

A Genomics / CNV Resource Built on GA4GH Standards

... also Implementation Driven Standards Development

Cancer Genomics Reference Resource

- **open** resource for oncogenomic profiles
- over **116'000** cancer CNV profiles
- more than **800** diagnostic types
- inclusion of reference datasets (e.g. TCGA)
- standardized encodings (e.g. NCIt, ICD-O 3)
- identifier mapping for PMID, GEO, Cellosaurus, TCGA, cBioPortal where appropriate
- core clinical data (TNM, sex, survival ...)
- data mapping services
- recent addition of SNV data for some series

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⁺

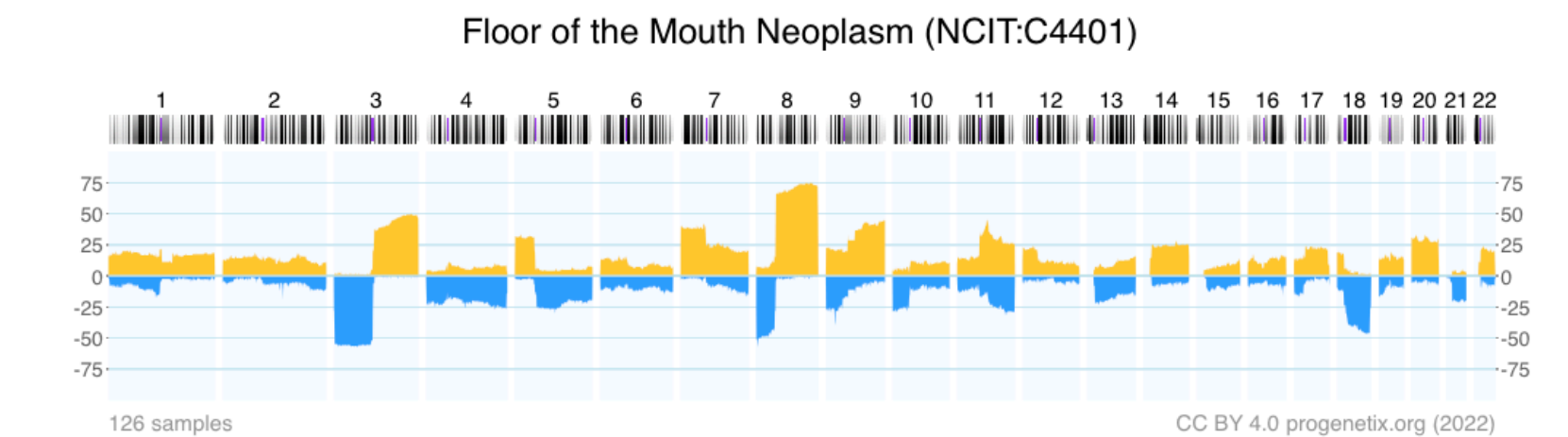
Documentation

News
Downloads & Use
Cases
Services & 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.



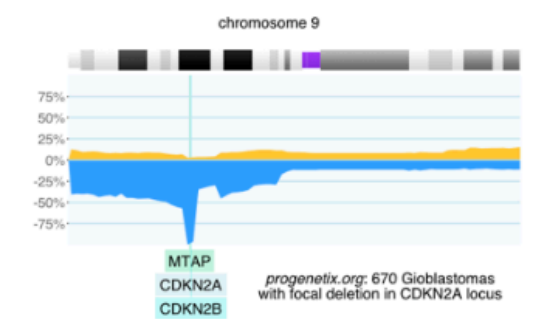
[Download SVG](#) | [Go to NCIT:C4401](#) | [Download CNV Frequencies](#)

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.



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.

Cancer Cell Lines

Cancer Genomics Reference Resource

- starting from >5000 cell line CNV profiles
 - 5754 samples | 2163 cell lines
 - 256 different NCIT codes
- genomic mapping of annotated variants and additional data from several resources (ClinVar, CCLE, Cellosaurus...)
 - 16178 cell lines
 - 400 different NCIT codes
- query and data delivery through Beacon v2 API

➔ integration in data federation approaches

The screenshot displays the cancerellines.org interface. On the left is a navigation sidebar with categories like 'Cancer Cell Lines', 'Search Cell Lines', 'Cell Line Listing', 'CNV Profiles by Cancer Type', 'Documentation', 'Progenetix', and 'Baudisgroup @ UZH'. The main content area is titled 'Cancer Cell Lines by Cellosaurus ID' and explains the hierarchical labeling of cell lines. Below this is a 'Cell Lines (with parental/derived hierarchies)' section with a search filter and a list of cell lines such as 'cellosaurus:CVCL_0312: HOS (204 samples)'. A 'Cell Line Details' pop-up for HOS (cellosaurus:CVCL_0312) shows 'Subset Type', 'Sample Counts' (204 samples, 57 direct matches, 21 CNV analyses), and 'Search Samples'. A 'Raw Data' section features a CNV plot for 21 samples across chromosomes 1-22. Below the plot is a table of 'Gene Matches' including ALK and AREG, with associated clinical notes and abstract links.

Assembly: GRCh38 Chro: NC_000007.14 Start: 140713328 End: 140924929

Type: SNV

cellz
Matched Samples: 1058
Retrieved Samples: 1000
Variants: 127
Calls: 1444

UCSC region
Variants in UCSC
Dataset Responses (JSON)

Visualization options

Digest	Gene	Pathogenicity	Variant type	Variant Instances
7:140834768-140834769:G>A	BRAF		Missense variant	V: pgxvar-63ce6abca24c83054b B: pgxbs-3DfBeeAC
7:140734714-140734715:G>A	BRAF		Missense variant	V: pgxvar-63ce6acda24c83054b B: pgxbs-3fB2a14B
7:140753334-140753339:T>TGTA	BRAF	Pathogenic		V: pgxvar-63ce6a903319d2172d2

Cell Line Details

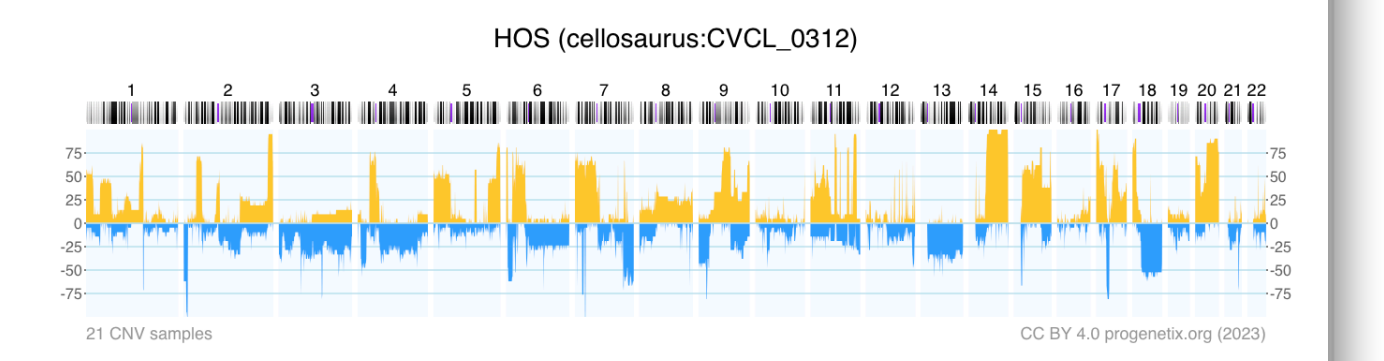
HOS (cellosaurus:CVCL_0312)

Subset Type
• Cellosaurus - a knowledge resource on cell lines [cellosaurus:CVCL_0312](#)

Sample Counts
• 204 samples
• 57 direct *cellosaurus:CVCL_0312* code matches
• 21 CNV analyses

Search Samples
Select *cellosaurus:CVCL_0312* samples in the [Search Form](#)

Raw Data (click to show/hide)



[Download SVG](#) | [Go to cellosaurus:CVCL_0312](#) | [Download CNV Frequencies](#)

Gene Matches Cytoband Matches Variants

Gene	Description	Abstract
ALK	. ABC-14 cells harbored no ALK mutations and were sensitive to ... crizotinib while also exhibiting MNGG HOS transforming gene (MET)	Rapid Acquisition of Alectinib Resistance in ALK-Positive Lung Cancer With High Tumor Mutation Burden (31374369) ABSTRACT

AREG	crizotinib while also exhibiting MNGG HOS	Rapid Acquisition of Alectinib Resistance ABSTRACT
------	---	--

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 (e.g. 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



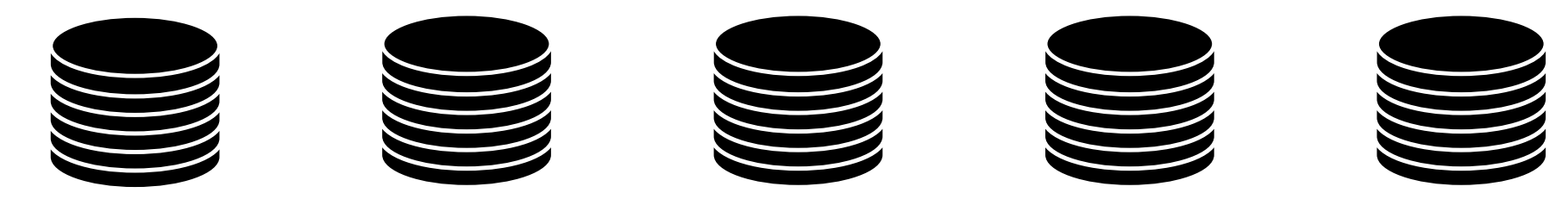
- *collations* contain pre-computed data (e.g. CNV frequencies, statistics) and information for all grouping entity instances and correspond to **filter values**
 - ▶ PMID:10027410, NCIT:C3222, pgx:cohort-TCGA, pgx:icdom-94703...
- *querybuffer* stores id values of all entities matched by a query and 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("5bab578d727983b2e00ch505")
]
```



variants analyses biosamples individuals

Entity collections



collations geolocs genespans publications qBuffer

Utility collections

github.com/progenetix/bycon/




Onboarding


Demonstrating Compliance

- Progenetix Beacon+ has served as implementation driver since 2016
- Beacon v2 as service with protocol-driven registries for federation
- GA4GH approved Beacon v2 in April 2022



Beacon v2 GA4GH Approval Registry

Beacons:    


 **European Genome-Phenome Archive (EGA)**

GA4GH Approval Beacon Test

This [Beacon](#) is based on the GA4GH Beacon [v2.0](#)

- Visit us
- Beacon API
- Contact us


BeaconMap	Matches the Spec
Bioinformatics analysis	Matches the Spec
Biological Sample	Matches the Spec
Cohort	Matches the Spec
Configuration	Matches the Spec
Dataset	Matches the Spec
EntryTypes	Matches the Spec
Genomic Variants	Matches the Spec
Individual	Matches the Spec
Info	Matches the Spec
Sequencing run	Matches the Spec

 **Theoretical Cytogenetics and Oncogenomics group at UZH and SIB**

Progenetix Cancer Genomics Beacon+ Beacon+ provides a forward looking implementation of the Beacon v2 API, with focus on structural genome variants and metadata based on the...

- Visit us
- Beacon UI
- Beacon API
- Contact us

BeaconMap	Matches the Spec
Bioinformatics analysis	Matches the Spec
Biological Sample	Matches the Spec
Cohort	Matches the Spec
Configuration	Matches the Spec
Dataset	Matches the Spec
EntryTypes	Matches the Spec
Genomic Variants	Matches the Spec
Individual	Matches the Spec
Info	Matches the Spec
Sequencing run	Matches the Spec


 **Centre Nacional Analisis Genomica (CNAG-CRG)**

Beacon @ RD-Connect

This [Beacon](#) is based on the GA4GH Beacon [v2.0](#)

- Visit us
- Beacon API
- Contact us

BeaconMap	Matches the Spec
Bioinformatics analysis	Matches the Spec
Biological Sample	Not Match the Spec
Cohort	Matches the Spec
Configuration	Matches the Spec
Dataset	Not Match the Spec
EntryTypes	Matches the Spec
Genomic Variants	Matches the Spec
Individual	Not Match the Spec
Info	Matches the Spec
Sequencing run	Matches the Spec

 **University of Leicester**

Cafe Variome Beacon v2

This [Beacon](#) is based on the GA4GH Beacon [v2.0](#)

- Beacon UI
- Beacon API
- Contact us

BeaconMap	Matches the Spec
Bioinformatics analysis	Matches the Spec
Biological Sample	Matches the Spec
Cohort	Matches the Spec
Configuration	Matches the Spec
Dataset	Matches the Spec
EntryTypes	Matches the Spec
Genomic Variants	Matches the Spec
Individual	Matches the Spec
Info	Matches the Spec
Sequencing run	Matches the Spec

✔ Matches the Spec
 ✘ Not Match the Spec
 ⚪ Not Implemented

Progenetix

Genomic resource utilizing Beacon v2 calls

- Progenetix uses Beacon v2 queries to drive its UI
- all individuals, biosamples, variants, analyses matched by a given query are stored by their object ids
- handovers for variant purposes (e.g. to retrieve all matched variants) are returned in the original response and asynchronously retrieved by the front end app

The screenshot displays the Progenetix web application interface. At the top, there is a search bar with the text "Edit Query". Below it, the assembly information is shown: "Assembly: GRCh38 Chro: 9 Start: 21500001-21975098 End: 21967753-22500000" and "Type: EFO:0030067 Filters: NCIT:C3058".

The main content area shows a "progenetix" search result. It includes statistics: "Matched Samples: 660", "Retrieved Samples: 660", "Variants: 279", and "Calls: 667". There are links for "UCSC region", "Variants in UCSC", and "Dataset Responses (JSON)". A "Visualization options" button is also present.

Below the statistics, there are tabs for "Results", "Biosamples", "Biosamples Map", "Variants", and "Annotated Variants". The "Results" tab is active, showing a table of matched subset codes, subset samples, matched samples, and subset match frequencies. The table has the following data:

Matched Subset Codes	Subset Samples	Matched Samples	Subset Match Frequencies
UBERON:0003031	4	1	0.250
pgx:icd...			
pgx:icd...			
NCIT:C3...			
UBERON:0018525	14	2	0.143
pgx:icdot-C71.1	14	2	0.143
UBERON...			
pgx:icdot-94423	84	4	0.048
NCIT:C3796	84	4	0.048
UBERON:0001869	1714	14	0.008
pgx:icdot-C71.0	1714	14	0.008

Below the table, there are download links for "Download 1-660", "Download Sample Data (JSON) 1-660", and "Download Sample Variants (JSON) 1-660".


Three API call overlays are shown:

- Yellow overlay:** `/beacon/biosamples/?requestedGranularity=record&limit=1000&skip=0&assemblyId=GRCh38&referenceName=9&variantType=EFO:0030067&start=21500000,21975098&end=21967753,22500000&filters=NCIT:C3058`
- Cyan overlay:** `/beacon/biosamples/?skip=0&limit=1000&accessid=fbffda57-0f41-4d6a-99fc-41d4cfdea9f6`
- Light blue overlay:** `/beacon/genomicVariations/?accessid=e2dadd91-9326-46de-97e4-6b88413b6bfe`
- Pink overlay:** `/services/sampleplots/?accessid=fbffda57-0f41-4d6a-99fc-41d4cfdea9f6&plotType=histogram&-size_plotimage_w_px=645`

The right side of the screenshot shows a browser's developer console with a table of network requests. The table has columns for Name, Do..., T, Transf..., T..., and time intervals (10.00s, 20.00s, 30.00s). The requests listed are: biosamples, biosamples, genomicVariations, genomicVariations, samplePlots.cgi, and collations.

Beacon v2 Conformity and Extensions in Progenetix

Putting the + into Beacon ...

- support & use of standard Beacon v2 PUT & GET variant queries, filters and meta parameters
 - ➔ variant parameters, geneld, lengths, EFO & VCF CNV types, pagination
 - ➔ widespread, self-scoping filter use for bio-, technical- and and id parameters with switch for descending terms use (globally or per term if using POST)
- extensive use of handovers
 - ➔ asynchronous delivery of e.g. variant and sample data, data plots
- + optional use of OR logic for filter combinations (global)
- + extension of query parameters
 - ➔ geographic queries incl. \$geonear and use of GeoJSON in schemas
-  no implementation of authentication on this open dataset

Progenetix provides a number of additional services and output formats which are initiated over the /services path or provided as request parameters and are not considered Beacon extensions (though they follow the syntax where possible).



progenetix / byconaut

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

bycon.progenetix.org
github.com/progenetix/bycon/

byconaut Public

main 2 branches

mbaudis get_plot_parameters

- bin
- docs
- exports
- imports
- local
- rsrc
- services
- tmp
- .gitignore
- LICENSE
- README.md
- __init__.py
- install.py
- install.yaml
- mkdocs.yaml

progenetix / beaconplus-web

Code Pull requests Actions Projects Security Insights Settings

beaconplus-web Public

forked from progenetix/progenetix-web

main 1 branch 0 tags

This branch is 44 commits ahead, 24 commits behind progenetix:main.

mbaudis code cleaning, no feature changes

- .github/workflows cleanup
- docs still first implementation clean-up
- extra documentation
- public graphic refinement
- src code cleaning, no feature changes
- .babelrc Simplify query generation and add
- .env.development first working version
- .env.local first working version
- .env.production env
- .env.staging env
- .eslintrc.json BioSubsetsPage perf optimisations

bycon Public

progenetix / bycon

Code Issues Pull requests 1 Actions Projects Wiki Security 3 Insights Settings

bycon Public

main 4 branches 25 tags

mbaudis 1.3.6 852 commits

.github/workflows	Create mk-bycon-docs.yaml	8 months ago
bycon	1.3.6	3 days ago
docs	1.3.6	3 days ago
local	1.3.5 preparation	2 weeks ago
.gitignore	Update .gitignore	3 months ago
LICENSE	Create LICENSE	3 years ago
MANIFEST.in	major library & install disentanglement	9 months ago
README.md	#### 2023-07-23 (v1.0.68)	4 months ago
install.py	1.3.6	3 days ago
install.yaml	v1.0.57	5 months ago
mkdocs.yaml	1.1.6	3 months ago
requirements.txt	1.3.6	3 days ago
setup.cfg	...	10 months ago
setup.py	1.3.6	3 days ago
updev.sh	1.3.6	3 days ago

beaconplus.progenetix.org
[.../progenetix/beaconplus-web/](https://github.com/progenetix/beaconplus-web/)

bycon.progenetix.org
github.com/progenetix/bycon/

About

Bycon - A Python Based Beacon API (beacon-project.io) implementation leveraging the Progenetix (progenetix.org) data model

- Readme
- CC0-1.0 license
- Activity
- 5 stars
- 4 watching
- 6 forks

Report repository

Releases

25 tags

[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Beacon v2

Federated Genomics

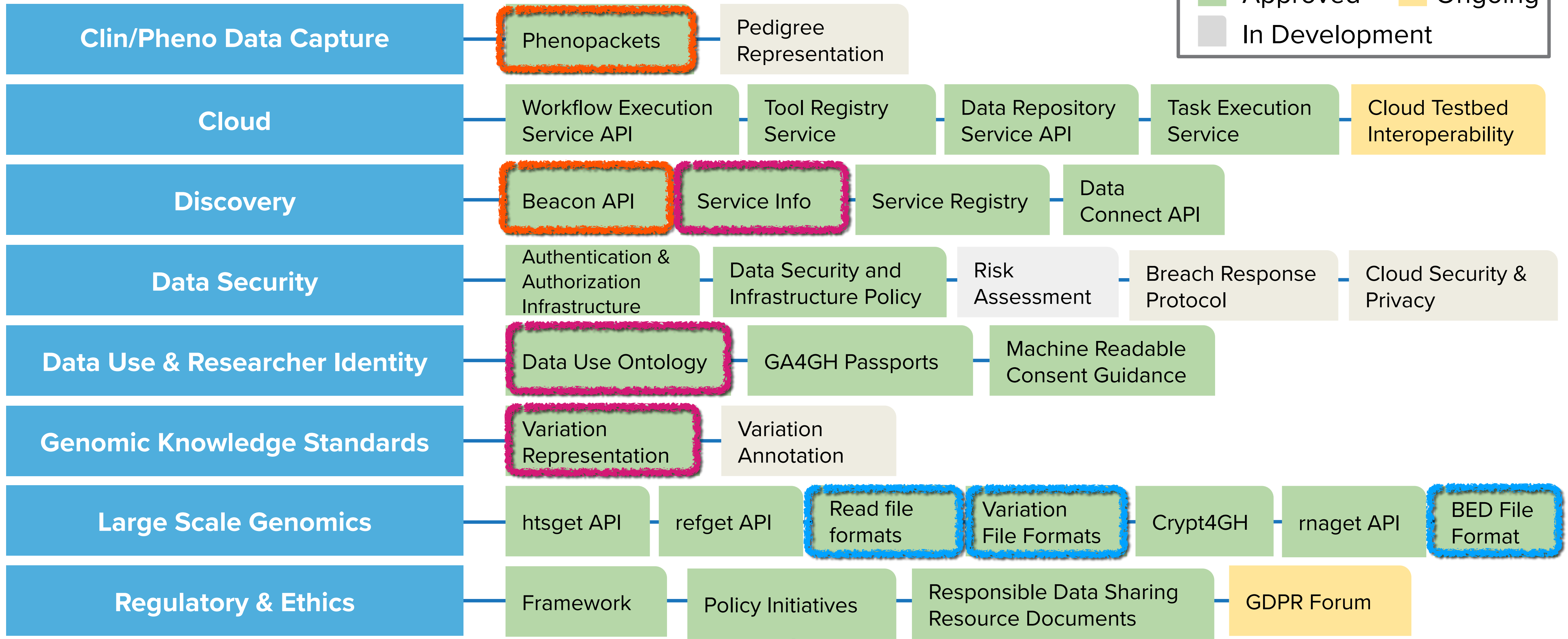


GA4GH 2020-2022 Strategic Roadmap



Global Alliance
for Genomics & Health

■ Approved ■ Ongoing
■ In Development



Ethics Review Recognition, Accountability, Consent, Privacy & Security Technical Standards & IP, Return of Results



Beacon



A **Beacon** answers a query for a specific genome variant against individual or aggregate genome collections

YES | **NO** | \0



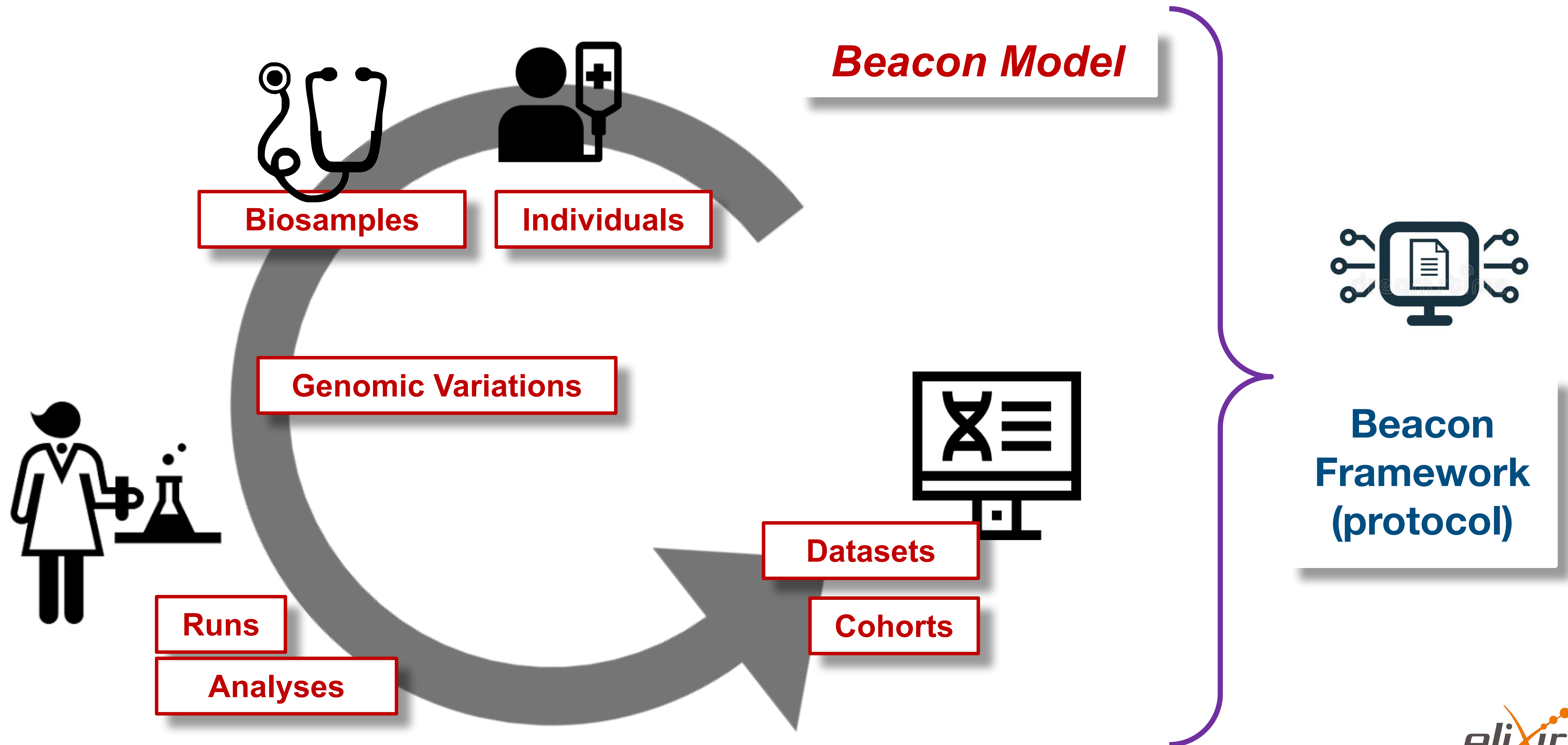
Have you seen this variant?
It came up in my patient
and we don't know if this is
a common SNP or worth
following up.

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**.

Beacon v2

docs.genomebeacons.org





Beacon v1 Development

Beacon v2 Development

Related ...

2014 GA4GH founding event; Jim Ostell proposes Beacon concept with "more features... version 2"

2015

- beacon-network.org aggregator created by DNASTack
- Beacon v0.3 release

2016

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

2017

- OpenAPI implementation
- integrating CNV parameters (e.g. "startMin, statMax")

2018

- 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

2022

- Beacon⁺ concept implemented on 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 ...
- discussions w/ clinical stakeholders

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

- further changes esp. in default model, aligning with Phenopackets and VRS

- unified beacon-v2 code & docs repository
- **Beacon v2 approved at Apr GA4GH Connect**

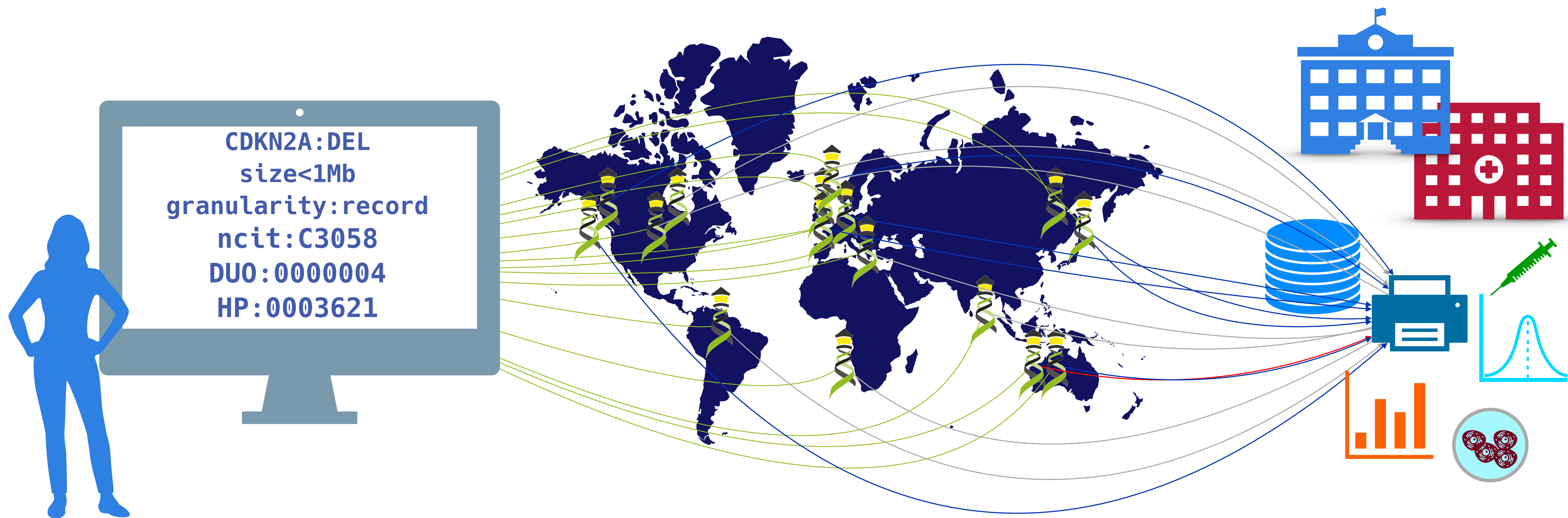
- ELIXIR starts Beacon project support

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

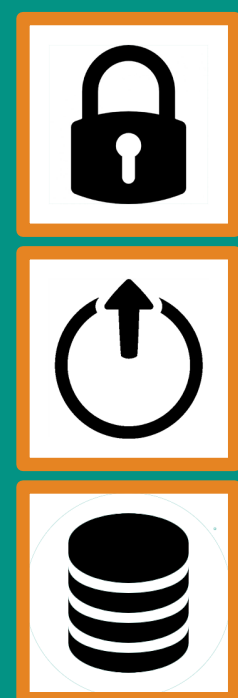
- new Beacon website (March)

- Beacon publication at Nature Biotechnology

- docs.genomebeacons.org



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 Queries

Implementation of Current Options

- (so far) the Beacon model does not define explicit query types
- disambiguation of parameters is left to implementers
- implicit query types:
 - ➔ allele/sequence query
 - ➔ range query, w/ or w/o additional parameters
 - ➔ bracket query (e.g. sized CNVs)
 - ➔ aminoacid, HGVS, gene

beaconplus.progenetix.org

Beacon Query Types

Sequence / Allele

CNV (Bracket)

Genomic Range

Aminoacid

Gene ID

HGVS

Sam

Dataset

Test Database - exemplez x

Chromosome *i*

Select...

Variant Type *i*

Select...

Start or Position *i*

19000001-21975098

Reference Base(s) *i*

N

Alternate Base(s)

A

Select Filters *i*

Select...

Query Database

Form Utilities

Gene Spans

Cytoband(s)

Query Examples

CNV Example

SNV Example

Range Example

Gene Match

Aminoacid Example

Identifier - HeLa

Beacon Queries

Implementation of Current Options

- (so far) the Beacon model does not define explicit query types
- disambiguation of parameters is left to implementers
- implicit query types:
 - ➔ allele/sequence query
 - ➔ range query, w/ or w/o additional parameters
 - ➔ bracket query (e.g. sized CNVs)
 - ➔ aminoacid, HGVS, gene

Beacon Query Types

Sequence / Allele | CNV (Bracket) | Genomic Range | Aminoacid | Gene ID | HGVS | Sarr

Dataset
Progenetix x Test Database - examplez x

Chromosome 17 (NC_000017.11) | **Variant Type** SO:0001059 (any sequence alteration - S... |

Start or Position 7577121

Reference Base(s) G | **Alternate Base(s)** A

Select Filters Select... |

Chromosome 17
7577121

Query Database

Form Utilities Gene Spans Cytoband(s)

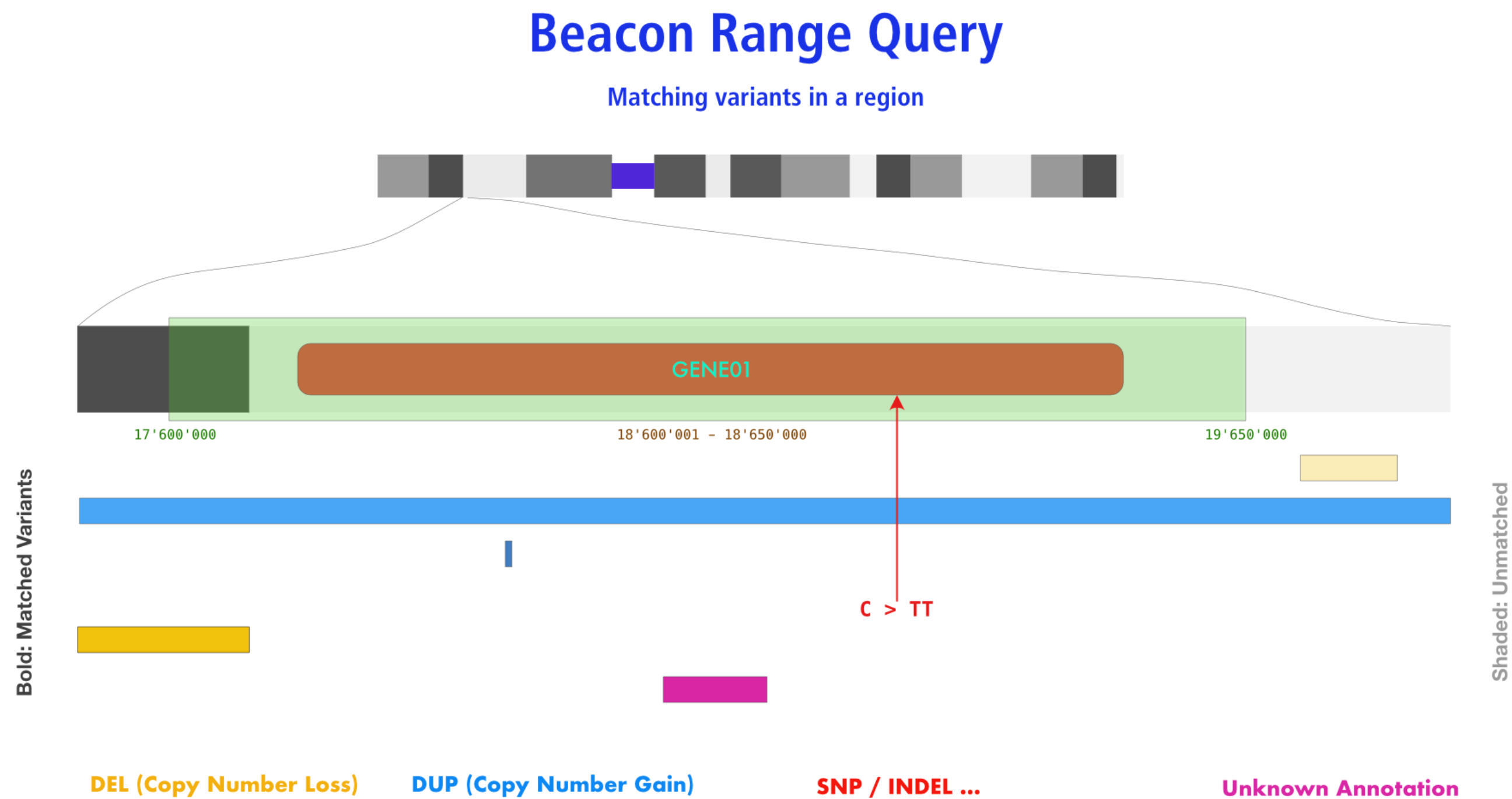
Query Examples CNV Example SNV Example Range Example Gene Match
Aminoacid Example Identifier - HeLa

This example shows a core Beacon query, against a specific mutation in the EIF4A1 gene, in the DIPG childhood brain tumor dataset.

Beacon Queries

Range ("anything goes") Request

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



beaconplus.progenetix.org

Beacon Query Types

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

Dataset

Test Database - exemplez x

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

Select...

Chromosome 17

7572826

7579005

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.

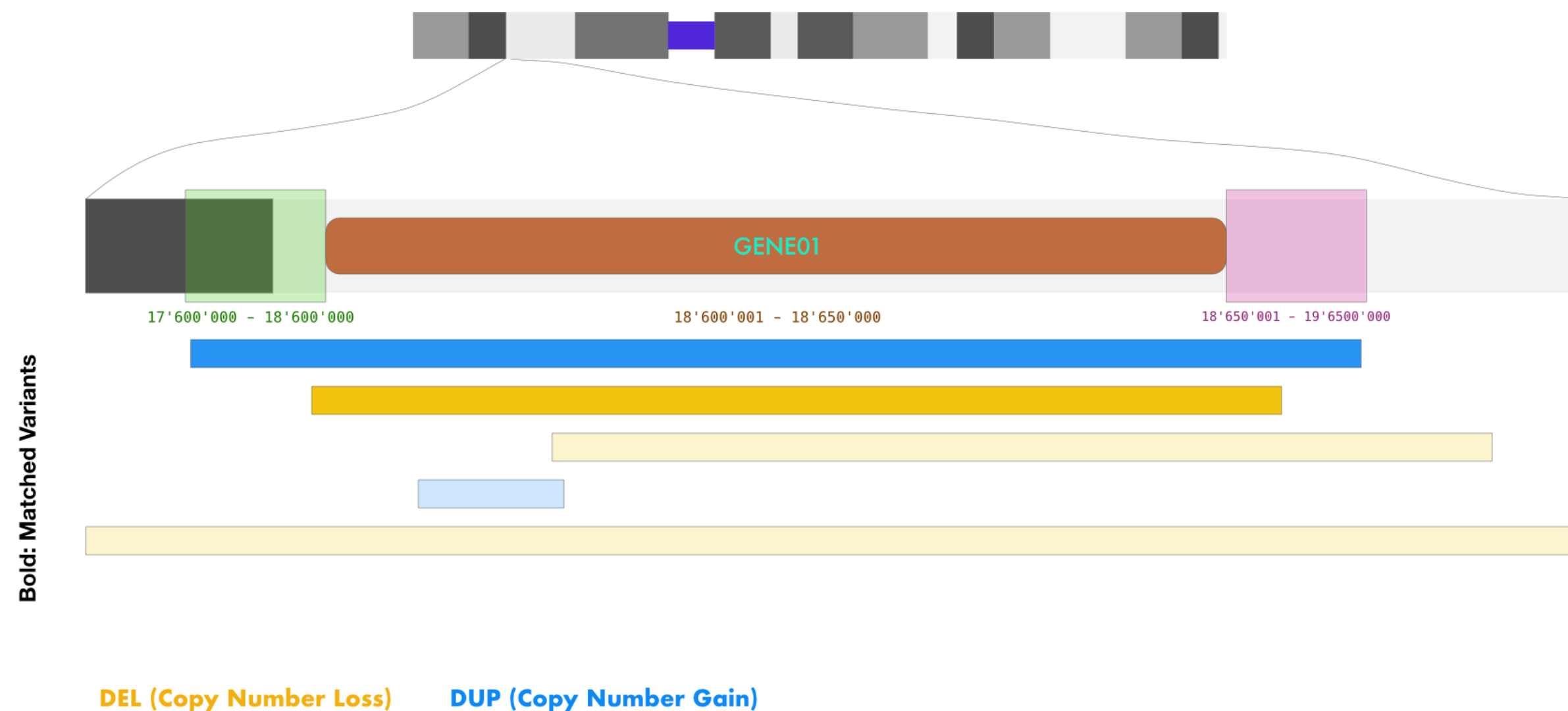
Beacon Queries

Bracket ("CNV") Query

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

Beacon Bracket Query

Example for complete regional match



Shaded: Unmatched

Beacon Query Types

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

Dataset: Test Database - examplez x | v

Chromosome: 9 (NC_000009.12) | v Variant Type: EFO:0030067 (copy number deletion) | v

Start or Position: 21000001-21975098 End (Range or Structural Var.): 21967753-23000000

Select Filters: NCIT:C3058: Glioblastoma (100) x | v

Chromosome 9

Query Database

Form Utilities: Gene Spans, Cytoband(s)

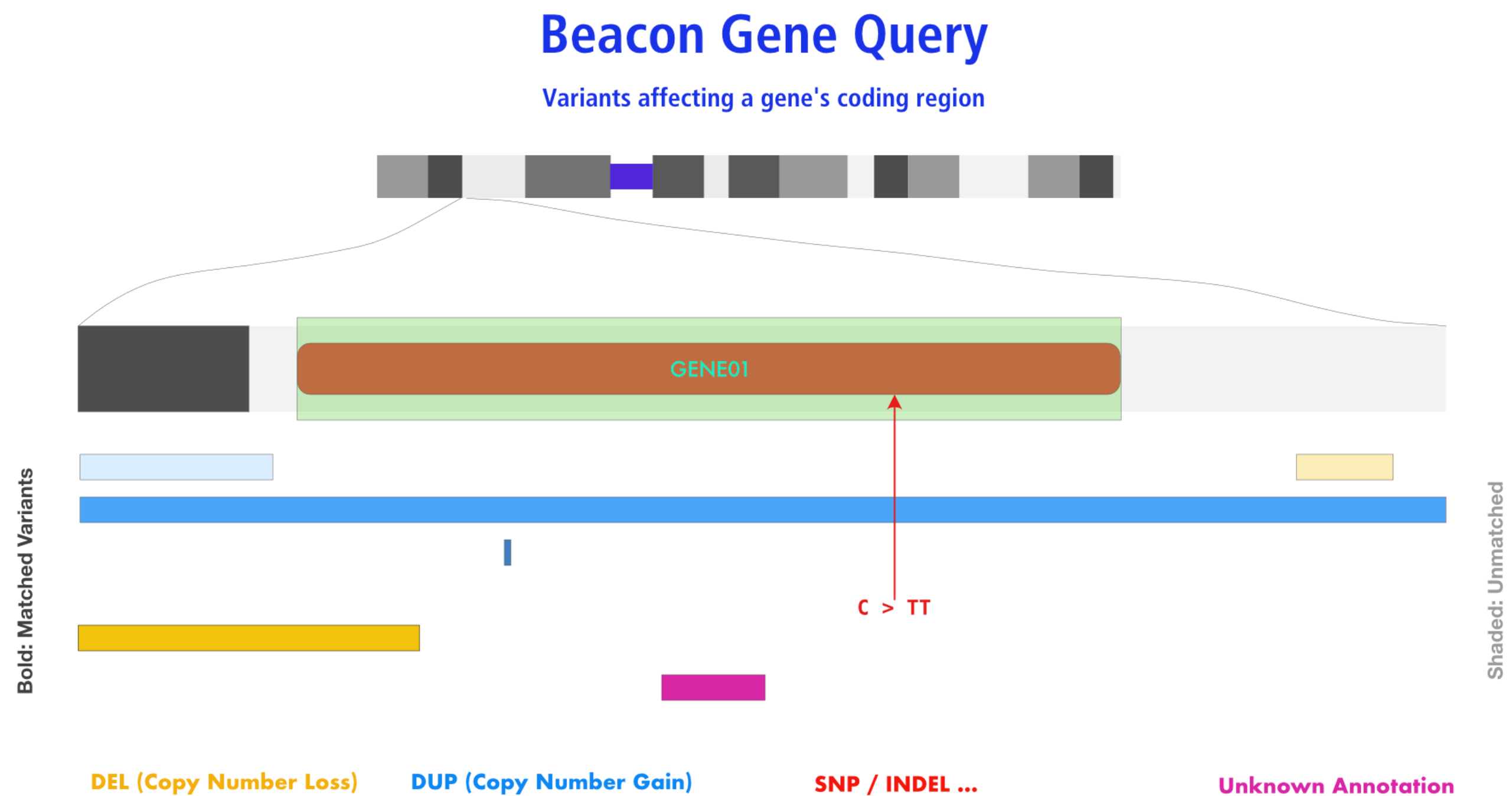
Query Examples: CNV Example, SNV Example, Range Example, Gene Match, Aminoacid Example, Identifier - HeLa

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. <= ~2Mbp in size). The query is against the examplez collection and can be modified e.g. through changing the position parameters or data source.

Beacon Queries

Gene Request

- defined through a (HUGO) gene symbol
- assuming hit on the gene's CDR but YMMV



beaconplus.progenetix.org

Beacon Query Types

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

Dataset
Cancer Cell Lines Collection x | v

Gene Symbol i
CDK2 (12:55966830-55972789) x | v

Variant Type i
Select... | v

Min Variant Length i | **Max Variant Length** i | **Alternate Base(s)**
 | | A

Select Filters i
Select... | v

Query Database

Form Utilities | Gene Spans | Cytoband(s)

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

Beacons in v2 can support the discovery of variants with overlap with the genomic location of a gene, indicated by its symbol (e.g. **CDK2**). Additional parameters can *optionally* be used to make matches more specific:

- **variantMinLength** and **variantMaxLength** to limit matched CNV sizes
- **genomicAlleleShortForm** (e.g. **V600E** with **BRAF**)
- **variantType** and **alternateBases** to specify variants

Beacon Queries

Missing or ill defined options

- **translocations** are in principle possible (start bracket with "referenceName" and end bracket with "mateName") but not yet documented / battle tested
- **functional elements?**
- exon hits beyond specifying individual ones by sequence
- tandem dups ...
- genomic **double hits**

Beacon Query Types

Sequence / Allele

CNV (Bracket)

Genomic Range

Aminoacid

Gene ID

HGVS

Sarr

Dataset

Test Database - exemplez x

Chromosome *i*

Select...

Variant Type *i*

Select...

Start or Position *i*

19000001-21975098

Reference Base(s) *i*

N

Alternate Base(s)

A

Select Filters *i*

Select...

Query Database

Form Utilities

Gene Spans

Cytoband(s)

Query Examples

CNV Example

SNV Example

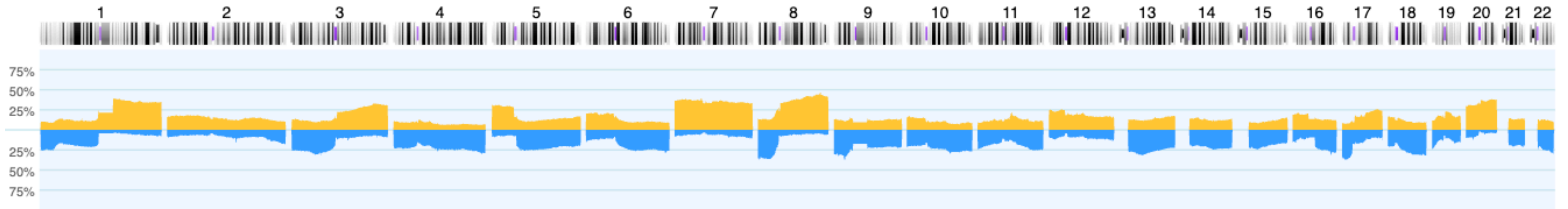
Range Example

Gene Match

Aminoacid Example

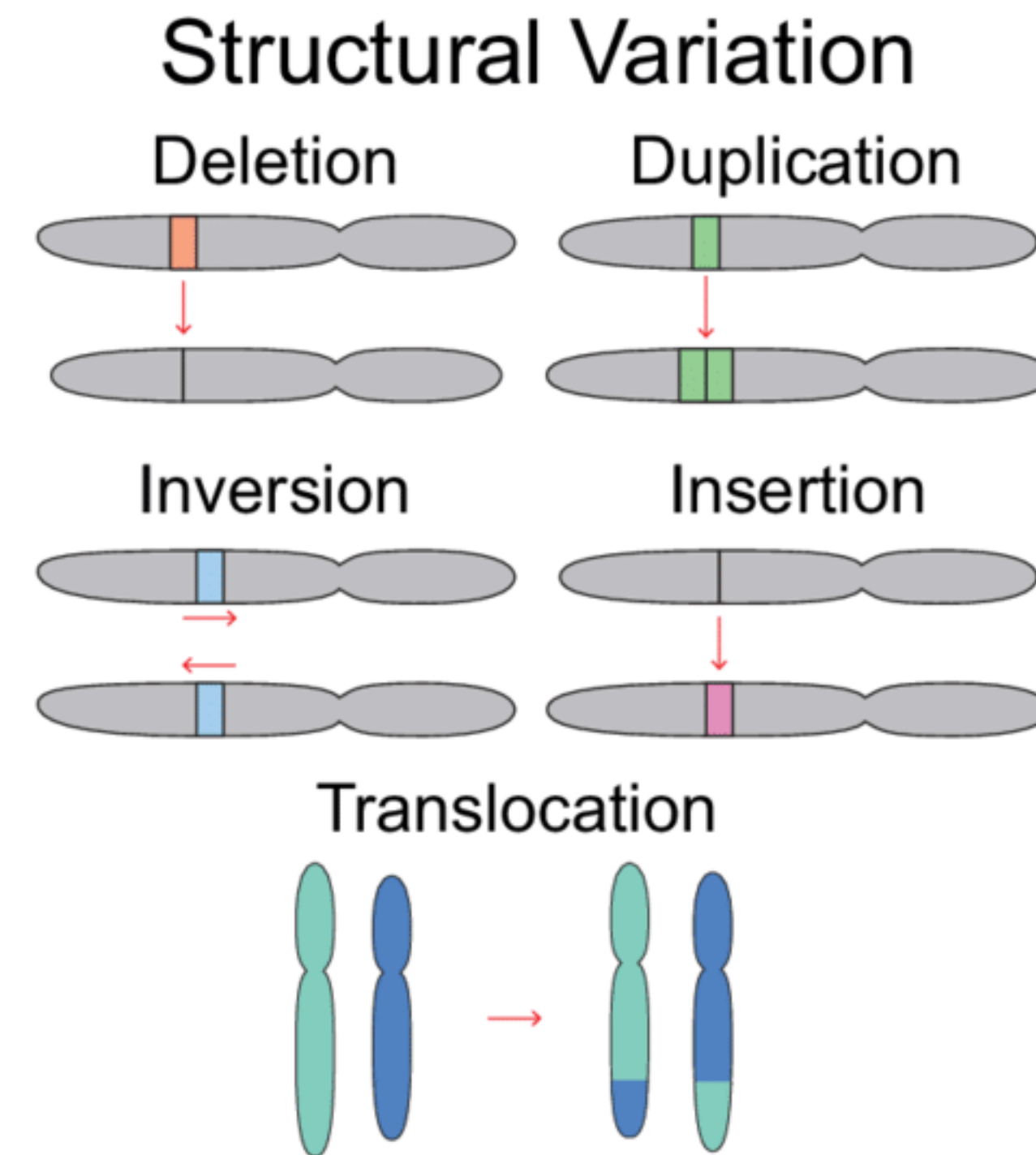
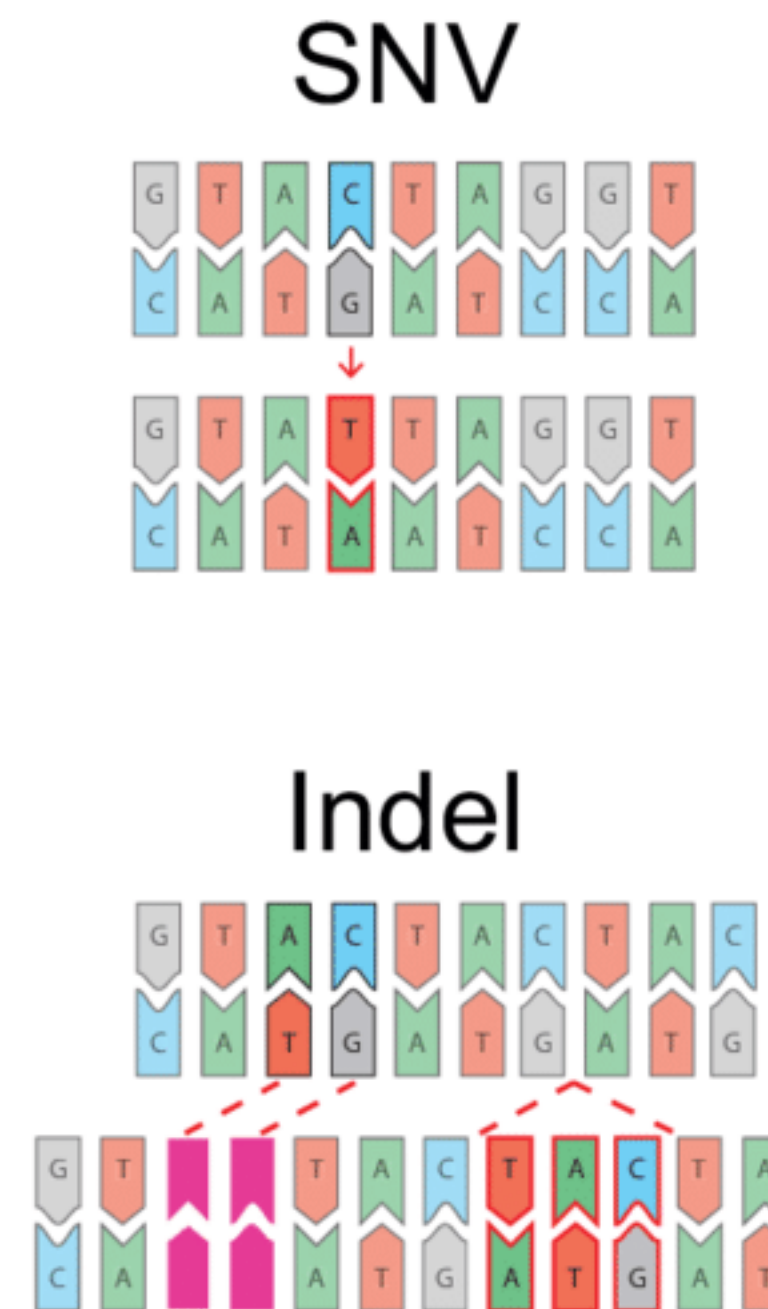
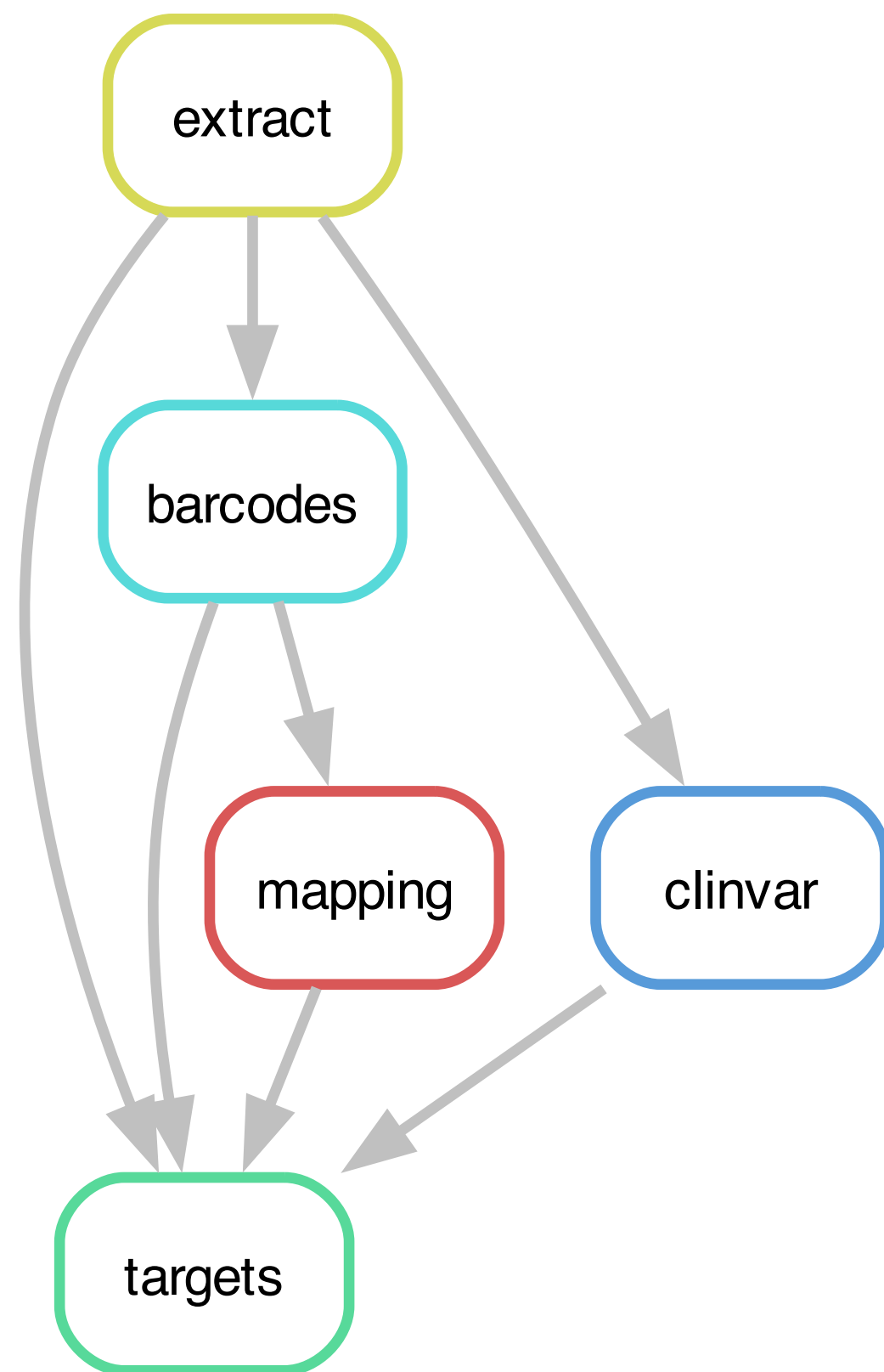
Identifier - HeLa

TCGA Cancer samples (pgx:cohort-TCGAcancers)



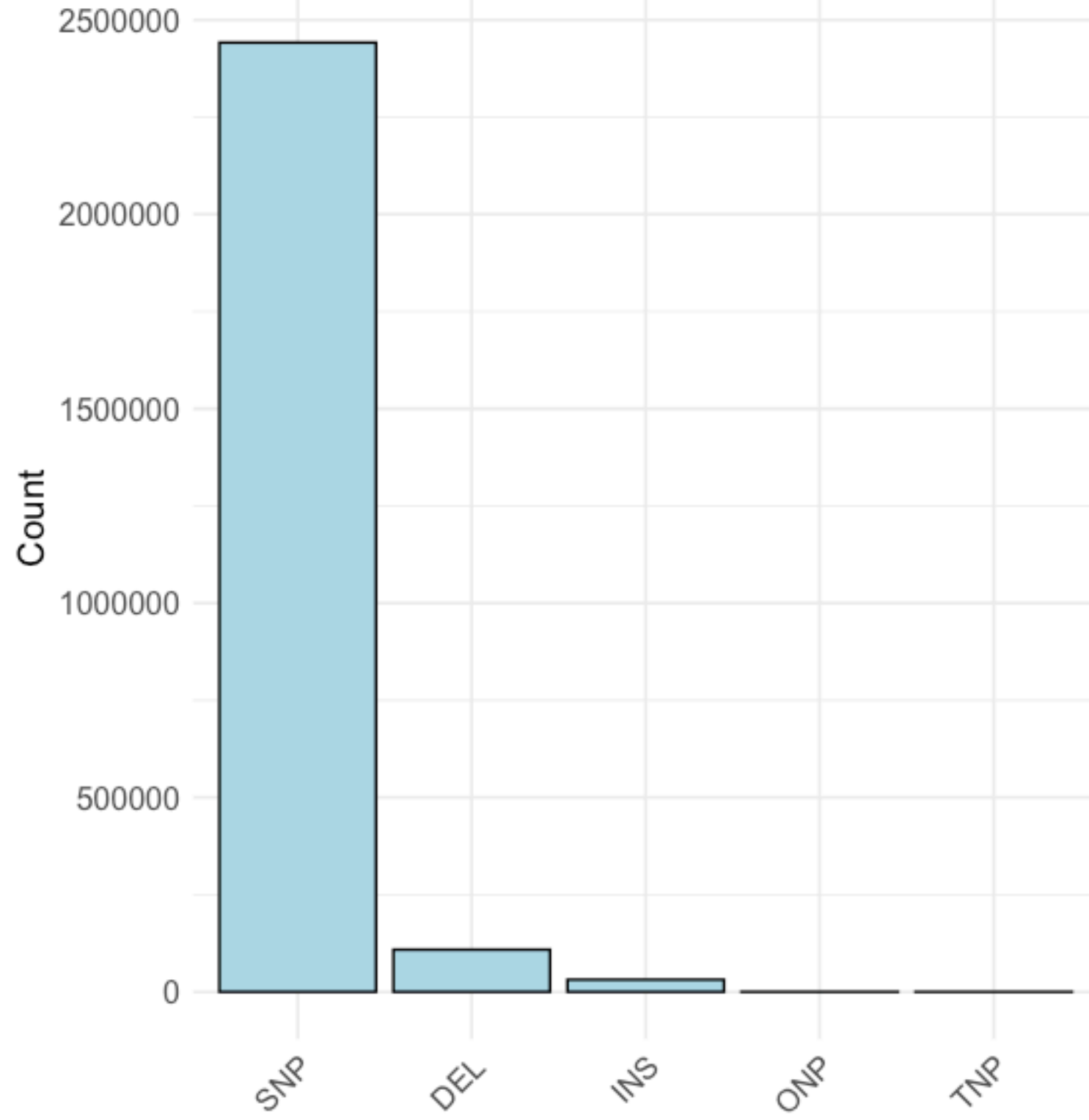
© CC-BY 2001 - 2023 progenetix.org

- **Integrating SNVs:**
 - Renewed interest due to technological advantages
 - Allow compound variant queries

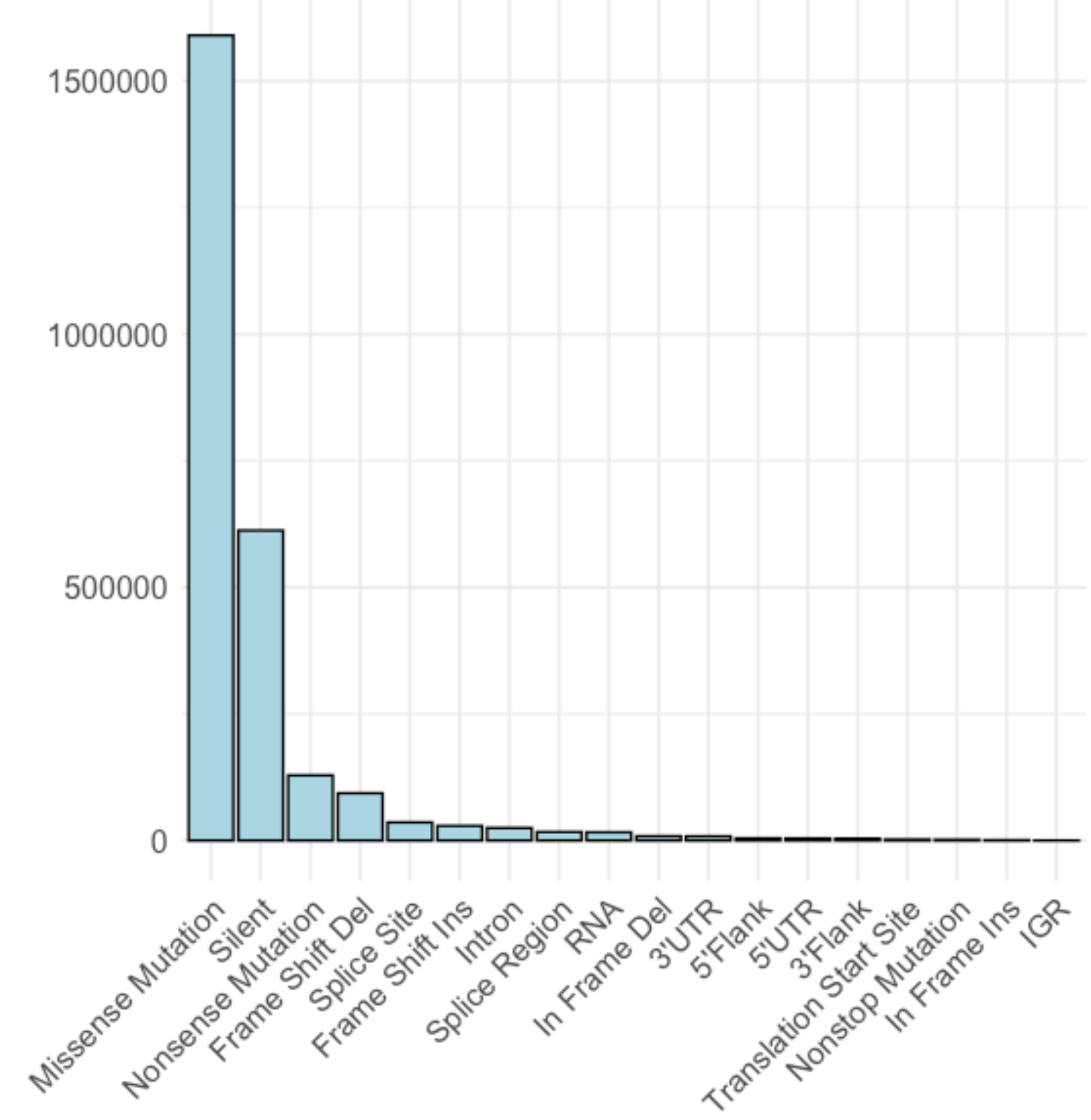


Nesta, Alex & Tafur, Denisse & Beck, Christine. (2020). Hotspots for Structural Variations in the Human Genome. Trends in Genetics

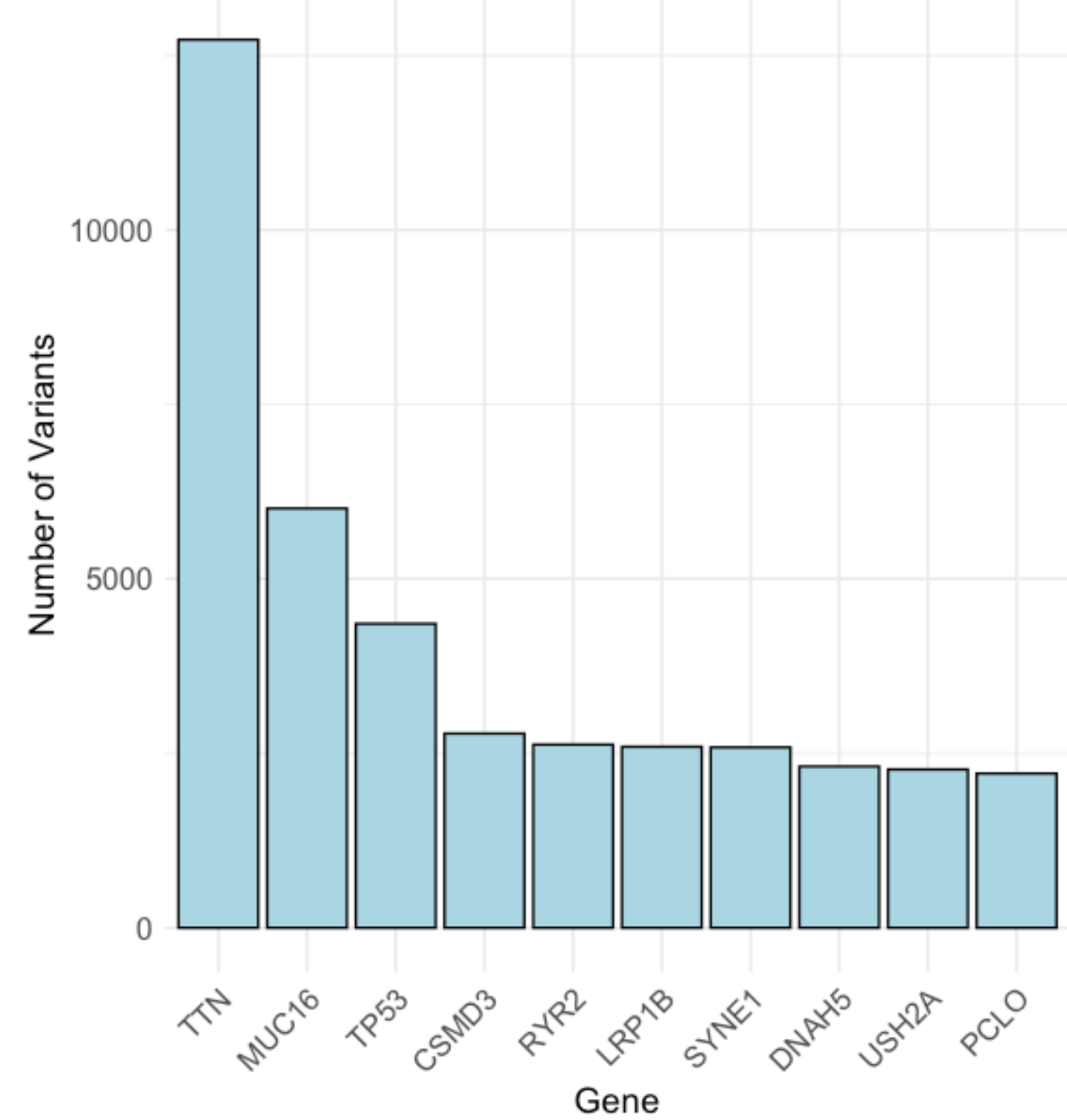
SNV types



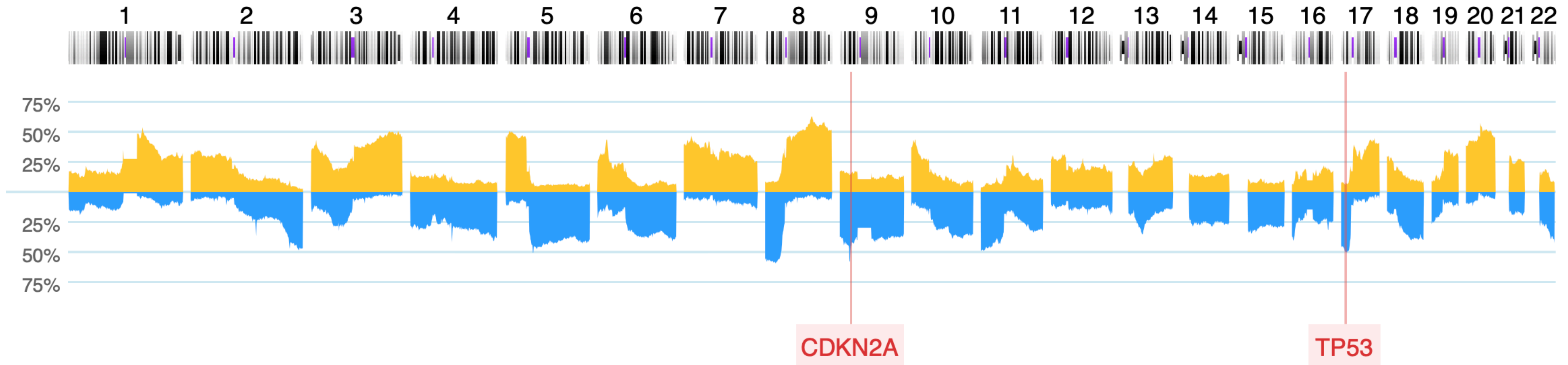
Variant Classes



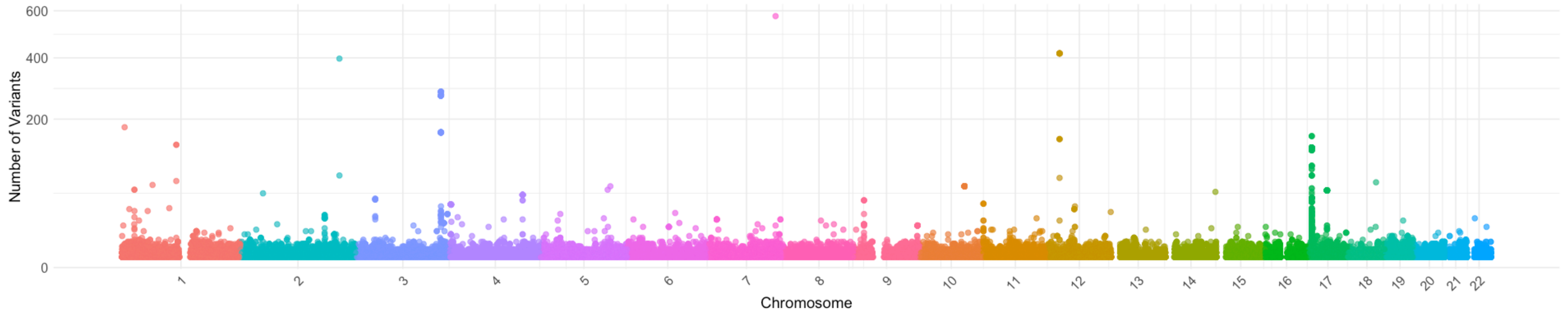
Top 10 Most Mutated Genes



TCGA BLCA project (pgx:TCGA.BLCA)



TCGA BLCA project
Variant Distribution on Genome Coordinates



Beacon Queries

Missing or ill defined options

- **translocations** are in principle possible (start bracket with "referenceName" and end bracket with "mateName") but not yet documented / battle tested
- **functional elements?**
- exon hits beyond specifying individual ones by sequence
- tandem dups ...
- genomic **double hits**

➔ **Beacon & hCNV Scout Team**

Beacon Query Types

Sequence / Allele

CNV (Bracket)

Genomic Range

Aminoacid

Gene ID

HGVS

Sarr

Dataset

Test Database - exemplez x

Chromosome *i*

Select...

Variant Type *i*

Select...

Start or Position *i*

19000001-21975098

Reference Base(s) *i*

N

Alternate Base(s)

A

Select Filters *i*

Select...

Query Database

Form Utilities

⚙ Gene Spans

⚙ Cytoband(s)

Query Examples

CNV Example

SNV Example

Range Example

Gene Match

Aminoacid Example

Identifier - HeLa

