A GENOME-WIDE ASSOCIATION STUDY OF RHEUMATOID ARTHRITIS IN BLACK SOUTH AFRICANS

Evans M. Mathebula[1,2], Ananyo Choudhury[2], Nimmisha Govind[3], Mohammed Tikly[3], Michele Ramsay[1,2]

[1]Sydney Brenner Institute for Molecular Biosciences, Faculty of Health Sciences, [2]Division of Human Genetics, School of Pathology and National Health Laboratory Services, Faculty of Health Sciences, and [3]Division of Rheumatology, University of Witwatersrand, Johannesburg, South Africa

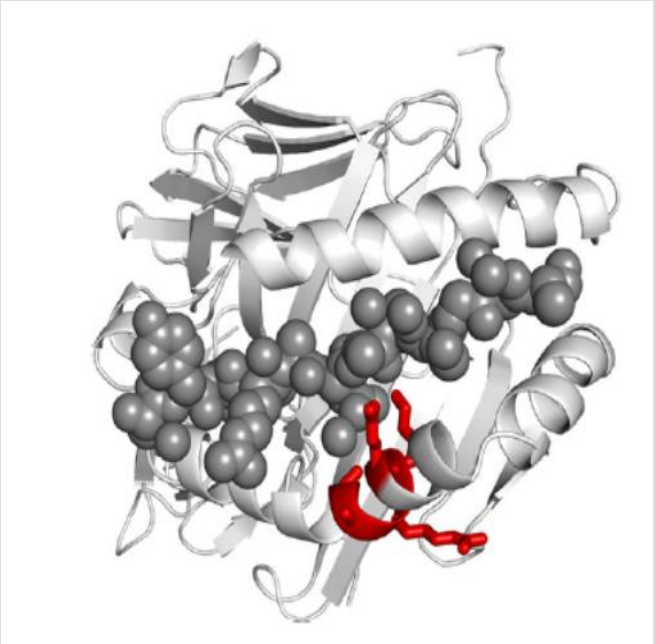
08 March 2019

RHEUMATOID ARTHRITIS

- Chronic, systemic inflammation primarily affecting small joints
- Affects ~1% adult population worldwide
- Genetic factors and environmental factors interaction
- E.g. Smoking

GENETICS OF RA

- Heritability ~40-60%
- HLA contribute 30% of risk
- HLA-DRB1 alleles – Shared Epitope
- Over 100 non-HLA SNPs
- PTPN22, PADI4, CTLA4, STAT4 …



Gregersen et al. 1987

Bax et al. 2011

AIM

- Using a genome-wide association approach the aim of the study was to identify genetic variants associated with susceptibility with seropositive RA in Black South Africans

PATIENTS

- Patients were recruited from the arthritis clinic at Chris Hani Baragwanath Academic Hospital (CHBAH), Soweto

- 647 unrelated, unselected
- 2010 ACR/EULAR classification criteria for RA
- Seropositive for RF/ACPA
- ≥18 years at disease onset
- 'Black'= if all 4 grandparents were black South African
- Written informed consent was obtained

- Study was approved by the University of the Witwatersrand Human Ethics department

CONTROLS

- 1612 Recruited from Soweto as part of the AWI-Gen project with minimal data (Sex, Age, Smoking)

- Ethnically and geographically matched

H3A SNP ARRAY

- GWAS SNP array
- H3Africa initiative, ASHG and Welcome trust
- African enriched
- Comprised of ~2.4M SNPs

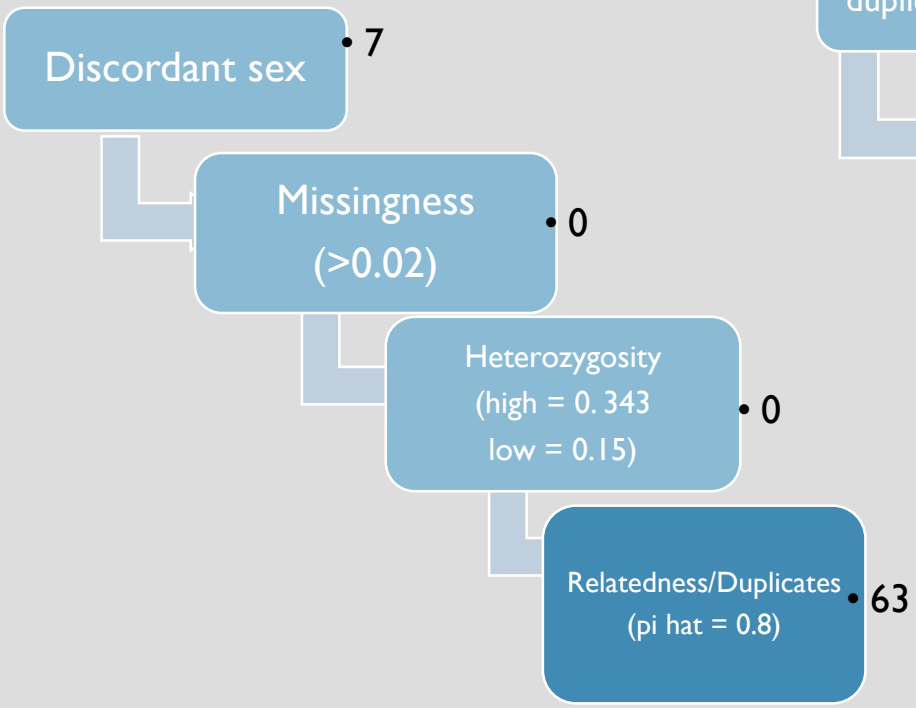
DEMOGRAPHIC AND CLINICAL CHARACTERISTICS

Variable	Cases (n=577)	Controls (n=1612)
Sex, Female(%)	86	58
Age, Mean (SD) Years	56±12.5	52±4.5
Disease duration, Mean(SD) Years	4.0±2.4	-
Smoking, Positive (%)	19	35.9

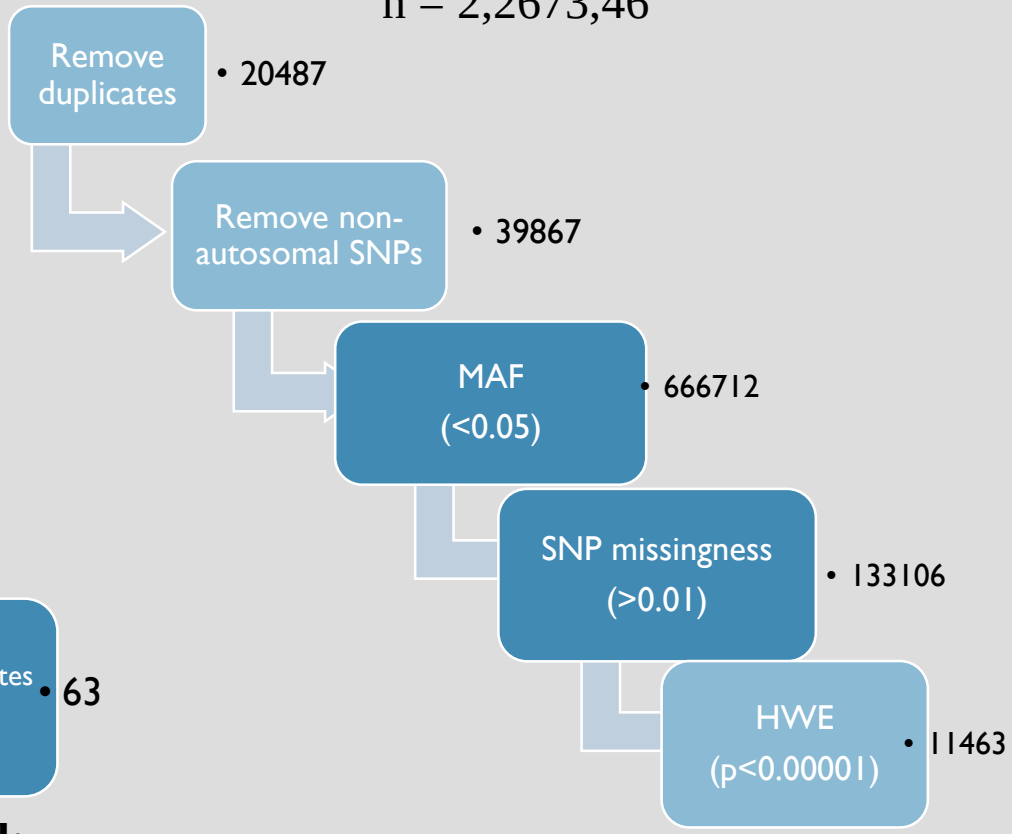
DATA QUALITY CONTROL

H3Africa GWAS QC pipeline (https://github.com/h3abionet/h3agwas)

Samples QC
n = 647



SNP QC
n = 2,2673,46



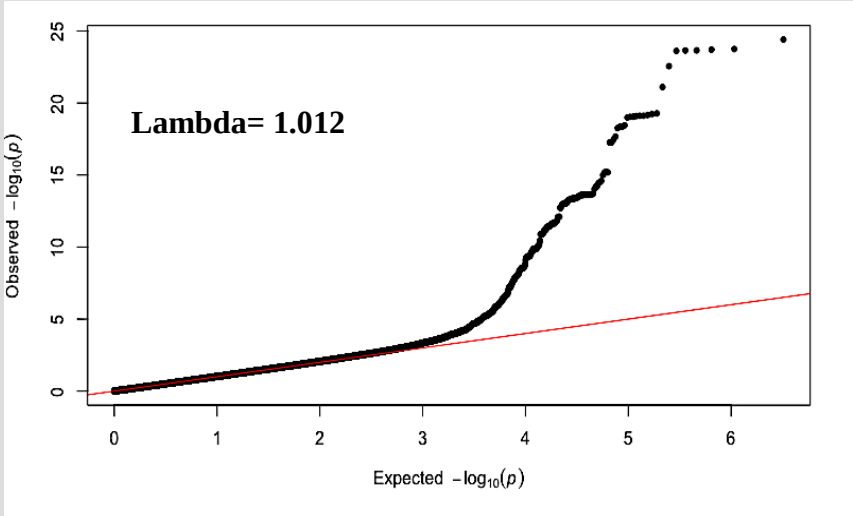
Total:
Cases = 577
SNPs = 1,395,711

STATISTICAL ANALYSIS

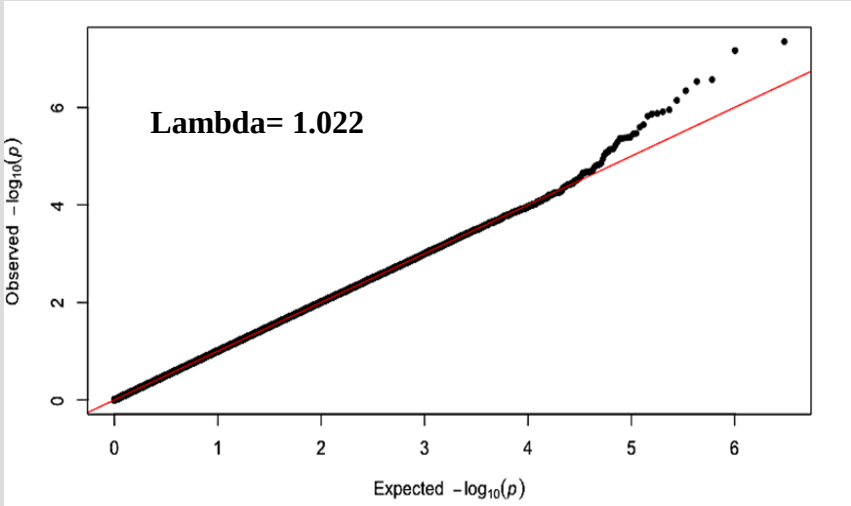
- Data analysis was performed in Plink v1.9 and R v3.5.1.
- Association was controlled for sex, age, smoking and the first 3 PCs
- A $p \leq 5 \times 10E\text{-}08$ was considered genome-wide significant

QUANTILE-QUANTILE PLOTS

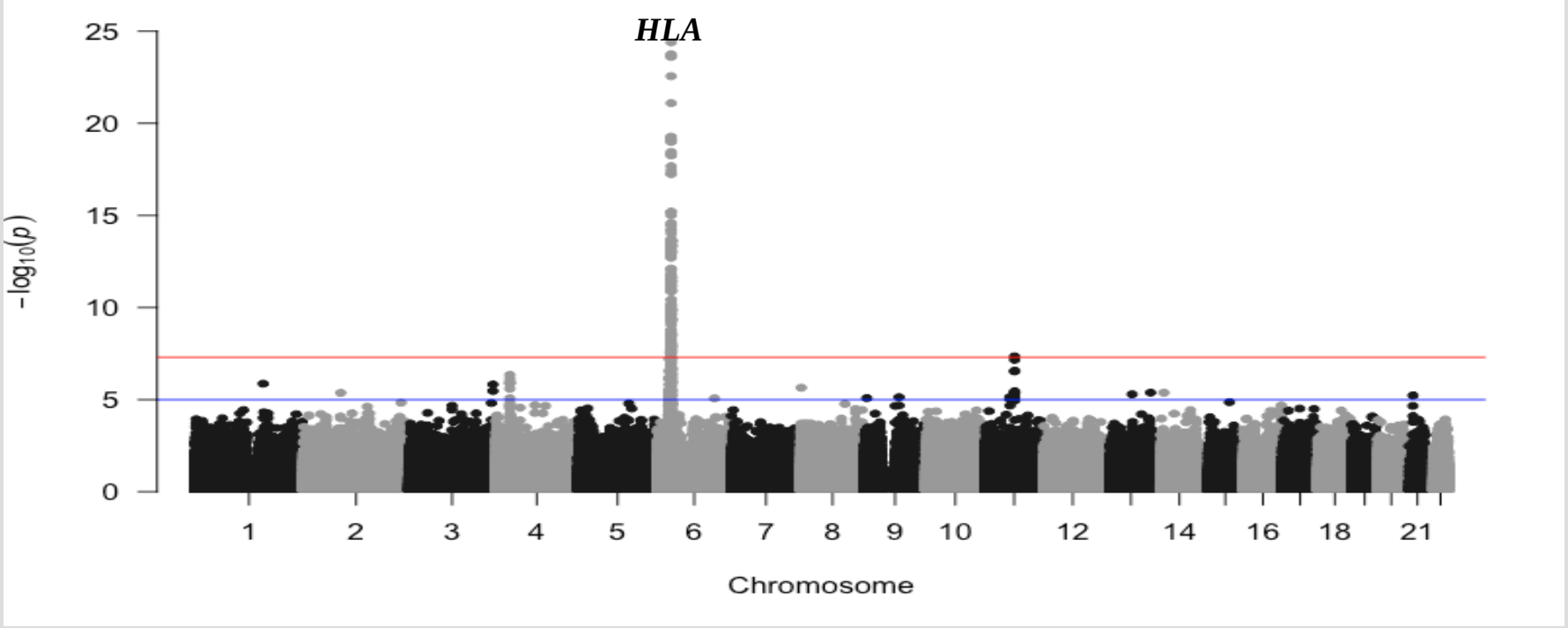
With CHR6



Without CHR6



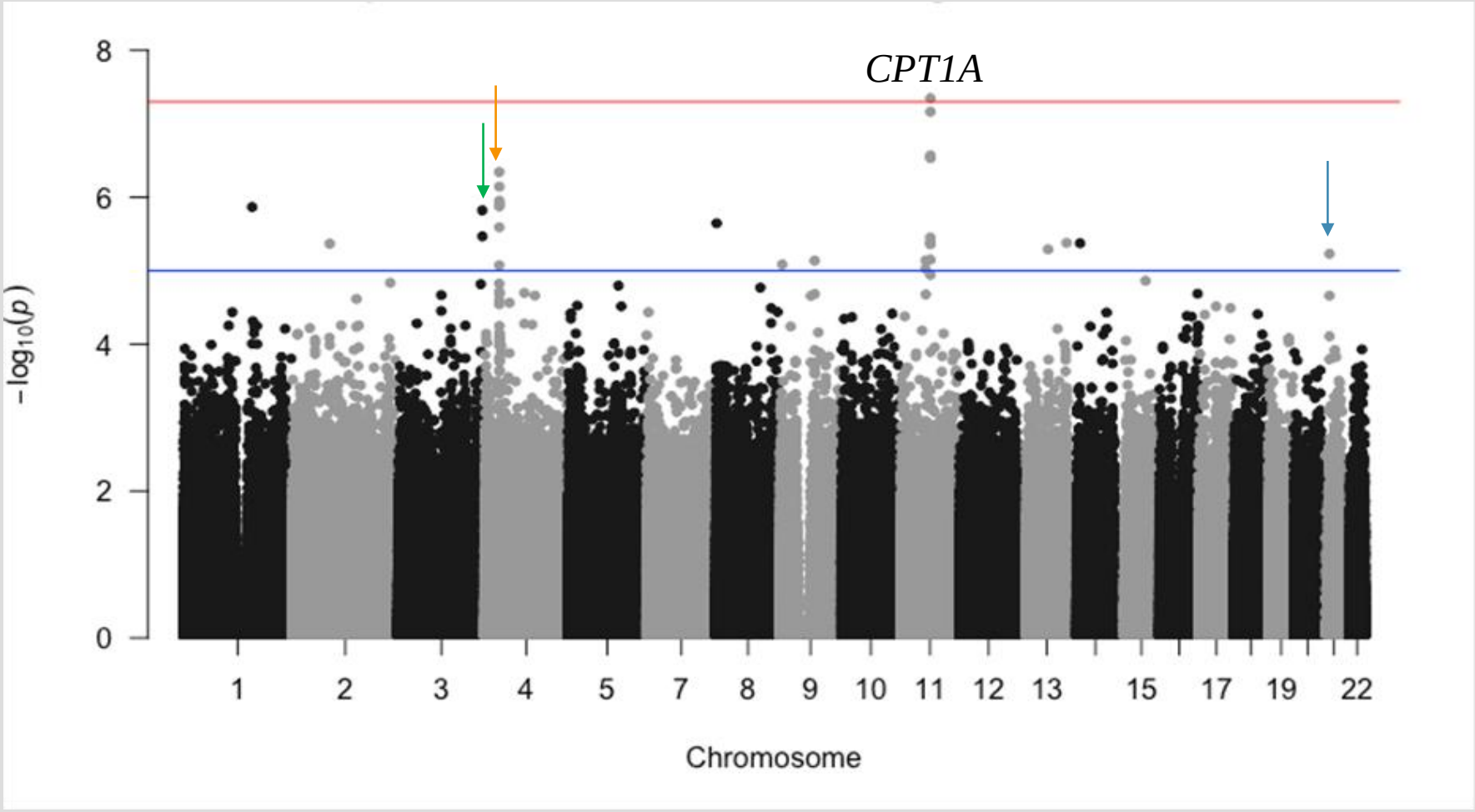
MANHATTAN PLOTS



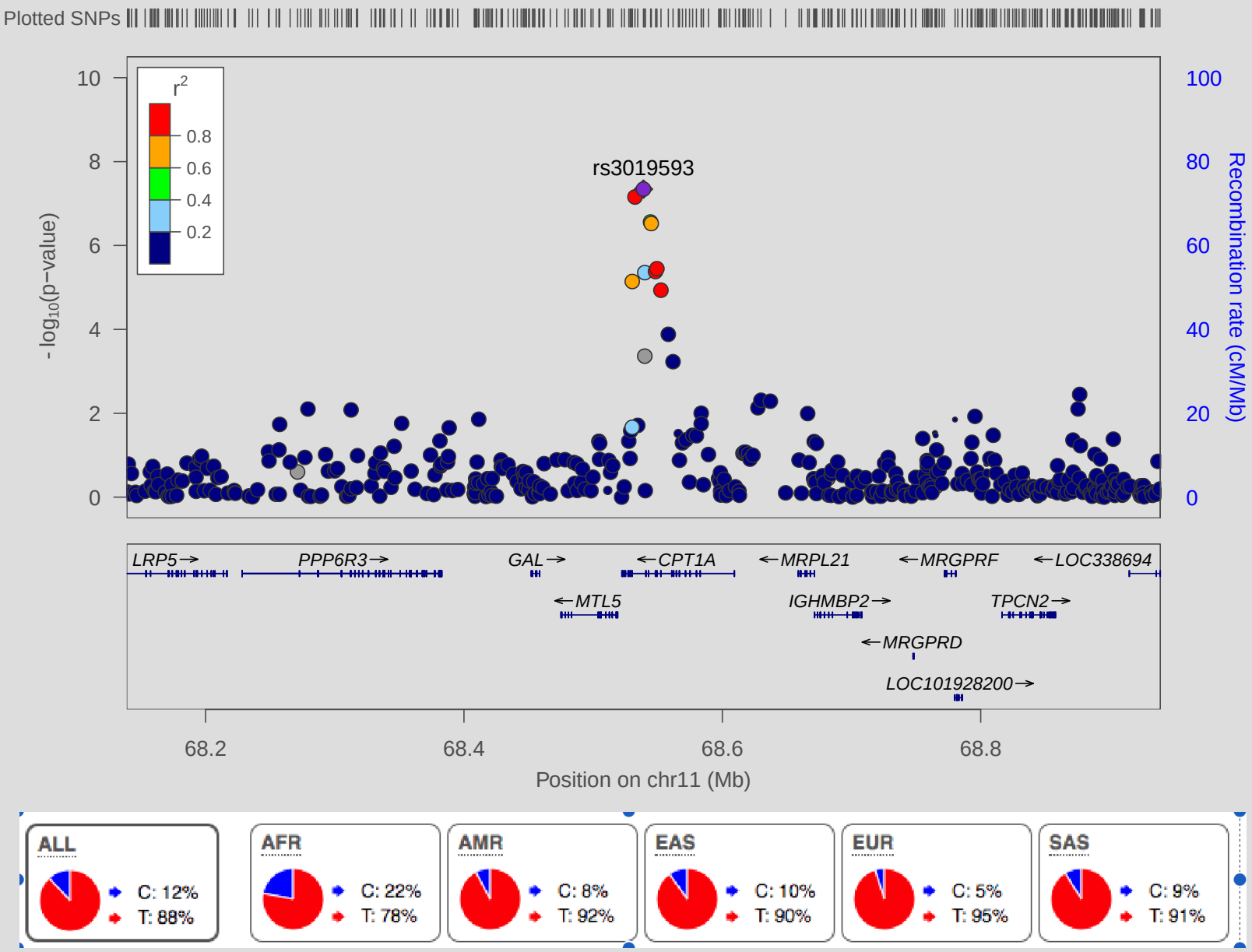
HLA LOCUS SIGNIFICANT SNPS

SNP	CHR	BP	A1	A2	MAF		OR	Genomic Region	Overlapping Gene	P-value
					Case	Control				
rs617578	6	32574603	A	G	0.2574	0.07227	3.309	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	3.988E-25
rs34855541	6	32559825	G	A	0.2457	0.066	3.352	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	1.788E-24
rs34291045	6	32560385	T	A	0.2452	0.06576	3.346	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	1.94E-24
rs35265698	6	32561334	G	C	0.2452	0.06592	3.337	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	2.467E-24
rs34350244	6	32561465	T	C	0.2439	0.0658	3.342	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	2.205E-24
rs34039593	6	32570311	G	T	0.2496	0.06917	3.289	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	2.289E-24
rs602457	6	32573562	C	T	0.2513	0.07878	2.927	Intergenic	*HLA-DRβ1* | *HLA-DQA1*	7.842E-22
rs1964995	6	32449411	C	T	0.4644	0.2557	2.164	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	5.153E-20
rs9391786	6	32448561	G	A	0.4635	0.2542	2.161	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	5.559E-20
rs9378264	6	32443451	A	G	0.4636	0.2551	2.156	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	6.592E-20
rs12195582	6	32444544	T	C	0.4636	0.2556	2.154	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	7.486E-20
rs9394099	6	32449160	T	G	0.4636	0.2556	2.154	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	7.486E-20
rs12194148	6	32444198	T	G	0.4635	0.255	2.152	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	7.813E-20
rs28895244	6	32443820	A	G	0.4636	0.2557	2.151	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	8.81E-20
rs9378212	6	32445691	T	C	0.4635	0.2554	2.154	Intergenic	*HLA-DRβ9* | *HLA-DRβ5*	8.914E-20

MANHATTAN PLOT WITHOUT CHROMOSOME 6



CHR11 LOCUSZOOM



NON-HLA SIGNIFICANT SNPS

SNP	CHR	Position	A1	A2	MAF		OR	Genomic Location	Overlapping gene	P-value
					Case	Controls				
rs3019593	11	68538939	C	T	0.2405	0.1813	1.699	intron	*CPT1A*	4.53E-08
rs3019596	11	68532388	G	A	0.2478	0.1909	1.675	intron	*CPT1A*	6.96E-08

NON-HLA FUNCTIONAL ANNOTATION

Gene	SNP	Genomic Location	Regulome	Regulatory Histone Marks				LofTool	CADD
				Promoter	Enhancer	DNAse	eQTL		
CPT1A	rs3019596	Intronic	1f	No	Yes	No	No	0.0338	2.175
CPT1A	rs3019593	Intronic	ND	No	No	Yes	No	0.0338	0.758

*ND – No data

CPT1A GENE ASSOCIATION WITH RHEUMATIC DISEASE

- *CPT1A* upregulated in RA and SLE (Abreu et al. 2018)

- *CPT1A* upregulated in RA, SLE and SSc (Hudson et al. 2017).

LIMITATIONS

- Small sample size
- Array not disease specific
- Non-coding SNPs elucidation

CONCLUSION

- This study further confirmed that chromosome 6, specifically HLA region, confers the strongest susceptibility genetic risk to seropositive RA in black South Africans.
- Two potentially novel SNP associations with RA were also identified on chromosome 11 intron region of *CPT1A*.
- In addition we have identified three non-HLA loci of interest with multiple hits on chromosome 3, 4 and 21.
- Further studies should look at functional characterisation and severity of the identified genetic variants.

ACKNOWLEDGEMENTS

Patients and controls

Mohammed Tikly

Michele Ramsay

Nimmisha Govind

Ananyo Choudhury

Audrey Thwala

Fellow Students and staff

NRF

Connective Tissue Disease Research Fund

University of Witwatersrand