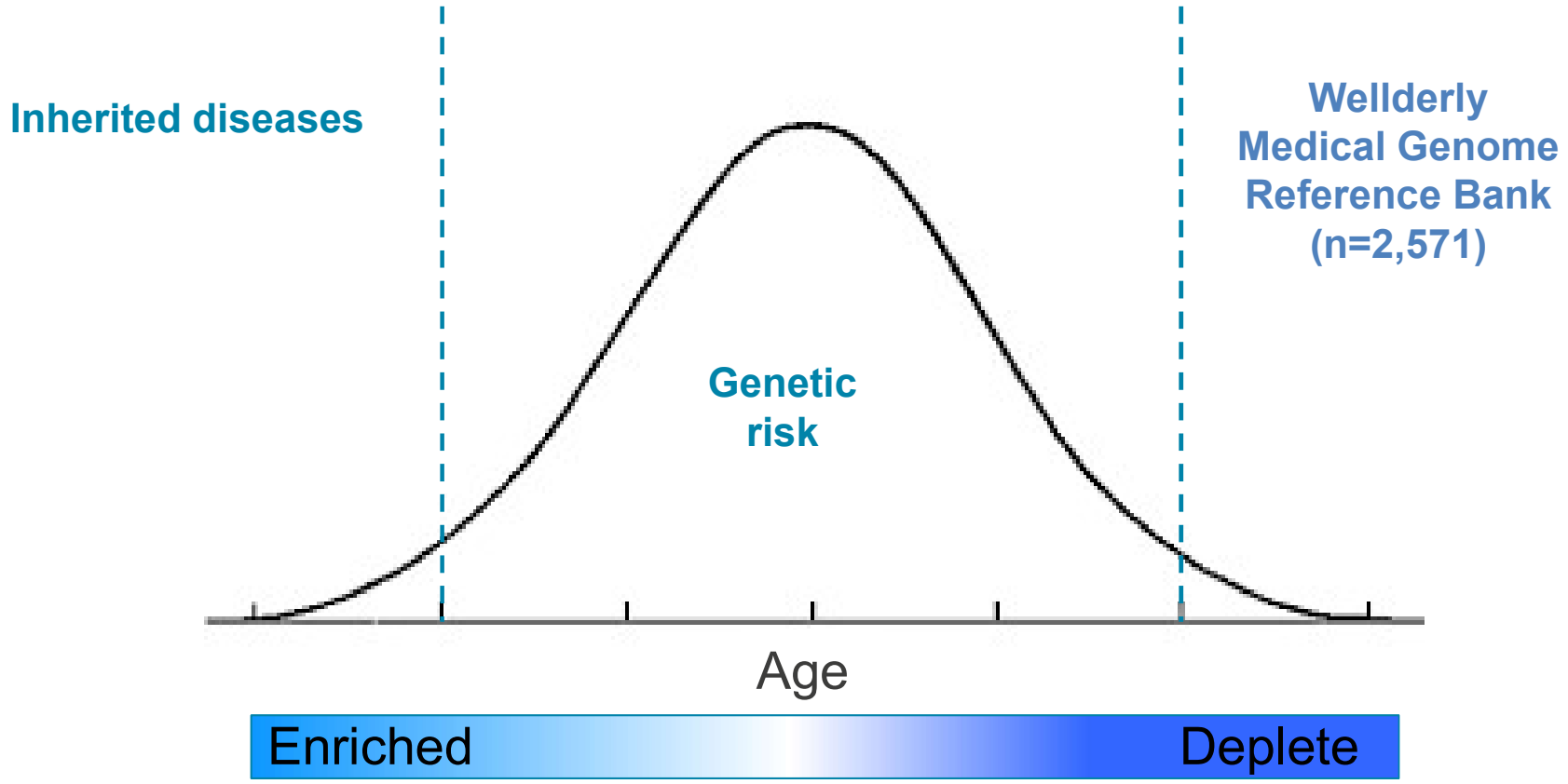
Garvan Institute
of Medical Research

The Medical Genome Reference Bank

A universal reference cohort of healthy aging

Extreme phenotype sampling design



The Medical Genome Reference Bank

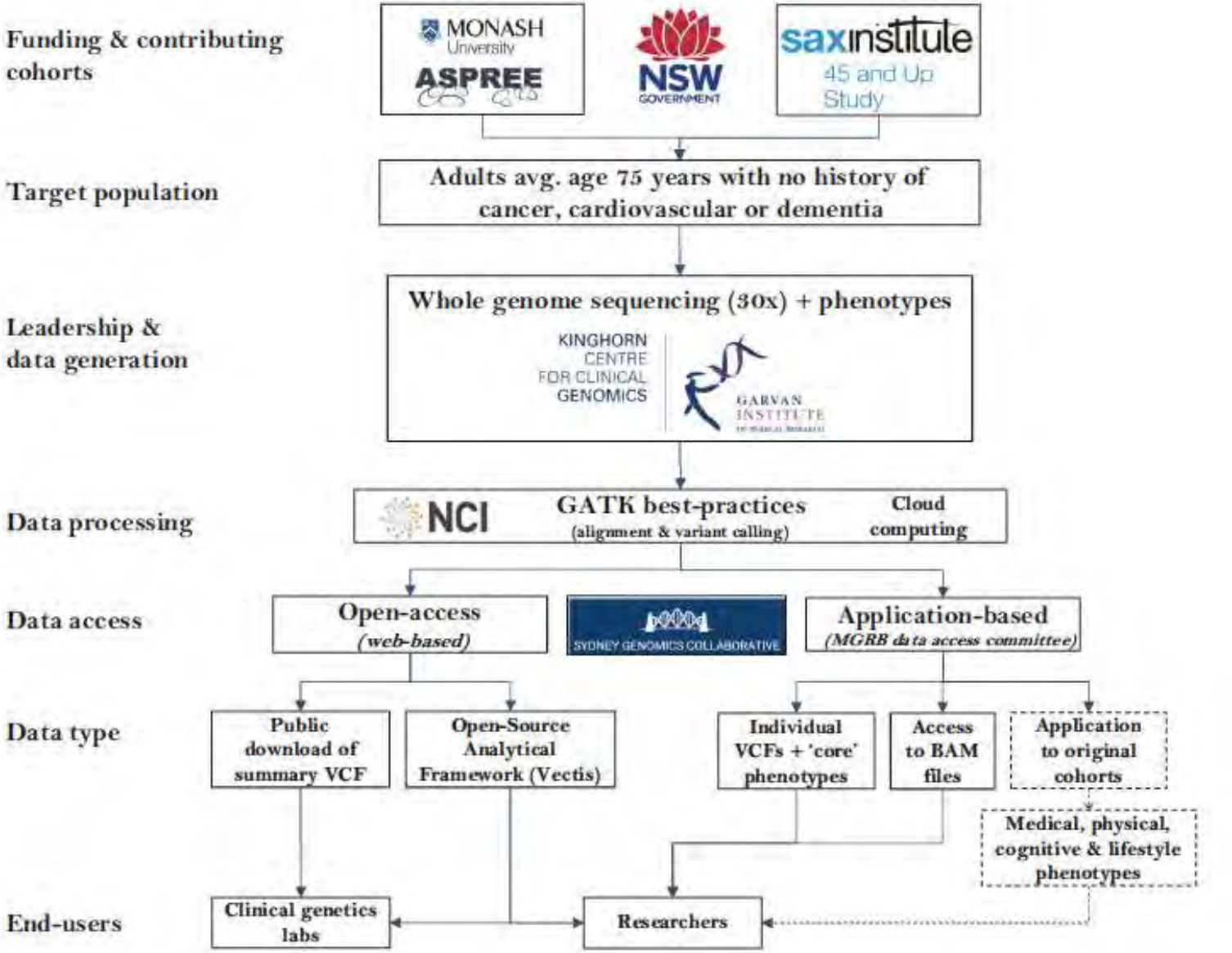
A publicly-available repository of genotypes of ~2,570 healthy elderly Australians.

Measure	***ASPREE***	***45 and Up***
Individuals	1,853	717
(percent female)	48.20%	59.30%
Age at blood draw (years)	79	70
	(75 – 95)	(64 – 91)
Height (m)	1.65	1.66
	(1.33 – 1.91)	(1.37 – 1.91)
Mass (kg)	74.5	72
	(33.4 – 127.1)	(36.0 – 147.0)
Mean sequencing depth (genome-wide)	38	39
	(26.8 – 46.0)	(27.3 – 45.5)
Genetic background		
Non-Finnish European	1,805	695
South and Central American	23	5
South Asian	14	6
Finnish European	10	7
East Asian	1	4

ASPREE

THE 45 AND UP STUDY
Research to improve health and wellbeing

Paul Lacaze, John McNeil, Martin McNamara, Sally Redman

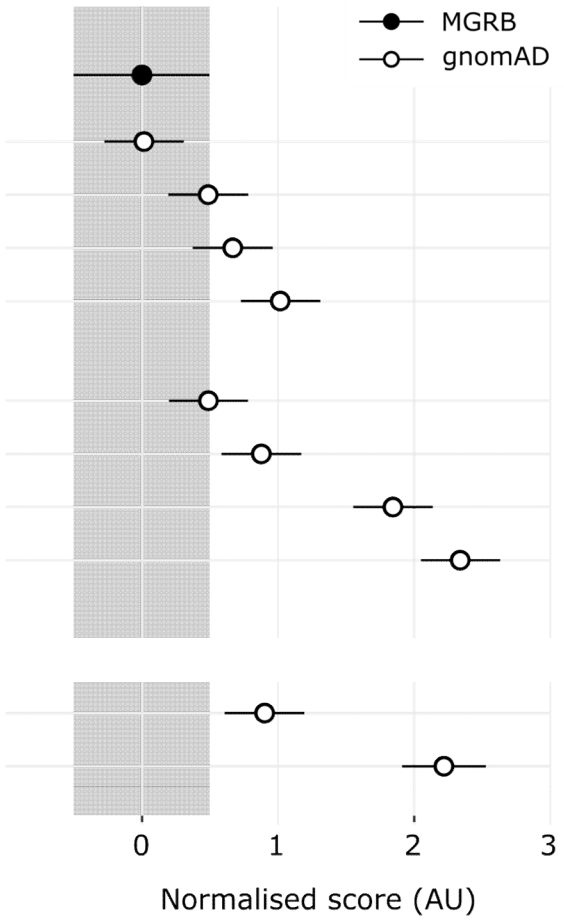

Funding & contributing cohorts
MONASH University
ASPREE
NSW GOVERNMENT
saxinstitute
45 and Up Study
Target population
Adults avg. age 75 years with no history of cancer, cardiovascular or dementia
Leadership & data generation
Whole genome sequencing (30x) + phenotypes
KINGHORN CENTRE FOR CLINICAL GENOMICS
GARVAN INSTITUTE
Data processing
NCI
GATK best-practices
(alignment & variant calling)
Cloud computing
Data access
Open-access
(web-based)
SYDNEY GENOMICS COLLABORATIVE
Application-based
(MGRB data access committee)
Data type
Public download of summary VCF
Open-Source Analytical Framework (Vectis)
Individual VCFs + 'core' phenotypes
Access to BAM files
Application to original cohorts
Medical, physical, cognitive & lifestyle phenotypes
End-users
Clinical genetics labs
Researchers

	MGRB	ExAC [4]	GnomAD [4]	UKBB SNPs [28]	HLI - JCVI [13]	Wellderly STSI [12]	SweGen [11]	HGVD [7]
Approx. cohort size (Feb 2018)	4000	60,000	140,000	500,000	10,000	600	1000	3200
Purpose-built cohort (versus data aggregation)	✓	X	X	✓	✓	✓	✓	✓
Whole genome sequencing	✓	X	✓	X	✓	✓	✓	X
Ability to detect complex and SV	✓	X	✓	X	✓	X	✓	X
Phenotype data to confirm absence of disease	✓	X	X	✓	✓	✓	X	?
Confirmed healthy elderly population	✓	X	X	X	X	✓	X	X
Allele frequencies made readily accessible	✓	✓	✓	X	X	X	✓	✓
Formal data access and approval policy	✓	X	X	✓	X	X	✓	X
Access provided to individual VCFs	✓	X	X	X	X	X	X	X
$n \geq 4000$ samples	✓	✓	✓	✓	✓	X	X	X
Consistent and compatible seq. technology	✓	✓	✓	X	✓	X	✓	✓

MGRB Medical Genome Reference Bank, *ExAC* Exome Aggregation Consortium, *GnomAD* Genome Aggregation Database, *UKBB SNPs* U.K. Biobank SNP data set, *HLI-JCVI* Human Longevity Inc - J. Craig Venter Institute, *STSI Wellderly* Scripps Translational Science Institute Wellderly study, *SweGen* Swedish Genome reference population project, *HGVD* Human Genetic Variation Database (Japan)

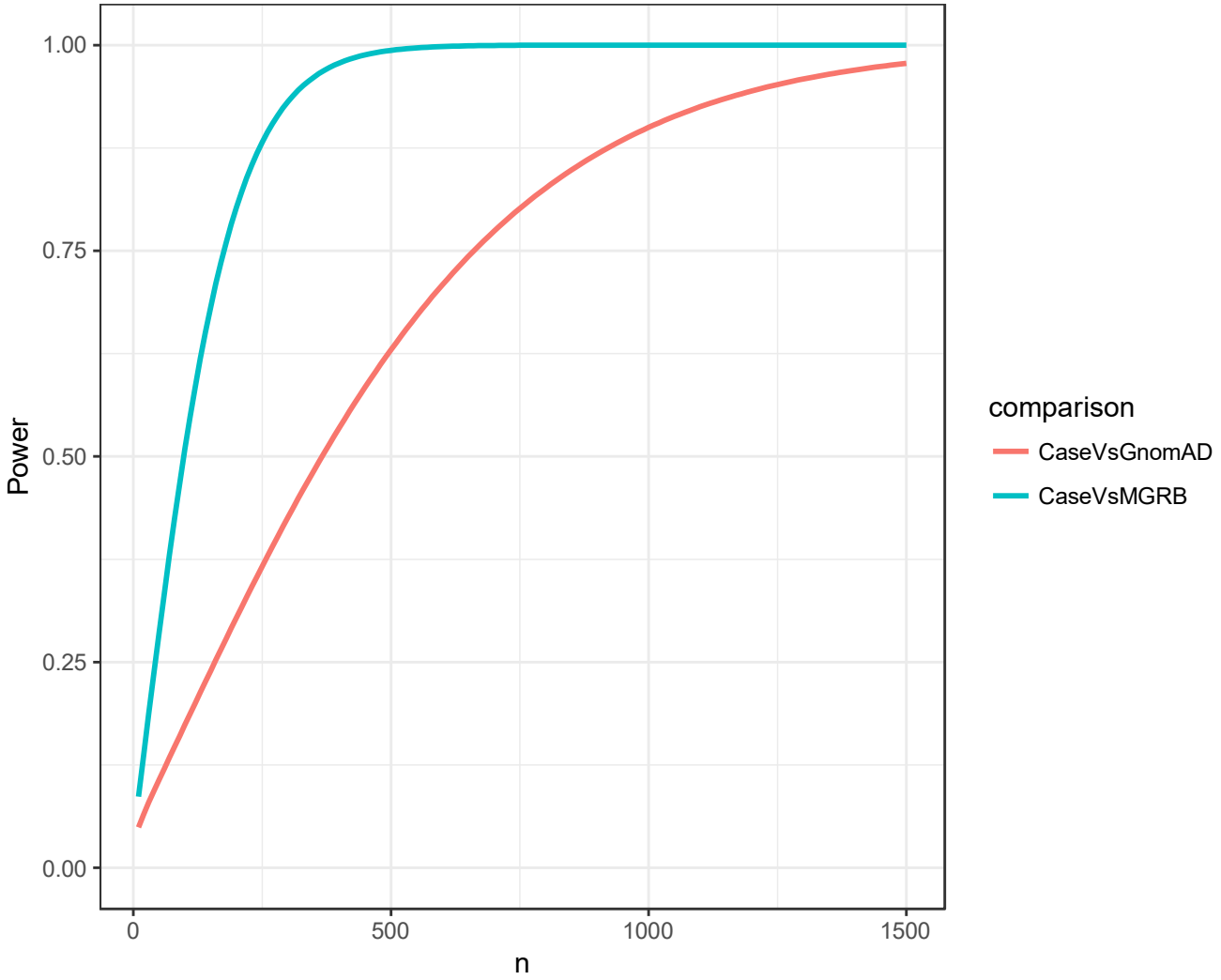
Common variant burden in the wellderly

Polygenic score	Normalised score	P-value
Colorectal cancer	0.01	1.000
Melanoma	0.49	0.392
Breast cancer	0.67	0.122
Prostate cancer	1.01	0.008
Systolic blood pressure	0.49	0.392
Pulse pressure	0.88	0.022
Diastolic pressure	1.84	< 0.001
EOCAD	2.34	< 0.001
Alzheimer's disease	0.90	0.020
Short lifespan	2.22	< 0.001



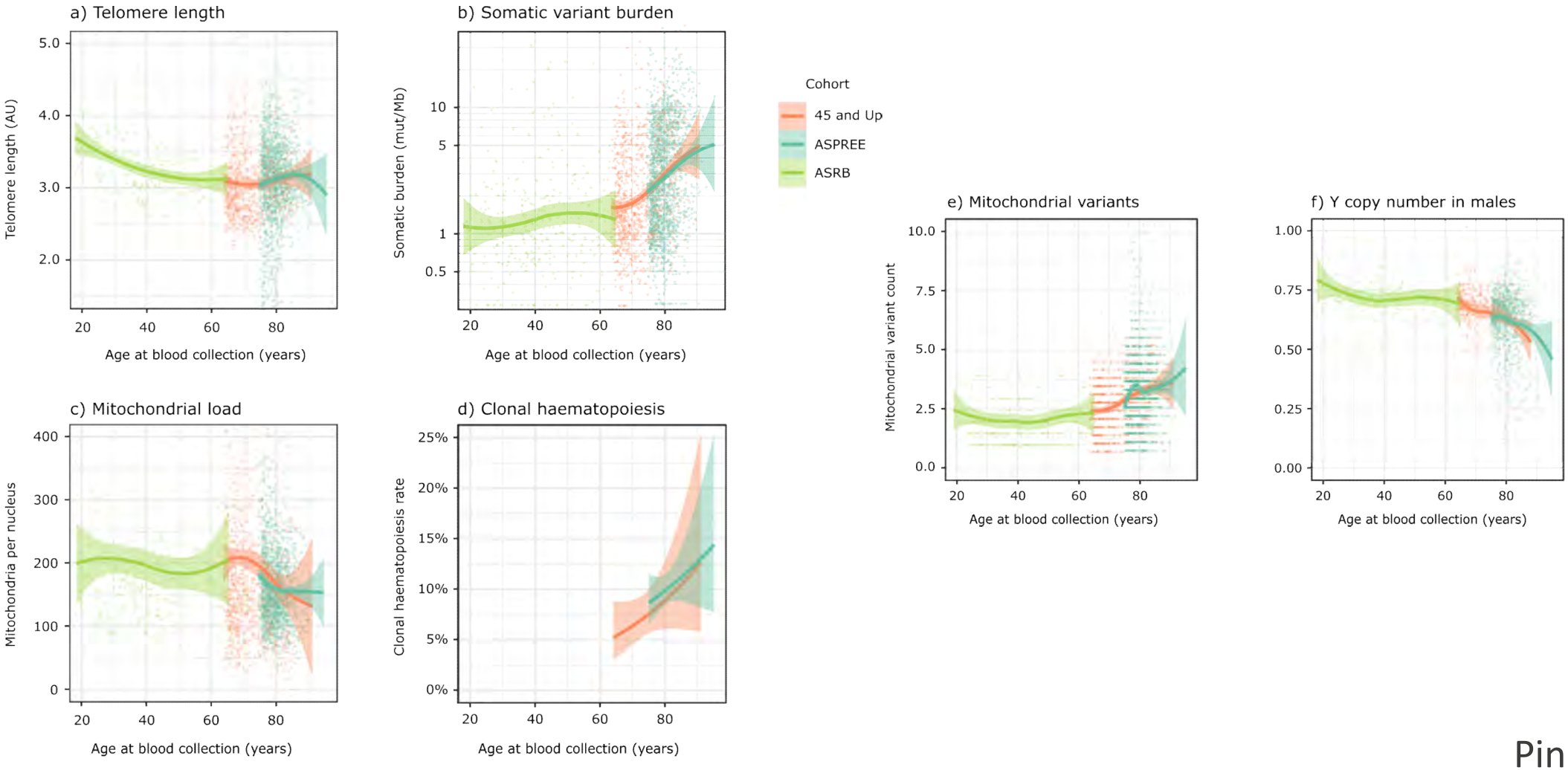
Pinese

The tripled power of extreme phenotype sampling



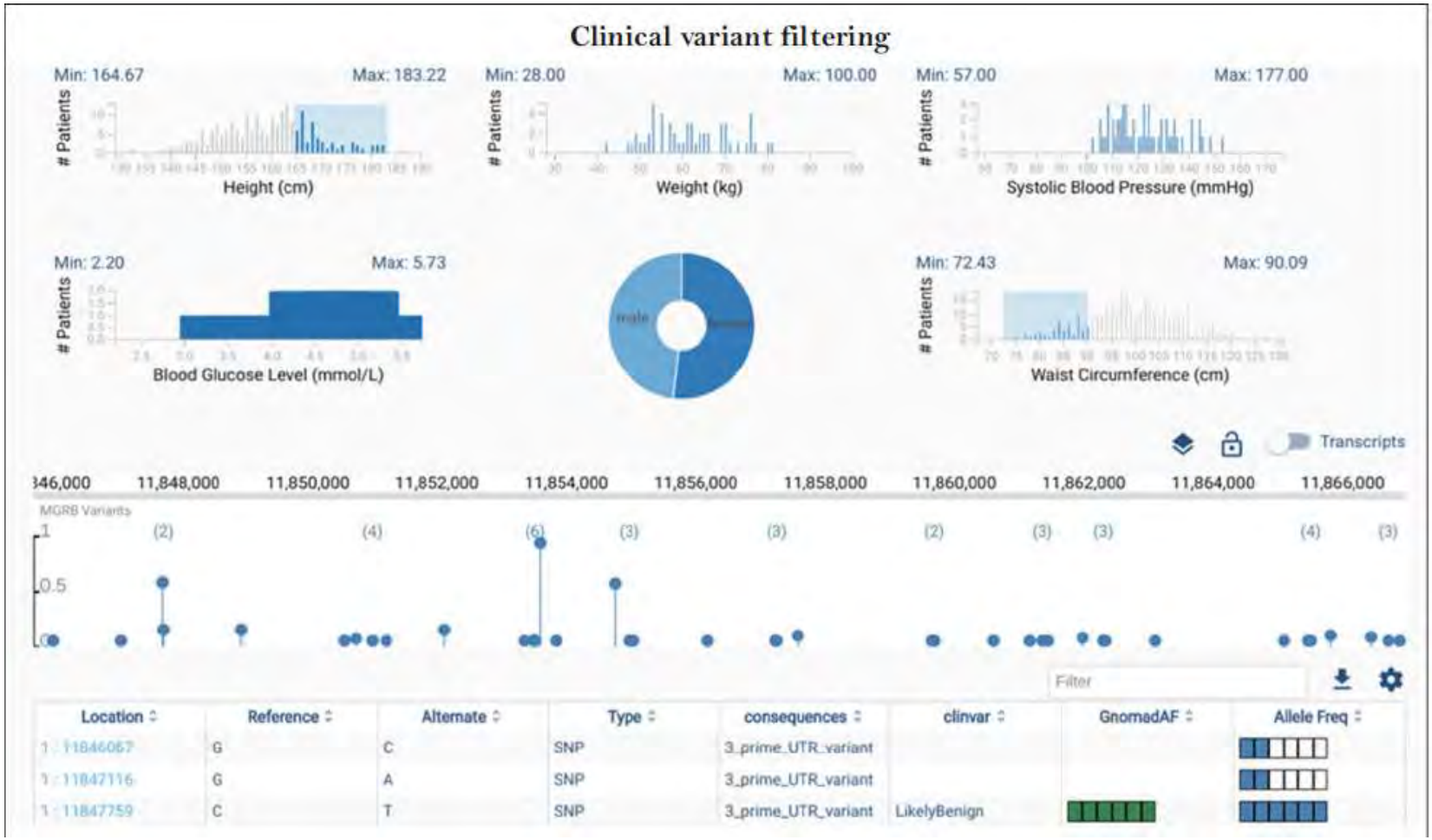
Pinese

Somatic variation and ageing



Pinese

Vectis: a searchable database of genetic variation




© Mapbox © OpenStreetMap Improve the underlying map
Canada
United States
Mexico
CU
Venezuela
Colombia
EC
Peru
Brazil
Bolivia
PY
UY
Argentina
CL
South Pacific Ocean
North Atlantic Ocean
South Atlantic Ocean
Scotia Sea
IS
Norwegian
SE
FI
NO
PL
BY
UA
ES
GR
MA
TN
Algeria
Libya
Egypt
MR
Mali
Niger
SN
NG
GH
TD
Sudan
CM
SS
ET
SO
KE
CD
TZ
AO
MZ
BW
South Africa
MG
Russia
Kazakhstan
Mongolia
UZ
KG
TM
TJ
China
Japan
Iran
IQ
KW
Saudi Arabia
AF
PK
NP
BD
India
MM
LA
TH
VN
PH
OM
Arabian Sea
RY
LK
MV
Malaysia
Indonesia
PG
Indian Ocean
Aust
Sea
Tasman Sea
NZ
Mapbox

Users of MGRB

- Tier 1 access publicly available (date)
- 40 applications to use the MGRB
  - 11 International
  - 12 Interstate
  - 17 NSW
- 3 manuscripts published or in press, several under way

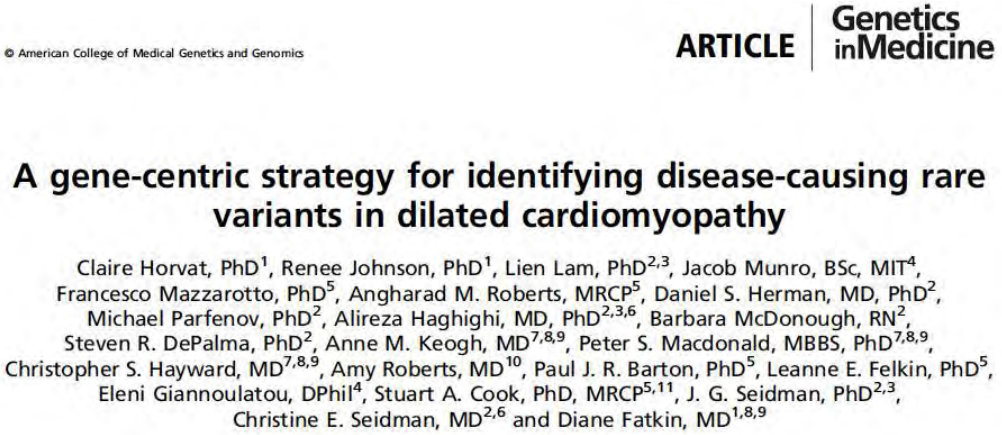
European Journal of Human Genetics (2019) 27:308–316
https://doi.org/10.1038/s41431-018-0279-z

ESHG

ARTICLE

The Medical Genome Reference Bank: a whole-genome data resource of 4000 healthy elderly individuals. Rationale and cohort design

Paul Lacaze[1] · Mark Pinese[2,3] · Warren Kaplan[2] · Andrew Stone[2] · Marie-Jo Brion[2] · Robyn L. Woods[1] · Martin McNamara[4] · John J. McNeil[1] · Marcel E. Dinger[2,3] · David M. Thomas[2,3]



© American College of Medical Genetics and Genomics

ARTICLE | Genetics inMedicine

A gene-centric strategy for identifying disease-causing rare variants in dilated cardiomyopathy

Claire Horvat, PhD[1], Renee Johnson, PhD[1], Lien Lam, PhD[2,3], Jacob Munro, BSc, MIT[4], Francesco Mazzarotto, PhD[5], Angharad M. Roberts, MRCP[5], Daniel S. Herman, MD, PhD[2], Michael Parfenov, PhD[2], Alireza Haghighi, MD, PhD[2,3,6], Barbara McDonough, RN[2], Steven R. DePalma, PhD[2], Anne M. Keogh, MD[7,8,9], Peter S. Macdonald, MBBS, PhD[7,8,9], Christopher S. Hayward, MD[7,8,9], Amy Roberts, MD[10], Paul J. R. Barton, PhD[5], Leanne E. Felkin, PhD[5], Eleni Giannoulatou, DPhil[4], Stuart A. Cook, PhD, MRCP[5,11], J. G. Seidman, PhD[2,3], Christine E. Seidman, MD[2,6] and Diane Fatkin, MD[1,8,9]