

Plataforma RD-Cat: módulo de datos genómicos

Gemma Bullich

Leslie Matalonga

Centre Nacional d'Anàlisi Genòmica, CNAG-CRG, Barcelona

Tècniques òmiques en el diagnòstic de malalties minoritàries,

16/11/2022

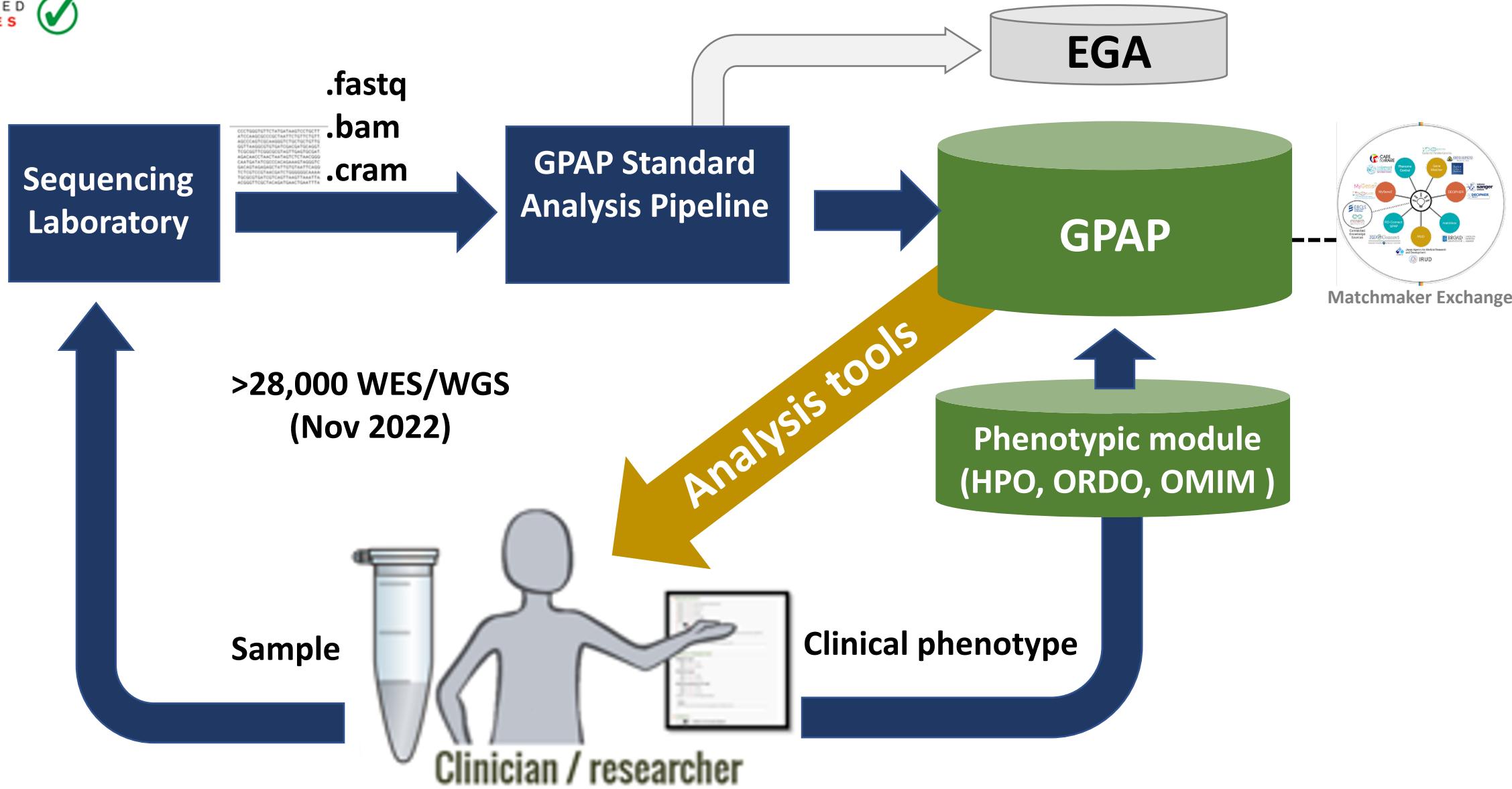


```
,123 0|0:123:123,123 0|1:123:123,123 0|1:49:52,5  
0:123:123,123 0|0:123:123,123 0|0:123:123,123 0|0:  
,123 0|0:123:123,123 0|0:123:123,123 0|0:123:123,123 0|0:  
1:123:123,123  
,123 1|0:123:123,123:56:0.0852854:21:19 0|0:  
,123 0|0:83:83,123 0|1:43:123,43 0|0:123:123,123 0|0:  
0:68:68,123 0|0:123:123,123 0|0:123:123,123 0|0:  
,51 0|0:43:43,123 0|0:87:123,87 0|0:114:123 0|0:  
0:37:37,123 0|0:123:123,123 0|0:123:123,123 0|0:  
,123 1|0:123:123,123  
,123 0|0:123:123,123:59:0.102882:5:3 0|0:113:123 0|0:  
0:123:123,123 0|0:123:123,123 0|0:76:105,76 0|0:  
,123 0|0:76:76,123 0|0:123:123,123 0|0:123:123,123 0|0:  
0:123:123,123 0|0:123:123,123 1|0:123:123,123 0|0:  
,123 1|0:123:123,123 0|1:106:123,106 0|0:  
0:113:123,113 0|0:123:123,123 0|0:  
0:123:1
```

```
tro@n8 indelcalling]$ cp /scratch/devel/fcastro/COPY_temp/indelcalling README_* .  
tro@n8 indelcalling]$ cp /scratch/devel/fcastro/data/1000genomes/indelcalling/CEU* .  
tro@n8 indelcalling]$ cp /scratch/devel/fcastro/data/1000genomes/indelcalling/CEU*.tbi CEU.  
tro@n8 indelcalling]$ ls  
tro@n8 indelcalling]$ RPO00031.2010_03.indels.genotypes.vcf.gz CEU.SRP00031.2010_03.indels.genotypes.vcf.gz.tbi  
tro@n8 indelcalling]$ cp /scratch/devel/fcastro/data/1000genomes/indelcalling/CEU* .  
tro@n8 indelcalling]$ pwd  
/scratch/devel/fcastro/COPY_temp/indelcalling  
tro@n8 indelcalling]$ cd /scratch/
```



Full integration with the GPAP



Molecular diagnostic and gene discovery challenge



GENOME or EXOME SEQUENCING

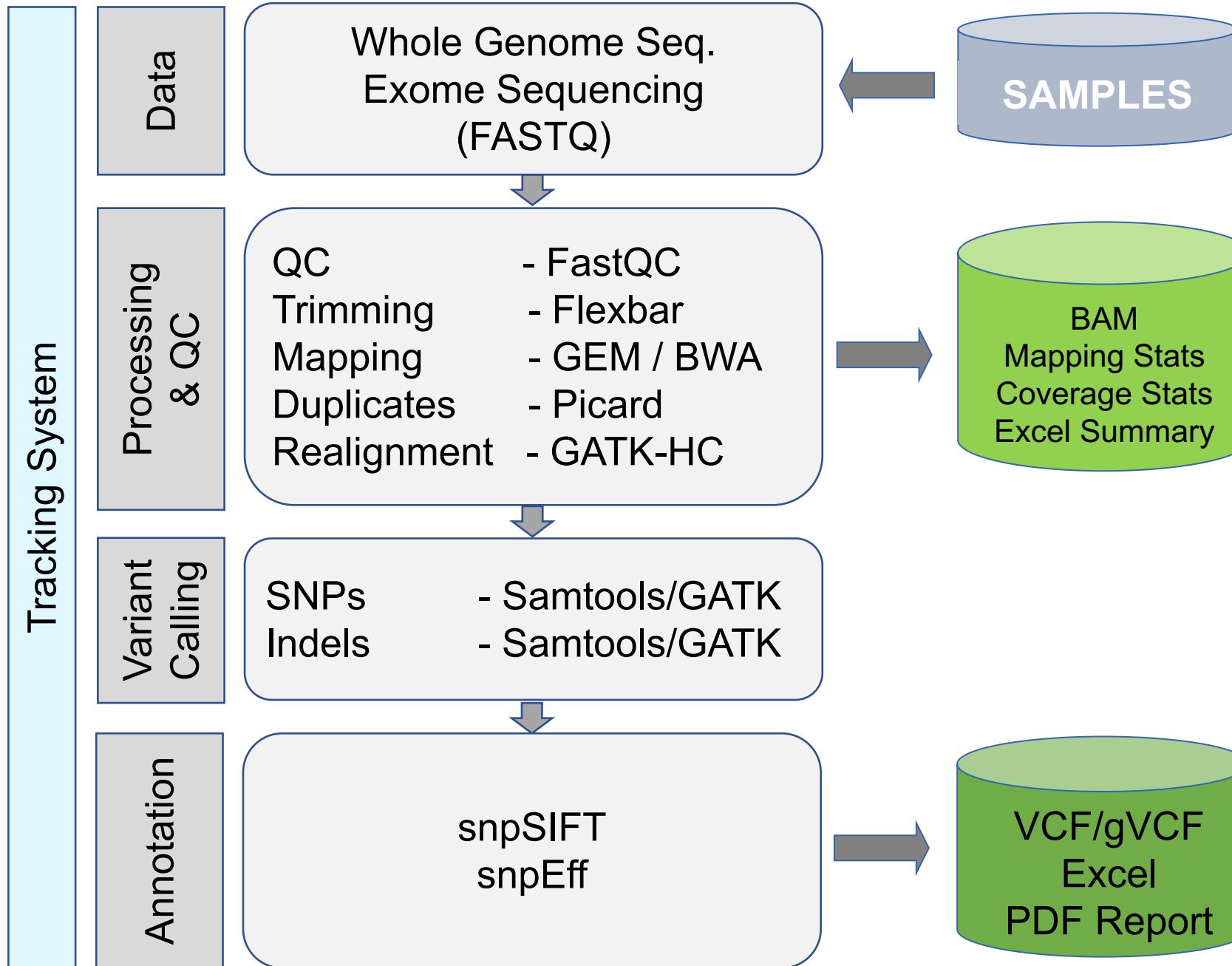


Variant calling
pipeline

- From fastq to gvcf (standard analysis pipeline Laurie et al., 2016– GATK best practices)

CNAG's Variant Calling Pipeline

R. Tonda, S. Derdak, S. Laurie, J. Camps, JR Trotta, S Beltran



Molecular diagnostic and gene discovery challenge



GENOME or EXOME SEQUENCING

Variant calling
pipeline

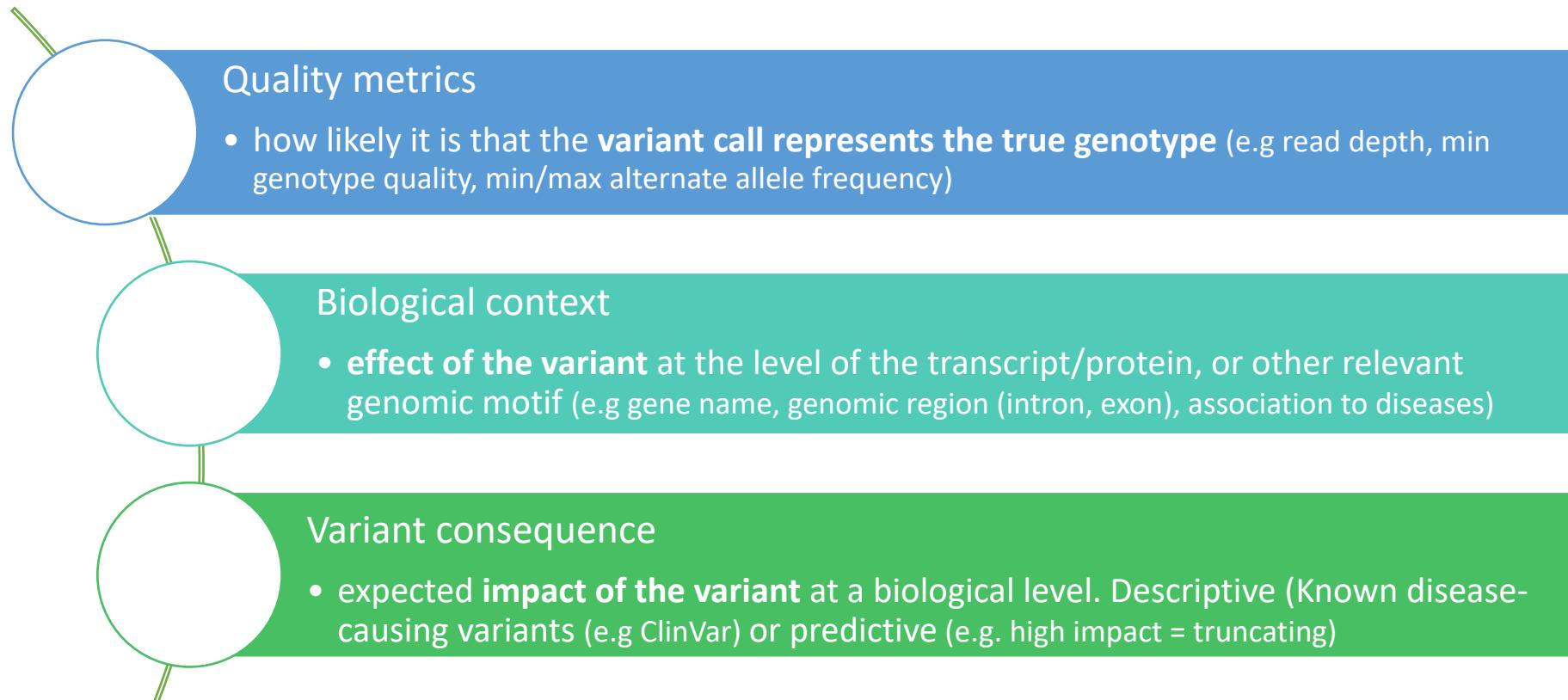
- From fastq to gvcf (standard analysis pipeline Laurie et al., 2016– GATK best practices)

Annotation

- Addition of information to a variant which helps us better understand the variant in terms of its biological context

Different class of variant annotation

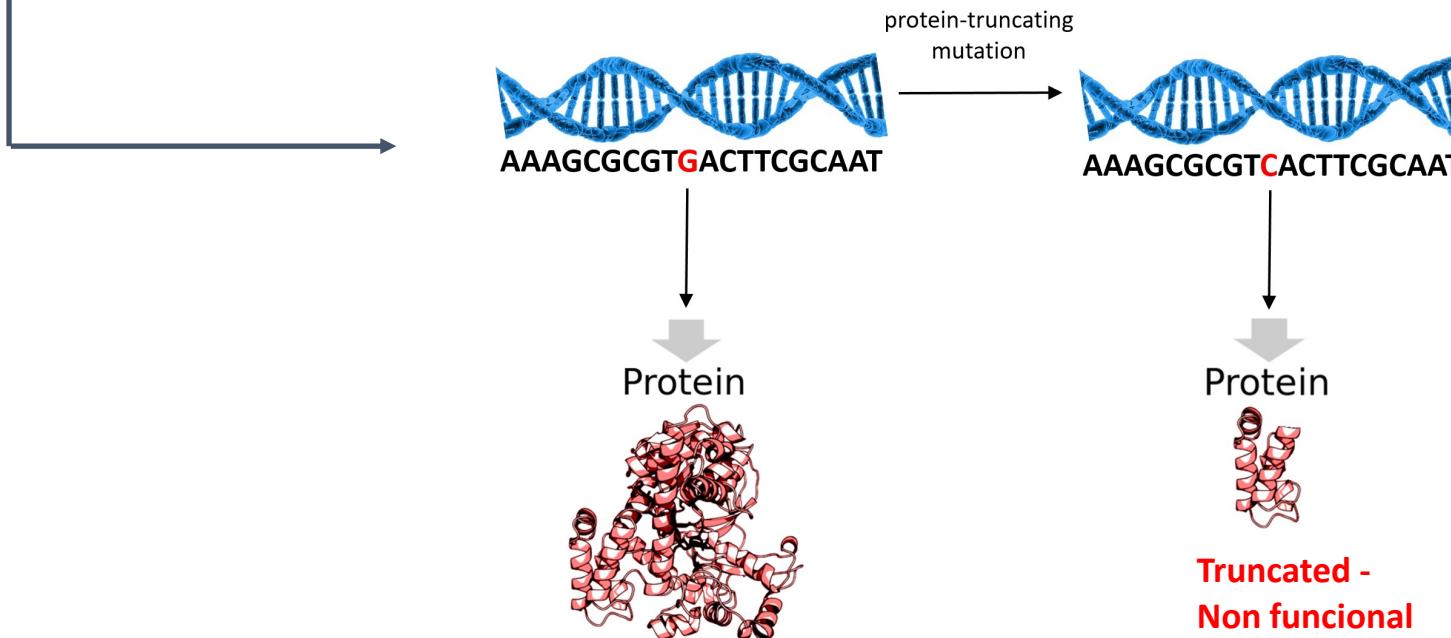
Addition of information to a variant which helps us better understand the variant in terms of its biological context



Filtering steps- variant consequence- SnpEff

SnpEff/ VEP categorise all variants into one of four impact categories (variant consequence annotation)

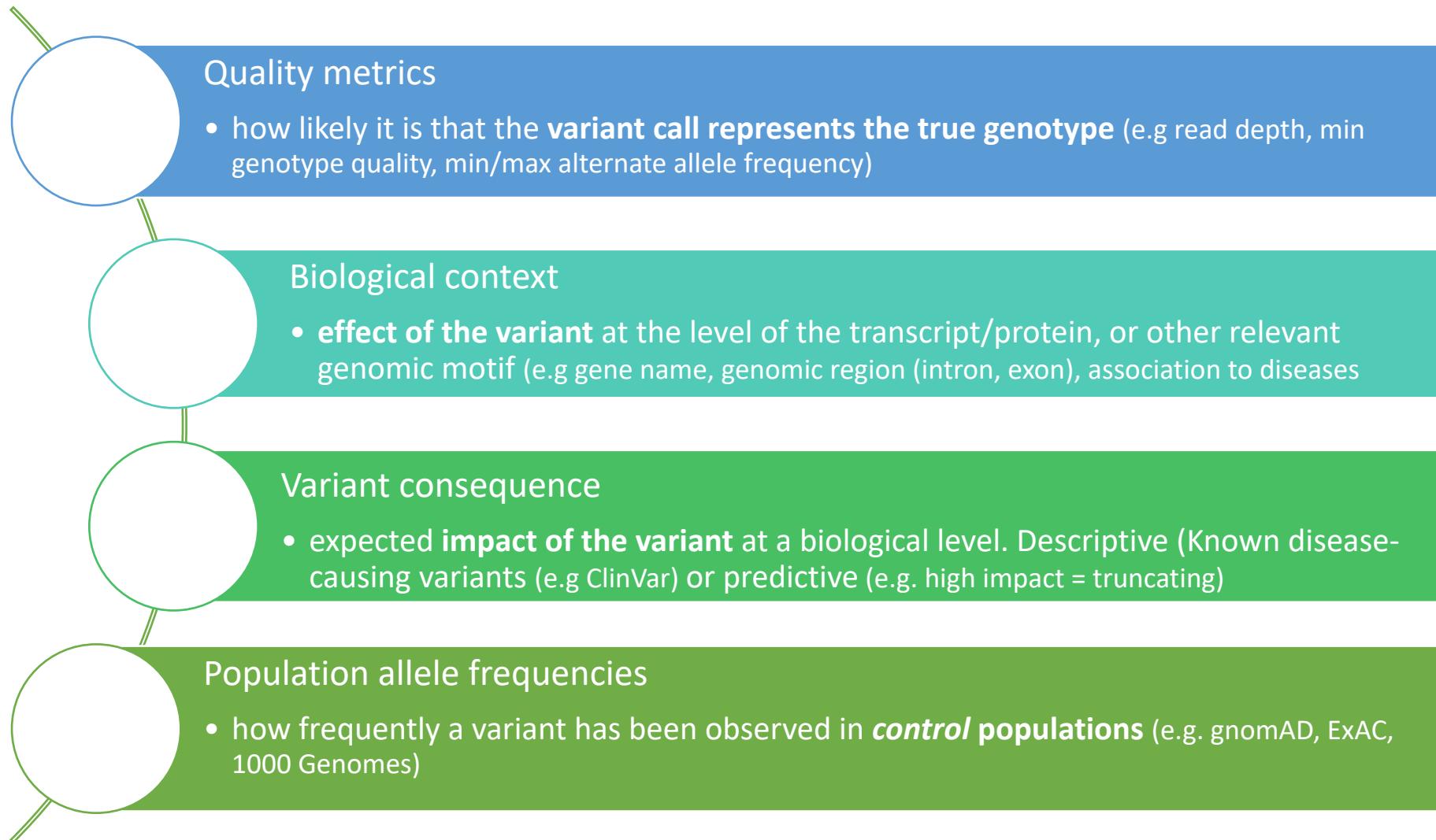
| | |
|-----------------|---|
| HIGH | Variants significantly affect protein structure e.g. nonsense, frameshift, canonical splice-sites |
| MODERATE | Mainly non-synonymous variants i.e. amino-acid change |
| LOW | Mainly synonymous variants i.e. no amino acid change |
| MODIFIER | Mutations in non-coding regions |



Source: <http://thedishonscience.stanford.edu/posts/essential-genes/>

Different class of variant annotation

Addition of information to a variant which helps us better understand the variant in terms of its biological context



Variants in the diploid genome

Single Nucleotide Variants (SNVs)

- > 3,000,000
- > 30,000-40,000 “exomic”

Short insertions and deletions (InDels)

- > 300,000-400,000

Copy number variants (CNVs)

- 5-10% of the genome

Other structural variants (SVs)

- ~10% of the genome

Variants in the diploid genome

Single Nucleotide Variants (SNVs)

- > 3,000,000
- > 30,000-40,000 “exomic”

Short insertions and deletions (InDels)

- > 300,000-400,000

Copy number variants (CNVs)

- 5-10% of the genome

Other structural variants (SVs)

- ~10% of the genome

Molecular diagnostic and gene discovery challenge



GENOME or EXOME SEQUENCING

Variant calling
pipeline

- From fastq to gvcf (standard analysis pipeline Laurie et al., 2016– GATK best practices)

Annotation

- Addition of information to a variant which helps us better understand the variant in terms of its biological context

Filtration

- Reduction in the number of variants of interest, based upon the annotations

Molecular diagnostic and gene discovery challenge



GENOME or EXOME SEQUENCING

Variant calling pipeline

- From fastq to gvcf (standard analysis pipeline Laurie et al., 2016– GATK best practices)

Annotation

- Addition of information to a variant which helps us better understand the variant in terms of its biological context

Filtration

- Reduction in the number of variants of interest, based upon the annotations

Prioritisation

- Upon phenotypic and molecular information

Molecular diagnostic and gene discovery challenge



GENOME or EXOME SEQUENCING

Variant calling pipeline

- From fastq to gvcf (standard analysis pipeline Laurie et al., 2016– GATK best practices)

Annotation

- Addition of information to a variant which helps us better understand the variant in terms of its biological context

Filtration

- Reduction in the number of variants of interest, based upon the annotations



Prioritisation

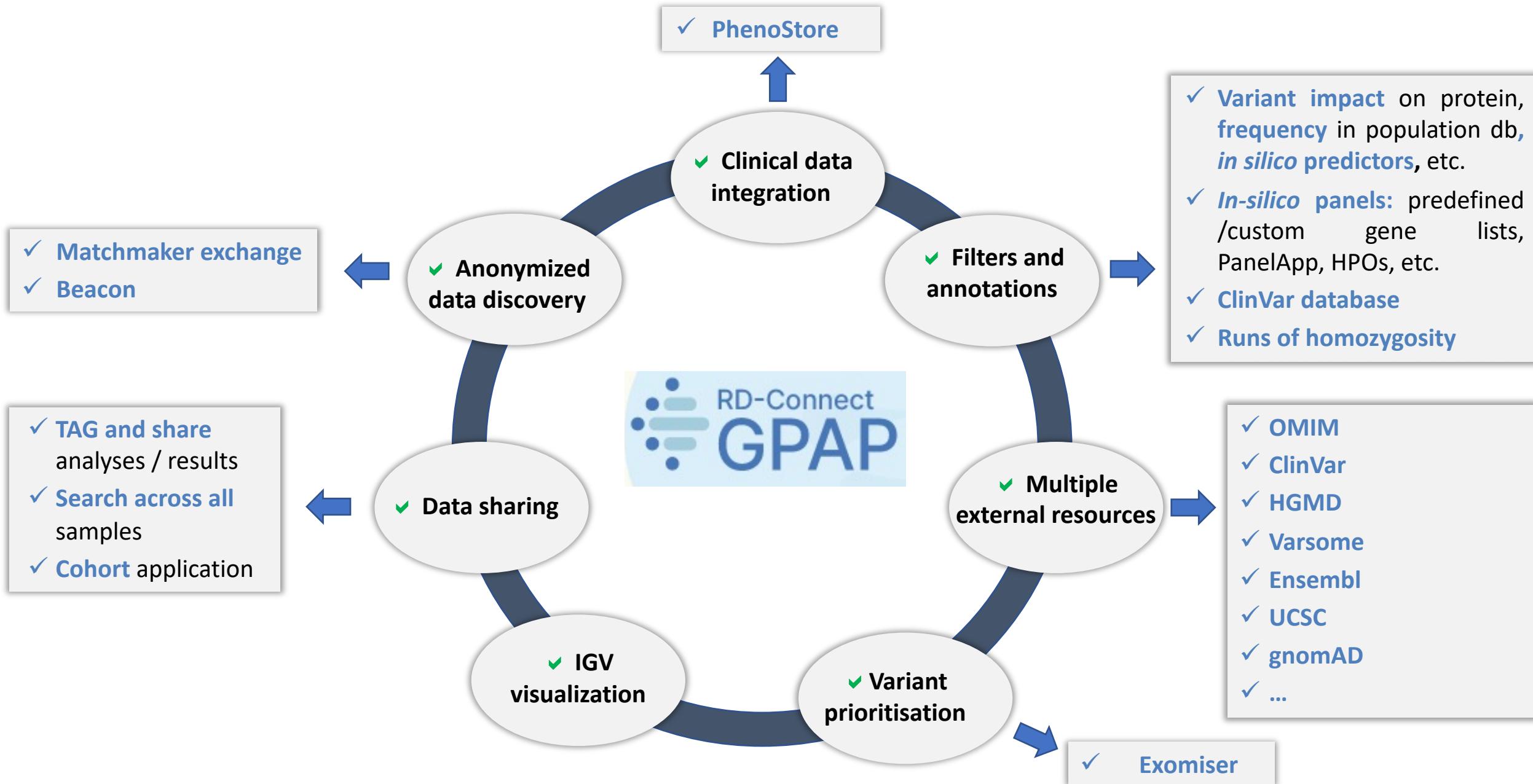
- Upon phenotypic and molecular information

Genomic data module

cnag

centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

CRG[®]
Centre
for Genomic
Regulation



Start analyzing

To provide you with the best experience, select the type of analysis that you want to perform

[Case Analysis](#)[Cohort Analysis](#)[Search Across All Experiments](#)[Patient Matchmaking](#)[Clinical Report](#)

Resume previous work

Visible studies created by you or your network

| Study Name | Description | Created on | Permission |
|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Search 7 records... |

Start analyzing

To provide you with the best experience, select the type of analysis that you want to perform

Case Analysis

Cohort Analysis

Search Across All Experiments

Patient Matchmaking

Clinical Report

Resume previous work

Visible studies created by you or your network

| Study Name | Description |
|---------------------|---------------------|
| Search 7 records... | Search 7 records... |



New GPAP Analysis - Case Study

Select experiments, inheritance type and filters to start

Hello! We are going to create the first analysis
of a GPAP Study

A study is a collection of analyses (experiments(s) and inheritance type). You can add the name and a short description for this study. Define the visibility level, and you are ready to go!

Study Name
 write a title for this new study

Description of the study

Select
 Private
Please select study visibility

 SKIP THIS STEP

New GPAP Analysis - Case Study

Select experiments, inheritance type and filters to start

1 2 3 4

Select Experiments Review Phenotypic and Genetic Information Inheritance Mode Apply Variant Filters

→ NEXT

SELECT EXPERIMENTS TO ANALYZE

Select Experiments by:

Local ExperimentID GPAP Experiment ID

Local Participant ID PhenoStore ID

Family ID

Automatically add family members

Search by Local Experiment ID

Enter Index Case code

→ Family members automatically added by default

No Experiments selected yet.

New GPAP Analysis - Case Study

Select experiments, inheritance type and filters to start



Select Experiments



Review Phenotypic and Genetic Information



Inheritance Mode



Apply Variant Filters

← BACK

→ NEXT

SELECTED EXPERIMENTS - PHENOTYPIC AND GENETIC INFORMATION

Participants included in the study

| | | |
|----------|----------|----------|
| | | |
| P0007507 | P0007508 | P0007509 |

Local ID: Case4C

CASE4C P0007507 Index

Sex: M

Affected: Affected

Life status: Not Available

Consanguinity: NA

Diagnosis: Not Available

Inheritance: Not Available

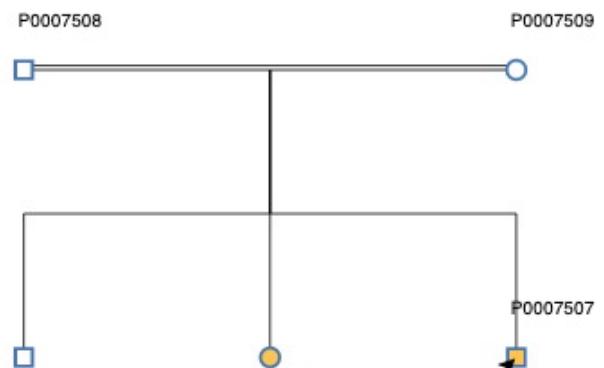
Phenotypes (obs.): Microcephaly, Seizures, Motor delay, Vomiting, Hepatomegaly, Motor deterioration, Developmental regression, Increased CSF lactate, Increased CSF protein, Lactic acidosis, Sensory axonal neuropathy, Myoclonic spasms, Decreased body weight, Intellectual disability, progressive, Motor axonal neuropathy, Neurodevelopmental delay, Cognitive impairment and Epileptic encephalopathy

Phenotypes (not obs.): Developmental stagnation

Disorders: Not Available

Genes: MT-ATP6(status: NA)

What you need to know



The index case has never been analyzed

Tagged variants: 1

Likely pathogenic MT-ATP6,chrom MT, pos 8860
21/9/2022

Additional Information

No runs of homozygosity were calculated for these experiments

New GPAP Analysis - Case Study

Select experiments, inheritance type and filters to start



Select Experiments



Review Phenotypic and Genetic Information



Inheritance Mode



Apply Variant Filters

← BACK

→ NEXT

INHERITANCE MODE

Please, bear in mind that the predefined inheritances take into account the reported affected status of the family members which in some cases may force unrealistic configurations. In such cases, please use the custom configuration to formulate your hypothesis

Select a simulated inheritance configuration:

- Autosomal Recessive (both simple and c-het)
- Autosomal Dominant
- Autosomal Dominant - de novo (parents required)
- X-Linked Recessive
- X-Linked Dominant

Precomputed inheritance modes (more than one can be selected)

or select yourself the configuration details

- Custom Inheritance

Configuration details

Autosomal Dominant - de novo (parents required)

Select whether you expect a variant to be homozygous, heterozygous or wild type for each family member

De Novo

| GPAP EXPERIMENT ID | CASE | AFFECTED | <input type="checkbox"/> REF/REF | <input checked="" type="checkbox"/> REF/ALT | <input type="checkbox"/> ALT/ALT | QUALITY SETTINGS 10 30 0.2 - 0.8 |
|--------------------|--------|----------------------------------|-------------------------------------|---|----------------------------------|-------------------------------------|
| Case4C | index | <input checked="" type="radio"/> | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | QUALITY SETTINGS 10 30 0.2 - 0.8 |
| Case4F | father | <input type="radio"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | QUALITY SETTINGS 10 30 0 - 0 |
| Case4M | mother | <input type="radio"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | QUALITY SETTINGS 10 30 0 - 0 |



4

Select Experiments

Review Phenotypic and Genetic Information

Inheritance Mode

Apply Variant Filters

[← BACK](#)[START ANALYSIS](#)

APPLY FILTERS

YOU CAN APPLY MORE FILTERS LATER. THESE FILTERS WILL ALLOW YOU TO REDUCE THE NUMBER OF VARIANTS FROM THE START.

› SELECT A PREDEFINED FILTER

Create or adjust filters

^ Variant Prioritization

Hover over the information icons to see more information about each filter

RESET VARIANT FILTERS

› VARIANT TYPE



› POPULATION



› SNV EFFECT PREDICTIONS



› POSITION SPECIFIC AND RUNS OF HOMOZYGOSITY



› Filter by genes

[← BACK](#)[START ANALYSIS](#)

APPLY FILTERS

YOU CAN APPLY MORE FILTERS LATER. THESE FILTERS WILL ALLOW YOU TO REDUCE THE NUMBER OF VARIANTS FROM THE START.

> SELECT A PREDEFINED FILTER

Create or adjust filters

> Variant Prioritization

^ Filter by genes

Filter by genes associated to HPOs, OMIM, pathways, or listed in PanelApp or upload your own lists

Operator

Union Intersection



CLEAR GENE LIST

In parenthesis, the number of genes selected in each section

Hover over the information icons to see more information about each filter

> PREDEFINED GENE LISTS (0)



> GENE SEARCH (HGNC DATABASE) AND UPLOAD (0)



> DISEASE-RELATED GENES (0)



> SYMPTOMS (0)



> PATHWAYS (0)





Study: NEW GPAP STUDY Analysis: AUTOSOMAL RECESSIVE HOMOZYGOUS Query: query_1 220540 variants

CHROMOSOME SUMMARY



Links

No variant selected

Frequent Links

dbSNP

gnomAD

ClinVar

VarSome

HGMD

UCSC

OMIM

Franklin

> Disease Information

> Variant Information

> Gene Information

> Pathway Information

> Data Discovery

AUTOSOMAL RECESSIVE HOMOZYGOUS X AUTOSOMAL RECESSIVE COMPOUND HETEROZ. X X-LINKED RECESSIVE - INDEX CASE MALE X

1 QUERY, 3 EXP.

1 QUERY, 3 EXP.

1 QUERY, 3 EXP.

QUERY_1(220540) X

NEW QUERY

VIEW APPLIED FILTERS

Analysis & queries included in the study

^ SNV and InDels (220540)

Variants table

VARIANTS

| Gene Info | | Variant Info | | | | Clinical Ass. | | Population | | Predictors | | | |
|----------------|----------------------|---------------|-------------------|------------------|---------------------|----------------------|---------|----------------|-----------|------------|------------|----------------------|-----------------------|
| Gene | Transcript Biotype | Effect Impact | Nucleotide change | Aminoacid Change | Consequence | OMIM | ClinVar | Internal Freq. | GnomAD AF | CADD Pred. | SIFT Pred. | Polyphen2 Hvar Pred. | Mutation Taster Pred. |
| RP11-206L10.9 | lincRNA | MODIFIER | n.142+318N>C | NA | intron_variant,n... | Retinitis pigment... | NA | 0.66279 | 0 | NA | NA | NA | NA |
| RP11-206L10.10 | processed_transcript | MODIFIER | NA | NA | downstream_gene... | Retinitis pigment... | NA | 0.66279 | 0 | NA | NA | NA | NA |
| RP11-206L10.8 | processed_transcript | MODIFIER | n.145+933N>G | NA | intron_variant,n... | Retinitis pigment... | NA | 0.66279 | 0 | NA | NA | NA | NA |
| AL669831.1 | protein_coding | MODIFIER | NA | NA | upstream_gene... | NA | NA | 0.66279 | 0 | NA | NA | NA | NA |



Study: NEW GPAP STUDY Analysis: AUTOSOMAL RECESSIVE HOMOZYGOUS Query: query_1 220540 variants



CHROMOSOME SUMMARY



External Links

Click to navigate to resource

Frequent Links

- dbSNP (rs)
- gnomAD (pos)
- ClinVar (pos)
- VarSome (pos)
- HGMD (gene)
- UCSC (pos)
- OMIM (gene)
- Franklin (pos)

> Disease Information

> Variant Information

> Gene Information

> Pathway Information

> Data Discovery

AUTOSOMAL RECESSIVE HOMOZYGOUS X AUTOSOMAL RECESSIVE COMPOUND HETEROZ. X X-LINKED RECESSIVE - INDEX CASE MALE X

1 QUERY, 3 EXP. 1 QUERY, 3 EXP. 1 QUERY, 3 EXP.

QUERY_1(220540) X

NEW QUERY



VIEW APPLIED FILTERS

^ SNV and InDels (220540)

Variant extra information

Gene Name: *C1orf170* CHR: 1 Pos.: 907170 REF: AG ALT: A

LABEL

TAG

EXPERIMENTS GT

POPULATION

PREDICTORS

PREVIOUS TAGS

IGV BROWSER

| | Gene | Transcript Biotype | Effect Impact | Nucleotide change | Aminoacid Change | Consequence | OMIM | ClinVar | Internal Freq. | GnomAD AF | CADD Pred. | SIFT Pred. | Polyphen2 Hvar Pred. | Mutation Taster Pred. |
|---|-----------------|----------------------|---------------|-------------------|------------------|---------------------|----------------------|---------|----------------|-----------|------------|------------|----------------------|-----------------------|
| ▼ | RP11-206L10.10 | processed_transcript | MODIFIER | NA | NA | downstream_gene... | Retinitis pigment... | NA | 0.66279 | 0 | NA | NA | NA | NA |
| ▼ | RP11-206L10.8 | processed_transcript | MODIFIER | n.145+933N>G | NA | intron_variant,n... | Retinitis pigment... | NA | 0.66279 | 0 | NA | NA | NA | NA |
| ▼ | AL669831.1 | protein_coding | MODIFIER | NA | NA | upstream_gene... | NA | NA | 0.66279 | 0 | NA | NA | NA | NA |
| ^ | C1orf170 | protein_coding | MODIFIER | NA | NA | downstream_gene... | NA | NA | 0.66279 | 0 | NA | NA | NA | NA |

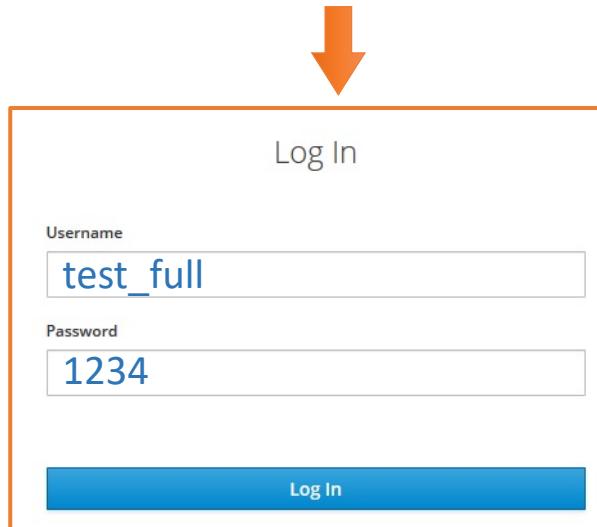
Genome-Phenome Analysis Platform

<https://playground.rd-connect.eu/nextgpap/#/>

DATA SUBMISSION DATA ANALYSIS DATA MANAGEMENT
GPAP Home / Data Analysis / Genomic Analysis



GUIDELINES CONTACT WELCOME GUEST LOGIN



An orange-bordered box containing a log-in form. At the top is the text "Log In". Below it is a "Username" field containing "test_full". Below that is a "Password" field containing "1234". At the bottom is a blue "Log In" button.

Log In

Username

test_full

Password

1234

Log In

cnag

centre nacional d'anàlisi genòmica
centro nacional de análisis genómico

 **CRG**
Centre
for Genomic
Regulation

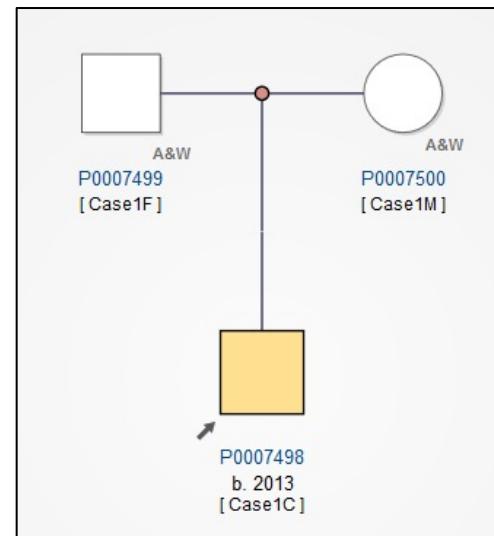


Case 1: description

Identifier: Case1C

| | |
|----------------------------|--------------------------------|
| Gender | Male |
| Age | 5 years |
| Referral | Congenital myasthenic syndrome |
| Onset | Congenital |
| Global pace of progression | Progressive (slow) |

| | |
|------------------------|---|
| Main clinical features | <ul style="list-style-type: none">• Neonatal hypotonia• Distal arthrogryposis• Inability to walk• Recurrent lower respiratory tract infections |
|------------------------|---|





Case 3: description

Identifier: Case3C

Gender Male

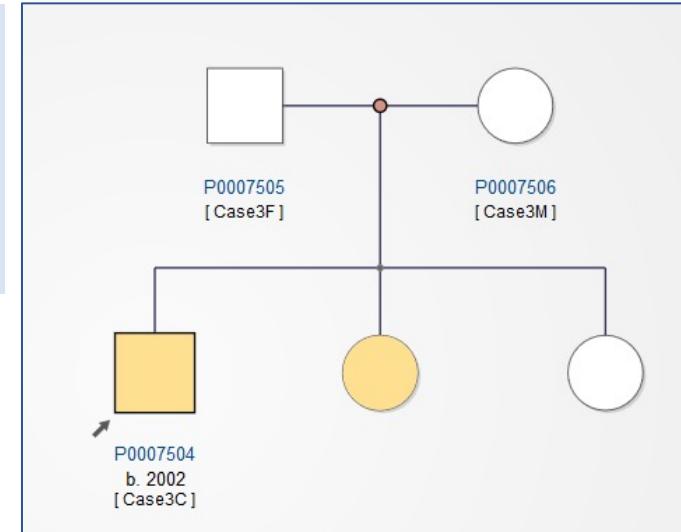
Age 16 years

Referral Muscular dystrophy

Onset Juvenile

Global pace of progression Progressive

- Main clinical features**
- Muscle weakness
 - Dystrophic muscle biopsy
 - Quadriceps muscle atrophy
 - Myalgia



ANSWERS: www.slido.com

Event code: 6293705



Case 15: description

RD-Connect identifier: Case15C

Gender Male

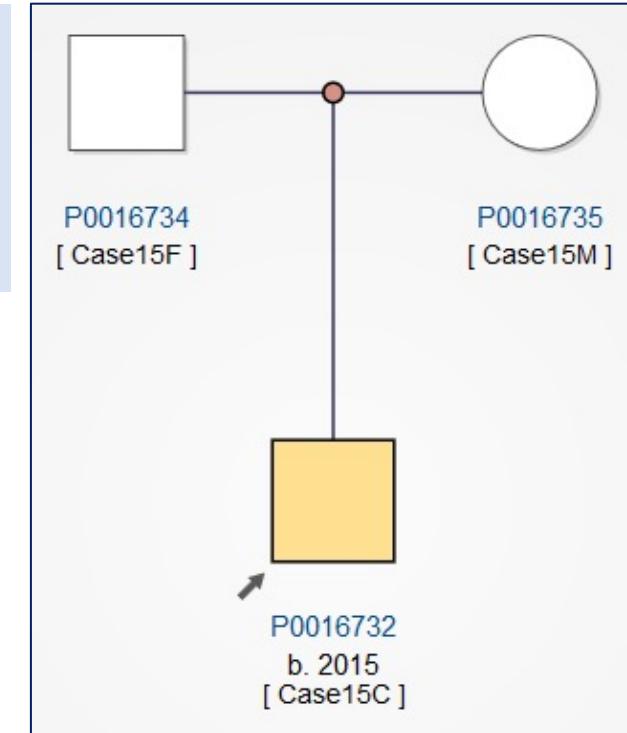
Age 4 years

Referral Metabolic diseases with epilepsy

Onset Infantile onset

Global pace of progression Progressive

- Main clinical features**
- Rhabdomyolysis
 - Metabolic acidosis
 - Seizures
 - Global developmental delay
 - Ventricular tachycardia





Case 2: description

Identifier: Case2C

Gender

Male

Age

38 years

Referral

Macular dystrophy

Onset

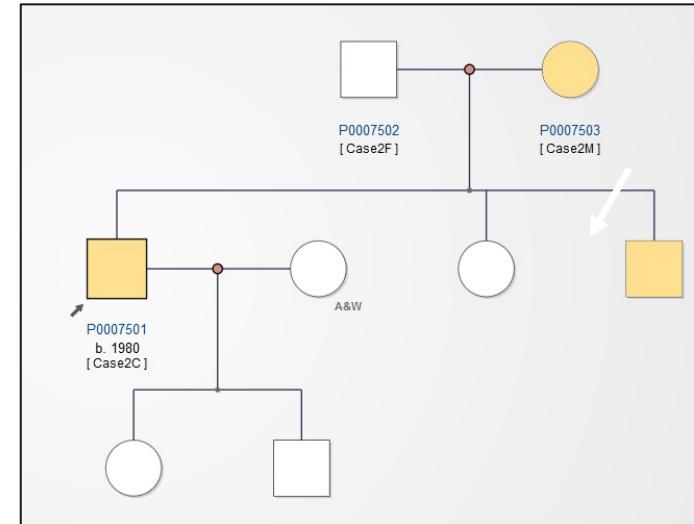
Adult

Global pace of progression

Progressive

Main clinical features

- Progressive visual loss
- Scotoma
- Abnormality of retinal pigmentation



ANSWERS: www.slido.com

Event code: 6293705

Acknowledgments



Bioinformatics Unit: Sergi Beltran

Gemma Bullich
Steven Laurie
Leslie Matalonga
Ida Paramonov
Marc Pauper
Pedro Rodríguez
Luca Zalatnai.

Raúl Ronda
Jordi Morata
Genís Parra
Jean-Rémi Trotta.

Anna Esteve
Marc Dabad
Beatriz Martín.

Davide Piscia
Alberto Corvo
Marcos Fernández
Alejandro García
Oriol López
Daniel Picó
Anastasios
Papakonstantinou.

Matthew Ingham
Jordi Camps
Eloi Casals
Cristina Frías
Óscar Mira



Juan Ramón González

Luis Pérez-Jurado
Mariona Bustamante
Leire Abarregui
Xavier Escribà
Natàlia Carreras



Bea Morte

H.U. Cruces: Isabel Tejada
H.U. Son Espases: Jordi Rosell
H. La Paz María Palomares
H.C.U. Lozano Blesa: Feliciano Ramos
H.C.U. Virgen Arrixaca: Encarna Guillén
H. Clínic: Susana Puig



IDIBAPS:

Antònia Ribes, Frederic Tort
Laia Rodríguez
Glòria Garrabou

IDIBELL-IGTP:

Aurora Pujol
Conxi Lázaro

IMIM:

Luis Pérez-Jurado

URDCat collaborators

