Practical exercise: **Understanding how *APOL1* Renal Risk 'Genotypes' are derived**

Purpose:

a. To use raw genotype data to derive *APOL1* risk alleles (G0, G1 and G2) and genotypes (G0/G0, G1/G0, G2/G0, G1/G2, G1/G1 and G2/G2)
b. To understand allele and genotype frequencies
c. To understand genetic association and causality

Background information:

It is made up of haplotypes that involve 3 different genetic variants. Two single nucleotide variants and one indel.
Each variant is defined by a 'rs' number, the nucleotides involved (A/G/C/T) and the amino acid changes (S-serine; G-glycine; I-isoleucine ; M-methionine; N-asparagine; and Y-tyrosine).

Three loci that are each bi-allelic are involved

Locus	Alleles	Genotypes
rs73885319	A or G	A/A or A/G or G/G
rs60910145	T or G	T/T or T/G or G/G
rs71785313	ins or del	ins/ins or ins/del or del/del

Haplotypes (defined by the sequence of alleles at the three loci on a single chromosome):

Haplotype	rs73885319 A/G S342G	rs60910145 T/G I384M	rs71785313 TTATAA(ins)/del N388Y389/del
G0	A	T	ins
G1	G	G	ins
G1	G	T	ins
G1	A	G	ins
G2	A	T	del

(The other haplotypes have not been observed)

APOL1 Genotypes – as usually referred to in the literature:
G0/G0, G1/G0, G2/G0, G1/G2, G1/G1 and G2/G2

Data as generated in the laboratory (either in a single genotype experiment or an array). The H3Africa array included rs73885319 and rs60910145, but not rs71785313. Therefore if using array data, rs71785313 has to be inferred (imputed) using a statistical probabilistic approach.

Part A: Converting genotype data into *APOL1* high risk 'genotypes'

Table 1: Raw genotypes – not phased

Individual	rs73885319 A/G	rs60910145 T/G	rs71785313 TTATAA/del
ID 1	A/A	T/T	ins/ins
ID 2	A/G	T/T	ins/ins
ID 3	A/A	T/T	ins/del
ID 4	A/A	T/T	del/del
ID 5	A/A	T/G	ins/ins
ID 6	A/G	T/G	Ins/ins
ID 7	A/G	T/T	Ins/del
ID 8	A/G	T/T	Ins/ins

Table 2: Use the data from table1 to work out the Haplotypes (2 filled in as examples – complete the rest of the table)

Individual	Haplotype 1	Haplotype 2	Diplotype	APOL1 risk 'genotype'
ID 1	A-T-ins	A-T-ins	A-T-ins/ A-T-ins	G0/G0
ID 2	A-T-ins	G-T-ins	A-T-ins/ G-T-ins	G0/G1
ID 3	A-T-ins	A-T-del	A-T-ins/A-T-del	G0/G2
ID 4	A-T-del	A-T-del	A-T-del/A-T-del	G2/G2
ID 5	A-T-ins	A-G-ins	A-T-ins/A-G-ins	G0/G1
ID 6	A-T-ins or A-G-ins	G-G-ins or G-T-ins		Either G0/G1 or G1/G1
ID 7	A-T-del	G-T-ins		G2/G1
ID 8	A-T-ins	G-T-ins		G0/G1

Questions:

1. Could you resolve all the 'genotypes' – actually haplotypes?

No, the 'genotype' for ID6 could not be resolved. There are 2 possibilities. When imputation is done, there will be a probability assigned to each option and the most probable one assigned.

2. How many individuals are 'high risk' *APOL1* genotypes?

Individuals ID4, ID7 and possible ID6 have two high risk alleles (G1 or G2)

3. What are the *APOL1* G0, G1 and G2 frequencies?

If we exclude ID 6 – that leaves us with 7 individuals and therefore 14 alleles. Allele frequencies: GO – 6/14 (0.43); G1 – 4/14 (0.29); G2 – 4/14 (0.29) (rounding off results in the frequencies not adding up to 1)

Part B: Effects of APOL1 'genotypes' on phenotype

Examine the following tables and figures and answer questions

Study 1: Brandenburg JT, Govender MA, Winkler CA, Boua PR, Agongo G, Fabian J, Ramsay M. Apolipoprotein L1 High-Risk Genotypes and Albuminuria in Sub-Saharan African Populations. Clin J Am Soc Nephrol. 2022 Jun;17(6):798-808. doi: 10.2215/CJN.14321121. Epub 2022 May 16. PMID: 35577564; PMCID: PMC9269651.

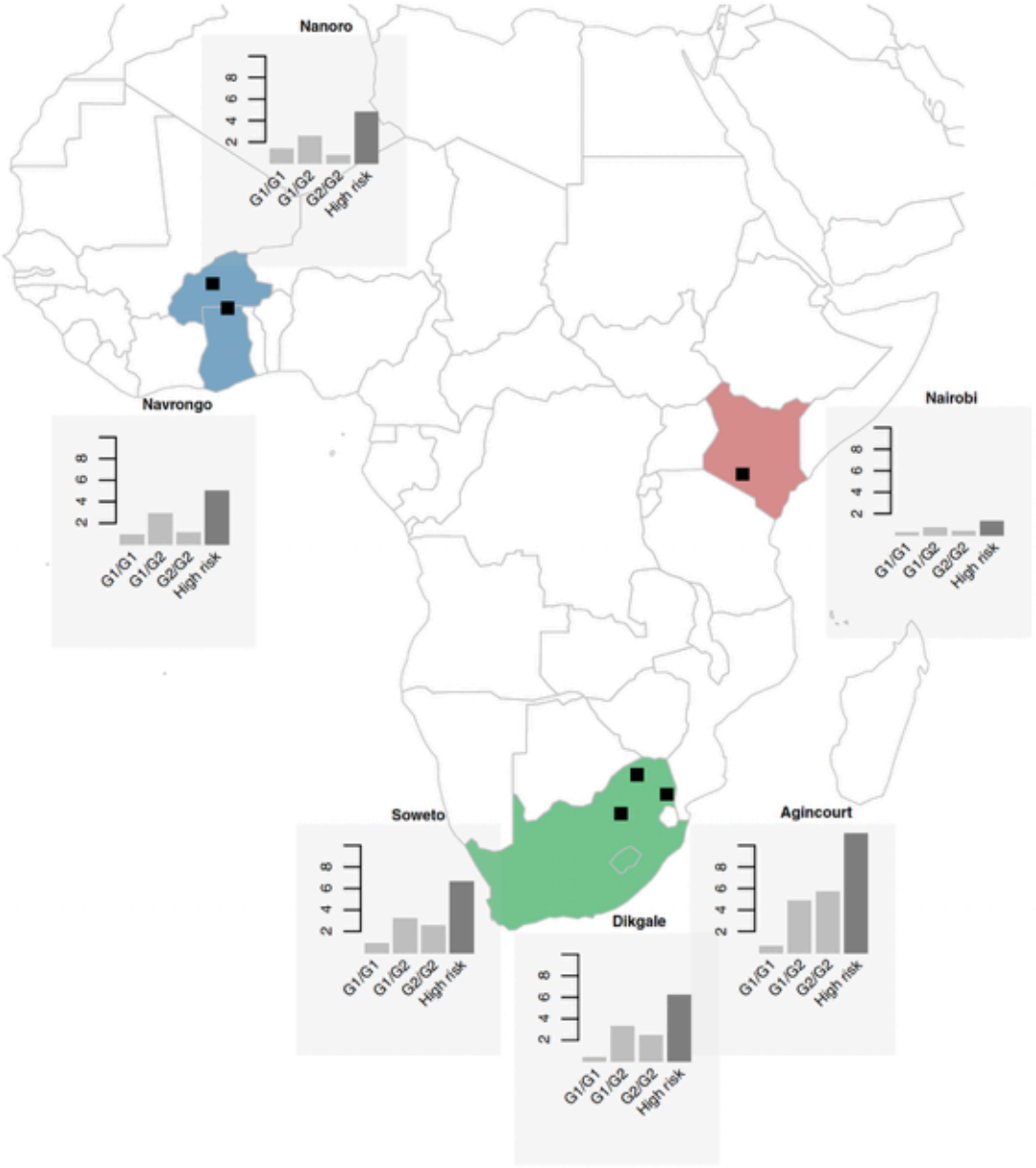


Figure 1. | Combined "high-risk" genotype frequencies are highest in West Africa or South Africa and lowest in East Africa, with distinct differences with regard to the G1/G1, G1/G2 and G2/G2 genotypes across regions. *APOL1* genotype (G1/G1, G1/G2, and G2/G2) frequencies (percentages) and combined "high-risk" genotypes (G1/G1, G1/G2, and G2/G2) for each sub-Saharan African region: Nanoro and Navrongo in West Africa; Nairobi in East Africa; and Soweto, Dikgale, and Agincourt study sites in South Africa.

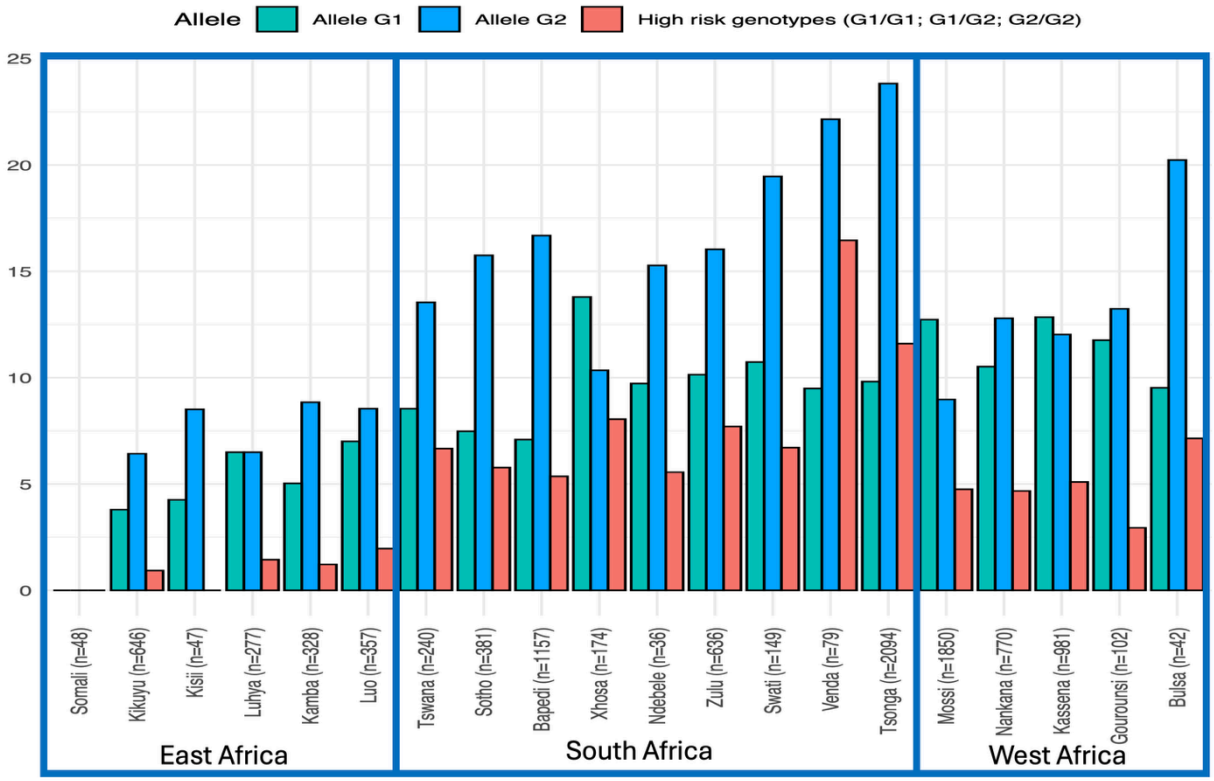

Allele
Allele G1
Allele G2
High risk genotypes (G1/G1; G1/G2; G2/G2)
25
20
15
10
5
0
Somali (n=48)
Kikuyu (n=646)
Kisii (n=47)
Luhya (n=277)
Kamba (n=328)
Luo (n=357)
East Africa
Tswana (n=240)
Sotho (n=381)
Bapedi (n=1157)
Xhosa (n=174)
Ndebele (n=36)
Zulu (n=636)
Swati (n=149)
Venda (n=79)
Tsonga (n=2094)
South Africa
Mossi (n=1850)
Nankana (n=770)
Kassena (n=981)
Gourounsi (n=102)
Bulsa (n=42)
West Africa

Table 4. Association between *APOL1* genotype and biomarkers of kidney disease stratified by region

Genotype[a]	All	East Africa	South Africa	West Africa
Low eGFR[b]				
G0/G1	1.03 (0.73 to 1.47)	1.00 (0.29 to 2.63)	1.06 (0.61 to 1.76)	1.01 (0.58 to 1.69)
G0/G2	0.76 (0.55 to 1.05)	0.77 (0.26 to 1.88)	0.82 (0.53 to 1.25)	0.74 (0.37 to 1.34)
G1/G1	0.52 (0.07 to 3.83)	NA[c]	0 (0 to 35,130.28)	0.93 (0.05 to 4.47)
G1/G2	0.81 (0.39 to 1.69)	4.0 (0.2 to 26.4)	1.13 (0.46 to 2.38)	NA[c]
G2/G2	0.89 (0.41 to 1.94)	15.09 (0.69 to 125.23)	0.99 (0.38 to 2.18)	NA[c]
Albuminuria[d]				
G0/G1	0.97 (0.77 to 1.22)	0.79 (0.38 to 1.48)	0.82 (0.58 to 1.14)	1.3 (0.9 to 1.8)
G0/G2	1.09 (0.9 to 1.31)	1.21 (0.73 to 1.95)	1.06 (0.84 to 1.35)	1.1 (0.7 to 1.6)
G1/G1	3.87 (2.16 to 6.93)[e]	NA[c]	3.67 (1.38 to 8.82)[e]	4.93 (2.15 to 10.26)[e]
G1/G2	1.24 (0.83 to 1.87)	3.24 (0.16 to 23.79)	1.20 (0.72 to 1.92)	1.28 (0.52 to 2.66)
G2/G2	1.65 (1.09 to 2.51)[e]	4.63 (0.62 to 23.40)	1.62 (1.01 to 2.51)[e]	1.09 (0.17 to 3.71)
Composite end point[f]				
G0/G1	0.97 (0.79 to 1.19)	0.84 (0.44 to 1.50)	0.83 (0.61 to 1.13)	1.22 (0.89 to 1.65)
G0/G2	0.99 (0.83 to 1.17)	1.12 (0.69 to 1.77)	0.98 (0.78 to 1.22)	0.95 (0.67 to 1.34)
G1/G1	2.73 (1.51 to 4.96)[e]	NA[c]	2.47 (0.87 to 6.18)	3.45 (1.52 to 7.13)[e]
G1/G2	1.06 (0.72 to 1.57)	2.56 (0.12 to 19.26)	1.11 (0.69 to 1.73)	0.89 (0.37 to 1.83)
G2/G2	1.40 (0.93 to 2.11)	4.32 (0.59 to 21.73)	1.41 (0.89 to 2.17)	0.74 (0.12 to 2.52)

The associations between *APOL1* genotype and biomarkers of kidney disease are presented as odds ratios and 95% confidence intervals. For biomarkers of kidney disease, repeat measures for those with low eGFR or albuminuria were not performed, thus preventing confirmation of CKD; therefore, we used "biomarkers of kidney disease" to define these measures on a single screening. Linear and logistic mixed models were adjusted for site as a random variable and age, sex, body mass index, diabetes mellitus status, hypertension status, and HIV status as fixed variables. NA, not applicable.

[a]*APOL1* genotypes G0/G1, G0/G2, G1/G1, G1/G2, and G2/G2 were compared with G0/G0 as the reference.

[b]Low eGFR is eGFR <60 ml/min per 1.73 m^2 calculated using the Chronic Kidney Disease Epidemiology Collaboration (creatinine) equation 2009 without adjusting for the African American coefficient.

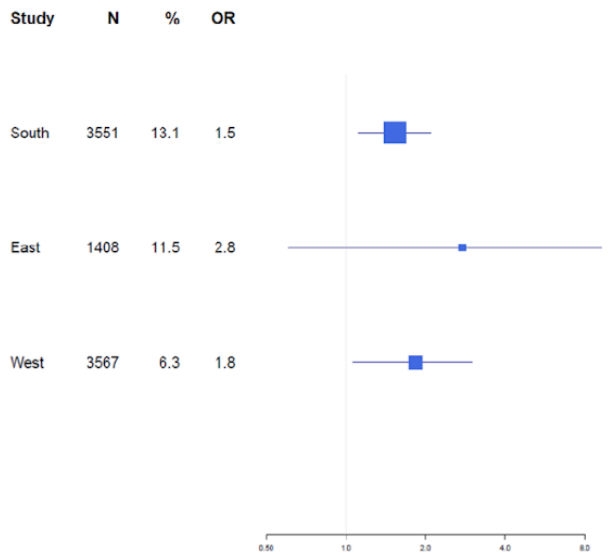
[c]Sample size was too small to perform the statistical test.

[d]Albuminuria is random spot urine albumin-creatinine ratio >3.0 mg/g.

[e]Corresponds at main section and in genotype at the total of individual with zero risk allele, one risk allele, and two risk alleles.

[f]The composite end point is low eGFR and/or albuminuria.

Supplemental Figure 1: Forest plot: association between high-risk *APOL1* genotypes (OR (95%CI)) and albuminuria by region



Regions: **East**: East Africa (Kenya); **West**: West Africa (Burkina Faso and Ghana); **South**: South Africa.
OR (odds ratio) with 95% CI (confidence interval).

Study 2: Gbadegesin RA, et al. H3Africa Kidney Disease Research Network. *APOL1* Bi- and Monoallelic Variants and Chronic Kidney Disease in West Africans. N Engl J Med. 2025 Jan 16;392(3):228-238. doi: 10.1056/NEJMoa2404211. Epub 2024 Oct 26. PMID: 39465900; PMCID: PMC11735277.

APOL1 risk variants	OR (95% CI)	
	Unadjusted	Adjusted†
All CKD cases (N=5578) vs. Controls (N=2777)		
APOL1 Risk Alleles: 2 vs. 1 and 0	1.34 (1.21–1.49)	1.25 (1.11–1.40)
G0/G1 vs G0/G0	1.16 (1.03–1.31)	1.19 (1.04–1.35)
G0/G2 vs G0/G0	1.18 (1.01–1.38)	1.19 (1.00–1.41)
G0/G1 and G0/G2 vs G0/G0	1.17 (1.05–1.30)	1.18 (1.04–1.33)
G1/G1 vs G0/G0	1.46 (1.26–1.69)	1.37 (1.16–1.61)
G1/G2 vs G0/G0	1.40 (1.18–1.65)	1.34 (1.12–1.61)
G2/G2 vs G0/G0	2.25 (1.52–3.34)	2.05 (1.35–3.13)

Associations of *APOL1* Risk Alleles and CKD among 8355 West Africans in the H3Africa Kidney Disease research network (covariates: Age, sex, BMI, MAP, HIV status, diabetes, clinical site, tobacco use and language group)

Questions:

1. What conclusions can you draw about genotype frequencies in different African populations?

The alleles and genotype frequencies vary from country to country and even between different ethnic groups within a country. This in turn could affect the ability to detect associations with kidney function traits.

2. Are the associations with the genotypes/haplotype or with the alleles?

Associations can be done either with alleles or with genotypes. Usually the associations with APOL1 are with the genotypes G0/G0, G1/G0, G2/G0, G1/G2, G1/G1 and G2/G2

3. Is the *APOL1* risk genotype causal of chronic kidney disease? Motivate you response

Individuals with the high risk genotypes have increased risk for some kidney function traits. However, alone they are not enough to cause kidney disease. When you look carefully at the data, the allele frequency in cases and controls are not vastly different and the Odds Ratios are modest. Note carefully when comparing studies whether the phenotypes are the same and what the study design was (e.g. population cross sectional study or case:control study)

4. Do G1 and G2 contribute equally to the kidney function phenotype?

No – look at the discussion in the papers and at the tables in multiple papers – especially the more recent ones.
e.g. for Albuminuria in Study 1 G1/G1 is more strongly associated with albuminuria than G2/G2 – which is also significantly associated. Study 2 has done different associations and compared specific genotype groups in different combinations. Two risk alleles vs 1 or none, and G2/G2 vs G0/G0.

Studies do not always agree, suggesting that there are many additional factors that contribute to risk – environmental and genetic.

NB: Revisit Jean-Tristan, Annette, Cassi and Walt's talks