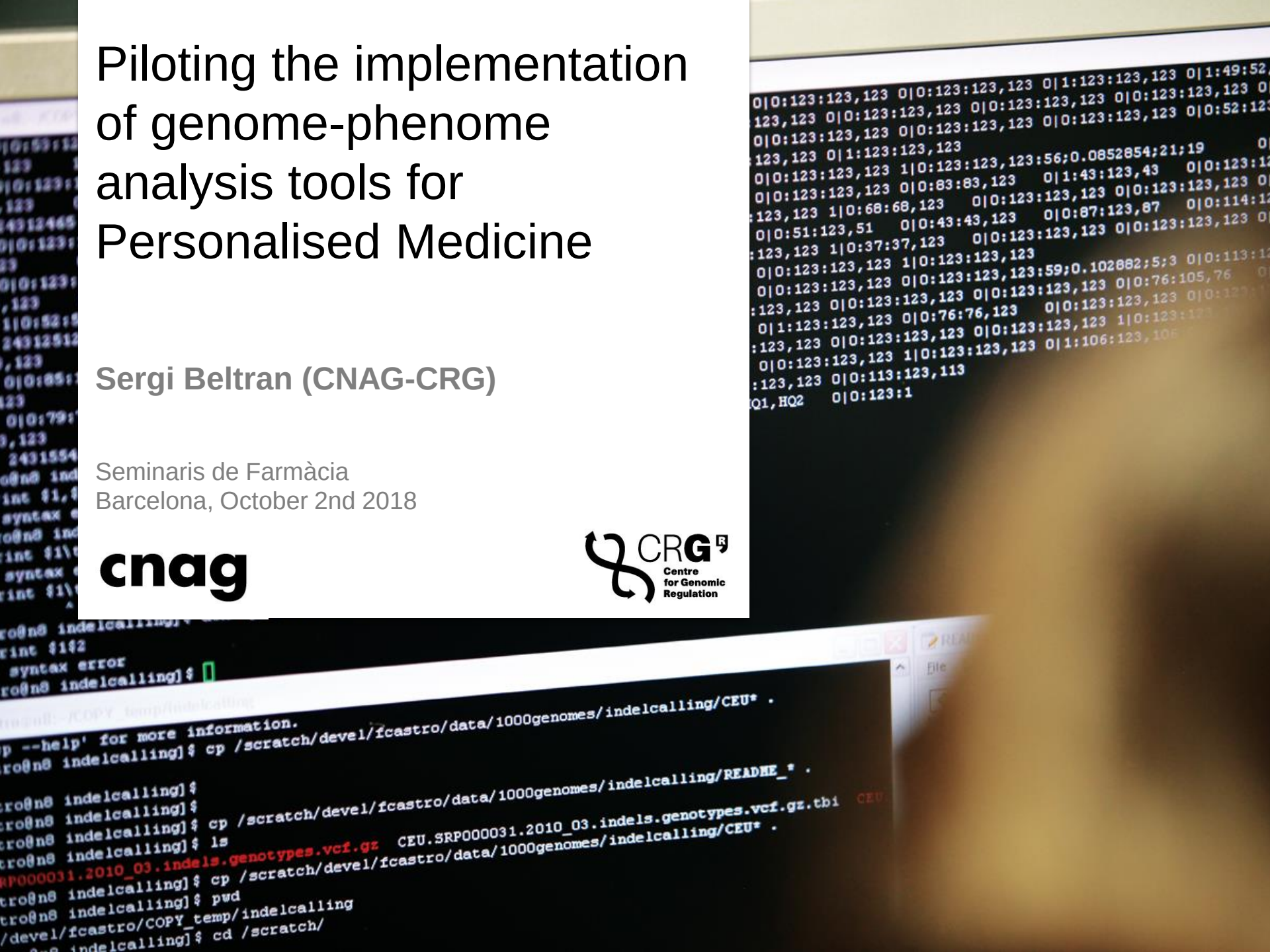
Piloting the implementation of genome-phenome analysis tools for Personalised Medicine

Sergi Beltran (CNAG-CRG)

Seminaris de Farmàcia
Barcelona, October 2nd 2018

cnag

CRG Centre for Genomic Regulation



CNAG Overview

The CNAG-CRG is a non-profit organization integrated in the CRG. It is funded by the Spanish Ministry of Economy and Competitiveness and the Catalan Government.

Our Mission

To carry out projects in genome analysis that will lead to significant improvements in people's health and quality of life, in collaboration with the Catalan, Spanish, European and International research and clinical community.

CNAG-CRG

Director: Ivo Gut

80 people, 60 bioinformaticians

ICTS – Infrastructure for OMICS Technologies (IOT)

The IOT includes the CNAG-CRG and the COS (Centre for Omics Sciences).

Infraestructuras Científicas y Técnicas Singulares

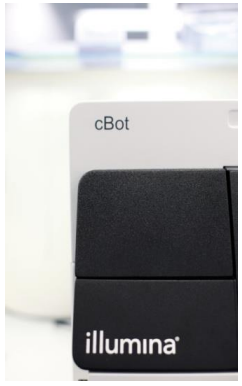
CNAG-CRG: Relevant partnerships

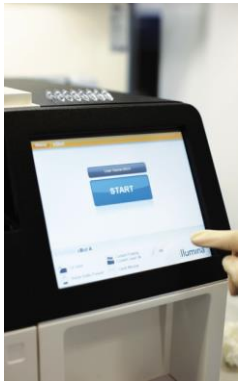
- ✓ Member of the **Global Alliance for Genomics and Health (GA4GH)**
- ✓ Participation, through the National Bioinformatics Institute (INB), in ELIXIR, the European bioinformatics infrastructure.
- ✓ Participation in the International Human Epigenome Consortium (IHEC)
- ✓ Participation in the International Cancer Genome Consortium (ICGC)
- ✓ Participation in the International Rare Diseases Research Consortium (IRDiRC)

The CNAG's Genomehenge 2018

Sequencing

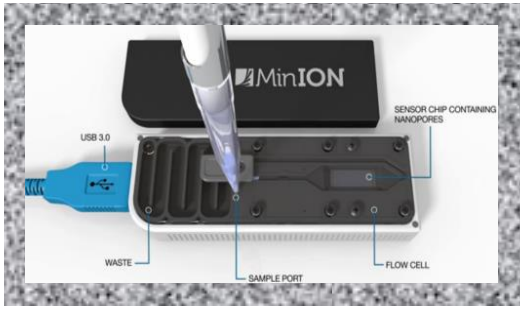
- 1 Illumina NovaSeq6000
- 2 Illumina HiSeq4000
- 4 Illumina HiSeq2000
- 1 Illumina MiSeq
- 3 Oxford Nanopores MinIons





Computing

- 3852 cores
- 3.7 PB disk + 3 PB tape
- 35,5 TB RAM
- Barcelona SuperComputing Center - 10 x 10 Gb/s



Sequencing Production

MiSeq

4.5x Coverage
2x300bp
50 hours

HiSeq2000

30x Coverage
2x100bp
11 days

HiSeq4000

30x Coverage
2x150bp
3 days

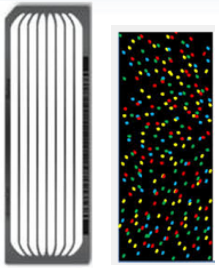
NovaSeq6000 (high-output run mode)

30x Coverage
2x150bp
44 hours

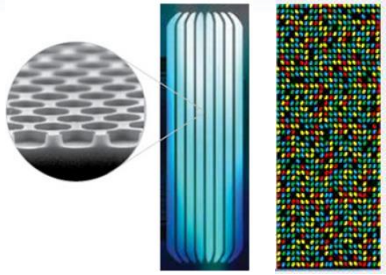
x24

x24

Non-Patterned Flow cell



Patterned Flow cell



illumina CSPro CERTIFIED SERVICE PROVIDER

CNAG's Workflow

BIOREPOSITORY

LABORATORY

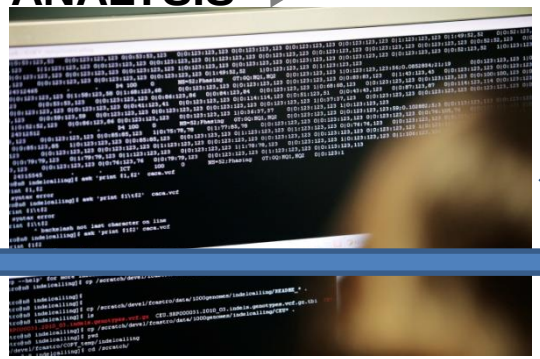
SEQUENCING

LIMS

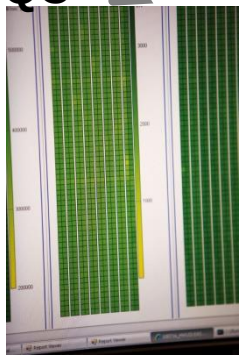
	Subproject	Status	Deadline (Countdown)	Details by Sample	Details by Library	Analysis	Sample Type	Application	Expected Capture Protocol	Numbers of Samples	Ready for Sample prep / Sequence	Samples	Sample Reception Dates	Aliquots	Libraries pass / total	Libraries Barcodes	Samples with a pass Library	Flowcells & Dates	Samples above Coverage	Samples above pfmr	Data Transfed by Sample
☐	_11	Open (Experimental)	2012-10-01 (26 days)	Details by Sample Sample Aggregate View	Details by Library Library Aggregate View	BAG	gDNA	ExomeCapture-Seq	Agilent Human All Exon 51Mb v4	6 8 8 6 4	yes	E061 E058 E055 E059 E056 E060 E057 E054	2012-07-02 2012-07-02	861C 948C 862C 863C 864C	3 / 4	861C 862C 863C 864C	3				0 sent, of 4 to be seq
☐	_01	Open (Experimental)	2012-10-02 (27 days)	Details by Sample Sample Aggregate View	Details by Library Library Aggregate View	BAG	gDNA	WG-Seq		2 4 4 4 2	yes	H367 H368 H368 H367	2012-07-03 2012-07-03	580C 581C	2 / 2	580C 581C	4	C0YHRACXX - 2012-07-27 C11CBACXX - 2012-08-09 C11D3ACXX - 2012-08-09			0 sent, of 2 to be seq
☐	_03	Open (Experimental)	2012-10-03 (28 days)	Details by Sample Sample Aggregate View	Details by Library Library Aggregate View	No	Exome library	ExomeCapture-Seq	Nimblegen SeqCap 64 Mb v3	16 80 80 0 80	yes	H349 H365 H346 H362 H343 H359 H340 H356 H337	2012-06-18 2012-07-04		80 / 80	I231 I229 I230 I227 I228 H342 H344 H346 H347	80	C0TRHACXX - 2012-06-19 C0TVGACXX - 2012-06-26 D113HACXX - 2012-06-26 D14PPACXX - 2012-07-25 C0YHRACXX - 2012-07-27	80		80 sent, of 80 to be seq

TRANSFER

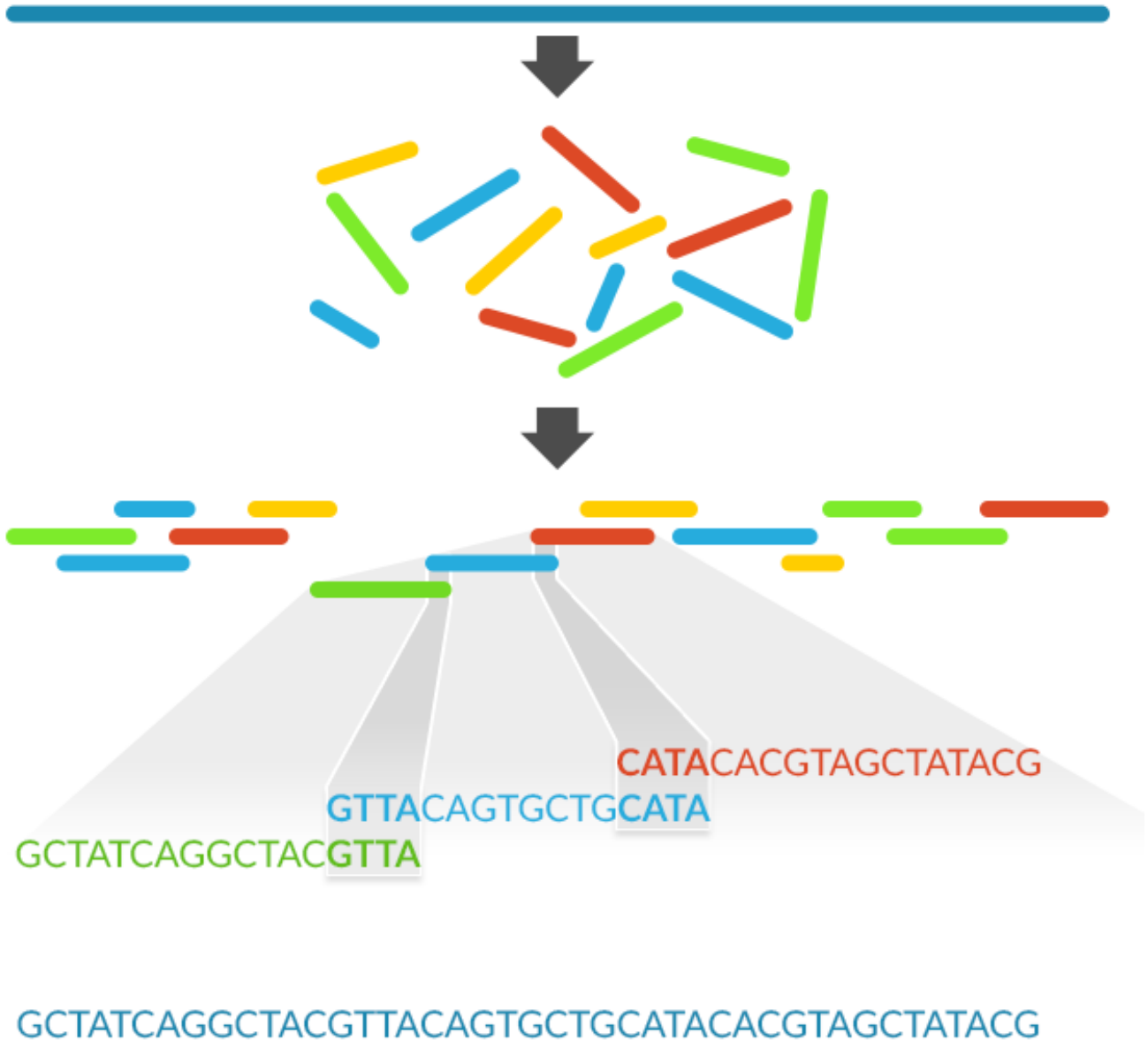
ANALYSIS



QC



Genome Sequencing



Adapted from http://www.micronautomata.com/bioinformatics

Lídia Àgueda
Head of Biorespository

Biological Resources → **Sequencing** → **Data analysis**

Marta Gut
Head of Sequencing

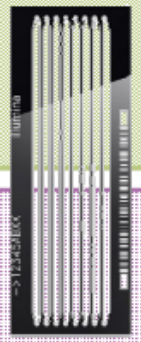
1 Library Preparation

Fragment DNA
Repair ends
Add A overhang
Ligate adapters
Purify

Julie Blanc,Sample Preparation manager

2 Cluster Generation

Hybridize to flow cell
Extend hybridized template
Perform bridge amplification
Prepare flow cell for sequencing



Katja Kahlem, Production manager

3 Sequencing

Perform sequencing
Generate base calls

1 Library Preparation

Fragment DNA
Repair ends
Add A overhang
Ligate adapters
Purify

Fragmentation

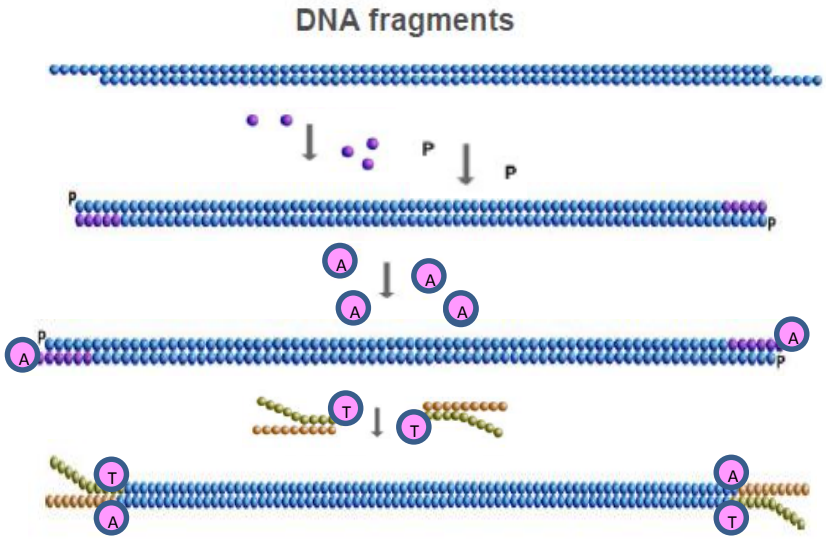
Repair Ends

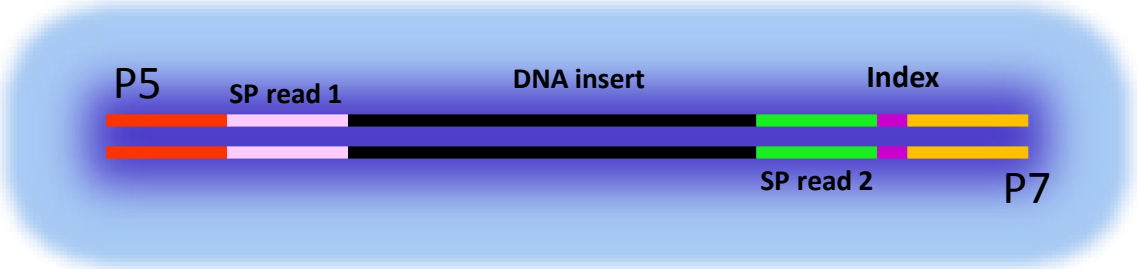
Add an “A” to the 3’ ends

Ligate Adapters

Size Selection

QC





1 Library Preparation

Fragment DNA
Repair ends
Add A overhang
Ligate adapters
Purify

WG _BS _SEq

mRNA Seq

Target capture

WG_Seq

…and many more

smallRNA_Seq

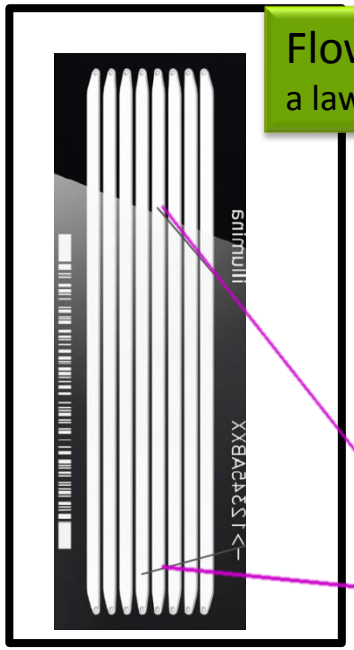
ChIP _Seq

2 Cluster Generation

Hybridize to flow cell
Extend hybridized template
Perform bridge amplification
Prepare flow cell for sequencing

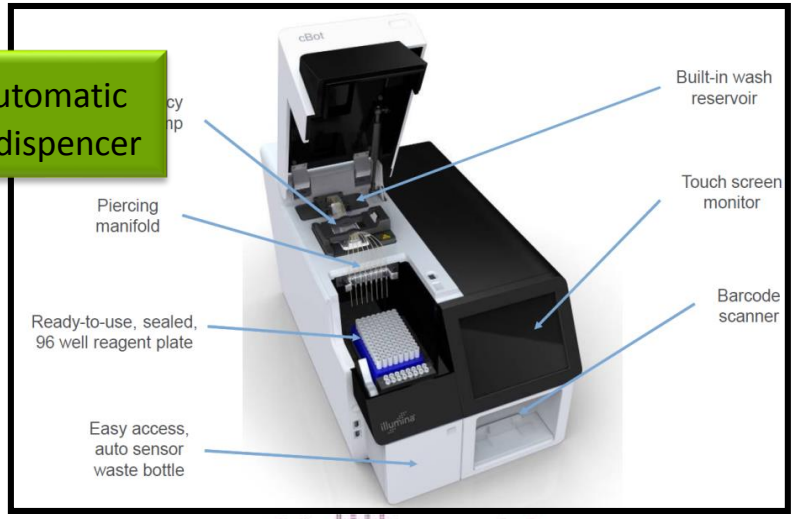
3 Sequencing

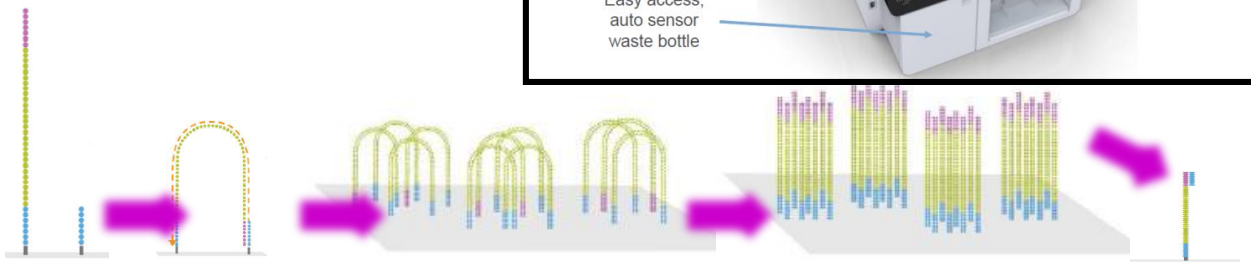
Perform sequencing
Generate base calls



Flow cell Glass slide with a lawn of oligonucleotides

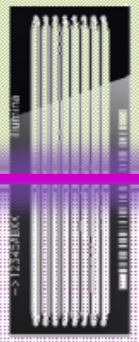
cBot Automatic reagent dispencer





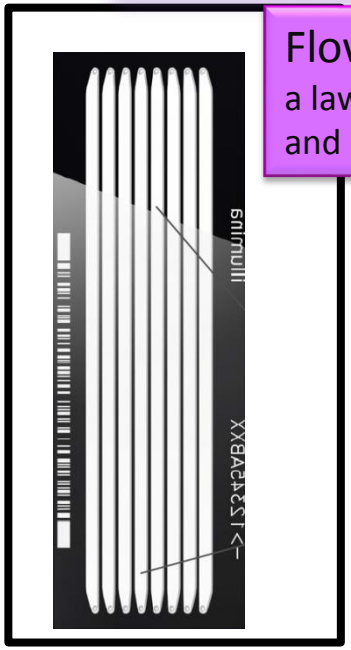
2 Cluster Generation

Hybridize to flow cell
Extend hybridized template
Perform bridge amplification
Prepare flow cell for sequencing



3 Sequencing

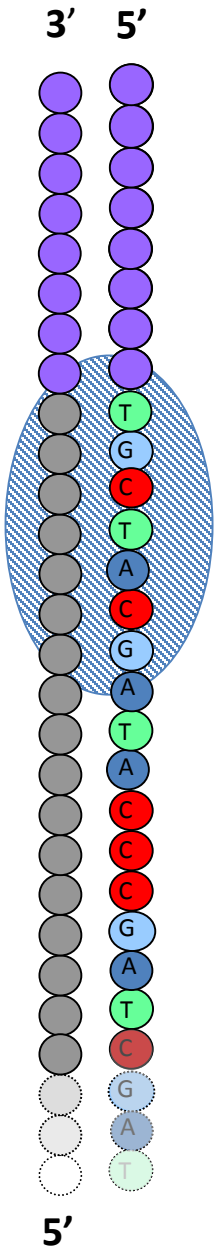
Perform sequencing
Generate base calls

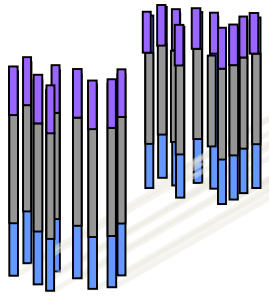


Flow cell Glass slide with a lawn of oligonucleotides and sequencing library

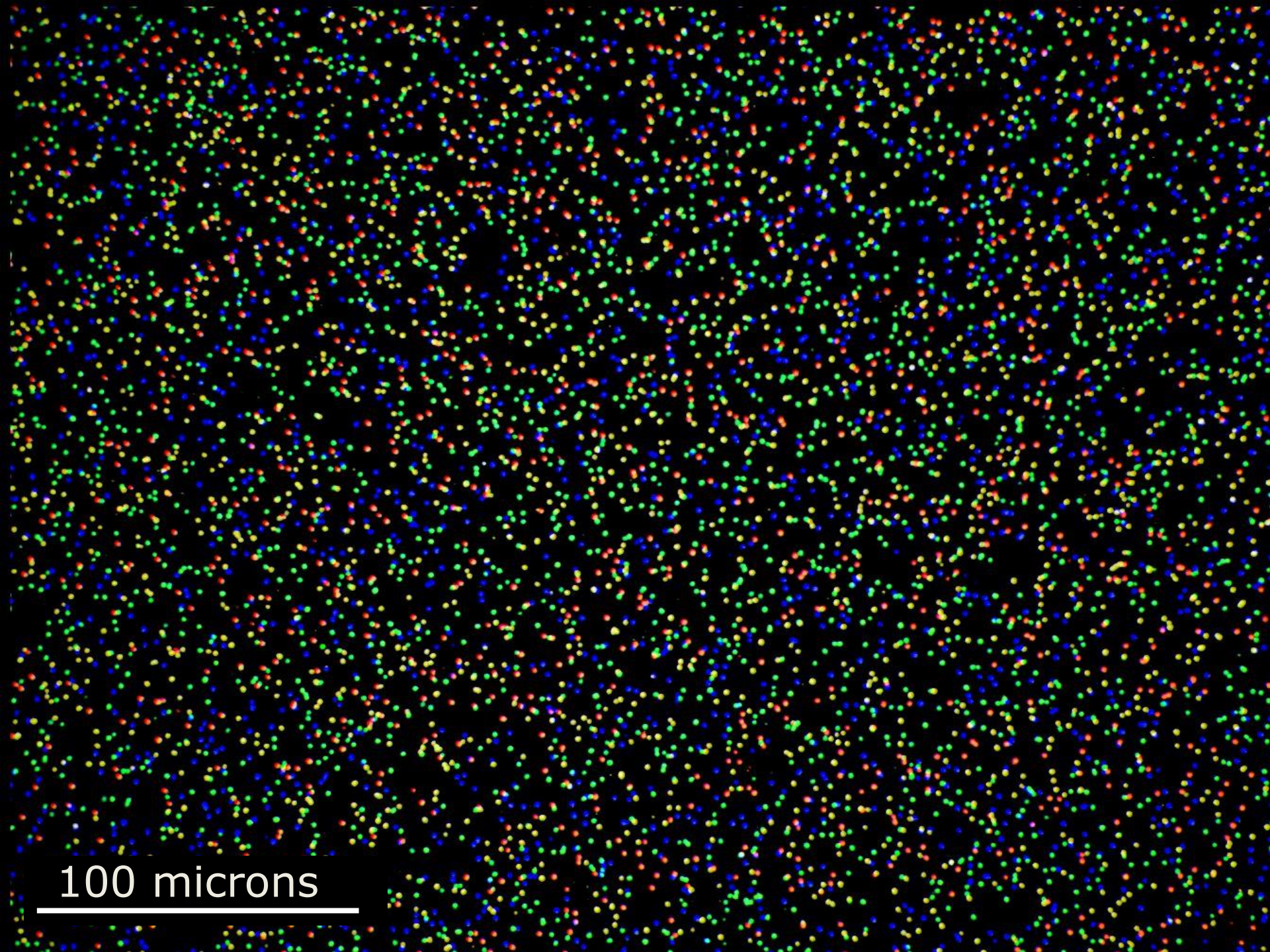
HiSeq2000 – the sequencer

Sequencing-by-synthesis (SBS)

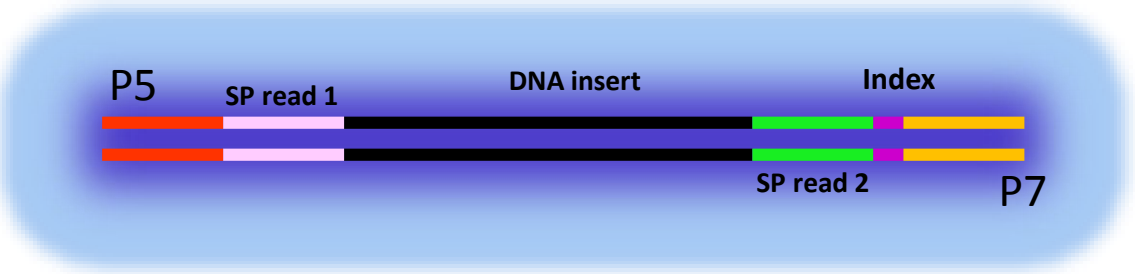


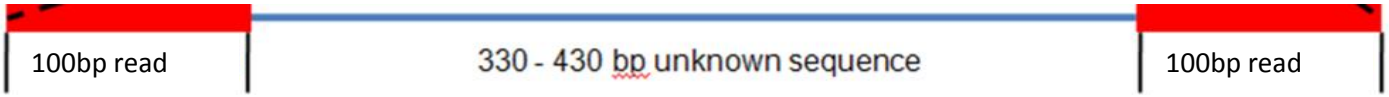


- All four labelled nucleotides in one reaction
- High accuracy
- Base-by-base sequencing
- No problems with homopolymer repeats


100 microns

Sequencing Output: FASTQ files

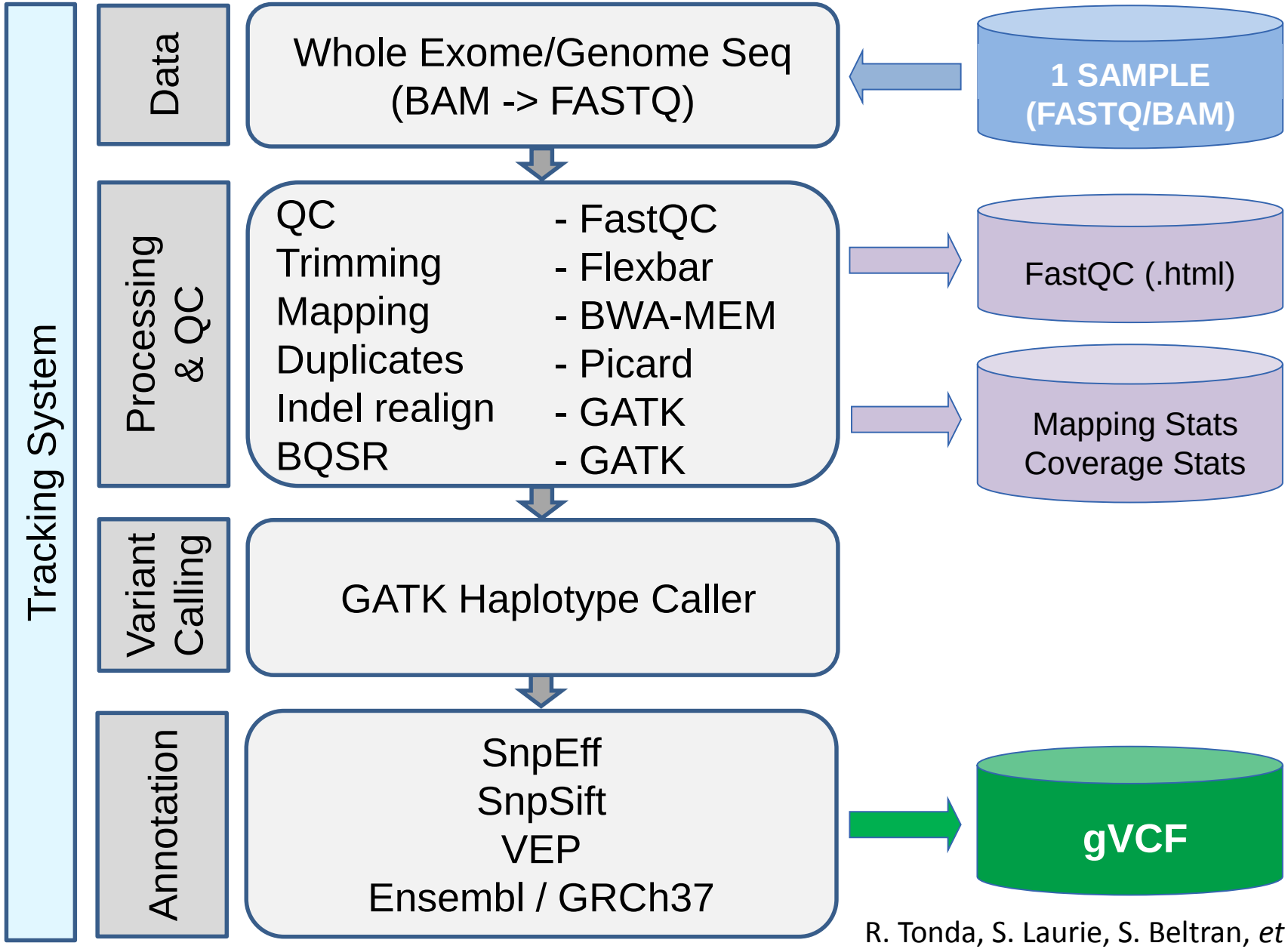





```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*(((***+))%%%++)(%%%%).1***-+*''))**55CCF>>>>>>CCCCCCC65
```

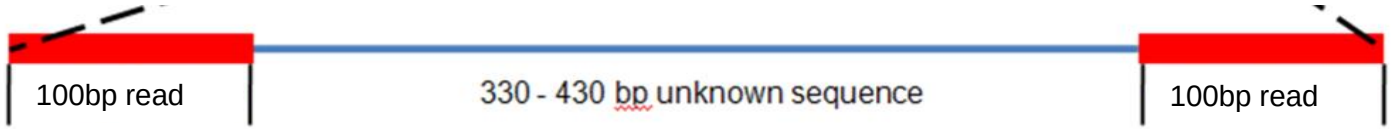
- Developed by the Wellcome Trust Sanger Institue
- Usually, each sequence (read) is split in 4 rows
- Sequence identifiers, description and quality encoding can be different

RD-Connect Genomics variant calling pipeline



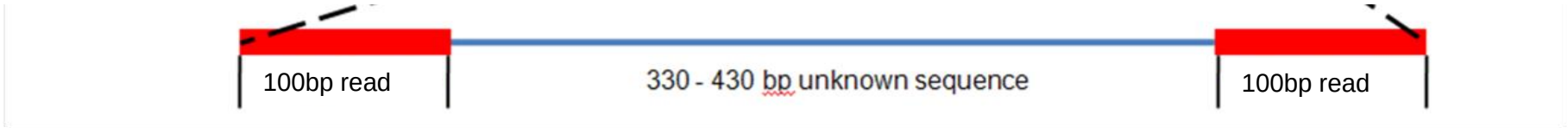
R. Tonda, S. Laurie, S. Beltran, *et al.*

Mapping to reference genome



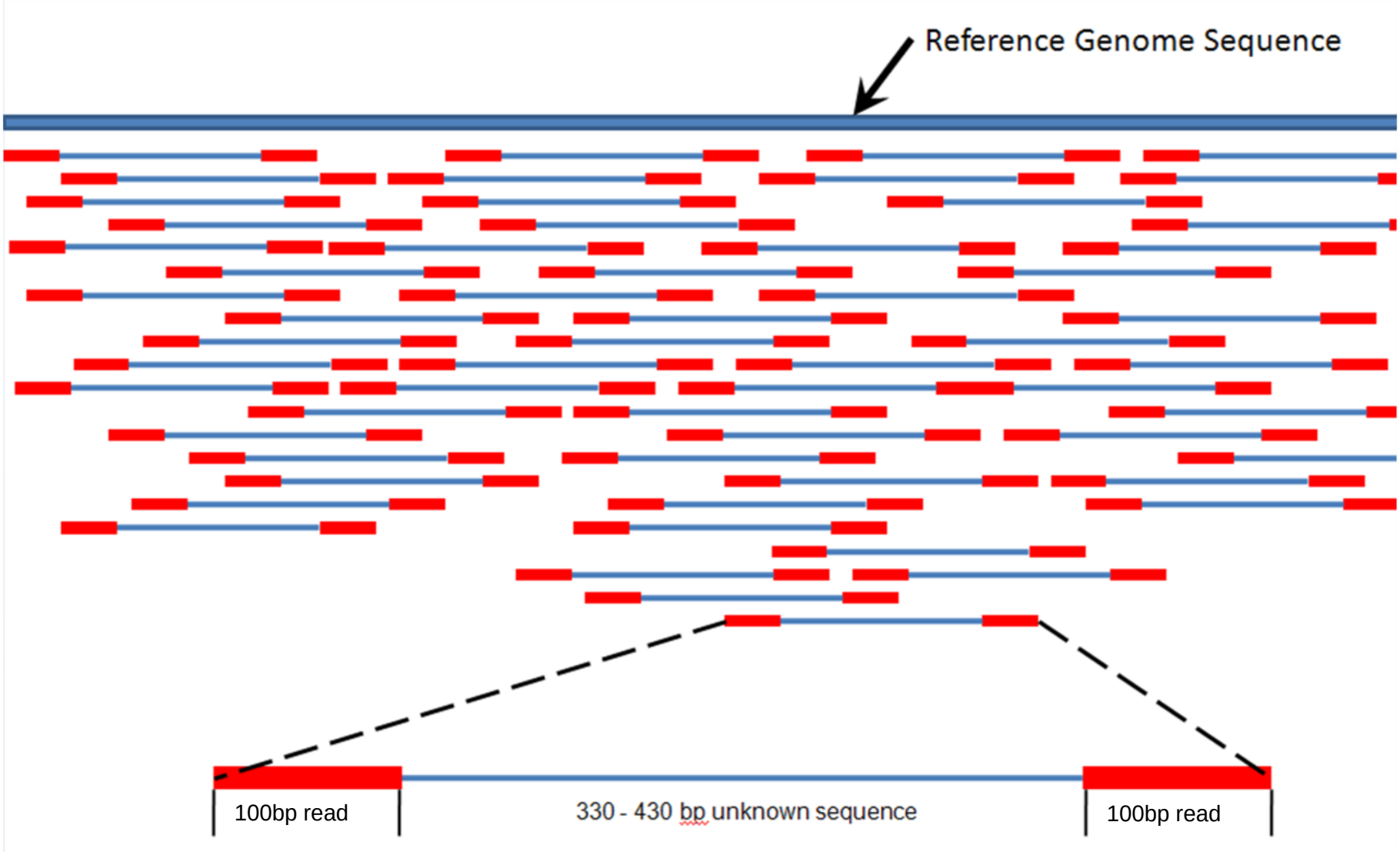
Adapted from wikipedia

Mapping to reference genome



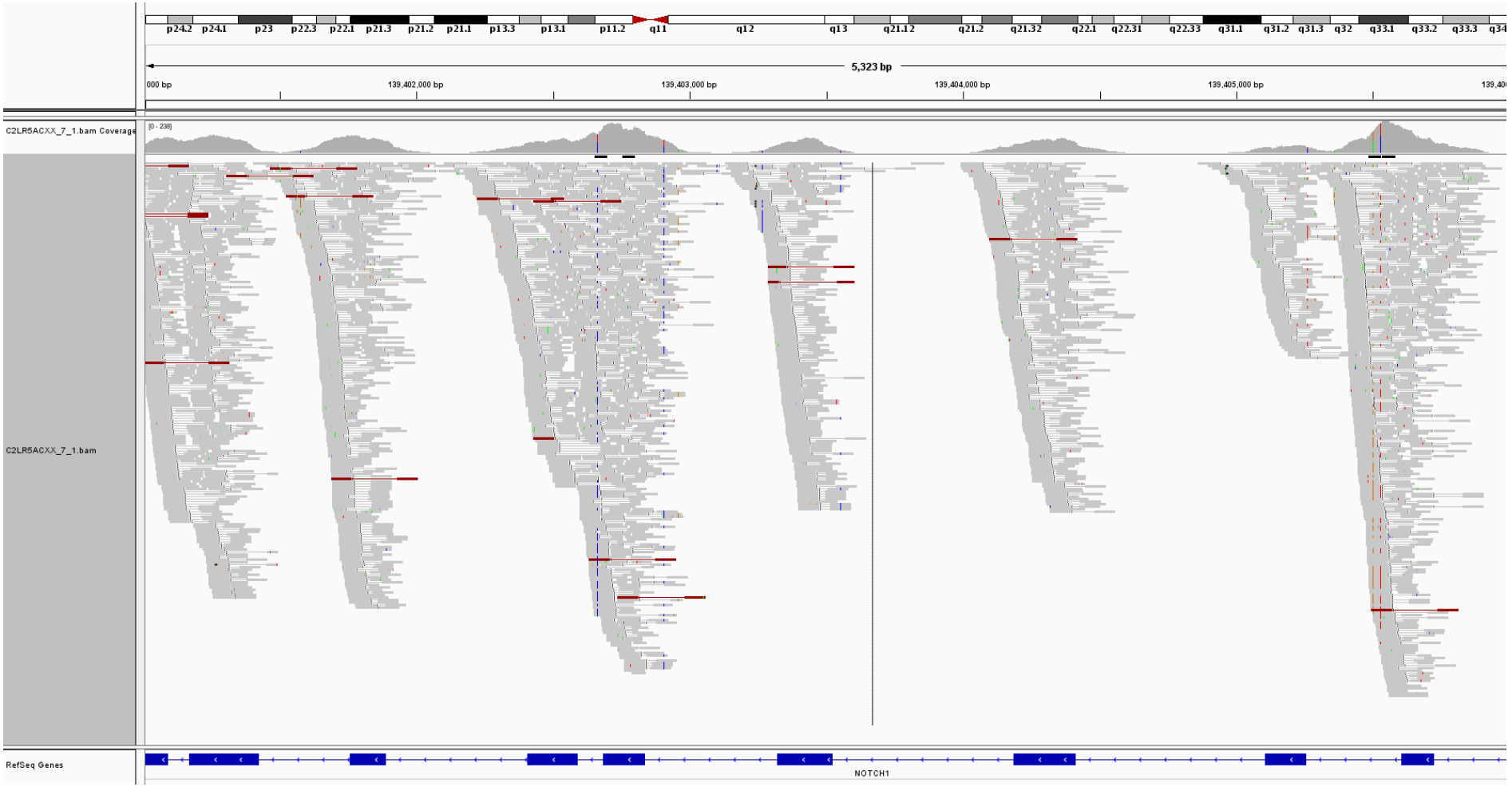
Adapted from wikipedia

Mapping to reference genome



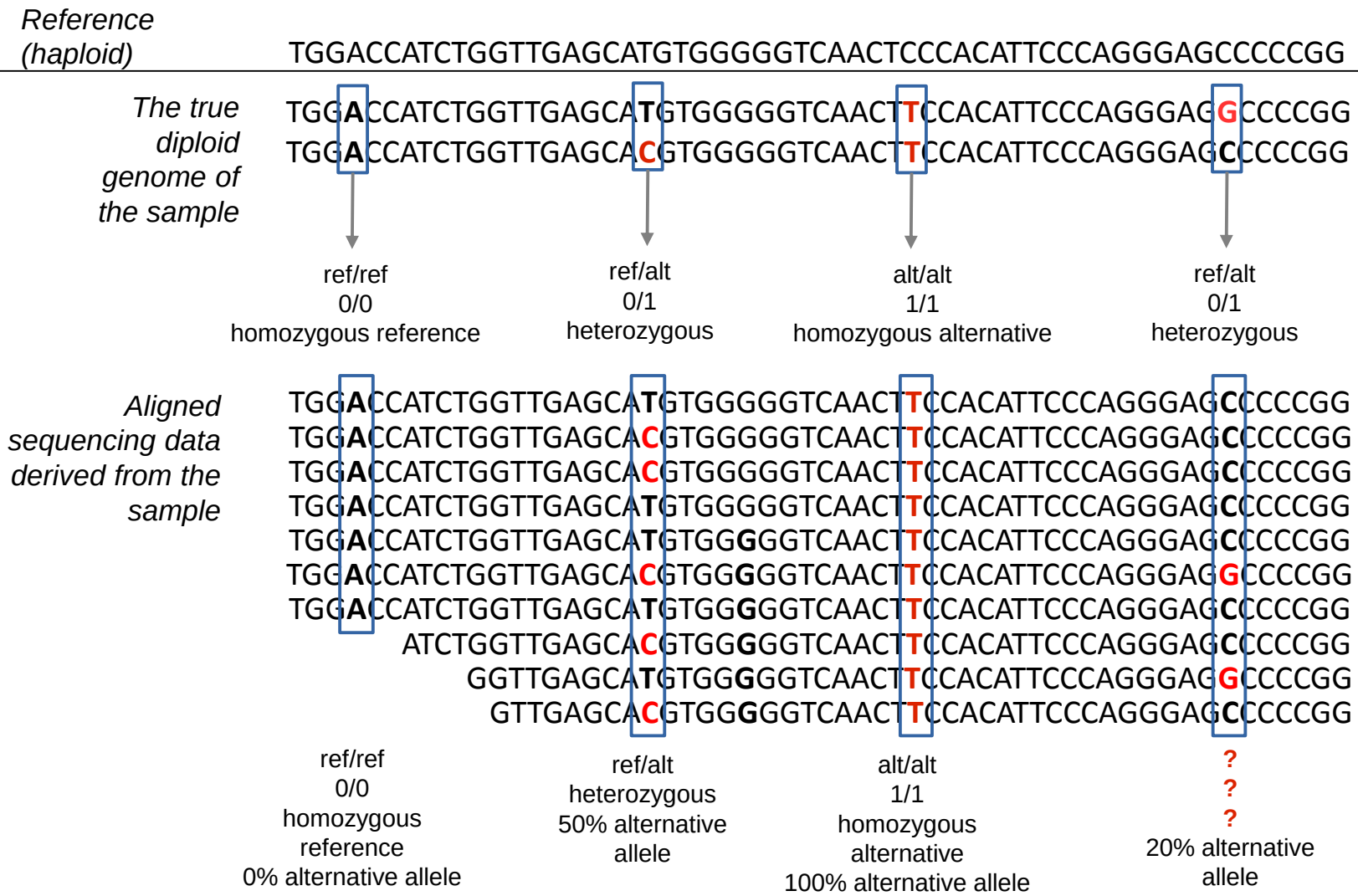
Adapted from wikipedia

Mapping: Exome sequence example

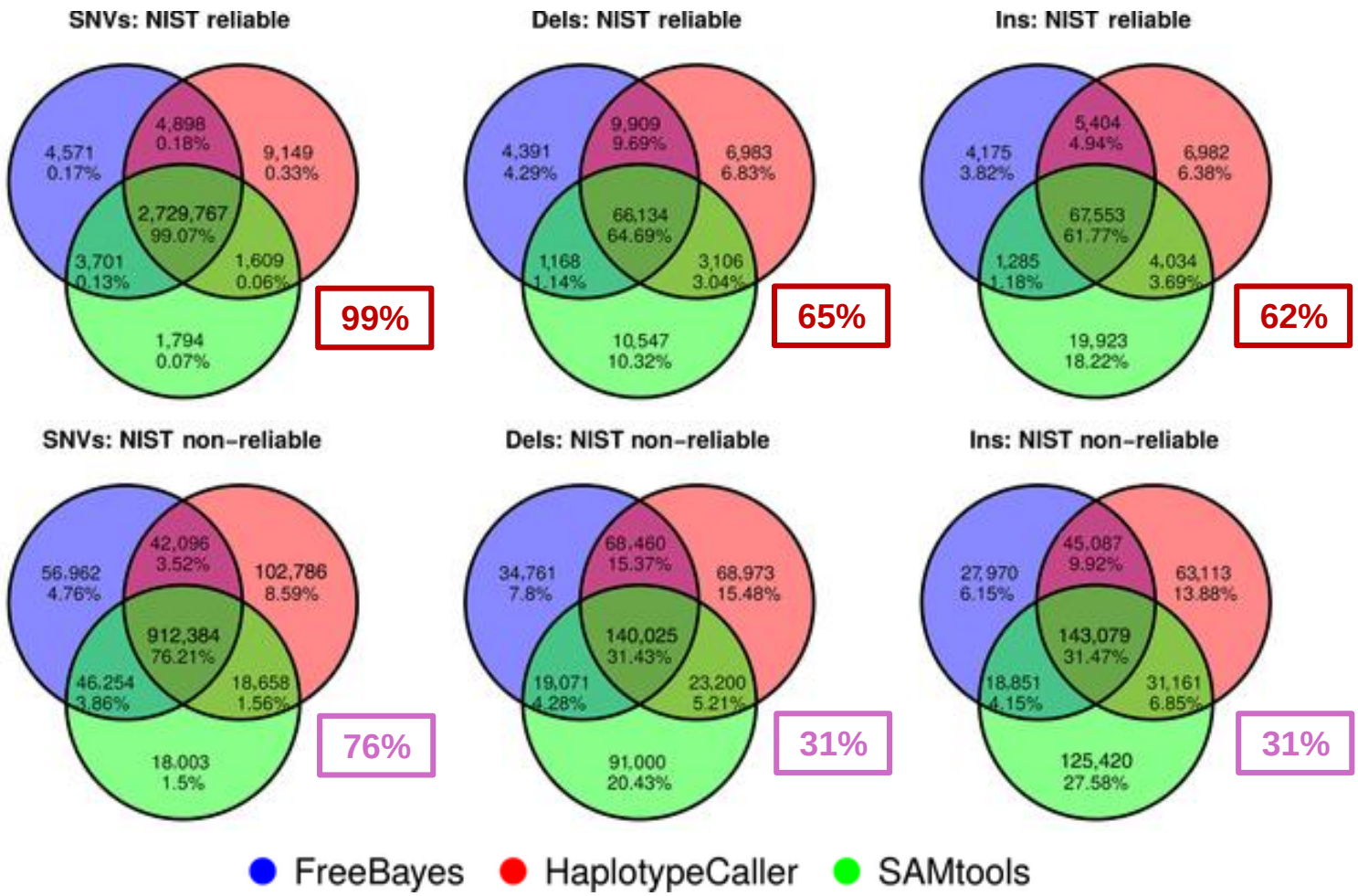


Alignments are stored in a **BAM file**, which is the binary version of **SAM** (Sequence/Alignment Map) format

Identification of genetic differences in comparison to a reference

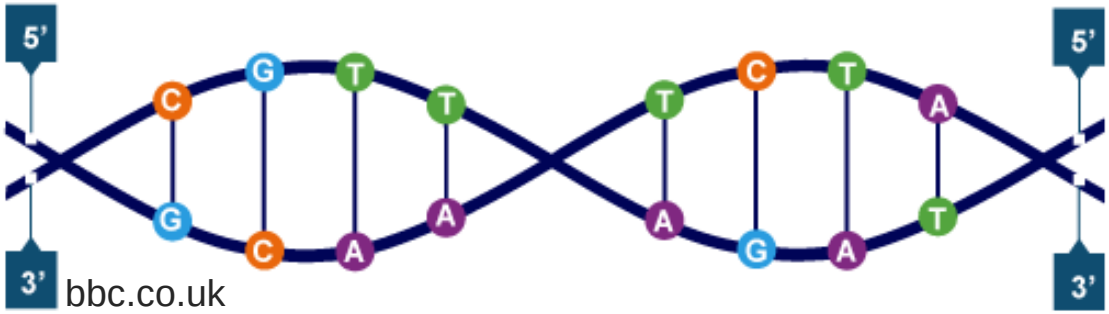


Benchmarking genomics pipelines



Laurie *et al*. Human Mutation, 2016

Rare diseases and genomics



bbc.co.uk

Human Genome:
3,200,000,000 bp
2 x 23 chromosomes

Each of us has
3.000.000 variants

Rare diseases and impact of reaching a molecular diagnosis

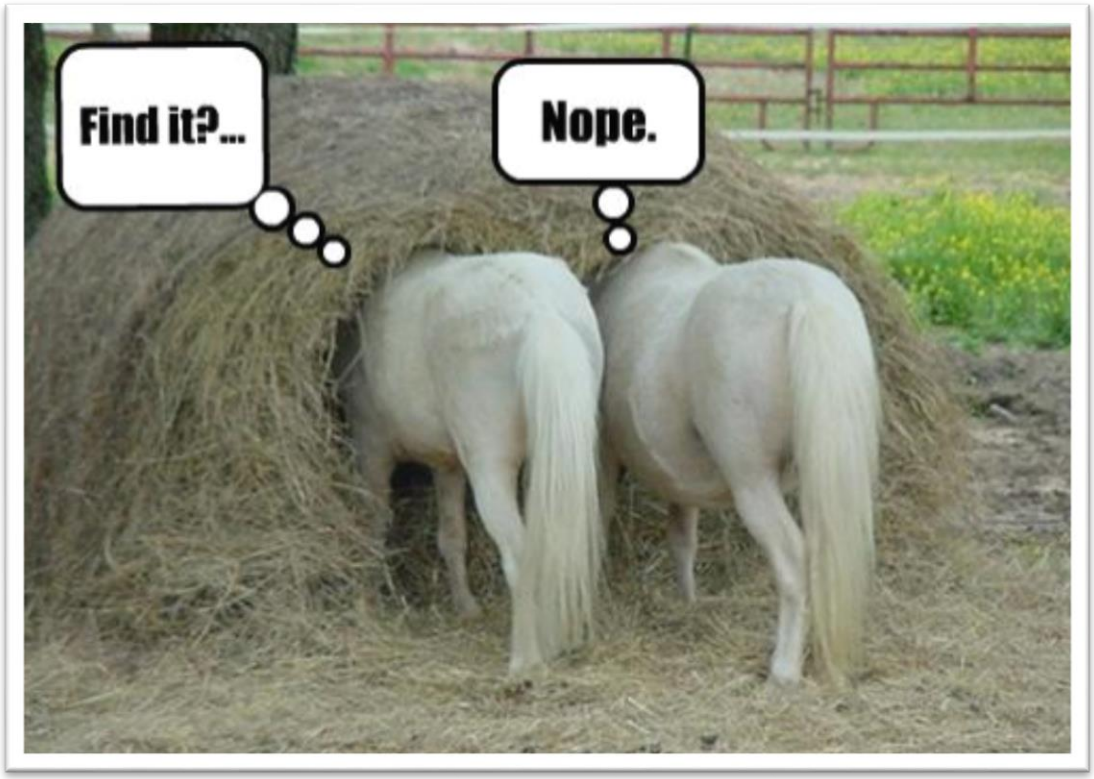
- **High phenotypic and genetic heterogeneity:**
  - **Challenge for diagnosis** mean average : 4.8 years.

- Importance of **reaching a molecular diagnosis**:

  - Treatment options
    - Curative (very few at this time)
    - Prevent progression: diet restriction metabolic disorders (e.g galactosemia)
    - Symptomatic and functional therapies
  - Progression and prognosis of the disease
  - Genetic counselling (prenatal / pre-implantational options)
  - Reduce patient and family anxiety

The diagnostic and gene discovery challenge

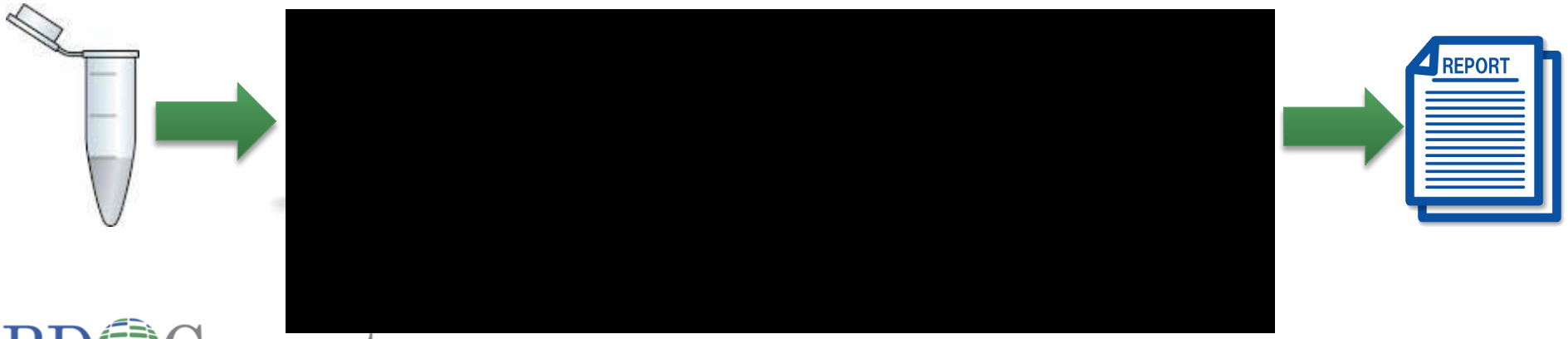
How do I find the pathogenic mutation within 3.000.000?



Interpretation is still difficult

Molecular diagnostics in NGS era

Sample in → Diagnosis out?



Interpretation is still difficult

Molecular diagnostics in NGS era

Sample in ~~> Diagnosis out?~~

RD-Connect Platform (platform.rd-connect.eu)

RD Connect

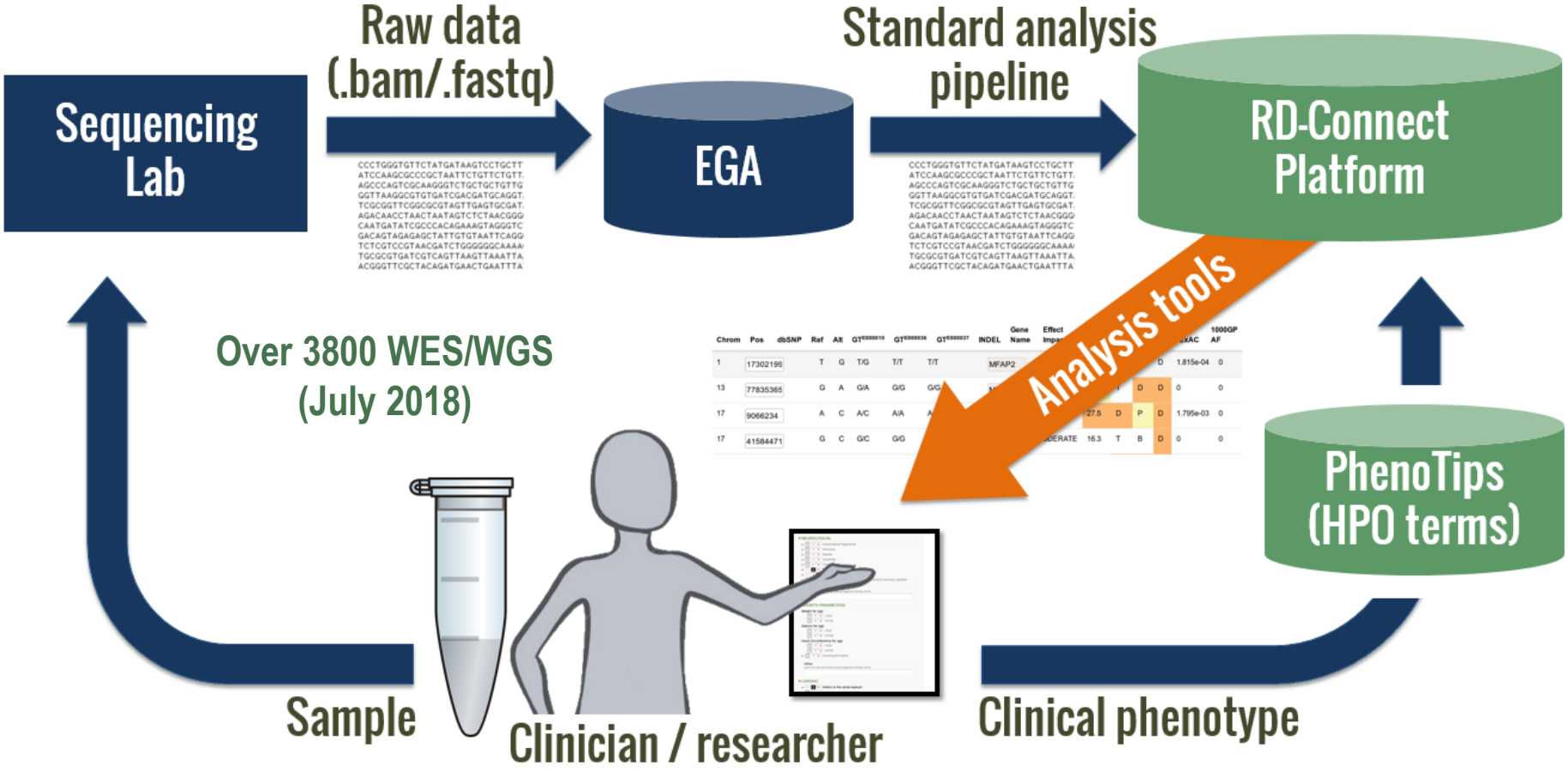
Home	Genome-Phenome Analysis Platform	PhenoTips	Registry and Biobank Finder	Biosample Catalogue

An integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research

Welcome to the central platform for access to data submitted by RD-Connect's partner projects. The online Genome-Phenome Analysis Platform is now open for submissions from all users. Our automated registration system will come online shortly, but if you would like to access the interface now please email platform@rd-connect.eu and we will contact you to request the information we need to set you up on the system.

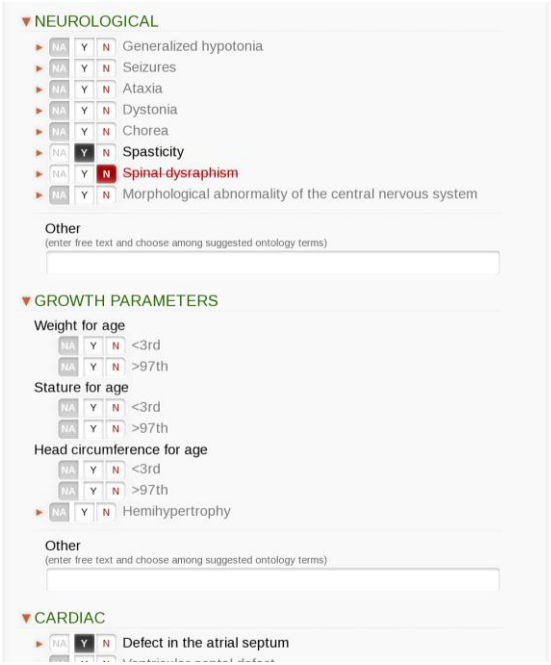
Get started today

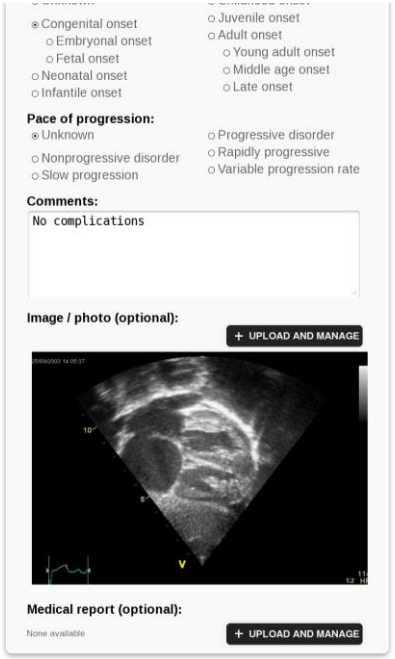
RD-Connect Genome-Phenome Analysis Platform (GPAP)



Ethics Committee Approval, GDPR Compliance, Code of Conduct, Data Access Committee, Security audited, User Activity logged

RD-Connect uses PhenoTips for collating phenotypic data



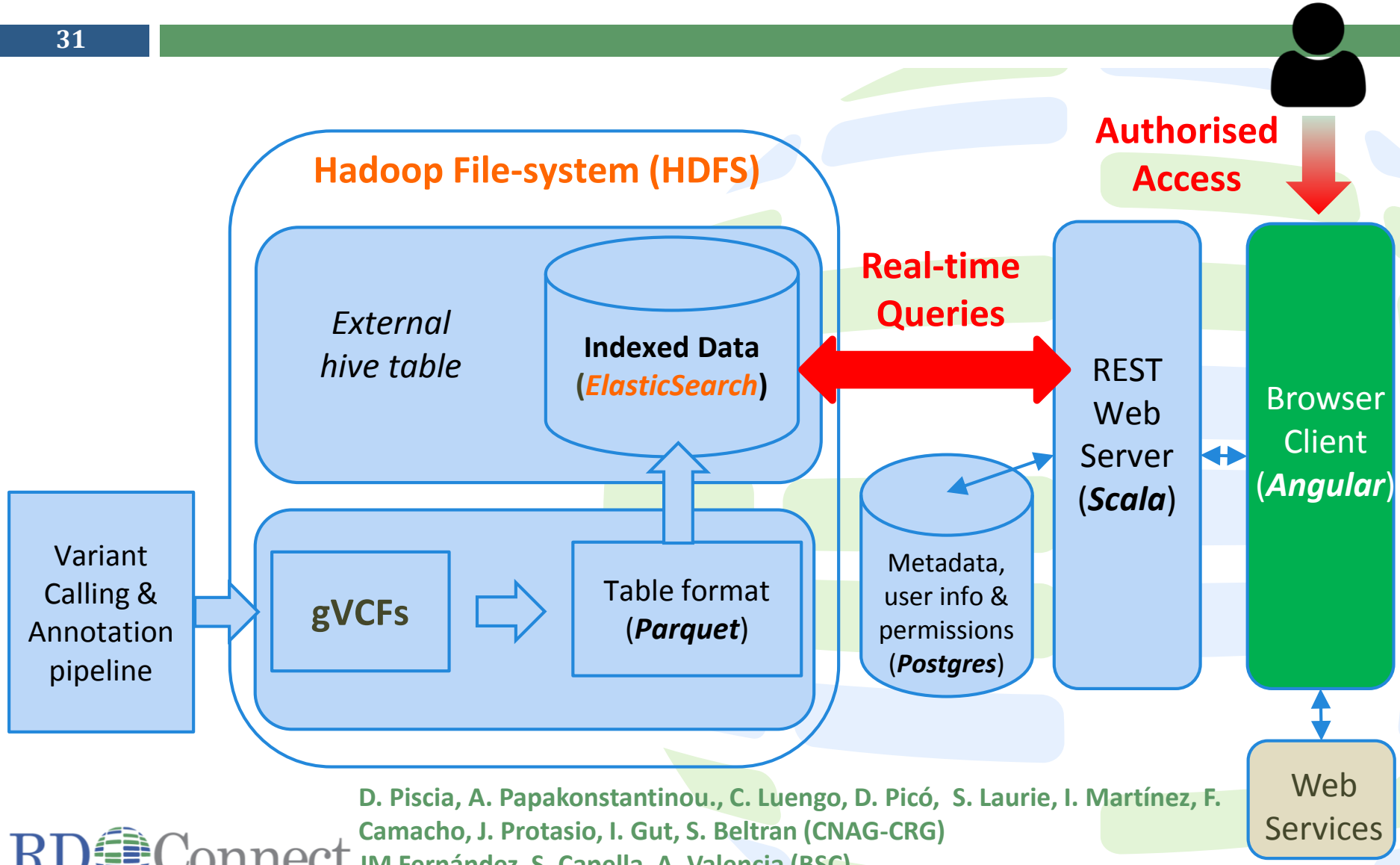


Deep phenotyping in PhenoTips (Brudno *et al*.) achieved using the Human Phenotype Ontology (HPO – Robinson, Köhler *et al*.)

Diseases classified using the Orphanet Rare Disease Ontology and OMIM identifiers

Information from PhenoTips can be sent directly to other tools within the platform (e.g. Exomiser, MME)

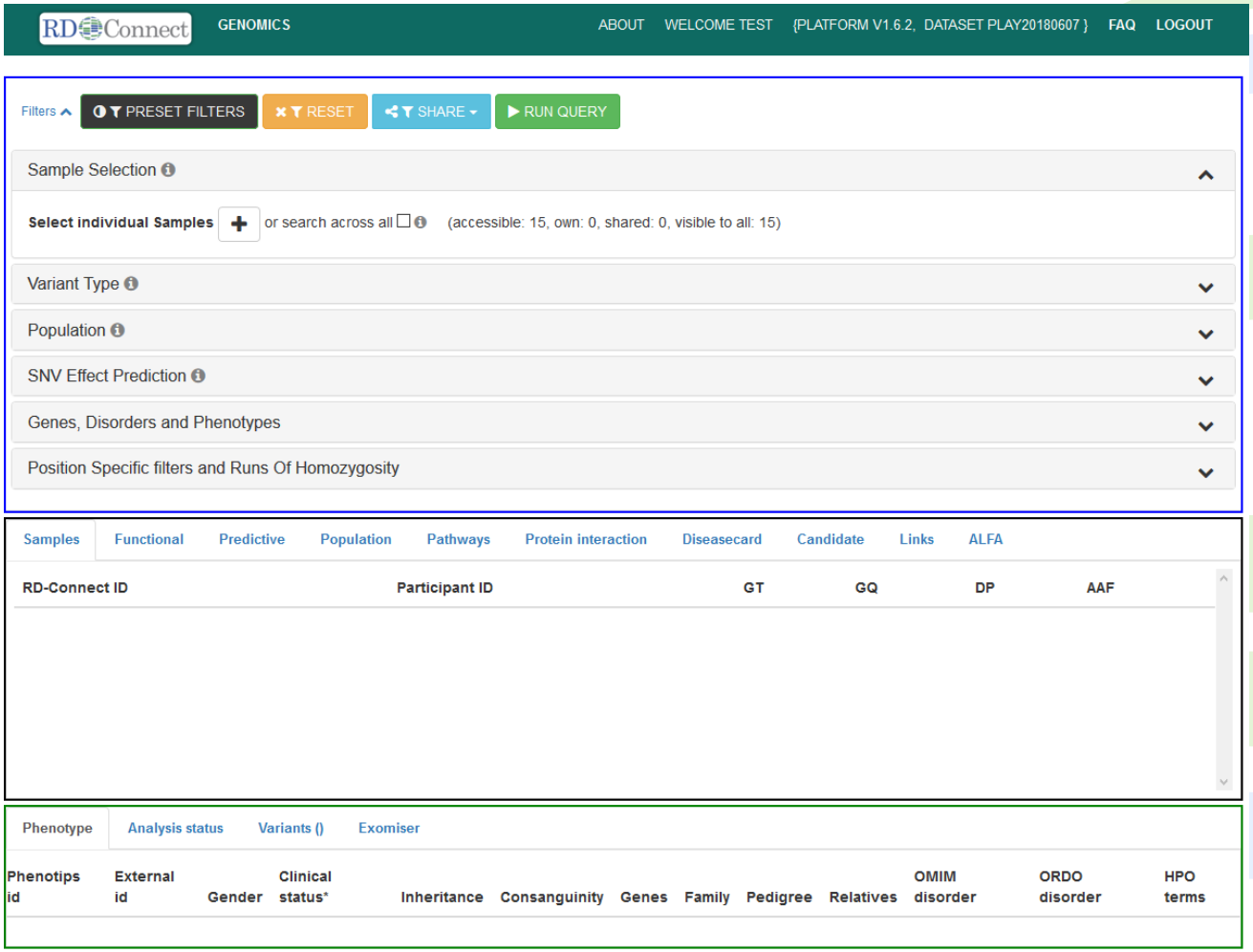
Big-data platform architecture



D. Piscia, A. Papakonstantinou., C. Luengo, D. Picó, S. Laurie, I. Martínez, F. Camacho, J. Protasio, I. Gut, S. Beltran (CNAG-CRG)
JM Fernández, S. Capella, A. Valencia (BSC)

RD Connect

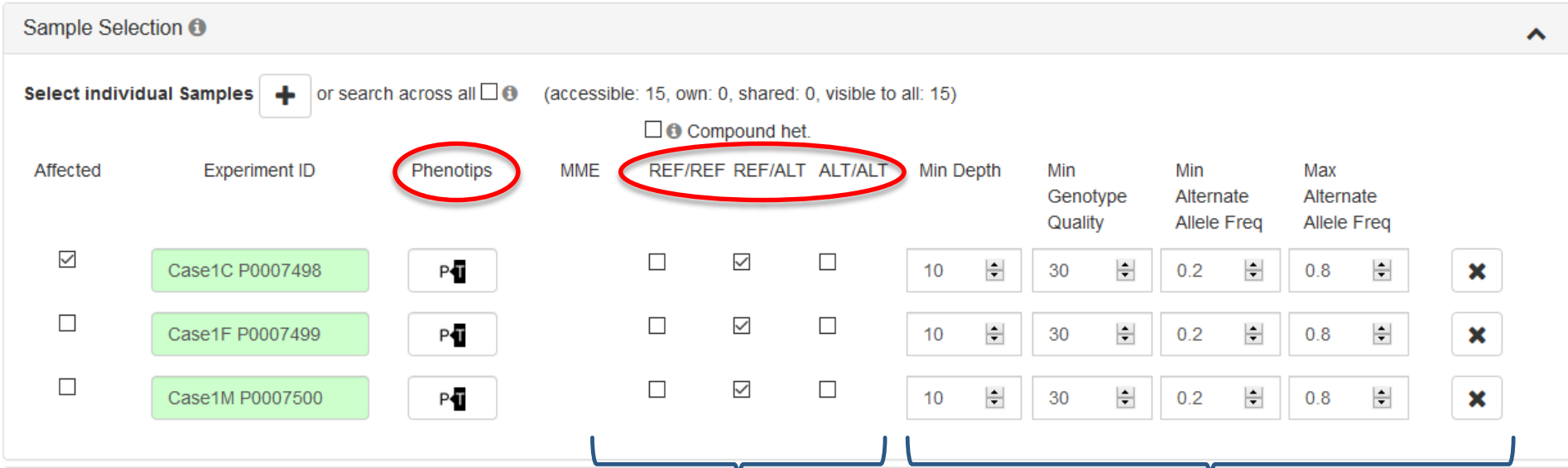
RD-Connect Genome- Phenome Analysis Platform



1. SAMPLES
2. FILTERS
3. RESULTS

RD Connect

Genomic platform: samples, inheritance and quality parameters

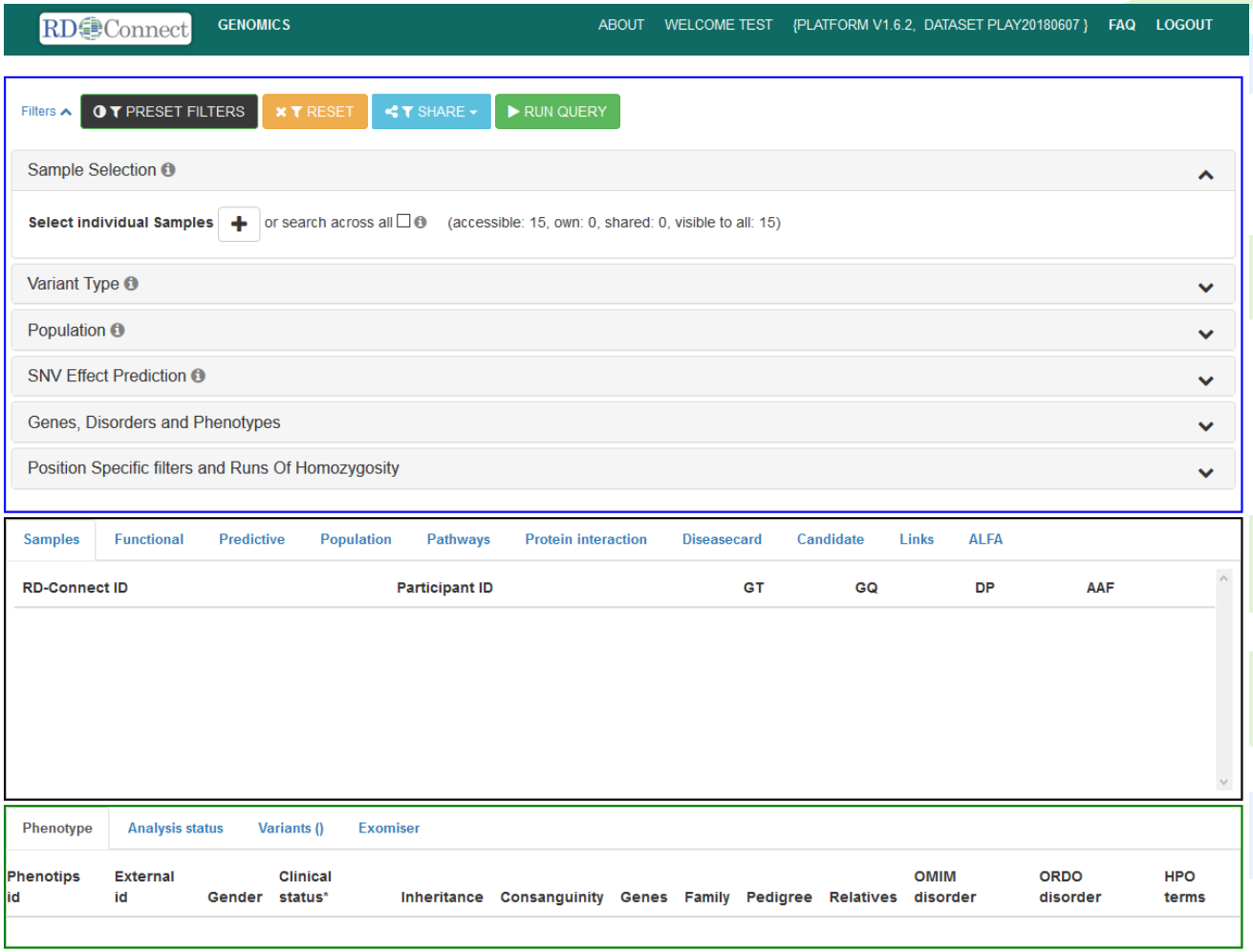


Genotypes – mode of inheritance

Quality parameters

RD Connect

RD-Connect Genome- Phenome Analysis Platform



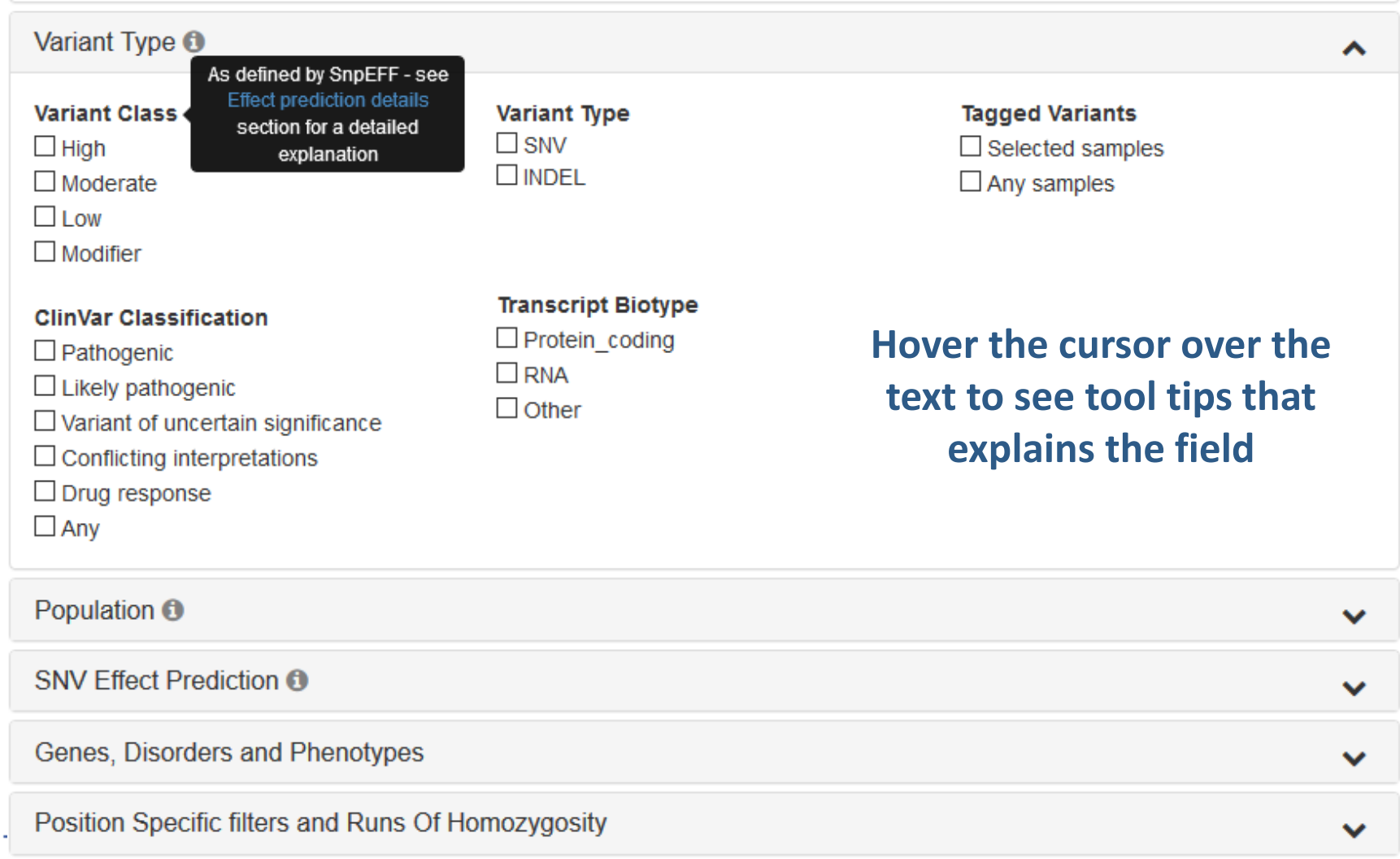
1. SAMPLES

2. FILTERS

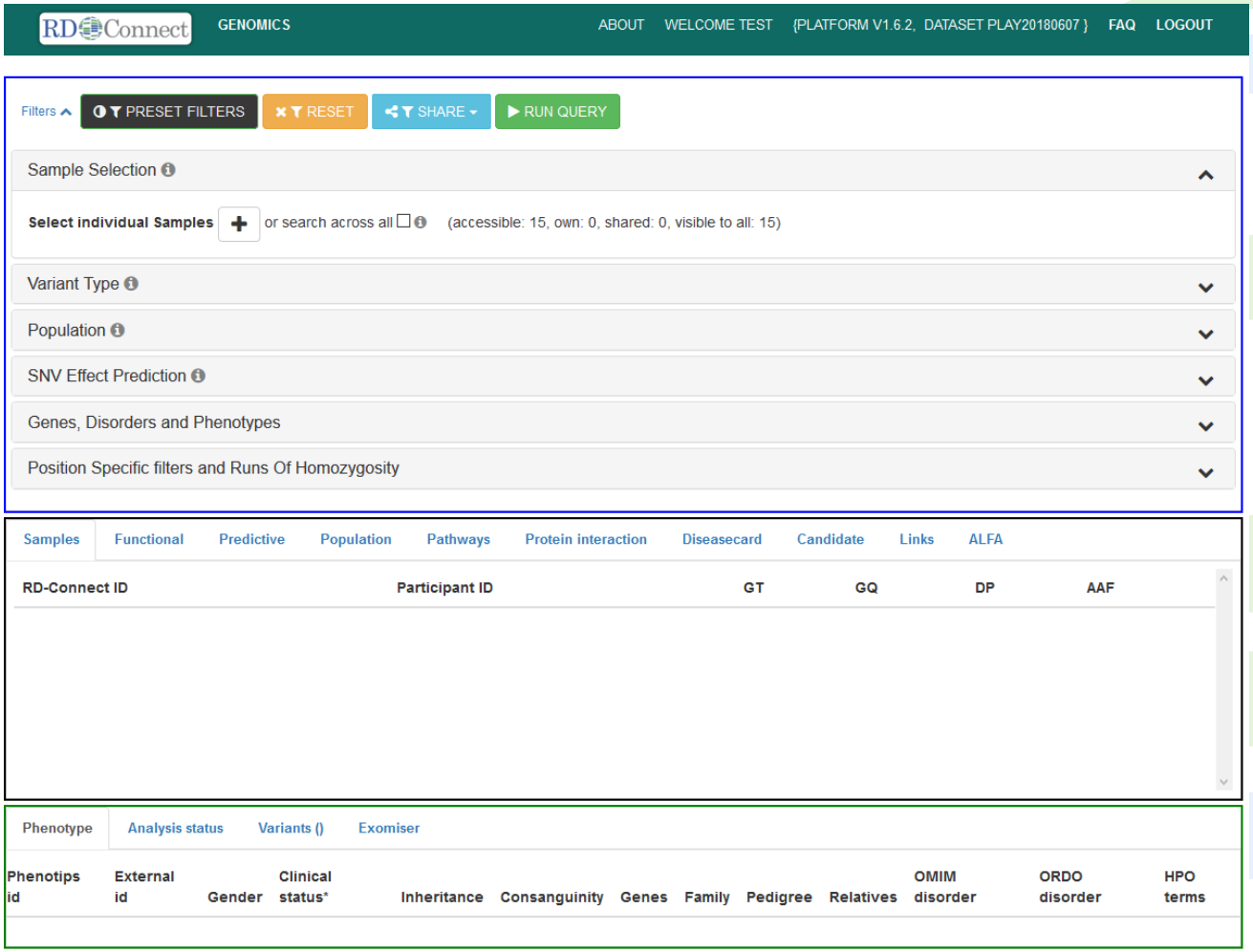
3. RESULTS

RD Connect

Genomic platform: filtering steps



RD-Connect genome phenome analysis platform:



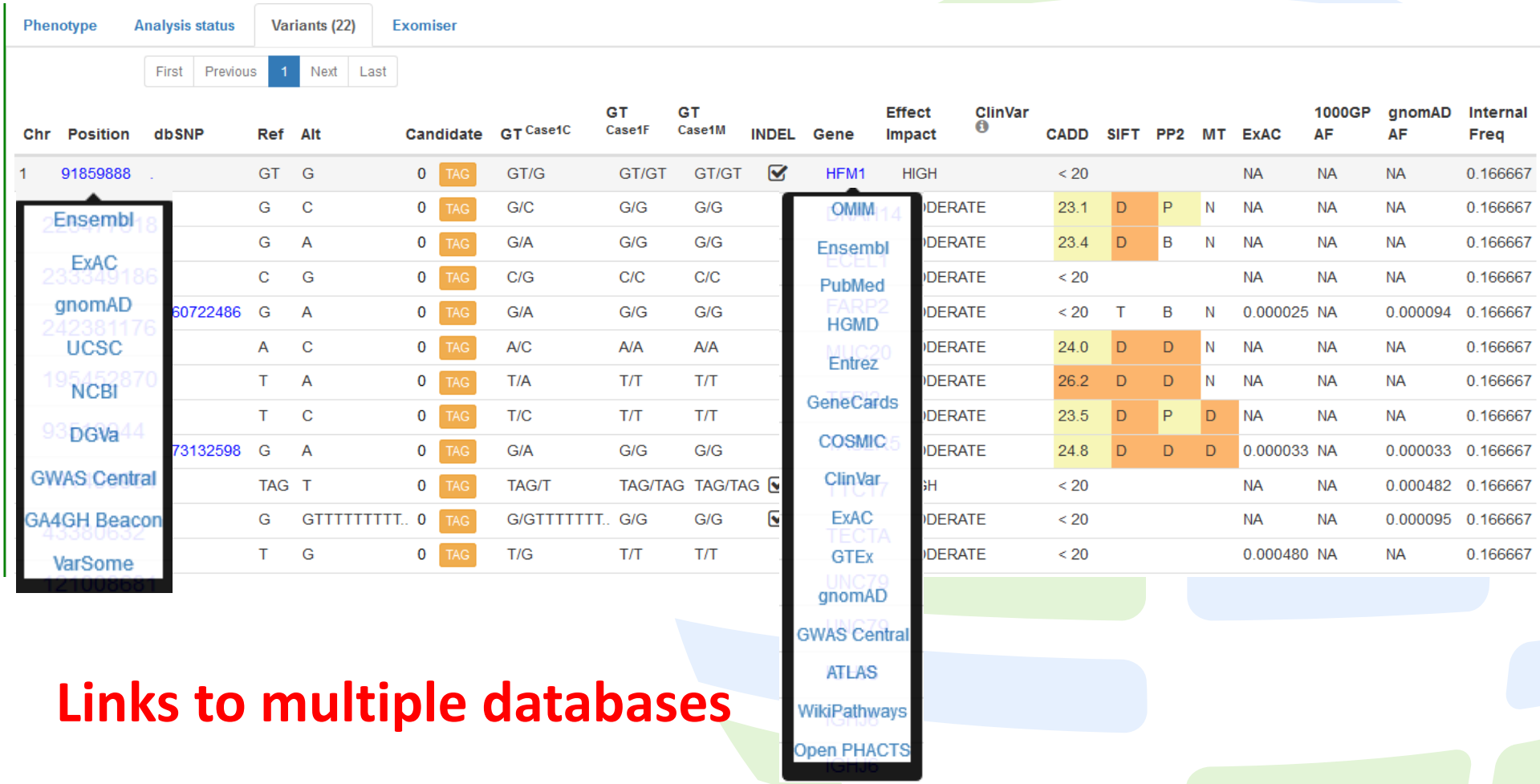
1. SAMPLES

2. FILTERS

3. RESULTS

RD-Connect

Genomic platform: results section



Links to multiple databases

RD Connect

Genomic platform: results section

Multiple tabs with detailed information for data interpretation

Samples | Functional | Predictive | Population | Pathways | Protein interaction | Diseasecard | Candidate | Links | ALFA

Gene Name	Transcript ID	Effect Impact	Consequence	Feature Type	HGVS coding	Amino Acid change	Amino Acid length	Exon Rank	CDS Position	Transcript BioType
HFM1	ENST00000370425	HIGH	frameshift_variant	transcript	c.255delA	p.Leu86Ter	1435	4/39	255/4308	protein_coding
HFM1	ENST00000427444	HIGH	frameshift_variant	transcript	c.129delA	p.Leu44Ter	196	3/4	129/591	protein_coding
HFM1	ENST00000455133	HIGH	frameshift_variant	transcript	c.255delA	p.Leu86Ter	138	4/4	255/417	protein_coding

Phenotype | Analysis status | Variants (22) | Exomiser

First | Previous | 1 | Next | Last

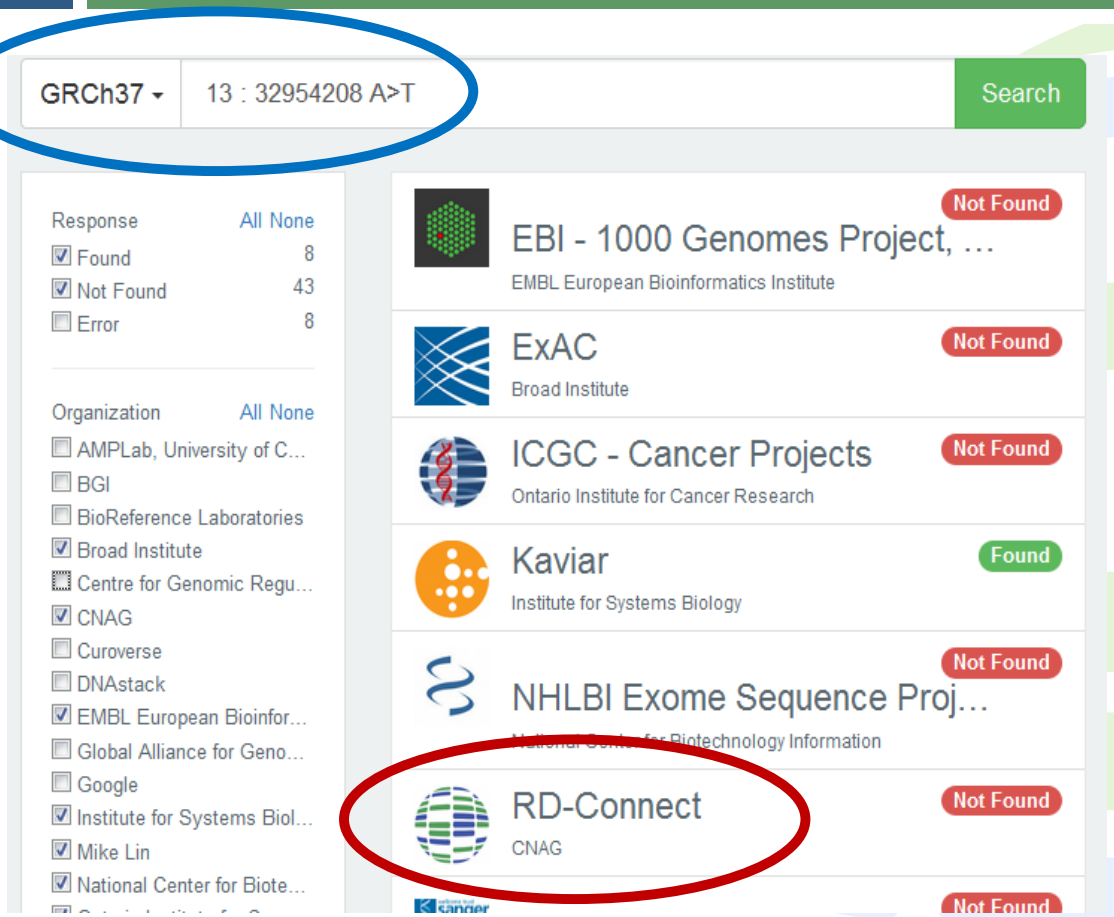
Chr	Position	dbSNP	Ref	Alt	Candidate	GT Case1C	GT Case1F	GT Case1M	INDEL	Gene	Effect Impact	ClinVar	CADD	SIFT	PP2	MT	ExAC	1000GP AF	gnomAD AF	Internal Freq
1	91859888	.	GT	G	0 TAG	GT/G	GT/GT	GT/GT	☑	HFM1	HIGH		< 20				NA	NA	NA	0.166667
1	225477618	.	G	C	0 TAG	G/C	G/G	G/G		DNAH14	MODERATE		23.1	D	P	N	NA	NA	NA	0.166667
2	233349186	.	G	A	0 TAG	G/A	G/G	G/G		ECEL1	MODERATE		23.4	D	B	N	NA	NA	NA	0.166667

RD Connect

RD-Connect GPAP functionalities

- ✓ **Standard filters and annotations** (variant impact on protein, frequency in control populations, frequency within database, *in silico* predictors, gene, genomic position, etc.)
- ✓ **Filter by genes of interest** (predefined/custom lists, associated to OMIM disease, HPO symptoms or Reactome)
- ✓ Filter by variants annotated in **ClinVar database**
- ✓ Filter to regions with observed long **Runs of homozygosity** (RoH)
- ✓ Direct link to **multiple external resources** (Ensembl, UCSC, gnomAD, HGMD, Human Splicing Finder, DiseaseCards, ALFA, etc.)
- ✓ **Phenotype driven strategy for variant prioritization** (Exomiser)
- ✓ **TAG and share candidate variants** using ACMG guidelines
- ✓ **Collaborative environment:** share data and queries
- ✓ **Anonymized data discovery** through Beacon and Matchmaker exchange

Data discovery: GA4GH Beacon

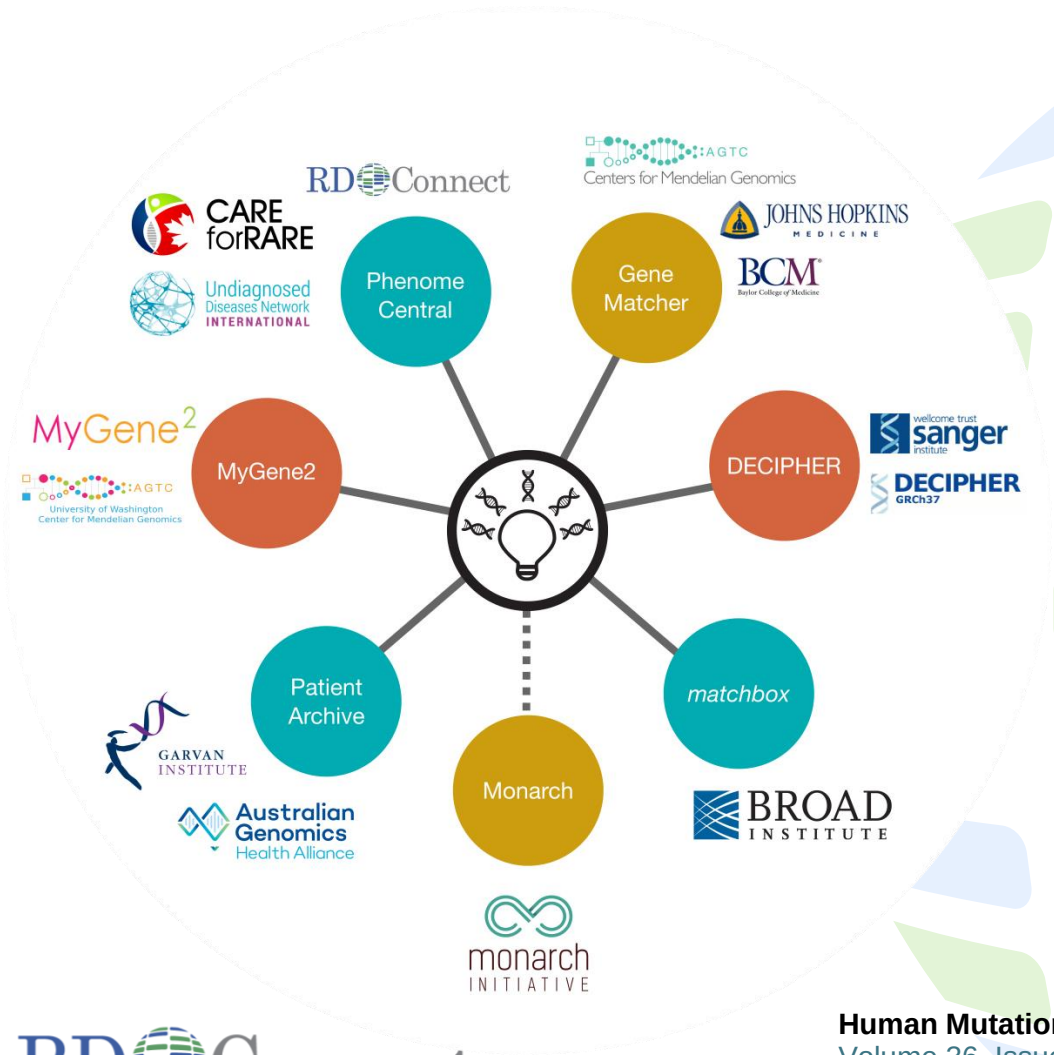


Question:
Does any sample in your database have **variant V**?

Answer:
Yes / No

https://beacon-network.org/#/

Data discovery: MatchMaker Exchange (MME, IRDiRC, GA4GH)



Question: Do you have a patient with similar phenotype and genotype as mine?

Human Mutation
Volume 36, Issue 10, pages 915-921, 17 SEP 2015 DOI: 10.1002/humu.22858
http://onlinelibrary.wiley.com/doi/10.1002/humu.22858/full#humu22858-fig-0001

2016 BBMRI-LPC Whole Exome Sequencing Call

Sequencing the exome of 900 rare disease samples
in collaboration with EuroBioBank and RD-Connect

Objectives

- ✓ to promote the usage of biobanks for rare diseases
  - samples deposited in the EuroBiobank network

- ✓ to promote the utilization of cutting-edge next-generation sequencing technology for the identification of novel causative variants and genes
  - free-of-charge sequencing of 900 exomes

- ✓ to molecularly diagnose rare disease patients
  - analysis through the RD-Connect platform

- ✓ to promote data sharing for rare disease research to enable future discovery
  - data sharing through the EGA and RD-Connect and phenotyping with HPO

2016 BBMRI-LPC Whole Exome Sequencing Call

Sequencing the exome of 900 rare disease samples
in collaboration with EuroBioBank and RD-Connect

17 transnational projects selected

899 samples for sequencing in total

10 projects sequenced at CNAG-CRG (**545** samples)

7 projects sequenced at WTSI (**354** samples)

A paradigm for best practice in data sharing in genomic studies?

bbmri-lpc
BBMRI – Large Prospective Cohorts

RD Connect

EuroBioBank

CONSENT

Sample

Phenotypic record

Biobank

EuroBioBank

INFRASTRUCTURE FOR SAMPLE SHARING

Sample can be reused for further research

PHENOTYPIC DATA STANDARDS

Diagnosis

INFRASTRUCTURE FOR DATA ANALYSIS

Sequencing

Online analysis

RD-Connect Platform

STANDARDIZED ALIGNMENT, CALLING & ANNOTATION

Data can be reused for further research

BBMRI-LPC clinician/ researcher

Combined genome-phenome database

INFRASTRUCTURE FOR DATA SHARING

EGA

The Undiagnosed Rare Disease Program of Catalonia (URDCat)

URDCAT is a research project embedded into the "Pla Estratègic de Recerca i Innovació en Salut (PERIS) 2016-2020" funded by the Health Department of the Generalitat de Catalunya.

MAIN OBJECTIVE:

Enable the Catalan Health System to **provide personalised genomic medicine as a fully integrated service for patients with RDs**, initially as a pilot project for RDs with neurologic involvement.

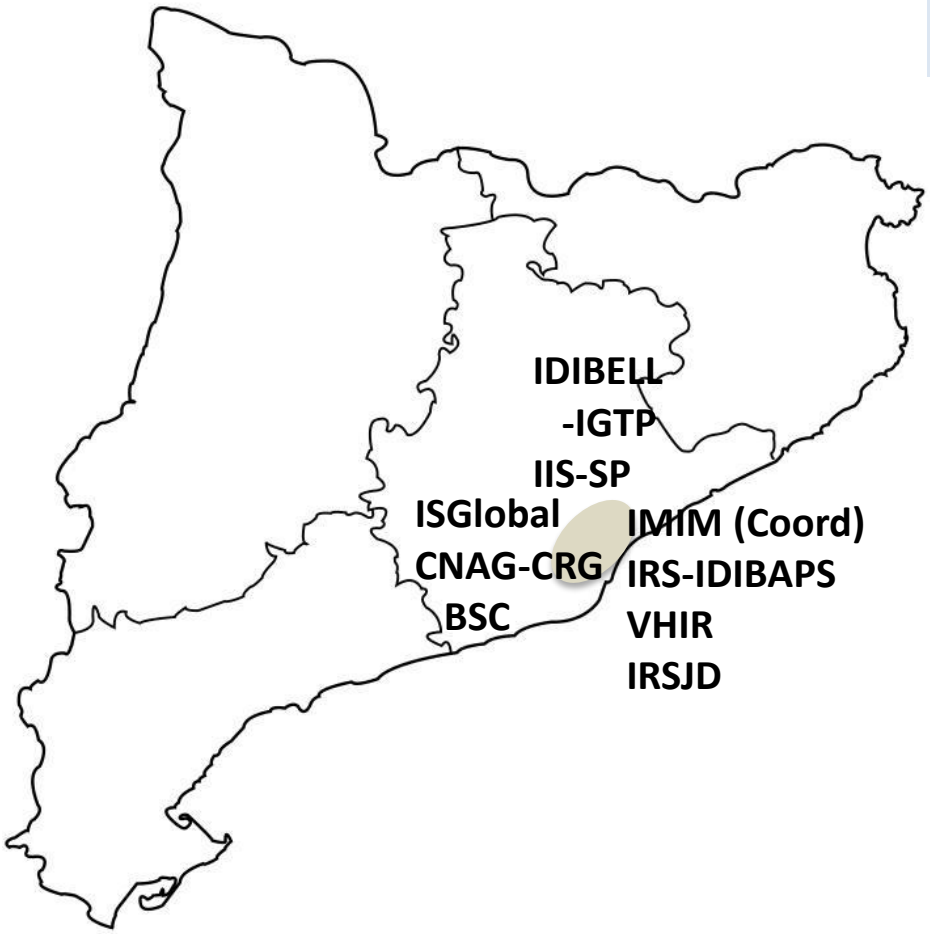
Specific objectives:

- **Standardise** the process of **analysis and integration of clinical and genomic data**
- Implement a **platform for analysis of genomics** data that is suitable for clinical practice
- **Identify causative genetic variants** of undiagnosed RD patients
- **Highlight the utility of genomics**, and advance its usage as a tool of **personalised medicine** for RD diagnostics
- **Educate and train stakeholders** in the value and use of genomic data

www.urdcat.cat

The Undiagnosed Rare Disease Program of Catalonia (URDCat)

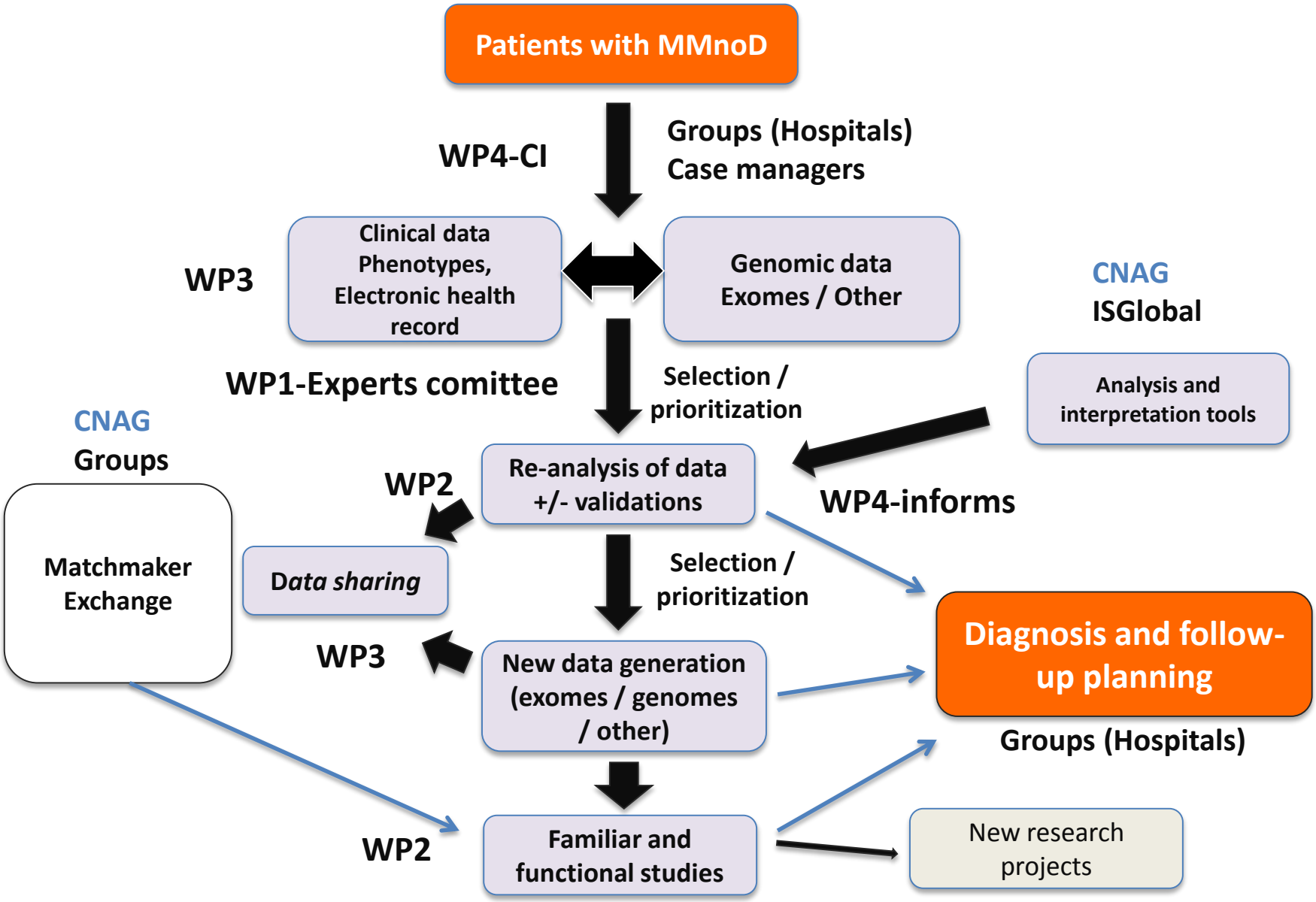
Coordinator: Luis Pérez Jurado (IMIM)



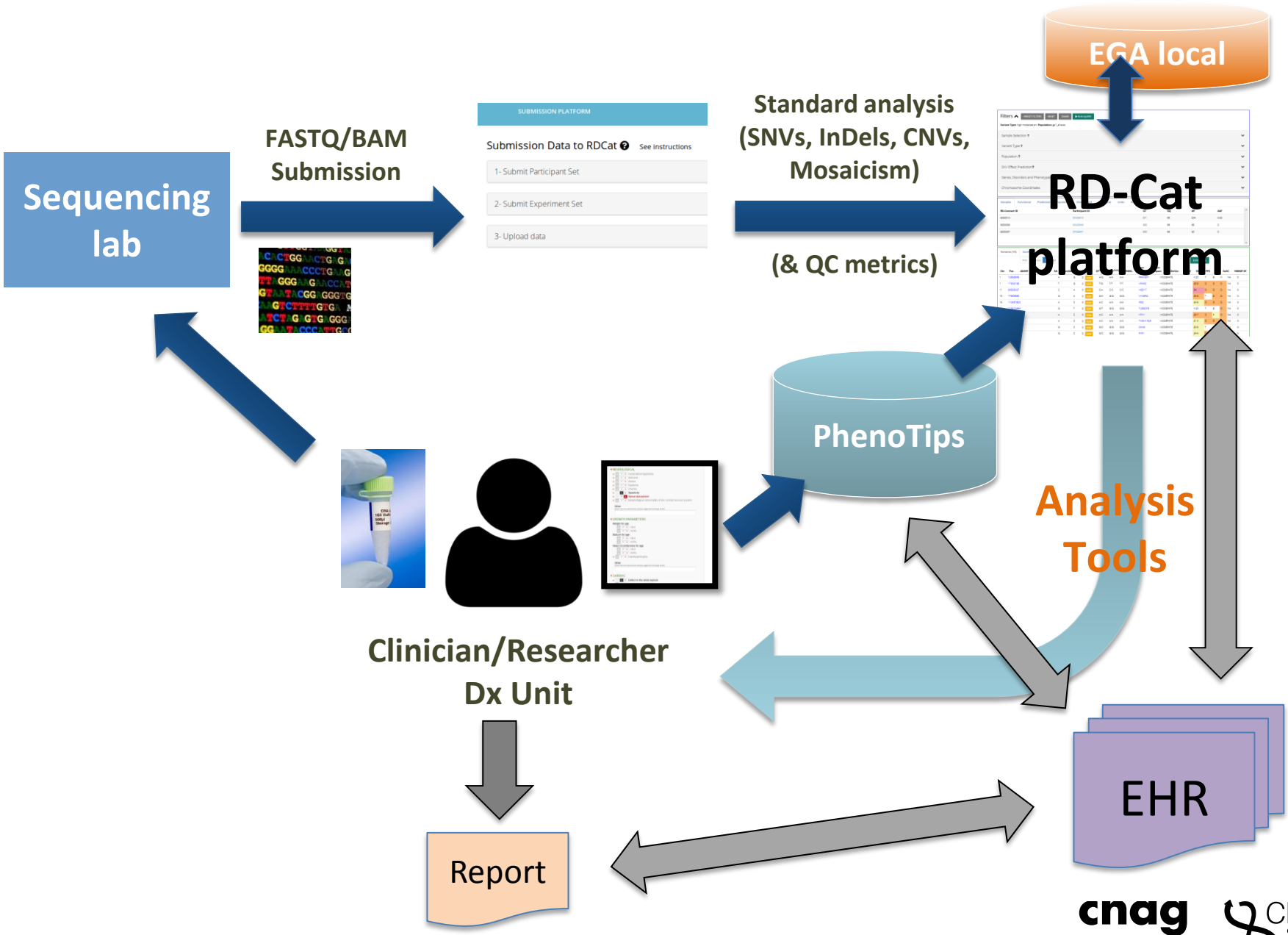
16 Groups
- 7 Hospitals (IIS)
- 3 CERCA centers
- FEDER (Patients)

WP1. Clinical data: patient selection and characterization (Coordinator: Alfons Macaya)
WP2. Genomics: sequencing, analysis, interpretation and validation (Coordinator: Antonia Ribes)
WP3. Data analysis and Interpretation platform (RDCat) (Coordinator: Sergi Beltran)
WP4. Ethical, Legal and and Social Issues (Coordinator: Francesc Palau)
WP5. Training (Coordinator: Luis Pérez Jurado)

The Undiagnosed Rare Disease Program of Catalonia (URDCat)



RDCAT platform workflow



cnag
centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

CRG
Centre for Genomic Regulation

Incidental or secondary findings

Genet Med. 2017 Feb;19(2):249-255. doi: 10.1038/gim.2016.190. Epub 2016 Nov 17.

Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics.

Kalia SS[1], Adelman K[2], Bale SJ[3], Chung WK[4,5], Eng C[6], Evans JP[7], Herman GE[8], Hufnagel SB[9], Klein TE[10], Korf BR[11], McKelvey KD[12,13], Ormond KE[10], Richards CS[14], Vlangos CN[15], Watson M[16], Martin CL[17], Miller DT[18].

Filters ^ | PRESET FILTERS | RESET | SHARE | RUN QUERY

Sample Selection ?

Select individual Samples + or search across all ? (accessible: 0, own: 568, shared: 0, visible to all: 568)

? Compound het.

Affected	Experiment ID	Phenotips	MME	REF/REF	REF/ALT	ALT/ALT	Min Depth	Min Genotype Quality	Min Alternate Allele Freq	Max Alternate Allele Freq	
☑	EPR532870 439	P[redacted]		☐	☑	☐	20	50	0,2	0,8	✖

Variant Type ?

Population ?

ACMG Medically Actionable Genes (n=59)
BabySeq Class A and B Genes (n=889)
Decipher all, July2017 (n=1,692)
Decipher confirmed, July2017 (n=1,098)
Digenic gene list, Feb2017 (n=136)
Imprinted-confirmed list, Feb2017 (n=80)
Imprinted-all, Feb2017 (n=253)
Intellectual Disabiltity-confirmed, Nov2017 (n=10
Intellectual Disability-all, Nov2017 (n=1648)
Leukodystrophy Associated (n=229)
Medically Interpretable Genome (n=5,419)
Mitocarta 2.0 (n=1158)
Muscle Gene Table, July2016 (n=416)
Primary Immune Deficiency, Nov2017 (n=387)
Tubingen HSP Version 6 (n=140)

59 genes for which the ACMG recommends reporting as incidental or secondary findings: Kalia SS et al (2016). Genetics in Medicine

Search OMIM

Search OMIM for clinical features,phenotypes...

Genes linked to : ✖ Remove All

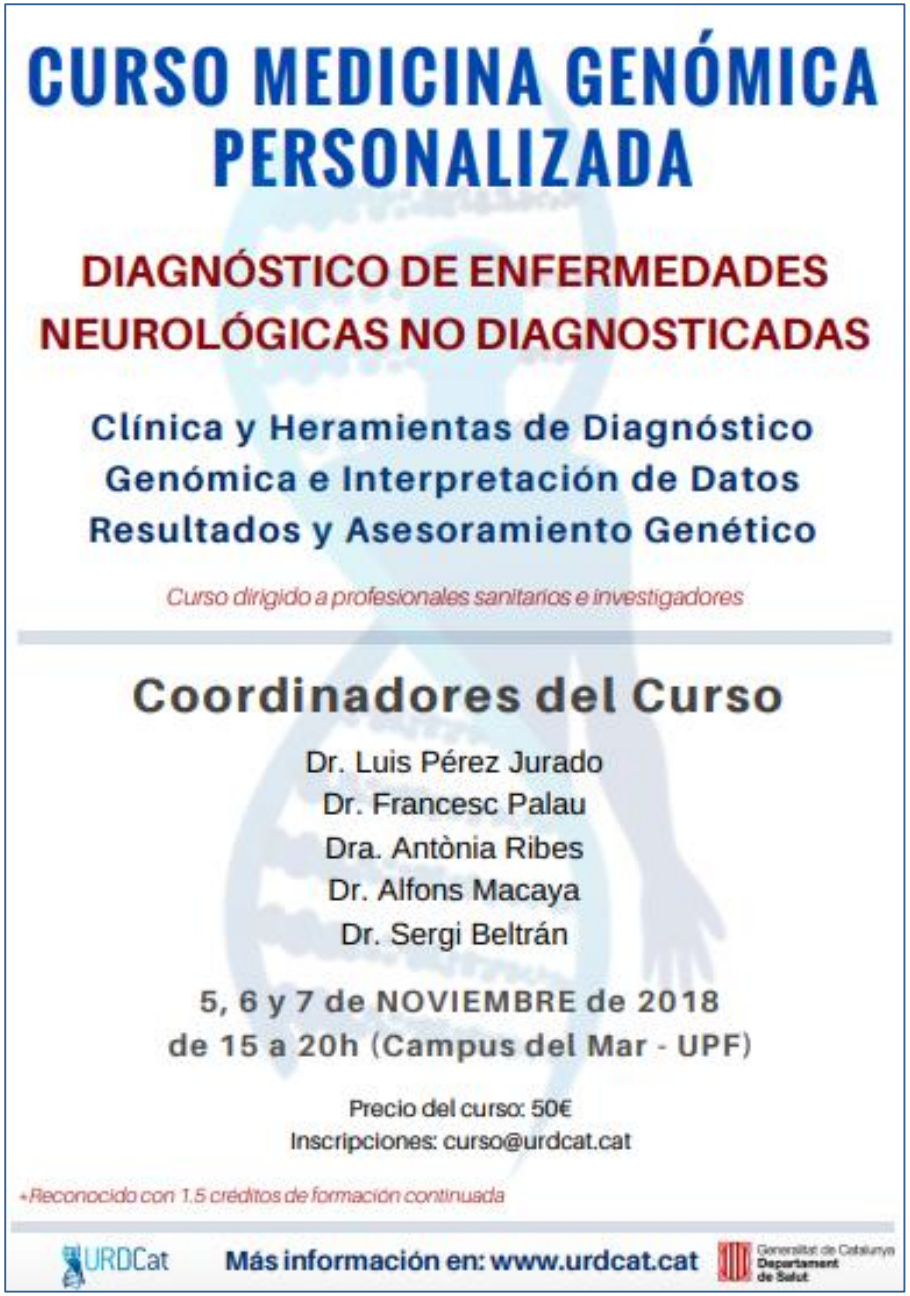
Search HPO

Enter search terms...

Fetch HPOs From PhenoTips | ✖ Remove All

Genes linked to :

Upload comma separated list of HGNC identifiers



www.urdcat.cat

NAVARRABIOMED
CENTRO DE INVESTIGACIÓN BIOMÉDICA

NAGEN – Proyecto Genoma Navarra

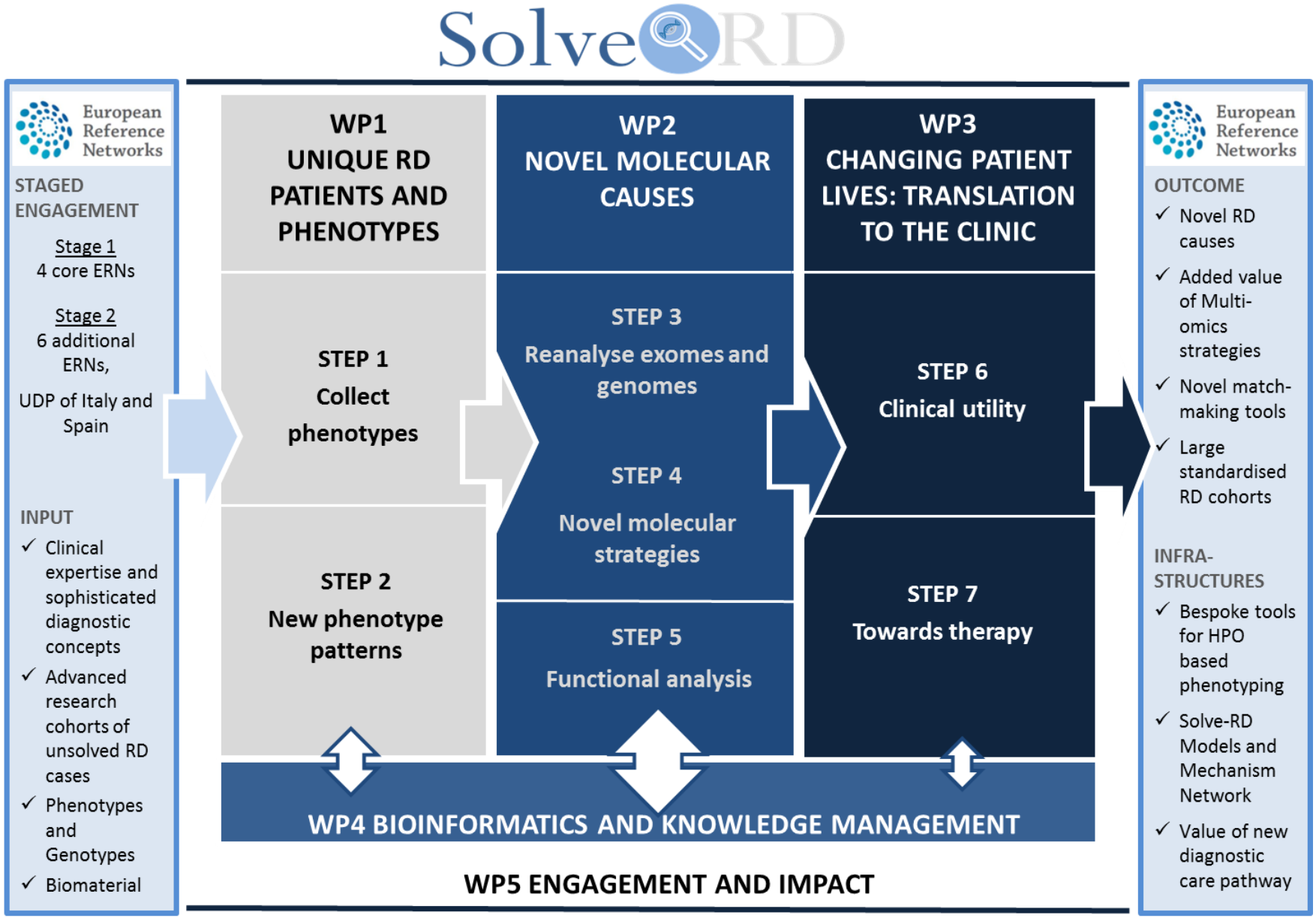
NAGEN es un proyecto piloto I+D+i promovido por el Centro de Investigación Biomédica Navarrabiomed en colaboración con la Dirección del Complejo Hospitalario de Navarra (CHN) para la implementación de la tecnología de secuenciación de Genoma Humano Completo (WGS) para su uso clínico en pacientes del CHN.

El objetivo de NAGEN es identificar las alteraciones del genoma de pacientes con más de 170 tipos de enfermedades raras o con ciertos tipos de cáncer de posible origen genético, cuya alteración genética causal no ha podido ser determinada a pesar de haberse realizado las pruebas genéticas indicadas para ello.

Empiece hoy

cnag
centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

CRG
Centre for Genomic Regulation


SolveRD
European Reference Networks
STAGED ENGAGEMENT
Stage 1
4 core ERNs
Stage 2
6 additional ERNs,
UDP of Italy and Spain
INPUT
✓ Clinical expertise and sophisticated diagnostic concepts
✓ Advanced research cohorts of unsolved RD cases
✓ Phenotypes and Genotypes
✓ Biomaterial
WP1 UNIQUE RD PATIENTS AND PHENOTYPES
WP2 NOVEL MOLECULAR CAUSES
WP3 CHANGING PATIENT LIVES: TRANSLATION TO THE CLINIC
STEP 1 Collect phenotypes
STEP 2 New phenotype patterns
STEP 3 Reanalyse exomes and genomes
STEP 4 Novel molecular strategies
STEP 5 Functional analysis
STEP 6 Clinical utility
STEP 7 Towards therapy
WP4 BIOINFORMATICS AND KNOWLEDGE MANAGEMENT
WP5 ENGAGEMENT AND IMPACT
European Reference Networks
OUTCOME
✓ Novel RD causes
✓ Added value of Multi-omics strategies
✓ Novel match-making tools
✓ Large standardised RD cohorts
INFRA-STRUCTURES
✓ Bespoke tools for HPO based phenotyping
✓ Solve-RD Models and Mechanism Network
✓ Value of new diagnostic care pathway

ELIXIR - the European life-science Infrastructure for Biological Information

ELIXIR consolidates Europe's **national centres, services, and core bioinformatics resources** into a single, **coordinated infrastructure.**

Platforms

Tools
Interoperabilty
Data
Compute
Training

Use Cases

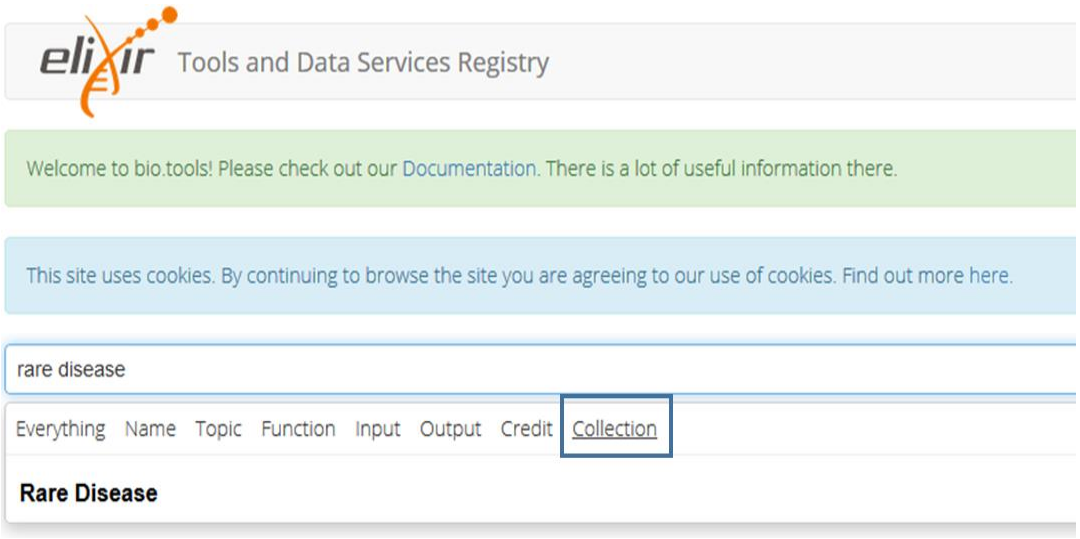
Human Data
Marine Metagenomics
Rare Disease
Plant science

5 year project
2015-2020

ELIXIR-EXCELERATE

Fast-track ELIXIR implementation and drive early user exploitation across the life-sciences

ELIXIR registry of data resources – bio.tools



List all tools useful for RD research

Ongoing work:

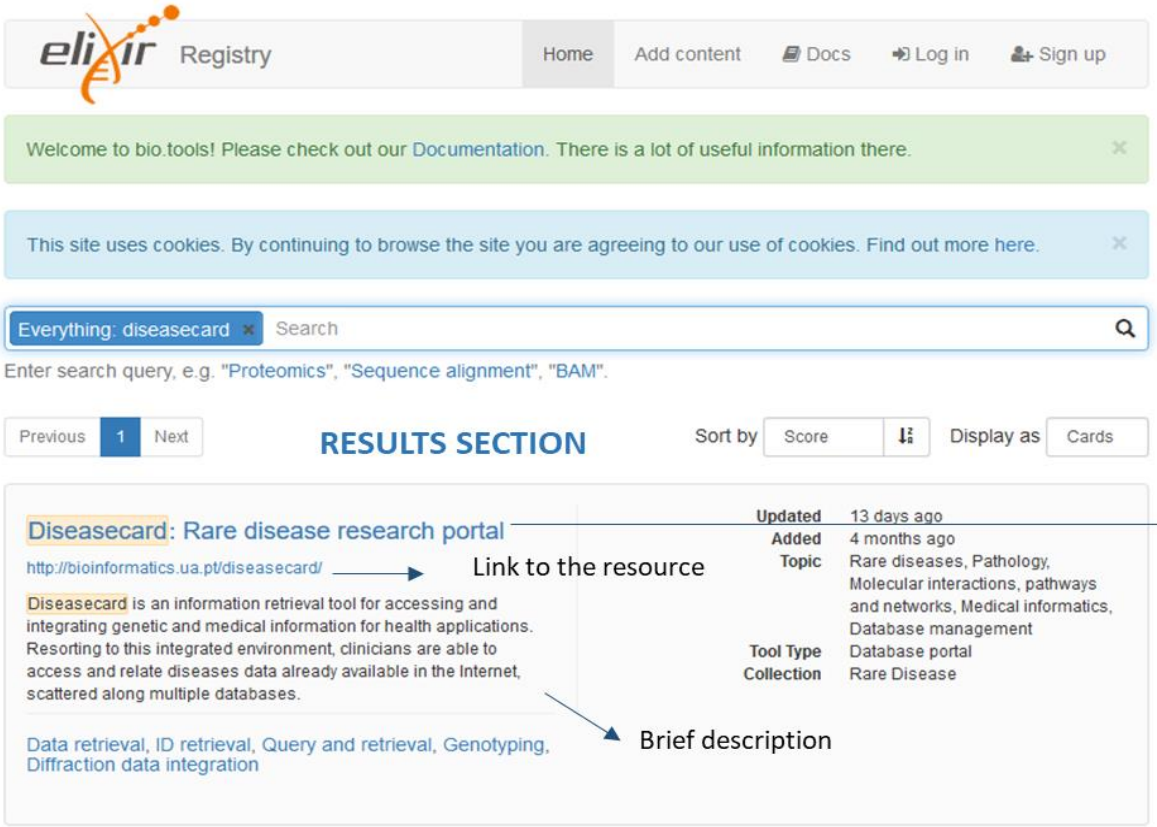
Link with benchmarking activities (WP2)

bio.tools API

Ongoing work:

RD-connect collection to pull up to date information concerning the tools integrated in the GPAP

ELIXIR registry of data resources



Search browser: type a resource/ tool of interest

Link to the resource

Tool name and link to more detailed information

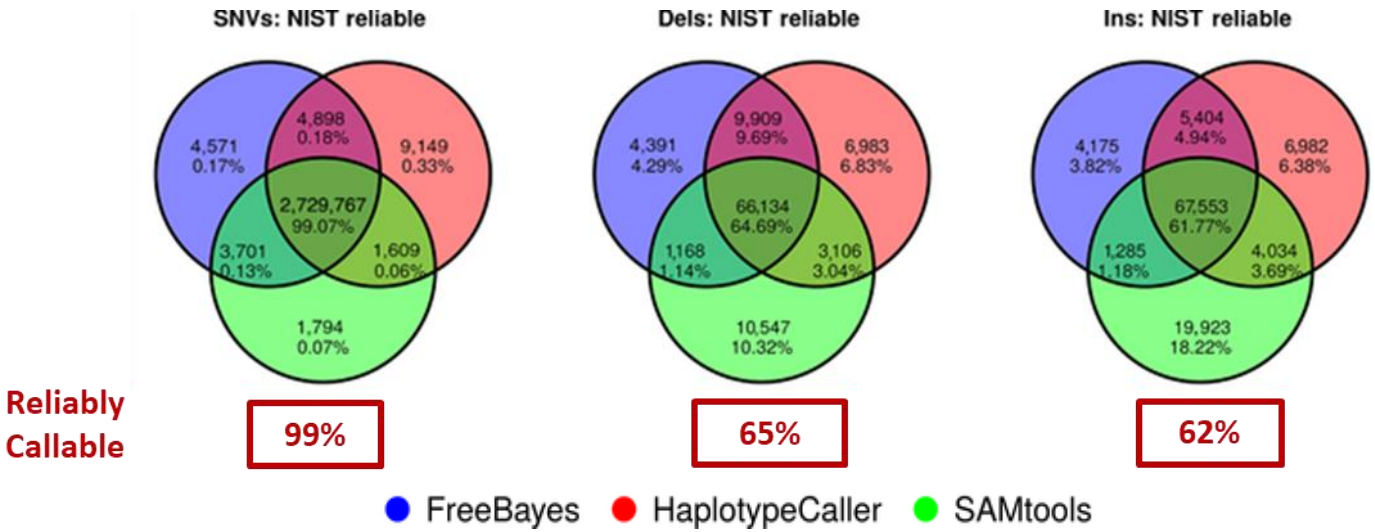
Brief description

Links to useful information concerning the tool.

Reference / source (direct link to pubmed entry)

Other actions

Benchmarking and standardizing genomics pipeline



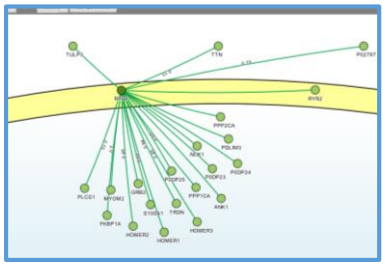
Laurie *et al*. Human Mutation, 2016

USEFUL FOR THE RD COMMUNITY
-> RD-connect variant calling pipeline

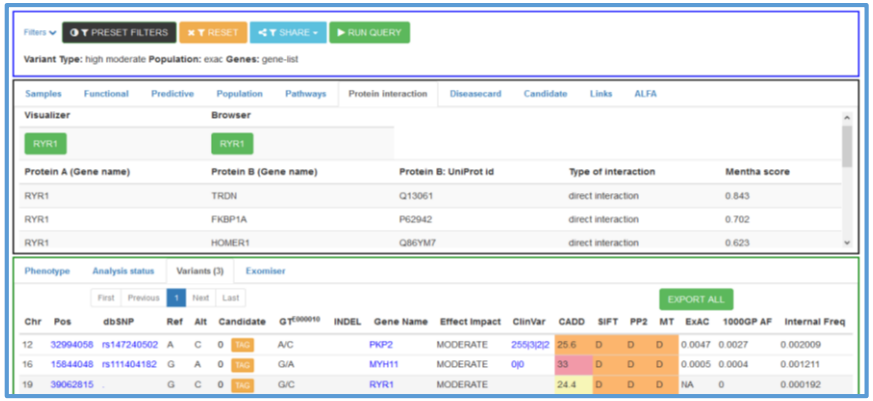
- Working in collaboration with the RD communities to implement the adequate quality reporting standards
- ELIXIR IS: "Development of Architecture for Software Containers at ELIXIR and its use by EXCELERATE use-case communities"

ELIXIR Implementation Study: Integration of ELIXIR-IIB in ELIXIR Rare Diseases activities

RD-Connect

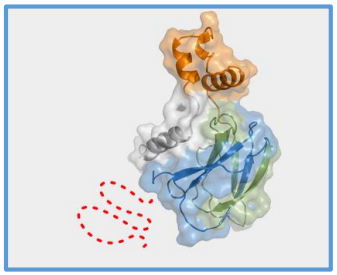


MINT / MENTHA (protein interaction databases)

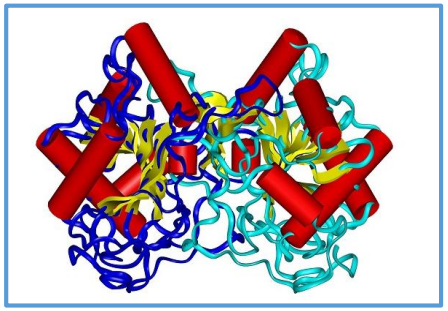


eDGAR

disease-gene Association database



VHLdb: A database of von Hippel-Lindau protein interactors and mutations



GALT-PROTEIN-DB 2.0

REDIportal

An ATLAS of A-to-I RNA editing events in human

HmtDB Human Mitochondrial DataBase

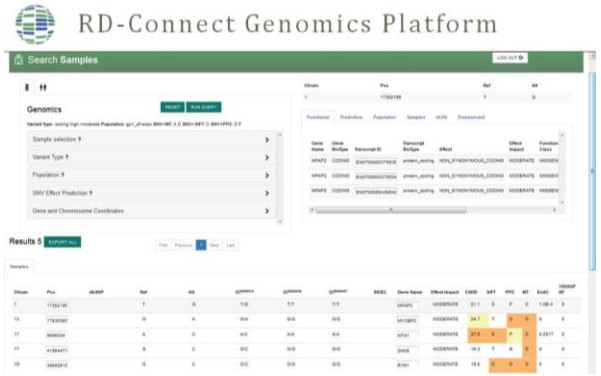
Human Mitochondrial DataBase

elixir

ELIXIR Implementation Study to visualize data deposited at the EGA in real-time

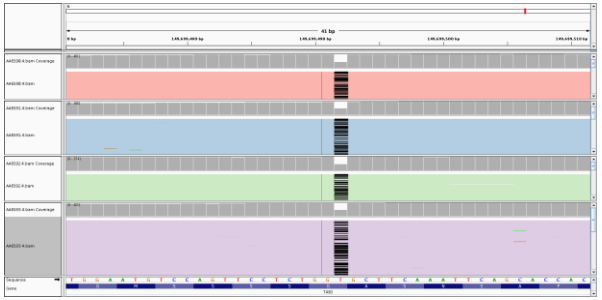
RD-Connect
BAM/CRAM

1 File submission to the EGA & ID-mapping



2 Request
- File ID/s
- Chr. coordinates

EGA
(BAM/CRAM)



3 Reply
- BAM/CRAM slice
- or
- Data representation

Htsget API

4 Rendering*

Secure and encrypted connections

*Broad's IGV screenshot

Adoptable by other projects and data types

Acknowledgements

RD-Connect

URDCat

cnag CRG Centre for Genomic Regulation

WP1: Coordination
Ivo Gut / Hanns Lochmüller
(CNAG-CRG)

WP2: Patient registries
Domenica Taruscio (ISS and EPIRARE)

WP3: Biobanks
Lucia Monaco
(Fondaz. Telethon & EuroBioBank)

WP4: Bioinformatics
Christophe Béroud
(INSERM Marseille)

WP5: Unified platform
Ivo Gut (CNAG Barcelona)

WP6 Ethical/legal/social
Mats Hansson (Uppsala)

WP7: Impact/Innovation
Kate Bushby
(Newcastle and EUCERD/ EJARD)

IMIM
Luis Pérez-Jurado
CLÍNIC
Antònia Ribes
Montse Milà
Francesc Cardellach
HSJD
Francesc Palau
Rafael Artuch
VHIR
Alfons Macaya
Ramón Martí
Eduardo Tizzano
IDIBELL
Aurora Pujol
Conxi Lázaro

HSP
Pia Gallano
ISGlobal
Juan R. González
UB
Daniel Grinberg
CNAG
Sergi Beltran

NAGEN PROYECTO GENOMA 1000 NAVARRA

Ángel Alonso
Sara Pasalodos
Josefa Salgado
Joaquín Dopazo
David Gómez

S. Beltran
S. Laurie
D. Piscia
L. Matalonga
A. Papakon.
F. Camacho
D. Picó
R. Tonda
J.R. Trotta
G. Parra
C. Luengo
J. Camps
I. Martinez
D. Ovelleiro
J. Protasio

M. Ingham
E. Casals
R. Alcántara

I. Gut
M. Gut
M. Bayès
B. Fusté
CNAG lab
CNAG admin

IRDiRC INTERNATIONAL RARE DISEASES RESEARCH CONSORTIUM

bbmri-lpc
BBMRI – Large Prospective Cohorts

Generalitat de Catalunya
Departament d'Economia i Coneixement

Generalitat de Catalunya
Departament de Salut

PERIS 2016 2020

SolveRD

elixir

excelerate

INB

ISC Instituto de Salud Carlos III

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

SEVENTH FRAMEWORK PROGRAMME