

**Aproximació molecular a l'estudi dels primats:
Evolució dels gens *RPS4Y* i
aplicació d'SNPs en conservació**

Memòria presentada per
Olga Andrés Viñas

Per optar al grau de
Doctora

Barcelona, 4 d'abril del 2007

INTRODUCCIÓ

1. Primates

1.1. Descripció general

Primates és l'ordre de mamífers dins del qual es troba situat l'home i els seus parents més propers. Pertany al regne Animalia, fílum Chordata i classe Mammalia. Els ancestres dels primats van aparèixer fa uns 70 milions d'anys i se'ls coneix amb el nom de prosimis primitius. Aquests primers primats eren nocturns i arboris, i moltes de les característiques que els definien estan relacionades amb aquesta primitiva vida nocturna als arbres. Molts d'aquests trets primitius van ser retinguts i van contribuir a l'èxit de la línia dels primats quan van adaptar-se a una vida menys arborícola i més terrestre.



Figura 1. Imatge de *Lagothrix cana*. Moltes espècies de primat són principalment arbòries. Les mans i els peus dels primats tenen el dit gros oponible, que els permet tenir una agafada forta. Les espècies amb cua la utilitzen com una cinquena extremitat. (L.C. Marigo/WPRC AV Archives)

Els primats es diferencien de la resta de mamífers perquè presenten una morfologia poc especialitzada i una gran plasticitat del comportament. L'increment del refinament de mans i peus per agafar objectes ha estat un tret definitori en l'evolució dels primats, que han mantingut el pentadactilisme primitiu dels mamífers (cinc dits a cada mà i a cada peu), tenen el dit polze oponible (que permet agafar de manera forta i precisa) i les gemmes dels dits molt sensibles al tacte. Els primats tenen una visió en color i binocular, de percepció profunda (percepció tridimensional real), però la visió perifèrica s'ha vist reduïda.

L'increment de les capacitats visuals ha comportat una disminució dels altres sentits, especialment de l'olfacte, de manera que s'ha produït una reducció progressiva de la mida del nas i de les

àrees olfactivas del cervell. Si els comparem amb la majoria d'animals, els primats tenen un cervell gran en relació a la resta del cos, amb una expansió especialment important de

les àrees implicades en el control de l'habilitat manual, de la coordinació ull-mà i de la visió estereoscòpica. Pel que fa a la reproducció, els primats han desenvolupat processos gestacionals cada vegada més eficients i llargs, amb una reducció de la mida de la ventrada, generalment amb un sol fill, i amb més atenció cap als infants.



Figura 2. Els primats són animals molt sociables que es relacionen amb el seu grup, per exemple: A) desparasitant-se (*Macaca fascicularis*), o B) amb exhibicions vocals (*Alouatta seniculus*). (R. Fontaine/WPRC AV Archives)

1.2. Hàbitat i distribució al planeta

Els primats no humans es distribueixen entre Amèrica del Sud i Central, Àfrica i Àsia (figura 3). La majoria de les espècies viuen als tròpics o subtòpics i depenen dels boscos per sobreviure. Hi ha molts tipus de bosc diferents segons el clima, l'altitud, el sòl, etc. A més, podem diferenciar entre boscos primaris, els que no han estat pertorbats durant centenars d'anys, i boscos secundaris, els que s'estan regenerant després d'haver patit una alteració, natural o provocada per l'home. Alguns primats poden viure únicament en boscos verges, mentre que d'altres prefereixen els boscos en regeneració. A més, dins el bosc les diferents espècies de primat tenen diferents preferències de nivell: algunes espècies es troben en el sòl, altres a mitja alçada i les més arbòries viuen a les capes més altes del bosc. Entre les espècies de primats n'hi ha que s'adapten bé a diferents hàbitats dins del seu rang mentre que altres són extremadament restrictives.

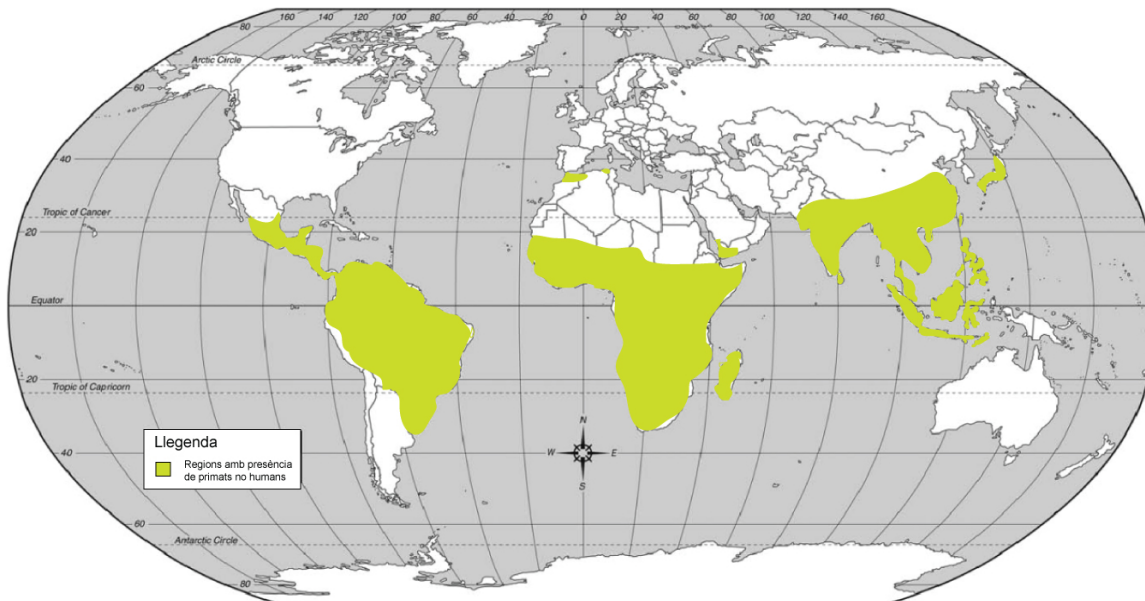


Figura 3. Distribució mundial actual dels primats no humans.

1.3. Taxonomia i filogènia

En general, la taxonomia sempre ha estat una qüestió difícil per als científics. Les categories jeràrquiques establertes per Linnée no són suficients per classificar els milions d'espècies vives conegudes actualment i s'han anat afegint noves categories intermèdies. A més, totes les classificacions pateixen modificacions a mida que la ciència avança, de manera que primer es tenien en compte la morfologia i distribució geogràfica de les espècies, després es va introduir el coneixement aportat pels fòssils, més endavant el derivat de l'estudi dels cromosomes i finalment s'ha produït la revolució dels estudis a nivell de DNA. D'altra banda, s'han donat etapes reduccionistes, fins a l'extrem de voler eliminar tota classificació, i etapes divisionistes, en què de qualsevol diferència en sorgia una espècie nova.

Dins d'aquest context, la taxonomia dels primats no ha estat una excepció i ha passat per totes les etapes. Avui dia encara hi ha diferències de criteri entre els primatòlegs, fins i tot a l'hora d'establir les categories més generals. Durant molt de temps s'havia diferenciat entre Prosimii i Anthropeidea, segons si els animals presentaven caràcters primitius o més evolucionats, respectivament, en una concepció antropocèntrica. Dins dels Prosimis s'inclouïa la família Tarsiidae, que és la de més difícil classificació, però s'obtenia una organització basada en graus i no en clades, ja que els tarsers s'agrupaven filogenèticament amb els Anthropeidea (veure figura 4). Una altra classificació que s'ha proposat és dividir l'ordre en 3 subordres: Prosimii, Tarsiiformes i Simiiformes (o Anthropeidea). La dificultat de la classificació dels tarsers prové del fet que tenen

característiques morfològiques i de comportament tant dels prosimis com dels simiiformes. Actualment, però, la taxonomia més àmpliament acceptada és la de Groves (2001), que divideix els primats en dos subordres: Strepsirrhini i Haplorrhini. La majoria de taxònoms han adoptat aquesta nomenclatura, basada en les característiques del nas, i han reclassificat els primats en primats “de nas humit” (Strepsirrhini), que inclou els antics Prosimii (excepte els Tarsiidae), i primats “de nas eixut” (Haplorrhini), que inclou els abans anomenats Anthropeoidea i la família Tarsiidae. La taula 1 mostra la classificació més acceptada actualment fins al nivell de família.

Taula 1. Classificació fins al nivell de família dels primats existents actualment. (Adaptat de Groves, 2001)

ORDRE PRIMATES

Subordre Strepsirrhini

Infraordre Lemuriformes

Superfamília Cheirogaleoidea

Família Cheirogaleidae

Superfamília Lemuroidea

Família Lemuridae

Família Megaladapidae

Família Indridae

Infraordre Chiromyiformes

Família Daubentoniidae

Infraordre Loriformes

Família Loridae

Família Galagonidae

Subordre Haplorrhini

Infraordre Tarsiiformes

Família Tarsiidae

Infraordre Simiiformes

Platyrrhini

Família Cebidae

Família Aotidae

Família Pitheciidae

Família Atelidae

Catarrhini

Superfamília Cercopithecoidea

Família Cercopithecidae

Superfamília Hominoidea

Família Hylobatidae

Família Hominidae

En primats, cada vegada és més evident que les dades morfològiques i moleculars convergeixen cap a un únic senyal filogenètic, de manera que és possible unir la informació obtinguda dels dos tipus de dades per donar més força a les filogènies (Yoder,

1994) i poder arribar així a una filogènia consens i fiable, a més de permetre datar les edats de divergència dels grans grups (veure figura 4). Les incongruències apareixen quan es vol tenir en compte la informació del registre fòssil, ja que aleshores sorgeixen tres nous arbres filogenètics possibles, que difereixen principalment en la posició dels tarsers (Shoshani *et al.*, 1996). La filogènia més acceptada actualment, però, és la basada en la combinació de dades morfològiques i genètiques obtingudes a partir de les espècies actuals (Shoshani *et al.*, 1996) (figura 4).

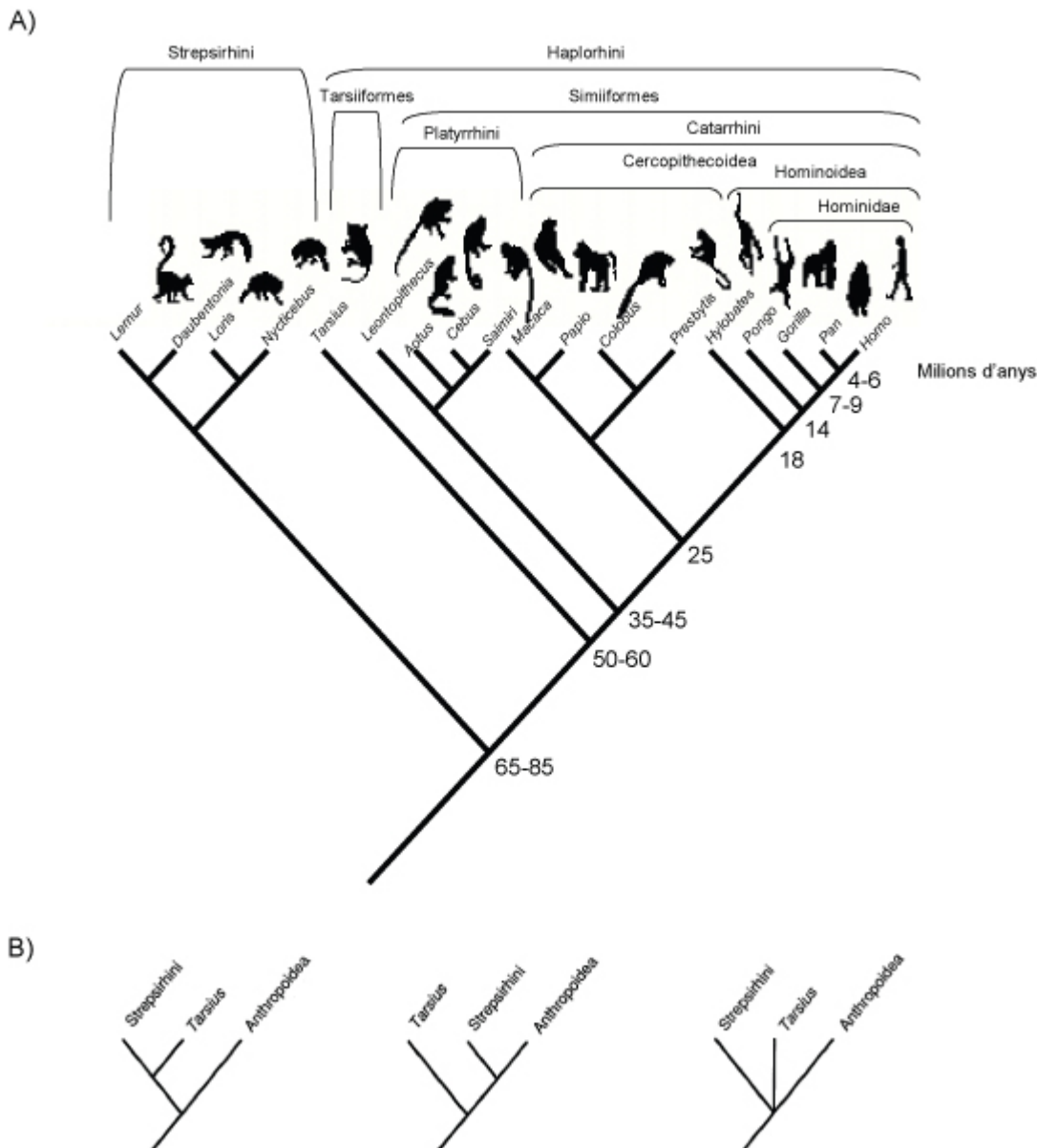


Figura 4. A) Filogènia dels primats actualment acceptada (Groves, 2001). Als nodes de les branques apareix el temps de divergència entre grups, en milions d'anys (estimacions extrems de Nahon, 2003). B) Controvèrsia en la posició dels tarsers quan es consideren dades de fòssils. (Figura adaptada de Shoshani *et al.*, 1996)

D'altra banda, el número d'espècies de primats vives actualment no està clar. El número varia segons si els grups més relacionats es consideren varietats de la mateixa espècie o espècies diferents. Segons els taxònoms més divisionistes els primats es classificarien en més de 350 espècies, mentre que pels taxònoms més reduccionistes hi hauria unes 190 espècies. A més, la qüestió es complica perquè cada any es descobreixen nous tàxons i els boscos tropicals potser encara amaguen espècies desconegudes. Pel que fa a la desaparició de tàxons, tot i que moltes espècies de primat es troben amenaçades, no s'havia registrat cap extinció durant el segle XX, fins que a finals de segle es va constatar la primera desaparició d'un tàxon, *Procolobus badius waldroni*, una subespècie de còlob vermell endèmica d'Àfrica de l'Oest (Oates *et al.*, 2000).

1.4. Estat de conservació

D'un terç a la meitat de les espècies de primat es troben amenaçades com a conseqüència de la sobreexplotació i la destrucció de l'hàbitat i aquesta situació pot ser encara molt pitjor en un futur no gaire llunyà si no es prenen mesures. Els primats són una peça essencial dins de l'ecosistema en el que viuen, de manera que la seva desaparició no només comporta la pèrdua d'un tàxon en concret sinó que pot tenir conseqüències greus i irreversibles per a tot l'ecosistema. La conservació dels primats, però, es fa difícil per una sèrie de factors, alguns inherents a les seves pròpies característiques. Els primats són animals tropicals, els seus hàbitats són els boscos i sabanes tropicals, de manera que viuen principalment en països en vies de desenvolupament amb una gran demanda de recursos (veure apartat 7 del Resum global). D'altra banda, els primats són animals amb un potencial reproductiu baix, les poblacions no poden créixer de pressa i passa molt de temps fins que es recuperen. A més, com que viuen en densitats poblacionals baixes, un increment de la mida de la població implica disposar d'un territori molt ampli.

La principal amenaça dels primats és la desforestació, causada principalment pel creixement de la població humana. La fragmentació de la selva a conseqüència de la tala massiva d'arbres ha provocat una reducció de la grandària de les poblacions de primats i l'aïllament físic, cosa que impedeix el flux gènic entre individus de diferents comunitats i augmenta la taxa de consanguinitat perquè no es poden produir els processos naturals d'emigració i immigració. Aquesta pèrdua d'hàbitat és causada per diferents factors, principalment pel desenvolupament de l'agricultura i la ramaderia, que ha comportat que actualment un 26% de la regió de terra del planeta s'utilitzi amb aquestes finalitats. A més, els productes de fusta, especialment els de fusta dels tròpics, tenen una gran demanda en

el mercat internacional i, tot i que existeix un comerç d'aquests productes controlat governamentalment, també hi ha una explotació il·legal en zones protegides, que a vegades pot arribar a ser més important que l'extracció legal. S'estima que es talen un total de 5 o 6 milions d'hectàrees cada any (Chapman, Peres, 2001). D'altra banda, la creació d'infraestructures, necessàries per dur a terme projectes de desenvolupament i d'explotació dels recursos naturals, també suposa una tala dels arbres i transformació dels hàbitats, ja que és necessari, per exemple, construir carreteres en llocs prèviament verges. Això comporta pèrdua d'hàbitat, erosió, pol·lució i facilita l'accés als caçadors furtius, a més d'augmentar la possibilitat de transmissió de malalties, d'introducció de depredadors, d'incendis i d'atropellaments. Els incendis, tant si comencen de manera natural o com a conseqüència de l'activitat humana, també són una amenaça per als primats perquè posen en perill els individus, destrueixen grans regions d'hàbitat i poden deixar infants i adolescents sense la protecció dels adults que encara necessiten (Chapman, Peres, 2001).

A més de la desforestació, hi ha altres pressions que s'afegeixen a aquesta, com la caça d'animals per vendre com a aliment o com a mascotes (veure Chapman, Peres, 2001). La caça de primats per aconseguir carn, tot i que està prohibida a la majoria de països, se segueix produint en tots els continents on habiten primats no humans i se sap que actualment es ven carn de primat també a Europa i Nord Amèrica. Durant segles, aquest tipus de caça ha format part de moltes cultures natives, però el creixement exponencial de la població humana ha fet que aquesta pràctica posi en perill l'existència de tots els primats. A més, l'ús recent d'armes de foc per caçar fa molt més efectiva aquesta pràctica i incrementa la pressió sobre les poblacions de primats. L'elevada demanda de carn al mercat fa que la caça furtiva sigui comú i les espècies de mida més gran són les que pateixen una pressió més forta (Refisch, Kone, 2005). Aquesta pràctica comporta també el problema afegit dels infants orfes, ja que els caçadors estan interessats en els adults. En alguns casos els infants orfes seran comercialitzats com a mascotes o bé moriran per falta d'atenció maternal. La caça il·legal de primats, però, no només té com objectiu l'obtenció d'aliment. Moltes vegades es cacen els animals per vendre'n la pell i parts del cos en mercats nacionals i internacionals com a trofeus o com a ingredients de medicines tradicionals. Gràcies a les lleis internacionals que prohibeixen el comerç d'espècies amenaçades, les exportacions de productes derivats d'espècies de primat han disminuït molt, però no han desaparegut.

Pel que fa al comerç d'animals vius, molts primats han estat, i encara són avui dia, capturats per comercialitzar-los il·legalment com a mascotes. Els bebès són les peces més preuades, i moltes vegades per aconseguir-los els caçadors furtius han de matar la mare i altres adults (Peres, 1991). D'altra banda, els primats no humans són els animals més propers a l'home i per això són molt valorats com a model per fer recerca biomèdica. Al 1989, el 80-90% dels primats utilitzats en biomedicina provenien principalment de la natura, eren caçats amb aquest objectiu i comercialitzats (Fitzgerald, 1989). Actualment, però, s'ha determinat que gairebé tots els primats que s'utilitzen en investigació provenen de granges de reproducció, tot i que encara n'hi ha una part que es comercialitza, especialment des d'Indonèsia i Filipines (Malik, 2000). A més, actualment tampoc és comú que es dugui a terme la captura d'animals per a zoològics i circs o per a la indústria cinematogràfica. Tanmateix, en tots aquests contextos, tant biomèdics com de la indústria de l'entreteniment, cal considerar qüestions ètiques i del benestar dels individus.

Finalment, el fet que els primats no humans siguin evolutivament molt propers als homes implica que les fisiologies siguin molt semblants, de manera que les malalties que afecten els uns també poden ser transmeses als altres. L'ecoturisme i l'expansió humana provoquen que els primats estiguin exposats a la presència de l'home i, doncs, a les malalties que pugui portar. D'aquesta manera, els primats no humans poden contreure malalties típicament humanes per a les que el seu sistema immunitari no està preparat i que poden resultar mortals. Hi ha casos recents que demostren la realitat d'aquesta amenaça, com per exemple la mort de 5.000 goril·les a la República del Congo infectats per una soca del virus d'Ebola provinent de l'home (Bermejo, 2005).

1.5. Protecció de les espècies

La *World Conservation Union* (IUCN), abans anomenada *International Union for the Protection of Nature* (IUPN), és la xarxa de conservació més extensa i important del món. Es va crear el 1948 amb l'objectiu d'influir, encoratjar i ajudar les societats arreu del planeta a conservar la integritat i diversitat de la natura i a assegurar que tots els usos dels recursos naturals es facin de manera equilibrada i ecològicament sostenible. A través de la *Species Survival Commission* (SSC), la IUCN valora a una escala global l'estatus de conservació de les espècies, subespècies, varietats i, fins i tot, poblacions especials, per tal de determinar quins tàxons estan amenaçats d'extinció i així poder promocionar la seva conservació. La informació de la situació dels diferents tàxons es recull a la Llista Vermella d'Espècies Amenaçades, on les espècies apareixen classificades segons el grau

de perill d'extinció i es fa especial èmfasi sobre les espècies que es troben en més perill, classificades com: “en perill crític” (CR), les que estan davant d'un risc d'extinció imminent a la natura; “en perill” (EN), les que es troben en un risc molt alt d'extinció a la natura; i “vulnerables” (VU), les que es troben en alt risc d'extinció a la natura.

El grup especialitzat en primats (*Primate Specialist Group*) de la SSC considera que gairebé la meitat de les espècies de primats han de ser tractades a nivell de conservació i que una de cada cinc espècies s'ha d'incloure a les categories d'espècies “en perill” o “en perill crític” d'extinció, ja que si no reben la protecció adequada desapareixeran en poques dècades. Al 2005, la IUCN reconeixia 230 primats amenaçats, un 26% del total de tàxons, dels quals 160 es trobaven a les categories “en perill crític” o “en perill”. A més, cal tenir en compte que molts dels tàxons es troben exclusivament en els anomenats punts calents (*hotspots*), en regions amenaçades que contenen una elevada biodiversitat, de manera que el 30% de la diversitat global dels primats es troba repartida en 14 punts calents (que conformen un 1% del total de l'àrea terrestre del planeta). D'aquests 14 punts calents, set són considerats d'elevada prioritat per als primats, ja que inclouen 106 tàxons en perill, un 46% de tots els primats amenaçats, de manera que el futur de gairebé la meitat del primats en perill està lligat al futur d'aquestes regions (Mittermeier, 2005).

La CITES (Convenció sobre el Comerç Internacional d'Espècies Amenaçades de Fauna i Flora Salvatges) és un acord internacional entre governs per assegurar que el comerç internacional d'espècimens d'animals i plantes silvestres no comporti una amenaça per a la seva supervivència. Les espècies emparades per la CITES es distribueixen en tres Apèndixs, segons el grau de protecció que necessiten. L'Apèndix I inclou les espècies que es troben en perill d'extinció i el seu comerç (inclosos els productes derivats) només pot ser autoritzat en circumstàncies excepcionals. L'Apèndix II inclou les espècies que no es troben en perill d'extinció però que necessiten que el seu comerç sigui controlat per evitar un ús incompatible amb la seva supervivència. L'Apèndix III inclou espècies que estan protegides en almenys un país que ha demanat ajuda a la resta per controlar-ne el comerç. Només es poden importar o exportar productes d'una espècie inclosa en els Apèndixs de la CITES si s'obtenen les llicències necessàries. Totes les espècies de primat es troben als Apèndix I o II de CITES. Gràcies a acords com la CITES s'ha limitat el comerç de primats, tant el comerç d'animals per ser utilitzats com a mascotes, que no és permès en cap cas, com pel seu ús en biomedicina, que s'ha d'intentar minimitzar i que, de totes maneres, es pot fer a partir d'animals reproduïts en captivitat per evitar noves captures. Els problemes més greus amb què s'enfronten els primats, però, són, d'una

banda, el comerç il·legal que, tot i la normativa, segueix existint, i, de l'altra, la destrucció de l'hàbitat, que és molt difícil de solucionar (discutit a l'apartat 7 del Resum global).

Per prevenir l'extinció de les espècies és necessari conèixer-les bé a tots els nivells, tant ecològic com morfològic com genètic, i així poder establir plans de conservació adequats. Des dels països desenvolupats es pot contribuir a fer avançar l'estudi dels primats per ajudar a la seva conservació, especialment des de la genètica, que no requereix que els treballs siguin realitzats *in situ* (veure apartat 2 d'aquesta Introducció). Així, cal invertir en recerca bàsica i aplicada dins del camp de la genètica de la conservació per tal d'aportar informació essencial que contribueixi a entendre les espècies i per dissenyar eines que permetin estudiar les poblacions amenaçades i controlar la seva situació.

1.6. Ximpanzés

1.6.1. Taxonomia i distribució geogràfica

El ximpanzé (*Pan troglodytes*) és una espècie de primat de la família Hominidae, a la qual també pertany l'home. Dins el gènere *Pan* s'inclou una altra espècie molt semblant al ximpanzé, el bonobo (*Pan paniscus*). Els llinatges de ximpanzés i bonobos van divergir fa només uns 2 milions d'anys (Stone *et al.*, 2002) i la branca evolutiva del gènere *Pan* es va separar de la de l'home entre 4 i 6 milions d'anys enrere (Groves, 2001), de manera que ximpanzés i bonobos són les espècies més properes a l'home.

Existeix una gran varietat morfològica dins les poblacions de ximpanzé, cosa que ha fet difícil la seva classificació. En alguns moments s'havien arribat a definir fins a 12 gèneres i encara més espècies i subespècies. Tradicionalment s'han distingit tres subespècies de ximpanzé: *Pan troglodytes troglodytes*, *Pan troglodytes verus* i *Pan troglodytes schweinfurthii*. Tot i que aquestes subespècies no es poden diferenciar morfològicament entre elles, estan ben caracteritzades a nivell molecular (Morin *et al.*, 1994). Fins i tot s'ha parlat d'ascendir la subespècie *P. t. verus* a la categoria d'espècie, ja que certs marcadors indiquen que està molt diferenciada genèticament de les altres dues subespècies, tant com ho estan altres mamífers considerats espècies diferents (Morin *et al.*, 1994). D'altra banda, les diferències genètiques que presenta una població de Nigèria i les barreres geogràfiques de la zona han suggerit l'existència d'una quarta subespècie, *Pan troglodytes vellerosus* (Gonder *et al.*, 1997), que sembla força acceptada actualment (Grubb *et al.*, 2003). Tanmateix, recentment s'ha proposat dividir l'espècie *P. troglodytes* en només dues subespècies: *P.t. vellerosus* i *P.t. troglodytes* (Gonder *et al.*, 2006).

L'hàbitat dels ximpanzés són les selves tropicals i les sabanes humides d'Àfrica Central i Occidental i s'estén des de Guinea fins a Tanzània (a Burundi, Camerun, República Democràtica del Congo, Congo, Gabon, Gàmbia, Ghana, Guinea Equatorial, Guinea Bissau, Costa d'Ivori, Libèria, Nigèria, Rwanda, Sierra Leone, Sudan, Tanzània, Uganda i Zaire). Fa temps les poblacions ocupaven la major part d'aquest territori, però en els últims anys el seu hàbitat ha estat reduït dràsticament com a conseqüència de la tala d'arbres i encara es troba en recessió (figura 5).

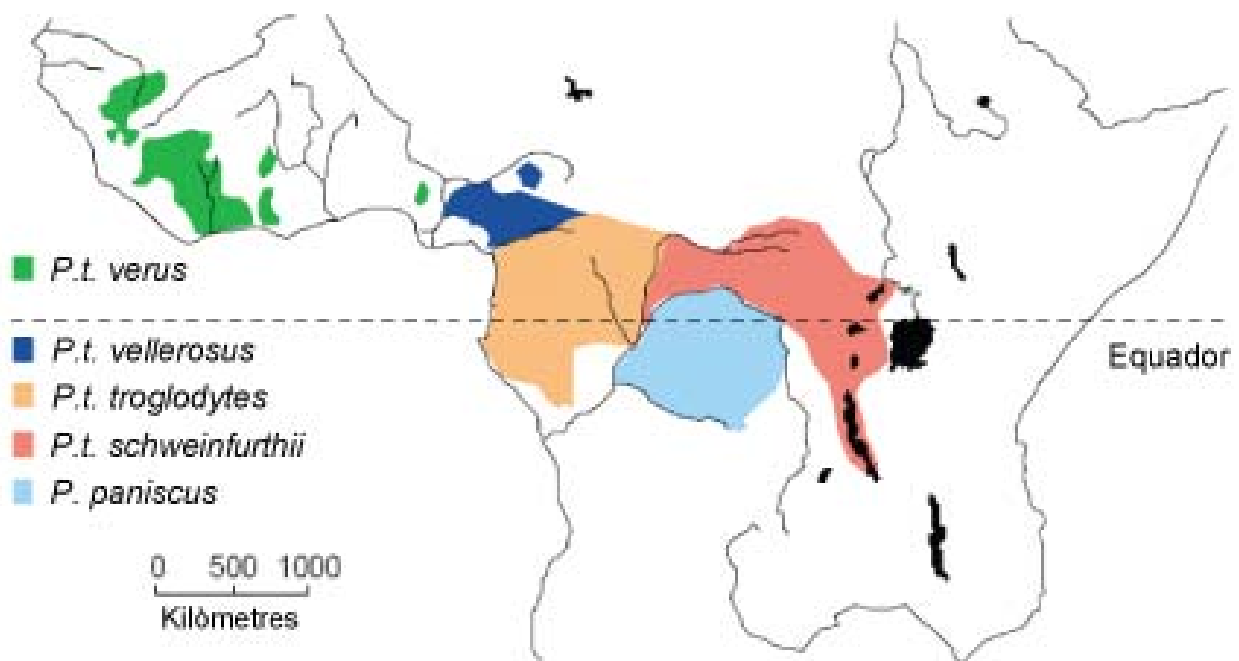


Figura 5. Àrea de distribució actual de *P. troglodytes* (amb totes les subespècies) i de *P. paniscus*. (Adaptada de Gagneux, 2002)

1.6.2. Morfologia, reproducció i alimentació

Els ximpanzés adults poden arribar a mesurar 130 cm, les femelles, i 160 cm, els mascles, i pesen entre 40 i 70 kg. Tenen el cos cobert de pelatge gruixut de color marró fosc, excepte la cara, els dits, els palmells de les mans i les plantes dels peus, que no tenen pèl. Quan neixen tenen la pell de color clar, que es va enfosquint a mida que creixen, i quan es fan grans el pèl de la barba se'ls pot tornar blanc. Els polzes i els dits grossos dels peus són oponibles, cosa que els permet tenir una agafada precisa. Les femelles solen tenir un fill cada cinc o sis anys i la gestació és de 8 mesos. Els infants deixen de mamar als 3 anys, però en general mantenen una relació propera amb la mare durant diversos anys més. Arriben a la pubertat als 8 (femelles) o 10 (mascles) anys. Els

ximpanzés poden viure fins als 40-45 anys a la natura, però en captivitat solen viure fins als 50 i alguns han arribat gairebé als 70 anys (MacDonald, 2001).

Els ximpanzés són omnívors i canvien la dieta segons l'estació de l'any (figura 6). Es tracta d'animals principalment frugívors, s'alimenten sobretot de fruita, fins i tot quan n'hi ha poca, però també de fulles, arrels, tubercles, flors, etc. Completen la dieta gairebé vegetariana amb insectes, mel, ous, ocells i mamífers petits o mitjans (rosegadors, antílops, porcs salvatges i, sobretot, altres primats més petits). Poden caçar per defensar-se o bé per obtenir aliment, i sovint ho fan en grup.

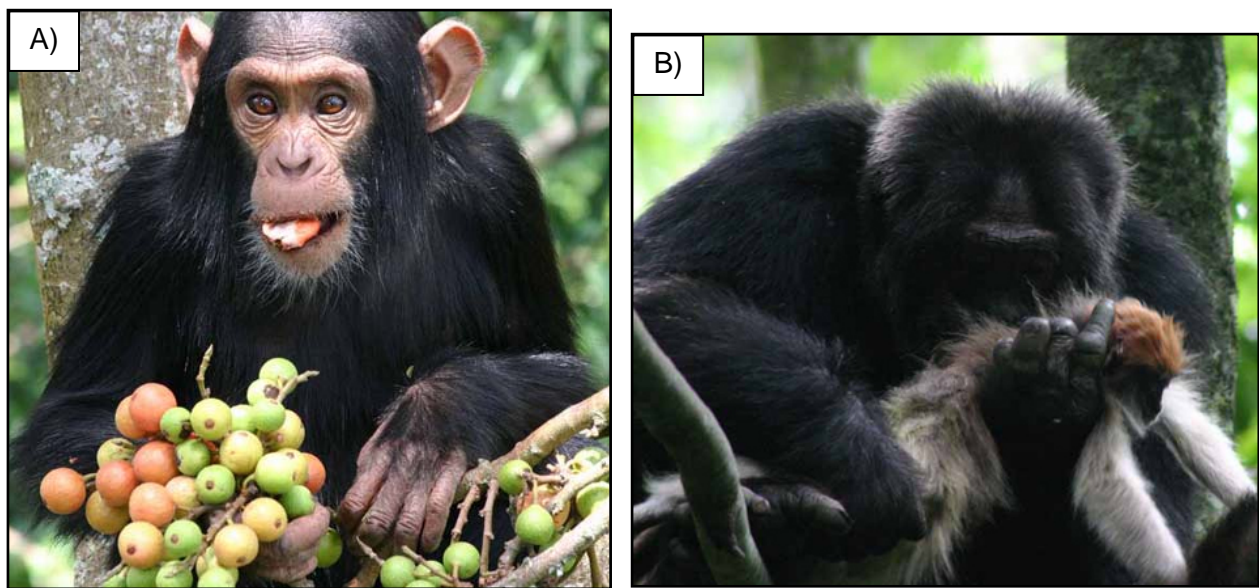


Figura 6. La dieta bàsica dels ximpanzés és frugívora (A), però la completen amb proteïnes extretes, per exemple, de la carn d'altres primats (B). (A. Houle/WPRC AV Archives)

1.6.3. Estructura social i comportament

Els ximpanzés són de vida gregària, viuen en comunitats multimasclle-multifemella de fusió-fissió. Els beneficis que els aporta viure en grup són la disminució de la probabilitat de ser depredats, la possibilitat de defensar els recursos i augmentar l'eficiència d'alimentar-se i l'increment en l'èxit d'aparellament perquè és més fàcil accedir a la parella (Boesch, 1996; Sakura, 1994). Hi ha una jerarquia de dominància ben definida entre els mascles, que a més són dominants sobre les femelles (Goldberg, Wrangham, 1997). Les femelles solen emigrar d'una comunitat a una altra a l'adolescència (Nishida *et al.*, 2003), mentre que els mascles romanen a la seva comunitat i patrulles d'individus emparentats vigilen les fronteres del seu territori, de manera que el grau de parentiu entre els mascles d'una comunitat és molt més elevat que el que hi ha entre les femelles (Morin *et al.*, 1994).

Les comunitats oscil·len entre els 20 i els 150 individus o més, amb diversos mascles, femelles i joves, tot i que es divideixen en famílies, amb un mascle dominant i vàries femelles amb les seves cries. Les femelles, però, tenen una vida més solitària, i la major part del temps estan aïllades amb les seves cries.

Els ximpanzés són tant arboris com terrestres, passen gran part del dia a terra, però cada nit construeixen nius amb fullaraca i branques per dormir als arbres. A més, durant el dia també poden construir nius per descansar i acostumen a menjar als arbres. Els nius poden fer-se de manera individual o pot ser que més d'un individu ajudi a preparar-lo i sigui compartit. La seva locomoció és quadrúpede, solen caminar a quatre potes utilitzant les plantes dels peus i descansant sobre els punys, però poden desplaçar-se de manera bípeda quan tenen les mans ocupades o en distàncies curtes.

Aquests animals són capaços de modificar i utilitzar objectes del seu voltant com a eines per aconseguir un objectiu concret. S'han descrit diverses activitats d'aquest tipus. Per exemple, els ximpanzés poden collir i adaptar bastons per introduir-los en els termiters i capturar els tèrmits i també fabriquen esponges amb fulles o molses per aconseguir aigua de les cavitats. Fins i tot se'ls ha observat utilitzant eines per caçar. D'altra banda, s'ha constatat que els ximpanzés tenen certes tradicions culturals que són pròpies de cada comunitat, de manera que les diferents comunitats tenen diferents costums, que es passen de generació en generació. S'han pogut descriure 39 tradicions diferents (Whiten *et al.*, 1999; Whiten *et al.*, 2001).

Pel que fa a la còpula, les relacions promíscues són molt comuns i les femelles poden copular amb mascles diferents oportunísticament durant el temps que dura l'estre. Però hi ha molts tipus d'estratègies diferents que comporten que les femelles puguin expandir el conjunt de mascles per triar sense perdre el benefici de la comunitat, mentre que als mascles els permet canviar d'estratègia segons la jerarquia que tinguin en cada etapa de la vida. Per exemple, el mascle dominant pot apartar-se de la resta de ximpanzés amb la femella en estre per evitar que altres mascles hi copulin; els mascles de rang inferior poden oferir carn a la femella en estre per aconseguir aparellar-se amb ella; i les femelles poden marxar de la seva comunitat i copular furtivament amb mascles d'una altra comunitat.

1.6.4. Estat de conservació

Els ximpanzés tenen una àmplia distribució a l'Àfrica tropical, però les seves poblacions estan patint una ràpida recessió principalment per l'alteració de l'hàbitat, el comerç de carn per menjar i el comerç d'infants vius per ser usats com a mascotes (Rowe, 1992). S'estima que la població total a la natura actualment està entre 100.000 i 150.000 segons la *World Association of Zoos and Aquariums* (WAZA), tot i que estimacions més optimistes diuen que les poblacions han augmentat i que l'any 2003 gairebé hi havia 300.000 individus (Oates, 2006). Actualment, més del 70% de l'hàbitat dels ximpanzés està alterat pel desenvolupament d'infraestructura i si la població humana segueix creixent al mateix ritme es calcula que cap al 2030 quedarà menys del 10% de l'hàbitat no modificat per l'home (Nellemén, 2002).

El ximpanzé es troba dins l'Apèndix I de CITES, de manera que el seu comerç està controlat i calen permisos especials per poder importar i exportar productes. A més, fins al 1995 el ximpanzé havia estat catalogat com a espècie "vulnerable" per la IUCN, però des del 1996 totes les subespècies estan classificades com "en perill d'extinció". Dins d'aquesta situació general, però, hi ha poblacions que estan en una situació molt pitjor, a punt de desaparèixer. Gràcies a les lleis nacionals i internacionals, les autoritats dels països on habita aquesta espècie confisquen regularment animals caçats il·legalment, però aleshores s'han d'enfrontar amb el problema del futur dels individus. Actualment existeixen diversos santuaris per a ximpanzés decomissats i orfes i s'estableixen plans d'alliberament, però és necessari estudiar bé els individus i les poblacions naturals perquè els programes de translocació siguin adequats i retornar els animals a la natura és una tasca difícil i no sempre és possible.

2. Genètica de la conservació

2.1. La conservació de les espècies

L'home és l'espècie animal que té un rang de distribució més ampli a la Terra i el seu èxit es deu, en part, al desenvolupament de l'agricultura i la ramaderia. Tanmateix, la substitució massiva de la natura per ambients humanitzats ha provocat la pèrdua constant d'espècies, tant animals com vegetals, i la desaparició d'ecosistemes sencers, de manera que el creixement de la població humana comporta una disminució accelerada de la diversitat de la Terra. S'ha estimat que de 3.000 a 30.000 espècies s'extingeixen cada any (Woodruff, 2001) i es considera que al voltant del 50% de les espècies animals estan "en perill crític", "en perill" o "vulnerables" a l'extinció (IUCN 2001). La desaparició de les espècies durant els últims segles s'està produint a una velocitat tan elevada que es parla de la sisena gran extinció (Delibes, 2004) i, si no es fa res per aturar aquesta tendència, el món en el seu estat actual podria desaparèixer. La societat és cada cop més conscient d'aquesta realitat i, doncs, comença a comprendre la importància de la conservació de les espècies.

Tanmateix, l'elecció de les espècies que cal protegir no és gens fàcil. Tradicionalment, aquesta selecció es basava en els coneixements ecològics i morfològics de què es disposava i moltes vegades era més una qüestió "folklòrica" que científica, ja que es protegien espècies emblemàtiques, prou carismàtiques com per commoure la societat. Avui en dia el coneixement científic permet anar una mica més enllà, es coneix millor el que es vol protegir i quines conseqüències se'n derivaran, tant per a l'espècie protegida com per a la resta d'espècies i l'ecosistema en general. Segons el rol que s'atribueix a les espècies, aquestes es poden classificar en quatre tipus: "espècies senyera", les que motiven la societat a invertir esforços per conservar-les, però que poden eclipsar les necessitats de la resta d'espècies; "espècies indicadores", el seguiment de les quals permet determinar l'estat de l'ecosistema; "espècies clau", que tenen un paper fonamental en el funcionament de la comunitat o de l'ecosistema on viuen; i "espècies paraigües", la conservació de les quals implica la protecció de tot l'ecosistema que les acull. Abans de començar un pla de protecció i seguiment cal ser prudent i estudiar bé quin és el rol d'una determinada espècie en el seu ecosistema i quines serien les conseqüències de la seva conservació a tots els nivells (Delibes, 2005).

Davant el gran nombre d'espècies que s'enfronten a una extinció imminent, els conservacionistes han de prendre decisions ràpides basant-se en la informació de què

disposin en aquell moment. És per això que el desenvolupament d'eines que generin moltes dades i la possibilitat d'analitzar-les ràpidament és essencial. Així, s'estan aplicant a la conservació d'espècies elements com la tecnologia del sistema de posicionament geogràfic (GPS), avenços matemàtics i la genètica. La genètica ha revolucionat moltes àrees de la ciència i de la vida en general. No és estrany, doncs, que aquesta branca de la biologia també comenci a ser fonamental dins el camp de la conservació biològica.

2.2. Genètica de la conservació

La genètica de la conservació és una nova disciplina que es basa en la informació genètica per tal d'establir els criteris de selecció en la protecció de les espècies. El seu objectiu és contribuir a la conservació de la diversitat genètica del planeta i a reduir el risc d'extinció de les espècies amenaçades. En primer lloc, doncs, és necessari determinar la quantitat de diversitat genètica i establir a quin nivell (individual, poblacional, d'espècie, de gènere, etc.) cal analitzar-la.

Darwin (1859) fou el primer a adonar-se de la importància de la variació existent en el manteniment de les poblacions naturals. Cap al 1970 va començar a preocupar el fet de mantenir les espècies vegetals que havien estat origen de les espècies cultivades i que estaven en recessió (Frankel, 1970; 1974), però no fou fins a la dècada de 1980 que es va aplicar la genètica de manera més generalitzada en conservació. Soulé (1980) té una importància fonamental en l'establiment de la biologia de la conservació a partir de 1980 com un camp multidisciplinari, que inclou la genètica, l'ecologia i la biologia de les espècies com a pilars. Actualment, gràcies al progrés de la genètica en general, s'estan fent grans avenços en genètica de la conservació: la incorporació de tècniques avançades de genotipatge permet generar moltes dades en poc temps; la seqüenciació de genomes complets ofereix molta informació bàsica; i la possibilitat d'usar mostres no invasives i forenses permet fer el seguiment de poblacions salvatges sense molestar els individus i també analitzar productes derivats d'espècies protegides.

Les espècies han de fer front a les alteracions constants a les que estan sotmesos els ambients, provocades principalment pel canvi climàtic, per la contaminació i per la introducció de nous competidors i noves malalties i paràsits. Davant d'aquests canvis, si les espècies no evolucionen per tal d'adaptar-se, s'extingeixen. La diversitat genètica d'una població és un reflex del seu potencial evolutiu i, doncs, de la seva capacitat d'adaptació, de manera que poblacions amb una bona diversitat poden respondre als

canvis mediambientals mentre que les poc diverses són incapaces d'adaptar-se i no poden sobreviure. Per això, la IUCN reconeix la necessitat de conservar la diversitat genètica com una de les tres prioritats de conservació global. D'altra banda, la pèrdua de diversitat genètica es pot relacionar amb una reducció de l'eficàcia biològica, ja que disminueix la taxa de reproducció i la de supervivència. És per això que en el seguiment de poblacions naturals i en el maneig de poblacions en captivitat es controla especialment el grau d'endogàmia i s'analitzen els nivells de diversitat genètica.

Les primeres mesures de diversitat genètica amb tècniques moleculars van ser possibles gràcies a la tècnica de l'electroforesi, que era capaç de separar les proteïnes segons la seva càrrega neta i el seu pes molecular i permetia detectar diferències entre les diferents formes d'una determinada proteïna. Els marcadors es van anomenar al·lozims, nom que deriva de la frase "variants al·lèliques dels enzims". Per poder aplicar aquesta tècnica, però, és necessari capturar l'animal d'estudi perquè requereix obtenir-ne sang o sacrificar-lo per aconseguir-ne el fetge o un ronyó. Pel que fa a l'estudi dels cromosomes, normalment mantenen un nombre, mida i forma constants dins d'una espècie, però diferent entre espècies, de manera que les tècniques cromosòmiques s'acostumen a utilitzar en anàlisis interespecífiques. Finalment, les anàlisis amb marcadors de DNA, que inicialment també requerien grans quantitats de mostra, actualment poden realitzar-se sobre mostres biològiques molt petites gràcies a la tècnica de la reacció en cadena de la polimerasa (PCR). En els últims anys s'han desenvolupat un gran nombre de mètodes per mesurar la diversitat genètica a partir de seqüències de DNA (veure apartat 2.4 d'aquesta Introducció).

La genètica ha estat essencial en conservació per tractar diverses qüestions crítiques:

- Gestionar poblacions petites, salvatges o en captivitat, per minimitzar la consanguinitat i la pèrdua de diversitat genètica.

La consanguinitat, que afecta poblacions petites i endogàmiques, produeix pèrdua de diversitat genètica i pot fer baixar la taxa de supervivència. Això es dona perquè l'homozigositat pot desemmascarar mutacions deletèries i reduir la capacitat d'adaptació de les poblacions, de manera que esdevenen molt més sensibles als canvis ambientals. En les poblacions captives s'ha observat que una de les causes principals de reducció de l'eficàcia biològica és la consanguinitat (Kalinowski *et al.*, 2000). Per tal de tractar de reduir l'endogàmia, en molts casos s'han dissenyat programes d'aparellament entre individus de diferents poblacions de manera que en

els descendents augmenti la diversitat genètica. En el cas de poblacions salvatges petites, que també es veuen afectades per la depressió per consanguinitat deguda als efectes acumulatius de la deriva genètica (Hedrick, Kalinowski, 2000), es poden establir plans d'alliberament d'individus d'altres poblacions genèticament diferents per fer augmentar la diversitat genètica i reduir així el risc d'extinció.

- Dissenyar programes de reintroducció definint les àrees adequades i els individus de genotip adient.

Abans de realitzar una reintroducció cal tot un estudi previ per establir la regió adequada per fer la reintroducció: cal assegurar-se que l'hàbitat no està contaminat, que disposa de suficients recursos i que no hi ha espècies invasives que puguin desplaçar l'espècie d'interès ni espècies que es puguin veure desplaçades per l'espècie monitoritzada. A més, cal seleccionar els individus que es volen reintroduir de manera que tinguin un genotip el més semblant possible al que hi havia a la regió, però prou diferent per poder establir una població amb una diversitat genètica suficient per mantenir-se estable. Després de la reintroducció és molt important fer un seguiment de la població per estudiar l'èxit del programa. Això mateix s'aplica a programes de reforçament, en què, a més, és necessari tenir en compte les característiques genètiques de la població simpàtrica perquè no sigui desplaçada sinó que s'aconsegueixi un increment de l'eficàcia biològica per la incorporació de nous genotips a la població original (Goossens *et al.*, 2002).

- Resoldre incerteses taxonòmiques i estructures poblacionals per poder definir les unitats de maneig i identificar poblacions amenaçades.

La definició de les unitats de maneig és un punt clau en conservació per poder determinar sense dubte quines poblacions, espècies, etc., cal protegir i establir així les bases legals per a la conservació (O'Brien, 1994). Per això és essencial caracteritzar bé les relacions filogenètiques i definir sense incerteses taxonòmiques les diferents categories sistemàtiques (subespècies, espècies, gèneres, etc.). Els conservacionistes han creat el terme d'unitat evolutivament significativa (ESU) per tal de poder establir plans de gestió per a la conservació de les unitats adequades (O'Brien, Mayr, 1991). En aquest sentit, la genètica pot ser fonamental per resoldre dubtes filogenètics, especialment quan la classificació morfològica és confusa.

- Detectar i predir l'efecte de fenòmens d'hibridació i introgressió. Detectar i definir les espècies invasives.

Una de les amenaces per a les espècies en perill d'extinció és la hibridació amb altres espècies properes, que provoca pèrdua d'identitat per la incorporació de gens aliens. La hibridació complica la delimitació i la definició de les diferents unitats evolutives i fa difícil establir plans de gestió (Haig *et al.*, 2004). Tot i que els híbrids entre espècies amenaçades no gaudeixen de protecció, s'ha vist que alguns programes de protecció s'estaven duent a terme sobre espècies híbrides. És important, doncs, detectar a temps possibles fenòmens d'introgressió i definir quines espècies poden ser invasives per tal d'evitar la hibridació amb les espècies amenaçades properes, o fins i tot la seva substitució.

- Ús en genètica forense.

Les tècniques moleculars, que són cada vegada més sensibles, permeten analitzar productes derivats de les espècies i saber de quina espècie es tracta, de manera que juguen un paper fonamental en el control de la comercialització de productes i d'espècies. D'aquesta manera és possible detectar delictes contra el medi ambient i ajudar en l'aplicació de la legislació (Domingo-Roura *et al.*, 2006; Wan, Fang, 2003; Yan *et al.*, 2005).

- Entendre la biologia de les espècies.

Trobar la relació entre l'eficàcia biològica o l'adaptació i els caràcters genètics dels individus o de les poblacions és bàsic per entendre el seu funcionament. La relació fenotip-genotip explica quins canvis genètics confereixen un avantatge adaptatiu que ha estat afavorit per la selecció natural. D'altra banda, la genètica permet estudiar els sistemes d'aparellament i de reproducció de les espècies, fins i tot de les més elusives i amenaçades, i fer estudis de parentiu dins una comunitat (p.e. Goossens *et al.*, 1998) identificant els seus membres individualment, a més de determinar la proporció sexual (*sex ratio*). És possible, també, establir censos de les poblacions a partir de mostres no invasives, és a dir, sense veure els animals ni haver de capturar-los, i estimar diversos paràmetres poblacionals que permeten descriure la història de les poblacions, com els patrons de migració de les espècies i processos de coll d'ampolla o d'expansió. A més, les tècniques moleculars permeten estimar la grandària poblacional

efectiva (N_e), és a dir, el número d'individus que contribuiran amb els seus al·lells a la següent generació. La N_e , que normalment és inferior al número real d'individus presents en una població, pot ser un indicador del grau de perill en què es troba aquesta població. Fins i tot s'ha definit una grandària mínima necessària perquè la població sigui viable (MVP, Shaffer, 1981).

2.3. Mostreig no invasiu

El desenvolupament de la tècnica de la PCR va obrir les portes a un tipus d'anàlisi menys agressiva per a les espècies perquè permetia obtenir moltes còpies de DNA a partir de molt poca quantitat inicial, de manera que no calia utilitzar teixits d'animals vius (biòpsies). Podem distingir tres mètodes diferents d'obtenció de mostres segons Taberlet et al. (1999):

- Mostreig destructiu: És aquell que implica la mort de l'animal. Va ser un mètode molt usat per a l'estudi d'al·lozims i en estudis de DNA mitocondrial abans de l'aparició de la PCR.
- Mostreig no destructiu: S'obté una mostra directament de l'animal, com biòpsies, extraccions de sang, plomes arrencades, etc., però no sempre és necessari capturar els individus (per exemple, amb balenes s'han utilitzat pistoles de dards per a biòpsies).
- Mostreig no invasiu: És el tipus de mostreig que permet obtenir la mostra sense necessitat de capturar ni molestar, i a vegades ni tan sols veure, l'animal d'estudi. Es tracta de recollir fonts de DNA que l'animal deixa darrera seu.

El mostreig no invasiu ha facilitat l'estudi d'espècies rares, elusives o en perill d'extinció en què el mostreig amb mètodes invasius no és factible, bé per dificultats tècniques (no es troba l'animal), bé pel perill que suposaria per a l'espècie. En espècies amenaçades, una pertorbació de la població deguda a la captura d'animals pot ser molt perjudicial.

Actualment és possible utilitzar una gran varietat de mostres obtingudes de manera no invasiva per extreure el DNA necessari per a les anàlisis genètiques. S'han utilitzat principalment mostres de femtes (p.e. Frantz *et al.*, 2004) i pèls (p.e. Taberlet *et al.*, 1997) perquè són fàcils d'obtenir en espècies salvatges, però també s'han pogut recollir altres mostres menys comuns, com orina (p.e. Valière, Taberlet, 2000), cèl·lules bucals (p.e. Brooks *et al.*, 2003) o plomes (p.e. Gautschi *et al.*, 2003). Pel que fa als primats, hi ha

estudis genètics realitzats a partir de femtes (p.e. Constable *et al.*, 2001; Vigilant *et al.*, 2001) i pèls (p.e. Gagneux *et al.*, 1999; Goldberg, Wrangham, 1997) i també a partir d'orina (p.e. Hayakawa, Takenaka, 1999), semen (Domingo-Roura *et al.*, 2004) i cèl·lules bucals obtingudes de restes de menjar mastegat (p.e. Sugiyama *et al.*, 1993) o de pals rosegats (Inoue, 2007).

D'altra banda, existeixen altres fonts per aconseguir mostres d'espècies que no impliquen pertorbar les poblacions, com són les carcasses d'animals trobats morts i els espècimens conservats en museus, que permeten estudiar les característiques poblacionals al llarg del temps. Els avenços tècnics han permès que es pugui obtenir DNA adequat per realitzar estudi genètics a partir de mostres antigues, des de restes de varies dècades d'antiguitat (p.e. Bellinger *et al.*, 2003) o de centenars d'anys (p.e. Rosenbaum *et al.*, 2000) fins a fòssils de milers d'anys, per exemple d'organismes ja extingits (p.e. Rogaev *et al.*, 2006). El desenvolupament de la genètica no invasiva ha estat una gran fita en la conservació de les espècies. La possibilitat de genotipar sense haver de capturar permet estudiar espècies que no podien ser analitzades d'una altra manera, a més de ser una estratègia coherent amb la pròpia definició de conservació. Tanmateix, l'ús de mostres no invasives no està exempt de problemes metodològics i limitacions. A partir de mostres no invasives se sol obtenir poca quantitat de DNA i acostuma a ser de baixa qualitat, és a dir, pot estar degradat, contaminat i contenir inhibidors de la PCR (a l'apartat 1 del Resum global es discuteixen les possibles solucions a aquestes limitacions).

La poca quantitat de DNA que es pot extreure de les mostres no invasives ha estat una de les principals limitacions, que ha fet molt difícil l'ús de marcadors nuclears i ha reduït la possibilitat de fer rèpliques, de manera que les anàlisis genètiques de mostres no invasives s'han fet principalment a partir de marcadors mitocondrials (més abundants). És per això que s'han anat desenvolupant mètodes d'amplificació total del genoma (*Whole Genome Amplification*, WGA) que contribueixen a superar aquesta limitació, ja que permeten l'obtenció d'un gran nombre de còpies de DNA de tot el genoma a partir de quantitats mínimes de DNA inicial. S'han desenvolupat diversos mètodes de WGA. Els mètodes basats en la PCR, com el degenerate oligonucleotide-primed PCR (Telenius *et al.*, 1992) i la primer extension pre-amplification (Zhang *et al.*, 1992), produeixen una cobertura parcial del genoma i tendeixen a produir una amplificació desequilibrada dels microsatèl·lits i dels SNPs, a més de poder introduir variació falsa en la seqüència dels productes. Recentment, s'han desenvolupat altres mètodes més fiables i eficients de WGA, com l'amplificació per desplaçament múltiple (*Multiple Displacement Amplification*,

MDA) (Dean *et al.*, 2002). La reacció d'MDA es basa en les característiques de la polimerasa del bacteriòfag phi29 que, a més de replicar, és capaç d'obrir la cadena de DNA i avançar de manera eficient i amb una taxa d'error molt baixa, ja que té activitat 3'→5' exonucleasa. Amb aquesta tècnica es generen milers, o fins i tot milions, de còpies de tot el genoma a partir de pocs nanograms de DNA original amb una bona cobertura de tot el genoma (figura 7). S'ha comprovat que aquest sistema té una gran fidelitat quan s'utilitza sobre DNA extret de mostres humanes, fins i tot a partir de molt poca quantitat de DNA inicial (Lasken, Egholm, 2003; Rook *et al.*, 2004; Sorensen *et al.*, 2004). Tot i el gran potencial d'aquest mètode, la seva fidelitat sobre mostres de primats no humans recol·lectades de manera no invasiva encara no havia estat testada.

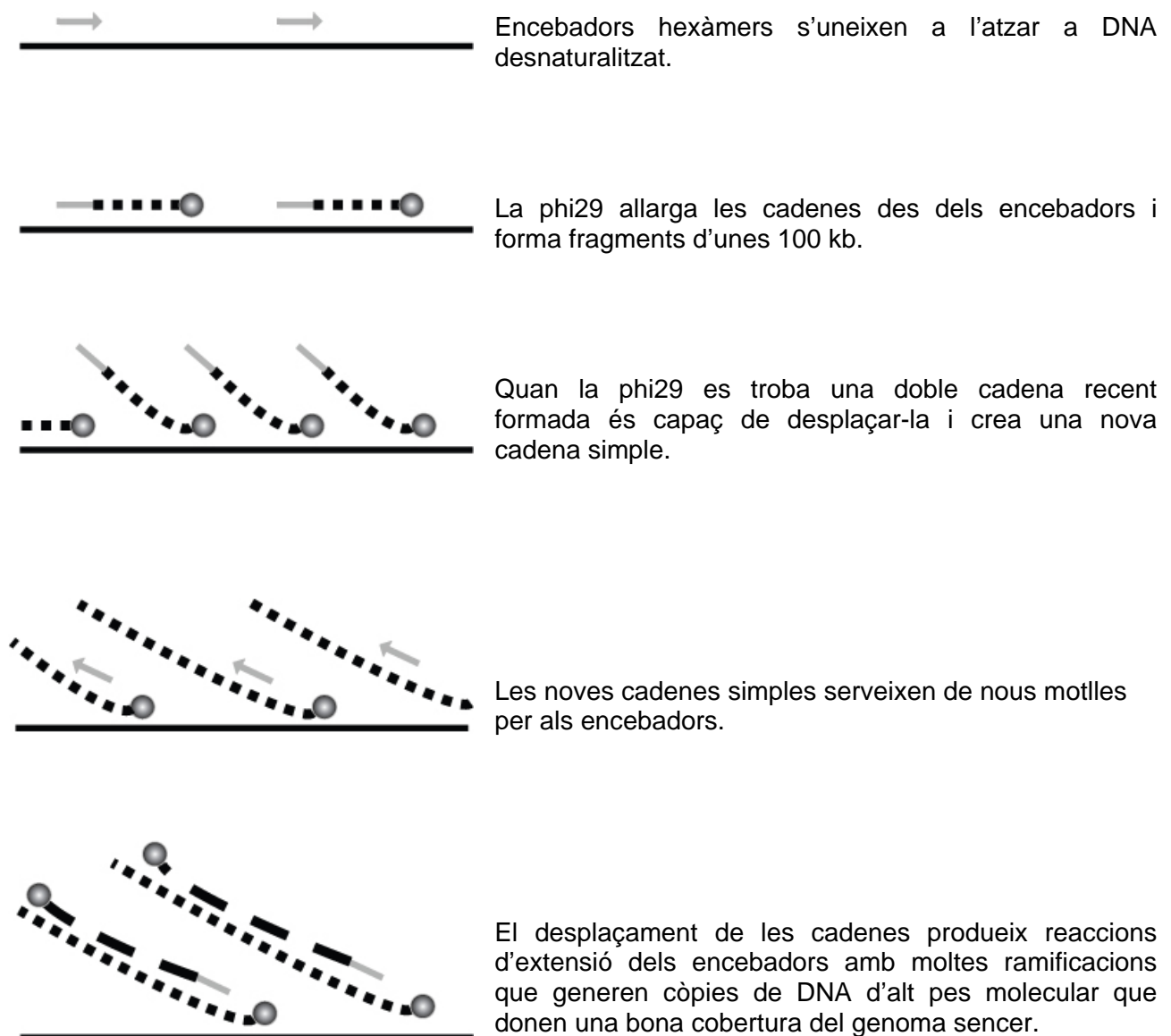


Figura 7. Esquema del procés de MDA. (Adaptat a partir del manual d'instruccions del fabricant del kit comercial, GE Healthcare).

2.4. El genoma eucariota

A les cèl·lules eucariotes el DNA es troba majoritàriament al nucli cel·lular organitzat en cromosomes, però una petita part del genoma és extranuclear i es troba localitzat en orgànuls citoplasmàtics, plastidis (p.e. cloroplasts) i mitocondris.

2.4.1. El DNA mitocondrial

El mitocondri és un orgànul de les cèl·lules eucariotes, limitat per una membrana unitària doble, que conté una petita fracció de DNA extranuclear. A diferència del DNA nuclear, el DNA mitocondrial és haploide, s'hereta via materna (de les mares a la descendència) en la majoria de les espècies i no recombina. El genoma mitocondrial en mamífers és una doble cadena circular que no acostuma a superar els 20.000 parells de bases i presenta l'ordre dels gens, que no contenen introns, molt conservat (Singer, 1993). En humans, el DNA mitocondrial està format per 16.569 parells de bases i conté 37 gens: 13 codifiquen per proteïnes, 22 per RNAs de transferència i dos per RNAs ribosomals. El genoma mitocondrial gairebé no conté seqüències no codificadores, només una petita zona anomenada regió control –de 870 a 1400 parells de bases en mamífers (Sbisà *et al.*, 1997)–, amb dues regions hipervariables que presenten una taxa de mutació molt superior a la de la resta de regions i molt més elevada que la taxa de mutació nuclear. El fet que els organismes presentin molts mitocondris, entre 100 i 10.000 còpies per cèl·lula, permet obtenir una gran quantitat de DNA mitocondrial, cosa que facilita l'ús de mostres recol·lectades de manera no invasiva. La informació que aporta el DNA mitocondrial serveix per traçar únicament la línia materna, de manera que pot mostrar patrons de migració específics. Aquest marcador s'aplica especialment per avaluar relacions taxonòmiques i diferències entre poblacions d'una mateixa espècie i ha estat molt utilitzat per estudiar l'evolució humana recent (revisat a Pakendorf, Stoneking, 2005).

2.4.2. El DNA nuclear

El nucli de les cèl·lules eucariotes és un orgànul envoltat per una doble membrana que conté la major part del material genètic de la cèl·lula, organitzat en cromosomes, que són múltiples molècules de DNA lineal unides a proteïnes estructurals. Els organismes diploides presenten dues còpies de cada cromosoma dins el nucli cel·lular, cadascuna d'un progenitor, ja que el DNA nuclear és d'herència mendeliana (biparental). Durant la meiosi el DNA nuclear pateix processos de recombinació, és a dir, els cromosomes

homòlegs poden intercanviar material genètic. La taxa de mutació del DNA nuclear és, en general, baixa.

Els nivells més elevats de diversitat genètica en el DNA es troben en aquelles bases que tenen poca importància funcional, és a dir, les que no codifiquen per productes funcionals o aquelles en les que les substitucions no canvien la funció de la molècula. Per altra banda, la diversitat genètica més baixa es troba en aquelles regions de la molècula que són funcionalment més importants, tot i que poden haver-hi excepcions. Per exemple, en els vertebrats, el complex major d'histocompatibilitat presenta una variació genètica molt elevada tot i tractar-se d'una regió codificadora perquè la selecció natural afavoreix aquesta variació, que és necessària per la funció d'aquest complex. La major part del DNA d'un organisme no codifica per productes funcionals, de manera que canvis en aquestes regions no tindrien un significat selectiu. A més, el 70% dels canvis dins dels gens són silenciosos i, doncs, tampoc solen estar afectats per la selecció natural. Les regions no codificadores inclouen tant les regions entre gens com les regions intragèniques que es transcriuen però no es tradueixen, els introns. Els introns són seqüències de DNA que interrompen la regió codificadora d'un gen (els exons) i que són eliminats de l'RNA missatger abans de la traducció. Tot i no contribuir a la formació de la proteïna, però, els introns poden contenir segments importants en la regulació del gen. D'altra banda, existeixen còpies de gens funcionals que han estat inactivades per mutació, els pseudogens, que no produeixen proteïnes completes (la majoria no es tradueixen o ni tan sols arriben a transcriure's). Tant les regions intergèniques com els introns i els pseudogens evolucionen de manera més ràpida que les regions codificadores.

2.4.3. Marcadors moleculars nuclears

Un marcador genètic aporta informació sobre la variació al·lèlica en un determinat locus. El procés d'escollir el marcador molecular adequat és un pas crític en genètica de la conservació, ja que la tria d'un marcador inapropiat pot comportar l'establiment d'accions de conservació incorrectes. Podem distingir entre quatre tipus de marcadors moleculars de DNA:

- Minisatèl·lits

Els minisatèl·lits o VNTR (*Variable Number Tandem Repeats*) són segments de DNA formats per seqüències de 10 a 100 bases, generalment riques en CG, repetides en tàndem i de mida variable, que es troben distribuïts per tot el genoma (Armour, 1999). La

variabilitat en el número de repeticions es genera per entrecreuaments desiguals. En humans, el 90% dels minisatèl·lits es troben a la regió subtelomèrica dels cromosomes i la seqüència dels telòmers és un minisatèl·lit.

- Microsatèl·lits

Els microsatèl·lits van ser descoberts per primera vegada a la dècada de 1980 (Miesfeld *et al.*, 1981) i de seguida es va veure el seu potencial com a marcadors moleculars gràcies al seu polimorfisme (Litt, Luty, 1989; revisat a Schlötterer, 2000b). Els loci microsatèl·lits són segments curts de DNA formats per seqüències d'un a sis nucleòtids de llargada repetides en tàndem. El número de repeticions en els microsatèl·lits és molt variable (de 5 a 100 aproximadament) com a conseqüència del mecanisme d'slippage. L'slippage és un procés mutacional específic de seqüències repetides en tàndem que té lloc durant la replicació i que genera un increment o una disminució normalment d'una sola unitat repetitiva. La taxa de mutació pot anar des de 10^{-2} fins a 10^{-6} per generació (Schlötterer, 2000a). En general, els microsatèl·lits amb més repeticions solen ser més variables que els més curts. Aquests marcadors es troben distribuïts de manera més o menys ubiqua per tot el genoma, tot i que la seva presència en regions codificadores està molt restringida. Els avantatges principals dels microsatèl·lits són l'elevada variabilitat que presenten, l'herència codominant nuclear i la possibilitat de tipar-los en mostres no invasives. Generalment han de ser desenvolupats per a cada espècie que s'estudia, però espècies molt properes solen compartir els microsatèl·lits i els encebadors solen funcionar bé d'una a l'altra. (A l'apartat 3 del Resum global es discuteixen els avantatges i inconvenients dels microsatèl·lits front els SNPs)

- Insercions i delecions (indels)

Els indels són insercions o delecions en la seqüència de DNA que poden afectar des d'una sola base fins a grans segments de DNA. Dins de les regions codificadores l'efecte dels indels sol ser devastador, ja que, a menys que la llargada de l'indel sigui un múltiple de tres (i afecti, doncs, codons sencers), s'altera la pauta de lectura i la proteïna que es formarà no serà funcional. Per això els indels són comuns en regions no codificadores però molt poc habituals en regions codificadores.

Els genomes eucariotes estan formats, en una gran proporció, per seqüències de DNA que s'han pogut moure a nous llocs dins del genoma. Aquestes grans insercions poden ser de dos tipus diferents:

- Transposons: Són elements que codifiquen per una transposasa que els escindeix del genoma i els permet ser inserits en un lloc nou. No es generen noves còpies.
 - Retroposons: Són elements transponibles que es mouen a través d'un intermediari d'RNA utilitzant una transcriptasa inversa, és a dir, fan còpies d'ells mateixos. Els LINEs (*Long Interspersed Elements*) i els SINEs (*Short Interspersed Elements*) són exemples de retroposons.
- Mutacions puntuals

Les mutacions puntuals són els canvis que suposen la substitució d'una sola base nucleotídica per una altra dins d'una seqüència de DNA. Sovint es consideren també mutacions puntuals les insercions o delecions que involucren una sola base nucleotídica. La taxa de mutació nucleotídica és de l'ordre de 10^{-9} (Graur, 1999) mutacions per posició per any. Les mutacions puntuals que es donen dins de seqüències codificadores es classifiquen en diferents tipus segons l'efecte que produeixen: mutacions sinònimes, que no provoquen canvi d'aminoàcid, i mutacions no sinònimes, que alteren la seqüència de la proteïna i que poden dividir-se en mutacions sense sentit (*nonsense*), les que generen un codó de terminació, i mutacions amb error de sentit (*missense*), les que generen un codó que codifica per un altre aminoàcid. En general, les mutacions puntuals produïdes en seqüències no codificadores no tenen conseqüències en la funcionalitat i s'anomenen mutacions silencioses, tot i que hi ha excepcions, com les mutacions que alteren el promotor, les regions d'empalmament dels introns o les regions reguladores. Recentment, s'ha començat a parlar de polimorfismes d'un sol nucleòtid (SNPs) per fer referència als canvis en la seqüència de DNA que afecten una sola base nucleotídica (veure apartat 2.5).

2.4.4. Tècniques de genotipatge

Els diferents marcadors moleculars poden ser analitzats amb diverses tècniques de genotipatge. El mecanisme de genotipatge ideal hauria de ser simple, de baix cost i sensible i hauria de generar un número elevat de dades fiables i comparables entre laboratoris. En els últims anys els mètodes de genotipatge han patit un gran

desenvolupament gràcies a les millores tècniques i actualment existeixen moltes eines disponibles per a l'estudi de la diversitat genètica. Les tècniques de genotipatge que s'han utilitzat més freqüentment són:

- RFLPs (*Restriction Fragment Length Polymorphisms*)

El descobriment dels enzims de restricció va permetre el desenvolupament de la tècnica dels RFLPs, polimorfismes en la llargada dels fragments de restricció. Aquesta tècnica permet l'estudi de la variabilitat a través de la detecció de substitucions d'una base a la seqüència de reconeixement d'un enzim de restricció. Per dur a terme la tècnica, el DNA purificat es talla amb un enzim de restricció i es corren els fragments resultants en un gel per separar-los segons la mida. Les cadenes dobles de DNA es desnaturalitzen i es fa una transferència de les cadenes simples a una membrana (transferència *Southern*). La membrana es posa en una solució amb sondes del locus d'interès marcades radioactivament i, després del procés d'hibridació, s'asseca i es fa una autoradiografia amb un film fotogràfic, de manera que es poden veure les diferències entre els patrons de restricció dels individus que evidencien variació en els llocs de tall dels enzims de restricció. Els RFLPs posen de manifest una herència codominant i avaluen la variació en zones conegudes. Però presenten desavantatges clars: cal que hi hagi sondes pel locus d'interès, la seva variabilitat és moderada i calen grans quantitats de DNA, de manera que no poden aplicar-se sobre mostres no invasives. Aquest mètode està sent substituït per mètodes basats en PCR. Per exemple, la PCR-RFLP és una tècnica que aplica exactament el mateix principi que els RFLPs però sobre fragments amplificats per PCR.

- DNA *fingerprinting*

El DNA *fingerprinting* és la tècnica utilitzada per revelar el patró de minisatèl·lits d'un individu. Per analitzar els minisatèl·lits cal purificar grans quantitats de DNA, tallar-lo amb enzims de restricció que alliberin el fragment minisatèl·lit i separar els fragments per la seva mida en un gel. Com amb els RFLPs, els fragments es desnaturalitzen per separar les dues cadenes de DNA i es transfereixen a una membrana que es posa en una solució amb sondes (cadenes simples de la unitat de repetició) marcades radioactivament. Després d'hibridar, rentar i assecar la membrana, es fa l'autoradiografia, que revela un patró de bandes de tipus codi de barres, l'anomenat DNA *fingerprint* o empremta genètica. El número de repeticions és tan variable que en general tots els membres d'una població tindran un patró de bandes propi i diferent al dels altres (excepte bessons univitel·lins). Els

avantatges dels minisatèl·lits són l'elevada variabilitat, que avalua la variació al llarg de molts loci escampats per tot el genoma i que no requereix un coneixement previ sobre la seqüència de DNA de l'espècie que es vol tipar. Però els loci no es poden identificar perquè provenen de diferents regions del genoma i no és possible aplicar la tècnica en mostres obtingudes de manera no invasiva perquè es requereix una gran quantitat de DNA. Per això actualment la tècnica del DNA *fingerprinting* també s'ha substituït per mètodes que es basen en la PCR, com els AFLPs o els RAPDs, i, doncs, que permeten l'obtenció d'una empremta genètica a partir de menys DNA inicial.

- RAPD (*Random Amplified Polymorphic DNA*)

La tècnica dels RAPDs es basa en l'ús d'encebadors inespecífics, de 10 a 20 bases, per amplificar per PCR fragments nuclears de DNA. Cada encebador pot generar diferents fragments, normalment de 100 a 200 bases de llarg, que se separen en gels d'agarosa o de seqüenciació i revelen un patró de presència-absència segons la variació que hi hagi en els llocs d'hibridació de l'encebador. El mètode dels RAPDs permet avaluar molts loci sense la necessitat de conèixer la seqüència per dissenyar encebadors específics i permet l'ús de mostres no invasives. Però el model presència/absència és d'herència dominant (presència dominant, absència recessiu) de manera que no obtenim tots els genotips existents. A més, la principal crítica a aquesta tècnica és que té una molt baixa reproductibilitat.

- AFLPs (*Amplified Fragment Length Polymorphisms*)

Els AFLPs són un mètode proper als RAPDs. El DNA genòmic es talla amb un enzim de restricció i fragments curts de DNA sintètic (adaptadors) de seqüència coneguda s'enganxen als extrems tallats. Llavors es realitza una PCR amb encebadors que complementen les seqüències de l'adaptador i dels llocs específics de restricció. S'obté un patró de presència/absència com amb els RAPDs, però s'aconsegueix una millor reproductibilitat.

- SSCP (*Single Strand Conformational Polymorphisms*)

La diversitat genètica entre productes de PCR de diferents individus es pot detectar també gràcies als SSCPs. En aquest mètode, les dues cadenes de DNA dels productes de PCR se separen a altes temperatures, es refreden immediatament i els productes de cadena simple es corren en un gel d'acrilamida a baixes temperatures. En aquestes condicions

les cadenes simples de DNA es pleguen sobre elles mateixes segons la seva seqüència i aquestes molècules corren en el gel segons la mida i la conformació, de manera que es pot detectar variació en les cadenes de DNA.

- Genotipatge de microsatèl·lits

La diversitat dels microsatèl·lits és detectada mitjançant les diferències en la mida de fragments de DNA amplificats, generalment fragments de 75 a 300 parells de bases. Es tracta de dissenyar encebadors a les regions conservades que flanquegen el microsatèl·lit, amplificar i separar segons la mida els fragments de DNA que en resulten en un gel d'agarosa o d'acrilamida. Es considera que la mida dels fragments és indicativa del número de repeticions del microsatèl·lit. Els fragments separats es poden detectar de diferents maneres: tenyint el gel amb bromur d'etidi, autoradiografiant el gel si els encebadors estaven marcats radioactivament o utilitzant un aparell autoseqüenciador si usem encebadors marcats amb fluorescència. Si un individu és heterozigot per un microsatèl·lit detectarem dues bandes de mida diferent, mentre que si és homozigot, només tindrà una banda per aquell microsatèl·lit.

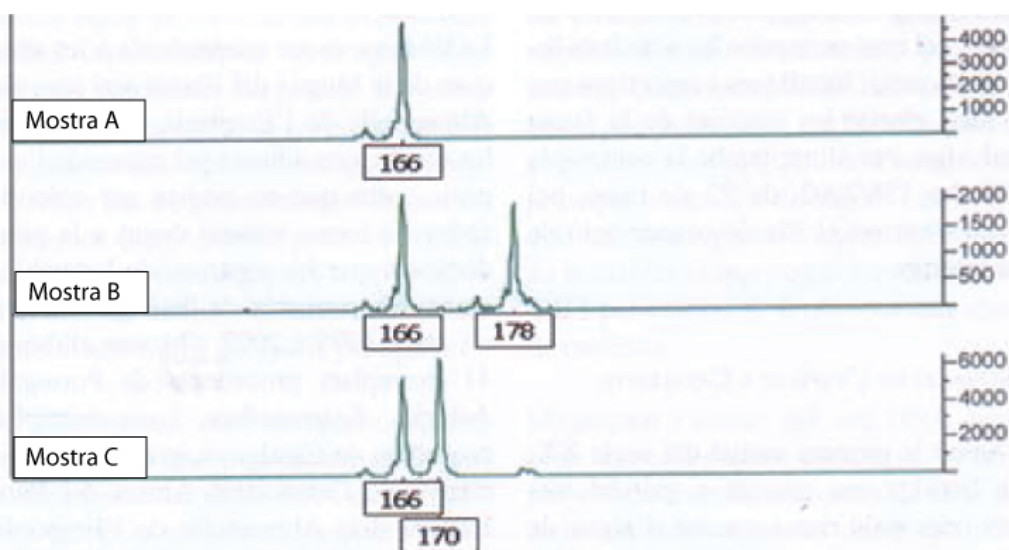


Figura 8. Resultat de l'anàlisi d'un marcador microsatèl·lit tetranucleotídic en tres mostres diferents, A, B i C. La mostra A és d'un individu homozigot per aquest marcador, mentre que els individus B i C són heterozigots, presenten un al·lel igual al de l'individu A però difereixen en l'altre al·lel. (A. Ferrando)

- Seqüenciació de DNA

La manera més directa de detectar variació genètica és determinar la seqüència de bases del DNA. S'acostuma a fer utilitzant seqüenciadors automàtics. En un inici es tractava d'un mètode molt car i lent, però les millores tècniques han permès disminuir considerablement

el cost econòmic i el temps i actualment es continua investigant per aconseguir millorar els diferents mètodes. Recentment ha començat a sorgir una nova tecnologia de seqüenciació, el piroseqüenciat, que promet abaratir i accelerar encara més aquest procés en un futur (veure apartat 4 del Resum global).

- Genotipatge d'SNPs (*Single Nucleotide Polymorphisms*)

Els SNPs, canvis introduïts per les mutacions puntuals en la seqüència de DNA, poden ser tipats a partir de molts mètodes diferents i se'n pot fer un genotipatge massiu aplicant tecnologia d'alt rendiment (veure apartat 2.5).

Comparació de les diferents tècniques

El mètode de DNA *fingerprinting* s'ha usat molt extensament en estudis forenses humans, tant en criminologia per identificar els autors de crims com en proves de paternitat, gràcies al patró exclusiu que presenta cada individu. Els RAPDs i els AFLPs han estat més utilitzats en estudis de plantes que en estudis d'animals. Aquestes tres tècniques, però, presenten dominància, de manera que és difícil comparar els resultats entre espècies i en poblacions amb diversitat genètica baixa serà difícil veure diferències en el patró de bandes. En general, doncs, les tècniques multiloci (DNA fingerprints, RAPD i AFLP) no s'utilitzen gaire actualment perquè són menys informatives tant pel que fa a distingir heterozigots d'homozigots (perquè són d'herència dominant) com pel que fa a la quantitat de variabilitat que es pot detectar.

És preferible, doncs, aplicar les tècniques que permeten observar una herència codominant perquè permeten distingir tots els al·lels en els individus heterozigots. La tria de la metodologia de genotipatge dependrà de diversos factors. En primer lloc, depèn del marcador que es vol aplicar i, després, dels recursos de què es disposa i de la qüestió a resoldre. Les característiques de les diferents tècniques de genotipatge de marcadors es troben resumides a la taula 2. Actualment el que més s'utilitza és el genotipatge de microsatèl·lits i el genotipatge d'SNPs (els avantatges i inconvenients d'aquests dos tipus de marcadors es troben discutits a l'apartat 3 de del Resum global).

Taula 2. Característiques dels diferents mètodes moleculars per avaluar la diversitat genètica (adaptada de Hartl D.L., 1997).

Mètodes	Font	Mostreig no invasiu	Cost	Temps de desenvolupament	Tipus d'herència	Anàlisi d'alt rendiment
Electroforesi	Sang, ronyó, fetge	No	Baix	Cap	Codominant	No
Genotipatge de microsatèl·lits	DNA	Sí	Moderat	Elevat	Codominant	Sí
DNA fingerprints	DNA	Sí	Moderat	Limitat	Dominant	No
RAPD	DNA	Sí	Baix-moderat	Limitat	Dominant	No
AFLP	DNA	Sí	Moderat-alt	Limitat	Dominant	No
RFLP	DNA	Sí	Moderat	Limitat	Codominant	No
SSCP	DNA	Sí	Moderat	Moderat	Codominant	No
Seqüenciació	DNA	Sí	Alt	Cap	Codominant	No
Genotipatge d'SNPs	DNA	Sí	Moderat-alt	Considerable	Codominant	Sí

2.5. SNPs

Un polimorfisme d'una sola base, o SNP, es produeix quan una posició en el DNA d'una espècie té dues bases alternatives amb una freqüència apreciable a la població (de més de l'1% la més baixa). Els polimorfismes d'un sol nucleòtid són els més abundants del genoma: els SNPs en els humans formen fins al 90% de les variacions genòmiques i es troben cada 100 o 300 parells de bases al llarg de tot el genoma. En poblacions naturals d'animals s'ha estimat que hi ha un SNP cada 500 parells de bases (Brumfield *et al.*, 2003; Morin *et al.*, 1994). Com que la taxa de mutació per a una determinada posició és baixa (uns 10^{-9} canvis per nucleòtid per any) (Graur, 1999), els SNPs solen tenir únicament dues variants, de manera que són marcadors bial·lèlics. Les mutacions puntuals poden ser transicions o transversions. Les transicions són la substitució d'una purina per una altra purina ($G \leftrightarrow A$) o d'una pirimidina per una altra pirimidina ($C \leftrightarrow T$), mentre que les transversions són la substitució d'una purina per una pirimidina (A o G per C o T) o d'una pirimidina per una purina (C o T per A o G). Tot i que hi ha el doble de transversions que de transicions possibles, els SNPs tendeixen a ser transicions com a conseqüència de la natura del procés mutacional (les transicions són molt més comuns) i del fet que les transversions en regions codificadores tenen una probabilitat més elevada de produir substitucions aminoacídiques que les transicions.

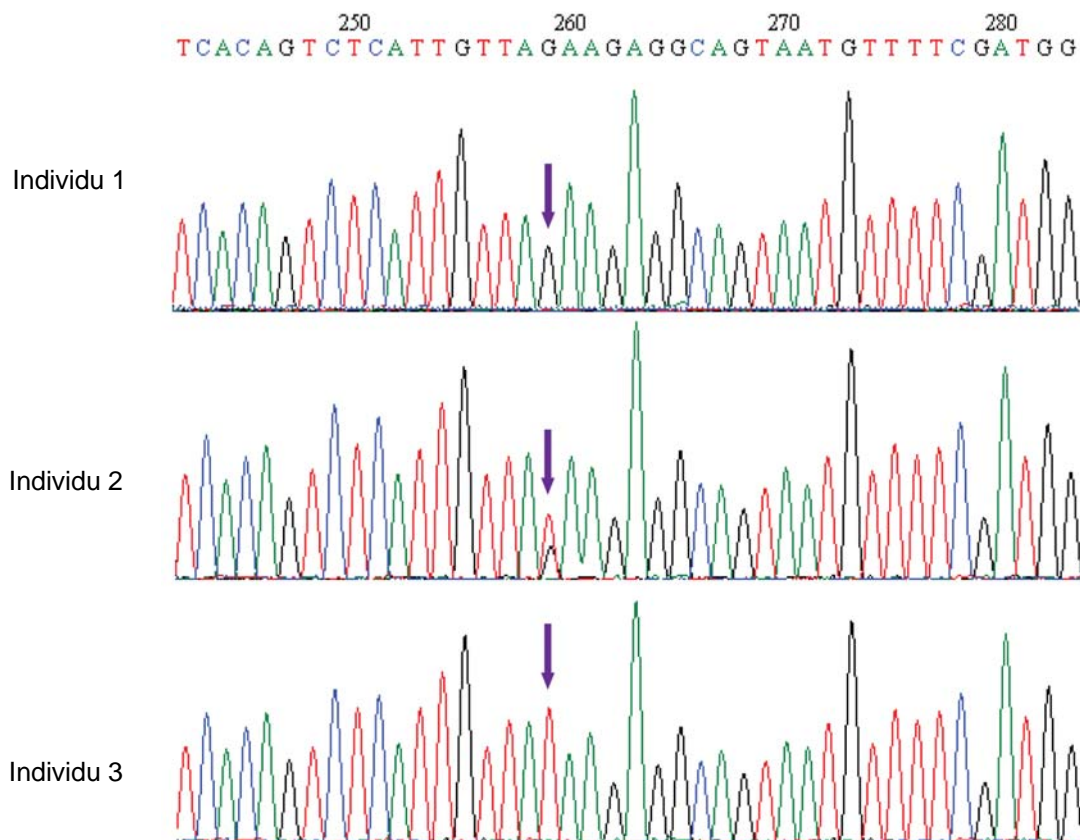


Figura 9. Cromatogrames de tres individus de la seqüència d'un locus amb polimorfisme G/T a la posició 259 en què es veu que l'individu 1 és homozigot pel nucleòtid G, l'individu 2 és heterozigot i l'individu 3 és homozigot pel nucleòtid T.

Els SNPs tenen un gran potencial per descriure la variació genètica en les poblacions naturals. De moment, han estat molt utilitzats per analitzar les diferències dins i entre poblacions humanes, especialment en l'estudi de gens relacionats amb malalties (p.e. Wang *et al.*, 1998) i en estudis d'evolució (p.e. Stephens *et al.*, 2001), gràcies als avantatges que aporten, com l'àmplia distribució per tot el genoma, l'elevada freqüència i, sobretot, que permeten fer genotipatges a gran escala. L'ús dels SNPs com a marcadors requereix dos passos principals: el descobriment dels loci i el genotipatge.

2.5.1. Descobriment d'SNPs

El descobriment dels SNPs és el procés d'identificar nous llocs polimòrfics en el genoma de l'espècie o població d'interès. En humans aquesta cerca molt sovint s'ha fet *in silico*, a partir de la informació genètica de molts individus emmagatzemada a les bases de dades públiques. Aquesta estratègia, però, no sempre és possible en altres espècies perquè no es disposa de la informació necessària i cal, doncs, identificar els SNPs a partir d'escanejors fets al laboratori. A més, els SNPs no es conserven d'una espècie a una altra, encara que sigui propera, de manera que els SNPs descoberts en una espècie no donen

informació de cap altra. No obstant, la informació de seqüència disponible per a diferents espècies model sí que pot ser útil per al descobriment de nous SNPs en espècies no model perquè permet aplicar diferents estratègies eficients. Per exemple, es poden dissenyar encebadors en regions homòlogues conservades, anomenades CATS (mètode desenvolupat per Lyons et al. (1997); aplicat a SNPs per Aitken et al. (2004)), o bé amplificar microsatèl·lits ja caracteritzats per analitzar-ne les regions flanquejants. Quan no existeix informació de seqüències útil per aplicar aquestes estratègies, l'alternativa és seqüenciar a l'atzar clons d'una llibreria genòmica de DNA de l'espècie d'interès i, a partir de les seqüències obtingudes, dissenyar encebadors específics (Primmer *et al.*, 2002). En qualsevol cas, una vegada aconseguits els encebadors adequats, és necessari seqüenciar diferents individus o conjunts de DNA per trobar els polimorfismes. Això suposa uns costos de seqüenciació molt grans i per això s'han aplicat mètodes alternatius per minimitzar l'esforç de seqüenciació, per exemple detectant en primer lloc els fragments que poden contenir polimorfismes. Aquest escaneig inicial es pot realitzar emprant diferents mètodes, com els AFLPs (Bensch *et al.*, 2002; Nicod, Largiader, 2003) o la DHPLC (Stone *et al.*, 2002; Underhill *et al.*, 1997).

DHPLC

Oefner & Underhill (1995) van desenvolupar la metodologia anomenada *Denaturing High-Performance Liquid Chromatography* (DHPLC) i van demostrar que és una tècnica eficient per comparar d'una manera automatitzada seqüències amplificades (Cotton, 1997; Ophoff *et al.*, 1996; Underhill *et al.*, 1996). La DHPLC es basa en la detecció d'heterodúplex i s'utilitza tant per identificar nous polimorfismes com per fer el genotipatge de les mostres. Aquesta tècnica únicament requereix disposar d'una seqüència de DNA adequada per dissenyar-hi encebadors, de manera que qualsevol segment de DNA d'uns quants centenars de bases de llarg pot ser amplificat per PCR i testat per detectar variació amb la DHPLC. Un cop s'obtenen els productes de PCR, s'ha de realitzar un procés de desnaturalització i rehibridació per permetre la formació d'heterodúplex. Si s'analitzen seqüències haploides, com en el cas del cromosoma Y, és necessari fer una barreja de productes de PCR abans de la desnaturalització. Segons si es formen només homodúplexs –no hi ha polimorfisme– o apareixen heterodúplexs –existeix polimorfisme– el patró observat serà diferent i molt fàcilment discernible (figura 10).

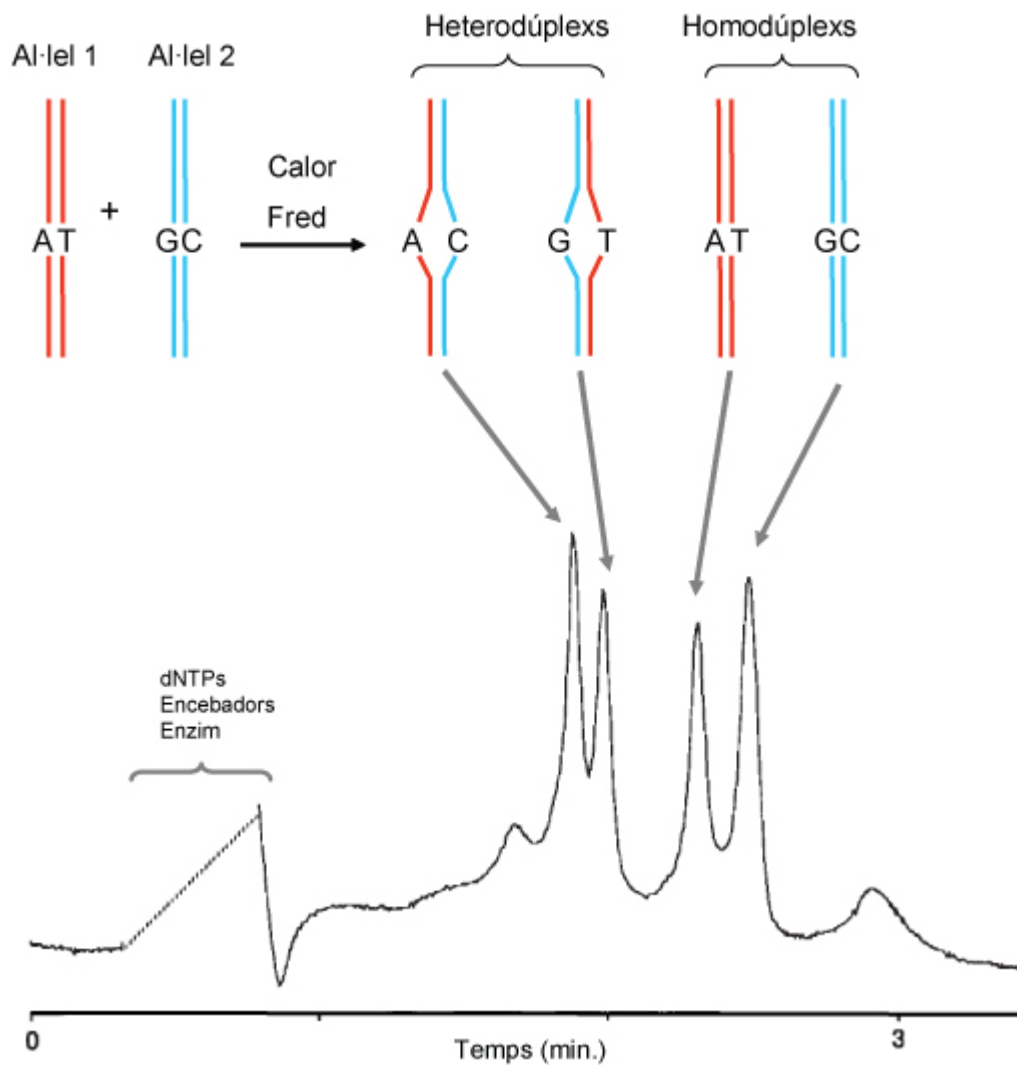


Figura 10. Esquema del resultat d'una anàlisi de DHPLC d'una mostra polimòrfica en què veiem quatre pics diferents corresponents als dos heterodúplexs i als dos homodúplexs possibles. (Adaptat de Xiao, Oefner, 2001)

Aquest mètode és especialment adequat quan s'estudien amb seqüències poc variables i, doncs, quan el patró que es detecta més freqüentment és l'homodúplex, com és el cas del cromosoma Y. En aquests casos no és necessari seqüenciar tots aquells fragments que hagin resultat monomòrfics en tots els individus comparats i només les mostres que hagin donat un patró d'heterodúplex seran seqüenciades per tal de caracteritzar el polimorfisme. La DHPLC, doncs, ofereix una opció eficient per revelar nous marcadors bial·lèlics en qualsevol genoma. D'altra banda, a més de participar en el seu descobriment, aquesta tècnica també permet determinar els perfils al·lèlics dels marcadors quan s'usa per realitzar assajos de genotipatge, tot i que per genotipar s'ha d'emprar amb més precaució per la possible incapacitació a l'hora de detectar una petita fracció dels loci polimòrfics.

2.5.2. Genotipatge d'SNPs

Una vegada identificat el número necessari d'SNPs per realitzar l'estudi d'interès, cal triar el mètode de genotipatge adequat. Aquesta tria depèn de molts factors (objectiu de l'estudi, recursos del laboratori, pressupost, etc.) i existeixen molts mètodes possibles (Vignal *et al.*, 2002). Els mètodes de PCR-RFLP i SSCP s'han utilitzat molt perquè únicament requereixen l'equipament bàsic d'un laboratori de biologia molecular, però no permeten fer genotipatge a gran escala i no són prou sensibles. Recentment s'ha desenvolupat un nou mètode, l'extensió d'encebador, que pot aplicar-se tant per extensió de base simple (SBE) com per extensió d'encebador específica d'al·lel (ASPE). Tant l'SBE com l'ASPE poden realitzar-se amb l'equipament estàndard d'un laboratori, però, alhora, es poden fer més eficients gràcies a tecnologia especialitzada, com els *microarrays* de DNA (p.e. Syvänen, 2001) o les anàlisis de citometria de flux (p.e. Taylor *et al.*, 2001) i així és possible aconseguir generar centenars o milers de genotips simultàniament. En l'SBE es dissenyen encebadors situats just a la posició 3' de l'SNP i es fa l'extensió amb dideoxinucleòtids, mentre que a l'ASPE es dissenyen dos encebadors, cadascun acabat en un dels dos nucleòtids possibles de l'SNP, i es fa l'extensió amb dideoxinucleòtids i una polimerasa d'alta fidelitat que amplifiqui només a partir de l'encebador que hibrida perfectament. A partir d'aquí, es poden aplicar diferents mètodes de detecció: electroforesi en gel (p.e. Makridakis, Reichardt, 2001), *microarrays* (p.e. Syvänen, 2001) o polarització fluorescent (p.e. Hsu *et al.*, 2001), per exemple. La detecció per electroforesi en gel ja ha estat comercialitzada per Applied Biosystems amb el kit *SNaPshot*, que utilitza l'equipament d'electroforesi fluorescent multicolor per detectar els ddNTPs incorporats.

Microarray de miniseqüència

Una metodologia prometedora que fa avançar cap als genotipatges a gran escala d'SNPs és l'ús de xarxes d'oligonucleòtids immobilitzats sobre *microarrays*. L'ús del format *microarray* permet reduir els costos de genotipar perquè s'analitzen simultàniament molts SNPs en cada mostra, moltes mostres alhora i permet utilitzar volums de reacció petits. Actualment, per genotipar SNPs s'utilitzen majoritàriament tres principis de reacció diferents: hibridació amb sondes d'oligonucleòtids específics d'al·lel (ASO) (Hacia *et al.*, 1998), lligació nucleotídica (Gerry *et al.*, 1999) i extensió d'encebador assistida per una DNA polimerasa (Pastinen *et al.*, 1997). Aquests tres tipus de reacció s'han utilitzat en un format de *microarray*, però les hibridacions ASO fetes simultàniament no permeten la discriminació dels genotips d'SNPs en genomes grans diploides (Wang *et al.*, 1998), de

manera que per genotipar SNPs de forma multiplex es prefereixen els mètodes assistits per enzims (Raitio *et al.*, 2001).

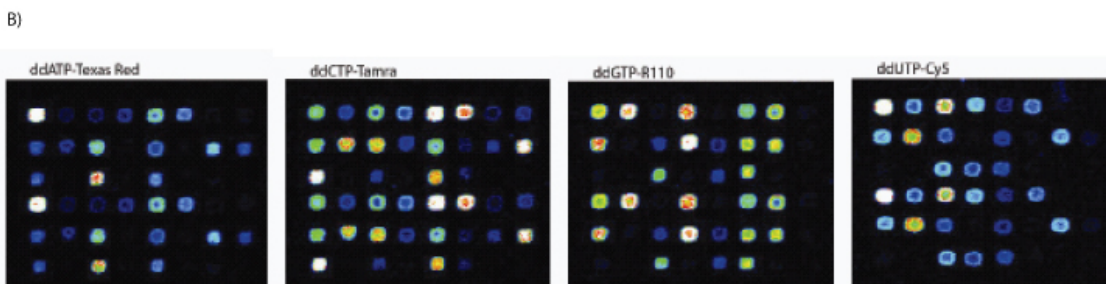
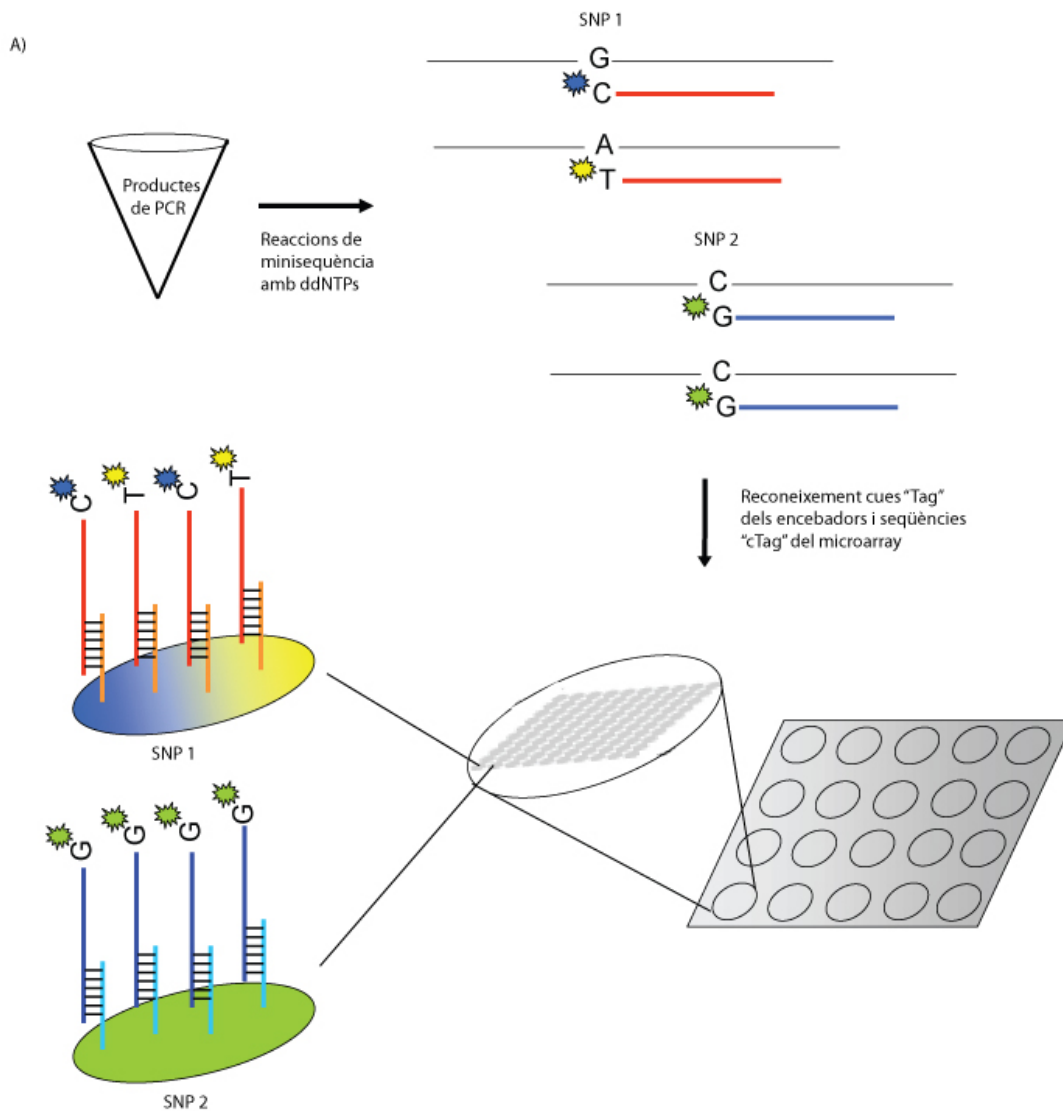


Figura 11. A) Principi del sistema del Tag-array de minisequència. B) Exemple de l'escaneig d'un genotipatge (Adaptat de Rönn *et al.*, 2006).

En l'extensió d'encebador assistida per una DNA polimerasa, es fa una reacció de minisequència, és a dir, la polimerasa estén els encebadors que hibriden adjacents a les posicions polimòrfiques amb la incorporació de dideoxinucleòtids marcats complementaris

als al·lels. En el format *microarray* els encebadors estan units covalentment al *microarray* i les reaccions es produeixen sobre aquesta plataforma. D'altra banda, existeix un sistema de miniseqüència alternatiu, el *Tag-array*, en el qual s'immobilitzen sobre el *microarray* oligonucleòtids genèrics ("cTags"), en lloc dels encebadors de detecció. Aquest sistema implica realitzar primer reaccions de miniseqüència cícliques en solució amb ddNTPs marcats amb fluorescència i utilitzant els encebadors de detecció als quals s'han afegit seqüències 5'-tag. Cada posició polimòrfica té el seu propi "tag" específic que és complementari a un dels "cTags" units al micorarray i així, quan els productes de les reaccions de miniseqüència s'apliquen sobre el *microarray*, cada "tag" hibridarà amb el seu corresponent "cTag". La fluorescència dels encebadors capturats es mesura en un escàner. A partir del coneixement de la posició dels diferents "c-Tags" es poden deduir els genotips dels SNPs (Lindroos *et al.*, 2002). Els ddNTPs poden estar marcats amb el mateix fluorocrom, però això implica haver de realitzar reaccions separades; per això és més econòmic i eficient utilitzar ddNTPs marcats amb quatre fluorocroms diferents, que poden ser analitzats a la mateixa reacció.

3. Duplicació i evolució dels genomes

3.1. Evolució: conceptes i mecanismes

L'evolució es pot definir com el canvi en el patrimoni genètic d'una població (o espècie) en el temps. Així, són les poblacions o les espècies les que evolucionen, no els organismes individualment, ja que un organisme té el mateix patrimoni de gens al llarg de tota l'existència. Cal no confondre evolució amb progrés: les poblacions s'adapten a l'ambient que hi ha en un lloc i temps determinats, i el que en un moment pot ser beneficiós pot deixar de ser-ho en el moment següent. Per poder actuar, l'evolució requereix la presència de variació genètica inicial, és a dir, és necessari que existeixin mecanismes que incrementin o generin variació genètica i mecanismes que la facin disminuir. Les mutacions són la font principal de nova variació genètica i la recombinació també afavoreix la creació de variació. La selecció natural, en canvi, actua sobre aquesta diversitat ja existent. Altres mecanismes que també operen sobre la variació genètica són la deriva genètica i el flux gènic.

La mutació es produeix quan la maquinària cel·lular que copia el DNA comet errors que alteren la seqüència genètica i que no són eliminats pels sistemes de reparació. Hi ha molts tipus diferents de mutació: mutacions puntuals, en què un nucleòtid és substituït per un altre (veure apartat 2.4.3); canvis en la llargada, per inserció o deleció de nucleòtids (veure apartat 2.4.3); i inversions o duplicacions de gens o regions genòmiques (veure apartat següent). La majoria de mutacions són neutres pel que fa a l'eficàcia biològica ja que només una petita part dels genomes eucariotes contenen segments codificadors i, tot i que algunes regions no codificadores són importants en regulació gènica o altres funcions cel·lulars, el més probable és que la majoria de canvis de base no tinguin cap efecte. D'altra banda, la majoria de mutacions que sí que comporten un canvi en l'eficàcia biològica solen ser deletèries, ja que les substitucions aminoacídiques poden fer variar l'estructura proteica i això pot comportar que la proteïna canviï o perdi la seva funció. Per tant, només un percentatge molt petit de les mutacions són beneficioses. Per tal que es produeixi la substitució d'un al·lel per un altre en una població cal que es doni un procés de dos passos: primer una mutació apareix en un individu, i es crea un nou al·lel, i després cal que la freqüència d'aquest al·lel augmenti fins arribar a la fixació a la població. La majoria de mutants, però, es perden, fins i tot els beneficiosos, ja que comencen a freqüències molt baixes (apareixen en un individu dins de tota una població).

La recombinació és el procés d'intercanvi genètic que es produeix entre seqüències homòlogues de DNA, amb total reciprocitat, sense que cap cromosoma guanyi o perdi material genètic. Un dels tipus més importants de recombinació és l'intercanvi de segments entre cromosomes homòlegs quan es troben aparellats durant el procés de meiosi. Aquest mecanisme causa redistribució massiva de gens materns i paterns, de manera que es creen noves combinacions al·lèliques sobre les que l'evolució pot actuar a la descendència.

La selecció natural, únic mecanisme d'evolució adaptativa, consisteix en un èxit reproductiu diferenciat de variants genètiques preexistents en el patrimoni genètic. Segons com actuï, la selecció natural pot fer incrementar o disminuir la variació genètica. El més comú és que la selecció elimini variants deletèries o parcialment deletèries (selecció negativa o purificadora), tot i que també pot succeir que una variant sigui beneficiosa i la selecció l'arrossegui fins a fixar-la (selecció positiva o direccional). En qualsevol d'aquests dos casos, el que es dona és una disminució de la variació genètica. No obstant, existeixen altres models de selecció que poden contribuir al manteniment de la variació genètica, com la selecció balancejadora, en què els individus heterozigots tenen avantatge sobre qualsevol dels dos homozigots. La selecció natural, doncs, actua sobre els canvis que apareixen a l'atzar i afavoreix els canvis beneficiosos, però no els indueix. Quan la selecció es produeix sobre factors que contribueixen a l'èxit reproductiu d'un organisme, es parla de selecció sexual, que pot ser molt important en aquelles espècies que presenten una elevada competència entre mascles (veure apartat 5 del Resum global).

La deriva genètica és el canvi en les freqüències al·lèliques que es dona a l'atzar. En fer un mostreig dins una població, la freqüència estimada dels al·lèls pot diferir lleugerament de la freqüència real simplement per atzar i fa augmentar o disminuir la freqüència d'un al·lèl concret. En poblacions petites la variància de les freqüències al·lèliques al llarg de les generacions és més gran que en poblacions grans. Canvis bruscos en la grandària d'una població poden fer canviar de forma molt important les freqüències al·lèliques, de manera que els al·lèls que sobreviuen no són representatius del patrimoni genètic. Aquest procés s'anomena coll d'ampolla o també efecte fundador, ja que és el mateix procés que es dona quan una petita població envaeix un nou territori. De manera que la deriva genètica fa disminuir la variació genètica. D'altra banda, existeix el flux genètic, que és el fenomen que es produeix quan individus d'una població arriben a una altra per migració i s'aparellen amb individus d'aquesta població, de manera que poden fer canviar el

patrimoni gènic local amb la incorporació de nous al·lèls o amb canvis en les freqüències al·lèliques.

Els diferents processos evolutius, que permeten que els organismes es vagin adaptant als diversos canvis ambientals, també són el motor de l'aparició i desaparició de les espècies. El procés d'especiació és el conjunt de mecanismes evolutius que condueixen a la formació de noves espècies. L'especiació pot ser al·lopàtrica o simpàtrica, segons la distribució geogràfica de les poblacions que intervenen en el procés. El més comú és l'especiació al·lopàtrica, en què una població queda dividida en dues (o més) subdivisions aïllades geogràficament, de manera que els individus no poden superar la barrera i els seus patrimonis gènics començaran a evolucionar independentment, fins al punt que ja no podrien encreuar-se. L'especiació simpàtrica es produeix quan dues subpoblacions esdevenen aïllades reproductivament sense haver estat primer aïllades geogràficament, però es tracta d'un procés menys comú.

3.2. Evolució per duplicació

La importància de la duplicació gènica com a mecanisme que aporta el material genètic bàsic per a l'evolució biològica ha estat reconeguda des de la dècada de 1930. L'any 1932, Haldane va suggerir que els esdeveniments de duplicació podrien ser favorables perquè produeixen gens que es poden alterar sense comportar cap desavantatge per a l'organisme, de manera que els organismes amb múltiples còpies de gens serien menys propensos a patir mutacions perjudicials (Haldane, 1932). El 1936, Bridges va descriure una de les primeres observacions de duplicació gènica lligades a variació morfològica i va proposar que la duplicació pot conduir a especiació (Bridges, 1936). El 1938, Serebrovsky va proposar que la selecció purificadora podria estar relaxada en aquells gens que es troben duplicats i també va parlar de la possibilitat que les dues còpies del gen estiguessin modificades (Serebrovsky, 1938). A partir d'aquí, el possible rol de la duplicació gènica en evolució es va continuar suggerint i es van anar proposant diversos escenaris possibles. El 1970 Ohno va afirmar que la duplicació gènica era el factor més important de l'evolució sense el qual la creació dels metazous, vertebrats i mamífers no hauria estat possible, ja que salts tan grans en evolució requerien la creació de nous loci gènics sense funcions prèvies (Ohno, 1970).

No va ser fins a la dècada del 1990, quan es van determinar i analitzar moltes seqüències, que la prevalença i la importància de les duplicacions gèniques va quedar clarament

demostrada. Les dades genòmiques que s'han anat obtenint recentment aporten evidències sòlides de l'abundància de gens duplicats en tots els organismes analitzats i indiquen que, en els tres dominis de la vida, una gran proporció dels gens van ser generats per duplicació gènica. Lynch (2003) van estimar que les duplicacions gèniques es generen (i es fixen) a una taxa aproximada de 10^{-2} a 10^{-3} per gen per milió d'anys en eucariotes com l'home, el ratolí, la drosòfila, l'arabidopsis o el llevat. Això vol dir que en una escala temporal de 100 milions d'anys s'espera que tots els gens d'un genoma s'hagin duplicat almenys una vegada. Aquesta taxa és més elevada que la taxa de substitució per posició nucleotídica, que és de 10^{-4} a 10^{-6} per gen per milió d'anys en genomes nuclears de vertebrats (Li, 1997). Els gens duplicats formen famílies gèniques de gens paràlegs i, si deixen de ser funcionals, es converteixen en pseudogens. En un inici, es pensava que les famílies multigèniques evolucionaven seguint el model d'evolució concertada, és a dir, tots els membres de la família evolucionaven com una única unitat, gràcies a entrecreuaments desiguals (Smith, 1976) o per l'acció de la conversió gènica (Jeffreys, 1979), com és el cas de la família dels gens d'RNA ribosomal (revisat a Nei, Rooney, 2005). Tanmateix, les recents evidències genòmiques indiquen que la majoria de famílies evolucionen per un procés de naixement i mort dels gens duplicats (Nei, Rooney, 2005). Aquest model assumeix que es creen nous gens per duplicació gènica i que alguns dels gens duplicats es mantenen al genoma durant molt de temps gràcies a l'adquisició de noves funcions, mentre que altres gens són suprimits o esdevenen no funcionals per l'acumulació de mutacions deletèries. S'han observat diferències en la grandària de les famílies gèniques, de manera que sembla que les proteïnes ribosomals i els factors de transcripció formen famílies més petites que altres proteïnes, com les que controlen el cicle cel·lular i el metabolisme (Conant, Wagner, 2002).

Generació de gens duplicats

La duplicació gènica pot sorgir d'entrecreuament desigual, retroposició o duplicació cromosòmica (o genòmica), que donen resultats molt diferents.

- 1) Els entrecreuaments desiguals acostumen a generar duplicacions gèniques en tàndem, de manera que els gens duplicats queden un al costat de l'altre dins el cromosoma. Segons on es doni l'entrecreuament, la regió duplicada pot contenir part d'un gen, un gen sencer o diversos gens (figura 12).

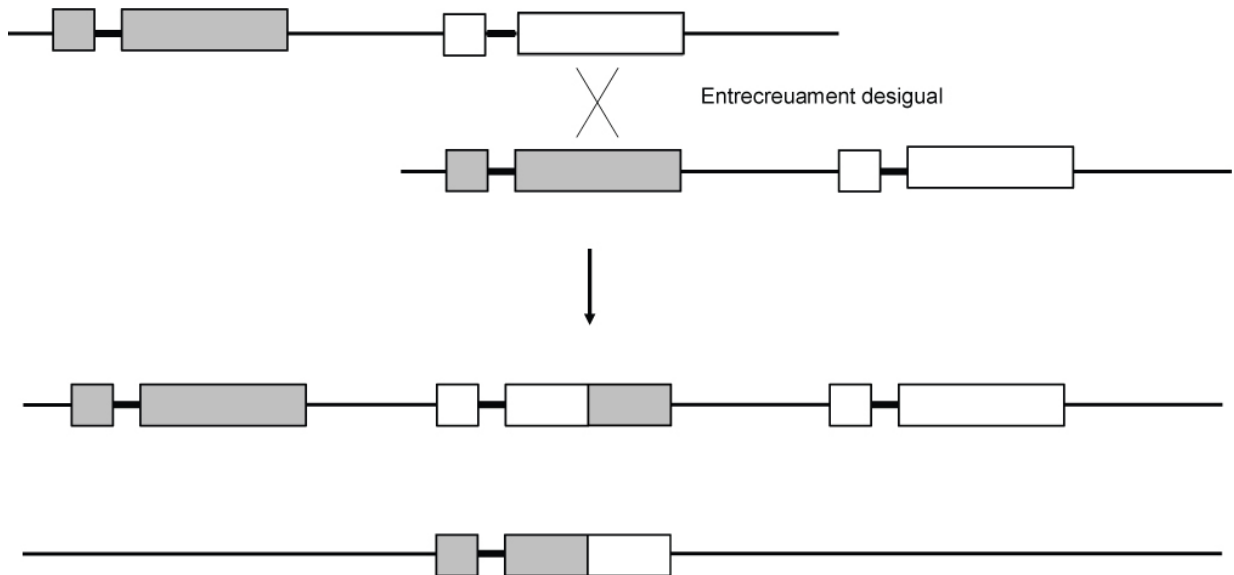


Figura 12. Esquema de l'entrecreuament desigual (adaptat de Zhang, 2003). Els rectangles blancs i grisos representen els exons i els quadrats negres, els introns.

- 2) En la retroposició, un RNA missatger es retrotranscriu a DNA complementari (cDNA) que s'insereix en el genoma. Aquest procés dona determinades característiques al nou gen: pèrdua dels introns i de les regions reguladores, presència de l'extensió poli-A i presència de repeticions directes curtes flanquejants, tot i que poden haver-hi desviacions d'aquest patró general (Long, 2001). El gen duplicat generat per retroposició no sol estar lligat al gen original sinó que la inserció del cDNA es dona a l'atzar dins del genoma. Com que el cDNA inserit no presenta la regió promotora ni les seqüències reguladores originals, la majoria de vegades no pot ser transcrit i ràpidament es converteix en un pseudogèn, a menys que per atzar la inserció s'hagi produït a continuació d'un promotor (Long, 2001) (figura 13).

- 3) Finalment, la duplicació cromosòmica o genòmica es produeix segurament per una manca de disjunció entre cromàtides germanes després de la replicació. Aquest mecanisme és freqüent en plantes però molt poc freqüent en animals (Li, 1997). Recentment, però, les anàlisis del genoma humà han revelat un altre tipus de duplicació a gran escala, la duplicació segmentària, que sovint implica de 1.000 a més de 200.000 nucleòtids (Samonte, Eichler, 2002). Encara no se sap quin mecanisme genera les duplicacions segmentàries, tot i que el fet que no estiguin en tàndem sembla excloure l'entrecreuament desigual com a responsable.

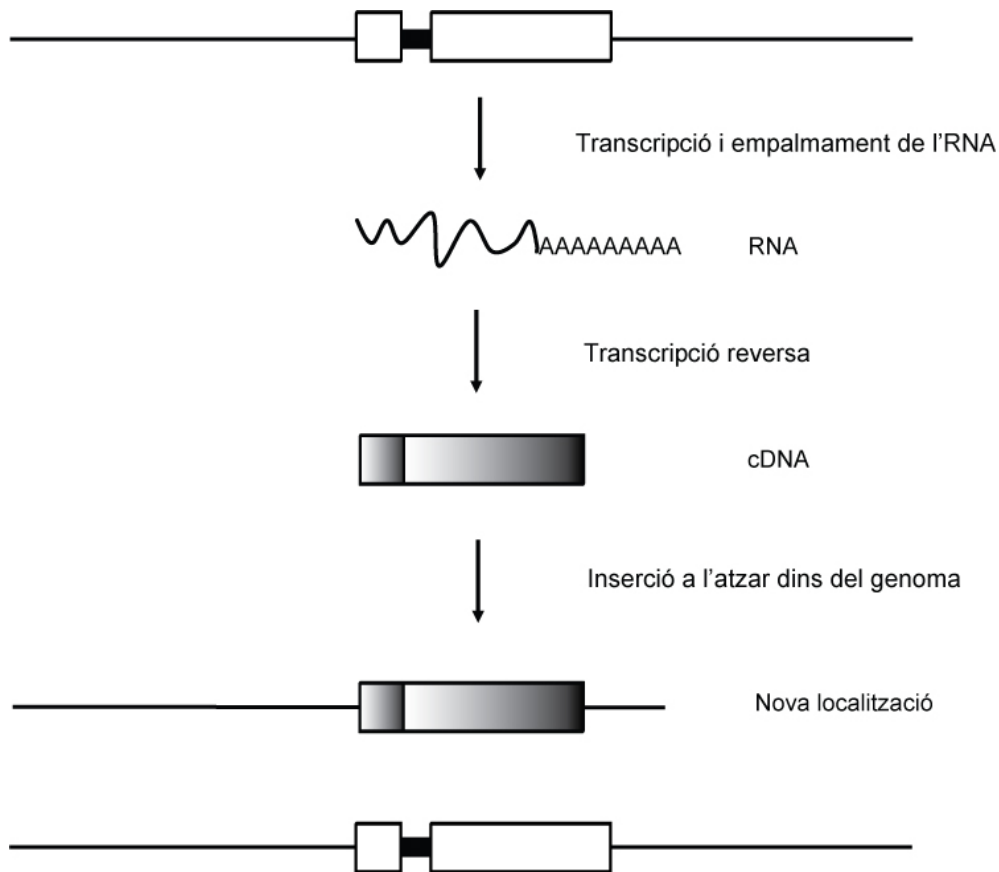


Figura 13. Esquema del procés de retroposició (adaptat de Zhang, 2003). Els rectangles blancs i grisos representen els exons i els quadrats negres, els introns.

3.3. Destí dels gens duplicats

Les duplicacions es donen en un individu i poden ser fixades o perdudes a la població d'una manera semblant a les mutacions puntuals. Si la presència de la duplicació és selectivament neutra, la probabilitat de fixació a la població és molt baixa ($1/2N$ en poblacions diploides, on N és la grandària poblacional), de manera que la majoria de gens duplicats es perdran. D'altra banda el destí evolutiu d'una duplicació a llarg terme serà determinat per les funcions dels gens duplicats. Es pot distingir entre quatre processos evolutius que determinen el destí dels gens duplicats:

- Pseudogenització

La duplicació gènica genera redundància funcional, de manera que mutacions que afectin l'estructura i la funció d'una de les dues còpies no seran deletèries perquè encara queda l'altre gen funcional i, per tant, la selecció no les eliminarà. L'acumulació de mutacions portarà una de les còpies del gen a patir un procés de pseudogenització fins a convertir-se en un pseudogèn, que no s'expressa o bé que dóna productes no funcionals. A la llarga,

els pseudogens són eliminats del genoma o bé esdevenen tan diferents als gens ancestrals que ja no són identificables, de manera que només els pseudogens relativament joves es poden reconèixer per similitud de seqüència. Quan el gen duplicat no està sota cap tipus de selecció, la pseudogenització té lloc durant els primers milions d'anys després de la duplicació i la vida mitja és d'uns 5 milions d'anys (Lynch, 2003). Tot i així, alguns gens duplicats s'han mantingut en el genoma durant molt de temps per alguna funció específica i només recentment s'han convertit en pseudogens per una relaxació de les constriccions funcionals. Un bon exemple és la família de gens del receptor olfactiu, que tant en humans com en ratolí està formada per uns 1.000 membres, però, mentre el percentatge de pseudogens en ratolins és només d'un 20%, en humans és del 60%, de manera que la majoria de pseudogens s'han originat a partir de la divergència dels primats, probablement perquè el desenvolupament d'una bona visió ha fet disminuir les constriccions funcionals sobre els gens relacionats amb l'olfacte (Rouquier *et al.*, 2000).

- Conservació de la funció

En alguns casos es mantenen les dues còpies d'un gen duplicat simplement perquè l'augment de quantitat de proteïna o de productes d'RNA que proporcionen suposa un avantatge. Això es pot aplicar principalment a gens que s'expressen de manera activa perquè els seus productes tenen una gran demanda, com els rRNAs i les histones. Existeixen dos mecanismes possibles que expliquen que dos gens paràlegs puguin mantenir-se amb la mateixa funció. Per una banda, una conversió gènica freqüent pot fer que dos gens paràlegs mantinguin les seqüències i les funcions sempre molt semblants. Aquest procés evolutiu es pot anomenar evolució concertada (Li, 1997). Per una altra banda, una selecció purificadora molt forta contra les mutacions que modifiquen la funció també pot evitar que els gens duplicats divergeixin. Alguns estudis suggereixen que l'acció de la selecció purificadora és molt més comú que la conversió gènica (Nei *et al.*, 2000; Piontkivska *et al.*, 2002) i que les condicions perquè la conversió gènica sigui afavorida selectivament són molt restrictives (Hurst, Smith, 1998).

- Subfuncionalització

En general, si la presència d'una quantitat extra de producte gènic no és avantatjosa, la redundància gènica no té sentit selectiu i no és probable que dos gens amb idèntica funció es mantinguin en el genoma de manera estable (Nowak *et al.*, 1997). S'ha vist, però, que

tots dos gens poden ser mantinguts si difereixen en algun aspecte de la seva funció, cosa que és possible a través d'un procés de subfuncionalització, en què cadascun dels gens duplicats adopta una part diferent de les funcions de l'únic gen ancestral i, per tant, és necessària l'acció combinada de les dues còpies per acomplir l'operativitat original. Force et al. (1999) van explicar la subfuncionalització a través del model de duplicació-degeneració-complementació (DDC), en què canvis deleteris complementaris en els gens duplicats comporten que junts puguin mantenir les funcions originals.

La subfuncionalització pot tenir lloc de diferents maneres. Una forma de subfuncionalització molt important és la diferenciació en el patró d'expressió gènica després de la duplicació, que s'ha pogut demostrar en diferents estudis. Per exemple, els gens *engrailed-1* i *engrailed-1b* en el peix zebra provenen d'una duplicació que tingué lloc al llinatge dels peixos amb aleta de rajada (Actinopterygii) i cadascun s'expressa en un teixit diferent, mentre que en altres espècies, com el ratolí, existeix un sol gen d'expressió ubiqua (Force et al., 1999). La variació en l'expressió gènica després de la duplicació sembla que és la norma general, més que una excepció, i acostuma a donar-se immediatament després de la duplicació (Gu et al., 2002). Aquests canvis poden afectar tant a l'especificitat de teixit com al moment del desenvolupament en què es dona l'expressió.

D'altra banda, la subfuncionalització també pot produir-se a nivell de la funció proteica i conduir a una especialització funcional quan una de les còpies del gen s'especialitza en una de les funcions originals del gen ancestral (Hughes, 1999). Un bon exemple és el de les RNases digestives del rinopitec d'Indoxina (*Pygathrix nemaeus*). En aquesta espècie s'ha detectat la presència d'un gen que codifica per un enzim digestiu especialitzat, *RNase1B*, que prové de la duplicació d'un gen bifuncional, *RNase1* (Zhang et al., 2002). En general, però, no està clar quin percentatge de gens duplicats evolucionen per subfuncionalització. Fins i tot hi ha autors que defensen que la subfuncionalització és un estat transitori entre el procés de duplicació i la neofuncionalització i no el destí final dels gens duplicats (Rastogi, Liberles, 2005).

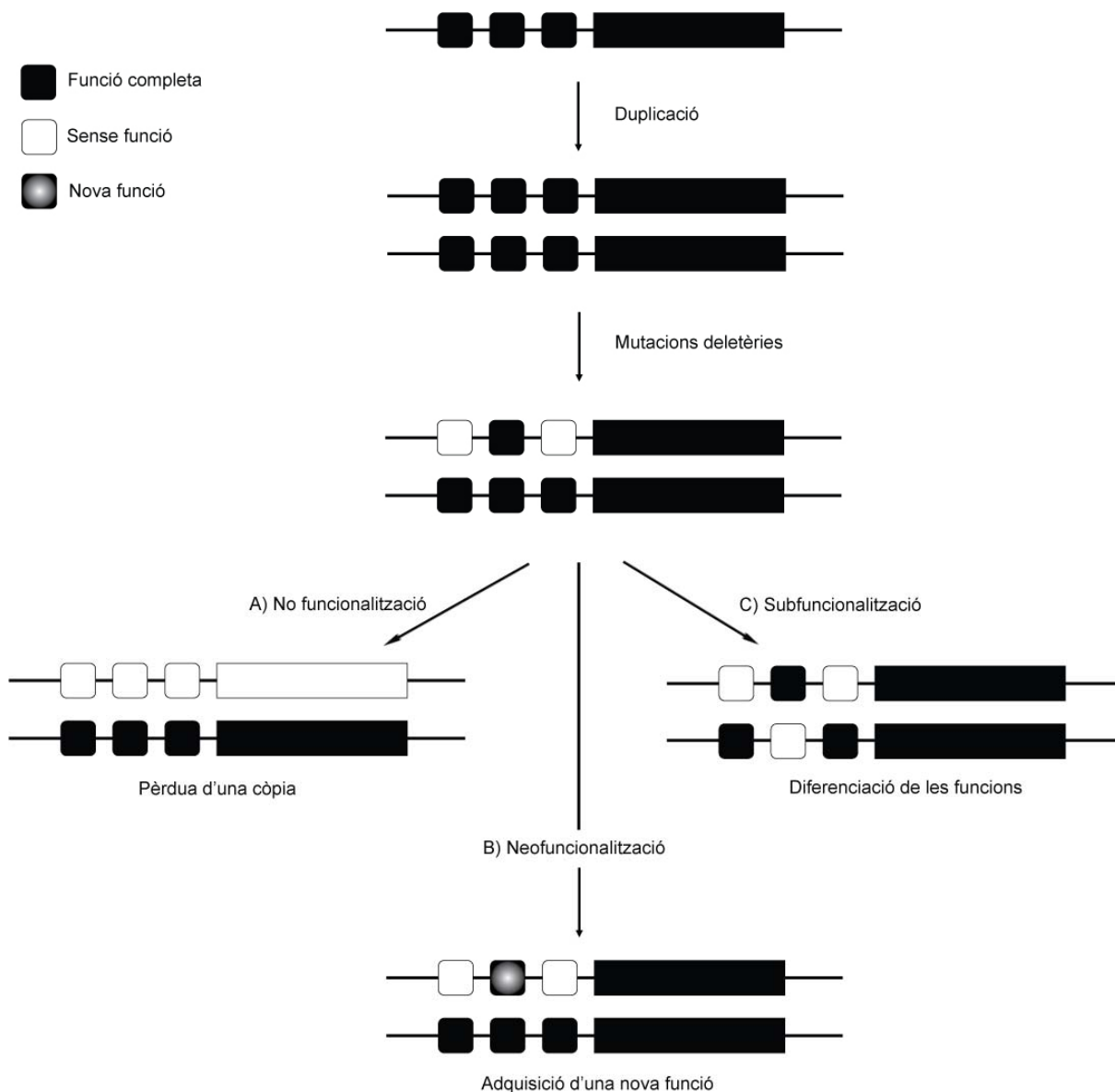


Figura 14. Possibles destins dels gens duplicats (adaptat de Force *et al.*, 1999). A) No funcionalització (o pseudogenització). B) Neofuncionalització. C) Subfuncionalització.

- Neofuncionalització

Finalment, un dels resultats més importants que pot donar la duplicació gènica és l'origen d'una nova funció: una de les còpies del gen muta i adopta una funció que no estava present en el gen ancestral, mentre que l'altra còpia manté la funció inicial, de manera que totes dues còpies perduren. El més probable és que la nova funció tingui certa relació amb la funció original i no es tracti d'una funció completament nova. Aquest és el cas dels gens per a les opsines sensibles al vermell i al verd dels humans, que després de la duplicació van divergir en la longitud d'ona màxima absorbida, de manera que permeten un rang de visió dels colors més ampli en els humans (Yokoyama, Yokoyama, 1989). Tanmateix, en alguns casos s'ha pogut descriure el sorgiment d'una funció completament

diferent a la del gen ancestral. Un bon exemple és el dels gens *EDN* i *ECP* presents als catarrins i pertanyents a la família de les RNases A, ja que la proteïna *ECP* ha adquirit una activitat antibacteriana, totalment independent a l'activitat ribonucleasa (Zhang *et al.*, 1998). La neofuncionalització de gens duplicats requereix un cert número de substitucions aminoacídiques. Quan el número de canvis genètics necessaris és elevat, encara no s'ha resolt quina funció adopta la proteïna durant els múltiples passos intermedis de canvi genètic.

3.4. Forces evolutives rere la divergència funcional dels gens duplicats

En l'evolució dels gens duplicats poden actuar diferents forces evolutives que determinaran el resultat de la duplicació. En el cas de la diferenciació en l'expressió gènica per subfuncionalització sembla que el més plausible és que s'hagin fixat mutacions deletèries complementàries a l'atzar sota unes condicions de relaxació de les constriccions funcionals que fan que la selecció purificadora no actuï, de manera que la diferenciació es dona sense l'acció de la selecció positiva. Pel que fa a l'especialització funcional i a la neofuncionalització, s'han descrit dos models diferents per explicar l'evolució dels gens duplicats (Zhang, 2003). El primer model es coneix com a l'efecte Dykhuizen-Hartl i no requereix l'acció de la selecció positiva, sinó que després de la duplicació la redundància gènica comporta unes condicions de reducció de les constriccions funcionals i de relaxació de la selecció purificadora que afavoreixen la fixació de mutacions a l'atzar en una de les còpies gèniques. Més endavant, aquestes mutacions fixades induiran un canvi en la funció gènica quan, per exemple es dona una alteració en l'ambient o en el fons genètic.

El segon model requereix l'acció de la selecció positiva. Hi ha dos escenaris possibles: i) després de la duplicació algunes substitucions neutres o quasi neutres creen una nova funció, encara poc activa, en una de les còpies gèniques i llavors la selecció positiva accelera la fixació de mutacions avantatjoses que potencien la nova funció gènica; o ii) el gen ancestral ja tenia dualitat funcional i la duplicació dona l'oportunitat que cadascuna de les còpies adquireixi una de les funcions originals i després noves substitucions seleccionades positivament poden millorar aquestes funcions.

Després d'un esdeveniment de duplicació, sovint s'observa una acceleració en l'evolució proteica, que pot ser explicada per qualsevol dels dos models evolutius. En general, en estudiar l'acció de les forces evolutives sempre s'assumeix com a hipòtesi nul·la l'evolució

neutra sota selecció purificadora relaxada. En el cas que s'observi una taxa de substitució no sinònima significativament més elevada que la taxa de substitució sinònima, la hipòtesi nul·la és rebutjada i s'accepta l'acció de la selecció positiva (veure apartat següent). En cas d'acceptar la hipòtesi nul·la, però, sempre s'infereix relaxació de la selecció purificadora, tot i que no hi hagi evidències. La baixa potència estadística dels mètodes de detecció de la selecció provoca que moltes vegades no sigui possible detectar l'acció de la selecció positiva i s'accepti incorrectament la relaxació de la selecció purificadora. De fet, s'ha determinat que perquè es doni la divergència funcional dels gens duplicats són necessàries tant la selecció positiva com la relaxació de la selecció purificadora (Zhang *et al.*, 2002).

3.5. Detecció de selecció positiva

La detecció de la selecció darwiniana positiva a nivell de seqüència de DNA és una qüestió que ha acaparat molta atenció dins de la biologia evolutiva. Tot i l'esforç invertit, la selecció positiva és difícil de detectar perquè sovint actua de manera esporàdica sobre poques posicions aminoacídiques i el senyal pot veure's emmascarat per la selecció negativa. La selecció pot ser analitzada a través d'estimacions de la proporció (ω) entre la taxa de substitució nucleotídica no sinònima (d_N) i la taxa de substitució nucleotídica sinònima (d_S) de seqüències gèniques codificadores homòlogues (Nei, 2000). Aplicant aquest test, una ω significativament més gran que 1 s'interpreta com una evidència de selecció positiva, una ω menor que 1 suggereix selecció purificadora i si ω és igual a 1 indica selecció neutra. Moltes vegades, però, aquest mètode no té prou potència estadística i la selecció positiva no es pot posar de manifest o bé apareix el problema contrari, l'obtenció de falsos positius.

En els últims anys s'han anat desenvolupant diferents mètodes per testar la selecció positiva d'una manera més específica. Alguns mètodes, els anomenats mètodes de branca, busquen la selecció en branques individuals, o en un set de branques, d'un arbre filogenètic (Messler, Stewart, 1997; Yang *et al.*, 1998; Yu, Irwin, 1996; Zhang *et al.*, 1997; Zhang *et al.*, 1998). D'altra banda, existeixen els mètodes de posició, que testen la selecció en posicions individuals dins dels codons (Nielsen, Yang, 1998; Suzuki, Gojobori, 1999; Yang *et al.*, 2000). Recentment, Yang & Nielsen (2002) van introduir el mètode de branca-posició per testar la selecció positiva en codons individuals al llarg dels llinatges i van dissenyar els models A i B. En aquests models, les branques de l'arbre són dividides *a priori* en llinatges *foreground* i *background* i es construeix un test de raó de

versemblança (LRT) comparant un model que permeti selecció positiva als llinatges *foreground* amb un model que no permeti aquesta variació. Simulacions per ordinador, però, van mostrar que aquests tests portaven freqüentment a descartar la hipòtesi nul·la per error i produïen falsos positius (Zhang, 2004). Més recentment, Zhang et al. (2005) han descrit una modificació del mètode de branca-posició proposat anteriorment i l'han utilitzat per construir dos nous LRTs, anomenats test 1 i test 2. Aquests autors han introduït variacions al model A de Yang & Nielsen (2002), de manera que el nou model A assumeix quatre classes de posicions possibles:

- la classe 0 inclou els codons conservats al llarg de l'arbre, amb una $0 < \omega_0 < 1$ estimada (l'antic model fixava $\omega_0 = 0$, de manera que no permetia l'existència de llocs conservats amb $0 < \omega_0 < 1$);
- la classe 1 inclou els codons que evolucionen de manera neutra al llarg de l'arbre amb una $\omega_1 = 1$;
- les classes 2a i 2b inclouen els codons que són conservats o neutres a les branques *background* però que estan sota selecció positiva a les branques *foreground*, amb una $\omega_2 > 1$ estimada a partir de les dades.

Aquest model A es pot utilitzar per construir dos LRTs:

- Test 1: El model A és la hipòtesi alternativa i la hipòtesi nul·la és el model M1a (gairebé neutre) (Yang *et al.*, 2000), que considera dues classes de posicions, amb $0 < \omega_0 < 1$ i amb $\omega_0 = 0$ per a totes les branques.
- Test 2: El model A també és la hipòtesi alternativa i com a hipòtesi nul·la s'utilitza un model A amb la ω_2 fixada a 1. Aquest model nul permet que les posicions que evolucionen sota selecció negativa als llinatges *background* siguin alliberades de les constriccions i puguin evolucionar de manera neutra als llinatges *foreground*. De manera que aquest és un test directe de selecció positiva als llinatges *foreground*. Simulacions per ordinador suggereixen que el test 2 pot distingir de manera precisa entre relaxació de les constriccions funcionals i selecció positiva, és més astringent, i sovint té més potència que el test 1 per detectar selecció positiva (Zhang *et al.*, 2005).

Dins del mètode de branca-posició, quan es rebutja el model nul, el model alternatiu (model A) permet identificar les posicions que estan sota selecció positiva al llarg del

l·linatge *foreground*. El mètode empíric de Bayes pot ser utilitzat per calcular les probabilitats posteriors del fet que cadascuna de les posicions pertanyin a la classe de posicions seleccionades positivament als l·linatges *foreground*. El mètode “naïve empirical Bayes” (NEB) (Yang, Nielsen, 2002), que utilitza les estimacions de màxima versemblança (MLE) dels paràmetres però ignora els errors de mostreig, dóna probabilitats menys ajustades que el mètode Bayes empirical Bayes (BEB) (Yang et al. 2005), que sí que té en compte de manera adequada les incerteses dins de les MLE i, doncs, dóna estimacions més precises.

PAML

El PAML, *Phylogenetic Analysis by Maximum Likelihood* (Yang, 1997), és un paquet de programes que permet dur a terme anàlisis filogenètiques de seqüències de DNA i de proteïna utilitzant el mètode de la màxima versemblança. Entre els programes que inclou el PAML es troba el *codeml*, que permet realitzar, entre altres, les anàlisis següents:

- Tests de màxima versemblança (LRTs) de diferents hipòtesis a través de comparar diferents models que té implementats.
- Reconstrucció per versemblança (Empirical Bayes) de les seqüències ancestrals utilitzant models nucleotídics, aminoacídics o de codons.
- Estimacions de taxes de substitució sinònimes i no sinònimes i detecció de selecció positiva en seqüències de DNA codificadores.

El gran avantatge del PAML és que ofereix una gran col·lecció de mètodes de substitució sofisticats que van sent actualitzats a mida que es milloren els models existents o quan en sorgeixen de nous. Així, a partir d'un conjunt de dades de seqüències es poden realitzar les anàlisis de màxima versemblança pels diferents models i aplicar els diferents tests, d'una manera ràpida i eficient. Gràcies a la seva potència, actualment el PAML és un dels programes més utilitzats en els estudis filogenètics i les anàlisis de selecció natural.

4. Cromosoma Y

4.1. Cromosomes sexuals en mamífers: Origen i evolució

Els cromosomes sexuals dels mamífers provenen d'un parell d'autosomes (Ohno, 1967) que fa uns 300 milions d'anys, després de la divergència dels mamífers, van començar a diferenciar-se. Després d'adquirir el gen determinant del sexe, el proto-Y va quedar aïllat genèticament de la resta del genoma per l'eliminació progressiva de la recombinació amb el proto-X. A mida que es produïa la supressió de la recombinació, el proto-Y patia una degradació paulatina, que actualment es pot detectar fins i tot a nivell citogenètic (per exemple, el cromosoma Y humà és tres vegades més petit que el cromosoma X i té una composició de la cromatina diferenciada). La pèrdua de gens en el proto-Y i el seu manteniment en el proto-X ha comportat un desequilibri genètic entre mascles i femelles, de manera que l'evolució ha desenvolupat mecanismes de compensació de dosi per restablir l'equivalència en l'expressió gènica dels gens del cromosoma X entre els dos sexes. Per altra banda, la diferenciació dels cromosomes sexuals també ha estat promoguda per l'acumulació de gens antagònics, beneficiosos per a un sexe i perjudicials per a l'altre. A la figura 15 es mostra un esquema de l'evolució dels cromosomes sexuals.

Els cromosomes X i Y són cromosomes molt atípics i molt diferenciats entre ells. En els mamífers placentaris, el cromosoma X té el mateix aspecte que els autosomes, és un cromosoma de mida mitjana (en els humans és dels més llargs, d'unes 165 Mb) i conté una proporció de gens estàndard. Sorprenentment, però, alguns treballs recents (p.e. Hurst, 2001) indiquen que el contingut gènic del cromosoma X no és homogeni sinó que hi ha certes classes de gens sobrerrepresentades: gens implicats en característiques sexuals i de reproducció, gens relacionats amb desenvolupament cerebral i intel·ligència i gens relacionats amb musculatura i esquelet. Per altra banda, el cromosoma Y és un cromosoma petit, parcialment heterocromàtic i amb pocs gens funcionals, que han evolucionat cap a funcions especialitzades en reproducció i caràcters sexuals masculins. A més, no recombinava gairebé en la seva totalitat. El cromosoma Y conté dos dominis diferents. Per una banda, presenta una petita regió pseudoautosòmica (PAR, situada a un o als dos extrems del cromosoma), que sí que s'aparella i recombinava amb el cromosoma X durant la meiosi. En els humans, per exemple, la regió PAR se situa als dos extrems del cromosoma Y i equival només a un 5% del total del cromosoma. Per altra banda, existeix una gran regió no-recombinant (NRY), que no recombinava amb l'X i que segueix una herència patrilineal. Actualment, però, es tendeix a anomenar aquesta part del

cromosoma Y “regió específica masculina” (MSY) pel fet que existeixen fenòmens de recombinació intracromosòmica.

Les característiques pròpies del cromosoma Y que el diferencien de la resta de cromosomes dels mamífers –la no recombinació en gairebé tota la seva llargada i el fet que només està present en el sexe masculí– probablement han tingut una gran influència en l'evolució de les funcions genètiques i del contingut genètic d'aquest cromosoma. Hi ha dues teories principals que expliquen l'evolució dels gens del cromosoma Y. La primera teoria es basa en el fet que l'absència de recombinació no permet la separació de les mutacions deletèries de les avantatjoses, de manera que condueix inevitablement cap a un deteriorament del contingut genètic de les regions no recombinants del cromosoma Y (Muller, 1964). L'altra teoria postula que l'especificitat masculina del cromosoma Y pot afavorir l'acumulació d'al·lels que potencien els mascles i deterioren les femelles (els al·lels antagònics), de manera que el cromosoma Y esdevindria un cromosoma especialitzat masculí (Fisher, 1931; Rice, 1987a). Aquestes dues teories no s'exclouen entre elles però són conceptualment contradictòries, ja que la primera defensa una pèrdua d'informació genètica mentre que la segona proposa un guany i conservació de la informació. És molt possible que el contingut genètic del cromosoma Y hagi estat modulats per tots dos fenòmens (veure l'apartat 5 del Resum global).

Lahn & Page (1999) van estudiar diferents parells de gens X-Y humans i els van classificar en quatre grans grups, quatre estrats del cromosoma X. Les seves observacions els van portar a proposar que l'evolució dels cromosomes sexuals humans han patit com a mínim quatre grans esdeveniments, cadascun dels quals eliminava la recombinació en un dels estrats, sense afectar l'ordre dels gens al cromosoma X. La diferenciació entre els gens dels cromosomes X i Y començaria després de la desaparició de la recombinació, probablement per inversions al cromosoma Y, de manera que s'hauria produït en un moment diferent per a cada estrat. El primer esdeveniment hauria estat contemporani amb el naixement dels cromosomes sexuals, just després de la divergència dels mamífers, i l'últim esdeveniment hauria succeït dins l'evolució dels primats.

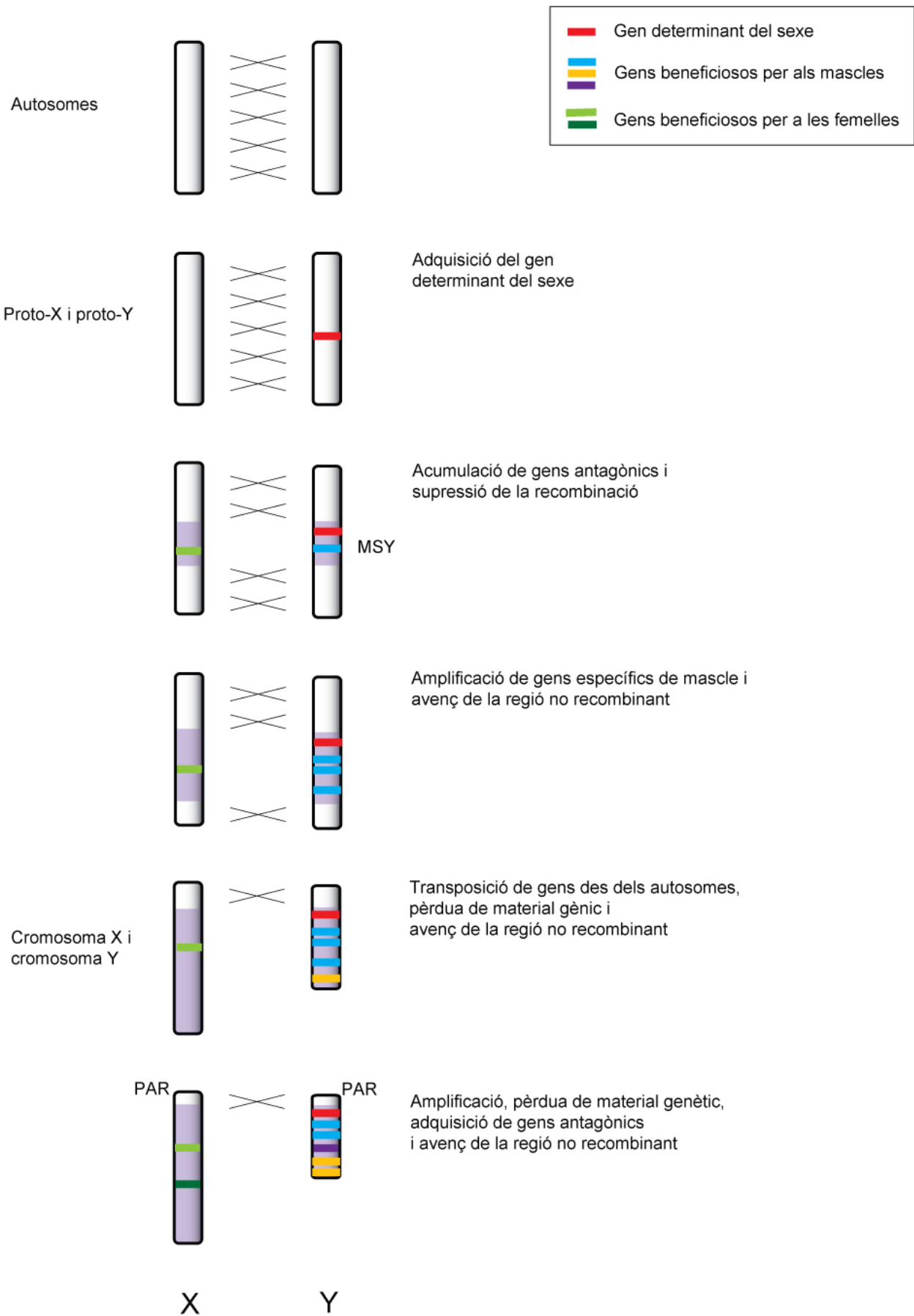


Figura 15. Esquema dels processos que han conduït a l'evolució dels cromosomes sexuals en mamífers (adaptat de Bachtrog, 2006).

Aquesta teoria ha estat reforçada per l'estudi realitzat per Skaletsky et al. (2003) a partir de la seqüenciació del cromosoma Y humà. Aquests autors defineixen tres grans classes de seqüències presents a l'eucromatina de la regió específica de mascle del cromosoma Y:

- i) seqüències transposades des de l'X, que són idèntiques en un 99% a seqüències de DNA de la regió Xq21 i que provenen d'una transposició de material del cromosoma X a l'Y que tingué lloc fa 3-4 milions d'anys a la branca de l'home; només conté dos gens (*TGIF2LY* i *PCDH11Y*);
- ii) segments degenerats de X, que contenen 27 gens o pseudogens de còpia única homòlegs a diferents gens del cromosoma X, amb una identitat nucleotídica del 60 al 96 %, i que semblen relíquies dels antics autosomes dels quals provenen els cromosomes sexuals; la majoria d'aquests gens s'expressen de manera ubíqua;
- iii) segments amplicònics, formats per seqüències que tenen una elevada similitud –fins a un 99,9% d'identitat– amb altres seqüències de la regió MSY; les unitats de repetició específiques de MSY s'anomenen amplicons, es troben repartides en set segments al llarg de l'eucromatina i contenen moltes famílies gèniques; la característica estructural més aparent és que inclouen vuit grans palíndroms amb una gran identitat de seqüència gràcies a la conversió gènica (Rozen *et al.*, 2003); tots els gens d'aquest grup són d'expressió predominant o exclusiva de testicle.

La classe de seqüències X-degenerades dóna suport a la teoria de la decadència gènica en absència de recombinació sexual, ja que existeixen molts pseudogens en el cromosoma Y de gens encara funcionals en el cromosoma X. Es pot assumir que centenars d'altres gens homòlegs a gens de l'X van ser eliminats durant l'evolució de la regió MSY sense deixar rastre en les seqüències actuals. D'altra banda, però, els segments amplicònics s'expliquen per la teoria de l'adquisició i conservació de gens que afavoreixen la fertilitat masculina. Són seqüències que s'han format a partir d'orígens genòmics diversos i per mecanismes diferents, però totes han evolucionat de la mateixa manera, formant diverses còpies gairebé idèntiques i amb un patró d'expressió específic de cèl·lules espermatogèniques.

Taula 3. Classificació dels gens del cromosoma Y, número de còpies i especificitat d'expressió (adaptat de Skaletsky *et al.*, 2003).

Tipus de seqüència	Símbol del gen	Nom del gen	Número còpies	Localització	Expressió en teixit	Homòleg lligat a X	Homòleg autosòmic
Transposat a l'X	TGIF2LY	<i>TGFB-induced factor homeobox 2-like Y</i>	1	Yp11.2	Testicle	TGIF2LX	–
	PCDH11Y	<i>Protocadherin 11 Y</i>	1	Yp11.2	Cervell fetal, cervell	PCDH11X	–
Total			2				
X-degenerat	SRY	<i>Sex determining region Y</i>	1	Yp11.31	Predominant a testicle	SOX3	–
	RPS4Y1	<i>Ribosomal protein S4 Y 1</i>	1	Yp11.31	Ubiqua	RPS4X	–
	ZFY	<i>Zinc finger Y</i>	1	Yp11.31	Ubiqua	ZFX	–
	AMELY	<i>Amelogenin Y</i>	1	Yp11.2	Dents	AMELX	–
	TBL1Y	<i>Transducin (beta)-like 1 protein Y</i>	1	Yp11.2	Cervell fetal, pròstata	TBL1X	–
	PRKY	<i>Protein kinase Y</i>	1	Yp11.2	Ubiqua	PRKX	–
	USP9Y	<i>Ubiquitin-specific protease 9 Y</i>	1	Yq11.21	Ubiqua	USP9X	–
	DDX3Y	<i>Dead box Y</i>	1	Yq11.21	Ubiqua	DBX	–
	UTY	<i>Ubiquitous TPR motif Y</i>	1	Yq11.21	Ubiqua	UTX	–
	TMSB4Y	<i>Thymosin (beta)-4 Y</i>	1	Yq11.221	Ubiqua	TMSB4X	–
	NLGN4Y	<i>Neurologin 4 isoform Y</i>	1	Yq11.221	Cervell, cerv. fetal, pròstata, testicle	NLGN4X	–
	CYorf15A	<i>Chromosome Y open reading frame 15A</i>	1	Yq11.222	Ubiqua	CXorf15	–
	CYorf15B	<i>Chromosome Y open reading frame 15B</i>	1	Yq11.222	Ubiqua	CXorf15	–
	JARID1D	<i>Jumonji, AT rich interactive domain 1D</i>	1	Yq11.222	Ubiqua	JARID1C	–
	EIF1AY	<i>Translation initiation factor 1A Y</i>	1	Yq11.223	Ubiqua	EIF1AX	–
	RPS4Y2	<i>Ribosomal protein S4 Y 2</i>	1	Yq11.223	Ubiqua	RPS4X	–
Total			16				
Amplicònic	TSPY	<i>Testis-specific protein Y</i>	~35	Yp11.2	Testicle	–	–
	VCY	<i>Variable charge Y</i>	2	Yq11.221	Testicle	VCX	–
	XKRY	<i>XK related Y</i>	2	Yq11.221	Testicle	–	–
	CDY	<i>Chromodomain Y</i>	4	Yq11.22	Testicle	–	CDYL
	HSFY	<i>Heat shock transcription factor Y</i>	2	Yq11.221	Testicle	–	–
	RBMV	<i>RNA-binding motif Y</i>	6	Yq11.2	Testicle	RBMX	–
	PRY	<i>PTP-BL related Y</i>	2	Yq11.223	Testicle	–	–
	BPY2	<i>Basic charge Y</i>	3	Yq11.223	Testicle	–	–
	DAZ	<i>Deleted in azoospermia</i>	4	Yq11.2	Testicle	–	DAZL
Total			~60				
Total global			~78				

El fet que el cromosoma Y contingui una sèrie de famílies gèniques amb múltiples còpies actives ha de tenir un sentit biològic. S'ha proposat que l'amplificació es dona en resposta a una reducció de l'expressió, de manera que una baixada en l'activitat d'un gen potenciaria l'acció de la selecció (o com a mínim establiria condicions permissives) per a la seva amplificació (Graves *et al.*, 1998). Com s'ha vist (apartat 3.3), el fet de disposar de diverses còpies és un avantatge davant una possible mutació que generi inactivació parcial o total de la funció. D'altra banda, s'ha observat que tots els gens que es troben en múltiples còpies són d'expressió predominant o exclusiva de testicle, de manera que l'amplificació potser ve donada per una necessitat d'augment de l'expressió en aquest teixit. A més, les còpies funcionals d'aquest grup de gens es troben situades majoritàriament en palíndroms, de manera que sovint pateixen fenòmens de conversió gènica (Rozen *et al.*, 2003) i aquest mecanisme podria estar evitant la seva degeneració perquè eliminaria les possibles mutacions deletèries (Skaletsky *et al.*, 2003).

4.2. Cromosoma Y en primats: Evolució

La comparació entre diferents mamífers placentaris revela conjunts de gens del cromosoma Y que coincideixen, però també posa de manifest que alhora hi ha una gran variació interespecífica en l'ordre dels blocs. Estudis d'hibridació *in situ* amb fluorescència (FISH) per investigar l'evolució del cromosoma Y de primats van mostrar que aquest cromosoma ha patit una ràpida evolució sense restriccions pel que fa al contingut de seqüències i l'organització (Archidiacono *et al.*, 1998) (figura 16).

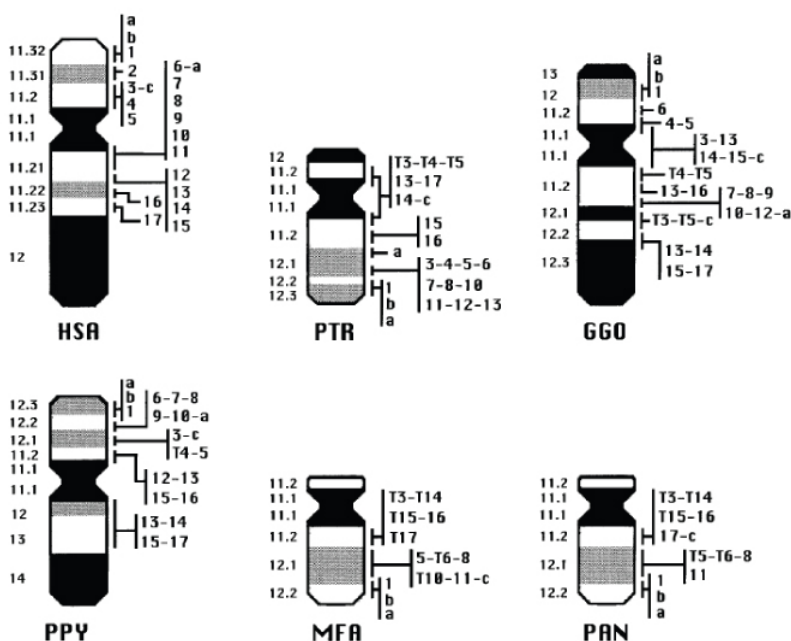


Figura 16. Esquema del cromosoma Y de diferents espècies de primat. Els números i lletres de la dreta indiquen l'ordre d'hibridació de les sondes, de manera que mostren les diferències quant a reorganitzacions cromosòmiques entre diferents espècies de primats. HSA, *Homo sapiens*; PTR, *Pan troglodytes*; GGO, *Gorilla gorilla*; PPYB, *Pongo pygmaeus pygmaeus*; MFA, *Macaca fascicularis*; PAN, *Papio anubis*. (Adaptat d'Archidiacono *et al.*, 1998)

Quan es comparen prometafases de cromosoma Y de diferents simis es posa de manifest d'una manera evident la gran variabilitat de mida que existeix. Els cromosomes Y d'humans i goril·les són els més grans i s'observen grans quantitats d'heterocromatina constitutiva en el seu braç llarg, mentre que el del tití *Callicebus moloch*, per exemple, és molt més petit i no presenta indicis d'heterocromatina constitutiva en estudiar-lo per FISH (Gläser *et al.*, 1998). Tot i així, tots els simis tenen en comú que el gen *SRY* es localitza molt a la prop de la regió pseudoautosòmica, a la frontera pseudoautosòmica (*pseudoautosomal boundary*, PABY). De manera que l'organització linear del gens de la regió PAR i el gen *SRY* està molt conservada i la frontera pseudoautosòmica és estable en tots els simis. En canvi, la part específica masculina del cromosoma Y ha estat més exposada a patir amplificacions, diversificacions i reorganitzacions específics d'espècie (Gläser *et al.*, 1998). La fixació ràpida de qualsevol d'aquestes variacions en l'evolució ha estat possible sempre que no interferís amb la fertilitat masculina. De fet, es pot considerar que les diversificacions específiques d'espècie en la seqüència de gens relacionats amb la fertilitat de la regió MSY podrien haver actuat com una barrera eficient contra la hibridació de manera que haurien promogut l'especiació (Gläser *et al.*, 1998). En els humans i grans simis s'ha vist que l'amplificació de gens i les reorganitzacions del cromosoma Y han estat processos molt importants durant l'evolució de la regió MSY. Aquests processos no només es detecten a nivell d'espècie sinó que també han fet evolucionar les subespècies, com és el cas de les dues subespècies d'orangutà, la de Borneo (*Pongo pygmaeus pygmaeus*) i la de Sumatra (*Pongo pygmaeus abelli*), que per les grans diferències que presenten es comencen a considerar com a espècies diferents (Xu, Arnason, 1996).

Recentment, s'ha pogut obtenir la seqüència d'aproximadament la meitat del cromosoma Y del ximpanzé, sense les regions repetitives (Hughes *et al.*, 2005; Kuroki *et al.*, 2006). Amb aquesta seqüenciació s'ha pogut determinar que la variació entre els cromosomes Y d'humà i ximpanzé és molt més elevada (1,78%) que la variació del genoma sencer (1,23%), cosa que suggereix una evolució accelerada del cromosoma Y respecte la resta del genoma (Kuroki *et al.*, 2006). S'han observat moltes diferències pel que fa als retroelements, com algunes insercions de retrovirus endògens específics de ximpanzé (CERV), i reorganitzacions estructurals a gran escala. A més, també s'han detectat diferències quant a inactivació de gens funcionals (veure apartat 5 del Resum global). Aquesta evolució ràpida del cromosoma Y pot ser deguda a una relaxació de les

restriccions selectives, a l'efecte de la selecció positiva o, més probablement, a l'acció dels dos factors.

4.3. Inactivació del cromosoma X – Deteriorament del cromosoma Y

Actualment se sap que l'organització del cromosoma Y és el resultat de la combinació de diversos processos evolutius: la degeneració de la massa de gens que inicialment eren comuns entre el parell de cromosomes sexuals primitius; l'adició de gens o regions genòmiques a la regió MSY; i l'acumulació de gens amb funcions específiques de mascle. Tot i que s'ha pogut reconstruir la història evolutiva dels cromosomes sexuals i s'ha pogut descriure el procés de degeneració que pateix el cromosoma Y a nivell molecular, encara se sap poc sobre la raó d'aquesta degeneració (veure apartat 5 del Resum global) i sobre el paper que juga el cromosoma X en aquesta evolució a través de la compensació de dosi. Els canvis en el número de gens actius entre els cromosomes sexuals poden venir donats tant per pèrdua o amplificació de gens al cromosoma Y com per la inactivació de gens del cromosoma X. Aquestes alteracions afecten la quantitat de producte gènic, a no ser que existeixi un mecanisme de compensació de l'expressió gènica (Graves *et al.*, 1998). Hi ha diversos escenaris possibles dins dels processos de desequilibri de dosi gènica, tal com es mostra a la figura 17.

Els gens que només es troben al cromosoma X pateixen inactivació en un dels cromosomes X de les femelles, mentre que els gens amb al·lels actius al cromosoma Y n'estan exempts (Disteché, 1997). La correlació entre dosi gènica i inactivació suggereix que la degradació progressiva del cromosoma Y està sent compensada per la inactivació al llarg del cromosoma X. Però, quin esdeveniment és primer? S'han defensat totes dues alternatives. Per un costat, s'ha proposat que primer hauria començat la decadència d'un gen al cromosoma Y, seguida d'una sobrerregulació de l'expressió del gen al cromosoma X i finalment la inactivació del gen en un dels cromosomes X en les femelles (Graves *et al.*, 1998; Jegalian, Page, 1998). En aquest cas, cal que la inactivació de l'X es doni gen a gen o en petits clústers, mai en tot el cromosoma alhora. Per altra banda, també s'ha defensat l'alternativa contrària, una inactivació al cromosoma X com a primer esdeveniment que alliberaria el cromosoma Y de les constriccions funcionals i el duria a una degeneració o a una especialització funcional i expressió específica (Graves *et al.*, 1998). Si la compensació de dosi ha evolucionat gen a gen, la primera alternativa és més plausible, quan un gen evoluciona a l'Y cap a un malfuncionament, la còpia homòloga a l'X s'inactiva per compensar la dosi. No obstant, també és possible pensar en una sobre-

regulació en blocs sencers de gens adjacents (per exemple, per canvis en l'estructura de la cromatina) que hauria provocat que no tots els gens s'expressessin al nivell òptim. D'aquesta manera, algun gen podria estar sobre-expressant-se a l'X mentre que encara era funcional a l'Y, és a dir, el gen al cromosoma Y es convertiria en redundat i podria acumular mutacions deletèries de manera neutra o fins i tot podria ser inactivat per selecció positiva per compensar la dosi (Bachtrog, 2006; Orr, Kim, 1998). Actualment, tot i disposar de molta informació, encara se sap poc sobre el mecanisme de compensació de dosi i la relació entre inactivació al cromosoma X i degeneració al cromosoma Y.

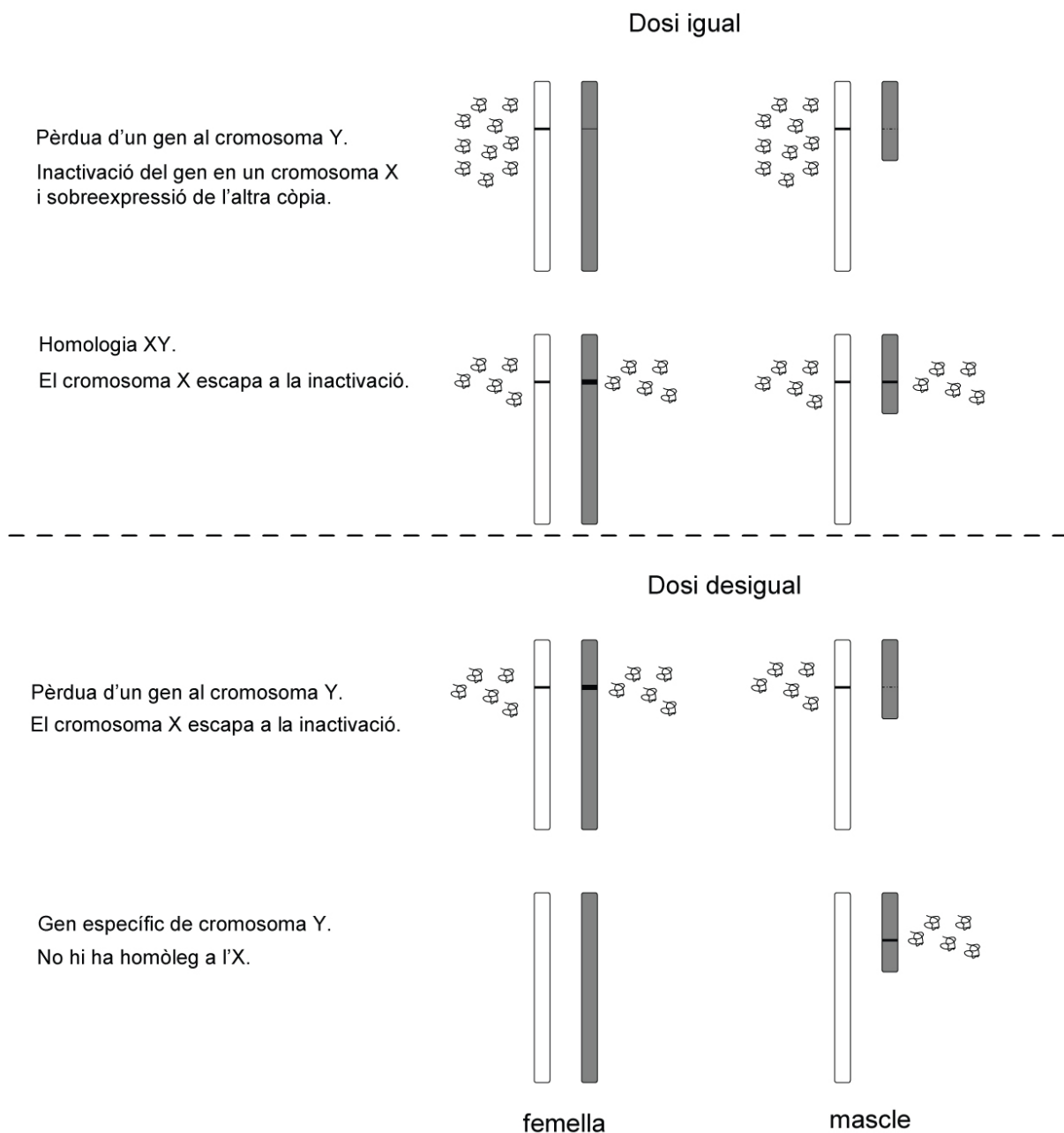


Figura 17. Esquema dels diversos processos de desequilibri de dosi gènica i de possible compensació (adaptat de Graves *et al.*, 1998).

4.4. Malalties associades a cromosoma Y

A la dècada del 1950 es relacionava la transmissió mendeliana de molts trets “masculins” amb el cromosoma Y (Stern, 1957). Poc després, el cromosoma Y va passar a ser considerat com un desert buit de gens, a excepció del gen determinant del sexe (*SRY*) (Stern, 1957). Durant les últimes dècades, les investigacions han anat posant de manifest el contingut gènic d'aquest cromosoma i la importància biomèdica de la regió MSY, que té un rol essencial en processos de determinació del sexe gonadal, creixement de l'esquelet i gènesi tumoral (p.e. Quintana-Murci *et al.*, 2001). Anàlisis de genòmica comparada del cromosoma Y humà amb espècies properes són essencials per resoldre moltes de les incògnites de les malalties relacionades amb el cromosoma Y.

Un dels desordres cromosòmics més comuns en dones i nenes és la síndrome de Turner, associada clàssicament a un cariotip 45,X (o 45,X0). Aquesta síndrome es caracteritza per una sèrie de trets facials típics, deformacions somàtiques variades i un desenvolupament defectuós dels ovaris que origina manca del desenvolupament dels caràcters secundaris en la pubertat, amb amenorrea primària i estatura baixa. Quan es va identificar el gen *SRY* (Sinclair *et al.*, 1990) es va pensar que aquest locus era l'únic responsable de la síndrome i que el seu estudi permetria entendre el mecanisme d'activació del gen i formació del testicle, però l'origen de la malaltia ha resultat ser més complex. S'ha hipotetitzat que l'haploinsuficiència d'alguns gens concrets presents tant en el cromosoma X com en l'Y seria la responsable d'alguns trets somàtics de la síndrome (Rao *et al.*, 1997; Zinn *et al.*, 1993), però per a la majoria de casos la identitat d'aquests gens encara és desconeguda. És molt probable que aquests gens que tenen un efecte en la síndrome de Turner siguin els classificats com a gens X-degenerats, ja que són els que estan presents tant al cromosoma Y en còpia simple com a l'X (Skaletsky *et al.*, 2003). S'ha descrit un possible rol de l'haploinsuficiència dels gens *RPS4X/RPS4Y* en alguns fenotips associats a la monosomia de l'X, com letalitat i anormalitats limfàtiques (Zinn, Ross, 1998), tot i que s'havia descartat que l'haploinsuficiència de l'*RPS4Y* fos l'única causa de la síndrome de Turner (Geerkens *et al.*, 1996).

Altres anomalies relacionades amb presència/absència del cromosoma Y són la disgènesi gonadal (o síndrome de Swyer), els mascles XYY i els avortaments espontanis recurrents. Per altra banda, també s'ha descrit que el cromosoma Y podria estar involucrat en oncogènesi, però el seu paper encara no està clar ni tothom l'accepta (McElreavey, Quintana-Murci, 2003).

Una altra vessant de la recerca biomèdica se centra en l'estudi de l'espermatogènesi i explora les bases genètiques de la infertilitat. Delecions i microdelecions en la regió MSY són la causa genètica principal dels problemes relacionats amb espermatogènesi en les poblacions humanes. Actualment s'accepta que delecions en tres zones de la regió MSY, anomenades AZFa, AZFb i AZFc, provoquen infertilitat (Vogt *et al.*, 1996). El coneixement de la seqüència de la regió MSY ha permès definir delecions recurrents, identificar els gens absents en les delecions (generalment gens membres de famílies específiques de testicle) i demostrar que la majoria de delecions provenen de recombinacions homòlogues entre amplicons gairebé idèntics (p.e. Kamp *et al.*, 2000; Kuroda-Kawaguchi *et al.*, 2001). Tot i els avenços aconseguits, encara no s'han pogut definir quin o quins gens són els realment responsables dels fenotips causats per les delecions (Ali, Hasnain, 2003).

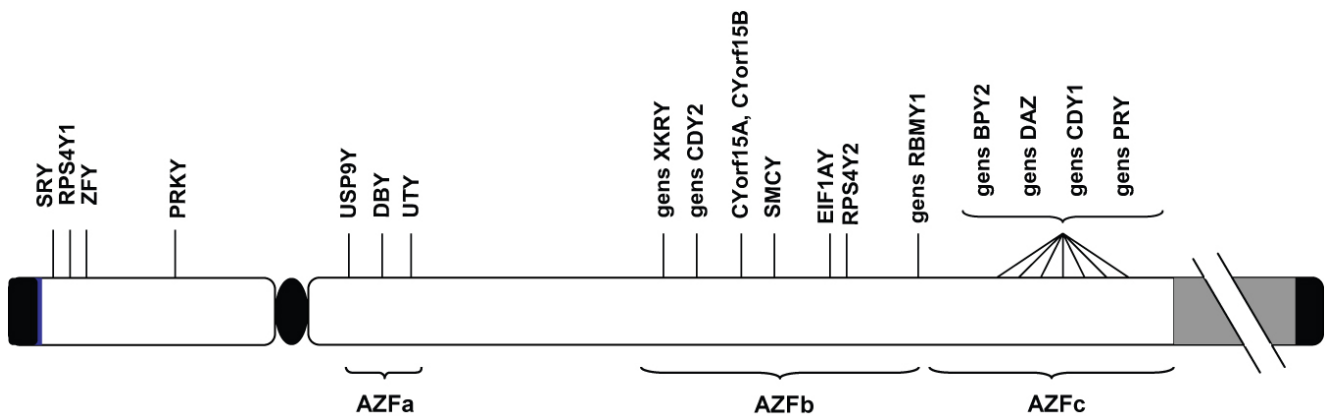


Figura 18. Esquema del cromosoma Y humà, amb els gens principals i les tres regions d'azoospermia (adaptat de Quintana-Murci, 2001).

4.5. Variabilitat i marcadors de DNA al cromosoma Y

Les primeres anàlisis genètiques del cromosoma Y van començar, com en la majoria dels casos, amb l'estudi dels humans. A banda de l'estudi dels gens del cromosoma Y per conèixer la seva implicació en fertilitat (p.e. Jobling, Tyler-Smith, 2000), aviat es va veure el potencial dels marcadors d'aquest cromosoma per investigar l'evolució recent dels humans des d'un punt de vista patrilíneal (p.e. Underhill *et al.*, 2000) i per aplicacions específiques en genètica forense (p.e. Corach *et al.*, 2001). L'ús de marcadors específics de cromosoma Y en genètica de poblacions d'altres mamífers encara no està molt estès a causa de la idea que el cromosoma Y conté poca variabilitat i per les dificultats tècniques per caracteritzar els marcadors en espècies no model (revisat a Petit *et al.*, 2002). Tanmateix, es comença a percebre el cromosoma Y com una important font d'informació sobre les poblacions (per exemple, per estudiar la història, patrons de migració, estructura poblacional o problemes de consanguinitat) i, gràcies als avenços tecnològics i a la

constant generació d'informació genètica de moltes espècies diferents, l'aplicació de marcadors de cromosoma Y és cada vegada més comú en camps com la genètica de la conservació.

El desenvolupament de marcadors pel cromosoma Y, doncs, s'ha donat principalment en l'espècie humana. Podem distingir entre tres classes diferents de marcadors pel cromosoma Y utilitzats en estudis forenses humans (Butler, 2003): i) els microsatèl·lits (especialment els de repeticions tetranucleotídiques), dels que ja se n'han descrit més de 200 (p.e. DYS19, DYS385) i s'han dissenyat kits per analitzar-los en sèrie de manera automatitzada; ii) els marcadors bial·lèlics, com els SNPs (se n'han descrit més de 250) i els polimorfismes d'insercions Alu (YAP) o altres indels; i iii) el marcador minisatèl·lit MSY1 (DYF155S1), format per 48-114 còpies amb una unitat de repetició de 25 parells de bases amb 5 variants de la seqüència de repetició diferents.

Per examinar marcadors SNP al cromosoma Y s'han utilitzat diferents mètodes, des de tecnologia per analitzar marcador a marcador, com la PCR a temps real (p.e. Lareu *et al.*, 2001), fins a mètodes capaços de dur a terme anàlisis multiplex, com l'espectrometria de masses "en temps de vol" (MALDI-TOF) (p.e. Paracchini *et al.*, 2002), *microarrays* (p.e. Raitio *et al.*, 2001) i *SNaPshot* (p.e. Inagaki, 2002), segons els objectius, les necessitats i la disponibilitat de la tecnologia. De totes maneres, en el context forense, el més plausible actualment és que els SNP del cromosoma Y s'apliquin de manera complementària als microsatèl·lits i no com un mètode aïllat per examinar la variació genètica masculina. En estudis poblacionals, l'ús de diferents marcadors del cromosoma Y amb diferents patrons i taxes de mutació permet inferir diversos esdeveniments poblacionals a diferents escales temporals (p.e. Brion *et al.*, 2003; veure apartat 2 del Resum global).

Fins a l'actualitat, els estudis més comuns de cromosoma Y realitzats en espècies no model han estat els de seqüenciació de determinats gens. Per exemple, diversos estudis en ximpanzé es basen en la seqüència humana per dissenyar els encebadors i es realitzen anàlisis comparatives home-ximpanzé dels nivells de diversitat (p.e. Burrows, Ryder, 1997; Deinard, Kidd, 2000; Hammer, 1995). D'altra banda, també s'han dut a terme estudis amb microsatèl·lits. En el cas del ximpanzé, s'han pogut aplicar els microsatèl·lits desenvolupats per als humans amb bons resultats d'amplificació i variabilitat, ja que es tracta d'una espècie propera (Gusmao *et al.*, 2002) i microsatèl·lits aïllats en humans també s'han pogut caracteritzar en altres primats (Erler *et al.*, 2004). Un altre exemple d'identificació i ús de microsatèl·lits del cromosoma Y en genètica de la

conservació és un estudi sobre diversitat en el llop escandinau (Sundqvist *et al.*, 2001). Pel que fa als SNPs de cromosoma Y, existeixen pocs exemples de la seva aplicació en conservació: Stone *et al.* (2002) van identificar 23 SNPs en cromosoma Y per estudiar la diversitat genètica en el gènere *Pan*, especialment en el ximpanzé, i Seddon *et al.* (2005) van dissenyar, a partir de seqüències de gos, assajos per estudiar 24 SNPs en poblacions de llop escandinau. D'altra banda, Hellborg & Ellegren (2003) van aplicar el mètode dels CATS (Lyons *et al.*, 1997) a l'estudi del cromosoma Y, de manera que van poder dissenyar una bateria de marcadors aplicables a diferents espècies de mamífers, que van anomenar YCATS. En aquest estudi es van dissenyar diferents encebadors exònics flanquejant 48 introns diferents basant-se en la seqüència humana i de ratolí i els van aplicar en 20 espècies de mamífers, amb un resultat d'uns 10 introns de mitjana amplificats per a cada espècie i un total de 100 kb de seqüència cromosòmica. L'estudi del cromosoma Y en genètica de la conservació, doncs, encara és un camp molt incipient que cal seguir desenvolupant.

