Genome-Wide Association Study of Immune Response to Twelve Common Pathogens

Maike Morrison, Dr. Linda Kachuri, Dr. Sara Rashkin, Dr. John Witte

Department of Biostatistics and Epidemiology, University of California San Francisco

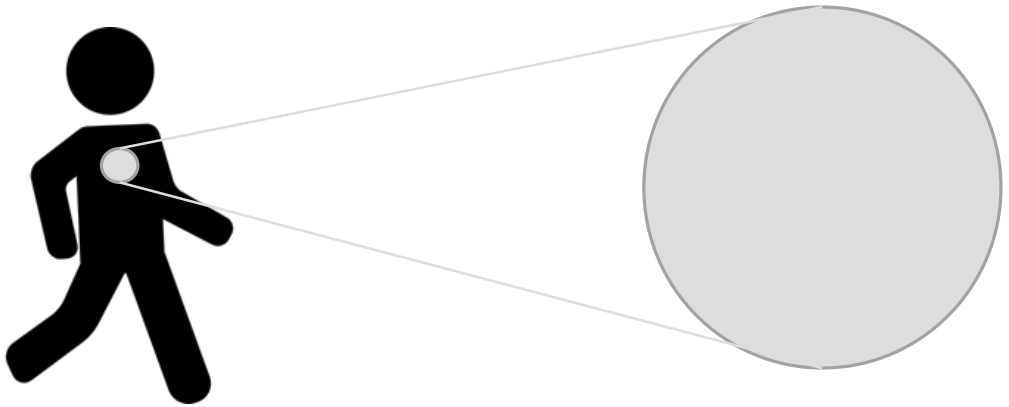
Summer Research Training Program 2019

UCSF Department of Epidemiology & Biostatistics

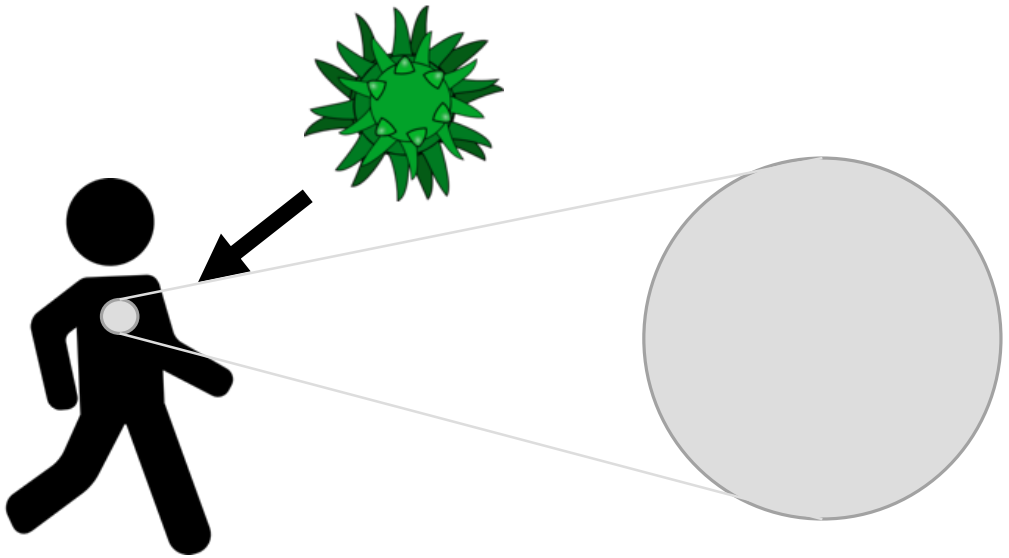
AMGEN Foundation

Amgen Scholars Program

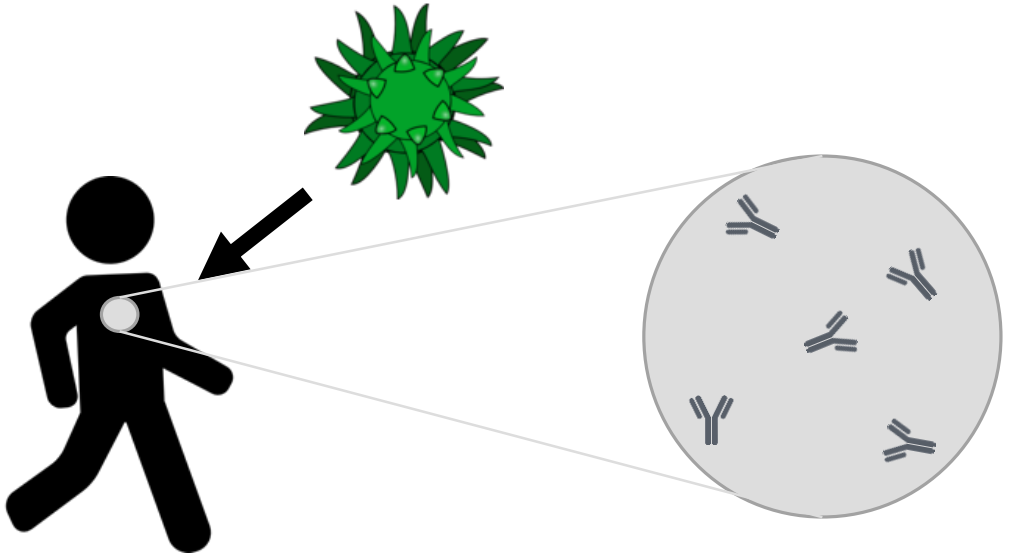
TEXAS The University of Texas at Austin



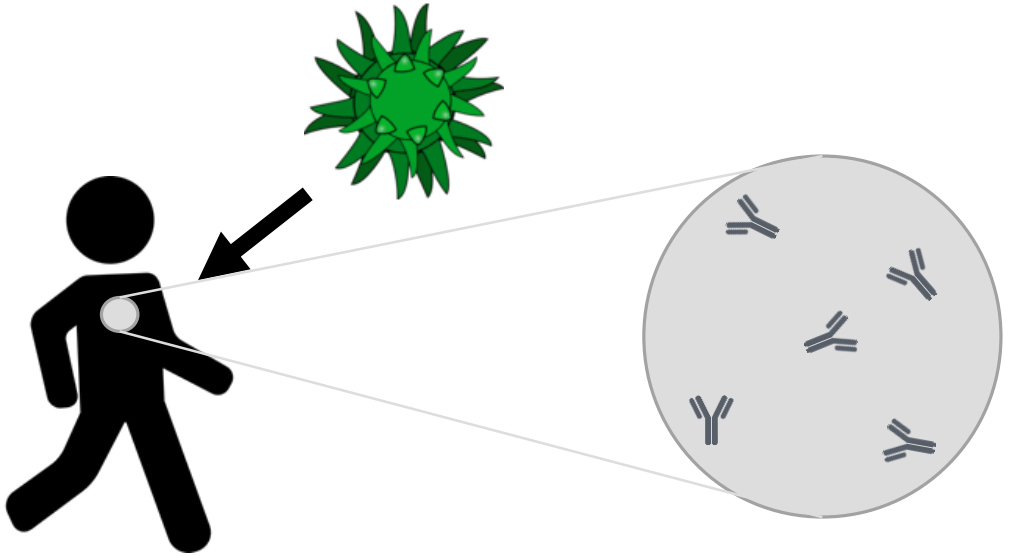
Immune response varies between individuals.

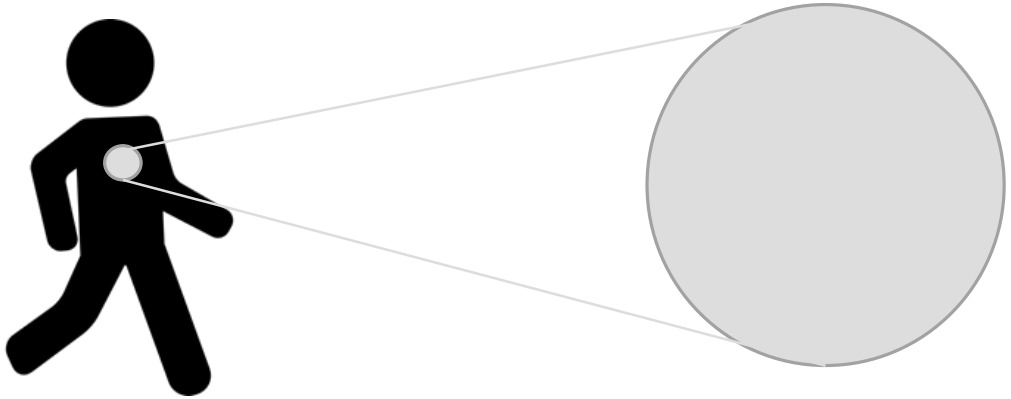


Immune response varies between individuals.

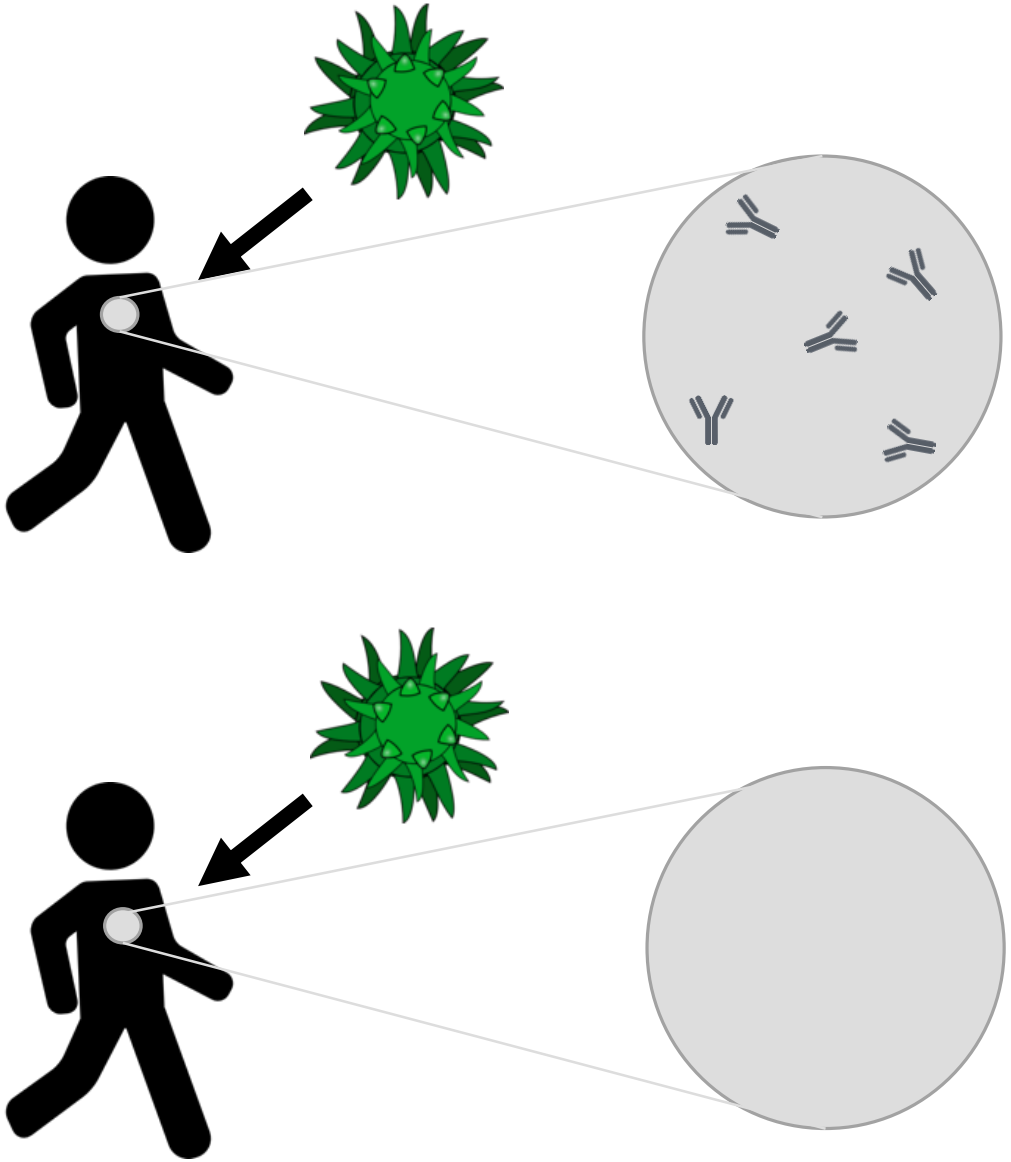


Immune response varies between individuals.

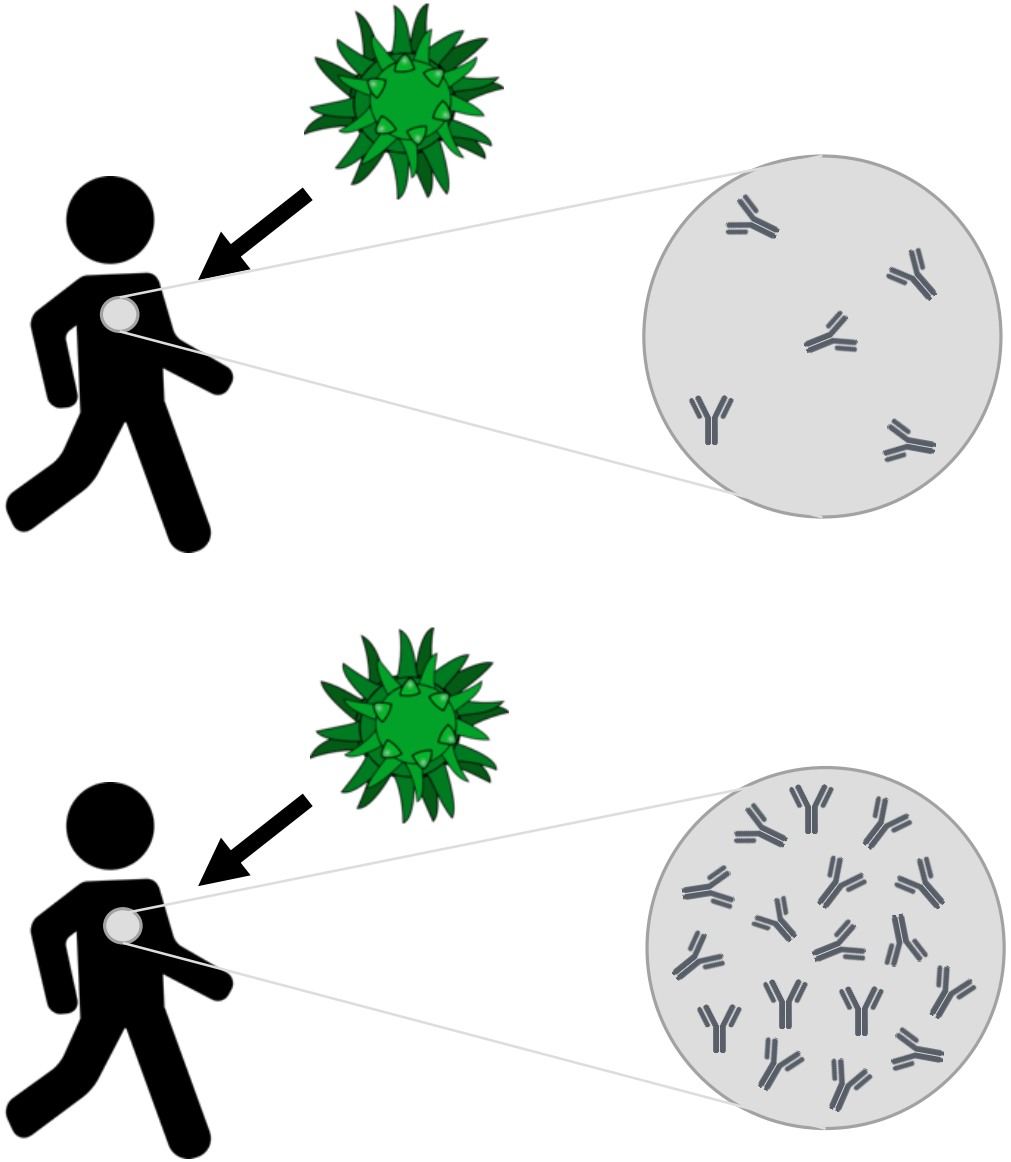




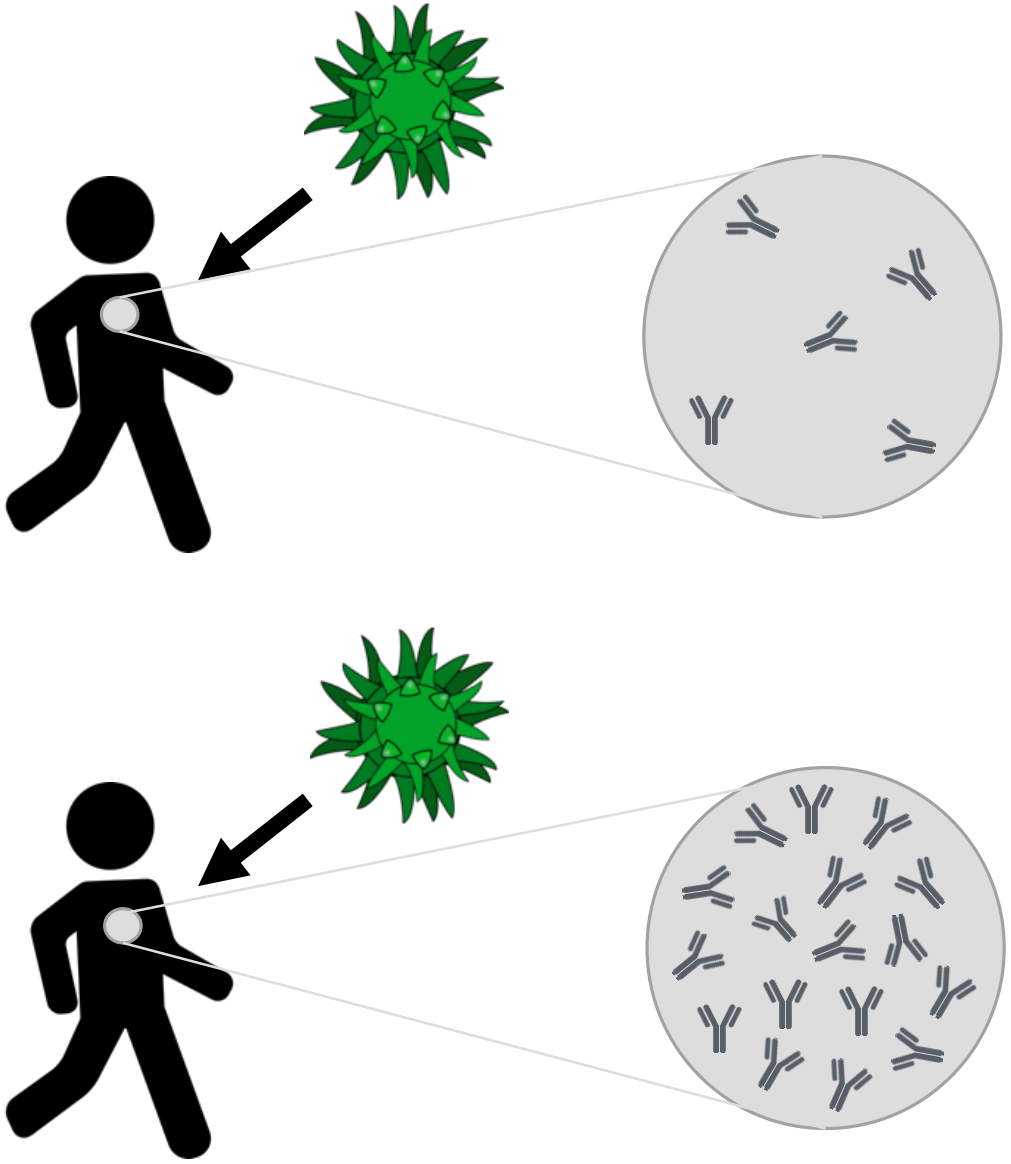
Immune response varies between individuals.



Immune response varies between individuals.

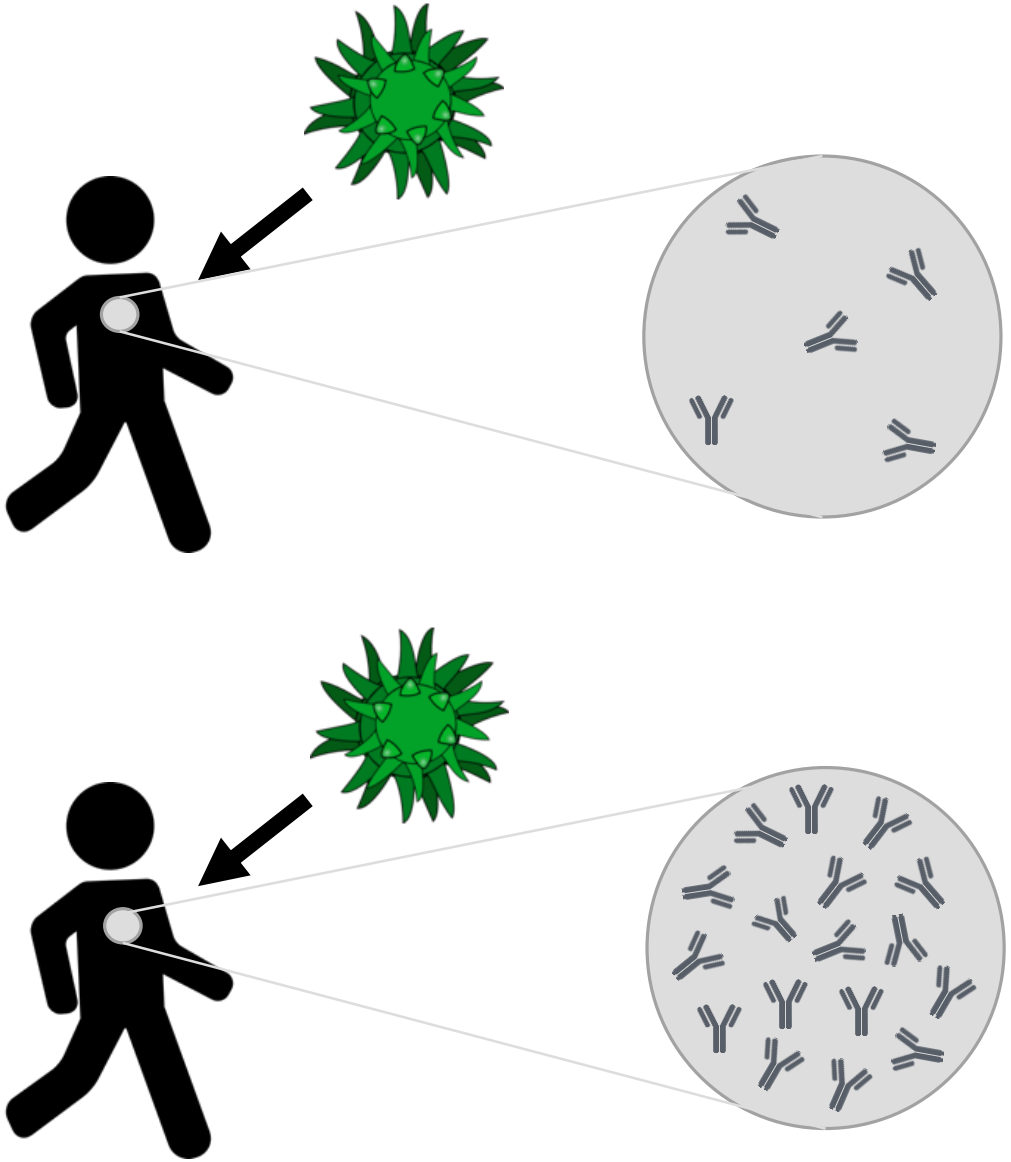


Immune response varies between individuals.



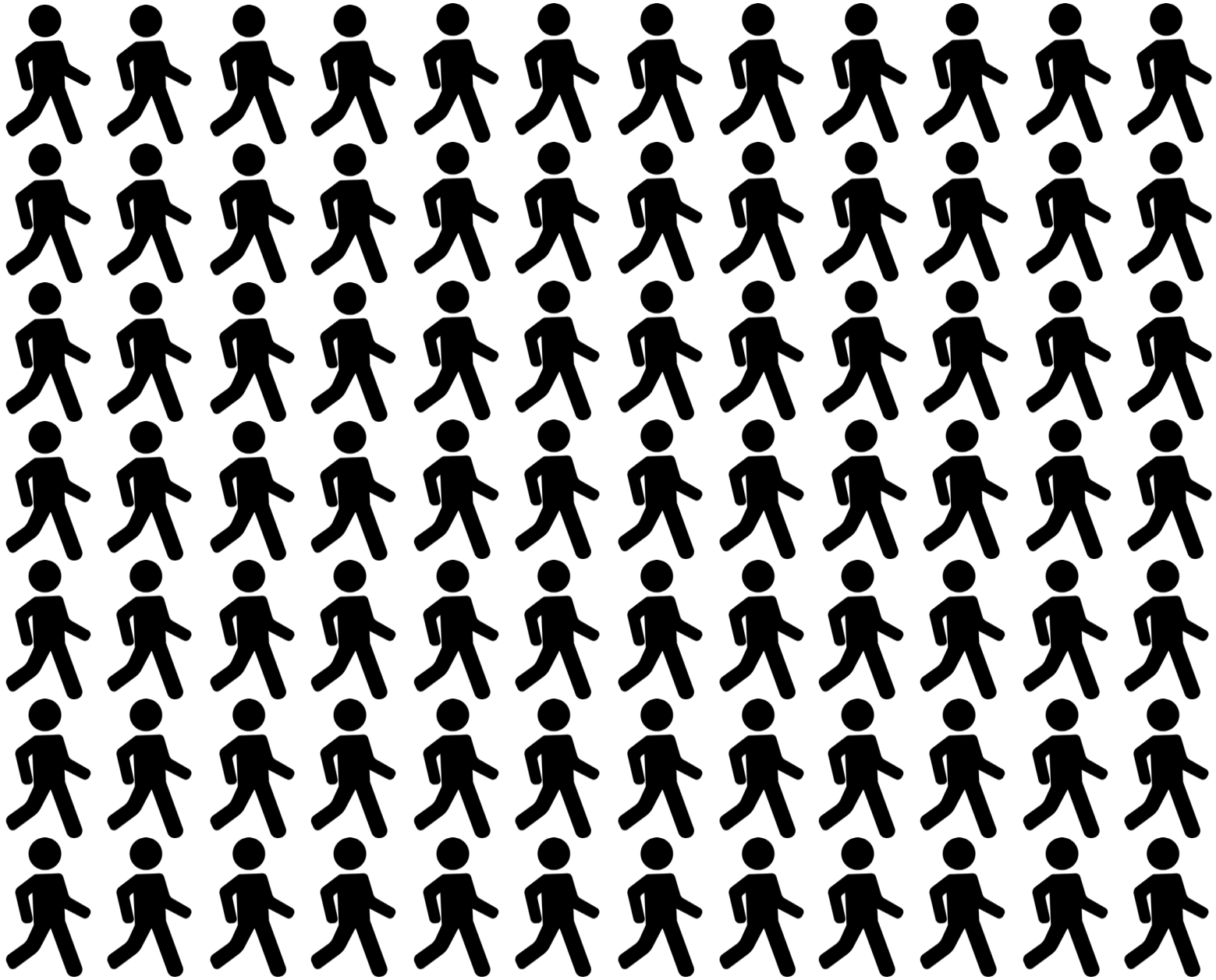
“Strength of Immune Response”

Immune response varies between individuals.

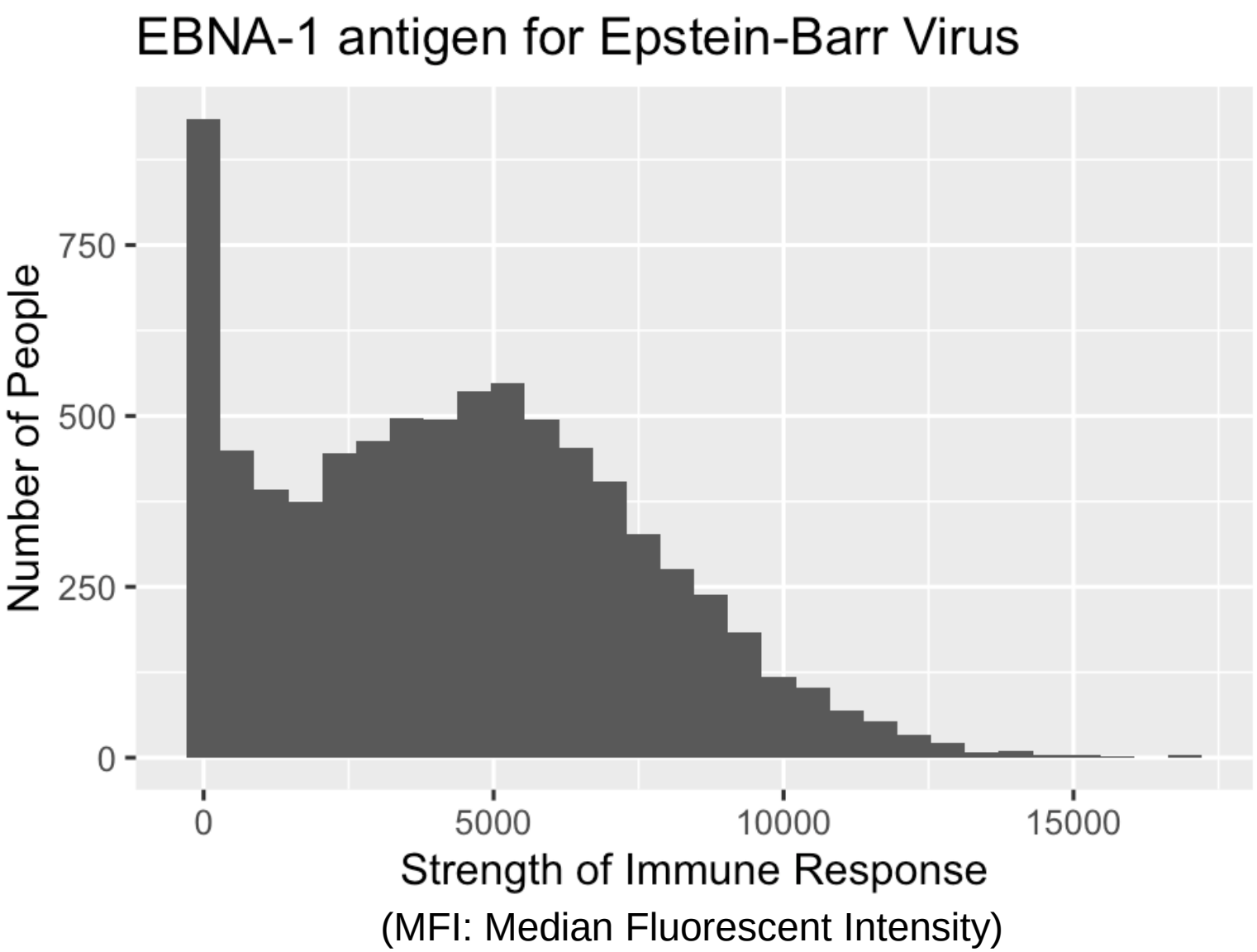


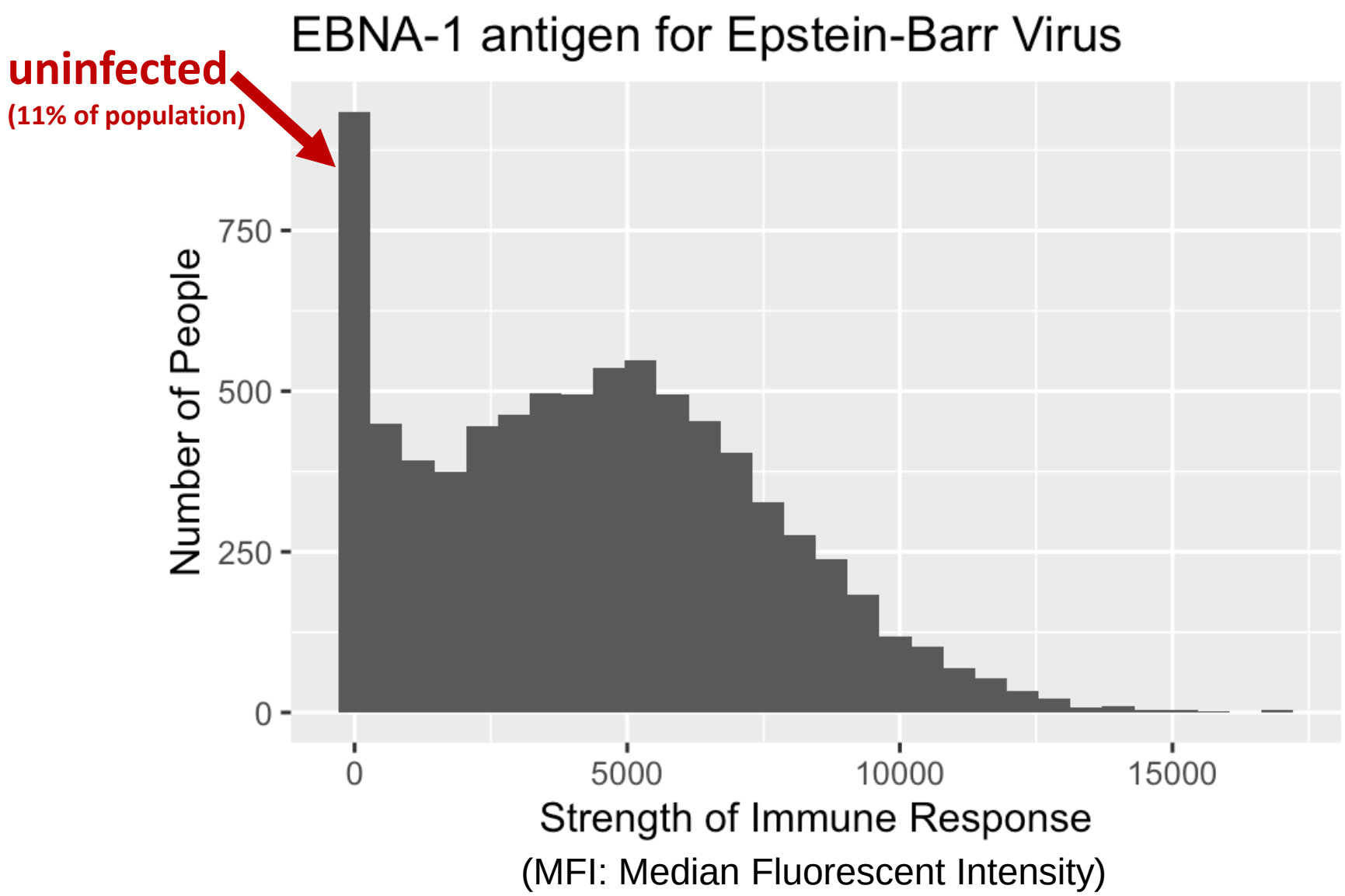
- Vaccine efficacy
- Pathogen-induced disease etiology
- Biological mechanisms of immune response
- Potential drug targets
- Immune-related cancer therapies

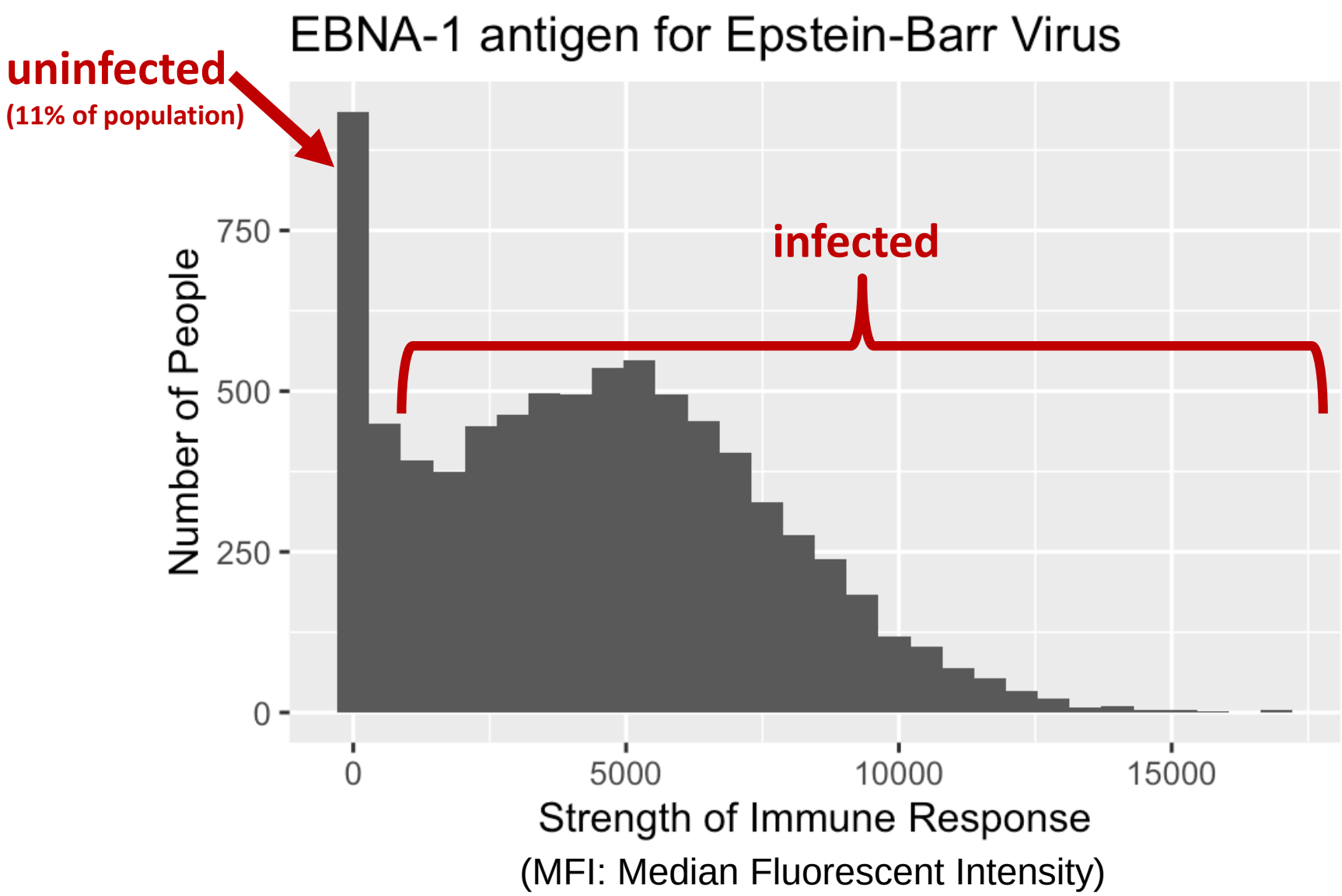
Immune response varies between individuals.

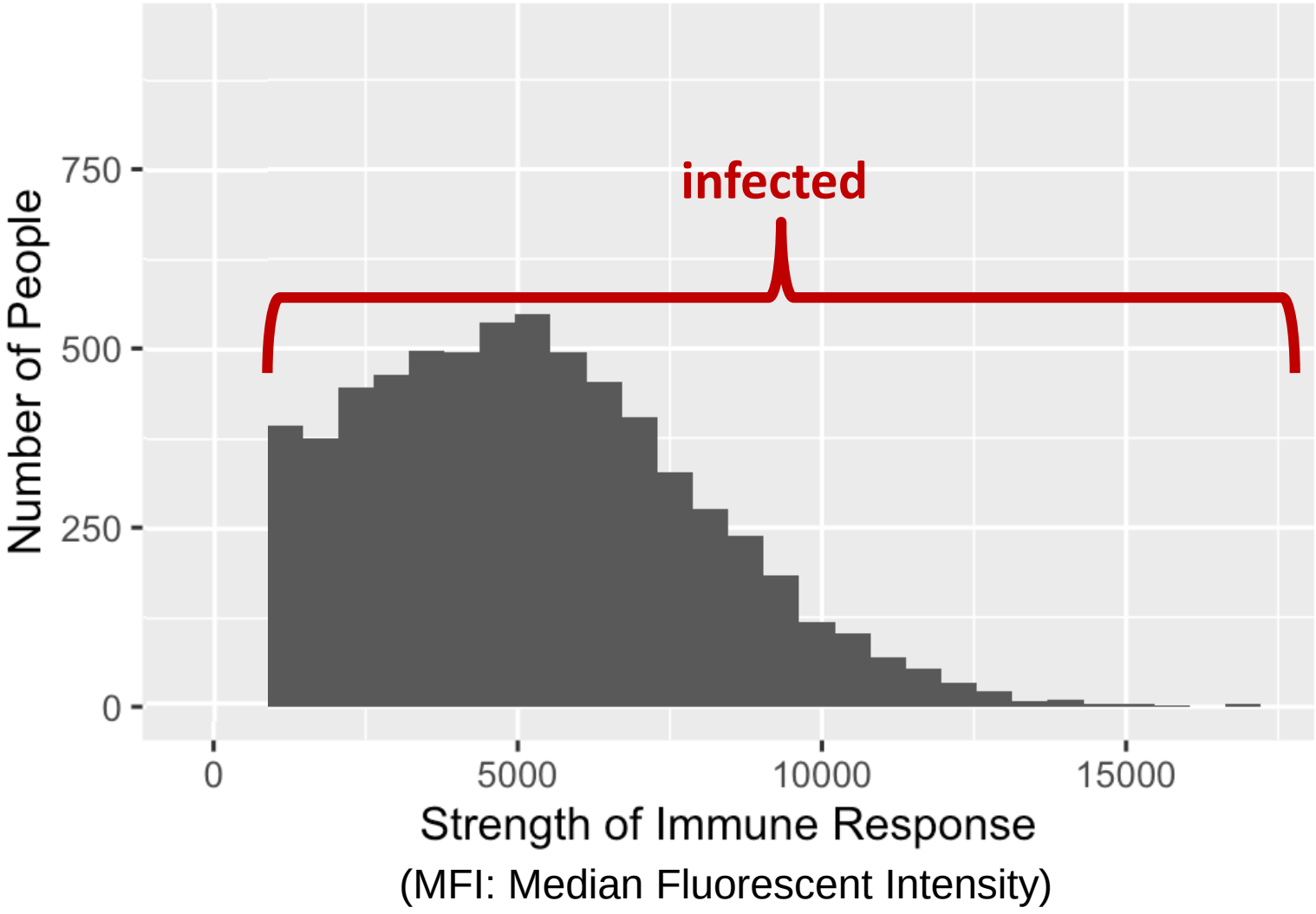


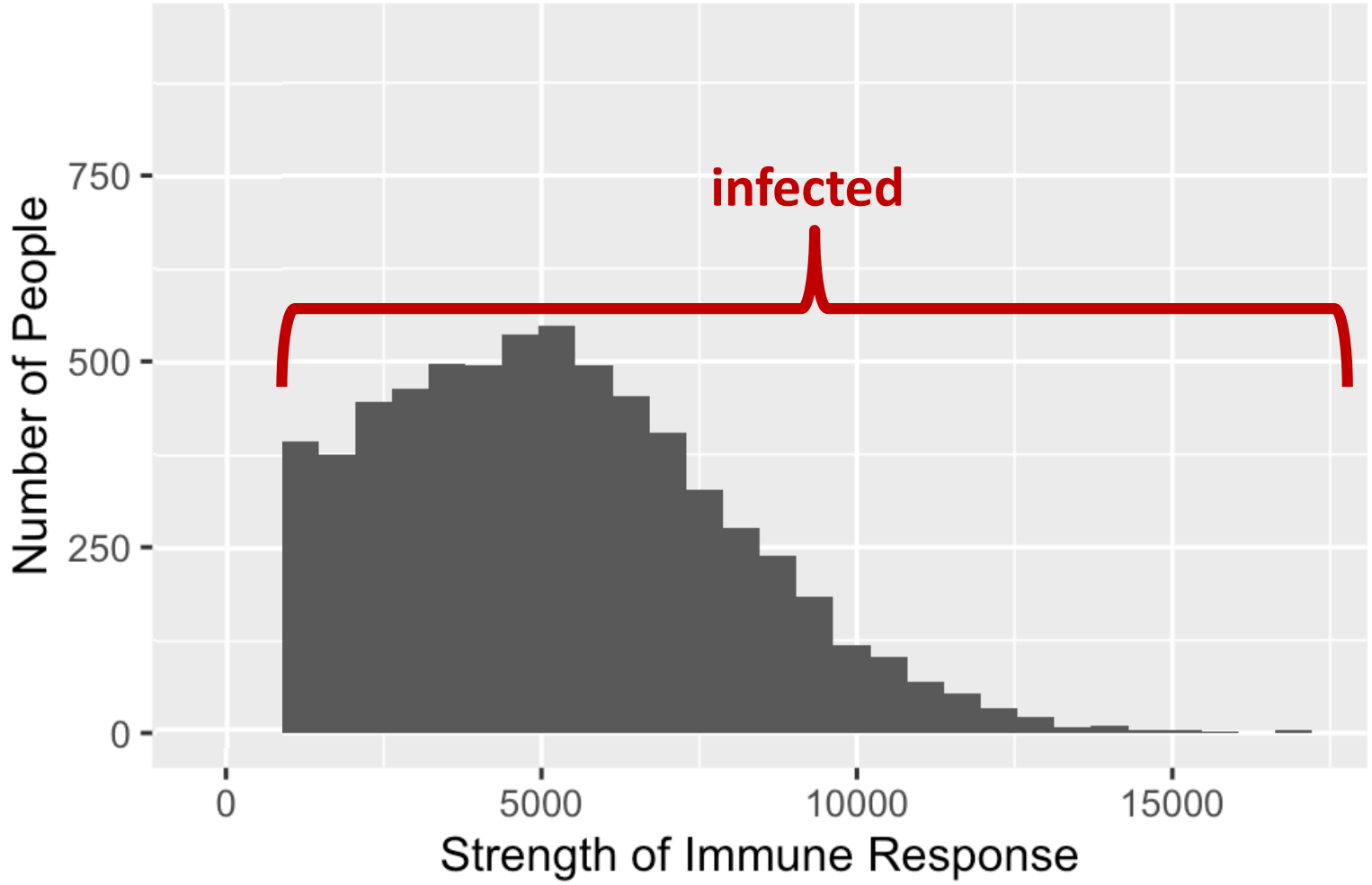
$N \approx 8{,}000$


EBNA-1 antigen for Epstein-Barr Virus
Number of People
750
500
250
0
0
5000
10000
15000
Strength of Immune Response
(MFI: Median Fluorescent Intensity)

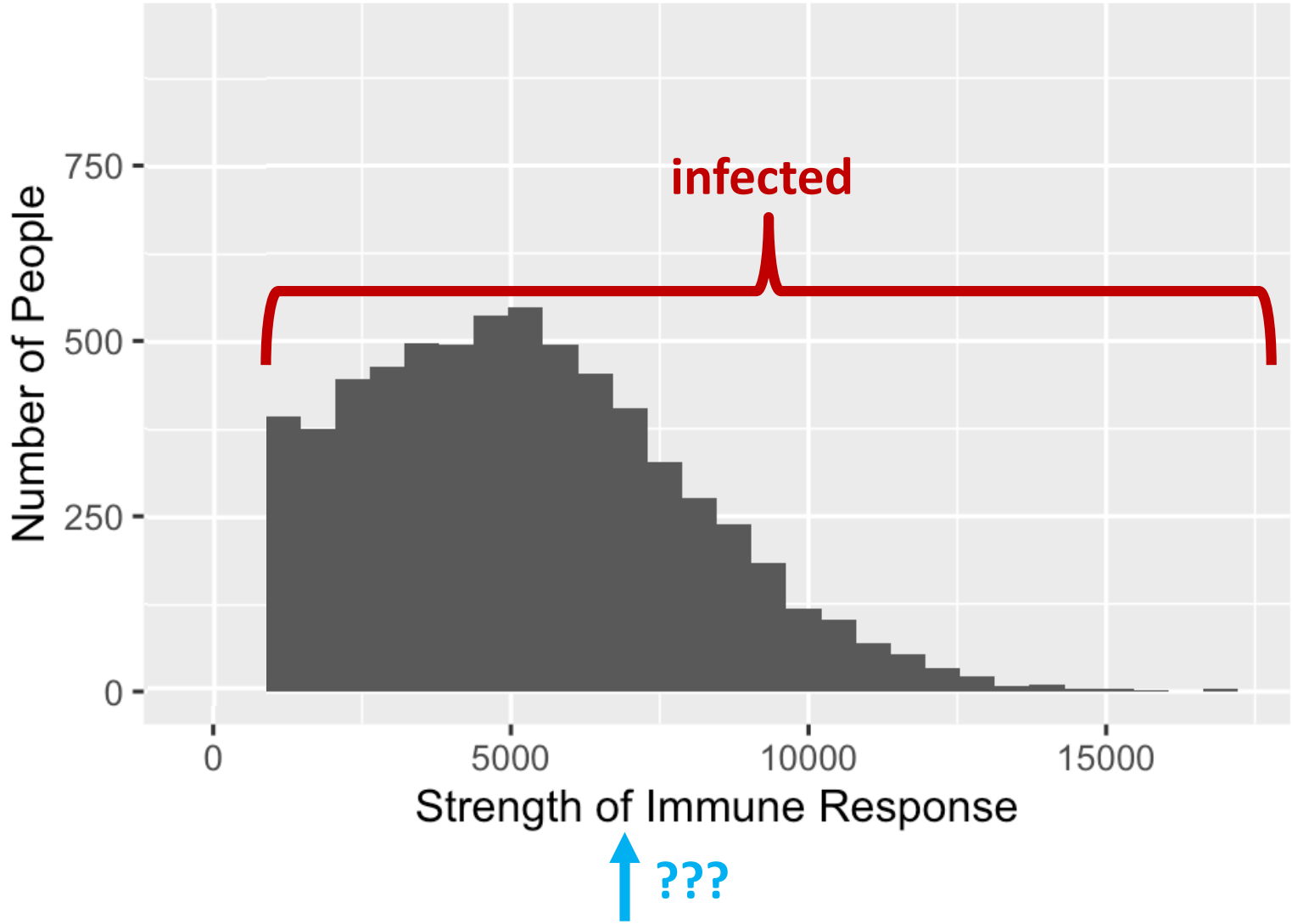

EBNA-1 antigen for Epstein-Barr Virus
uninfected
(11% of population)
Number of People
750
500
250
0
0
5000
10000
15000
Strength of Immune Response
(MFI: Median Fluorescent Intensity)


EBNA-1 antigen for Epstein-Barr Virus
uninfected
(11% of population)
infected
Number of People
750
500
250
0
0
5000
10000
15000
Strength of Immune Response
(MFI: Median Fluorescent Intensity)


infected
750
500
250
0
Number of People
0
5000
10000
15000
Strength of Immune Response
(MFI: Median Fluorescent Intensity)


infected
750
500
250
0
Number of People
0
5000
10000
15000
Strength of Immune Response

Among infected individuals, what is the genetic basis of variability in immune response?



Genome-Wide Association Study

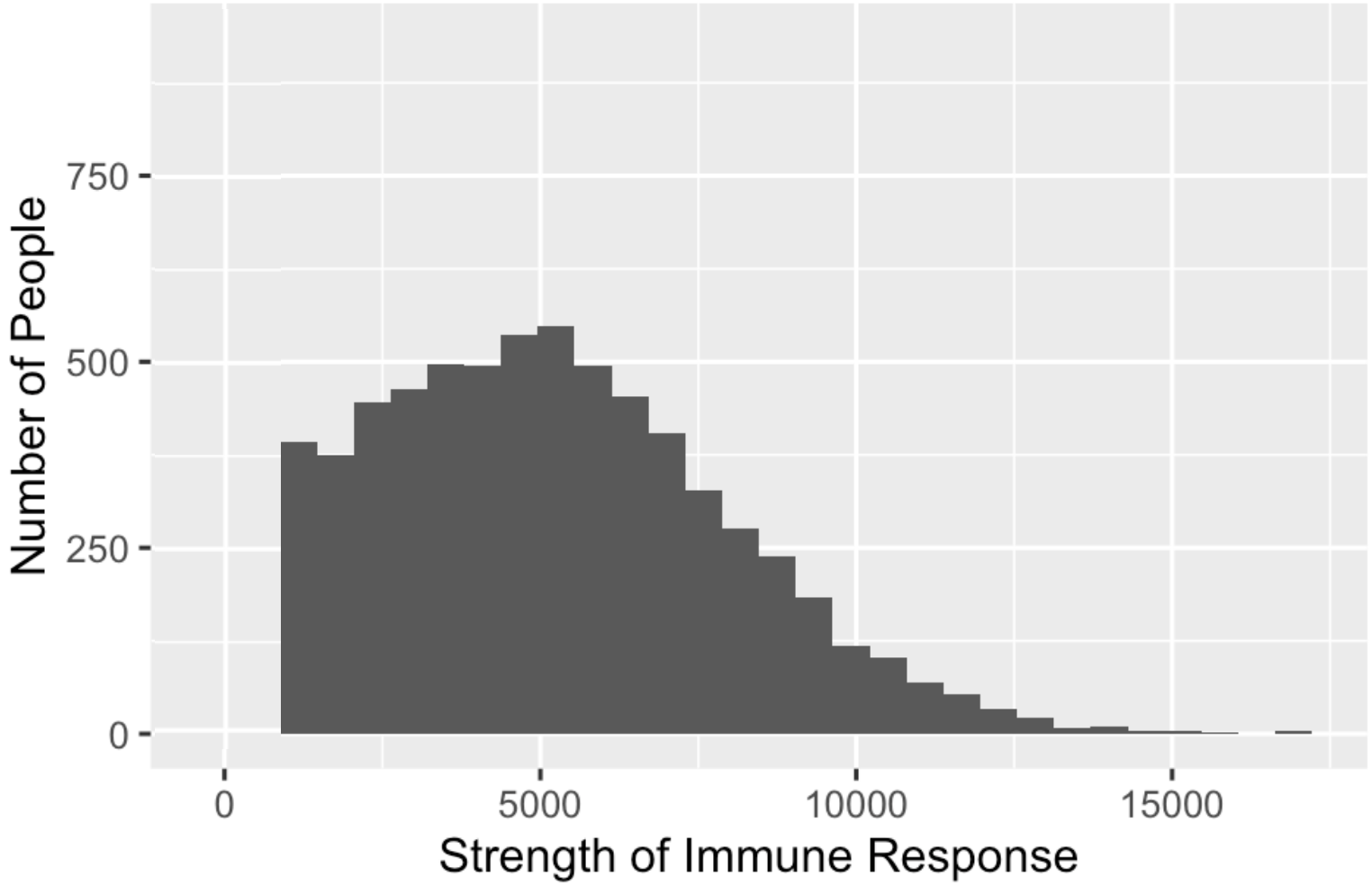
Among infected individuals, what is the genetic basis of variability in immune response?

Single-Nucleotide Polymorphism (“SNP”)

Single-Nucleotide Polymorphism (“SNP”)

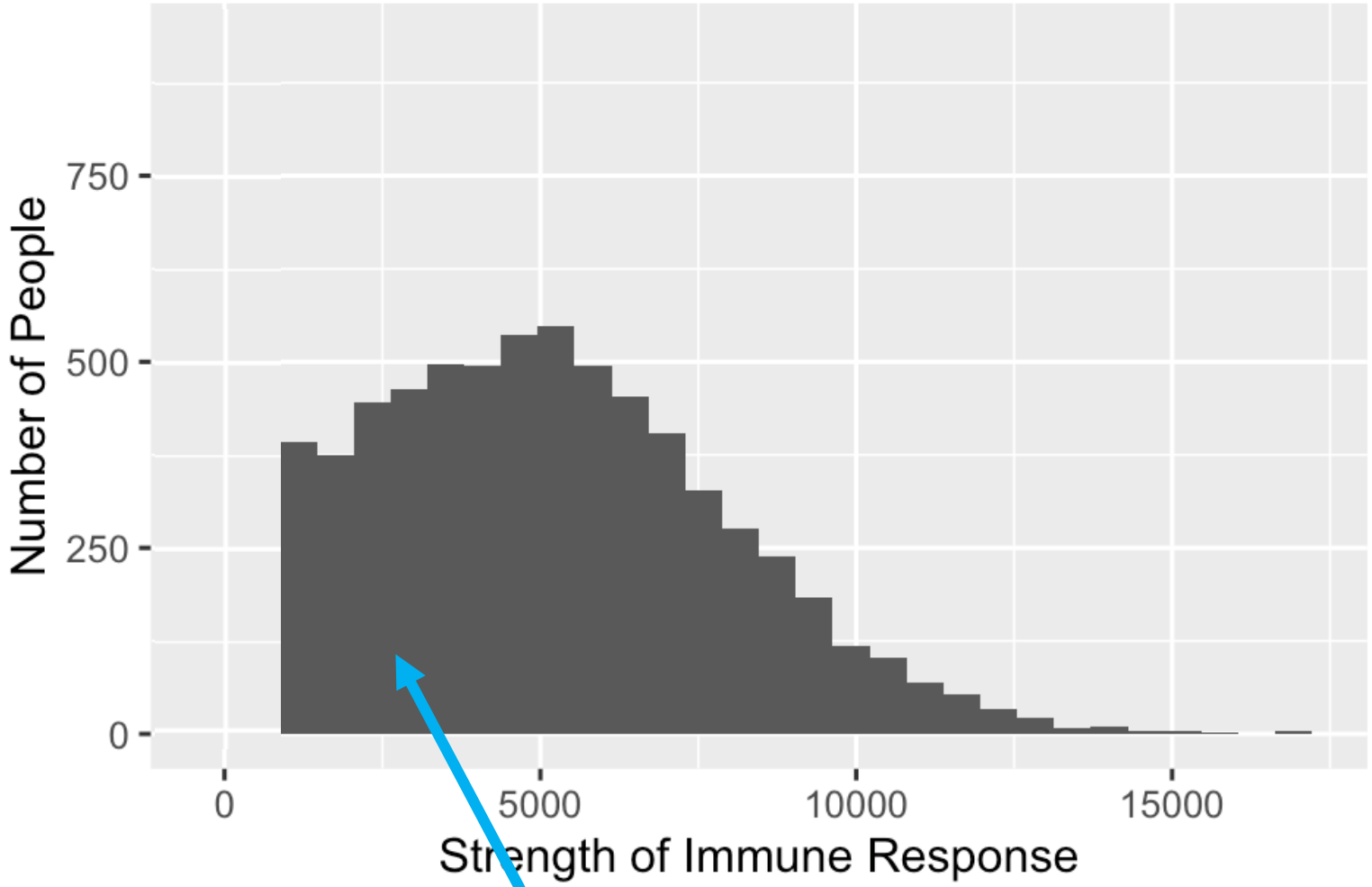
Single-Nucleotide Polymorphism (“SNP”)

Single-Nucleotide Polymorphism (“SNP”)



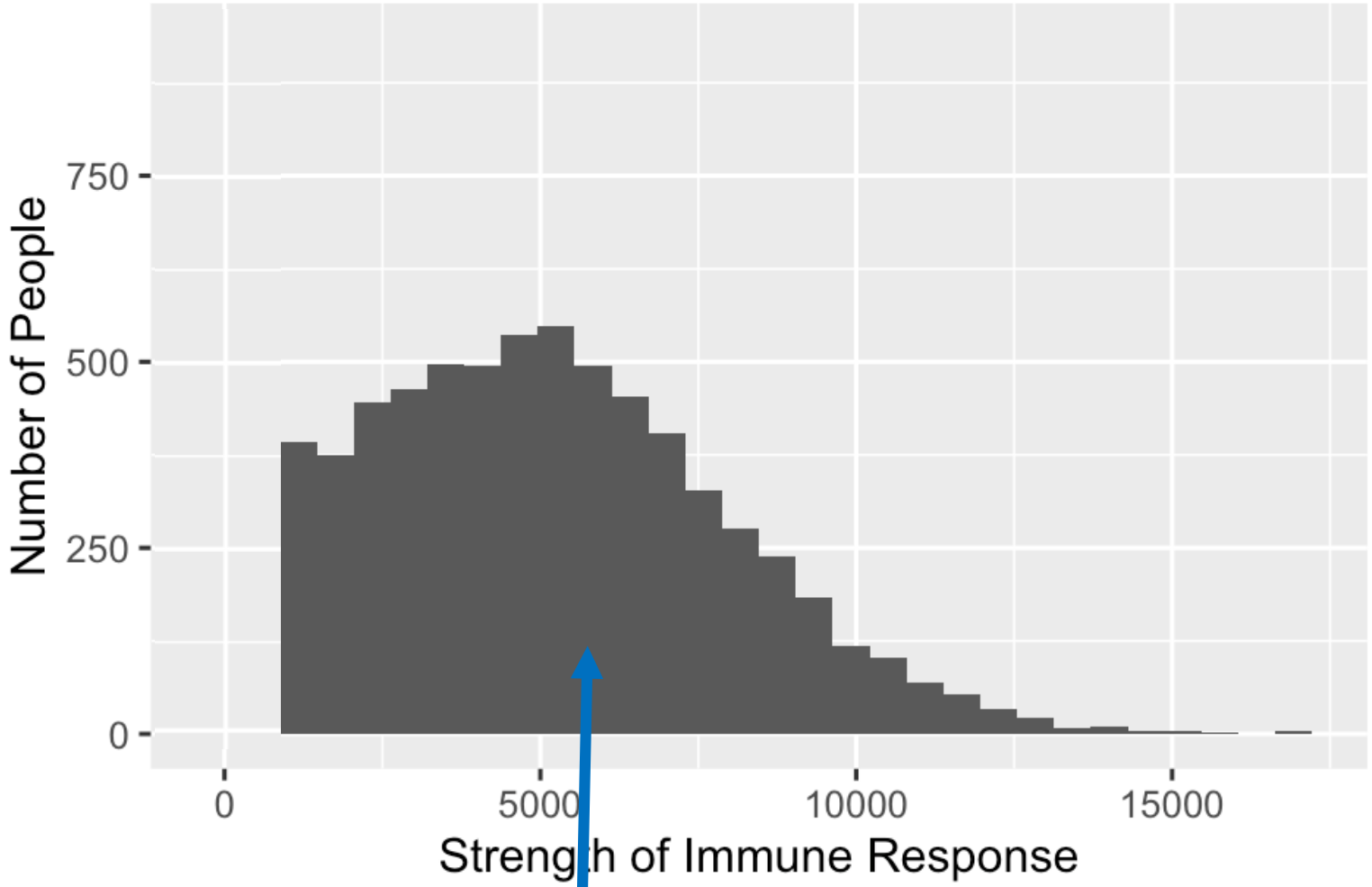
Single-Nucleotide Polymorphism (“SNP”)

AA



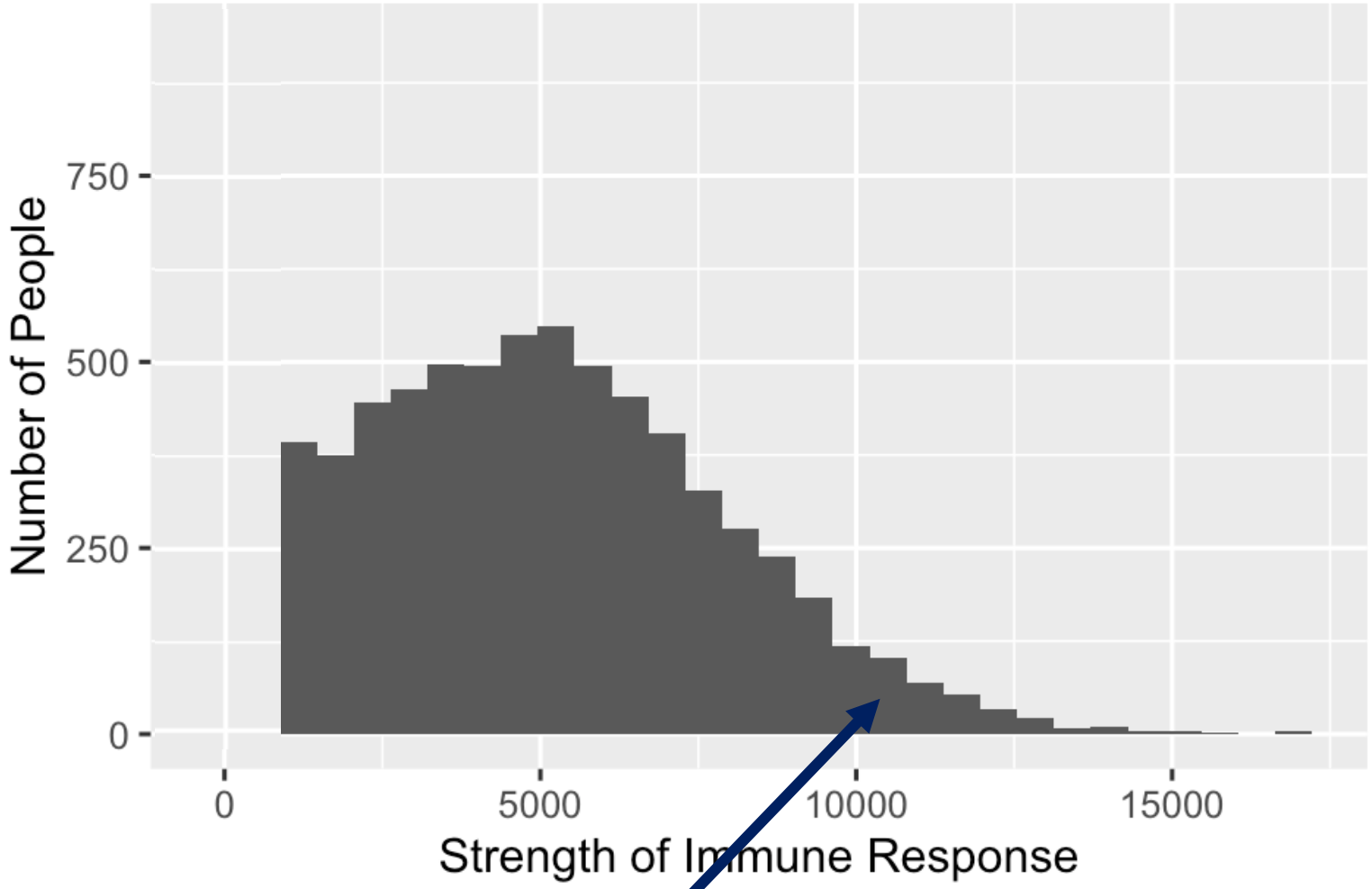
Single-Nucleotide Polymorphism (“SNP”)

AA



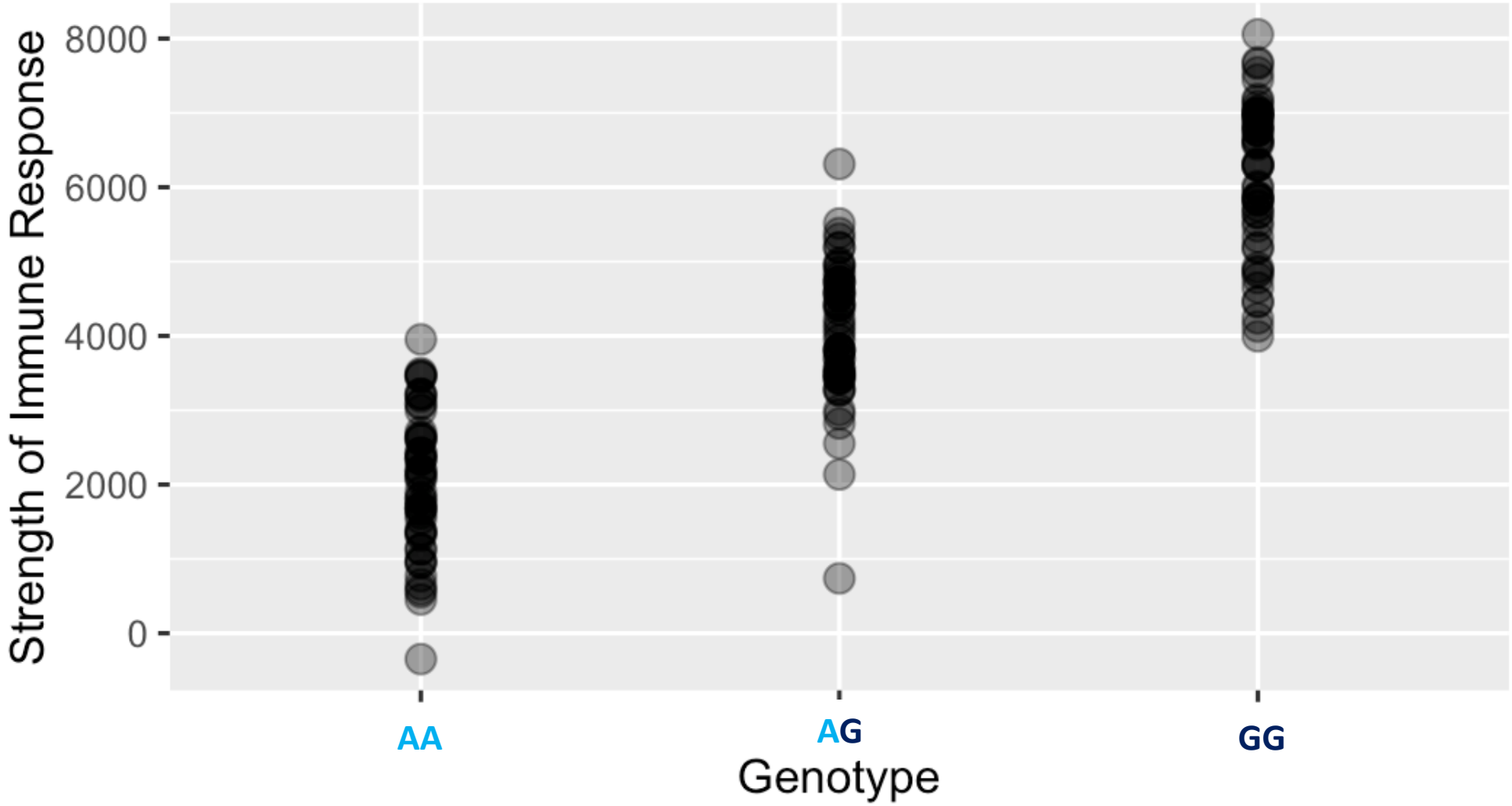
Single-Nucleotide Polymorphism (“SNP”)

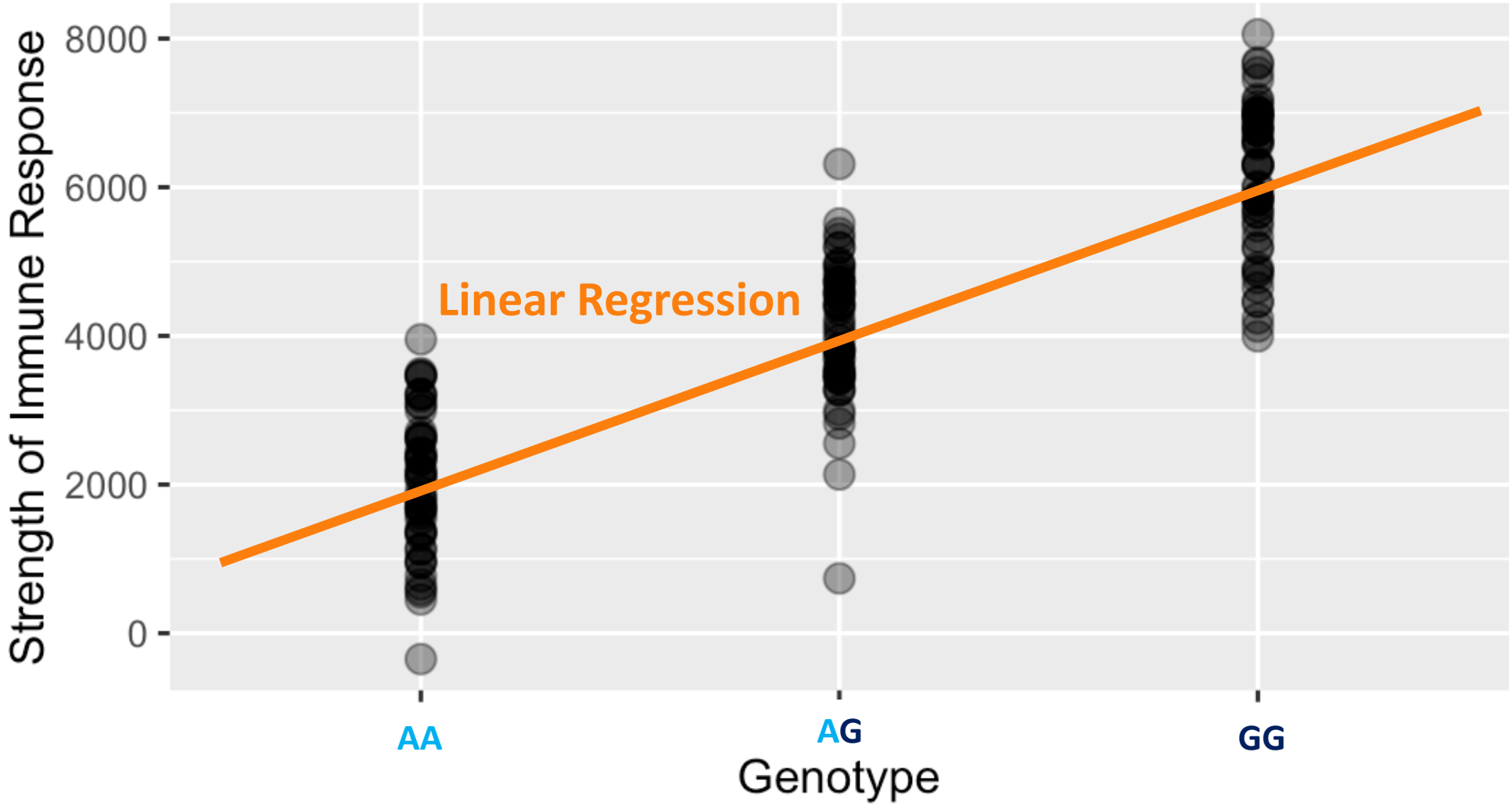
AG

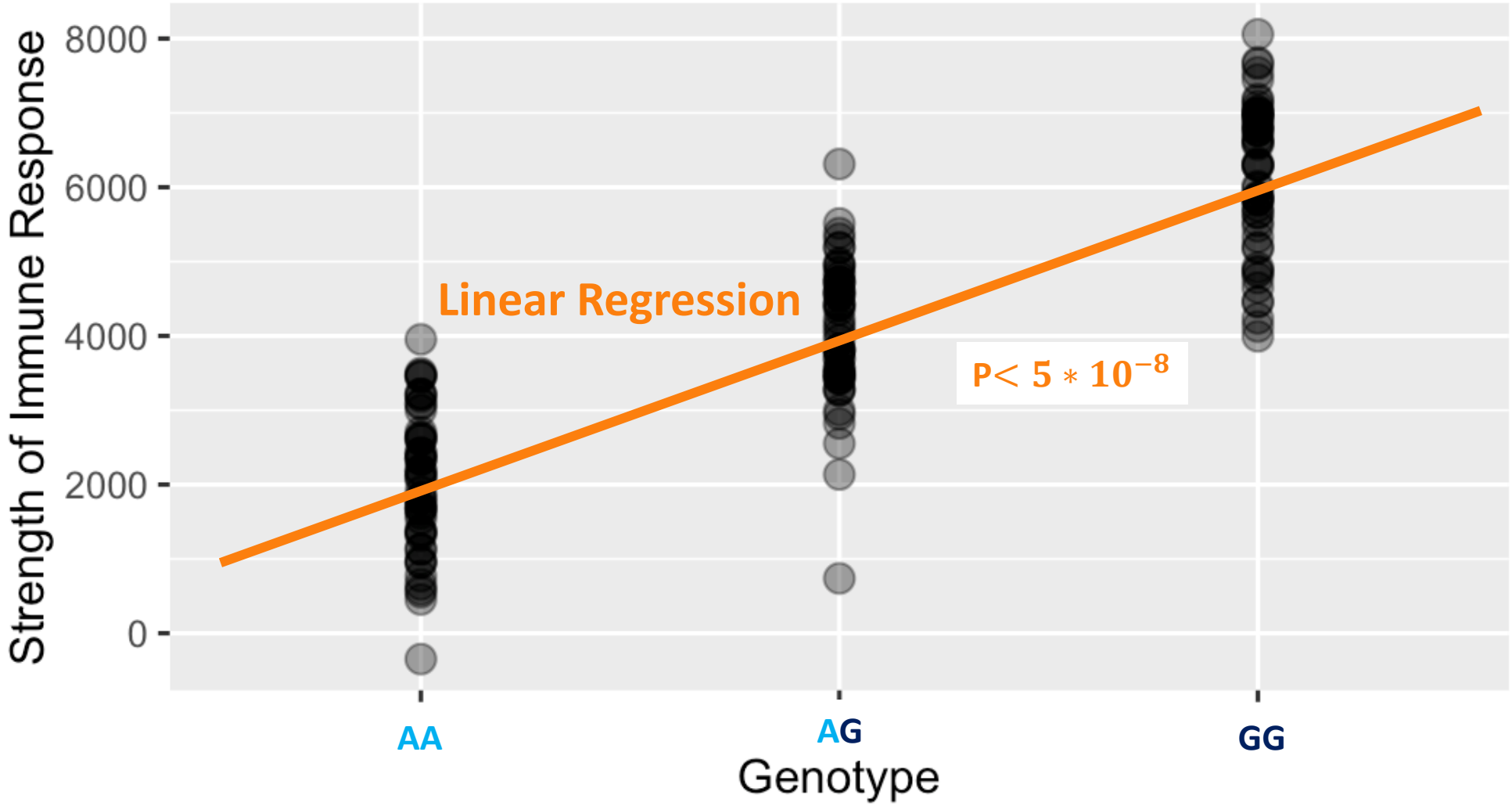


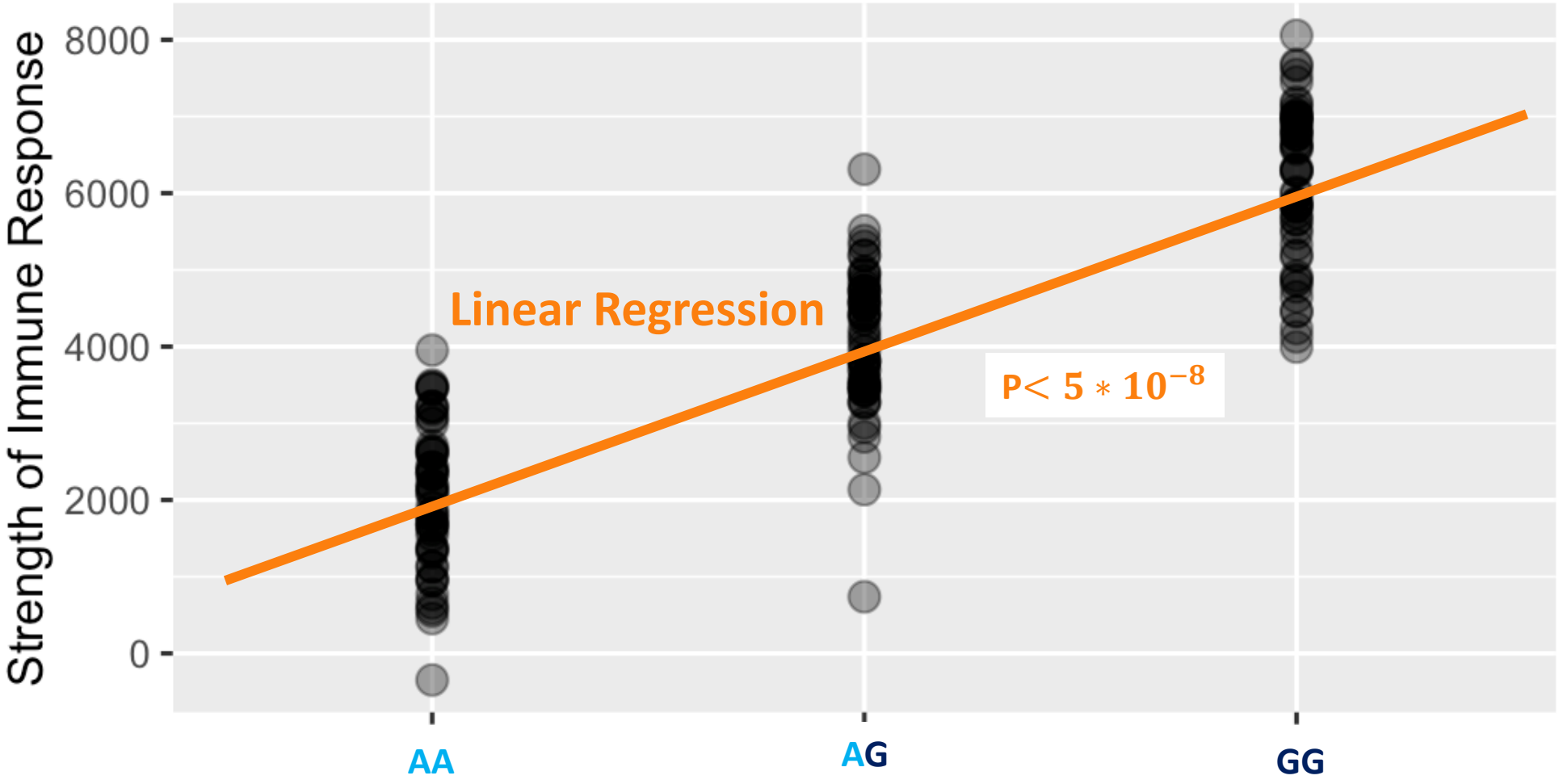
Single-Nucleotide Polymorphism (“SNP”)

GG

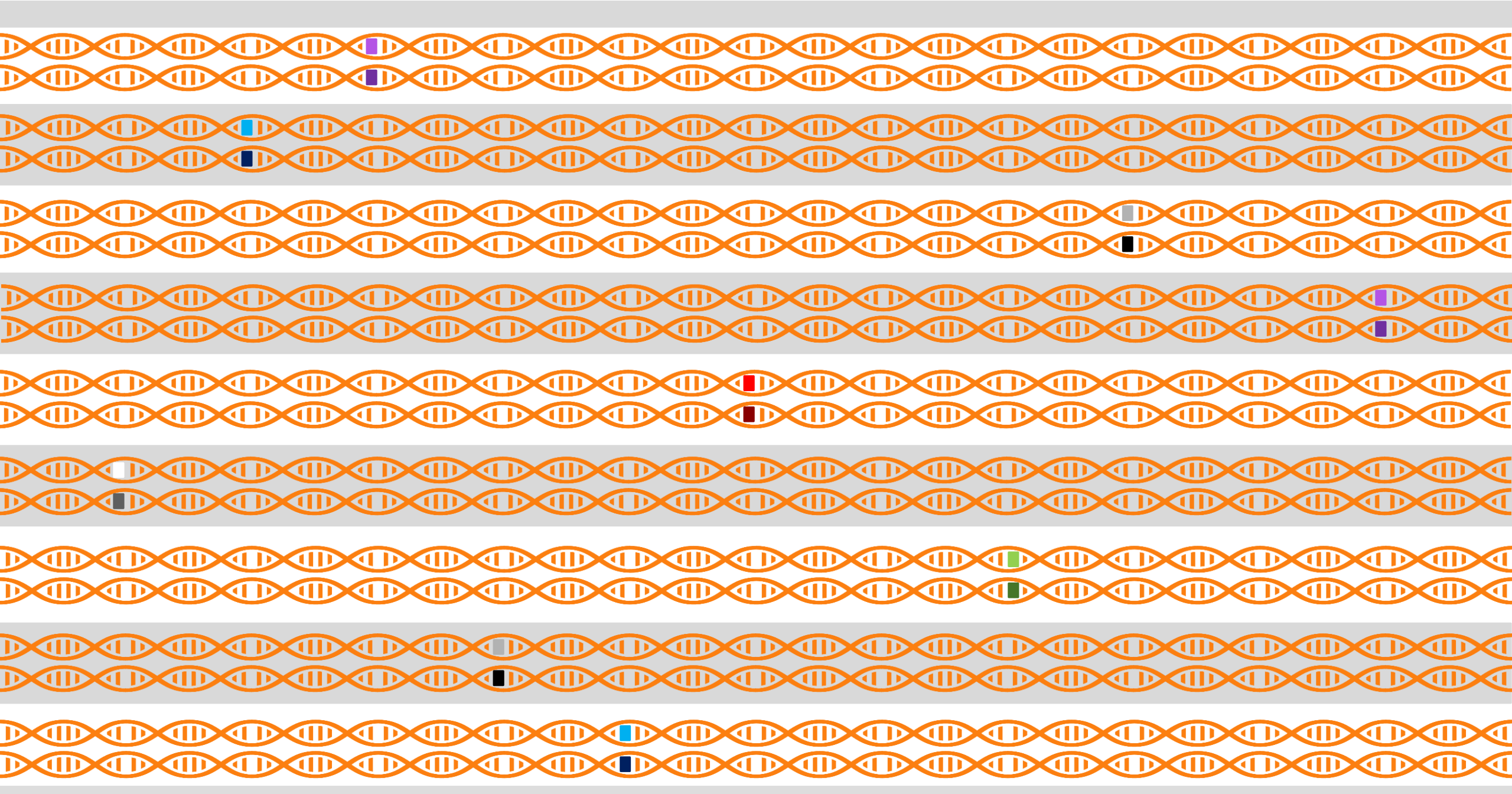

Strength of Immune Response
8000
6000
4000
2000
0
AA
AG
GG
Genotype


Strength of Immune Response
8000
6000
4000
2000
0
Linear Regression
AA
AG
GG
Genotype


Strength of Immune Response
8000
6000
4000
2000
0
Linear Regression
P< 5 ∗ 10^−8
AA
AG
GG
Genotype



$$y = \beta_0 + \boldsymbol{\beta_1} \times (\mathrm{G} \in \{0, 1, 2\}) + \beta_{\mathrm{Age}} \times (\mathrm{Age}) + \cdots + \epsilon$$

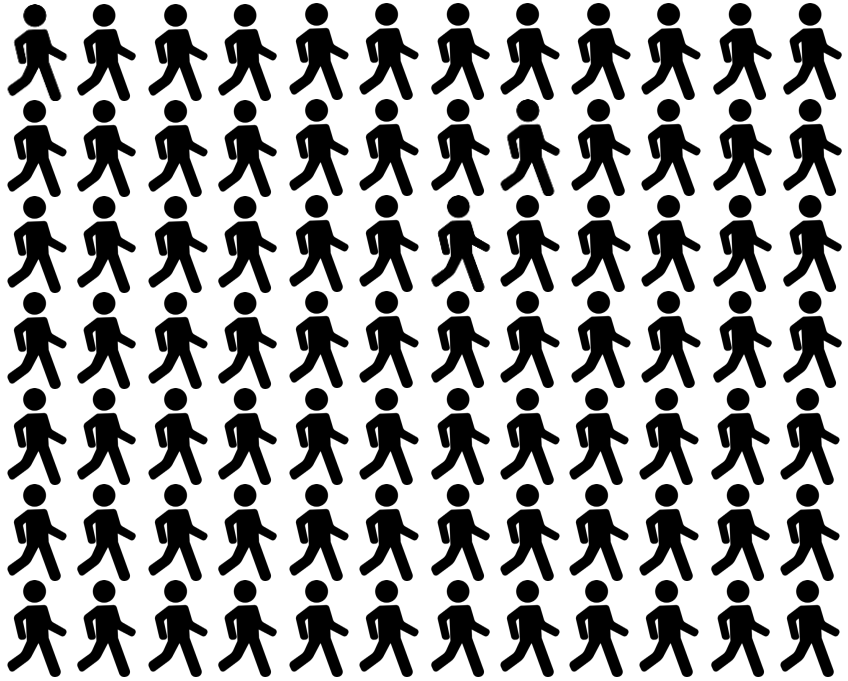


We analyzed many different SNPs throughout the genome.

Data

UK Biobank:

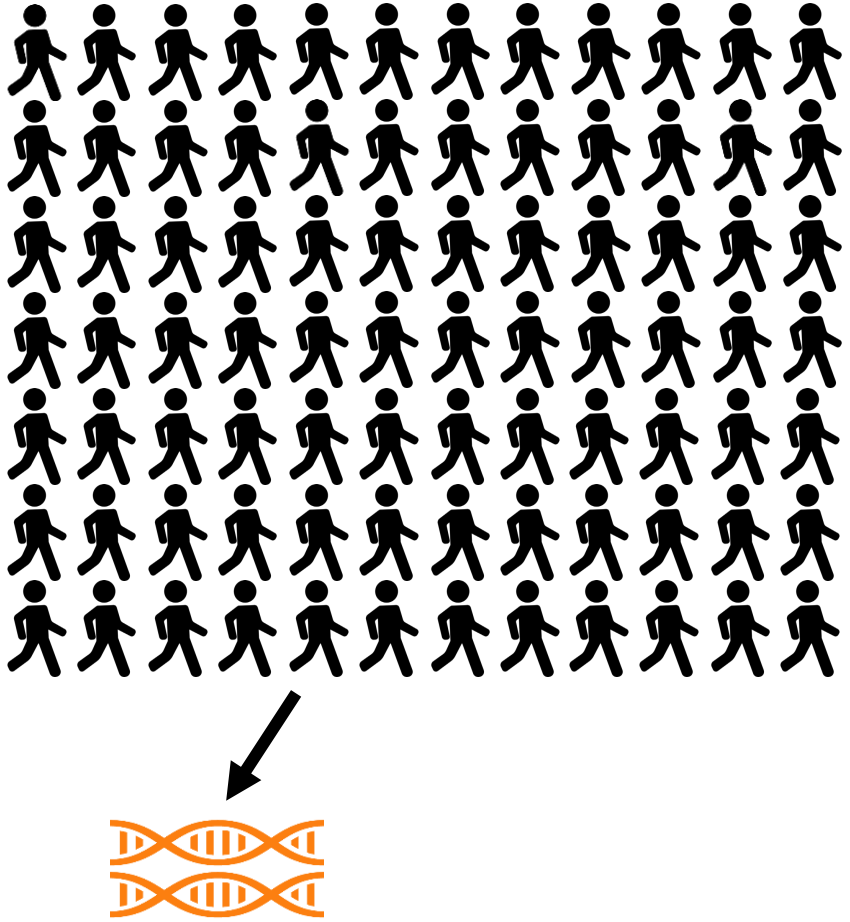
- ~500,000 individuals recruited (2006-2010)
- Aged 40-69 at time of recruitment



Data

UK Biobank:

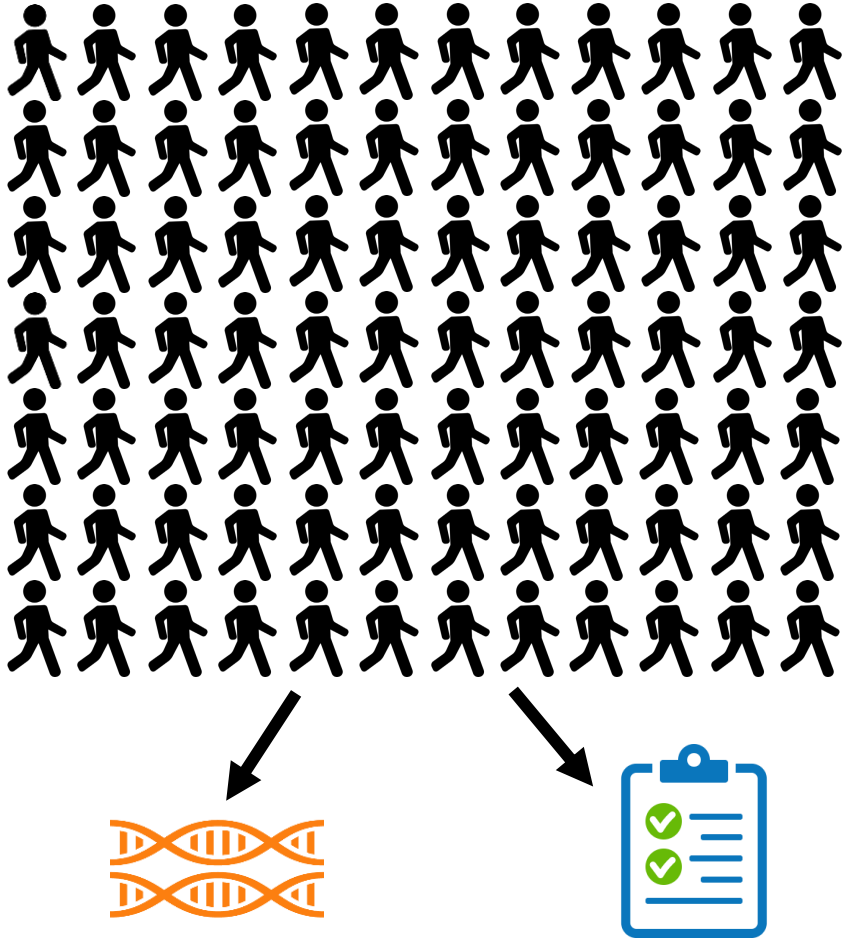
- ~500,000 individuals recruited (2006-2010)
- Aged 40-69 at time of recruitment
- ~9 million SNPs (minor allele frequency >0.01)



Data

UK Biobank:

- ~500,000 individuals recruited (2006-2010)
- Aged 40-69 at time of recruitment
- ~9 million SNPs (minor allele frequency >0.01)
- Thousands of phenotypes



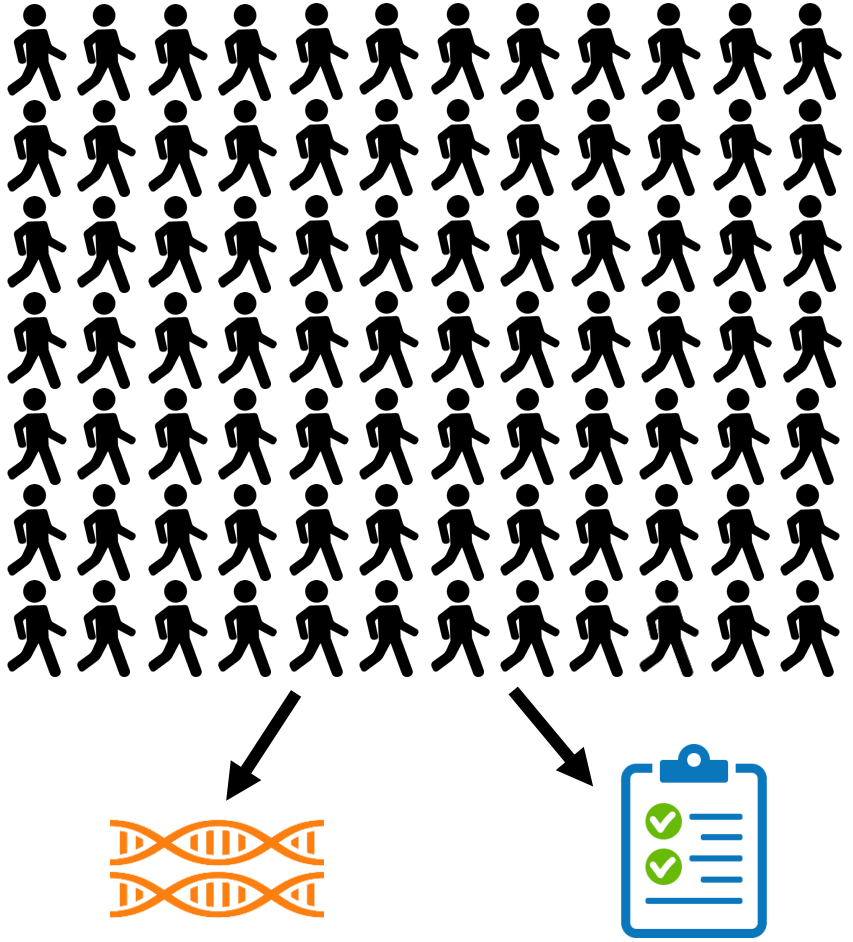
Data

UK Biobank:

- ~500,000 individuals recruited (2006-2010)
- Aged 40-69 at time of recruitment
- ~9 million SNPs (minor allele frequency >0.01)
- Thousands of phenotypes

Pilot Study:

- ~8,000 individuals of European ancestry



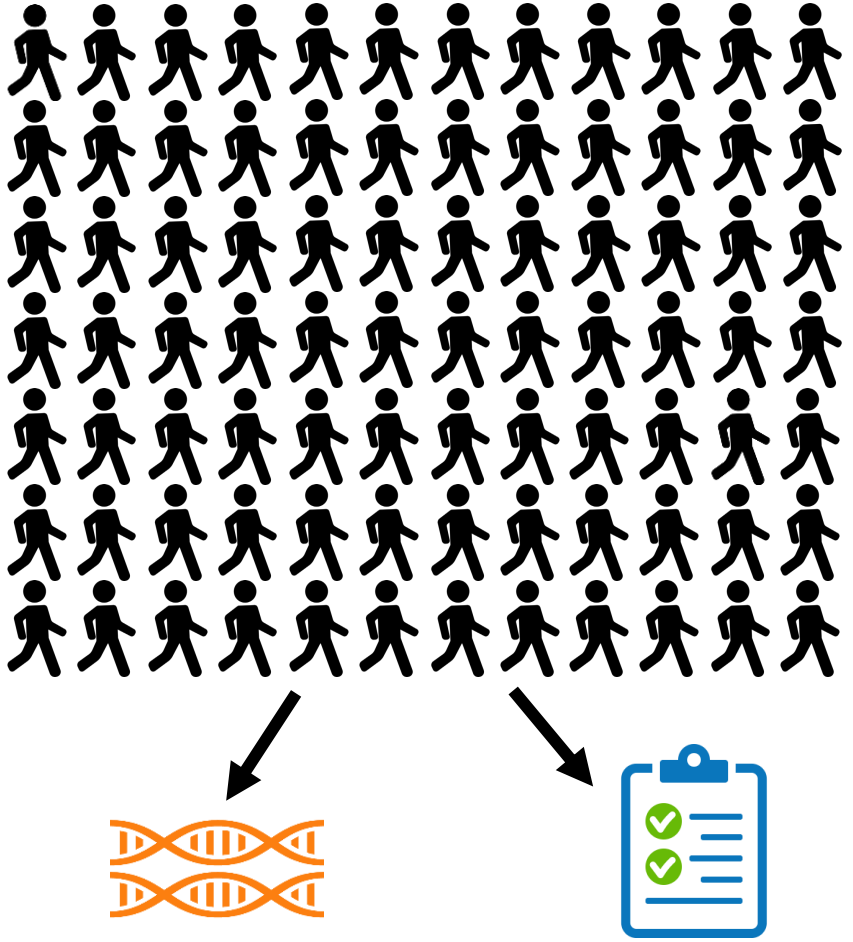
Data

UK Biobank:

- ~500,000 individuals recruited (2006-2010)
- Aged 40-69 at time of recruitment
- ~9 million SNPs (minor allele frequency >0.01)
- Thousands of phenotypes

Pilot Study:

- ~8,000 individuals of European ancestry
- Hundreds of biomarkers
  - Includes 44 antibodies
  - 22 antibodies with seroprevalence >20%



Data

12 Pathogens:

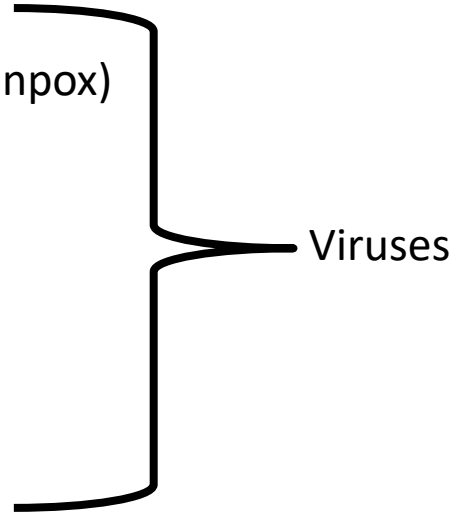
- Herpes Simplex virus-1
- Varicella Zoster Virus (Chickenpox)
- Epstein-Barr Virus
- Human Cytomegalovirus
- Human Herpesvirus-6
- Human Herpesvirus-7
- Human Polyomavirus BKV
- Human Polyomavirus JCV
- Merkel Cell Polyomavirus
- *Chlamydia trachomatis*
- *Helicobacter pylori*
- *Toxoplasma gondii*

Total of 22 antibodies against these pathogens
Each antigen's seroprevalence >20%

Data

12 Pathogens:

- Herpes Simplex virus-1
- Varicella Zoster Virus (Chickenpox)
- Epstein-Barr Virus
- Human Cytomegalovirus
- Human Herpesvirus-6
- Human Herpesvirus-7
- Human Polyomavirus BKV
- Human Polyomavirus JCV
- Merkel Cell Polyomavirus
- *Chlamydia trachomatis*
- *Helicobacter pylori*
- *Toxoplasma gondii*

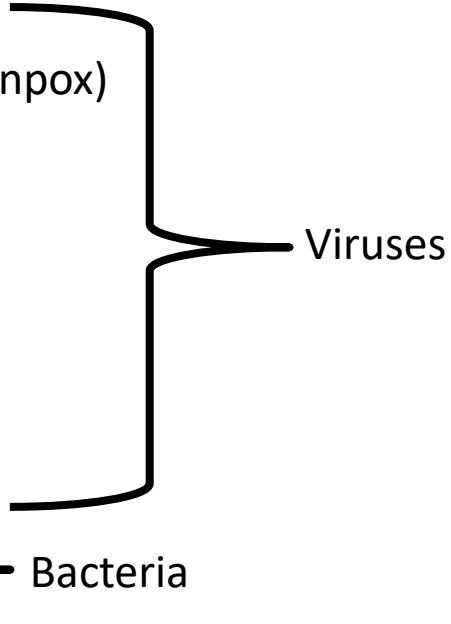


Total of 22 antibodies against these pathogens
Each antigen's seroprevalence >20%

Data

12 Pathogens:

- Herpes Simplex virus-1
- Varicella Zoster Virus (Chickenpox)
- Epstein-Barr Virus
- Human Cytomegalovirus
- Human Herpesvirus-6
- Human Herpesvirus-7
- Human Polyomavirus BKV
- Human Polyomavirus JCV
- Merkel Cell Polyomavirus
- *Chlamydia trachomatis*
- *Helicobacter pylori*
- *Toxoplasma gondii*

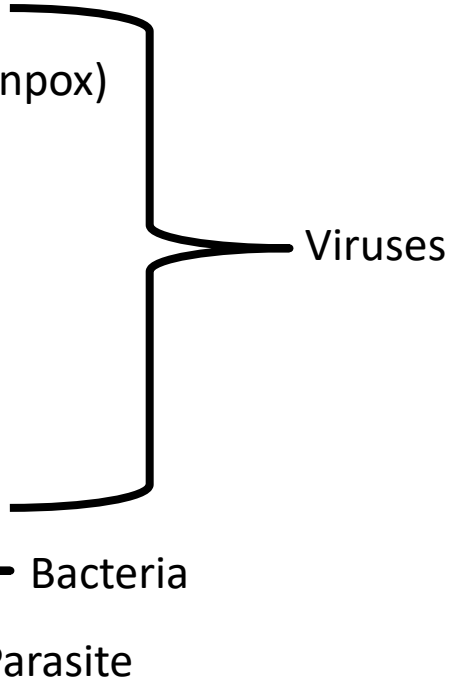


Total of 22 antibodies against these pathogens
Each antigen's seroprevalence >20%

Data

12 Pathogens:

- Herpes Simplex virus-1
- Varicella Zoster Virus (Chickenpox)
- Epstein-Barr Virus
- Human Cytomegalovirus
- Human Herpesvirus-6
- Human Herpesvirus-7
- Human Polyomavirus BKV
- Human Polyomavirus JCV
- Merkel Cell Polyomavirus
- *Chlamydia trachomatis*
- *Helicobacter pylori*
- *Toxoplasma gondii*

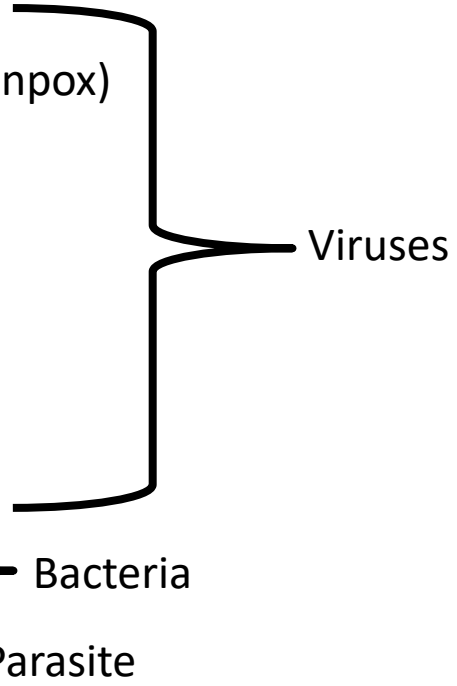


Total of 22 antibodies against these pathogens
Each antigen's seroprevalence >20%

Data

12 Pathogens:

- Herpes Simplex virus-1
- Varicella Zoster Virus (Chickenpox)
- Epstein-Barr Virus
- Human Cytomegalovirus
- Human Herpesvirus-6
- Human Herpesvirus-7
- Human Polyomavirus BKV
- Human Polyomavirus JCV
- Merkel Cell Polyomavirus
- *Chlamydia trachomatis*
- *Helicobacter pylori*
- *Toxoplasma gondii*



Total of 22 antibodies against these pathogens
Each antigen's seroprevalence >20%

These pathogens are risk factors for cancer and cardiovascular or neurodegenerative diseases.

Data

Covariates for all regression models:

- Age
- Sex
- Body mass index
- Genotyping array
- Antigen assay date
- Antigen quality control indicator

Data

Covariates for all regression models:

- Age
- Sex
- Body mass index
- Genotyping array
- Antigen assay date
- Antigen quality control indicator

Additional covariates for sexually-transmitted pathogens:

- Categorized number of lifetime sexual partners

Data

Covariates for all regression models:

- Age
- Sex
- Body mass index
- Genotyping array
- Antigen assay date
- Antigen quality control indicator

Additional covariates for sexually-transmitted pathogens:

- Categorized number of lifetime sexual partners

$$y = \beta_0 + \boldsymbol{\beta_1} \times (\mathrm{G} \in \{0, 1, 2\}) + \beta_{\mathrm{Age}} \times (\mathrm{Age}) + \cdots + \epsilon$$

Data

Covariates for all regression models:

- Age
- Sex
- Body mass index
- Genotyping array
- Antigen assay date
- Antigen quality control indicator

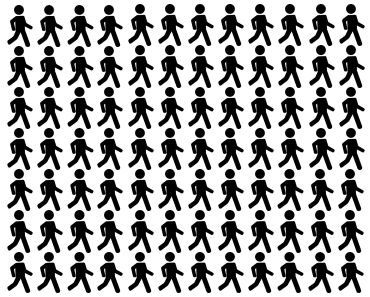
Additional covariates for sexually-transmitted pathogens:

- Categorized number of lifetime sexual partners

$$y = \beta_0 + \beta_1 \times (\mathrm{G} \in \{0, 1, 2\}) + \underbrace{\beta_{\mathrm{Age}} \times (\mathrm{Age}) + \cdots}_{\text{Covariates}} + \epsilon$$

Methods - Summary

UK Biobank Data

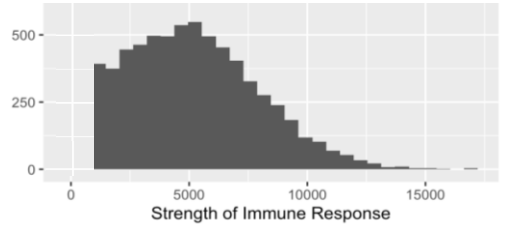


N=8,000

22 common antibodies (12 pathogens)

Genome-Wide Association Study

$$y = \beta_0 + \boldsymbol{\beta_1} \times (\mathrm{G} \in \{0, 1, 2\}) + \cdots + \epsilon$$



Results

- Found associated SNPs for **17 antigens** against **10 pathogens**

Results

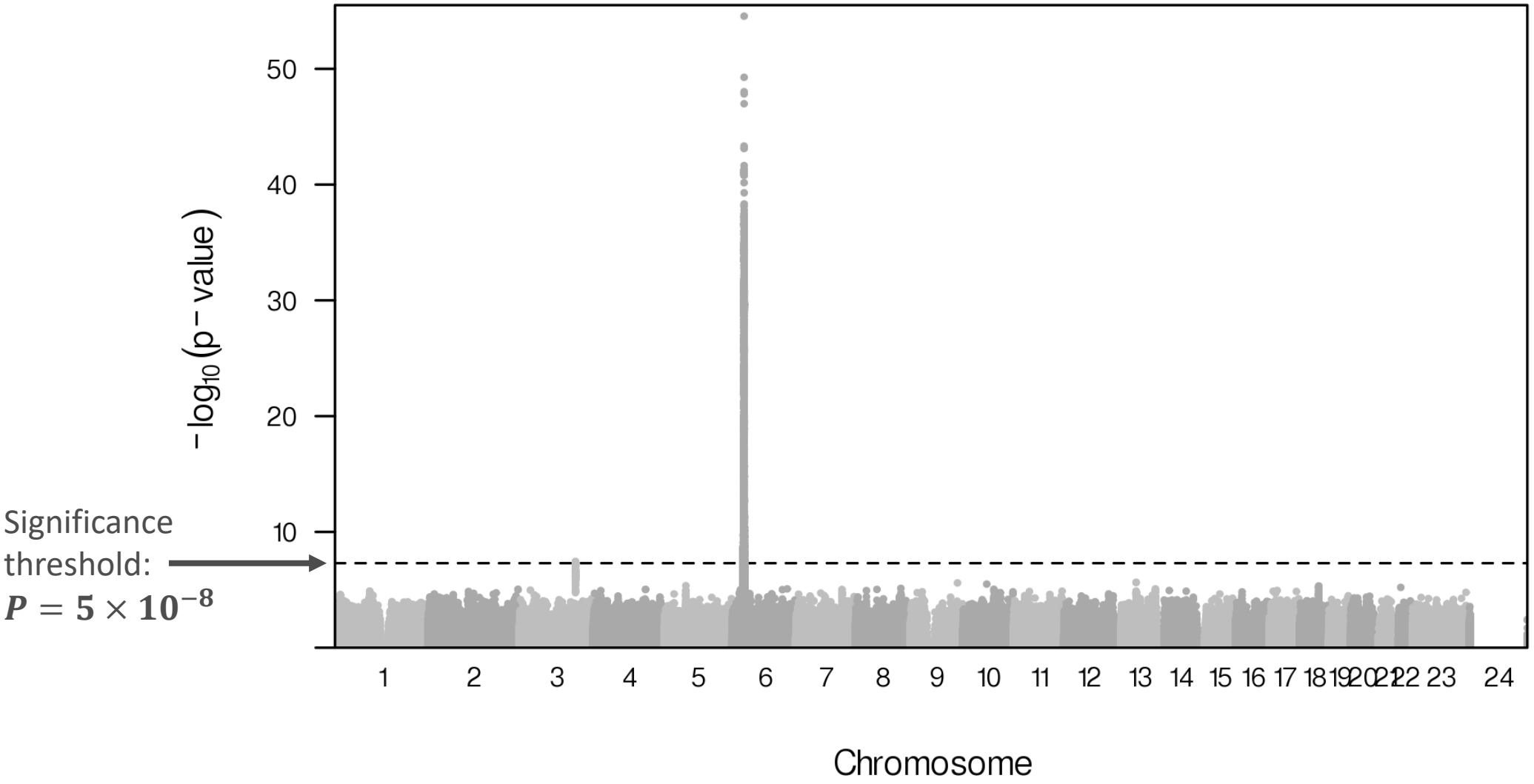
- Found associated SNPs for **17 antigens** against **10 pathogens**
- **99** independent ($R^2 < 0.05$) SNPs ($P < 5 \times 10^{-8}$)

Results

- Found associated SNPs for **17 antigens** against **10 pathogens**
- **99** independent ($R^2 < 0.05$) SNPs ($P < 5 \times 10^{-8}$)
- **47** of these SNPs lie within genes (45 distinct genes)
  - Many genes associated with autoimmunity, inflammation, and vaccine response

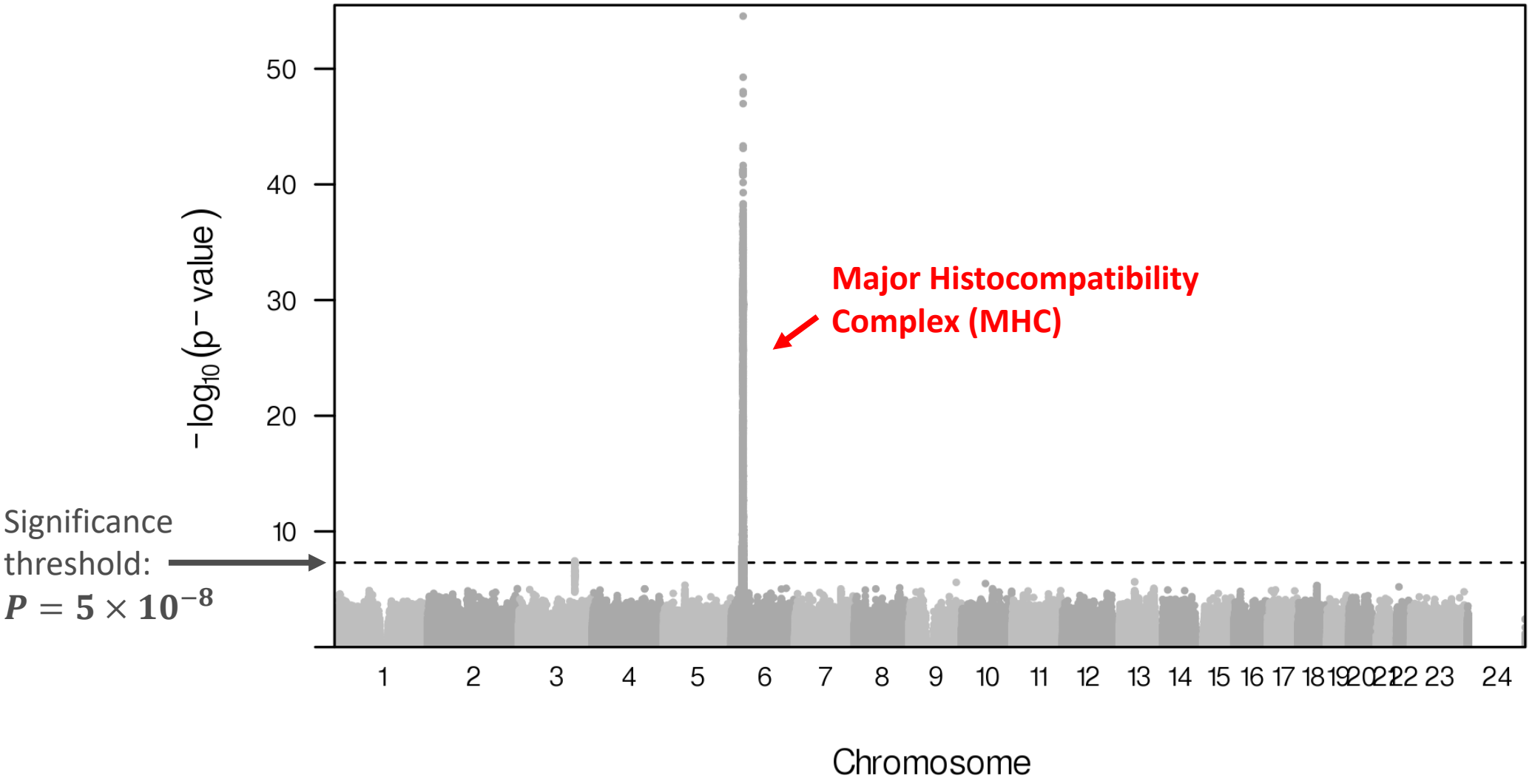
EBNA-1 Antigen for **N = 6845**

Epstein-Barr Virus

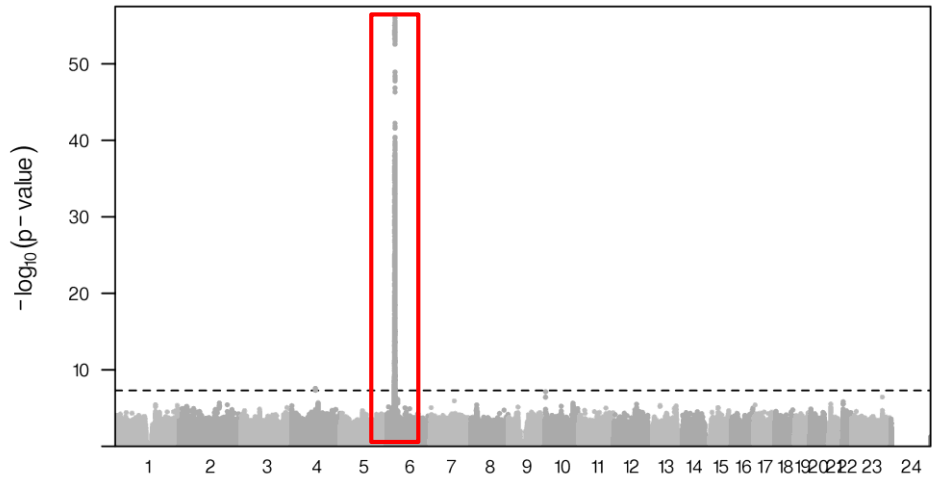


EBNA-1 Antigen for **N = 6845**

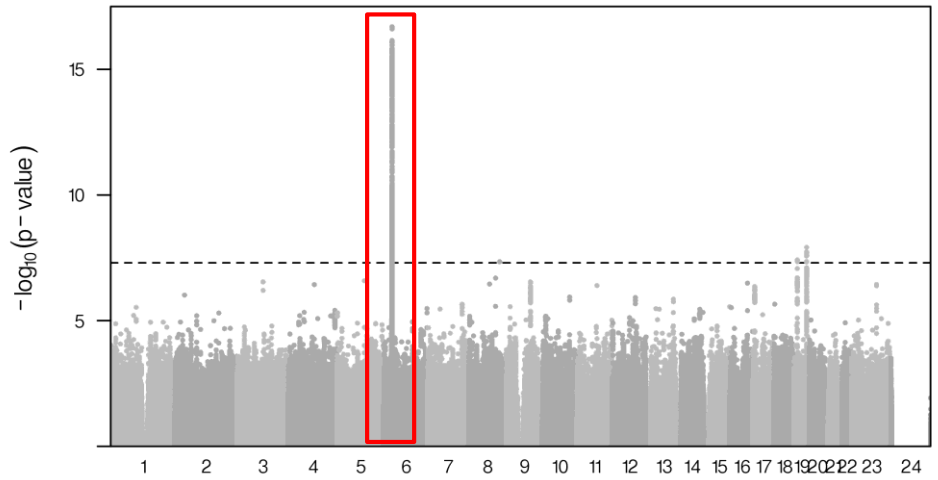
Epstein-Barr Virus



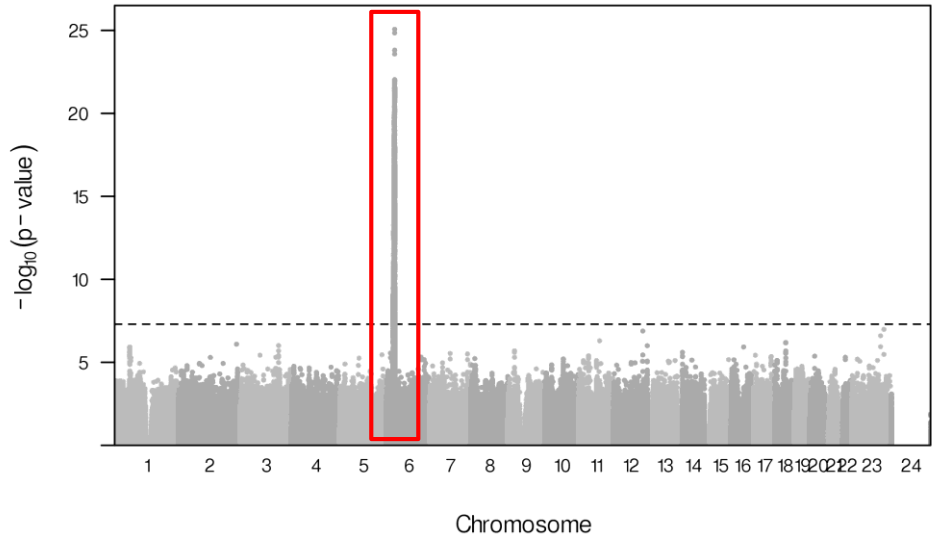
ZEBRA Antigen for **Epstein-Barr Virus** **N = 7033**



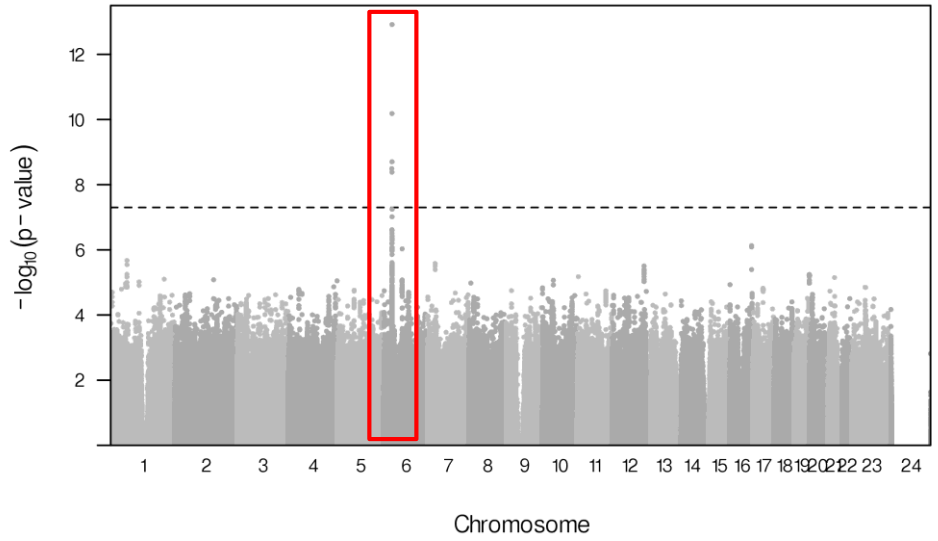
JC VP1 Antigen for **Human Polyomavirus JCV** **N = 4367**



gE/gI Antigen for **Varicella Zoster Virus** **N = 7124**



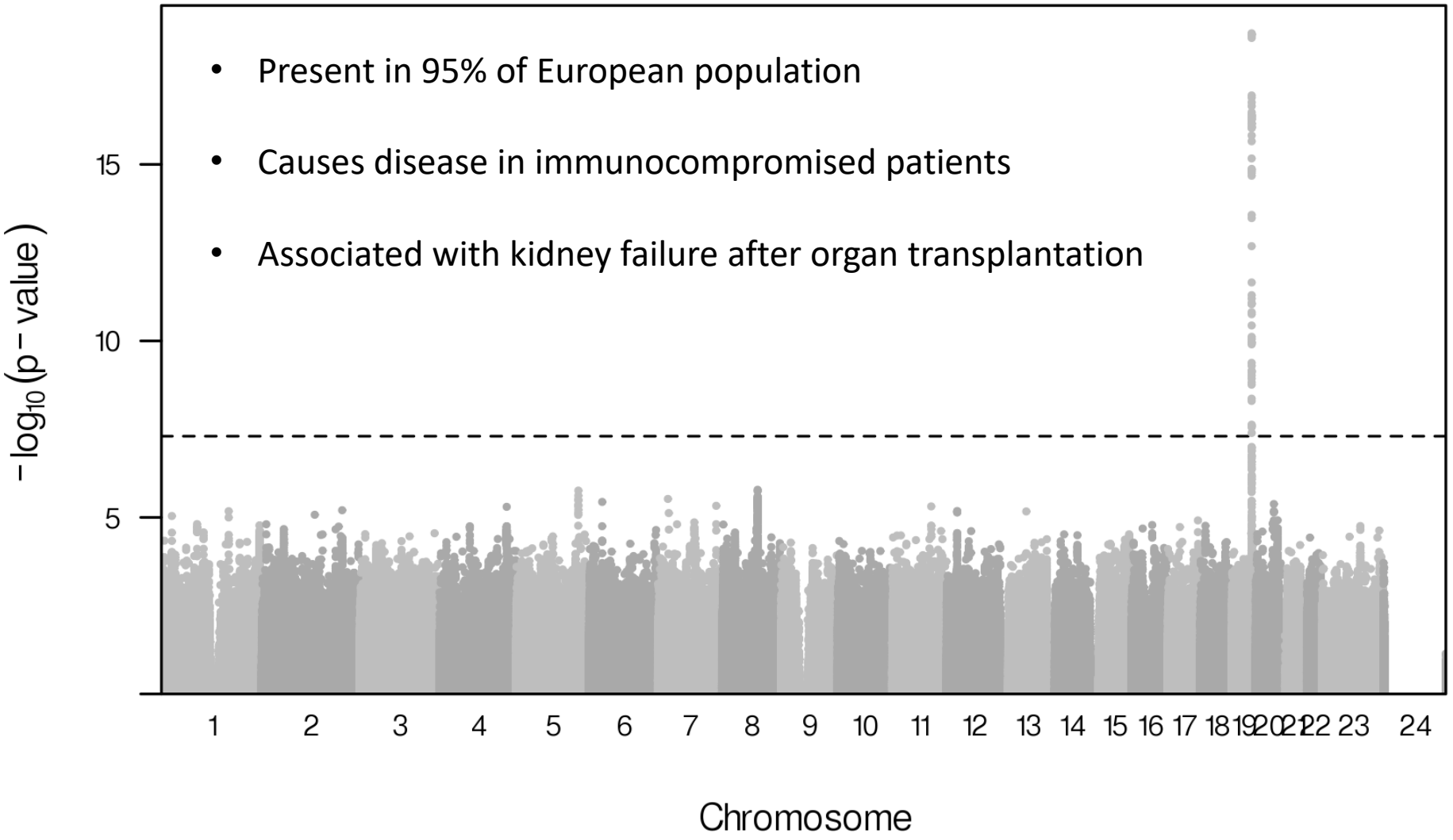
pp52 Antigen for **Human Cytomegalovirus** **N = 4881**



We identified **45** independent SNPs in the Major Histocompatibility Complex

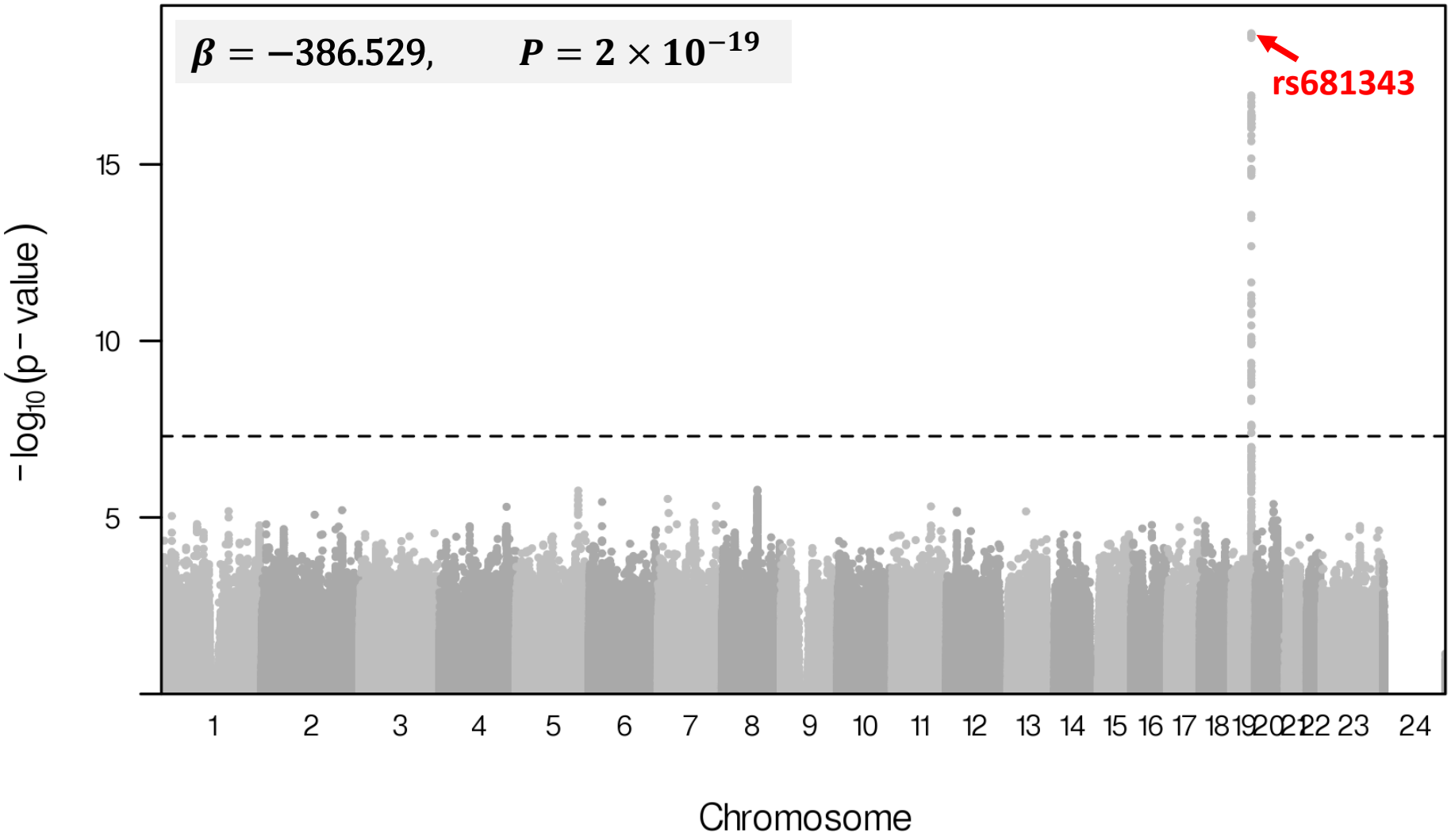
BK VP1 Antigen for **N = 7349**

Human Polyomavirus BKV (PBKV)



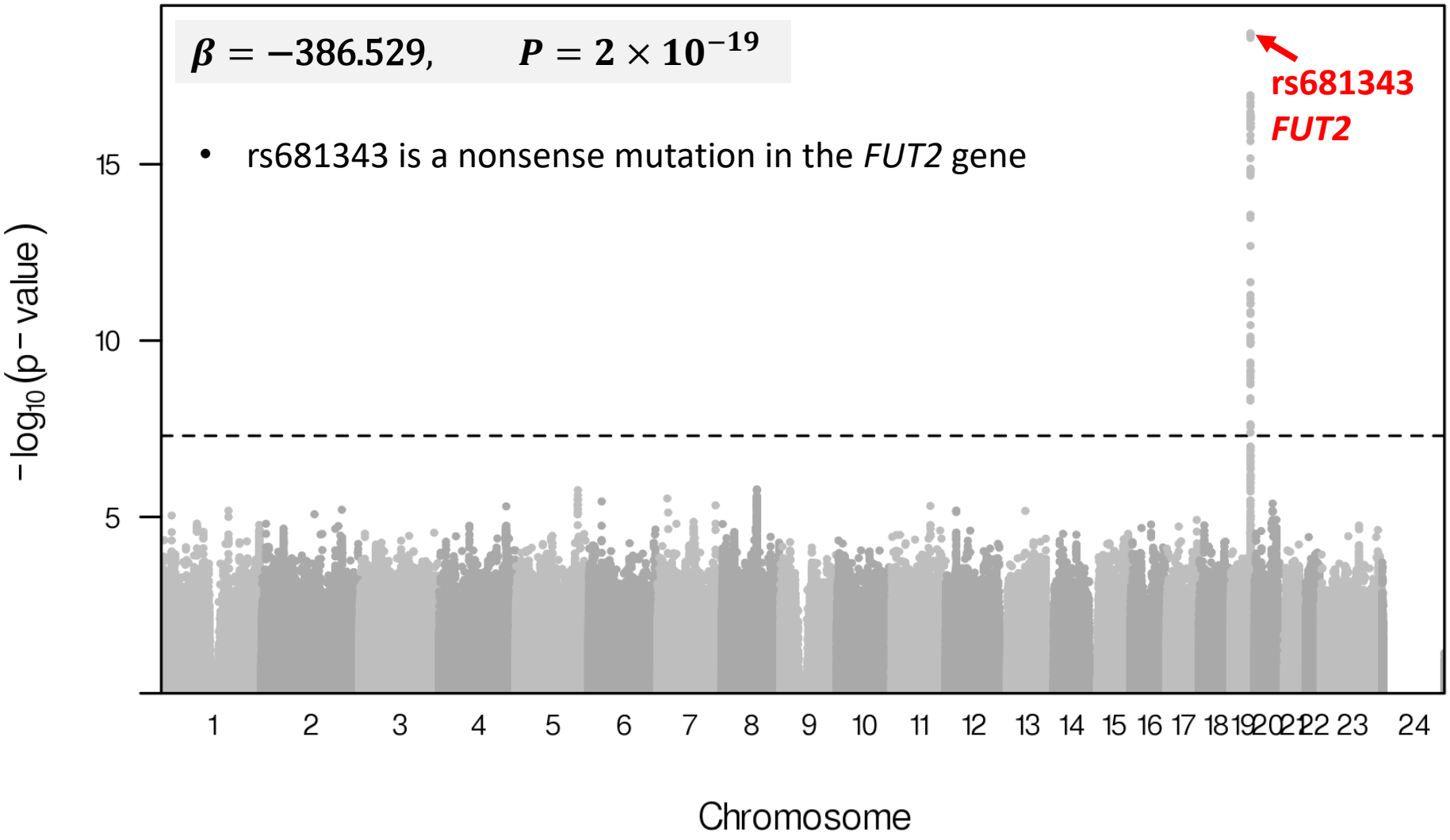
BK VP1 Antigen for **N = 7349**

Human Polyomavirus BKV (PBKV)



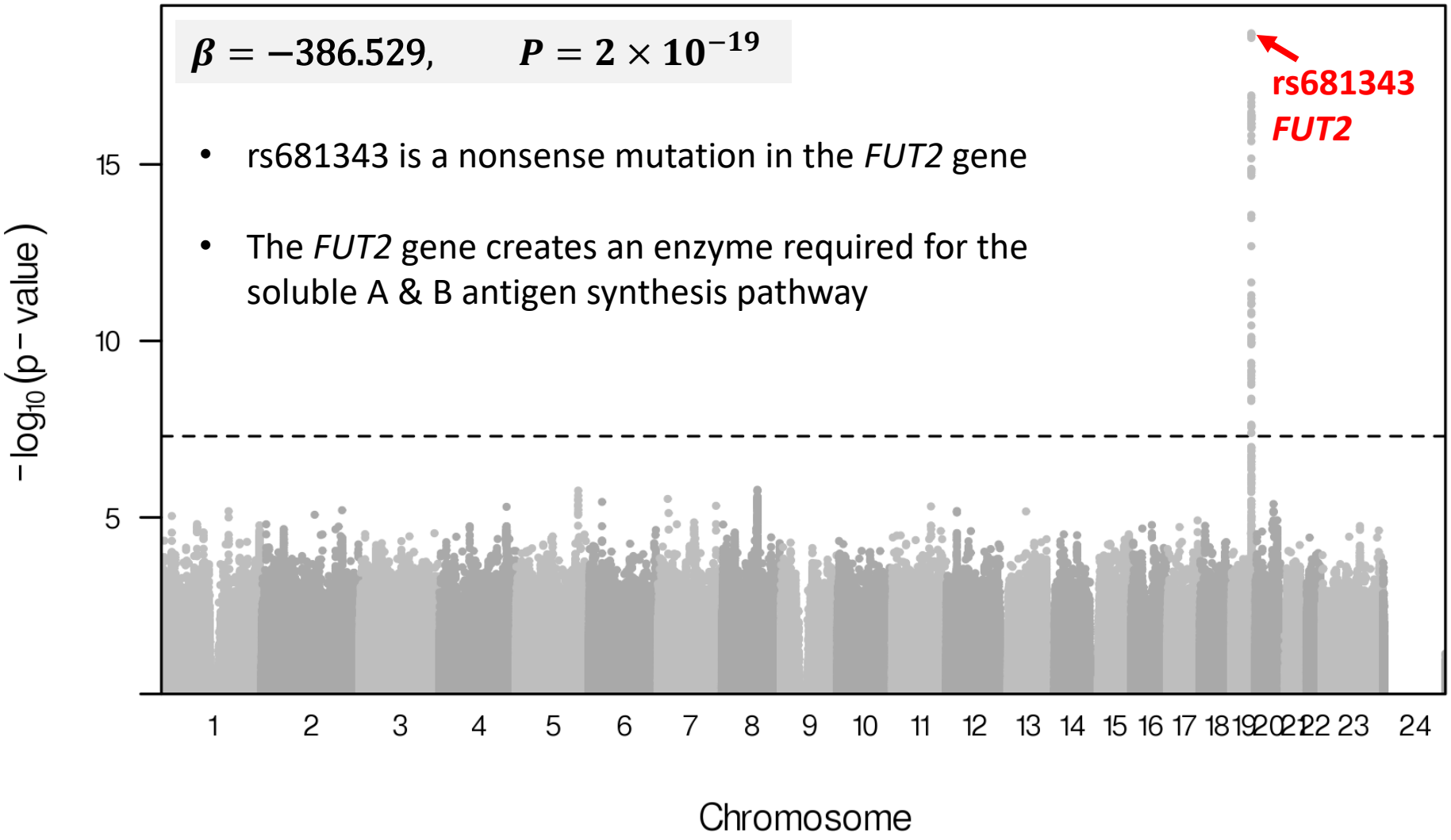
BK VP1 Antigen for **N = 7349**

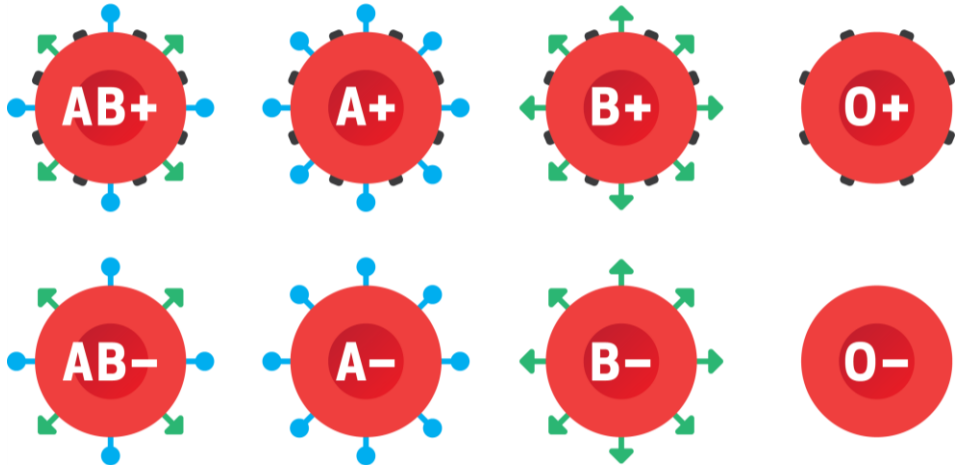
Human Polyomavirus BKV (PBKV)



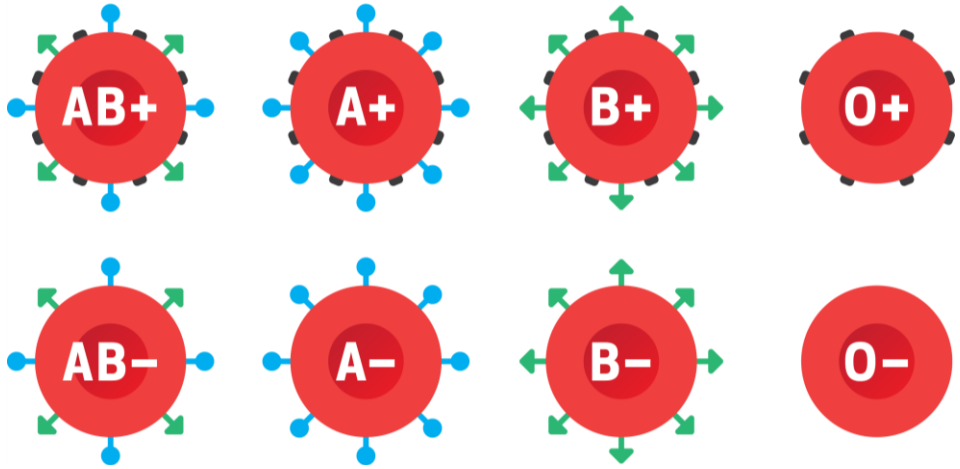
BK VP1 Antigen for **N = 7349**

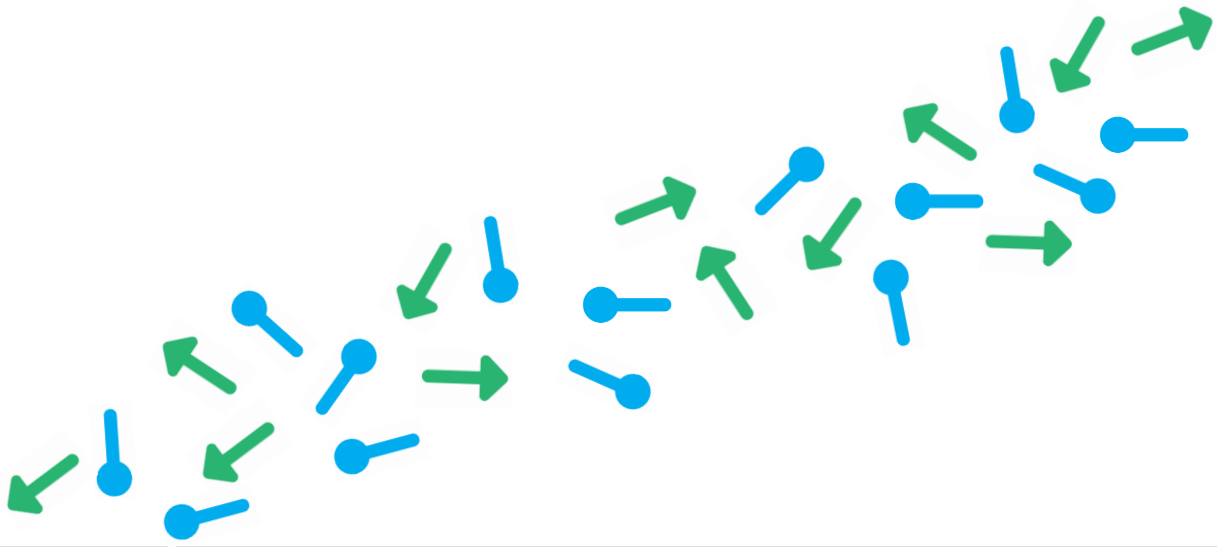
Human Polyomavirus BKV (PBKV)



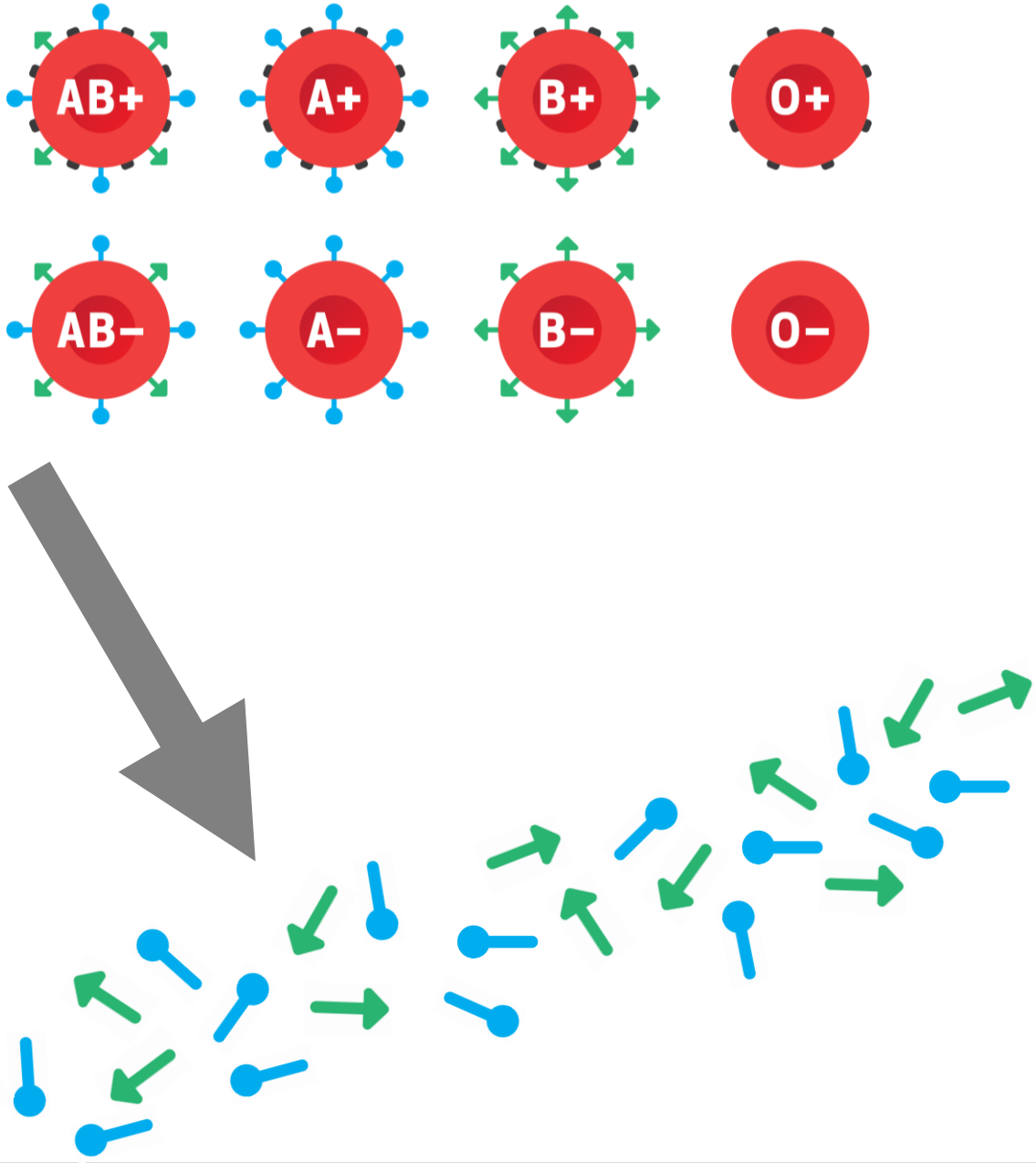


Graphics adapted from blood.ca

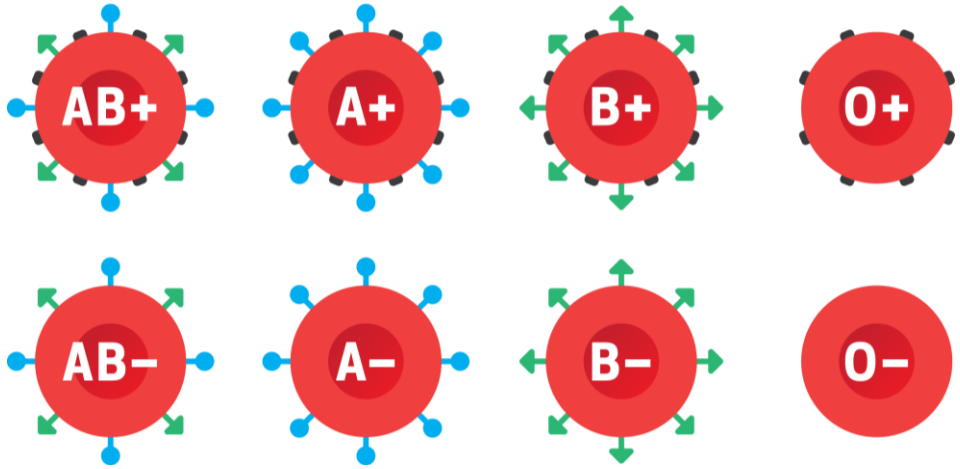




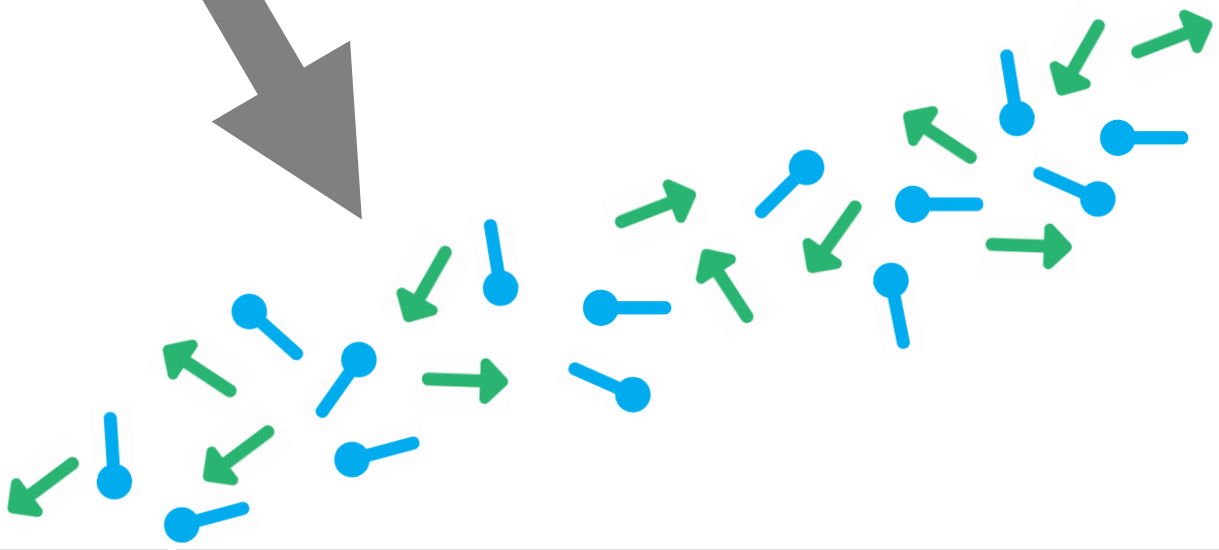
Graphics adapted from blood.ca



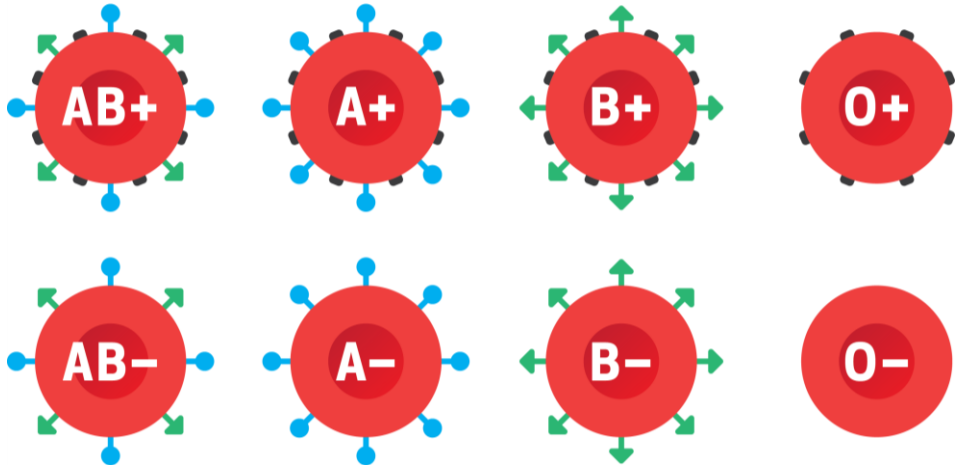
Graphics adapted from blood.ca



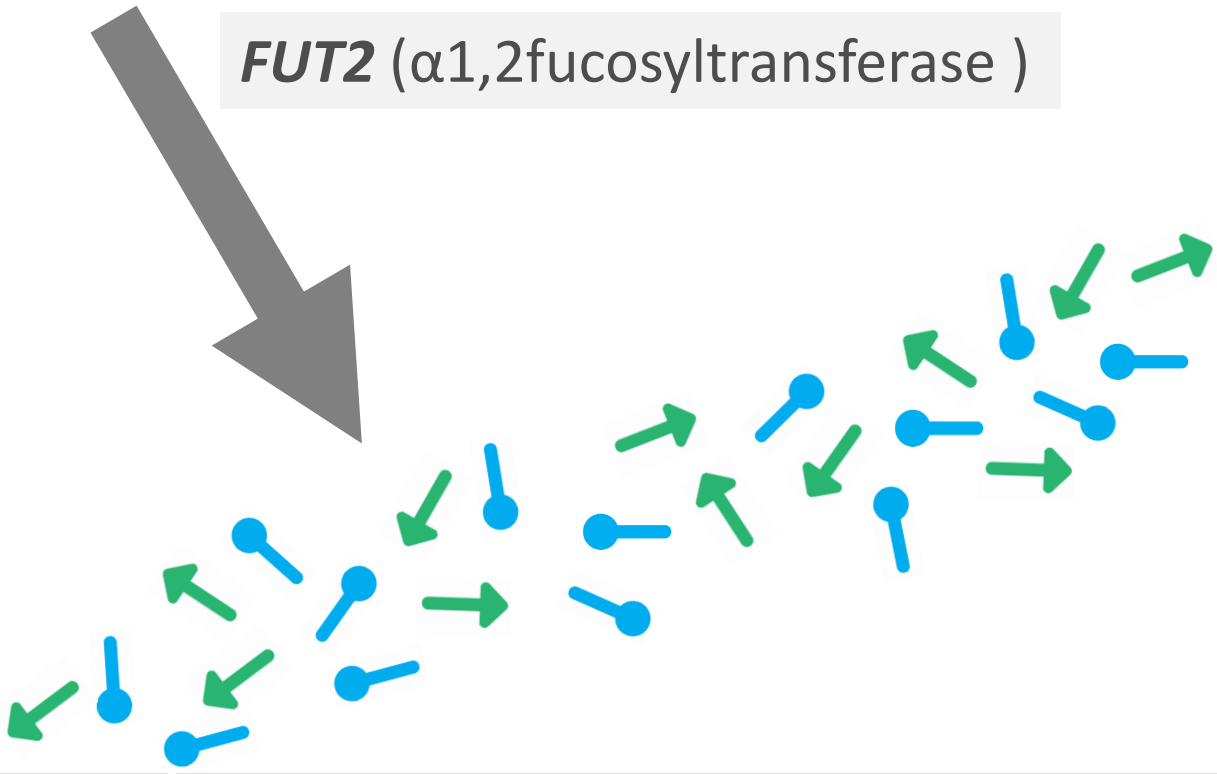
FUT2 (α1,2fucosyltransferase)



Graphics adapted from blood.ca



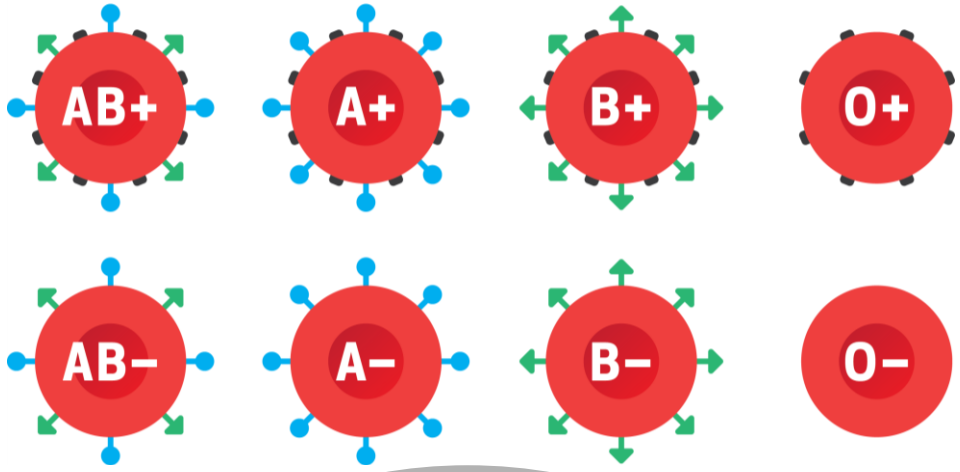
FUT2 (α1,2fucosyltransferase)



20% of Europeans are homozygous for a **recessive nonsense mutation** (FUT2-/-) which results in a **nonfunctional enzyme.**

PMID:7876235

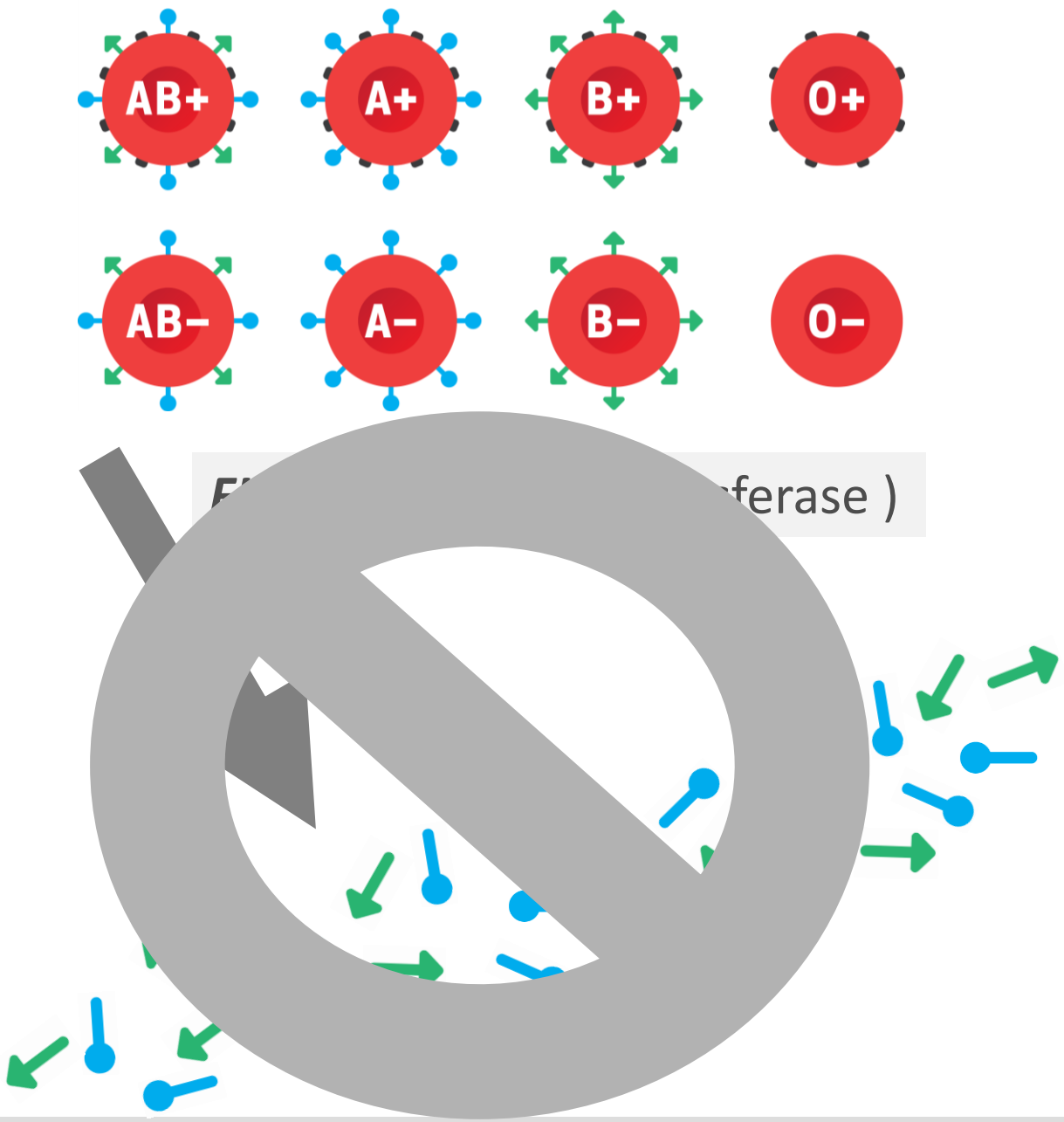
Graphics adapted from blood.ca



20% of Europeans are homozygous for a **recessive nonsense mutation** (FUT2-/-) which results in a **nonfunctional enzyme.**

PMID:7876235

Graphics adapted from blood.ca



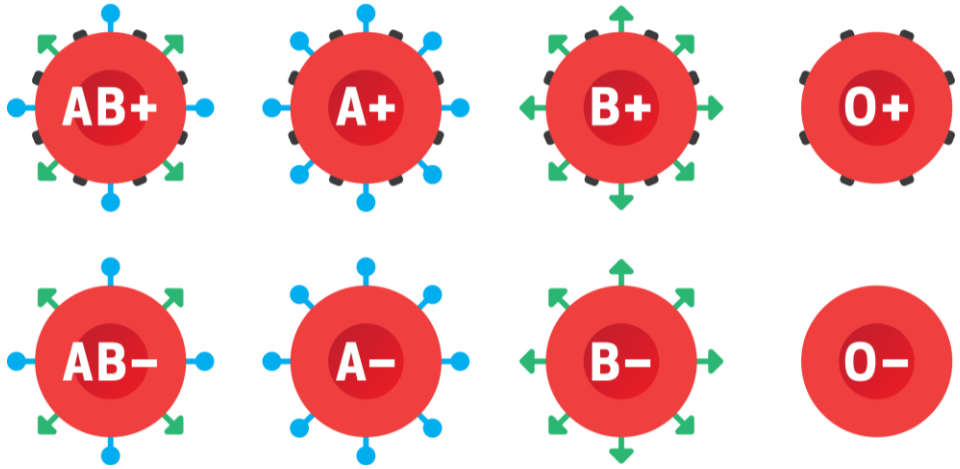
20% of Europeans are homozygous for a **recessive nonsense mutation** (FUT2-/-) which results in a **nonfunctional enzyme.**

PMID:7876235

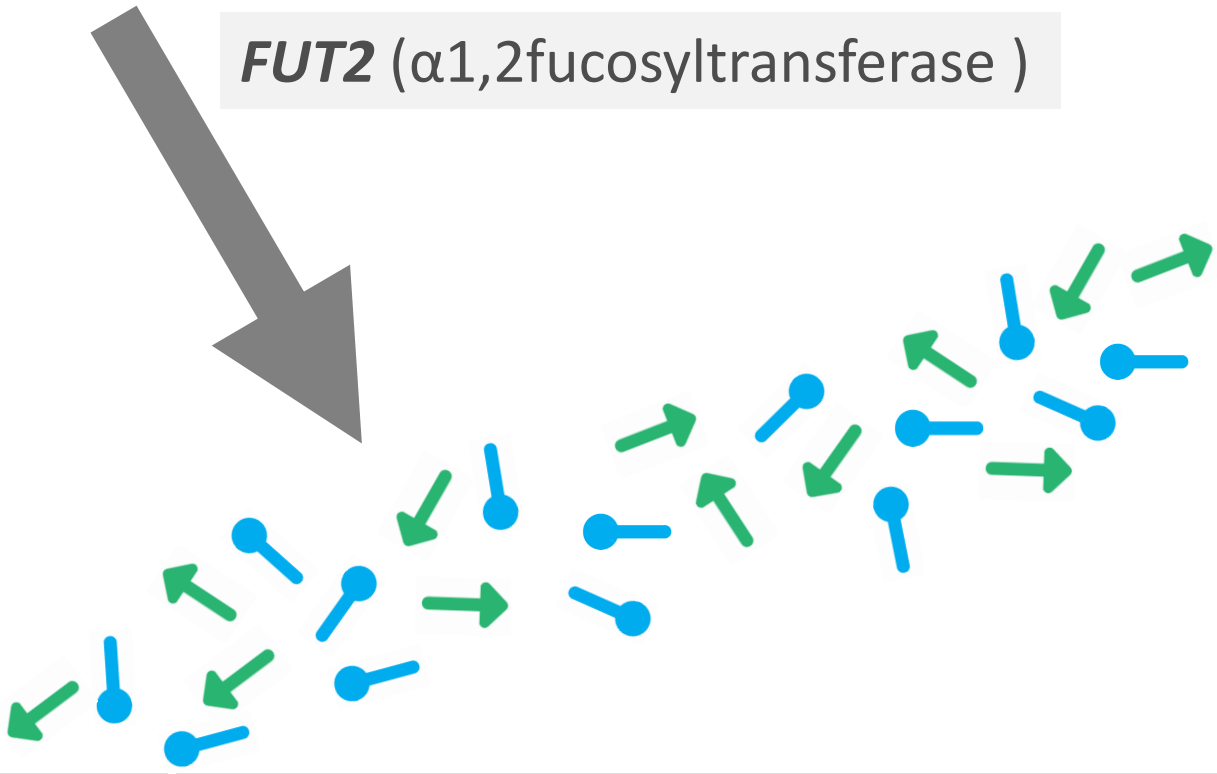
These individuals are **highly resistant** to **norovirus** infection.

PMID: 16306606, 16973373

Graphics adapted from blood.ca



FUT2 (α1,2fucosyltransferase)



20% of Europeans are homozygous for a **recessive nonsense mutation** (FUT2-/-) which results in a **nonfunctional enzyme.**

PMID:7876235

These individuals are **highly resistant** to **norovirus** infection.

PMID: 16306606, 16973373

The association of *FUT2* with Polyomavirus BKV suggests that **secreted histo-blood group antigens could play a role in polyomavirus response.**

Graphics adapted from blood.ca

Conclusions

Different genes are associated with different antibodies.

Conclusions

Different genes are associated with different antibodies.

Many of these genes are also associated with autoimmunity, inflammation, and vaccine response.

Conclusions

Different genes are associated with different antibodies.

Many of these genes are also associated with autoimmunity, inflammation, and vaccine response.

Some of these genes could elucidate how the immune system interacts with specific pathogens.

Future Directions

Genetic basis of pathogen-induced cancers & immune response

Future Directions

Genetic basis of pathogen-induced cancers & immune response

Biological function of immune-response-associated genes

Future Directions

Genetic basis of pathogen-induced cancers & immune response

Biological function of immune-response-associated genes

Personalized prevention (vaccines) & intervention (drug targets)

Thank you!

- Dr. Linda Kachuri & Dr. Sara Rashkin
- Dr. John Witte & The Witte Lab
- UCSF SRTP: Julia Clark & Dr. D'Anne Duncan
- Amgen Foundation

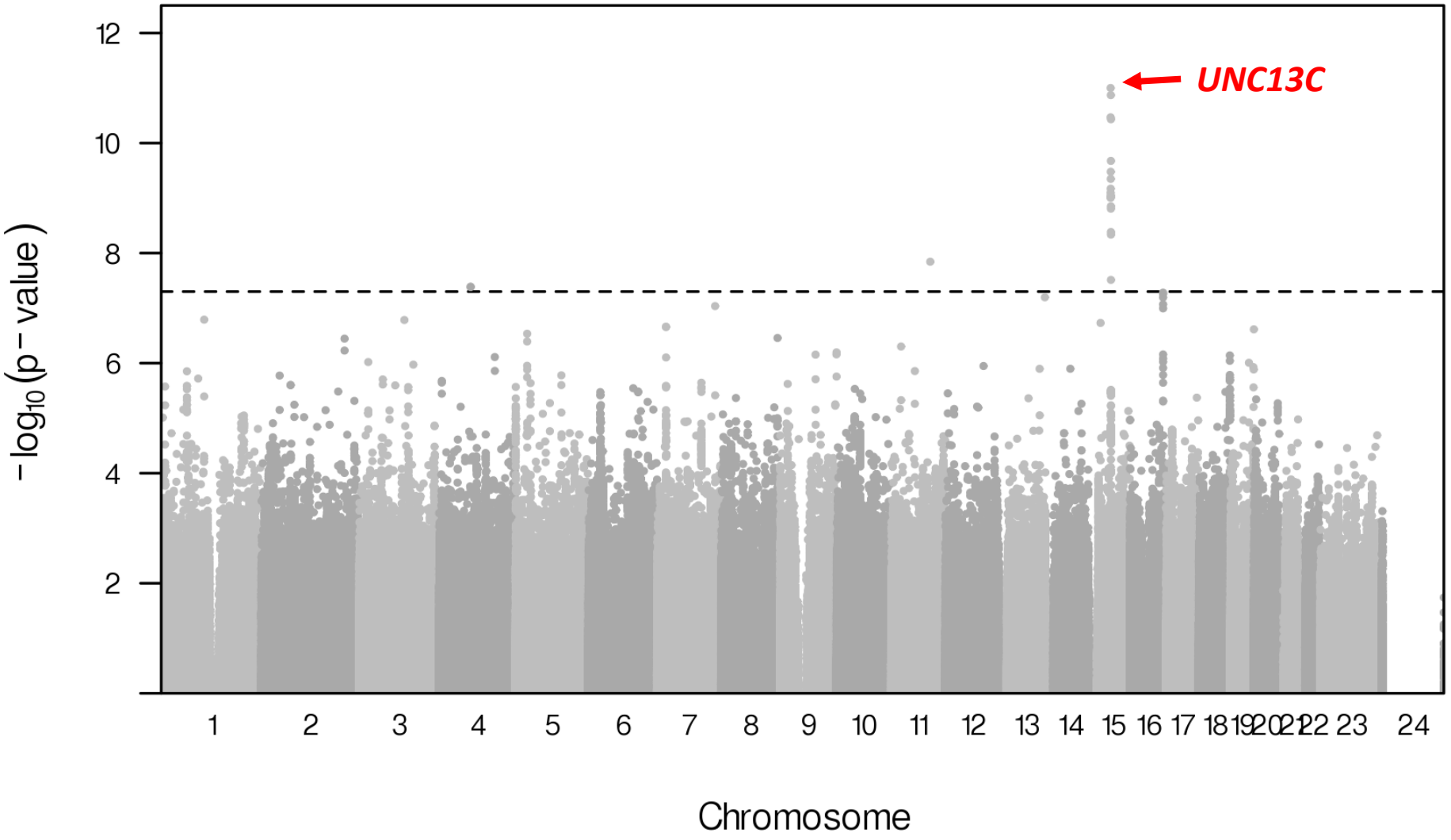
UCSF Department of Epidemiology & Biostatistics

AMGEN® Foundation

Amgen Scholars Program

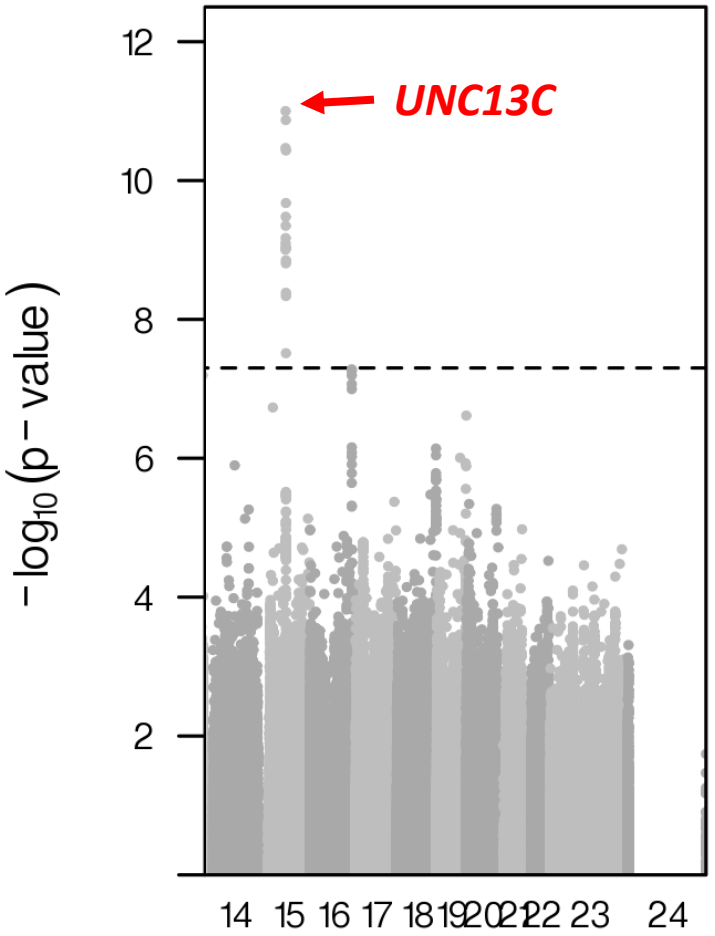
IE1B Antigen for **N = 6095**

Human Herpesvirus-6



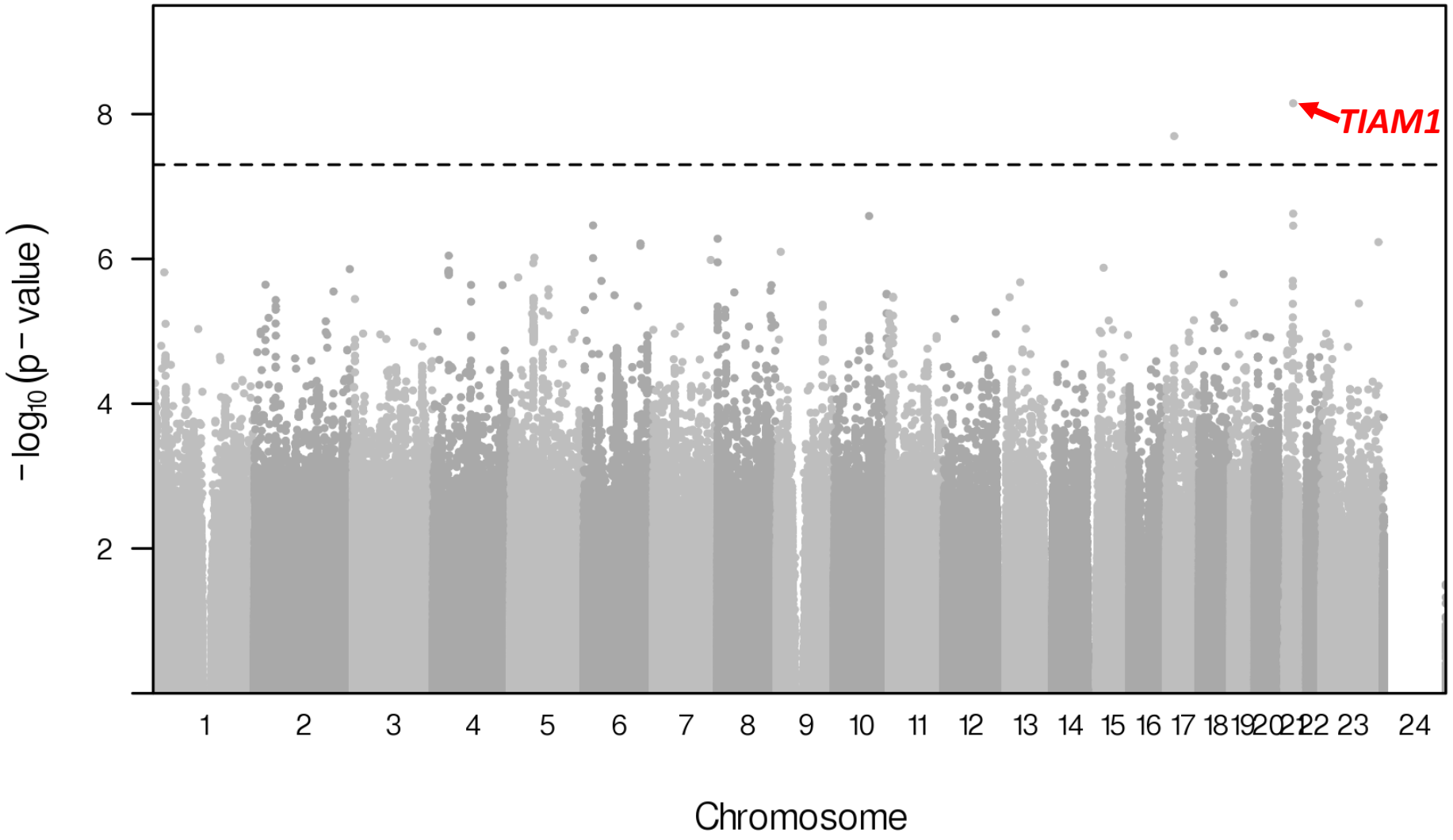
rs118162566 – *UNC13C* Gene

- Related to immune response to IE1B antigen for Human Herpesvirus-6 ($\beta = -506.2, P = 10^{-11}$)
- *UNC13C* may play a role in vesicle maturation during exocytosis
- Appeared in same GWAS of antibody response to the **smallpox vaccine** (PMID: 22542470, $P = 8 \times 10^{-7}$)



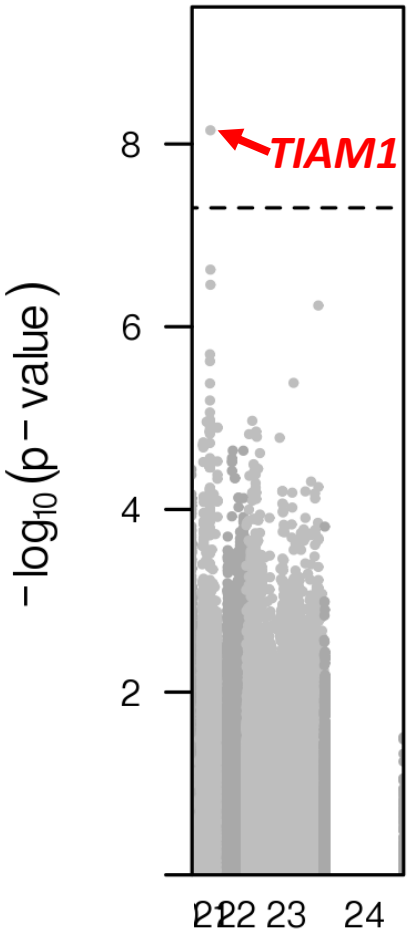
UreA Antigen for **N = 1891**

Helicobacter pylori



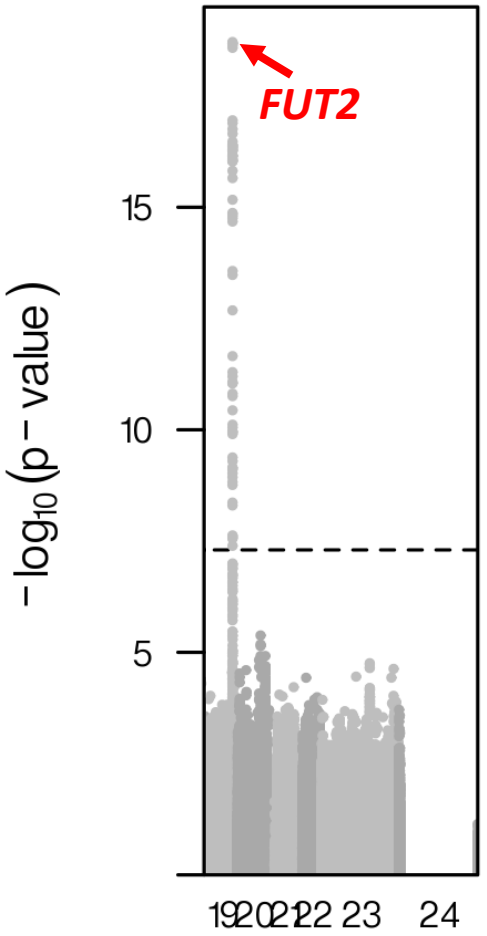
rs113552662 – *TIAM1* Gene

- Related to immune response to UreA antigen for *Helicobacter pylori* ($\beta = -2023, P = 7 \times 10^{-9}$)
- *TIAM1* (**T Cell Lymphoma Invasion and Metastasis 1**) plays a role in cell invasion, metastasis, and carcinogenesis
- Appeared in GWAS of **HIV-1 virologic failure** (PMID: 25461247, $P = 10^{-6}$)



rs681343 - Fucosyltransferase 2

- Related to immune response to BK VP1 antigen for Human Polyomavirus BKV ($\beta = -386.529, P = 2 \times 10^{-19}$)
- *FUT2* is an enzyme required for the soluble A & B antigen synthesis pathway

Also associated with:

- **Crohn Disease** (PMID: 26192919)
- Several **autoimmune diseases** (PMID: 26301688)
- **Infectious diseases** such as rotavirus, norovirus, and pneumonia (PMID: 24277741, 16306606, 2876155)