

Schema for UCSC Genes - UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

Database: mm9 **Primary Table:** knownGene **Row Count:** 55,419

Format description: Genes based on RefSeq, GenBank, and UniProt.

field	example	SQL type	info	description
name	uc007aet.1	varchar(255)	values	Name of gene
chrom	chr1	varchar(255)	values	Reference sequence chromosome or scaffold
strand	-	char(1)	values	+ or - for strand
txStart	3195984	int(10) unsigned	range	Transcription start position
txEnd	3205713	int(10) unsigned	range	Transcription end position
cdsStart	3195984	int(10) unsigned	range	Coding region start
cdsEnd	3195984	int(10) unsigned	range	Coding region end
exonCount	2	int(10) unsigned	range	Number of exons
exonStarts	3195984,3203519,	longblob		Exon start positions
exonEnds	3197398,3205713,	longblob		Exon end positions
proteinID		varchar(40)	values	UniProt display ID for Known Genes, UniProt accession or RefSeq protein ID for UCSC Genes
alignID	uc007aet.1	varchar(255)	values	Unique identifier for each (known gene, alignment position) pair