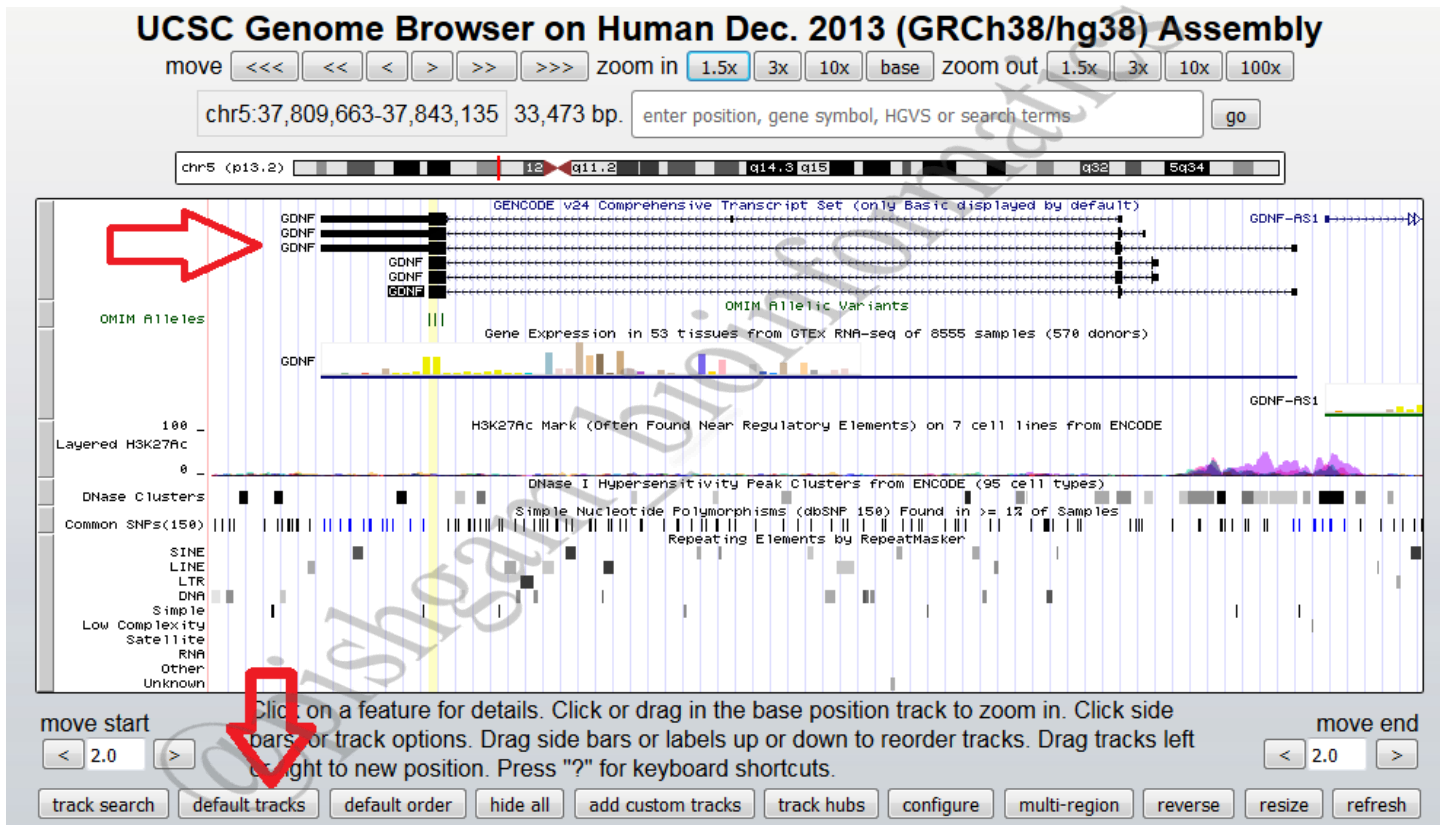
GFRA2 (uc003wzu.2) at chr8:21690403-21788810 - Homo sapiens GDNF family receptor alpha 2 (GFRA2), transcript variant 1, mRNA. (from RefSeq NM\_001162732.4)

GZF1 (uc002wsz.4) at chr20:23864384-23873045 - Homo sapiens GDNF inducible zinc finger protein 1 (GZF1), transcript variant 2, mRNA. (from RefSeq NM\_001162732.4)

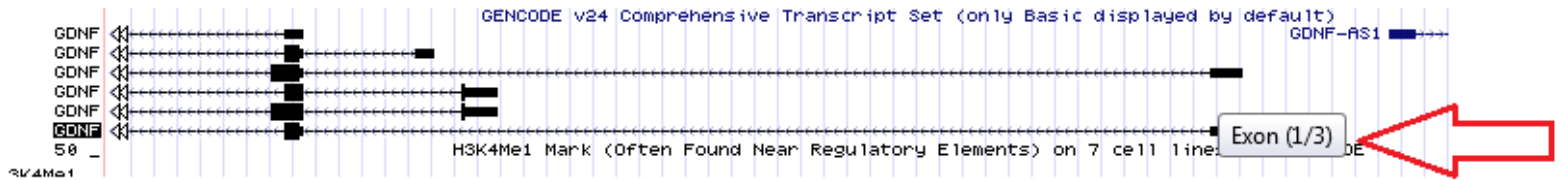
GFRA4 (uc002win.3) at chr20:3659292-3663399 - Homo sapiens GDNF family receptor alpha 4 (GFRA4), transcript variant 1, mRNA. (from RefSeq NM\_001162732.4)

GFRA1 (uc001lcj.4) at chr10:116063764-116273467 - Homo sapiens GDNF family receptor alpha 1 (GFRA1), transcript variant 1, mRNA. (from RefSeq NM\_001162732.4)

در صفحه باز شده سایت UCSC چندین ژن را برای عبارت جستجو شده در قسمت Known Genes نشان می‌دهد که در این قسمت باید بر روی ژن موردنظر خود کلیک نمایید.



با کلیک بر روی ژن موردنظر به صفحه فوق منتقل خواهید شد که در آن می‌توانید ژن انتخاب شده خود را مشاهده نمایید. در صورتی که در پنجره فوق اسم ژن موردنظر خود را مشاهده نمودید بر روی گزینه default tracks کلیک نمایید. در مورد اطلاعاتی که در پنجره فوق به نمایش در آمده است در ادامه توضیحاتی را ارائه خواهیم داد.



برای مشاهده اطلاعات هر ژن همانند شکل فوق بر روی شکل گرافیکی ژن ها کلیک نمایید و با کلیک بر روی شکل ژن ها به صفحه زیر منتقل خواهید شد.

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

### Human Gene GDNF (ENST00000326524.6) Description and Page Index

**Description:** Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 1, mRNA. (from RefSeq NM\_000514)

**RefSeq Summary (NM\_000514):** This gene encodes a secreted ligand of the TGF-beta (transforming growth factor-beta) superfamily of pro Ligands of this family bind various TGF-beta receptors leading to recruitment and activation of SMAD family transcription factors that regulate expression. The encoded preproprotein is proteolytically processed to generate each subunit of the disulfide-linked homodimer. The recombinant form of this protein, a highly conserved neurotrophic factor, was shown to promote the survival and differentiation of dopaminergic neurons and was able to prevent apoptosis of motor neurons induced by axotomy. This protein is a ligand for the product of the RET (rearranged during transfection) protooncogene. Mutations in this gene may be associated with Hirschsprung disease and Tourette syndrome. This gene encodes multiple protein isoforms that may undergo similar proteolytic processing. [provided by RefSeq, Aug 2016].

**Gencode Transcript:** ENST00000326524.6  
**Gencode Gene:** ENSG00000168621.14

**Transcript (Including UTRs)**  
**Position:** hg38 chr5:37,812,677-37,839,680 **Size:** 27,004 **Total Exon Count:** 3 **Strand:** -

**Coding Region**  
**Position:** hg38 chr5:37,815,651-37,834,796 **Size:** 19,146 **Coding Exon Count:** 2

در شکل فوق، توضیحاتی در مورد ژن مورد بررسی، نقش آن، اندازه و تعداد آگزون ها مشاهده می نماید.

Page Index	Sequence and Links	UniProtKB Comments	MalaCards	CTD	RNA-Seq Expression
Microarray Expression	RNA Structure	Protein Structure	Other Species	GO Annotations	mRNA Descriptions
Pathways	Other Names	GeneReviews	Methods		

Data last updated: 2016-03-28

در جدول فوق می توانید اطلاعات زیر را مشاهده نمایید:

**Microarray Expression:** مشاهده بیان ژن در بافت ها و نمونه های مختلف با تکنیک میکروارنا

**Protein Structure:** مشاهده ساختار سه بعدی و همچنین دومین های پروتئین



GO Annotations: مشاهده عملکرد مولکولی و فرایند زیستی که ژن مورد بررسی در آن ها دخیل می باشد.

RNA-Seq Expression: مشاهده بیان ژن در بافت ها و نمونه های مختلف با تکنیک RNA-Seq

Pathways: مشاهده مسیرهای سیگنالی که ژن مربوطه در آن دخیل می باشد.

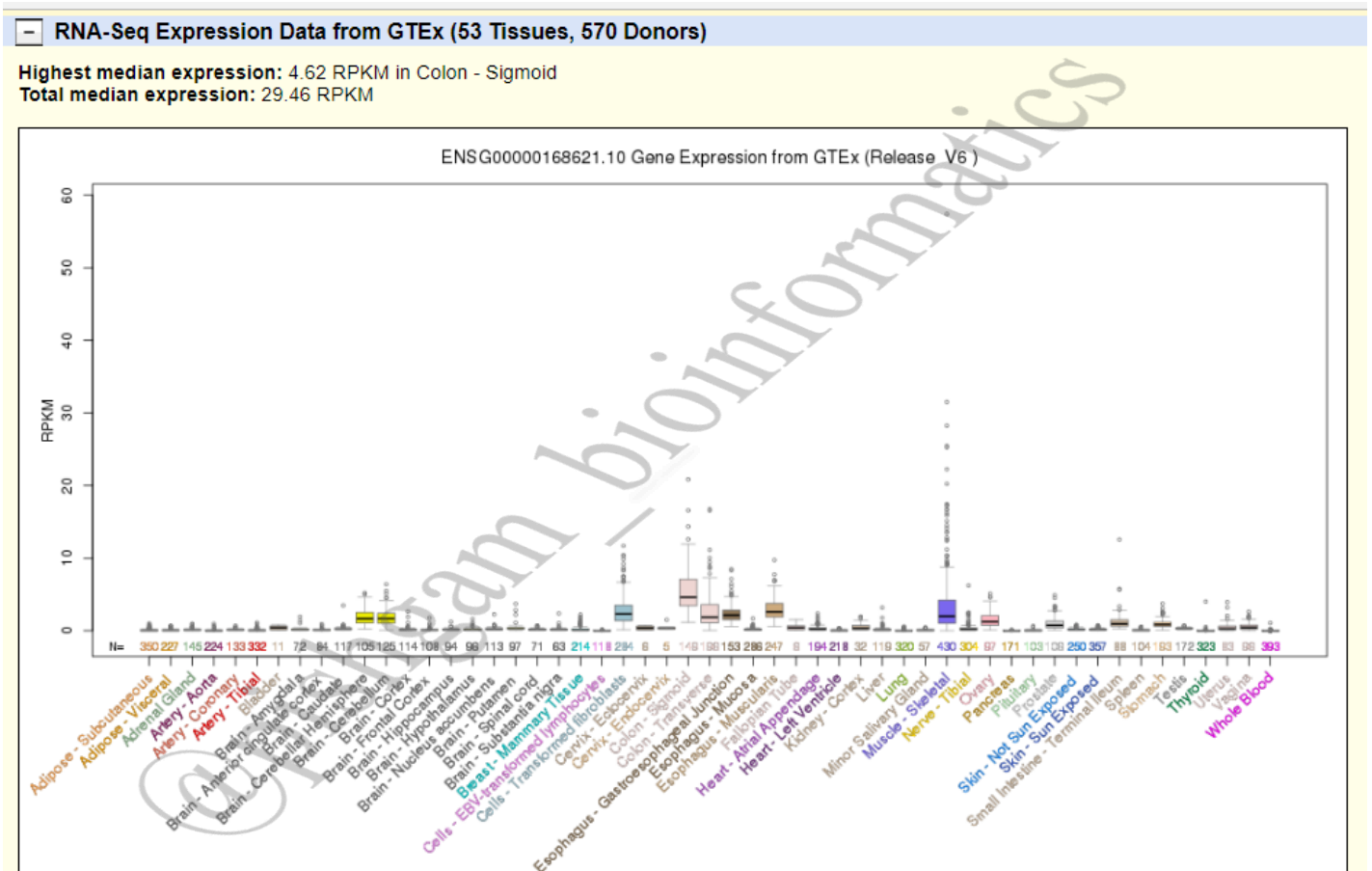
Other Names: مشاهده سایر اسم های ژن مورد بررسی



- Sequence and Links to Tools and Databases					
Genomic Sequence (chr5:37,812,677-37,839,680)		mRNA (may differ from genome)		Protein (211 aa)	
Gene Sorter	Genome Browser	Other Species FASTA	Gene interactions	Table Schema	BioGPS
CGAP	Ensembl	Entrez Gene	ExonPrimer	Gencode	GeneCards
Gepis Tissue	HGNC	HPRD	Lynx	MGI	MOPED
neXtProt	OMIM	PubMed	Reactome	Stanford SOURCE	UniProtKB
Wikipedia					

در جدول فوق نیز می توانید با کلیک بر روی هر پایگاه داده اطلاعات مربوط به ژن مورد بررسی را در سایر پایگاه های زیستی مشاهده نمایید.

### بررسی بیان ژن ها در بافت های مختلف با داده های RNA-seq در سایت ucsc





در بخش RNA-Seq Expression می توان اصلاحات مربوط به بیان ژن مربوطه را در نمونه های مختلف که با روش RNA-Seq بررسی شده است مشاهده نمود. در نمودار فوق داده بیان ژن برای نمونه ها و بافت های مختلف به صورت جعبه ای نمایش داده می شود میزان بیان ژن ها به صورت RPKM نمایش داده شده است که مقادیر بالاتر به معنی بیان بالاتر ژن در آن بافت می باشد.

تعیین عملکرد مولکولی و هم چنین فرایندهای زیستی ژن ها

@pishgam\_bioinformatics

## Gene Ontology (GO) Annotations with Structured Vocabulary

### Molecular Function:

[GO:0005102](#) receptor binding  
[GO:0008083](#) growth factor activity  
[GO:0042803](#) protein homodimerization activity

### Biological Process:

[GO:0001656](#) metanephros development  
[GO:0001657](#) ureteric bud development  
[GO:0001658](#) branching involved in ureteric bud morphogenesis  
[GO:0001755](#) neural crest cell migration  
[GO:0001759](#) organ induction  
[GO:0001941](#) postsynaptic membrane organization  
[GO:0003337](#) mesenchymal to epithelial transition involved in metanephros morphogenesis  
[GO:0007165](#) signal transduction  
[GO:0007399](#) nervous system development  
[GO:0007411](#) axon guidance  
[GO:0007422](#) peripheral nervous system development  
[GO:0008344](#) adult locomotory behavior  
[GO:0021784](#) postganglionic parasympathetic nervous system development  
[GO:0030432](#) peristalsis  
[GO:0031175](#) neuron projection development  
[GO:0032770](#) positive regulation of monooxygenase activity  
[GO:0033603](#) positive regulation of dopamine secretion  
[GO:0043066](#) negative regulation of apoptotic process  
[GO:0043524](#) negative regulation of neuron apoptotic process  
[GO:0045944](#) positive regulation of transcription from RNA polymerase II promoter  
[GO:0048255](#) mRNA stabilization  
[GO:0048484](#) enteric nervous system development  
[GO:0048485](#) sympathetic nervous system development  
[GO:0051584](#) regulation of dopamine uptake involved in synaptic transmission  
[GO:0060676](#) ureteric bud formation  
[GO:0060688](#) regulation of morphogenesis of a branching structure  
[GO:0072107](#) positive regulation of ureteric bud formation  
[GO:0072108](#) positive regulation of mesenchymal to epithelial transition involved in metanephros morphogenesis  
[GO:0090190](#) positive regulation of branching involved in ureteric bud morphogenesis  
[GO:2001240](#) negative regulation of extrinsic apoptotic signaling pathway in absence of ligand

برای تعیین عملکرد مولکولی و هم چنین فرایندهای زیستی که ژن مورد نظر در آن شرکت دارد می

توان در قسمت Gene Ontology (GO) Annotations with Structured Vocabulary

لیست عملکرد مولکولی (در بخش Molecular Function) و فرایندهای زیستی که ژن مورد نظر

در آن شرکت دارد (در بخش Biological Process) را مشاهده نمود.

## تعیین مسیرهای سیگنالی مرتبط برای ژن ها

### - Biochemical and Signaling Pathways

#### Reactome (by CSHL, EBI, and GO)

Protein P39905 ([Reactome details](#)) participates in the following event(s):

[R-HSA-375144](#) Interaction of NCAM1:GFRalpha-1 with GDNF  
[R-HSA-5672965](#) RAS GEFs promote RAS nucleotide exchange  
[R-HSA-419037](#) NCAM1 interactions  
[R-HSA-375165](#) NCAM signaling for neurite out-growth  
[R-HSA-422475](#) Axon guidance  
[R-HSA-1266738](#) Developmental Biology  
[R-HSA-5673001](#) RAF/MAP kinase cascade  
[R-HSA-112412](#) SOS-mediated signalling  
[R-HSA-167044](#) Signalling to RAS  
[R-HSA-170968](#) Frs2-mediated activation  
[R-HSA-170984](#) ARMS-mediated activation  
[R-HSA-179812](#) GRB2 events in EGFR signaling  
[R-HSA-180336](#) SHC1 events in EGFR signaling  
[R-HSA-186763](#) Downstream signal transduction  
[R-HSA-187706](#) Signalling to p38 via RIT and RIN  
[R-HSA-451927](#) Interleukin-2 signaling  
[R-HSA-881907](#) Gastrin-CREB signalling pathway via PKC and MAPK  
[R-HSA-912526](#) Interleukin receptor SHC signaling  
[R-HSA-1250347](#) SHC1 events in ERBB4 signaling  
[R-HSA-1433557](#) Signaling by SCF-KIT  
[R-HSA-2424491](#) DAP12 signaling  
[R-HSA-2586552](#) Signaling by Leptin  
[R-HSA-2871796](#) FCERI mediated MAPK activation  
[R-HSA-5218921](#) VEGFR2 mediated cell proliferation  
[R-HSA-5654693](#) FRS-mediated FGFR1 signaling  
[R-HSA-5654700](#) FRS-mediated FGFR2 signaling  
[R-HSA-5654706](#) FRS-mediated FGFR3 signaling  
[R-HSA-5654712](#) FRS-mediated FGFR4 signaling  
[R-HSA-5684996](#) MAPK1/MAPK3 signaling  
[R-HSA-112399](#) IRS-mediated signalling  
[R-HSA-187687](#) Signalling to ERKs  
[R-HSA-169893](#) Prolonged ERK activation events  
[R-HSA-177929](#) Signaling by EGFR  
[R-HSA-186797](#) Signaling by PDGF

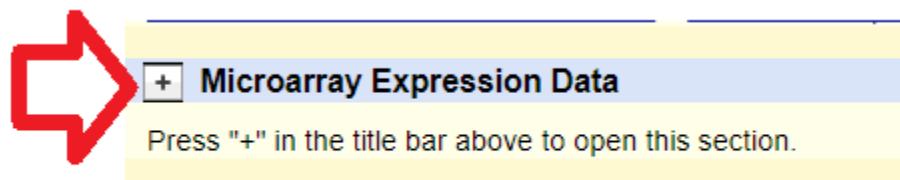
برای تعیین مسیرهای سیگنالی که ژن مورد نظر در آن شرکت دارد می توان در قسمت **Biochemical and Signaling Pathways** لیست مسیرهای سلولی که ژن مورد نظر در آن شرکت دارد را مشاهده نمود.

**- Other Names for This Gene**

**Alternate Gene Symbols:** B7WPK7, GDNF\_HUMAN, NM\_000514, O95448, O95449, O95986, P39905, Q6FH33, Q96L44, Q9UD32, Q9UD33, Q9UMV2, Q9UP67, Q9UP97  
**UCSC ID:** uc011cpl.2  
**RefSeq Accession:** [NM\\_000514](#)  
**Protein:** [P39905](#) (aka GDNF\_HUMAN)  
**CCDS:** [CCDS3922.1](#)

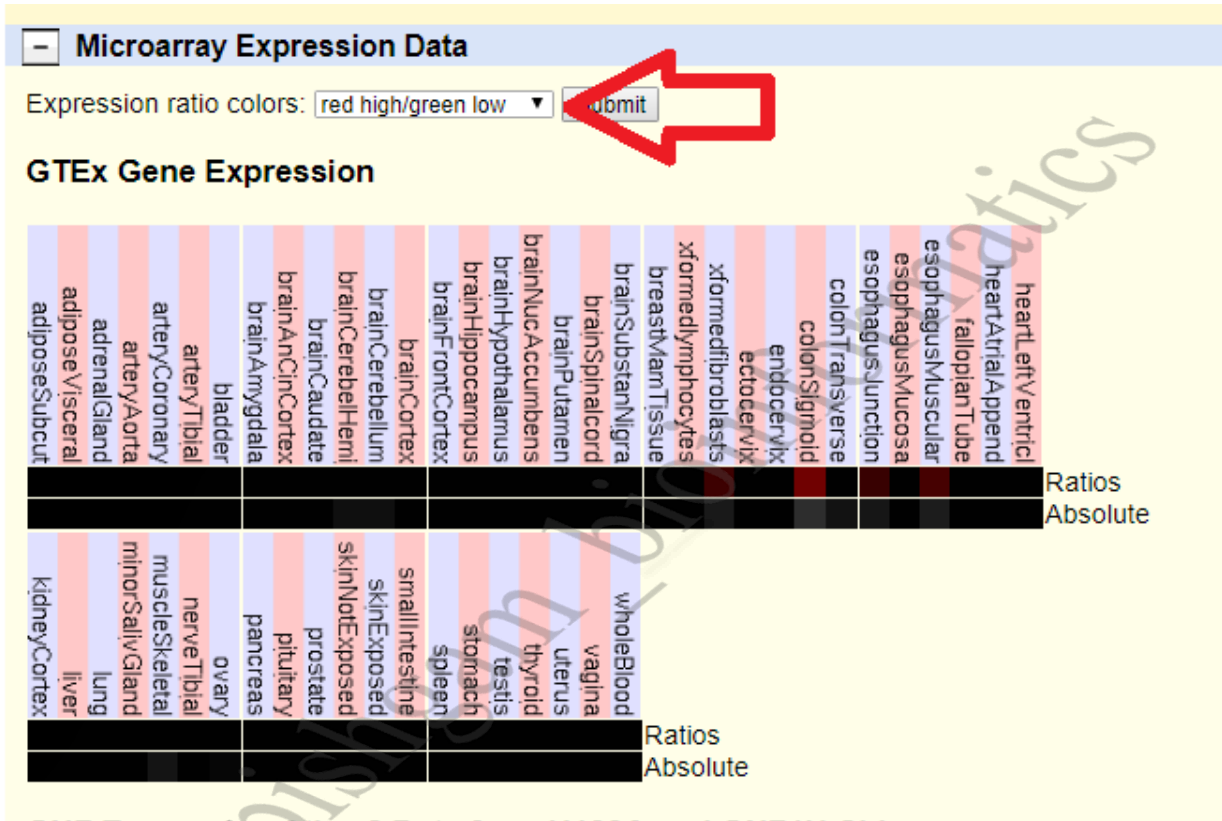
در قسمت **Other Names for This Gene** می توان سایر اسم های ژن مورد نظر را مشاهده نمود.

بررسی بیان ژن ها در بافت های مختلف با داده های میکروآرای در سایت **UCSC**



**+ Microarray Expression Data**  
Press "+" in the title bar above to open this section.

در بخش **Microarray Expression Data** می توان اطلاعات مربوط به بیان ژن مربوطه را در نمونه های مختلف که با روش میکروآرای بررسی شده است مشاهده نمود برای مشاهده داده های بیان ژن بر روی علامت + کلیک نمایید تا داده های زیر نمایش داده شود.



داده بیان ژن برای نمونه ها و بافت های مختلف به صورت heatmap نمایش داده می شود در قسمت Expression ratio colors: می توان تعیین نمود که بیان بالا و بیان پایین با چه رنگی نمایش داده شود سپس در قسمت Ratio می توان بیان ژن مورد نظر را در بین نمونه های مختلف مشاهده نمود که در این مثال رنگ قرمز بیانگر بیان بیشتر ژن می باشد که ژن مورد بررسی در بافت Colon sigmoid دارای بیان بیشتری نسبت به سایر بافت ها می باشد.

## شناسایی دومین های پروتئین و ساختار سه بعدی پروتئین ها

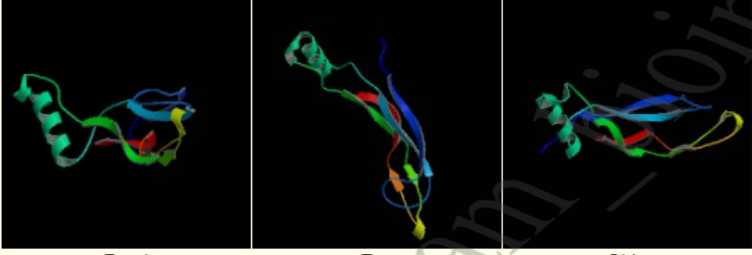
**Protein Domain and Structure Information**

**InterPro Domains:** [Graphical view of domain structure](#)  
[IPR029034](#) - Cystine-knot\_cytokine  
[IPR016649](#) - GDNF  
[IPR001839](#) - TGF-b\_C

**Pfam Domains:**  
[PF00019](#) - Transforming growth factor beta like domain

**Protein Data Bank (PDB) 3-D Structure**  
[2V5E](#) - X-ray [Chimera3FUB](#) - X-ray [Chimera4UX8](#) - EM [Chimera](#)  
[Chimera help](#)

**ModBase Predicted Comparative 3D Structure on [P39905](#)**



دومین ها پروتئین

ساختار سه بعدی پروتئین

Front Top Side

The pictures above may be empty if there is no ModBase structure for the protein. The ModBase structure frequently covers just a fragment of the protein. You may be asked to log onto ModBase the first time you click on the pictures. It is simplest after logging in to just click on the picture again to get to the specific info on that model.

برای شناسایی خانواده، دومین ها و ساختار سه بعدی پروتئین ها می توان از بخش Protein Domain and Structure Information استفاده نمود به این صورت که در بخش InterPro Domains می توان خانواده دومین های پروتئینی و در بخش Protein Data Bank (PDB) 3-D Structure ساختار سه بعدی پروتئین را مشاهده نمود.

**- Protein Domain and Structure Information**

**InterPro Domains:** [Graphical view of domain structure](#)  
[IPR029034](#) - Cystine-knot\_cytokine  
[IPR016649](#) - GDNF  
[IPR001839](#) - TGF-b\_C

**Pfam Domains:**  
[PF00019](#) - Transforming growth factor beta like domain

با کلیک بر روی دومین ها همانند شکل فوق توضیحات مربوط به آن دومین و یا خانواده همانند شکل زیر نمایش داده خواهد شد.

## Homologous Superfamily


### Cystine-knot cytokine (IPR029034)

Short name: Cystine-knot\_cytokine

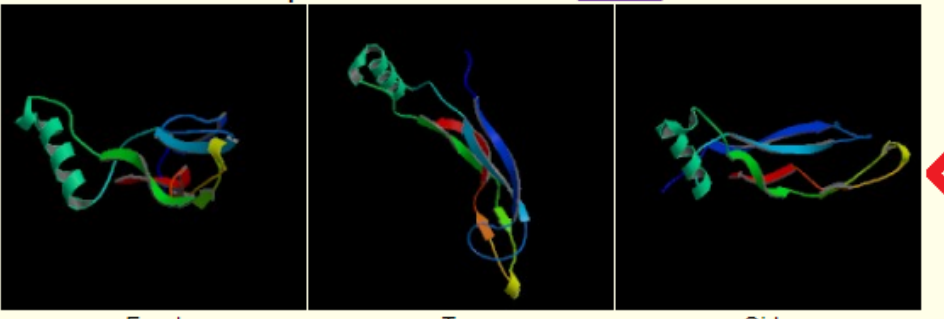
#### Description

This cystine-knot cytokine domain is a disulfide-rich, all-beta core domain that is found in glycoprotein hormones [PMID: 9218442], in members of the TGF-beta family [PMID: 21369710], IL-17 family [PMID: 21337007] and VEGF growth factor family [PMID: 21917115] among others.



**Protein Data Bank (PDB) 3-D Structure**  
  
[2V5E - X-ray Chimera](#)[3FUB - X-ray Chimera](#)[4UX8 - EM Chimera](#)  
[Chimera help](#)

**ModBase Predicted Comparative 3D Structure on P39905**




Front Top Side

















The pictures above may be empty if there is no ModBase structure for the protein. The M  
 You may be asked to log onto ModBase the first time you click on the pictures. It is simpl  
 specific info on that model.

هم چنین با کلیک نمودن بر روی تصاویر سه بعدی پروتئین صفحه زیر باز خواهد شد که در آن می توانید مشخصات ساختار سه بعدی را مشاهده نمایید.


**Model Information**

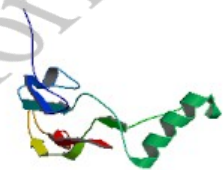
Perform action on this model 


Quality criteria indicate whether the model is considered reliable (green) or unreliable (red).

 Target Region	109-211
 Protein Length	211
 Template PDB Code	<a href="#">3fubD</a> 
 Template Region	32-134
 Sequence Identity	100.00%
 E-Value	0
 GA341	1.00
 MPQS	1.52495
 z-DOPE	0.58
 TSVM Method	MSALL
 TSVM RMSD	1.398
 TSVM NO35	0.968
 Dataset	MW-Human_D11
 ModPipe Version	SVN.r1600
 Model Date	2016-06-12

Mouse over the images for more information

Sequence Model Coverage Summary for all Models of this Sequence 



Filtered models for current sequence  ([Show all models](#))



شکل فوق بیانگر ویژگی های ساختار سه بعدی پروتئین ها می باشد برای دانلود فایل ساختار سه بعدی پروتئین در قسمت Template PDB Code بر روی لینک نمایش داده شده کلیک نمایید تا صفحه زیر باز شود.



سپس در صفحه فوق زبانه 3D view را کلیک نمایید تا ساختار سه بعدی پروتئین نمایش داده شود و با کلیک بر روی گزینه Download Files فایل PDB مربوط به ساختار سه بعدی پروتئین دانلود خواهد شد.