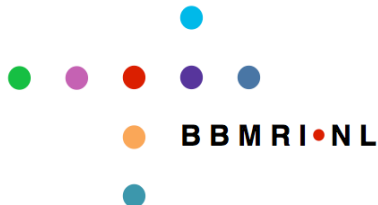


# Population structure based on common and rare variation in the Netherlands

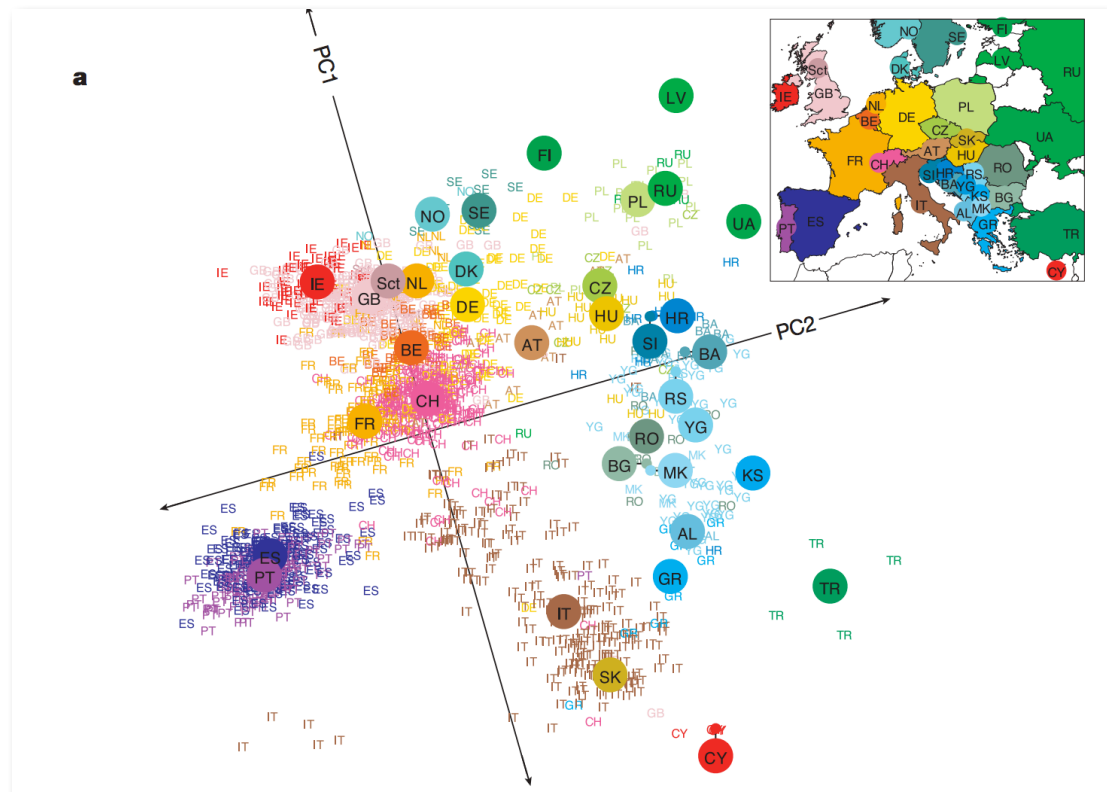
Sara Pulit  
de Bakker Lab  
UMC Utrecht, Netherlands



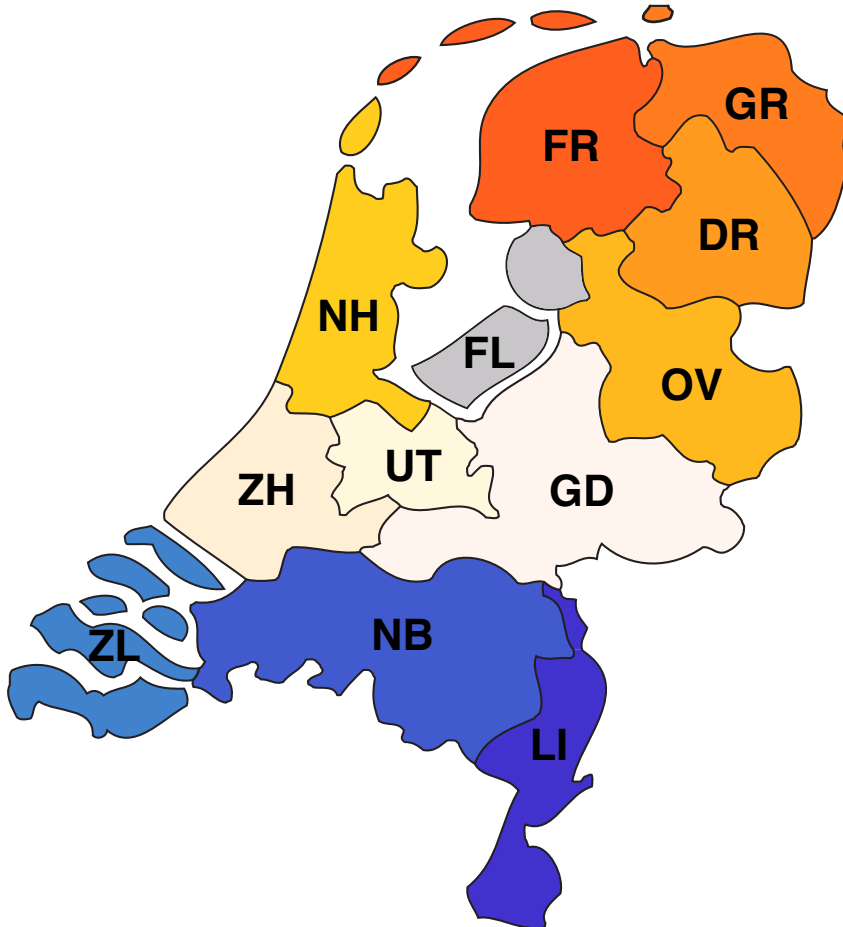
## LETTERS

## Genes mirror geography within Europe

John Novembre<sup>1,2</sup>, Toby Johnson<sup>4,5,6</sup>, Katarzyna Bryc<sup>7</sup>, Zoltán Kutalik<sup>4,6</sup>, Adam R. Boyko<sup>7</sup>, Adam Auton<sup>7</sup>, Amit Indap<sup>7</sup>, Karen S. King<sup>8</sup>, Sven Bergmann<sup>4,6</sup>, Matthew R. Nelson<sup>8</sup>, Matthew Stephens<sup>2,3</sup> & Carlos D. Bustamante<sup>7</sup>

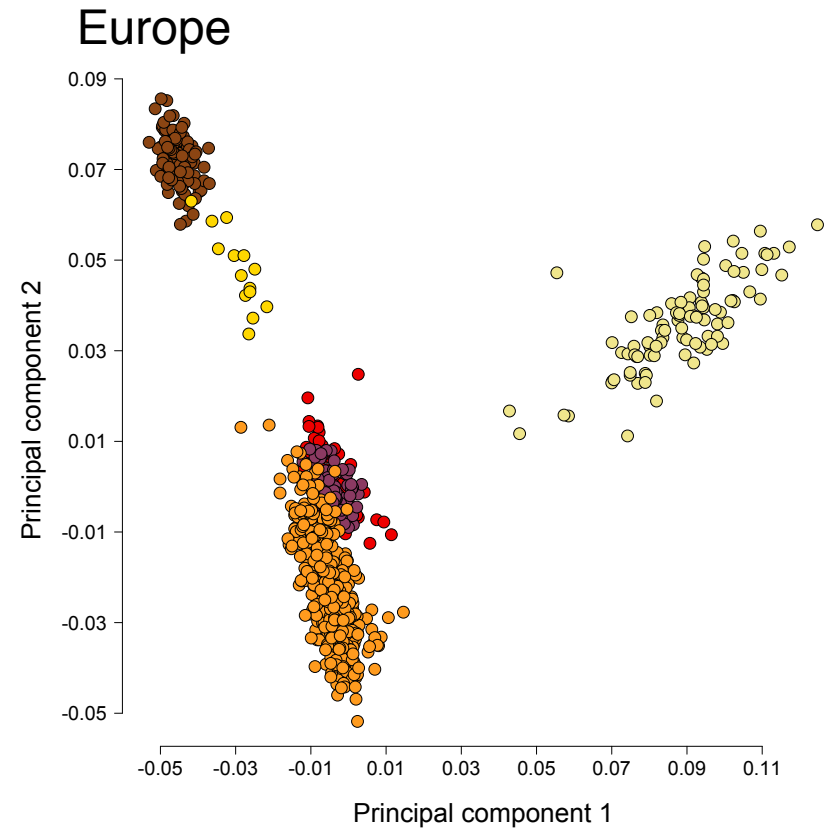
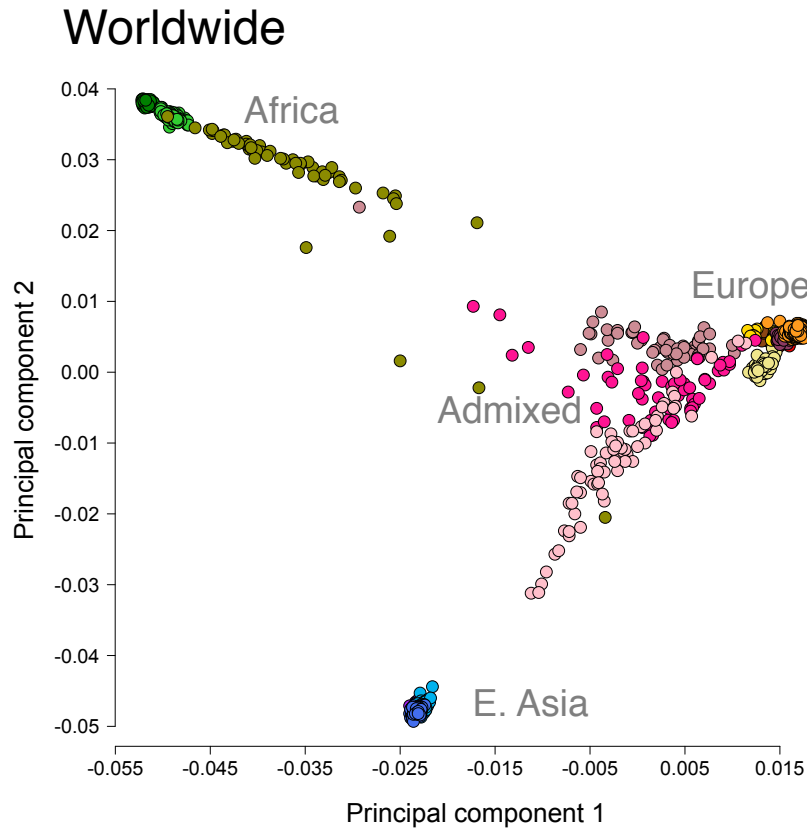


# Genome of the Netherlands (GoNL)

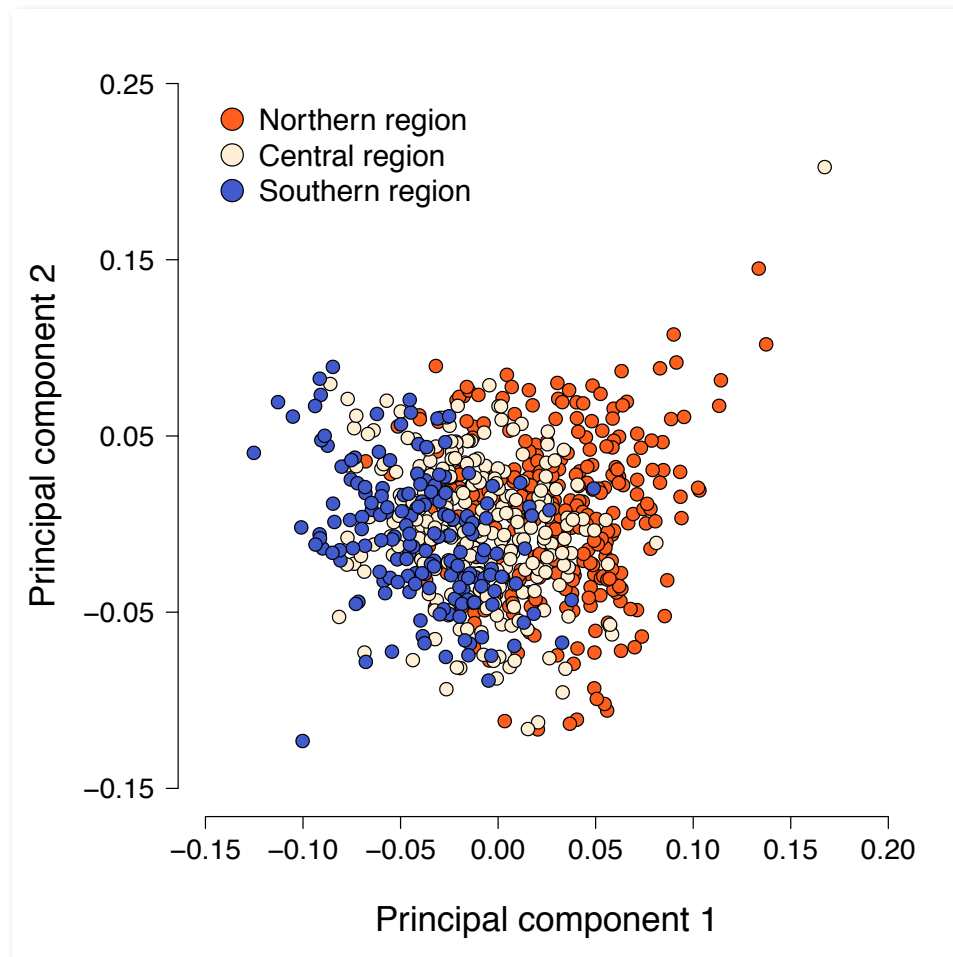


Province	Code	N
Friesland	FR	62
Groningen	GR	57
Drenthe	DR	56
Overijssel	OV	57
Noord-Holland	NH	91
Utrecht	UT	48
Gelderland	GD	57
Zuid-Holland	ZH	168
Zeeland	ZL	46
Noord-Brabant	NB	67
Limburg	LI	58

# Global and continental population differentiation

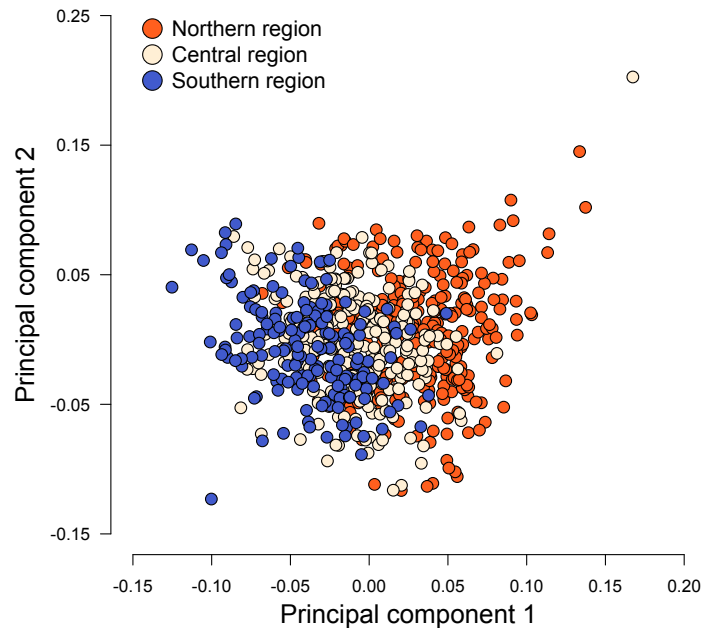


# North – South structure in GoNL

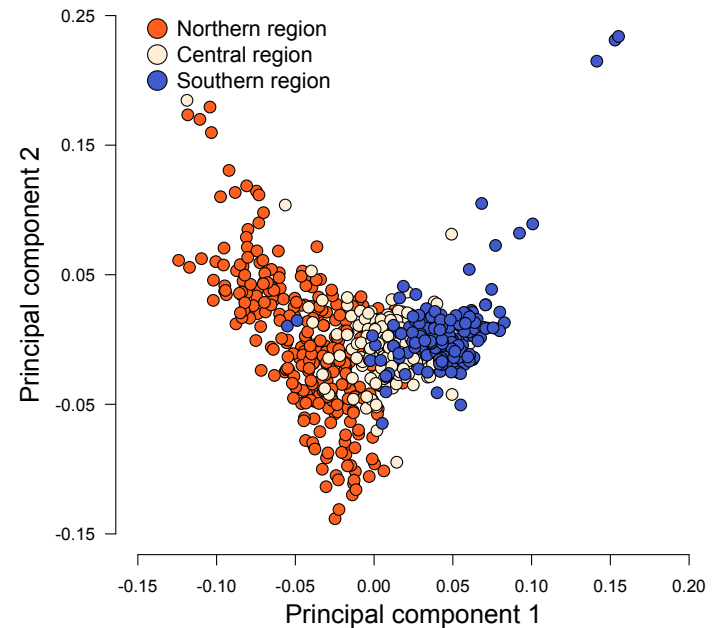


# Rare variant structure not captured by PCA

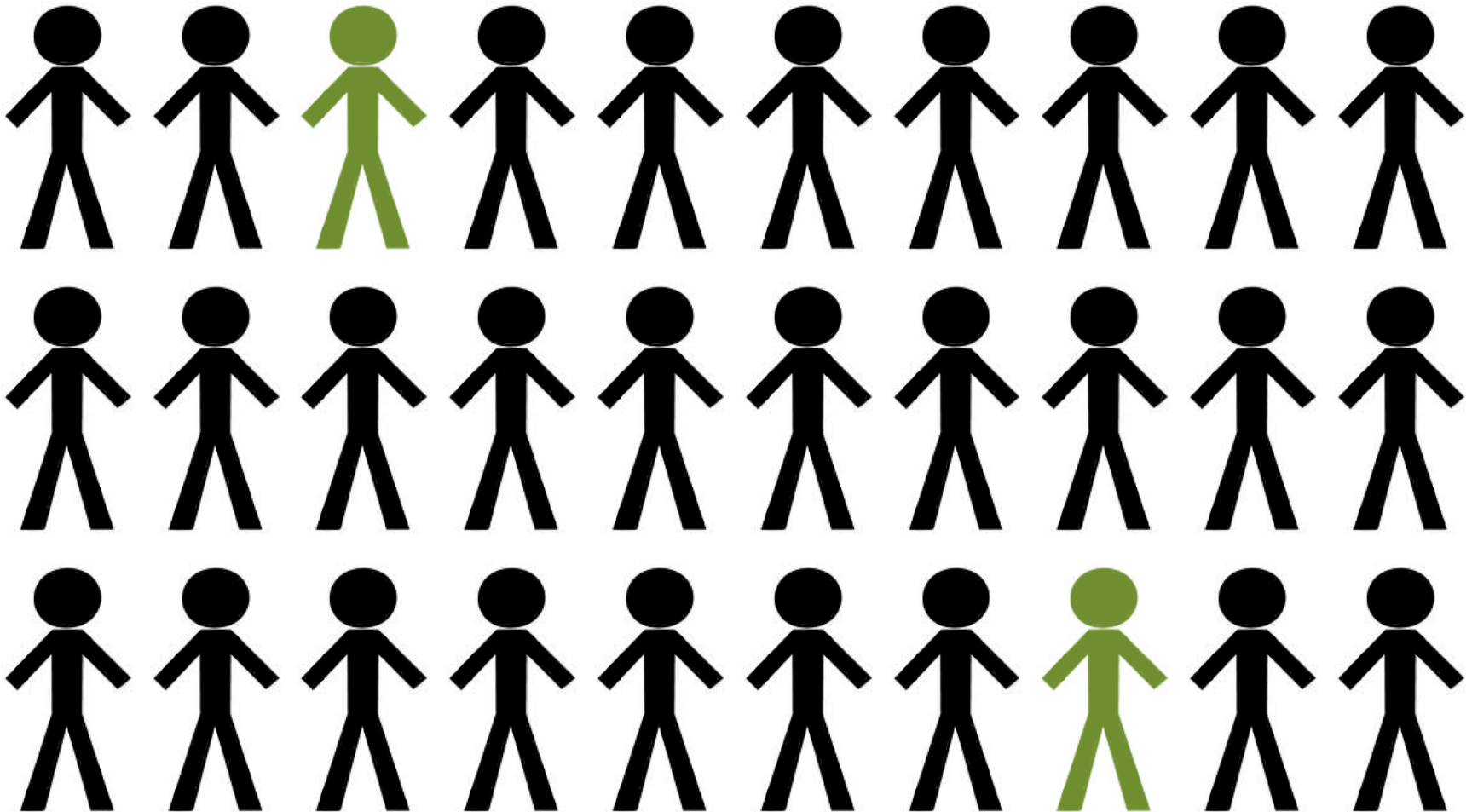
Variation: >5%



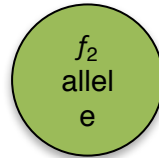
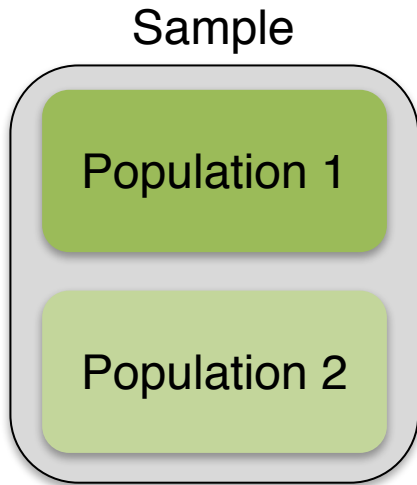
Variation: >0.5%



# $f_2$ variants

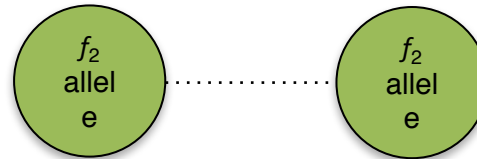
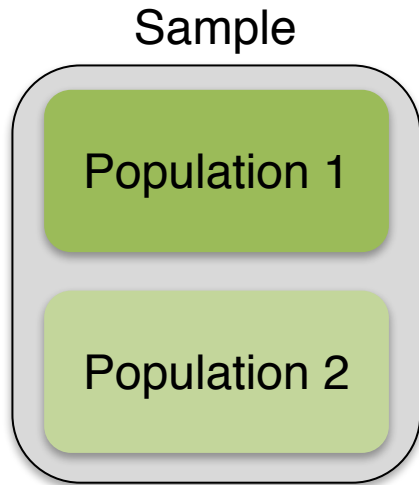


# Connections within and across populations

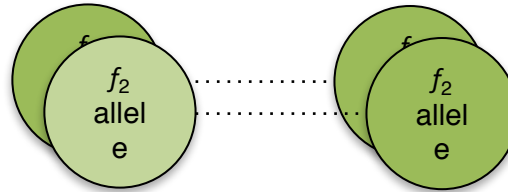
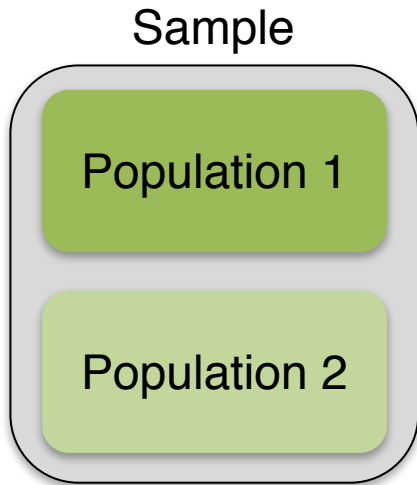




# Connections within and across populations



# Connections within and across populations

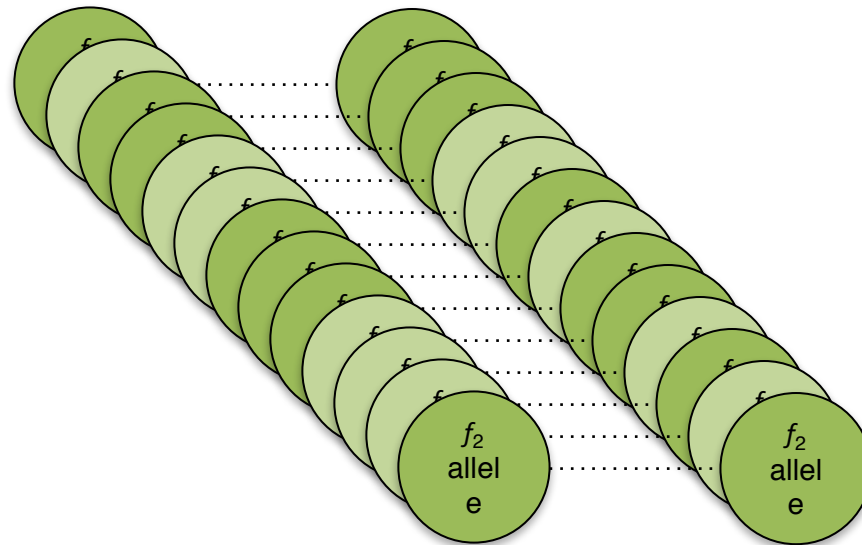


# Connections within and across populations

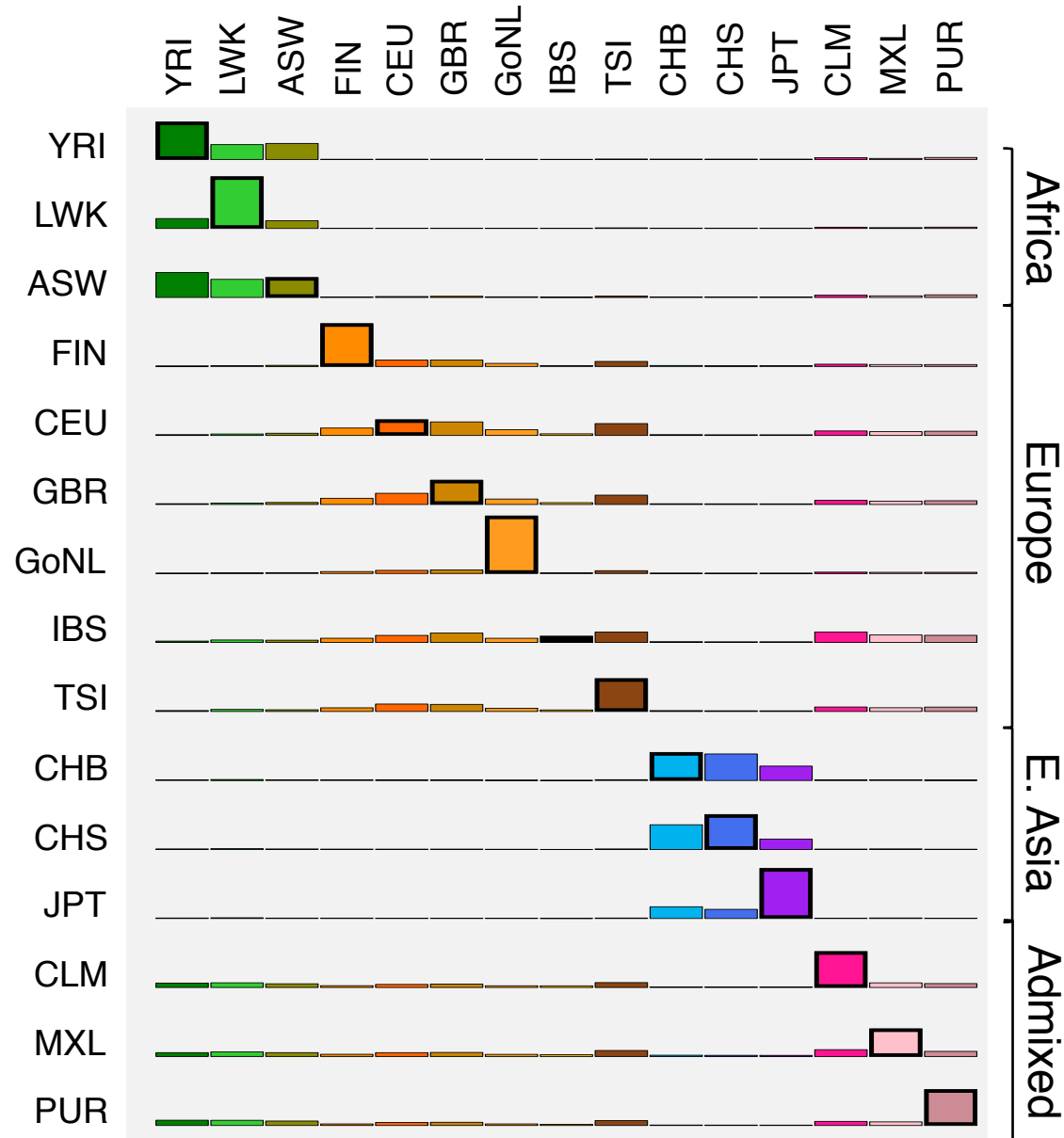
Sample

Population 1

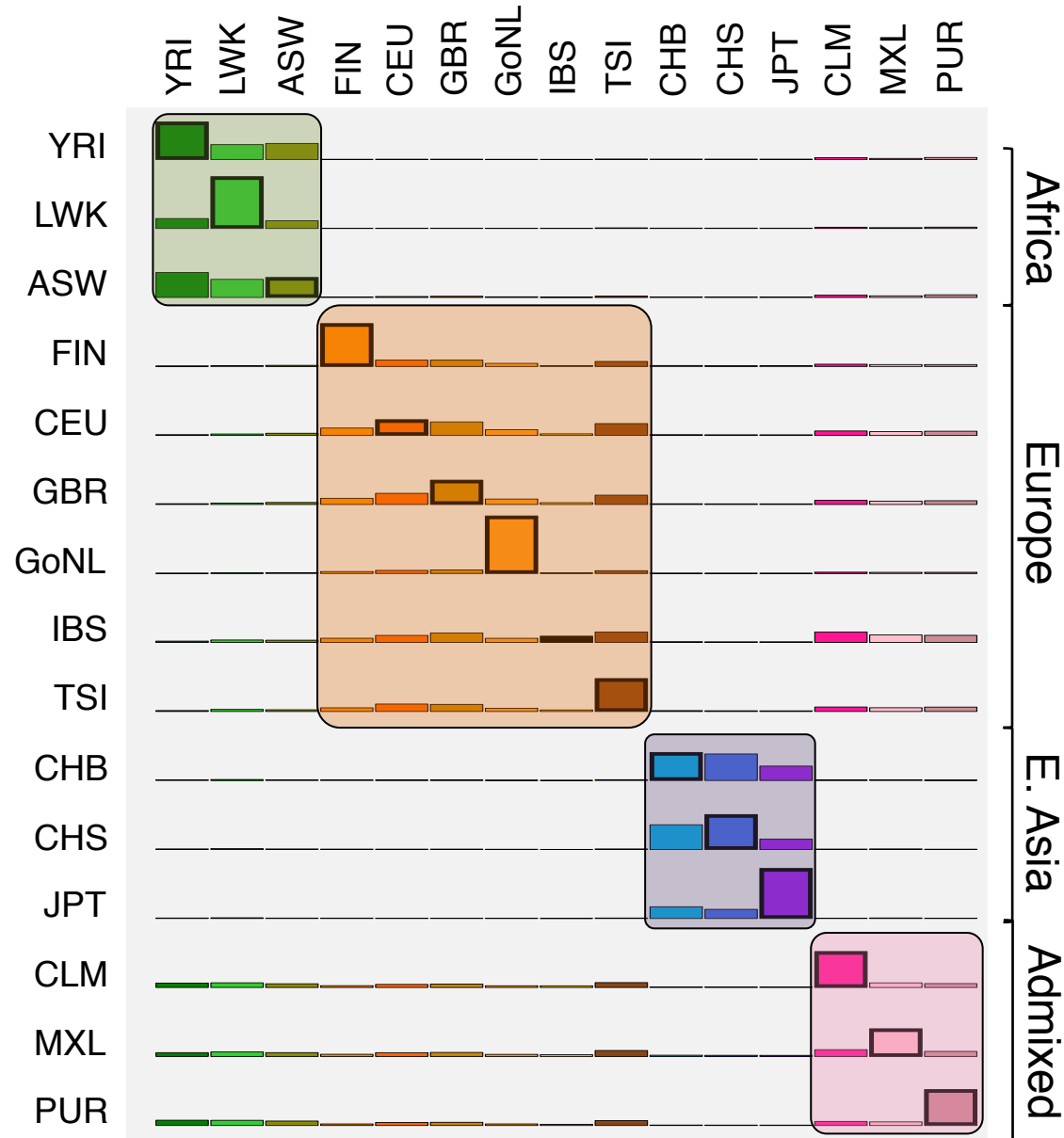
Population 2



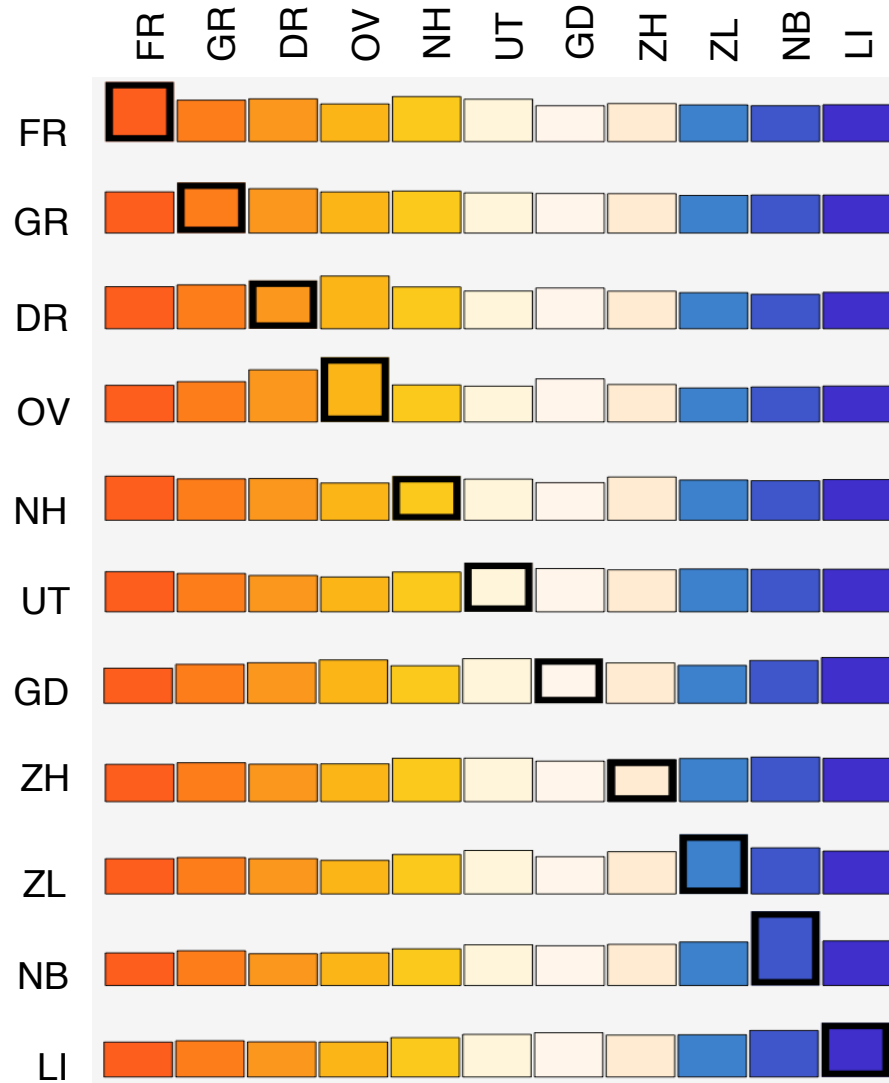
# $f_2$ variants in a global sample



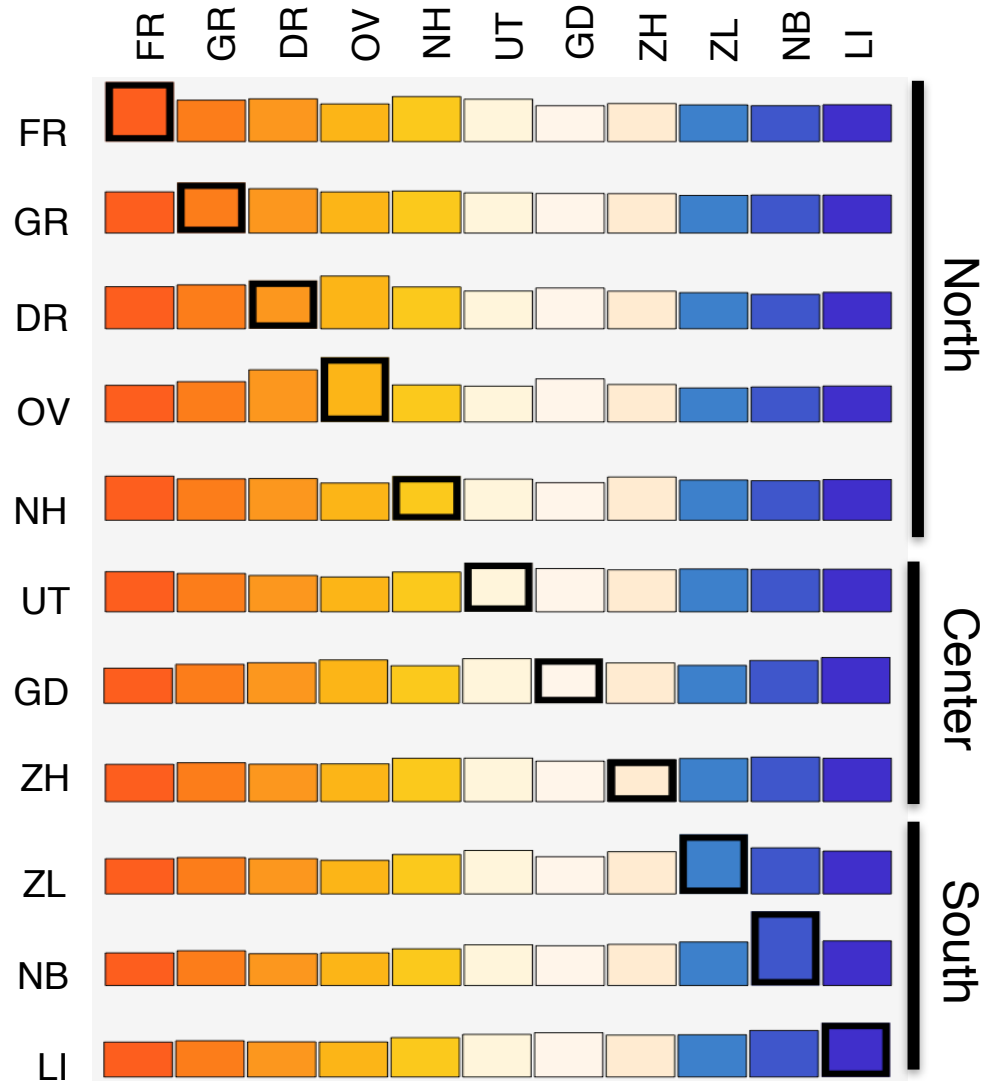
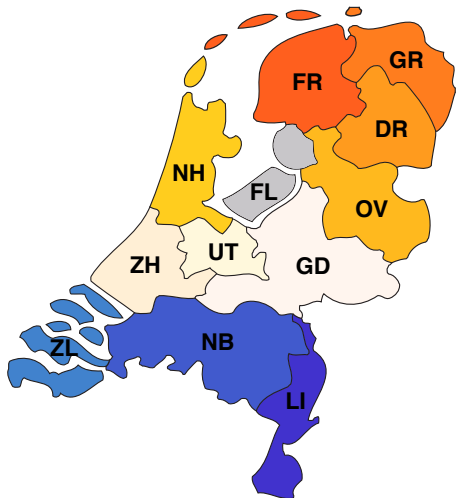
# $f_2$ variants in a global sample



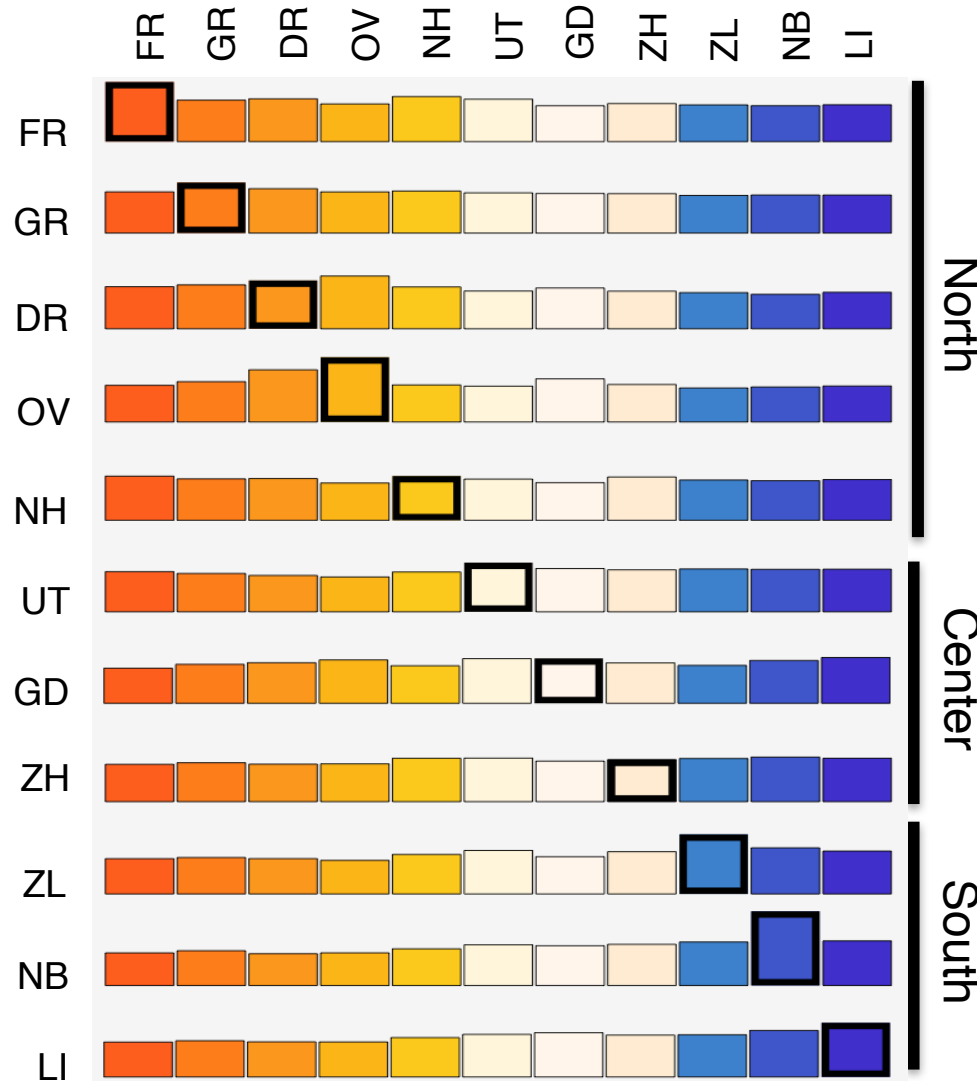
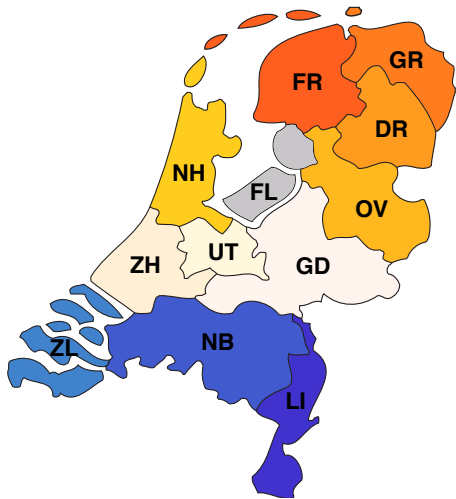
# $f_2$ variants in GoNL



# $f_2$ variants in GoNL



# $f_2$ variants in GoNL



North

Center

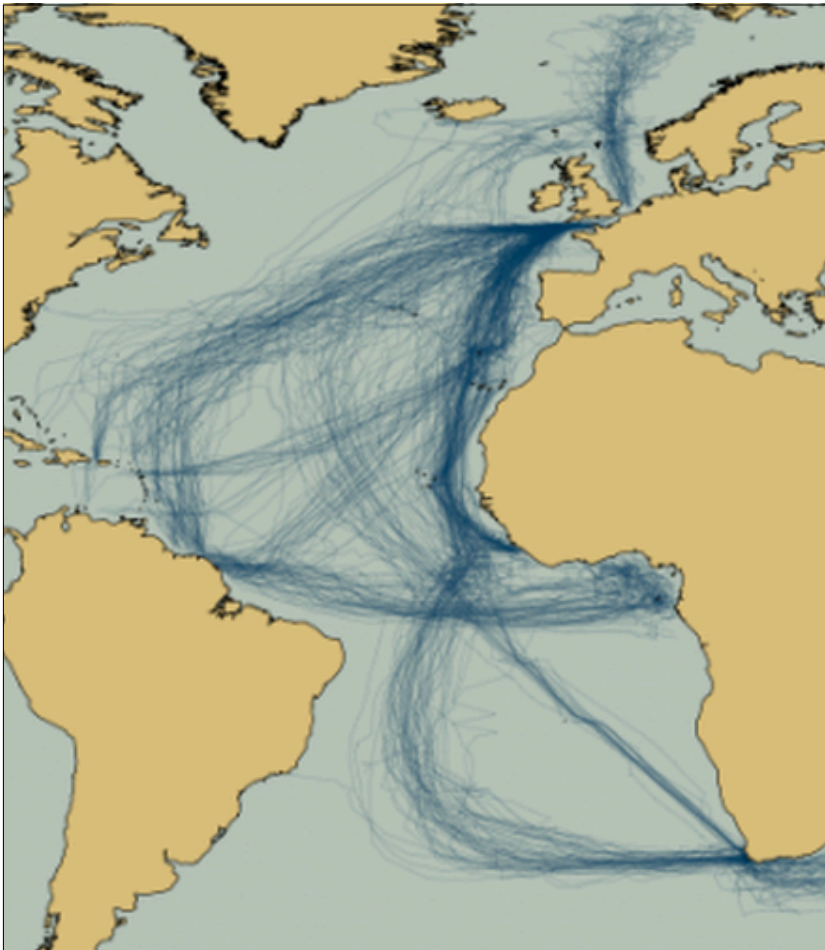
South

$\chi^2=27075$ ,  
 $p<10^{-200}$

$\chi^2=1450$ ,  
 $p<10^{-200}$

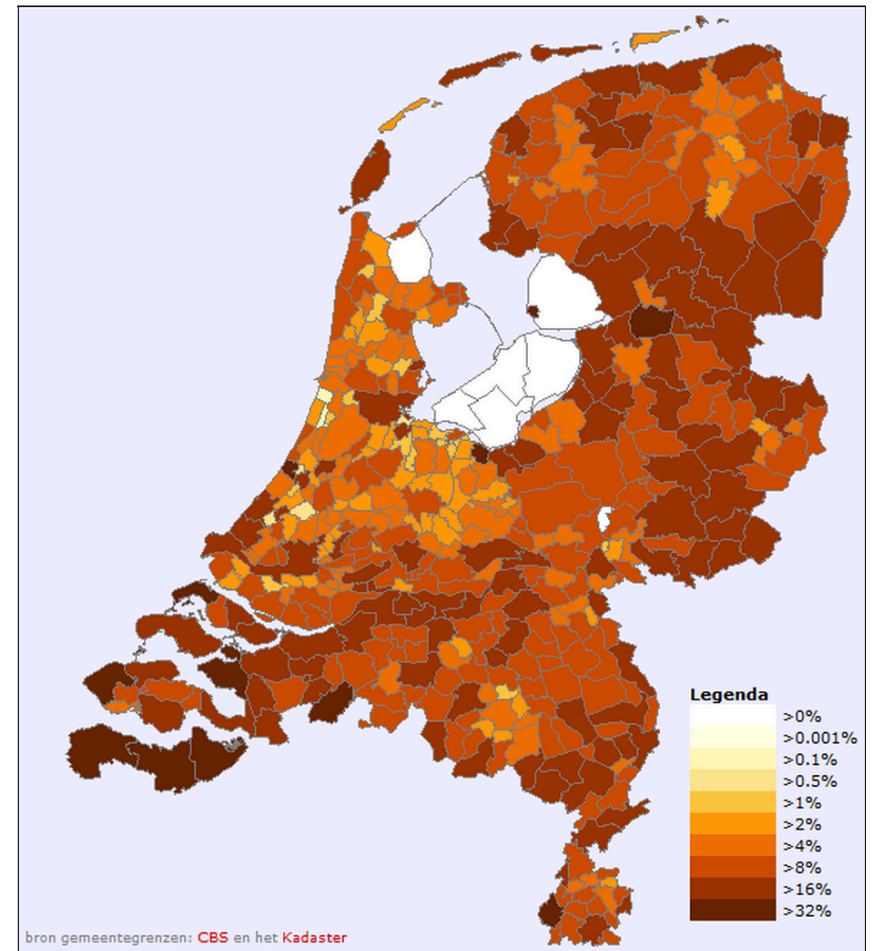


## Central provinces: seat of the Dutch economy



Map: [spatialanalysis.co.uk](http://spatialanalysis.co.uk)  
Data: Climatological Database for the World's Oceans

## Limited migration in recent centuries



Merteens Instituut

# Conclusions and next steps

Common variation mirrors geography in a homogeneous sample drawn from a small area but does not capture provincial clustering

Sharing of rare variants reveals fine-scaled population substructure

The potential impact of rare variant structure on rare variant association studies requires further investigation

# Acknowledgements

## UMC Utrecht



Androniki Menelaou  
Clara C. Elbers  
Laurent C. Francioli  
Paul I.W. de Bakker

## Columbia University



Pier F. Palamara  
Itsik Pe'er

## GoNL Steering Committee

Cisca Wijmenga  
Morris A. Swertz  
Cornelia M. van Duijn  
Dorret I. Boomsma  
Eline Slagboom  
Gertjan B. van Ommen  
Paul I.W. de Bakker

## UMC Groningen and UMC Leiden

Morris Swertz  
Freerk van Dijk  
Pieter B.T. Nieerincx  
Victor Guryev  
Kai Ye

