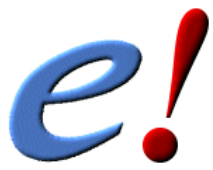




Variazioni nel Ensembl

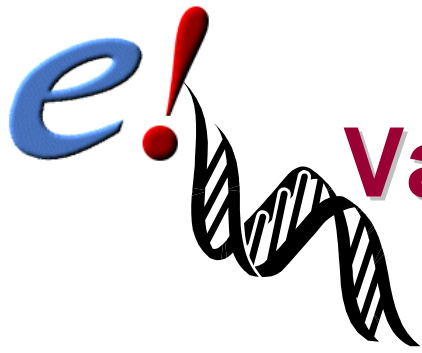


Giulietta Spudich
Feb, 2007



Ordine

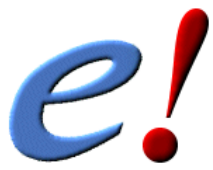
- Variabilità della genoma di una specie
(*SNPs: single nucleotide polymorphisms*)
- Haplotypes (progetto HapMap)
- Linkage Disequilibrium
- SNPs nel 'strains' di *Mus musculus*



Varietà nella Genoma Umana

**Mutazioni genetiche
(sostituzioni (SNPs) e
insertion/deletions (InDels)
sono le cause di variabilità
nel
uno di ogni 300 bp (umano)**

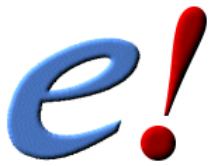
**Ci sono 3 billioni 'base pairs' nella
genoma umana.**



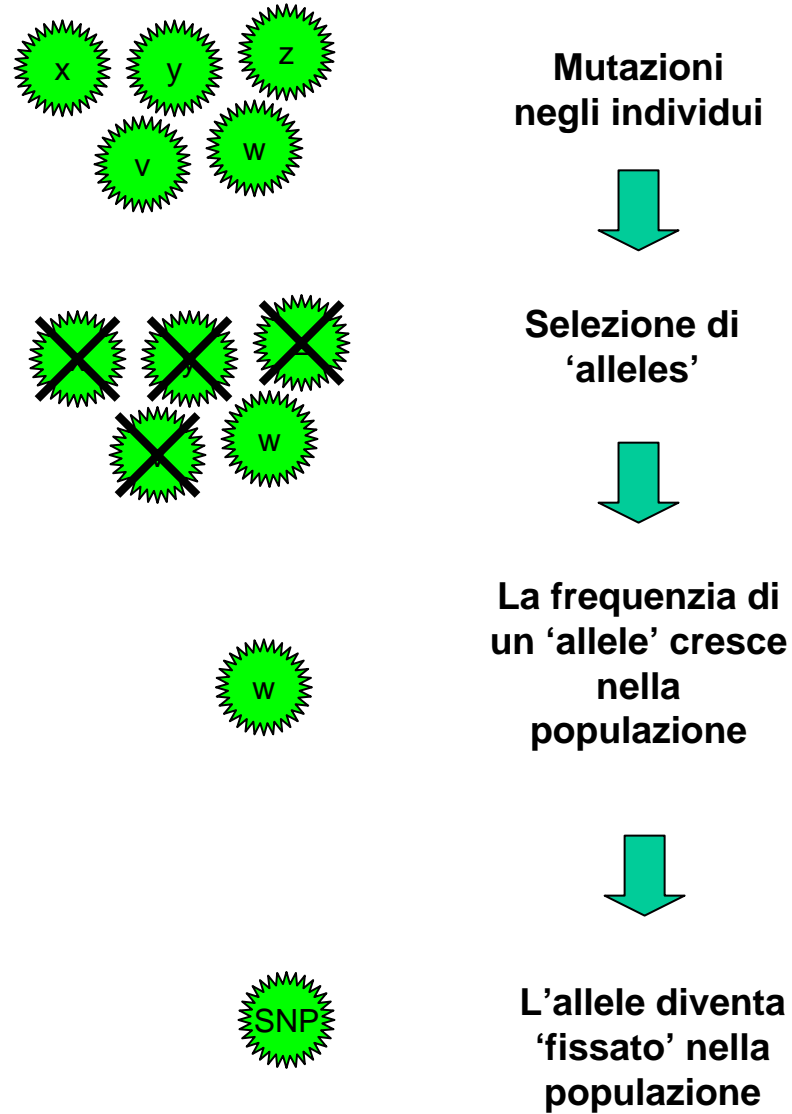
Single nucleotide polymorphisms (SNPs)

- **Estimato 10 milioni SNPs nel genoma umana.**
- **Polymorphism: una variazione che esista in almeno 1% della popolazione.**
- **La maggioranza (90%) dei 'polymorphisms' sono SNPs: cambiano un 'nucleotide'.**





La nascita di SNPs





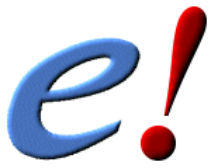
Le conseguenze

Tipo	Conseguenze
SNPs in una regione 'coding' puo cambiare la sequenza della proteina. 'non-synonymous'	La causa della maggioranza di malattia monogenica, e.g: Cystic fibrosis (CFTR) emofilia (F8)
SNPs in una regione 'coding' che non cambiano la sequenza della proteina. 'synonymous'	Puo influire lo 'splicing'
SNPs nel 'promoter' o regioni di regolazione	Puo cambiare la quantità, locazione o tempismo del 'gene expression'
SNPs nelle altre regioni	Non sappiamo un impatto Sono utili come 'markers'



Lo studio delle variazioni- perchè?

- Prevedere il rischio di una malattia
- Prevedere le reazioni ai cambiamenti di ambiente o ai medicinali (pharmacogenomics)
- Biological markers
- Medicina forense/ criminali
- Evoluzione
- Nel lab: per progettare 'primers', studi del hybridization



Applicazioni Pratiche

Ferro eccessivo nel sangue
Mutazione nel gene HFE
0.5% della popolazione del USA ha due mutated 'alleles'.

Medicina personalizzata

Evoluzione

Table of Contents
 C&EN Classifieds
 News of the Week
 Cover Story
 Editor's Page
 Business
 Government & Policy
 Science/Technology
 Concentrates
 Business
 Government & Policy
 Science/Technology
 Education
 ACS News
 Calendars
 Books
 Digital Briefs
 ACS Comments
 Career & Employment
 Special Reports
 Letters
 New Products
 Meetings
 Newsbriefs
 Nanotechnology
 What's That Stuff?
 Pharmaceutical Century

Hot Articles
 Safety Letters
 Chemovlopedia

Back Issues
 2001

How to Subscribe
 Subscription Changes
 About C&EN
 Copyright Permission
 E-mail webmaster

COVER STORY

August 13, 2001

Volume 79, Number 33
 CENEAR 79 33 pp. 37-42
 ISSN 0009-2347

[\[Previous Story\]](#) [\[Next Story\]](#)

CHEMICAL
 & Engineering News
 NEW VOICES IN CHEMISTRY
 CLICK HERE

Millennium
 Special

editorial

Gen
 lea
 CEL

FO
 high
 COUR

What
 mak
 pres
 indi
 test
 dete
 wou
 not
 pha

THE GENOGRAPHIC PROJECT POWERED BY IBM.

MAIN MENU GENETICS OVERVIEW ATLAS OF THE HUMAN JOURNEY YOUR GENETIC JOURNEY

INTRODUCTION ABOUT THE PROJECT HOW TO PARTICIPATE FREQUENTLY ASKED QUESTIONS CREDITS

A LANDMARK STUDY OF THE HUMAN JOURNEY

Who was your first ancestor? New DNA studies say that all humans descended from an African ancestor who lived only 60,000 years ago. Uncover the specific paths that led from him to you—the ultimate human history, as written in our genes.

NEWS AND RESOURCES

- Humans to Blame for Ice Age Extinctions, Study Says
- More Related News >>
- Related Web Resources >>

YOUR GENETIC JOURNEY

Explore your own genetic journey with Dr. Spencer Wells. DNA analysis includes a depiction of your ancient ancestors and an interactive map tracing your genetic lineage around the world and through the ages.

Enter Your Genographic Project Kit ID Here

LOG IN

Remember my log in (optional)

ALSO SEE

- Help Support the Genographic Project Field Research
- Video: Indigenous Representatives Talk About Their Migratory Histories (Download Windows Media)

GENETICS OVERVIEW

Delve inside and explore the basics of genetics, from chromosomes to natural selection.

Question of the Week

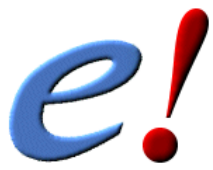
- What happens once I order a Genographic Project Participation Kit? >>

THE WAITT FAMILY FOUNDATION
 Global field science supported by the Waitt Family Foundation

NATIONAL GEOGRAPHIC IBM
 A research partnership of National Geographic and IBM

NATIONALGEOGRAPHIC.COM
 © 1996-2005 National Geographic Society. All rights reserved.

Home | Site Index | Search | Free Newsletters | Subscriptions
 Shopping | Contact Us | Advertise With Us | New Privacy Policy | Press Room



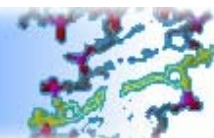
SNPs – tipi nel Ensembl

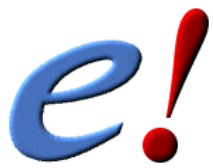
- **Coding SNPs:**
 - Non-synonymous (cambiano la seq aa)
 - Synonymous (non cambiano la seq aa)
- **Non-coding SNPs:**
 - UTR (Untranslated Regions)
 - Regulatory Region
 - Intronic
- **Intergenic (5Kb)** (tra I geni)



dbSNP (NCBI)

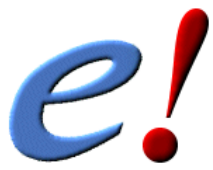
- Database di SNPs (e polymorphisms come in-dels)
- 27.8 Millions sottomissioni (5.6 Millions 'reference SNPs' (rs)).
- <http://www.ncbi.nlm.nih.gov/SNP>



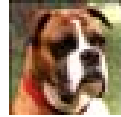


SNPs in Ensembl

- La maggioranza di SNPs sono di dbSNP (ID: rs.....):
- Per umano anche dal:
 - HGVbase
 - TSC
 - Affy GeneChip Mapping Array (100K, 500K)
- Per topolino che anche:
 - Sanger (qualche 'strains')



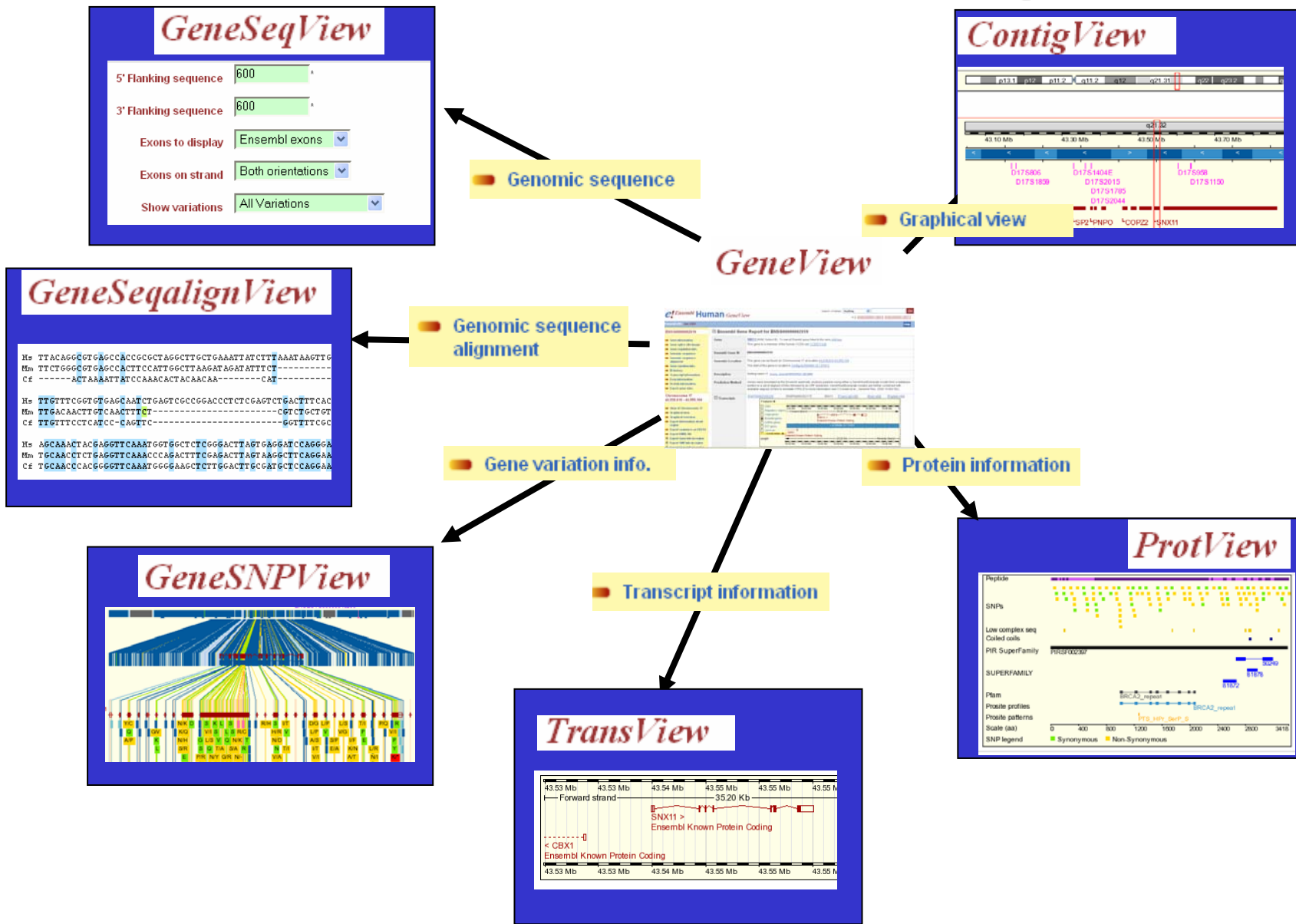
SNPs nel Ensembl - Speci



- umano
- topolino
- ratto
- cane
- gallo
- pesce zebra
- zanzare



SNPs nel Ensembl - Pagine





GeneSNPView

Quale SNPs sono nel mio gene?

Choose SNP type

SNPs and variations in region of gene ENSG00000130427

Assembly

Transcript zoomed

SNPs

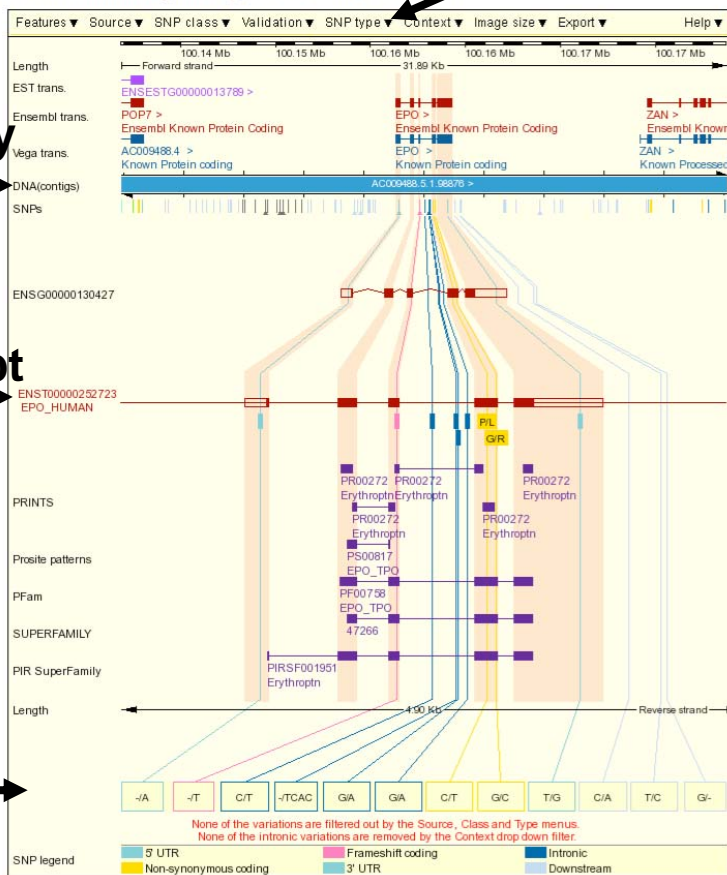


Table of Variations

ID	Type	Chr: bp	Alleles	Ambiguity	AA change	AA co-ordinate	Class	Source	Validation
rs34144627	SPRIME_UTR	7: 100156486-100156485	-/A	-	-	-	insertion dbSNP	-	-
rs34937405	FRAMESHIFT_CODING	7: 100157590-100157589	-/T	-	-	77 (1)	insertion dbSNP	-	-
rs507392	INTRONIC	7: 100157872	C/T	Y	-	-	snp	HGVbase, dbSNP	cluster, doublehit
rs33978705	INTRONIC	7: 100158068-100158067	-/TCAC	-	-	-	insertion dbSNP	-	-
rs484199	INTRONIC	7: 100158087	G/A	R	-	-	snp	HGVbase, dbSNP	-
rs7789679	INTRONIC	7: 100158157	G/A	R	-	-	snp	dbSNP	-
rs11976235	NON_SYNONYMOUS_CODING	7: 100158317	C/T	Y	P/L	114 (2)	snp	dbSNP	cluster
rs1126887	NON_SYNONYMOUS_CODING	7: 100158394	G/C	S	G/R	140 (1)	snp	dbSNP	-
rs564449	3PRIME_UTR	7: 100159074	T/G	K	-	-	snp	HGVbase, dbSNP, TSC, Affy GeneChip 500K Mapping Array	-
rs551238	DOWNSTREAM	7: 100159464	C/A	M	-	-	snp	HGVbase, dbSNP, TSC	cluster, freq, doublehit
rs4729606	DOWNSTREAM	7: 100159726	T/C	Y	-	-	snp	HGVbase, dbSNP	cluster, doublehit
rs35970223	DOWNSTREAM	7: 100159780	G/-	-	-	-	deletion dbSNP	-	-

© 2006 WTSI / EBI. Ensembl is available to download for public use - please see the code licence for details.

- Gene tree info
- Gene variation info.
- ID history



TransView

SNPs nella sequenza

```

1  CCCGGAGCCGGACCGGGGCCACCGCGCCCGCTCTGCTCCGACACCAGCCCGCCCTGGACAG
.....

61  CCGCCCTCTCTCCAGGCCCGTGGGGCTGGCCCTGCACCCGCGAGCTTCCCGGGATGAGG
.....

121  GCGCCCGGTTGTTGCTACCCGGCGCGCCCGAGGTCGCTGAGGGGACCCCGGCGAGGCGCGGA
.....

181  GATGGGGGTGCACGAATGTCCTGGCTGGCTGCGCTTCTCTGTCCTGCTGCTGGCTGCC
..M--G--V--H--E--C--P--A--W--L--W--L--L--L--S--L--L--S--L--P
.....

241  TCTGGGCTCCAGTCTTGGGCGCCCGCCACCGCTCATCTGTGACAGCCGAGTCTCTGGA
20 --L--G--L--P--V--L--G--A--P--P--R--L--I--C--D--S--R--V--L--E
.....

301  GAGGTACCTCTTGGAGGCCAAGGAGGCCGAGAAATTCACGACGGGCTGTGCTGAACACTG
40 --R--Y--L--L--Y--A--K--E--A--I--N--I--T--T--G--C--A--Y--H--C
.....

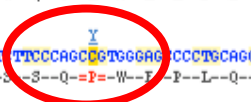
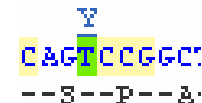
361  CAGCTTGAAATGAGAATACACTGTCCAGACACCAAGTAAATTTCTATGCTGGAGAGG
60 --S--L--N--E--N--I--T--V--P--D--T--K--V--N--Y--A--W--K--R
.....

421  GATGGAGGTCGGGCAGCAGGCCGTGAAAGTCTGGCAGGGCCCTGGCCCTGCTGTGGGAGG
80 --M--E--V--G--Q--A--V--E--V--W--Q--G--L--A--L--L--S--E--A
.....

481  TGTCTGCGGGGCCAGGCCCTGTGGTCAACTTTCCAGCCCTGGGAGCCCTGCAGACT
100 --V--L--R--G--Q--A--L--L--V--N--S--Q--P--W--F--P--L--Q--L
.....

541  GCATGTGATTAAGCCGTCAGTGGCTTTCGAGCCTCACCCTCTGCTTCGGGCTCTGG
120 --H--V--D--K--A--V--S--G--L--R--S--L--T--T--L--L--R--A--L--G

```



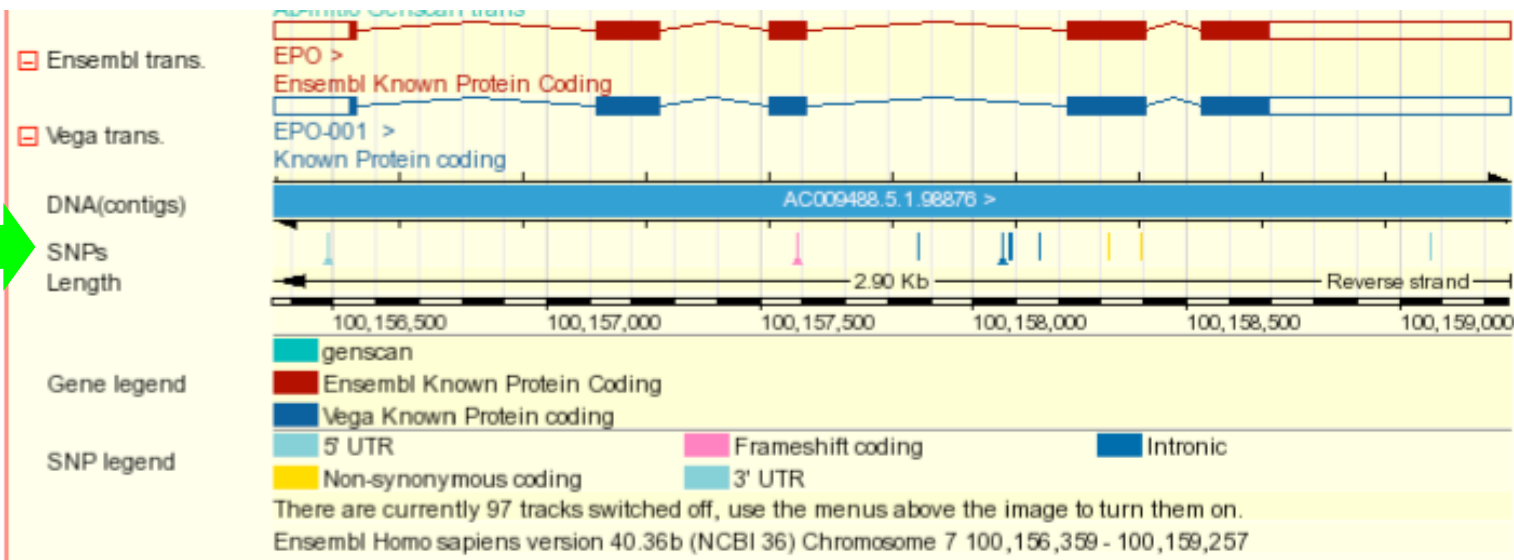


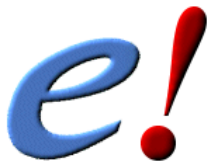
ContigView

SNPs in una regione cromosomale

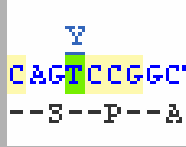
Features ▼ Comparative ▼

- SNPs
- Genotyped SNPs
- Affy 100k SNP
- Affy 500k SNP





SNPs nel Ensembl (Pagine)



- *TransView*

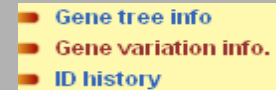
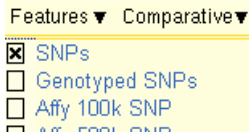
- *ProtView*

- *MapView*

Features:
SNPs in roll-down menus

- *ContigView*

- *GeneSNPView*

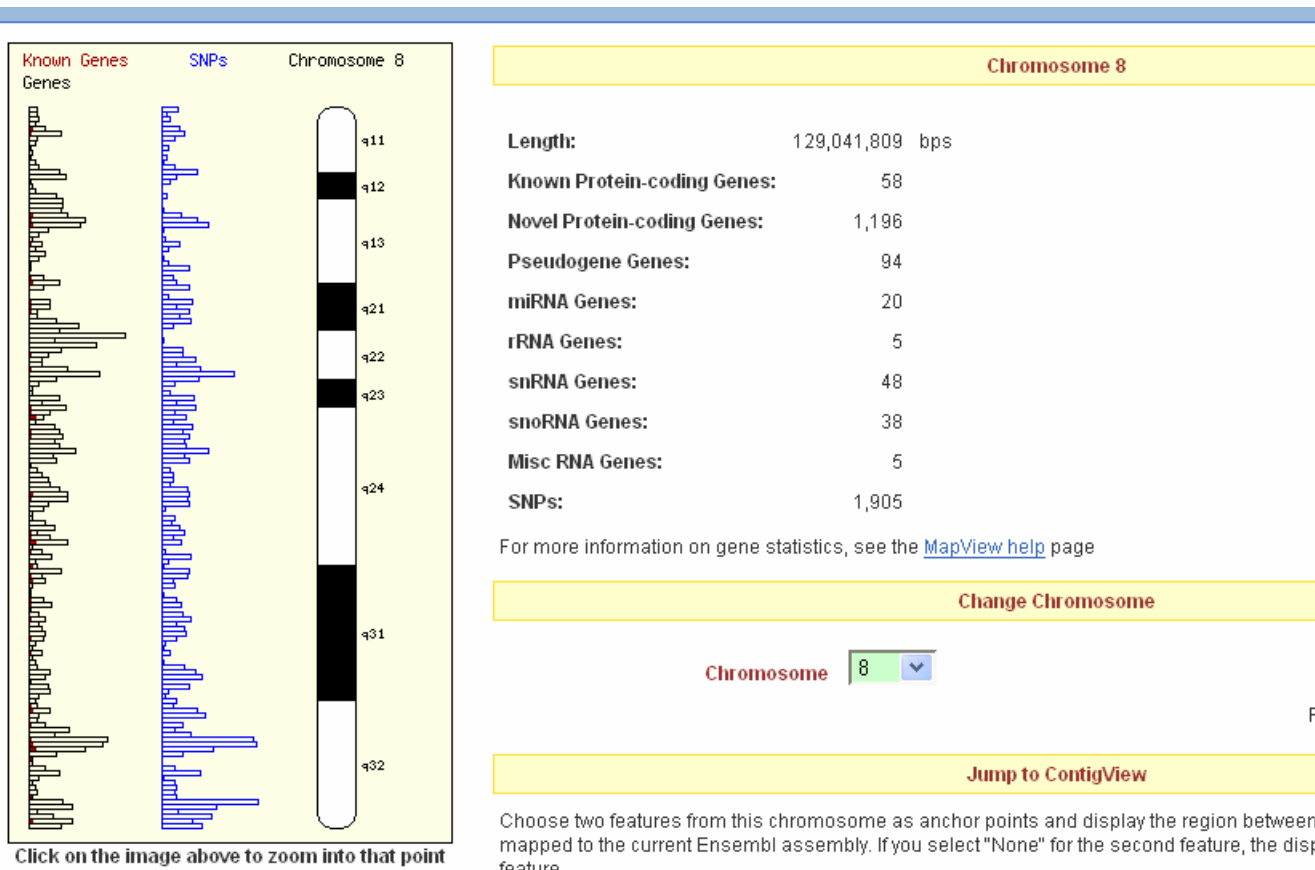


- *SNPView*

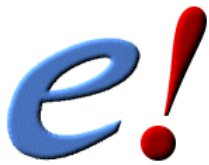


MapView

La densita dei SNPs su una cromosoma



Esempio: Ratto cromosoma 8



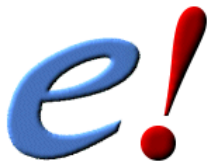
GeneSeqView

SNPs nella sequenza genomica

```

UUUAGTGTATCAATAACATTATTTTTTTTATAGAAAAATAAAUCAAUCCAAAGGAAAAATGGTG
GGCAGGTCCTGGTGAATATGGCTGTGATAATTATATTAGCAATCTCTTTGGCTAATATTT
GAAGCCCAAATAATTGAATCACAAATGATCTCTCCCCAGAAAAATATATAAAATGCACCTTG
GAATCTAGAAGGCCTTTTAGTCTGCCAAAAGAAAACCTTCTTATCATAAGCAGCAGAAGTC
CCATTTACCAAATTGGAAAAGTTAAAGTTACAAAAGCATCAATCATAAGACTTCCATTTCAGG
GATGGCAATTGGGAGTAAGACTTTTTAGTAAAGAAAACTAACACAAAAGTATTAGACTCT
GTAAAAGTCTTACCAAATTTGATTCTGGAACACCTATTCTATTTGGTAAAAGATGATGAA
TTCCGGAGCCAAATGTTCTTTTCATGAAGGATTTGAAAACCTGCCATGAAAAAATACGCCAA
TCAAACCTTTTAGCTTGAGACTCTATTCACTGATTAGATTTTTTTAAATACTGATGGGCCCT
GCTTCTCAGAAGTGACAAGGATGGCCCTCAATCTCAATTTTTGTAATACATGTTCCATTT
GCCAATGAGAAAATATCAGGTTACTAATTTTTCTTCTATTTTTCTAGTGCCTTTCCATGT
GGAGAGTTTCTGTTTCACAACTTCTAGCTCACCCGCTGAGACTGTTTTTCCTGAT |{base 19922:C/T} |{base 19965:A/C}
GTGGACTATGTAATTTCTACTGAGCTGAACCATTTTGGATACCTCCTCAAGCACC
CAATCATTARTGACTTCACTCGGGTTGTTGGTGGAGAGATGCCAAACCAGGTCATTC
CCTTGGCAGGTACTTTATACTGATGGTGTGTCAAAAACCTGGAGCTCAGCTGGCAAAGACACA
GGCCAGGTGGGAGACTCAGGCTATTTACTACACAGACCTATTGGGATGTGAGAAGTATT
TAGGCAAGTTTTCAGCACTAACCAATGTGAGAAGGCCCTCCAGAGATGACCAAGTTGGTGAAA
GAGAGGCTCAAAAACAGCTACCATAACAGGTCAAGAAGAATTTGGCATTAAAGGAAAACAGCA
TAGCAGGATTCAGACAGGCAACTGGTCAACAACATGAAGCTCTGGAAGAAAGGTCCGAC
GTACTCAGGTTTCAGGGCACTACTTCAGCTTCAGCCCTTGCAAAAAACTGGTGAAGTTGGA
AAGTCTTTAGGCTAAAGAAAAATGGATTAATTTAAAAGGGGTAAGAAAAGGCACTCAAGG
AGGAAGGATTAAGGCAAGAAGACTAGGTTCCAAGAAAACAGGGCATGAGAGAGAGTCTTGATC
TACCACTATAGTTCTCGTGGTAGCATCAGAAATCACCTGGGAACGTAGAAAATGCAAAATTCT
CCTGCTCTACACTAGACCTACCAAATCAGAATATCTAGGGGCTGGGCCCCAGCAGTCTGT
GCCCAAACAAGCACTGCAGGTGATTTTGATGCACATTATAGTTTGAAAACCTAGGCCAGGT
GCAGTGGCTCATGCCAATAATCCCAGCACTTTGGGAGACTGAGACGGGAGGATTGCTTAA
ACCCAGGAGTTTGACACCAGCCTGGGCAACACGGCGAAAACCCACCTCTAATTAAAAAAAA
ATACAAAAAATTAGCTAGGTGTGATGGCTCCCACTCTGCTCCACAGTATTCAGGAGGCTG
AGGTGGGAGAAATCACCTGACCTGGAAAAGTCGAGGCTGCAGTGAATTTGATCACACCAC
TGCCTTCAGCCTGAGTGACAGAGTAAGACCCTATCTCAAAAAACAGAAAAAGAAAAACA

```



SNPView

Informazione di un SNP specifico:

Di dov'è?... dbSNP, TSC, HGVbase, Affy Chip Data

Com'era confermato 'validated'? .. Lo stesso di un un altro SNP...

Genotype e frequenza di un 'allele' in una popolazione.

Localizioni nei transcripts (un SNP può essere in piu di uno transcript.)

Tipo: coding/noncoding

SNP Report

SNP rs4149751 (dbSNP124)

Synonyms HGVbase [SNP00448318](#)

Alleles AC (ambiguity code **M**)

Validation status Unknown

Linkage disequilibrium data Links to LDview per population: [PERLEGENAFD_CHN_PANEL](#) [PERLEGENAFD_AFR_PANEL](#) [CHL-HAPMAP_HapMap-CEU](#) [PERLEGENAFD_EUR_PANEL](#)

Sequence region
 CTTGTTCACGACGACACATGCTTGTGATCTACAAGATTCCACCTATAACACATGTC
 TGTCTGTGCTTCATGAAAGAGATGATTCATCTCAAGAGAGATGAGGGACCCCAT
 GTTACTGAAATGGAGGGAGGCGCTTTTACTGCGATATTATGGTGGGTTGAAAGT
 GCAATGAAAGCAATATGAAATATATACCAAGATTCGGATATGTCACGATTAAG
 GAAAAACAGCTCTTACTGAAAGATGATTTCCAGGTTAATTTATTGAAATGAA
 AATTAAGAGGCTTCTACTGAAATGATCTCTGCGATTTGTGATTTGAAATGAT
 ACATTTGATGATGCTTTGTTTGTGTTAAGAGGAAATTTGATATTATCTAGCA
 AATTTGAAAGATGAGCTAGAGGATATATTTGAGAAATTAATGATGATTT
 CTAAAGCCACGCTTCAAAATCTGAA (SNP [highlights](#))

Genotype frequencies per population

Population	Genotypes	Genotypes AC	Genotypes AA	Description
PDA-UNFHCRD-PDA/AFRICAN-PANEL	0.043	0.957		This population of 24 individuals (12 male/12 female) is composed of DNA available from the Coriell Cell Repository. These individuals were selected from the human variation panel of 50 African Americans (005AA). CCR ID PDA-UNFHCRD ID SEX:NA1701 0001 M

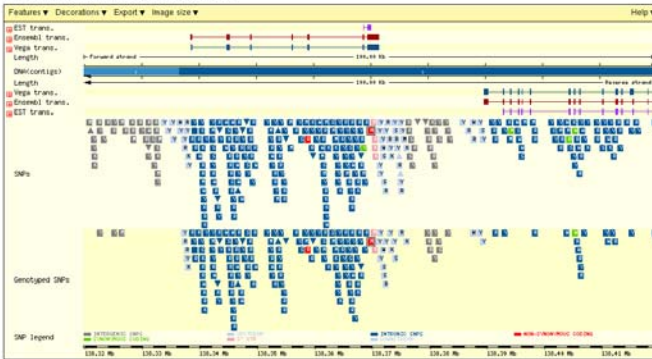
Allele frequencies per population

Population	Alleles	Alleles C	Alleles A	Description
PDA-UNFHCRD-PDA/AFRICAN-PANEL	0.022	0.978		This population of 24 individuals (12 male/12 female) is composed of DNA available from the Coriell Cell Repository. These individuals were selected from the human variation panel of 50 African Americans (005AA). CCR ID PDA-UNFHCRD ID SEX:NA1701 0001 M

SNP rs4149751 is located in the following transcripts

Genomic location (strand)	Transcript: start-end	Gene/SNPView link	Translation: start-end	Peptide allele	Consequence
X:138369745-138369745 (1)	ENST0000218099 : 1410-1410	SNP in gene context	ENSP0000218099: 481-481	TIP	NON_SYNONYMOUS_CODING

SNP Context - chromosome X 138369745



Individual genotypes for SNP rs4149751

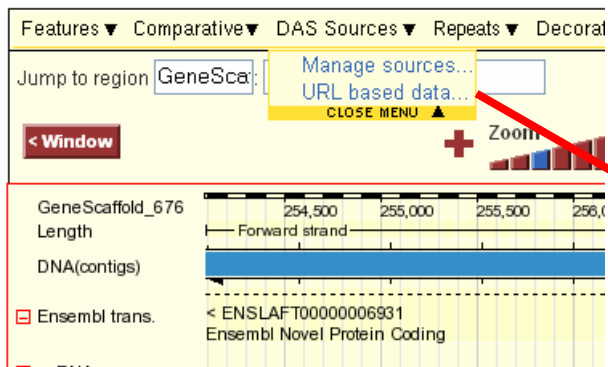
Individual (Gender)	Genotype	Description	Populations	Father	Mother
D006 (Male)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 50	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D012 (Female)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 100	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D034 (Female)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 100	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D013 (Female)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 50	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D007 (Male)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 50	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D040 (Female)	AC	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 50	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D004 (Male)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 50	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		
D005 (Male)	AA	HUMAN VARIATION PANEL, AFRICAN AMERICAN PANEL OF 100	PDA-UNFHCRD-PDA/AFRICAN-PANEL , TSC-CEL-CEL , strum_amosus , TSC-CEL-MOT , strum_amosus , TSC-CEL-WC202 , strum_amosus , MYRISLAB-STANFORD-PUB-HOSMA , IFDA-WEISS-MARTINEZ-02 , AFF-AMM , PARC-PARC/AFRICAN-PANEL , PERLEGENAFD_AFR_PANEL		



I Vostri Dati

Use Ensembl to...

- Run a BLAST search
- Search Ensembl database
- Data mining [BioMart]
- Display your own data
- Export data
- Download data

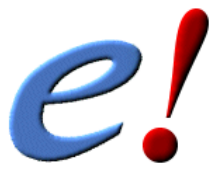


Manage Sources

- Add Data Source
- Upload your data

**Mostrate i vostri dati nelle pagine:
ContigView, CytoView, GeneView, TransView, ProtView and KaryoView**

**O fate un consegna (submission) a dbSNP
<http://www.ncbi.nih.gov/SNP/>**



Ordine

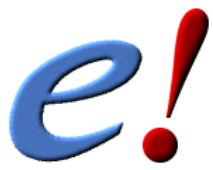
- Variabilità nella genoma di una specie
(*SNPs: single nucleotide polymorphisms*)
- **Haplotypes (progetto HapMap)**
- Linkage Disequilibrium
- SNPs nel 'strains' di *Mus musculus*



HapMap

- Una campagna internazionale per identificare e catalogare cos'è simile e cos'è diverso nella genoma umana.
- Collaborazione tra scienziati e le agenzie del Giappone, Grand Bretagna, Canada, Cina, Nigeria, and l'USA.
- Tutta l'informazione del progetto è accessibile dal pubblico.
- <http://www.hapmap.org/>





Haplotype

- una raccolta di SNPs (di solito ~25 kb)
- ereditato insieme durante il tempo



HapMap

- Fase I:
1 milione SNPs umani
Densità: 1/3,000 'bases'
- Fase II:
Un altro 4.6 milioni SNPs dai
database pubblici
Densità: 1/600 'bases'



HapMap



- Sequenze delle popolazioni africane, asiatiche e con progenitori europei



- 270 DNA sequenze di 4 popolazioni:
 - 30 'trio' (2 genitori e figlio/a adulto/a) della popolazione Yoruba di Ibadan, Nigeria
 - 45 giapponesi dal regione di Tokyo
 - 45 cinesi (Han) di Beijing
 - 30 'trio' di Utah (USA) con progenitori di Europa (nord ed ovest) (gruppo CEPH)



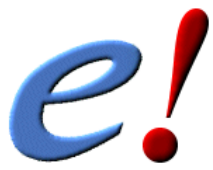
Beijing





SNPView

Population:	CSHL-HAPMAP:HapMap-CEU [size: unknown]
Description:	30 mother-father-child trios from the CEPH collection (Utah residents with ancestry,from northern and western Europe), representing one of the populations studied,in the International HapMap project (http://www.hapmap.org)
SNP in tagged set for this population:	Yes
Super-population:	EUROPE [size: unknown]
Description:	Samples from Europe north and west of Caucasus Mountains, Scandinavia, Atlantic Islands.



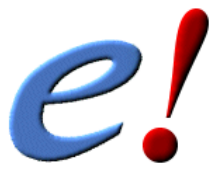
Linkage Disequilibrium

Linkage disequilibrium (LD)

Una misura della frequenza in cui due 'SNPs' sono ereditati insieme.

(Casuale? O collegato?)

$\text{Freq A} * \text{freq B} = \text{Freq(AB)}?$



La misura 'LD'

$$D = p(AB) - p(A)p(B)$$

- $D' = D / D_{\max}$

- $D' = 1$

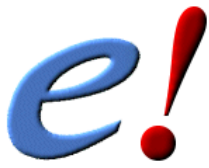
LD 'perfetto'

(~~'allele' A e B sono sempre insieme~~)

- $r^2 = D^2 / p(A)p(B)p(a)p(b)$

- $r^2 = 1$

LD 'perfetto'



Linkage Disequilibrium

LDTableView

Pairwise linkage disequilibrium values

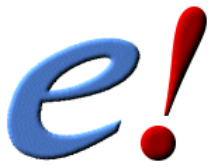
Pairwise r2 values for 7:100133317-100183316. Population: CSHL-HAPMAP:HapMap-CEU

SNPs : bp position	rs221789: 100109457	rs4727459: 100114757	rs221793: 100118832	rs221794: 100119704	rs221795: 100121197	rs2272572: 100123824	rs221797: 100123910	rs10233533: 100124605	rs221801: 100127967	rs11979818: 100131575	rs221780: 100134971	rs6465774: 100138384	rs221770: 100140030	rs10277087: 100145474
rs4727459: 100114757	0.061	-	-	-	-	-	-	-	-	-	-	-	-	-
rs221793: 100118832	0.609	-	-	-	-	-	-	-	-	-	-	-	-	-
rs221794: 100119704	1.000	0.057	0.611	-	-	-	-	-	-	-	-	-	-	-
rs221795: 100121197	0.130	0.507	0.300	0.142	-	-	-	-	-	-	-	-	-	-
rs2272572: 100123824	0.061	0.954	-	0.055	0.484	-	-	-	-	-	-	-	-	-
rs221797: 100123910	0.704	-	0.870	0.818	0.256	-	-	-	-	-	-	-	-	-
rs10233533: 100124605	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs221801: 100127967	0.397	0.448	0.239	0.290	0.898	0.427	0.237	-	-	-	-	-	-	-
rs11979818: 100131575	0.061	0.955	-	0.060	0.531	0.911	-	0.470	-	-	-	-	-	-
rs221780: 100134971	1.000	0.057	0.611	1.000	0.142	0.055	0.818	0.290	0.060	-	-	-	-	-
rs6465774: 100138384	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs221770: 100140030	0.704	-	0.870	0.818	0.256	-	1.000	0.237	-	0.818	-	-	-	-
rs10277087: 100145474	0.397	0.457	0.239	0.282	0.900	0.435	0.231	1.000	0.478	0.282	-	-	0.231	-
rs506597: 100151356	0.682	-	0.856	0.748	0.226	-	0.924	0.209	-	0.748	-	0.924	0.203	-

Rosso = Massimo LD

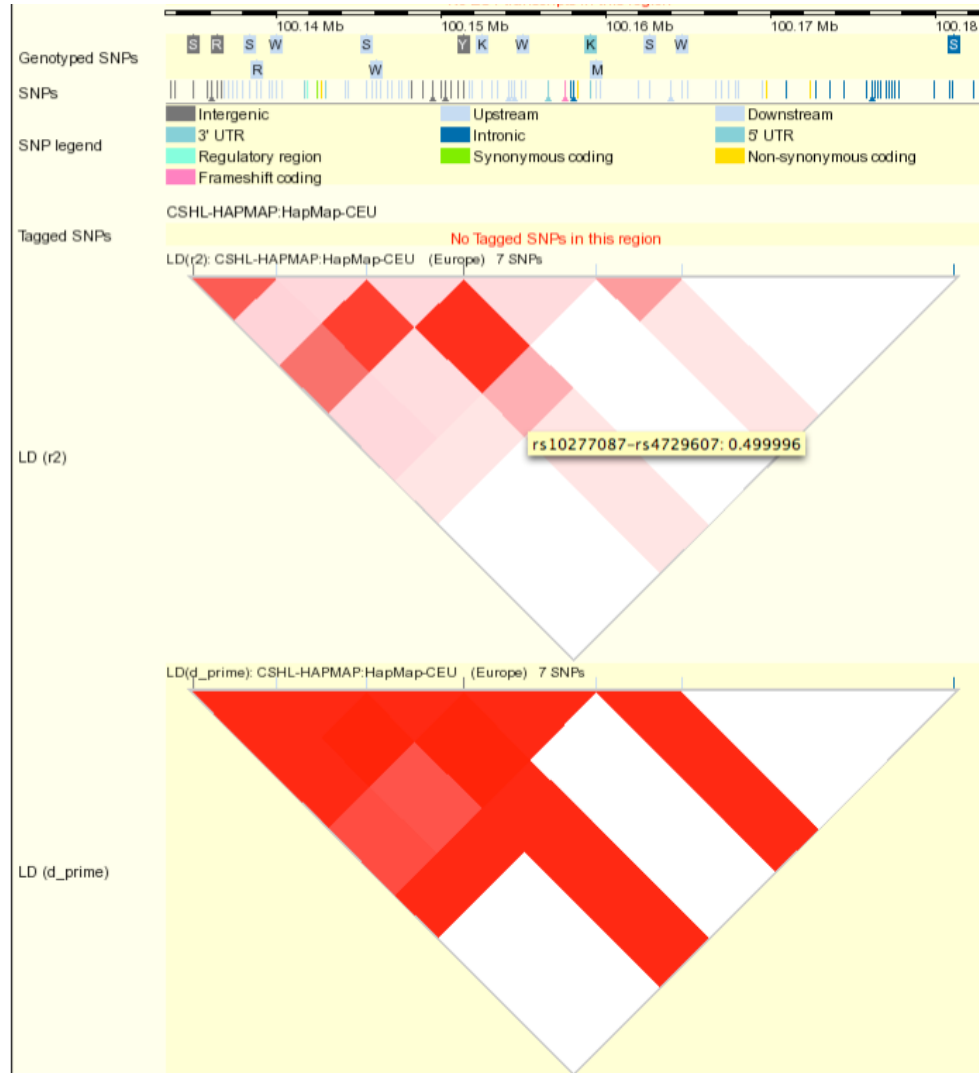
Bianco = Minimo/zero LD





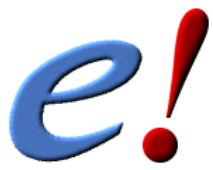
Linkage Disequilibrium

LDView



Rosso (LD alto)

Bianco (LD basso/zero)



Ordine

- Variabilità dell genoma di una specie
(*SNPs: single nucleotide polymorphisms*)
- Haplotypes (progetto HapMap)
- Linkage Disequilibrium
- **SNPs nel BioMart e 'strains' di *Mus musculus***



SNPs nel BioMart

Scegli la database 'Ensembl Variation'



HOME DOCS NEWS CONTACT CREDITS **MARTVIEW**

New XML Help Count Results

» **Dataset:**
Homo sapiens SNPs (dbSNP 126;
HGvbase 15; TSC 1; Affy GeneChip
500K Mapping Array ; Affy GeneChip 10
» **Attributes** (SNPs)
Chromosome Name
Start Position (bp)
Reference ID
» **Filters**
[None selected]

» **Dataset:**
[None Selected]

Database:

Dataset:

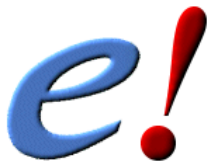
< >

Choose a **Dataset** above, then use the left panel to navigate through the **Attributes** and **Filters** making your selections in this main panel. To preview the results click the **Results** button in the top panel.

[Tutorial](#)

biomart version 0.5

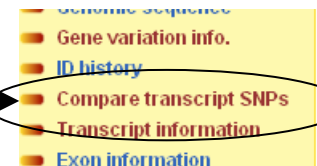
Anche il sito di HapMap ha un BioMart!



SNPs specifici per un 'strain'

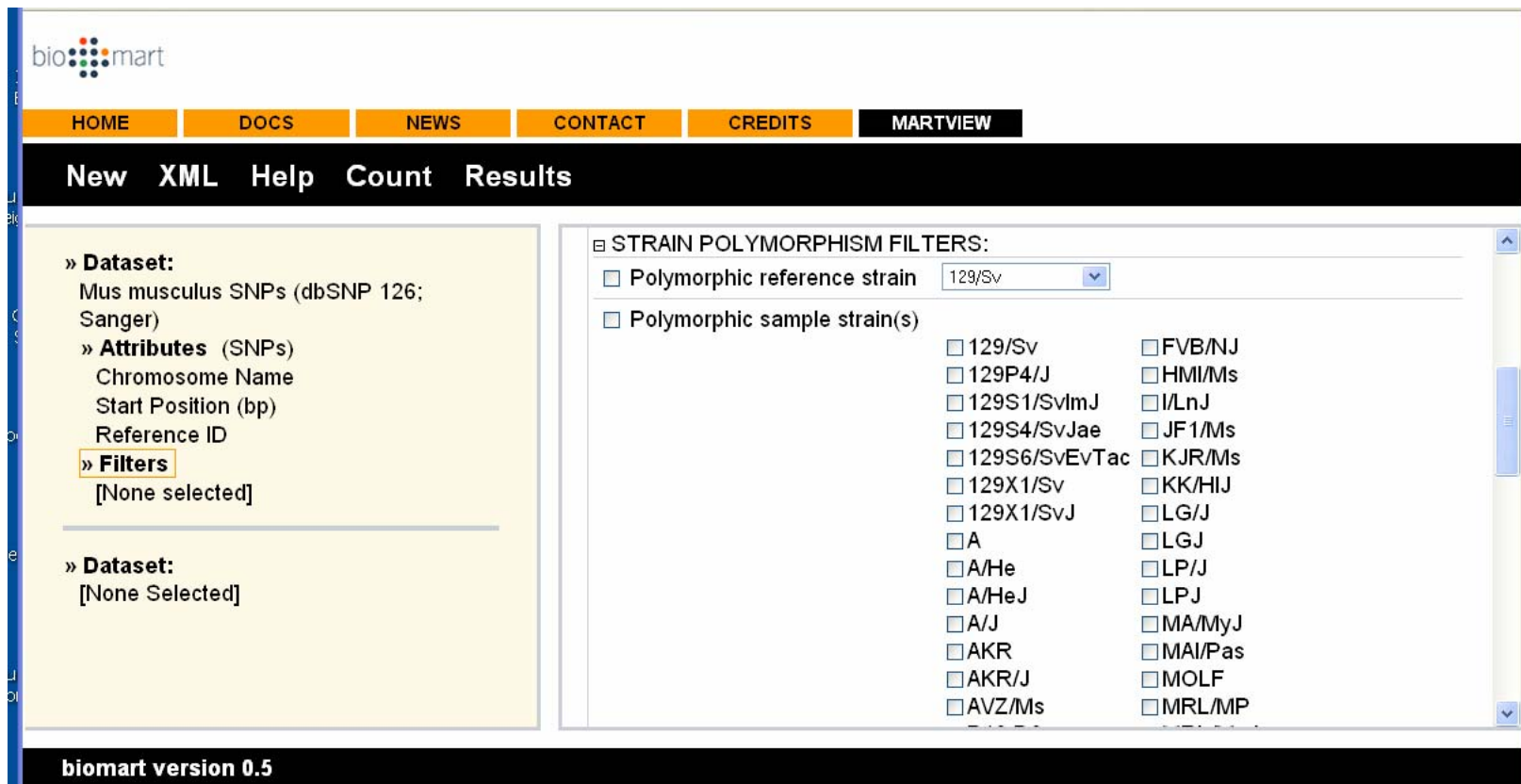
- La sequenza genomica del *Mus musculus* nel Ensembl corrisponde al strain C57BL/6J
- Però c'è l'informazione per i SNPs nei più di 30 altri 'strains'.

- Per vederli nel sito web, 'click' qui:





SNPs per i 'strains' nel BioMart



The screenshot shows the BioMart web interface. At the top, there is a navigation bar with links for HOME, DOCS, NEWS, CONTACT, CREDITS, and MARTVIEW. Below this is a secondary navigation bar with links for New, XML, Help, Count, and Results. The main content area is divided into two panels. The left panel displays the dataset information for 'Mus musculus SNPs (dbSNP 126; Sanger)', including a list of attributes (SNPs, Chromosome Name, Start Position (bp), Reference ID) and a 'Filters' section which is currently empty. The right panel is titled 'STRAIN POLYMORPHISM FILTERS:' and contains a dropdown menu for 'Polymorphic reference strain' set to '129/Sv'. Below this is a list of checkboxes for various mouse strains, including 129/Sv, FVB/NJ, 129P4/J, HMI/Ms, 129S1/SvImJ, I/LnJ, 129S4/SvJae, JF1/Ms, 129S6/SvEvTac, KJR/Ms, 129X1/Sv, KK/HIJ, 129X1/SvJ, LG/J, A, LGJ, A/He, LP/J, A/HeJ, LPJ, A/J, MA/MyJ, AKR, MAI/Pas, AKR/J, MOLF, and AVZ/Ms, MRL/MP.

bio::mart

HOME DOCS NEWS CONTACT CREDITS MARTVIEW

New XML Help Count Results

» **Dataset:**
Mus musculus SNPs (dbSNP 126;
Sanger)

» **Attributes** (SNPs)
Chromosome Name
Start Position (bp)
Reference ID

» **Filters**
[None selected]

» **Dataset:**
[None Selected]

STRAIN POLYMORPHISM FILTERS:

Polymorphic reference strain 129/Sv

Polymorphic sample strain(s)

<input type="checkbox"/> 129/Sv	<input type="checkbox"/> FVB/NJ
<input type="checkbox"/> 129P4/J	<input type="checkbox"/> HMI/Ms
<input type="checkbox"/> 129S1/SvImJ	<input type="checkbox"/> I/LnJ
<input type="checkbox"/> 129S4/SvJae	<input type="checkbox"/> JF1/Ms
<input type="checkbox"/> 129S6/SvEvTac	<input type="checkbox"/> KJR/Ms
<input type="checkbox"/> 129X1/Sv	<input type="checkbox"/> KK/HIJ
<input type="checkbox"/> 129X1/SvJ	<input type="checkbox"/> LG/J
<input type="checkbox"/> A	<input type="checkbox"/> LGJ
<input type="checkbox"/> A/He	<input type="checkbox"/> LP/J
<input type="checkbox"/> A/HeJ	<input type="checkbox"/> LPJ
<input type="checkbox"/> A/J	<input type="checkbox"/> MA/MyJ
<input type="checkbox"/> AKR	<input type="checkbox"/> MAI/Pas
<input type="checkbox"/> AKR/J	<input type="checkbox"/> MOLF
<input type="checkbox"/> AVZ/Ms	<input type="checkbox"/> MRL/MP

biomart version 0.5



Gruppo Ensembl

Leaders [Ewan Birney \(EBI\)](#), [Tim Hubbard \(Sanger Institute\)](#)

[Glenn Proctor](#), Ian Longden, Patrick Meidl, Andreas Kähäri

[Arek Kasprzyk](#), Syed Haider, Richard Holland, Damian Smedley, Benoit Ballester

Eugene Kulesha

[Xosé M Fernández](#), Bert Overduin, Michael Schuster, Giulietta Spudich

[James Smith](#), Fiona Cunningham, Anne Parker, Bethan Pritchard, Stephen Rice, Steve Trevanion, Matt Wood

[Abel Ureta-Vidal](#), Benoit Ballester, Kathryn Beal, Stephen Fitzgerald, Javier Herrero Sánchez, Albert Vilella

[Val Curwen](#), [Steve Searle](#), Bronwen Aken, Julio Banet, Laura Clarke, Sarah Dyer, Kevin Howe, Felix Kokocinski, Jan-Hinnerck Vogel, Simon White

[Paul Flicek](#), Yuan Chen, Stefan Gräf, Nathan Johnson, Daniel Rios

[Martin Hammond](#), Dan Lawson, Karyn Megy

