

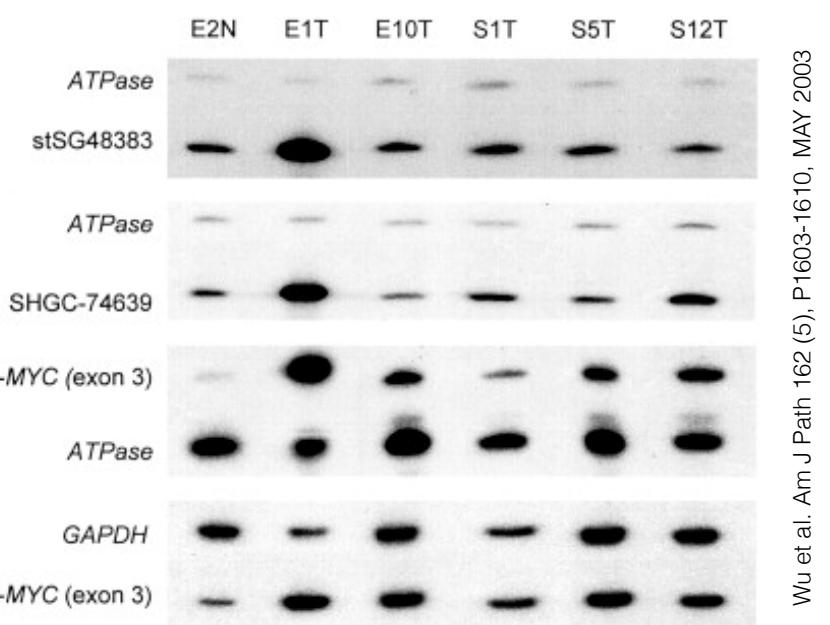
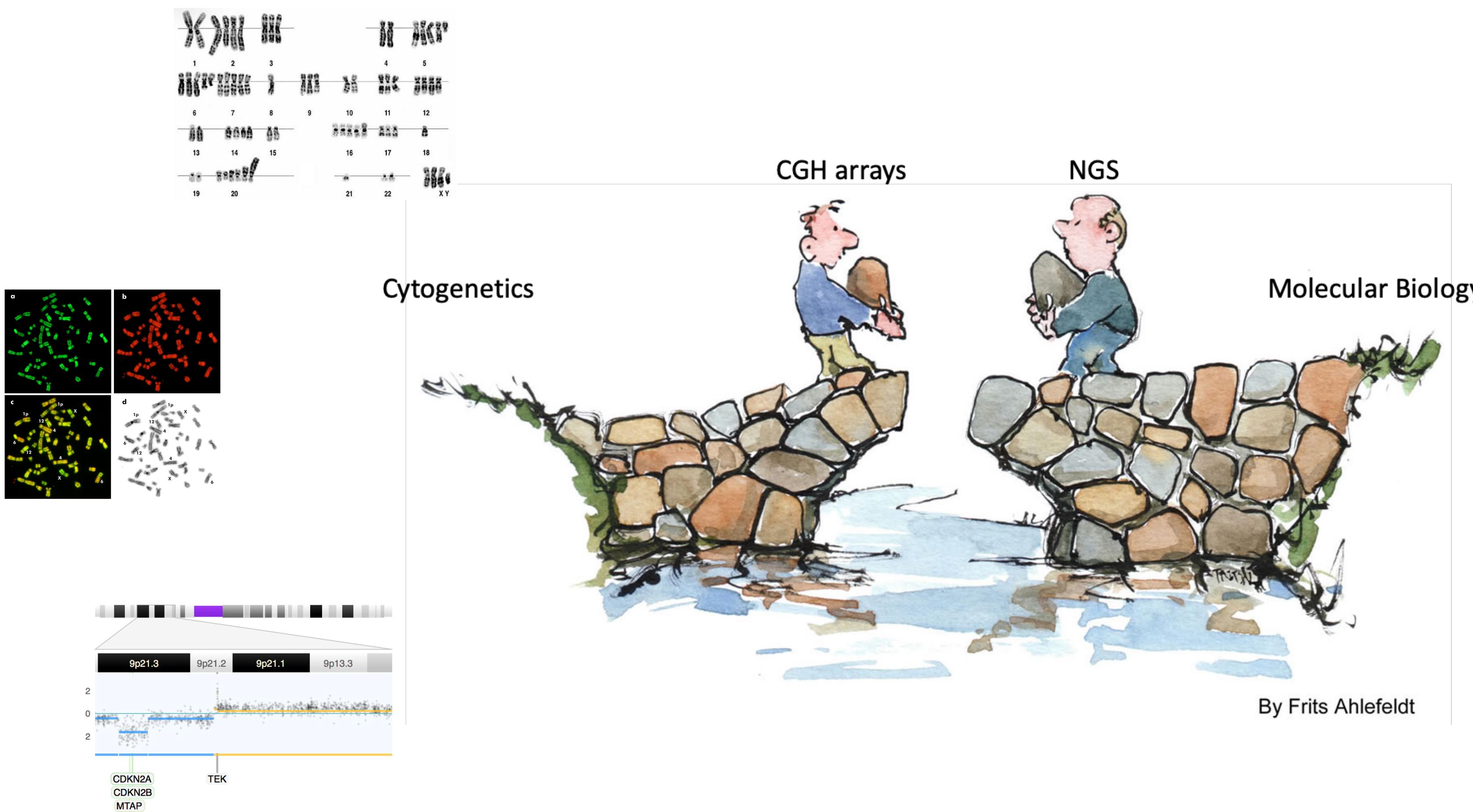


ELIXIR hCNV Community

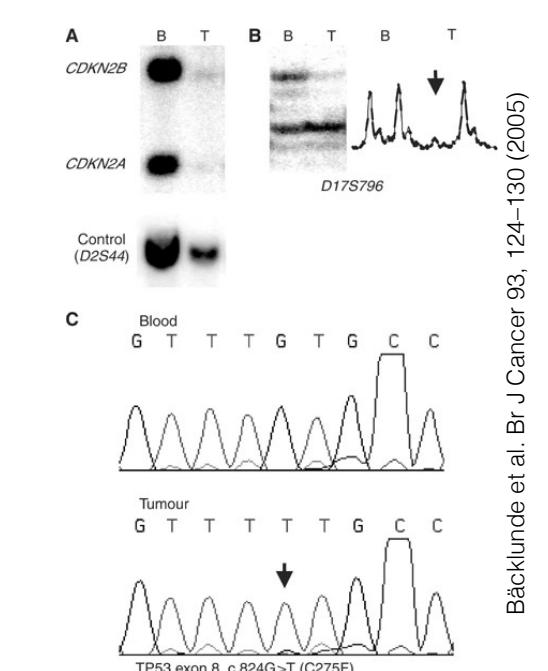
www.elixir-europe.org

h-CNV scientific context

- Structural variants have been the first ones to be detected in humans (late 1950s)
- Genes' mutations shortly followed (Ingram et al. 1957)



Wu et al. Am J Path 162 (5), P1603-1610, May 2003

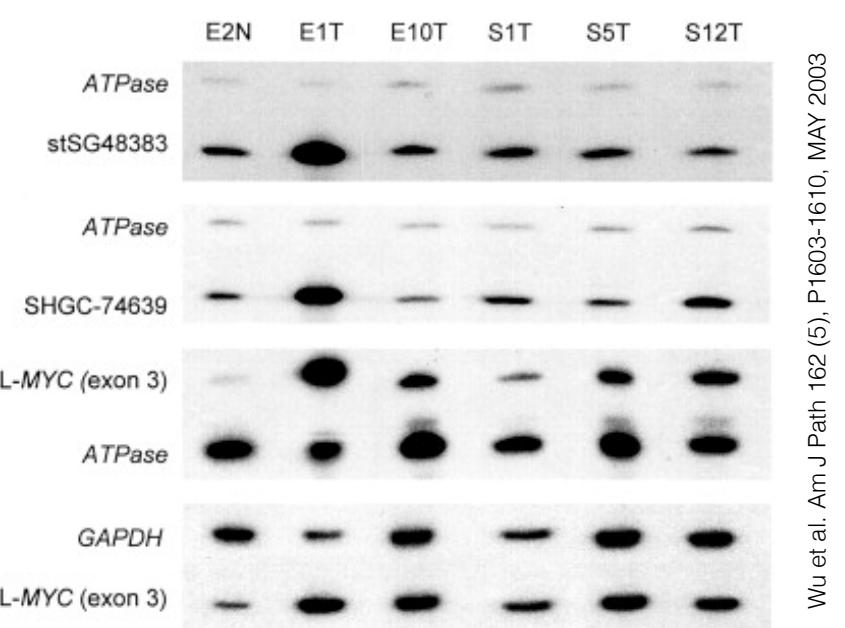
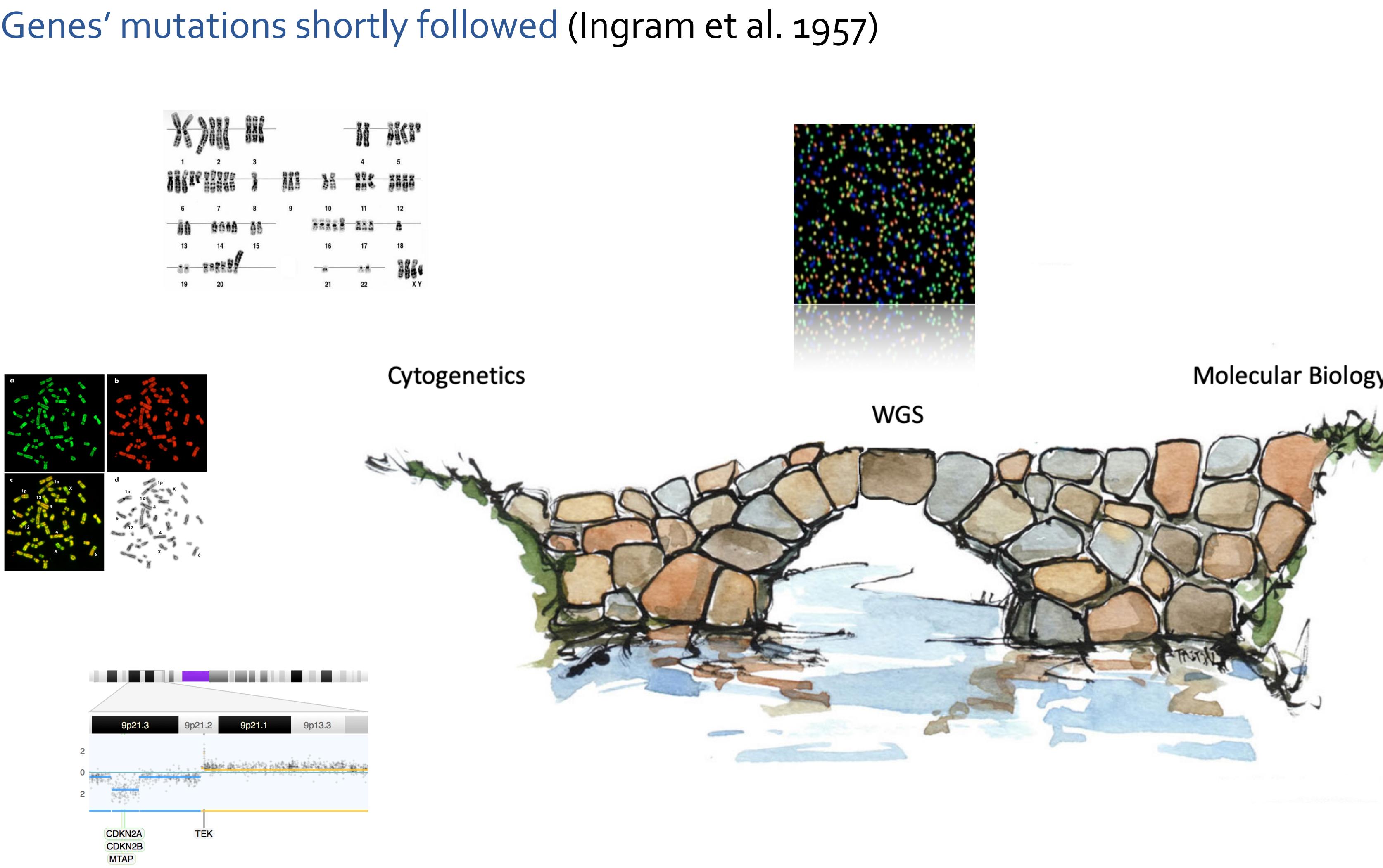


Bäcklund et al. Br J Cancer 93, 124-130 (2005)

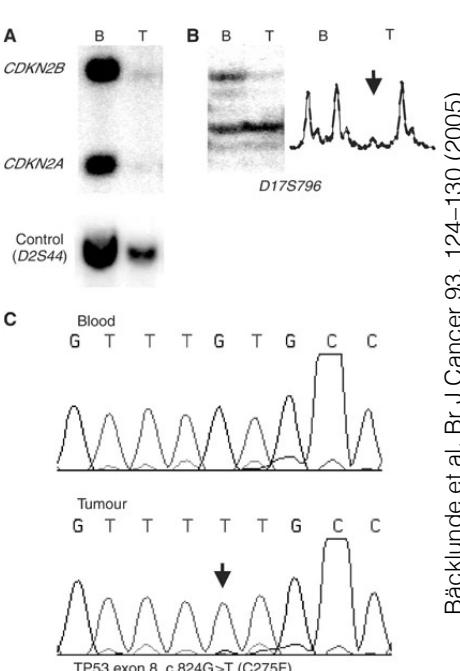


h-CNV scientific context

- Structural variants have been the first ones to be detected in humans (late 1950s)
- Genes' mutations shortly followed (Ingram et al. 1957)



Wu et al. Am J Path 162 (5), P1603-1610, May 2003



Bäcklund et al. Br J Cancer 93, 124-130 (2005)

h-CNV scientific context



- Structural variants have been the first ones to be detected in humans (late 1950s)
- Genes' mutations shortly followed (Ingram et al. 1957)
- CNV (Copy Number Variation) are the most frequent genetic cause of diseases
 - Submicroscopic aberration in 1:270 fetuses (general population - Srebnik et al. 2017)
 - Account for 5 to 60% of mutations in RD genes
 - Widespread **somatic mutations** (deletions, duplications, high-level amplifications...) in **cancer genomes**
 - ~20% of "average" cancer genome is imbalanced (own data @ progenetix.org)
 - Involvement in common diseases (obesity, cardiovascular diseases ...)
 - Play a critical role in species phenotypic diversity (comparison between ancient DNA and distinct human populations - Sudmant et al. Science. 2015)
- Detection, annotation and interpretation is not as efficient as for SNV
 - **Limited standards** for analysis & reporting due to complex & quantitative nature
 - Need for improvements both for **research** and **clinical practice**

hCNV Community



1st community meeting held in September 2018 - Hinxton

Community officially approved in February 2019

11 ELIXIR nodes, represents around 35 people



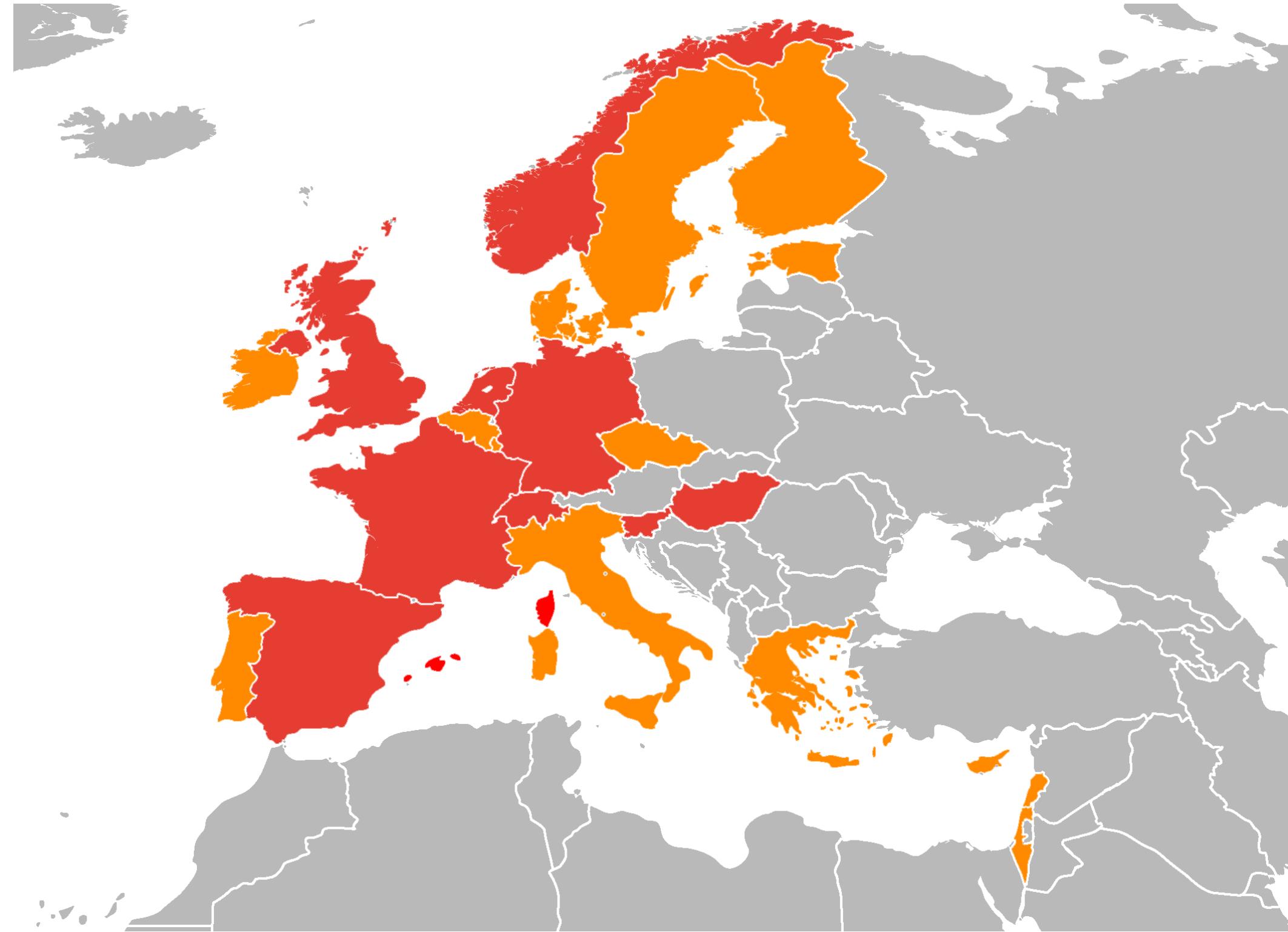
Christophe Béroud
(ELIXIR France)



David Salgado
(ELIXIR France)



Michael Baudis
(ELIXIR Switzerland)



Mission statement

Despite the fact that **Copy Number Variations** are the **most prevalent genetic mutation type**, identifying and interpreting them is still a major challenge. The ELIXIR human Copy Number Variation (hCNV) Community aims to implement processes to make the **detection, annotation, interpretation and sharing** of these variations easier.

<https://elixir-europe.org/communities/hcnv>

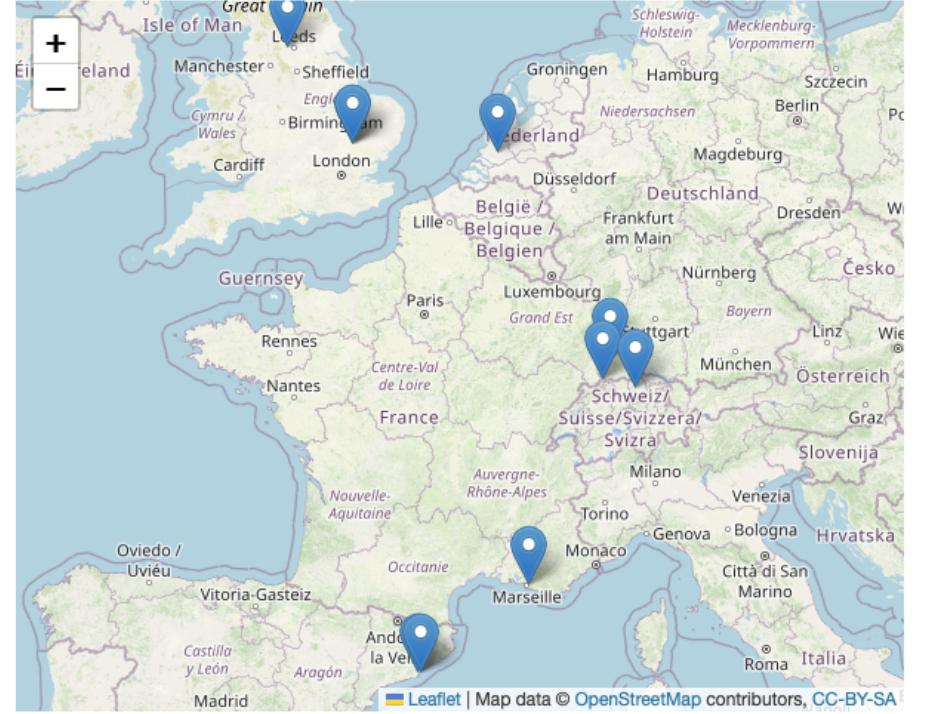


Community organisation

- Bi-weekly conference calls (1 general / 1 technical) → future 1 TC / month
- Participation to the HDC conference calls and to ELIXIR 2019-2023 Programme (back-and-forth reporting)
- Community workshop: 1st semester 2023
- Information through ELIXIR-hCNV mailing list - hcnv@elixir-europe.org
- ELIXIR hCNV community webpage
 - elixir-europe.org/communities/hcnv
- Dedicated hCNV custom domain & website
 - cnvar.org
- hCNV GitHub repository for visibility, communication and sharing
 - github.com/hcnv

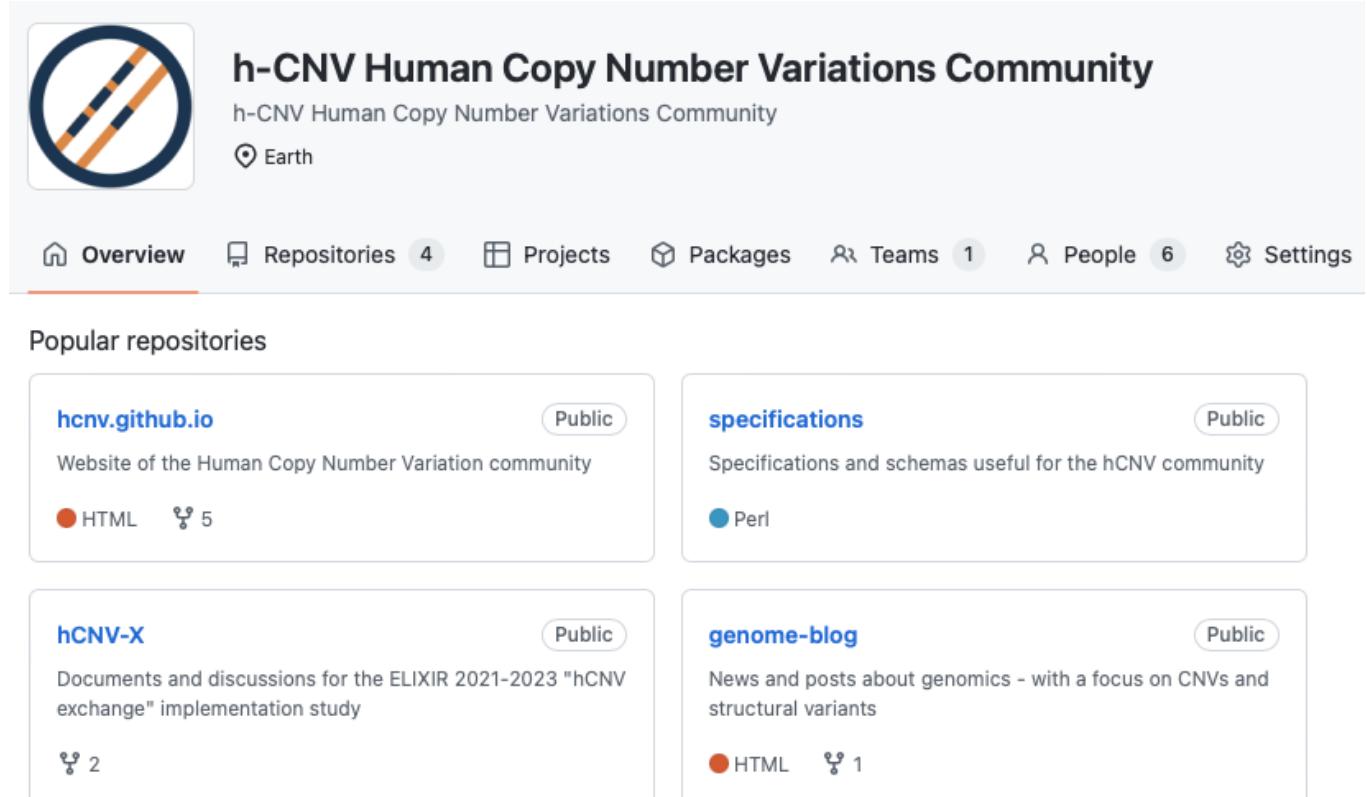
h-CNV Community
hCNV Home
About ELIXIR h-CNV
h-CNV Projects
CNV Annotation Standards
Databases & Resources
Genome Blog
Contacts
h-CNV @ ELIXIR
Beacon Project

The [website](#) of the *Human Copy Number Variation Community (hCNV)* is a resource originating in ELIXIR's h-CNV Community Implementation Study (2019-2021).



[CNV News and Announcements](#)

[CNV Annotation Standards](#)



h-CNV Human Copy Number Variations Community
h-CNV Human Copy Number Variations Community 

[Overview](#) [Repositories 4](#) [Projects](#) [Packages](#) [Teams 1](#) [People 6](#) [Settings](#)

Popular repositories

hcnv.github.io   5	specifications  Specifications and schemas useful for the hCNV community 
hCNV-X  Documents and discussions for the ELIXIR 2021-2023 "hCNV exchange" implementation study  2	genome-blog  News and posts about genomics - with a focus on CNVs and structural variants  1



Main achievements since community creation



Dissemination



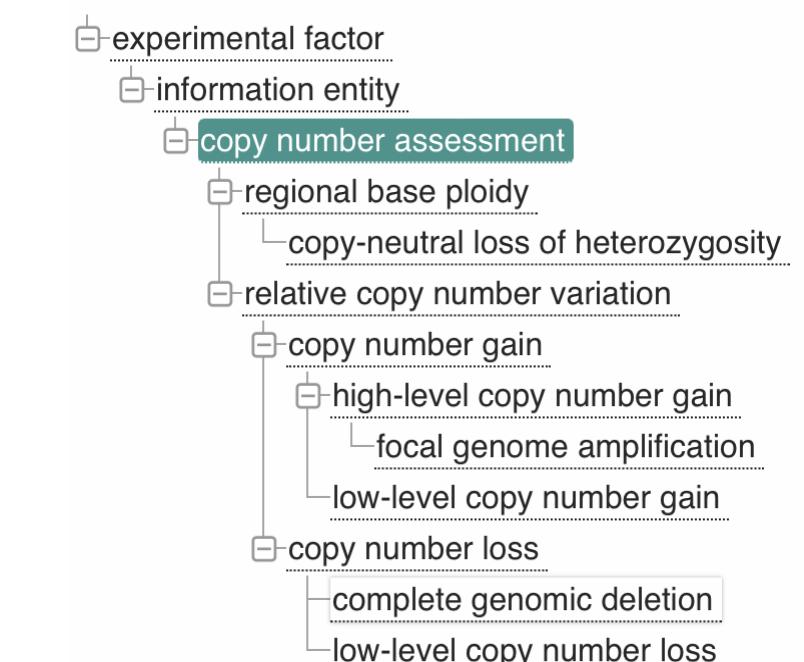
- Publication of a F1000 paper describing the hCNV community and its objectives (PMID 34367618)
- Organization of Bio Hackathons → Instrumental to commissioned services
 - 2019 "Beacon4CNV proof of concept"
 - 2020 "CNV detection tools and benchmark"
 - 2021 "Galaxy training resources for CNVs detection software"
- CNV workshop: HGVS satellite meeting of the ESHG meeting, 2018, Göteborg
- Community presentations outside ELIXIR (extract)
 - 12th CeBiTec Symposium "Big Data in Medicine and Biotechnology". March 2018, Bielefeld, Germany.
 - Achromatic Puce Meeting, 2019, France, GA4GH VRS taskforce 2021, Progenetix @ Cancer Genomics Consortium Annual Meeting 2021, GA4GH Connect Beacon Workshop 2022 ...
- Community presentations within ELIXIR (extract)
 - @ELIXIR All Hands
 - 2018, Germany Berlin; 2019, Portugal, Lisboa (dedicated Community Workshop); 2021 (virtual); 2022 (in Beacon, Machine Learning, OpenEbench workshops)



Standards



- Document on recommendations for CNV annotations ([document](#))
- Contribution to Beacon specifications
 - h-CNV members were part of the Beacon v2 variant query scout team
 - definition of [Beacon v2 query types](#) (Bracket Query and range Query) suitable for CNV discovery
- Document on Phenotypic and Ontologies for CNVs ([document](#))
- Survey on CNV standards ([document](#))
- Addition of specific terms to the EDAM ontology for CNV detection tools description
- Addition of CNV result & platform terms to EFO ontology
- Exploration of use of Beacon on CNV resources
 - [Progenetix](#) (CH, cancer)
 - [BANCCO](#) (FR, rare diseases)



Reference materials



- Web page listing documents generated in the study ([link](#))
- Listing of current available reference datasets from published studies ([document](#)) (15 studies)
- Addition of CNV tools to the bio.tools catalogue and creation of a bio.tools hCNV subdomain (<https://elixir-hcnv.bio.tools/>) (110 in bio.tools)

Links with ELIXIR platforms and communities

- Involvement in Galaxy activities (prototyping of CNV workflows for benchmarking, hCNV galaxy training material)
- Involvement in Containers activities, Tools/Compute platform - SIS BioContainers (Containerization and Galaxy Wrapping of hCNV tools) - [containers](#), [bio.tools](#))
- Prototype of CNV benchmark activities in OpenEBench
- Contribution to ELIXIR and GA4GH Beacons activities
- Open science strategy adoption by the community - FAIRification of CNV resources



Links with other scientific communities and users groups



- EOSC-Pillar, contribution - CNV shared repository for Gold Standard CNV detection (ELIXIR-FR)
- Interactions with the TransBioNet project ([CNV detection in Diagnostic context in Spain](#)) (ELIXIR-ES) and AChroPuce Network ([in France](#)) (ELIXIR-FR)
- hCNV members participation to GA4GH projects and work streams (VRS, Future of VCF, GA4GH Beacon SV scout team, Phenopackets)
 - Participation to the GA4GH Connect Beacon Workshop 2022 (Michael Baudis, David Salgado)
 - Participation to VCF specification and adaptation to CNV (Timothée Cezard)
- Bioinformatics research team @ Imagine Institute (Antonio Rausell, Paris, France) to build a Machine Learning based classifier to predict the pathogenicity of CNV (paper submitted at AJHG)

and now



Community led Implementation Studies - 2021-2023



Reference hCNV datasets, use-case workflows and benchmarking

Lead: Christophe Béroud and David Salgado, ELIXIR-FR

Partners: Michael Baudis, ELIXIR-CH
Sergi Beltran, Steven Laurie, ELIXIR-ES
Salvador Capella-Gutierrez, ELIXIR-ES
Björn Grüning, ELIXIR-DE
Krzysztof Poterlowicz, ELIXIR-UK

Beacon and beyond - Implementation-driven standards and protocols for CNV discovery and data exchange

Lead: Michael Baudis, Alexander Kanitz, ELIXIR-CH

Partners: Christophe Béroud, David Salgado, ELIXIR-FR
Babita Singh, Jordi Rambla, Sergi Beltran, Steven Laurie, ELIXIR-ES
Timothée Cezard, Kirill Tsukanov, EMBL-EBI
Tim Beck, Anthony Brookes, Krzysztof Poterlowicz, ELIXIR-UK
Björn Grüning, ELIXIR-DE

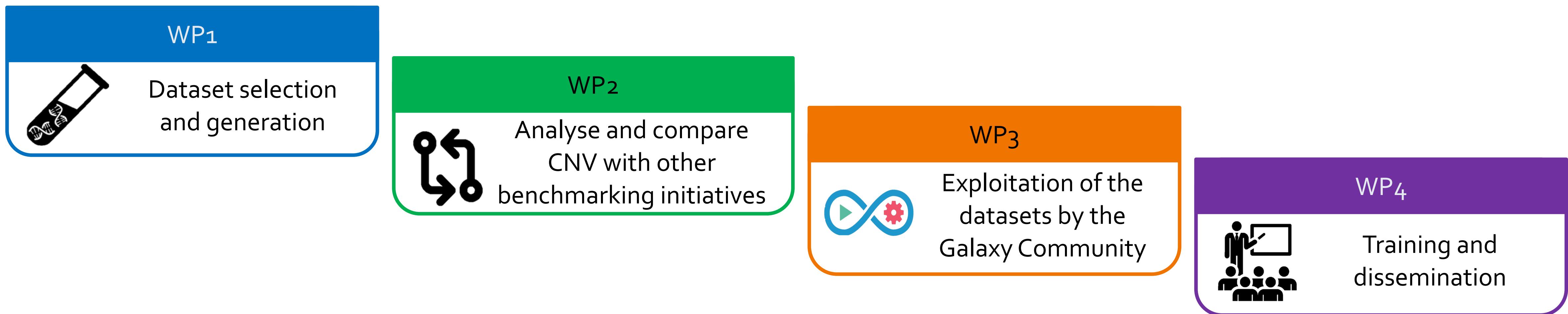


To address the needs

6 partners from 5 different ELIXIR Nodes

- ELIXIR-FR, ELIXIR-UK, ELIXIR-CH, ELIXIR-ES, ELIXIR-DE
- Involved in many Elixir communities (hCNV, Rare-diseases, Galaxy, Containers, Beacon, ...)
- Platforms (Tools, Training, ...)

Proposed a plan in 4 work packages





biological samples selection (G. Collod-Béroud)

- Use of the CNV dataset list of publications from the WP2 first hCNV IS. Enriched by additional publications of the existing reference materials for WGS, WES and gene panels (26 publications)
- Identification of the most frequently used individuals with biological samples available at the **Coriell Institute Biobank** (<https://coriell.org/1/Browse/Biobanks>) cell-line provider as **extracted DNA**.

Coriell references	Gender	Family numbers	Origine	Number of recurrence
NA12878	F	1463	UTAH/MORMON	14
NA24385	M	3140	White	6
NA19240	F	Y117	Nigerian Yoruba	5
NA10847	F	1334	UTAH/MORMON	4
NA10851	M	1334	UTAH/MORMON	4
NA11993	F	1362	UTAH/MORMON	4
NA12156	F	1408	UTAH/MORMON	4
HG00733	F	PR05	Porto Rican	4



Most cited samples in the 26 selected publications



- Search for relatives → trio (27 samples)

Coriell ID	Gender	Family ID	Origine	Relationship	Recurence	HapMap and 1000 genomes Project
NA12878	F	1463	UTAH/MORMON	Mother	14	HapMap et 1000 Genomes Phase 1
NA12891	M	1463	UTAH/MORMON	Maternal grandfather	4	HapMap et 1000 Genomes Phase 1
NA12892	F	1463	UTAH/MORMON	Maternal grandmother	4	HapMap et 1000 Genomes Phase 1
NA24385	M	3140	White	Son	6	
NA24143	F	3140	White	Mother	1	
NA24149	M	3140	White	Father	1	
NA19240	F	Y117	Nigerian Yoruba	Daughter	5	HapMap et 1000 Genomes Phase 1
NA19238	F	Y117	Nigerian Yoruba	Mother	3	HapMap et 1000 Genomes Phase 1
NA19239	M	Y117	Nigerian Yoruba	Father	3	HapMap et 1000 Genomes Phase 1
NA10847	F	1334	UTAH/MORMON	Mother	4	HapMap et 1000 Genomes Phase 1
NA12239	F	1334	UTAH/MORMON	Maternal grandmother	2	HapMap
NA12146	M	1334	UTAH/MORMON	Maternal grandfather	2	HapMap
NA10851	M	1334	UTAH/MORMON	Father	4	HapMap et 1000 Genomes Phase 1
NA12144	M	1334	UTAH/MORMON	Paternal grandfather	3	HapMap et 1000 Genomes Phase 1
NA12145	F	1334	UTAH/MORMON	Paternal grandmother	2	HapMap
NA11993	F	1362	UTAH/MORMON	Paternal grandmother	4	HapMap et 1000 Genomes Phase 1
NA11992	M	1362	UTAH/MORMON	Paternal grandfather	3	HapMap et 1000 Genomes Phase 1
NA10860	M	1362	UTAH/MORMON	Father	2	HapMap
NA12156	F	1408	UTAH/MORMON	Maternal grandmother	4	HapMap et 1000 Genomes Phase 1
NA12155	M	1408	UTAH/MORMON	Maternal grandfather	3	HapMap et 1000 Genomes Phase 1
NA10831	F	1408	UTAH/MORMON	Mother	1	HapMap
HG00733	F	PR05	Porto Rican	Daughter	4	
HG00731	M	PR05	Porto Rican	Father	2	1000 Genome Phase 1
HG00732	F	PR05	Porto Rican	Mother	2	1000 Genome Phase 1

+ Chinese (NA24631, NA24694, NA24695) trio

+ 3 patients from Copy Number Variation Panel - CNVPANEL01 or CNVPANEL02

<https://www.coriell.org/Search?q=copy%20number> NA04327 DMD dup5-7; NA05090 DMD del 3-5; NA04520 TSC2 del 1-14



WP2

Benchmarking process (in progress)



Analyse and compare CNV with other benchmarking initiatives

Assemble CNV detection tools to Workflows
(containers, wrapping, ...)

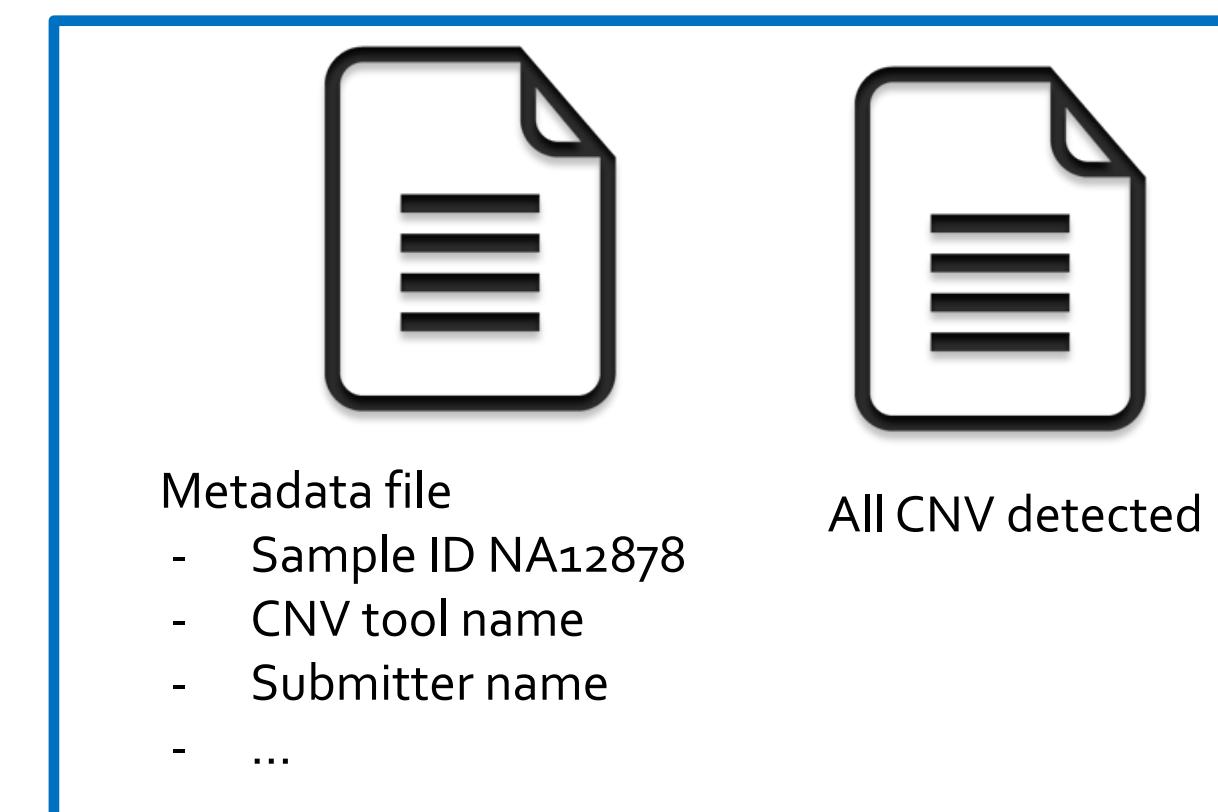
Galaxy **SnakeMake** **nextflow**

Galaxy
WF shared through


Consensus CNV
detected by hCNV
community

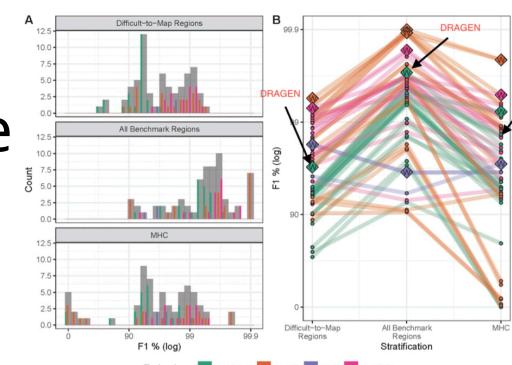
CNV detected from
published studies
(e.g DGVa datasets)

Will allow users to run WF on shared or private instances
(on sample shared either between hCNV ref datasets or
published studies)



OpenEbench
API client

- Compute Metrics (need to be defined)
- Display comparison on each reference datasets
- Store & share results





Tools containerized and/or wrapped in Galaxy

WES detection tools

Containerized										Galaxy integration									
	Title	Description	Source	Tool version	CONDA	CONDA source	Wrapper source	Wrapper owner	Wrapper version	Comment	Pull requests (Linked)	Status	Assignee						
1	AbsCN-seq	AbsCN-seq is an R package	https://bioinformaticshome.org/	1.0	NO							To add							
2	Accuracy	Accuracy is a software tool	https://www.yfish.org/display/accuracy		NO					(whole exome may work too)		To add							
3	ADTEX	A tool for detection of copy number variants	https://bioinformaticshome.org/	2.0	NO							To add							
4	aCNViewer	aCNViewer is a tool for visualizing CNVs	https://bioinformaticshome.org/	2.2	NO							To add							
5	Bamgineer	A tool to simulate haplotypes	https://bioinformaticshome.org/	2.0	NO							To add							
6	BubbleTree	BubbleTree is an R package	https://bioinformaticshome.org/	2.14.0	Yes	https://anaconda.org/search?q=bubbletree						To add							
7	CANOES	CANOES is a tool for detecting CNVs	https://bioinformaticshome.org/	*** the tool is not available	NO							To add							
8	CODEX2	The tool uses depth information to detect CNVs	https://github.com/yuchaoji		NO							To add							
9	Control-FREEC	a tool for assessing copy number variants	https://github.com/BoevaLab	v11.6	Yes	https://anaconda.org/search?q=control-freec	https://github.com/galaxyprj	iuc	11.6			Up-to-date							
10	CoNIFER	Uses exome sequencing data to detect CNVs	https://sourceforge.net/projects/conifer	0.2.1	Yes	https://anaconda.org/search?q=conifer						To add							
11	CONTRA	a tool for copy number variation detection	https://contra-cnv.sourceforge.net	v2.0.8	NO		https://toolshed.g2.bx.psu.edu/shed/module/fcaramia					To update							
12	CoNVaDING	A tool for identification of CNVs	https://bioinformaticshome.org/	1.2.0	NO							To add							
13	CloneCNA	CloneCNA is a tool to detect CNVs	https://bioinformaticshome.org/	2.0	NO							To add							
14	CNVkit	is a Python library and command-line tool for CNV detection	https://github.com/etal/cnvl	0.9.9	Yes	https://anaconda.org/search?q=cnvkit						To add							
15	cn.mops	cn.mops (Copy Number mops) is a Python script	https://bioinformaticshome.org/	1.30.0	Yes	https://anaconda.org/search?q=cn.mops						To add							
16	CNspector	CNspector is a web-based tool for CNV detection	https://bioinformaticshome.org/		NO							special methods for WES	To add						
17	CNVfinder	CNVfinder is a tool to detect CNVs	https://bioinformaticshome.org/	1.0.0b5	NO							To add							
18	CNVnator	A tool for discovery and characterization of CNVs	https://bioinformaticshome.org/		Yes	https://anaconda.org/search?q=cnvnator	https://toolshed.g2.bx.psu.edu/shed/module/fanruimeng			It was used in a paper to detect CNVs		Neglected							
19	cnvOffSeq	cnvOffSeq is specifically designed for CNV detection	https://bioinformaticshome.org/	0.1.2	NO							To add							
20	CN_Learn	CN_Learn is a tool for copy number variation detection	https://bioinformaticshome.org/		NO							To add							
21	DeAnnCNV	DeAnnCNV (Detection and annotation of CNVs)	https://bioinformaticshome.org/		NO							To add							
22	EXCAVATOR2	EXCAVATOR2 is a tool for CNV detection	https://bioinformaticshome.org/	11.2	NO							To add							
23	ExCNVSS	ExCNVSS is a software tool for CNV detection	https://bioinformaticshome.org/		NO							To add							
24	ExoCNVTest !!!!!	an exome sequencing analysis tool for CNV detection	https://www.ncbi.nlm.nih.gov		NO							To add							
25	exomeCopy	exomeCopy R package is for CNV detection	https://bioinformaticshome.org/	1.30.0	Yes	https://anaconda.org/search?q=exomecopy						To add							
26	ExomeCNV	a tool for detection of copy number variants	https://bioinformaticshome.org/	1.4	NO							To add							
27	ExomeDepth	A tool for calling copy number variants	https://github.com/vplagnol	1.1.11	Yes	https://anaconda.org/search?q=exomedepth	https://toolshed.g2.bx.psu.edu/shed/module/crs4	crs4	1.1.0			To update							
28	Genovar	Genovar is a tool for detecting CNVs	https://bioinformaticshome.org/	0.951b	NO							To add							
29	HMZDelFinder	HMZDelFinder is a tool for CNV detection	https://bioinformaticshome.org/	3.2.1	NO							To add							
30	ichorCNA	ichorCNA is an R tool for CNV detection	https://bioinformaticshome.org/	0.1.0	Yes	https://anaconda.org/search?q=ichorcna						To add							
31	iCNV	A tool for detecting copy number variants	https://bioinformaticshome.org/	1.4.0	Yes	https://anaconda.org/search?q=icnv						To add	khaled'						
32	MATCHCLIP	A legacy tool for detecting CNVs	https://bioinformaticshome.org/		NO							To add							
33	matchclips2	An updated version of matchclips	https://bioinformaticshome.org/		NO							To add							

<https://github.com/users/khaled196/projects/1/views/1>





Training material

Galaxy Training! Variant Analysis Languages Help Extras Search Tutorials

Somatic Variant Discovery from WES Data Using Control-FREEC

Authors: Khaled Jum'ah Katarzyna Murat Wolfgang Maier David Salgado Krzysztof Poterlowicz

Overview

Questions:

- What are the specific challenges in locating human Copy Number Variations (hCNVs)?
- How to preprocess the sequenced reads for hCNVs detection?
- How can you detect the hCNVs in/from tumor and normal tissue of the same individual?
- How can you visualise the hCNVs' findings and compare them for specific regions?

Objectives:

- Use Control-FreeC for hCNV Identification in tumor tissue.
- Visualise the detected hCNVs in specific chromosomes.

Requirements:

- Introduction to Galaxy Analyses
- Sequence analysis
 - Quality Control: slides - hands-on
 - Mapping: slides - hands-on

Time estimation: 3 hours

Supporting Materials:

Datasets Workflows FAQs Available on these Galaxies

Last modification: Jan 12, 2023

License: Tutorial Content is licensed under Creative Commons Attribution 4.0 International License The GTN Framework is licensed under MIT

Introduction
Data Preparation
Quality control and mapping of NGS reads
Copy Number Variation detection (hCNV).
Conclusion
Frequently Asked Questions
References
Feedback

Introduction

Human Copy Number Variations (hCNVs) are the result of structural genomic rearrangements that result in the duplication or deletion of DNA segments. These changes contribute significantly to human genetic variability, diseases, and somatic genome variations in cancer and other diseases [Nam et al. 2015](#). hCNVs can be routinely investigated by genomic hybridisation and sequencing technologies. There is a range of software tools that can be used to identify and quantify hCNVs. Unfortunately, locating hCNVs is still a challenge in standardising formats for data representation and exchange. Furthermore, the sensitivity, specificity, reproducibility, and reusability of hCNV detection and analysis research software varies. As a result, there is a need for the adoption of community-developed standards for data discovery and exchange. To address this, ELIXIR developed [Beacon](#) protocol and [for Genomics and Health](#) standards, as well as mechanisms for annotating, benchmarking, creating reproducible and sharable tools and workflows, such as [WorkflowHub](#), Galaxy, and [ELIXIR](#), and, most importantly, accessible training resources and infrastructure.



<https://training.galaxyproject.org/training-material/topics/variant-analysis/tutorials/somatic-variant-discovery/tutorial.html>



Current achievements IS2 Beacon & Beyond

- Implementation of CNV resources over Beacon v2
 - Progenetix as reference implementation
 - Beacon v2 queries & model
 - handovers for asynchronous data retrieval => blueprint for sandboxing
 - open source, open access
 - BANCCO as example for rare disease data in protected context with input for query model development



Assembly: GRCh38 Chro: 9 Start: 21500001-21975098 End: 21967753-22500000
Type: EFO:0030067 Filters: NCIT:C3058

progenetix

Matched Samples: 660 Retrieved Samples: 660 Variants: 279 Calls: 667

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

Visualization options

Results Biosamples Biosamples Map Variants Annotated Variants

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

75% 50% 25% 0% -25% -50% -75%

progenetix: 662 samples CC BY 4.0 progenetix.org (2022)

Matched Subset Codes	Subset Samples	Matched Samples	Subset Match Frequencies
UBERON:0002021	4	1	0.250
pgx:icdot-C71.4	4	1	0.250
pgx:icdom-94403	4286	656	0.153
NCIT:C3058	4370	656	0.150
UBERON:0016525	14	2	0.143
pgx:icdot-C71.1	14	2	0.143
UBERON:0000955	7199	643	0.089
pgx:icdot-C71.9	7204	643	0.089
pgx:icdom-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

Download Sample Data (TSV)
1-660 [\[link\]](#)

Download Sample Data (JSON)
1-660 [\[link\]](#)

Download Sample Variants (JSON)
1-660 [\[link\]](#)

Network traffic analysis in the background:

Name	Do...	T Transf...	T...
biosamples	pro...	5.14 KB	2...
biosamples	pro...	52.60...	1...
genomicVariations	pro...	25.99...	1...
genomicVariations	pro...	3.98 KB	8...
samplePlots.cgi	pro...	26.13 ...	2...
collations	pro...	199.4...	1...

1 6 7.04 MB 313.3 KB 0 116ms

Auto — Page 1

Current achievements IS2 Beacon & Beyond

- Dissemination and documentation
 - Beacon v2 reference documentation (docs.genomebeacons.org)
 - detailed documentation of CNV specific query options
- Community engagement
 - CNV resource & Beacon demonstration at various ELIXIR & external events
- Demonstration of Beacon integration w/ Galaxy

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

Genomic Queries

Search

beacon-v2 ☆ 12 9

elixir

Beacon Range Queries ↗

Beacon Range Queries are supposed to return matches of any variant with at least partial overlap of the sequence range specified by `reference_name`, `start` and `end` parameters.

Beacon Range Query
Matching variants in a region

Beacon v2 Documentation

Beacon Types

Beacon Flavours

Security

Networks

Implement & Deploy a Beacon

Implementations Options

Standards Integration

Filters

Beacon Implementations >

Beacon Components

Framework

Models >

Querying Beacons

REST API

Genomic Queries

- Beacon Sequence Queries
- Beacon Range Queries**
- Beacon Geneld Queries
- Beacon Bracket Queries
- Genomic Allele Query (Short Form)
- Aminoacid Change Query
- variantType Parameter Interpretation

⚠ Use of `start` and `end`

Bracket queries require the use of **two** `start` and `end` parameters, in contrast to *Range Queries*.

Example: CNV Query - TP53 Deletion Query by Coordinates

The following example shows a "bracket query" for focal deletions of the *TP53* gene locus:

- The start of the deletion has to occur anywhere from approx. 2.5Mb 5' of the CDR start to just before the end of the CDR.
- The end of the matched CNVs has to be anywhere from the start of the gene locus to approx. 2.5Mb 3' of its end.

This leads to matching of deletion CNVs which have at least some base overlap with the gene locus but are not larger than approx. 5Mb (operational definitions of focality vary between 1 and 5Mb).

Beacon v2 GET Beacon v2 POST Beacon v1 Beacon v0.3

?datasetIds=TEST&referenceName=NC_000017.11&variantType=DEL&start=5000000,7676592&end=7669607,10000000



Current achievements IS2 Beacon & Beyond

- Implementation-driven standards development
 - cross-alignment of genomic annotation standards through participation in different projects (GA4GH VRS, Beacon, VCF)
 - communities w/ different practices contribute to standards as needed for their own resources (RD, cancer, genetics vs. genomics...)
 - implementation in reference projects allows to challenge and evolve standards (e.g. VCF4.4)

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

This table corresponds to the [Beacon v2 documentation](#).

Beacon	VCF	SO	EFO	VRS	Notes
DUP ¹ or EFO:0030070	DUP SVCLAIM=D ²	SO:0001742 copy_number_gain	EFO:0030070 copy number gain	low-level gain (implicit)	a sequence alteration whereby the copy number of a given genomic region is greater than the reference sequence
DUP ¹ or EFO:0030071	DUP SVCLAIM=D ²	SO:0001742 copy_number_gain	EFO:0030071 low-level copy number gain	low-level gain	
DUP ¹ or EFO:0030072	DUP SVCLAIM=D ²	SO:0001742 copy_number_gain	EFO:0030072 high-level copy number gain	high-level gain	commonly but not consistently used for >=5 copies on a bi-allelic genome region
DUP ¹ or EFO:0030073	DUP SVCLAIM=D ²	SO:0001742 copy_number_gain	EFO:0030073 focal genome amplification	high-level gain	commonly but not consistently used for >=5 copies on a bi-allelic genome region, of limited size (operationally max. 1-5Mb)
DEL ¹ or EFO:0030067	DEL SVCLAIM=D ²	SO:0001743 copy_number_loss	EFO:0030067 copy number loss	partial loss (implicit)	a sequence alteration whereby the copy number of a given genomic region is smaller than the reference sequence
DEL ¹ or EFO:0030068	DEL SVCLAIM=D ²	SO:0001743 copy_number_loss	EFO:0030068 low-level copy number loss	partial loss	
DEL ¹ or EFO:0030069	DEL SVCLAIM=D ²	SO:0001743 copy_number_loss	EFO:0030069 complete genomic deletion	complete loss	complete genomic deletion (e.g. homozygous deletion on a bi-allelic genome region)



Community evaluation SWOT analysis





INTERNAL

HELPFUL

Community of experts from multiple countries
Strong support from ELIXIR Hub
Members are key contributors of ELIXIR, GA4GH
Various skills (Data, DB, Interpretation)
All fields are covered (RD, Onco, CD)
Interactions with ELIXIR platforms
Implementation studies (x3)

S

EXTERNAL

HARMFUL

Only a subset of partners are really actives
Difficulties to interact with platforms
Decreased interest over time
Some topics may not be addressed properly
Lack of funding for dedicated positions

W

Links to international initiatives and standards development (VCF, Beacon)
Interactions with other communities
National projects (Diagnosis, Research)
Reference materials & guidelines
Communication at international meetings
Industry

O

T

Other international initiatives
Limited adoption of recommendations
Difficulties to share human data
Lack of funding
Administrative difficulties in ELIXIR-FR
(*recruitments & reference dataset sequencing*)

hCNV Community Looking Forward

- engaging / onboarding external partners through participation in leading edge, standard definition / implementation projects
- identification of lead issue for (well supported) ELIXIR project
 - community identified need for resource development
 - ➡ germline CNV profiling catalogue, including many resources with proper technological and population provenance and fully using advanced standards (Beacon v2, VCF4.4, GA4GH VRS...)
 - seen as initiator for ELIXIR reference resource & **integrator** of different initiatives
- wish for high-level cross-community stewardship, as well as top-down requests for participation
 - critical review of hCNV intersections w/ other communities & "forced engagement"
 - including hCNV in planning & execution of large-scale genomics infrastructure projects





Thanks to all hCNV community members and beyond



www.elixir-europe.org