hg38 NCBI RefSeq genes, curated subset (NM_*, NR_*, NP_* or YP_*) - NM_001278555.2

RefSeq Gene UBE2E3

RefSeq: NM_001278555.2 **Status:** Reviewed
Description: ubiquitin conjugating enzyme E2 E3, transcript variant 4
Molecule type: mRNA
Source: BestRefSeq
Biotype: protein_coding
Synonyms: UBCH9,UbcM2
OMIM: 604151
Protein: NP_001265484.1
HGNC: 12479
Entrez Gene: 10477
GeneCards: UBE2E3
AceView: UBE2E3

Summary of **UBE2E3**
The modification of proteins with ubiquitin is an important cellular mechanism for targeting abnormal or short-lived proteins for degradation. Ubiquitination involves at least three classes of enzymes: ubiquitin-activating enzymes, or E1s, ubiquitin-conjugating enzymes, or E2s, and ubiquitin-protein ligases, or E3s. This gene encodes a member of the E2 ubiquitin-conjugating enzyme family. The encoded protein shares 100% sequence identity with the mouse and rat counterparts, which indicates that this enzyme is highly conserved in eukaryotes. Multiple alternatively spliced transcript variants encoding the same protein have been found for this gene. [provided by RefSeq, Jun 2013].

---

mRNA/Genomic Alignments (NM_001278555.2)

BROWSER	SIZE	IDENTITY	CHROMOSOME	STRAND	START	END	QUERY	START	END	TOTAL
browser	1373	100.0%	2	+	180980360	181063425	NM_001278555.2	1	1373	1373

View details of parts of alignment within browser window.

Position: chr2:180980360-181063425
Band: 2q31.3
Genomic Size: 83066
Strand: +
Gene Symbol: UBE2E3
CDS Start: complete
CDS End: complete

Links to sequence:

- Predicted Protein
- Predicted mRNA may be different from the genomic sequence.
- Genomic Sequence from assembly

Data schema/format description and download

Go to NCBI RefSeq track controls

Source data version: NCBI RefSeq GCF_000001405.40-RS_2024_08 (2024-08-27)
Data last updated at UCSC: 2024-09-11

Description

The NCBI RefSeq Genes composite track shows human protein-coding and non-protein-coding genes taken from the NCBI RNA reference sequences collection (RefSeq). All subtracks use coordinates provided by RefSeq, except for the *UCSC RefSeq* track, which UCSC produces by realigning the RefSeq RNAs to the genome. This realignment may result in occasional differences between the annotation coordinates provided by UCSC and NCBI. For RNA-seq analysis, we advise using NCBI aligned tables like RefSeq All or RefSeq Curated. See the Methods section for more details about how the different tracks were created.

Please visit NCBI's Feedback for Gene and Reference Sequences (RefSeq) page to make suggestions, submit additions and corrections, or ask for help concerning RefSeq records.

For more information on the different gene tracks, see our Genes FAQ.

Display Conventions and Configuration

This track is a composite track that contains differing data sets. To show only a selected set of subtracks, uncheck the boxes next to the tracks that you wish to hide. **Note:** Not all subtracts are available on all assemblies.

The possible subtracks include:

RefSeq aligned annotations and UCSC alignment of RefSeq annotations

- *RefSeq All* – all curated and predicted annotations provided by RefSeq.
- *RefSeq Curated* – subset of *RefSeq All* that includes only those annotations whose accessions begin with NM, NR, NP or YP. (NP and YP are used only for protein-coding genes on the mitochondrion; YP is used for human only.)
- *RefSeq Predicted* – subset of RefSeq All that includes those annotations whose accessions begin with XM or XR.
- *RefSeq Other* – all other annotations produced by the RefSeq group that do not fit the requirements for inclusion in the *RefSeq Curated* or the *RefSeq Predicted* tracks, as they do not have a product and therefore no RefSeq accession. More than 90% are pseudogenes, T-cell receptor or immunoglobulin segments. The few remaining entries are gene clusters (e.g. protocadherin).
- *RefSeq Alignments* – alignments of RefSeq RNAs to the human genome provided by the RefSeq group, following the display conventions for PSL tracks.
- *RefSeq Diffs* – alignment differences between the human reference genome(s) and RefSeq transcripts. (Track not currently available for every assembly.)
- *UCSC RefSeq* – annotations generated from UCSC's realignment of RNAs with NM and NR accessions to the human genome. This track was previously known as the "RefSeq Genes" track.
- *RefSeq Select+MANE (subset)* – Subset of RefSeq Curated, transcripts marked as RefSeq Select or MANE Select. A single *Select* transcript is chosen as representative for each protein-coding gene. This

track includes transcripts categorized as MANE, which are further agreed upon as representative by both NCBI RefSeq and Ensembl/GENCODE, and have a 100% identical match to a transcript in the Ensembl annotation. See NCBI RefSeq Select. Note that we provide a separate track, MANE (hg38), which contains only the MANE transcripts.

- *RefSeq HGMD (subset)* – Subset of RefSeq Curated, transcripts annotated by the Human Gene Mutation Database. This track is only available on the human genomes hg19 and hg38. It is the most restricted RefSeq subset, targeting clinical diagnostics.

The *RefSeq All*, *RefSeq Curated*, *RefSeq Predicted*, *RefSeq HGMD*, *RefSeq Select/MANE* and *UCSC RefSeq* tracks follow the display conventions for gene prediction tracks. The color shading indicates the level of review the RefSeq record has undergone: predicted (light), provisional (medium), or reviewed (dark), as defined by RefSeq.

Color	Level of review
	Reviewed: the RefSeq record has been reviewed by NCBI staff or by a collaborator. The NCBI review process includes assessing available sequence data and the literature. Some RefSeq records may incorporate expanded sequence and annotation information.
	Provisional: the RefSeq record has not yet been subject to individual review. The initial sequence-to-gene association has been established by outside collaborators or NCBI staff.
	Predicted: the RefSeq record has not yet been subject to individual review, and some aspect of the RefSeq record is predicted.

The item labels and codon display properties for features within this track can be configured through the check-box controls at the top of the track description page. To adjust the settings for an individual subtrack, click the wrench icon next to the track name in the subtrack list .

- **Label:** By default, items are labeled by gene name. Click the appropriate Label option to display the accession name or OMIM identifier instead of the gene name, show all or a subset of these labels including the gene name, OMIM identifier and accession names, or turn off the label completely.
- **Codon coloring:** This track has an optional codon coloring feature that allows users to quickly validate and compare gene predictions. To display codon colors, select the *genomic codons* option from the *Color track by codons* pull-down menu. For more information about this feature, go to the Coloring Gene Predictions and Annotations by Codon page.

The *RefSeq Diffs* track contains five different types of inconsistency between the reference genome sequence and the RefSeq transcript sequences. The five types of differences are as follows:

- *mismatch* – aligned but mismatching bases, plus HGVS g. to show the genomic change required to match the transcript and HGVS c./n. to show the transcript change required to match the genome.
- *short gap* – genomic gaps that are too small to be introns (arbitrary cutoff of < 45 bp), most likely insertions/deletion variants or errors, with HGVS g. and c./n. showing differences.
- *shift gap* – shortGap items whose placement could be shifted left and/or right on the genome due to repetitive sequence, with HGVS c./n. position range of ambiguous region in transcript. Here, thin and thick lines are used -- the thin line shows the span of the repetitive sequence, and the thick line shows the rightmost shifted gap.
- *double gap* – genomic gaps that are long enough to be introns but that skip over transcript sequence (invisible in default setting), with HGVS c./n. deletion.
- *skipped* – sequence at the beginning or end of a transcript that is not aligned to the genome (invisible in default setting), with HGVS c./n. deletion

HGVS Terminology (Human Genome Variation Society): g. = genomic sequence ; c. = coding DNA sequence ; n. = non-coding RNA reference sequence.

When reporting HGVS with RefSeq sequences, to make sure that results from research articles can be mapped to the genome unambiguously, please specify the RefSeq annotation release displayed on the transcript's Genome Browser details page and also the RefSeq transcript ID with version (e.g. NM_012309.4 not NM_012309).

Methods

Tracks contained in the RefSeq annotation and RefSeq RNA alignment tracks were created at UCSC using data from the NCBI RefSeq project. Data files were downloaded from RefSeq in GFF file format and converted to the genePred and PSL table formats for display in the Genome Browser. Information about the NCBI annotation pipeline can be found here.

The RefSeq Diffs track is generated by UCSC using NCBI's RefSeq RNA alignments.

The UCSC RefSeq Genes track is constructed using the same methods as previous RefSeq Genes tracks. RefSeq RNAs were aligned against the human genome using BLAT. Those with an alignment of less than 15% were discarded. When a single RNA aligned in multiple places, the alignment having the highest base identity was identified. Only alignments having a base identity level within 0.1% of the best and at least 96% base identity with the genomic sequence were kept.

Data Access

The raw data for these tracks can be accessed in multiple ways. It can be explored interactively using the REST API, Table Browser or Data Integrator. The tables can also be accessed programmatically through our public MySQL server or downloaded from our downloads server for local processing. The previous track versions are available in the archives of our downloads server. You can also access any RefSeq table entries in JSON format through our JSON API.

The data in the *RefSeq Other* and *RefSeq Diffs* tracks are organized in bigBed file format; more information about accessing the information in this bigBed file can be found below. The other subtracks are associated with database tables as follows:

genePred format:

- RefSeq All - `ncbiRefSeq`
- RefSeq Curated - `ncbiRefSeqCurated`
- RefSeq Predicted - `ncbiRefSeqPredicted`
- RefSeq HGMD - `ncbiRefSeqHgmd`
- RefSeq Select+MANE - `ncbiRefSeqSelect`
- UCSC RefSeq - `refGene`

PSL format:

- RefSeq Alignments - `ncbiRefSeqPsl`

The first column of each of these tables is "bin". This column is designed to speed up access for display in the Genome Browser, but can be safely ignored in downstream analysis. You can read more about the bin indexing system here.

The annotations in the *RefSeqOther* and *RefSeqDiffs* tracks are stored in bigBed files, which can be obtained from our downloads server here, `ncbiRefSeqOther.bb` and `ncbiRefSeqDiffs.bb`. Individual regions or the whole set of genome-wide annotations can be obtained using our tool `bigBedToBed` which can be compiled from the source code or downloaded as a precompiled binary for your system from the utilities directory linked below. For example, to extract only annotations in a given region, you could use the following command:

```
bigBedToBed http://hgdownload.soe.ucsc.edu/gbdb/hg38/ncbiRefSeq/ncbiRefSeqOther.bb -chrom=chr16 -
start=34990190 -end=36727467 stdout
```

You can download a GTF format version of the RefSeq All table from the GTF downloads directory. The genePred format tracks can also be converted to GTF format using the `genePredToGtf` utility, available from the utilities directory on the UCSC downloads server. The utility can be run from the command line like so:

```
genePredToGtf hg38 ncbiRefSeqPredicted ncbiRefSeqPredicted.gtf
```

Note that using genePredToGtf in this manner accesses our public MySQL server, and you therefore must set up your hg.conf as described on the MySQL page linked near the beginning of the Data Access section.

A file containing the RNA sequences in FASTA format for all items in the *RefSeq All*, *RefSeq Curated*, and *RefSeq Predicted* tracks can be found on our downloads server here.

Please refer to our mailing list archives for questions.

Previous versions of the ncbiRefSeq set of tracks can be found on our archive download server.

Credits

This track was produced at UCSC from data generated by scientists worldwide and curated by the NCBI RefSeq project.

References

Kent WJ. BLAT - the BLAST-like alignment tool. *Genome Res.* 2002 Apr;12(4):656-64. PMID: 11932250; PMC: PMC187518

Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, Farrell CM, Hart J, Landrum MJ, McGarvey KM *et al.* RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* 2014 Jan;42(Database issue):D756-63. PMID: 24259432; PMC: PMC3965018

Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 2005 Jan 1;33(Database issue):D501-4. PMID: 15608248; PMC: PMC539979