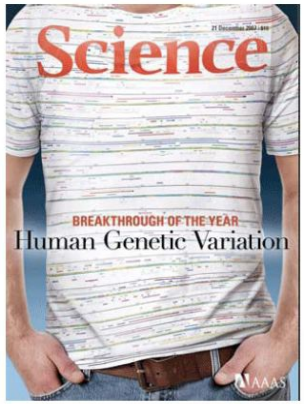
Genomics & Medicine

http://biochem118.stanford.edu/

Single Nucleotide Polymorphisms (SNPs)

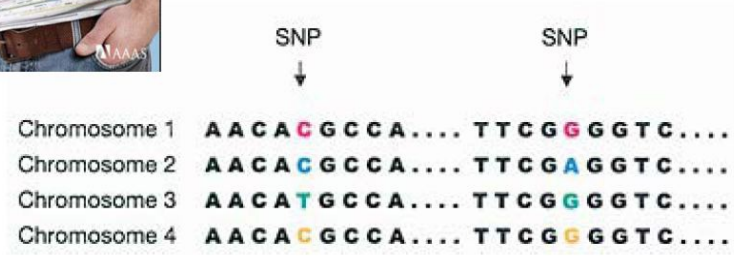
2007 SCIENTIFIC BREAKTHROUGH OF THE YEAR

Science Magazine, December 21, 2007



"It's all about me!"

Single Nucleotide Polymorphisms (SNPs)



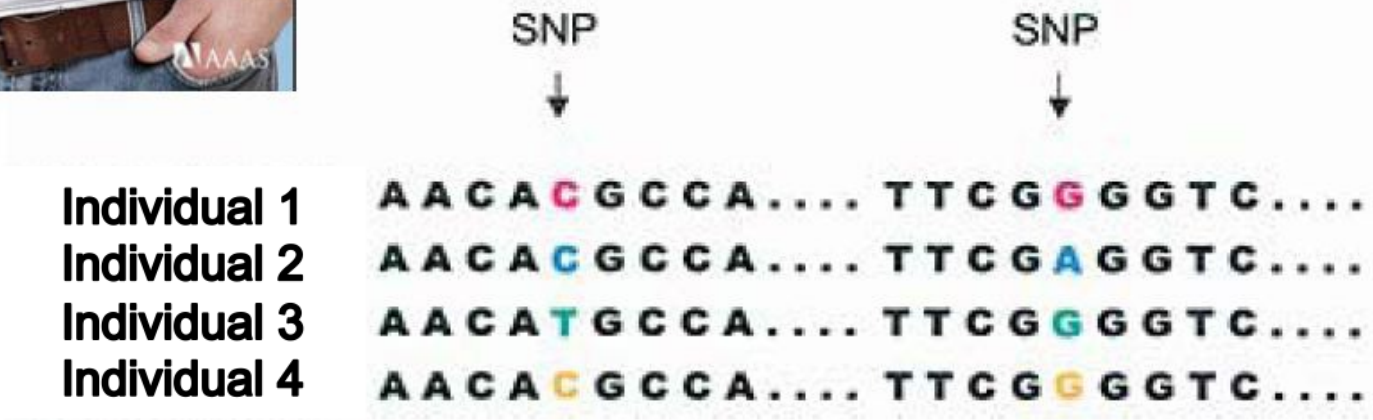
Doug Brutlag, Professor Emeritus of Biochemistry & Medicine (by courtesy)
Stanford University School of Medicine

2007 SCIENTIFIC BREAKTHROUGH OF THE YEAR

Science Magazine, December 21, 2007

"It's all about me!"

Single Nucleotide Polymorphisms (SNPs)

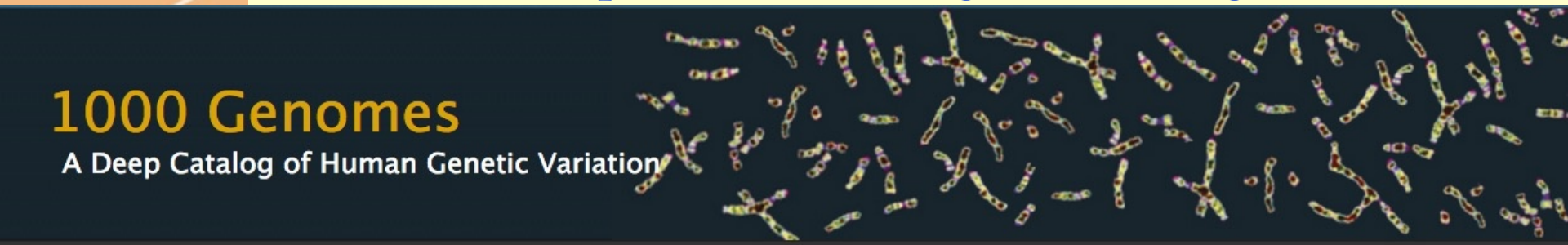


International HapMap Project

http://www.hapmap.org/

Thousand Genomes Project

http://www.1000genomes.org/



Home About Data Analysis Participants Contact Browser Wiki FTP search Search

LATEST ANNOUNCEMENTS

WEDNESDAY SEPTEMBER 30, 2015

A global reference for human genetic variation

The Phase 3 publication, A global reference for human genetic variation and the Phase 3 Structural variation publication, An integrated map of structural variation in 2,504 human genomes are now available from Nature alongside a celebration of 25 years of the Human Genome Project

The variants from the Phase 3 analysis are available in ftp/release/20130502/ and extended information about the SV dataset can be found in ftp/phase3/integrated_sv_map/.

Both these papers are open access and should be free for everyone to read and download.

If you have any questions about the data these papers are based on or how to access it please email info@1000genomes.org

http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/

Recent project announcements

FRIDAY OCTOBER 16, 2015

GRCh38 mapping of the Illumina Platinum Genomes CEU pedigree

NAVIGATION

- Frequently Asked Questions

LINKS

All Project Announcements

Sample and Project Information

Media Archive

Find the 1000 Genomes Project Publications

1000 Genomes
A Deep Catalog of Human Genetic Variation

A Global Reference for Human Genetic Variation

http://www.nature.com/nature/journal/v526/n7571/full/nature15393.html

A global reference for human genetic variation

The 1000 Genomes Project Consortium*

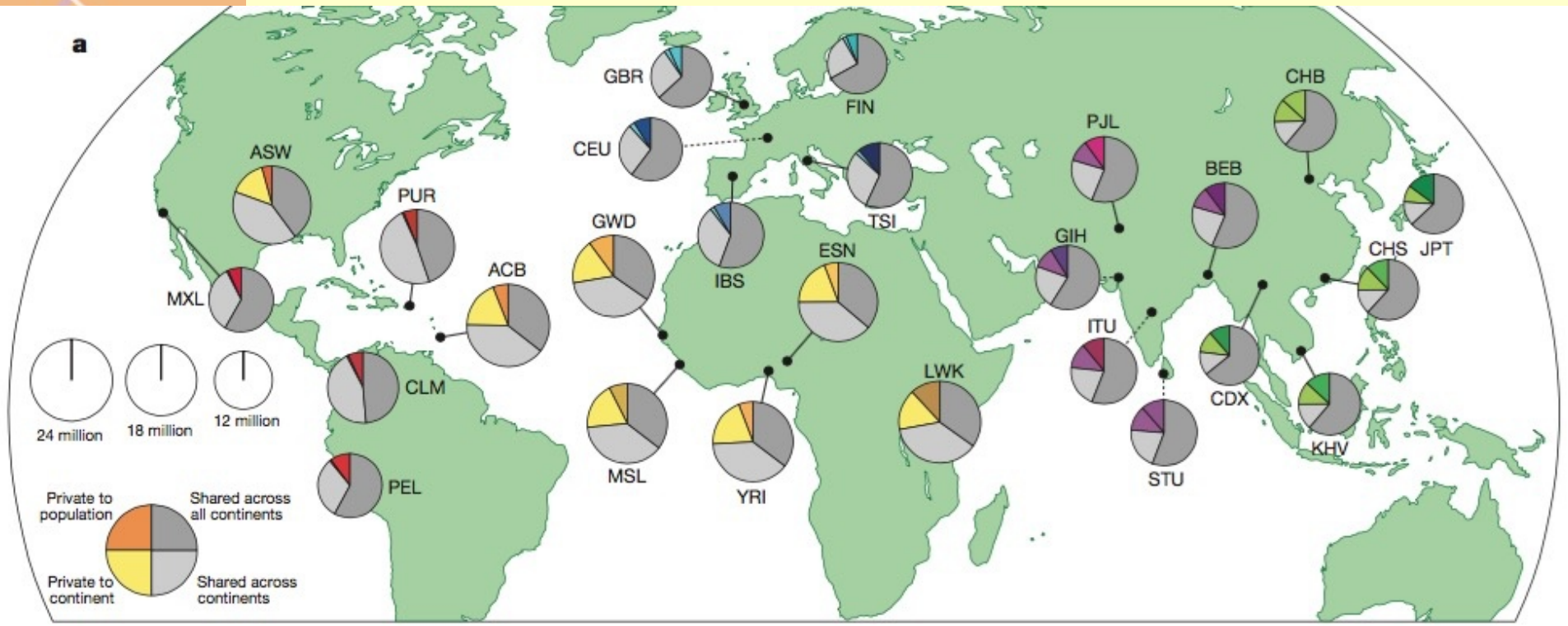
The 1000 Genomes Project set out to provide a comprehensive description of common human genetic variation by applying whole-genome sequencing to a diverse set of individuals from multiple populations. Here we report completion of the project, having reconstructed the genomes of 2,504 individuals from 26 populations using a combination of low-coverage whole-genome sequencing, deep exome sequencing, and dense microarray genotyping. We characterized a broad spectrum of genetic variation, in total over 88 million variants (84.7 million single nucleotide polymorphisms (SNPs), 3.6 million short insertions/deletions (indels), and 60,000 structural variants), all phased onto high-quality haplotypes. This resource includes >99% of SNP variants with a frequency of >1% for a variety of ancestries. We describe the distribution of genetic variation across the global sample, and discuss the implications for common disease studies.

Nature **526, 68-74 (October 1, 2015)**

1000 Genomes
A Deep Catalog of Human Genetic Variation

A Global Reference for Human Genetic Variation

http://www.nature.com/nature/journal/v526/n7571/full/nature15393.html



Nature **526, 68-74 (October 1, 2015)**

An Integrated Map of Structural Variation in 2,504 Human Genomes

An integrated map of structural variation in 2,504 human genomes

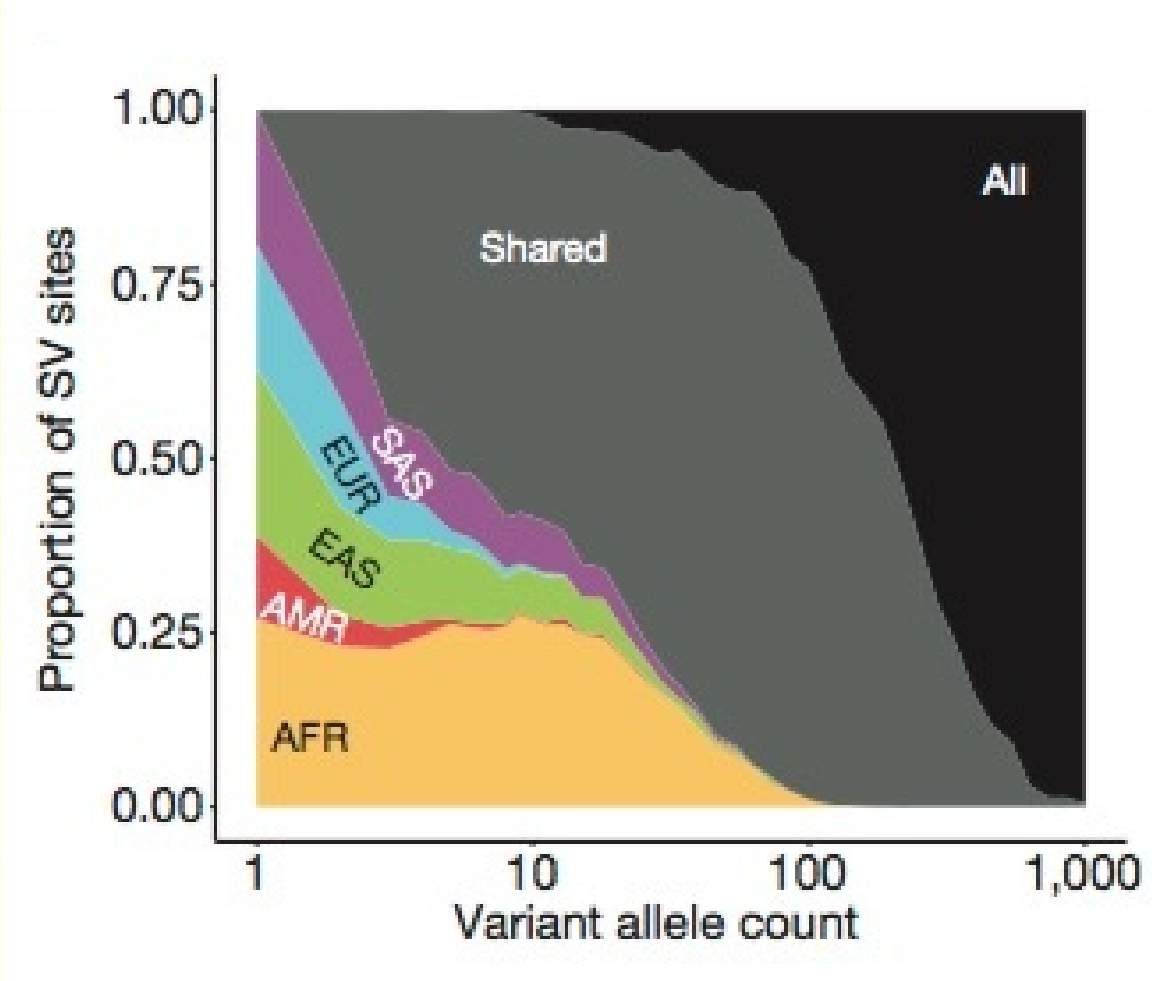
A list of authors and their affiliations appears at the end of the paper.

Structural variants are implicated in numerous diseases and make up the majority of varying nucleotides among human genomes. Here we describe an integrated set of eight structural variant classes comprising both balanced and unbalanced variants, which we constructed using short-read DNA sequencing data and statistically phased onto haplotype blocks in 26 human populations. Analysing this set, we identify numerous gene-intersecting structural variants exhibiting population stratification and describe naturally occurring homozygous gene knockouts that suggest the dispensability of a variety of human genes. We demonstrate that structural variants are enriched on haplotypes identified by genome-wide association studies and exhibit enrichment for expression quantitative trait loci. Additionally, we uncover appreciable levels of structural variant complexity at different scales, including genic loci subject to clusters of repeated rearrangement and complex structural variants with multiple breakpoints likely to have formed through individual mutational events. Our catalogue will enhance future studies into structural variant demography, functional impact and disease association.

Nature **526, 75-81 (October 1, 2015)**

1000 Genomes
A Deep Catalog of Human Genetic Variation

An Integrated Map of Structural Variation in 2,504 Human Genomes



Nature **526, 75-81 (October 1, 2015)**

NHS 100,000 Genomics England Project

http://www.genomicsengland.co.uk/

Genomics england

Home | About Genomics England | The 100,000 Genomes Project | GeCIP | Library and resources | News | Contact us

Genomics England, with the consent of participants and the support of the public, is creating a lasting legacy for patients, the NHS and the UK economy through the sequencing of 100,000 genomes: the 100,000 Genomes Project.

Genomics England was set up by the Department of Health to deliver the 100,000 Genomes Project. Initially the focus will be on rare disease, cancer and infectious disease. The project is currently in its pilot phase and will be completed by the end of 2017.

Read more...

NIH Precision Medicine Initiative

http://www.nih.gov/precisionmedicine/

National Institutes of Health
Turning Discovery Into Health

Search

For Employees | Staff Directory | En Español

Health Information | Grants & Funding | News & Events | Research & Training | Institutes at NIH | About NIH

NIH Home > Research & Training

PRECISION MEDICINE INITIATIVE

Precision Medicine Initiative

- What are the near-term goals?
- What are the longer-term goals?
- How is it different?
- Who will participate?
- NIH Workshop

Precision Medicine Initiative

Far too many diseases do not have a proven means of prevention or effective treatments. We must gain better insights into the biology of these diseases to make a difference for the millions of Americans who suffer from them. Precision medicine is an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. While significant advances in precision medicine have been made for select cancers, the practice is not currently in use for most diseases. Many efforts are underway to help make precision medicine the norm rather than the exception. To accelerate the pace, President Obama has now unveiled the Precision Medicine Initiative – a bold new enterprise to revolutionize medicine and generate the scientific evidence needed to move the concept of precision medicine into every day clinical practice.

Email Updates

To sign up for updates please enter your e-mail address.

Submit

Related Links

- NEJM Perspective: A New Initiative on Precision Medicine
- White House Precision Medicine Web Page
- White House Fact Sheet: President Obama's Precision Medicine Initiative
- Precision Medicine Initiative and Cancer Research

Single Nucleotide Polymorphisms (SNPs) in the Human Genome

GCTGTATGACTAGAAGATCGAT
GCTGTATGACGAGAAGATCGAT

- About 85 million sites in the human genome where sequence variations have occurred
- About 30 million sites where variation exceeds 1% of a particular population (MAF > 1%)
- Each ethnicity has its own distribution of SNPs
- About 4.1 to 5 million sites where any individual varies from the reference human genome.
- Each person differs from others in 4.1 million places (about 0.14% of the genome)
- SNP sequence variations are common, unlike disease causing mutations which are rare.

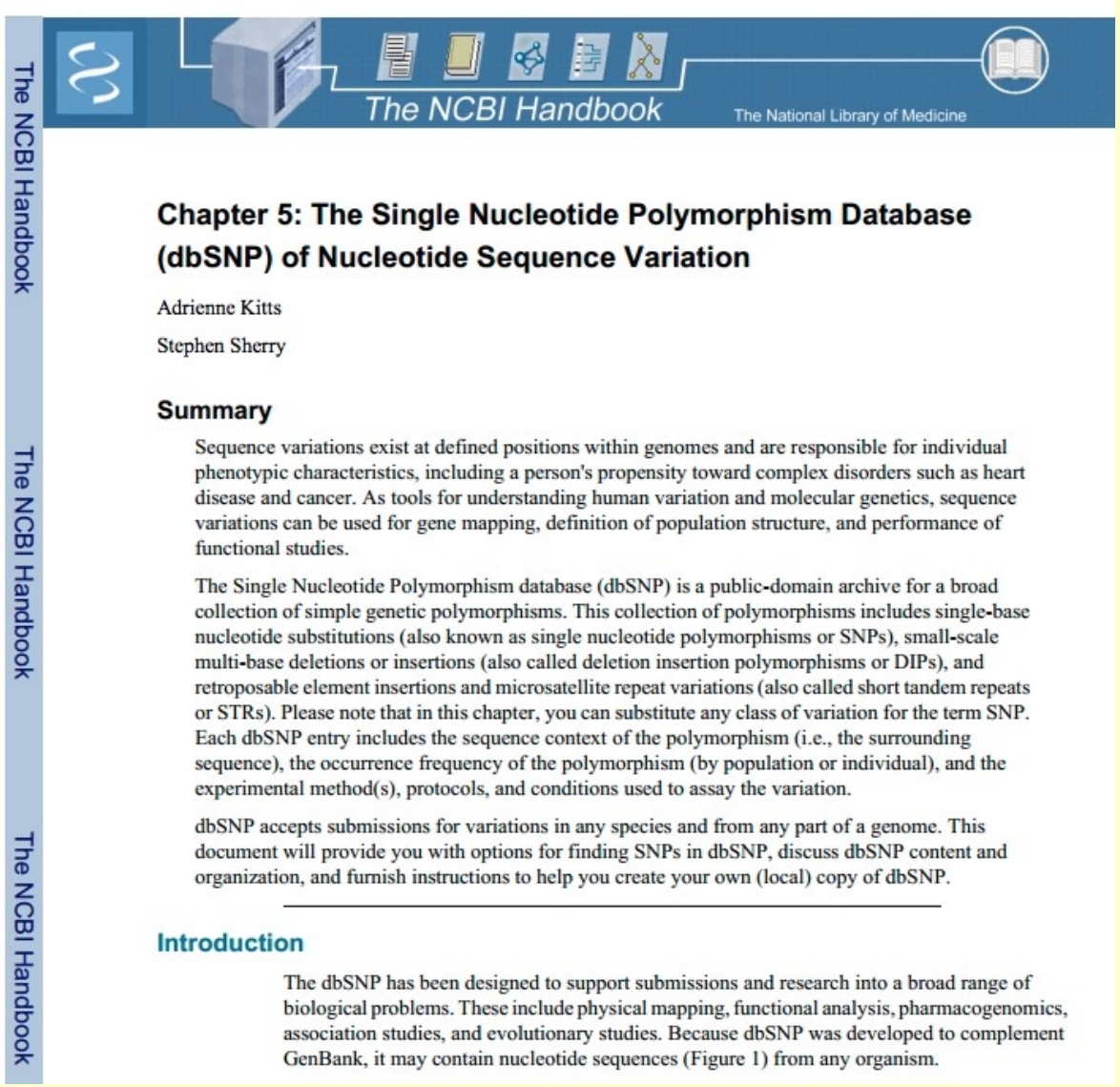
Single Nucleotide Polymorphisms (SNPs) in the Human Genome

GCTGTATGACTAGAAGATCGAT
GCTGTATGACGAGAAGATCGAT

- SNPs can be used for identifying individuals and forensics
- SNPs are used for mapping & genome-wide association studies of complex diseases
- SNPs are used for ancestry tracking & family relationships
- SNPs are used to predict risk of common genetic diseases
- SNPs are used for classifying patients in clinical trials
- SNPs are used to predict drug sensitivity an adverse reactions
- SNPs are used for personalized medicine & pharmacogenomics
- While SNPs are linked with disease, they do not cause disease
- In short, SNPs are used as genetic markers

The dbSNP Database

http://www.ncbi.nlm.nih.gov/books/NBK21088/pdf/ch5.pdf



The NCBI Handbook

The National Library of Medicine

Chapter 5: The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation

Adrienne Kitts

Stephen Sherry

Summary

Sequence variations exist at defined positions within genomes and are responsible for individual phenotypic characteristics, including a person's propensity toward complex disorders such as heart disease and cancer. As tools for understanding human variation and molecular genetics, sequence variations can be used for gene mapping, definition of population structure, and performance of functional studies.

The Single Nucleotide Polymorphism database (dbSNP) is a public-domain archive for a broad collection of simple genetic polymorphisms. This collection of polymorphisms includes single-base nucleotide substitutions (also known as single nucleotide polymorphisms or SNPs), small-scale multi-base deletions or insertions (also called deletion insertion polymorphisms or DIPs), and retroposable element insertions and microsatellite repeat variations (also called short tandem repeats or STRs). Please note that in this chapter, you can substitute any class of variation for the term SNP. Each dbSNP entry includes the sequence context of the polymorphism (i.e., the surrounding sequence), the occurrence frequency of the polymorphism (by population or individual), and the experimental method(s), protocols, and conditions used to assay the variation.

dbSNP accepts submissions for variations in any species and from any part of a genome. This document will provide you with options for finding SNPs in dbSNP, discuss dbSNP content and organization, and furnish instructions to help you create your own (local) copy of dbSNP.

Introduction

The dbSNP has been designed to support submissions and research into a broad range of biological problems. These include physical mapping, functional analysis, pharmacogenomics, association studies, and evolutionary studies. Because dbSNP was developed to complement GenBank, it may contain nucleotide sequences (Figure 1) from any organism.

The dbSNP Database

http://www.ncbi.nlm.nih.gov/books/NBK174586/

The Database of Short Genetic Variation (dbSNP)

Kitts A, Phan L, Ward M, et al.

Scope

Sequence variation is of scientific interest to population geneticists, genetic mappers, and those investigating relationships among variation and phenotype. These variations can be of several types, from simple substitutions that do not affect sequence length, to those that result in minor length differences, to those that affect multiple genes and multiple chromosomes. Variations can also be categorized with respect to their frequency within a population, from a variation with a single allele to a variation that is highly polymorphic.

Although SNP is the abbreviation for "single nucleotide polymorphism," dbSNP is a public archive of all short sequence variation, not just single nucleotide substitutions that occur frequently enough in a population to be termed polymorphic. dbSNP includes a broad collection of simple genetic variations such as single-base nucleotide substitutions, small-scale multi-base deletions or insertions, and microsatellite repeats. Data submitted to dbSNP can be from any organism, from any part of a genome, and can include genotype and allele frequency data if those data are available. dbSNP accepts submissions for all classes of simple sequence variation, and provides access to variations of germline or somatic origin that are clinically significant.

In order to emphasize the comprehensive nature of dbSNP's content, the full name of the database was changed from "database of Single Nucleotide Polymorphism" to the more inclusive "database of Short Genetic Variation" in July of 2011. The acronym that represents the database will remain "dbSNP" to avoid any confusion that might arise from a complete name change.

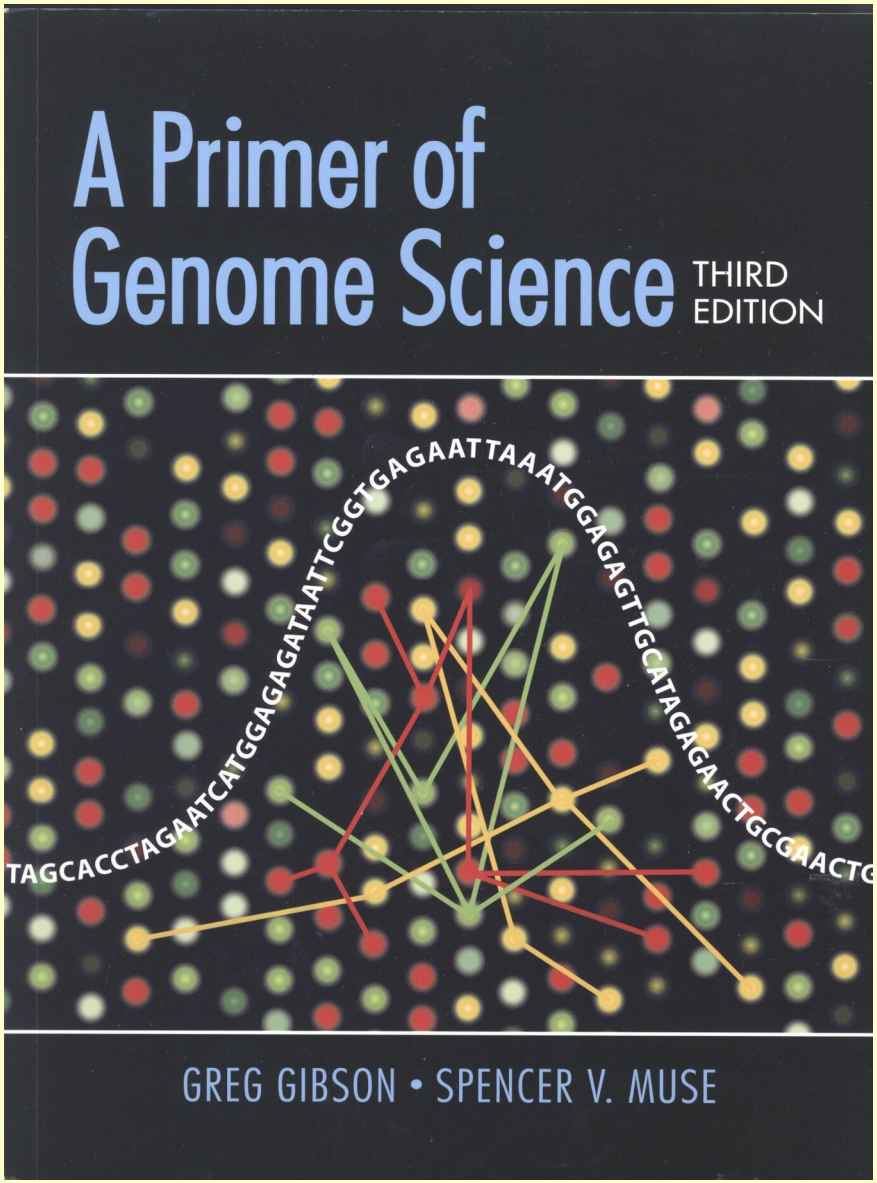
Each record in dbSNP includes the sequence context of the variant, the frequency of the polymorphism in a population if available, its zygosity if available, and the experimental method(s), protocols, and conditions used to assay the variation by each submitter. Individual submissions are clustered into dbSNP reference records (rs#) that contain summary data which may include clinical significance from ClinVar, association with phenotype from dbGaP, variation false positive status, allele origin (germline or somatic), and submitter attributes.

The dbSNP has been designed to support submissions and research into a broad range of biological problems that include the identification of genotype-phenotype relationships, genetic and physical mapping, functional analysis, pharmacogenomics, and association studies.

Medical Genetics

Advances in next-generation sequencing technologies allow

A Primer of Genome Science
Chapter 3 Genomic Variation

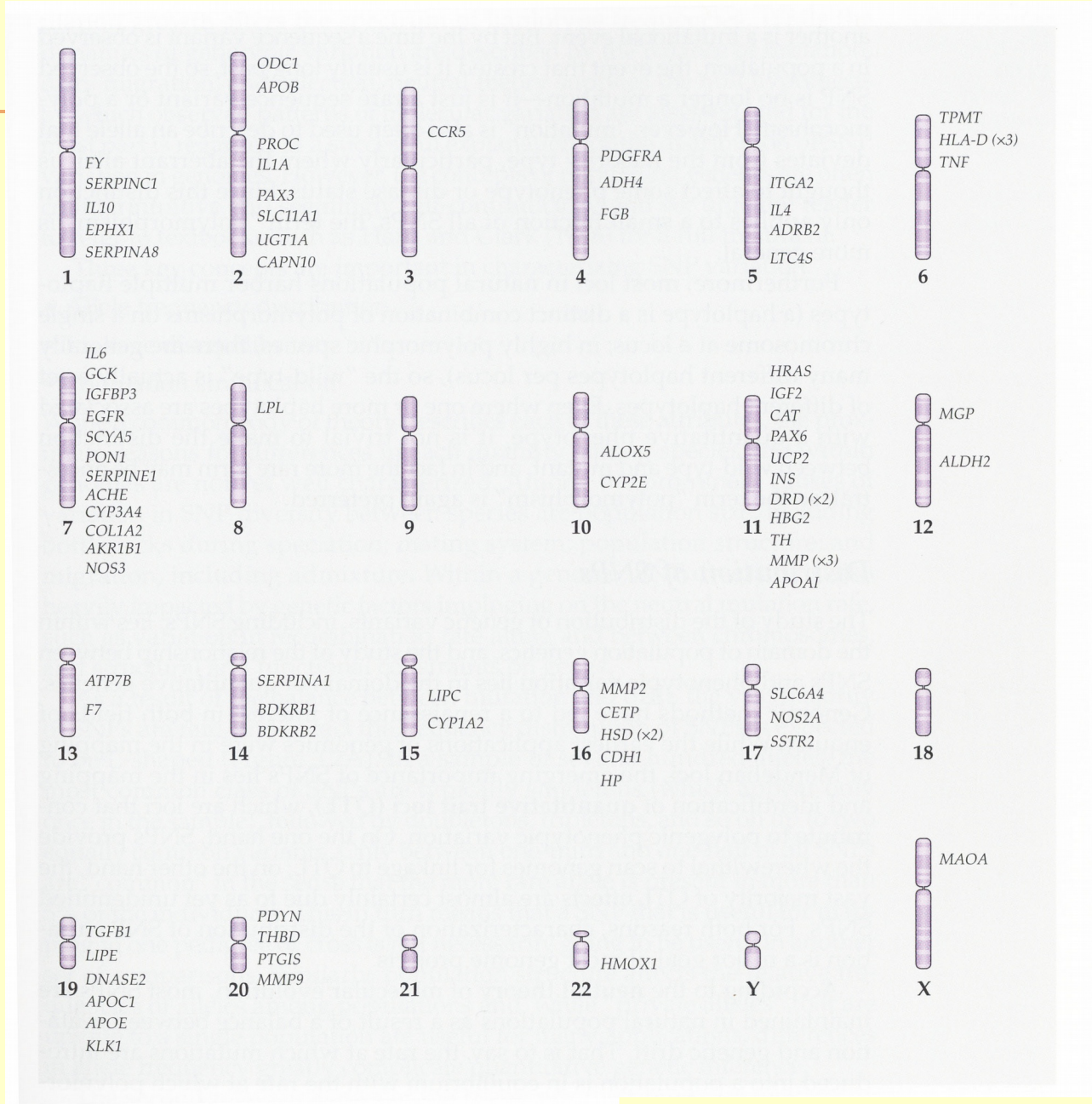


Types of SNPs

http://www.ncbi.nlm.nih.gov/books/NBK174586/

- Noncoding SNPs
  - Promoters
  - 5′ UTR
  - 3′ UTR
  - Introns
  - Intergenic Regions
  - Pseudogenes
  - Regulatory
    - Splicing
    - Transcriptional regulation (promoter & transcription factor binding sites)
    - Translational regulation (initiation or termination)
    - Regulatory miRNA target sites
- Coding SNPs
  - Synonymous SNPs (third position variation)
  - Replacement SNPs (change amino acid)
    - Functional SNPs (acceptable amino acid replacement)
    - Non-functional SNPs (traits & diseases)

SNPs in Human Promoters that Cause Disease



Human βHemoglobin Gene

http://www.ncbi.nlm.nih.gov/gene/3043

NCBI Resources How To brutlag My NCBI

Gene Gene Search

Limits Advanced

Display Settings: Full Report Send to:

HBB hemoglobin, beta [*Homo sapiens*]

Gene ID: 3043, updated on 14-Oct-2012

Summary

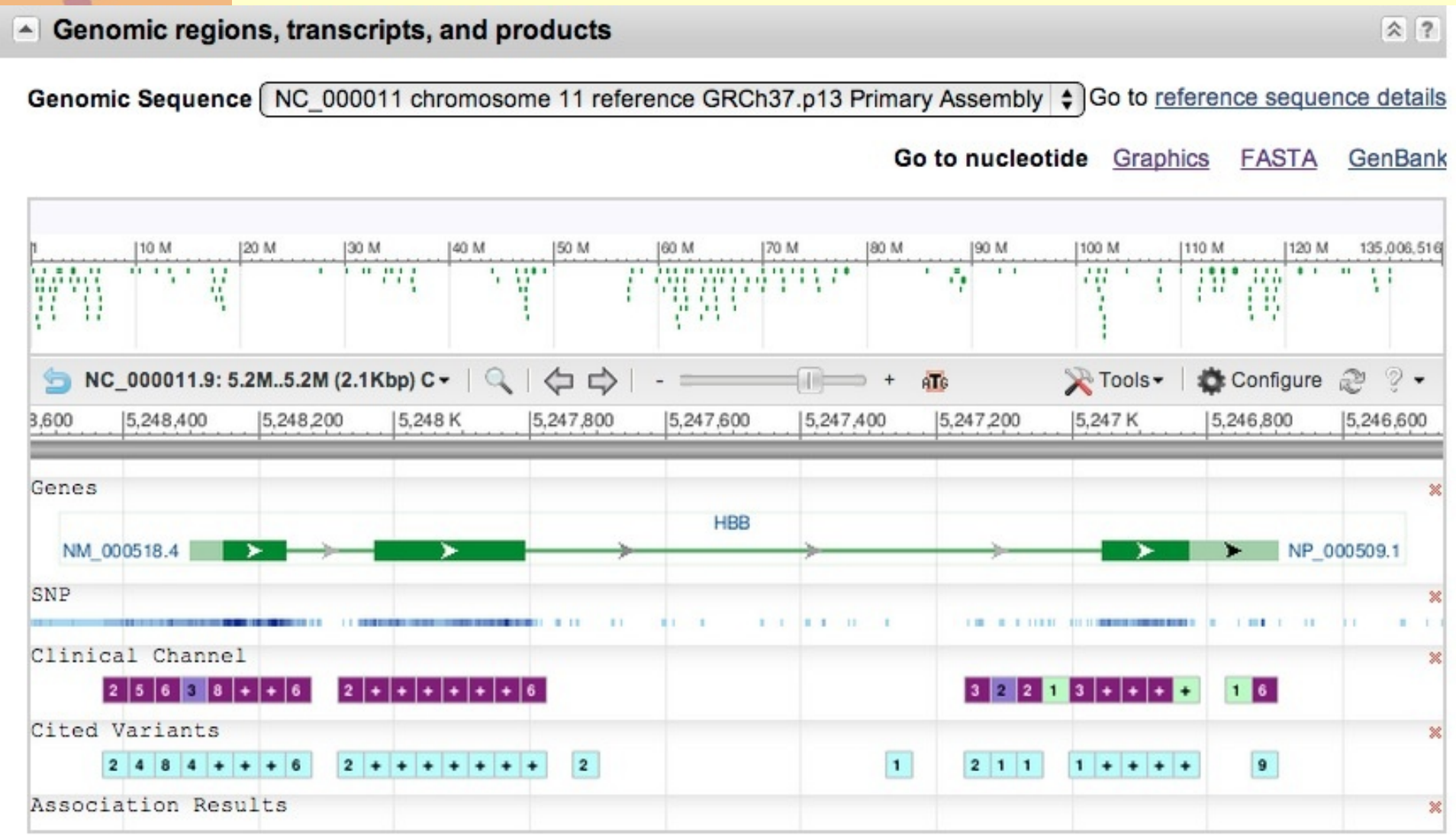
Official Symbol	HBB provided by HGNC
Official Full Name	hemoglobin, beta provided by HGNC
Primary source	HGNC:4827
See related	Ensembl:ENSG00000244734; HPRD:00786; MIM:141900; Vega:OTTHUMG00000066678
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	CD113t-C; beta-globin
Summary	The alpha (HBA) and beta (HBB) loci determine the structure of the 2 types of polypeptide chains in adult hemoglobin, Hb A. The normal adult hemoglobin tetramer consists of two alpha chains and two beta chains. Mutant beta globin causes sickle cell anemia. Absence of beta chain causes beta-zero-thalassemia. Reduced amounts of detectable beta globin causes beta-plus-thalassemia. The order of the genes in the beta-globin cluster is 5'-epsilon -- gamma-G -- gamma-A -- delta -- beta--3'. [provided by RefSeq, Jul 2008]

Table of contents

Related information

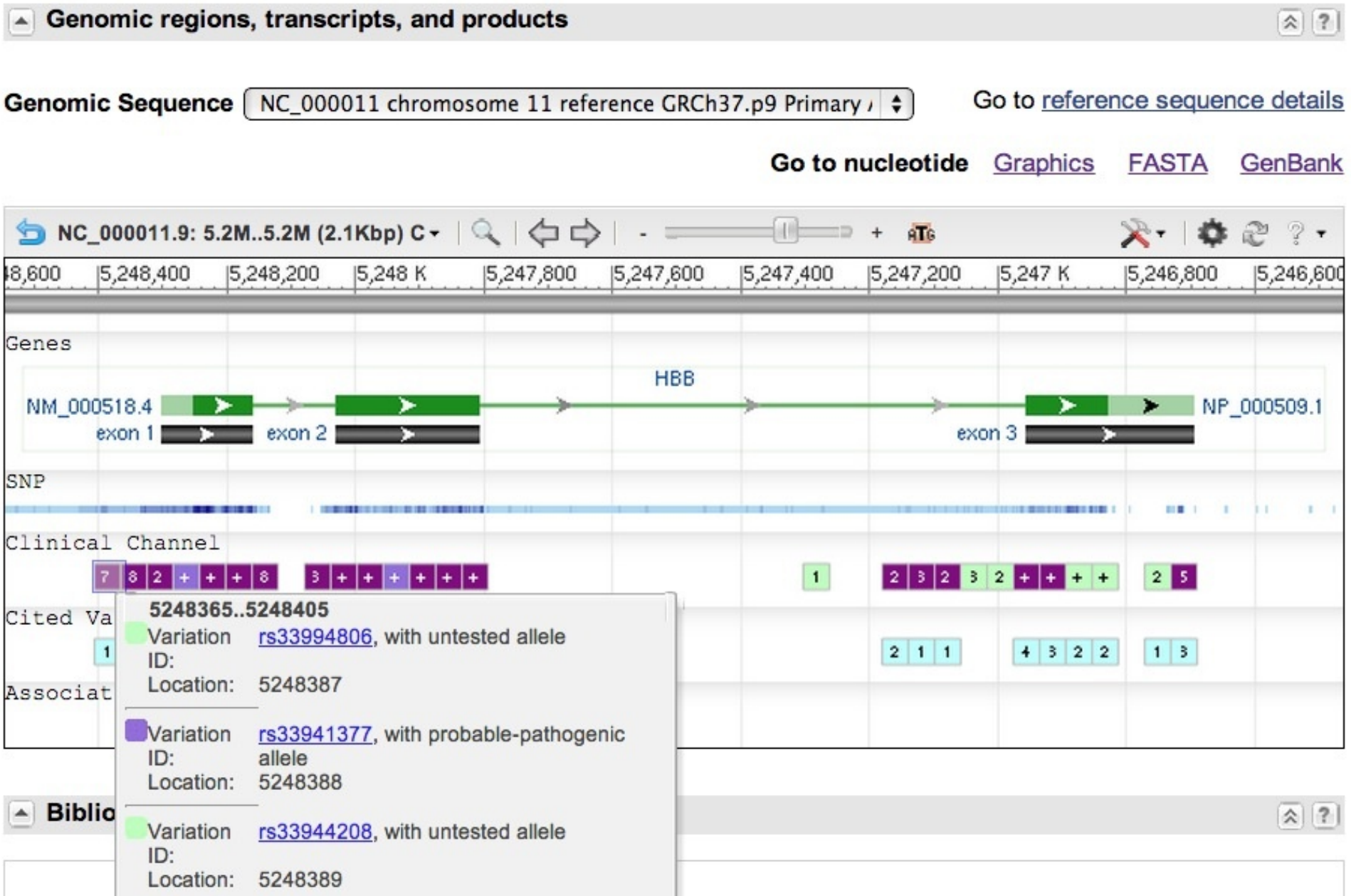
Human βHemoglobin Gene

http://www.ncbi.nlm.nih.gov/gene/3043



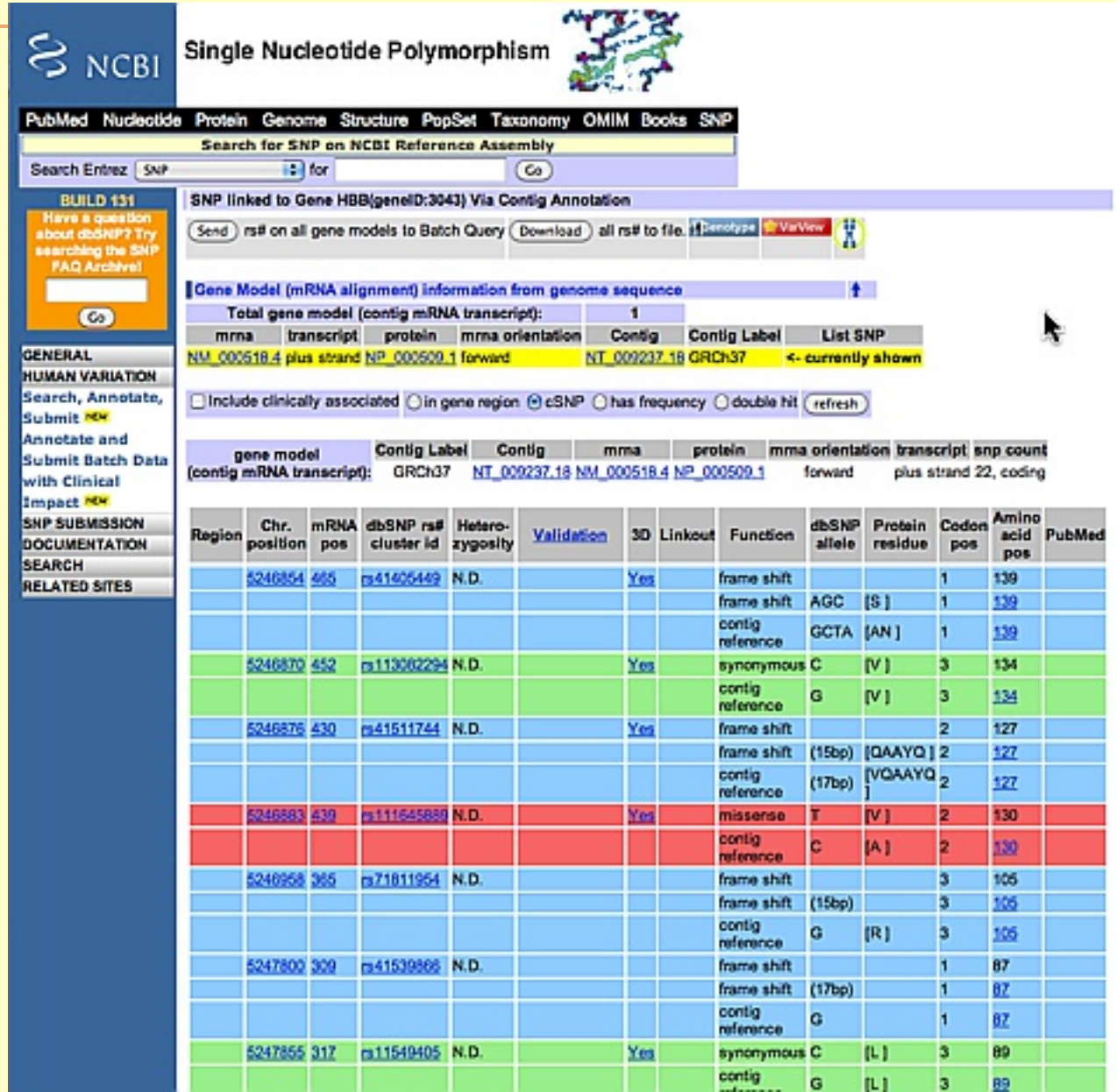
Human βHemoglobin Gene

http://www.ncbi.nlm.nih.gov/gene/3043



Human βHemoglobin Gene SNPs

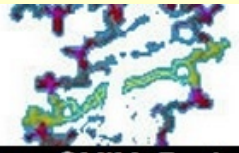
http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=3043



βHemoglobin Gene SNP rs11549406

http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=11549406

dbSNP
Short Genetic Variations



Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search for SNP on NCBI Reference Assembly

for Go

Reference SNP(refSNP) Cluster Report: rs11549406

RefSNP	
Organism:	human (Homo sapiens)
Molecule Type:	cDNA
Created/Updated in build:	120/137
Map to Genome Build:	37.3
Validation Status:	

Allele	
Variation Class:	SNV: single nucleotide variation
RefSNP Alleles:	C/G
Allele Origin:	
Ancestral Allele:	C
Clinical Channel:	unknown
Clinical Significance:	NA
MAF/MinorAlleleCount:	NA
MAF Source:	

HGVS Names
NC_000011.9:g.5247878G>C
NG_000007.3:g.70968C>G
NM_000518.4:c.244C>G
NP_000509.1:p.Leu82Val
NT_009237.18:g.5187878G>C
NW_001838021.1:g.876426G>C

SNP Details are organized in the following sections:

GeneView | Map | Submission | Fasta | Resource | Diversity | Validation

Integrated Maps (Hint: click on 'Chr Pos' or 'Contig Pos' column value to see variation in NCBI sequence viewer)

Assembly	Genome Build	Chr	Chr Pos	Contig	Contig Pos	SNP to Chr	Contig allele	Contig to Chr
GRCh37.p5	37.3	11	5247878	NT_009237.18	5187878	Rev	G	Fwd
reference	36.3	11	5204454	NT_009237.17	4035119	Rev	G	Fwd
Celera	36.3	11	5366549	NW_925006.1	866873	Rev	G	Fwd
HuRef	37.3	11	4907051	NW_001838021.1	876426	Rev	G	Fwd
HuRef	36.3	11	4907051	NW_001838021.1	876426	Rev	G	Fwd

βHemoglobin Gene SNP rs11549406

http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=11549406

The submission **ss16249024** has the longest flanking sequence of all cluster members and was used to instantiate sequence for **rs115**

NCBI Assay ID	Handle|Submitter ID	Validation Status	ss to rs Orientation /Strand	Alleles	5' Near Seq 30 bp	3' Near Seq 30 bp
ss16249024	CGAP-GAI|1496874		fwd/B	C/G	tttagtgatggcctggctcacctggacaactcaagggcacctttgccacactgagtgagc	

Fasta sequence (Legend)

```
>gnl|dbSNP|rs11549406|allelePos=51|totalLen=101|taxid=9606|snpclass=1|alleles='C/G'|mol=cDNA|build=120
GCAAGAAAGT GCTCGGTGCC TTTAGTGATG GCCTGGCTCA CCTGGACAAC
S
TCAAGGGCAC CTTTGCCACA CTGAGTGAGC TGCACTGTGA CAAGCTGCAC
```

NCBI Resource Links

Submitter-Referenced	dbSNP Blast Analysis	UniGene Cluster ID	3D structure mapping
GenBank		523443	NP_000509
BI602394 BG706169			

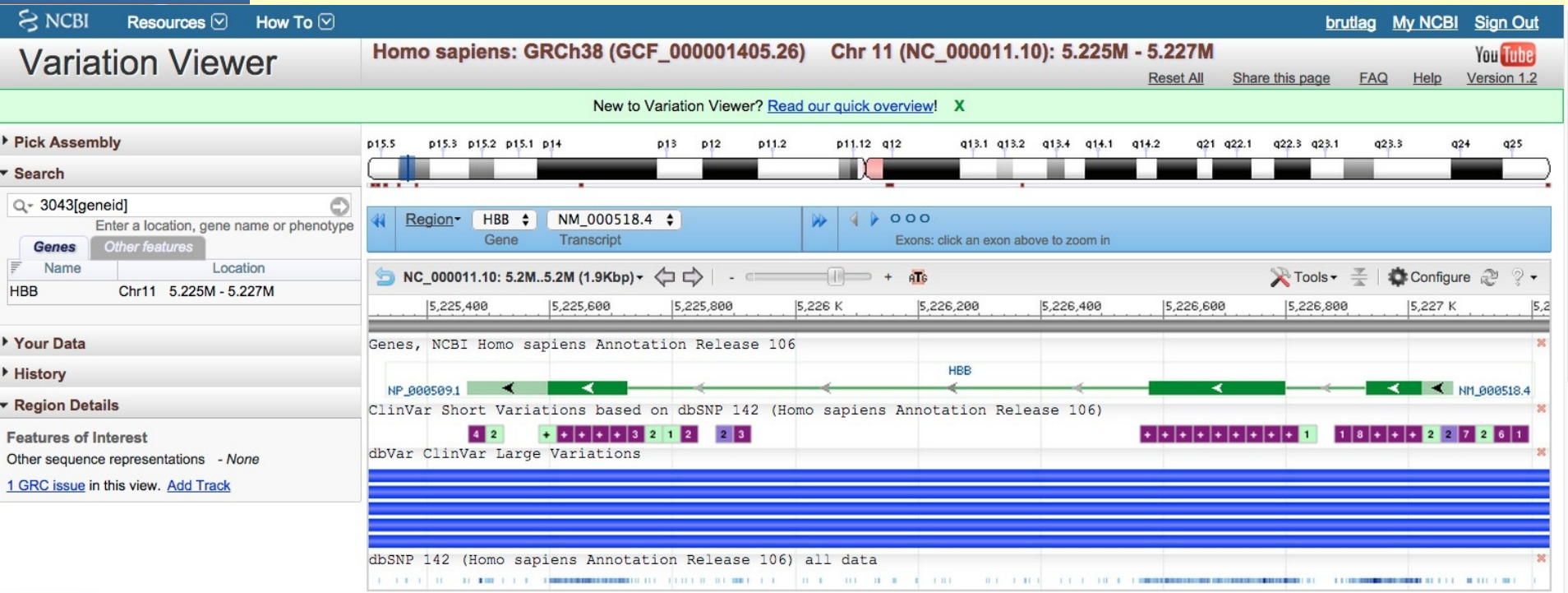
Population Diversity

	Sample Ascertainment				Genotype Detail			Alleles	
ss#	Population	Individual Group	Chrom. Sample Cnt.	Source	C/C	C/G	HWP	C	G
ss16249024	HapMap-CEU	European	112	IG	0.982	0.018	1.000	0.991	0.009
	HapMap-HCB	Asian	86	IG	1.000			1.000	
	HapMap-JPT	Asian	86	IG	1.000			1.000	
	HapMap-YRI	Sub-Saharan African	114	IG	1.000			1.000	
	ENSEMBL_Watson		2	IG	1.000			1.000	
	ENSEMBL_Venter		2	IG	1.000			1.000	

Summary	Average Het.+/- std err:	Individual Count	Founders Count	Individual Overlap	Genotype Conflict
	0.005+/-0.050	527	411	0	0

Human β-Hemoglobin Variation Viewer

http://www.ncbi.nlm.nih.gov/variation/view/



Variation Data

Filter by

Source database
- dbSNP (632)
- dbVar (37)

In ClinVar
- Yes (378)
- No (291)

Worst clinical significance
- Pathogenic (109)
- Likely pathogenic (11)
- drug response (0)
- other (248)
- risk factor (0)

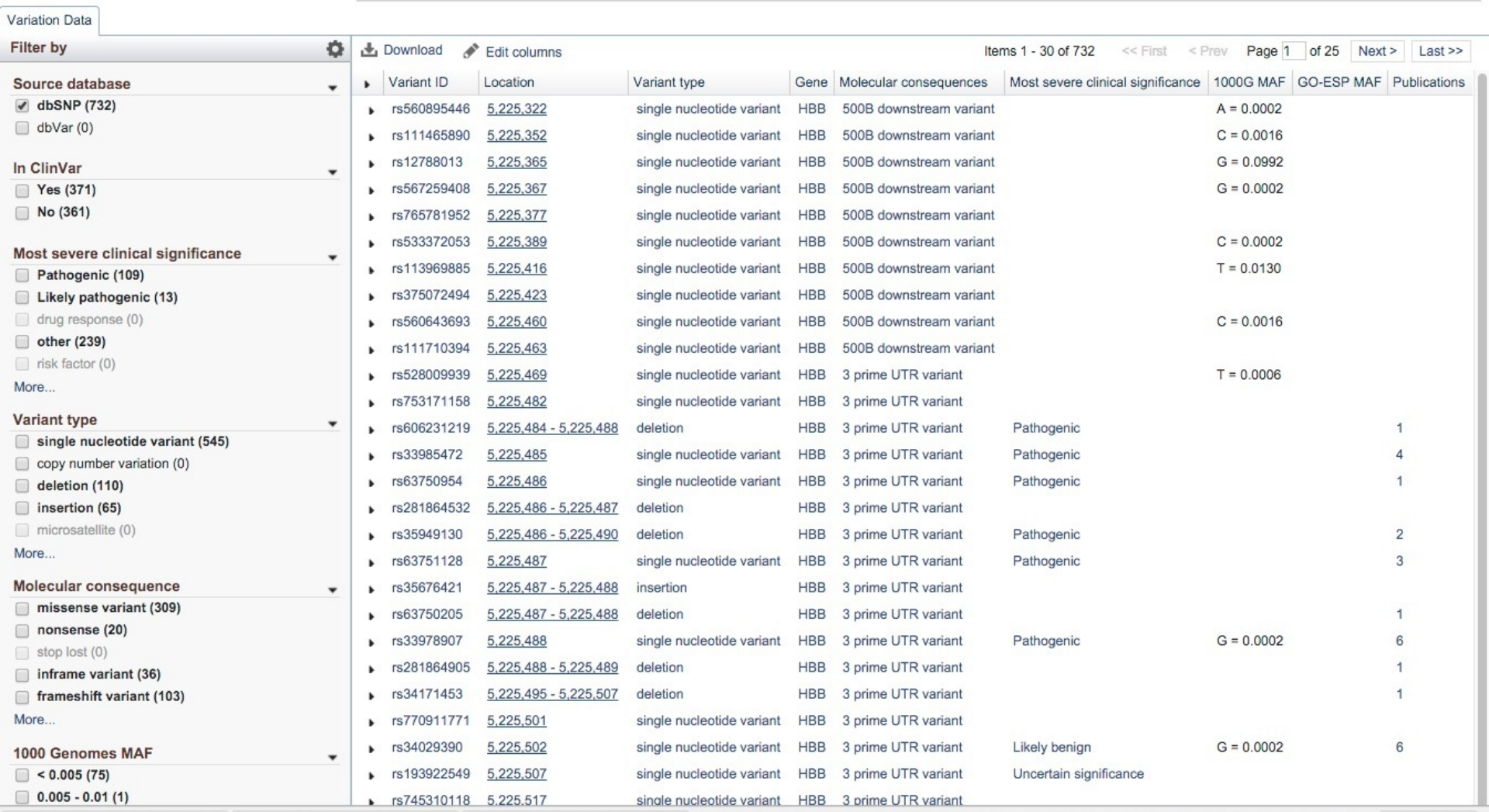
More...

Download | Edit columns | Items 1 - 30 of 669 | << First | < Prev | Page 1 of 23 | Next > | Last >>

Variant ID	Location	Variant type	Gene	Molecular consequences	Worst clinical significance	1000G MAF	GO-ESP MAF	Publications
nsv931147	61,793 - 10,727,969	copy number variation	PNPLA2 and 268 more		Pathogenic			1
nsv915986	196,855 - 5,321,874	copy number variation	PNPLA2 and 153 more		Pathogenic			1
nsv984845	198,510 - 135,074,876	copy number variation	SPTBN2 and 1515 more					1
nsv532276	202,758 - 31,726,224	copy number variation	TRIM5 and 387 more		Pathogenic			1
nsv1054121	205,983 - 6,415,299	copy number variation	TRIM5 and 192 more					1
nsv1048536	205,983 - 17,160,103	copy number variation	TRIM5 and 304 more					1
nsv1037023	205,983 - 30,840,538	copy number variation	TRIM5 and 382 more					1

Human β-Hemoglobin Variation Viewer

http://www.ncbi.nlm.nih.gov/variation/view/



Variation Data

Filter by

Source database
- dbSNP (732)
- dbVar (0)

In ClinVar
- Yes (371)
- No (361)

Most severe clinical significance
- Pathogenic (109)
- Likely pathogenic (13)
- drug response (0)
- other (239)
- risk factor (0)

More...

Variant type
- single nucleotide variant (545)
- copy number variation (0)
- deletion (110)
- insertion (65)
- microsatellite (0)

More...

Molecular consequence
- missense variant (309)
- nonsense (20)
- stop lost (0)
- inframe variant (36)
- frameshift variant (103)

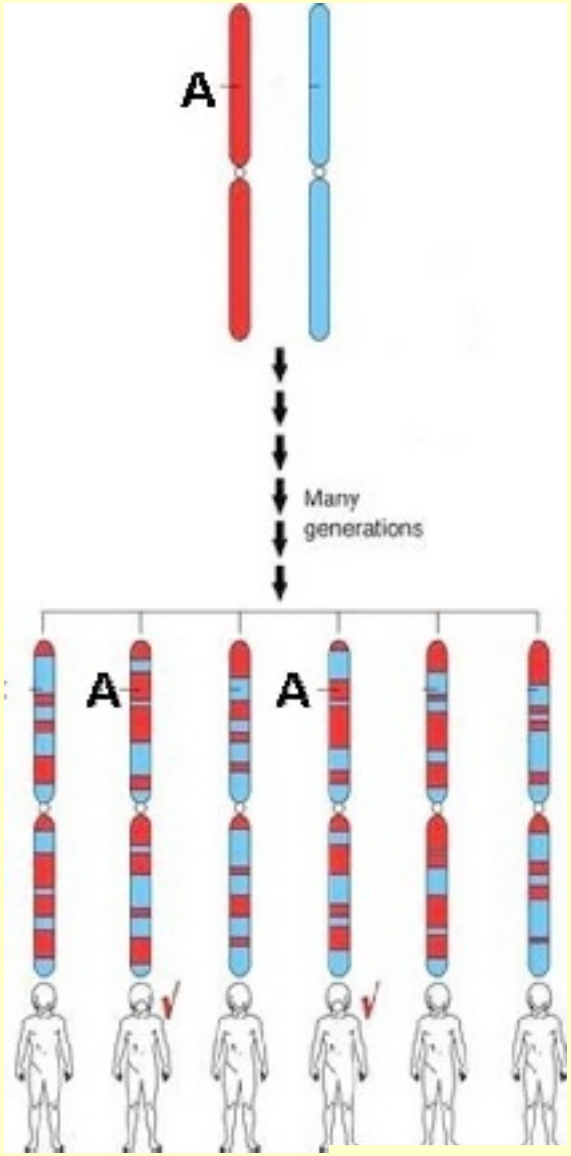
More...

1000 Genomes MAF
- < 0.005 (75)
- 0.005 - 0.01 (1)

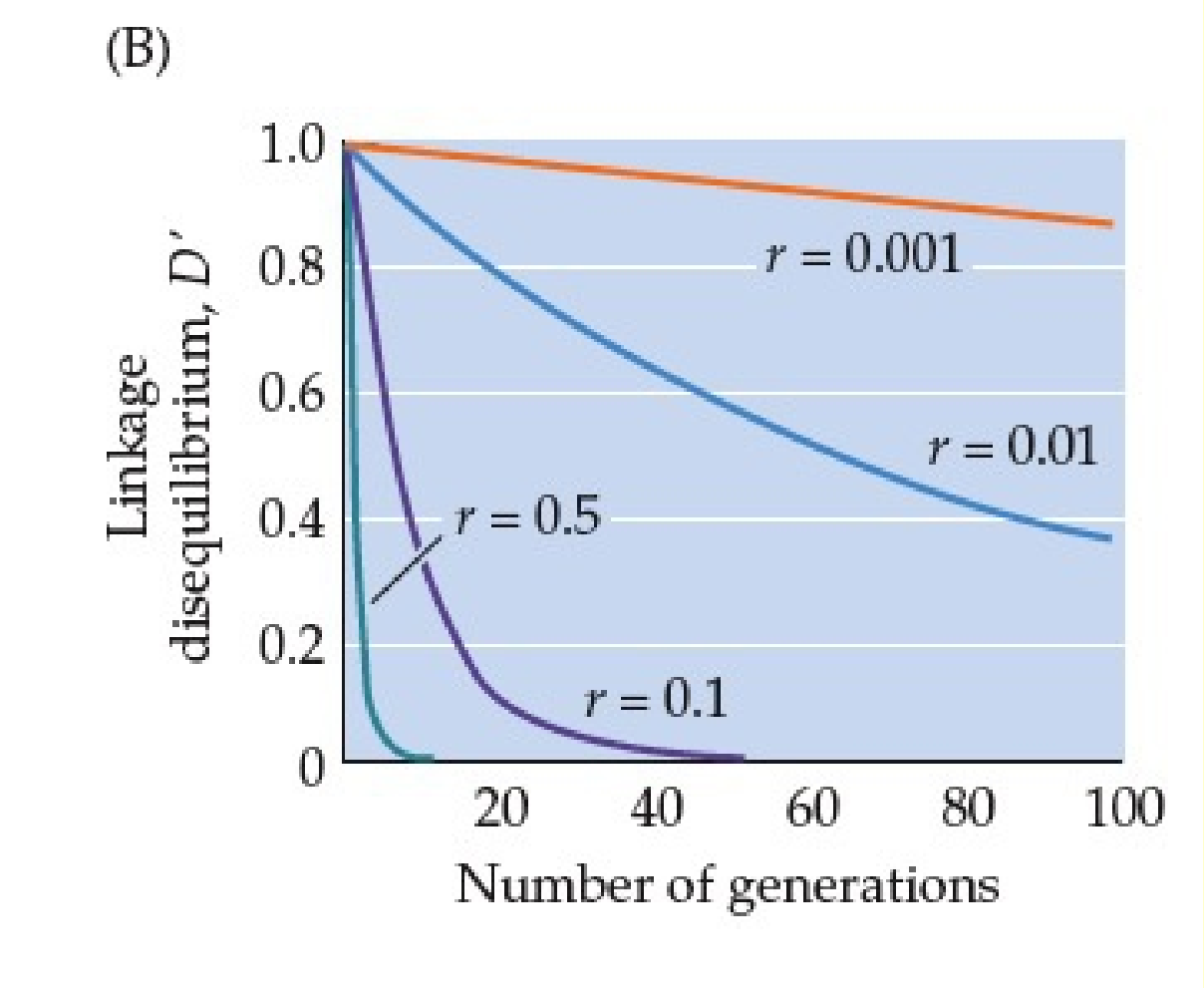
Download | Edit columns | Items 1 - 30 of 732 | << First | < Prev | Page 1 of 25 | Next > | Last >>

Variant ID	Location	Variant type	Gene	Molecular consequences	Most severe clinical significance	1000G MAF	GO-ESP MAF	Publications
rs560895446	5,225,322	single nucleotide variant	HBB	500B downstream variant		A = 0.0002		
rs111465890	5,225,352	single nucleotide variant	HBB	500B downstream variant		C = 0.0016		
rs12788013	5,225,365	single nucleotide variant	HBB	500B downstream variant		G = 0.0992		
rs567259408	5,225,367	single nucleotide variant	HBB	500B downstream variant		G = 0.0002		
rs765781952	5,225,377	single nucleotide variant	HBB	500B downstream variant				
rs533372053	5,225,389	single nucleotide variant	HBB	500B downstream variant		C = 0.0002		
rs113969885	5,225,416	single nucleotide variant	HBB	500B downstream variant		T = 0.0130		
rs375072494	5,225,423	single nucleotide variant	HBB	500B downstream variant				
rs560643693	5,225,460	single nucleotide variant	HBB	500B downstream variant		C = 0.0016		
rs111710394	5,225,463	single nucleotide variant	HBB	500B downstream variant				
rs528009939	5,225,469	single nucleotide variant	HBB	3 prime UTR variant		T = 0.0006		
rs753171158	5,225,482	single nucleotide variant	HBB	3 prime UTR variant				
rs606231219	5,225,484 - 5,225,488	deletion	HBB	3 prime UTR variant	Pathogenic			1
rs33985472	5,225,485	single nucleotide variant	HBB	3 prime UTR variant	Pathogenic			4
rs63750954	5,225,486	single nucleotide variant	HBB	3 prime UTR variant	Pathogenic			1
rs281864532	5,225,486 - 5,225,487	deletion	HBB	3 prime UTR variant				
rs35949130	5,225,486 - 5,225,490	deletion	HBB	3 prime UTR variant	Pathogenic			2
rs63751128	5,225,487	single nucleotide variant	HBB	3 prime UTR variant	Pathogenic			3
rs35676421	5,225,487 - 5,225,488	insertion	HBB	3 prime UTR variant				
rs63750205	5,225,487 - 5,225,488	deletion	HBB	3 prime UTR variant				1
rs33978907	5,225,488	single nucleotide variant	HBB	3 prime UTR variant	Pathogenic	G = 0.0002		6
rs281864905	5,225,488 - 5,225,489	deletion	HBB	3 prime UTR variant				1
rs34171453	5,225,495 - 5,225,507	deletion	HBB	3 prime UTR variant				1
rs770911771	5,225,501	single nucleotide variant	HBB	3 prime UTR variant				
rs34029390	5,225,502	single nucleotide variant	HBB	3 prime UTR variant	Likely benign	G = 0.0002		6
rs193922549	5,225,507	single nucleotide variant	HBB	3 prime UTR variant	Uncertain significance			
rs745310118	5,225,517	single nucleotide variant	HBB	3 prime UTR variant				

Origin of Haplotypes



Linkage Disequilibrium and Recombination Rate



Linkage Disequilibrium (LD) Across the Human LPL Gene

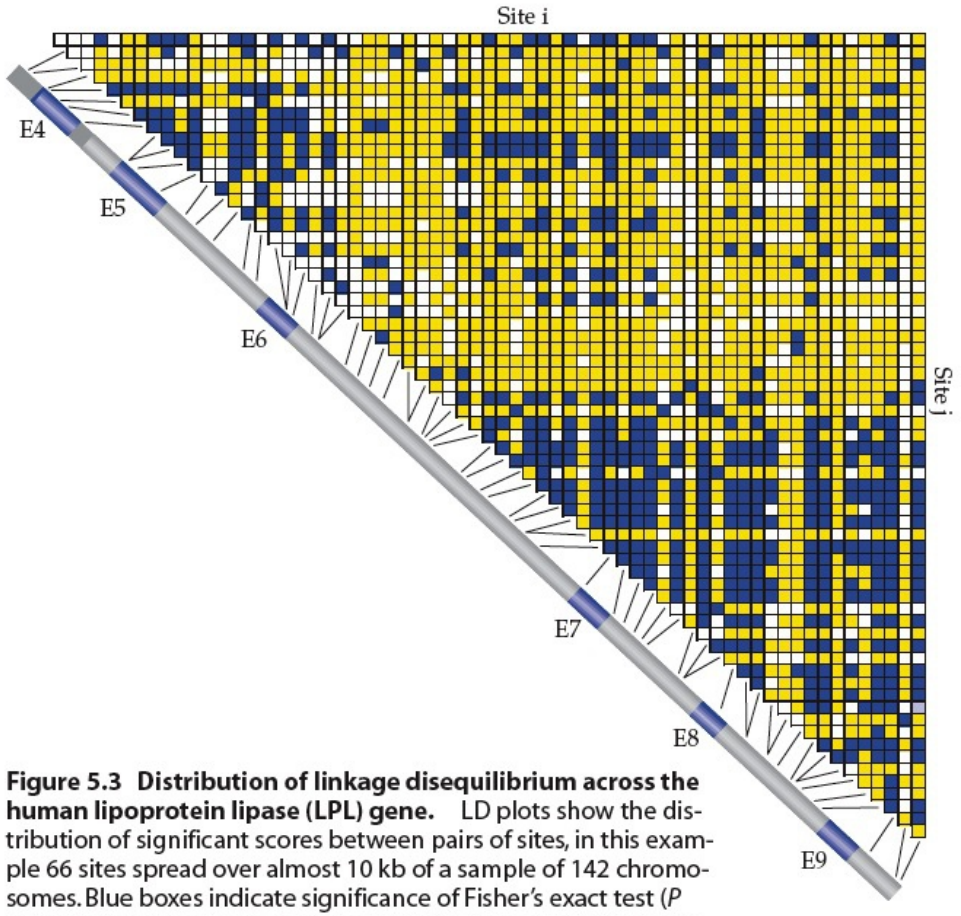
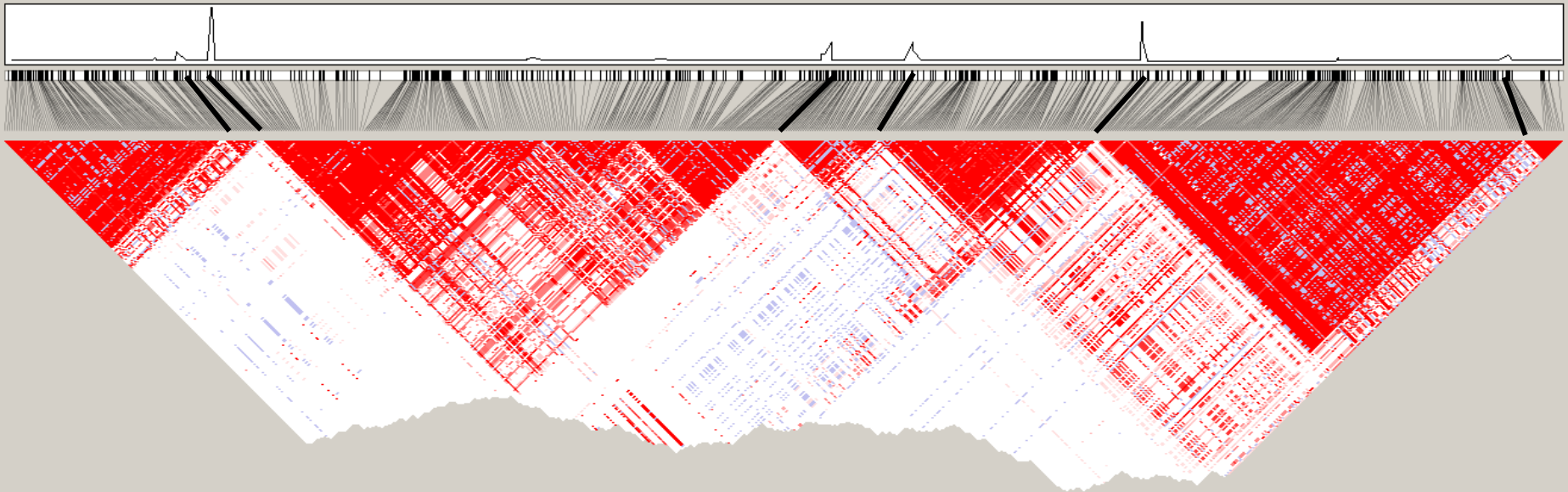


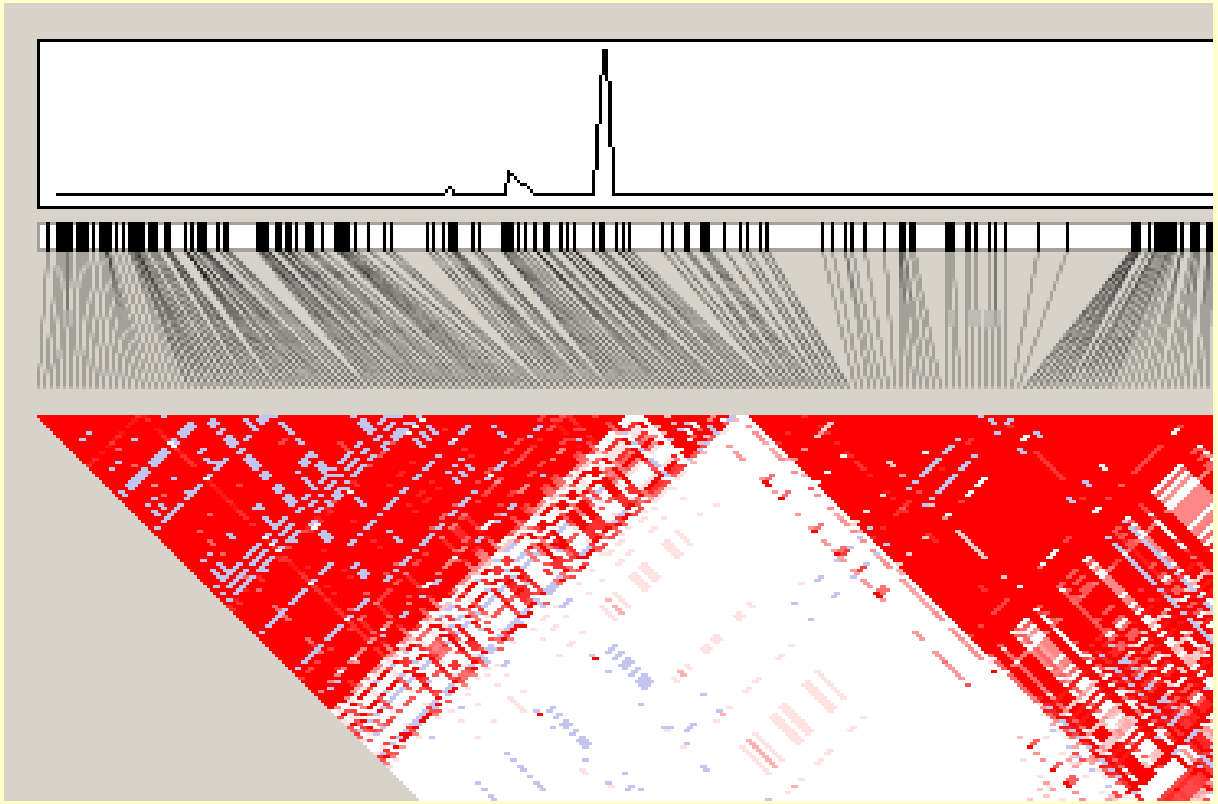
Figure 5.3 Distribution of linkage disequilibrium across the human lipoprotein lipase (LPL) gene. LD plots show the distribution of significant scores between pairs of sites, in this example 66 sites spread over almost 10 kb of a sample of 142 chromosomes. Blue boxes indicate significance of Fisher's exact test ($P < 0.001$), yellow boxes indicate nonsignificance, and white boxes are cases where there was insufficient power to test for LD at this level. Note that the extent of LD varies across the locus, and is not restricted to exon sequences. (Redrawn from Clark et al. 1998.)

Recombination hotspots are widespread and account for linkage disequilibrium structure



7q21

Recombination hotspots are widespread and account for linkage disequilibrium structure



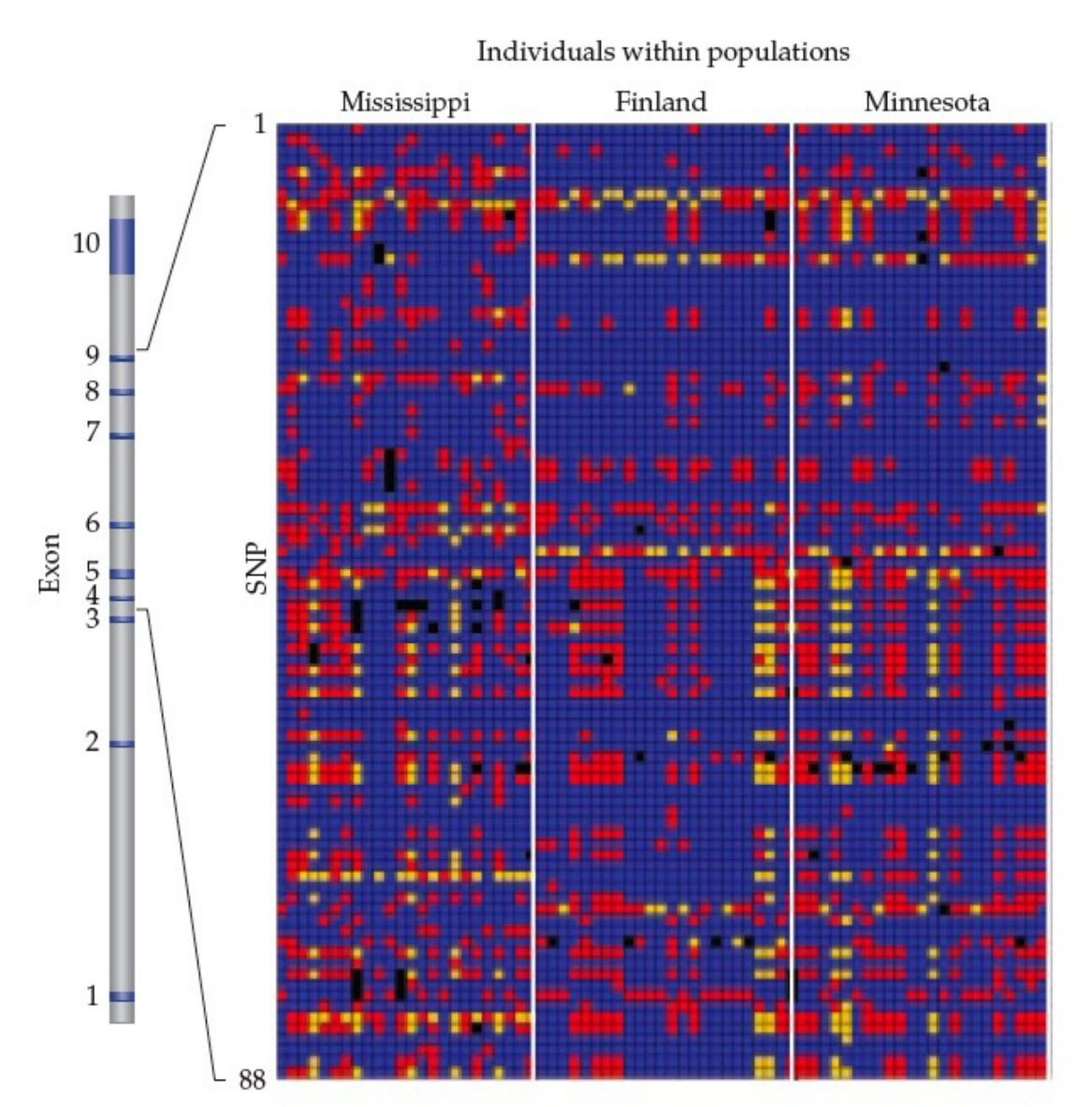
7q21

Consensus binding site for PRDM9

http://www.ncbi.nlm.nih.gov/gene/56979

Observation of Haplotypes

SNPs in Populations



Sequence and Distance-Based Phylogenies (evolutionary trees)

- Distance-based methods
  - Branch lengths = D(i,j)/2 for sequences i, j
  - Distances must be metric
  - Distances can reflect time or number of changes
  - Distances must be relatively constant per unit branch length