1000 Genomes Browser Orientation and Pilot Data

http://browser.1000genomes.org

Based on Pilot Project Data

Paul Flicek

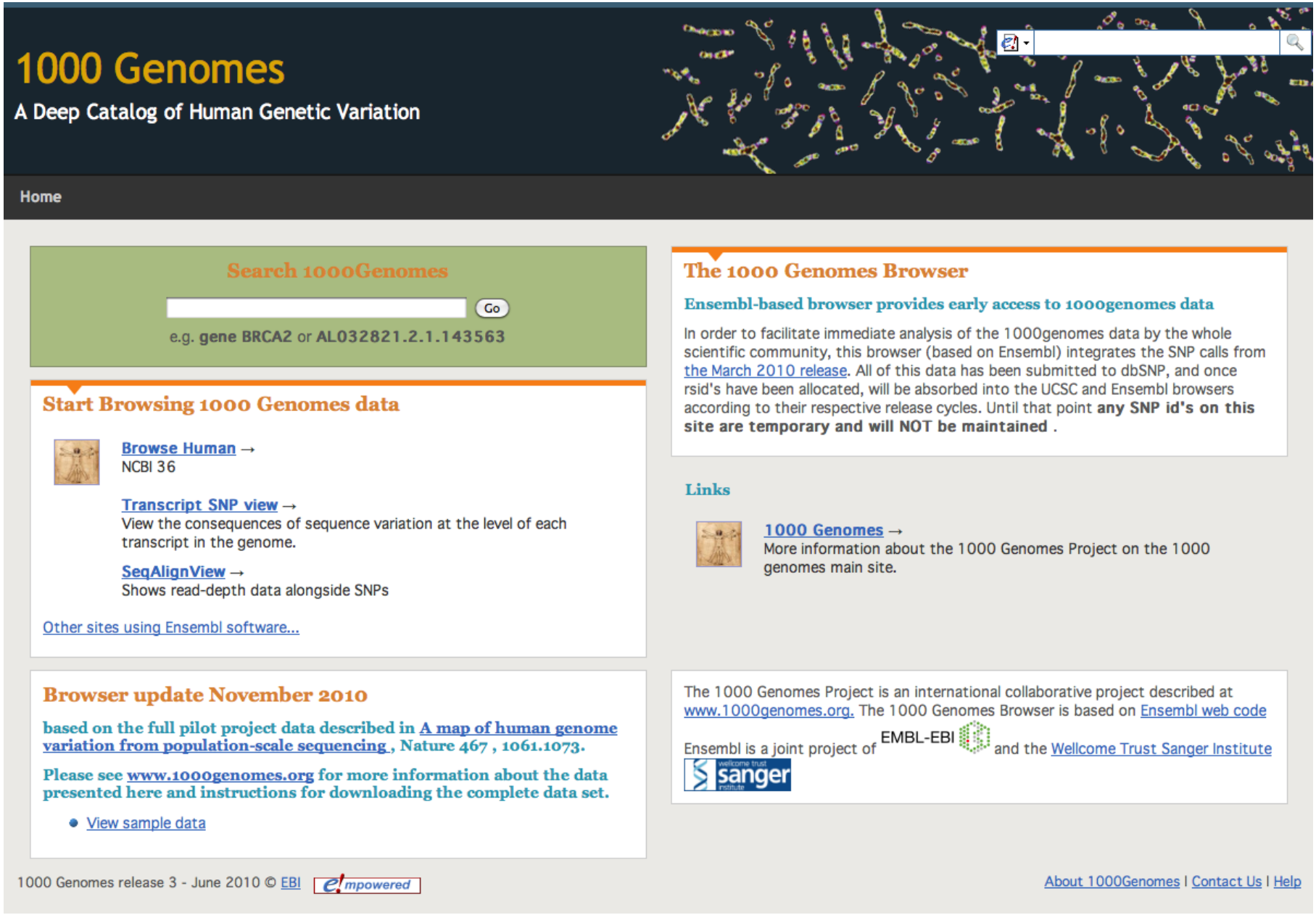
European Bioinformatics Institute

3 November 2010

Overview

- Based on version 54 of Ensembl code and NCBI36 version of the human genome assembly
- Contains all of the gene information normally present in Ensembl
  - Gene and transcript annotation, external references, sequence data
- Incorporates essentially all of the 1000 Genomes Pilot data
  - Some details and additional configuration options will be added over the coming weeks
- Please send questions to info@1000genomes.org

- 1000 Genomes full project data will be released in a new browser as the data becomes available

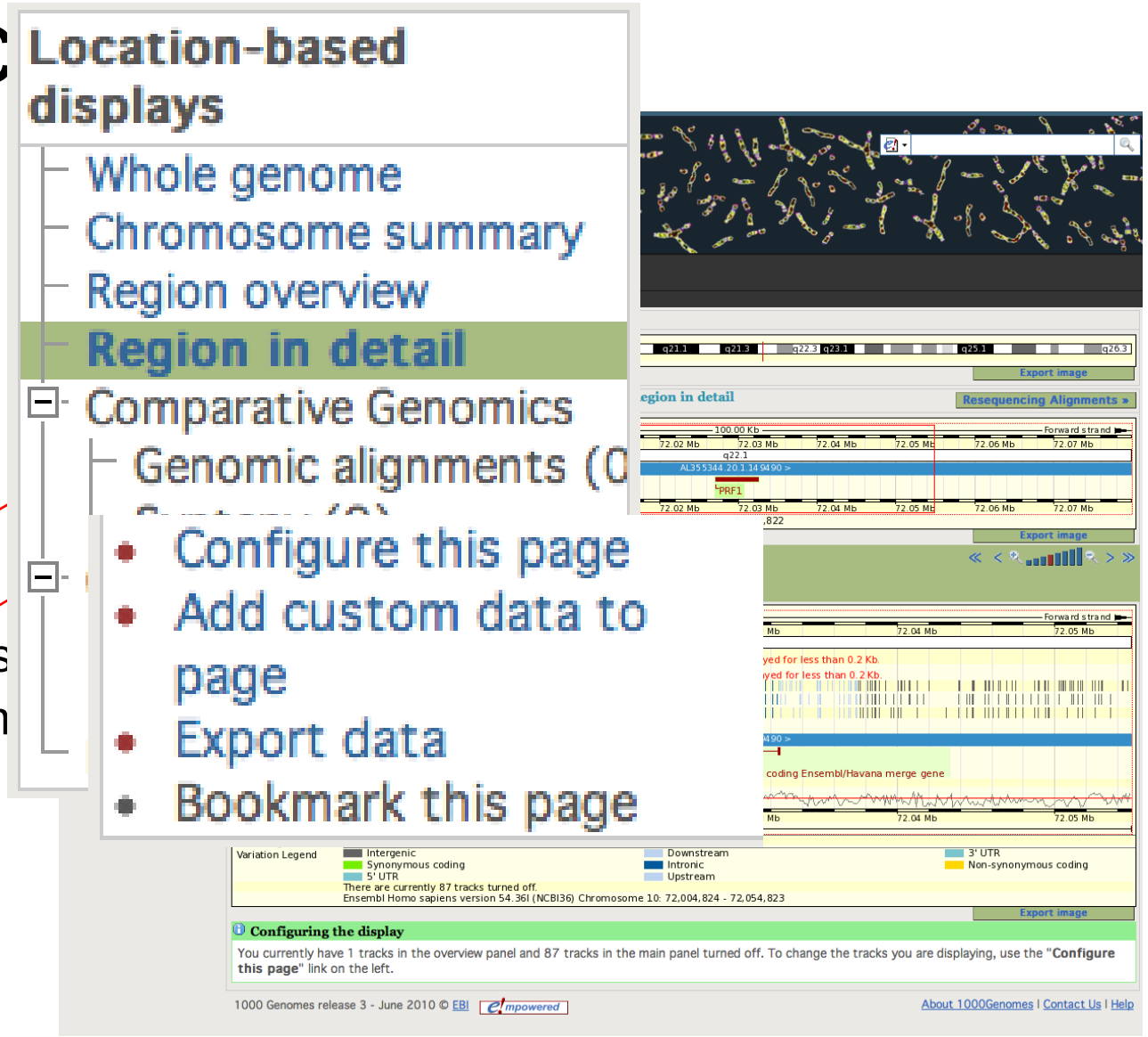


1000 Genomes Browser Home Page

http://browser.1000genomes.org

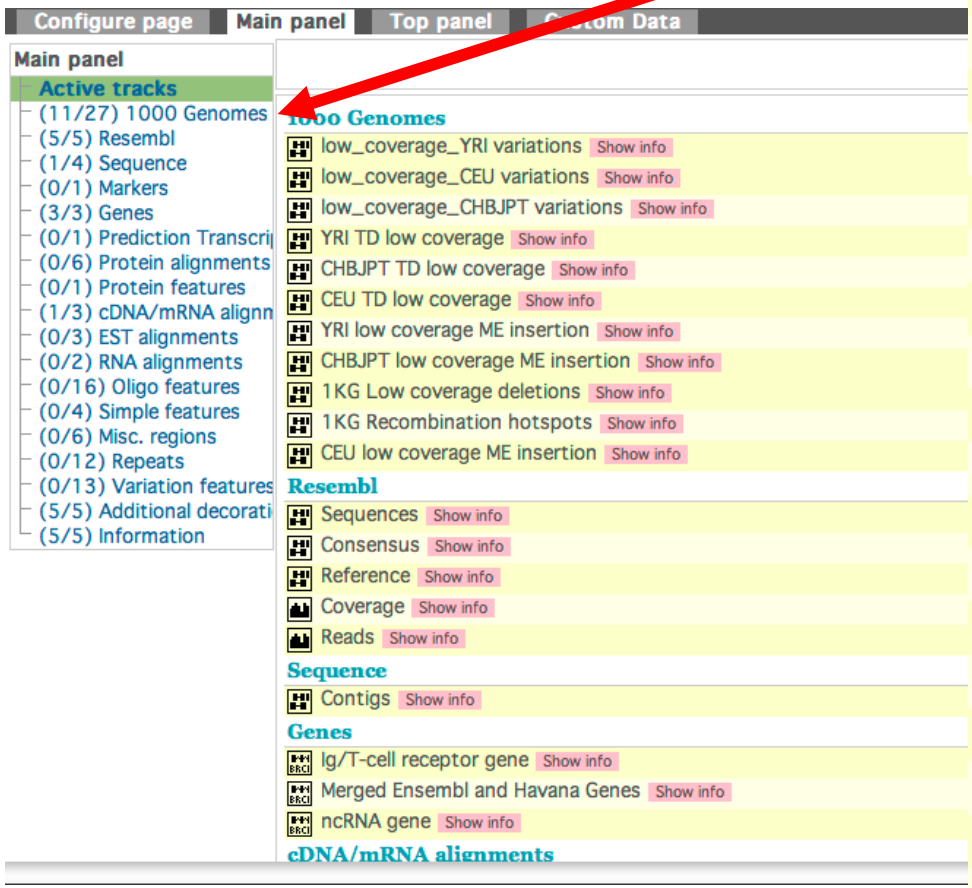
Main loc

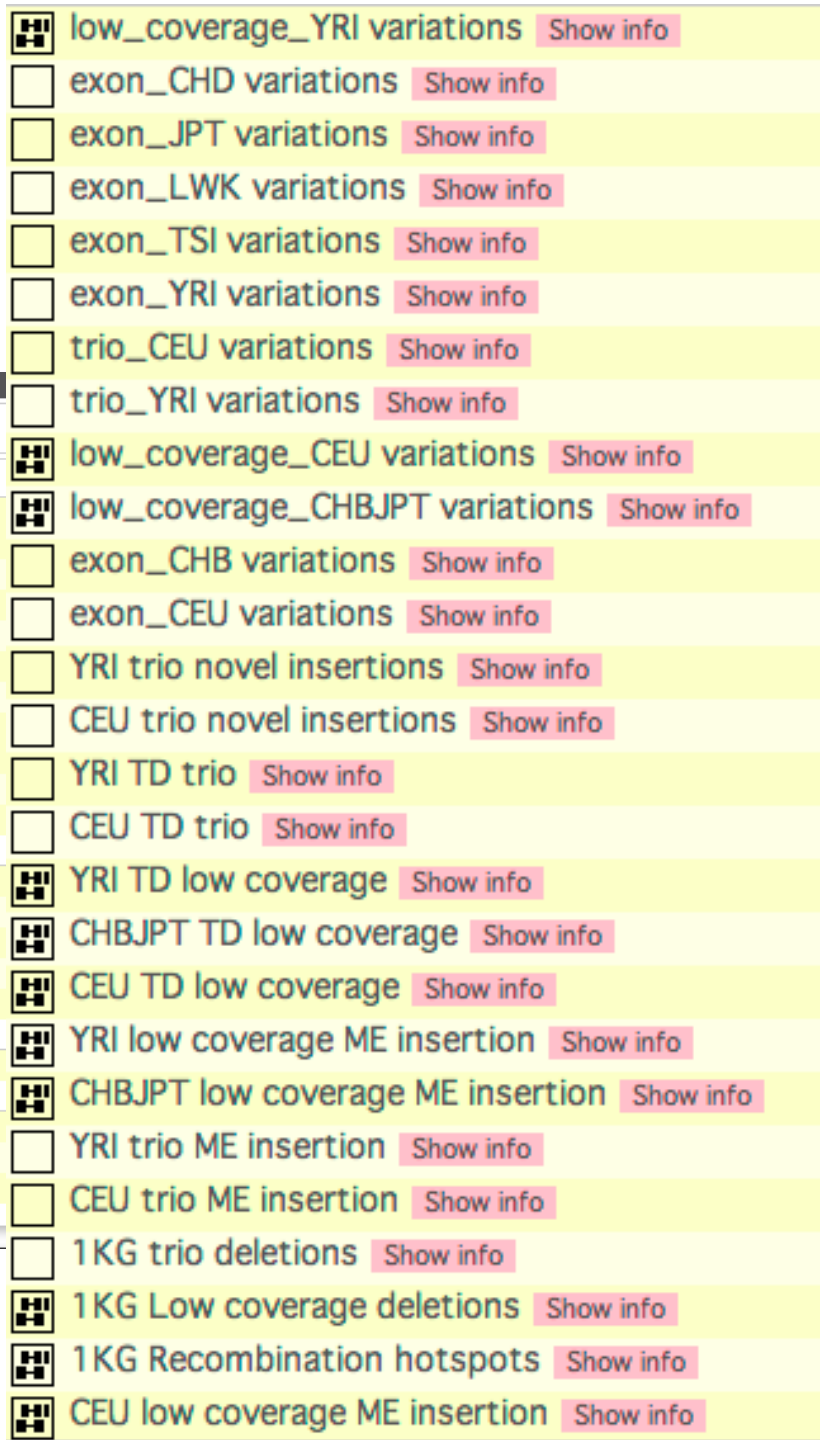
- Built on Ensembl
- Navigation is on the left hand side
- Page configuration and data expert links
- Includes only human data in current release



http://browser.1000genomes.org

Page configuration

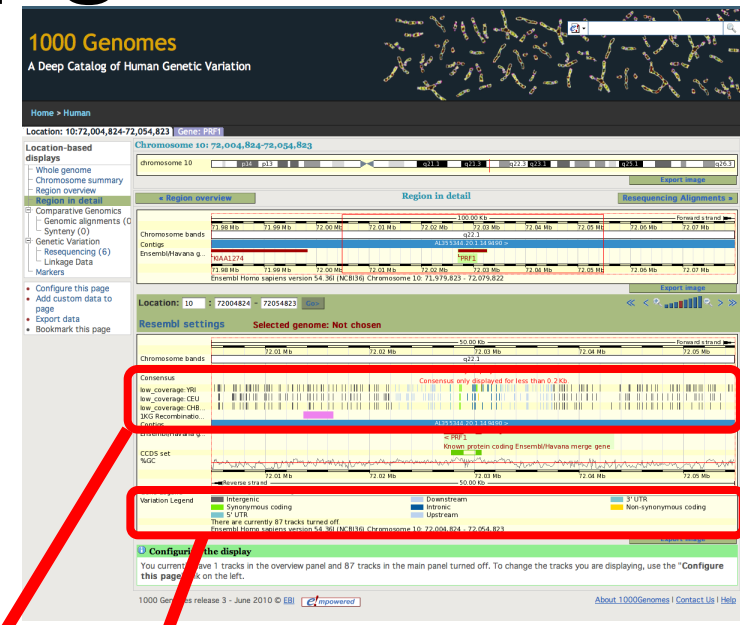


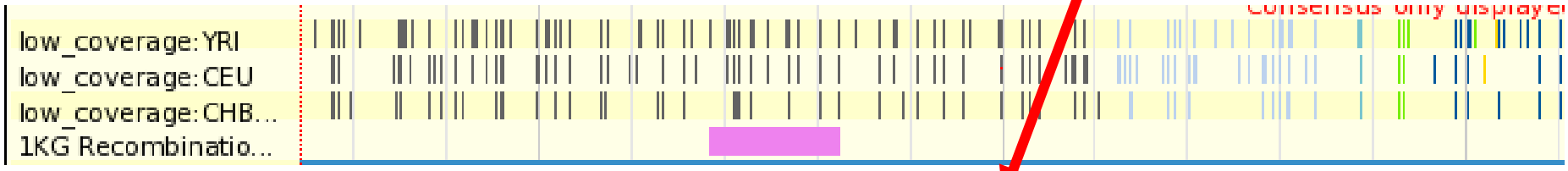


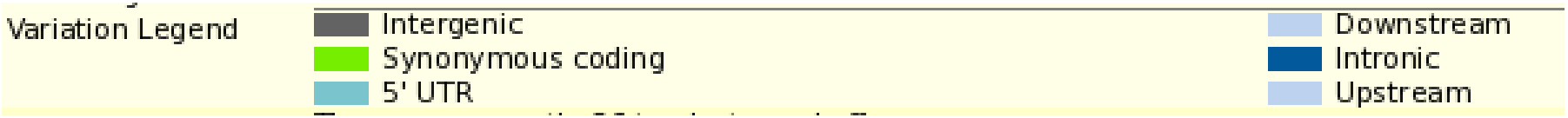
http://browser.1000genomes.org

1000 Genomes SNPs

- The high coverage trios, low coverage populations and exon capture data can be viewed in specific tracks on location pages
- These are selected from the “Configure this page” menu and appear as tracks near the middle of the display
- Tracks for all SNPs from dbSNP build 129 are also available



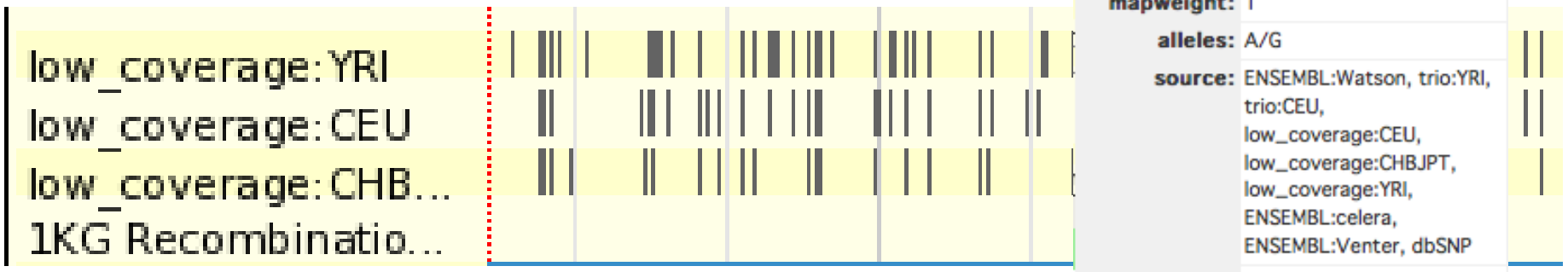




http://browser.1000genomes.org

SNP Information

- SNPs are clickable which brings up a small window with basic information
- The “SNP properties” link leads to a dedicated page for the SNP with detailed information (including information imported from dbSNP)



X

Variation: rs10740355

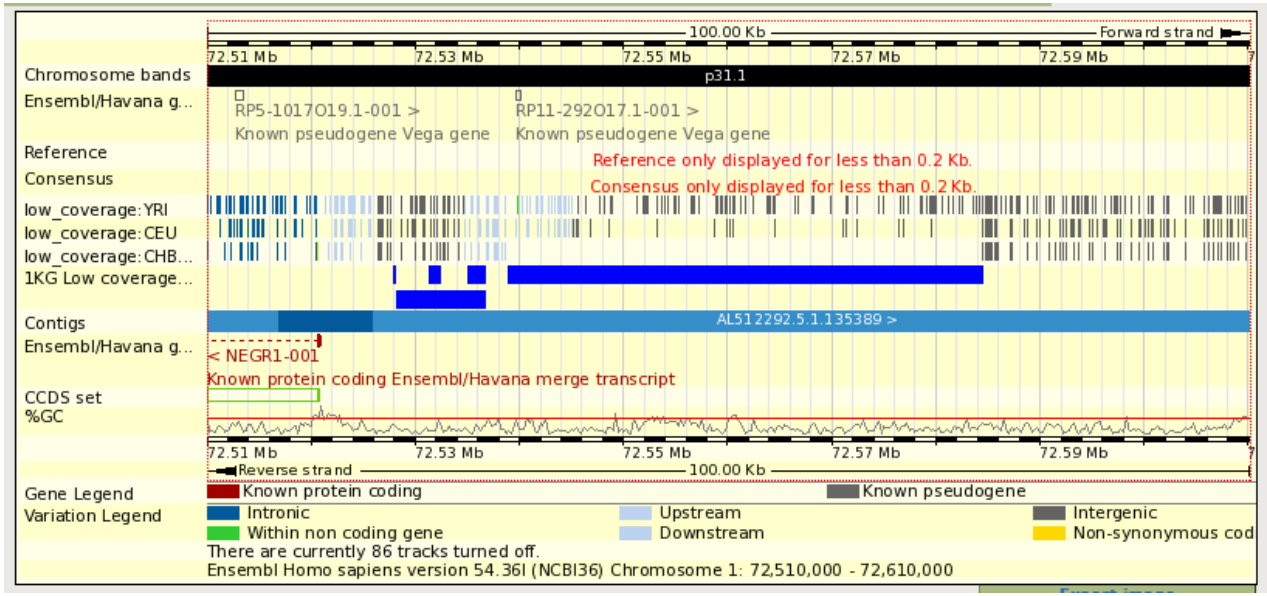
Variation Properties

bp:	72019982
status:	cluster, freq, hapmap
class:	snp
ambiguity code:	R
mapweight:	1
alleles:	A/G
source:	ENSEMBL:Watson, trio:YRI, trio:CEU, low_coverage:CEU, low_coverage:CHBJPT, low_coverage:YRI, ENSEMBL:celera, ENSEMBL:Venter, dbSNP
type:	INTERGENIC

http://browser.1000genomes.org

SV Information

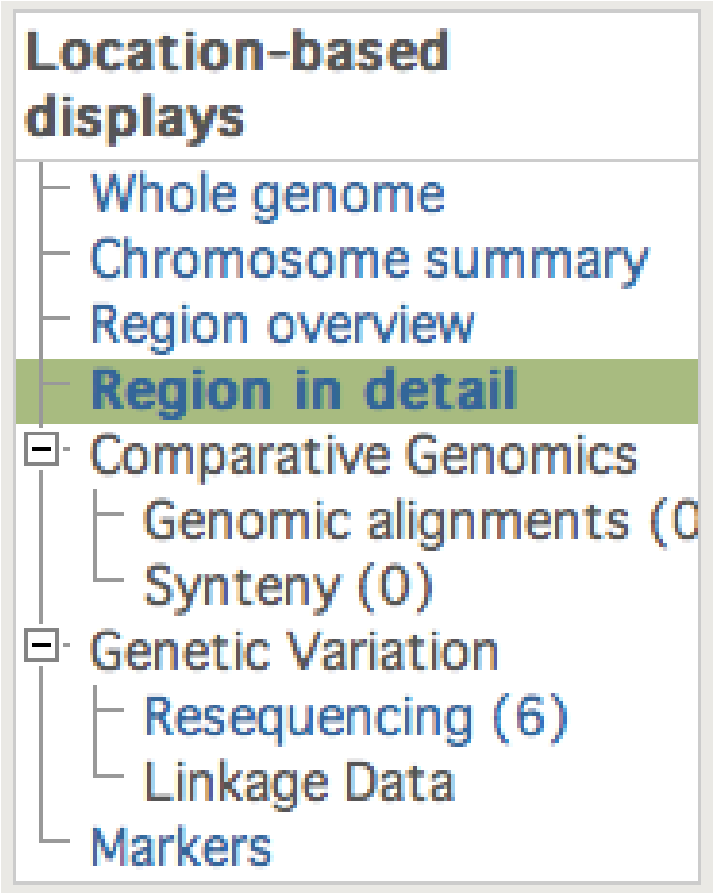
- Structural variants of specific types are selectable via the “configure this page” link
- Pop-up menus from clicking on the variant list whether the variant has been validated



http://browser.1000genomes.org

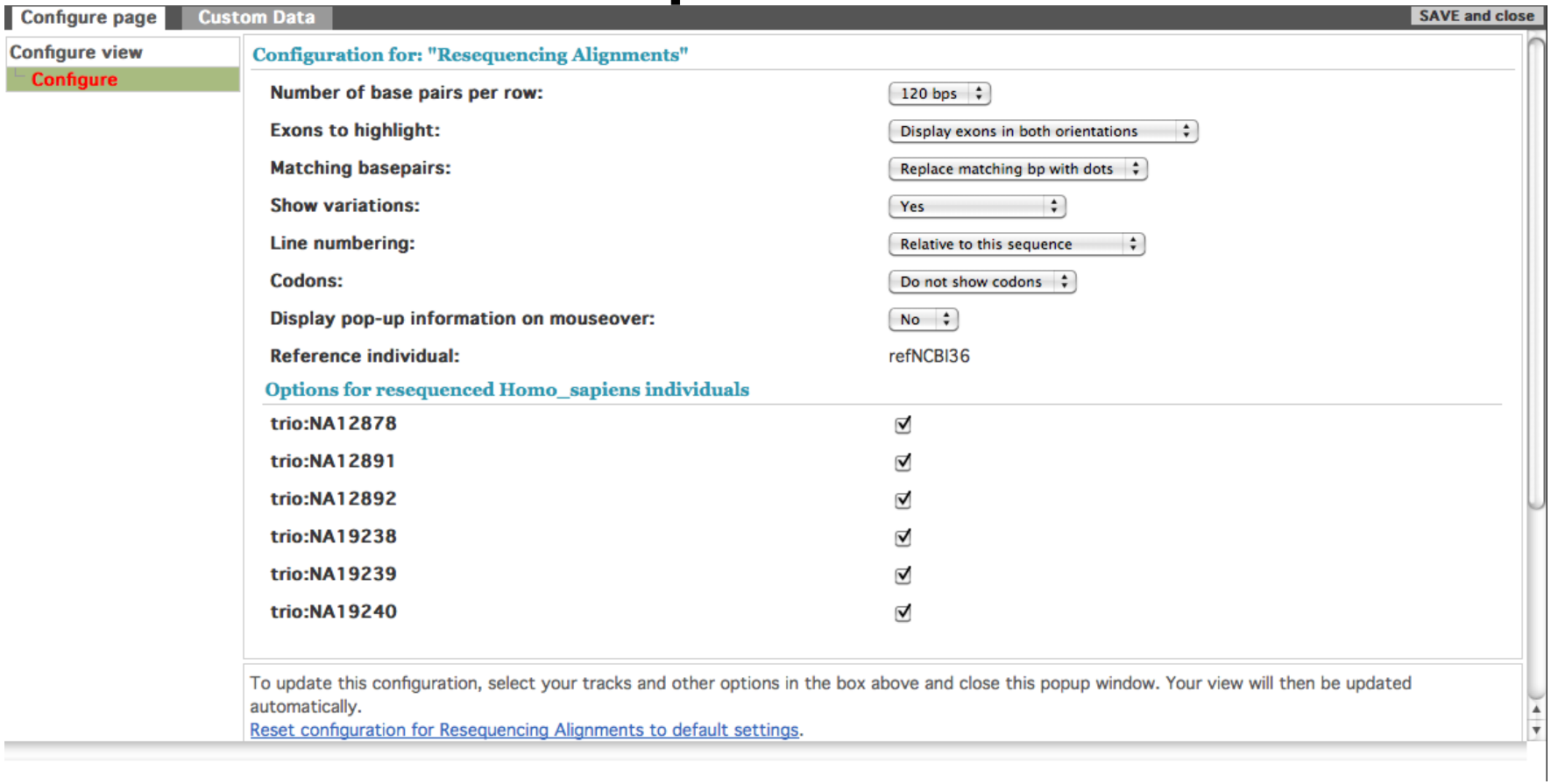
Resequencing alignment

- View any region of the genome in alignment with reference for the 6 high coverage trio individuals
- Assumption made that if there is sequence coverage and not a SNP called, the base is the same as the reference
  - Not a de novo assembly
- Use “Resequencing” link on the left side of pages to access view



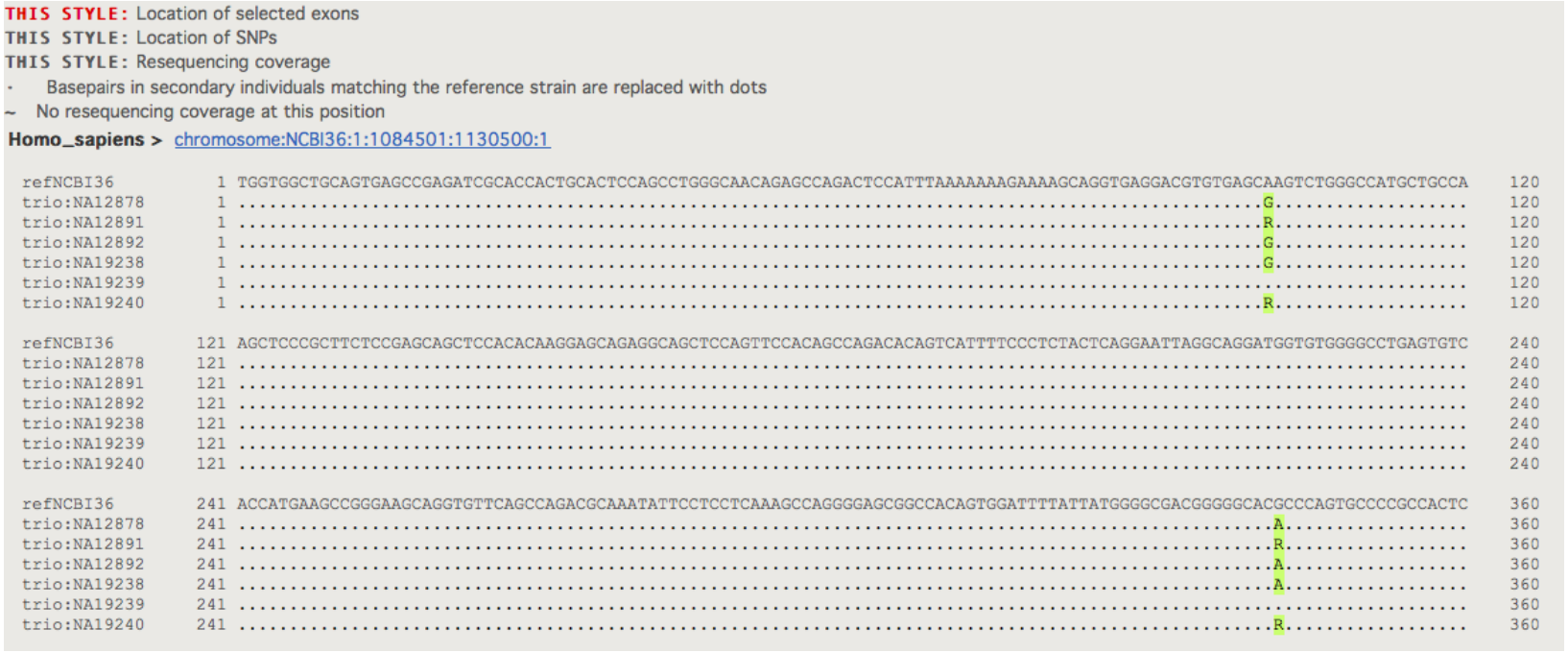
http://browser.1000genomes.org

Resequencing alignment options



http://browser.1000genomes.org

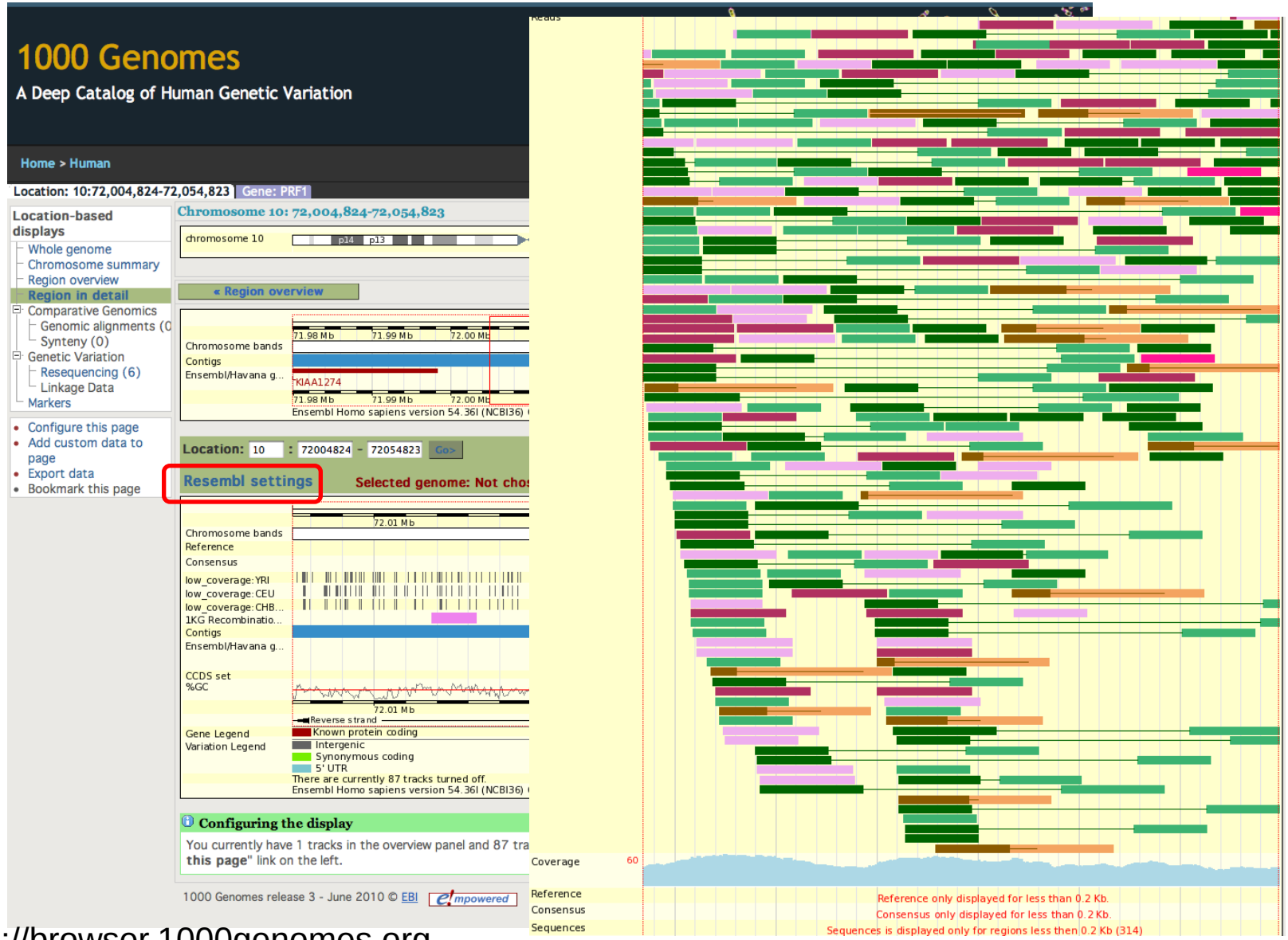
Resequencing alignment output



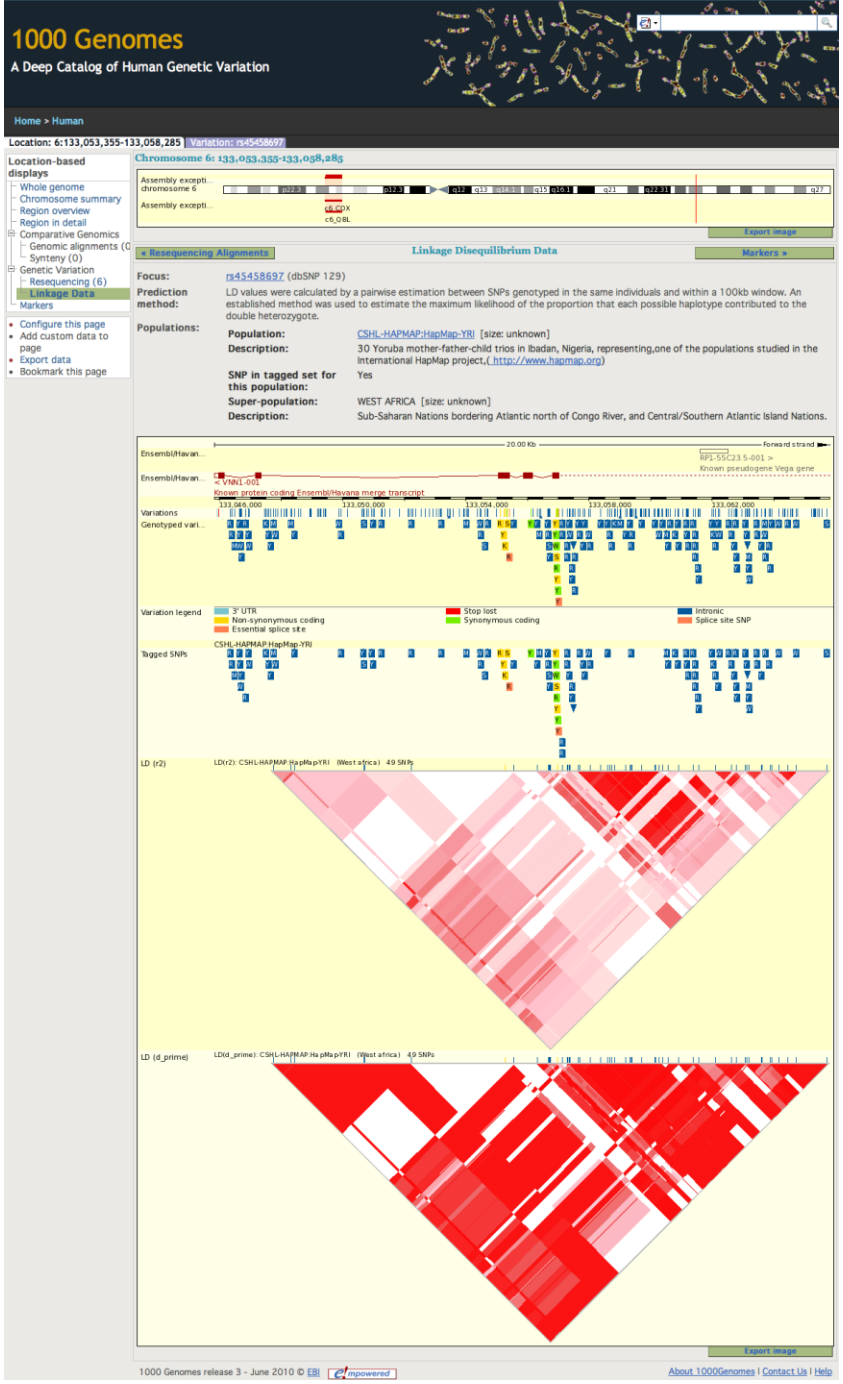
- SNPs are annotated
  - Heterozygous SNPs with ambiguity codes
- “~” means no coverage

http://browser.1000genomes.org

Sequence level data (trios only)



http://browser.1000genomes.org

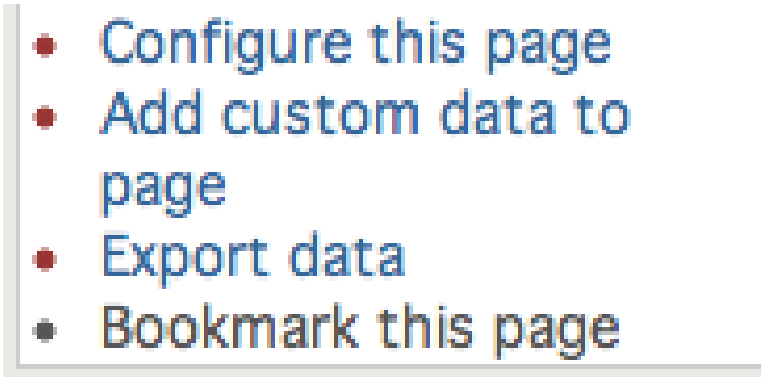


LD Information

- Currently based on data from HapMap and Perlegen populations
- Populations selectable from drop down tab

http://browser.1000genomes.org

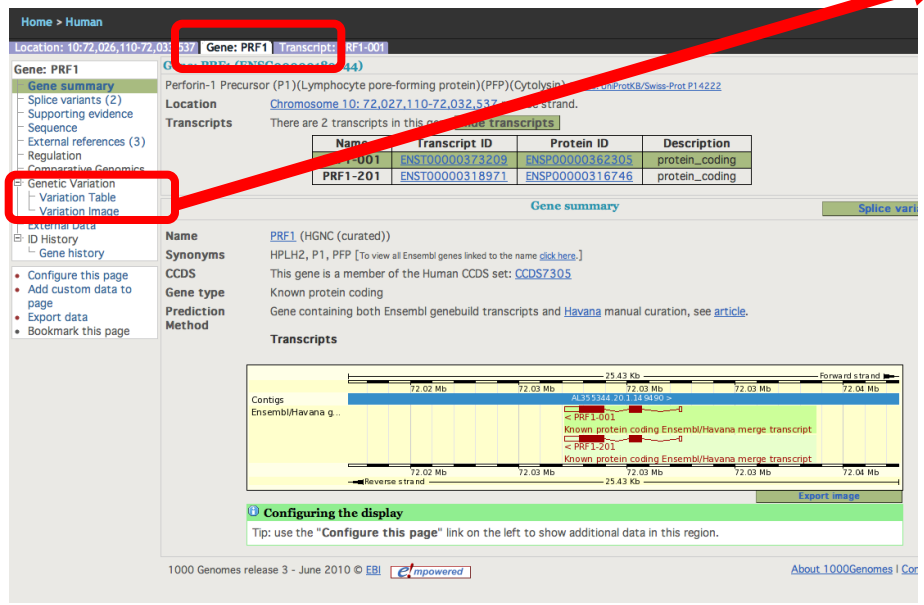
Data Export

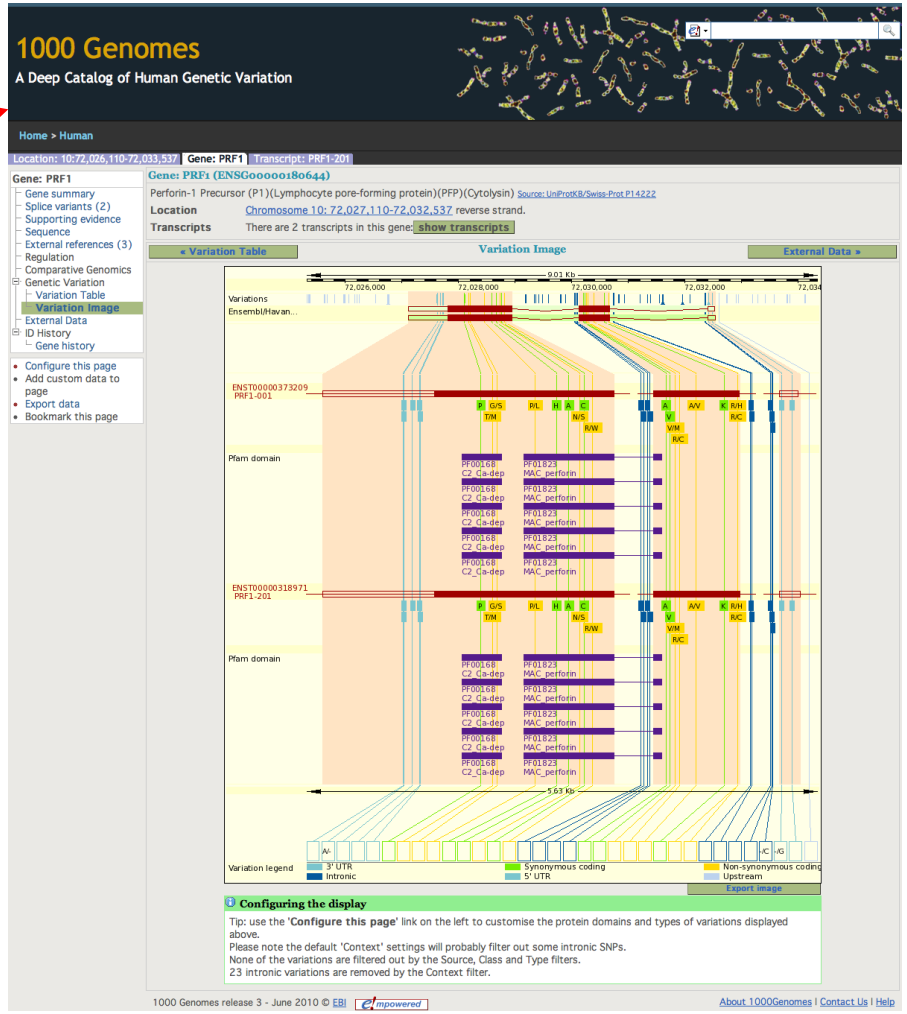


- Summary data from the region being viewed can be exported
- Additional export configuration options coming soon

More variation displays

Gene Page

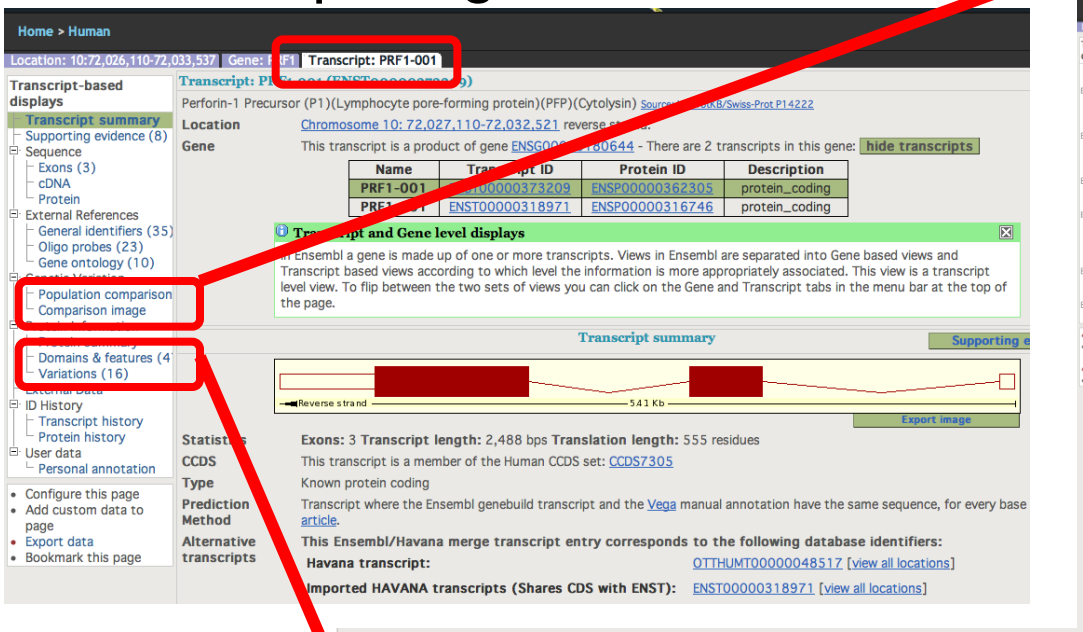


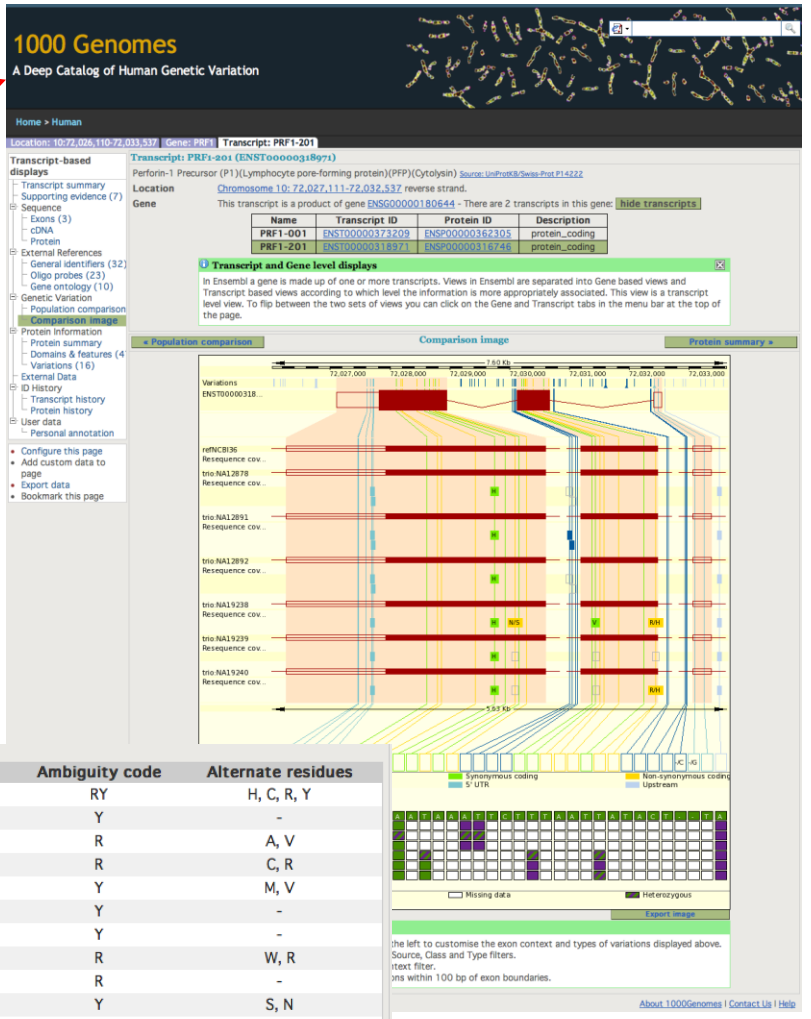


http://browser.1000genomes.org

More variation displays

Transcript Page





Residue	SNP ID	SNP type	Alleles	Ambiguity code	Alternate residues
4	rs12161733	Non-synonymous	A/G	RY	H, C, R, Y
32	rs2228018	Synonymous	T/C	Y	-
91	rs35947132	Non-synonymous	A/G	R	A, V
126	rs34279237	Non-synonymous	A/G	R	C, R
135	rs12263572	Non-synonymous	T/C	Y	M, V
145	trio:YRI:10:72030230	Synonymous	T/C	Y	-
154	low_coverage:YRI:10:72030203	Synonymous	T/C	Y	-
225	rs28933973	Non-synonymous	A/G	R	W, R
242	rs35329429	Synonymous	A/G	R	-
252	rs28933375	Non-synonymous	T/C	Y	S, N
274	rs885821	Synonymous	A/G	R	-
300	rs885822	Synonymous	A/G	R	-
345	rs28933374	Non-synonymous	A/G	R	P, L
426	rs1042652	Non-synonymous	T/C	Y	S, G
435	rs28933376	Non-synonymous	A/G	R	T, M
459	rs2228019	Synonymous	A/G	R	-

http://browser.1000genomes.org

Credits

- Eugene Kulesha, Glenn Proctor
- Natassa Spiridou
- Will McLaren, Fiona Cunningham
- Laura Clarke, Holly Zheng-Bradley, Rick Smith

- info@1000genomes.org