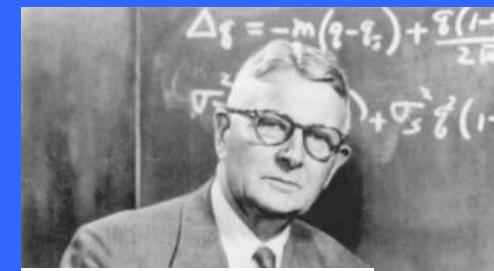


Una Guía para Principiantes a la F_{ST} : Interpretación y algunos patrones.

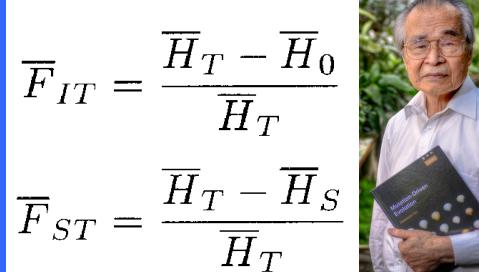
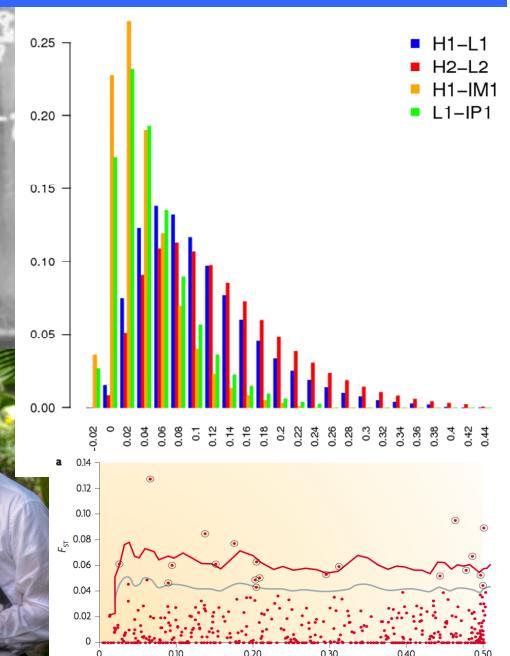
Luis E. Eguiarte
Laboratorio de
Evolución
Molecular y
Experimental
Instituto de Ecología,
UNAM



$$\overline{F}_{IS} = \frac{\overline{H}_S - \overline{H}_0}{\overline{H}_S}$$

$$\overline{F}_{IT} = \frac{\overline{H}_T - \overline{H}_0}{\overline{H}_T}$$

$$\overline{F}_{ST} = \frac{\overline{H}_T - \overline{H}_S}{\overline{H}_T}$$



Estadísticos o coeficientes F : F_{IS} , F_{ST} , F_{IT}

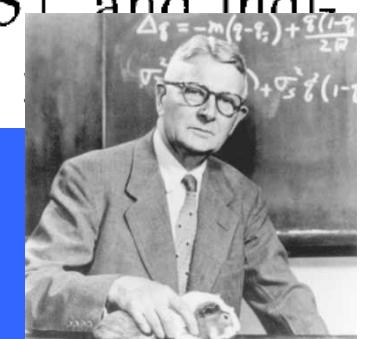
Los estimadores más importantes de la mayoría de los estudios empíricos en genética y genómica de poblaciones.

Una herramienta básica, como vamos a ver!

Sewall Wright 1951, 1965

Several different approaches have been used to estimate the amount of differentiation in the subdivisions of a population. Most importantly, Wright (1951, 1965b) developed an approach to partition the genetic variation in a subdivided population that is commonly used and provides an obvious description of differentiation. This approach consists of three different F coefficients (these are correlation coefficients and are different from the F statistics used in the analysis of variance) used to allocate the genetic variability to the total population level (T), subpopulations (S) and individuals (I). These three coefficients, F_{ST} , F_{IT} , and F_{IS} , are

Dos formas de verlos: partición de la varianza o correlaciones.



H_o , promedio de la heterocigosis **observadas** en c/población

H_s promedio de la heterocigosis **esperadas** en cada población, de sus p y q de cada una

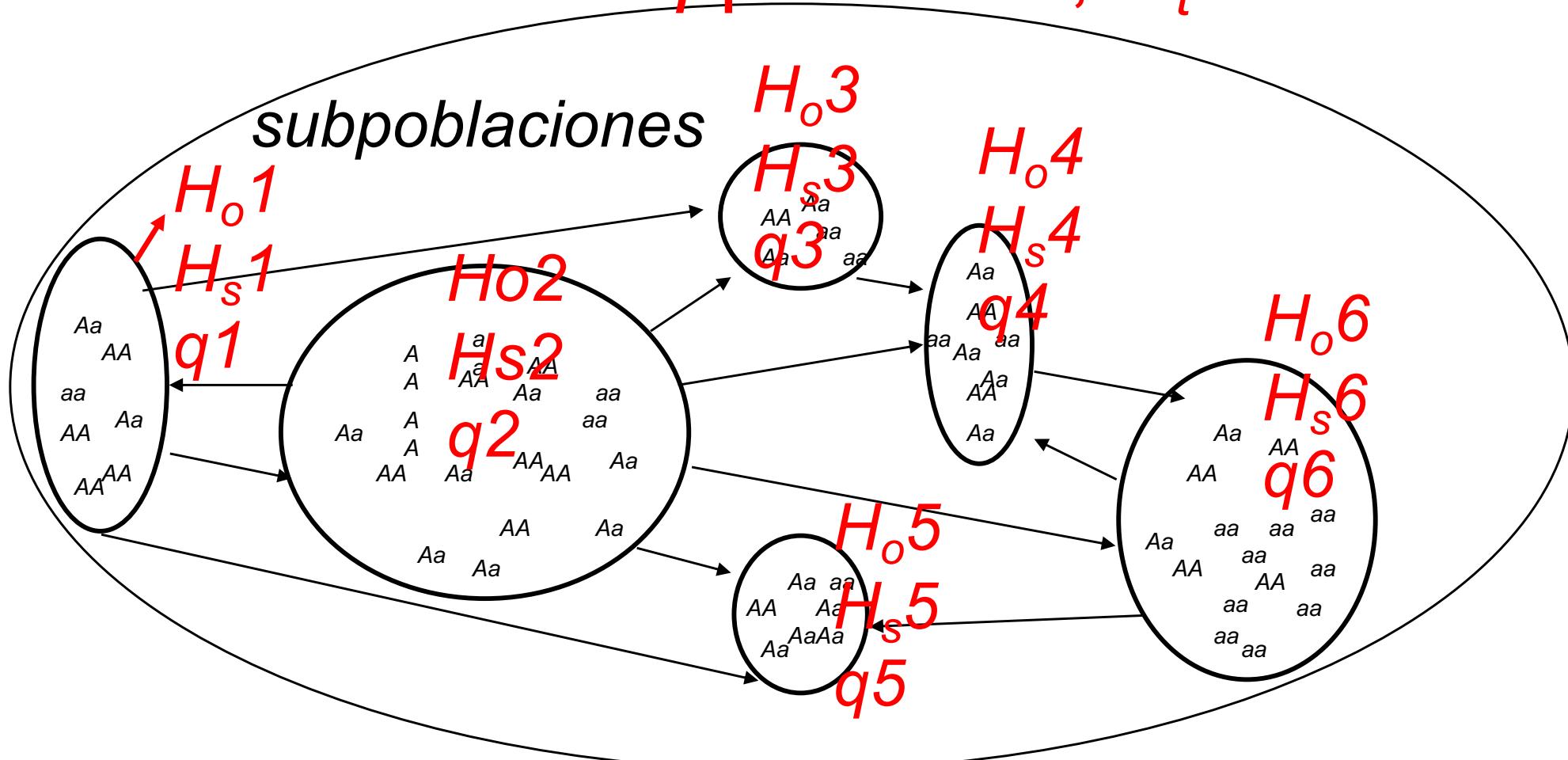
H_t la heterocigosis del promedio de las p y q de toda la población como un todo!

$$\overline{F}_{IS} = \frac{\overline{H}_S - \overline{H}_0}{\overline{H}_S}$$

$$\overline{F}_{IT} = \frac{\overline{H}_T - \overline{H}_0}{\overline{H}_T}$$

$$\overline{F}_{ST} = \frac{\overline{H}_T - \overline{H}_S}{\overline{H}_T}$$

ESPECIE q promedio, H_t



$$F_{it} = F_{is} + (1-F_{is}) F_{st}$$



Nei (1977) also extended this analysis to multiple loci.

$$\bar{F}_{IS} = \frac{\bar{H}_S - \bar{H}_0}{\bar{H}_S}$$

\bar{H}_0 promedio de la heterocigosis observadas.

$$\bar{F}_{IT} = \frac{\bar{H}_T - \bar{H}_0}{\bar{H}_T}$$

H_s promedio de la heterocigosis esperadas en cada población, de sus p y q de cada una.

$$\bar{F}_{ST} = \frac{\bar{H}_T - \bar{H}_S}{\bar{H}_T}$$

H_t la heterocigosis del promedio de las p y q de toda la población como un todo!

$$F_{it} = F_{is} + (1-F_{is}) F_{st}$$

Nei (1977) also extended this analysis to multiple loci.

$$\bar{F}_{IS} = \frac{\bar{H}_S - \bar{H}_0}{\bar{H}_S}$$

$$\bar{F}_{IT} = \frac{\bar{H}_T - \bar{H}_0}{\bar{H}_T}$$

$$\bar{F}_{ST} = \frac{\bar{H}_T - \bar{H}_S}{\bar{H}_T}$$

F_{is} : efectos por endogamia/
sistemas reproductivos:
-1 puros hetero.,
0 HW, 1 endogamia total

F_{st} : diferenciación por deriva vs.
migración o selección:
0 idénticas en f. alelicas, 1
totalmente diferentes.

F_{it} = endogamia + deriva, -1 a 1

F_{is} : efectos por endogamia/sistemas reproductivos:
-1 puros hetero., 0 HW, 1 endogamia total
 F_{st} : diferenciación por deriva vs.migración o selección:
0 idénticas en f. alelicas, 1 totalmente diferentes. O
proporción de la var. genética entre poblaciones, de 0
(toda en las pob.) a 1 (toda ente pob.)
 $F_{it} = \text{endogamia} + \text{deriva}$, -1 a 1

F_{ST} is a measure of the genetic differentiation over subpopulations and is always positive. F_{IS} and F_{IT} are measures of the deviation from Hardy-Weinberg proportions within subpopulations and in the total population, respectively, where positive values indicate a deficiency of heterozygotes and negative values indicate an excess of heterozygotes. There has been

$$F_{it} = F_{is} + (1-F_{is}) F_{st}$$

Estadísticos o coeficientes F : F_{is} , F_{st} , F_{it}

Relaciones entre las F s

$$1 - F_{IT} = (1 - F_{ST})(1 - F_{IS})$$

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

$$F_{it} = F_{is} + (1 - F_{is}) F_{st}$$

The significance of F_{IS} can be calculated from a χ^2 test

$$\chi^2 = NF_{IS}^2$$

where N is the number of individuals in the sample and there is one degree of freedom.

Significancia F_{is} : depende de la N
y se distribuye como chi-cuadrada,
con un grado de libertad.

Otra forma de ver a la F_{st}

$$F_{ST} = \frac{V(q)}{\bar{q}(1 - \bar{q})}$$

VARIANZA
DE LAS
FREC.
ALELICAS
ENTRE
POBLACIONES

Significancia F_{st}

$$\chi^2 = 2N F_{ST}$$

$2N$ is the number of gametes in the sample and there is one degree of freedom

Un par de ejemplos sencillos:

<i>Subpopulation</i>	A_1A_1	A_1A_2	A_2A_2	q
1	0.25	0.5	0.25	0.5
2	0.35	0.3	0.35	0.5
	$F_{IS} = 0.2$	$F_{IT} = 0.2$	$F_{ST} = 0.0$	
1	0.25	0.5	0.25	0.5
2	0.49	0.42	0.09	0.3
	$F_{IS} = 0.0$	$F_{IT} = 0.0417$	$F_{ST} = 0.0417$	

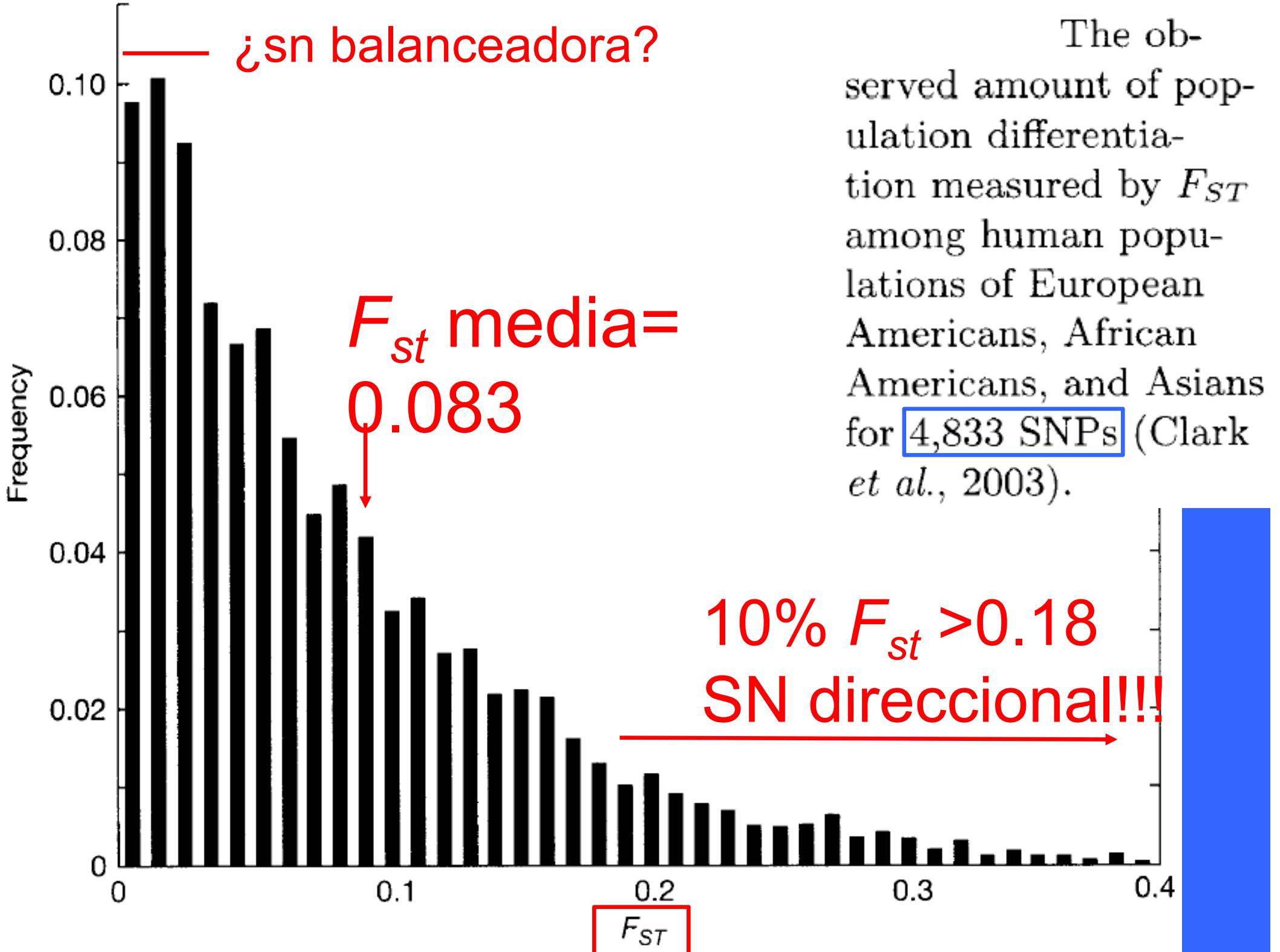
the same allele frequencies so that there is no genetic differentiation among subpopulations, making $F_{ST} = 0$. However, both F_{IS} and F_{IT} are positive because of the deficiency of heterozygotes in subpopulation 2. In the bottom example, both subpopulations are in Hardy–Weinberg proportions so that $F_{IS} = 0$. However, because of the variation in allele frequencies between subpopulations, both F_{IT} and F_{ST} are positive.

$$F_{it} = F_{is} + (1-F_{is}) F_{st}$$

Diferencia en F_{st} entre loci: ¡sugiere selección!

Clark *et al.* (2003) estimated the amount of population differentiation among European Americans, African Americans, and Asians for 4833 SNPs (Figure 9.7). A large proportion of the SNPs indicated very low differentiation, and the mean F_{ST} value was 0.083. However, the distribution had a long tail with 10% of the SNPs having an F_{ST} value of > 0.18. The SNPs in the long tail indicate that there is strong differentiation among the groups for the genetic regions marked by these SNPs, potentially pointing out past selective events acting differentially in these populations.

10% de los genes en humanos,
 F_{st} mayor de 0.18= selección diferencial
(otros menos diferenciación, selección balanceadora!)



G_{st} : estimador de la F_{st} a partir de H

As an estimate of F_{ST} , and assuming Hardy–Weinberg proportions, Nei (1973) defined the coefficient of gene differentiation as

$$G_{ST} = \frac{H_T - H_S}{H_T}$$

where H_S is the average subpopulation Hardy–Weinberg heterozygosity and $H_T = 1 - \sum \bar{p}_i^2$ for any number of alleles. Nei (1973, 1987) pointed out

that although G_{ST} is a good measure of the relative differentiation among subpopulations, it is highly dependent on the amount of variation within subpopulations and in the total population.

depende del total de
variación...

$$G_{ST} = \frac{H_T - H_S}{H_T}$$

Problema con microsatélites, H_S y H_T cerca de 1...

Si hay mucha variación, la G_{st} puede ser MUY baja, pero sin que se comparten alelos!

The dependence of G_{ST} on the amount of genetic variation is particularly true for highly variable loci such as microsatellite loci where both H_S and H_T can approach unity. As a result, the G_{ST} can be very small even if the subpopulations have nonoverlapping sets of alleles (Hedrick, 1999b; see also Charlesworth, 1998). This seems counterintuitive because in the two-allele case, when the subpopulations are monomorphic for different alleles, $F_{ST} = 1$. However, G_{ST} measures the proportional amount of variation within subpopulations as compared with the total population and does not specify the identity of the alleles involved. The magnitude of G_{ST} can also be written as

$$G_{ST} = 1 - \frac{H_S}{H_T}$$
$$< 1 - H_S$$

where $1 - H_S$ is the average within population homozygosity. From this, it is obvious that the differentiation cannot exceed the level of homozygosity, no matter what evolutionary factor is influencing the amount and pattern of variation. Obviously, when using highly polymorphic makers that make the level of homozygosity low, then the maximum G_{ST} must also be greatly reduced.

Si hay muchas variación, baja homocigosis y la G_{st} aún más chica, baja su utilidad... así, se han propuestos otros índices mas robustos, como $G^{\prime st}$ de Hedrick, o la $DJost$

Otros niveles de partición se pueden proponer
n niveles, de mas local a más global

there is a logical regional level into which subpopulations can be placed, then we can calculate the additional measures

$$F_{SR} = \frac{H_R - H_S}{H_R}$$

$$F_{RT} = \frac{H_T - H_R}{H_T}$$

which partition the variation into the diversity among subpopulations within region and that among regions for the total population. With such hierarchical partitioning, it is possible to see at which level the largest amount of variation can be explained. For example, most of the variation may be among subpopulations in some species, whereas in other species, most of the variation may be among regional groups (see Example 9.4 for

7 poblaciones, de tres regiones, un locus

TABLE 9.6 The frequency of alleles at the LL53 microsatellite locus for the remaining natural populations of the Gila topminnow from three regional groups, Bylas Springs, Sonoita Creek Springs, and Sonoita Creek, and where — indicates allele is absent.

	Allele							
	138	142	144	146	148	150	164	H
Bylas Springs								
Bylas Spring I	—	—	—	1.000	—	—	—	0.000
Bylas Spring II	—	—	0.115	0.885	—	—	—	0.204
Mean	—	—	0.058	0.942	—	—	—	0.109
Sonoita Creek Springs								
Cottonwood Spring	—	0.278	0.722	—	—	—	—	0.401
Monkey Spring	—	0.988	—	—	—	0.012	—	0.024
Mean	—	0.633	0.361	—	—	0.006	—	0.469
Sonoita Creek								
Coalmine Canyon	—	0.725	0.250	—	0.025	—	—	0.411
Sonoita Creek	—	0.759	0.241	—	—	—	—	0.366
Red Rock Falls	0.025	0.700	—	—	—	—	0.275	0.434
Mean	0.008	0.728	0.164	—	0.008	—	0.092	0.435
Total mean	0.004	0.493	0.190	0.269	0.004	0.002	0.039	0.647

$$H_s = 0.263$$

$$H_r = 0.357$$

$$H_t = 0.647$$

$$F_{sr} = 0.253$$

$$F_{rt} = 0.456$$



$$\uparrow \quad H_t$$

mean of these seven values is $H_S = 0.263$. Next calculate the mean allele frequency within each regional group. Using these mean allele frequencies, the Hardy–Weinberg heterozygosity for each region can be calculated (also in the rightmost column). The weighted mean of these values is $H_R = [2(0.109) + 2(0.469) + 3(0.435)]/7 = 0.352$. Finally, the mean allele frequency is calculated (bottom row), and the Hardy–Weinberg heterozygosity of these frequencies is $H_T = 0.647$. With these values, the proportion of variation among subpopulations within regions is $F_{SR} = 0.253$, and the proportion of variation among regions is $F_{RT} = 0.456$. Therefore, nearly twice as much variation is partitioned among the regions as among the subpopulations within groups. Note that we have calculated these values for

$$H_S = 0.263$$

$$H_R = 0.357$$

$$H_T = 0.647$$

$$F_{sr} = 0.253$$

$$F_{rt} = 0.456$$

	Allele							
	138	142	144	146	148	150	164	H
Bylas Springs								
Bylas Spring I	—	—	—	1.000	—	—	—	0.000
Bylas Spring II	—	—	0.115	0.885	—	—	—	0.204
Mean	—	—	0.058	0.942	—	—	—	0.109
Sonoita Creek Springs								
Cottonwood Spring	—	0.278	0.722	—	—	—	—	0.401
Monkey Spring	—	0.988	—	—	—	0.012	—	0.024
Mean	—	0.633	0.361	—	—	0.006	—	0.469
Sonoita Creek								
Coalmine Canyon	—	0.725	0.250	—	0.025	—	—	0.411
Sonoita Creek	—	0.759	0.241	—	—	—	—	0.366
Red Rock Falls	0.025	0.700	—	—	—	—	0.275	0.434
Mean	0.008	0.728	0.164	—	0.008	—	0.092	0.435
Total mean	0.004	0.493	0.190	0.269	0.004	0.002	0.039	0.647

$$\frac{m_p}{m_s} = \frac{F_{ST}(m) - 2F_{ST} + F_{ST}F_{ST}(m)}{F_{ST}(1 - F_{ST}(m))}$$

TABLE 9.7 The overall (biparental) and maternal F_{ST} values for eight tree species, one oak (*Quercus*) and five pines (*Pinus*), and two mountain ash (*Sorbus*), and the estimated ratio of pollen to seed gene flow (Ennos, 1994; Latta and Mitton, 1997; Oddou Muratorio *et al.*, 2001; Bacles *et al.*, 2004).

Species	F_{ST}	$F_{ST}(m)$	m_p/m_s
<i>Quercus petraea/Q. robur</i>	0.037	0.88	196
<i>Pinus contorta</i>	0.061	0.66	28
<i>P. radiata</i>	0.13	0.83	31
<i>P. attenuata</i>	0.12	0.86	44
<i>P. muricata</i>	0.22	0.88	24
<i>P. flexilis</i>	0.013	0.68	159
<i>Sorbus aucuparia</i>	0.043	0.13	1.4
<i>S. torminalis</i>	0.11	0.34	2.2

Rango F_{st} de 0.013 a 0.22 en estos árboles
 cocientes de m_p/m_s de 196 y 159 pinos y encinos
 semillas dispersadas por viento/gravedad,
 a 1.4 *Sorbus*, dispersados por aves

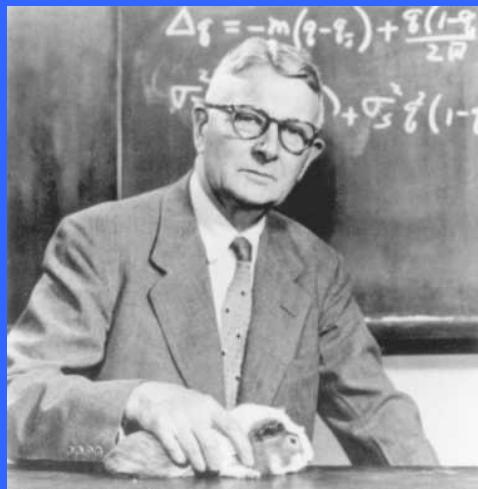


Estructura genética y deriva: balance entre migración y deriva

El flujo génico homogeniza las f. alélicas,
la deriva hace que diverjan...

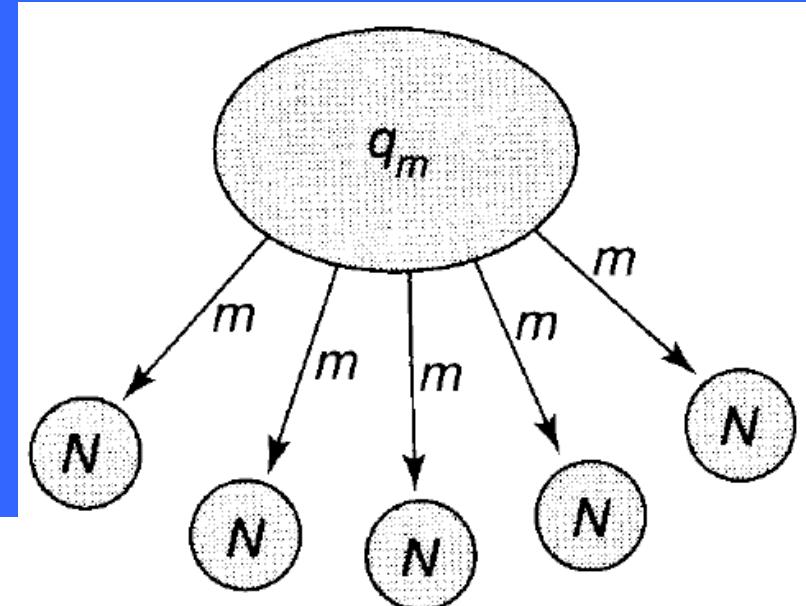
The effect of gene flow is to keep the allele frequencies in different subpopulations similar. However, if the subpopulations are finite in size, then genetic drift may result in random differences among them, even with gene flow. The simplest model to examine the joint effects of gene flow and genetic drift assumes that migrants enter a number of equal-sized, finite populations in equal proportions from a source population. More realistically, subpopulations are distributed over space, and gene flow between them must depend to some extent on their distance from each other. For example, distance-dependent gene flow can be included in models where individuals are distributed in discrete groups, colonies, or villages; these are generally known as stepping-stone models. The general model of pop-

El modelo de islas, o continente-islas de S. Wright, 1940



The Continent–Island or Island Model

Balance derivado de la migración si se mantiene la tasa m y N_e mucho tiempo



Wright (1940) called this the **island**

model because he assumed that there were *many finite subpopulations (equivalent to the continent) that were the source of migrants as well as receive them*. When the amount of gene flow and the population size on the islands are both large, then the allele frequency on the islands will soon become similar to that on the continent—essentially the situation that we discussed earlier. However, if the population size on the islands is small and/or the rate of gene flow is low, then it is expected that genetic drift could result in chance changes in allele frequencies. As a result, the allele frequencies on the islands may differ significantly from each other and from the allele frequency in the migrants.

Let us assume that there is a probability $1/(2N)$ that two alleles are identical by descent in the previous generation $t - 1$ and a probability $1 - 1/(2N)$ that they are descended from different alleles in the previous generation (N is assumed to be the effective population size here). The expected homozygosity in generation t is then

¿cómo cambia
la homocigosis?

El incremento + la anterior

$$f_t = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{t-1}$$

The probability of identity is modified by the probability that both alleles are not migrants, or $(1 - m)^2$, so

$$f_t = \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{t-1} \right] (1 - m)^2$$

En el equilibrio, si se mantiene ct. m , la f ya no cambia

$$f_t = \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{t-1} \right] (1-m)^2$$

If it is assumed that there is an equilibrium between gene flow bringing in new variation and finite population size reducing variation, then $f = f_t = f_{t-1}$. Furthermore, if we assume that f is equal to the equilibrium fixation index F_e or F_{ST} , then

Relación entre Fst y m.

$$F_{ST} = \frac{(1-m)^2}{2N - (2N-1)(1-m)^2}$$

When $m = 0$, $F_{ST} = 1$ and when $m = 1$, $F_{ST} = 0$. If the terms with m^2 are ignored, then

Primera aproximación

$$F_{ST} = \frac{1-2m}{4Nm+1-2m}$$

$$F_{ST} = \frac{1 - 2m}{4Nm + 1 - 2m}$$

and when we ignore $2m$ in both the numerator and denominator, then

si m es pequeña
Segunda aproximación

$$F_{ST} \approx \frac{1}{4Nm + 1}$$

When $m < 0.01$, expressions 9.12b and 9.12c give quite similar values

Si m pequeña, las 3 resultados similares

If we assume that there are k equivalent subpopulations,

$$G_{ST} = \frac{1}{4Nm \left(\frac{k}{k-1} \right)^2 + 1}$$

una estimación de G_{st} si hay k poblaciones

importante si analizamos pocas poblaciones

**Nm = migrantes efectivos, con 1 que llegue,
se evita la deriva... o no...**

In general, it has been suggested that **one migrant per generation**, $Nm = 1$, is enough to prevent the effects of genetic drift among populations. If $Nm = 1$, then $F_{ST} = 0.2$ in equation 9.12c, a significant level of differentiation, even for very small sample sizes. This fairly substantial value and other considerations led Mills and Allendorf (1996) to recommend that $Nm = 1$ may be inadequate connectivity for natural populations, and **they recommended higher levels of gene flow** for management of endangered species. However, Wang (2004) has considered the theoretical assumptions underlying the one-migrant-per-generation recommendation and has found that they are generally robust when the effective number of migrants, $N_e m_e$, where the **effective rate of gene flow m_e** , which takes into account variance in migration, is substituted for m .

Estimación indirecta de Nm

Expression 9.12c can be solved for an estimate of the number of migrants per generation as follows:

$$Nm = \frac{1 - F_{ST}}{4F_{ST}}$$

This relationship has been widely used to estimate the number of migrants between populations. It is an approximation of a particular theoretical

model at equilibrium and therefore should be used only as a general guideline to estimate the number of migrants (see the discussion in Waples, 1998; Gaggiotti *et al.*, 1999; Whitlock and McCauley, 1999; Neigel, 2002). In ad-

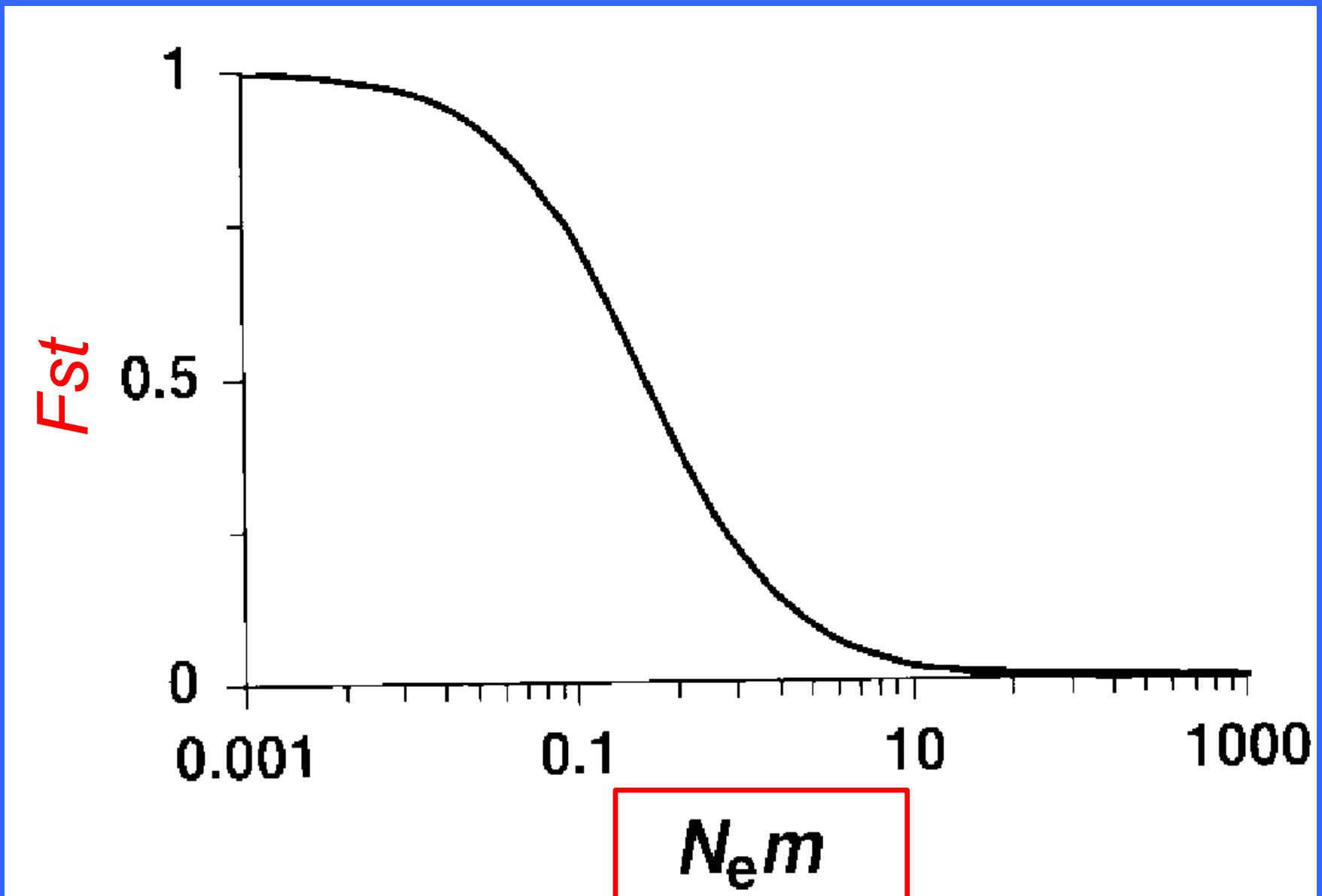
Una guía general, a “orden de magnitud”

$$Nm = \frac{1 - F_{ST}}{4F_{ST}}$$

dition, when F_{ST} is small, then there can be bias in the estimation of Nm . For example, if $F_{ST} = 0.01$, then expression 9.12e gives an estimate of $Nm = 24.8$. Using expression 9.11b and assuming that $N = 50$, we get an estimate of $m = 0.29$ and $Nm = 14.6$ a value over 40% lower.

pero puede tener sesgos
si F_{st} es chica... ca. 0.01, Nm de 24.8
a 14.6

Al disminuir N_e y/o m (tasa de migración) aumenta la diferenciación entre poblaciones medida como F_{st} ya que $F_{st} = 1 / (4Nm + 1)$ en el equilibrio.



¿En cuanto **tiempo** se llega al equilibrio Nm / F_{st} ?

One concern about using estimates based on variation in allele frequency over groups is that they may be strongly influenced by the history of the populations and may not be at equilibrium. First, Wright (1943b) showed that if there is no gene flow between populations ($m = 0$ in expression 9.12a), then

$$F_{ST} = 1 - e^{-t/2N}$$

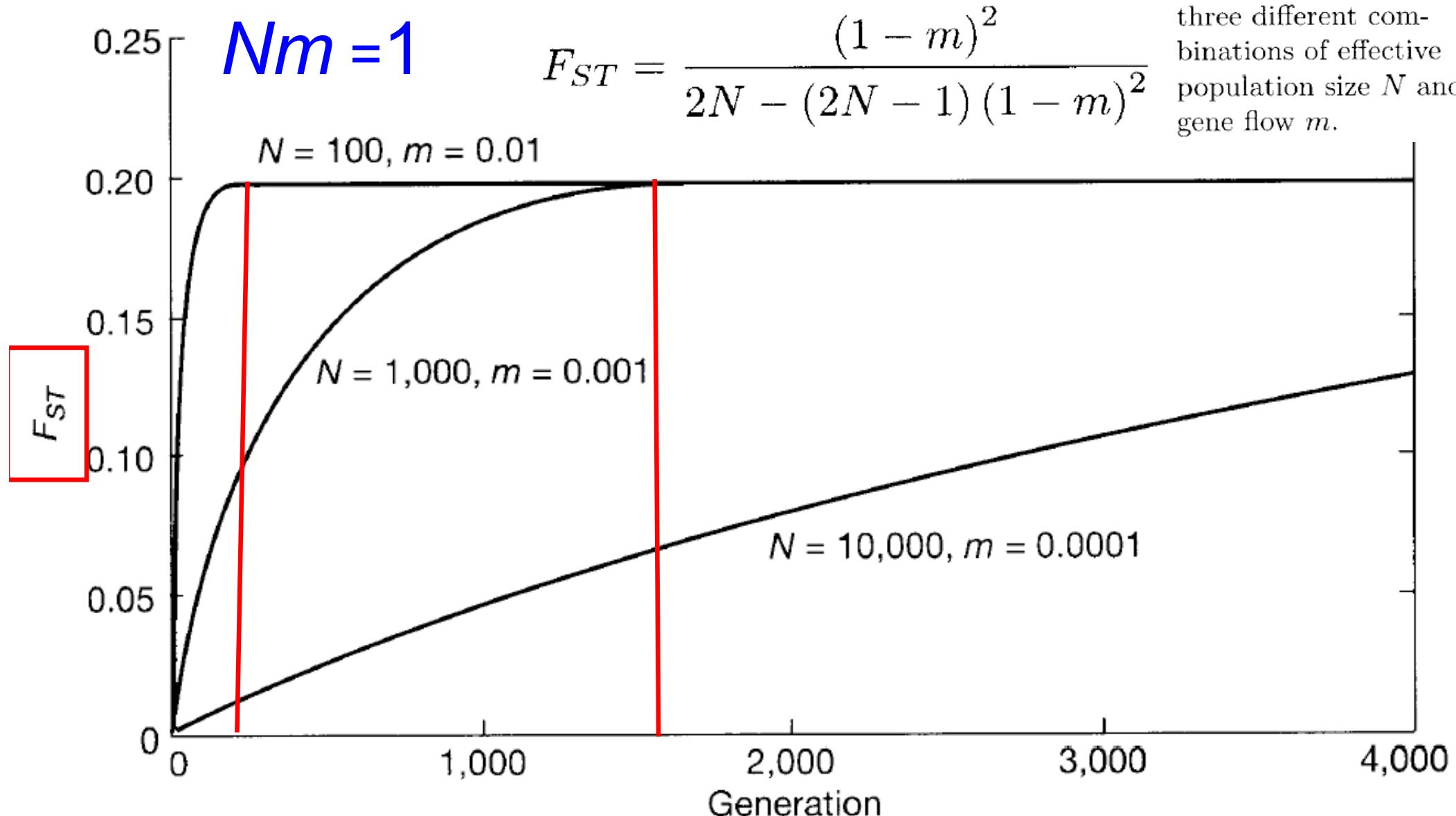
Dinámica **pura deriva (sin flujo)**
de 0 primero, llega a 1..., según t y N

This expression ranges from near 0 in the early generations and approaches unity when genetic drift over time has resulted in complete divergence between the populations. From this expression, the amount of F_{ST} is expected to increase at a nearly linear rate at low values and then asymptotically approach unity at high values.

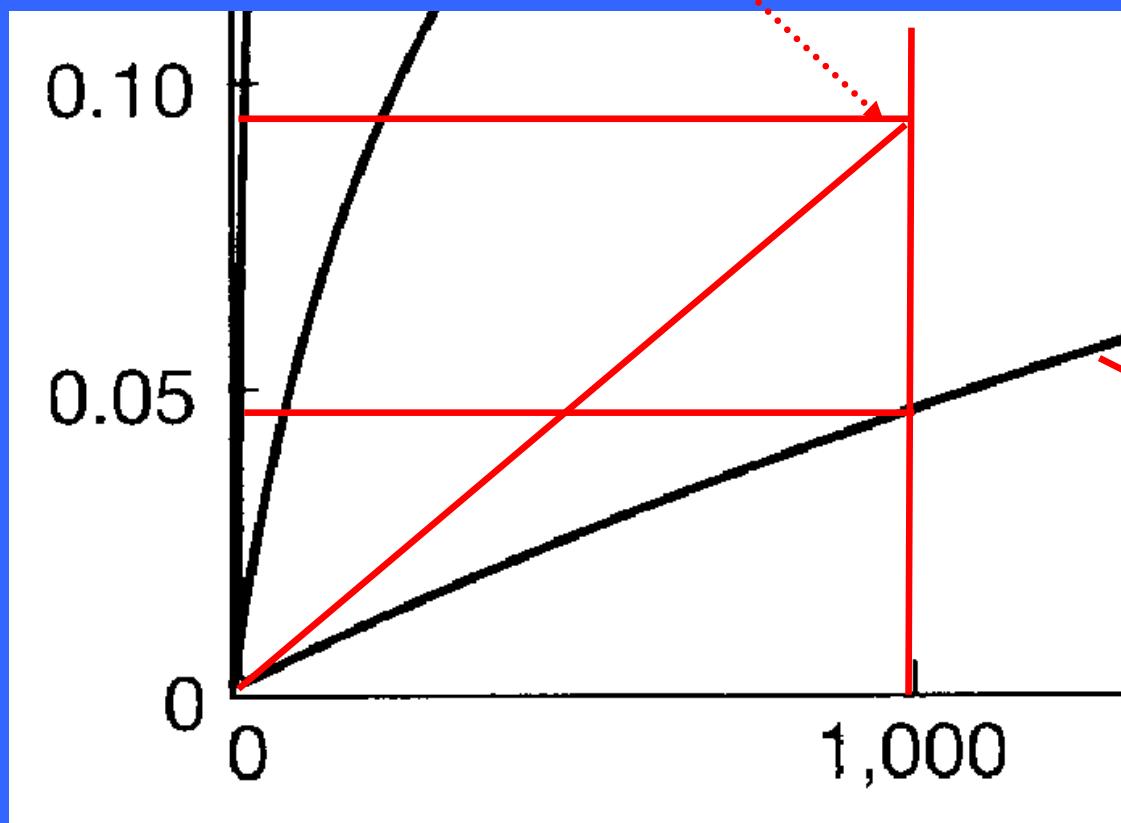
combinations of N and m . Figure 9.11 gives F_{ST} over time when the initial value of $F_{ST} = 0$ and $Nm = 1$. Obviously, when the effects of genetic drift

Se puede llegar rápido a la F_{st} en eq.
si m es grande o N chico. Iterando la fórmula

Figure 9.11. The amount of differentiation (F_{ST}) expected over generations for three different combinations of effective population size N and gene flow m .



difference in F_{ST} at a given point in time with and without gene flow. For example, with $N = 10,000$ using expression 9.13a after 1000 generations, $F_{ST} = 0.095$. With $m = 0.0001$ as in Figure 9.11, after 1000 generations $F_{ST} = 0.044$, only 46% as much.



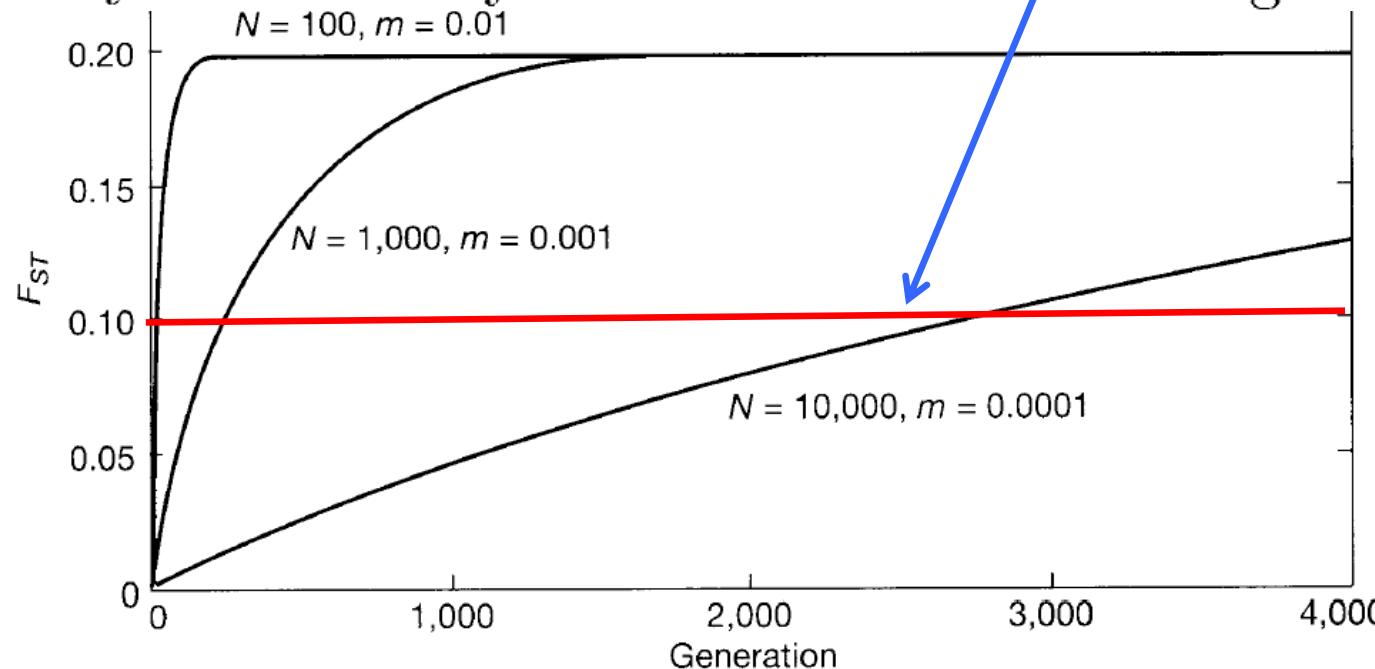
Poco flujo génico
($m = 0.001$) hace
mucho más **lenta**
la diferenciación
que SIN flujo,
 $N = 10,000$

$$F_{ST} \approx \frac{1}{4Nm + 1}$$

Crow and Aoki (1984) showed that the time for G_{ST} to go half way to equilibrium is approximately

$$t_{0.5} \approx \frac{\ln(2)}{(2m + 1/2N)}$$

showing explicitly that the rate of approach to equilibrium is faster as the values of m and N increase, that is, as the effects of gene flow and genetic drift increase. For example using this expression, when $N = 100$ and $m = 0.01$, $N = 1000$ and $m = 0.001$, and $N = 10,000$ and $m = 0.0001$, then it takes about 28, 277, and 2773 generations, respectively, to go halfway to the equilibrium frequencies (see also Figure 9.11). As is apparent for $N = 1000$ and $m = 0.001$ in Figure 9.11, the final approach to the equilibrium may be relatively slower than the time to go halfway to the equilibrium



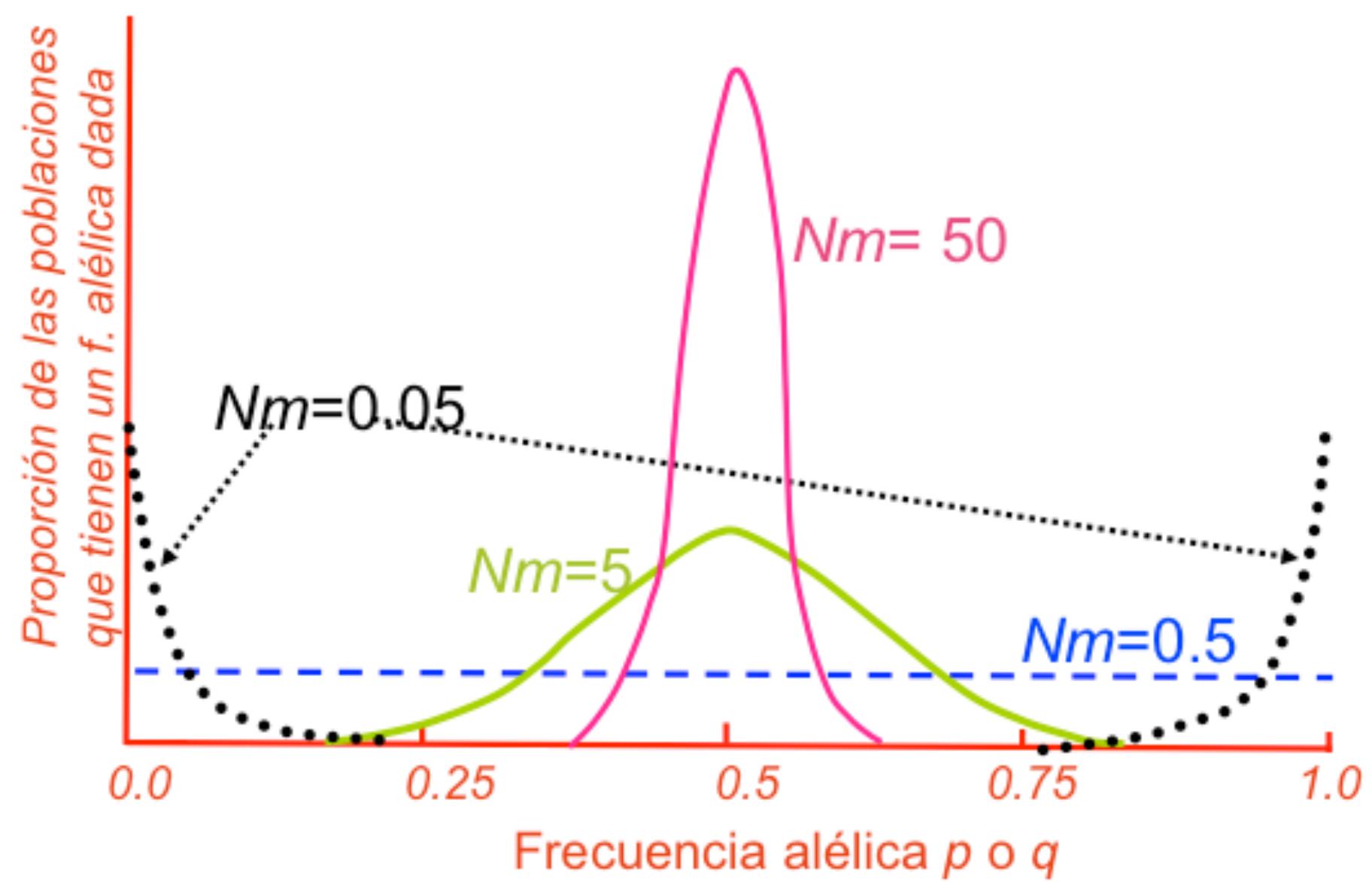
Tiempo
a $\frac{1}{2}$ de F_{st}
si $m=0.1$ 28 g
 $m=.001$, 277g
 $m=.0001$, 2773

Resultados de Wright 1940 en el equilibrio flujo/deriva si A_2 se mantiene constante = q_m Pico en q_m , si $4Nm$ mucho mayor que 1

Wright (1940) gave an explicit way of combining the effects of gene flow and genetic drift to predict the distribution of allele frequencies over islands. Assume that the frequency of A_2 in the migrants is constant and is equal to q_m . If we examine a large number of island populations, their average allele frequency will be q_m , but depending on the population size and the amount of gene flow, the distribution of allele frequency over islands will vary. For

N and m are large,

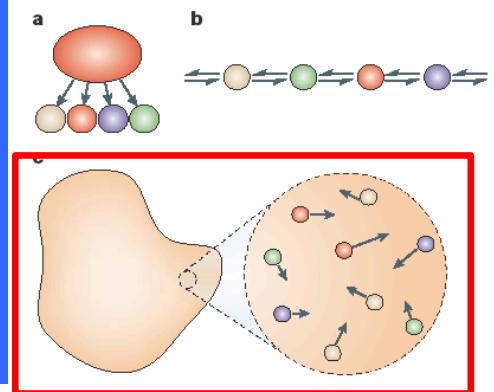
approach q_m , and their distribution will have a large peak around q_m . The shape of the distribution of allele frequencies over islands is related to the size of $4Nm q_m$ and $4Nm(1 - q_m)$. If these values are both much greater than one, then the island frequencies will be very close to each other and to that of the continent. In fact, if $4Nm \gg 1$ —that is, $m \gg 1/4N$ —or there is much more than one migrant every four generations where Nm is the number of migrants, then there will be virtually no differentiation among the island populations (e.g., $Nm = 50$ in Figure 9.12). On the other



Una Nm alta, más de 5 (ca. $F_{st} = 0.05$ o menor), mantiene la variación (domina flujo). Una Nm chica, menor de 1 (ca. $F_{st} = 0.2$ o mayor), domina la deriva y se pierde la variación genética.

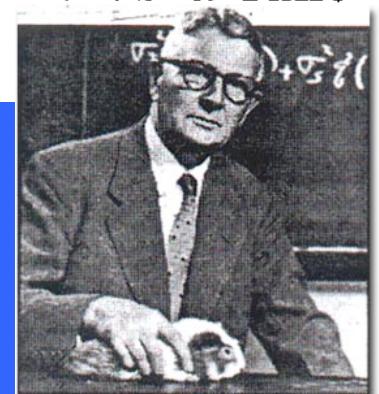
Isolation by Distance: el aislamiento por distancia

S. Wright 1943



gene flow when individuals were randomly distributed and there was **isolation by distance** between individuals. Slatkin (1991) and Rousset (1997) have suggested that the amount of genetic divergence as estimated by Nm (expression 9.12e) or $F_{ST}/(1 - F_{ST})$, respectively, should change in a linear fashion with the inverse of geographic distance and the geographic distance, respectively, between pairs of populations along a linear habitat (these measures are scaled inverses of each other). There are a number genetic **distance measures** that have been used to examine the relationship with geographic distance (Paetkau *et al.*, 1997; Hardy *et al.*, 2003) and a variety of statistical approaches to examine genetic divergence as a function of geographic distance (Epperson, 2003). Example 9.8

Nm , $F_{st}/(1-F_{st})$ o D de Nei vs.
distancia geográfica (Mantel test)



populations in Arizona and southern California, Gutiérrez-Espeleta *et al.* (2000) determined genetic variation at 10 highly variable microsatellite loci.

Overall, they found extensive genetic variation in most of the populations, nearly as much as in a population from Alberta that had not undergone a great reduction in numbers. When the amount of genetic divergence between these groups is primarily determined by isolation by distance and not greatly influenced by subspecies designation.

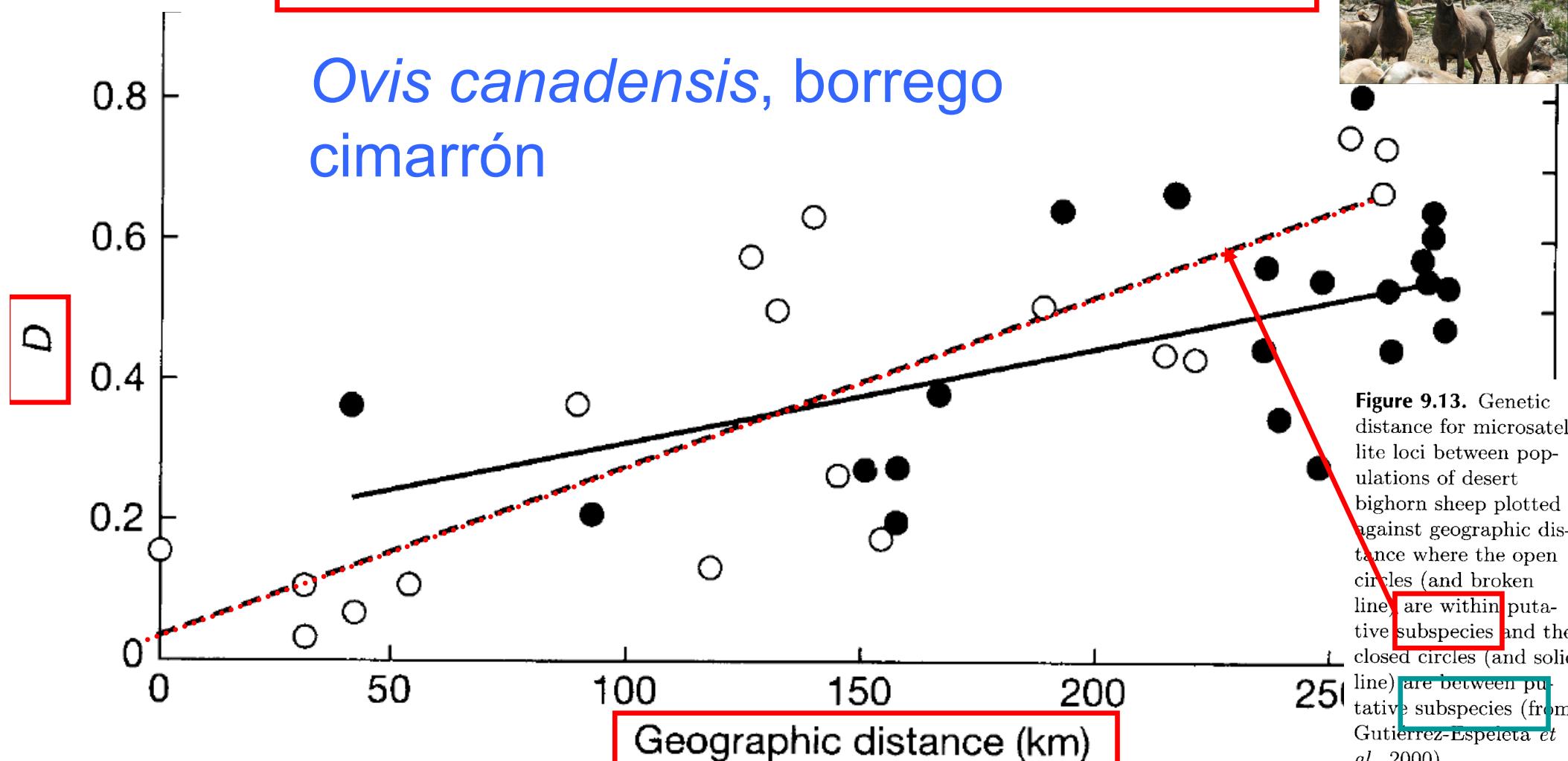
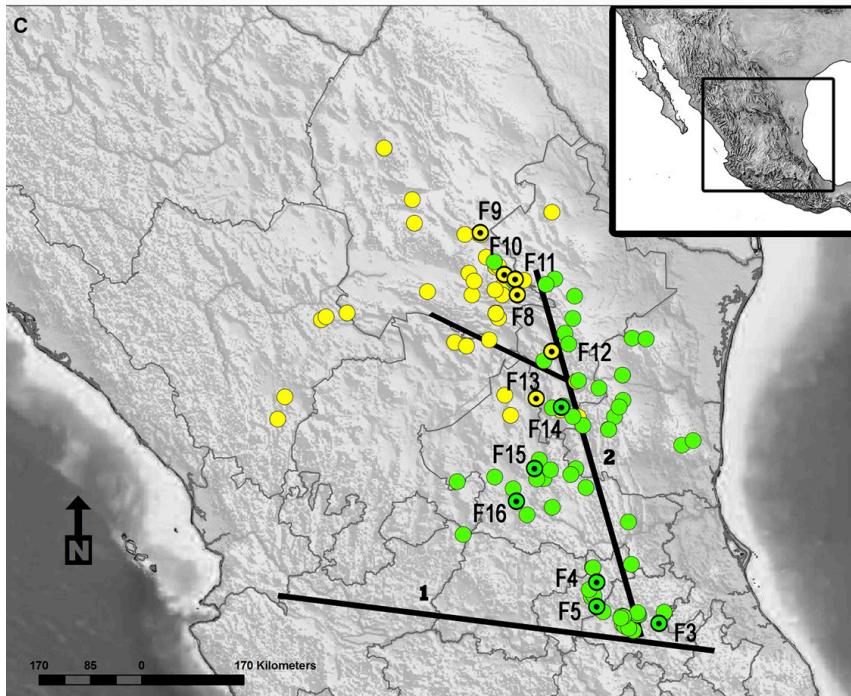


Figure 9.13. Genetic distance for microsatellite loci between populations of desert bighorn sheep plotted against geographic distance where the open circles (and broken line) are within putative subspecies and the closed circles (and solid line) are between putative subspecies (from Gutiérrez-Espeleta *et al.*, 2000).

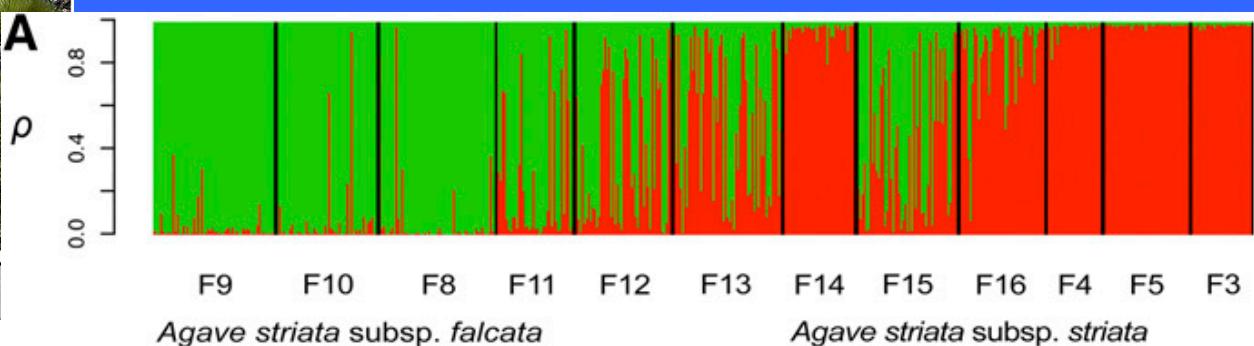


Population genetic analysis and bioclimatic modeling in *Agave striata* in the Chihuahuan Desert indicate higher genetic variation and lower differentiation in drier and more variable environments¹

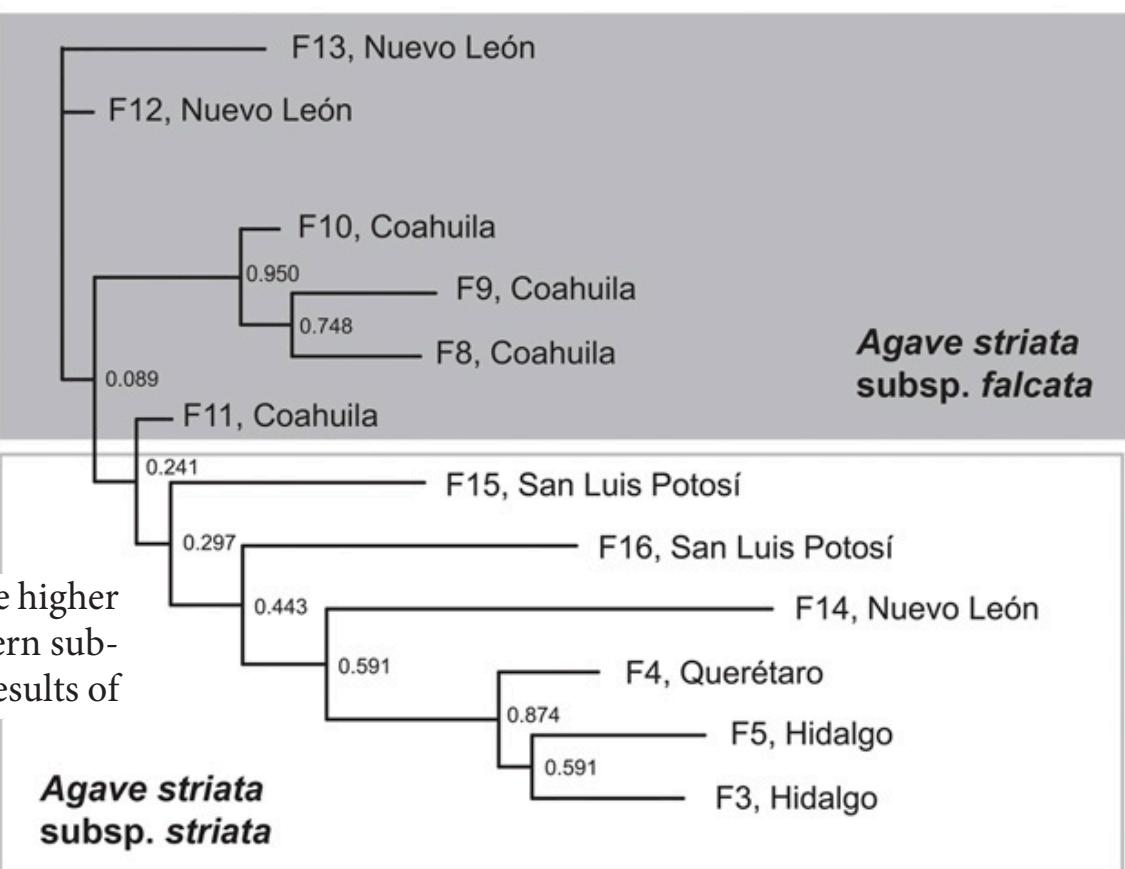
Laura Trejo^{2,5}, Leonardo O. Alvarado-Cárdenas³, Enrique Scheinvar⁴, and Luis E. Eguiarte^{4,5}



southern subspecies, *A. striata* subsp. *striata*, had on average higher population differentiation ($\Theta = 0.4085 \pm 0.010$) than northern subspecies, *A. striata* subsp. *falcata* ($\Theta = 0.1362 \pm 0.011$). The results of



F9 F10 F8 F11 F12 F13 F14 F15 F16 F4 F5 F3
Agave striata subsp. falcata Agave striata subsp. striata



ISSRs

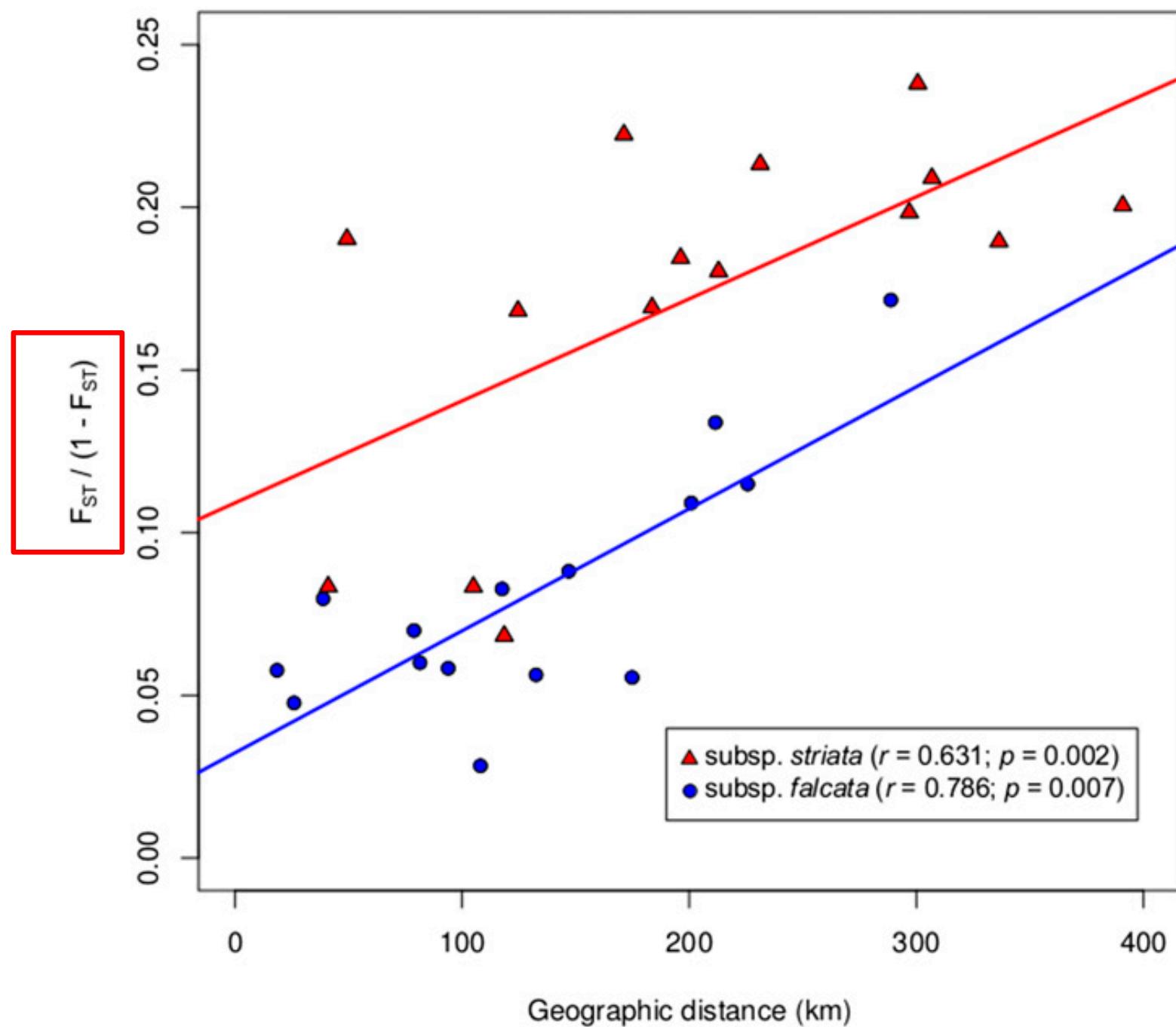
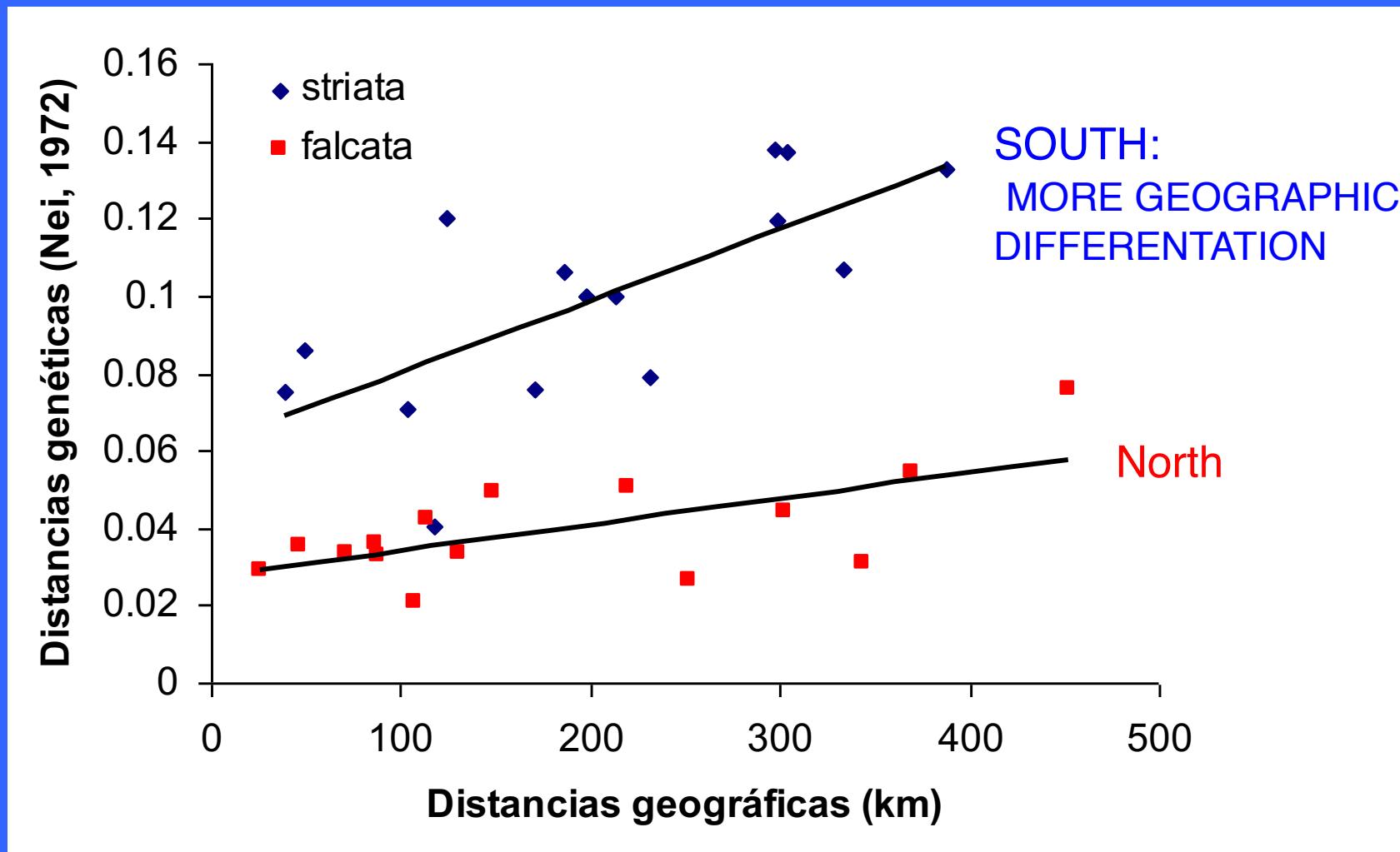


FIGURE 4 Mantel test analysis to evaluate isolation by distance within each subspecies of *Agave striata*.

¿Active speciation? Isolation by distance, Distancias de Nei



$$D = -\ln(I)$$
$$= J_{xy} / (J_x J_y)^{1/2}$$

Agave striata striata & A.s.falcata

Hay correlaciones taxonómicas y ecológicas con la diferenciación

F_{st}

Una Nm alta, más de 5 (ca.

$F_{st} = 0.05$ o menor), mantiene la variación (domina flujo). Una Nm chica, menor de 1 (ca. $F_{st} = 0.2$ o mayor), domina la deriva y se pierde la variación genética.

Species	F_{ST}
Mammals (57 species)	0.24
Birds (23 species)	0.05
Reptiles (22 species)	0.26
Amphibians (33 species)	0.32
Fish (79 species)	0.14
Insects (46 species)	0.10
Plants	
Selfing	0.51
Mixed selfing and outcrossing	
Animal pollination	0.22
Wind pollination	0.10
Outbreeding	
Animal pollination	0.20
Wind pollination	0.10

Revisiones: Isoenzimas:

J. L. HAMRICK AND M.J. W. GODT

Phil. Trans. R. Soc. Lond. B (1996) 351, 1291-1298

Effects of life history traits on genetic diversity
in plant species

Life form, geographic range, breeding system and taxonomic status had significant effects on the partitioning of genetic diversity within and among plant populations.

¿Correlaciona la ecología y evolución con la variación genética??

1996: More than 2,200 studies have reported allozyme variation for seed plants.

735 entries (species/ studies) supplied useful data.

Five traits that had the greatest influence on the levels and distribution of genetic diversity. (1) breeding system, (2) seed dispersal mechanism, (3) life form, (4) geographic range and (5) taxonomic status

Table 1. *Categories of each of the five life history traits used to produce the two-trait combinations*

(See Hamrick & Godt 1989 for more complete explanation of the traits.)

categories for each trait

breeding system	seed dispersal mechanism	life form	geographic range	taxonomic status
outcrossing	attached	annual	endemic	gymnosperm
mixed mating	gravity	short-lived perennial	narrow	dicotyledon
selfing	ingested	long-lived perennial	regional	monocotyledon
	wind		widespread	

Table 3. Mean levels of genetic variation within species and its distribution among populations for combined categories of breeding system and seed dispersal mechanism

trait combination	<i>n</i>	P_s (%)	H_{es}	G_{ST}
outcrossing				diferenciación
attached	63	67.9 ^a	0.188 ^b	0.114 ^d
gravity	178	50.2 ^b	0.152 ^{cde}	0.189 ^c
ingested	54	52.4 ^b	0.200 ^{ab}	0.223 ^c
wind	186	62.4 ^a	0.157 ^{cd}	0.101 ^d
mixed-mating				
gravity	63	52.7 ^b	0.174 ^{bc}	0.248 ^c
ingested	17	34.1 ^e	0.108 ^{ef}	0.269 ^c
wind	62	42.0 ^{bc}	0.118 ^{ef}	0.175 ^{cd}
selfing				
attached	29	64.7 ^a	0.236 ^a	0.426 ^b
gravity	94	34.5 ^e	0.097 ^f	0.533 ^a

Table 10. *Mean levels of genetic variation within species and its distribution among populations for several plant families*

family	<i>n</i>	P_s (%)	H_{es}	G_{ST}
Asteraceae	101	45.3	0.127	0.204
Chenopodiaceae	22	40.6	0.099	0.540
Cucurbitaceae	23	40.4	0.168	0.397
Onagraceae	23	34.4	0.106	0.338
Orchidaceae	16	44.8	0.137	0.087
Schrophulariaceae	16	37.2	0.123	0.372
Solanaceae	23	32.0	0.094	0.426
Poaceae	91	62.7	0.201	0.284
Fabaceae	48	59.6	0.184	0.277
Myrtaceae ^a	14	81.8	0.222	0.134
Fagaceae ^a	27	65.3	0.198	0.085
Pinaceae ^a	103	73.0	0.176	0.073

^a Families with predominantly woody taxa.

Comparación Marcadores derivados del PCR: ISSR, RAPDS, AFLPs e isoenzimas

HILDE NYBOM

Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants
Molecular Ecology (2004) 13, 1143–1155

The grand mean for RAPD-derived within-population gene diversity H was 0.214, which is close to the allozyme-derived $H_S = 0.230$ (Hamrick & Godt (1989)).
o sea, se portan parecido...

Table 1 Database of the review on life history traits and diversity obtained with RAPD, AFLP and ISSR markers. Number of studies, mean and standard deviation are given for each of four sampling strategy parameters: number of populations, number of plants per population, maximum geographical distance between sampled populations and number of polymorphic markers, and for three genetic parameters: AMOVA-derived F_{ST} , Nei's G_{ST} , and mean within-population diversity, H_{pop}

Parameter	RAPD N	Mean ± SD	AFLP N	Mean ± SD	ISSR N	Mean ± SD
Populations	158	7.9 ± 6.4	27	14.0 ± 14.8	13	10.3 ± 7.7
Plants	156	18.5 ± 14.6	27	14.5 ± 9.0	12	17.5 ± 8.0
Distance (km)	152	956 ± 1880	26	1547 ± 2664	12	1315 ± 2335
Markers	157	72.3 ± 59.5	27	238.1 ± 277.9	13	54.9 ± 19.7
Φ_{ST}	116	0.34 ± 0.24	21	0.35 ± 0.18	9	0.35 ± 0.25
G_{ST}	46	0.27 ± 0.21	12	0.21 ± 0.14	6	0.34 ± 0.29
H_{pop}	60	0.22 ± 0.12	13	0.23 ± 0.08	4	0.22 ± 0.08

Se comportan de manera parecida entre ellos y con la isoenzimas pero son dominantes, y solo tienen dos “alelos” presencia y ausencia

GTGTGCAGACTGGCCTCTTCC

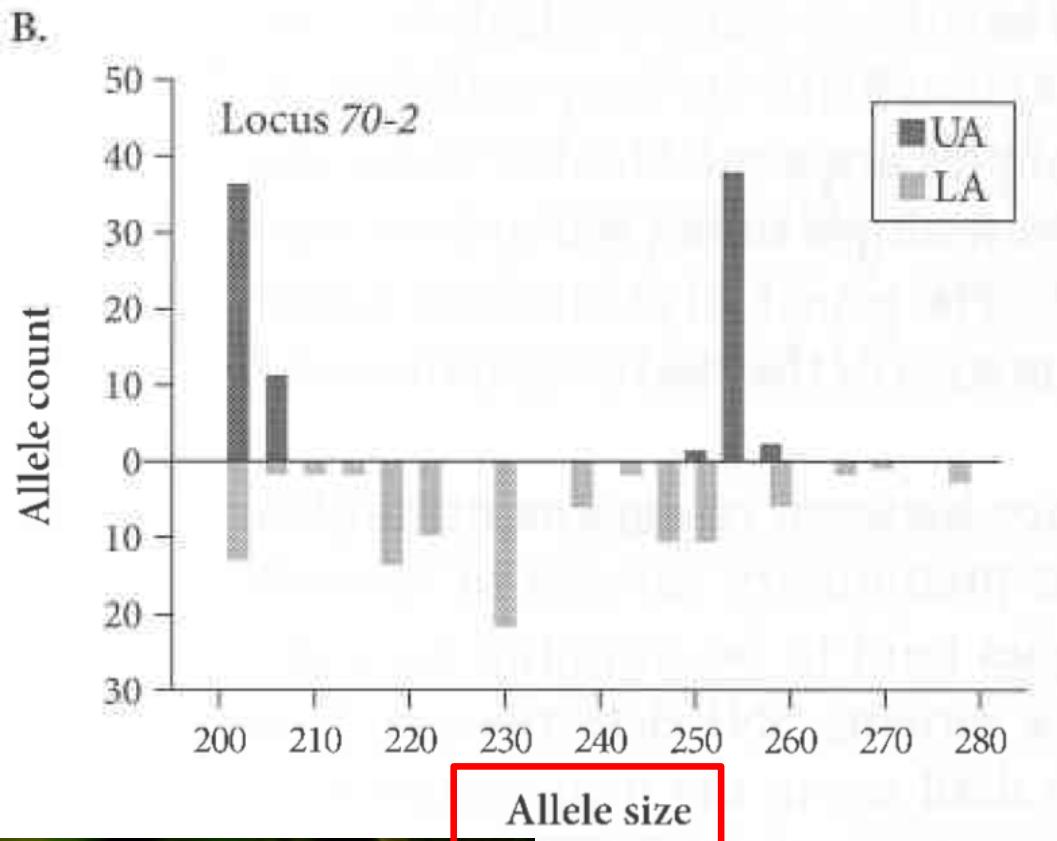
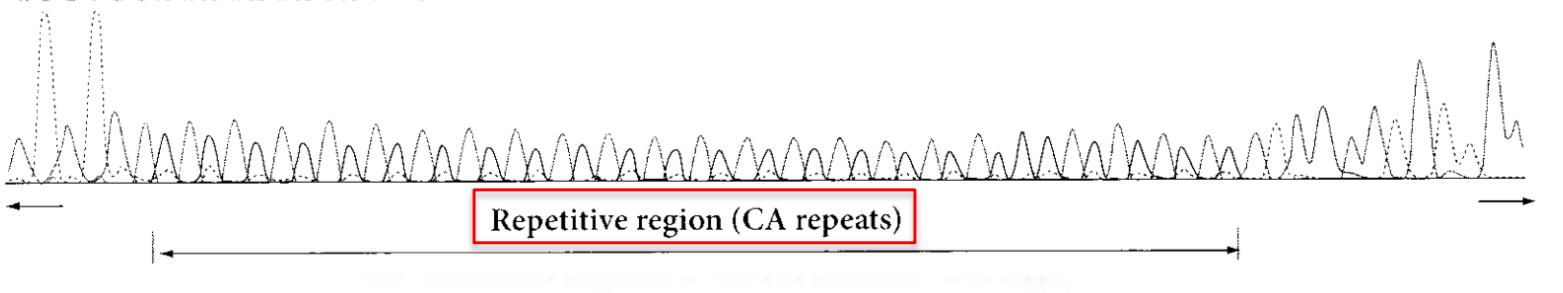


FIGURE 1.9 Microsatellite variability. **A.** Sequence read of a single allele of a microsatellite repeat region in the sea louse *Lepeophtheirus salmonis*. The figure shows the output of a DNA sequencing run. This allele has a succession of alternating CA bases, repeated 12 times (center part of the figure), while at the two ends some unique, nonrepetitive flanking sequence is seen. [Adapted from www.st-andrews.ac.uk/~merg/sea%20lice.htm.] **B.** Distribution of repeat numbers of a tetranucleotide microsatellite repeat in alleles sampled from natural populations of a fish (the guppy, *Poecilia reticulata*). The black bars show results from the upper Aripo River (Trinidad), and the grey bars show data from the lower river. [Adapted from Figure 1 of Oosterhout et al. (2006).]

Microsatélites =
SSR: *short sequence repeats*
STMS: *sequence tagged microsatellite sites*
Codominantes, se detectan heterocigos

Microsatélites, H , Nybom, 2004

Parameter	N	Mean ± SD
Populations	106	4.1 ± 6.1
Plants	104	51.5 ± 55.4
Distance (km)	37	1103 ± 2694
Loci	105	8.4 ± 6.7
Alleles	90	83.2 ± 63.0
F_{ST}	33	0.26 ± 0.17
R_{ST}	18	0.24 ± 0.21
H_E	104	0.61 ± 0.21
H_O	80	0.58 ± 0.22

Table 2 Database of the review on life history traits and STMS marker diversity. Number of studies, mean and standard deviation are given for each of five sampling strategy parameters: number of populations, number of plants per population, maximum geographical distance between sampled populations, number of polymorphic loci and number of polymorphic alleles, and for four genetic parameters: population differentiation measured with F_{ST} and R_{ST} , and mean within-population diversity measured as H_E and H_O

Ventajas: Más alta H que los otros marcadores, que en promedio su H eran 0.225, debido a que en promedio son 9.9 alelos por loci.
Muy altas tasas de mutación!
Desventajas: 1) Caros y difíciles de desarrollar.

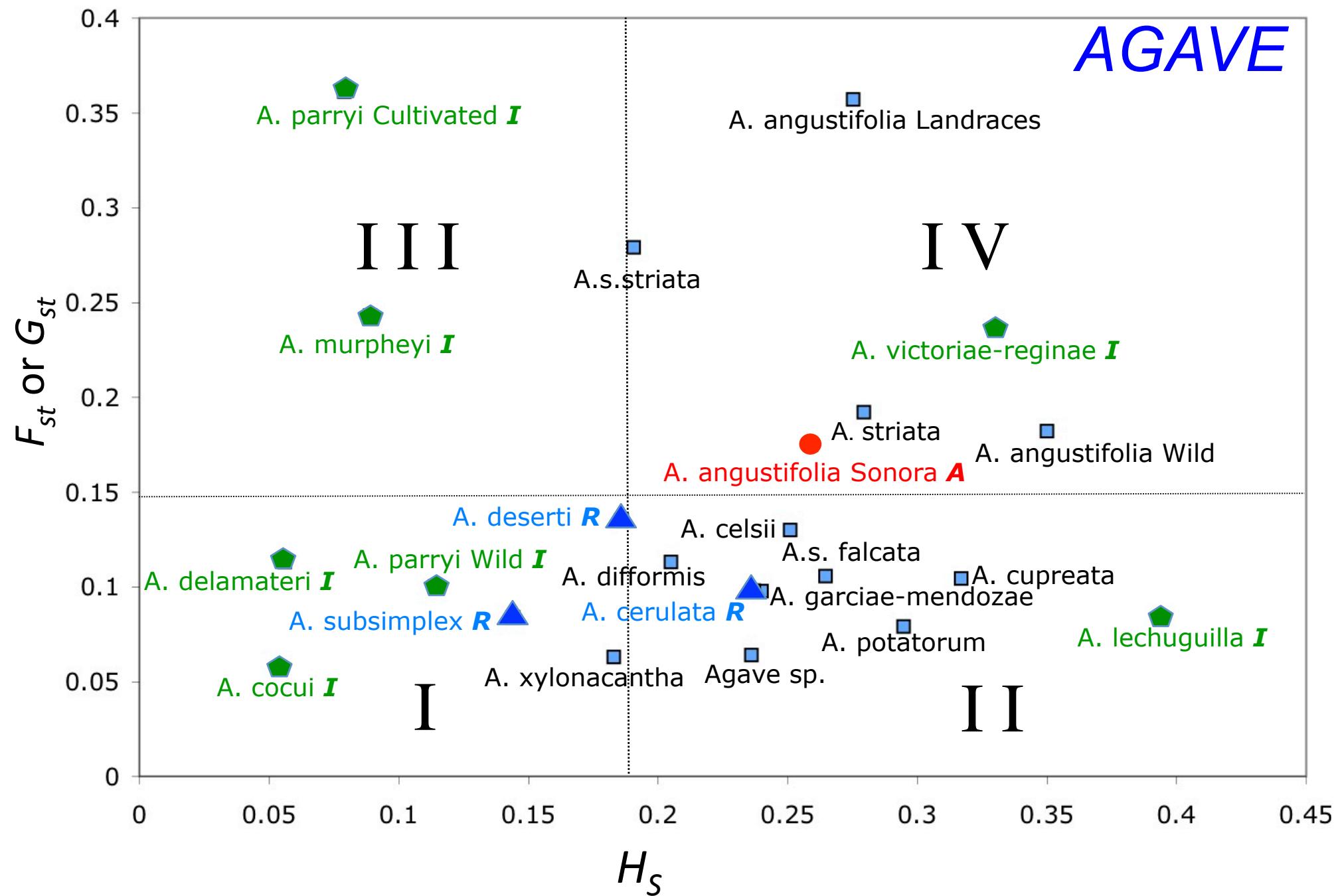
2) Homoplasia:
Puede haber convergencias
3) Alelos nulos,
parece haber más homocigos

Table 5 Estimates of within-population variation and/or among-population variation obtained with at least two different marker methods for the same set of samples (sometimes instead two overlapping but not totally identical subsets of samples). All of these studies except those concerning *Camellia sinensis* (one set of cultivated accessions), *Glycine max + G. soja* (12 genotypes), *Oryza sativa* (three groups of cultivated accessions) and the apomictic *Limonium dufourii* were also included in the main data compilation

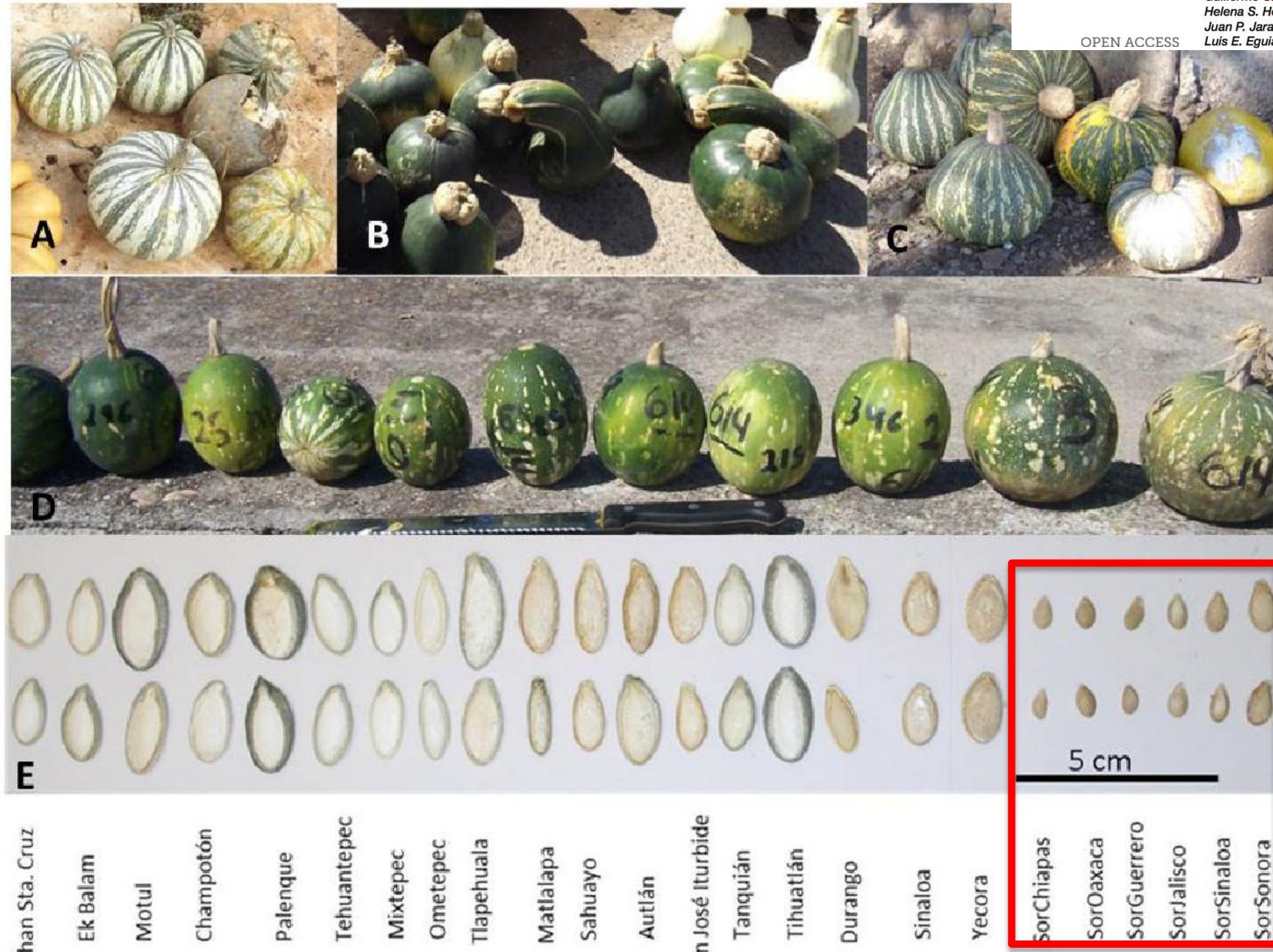
Species	Within-population variation	Among-population variation
<i>Avicennia marina</i>	0.19 (AFLP, H_E) 0.78 (STMS, H_F)	0.20 (AFLP, Φ_{ST}) 0.54 (STMS, Φ_{ST}) – –
<i>Camellia sinensis</i>	0.36 (AFLP, H) 0.31 (RAPD, H)	– –
<i>Elymus fibrosus</i>	0.10 (RAPD, H_E) 0.25 (STMS, H_E)	0.65 (RAPD, G_{ST}) 0.54 (STMS, G_{ST}) – –
<i>Glyine max + G. soja</i>	0.32 (AFLP, H_E) 0.31 (RAPD, H_E) 0.60 (STMS, H_E)	– – –
<i>Hordeum spontaneum</i>	0.16 (AFLP, H_E) 0.47 (STMS, H_E)	0.31 (AFLP, G_{ST}) 0.36 (STMS, G_{ST})
<i>Leucopogon obtectus</i>	–	0.10 (AFLP, Φ_{ST}) 0.13 (RAPD, Φ_{ST})
<i>Limonium dufourii</i>	–	0.53 (AFLP, Φ_{ST}) 0.52 (RAPD, Φ_{ST})
<i>Oryza granulata</i>	–	0.84 (ISSR, Φ_{ST}) 0.90 (RAPD, Φ_{ST})
<i>Oryza sativa</i>	–	0.63 (AFLP, Φ_{ST}) 0.34 (ISSR, Φ_{ST}) 0.42 (RAPD, Φ_{ST})
<i>Pinus contorta</i>	0.43 (RAPD, H_E) 0.73 (STMS, H_E)	0.94 (RAPD, F_{ST}) 0.97 (STMS, F_{ST})
<i>Pinus oocarpa</i>	0.34 (AFLP, H_{pop}) 0.36 (RAPD, H_{pop})	0.07 (AFLP, G_{ST}) 0.1 L (RAPD, G_{ST})
<i>Pinus pinaster</i>	0.160 (AFLP, H_S) 0.734 (STMS, H_S)	0.10 (AFLP, G_{ST}) 0.11 (STMS, G_{ST})
<i>Senecio layneae</i>	–	0.38 (ISSR, Φ_{ST}) 0.26 (RAPD, Φ_{ST})
<i>Swietenia macrophylla</i>	–	0.38 (AFLP, Φ_{ST}) 0.24 (STMS, F_{ST})

Micros más H
STMP:sequence
tagged
microsatellite
sites = micros
Todo lo demás
similar en los
diferentes
marcadores
(la diferenciación
similar en
distintos
marcadores)

AGAVE



Cucurbita argyropserma ssp. *sororia* silvestre y *C.a.* ssp. *argyropsermacultivada*, usada en semillas para moles como el pipian.

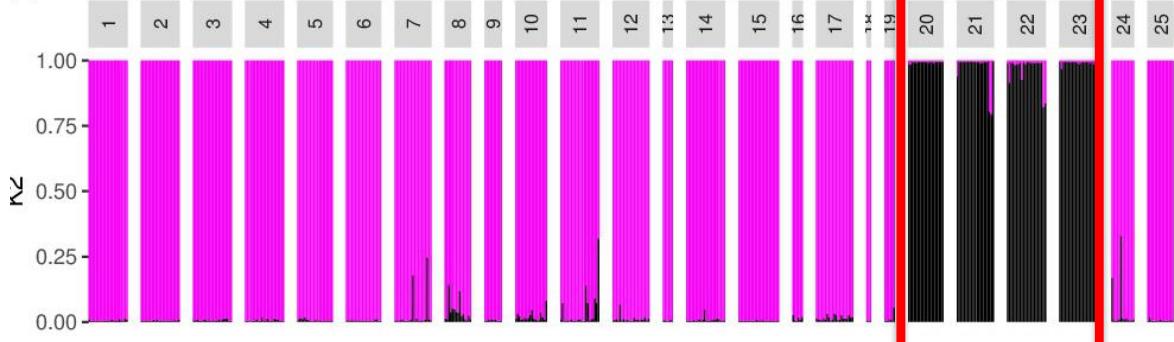
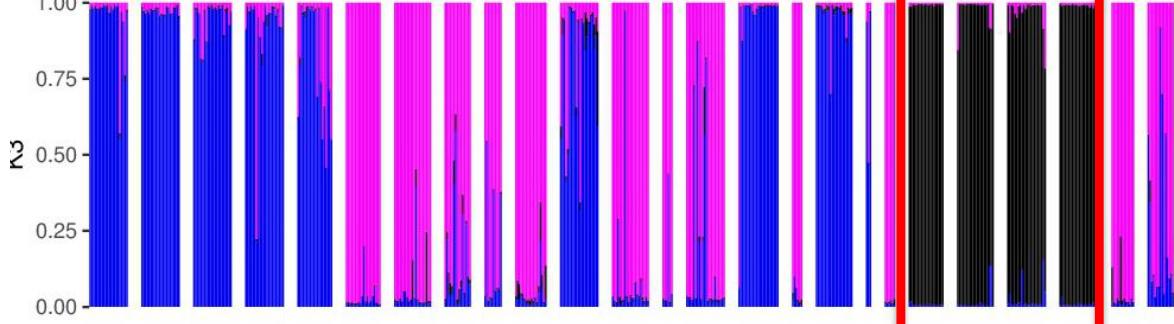
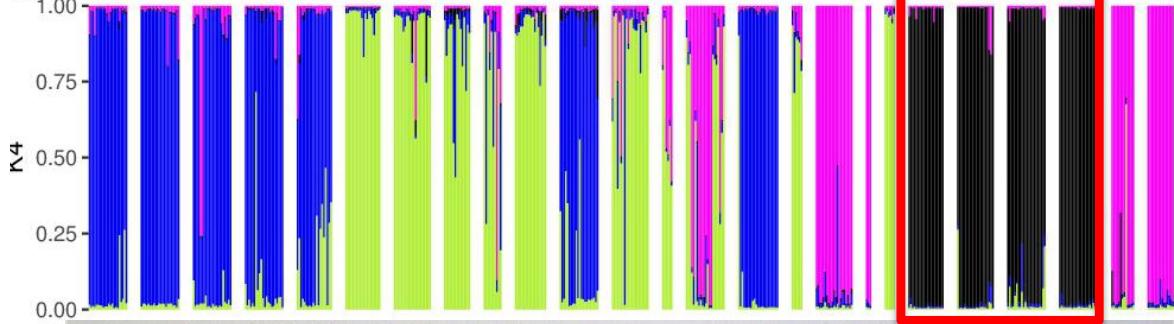


Genetic Resources in the “Calabaza Pipiana” Squash (*Cucurbita argyrosperma*) in Mexico: Genetic Diversity, Genetic Differentiation and Distribution Models

Guillermo Sánchez-de la Vega¹, Gabriela Castellanos-Morales^{1,2,3}, Niza Gámez¹, Helena S. Hernández-Rosales¹, Alejandra Vázquez-Lobo^{1,4}, Erika Aguirre-Planter¹, Juan P. Jaramillo-Correa¹, Salvador Montes-Hernández⁵, Rafael Lira-Saade^{2*} and Luis E. Eguiarte^{1*}

OPEN ACCESS



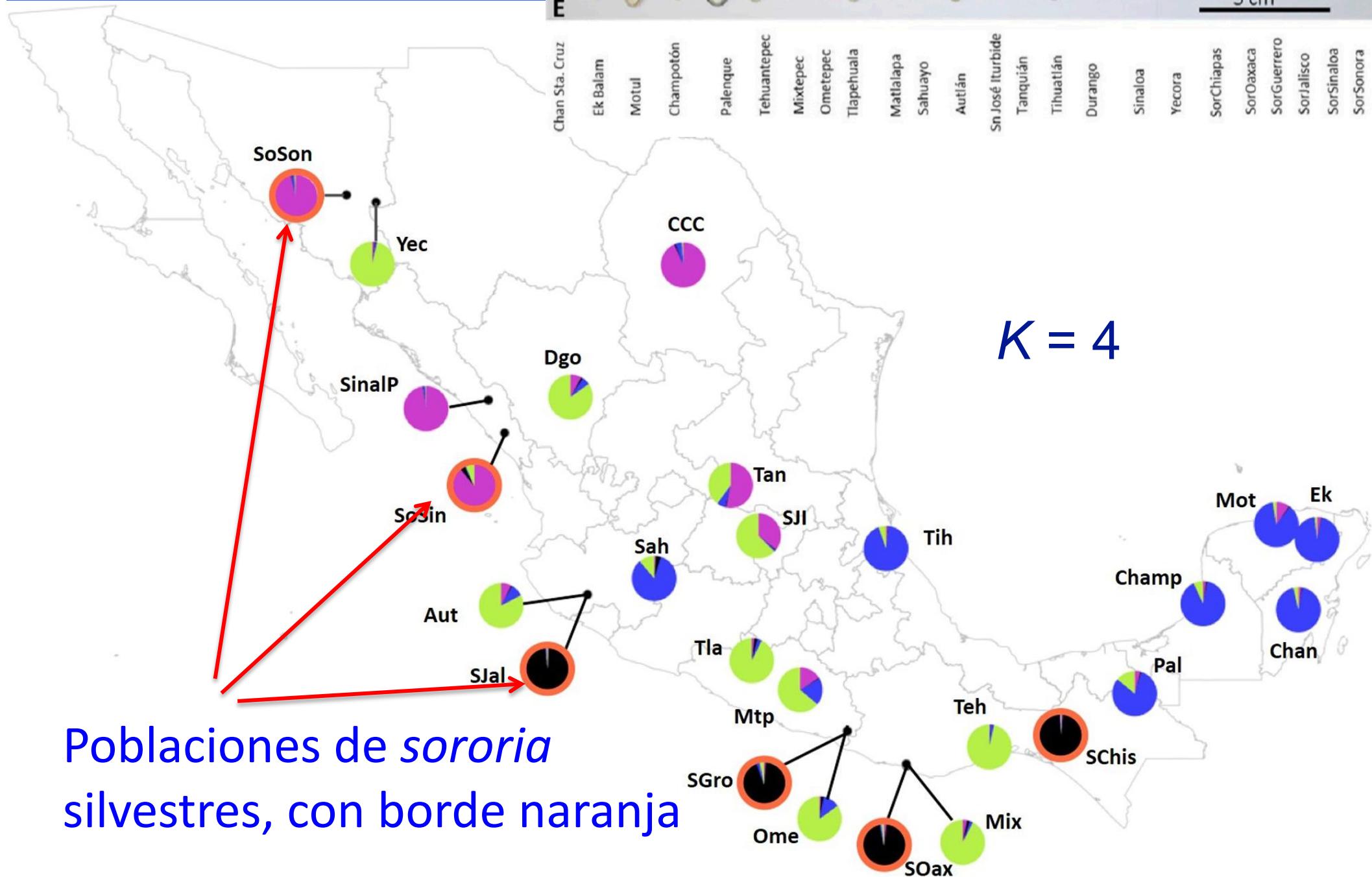
A**B****C****E**

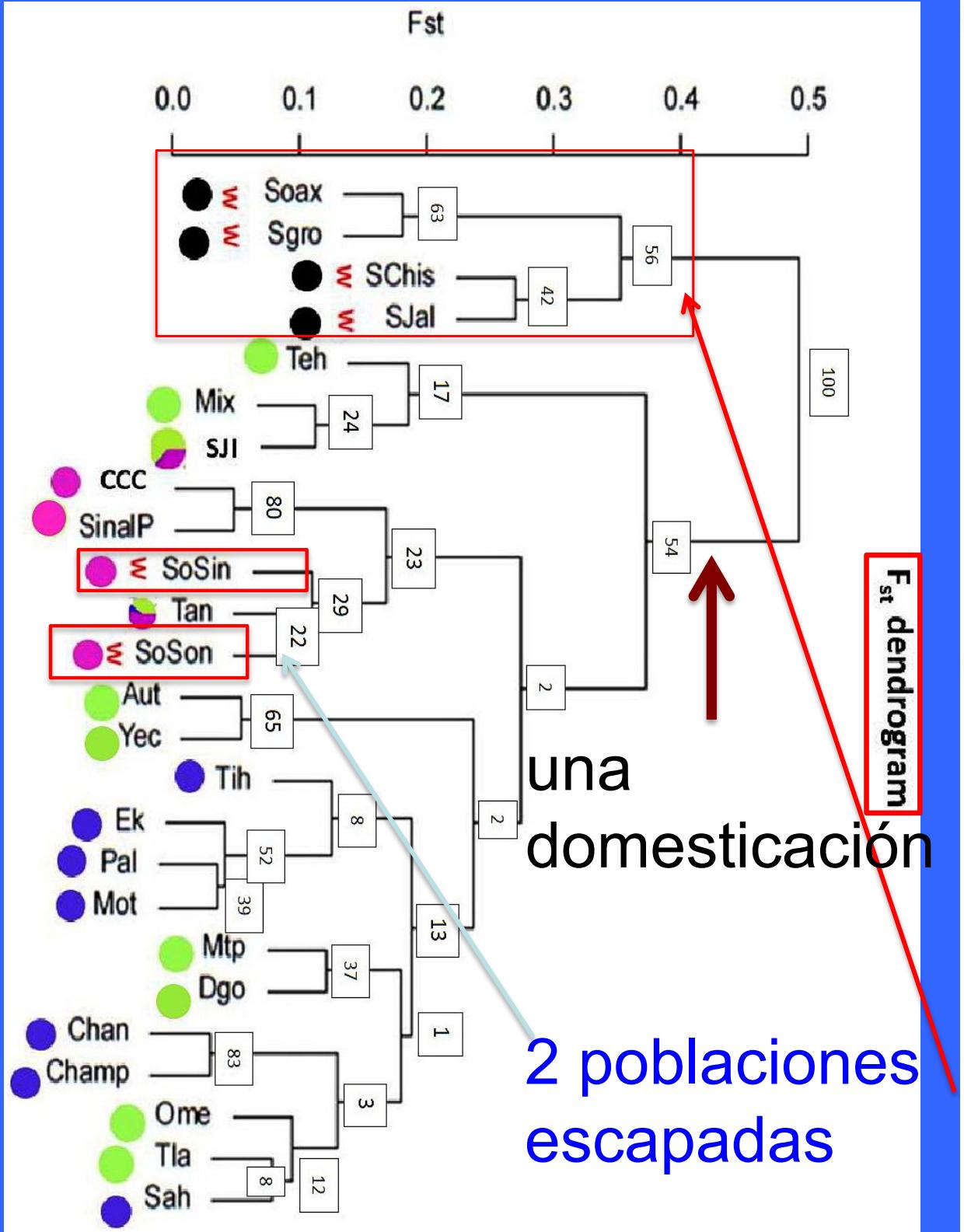
Análisis Structure
 $K=2$

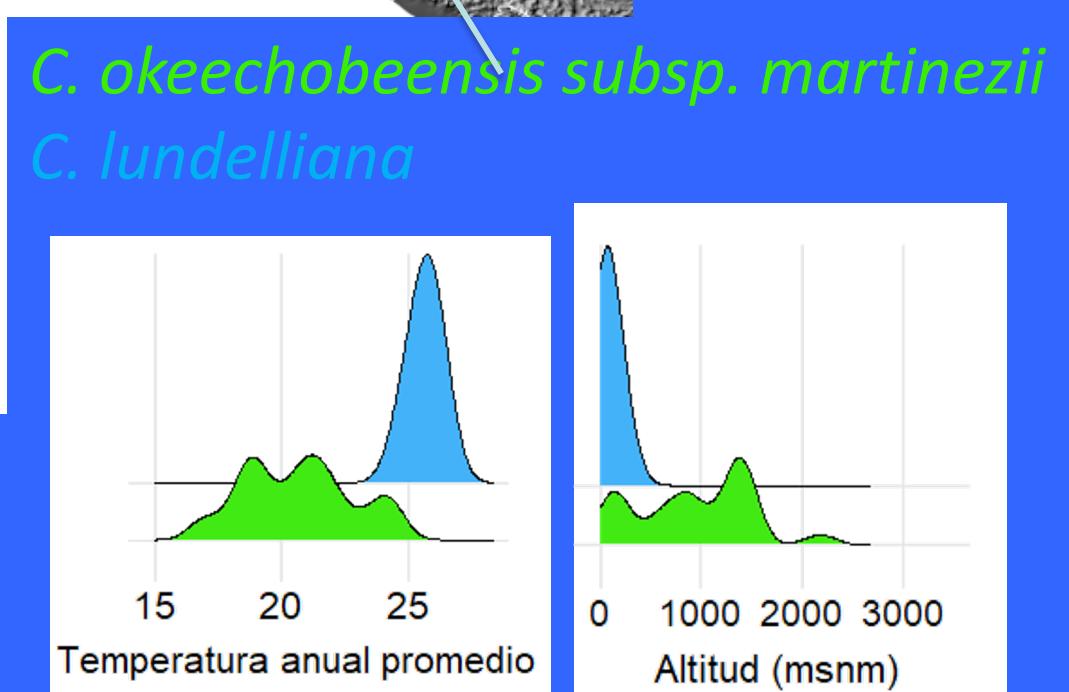
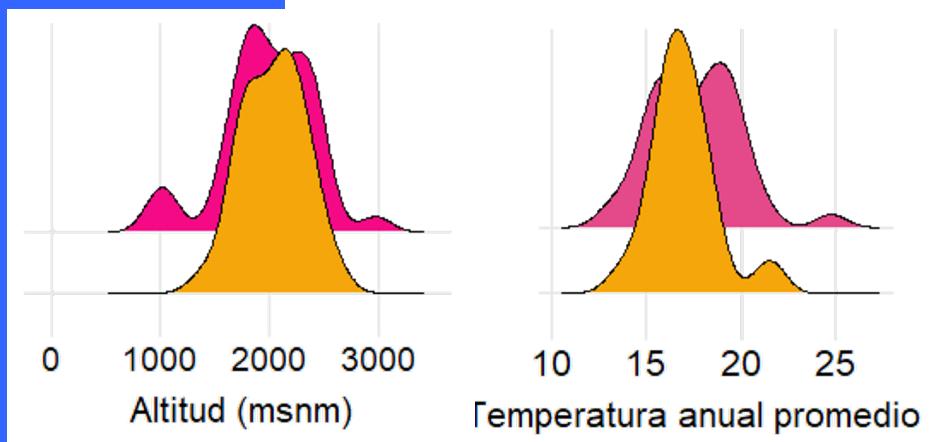
Cucurbita
argyrosperma
 $K=3$ 440 individuos
84 alelos
9 loci
 $K=4$ microsatélites.

En negro, las poblaciones
realmente silvestres:
rosa escapadas...

Cucurbita argyrosperma

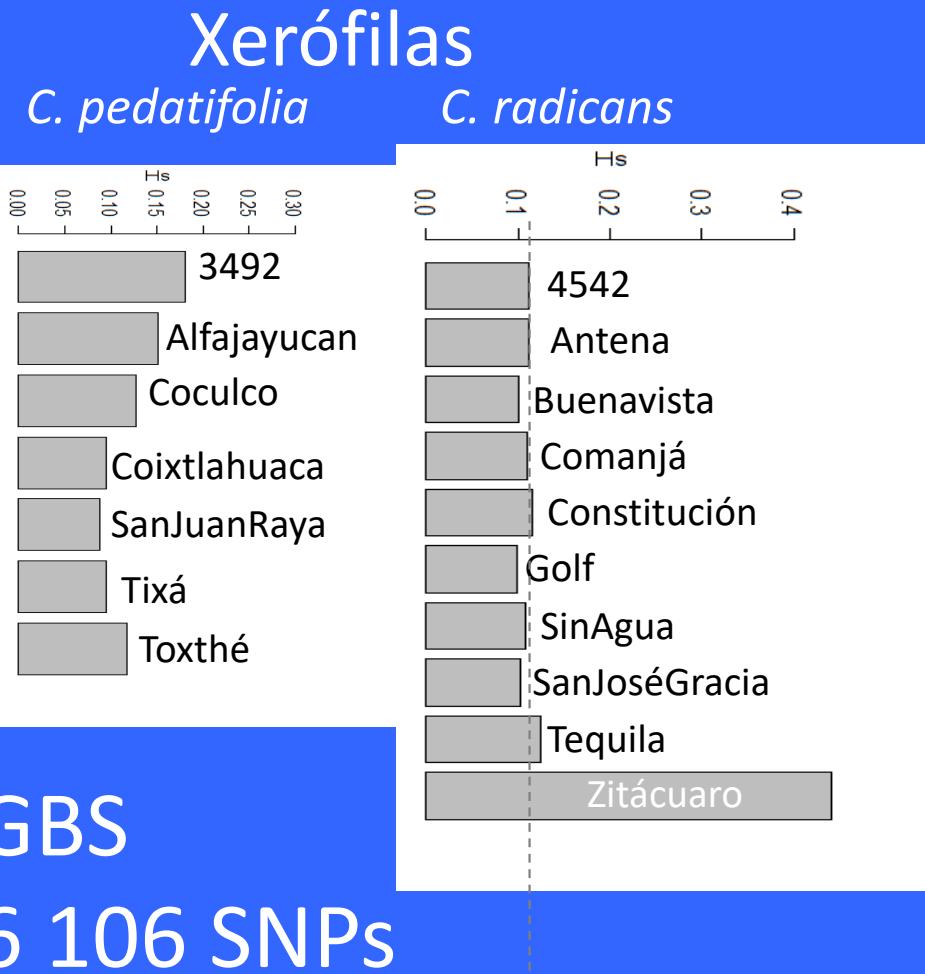
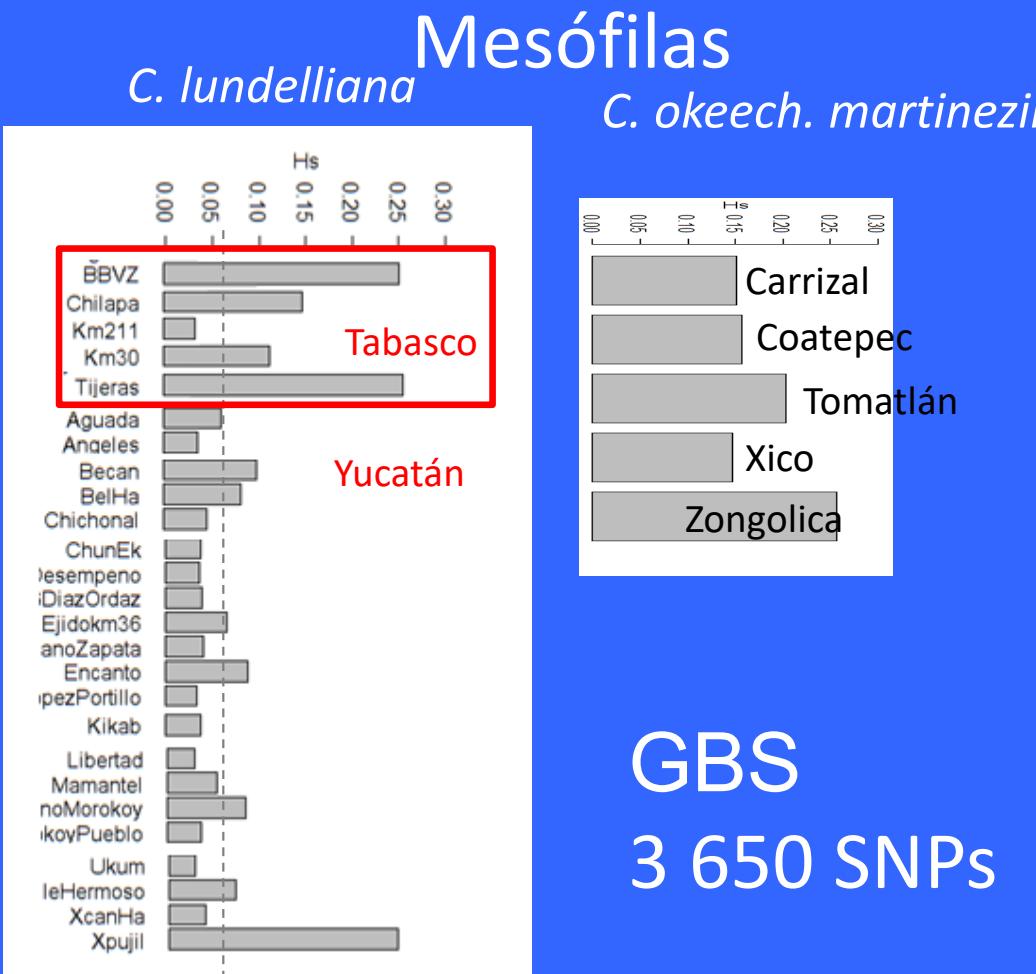






Patrones de diversidad; datos preliminares

Cucurbita silvestres



H_S 0.0567

0.2153

0.1433

0.1292

F_{ST} 0.1808

0.0940

0.2920

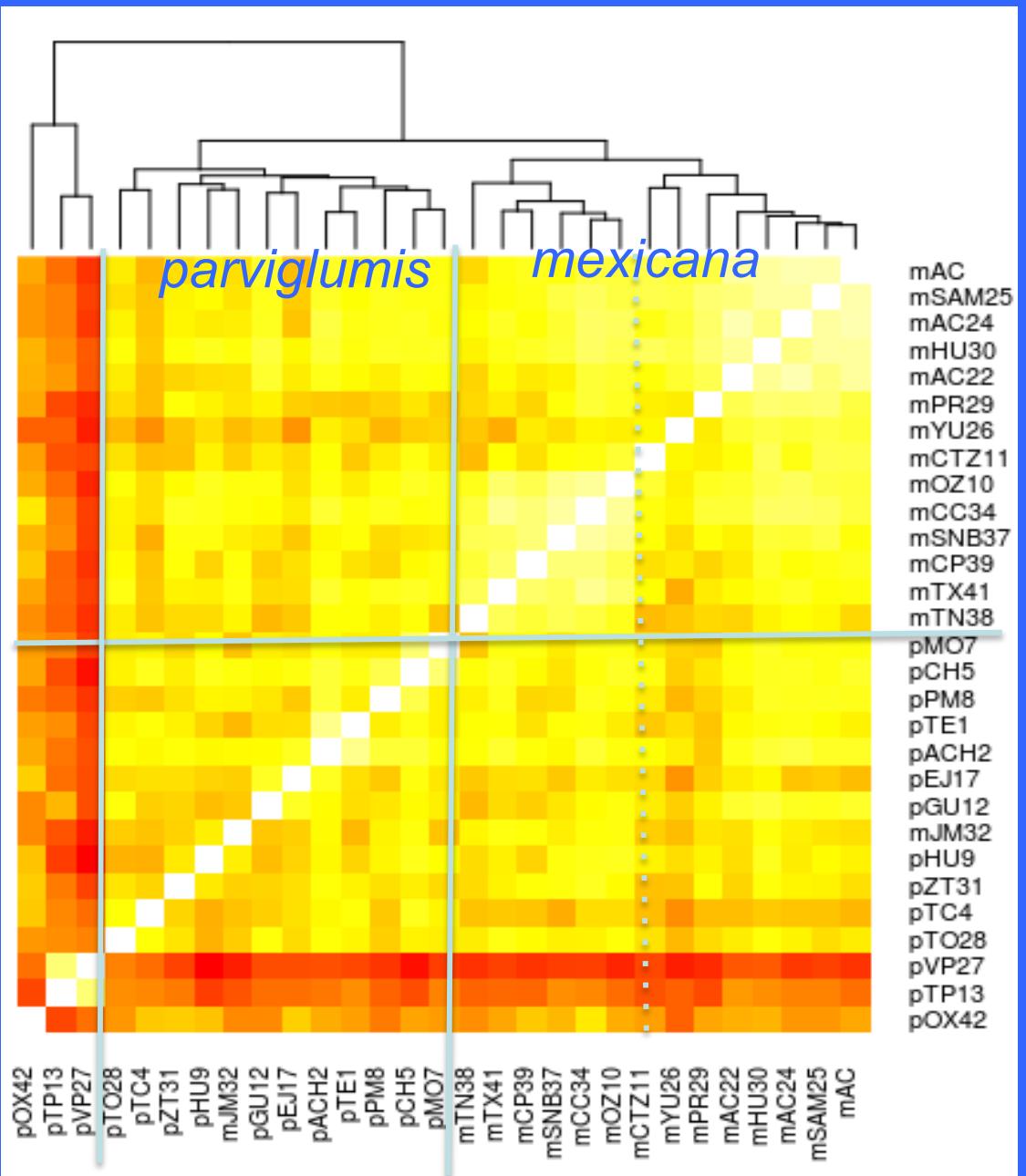
0.1438

F_{IS} -0.0449

-0.3141

0.0885

0.0151



Teosintes, 22 loci
microsatélites
 D_{st} de Jost
pareadas
Elevada
estructura dentro
y entre
subespecies de
3 grandes grupos
+2 subgrupos en
mexicana

Population Structure and Its Effects on Patterns of Nucleotide Polymorphism in Teosinte (*Zea mays* ssp. *parviglumis*)

David A. Moeller,^{*,1} Maud I. Tenaillon[†] and Peter Tiffin^{*,2}

7 populations 6-18 individuals per population, 84 in total.

5 nuclear loci,
2 chloroplast seq.

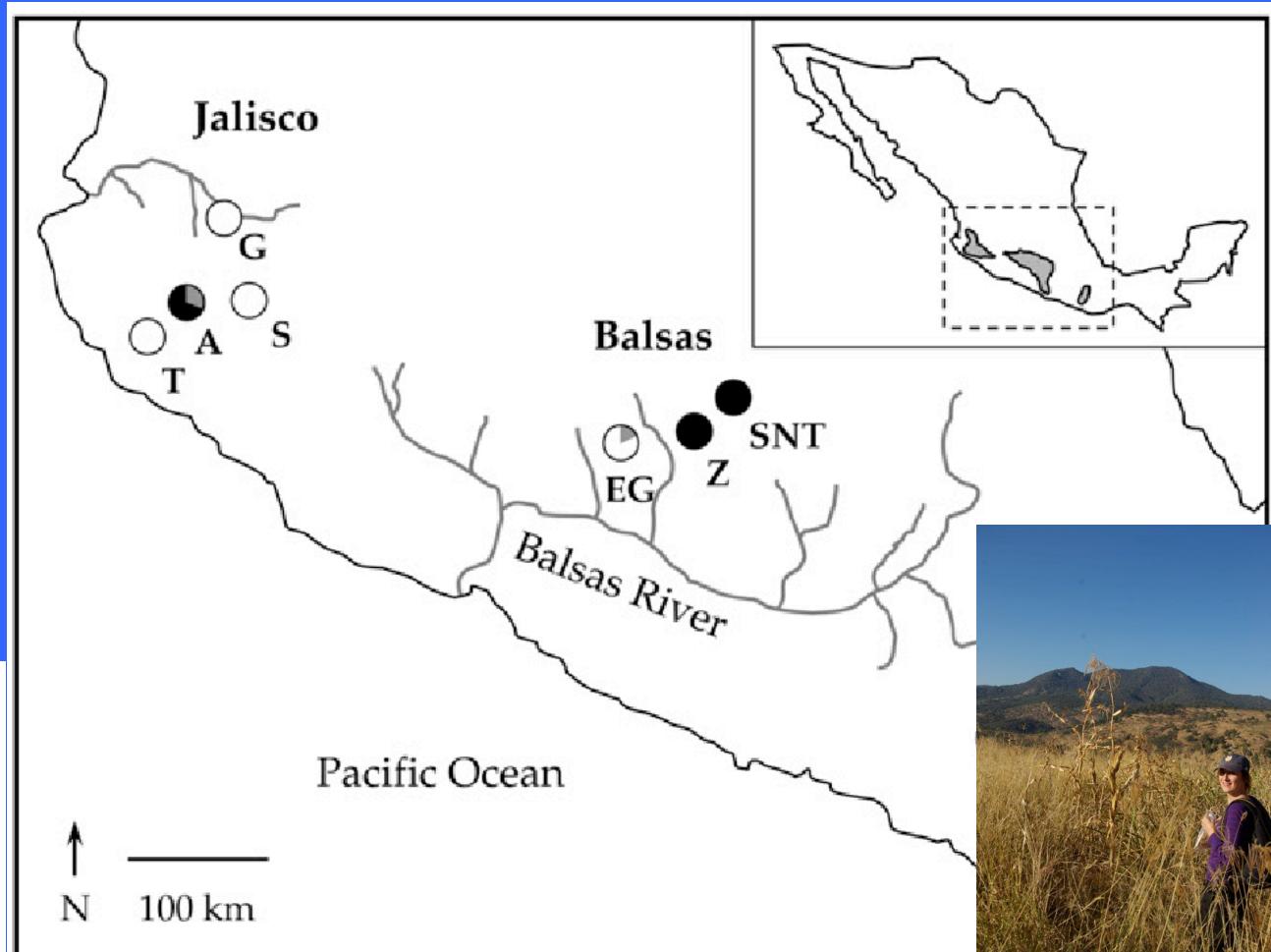


FIGURE 1.—Geographic distribution of the seven subpopulations of *Zea mays* ssp. *parviglumis* included in this study. The four western subpopulations are found in the state of Jalisco and the three eastern subpopulations are found in the Balsas River region of the states of Mexico and Michoacán. The pie diagrams show the proportion of each of the three chloroplast haplotypes found in each subpopulation. The inset map of Mexico shows the entire geographic distribution of the taxon.

π highest values in red

π lower values in blue

Nucleotide polymorphism, measured by θ and π , for subpopulation-specific and species-wide samples of five nuclear genes

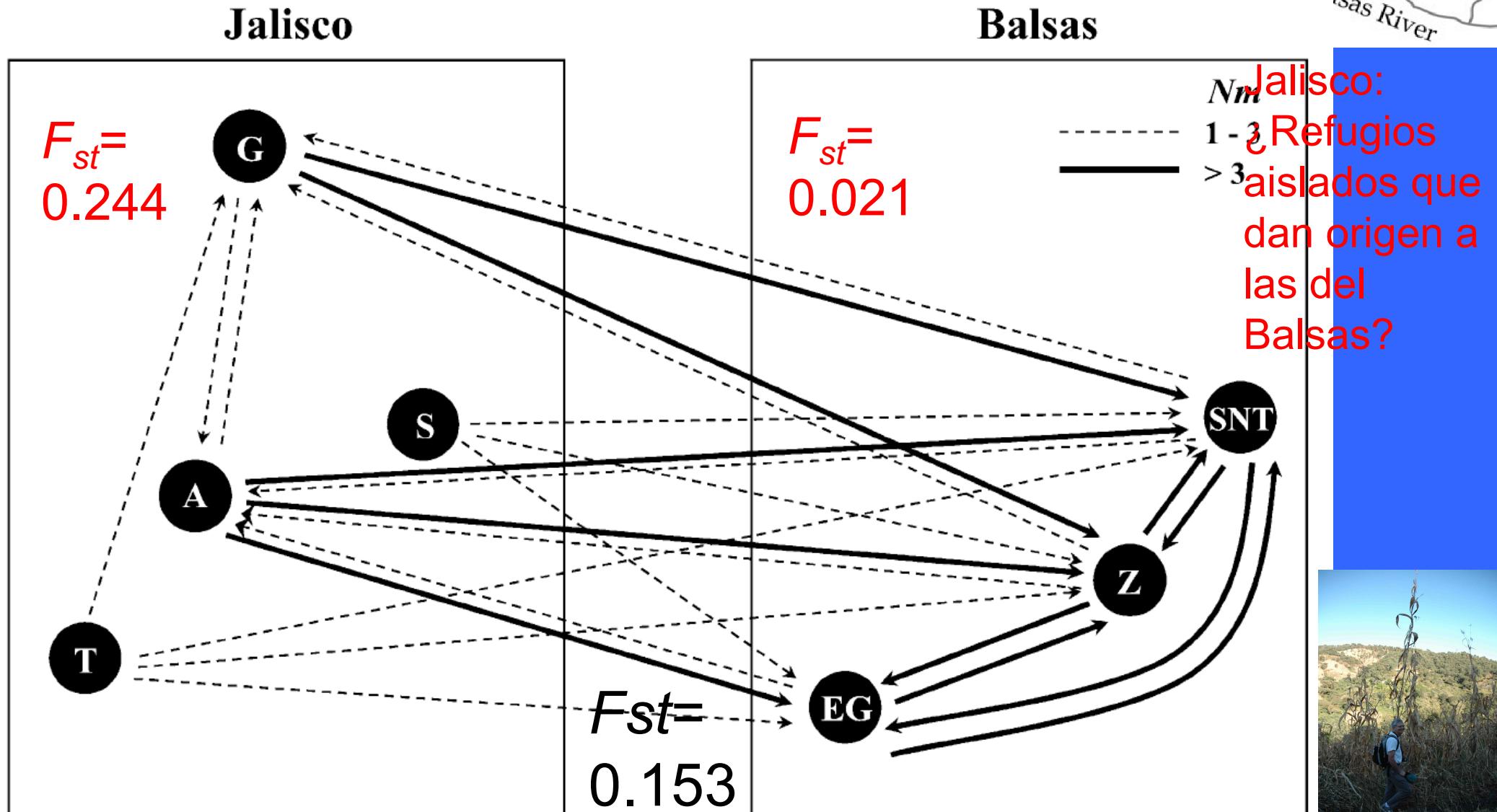
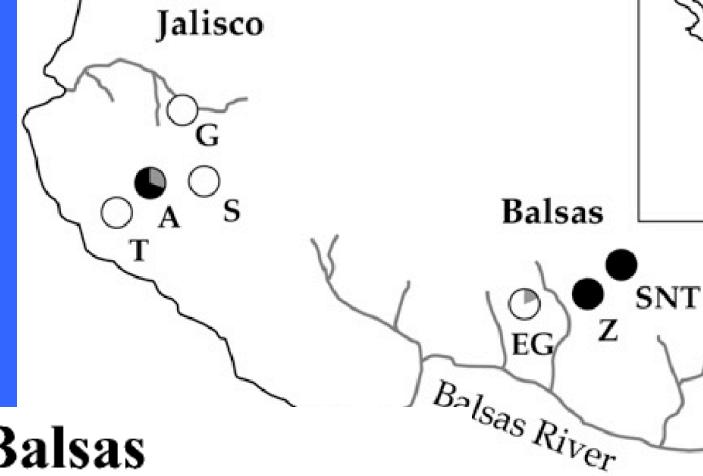
Population	<i>adh1</i>		<i>asg65</i>		<i>bnl7</i>		<i>glb1</i>		<i>waxy</i>	
	θ	π	θ	π	θ	π	θ	π	θ	π
Jalisco subpopulations										
A	0.0158	0.0157	0.0105	0.0111	0.0023*	0.0091	0.0227*	0.0185*	0.0094	0.0080*
G	0.0141	0.0157	0.0090	0.0087	0.0086	0.0092	0.0189*	0.0212*	0.0085*	0.0066*
S	0.0093*	0.0082*	0.0092	0.0101	0.0000*	0*	0.0157*	0.0189*	0.0048*	0.0028*
T	0.0078*	0.0100*	0.0055*	0.0054*	0.0039*	0.0032*	0.0105*	0.0112*	0.0044*	0.0025*
Balsas subpopulations										
EG	0.0211	0.0203	0.0098	0.0087	0.0074	0.0059	0.0252*	0.0218*	0.0079*	0.0076*
SNT	0.0205	0.0203	0.0094	0.0072	0.0050*	0.0035*	0.0302	0.0235*	0.0104	0.0091
Z	0.0194	0.0186	0.0113	0.0096	0.0121	0.0107	0.0260*	0.0216*	0.0121	0.0098
Species-wide	0.0185	0.0173	0.0147	0.0097	0.0128	0.0094	0.0412	0.0214*	0.0153	0.0091

Population-specific estimates shown in italics differed significantly from the resampled distribution using species-wide sampling. Asterisks denote values that differ significantly from the coalescent distributions ($P < 0.05$). For all significant differences, population-specific estimates were less than resampled and coalescent distributions.

less variation in S and T in Jalisco
more variation in Balsas in general



FIGURE 2.—Asymmetric migration rates between pairs of subpopulations within and between the two geographic regions (indicated by boxes) estimated using the Bayesian version of LAMARC. Arrows indicate the direction of migration inferred from the analysis. Migration rates less than one are not shown.



MaizeSNP50

BeadChip

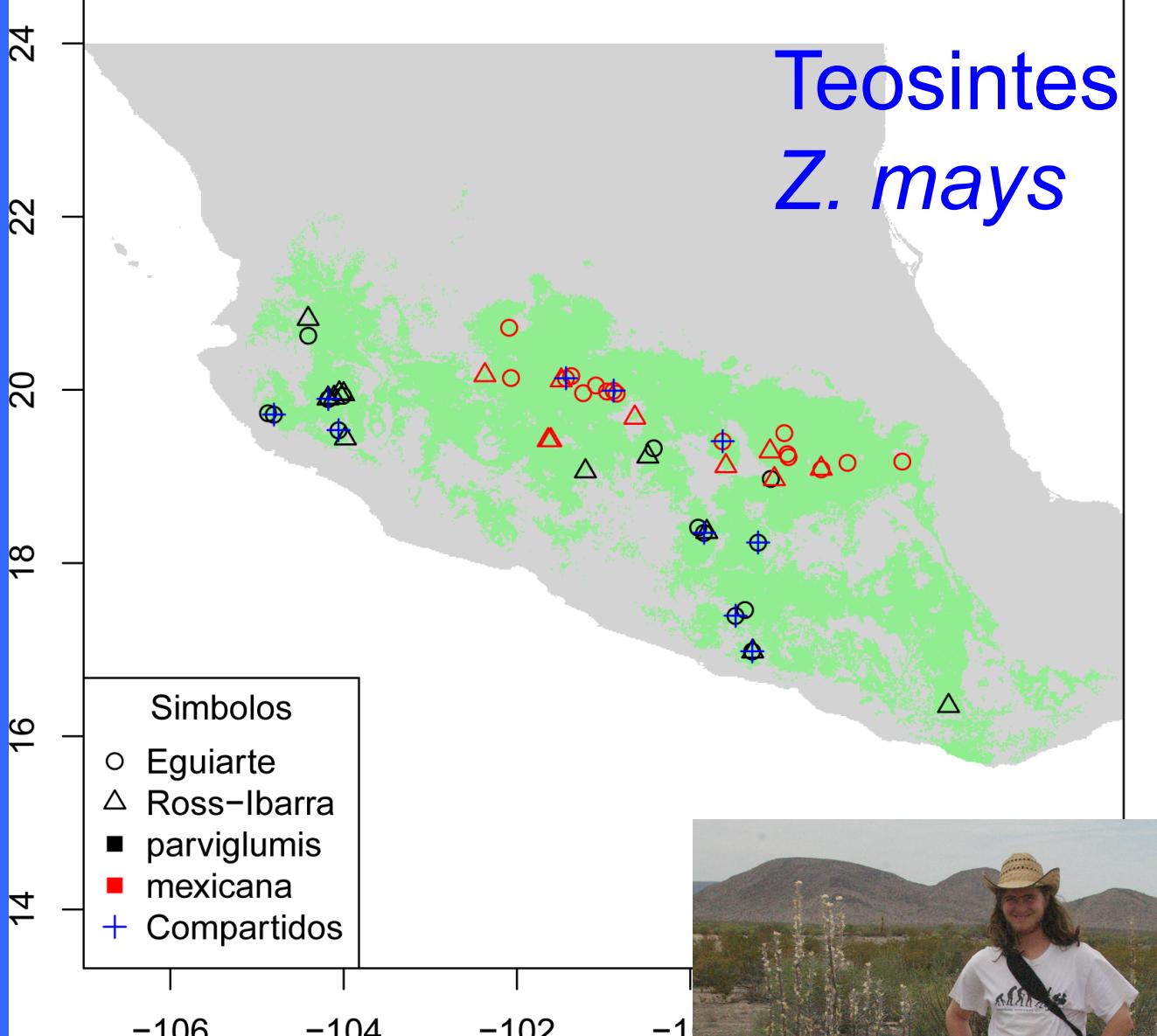
Illumina

36,000 SNPs

(7 poblaciones de
Z.m. parviglumis
y 2 de *Z.m.*
mexicana)

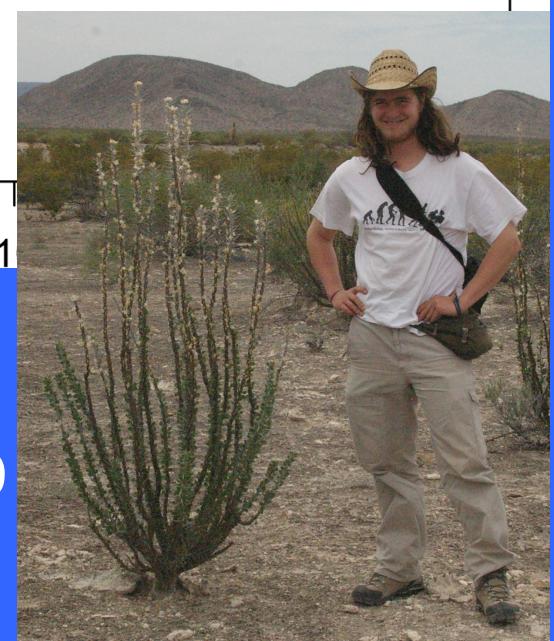
+ Pyhajarvi et al.

2013. *Genome
Biology and Evolution:*
10 *parviglumis*
11 *mexicana*

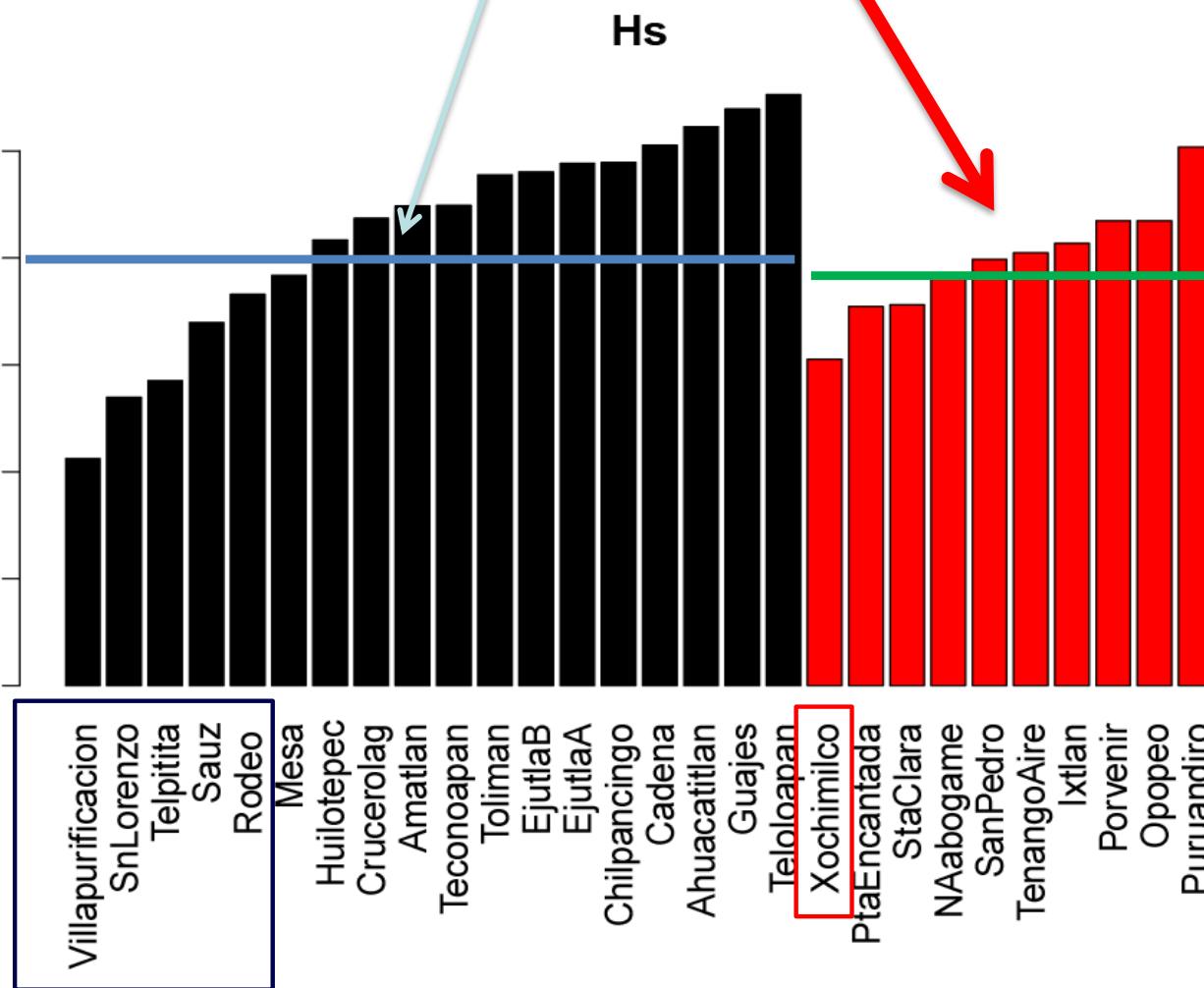
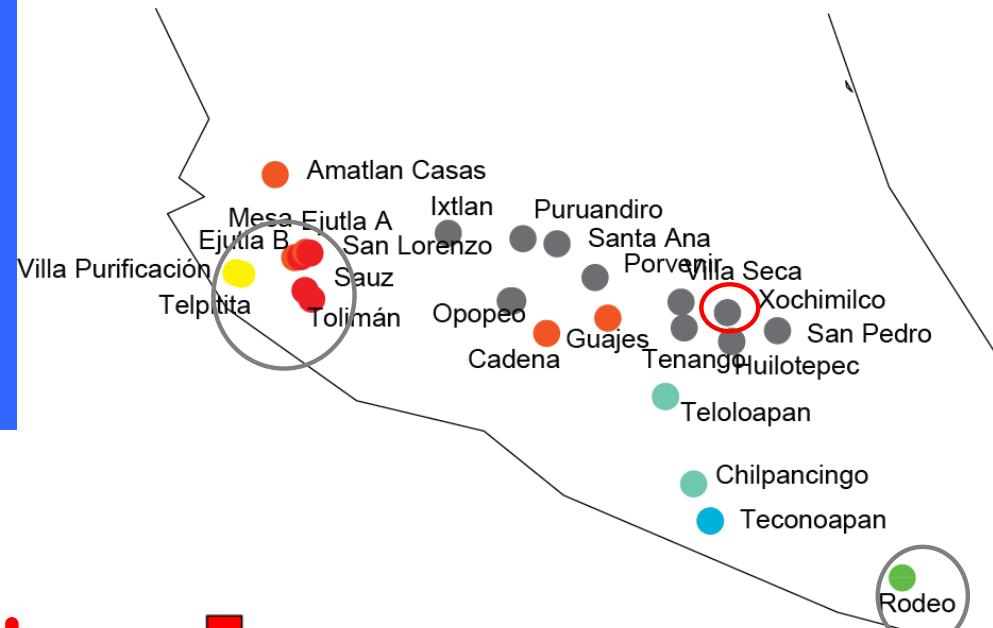


Teosintes
Z. mays

Jonás Aguirre
Tesis doctorado



10,000 SNPs aleatorios
 18 poblaciones *parviflumis*
 10 poblaciones *mexicana*



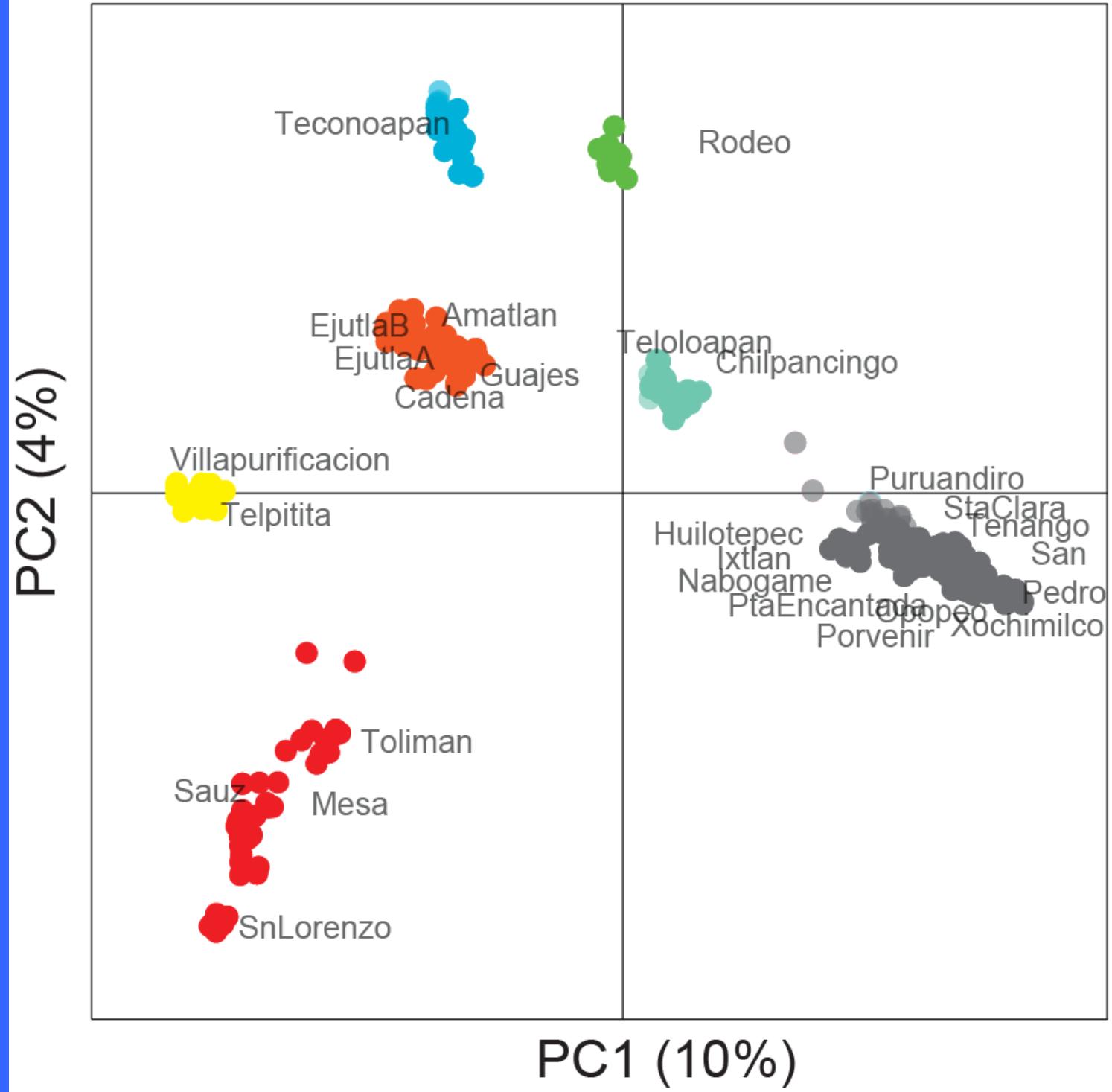
Un poco
 menos
 variación
 genética en
mexicana.
 rojo.

PCA of teosintes data set

Alta
diferenciación
genética :
10,000 SNPs
10 a 11
plantas por
población

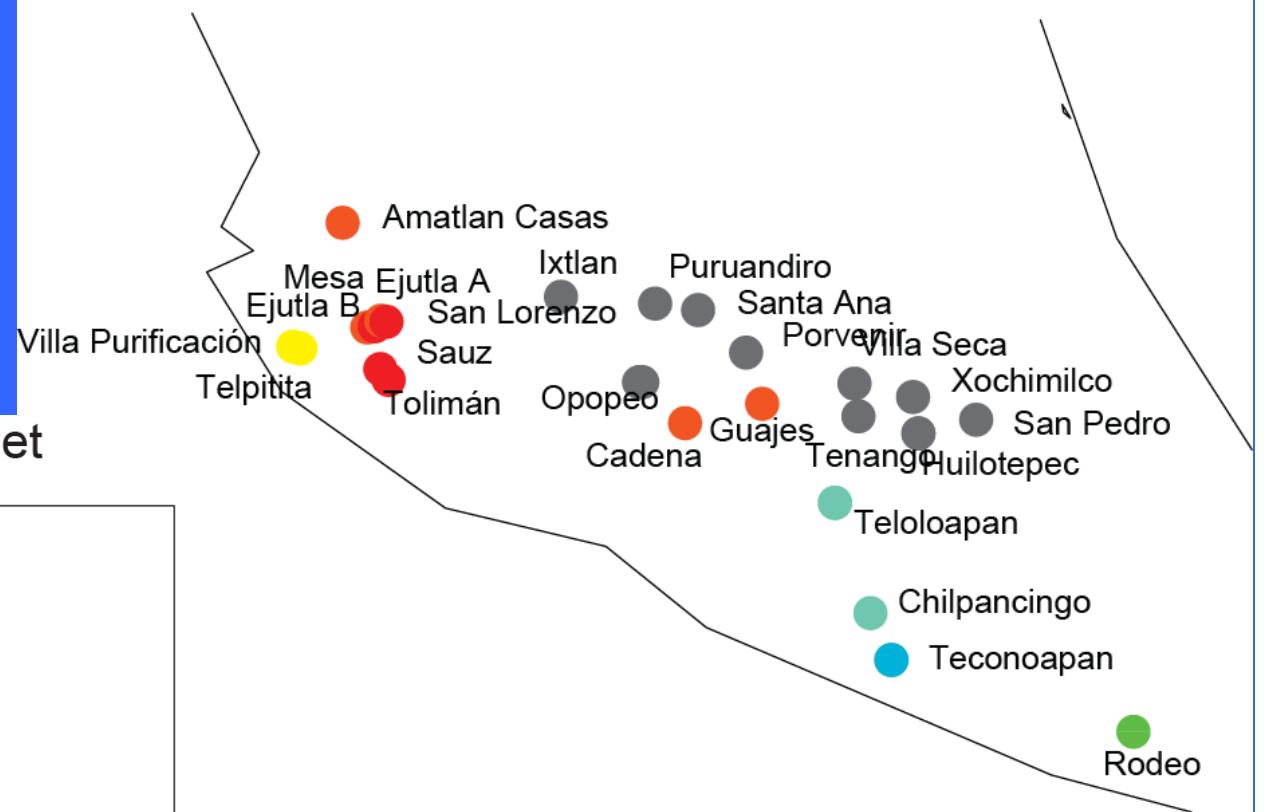
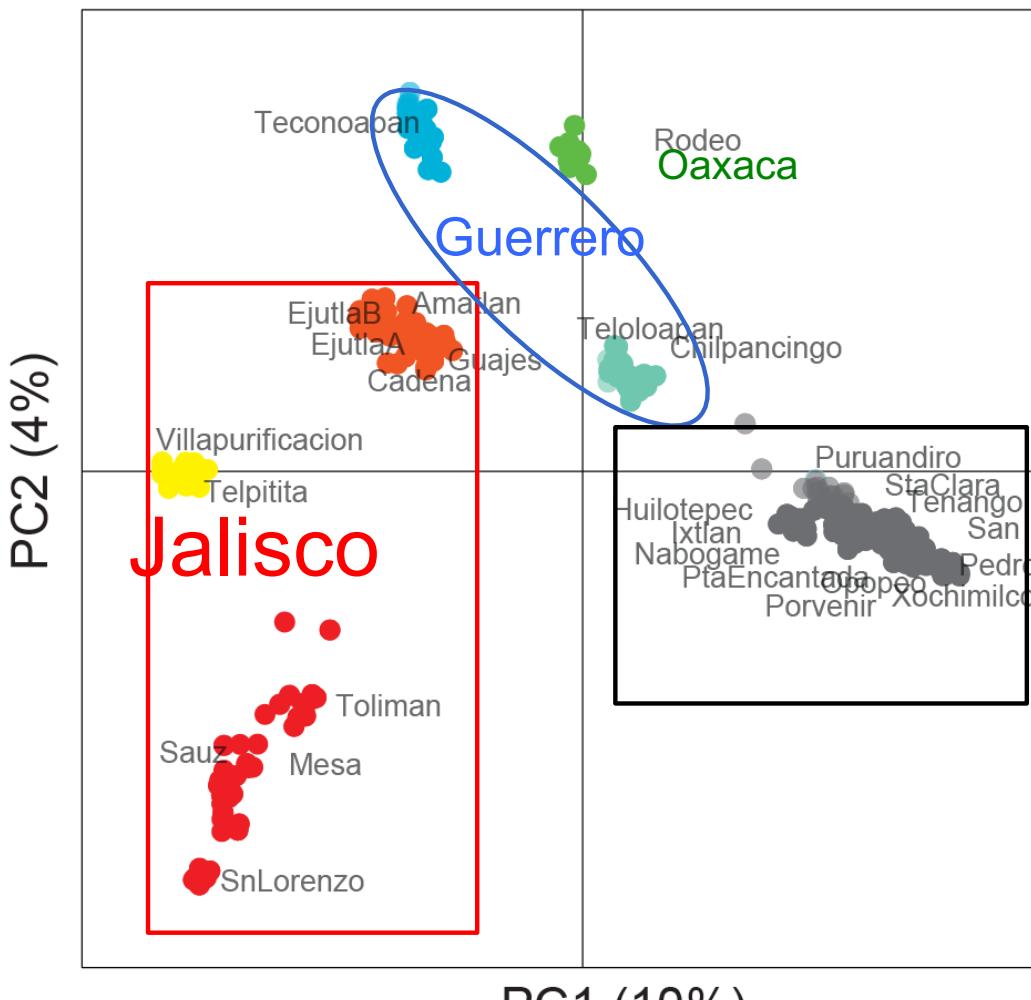
7 *parviflumis*
+ 21

Pyhajarvi et al.

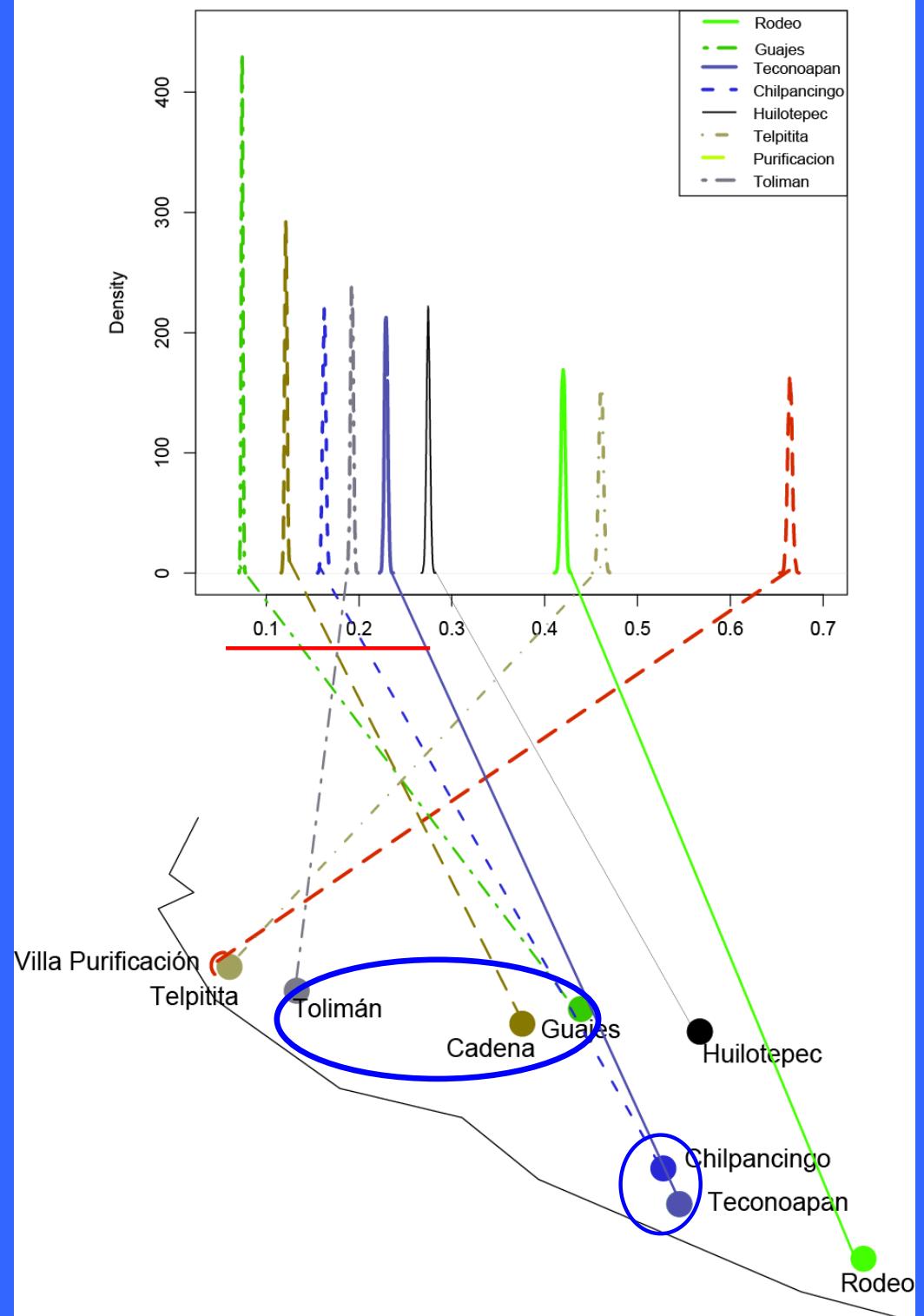


PCA por región 10,000 SNPs 28 poblaciones

PCA of teosintes data set



FST Posterior Distribution



Mayor
diferenciación en
los extremos,
menor en las
centrales.
Solo se muestra pobl.
selectas de
parviglumis
 F_{ST} medida de
diferenciación, 0 a 1
distribución posterior
(Beta) Bayescan

Aislamiento
por
distancia
parviflumis:
15 poblaciones

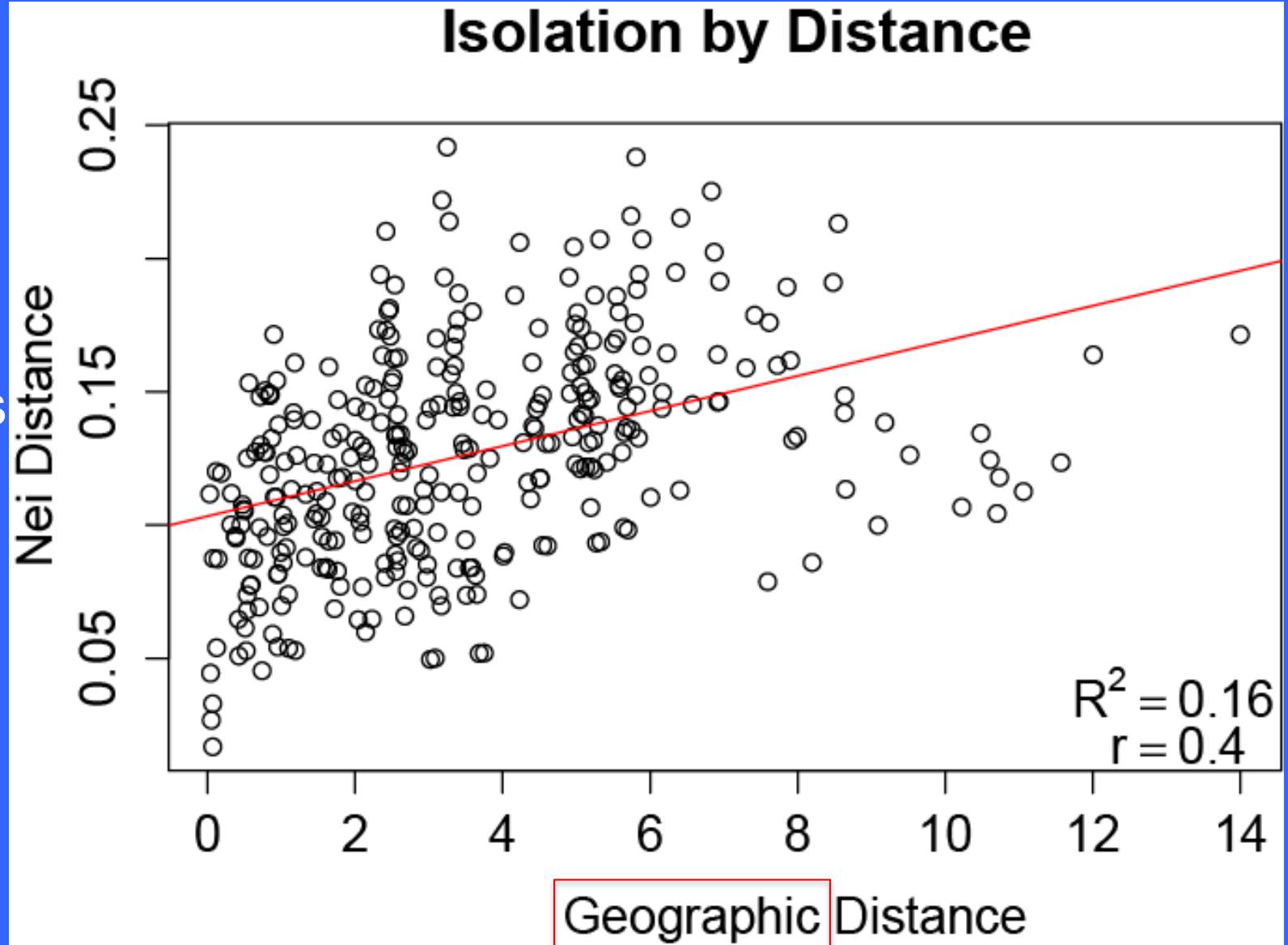
mexicana: 13
poblaciones

10,000

SNPs

vegan library R
 $p < 0.00001$

Prueba de
Mantel



Las poblaciones cercanas
geográficamente se parecen
genéticamente: flujo génico, dispersión
limitada, historia.

Aislamiento ambiental

Distancias Euclideanas entre pares de sitios, basadas en los 2 primeros componentes principales del clima

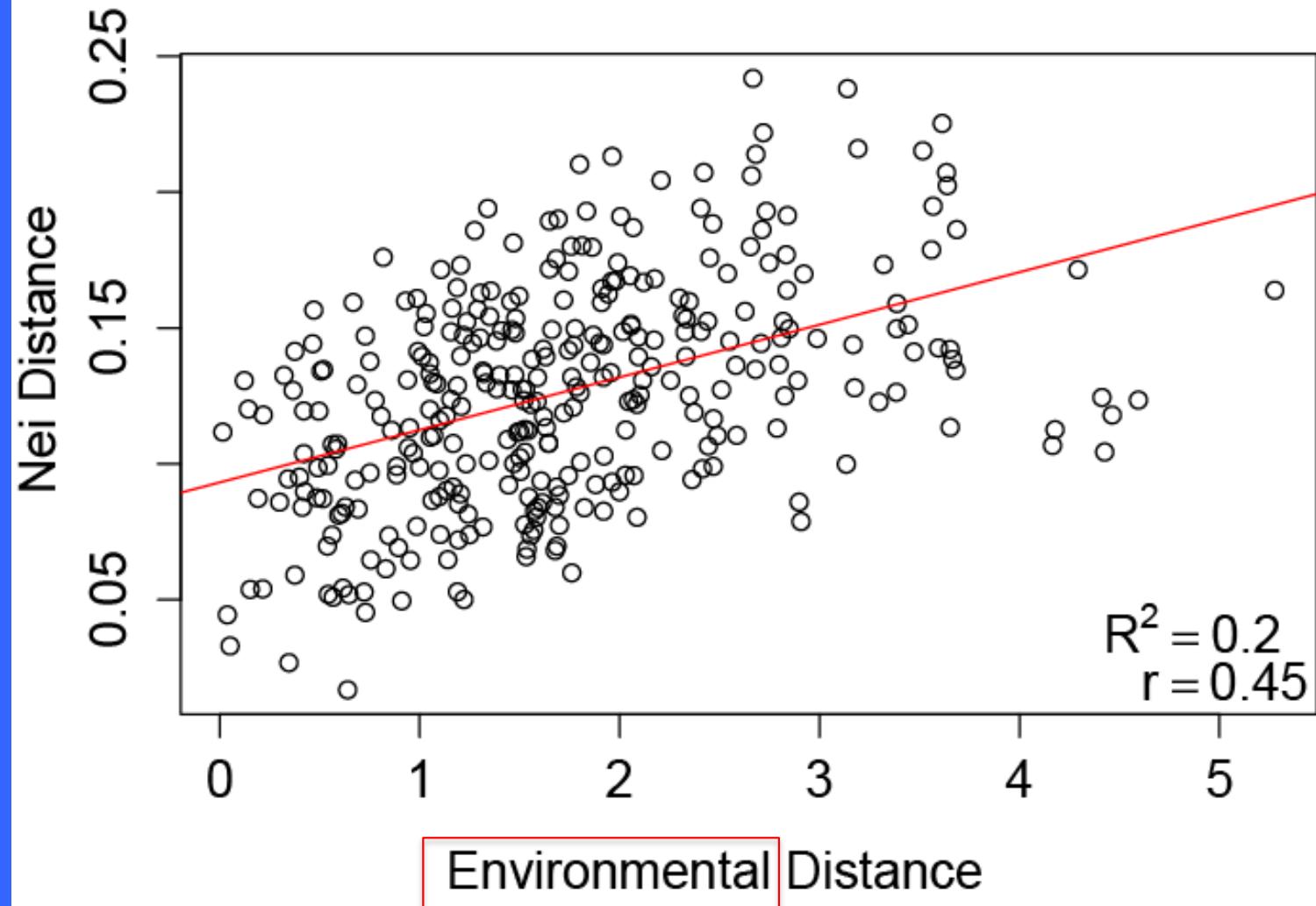
parviflumis: 15 poblaciones
mexicana: 13 poblaciones

Mantel

$p < 0.00001$

Más fuerte que el aislamiento por distancia

Isolation by Environment



Las poblaciones climáticamente parecidas se parecen genéticamente: adaptación local al clima!!

Correlaciones distancia genética vs. distancias geográfica y climática: más importante ambiente

Modelo	p	R ²
Geografía	<0.00001	0.18
Ambiente (clima)	<0.00001	0.2
Geografía+ Ambiente	<0.00001	0.22

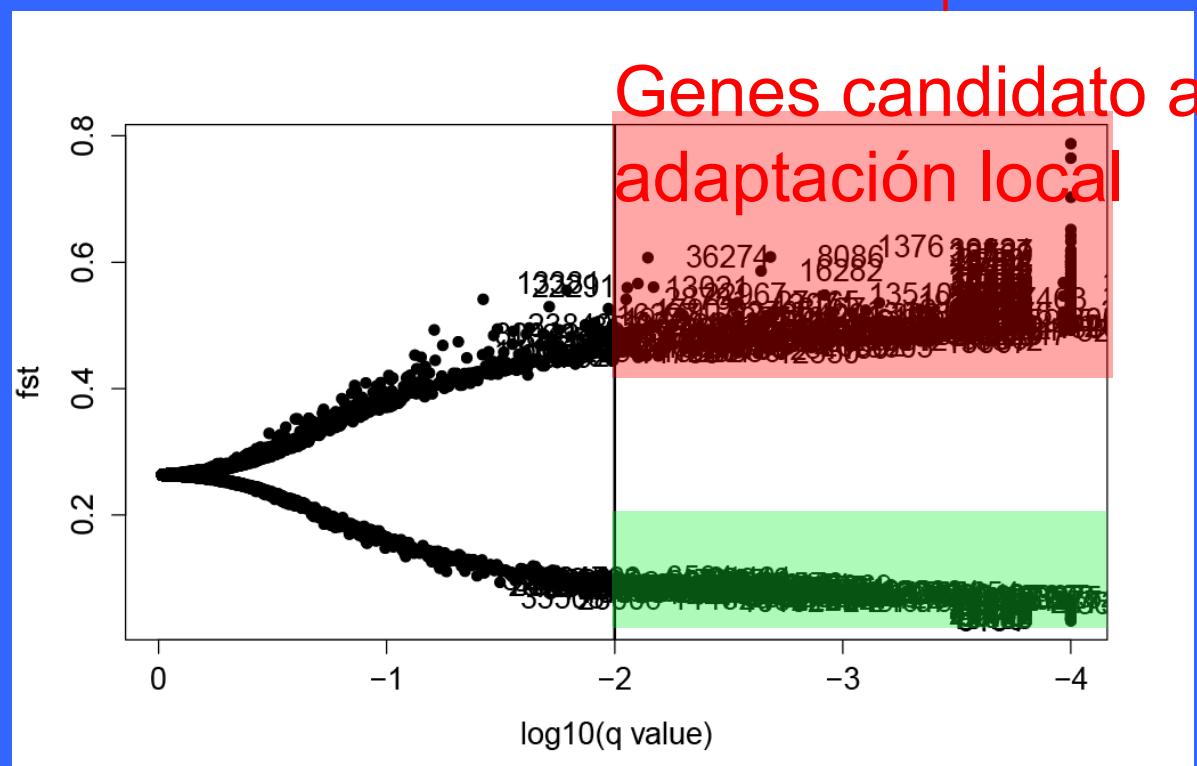
Pruebas de Mantel parciales

Modelo	p
Genética y Geografía, cte. Ambiente	0.06
Genética y Ambiente, cte. Geografía	0.0027
Ambiente y Geografía cte. Génetica	<0.00001



Zea mays spp. *parviglumis*: Bayescan (Foll y Gaggiotti 2008), 36,000 SNPs, 18 poblaciones, 163 genes candidato purificadora, 248 adaptación local.

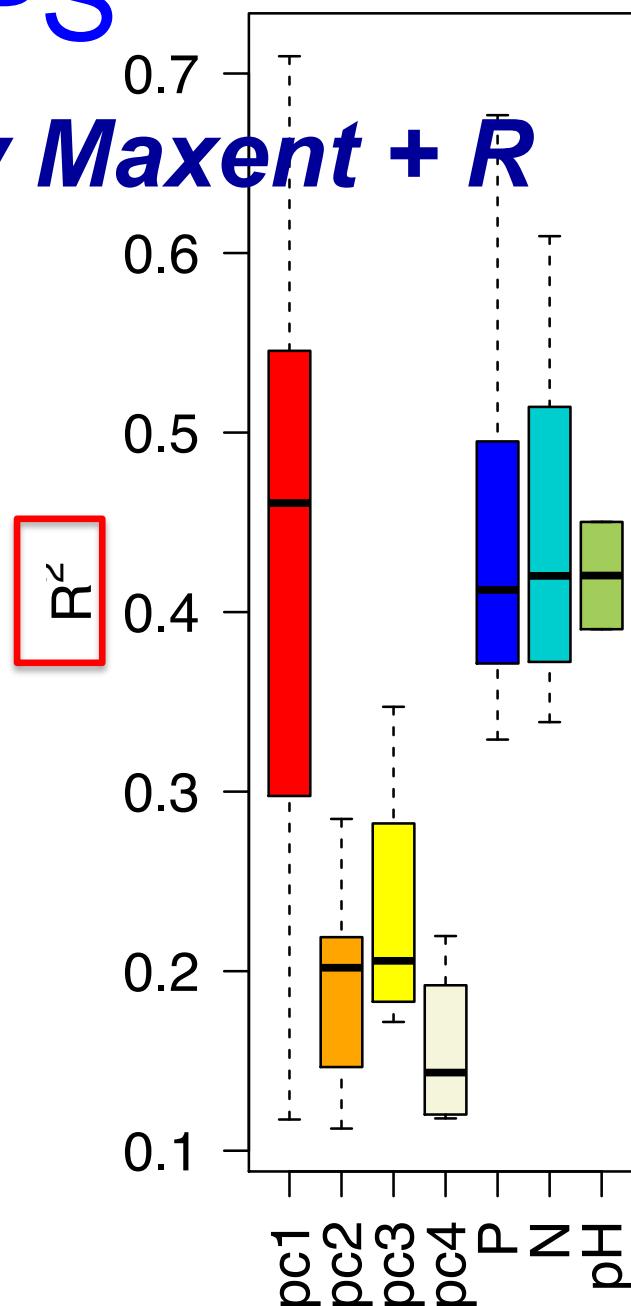
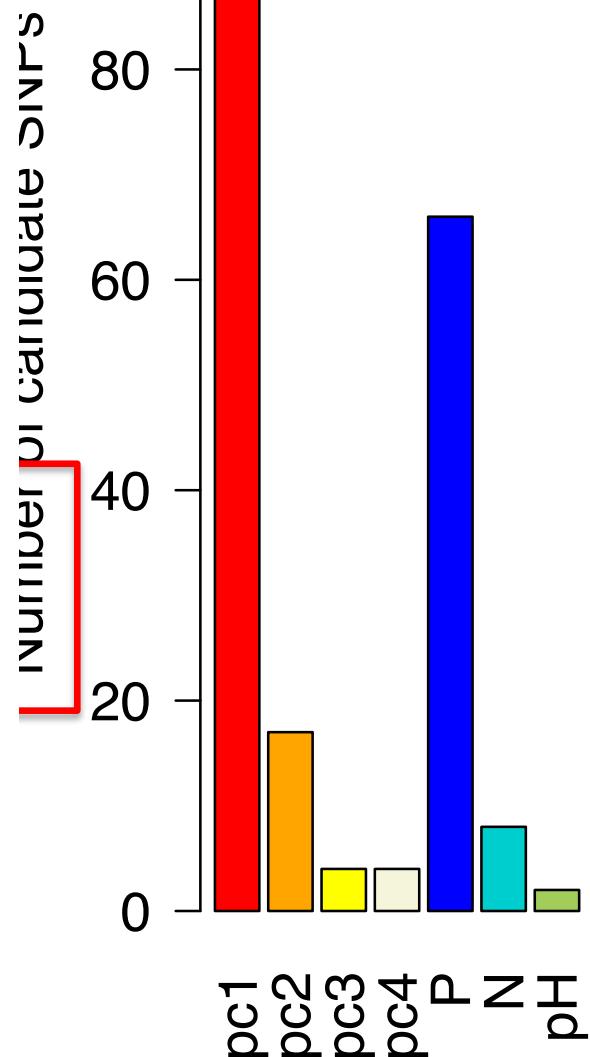
Tipo de selección	Número de SNPs
Direccional local	248
Balanceadora, pur.	163



Factores ambientales, clima y suelo

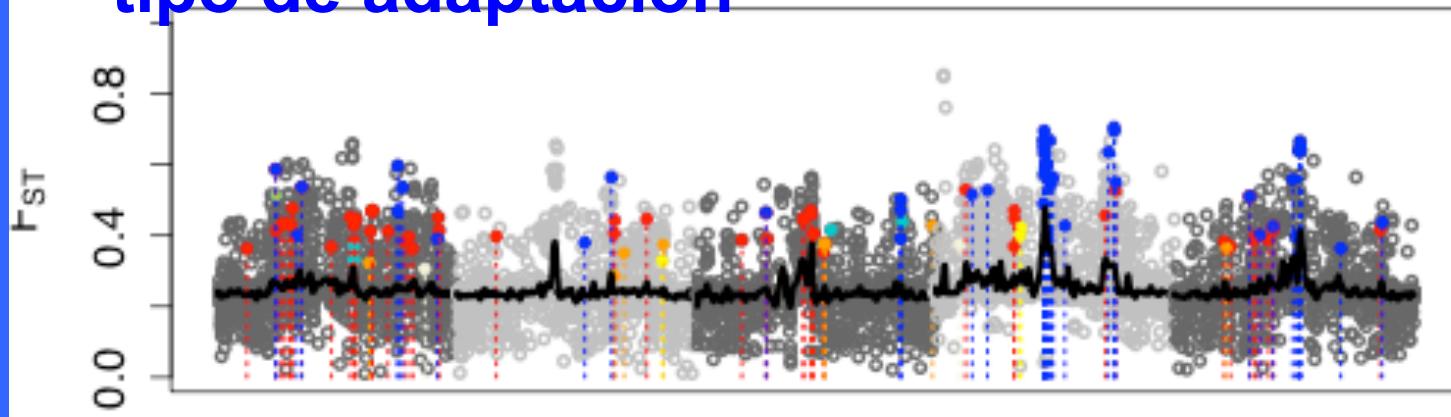
36 mil SNPs

Bayescenv y Maxent + R

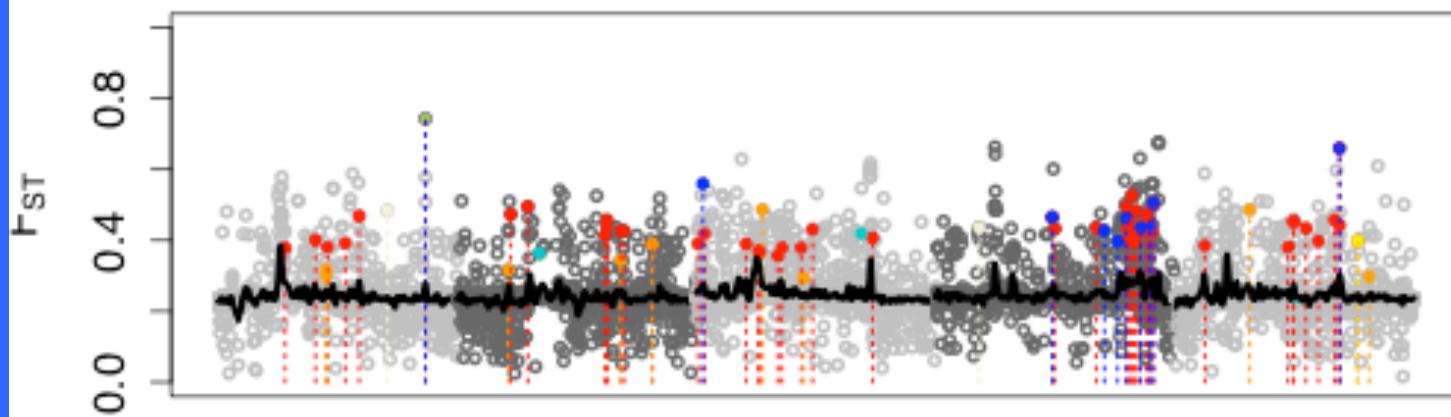


Temperatura
precipitación
y fósforo son
los factores
principales de
selección y
adaptación
local.

Posiciones SNPs con selección, por tipo de adaptación



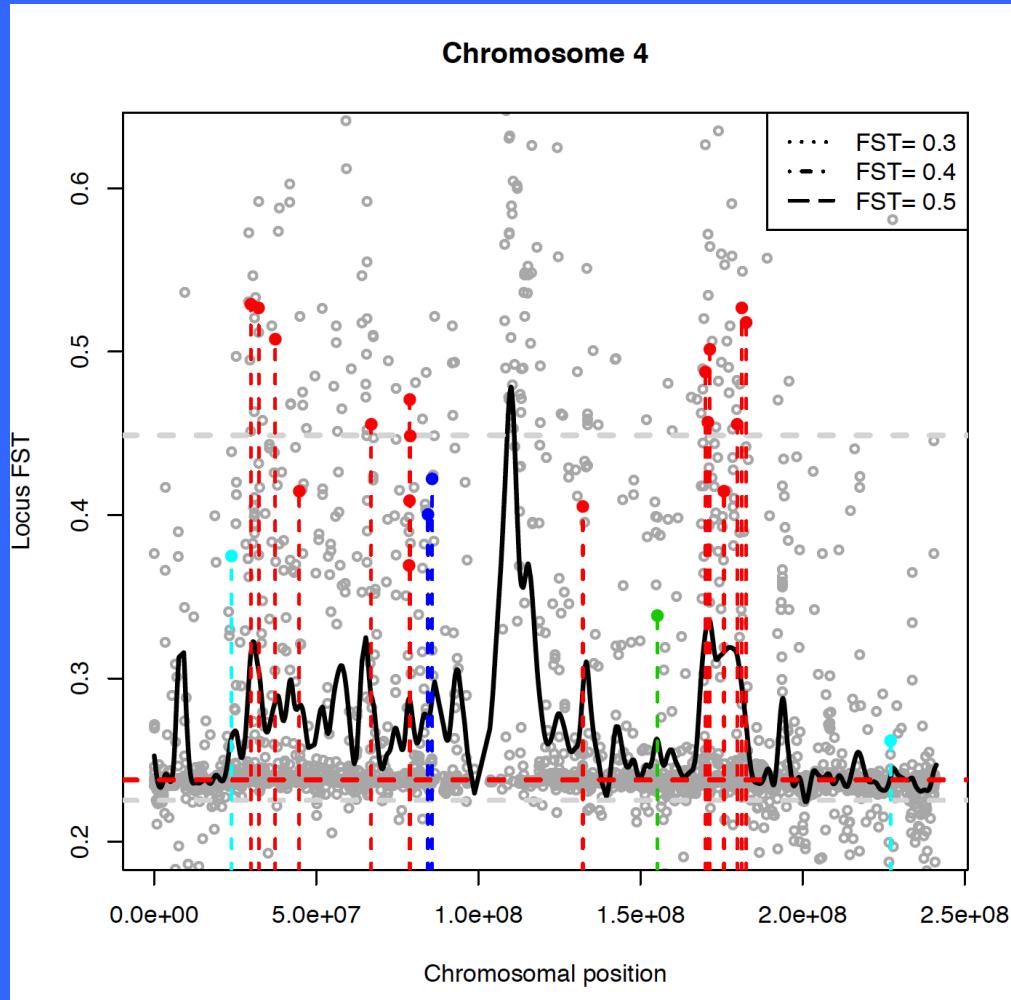
Chr 1 Chr 2 Chr 3 Chr 4 Chr 5



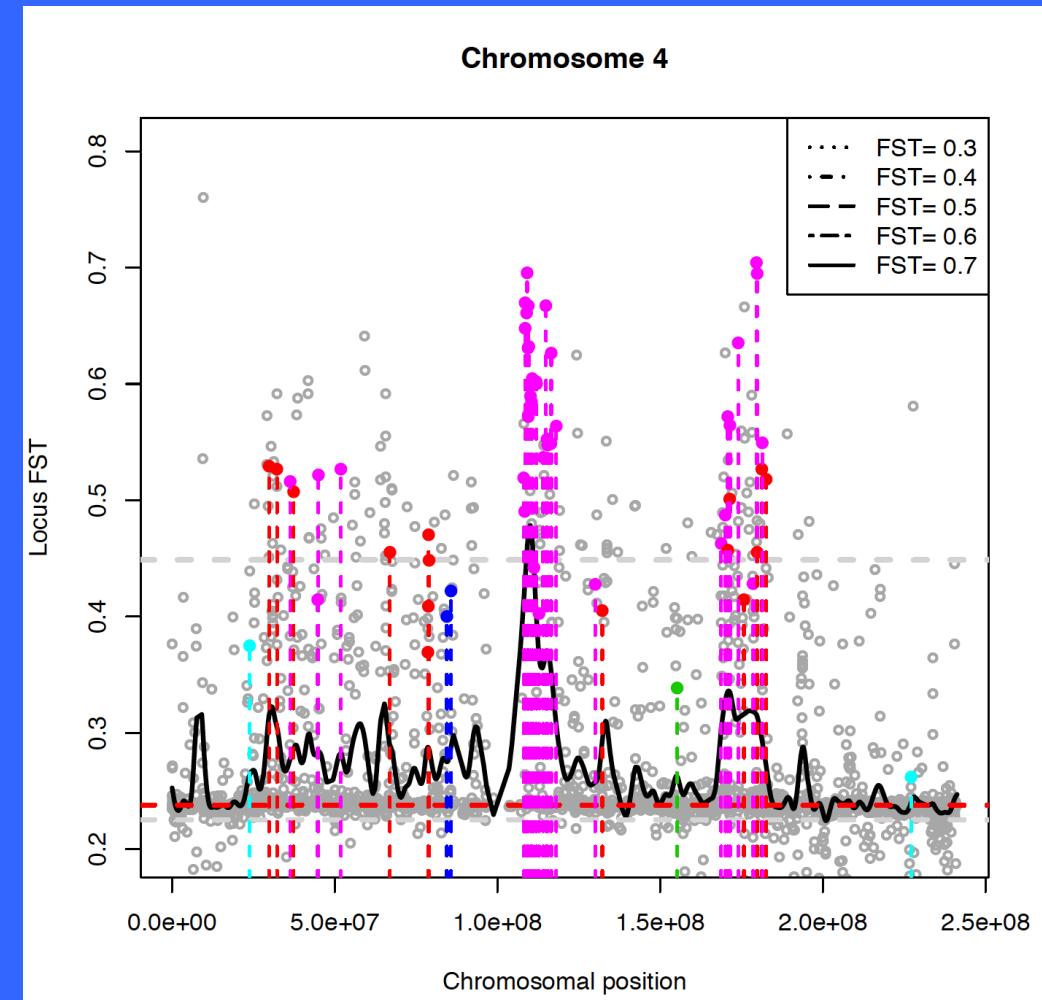
Chr 6 Chr 7 Chr 8 Chr 9 Chr 10

Bayescenv: una herramienta poderosa

Sólo clima



Clima + Suelo



Usando genomas completos

Signatures of local adaptation in lowland and highland teosintes from whole-genome sequencing of pooled samples

M.-A. FUSTIER,* J.-T. BRANDENBURG,* S. BOITARD,† J. LAPEYRONNIE,† L. E. EGUIARTE,‡
Y. VIGOUROUX,§ D. MANICACCI* and M. I. TENAILLON*

Table 1. Sequencing statistics for 6 populations

Population	# reads ^a	Coverage ^b (%)	Depth ^c	mapped reads (%)
H1	377,317,315	58.83	22.1	35.1
L1	312,979,394	59.33	22.9	36.2
H2	317,647,367	55.73	21.5	43.3
L2	437,425,261	53.88	26.6	35.6
IM1	196,271,307	54.14	17.6	35.8
IP1	194,443,385	60.89	16.6	35.0

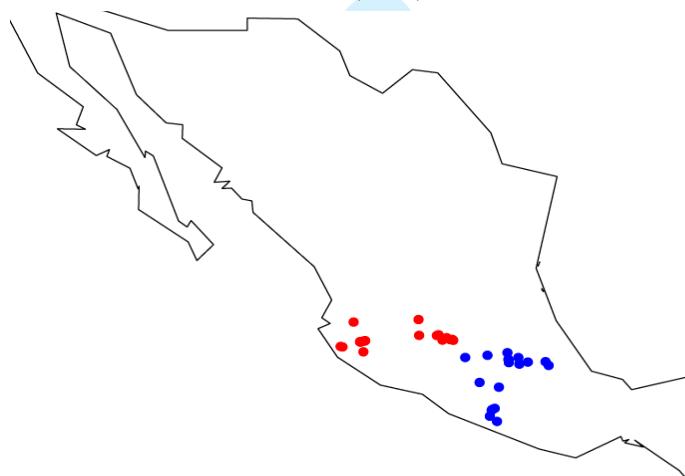


Table 2. Number of SNPs detected for each method and in combination with BAYENV2

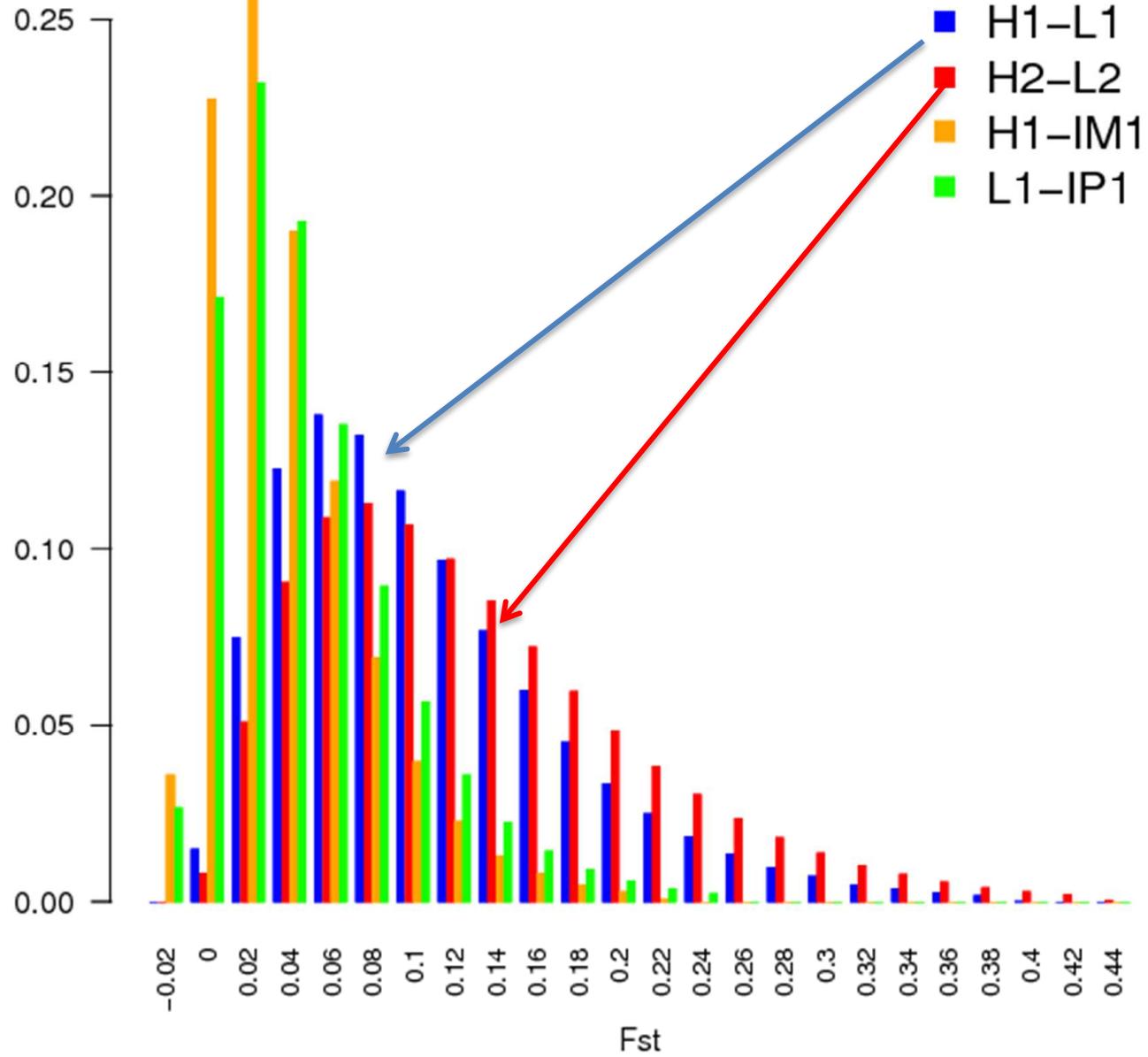
Method	# SNPs detected ^a	# SNPs detected with BAYENV2 ^b
Bayescan	1028 (596)	17 (1.65%)
<i>Fst</i>	3662 (3228)	55 (2.07%)
PoolHMM	50 windows, 161371 (80302)	229 (0.14%)
None	-	58 (0.05%)

^a Number of SNPs detected using a single method, in bracket the number of genic SNPs

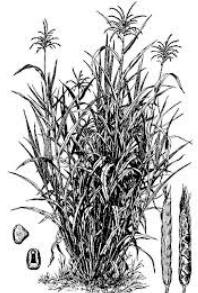
^b In bracket the percent of SNPs significant with BAYENV2 among those detected by each method



Comparación de 6
poblaciones
altas y bajas en los
dos gradientes
ambientales



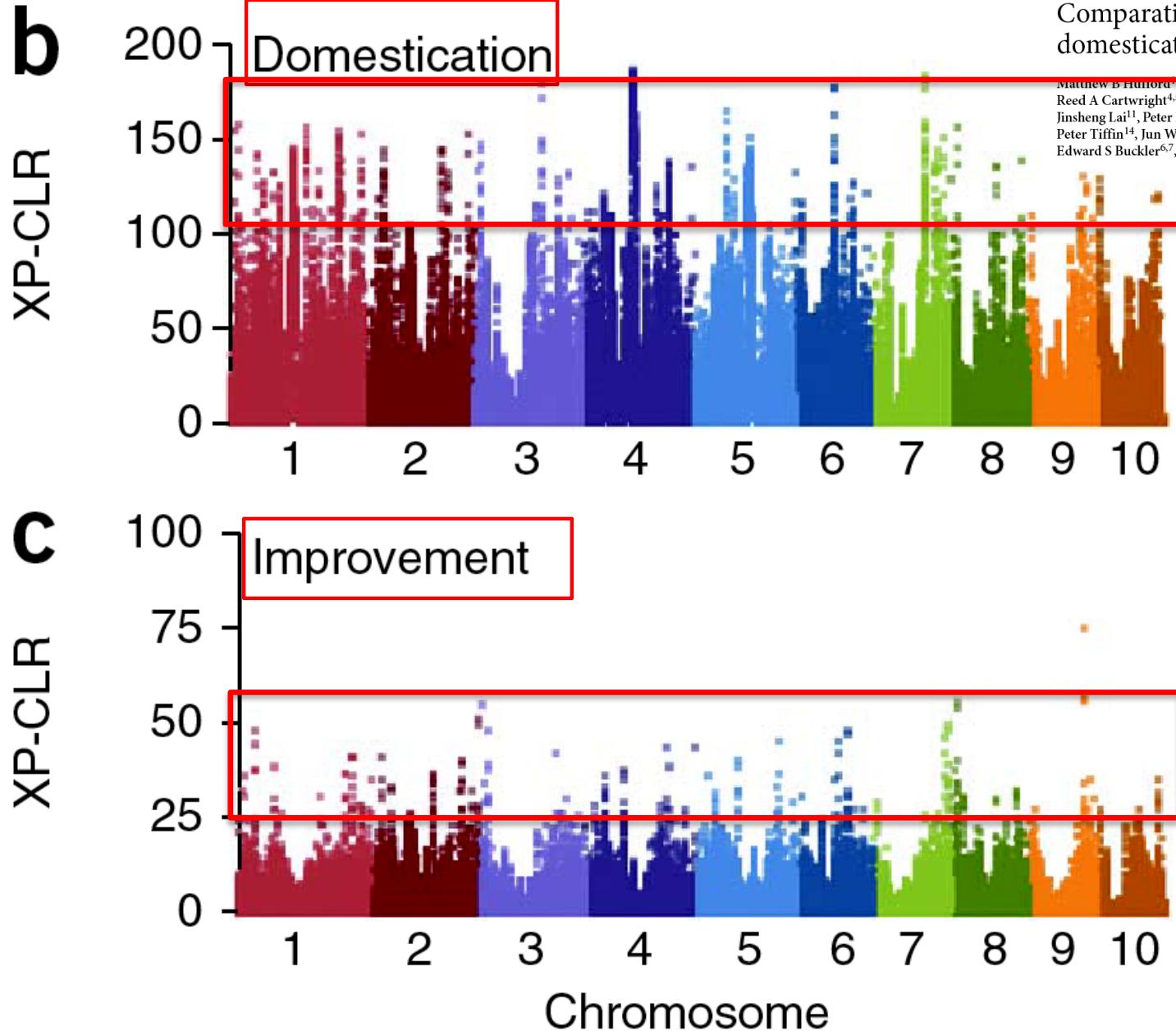
Distribución de la F_{st} entre pares de poblaciones analizadas con 8,479,581 SNPs.
Más diferentes los extremos de las distribuciones.



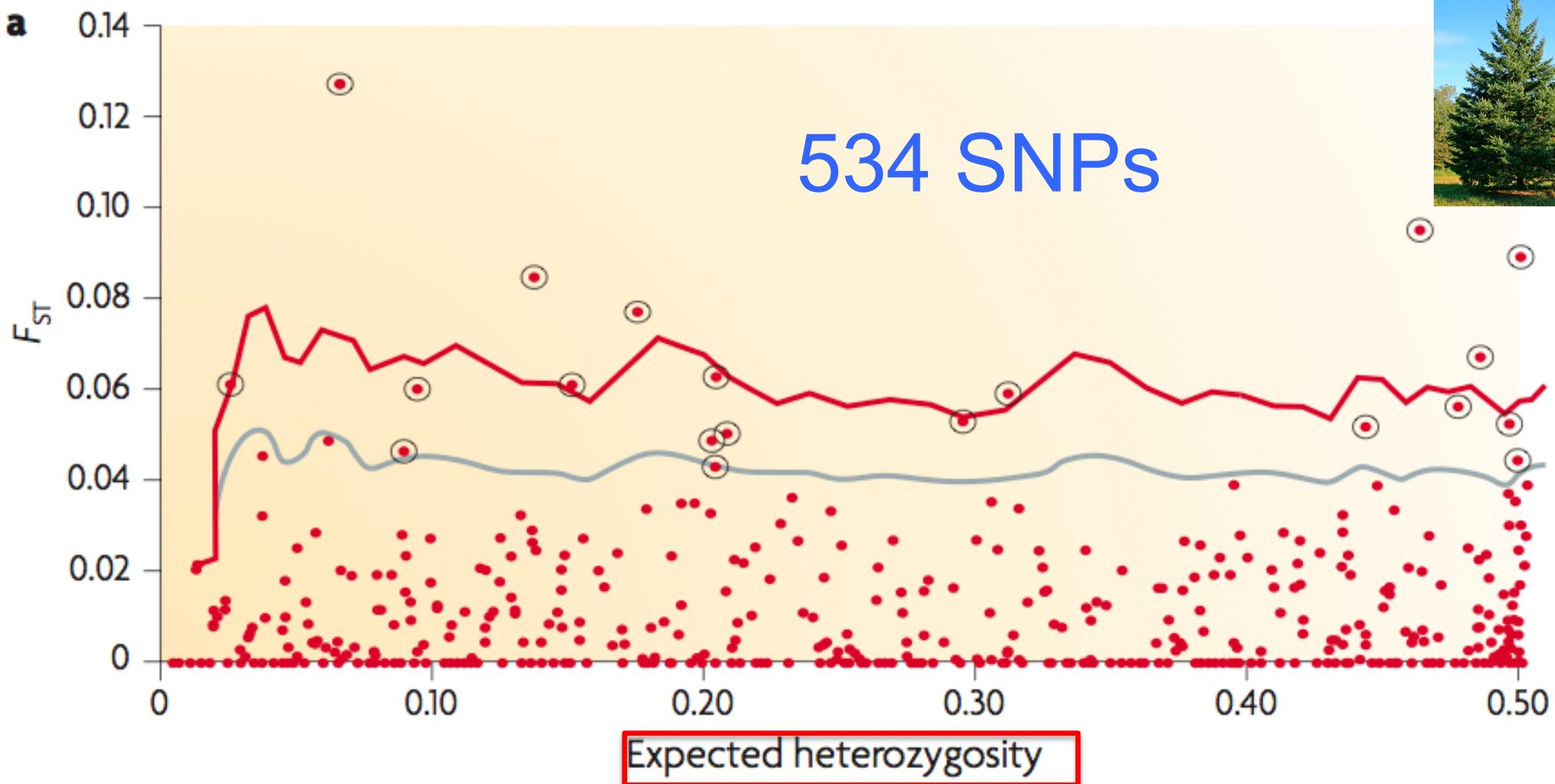
Matthew B Flinck^{1,7}, Xun Xu^{2,17}, Joost van Heerwaarden^{1,17}, Tanja Pyhäjärvi^{1,17}, Jer-Ming Chia³,
Reed A Cartwright^{4,5}, Robert J Elshire⁶, Jeffrey C Glaubitz⁶, Kate E Guill^{7,8}, Shawn M Kaepler^{9,10},
Jinsheng Lai¹¹, Peter I Morrell¹², Laura M Shannon¹³, Chi Song², Nathan M Springer¹⁴, Ruth A Swanson-Wagner¹⁴,
Peter Tiffin¹⁴, Jun Wang², Gengyun Zhang², John Doebley¹³, Michael D McMullen^{7,8}, Doreen Ware^{3,7},
Edward S Buckler^{6,7}, Huang Yang² & Jeffrev Ross-Ibarra^{1,15,16}

nature
genetics

Genomas
MAIZ
21
millones
de snps

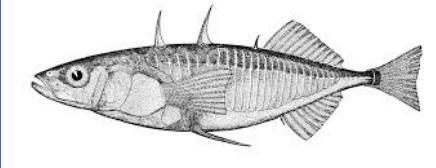
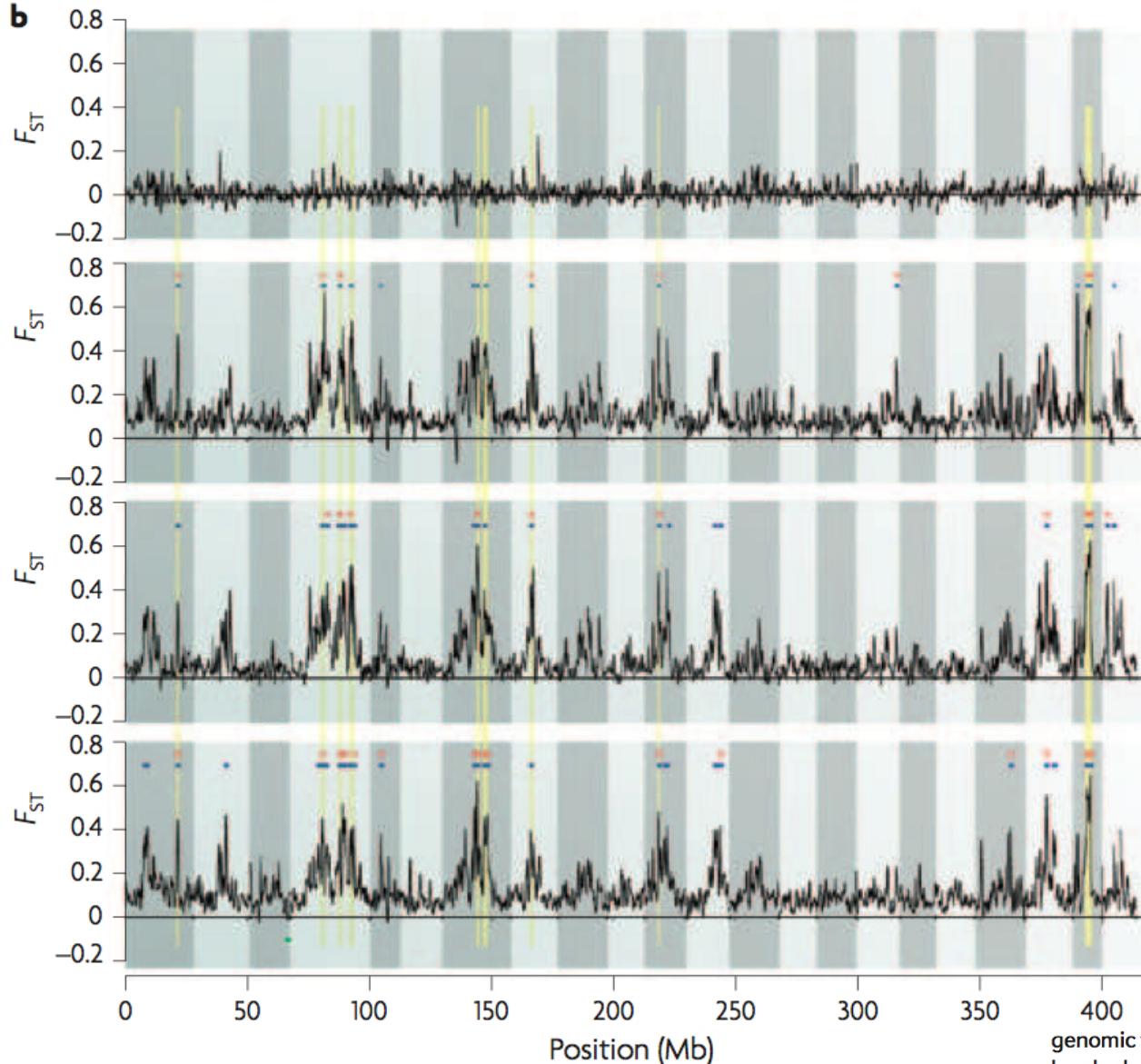


(b,c) Genome-wide likelihood (XP-CLR) values for selection during domestication (b) and improvement (c), with chromosome number indicate along the x axis. (d,e) Distributions of feature size (d) and gene counts

a

Genome scans for selection can focus on either candidate loci or anonymous loci.

Namroud et al.⁶² sampled white spruce (*Picea glauca*) from 6 populations in Quebec and genotyped 534 SNPs located on 345 candidate genes. Part a of the figure shows their F_{ST} outlier analysis of these data, based on the relationship between F_{ST} and expected heterozygosity⁵⁹; the grey and red lines represent the 95% and 99% confidence levels, respectively. Against a background of little population differentiation ($F_{ST} = 0.006$), this analysis identified 20 SNPs (circled dots) in 19 genes above the 95% confidence level. New



2 marinas

marinas vs.
c/u agua dulce

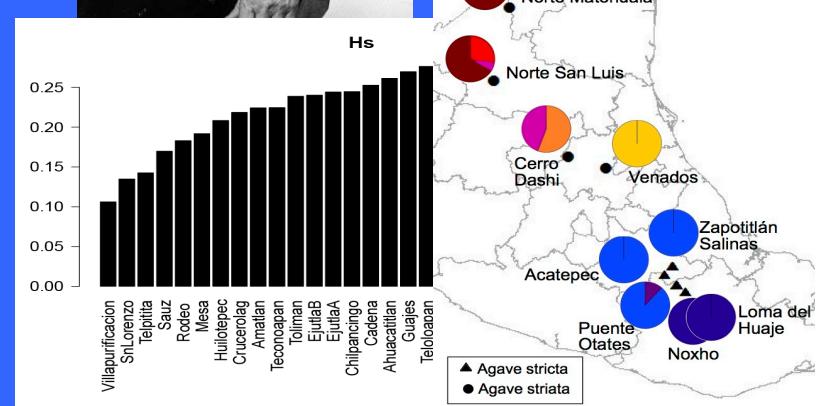
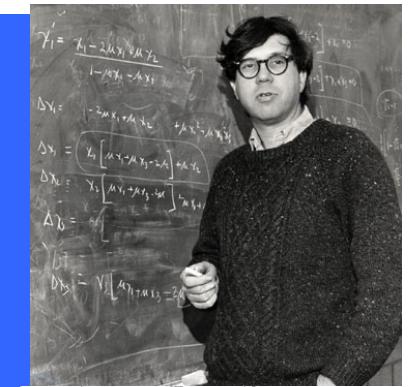
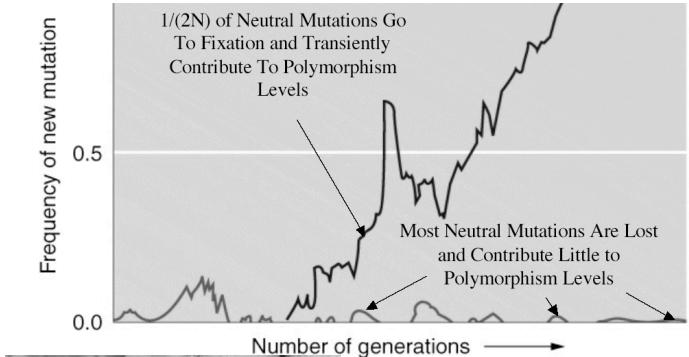
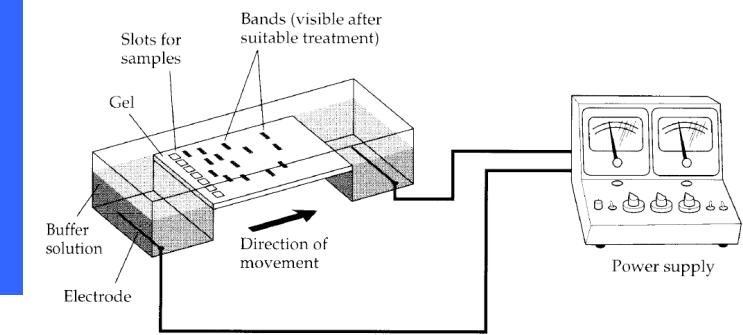
azul $p < 10^{-5}$
naranja $p < 10^{-7}$

45 mil SNP *Gasterosteus aculeatus* threespine stickleback

Hohenlohe PA, et al.. (2010) PLoS Genet 6(2): e1000862. doi:10.1371/journal.pgen.1000862

genomic tools also allow anonymous markers to be assayed across the genome to identify local adaptation; for example, Hohenlohe et al.⁶⁶ sampled 100 threespine stickleback individuals across 5 populations in Alaska. They used sequencing of restriction-site-associated DNA (RAD) tags¹² to simultaneously identify and genotype over 45,000 SNPs across the genome. This density of markers allows population genetic statistics, such as F_{ST} , to be visualized as continuous distributions along chromosomes. In part b of the figure, the top panel shows F_{ST} between the two marine populations. The next three panels show differentiation between each of the three freshwater populations and the two marine populations. Coloured bars above each graph show regions of significantly elevated F_{ST} as indicated by bootstrap resampling (blue, $p \leq 10^{-5}$; red, $p \leq 10^{-7}$). Vertical grey shading indicates the chromosomes, and yellow shading indicates the nine most significant and consistent peaks of freshwater-versus-marine differentiation. Common patterns of population differentiation (yellow shading shared among the three populations) indicate genomic regions that have responded to divergent selection in parallel across

¡Gracias!

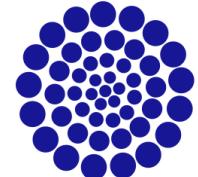


Agradecimientos: Apoyos:

CONABIO – KE004 y PE001/18 Rafael Lira y Luis Eguiarte
Conacyt CB2011/167826
Conacyt Problemas nacionales, Proyecto Milpa Daniel Piñero,
Proyecto Fronteras número 177, LEE, SEP-CONACYT-ANUIES-
EC OS Francia M12-A03, CONACYT-ANUIES 207571
Papiit, UNAM, IG200215, IN202712, IN224309

Personas: Jonás Aguirre, Alejandra Moreno Letelier, Erika Aguirre-Planter, Laura Espinosa Asuar, Daniel Piñero, Rafael Lira Saade, Salvador Montes, Alberto Bürquez, Carlos Martínez del Rio, Mike Travisano, Peter Tiffin, Brandon S. Gaut, Maud Tenaillon, Roberto Trejo, Yocelyn Gutierrez Guerrero, Nuri Flores, Felipe y María Eguiarte, Alejandra Vázquez-Lobo, Gabriela Castellanos, Leslie Paredes, Josué Barrera, Guillermo Sánchez, Helena Hernández, Silvia Barrientos, Paulina Hernández, Karen Ruiz, decenas de amigos y estudiantes.

¡GRACIAS A TODOS!!!



CONACYT

Consejo Nacional de Ciencia y Tecnología



Dirección General de Asuntos
del Personal Académico



INSTITUTO DE ECOLOGIA
UNAM



