Beacon: A Protocol for Federated Discovery and Sharing of Genomic Data

Motivation

- Analysis of large volumes of genomic data has the potential to drive discoveries and applications in medicine.
- No institution has enough resources to capture all human variation.
- > Standards for interoperability enabling sharing across a federated network of genomic data resources are needed.

Beacon

- Standard defining a web service for discovery and sharing of information about alleles.
- ▶ Demonstration of willingness of international organizations to share data.
- ► Initiative of the Global Alliance for Genomics & Health.
- More information: https://ga4gh.org/#/beacon.

API

- Request: "Do you have information about this mutation?".
 - "referenceName": "13",
 "start": 32936731,
 "referenceBases": "G",
 "alternateBases": "C",
 "assemblyId": "GRCh37"
- Response: "yes"/"no" (optionally additional information about the mutation).

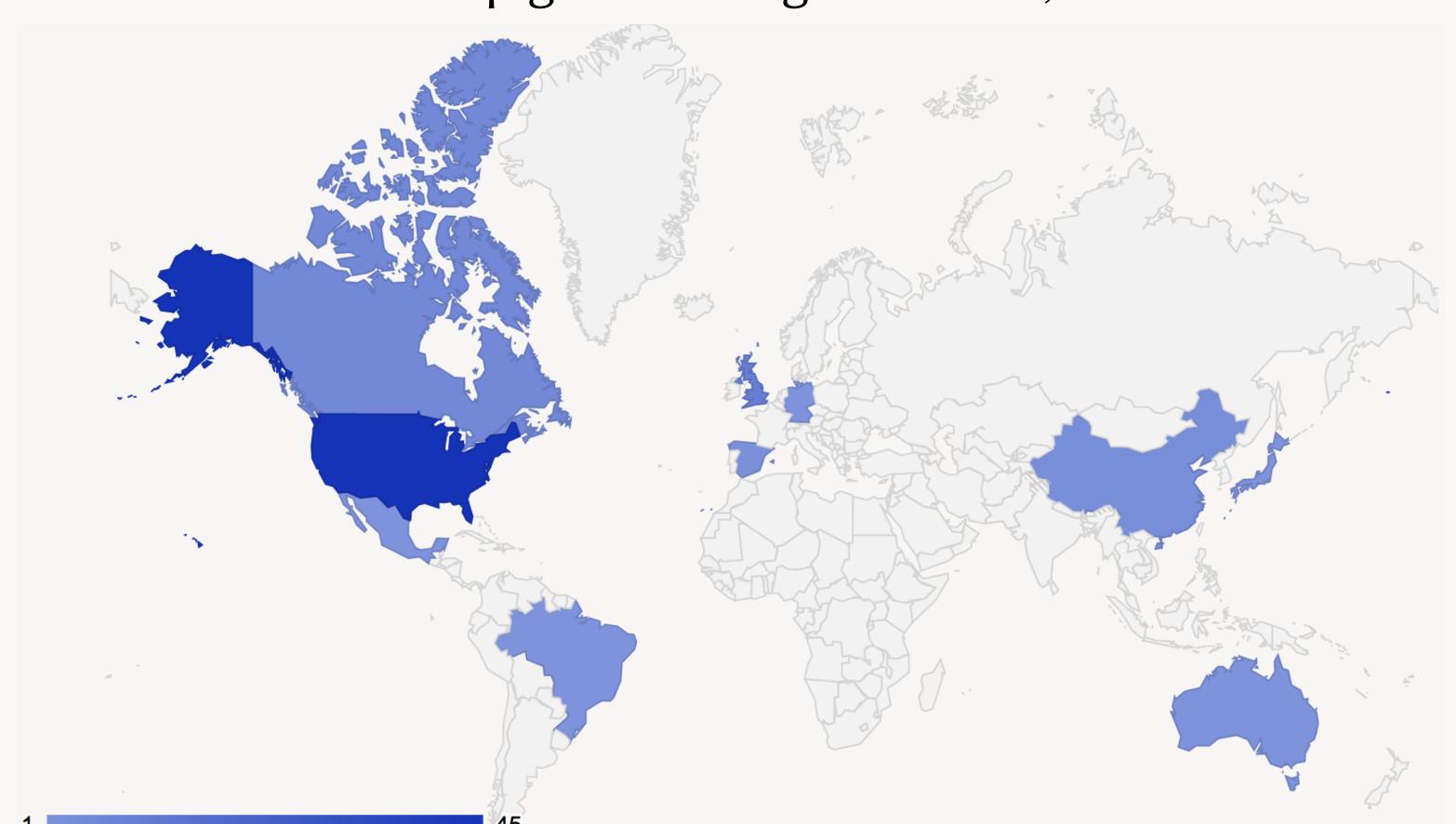
```
"exists": true,
"datasetAlleleResponses": [
    {
       "exists": true,
       "info": {
```

- Supported by a variety of tools, e.g. reference implementation, command line client, query library, sample beacons in multiple programming languages, adapters for various sources of genomic data.
- More information: https://github.com/ga4gh/beacon-team.

Participants

• • •

25+ of the world's top genomic organizations, 60+ beacons.



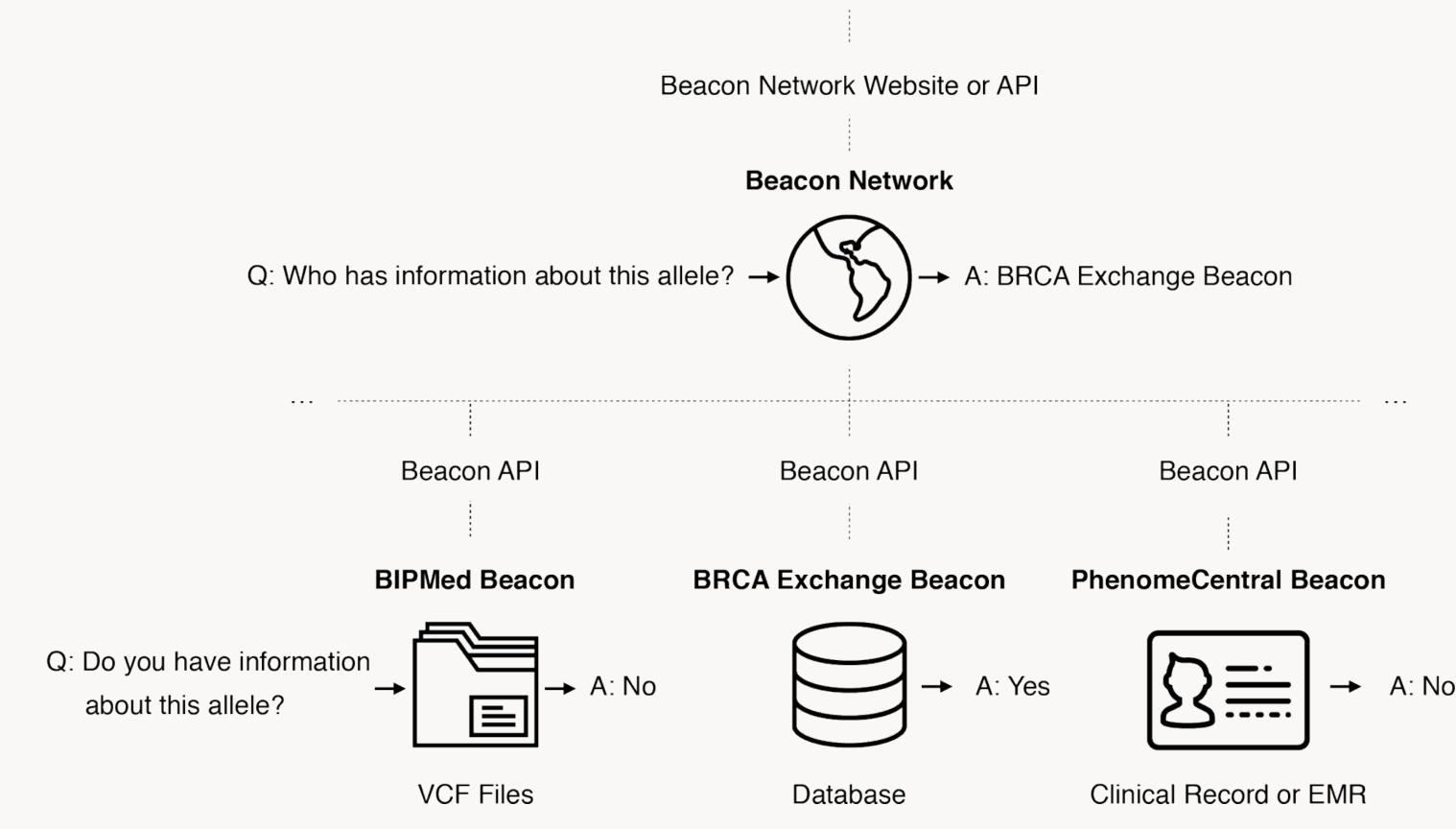
► 160+ datasets of various origins, e.g. sequence observations from patients, population studies, in-silico predictions, curated/crowd-sourced databases, scientific literature.

Beacon Network

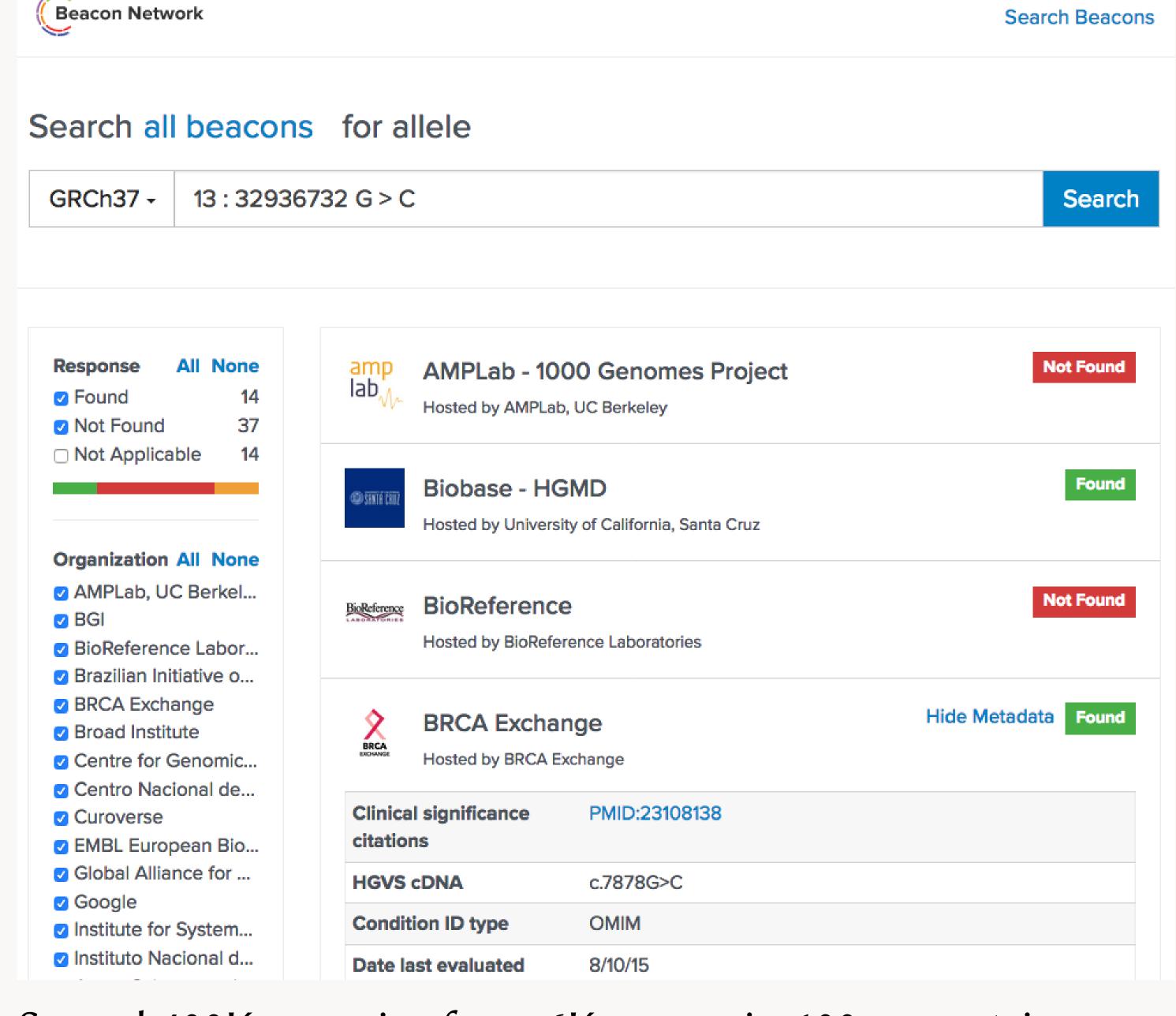
(Beacon Network

Search engine across the world's beacons.

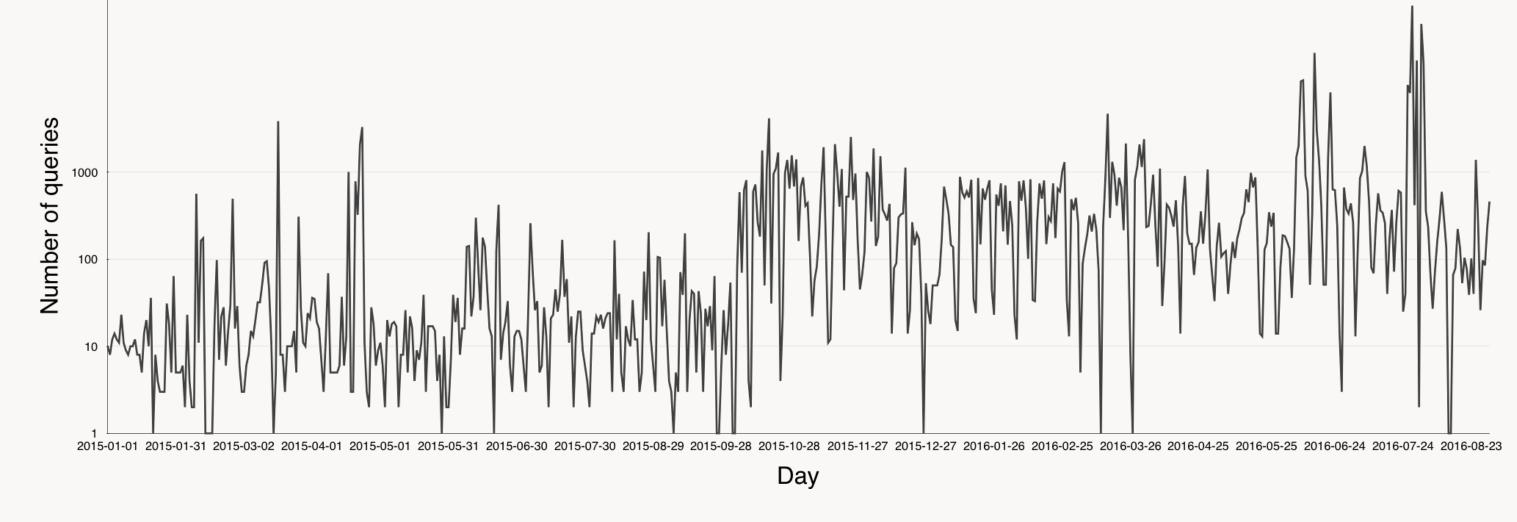




Translates genomic queries for each beacon, intelligently distributes the queries through the network, and aggregates the results.



➤ Served 400K+ queries from 6K+ users in 100+ countries, resulting in 2M+ queries dispatched to the participants of the network.



More information: http://beacon-network.org.

