

ARTÍCULO DE REVISIÓN

Avances metodológicos para el estudio conjunto de la información genética, genealógica y geográfica en análisis evolutivos y de distribución

Methodological advances for the combined study of genetic, genealogical and geographical information in evolutionary and distributional analyses

ELLA VÁZQUEZ-DOMÍNGUEZ*, SUSETTE CASTAÑEDA-RICO, TANIA GARRIDO-GARDUÑO
& TANIA A. GUTIÉRREZ-GARCÍA

Departamento de Ecología de la Biodiversidad, Instituto de Ecología, Universidad Nacional Autónoma de México, Apartado Postal 70-275, Ciudad Universitaria, México DF, 04510, México

* Autor correspondiente: evazquez@ecologia.unam.mx

RESUMEN

El análisis de la distribución de linajes de genes a lo largo de su distribución geográfica, para elucidar patrones y procesos históricos, evolutivos y ecológicos que determinan la distribución de la biota en el planeta, conforma la disciplina de estudio conocida como filogeografía. En esta revisión presentamos un resumen de la historia, definición y progresos de la filogeografía, profundizando sobre las herramientas, métodos y técnicas utilizadas para desarrollar estudios filogeográficos. Sintetizamos los fundamentos y avances analíticos y teóricos, con explicaciones mediante ejemplos relevantes que permitirán a los lectores obtener el sentido y alcances de esta área de investigación. Finalmente, indicamos las nuevas líneas de análisis y estudio de la filogeografía.

Palabras clave: ADN, coalescencia, estadísticos de resumen, filogeografía, genealogías de genes.

ABSTRACT

Phylogeography is a discipline that involves the analysis of the distribution of gene lineages throughout their geographical range, with the objective of evaluating and elucidating the historic, evolutionary and ecological patterns and processes that determine the global distribution of the biota. In the present review we present a summary of the history, definition and progresses of phylogeography, deepening about the tools, methods and techniques used to perform phylogeographical studies. We synthesize the fundamentals and theoretical and analytical advances, with explanations based on relevant examples that will allow readers to acquire the sense and scope of this research field. Finally, we point out the novel routes of phylogeographic analysis and study.

Key words: DNA, coalescence, gene genealogies, phylogeography, summary statistics.

INTRODUCCIÓN

La filogeografía, a 20 años de su nacimiento como disciplina de estudio (Avice et al. 1987), ha acuñado diversas definiciones, de las cuales la más ampliamente utilizada es la que formuló John Avice, en el 2000, en su libro *Phylogeography, the history and formation of species*: “es el campo de estudio relacionado con los principios y procesos que gobiernan la distribución geográfica de linajes de genes, sobre todo aquellos entre y dentro de especies

cercanamente relacionadas” (Avice 2000). Sin embargo, la filogeografía, como la entendemos hoy, se enmarca en un área más amplia denominada ecología molecular, la cual se define, de manera general, como el conocimiento y aplicación de marcadores genéticos moleculares para explorar preguntas y problemas en ecología y evolución. La naturaleza y la escala de los estudios de ecología molecular están definidas por el hecho de que se estudian las relaciones genéticas entre individuos, poblaciones y especies, que se

emplean marcadores genéticos moleculares, y que es necesario contar con información cualitativa y cuantitativa de cambios en la estructura o composición genética (Carvalho 1998, Freeland 2005, Vázquez-Domínguez 2007a).

La filogeografía trabaja con los componentes históricos, filogenéticos, de la distribución espacial de linajes de genes, y considera como ejes el tiempo y el espacio, en los cuales (idealmente) se mapean las genealogías de estudio. Al comparar las relaciones evolutivas de los linajes con su distribución geográfica, es posible tener un mejor entendimiento de cuáles factores han influido más en la distribución de la variación genética. Así, la filogeografía engloba aspectos tanto temporales (relaciones evolutivas) como espaciales (distribución geográfica). Asimismo, enfatiza aspectos históricos de la distribución espacial actual de linajes de genes, utilizando información temporal (histórica) que ha sido posible obtener a partir del uso y análisis de la información contenida en las moléculas de ácido desoxirribonucleico (ADN); en otra escala (ecológica), enfatiza el papel de las presiones o factores ecológicos contemporáneos en la conformación de la distribución espacial de caracteres. Por lo tanto, la filogeografía trata de interpretar el modo y la extensión en que los procesos históricos en la demografía pudieron dejar marcas evolutivas en la distribución geográfica actual de caracteres genéticamente basados (Avice 2000, Vázquez-Domínguez 2002, Freeland 2005).

Marcadores moleculares en filogeografía

El acelerado desarrollo de técnicas moleculares en general, así como sobre el conocimiento de la molécula de ADN mitocondrial (ADNmt) en particular (mediados-finales de los 70), contribuyó significativamente para la consolidación de los análisis filogeográficos. Entre las características moleculares que hacen especial al ADNmt están que es una molécula circular covalentemente cerrada, de tamaño pequeño (ca. 16-20 kilobases), conformada por un total de 37 genes (13 RNA mensajeros, 2 RNA ribosomales y 22 RNA de transferencia), además de una región conocida, en vertebrados y equinodermos, como región control (ca. 1 kb) o d-loop (fragmento más pequeño dentro de la región control misma), que controla la

replicación y transcripción en la molécula y tiene una tasa excepcionalmente alta de sustitución y de polimorfismos en muchos taxones (para una descripción detallada del ADNmt ver Brown 1985, Wainscoat 1987, Lunt et al. 1998). Las propiedades más interesantes en términos filogeográficos y particularmente del ADNmt en animales son su alta tasa de evolución (sustitución) a nivel de secuencias de nucleótidos, su prácticamente nula recombinación, gran variación intraespecífica, y más importante, su herencia estrictamente materna (con muy escasas excepciones), que permite describir la historia matrilineal de organismos coespecíficos y con ello el poder aplicar estimaciones de reloj molecular y hacer análisis coalescentes (ver más adelante). El ADNmt de plantas, aunque habita también en el citoplasma y codifica grupos similares de genes, no está regido por las mismas reglas que el animal; evoluciona de manera rápida respecto al orden de los genes, pero despacio en secuencias nucleotídicas (Palmer 1990). El ADN de cloroplasto (ADNcp), aunque con ciertos problemas como la posibilidad de tener herencia biparental en algunos taxones y tasas de evolución de moderadas a bajas, ha sido la opción más exitosa para desarrollar estudios filogeográficos con plantas (Soltis et al. 1992). Estas razones, además de que existe un gran número de oligonucleótidos ('primers') mitocondriales universales para animales, ha tenido como consecuencia que los estudios filogeográficos de animales hayan históricamente quintuplicado aquellos de plantas.

Sin embargo, el uso de ADNmt representa problemas que hay que tomar en cuenta: dado que implica un único locus, tiene limitaciones en cuanto a la reconstrucción de historias poblacionales, sobre todo si ese locus ha estado sujeto a selección o algún otro proceso, o si el ADNmt pasó recientemente de una especie a otra por hibridización. También son de considerarse su alta sensibilidad a cuellos de botella y la posibilidad de que su herencia materna resulte en una reconstrucción incompleta de la historia poblacional si hembras y machos tienen patrones de dispersión diferentes (Avice 2000, Freeland 2005). La forma más adecuada de probar si una genealogía de ADNmt refleja correctamente la historia es evaluando su concordancia con

genealogías inferidas con regiones de ADN en otros genomas: en plantas normalmente se comparan con plástidos (orgánulos limitados por membrana que se encuentran solamente en las células de plantas y algas) y regiones nucleares, mientras que en animales solo es posible con loci nucleares. Ello conlleva problemas inherentes, sobre todo por la posibilidad de recombinación que implica que diferentes partes de un mismo locus pueden tener diferente historia evolutiva; sin embargo, también se ha encontrado, en un buen número de estudios filogeográficos con genes nucleares, que la recombinación no es necesariamente un problema insalvable (Hare 2001, Husmeier & Wright 2001).

Por otro lado, dados los cientos o miles de moléculas de ADNmt que tiene un individuo, la heteroplasmia –que es la presencia de dos o más haplotipos mitocondriales dentro de una misma célula– parecía un problema potencial en estudios filogeográficos. Contrario a lo esperado, se conoce que los casos de heteroplasmia son mínimos y no representan un problema en estudios rutinarios con ADNmt en animales. Por el contrario, se reconoce que existe la posibilidad de efectos de cambios homoplásicos (caracteres similares pero no derivados de un ancestro común) en estudios con ADNmt, por ejemplo cambios paralelos o convergentes en sitios de restricción, sobre todo en términos de ambigüedad en la construcción de los árboles o redes de haplotipos. Cabe mencionar que no solo existen ya formas de evaluar y lidiar con la homoplasia (ver Templeton 1983), sino que los problemas aquí señalados cuando se usa ADNmt son actualmente fáciles de identificar por la gran experiencia acumulada a través de los múltiples estudios filogeográficos publicados (para una lista de ejemplos en animales ver Avise 2000, Vázquez-Domínguez 2007b y en plantas Freeland 2005, Eguiarte et al. 2007), así como el impresionante desarrollo que han tenido los métodos de análisis (extensa revisión en Emerson et al. 2001, Hare 2001, Posada & Crandall 2001, Knowles & Maddison 2002, Hey & Machado 2003, Templeton 2004, Posada et al. 2006, Eguiarte et al. 2007). Finalmente, la mayoría de los estudios filogeográficos que han utilizado genes nucleares se han basado en la secuenciación de genes específicos (e.g. microsátelites). Asimismo, en filogeografía

existe un interés creciente por usar polimorfismos de nucleótidos únicos (single nucleotide polymorphisms, SNPs) de múltiples loci para reconstruir historias poblacionales, dado que los SNPs representan la forma más prevaleciente de variación genética (Brumfield et al. 2003).

Un ejemplo con ADNmt donde se muestra una clara estructura filogeográfica es el trabajo de Graziotin et al. (2006), donde emplean citocromo b para analizar un complejo de serpientes (Familia Viperidae) que se distribuye en el bosque brasileño del Atlántico y que incluye a una especie continental *Bothrops jararaca* (Wied-Neuwied, 1824), y dos especies insulares *B. insularis* (Amaral, 1922) y *B. alcatraz* (Marques, Martins & Sazima, 2002). Los resultados indican que estas dos últimas especies se originaron recientemente a partir de la primera, pero por distintos eventos durante el Pleistoceno (origen parafilético). Por otro lado, se identificaron dos filogrupos continentales (norte y sur) con un tiempo de divergencia de aproximadamente 3.8 millones de años (MA) (Plioceno). El grupo del sur muestra una mayor fluctuación del tamaño poblacional y menor diversidad genética que el norte, lo que se atribuye a que el sur de Brasil tuvo un mayor impacto de las oscilaciones climáticas durante el Pleistoceno (e.g. Behling & Lichte 1997, Behling & Negrelle 2001). En otro ejemplo con citocromo b, se ha podido confirmar que la ruta migratoria del roedor *Ototylomys phillotis* (Merriam, 1901) fue de sur a norte (Centroamérica hacia sur de México), durante el gran intercambio biótico que se dio al formarse el istmo de Panamá hace aproximadamente 3 MA (Gutiérrez-García & Vázquez-Domínguez resultados no publicados).

No solo el ADNmt ha mostrado óptimos resultados en estudios filogeográficos, por ejemplo en un estudio con *Arabis alpina* (L.) donde se empleó ADN de cloroplasto (región trnL-F) y nuclear (ITS), ambos marcadores mostraron, con ligeros cambios, una clara diferenciación genética y estructura filogeográfica a lo largo de la distribución de la especie, reconociéndose básicamente tres grupos: este de África, Asia y el Mediterráneo-Europa (Koch et al. 2006). Se pudo determinar que hubo tres procesos de dispersión a partir de las poblaciones de Asia (Medio Este), poblaciones además con altas tasas de

diversidad genética: i) hacia África vía la península Arábiga, ii) por el oeste hacia Europa y el norte de África y más tarde hacia el norte, iii) una nueva migración hacia el este de África (Koch et al. 2006). Algunos de estos patrones se han encontrado para otras especies de plantas (e.g. Mummenhoff et al. 2001, Bleeker et al. 2002).

Reloj molecular

Una forma sencilla de evaluar las relaciones evolutivas entre diferentes alelos es estimar cuánto difieren dos secuencias homólogas. Esto es, se puede estimar hace cuánto tiempo divergieron dos secuencias con base en qué tan diferentes son una de otra, pues secuencias similares en general habrán divergido recientemente y secuencias muy diferentes habrán sido evolutivamente independientes por un tiempo relativamente largo. Dicha estimación se hace normalmente con base en el número de diferencias en las bases y se expresa como el porcentaje de sitios variables entre pares de secuencias, aunque modelos más complejos toman en cuenta procesos mutacionales, por ejemplo dando diferente peso a las transiciones y transversiones o a las substitutiones sinónimas respecto a las no-sinónimas.

En los años 1960 se tenía la idea de que las secuencias de ADN evolucionaban a tasas casi constantes y por lo tanto que la diferencia entre dos secuencias podía utilizarse para estimar directamente el tiempo de divergencia entre estas, a lo que se le conoce como reloj molecular (Zuckerlandl & Pauling 1965). La calibración del reloj molecular se basa en la fecha aproximada cuando dos linajes genéticos divergieron, fecha que idealmente debería obtenerse de información independiente de la molecular, por ejemplo del registro fósil o de un evento geológico conocido (e.g. surgimiento de una isla). Posteriormente se calcula el valor de divergencia (de las secuencias) que ha habido desde esa fecha, ello al dividir la cantidad de divergencia que ha ocurrido durante ese tiempo entre el tiempo estimado desde que divergieron, con lo que se obtiene una estimación de la tasa a la que ha ocurrido la evolución molecular (o lo que es lo mismo, la tasa a la que el reloj molecular ha ido caminando).

Por ejemplo, si se tienen dos secuencias de 500 pares de bases (pb), las cuales difieren en 20 pb y se conoce un evento geológico que sucedió hace 1 millón de años, se tiene que:

Fechado = 1.000.000 años

Cantidad de divergencia = $20/500 = 0,04 (= 4 \%)$

Calibración: dado que normalmente se expresa en términos de porcentaje de divergencia por millón de años, en este ejemplo la calibración del reloj molecular sería de 4 % de divergencia por cada millón de años.

Tomando en cuenta que los diferentes haplotipos dentro de una población registran la historia matrilineal de eventos mutacionales, la filogeografía ha podido emplear estimaciones del reloj molecular, ya sea de ADNmt o de marcadores nucleares. El ADNmt tiene una tasa de diferenciación alta, pero muy variable entre grupos taxonómicos; el reloj molecular más ampliamente utilizado es el de vertebrados –y más específicamente mamíferos– de aproximadamente 2 % de divergencia de secuencias por cada millón de años (Brown et al. 1979). Sin embargo, este reloj constante es atribuible solo a cambios en ciertas posiciones de nucleótidos específicas, además de que diferentes regiones y genes del ADNmt mismo presentan variadas tasas de evolución dentro de un linaje. Por ejemplo, la región control ha sido particularmente útil para análisis filogeográficos en escalas muy pequeñas de tiempo microevolutivo (e.g. miles o cientos de miles de años), ya que tiene una tasa excepcionalmente rápida de sustitución de nucleótidos y altos niveles de polimorfismo intraespecífico (Lunt et al. 1998). Aunque no existen reglas específicas, la tasa de mutación del ADNmt parece variar de acuerdo con ciertas reglas taxonómicas, como hábito termal, tiempo generacional y tasa metabólica (Martin & Palumbi 1993), por lo que es mejor usar un reloj molecular que haya sido calibrado dentro del grupo taxonómico y región génica de estudio. Se ha observado que existe variación temporal en cuanto a la tasa de sustitución, y que esta tasa es además específica a cada gen, lo cual debe considerarse al hacer estimaciones de tasas evolutivas (Bedford & Hartl 2008).

García-Moreno et al. (2004) utilizan reloj molecular con la tasa de diferenciación del 2 % para estimar tiempos de divergencia de

Chlorospingus ophthalmicus (Du Bus de Gisignes, 1847), especie de ave restringida a bosques mesófilos en México, entre diferentes localidades que corresponden a complejos montañosos. Así, sus resultados muestran que las poblaciones del norte y sur de la Sierra Madre Occidental se diferenciaron hace 450.000 años. Por otro lado, Devitt (2006) realizó un análisis filogeográfico para evaluar la historia evolutiva y biogeográfica de *Trimorphodon biscutatus* (Duméril Bibron & Duméril, 1854) (serpiente) a lo largo de su distribución entre la zona de transición neártica y neotropical de México. Para ello, utilizó un reloj molecular calculado con un método Bayesiano para obtener tiempos de divergencia. El método requiere de diferentes puntos de calibración (superior e inferior) en ausencia de registro fósil, para lo cual se consideró la edad (11 MA) y la velocidad de evolución (0,07 sustituciones/sitio/millones de años) del nodo del grupo externo y el rango de edad de dos nodos internos (14-7 MA y 8-4 MA), basados en datos geológicos del Eje Neovolcánico Transversal y de la península de Baja California. Con ello pudieron datar con certeza los cinco clados obtenidos, cuya diferenciación es consistente con un modelo de vicarianza durante el Neógeno.

Demografía y coalescencia

Bajo neutralidad, en una población de tamaño constante y a lo largo de las generaciones, surgen de forma permanente nuevos alelos por mutación y se pierden otros por deriva génica, de tal forma que todos los alelos de un gen en una generación derivan de –o coalescen hacia– un único alelo ancestral. Dado que la mayoría de las mutaciones ocurren en un único punto en el tiempo y el espacio, es posible hacer inferencias de eventos pasados así, la propagación de cada nueva mutación (alelo) es afectada por patrones de dispersión, tamaño poblacional, selección natural y otros procesos que pueden deducirse a partir de la distribución actual de dichas mutaciones (Freeland 2005). De esta forma, al ver hacia atrás en el tiempo, los haplotipos de ADNmt o de cualquier gen nuclear, eventualmente coalescen hacia el ancestro común más reciente (ACMR). Se puede identificar el ACMR utilizando la coalescencia (teoría de coalescencia), la cual se basa en una teoría matemática desarrollada por

Kingman (1982) para describir la genealogía de genes neutrales hacia atrás en el tiempo (para una revisión amplia ver Hudson 1990, 1998). Esto es, si aplicamos la coalescencia a las secuencias de múltiples alelos de un locus particular, se pueden describir las historias evolutivas de esos alelos al rastrear hacia atrás el punto en el cual coalescen (se unen). Así, los procesos de coalescencia en poblaciones naturales “garantizan” las conexiones filogenéticas entre genotipos dentro de una especie, vía sendas verticales de descendencia. La teoría de coalescencia moderna es la forma de evaluar los procesos de ramificación de árboles genealógicos (Hudson 1990, 1998, Nordborg 2000).

Aunque la teoría es complicada, el concepto general de coalescencia es relativamente sencillo: líneas matrilineales semejantes a un ‘pedigree’ (ver Fig. 1) muestran cómo, partiendo de información del presente (haplotipos diferentes), es posible trabajar hacia atrás en el tiempo (n número de generaciones) para reconstruir la historia de linajes genéticos (n número de individuos), dentro de una población en particular. De los tres linajes que se muestran en oscuro en la Figura 1, los haplotipos A y B coalescen primero, mientras que el ACMR de los tres linajes está ubicado más atrás en el tiempo. Eventualmente, todos los alelos dentro de cualquier población (sin considerar migrantes recientes) deberán coalescer a un único alelo ancestral, sin embargo, cuánto tiempo haya tenido que transcurrir dependerá de múltiples factores, pero principalmente del tamaño efectivo de la población (N_e): la relación entre el tiempo de coalescencia y el tamaño poblacional es directa, porque la probabilidad de que un par de secuencias coalescan en un punto –histórico– está en función del tamaño de la población, y mientras más pequeño sea éste, más rápidamente coalescerán.

Si una población tiene un tamaño constante de N_e y los individuos dentro de la población tienen reproducción aleatoria durante cada generación, la probabilidad de que dos haplotipos diferentes ‘elijan’ al mismo ‘padre’ (haplotipo ancestral) en la generación inmediata anterior y coalescan es $1/2N_e$ para un locus nuclear diploide y, en la mayoría de los casos, $1/N_{ef}$ para ADNmt (N_{ef} es el tamaño efectivo de hembras de la población). Asimismo, la probabilidad de que elijan padres

diferentes (i.e. no coalescan) es de $1 - 1/2N_e$ y $1 - 1/N_{ef}$. El tiempo de coalescencia promedio de todos los haplotipos en una población es $4N_e$ y N_e generaciones, para genes diploides y mitocondriales, respectivamente. Si consideramos que en las poblaciones naturales los factores demográficos (e.g. tasa de fecundidad, tasa de mortandad, etc.) determinan los tamaños poblacionales, se entiende que existe una intrincada relación entre demografía y la filogenia de genes, de manera que las variables demográficas son factores decisivos que gobiernan la profundidad, la forma y los patrones demográficos de los árboles de genes. Al mismo tiempo, demografía, filogenia y coalescencia se relacionan íntimamente, de manera que la demografía determina la dinámica de los procesos de ramificación y coalescencia de linajes, como puede apreciarse en la Figura 2: se pueden inferir cambios históricos del tamaño poblacional a partir de secuencias de poblaciones actuales, y

combinando la información de árboles filogenéticos y genealogías es posible determinar si la población se ha mantenido relativamente constante (Fig. 2A), o si, por el contrario, ha habido un crecimiento (Fig. 2B) o decremento (Fig. 2C) exponencial. Una excelente revisión sobre demografía y filogeografía se puede consultar en Avise (2000). Se conoce que el tiempo de coalescencia se ve afectado no solo por N_e , sino por factores como fluctuaciones del tamaño poblacional, selección natural e inmigración, entre otros. Por ello, modelos matemáticos y estadísticos basados en la teoría de coalescencia deben considerar diversos parámetros demográficos, evolutivos y ecológicos (Nordborg 2000).

La coalescencia, junto con análisis demográficos históricos, puede ayudar a determinar la existencia de periodos de expansión o reducción poblacional. Ejemplo de ello son los estudios en los que se determina el tiempo al

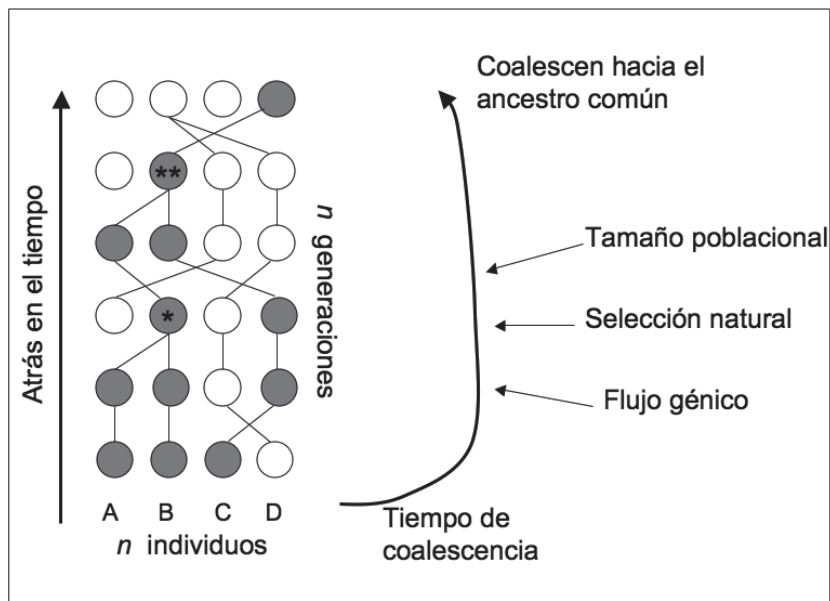


Fig. 1: Relación evolutiva (genealógica) de cuatro haplotipos (individuos) dentro de una población. Los linajes de los haplotipos pueden rastrearse hacia atrás en el tiempo y así identificar los eventos coalescentes. Por ejemplo, los haplotipos A y B tienen su ancestro común más reciente (ACMR) dos generaciones atrás (primer evento coalescente; ancestro señalado con un asterisco), mientras que el ACMR de los haplotipos A, B y C (doble asterisco) se da cuatro generaciones atrás (segundo evento coalescente). Características como tamaño poblacional, selección natural y flujo génico determinan en gran medida los patrones de coalescencia.

Genealogical relationship of four haplotypes (individuals) within a population. Haplotype lineages can be traced back on time to identify coalescent events. For example, haplotypes A and B have their most recent common ancestor (MRCA) two generations back (first coalescent event; ancestor indicated by an *). While the MCRA of haplotypes A, B and C (double asterisk) is four generations back (second coalescent event). Coalescent patterns greatly depend on factors such as population size, natural selection and gene flow.

ancestro común más reciente y además se identifica el comportamiento del tamaño de la población a través del tiempo, como el trabajo con modelos de evolución humana de Templeton (2002). En este caso el autor comparó los patrones de distribución histórica de las poblaciones humanas de acuerdo con varios genes nucleares y mitocondriales (recordemos que cada gen tendrá diferente tiempo de coalescencia esperado según su forma de herencia). Realizó ajustes en los cálculos de tiempo de divergencia para cada gen a partir de nueva evidencia, y pudo

así identificar la existencia de varios eventos de expansión poblacional a partir de África y su dirección (ya sea hacia Asia o el Pacífico). Pudo además describir los procesos asociados con dichos eventos, como aislamiento por distancia, patrones de flujo génico y expansión a través de áreas continuas o discontinuas, y determinar el momento en que posiblemente ocurrieron. También es posible determinar los efectos de procesos vicariantes en la distribución de la estructura genética de poblaciones y especies (Smit et al. 2007).

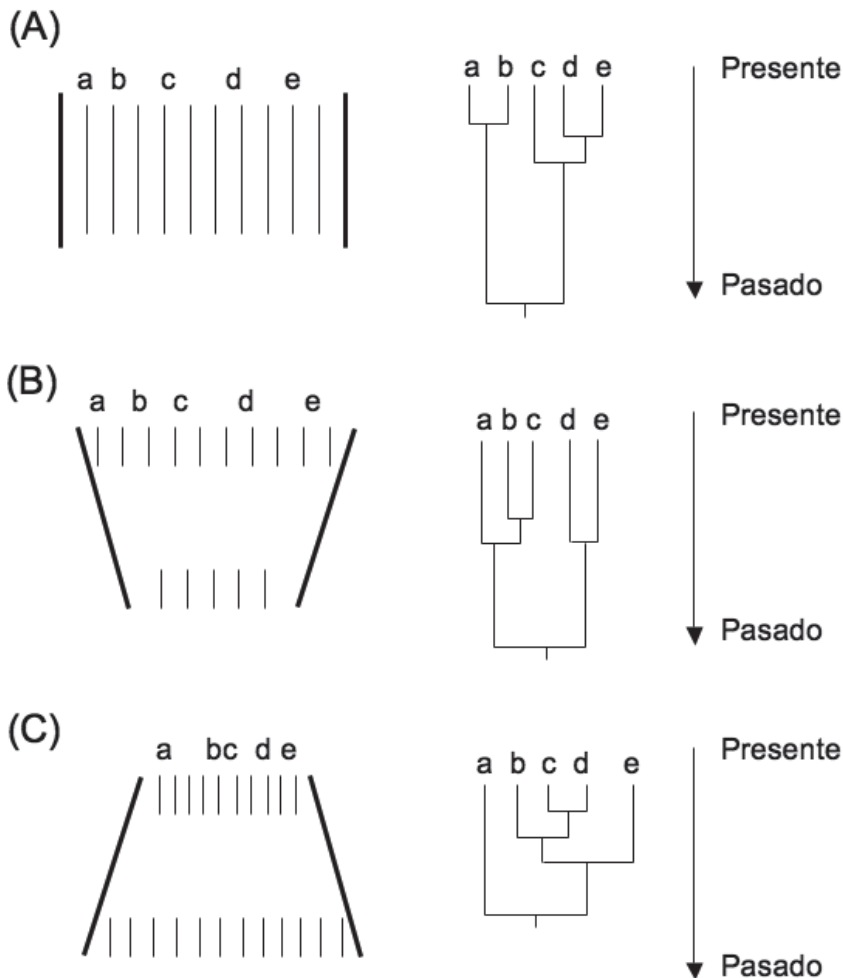


Fig. 2: Se muestran genealogías hipotéticas de tres poblaciones, con sus árboles filogenéticos correspondientes, de (A) población de tamaño constante, (B) población con crecimiento exponencial y (C) población con decremento exponencial. En los tres casos, las poblaciones actuales son todas del mismo tamaño ($n = 10$), de las cuales se denotan cinco haplotipos (a, b, c, d, y e). Las diferencias en la topología de los árboles son resultado de la selección ('muestreo') de haplotipos de poblaciones con dinámicas diferentes (ver Emerson et al. 2001).

Hypothetical genealogies are shown for three populations, with their corresponding phylogenetic tree, for (A) a population with constant size, (B) with exponential growth, and (C) with exponential decrease. Populations from all three cases have the same size ($n = 10$), from which five haplotypes are shown (a, b, c, d, and e). Differences in tree topology are the result of the selection ('sampling') of haplotypes from populations with different dynamics (see Emerson et al. 2001).

Peters et al. (2005) utilizaron un método de coalescencia (aislamiento con migración) para determinar, en el pato de bosque *Aix sponsa* (Linnaeus, 1758), si existía diferenciación entre las poblaciones del este y las del oeste de Norteamérica. Para ello calcularon el tamaño poblacional, el tiempo de divergencia y el flujo genético materno. Los resultados mostraron una clara diferenciación entre las poblaciones ($\Phi_{st} = 0.31$), así como un tiempo de divergencia de al menos 10.000 años, resultado principalmente del escaso flujo génico entre ellas y por los distintos eventos de glaciación, en los que al parecer las poblaciones han ocupado al menos dos refugios durante la última glaciación del Pleistoceno tardío. Finalmente, encontraron que el tamaño poblacional es mayor en el este –además de que presenta una expansión reciente–, contrario a lo que sucede en la población del oeste que mantiene un tamaño pequeño semejante a la población ancestral y no muestra expansión reciente.

MÉTODOS DE ANÁLISIS

Los métodos moleculares actuales permiten no solo la estimación de las frecuencias alélicas, sino también de las relaciones genealógicas entre alelos, con lo cual es posible considerar el flujo génico histórico y la fragmentación de poblaciones en un marco conceptual más evolutivo y geográfico, ya que se pueden comparar las relaciones evolutivas de los linajes genéticos con su ubicación geográfica (Freeland 2005). Sin embargo, dado que los métodos tradicionales de la genética de poblaciones no consideran los factores históricos, dichos métodos deben adecuarse para poder hacer uso de la información histórica contenida en los estimadores moleculares. Wright (1931, 1943) introdujo los estadísticos de F , que han permitido cuantificar la subdivisión –estructura– de las poblaciones, así como el flujo génico, a partir de las frecuencias alélicas obtenidas en diferentes localidades geográficas. Actualmente, al poder utilizar datos de secuencias de ADN y representarlos en un árbol de alelos o haplotipos, se tiene no sólo la escala geográfica, sino una dimensión más: el tiempo evolutivo. Así, mientras que los estadísticos de F no hacen uso de la información temporal,

nuevos procedimientos estadísticos basados en la información de ADN pueden separar los eventos históricos de aquellos a nivel de la estructura poblacional (Hudson et al. 1992, Templeton et al. 1995, Templeton 2004).

Como veremos más adelante, se han desarrollado métodos para examinar el efecto de los patrones históricos de migración y dispersión sobre la distribución de genes, los cuales integran la información genealógica y de distribución para hacer inferencias sobre los patrones históricos de flujo génico (e.g. Hudson 1990, Templeton et al. 1995, Avise 2000, Cruzan & Templeton 2000, Templeton 2004). En particular, modelos de genética de poblaciones basados en la teoría de coalescencia proveen un marco conceptual estadístico para el análisis de, entre otros, la información filogeográfica (e.g. Modelo de Niegel; Avise 2000), lo que ha llevado a compilar dicha información metodológica bajo el concepto de ‘estadística filogeográfica’ (Knowles & Maddison 2002).

El análisis filogeográfico parte, en su nivel más fino tal vez, de qué tan estructuradas (diferenciadas) están las poblaciones; sin embargo, y como ya hemos mencionado antes, el enfoque filogeográfico tiene que analizar información tanto espacial como temporal, en escalas históricas (evolutivas) y ecológicas. Por lo tanto, en filogeografía es necesario primero establecer las relaciones genealógicas entre haplotipos, y posteriormente identificar qué factores históricos y geográficos pudieron haber influido o determinado la distribución actual de los haplotipos. Para ello, se utilizan básicamente dos métodos de análisis, uno fundamentalmente gráfico, basado en árboles de genes o en coalescencia, y otro que considera estadísticos de resumen y parámetros demográficos (Emerson et al. 2001, Posada & Crandall 2001, Knowles & Maddison 2002, Hey & Machado 2003; Tabla 1).

Métodos gráficos (basados en árboles de genes)

Métodos filogenéticos tradicionales

En filogeografía se aplican los mismos algoritmos que se emplean para reconstruir árboles de especies o taxones superiores; sin embargo, como en el caso particular de la filogeografía, las OTUs del árbol filogenético

TABLA 1

Resumen de los métodos de análisis filogeográfico incluidos en la presente revisión, donde se señalan las ventajas y desventajas principales, así como las referencias más importantes. Para una explicación más extensa de cada método, ver el texto.

Summary of the phylogeographical methods of analyses included in the present review, where main advantages and disadvantages are indicated, with their corresponding references. For a more extensive explanation of each method refer to the main text.

Método	Ventajas	Desventajas	Referencias
Métodos gráficos (basados en árboles de genes)			
1. Métodos filogenéticos tradicionales	Se requiere conocer tiempos de divergencia. Describe relaciones interespecíficas, jerárquicas. Es posible utilizar información de genes nucleares.	Evalúan relaciones jerárquicas y no genealógicas (haplotipos). No permite visualizar frecuencias ni pasos mutacionales	Avice 2000, Freeland 2005, Hawkins 2006, Kelchnes & Thomas 2006, Eguarte et al. 2007, Spellman & Klicka 2007
2. Redes de haplotipos	Describe las relaciones intraespecíficas en una red, lo que permite describir relaciones reticuladas entre haplotipos (y no jerárquicas). Incluye la frecuencia gráfica de los haplotipos e infiere haplotipos faltantes.	No permite visualizar distancias genéticas en forma gráfica. No describe relaciones evolutivas. Límite de confianza sensible al número y la distribución geográfica de las muestras.	Templeton 1992, Clement 2000, Posada & Crandall 2001, Muwakina et al. 2003, Excoffier et al. 2005
3. Análisis de clados anidados	Permite tener análisis estadísticos de la relación geográfica y genealógica. Permite evaluar si el muestreo es adecuado.	Muy sensible a la resolución genética. Interpretaciones limitadas por el tamaño de muestra. No considera límites de confianza.	Crandall & Templeton 1993, Templeton & Sing 1993, Templeton et al. 1995, Templeton 1998, Cruzan & Templeton 2000, Posada et al. 2000, Frutos & Van Den Bussche 2002, Knowles & Maddison 2002, Templeton 2004, Hasbún et al. 2005, Posada et al. 2006
Métodos gráficos (basados en coalescencia)			
1. Gráficas de linajes a través del tiempo	Permite inferir cambios históricos de tamaño poblacional. Incluye análisis estadístico.	Sensible a delimitaciones genéticas y taxonómicas en la definición de las especies. Error al identificar especies hermanas por falta de muestreo. Problemas por baja variabilidad genética, persistencia de haplotipos ancestrales y multibifurcaciones debido a descendientes de varios linajes.	Emerson et al. 2001, Hawkins 2006
2. Gráficas de líneas de cielo	Permite utilizar el tamaño efectivo de la población (N_e). Mayor poder discriminante entre grupos.	Es necesaria una alta variabilidad, ya que se requiere conocer el tiempo de divergencia e inferir un árbol de genes. Asume reloj molecular y muestreo al azar.	Emerson et al. 2001, Strimmer & Pybus 2001, Holmes et al. 1995, 1999, Burbrink et al. 2008
Estadísticos de resumen			
1. Distribución de diferencias apareadas	Se desarrollaron a partir de modelos neutrales. Utilizan la tasa de mutación poblacional (θ). Permite estimar cambios en el tamaño poblacional histórico (e.g. distribución mismatch, D de Tajima).	No considera la estructura genealógica de los datos. Se ajustan a un modelo demográfico (e.g. modelos de islas, de aislamiento por distancia) y no todos pueden evaluarse estadísticamente.	Tajima 1983, Fu 1994
2. Parámetros demográficos basados en coalescencia	Incluyen estimadores de máxima verosimilitud (e.g. Monte Carlo, Cadenas de Markov). Se puede obtener una medida de incertidumbre. Permite estimar tasa de migración, crecimiento poblacional, identificar al ancestro común más reciente, determinar la edad y ubicación de subpoblaciones ancestrales.	Requieren de un trabajo computacional intensivo.	Bahlo & Griffiths 2000, Emerson et al. 2001

son los haplotipos, se obtiene un árbol de genes, que se define como un diagrama ramificante que describe el patrón de ancestría entre secuencias homólogas de ADN de diferentes individuos de una población o una especie (Hey & Machado 2003), y donde las ramas de diferentes longitudes expresan la cantidad de cambio evolutivo (en términos, por ejemplo, de número de pasos mutacionales, número de sustituciones). Las variantes o haplotipos del ADNmt registran la historia matrilineal de eventos mutacionales (Fig. 3A), y pueden conectarse filogenéticamente en un filograma (Fig. 3B), el cual puede a su vez

sobreponerse a la distribución geográfica del grupo de estudio (Fig. 3C), a fin de tener una primera –y meramente gráfica– aproximación a los procesos evolutivos responsables de la distribución actual de la especie o grupo de especies (Avice 2000).

Aunque los árboles de genes, que describen el patrón de ancestría del ADN de una población, son diferentes de los árboles filogenéticos, que se refieren al patrón de ancestría de un taxón, los primeros pueden analizarse con los métodos filogenéticos tradicionales, como parsimonia, o aquellos que involucran distancias genéticas o la

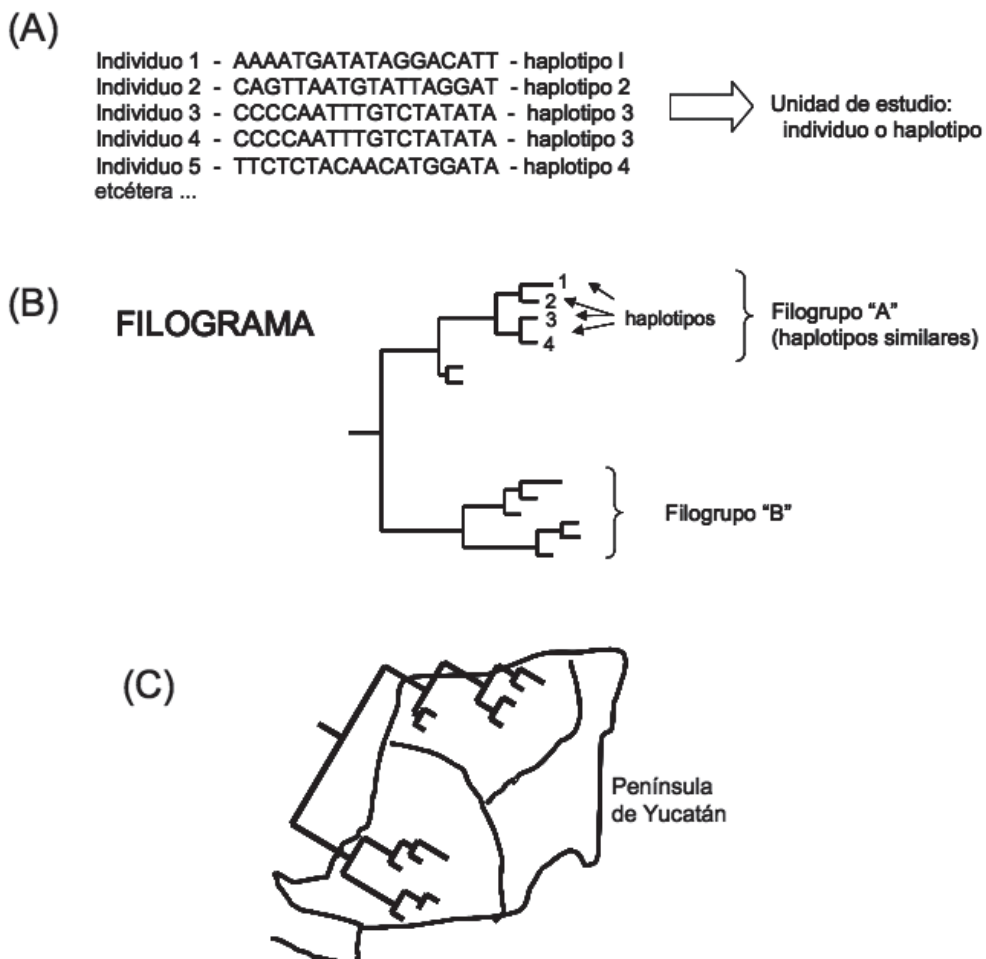


Fig. 3: (A) haplotipos individuales (secuencias de ADN mitocondrial) de una especie, (B) que se conectan filogenéticamente en un filograma; grupos de haplotipos diferenciados (monofiléticos) se denominan filogrupos, (C) el filograma puede sobreponerse a la distribución geográfica del grupo de estudio.

(A) individual haplotypes (mitochondrial DNA sequences) from a species, (B) that can be connected phylogenetically in a phylogram; differentiated haplotype groups (monophyletic) are called phylogroups, (C) the phylogram can be overlaid onto the geographical distribution of the group of study.

posibilidad de aplicación de modelos evolutivos específicos, como neighbor-joining, mínima evolución, máxima verosimilitud y métodos bayesianos. A través de estos análisis filogenéticos es posible determinar filogrupos diferentes (Fig. 3); si existe concordancia entre la distribución geográfica de los haplotipos y los haplogrupos reflejados en la filogenia, dicha información sirve de apoyo (o como base) de otros análisis que evalúen directamente concordancia (e.g. clados anidados; ver más adelante). Los análisis filogenéticos permiten estimar tiempos de divergencia, información que es de vital importancia en filogeografía (Tabla 1). Spellman & Klicka (2007) utilizaron dos criterios filogenéticos (parsimonia y máxima verosimilitud) para evaluar hipótesis históricas sobre la evolución de *Sitta carolinensis*, ave que se distribuye ampliamente en Norteamérica. Obtuvieron un filograma con cuatro filogrupos monofiléticos (Este, Pacífico, este de la sierra Nevada y las montañas Rocallosas, la Gran Cuenca y México), filogrupos sorprendentemente concordantes con la distribución de los bosques de encino-pino. Los resultados apoyan la hipótesis de que esta especie evolucionó dentro de este tipo de hábitat, contrario a otra hipótesis propuesta sobre una expansión reciente, y que el aislamiento posterior de sus poblaciones ha resultado en una diferenciación significativa entre poblaciones y variación genética críptica.

Existe variada literatura que puede revisarse para profundizar sobre métodos filogenéticos; en particular sobre su aplicación en filogeografía ver Freeland (2005), Hawkins (2006), Kelchner & Thomas (2006), Eguiarte et al. (2007).

Redes de haplotipos

En estudios filogeográficos se utilizan con mayor frecuencia métodos de reconstrucción filogenética que consideran la coalescencia, los cuales se emplean ya sea en combinación con los filogenéticos tradicionales o, por el contrario, como el único análisis de estudio (Tabla 1). Así, a diferencia de los árboles bifurcados tradicionales, estos métodos permiten describir las relaciones evolutivas en forma de árboles multibifurcantes, en los cuales un único

haplotipo puede dar origen a muchos haplotipos, creando lo que se conoce como 'redes', particularmente como redes mínimas de haplotipos (conexiones entre haplotipos dadas por el mínimo número posible de mutaciones entre ellos). Dado que se trabaja con filogenias de haplotipos a nivel intraespecífico (y no a nivel interespecífico), las cuales casi no presentan recombinación, pueden tener muy poca variación e implican un tiempo evolutivo mucho menor, Templeton et al. (1992) propusieron un método que toma en cuenta estas particularidades, conocido como parsimonia estadística. Clement et al. (2000) desarrollaron el programa de cómputo TCS ("Phylogenetic network estimation using statistical parsimony") que permite estimar redes filogenéticas a partir de secuencias de ADN o de matrices de distancias nucleotídicas utilizando el algoritmo de Templeton et al. (1992). Este algoritmo estima el número máximo de diferencias de pares de bases entre secuencias que pueden atribuirse a una serie de mutaciones únicas en cada sitio de la secuencia; número al que se conoce como límite de parsimonia (con una confianza del 95 %). El programa no conecta a la red los haplotipos que difieren por un número de pares de bases que exceda dicho límite, por el supuesto de posible homoplasia. Sin embargo, el límite puede excederse por otras razones, como un tamaño pequeño de muestra (e.g. número de haplotipos), o un muestreo sesgado geográficamente (i.e. falta de haplotipos o huecos en la distribución del taxón de estudio). Una vez que se estima el límite de parsimonia, el algoritmo conecta los haplotipos que difieren por una sola mutación, seguido de los que difieren por dos mutaciones, luego tres y así sucesivamente, hasta que se alcance el límite o hasta que se unan todos los haplotipos.

El resultado final es una red que muestra las relaciones entre todos los haplotipos con base en el número mínimo de mutaciones (los pasos a seguir para el uso del programa se pueden ver en Clement et al. 2000). Las conexiones entre los haplotipos representan eventos coalescentes por lo que pueden hacerse, a partir de la red de parsimonia y con base en los principios de coalescencia, las siguientes predicciones (ver Clement et al. 2000) (Fig. 4): i) los haplotipos con mayor frecuencia tienen alta probabilidad de ser alelos ancestrales, ii) los haplotipos que tienen múltiples conexiones muy probablemente

son ancestrales, iii) dentro de la red, los haplotipos ancestrales son de interior y los haplotipos nuevos tienen mayor probabilidad de ser periféricos ('de punta'), iv) se espera que los haplotipos ancestrales tengan una amplia distribución geográfica (básicamente porque los

individuos que tienen esos haplotipos han tenido mucho tiempo para dispersarse), y v) los haplotipos con una única conexión se prevé estén unidos a haplotipos de la misma población (dado que han evolucionado recientemente y por lo tanto sin tiempo para dispersarse).

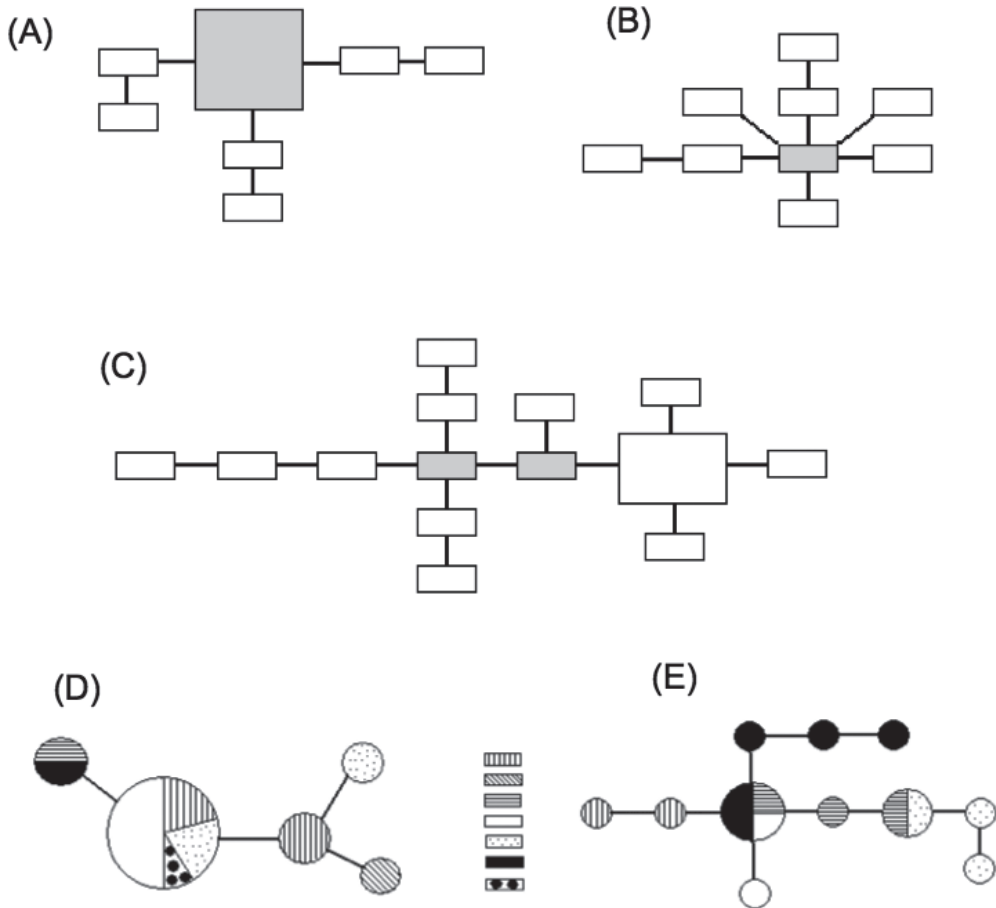


Fig. 4: Red mínima de haplotipos en las que se muestran las conexiones entre haplotipos, unidos por un paso mutacional. El tamaño de los rectángulos es proporcional a la frecuencia de los haplotipos en la población. Se muestran las diferentes predicciones que pueden hacerse a partir de la red: (A) los haplotipos con mayor frecuencia tienen alta probabilidad de ser alelos ancestrales (rectángulo sombreado), (B) los haplotipos que tienen múltiples conexiones muy probablemente son ancestrales (rectángulo sombreado), (C) dentro de la red, los haplotipos ancestrales son de interior (sombreados) y los haplotipos nuevos tienen mayor probabilidad de ser periféricos ('de punta'), (D) se espera que los haplotipos ancestrales tengan una amplia distribución geográfica (haplotipo central, con alta frecuencia en cuatro diferentes localidades), y (E) los haplotipos con una única conexión se prevé estén unidos a haplotipos de la misma población (círculos oscuros).

Haplotype minimum network where connections among haplotypes, joined by one mutational step, are shown. Rectangle size is proportional to haplotype frequency in the population. The different predictions that can be asserted from the network are shown: (A) haplotypes with the highest frequency have more probability of being ancestral alleles (gray rectangle), (B) haplotypes with multiple connections are most probably ancestral (gray rectangle), (C) within the network, ancestral haplotypes are interior (shaded) and new haplotypes have higher probability of being peripheral ('point haplotypes'), (D) ancestral haplotypes are expected to have ample geographical distribution (central haplotype with high frequency in four different locations), and (E) haplotypes with only one connection are expected to be connected with haplotypes from the same population (dark circles).

Un ejemplo donde se utilizó una red mínima de haplotipos es un estudio con el jabalí verrugoso *Pacochoerus africanus* (Paillas, 1766) en África (Muwakina et al. 2003). Los resultados mostraron tres clados altamente diferenciados entre sí: un clado Oeste, separado de un clado Sur por 17 pasos mutacionales y un clado Este separado del clado Sur por 10 pasos, además de estructuración a escala más fina con haplotipos que divergen dentro de los clados mismos. Los autores concluyen que dicha estructura es resultado de las fluctuaciones ambientales del Plioceno-Pleistoceno, que ocasionaron que las poblaciones de jabalí se aislaran en tres refugios diferentes.

Pueden utilizarse otros algoritmos aparte de TCS, siempre y cuando sea posible construir redes de haplotipos. Por ejemplo, Cassens et al. (2003) en un estudio con el delfín *Lagenorhynchus obscurus* (Gray, 1828) del hemisferio Sur, utilizaron diferentes formas de estimar redes, e.g. 'minimum spanning network' (con el software Arlequin), y los métodos de 'split decomposition' (con SplitsTree) y 'median-joining networks' (con Network) (Huson 1998, Excoffier et al. 2005). Con esta combinación de análisis encontraron una diferenciación significativa entre los haplotipos del Atlántico (Perú) de los del Pacífico (Argentina y Sudáfrica).

Análisis de clados anidados

Para poder estimar los factores históricos y geográficos que han determinado la distribución actual de haplotipos, en filogeografía se utiliza con considerable frecuencia el análisis de clados anidados (NCA, por sus siglas en inglés; Templeton et al. 1995). Su popularidad descansa en tres razones principales: i) permite hacer un análisis estadístico objetivo con el que se puede evaluar la hipótesis nula de que los haplotipos de los clados se distribuyen geográficamente al azar (i.e. no asociación entre la variación haplotípica y la geografía), ii) permite detectar y probar estadísticamente los mecanismos evolutivos responsables de la distribución espacial de los patrones de variación genética observada, y iii) permite interpretar los patrones significativos utilizando criterios explícitos, con los que es posible evaluar si el muestreo es adecuado para detectar una

asociación significativa entre los clados y la geografía, y si es adecuado además para interpretar biológicamente dicha asociación (Tabla 1).

El primer paso para hacer un NCA es el contar con un red de haplotipos como las explicadas arriba; posteriormente se utilizan reglas específicas para definir una serie de clados anidados jerárquicamente dentro de la red (Reglas de anidación; Templeton & Sing 1993, Templeton 2004) (Fig. 5). En el primer nivel se definen los 'clados de un paso', que agrupan a los haplotipos más exteriores (haplotipos de punta) con el haplotipo inmediato anterior separado por una sola mutación (haplotipos de interior); estos clados de un paso se anidan en 'clados de dos pasos', que contienen al clado de nivel uno unido al haplotipo inmediato anterior separado por una sola mutación, y así sucesivamente. Se continúa el anidamiento hasta que el siguiente nivel de anidamiento más alto resulta en un único clado que contiene a toda la red (Fig. 5). Como resultado, se tiene que los niveles más altos del anidamiento corresponden a los eventos coalescentes más antiguos y viceversa.

Posteriormente en el NCA se analizan en forma conjunta el cladograma anidado y su distribución geográfica, considerando las frecuencias de cada uno de los haplotipos así como dos medidas de distancia: distancia del clado (D_c) que representa la distancia promedio de los miembros del clado al centro geográfico de ese clado (es decir, la dispersión geográfica de los haplotipos de un clado); y distancia del clado anidado (D_n) que mide la distancia promedio de los miembros del clado anidado al centro geográfico del clado anidado (es decir, cómo se distribuye un clado en relación con otros clados en el mismo nivel de anidamiento); se estima también una medida de interior-punta dentro de cada categoría anidada, como la diferencia entre la distancia interior promedio menos la distancia de punta promedio (clado de interior y clado de punta en cada nivel jerárquico; ver Crandall & Templeton 1993, Cruzan & Templeton 2000 y Posada et al. 2006 para detalles de los cálculos). La evaluación estadística de estas medidas de distancia se realiza por medio de un procedimiento de Monte Carlo, que son algoritmos usados en métodos estocásticos de simulación computacionalmente intensivos, que se basan

en un muestreo repetido y aleatorio de datos simulados obtenidos a partir de los datos empíricos (Luikart & England 1999). En el caso específico del NCA esto se hace a través de la construcción de distribuciones nulas: se construye una tabla de datos de contingencia para cada clado y nivel de anidamiento, las cuales se aleatorizan y de las cuales se vuelven a calcular las medidas de distancia, con las que se comparan iterativamente los datos empíricos. Existe una forma de hacer estos cálculos de manera automatizada, con el programa GeoDis escrito por Posada y colaboradores (Posada et al. 2000, 2006). Los resultados que da el programa incluyen las comparaciones (D_c y D_n) que son significativamente diferentes de lo esperado por el azar (junto con los valores de significancia obtenidos con pruebas de chi-cuadrado)

El objetivo de estimar aquellas asociaciones significativas entre haplotipos y su distribución geográfica es poder aceptar o rechazar –estadísticamente– la hipótesis de no asociación entre la variación haplotípica y la geografía. Ello se hace analizando los

resultados obtenidos con GeoDis por medio de una clave de inferencia (‘clave de Templeton’; Templeton 2004) para determinar cuál o cuáles escenarios explican más acertadamente los patrones observados en la distribución de los haplotipos; se parte de tres procesos posibles: i) aislamiento por distancia, debido a flujo génico restringido, ii) fragmentación histórica, y iii) expansión del área de distribución de las poblaciones, incluyendo eventos de colonización a distancia. La descripción detallada del análisis de clados anidados, así como las herramientas para llevarlo a cabo, se incluyen en Templeton et al. (1995), Templeton (1998) y Templeton (2004). Existe ya un programa con el que puede hacerse todo este procedimiento de manera automatizada (ANeCA; Panchal 2007).

El NCA tiene ciertas limitaciones que deben considerarse, como que es sensible a la resolución genética para poder detectar los procesos históricos, además de que las interpretaciones biológicas que pueden hacerse están significativamente limitadas por el tamaño de muestra (individuos por población) y

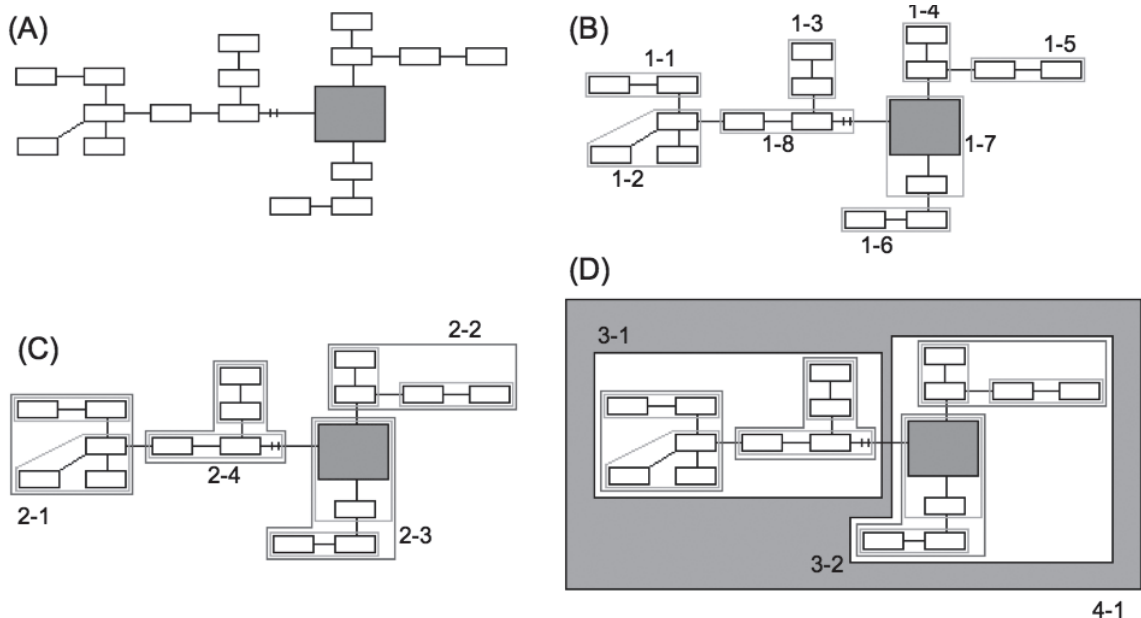


Fig. 5: Diferentes niveles de anidación de (A) una red mínima de haplotipos. (B) clados de un paso: en el primer nivel, se anidan (‘unen’) haplotipos de punta a su haplotipo inmediato anterior separado por un solo paso mutacional, seguidos de los clados de segundo (C) y tercer nivel (D), hasta que el nivel final incluye a toda la red en un único cuarto nivel (D).

Different nesting levels of (A) a minimum haplotype network. (B) one-step clades: on the first level, point haplotypes are nested (‘joined’) to their immediate anterior haplotype, separated by one mutational step, followed by second (C) and third level clades (D), until the last level includes the whole network in a unique fourth level clade (D).

por la amplitud del muestreo (sitios muestreados que deben cubrir adecuadamente la distribución geográfica de la especie de estudio), así como por el hecho de que la clave de inferencia solo provee respuestas de 'sí' y 'no', sin considerar límites de confianza (Knowles & Maddison 2002, Templeton 2004). Existe una fuerte controversia alrededor de este método, lo que ha generado una serie de comunicaciones entre autores y opositores (ver Petit 2008 y referencias incluidas), particularmente por quienes señalan que el NCA no es confiable porque resulta en un alto porcentaje de falsos positivos y de falsos negativos. Sin embargo, Templeton (2008) ha demostrado que dichos porcentajes son mucho menores cuando se evalúan casos reales, habiendo validado el método con una enorme base de datos de controles positivos, en comparación a cuando se prueban a través de datos no reales, sino obtenidos con modelos matemáticos.

Un ejemplo de análisis que integra una red mínima de haplotipos y el posterior NCA es el estudio de Frutos & Van Den Bussche (2002) con poblaciones de armadillos de Paraguay y secuencias de citocromo b. En este caso, obtuvieron la frecuencia de los haplotipos, hicieron la red mínima y el análisis de clados y posteriormente utilizaron la clave de Templeton. El análisis permitió determinar que el muestreo es inadecuado para ciertos clados (localidades geográficas), que uno de los clados se encuentra en expansión y otro tiene flujo génico restringido. Las conclusiones del estudio incluyen la identificación de movimiento por parte de las hembras y el efecto de un río como barrera a la dispersión. Con la misma metodología, Hasbún et al. (2005) analizan la filogeografía de una lagartija de América Central y muestran la secuencia de colonización de las poblaciones del género *Ctenosaura* a lo largo de su rango de distribución, atribuible a la historia geológica de la región. A partir de la clave de Templeton sugieren procesos de fragmentación en el pasado para el cladograma completo (i.e. todo el rango de distribución), además de eventos de fragmentación y expansión para los clados que lo integran (i.e. algunas subpoblaciones).

Después de esta explicación de los métodos gráficos basados en árboles, es importante mencionar que una de las limitantes más

importantes de dichos métodos es que la mayoría de las historias demográficas dan como resultado árboles de genes que varían grandemente en cuanto a topología y longitud de ramas, varianza conocida como varianza estocástica. Dicha varianza está asociada a la variación inherente observada en las historias genealógicas obtenidas con diferentes genes (e.g. nucleares versus mitocondriales), que puede dar como resultado historias diferentes para una misma población. Otro problema de los métodos basados en árboles es que están limitados por la forma de estimación del árbol de genes y no se puede siempre asegurar que represente la verdadera historia de la población, por ejemplo debido a mutación recurrente en el ADNmt o por recombinación en genes nucleares (Hey & Machado 2003). Para lidiar con estas limitantes se ha hecho cotidiano, por un lado, el uso de diferentes análisis en un mismo estudio (por ejemplo, métodos filogenéticos junto con clados anidados y estimadores de genética de poblaciones y estructuración genética), y por otro, ha motivado el desarrollo de otros métodos directamente enfocados a descifrar las historias demográficas de las poblaciones, ya sea gráficos basados en coalescencia y matemáticos basados en estadísticos de resumen demográfico, como veremos a continuación.

Métodos gráficos (basados en coalescencia)

Gráficas de linajes a través del tiempo

Con análisis de coalescencia se ha demostrado que secuencias homólogas de una población contienen información sobre la historia demográfica de esa población. Cuando dichas secuencias se analizan a través de un árbol genealógico, cada nodo interno corresponde, por definición, a un evento de coalescencia, y por ello se pueden inferir cambios históricos en el tamaño poblacional a partir de la distribución temporal de los nodos en la genealogía. Así, se puede graficar la distribución acumulada de los eventos coalescentes contra el tiempo (Fig. 6A) y la forma de la curva resultante –o gráfica de linajes a través del tiempo (LTT por sus siglas en inglés)– se puede usar para probar hipótesis específicas de la historia de la población, las cuales se derivan de las curvas esperadas bajo coalescencia (simulaciones); posteriormente se

evalúa el ajuste entre las curvas teóricas obtenidas con LTT y la curva observada con los datos empíricos, para poder elegir entre las hipótesis alternativas (ver Emerson et al. 2001 y Hawkins 2006 para una revisión de este método; Tabla 1).

Gráficas de líneas de cielo ('skylines')

En las gráficas skyline, generalizadas o no, se utiliza el tamaño efectivo de la población (N_e) en lugar de los eventos coalescentes, por lo que son más fáciles de interpretar que los LTT. Para ello, se estima la media armónica del tamaño efectivo para cada intervalo internodos, medidos en eventos mutacionales que corresponden a periodos de tiempo (M_i), la cual se grafica contra el tiempo (Fig. 6B) para tener una representación no paramétrica de tamaño efectivo de la

población a lo largo del tiempo. Finalmente se ajustan modelos demográficos específicos y se estiman los parámetros correspondientes utilizando máxima verosimilitud (Emerson et al. 2001, Strimmer & Pybus 2001). La ventaja de los skylines generalizados es que pueden aplicarse en casos en los que el árbol genealógico no está totalmente resuelto y/o los datos no sean muy variables (Tabla 1).

Estos métodos han sido utilizados, por ejemplo, en estudios de HIV-1: las gráficas skyline han mostrado que en la población humana el subtipo A ha tenido un aumento exponencial a tasa constante, y que el crecimiento ha sido logístico para el subtipo B, mientras que los LTT indican crecimiento exponencial para ambos subtipos. Estos resultados sugieren que para el HIV-1 los skylines tienen mayor poder discriminante

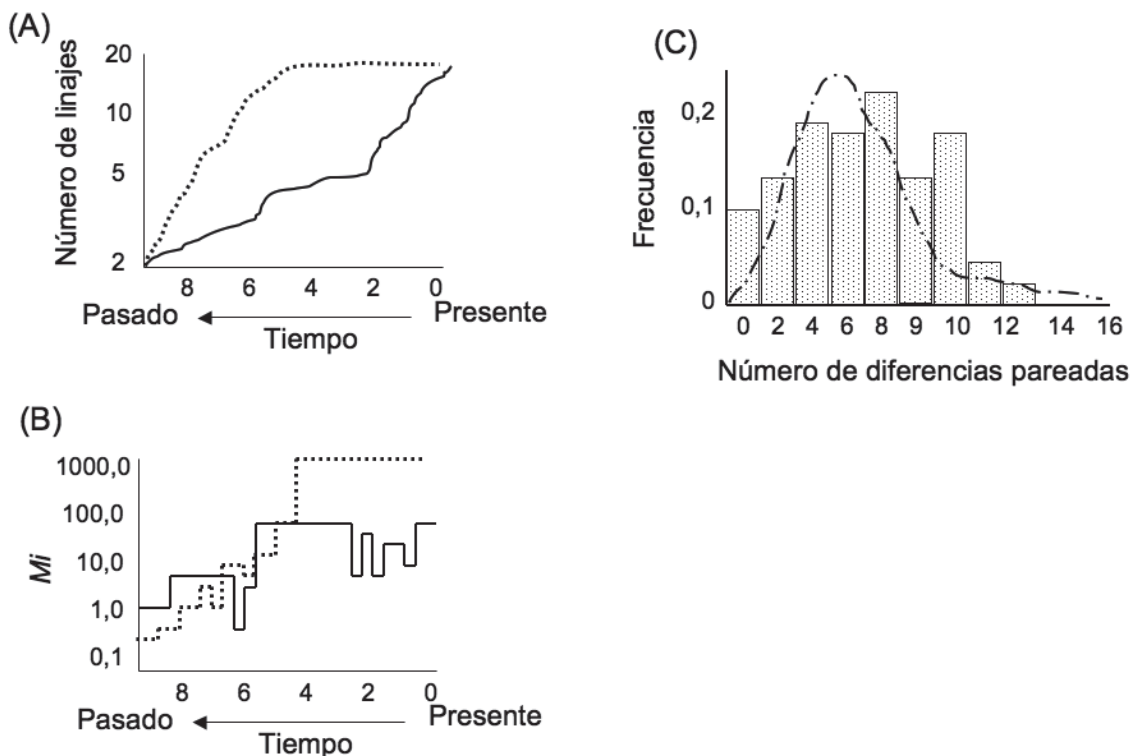


Fig. 6: (A) gráfica de linajes a través del tiempo (LTT) y (B) gráfica 'skyline', ambas en escala logarítmica, donde la línea punteada muestra una población en crecimiento y la línea continua una población de tamaño relativamente constante a lo largo del tiempo. (C) en la gráfica 'mismatch' se muestra la distribución de frecuencias del número de diferencias pareadas entre secuencias (barras) y la curva teórica esperada de una población con expansión poblacional.

(A) lineage-through-time plot (LTT) and (B) skyline plot, both in logarithmic scale, where the dotted line indicates a growing population and the continuous line shows a population with constant size through time. (C) mismatch graph showing the frequency distribution of the number of pairwise differences between sequences (bars) and the expected theoretical curve from a population with exponential growth.

(Holmes et al. 1995, 1999). En otro ejemplo, Burbrink et al. (2008) pudieron determinar con este método cuáles poblaciones de la serpiente *Coluber constrictor* (Linnaeus, 1758) ampliamente distribuida en Norteamérica, tuvieron un crecimiento poblacional hace alrededor de 200.000 años, y que no se vieron afectadas por los ciclos glaciales posteriores.

Estadísticos de resumen (distribución de diferencias pareadas)

Los métodos que consideran un esquema estadístico para analizar secuencias de ADN se desarrollaron a partir de modelos neutrales (sin selección natural), desarrollados previo al advenimiento de las secuencias y la coalescencia, pero que son directamente intercambiables. Los estadísticos de resumen más comúnmente utilizados se enfocan en la tasa de mutación poblacional, denotada típicamente como θ , que es igual a:

$$\theta = 4Nu$$

y donde N es el tamaño poblacional y u es la tasa de mutación neutral del gen en cuestión. Así, el número de mutaciones neutrales que separa a dos secuencias es $4N \times u$, en un organismo diploide. Dicho parámetro puede estimarse usando el número de sitios polimórficos en la muestra (S), el número promedio de diferencias pareadas entre las secuencias (k) o el número de mutaciones singletons (η) (Fu 1994). Con base en este estadístico es posible estimar cambios en el tamaño poblacional histórico, evaluando si los datos empíricos se ajustan a un modelo de tamaño poblacional constante. Para ello se utilizan comúnmente pruebas de neutralidad, entre las cuales tenemos los estimadores de Fu & Li, D^* y F^* , que consideran S y k como estimadores no sesgados de θ (Fu & Li 1993), y la D de Tajima (Tajima 1983), que se basa en la diferencia entre S y k. Otra forma comúnmente utilizada para estas estimaciones es la conocida como 'distribución mismatch', un procedimiento basado en distancias genéticas (Rogers & Harpending 1992) que evalúa si la distribución observada de S o k en la población de estudio es estadísticamente diferente de la esperada bajo la hipótesis de crecimiento poblacional (Fig. 6C).

La principal limitante de estos estadísticos de resumen es que hacen uso solo de una parte de la información posible, ya que no consideran la estructura genealógica de los datos; además dependen del modelo demográfico que se elige como al que mejor se ajustan los datos empíricos (e.g. modelo de islas, de aislamiento por distancia), y sólo existe un número determinado de modelos que pueden evaluarse estadísticamente (Tabla 1).

Estadísticos de resumen (parámetros demográficos basados en coalescencia)

Para lidiar con las limitantes de los estadísticos de resumen mencionadas arriba, recientemente se ha desarrollado el uso de estimadores de máxima verosimilitud (Emerson et al. 2001) que incorporan, entre otros:

a) un muestreo 'Metropolis-Hastings' (MMH) junto con la integración de diferentes filogenias. El MMH es un método de Monte Carlo ampliamente utilizado que se basa en una modificación del esquema original de Monte Carlo, pues usa cadenas de Markov para muestrear, de entre los diferentes estados del proceso estocástico bajo estudio, sólo los estados más probables y hacerlo además con una probabilidad mayor. Con ello evita tener que muestrear todos los estados (cuyo número puede ser extremadamente alto y hacer el análisis irrealizable). Así, es posible analizar tanto el árbol genealógico que mejor describe las secuencias como otros menos probables, y tener entonces una medida de incertidumbre, para evaluar los resultados respecto a una tasa de crecimiento (o decremento) exponencial de la población.

b) uso de ecuaciones recursivas para calcular curvas de verosimilitud para el parámetro q. Las ecuaciones se usan para construir una cadena de Markov; se estima la función de verosimilitud para un valor dado de q, al simular de manera independiente y repetida la cadena de Markov y al tomar la media de los valores simulados. Posteriormente se construyen curvas de verosimilitud para el rango de valores de q. Este método ha sido muy útil para analizar poblaciones subdivididas y estimar tasa de migración, crecimiento poblacional, identificación de dónde ocurrió en la población el ancestro común más reciente, así como determinar la edad y ubicación de

subpoblaciones ancestrales (Bahlo & Griffiths 2000). Como puede deducirse entonces, es una aplicación de gran utilidad en estudios filogeográficos.

Existen otras propuestas para el análisis de poblaciones subdivididas que incorporan parámetros demográficos sin embargo; para los propósitos de estudios filogeográficos, los descritos en esta última parte de la presente revisión son hasta ahora las mejores herramientas disponibles. Nuevamente, existe variada literatura que puede consultarse para ahondar tanto como sea necesario en estos métodos y sus aplicaciones. Para una revisión general de la cual partir, y muy importante también, para una descripción de los programas estadísticos y de cómputo disponibles para aplicar los métodos descritos y que suman ya una extraordinariamente larga lista, ver Gaggiotti et al. (1999), Luikart & England (1999), Emerson et al. (2001), Hare (2001), Posada & Crandall (2001), Knowles & Maddison (2002), Hey & Machado (2003), Knowles (2003), Crisp & Cook (2005), Kelchner & Thomas (2006), Posada (2006), Allendorf & Luikart (2007), Eguiarte et al. (2007).

Es importante enfatizar que las preguntas que pueden analizarse bajo un enfoque filogeográfico son increíblemente variadas, desde la evaluación básica de si las poblaciones de una especie tienen un patrón filogeográfico definido, es decir, poblaciones diferenciadas a lo largo de su distribución geográfica y cuya estructuración responde a procesos históricos y demográficos reconocibles, como por ejemplo vicarianza y aislamiento por distancia (Riddle et al. 2000), lo mismo que estudios que involucran evaluaciones taxonómicas (Harris et al. 2000), hibridización (Godinho et al. 2006), poblaciones de origen (Grazziotin et al. 2006, Vázquez-Domínguez et al. 2009), aplicaciones en conservación (Smith et al. 2000), refugios pleistocénicos (Avice & Walker 1998), procesos de colonización (Richardson et al. 2002) y expansión (Lessa et al. 2003), hasta evolución humana (Templeton 2002); y no sólo hay estudios en vertebrados y plantas, pues se han descrito patrones filogeográficos en parásitos (Mejía-Madrid et al. 2007), bacterias (Carbone & Kohn 2001) y hongos (Reeb et al. 2007). Y esta no es una lista exhaustiva.

Las herramientas y métodos de análisis filogeográfico están en pleno auge, y es seguro

que su desarrollo, perfeccionamiento y alcance se verá maximizado en el corto plazo. Indudablemente se enfatizará el desarrollo de análisis que incorporan la coalescencia, en particular se está trabajando en integrar el modelaje de nicho ecológico y de sistemas de información geográfica a los métodos de coalescencia, para poder evaluar conjuntamente la distribución geográfica y la genealogía (ver Carstens & Richards 2008). Se están desarrollando también métodos bayesianos que incorporan coalescencia y permiten estimar demografía histórica (Drummond et al. 2005). Asimismo, los estudios de filogeografía comparada son cada vez más frecuentes; no existe una metodología específica para evaluar la congruencia filogeográfica de diferentes taxa, por lo que esta es un área que también está en pleno auge (Hickerson et al. 2006). Aunque la filogeografía tiene ya más de 20 años, las enormes posibilidades de estudio presentados en esta revisión, se verán maximizadas con los avances metodológicos en puerta para el estudio conjunto de la información genética, genealógica y geográfica en análisis evolutivos y de distribución.

LITERATURA CITADA

- ALLENDORF F & G LUIKART (2007) Conservation and the genetics of populations. Blackwell Publishing, Oxford, United Kingdom. 642 pp.
- AVISE JC, J ARNOLD, RM BALL, E BIRMINGHAM, T LAMB, JE NIEGEL, CA REEB & NC SAUNDERS (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18: 489-522.
- AVISE JC & D WALKER (1998) Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London B* 265: 457-463.
- AVISE JC (2000) Phylogeography. The history and formation of species. Harvard University Press, Cambridge, United Kingdom. 447 pp.
- BAHLO M & RC GRIFFITHS (2000) Inference from gene trees in a subdivided population. *Theoretical Population Biology* 57: 79-95.
- BEDFORD T & DL HARTL (2008) Overdispersion of the molecular clock: temporal variation of gene-specific substitution rate in *Drosophila*. *Molecular Biology and Evolution* 25: 1631-1638.
- BEHLING H & M LICHTER (1997) Evidence of dry and cold climatic conditions at glacial times in tropical southeastern Brazil. *Quaternary Research* 48: 348-358.
- BEHLING H & RRB NEGRELLE (2001) Tropical rain forest and climate dynamics of the Atlantic lowland, Southern Brazil, during the Late Quaternary. *Quaternary Research* 56: 383-389.

- BLEEKER W, C WEBER-SPARENBERG & H HURKA (2002) Chloroplast DNA variation and biogeography in the genus *Rorippa* Scop. (Brassicaceae). *Plant Biology* 4: 104-111.
- BROWN WM (1985) The mitochondrial genome of animals. En: MacIntyre RJ (ed) *Molecular evolutionary genetics*: 95-130. Plenum, New York, USA.
- BROWN WM, M GEORGE & AC WILSON (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences USA* 76: 1967-1971.
- BRUMFIELD RT, P BEERLI, DA NICKERSON & SV EDWARDS (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution* 18: 249-256.
- BURBRINK FT, F FONTANELLA, R PYRON, TQ GUIHER & C JIMÉNEZ (2008) Phylogeography across a continent: the evolutionary and demographic history of the North American racer (Serpentes: Colubridae: *Coluber constrictor*). *Molecular Phylogenetics and Evolution* 47: 274-288.
- CARBONE I & LM KOHN (2001) A microbial population-species interface: nested cladistics and coalescent inference with multilocus data. *Molecular Ecology* 10: 947-964.
- CARSTENS BC & CL RICHARDS (2008) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution* 61: 1439-1454.
- CARVALHO GR (ed) (1998) *Advances in molecular ecology*. IOS Press, Amsterdam, Netherlands. 314 pp.
- CASSENS I, K VAN WAEREBEEK, B BEST, EA CRESPO, J REYES & MC MILINKOVITCH (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): a critical examination of network methods and rooting procedures. *Molecular Ecology* 12: 1781-1792.
- CLEMENT M, D POSADA & KA CRANDALL (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657-1660.
- CRANDALL KA & AR TEMPLETON (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134: 959-969.
- CRISP MD & LG COOK (2005) Do early branching lineages signify ancestral traits? *Trends in Ecology & Evolution* 20: 122-128.
- CRUZAN MB & AR TEMPLETON (2000) Paleogeology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology & Evolution* 15: 491-496.
- DEVITT TJ (2006) Phylogeography of the Western Lyresnake (*Trimorphodon biscutatus*): testing aridland biogeographical hypotheses across the Neartic-Neotropical transition. *Molecular Ecology* 15: 4387-4407.
- DRUMMOND AJ, A RAMBAUT, B SHAPIRO & OG PYBUS (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22: 1185-1192.
- EGUIARTE L, V SOUZA & X AGUIRRE (2007) *La ecología molecular de plantas y animales*. INE/Conabio, México. 592 pp.
- EMERSON BC, E PARADIS & C THÉBAUD (2001) Revealing the demographic histories of species using DNA sequences. *Trends in Ecology & Evolution* 16: 707-716.
- EXCOFFIER L, G LAVAL & S SCHNEIDER (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.
- FREELAND JR (2005) *Molecular ecology*. John Wiley & Sons, Ltd., London, United Kingdom. 388 pp.
- FRUTOS SD & RA VAN DEN BUSSCHE (2002) Genetic diversity and gene flow in nine-banded armadillos in Paraguay. *Journal of Mammalogy* 83: 815-823.
- FU YX (1994) Estimating effective population size or mutation rate using the frequencies of mutations of various classes in a sample of DNA sequences. *Genetics* 138: 1375-1386.
- FU YX & WH LI (1993) Statistical tests of neutrality of mutations. *Genetics* 133: 693-709.
- GAGGIOTTI OE, O LANGE, K RASSMANN & C GLIDDONS (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* 8: 1513-1520.
- GARCÍA-MORENO J, AG NAVARRO-SIGÜENZA, AT PETERSON & LA SÁNCHEZ-GONZÁLEZ (2004) Genetic variation coincides with geographic structure in the common bush-tanager (*Chlorospingus ophthalmicus*) complex from Mexico. *Molecular Phylogenetics and Evolution* 33: 186-196.
- GODINHO R, B MENDONÇA, EG CRESPO & N FERRAND (2006) Genealogy of the nuclear b-fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. *Heredity* 96: 454-463.
- GRAZZIOTIN FG, M MONZEL, S ECHEVERRIGARAY & SL BONIATO (2006) Phylogeography of the *Bothrops jararaca* complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. *Molecular Ecology* 15: 3969-3982.
- HARE MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* 16: 700-706.
- HARRIS D, DS ROGERS & J SULLIVAN (2000) Phylogeography of *Peromyscus fuvvus* (Rodentia; Muridae) based on cytochrome b sequence data. *Molecular Ecology* 9: 2129-2135.
- HASBÚN CR, A GÓMEZ, G KHÖLER & DH LUNA (2005) Mitochondrial DNA phylogeography of the Mesoamerican spiny-tailed lizards (*Ctenosaura quinquecarinata* complex): historical biogeography, species status and conservation. *Molecular Ecology* 14: 3095-3107.
- HAWKINS JA (2006) Using phylogeny to investigate the origins of the Cape flora: the importance of taxonomic, gene and genome sampling strategies. *Diversity and Distributions* 12: 27-33.
- HEY J & CA MACHADO (2003) The study of structured populations – new hope for a difficult and divided science. *Genetics* 4: 535-543.
- HICKERSON MJ, G DOLMAN & C MORITZ (2006) Comparative phylogeographic summary statistics for testing simultaneous vicariance. *Molecular Ecology* 15: 209-223.
- HOLMES EC, S NEE, A RAMBAUT, GP GARNETT & PH HARVEY (1995) Revealing the history of infectious disease epidemics through phylogenetic trees. *Philosophical Transactions of the Royal Society of London B* 349: 33-40.
- HOLMES EC, OG PYBUS & PH HARVEY (1999) The molecular population genetics of HIV-1. En:

- Crandall KA (ed) The evolution of HIV: 177-207. Johns Hopkins University Press, Baltimore, USA.
- HUDSON RR (1990) Gene genealogies and the coalescent process. En: Futuyama D & J Antonovics (eds) Oxford surveys in evolutionary biology: 1-44. Oxford University Press, Oxford, United Kingdom.
- HUDSON RR (1998) Island models and the coalescent process. *Molecular Ecology* 7: 413-418.
- HUDSON RR, M SLATKIN & WP MADDISON (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* 132: 583-589.
- HUSMEIER D & F WRIGHT (2001) Probabilistic divergence measures for detecting interspecies recombination. *Bioinformatics* 17: S123-S131.
- HUSON DH (1998) SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* 14: 68-73.
- KELCHNER SA & MA THOMAS (2006) Model use in phylogenetics: nine key questions. *Trends in Ecology & Evolution* 22: 87-94.
- KINGMAN JFC (1982) The coalescent. *Stochastic Processes and their Applications* 13: 235-248.
- KNOWLES LL & WP MADDISON (2002) Statistical phylogeography. *Molecular Ecology* 11: 2623-2635.
- KNOWLES LL (2003) The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology* 17: 1-10.
- KOCH MA, C KIEFER, D EHRICH, J VOGEL, C BROCHMANN & K MUMMENHOFF (2006) Three times out of Asia Minor: the phylogeography of *Arabis alpine* L. (Brassicaceae). *Molecular Ecology* 15: 825-839.
- LESSA EP, JA COOK & JL PATTON (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences* 100: 10331-10334.
- LUIKART G & PR ENGLAND (1999) Statistical analysis of microsatellite DNA data. *Trends in Ecology & Evolution* 14: 253-256.
- LUNT DH, LE WHIPPLE & BC HYMAN (1998) Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. *Molecular Ecology* 7: 1441-1455.
- MARTIN AP & SR PALUMBI (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences USA* 90: 4087-4091.
- MEJÍA-MADRID H, E VÁZQUEZ-DOMÍNGUEZ & G PÉREZ-PONCE DE LEÓN (2007) Phylogeography and freshwater basins in Central Mexico: recent history as revealed by the fish parasite *Rhabdochona lichtenfelsi* (Nematoda). *Journal of Biogeography* 34: 787-801.
- MUWANIKA VB, S NYAKAANA, HR SIEGISMUND & P ARCTANDER (2003) Phylogeography and population structure of the common warthog (*Phacochoerus africanus*) inferred from variation in mitochondrial DNA sequences and microsatellite loci. *Heredity* 91: 361-372.
- MUMMENHOFF K, H BRÜGGEMANN & JL BOWMAN (2001) Chloroplast DNA phylogeny and biogeography of *Lepidium* (Brassicaceae). *American Journal of Botany* 88: 2051-2063.
- NORDBORG M (2000) Coalescent theory. En: Balding DJ, MJ Bishop & C Cannings (eds) *Handbook of statistical genetics*: 1-37. John Wiley & Sons, London, United Kingdom.
- PALMER JD (1990) Contrasting modes and tempos of genome evolution in plants and algae. *Trends in Genetics* 6: 115-120.
- PANCHAL M (2007) The automation of nested clade phylogeographic analysis. *Bioinformatics* 23: 509-510.
- PETERS JL, W GRETES & KE OMLAND (2005) Late Pleistocene divergence between eastern and western populations of wood ducks (*Aix sponsa*) inferred by the "isolation with migration" coalescent method. *Molecular Ecology* 14: 3407-3418.
- PETIT RJ (2008) The coup de grâce for the nested clade phylogeographic analysis? *Molecular Ecology* 17: 516-518.
- POSADA D (2006) ModelTestServer: a web-based tool for the statistical selection of models of nucleotide substitution online. *Nucleic Acids Research* 34: W700-W703.
- POSADA D & KA CRANDALL (2001) Intraspecific gene genealogies: tress grafting into networks. *Trends in Ecology & Evolution* 16: 37-45.
- POSADA D, KA CRANDALL & AR TEMPLETON (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* 9: 487-488.
- POSADA D, KA CRANDALL & AR TEMPLETON (2006) Nested clade analysis statistics. *Molecular Ecology Notes* 6: 590-593.
- REEB V, P HAUGEN, D BHATTACHARYA & F LUTZONI (2007) Evolution of *Pleopsidium* (Lichenized Ascomycota) S943 group I introns and the phylogeography of an intron-encoded putative homing endonuclease. *Journal of Molecular Evolution* 64: 1432-1432.
- RICHARDSON BA, J BRUSFELD & B KLOPFENSTEIN (2002) DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology* 11: 215-227.
- RIDDLE BR, DJ HAFNER, LF ALEXANDER & JR JAEGER (2000) Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. *Proceedings of the National Academy of Sciences USA* 97: 14438-14443.
- ROGERS AR & H HARPENDING (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9: 552-569.
- SMITH TB, K HOLDER, D GIRMAN, K O'KEEFE, B LARISON & T CHAN (2000) Comparative avian phylogeography of Cameroon and Equatorial Guinea mountains: implications for conservation. *Molecular Ecology* 9: 1505-1516.
- SMIT HA, TJ ROBINSON & BJ VAN VUUREN (2007) Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular Ecology* 16: 2680-2692.
- SOLTIS DE, PS SOLTIS & BJ MILLIGAN (1992) Intraspecific chloroplast DNA variation: systematic and phylogenetic implications. En: Soltis PS & JJ Doyle (eds) *Molecular systematics of plants*: 117-150. Chapman & Hall, New York, USA.
- SPELLMAN G & J KLICKA (2007) Phylogeography of the White-breasted Nuthatch (*Sitta carolinensis*): diversification in North American pine and oak woodlands. *Molecular Ecology* 16: 1729-1740.
- STRIMMER K & OG PYBUS (2001) Exploring the demographic history of DNA sequences using the generalized skyline plot. *Molecular Biology and Evolution* 18: 2298-2305.
- TAJIMA F (1983) Evolutionary relationships of DNA

- sequences in finite populations. *Genetics* 105: 437-460.
- TEMPLETON AR (1983) Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of man and the apes. *Evolution* 37: 221-244.
- TEMPLETON AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* 7: 381-397.
- TEMPLETON AR (2002) Out of Africa again and again. *Nature* 416: 45-51.
- TEMPLETON AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* 13: 789-809.
- TEMPLETON AR (2008) Nested clade analysis: an extensively validated method for strong phylogeographic inference. *Molecular Ecology* 17: 1877-1880.
- TEMPLETON AR & CF SING (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134: 659-669.
- TEMPLETON AR, KA CRANDALL & CF SING (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619-633.
- TEMPLETON AR, E ROUTMAN & C PHILLIPS (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* 140: 767-782.
- VÁZQUEZ-DOMÍNGUEZ E (2002) Phylogeography, historical patterns and conservation of natural areas. En: Nelson G, JC Day, LM Sportza, J Loucky & C Vásquez (eds) *Protected areas and the regional planning imperative in North America*: 369-378. University of Calgary Press, Calgary, Canada.
- VÁZQUEZ-DOMÍNGUEZ E (2007a) ¿Cuál es el enfoque de la ecología molecular? Ejemplos con un ratón espinoso. En: Sánchez-Rojas G & AE Rojas-Martínez (eds) *Tópicos en sistemática, biogeografía, ecología y conservación de mamíferos*: 59-70. Universidad Autónoma del Estado de Hidalgo, Serie Ciencia al día 1, México.
- VÁZQUEZ-DOMÍNGUEZ E (2007b) Filogeografía y vertebrados. En: Eguiarte L, V Souza & X Aguirre (eds) *La ecología molecular de plantas y animales*: 441-466. INE/Conabio, México.
- VÁZQUEZ-DOMÍNGUEZ E, A MEJÍA-MADRID & R VEGA (2009) *Oryzomys couesi* en el sureste de México: estimaciones genéticas y filogeográficas. En: Cervantes FA, J Vargas-Cuenca & Y Hortelano-Moncada (eds) *60 años de la Colección Nacional de Mamíferos del Instituto de Biología, UNAM. Aportaciones al conocimiento y conservación de los mamíferos mexicanos*. Instituto de Biología, UNAM, México.
- WAINSCOAT J (1987) Out of the garden of Eden. *Science* 234: 194-196.
- WRIGHT S (1931) Evolution in Mendelian populations. *Genetics* 16: 97-159.
- WRIGHT S (1943) Isolation by distance. *Genetics* 28: 114-138.
- ZUCKERKANDL E & L PAULING (1965) Evolutionary divergence and convergence in proteins. En: Bryson V & HJ Voegel (eds) *Evolving genes and proteins*: 97-166. Academic Press, New York, USA.

Editor Asociado: Elie Poulin

Recibido el 29 de febrero de 2008; aceptado el 25 de marzo de 2009

