

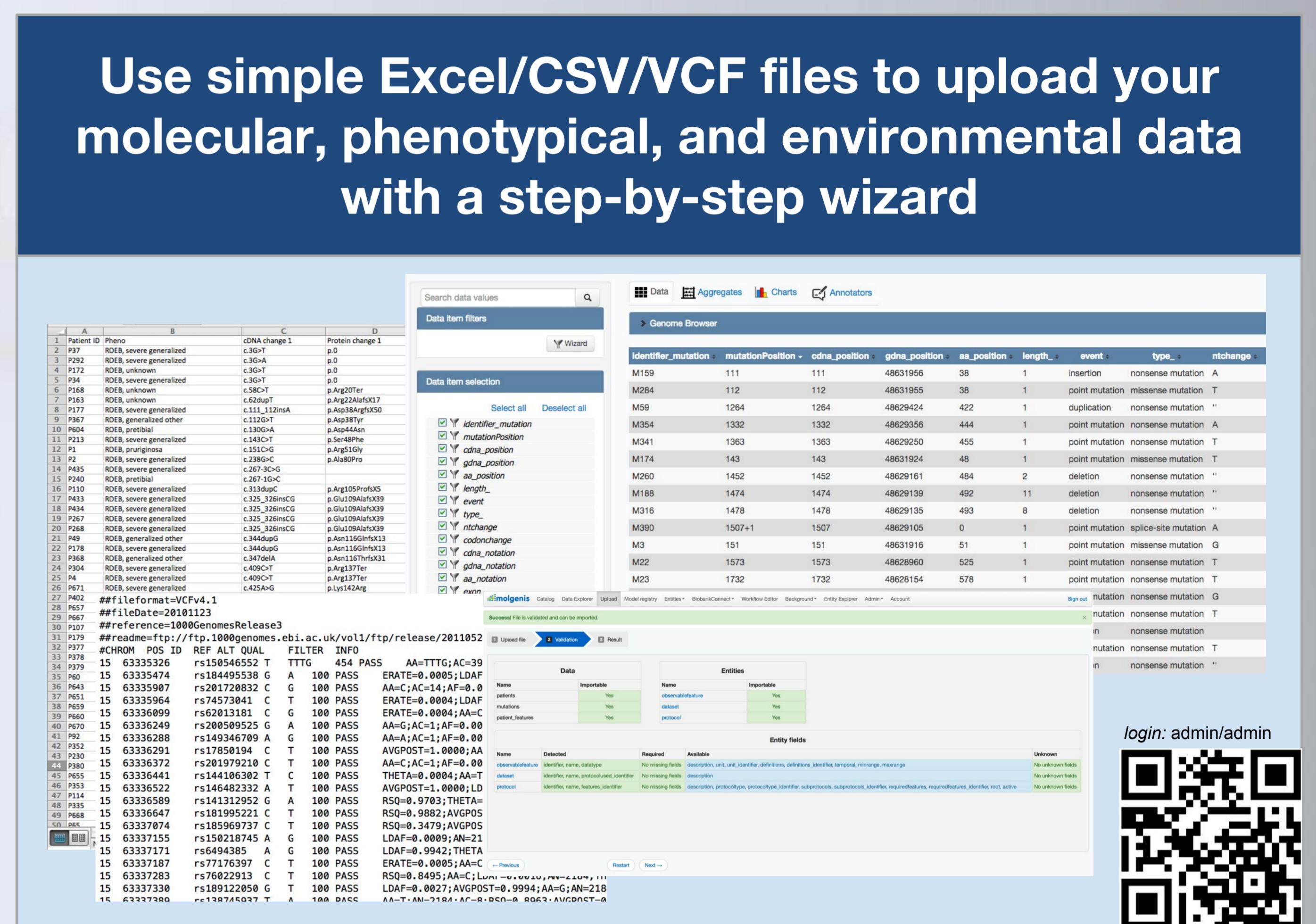
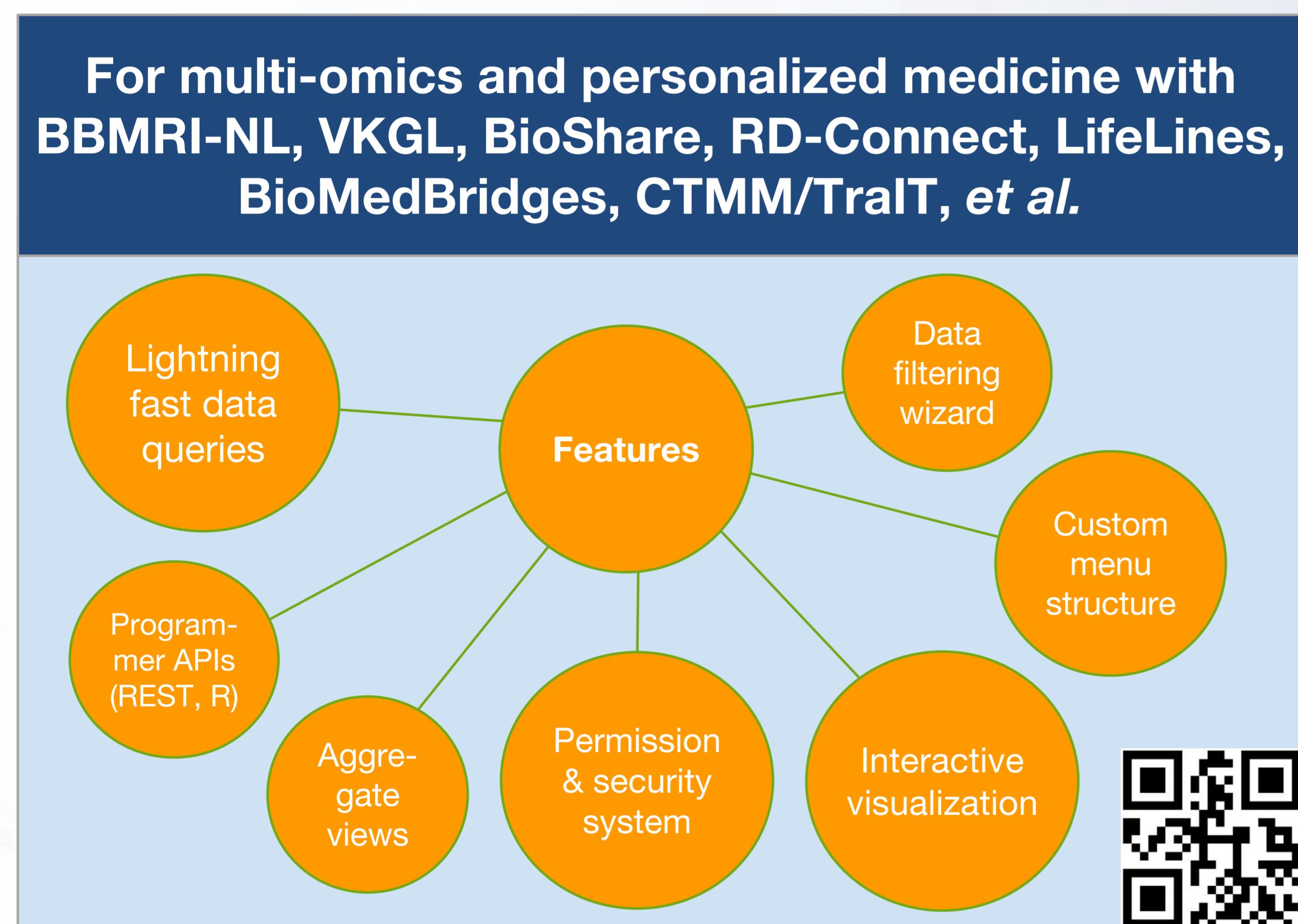
# BBMRI.nl

Biobanking and  
BioMolecular resources  
Research Infrastructure  
The Netherlands

# Towards a personalized genomics workbench using MOLGENIS

K. Joeri van der Velde<sup>1,2</sup>, Bart Charbon<sup>1</sup>, Dennis Hendriksen<sup>1</sup>, Mark de Haan<sup>1</sup>, Cisca Wijmenga<sup>2</sup>, Richard J. Sinke<sup>2</sup>, Tom J. de Koning<sup>2</sup>, Rolf H. Sijmons<sup>2</sup>, BioMedBridges, BBMRI-NL, VKGL data sharing group & MOLGENIS team<sup>1,2</sup>, Imre Vastrik<sup>3</sup>, and Morris A. Swertz<sup>1,2</sup>

1 University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, The Netherlands, 2 University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands, 3 University of Helsinki, Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland



**Enrich your data with information from public molecular, genetic, clinical and pathogenicity knowledge bases using ‘annotators’**

The screenshot shows the molgenis Annotators interface. At the top, there are tabs for Data, Aggregates, Charts, Annotators, and Disease Matcher. The Annotators tab is active.

**Annotators available:**

- KEGG i
- HGNC-Symbol i
- Clinical Genomic Database i
- OmimHpo i

**Annotations not available:**

- CADD i
- Cosmic i
- dbNSFP-Variant i
- Clinvar i
- dbNSFP-Gene i
- EBI-CHeMBL i

**Run annotation**  Copy before annotating

A central box highlights the dbNSFP annotation tool, which is part of the COSMIC suite. It includes logos for COSMIC, OMIM, ChEMBL, dbNSFP, CGD, and KEGG. The dbNSFP section states: "Variant annotation based on 23 different sources". Below this is a section for "Clinically relevant disease gene associations" and "High-level functions and utilities of the biological system, e.g. pathways". A large "And more!" button is at the bottom.

**Detailed view of dbNSFP annotations:**

The right side shows a detailed view of dbNSFP annotations for item 1402. It includes a search bar, data item filters, and a data item selection list. A black arrow points to the "INFO\_GONLMAF" entry in the selection list, which is checked. Other checked entries include INFO\_EXACMAF, INFO\_1KGMAF, and INFO\_PHENOMIZERPVAL.

**Molgenis Header:**

molgenis Upload Catalogue Data Explorer Mapping Service

1402

Search data values

Data item filters  Wizard

Data item selection

Select all Deselect all

- INFO\_TDT
- INFO\_ANN
- INFO\_LOF
- INFO\_NMD
- INFO\_GONLMAF
- INFO\_GONLGTC
- INFO\_CGDCOND
- INFO\_CGDAGE
- INFO\_CGDINH
- INFO\_CGDDGIN
- INFO\_EXACMAF
- INFO\_1KGMAF
- INFO\_PHENOMIZERPVAL
- INFO\_PHENOMIZEROMIM
- INFO\_CLINVAR\_CLNSIG
- INFO\_MONGENDISCAND
- INFO\_HPOIDS
- INFO\_HPOTERMS

Genome Browser Human GRCh38 Genome Genes Repeats Conservation 1402 1407 Powered by Biodall

#CHROM	ALT
1	A
1	T
1	G
1	A
1	C

# Plug-ins for detailed research and clinic use: WikiPathways, candidate gene reporting, patient symptom-to-disease matching, etc.

The screenshot displays the Molgenis interface for a monogenic disease candidate report. The top navigation bar includes links for Upload, Catalogue, Data Explorer, Mapping Service, Plugins, Entities, Admin, About, Account, and Sign out. Below the navigation is a toolbar with Data, Aggregates, Charts, Annotators, and 5GPM.

The main content area features a title "Monogenic disease candidate report for" followed by a section titled "Candidate genes". A legend indicates gene status: Green for strong match ( $p < 0.05$ ), Light green for weak match ( $p > 0.05$ ), and Grey for no match. A note states: "Green genes have a strong Phenomizer symptom match ( $p < 0.05$ ), light green is a weak symptom match ( $p > 0.05$ ), and grey genes do not a match. Hover over a gene to see details and the variants for this candidate below. Genes in bold font appear in multiple categories. Click on a gene to 'exclude' this candidate by flagging it with a red color."

The "Candidate genes" table is organized into three columns: Dominant, Recessive, and Compound. It lists genes categorized by impact: High impact, Moderate impact, and Other. The "Dominant" column includes FAM58A, KAL1, HUWE1, CES1, ABCA1, NEB, COL18A1, SCARF2, ABCC11, EPG5, GF1, LCR, PRDM16, ACAN, JPH3, MC1R, MYH2, VWF, GNAL, ATN1, ASCL1, GIGYF2, KIF1A, HTT, MAP3K1, FOXC1, PCSK9, ZIC3, TMIE, ADAMTS2, DMD, ZNF81, IGSF1, SERPINA1, SMPD1, SEC23B, LOXHD1, SYNE1, ZNF141, NEB, BCAM, KEL, AQP1.

Below the table are sections for "Gene details" and "Variant details". The "Gene details" table shows KIF1A as the disorder gene, with inheritance AD/AR, Generalized inh. DOM\_OR\_REC, Onset N/A, and Phenomizer 0.029. The "Variant details" table shows a disruptive\_inframe\_deletion at position 241696840.

The central part of the page contains a diagram titled "A Model of Triggering of Striated Muscle" showing the interaction between Tropomyosin, Troponin-T, Troponin-I, and Troponin-C, leading to contraction by Ca++. Another diagram shows the Myosin motor protein structure.

The right side of the page displays symptoms: HP:0000707, HP:0100022, HP:0001300, HP:0007325, HP:0002015, HP:0001336, HP:0000496, HP:0001250. A detailed clinical synopsis for these symptoms is provided, mentioning perfect matches (9 items) and similar matches (16 items). The perfect matches include Cockayne syndrome, type A; Cohen syndrome; and mitochondrial DNA depletion syndrome 4A (Alpers type). The similar matches include various disorders like combined oxidative phosphorylation deficiency 1, congenital disorders of glycosylation, and Fanconi anemia.

At the bottom left, a table of variants is shown:

CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	ref
1	78383467	rs1749913	G	A	.	PASS	AC=21;AN=22;GTC=0,1,10	0
1	78392121	.	T	C	.	PASS	AC=1;AN=2;GTC=0,1,0	0
1	78392391	.	G	A	.	PASS	AC=1;AN=2;GTC=0,1,0	0
1	78392445	.	C	A	.	PASS	AC=1;AN=2;GTC=0,1,0	0
1	78392446	rs1166698	G	A	.	PASS	AC=48;AN=84;GTC=0,36,6	0
1	78395131	.	A	C	.	PASS	AC=1;AN=2;GTC=0,1,0	0

# Add your statistics: create and use ‘on the spot’ R analyses with built-in R framework and R-API (example: A.S.E. @ [www.molgenis.org/ase](http://www.molgenis.org/ase))

Add Script

Type \* r

Generate security token

Content \*

```
9 if (!require("RColorBrewer")) {  
10   install.packages("RColorBrewer", dependencies = TRUE)  
11   library(RColorBrewer)  
12 }  
13  
14 #####  
15 ### B) Reading in data and transform it into matrix format  
16 #####  
17  
18 data <- read.csv("../datasets/heatmaps_in_r.csv", comment.char="#")  
19 rnames <- data[,1] # assign labels in column 1 to "rn"  
20 mat_data <- data.matrix(data[,2:ncol(data)]) # transform column 2-5 into a matrix  
21 rownames(mat_data) <- rnames # assign row names  
22  
23
```

Name \* ASE plot

Parameters Select ScriptParam

Result file extension resultFileExtension

Cancel Save

1.12E-4	Chromosome	1	SNP_ID
1.49E1	Position	11087524	Reference_allele
0.465	Samples	28	Alternative_allele

Gene\_symbol MASP2 Biotype protein\_coding

rs1782455 - ENSG00000009724 p-value:  $1.12 \times 10^{-4}$

Percentage alternative (A)

Reference allele count (G)

Total number of variants

QR code

# Query all shared variant data from e.g. VKGL\* diagnostic labs in one genome browser

\* Richard Sinke (UMCG), Rien Blok (MUMC), Dennis Dooijes (UMCU), Martin Elfrink (UMCU), Marielle van Gijn (UMCU), Jan Jongbloed (UMCG), Ronald Lekanne (AMC), Renee Niessen (UMCG), Rolf Pfundt (RUMC), Claudia Ruivenkamp (LUMC), Jasper Saris (EMC), Rolf Sijmons (UMCG, VKGN), Maartje Vogel (NKI-AVL), Morris Swertz (UMCG), Bart Charbon (UMCG), Christian Gilissen (RUMC), Bart de Koning (MUMC), Jeroen Laros (LUMC), Thiery Loones (Cartagenia), Pieter Neerincx (UMCG), Bram Miedema (UMCG), Ies Nijman (UMCU), Jonathan Taminau (Cartagenia), Peter Taschner (LUMC), Quinten Waisfisz (VUMC), Steven van Vooren (Cartagenia)

Search data values ✖️ 🔍

Data Aggregates Charts

Genome Browser

Human GRCh37/hg19 2:179,262,834..179,896,105 🔍 ✖️ ➕ ✖️ ✖️

Genome 179,300,000 179,350,000 179,400,000 179,450,000 179,500,000 179,550,000 179,600,000 179,650,000 179,700,000 179,750,000 179,800,000

Genes >RP11-65L3.1 >DFNB59 >PLEXHA3 <TTN <RNUT7-104P <CCDC141  
>AC009948.5 >AC009948.7 >TTN-AS1 >RP11-65L24.4  
>PRKRA <FKBP7 >RP11-65L3.4 >RP11-17112.5 >RP11-17112.4 >RP11-17112.1  
>RP11-65L3.2 >RP11-17112.3 >RP11-17112.2

Merged UMCG UMCU GoNL\_Cardio LOVD\_TTN UMCG\_Cardio\_ManVar\_Benign

GTC(UMCG) ▾ GTC(UMCU) ▾

POS	REF	ALT
236910942	C	G
236910983	G	A
236911122	C	T
236911137	A	G
236911181	C	T
236911221	G	T

0,10,0 138,8,0

0,10,1

0,10,2 203,12,0

0,10,2 203,12,0

0,10,2 203,12,0

0,10,2 179,11,0

QR code