

**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

RESUM GLOBAL

Aquesta tesi inclou tres capítols de Resultats en forma d'article, en els que es tracten principalment els aspectes següents: i) la possibilitat de l'ús de productes d'MDA en estudis genètics amb mostres no invasives; ii) la utilitat de la tecnologia dels *microarrays* per a l'estudi de la variabilitat de poblacions de ximpanzés; i iii) l'evolució de la família gènica *RPS4Y* i els mecanismes que han actuat per mantenir les diferents còpies del gen. Cadascun d'aquests treballs conté una discussió específica corresponent a l'estudi en concret. En aquest apartat, doncs, es farà una discussió més extensa i general dels resultats obtinguts i s'englobaran aquests treballs en un context més ampli. Així, es tractarà sobre la potencialitat del mostreig no invasiu i s'avaluarà la manera com cal abordar l'estudi de la variació del cromosoma Y, quina informació aporta aquesta recerca i com serà la investigació del futur. A més s'oferirà una visió general de l'evolució del cromosoma Y i del seu possible destí. Finalment, es farà una reflexió sobre el futur que es planteja en genètica de la conservació i sobre el destí que s'albira per als primats.

1. Potenciar el mostreig no invasiu

En diferents àrees de la genètica, especialment en estudis clínics, forenses i de conservació de les espècies, cada vegada es fa més evident la necessitat de desenvolupar tècniques que permetin l'ús de mostres de les que s'obtenen quantitats molt reduïdes de DNA, com les mostres preservades en parafina o formol, espècimens de museu, productes biològics derivats o mostres recol·lectades de manera no invasiva. Dins de la genètica de la conservació, la possibilitat d'utilitzar DNA extret de mostres no invasives és un aspecte fonamental per poder estudiar les espècies sense pertorbar les poblacions. Aquest tipus de mostreig, però, té una sèrie de limitacions que compliquen la seva aplicació, principalment la contaminació i la poca quantitat i qualitat del DNA que se n'obté (veure apartat 2.3 de la Introducció). Per potenciar l'ús del mostreig no invasiu, doncs, és imprescindible superar aquestes limitacions i millorar la metodologia de mostreig, des del moment de la recol·lecta fins a l'extracció del DNA.

Una de les tecnologies que recentment s'ha convertit en essencial en els estudis genètics és l'amplificació de genomes sencers (WGA), ja que proporciona un gran increment de les quantitats inicials de DNA. Dins d'aquest camp hi ha moltes

tècniques diferents basades en diversos mètodes, però sembla que l'amplificació per desplaçament múltiple és una de les que dona millors resultats (revisat a Lovmar, Syvanen, 2006). És per això que aquesta tecnologia sembla la clau de la solució a la limitació de la poca quantitat de DNA que s'obté de les mostres recollides de manera no invasiva. Tanmateix, alguns estudis suggereixen que l'MDA perd eficiència si és utilitzat sobre DNA degradat (Lage *et al.*, 2003). Calia, doncs, demostrar que aquest mètode era eficient i no introduïa canvis o mutacions si s'utilitzava DNA extret de mostres no invasives, que molt possiblement presenta degradació. En el primer treball presentat en aquesta tesi, doncs, s'ha volgut comprovar la possibilitat d'utilitzar l'MDA sobre mostres de semen de macaco obtingudes de manera no invasiva i verificar que els productes obtinguts fossin adequats per dur a terme anàlisis genètiques.

Els resultats obtinguts en aquest primer treball han estat encoratjadors. És per això que es va decidir ampliar l'espectre de mostres no invasives que calia testar i en un treball posterior realitzat conjuntament amb altres grups, s'ha pogut demostrar que els productes d'MDA obtinguts a partir de diverses mostres de primats també són adequats per a anàlisis genètiques (Rönn *et al.*, 2006). En aquest darrer treball, però, no es van incloure mostres de femtes, que són un dels tipus de mostres no invasives més utilitzats i que presenten més problemes de contaminació i degradació. Seria molt interessant, doncs, testar la tècnica de l'MDA sobre mostres de femtes per tal d'ampliar l'aplicabilitat d'aquesta font de DNA. Actualment s'està produint una contínua millora de les tècniques d'amplificació de genomes sencers gràcies a la importància d'aquestes metodologies en l'estudi de patologies humanes (Wang *et al.*, 2004), de manera que, en el cas que l'MDA no donés bons resultats amb el DNA extret de femtes, es podrien analitzar altres tècniques. El desenvolupament de tecnologies que incrementin les possibilitats d'utilització de mostres obtingudes de manera no invasiva és un dels punts essencials en genètica de la conservació.

D'altra banda, existeixen diverses mesures que es poden aplicar per superar les altres limitacions de les mostres obtingudes de manera no invasiva. La degradació del DNA –que comença just després de la mort d'un individu (o d'una cèl·lula) com a conseqüència, principalment, de la hidròlisi i l'oxidació, dels efectes de les

radiacions, de la presència de microorganismes i de les altes temperatures (Carter, 2000)–, pot minimitzar-se realitzant el mostreig amb el menor lapse de temps entre la producció de la mostra i la seva recol·lecció i emmagatzemant la mostra en condicions òptimes de preservació (veure Prendini, 2002). Pel que fa a les contaminacions, l'amplificació de DNA contaminant inherent a la mostra podria ser evitat mitjançant un bon disseny d'encebadors específics per al tàxon d'interès. A més, la poca quantitat de DNA present a les mostres augmenta el perill de les contaminacions durant tot el procés de manipulació del DNA. Aquest risc es pot reduir realitzant els processos de pre-PCR i post-PCR en cambres separades, fent ús de material estèril o autoclavant els estris i solucions abans de cada ús, duent a terme les anàlisis en laboratoris que no treballin habitualment amb els tàxons d'interès i afegint sempre controls negatius en totes les fases del procés (p.e. Taberlet *et al.*, 1996; Taberlet *et al.*, 1999). Per últim, la inhibició de les reaccions de PCR també pot ser un problema, però actualment existeixen protocols dissenyats per disminuir l'efecte dels possibles inhibidors, com els mètodes d'extracció de DNA basats en sílica (Yang *et al.*, 1998) o en Chelex (Singer-Sam, 1989).

Inicialment, amb les mostres recol·lectades de manera no invasiva només era possible realitzar anàlisis genètiques utilitzant marcadors mitocondrials, ja que les cèl·lules presenten múltiples còpies de DNA mitocondrial (veure apartat 2.3 de la Introducció). Les limitacions del mostreig no invasiu havien estat encara més accentuades en els estudis del cromosoma Y donada la seva haploïdia. Tanmateix, tots els avenços tecnològics que s'han anat produint en els darrers anys han possibilitat i potenciat l'estudi de diferents regions del genoma nuclear (veure apartat 2.3 de la Introducció), fins i tot del cromosoma Y (com és el cas del segon treball presentat en aquesta tesi).

2. Cromosoma Y com a font d'informació

L'estudi del cromosoma Y ha seguit la mateixa evolució que l'estudi de la resta del genoma (sempre, però, de manera més lenta), és a dir, primer des de les observacions fenotípiques, després a partir de les proteïnes i finalment a nivell de DNA (citogenètica, seqüència, marcadors). Aquest cromosoma comparteix una sèrie de característiques amb el DNA mitocondrial que els diferencien de la resta del genoma: no recombinen, estan en estat haploide i són d'herència uniparental. D'altra

banda, els estudis de variabilitat en cromosoma Y són la complementació als estudis amb DNA mitocondrial, ja que el DNA mitocondrial revela la història del llinatge femení mentre que el cromosoma Y és específic del llinatge masculí. Poder disposar de la informació tant de DNA mitocondrial com de cromosoma Y és especialment interessant en aquelles espècies en què els mascles i les femelles presenten comportaments diferents pel que fa als patrons de migració i dispersió, com és el cas dels ximpanzés. És per això que en el segon treball presentat en aquesta tesi es va creure interessant desenvolupar una eina que permetés estudiar els dos tipus de marcadors alhora, de manera ràpida i automatitzada, en ximpanzés. Es va decidir fer una cerca a les bases de dades públiques per trobar polimorfismes d'una sola base en el DNA mitocondrial i es va dur a terme el descobriment de nous SNPs en cromosoma Y per completar els que hi havia a la literatura, amb l'objectiu final de crear un *microarray* que permetés realitzar el genotipatge de tots els SNPs alhora. D'aquesta manera, amb el *microarray* desenvolupat s'obté informació del llinatge matern i patern de les poblacions de ximpanzé en un sol assaig.

Els estudis basats en DNA mitocondrial han estat molt populars tradicionalment, tant en estudis de filogeografia pel fet de ser loci no recombinants i d'herència uniparental, com en genètica no invasiva pel fet que a partir de la mateixa mostra s'obté una quantitat de DNA mitocondrial molt més elevada, i en millor estat, que de DNA nuclear. El DNA mitocondrial, però, té una sèrie d'inconvenients que cal tenir en compte. D'una banda, existeix un nombre creixent de còpies de gens mitocondrials que es transposen al nucli i formen una superfamília de DNA mitocondrial nuclear (*Numt*), que són considerats pseudogens a causa de les diferències entre el codi genètic mitocondrial i citoplasmàtic (revisat a Bensasson *et al.*, 2001). Els *Numts*, doncs, poden ser un problema greu en les anàlisis de variabilitat perquè introdueixen errors i per això s'han desenvolupat tècniques per minimitzar la seva interferència, tot i que no sempre tenen prou potència per solucionar completament el problema (Bensasson *et al.*, 2001). A més, existeix la qüestió de l'heteroplàsmia (revisat a Kmiec *et al.*, 2006), la barreja de diversos tipus de mitocondri en un sol individu, que comporta que un únic individu contingui seqüències de DNA mitocondrial portadores de diferents mutacions. D'altra banda, els marcadors mitocondrials no són útils per investigar la pèrdua recent de variabilitat genètica ni esdeveniments a nivell individual, com la identitat, dispersió individual i

sistemes d'aparellament. Tot i així, el DNA mitocondrial segueix sent una bona eina en genètica de la conservació per resoldre incerteses taxonòmiques en les categories inferiors (família, gènere, espècie) gràcies a la seva evolució ràpida. A més, el DNA mitocondrial és un bon marcador auxiliar al DNA nuclear perquè permet estudiar la variabilitat genètica d'herència materna.

El cromosoma Y, a més d'aportar informació complementària a la del DNA mitocondrial perquè dóna la visió del llinatge patern, té propietats específiques que el fan més interessant. D'una banda, un sol cromosoma Y és transmès a través de la línia germinal i n'hi ha un per cèl·lula, de manera que no pateix heteroplàsmia. D'altra banda, el cromosoma Y conté una varietat més gran de marcadors polimòrfics que el DNA mitocondrial i aquests marcadors cobreixen tot el rang de taxes de mutació, des de les més baixes dels SNPs (de Knijff, 2000) fins a les més elevades dels microsatèl·lits (Kayser *et al.*, 2000). Pels estudis poblacionals, però, l'ideal és combinar marcadors tant del DNA mitocondrial i del cromosoma Y com de la resta del genoma nuclear, per poder comparar informació general de la població amb informació específica de llinatge matern i patern i poder deduir així la dinàmica poblacional.

3. Marcadors: SNPs versus Microsatèl·lits

El segon estudi presentat en aquesta tesi tenia com objectiu principal analitzar en un estudi pilot la possibilitat de transferir la tecnologia desenvolupada en humans a l'estudi d'espècies no model. Per assolir aquest fi, com ja s'ha vist, es va decidir desenvolupar una eina que permetés analitzar la variabilitat dels ximpanzés a partir de polimorfismes d'una sola base de cromosoma Y i de DNA mitocondrial i que contribuís així a l'estudi i conservació d'aquesta espècie. Es va optar per la utilització d'SNPs pels avantatges que ofereixen aquests marcadors front altres marcadors més clàssics. Com veïem a la Introducció d'aquesta tesi (apartat 4.5.), dins del cromosoma Y hi ha diferents tipus de marcadors moleculars. En general, no existeix un marcador "universal" i la tria del marcador a utilitzar en cada moment depèn de diversos factors, com el temps, el cost, la disponibilitat i, principalment, la qüestió a resoldre. A més, en el cas del cromosoma Y, existeix la dificultat afegida que fins fa poc hi havia un número molt baix de marcadors desenvolupats en humans i que

encara ara n'hi ha molt pocs per a espècies no model (veure apartat 4.5. de la Introducció). Inicialment es van aïllar microsatèl·lits del cromosoma Y d'humans (Ayub, 2000), que han estat molt utilitzats en genètica forense (p.e. Corach *et al.*, 2001) i també s'han aplicat amb èxit en l'estudi d'espècies properes, com el ximpanzé (p.e. Gusmao *et al.*, 2002) (veure apartat 4.5. de la Introducció). Per avaluar específicament les migracions històriques, però, és essencial l'ús de polimorfismes amb baixa taxa de mutació, que permeten reconstruir l'estat ancestral i que clarament preserven la informació haplotípica específica de la població al llarg d'una escala temporal que cobreixi tota la història de l'espècie. En aquest cas, els polimorfismes bial·lèlics associats a la regió específica de mascle del cromosoma Y són un bon marcador haploide que reflecteix l'ancestre parental. El potencial del cromosoma Y per explorar qüestions relacionades amb l'evolució dels humans moderns ha estat reconegut des de fa temps (p.e. Casanova *et al.*, 1985; Ngo *et al.*, 1986), però l'ús d'SNPs del cromosoma Y com a marcadors en estudis evolutius és recent en humans (p.e. Underhill *et al.*, 1996) i encara és poc comú en estudis amb altres espècies (p.e. Meadows *et al.*, 2004; Natanaelsson *et al.*, 2006) i en estudis de conservació (p.e. Stone *et al.*, 2002).

L'elevada taxa de mutació dels microsatèl·lits fa que siguin bons marcadors per analitzar esdeveniments històrics recents i per resoldre qüestions de relació, estructura i classificació a nivell poblacional i individual. Tanmateix, els microsatèl·lits tenen certes limitacions com a marcadors. Tot i que en general es poden utilitzar microsatèl·lits aïllats en espècies properes, el fet que aquests marcadors es trobin principalment en regions no codificadores (on la taxa de substitució nucleotídica és elevada) comporta que fàcilment es produeixin canvis en la regió flanquejant del microsatèl·lit, que poden impedir que l'encebador hibridi, i canvis dins la seqüència del microsatèl·lit, que poden fer que un microsatèl·lit molt variable en una espècie sigui monomòrfic en una altra. Així doncs, és recomanable fer aïllaments *de novo* per a cada espècie. El fet que la variació dels microsatèl·lits depengui del número de repeticions i que la variació sigui detectada per electroforesi comporta també certes limitacions. A la regió flanquejant es poden generar indels que canviïn la mida del fragment i que el facin migrar junt amb fragments amb un número de repeticions diferent, és a dir, es pot donar un problema d'homoplàsia (veure López-Giráldez, 2006). A més, les mutacions puntuals que es produeixen a les regions flanquejants

poden impedir la unió de l'encebador i provocar la no amplificació d'un al·lel (al·lel nul), de manera que es detectarien falsos homozigots que farien disminuir l'estimació dels heterozigots.

L'homoplàsia i els al·lells nuls, que fan subestimar la diversitat genètica, tenen un efecte més o menys pronunciat en funció de la grandària de la població. Aquests factors no suposen un problema greu en poblacions grans perquè els microsatèl·lits són molt variables. En poblacions en perill d'extinció, però, aquestes limitacions sí que són una qüestió problemàtica perquè en les poblacions petites, en declivi, acostuma a haver-hi una pèrdua d'al·lells important (s'ha calculat que el 87,5% dels animals en perill tenen menys de 5 al·lells per locus). A més, els problemes dels microsatèl·lits es fan més importants quan s'utilitzen mostres no invasives, com pèl i femtes, de les que s'obté molt poc DNA, degradat i contaminat, cosa que afegeix més limitacions. Així, es poden donar fenòmens de no amplificació d'un al·lel perquè per atzar l'al·lel no està representat a la reacció de PCR (*allelic dropout*) i fenòmens d'amplificació de falsos al·lells a causa de la contaminació, que s'han intentat solucionar amb el mètode de múltiples-tubs (Navidi *et al.*, 1992; Taberlet *et al.*, 1996). Finalment, les dades generades de l'anàlisi de microsatèl·lits en diferents laboratoris no sempre són comparables perquè poden haver-hi diferències en la mida dels al·lells degudes al mètode utilitzat, a la màquina seqüenciadora o a les condicions de PCR (per exemple, pot ser que s'afegeixi una adenina a 3' del producte amplificat).

Pel que fa als polimorfismes d'una sola base (SNPs), el principal problema d'aquests marcadors està en la primera fase, en el descobriment i la tria d'SNPs. Actualment hi ha una manca d'informació generalitzada: hi ha pocs SNPs identificats, poques seqüències per cercar-los i per a poques espècies. La taxa de mutació dels SNPs, de l'ordre de 10^{-9} (Graur, 1999), és molt més baixa que la dels microsatèl·lits, que és de l'ordre de $10^{-2} - 10^{-6}$ (Schlötterer, 2000a), de manera que els SNPs no solen ser recurrents ni es troben conservats entre espècies i això implica que per a cada espècie estudiada sempre s'han de descobrir *de novo*. El procés de descobriment d'SNPs és lent i d'un cost elevat. D'altra banda, un dels greus problemes dels SNPs és el biaix de composició (*ascertainment bias*), que resulta de la selecció de loci a partir d'una mostra d'individus no representativa, o utilitzant un mètode particular, i

que dona un espectre no representatiu de les freqüències al·lèliques a la població. Es tracta d'una qüestió crítica en moltes aplicacions dels SNPs (Morin *et al.*, 2004) perquè pot introduir un biaix sistemàtic en les estimes de la variació dins i entre poblacions. És molt important, doncs, aplicar un bon protocol per a la identificació d'SNPs i tenir en compte també el número i origen dels individus analitzats per tal d'evitar un efecte de biaix de composició.

Tot i que el biaix de composició pot ser un problema per a algunes aplicacions dels SNPs, aquests marcadors poden ser una eina eficient i útil gràcies a les seves característiques, com la gran potència estadística que generen, l'extensa cobertura del genoma que ofereixen, la qualitat de les dades que se n'obté i la possibilitat de tipar molts marcadors alhora. A diferència dels microsatèl·lits, les dades generades a partir de l'anàlisi d'SNPs són perfectament comparables entre laboratoris diferents gràcies al fet que els SNPs són marcadors bial·lèlics i el resultat de tipar-los, que sempre és homogeni, pot ser codificat com a *Sí* o *No*, i això simplifica l'anàlisi posterior, que pot ser automatitzada fàcilment. Per poder construir bons models estadístics cal conèixer bé el procés de generació de les dades, especialment quins són els errors de genotipatge que es poden produir. Les anàlisis dels SNPs generen només dos tipus possibles d'error, la no amplificació d'un al·lel o el procés contrari, considerar un homozigot com a heterozigot, però no es poden generar falsos al·lèls, cosa que facilita el disseny de models estadístics. A més, hi ha anàlisis que impliquen l'aplicació d'un model mutacional i es coneix millor el model mutacional pels SNPs, model d'al·lèls infinits, que el model pels microsatèl·lits, que sembla que és el model mutacional per *stepwise*. Quan es treballa amb mostres no invasives, un gran avantatge dels SNPs és que poden ser utilitzats amb DNA parcialment degradat perquè permeten treballar amb fragments curts, ja que requereixen una PCR de només unes 50 bases (els encebadors que flanquegen l'SNP han de ser d'unues 20 bases cadascun). Ara bé, quan es treballa amb polimorfismes d'una sola base cal tenir en compte que sempre és necessari un número molt més elevat d'SNPs que de microsatèl·lits per obtenir la mateixa informació, ja que els microsatèl·lits tenen molts al·lèls per a cada locus mentre que els SNPs són bial·lèlics (per exemple, per a anàlisis de relació cal utilitzar un nombre molt elevat d'SNPs independents, de 40 a 100, amb freqüències al·lèliques elevades, mentre que amb 7-14 microsatèl·lits s'obtindria la mateixa informació (Morin *et al.*, 2004)).

Les propietats dels SNPs i el recent desenvolupament de tecnologia d'alt rendiment, podrien fer que en pocs anys els SNPs substituïssin els microsatèl·lits com a marcador preferit en moltes aplicacions en genètica de la conservació. Cal no oblidar, però, que les utilitats i limitacions dels SNPs front els microsatèl·lits en genètica de poblacions depenen, en gran part, de les qüestions que es volen resoldre i que sempre cal anar amb compte amb el possible biaix de composició (Schlötterer, 2004). Així, els SNPs poden donar molt bons resultats en estimes de la variació genètica i en identificació individual, de parentiu i de relació (Krawczak, 1999; Manel *et al.*, 2002). A més, en anàlisis d'estructura poblacional, els SNPs ofereixen una millor representació de tot el genoma i permeten identificar *outliers*, que podrien ser loci sota selecció, mentre que l'elevada taxa de mutació dels microsatèl·lits fa que l'homoplàsia sigui comú i les estimes incorrectes (però, en cas d'existir, el biaix de composició dels SNPs és un greu problema en aquest tipus d'estudis). Com que les estimacions de la grandària poblacional se solen calcular a partir de mostres no invasives, l'aplicació dels SNPs pot donar millors resultats que els microsatèl·lits. El possible biaix de composició té poc efecte en les estimes de N_e a curt termini mentre que afecta més a les de llarg termini perquè tenen en compte l'heterozigotitat a la població. Pel que fa a estudis dels canvis en la grandària poblacional, en canvi, en general els SNPs són menys útils que els microsatèl·lits perquè pel fet de ser bial·lèlics la potència per detectar la pèrdua de variabilitat genètica o un coll d'ampolla recent és menor (Brumfield *et al.*, 2003). A més, els SNPs tampoc són útils per detectar expansions poblacionals perquè la fixació de nous al·lels és molt lenta. Actuant, en alguns estudis de conservació ja s'han començat a utilitzar els SNPs (p.e. Batley, Hayes, 2003; Belfiore *et al.*, 2003; Seddon *et al.*, 2005) i en pocs anys aquest ús pot convertir-se en habitual. Els SNPs seran més populars en conservació quan s'hagin desenvolupat tècniques que permetin l'aïllament i l'escaneig d'SNPs ràpid i econòmic en moltes espècies.

Dins del gran ventall de tècniques possibles de genotipatge d'SNPs, en el segon treball d'aquesta tesi es va seleccionar el *microarray* de miniseqüència perquè és una tècnica àmpliament utilitzada en humans pels seus grans avantatges. Com ja s'ha vist, la tecnologia del *microarray* ofereix la possibilitat d'analitzar molts SNPs alhora en moltes mostres simultàniament, de manera que és un mètode molt eficient. L'inconvenient és que el seu desenvolupament requereix un gran esforç inicial de

descoberta d'SNPs, que s'ha de fer per a cada nova espècie que es vol estudiar, i en espècies no model és un procés força lent i costós. El nostre estudi sobre una població captiva de ximpanzés ha demostrat que el *microarray* desenvolupat és fiable ja que els resultats de l'anàlisi han coincidit amb l'arbre genealògic conegut. Així, aquest sistema pot ser una eina molt útil en la gestió d'espècies captives, per exemple per dur a terme tests de paternitat o programes de reproducció. D'altra banda, l'anàlisi de dues poblacions salvatges ha determinat que el *microarray* permet definir adequadament l'estructura poblacional i la història demogràfica pel llinatge matern, tot i que la història evolutiva de les poblacions no pot ser ben resolta. Amb aquest sistema, però, no s'ha pogut definir l'estructura poblacional esperada segons les dades ecològiques per la línia paterna. Les limitacions trobades probablement són degudes a un problema de mida mostral ja que es disposava d'un número molt baix de mascles –limitació que pot ser freqüent en estudiar poblacions salvatges i que no sempre es podrà evitar– i a un número insuficient d'SNPs analitzats. Tot i que en humans s'ha estimat que de 30 a 50 SNPs serien suficients per a la identificació individual (Chakraborty *et al.*, 1999), els SNPs de regions que no recombinen són menys informatius pel fet d'estar lligats. Per tant, per tal que aquest *microarray* pugui ser una eina útil per a l'estudi i gestió de poblacions salvatges caldria optimitzar-lo incrementant el número d'SNPs. A més, si s'afegissin SNPs de regions autosòmiques el *microarray* aportaria informació de totes les regions genòmiques i es podria convertir en una eina de gran suport en l'estudi i conservació de poblacions de ximpanzé, fins al punt que podria desplaçar altres tècniques i marcadors.

4. Marcadors i genotipatge del futur (en genètica de la conservació)

Durant els darrers anys, els avenços tecnològics que s'han anat produint dins el camp de la genètica han fet avançar molt els estudis científics i la manera de fer ciència està canviant cap a una concepció més global, a nivell genòmic i d'interaccions de sistemes. Com s'ha anat veient, el potencial de l'ús de la genètica en conservació augmenta gràcies als avenços tecnològics que permeten l'escaneig massiu d'un gran nombre de marcadors de DNA d'una manera econòmica i ràpida per a un nombre d'espècies cada vegada més gran. Apareixen nous mètodes, com

els *microarrays* de DNA, i noves disciplines, com la genòmica, que revolucionaran el camp de la genètica de la conservació en pocs anys.

Tot i la importància del mostreig no invasiu i el continu desenvolupament de tècniques per potenciar el seu ús (com s'ha vist en el primer apartat del Resum global), també cal tenir en compte que és essencial la creació de col·leccions i l'emmagatzematge en bones condicions de material biològic (teixits, DNA, línies cel·lulars, etc.) disponible per desenvolupar projectes d'investigació. D'aquesta manera s'aconseguiria evitar el remostreig innecessari, es disminuiria l'ús d'animals vius en recerca i s'ajudaria en l'estandardització d'eines i protocols moleculars. Aquest tipus d'iniciativa és especialment important en primats, ja que són el focus de molts estudis, tant en biomedicina com en conservació. Per això ja existeixen alguns projectes, tant a la Comunitat Europea (www.INPRIMAT.org) com a Estats Units (www.IPBIR.org), que tenen com a objectiu la creació de bancs de material biològic de primat i bases de dades de les col·leccions per fer accessible tot el material a la comunitat científica (Ryder, 2005). D'altra banda, els mètodes d'amplificació genomes sencers, que ja s'ha vist que són essencials per potenciar el mostreig no invasiu, també poden ser molt útils per crear reservoris de material genètic pràcticament inesgotables que contribuirien a evitar el remostreig innecessari. Aquestes tècniques presenten diversos avantatges metodològics respecte altres estratègies utilitzades actualment, com la immortalització de cèl·lules (*cell lines*), ja que la seva aplicació és molt més senzilla, permeten partir de gairebé qualsevol tipus de mostra i no comporten fenòmens de reorganització cromosòmica.

Dins del camp de la tecnologia d'alt rendiment els *microarrays* s'estan convertint en una eina popular. Els *microarrays* de DNA generalment són milers de fragments de DNA (de 20 a 25 bases) limitats dins un patró en xarxa sobre un portaobjectes o un paper de filtre de niló de la mida d'una targeta de crèdit. El concepte bàsic dels *microarrays* s'ha anat adaptant i combinant amb altres tècniques i s'ha escampat per moltes àrees dels estudis genètics (revisat a Hoheisel, 2006). El més comú són els *microarrays* per detectar expressió gènica, que permeten estudiar quins gens estan relacionats, per exemple, amb adaptació a un determinat canvi ambiental o a malalties produïdes per un patogen. Recentment, però, s'ha vist que els *microarrays* també poden ser utilitzats per analitzar SNPs i que permeten el genotipatge de

milers d'SNPs simultàniament (Syvänen, 2001). Ja que la precisió i el rendiment dels *microarrays* ja està avançant molt (Jobs *et al.*, 2003), el futur dels *microarrays* de genotipatge està en el desenvolupament de tècniques que permetin augmentar la sensibilitat i disminuir el temps de preparació de les mostres, cosa que es pot aconseguir utilitzant mètodes de detecció que no es basin en l'ús de sondes (p.e. Arntz *et al.*, 2003; Brandt *et al.*, 2003; Hahm, Lieber, 2004). D'altra banda, un altre gran avenç és el disseny d'una base universal per a tots els *microarrays*, és a dir, que no s'hagi d'establir i produir una nova plataforma per a cada nova aplicació (Gerry *et al.*, 1999), idea que ja s'havia començat a desenvolupar i aplicar en diferents estudis (Shoemaker *et al.*, 1996; Urata *et al.*, 1992). Aquesta aplicació és la que ha estat testada en el segon treball d'aquesta tesi.

Tot i la potencialitat de la tecnologia del *microarray*, el futur de l'ús dels SNPs en les espècies no model es troba en la genòmica i en la possibilitat de seqüenciar de manera eficient, ja que el procés d'elaboració d'un *microarray* és lent i car i només seria rendible, per exemple, si un sol *microarray* produït a gran escala servís per respondre moltes qüestions alhora (Rubenstein, 2003) o quan fos necessari testar moltes mostres alhora de manera rutinària, com en el control de qualitat de productes animals, com làctics o de la carn (Peter *et al.*, 2004), o de productes vegetals (Ronning *et al.*, 2005). Així doncs, sembla que la seqüenciació és la clau en conservació. Actualment s'està millorant una tecnologia molt prometedora, el piroseqüenciat, que en el futur permetrà seqüenciar de manera massiva el genoma de qualsevol organisme. Aquesta metodologia es basa en el principi de "seqüenciar a través de sintetitzar", que va desenvolupar a la dècada de 1990 el grup de Nyrén (actualitzat a Gharizadeh *et al.*, 2006), i que es basa en la reacció enzimàtica quimiluminiscent que s'origina quan hi ha un reconeixement molecular. El mètode permet seqüenciar una cadena simple de DNA a mida que se sintetitza la cadena complementària i cada vegada que un nucleòtid és incorporat a la cadena creixent la cascada de reaccions enzimàtiques que es genera dona com a resultat un senyal lumínic, que pot ser detectat. El piroseqüenciat és un procés completament automàtic, fiable i de precisió, que permet analitzar moltes mostres en molt poc temps i és el mètode actual de seqüenciació més ràpid i eficient. Tanmateix, de moment aquesta tècnica té el gran inconvenient que tan sols es poden seqüenciar fragments curts de DNA, i per això actualment només està sent utilitzada en re-

seqüenciacions o per seqüenciar genomes d'espècies properes a altres de les que ja s'ha obtingut la seqüència del genoma. D'altra banda, el piroseqüenciat també s'ha començat a aplicar per al descobriment i genotipatge d'SNPs amb molt bons resultats (Fakhrai-Rad *et al.*, 2002).

En qüestions de gestió de les poblacions és necessari obtenir informació de les anàlisis de manera ràpida per poder actuar amb eficiència. Per això s'està invertint en el desenvolupament de sistemes d'identificació ràpida i es fan grans esforços per millorar la tecnologia d'alt rendiment. Dins aquest context, l'any 2003 Hebert *et al.* (Hebert *et al.*, 2003) van proposar la tècnica del DNA *barcoding*, que defensa que caldria caracteritzar tantes espècies com fos possible pel mateix marcador genètic i que tota la informació referent a l'espècimen i a la seqüència hauria d'estar disponible en una base de dades pública, de manera que la potència del projecte incrementaria a mida que s'anessin caracteritzant més tàxons. Per als eucariotes es va triar com a codi de barres potencial una regió de 648 parells de bases del gen mitocondrial per a la subunitat 1 de la citocrom oxidasa C (COI). Aquesta proposta ha obert un gran debat sobre la seva aplicabilitat real en conservació, amb grans defensors de la idea (DeSalle, 2006; Janzen, 2004) i grans detractors (p.e. Rubinoff, 2006; volum de febrer del 2003 de la revista *Trends in Ecology and Evolution*). Recentment s'ha testat la possible utilitat del DNA *barcoding* en la diagnosi de material biològic de primat (Lorenz *et al.*, 2005) i s'ha demostrat el gran paper que podria jugar, per exemple, en conservació i en el control de l'aplicació de les lleis de protecció de les espècies. Aquest estudi també posa de manifest la importància de disposar de bones seqüències de referència ben caracteritzades a tots els nivells (informació sobre la mostra, l'individu, l'espècie, etc.) per poder comparar-les amb les de les mostres a estudiar.

Podem predir que en un futur no gaire llunyà els sets de dades seran molt més grans que actualment, amb molts més loci, individus i poblacions, cosa que permetrà que les estimes dels paràmetres poblacionals siguin molt més acurades i precises. Aquests grans sets de dades seran comuns gràcies a avenços com l'automatització i les tecnologies d'alt rendiment, com l'ús de robots pipetejadors i seqüenciadors de capil·lars més ràpids que els actuals. La PCR multiplex, que és la co-amplificació per PCR de més d'un locus en un sol tub utilitzant parells d'encebadors per a més d'un

locus, pot reduir molt el temps i els costos que impliquen produir grans quantitats de dades. L'estalvi també pot venir de l'eliminació dels passos post-PCR de tractament de productes de PCR. A més, dur a terme diverses anàlisis en paral·lel també permet aconseguir un alt rendiment (amb els *microarrays*, per exemple). El que és essencial, però, és fer un esforç de recopilació i emmagatzematge de tota la informació que es va generant per fer-la pública i que tota la comunitat científica pugui utilitzar-la. Per això són primordials les bases de dades de lliure accés (p.e. GenBank, EMBL), que recullen informació de seqüències parcials, i projectes de catalogació de marcadors, com el HapMap (Gibbs *et al.*, 2003), que és un catàleg de les variants genètiques comunes en els humans i de la seva informació poblacional i genòmica.

La genòmica es pot definir com l'estudi de l'estructura i funció d'un nombre molt elevat de gens i la genòmica de poblacions fa referència a l'estudi de molts marcadors de DNA en molts individus de diferents poblacions (Luikart *et al.*, 2003). Els grans avantatges d'estudiar molts loci alhora són l'increment de potència estadística en la majoria d'anàlisis i la millora en la cobertura del genoma, cosa que permet fer inferències menys esbiaixades i identificar possibles loci que es comportin de manera diferent i que, doncs, podrien estar afectats per la selecció i tenir importància, per exemple, en eficàcia biològica, adaptació local o especiació. Els estudis de genòmica de poblacions per ser informatius requereixen: 1) el desenvolupament de recursos genòmics, com un gran nombre de marcadors de DNA (millor si estan mapats) o seqüències de fragments de gens; i 2) capacitat de genotipatge a gran escala facilitat per la PCR multiplex o per assajos de genotipatge ràpids.

Tal com s'ha anat comentant al llarg d'aquest Resum global, sembla que el que realment revolucionarà el camp de la genètica –i de la genètica de la conservació– són els avenços que s'estan produint per aconseguir una seqüenciació ràpida i econòmica de genomes complets. L'any 2001, l'*International Human Genome Sequencing Consortium* va publicar els primers resultats de la seqüenciació del genoma humà complet. Des de llavors s'ha treballat per millorar i completar els buits d'aquest primer esbós i s'han dut a terme projectes per obtenir els genomes complets de moltes altres espècies animals, tant d'animals models (p.e. la mona

Rhesus, *Macaca mulatta*, o el ratolí comú, *Mus musculus*) com d'espècies importants en producció animal (p.e. el gall, *Gallus gallus*, o el bou, *Bos taurus*) o per la seva posició dins l'arbre filogenètic (p.e. el ximpazé, *Pan troglodytes*, o el gos domèstic, *Canis familiaris*). Totes les dades que es van generant es fan públiques i es posen a l'abast de tothom a través de les pàgines web dels projectes (p.e. www.ensembl.org). Aquesta informació té un gran valor en l'estudi de les espècies seqüenciades i per fer anàlisis de genòmica comparada, però, a més, també és bàsica a l'hora d'estudiar espècies properes de les que encara no s'han generat seqüències. Per exemple, a partir de les seqüències disponibles es poden dissenyar encebadors universals per a un grup determinat d'espècies properes filogenèticament a l'espècie model, tal com vèiem a la Introducció (apartat 2.5.1). Així, tots aquests coneixements són de gran utilitat en genètica de la conservació ja que serveixen de base per dissenyar estudis genètics.

Un exemple clar de la rapidesa amb què es produeixen els avenços dins el camp de la genètica i la genòmica és el cas de l'estudi sobre variabilitat en el cromosoma Y de ximpanzé presentat al segon capítol d'aquesta tesi. En el moment de començar l'estudi no estava disponible la seqüència del genoma d'aquesta espècie i vam haver de plantejar la recerca a partir de la informació existent. D'aquesta manera, es va decidir dissenyar encebadors sobre la seqüència de cromosoma Y humana per amplificar fragments intrònics de ximpanzé i aplicar la tecnologia de la DHPLC (Oefner, Underhill, 1995) per tal de portar a terme el descobriment dels SNPs. Recentment, però, el cromosoma Y de ximpanzé ha estat seqüenciat i, a més, s'ha fet per duplicat, en dos individus diferents, cosa que ha permès comparar les dues seqüències i determinar les diferències existents. Aquestes diferències poden ser, en part, errors de seqüència, però la majoria seran polimorfismes. En aquests moments, doncs, el plantejament d'un estudi com el realitzat en aquesta tesi podria ser completament diferent gràcies a la nova informació disponible.

La gran massa d'informació que s'està generant, i que incrementa exponencialment amb el temps, ha fet aparèixer recentment una nova disciplina, la bioinformàtica, que és l'aplicació d'eines computacionals a la gestió i l'anàlisi de dades biològiques. El desenvolupament d'aquest camp ha estat espectacular i pateix un avenç continu per ajustar-se a les demandes de la recerca, que necessita la creació de mètodes

analítics i programes informàtics cada vegada més ràpids, potents i fiables, i que permetin el tractament d'una gran quantitat de dades. Dins la bioinformàtica, és necessari també seguir treballant en la millora dels mètodes estadístics, per exemple, dissenyant mètodes de detecció de la selecció positiva més sensibles que permetin discriminar l'acció de la selecció quan aquesta és dèbil. El desenvolupament d'aplicacions informàtiques ha estat bàsic en tots els camps de la biologia, i entre ells en la genètica de la conservació. Inicialment el focus de la bioinformàtica estava en la creació d'eines i tècniques analítiques pel tractament de seqüències nucleotídiques i proteiques, però actualment s'avança cap a un intent d'integració de molts tipus de dades diferents i d'àrees diverses per poder fer anàlisis més globals que tinguin en compte tots els factors necessaris. Encara falta fer un esforç perquè les dades de tipus ecològic estiguin disponibles en bases de dades de la mateixa manera com ho estan les dades genètiques. La possibilitat d'integrar informació ecològica i genètica és una de les fites que més contribuiran a la gestió dels recursos de la terra i de la biodiversitat (revisat a Jones *et al.*, 2006).

5. Evolució del cromosoma Y. Efecte de la selecció sexual

Fins aquí s'ha anat veient que el cromosoma Y és una font d'informació interessant per a l'estudi de les poblacions ja que permet definir l'estructura poblacional paterna i determinar si existeix dispersió diferencial entre mascles i femelles. La tecnologia d'alt rendiment, que pot ser aplicada a aquests estudis, ha facilitat i optimitzat l'obtenció de resultats. Aquestes millores tècniques i l'adquisició d'informació són essencials perquè la genètica de la conservació avanci en el seu objectiu de preservar les espècies. El cromosoma Y, però, no només és interessant com a marcador sinó que, a partir del coneixement que aquest cromosoma no era un desert mancat de gens, el seu estudi ha centrat l'atenció dels científics gràcies al fet que conté el gen que determina el sexe en els mamífers i a les característiques que el fan un cromosoma únic. Entre aquestes propietats cal destacar el seu estat haploide, la presència exclusiva en un dels sexes, la manca de recombinació gairebé total i el fet que s'hi ha produït una acumulació de gens, i famílies gèniques, relacionats amb espermatogènesi i amb expressió específica a testicle. Donada la

proximitat filogenètica entre l'home i la resta de primats, l'estudi de l'evolució del cromosoma Y en primats i de la relació entre mutacions gèniques i infertilitat apareix com una àrea de recerca fonamental. El tercer treball presentat en aquesta tesi s'engloba dins d'aquest context. Aquest treball ha permès aclarir la història evolutiva dels gens *RPS4Y* en primats. S'ha pogut confirmar que tots els infraordres de primat presenten una còpia del gen *RPS4* en el cromosoma Y, el gen *RPS4Y*. D'altra banda, també ha estat possible determinar que la duplicació del gen *RPS4Y1* que ha generat el gen *RPS4Y2* durant l'evolució dels primats es va produir entre la divergència de les mones del Nou Món i les mones del Vell Món, fa uns 35 milions d'anys. És sabut que aquesta segona còpia lligada a cromosoma Y és funcional en els humans, amb una expressió restringida a testicle i pròstata. Pel que fa a la seva funcionalitat i expressió a la resta de primats encara no hi ha cap informació, tot i que les seqüències de cDNA de ximpanzé i macaco no mostren alteracions en forma de codons de terminació ni canvis en la pauta de lectura, de manera que possiblement també es tracta d'un gen actiu en aquests llinatges. Ara bé, la seqüència del gen *RPS4Y2* en humans ha acumulat una sèrie de mutacions no sinònimes exclusives que, a més, estan afectades per la selecció positiva, mentre que aquesta selecció no és detectada en els altres llinatges.

Com s'ha vist a la Introducció (apartat 4.1.), els cromosomes sexuals provenen d'un parell de cromosomes autosòmics, però, mentre que el cromosoma X ha mantingut la major part del seu contingut gènic, el cromosoma Y ha perdut gran part dels gens, especialment els gens que no són específics de mascle. De fet, en el cromosoma Y la desaparició de gens, tant per dissolució dins el genoma a conseqüència de mutacions com per escissió de material genètic, ha estat un procés constant durant l'evolució. Un exemple d'aquesta pèrdua de gens *house-keeping* és la no presència del gen *RPS4Y* en mamífers. Tanmateix, els primats han mantingut aquesta còpia funcional i, per això, l'origen del gen *RPS4Y* no ha estat fàcil d'elucidar. Inicialment, quan Bergen et al. (1998) van definir aquesta còpia van trobar-la exclusivament en primats i no van ser capaços d'amplificar-la en cap de les altres espècies que van provar, totes elles de rosegadors. Jegalian & Page (1998), però, van trobar una còpia lligada a cromosoma Y en una espècie no primat, la *Monodelphis domestica*. Recentment, quan es van definir les diferents classes de gens presents en el cromosoma Y humà, tant l'*RPS4Y1* com l'*RPS4Y2* es van considerar X-

degenerades, és a dir, que provenien del cromosoma ancestral que va donar lloc als cromosomes sexuals (Skaletsky *et al.*, 2003). Aquesta classificació, però, només té sentit per al gen *RPS4Y1*, ja que en el cromosoma X hi ha un sol gen homòleg. De fet, en diferents publicacions apareix la hipòtesi de la presència d'un gen *RPS4Y* en l'ancestre del cromosoma Y de mamífers, que després hauria desaparegut en tots els llinatges excepte en primats i marsupials (Bachtrog, 2001; Graves, 2002). Alguns autors defensen la presència d'una còpia lligada a cromosoma Y en bovins (Graves, 2006; Graves *et al.*, 2006), segurament basats en seqüències dipositades a GenBank. Tanmateix, si s'analitzen amb atenció les seqüències de *Bos taurus* de les bases de dades públiques es pot arribar a la conclusió que, de fet, es tracta de seqüències que corresponen a la còpia del gen del cromosoma X. La hipòtesi més plausible, doncs, és que el gen *RPS4Y* es trobava inicialment en tots els mamífers, ja que hi ha evidències que posen de manifest que formava part del primer estrat en l'evolució del cromosoma Y (Lahn, Page, 1999), i va ser eliminat, o va degenerar fins desaparèixer, en tots els llinatges excepte en primats i marsupials. La hipòtesi de dos processos independents de nova incorporació gènica des del cromosoma X cap al cromosoma Y que donen el mateix resultat en primats i marsupials és evolutivament poc parsimoniós. Tanmateix, és estrany constatar que en el genoma de tots els primats hi ha com a mínim un pseudogen provinent d'una còpia no funcional de mRNA retrotranscrit i inserit al genoma del gen *RPS4Y*, mentre que al genoma de la resta de mamífers només es poden detectar retrotranscrits del gen *RPS4X*, cosa que indicaria que la inactivació i desaparició d'aquest gen al cromosoma Y va ser molt ràpida.

En mamífers les forces que porten cap a l'acumulació de canvis en el cromosoma Y són, principalment, l'elevada taxa de mutació i la ineficiència de selecció deguda a la no recombinació. El cromosoma Y està subjecte a més canvis que la resta del genoma perquè en totes les generacions les meiosis es produeixen al testicle, mentre que els autosomes i el cromosoma X passen pel testicle només la meitat o un terç de les vegades. El testicle és un ambient hostil ja que l'esperma és un medi oxidatiu que no disposa de mecanismes de reparació i, a més, calen moltes més divisions cel·lulars per fer un espermatozoide que per fer un òvul, és a dir, es produeixen més oportunitats per a l'aparició de dany (Aitken, Graves, 2002). Aquestes característiques podrien explicar l'observació que la majoria de malalties

genètiques dominants que apareixen *de novo* vénen dels cromosomes del pare (Makova, Li, 2002).

Tot i que s'ha fet un gran progrés en l'estudi dels processos moleculars que fan que el cromosoma Y degeneri, encara falta determinar per què es produeix aquesta degeneració. La reducció en l'activitat dels gens per acumulació de mutacions deletèries pot conduir a una pèrdua total de funció i a una desaparició posterior dels gens. L'acumulació de mutacions deletèries en el cromosoma Y, i a llarg terme la completa degeneració, sembla que prové de la interferència entre les mutacions sotmeses a selecció diferencial que s'hi produeixen, que té relació amb la falta de recombinació en aquest cromosoma. De manera simple podríem dir que les poblacions naturals estan subjectes a mutacions recurrents amb efectes beneficiosos i deleteris sobre les que actua la selecció natural per tal d'incorporar les beneficioses i eliminar les deletèries. En absència de recombinació, però, l'acció de la selecció direccional en un locus pot interferir amb l'acció direccional que es produeixi en un altre locus lligat. De manera que l'eficàcia neta de la selecció natural en un cromosoma no recombinant, com és el cas del cromosoma Y, disminueix: les mutacions deletèries s'eliminen de manera menys eficient i les mutacions beneficioses no s'incorporen de forma eficient. Aquesta és la base de la majoria de models genètics que s'han proposat per explicar la degeneració que s'observa al cromosoma Y. Tant els models de selecció purificadora contra les mutacions deletèries com els de fixació de mutacions beneficioses poden donar com a resultat la fixació de mutacions deletèries que observem en el cromosoma Y. Charlesworth (1996) va proposar l'existència d'una acumulació de mutacions lleugerament deletèries per efecte de selecció de fons (*background selection*), de manera que en una determinada població únicament aquells cromosomes Y lliures de mutacions fortament deletèries podrien passar a les futures generacions; en poblacions petites s'incrementaria la probabilitat de fixació de mutacions lleugerament deletèries. Segons la hipòtesi anomenada *Muller's ratchet*, un procés estocàstic provocaria que es perdessin el grup de cromosomes Y sense mutacions deletèries que, per la falta de recombinació, mai més podrien ser restablerts, de manera que el següent millor grup, el que tingués menys mutacions, seria el que es fixaria a la població (Charlesworth, Charlesworth, 1997; Muller, 1964). Un altre model que provocaria acumulació de mutacions deletèries és l'arrossegament selectiu (*genetic hitchhiking*)

produït per mutacions avantatjoses, és a dir, la proliferació dins de la població d'una mutació favorable en el cromosoma Y pot arrossegar fins a la fixació qualsevol al·lel deleteri associat a ella, sempre que el cromosoma mantingui un avantatge net en l'eficàcia biològica (*fitness*). Segons aquest model la degeneració del cromosoma Y és un reflex de l'adaptació en uns loci en concret, a expenses de la majoria dels altres gens del cromosoma (Rice, 1987b). D'altra banda, existeix el model anomenat *Ruby in the rubbish*, que explicaria la manca d'adaptació al cromosoma Y per pèrdua de mutacions avantatjoses en cas que vagin lligades a mutacions deletèries de coeficient de selecció de magnitud igual o superior (Peck, 1994). Així, només aquelles mutacions avantatjoses lliures de mutacions deletèries podran ser seleccionades i retingudes al cromosoma Y per contribuir a l'adaptació (Peck, 1994). El resultat de tots quatre models en l'evolució del cromosoma Y hauria comportat que el més avantatjós fos afavorir i incrementar l'activitat dels gens transcrits en el proto-X (és a dir, desenvolupar la compensació de dosi) mentre que l'activitat de l'homòleg al proto-Y es veuria reduïda. Aquest sembla ser el cas de la parella de gens *RPS4X/RPS4Y* en primats, ja que, del total de proteïna RPS4 present en els complexos ribosomals, només un 15% prové del gen *RPS4Y* (Zinn *et al.*, 1994).

Com s'ha vist al Capítol 3 de Resultats, el gen *RPS4Y2* en humans mostra una acceleració de la taxa de mutació no sinònima i es troba evolutivament sota l'acció de la selecció positiva. El cromosoma Y ha estat molt estudiat per detectar-hi l'acció de la selecció adaptativa en un intent d'explicar l'evolució d'aquest cromosoma. En comparacions entre els genomes d'humà i ximpanzé s'ha pogut determinar l'existència d'un excés de gens afectats per selecció positiva entre els gens que tenen una expressió màxima en testicle (Nielsen *et al.*, 2005). Al mateix temps, els gens relacionats amb espermatogènesi també presenten aquest tret (Nielsen *et al.*, 2005). Aquestes observacions concorden amb les evidències que la selecció positiva ha estat clau en l'evolució d'alguns gens del cromosoma Y en mamífers (Bielawski, Yang, 2001; Dorus *et al.*, 2003; Gerrard, Filatov, 2005; Jobling *et al.*, 1998; Pamilo, O'Neill, 1997). D'altra banda s'ha proposat que la selecció purificadora és menys eficient en els gens del cromosoma Y que en els de l'X (Gerrard, Filatov, 2005). L'acció de la selecció sexual, que ha estat descrita en l'adaptació de diversos caràcters relacionats amb la reproducció, s'ha defensat com a mecanisme fonamental en l'evolució del contingut gènic del cromosoma Y (Roldan, Gomendio,

1999), ja que aquest cromosoma és ric en gens relacionats amb espermatogènesi. Aquestes observacions donen suport a la hipòtesi que les mutacions beneficioses per als mascles però alhora perjudicials per a les femelles s'haurien acumulat en aquest cromosoma, ja que els gens afectats només serien transmesos al fill (on són afavorits) i no a les filles (on haurien estat seleccionats negativament), de manera que els trets que assegurin una bona capacitat reproductiva en els mascles haurien estat sota una forta pressió selectiva (Roldan, Gomendio, 1999).

Es coneix que el gen *RPS4Y2* en humans presenta una expressió restringida a testicle. Aquest patró d'expressió pot ser degut a mutacions en el promotor exclusives també del gen humà. D'altra banda, aquest gen es troba situat a una regió relacionada amb azoospermia, la regió AZFb, és a dir, mutacions en aquesta zona del cromosoma Y provoquen infertilitat. Aquestes característiques fan pensar que aquest gen pot haver adquirit una nova funció en el llinatge humà i que aquesta nova funció possiblement està relacionada amb el procés d'espermatogènesi. En cas de tractar-se d'un gen amb funció espermàtica, la selecció sexual pot haver jugat un paper important sobre el caràcter definit pel nou gen. Tanmateix, per poder confirmar aquesta hipòtesi caldria fer anàlisis bioquímiques i funcionals amb la proteïna RPS4Y2 per comprovar que ja no té funció com a proteïna ribosomal i per analitzar si és essencial per a la fertilitat masculina.

La selecció sexual juga un paper important en l'efecte dels mecanismes d'evolució gènica en determinades espècies, ja que s'ha observat una correlació entre selecció sexual i el tipus de sistema reproductiu. Així, la selecció sexual sembla que exerceix una pressió més forta en aquelles espècies promíscues, amb sistemes de reproducció dispersos o multi-masclle/multi-femella, és a dir, espècies en què la femella pot copular amb diversos mascles i, doncs, els mascles han de competir entre ells, cosa que s'ha observat, per exemple, en diverses espècies de primat (Dixson, Anderson, 2002). La competència entre els mascles es pot donar a diferents nivells, de manera que la selecció sexual pot afectar caràcters molt diversos (Martín-Vivaldi, 2002). En la competència pre-còpula, els mascles competeixen per aconseguir recursos i territori, per exemple marcant el terreny amb exhibicions vocals o visuals, o lluiten directament entre ells per les femelles, utilitzant les "armes" que han desenvolupat, com les banyes. Aquest tipus d'enfrontaments

són comuns en mamífers, de manera que la selecció sexual és intensa sobre caràcters com la mida del cos, l'armament, les exhibicions de força i el comportament agressiu. A més, els mascles també poden desenvolupar trets per atreure l'atenció i la resposta sexual de les femelles, especialment en aus, insectes i peixos, en què es troben evidències d'aquest tipus de competència en el color, sons, productes químics, mida, etc. En molts casos, els dos tipus de competència pre-còpula estan relacionats i podem observar que les femelles mostren preferència pels mascles que demostren ser més dominants. D'altra banda, existeix la competència espermàtica, ja que en moltes de les espècies animals estudiades les femelles tenen relacions sexuals amb més d'un mascle, de manera que, pels mascles, aconseguir aparellar-se no vol dir necessàriament deixar descendència. La competència espermàtica que s'estableix per fecundar l'òvul pot afavorir caràcters ben diferents als afavorits per competència pre-còpula. En un inici es pensava que en mamífers no es donava aquest tipus de competència perquè els espermatozoides estan vius dins el tracte reproductor femení durant un període de temps molt curt, però es va veure que aquesta idea era errònia ja que és comú observar que del mateix part sorgeixen cries de diferents pares o, en espècies amb una sola cria, s'ha vist que aquesta pot provenir d'un mascle diferent al que s'ha observat copulant amb la femella. Finalment, existeix la competència post-còpula exercida pels mascles, que poden desenvolupar diferents estratègies per evitar que la femella deixi descendència d'altres mascles. Un cop acabada la còpula, els mascles d'algunes espècies (per exemple, en macacos i ximpanzés) deixen a la vagina de les femelles un tap copulatori, una massa dura elàstica formada pel semen ejaculat, que evita que altres mascles puguin inseminar-les (tot i que s'han proposat hipòtesis alternatives per a la funció d'aquest tap, com la prevenció d'un possible reflux). D'altra banda, en diverses espècies de mamífers s'ha descrit que els mascles poden matar les cries dels seus competidors, tant amb actes d'infanticidi com provocant l'avortament de la femella prenyada (per exemple, en lleons i rosegadors).

La relació entre selecció sexual i sistema d'aparellament s'ha descrit pels diferents tipus de competència. Aquesta relació pot ser estudiada en els hominoides, ja que presenten una gran diversitat en el comportament social i sexual. Els gibons han estat descrits tradicionalment com a espècies monògames (Brockelman, 1998; tot i que hi ha discrepàncies, veure MacKinnon, 1977), de manera que la competència

espermàtica no és important i aquest fet es veu reflectit en la manca de dimorfisme sexual i en la mida petita dels testicles (Dixson, 1998). En l'extrem oposat es trobarien els ximpanzés i els bonobos, que presenten una elevada competència espermàtica perquè el seu sistema d'aparellament és de tipus multi-masclle/multi-femella (Goodall, 1986) amb molta promiscuïtat. Aquestes espècies són les que tenen els testicles més grossos dins dels primats i són de les que tenen una proporció de la mida testicle-cos més elevada (Harvey, 1984). A més, els ximpanzé són els únics hominoides en què s'ha observat el tap copulatori (Dixson, 1998). Entre aquests dos extrems es troba el sistema social dels goril·les, que viuen en grups petits i estables en què un o dos mascles dominants monopolitzen totes les còpules (Robbins, 1995). Aquesta manca de competència espermàtica es reflecteix en la mida petita dels testicles i de les vesícules seminals d'aquests animals (Dixson, 1998). El comportament social i sexual dels orangutans no està tan ben caracteritzat, tot i que sembla que presenten un sistema d'aparellament dispers que comporta promiscuïtat de les femelles (van Schaik, 1996). Tot i així, la mida petita dels testicles en els mascles d'orangutan comparada amb la gran diferència dimòrfica de la mida del cos no permet aclarir fins a quin punt la competència espermàtica és important (Dixson, 1998). Pel que fa als humans, és més difícil determinar el sistema reproductiu natural com a conseqüència de la gran influència de la cultura.

La selecció sexual s'ha proposat, doncs, com el mecanisme principal en l'evolució dels genitals en moltes espècies, tot i que els mecanismes exactes involucrats no estan tan clars (Hosken, Stockley, 2004). Des d'un punt de vista molecular, ja hi ha estudis que demostren l'acció de la selecció natural en l'evolució de caràcters lligats a reproducció en relació amb el sistema d'aparellament. Un exemple ben clar és el de l'evolució en els primats dels gens que codifiquen per les proteïnes que formen part del semen ejaculat, especialment els gens de la semenogelina, *SEMG1* i *SEMG2*, que és la principal proteïna estructural del semen (Dorus *et al.*, 2004; Jensen-Seaman, Li, 2003). S'ha trobat una correlació positiva de la relació entre taxa de substitució no sinònima i sinònima (ω) del gen *SEMG2* amb la mida relativa dels testicles en diferents espècies de primat estudiades, de manera que la taxa evolutiva correlaciona de manera estreta amb els nivells de poliàndria i, així, les espècies amb femelles més promíscues són les que presenten una ω més elevada (Dorus *et al.*, 2004). L'acció de la selecció positiva ha estat detectada també en set altres

proteïnes relacionades amb la producció de semen en l'evolució dels primats (Clark, Swanson, 2005). En el cas dels gens *RPS4Y*, les evidències suggereixen que després de la duplicació la nova còpia, lliure de constriccions, hauria començat un procés evolutiu de neofuncionalització, possiblement amb la participació de la selecció sexual, com a mínim en humans. L'adquisició d'una nova funció, probablement lligada a espermatogènesi, és, de fet, la hipòtesi més plausible per explicar el manteniment de les dues còpies d'aquest gen en el llinatge humà.

En comparar la seqüència del genoma humà amb la seqüència del genoma del ximpanzé ha sorprès detectar que més de 100 gens presenten interrupcions a la seqüència codificadora específiques de llinatge, en forma de mutacions de codons de terminació o de llocs d'empalmament, de canvis de pauta de lectura per insercions o delecions o d'escissió de gens (Mikkelsen *et al.*, 2005; Wang *et al.*, 2006). Les pèrdues gèniques específiques de llinatge que es produeixen al cromosoma Y són especialment interessants perquè en aquest cromosoma hi ha una acumulació de gens implicats en espermatogènesi (Lahn, Page, 1997; Skaletsky *et al.*, 2003), de manera que el seu estudi pot revelar la història dels canvis evolutius que s'han produït en els sistemes de reproducció i fertilitat entre els diferents llinatges. Com s'ha anat veient, el cromosoma Y, per la seva manca de recombinació gairebé total, és especialment propens a patir pèrdues gèniques per diferents possibles mecanismes. La recent publicació de la seqüència del cromosoma Y del ximpanzé en dos estudis independents (Hughes *et al.*, 2005; Kuroki *et al.*, 2006) ha permès fer anàlisis comparatives entre aquest cromosoma i l'humà. Aquestes comparacions han revelat que dels 16 gens de còpia simple (els X-degenerats, segons la classificació d'Skaletsky *et al.*, 2003) funcionals al cromosoma Y humà, n'hi ha quatre (*CYorf15B*, *TBL1Y*, *TMSB4Y* i *USP9Y*) que han estat inactivats per mutació al llinatge del ximpanzé, mentre que no hi ha evidències de l'esdeveniment oposat (Hughes *et al.*, 2005; Kuroki *et al.*, 2006; Tyler-Smith *et al.*, 2006). Aquests quatre gens contenen set interrupcions en total, que han estat estudiades més en profunditat i s'ha pogut determinar que el goril·la presenta la mateixa seqüència que els humans mentre que el bonobo presenta totes les mutacions disruptives del ximpanzé, excepte la del gen *TBL1Y*, que està, però, present en totes les subespècies de ximpanzé (Perry, 2007). Aquestes observacions, però, no impliquen que no pugui haver-hi altres mutacions disruptives en altres regions d'aquests gens,

o en altres gens, en els llinatges del goril·la i el bonobo. D'altra banda, que un gen tingui una mutació disruptiva no vol dir necessàriament que deixi de ser funcional, i en el cas d'aquests gens s'han pogut detectar RNAs missatgers (mRNAs) per a tres d'ells (Hughes *et al.*, 2005), de manera que aquests mRNAs podrien tenir una funció reguladora o podrien ser traduïts a proteïnes parcials. Tanmateix, no s'han detectat diferències quant a diversitat nucleotídica entre la regió d'abans i la de després de la mutació, com s'esperaria en cas que el gen fos funcional (la d'abans encara estaria subjecta a la selecció, mentre que la de després podria actuar com un pseudogèn) (Perry, 2007).

Les diferències observades entre els llinatges poden intentar explicar-se a través de la possible relació entre l'evolució del cromosoma Y i els sistemes d'aparellament. Les mutacions en els gens del cromosoma Y de ximpanzé podrien ser conseqüència de fenòmens d'arrossegament genètic per l'efecte d'una forta selecció positiva (selecció sexual) en alguna altra regió del cromosoma. En el llinatge que condueix als humans la pressió de la selecció positiva hauria estat menor o bé el fet de danyar els gens podria haver tingut conseqüències molt més deletèries. Així doncs, ja que hi ha un gran enriquiment de gens amb funció espermàtica en el cromosoma Y, sembla raonable pensar que poden haver-se produït molts processos de *selective sweep* durant l'evolució del cromosoma Y dels ximpanzés com a reflex de la importància del paper que juga la selecció sexual en aquesta espècie a causa de la competència espermàtica. D'altra banda, també és possible que algunes mutacions disruptives fossin avantatjoses per elles mateixes, segons la hipòtesi de "menys és més" d'Olson (1999), que va proposar que la pèrdua de la funció gènica durant l'evolució dels homínids ha comportat en alguns casos un benefici en l'eficàcia biològica. De moment, però, hi ha pocs exemples que donin suport a aquesta teoria (Stedman *et al.*, 2004; Wang *et al.*, 2006; Xue *et al.*, 2006) i alguns d'ells han estat molt discutits (McCollum *et al.*, 2006; Perry *et al.*, 2005). En el cas del gen *USP9Y*, que presenta disruptcions en el llinatge del gènere *Pan* (Perry, 2007), s'han trobat dues mutacions disruptives diferents en l'espècie *Ateles geoffroyi* (Gerrard, Filatov, 2005), que també té un sistema reproductiu multi-masclle/multi-femella (Chapman *et al.*, 1995), cosa que fa pensar en la possibilitat que eliminar la funcionalitat d'aquest gen podria ser beneficiós en els primats amb una alta competència espermàtica. De tota manera, no es poden descartar les opcions que aquestes espècies tinguin altres gens no

presentes en els humans per compensar aquesta pèrdua o que en humans el gen *USP9Y* hagi adquirit noves funcions, ja que mutacions en el gen humà comporten infertilitat (Blagosklonova *et al.*, 2000; Sun *et al.*, 1999).

6. Destí del cromosoma Y

A mida que s'obtenen més coneixements sobre l'estructura i la possible evolució del cromosoma Y en mamífers i, més concretament, en primats, van apareixent nous dubtes. Des dels primers estudis moleculars del cromosoma Y, el debat sobre el futur d'aquest cromosoma està obert. Com s'ha anat veient al llarg d'aquest Resum global, l'evolució del cromosoma Y és un procés constant de pèrdua gènica i, alhora, d'acumulació i amplificació de gens. Les evidències actuals suggereixen que el gen *RPS4Y* ha desaparegut en la majoria de llinatges dels mamífers, de manera que és un bon exemple de la pèrdua gènica que es va produint en aquest cromosoma. En canvi, l'aparició del gen *RPS4Y2* durant l'evolució dels primats és un cas de nou material que s'acumula al cromosoma Y, provinent d'una altra regió del mateix cromosoma. Així doncs, la gran pregunta que es plantegen els científics és: degenerarà el cromosoma Y fins a desaparèixer?

En un càlcul groller de la mitjana de la taxa de pèrdua de gens en el cromosoma Y humà, Aitken & Graves (2002) van predir que al cromosoma Y li restaven uns 10 milions d'anys de vida. Recentment, en conèixer-se que el cromosoma X presenta 1.000 gens i no 1.600 com es pensava inicialment, Graves (2006) ha afinat els càlculs i proposa que han de passar uns 14 milions d'anys perquè el cromosoma Y s'extingeixi. Aquestes consideracions han estat molt debatudes. Alguns autors (p.e. Graves, 2005) veuen el cromosoma Y com un element molt dinàmic que pot aparèixer, desaparèixer i reaparèixer en els genomes dels eucariotes. Les observacions fetes en espècies de *Drosophila* donen suport a aquesta hipòtesi. La manca d'homologia entre els cromosomes X i Y en moltes espècies de *Drosophila* ha portat diversos autors (Carvalho, 2002; Hackstein *et al.*, 1996) a suggerir que el cromosoma Y d'aquest gènere no es va originar a partir de l'homòleg del cromosoma X, que segurament va degenerar fins a perdre's, sinó que es va generar a partir d'un cromosoma B (tot i que no es pot descartar la possibilitat que la degeneració fos tan forta que no hagi deixat signes d'homologia). El cromosoma supernumerari B hauria

aconseguit alinear-se amb el cromosoma X, hauria adquirit gens específics de mascle d'altres llocs del genoma i hauria donat lloc a un cromosoma Y ancestral, que seria el que ha estat retingut en la majoria d'espècies de *Drosophila* actuals, com *D. melanogaster*. Tanmateix, el cromosoma Y de les espècies del llinatge de *D. pseudoobscura* no és homòleg al de *D. melanogaster*, de fet, no té res a veure amb el cromosoma Y ancestral, de manera que es tracta d'una generació *de novo* que va tenir lloc en el llinatge del grup *pseudoobscura*, quan fa uns 18 milions d'anys el cromosoma Y ancestral es va translocar cap a un autosoma i va aparèixer un cromosoma Y completament nou (Carvalho, Clark, 2005). A més, es pot observar una reducció de l'heterocromatina i de la mida dels introns en els gens que havien estat lligats a Y i que van passar a ser autosòmics (Carvalho, Clark, 2005). D'altra banda, l'espècie *D. miranda* presenta un nou cromosoma Y molt jove, que es va formar fa només aproximadament 1 milió d'anys per la fusió del cromosoma Y ancestral amb un autosoma, cosa que va fer que uns 2.000 gens esdevinguessin lligats a sexe. S'ha pogut observar que les regions genòmiques del nou X gairebé no presenten canvis, mentre que la regió neo-Y, que mostra els primers estadis de l'evolució d'un cromosoma Y, ja conté signes clars de degeneració (mutacions que introdueixen codons de terminació o canvis en la pauta de lectura) i acumulació d'elements transponibles, delecions i reorganitzacions estructurals (Bachtrog, 2005).

El procés de degeneració del cromosoma Y fins a la seva completa desaparició no només se sosté en les observacions fetes a *Drosophila*, sinó que també es veu reforçat per descobriments fets en mamífers. S'han descrit dos gèneres de rosegadors que inclouen espècies que no tenen cromosoma Y, la rateta-talp (*Ellobius*) (Fredga, 1988; Just *et al.*, 1995) i la rata de camp espinosa japonesa (*Tokudaia*). L'espècie *Ellobius lutescens* té un cariotip XO en tots dos sexes, com la rata de camp espinosa japonesa (Arakawa *et al.*, 2002), mentre que l'espècie *E. tancrei* és XX pels dos sexes (Vogel *et al.*, 1998). La rateta-talp no presenta cap traça del gen *Sry* ni de cap altra marca de cromosoma Y, de manera que el sistema de determinació del sexe ha canviat completament i els factors de fertilitat del cromosoma Y també han estat substituïts, però el nou sistema encara és del tot desconegut (Just *et al.*, 2002). En el cas del gènere *Tokudaia*, tampoc s'han detectat restes de *Sry*, però sí que sembla que hi ha presència en tots dos sexes de, com a mínim, dos gens (*Zfy* i *Tspy*) provinents del cromosoma Y, de manera que sembla

que una regió del cromosoma Y s'ha translocat cap a l'X o cap a un autosoma. Fins i tot dins dels primats, alguns autors han descrit cariotips sexuals de tipus XX/XO com a conseqüència de la translocació del cromosoma Y cap a un dels autosomes en espècies del gènere *Aotus* (*A. infulatus*, *A. azarae* i *A. nigriceps*) (Ma, 1981; Ma *et al.*, 1975; Mudry *et al.*, 1998; Pieczarka, Nagamachi, 1988). Aquestes espècies de mamífers, doncs, han estat capaces d'existir sense cromosoma Y, però la manera com ho aconsegueixen encara avui en dia és un enigma i no s'ha pogut aclarir si el nou sistema de determinació del sexe depèn de l'acció d'un nou gen en un determinat cromosoma (i a partir d'aquí potser es començaria a desenvolupar un nou proto-Y) o bé si s'ha passat a un sistema dependent de l'ambient, per exemple de la temperatura, com el d'alguns llinatges de rèptils (Janzen, Phillips, 2006).

Podria ser, doncs, que la desaparició total del cromosoma Y fos l'últim estadi en l'evolució d'aquest cromosoma, però diferents autors opinen que aquesta desaparició portaria a l'espècie cap a l'extinció pel paper essencial que juga el cromosoma Y en la determinació del sexe (Aitken, Graves, 2002; Sykes, 2003). Tanmateix, aquesta visió més apocalíptica del destí del cromosoma Y sembla poc realista. Com s'ha anat veient, la desaparició del cromosoma Y no implica necessàriament l'extinció de l'espècie, sinó que hi ha diferents alternatives. D'altra banda, molts autors defensen el manteniment d'aquest cromosoma dins del genoma gràcies a diversos mecanismes. L'eficàcia dels diferents processos que causen degeneració genètica depèn del número de gens lligats en el cromosoma Y (Charlesworth, Charlesworth, 2000), de manera que amb el número tan reduït de gens funcionals que han sobreviscut en el cromosoma Y modern dels mamífers (Skaletsky *et al.*, 2003) la taxa de mutació deletèria segurament és massa baixa perquè pugui actuar la selecció de fons, de la mateixa manera que segurament les mutacions avantatjoses també són massa escasses perquè els *selective sweeps* puguin tenir un paper important en l'acumulació de mutacions deletèries (Gerrard, Filatov, 2005). El fet que la diversitat genètica en el cromosoma Y humà sigui només lleugerament inferior a la del cromosoma X concorda amb la hipòtesi que després de 230-320 milions d'anys de degeneració aquest cromosoma conté un número massa baix de gens com perquè els processos de degeneració puguin continuar més enllà (Lahn, Page, 1999).

D'altra banda, la manca de recombinació es veu com un dels factors més influents en la degeneració del cromosoma Y, però la recombinació no és l'únic mecanisme de redistribució genòmica existent. En el cromosoma Y es dona un procés continu d'esdeveniments de conversió gènica, recombinació homòloga intracromosòmica, inversió i deleció que, junt amb l'acció de la selecció purificadora –que s'ha vist que en humans té un efecte més important del que es pensava inicialment (Hughes *et al.*, 2005)–, podrien tenir suficient potència per evitar la desaparició del cromosoma Y (de Knijff, 2006). Aquesta visió donaria un sentit a la natura repetitiva de l'estructura del cromosoma Y, ja que els elements repetitius serien essencials per al manteniment del cromosoma. A més, les famílies multigèniques del cromosoma Y, totes formades per gens d'expressió específica de testicle, és a dir, probablement fonamentals en espermatogènesi, mantenen la seqüència dels gens conservada gràcies, en part, al sistema de naixement i mort de gens. A part d'aquestes multiplicacions en grans famílies gèniques, es donen també noves aportacions per duplicació i transposició de gens des d'altres llocs del genoma, com en el cas dels gens *TGIF2LY* i *PCDH11Y* en humans (Skaletsky *et al.*, 2003), el gen *RPS4Y2* dins la filogènia dels primats (tercer capítol d'aquesta tesi) o nous gens detectats en carnívors (Murphy *et al.*, 2006).

Com s'ha anat veient, el cromosoma Y ha despertat un gran interès entre els científics i ha estat la diana de molts estudis, però, tot i que en els darrers anys s'ha avançat en l'obtenció de coneixements, aquest cromosoma encara ara és un misteri en molts aspectes. La història evolutiva del gen *RPS4Y* al llarg de la radiació dels mamífers, per exemple, encara s'ha d'acabar de completar. Per aclarir l'escenari evolutiu d'aquet gen és imprescindible disposar de la seqüència del cromosoma Y de diferents espècies de mamífers que cobreixin tots els ordres i així poder determinar amb exactitud en quins llinatges s'ha mantingut realment aquest gen. La manca d'informació sobre el cromosoma Y prové de les enormes dificultats que comporta el seu estudi, especialment per la gran quantitat d'elements repetitius i regions repetides que formen aquest cromosoma. Els projectes de seqüenciació de genomes complets mostren clarament aquesta dificultat, ja que el cromosoma Y sempre és l'última peça que es seqüencia i s'aconsegueix després d'un esforç especial. Així, recentment s'ha publicat l'esbós de la seqüència del genoma del ximpanzé (Mikkelsen *et al.*, 2005) i, a part, la seqüència refinada d'una part del

cromosoma Y d'aquesta espècie (de la regió no repetitiva), que ha requerit un esforç específic dut a terme per dos grups en paral·lel (Hughes *et al.*, 2005; Kuroki *et al.*, 2006), de la mateixa manera que primer es va publicar l'esbós de la seqüència del genoma humà l'any 2001 (Lander *et al.*, 2001; Venter *et al.*, 2001), el genoma ja més ben organitzat a l'abril del 2003 i pocs mesos després la seqüència del cromosoma Y (Skaletsky *et al.*, 2003). Així, dels vuit genomes de mamífers que han estat seqüenciats i que ja s'han aconseguit organitzar, només es disposa del cromosoma Y de tres espècies: l'home, que presenta el cromosoma Y complet, i el ximpanzé i el ratolí, dels que es disposa de la seqüència parcial. Davant d'aquesta situació, el *National Human Genome Research Institute* (NHGRI) ha establert com una de les prioritats de seqüenciació el refinament de la seqüència del cromosoma Y de cinc mamífers –*Bos taurus*, *Rattus norvegicus*, *Monodelphis domestica*, *Macaca mulatta* i *Callithrix jacchus*–, entre ells dues espècies de primat (les dades són del juny del 2006). El NHGRI s'ha adonat de la importància de la genòmica comparada dins la recerca biològica i ha apostat per la seqüenciació de molts organismes diferents. Aquest centre permet que els investigadors facin propostes sobre les espècies que creuen que caldria seqüenciar a través del sistema dels *White Paper Proposal*, articles que han de justificar el suggeriment. Rozen *et al.* (2006) han fet una proposta (*White Paper*) de seqüenciació del cromosoma Y en set espècies de mamífers, ja que afegeixen a les cinc ja enumerades el gat (*Felis catus*) i el gos (*Canis familiaris*).

Actualment ja es disposa de la seqüència del genoma complet de dos primats no humans, el ximpanzé i el macaco (tot i que la d'aquest darrer encara s'ha de refinar), i hi ha propostes acceptades per seqüenciar set espècies més, algunes de les quals ja es troben ja en progrés o estan a punt de començar a ser seqüenciades, i hi ha dues espècies més recomanades. D'altra banda, també s'ha acceptat seqüenciar de manera refinada el cromosoma Y de macaco i del tití comú. Això denota la importància dels primats en la genòmica comparada, ja que són la clau per comprendre qüestions de biologia evolutiva i com a models en biomedicina per trobar les causes de malalties que afecten els humans. En aquest context, l'estudi comparat del cromosoma Y en primats, especialment en els gran simis, i en mamífers en general, és essencial per revelar qüestions crítiques sobre la infertilitat en els homes i altres aspectes que tenen la base genètica en aquest cromosoma.

Com s'ha vist, per poder aprofundir en la relació entre selecció sexual i sistemes d'aparellament cal que puguem comparar les seqüències dels gens afectats, i preferiblement de tot el cromosoma Y, d'espècies que representin tots els exemples de sistemes reproductius, de manera que seria molt interessant poder disposar de les seqüències completes de cromosoma Y de tots els hominoides.

7. Genètica de la conservació i conservació de les espècies. Futur

La genètica de la conservació ha sobreviscut a tres grans reptes en l'intent d'ajudar en les decisions de gestió de la conservació (DeSalle, Amato, 2004). El primer repte va aparèixer en relació a les definicions d'espècie i de fronteres i en el rol que la genètica havia de jugar en la delimitació de les unitats de conservació (Ryder, 1986). El segon repte va venir de la crítica de Caughley (1994), que deia que es donava massa importància a les qüestions tècniques en conservació (incloent la genètica de la conservació) en detriment de qüestions que considerava més importants, com l'amenaça de l'hàbitat i les malalties. Finalment, el tercer repte és un dels aspectes més controvertits de la biologia de la conservació: quin és el rol real que juguen els factors genètics en l'extinció i el perill de la vida salvatge (DeSalle, 2005). Aquest dubte sobre la importància dels factors genètics en el procés d'extinció de les espècies va començar amb una publicació de Lande (1988) que suggeria que els efectes dels factors demogràfics sobre les espècies eren molt més importants que els factors genètics i que, de fet, les espècies s'extingien abans de poder detectar canvis en la diversitat genètica. Spielman et al. (2004) van decidir fer un estudi global per determinar si en les espècies amenaçades es podia detectar una reducció de la diversitat genètica en comparació amb altres espècies relacionades no amenaçades i van poder comprovar que sí que hi havia diferències, de manera que la pèrdua de variabilitat genètica sí pot ser un dels factors clau en l'evolució de les poblacions en perill d'extinció. D'altra banda, quan es tracta amb espècies amenaçades els factors com la destrucció de l'hàbitat, la fragmentació, la pol·lució, la sobreexplotació i les espècies exòtiques introduïdes es perceben fàcilment com a qüestions crítiques directes i és fàcil, doncs, plantejar-se si el deteriorament genètic és realment un factor a tenir en compte i si és o no prioritari. De fet, fins fa pocs anys els problemes genètics es veien simplement com una possibilitat teòrica. Actualment, però, l'empobriment genètic es percep com un problema real per a

moltes espècies amenaçades i en perill ja que la resposta d'una població genèticament poc diversa front al canvi és pitjor que la de poblacions diverses (Frankham, 2002), de manera que la genètica cada vegada es té més en compte en conservació per ajudar en les decisions de gestió de les espècies. A més, s'estan desenvolupant eines genètiques i computacionals per contribuir al control de les poblacions que cal gestionar i conservar (Schwartz *et al.*, 2007), des de l'estadi inicial de definició de les unitats de maneig i estudi de la diversitat d'una població en concret fins al seguiment a llarg terme de la dinàmica de la població, a més d'analitzar la possibilitat de portar a terme programes de reintroducció o qualsevol altre tipus de translocació.

L'objectiu final d'un procés de conservació és aconseguir que l'espècie en concret sigui suficientment abundant i ben estesa en llibertat i conservi prou variabilitat genètica com per poder mantenir la seva funcionalitat, garantir la seva capacitat d'adaptació als canvis ambientals i conservar les seves potencialitats evolutives sense necessitat d'ajuda d'origen humà. Per maximitzar el potencial evolutiu d'una espècie i evitar el perill d'extinció a llarg terme és essencial mantenir un nombre elevat de poblacions diferents i no protegir una sola població. Al llarg d'aquest Resum global s'ha anat veient que els grans avenços que es produeixen en genòmica, bioinformàtica, tecnologia d'alt rendiment, etc., fan avançar molt de pressa la biologia de la conservació i milloren les possibilitats de protegir les espècies. Tanmateix, a part de tots els progressos que s'estan duent a terme en genètica, no es pot oblidar que la conservació de les espècies no té cap sentit si no es preserven també els ecosistemes, tal com apuntava Caughley (1994). Cal que els hàbitats de les espècies siguin prou extensos i connectats entre ells per permetre el desplaçament i l'adaptació en cas de canvi. No té sentit salvar una espècie si el seu hàbitat desapareix, de manera que els plans de conservació han de tenir en compte molts aspectes diferents que cal potenciar. Entendre la dinàmica dels ecosistemes és fonamental per poder dur a terme plans per a la seva protecció. En molts casos, protegir un ecosistema significa protegir totes les espècies que hi viuen, tot i que les més amenaçades potser requereixen esforços específics.

La gran dificultat de protegir els hàbitats és la seva distribució dins del planeta, ja que les àrees amb més biodiversitat es troben en països en vies de

desenvolupament mancats de molts recursos i amb altres prioritats, tal com vèiem a la Introducció d'aquesta tesi pels primats (apartat 1.3). En un estudi fet a Àfrica (Rondinini *et al.*, 2006) s'han identificat els llocs prioritaris que cal protegir per actuar sobre la conservació dels vertebrats africans, molts dels quals estan amenaçats, i s'ha determinat que més de la meitat dels llocs irremplaçables són regions amb una elevada densitat humana. En molts casos, la destrucció dels boscos no té cap sentit econòmic, però a curt terme aporta ajuda immediata, de manera que una part crucial de la solució es troba en l'educació, és essencial fer entendre a les poblacions humanes que és molt més valuós a llarg terme conservar els hàbitats que destruir-los. Així, la creació de reserves és una estratègia necessària, però és fonamental trobar la manera de conservar els ecosistemes sense que això tingui un efecte negatiu per a les poblacions humanes que també hi habiten. Hi ha exemples encoratjadors que han tingut èxit combinant les necessitats dels animals i les necessitats dels habitants de la regió, com és el cas del turisme de goril·les a Rwanda, on aquesta activitat s'ha convertit en la segona gran font d'ingressos del país i, tot i que humans i goril·les entren en contacte i hi ha hagut milers de visitants, no s'ha produït cap incident. En conservació, doncs, cal considerar sempre tant factors ecològics com socioeconòmics de cada regió en concret perquè els plans de gestió que s'hi estableixin tinguin èxit (Balmford *et al.*, 2002; Balvanera *et al.*, 2001). Per exemple, en el cas dels primats no humans, per poder evitar el comerç de carn és necessari que els països que en són consumidors trobin fonts de proteïna alternatives que permetin sostenir la població humana sense haver de caçar primats (Bowen-Jones, Pendry, 1999). És important també no oblidar que els beneficis de conservar els ecosistemes són internacionals, però que els costos són per als països on es troben els boscos, que, a més, són els països més pobres del planeta (World Parks Congress a Durban, Sud Àfrica, 2003). És per això que cal que els països desenvolupats contribueixin d'alguna manera, per exemple condonant el deute exterior d'un país si s'hi estableixen reserves naturals.

La recuperació d'espècies amenaçades és un procés, doncs, molt complex en què intervenen tant factors biològics i ecològics com factors de caràcter social, polític i econòmic. A més, els recursos són finits, de manera que no és possible protegir-ho tot, ni té sentit seguir la tàctica de protegir per protegir, sinó que cal definir amb criteri unes prioritats bàsiques. Fins que els objectius científics i polítics, amb la societat

entremig, no arribin a integrar-se, la majoria d'extincions no es podran evitar. És evident, doncs, que la conservació té dues vessants clares: per una banda, els científics han de detectar el problema, estudiar-lo i suggerir solucions; per una altra banda, l'aplicació pràctica de les mesures i el seu èxit depenen d'elements sociopolítics. Davant d'aquesta situació, a l'hora de triar les espècies que es volen recuperar cal que els científics determinin quines són les necessitats més immediates per protegir la biosfera i que tinguin en compte tots els factors que intervenen, per exemple, quin és el rol que juguen les diferents espècies dins la societat i dins l'ecosistema (veure apartat 2.1. de la Introducció).

Els primats no humans sempre han causat fascinació en l'home, ja que presenten característiques que els fan excepcionals –són un grup increïblement divers d'animals molt intel·ligents, que inclou més de 600 tàxons amb el rang de mides corporals més ampli de tots els mamífers. Aquesta atracció s'ha intentat utilitzar per potenciar la conservació dels boscos tropicals, fent servir els primats com a “espècies senyera” per atreure l'atenció de la població cap a la necessitat de protegir els seus hàbitats. Com s'ha vist, l'ordre dels primats inclou l'home i els tàxons vius més propers a la nostra espècie i aquest fet ha conduït al gran interès que actualment existeix a estudiar els primats no humans, tant a nivell genètic com de comportament, per tal d'entendre l'evolució humana i per buscar l'origen genètic de determinades malalties, per exemple. S'han invertit molts esforços en estudis de genòmica comparada de primats per tal de determinar què és allò que ens fa humans. Tot i que aquests estudis permetran ampliar el coneixement sobre l'home, no només s'ha de pensar des d'un punt de vista antropocèntric, sinó que comprendre la diversitat genètica dins de l'ordre dels primats mostrarà com ha canviat el genoma dins d'aquest llinatge al llarg del temps i revelarà, per exemple, com cadascuna de les diferents espècies s'ha adaptat al seu hàbitat, i com el comportament i el sistema social correlacionen amb l'estructura genètica de les poblacions. Tota aquesta informació, potser, ajudarà en la tasca de la conservació d'aquests animals. En aquests moments, més d'un terç dels primats es troba amenaçat i la situació és extrema per a algunes de les espècies, principalment com a conseqüència de la gran desforestació que pateixen les selves tropicals, de la persecució de què són víctimes per vendre'n la carn o els infants com a mascotes i de les noves malalties transmeses pels humans. Com s'ha anat veient al llarg

d'aquesta tesi, la genètica de la conservació pot ajudar a determinar la situació de diversitat en què es troben les espècies i a definir les unitats de maneig, a més de ser essencial en plans de reintroducció o de reproducció. No obstant, tots els esforços són en va si no es troben integrats dins d'un procés global de protecció dels ecosistemes. Si la situació segueix com fins ara, els primats continuaran estant en perill i la seva situació serà cada vegada més crítica, fins arribar a un procés d'extinció global.

