

Research Discussion Group: Tell us about yourself!

Name: Alejandro (Alex) Ochoa

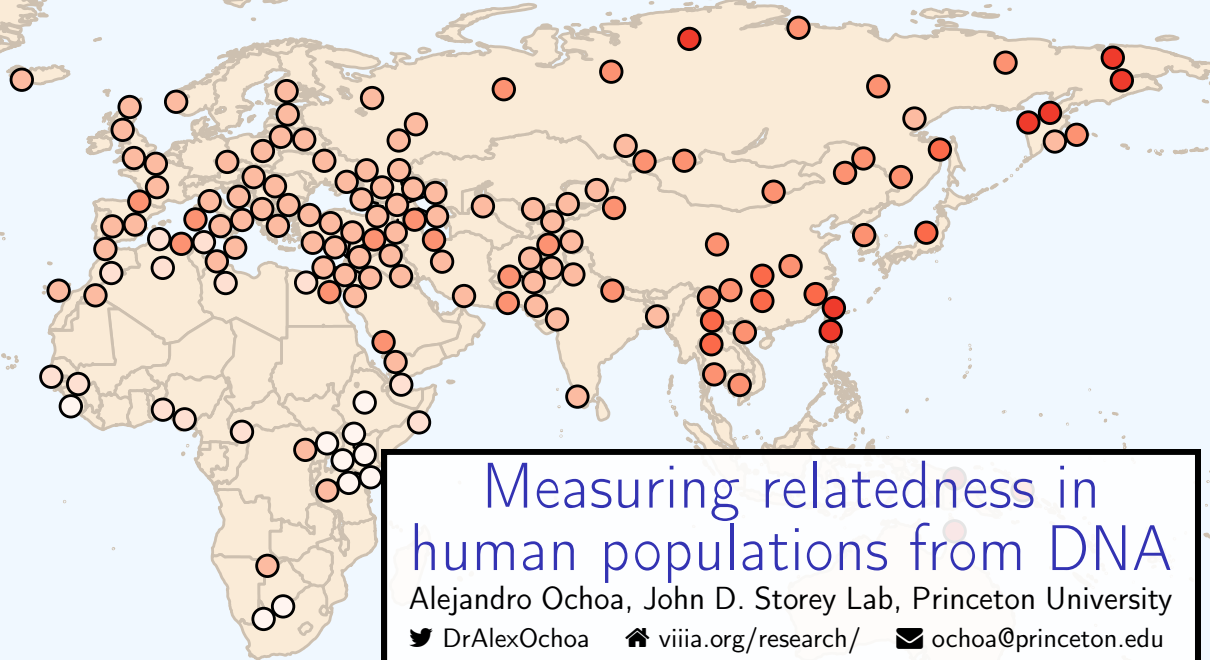
Lab: Postdoc in John Storey's lab (genomics and statistics)

After Princeton: Hope to be a professor leading a research lab

Home town: El Paso, TX / Ciudad Juárez, México

Fun facts:

- ▶ Play guitar, was in a grad student rock band.
- ▶ Languages: English, Spanish, French; Latin, Medieval Spanish, Nahuatl, and more...



Measuring relatedness in human populations from DNA

Alejandro Ochoa, John D. Storey Lab, Princeton University

🐦 DrAlexOchoa

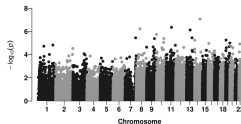
🏠 viiia.org/research/

✉️ ochoa@princeton.edu

Why study relatedness?



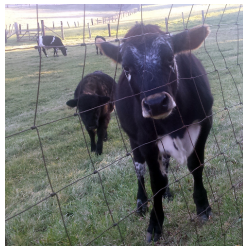
Human genetics
is fascinating!



Search for
disease-causing
genetic variants

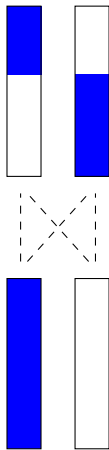
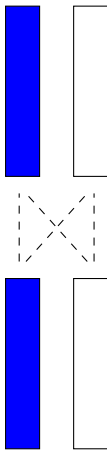
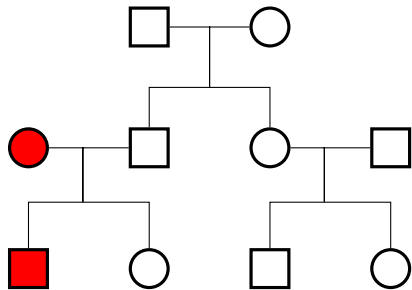


Heritability of
complex traits

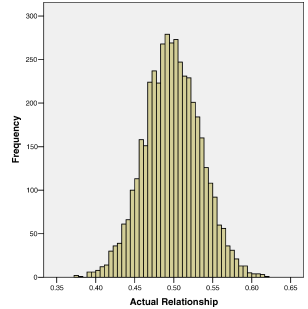
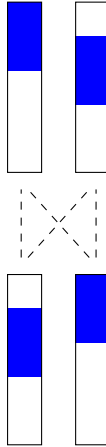
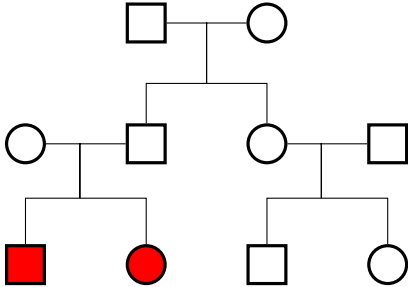


Animal and plant
breeding

The kinship coefficient for parent-child: $\frac{1}{4}$

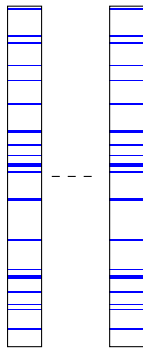
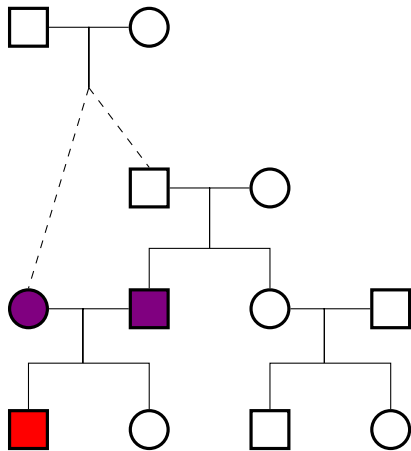


The kinship coefficient for siblings: $\frac{1}{4}$ on average



Visscher *et al.* (2006)

The inbreeding coefficient in populations



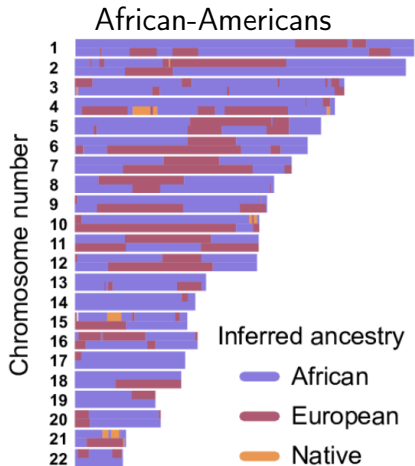
Measurements relative
to a reference pop.:

Inbreeding = 0 in the local population

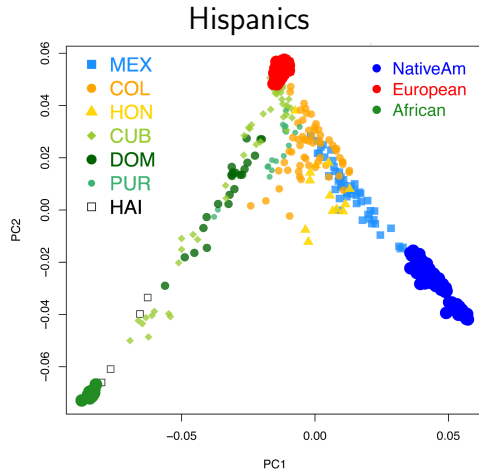
Inbreeding ≥ 0 relative to a distant ancestral population

Better measured using covariance

Recently admixed populations



Baharian *et al.* (2016)



Moreno-Estrada *et al.* (2013)

Kinship model for genotypes

symbol	meaning
T	ref ancestral population
i	locus index
j, k	individual indexes
p_i^T	ref allele frequency
x_{ij}	genotype (num ref alleles)
φ_{jk}^T	kinship of j, k
f_j^T	inbreeding of j

Statistical model:

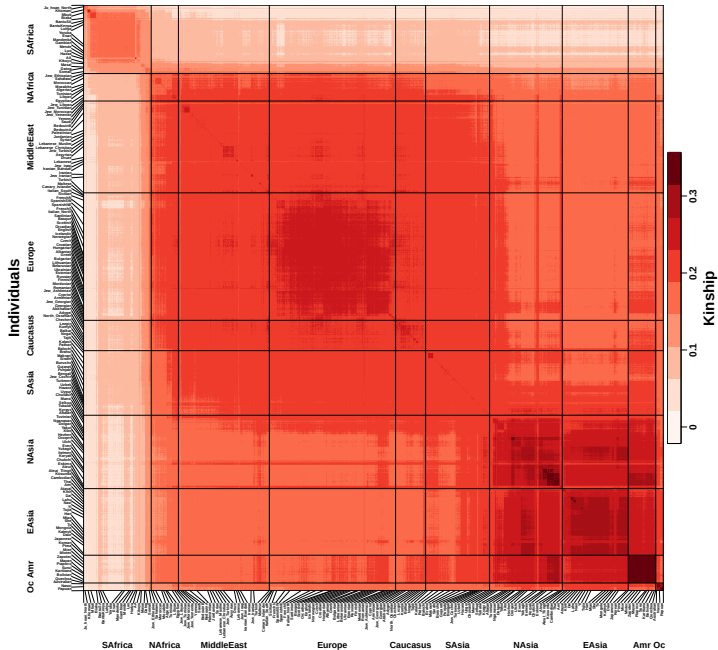
$$E[x_{ij} | T] = 2p_i^T,$$

$$\text{Var}(x_{ij} | T) = 2p_i^T (1 - p_i^T) (1 + f_j^T),$$

$$\text{Cov}(x_{ij}, x_{ik} | T) = 4p_i^T (1 - p_i^T) \varphi_{jk}^T.$$

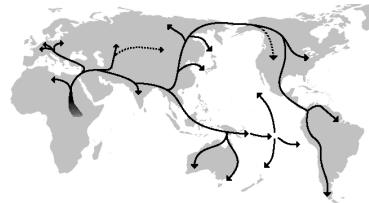
(Wright 1921, 1951; Malécot 1948; Jacquard 1970).

We developed a new kinship estimation framework that works for arbitrary population structures!

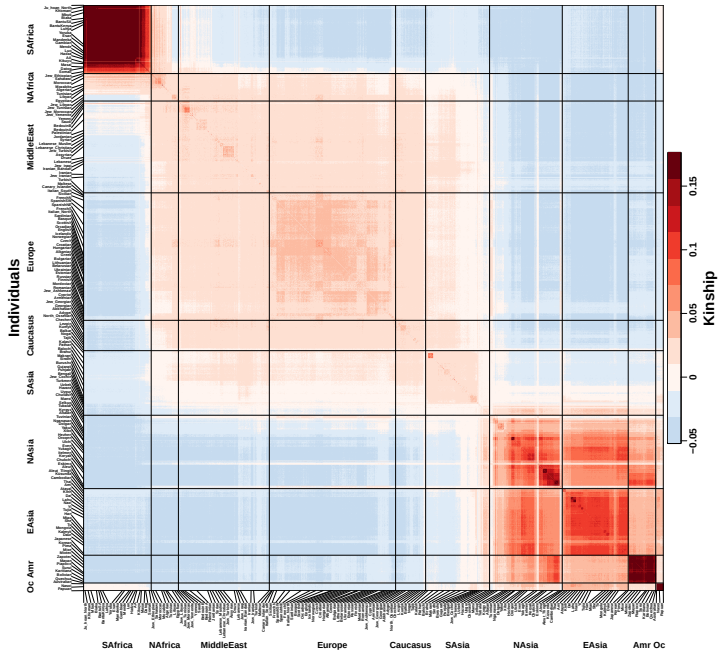


New kinship estimates

Genotypes from "Human Origins" (Lazaridis et al. 2014, 2016)

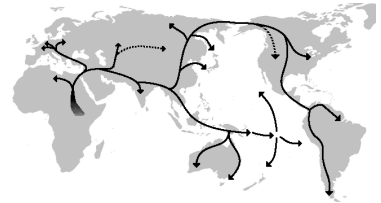


Edited from Ephert [CC BY-SA 3.0], via
Wikimedia Commons



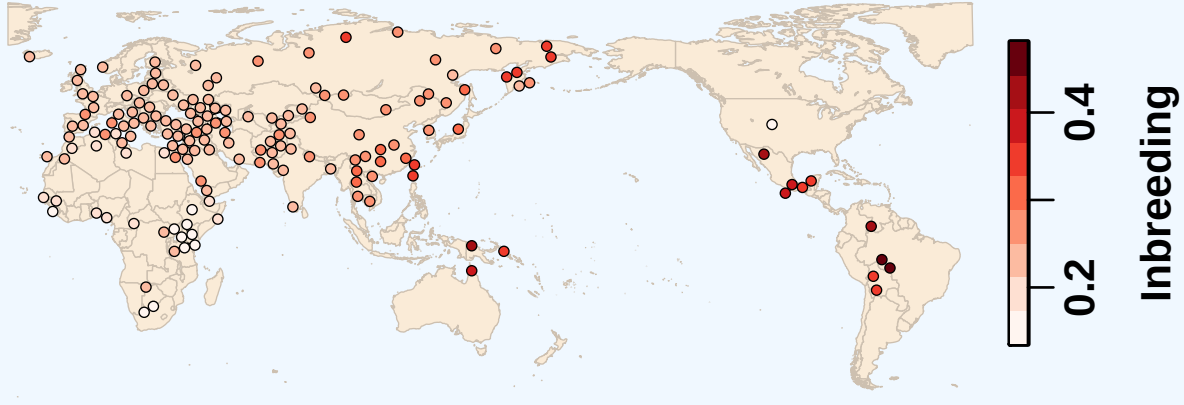
Standard kinship estimates

Genotypes from "Human Origins" (Lazaridis et al. 2014, 2016)



Edited from Ephert [CC BY-SA 3.0], via
Wikimedia Commons

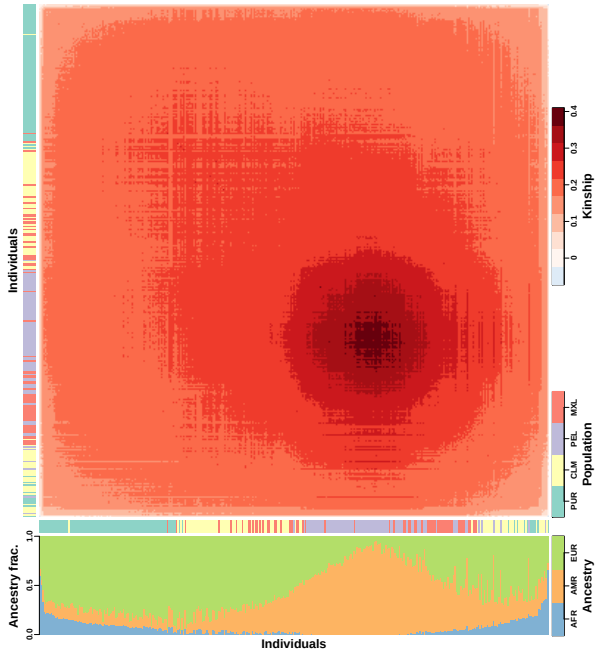
Population-level inbreeding



Kinship driven by admixture in Hispanics

New kinship estimates

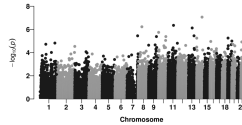
Genotypes from the 1000 Genomes Project (2012)



Improved relatedness has repercussions across genetics!



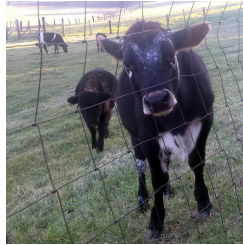
Easy to measure
routinely



Search for
disease-causing
genetic variants



Heritability of
complex traits



Animal and plant
breeding

Acknowledgments

John D. Storey

Andrew Bass

Irineo Cabrereros

Chee Chen

Wei Hao

Emily Nelson

Riley Skeen-Gaar

Neo Christopher Chung

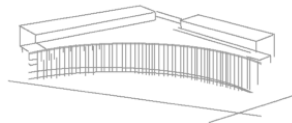
Institute of Informatics

University of Warsaw

Funding:

National Institutes of Health

Otsuka Pharmaceuticals



Lewis-Sigler Institute for Integrative Genomics

