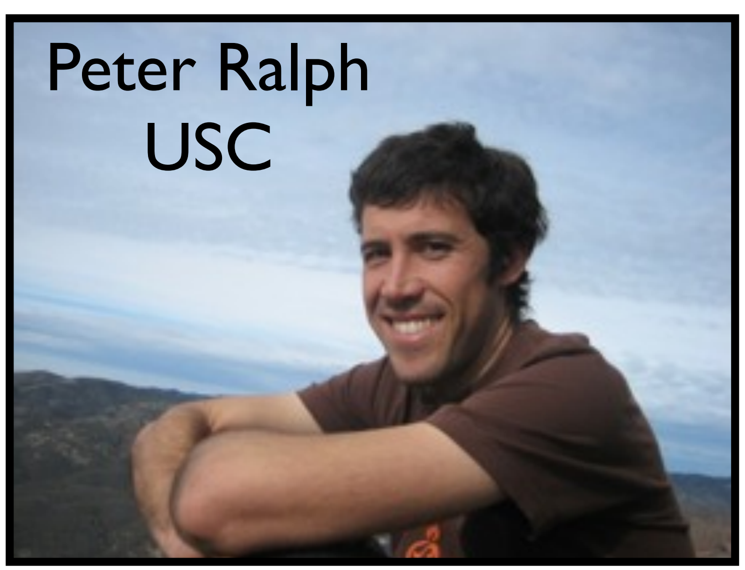
The Geography of Genetic Structure and Admixture

Gideon Bradburd
UC Davis
@gbradburd
genescape.org

The Geography of Genetic Structure and Admixture

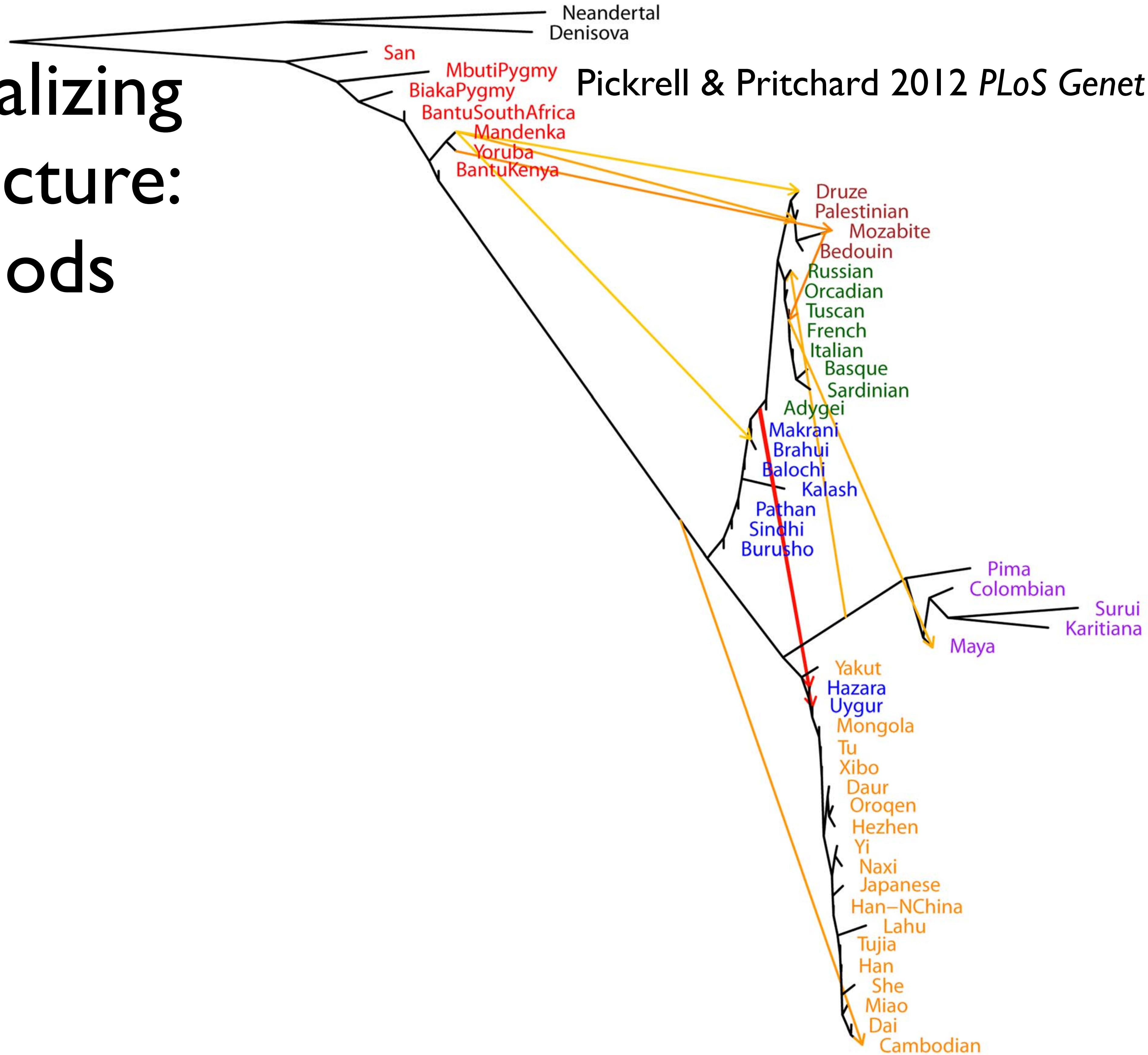


Gideon Bradburd
UC Davis
@gbradburd
genescape.org

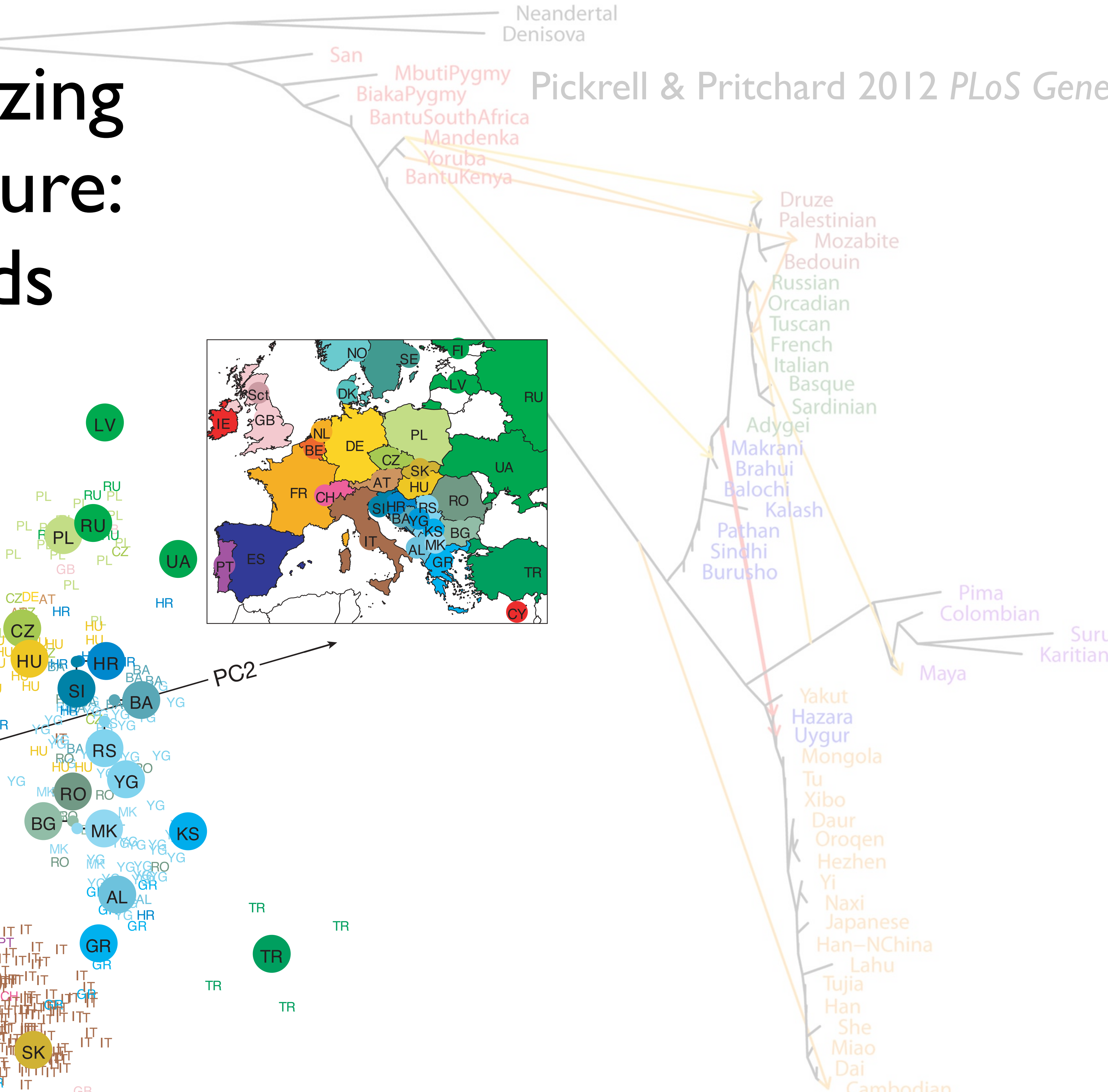
inferring & visualizing population structure: current methods

inferring & visualizing population structure: current methods

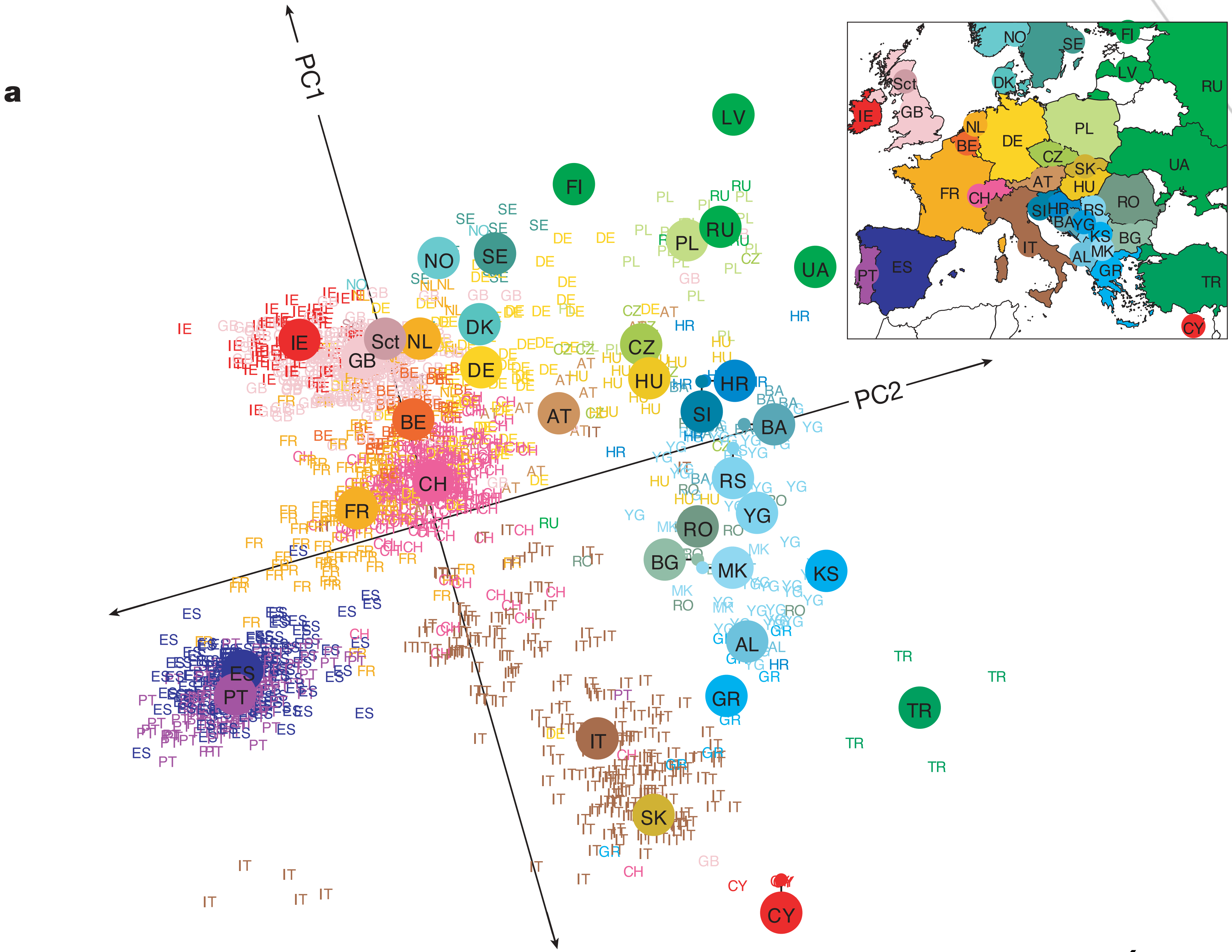
Pickrell & Pritchard 2012 *PLoS Genet*



inferring & visualizing population structure: current methods



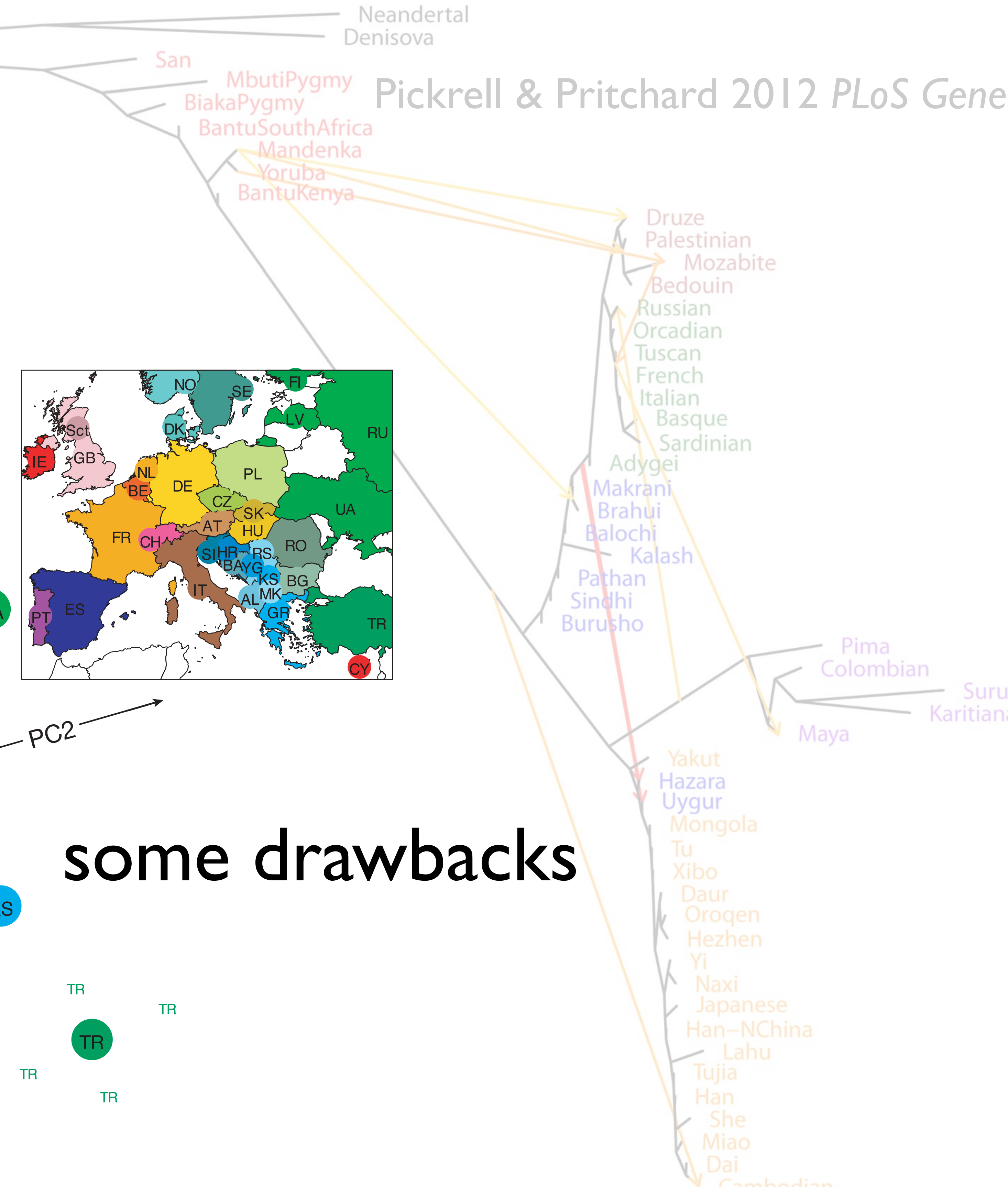
Pickrell & Pritchard 2012 *PLoS Genet*



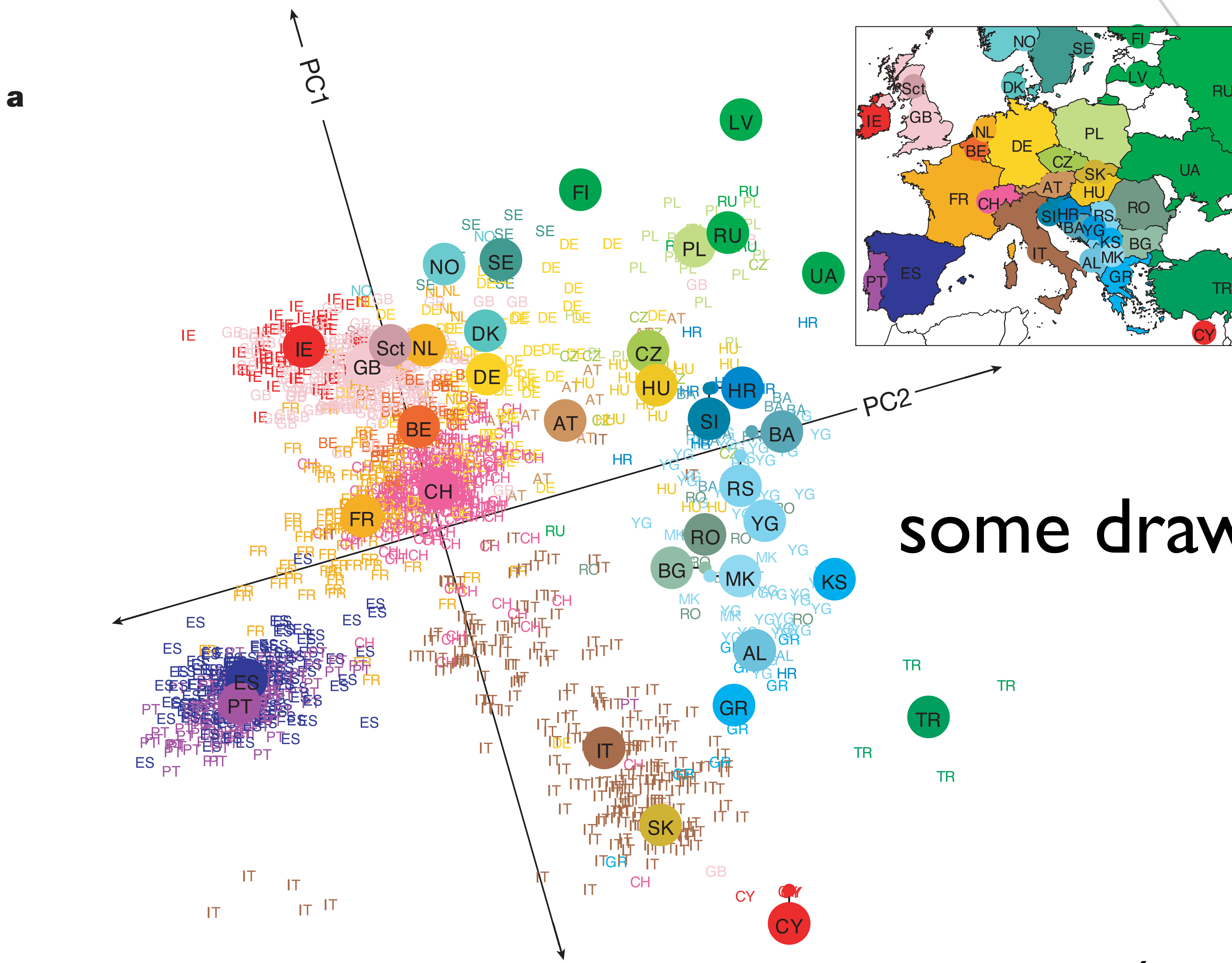
Novembre *et al.* 2008 *Nature*

(see also: SPA,EEMS)

inferring & visualizing population structure: current methods



Pickrell & Pritchard 2012 *PLoS Genet*

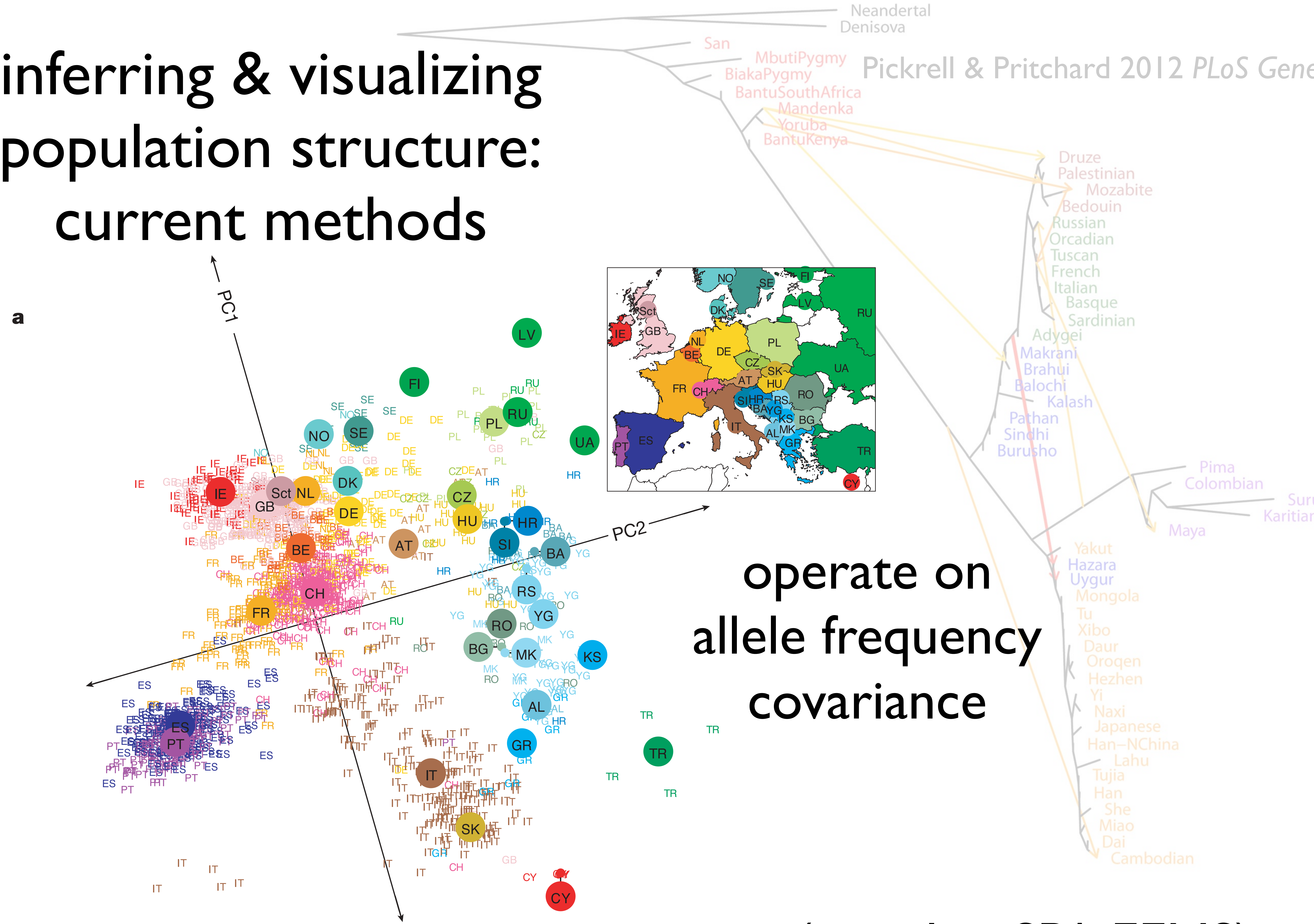


Novembre *et al.* 2008 *Nature*

some drawbacks

(see also: SPA,EEMS)

inferring & visualizing population structure: current methods



Novembre *et al.* 2008 *Nature*

Pickrell & Pritchard 2012 *PLoS Genet*

operate on allele frequency covariance

(see also: SPA,EEMS)

SpaceMix Data:

$$\hat{f}_{\ell,k} = \text{sample frequency at locus } \ell \text{ in population } k$$

$$X_{\ell,k} = \frac{\hat{f}_{\ell,k} - \epsilon_\ell}{\sqrt{\epsilon_\ell(1-\epsilon_\ell)}} \quad = \quad \begin{array}{r}\text{standardized sample frequency} \\ \text{at locus } \ell \text{ in population } k\end{array}$$

$$\widehat{\Omega} = \frac{1}{L} X X^T \quad = \quad \begin{array}{r}\text{standardized sample covariance} \\ \text{across all } L \text{ loci}\end{array}$$

SpaceMix Model: Isolation by Distance

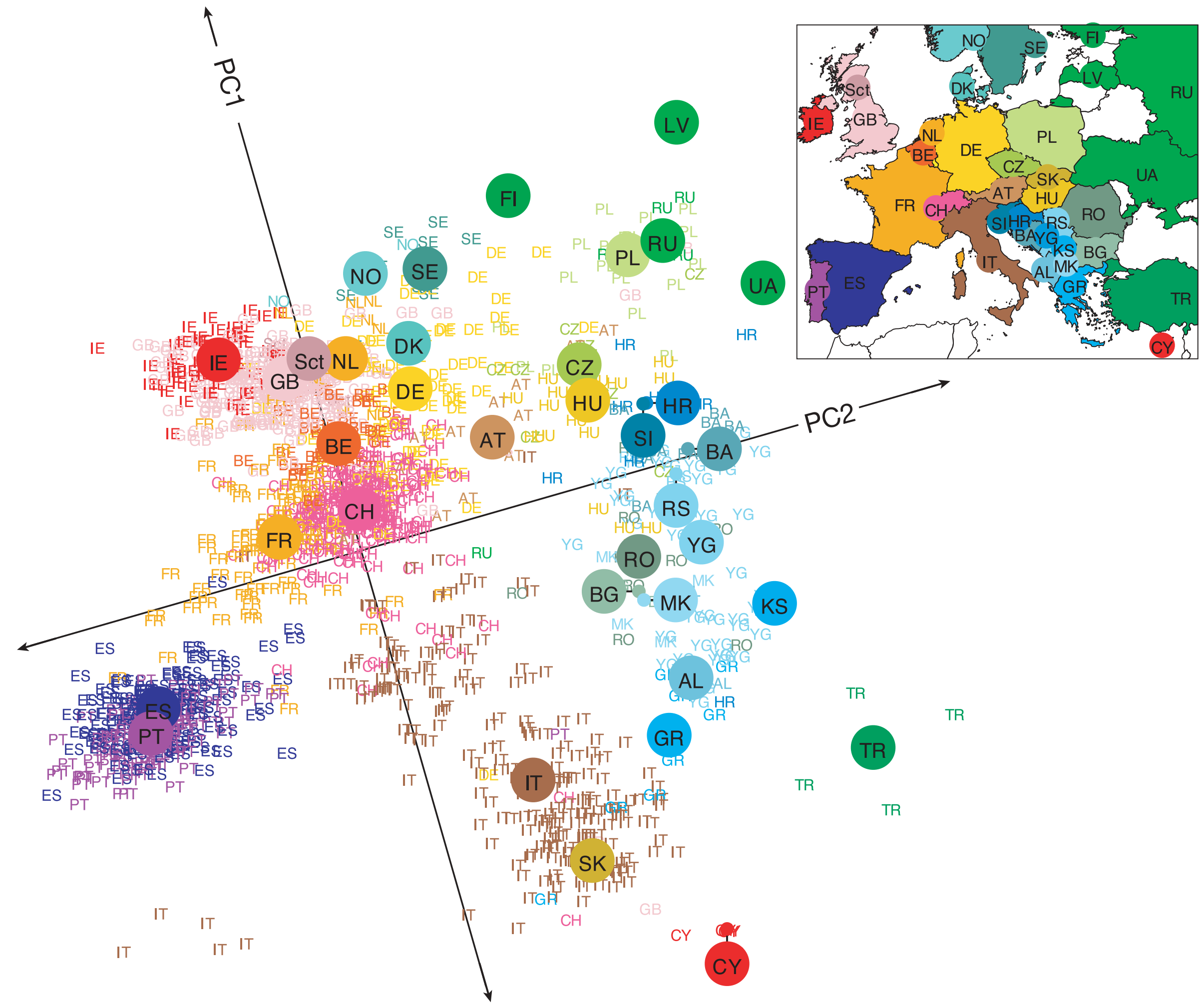
$$\hat{f}_{\ell,k} = \text{sample frequency at locus } \ell \text{ in population } k$$

$$X_{\ell,k} = \frac{\hat{f}_{\ell,k} - \epsilon_\ell}{\sqrt{\epsilon_\ell(1-\epsilon_\ell)}} \quad = \quad \text{standardized sample frequency at locus } \ell \text{ in population } k$$

$$\widehat{\Omega} = \frac{1}{L} X X^T \quad = \quad \text{standardized sample covariance across all } L \text{ loci}$$

Isolation by Distance

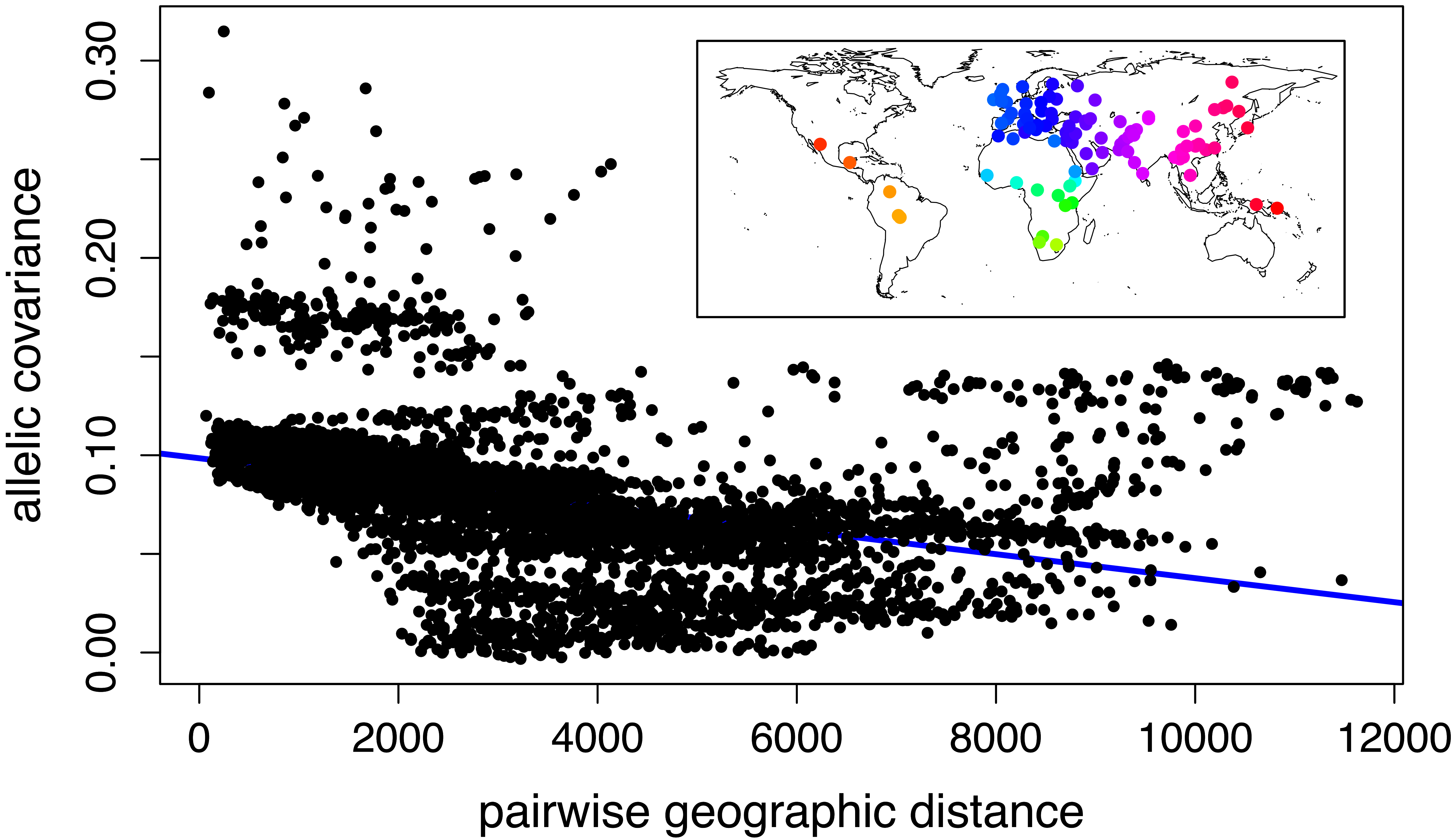
Wright 1943



“Everything is related to everything else,
but near things are more related than distant things”

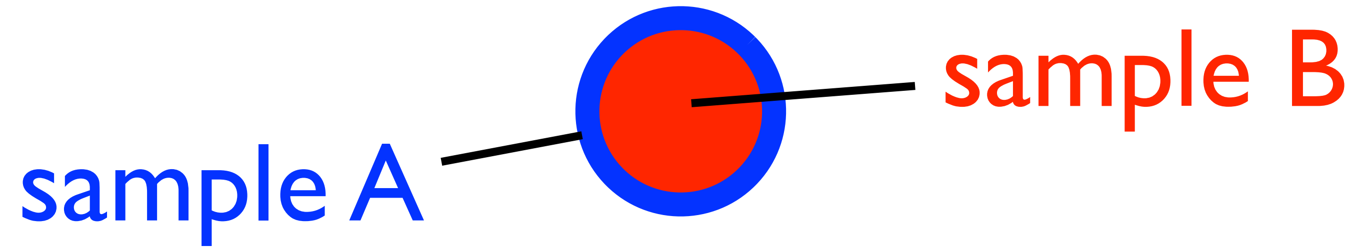
-Waldo Tobler 1970

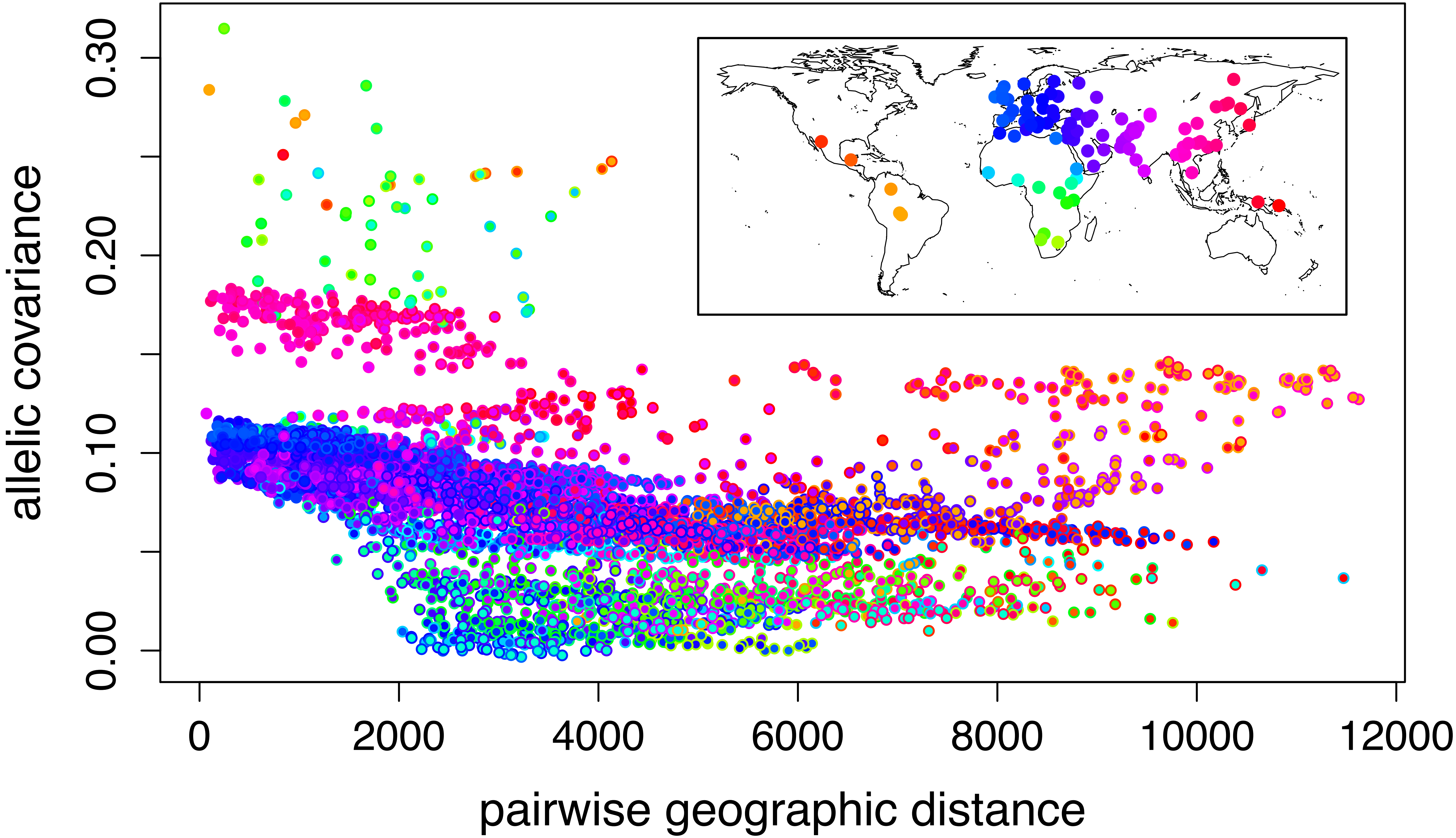
Isolation by Distance: empirical example



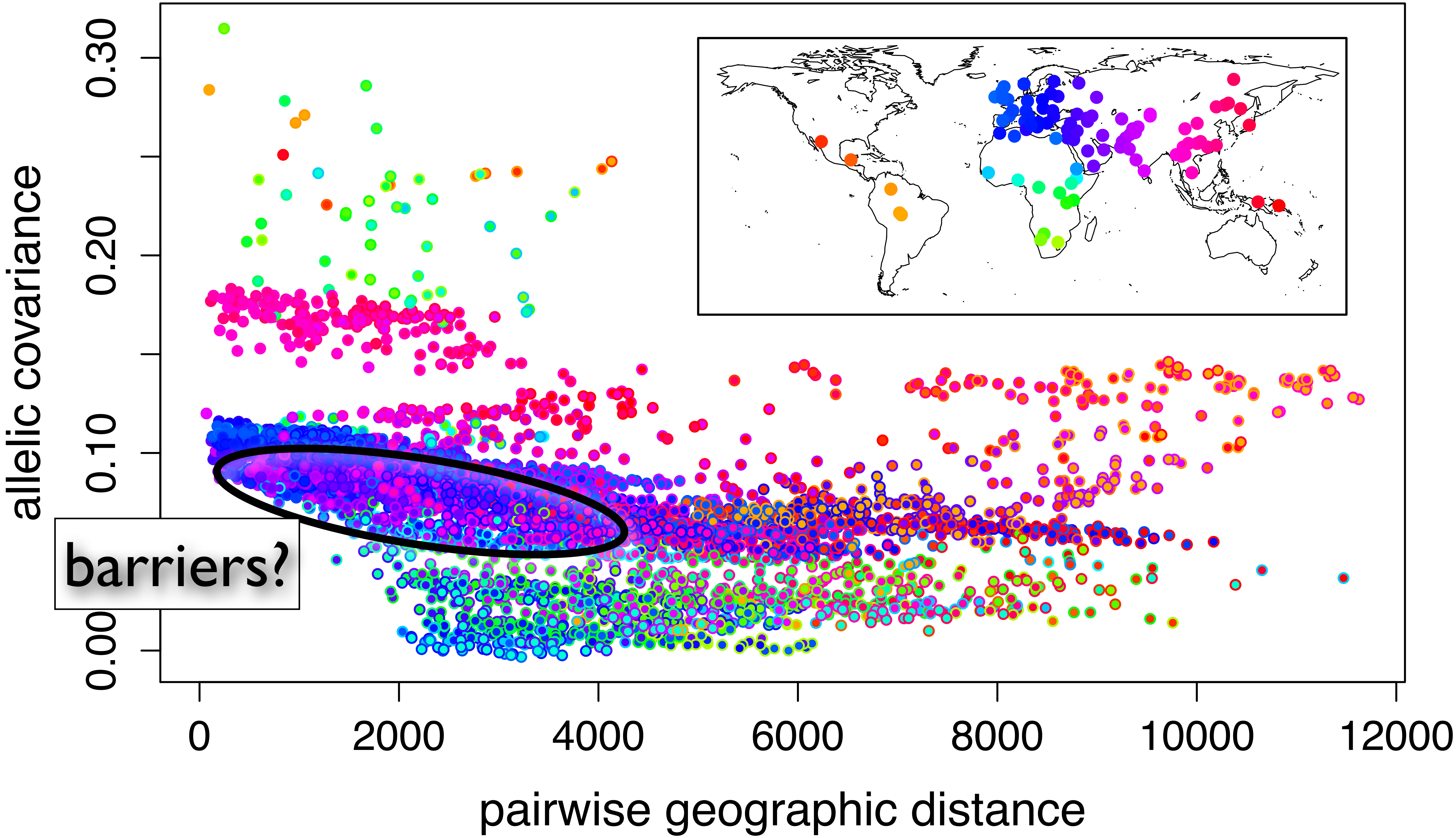
Hellenthal *et al.* 2014 *Science*

Isolation by Distance: empirical example

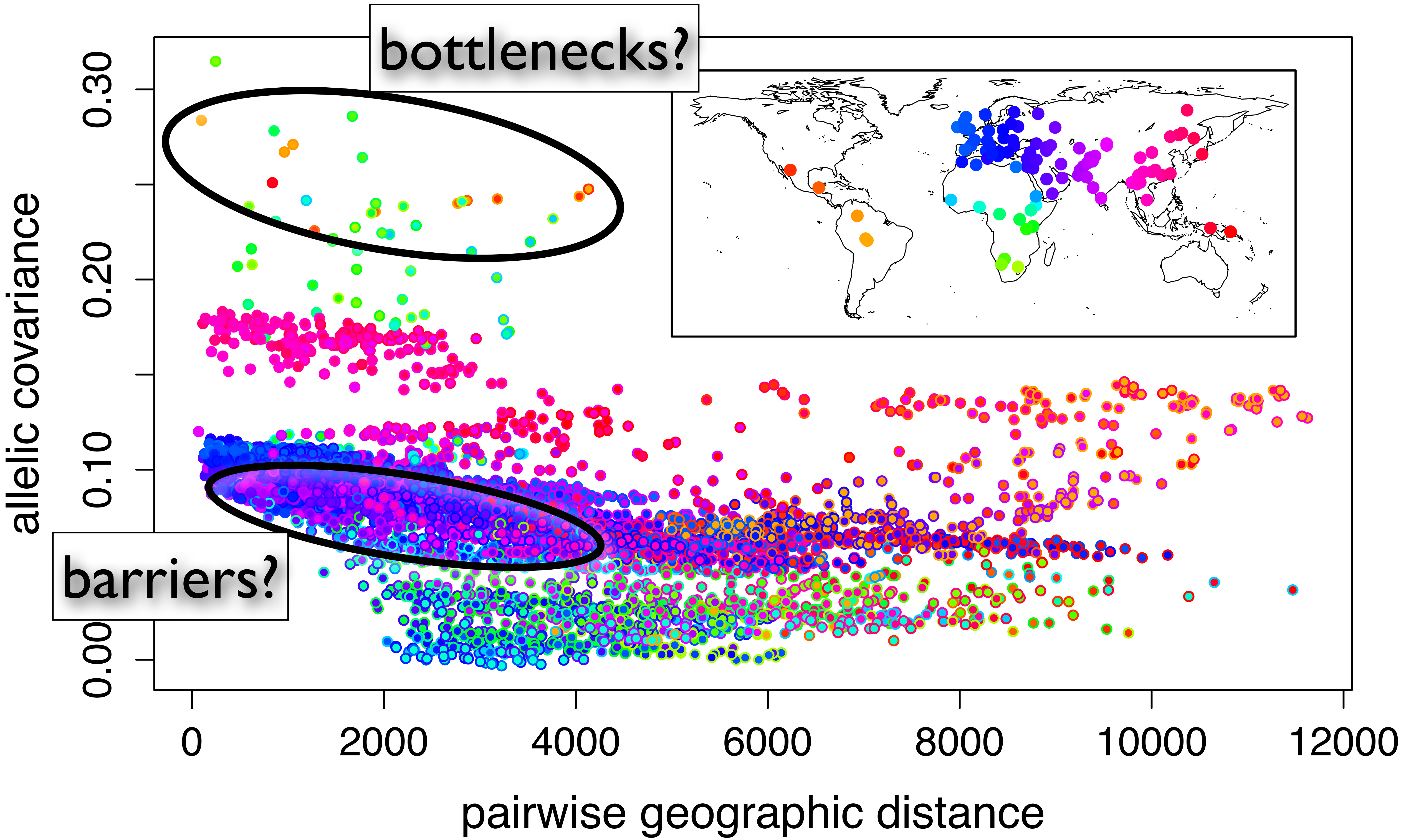




Isolation by Distance: empirical example



Isolation by Distance: empirical example



SpaceMix Model

$$X \sim \text{spatial Gaussian process}$$

SpaceMix Model

$X \sim$ spatial Gaussian process

$$\Omega_{i,j} = \frac{1}{\alpha_0} \exp\left(-(\alpha_1 D_{i,j})^{\alpha_2}\right)$$

SpaceMix Model

$$X \sim \text{spatial Gaussian process}$$

$$\Omega_{i,j} = \frac{1}{\alpha_0} \exp\left(-(\alpha_1 D_{i,j})^{\alpha_2}\right)$$

$$P(\widehat{\Omega} \mid \Omega) = \mathcal{W}\left(L\widehat{\Omega} \mid \Omega, L\right)$$

SpaceMix Model

$$X \sim \text{spatial Gaussian process}$$

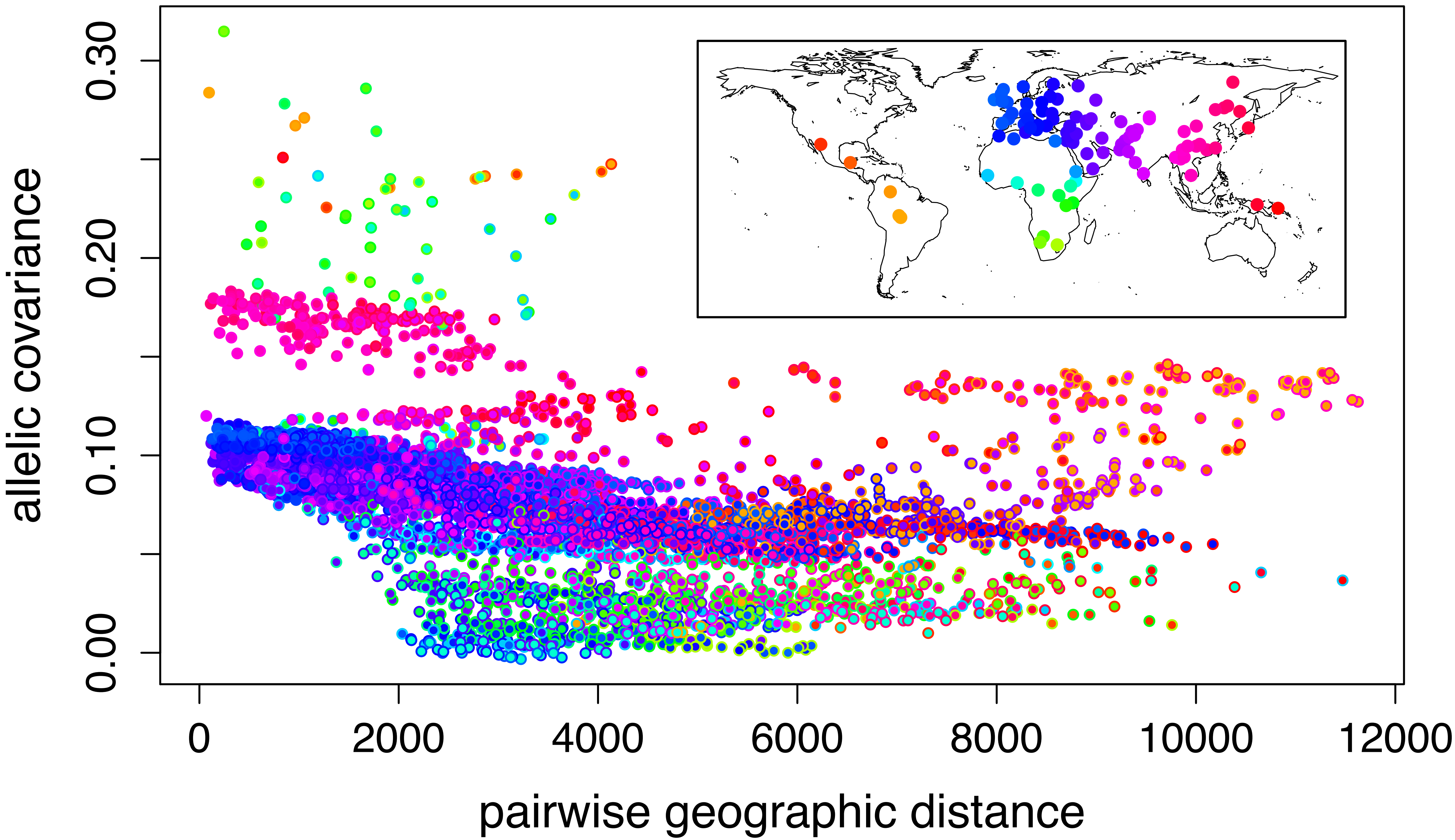
$$\Omega_{i,j} = \frac{1}{\alpha_0} \exp\left(-(\alpha_1 D_{i,j})^{\alpha_2}\right)$$

$$P(\widehat{\Omega} \mid \Omega) = \mathcal{W}\left(L\widehat{\Omega} \mid \Omega, L\right)$$

$$\Omega = \text{function}(\alpha_0, \alpha_1, \alpha_2)$$

$$P\left(\Omega \mid \widehat{\Omega}\right) \propto P(\widehat{\Omega} \mid \Omega, L) P(\vec{\alpha})$$

Isolation by Distance



SpaceMix Model

$$X \sim \text{spatial Gaussian process}$$

$$\Omega_{i,j} = \frac{1}{\alpha_0} \exp\left(-(\alpha_1 D_{i,j})^{\alpha_2}\right)$$

$$P(\widehat{\Omega} \mid \Omega) = \mathcal{W}\left(L\widehat{\Omega} \mid \Omega, L\right)$$

$$\Omega = \text{function}(\alpha_0, \alpha_1, \alpha_2)$$

$$P\left(\Omega \mid \widehat{\Omega}\right) \propto P(\widehat{\Omega} \mid \Omega, L)P(\vec{\alpha})$$

SpaceMix Model

$$X \sim \text{spatial Gaussian process}$$

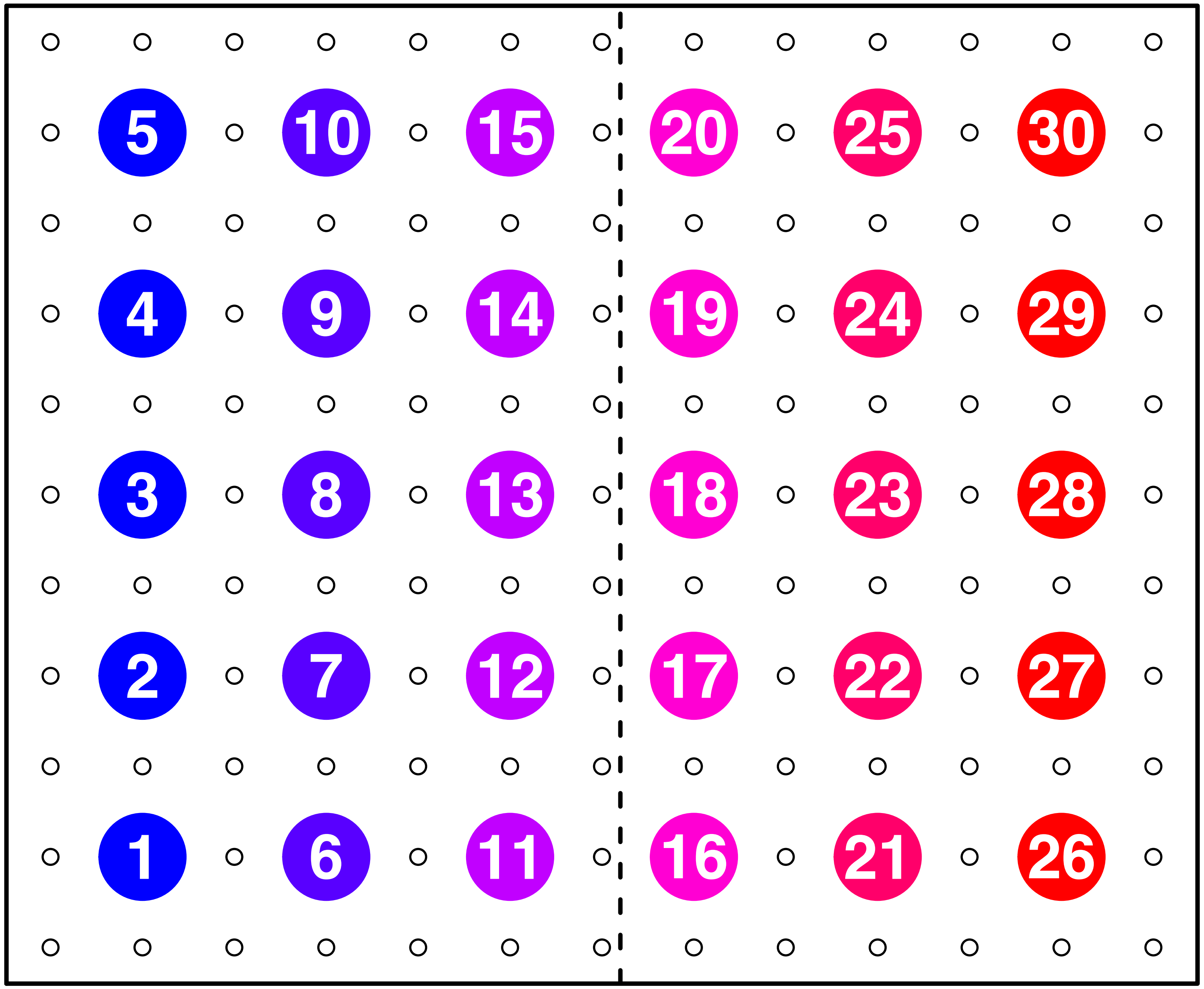
$$\Omega_{i,j} = \frac{1}{\alpha_0} \exp\left(-(\alpha_1 D_{i,j})^{\alpha_2}\right)$$

$$P(\widehat{\Omega} \mid \Omega) = \mathcal{W}\left(L\widehat{\Omega} \mid \Omega, L\right)$$

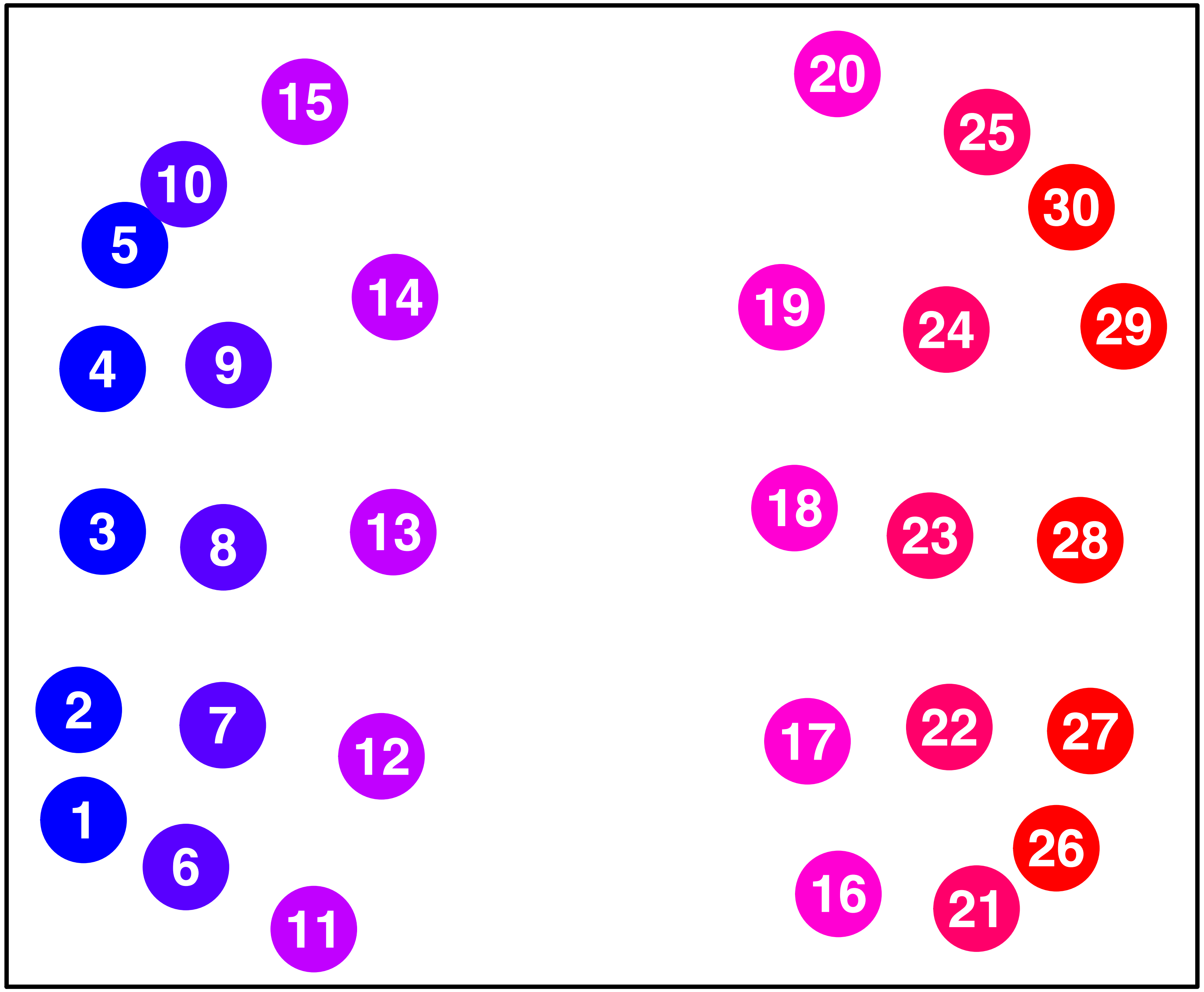
$$\Omega = \text{function}(\alpha_0, \alpha_1, \alpha_2, D(G'))$$

$$P\left(\Omega \mid \widehat{\Omega}\right) \propto P(\widehat{\Omega} \mid \Omega, L) P(\vec{\alpha}) P(G')$$

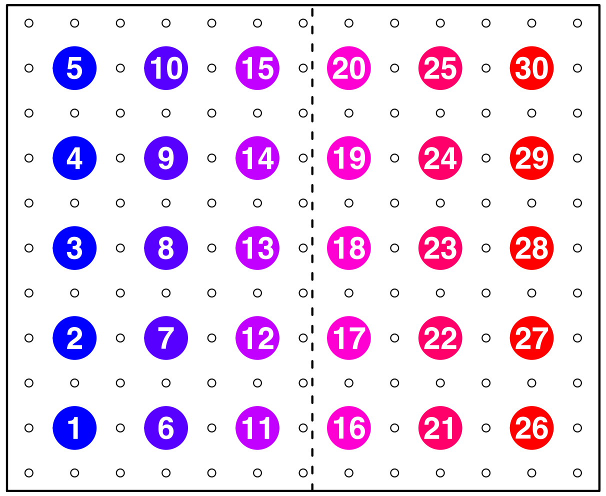
Take this out for a spin: map with a gap

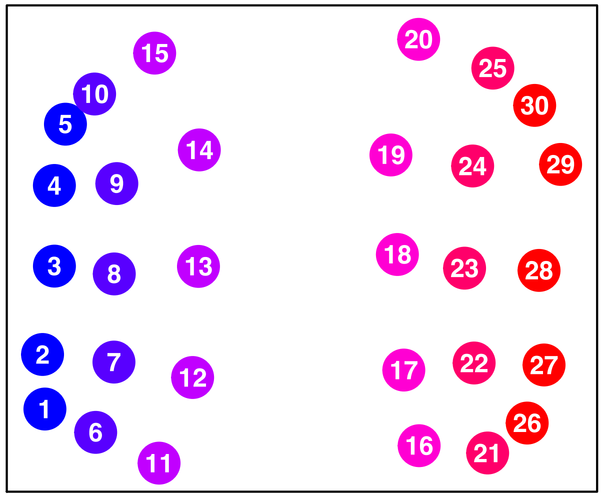


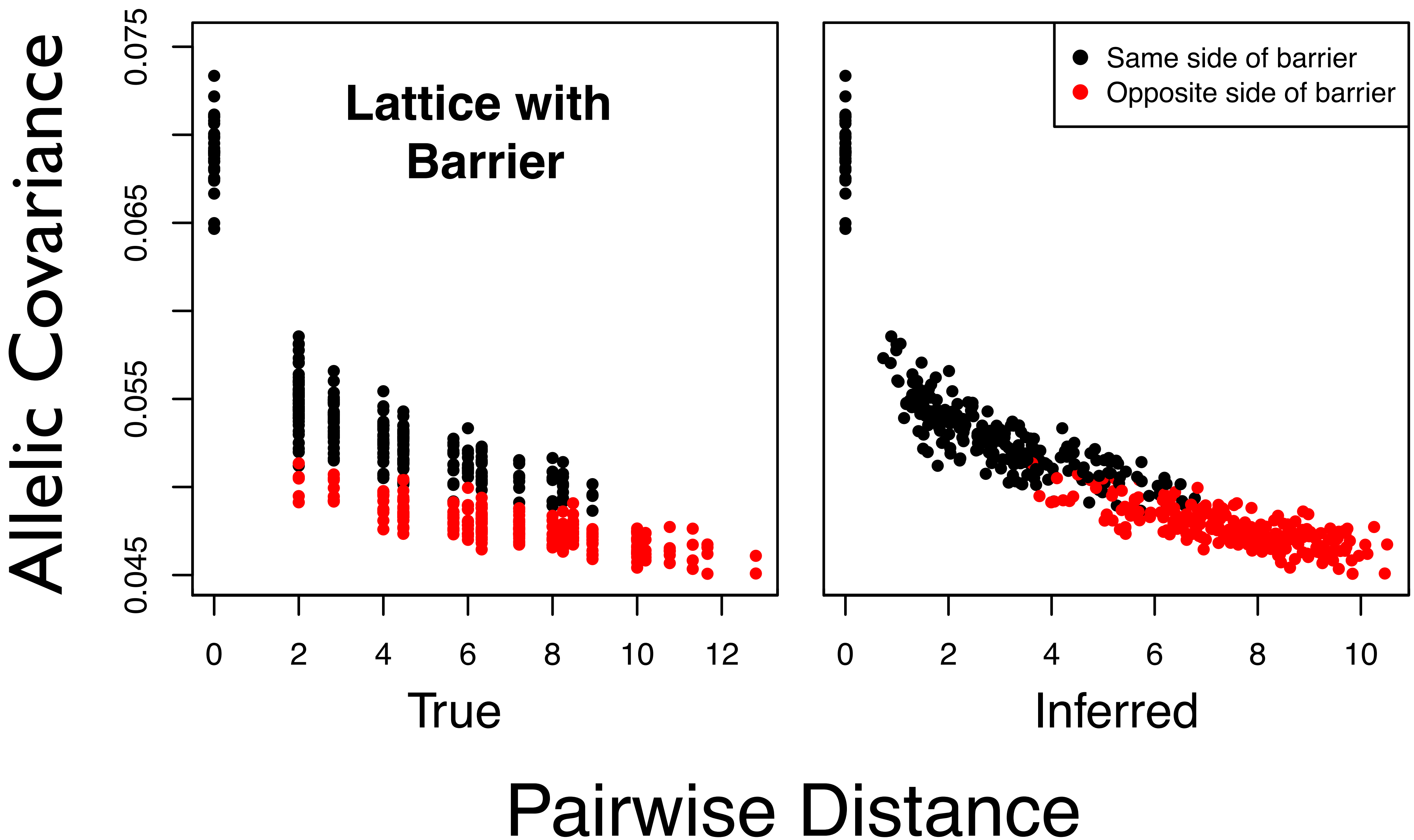
Take this out for a spin: map with a gap



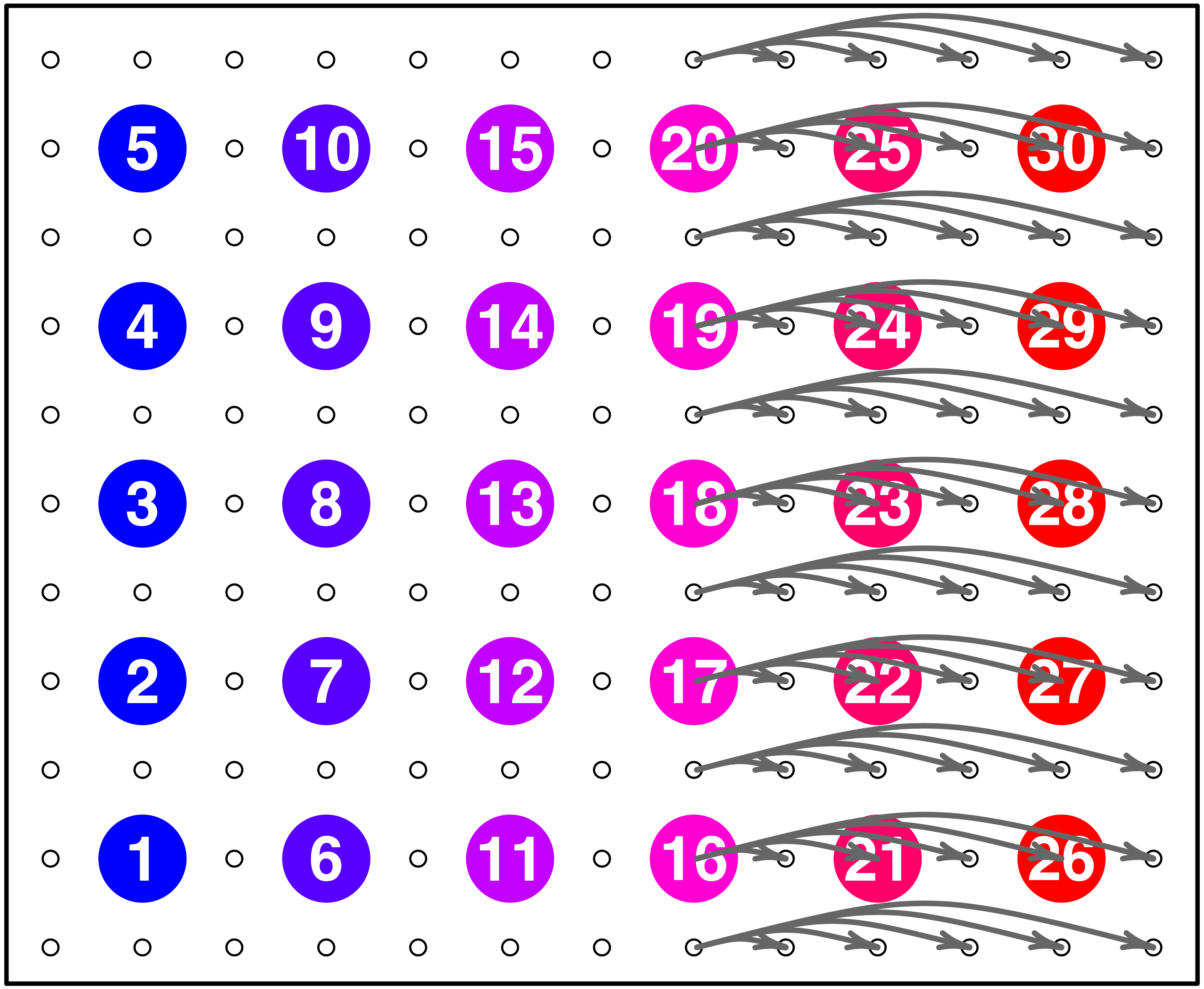
Take this out for a spin: map with a gap



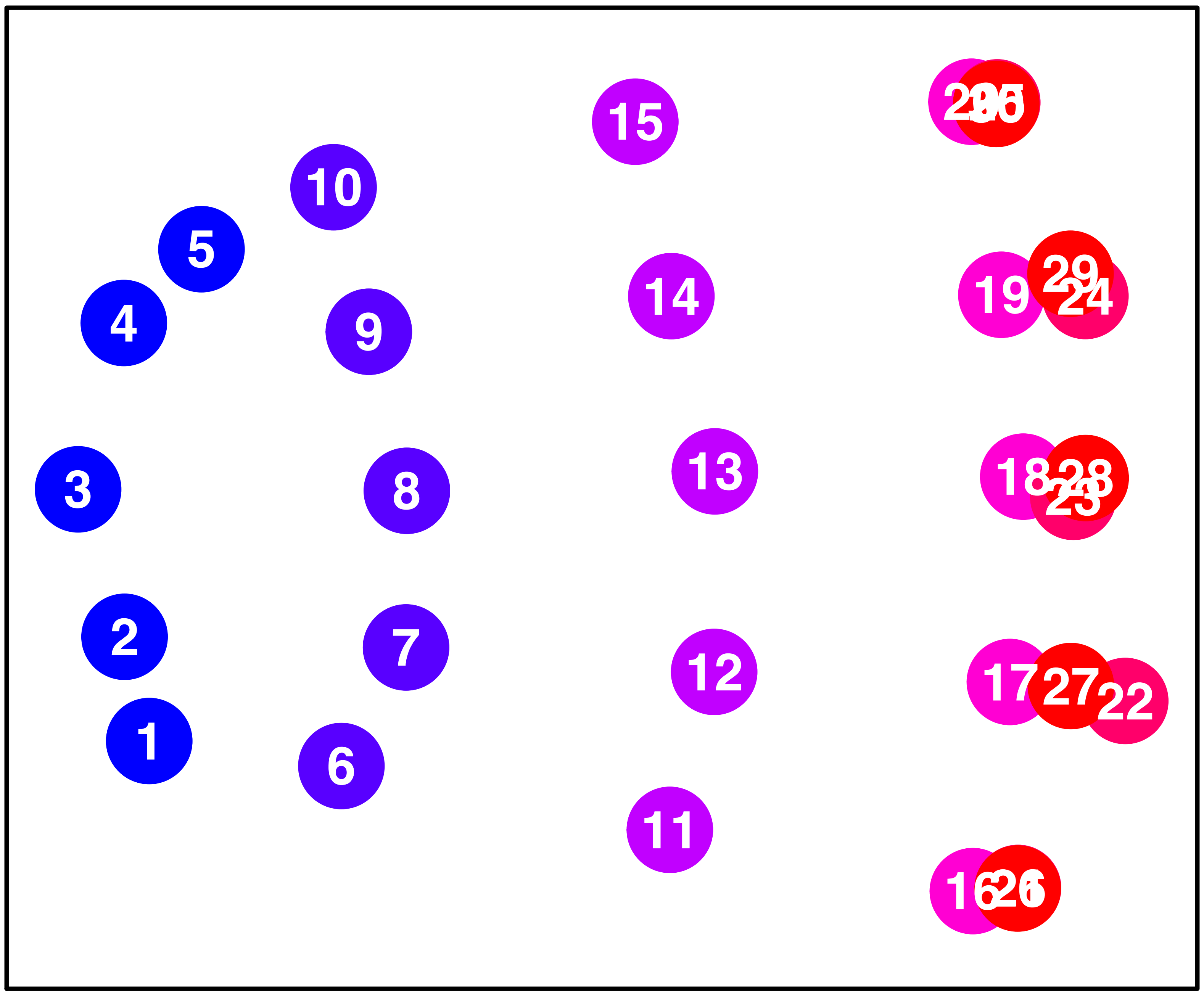




Take this out for a spin: expansion event

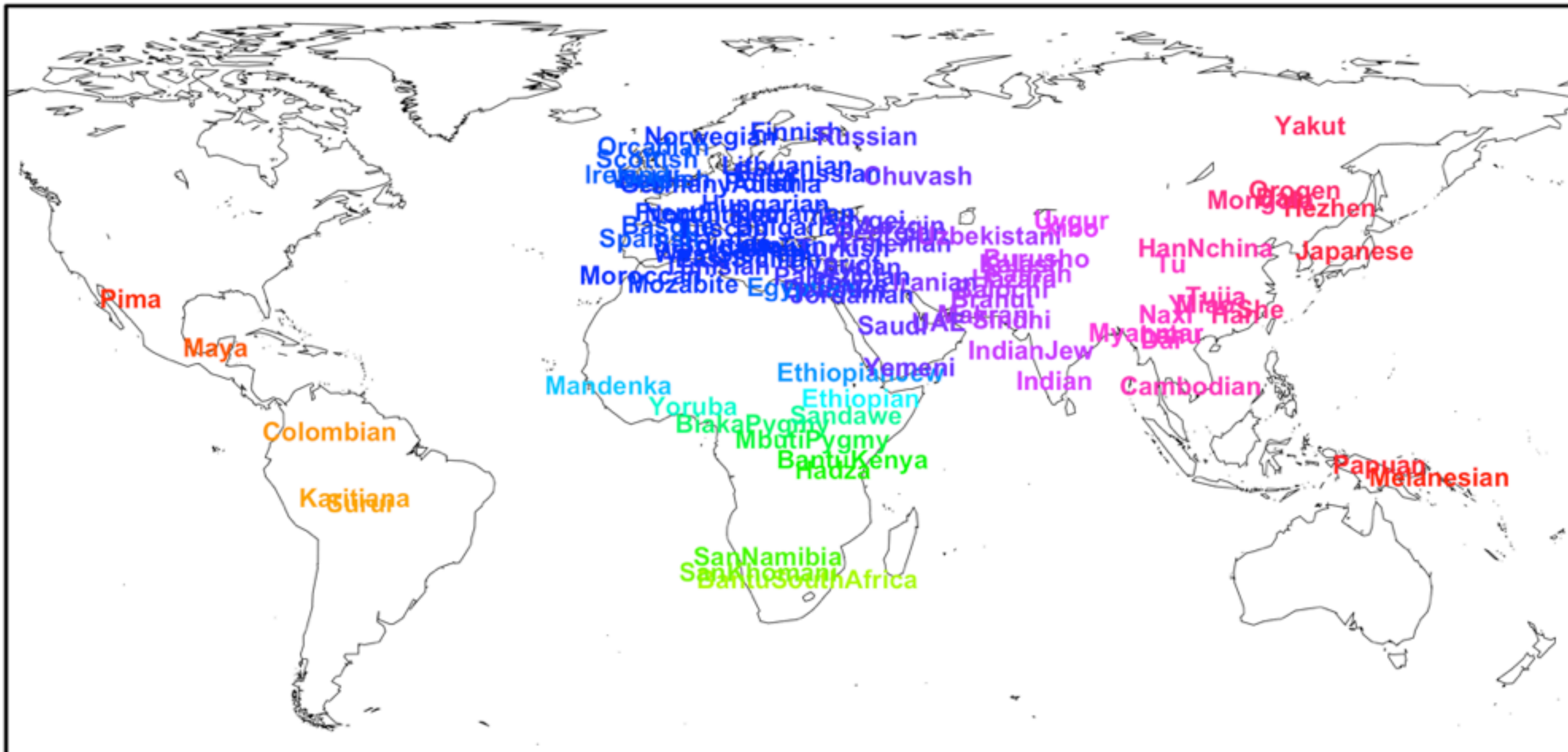


Take this out for a spin: expansion event



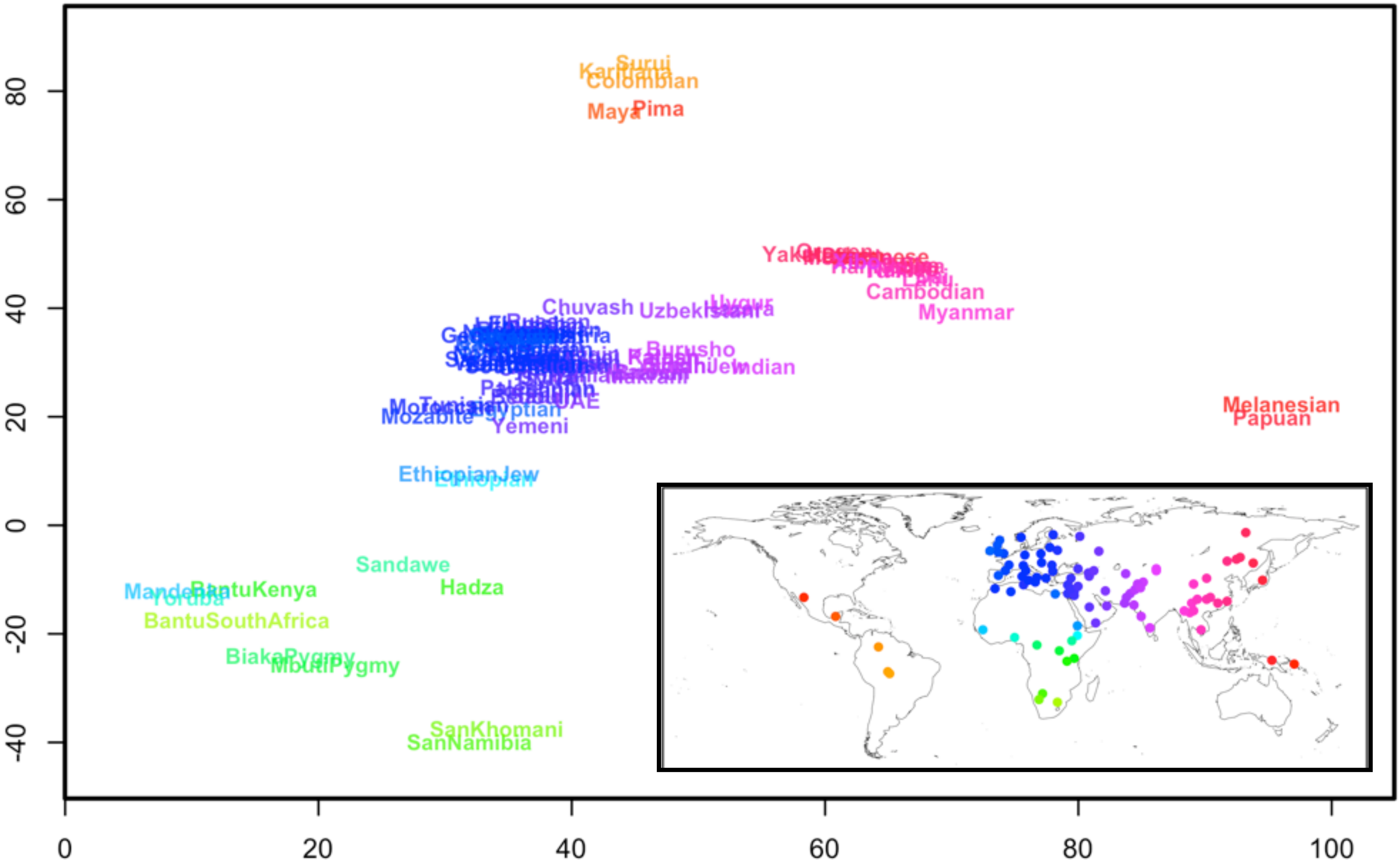
Empirical Application: Humans

Empirical Application: Humans

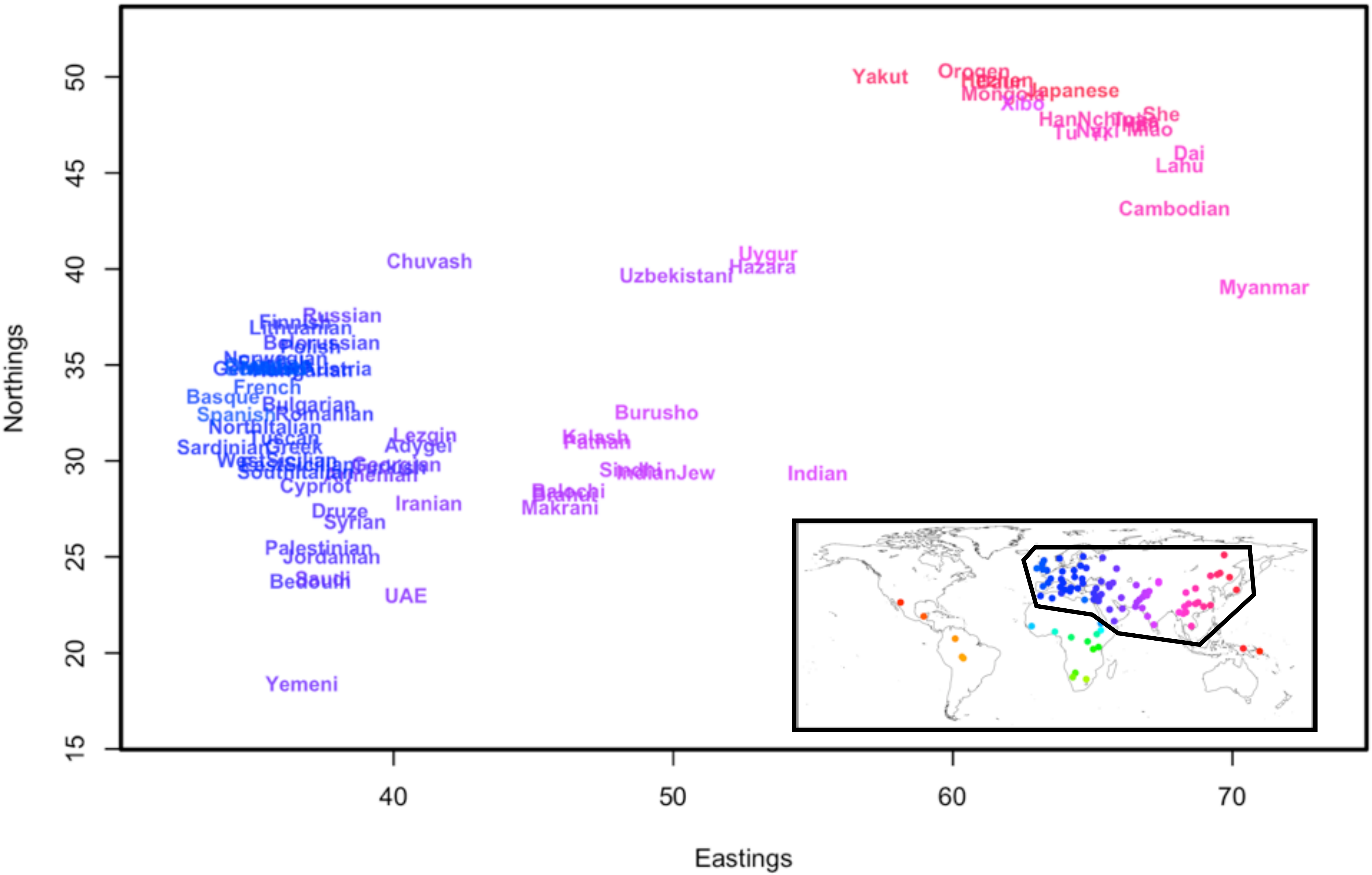


Hellenthal *et al.* 2014 *Science*

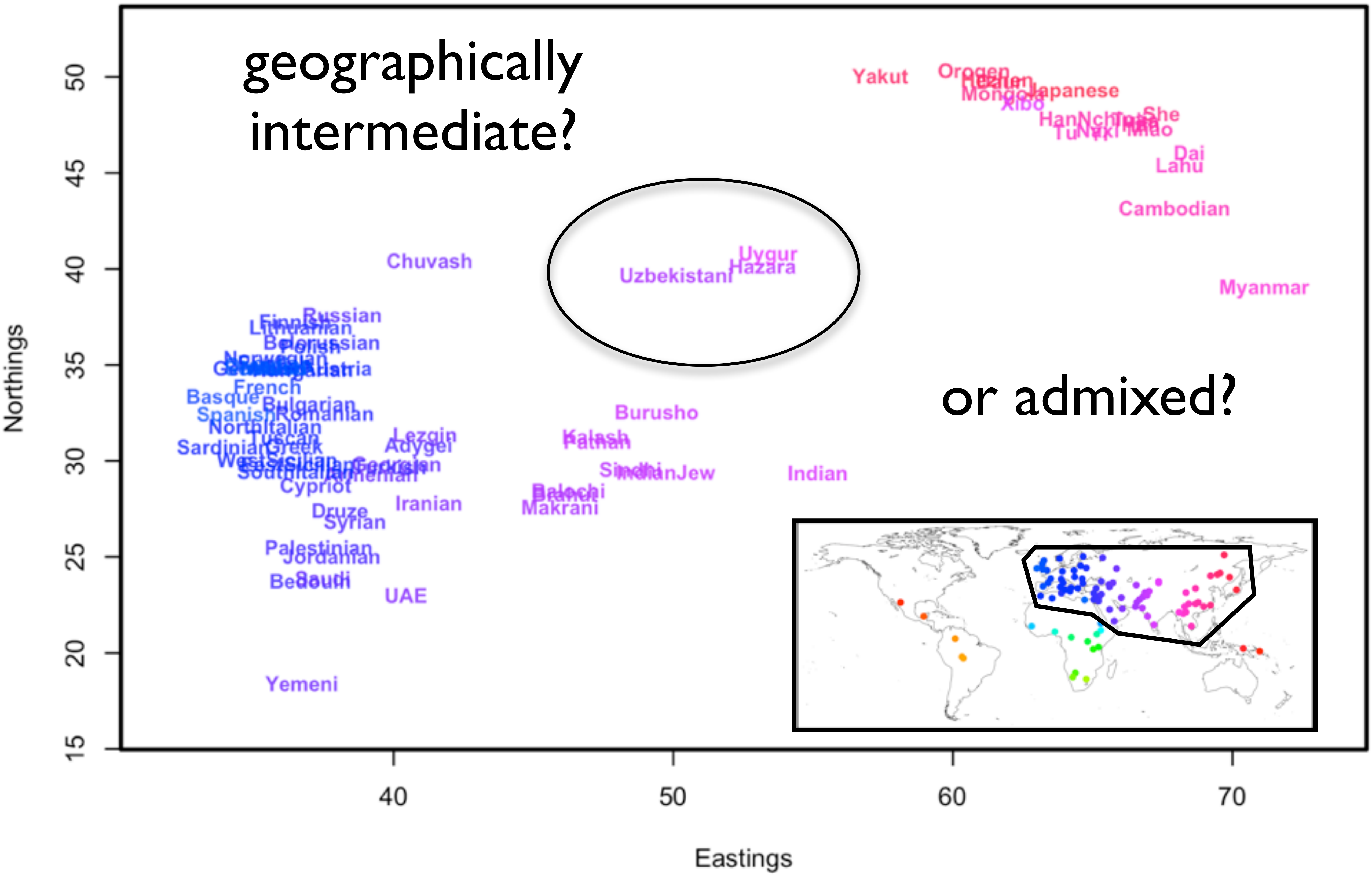
SpaceMix Map: Humans



SpaceMix Map: Humans

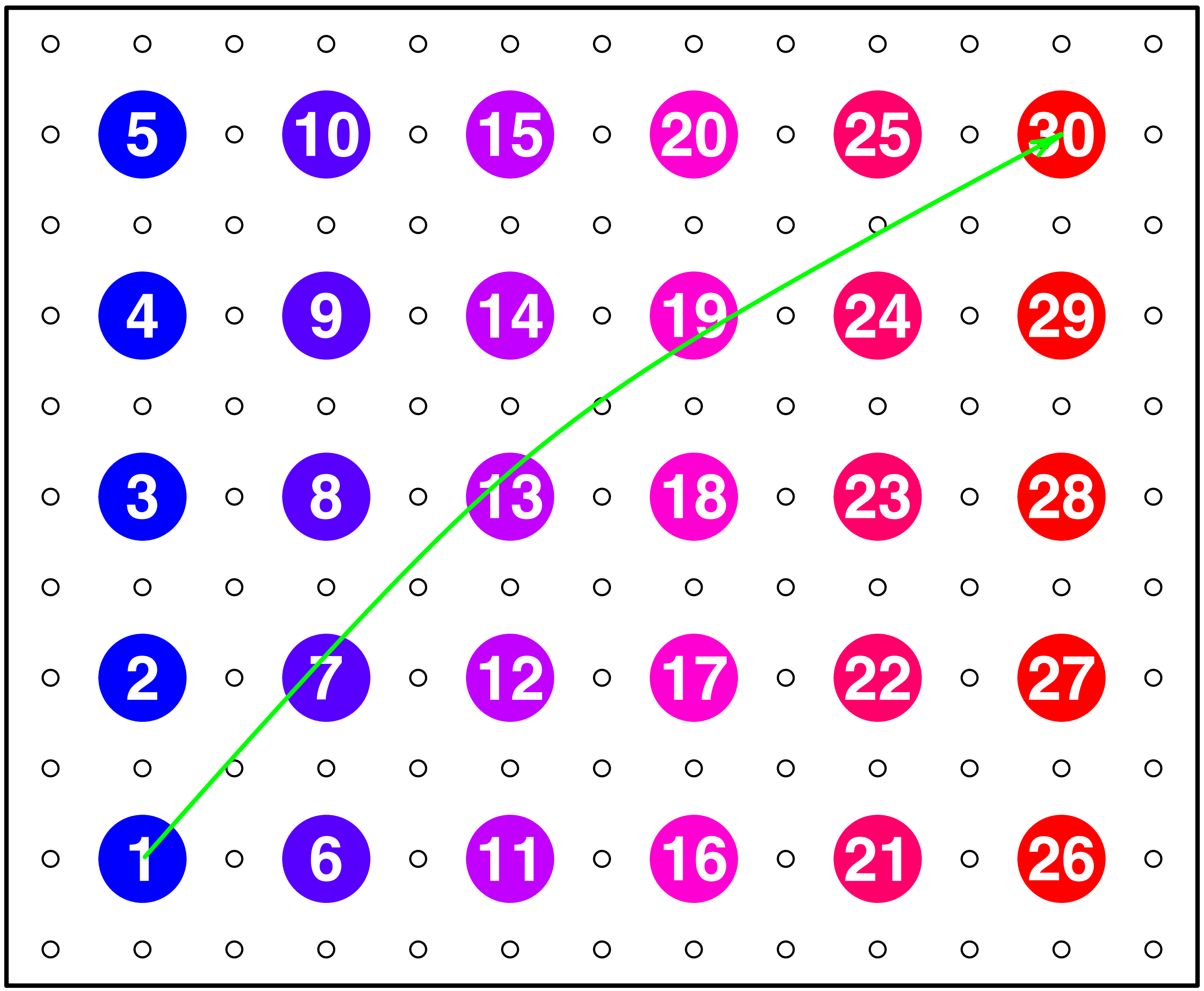


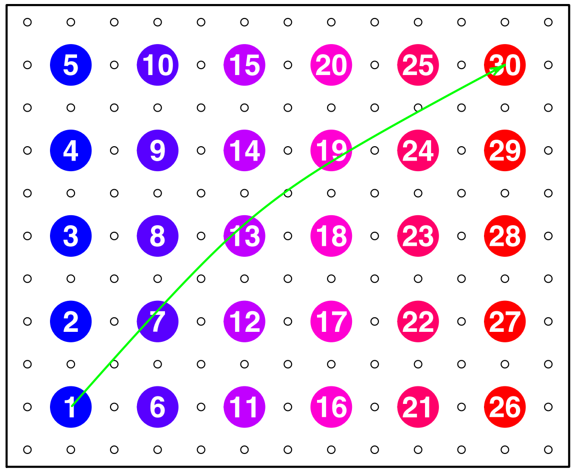
SpaceMix Map: Humans



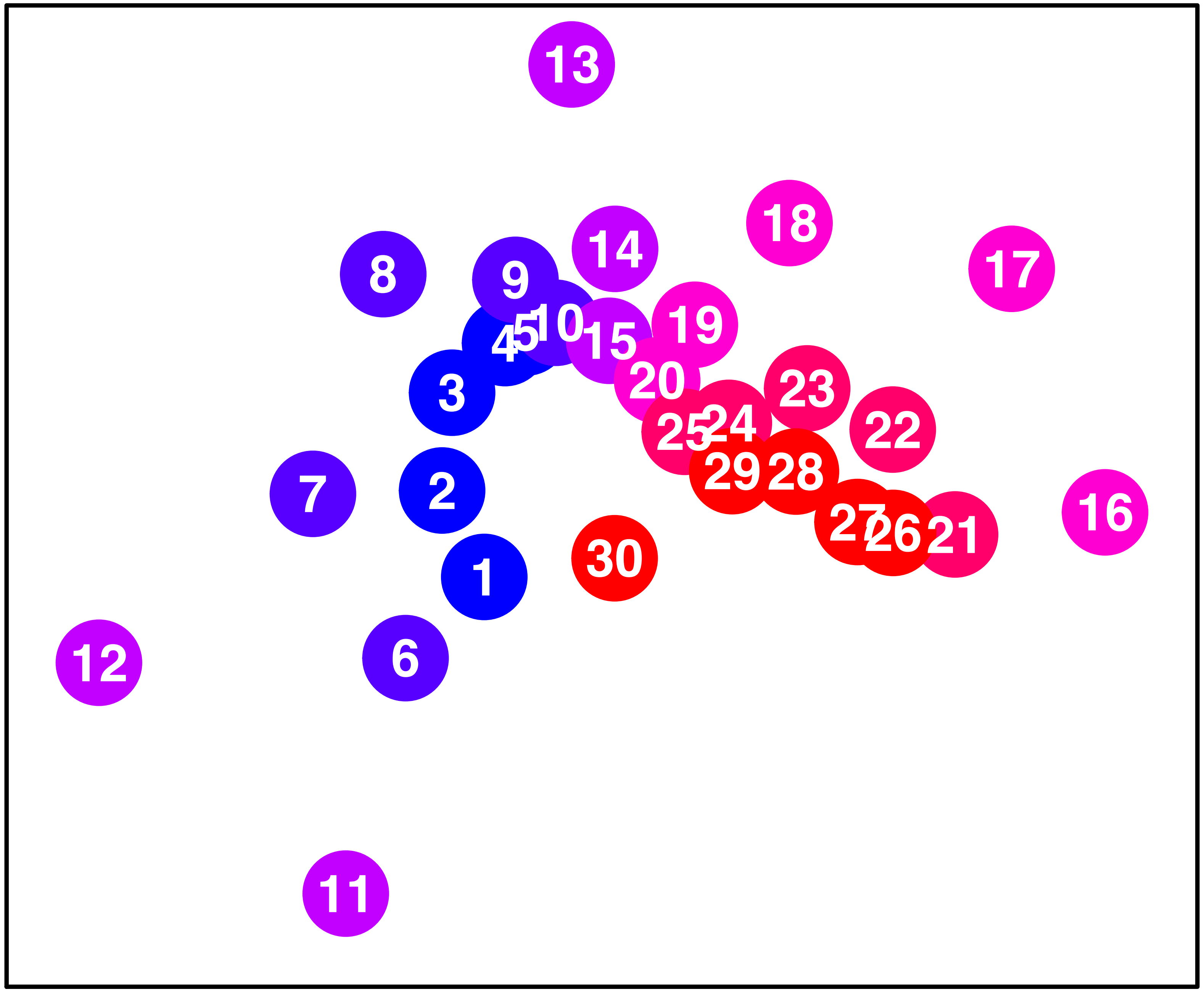
What happens if there's admixture?

What happens if there's admixture?

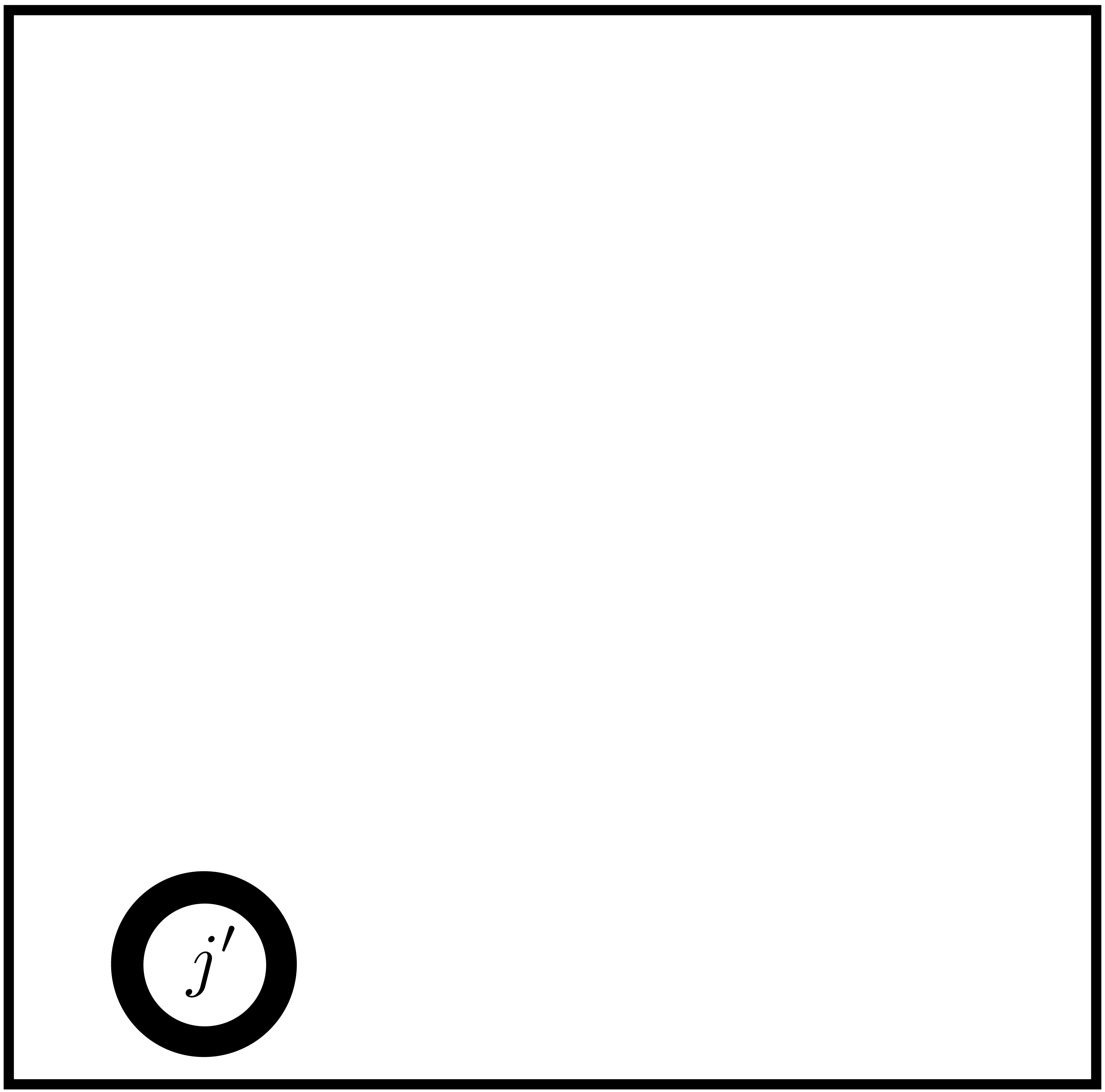




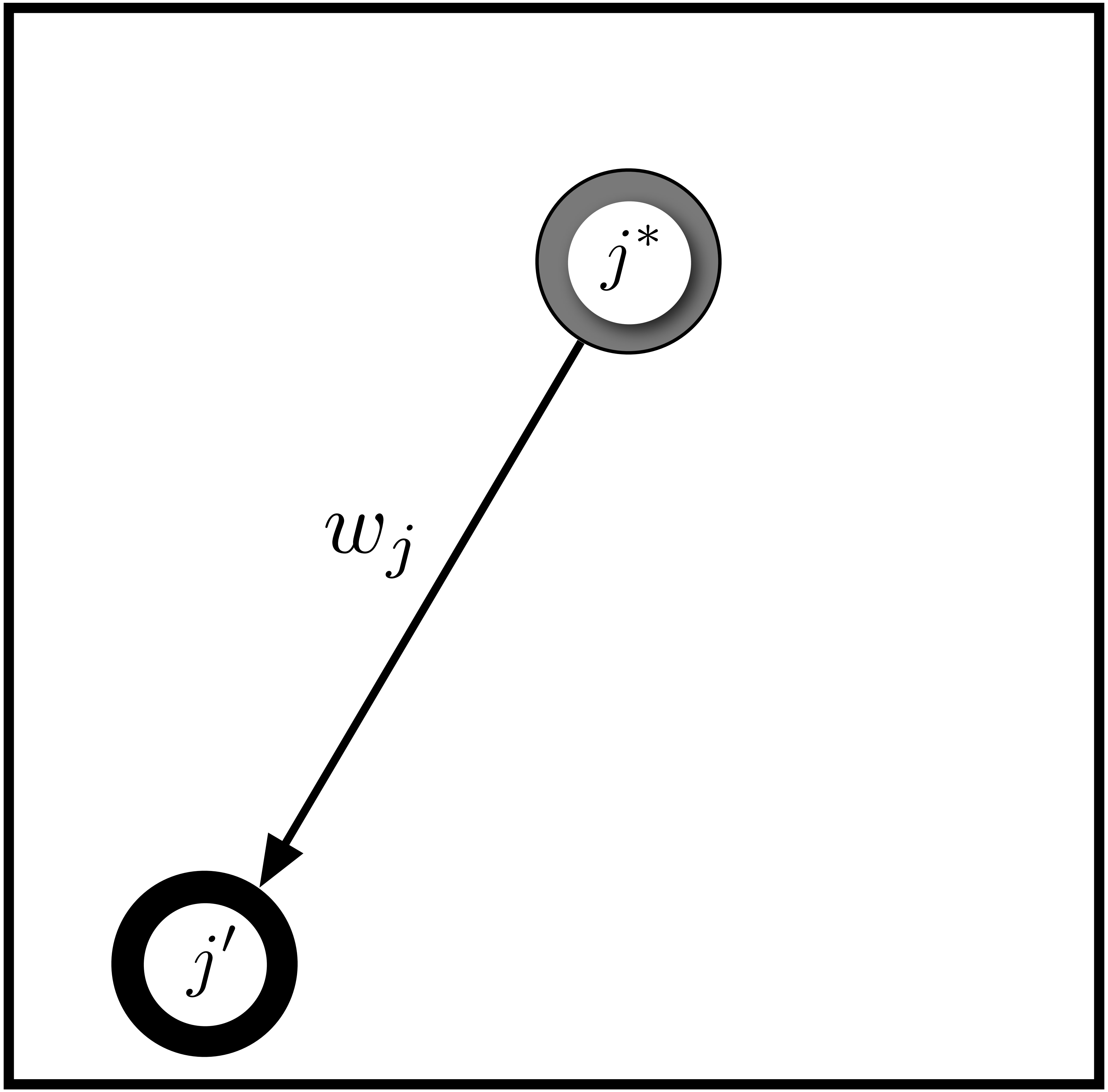
Admixture Simulation: corner admixture event



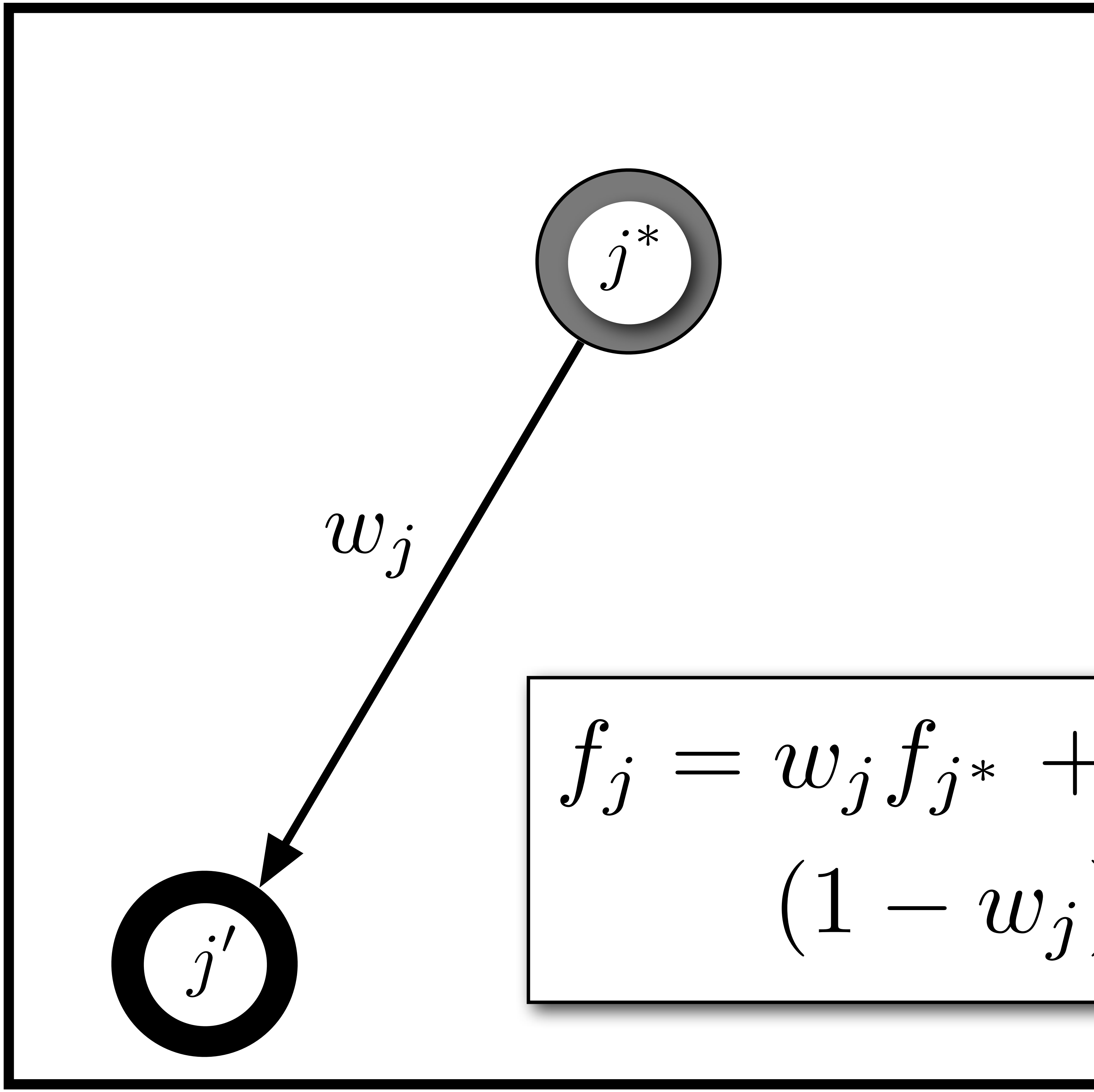
Explicitly modeling admixture



Explicitly modeling admixture

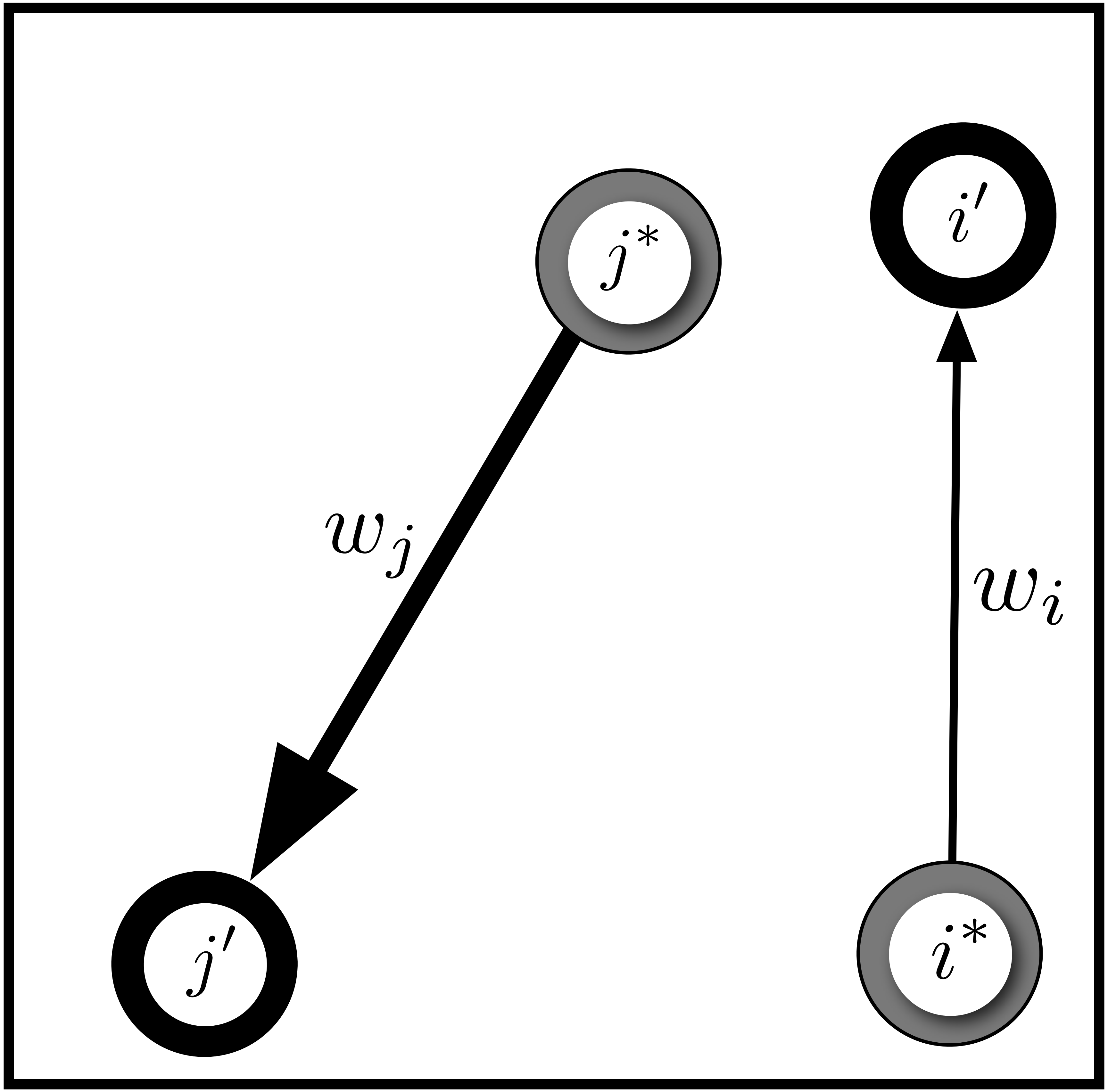


Explicitly modeling admixture



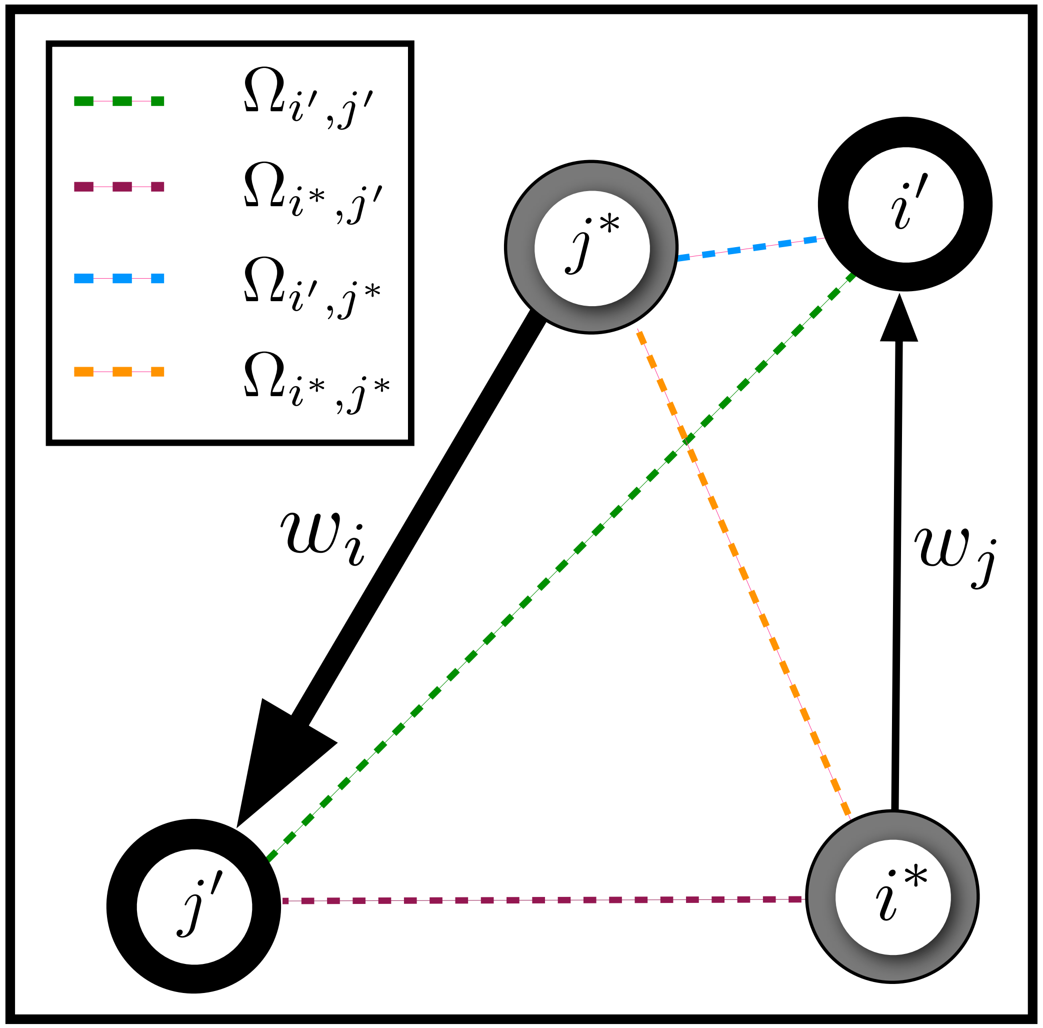
$$f_j = w_j f_{j^*} + (1 - w_j) f_{j'}$$

Explicitly modeling admixture



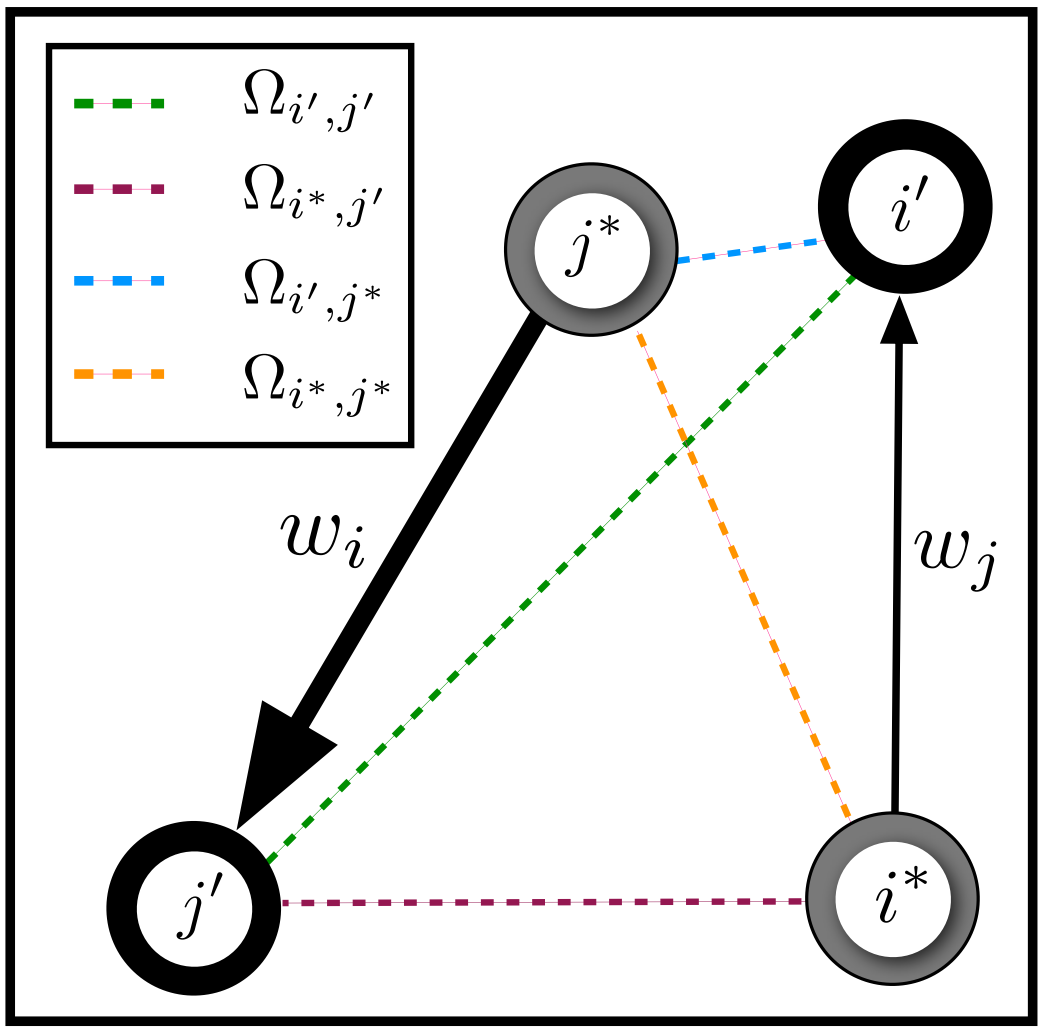
Explicitly modeling admixture

$$\Omega^*_{i,j} = (1 - w_i)(1 - w_j)\Omega(D_{i',j'}) + w_i(1 - w_j)\Omega(D_{i^*,j'}) + w_j(1 - w_i)\Omega(D_{i',j^*}) + w_i w_j \Omega(D_{i^*,j^*})$$



Explicitly modeling admixture

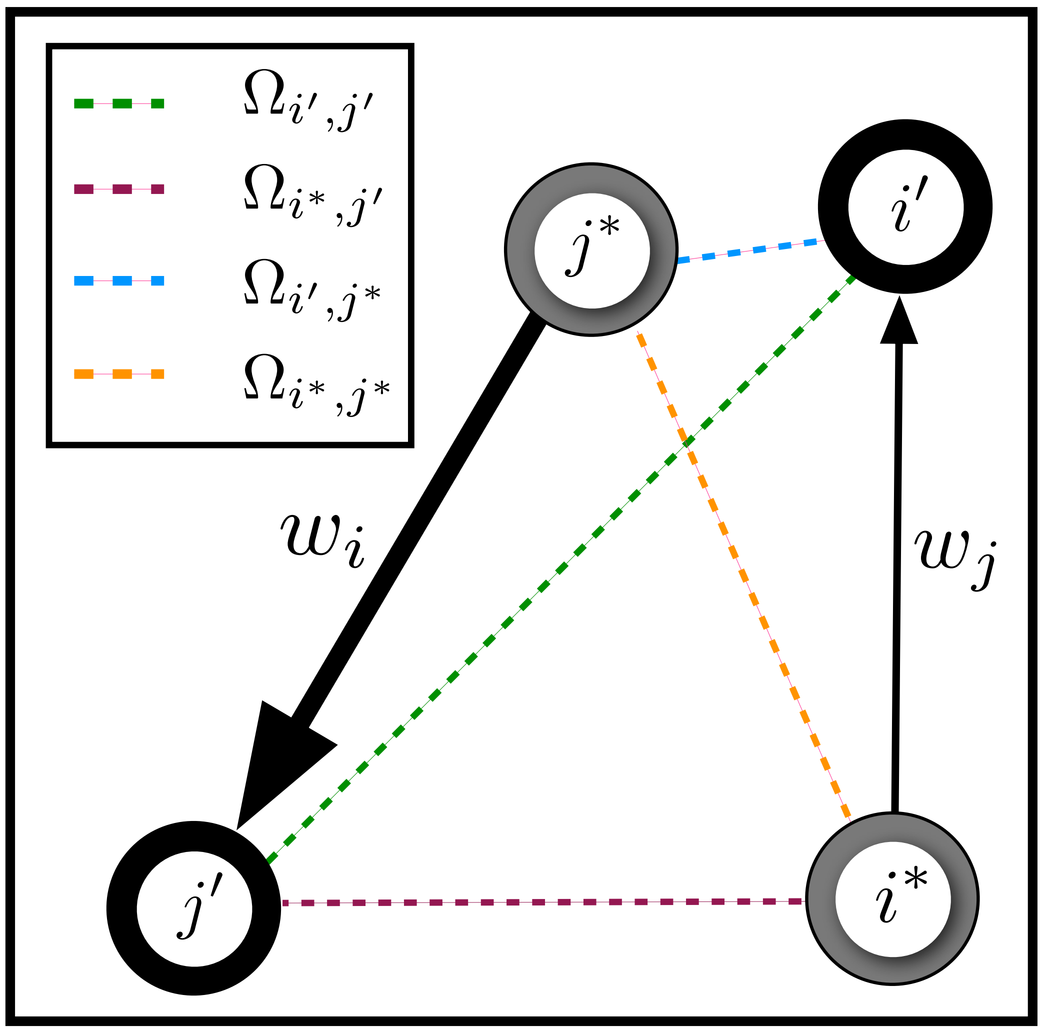
$$\Omega^*_{i,j} = (1-w_i)(1-w_j)\Omega(D_{i',j'}) + w_i(1-w_j)\Omega(D_{i^*,j'}) + w_j(1-w_i)\Omega(D_{i',j^*}) + w_i w_j \Omega(D_{i^*,j^*})$$



$$P\left(\Omega^* \mid \widehat{\Omega}\right) \propto P\left(\widehat{\Omega} \mid \Omega^*, L\right) P(\vec{\alpha}) P(G')$$

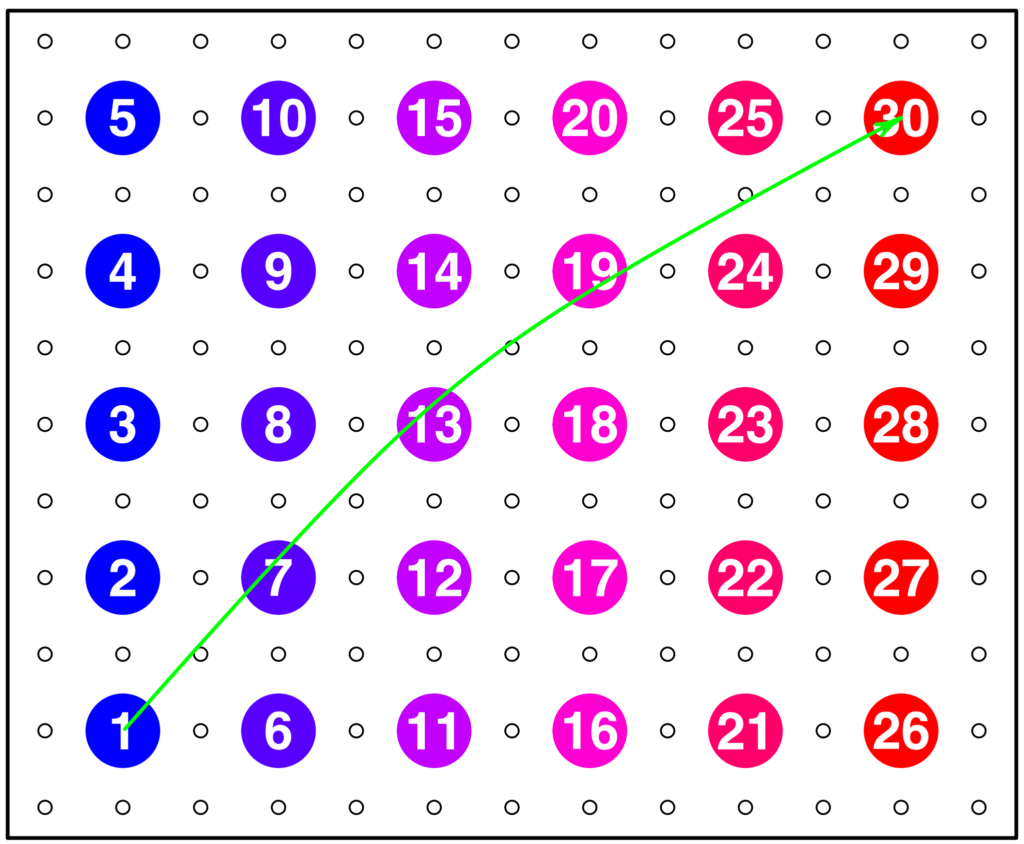
Explicitly modeling admixture

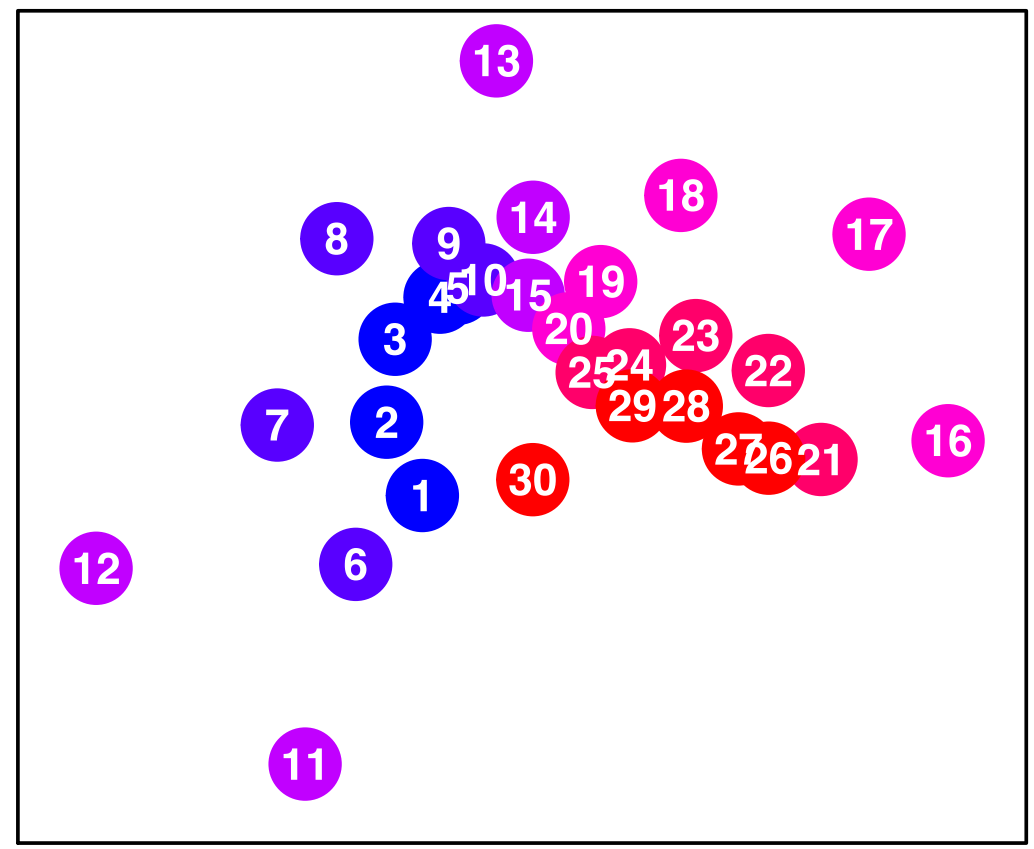
$$\Omega^*_{i,j} = (1-w_i)(1-w_j)\Omega(D_{i',j'}) + w_i(1-w_j)\Omega(D_{i^*,j'}) + w_j(1-w_i)\Omega(D_{i',j^*}) + w_i w_j \Omega(D_{i^*,j^*})$$



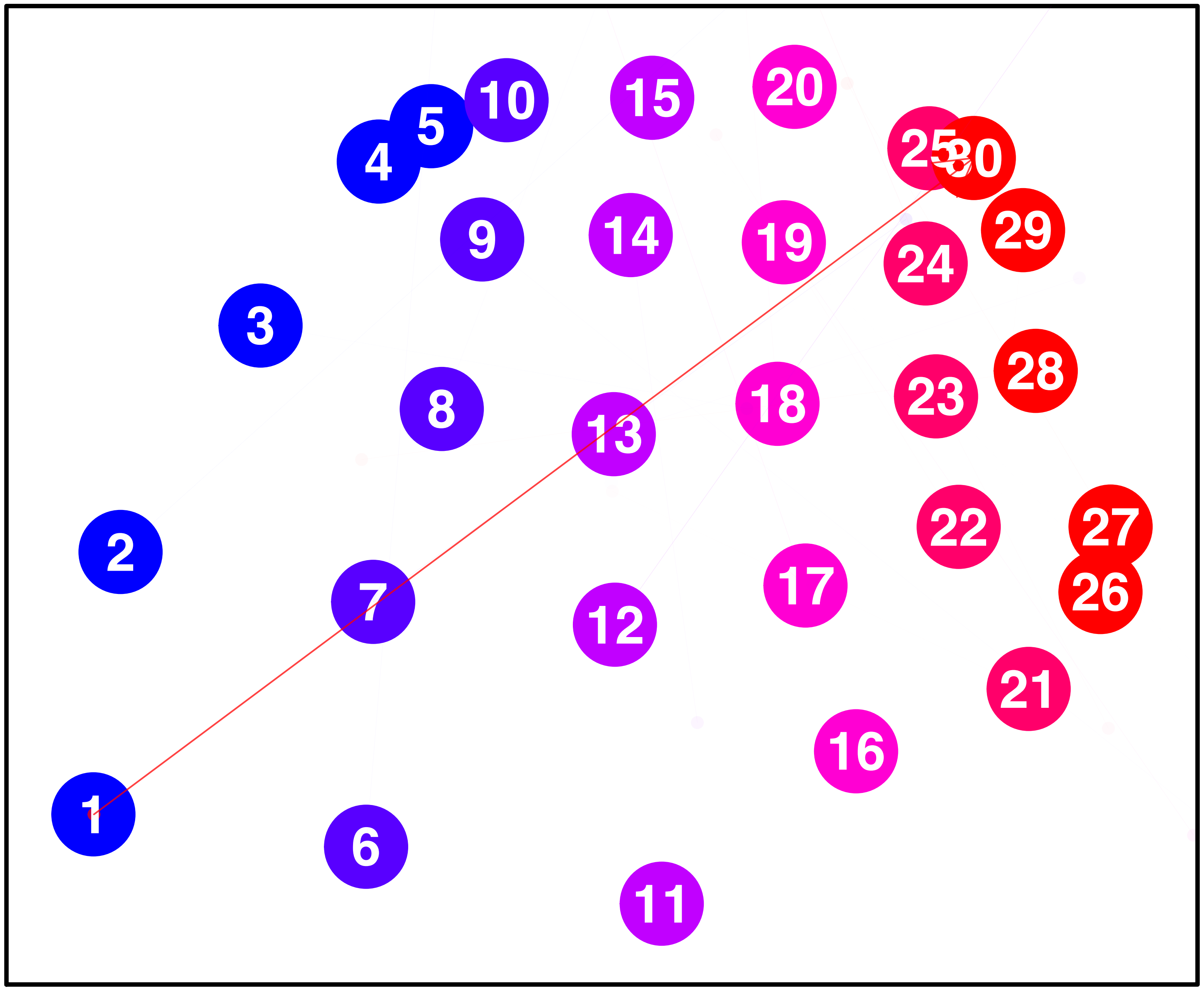
$$P\left(\Omega^* \mid \widehat{\Omega}\right) \propto P\left(\widehat{\Omega} \mid \Omega^*, L\right) P(\vec{\alpha}) P(G') P(G^*) P(w)$$

Admixture Simulation:
corner admixture event

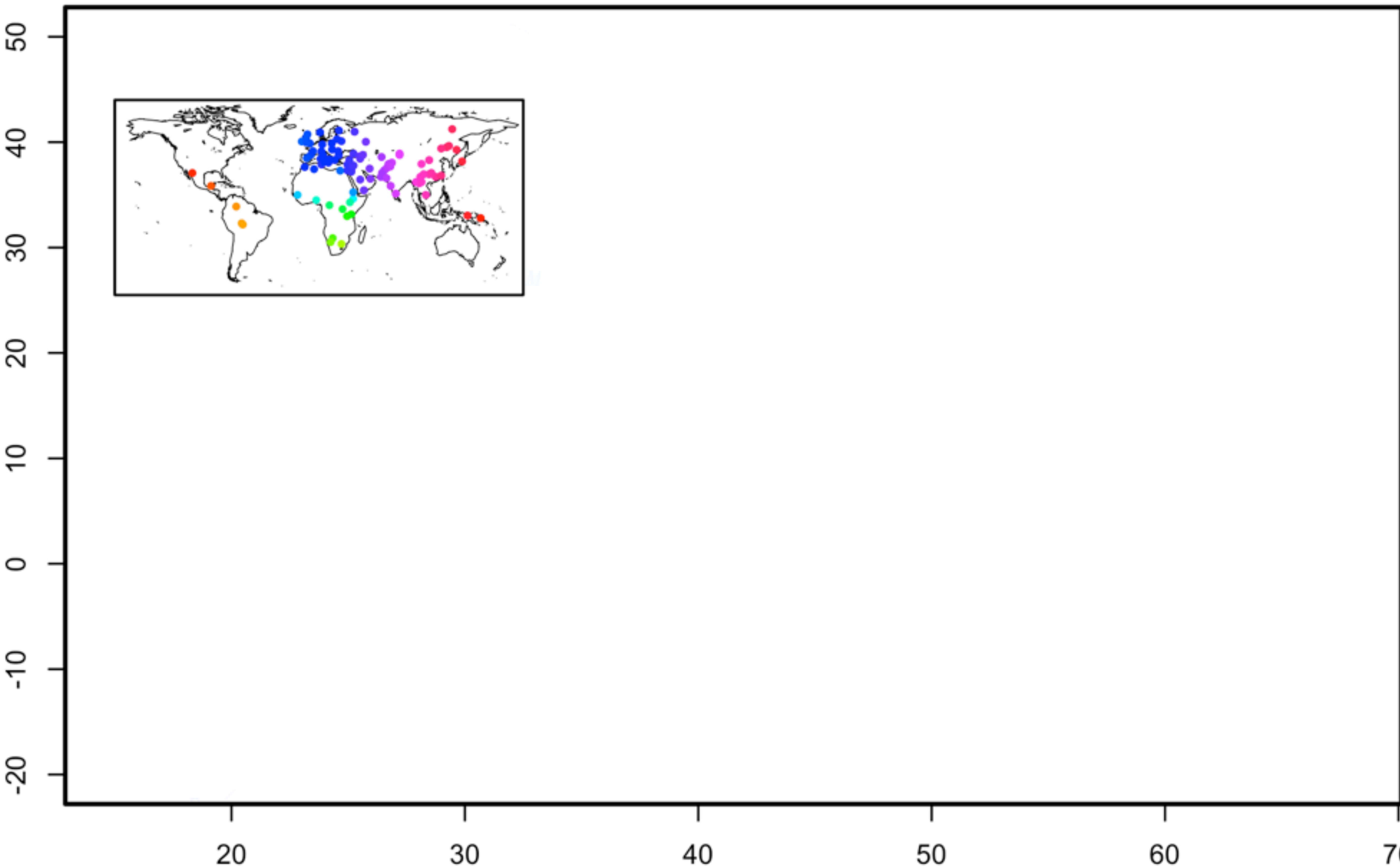




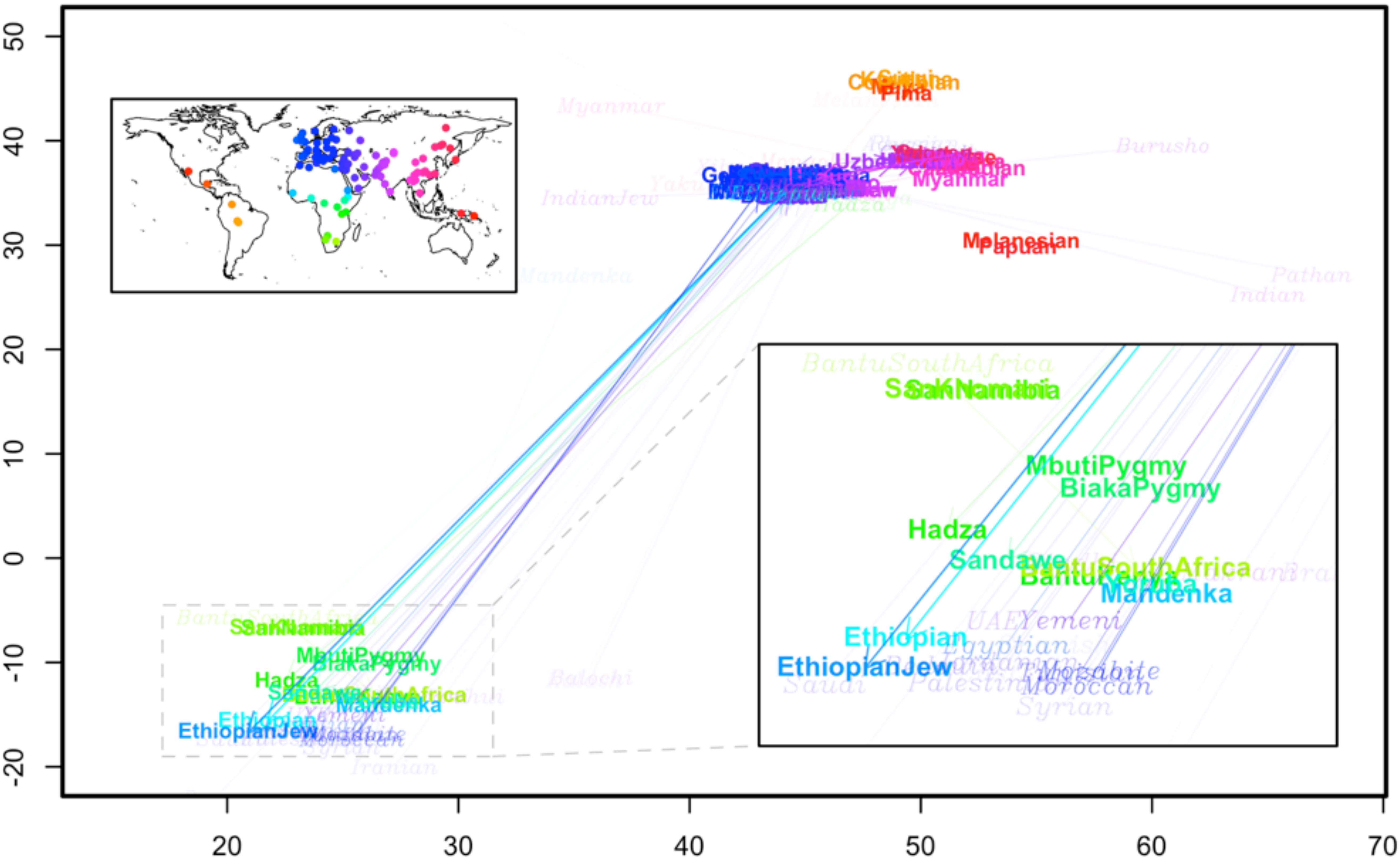
Modeling admixture: corner admixture event



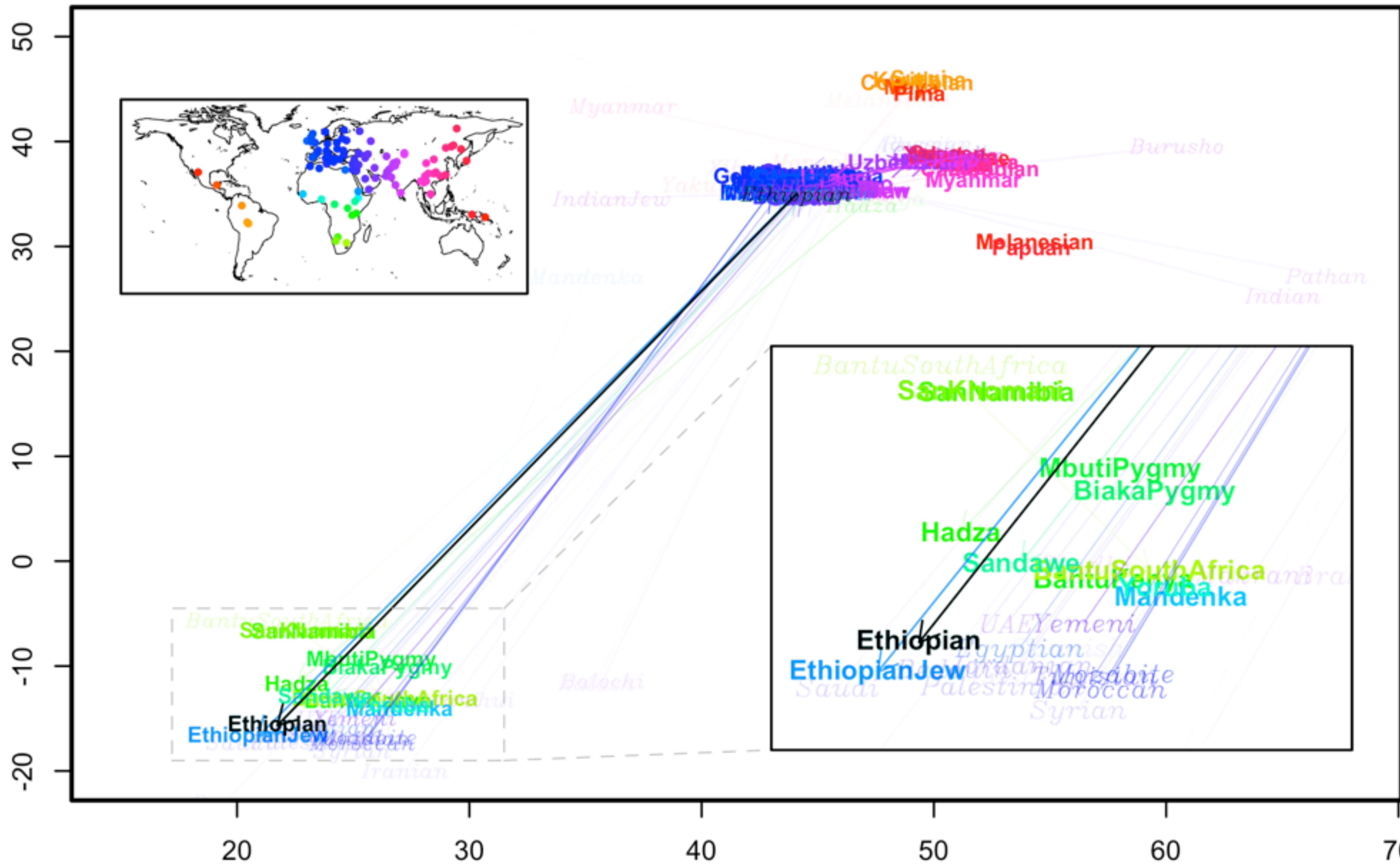
SpaceMix Map: Humans with Admixture



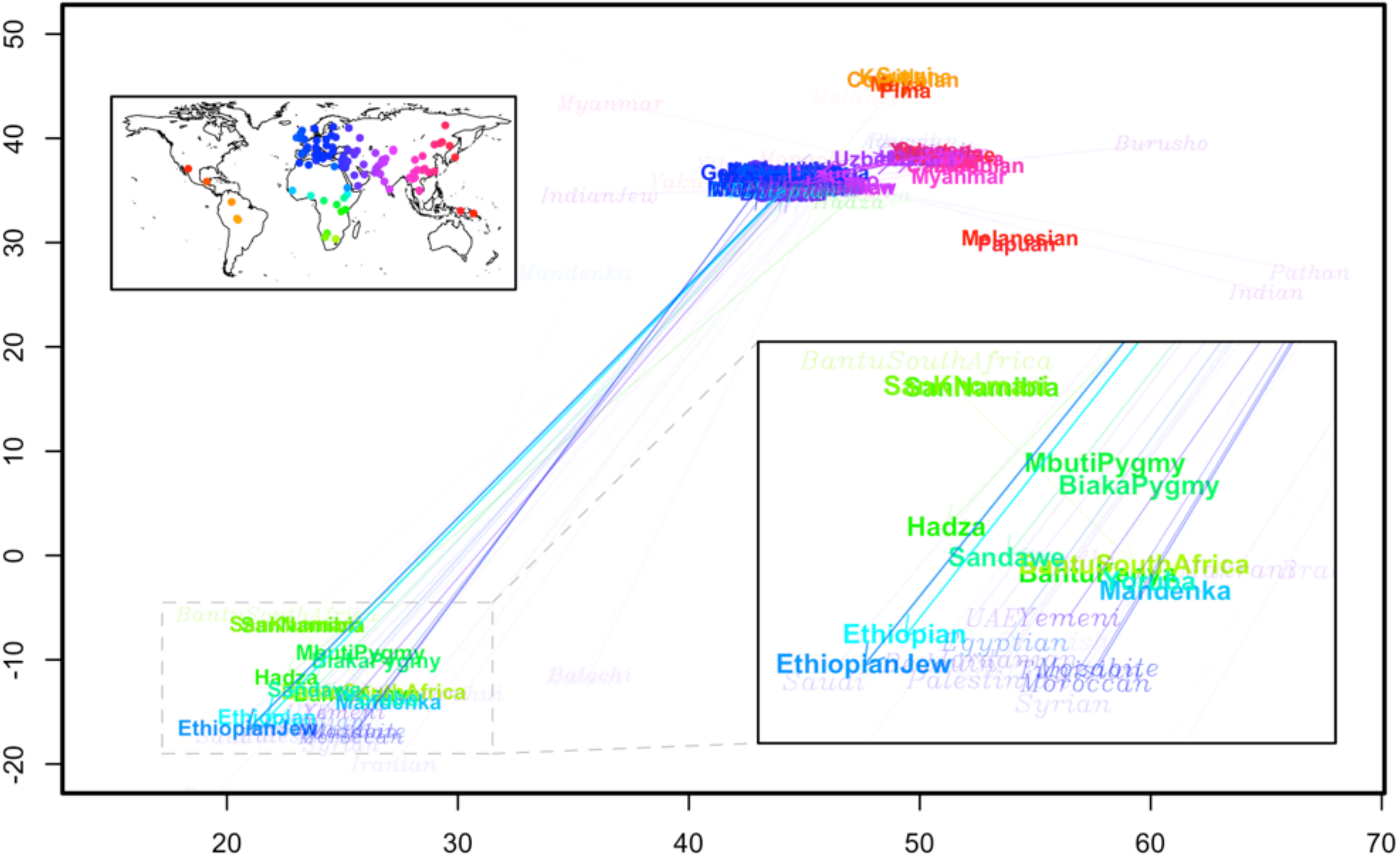
SpaceMix Map: Humans with Admixture

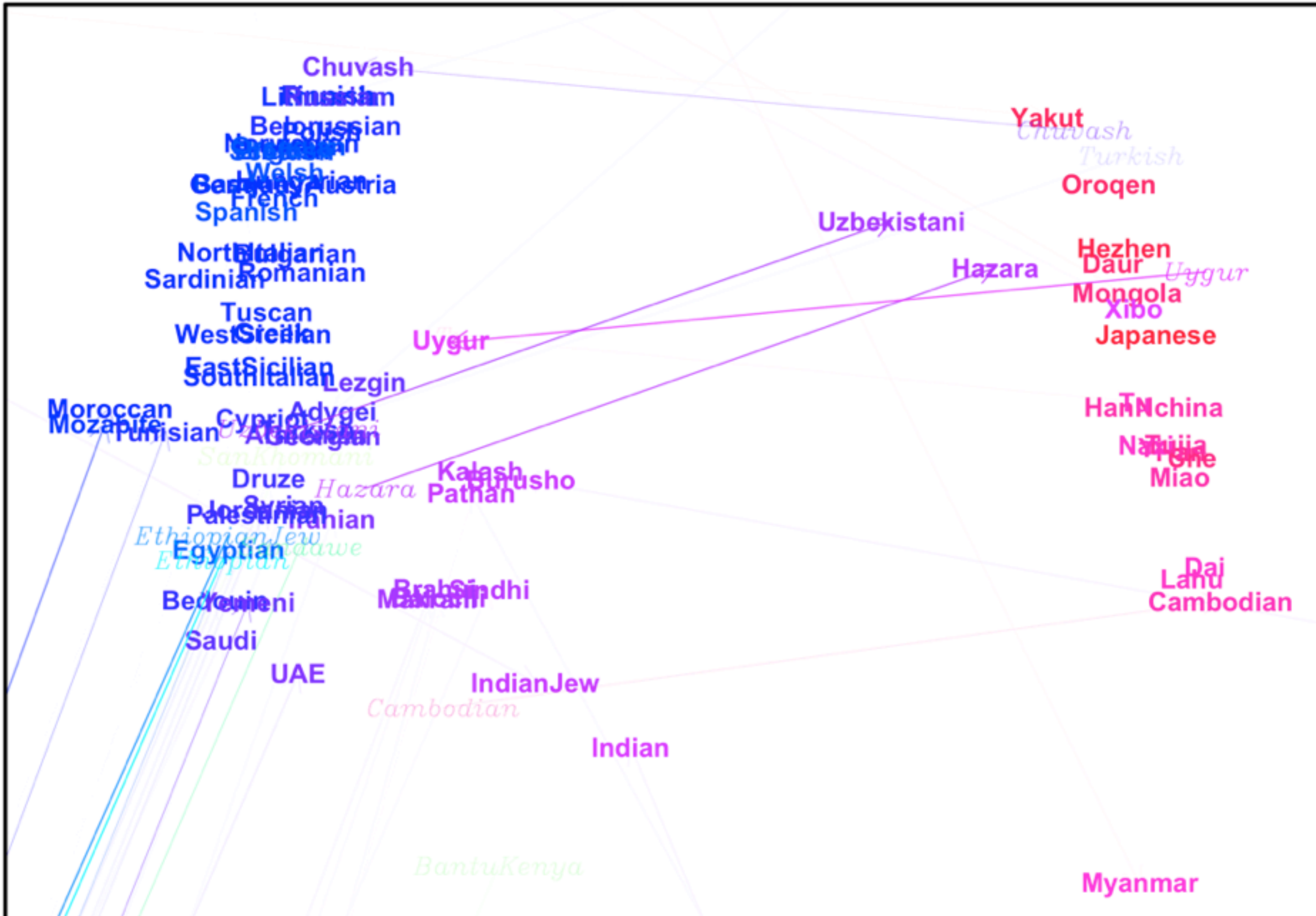


SpaceMix Map: Humans with Admixture



SpaceMix Map: Humans with Admixture




Chuvash
Belorussian
Welsh
French
Spanish
Sardinian
Romanian
Tuscan
Lezgin
Moroccan
Tunisian
Adygei
Druze
Egyptian
Bedouin
Saudi
UAE
Kalash
Burusho
Pathan
Sindhi
IndianJew
Indian
Uygur
Uzbekistani
Hazara
Yakut
Oroqen
Hezhen
Daur
Mongola
Xibo
Japanese
Miao
Lahu
Dai
Cambodian
Myanmar
Chuvash
Turkish
Uygur
Hazara
Cambodian
BantuKenya
EthiopianJew
Ethiopian
SanKhomani

SpaceMix

Conclusions

Extensions

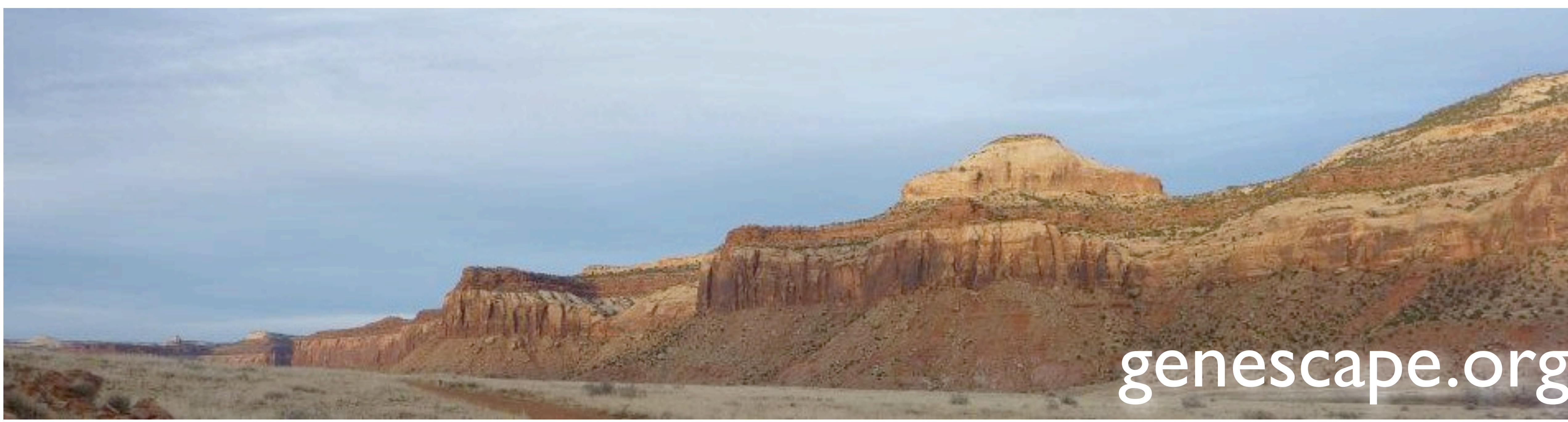
Thanks To:

Collaborators:
- Graham Coop
- Peter Ralph

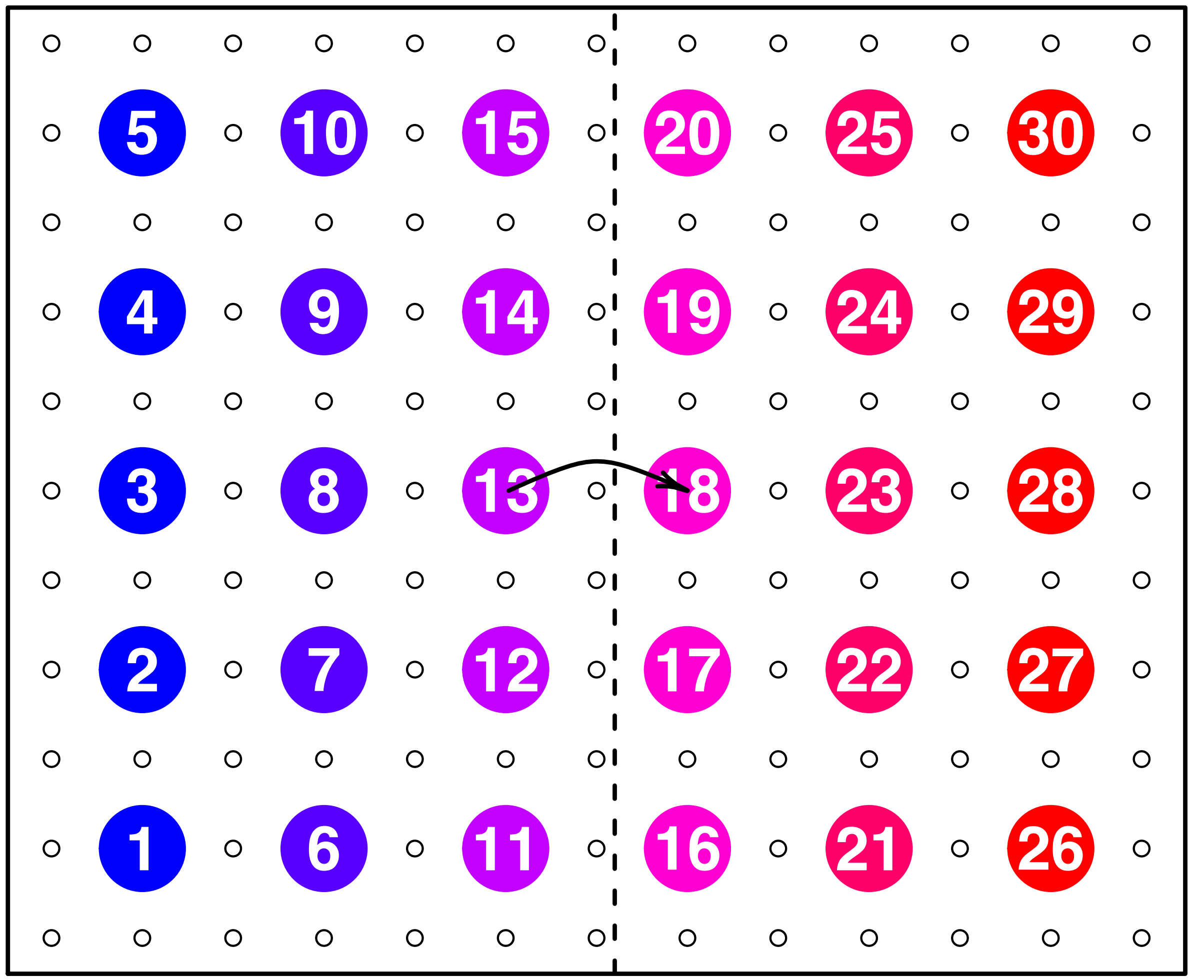
Funding:
- NSF, CPB

Wisdom:
- Yaniv Brandvain
- Marjorie Weber
- Luke Mahler
- Will Wetzel
- Coop Lab
- Brad Shaffer

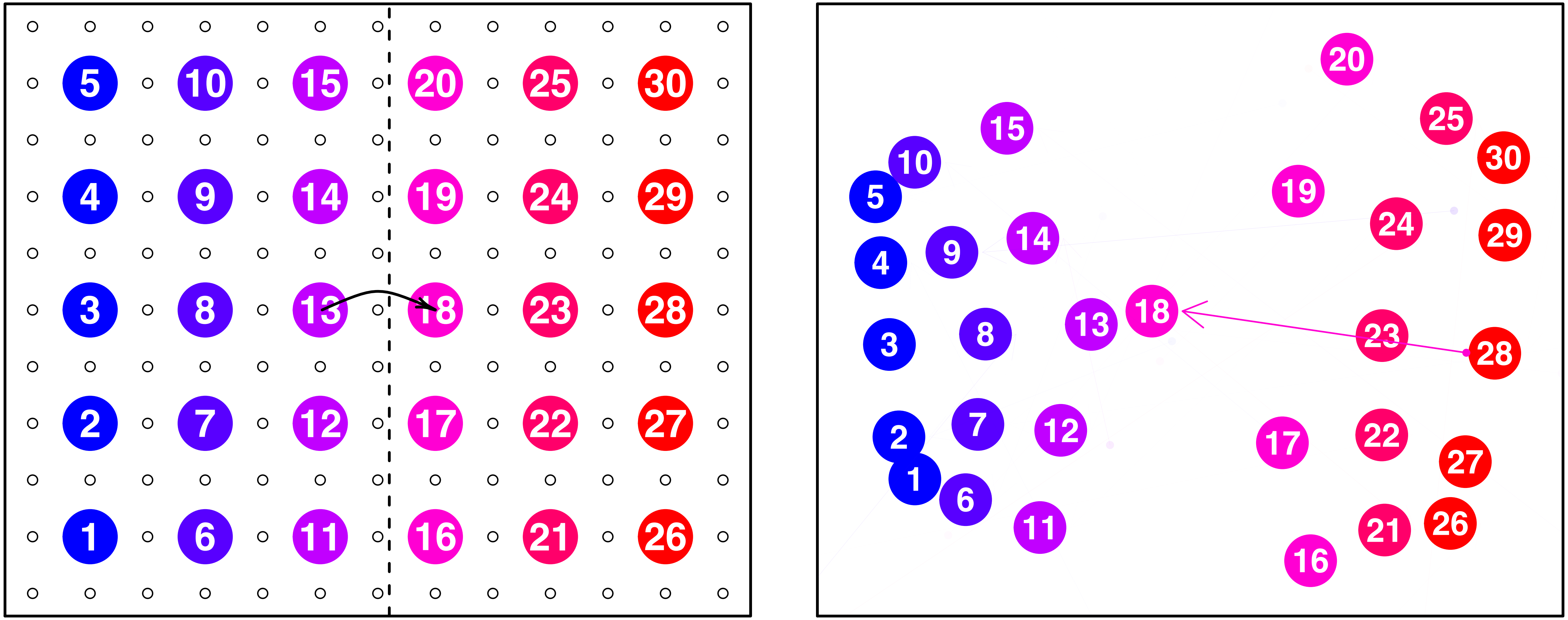
Data:
- Garrett Hellenthal



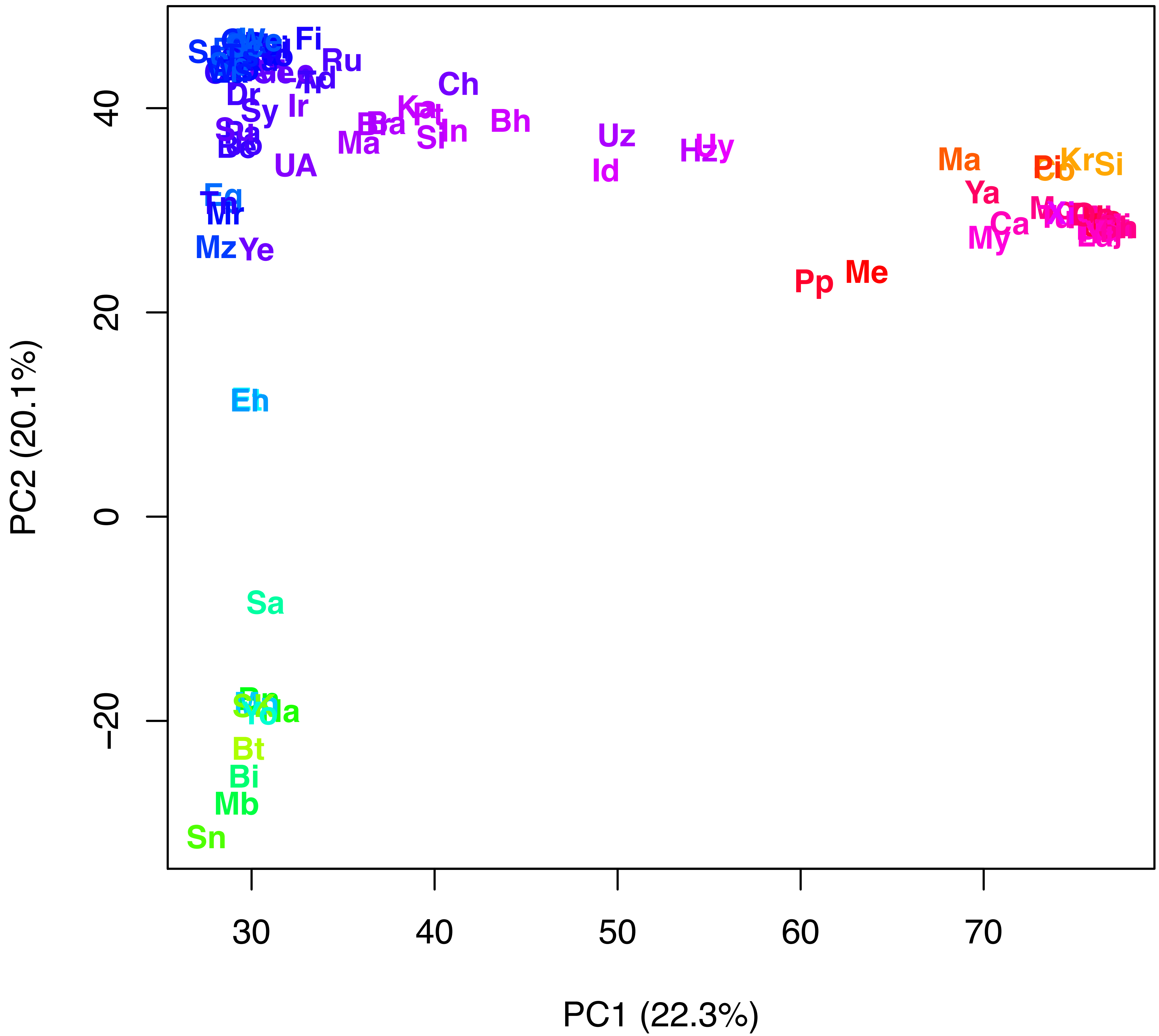
Modeling admixture: corner admixture event

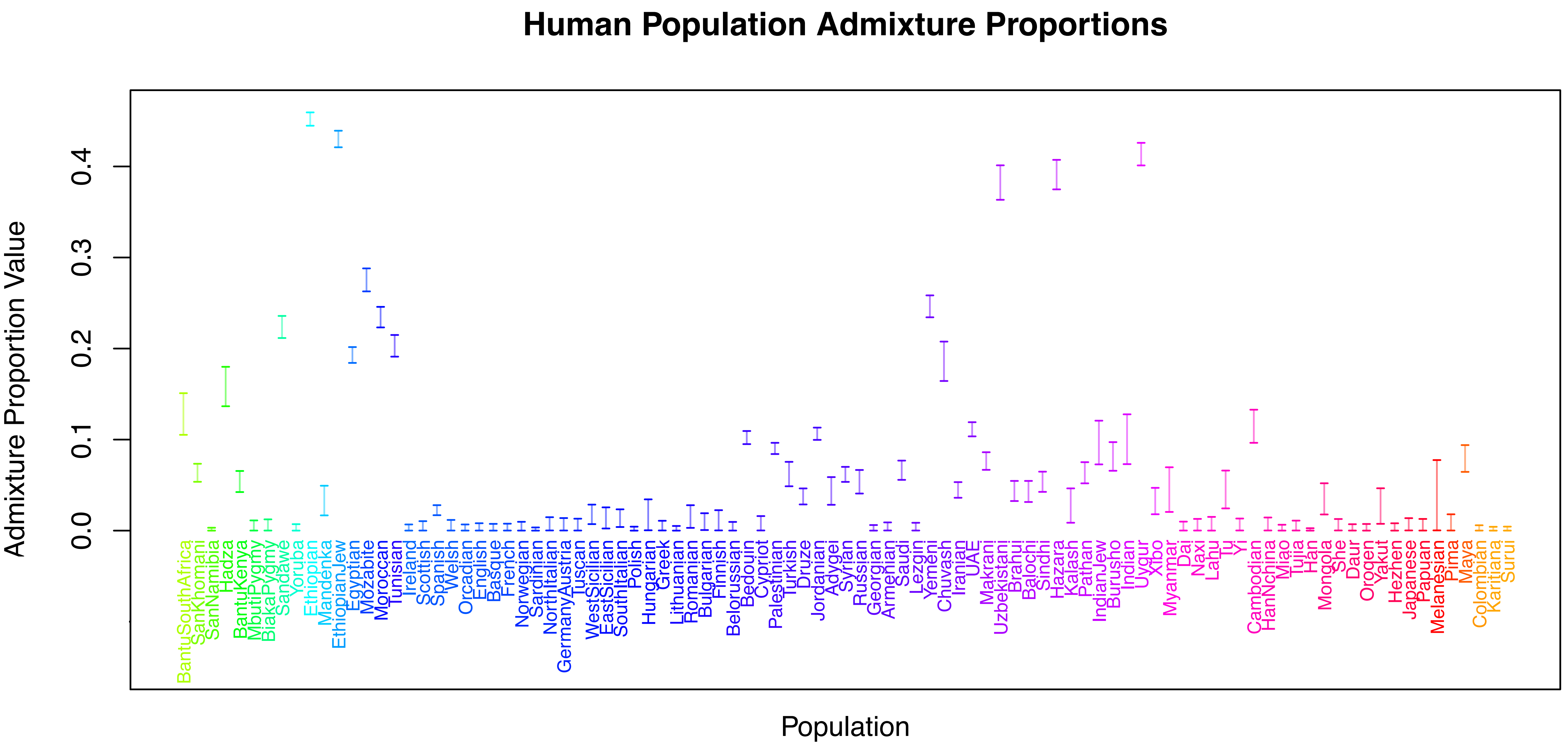


Modeling admixture: corner admixture event



PCA map of Human samples




Human Population Admixture Proportions
Admixture Proportion Value
0.0
0.1
0.2
0.3
0.4
Population
BantuSouthAfrica
SanKhomani
SanNamibia
Hadza
BantuKenya
MbutiPygmy
BiakaPygmy
Sandawe
Yoruba
Ethiopian
Mandenka
EthiopianJew
Egyptian
Mozabite
Moroccan
Tunisian
Ireland
Scottish
Spanish
Welsh
Orcadian
English
Basque
French
Norwegian
Sardinian
NorthItalian
GermanyAustria
Tuscan
WestSicilian
EastSicilian
SouthItalian
Polish
Hungarian
Greek
Lithuanian
Romanian
Bulgarian
Finnish
Belorussian
Bedouin
Cypriot
Palestinian
Turkish
Druze
Jordanian
Adygei
Syrian
Russian
Georgian
Armenian
Saudi
Lezgin
Yemeni
Chuvash
Iranian
UAE
Makrani
Uzbekistani
Brahui
Balochi
Sindhi
Hazara
Kalash
Pathan
IndianJew
Burusho
Indian
Uygur
Xibo
Myanmar
Dai
Naxi
Lahu
Tu
Yi
Cambodian
HanNchina
Miao
Tujia
Han
Mongola
She
Daur
Oroqen
Yakut
Hezhen
Japanese
Papuan
Melanesian
Pima
Maya
Colombian
Karitiana
Surui