

Understanding *M. tuberculosis* variants using Galaxy and the

COMBAT TB Explorer

Thoba Lose

Software Developer



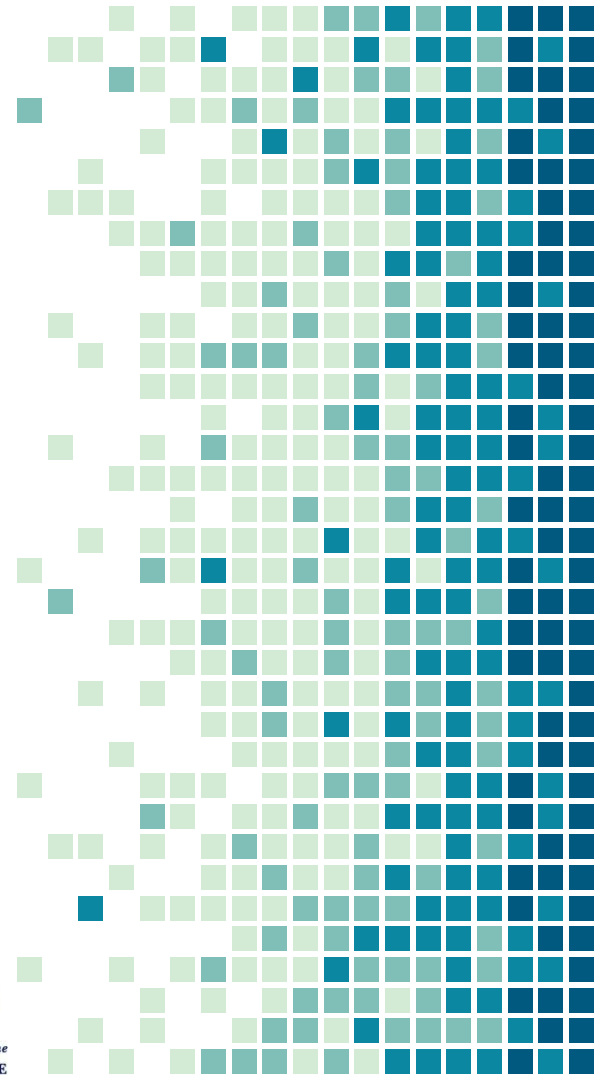
SANBI
South African National
Bioinformatics Institute

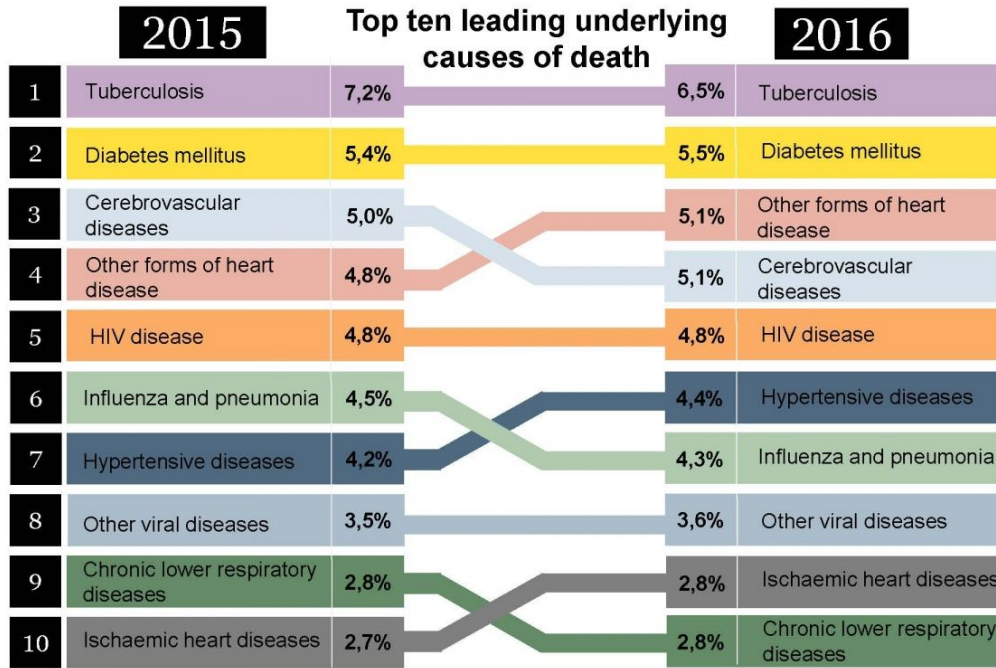


**UNIVERSITY of the
WESTERN CAPE**



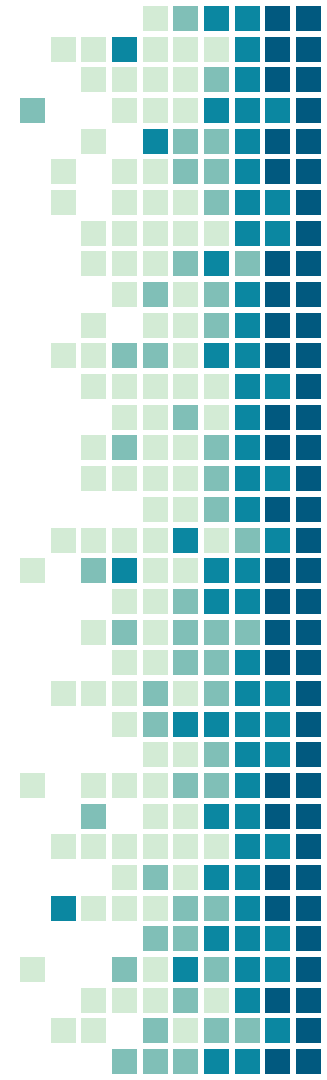
[#GalaxyAfrica](#)





THE SOUTH AFRICA I KNOW, THE HOME I UNDERSTAND

Top 10 leading in causes of death in South Africa



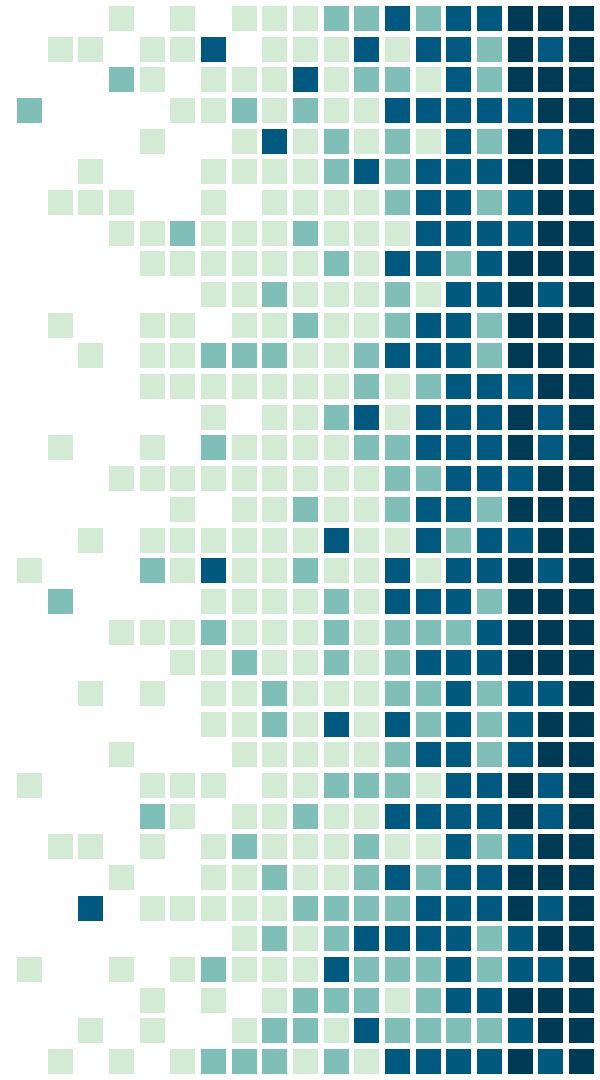
“ While workflows and associated software to process raw data and produce high-confidence variant calls have significantly improved, filtering of candidates to identify a subset relevant to a specific study is still a complex exercise...

A practical guide to filtering and prioritizing genetic variants

Mahjoubeh Jalali Sefid Dashti and Junaid Gamielidien
South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa

BioTechniques 62:18-30 (January 2017) doi 10.2144/000114492

Keywords: whole-genome sequencing; whole-exome sequencing; functional variants; variant prioritization



Questions a biologist might ask



What part of the gene is this variant located in?

What is the predicted effect of the variant?

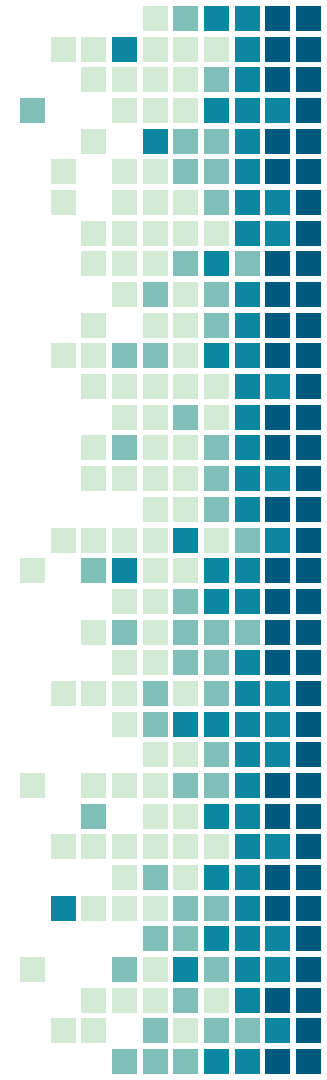
Is the variant in a gene associated with an essential pathway?

Is the variant in a gene that encodes a protein that interacts with a known disease protein?

Is the variant in a gene that has a function that coincides with the pathology of TB?

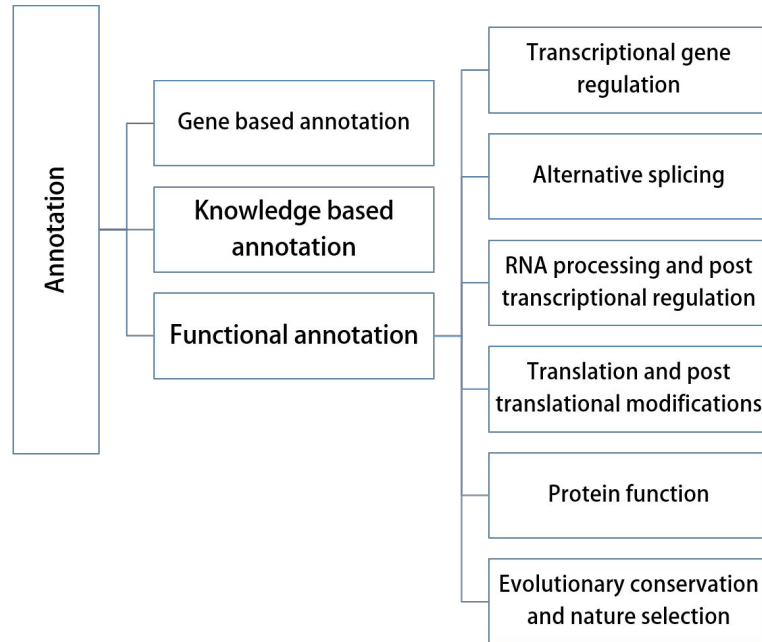
Is the variant in a gene that encodes a target protein (drug target)?

There is a possibility that genes associated with IPR000873 activate fatty acids during dormancy



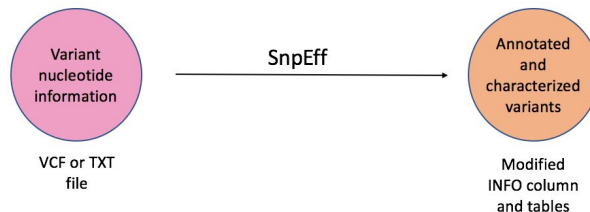
Variant Annotation

- Involves adding metadata and knowledge in effort to enhance assessment of variants likely to impact function.
- SnpEff is a variant annotation and effect prediction tool.



SnEff

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
12	1290	rs6875309	T	C	56	PASS	NS=4;DP=20;AF=0.045
9	680	rs44309	G	A	2	q15	NS=1;DP=3;AF=0.78

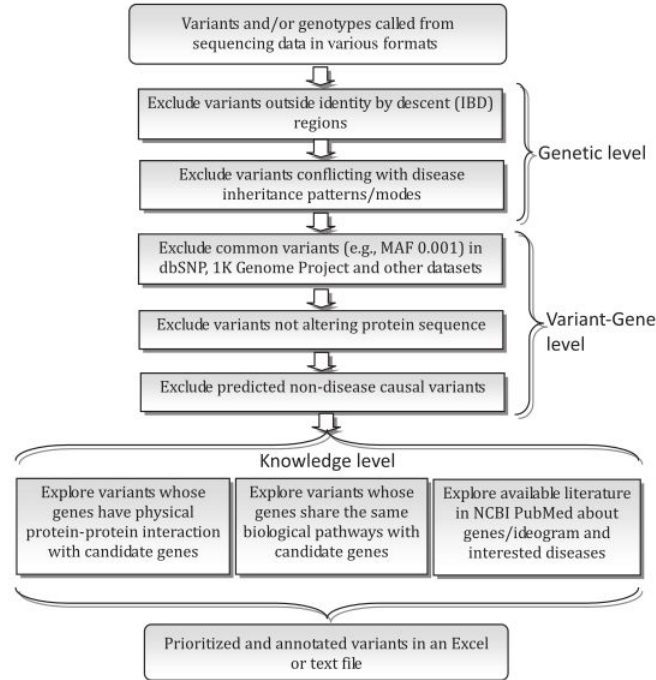


#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
12	1290	rs6875309	T	C	56	PASS	NS=4;DP=20;AF=0.045 ;EFF=STOP_LOSS(HIGHINONSENSEITag/CogIF2*1 3449 TRP2LIICODINGINM_97375I)
9	680	rs44309	G	A	2	q15	NS=1;DP=3;AF=0.78 ;EFF=STOP_GAIN(HIGHIFRAMESHIFTITag/CogIF2*1 689 FRC2LIICODINGINM_454648I)

Variant Prioritization



- It is possible to rank variants based on predicted impact. However, identifying the strongest candidates is not obvious
- Assessing candidate genes in the context of extant knowledge is vital.

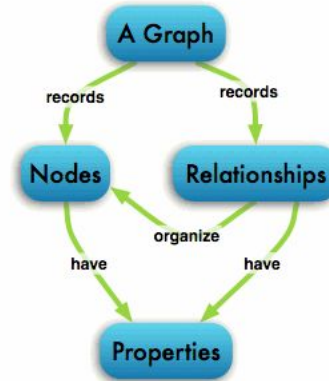
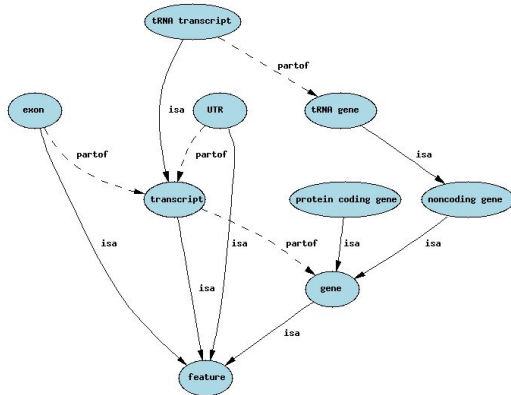


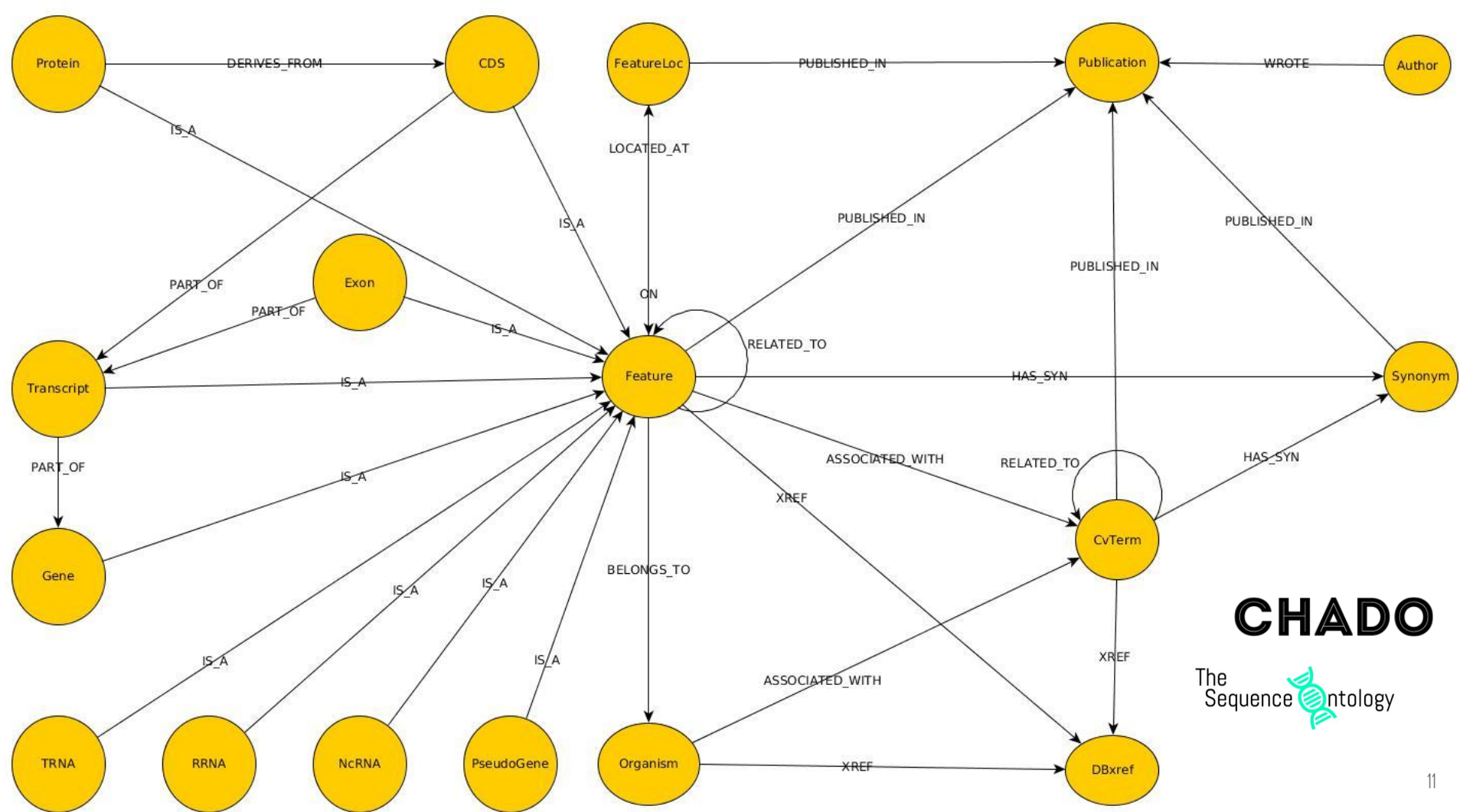
<https://dx.doi.org/10.1093/2Fnar%2Fgkr1257>

Neo4j



- A highly scalable graph database with a declarative query language called Cypher
 - Property Graph Model
 - Whiteboard friendly
 - Schema-less
 - Exposes a RESTful API





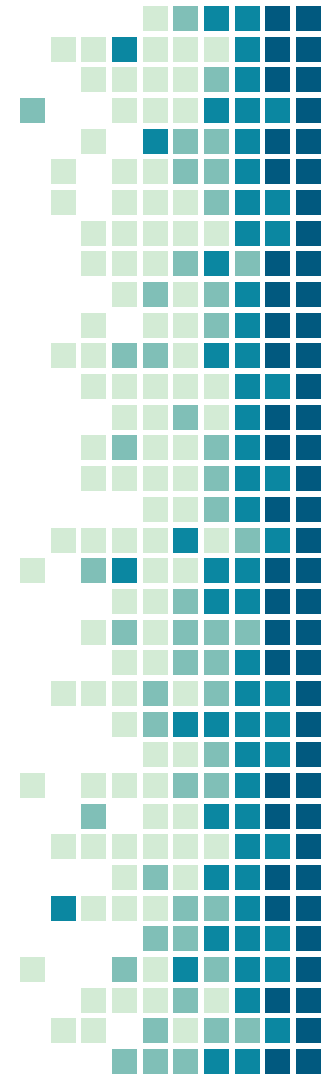
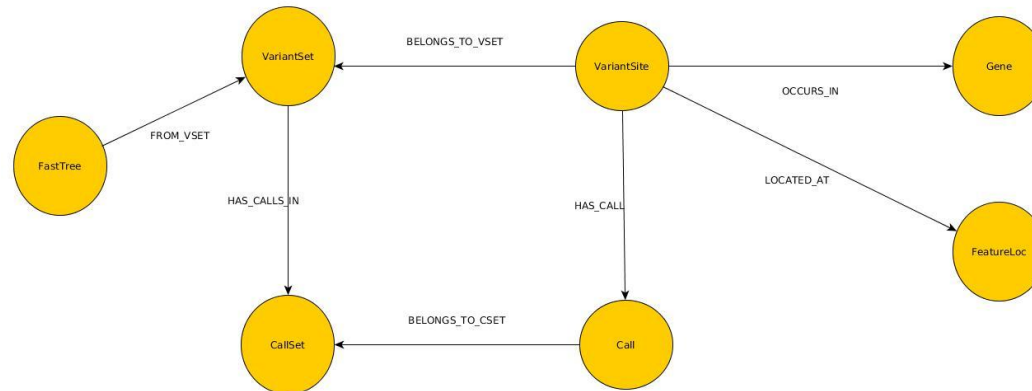
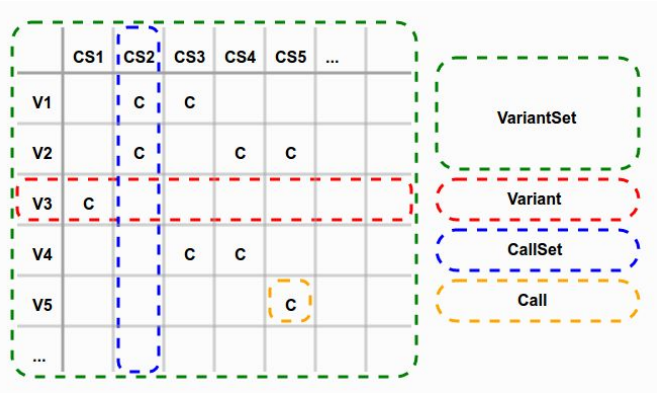
CHADO

The Sequence  ntology

GA4GH Variant Model

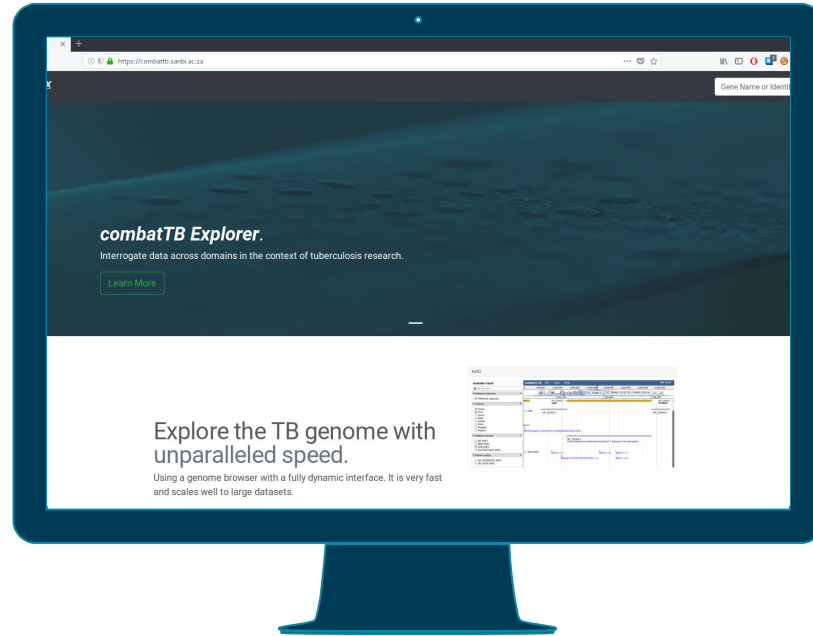


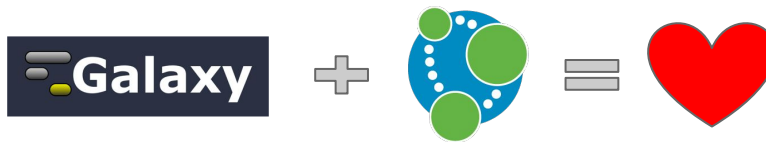
Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.



COMBAT TB Explorer

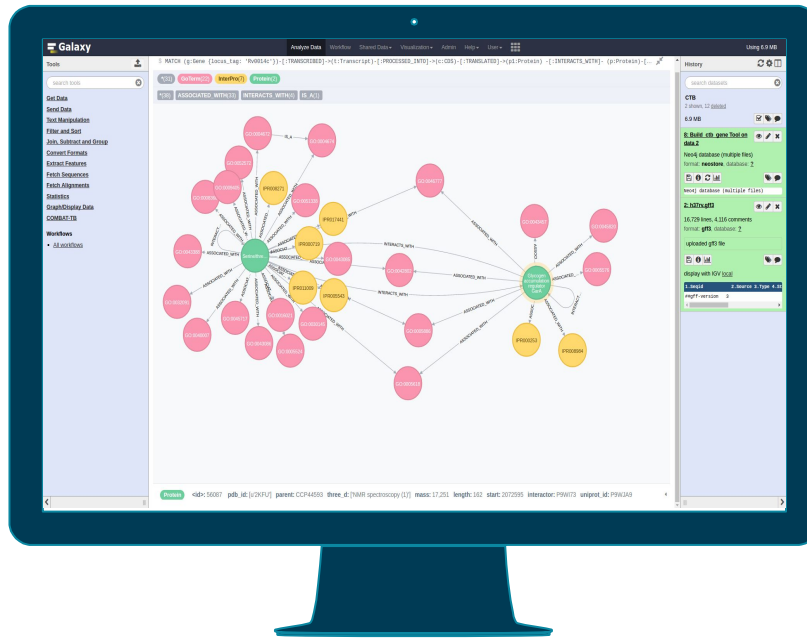
An integrated knowledge management system that allows researchers to interrogate data in the context to TB research and data available in public repositories





Neo4j Galaxy IE

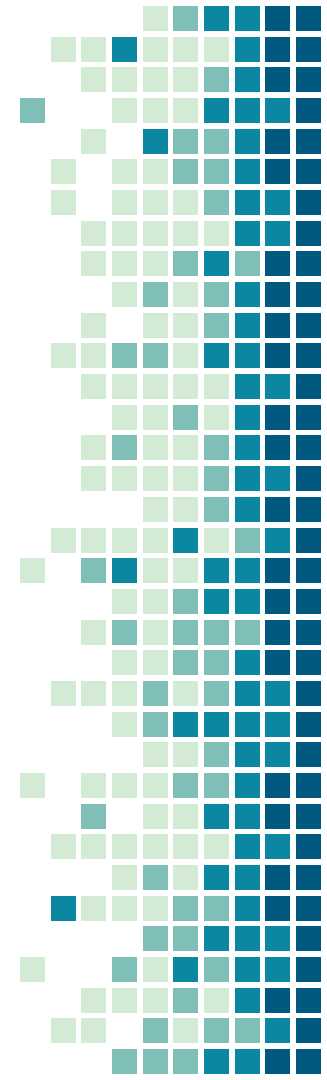
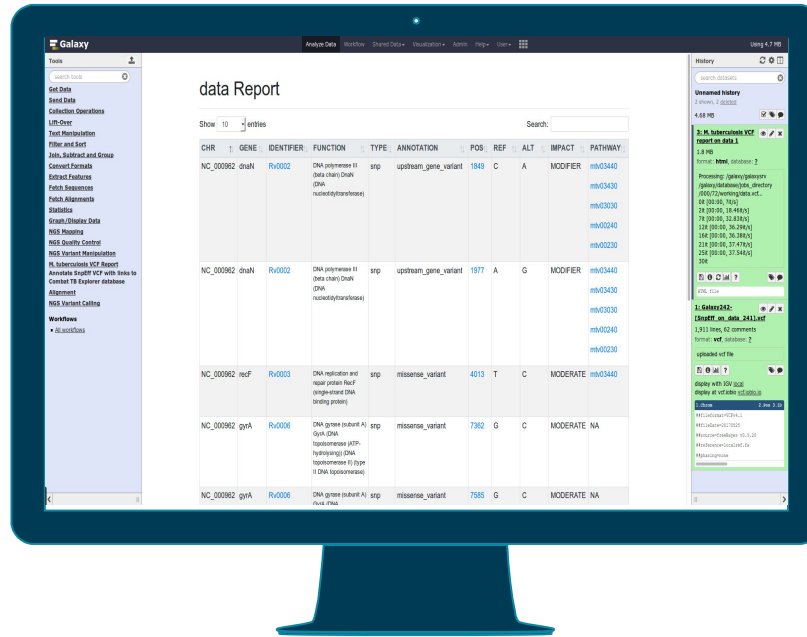
Based on a customized Neo4j Docker image and utilises the **neostore** datatype.





tbvcfreport

A tool that utilises the Neo4j database to generate an interactive HTML-based VCF report, from SnpEff produced VCF files, with links to the COMBAT TB Explorer



Conclusion

- The integration of extant biological resources in the context of TB, the Combat TB Explorer combined with Galaxy provide a single, easy to use, integrated approach to prioritize candidate variants.



The Team

Christoffels Lab:

- L. Abdoll
- T. Calvert-Joshua
- A. Christoffels
- K. Direko
- Z. Mashologu
- J. Southgate
- P. van Heusden



Funding

This project is funded by:

- South African Medical Research Council
- South African Research Chairs Initiative of the Department of Science and Technology and
- National Research Foundation



THANKS!

Any questions?

You can find me at:
thoba@sanbi.ac.za



 Galaxy
Africa



UNIVERSITY of the
WESTERN CAPE



science
& technology

Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA



National
Research
Foundation