WEEK 3 ASSIGNMENT

Introduction to Computational Biology – BIOL 509000 | Fall 1 2020

Christina Morgenstern

Question #1: On what chromosome is the Ccnd1 gene located? Identify the span of bases where the gene is located on this chromosome, reporting as "ChrX: Y-Z."

The gene *Homo sapiens* cyclin D1 (CCND1) is located on chromosome 11: chr11:69,455,873-69,469,242

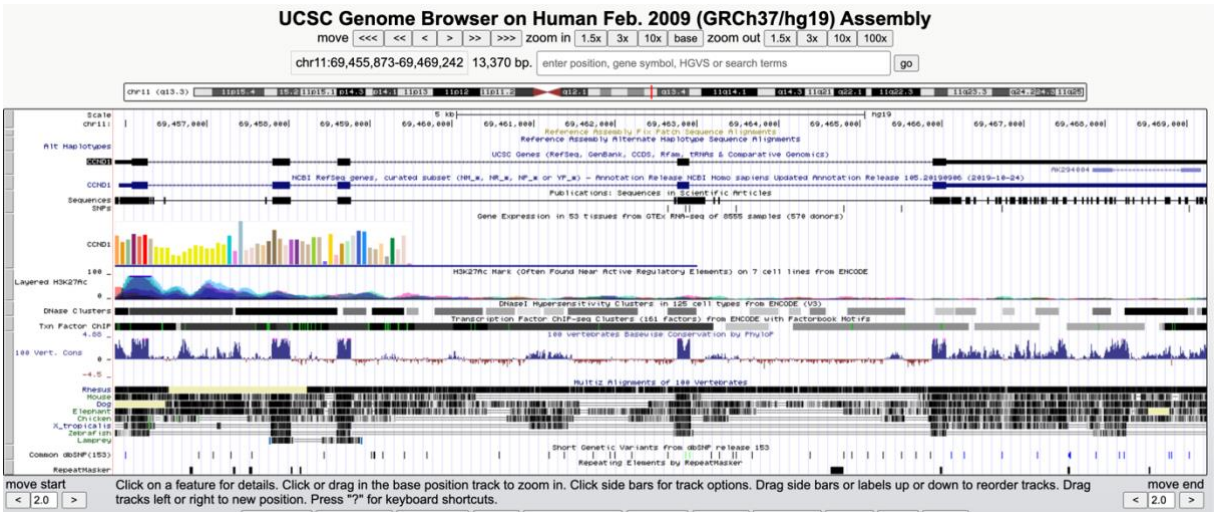


Figure 1. Result for CCND1 in UCSC Genome Browser.

Question #2: How many exons are reported in the "NCBI RefSeq Genes" bar for Ccnd1? How many introns?

There are 5 exons and 4 introns reported for CCND1.

Question #3: Is the Ccnd1 gene encoded in the plus or minus direction?

The CCND1 gene is encoded on the plus (+) strand.

Question #4: Before you hide the "UCSC Genes" track as prompted in the tutorial video, how many isoforms are listed under "UCSC Genes" for Ccnd3? How many are listed under "NCBI RefSeq"?

There are 7 isoforms for CCND3 listed under the UCSC Genes, and 6 isoforms under the NCBI RefSeq.

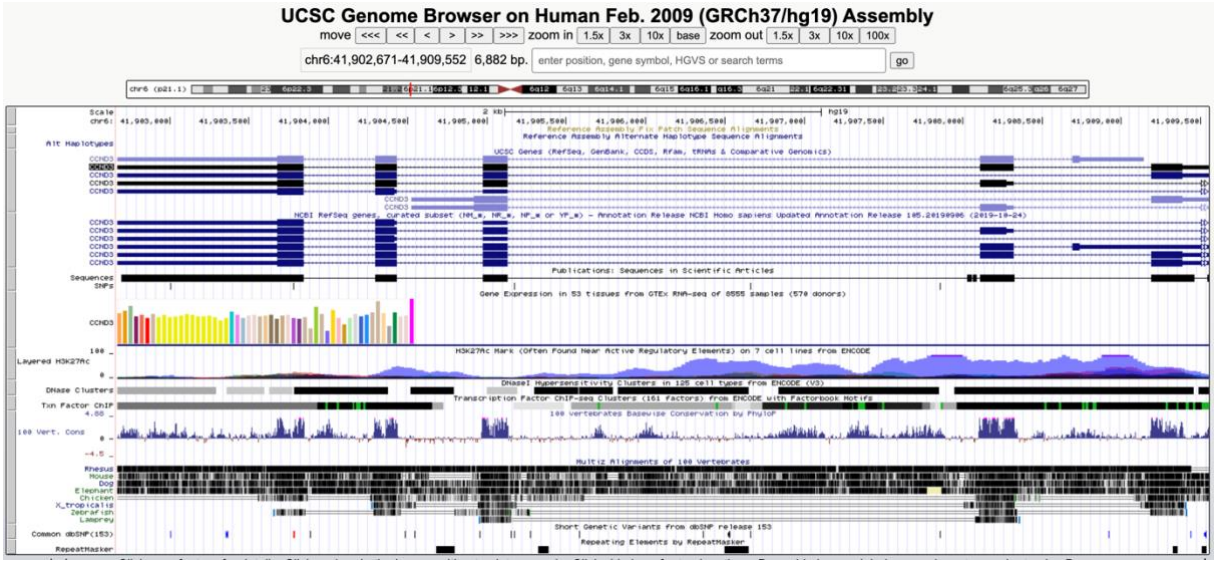


Figure 2. Result for CCND3 in UCSC Genome Browser.

Question #5: Is the Ccnd3 gene encoded in the plus or minus direction?

The CCND3 gene is encoded on the minus (-) strand.

Question #6: Consider the "100 Vert. Cons" track. Where do you see similarity spikes located? Are they limited exclusively to exon regions? Describe your findings.

The spikes mainly correspond to the exons denoting highly conserved sequences. But there are also spikes in the introns (mainly introns 1 and 3) albeit smaller ones. These intronic conserved regions might also be biologically relevant.

Question #7: Consider the "Multiz Alignments" track. The listed organisms have varying degrees of conservation in this gene when compared to humans. What animals have the top four hits (most conserved regions) when their Ccnd3 gene is compared to the human sequence? Which organism's nucleotide sequence is the least conserved?

Rhesus monkey, mouse, dog and elephant show a high degree of conservation with the human CCND3 gene. The least conserved gene when compared to the human variant is the lamprey CCND3 (see Fig. 2 Multiz Alignments of 100 vertebrates track).

Question #8: For the rat (Rattus norvegicus) alignment, on what chromosome is the alignment region found?

The alignment region is found on chromosome 9.

Question #9: The orthologue summary table reports the number of species in each grouping that is considered 1:1 (where there is only one copy of the gene found in each species). For the rodent category, how many are reported as 1:1?

All listed 7 rodents report a 1-to-1 type of CCND3 orthologue.

Show All entries | Show/hide columns | Filter

Species	Type	dN/dS	Ensembl identifier & gene name	Compare	Location	Target %id	Query %id
Guinea Pig (*Cavia porcellus*)	1-to-1	0.11806	ENSCPOG00000025856 CCND3 cyclin D3 [Source:HGNC Symbol;Acc:1585]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	scaffold_15:32127301-32131465:1	89	74
Kangaroo rat (*Dipodomys ordii*)	1-to-1	n/a	ENSDORG00000010969 Ccnd3 cyclin D3 Gene [Source:MGI Symbol;Acc:MGI:88315]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	GeneScaffold_2419:28308-33715:-1	95	95
Mouse (*Mus musculus*)	1-to-1	0.05837	ENSMUSG00000034165 Ccnd3 cyclin D3 [Source:MGI Symbol;Acc:MGI:88315]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	17:47642000-47736632:1	95	95
Pika (*Ochotona princeps*)	1-to-1	n/a	ENSOPRG00000015849 CCND3 cyclin D3 [Source:HGNC Symbol;Acc:1585]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	GeneScaffold_1730:62052-84458:-1	96	96
Rabbit (*Oryctolagus cuniculus*)	1-to-1	0.06207	ENSOCUG00000017573 Q0GMK7_RABIT Cyclin D3 [Source: UniProtKB/TrEMBL; acc: Q0GMK7]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	12:31536951-31635927:-1	97	75
Squirrel (*Spermophilus tridecemlineatus*)	1-to-1	0.05659	ENSSTOG00000021553 CCND3 No description	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	JH393286.1:17880077-17885855:1	96	96
Tree Shrew (*Tupaia belangeri*)	1-to-1	n/a	ENSTBEG00000009546 CCND3 cyclin D3 [Source:HGNC Symbol;Acc:1585]	Region Comparison, Alignment (protein), Alignment (cDNA), Gene Tree (image)	GeneScaffold_1972:39883-45223:-1	96	76

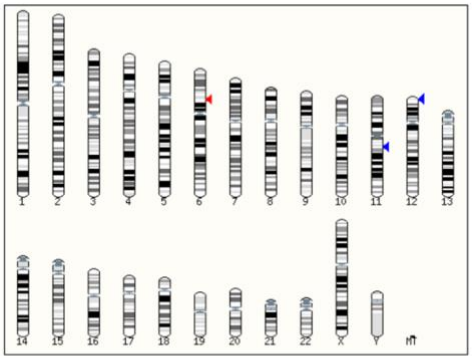
Figure 3. CCND3 orthologues in rodents.

Question #10: The first listed Family ID contains 3 genes. What is the consensus annotation for this family? What are the three genes included in this family? Provide the gene names in this family.

In humans, three genes belong to this family that is of the G1/S specific cyclin consensus annotation. The three genes are CCND3, CCND1 and CCND2 (see Fig. 4).

HUMAN genes in this family

Ensembl genes containing proteins in family ENSFM00250000001580



Gene ID and Location	Gene Name	Description(if known)
ENSG00000112576 Chromosome 6: 41.90m	CCND3	cyclin D3 [Source:HGNC Symbol;Acc:1585]
ENSG00000110092 Chromosome 11: 69.46m	CCND1	cyclin D1 [Source:HGNC Symbol;Acc:1582]
ENSG00000118971 Chromosome 12: 4.38m	CCND2	cyclin D2 [Source:HGNC Symbol;Acc:1583]

Figure 4. CCND3 protein family members.