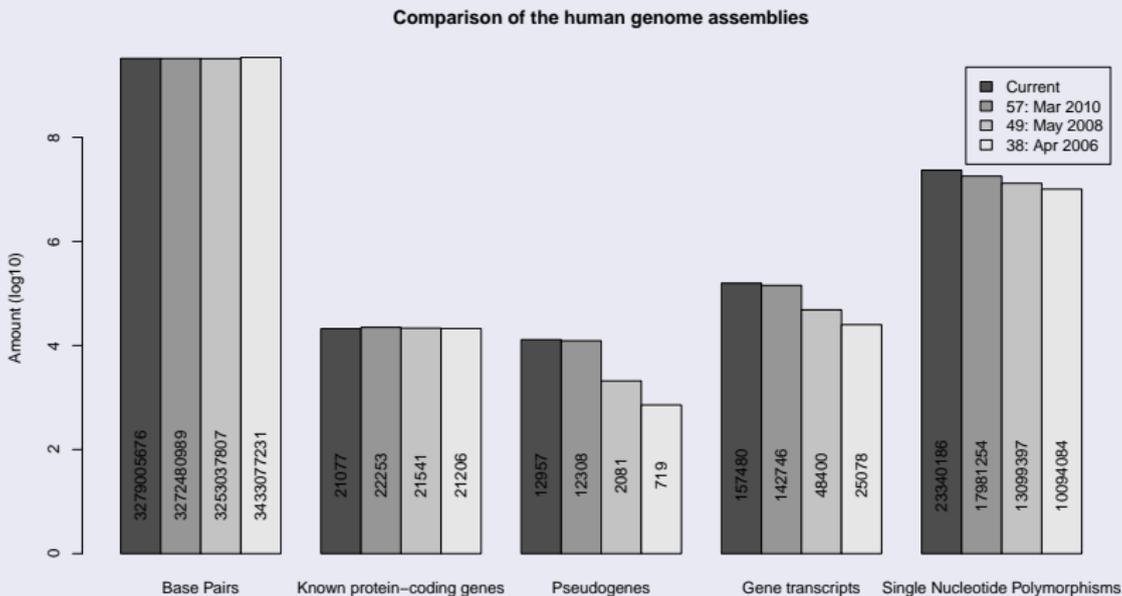


Exercise 26: Web Hunt

Andreas Loibl

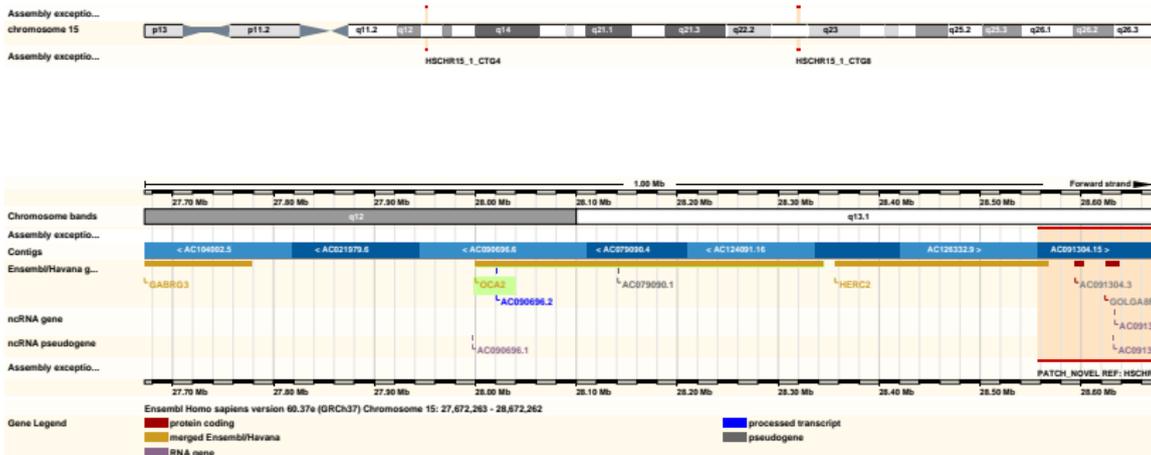
November 17, 2010

- a) Plot a bar graph to compare [several amounts] between the human genome assemblies
- b) EnSEMBL 38: Apr 2006; 49: Mar 2008; 57: Mar 2010



c) Search for the OCA2 gene locus

The OCA2 gene is located at chromosome 15 at 28,000,021bp-28,344,504bp.



d) At which position is the next pseudogene?

The next pseudogene after the OCA2 gene is located at
28,140,850bp - 28,141,728bp

Location-based displays

- Whole genome
- Chromosome summary
- Region overview
- Region in detail**
- Comparative Genomics
 - Alignments (image) (52)
 - Alignments (text) (52)
 - Multi-species view (48)
 - Synteny (15)
- Genetic Variation
 - Resequencing (2)
 - Linkage Data
- Markers
- Other genome browsers
 - UCSC
 - NCBI
 - Vega

Configure this page

Manage your data

Export data

Bookmark this page

Chromosome 15: 27,999,039-28,016,186

Assembly exceptio... chromosome 15

Assembly exceptio... HSCR15_1_CT64 HSCR15_1_CT68

[Export image](#)

Region in detail [help](#) [Alignments \(image\)](#)

Chromosome bands

Contigs

Ensembl/Havana g...

ncRNA gene

ncRNA pseudogene

Gene Legend

Gene-based (Vega): AC079090.1

Gene ENS00000221992

Location Chromosome 15
28,140,850-28,141,728

Gene type Known pseudogene

Strand Forward

Analysis Havana gene

Manually annotated transcripts (etermined on a case-by-case basis) from the Havana project.

Location: 15:27999039-28016186

Gene:

- ▼ Gene tree (images)
- └ Gene Tree (text)
- └ Gene Tree (alignment)
- └ Orthologues
- └ Paralogs
- └ Protein families
- Genetic Variation
- └ Variation Table
- └ Variation Image
- External Data
- └ Personal annotation
- ID History
- └ Gene history

Configure this page

Manage your data

Export data

Bookmark this page

Name AC079090.1 (Clone-based (Vega))

Gene type Known pseudogene

Prediction Method Manually annotated transcripts (determined on a case-by-case basis) from the [Havana](#) project.

Alternative genes This gene corresponds to the following database identifiers:
Havana gene: [OTTHUMG00000128936](#) [\[view all locations\]](#)



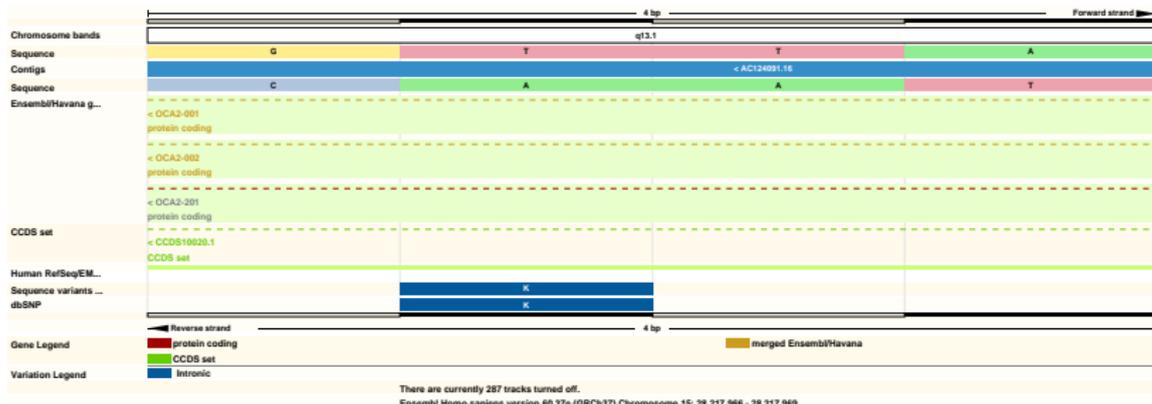
Export image

Configuring the display

Tip: use the 'Configure this page' link on the left to show additional data in this region.

e) A SNP is annotated in position 28217967bp. Which nucleotides are given for the different alleles?

The SNP at 28,217,967bp has the ambiguity code “K”, which means that its alleles are “T” or “G”.



SNPs (Single Nucleotide Polymorphisms)

4 bp Forward strand

Chromosome bands: q13.1

Sequence: G T T A

Contigs: < AC124091.16

Sequence: C A A T

Ensembl/Havana genes: < OCA2-001 protein coding, < OCA2-002 protein coding, < OCA2-201 protein coding

CCDS set: < CCDS10020.1 CCDS set

Human RefSeq/EMBL: Variation: rs59410734

dbSNP: Variation Properties

Gene Legend: protein coding (red), CCDS set (green), intronic (blue)

Variation Legend: There are current Ensembl Homo

Configuring the display

You currently have 127 tracks in the overview panel and 287 tracks in the main panel on the left.

Alleles: T/G
source: dbSNP
Type: INTRONIC

Export image

Ensembl release 90 - Nov 2010 © WTSI / FRI

About Ensembl | Contact Us | Help

Ensembl
[Login](#) | [Register](#) | [BI & ST/BI & T](#) | [BioMed](#) | [Tools](#) | [Downloads](#) | [Help](#) | [Documentation](#) | [Mirrors](#)

Human (GRCh37) | Location: 15:28,217,966-28,217,969 | Gene: OCA2 | Variation: rs59410734

Variation displays

- Summary
- Gene/Transcript (3)
- Population genetics
- Individual genotypes
- Context
- Linked variations
- Phenotype Data
- Phylogenetic Context (4)
- External Data

Configure this page | Manage your data | Export data | Bookmark this page

Variation: rs59410734

Variation class SNP ([rs59410734](#) source [dbSNP_131](#) - Variants (including SNPs and indels) imported from dbSNP)

Synonyms None currently in the database

Alleles T/G (Ambiguity code: K)

Location This feature maps to 15:28217967 (forward strand) | [View in location tab](#)

Variation summary [help](#) [Gene/Transcript >](#)

Validation status Unknown

HGVS names [ENST00000354638.3](#) c.1504-5999A>C
[ENST00000353809.5](#) c.1432-5999A>C
[ENST00000382996.2](#) c.1504-5999A>C

Linkage disequilibrium data No linkage data for this variant

Flanking Sequence (reference and dbSNP)

```
CCACAGTATTTCCATAAGGGTTGTATTAATTTACATTCAAACAGCAGTGATAAGCAT
TCCCTTTTACCACATCTGTGCCAACATCTATGGTTTTTGACTTTTAAATAATGGCCAT
AATGGCCATTCTGGCTGTAATAAGGTGATATCTCATTCCTTTAAATTTGATTTCCGGATG
ATTAGTAGTGTTCAGCATTTTTCATATGTTCTGTGGCAGTGGCATATCTCTCCAAAA
ATGCTACTCATGTCATTTGCCTACTTTTATTGGGATATTTTTTTTCTTGCATATT
TGTTCAGTTCCTATAGATCTGGATATTAGTCTTTGTGAGTTCATAGTTGCAAAC
ATTGCTCCCATCTGTAGGTTTTCTGTTTACTTTGATGGKTAATTTTATTTTATTTTAT
TTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTGACAGAGTCTCTCT
TCTCACCAAGCTGGAGTGCAGTGGTCAATCTCGGCTCACTGCAACCTCTGCCTCCCTG
GTTCAAGTGATCTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAAGCATGCAACACC
ACGCCAGCTATTTTTTATTTATTTATTTAGTAGAGATGGGGTTTCACTGCTTTGGCCAG
GCTGGTCTCTAATCTGACCTGAAATGATCTCCACCTTGGCTCCCAAAGTGGTGGG
ATTACAGGCTGAGCCACCATGCCAGCTGCATATGCTTTTCAGTTAATTAGATCCCA
TTATTTTGTTCGACGCA
```

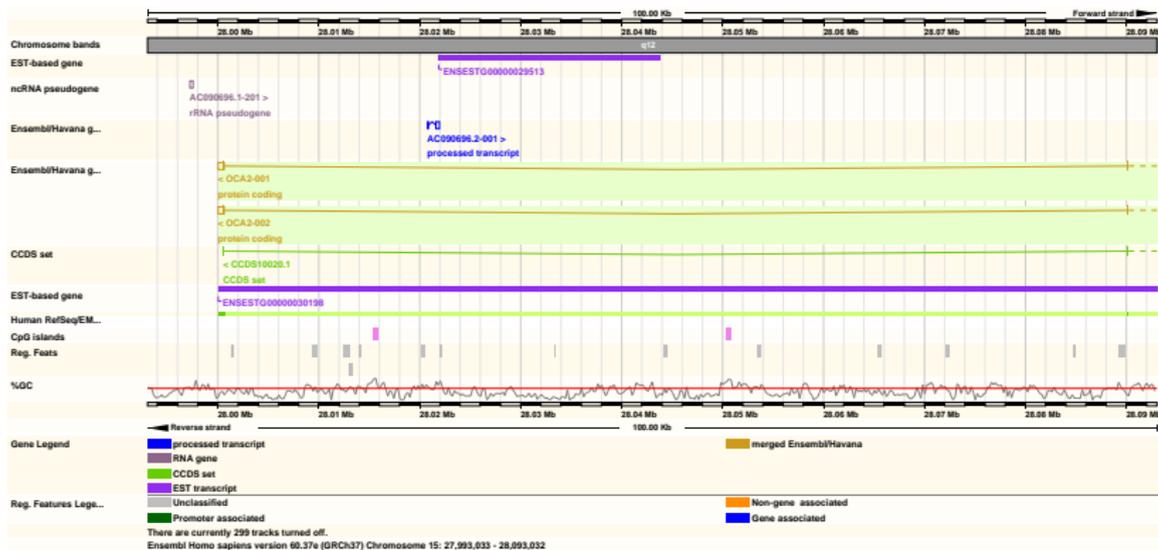
(Variant highlighted)

Ensembl release 60 - Nov 2010 ([WTS](#) / [EBI](#)) [About Ensembl](#) | [Contact Us](#) | [Help](#)

[Permanent link](#) - [View in archive site](#)

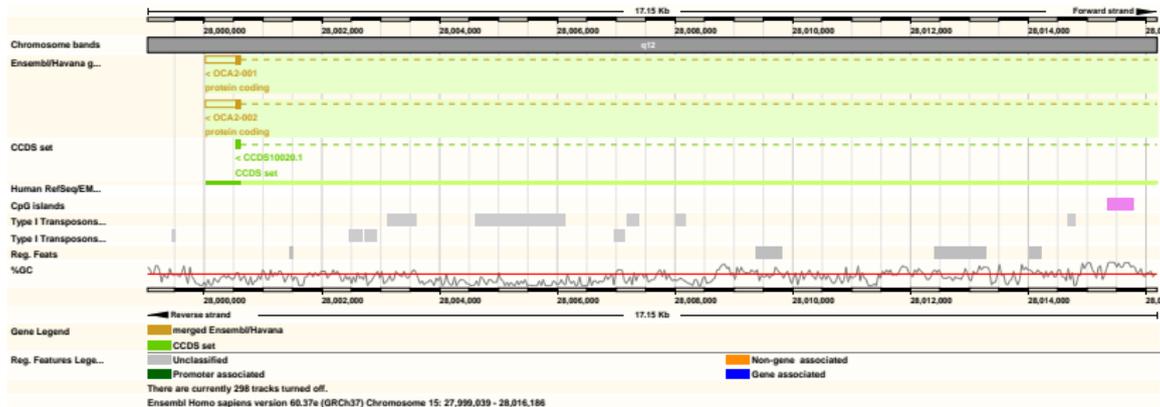
f) There are 7 GpC islands reported within the OCA2 gene. How many ESTs are annotated within the region between CpG islands 1 & 2?

There is only **one** EST between CpG island 1 & 2:
ENSESTG00000029513



g) How many LINES are annotated within the gene and upstream of the first CpG island?

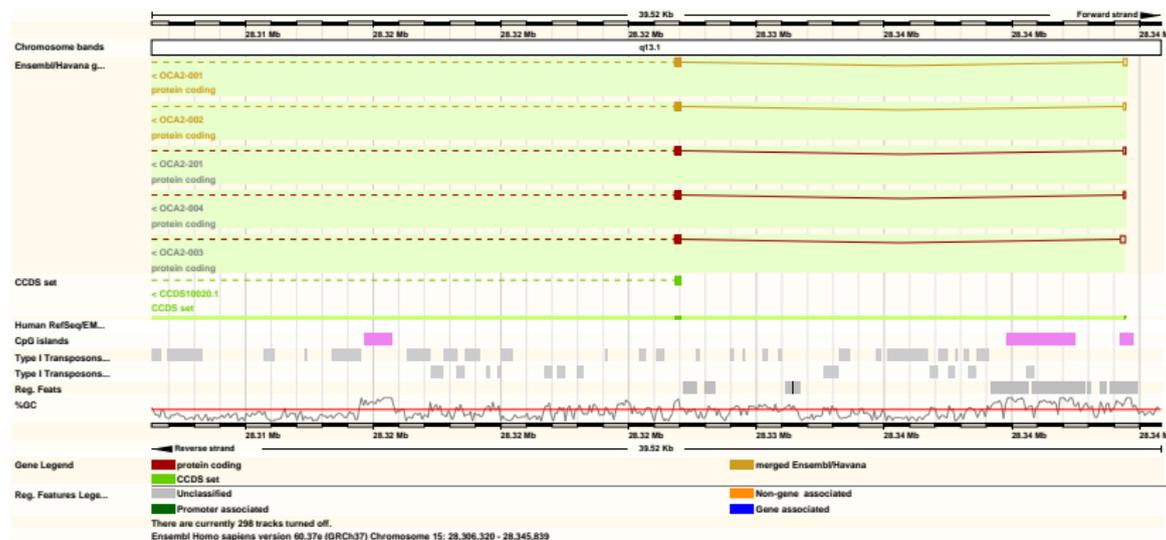
There are **five** LINES annotated in the given range.



Note: The first row starting with “Type 1 Transposons...” are LINES, the second SINES.

h) How many SINES are annotated within the region between the CpG islands 5 & 6?

There are **13** SINES annotated in the given range.



Note: The first row starting with "Type 1 Transposons..." are LINES, the second SINES.