

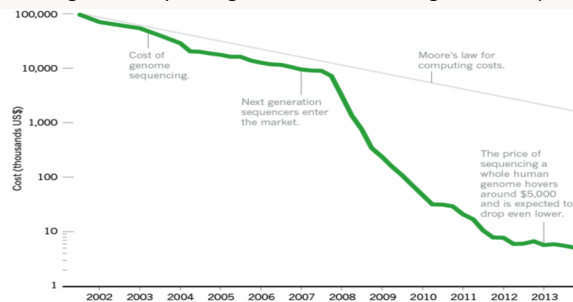
Beacon Network: A System for Global Genomic Data Sharing

Motivation

"The only way to deal with variants of unknown significance is to create a large-scale genomic knowledge network that includes all the genetic variants observed worldwide, and their associated health observations. This requires a global network of shared data. The vast majority of human genetic variants are individually rare but collectively common, so this effort will have health impact on almost everybody."

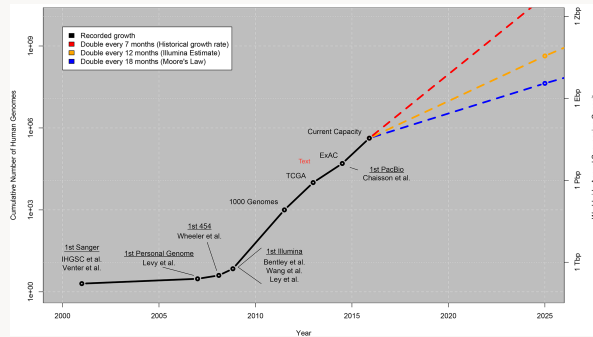
David Haussler, Director of UCSC Genomics Institute
<https://genomics.soe.ucsc.edu/news/article/172>

- ▶ Genomics is the key to understanding human health.
- ▶ Data generation is exponential.
 - Cost of genome sequencing down 1M times during the recent years.



Hayden, E. C. (2014). Technology: The \$1,000 genome. *Nature* 507(7492): 294-295.

- 100M-2B humans sequenced by 2025 (4-5 orders of magnitude rise).



Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, et al. (2015). Big Data: Astronomical or Genomical? PLoS Biol 13(7): e1002195.

- ▶ Aggregating data in a centralized site is no longer enough.
- ▶ A key obstacle to solving diseases is the difficulty of discovering additional affected families efficiently.


Beacon

- ▶ Web service for genetic data sharing.
- ▶ Receives questions of the form *"Do you have information about this mutation?"*.
- ▶ Responds with *"yes"* or *"no"*, optionally with additional information about the mutation.
- ▶ Initiative of the Global Alliance for Genomics & Health.
 - Coalition of over 400 institutions developing standards for sharing of genomic and clinical data.

More information: <https://genomicsandhealth.org/work-products-demonstration-projects/beacon-project-0>.

Beacon Network

- ▶ Search engine across the world's public beacons.
- ▶ Answers questions of the form *"Who has information about this mutation?"*.
- ▶ Works by translating genomic queries for each beacon, intelligently distributing the queries through the network, and aggregating the results.


Beacon Network

Search Beacon

Search **all beacons** for allele

GRCh37 -

13 : 32936732 G > C


Search

Response
All None


☒ Found 14
☒ Not Found 37
☐ Not Applicable 14

Organization
All None


☒ AMPLab, UC Berkel...
☒ BGI
☒ BioReference Labor...
☒ Brazilian Initiative o...
☒ BRCA Exchange
☒ Broad Institute
☒ Centre for Genomic...
☒ Centro Nacional de...
☒ Curoverse
☒ EMBL European Bio...
☒ Global Alliance for ...
☒ Google
☒ Institute for System...
☒ Instituto Nacional d...


AMPLab - 1000 Genomes Project
Hosted by AMPLab, UC Berkeley


Not Found


Biobase - HGMD
Hosted by University of California, Santa Cruz

Found


BioReference
Hosted by BioReference Laboratories

Not Found


BRCA Exchange
Hosted by BRCA Exchange

Hide Metadata

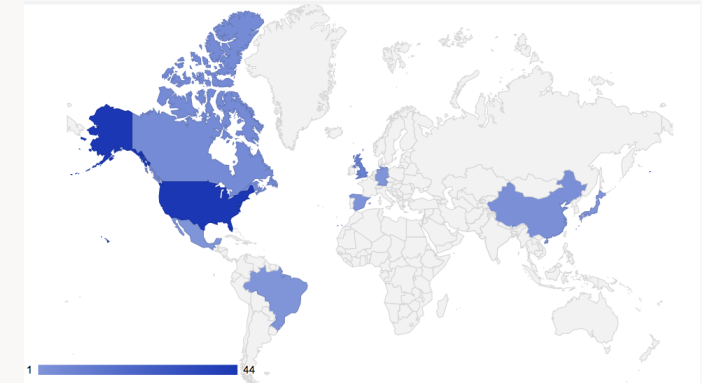
Found

Clinical significance citations	PMID:23108138
HGVS cDNA	c.7878G>C
Condition ID type	OMIM
Date last evaluated	8/10/15

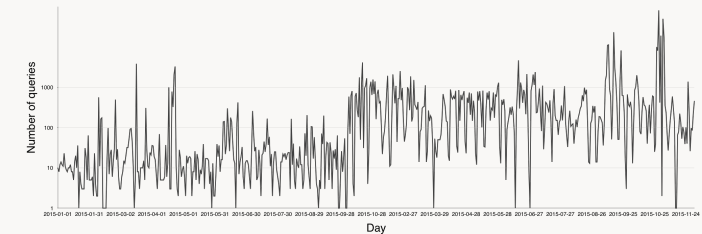
More information: <https://beacon-network.org>.

Status today

- ▶ Biggest search and discovery engine of human genomic variation in the world.
- ▶ 25+ of the world's top genomic organizations, 60+ beacons.



- ▶ Served 400K+ genomic queries resulting in 2M+ queries dispatched to the participants of the network.



- ▶ Search spanning est. 2M+ samples and 2B+ mutations.
- ▶ 6K+ users from 100+ countries.
- ▶ Queries targeting harmful genetic variations, rare diseases and otherwise interesting or unusual genes and phenotypes.
- ▶ Mentioned in several scientific publications and press articles, presented at major conferences (e.g. ASHG, BioData World).
- ▶ Integrated in commercial products (e.g. DNASTack, Omnicia).
- ▶ Influences standards for how genomic data is shared, drives development of the Beacon API specification.
- ▶ Now REST API, web frontend, command line client, query library, compatible beacon implementations in multiple programming languages, and adapters for various sources of genomic data.