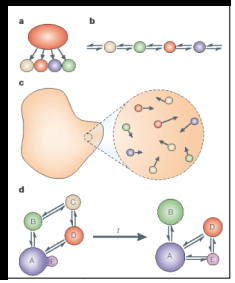
El flujo génico y la estructura de las poblaciones.

Quinta Sesión
GENETICA DE POBLACIONES:

Luis Eguiarte y
Valeria Souza



El flujo génico y la estructura de las poblaciones.

El modelo continente-isla de flujo génico.
Estimaciones directas e indirectas de flujo génico.
El efecto Wahlund.
Los estadísticos *F* de Wright.
Flujo génico y deriva.

FLUJO GENETICO gene flow:

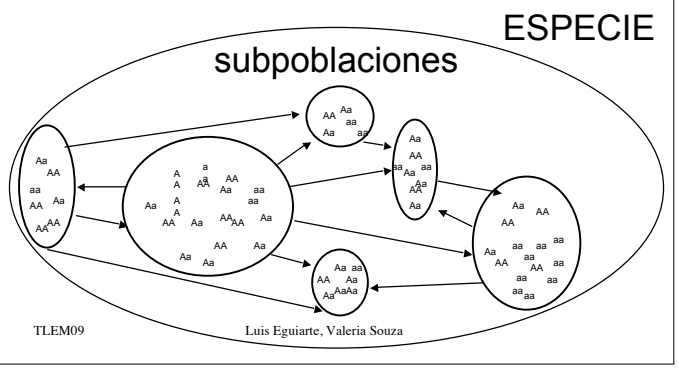
Es la violación al supuesto del equilibrio de Hardy-Weinberg que se refiere al "aislamiento" de la población.

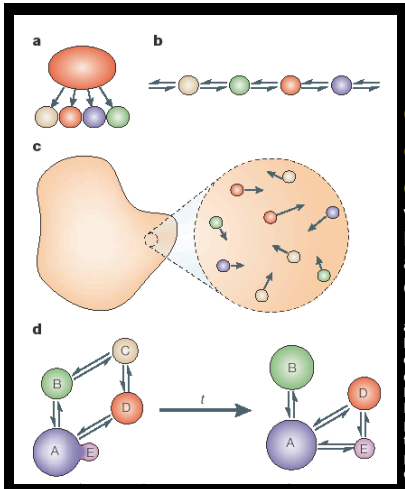
Es la incorporación de genes a nuestra poza génica provenientes de una o más poblaciones diferentes (Futuyma, 1986).

En genética de poblaciones también lo llamamos **migración**

***m*= migration rate, la tasa de migración**
la probabilidad de que un gen tomado al azar de una subpoblación sea un migrante

se considera que la especie esta formada por ***k*** subpoblaciones, con diferentes frecuencias alélicas, conectadas por diferentes **tasas de migración *m***





Existen diferentes posibles modelos de la estructura geográfica de las especies.

Violaciones al modelo básico,
ayudan a entender el papel de la geografía...

a | Island model of migration.
b | Stepping-stone model.
c | Isolation by distance model100,101.
d | Metapopulation model.
In nature, not only do individuals move between populations, but also individual populations come and go over time (*t*) with the founding and extinction of entire populations being an importantcomponent of population structure

El flujo génico **homogeniza** las diferentes poblaciones que forman a las especies.

así, el flujo génico se puede pensar que actúa contra la deriva génica y la selección... PUEDE DISMINUIR LA ADAPTACION LOCAL

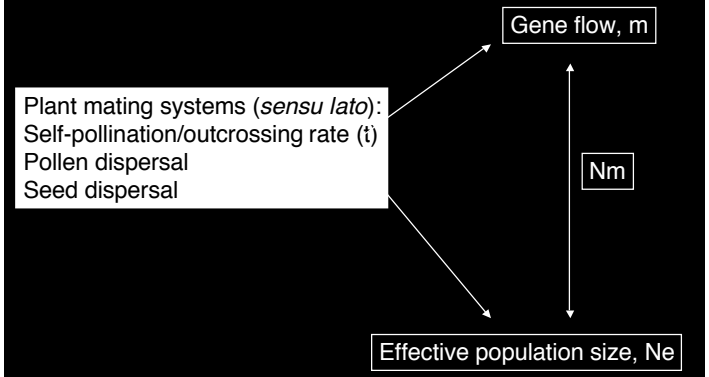
pero, por otra parte, incrementa la variación genética, sobre la que puede actuar la selección y reducir la carga genética

y podría jalar a las poblaciones los picos adaptativos más altos del modelo del Shifting balance.

Gene flow:

In practice is closely related to (and difficult to separate from):

Mating systems (t, outcrossing rates, etc.).
Pollen and seed dispersal (within a population).
Effective population size.
Hybridization



factors. Because of the confusion surrounding these terms, Endler (1977) made an effort to distinguish among gene flow, migration, and dispersal. Following his recommendation, we use **gene flow** to indicate *movement between groups that results in genetic exchange*.

incorporación de alelos a la poza génica

POPULATION STRUCTURE

A population may have substructure—differences in genetic variation among its constituent parts—for several different evolutionary reasons. For example, a population may have localized subpopulations in which there is genetic drift. Exchange of individuals may not have equal probabilities

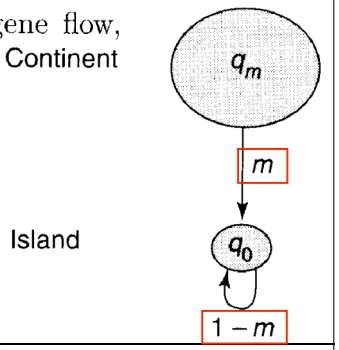
estructura genética = a la diferencia en frecuencias alélicas en el espacio

The Continent–Island Model

effectively unidirectional gene flow,

en una sola dirección,
de una fuente: continente
a un sumidero: isla

m = tasa de migración
proporción de migrantes
del continente a la isla
va de 0 a 1



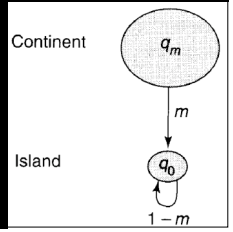
portion of migrants moving into the island population each generation be m and the proportion of nonmigrants (residents) be $1-m$. If the frequency of A_2 in the migrants (the continent) is q_m and the frequency of A_2 on the island before gene flow is q_0, the allele frequency after gene flow is

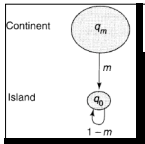
portion of migrants moving into the island population each generation be m and the proportion of nonmigrants (residents) be $1-m$. If the frequency of A_2 in the migrants (the continent) is q_m and the frequency of A_2 on the island before gene flow is q_0, the allele frequency after gene flow is

$$q_1 = (1-m)q_0 + mq_m$$

los que se quedan los que se llegan

$$= q_0 - m(q_0 - q_m)$$





$$= q_0 - m(q_0 - q_m)$$

$$\Delta q = q_1 - q_0$$

$$= -m(q_0 - q_m)$$

will be no change in allele frequency if $m = 0$ or if $q_0 = q_m$

stable equilibrium frequency of A_2 at $q_m = q_0$

en el equilibrio, la ISLA llega a la ***q* del continente**...

fórmula para el cambio en n generaciones, continente-isla

$$q_2 = (1-m)\,q_1 + mq_m$$

$$q_2 = (1-m)^2\,q_0 + \left[1-(1-m)^2\right]q_m$$

$$q_t = (1-m)^t q_0 + \left[1-(1-m)^t\right]q_m$$

As t increases, the first term in this expression approaches zero, and the second term approaches q_m so that q_t asymptotically approaches the equilibrium value of q_m.

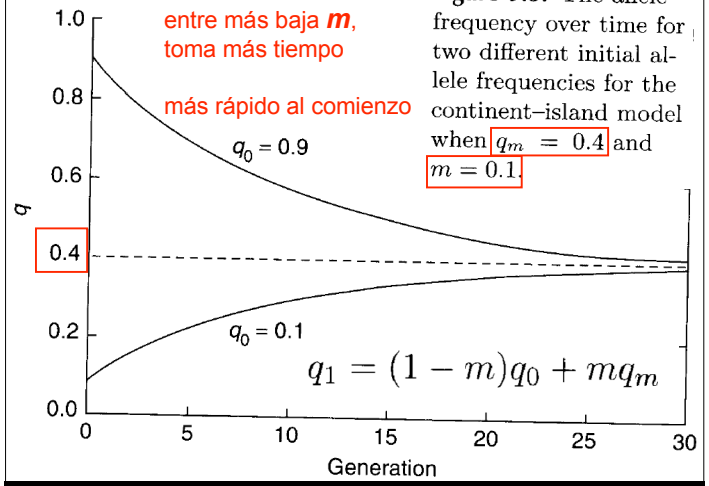


Figure 9.3. The allele frequency over time for two different initial allele frequencies for the continent–island model when $q_m = 0.4$ and $m = 0.1$.

LOBO ROJO
hibridiza con el coyote!

Example 9.1. Red wolves (*Canis rufus*) historically occurred throughout southeastern North America, but by the 1960s, they were confined to a small population in Louisiana and Texas where there was hybridization with the much more abundant coyote (*C. latrans*). Wolves from this population were captured to start a captive population, and in 1987, this captive population was used to establish a wild population in eastern North Carolina (Phillips *et al.*, 2003). However, over the next decade, coyotes colonized this area, and in the late 1990s, it was estimated that approximately 15% of the litters in the newly established population were hybrid.

m = ca. 0.075 de coyote

that $q_0 = 1$—that is, initially that there is 100% red wolf ancestry in the red wolves—and $q_m = 0$—all of the coyote genes do not reflect red wolf ancestry. In hybrid litters, half the ancestry is from red wolves and half from coyotes; thus, let us assume that $m = 0.075$, or half the rate of observed hybrid litters. Assuming that the generation length in red wolves is approximately 5 years, then Figure 9.4 shows that the proportion of red wolf ancestry is expected to drop quickly from 100% to approximately 46% after 50 years.

usando m= 0.075, en cuanto tiempo se "vuelven coyotes"??

en 50 años, 46% queda del lobo...

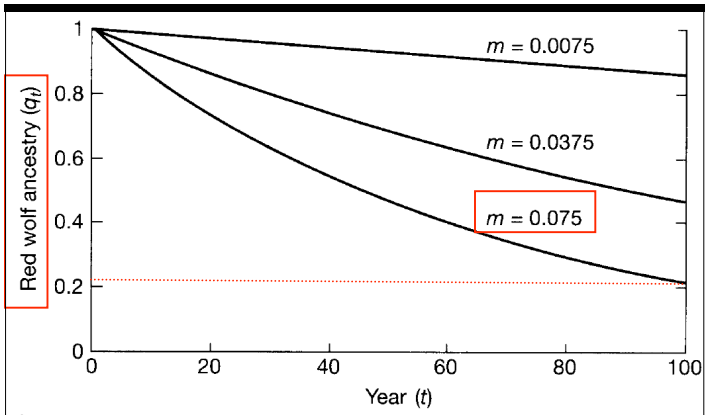
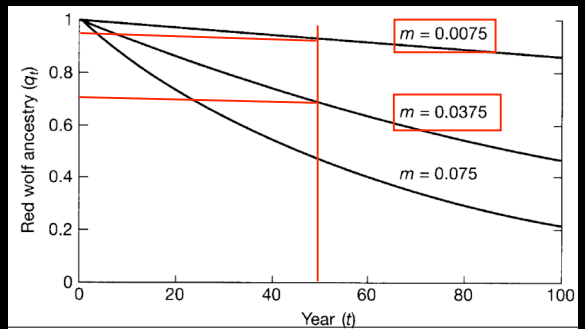


Figure 9.4. The predicted proportion of red wolf ancestry over time when there is gene flow from coyotes as might occur in a continent (coyote)–island (red wolf) model. The proportion of gene flow is assumed to be 0.075 without management intervention and either 0.0375 or 0.0075 if 50% or 90% of the hybrid litters are identified and eliminated.

¿MANEJO?

However, if there are management actions to identify (Miller *et al.*, 2003) and eliminate hybrid litters, then the rate of gene flow can be significantly reduced. For example, if 50% or 90% of the hybrid litters are eliminated so that $m = 0.0375$ or 0.0075, then the red wolf ancestry remains approximately 68% and 93%, respectively, after 50 years. This heuristic



ESTIMATION OF GENE FLOW AND POPULATION STRUCTURE

Estimating the amount of gene flow in most situations is rather difficult.

Direct estimates of the amount of gene flow can be obtained in organisms where different individuals can be identified. Direct observation of

muy difícil, problemas de n, muestreo, estimación

Indirect measures of gene flow using genetic markers are useful to confirm behavioral or other observations or when these observations are inconclusive or impossible. The simplest indirect estimate is for hybrid

Estimación Indirecta en Poblaciones Híbridas, 2 poblaciones

Hybrid populations—***those that receive migrants from two or more other populations***—are often of interest because they are at a border between two species or they may be undergoing rapid genetic change. Such populations are generally the result of intermating or the mixture of two or more parental populations over an extended period of time. If there are only two parental populations, then the ***proportion of gene flow from the outside population*** (or the continent), sometimes called **admixture**, in one generation can be estimated by rearranging equation 9.1*a* so that

se necesita saber las f alélicas originales, la final en la población híbrida y el tiempo

$$\hat{m} = \frac{q_0 - q_1}{q_0 - q_m}$$

$q_1 = (1 - m)q_0 + mq_m$

q_o= antes, q_1 =después, q_m = la de los migrantes

If gene flow has occurred over a number of generations, the symbols in this expression can be changed so that

isla antes

$$\hat{M} = \frac{q_A - q_H}{q_A - q_B}$$

"continente"

where $\hat{M}$ is the estimate of the total amount of gene flow of individuals from parental population B in the hybrid population H ($1 - \hat{M}$ is the proportion from parental population A) (Cavalli-Sforza and Bodmer, 1971). The allele frequencies in parental populations A and B and the hybrid population H are q_A, q_B, and q_H, respectively.

$$(1 - \hat{m})^t = \frac{q_t - q_m}{q_0 - q_m}$$

This can be rewritten to reflect gene flow of populations A and B into population H as

$$(1 - \hat{m})^t = \frac{q_H - q_B}{q_A - q_B}$$

$$\hat{m} = 1 - e^{\ln\left(1 - \hat{M}\right)/t}$$

where $1 - \hat{M} = (q_H - q_B)/(q_A - q_B)$. The value of $\hat{m}$ here is the estimate of the proportion of gene flow per generation into the hybrid population from parental population B. Of course, this estimate of gene flow assumes

Datos humanos, poblaciones en EUA de origen Africano

samples with predominantly African ancestry. The most extreme frequency difference is for the *FY-NULL* 1* marker, which was present in all of the Europeans examined and in none of the African individuals. The estimate of total gene flow (and the standard error) in the rightmost column is similar to that given in expression 9.5*a* but incorporates the effect of genetic drift and sampling variance (Long, 1991) and is combined over 10 loci.

TABLE 9.3 The allele frequencies of the four most extreme population-specific alleles in Africans, Europeans, and six U.S. populations of African descent (Parra *et al.*, 1998). In the rightmost column is the estimated European ancestral proportion for the populations of African descent, averaged over 10 population-specific alleles.

Population	*FY-NULL*1*	*OCA2*1*	*RB2300*1*	*GC-1F*	$\hat{M}^*$
African	0.000	0.098	0.920	0.824	
European	1.000	0.769	0.333	0.156	
African-American					
Maywood, IL	0.185	0.203	0.776	0.710	0.188 ± 0.014
New York	0.210	0.220	0.821	0.738	0.198 ± 0.021
Philadelphia	0.160	0.137	0.802	0.771	0.138 ± 0.019
Charleston, SC	0.112	0.208	0.888	0.765	0.116 ± 0.013
New Orleans	0.200	0.284	0.842	0.669	0.225 ± 0.016
Jamaica	0.065	0.091	0.870	0.790	0.068 ± 0.013

$$\hat{M} = \frac{q_A - q_H}{q_A - q_B}$$

Estimación indirecta de la tasa de migración, cont.

generation length is about 25 years, then the elapsed number of generations (t) during which gene flow has occurred is approximately 12. With these assumptions, $\hat{m}$ from expression 9.5*b* would be 0.021 for New Orleans and 0.010 for Charleston. The hypothetical pattern of gene flow over time for *FY-NULL*1* with these values would be as given in Figure 9.6. As expected,

si 12 generaciones, *m*= 0.021 a 0.010

$$\hat{m} = 1 - e^{\ln\left(1-\hat{M}\right)/t}$$

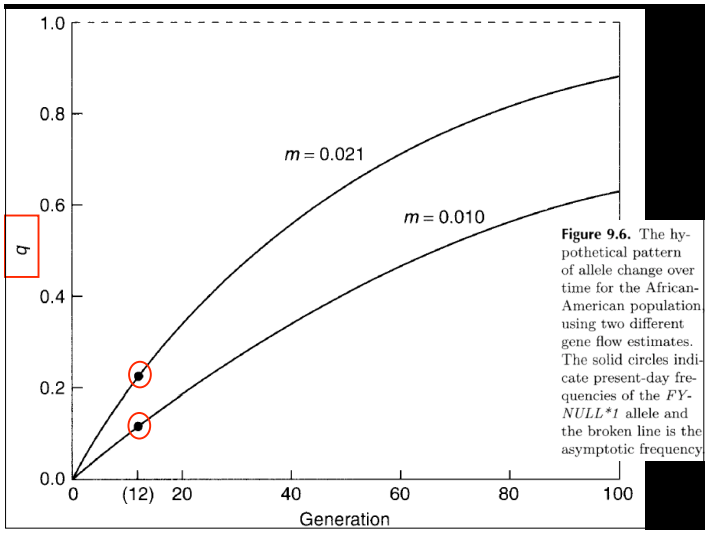


Figure 9.6. The hypothetical pattern of allele change over time for the African-American population using two different gene flow estimates. The solid circles indicate present-day frequencies of the *FY-NULL*1* allele and the broken line is the asymptotic frequency

by examining mtDNA and Y chromosome variants. For the nine United States samples for which there were data for both types of markers, estimated gene flow for mtDNA was 0.140 and for Y chromosomes was 0.248. Thus there appears to have been, historically, a sex-biased contribution of European ancestry, with more matings between European-American males and African-American females than vice versa.

diferencias mitocondria (m = 0.14) y Y (m= 0.248)...

Híbridas de varias fuentes

$$q_i' = m_{i1}\,q_1 + m_{i2}\,q_2 + \ldots m_{ij}\,q_j + \ldots m_{ik}\,q_k = \sum_{j=1}^{k} m_{ij}\,q_j$$

When there are several ancestral sources for a population, then the allele frequency can be calculated using expression 9.2*a*. If a population has ancestral contributions from several populations, then

$$q_H = M_1 q_1 + M_2 q_2 + M_3 q_3 \cdots M_i q_i$$

where M_i and q_i are the contributions and allele frequencies from the ith population.

CODIS MÉXICO 13 loci

For example, Cerda-Flores *et al.* (2002) estimated allele frequencies for the 13 core CODIS (Combined DNA Index System of the U.S. Federal Bureau of Investigation, see p. 639) microsatellite loci in a Mestizo (mixed race) population from northeastern Mexico. Previous studies have shown that in such populations there is ancestry from Europeans, Amerindians, and Africans. Table 9.4 gives the allele frequencies for two of these loci in the Mestizo population and from samples from Spain, Nigeria, and southwestern United States Amerindians. There is a general overlap in the presence of the alleles, but the representative ancestral populations differ in allele frequencies. Overall, Cerda-Flores *et al.* (2002) estimated that in this population 55%, 40%, and 5% of the ancestry are Spanish, Amerindian, and African, respectively. For example, if we examine allele 15 at locus D3S1358 (boldface in Table 9.4) using expression 9.5*c*, then the predicted $q_H = (0.55)(0.228) + (0.40)(0.653) + (0.05)(0.312) = 0.0402$ and the observed $q_H = 0.0.409$, only slightly different.

España= 55%, Amerindios=40% , Nigeria= 5%

TABLE 9.4 The estimated allele frequencies for two microsatellite loci (D3S1358 and VWA) from a Mestizo (mixed race) population from northeastern Mexico and from representative ancestral populations from Spain, Amerindians, and Africa (Cerda-Flores *et al.*, 2002).

	D3S1358				*VWA*			
Allele	*Mestizo*	*Spain*	*Amerindian*	*Africa*	*Mestizo*	*Spain*	*Amerindian*	*Africa*
11	—	—	—	0.010				
12	0.004	—	—	—				
13	0.010	0.011	0.008	—	—	—	0.001	—
14	0.073	0.080	0.061	0.135	0.087	0.116	0.045	0.087
15	**0.409**	**0.228**	**0.653**	**0.312**	0.105	0.167	0.036	0.304
16	0.238	0.243	0.168	0.312	0.318	0.268	0.439	0.239
17	0.147	0.217	0.077	0.177	0.297	0.173	0.254	0.174
18	0.115	0.199	0.032	0.052	0.136	0.174	0.145	0.087
19	0.004	0.022	0.001	—	0.056	0.076	0.066	0.065
20					—	0.025	0.014	0.011
21					—	—	—	0.022
22					—	—	—	0.011

Principio de Wahlund 1928, una consecuencia de la estructura, útil para definir los estadísticos **F**

Sometimes population substructuring is not obvious, and as a result, a sample may consist of a group of heterogeneous subsamples from a population. For example, subpopulations may be separated by subtle physical or ecological barriers that limit movement between groups. ***When these subpopulations are lumped together and if there are differences in allele frequencies among these subsamples, there will be a deficiency of heterozygotes and an excess of homozygotes***, even if Hardy–Weinberg proportions exist within each subsample (Wahlund, 1928). As is

a veces no sabemos que hay subestructura, pero si c/ subpoblacion frec. alélicas diferentes va a parecer que hay déficit de heterócigos

Sometimes population substructuring is not obvious, and as a result, a sample may consist of a group of heterogeneous subsamples from a population. For example, subpopulations may be separated by subtle physical or ecological barriers that limit movement between groups. ***When these subpopulations are lumped together and if there are differences in allele frequencies among these subsamples, there will be a deficiency of heterozygotes and an excess of homozygotes***, even if Hardy–Weinberg proportions exist within each subsample (Wahlund, 1928). As is

pob A		pob B		total		ESPERADOS
p=.25	q=.75	p=.75	q=.25	p=.5	q=.5	
AA	6	56		62	>	50
Aa	38	38		76	<	100
aa	56	6		62	>	50

SI NO SE SABE QUE SON DOS POBLACIONES, AUNQUE CADA UNA ESTA EN HW. EN TOTAL PARECE QUE FALTAN HETEROCIGOS!

ESTADISTICOS F DE WRIGHT, a partir de Walhund

F Coefficients and Other Differentiation Measures

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 interrelated

los estimadores más importantes de dif. en estudios empíricos

Fit = Fis + (1-Fis) Fst

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

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

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

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

Ho promedio de la heterocigosis observadas

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

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

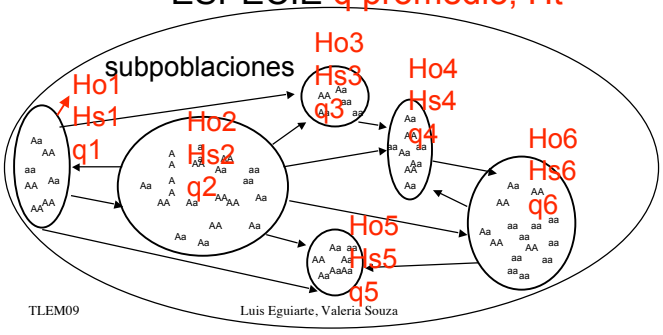
Ho promedio de la heterocigosis observadas en c/población

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

Ht 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, Ht



TLEM09 Luis Eguiarte, Valeria Souza

Fit = Fis + (1-Fis) Fst

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

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

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

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

Fis: efectos por endogamia/ sistemas reproductivos: -1 puros hetero., 0 HW, 1 endogamia total

Fst: diferenciación por deriva vs. migración o selección, 0 idénticas en f. alelicas, 1 totalmente diferentes

Fit= endogamia + deriva, -1 a 1

F_{ST}, F_{IT}, and F_{IS},

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

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

Fit = Fis + (1-Fis) Fst

Fit = Fis + (1-Fis) Fst

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

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

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

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

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

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

TABLE 9.5 Two hypothetical examples to illustrate the meaning of F coefficients.

Subpopulation	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.

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 selección diferencial
(otros menos diferenciación, selección balanceadora!)

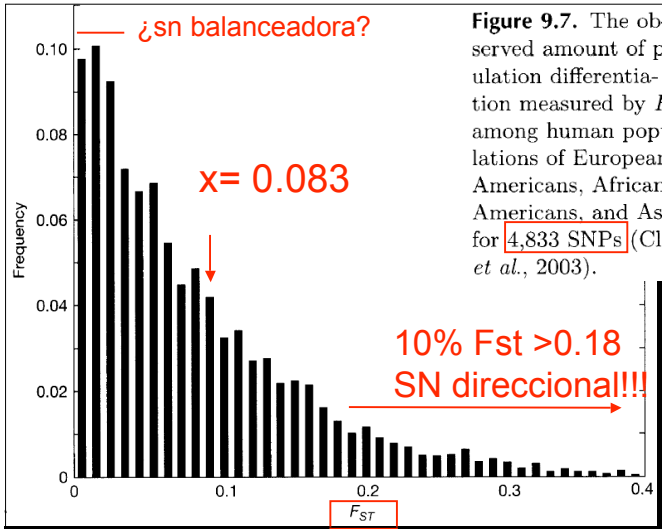


Figure 9.7. The observed amount of population differentiation measured by F_{ST} among human populations of European Americans, African Americans, and Asians for 4,833 SNPs (Clark *et al.*, 2003).

Gst: estimador de la Fst 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 $H_T = 1 - \sum \overline{p}_i^2$ 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 micros

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

Si hay mucha variación, la Gst puede ser MUY baja sin que se compartan alelos!

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

$$< 1 - H_S$$

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...

Otros niveles de partición! se pueden proponer n

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

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

Hs=0.263
Hr=0.357
Ht=0.647

Fsr= 0.253
Frt= 0.456

Ht

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

Fsr= 0.253

Frt= 0.456

casi el doble de la variación se encuentra entre regiones que las subpoblaciones dentro de un grupo!

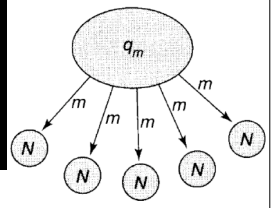
POPULATION STRUCTURE AND GENETIC DRIFT

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-

Balance Deriva- Migración

The Continent–Island or Island Model



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

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

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

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

el incremento de f por deriva , dado que ninguno de los alelos sea migrante

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

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

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

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

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

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

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

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

si *m* es pequeña

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

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

una guía general, a "orden de magnitud"

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

pero puede tener sesgos si F_{st} es chica... ca. 0.01

¿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 9.12*a*), then

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

dinámica pura deriva, de 0 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 Fst en eq.
si m es grande o N chica

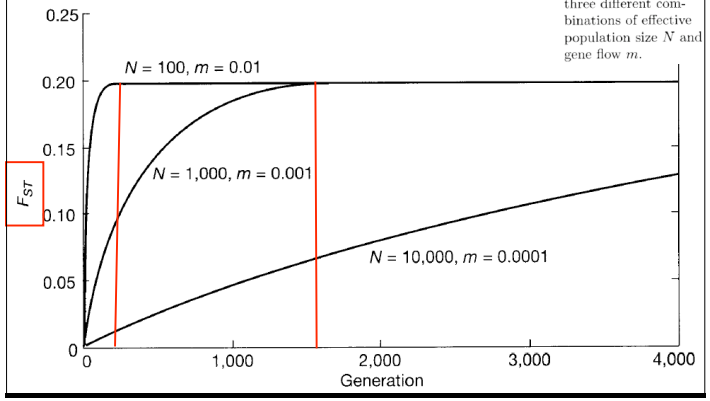
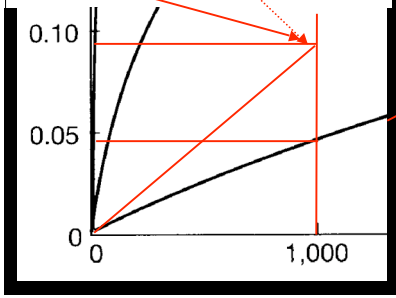


Figure 9.11. The amount of differentiation (F_{ST}) expected over generations for three different combinations of effective population size N and gene flow m.

difference in F_{ST} at a given point in time with and without gene flow. For example, with $N = 10,000$ using expression 9.13*a* 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.

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



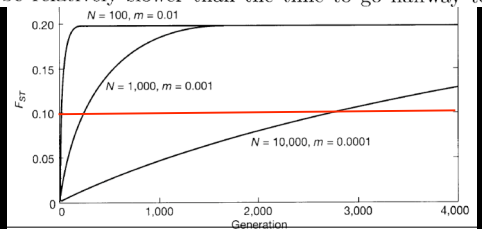
$$F_{ST} \approx \frac{1}{4Nm + 1}$$

poco flujo génico (m=0.001) hace mucho más **lenta la diferenciación**!

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 ½ de Fst

Efecto de la subdivisión en N_e

It is also useful to point out the effect of subdivision on the effective population size. Wright (1943b) showed that for the island model

$$N_e = kN\left[1 + \frac{(k-1)^2}{4Nmk^2}\right] = \frac{kN}{1 - F_{ST}}$$

Nm, is large or F_{ST} is low, then $N_e \approx kN$

si Nm grande, Ne casi kN

$$N_e = kN\left[1 + \frac{(k-1)^2}{4Nmk^2}\right] = \frac{kN}{1 - F_{ST}}$$

Nm, is large or F_{ST} is low, then $N_e \approx kN$

flow is low, $4Nm < 1$, then the effective population size may be larger than kN. if $4Nm = 0.25$ ($F_{ST} = 0.8$), then $N_e \approx 5kN$

low level of gene flow allows each subpopulation to evolve independently

pero si *Nm* pequeña, puede se mucho más grande que *kN*

Resultados de Wright 1940 en el equilibrio flujo/deriva

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 $4Nmq_m$ 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 hand, if $4Nmq_m$ and $4Nm(1-q_m)$ are less than one and q_m is nearly 0.5, then the distribution of alleles over island populations becomes U shaped, and most of the islands become fixed or nearly fixed for either A_1 or A_2 (e.g., $Nm = 0.05$ in Figure 9.12). These distributions are examples of stable

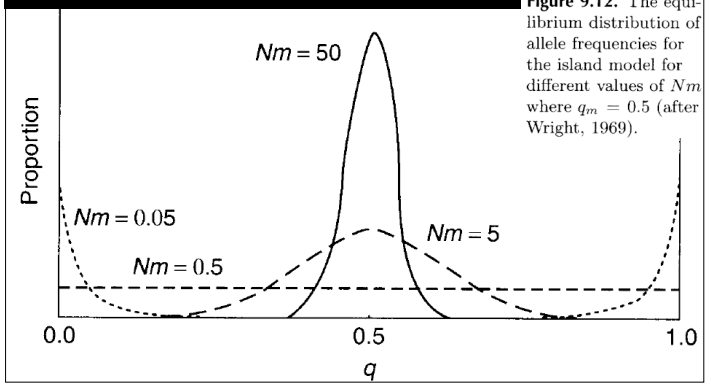
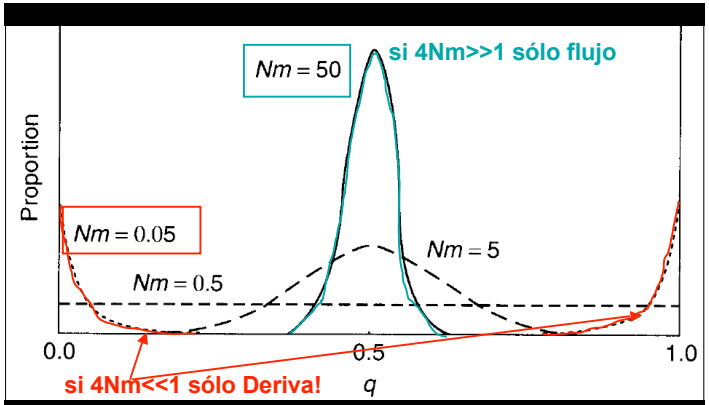


Figure 9.12. The equilibrium distribution of allele frequencies for the island model for different values of Nm where $q_m = 0.5$ (after Wright, 1969).



Isolation by Distance:
el aislamiento por distancia de 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 discusses the

Nm, Fst/(1-Fst) 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.

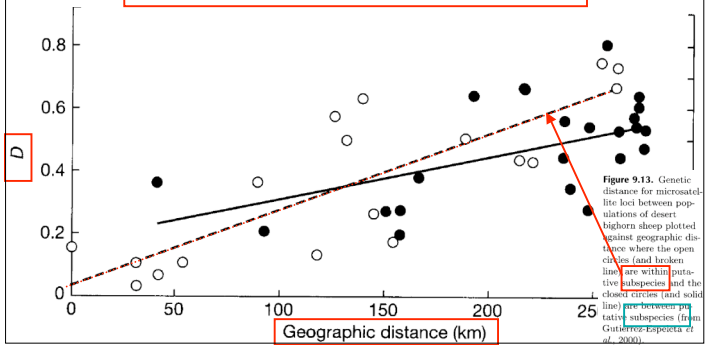


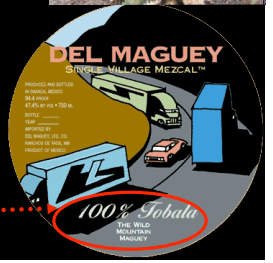
Figure 9.13. Genetic distance for microsatellite loci between populations of desert bighorn sheep plotted against geographic distance where the open circles (and broken line) are within putative subspecies and the closed circles (and solid line) are between putative subspecies (from Gutiérrez-Espeleta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

Dos especies mezcaleras muy cercanas

***Agave potatorum* y *A. cupreata*:**

"Papalometl" (ancho) y "Tobalá"

ENRIQUE SCHEINVAR!

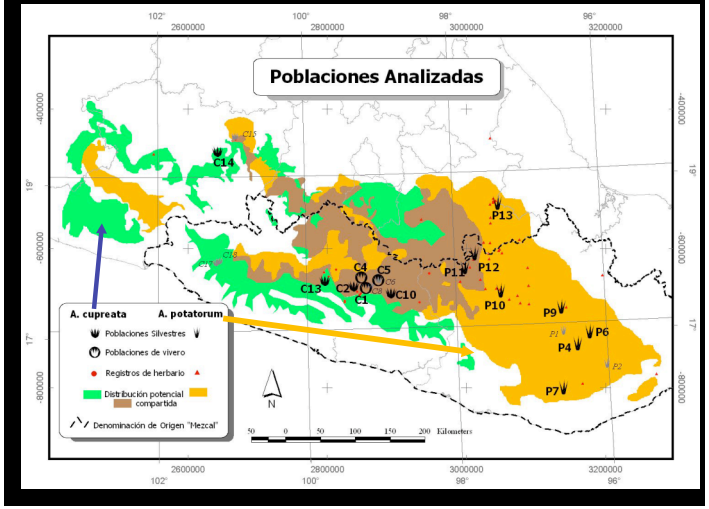


Mezcal papalote / tobalá: **90 loci, ISSRs**

Xitlali Aguirre y Enrique Scheinvar

Oaxaca *A. potatorum*

Guerrero y Michoacán *A. cupreata*

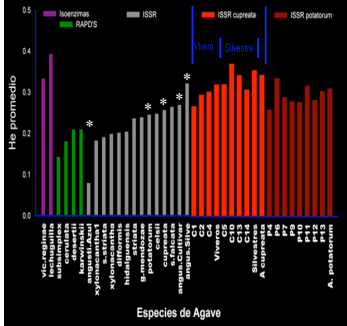


Variación Genética en *AGAVE CUPRETA* y *A. POTAROUM*
ligeramente más alta en *cupreata*

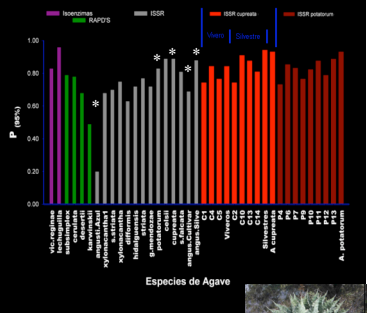
Población	Ubicación	N	He ($V_{(He)}$)		$P_{(95\%)}$ ($V_{(P)}$)	
C1 Vivero "Ayahualco"	Chilapa, Gro.	25	0.2689	(0.0008)	74.44	(0.192)
C4 Vivero "La Esperanza"	Chilapa, Gro.	34	0.3043	(0.0010)	84.44	(0.133)
C5 Vivero "Trapiche Viejo"	Chilapa, Gro.	41	0.3225	(0.0012)	76.67	(0.181)
C2 "Ayahualco"	Chilapa, Gro.	32	0.2969	(0.0010)	74.44	(0.192)
C10 "Mesones"	Tlapa, Gro.	41	0.3723	(0.0015)	91.11	(0.082)
C13 "La Laguna"	Xochipala, Gro.	28	0.3445	(0.0013)	87.78	(0.108)
C14 "Etucuaro"	V. Madera, Mich.	33	0.3096	(0.0011)	81.11	(0.155)
Viveros	**Chilapa, Gro.**	**100**	**0.3219**	**(0.0012)**	**84.44**	**(0.133)**
Silvestres	**Chilapa, Tlapa, Villa Madera**	**134**	**0.3566**	**(0.0014)**	**94.44**	**(0.053)**
A. cupreata	**Gro-Mich.**	**234**	**0.3452**	**(0.0013)**	**93.33**	**(0.063)**
P4 "San Dionisio"	Sn. Dionisio Ocotepec, Oax.	22	0.2605	(0.0007)	73.33	(0.198)
P6 "Albarradas"	Sn. Lorenzo Albarradas, Oax.	31	0.3374	(0.0013)	85.56	(0.125)
P7 "Miahatlán"	Sto. Tomás Tamazulapa, Oax.	49	0.2909	(0.0009)	83.33	(0.140)
P9 "Sta. Catarina"	Sta. Catarina, Oax.	48	0.2813	(0.0009)	76.67	(0.180)
P10 "Yanhutitlan"	Sto. Domingo Yanhuitlan, Oax.	31	0.2794	(0.0009)	82.44	(0.133)
P11 "Zapotitlán"	Zapotitlán Palmas, Oax.	46	0.3183	(0.0011)	87.78	(0.108)
P12 "Tequistepec"	Sn. Pedro yPablo Tequixtepec, Oax.	46	0.2837	(0.0009)	78.98	(0.168)
P13 "Azumbilla"	Chapulco, Puebla	38	0.306	(0.0010)	88.89	(0.100)
A. potatorum	**Oax.-Pue**	**311**	**0.3122**	**(0.0011)**	**93.33**	**(0.063)**
Total todas	**Gro-Oax-Mich-Pue**	**545**	**0.331**	**(0.0012)**	**96.67**	**(0.032)**

Variación Genética en *Agave*, especialmente *A. cupreata y A. potarorum*

He, heterocigosis eserada

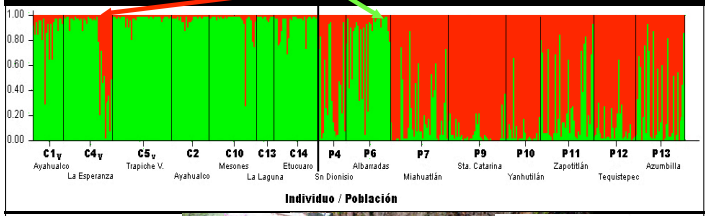


P, *proporción de loci polimórficos*



A pesar de manejo, no se ha perdido variación genética, ni siquiera en los viveros de Chilapa, Guerrero
pero parece estar promoviendo la hibridización entre las dos especies...

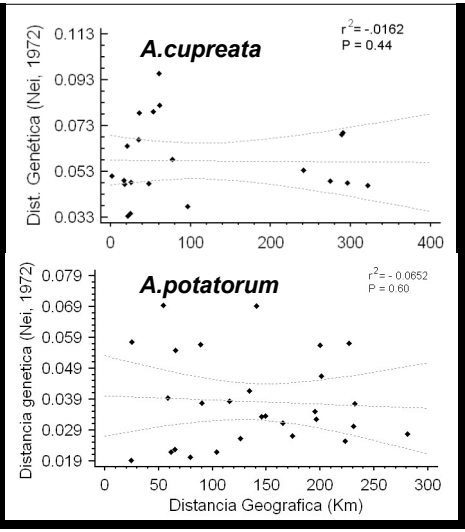
Structure: Análsis Bayesiano para asignar grupos
Agave cupreata verde *Agave potatorum* rojo



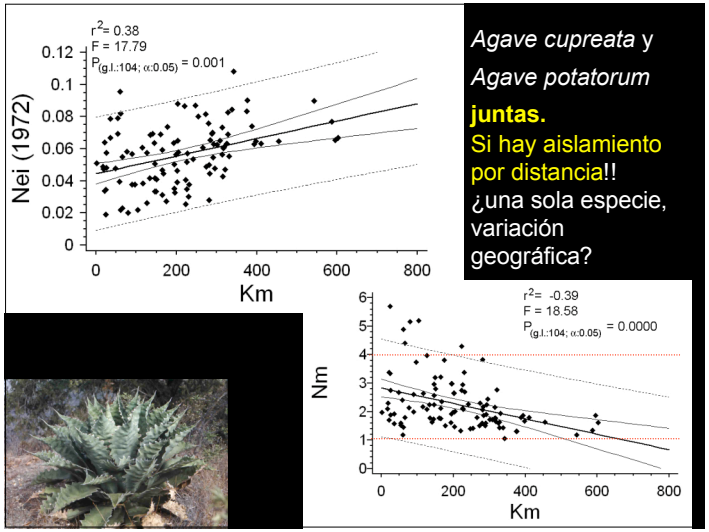
Diferenciación genética: mayor en *A. cupreata*
F_{st} = 0.1044, *Nm* =1.58,

que en *A. potatorum*
F_{st} = 0.0791, *Nm* = 2.23

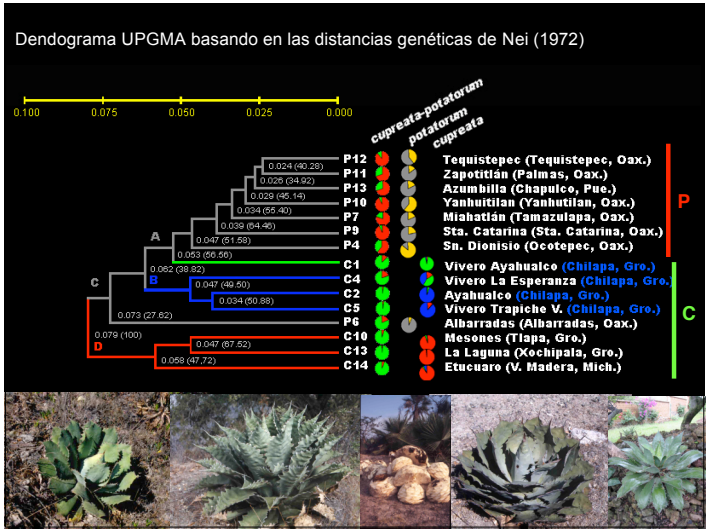
esto es relevante para el manejo con conservación de los recurso genéticos y puede servir para definir denomicnaciones de origen finas...

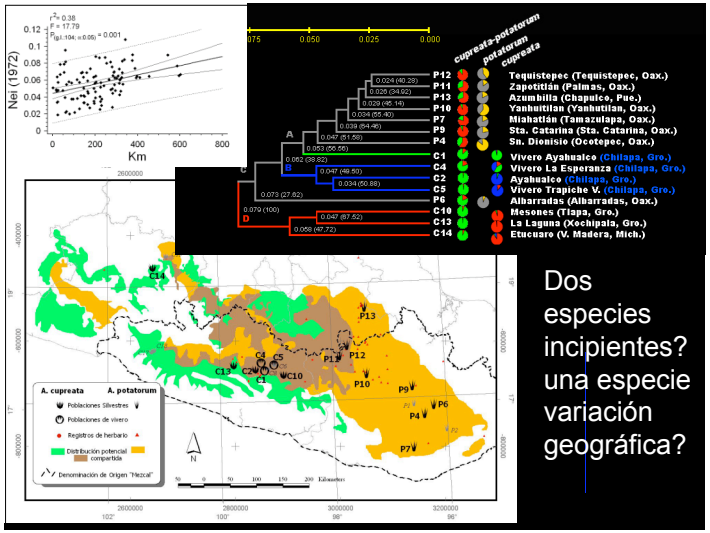


No hay aislamiento por distancia dentro de cada especie



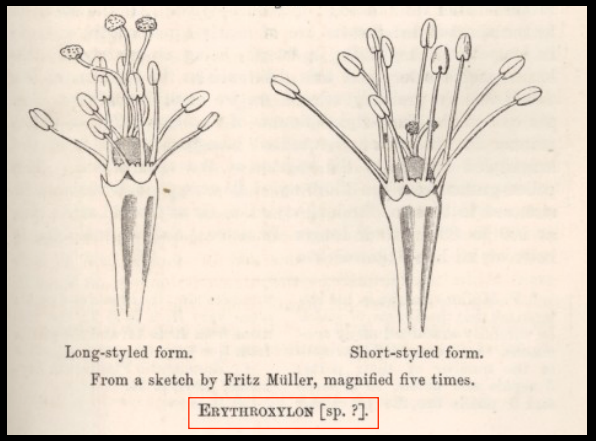
Agave cupreata y *Agave potatorum* **juntas.**
Si hay aislamiento por distancia!!
¿una sola especie, variación geográfica?





Dos especies incipientes? una especie variación geográfica?

Darwin Heterostilia



New Phytologist

New Phytologist (2005) **166**: 663–672

Research

Local genetic differentiation among populations of the mass-flowering tropical shrub *Erythroxylum havanense* (Erythroxylaceae)

César A. Domínguez[1], César A. Abarca[1], Luis E. Eguiarte[1] and Francisco Molina-Freaner[2]

Rapd´s DOMINANTES
5 primers, 51 "loci"

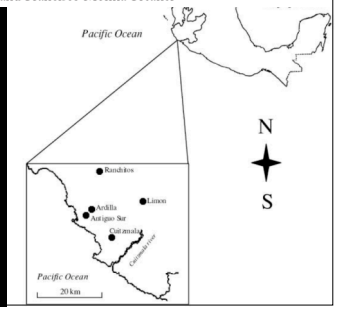
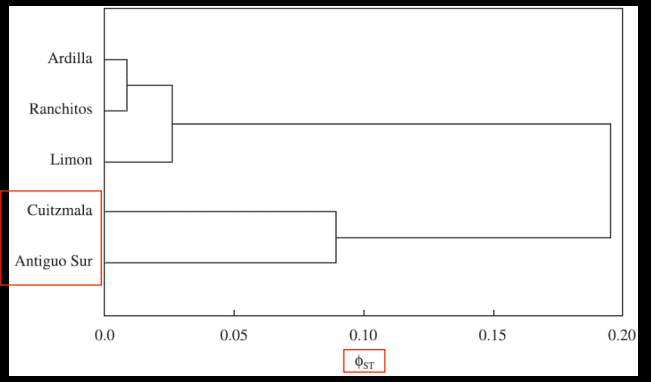


Table 1 Estimates of genetic diversity for five closely located populations of *Erythroxylum havanense* from the Pacific coast of Mexico

Population	*n*	H_E	*H'*
Ranchitos	27	0.437 (0.0080)	0.57
Ardilla	36	0.475 (0.0037)	0.73
Antiguo Sur	23	0.463 (0.0049)	0.87
Cuitzmala	22	0.417 (0.0142)	0.82
Limon	36	0.363 (0.0271)	0.47



mucha diferenciación

Table 2 Summary of the analysis of molecular variance (AMOVA) performed on five populations of *Erythroxylum havanense* from western Mexico

Source of variation	df	VC	% Total	*P*
Among populations	4	0.67	13.76	< 0.0002
Within populations	139	4.19	86.24	< 0.0002
Total	143			

VC, variance-components estimates; % Total, percentages of the total variance contributed by each component.

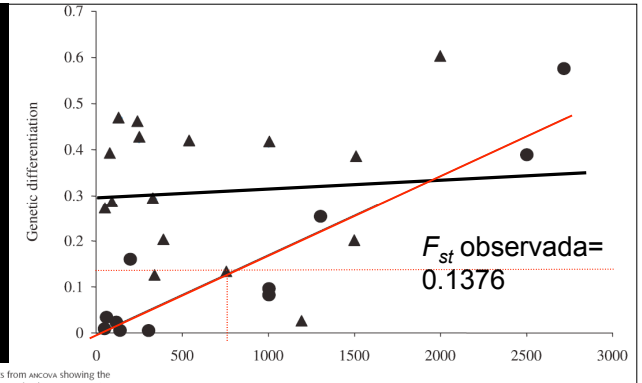
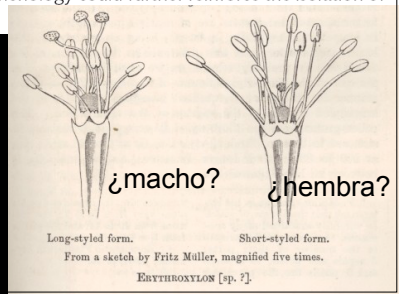


Fig. 4 Results from ANCOVA showing the predicted relationship between genetic differentiation (Φ_{ST}, G_{ST}, or estimations based on Shannon's information index) and the maximum geographic distance between sampled populations of 27 tropical perennial outcrossing species. The solid line indicates the adjusted model for species with undisturbed natural populations (circles, −0.000082 + 0.0002 × distance), while the dotted line accounts for species with evidence of disturbance or a special status (triangles, 0.3028 + 0.000019 × distance).

mucha más diferenciación de la esperada a 20 km comprando con especies de ecología similar… ¿floración sincróncia por parches? ¿limitada dispersión?

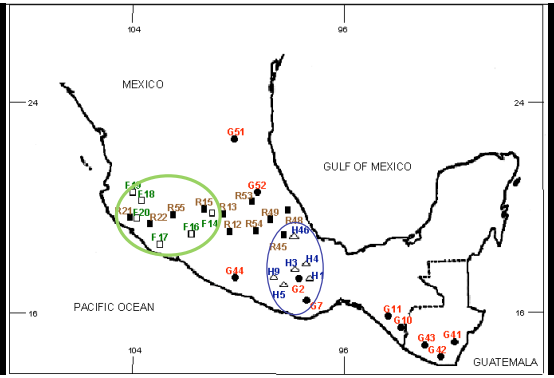
- High levels of genetic variation within populations and significant differentiation among populations located very near to each other were found. Furthermore, spatial autocorrelation analysis indicated the presence of significant genetic structure at short spatial distances.
- We suggest that by influencing the foraging behavior of pollinators and frugivores, mass flowering may produce the observed patterns of genetic structure, while small differences in flowering or fruiting phenology could further reinforce the isolation of nearby populations.

otro datos sugieren una posible **evolución al diocismo**, sexos separados



Analysis in the four species and 33 populations of firs, genus ***Abies*** **(Pinaceae)** from southern Mexico.

Aguirre-Planter et al. (2000, Am. J. of Botany 87: 362-371).
Jaramillo-Correa, J. P. et al. (2008 Molecular Ecology. **17**: 2476–2490).



Distribution of the Southern *Abies*:
Abies guatemalensis G, Wide N-S also El Salvador and Honduras
A. religiosa R, Wide E-W
A. hickeli H Restricted S
A. flinckii F, Restricted W

Different indirect methods for the estimate of $N_e m$ from genetic variation:

a) From $\mathbf{F_{st} = 1/ (4 N_e m +1)}$

b) From the **private alleles** method of Slatkin (1985):
Average frequency of alleles found in only one population
$\log10[p(1)] = a \log 10\ N_e m + b$
a and b are constants derived by simulated data
Logic: with high gene flow, the rare alleles should be very rare

c) **Isolation by distance analysis** (genetic differentiation as a function of the geographic distance), using estimates of $N_e m$ for pairs of populations (M), following Slatkin (1993)

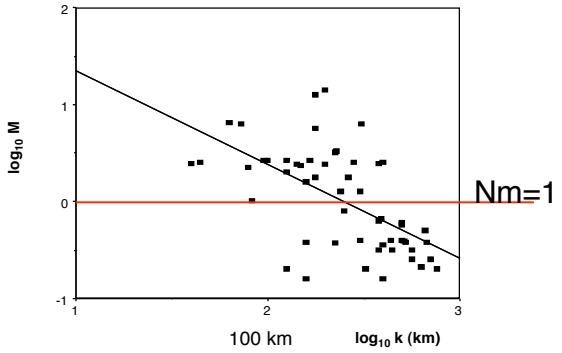
Estimates from F_{st} and private alleles:

16 allozymic loci	F_{st}	F_{st} $N_e m$	Private alleles $N_e m$
A. flinckii R W	0.271	0.672	3.42
A. guatemalensisW N-S	0.122	1.8	2.88
A. hickeli R S	0.073	3.17	2.70
A. religiosa W E-W	0.250	0.75	1.67

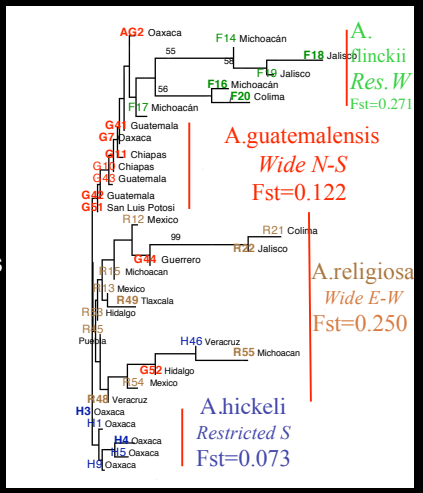
Higher in private alleles, but both suggest relatively **high levels of gene flow** among populations.

Isolation by distance (M= Nm pareado) analysis:

Indicates only a clear pattern of genetic differentiation as a function of distance in *A. religiosa* (11 populations).



Allozymes **Neighbor-joining** analysis using Nei´s distance, *Abies* of southern Mexico



Chloroplast microsatellites (cpSSRs)

Sampling:
19 populations:

Abies guatemalensis G, Wide N-S	8
A.religiosa R, Wide E-W	4
A.hickeli H Restricted S	4
***A.flinckii* F, Restricted W**	**3**

16 to 29 individuals per population (average 20 per population)

wind pollinated, **chloroplast travels with pollen (male part of the history)** (female part seed, mitochondria...).

Variación genética en los microsatélites de cloroplasto de *Abies*:

Total 51 haplotypes, 7 unique haplotypes in *A. flinckii*, the other species share most of them

Species	*Nind*	*Nhapl*	*hap/ind*	*He*	*Fst*	*Rst*
A.guatemalensis	160	31	0.19	0.866	Fst= 0.131	Rst= 0.104
A. religiosa	74	21	0.28	0.888	Fst= 0.075	Rst= 0.136
A. flinckii	70	10	0.14	0.750	Fst= 0.029	Rst= 0.116
A. hickeli	69	21	0.30	0.946	Fst= 0.065	Rst= 0.040

R_{st} theoretically better for microsatellites (stepwise model vs. infinite alleles model), but poor performance in empirical studies (very large sample sizes and time needed, Goldstein et al., 1995; Gaggiotti et al. 1999)

Microsatellite Pt 30204 mutation hotspot in *Abies*?
(Vendramin et al, 1999)

	Chloroplast				Nuclear
A.guatemalensis	Fst= 0.131	Nm= 3.32	Rst= 0.104	Nm=4.32	Nm=1.8
A. religiosa	Fst= 0.075	Nm=6.09	Rst= 0.136	Nm=0.136	Nm=0.75
A. flinckii	Fst= 0.029	Nm=16.74	Rst= 0.116	Nm=3.81	Nm=0.67
A. hickeli	Fst= 0.065	Nm=7.19	Rst= 0.040	Nm=11.86	Nm=3.17

Always higher *Nm* for Chloroplast

A.guatemalensis	*1.8 times higher chl/nuclear*
A. religiosa	*9.8*
A. flinckii	*25*
A. hickeli	*2.3*

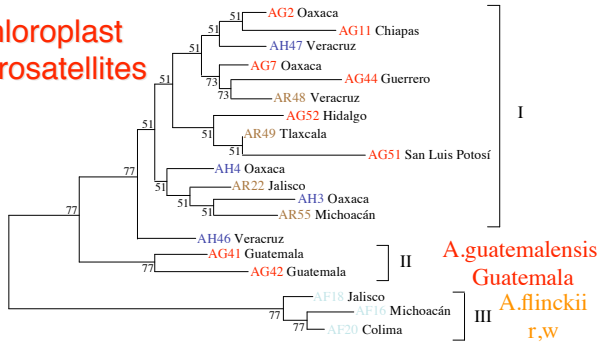
AMOVA chloroplast (Excoffier et al., 1992)

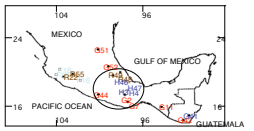
88.4% variation within populations
5.50% among population of the same species
5.66% among species

Congruent with the allozyme study
(5.6% among species!)

This is clearer in the NJ diagram

Chloroplast microsatellites





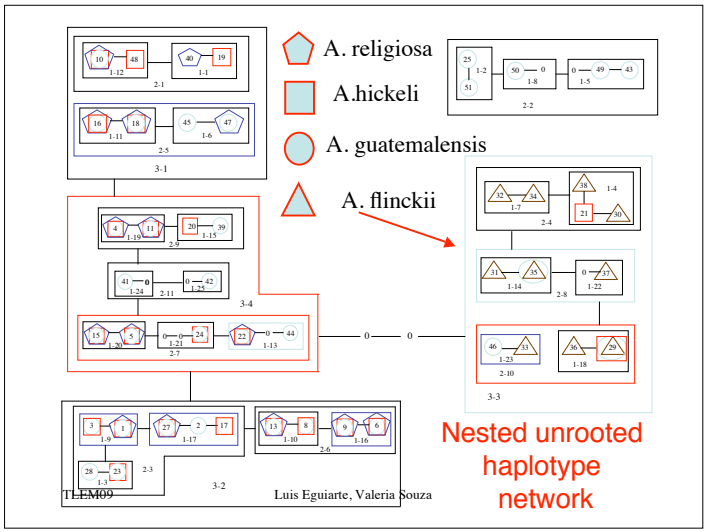
A.guatemalensis
Guatemala
A.flinckii
r,w

Indicate **complex gene flow by pollen (introgression)** among species in central Mexico....

Nested clade analysis chloroplast microsatellites (cpSSRs)

Templeton et al. 1987, 1992, 1995, 2001
TCS: Construct network (Clement et al. (2000))
GEODIS 2.0: Exact permutation contingency test and distance analyses (Posada et al. (2000))

All species and populations together



Analysis based on the key by Templeton et al. 1995:

Fragmentation events: in clades 1-9, 1-16, 1-17, 1-23, 2-5

Restricted gene flow from isolation by distance clades 2-7, 2-10, 3-4

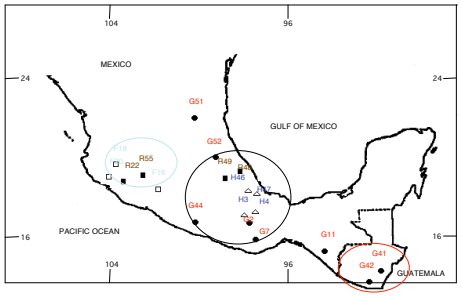
Rage expansion: clade 1-13

Ancient haplotypes located at the center of the nested network and shared by *A. guatemalensis*, *A. hickeli* and *A. religiosa*, persistence of ancestral polymorphisms

I.e., a complex history for the haplotypes of the chloroplast in Mexican *Abies*...

Chloroplast genome shared among central Mexican populations regardless of the species.

Differences in *A. flinckii* and the southern *A. guatemalensis* populations



American Journal of Botany 89(7): 1156–1163. 2002.

GENETIC STRUCTURE AND INDIRECT ESTIMATES OF GENE FLOW IN THREE TAXA OF *CUCURBITA* (CUCURBITACEAE) IN WESTERN MEXICO[1]

SALVADOR MONTES-HERNANDEZ[2,3,4] AND LUIS E. EGUIARTE[3]

The *Cucurbita* complex in traditional agroecosystems. Salvador Montes

A study of several cultivated and wild taxa form western Mexico (Jalisco: Mun. Autlán, Ejutla, El Limón, El Grullo):

3 taxa:
- The cultivated *Cucurbita moschata*.
- The cultivated *C. argyrosperma* spp. *argyrosperma*.
- The wild *C. argyrosperma* spp. *sororia*.

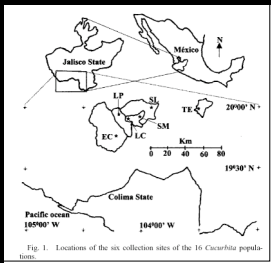
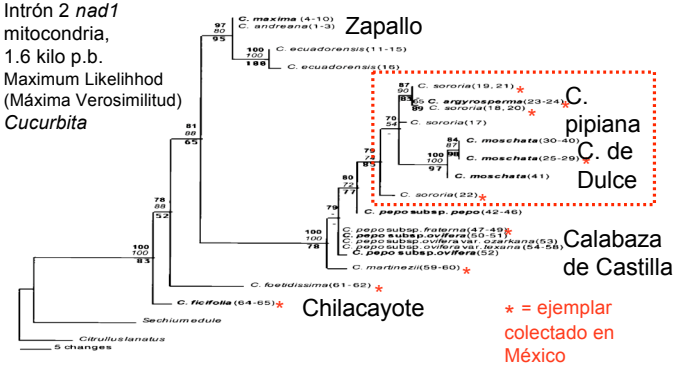


Fig. 1. Locations of the six collection sites of the 16 *Cucurbita* populations.

Intrón 2 *nad1* mitocondria, 1.6 kilo p.b.
Maximum Likelihhod (Máxima Verosimilitud)
Cucurbita



* = ejemplar colectado en México

Sanjur O. I., D. R. Piperno, T. C. Andres y L. Wessel-Beaver. 2002. Phylogenetic relationships among domesticated and wild species of *Cucurbita* (Cucurbitaceae) inferred from a mitochondrial gene: Implications for crop plant evolution and areas of origin. PNAS 99:535–540.

Indirect analysis:

12 allozymic loci, 15 populations from 6 localities.

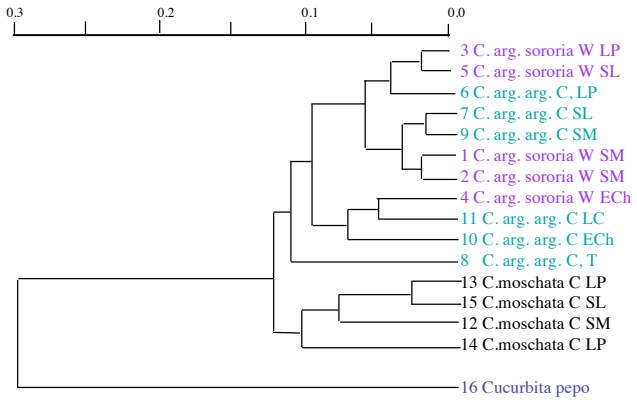
	F_{is}	F_{st}	N_em
Mean	-0.061	0.087	2.62
95%CI	(-0.15, 0.17)	(0.05, 0.12)	(1.17, 4.66)

This general analysis suggest gene flow among populations in the complex.

Analyzing each taxa:	N	F_{is}	F_{st}	N_em
Cult.*Cucurbita moschata.*	4	-0.162	0.077*	3.00
Cult. *C. a.* spp.*argyrosperma*	6	-0.042	0.096*	2.34
Wild*C. a.* spp. *sororia.*	5	0.004	0.040*	5.95

The gene flow seem to be as high within a taxa than in than for the complete complex.

An UPGMA analysis suggests patterns of gene flow among taxa, in particular within a locality in *C. argyrosperma.*



Gene flow within localities:

To test for gene flow within localites, we estimated the F_{st} and N_em in the sites were we have the 3 taxa:

	F_{st}	N_em
San Miguel	0.058	4.06
San Lorenzo	0.062	3.76
El Chante	0.099	2.27
Los Parajitos	0.116	1.90

Intense gene flow among taxa in some localities (San Miguel, San Lorenzo)?:

Direct analyses of gene flow.

TOTAL	F_{is}	F_{st}	N_em
Mean	-0.061	0.087	2.62
95%CI	(-0.15, 0.17)	(0.05, 0.12)	(1.17, 4.66)

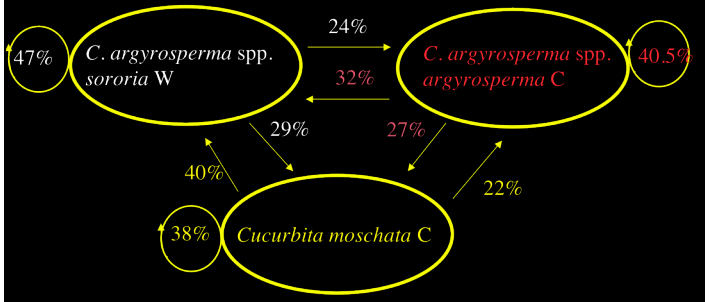
Direct analyses of gene flow:

Bees in the flowers (expected), El Chante, 1999

Bee species	C. moschata(C)	C. argy. argy.(C)	C.argy. sor.(W)
Peponapis azteca	133(156)	177(171)	209(191)
Xenoglossa gabbi	37(24)	17(27)	27(30)
Xenoglossa fulva	10 (5)	5(5)	0(6)
Apis mellifera	13(10)	11 (11)	9(12)
Augochlora smaragdina	11(11)	14(12)	10(13)
Otros	19 (17)	20(19)	17(20)

Although there are some preferences by the bees, the 3 taxa share the same pollinators, suggesting gene flow among them.

Direct analyses of gene flow.
Analyses using fluorescent dyes
Five days of study in a plot in El Chante, 1999.
We analyzed 198 stigmas.
More than 50% of the pollen moves to other taxa



fin Quinta Sesión