#### MIXED-MODEL ANALYSIS OF COMMON VARIATION REVEALS PATHWAYS EXPLAINING VARIANCE IN AMD RISK

JAKE HALL CASE WESTERN RESERVE UNIVERSITY - BUSH LAB IGES 2014

QUESTION

**M**ETHODS

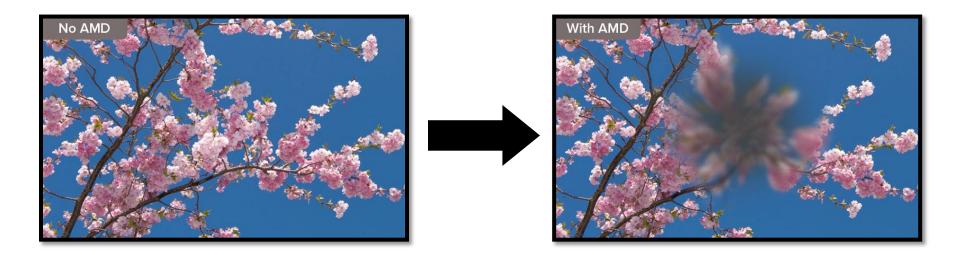
RESULTS

CONCLUSIONS



#### Age-related Macular Degeneration

- Progressive, neurodegenerative disease
- Loss of central vision
- A leading cause of blindness
  - **30+ million** affected worldwide



QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

# **GENETICS OF AMD**

- Heritability 45 70%\*
- Most strongly associated genes:
  - **CFH** [Complement Factor H]
  - **C2** [Complement Component 2]
  - **C3** [Complement Component 3]
  - **CFB** [Complement Factor B]
  - **ARMS2/HTRA1** [Age-Related Maculopathy Susceptibility 2 / High-Temperature Requirement A Serine Peptidase 1]
- 10 30% of heritability explained by 19 known risk SNPs\*

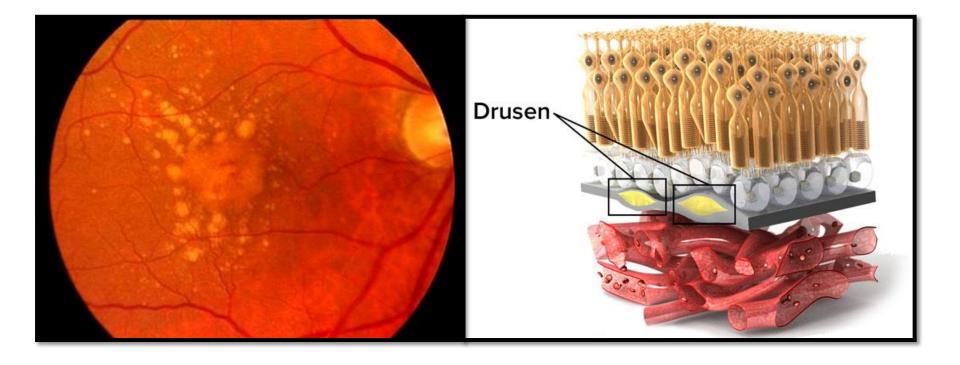
QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

## SIGNS OF AMD



#### AMD Pathogenesis...?

QUESTION

**M**ETHODS

RESULTS

**CONCLUSIONS** 

### **AMD PATHOGENESIS IN LITERATURE**



QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

## **AMD PATHOGENESIS IN LITERATURE**

8 POTENTIAL MECHANISMS				
Angiogenesis	Inflammation			
Antioxidants	Nicotine/Smoking			
Apoptosis	<b>Oxidative</b> Phosphorylation			
<b>Complement Activation</b>	Tricarboxylic Acid Cycle			

QUESTION

**M**ETHODS

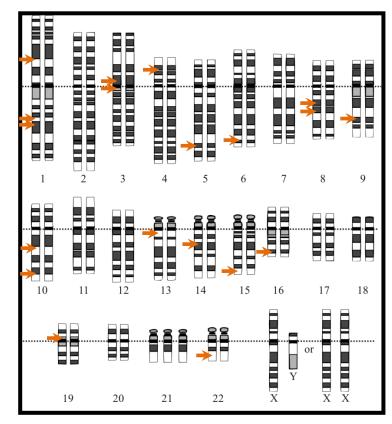
RESULTS

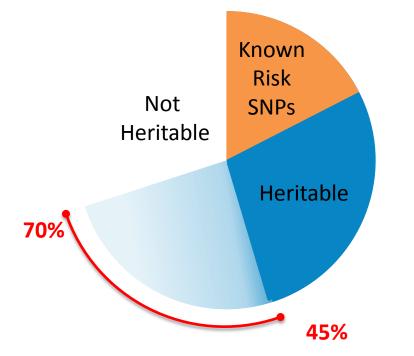
CONCLUSIONS

## MOTIVATION



8 POTENTIAL MECHANISMS				
Inflammation				
Nicotine/Smoking				
<b>Oxidative</b> Phosphorylation				
Tricarboxylic Acid Cycle				





INTRODUCTION QUESTION

METHODS

RESULTS

**C**ONCLUSIONS

#### QUESTION

## How much disease risk is explained by SNPs in potentially AMD-related pathways?

QUESTION

**M**ETHODS

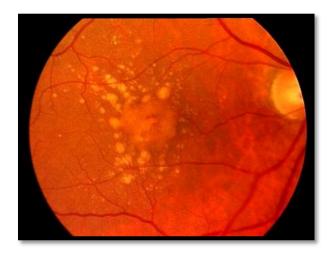
RESULTS

CONCLUSIONS

## DATASET

#### **Case/Control Design**

- Subjects ascertained in clinics at:
  - Vanderbilt University
  - o Duke University
  - University of Miami Health System
- All European descent
- Fundus photography used to confirm case status
- Primary genotyping: Affymetrix 6.0
- Secondary genotyping: Sequenom & TaqMan



INTRODUCTION QUE

QUESTION

**M**ETHODS

RESULTS

# **QC STEPS**

#### **SNPs Excluded**

- -Non-autosomal
- -Genotyping efficiency < 95%
- -Sample efficiency < 90%
- Minor allele frequency < 2%</p>
- -HWE p-value < 1  $\times$  10<sup>-6</sup>

#### **Covariates Required**

- -Age (Years)
- -Sex (M/F)

#### Post-QC

#### Total SNPs: 659,183

Affymetrix 6.0: **659,108** Custom Sequenom: **71** Custom TaqMan: **4** 

Cases: **1,145** Controls: **668** 

QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

8 POTENTIAL MECHANISMS				
Angiogenesis	Inflammation			
Antioxidants	Nicotine/Smoking			
Apoptosis	<b>Oxidative</b> Phosphorylation			
<b>Complement Activation</b>	Tricarboxylic Acid Cycle			

INTRODUCTION QUESTION

**METHODS** 

RESULTS

CONCLUSIONS

## **DEFINING PATHWAYS**

#### The Gene Ontology [GO]

AmiGO browser or download database

Searched for GO Term closes to our mechanism of interest

I GO:0008150 biological_process			
GO:0050896 response to stimulus			
GO:0042221 response to chemical			
GO:0009719 response to endogenous s	timulus		
<ul> <li>GO:1901698 response to nitrogen comp</li> <li>GO:0010033 response to organic substance</li> </ul>		GO ID	# Genes
GO:0010243 response to organonitro		GO:0001525	379
GO:0043279 response to alkaloid	Antioxidant Activity	GO:0016209	69
GO:0014070 response to organic c GO:0035094 response to nico	Δηροητοτις δισηριήσ	GO:0097190	1,635
GO:0035095 behavioral response	<b>Complement Activation</b>	GO:0006956	184
GO:0071316 cellular response	Inflammatory Response	GO:0006954	534
	Response to Nicotine	GO:0035094	31
	<b>Oxidative Phosphorylation</b>	GO:0006119	78
	Tricarboxylic Acid Cycle	GO:0006099	33

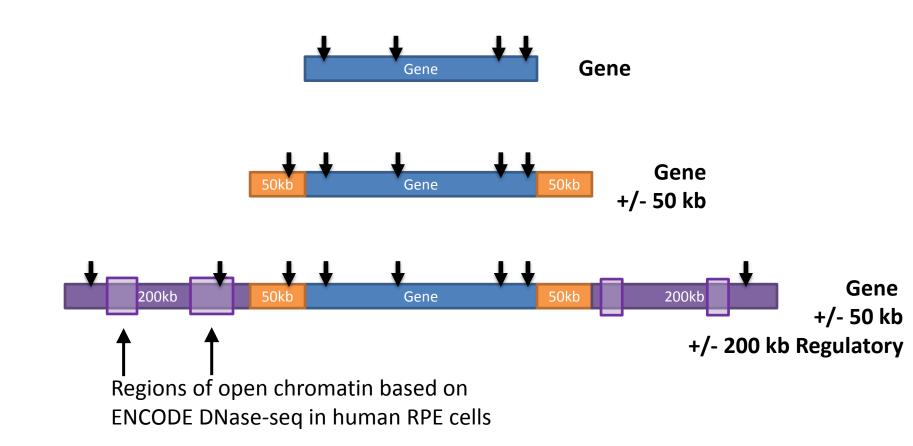
INTRODUCTION QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

## PATHWAY REGIONS



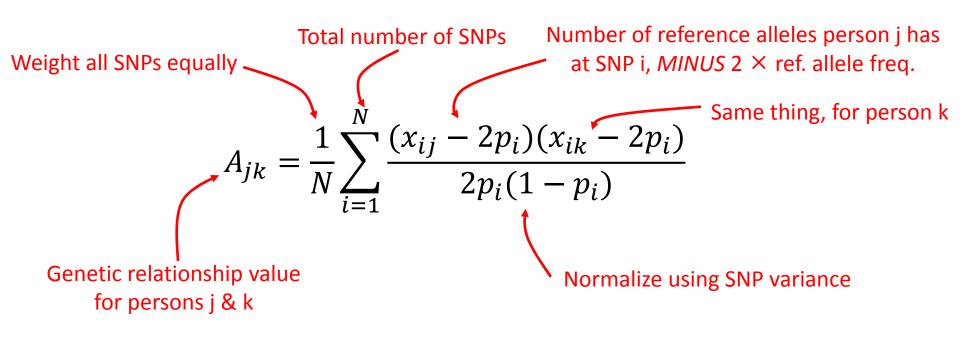
INTRODUCTION QUESTION N

METHODS

RESULTS

# MIXED MODEL ANALYSIS

- GCTA\* Genetic Relationship Matrices [GRMs] estimate genetic sharing among individuals in dataset using variance-covariance matrix
- Restricted Maximum Likelihood [REML] estimates the genetic contribution of each pathway (GRM) on AMD risk → Proportion of AMD Risk Explained



QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

## **MODEL SIGNIFICANCE**

#### Likelihood Ratio Test (LRT)

Full:pathway GRM + rest GRM + Covariates = LikelihoodReduced:rest GRM + Covariates = LikelihoodReduced:rest GRM + Covariates = Likelihood

$$D = -2\ln\left(\frac{Likelihood_{Full}}{Likelihood_{Reduced}}\right)$$

D statistic used to determine the probability that the Pathway/GRM impacts risk for AMD

(All analyses include Age, Sex, & 2 Principal Components)

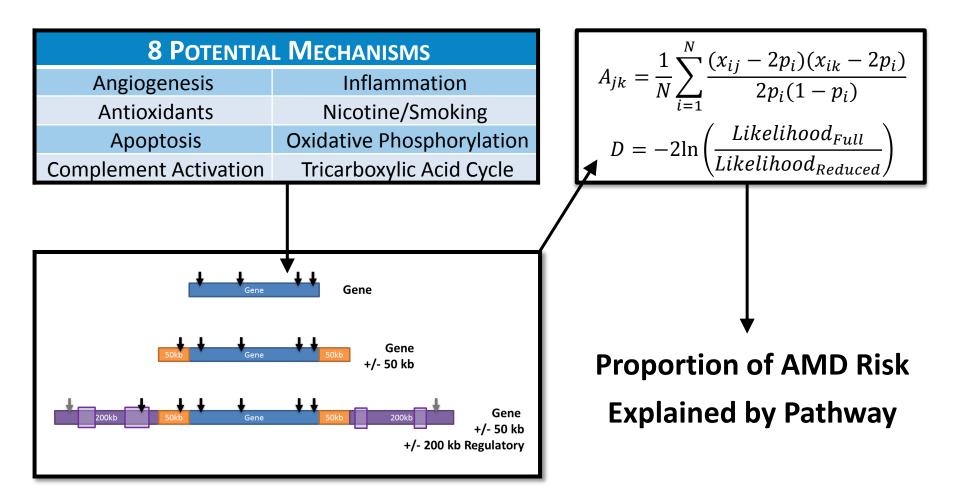
INTRODUCTION QUESTION

TION

METHODS

RESULTS

## **OVERVIEW**



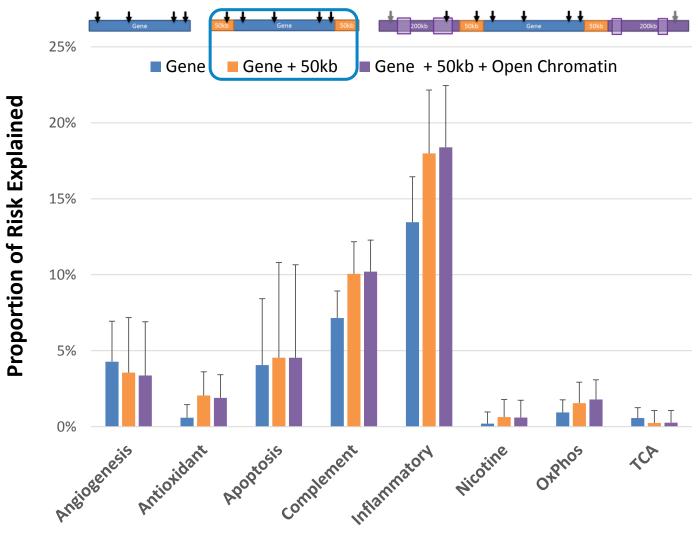
QUESTION

**M**ETHODS

**C**ONCLUSIONS

**R**ESULTS

#### RESULTS **PATHWAY REGIONS**



...Known Risk SNPs ?

QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS

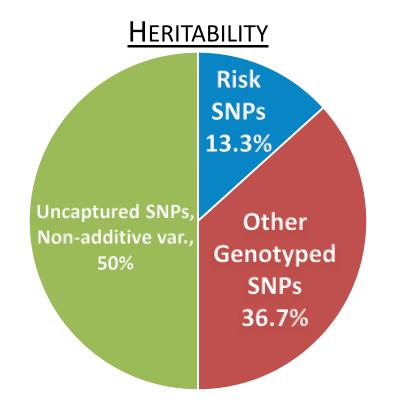
## **RESULTS** | **RISK SNPS**

#### Known Risk GRM + Rest GRM + Covariates = Likelihood<sub>Full</sub>

#### **Proportion of AMD risk explained**

19 Risk SNPs: 13.3% (p=1.4-61)

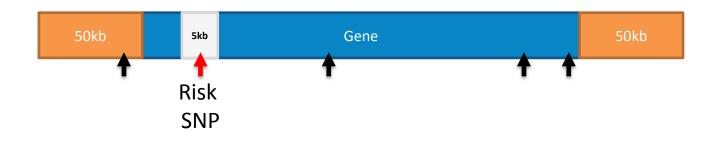
+ 5 kb flanking: 15.4% (p=1.6<sup>-53</sup>)



INTRODUCTION QUESTION METHODS RESULTS CONCLUSIONS

 RESULTS
 PRIMARY ANALYSIS PARAMETERS

#### [Genes ± 50 kb]



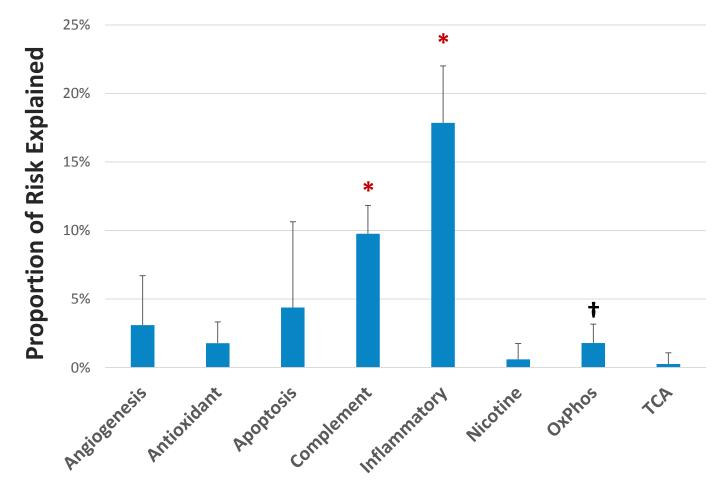
QUESTION

METHODS

**RESULTS** CONCLUSIONS

### **RESULTS** PRIMARY ANALYSIS [EXCLUDING KNOWN RISK]

... # Genes / Pathway?

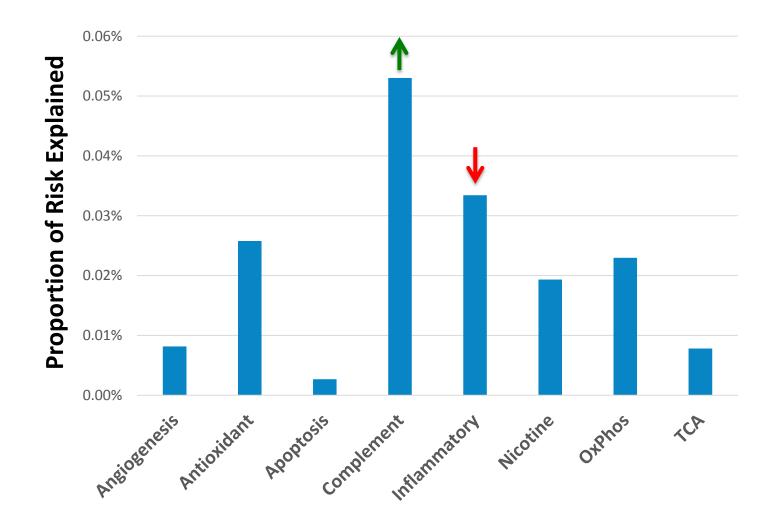


\* Significant p-value (Complement: 6.8 × 10<sup>-26</sup> / Inflammatory: 9.5 × 10<sup>-8</sup>)

**†** Oxidative Phosphorylation p-value: **0.08** 

INTRODUCTION QUESTION METHODS RESULTS CONCLUSIONS

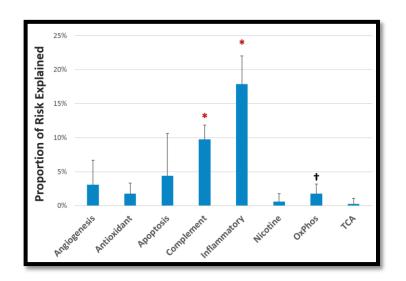
 RESULTS
 RISK EXPLAINED / GENE

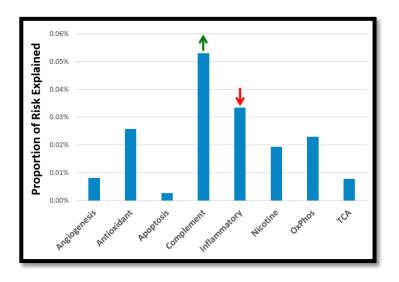


INTRODUCTION QUESTION

## CONCLUSIONS

- Accounting for known risk SNPs, **complement** pathway contains fewer additional SNPs with higher average AMD risk explained
- Inflammatory pathway contains more SNPs with smaller effects, but more overall AMD risk explained
- Variants in open chromatin regions 50kb 250kb away from gene explain *little* AMD risk





#### ACKNOWLEDGEMENTS

#### **CASE WESTERN RESERVE UNIVERSITY**

WILLIAM S. BUSH (ADVISOR) JONATHAN L. HAINES

VANDERBILT UNIVERSITY MILAM A. BRANTLEY

INITIAL FUNDING SOURCE NIH TRAINING GRANT 1T32EY021453-01

#### THE UNIVERSITY OF MIAMI

Anita Agarwal Jacklyn L. Kovach Margaret A. Pericak-Vance Stephen D. Schwartz William K. Scott

- jakehall@case.edu -

QUESTION

**M**ETHODS

RESULTS

**C**ONCLUSIONS

## PATHWAY SNP OVERLAP

PATHWAY	Angiogenesis	Antioxidant	Apoptotic	Complement	Inflammatory	Nicotine	OxPhos	тса
Angiogenesis	100.0%	0.0%	8.2%	1.3%	8.2%	0.6%	0.2%	0.0%
Antioxidant	0.0%	100.0%	0.2%	0.0%	0.7%	0.0%	0.4%	0.0%
Apoptotic	8.2%	0.2%	100.0%	0.2%	7.3%	0.5%	0.4%	0.0%
Complement	1.3%	0.0%	0.2%	100.0%	6.1%	0.0%	0.0%	0.0%
Inflammatory	8.2%	0.7%	7.3%	6.1%	100.0%	1.6%	0.2%	0.0%
Nicotine	0.6%	0.0%	0.5%	0.0%	1.6%	100.0%	0.0%	0.0%
OxPhos	0.2%	0.4%	0.4%	0.0%	0.2%	0.0%	100.0%	1.1%
ТСА	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.1%	100.0%

		Nearby Gene(s)					
Index SNP Symbol		Name	Distance to index SNP [kb] / Location				
rs10490924	ARMS2	Age-related maculopathy susceptibility 2	0 / coding sequence				
1510490924	HTRA1	HtrA serine peptidase 1	6.6 / upstream				
rs10737680	CFH	Complement factor H	0 / intronic				
	C2	Complement component 2	17 / downstream				
rs429608	CFB	Complement factor B	10.6 / downstream				
	SKIV2L	Superkiller viralicidic activity 2-like (S. cerevisiae)	0 / intronic				
rs2230199	СЗ	Complement component 3	0 / coding sequence				
rs5749482	TIMP3	TIMP metallopeptidase inhibitor 3	137.1 / upstream				
153749402	SYN3	Synapsin III	0 / intronic				
rs4420638	APOE	Apolipoprotein E	10.3 / downstream				
154420050	APOC1	Apolipoprotein C-I	5.0 / downstream				
rs1864163	CETP	Cholesteryl ester transfer protein, plasma	0 / intronic				
rs943080	VEGFA	Vascular endothelial growth factor A	72.4 / downstream				
rs13278062	TNFRSF10A	Tumor necrosis factor receptor superfamily, member 10a	0.3 / upstream				
rs920915	LIPC	Lipase, hepatic	35.7 / upstream				
CFI		Complement factor I	71.4 / downstream				
rs4698775	CCDC109B	Coiled-coil domain containing 109B	0 / intronic				
rs3812111	COL10A1	Collagen, type X, alpha 1	0 / intronic				
rs13081855	COL8A1	Collagen, type VIII, alpha 1	0 / intronic				
rs3130783	IER3	Immediate early response 3	62 / upstream				
155150785	DDR1	Discoidin domain receptor tyrosine kinase 1	77.5 / upstream				
rs8135665	SLC16A8	Solute carrier family 16, member 8 (monocarboxylic acid transporter 3)	0 / intronic				
rs334353	TGFBR1	Transforming growth factor, beta receptor 1	0 / intronic				
rs8017304	RAD51B	RAD51 homolog B (S. cerevisiae)	0 / intronic				
	ADAMTS9 ADAM metallopeptidase with thrombospondin type 1 motif, 9		32 / upstream				
rs6795735	ADAMTS9-AS2	ADAMTS9 antisense RNA 2 (non-protein coding)	0 / intronic				
	MIR548A2	microRNA 548a-2	0.3 / upstream				
rs9542236	<b>B3GALTL</b>	Beta 1,3-galactosyltransferase-like	0 / intronic				

INTRODUCTION QUESTION

METHODS

RESULTS

CONCLUSIONS

## LIMITATIONS

- Pathway definitions
- Same platform used in meta
- Europeans only
- SNPs overlapping between pathways
- Additional genetic risk for AMD may exist (non-additive variation, SNPs not genotyped, etc.)

## DATASET

#### Pre-QC:

Case / Control study design

- Cases: 1,247
- Controls: 708

Total SNPs: 906,688

- Affymetrix 6.0: 906,600
- Custom Sequenom: 84
- Custom TaqMan: 4

#### QUESTION

**M**ETHODS

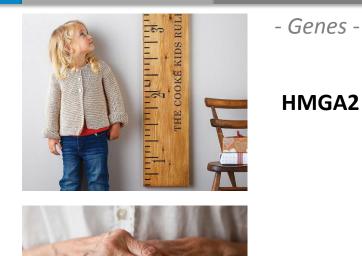
#### RESULTS CONCLUSIONS

- Trait Variance / Risk -

- Heritability -81%

55%

51%



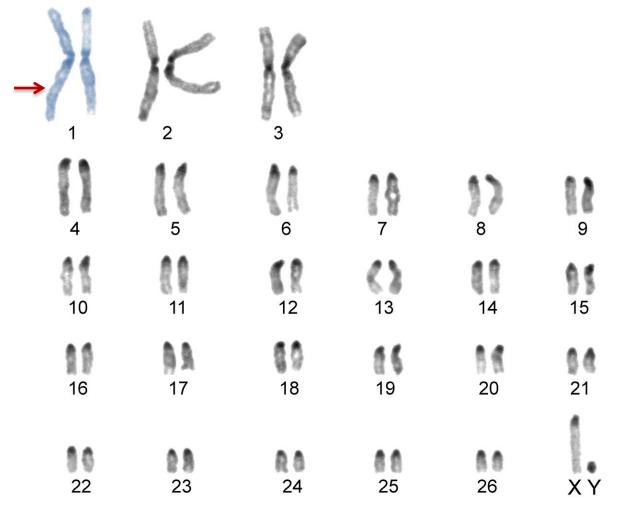
HMGA2

HLA-DRB1	Known Heritability Unknown
FTO TMEM18 TFAP2B MC4R	Environmental
CSTM1	

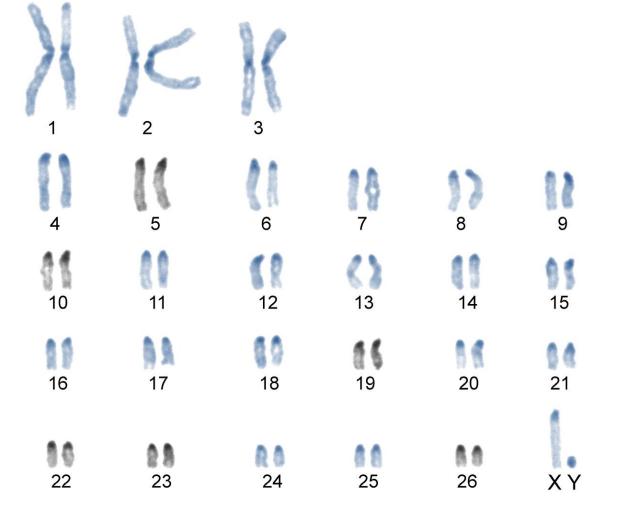
30%

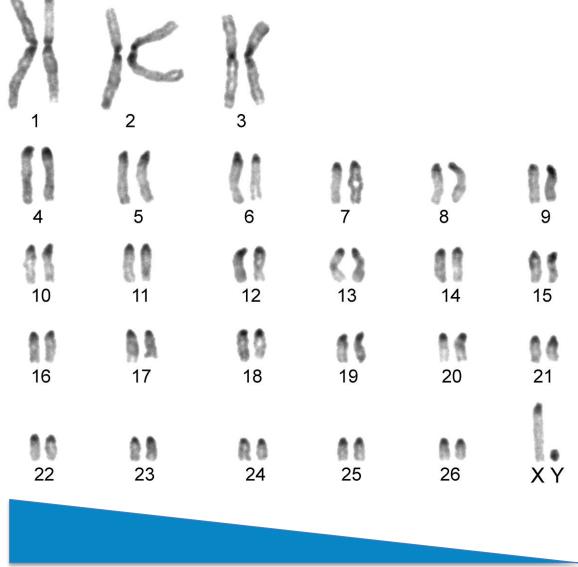


C2INIT **IL10** LTC4S > 100 genes



CONCLUSIONS





Chromosome, Chromosome Region, Pathway, Gene, SNP

#### 16/19 SNPs – Logistic Regression 639,825 SNPs; 1955 total individuals Age, Sex, 2 PCs covariates

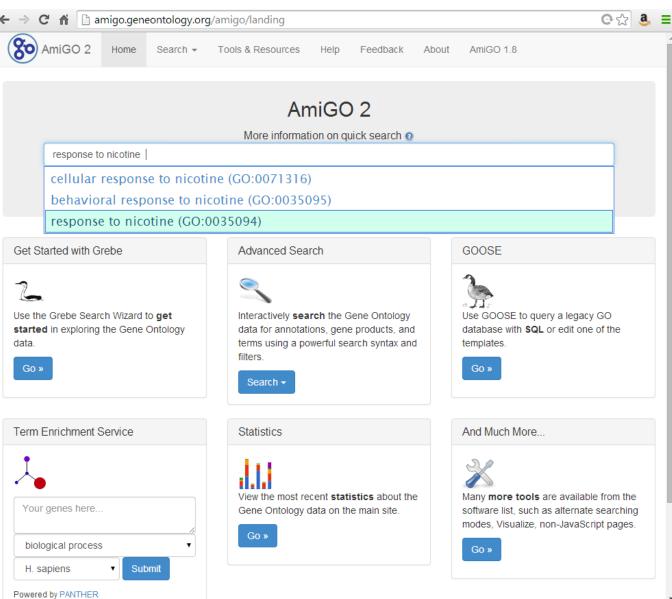
Gene Symbol	Chr.	rs#	P-value
CFH	1	rs10737680	1.67 E-25
ARMS2-HTRA	10	rs10490924	9.85 E-24
C2-CFB	6	rs429608	8.69 E-13
С3	19	rs2230199	2.57 E-06
CEPT	16	rs1864163	2.02 E-05
COL8A1	3	rs13081855	0.001
B3GALTL	13	rs9542236	0.001
APOE-APOC1	19	rs4420638	0.002
ADAMTS9-MIR548A2	3	rs6795735	0.018
TNFRSF10A	8	rs13278062	0.057
RAD51B	14	rs8017304	0.095
VEGFA	6	rs943080	0.103
TIMP3-SYN3	22	rs5749482	0.121
TGFBR1	9	rs334353	0.127
CFI	4	rs4698775	0.296
LIPC	15	rs920915	0.449

QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS



QUESTION

**M**ETHODS

RESULTS

Accession	GO:0035094
Name	response to nicotine
Ontology	biological_process
Synonyms	None
Definition	Any process that results in a change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a nicotine stimulus. <i>Source:</i> CHEBI:17688, ISBN:0198506732, ISBN:0582227089, GOC:bf, GOC:ef
Comment	None
History	See term history for GO:0035094 at QuickGO
Subset	None
Community	GN Add usage comments for this term on the GONUTS wiki.
Related	Link to all genes and gene products associated to response to nicotine.
	Link to all direct and indirect annotations to response to nicotine.
	Link to all direct and indirect <b>annotations download</b> (limited to first 10,000) for response to
	nicotine

QUESTION

**M**ETHODS

RESULTS

**C**ONCLUSIONS

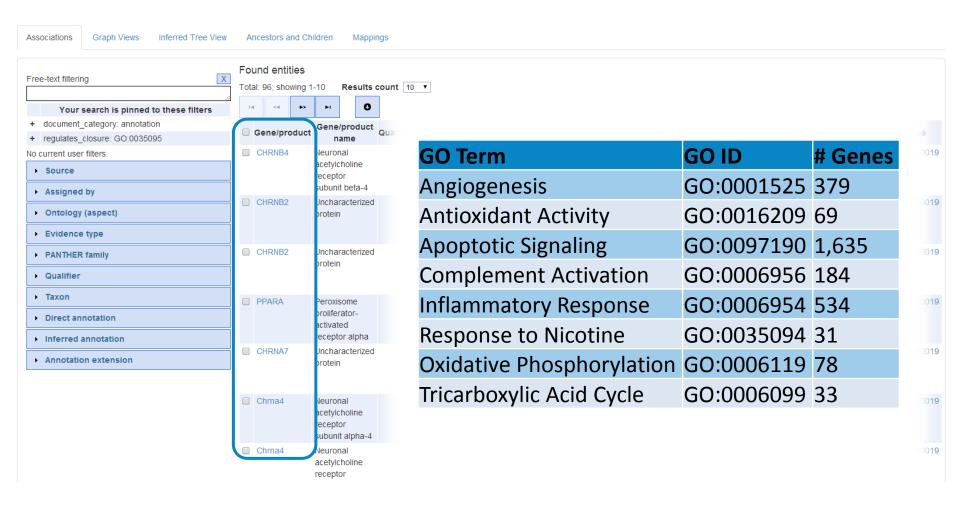
Associations	Graph Views	Inferred Tree View	Ancestors and Children	Mappings
I GO:000815	0 biological_proces	SS		
I GO:0050	896 response to s	timulus		
I GO:00	42221 response to	o chemical		
I GO	:0009719 response	e to endogenous stimulu	IS	
I GO	1901698 response	e to nitrogen compound		
I GO	:0010033 response	e to organic substance		
IG	GO:0010243 respo	nse to organonitrogen c	ompound	
1	GO:0043279 res	ponse to alkaloid		
1	GO:0014070 res	, ponse to organic cyclic (	compound	
(	▼ GO:0035094 I	response to nicotine		
	I GO:003509	5 behavioral response t	o nicotine	
		6 cellular response to ni		
	_			

QUESTION

**M**ETHODS

RESULTS

CONCLUSIONS



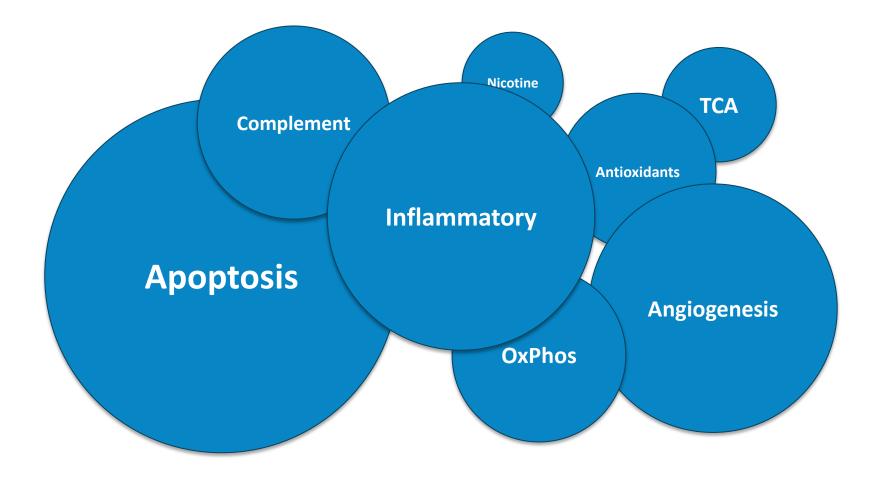
QUESTION

Methods

RESULTS

CONCLUSIONS

## **AMD PATHOGENESIS IN LITERATURE**



#### Extras

