Your Genes and Your Health
http://bio84.stanford.edu/

Stanford Continuing Studies
http://continuingstudies.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

© Doug Brutlag 2016
Far too many diseases do not have a proven means of prevention or effective treatments. We must gain better insights into the biological, environmental, and behavioral influences on 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.
Genomics England
http://www.genomicsengland.co.uk/

Genomics England is delivering the 100,000 Genomes Project.

We are creating a new genomic medicine service with the NHS – to support better diagnosis and better treatments for patients. We are also enabling medical research.

More Information about the 100,000 Genomes Project

News

Latest from the 100,000 Genomes Project

Across England patients with cancer are now being recruited to the main phase of the 100,000 Genomes Project, giving the fight against cancer a significant boost. Minister for Life Sciences George Freeman MP announced

Latest Videos

Participant Stories

Genome sequencing has given Jessica a diagnosis for her rare condition.

Twitter

This website uses cookies to improve your experience. By continuing to use the site, you agree to the use of cookies.
Francis Crick Institute
http://www.crick.ac.uk/

The Francis Crick Institute will be an inter-disciplinary medical research institute. Its work will help understand why disease develops and find new ways to treat, diagnose and prevent illnesses such as cancer, heart disease, infections, and neurodegenerative diseases.

>> Find out more

Strategy
The Francis Crick Institute will be an entirely new institute with a distinctive vision of how medical research should be conducted. It will play a key role in creating the foundation of knowledge on which this century’s improvements in health will be based.

Construction
The building of the Francis Crick Institute is underway with appointed contractor, Laing O’Rourke, leading the works. Designed by HOK with PLP Architecture, the building is scheduled for completion in 2015.

NEWS
New Knowledge Quarter launched by the Chancellor
Crick awards Hard FM contract
The Crick hosts the London Evening Standard’s 1000 event

View all news

SCIENCE NEWS
Bacteria could contribute to development of wound-induced skin cancer
Study sheds light on link between gut microorganisms and nervous system
Plans for an Australian 100,000 Genomes Project

Dr Philippa Brice
Sunday, 6 December 2015

Australia may be moving towards a home-grown version of the UK 100,000 Genomes Project.

According to *The Australian*, the Garvan Institute of Medical Research is reportedly in talks with the federal government and other public and private research institutions including major telecommunications company Telstra, to see whether they can create such a project by working in partnership. Telstra already has a dedicated health division.
All humans are ~99.9% identical at the DNA sequence level, and yet...

all of us carry a significant number of ‘glitches’ in our genomes.
The International HapMap Project is a partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States to develop a public resource that will help researchers find genes associated with human disease and response to pharmaceuticals. See “About the International HapMap Project” for more information.

- 2013-06-14: HapMap data conversion tool
  There are several inquires for a conversion tool to convert HapMap data into the VCF format. Please take a look of The Genome Analysis Toolkit (by Broad Institute).

- 2012-12-06: Downtime for hardware maintenance
  From December 15 - 16, Hapmap site will be taken offline for an internal hardware maintenance. Sorry for the inconvenience.

- 2011-06-13: HapMap help desk announcement
  There was a problem with the HapMap help desk system. In the past several weeks, emails sent to hapmap-help@ncbi.nlm.nih.gov did not reach the help desk, and thus user requests were not addressed. Please resend your email request if you sent emails to the HapMap help desk in the past several weeks. Sorry for the inconvenience.

- 2011-04-20: Hapmap help desk service interruption notice
  There will be no help desk support from 05/03/2011 to 05/23/2011. Sorry for the inconvenience.

- 2011-02-02: Haplovie issues with rel 28 data
  Recently, there are several questions about Haplovie data format errors when users tried to analyze HapMap release 28 data. The current Haplovie version (4.2) does not recognize the new individuals in release 28 and the software will generate an error similar to "Hapmap data format error: NA18876" when trying to open the data.

Haplovie is developed and maintained by an organization different from HapMap. Please contact Haplovie help desk (haplovie@broadinstitute.org) for questions specific to this software.

- 2011-01-19: HapMap phase II recombination rate on GRCh37
  The liftover of the HapMap II genetic map from human genome build b35 to GRCh37 is available. Data is available for bulk download.
2007 SCIENTIFIC BREAKTHROUGH OF THE YEAR
Science Magazine, December 21, 2007

“It’s all about me!”

Single Nucleotide Polymorphisms (SNPs)

Individual 1


Individual 2


Individual 3


Individual 4

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_svs_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

GRCh38 alignments for Exome and High Coverage 1000 Genomes Data
An integrated map of structural variation in 2,504 human genomes


Affiliations | Contributions | Corresponding authors

Nature 526, 75–81 (01 October 2015) | doi:10.1038/nature15394
Received 19 February 2015 | Accepted 20 August 2015 | Published online 30 September 2015

Nature 526, 75-81. (30 September 2015)
Single Nucleotide Polymorphisms (SNPs) in the Human Genome

About 38 million sites in the human genome where sequence variations have occurred

About 15 million sites where variation exceeds 1% of a particular population (MAF > 1%)

Each ethnic group has its own distribution of SNPs

About 3 million sites where any individual varies from the consensus human genome.

Each person differs from each other in 3 million places (about 0.1% of the genome)

SNP sequence variations are common, unlike disease causing mutations which are rare.
Single (Simple) Nucleotide Polymorphisms (SNPs)

GCTGTATGACTAGAAGATCGAT
GCTGTATGACGAGAAGATCGAT

• SNPs can be used for identifying individuals and forensics
• SNPs are used for mapping & genome-wide association studies of complex disease
• SNPs are used for ancestry tracking & family relationships
• SNPs are used for estimating predisposition to disease
• SNPs are used to predict risk of common genetic diseases
• While SNPs are linked with disease, they do not cause disease
• SNPs are used for personalized medicine and genomics
• SNPs are used for classifying patients in clinical trials
• In short, SNPs are used as genetic markers to map human diseases and traits and migrations.
Chapter 5: The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation

Adrienne Kitsis
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 breadth 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 breadth 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.
Gene Transcription into RNA and Translation into Protein

Transcription

Splicing

Translation

Exon

Intron

Exon

Gene (DNA)

Transcript (RNA)

mRNA (RNA)

Protein
The Genetic Code
Human β-Hemoglobin Gene


HBB hemoglobin, beta [Homo sapiens]
Gene ID: 3043, updated on 24-Oct-2010

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]
Human β-Hemoglobin Gene

Human β-Hemoglobin Gene SNPs

<table>
<thead>
<tr>
<th>Region</th>
<th>Chr. position</th>
<th>mRNA pos</th>
<th>dbSNP rs#</th>
<th>cluster id</th>
<th>Heterozygosity</th>
<th>Validation</th>
<th>3D Linkout</th>
<th>Function</th>
<th>dbSNP allele</th>
<th>Protein residue</th>
<th>Codon pos</th>
<th>Amino acids pos</th>
<th>PubMed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs1405464</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>frame shift</td>
<td>AGC</td>
<td>[S]</td>
<td>1</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs113062294</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>frame shift</td>
<td>[AN]</td>
<td>1</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs11511174</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>frame shift</td>
<td>[17bp]</td>
<td>[Q]</td>
<td>1</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs115465889</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>missense</td>
<td>[A]</td>
<td>2</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs71811954</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>frame shift</td>
<td>(15bp)</td>
<td>[R]</td>
<td>3</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs14539966</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>frame shift</td>
<td>(17bp)</td>
<td>[L]</td>
<td>1</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rs11549405</td>
<td>N.D.</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td>synonymous</td>
<td>C</td>
<td>3</td>
<td>89</td>
<td></td>
</tr>
</tbody>
</table>
Origin of Haplotypes
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 SNP linkage structure
Recombination hotspots are widespread and account for SNP linkage structure.
SNPs in Populations

Individuals within populations

Mississippi | Finland | Minnesota

Exon

 SNP

© Gibson & Muse, A Primer of Genome Science
Sequence and Distance-Based Evolutionary Trees

• Sequence-Based Methods
  – Assigns mutations to branches
  – Minimize number of changes
  – Topology maximizes similarity of neighboring leaves
Human Prehistory 101

Prologue
Portrait of a Glitch

- Revere La Noue, MFA, Stanford, 2005
- What is this film about?
- What classes of glitches are mentioned?
- What do these glitches cause?
- Why did I show this film?
Portrait of a Glitch

Revere La Noue, MFA, Stanford, 2005

What is this film about?

What classes of glitches are mentioned?

What do these glitches cause?

Why did I show this film?