

# Using Entrez GENE

March 23-24, 2009

To access Entrez GENE: Click on “Genomic Biology” category

**NCBI**  
National Center for Biotechnology Information  
[National Library of Medicine](#) [National Institutes of Health](#)

PubMed All Databases BLAST OMIM Books TaxBrowser Structure

Search  for

**SITE MAP**  
Alphabetical List  
Resource Guide

**About NCBI**  
An introduction to NCBI

**GenBank**  
Sequence submission support and software

**Literature databases**  
PubMed, OMIM, Books, and PubMed Central

**Molecular databases**  
Sequences, structures, and taxonomy

**Genomic biology**  
The human genome, whole genomes, and related resources

**What does NCBI do?**  
Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. [More about NCBI...](#)

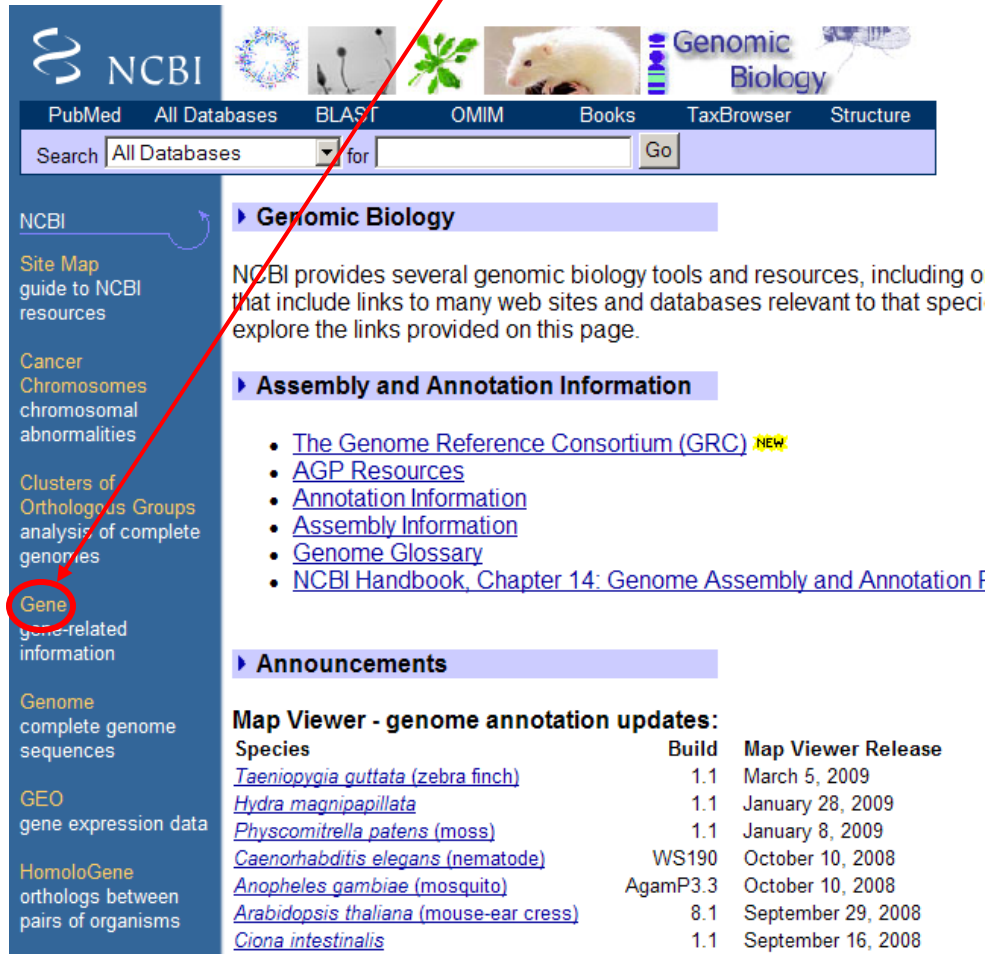
**Hot Spots**

- ▶ Clusters of orthologous groups
- ▶ Coffee Break, Genes & Disease, NCBI Handbook
- ▶ Electronic PCR
- ▶ Entrez Home
- ▶ Entrez Tools
- ▶ Gene expression omnibus (GEO)
- ▶ Human genome resources
- ▶ Influenza Virus Resource
- ▶ Map Viewer
- ▶ dbMHC
- ▶ Mouse genome resources

**Genome Reference Consortium**  
The [Genome Reference Consortium](#) (GRC) has been formed to continue the improvement of the human and mouse genome reference assemblies. The goal of the GRC is to fix the small number of loci that may be misrepresented in the reference assembly, fill the remaining gaps, and to produce alternate representations of complex loci.

**PubMed Central**  
[PubMed Central](#) is an archive of biomedical and life sciences journals.

Click on "Gene"



The image shows a screenshot of the NCBI Genomic Biology page. A red arrow points from the text "Click on 'Gene'" to the "Gene" link in the left sidebar. The page features a search bar at the top, a navigation menu, and several sections of content including "Genomic Biology", "Assembly and Annotation Information", and "Announcements".

NCBI

Site Map  
guide to NCBI resources

Cancer  
Chromosomes  
chromosomal abnormalities

Clusters of  
Orthologous Groups  
analysis of complete genomes

Gene  
gene-related information

Genome  
complete genome sequences

GEO  
gene expression data

HomoloGene  
orthologs between pairs of organisms

PubMed All Databases BLAST OMIM Books TaxBrowser Structure

Search All Databases for Go

► Genomic Biology

NCBI provides several genomic biology tools and resources, including o that include links to many web sites and databases relevant to that speci explore the links provided on this page.

► Assembly and Annotation Information

- [The Genome Reference Consortium \(GRC\)](#) NEW
- [AGP Resources](#)
- [Annotation Information](#)
- [Assembly Information](#)
- [Genome Glossary](#)
- [NCBI Handbook, Chapter 14: Genome Assembly and Annotation F](#)

► Announcements

**Map Viewer - genome annotation updates:**

Species	Build	Map Viewer Release
<a href="#">Taeniopygia guttata (zebra finch)</a>	1.1	March 5, 2009
<a href="#">Hydra magnipapillata</a>	1.1	January 28, 2009
<a href="#">Physcomitrella patens (moss)</a>	1.1	January 8, 2009
<a href="#">Caenorhabditis elegans (nematode)</a>	WS190	October 10, 2008
<a href="#">Anopheles gambiae (mosquito)</a>	AgamP3.3	October 10, 2008
<a href="#">Arabidopsis thaliana (mouse-ear cress)</a>	8.1	September 29, 2008
<a href="#">Ciona intestinalis</a>	1.1	September 16, 2008

You will use Entrez Gene to answer question #3 on the datasheet. The following screenshots provide clues on where to look for the answers.

You can search by name of disease, gene name, gene symbol. Remember to limit your search to human [orgn]

NCBI Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure OMIM

Search Gene for **cx26 [sym] AND human [orgn]** Go Clear

Limits Preview/Index History Clipboard Details

Entrez Gene is a searchable database of genes, from [RefSeq](#) genomes, a [Map Viewer](#)

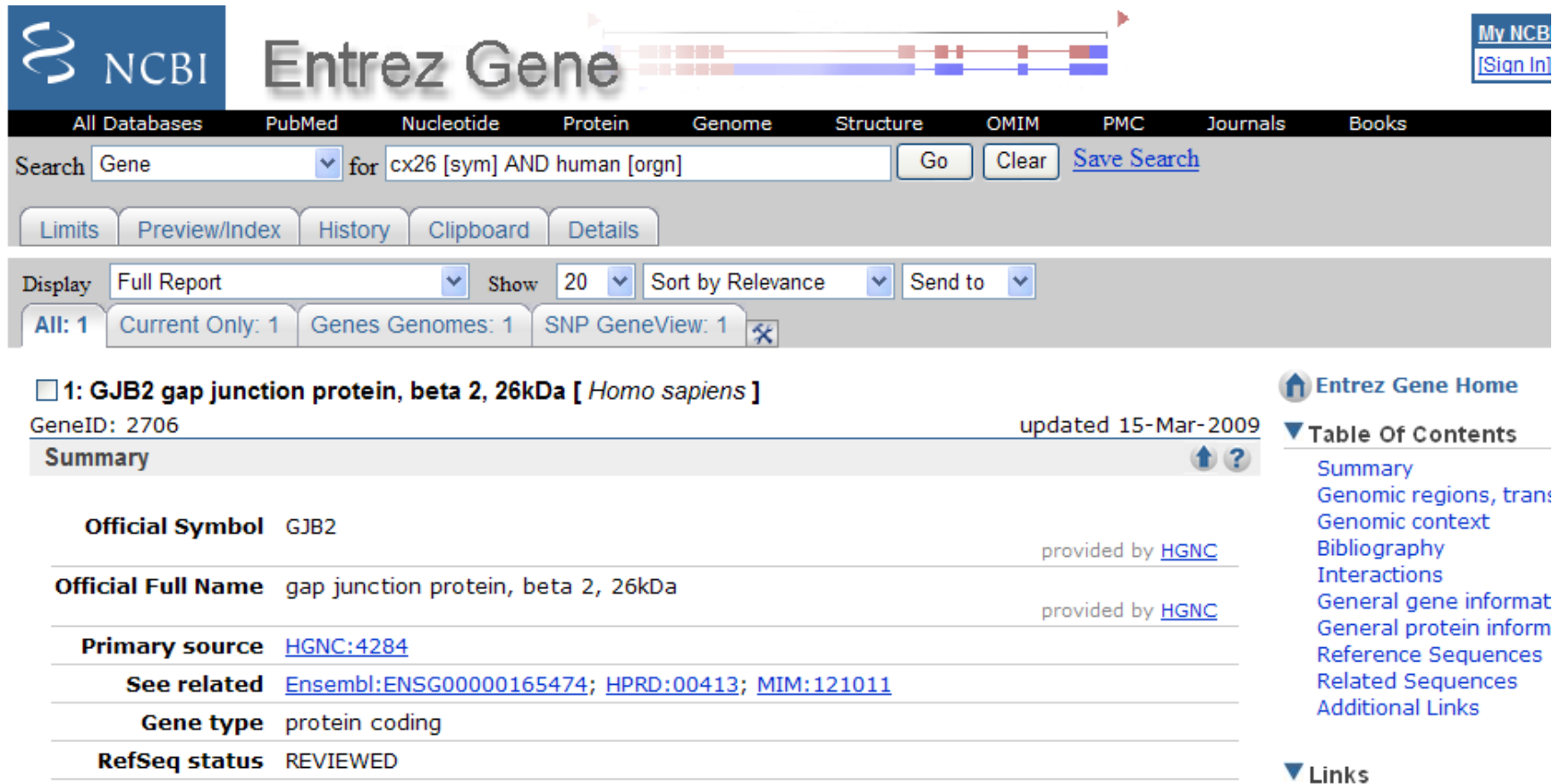
**News** Sort by chromosome; Search by preferred symbol; Property "of

**Sample Searches**

Find genes by...	Search text
free text	<a href="#">human mus</a>
partial name and multiple species	<a href="#">transporter[</a>

What chromosome is your disease gene found on?  
What is the map location for this gene?

Scroll down to the section labeled “Genomic Context.”



NCBI Entrez Gene

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search Gene for cx26 [sym] AND human [orgn] Go Clear Save Search

Limits Preview/Index History Clipboard Details

Display Full Report Show 20 Sort by Relevance Send to

All: 1 Current Only: 1 Genes Genomes: 1 SNP GeneView: 1

1: GJB2 gap junction protein, beta 2, 26kDa [ *Homo sapiens* ]

GeneID: 2706 updated 15-Mar-2009

Summary

<b>Official Symbol</b>	GJB2	provided by <a href="#">HGNC</a>
<b>Official Full Name</b>	gap junction protein, beta 2, 26kDa	provided by <a href="#">HGNC</a>
<b>Primary source</b>	<a href="#">HGNC:4284</a>	
<b>See related</b>	<a href="#">Ensembl:ENSG00000165474</a> ; <a href="#">HPRD:00413</a> ; <a href="#">MIM:121011</a>	
<b>Gene type</b>	protein coding	
<b>RefSeq status</b>	REVIEWED	

Entrez Gene Home

Table Of Contents

- Summary
- Genomic regions, trans
- Genomic context
- Bibliography
- Interactions
- General gene informat
- General protein inform
- Reference Sequences
- Related Sequences
- Additional Links

Links

**Also known as** HID; KID; PPK; CX26; DFNA3; DFNB1; NSRD1; DFNA3A; DFNB1A; GJB2

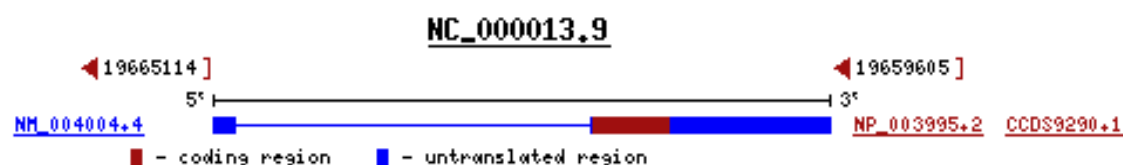
**Summary** This gene encodes a member of the gap junction protein family. The gap junctions were first characterized by electron microscopy as regionally specialized structures on plasma membranes of contacting adherent cells. These structures were shown to consist of cell-to-cell channels that facilitate the transfer of ions and small molecules between cells. The gap junction proteins, also known as connexins, purified from fractions of enriched gap junctions from different tissues differ. According to sequence similarities at the nucleotide and amino acid levels, the gap junction proteins are divided into two categories, alpha and beta. Mutations in this gene are responsible for as much as 50% of pre-lingual, recessive deafness. [provided by RefSeq]

## Genomic regions, transcripts, and products



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

[Try our new Sequence Viewer](#)



## Genomic context



**chromosome: 13; Location: 13q11-q12**

[See GJB2 in MapViewer](#)

How many nucleotides is this gene's coding region?

How many nucleotides are in the gene's transcript?

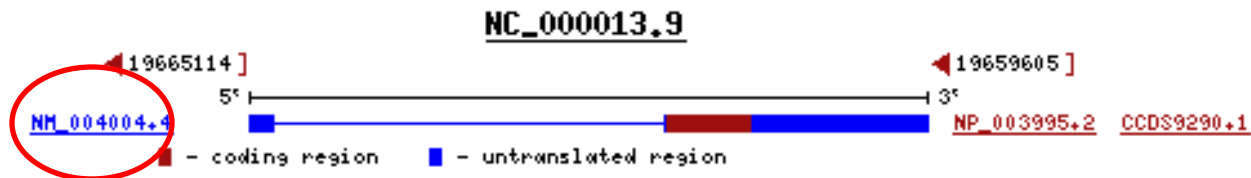
Click on the NM link. Click on GenBank. This takes you to the mRNA record in Gen Bank.

junction proteins are divided into two categories, alpha and beta. Mutations in this gene are responsible for as much as 50% of pre-lingual, recessive deafness.  
[provided by RefSeq]

## Genomic regions, transcripts, and products

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

[Try our new Sequence Viewer](#)



## Genomic context

chromosome: 13; Location: 13q11-q12

[See GJB2 in MapViewer](#)

If your mRNA has a replacement record, click on that link.

1: [NM\\_004004](#). Reports ...[gi:118572604] The record has been replaced by [NM\\_004004.5](#)

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

LOCUS NM\_004004 2331 bp mRNA linear PRI 20-JUL-2008  
DEFINITION Homo sapiens gap junction protein, beta 2, 26kDa (GJB2), mRNA.  
ACCESSION NM\_004004  
VERSION NM\_004004.4 GI:118572604  
KEYWORDS .  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2331)  
AUTHORS Schimmenti,L.A., Martinez,A., Telatar,M., Lai,C.H., Shapiro,N.,  
Fox,M., Warman,B., McCarra,M., Crandall,B., Slinger,Y., Grody,W.W.

This is the GenBank record for the mRNA.

How many nucleotides are in the gene's transcript?

Format: [GenBank](#) [FASTA](#) [Graphics](#) [More Formats](#) [Downl](#)

---

NCBI Reference Sequence: NM\_004004.5

## Homo sapiens gap junction protein, beta 2, 26kDa (GJB2), mRNA

[Ch](#)  
[Cu](#)  
[Pic](#)  
Des  
usir  
[Art](#)  
▶ F  
n  
▶ F  
n  
▶ F  
C

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

LOCUS	NM_004004	2347 bp	mRNA	linear	PRI 13-MAR-2009
DEFINITION	Homo sapiens gap junction protein, beta 2, 26kDa (GJB2), mRNA.				
ACCESSION	NM_004004				
VERSION	NM_004004.5 GI:195539329				
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 2347)				

Now, scroll to the "Features" section of the record. Look for the section labeled "CDS". CDS stands for coding sequence---the region of nucleotides that corresponds with the sequence of amino acids in a protein (location includes start and stop codons).

QUESTION: How many nucleotides is this gene's coding region? Subtract the smaller number from the larger number, so  $896-216 = 680$  nucleotides in the coding region.

Use the back arrow to return to the record in EntrezGene

```
EXON      194..2334
          /gene="GJB2"
          /gene_synonym="CX26; DFNA3; DFNA3A; DFNB1; DFNB1A; HID;
          KID; NSRD1; PPK"
          /inference="alignment:Splign"
          /number=2
CDS       216..896
          /gene="GJB2"
          /gene_synonym="CX26; DFNA3; DFNA3A; DFNB1; DFNB1A; HID;
          KID; NSRD1; PPK"
          /note="connexin 26; gap junction protein, beta 2, 26kD
          (connexin 26)"
          /codon_start=1
          /product="gap junction protein beta 2"
          /protein_id="NP_003995.2"
          /db_xref="GI:42558283"
          /db_xref="CCDS:CCDS9290.1"
          /db_xref="GeneID:2706"
          /db_xref="HGNC:4284"
          /db_xref="HPRD:00413"
          /db_xref="MIM:121011"
          /translation="MDWGTLQITILGGVNHSTSIGKIWLTVLFFRIMILVVAKEVW
          GDEQADFVCNTLQPGCKNVCYDHYFPISHIRLWALQLIFVSTPALLVAMHVAYRRHEK
          KRKFIKGEIKSEFKDIEEIKTQKVRIEGLWWTYTSSIFFRVIFEAAFMYVFVYMYDG
          FSMQRLVKCNAWPCPNTVDCFVSRPTEKTVFTVFEMIAVSGICILLNVTELCYLLIRYC
          SCSQZWDV"
```

How many amino acids are in the gene's peptide? To find the answer, click on the NP link and choose "GENPEPT"

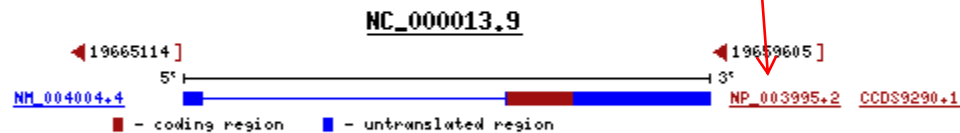
junction proteins are divided into two categories, alpha and beta. Mutations in this gene are responsible for as much as 50% of pre-lingual, recessive deafness. [provided by RefSeq]

### Genomic regions, transcripts, and products



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

[Try our new Sequence Viewer](#)



### Genomic context



chromosome: 13; Location: 13q11-q12

[See GJB2 in MapViewer](#)

This will take you the record for the protein in the Entrez Protein database.

NCBI Sequence Viewer v2.0 - Microsoft Internet Explorer provided by ISU Library

http://www.ncbi.nlm.nih.gov.proxy.lib.iastate.edu:2048/entrez/viewer.fcgi?val=NP\_003995.2&dopt=gp

NCBI Entrez Protein

My NCBI [Sign In] [Registered]

Search Protein for [gap junction prot...] Go Clear

Display GenPept Show 5 Send to

Range: from begin to end Features:  SNP  CDD Refresh

1: NP\_003995. Reports gap junction prot... [gi:42558283] BLink, Conserved Domains, Links

Comment Features Sequence

LOCUS NP\_003995 226 aa linear PRI 13-MAR-2009

DEFINITION gap junction protein beta 2 [Homo sapiens].

ACCESSION NP\_003995

VERSION NP\_003995.2 GI:42558283

DBSOURCE REFSEQ: accession [NM\\_004004.5](#)

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.

REFERENCE 1 (residues 1 to 226)

AUTHORS Joseph, A.Y. and Rasool, T.J.

TITLE High frequency of connexin26 (GJB2) mutations associated with nonsyndromic hearing loss in the population of Kerala, India

JOURNAL Int. J. Pediatr. Otorhinolaryngol. 73 (3), 437-443 (2009)

PUBMED [19157576](#)

REMARK GeneRIF: Observational study of gene-disease association. (HuGE Navigator)

REFERENCE 2 (residues 1 to 226)



How many amino acids are in the gene's peptide?

Double click to change security settings

Done Internet 100%

•What is the function of this gene's peptide?

There are several places in the Entrez GENE record which provide information about the peptide's function.

Summary		 
<b>Official Symbol</b>	GJB2	provided by <a href="#">HGNC</a>
<b>Official Full Name</b>	gap junction protein, beta 2, 26kDa	provided by <a href="#">HGNC</a>
<b>Primary source</b>	<a href="#">HGNC:4284</a>	
<b>See related</b>	<a href="#">Ensembl:ENSG00000165474</a> ; <a href="#">HPRD:00413</a> ; <a href="#">MIM:121011</a>	
<b>Gene type</b>	protein coding	
<b>RefSeq status</b>	REVIEWED	
<b>Organism</b>	<a href="#">Homo sapiens</a>	
<b>Lineage</b>	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>	
<b>Also known as</b>	HID; KID; PPK; CX26; DFNA3; DFNB1; NSRD1; DFNA3A; DFNB1A; GJB2	
<b>Summary</b>	This gene encodes a member of the gap junction protein family. The gap junctions were first characterized by electron microscopy as regionally specialized structures on plasma membranes of contacting adherent cells. These structures were shown to consist of cell-to-cell channels that facilitate the transfer of ions and small molecules between cells. The gap junction proteins, also known as connexins, purified from fractions of enriched gap junctions from different tissues differ. According to sequence similarities at the nucleotide and amino acid levels, the gap junction proteins are divided into two categories, alpha and beta. Mutations in this gene are responsible for as much as 50% of pre-lingual, recessive deafness. [provided by RefSeq]	

The summary may contain information.

KEGG pathway: Cell Communication

[01430](#)

Reactome Event: Gap junction trafficking and regulation

[157858](#)

## Homology

Mouse, Rat

[Map Viewer](#)

More information about function in the GeneOntology section

## GeneOntology

Provided by [GOA](#)

Function	Evidence
<a href="#">gap junction channel activity</a>	IEA
<a href="#">protein binding</a>	IEA

Refers to molecular function of the gene or gene product.

Process	Evidence
<a href="#">cell-cell signaling</a>	TAS <a href="#">PubMed</a>
<a href="#">male genitalia development</a>	IEA
<a href="#">sensory perception of sound</a>	TAS <a href="#">PubMed</a>
<a href="#">transport</a>	TAS <a href="#">PubMed</a>

Refers to biological processes that have been linked to the gene or gene product.

Component	Evidence
<a href="#">Golgi membrane</a>	EXP <a href="#">PubMed</a>
<a href="#">cell junction</a>	IEA
<a href="#">connexon complex</a>	TAS <a href="#">PubMed</a>
<a href="#">integral to membrane</a>	IEA
<a href="#">lateral plasma membrane</a>	IEA
<a href="#">plasma membrane</a>	EXP <a href="#">PubMed</a>
<a href="#">plasma membrane</a>	TAS <a href="#">PubMed</a>

Refers to the location in the cell in which the gene product acts.

Scroll to the very bottom of the page to the “Additional Links” section. There are links to resources that were discussed earlier: the MIM link goes to OMIM, Genes & Disease book, as well as other resources. Both of these resources may give information about the peptide’s function.

#### Additional Links



- MIM [121011](#)
- UMD Locus Specific Databases [GJB2/](#)
- GeneTests for MIM: [121011](#)
- Genes and Disease [deafness.html](#)
- HPRD [00413](#)
- Hereditary Hearing Loss Homepage [Hereditary Hearing Loss Homepage](#)
- The Connexin-deafness homepage [The Connexin-deafness homepage](#)
- UCSC [UCSC](#)
- UniGene [Hs.524894](#)
- UniGene [Hs.714494](#)

Display  Show  Sort by Relevance  Send to

Where in the body is this peptide typically found?

HPRD stands for Human Protein Reference Database. It includes information about where the peptide is expressed in the body.

Unfortunately, the links to HPRD are not working quite right. So, click on the HPRD link below. Also, note the record number 00413.

#### Additional Links



- MIM [121011](#)
- UMD Locus Specific Databases [GJB2/](#)
- GeneTests for MIM: [121011](#)
- Genes and Disease [deafness.html](#)
- HPRD [00413](#)
- Hereditary Hearing Loss Homepage [Hereditary Hearing Loss Homepage](#)
- The Connexin-deafness homepage [The Connexin-deafness homepage](#)
- UCSC [UCSC](#)
- UniGene [Hs.524894](#)
- UniGene [Hs.714494](#)

Display  Show  Sort by Relevance  Send to

Then, click on HPRD. This will take you to the homepage of the database.

The screenshot shows a Windows Internet Explorer browser window with the address bar displaying <http://www.hprd.org/protein/00413>. The page content includes a breadcrumb trail: "You are at: [HPRD](#)". Below this is a "Site Error" message: "An error was encountered while publishing this resource. Resource not found. Sorry, the requested resource does not exist. Check the URL and try again. Resource: <http://www.hprd.org/protein>".

Under the heading "Troubleshooting Suggestions", there are three bullet points:

- The URL may be incorrect.
- The parameters passed to this resource may be incorrect.
- A resource that this resource relies on may be encountering an error.

A link to "Molecule Authority" is provided for more detailed information. The page concludes with a request to send questions to [help](#), a copyright notice for Johns Hopkins University and the Institute of Bioinformatics, and a statement that it is a joint project between two entities, represented by broken image icons.

Click on Query.

The screenshot shows the Human Protein Reference Database (HPRD) website. The browser window title is "Human Protein Reference Database - Windows Internet Explorer" and the address bar shows "http://www.hprd.org/". The page features a navigation sidebar on the left with buttons for "Query", "Browse", "Blast", "FAQs", "Download", "Human Proteinpedia", "Pathways", "PhosphoMotif Finder", and "Become a 'Molecule Authority'". The main content area includes a "News" section with two articles, a "Highlights" section with "PhosphoMotif Finder" and "Pathways" information, and a "Statistics" table. A large "About HPRD" section at the bottom contains a licensing notice and a citation for the database.

**Human Protein Reference Database**

You are at: HPRD

**Query**

**News**

- name: biotechnology **"Human Proteinpedia enables data sharing of human proteins" in February 2008 issue of Nature Biotechnology**
- name: biotechnology **PhosphoMotif Finder, published in February 2007 issue of Nature Biotechnology**
- BMC Bioinformatics **Comparison of Protein-Protein Interaction Databases, published in BMC Bioinformatics**

**Highlights**

**PhosphoMotif Finder**  
Allows you to check if your protein contains any phosphorylation motif described in the literature

**Pathways**  
A set of 20 curated signaling pathways are available as part of a new pathway resource that we have developed called 'NetPath.'

**HPRD Release 7 <sup>New</sup>**  
The latest Release 7 is available for download. [Click here...](#)

**Statistics**

<b>Protein Entries</b>	25,661
<b>Protein-Protein Interactions</b>	38,167
<b>Domains</b>	455
<b>PTMs</b>	16,972
<b>PubMed Links</b>	270,466

**About HPRD**

COMMERCIAL ENTITIES MAY NOT USE THIS SITE WITHOUT PRIOR LICENSING AUTHORIZATION. PLEASE SEND AN [E-MAIL](#) FOR FURTHER INFORMATION ABOUT LICENSING.

The Human Protein Reference Database represents a centralized platform to visually depict and integrate information pertaining to domain architecture, post-translational modifications, interaction networks and disease association for each protein in the human proteome. All the information in HPRD has been manually extracted from the literature by expert biologists who read, interpret and analyze the published data. HPRD has been created using an object oriented database in Zope, an open source web application server, that provides versatility in query functions and allows data to be displayed dynamically.

Please cite the following reference for this database:  
Prasad, T. S. K. *et al.* (2008) Human Protein Reference Database - 2009 Update. *Nucleic Acids Research*. doi:10.1093/nar/gkn892.

Please send any questions or comments about the Human Protein Reference Database to [help](#)

Copyright © Johns Hopkins University and the Institute of Bioinformatics.

This is a joint project between:

Enter the HPRD Identifier number, 00413, and click the Search button.

Human Protein Reference Database - Windows Internet Explorer

http://www.hprd.org/query

File Edit View Favorites Tools Help

Human Protein Reference Database

**Human Protein Reference Database**

You are at HPRD >> Query

**Query**

The default behavior if more than one term is entered within a field is 'AND.' e.g. entering 'SH2 SH3' in 'Domain' search field will search for all the proteins that have both SH2 and SH3 domains. Similarly, if more than one field is filled in, it will be treated as an 'AND' query. For more information go to the [FAQ](#)

Protein Name

Accession Number  RefSeq

HPRD Identifier

Gene Symbol

Chromosome Locus

Molecular Class  [See List](#)

PTMs  [See List](#)

Cellular Component  [See List](#)

Domain Name  [See List](#)

Motif  [See List](#)

Expression  [See List](#)

Length of Protein Sequence From  to  in amino acids

Molecular Weight From  to  in kDa

Diseases

Please send any questions or comments about the Human Protein Reference Database to [help](#)

Copyright © Johns Hopkins University and the Institute of Bioinformatics.

This is a joint project between:

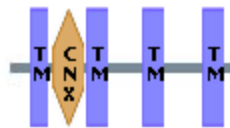
PandeyLab and Institute of Bioinformatics

Error on page. Internet 100%

# Connexin 26

Molecular Class	Membrane transport protein
Molecular Function	Auxiliary transport protein activity
Biological Process	Transport

This protein is expressed in the cochlea, liver, and skin.



## ALTERNATE NAMES

## DISEASES

## PTMs & SUBSTRATES

## SUMMARY

## SEQUENCE

## INTERACTIONS

## EXTERNAL LINKS

### General

Gene Symbol: [GJB2](#)    Molecular Weight (Da): 26275    Gene Map Locus: 13q11-q12

### Localization

Primary: [Plasma membrane](#) [GO](#)    Alternate:

### Domains and Motifs

### Expression

Domains	Motifs
<a href="#">TM</a> 132 - 154	
<a href="#">TM</a> 193 - 215	
<a href="#">TM</a> 21 - 40	
<a href="#">CNX</a> 42 - 75	

Site of Expression
<a href="#">Cochlea</a>
<a href="#">Liver</a>
<a href="#">Skin</a>