Assessing the Genetic Burden of Familial Hypercholesterolemia in a large Middle Eastern Biobank

Geethanjali Devadoss Gandhi, Waleed Aamer, Navaneethakrishnan Krishnamoorthy, Najeeb Syed, Elbay Aliyev, Aljazi Al-Maraghi, Mohammed Kohailan, Jamil Alenbawi, Mohammed Elanbari, Qatar Genome Program Research Consortium (QGPRC), Borbala Mifsud , Younes Mokrab, Charbel Abi Khalil, and Khalid A. Fakhro *

BACKGROUND

- The genetic architecture underlying Familial Hypercholesterolemia (FH) in Middle Eastern Arabs is yet to be fully described, and approaches to assess this from population-wide biobanks are important for public health planning and personalized medicine.

METHODS

- We evaluate the pilot phase cohort (n=6,140 adults) of the Qatar Biobank (QBB) for FH using the Dutch Lipid Clinic Network (DLCN) criteria, followed by an in-depth characterization of all genetic alleles in known dominant (*LDLR*, *APOB*, and *PCSK9*) and recessive (*LDLRAP1*, *ABCG5*, *ABCG8*, and *LIPA*) FH-causing genes derived from whole-genome sequencing (WGS) of Qatar Genome Program (QGP).
- We also investigate the utility of a globally established 12-SNP polygenic risk score to predict FH individuals in this cohort with Arab ancestry.

RESULTS

- Using DLCN criteria, we identify eight ‘definite’, 41 ‘probable’ and 334 ‘possible’ FH individuals, estimating a prevalence of FH ‘(definite or probable’) in the Qatari cohort of **~1:125.**
- We identify ten previously known pathogenic single-nucleotide variants (SNVs) and 14 putatively novel SNVs, as well as one novel copy number variant in *PCSK9*.
- Further, despite the modest sample size, we identify one homozygote for a known pathogenic variant (*ABCG8*, p. Gly574Arg) associated with Sitosterolemia 2.
- Finally, calculation of polygenic risk scores found that individuals with 'definite or probable' FH have a significantly higher LDL-C SNP score than ‘unlikely’ individuals (p=0.0003), demonstrating its utility in Arab populations.

Table 1. DLCN FH diagnostic criteria as modified and used in this study.

Dutch Lipid Clinic Network (DLCN)	Points	No. of QBB participants (n=6140)
Family History		
First-degree relative with known coronary and vascular disease	1	862
Clinical history		
Patient with premature coronary artery disease	2	50
Patient with premature cerebral or peripheral vascular disease	1	6
LDL-C (mmol/liter)		
LDL-C (≥ 8.5)	8	28
LDL-C (6.5–8.4)	5	75
LDL-C (5.0–6.4)	3	264
LDL-C (4.0–4.9)	1	801
Diagnosis		
Definite FH	> 8	8
Probable FH	6 – 8	41
Possible FH	3 – 5	334
Unlikely FH	<3	5757

Figure 1. Mapping of key regions in the 3D structure of PCSK9 and LDLR.

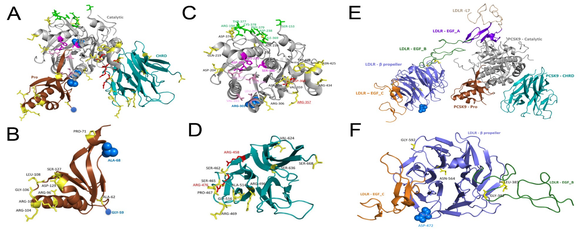


Figure 3. LDL-C SNP score for DLCN criteria.

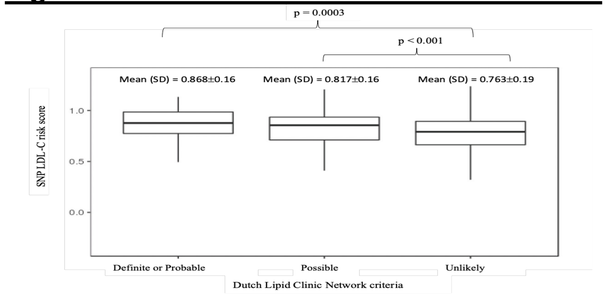
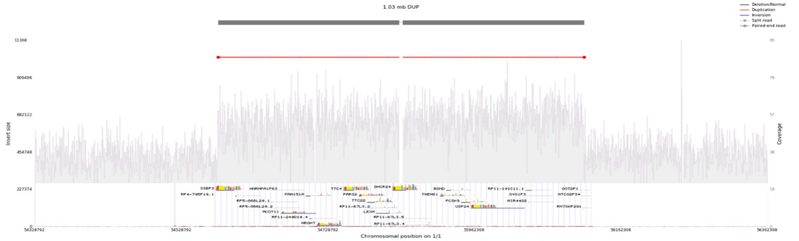


Table 2. Pathogenic variants associated with FH in the QGP.

Gene	Amino-acid change	QGP AC	Estimated clinical penetrance	OMIM Phenotype
ABCG5	p.Arg446*	27	.	Sitosterolemia 1
ABCG8	p.Arg412*	1	.	Sitosterolemia 2
ABCG8	p.Gly574Arg	6	100% (1/1)*	Sitosterolemia 2
APOB	.	2	.	Hypobetalipoproteinemia
APOB	p.Arg490Trp	2	.	Hypobetalipoproteinemia
LDLR	p.Leu385Arg	1	0% (0/1)	Familial Hypercholesterolemia
LDLR	p.Asn564Ser	1	100% (1/1)	Familial Hypercholesterolemia
LDLR	p.Asp90Gly	3	67% (2/3)	Familial Hypercholesterolemia
LDLR	.	6	83% (5/6)	Familial Hypercholesterolemia
LIPA	p.Thr288Ile	1	.	Lysosomal acid lipase deficiency

Figure 2. Structural variant analysis of loci 1:54828792-55862308 showing gene duplication in *PCSK9* gene.



CONCLUSION

- We design and implement a standardized approach to phenotyping a population biobank for FH risk followed by systematically identifying known variants and assessing putative novel variants contributing to FH burden in Qatar.
- **Our results motivate similar studies in population-level biobanks – especially those with globally under-represented ancestries – and highlight the importance of genetic screening programs for early detection and management of individuals with high FH risk in health systems**.