1000 Genomes: A Deep Catalog of Human Genetic Variation

RRID:SCR_006828
Type: Tool

Proper Citation

1000 Genomes: A Deep Catalog of Human Genetic Variation (RRID:SCR_006828)

Resource Information

URL: http://www.1000genomes.org/

Proper Citation: 1000 Genomes: A Deep Catalog of Human Genetic Variation (RRID:SCR_006828)

Description: International collaboration producing an extensive public catalog of human genetic variation, including SNPs and structural variants, and their haplotype contexts, in an effort to provide a foundation for investigating the relationship between genotype and phenotype. The genomes of about 2500 unidentified people from about 25 populations around the world were sequenced using next-generation sequencing technologies. Redundant sequencing on various platforms and by different groups of scientists of the same samples can be compared. The results of the study are freely and publicly accessible to researchers worldwide. The consortium identified the following populations whose DNA will be sequenced: Yoruba in Ibadan, Nigeria; Japanese in Tokyo; Chinese in Beijing; Utah residents with ancestry from northern and western Europe; Luhya in Webuye, Kenya; Maasai in Kinyawa, Kenya; Toscani in Italy; Gujarati Indians in Houston; Chinese in metropolitan Denver; people of Mexican ancestry in Los Angeles; and people of African ancestry in the southwestern United States. The goal Project is to find most genetic variants that have frequencies of at least 1% in the populations studied. Sequencing is still too expensive to deeply sequence the many samples being studied for this project. However, any particular region of the genome generally contains a limited number of haplotypes. Data can be combined across many samples to allow efficient detection of most of the variants in a region. The Project currently plans to sequence each sample to about 4X coverage; at this depth sequencing cannot provide the complete genotype of each sample, but should allow the detection of most variants with frequencies as low as 1%. Combining the data from 2500 samples should allow highly accurate estimation (imputation) of the variants and genotypes for each sample that were not seen directly by the light sequencing. All samples from the
1000 genomes are available as lymphoblastoid cell lines (LCLs) and LCL derived DNA from the Coriell Cell Repository as part of the NHGRI Catalog. The sequence and alignment data generated by the 1000genomes project is made available as quickly as possible via their mirrored ftp sites. ftp://ftp.1000genomes.ebi.ac.uk ftp://ftp-trace.ncbi.nlm.nih.gov/1000genomes

**Abbreviations:** 1000 Genomes

**Synonyms:** International 1000 Genomes Project, 1000 Genomes Project

**Resource Type:** database, consortium, data or information resource, portal, organization portal, data set

**Keywords:** genetic variation, gene, next-generation sequencing, sequence, alignment, genome, single-nucleotide polymorphism, structural variant, haplotype, genome-wide association study, pharmacology, genetics, biomarker, consortium, data sharing, genotype, phenotype, FASEB list

**Funding Agency:** Wellcome Trust Sanger Institute; Hinxton; United Kingdom, Beijing Genomics Institute; Shenzhen; China, NHGRI, 454 Life Sciences Roche, Life Technologies, Illumina

**Availability:** Free, Public, Restrictions apply, Http://www.1000genomes.org/data#DataAccess

**Resource Name:** 1000 Genomes: A Deep Catalog of Human Genetic Variation

**Resource ID:** SCR_006828

**Alternate IDs:** nlx_143819, OMICS_00261

---

### Ratings and Alerts

No rating or validation information has been found for 1000 Genomes: A Deep Catalog of Human Genetic Variation.

No alerts have been found for 1000 Genomes: A Deep Catalog of Human Genetic Variation.

---

### Data and Source Information

**Source:** SciCrunch Registry

---

### Usage and Citation Metrics

We found 4107 mentions in open access literature.
Listed below are recent publications. The full list is available at RRID.


Melton HJ, et al. (2023) MIMOSA: A resource consisting of improved methylome imputation models increases power to identify DNA methylation-phenotype associations. medRxiv : the preprint server for health sciences.

Han SK, et al. (2023) Mapping genomic regulation of kidney disease and traits through high-resolution and interpretable eQTLs. Nature communications, 14(1), 2229.

Downie ML, et al. (2023) Shared genetic risk across different presentations of gene test-negative idiopathic nephrotic syndrome. Pediatric nephrology (Berlin, Germany), 38(6), 1793.

de Mol CL, et al. (2023) Multiple sclerosis risk variants influence the peripheral B-cell compartment early in life in the general population. European journal of neurology, 30(2), 434.


Mostad P, et al. (2023) Improved computations for relationship inference using low-coverage
sequencing data. BMC bioinformatics, 24(1), 90.


