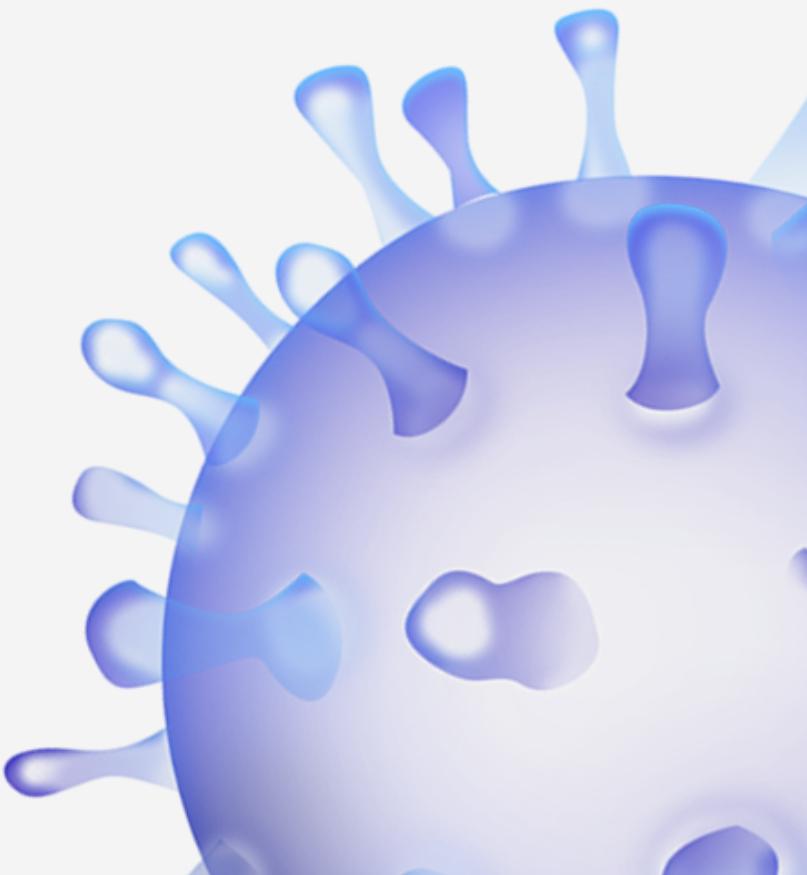
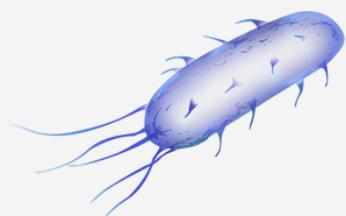
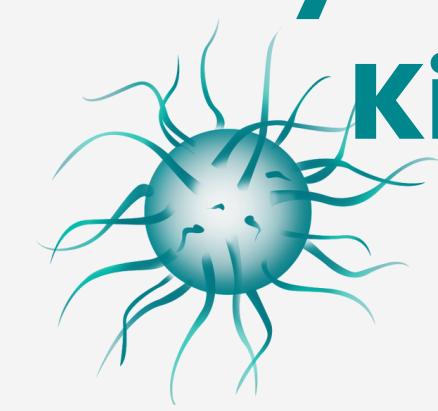




# **INVESTIGATING MEASLES SUSCEPTIBILITY GENES ON A GLOBAL SCALE**

**Alysha Puti Maulidina  
Kimberly Mazel**



# Contents

- 01** Introduction
- 02** Problems
- 03** Proposed method
- 04** Results
- 05** Conclusion and recommendation



# Introduction

- Measles is a highly contagious viral disease that persists as a global threat to public health despite the introduction of effective vaccines.
- Measles occur globally but the epidemiology of the infections varies depending on the geographical regions.

It is said to be prevalent in **Africa** and **Southeast Asia**

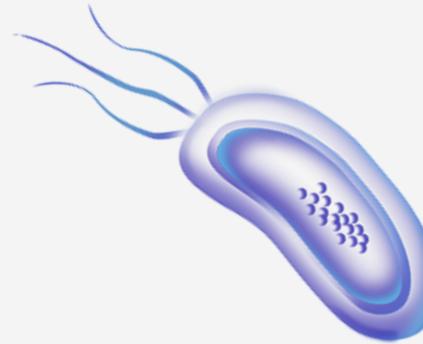
**The goal of this study is to identify genetic variations that may contribute to measles disease transmission and their distributions across population.**

Does one population share a measles gene exclusive to their population?



Primary symptoms of Measles:  
cold-like symptoms (**flu, cough**)  
and **skin rashes**

# Proposed Solution



## Dataset & SNPs

We plan to integrate a global dataset containing genomic information along with SNPs that are associated with measles to identify which variations are susceptible to them.

## Susceptibility Mapping

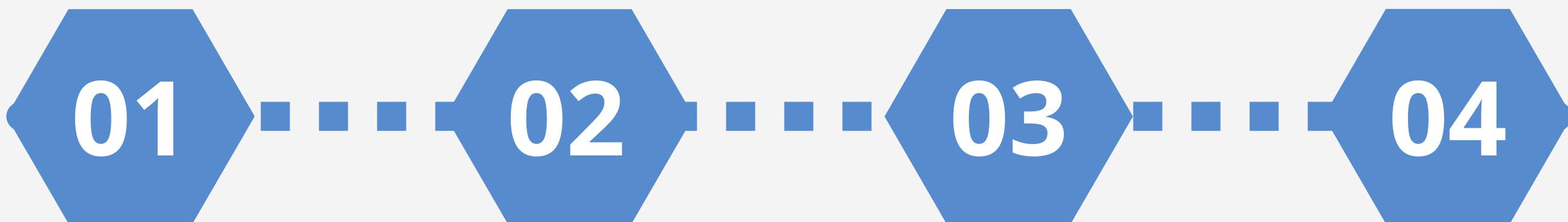
The locations of susceptibility to measles can be mapped out on a global scale, providing insight on the distribution of susceptible genetic variants across different populations of the world.

## How can this be useful?

- Improve vaccination strategies
- Personalized plans for individuals with higher genetic risk factors
- Further studies on why some areas are more susceptible to measles



# Methodology

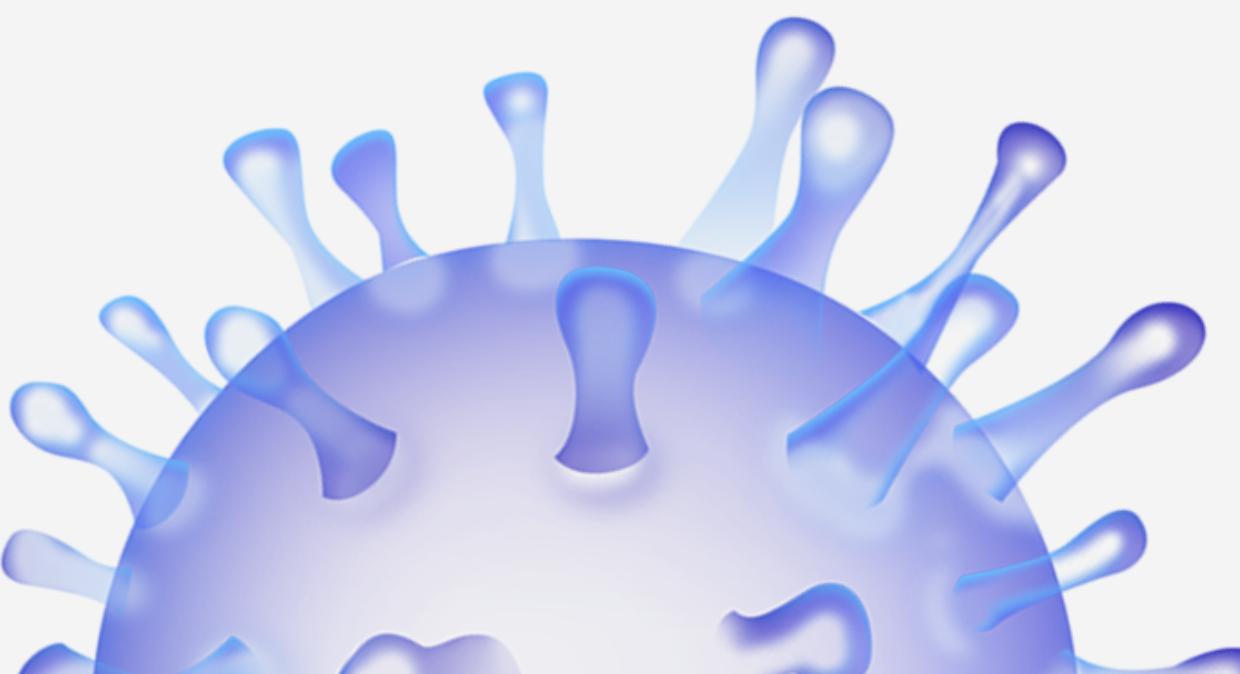


**Identification of measles-associated SNPs**

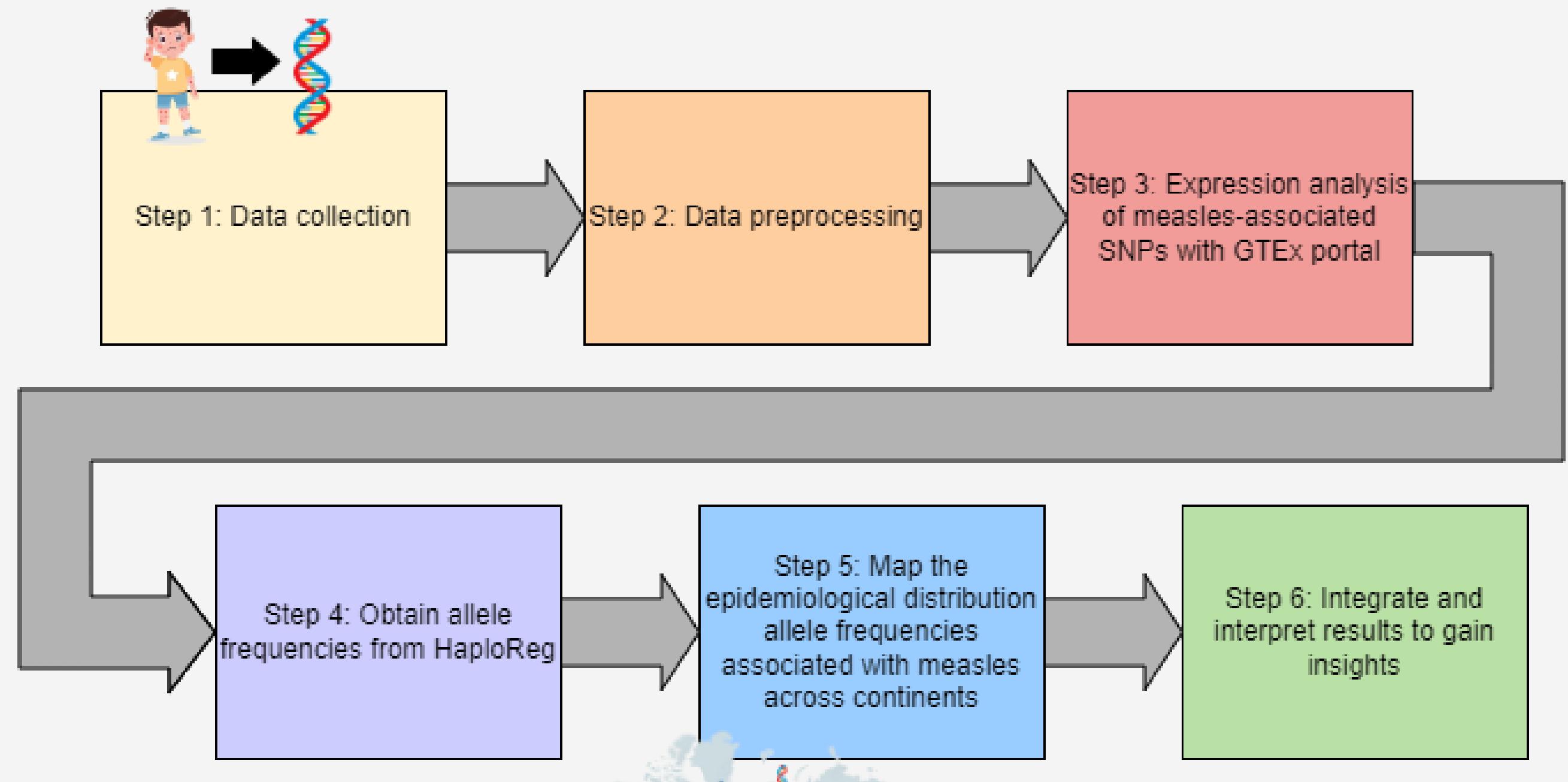
**Using GTEx Portal to see where the SNPs appear in our body**

**HaploReg and Ensembl: Understanding if these measles-related DNA changes are more common in some populations than others.**

**Using PCA observe how genetic variants cluster or disperse across different populations**



# Bioinformatics Pipeline



# Results - GWAS

U	V
MAPPED_GENE	SNPS
NUDT12 - NIHCOLE	rs500141
MIR548A1HG - RPL21P61	rs71868007
PHGDH	rs72988658
SGCG	rs74754407
HLA-DOA - HLA-DPA1	rs78331658
MYO1E	rs78595738
NCK2	rs80286292
YPEL2	rs8082454
PPIAP34 - ZBTB40	rs11584259
RNU6-279P - RNU6-871P	rs1353279
KCNIP4	rs144550967
FAM83D	rs186508932
RNU1-14P - RN7SKP218	rs191080780
HUNK	rs2833560
FTH1P14 - SRSF2P1	rs5928516
DELEC1 - U2	rs75327648
XIRP2	rs79887329
BBX	rs9879864
LINC01938 - RN7SKP60	rs116236449
PAPPA2	rs11806613
C7orf31	rs12700593
MIR4675 - NEBL	rs141236446
CTNND2	rs143835899
GC	rs144616029
ADGRL3	rs147407934
LINC02647	rs149731223
DLG2	rs190500889
RNF138P2 - RNU1-14P	rs190844313
LINC00210 - LINC01653	rs191590039
RN7SKP218 - LINC01446	rs192648543
KIAA1671	rs200507792
MRM3P2 - PRPS1L1	rs4568518

Purpose: to see the genes  
that make humans more  
susceptible to measles

# Results - GTEx Portal

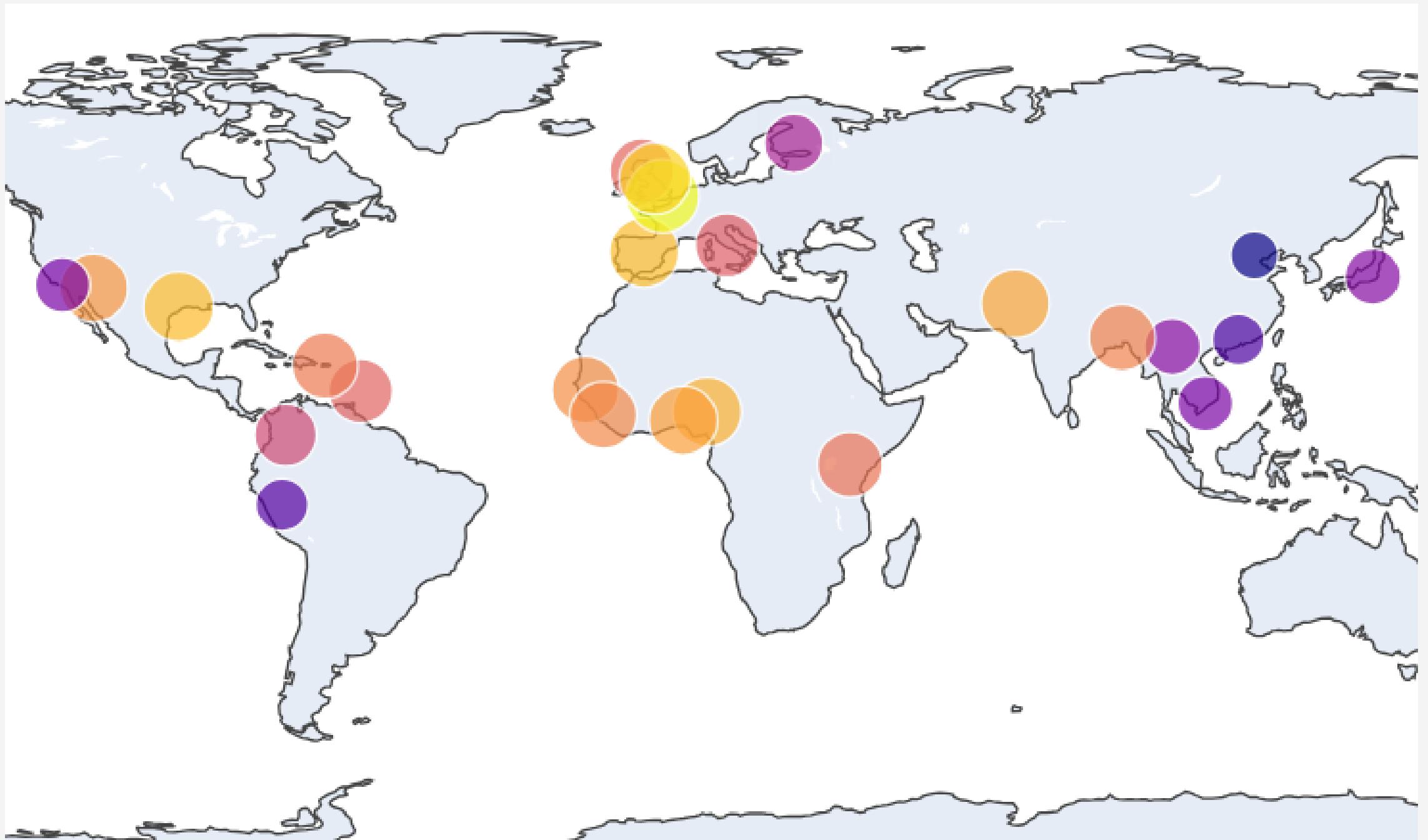
Gencode Id	Gene Symbol	Variant Id	SNP Id	P-Value	NES	Tissue
ENSG00000108395.13	TRIM37	chr17_59363163_G_T_b38	rs8082454	0.000015	-0.20	Adipose - Visceral (Omentum)
ENSG00000218510.6	LINC00339	chr1_22405893_G_C_b38	rs11584259	0.000022	0.29	Adipose - Subcutaneous
ENSG00000142794.18	NBPF3	chr1_22405893_G_C_b38	rs11584259	0.000063	0.29	Lung
ENSG00000162552.14	WNT4	chr1_22405893_G_C_b38	rs11584259	0.000081	0.18	Thyroid
ENSG00000118596.11	SLC16A7	chr12_59422765_T_C_b38	rs1353279	0.000033	-0.25	Liver
ENSG00000114439.18	BBX	chr3_107664490_A_G_b38	rs9879864	0.000061	0.083	Whole Blood
ENSG00000116183.10	PAPPA2	chr1_176566002_A_G_b38	rs11806613	0.0000000051	0.44	Pancreas
ENSG00000153790.11	C7orf31	chr7_25176497_T_C_b38	rs12700593	0.0000020	-0.11	Cells - Cultured fibroblasts
ENSG00000153790.11	C7orf31	chr7_25176497_T_C_b38	rs12700593	0.0000021	-0.16	Adipose - Subcutaneous
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000000000084	0.22	Thyroid
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.00000000065	0.25	Esophagus - Muscularis
ENSG00000071189.21	SNX13	chr7_18001462_G_A_b38	rs4568518	0.0000000028	-0.17	Testis
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000000093	0.24	Artery - Tibial
ENSG00000071189.21	SNX13	chr7_18001462_G_A_b38	rs4568518	0.0000000017	0.16	Colon - Transverse
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.000000029	0.20	Nerve - Tibial
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000023	0.20	Skin - Sun Exposed (Lower leg)
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000046	0.20	Artery - Aorta
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000052	0.18	Esophagus - Mucosa
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.0000085	0.17	Colon - Transverse
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.000017	0.21	Stomach
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.000026	0.24	Prostate
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.000030	0.18	Colon - Sigmoid
ENSG00000071189.21	SNX13	chr7_18001462_G_A_b38	rs4568518	0.000037	-0.10	Nerve - Tibial
ENSG00000279048.1	RP11-511H23.2	chr7_18001462_G_A_b38	rs4568518	0.000043	0.16	Cells - Cultured fibroblasts

Purpose: to see where in the body these SNPs show up and how they affect what our genes do. Since measles affects the skin, we paid extra attention to SNPs that change how genes work in skin cells.

# Results - Ensembl

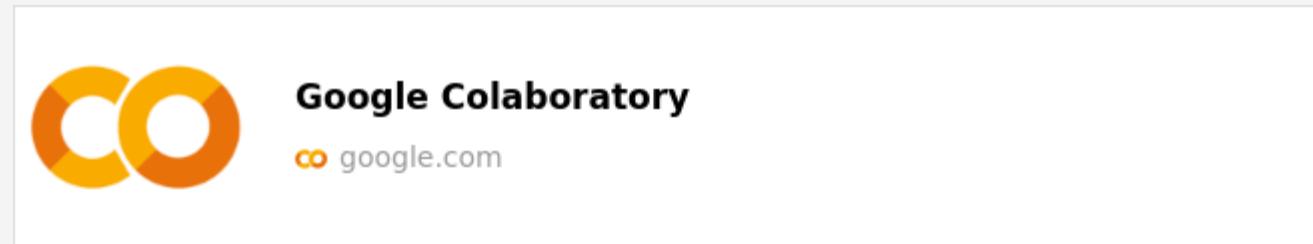
We mapped the population genetics of each chosen measles related SNP based on the results from **Ensembl**.

**Ensembl** is a genome browser containing information on SNPs. This allows us to compare known population distributions to the sample dataset we used.



# Results

Each SNP has information on their reference and alternative allele.



**rs9879864 SNP**

Most severe consequence | [intron variant](#) | [See all p](#)

Alleles | [A/G](#) | Ancestral: A | MAF

Change tolerance | [GDP: G>A,472](#) | [GERP: 472](#)

Location | [Reference/Alternative alleles \(Forward strand\)](#)

- Reference alleles are those that are most commonly found in the population.
- Alternative alleles are any alleles other than the reference allele at the given genomic position.

We cannot make conclusions that measles is more susceptible in one allele or the other, but plotting the distribution SNPs in these alleles can reveal a pattern that can be further studied.

# Results - Pre-processing for PCA

## Genetic Variation

**Data:** raw data consisting of measurements of genetic variations across different individuals from various populations in VCF format (source: 1000genomes)

**Input**

**pysam:** raw, VCF data is parsed to extract the necessary information (i.e. Sample code, variants associated with measles, population code).

**Building genotype matrix:** entries represent the genotype for each variant, row = sample, column = variants.

**Process**

a CSV file is generated, containing the original genotype matrix with samples as rows, variants as columns, and an additional column for the population codes.

**Output**

# Results - PCA analysis



Google Colaboratory  
co google.com

**Dataset Matrix:** processed data arranged in a matrix where rows represent individual samples (people) and columns represent different SNPs.

**Population Labels:** labels indicating each sample's population group (**African, American, Asian, European**).

**Input**

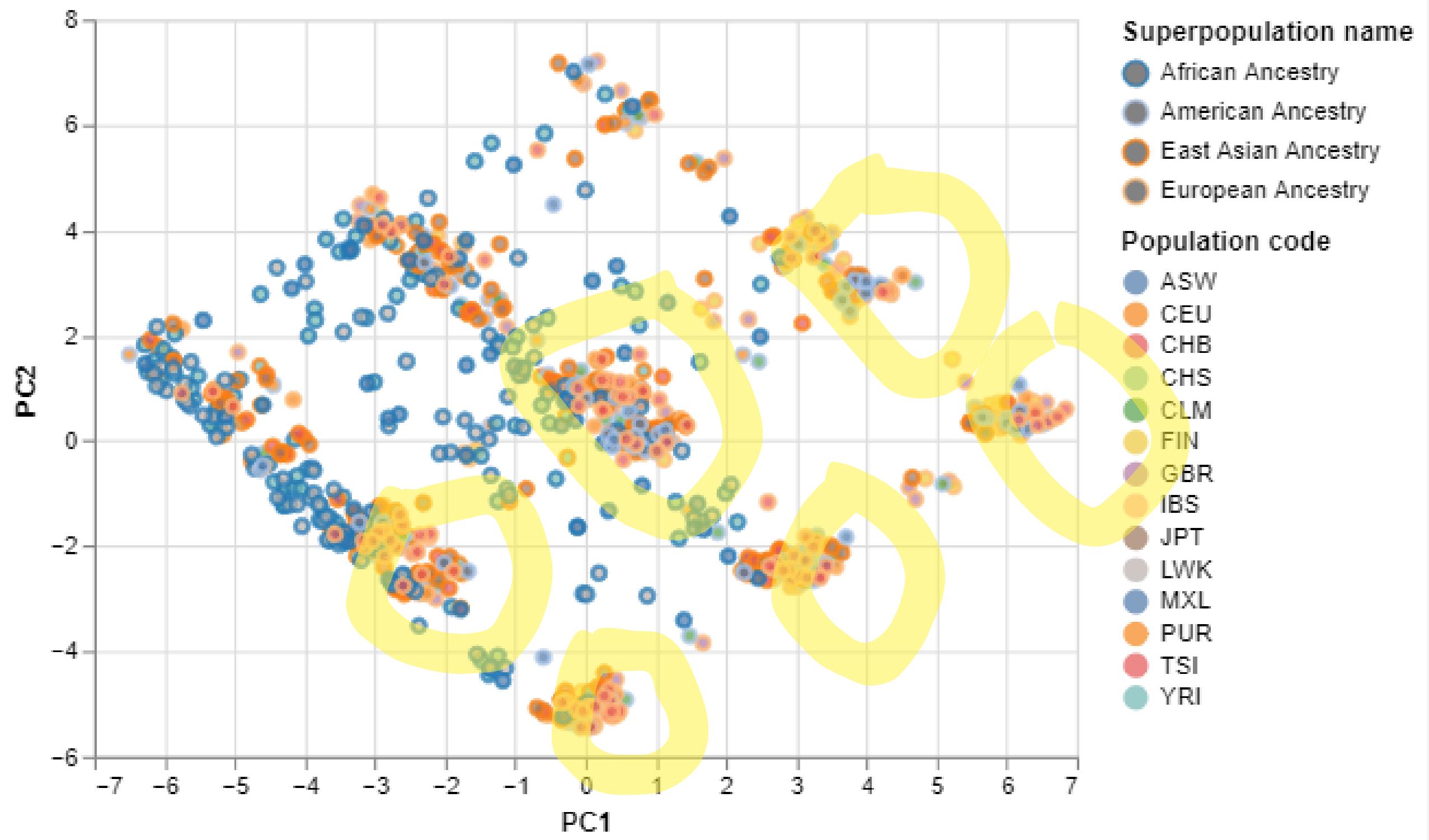
- 1. Calculate the covariance matrix** - shows how each SNP is correlated with every other SNP.
- 2. Eigenvalue decomposition** - helps PCA view the most important patterns or directions in the data
- 3. Principle component** - choosing the top features based on the direction to reduce variables
- 4. Transforming the data** - simpler set of data based on the PC

**Process**

Visualization to show how genetic variation related to susceptibility varies across the different population.

**Output**

# Results - PCA analysis



We can see that...

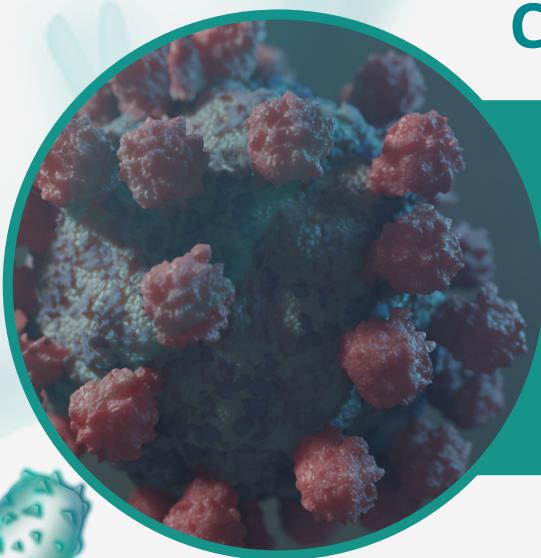
There is a significant overlap between individuals of different ancestries, particularly between those of European, East Asian, and American ancestries.

This could indicate that there are common genetic variants among these individuals that contribute to susceptibility.

Susceptibility to measles may not be strictly confined to genetic ancestry/population.

# Conclusions

## Conclusion 01



The main genes affecting Measles are **NBPF3** and **RP11-511H23.2**. These genes affected the measles risk in individuals and populations.

## Conclusion 02



The variants **rs11584259** and **rs4568518** affected NBPF3 and RP11-511H23.2 respectively.

## Conclusion 03



PCA: SNPs associated with measles susceptibility **might not be unique to one specific ancestry** and could be shared across these populations.

Takeaway and future research:

- Studying the correlation between SNPs and susceptibility -> **We need to know which samples has had measles.**
- Creating effective vaccine strategies based on geography and genetic ancestry -> **e.g. The diversity within certain groups, particularly Africans, could imply variability in how individuals respond to measles infection and vaccination.**

# THANK YOU



Alysha Puti Maulidina



Kimberly Mazel

# REFERENCES

Conis, E. (2019). Measles and the modern history of vaccination. *Public Health Reports*, 134(2), 118–125.  
<https://doi.org/10.1177/0033354919826558>

Kondamudi, N. P., & Waymack, J. R. (2023, August 12). Measles. NCBI Bookshelf.  
<https://www.ncbi.nlm.nih.gov/books/NBK448068/>

Krawiec, C., & Hinson, J. W. (2023, January 16). Rubeola (measles). StatPearls - NCBI Bookshelf.  
<https://www.ncbi.nlm.nih.gov/books/NBK557716/>

Ludlow, M., Lemon, K., De Vries, R. D., McQuaid, S., El, M., G, V. A., Yüksel, S., Rj, V., Osterhaus, A. D. M. E., De Swart, R. L., & Duprex, W. P. (2013). Measles virus infection of epithelial cells in the macaque upper respiratory tract is mediated by subepithelial immune cells. *Journal of Virology*, 87(7), 4033–4042.  
<https://doi.org/10.1128/jvi.03258-12>

Misin, A., Antonello, R. M., Di Bella, S., Campisciano, G., Zanotta, N., Giacobbe, D. R., Comar, M., & Luzzati, R. (2020). Measles: An Overview of a Re-Emerging Disease in Children and Immunocompromised patients. *Microorganisms*, 8(2), 276. <https://doi.org/10.3390/microorganisms8020276>

Shastry, B. S. (2007). SNPs in disease gene mapping, medicinal drug development and evolution. *Journal of Human Genetics*, 52(11), 871–880. <https://doi.org/10.1007/s10038-007-0200-z>

World Health Organization. (2023, August 9). Measles. WHO. <https://www.who.int/news-room/fact-sheets/detail/measles>

