Thailand National Infrastructure in Genome Informatics/Gene Bank

Sissades Tongsima, Ph.D.

National Biobank of Thailand (NBT) and
National Center for Genetic Engineering and Biotechnology

26 March 2019

เศรษฐกิจฐานชีวภาพ

Bio-based Economy

นวัตกรรมที่สร้างคุณค่าจากข้อมูลชีวภาพ 4 ด้าน

ด้านอุตสาหกรรมเทคโนโลยีชีวภาพ

ข้อมูลสารสำคัญและด้านศักยภาพของพืช จุลินทรีย์ เช่น เอนไซม์ สารชีวภัณฑ์ วัคซีน ยา

ด้านการเกษตรและอาหาร

ข้อมูลจีโนไทป์-ฟีโนไทป์เพื่อการปรับปรุงพันธุ์พืชและสัตว์ อาหารฟังก์ชัน

ด้านชุมชนและสิ่งแวดล้อม

ข้อมูลนิเวศวิทยา เพื่อการอนุรักษ์และฟื้นฟูป่า

ด้านการแพทย์และสุขภาพ

ข้อมูลพันธุกรรมอ้างอิงของประชากรไทย เพื่อสนับสนุนการแพทย์แม่นยำ

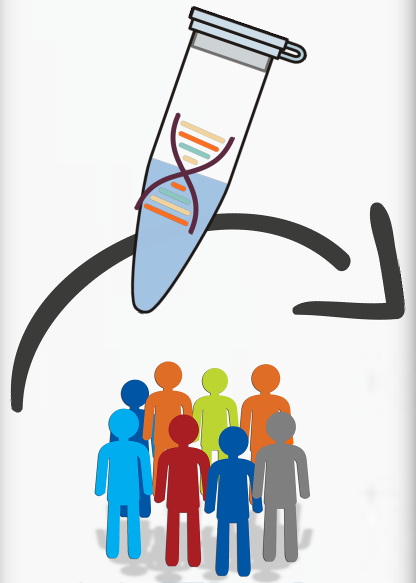
Dr. Chumpol

2003

ABI model 3730 DNA Sequencer

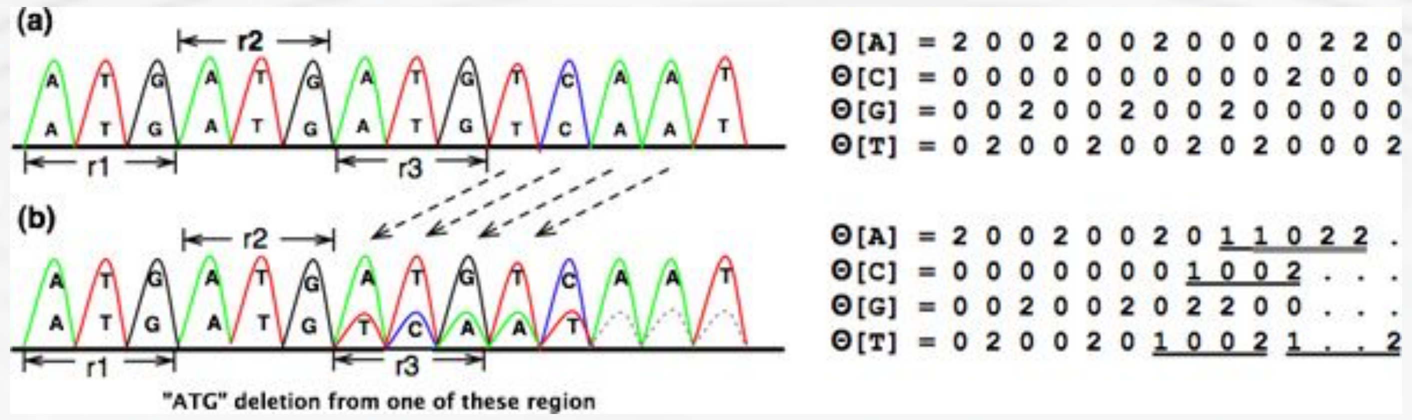
ThaiSNP

Thailand SNP Discovery Project



32 Thais

VarDetect



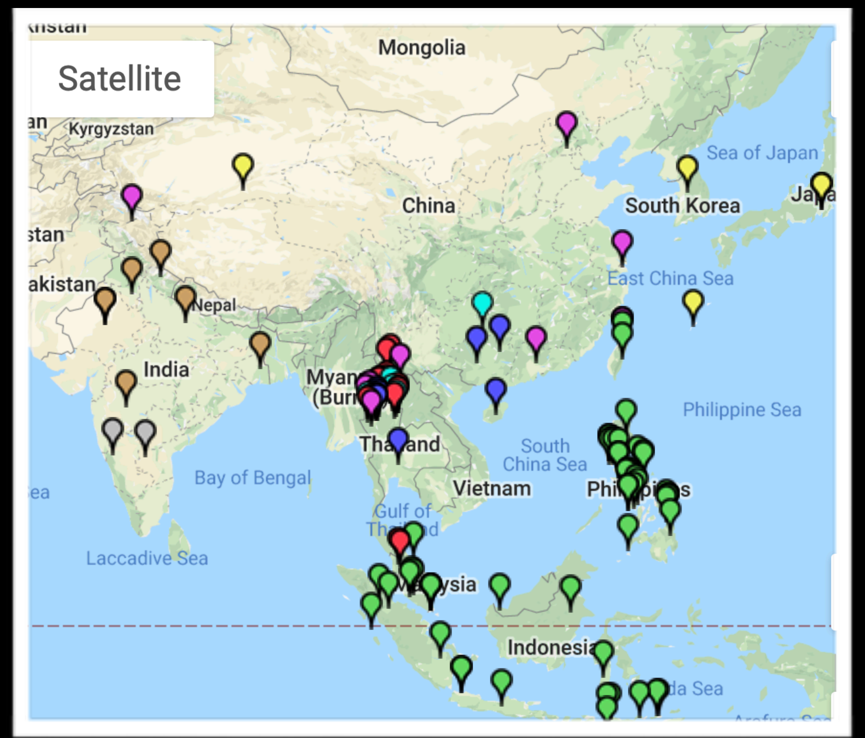
Ngamphiw *et al.* 2008 *BMC Bioinformatics*

2008

510K SNPArray

ThaiSNP2

2009-11

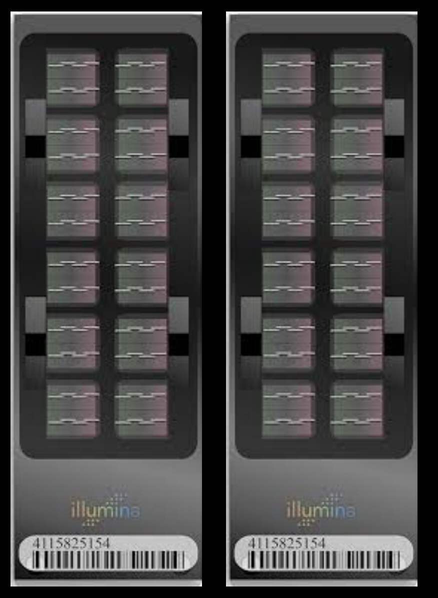


HUGO Pan-Asian SNP Consortium, *Science* (2009)

Ngamphiw *et al.* 2011 *PLOS One*

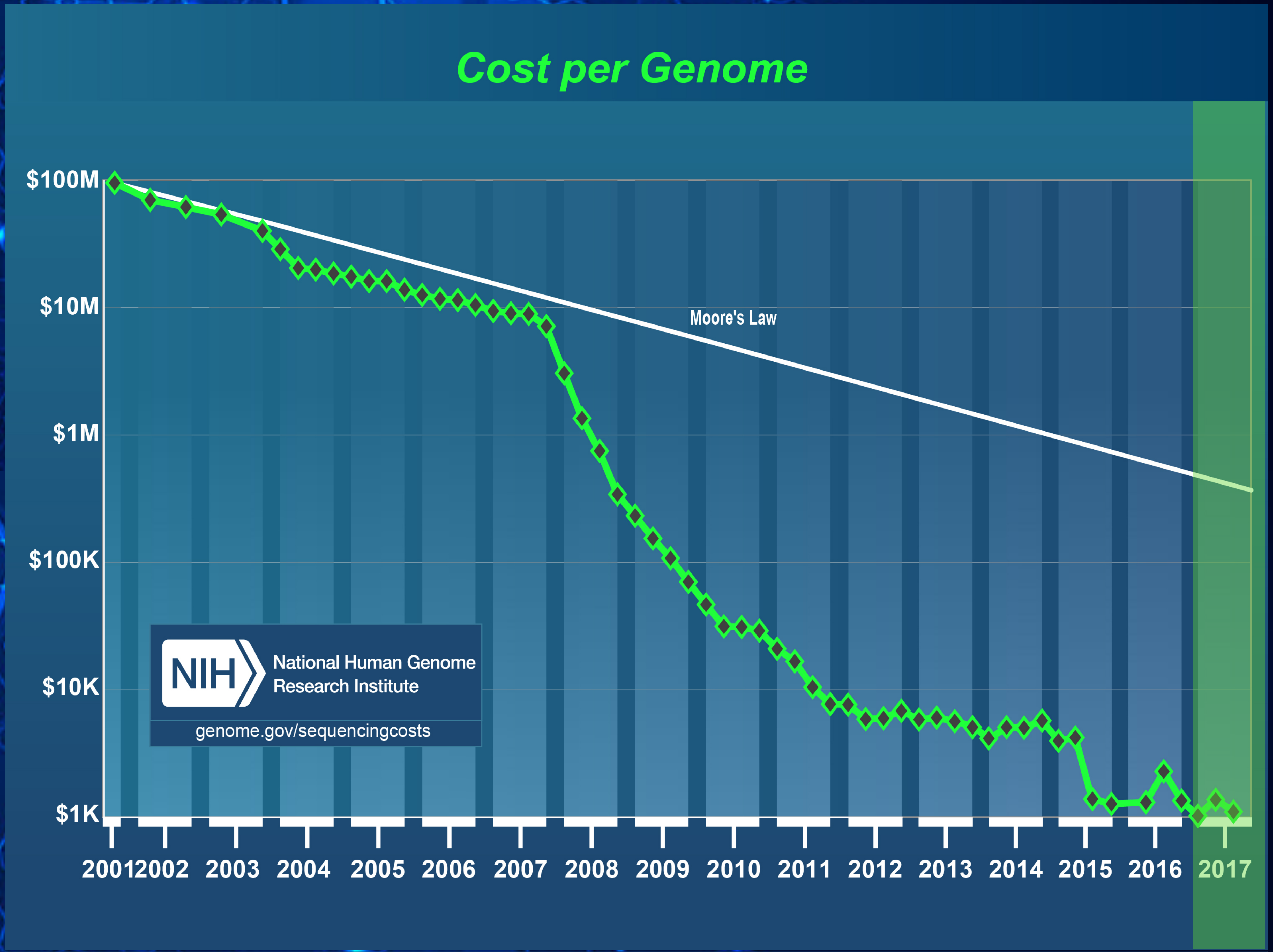
PASNP

2009



illumina

Infinium610-Quad



Genomics England

2014

All of US

2015

GENOMICS THAILAND

2017

GENOMICS THAILAND

เปลี่ยน แนวคิดทางการแพทย์

- การวินิจฉัยโรค
- การป้องกัน
- การรักษา

Mission

การวิจัย

ฐานข้อมูลอ้างอิงพันธุกรรมประชากรไทย

ให้บริการ

บริการตรวจหาความผิดปกติทางพันธุกรรม

วิเคราะห์และจัดการข้อมูล

ELSI, พัฒนาบุคลากร, พัฒนาอุตสาหกรรมใหม่

NBT

Duties: สร้างฐานข้อมูลกลางของประเทศ

- ฐานข้อมูลอ้างอิงความหลากหลายทางพันธุกรรมของประชากรไทย
- แพลตฟอร์มประมวลผลข้อมูลพันธุกรรม

④ Knowledge

- ✔ โรคมะเร็ง
- ✔ โรคหายากหรือไม่สามารถวินิจฉัย
- ✔ โรคไม่ติดต่อเรื้อรัง
- ✔ โรคติดเชื้อ
- ✔ เภสัชพันธุศาสตร์

③ Information

Query

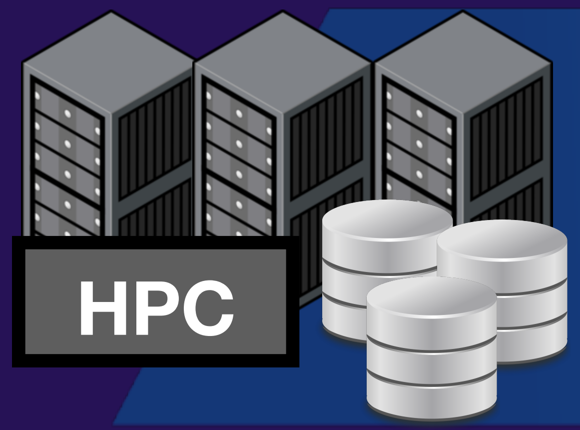
ฐานข้อมูลอ้างอิงพันธุกรรมของประชากรไทยและ genomic landscape ที่สามารถสืบค้นและแสดงผลผ่าน web portal

② Data Processing

Workflow

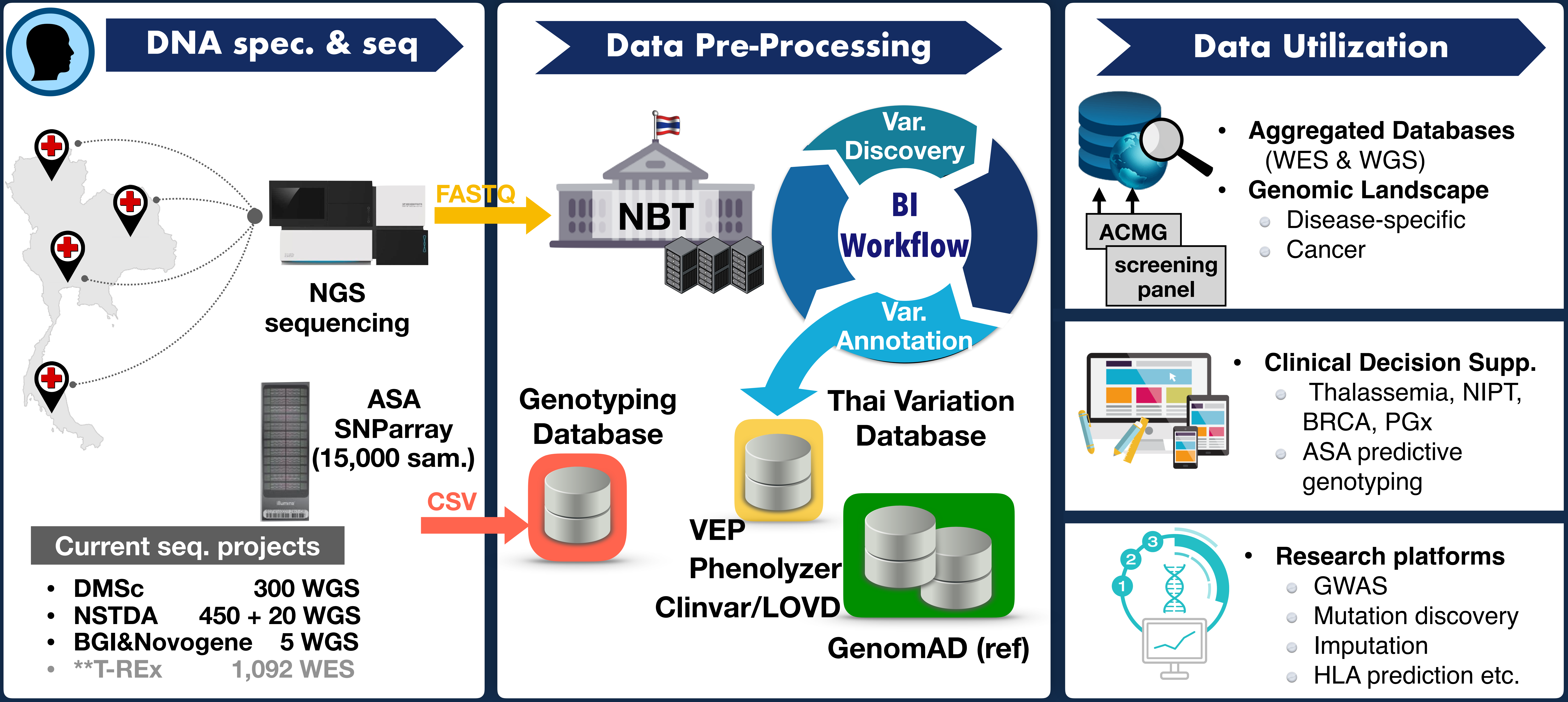
กระบวนการสกัดข้อมูลจีโนมเพื่อค้นหาตำแหน่งความหลากหลายทางพันธุกรรมของประชากรไทย

① Data Storage

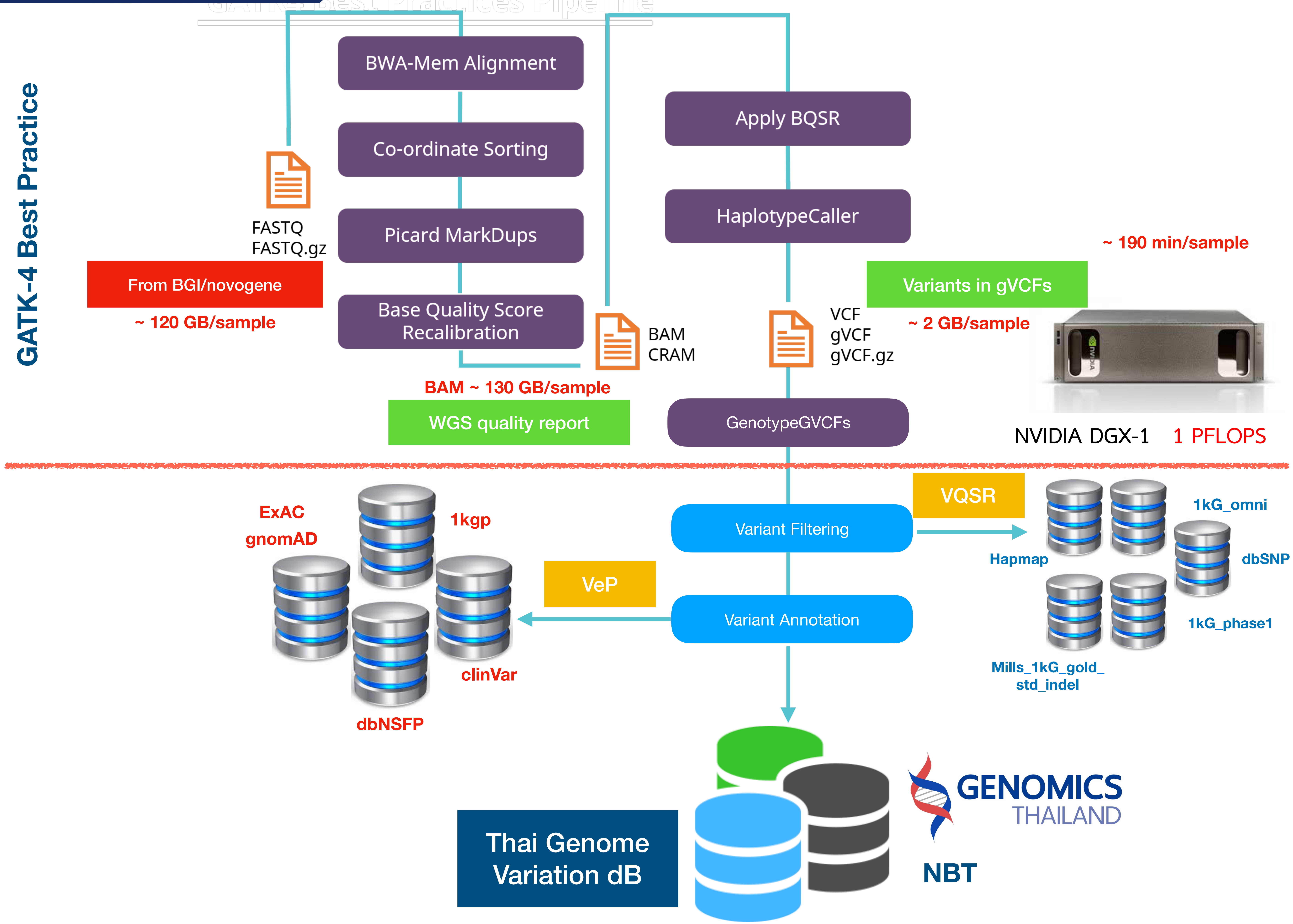


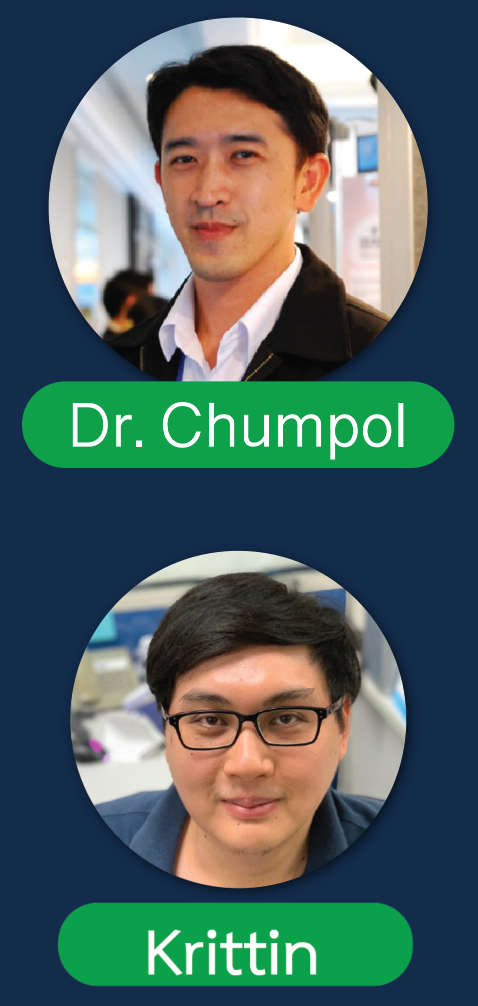
ระบบจัดเก็บและบริหารจัดการข้อมูลจีโนมมนุษย์ขนาดใหญ่ 5 หมื่นคน (~ 15 petabytes) ในอีก 5 ปีข้างหน้า

Data Flow and Management



Data Pre-Processing



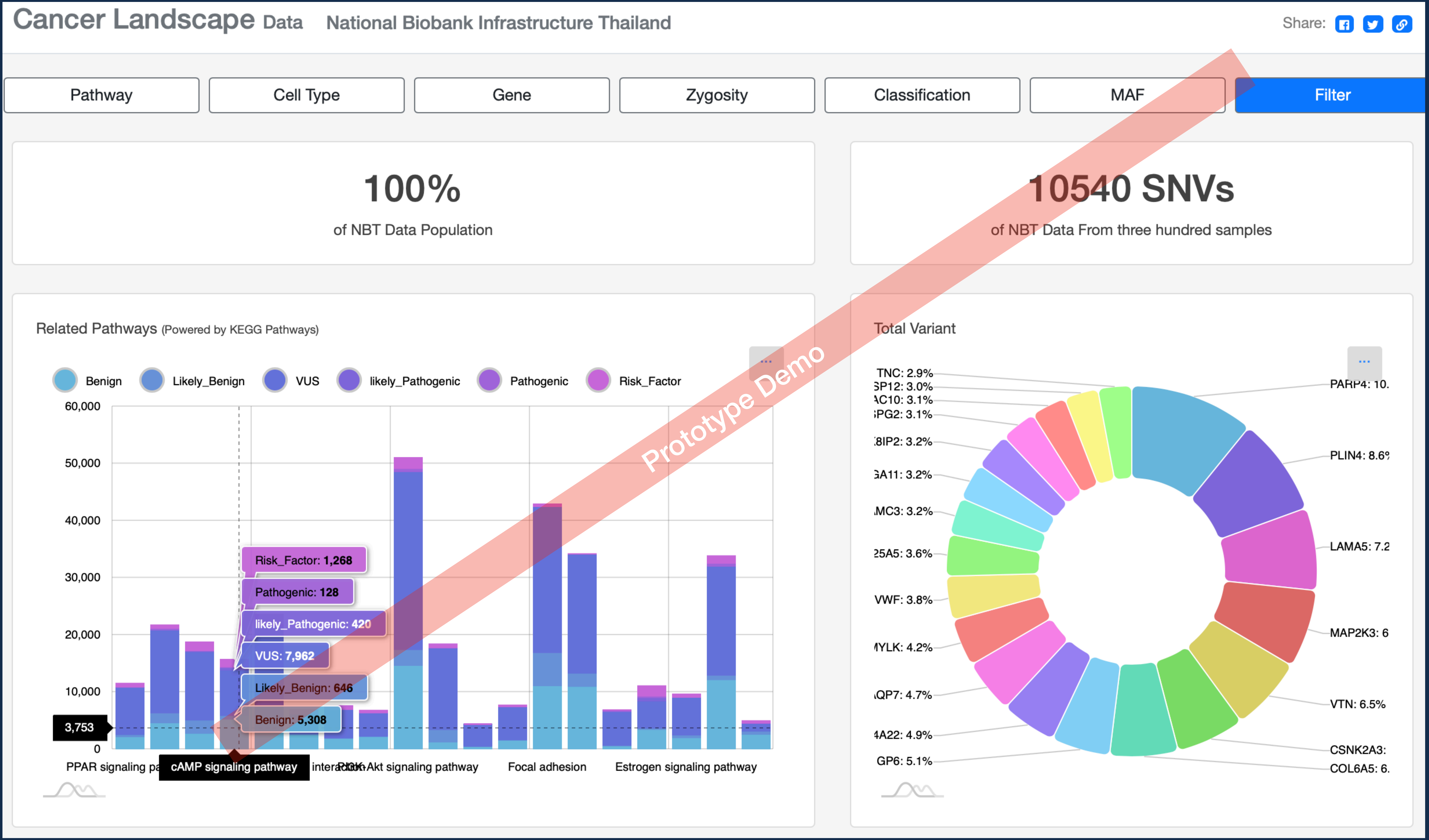


Demonstration

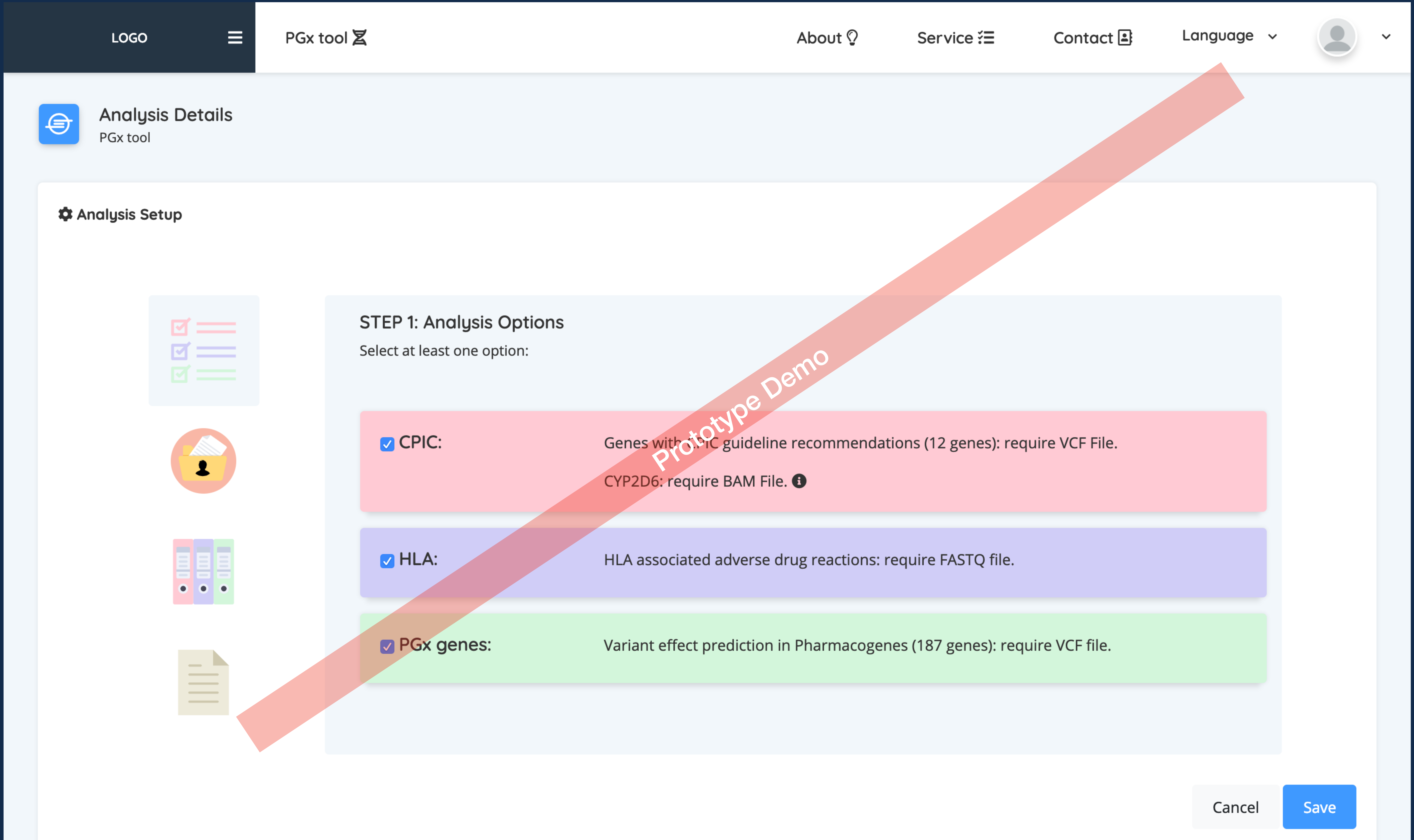
CDSS and research platforms

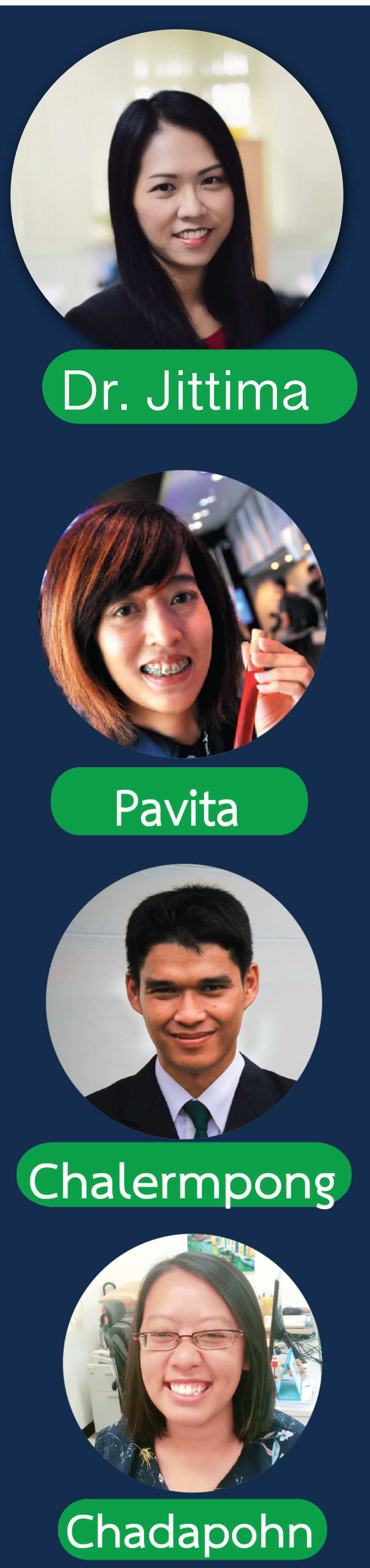
- Cancer genomic landscape
- Phamacogenomics
- Genome-wide association study (GWAS)
- Thalassemia
- Non-Invasive Prenatal Testing (NIPT)
- Structural variants discovery

Cancer Genomic Landscape

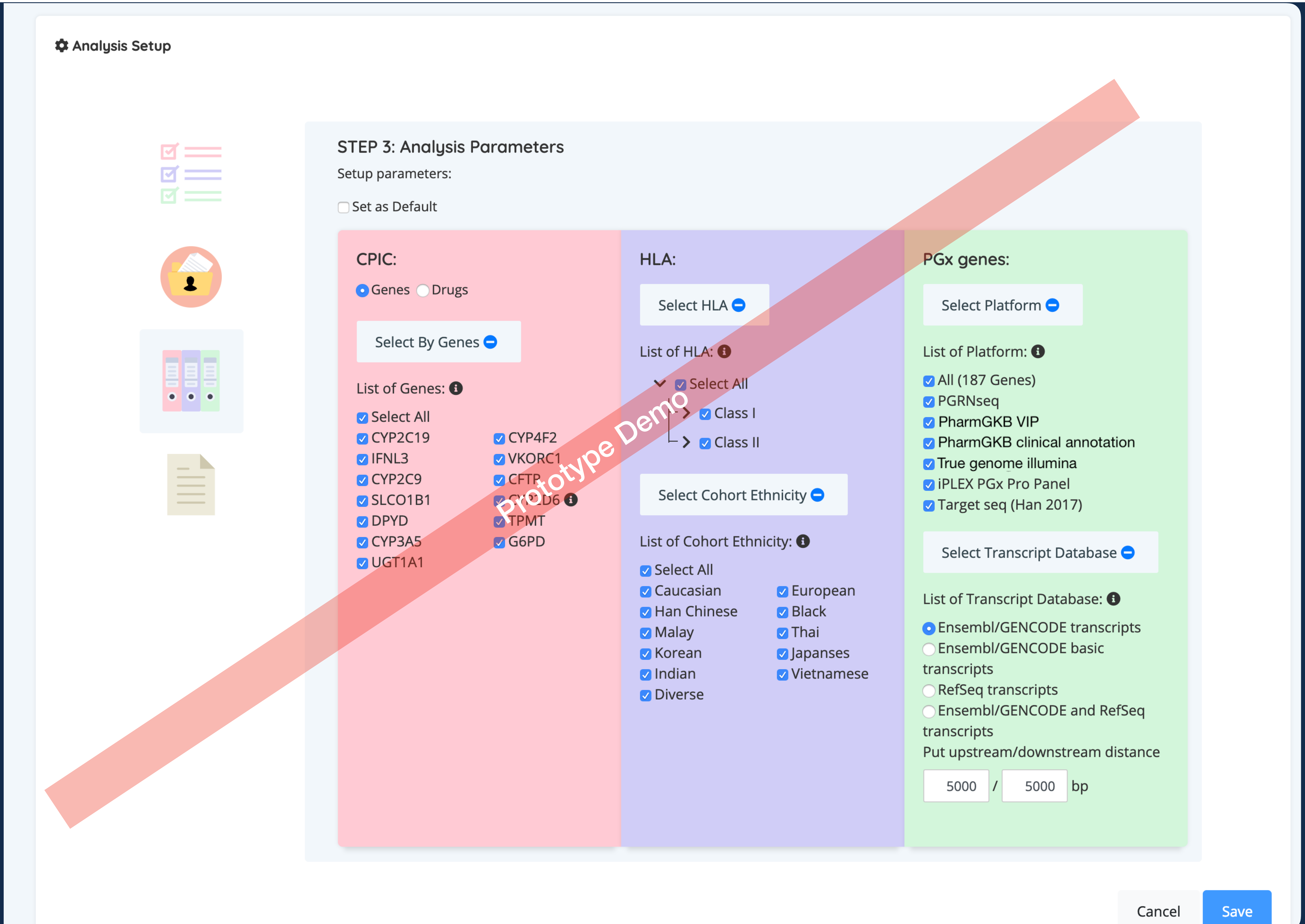


Pharmacogenomics

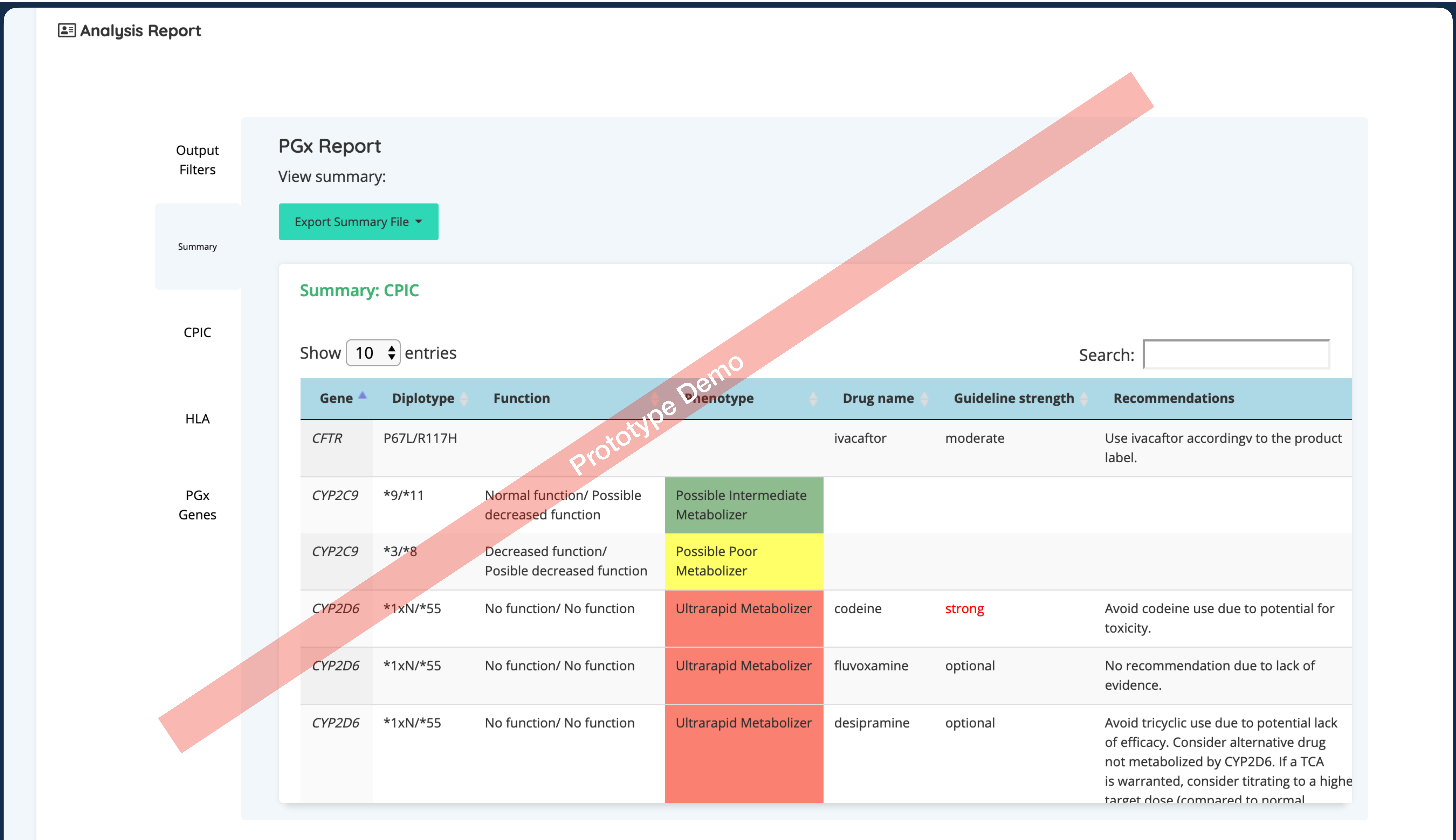




Pharmacogenomics *(cont'd)*



Pharmacogenomics *(cont'd)*



Analysis Report

Output Filters

Summary

CPIC

HLA

PGx Genes

PGx Report

View summary:

Export Summary File

Summary: CPIC

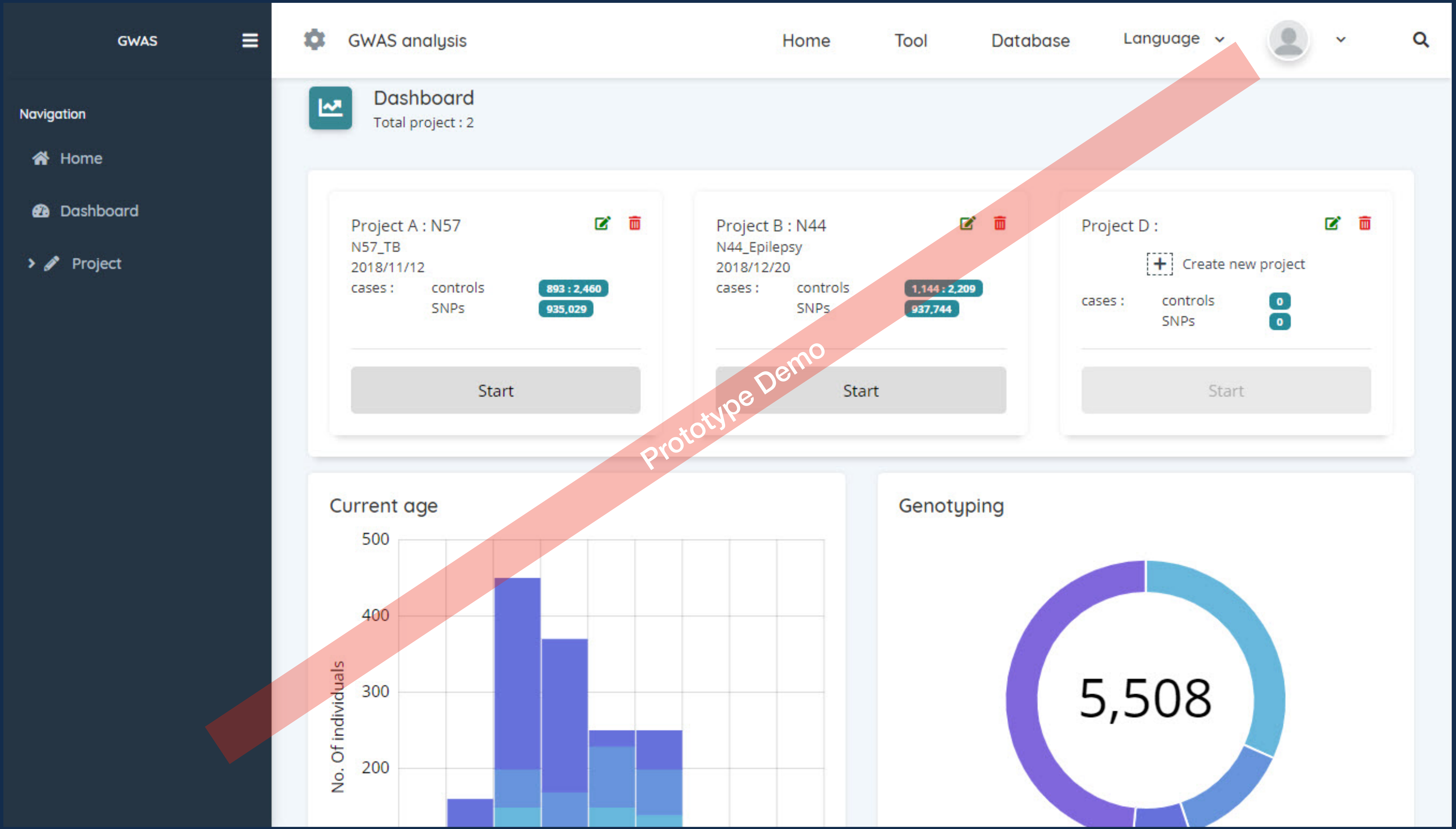
Show 10 entries

Search:

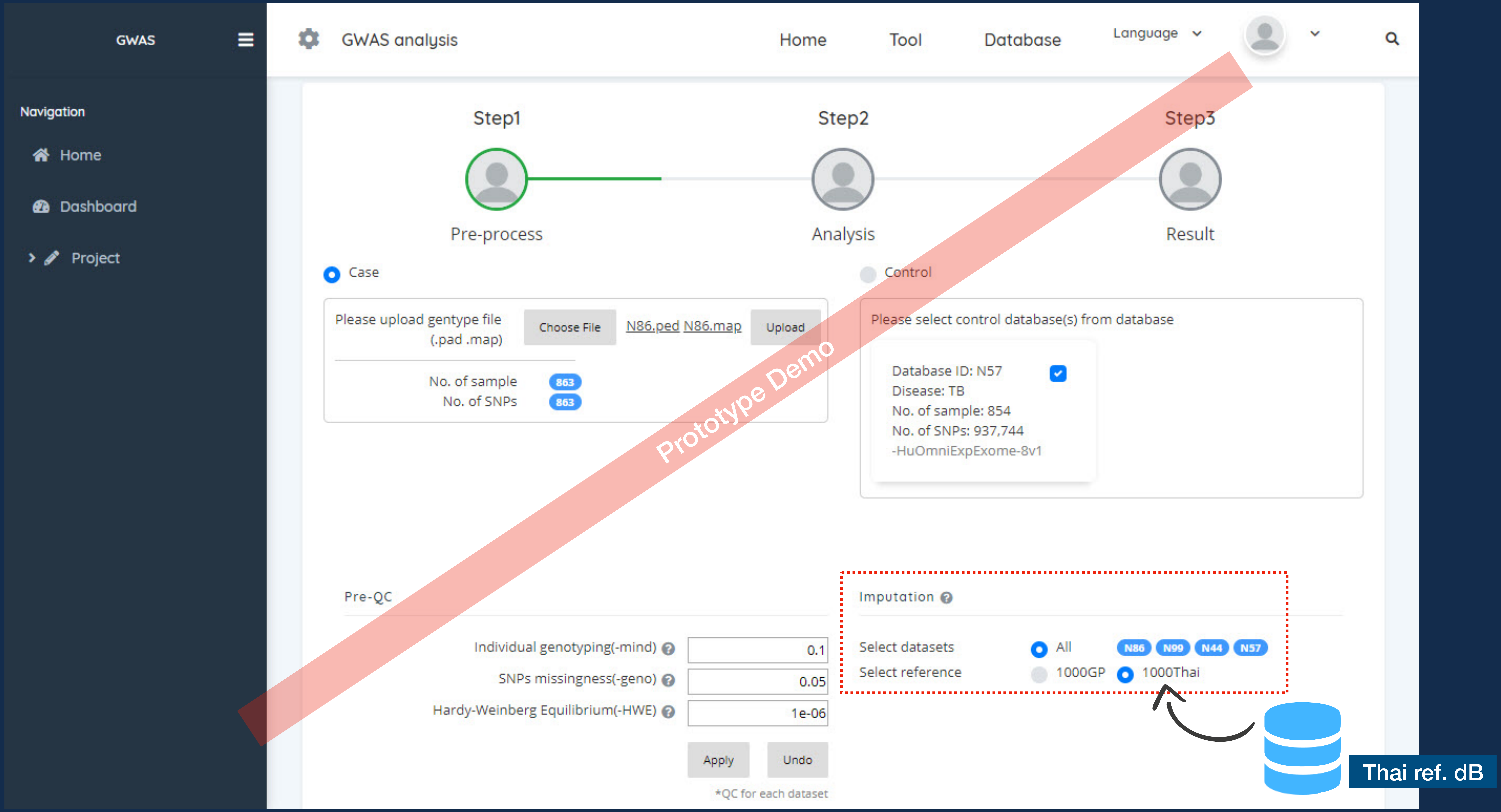
Prototype Demo

Gene	Diplotype	Function	Phenotype	Drug name	Guideline strength	Recommendations
CFTR	P67L/R117H			ivacaftor	moderate	Use ivacaftor accordingv to the product label.
CYP2C9	*9/*11	Normal function/ Possible decreased function	Possible Intermediate Metabolizer			
CYP2C9	*3/*8	Decreased function/ Posible decreased function	Possible Poor Metabolizer			
CYP2D6	*1xN/*55	No function/ No function	Ultrarapid Metabolizer	codeine	strong	Avoid codeine use due to potential for toxicity.
CYP2D6	*1xN/*55	No function/ No function	Ultrarapid Metabolizer	fluvoxamine	optional	No recommendation due to lack of evidence.
CYP2D6	*1xN/*55	No function/ No function	Ultrarapid Metabolizer	desipramine	optional	Avoid tricyclic use due to potential lack of efficacy. Consider alternative drug not metabolized by CYP2D6. If a TCA is warranted, consider titrating to a highe target dose (compared to normal

Genome-wide association study (GWAS)

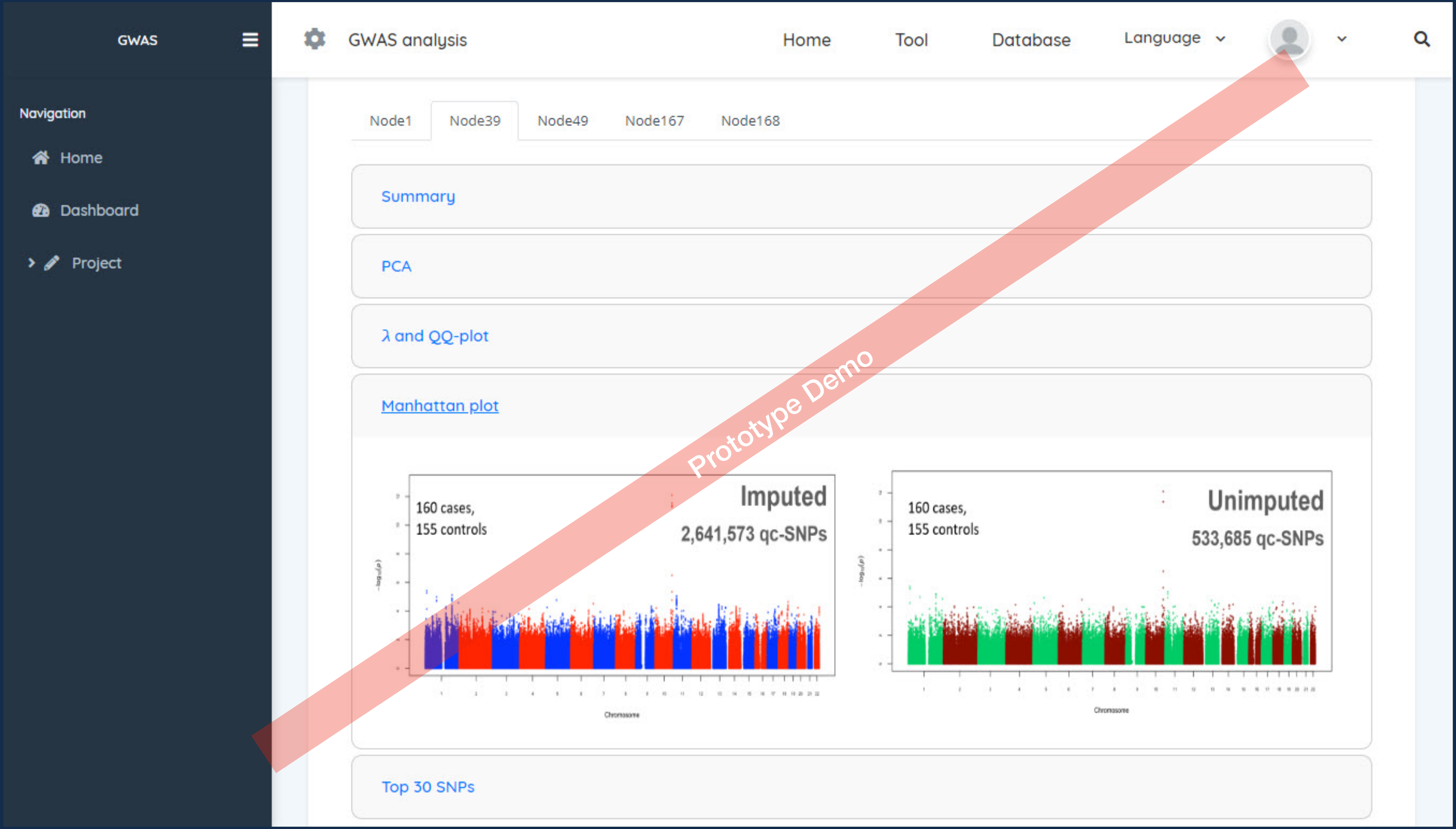


GWAS (cont'd)



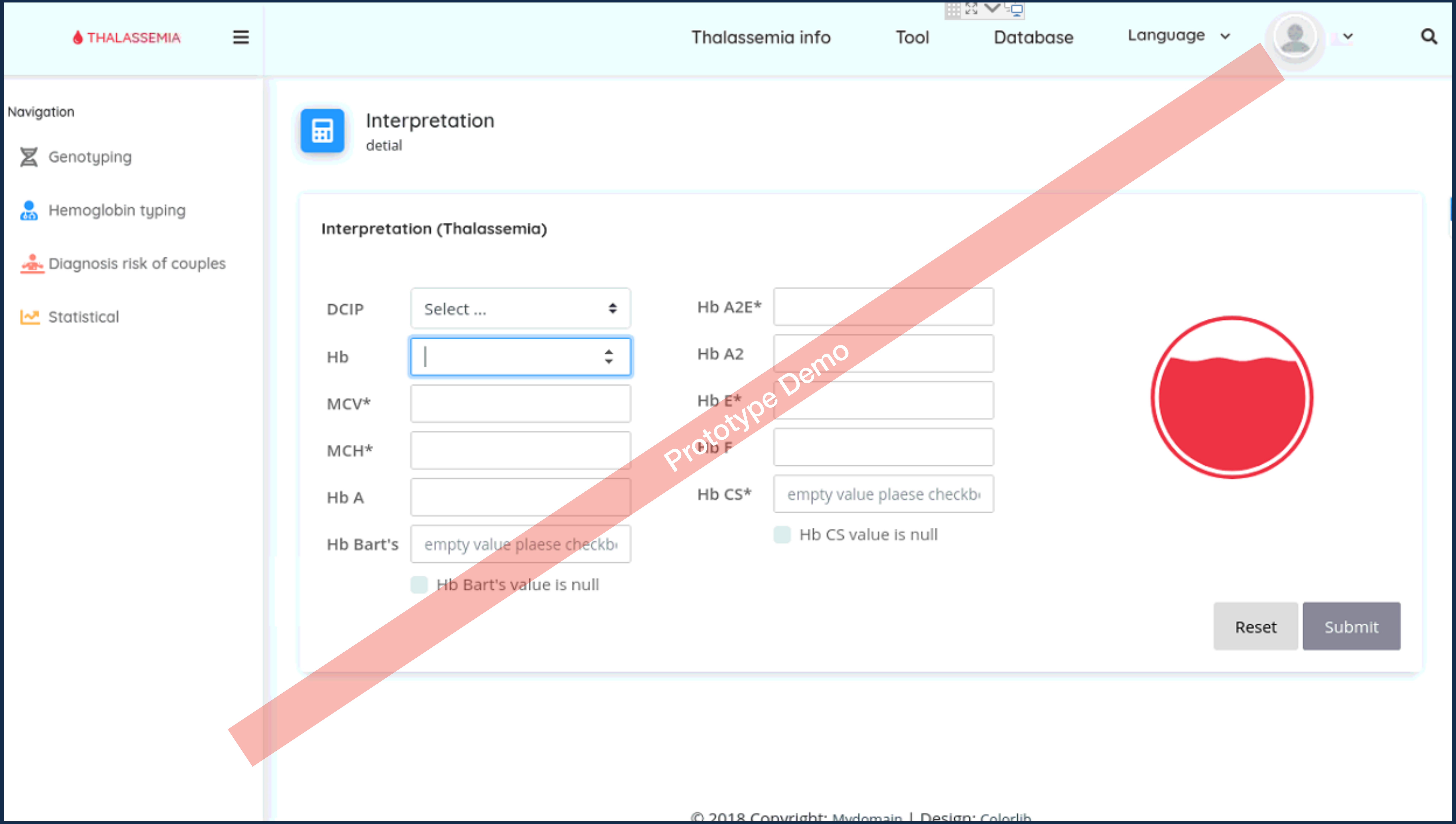
GWAS
Navigation
Home
Dashboard
Project
GWAS analysis
Home
Tool
Database
Language
Step1
Step2
Step3
Pre-process
Analysis
Result
Case
Control
Please upload gentype file (.pad .map)
Choose File
N86.ped N86.map
Upload
No. of sample 863
No. of SNPs 863
Please select control database(s) from database
Database ID: N57
Disease: TB
No. of sample: 854
No. of SNPs: 937,744
-HuOmniExpExome-8v1
Prototype Demo
Pre-QC
Individual genotyping(-mind) 0.1
SNPs missingness(-geno) 0.05
Hardy-Weinberg Equilibrium(-HWE) 1e-06
Apply
Undo
*QC for each dataset
Imputation
Select datasets
All
N86 N99 N44 N57
Select reference
1000GP
1000Thai
Thai ref. dB

GWAS (cont'd)



GWAS
Navigation
Home
Dashboard
Project
GWAS analysis
Home
Tool
Database
Language
Node1
Node39
Node49
Node167
Node168
Summary
PCA
λ and QQ-plot
Manhattan plot
160 cases,
155 controls
Imputed
2,641,573 qc-SNPs
160 cases,
155 controls
Unimputed
533,685 qc-SNPs
Top 30 SNPs
Prototype Demo

Thalassemia

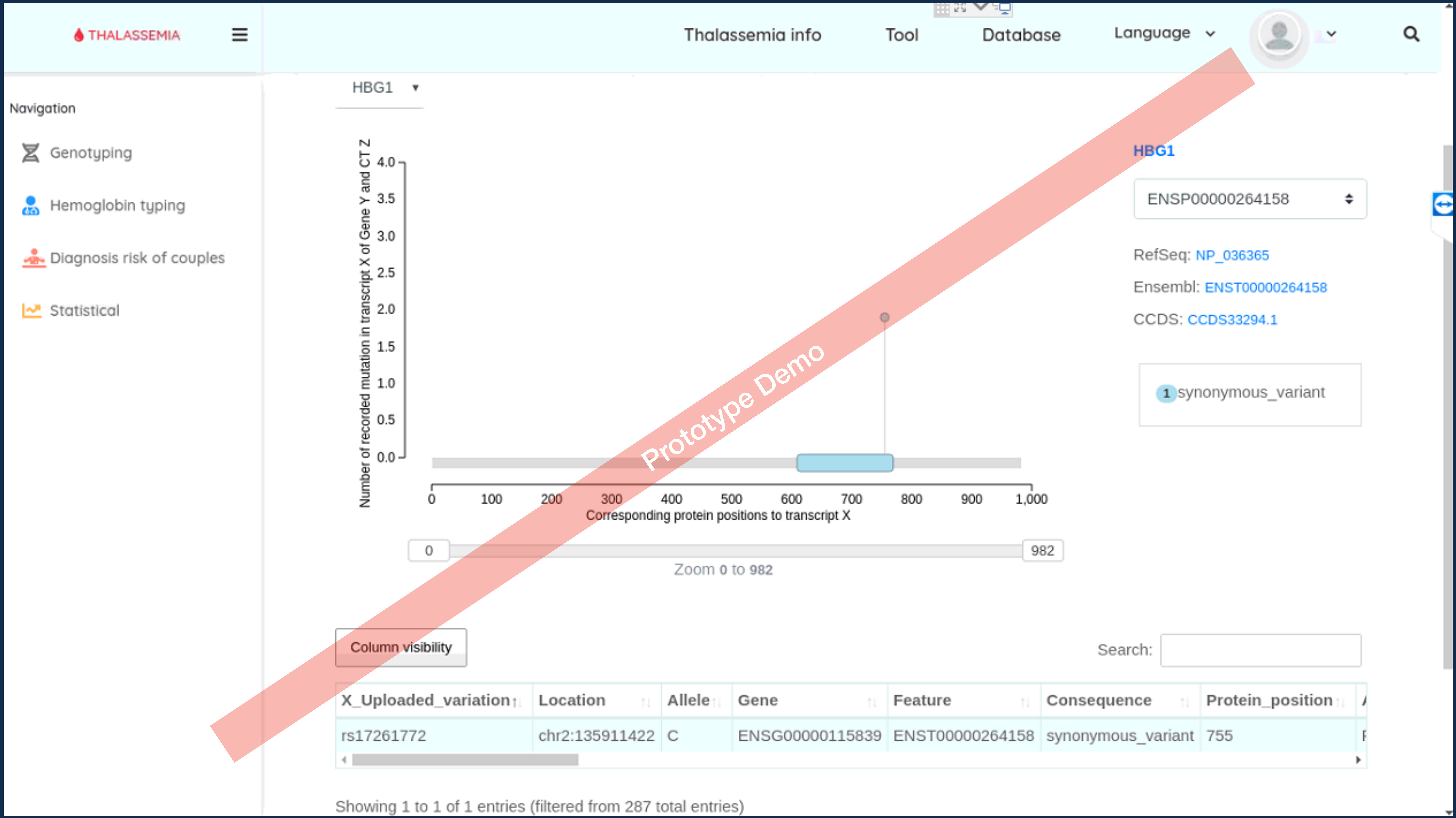


Wanwisa

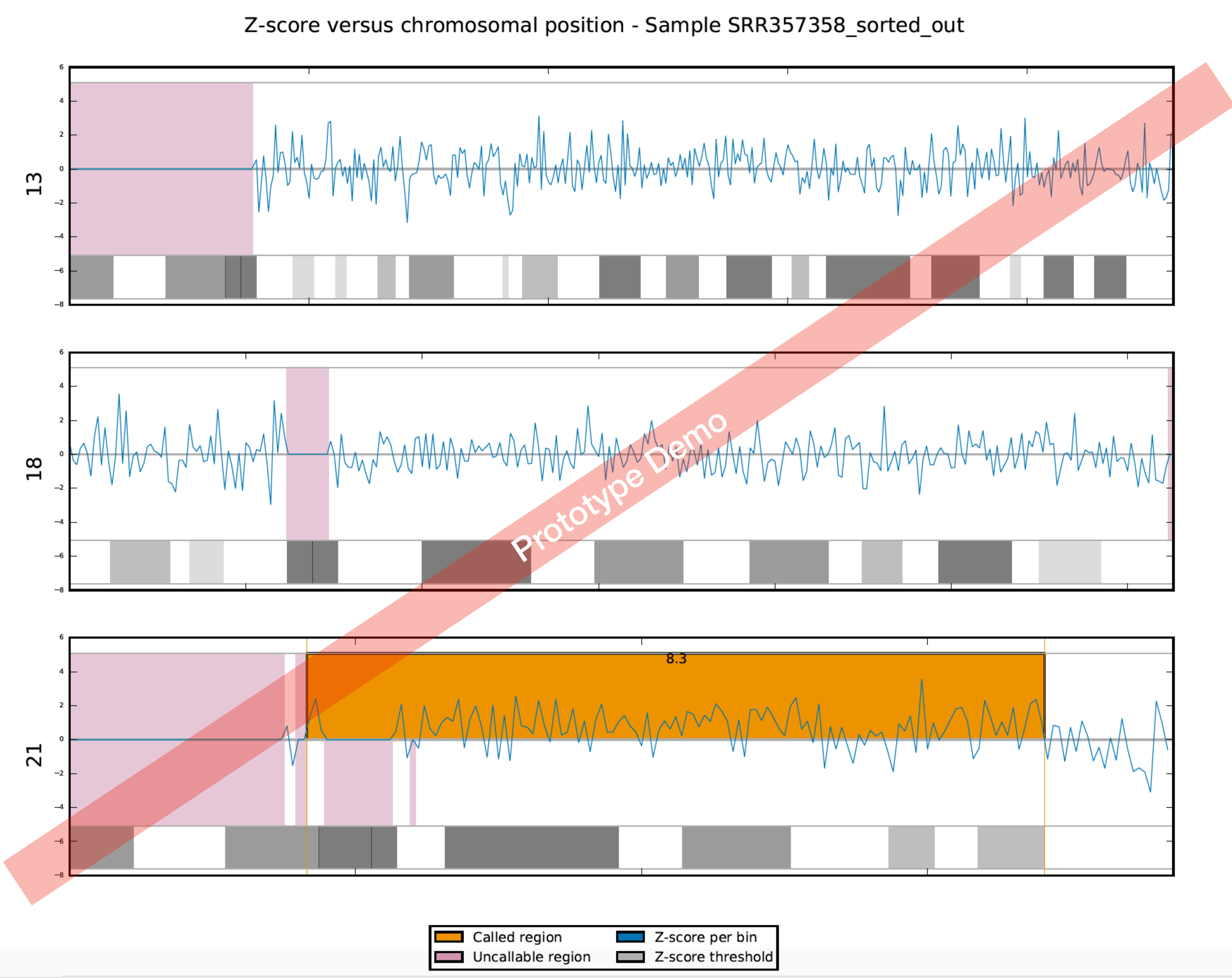
Watcharapot

Peerapat

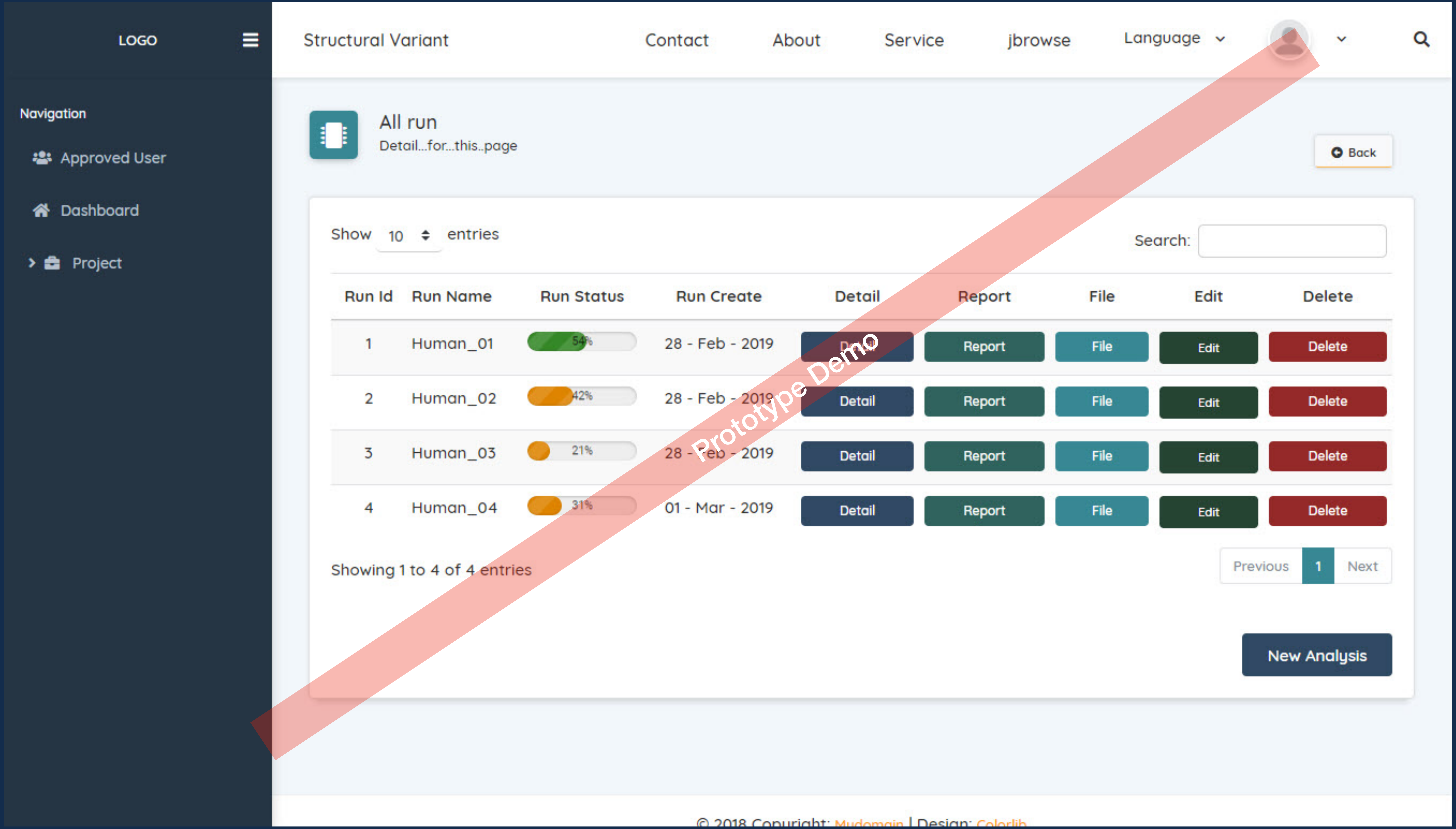
Thalassemia *(cont'd)*



Non-Invasive Prenatal Testing (NIPT)



Structural variants discovery




Worawich


Dr. Chumpol


Athiphat


Watcharapot

Structural variants discovery *(cont'd)*

Structural Variant | Contact | About | Service | jbrowse | Language

Circos graph

Detail...for...this..page

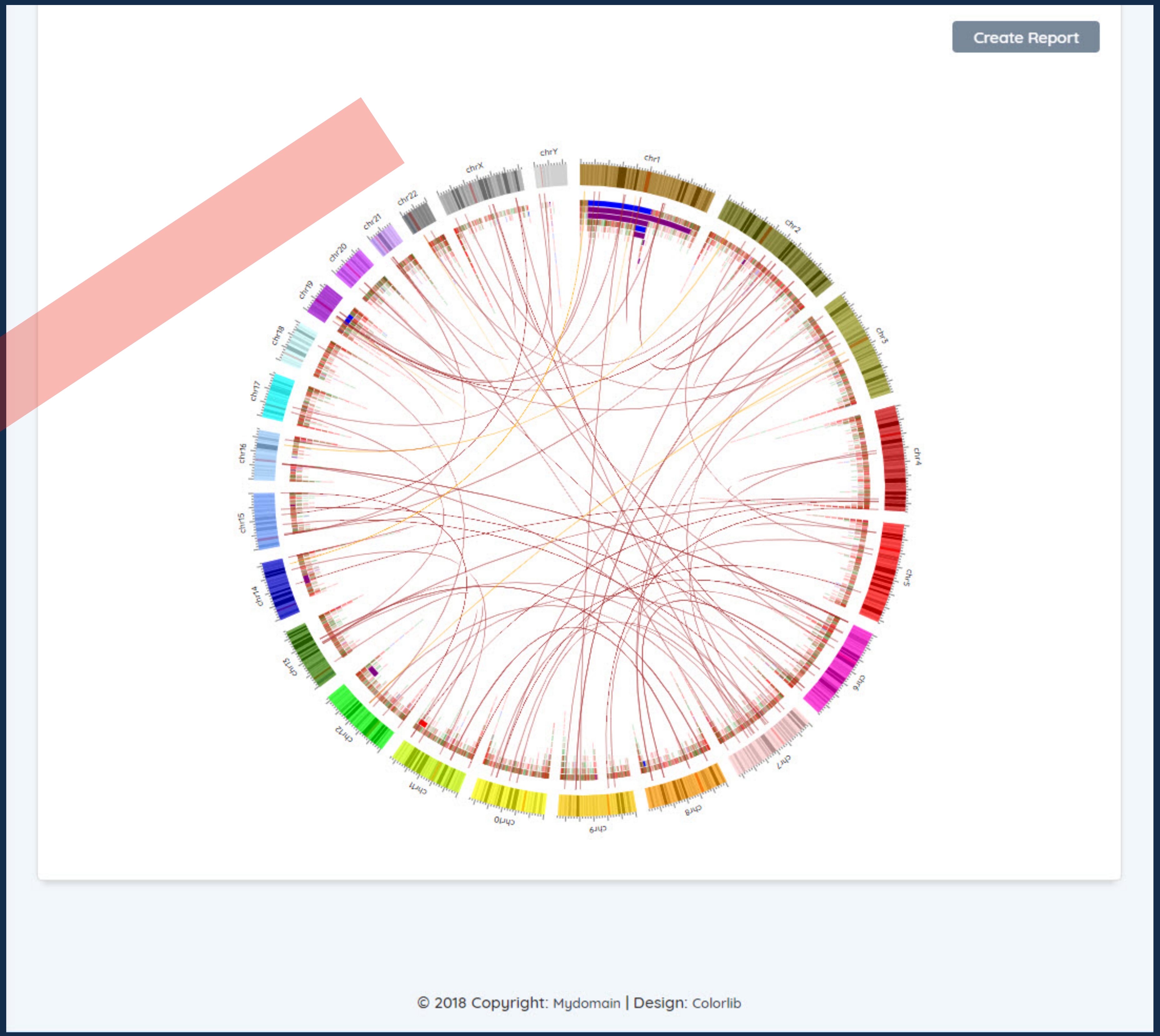
Back

Show 10 entries | Show Filter | Clear Filter | Search:

CHR	POS	QUAL	SVTYPE	SVLEN	CIPOS	GENES	JBrowse	Detail
chr1	766593	112	DEL	-2519	0, 3	LINC00115, LINC01128	JBrowse	Detail
chr1	823830	440	DEL	-87		LOC100130417, FAM41C	JBrowse	Detail
chr1	823929	413	DEL	-72		LOC100130417, FAM41C	JBrowse	Detail
chr1	869368	578	DEL	-225		SAMD11	JBrowse	Detail
chr1	869701	152	DEL	-616	0, 25	SAMD11	JBrowse	Detail
chr1	912125	144	DEL	-162		PERM1, PLEKHN1	JBrowse	Detail
chr1	934143	475	INS	58		PERM1, HES4	JBrowse	Detail
chr1	965600	162	DEL	-163		AGRN	JBrowse	Detail
chr1	1004204	487	INS	78	0, 10	AGRN, RNF223	JBrowse	Detail
chr1	1142719	171	DEL	-421		TNFRSF4, TNFRSF18	JBrowse	Detail

Showing 1 to 10 of 10,321 entries

Previous 1 2 3 4 5 ... 1033 Next

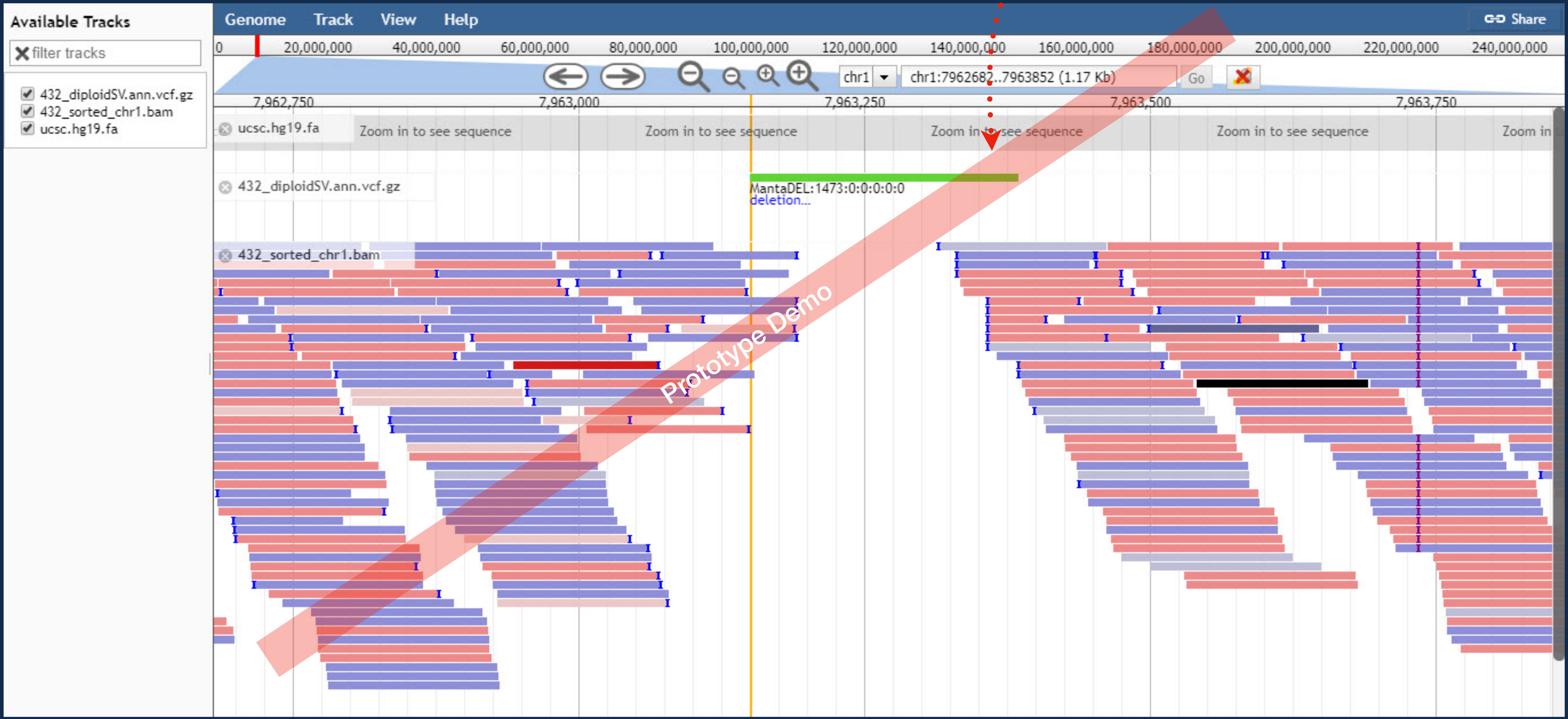


SVs

- Indels
- duplications
- inversions
- translocations

Structural variants discovery *(cont'd)*

Large deletion



Acknowledgement

All collaborators
NBT members