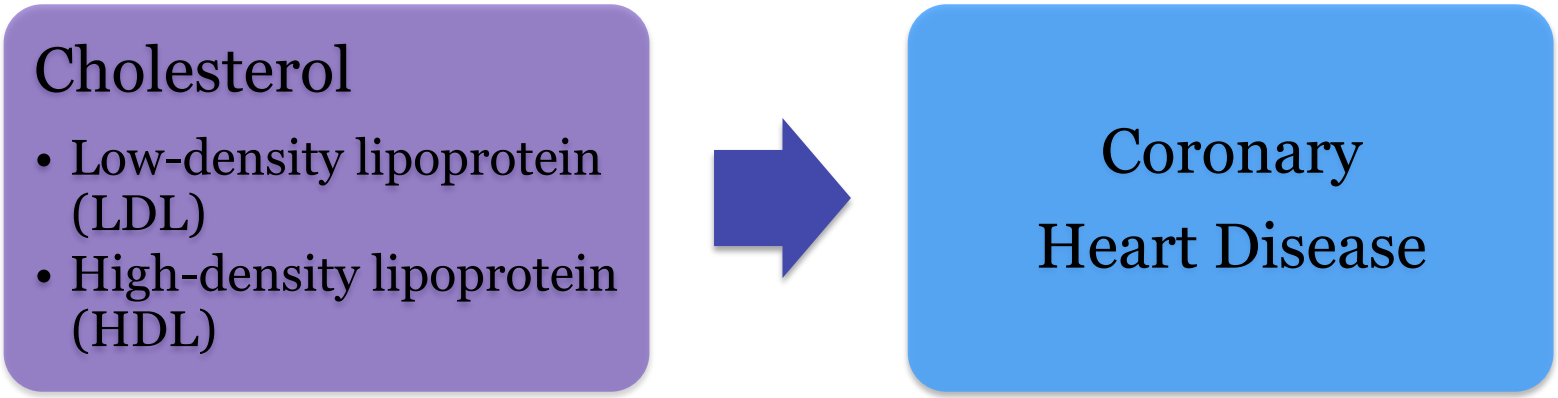
Exome Sequencing to Discover Novel Causes of Mendelian Dyslipidemias

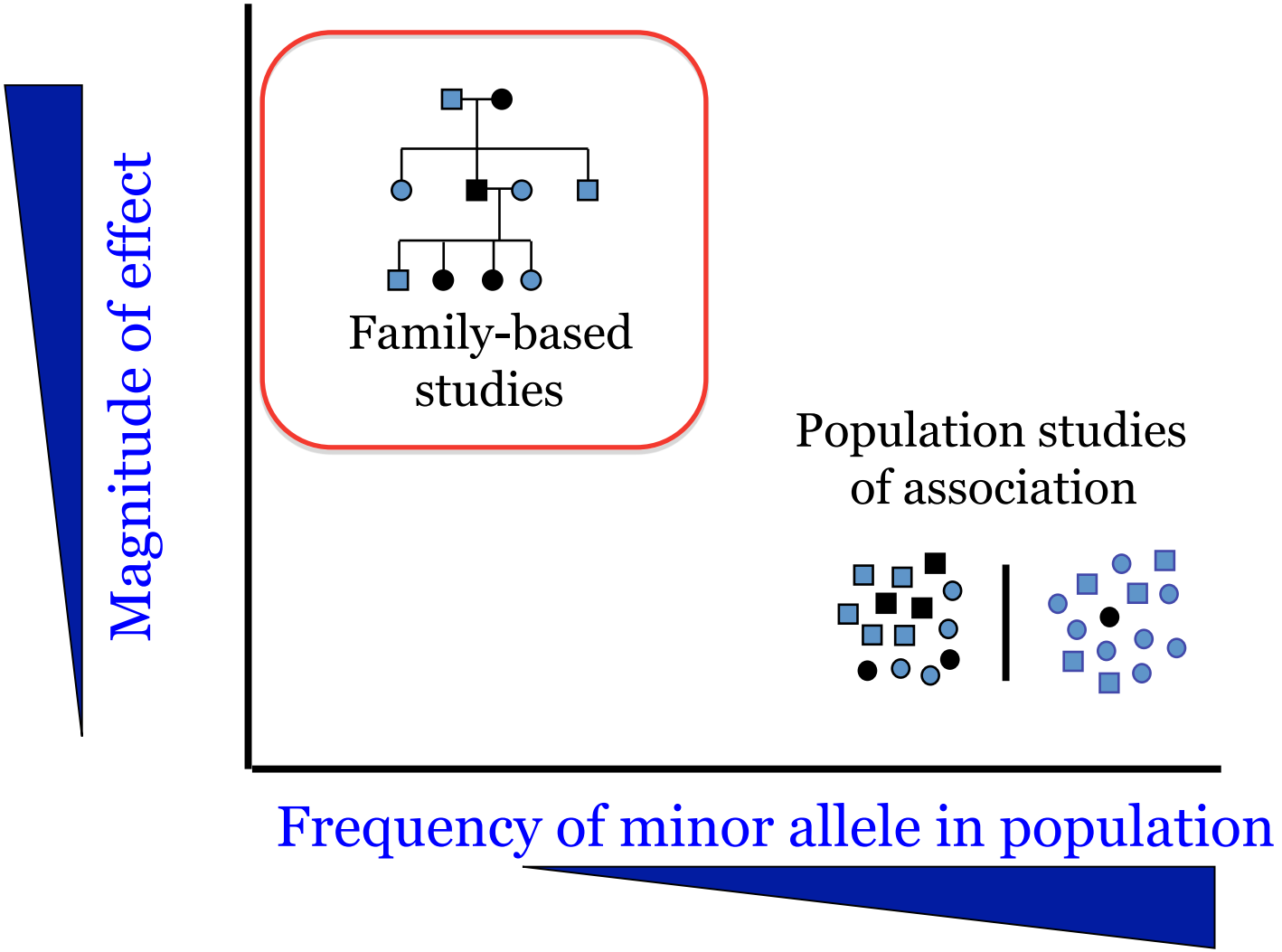
Alessa Moscoso
Harvard University
The Broad Institute of Harvard and MIT
Massachusetts General Hospital
In collaboration with Sekar Kathiresan and
Nathan Stitziel

Coronary heart disease is the leading cause of death in the US



Lipid levels are highly heritable

Mapping strategies to discover genes for lipid levels



Known genes for Mendelian lipid disorders

Gene	Locus	GWAS SNP	Disorder and lipid phenotype
ABCA1	9q31.1	rs1883025	Tangier disease: low HDL
ABCG5	2p21	rs4299376	Sitosterolemia: high LDL
ABCG8	2p21	rs4299376	Sitosterolemia: high LDL
APOA1	11q23-q24	rs964184	ApoA-I deficiency: low HDL
APOA5	11q23	rs964184	ApoA-V deficiency: high VLDL and chylomicrons
APOB	2p24	rs515135	Familial hypobetalipoproteinemia: low LDL
			Familial defective ApoB-100: high LDL
APOC2	19q13	rs4420638	Familial ApoC-II deficiency: high chylomicrons
APOE	19q13	rs4420638	Familial dysbetalipoproteinemia: high VLDL remnants and chylomicrons
CETP	16q13	rs173539	Cholesteryl ester transfer protein deficiency: high HDL
LCAT	16q22	rs2271293	Lecithin-cholesterol acyltransferase deficiency (fish-eye disease): low HDL
LDLR	19p13	rs6511720	Familial hypercholesterolemia: high LDL
LDLRAP1	1p36-p35	rs12027135	Autosomal recessive hypercholesterolemia: high LDL
LIPC	15q22	rs10468017	Familial hepatic lipase deficiency: high VLDL remnants
LPL	8p21	rs12678919	Lipoprotein lipase deficiency: high chylomicrons
MTTP	4q24	N/A	Abetalipoproteinemia: low LDL
PCSK9	1p32	rs11206510	Autosomal-dominant hypercholesterolemia: high LDL
			PCSK9 deficiency: low LDL

Known genes for Mendelian lipid disorders

1	Locus	GWAS SNP	Disorder and lipid phenotype
ABCA1	9q31.1	rs1883025	Tangier disease: low HDL
ABCG5	2p21	rs4299376	Sitosterolemia: high LDL
ABCG8	2p21	rs4299376	Sitosterolemia: high LDL
APOA1	11q23-q24	rs964184	ApoA-I deficiency: low HDL
APOA5	11q23	rs964184	ApoA-V deficiency: high VLDL and chylomicrons
APOB	2p24	rs515135	Familial hypobetalipoproteinemia: low LDL
			Familial defective ApoB-100: high LDL
APOC2	19q13	rs4420638	Familial ApoC-II deficiency: high chylomicrons
APOE	19q13	rs4420638	Familial dysbetalipoproteinemia: high VLDL remnants and chylomicrons
CETP	16q13	rs3764261	Cholesteryl ester transfer protein deficiency: high HDL
LCAT	16q22	rs2271293	Lecithin-cholesterol acyltransferase deficiency (fish-eye disease): low HDL
LDLR	19p13	rs6511720	Familial hypercholesterolemia: high LDL
LDLRAP1	1p36-p35	rs12027135	Autosomal recessive hypercholesterolemia: high LDL
LIPC	15q22	rs10468017	Familial hepatic lipase deficiency: high VLDL remnants
LPL	8p21	rs12678919	Lipoprotein lipase deficiency: high chylomicrons
MTTP	4q24	N/A	Abetalipoproteinemia: low LDL
PCSK9	1p32	rs11206510	Autosomal-dominant hypercholesterolemia: high LDL
			PCSK9 deficiency: low LDL

However, there are families where these genes do not explain the lipid disorder

Study population

Studied 6 Dutch families with dyslipidemias

- 3 families with high LDL (Hypercholesterolemia)
- 2 families with low HDL (Hypo-alphalipoproteinemia)
- 1 family with high HDL (Hyper-alphalipoproteinemia)

Hypothesis

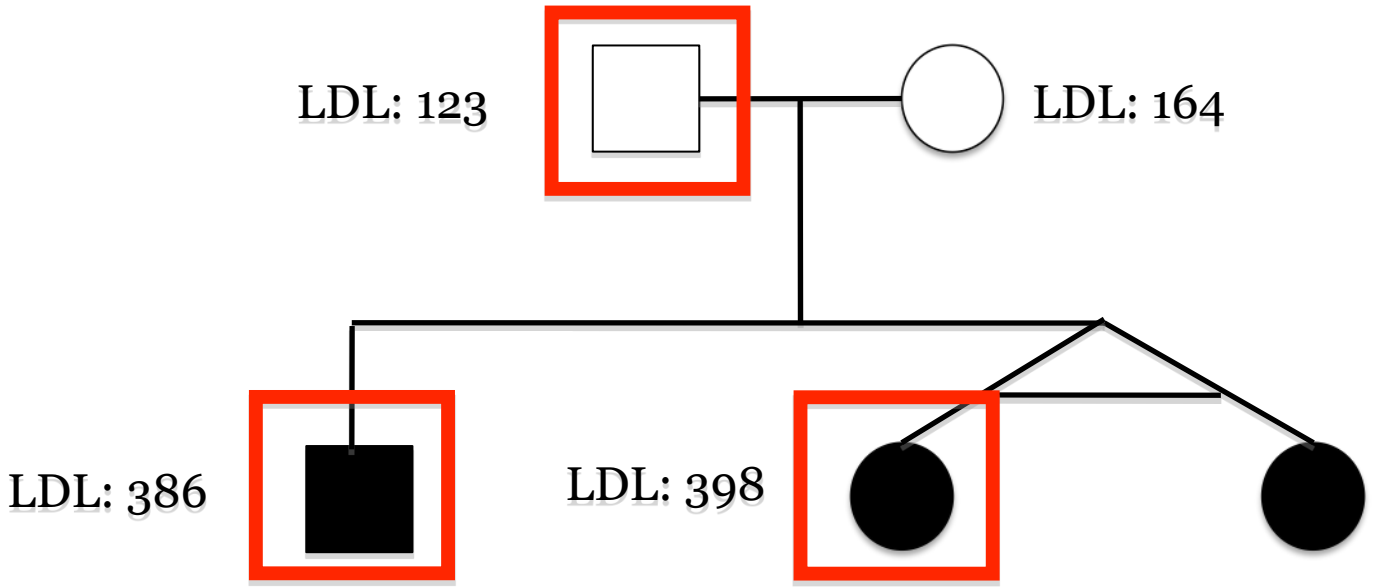
Novel genes are responsible for the Mendelian segregation of dyslipidemia in these families

How do we find the causal variant for each family among all genes in the genome?

Approach: exome sequencing

- 'Exome' refers to all of the protein-coding regions in the genome
- ~18,000 genes, ~180,000 exons
- Covers ~30,000,000 bases out of the 3 billion in the human genome
- Now possible to select and sequence all of the 30,000,000 bases from exons

FH4-0139 Autosomal recessive



All offspring are affected with high LDL

LDL level: (mg/dL)

Sequencing results

Sequencing Metrics	Per sample
Total bases	8.9×10^9
Target coverage	123 times
% bases > 20x covered	85%
Total SNPs	18,964
Heterozygous to homozygous ratio	1.6
Transition to transversion ratio	2.4

Variants found in family FH4-0139

High-quality variants: 53,479

How do we go from 54,000 variants to only 1 variant?

3 assumptions

1. Complete penetrance
2. Recessive model
3. Functional mutations

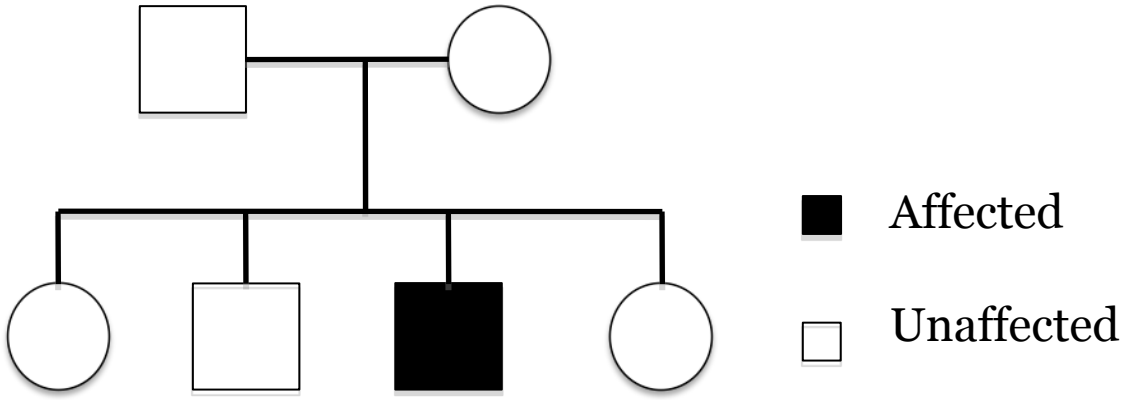
1. Complete penetrance

	Sample High LDL	Control Low LDL
SNP 1	A G	A A
SNP 2	C G	C G

Controls

- Approximately 250 control exomes for LDL family
- Final filter of 2500 individuals sequenced for various phenotypes in ESP database

2. Recessive model

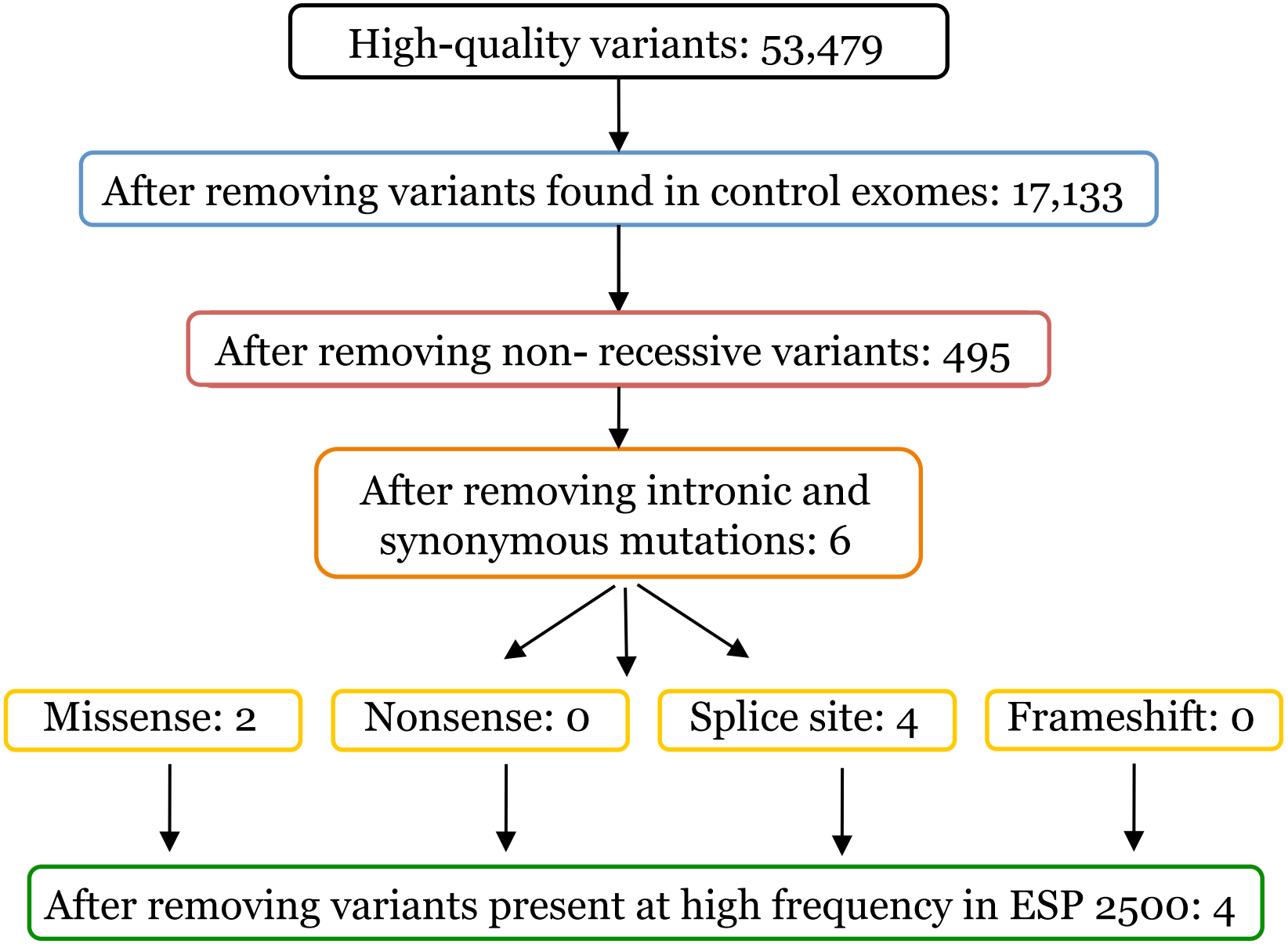


-Parents required to be heterozygous
-Affecteds required to be homozygous

3. Functional mutations

- Nonsense
- Frameshift
- Splice
- Missense

Single variants found in family FH4-0139



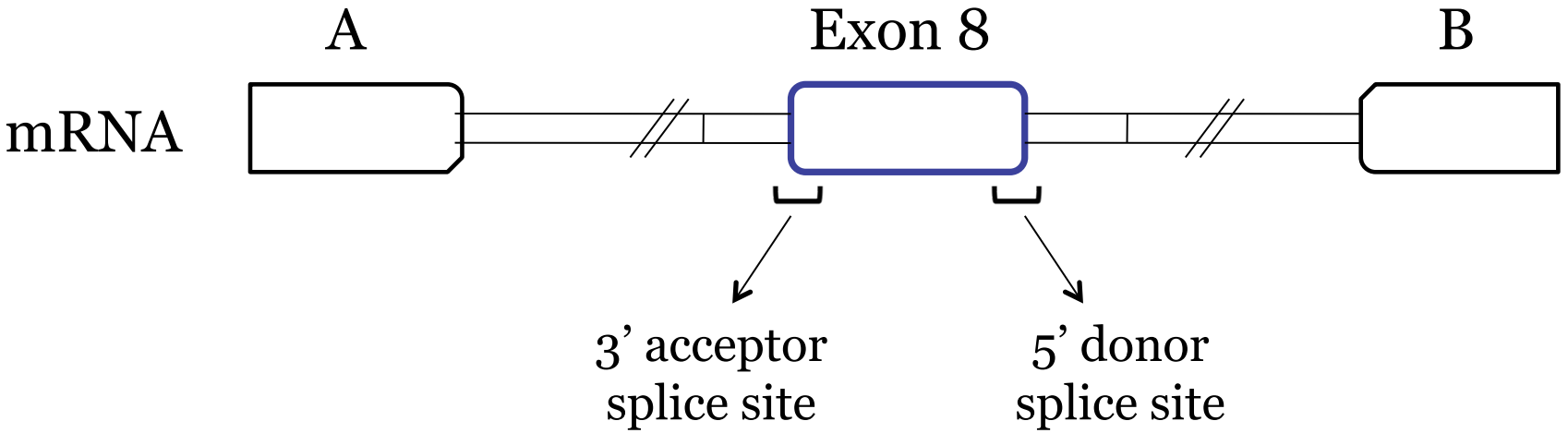
Variant in family FH4-0139 found in LIPA gene

CHR	POSITION	GENE NAME	CLASS	MINOR ALLELE FREQUENCY
10	90982268	LIPA	splice	0.022%, (1/4550) in ESP

LIPA is involved with lipid catabolism

- encodes enzyme lysosomal acid lipase
- mutations in *LIPA* cause cholesteryl ester storage disease (CESD)

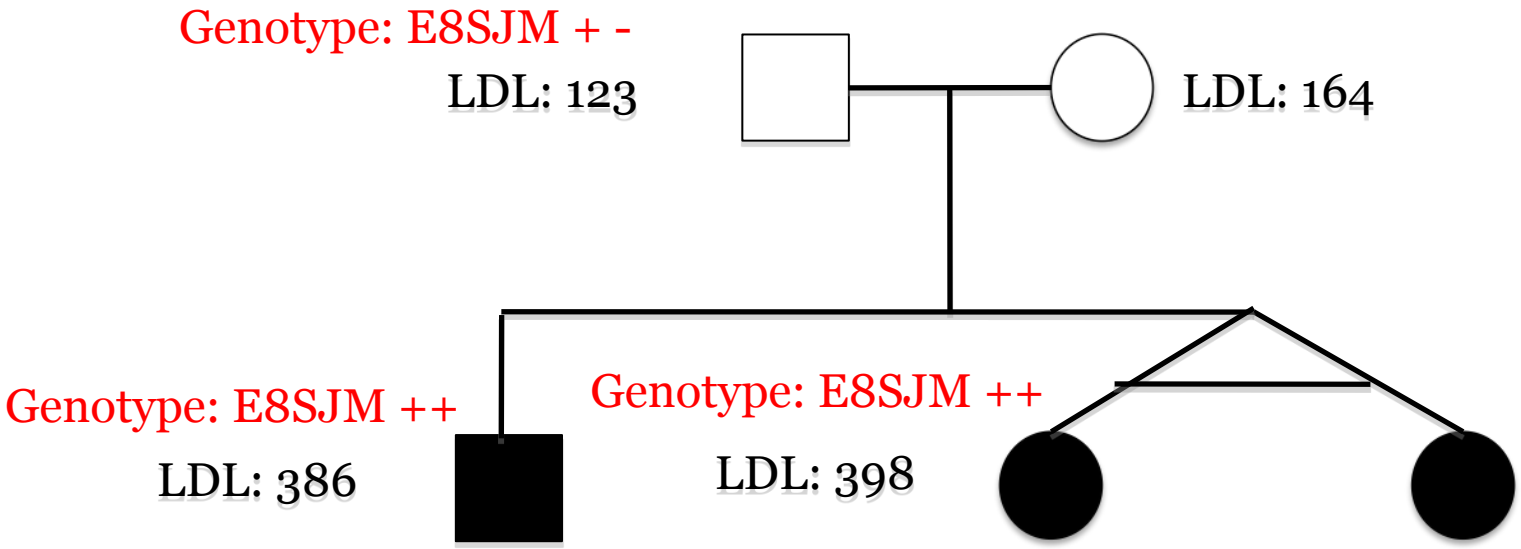
Splice Site Exon 8 Junction Mutation in *LIPA* gene



Wildtype	ag/T	CAG/gt
E8SJM	ag/T	CA<u>A</u>/gt

Mutation results in an in-frame deletion of 24 amino acids from the enzyme

FH4-0139 Autosomal Recessive



Affecteds are homozygous for E8SJM

LDL level: (mg/dL)

Conclusions

- Exome sequencing is an effective technique to identify the causal variant for Mendelian syndromes

- LIPA E8SJM is the likely cause of the hypercholesterolemia seen in family FH4-013

- Additional phenotyping (e.g., LIPA enzyme activity) planned for affected family members

Acknowledgements

- Dr. Sekar Kathiresan
- Dr. Nathan Stitziel
- Dr. Kastelein
- Summer Research Program in Genomics
- The Broad Institute
- NHLBI ESP
- Dr. Gina Peloso
- Sara Pullit
- Eboney Smith
- Lori Thomae
- Bruce Birren
- Nicole Edmonds