Bioinformatica II

LM Biologia Evoluzionistica, Università di Padova

Docenti: Dr. Giorgio Valle, Dr. Stefania Bortoluzzi

Esercitazione 1

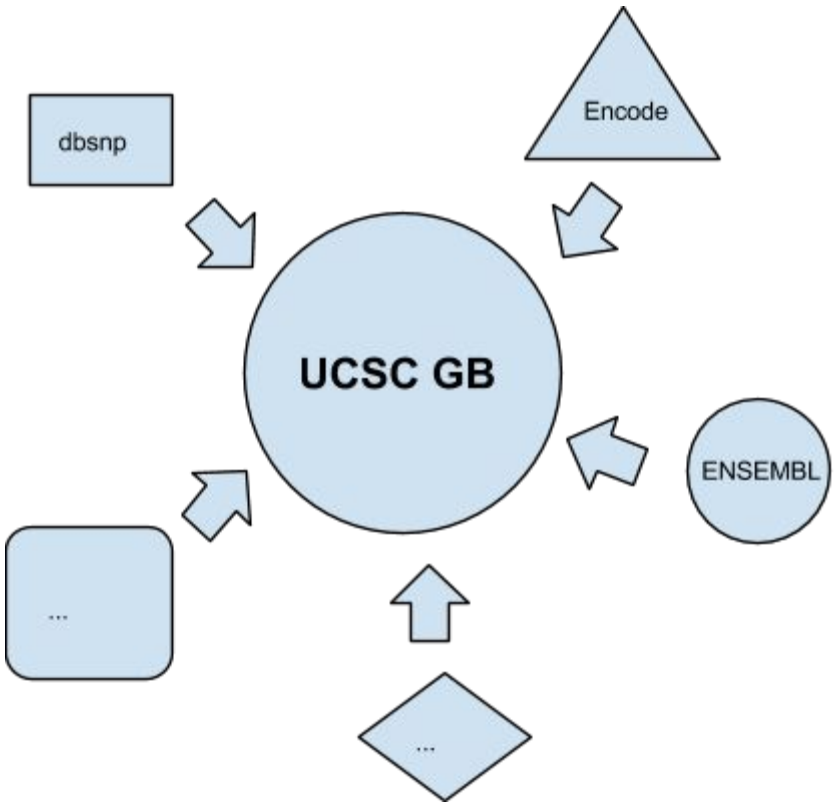
Padova, 10 novembre 2015

GUIDA

Introduzione all' UCSC Genome Browser

Obbiettivo dell'esercitazione

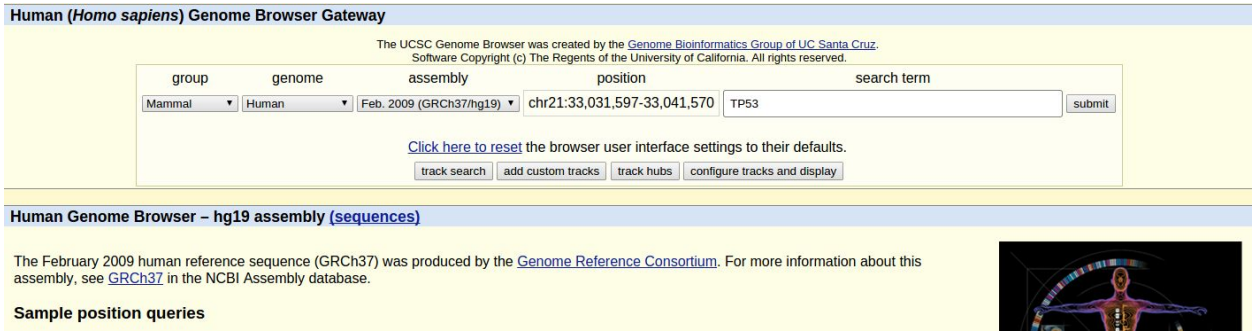
Capire il funzionamento e l'utilità di un genome browser attraverso l'utilizzo dell' UCSC Genome Browser (http://genome.ucsc.edu/)



Ricerche di base

1 - Cliccare su **Genome Browser** nella barra laterale oppure **Genomes** nella barra superiore.

2 - Cercare il gene tp53 (umano) utilizzando la versione di febbraio 2009 dell'assemblaggio



3 - Nella pagina dei risultati della ricerca cliccare la entry "TP53 (uc002gij.3) at chr17:7571720-7590868"

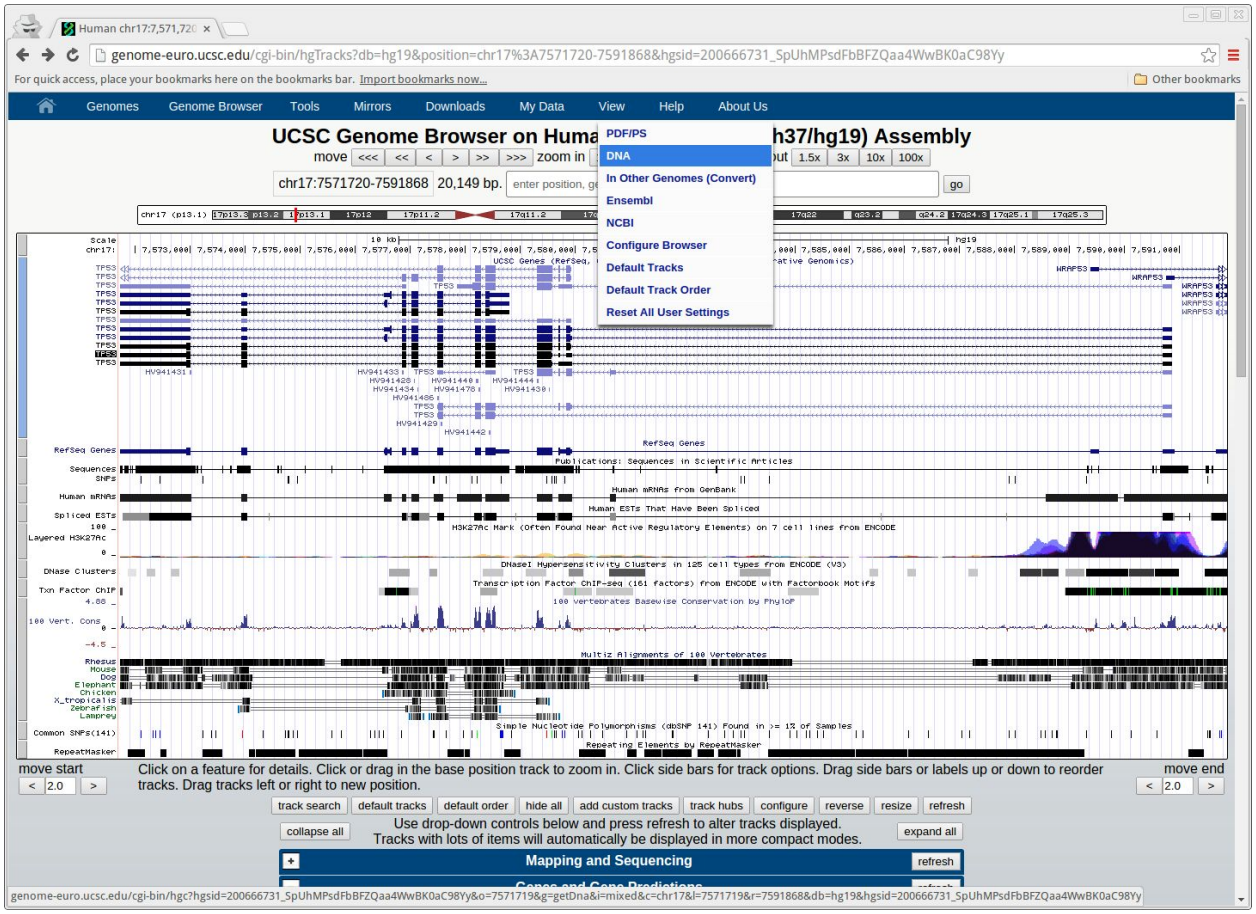
4 - Includere 1000 basi a valle (promotore putativo) del trascritto scelto: "chr17:7,571,720-7,591,868"

5 - Provare a cambiare l'ordine con cui le track vengono visualizzate nel "viewer"

6 - Cambiare i diversi livelli di visualizzazione della track Spliced ESTs

7 - Visualizzare la pagina di dettagli dell'isoforma da noi scelta cliccando sulla riga con il corrispondente trascritto.

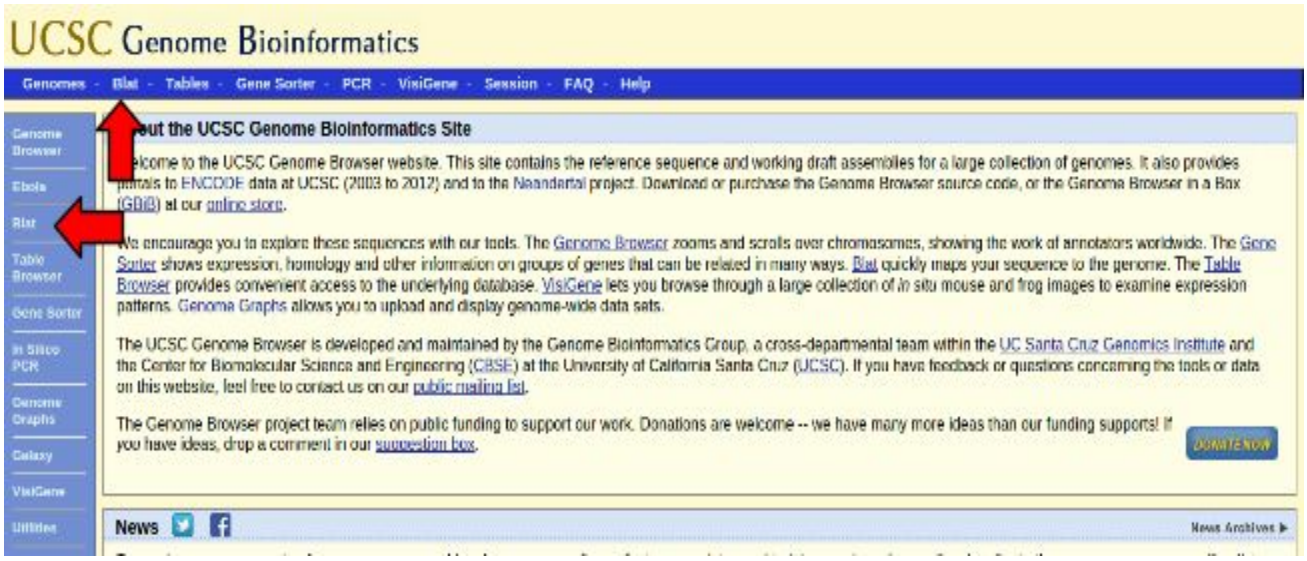
8 - Visualizzare la Sequenza di DNA corrispondente alla regione nel "viewer".



9 - A partire da Extended DNA Case/Color Options scegliere di visualizzare UCSC Genes in rosso (254) e Spliced ESTs in verde (254). Cosa rappresentano le regioni colorate in giallo?

10 - A partire dalla pagina di dettagli dell'isoforma scelta (punto 7) ottenere la sequenza proteica in formato FASTA.

Ricerche a partire da dati di sequenza



1 - Usare BLAT per fare una ricerca di similarità a partire dalle seguenti 2 sequenze:

```
>Seq1
gatgggattggggttttctttaaacgtgtgctcaagactggcgctaaaagttttgagcttctcaaaagtctagag
ccaccgtccagggagcaggtagctgctgggctccggggacactttgcgttcgggctgggagcgtgctttccacga
cggtgacacgcttccctggattggcagccagactgccttccgggtcactgccatggaggagccgcagtcagatcc
tagcgtcgagccccctctgagtcaggaaacattttcagacctatggaaactacttcctgaaaacaacgttctgtc
ccccttgccgtcccaagcaatggatgatttgatgctgtccccggacgatattgaacaatggttcactgaagaccc
aggtccagatgaagctcccagaatgccagaggctgctccccccgtggcccctgcaccagcagctcctacaccggc
ggcccctgcaccagccccctcctggcccctgtcatcttctgtcccttcccagaaaacctaccaggggacctacgg
tttccgtctgggcttcttgcattctgggacagccaagtctgtgacttgcacgtactcccctgccctcaacaagat
gttttgccaactggccaagacctgccctgtgcagctgtgggttgattccacacccccgcccggcacccgcgtccg
cgccatggccatctacaagcagtcacagcacatgacggaggttgtgaggcgctgcccccaccatgagcgctgctc
agatagcgatggtctggcccctcctcagcatcttatccgagtggaaggaaatttgcgtgtggagtatttggatga
cagaaacacttttcgacatagtgtggtggtgccctttaccatgctgaggttggctctgactgtaccaccatcca
ctacaactacatgtgtaacagttcctgcatgggcggcatgaaccggaggcccatcctcaccatcatcacactgga
agactccagtggtaatctactgggacggaacagctttgaggtgcgtgtttgtgcctgtcctgggagagaccggcg
cacagaggaagagaatctccgcaagaaaggggagcctcaccacgagctgcccccagggagcactaagcgagcact
gcccaacaacaccag
>Seq2
tttgttgtggggaggggatggggagtaggacataccagcttagattttaaggtttttactgtgagggatgtttg
ggagatgtaagaaatgttcttgcagttaagggttagtttacaatcagccacattctaggtaggggcccacttcac
cgtactaaccagggaagctgtccctcactgttgaattttctctaacttcaaggcccatatctgtgaaatgctggc
atttgcacctacctcacagagtgcattgtgagggttaatgaaataatgtacatctggccttgaaaccacctttta
```

ttacatggggtctagaactttacccatttaggatttagtttatcctctccctgttggtcggtgggttggtagttt
ctacagttgggcagctggttaggtagagggagttgtcaagtctctgctggcccagccaaaccctgtctgacaacc
tcttggtgaaccttagtacctaaaaggaaatctcaccccatcccacaccctggaggatttcatctcttgtatatg
atgatctggatccaccaagacttgttttatgctcagggtcaatttctttttctttttttttttttttttctttt
ttctttgagactgggtctcgctttgttgcccaggctggagtggagtggcgtgatcttggcttactgctttaaacc
tttttaactttagggtttacccacctcagcctccggagtagctgggaccacaggttcatgccaccatggccagcc
aacttttgcatgttttgtagagatggggtctcacagtgttgcccaggctggtctcaaactcctgggctcaggcga
tccacctgtctcagcctcccagagtgctgggattacaattgtgagccaccacgtccagctggaagggtcaacatc
ttttacattctgcaagcacatctgcattttcaccccacccttcccctccttctccctttttatatcccatttta
tatcgatctcttattttacaataaaactttgctgccacctgtgtgtctgaggggtg

Esercizio di riepilogo

Vogliamo ora identificare tutti gli SNP Coding - Non-Synonymous nel gene NOTCH1 di topo e visualizzarli in rosso.