

OpenCB Home

An integrated suite of high-performance (big data) applications for the management and analysis of population-scale genomic data.

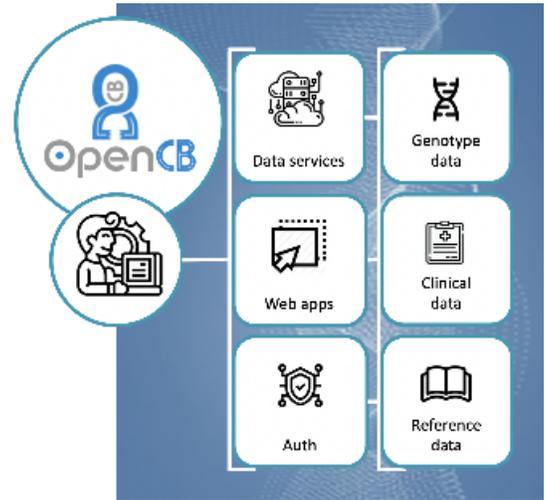
By replacing datafiles with databases our software promotes:

- Scalability; the ability to run queries in real-time across hundreds of thousands of genomes.
- Accessibility; the ability to analyse from anywhere on the web without needing access to the local filesystem.
- Integration; the ability to link data together using flexible data models covering reference data, variant data and clinical metadata.
- Security; the ability to protect data using federated (e.g. SSO) authentication and role-based authorisation schemes.

OpenCB solutions are typically used as:

- The storage target of secondary analysis pipelines.
- The data source for tertiary analysis workflows.

The open-source OpenCB software is developed and maintained by researchers from multiple organisations and made freely available at <https://github.com/opencb>



Flagship Projects

CellBase

The "unified reference"

CellBase aggregates over 10 TB reference data from over 20 data sources (and counting). Data are exposed via a single, consistent API. Users can use the public instance hosted by the University of Cambridge or install their own copy. Those in the latter camp can use CellBase to manage their own reference collection.

OpenCGA

The "VCF database"

OpenCGA is software for storage and retrieval of genotype data and associated clinical and operational metadata. Its integration with CellBase provides for powerful variant annotation functionality. It provides extensive web services, APIs for R & Python, and its own command line interface.

IVA

The "web application"

The Interactive Variant Analysis (IVA) makes it easy to work with the variant information stored in OpenCGA and annotated by CellBase. It has tools for browsing, filtering, analysis and interpretation that are tailored for studies of population genomics and genomic medicine.

- [Genome Maps](#)
- [BioNetDB](#)
- [HGVA](#)
- [JSorolla](#)
- [oskar](#)

- [OpenCB People](#)
- [CellBase Contributors](#)
- [OpenCGA Contributors](#)
- [IVA Contributors](#)
- ...and many, many more

[Nacho Medina <igmechas@gmail.com>](mailto:igmechas@gmail.com)