

***Flujo génico, diferenciación y estructura genética de las poblaciones, con ejemplos en especies de plantas mexicanas***

*Versión 4 de octubre del 2010*

*Luis E. Eguiarte, Erika Aguirre-Planter, Enrique Scheinvar, Andrea González González y Valeria Souza.*

Laboratorio de Evolución Molecular y Experimental, Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México.  
[fruns@servidor.unam.mx](mailto:fruns@servidor.unam.mx)

**Índice:**

<b>Introducción</b>	<b>2</b>
<b>La estructura genética de las poblaciones</b>	<b>3</b>
<b>Frecuencias alélicas</b>	<b>4</b>
<b>Diferenciación entre poblaciones</b>	<b>4</b>
<b><i>Los estadísticos <math>F</math> de Wright</i></b>	<b>5</b>
<b>Valores de <math>F_{st}</math> en poblaciones naturales</b>	<b>6</b>
<b>Cuidado en el uso e interpretación de la <math>F_{st}</math> y otros métodos relacionados</b>	<b>6</b>
<b>Estimaciones de endogamia: <math>F_{is}</math></b>	<b>8</b>
<b>Flujo génico</b>	<b>9</b>
<b>Métodos directos para estimar el flujo génico</b>	<b>10</b>
<b>Métodos indirectos para estimar el flujo génico</b>	<b>11</b>
<b>¿En cuanto tiempo la <math>F_{st}</math> llega al equilibrio?</b>	<b>12</b>
<b>Otros métodos para estimar indirectamente el flujo génico</b>	<b>13</b>
<b>Aislamiento por distancia</b>	<b>13</b>
<b>Estructura genética fina</b>	<b>14</b>
<b>Paquetes y programas para el estudio del flujo génico y la estructura genética</b>	<b>15</b>
<b>Patrones y perspectivas</b>	<b>15</b>
<b>Agradecimientos</b>	<b>15</b>
<b>Referencias citadas en el texto (incluyendo figuras y recuadros)</b>	<b>16</b>
<b>Recuadro 1: Otros índices que miden la diferenciación genética entre poblaciones relacionados con la <math>F_{ST}</math></b>	<b>20</b>
<b>Recuadro 2. El método de los alelos privados (únicos)</b>	<b>20</b>
<b>Recuadro 3. Las distancias genéticas</b>	<b>21</b>
<b>Pies de figuras</b>	<b>22</b>
<b>Tabla 1</b>	<b>24</b>
<b>Tabla 2</b>	<b>28</b>
<b>Tabla 3</b>	<b>29</b>
<b>Tabla 4</b>	<b>30</b>

## Introducción

Las frecuencias alélicas, o sea la proporción de cada uno de los alelos en un gene dado (usualmente denotadas por las letras  $p$  y  $q$ , ver Hedrick, 2005, Eguiarte, 2009) generalmente son diferentes entre las poblaciones que forman a una especie. En la Fig. 1 ilustramos esta idea: tenemos seis poblaciones, cada una formada por distintos números de individuos homócigos  $AA$  o  $aa$  o heterócigos  $Aa$ . Estas poblaciones están conectadas entre si por diferentes tasas de flujo génico o tasas de migración,  $m$ , representadas por las flechas que conectan a estas poblaciones y cada población tiene distintas proporciones de los genotipos y de los alelos  $A$  y  $a$ .

Las diferencias en las frecuencias alélicas entre las poblaciones son consecuencia de las acción de las fuerzas evolutivas a lo largo del tiempo, y por eso funcionan como un archivo de la historia evolutiva de las poblaciones. La selección natural, al adaptar a cada una de las poblaciones a sus condiciones locales, generalmente va a aumentar la diferenciación, siempre y cuando las condiciones sean diferentes en cada localidad y el gen que se estudia tenga que ver con esta adaptación o se encuentre cercano en el cromosoma (ligado) a un gen seleccionado. Los otros efectos de la selección natural se exploran brevemente en Eguiarte (2009) y en diferentes textos, por ejemplo en Futuyma (1998 págs. 337 a 398) y Hedrick (2005 págs. 213 a 236). La deriva génica es un proceso azaroso que siempre produce que se diferencien (que diverjan) las diferentes poblaciones de una especie y va a ser más importante entre más pequeñas sean las poblaciones (Hedrick, 2005; Eguiarte, 2009). La mutación también incrementa la diferenciación genética, pero debido a que usualmente las tasas de mutación son muy bajas, su papel en generar diferenciación dentro de una especie es mínimo. A contrario, el flujo génico, también llamado migración, al mover genes entre las poblaciones que forma a una especie evita que diverjan al homogenizar las frecuencias alélicas.

El flujo génico es la violación al supuesto del equilibrio de Hardy-Weinberg que se refiere al “aislamiento” de la población. Formalmente, el flujo génico se define como la incorporación de genes a una poza génica provenientes de una o más poblaciones diferentes dentro de una especies (Fig. 1). Así, el flujo génico funciona como una fuerza evolutiva cohesiva, manteniendo como un todo evolutivo a las especies y evitando que sean demasiado diferentes cada una de las poblaciones que la forman (Fig. 1). La tasa de migración  $m$  es el parámetro que necesitamos conocer para cuantificar la relevancia de este proceso en una población o conjunto de poblaciones, así como debemos saber los valores de los coeficientes de selección  $s$  o las adecuaciones  $w$  para saber la intensidad de la selección o el tamaño efectivo para conocer la relevancia de la deriva génica (el azar) en la evolución. La tasa de migración  $m$  se define la probabilidad de que un gen tomado al azar de una población sea migrante, en otras palabras, que no sea nativo a la población y haya llegado en esa generación de otra población de la misma especie.

En este capítulo vamos a revisar los aspectos más importante en relación a la biología evolutiva y la genética de poblaciones del flujo génico y el estudio de la divergencia entre las poblaciones que forman a una especie dada.

### **La estructura genética de las poblaciones**

Las diferencias en las frecuencia alélicas entre las poblaciones constituyen la *estructura genética*. Cuando hablamos de una especie con *alta estructura genética*, nos referimos a que se pueden detectar fuertes diferencias en las frecuencias alélicas entre las poblaciones. En contraste, en una especie con *baja estructura genética*, las poblaciones que la constituyen son casi idénticas, con nulas o muy pocas diferencias en las frecuencias alélicas.

Entender las causas de la estructura genética, o sea las diferencias en las frecuencias alélicas, es la principal labor de la genética de poblaciones experimental. En términos generales, si detectamos que las poblaciones de una especies son muy distintas genéticamente (en otras palabras, que tienen una alta estructura genética) indica que los tamaños de las poblaciones han sido pequeños (o sea que se ha habido intensa deriva génica) y/o que ha existido poco flujo génico (migración) entre estas poblaciones por mucho tiempo. La selección natural también puede causar las diferencias, pero como ya mencionamos se esperaría que sólo generara diferencias en los genes relacionados con las adaptaciones o en genes ligados (muy cercanos en el cromosoma) a ellos.

Existen varios métodos para estimar la diferenciación y estructura genética poblacional, pero es importante recalcar que las diferentes medidas de estructura genética están relacionadas entre si, y simplemente se basan en analizar las diferencias en las frecuencias alélicas. Así, los métodos más sencillos simplemente comparan estadísticamente las diferencias frecuencias alélicas. Otros métodos se basan en estimar la proporción de variación genética que se encuentran dentro y entre las poblaciones, usualmente empleando el estadístico  $F_{st}$  de Wright o sus análogos, que permiten comparar de manera clara y cuantitativa las diferentes especies. Estos estimadores en principio tienen la ventaja adicional de que nos pueden dar el estimado conjunto de la importancia de la deriva génica y el flujo, la  $Nm$ , que como demostró Sewall Wright nos indica si la deriva génica es importante (si  $Nm$  es muy pequeño, usualmente menor de 1) o si el flujo génico es el proceso evolutivo determinante (i.e., si  $Nm$  es “grande”, usualmente mayor de 4). Por último, hay diferentes medidas de *distancias genética* entre pares de poblaciones. La más conocida es la distancia genética de Nei, que refleja el número de substituciones nucleotídicas a nivel ADN y se ha usado ampliamente en diferentes grupos de organismos ya que permite comparaciones de la diferenciación genética a diferentes niveles (por ejemplo, entre poblaciones, entre variedades o subespecies, o entre especies y aún entre géneros, familias, etc.). Adicionalmente, podemos usar las distancia genéticas entre pares de poblaciones para la reconstrucción de las genealogías de poblaciones, ya sea usando métodos como el UPGMA o Neighbor-joining, o usar estas distancias en análisis de aislamiento por distancia, graficando las distancias genéticas

pareada como función de la distancia geográfica, a veces con alguna transformación, y usualmente analizando la confiabilidad con una prueba de Mantel, relacionadas con el pujante campo de la filogeografía (Avice, 2000; para revisiones recientes en español ver Eguiarte et al. 2007).

A continuación vamos a discutir con mayor detalle algunos puntos conceptuales y metodológicos sobre la estructura genética, la diferenciación, así como del papel que el flujo génico en la evolución, usando para esto ejemplos de poblaciones naturales de plantas de México. Así mismo analizaremos otros puntos relacionados con la estructura genética, como la estimación de la endogamia, el grado de clonalidad y otros estudios de estructura genética a nivel finos, los cuales utilizan métodos tales como el de las autocorrelaciones espaciales.

### **Frecuencias alélicas**

El primer paso para analizar la estructura y diferenciación genética es la obtención de las frecuencias alélicas. En genes codominantes (en los que se pueden diferenciar los heterócigos de los homócigos), como isoenzimas, microsatélites, o secuencias de genes nucleares de copia única, es relativamente sencilla la estimación de las frecuencias alélicas y se desprende de los principios básicos de la genética de poblaciones, como ilustramos en Eguiarte (2009). Para detalles técnicos y las pruebas estadísticas relacionadas con su estimación se sugiere revisar textos de genética de poblaciones y evolución como Avice (1994) o Hedrick (2005).

En el caso de marcadores dominantes, como RAPDs, AFLPs e ISSRs, el problema se complica, ya que los heterócigos no pueden ser distinguidos de los homócigos dominantes. A pesar de este problema, dicho marcadores dominantes son actualmente muy populares, ya que permiten analizar directamente (a nivel ADN) grandes secciones neutras del genoma (muchos loci) a un costo muy bajo y en especies para las cuales no hay datos genómicos previos. En estos casos, para calcular sus frecuencias alélicas usualmente se considera que las poblaciones están cerca del equilibrio de Hardy-Weinberg, y se estima la frecuencia de los alelos recesivos ( $q$ ) a partir de los individuos “sin banda”, (los homócigos recesivos) como la raíz cuadrada de la frecuencia de las ausencias, esto es,  $x = q^2$ , siendo  $q = x^{1/2}$ . Por otro lado, la frecuencia de los alelos dominantes ( $p$ ) es igual a  $1 - q$  (pero ver Lynch y Milligan (1994)). Otra opción es ignorar las frecuencias alélicas y trabajar con otros estimados de la diversidad, que permiten separar a la variación genética en componentes dentro y entre las poblaciones, como el estimador de Shannon (Lewontin, 1972; ver por ejemplo Domínguez et al. 2005).

### **Diferenciación entre poblaciones**

Para detectar diferencias estadísticas significativas en las frecuencias alélicas entre las poblaciones se puede realizar diferentes pruebas estadísticas. Por ejemplo, la prueba de heterogeneidad en las frecuencias alélicas de Workman y Niswander (1970; ver también Hedrick (2005)) la cuál está basada en una Ji-cuadrad. En la palma tropical *Astrocaryum mexicanum* (Tabla 1) en Los Tuxtlas,

Veracruz, se aplicó esta prueba para analizar la diferenciación entre sitios, y se detectó mayores diferencias en las frecuencias alélicas en los adultos (diferenciación genética significativa en 4 de 5 loci ( $P$  menor de 0.02)), mientras que en las semillas fue muy baja (sólo fue significativo un loci), sugiriendo alto flujo génico en la polinización y posterior selección natural a lo largo del ciclo de vida de esta palma que genera alta diferenciación genética. O por ejemplo, en un estudio en considerando las tres especies de *Agave* más comunes del desierto Sonorense (Tabla 2) se utilizó la prueba combinada de Fisher (Sokal y Rohlf, 1995). Cuando se analizan las tres especies de manera conjunta (dado que son muy cercanas evolutivamente), se detectan diferencias en las frecuencias alélicas en todos los loci, pero esta diferenciación se reduce a sólo algunos de los loci en dos especies cuando se estudia cada una por separado: las especies son genéticamente diferentes, pero las poblaciones de una especie dada aún no han divergido suficiente, ya sea porque se mantienen cohesivas debido flujo génico, o porque has pasado muy poco tiempo y no han llegado al equilibrio diferenciación genética/flujo génico, como explicamos más adelante. En otro estudio se compararon las frecuencias alélicas en las poblaciones del maguey mezcalero *A. cupreta* en Guerrero (Eguiarte *et al.*, 2006), usando la prueba exacta de Raymond y Rousset (1995). En este casi se detectaron diferencias significativas entre poblaciones en la mayoría de los genes (23 de los 28 loci con una  $P < 0.05$ ): son diferentes entre si las poblaciones, lo cual indica que se pueden distinguir variedades genéticas en al especie, y que no se deben de plantar semillas de un lugar en otro lados (ya que son diferentes genéticamente), aunque las diferencias totales entre todas las muestras no es muy grandes ( $F_{st} = 0.11$ , más adelante explicaremos éste índice).

### **Los estadísticos $F$ de Wright**

Sewall Wright (1951) introdujo un método para partir el coeficiente de endogamia en una población subdividida ( $F_{it}$ ) entre el componente debido a apareamientos no-aleatorios dentro de poblaciones ( $F_{is}$ ) y a la subdivisión entre poblaciones ( $F_{st}$ ). Así, la endogamia total tendría un componente generado por la cruce entre parientes dentro de una población ( $F_{is}$ ) y otro por el balance entre la deriva génica y el flujo génico ( $F_{st}$ ). La definición original de Wright se basa en el coeficiente de endogamia. Así, los estadísticos  $F$  pueden ser vistos como la correlación entre genes homólogos tomados de un nivel de la subdivisión, en relación con cualquier otro nivel superior. La correlación entre los genes dentro de los individuos (I) en relación con los genes de la población total (T) es representada por  $F_{it}$ , que corresponde con la endogamia total. La correlación entre genes dentro de los individuos en relación con los de la subpoblación (S) es representada por  $F_{is}$ , mientras que la correlación entre los genes dentro de la subpoblación en relación con los de la población total está representada por  $F_{st}$ , que es igual a la probabilidad de que dos alelos idénticos por descendencia (*identical by descent*, provenientes de una población ancestral) se combinen en un cigoto.

Los estadísticos  $F$  se relacionan entre si de la siguiente manera:  
 $(1 - F_{it}) = (1 - F_{st})(1 - F_{is})$  y por lo tanto  $F_{st} = (F_{it} - F_{is}) / (1 - F_{is})$ .

La estimación de la  $F_{st}$  es más fácil de visualizar siguiendo la definición de Nei (1973) :  $F_{st} = (H_t - H_s) / H_t$ , donde  $H_t$  es el promedio de la heterocigosis esperada en la población total, para todos los loci y  $H_s$  es el promedio de la heterocigosis esperada dentro de subpoblaciones para todos los loci.  $F_{st}$  mide la reducción en la heterocigosis debida a diferenciación genética entre poblaciones.

También se puede definir la  $F_{st}$  en términos de las varianzas en las frecuencias alélicas entre las poblaciones, lo cual puede ser más intuitivo:  $F_{st} = \text{Var}(p) / (p(1-p))$ , entre mayor haya sido la deriva génica, mayor será la varianza (las diferencias entre) las frecuencias alélicas entre poblaciones.

Adicionalmente, es importante mencionar que en la literatura existen variantes en la estimación de la  $F_{st}$  que se mencionan en el Recuadro 1.

La  $F_{st}$  tiene varias características que la han hecho atractiva para los biólogos evolutivos. En primer lugar, su interpretación es relativamente sencilla: Si la  $F_{st} = 0$ , indica que las frecuencia alélicas son idénticas en todas las poblaciones estudiadas, o sea que aún no ha habido nada de diferenciación entre ellas (i.e.,  $H_t = H_s$ ). El máximo posible en la  $F_{st}$  es de 1, al que se llega cuando cada población esta fija en alelos diferentes (i.e.,  $H_s=0$ ), o sea que son completamente diferentes las frecuencias alélicas de las poblaciones que estudiamos.

Otra ventaja de la  $F_{st}$  es que, como ya mencionamos, a partir de la ella se puede obtener una aproximación de  $Nm$  si se ha llegado al equilibrio entre los procesos de la deriva génica y la migración.

### Valores de $F_{st}$ en poblaciones naturales

La  $F_{st}$  se ha calculado en gran cantidad de organismos, por lo cual se facilita su comparación. En la Tabla 1 mostramos algunos valores en plantas mexicanas.

Por ejemplo, el primer estudio donde se evaluó la  $F_{st}$  en México, fué en las poblaciones silvestres del frijol ayocote (*Phaseolus coccineus*) del Centro de México. En esta planta se encontró una  $F_{st} = 0.203$ , que representa un valor intermedio para una herbácea: hay bastante diferenciación entre las poblaciones, posiblemente debido a que sus tamaños efectivos son relativamente pequeños y que sus genes se mueven poco, ya sea porque hay baja dispersión de sus semillas o de su polen. Podemos comparar esta  $F_{st}$  con los valores de otras especies, como los árboles tropicales (Tabla 1), que tienen un rango que va desde valores muy pequeños de diferenciación, como  $F_{st} = 0.026$  en *Psychotria faxlucens* y 0.029 en *Cecropia obtusifolia* (ambos de la región de Los Tuxtlas, Veracruz), que indican mucho flujo génico y poca deriva génica, a valores similares, como de 0.23 en el mangle rojo (*Rhizophora mangle*, datos de ambas costas de México) o aún más altos, como 0.51 en *Antirhea aromatica*, una planta microendémica en peligro de extinción que seguramente tiene muy pequeños tamaños efectivos y poca dispersión de sus semillas y polen.

En la Tabla 3 mostramos para cuatro especies de oyameles (*Abies*) mexicanos las estimaciones de diferenciación genética usando diferentes marcadores genéticos (nucleares (isoenzimas) y cloroplasto (microsatélites)) y

dos índices ( $F_{st}$  y la  $R_{st}$ , ver Recuadro 1) y el método de los alelos privados (ver Recuadro 2). Las diferenciación en general es entre baja a moderada, como usualmente se encuentran en árboles, y similares entre las especies, lo cual sugieren elevado flujo génico entre las poblaciones que forman cada especie. Las  $F_{st}$  nucleares tienen a ser más elevadas que las de cloroplasto, lo cual hace sentido, ya que el cloroplasto se dispersa con el polen, el cual se puede mover mucho en estas especies polinizadas por el viento. En general, las estimaciones de diferenciación son parecidas entre los dos marcadores (nucleares y cloroplasto) y en el cloroplasto con los dos métodos (excepto en *A. flinckii* y *A. religiosa*, donde la  $F_{st}$  en cloroplasto parece muy baja, comparada con la nuclear y la  $R_{st}$  de cloroplasto).

### **Cuidado en el uso e interpretación de la $F_{st}$ y otros métodos relacionados**

Consideramos importante señalar que hay tener cuidado al interpretar la  $F_{st}$  en comparaciones entre especies. Como hicimos en la sección anterior, ya que el muestreo en distintos estudios pueden ser muy contrastante. Por ejemplo, algunos estudios se realizan en una sola localidad, comparando partes de una población o manchones de individuos aislados por metros, como es el caso de los estudios de árboles tropicales en Los Tuxtlas que mencionamos arriba. Por otro lado, en algunos estudios se comparan poblaciones separadas por cientos a aún miles de kilómetros, como el estudio del mangle rojo o de los *Abies* que indicamos en los párrafos anteriores. Ciertamente, aunque la diferenciación genética depende de los detalles de biología e historia de cada especie, en los estudio que contemplan una menor área geográfica, se debe esperar menor diferenciación (medida como  $F_{st}$ ) que en estudios que abarcan una área muy grande de cientos o miles de kilómetros.

Para marcadores dominantes, se ha sugerido que es mejor realizar la partición de la varianza del AMOVA con los haplotipos (ver Recuadro 1), como hicieron Peakall et al. (1995) con muestras Tejanas y Mexicanas del pasto *Buchloe dactyloides*, o usar estimaciones de diversidad independientes de una estimación de la frecuencias alélicas (usando por ejemplo las estimaciones de diversidad genética usando el índice de Shannon que mencionamos arriba), como llevaron a cabo Domínguez et al. (2005) en un árbol pequeño de las selvas secas de la región de Chamela, *Erythroxylum havanense*, pariente cercano de la planta de la coca. En este último estudio, estimaron el equivalente a la  $F_{st}$  de 0.094, mientras que un AMOVA indicó una diferenciación entre poblaciones del 0.137. En otras palabras, del total de la variación genética que tiene la especie, entre el 9 y el 13.7% se debe a diferencias entre las poblaciones (mientras que entre el 91.6 al 86.3% de la variación se encuentra dentro de las poblaciones) O por ejemplo, en el estudio mencionado arriba de las cuatro especies de oyamel (*Abies*), Aguirre-Planter (2005) estimó la varianza a diferentes niveles usando microsatélites de cloroplasto, y encontró con un AMOVA que el 5.50% se debía a las diferencias entre las poblaciones de una misma especie (este es el equivalente a la  $F_{st}$ ) y que solo 5.66% de la variación genética se podía adscribir a diferencias entre especies (o sea, las especies son muy parecidas entre sí, tal vez porque se originaron hace muy

poco), mientras que la mayor parte la variación (88.4%) se encontraba dentro de las poblaciones, similar a los describimos arriba para *Erythroxylum havanense*. Estos valores en los oyameles son congruente con los análisis de genes nucleares (isoenzimas), donde sólo el 5.6% de la variación se debía las diferencias entre especies (Aguirre-Planter et al. (2000); ver también Tablas 1 y 3 y Jaramillo-Correa et al. (2008) para otros estudios similares).

### Estimaciones de endogamia: $F_{is}$

Asociado a las  $F_{st}$ , cuando los marcadores son codominantes generalmente se calculan los otros índices  $F$  de Wright, (si son dominantes, no se pueden calcular estas otras medias, ya que se requiere conocer la heterocigosis observada, la heterocigosis real), siendo el más relevante la  $F_{is}$ .

En términos generales, podemos definir que  $F_{is} = (H_s - H_o)/H_s$ , donde  $H_s$  es la heterocigosis promedio por subpoblación, y la  $H_o$  es la heterocigosis observada. Si se calcula para una sola población, la  $F_{is}$  corresponde al índice de fijación o de endogamia,  $F$  (Hedrick, 2005). Para revisiones sencillas en español sobre la biología de la endogamia, ver Eguiarte y Piñero, 1999 o Eguiarte et al., 1999).

La  $F_{is}$  describe la distribución de los genotipos dentro de las poblaciones, y nos indica qué tan lejos se encuentra una población del equilibrio de Hardy-Weinberg (i.e., las frecuencias de genotipos que se obtendrían si los apareamientos son al azar). La  $F_{is}$  es de 0 si las poblaciones se encuentran en las proporciones esperadas si los apareamientos fueran estrictamente al azar (Hardy-Weinberg), y puede llegar hasta 1 si sólo se encuentran individuos homocigos en la población (lo cual generalmente es una consecuencia de la endogamia extrema). Si la  $F_{is}$  es negativa, quiere decir que se tiene un exceso de heterocigos. Por ejemplo, en *Astrocaryum mexicanum* (Tabla 1) se estimó una  $F_{is}$  promedio en adultos de -0.42, y en semillas de -0.2. En ambos casos es negativa, pero el exceso de heterocigos es mayor en los adultos, posiblemente como consecuencia de que los individuos heterocigos funcionan y sobreviven mejor que los homocigos (o sea, por selección natural a favor de los heterocigos).

Otros valores de  $F_{is}$  se muestran en la Tabla 1, donde tenemos especies con valores altos y positivos, que indican una alta endogamia, resultado de una autofecundación regular, como el frijol común, *P. vulgaris* ( $F_{is} = 0.6$ ), o por ejemplo el mangle rojo, que tiene una  $F_{is} = 0.45$ , o el cactus endémico de las costas de Baja California, *Stenocereus eruca*  $F_{is} = 0.739$ , o algunas compuestas en el Pedregal de San Angel, que llegan a una  $F_{is} = 0.895$ .

Por otra parte, conviene indicar que el principal determinante de la endogamia en las poblaciones vegetales es la probabilidad de que una semilla sea el resultado de polinización cruzada (*outcrossing rate*),  $t$  (para una discusión en español, ver Eguiarte et al. 1999). Indirectamente, si toda la endogamia se debe a autofertilización se puede calcular la  $t = 1 - F / 1 + F$ .

### Flujo génico



Una vez definidas las medidas básicas de diferenciación genética, podemos regresar a discutir con cuidado la biología evolutiva del flujo génico o migración. Fundamentalmente, el flujo génico consiste en el movimiento de genes de una población a otra. Puede generarse por el movimiento de gametos, de semillas o de individuos. O pensando en genomas, puede deberse a la incorporación a una población de genes nucleares o de genomas uniparentales, como el de la mitocondria o el cloroplasto, que generalmente son heredados sólo por la vía materna. También puede referirse a la extinción y recolonización de poblaciones enteras (procesos llamados de metapoblaciones) (Slatkin, 1985a, ver Hedrick 2005 para una discusión actualizada en términos de genética de poblaciones).

Durante un tiempo se pensaba que el flujo génico en general era muy restringido y de poca importancia evolutiva, y que la selección natural era la fuerza evolutiva que mantenía unificadas (cohesivas) a las especies (ver por ejemplo la revisión de Levin (1981) y discusiones en Futuyma (1998). Sin embargo, al estimar con métodos moleculares y ecológicos experimentalmente el flujo génico en distintas especies, se ha observado que los niveles pueden ser bastante altos y que de esta manera la migración pueden actuar como una fuerza que mantiene cohesionadas a las especies.

El flujo génico es un componente fundamental de la estructura poblacional, ya que sus patrones y niveles determinan hasta qué grado cada población de una especie es (o no) una unidad evolutiva independiente y la tasa de movimiento de genes de una población a otra puede determinar si la especie evoluciona al azar (por deriva génica) o responde de manera adaptativa (por selección natural) y así puede afectar de manera importante a la evolución de las especies (Futuyma, 1998; Eguiarte, 2009). De esta manera, el flujo génico puede determinar la persistencia y adaptación de poblaciones locales, las tasas de extinción de las poblaciones y especies, la evolución de los rangos de adaptación y distribución de las especies y de muchas propiedades ecológicas. Si el flujo génico entre poblaciones de una especie es alto, entonces todas las poblaciones evolucionan de manera conjunta, pero si es muy bajo, las poblaciones de una especie empiezan a divergir y pueden evolucionar casi independientemente. Si continua la diferenciación, puede llegar a surgir aislamiento reproductivo y en consecuencia el establecimiento de linajes evolutivamente independientes (o sea, se da un proceso de especiación). Qué tanto flujo génico es necesario para prevenir la evolución independiente de las poblaciones de una especie dependerá de la intensidad de las otras fuerzas evolutivas (Hedrick, 2005; Eguiarte, 2009).

Una medida que resulta fácil de calcular, como veremos más adelante, y que es conceptualmente muy útil es la  $Nm$ , que es la multiplicación del tamaño efectivo, ( $N_e$ ) por la tasa de migración ( $m$ ), y nos habla del número de *migrantes efectivos*, o sea del número de organismos que llegan a una población y se incorporan a su poza génica, que es una cantidad interesante, ya que es un número absoluto, y no necesitamos conocer el tamaño efectivo, que puede ser muy complicado de calcular. Si  $Nm$  es alto (generalmente se considera 4, como vamos a ver adelante), el flujo génico supera los efectos de la deriva génica y previene la diferenciación local. Si  $Nm$  es pequeño (menor a 1), entonces se

puede decir que la deriva actúa independientemente en cada una de las poblaciones.

Se han utilizado distintos métodos directos e indirectos para tratar de medir estas tasas de flujo génico que revisaremos brevemente a continuación.

### **Métodos directos para estimar el flujo génico**

Se basan en observaciones o experimentos que miden el grado de dispersión de gametos o individuos para tratar de estimar a partir de estos datos ecológicos y demográficos directamente la tasa de migración ( $m$ ). Por ejemplo, se puede estudiar la dispersión de individuos marcados con etiquetas, anillos o radiotransmisores o de sus gametos (por ejemplo de granos de polen) marcados (el polen se puede marcar con polvos fluorescentes u otros tintes o radiactivamente, por ejemplo). Métodos alternativos usando marcadores moleculares han dado en parte origen al floreciente campo de la ecología molecular (Eguiarte et al. 2007), donde ya es rutinario realizar análisis de parentesco con marcadores variables que permiten identificar a los padres y después cuantificar el patrón del movimiento y los niveles de dispersión de los genes. Por ejemplo, en su estudio pionero con monocotiledónea *Chamaelirium luteum*, Meagher (1986) pudo cuantificar la varianza en el éxito reproductivo como una función de la distancia entre individuos reproductivo a partir de datos de análisis de paternidad. Modificaciones subsecuentes a estas ideas han sido utilizados para el estudio detallado del movimiento de genes en poblaciones (Devlin y Ellstrand, 1990; Roeder et al., 1989; Smouse y Meagher, 1994). Análisis detallados de parentesco nos permiten estimar la distribución de las distancias de dispersión y examinar el movimiento de genes por polen y semilla dentro de una población.

Las estimaciones directas obtenidas con datos ecológicos son una importante herramienta para predecir y corroborar otras estimaciones directas experimentales y las indirectas inferidas a partir de las distribución de las frecuencias alélicas. De cualquier forma, vale la pena mencionar que las predicciones de las medidas directas y las estimaciones indirectas pueden diferir por varias razones. Las medidas directas estiman a la migración con una muestra (que puede ser muy pequeña y sesgada) dentro de una área determinada, misma que no es necesariamente representativa de una población. Por otra parte, en los experimentos ecológicos usualmente no se sabe si el migrante se reproduce exitosamente y obviamente estas medidas directas no pueden ser entendidas directamente como flujo génico a largo plazo, ya que sólo son valores del flujo génico que ocurre en el periodo en el que se midió, generalmente uno o pocos eventos reproductivos (Whitlock y McCauley, 1999), y subestiman la frecuencia de la dispersión a larga distancia (ya que estas pueden ser muy difíciles de detectar, pero pueden ser importantes), y no estiman la importancia de extinciones y recolonizaciones como una fuente de flujo génico (Slatkin, 1985a),

### **Métodos indirectos para estimar el flujo génico**

Los métodos indirectos se basan en el análisis de la distribución espacial de alelos en las poblaciones y de ésta manera se hacen inferencias de los niveles o patrones de flujo génico en las poblaciones usando diferentes modelos, algunos de los cuales se ilustran en las Figs. 1 y 2 (Slatkin, 1985a).

La mayoría de los modelos teóricos de flujo génico surgen de los conceptos desarrollados por Sewall Wright, basados en poblaciones continuas y utilizando un enfoque de aislamiento por distancia o en poblaciones que funcionan como islas que se diferencian por mutación y deriva génica (Wright, 1943). El modelo usado comúnmente para estimar flujo génico es el modelo de islas infinitas (*infinite island model*) de Wright (1951) (ver Fig. 2a). Este modelo considera condiciones en equilibrio, o sea que tan diferentes son las frecuencias alélicas que alcanzan después de mucho tiempo un grupo de poblaciones en las que la deriva génica elimina variación genética y el flujo génico la incrementa; si se mantiene las tasas de flujo génico y el tamaño efectivo, después de un tiempo se llega a un equilibrio dinámico. El modelo inicial considera un número muy grande (infinito) de islas o poblaciones toda con exactamente el mismo tamaño poblacional y que intercambian migrantes entre cualquiera de las islas con la misma probabilidad a una tasa de migración  $m$  que no cambia en el tiempo. Las poblaciones pueden ser tratadas como réplicas, y el modelo se puede caracterizar con sólo dos parámetros: el tamaño poblacional ( $N$ ) y tasa de migración ( $m$ ). La importancia de la deriva génica es proporcional a  $1/N$ , mientras que la importancia del flujo génico es proporcional a  $m$  (Slatkin, 1985a).

Otro modelo de flujo génico el las piedras de paso (*stepping-stone*), introducido por Kimura (1953), que ilustramos en las Figs. 2 b y c. En este modelo las poblaciones se localizan en una especie de enrejado de una, dos o tres dimensiones y los individuos sólo pueden moverse entre poblaciones adyacentes, o sea el flujo sólo puede ser entre las poblaciones cercanas en una generación, pero los resultado del equilibrio, aunque análogos a los del modelo de Wright, son más complicados y menos generales.

Para estimar el flujo génico con un método indirecto usualmente se emplea la fórmula de Wright (1951) cuando la  $F_{st}$  llega al equilibrio entre deriva génica y migración en el modelo de islas infinitas:  $F_{st} \approx 1 / (4 Nm + 1)$ ; despejando,  $Nm \approx (1/F_{st} - 1) / 4F_{st}$ . Bajo este modelo, Wright (1969), encontró que una tasa de migración  $>1$  cada generación es suficiente para contrarrestar la diferenciación genética debida a deriva génica (en la Fig. 3 se ilustra el equilibrio para  $Nm = 5$  y  $50$ ), o sea si  $Nm$  es mayor de 1, la deriva génica es despreciable y el flujo génico hace que todas poblaciones evolucionen de manera cohesiva como un conjunto, mientras que si es menor de 1, la deriva génica juega un papel cada vez más importante, hasta que domina todo el proceso, como ilustramos en la Fig. 3 para  $Nm = 0.5$  y  $Nm = 0.05$ .

Por ejemplo, en la Tabla 2 se estimó la  $Nm$  en *Agave* en el desierto de Sonora con dos métodos, el descrito arriba partir de la  $F_{st}$  y otro método que se describe en el Recuadro 2 (Alelos privados). En todas las estimaciones  $Nm$  fue mayor de 1, y llego hasta 6.14, y en general fue más alta dentro de cada una de las especies que en total (las tres especies juntas). Esto indica que el flujo génico es una fuerza importante en estas especies, lo suficiente alta para hacer

que la deriva génica (y así el azar) no sea una fuerza evolutiva relevante, y que obviamente, hay más flujo dentro de una especie que entre especies (aunque debido al alto valor de  $Nm$ , en el artículo se especula que tal vez no sean tres especies, sino solo una especie con diferenciación geográfica). El caso de los oyameles (*Abies*, Tabla 3), los estimadores para los genes nucleares indican menor flujo génico que en el caso anterior, ya que las estimaciones de  $Nm$  a partir de la  $F_{st}$  fueron de entre 0.75 a 3.17, y con el método de los Alelos privados fueron de entre 1.67 a 3.42, sugiriendo que en algunos casos la deriva génica podría tener un papel en la evolución de estas especies, al obtenerse valores menores y cercanos a 1. Para el cloroplasto, los estimadores de flujo génico son más altos, como se podría esperar debido a que en estas especies el cloroplasto se mueve con el polen, y como comentamos arriba son polinizadas por el viento, aunque la dispersión de los estimadores es muy grande, ya que van de entre 0.136 a 16.74, y sugieren que el estimador  $R_{st}$  a veces no es adecuado (ver Recuadro 1).

Considerando a la mutación y un número finito de islas (*N-island model*), Crow y Aoki (1984) encontraron que en el equilibrio, el equivalente con alelos múltiples de la  $F_{st}$  ( $G_{st}$  de Nei, 1973, ver Recuadro 1),  $G_{st} \approx 1 / (4Nm\alpha + 1)$  donde  $\alpha = [n / (n-1)]^2$ , y  $n$  es el número de subpoblaciones. Esta corrección es relevante si se estudian pocas poblaciones; si son muchas las poblaciones analizadas, ambas estimaciones (i.e., Wright (1951) y Crow y Aoki (1984)) dan valores similares. Slatkin y Barton (1989) demostraron que esta estimación de la  $Nm$  a partir de la  $F_{st}$  es uno de los métodos más robustos para inferir indirectamente la tasa de migración y la migración efectiva (comprada con otros métodos como el uso de alelos raros o únicos, ver Recuadro 2) y por eso es el método que generalmente se utiliza para aproximar indirectamente la  $Nm$ .

No hay que olvidar que los valores calculados de  $Nm$  dependen directamente de nuestra estimación de la  $F_{st}$  y la estimación final de  $Nm$  reflejará los errores que haya en la estimación de la  $F_{st}$ ; por esta y otras razones, la utilidad de estos métodos ha sido severamente cuestionada por autores (Whitlock y McCauley, 1999; Neigel, 2002). Sin embargo, la mayor parte de los genetistas de poblaciones las consideramos aproximaciones sencillas y útiles (ver discusión en Hedrick, 2005).

### ¿En cuanto tiempo la $F_{st}$ llega al equilibrio?

Este es un problema relevante para nuestras inferencias y para la filogeografía, por lo que lo mostramos en la Fig. 4. Si se tarda mucho tiempo en llegar al equilibrio, es posible que las poblaciones que estamos estudiando no hayan llegado al equilibrio, y que sean en este momento más parecidas genéticamente de lo que van a ser una vez que lleguen al equilibrio, y así nuestra estimación de  $Nm$  sugeriría tamaños efectos y/o tasa de migración mayores que las reales. En términos generales, se llega al equilibrio en unas 150 generaciones si la tasa de migración es relativamente alta ( $m=0.01$ ) y el tamaño efectivo es pequeño ( $N=100$ ) (Fig. 3, en color rosa), mientras que si la tasa de migración es más pequeña ( $m=0.001$ ) y el tamaño efectivo mayor ( $N=1,000$ ) tomará un orden de magnitud más, o sea unas 1,500 generaciones, en llegar al equilibrio, como se

ve en color verde en la Fig. 3 (Crow y Aoki, 1984; Hedrick, 2005). Los casos en los cuales el tamaño efectivo es muy grande y la tasa de migración muy pequeña, la  $F_{st}$  va a dar sobrestimaciones del flujo génico en la mayoría de los casos (porque la  $F_{st}$  es aún muy pequeña, ya que no se ha llegado al equilibrio). Posiblemente este sea el caso por el que muchas poblaciones tienen valores de  $F_{st}$  mucho más bajos de lo esperable según su ecología e historia natural. Estos casos de incongruencia, muy frecuentes por ejemplo en zonas templadas y a altas latitudes, seguramente tienen que ver con eventos de colonización reciente (que posiblemente han sucedido después de las últimas glaciaciones).

### **Otros métodos para estimar indirectamente el flujo génico**

Otras maneras relativamente populares de estimar el flujo génico ( $Nm$ ) y la diferenciación es a través del método de los Alelos Privados de Slatkin (1981, 1985b) y de las Distancias Genéticas entre pares de poblaciones, que se discuten en los Recuadros 2 y 3, respectivamente, y se ilustran en los ejemplos de las Tablas 2 y 3 y en la Figura 5.

### **Aislamiento por distancia**

Cuando una especie presenta una distribución geográfica amplia, generalmente se espera que las poblaciones más cercanas geográficamente también sean las más parecidas genéticamente, ya que se supone que el flujo génico debe ser más probable entre las poblaciones contiguas que entre las más lejanas. A esta idea se le conoce como “aislamiento por distancia” (Wright, 1943) y nos dice que los individuos tienden a aparearse más con aquellos más cercanos geográficamente de lo que se esperaría si los cruzamientos fueran al azar en toda la distribución de la especie.

Con el fin de conocer si una especie sigue un modelo de aislamiento por distancia, es decir, si existe una correlación entre las distancias geográficas y las distancias genéticas (ver Recuadro 3) entre pares de poblaciones, se pueden llevar a cabo análisis estadísticos formales como la prueba de Mantel (1967, ver Fig. 5): Si existe una correlación positiva significativa entre la distancia geográfica y la genética, existe evidencia de aislamiento por distancia, esto es, que el flujo génico es limitado y localizado, lo cual origina que las poblaciones más cercanas en el espacio sean más cercanas y tengan más flujo génico que las más alejadas.

La forma específica de graficar los datos en un análisis de aislamiento por distancia puede permitir hacer entender mejor los posibles modelos finos de estructura genética que siguen las poblaciones y hacer inferencias sobre la magnitud de su tasa de migración (Slatkin, 1994; Rousset, 2001). Por ejemplo, la distancia genética usada puede ser directamente la distancia genética de Nei, como hizo Navarro-Quezada et al. (2003, su Fig.5), o la estimaciones de  $Nm$  pareadas ( $M$ ) en logaritmo (según recomienda Slatkin (1993, 1994)), como analizaron Aguirre-Planter et al. (2000), o seguir la sugerencia de Rousset (2001) y graficar  $F_{st} / (1 - F_{st})$ .

En la Fig. 5 ilustramos los análisis usando 90 loci de ISSRs de la tesis de maestría de Enrique Scheinvar (2008) para un conjunto de 15 poblaciones del complejo *Agave cupreata* y *A. potatorum* del Michoacán, Guerrero, Oaxaca y Puebla. Cada punto representa una comparación entre dos de las poblaciones. En la Fig. 5 a. mostramos las distancias genéticas de Nei (1972, ver Recuadro 3) y en la Fig. 5. b. las estimaciones de  $Nm$  a partir de la  $F_{st}$ . Es claro que entre mayor es la distancia geográfica, estadísticamente aumenta la distancia genética de Nei y como consecuencias disminuye la  $Nm$ , aunque también hay mucha dispersión en los puntos, lo cual indica que hay poblaciones cercanas con poco flujo génico y otras más lejanas con niveles de flujo relativamente elevados, lo cual sugiere que la orografía y la ecología también son importantes en determinar los patrones de flujo y dispersión, no solo la distancia espacial. En términos generales, en estos agaves las poblaciones a menos de 600 km tienen una  $Nm$  mayor de 1, y como no hay poblaciones a mayor distancia, los modelos básicos de Wright que vimos arriba indican que todas las poblaciones tienen suficiente flujo génico para no divergir entre ellas por deriva génica, permitiendo la acción efectiva de la selección natural al generar muy altos tamaños efectivos (Hedrick, 2005; Eguiarte, 2009).

### Estructura genética fina

Dentro de una población se puede estimar el grado de diferenciación entre los individuos que la forman con diferentes metodologías. Lo más sencillo es tomar diferentes muestras dentro de una población continua, como se hizo en Eguiarte et al. (1992, 1993) en *Astrocaryum mexicanum* en Los Tuxtlas, donde se tenían mapeados a todos los organismos analizados, y se hicieron inferencias sobre el flujo génico, la vecindad genética y el tamaño efectivo de manera directa e indirecta. En *Cecropia obtusifolia*, también en Los Tuxtlas, Epperson y Alvarez-Buylla (1997) analizar las auto-correlaciones espaciales en las frecuencias alélicas. Más recientemente, Domínguez et al. (2005) analizaron la estructura especial fina en una población de *Erythroxylum havanense*, usando el método de autocorrelación de Smouse y Peakall (1999) y encontraron que hay una estructura fina significativa a distancias muy pequeñas: a distancias menores de 5 metros la correlación genética es positiva, sugiriendo endogamia y poca dispersión, para volverse negativa a unos 35 m de distancia.

También se ha estudiado con método análogos la clonalidad en la cactácea endémica en peligro de extinción *Steneocereus eruca* de Baja California (Clark-Tapia et al., 2005), donde se mapearon los genotipos con 57 loci de RAPDs y se encontró un cociente genotipos / total de individuos analizado ( $G/N$ ) = 0.83, en otras palabras, la planta, aunque con clonal, presenta una diversidad grande de genotipos. Alfonso-Corrado et al. (2004) analizaron con RAPDs dos especies de encinos, *Quercus eduardii* y *Quercus potosina*, y realizaron análisis de autocorrelaciones y encontraron que aunque clonales, hay bastante diversidad de genotipos, (*Q. eduardii*:  $G/N=0.60$ ; *Q. potosina*:  $G/N=0.65$ ) y presentan bajos niveles de diferenciación genética (*Q. eduardii*  $F_{st}$  = 0.19, *Q. potosina*  $F_{st}$  = 0.13). O por ejemplo, recientemente Celeste Rives (2009) llevó a cabo un detallado estudio en dos poblaciones de *Agave striata* en

Metztitlán usando ISSR, descubriendo que aunque hay un grado de clonalidad, se detectan muchos más genotipos que lo que se había especulado previamente para la especie.

### **Paquetes y programas para el estudio del flujo génico y la estructura genética**

El uso de diferentes métodos para analizar la estructura genética sigue ciclos que parecen seguir más modas que tener bases científicas razonadas. En muchos casos el uso de un método particular depende de que exista un algoritmo sencillo o un programa de computadora accesible y disponible para su cálculo. Actualmente existen varios paquetes computacionales populares y de relativo fácil uso. En la Tabla 4 se da una lista de algunos programas que ha probado ser útiles y sitios de Internet con información útil y consideramos que es un buen punto de entrada para los alumnos interesados en aplicar estos métodos a sus datos, sin pretender ser de ninguna forma una tabla exhaustiva de todos los recursos disponibles en estos formatos.

### **Patrones y perspectivas**

En la Tabla 1 se presentan los valores de variación genética ( $H$ ) y los estadísticos  $F$  para varios estudios de plantas Mexicanas, tanto angiospermas como coníferas. En general, el rango de diferenciación, aún dentro de un grupo de historias de vida, formas de crecimiento o grupo filogenético es muy amplio. Algunos grupos presentan muy poca diferenciación genética, como es el caso de las poblaciones de la biznaga *Echinocactus platyacanthus* en Tehuacán, el de algunas cactáceas columnares del desierto de Sonora, o de ciertos árboles tropicales, mientras que otras especies tienen niveles muy altos de diferenciación genética, destacando algunas especies en géneros de árboles como *Pinus* o algunos árboles tropicales (*Anthirrea aromatica*, *Rhizophora mangle*, *Bursera spp.*), que en otras partes del mundo generalmente tienen valores muy bajos de  $F_{st}$ . Estos estudios, acoplados a futuros estudios detallados de flujo génico y marcadores uniparentales que nos hablen de los patrones diferenciales de flujo génico por semilla y por polen, junto con simulaciones de coalescencia, datos de alta resolución de genes particulares a nivel de secuencia de ADN y análisis estadísticos robustos, nos ayudarán a entender los tiempos de origen de las plantas mexicanas, sus patrones y causas de su especiación, para así definir estrategias para su manejo y conservación genética.

### **Agradecimientos**

Este trabajo se escribió durante una estancia sabática de LEE y VS en la Universidad de California en Irvine (UCI), bajo la dirección de Brandon Gaut, con apoyo de la fundación UC-Mexus, del Conacyt y la DGAPA, UNAM. Los estudios descritos se han realizado con diferentes apoyos de la DGAPA, UNAM (en particular Papiit IN224309), Conacyt (en particular SEP-2004-C01-46475-Q), SEMARNAT- CONACYT (2002 C01- 0246) y Conabio (CS016).

## Referencias citadas en el texto (incluyendo figuras y recuadros)

- Aguirre Planter, E. 2005. Historia evolutiva de *Abies guatemalensis*, *A. religiosa* y *A. hickeli*, especies con distribuciones y tamaños poblaciones contrastantes: estructura genética, flujo génico y conservación. Tesis de doctorado. Instituto de Ecología, UNAM, D.F., México.
- Aguirre-Planter E., G. R. Furnier y L. E. Eguiarte. 2000. Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany* 87: 362-371.
- Alfonso-Corrado C, Esteban-Jimenez R, Clark-Tapia R, Piñero D, Campos J, Mendoza A. 2004. Clonal and genetic structure of two Mexican oaks: *Quercus eduardii* and *Quercus potosina* (Fagaceae). *Evolutionary Ecology* 18: 585-599.
- Avise, J. 1994. Molecular markers. Natural History and Evolution. Chapman & Hall, New York, USA, 511 págs.
- Avise, J. C. 2000. Phylogeography ; the history and formation of species Harvard University Press, Cambridge, Mass. USA, 447 págs.
- Clark-Tapia, R., C. Alfonso-Corrado, L.E. Eguiarte y F. Molina-Freaner. 2005. Clonal diveristy and disitribution in *Stenocrreus eruca* (Cactaceae), a narrow endemic cactus of Sonora desert. *American Journal of Botany*. 92: 272-278.
- Crow, J.F. y K. Aoki. 1984. Group selection for a polygenic behavioural trait: estimating the degree of population subdivision. *Proc. Natl. Acad. Sci. USA*. 81: 6073-6077.
- Devlin, B. y N.C. Ellstrand. 1990. The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* 44: 248-259.
- Domínguez, C.A., C. Abarca, R. Cueva, F. Molina y L. E. Eguiarte. 2005. Local genetic differentiation among populations of the mass-flowering shrub *Erythroxylum havanense* (Erythroxylaceae). *New Phytologist* 166: 663-672.
- Eguiarte, L.E. 2009. Nueva guía para principiantes a la genética de poblaciones, en Morrone J.J. y P. Magaña (editores). *Evolución Biológica*. Facultad de Ciencias, UNAM, D.F., México, págs. 83- 102.
- Eguiarte, L., Pérez Nasser, N. y Piñero, D. 1992. Genetic structure, outcrossing rate and heterosis in *Astrocaryum mexicanum* (tropical palm): Implications for evolution and conservation. *Heredity* 69: 217-228.
- Eguiarte, L. E., A. Búrquez, J. Rodríguez, M. Martínez-Ramos, J. Sarukhán y D. Piñero. 1993. Direct and indirect estimates of neighborhood and effective population size in a tropical palm, *Astrocaryum mexicanum*. *Evolution* 47: 75-87.
- Eguiarte, L. E., J. Nuñez-Farfán, C. Domínguez y C. Cordero. 1999. Biología evolutiva de la reproducción en plantas. En Nuñez-Farfán J. y L. E. Eguiarte (editores) *La Evolución Biológica*. Facultad de Ciencias, Instituto de Ecología, UNAM, CONABIO, D.F., México págs. 117-152.



- Eguiarte, L. E. y D. Piñero. 1999. Genética de la conservación: leones vemos, genes no sabemos, en Nuñez-Farfán J. y L. E. Eguiarte (editores) La Evolución Biológica. Facultad de Ciencias, Instituto de Ecología, UNAM, CONABIO, D.F., México, págs. 371-398.
- Eguiarte Fruns, L. E., A. González González y E. Sheinvar Gottdiener. 2006. Genética de poblaciones en viveros de *A. cupreata* e impactos de los planes de manejo en la diversidad y estructuración de esta especie” Proyecto Conabio CS016, Abril 2006. D.F., México.
- Eguiarte, L.E., V. Souza y X. Aguirre, compiladores. 2007. Ecología Molecular. INE, SEMARNAT, CONABIO, UNAM. México, D.F., México, 592 págs.
- Epperson B. y E. Alvarez-Buylla. 1997. Limited seed dispersal and genetic structure in life stages of *Cecropia obtusifolia*. *Evolution* 51: 275-282.
- Excoffier, L. 2001. Analysis of population subdivision. En: Baldwin D.J., Bishop M. y Cannings L. (edits.) *Handbook of statistical genetics*. John Wiley & Sons. Chichester, U.K., págs. 271-307.
- Gentry, H.S. 1982. *Agaves of continental North America*. The University of Arizona press. Tucson, Arizona, USA. 670 págs.
- Good-Avila, S.V., V. Souza, B. S. Gaut y L. E. Eguiarte. 2006. Timing and rate of speciation in *Agave* (Agavaceae). *Proc. Natl. Acad. Sci. USA* 103: 9124-9129.
- Futuyma, D.J. 1998. *Evolutionary Biology*. Third edition. Sinaure Associates, Inc. Sunderland Massachusetts. 763 págs.
- Hedrick, P.W. 2005. *Genetics of populations*. Third edition. Jones and Bartlett publishers. Sudbury, Massachusetts. 737 págs.
- Jaramillo-Correa, J. P. , E. Aguirre-Planter, D. P. Khasa, L. E. Eguiarte, D. Piñero, G. R. Furnier, J. Bousquet. 2008. Ancestry and Holocene divergence of subtropical montane forest isolates: molecular biogeography of the genus *Abies* (Pinaceae) in southern México and Guatemala. *Molecular Ecology* 17: 2476–2490.
- Kimura, M. 1953. “Stepping-stone” model of population. *Ann. Rep. Natl. Inst. Genet. Japan* 3: 62-63.
- Levin, D. A. . 1981. Dispersal versus gene flow in plants. *Ann Missouri Bot. Gard.*, 68: 233-253.
- Lewontin, R. 1972. The apportionment of human variation. *Evol. Biol.* 6: 381-398.
- Lynch, M. y B. Milligan. 1994. Analysis of population genetic-structure with RAPD markers. *Molecular Ecology* 3 : 91 199.
- Mantel, N. 1967. The detection of disease clustering and generalized regression approach. *Cancer Research*, 27, 209-220.
- Meagher, T.R. 1986. Analysis of paternity within a natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents. *The American Naturalist* 128: 199-215.
- Navarro-Quezada, A., González, R., F. Molina-Freaner y L. E. Eguiarte. 2003. Genetic differentiation in the *Agave deserti* (Agavaceae) complex in the Sonoran Desert. *Heredity* 90: 220-227.
- Nei, M. 1972. Genetic distance between populations. *American Naturalist* 106: 283-292.

- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70: 3321-3323.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press. New York, 512 páginas.
- Neigel, J.E. 2002. Is  $F_{ST}$  obsolete? *Conservation Genetics* 3: 167-173.
- Núñez-Farfán, J., C.A. Domínguez, L. E. Eguiarte, A. Cornejo, M. Quijano, J. Vargas y R. Dirzo. 2002. Genetic divergence among Mexican populations of red mangrove (*Rhizophora mangle*): geographic and historic effects. *Evolutionary Ecology Research* 4: 1049-1064.
- Peakall, R., S.E. Smouse y D.R. Huff. 1995. Evolutionary implication of allozyme and RAPD variation in diploid population of dioecious buffalograss *Buchloe dactyloides*. *Molecular Ecology* 4: 135-147.
- Raymond, M. y F. Rousset. 1995. GENEPOP Version 1.2: population genetics software for exact tests and ecumenism. *Journal of Heredity* 86: 248-249.
- Rives, G., R. C. 2009. Diversidad clonal y estructura genética espacial en escala fina de *Agave striata* Zucc. Tesis de Licenciatura. Facultad de Ciencias, UNAM, México D.F., México, 89 págs.
- Roeder K., Devlin B. y B.G. Lindsay. 1989. Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. *Biometrics* 45: 363-379.
- Rousset, F. 2001. Inferences from spatial population genetics. En: Baldwin D.J., Bishop M. y Cannings L. (eds.) *Handbook of statistical genetics*. John Wiley & Sons. Chichester, U.K. págs. 239-262.
- Scheinvar, G., E. 2008. Genética de poblaciones silvestres y cultivadas de dos especies mezcaleras: *Agave cupreata* y *Agave potatorum*. Maestría en Ciencias Biológicas, Instituto de Ecología, UNAM, D.F., México.
- Slatkin, M. 1981. Estimating levels of gene flow in natural populations. *Genetics* 99: 323-335.
- Slatkin, M. 1985a. Gene flow in natural populations. *Ann. Rev. Ecol. Syst.* 16: 393-430.
- Slatkin, M. 1985b. Rare alleles as indicators of gene flow. *Evolution* 39: 53-65.
- Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47: 264-279.
- Slatkin, M. 1994. Gene flow and population structure. En: Real, L.A., ed. *Ecological Genetics*, Princeton University Press, Princeton N.J., USA. pp. 3-18.
- Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.
- Slatkin, M., y N.H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43: 1349-1368.
- Smouse, P.E. y T.R. Meagher. 1994. Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). *Genetics* 136: 313-322.
- Smouse, P.E. y R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82: 561-573.
- Sokal, R.R. y F.J. Rohlf. 1995. *Biometry*. 3rd ed. W.H. Freeman, New York.

- Weir, B.S. y C.C. Cockerham. 1984. Estimating  $F$  –statistics for the analysis of population structure. *Evolution* 38: 1358-1370.
- Whitlock, M.C. y D.E. McCauley. 1999. Indirect measures of gene flow and migration:  $F_{ST}$  not equal to  $1/(4Nm + 1)$ . *Heredity* 82: 117-125.
- Workman, P.L. y J.D. Niswander. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. *Amer. J. Hum. Genet.* 22: 24-29.
- Wright, S. 1943. Isolation by distance. *Genetics* 28: 114-138.
- Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics* 15:323-354.
- Wright, S. 1969. *Evolution and Genetics of Populations, Vol.2, The theory of gene frequencies*. University of Chicago Press, Chicago, USA.

=====

## **RECUADRO 1: OTROS INDICES QUE MIDEN LA DIFERENCIACION GENETICA ENTRE POBLACIONES RELACIONADOS CON LA $F_{st}$**

Inicialmente Wright (1951) definió la  $F_{st}$  sólo para un locus con dos alelos. La formulación de Nei (1973) permite incluir un número ilimitado de alelos por loci, en términos de la comparación en la heterocigotía y llamó a este índice  $G_{st}$ .

Posteriormente, Weir y Cockerham (1984) desarrollaron un algoritmo que hace una partición de la varianza análoga a un análisis de la varianza (ANOVA), y a esta estimación la llamaron coeficiente de coancestría o theta, que tiene la ventaja de estar menos sesgado que las otras dos estimaciones y que permite manejar de manera adecuada diferencias en los número de individuos y loci entre localidades.

Más recientemente, Excoffier (ver revisión en 2001) generalizó el método anterior en el AMOVA (Analysis of Molecular Variance), que tiene menos requerimientos estadísticos y resulta menos sesgada que una estimación de  $F_{st}$  normal, y se puede estimar a varios niveles jerárquicos (subpoblaciones, poblaciones en una región, entre regiones, subespecies, etc.) y es especialmente útil para datos dominantes (como RAPDs, ISSRs, AFLPs, etc.).

Para otros modelos de mutación, como el *stepwise mutation model* (que se puede traducir como “modelo mutacional de un paso”, que se en teoría siguen los microsatélites, se han propuesto la  $R_{st}$ , o para datos tipo secuencias de ADN, como la  $N_{st}$  (ver revisión en Excoffier, 2001).

=====

## **RECUADRO 2. EL METODO DE LOS ALELOS PRIVADOS (UNICOS)**

La idea básica de este método es sencilla: si hay mucho flujo génico, todos los alelos tendrán frecuencias parecidas en todas las poblaciones, mientras que si hay poco flujo, algunos alelos solo existirán en las poblaciones donde se originaron por mutación y en esas poblaciones pueden llegar a frecuencias altas. A partir de simulaciones, Slatkin analizó la relación entre la frecuencia de estos alelos que sólo existen en una población (los alelos únicos o “privados”, como los llamo Slatkin) y el flujo génico y obtuvo ecuaciones sencillas. Así, en este método solo se necesita estimar el promedio de las frecuencias alélicas de los alelos únicos (= privados) en diferentes poblaciones que se denomina  $p(1)$ . Slatkin (1985b) encontró que para el modelo de islas y el de piedras de paso el  $\log_{10}[p(1)]$  está linealmente relacionado a  $\log_{10}(Nm)$ , entonces  $\log_{10}[p(1)] = (a \log_{10}(Nm) + b)$  donde  $p(1)$  es la frecuencia promedio de los alelos privados y  $a$  y  $b$  son constantes determinadas a partir de datos simulados y dependen del número de individuos muestreados en cada población (para detalles ver Slatkin, 1985b).

Este método y la estimación de  $Nm$  a partir de  $F_{st}$  son parecidos, ya que ambos se basan en analizar la dispersión de la distribución de las frecuencias alélicas entre poblaciones. Sin embargo, simulaciones han demostrado que el

método de alelos privados es más sensible a errores en la estimación y muestreo de las frecuencias alélicas (Slatkin y Barton, 1989).

Por ejemplo, Aguirre-Planter et al. (2000; Tabla 3) estimaron  $Nm$  con diferentes métodos relacionados a la  $F_{st}$  para isoenzimas y con los alelos privados y obtuvo estimaciones similares, lo mismo que Navarro-Quezada et al. (2003) trabajando en el grupo de especies relacionados a *A. deserti* con datos de RAPDs (Tabla 2). Navarro-Quezada et al. (2003) también usaron otro método cualitativo sugerido por Slatkin (1981), que consiste en graficar las frecuencias alélicas promedio por población como función de la proporción de poblaciones que lo presentan, y comparar con las simulaciones obtenidas por Slatkin (1981) para modelos de islas y de pérdidas de paso. En éste análisis, *A. subsimplex* y *A. cerulata* se aproximaron a una  $m = 0.1$ , mientras que en *A. deserti* fue más bajo, cercano a  $m = 0.05$ .

---

### RECUADRO 3. LAS DISTANCIAS GENÉTICAS

Una medida particularmente útil de diferenciación genética (e indirectamente de flujo genético) son los índices de distancia genética, que se calculan entre pares de poblaciones y describen el grado de diferenciación genética entre estas dos poblaciones. Así, cuando dos poblaciones tienen una distancia genética alta, las poblaciones están muy diferenciadas, pues sus frecuencias alélicas son diferentes. Por ejemplo, puede calcular directamente la  $F_{st}$  entre dos poblaciones y usar esta medida como distancias genética o usar la estimación indirecta de  $Nm$  ( $M$  en la nomenclatura de Slatkin (1993, 1994)).

Además de la  $F_{st}$  existen otros índices de distancia genética, mucho de ellos diseñados originalmente para análisis cuantitativo de datos multivariados. Nei (1987, págs. 208 a 253) hace una buena revisión de esta extensa literatura.

La medida de distancia genética más usada es la Distancia de Nei (1972, 1973) de la que existen a su vez varias modificaciones (Nei, 1987). Este índice pretende estimar el número de mutaciones que a nivel nucleotídico se han acumulado en las secuencias de dos linajes durante el tiempo que ha transcurrido a partir de su divergencia original. La idea (Nei, 1987, págs. 218 a 219) es que los diferentes alelos (o la presencia o ausencia de los "loci" en el caso de marcadores relacionados al PCR como RAPDs o ISSRs) es originada gracias a que un codón (o una base nucleotídica en el caso de marcadores relacionado al PCR) es distinto. Por lo tanto, a partir de datos de frecuencias alélicas, debería ser posible calcular estadísticamente el número promedio de las diferencias en los codones o en las bases nucleotídicas por locus. Dado que éste número es una medida directa de las diferencias genéticas entre dos poblaciones, se le considera como una medida de distancia genética (Nei, 1972, 1973).

Así, según Nei, el promedio en el número neto de sustituciones nucleotídicas  $D$  estaría dado por  $D = -\log_e I$  donde  $I$  (la Identidad Genética) =  $J_{xy} / (J_x J_y)^{1/2}$ , y  $J_x$  = sumatoria de  $x_i^2$  (la suma de las frecuencias alélicas al cuadrado en la población 1),  $J_y$  = sum  $y_i^2$  (la sumatoria de las frecuencia alélicas al

cuadrado de la población 2), y  $J_{xy}$  = sumatoria del producto, para cada alelo, de las frecuencia alélicas en la población 1 por la frecuencia alélica de ese mismo alelo en la otra población. Si las frecuencias alélicas son las mismas,  $J_x = J_y$ , y la  $I$  nos da 1 (la identidad  $I$  es lo máximo posible), mientras que si no comparten ninguno de los alelos, las  $J_{xy}$  va a ser 0, y la  $I$  nos da de 0 (la identidad  $I$  es la menor posible). La  $D$  por lo tanto toma valores de 0, si son idénticas las poblaciones, a infinito, si son completamente diferentes.

Si no ha pasado mucho tiempo, digamos que menos de un millón de años Nei (1987, pág. 237)) sugiere que la  $D = 2\alpha t$ , donde  $\alpha$  = la tasa de mutación por locus por año. Así, Nei (1987) propone que si la  $\alpha$  del marcador genético es  $1 \times 10^{-7}$ , el tiempo de separación entre dos poblaciones sería  $t = 5 \times 10^6 D$  (ver Nei, 1987, pág. 237).

Por ejemplo, en el complejo de *Agave deserti* en el desierto de Sonora, al comparar las poblaciones dentro de cada una de las especies, Navarro-Quezada et al. (2003) encontraron una  $D$  de Nei = 0.032 (SE+/-0.004, 23 comparaciones), que con la fórmula anterior daría 160 mil años, mientras que si comparamos a las tres especies (*A. cerulata*, *A. deserti* y *A. subsimplex*)  $D$  aumenta a 0.041 (SE +/-0.003, 63 comparaciones), que nos da 205 mil años de separación, substancialmente menor que la fecha previamente sugerida de unos 5 millones de años considerando la separación de la península de Baja California, (Gentry, 1982), lo cual indica que las poblaciones migraron a la península mucho tiempo después de que esta se originara y se han diferenciado muy recientemente, lo cual también corresponde con nuestras estimaciones con una fecha de entre 7.8 y 10.1 millones de años para el origen del género completo (Good-Avila et al., 2006). Claro que estas estimaciones tendrán que calibrarse con tasas de mutación mas específicas, ya que no sabemos exactamente la tasa de mutación para los marcadores usados en estos estudios, RAPDs en particular.

=====

### Pies de figuras:

Fig. 1. Un modelo conceptual básico en genética de poblaciones. La especie, delimitada por un óvalo rojo, esta formada por una serie de poblaciones en azul (seis en este caso) cada una con diferentes números y proporciones de individuos homócigos ( $AA$  y  $aa$ ) y hetreócigos ( $Aa$ ). El flujo génico, representado por las flechas, homogeniza e impide que diverjan demasiado las poblaciones, dando coherencia y permitiendo que la especie evolucione en conjunto.

Fig.2: Dos posibles modelos de flujo génico. Fig. 2a Ilustra un modelo de islas de para el caso de 4 poblaciones. Cada población esta conectada a las otras por las mismas tasas de flujo génico,  $m$ . Fig.2b. El modelo de flujo génico de piedras de paso (*stepping-stone*) de Kimura (1953) en dos dimensiones. En cada generación los genes sólo pueden moverse entre las poblaciones adyacentes. Este podría ser el caso de poblaciones a lo largo de ríos, como los ahuehuetes, o las poblaciones en las costas del mangle rojo que describimos en el texto (ver Nuñez-Farfán et al., 2002).

Fig. 2c. El modelo de flujo génico de piedras de paso en tres dimensiones, al que se aproximan muchas poblaciones naturales.

Fig. 3. El balance entre flujo génico y deriva según Sewall Wright (1969). Se muestran las distribuciones en el equilibrio (después de mucho tiempo) que alcanzan  $n$ -poblaciones derivadas de una sola población original para diferentes escenarios en los que se mantienen constantes sus tamaños efectivos ( $N_e$ ) y las tasas de migración ( $m$ ) entre ellas, y todas inician con las mismas frecuencias alélicas,  $p=q=0.5$  en este ejemplo. Si hay elevado flujo génico, como en el caso de  $Nm=50$ , la mayoría de las poblaciones permanece cerca de la frecuencia alélica original, de 0.5, y se genera una campana muy alta alrededor de la frecuencia alélica original, que se muestra con la curva de color rosa e indica que el flujo génico es la fuerza dominante y ha habido poca diferenciación entre ellas. Si el flujo génico o el tamaño efectivo son menores, comienza a actuar la deriva génica, y se obtiene una campana más plana, como se ve en  $Nm=5$ , en color verde. Si el  $N_e$  es aún menor o la migración es más pequeña, la deriva génica y el flujo son similares, y se produce una distribución plana que mostramos con color azul en una línea partida para una  $Nm=0.5$ . Este quiere decir que las poblaciones pueden estar en cualquier frecuencia alélica. Si el  $Nm$  es aún menor, el proceso queda determinado por la deriva génica, y las poblaciones van perdiendo su variación genética, y la mitad se encuentra cerca de un  $q=1$  y la otra mitad alrededor de una  $q=0$ , como se muestra con la línea negra punteada para  $Nm=0.05$ .

Fig. 4. El incremento de la diferenciación entre poblaciones, definido como la  $F_{st}$  en el tiempo en generaciones, como función del tamaño efectivo,  $N_e$  y la tasa de migración entre ellas,  $m$ . Si los tamaños efectivos  $N_e$  son pequeños (menos de 100) y las tasas de flujo génico  $m$  son altas (mayores de 0.01) se llega a la  $F_{st}$  en equilibrio en cerca de unas 150 generaciones. Si los tamaños efectivos son mayores, tarda mucho más tiempo, por ejemplo si  $N_e$  es 1000 y  $m$  es de 0.001, tarda un orden de magnitud más, cerca de 1,500 generaciones en llegar al equilibrio. Así, con tamaños efectivos muy grandes y tasas de migración muy pequeñas puede tardar muchos miles o millones de generaciones en llegar al equilibrio en la diferenciación. Modificado de Hedrick (2005).

Fig. 5. Análisis de aislamiento por distancias en un conjunto de 15 poblaciones del complejo *A. cupreata* y *A. potatorum* de Michoacán, Guerrero, Oaxaca y Puebla estimados con 90 loci de ISSRs (Scheinvar (2008)). Cada punto representa una comparación entre dos de las poblaciones. Fig. 5 a. Las distancias genéticas de Nei (1972, ver Recuadro 3); Fig. 5. b. Estimaciones de  $Nm$  a partir de la  $F_{st}$ ; como referencia se marcan los umbrales de  $Nm=1$  y  $Nm=4$ . Fig. 5. c. *A. cupreata* en Guerrero. Fig. 5. d. *A. potarorum* en Oaxaca.

=====

**Tabla 1.** Estimaciones de variación genética en especies mexicanas.  $H_e$ = Diversidad genética, definida como Heterocigosis esperada en Hardy-Weinberg (rango 0, si no hay variación a 1, el máximo posible).  $F_{is}$ = Estadístico de Wright que indica endogamia, es 0 si la población esta en Hardy-Weinberg, y si es negativo indica exceso de heterocigos, mientras que si es positivo indica exceso de homocigos, (rango de -1 a 1).  $F_{st}$ = Estadístico de Wright que se refiere a la diferenciación genética entre poblaciones debido a deriva génica (va de 0 si no hay diferenciación a 1 si son completamente distintas las poblaciones).  $F_{it}$ = Estadístico de Wright que señala el efecto conjunto de la endogamia y la deriva génica en la diferenciación total de las poblaciones (va de -1 a +1). Si no se indica nada, los datos son de isoenzimas. Las otras opciones son MC= microsatélites de cloroplasto, MN= microsatélites nucleares, RAPDs. Polip. indica que la especie es poliploide, y no se pudo calcular la  $F_{is}$

	$H_e$	$F_{is}$	$F_{st}$	$F_{it}$	Referencias
=====					
<u>Frijol (<i>Phaseolus</i> spp.):</u>					
<i>P. coccineus</i>	0.305	0.25	0.20	0.40	Escalante et al. 1994
<i>P. coccineus</i>	0.24	-0.368	----	----	Souza et al. 1997
<i>P. vulgaris</i> cultivado	0.04	0.60	----	----	Escalante et al., 1994
<i>P. vulgaris</i> silvestre	0.025	----	----	----	Souza et al. 1997
<i>P. vulgaris</i> cultivado	0.016 a 0.023	----	----	----	Souza et al. 1997
-----					
<u>Arboles y palmas tropicales</u>					
<u><i>Astrocaryum</i></u>					
<i>mexicanum</i>	0.15	-0.45	0.042	-0.39	Eguiarte et al. 1992
<i>Psychotria faxlucens</i>	0.20	0.115	0.026	0.138	Pérez-N. et al., 1993
<i>Cecropia obtusifolia</i>	0.05	0.034	0.029	0.061	Alvarez-Buylla y Garay 1994
<i>Antirhea aromatica</i>	0.3081	0.2689	0.5105	0.6383	González-A. y Castillo 2005
<i>Chamaedorea elatior</i>	0.316	0.422	0.065	0.479	Luna, 1999
<i>Ch. alternans</i> RAPDs	0.35	----	0.018	----	Otero, 1998
<u><i>Desmoncus</i></u>					
<i>orthacanthus</i>	0.36	-0.370	0.126	-0.197	Escalante et al. ms.
<i>Erythroxylum</i>					
<i>havanense</i> RAPDS	0.36-0.48	----	0.094	----	Domínguez et al. 2005
<i>Rhizophora mangle</i>					Núñez-F. et al. 2002
Costa Caribe y Golfo	0.118	0.48	0.25	0.61	
Costa Pacífico	0.170	0.44	0.06	0.47	
Total	0.141	0.45	0.23	0.58	
<i>Bursera cuneata</i>	0.22	0.26	0.14	0.37	Del Valle, 1997
<i>Bursera microphylla</i>	0.183	0.387	0.179	0.497	Hernández, 1999
<i>Bursera hindsiana</i>	0.297	0.118	0.169	0.267	Vargas, 2000
<i>Brongniartia vazquezii</i>	0.201	0.533	0.101	0.583	González-A. y Núñez-F. 2001
<i>Leucaena esculenta</i>	0.147-0.411	0.304	0.142	0.402	Zárate, 1999
-----					
<u>Coníferas:</u>					
<i>P. rzedowskii</i>	0.22	0.27	0.17	0.40	Delgado et al. 1999
<i>P. lagunae</i>	0.386	0.534	0.188	0.622	Molina et al. 2001
<i>P. muricata</i>	0.346	0.307	0.161	0.418	Molina et al. 2001
<i>P. maximartinezii</i>	0.122	0.081	----	----	Ledig et al. 2001
<i>P. maximartinezii</i> MC	0.27	0.046	----	----	Moreno 2002



<i>P.nelsonii</i>	MC	0.73	----	0.13	----	Cuenca 2001
<i>P.pinceana</i>		0.374	0.458	0.257	0.592	Molina et al.2001
<i>P. pinceana</i>		0.174	0.14	0.152	0.271	Ledig et al. 2001.
<i>P.pinceana</i>	MC	0.824	----	0.78	----	Escalante 2001
<i>P. cembroides</i>	MC	0.59	----	0.0425	----	Cuenca 2003
<i>P.discolor-</i>						
<i>P.johannis</i>	MC	0.77	----	0.126	----	Cuenca 2003
<i>P.pseudostrobus</i>	MN	0.203	0.038	0.022	----	Delgado, 2002
<i>P. pseudostrobus</i>	MC	0.316	----	0.146	----	Delgado, 2002
<i>P. montezumae</i>	MN	0.228	0.138	0.122	----	Delgado, 2002
<i>P.montezumae</i>	MC	0.281	----	0.27	----	Delgado, 2002
Híbridos						
<i>P.s. x P.m.</i>	MN	0.213	0.049	0.082	----	Delgado, 2002
Híbridos						
<i>P.s. x P.m</i>	MC	0.393	----	0.27	----	Delgado, 2002
-----						
Cuatro <i>Abies</i> del norte de México*		0.099	0.36	0.21	0.49	Keiman, 1997, Furnier y Eguiarte1997
<i>A. guatemalensis</i>		0.077	0.19	0.21	0.36	Aguirre et al. 2000
<i>A. hickelii</i>		0.088	0.44	0.02	0.46	Aguirre et al. 2000
<i>A. religiosa</i>		0.108	0.22	0.24	0.41	Aguirre et al. 2000
<i>A.flinckii</i>		0.110	0.07	0.28	0.33	Aguirre et al. 2000
<i>A. guatemalensis</i> MC		0.866	----	0.131	----	Aguirre 2005.
<i>A. hickelii</i> MC		0.946	----	0.065	----	Aguirre 2005.
<i>A. religiosa</i> MC		0.888	----	0.075	----	Aguirre 2005.
<i>A.flinckii</i> MC		0.750	----	0.029	-----	Aguirre 2005
-----						
*( <i>A. concolor</i> , <i>A. durangensis</i> , <i>A. d var. coahuilensis</i> y <i>A. vejari</i> )						
-----						
Plantas de zonas áridas						
<i>A. victoriae-reginae</i>		0.335	0.055	0.236	0.281	Martínez-Palacios et al. 1999
<i>A. lechugilla</i>		0.394	0.105	0.083	0.179	Silva y Eguiarte, 2003
<i>A.subsimplex</i>		0.28	0.18	0.31	----	Eguiarte et al. 2000
<i>A. subsimplex</i> RAPDS		0.144	----	0.084	----	Navarro et al. 2003
<i>A. deserti</i> RAPDS		0.186	----	0.135	----	Navarro et al. 2003
<i>A. cerulata</i> RAPDS		0.237	----	0.098	----	Navarro et al. 2003
<i>Manfreda</i>						
<i>brachystachya</i>		0.48	-0.36	0.03	-0.31	Eguiarte et al. 2000
<i>P. glandulosa</i>		0.45	-0.44	----	----	Golubov et al. 1999a
<i>Flourensia</i>						
<i>cernua</i>		0.46	-0.146	0.08	-0-013	Ferrer et al. 2004.
<i>Lophocereus schotti</i>		0.126	-0.187	0.130	-0.031	Parker y Hamrick, 1992
<i>Echinocactus</i>						
<i>platyacanthus</i>		0.026	0.406	0.0021	0.407	Jiménez, 2008.
<i>Stenocereus eruca</i>		0.154	0.739	0.069	0.757	Clark, 2000.
<i>Stenocereus gummosus</i>	0.261		0.608	0.102	0.648	Clark y Molina-Freaner. 2003
<i>Carnigea gigantea</i>	0.116		0.057	0.075	----	Hamrick et al. 2002
<i>Pachycereus pringeli</i>	0.200		Polip	0.076	----	Hamrick et al. 2002
<i>Stenocrereus griseus</i>	0.167		0.202	0.096	----	Hamrick et al. 2002
<i>Stenocreus thurberi</i>	0.169		0.036	0.128	----	Hamrick et al. 2002
-----						
Otras plantas						
<i>Cucurbita argyrosperma</i>						
<i>spp. sororia</i>	0.426		0.004	0.04	0.044	Montes y Eguiarte, 2002
<i>C. argyrosperma</i>						
<i>spp. argyrosperma</i>	0.391		-0.042	0.096	0.058	Montes y Eguiarte, 2002

<i>C. moscata</i>	0.416	-0.162	0.077	-0.073	Montes y Eguiarte, 2002
<i>Aechmea tuitensis</i>	0.155	0.63	-----	-----	Izquierdo y Piñero, 2000
<i>Aechmea macvaughii</i>	0.15	0.35	-----	-----	Izquierdo y Piñero, 1998
<i>A. lueddemanniana</i>	0.24	0.67	0.32	0.78	Izquierdo y Piñero, 1998
<i>A. mexicana</i>	0.16	0.64	0.62	0.86	Izquierdo y Piñero, 1998
<i>Bdallophytum bambusarum</i>	0.43	-0.05	0.10	0.05	García-Franco et al. 1998
<i>Lacandonia schismatica</i>	0.0	-----	-----	-----	Coello et al., 1993
<i>Datura stramonium</i>	0.0	-----	-----	-----	Núñez-Farfán, 1991
11 especies de <i>Tithonia</i>					
	0.092-0.139	-0.15	0.52	0.44	Morales, 1996
<i>Senecio praecox</i>	0.117	0.561	0.065	0.561	Martínez del Río, 2003.
<i>Verbesina virgata</i>	0.083	0.809	0.16	0.904	Martínez del Río, 2003.
<i>Eupatorium petiolare</i>	0.101	0.895	0.192	0.928	Martínez del Río, 2003.
<i>Dahlia coccinea</i>	0.249	0.586	0.081	0.623	Martínez del Río, 2003
<i>Tagetes lunulata</i>	0.022	0.754	0.061	0.769	Martínez del Río, 2003.
<i>Dieffenbachia seguine</i>	0.3323	0.30	0.31	0.51	Cuartas, 2002
-----					
Plantas introducidas/ cultivadas en México:					
<i>Schinus molle</i>	0.093	0.076	0.115	0.183	Aguirre, 1994
<i>Reseda luteola</i>	0.374	0.361	0.024	0.376	Huerta, 1994
<i>Cocos nucifera</i>	0.087	0.399	0.364	0.618	Zizumbo y Piñero 2002

#### Referencias de la Tabla 1 no incluidas en las referencias generales.

- Aguirre Planter, E. 1994. Análisis comparativo de la estructura genética de poblaciones nativas de Perú e introducidas a México y España de *Schinus molle* L. (Anacardiaceae). Licenciatura en Biología. Facultad de Ciencias, UNAM.
- Alvarez-Buylla, E.R. y A. Garay. 1994. Population genetic structure of *Cecropia obtusifolia*, a tropical pioneer tree species. *Evolution* 48: 437-453.
- Alvarez-Buylla, E., Chaos, A., Piñero, D. y Garay, A. 1996. Demographic genetics of a pioneer tropical tree species: Patch dynamics, seed dispersal and seed banks, *Evolution* 50: 1155-1166
- Clark Tapia, Ricardo. 2000. Estructura genética de dos cactáceas columnares del desierto sonorense: *Stenocereus gummosus* y *S. eruca* (Cactaceae). Tesis de maestría, UACPyP del CCH, Instituto de Ecología, UNAM, D.F., México.
- Clark-Tapia R. y F. Molina-Freaner. 2003. The genetic structure of a columnar cactus with a disjunct distribution: *Stenocereus gummosus* in the Sonoran desert. *Heredity* 90: 443-450.
- Coello G., A. Escalante y J. Soberón. 1993. Lack of genetic variation in *Lacandonia schismatica* (Lacandoniaceae: Triuridales) in its only know locality. *Ann. Missouri Bot. Gard.* 80: 898-901.
- Cuartas H., S. E. 2002. Efectos ecológicos y genéticos de la fragmentación en las poblaciones de *Dieffenbachia seguine* (Araceae) en Los Tuxtlas. Maestría en Ciencias Biológicas (Ambientales), Posgrado en Ciencias Biológicas, UNAM, México, Distrito Federal.
- Cuenca Navarro, A. 2001. Variación y estructura genética de una especie de pino endémica de México (*Pinus nelsonii* SHAW). Tesis de licenciatura en Biología. Facultad de Ciencias, UNAM, México, D.F.
- Cuenca Navarro, A. 2003. Evidencia de dos linajes genéticos en *Pinus cembroides* revelada por microsatélites de cloroplasto. Tesis de Maestría en Ciencias Biológicas, Posgrado en Ciencias Biológicas, UNAM, México, D.F.
- Delgado Valerio, P. 2002. Introgresión e hibridización entre dos especies de pinos, *Pinus montezumae* y *P. pseudostrobus*. Tesis de Doctorado en Ciencias Biológicas, UNAM.

- Delgado, P., Piñero, D., Chaos, A., Pérez-Nasser, N. y Alvarez-Buylla, E. 1999. High population differentiation and genetic variation in the endangered Mexican pine, *Pinus rzedowskii* (Pinaceae). *American Journal of Botany*. 86: 669-676.
- Delgado, P., A. Cuenca, A. E. Escalante, F. Molina-Freaner y Piñero, D. 2002. Comparative genetic structure in pines: Evolutionary and conservation consequences. *Revista Chilena de Historia Natural*. 75: 27-37.
- Del Valle, M. 1997. Estructura genética de *Bursera cuneata* en el Centro de México. Tesis biología, Facultad de Ciencias, UNAM México, Distrito Federal.
- Eguiarte, L.E., A. Silva y V. Souza. 2000. Biología evolutiva de la familia Agavaceae: biología reproductiva, genética de poblaciones y filogenia. Boletín de la Sociedad Botánica de México. 166: 131-150.
- Escalante Hernández, A. E. 2001. Estructura genética de poblaciones de *Pinus pinceana* usando como marcadores moleculares microsatélites de cloroplasto cpSSR's. Tesis de Licenciatura en Biología. Facultad de Ciencias, UNAM.
- Escalante, A.M., Coello, G., Eguiarte, L.E. y Piñero, D. 1994. Genetic structure and mating systems in wild and cultivated populations of *Phaseolus coccineus* and *P. vulgaris* (Fabaceae). *American Journal of Botany*. 81: 1096-1103.
- Escalante, S., R. Orellana y L.E. Eguiarte. ms. Genetic structure of the tropical palm *Desmoncus orthacanthos*: implication for management and conservation. Manuscript.
- Ferrer, M.M., L.E. Eguiarte y C. Montaña. 2004. Genetic structure and outcrossing rates in *Flourensia cernua* (Asteraceae) growing at different densities in the South-western Chihuahuan Desert. *Annals of Botany* 94: 419-426.
- Furnier G. R. y L.E. Eguiarte. 1997. Niveles y patrones de variación genética en el género *Abies* en México. Informe final. CONABIO N. de referencia B138. México D.F., México.
- García Franco, J.G., V. Souza, L. E. Eguiarte y V. Rico-Gray. 1998. Genetic variation, genetic structure and effective population size in the tropical holoparasitic endophyte *Bdallophytum bambusarum* (Rafflesiaceae). *Plant Systematics and Evolution* 210: 271-288.
- Golubov J., L. E. Eguiarte, M. C. Mandujano J. Lopez-Portillo y C. Montaña. 1999. Why be a honeyless honey mesquite? Reproduction and mating system of nectarful and nectarless individuals *American Journal of Botany* 86: 955-963.
- González-Astorga, J. y J. Núñez-Farfán. 2001. Effect of habitat fragmentation on the genetic structure of the narrow endemic *Brongniartia vazquezii*. *Evolutionary Ecology Research* 3: 1-12.
- Gonzalez-Astorga, J. y G. Castillo-Campos. 2005. Genetic variability of the narrow endemic tree *Antirhea aromatica* (Rubiaceae, Guettardeae) in a tropical forest of Mexico. *Annals of Botany*. 93: 521-528.
- Hamrick, J.L., J.D. Nason y T.H. Fleming. 2002. Genetic diversity in columnar cacti. En Fleming T.H., Valiente-Banuet A. (eds) *Evolution, Ecology and Conservation of Columnar Cacti and Their Mutualists*, University of Arizona Press: Tucson, AZ. pp 122-133.
- Hernández, A. 1999. Diferenciación genética entre poblaciones de *Bursera microphylla* en Sonora Baja California e Islas del Mar de Cortés. Tesis Profesional, Biología, Facultad de Ciencias, UNAM, México, Distrito Federal, México.
- Huerta Ocampo, E. 1994. Análisis comparativo en la estructura genética y tasas de entrecruzamiento en poblaciones de *Roseda luteola* (Resedaceae) en México y España. Tesis de licenciatura en Biología. Facultad de Ciencias, UNAM.
- Izquierdo, L.Y. y Piñero, D. 1998. Allozyme divergence among four species of *Podaechmea* sensu lato and the status of *Ursulaea* (Bromeliaceae, Bromelioideae), *Plant Systematics and Evolution*, 213: 207-215.
- Izquierdo, L.Y. y Piñero, D. 2000. High genetic diversity in the only known population of *Aechmea tuitensis* (Bromeliaceae). *Australian Journal of Botany*. 48:645-650.
- Jiménez Sierra, C.L. 2008. Demografía y biología de *Echinocactus platyacanthus* Link et Otto, en el Valle de Tehuacán, Puebla. Tesis de doctorado, Facultad de Ciencias.
- Keiman, A. F. 1997. Niveles y patrones de variación isoenzimática en el género *Abies* del Norte de México. Tesis de Licenciatura, Facultad de Ciencias, UNAM.
- Ledig, F.T., Capo-Arteaga M.A., Hodgskiss P.D., Sbay H., Flores-Lopez C., Conkle M.T., Bermejo-Velazquez B. 2001. Genetic diversity and the mating system of a rare Mexican pinon, *Pinus pinceana*, and a comparison with *Pinus maximartinezii* (Pinaceae). *American Journal of Botany* 88: 1977-1987.
- Luna Reyes, R. 1999. Demografía y genética de poblaciones de *Chamaedorea elatio* en la selva de Los Tuxtlas, Veracruz. Tesis de licenciatura, Facultad de Ciencias, UNAM, México, D.F.
- Martínez del Río, A. E. 2003. Análisis de la variación genética de cinco especies (*Dahlia coccinea*, *Eupatorium petiolare*, *Senecio praecox*, *Tagetes lunulata* y *Verbesina virgata*) de compuesta del Pedregal de San Angel, C.U., México, D.F. Tesis de Licenciatura, Facultad de Ciencias, UNAM, D.F., México.

- Martínez-Palacios, A., L. E. Eguiarte y G. R. Furnier. 1999. Genetic diversity of the endangered endemic *Agave victoriae-reginae* (Agavaceae) in the Chihuahuan desert. *American Journal of Botany* 86: 1093-1097.
- Molina-Freaner, F., Delgado, P., Pérez-Nasser, N., Piñero, D. y Alvarez-Buylla, E. 2001. Do rare pines need different conservation strategies?: Evidence from three Mexican species. *Canadian Journal of Botany* 79: 131-138.
- Montes-Hernández, S. y L.E. Eguiarte. 2002. Genetic structure and indirect estimates of gene flow in three taxa of *Cucurbita* in Western Mexico. *American Journal of Botany* 89: 1156-1163.
- Morales Guillaumin, E. 1996. El método comparativo en los estudios de evolución de historias de vida: Un ejemplo con el género *Tithonia* (Asteraceae). Tesis de Doctorado en Ecología. UACPyP-CCH-Centro de Ecología, UNAM.
- Moreno Letelier, A. 2002. Sistema de apareamiento y variación genética en *Pinus maximartinezii* Rzedowskii. Licenciatura en Biología. Facultad de Ciencias, UNAM.
- Núñez-Farfán, J. 1991. Biología evolutiva de *Datura stramonium* L. en el Centro de México: Selección natural de la resistencia a los herbívoros, sistema de cruzamiento y variación genética intra e interpoblacional. Tesis de doctorado. UACPyP del CCH, Centro de Ecología. UNAM. D.F., México.
- Otero, Arnaiz, M.A. 1998. Variación genética y biología reproductiva de *Chamaedora alternans* mediante el uso de los marcadores oclulares (RAPDs) en las selva de Los Tuxtlas, Veracruz, México. Tesis de Mestría, Facultad de Ciencias, UNAM. D.F., México.
- Parker, K.C. y J.L. Hamrick. 1992. Geneti diversity and clonal structure in a column cactus, *Lophocercus schottii*. *American Journal of Botany* 79: 86-96.
- Pérez Nasser, N., L. Eguiarte, L. y D. Piñero. 1993. Mating system and genetic structure of the distylous tropical tree *Psychotria faxlucens* Lorence & Dwyer (Rubiaceae). *American Journal of Botany* 80:45-56.
- Silva-Montellano A. y L. E. Eguiarte. 2003. Geographical patterns in the reproductive ecology of *Agave lechuguilla* (Agavaceae) in the Chihuahuan desert: II. Genetic variation, differentiation and inbreeding estimates. *Am. J. of Botany* 90: 700-707.
- Souza, V., J. Bain, C. Silva, V. Bouchet, A. Valera, E. Márquez y L. Eguiarte. 1997. Ethnomicrobiology: Do Agricultural practices modify the population structure of the nitrogen fixing bacteria, *Rhizobium etli* biovar *phaseoli*? *Journal of Ethnobiology* 17(2): 249-266.
- Vargas García, J. 2000. Impacto de la formación de la Península de Baja California sobre la estructura genética de *Bursera hindsiana*. Tesis de licenciatura, Biología, Facultad de Ciencias, UNAM, México, D.F., México.
- Silva-Montellano A. y L. E. Eguiarte. 2003. Geographical patterns in the reproductive ecology of *Agave lechuguilla* (Agavaceae) in the Chihuahuan desert: II. Genetic variation, differentiation and inbreeding estimates. *Am. J. of Botany* 90: 700-707.
- Zárate Pedroche, S. 1999. Estudios sistemáticos del proceso de domesticación del género *Leucaena* en México. Tesis de Doctorado. Facultad de Ciencias, UNAM
- Zizumbo-Villarreal, D. y D. Piñero. 2002. Diversity and phylogenetic analysis in *Cocos nucifera* L. in México. *Genetic Resources and Crop Evolution*, 49: 237-245.

**Tabla 2.** Diferentes medidas de diferenciación genética en 14 poblaciones de 3 especies de *Agave* del desierto de Sonora, obtenidas con 41 loci con la técnica RAPDs. En primer lugar tenemos una comparación de las frecuencias alélicas con la prueba combinada de Fisher ( $P$  igual o menor a 0.001), theta de Weir y Cockerham (1984, equivalente a la  $F_{st}$ ), la  $G_{st}$  (Nei, 1973, ver Recuadro 1), y estimaciones de  $Nm$  a partir de  $F_{st}$  (theta, ver Recuadro 1) y del modelo de los alelos privados de Slatkin (1981, ver Recuadro 2).

	Prueba combinada de Fisher	$F_{ST}$	$G_{ST}$	$Nm$ de $F_{st}$	$Nm$ alelos priv.
Todas las 14 pobl.	Todos los loci diferentes	0.137	0.113	2.91	3.51
<i>A. subsimplex</i>	11/41 loci diferentes	0.084	0.07	6.14	3.83
<i>A. deserti</i>	Todos los loci diferente	0.135	0.149	2.99	no hay
<i>A. cerulata</i>	35/41 loci diferentes	0.098	0.095	4.38	3.36

**Tabla 3.** Estimaciones de diferenciación genética y  $Nm$  en 4 especies de oyamel, *Abies*, en poblaciones del sur de México y Guatemala, usando 16 loci nucleares (isoenzimáticos) y microsatélites de cloroplasto. Ver texto para detalles. Datos de Aguirre-Planter et al. (2000) y Aguirre-Planter (2005)

	Núcleo (Isoenzimas)			Microsatélites cloroplasto			
	$F_{st}$		Alelos Privados				
	$F_{st}$	$Nm$		$F_{st}$	$Nm$	$R_{st}$	$Nm$
<i>A. flinckii</i>	0.271	0.672	3.42	0.029	16.74	0.116	3.81
<i>A. guatemalensis</i>	0.122	1.8	2.88	0.131	3.32	0.104	4.32
<i>A. hickeli</i>	0.073	3.17	2.70	0.065	7.19	0.040	11.86
<i>A. religiosa</i>	0.250	0.75	1.67	0.075	6.09	0.136	0.136

**Tabla 4.** Algunos paquetes y sitios de internet útiles para análisis de flujo génico, y estructura genética y genética de poblaciones.

<b>Programa</b>	<b>Autores</b>	<b>Sitio Internet</b>
TFPGA: Tools For Population Genetic Analysis	Miller, Mark 1997	<a href="http://www.marksgeneticsoftware.net/">http://www.marksgeneticsoftware.net/</a>
Arlequin: A software for population genetics data análisis	S. Schneider, D. Roessli y L. Excoffier, 2000	<a href="http://lgb.unige.ch/arlequin/">http://lgb.unige.ch/arlequin/</a>
Partition (Métodos Bayesianos)	Khalid Belkhir y Kevin J. Dawson	<a href="http://www.univ-montp2.fr/~genetix/partition/partition.htm">http://www.univ-montp2.fr/~genetix/partition/partition.htm</a>
GDA: Genetic Data Analysis	Paul O. Lewis y Dmitri Zaykin	<a href="http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php">http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php</a>
Hickory: Software for analysis of geographic structure in genetic data	Kent Holsinger y Paul Lewis	<a href="http://darwin.eeb.uconn.edu/hickory/hickory.html">http://darwin.eeb.uconn.edu/hickory/hickory.html</a>
GENEPOP	Michel Raymond y Francois Rousset	<a href="http://wbiomed.curtin.edu.au/genepop/">http://wbiomed.curtin.edu.au/genepop/</a>
AMOVA-PREP	Mark P. Miller	<a href="http://www.marksgeneticsoftware.net/">http://www.marksgeneticsoftware.net/</a>
STRUCTURE (utiliza genotipo multilocus para encontrar estructura poblacional. Métodos bayesianos).	Hubisz, Falush, Stephens and Pritchard	<a href="http://pritch.bsd.uchicago.edu/structure.html">http://pritch.bsd.uchicago.edu/structure.html</a>
SPAGeDi: Spatial Pattern Analysis of Genetic Diversity	Olivier Hardy y Xavier Vekemans	<a href="http://ebe.ulb.ac.be/ebe/Software.html">http://ebe.ulb.ac.be/ebe/Software.html</a>
Geneland	Gilles Guillot, Filipe Santos y Arnaud Estoup	<a href="http://www2.imm.dtu.dk/~gigu/Geneland/">http://www2.imm.dtu.dk/~gigu/Geneland/</a>
<hr/>		
<a href="http://evolution.genetics.washington.edu/phylip/software.html">http://evolution.genetics.washington.edu/phylip/software.html</a> Página con ligas a varios programas		
<a href="http://www.ecologia.unam.mx/laboratorios/evolucionmolecular/">http://www.ecologia.unam.mx/laboratorios/evolucionmolecular/</a> Página de nuestro laboratorio (Evolución Molecular y Experimental, Instituto de Ecología, UNAM), donde se pueden encontrar los pdf de muchos de los artículos citados aquí escritos por nosotros, tesis, e información importante sobre métodos y ligas a otras páginas.		