



hCNV Community

Standards and protocols for CNV discovery and data exchange



*Andrew Stubbs
Antonio Rausell
Krzysztof Poterlowicz
Michael Baudis*



www.elixir-europe.org

hCNV products

- Resource (<https://refcnv.org/>) : Progenetix cancer CNV reference resource and upcoming refCNV for reference copy number data accessible via Beacon API - Michael
- Resource (htsget2Galaxy) : FAIR access to GA4GH databases - Andrew
- Resource (CNV 2 Galaxy Beacon) : integrate rare disease workflow with external data (e.g hCNV from refcnv) and Cancer data in Genomic Beacons. – Khaled / Krzys
- Resource: - bioinformatics and ML tools for CNV assessment, including a couple of slides opening to non-coding SNVs at the very end (which can be of interest for the rare-disease group even if is beyond the CNV scope). - Antonio
- Standards: Beacon, VRS and VCF support for CNV representation and querying



progenetix.org

Cancer Genomics Reference Resource

- **open** resource for oncogenomic profiles
- over **150'000** cancer **CNV** profiles
- more than **900** diagnostic types
- runs on a **Beacon API**
- inclusion of reference datasets (e.g. TCGA)
- support for SNV data (TCGA, cell lines...)
- 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



CNV Profiles by Cancer Type

NCIT Neoplasia Codes
ICD-O Morphologies
ICD-O Organ Sites
TNM & Grade

Search Samples

Data Cohorts

arrayMap
TCGA Cancer Samples
cBioPortal Studies

Cancer Cell Lines^o

Publication DB

Genome Profiling
Progenetix Use

Services

NCIt Mappings
UBERON Mappings

Upload & Plot

OpenAPI Paths and Examples

Beacon⁺

Documentation

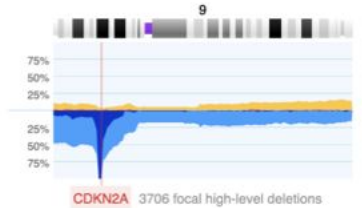
News
Downloads & Use Cases
Services & API

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* of currently **156871** samples from **912** different cancer types (NCIt neoplasm classification)

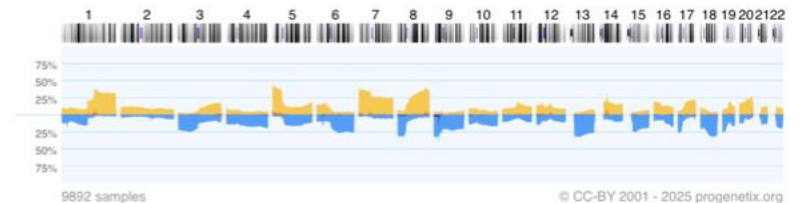
Local CNV Frequencies^o

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^o

Frequency profiles of regional genomic gains and losses for all categories (diagnostic entity, publication, cohort ...) can be accessed through the respective Cancer Types pages with visualization and sample retrieval options. Below is a typical example of the aggregated CNV data in 9087 samples in Lung Non-Small Cell Carcinoma with the frequency of regional **copy number gains (high level)** and **losses (high level)** displayed for the 22 autosomes.



[Download SVG](#) | [Go to NCIT:C2926](#) | [Download CNV Frequencies](#)

Cancer Genomics Publications^o

Through the [[Publications](#)] page Progenetix provides 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.

- 
- Universität
Zürich ^{UZH}

Sample selection follows a hierarchical system in which samples matching the child terms of a selected class are included in the response.

Filter subsets e.g. by prefix Hierarchy Depth: 4 levels

No Selection

Glioblastoma (NCIT:C3058)

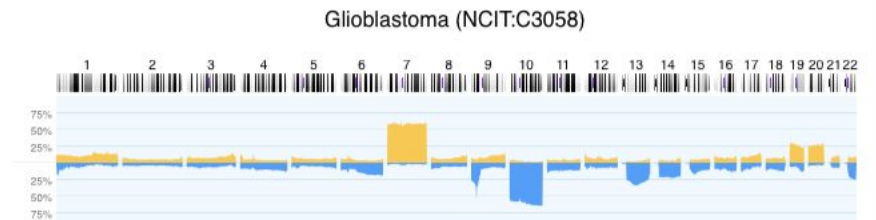
Sample Counts

- 4370 samples
- 4286 direct *NCIT:C3058* code matches
- 4384 CNV analyses

Search Samples

Select *NCIT:C3058* samples in the [Search Form](#)

Raw Data (click to show/hide)



© CC-BY 2001 - 2023 progenetix.org

[Download SVG](#) | [Go to NCIT:C3058](#) | [Download CNV Frequencies](#)

- NCIT:C4822: Malignant Glioma (5598 samples, 5418 CNV profiles)
- NCIT:C6770: Ependymal Tumor (627 samples, 627 CNV profiles)
- NCIT:C6958: Astrocytic Tumor (5882 samples, 5896 CNV profiles)
- NCIT:C6960: Oligodendroglial Tumor (703 samples, 703 CNV profiles)
- NCIT:C8501: Brain Stem Glioma (2 samples, 2 CNV profiles)
- NCIT:C3716: Primitive Neuroectodermal T... (2213 samples, 2214 CNV profiles)
- NCIT:C4747: Glioneuronal and Neuronal Tumors (89 samples, 89 CNV profiles)
- NCIT:C6965: Pineal Parenchymal Cell Neoplasm (51 samples, 51 CNV profiles)

Cancer Genomics Reference Resource

- **open** resource for oncogenomic profiles
- over **150'000 cancer CNV profiles**
- more than **900 diagnostic types**
- runs on a **Beacon API**
- inclusion of reference datasets (e.g. TCGA)
- support for SNV data (TCGA, cell lines...)
- 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

Edit Query

Assembly: GRCh38 Chro: refseq:NC_000009.12 Start: 21500001-21975098
End: 21967753-22500000 Type: EFO:0030067 Filters: NCIT:C3058

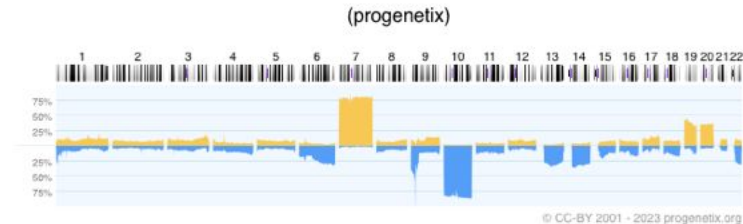
progenetix

Matched Samples: 657
Retrieved Samples:
Variants: 276
Calls: 659

[UCSC region](#)
[Variants in UCSC](#)
[Dataset Responses \(JSON\)](#)

Visualization options

Results Biosamples Biosamples Map Variants



[Reload histogram in new window](#)

Matched Subset Codes 1	Subset Samples 1	Matched Samples 1	Subset Match Frequencies 1
pgx:icdot-C71.4	4	1	0.250
pgx:icdom-94403	4286	653	0.152
NCIT:C3058	4370	653	0.149
pgx:icdot-C71.1	14	2	0.143
pgx:icdot-C71.9	7204	640	0.089
NCIT:C3796	84	4	0.048
pgx:icdom-94423	84	4	0.048
pgx:icdot-C71.0	1714	14	0.008

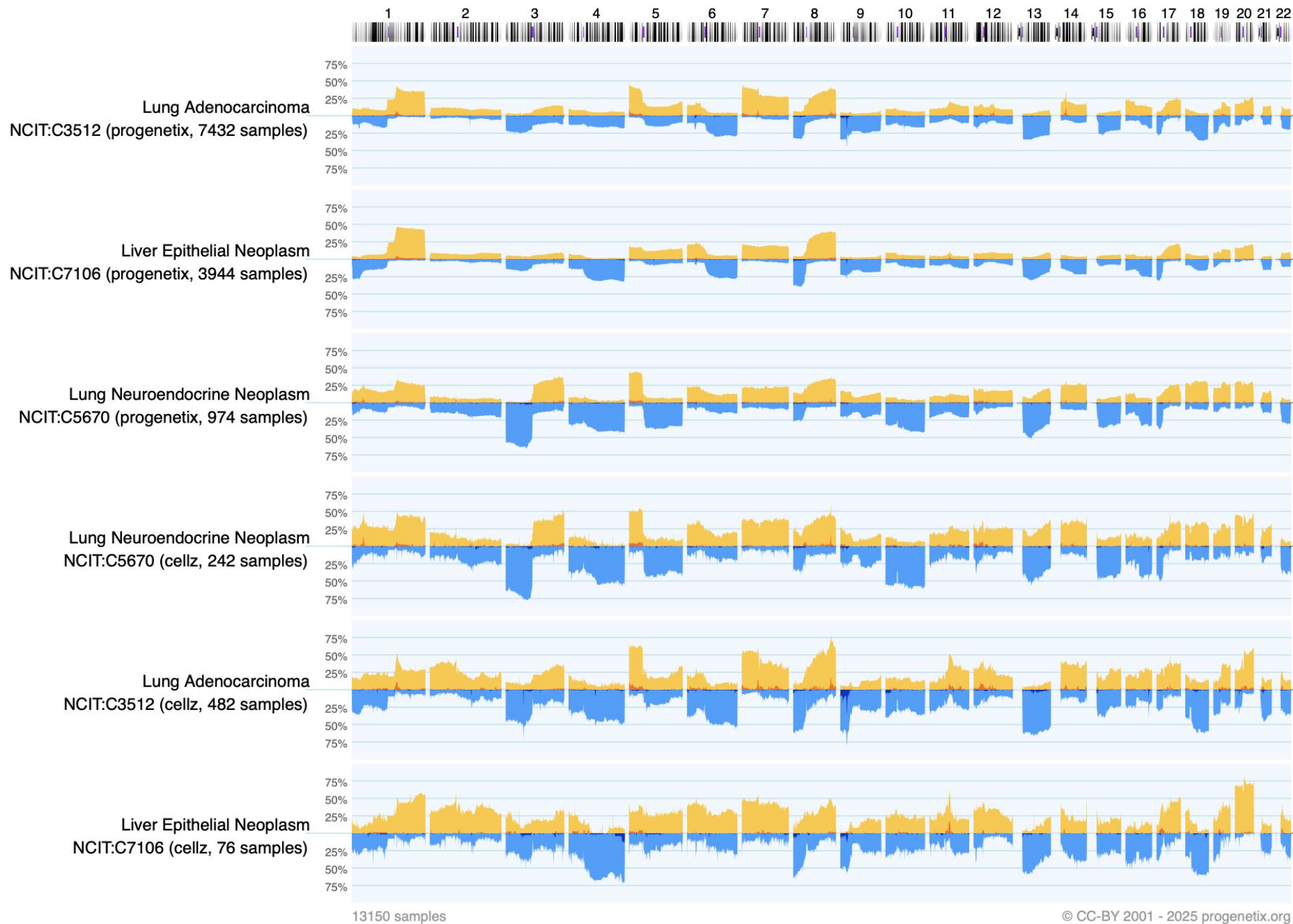
Download Sample Data (TSV)

1-657 [Download](#)

Download Sample Data (JSON)

1-657 [Download](#)

- 
- Universität
Zürich ^{UZH}



refCNV

Germline CNVs...

- germline CNVs vary widely by genomic background but importantly also by technical assessment
 - sequencing & array
 - bioinformatics workflows
- frequency based information can be misleading
- starting from a benchmarking project (A. Stubbs et al.) we started to do a "non-judgemental" collection w/ technical annotations
- Beacon API, obviously ...



CNV Profiles by
Platform

CNV Profiles by
Analysis Pipeline

Search Samples

Beacon⁺

Documentation

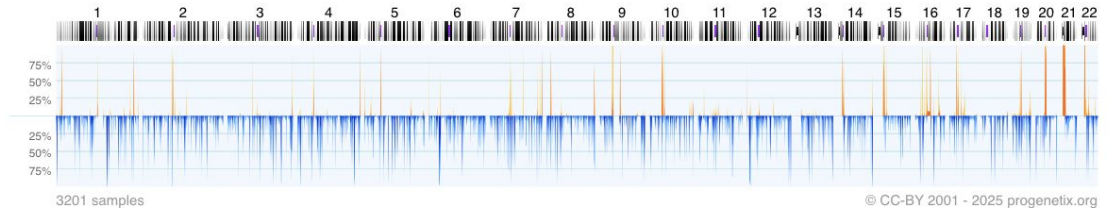
Baudisgroup @ UZH

Genomic Copy Number Variation (CNV) data from reference samples

i Under Construction

This site is currently under construction, with contributions by groups from the University of Zurich and Erasmus MC. Neither data content nor representation have been finalized. PLEASE DO NOT USE FOR ANY RESEARCH OR REFERENCE PURPOSES!

Frequency profiles of regional genomic gains and losses for all categories (diagnostic entity, publication, cohort ...) can be accessed through the respective Cancer Types pages with visualization and sample retrieval options. Below is a typical example of the aggregated CNV data in 3201 samples of the 1000 Genomes Dragen CNV analysis set. The frequency of regional **copy number gains (high level)** and **losses (high level)** displayed for the 22 autosomes as occurrence of any of these CNVs in the 1Mb binned intervals.



[Download SVG](#) | [Go to DRAGEN-CNV](#) | [Download CNV Frequencies](#)

The repository contains CNV tracks for many of the 1000 Genomes samples, analyzed by different platforms or data pipelines and therefore allows to compare private analysis data to results from these different call sets, to avoid interpretation biases from using reference data with a different analysis profile from the one used in your study. The plot below shows analysis specific CNV tracks for chromosome 13 in the HG01572 sample from the 1000 Genomes set, for several calling pipelines.



Please be aware that the small size of most CNVs is not correctly represented at this zoom level (overplotting due to limited resolution).

refCNV

Germline CNVs...

- germline CNVs vary widely by genomic background but importantly also by technical assessment
 - sequencing & array
 - bioinformatics workflows
- frequency based information can be misleading
- starting from a benchmarking project (A. Stubbs et al.) we started to do a "non-judgemental" collection w/ technical annotations
- Beacon API, obviously ...

Analysis Pipelines

This page represents samples with analyses from different pipelines.

Analysis Pipelines

Filter subsets e.g. by prefix

Hierarchy Depth:

3 levels



No Selection

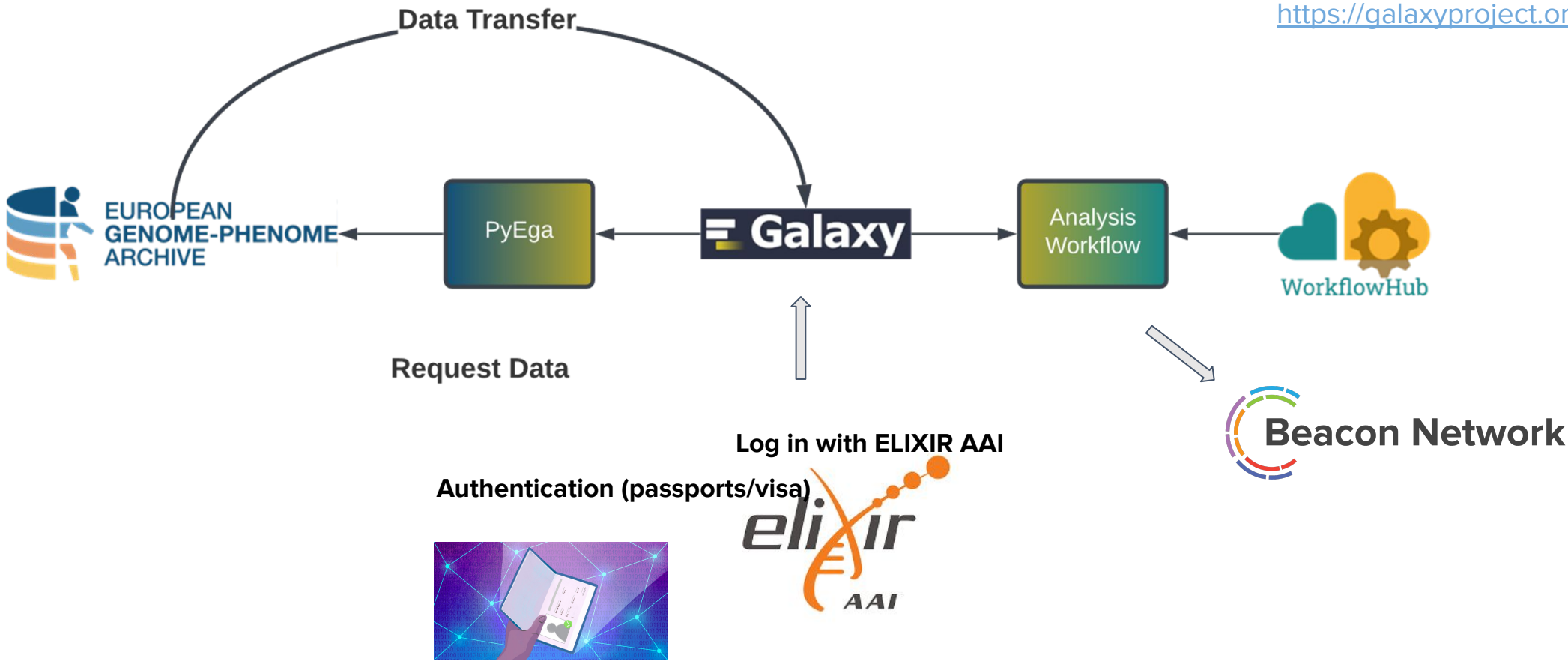
- ☐ [1000_Genomes_Consortium_Phase_1](#): 1000_Genomes_Consortium_Phase_1 (1055 samples, 1055 CNV profiles)
- ☐ [1000_Genomes_Consortium_Phase_3](#): 1000_Genomes_Consortium_Phase_3 (2504 samples, 2504 CNV profiles)
- ☐ [ADM2](#): ADM2 (2092 samples, 2092 CNV profiles)
- ☐ [Birdseye](#): Birdseye (270 samples, 270 CNV profiles)
- ☐ [Custom HMM algorithm](#): Custom HMM algorithm (1552 samples, 1552 CNV profiles)
- ☐ [DRAGEN-CNV](#): DRAGEN-CNV (3201 samples, 3201 CNV profiles)
- ☐ [ELAND \(PEM\)](#): ELAND (PEM) (1 sample, 1 CNV profile)
- ☐ [HMMSeg](#): HMMSeg (9 samples, 9 CNV profiles)
- ☐ [PEM](#): PEM (1 sample, 1 CNV profile)
- ☐ [VAMP](#): VAMP (1 sample, 1 CNV profile)
- ☐ [era-EnsembleCNV](#): era-EnsembleCNV (2001 samples, 2001 CNV profiles)
- ☐ [era-PennCNV](#): era-PennCNV (1993 samples, 1993 CNV profiles)
- ☐ [era-QuantiSNP](#): era-QuantiSNP (2001 samples, 2001 CNV profiles)
- ☐ [era-iPattern](#): era-iPattern (1866 samples, 1866 CNV profiles)
- ☐ [labelSeg-based calibration](#): labelSeg-based calibration (16671 samples, 16671 CNV profiles)
- ☐ [mrFAST \(Read Depth\)](#): mrFAST (Read Depth) (3 samples, 3 CNV profiles)



CINECA

HTSgetGalaxy :FAIR CLOUD Analysis

Galaxy & GA4GH alignment:
<https://galaxyproject.org/ga4gh/>



NOTE: Pilot implementation uses username/password for authentication, future release of Galaxy will support use of passports & visas



Data analysis & reporting: Galaxy clin.iobio trio analysis

- Galaxy Interactive Tool
- Docker based
- Interactive variant exploration

Galaxy Galaxy / Live Europe

Workflow Visualize Shared Data Help Login or Register

CLIN IOBIO
A comprehensive clinical variant analysis workflow

BAM IOBIO
Examine sequence alignment in seconds

VCF IOBIO
Examine your variant file in seconds

Gene DLL4 External links Other tools

Variants in DLL4 Proband variants NAI2878

Proband coverage

Variant in DLL4 External links HGVS = rs533126562 SNP 15:41229631 T->G Cys to Trp at 653

Quality Sufficient depth and allele counts

Gene Associations #193 GTR Hydrocephalus #173 Phars hydrocephalus

Pathogenicity Missense variant 0.785 REVEL

Pop Freq in gnomAD 0% Allele freq

Inheritance de novo 0 of 57 1 of 55

Conservation Highly conserved 5.03 phyloP scores

Phenotypes Input: MPH; Megalocephaly-Polymicrogyria-Polydactyly-Hydrocephalus syndrome; multiple congenital anomalies; tetralogy of fallot; brain anomalies consisting of bilateral polymicrogyria and cortical dysplasia; post axial polysyndactyly of hand and feet; macrosomia affecting head and length; hypotonic with global developmental delays

Phenotypes terms hydrocephalus, megalencephaly, polymicrogyria

Reviewed Variants Significant Variant in DLL4 HGVS rs533126562 SNP chr15:41229631 T->G Cys to Trp at 653

Quality Sufficient depth and allele counts

Gene Associations #193 GTR Hydrocephalus #173 Phars hydrocephalus

Pathogenicity Missense variant 0.785 REVEL

Frequency 0% Allele freq

Inheritance de novo 0 of 57 1 of 55

Conservation Highly conserved 5.03 phyloP scores

HTSgetGalaxy : Summary

We have implemented:

- Secure access to GA4GH EGA service using PyEGA3 Galaxy service
- Standard analysis for B1MG synthetic data (Filtering etc...)
- Interactive gene variant detection for trio-analysis with clin.iobio in Galaxy
- Related Galaxy Training Network (GTN)
- Single source of Cancer analysis tooling for scientific community – Cancer Galaxy
(<https://cancer.usegalaxy.eu/>)



(GIGA)ⁿ
SCIENCE

GigaScience, 2023, 1–6

doi: [xx.xxxx/xxxx](https://doi.org/10.1093/gigascience/giag001)
Manuscript in Preparation
Paper

PAPER

Global Alliance for Genomics & Health Compliant FAIR Data Access for Genomics in Galaxy

Jasper Ouwerkerk^{1,*†}, Helena Rasche^{1,†}, Dylan Spalding², Saskia Hiltermann^{1,†} and Andrew P. Stubbs¹



A step towards creating a
generalized Omics platform for
FAIR data analysis



GLOBAL ALLIANCE FOR GENOMICS & HEALTH COMPLIANT FINDABLE ACCESSIBLE INTEROPERABLE
AND REUSEABLE DATA ACCESS FOR GENOMIC ANALYSIS IN GALAXY

Jasper Ouwerkerk¹, Helena Rasche¹, Dylan Spalding², Saskia Hiltermann¹, and Andrew P. Stubbs¹

¹Erasmus Medical Center, Clinical Bioinformatics Group, Department of Pathology, Rotterdam, The Netherlands

²CSC-IT Center for Science, Espoo, Finland

Methods

- Implemented PyEGA3 in Galaxy
- Implemented a new trio-analysis tool, *clin.iobio*, in Galaxy and compared it to the existing tool GEMINI.
- Built a trio-analysis workflow and tutorial in Galaxy
- Validated workflow on 6 synthetic B1MG use cases.

B1MG Family Trio

- Case 5 out of 6
- Gender: Female
- Age: 35 years
- Referral: Breast Cancer
- Main Clinical Features:
 - Breast carcinoma
 - Neoplasm of the breast

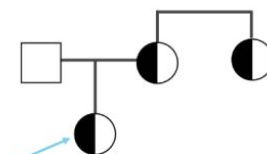


Figure 1: The pedigree of the case to be analyzed

Results

- Created a FAIR trio-analysis workflow in Galaxy, see Figure 2.
- We found that GEMINI reports 244 false positive

FAIR¹

- Findable by both machines and humans
- Accessible using a standard open protocol
- Interoperable so it can easily be processed and analysed
- Reusable so the data can be understood by anyone and make analyses reproducible

Adoption
EGA

EGA: European Genome-Phenome Archive²

- Archive for FAIR data
- Data owners → Data Access Committee

Secure Access
PyEGA3

hCNV Galaxy and Nextflow



Galaxy

- Structural genomic variant calling tools with user-friendly interface for running tools and managing data.
- Galaxy is also a powerful platform for developing and sharing bioinformatics workflows deposited in the WorkflowHub.
- Galaxy training network tutorial (slides, hands on, workflows)



Nextflow

- Nextflow workflows/modules available to be used on a variety of compute platforms, including local machines, clusters, and clouds.
- Nextflow can be used to run hCNV on multiple samples in parallel, which can significantly reduce the amount of time required to analyze a large dataset.

Galaxy Training Network



UNIVERSITY of
BRADFORD



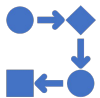
Galaxy Training!



- Define the specific challenges in locating human Copy Number Variances (hCNVs).
- Show how to pre-process the sequenced reads for hCNVs detection.
- Show an example on using Control-FREEC to detect the hCNVs from cancer dataset.
- Visualise the hCNVs' findings.
- Publish the workflow on Galaxy and Workflow hub to be accessible by the community



Control-FREEC Training material



Control-FREEC workflow

Working with Beacon V2: A Comprehensive Guide to Creating, Uploading, and Searching for Variants with Beacons

Author(s)



Khaled Jum'ah



Katarzyna Kamieniecka



Krzysztof Poterłowicz

Reviewers



Overview

Questions:

- What does the term "Beacon" refer to?
- How can MongoDB be employed to establish a Beacon tailored for your institution?
- In what manner can variant data and metadata be readied into a format compatible with Beacons?
- What are the steps involved in importing data into a Beacon seamlessly?
- How does one perform queries on a Beacon to retrieve information about variants?

Objectives:

- Comprehend the fundamental concepts and applications of Beacons
- Apply skills in utilizing MongoDB to construct and manage Beacons
- Analyze and transform variants and metadata into structures compatible with Beacon requirements
- Execute a step-by-step process to import data into Beacons
- Develop the ability to query Beacons for variants



Beacon v2 tools



UNIVERSITY of
BRADFORD

Beacon2 Import

Import JSON formatted datasets to beacon database
(Galaxy Version 1.0.7+galaxy0)



Run Tool

Tool Parameters

DATABASE HOST - optional

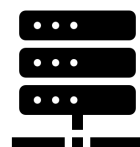
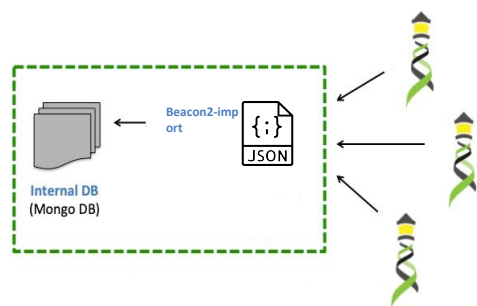
Hostname/IP of the beacon database (--db-host)

DATABASE PORT *

Port of the beacon database (--db-port)

DATABASE PASSWORD - optional

Login password for the beacon database (--db-pass)



Beacon2 Bracket

Specifies a sequence ranges
for both the start and end
positions of a genomic
variation
(Galaxy Version
1.0.7+galaxy0)



Run Tool

Tool Parameters

DATABASE HOST - optional

Hostname/IP of the beacon database (--db-host)

Beacon2 Gene

Queries the beacon
database and retrieve the
genomic variants matching
gene symbol
(Galaxy Version
1.0.7+galaxy0)



Run Tool

Tool Parameters

DATABASE HOST - optional

Hostname/IP of the beacon database (--db-host)

Beacon2 Sequence

Query for the existence of a
specified sequence at a
given genomic position
(Galaxy Version
1.0.7+galaxy0)



Run Tool

Tool Parameters

DATABASE HOST - optional

Hostname/IP of the beacon database (--db-host)

Beacon2 Range

Retrieve the genomic
variants from the beacon
database by specifying start
and end positions
(Galaxy Version
1.0.7+galaxy0)



Run Tool

Tool Parameters

DATABASE HOST - optional

Hostname/IP of the beacon database (--db-host)

hCNV Galaxy and Nextflow ecosystem



hCNV data analyses with Nextflow modules



UNIVERSITY of
BRADFORD

modules

Public

Forked from [nf-core/modules](#)

Repository to host tool-specific module files for the Nextflow DSL2 community!



biohackeu23



nf-core



Nextflow MIT License Updated last month



new module: beacon2-ri-tools #4264

Open

kkamieniecka opened this issue last month · 0 comments



kkamieniecka commented last month

Member



beacon2-ri-tools transform genomic variations data (VCF) to queryable data (MongoDB)

See extended documentation [here](#).

+ Add tasklist



kkamieniecka added the **new module** label last month



bioinformatics and ML tools for CNV assessment

... Antonio



CNV annotation & Query Standards Development

GA4GH VRS & Beacon Scouts

- unambiguous representation and recovery of genomic variation data still represent major problems in large scale genomic studies (meta-analyses and data federated data discovery)
- structural genome variants present additional challenges both on epistemic and technical levels
- GA4GH has emerged as the major driver for standards development and harmonization
- ELIXIR hCNV members contribute to the development and improvement of CNV representation and query options as well as associated vocabularies
 - VCF 4.4+ D and J flags (e.g. indicating dosage w/o locus assignment)
 - EFO relative copy number count term tree
 - VRS v1.3/v2 representation of copyNumberChange and copyNumberCount
 - Beacon query types specifically suited for CNV or "fuzzy location" events



CNV annotation & Query Standards Development

GA4GH VRS & Beacon Scouts: CNV Classes

GA4GH VRS1.3+	Beacon	VCF v4.4+	SO
EFO:0030070 gain	DUP or EFO:0030070	DUP SVCLAIM=D	SO:0001742 copy_number_gain
EFO:0030071 low-level gain	DUP or EFO:0030071	DUP SVCLAIM=D	SO:0001742 copy_number_gain
EFO:0030072 high-level gain	DUP or EFO:0030072	DUP SVCLAIM=D	SO:0001742 copy_number_gain
EFO:0030072 high-level gain	DUP or EFO:0030073	DUP SVCLAIM=D	SO:0001742 copy_number_gain
EFO:0030067 loss	DEL or EFO:0030067	DEL SVCLAIM=D	SO:0001743 copy_number_loss
EFO:0030068 low-level loss	DEL or EFO:0030068	DEL SVCLAIM=D	SO:0001743 copy_number_loss
EFO:0020073 high-level loss	DEL or EFO:0020073	DEL SVCLAIM=D	SO:0001743 copy_number_loss
EFO:0030069 complete genomic loss	DEL or EFO:0030069	DEL SVCLAIM=D	SO:0001743 copy_number_loss



CNV annotation & Query Standards Development

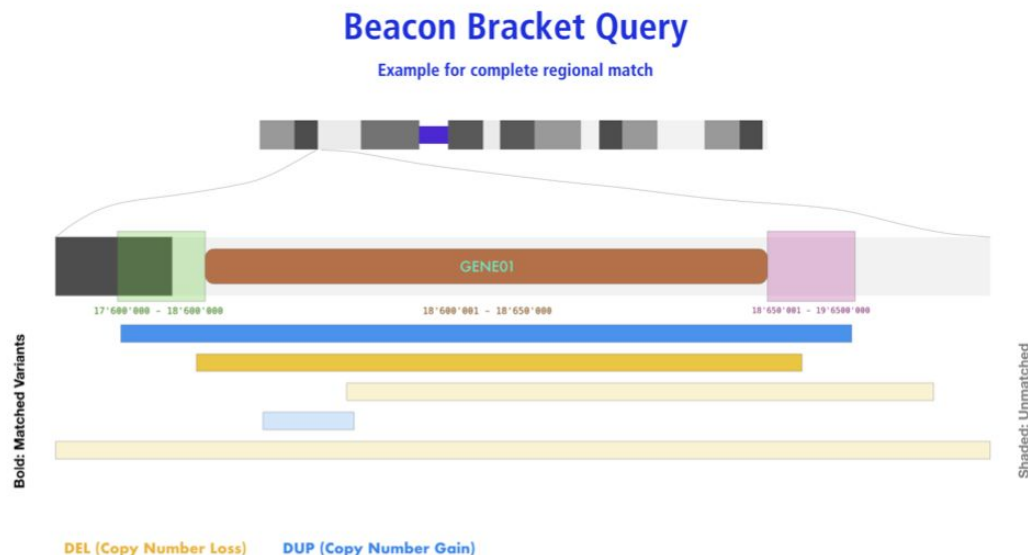
Beacon Scouts: CNV Queries



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 Sarr

Dataset
Test Database - examplez x

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

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

Select Filters NCIT:C3058: Glioblastoma (100) x

Chromosome 9 21000001 21975098 21967753 23000000

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.



CNV annotation & Query Standards Development

Beacon Scouts: Moving Ahead



Beacon Scouts

Finding the Paths to Beacon's Future

- Genomic Variation Scouts
 - ➔ extension to the query model based on assessed needs
 - ▶ fusions/breakpoints, cytogenetic annotations, repeats, categorical variants...
 - ➔ adoption of evolving VRS... standards for variant representation
 - ▶ adjacency, repeats...
 - ▶ re-use of parameters where clear (e.g. **sequenceLength** instead of **variantMinLength** + **variantMaxLength**)

GA4GH Beacon Genomic Variation Query Standards

Beacon VQS Requests

The `VQRequest` type represents the generic collection of variant parameters supported in Beacon v2+ requests. These include parameters with close alignment to VRS v2 concepts and replacing some Beacon v1/v2 generics with tighter definitions (e.g. `referenceAccession` instead of `referenceName` and `accession` or `copyChange` for a specific subset of former `variantType` values) but also keep some concepts beyond VRS scope or specifically geared towards query applications (`geneId`, `sequenceLength`)

For the parameter definitions please see the [requestParameterComponents](#) page.

VQRequest Parameters

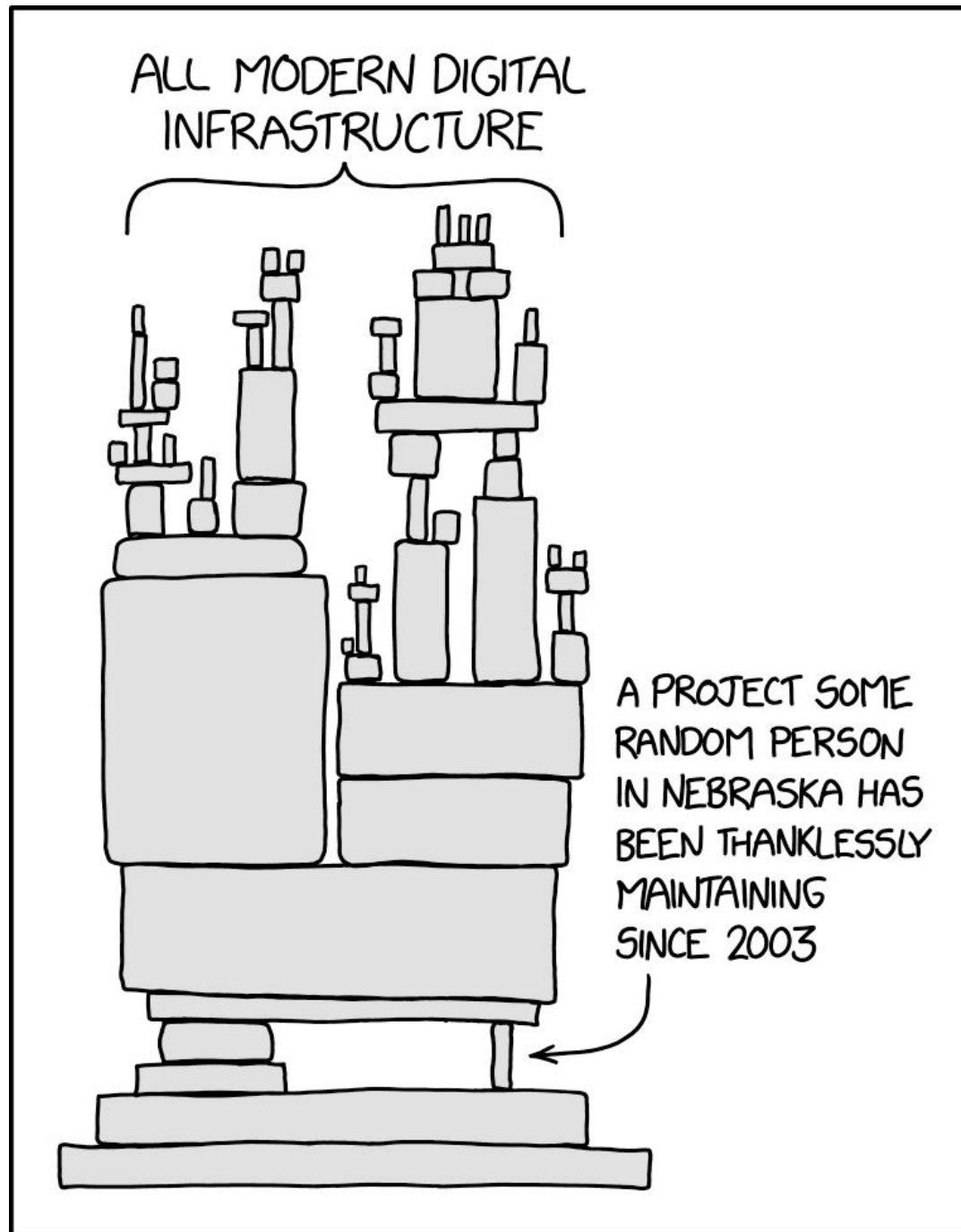
```
requestProfileId : ./requestParameterComponents.yaml#/$defs/RequestProfileId
referenceAccession : ./requestParameterComponents.yaml#/$defs/RefGetAccession
start : ./requestParameterComponents.yaml#/$defs/SequenceStart
end : ./requestParameterComponents.yaml#/$defs/SequenceEnd
sequence : ./requestParameterComponents.yaml#/$defs/Sequence
copyChange : ./requestParameterComponents.yaml#/$defs/CopyChange
adjacencyAccession : ./requestParameterComponents.yaml#/$defs/AdjacencyAccession
adjacencyStart : ./requestParameterComponents.yaml#/$defs/AdjacencyStart
adjacencyEnd : ./requestParameterComponents.yaml#/$defs/AdjacencyEnd
repeatSubunitCount : ./requestParameterComponents.yaml#/$defs/RepeatSubunitCount
repeatSubunitLength : ./requestParameterComponents.yaml#/$defs/RepeatSubunitLength
geneId : ./requestParameterComponents.yaml#/$defs/GeneId
aminoacidChange : ./requestParameterComponents.yaml#/$defs/AminoacidChange
genomicAlleleShortForm :
./requestParameterComponents.yaml#/$defs/GenomicAlleleShortForm
sequenceLength : ./requestParameterComponents.yaml#/$defs/SequenceLength
vrsType : ./requestParameterComponents.yaml#/$defs/VRSType
```

Table of contents

- VQRequest Parameters
- Beacon v2+/VQS "VRSified" Request Examples
- Copy number gains involving the whole locus chr2:54,700,000-63,900,000
- Focal high-level deletion involving the CDKN2A locus
- Find t(8;14)(q24;q32) translocations
- CAG repeat in the first exon of the huntingtin gene (HTT)
- CAG repeat in the first exon of the huntingtin gene (HTT)
- CGG trinucleotide repeat expansion in the FMR1 gene
- Query for a focal deletion involving TP53



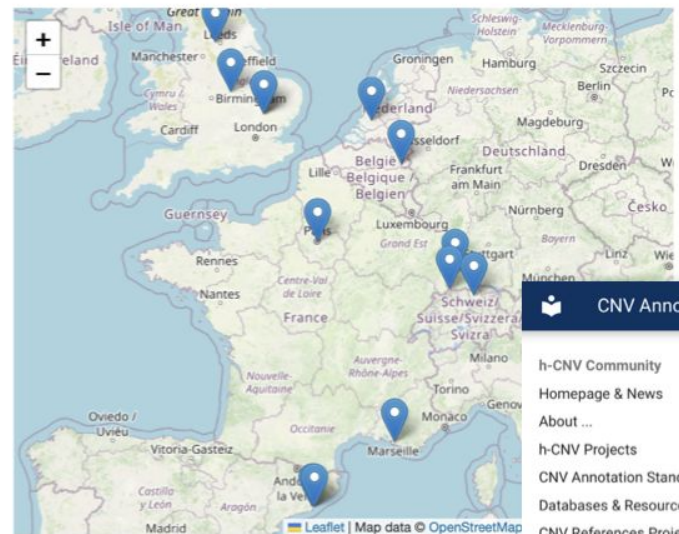
Contribute
Participate
Exchange
Maintain



- h-CNV Community
- Homepage & News
- About ...
- h-CNV Projects
- CNV Annotation Standards
- Databases & Resources
- CNV References Project
- Contacts
- Genome Blog
- h-CNV @ ELIXIR
- Beacon Project

ELIXIR Human Copy Number Variation community

Among the different types of inherited and acquired genomic variants, regional genomic copy number variations (CNV) contribute - if measured by affected genomic sequences - contribute by far the largest amount of genomic changes, contributing both to many syndromic diseases as well as the vast majority of human cancers. The website of the *Human Copy Number Variation Community* (hCNV) is a resource originated in ELIXIR's h-CNV Community Implementation Study (2019-2021) with the aim to provide a resource hub and knowledge exchange space for scientists and practitioners working with - or being interested in - genomic copy number variations in health and diseases. However, the scope of the community extends beyond CNVs and includes definition of and work with other types of genomic variations with a focus on structural variants.



ELIXIR hCNV Community

<https://cnvar.org/>

- h-CNV Community
- Homepage & News
- About ...
- h-CNV Projects
- CNV Annotation Standards
- Databases & Resources
- CNV References Project
- Contacts
- Genome Blog
- h-CNV @ ELIXIR
- Beacon Project

CNV Term Use Comparison in Computational (File/Schema) Formats

This table is maintained in parallel with the [Beacon v2 documentation](#).

EFO	Beacon	VCF	SO	GA4GH VRS ¹	Notes
EFO:0030070 copy number gain	DUP ² or EFO:0030070	DUP SVCLAIM=D ³	SO:0001742 copy_number_gain	EFO:0030070 gain	a sequence alteration whereby the copy number of a given genomic region is greater than the reference sequence
EFO:0030071 low-level copy number gain	DUP ² or EFO:0030071	DUP SVCLAIM=D ³	SO:0001742 copy_number_gain	EFO:0030071 low-level gain	
EFO:0030072 high-level copy number gain	DUP ² or EFO:0030072	DUP SVCLAIM=D ³	SO:0001742 copy_number_gain	EFO:0030072 high-level gain	commonly but not consistently used for >=5 copies on a bi-allelic genome region
EFO:0030073 focal genome amplification	DUP ² or EFO:0030073	DUP SVCLAIM=D ³	SO:0001742 copy_number_gain	EFO:0030072 high-level gain ⁴	commonly but not consistently used for >=5 copies on a bi-allelic genome region, of limited size (operationally max. 1-5Mb)
EFO:0030067 copy number loss	DEL ² or EFO:0030067	DEL SVCLAIM=D ³	SO:0001743 copy_number_loss	EFO:0030067 loss	a sequence alteration whereby the copy number of a given genomic region is smaller than the reference sequence
EFO:0030068 low-level copy number loss	DEL ² or EFO:0030068	DEL SVCLAIM=D ³	SO:0001743 copy_number_loss	EFO:0030068 low-level loss	
EFO:0020073 high-level copy number loss	DEL ² or EFO:0020073	DEL SVCLAIM=D ³	SO:0001743 copy_number_loss	EFO:0020073 high-level loss	a loss of several copies; also used in cases where a complete genomic deletion cannot be asserted

display a menu



Thank you



www.elixir-europe.org

