SUPPLEMENTAL MATERIAL

Methods

Data sources and study population

The individual level data will not be made available to other researchers without IRB approval from the VA.

Participants of multiple ethnicities were recruited from approximately 50 VA healthcare facilities across the United States.[1] Individuals consented to a blood draw and to have their DNA extracted for genomic profiling and linked to their full electronic health record within the VA. Both MVP biobank and this analysis were approved by the VA institutional review boards.

Phenotypic measures

We obtained demographic variables such as age and sex from the MVP enrollment questionnaire and participants' electronic health record (EHR) data. We secured all measures of LDL-C performed in VA laboratories using standardized assays for up to 15 years. The MVP participants had a median of 11 measurements (1st quartile of 6 and 3rd quartile of 19) of LDL-C, with 4.4% having only one measurement of LDL-C. When more than one measure was available, we extracted the maximal level of LDL-C to approximate the most likely level of untreated LDL-C.[2] We also analyzed the single outpatient LDL-C value obtained closest to the time of enrollment into MVP during 2011-2016 and statin prescriptions going back to 2002. Evidence of statin use was obtained from prescribing and pharmacy records in the EHR.

Genotypic measures and identification of FH variants

Blood samples drawn from consenting MVP participants were shipped to a central biorepository in Boston, Massachusetts, where DNA was extracted and shipped to two external genotyping centers for genotyping on an Affymetrix biobank array designed specifically for the

MVP. Genotyping was performed at two sites using the same Affymetrix best practices pipeline including several quality control (QC) procedures targeting genotyping sites and batches. Genotype calling was performed on all samples together in batches grouped by site and sample processing date. Standard Axiom genotyping quality matrix (dish quality control (DQC) >0.95, QC call rate >0.97) were comparable across the batches and sites. Furthermore, an advanced marker and sample QC procedure was used to clean and harmonize genotype calls. Probe sets were analyzed for any inconsistencies across all batches and removed. QC metrics like the final sample call rates, non-missingness, minor allele frequencies, Ratio A Allele frequencies were all determined to be consistent across sites and batches. The MVP genomics working group applied genotype calling algorithms to the data in batches using the Affymetrix Power Tools Suite (v1.18). Standard quality control pipelines were used to exclude duplicate samples, samples with more heterozygosity than expected, or discordance between sex inferred by genotyping versus self-report. We also excluded related individuals (halfway between 2nd and 3rd degree relatives or closer) as measured by the KING [3] software.

We queried the ClinVar database[4] (archive of June 2017) for "pathogenic" or "likely pathogenic" variants linked to familial hypercholesterolemia. We considered variants with conflicting interpretations of pathogenicity but at least one "pathogenic" or "likely pathogenic" assertion. These ClinVar variants were matched by unique identifiers (*i.e.,* rs ID) and chromosomal position to variants genotyped on the MVP biobank array. Additional annotation was obtained from the Human Gene Mutation Database[5] to further understand the molecular and functional effects of selected FH variants.

Statistical analysis

We used ADMIXTURE [6] analysis with the 1000 Genome phase 3 samples and known

continental ancestries to assign genotyped individuals to European (>80% European Ancestry) and admixed African Americans (>50% African Ancestry). Global and ethnicity-specific principal component analyses were performed using flashPCA software.[7] We then calculated carrier frequencies within each of two main ancestry groups.

We used linear regression models to estimate the effects of FH variants on untransformed LDL-C, adjusted for age, age^2 and sex for FH variants with a carrier count of 30 or greater (a carrier frequency of ~0.009%). For these variants, we considered associations to be significant if the p was < 0.05/the number of tested variants. FH variants with fewer than 30 carriers were combined into total burden and gene-specific scores (i.e., *LDLR, APOB* and *PCSK9*) to assess associations with maxLDL but they were not considered individually. All association analyses were performed in the R statistical environment version 3.2.5 (http://www.r-project.org/).

Phenome-wide association analysis of FH variants with clinical outcomes

Genotyped MVP participants were included in the phenome-wide association study (PheWAS) if the electronic health record (EHR) reflected two or more separate encounters in the VA Healthcare System in each of the two years prior to enrollment in the MVP. We identified 277,531 veterans and 21,209,658 prevalent ICD-9 diagnosis codes were available for analysis. ICD-9 diagnosis codes were collapsed to clinical disease groups and corresponding controls.[8] Diseases were required to have a prevalence of more than 400 cases to be included in the PheWAS analysis. Among participants with European ancestry, carriers of FH variants were compared to the other participants using logistic regression adjusting for age, sex, and ten principal components using the PheWAS R package.[9] A total of 1,171 disease phenotypes were available for analysis. Additionally, we identified individuals with clinical CHD or peripheral artery disease (PAD) using inpatient and outpatient ICD-9 and Current Procedural Terminology

codes available in EHR data from up to 15 years prior to enrollment in the MVP.

References

1. Gaziano JM, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol*. 2016;70:214-23.
2. Abul-Husn NS, et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science*. 2016;354.
3. Manichaikul A, et al. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26:2867-73.
4. Landrum MJ, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44:D862-8.
5. Stenson PD, et al. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet*. 2017;136:665-677.
6. Alexander DH, et al. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655-64.
7. Tsai TH, et al. LC-MS/MS-based serum proteomics for identification of candidate biomarkers for hepatocellular carcinoma. *Proteomics*. 2015;15:2369-81.
8. Denny JC, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol*. 2013;31:1102-10.
9. Denny JC, et al. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. 2010;26:1205-10.

Supplemental Table 1: Distribution of P/LP FH Variants (obtained from ClinVar database) among multi-ethnic MVP participants.

probeset_id	rs	AlleleID	Name	Gene	Clinical significance Last reviewed	GRCh37Chr	GRCh37Loc	GRCh38Ch	GRCh38Loc	VariationID	# of carriers
AX-86662922	rs564427867	281418	NM_174936.3(PCSK9):c.94G>A (p.Glu32Lys)	PCSK9	Conflicting interpretations of pathogenicity(Last reviewed: Jun 14, 2016)	1	55505604	1	55039931	297692	8
AX-83587852	rs143275858	354066	NM_174936.3(PCSK9):c.1251C>A (p.H417Q)	PCSK9	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	1	55523779	1	55058106	369875	344
AX-83302687	rs141502002	196756	NM_174936.3(PCSK9):c.1405C>T (p.Arg469Trp)	PCSK9	Conflicting interpretations of pathogenicity(Last reviewed: Mar 28, 2016)	1	55524222	1	55058549	201128	1147
AX-86659670	rs374603772	196758	NM_174936.3(PCSK9):c.1486C>T (p.Arg496Trp)	PCSK9	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	1	55524303	1	55058630	201129	14
AX-83159740	rs12713559	32936	NM_000384.2(APOB):c.10672C>T (p.Arg3558Cys)	APOB	Conflicting interpretations of pathogenicity(Last reviewed: Dec 8, 2016)	2	21229068	2	21006196	17897	366
AX-83321223	rs5742904	32929	NM_000384.2(APOB):c.10580G>A (p.Arg3527Gln)	APOB	Conflicting interpretations of pathogenicity(Last reviewed: Jan 10, 2017)	2	21229160	2	21006288	17890	256
AX-86618968	rs144467873	48709	NM_000384.2(APOB):c.10579C>T (p.Arg3527Trp)	APOB	Pathogenic/Likely pathogenic(Last reviewed: Dec 8, 2016)	2	21229161	2	21006289	40223	19
AX-98075214	rs267607213	18780	NM_000527.4(LDLR):c.131G>A (p.Trp44Ter)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11210962	19	11100286	3741	1
AX-98033604	rs121908041	18781	NM_000527.4(LDLR):c.137G>C (p.Cys46Ser)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Mar 25, 2016)	19	11210968	19	11100292	3742	1
AX-86597121	rs376207800	362667	NM_000527.4(LDLR):c.185C>G (p.Thr62Arg)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11211016	19	11100340	375775	84
AX-86667966	rs370860696	171197	NM_000527.4(LDLR):c.232C>T (p.Arg78Cys)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Mar 25, 2016)	19	11213381	19	11102705	161289	15
AX-86641078	rs121908025	18724	NM_000527.4(LDLR):c.259T>G (p.Trp87Gly)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11213408	19	11102732	3685	8
AX-86712213	rs749038326	245449	NM_000527.4(LDLR):c.268G>T (p.Asp90Tyr)	LDLR	Likely pathogenic(Last reviewed: Nov 5, 2016)	19	11213417	19	11102741	251106	7
AX-86574463	rs750474121	245462	NM_000527.4(LDLR):c.292G>C (p.Gly98Arg)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Mar 25, 2016)	19	11213441	19	11102765	251119	19
AX-86737971	rs377271627	171198	NM_000527.4(LDLR):c.296C>G (p.Ser99Ter)	LDLR	Pathogenic(Last reviewed: Dec 9, 2016)	19	11213445	19	11102769	161269	7
AX-82987078	rs144172724	171199	NM_000527.4(LDLR):c.301G>A (p.Glu101Lys)	LDLR	Pathogenic/Likely pathogenic, not provided(Last reviewed: Dec 16, 2016)	19	11213450	19	11102774	161266	12
AX-86732996	rs774723292	245500	NM_000527.4(LDLR):c.343C>T (p.Arg115Cys)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11215925	19	11105249	251162	12
AX-82996623	rs201102461	227401	NM_000527.4(LDLR):c.344G>A (p.Arg115His)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11215926	19	11105250	225402	1
AX-98033792	rs193922571	245515	NM_000527.4(LDLR):c.362G>T (p.Cys121Phe)	LDLR	Likely pathogenic(Last reviewed: Mar 25, 2016)	19	11215944	19	11105268	251177	4
AX-98067814	rs121908044	18786	NM_000527.4(LDLR):c.621C>T (p.Gly207=)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11216203	19	11105527	3747	3
AX-86702908	rs373822756	181233	NM_000527.4(LDLR):c.662A>G (p.Asp221Gly)	LDLR	Pathogenic/Likely pathogenic, not provided(Last reviewed: Jan 23, 2017)	19	11216244	19	11105568	183092	1
AX-83535358	rs121908035	245738	NM_000527.4(LDLR):c.693C>G (p.Cys231Trp)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11216275	19	11105599	251400	8
AX-86663356	rs768563000	245758	NM_000527.4(LDLR):c.718G>T (p.Glu240Ter)	LDLR	Pathogenic(Last reviewed: Nov 26, 2016)	19	11217264	19	11106588	251422	34
AX-83503525	rs150673992	171200	NM_000527.4(LDLR):c.757C>T (p.Arg253Trp)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Aug 31, 2016)	19	11217303	19	11106627	161261	235
AX-86638837	rs121908040	362692	NM_000527.4(LDLR):c.782G>A (p.Cys261Tyr)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11217328	19	11106652	375800	4
AX-83464547	rs139043155	171201	NM_000527.4(LDLR):c.798T>A (p.Asp266Glu)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Dec 16, 2016)	19	11217344	19	11106668	161287	5
AX-86721456	rs751122998	354088	NM_000527.4(LDLR):c.820delA (p.T274Hfs*95)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Oct 9, 2016)	19	11218070	19	11107394	369863	4
AX-86619878	rs368657165	171204	NM_000527.4(LDLR):c.862G>A (p.Glu288Lys)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11218112	19	11107436	161268	20
AX-82929130	rs151207122	171205	NM_000527.4(LDLR):c.907C>T (p.Arg303Trp)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Mar 25, 2016)	19	11218157	19	11107481	161281	49
AX-83137728	rs121908030	245850	NM_000527.4(LDLR):c.910G>T (p.Asp304Tyr)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11218160	19	11107484	251517	30
AX-86707238	rs373869746	171206	NM_000527.4(LDLR):c.967G>A (p.Gly323Ser)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11221354	19	11110678	161282	5
AX-54781084	rs72658860	171207	NM_000527.4(LDLR):c.970G>A (p.Gly324Ser)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Jan 24, 2017)	19	11221357	19	11110681	161263	1629
AX-83216162	rs139361635	245927	NM_000527.4(LDLR):c.1024G>T (p.Asp342Tyr)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Nov 5, 2016)	19	11221411	19	11110735	251603	798
AX-98067994	rs193922566	45113	NM_000527.4(LDLR):c.1055G>A (p.Cys352Tyr)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 12, 2016)	19	11221442	19	11110766	36450	6
AX-86626676	rs552422789	230981	NM_000527.4(LDLR):c.1217G>A (p.Arg406Gln)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11223984	19	11113308	228798	9
AX-86622889	rs137943601	45116	NM_000527.4(LDLR):c.1222G>A (p.Glu408Lys)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11223989	19	11113313	36453	3
AX-86631429	rs368562025	171212	NM_000527.4(LDLR):c.1238C>T (p.Thr413Met)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Jan 3, 2017)	19	11224005	19	11113329	161276	9
AX-86723634	rs748554592	246061	NM_000527.4(LDLR):c.1241T>G (p.Leu414Arg)	LDLR	Likely pathogenic(Last reviewed: Mar 25, 2016)	19	11224008	19	11113332	251747	1
AX-86703204	rs773658037	246066	NM_000527.4(LDLR):c.1247G>A (p.Arg416Gln)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 13, 2016)	19	11224014	19	11113338	251752	14
AX-98067995	rs193922567	354095	NM_000527.4(LDLR):c.1358+2T>C	LDLR	Likely pathogenic	19	11224127	19	11113451	369862	3
AX-86593083	rs370777955	171215	NM_000527.4(LDLR):c.1467C>G (p.Tyr489Ter)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Mar 25, 2016)	19	11224319	19	11113643	161270	4
AX-86670684	rs730882106	181262	NM_000527.4(LDLR):c.1576C>T (p.Pro526Ser)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Dec 16, 2016)	19	11224428	19	11113752	183120	8
AX-83471866	rs137929307	171217	NM_000527.4(LDLR):c.1775G>A (p.Gly592Glu)	LDLR	Pathogenic/Likely pathogenic, not provided(Last reviewed: Dec 16, 2016)	19	11227604	19	11116928	161271	27
AX-86723961	rs373371572	171218	NM_000527.4(LDLR):c.1783C>T (p.Arg595Trp)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Feb 10, 2017)	19	11227612	19	11116936	161290	8
AX-86720647	rs201102492	246329	NM_000527.4(LDLR):c.1784G>T (p.Arg595Leu)	LDLR	Likely pathogenic(Last reviewed: Mar 25, 2016)	19	11227613	19	11116937	252029	9
AX-86634082	rs747134711	246378	NM_000527.4(LDLR):c.1855T>C (p.Phe619Leu)	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 16, 2016)	19	11230777	19	11120101	252083	2
AX-83048349	rs139791325	181269	NM_000527.4(LDLR):c.1876G>A (p.Glu626Lys)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Feb 3, 2017)	19	11230798	19	11120122	183127	167
AX-98033790	rs193922569	45121	NM_000527.4(LDLR):c.1978C>T (p.Gln660Ter)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11230900	19	11120224	36458	4
AX-86618749	rs150021927	246458	NM_000527.4(LDLR):c.1999T>C (p.Cys667Arg)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11231057	19	11120381	252163	3
AX-98067698	rs28942083	18728	NM_000527.4(LDLR):c.2000G>A (p.Cys667Tyr)	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11231058	19	11120382	3689	8
AX-86581791	rs774730452	246486	NM_000527.4(LDLR):c.2050G>A (p.Ala684Thr)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11231108	19	11120432	252192	0
AX-86738370	rs751228587	246496	NM_000527.4(LDLR):c.2060dupT (p.Asn688Glnfs)	LDLR	Pathogenic(Last reviewed: Mar 25, 2016)	19	11231118	19	11120442	252201	1
AX-83379229	rs201573863	246513	NM_000527.4(LDLR):c.2096C>T (p.Pro699Leu)	LDLR	Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11231154	19	11120478	252219	51
AX-86688504	rs368838866	181272	NM_000527.4(LDLR):c.2101G>A (p.Gly701Ser)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Nov 5, 2016)	19	11231159	19	11120483	183130	25
AX-83408984	rs145787161	18783	NM_000527.4(LDLR):c.2140+1G>A	LDLR	Pathogenic/Likely pathogenic(Last reviewed: Dec 16, 2016)	19	11231199	19	11120523	3744	7
AX-86582518	rs72658867	45123	NM_000527.4(LDLR):c.2140+5G>A	LDLR	Conflicting interpretations of pathogenicity(Last reviewed: Dec 25, 2016)	19	11231203	19	11120527	36460	6501
AX-83148620	rs137853964	45125	NM_000527.4(LDLR):c.2479G>A (p.Val827Ile)	LDLR	Conflicting interpretations of pathogenicity, not provided(Last reviewed: Jan 5, 2017)	19	11240278	19	11129602	36462	328
AX-86643801	rs377437226	246635	NM_000527.4(LDLR):c.2546C>A (p.Ser849Ter)	LDLR	Pathogenic(Last reviewed: Mar 25, 2016)	19	11240345	19	11129669	252350	38