

# Beacon v2 and Beyond

## A Standard Framework for Data Discovery and Sharing in Biomedical Genomics

**Michael Baudis**

Professor of Bioinformatics

University of Zürich

Swiss Institute of Bioinformatics **SIB**

Member GA4GH Strategic Leadership Committee

GA4GH Workstream Co-lead *DISCOVERY*

Co-lead ELIXIR Beacon API Development

Co-lead ELIXIR hCNV Community



Universität  
Zürich<sup>UZH</sup>



**SIB**  
Swiss Institute of  
Bioinformatics



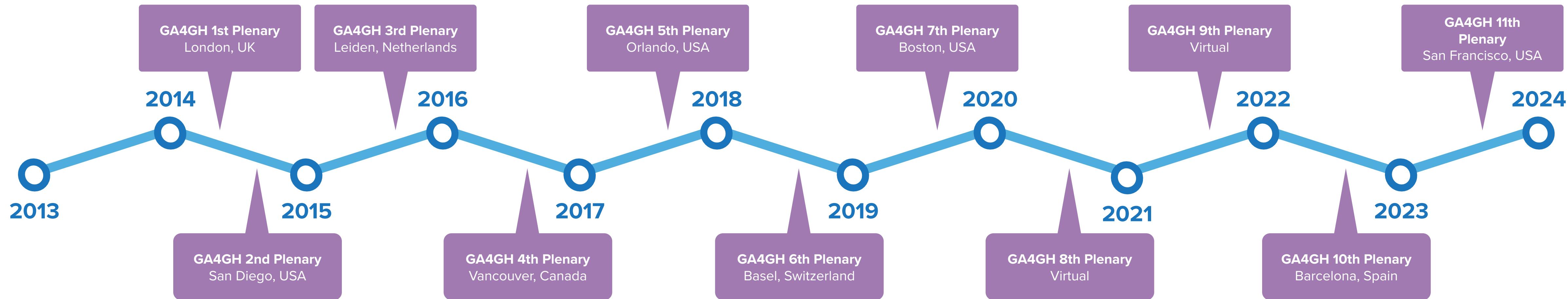
**Global Alliance**  
for Genomics & Health  
Collaborate. Innovate. Accelerate.



# GA4GH timeline



Global Alliance  
for Genomics & Health



Pre-launch	Building momentum	GA4GH Connect	Gap analysis	Strategic Refresh
 <p>73 partners sign a letter of intent to form an alliance</p>	 <p><b>Global Alliance</b> for Genomics &amp; Health Collaborate. Innovate. Accelerate.</p> <p><b>Formal launch of GA4GH</b></p> <p>Published <i>Framework for Responsible Sharing of Genomic and Health-Related Data</i></p> <p>Formed four working groups</p> <p>Developed three demonstration projects</p>	 <p>Launch of <b>GA4GH Connect</b> and Strategic Roadmap</p> <p>Formation of new organizational structure consisting of eight Work Streams and over twenty Driver Projects</p>	<p><b>Gap analysis</b> identifies three community imperatives</p> <ul style="list-style-type: none"><li> Interoperability and alignment</li><li> Implementation support</li><li> Engaging with healthcare and clinical standards</li></ul>	 <p><b>Strategic refresh</b> introduces updates to GA4GH to better meet the three community imperatives</p>

## Beacon v1 Development

2014

GA4GH founding event; Jim Ostell proposes Beacon concept including "more features ... version 2"

2015

- beacon-network.org aggregator created by DNAstack

2016

- Beacon v0.3 release
- work on queries for structural variants (brackets for fuzzy start and end parameters...)

2017

- OpenAPI implementation
- integrating **CNV parameters** (e.g. "startMin, statMax")
- Beacon v0.4 release in January; feature release for GA4GH approval process
- **GA4GH Beacon v1 approved** at Oct plenary

2019

- ELIXIR Beacon Network

2020



2021

## Beacon v2 Development

- Beacon+ concept implemented @ [progenetix.org](#)
- concepts from GA4GH Metadata (ontologies...)
- entity-scoped query parameters ("individual.age")

- Beacon+ demos "handover" concept

- Beacon hackathon Stockholm; settling on **filters**
- Barcelona goes Zurich developers meeting
- Beacon API v2 Kick off
- adopting "handover" concept

- "**Scouts**" teams working on different aspects - filters, genomic variants, compliance ...

- **framework + models** concept implemented
- range and bracket queries, variant length
- starting of GA4GH review process

- changes in default model, aligning with Phenopackets and VRS
- unified beacon-v2 code & docs repository
- **Beacon v2 approved** at April GA4GH Connect

2022

## Related ...

- ELIXIR starts Beacon project support

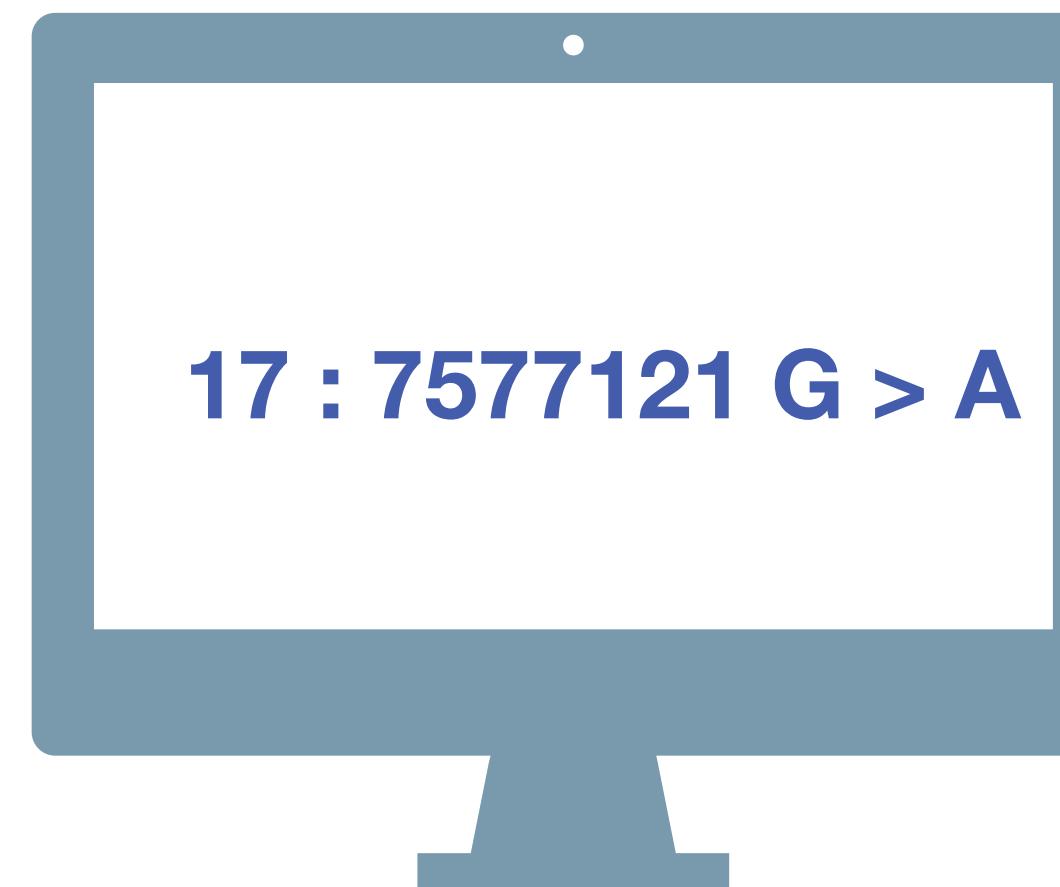
- GA4GH re-structuring (workstreams...)
- Beacon part of Discovery WS

- new Beacon website (March)

- Beacon publication at Nature Biotechnology

- Phenopackets v2 approved

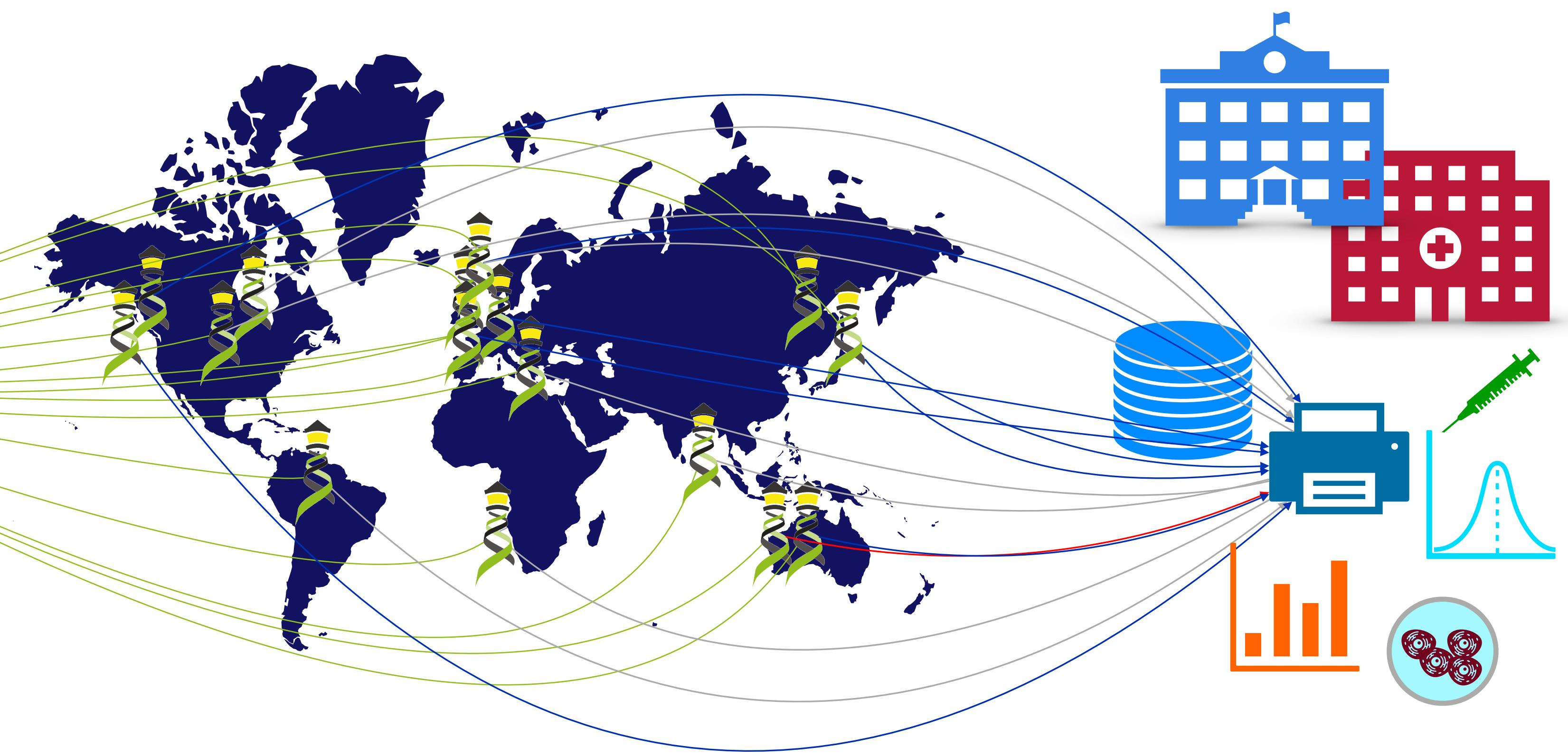
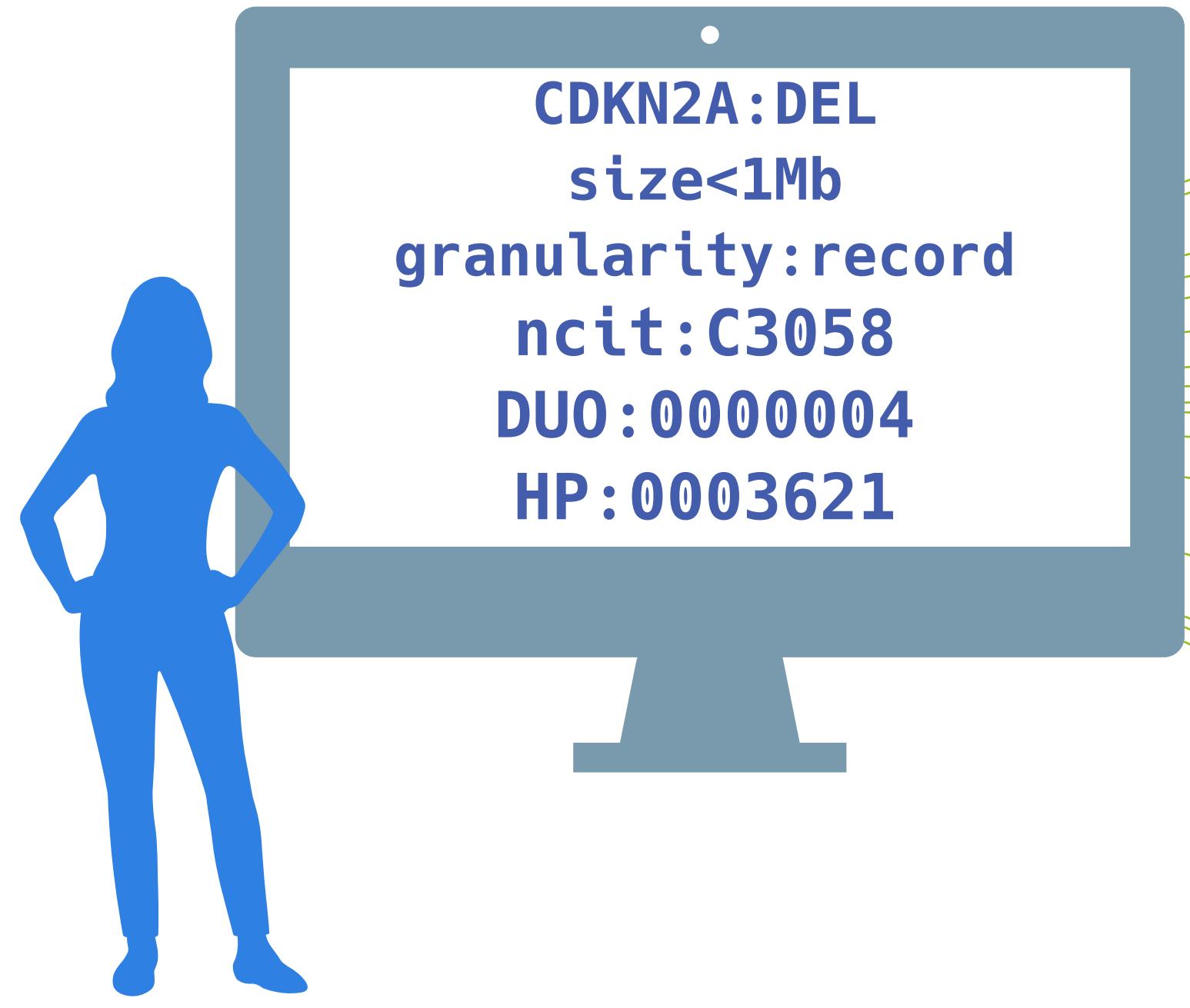
- [docs.genomebeacons.org](#)



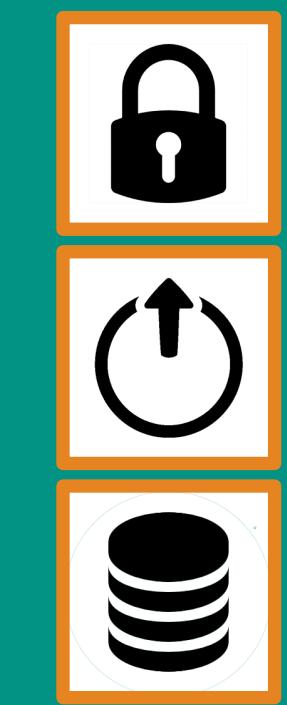
Have you seen this variant?  
It came up in my patient  
and we don't know if this is  
a common SNP or worth  
following up.

A Beacon network federates  
genome variant queries  
across databases that  
support the **Beacon API**

Here: The variant has  
been found in **few**  
resources, and those  
are from **disease**  
specific **collections**.



Can you provide data about focal deletions in CDKN2A in Glioblastomas from juvenile patients with unrestricted access?

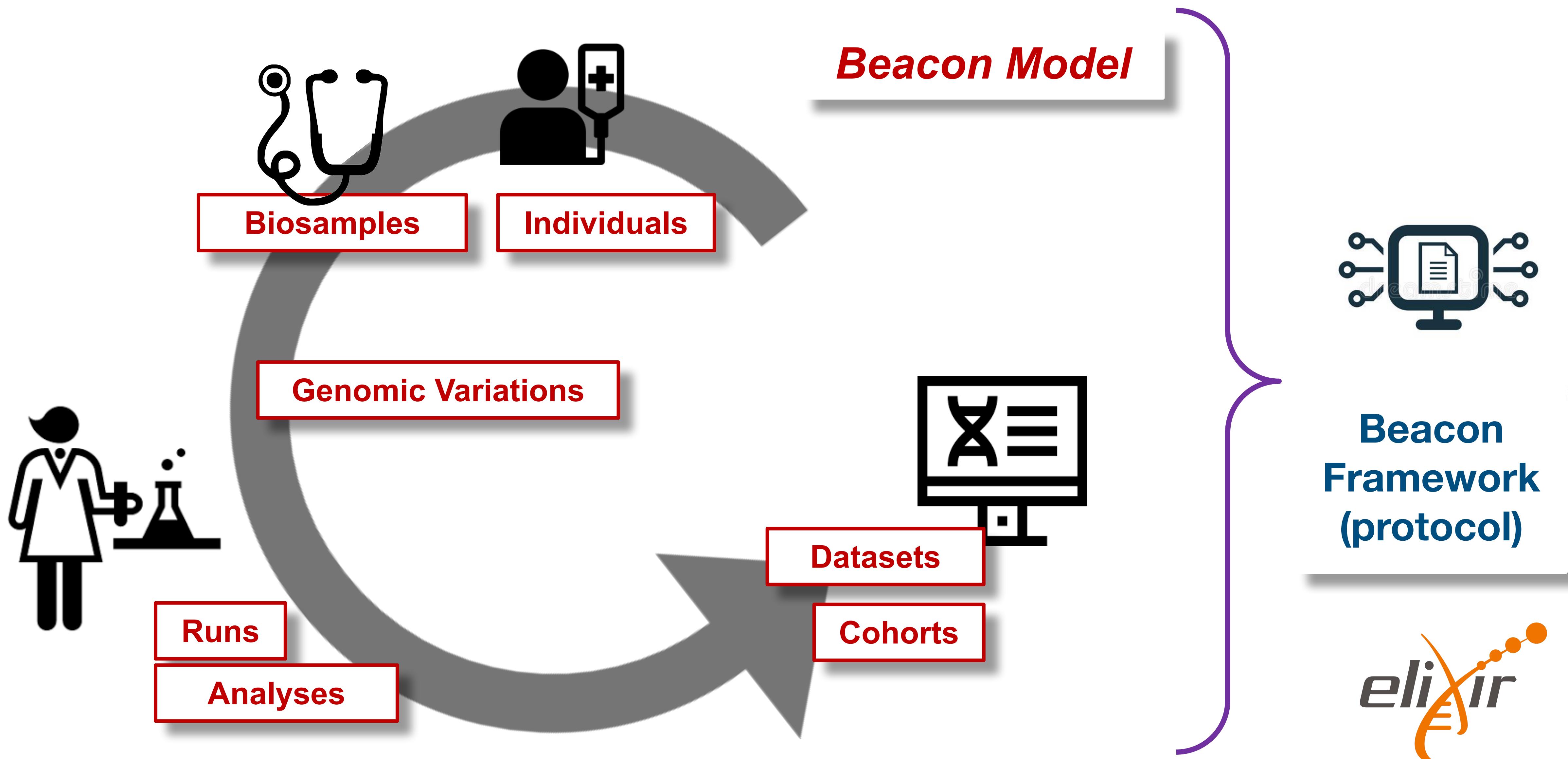


## Beacon v2 API

The Beacon API v2 represents a simple but powerful **genomics API** for **federated** data discovery and retrieval

# Beacon v2

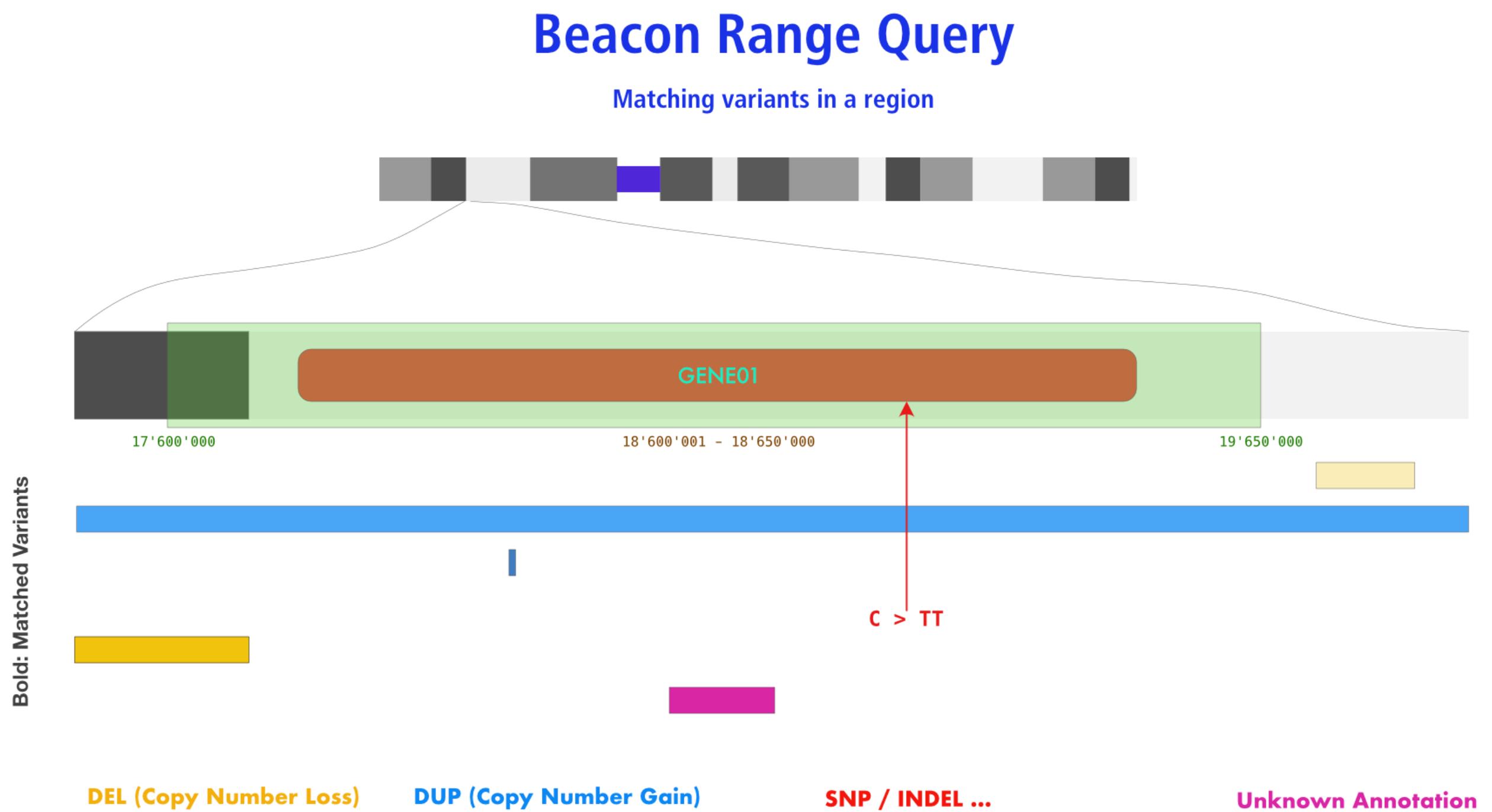
docs.genomebeacons.org



# Variation Queries

## Range ("anything goes") Request

- defined through the use of 1 start, 1 end
- any variant... but can be limited by type etc.



## Beacon Query Types

Sequence / Allele   CNV (Bracket)   **Genomic Range**   Aminoacid   Gene ID   HGVS   Sam

Dataset: Test Database - examplez

Chromosome: 17 (NC\_000017.11)

Variant Type: SO:0001059 (any sequence alteration - S...)

Start or Position: 7572826

End (Range or Structural Var.): 7579005

Reference Base(s): N

Alternate Base(s): A

Select Filters: Chromosome 17

Query Database

Form Utilities: Gene Spans, Cytoband(s)

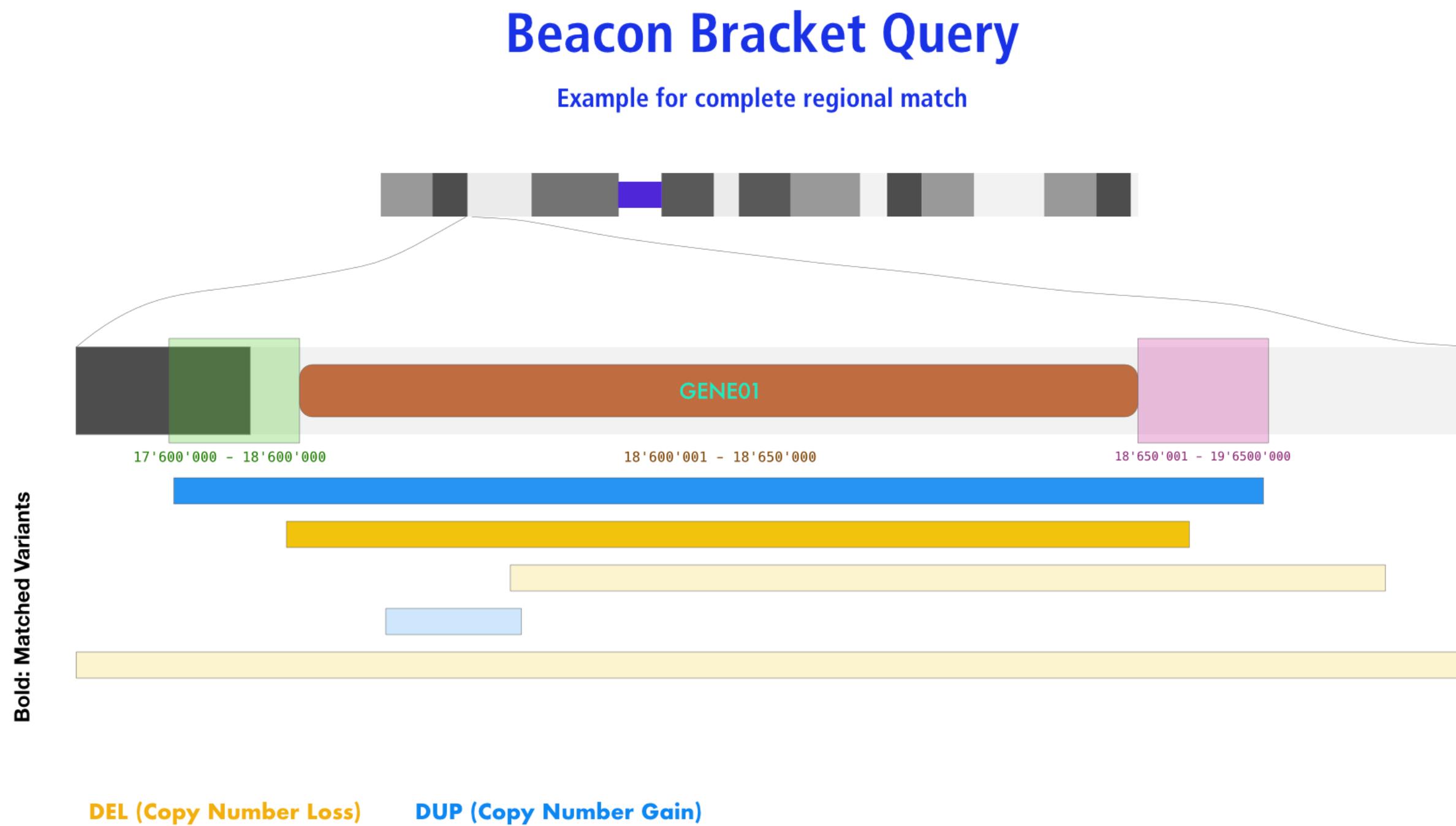
Query Examples: CNV Example, SNV Example, Range Example, Gene Match, Aminoacid Example, Identifier - HeLa

As in the standard SNV query, this example shows a Beacon query against mutations in the EIF4A1 gene in the DIPG childhood brain tumor dataset. However, this range + wildcard query will return any variant with alternate bases (indicated through "N"). Since parameters will be interpreted using an "AND" paradigm, either Alternate Bases OR Variant Type should be specified. The exact variants which were being found can be retrieved through the variant handover [H->O] link.

# Variation Queries

## Bracket ("CNV") Query

- defined through the use of 2 start, 2 end
- any contiguous variant...



### Beacon Query Types

Sequence / Allele   **CNV (Bracket)**   Genomic Range   Aminoacid   Gene ID   HGVS   Sam

#### Dataset

Test Database - examplez X | ▼

#### Chromosome i

9 (NC\_000009.12) | ▼

#### Variant Type i

EFO:0030067 (copy number deletion) | ▼

#### Start or Position i

21000001-21975098

#### End (Range or Structural Var.) i

21967753-23000000

#### Select Filters i

NCIT:C3058: Glioblastoma (100) X | ▼

#### Chromosome 9 i

21000001-21975098



### Query Database

#### Form Utilities

⚙️ Gene Spans

⚙️ Cytoband(s)

#### Query Examples

[CNV Example](#)

[SNV Example](#)

[Range Example](#)

[Gene Match](#)

[Aminoacid Example](#)

[Identifier - HeLa](#)

This example shows the query for CNV deletion variants overlapping the CDKN2A gene's coding region with at least a single base, but limited to "focal" hits (here i.e.  $\leq \sim 2\text{Mbp}$  in size). The query is against the examplez collection and can be modified e.g. through changing the position parameters or data source.

# Contributing to Standards Development: CNV Terms

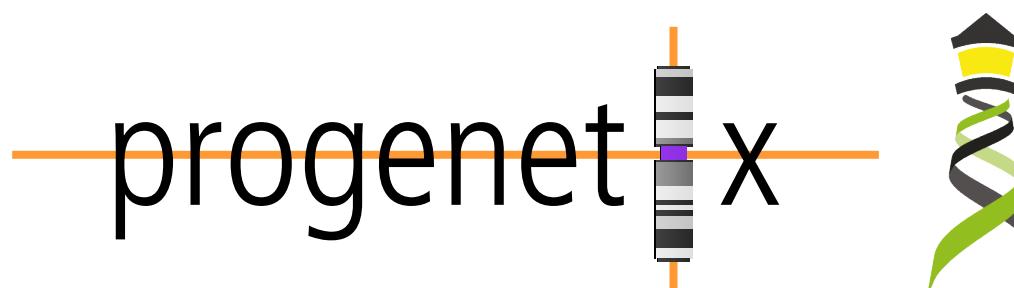
## in computational (file/schema) formats

GA4GH VRS1.3+	EFO	Beacon	VCF	SO
<a href="#">EFO:0030070</a> gain	<a href="#">EFO:0030070</a> copy number gain	DUP or <a href="#">EFO:0030070</a>	DUP SVCLAIM=D	<a href="#">SO:0001742</a> copy_number_gain
<a href="#">EFO:0030071</a> low-level gain	<a href="#">EFO:0030071</a> low-level copy number gain	DUP or <a href="#">EFO:0030071</a>	DUP SVCLAIM=D	<a href="#">SO:0001742</a> copy_number_gain
<a href="#">EFO:0030072</a> high-level gain	<a href="#">EFO:0030072</a> high-level copy number gain	DUP or <a href="#">EFO:0030072</a>	DUP SVCLAIM=D	<a href="#">SO:0001742</a> copy_number_gain
<a href="#">EFO:0030072</a> high-level gain	<a href="#">EFO:0030073</a> focal genome amplification	DUP or <a href="#">EFO:0030073</a>	DUP SVCLAIM=D	<a href="#">SO:0001742</a> copy_number_gain
<a href="#">EFO:0030067</a> loss	<a href="#">EFO:0030067</a> copy number loss	DEL or <a href="#">EFO:0030067</a>	DEL SVCLAIM=D	<a href="#">SO:0001743</a> copy_number_loss
<a href="#">EFO:0030068</a> low-level loss	<a href="#">EFO:0030068</a> low-level copy number loss	DEL or <a href="#">EFO:0030068</a>	DEL SVCLAIM=D	<a href="#">SO:0001743</a> copy_number_loss
<a href="#">EFO:0020073</a> high-level loss	<a href="#">EFO:0020073</a> high-level copy number loss	DEL or <a href="#">EFO:0020073</a>	DEL SVCLAIM=D	<a href="#">SO:0001743</a> copy_number_loss
<a href="#">EFO:0030069</a> complete genomic loss	<a href="#">EFO:0030069</a> complete genomic deletion	DEL or <a href="#">EFO:0030069</a>	DEL SVCLAIM=D	<a href="#">SO:0001743</a> copy_number_loss

# Beacon v2 Filters

# **Example: Use of hierarchical classification systems (here NCI neoplasm core)**

- Beacon v2 relies heavily on "filters"
    - ontology term / CURIE
    - alphanumeric
    - custom
  - Beacon v2 "filters" assumes inclusion of child terms when using hierarchical classifications
    - implicit *OR* with otherwise assumed *AND*
  - implementation of hierarchical annotations overcomes some limitations of "fuzzy" disease annotations



Beacon<sup>+</sup> specific: Multiple term selection with OR logic

<input checked="" type="checkbox"/>	> <a href="#">NCIT:C4914: Skin Carcinoma</a>	213
<input type="checkbox"/>	> <a href="#">NCIT:C4475: Dermal Neoplasm</a>	109
<input checked="" type="checkbox"/>	> <a href="#">NCIT:C45240: Cutaneous Hematopoietic and Lymphoid Cell Neoplasm</a>	310

**Filters:** NCIT:C4914, NCIT:C4819, NCIT:C9231, NCIT:C2921, NCIT:C45240, NCIT:C6858, NCIT:C3467, NCIT:C45340, NCIT:C7195, NCIT:C3246, NCIT:C7217

progenetix

Variants: 0    falleles: 0    Callsets Variants ↗    UCSC region ↗    Calls: 0    Legacy Interface ↗    Samples: 523    [Show JSON Response](#)

Results    **Biosamples**

<b>Id</b>	<b>Description</b>	<b>Classifications</b>	<b>Identifiers</b>	<b>DEL</b>	<b>DUP</b>	<b>CNV</b>
PGX_AM_BS_MCC01	Merkel cell carcinoma	<a href="#">icdot-C44.9</a> Skin, NOS <a href="#">icdom-82473</a> Merkel cell carcinoma <a href="#">NCIT:C9231</a> Merkel Cell Carcinoma	<a href="#">PMID:9537255</a>	0.116	0.104	0.22
PGX_AM_BS_MCC02	Merkel cell carcinoma	<a href="#">icdot-C44.9</a> Skin, NOS <a href="#">icdom-82473</a> Merkel cell carcinoma <a href="#">NCIT:C9231</a> Merkel Cell Carcinoma	<a href="#">PMID:9537255</a>	0.154	0.056	0.21
PGX_AM_BS_MCC03	Merkel cell carcinoma	<a href="#">icdot-C44.9</a> Skin, NOS <a href="#">icdom-82473</a> Merkel cell carcinoma <a href="#">NCIT:C9231</a> Merkel Cell Carcinoma	<a href="#">PMID:9537255</a>	0.137	0.21	0.347
PGX_AM_BS_MCC04	Merkel cell carcinoma	<a href="#">icdot-C44.9</a> Skin, NOS <a href="#">icdom-82473</a> Merkel cell carcinoma <a href="#">NCIT:C9231</a> Merkel Cell Carcinoma	<a href="#">PMID:9537255</a>	0.158	0.056	0.214
PGX_AM_BS_MCC05	Merkel cell carcinoma	<a href="#">icdot-C44.9</a> Skin, NOS <a href="#">icdom-82473</a> Merkel cell carcinoma <a href="#">NCIT:C9231</a> Merkel Cell Carcinoma	<a href="#">PMID:9537255</a>	0.107	0.327	0.434

« < > »

Page 1 of 105

# Begriffsbestimmung

## The right expressions help to conceptualize...

- **Beacon:** The protocol/API, with framework and default model
- **beacon:** Implementation of Beacon
  - using the Beacon v2 framework & supporting at minimum boolean responses
  - suggested support of Beacon v2 default model but can choose other
- Beacon **Aggregator:** service distributes queries to beacons and aggregates responses into a single Beacon response
  - potential to liftover genomes, remap filtering terms, translate between protocol versions...
  - entry point to or potentially itself node in a ...
- Beacon **Network:** Set of beacons with shared entry point for distributed queries and aggregated response delivery
  - "true" beacon networks should have managed aspects - scope, term use...
  - networks may combine mixes of internal (protected, rich data, additional extensions...) and external interfaces

## Response

## Summary

## Meta

### "Aggregator Summaries Response"

requestedGranularity:  
boolean/count

exists  
numTotalResults  
maximumGranularity  
minimumGranularity  
...

resultSets  
- id  
  beaconId  
  infoUrl  
  exists  
  resultsCount  
- id  
  beaconId  
  infoUrl  
  exists  
- id  
  beaconId  
  ...  
...

### Boolean/Count Response

requestedGranularity:  
boolean/count

exists  
(numTotalResults)

### Beacon v2 Aggregator / Network Responses

### Resultsets Response

requestedGranularity:  
record  
pagination

exists  
numTotalResults  
...

resultSets  
- id  
  setType  
  exists  
  resultsCount  
  results []  
- id  
  setType  
  exists  
  resultsCount  
  results []  
- id  
  setType  
  ...  
...

### "Aggregator Records Response"

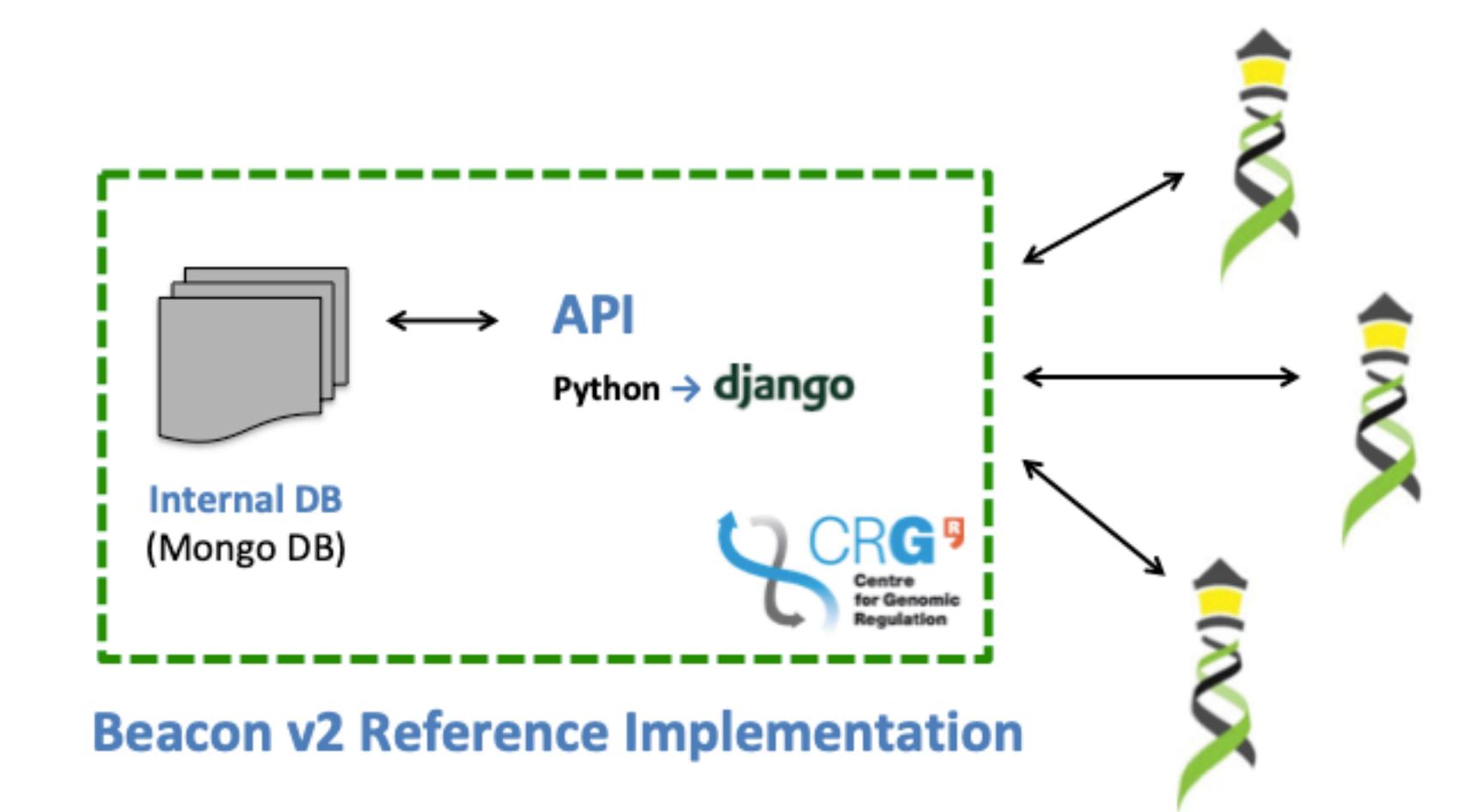
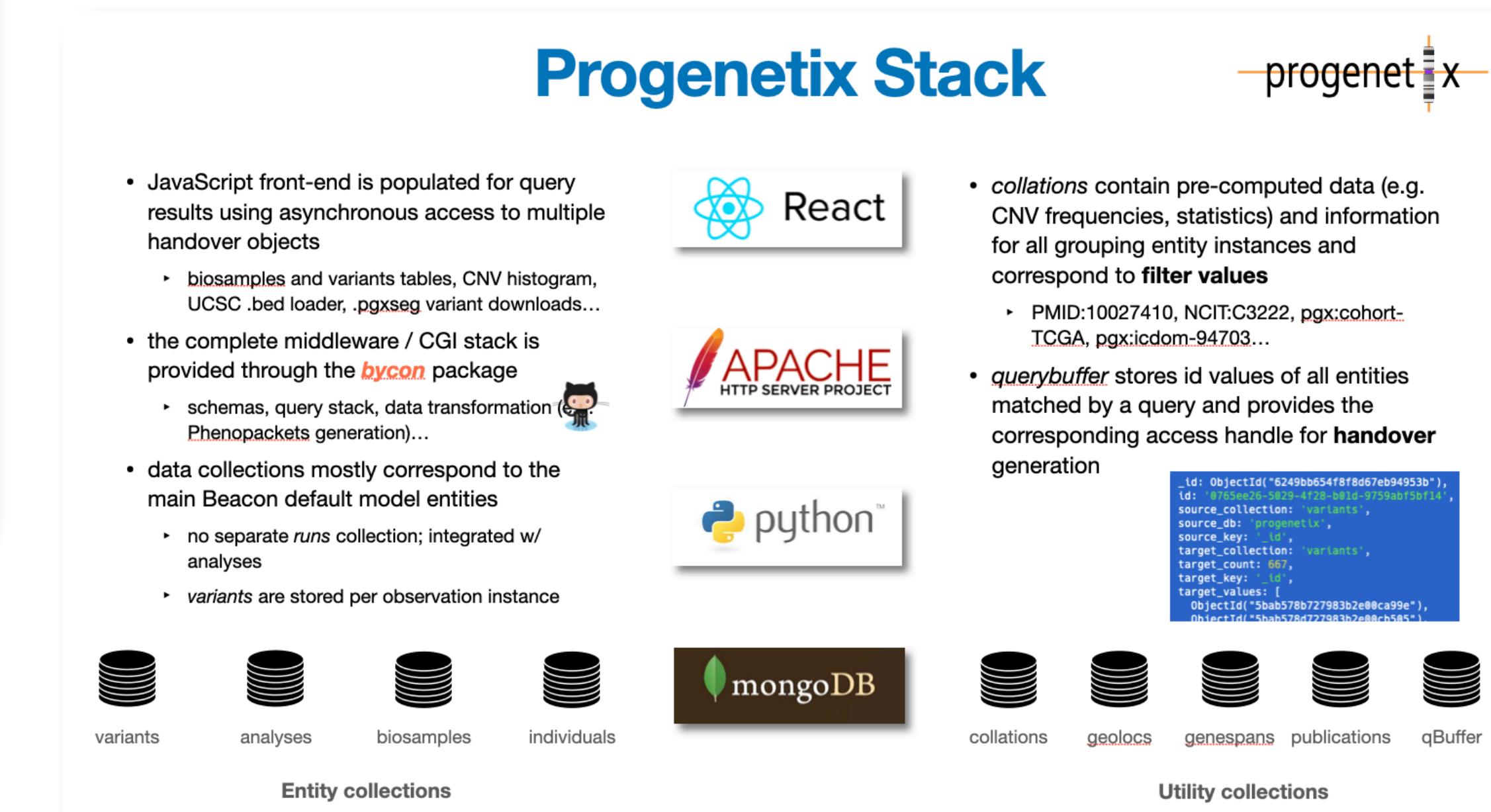
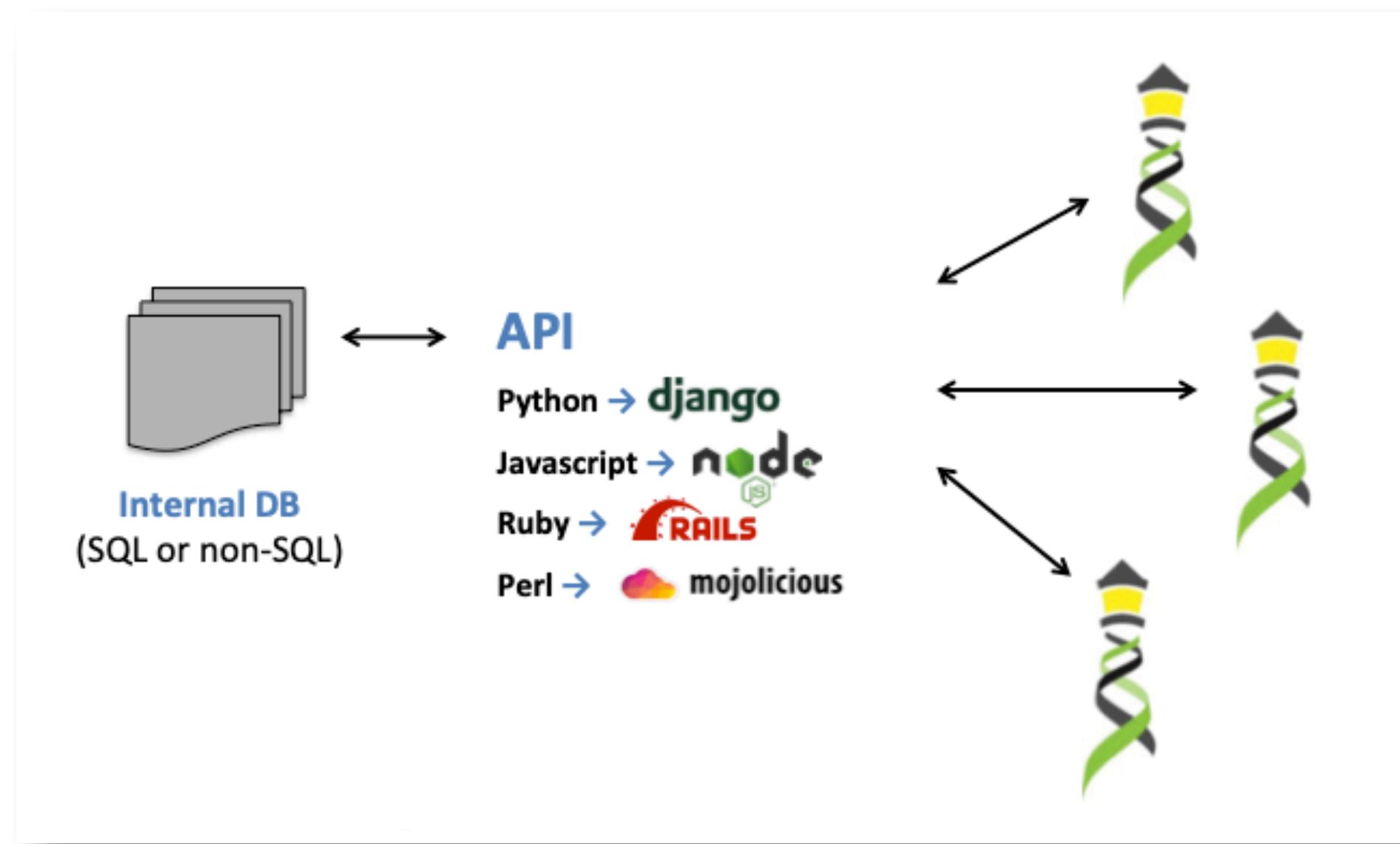
requestedGranularity:  
record  
pagination

exists  
numTotalResults  
maximumGranularity  
minimumGranularity  
...

resultSets  
- id  
  beaconId  
  infoUrl  
  setType  
  exists  
  resultsCount  
  returnedGranularity  
  ...  
  results []  
- id  
  beaconId  
  infoUrl  
  exists  
  ...

# Implementing Beacon v2

... its just code \\_(\_ツ)\_/



# *bycon* for GA4GH Beacon

## Implementation driven development of a GA4GH standard

## Cancer Genomics Reference Resource

- **open** resource for oncogenomic profiles
- over **140'000 cancer CNV profiles**
- SNV data for some series (e.g. TCGA)
- more than **900 diagnostic types**
- inclusion of reference datasets (e.g. TCGA)
- standardized encodings (e.g. NCIIt, ICD-O 3)
- identifier mapping for PMID, GEO, Cellosaurus, TCGA, cBioPortal where appropriate
- core clinical data (TNM, sex, survival ...)
- data mapping services



### Cancer CNV Profiles

ICD-O Morphologies  
ICD-O Organ Sites  
Cancer Cell Lines  
Clinical Categories

### Search Samples

arrayMap  
TCGA Samples  
1000 Genomes  
Reference Samples  
DIPG Samples  
cBioPortal Studies  
Gao & Baudis, 2021

### Publication DB

Genome Profiling  
Progenetix Use

### Services

NCIt Mappings  
UBERON Mappings

### Upload & Plot

### Beacon<sup>+</sup>

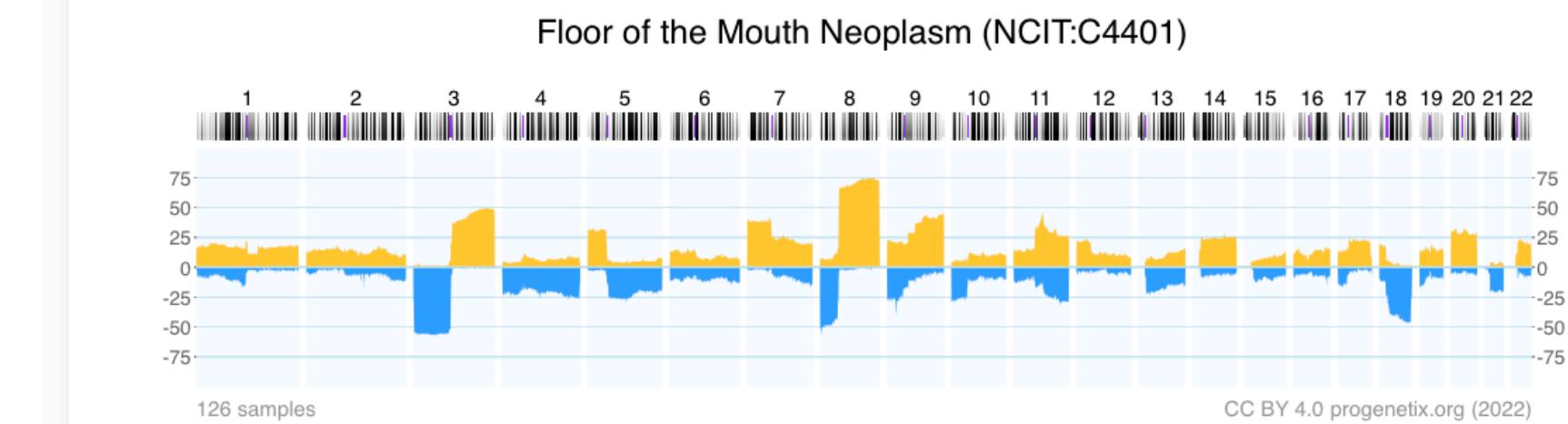
### Documentation

News  
Downloads & Use  
Cases  
Sevices & API

### Baudisgroup @ UZH

## Cancer genome data @ progenetix.org

The Progenetix database provides an overview of mutation data in cancer, with a focus on copy number abnormalities (CNV / CNA), for all types of human malignancies. The data is based on *individual sample data* from currently **142063** samples.



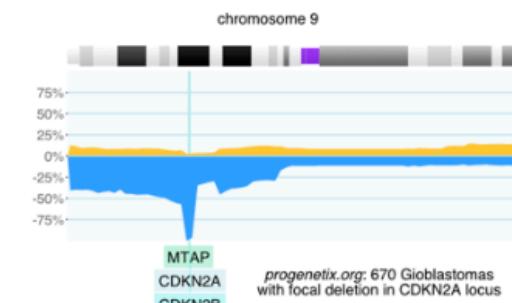
[Download SVG](#) | [Go to NCIT:C4401](#) | [Download CNV Frequencies](#)

Example for aggregated CNV data in 126 samples in Floor of the Mouth Neoplasm.  
Here the frequency of regional **copy number gains** and **losses** are displayed for all 22 autosomes.

### Progenetix Use Cases

#### Local CNV Frequencies

A typical use case on Progenetix is the search for local copy number aberrations - e.g. involving a gene - and the exploration of cancer types with these CNVs. The [\[ Search Page \]](#) provides example use cases for designing queries. Results contain basic statistics as well as visualization and download options.



progenetix.org: 670 Glioblastomas with local deletion in CDKN2A locus

#### Cancer CNV Profiles

The progenetix resource contains data of **834** different cancer types (NCIt neoplasm classification), mapped to a variety of biological and technical categories. Frequency profiles of regional genomic gains and losses for all categories (diagnostic entity, publication, cohort ...) can be accessed through the [\[ Cancer Types \]](#) page with direct visualization and options for sample retrieval and plotting options.

#### Cancer Genomics Publications

Through the [\[ Publications \]](#) page Progenetix provides **4164** annotated references to research articles from cancer genome screening experiments (WGS, WES, aCGH, cCGH). The numbers of analyzed samples and possible availability in the Progenetix sample collection are indicated.

# bycon Beacon

## Implementation driven standards development

- Progenetix' Beacon+ has served as implementation driver since 2016
- the *bycon* package is used to prototype advanced Beacon features such as
  - structural variant queries
  - data handovers
  - Phenopackets integration
  - variant co-occurrences
  - ...

Beacon protocol response verifier at time of GA4GH approval Spring 2022

Beacon v2 GA4GH Approval Registry

Beacons:    

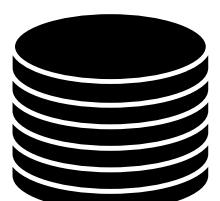
Category	EGA	progenetix	cnag	University of Leicester
BeaconMap	Green	Green	Green	Green
Bioinformatics analysis	Green	Green	Green	Green
Biological Sample	Green	Red	Red	Green
Cohort	Green	Green	Green	Green
Configuration	Green	Green	Green	Green
Dataset	Green	Red	Red	Green
EntryTypes	Green	Green	Green	Green
Genomic Variants	Green	Green	Green	Green
Individual	Green	Red	Red	Green
Info	Green	Red	Red	Green
Sequencing run	Green	Green	Green	Green

Legend:  Matches the Spec  Not Match the Spec  Not Implemented

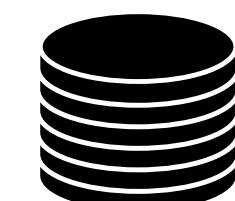
# *bycon* based Progenetix Stack



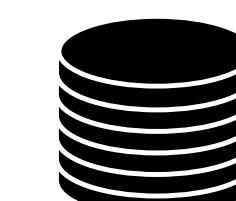
- JavaScript front-end is populated for query results using asynchronous access to multiple handover objects
  - ▶ biosamples and variants tables, CNV histogram, UCSC .bed loader, .pgxseg variant downloads...
- the complete middleware / CGI stack is provided through the *bycon* package
  - ▶ schemas, query stack, data transformation (Phenopackets generation)...
- data collections mostly correspond to the main Beacon default model entities
  - ▶ no separate *runs* collection; integrated w/ analyses
  - ▶ *variants* are stored per observation instance



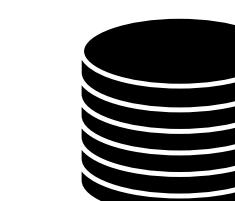
variants



analyses



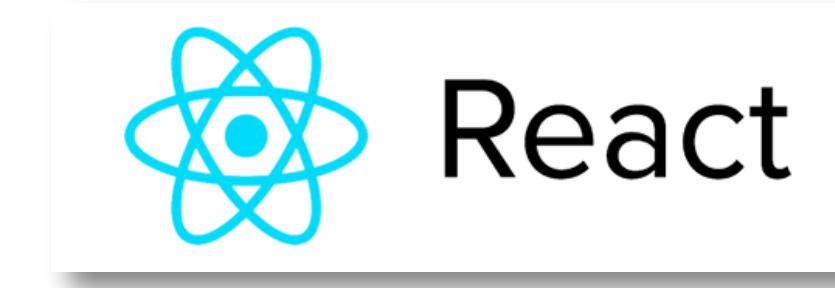
biosamples



individuals

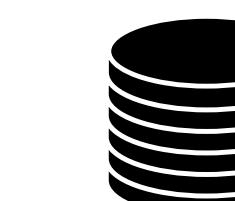


Entity collections

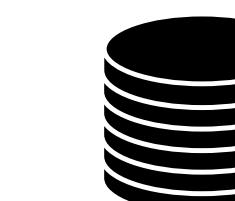


- *collations* contain pre-computed data (e.g. CNV frequencies, statistics) and information for all grouping entity instances and correspond to **filter values**
  - ▶ PMID:10027410, NCIT:C3222, pgx:cohort-TCGA, pgx:icdom-94703...
- *querybuffer* stores id values of all entities matched by a query and provides the corresponding access handle for **handover** generation

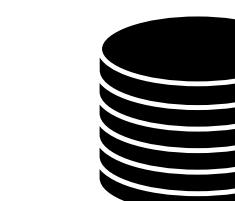
```
_id: ObjectId("6249bb654f8f8d67eb94953b"),
id: '0765ee26-5029-4f28-b01d-9759abf5bf14',
source_collection: 'variants',
source_db: 'progenetix',
source_key: '_id',
target_collection: 'variants',
target_count: 667,
target_key: '_id',
target_values: [
  ObjectId("5bab578b727983b2e00ca99e"),
  ObjectId("5bab578d727983b2e00cb505")]
```



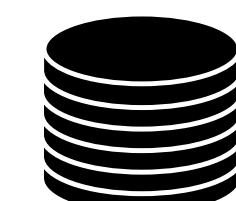
collations



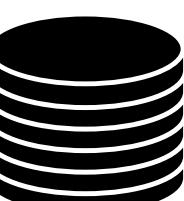
geolocs



genespans



publications



qBuffer

Utility collections

# Client for Accessing Progenetix

## pgxRpi: an R/Bioconductor package

- **Query and export variants**

[https://progenetix.org/beacon/biosamples/pgxbs-kftvh94d/g\\_variants](https://progenetix.org/beacon/biosamples/pgxbs-kftvh94d/g_variants)

```
> variants <- pgxLoader(type="variant",biosample_id="pgxbs-kftvh94d")
```

- **Query metadata of biosamples and individuals by filters (e.g. NCIt, PMID)**

<http://progenetix.org/services/sampletable/?filters=NCIT:C3697>

```
> biosamples <- pgxLoader(type="biosample",filters="NCIT:C3697")
```

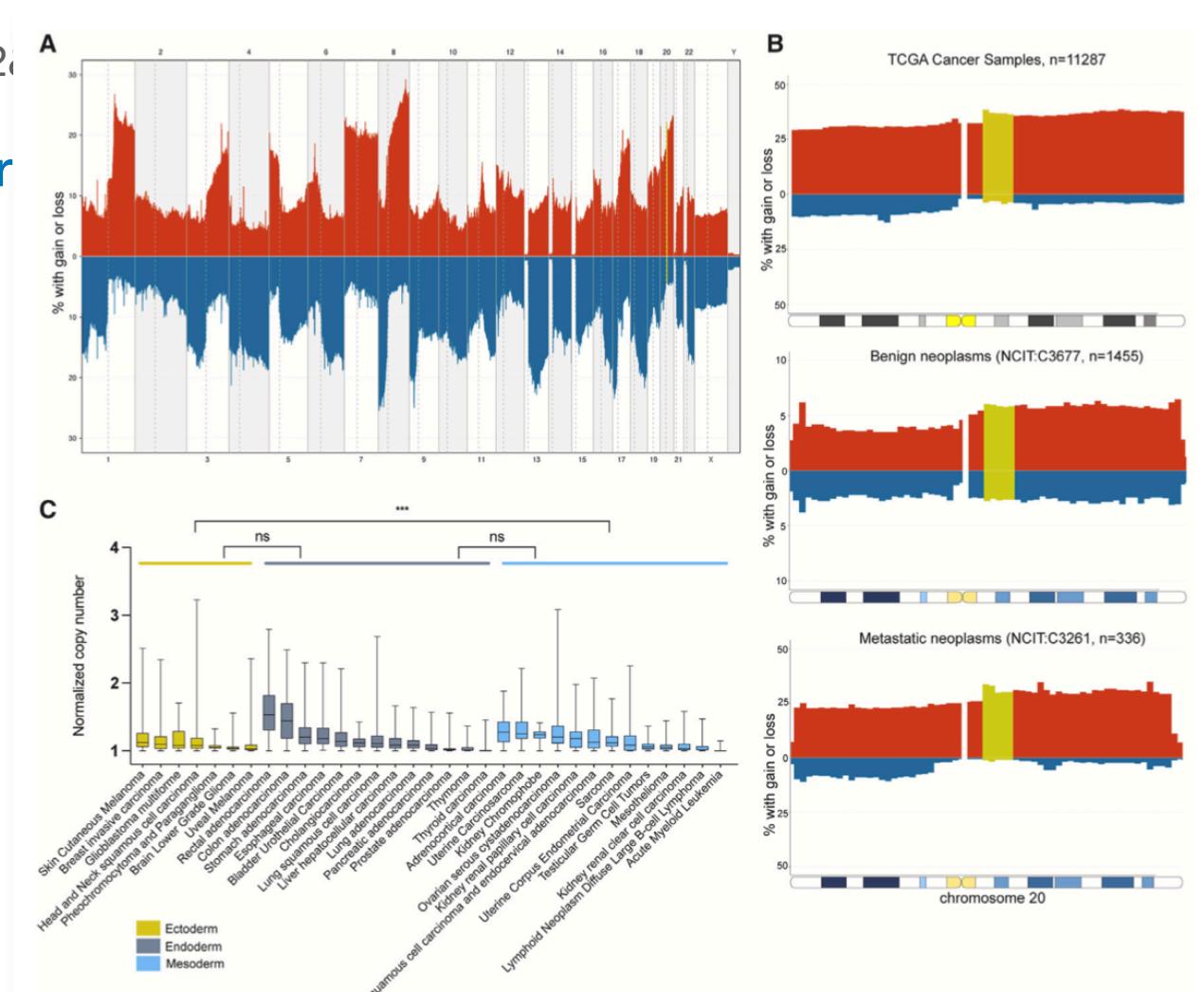
- **Query and visualize CNV frequency by filters**

<http://www.progenetix.org/services/intervalFrequencies/?filters=NCIT:C35126>

```
> freq <- pgxLoader(type="frequency",output="pgxfreq",filter  
> pgxFreqplot(freq)
```

- **Process local .pgxseg files**

```
> info <- pgxSegprocess(file=file, show_KM_plot = T,  
return_seg = T, return_metadata = T, return_frequency = T)
```



## pgxRpi

This is the **development** version of pgxRpi; for the stable release version, see [pgxRpi](#).

### R wrapper for Progenetix

platforms all rank 2178 / 2266 support 0 / 0 in Bioc < 6 months build unknown updated < 1 month dependencies 137  
DOI: [10.18129/B9.bioc.pgxRpi](https://doi.org/10.18129/B9.bioc.pgxRpi)

**Bioconductor version:** Development (3.20)

The package is an R wrapper for Progenetix REST API built upon the Beacon v2 protocol. Its purpose is to provide a seamless way for retrieving genomic data from Progenetix database—an open resource dedicated to curated oncogenomic profiles. Empowered by this package, users can effortlessly access and visualize data from Progenetix.

**Author:** Hangjia Zhao [aut, cre] Michael Baudis [aut]

**Maintainer:** Hangjia Zhao <[hangjia.zhao@uzh.ch](mailto:hangjia.zhao@uzh.ch)>

Use case: 2024 article using Progenetix' **pgxRpi** to retrieve & visualize 117'587 cancer CNV profiles for a study into pluripotent stem cells' genomics

## Stem Cell Reports Review



OPEN ACCESS

### Gains of 20q11.21 in human pluripotent stem cells: Insights from cancer research

Nuša Krivec,<sup>1,2</sup> Manjusha S. Ghosh,<sup>1,2</sup> and Claudia Spits<sup>1,2,\*</sup>

<sup>1</sup>Research Group Reproduction and Genetics, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel, Brussels, Laarbeeklaan 103, 1090 Brussels, Belgium

<sup>2</sup>These authors contributed equally

\*Correspondence: [claudia.spits@vub.be](mailto:claudia.spits@vub.be)

<https://doi.org/10.1016/j.stemcr.2023.11.013>

#### Figure 2. Copy-number alterations of human chromosome 20q11.21 in cancers

(A) Aggregated copy-number variation (CNV) data of 117,587 neoplasms (NCIT: C3262) from the Progenetix database (Huang et al., 2021) were plotted using R library pgxRpi. The percentage of samples with aberrations (red, gain; blue, loss) for the whole chromosome are indicated on the y axis. Chromosomal regions are depicted on the x axis; the minimal region of interest at chr20:31216079–35871578 is marked in moss green. NCIT, National Cancer Institute Thesaurus.

(B) Top to bottom: Aggregated CNV data of 11,287 TCGA cancer samples, 336 metastatic neoplasms (NCIT: C3261), and 1,455 benign neoplasms (NCIT: C3677) from the Progenetix database (Huang et al., 2021), respectively, were plotted using R library pgxRpi. The percentage of samples with aberrations (red, gain; blue, loss) for the whole chromosome are indicated on the y axis. Chromosomal regions are depicted on the x axis; the minimal region of interest at chr20:31216079–35871578 is marked in moss green.

# **Beacon Security**



# Making Beacons Biomedical - Beacon v2

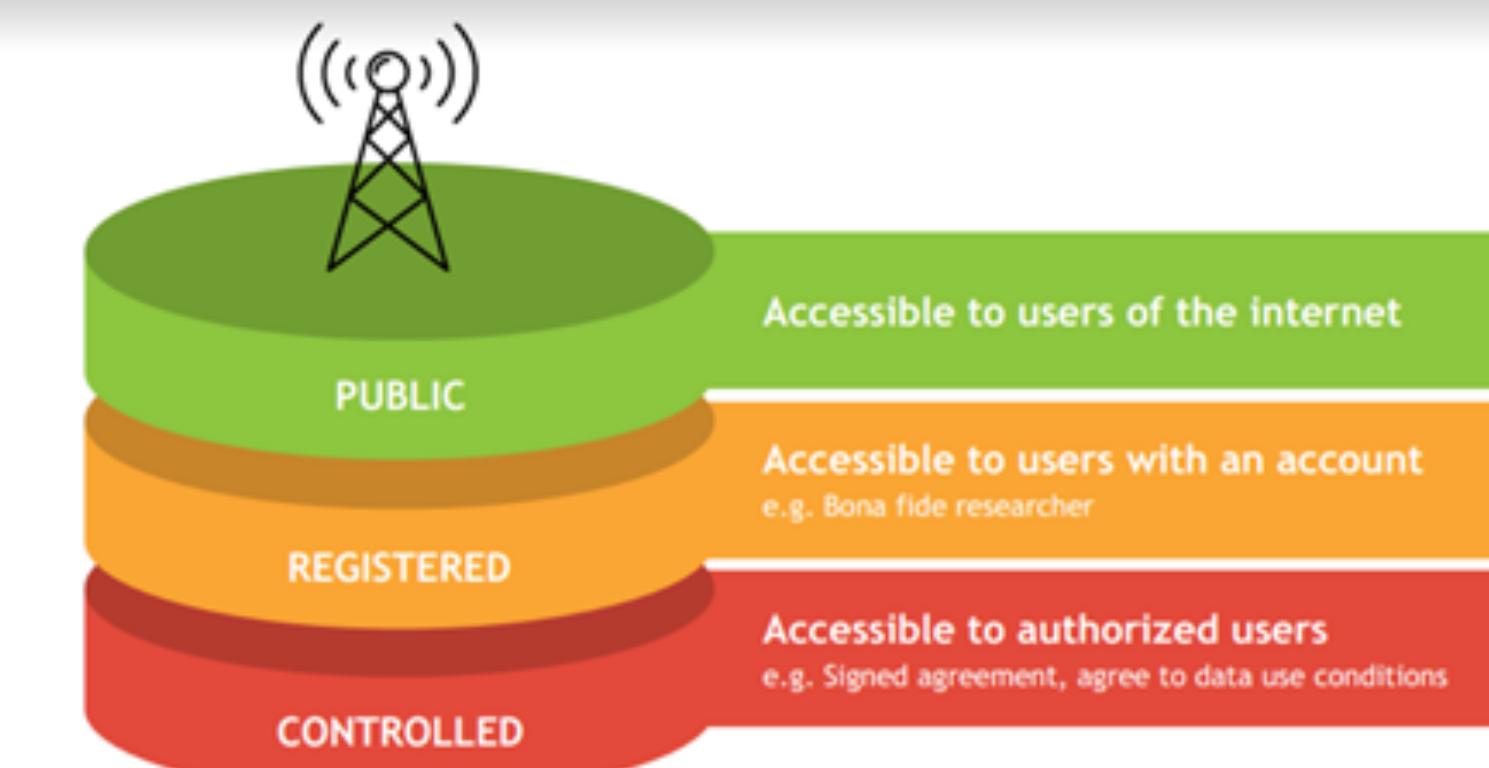
- Scoping queries through "biodata" parameters
- Extending the queries towards clinically ubiquitous variant formats
  - cytogenetic annotations, named variants, variant effects
- Beacon queries as entry for **data delivery**
  - Beacon v2 permissive to respond with variety of data types
    - Phenopackets, biosample data, cohort information ...
  - handover to stream and download using htsget, VCF, EHRs
- Interacting with EHR standards
  - FHIR translations for queries and handover ...
- Beacons as part of local, secure environments
- Authentication to enable non-aggregate, patient derived datasets
  - ELIXIR AAI with compatibility to other providers (OAuth...)

Definitely breaks the  
"Relative Security  
by Design"  
Concept!

# Beacon Security

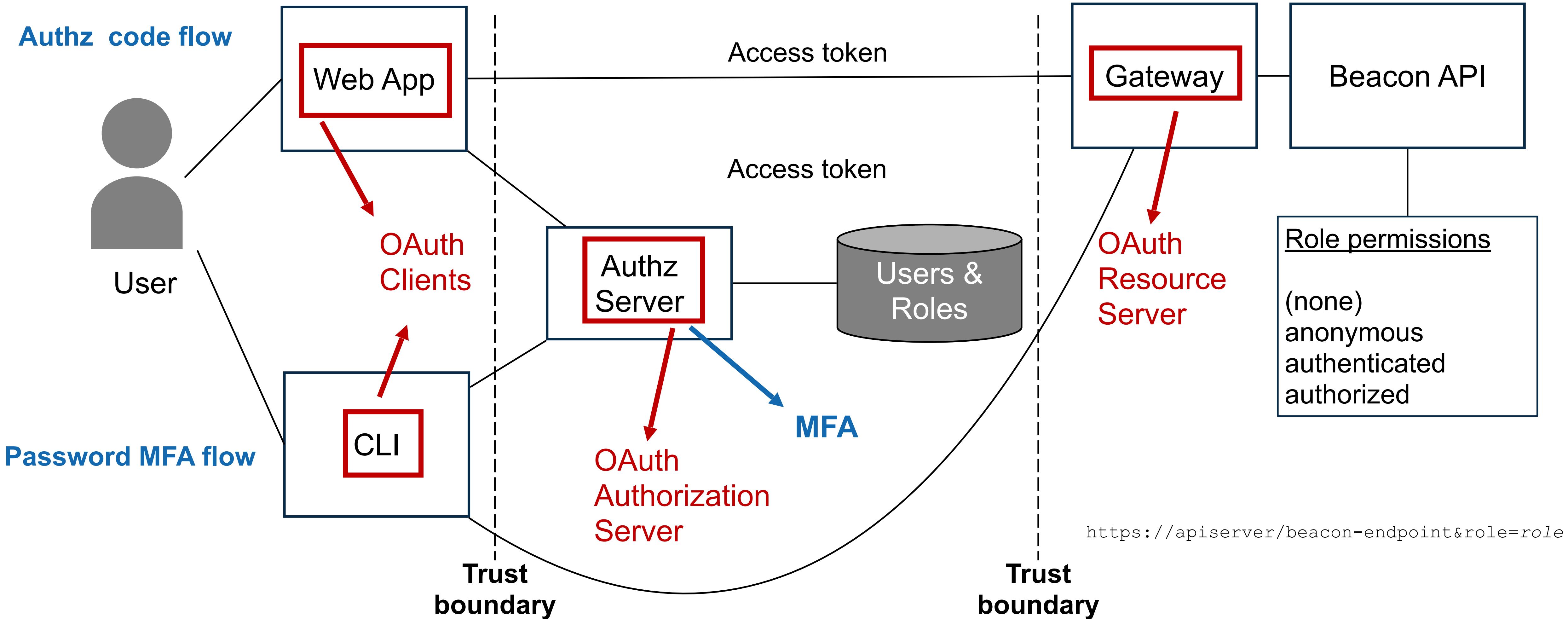
## Security by Design ... if Implemented in the Environment

- the beacon API specification does not implement explicit security (e.g. checking user authentication and authorization)
- the framework implements different levels of response granularity which can be mapped to authorization levels (**boolean** / **count** / **record** level responses)
- implementations can have beacons running in secure environments with a **gatekeeper** service managing authentication and authorization levels, and potentially can filter responses for escalated levels
- the backend can implement additional access reduction, on a user <-> dataset level if needed



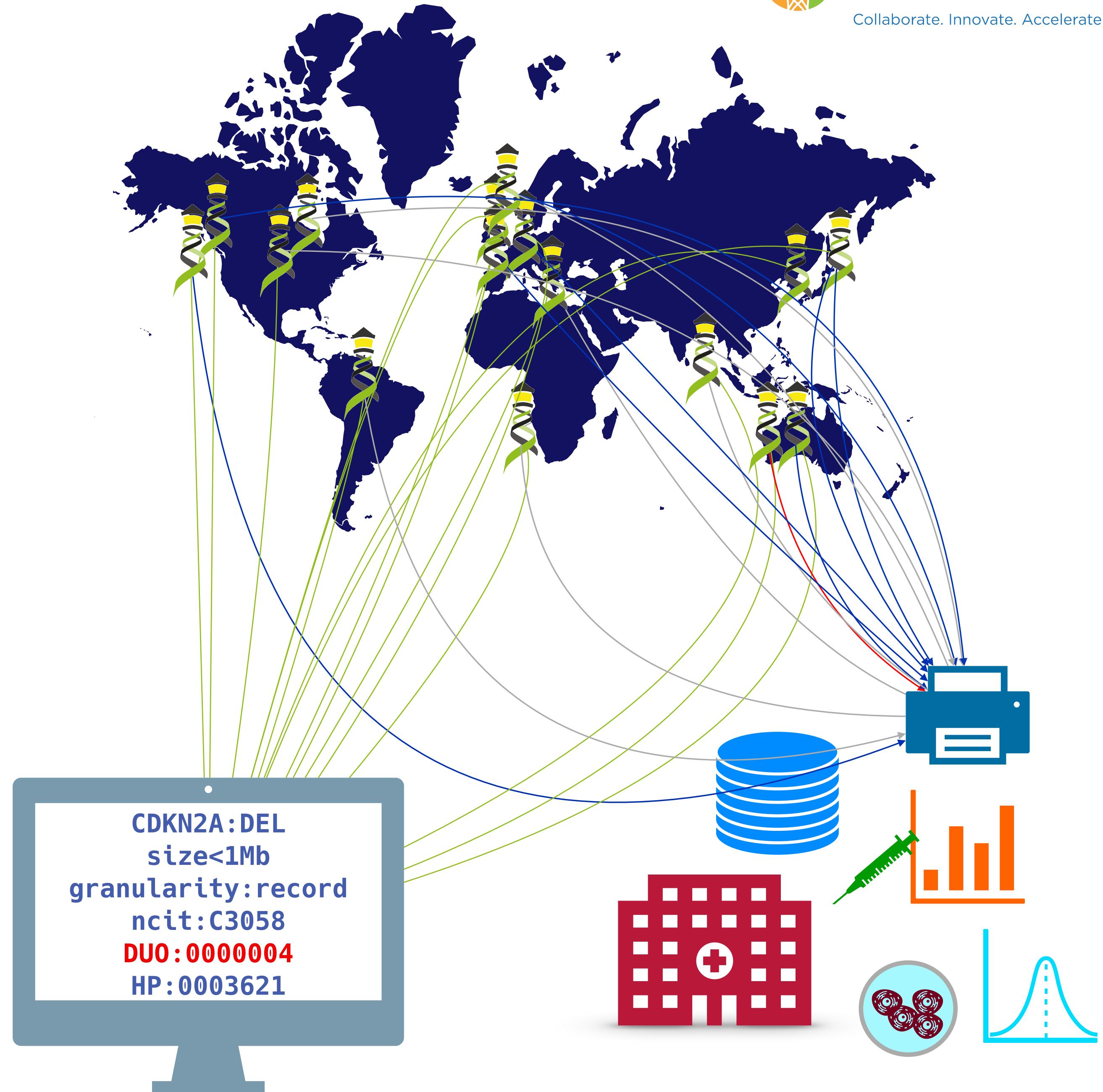
# Architecture

## Running the *bycon* stack in a secure environment



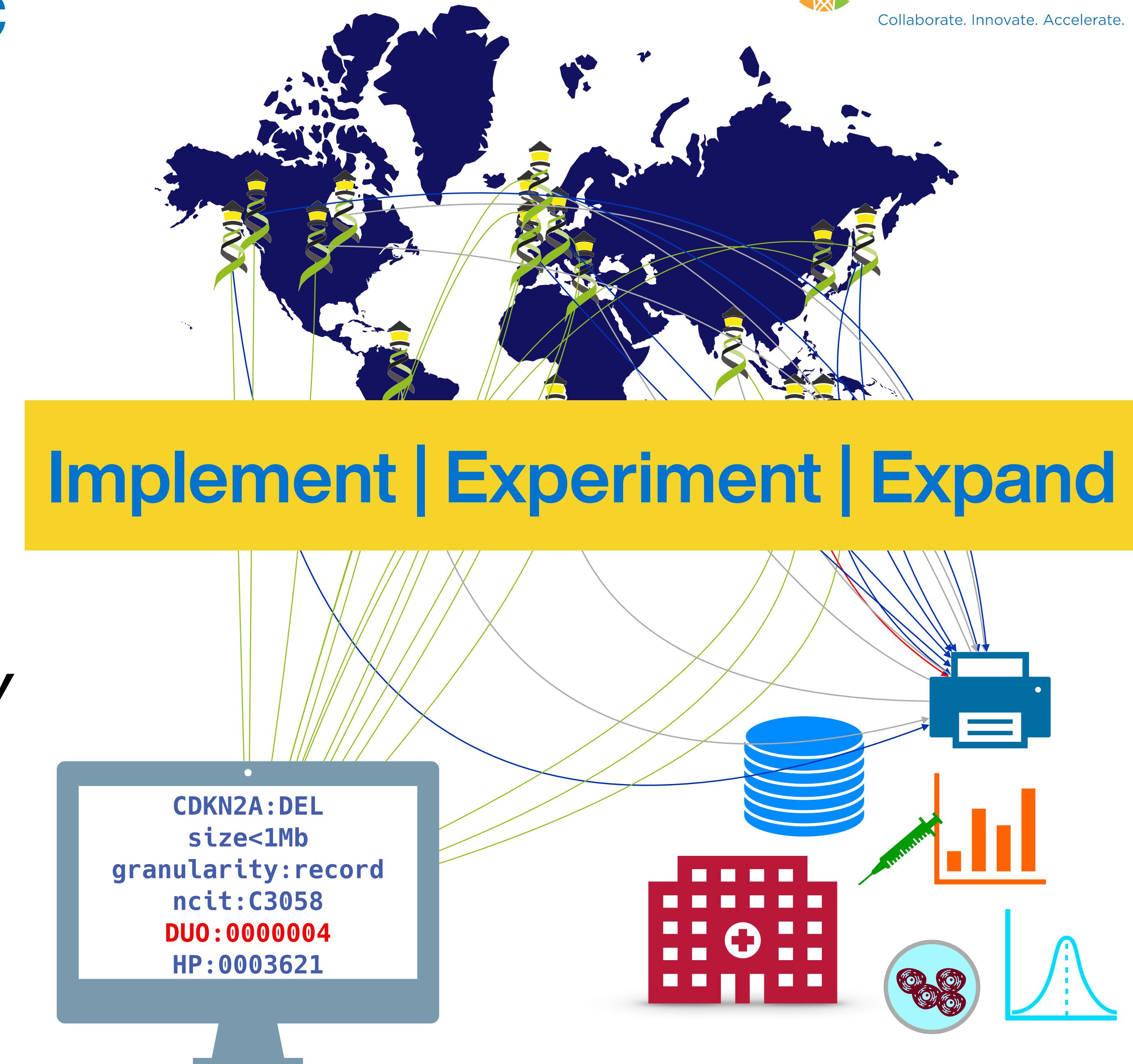
# What Can You Do?

- find a way to make your (patients') **data discoverable** - through adding *at least* the relevant metadata to national or project centric repositories
- use forward looking consent and data protection models (**ORD** principle "as secure as necessary, as open as possible")
- **support** and/or get involved with international **data standards** efforts and projects



# Beacon for Genomic Discovery Proxies

- Feature beacons for privacy protecting data discovery
  - privacy protection through aggregated data, cohorts
  - alternative is "**horizontal gatekeeping**": separate Beacons for ***discovery*** of e.g. genomic and phenotypic data and ***data delivery*** upon request / authentication
  - We'd love to help launching your beacon (especially as a **bycon...**)



# Save the dates!



**Global Alliance**  
for Genomics & Health

## April Connect 2025

1 to 4 April 2025

Broad Institute, Cambridge, USA

[Registration Open Now](#)



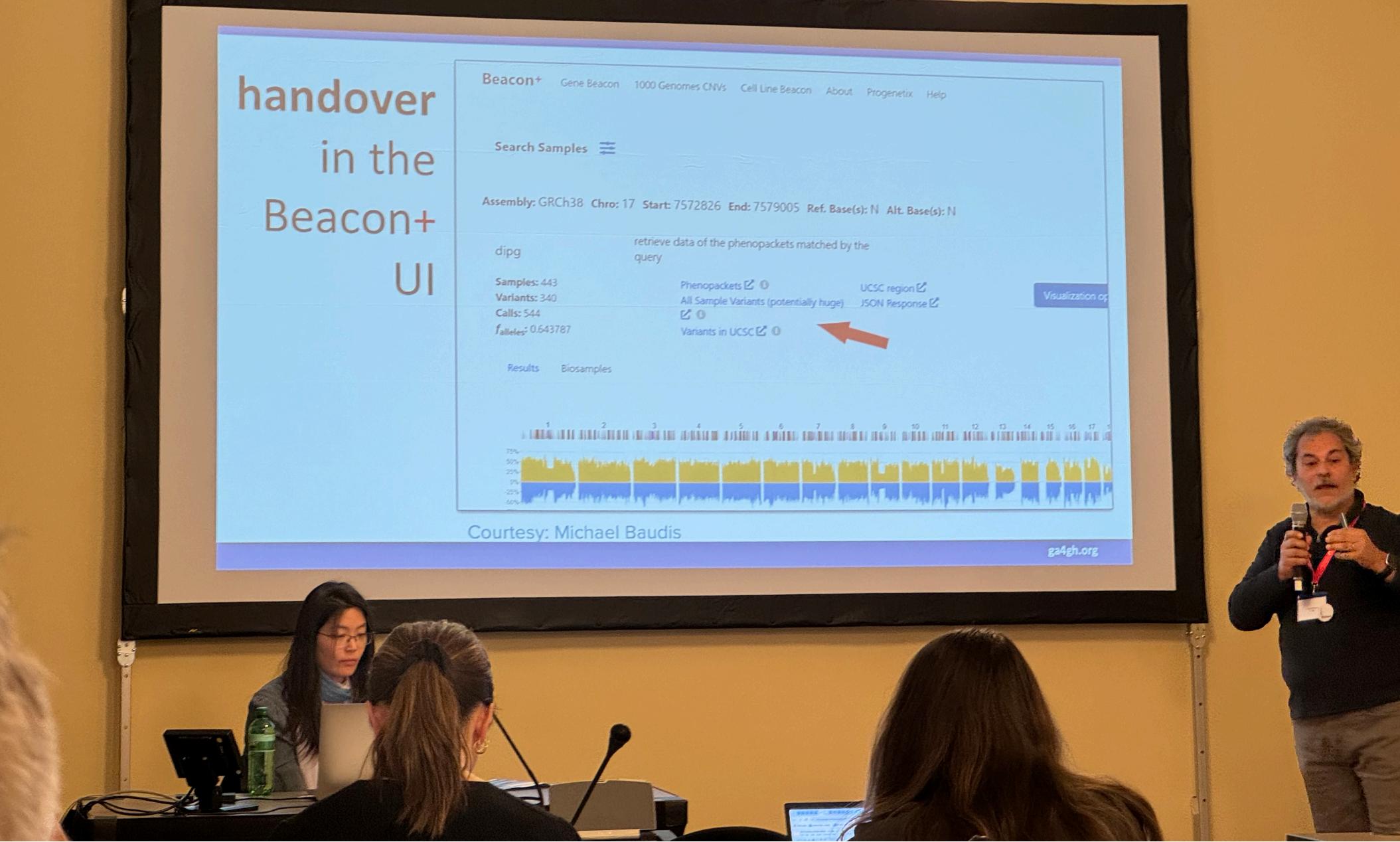
## 13th Plenary

6 to 10 October

UKK, Uppsala, Sweden

Registration Opening Soon





The Global Alliance for Genomics and Health (GA4GH) gathered for the 2024 [April Connect meeting](#) in Ascona, Switzerland and online from 21 to 24 April. The GA4GH Connect meetings provide an opportunity for contributors to advance the GA4GH Road Map, showcase GA4GH standards and policies in action, and gather feedback on product development and community needs. The meeting brought together 103 in-person attendees and 312 virtual attendees for updates from Work Streams and Driver Projects, breakout sessions, and themed events.





University of  
Zurich UZH

UNIVERSITY OF  
LEICESTER



Swiss Institute of  
Bioinformatics

Michael Baudis

Hangjia Zhao

Ziying Yang

Ramon Benitez

Brito

Rahel Paloots

Bo Gao

Qingyao Huang



Jordi Rambla

Arcadi Navarro

Roberto Ariosa

Manuel Rueda

Lauren Fromont

Mauricio Moldes

Claudia Vasallo

Babita Singh

Sabela de la Torre

Fred Haziza

Cafe  
CV Variome  
Central

Tony Brookes

Tim Beck

Colin Veal

Tom Shorter



Juha Törnroos

Teemu Kataja

Ikkka Lappalainen

Dylan Spalding



Augusto Rendon

Ignacio Medina

Javier López

Jacobo Coll

Antonio Rueda

cnag

centre nacional d'anàlisi genòmica  
centro nacional de análisis genómico

Sergi Beltran

Carles Hernandez

Inserm

Institut national  
de la santé et de la recherche médicale

David Salgado

BSC  
Barcelona  
Supercomputing  
Center  
Centro Nacional de Supercomputación

Salvador Capella

Dmitry Repchevski

JM Fernández

DisGeNET

Laura Furlong

Janet Piñero



Serena Scollen

Gary Saunders

Giselle Kerry

David Lloyd

H3Africa  
Human Heredity & Health in Africa

Nicola Mulder

Mamana

Mbiyavanga

Ziyaad Parker

•EU Can

CAN.

David  
Torrents

AUTISM  
SPEAKS®  
Dean Hartley



Fundación Progreso y Salud  
CONSEJERÍA DE SALUD

Joaquin Dopazo

Javier Pérez

J.L. Fernández

Gema Roldan

CINECA

Thomas Keane

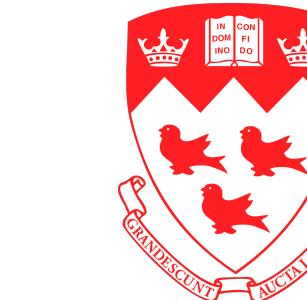
Melanie Courtot

Jonathan Dursi

Heidi Rehm  
Ben Hutton

GEM Japan

Toshiaki  
Katayama



Stephane Dyke

DNA STACK

Marc Fiume

Miro Cupak

BRCA  
EXCHANGE

Melissa Cline

ENA

EMBL-EBI

Diana Lemos

EUROPEAN JOINT PROGRAMME  
RARE DISEASES

VICC Variant Interpretation  
for Cancer Consortium

GA4GH Phenopackets

Peter Robinson

Jules Jacobsen

GA4GH VRS  
Alex Wagner  
Reece Hart

Beacon PRC

Alex Wagner

Jonathan Dursi

Mamana Mbiyavanga

Alice Mann

Neerjah Skantharajah

elixir

# The Beacon team through the ages