GeneGrid: finding disease-causing variants in NGS data

BTEP Exome-Seq Workshop
Justin Lack, CCBR
Feb 22, 2017

Slides provided by Christian Zinser, Genomatix GmbH

Basis

Genomic variants like SNPs or InDels are of major interest to biologists and clinicians

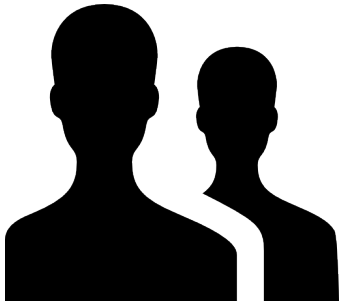
Identifying causal variants is crucial for the diagnostics of rare and common diseases

With NGS technology it is possible to detect millions of variants within an individual genome

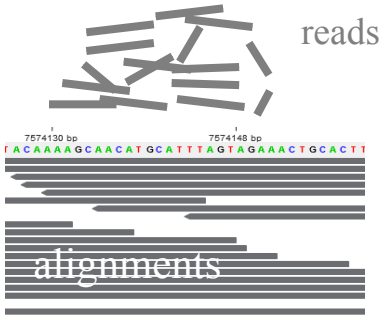
Which are the relevant ones?

General workflow

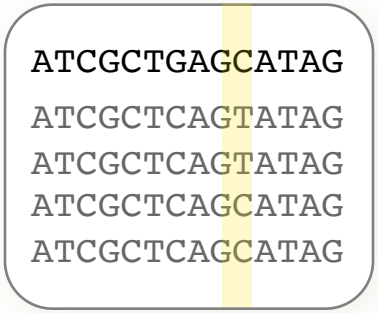
Patient sampling → **Sequencing and alignment** → **Variant calling** → **Variant analysis in GeneGrid**



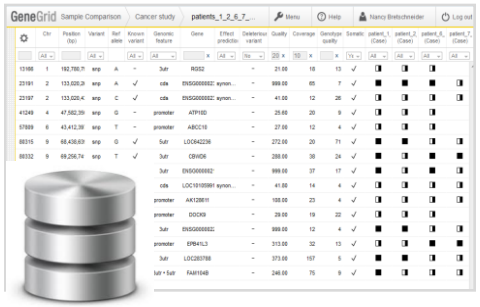
Sequencing library



BAM file with read positions



VCF file with variant information



Short list of potentially interesting variants

"Icon by Daniel Bruce from Flaticon.com"

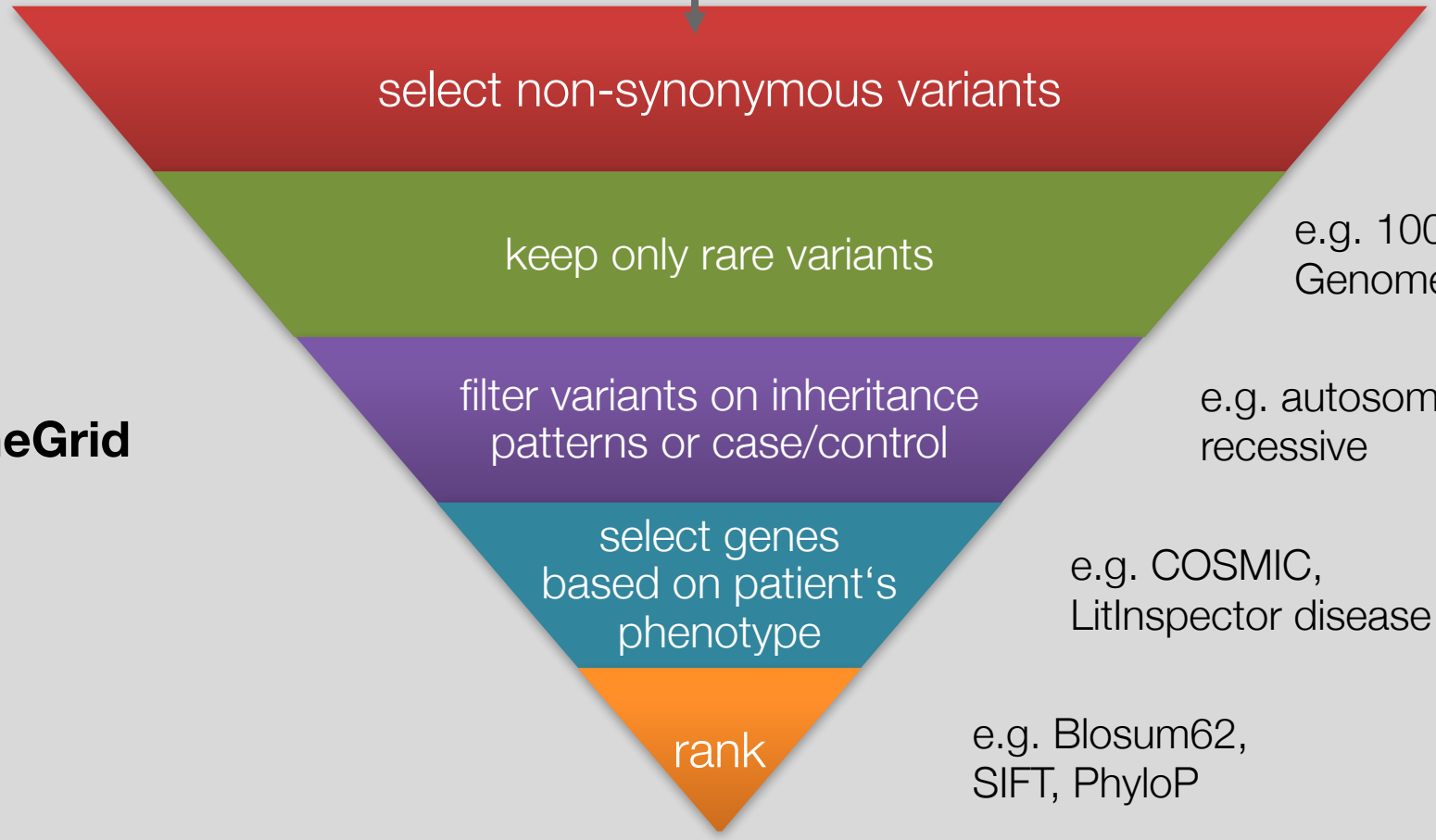
Finding the needle in the haystack

Variant caller

VCF file
SNPs and smal lInDels

up to millions of variants

GeneGrid



e.g. 1000 Genomes Project

e.g. autosomal recessive

e.g. COSMIC, LitInspector disease

e.g. Blosum62, SIFT, PhyloP

10s of causal variants

Selected GeneGrid filter fields

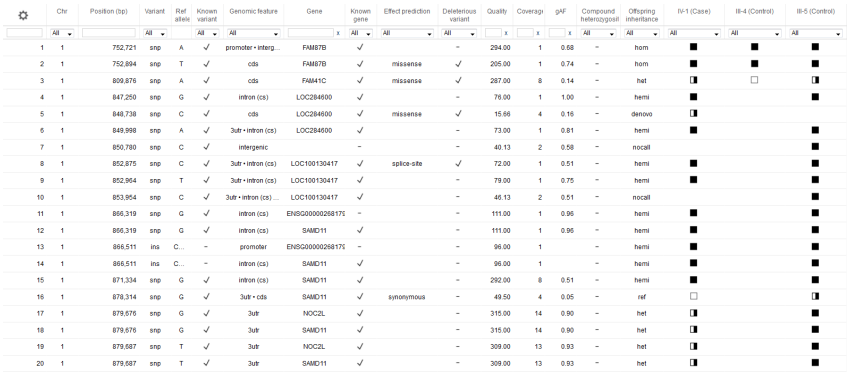
Field	Description
Known gene	Variant is in an annotated gene
Genotype quality	Variant confidence score; higher is better; max 999
gAF	Global minor allele frequency from 1000 Genomes Project; also separate for African, American, Asian, European; <0.01: rare
espMAF	Minor allele frequency from Exome Sequencing Project; also separate for African American, European American; <0.01: rare
exacAF	Alternative allele frequency from the Exome Aggregation Consortium (ExAC) project; <0.01: rare
BLOSUM	AA substitution score; <0: rare; >0: common
SIFT	Sorting Intolerant From Tolerant: predicts effect of AA substitution on protein function; <0.05: affects function
SIFT pred	Yes: damaging prediction based on SIFT score
PhyloP	DNA conservation score; >0: conserved
GERP	Genomic Evolutionary Rate Profiling; RS Score; >0: fewer substitutions than expected -> evolutionary constraint

Selected GeneGrid filter fields

Field	Description
SiPhy Pi, Omega	Overlap with evolutionary constrained elements (two parameter sets)
Literature diseases	Associated diseases based on gene co-citations
Literature tissues	Associated tissues based on gene co-citations
Somatic mutation tissues	Associated tissues; source: COSMIC
Clinical diseases	Associated clinical diseases; source: ClinVar and OMIM
Diagnostic tests	Number of diagnostic tests available for gene
Diagnostic diseases	Associated disease terms based on diagnostic gene tests (gene level)
Clinical significance	Summary of clinical significance based on ClinVar database (position level)
GO processes	Associated GO terms for the domain of cellular components
GO functions	Associated GO terms for the domain of molecular functions
GO components	Associated GO terms for the domain of biological processes

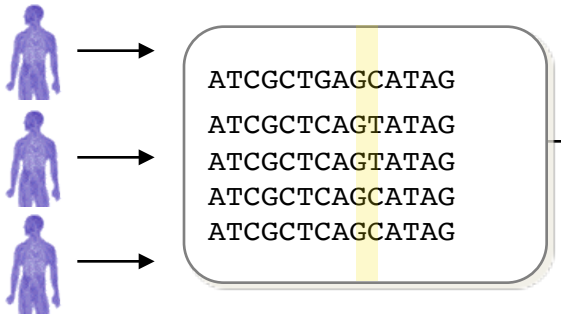
The Genomatix GeneGrid technology

available online



view, compare
filter, sort

Pre-knowledge database
annotations
disease & medical info



import

annotate

SNPs/InDels

GeneGrid database

Results &
Reports

Data sources

Internal sources

Variant annotation	Genomatix
Genome annotation	ElDorado
Text mining (PubMed)	LitInspector
Combined thesaurus (MeSH, NCIt, UMLS)	Genomatix Thesaurus
Pathways and networks	GePS

Data sources

External sources

Alleles and allele frequencies	dbSNP 1000 Genomes Project ESP6500 ExAC
Protein effect predictions	BLOSUM SIFT
Evolutionary conservation	PhyloP GERP 29 Mammals Project
Diagnostic annotation	GTR
Phenotype annotation	ClinVar COSMIC OMIM
Gene ontology	GO
Sequence variant description	HGVS
Regulatory regions	Ensembl Regulatory Build

GeneGrid example 1

Trio analysis

Leber congenital amaurosis

Inherited eye disease

Onset at birth or in early childhood

Blindness or impaired vision with loss of central vision

Genetically heterogeneous

e.g. LCA1 caused by homozygous mutation in GUCY2D
LCA2: RPE65
LCA3: SPATA7
LCA4: AIPL1
etc. (18 types known)

Leber congenital amaurosis

Published in final edited form as:
Nat Genet. 2012 September ; 44(9): 1040–1045. doi:10.1038/ng.2361.

NMNAT1 mutations cause Leber congenital amaurosis

Marni J Falk[1,2,22], **Qi Zhang**[3,4,22], **Eiko Nakamaru-Ogiso**[5], **Chitra Kannabiran**[6], **Zoe Fonseca-Kelly**[3,4], **Christina Chakarova**[7], **Isabelle Audo**[8,9,10,11], **Donna S Mackay**[7], **Christina Zeitz**[8,9,10], **Arundhati Dev Borman**[7,12], **Magdalena Staniszewska**[3,4], **Rachna Shukla**[6], **Lakshmi Palavalli**[6], **Saddek Mohand-Said**[8,9,10,11], **Naushin H Waseem**[7], **Subhadra Jalali**[6,13], **Juan C Perin**[14], **Emily Place**[1,3,4], **Julian Ostrovsky**[1], **Rui Xiao**[15], **Shomi S Bhattacharya**[7,16], **Mark Consugar**[3,4], **Andrew R Webster**[7,12], **José-Alain Sahel**[8,9,10,11,17,18], **Anthony T Moore**[7,12,19], **Eliot L Berson**[4], **Qin Liu**[3,4], **Xiaowu Gai**[20,21,23], and **Eric A. Pierce**[3,4,23]

[1]Division of Human Genetics, Department of Pediatrics, The Children's Hospital of Philadelphia and University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

[2]Division of Child Development and Metabolic Disease, Department of Pediatrics, The Children's Hospital of Philadelphia and University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

[3]Ocular Genomics Institute, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts, USA

[4]Berman-Gund Laboratory for the Study of Retinal Degenerations, Department of Ophthalmology, Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts, USA

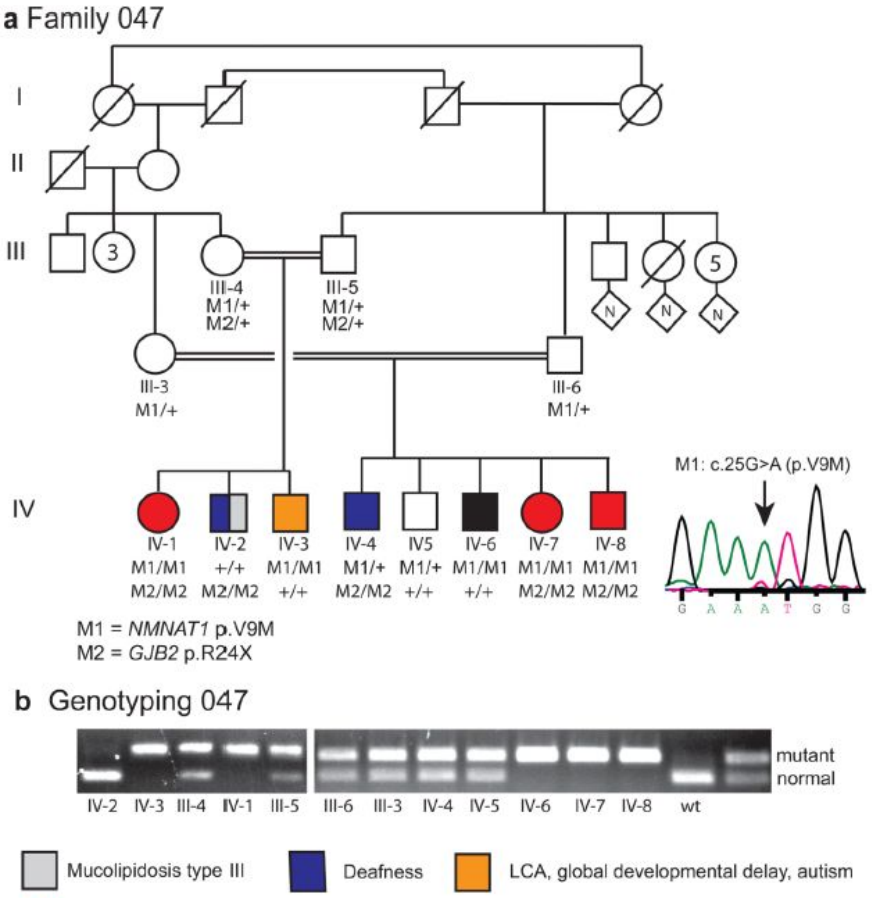
[5]Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

[6]Kallam Anji Reddy Molecular Genetics Laboratory, LV Prasad Eye Institute (LVPEI), Kallam Anji Reddy Campus, LV Prasad Marg, Hyderabad, India

[7]Institute of Ophthalmology, University College of London, London, UK

[8]Institut National de la Santé et de la Recherche Médicale, U968, Paris, France

Consanguineous family



LCA, congenital deafness, global developmental delay, autism

Step 1: load and annotate VCF files

Getting started

The Genomatix GeneGrid technology enables you to quickly reduce millions of small variants to the few or even the single relevant one(s). All known & novel SNPs in your results can be annotated using our extensive annotation. You can filter the list for those variants of interest to you, perform trio analyses, compare case and control sets (using multiple samples) or identify somatic SNPs within minutes. Read more

Variant Annotation

Load your VCF files with samples into GeneGrid to be automatically annotated.

Sample Comparison

Find the relevant small variants and identify disease-causing mutations by comparing samples.

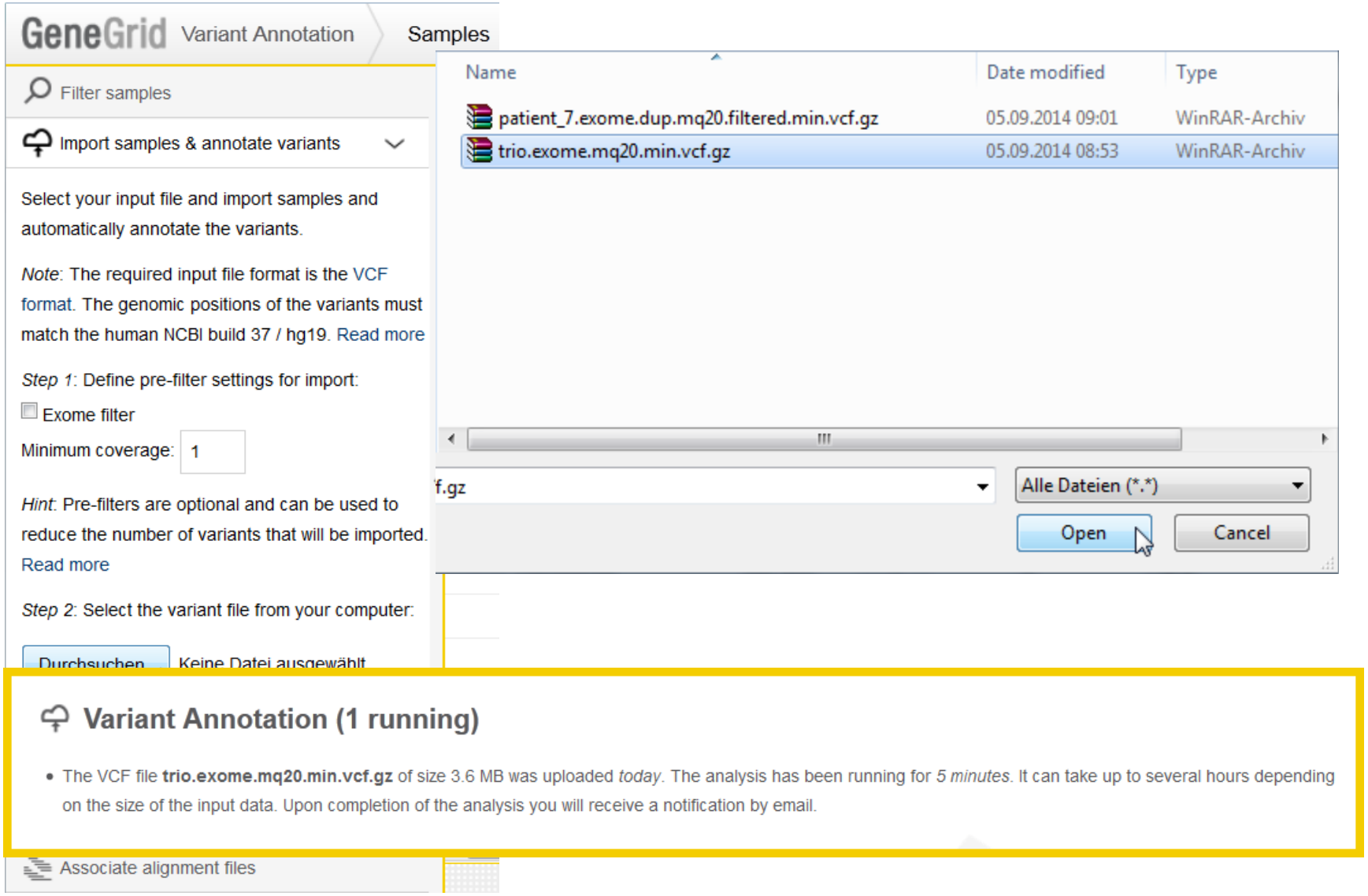
Genome Browser

Browse the human genome in context of your variants of interest and explore publicly available data.

Pathway System (GePS)

Browse, search and load canonical pathways and visualize affected genes on pathway level.

Step 1: load and annotate VCF files



Step 2: sample comparison

Getting started

The Genomatix GeneGrid technology enables you to quickly reduce millions of small variants to the few or even the single relevant one(s). All known & novel SNPs in your results can be annotated using our extensive annotation. You can filter the list for those variants of interest to you, perform trio analyses, compare case and control sets (using multiple samples) or identify somatic SNPs within minutes. Read more

Variant Annotation

Load your VCF files with samples into GeneGrid to be automatically annotated.

Sample Comparison

Find the relevant small variants and identify disease-causing mutations by comparing samples.

Genome Browser

Browse the human genome in context of your variants of interest and explore publicly available data.

Pathway System (GePS)

Browse, search and load canonical pathways and visualize affected genes on pathway level.

Step 2: sample comparison

Sample ID	Input file	Sample	Number of non-ref variants	Class	Activated	Associated alignments	Exome filter	Minimum coverage
	x	x	x	x	All	All	All	All	x
Open	LCA047_Trio_Demo.vcf	III-5			✓	✓	-	1
1277	LCA047_Trio_Demo.vcf	III-4			✓	✓	-	1
1276	LCA047_Trio_Demo.vcf	IV-1			✓	✓	-	1

Step 1: Select the type of comparison study:

Trio | Cancer | Other

Step 2: Assign the samples to the groups:

Hint: Just drag an activated sample from the table on the right side and drop it in one of the two groups below. Read more

Offspring (0 assigned, requires 1 more)

	Sample	#	ID

Parents (0 assigned, requires 2 more)

	Sample	#	ID

Study name:

Please enter a result name

Submit

Compare samples

Step 1: Select the type of comparison study:

Trio | Cancer | Other

Step 2: Assign the samples to the groups:

Offspring (1 assigned)

	Sample	#	ID
1	IV-1	77,084	1276

Parents (2 assigned)

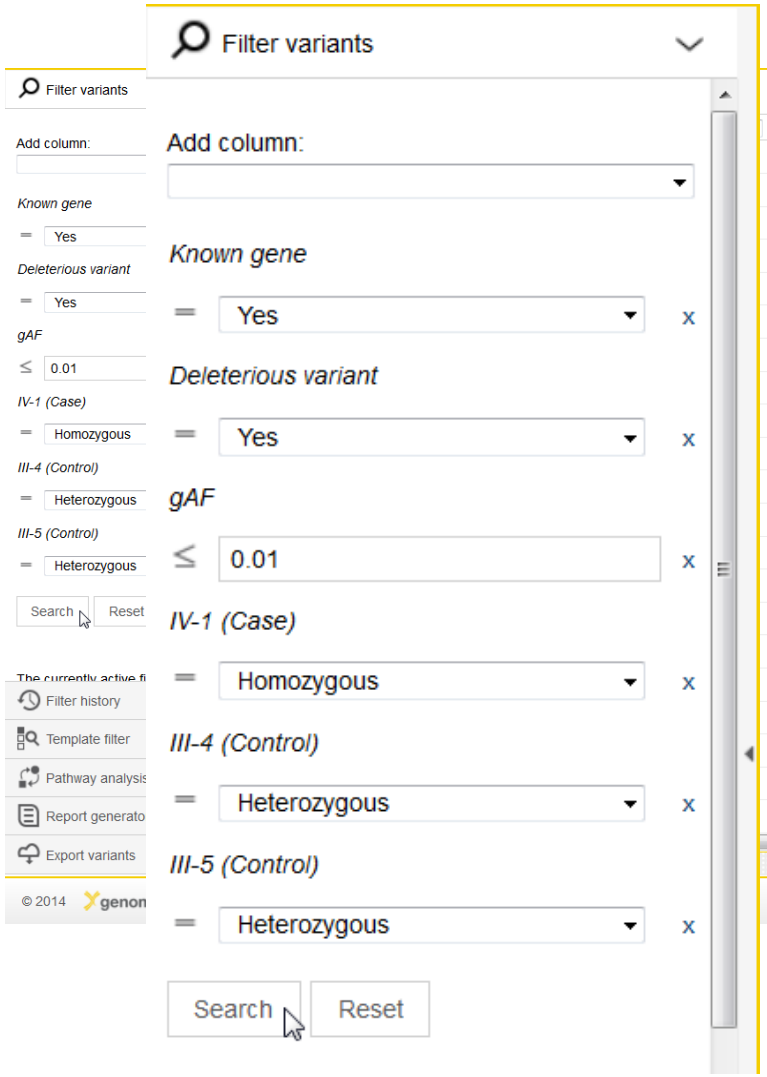
	Sample	#	ID
1	III-4	41,217	1277
2	III-5	76,762	1278

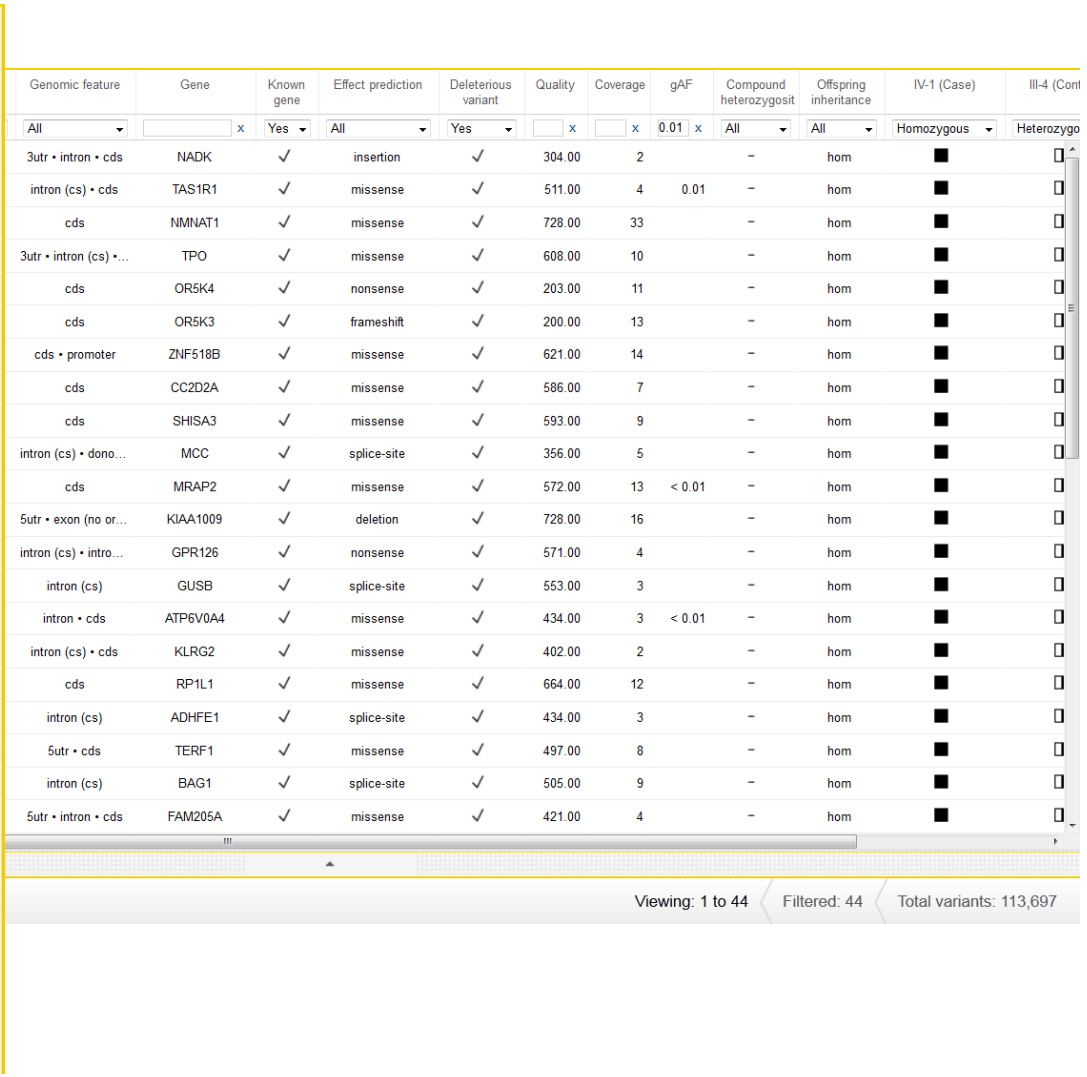
Study name:

LCA_Trio

Submit

Step 3: filtering



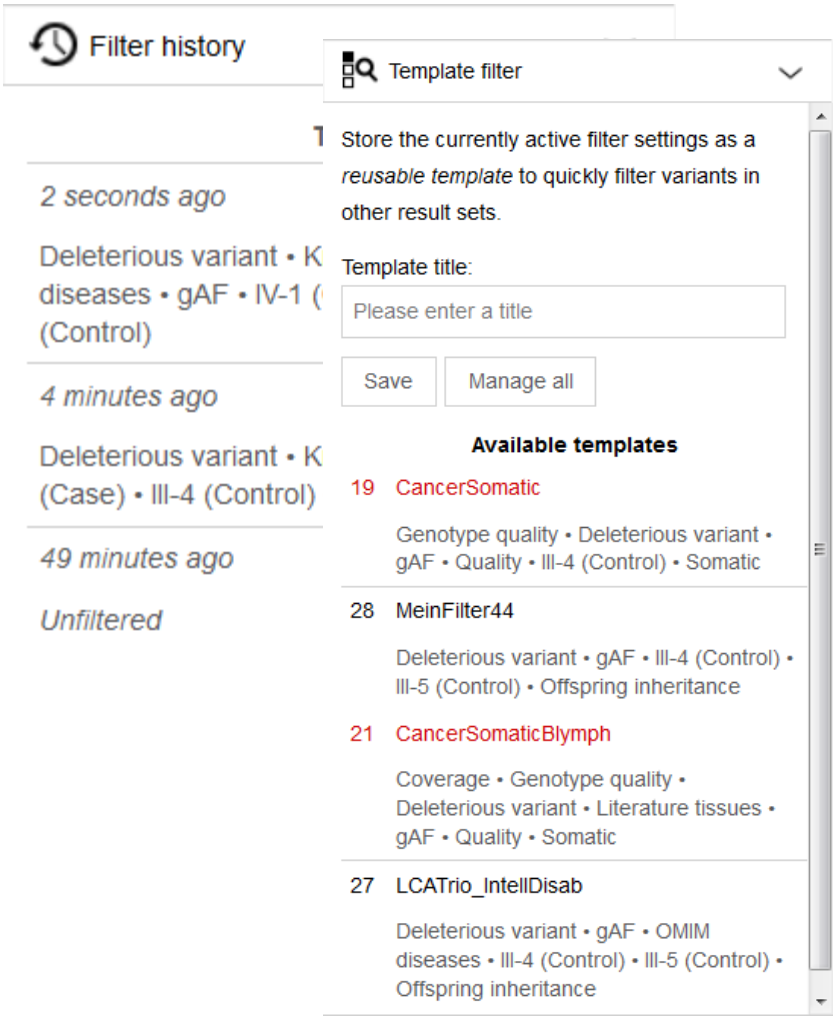


Genomic feature	Gene	Known gene	Effect prediction	Deleterious variant	Quality	Coverage	gAF	Compound heterozygosit	Offspring inheritance	IV-1 (Case)	III-4 (Cont
All		Yes	All	Yes			0.01	All	All	Homozygous	Heterozygo
3utr • intron • cds	NADK	✓	insertion	✓	304.00	2		-	hom	■	
intron (cs) • cds	TAS1R1	✓	missense	✓	511.00	4	0.01	-	hom	■	
cds	NMNAT1	✓	missense	✓	728.00	33		-	hom	■	
3utr • intron (cs) •...	TPO	✓	missense	✓	608.00	10		-	hom	■	
cds	OR5K4	✓	nonsense	✓	203.00	11		-	hom	■	
cds	OR5K3	✓	frameshift	✓	200.00	13		-	hom	■	
cds • promoter	ZNF518B	✓	missense	✓	621.00	14		-	hom	■	
cds	CC2D2A	✓	missense	✓	586.00	7		-	hom	■	
cds	SHISA3	✓	missense	✓	593.00	9		-	hom	■	
intron (cs) • dono...	MCC	✓	splice-site	✓	356.00	5		-	hom	■	
cds	MRAP2	✓	missense	✓	572.00	13	< 0.01	-	hom	■	
5utr • exon (no or...	KIAA1009	✓	deletion	✓	728.00	16		-	hom	■	
intron (cs) • intro...	GPR126	✓	nonsense	✓	571.00	4		-	hom	■	
intron (cs)	GUSB	✓	splice-site	✓	553.00	3		-	hom	■	
intron • cds	ATP6V0A4	✓	missense	✓	434.00	3	< 0.01	-	hom	■	
intron (cs) • cds	KLRG2	✓	missense	✓	402.00	2		-	hom	■	
cds	RP1L1	✓	missense	✓	664.00	12		-	hom	■	
intron (cs)	ADHFE1	✓	splice-site	✓	434.00	3		-	hom	■	
5utr • cds	TERF1	✓	missense	✓	497.00	8		-	hom	■	
intron (cs)	BAG1	✓	splice-site	✓	505.00	9		-	hom	■	
5utr • intron • cds	FAM205A	✓	missense	✓	421.00	4		-	hom	■	

Viewing: 1 to 44 | Filtered: 44 | Total variants: 113,697

Filter history and template filters

Speed up the analysis and facilitate sample comparisons



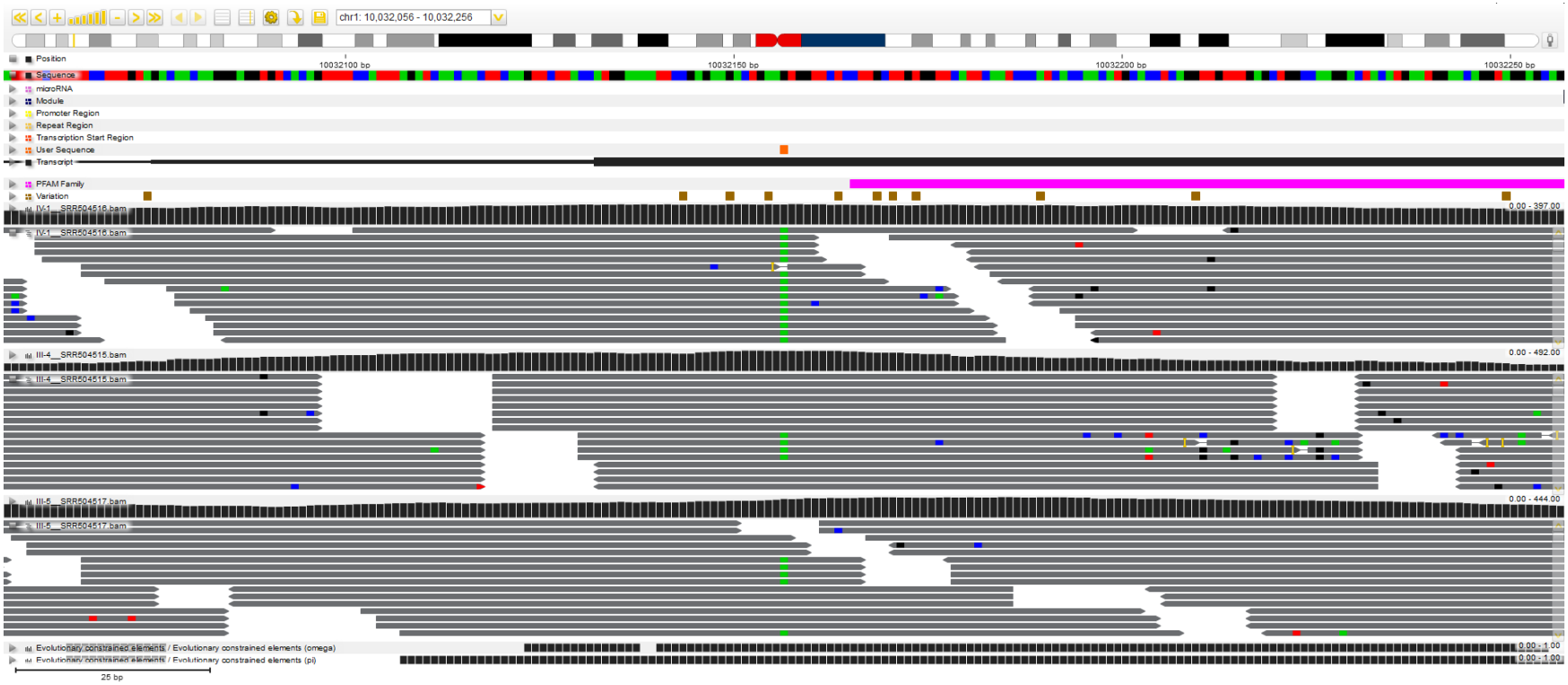
Combine filters that are routinely used, e.g.

- Minimum read coverage
- Genotype quality
- Allele frequency (gAF)
- Phenotype annotation
- Inheritance patterns
- Case / control genotypes

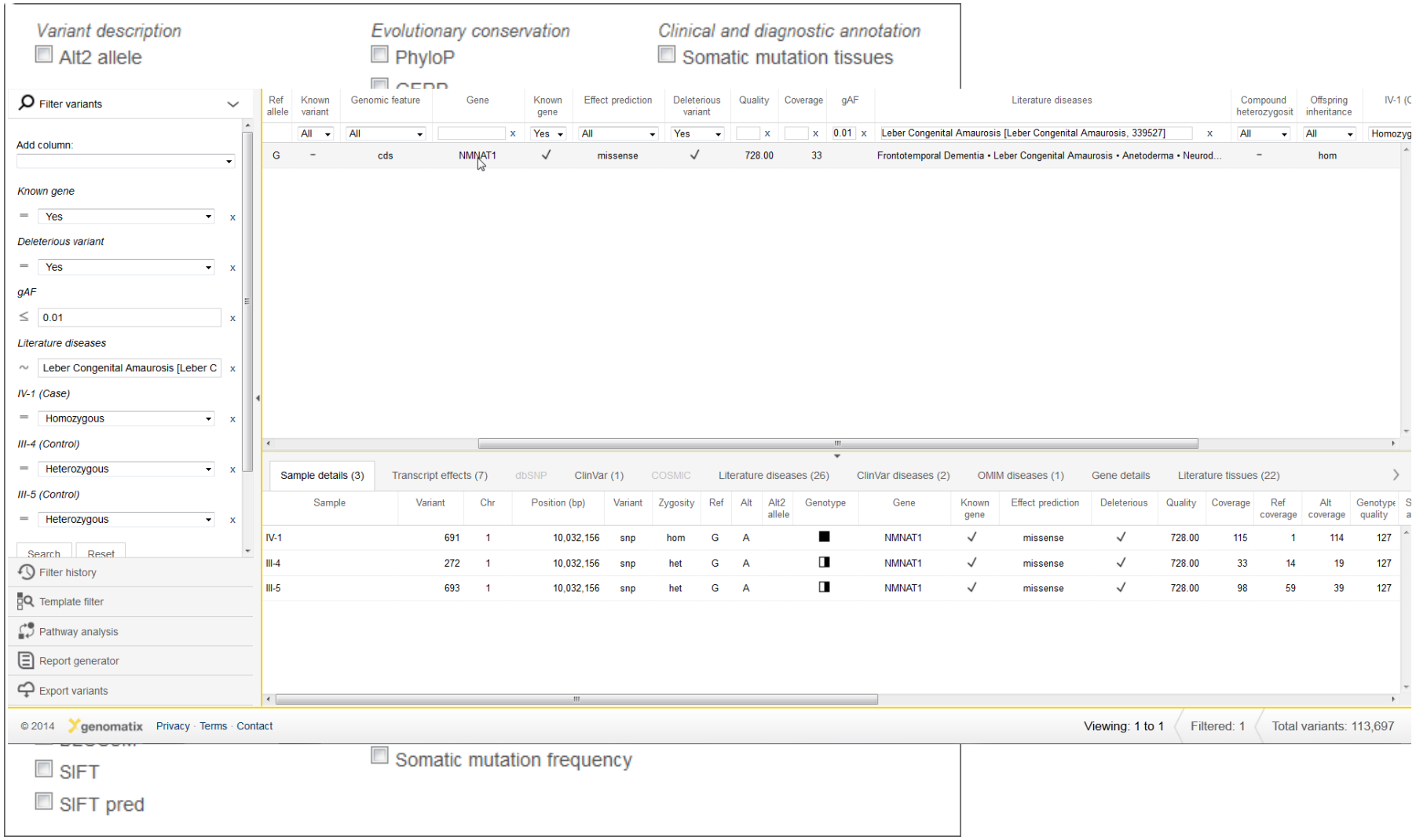
and save as template for use in other comparisons

Vizualisation in genome browser

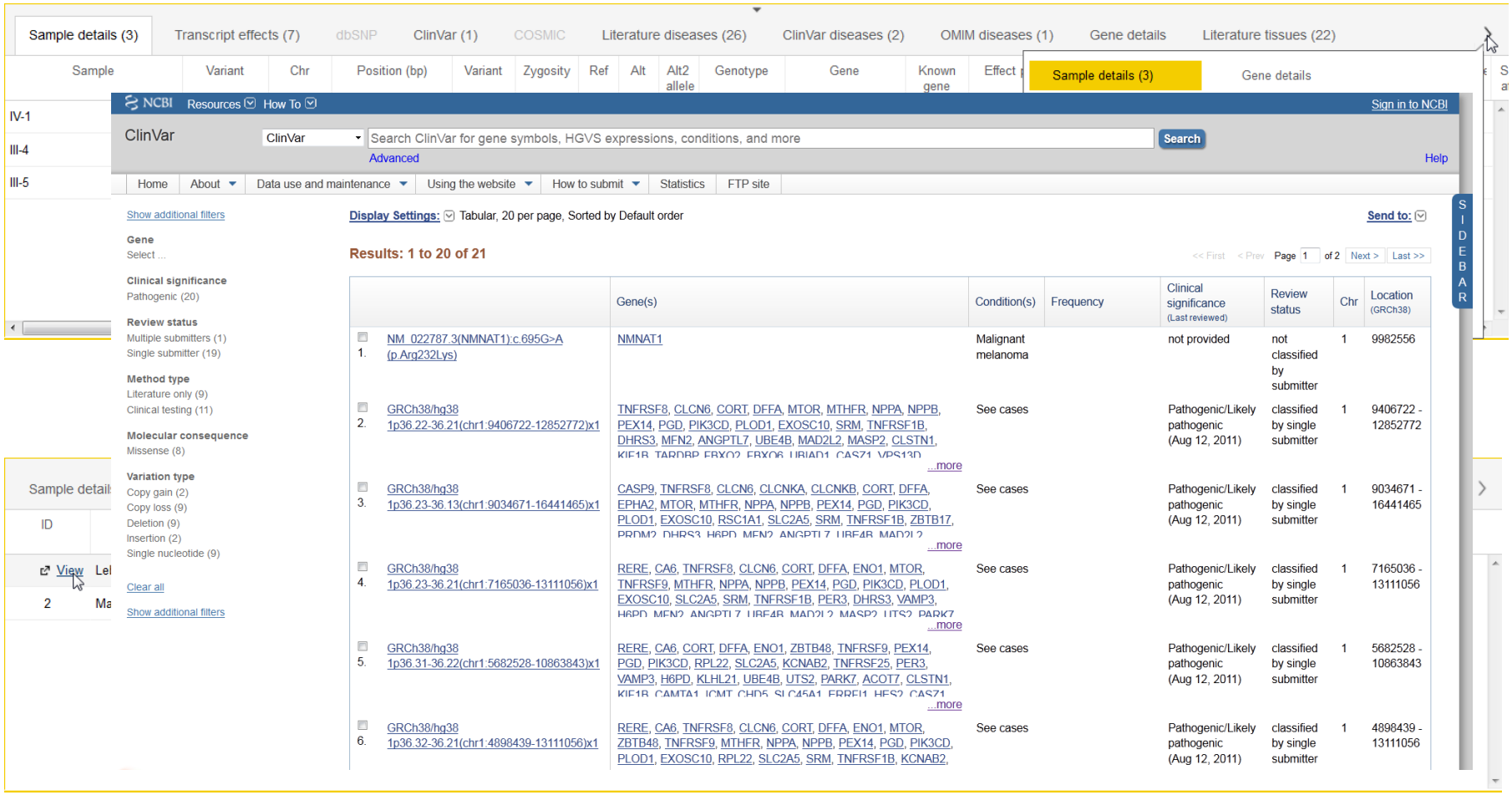
	Chr	Position (bp)	Variant	Ref allele	Known variant	Genomic feature	Gene	Known gene	Effect prediction	Deleterious variant	Quality	Coverage	gAF	Literature diseases
	All		All		All	All	x	All	All	Yes	x	x	0.01 x	Leber Congenital Amaurosis [Leber Congenital Amaurosis, 339
Browse	1	10,032,156	snp	G	-	cds	NMNAT1	✓	missense	✓	728.00	33		Frontotemporal Dementia • Leber Congenital Amaurosis • Anet



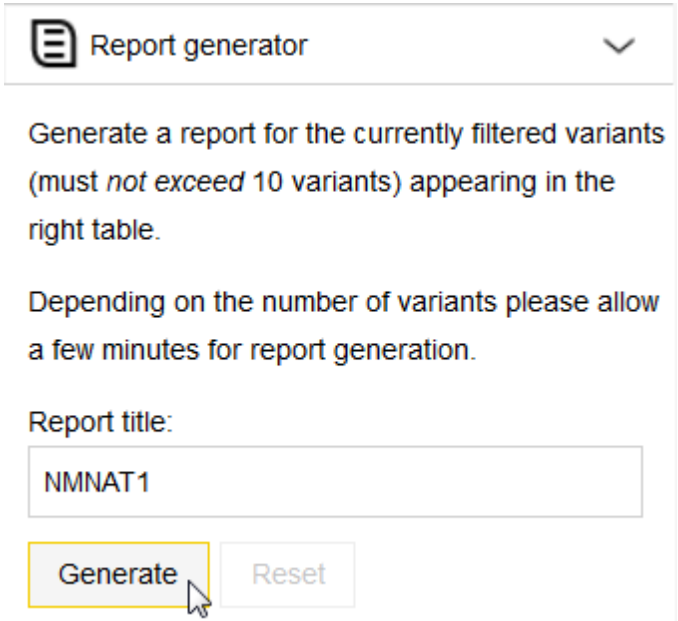
Additional annotation / filter columns



Annotation detail views and links



Report generator



NMNAT1

October 27, 2014

DNA variants

Summary

This report consists of one variant:

1. *NMNAT1* | SNV | c.25G>A / p.(Val9Met) | pathogenic

p.(Val9Met) in *NMNAT1*

Variant description

The indicated single nucleotide variant (SNV) is located on chromosome 1 at position 10,032,156 bp. It overlaps the coding sequence of at least one transcript of gene *NMNAT1* . The reference allele for this variant is G, whereas the alternative allele is A.

Variant quality (smallest value across all samples in comparison)

- Minimal depth of coverage (COV): 33 reads
- Minimal quality value for the assertion of the alternative allele (QUAL): 728.00
- Minimal conditional genotype quality for this site being a variant (GQ): 127

Case distribution (1/1)

- Homozygous variant:
  1. IV-1 (COV: 115 reads, QUAL: 728.00, GQ: 127)

Control distribution (2/2)

- Heterozygous variant:
  1. III-4 (COV: 33 reads, QUAL: 728.00, GQ: 127)
  2. III-5 (COV: 98 reads, QUAL: 728.00, GQ: 127)

Predicted molecular effects on protein

This variant is predicted to be a missense mutation which alters the protein's amino acid from valine (Val) to methionine (Met). The prediction for p.(Val9Met) is based on 7 annotated transcripts for that gene locus. The BLOSUM62 substitution matrix reports a score of 1 for this alteration.

Known variant

The variant was not previously reported in dbSNP.

Clinical significance

The mutation g.10032156G>A was clinically associated with Leber congenital amaurosis 9 and classified as pathogenic (ID RCV000030771).

Allele frequencies in populations

There was no alternative allele frequency listed for this variant in the 1000 genomes data set. There was no minor allele frequency listed for this variant in the NHLBI GO Exome Sequencing Project (ESP6500) data set.

Computational protein effect prediction

The amino acid substitution is predicted to be damaging (SIFT score: 0.01). The sequence alteration is evaluated to be disease-causing (MutationTaster score: 0.94). The prediction that this mutation is in highly conserved amino acids within protein-coding sequences is neutral (LRT score: < 0.01). The accuracy of these tools is unknown.

Evolutionary conservation

Variant overlaps with evolutionary constrained element (detected using SiPhy-ω and SiPhy-π statistics). The conservation across 28 species is described with PhyloP (score: 0.98). GERP identifies constrained elements in multiple alignments by quantifying substitution deficits (score: 2.99).

NMNAT1

General information

Data source: EIDorado

Gene description

nicotinamide nucleotide adenylyltransferase 1 (ID 64802)

NMNAT1 is also known as *LCA9, NMNAT, NMNAT1, PNAT1* and located on locus 1p36.22.

Brief summary

This gene encodes an enzyme which catalyzes a key step in the biosynthesis of the coenzyme NAD. The encoded protein is one of several nicotinamide nucleotide adenylyltransferases. Studies in Drosophila and mammalian neurons have shown the encoded protein can confer protection to damaged neurons. This protection requires enzymatic activity which increases NAD levels and activates a nuclear deacetylase which is the protective molecule. Pseudogenes of this gene are located on chromosomes 1, 3, 4, 14 and 15. [provided by RefSeq, Dec 2011]

Export of annotated variants

Export variants

Export the currently filtered list of variants appearing in the right table.

File format:

TSV VCF

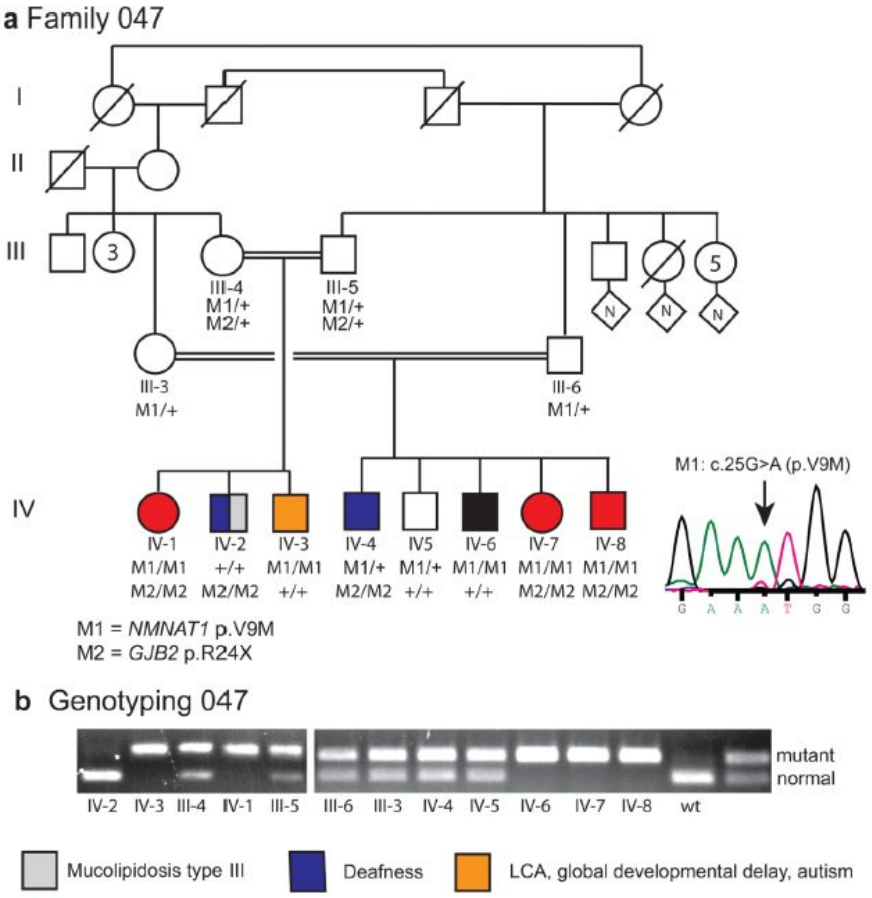
##Number: Consecutively numbered variants
##Contig: Contig numbers in karyotypic order, Genomatix
##Position: Position on the contig, Customer
##Type: Type of variation, Customer
##Reference: Sequence for reference, Customer
##KnownVariant: Known variant, dbSNP, Build 138
##Function: Genomic feature, ElDorado, May 2013
##Gene: Preferred symbol for gene, ElDorado, May 2013
##Category: Predicted effects, Genomatix
##IsDeleterious: Deleterious variant effect prediction, Genomatix
##Quality: Quality value for the assertion of the alternative allele (minimum), Customer
##Coverage: Coverage (minimum), Customer
##MafGenomes: Alternative allele frequency from the whole 1000 genomes project, 1000 Genomes Project, Apr 2012
##LiteratureDiseases: Associated disease terms based on literature mining by Genomatix (gene level), LitInspector, Sep 2014
##ClinVarDiseases: Associated clinical disease terms based on ClinVar database (gene level), ClinVar, Oct 2014
##OmimDiseases: Associated disease terms based on OMIM database (gene level), OMIM, Oct 2014
##SampleCase: IV-1 (Case)
##SampleControl: III-4 (Control)
##SampleControl: III-5 (Control)

Number	Contig_ChrL	Contig_Cont	Position	Type	Reference	KnownVaria	Function	Gene_Genel	Gene_Genes	Category	IsDeleteriou	Quality	Coverage	MafGenome	LiteratureDiseases	ClinVarDiseases	OmimDiseases	SampleCase	SampleConti	SampleControl_2
290	1	1	1684347	insertion	CCCTCCTCCT	0	3utr,intron,c	65220	NADK	insertion	1	304	2	0	Tuberculosis:Dysplasti	Lung cancer		hom	het	het
607	1	1	6637056	snp	G	1	intron_cs,cd	80835	TAS1R1	missense	1	511	4	0.0093	Ageusia:Herpes Zoster	Malignant melanoma		hom	het	het
746	1	1	10032156	snp	G	0	cds	64802	NMNAT1	missense	1	728	33	0	Frontotemporal Deme	Leber congenital am	Leber congenital am	hom	het	het
11097	2	2	1440078	snp	C	1	3utr,intron_(	7173	TPO	missense	1	608	10	0	Multiple System Atrop	Deficiency of iodide	Thyroid dyshormono	hom	het	het
21791	3	3	98073591	deletion	TAAAAAAA	0	cds	403278	OR5K4	nonsense	1	203	11	0		Malignant melanoma		hom	het	het
21793	3	3	98110406	insertion	GAAAAAAA	1	cds	403277	OR5K3	frameshift	1	200	13	0				hom	het	het
25841	4	4	10446604	snp	T	0	cds,promote	85460	ZNF518B	missense	1	621	14	0		Malignant melanoma		hom	het	het
25872	4	4	15542617	snp	C	1	cds	57545	CC2D2A	missense	1	586	7	0	Meckel syndrome type	Meckel syndrome, ty	COACH syndrome:M	hom	het	het
26408	4	4	42403128	snp	C	1	cds	152573	SHISA3	missense	1	593	9	0	Neoplasms:Cell Transfi	Malignant melanoma		hom	het	het
31745	5	5	112676428	snp	C	1	intron_cs,cs_	4163	MCC	splice-site	1	356	5	0	Colorectal carcinoma:N	Carcinoma of colon:L	Colorectal cancer, so	hom	het	het
39823	6	6	84798956	snp	G	1	cds	112609	MRAP2	missense	1	572	13	0.0005	Obesity:Adrenocortica	Malignant melanoma	Obesity, susceptibili	hom	het	het
39832	6	6	84896313	deletion	TTTCTT	0	5utr,exon,cd	22832	KIAA1009	deletion	1	728	16	0	Communicable Diseases			hom	het	het
40963	6	6	142738314	deletion	CTTCTTTTC	0	intron_cs,int	57211	GPR126	nonsense	1	571	4	0	Peripheral neuropathy	Lung cancer		hom	het	het
44143	7	7	65444359	insertion	TGAGAG	0	intron_cs	2990	GUSB	splice-site	1	553	3	0	Mucopolysaccharidosis	Mucopolysaccharido	Mucopolysaccharido	hom	het	het
46432	7	7	138424359	snp	T	1	intron,cds	50617	ATP6V0A4	missense	1	434	3	0.0009	Renal Tubular Acidosis	Renal tubular acidosi	Renal tubular acidosi	hom	het	het
46467	7	7	139164446	snp	G	1	intron_cs,cd	346689	KLRG2	missense	1	402	2	0	Prostate Cancer			hom	het	het
47886	8	8	10467589	snp	T	1	cds	94137	RP1L1	missense	1	664	12	0	Occult macular dystrop	Malignant melanoma	Occult macular dystr	hom	het	het
49527	8	8	67369047	snp	T	1	intron_cs	137872	ADHFE1	splice-site	1	434	3	0	D-2-hydroxyglutaric ac	Malignant melanoma		hom	het	het
49733	8	8	73937076	snp	T	0	5utr,cds	7013	TERF1	missense	1	497	8	0	Low Grade Prostatic Intraepithelial Neoplasia:Acute leukemias:A			hom	het	het
52162	9	9	33255933	snp	A	0	intron_cs	573	BAG1	splice-site	1	505	9	0	Breast cancer:Invasive Breast Carcinoma:Periampullary Adenocar			hom	het	het

GeneGrid example 2

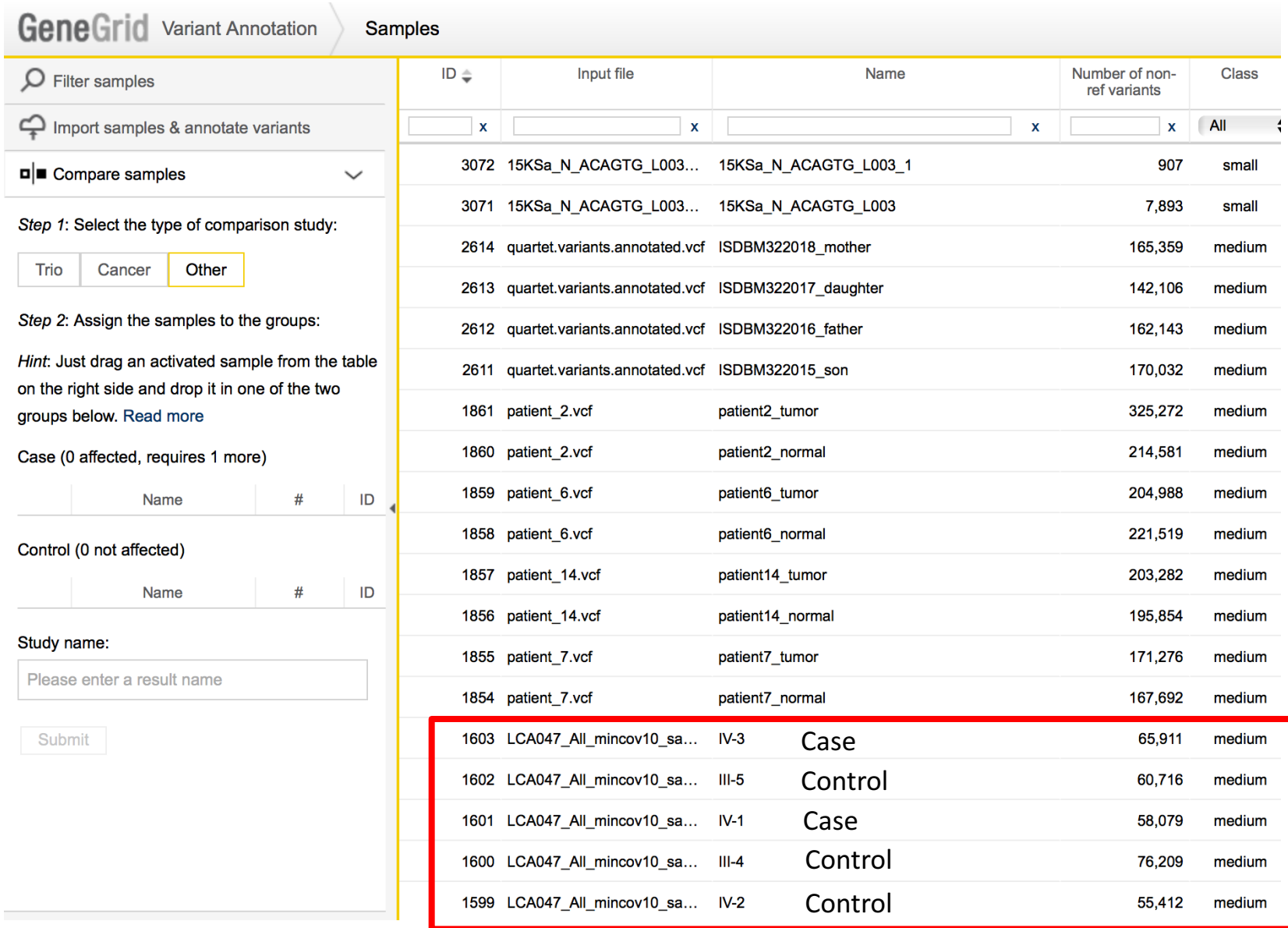
Familial autism analysis

Consanguineous family

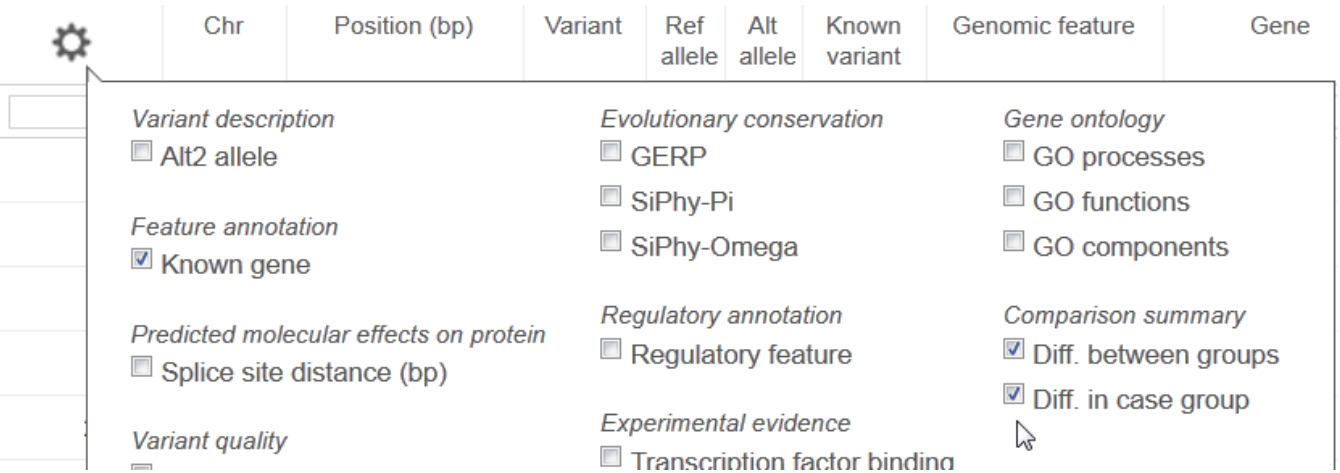


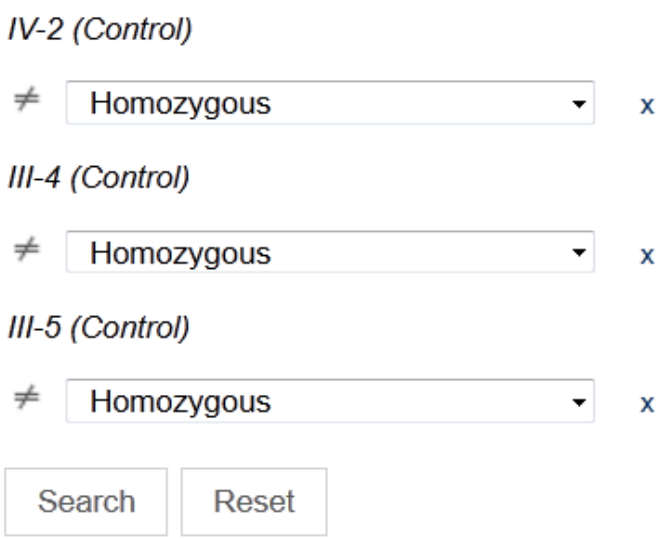
LCA, congenital deafness, global developmental delay, autism

Step 2: sample comparison



Step 2: sample comparison



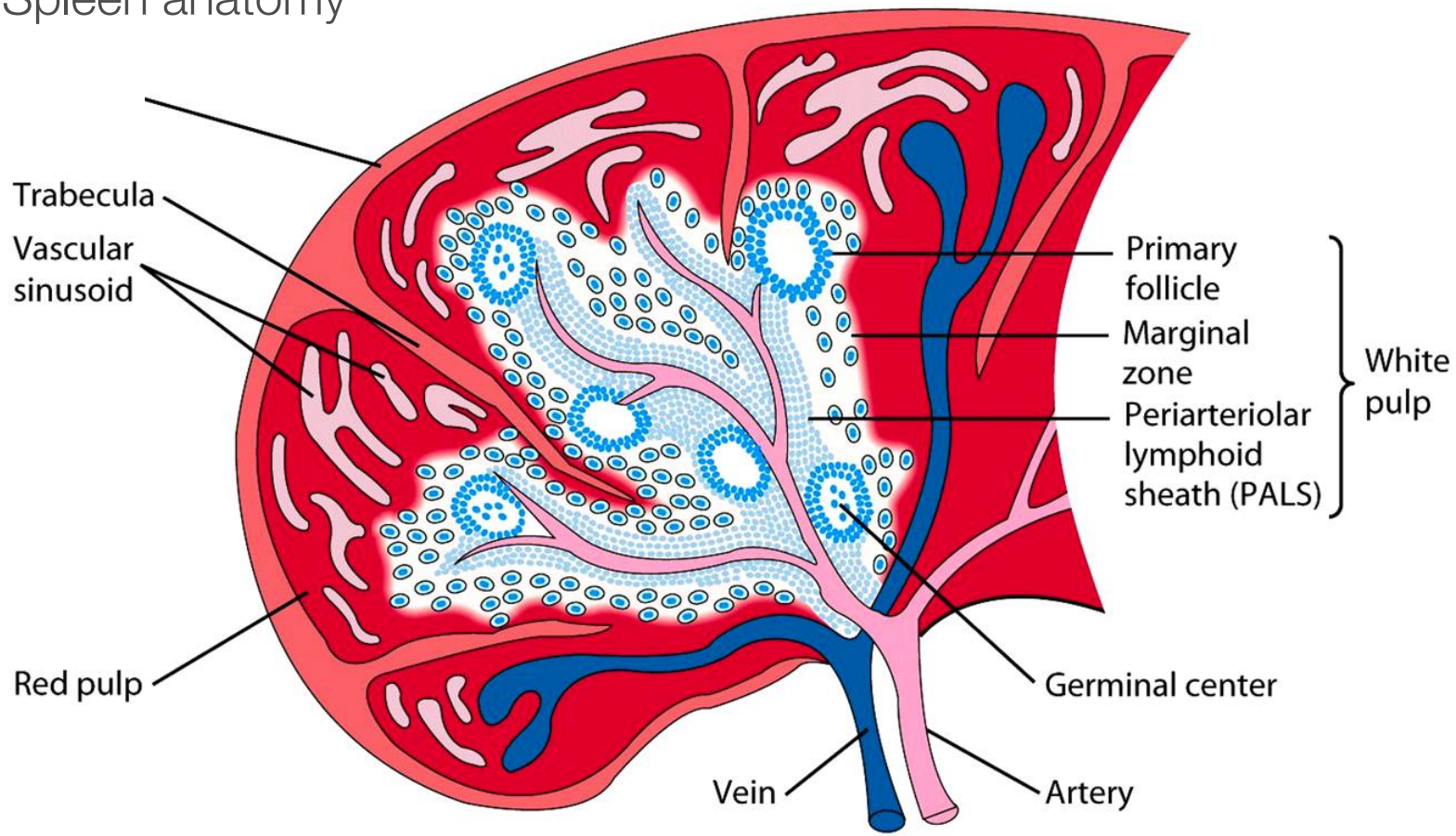


GeneGrid example 3

Cancer analysis

Splenic marginal zone lymphoma (SMZL)

Spleen anatomy

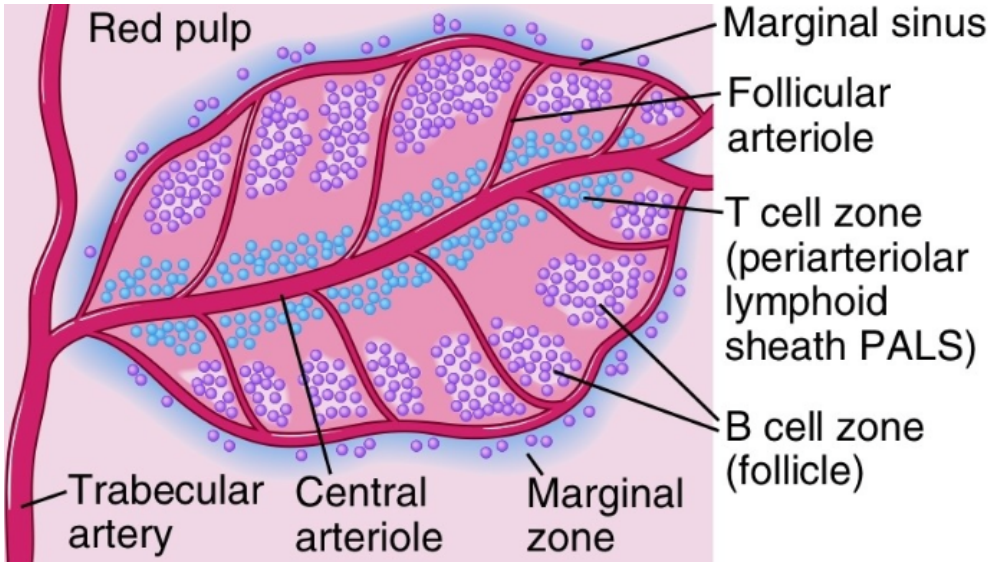


Source: Northern Arizona University http://www2.nau.edu/~fpm/immunology/spleen1.jpg

Splenic marginal zone lymphoma (SMZL)

B cell lymphoma consisting of small lymphocytes and larger blasts

- Starts in white pulp of the spleen
- Invades mantle zone of splenic follicles
- Erodes marginal zone
- Ultimately invades red pulp of the spleen



Frequent clonal rearrangements of immunoglobulin genes

Splenic marginal zone lymphoma (SMZL)

Leukemia (2014) **28**, 1334–1340

www.nature.com/leu

ORIGINAL ARTICLE

Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation

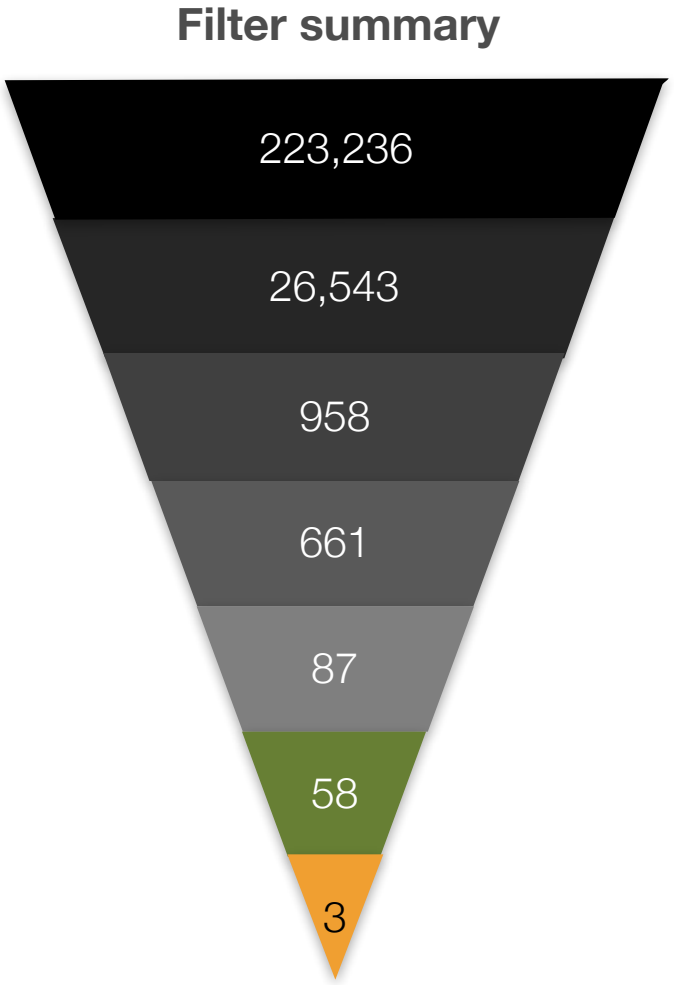
N Martínez[1], C Almaraz[1], JP Vaqué[1], I Varela[2], S Derdak[3], S Beltran[3], M Mollejo[4], Y Campos-Martin[4], L Agueda[3], A Rinaldi[5], I Kwee[5,6,7], M Gut[3], J Blanc[3], D Oscier[8], JC Strefford[9], J Martinez-Lopez[10], A Salar[11], F Sole[12], JL Rodriguez-Peralto[13], C Diez-Tascón[14], JF García[15], M Fraga[16], E Sebastián[17], J Alvés[18], J Menárguez[19], J González-Carreró[20], LF Casado[4], M Bayes[3], F Bertoni[5,21], I Gut[3] and MA Piris[1,22]

Splenic marginal zone lymphoma (SMZL) is a B-cell neoplasm whose molecular pathogenesis remains fundamentally unexplained, requiring more precise diagnostic markers. Previous molecular studies have revealed 7q loss and mutations of nuclear factor κB (NF-κB), B-cell receptor (BCR) and Notch signalling genes. We performed whole-exome sequencing in a series of SMZL cases. Results confirmed that SMZL is an entity distinct from other low-grade B-cell lymphomas, and identified mutations in multiple genes involved in marginal zone development, and others involved in NF-κB, BCR, chromatin remodelling and the cytoskeleton.

Leukemia (2014) **28,** 1334–1340; doi:10.1038/leu.2013.365

Keywords: whole-exome sequencing; splenic; lymphoma; marginal zone

Analysis example: SMZL in patient 7



total number

somatic variants

select deleterious variants

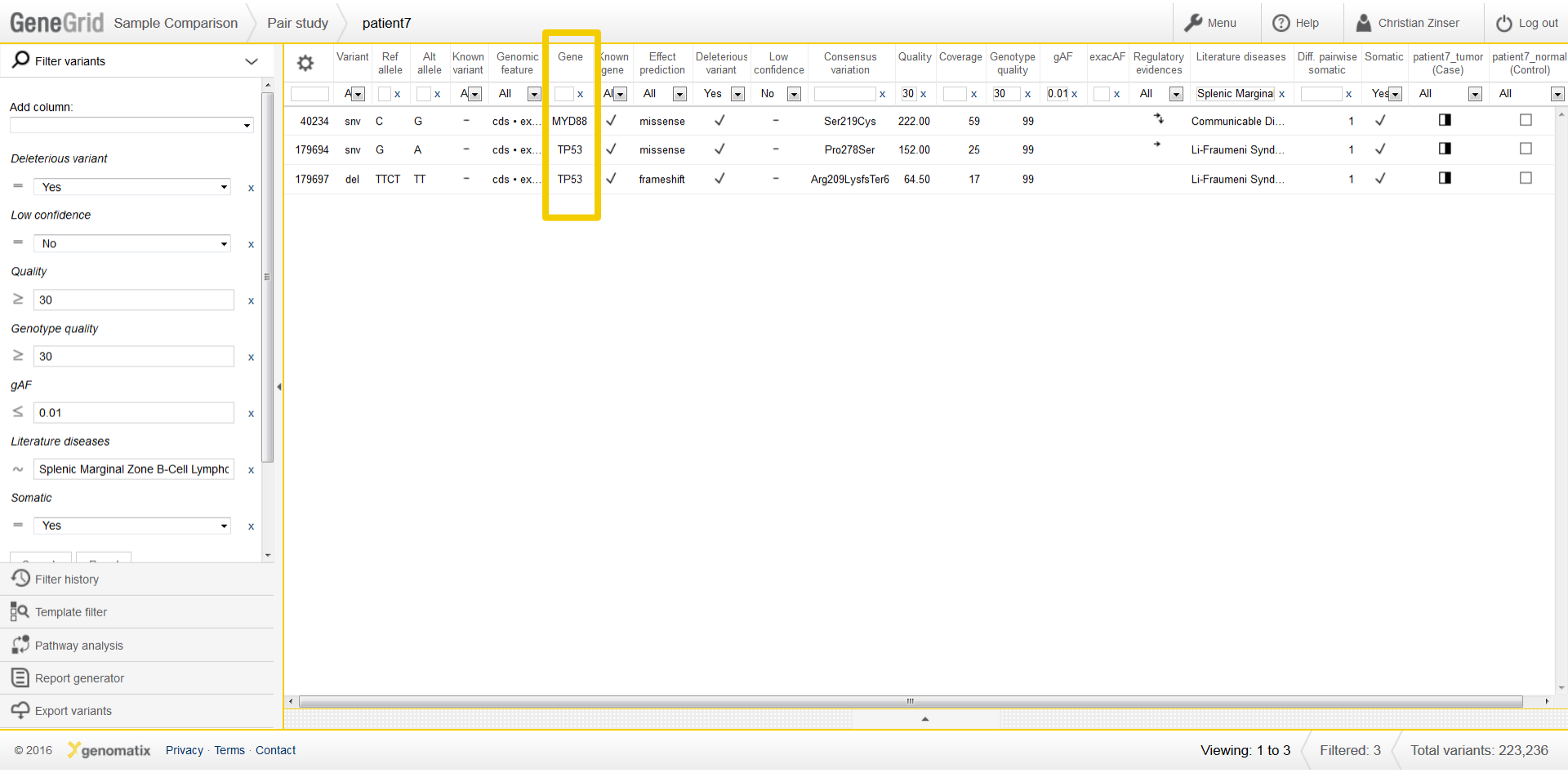
select rare variants (gAF < 1%)

quality and genotype quality > 30

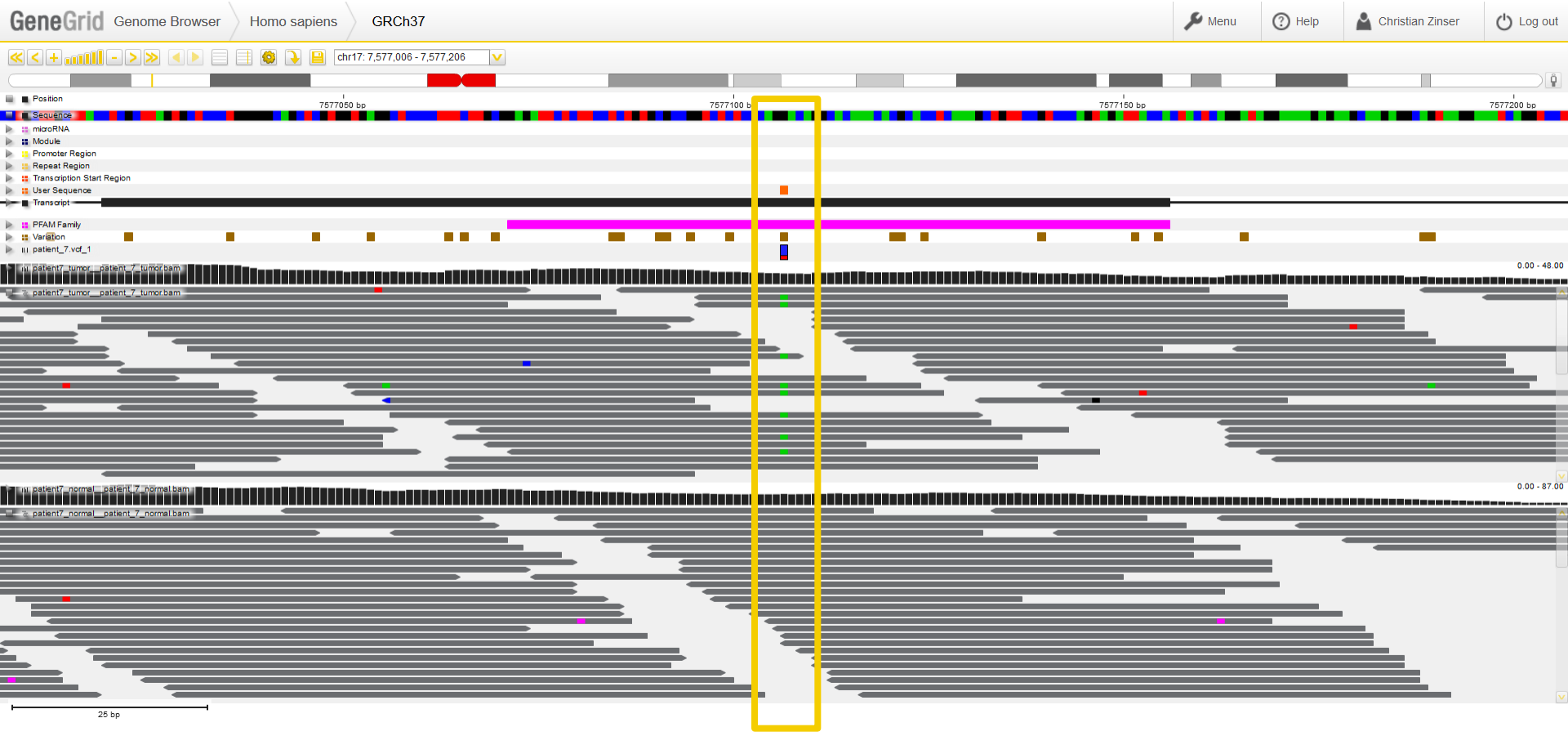
only high confidence predicted protein effects

select genes associated with
“Splenic Marginal Zone B-Cell Lymphoma” in the literature

Filtering result



Visualization



Read coverage at SNP position in TP53 for tumor and control

Visualization

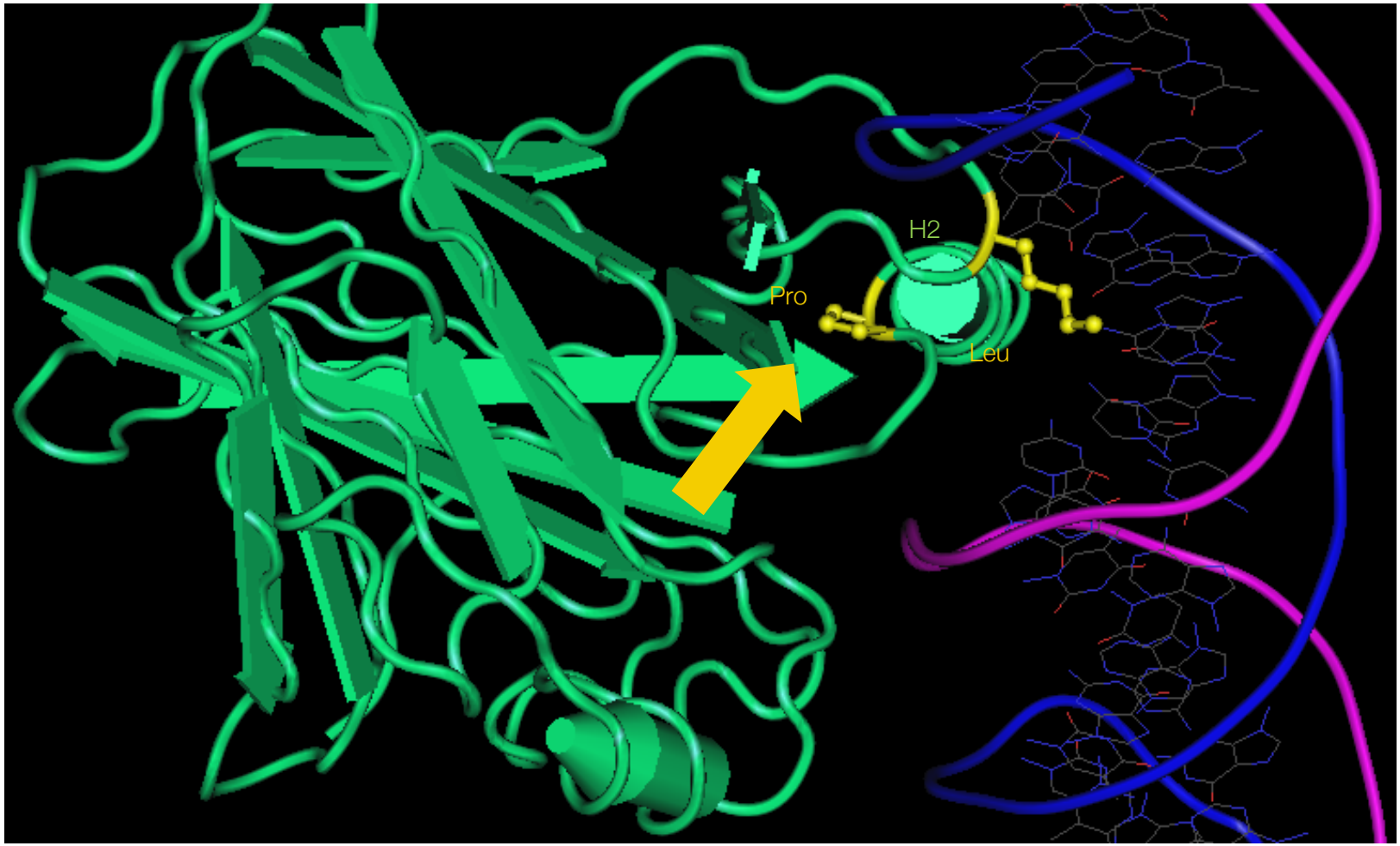
The SNP is located in the DNA-binding domain

Visualization

Sample details (2) | Transcript effects (22) | dbSNP | ClinVar | Somatic mutations (19) | COSMIC (12) | Literature diseases (100+) | Clinical diseases (39) | Gene details | Citations (94)

Ref allele	Alt allele	Ref protein sequence	Alt protein sequence	Gene	Transcript accession	Transcript version	Strand	Type	Genomic feature	Effect prediction	Deleterious	Low confidence	Variation coding	Variation protein	Coding position (bp)	Relative coding position	BLOSUM	SIFT score
G	A	P	S	TP53	ENST00000455263	2	-	protein coding	cds • exon	missense	✓	–	832C>T	Pro278Ser	832	0.80	-1	0.02
G	A	P	S	TP53	ENST00000509690	1	-	protein coding	cds • exon	missense	✓	✓	436C>T	Pro146Ser	436	0.73	-1	0.03
G	A	P	S	TP53	NM_001126112	2	-	protein coding	cds • exon	missense	✓	–	832C>T	Pro278Ser	832	0.70	-1	0.03
G	A	P	S	TP53	NM_001276761	1	-	protein coding	cds • exon	missense	✓	–	715C>T	Pro239Ser	715	0.67	-1	0.02
G	A	P	S	TP53	NM_000546	5	-	protein coding	cds • exon	missense	✓	–	832C>T	Pro278Ser	832	0.70	-1	0.03
G	A	P	S	TP53	NM_001276760	1	-	protein coding	cds • exon	missense	✓	–	715C>T	Pro239Ser	715	0.67	-1	0.02

P53 core domain in complex with DNA



Source: NCBI structure, MMDB ID: 106061

- the identified mutation could affect DNA binding or DNA affinity