

TESI DOCTORAL
UNIVERSITAT DE BARCELONA

FACULTAT DE BIOLOGIA
DEPARTAMENT DE GENÈTICA
Programa de doctorat bienni 2005-2007

**Estudi de la recombinació en *D. subobscura* a
nivell genòmic**

Memòria presentada per Cinta Pegueroles Queralt, realitzada al Departament de Genètica, per accedir al títol de Doctor per la Universitat de Barcelona, sota la direcció dels doctors Marta Pascual Berniola i Francesc Mestres Naval.

Cinta Pegueroles Queralt
Barcelona, Març de 2010

VIST-I-PLAU
LA DIRECTORA DE LA TESI
Dra. Marta Pascual Berniola
Professora Titular de Universitat
Facultat de Biologia

VIST-I-PLAU
EL DIRECTOR DE LA TESI
Dr. Francesc Mestres Naval
Professor Titular de Universitat
Facultat de Biologia

Des de ben petita tenia clar què volia ser de gran: volia ser zoòloga i viatjar per les selves i oceans de tot el món per estudiar els animals. Suposo que els estius que vaig passar a “l’Ampolla”, a la vora del mar i rodejada d’oliveres, animals (sobretot, peixos, insectes, gats i demes bèsties) i el meu germà (un animalet més) van tenir molt a veure. O potser van ser les excursions per anar a buscar “pedretes” en família, o les acampades al Mas de la Franqueta, on els melons es refresquen al riu i a més es pot visitar la “pista d’aterratge d’ovnis”. Segur que tot això hi va tenir molt a veure...

Vaig anar creixent, i els gustos i la manera de vestir van anar canviant (per sort), però seguia tenint clar que jo volia ser biòloga, tot i que la majoria dels meus amics eren de lletres: l’Aina, l’Ari, la Laia, l’Alba, la Tere, la Clara, el Jordi... Però no va ser fins a COU, gràcies al meu professor Antonio Arregui, que vaig saber que el que més m’apassionava era la genètica. Però per entrar a la universitat primer s’havia de fer un examen, la selectivitat. Sabia que “la sele” era un examen molt temut, però si havia passat els quatre anys d’institut sense estudiar més d’una tarda per examen, no veia perquè aquell havia de ser diferent. La veritat es que ho vaig entendre de cop. Però no hi ha mal que per bé no vingui, i a la segona la “sele” va anar molt millor! Després d’una visita fugaç per la facultat de Química de la UAB, vaig confirmar encara més les ganes d’estudiar biologia. A més, remenant per la web vaig trobar un curs interessant: I tu? Jo, bioquímica, organitzat per Josep M. Fernández-Novell. La veritat és que va estar molt bé i animo a que el segueixin organitzant durant molts anys més. Per primer cop vaig entrar una mica en contacte amb el món de la biologia molecular i la recerca, i vaig fer la meva primera PCR!

Finalment, vaig entrar a estudiar biologia. La carrera em va encantar. La veritat és que tinc el record de molts bons professors que em van fer gaudir amb matèries molt diverses. També vaig fer un grapat de bons amics: Andrea, Nat, Neus, Sara, Vicki, Victor, Dani, Rosa, Amanda, Jordi... que a més de fer-me costat i riure, durant la carrera entre tots em recordàveu quan tenia pràctiques! (ja se sap que els horaris de biologia són complicats i una és una mica desastrilla, jejeje). La carrera estava molt bé, però les pràctiques eren poques i massificades, i jo tenia “mono” de laboratori... Així que durant dos estius vaig fer estades a l’Hospital Clínic, als departaments de Bioquímica i Microbiologia. Moltes gràcies Ramon Deulofeu i Jordi Costa per haver-me acollit! I

molt especialment dono les gràcies a la Teresa, ja que amb ella vaig aprendre tot el que necessitava saber per treballar en un laboratori.

Els anys van passar molt ràpid i finalment va arribar l'últim any de carrera. Jo tenia ganes de col·laborar en algun laboratori... Després de fer un parell d'incursions fallides, vaig anar a parar amb qui serien els meus jefes i companys d'aventures i desventures durant quatre anys i pico: el Francesc (alies "El maestro") i la Marta (alies "l'experta en microsatèl·lits"). Des del primer dia que vam parlar, vam connectar, així que no va ser gens difícil de prendre la decisió de quedar-me. Treballar allí m'ha permès conèixer a molt bona gent. Al principi érem poquets, estava el Lluís (alies "El gran jefe"), al que sempre he anat a demanar consell quan més ho necessitava, el Joan, que sempre ha estat a punt per solucionar-me qualsevol problema relacionat amb *D. subobscura* i inversions, el Pedro (alies "Pedrito"), que sempre estava allí per explicar-me les seves històries, animar-me i ajudar quan feia falta, l'Hèctor, tot i anar molt atrafegat sempre tenia temps per col·laborar, fer-nos riure amb els seus acudits i preparar alguna bona paella i el Ferran, que és el meu germanet gran de laboratori, el meu okupa preferit i un gran amic. Després la família va anar creixent, vaig conèixer la Josiane, la meva okupa femenina preferida, el Pedro (alies "el Argentino") i la Isabelle, que van fer estades curtes però molt intenses, el Víctor H, el nostre proveïdor de caramels de cafè colombià amb llet, la Gemma, la meva companya de captures sempre a punt per riure una estona, la Celia, la nostra "danesa" i organitzadora d'excursions preferida, el Víctor, sempre molt atent i amable amb tothom, el Pedro (alies "el portugués"), que com a bon postdoc sempre està a punt de resoldre els nostres dubtes, el João, un altre típic portuguès de metro noranta amb moltes inquietuds i l'Albert, un boníssim ajudant i company, sempre curiós i amb ganes d'aprendre noves coses. També he conegut a gent d'altres laboratoris, l'Edu, amb qui he passat molts bons moments, la Isa, sempre molt amable i col·laboradora, el Marc, un invasor marí molt hospitalari, l'Enrique, un molt bon revisor, el Freddy i la Montse, els nostres nous companys... i en general a tot el departament de genètica. També he d'agrair molt especialment l'ajuda de l'Empar i l'Ana, que a més de fer la seva feina de manera impecable, sempre transmeten alegria i bon humor. A més, sempre m'heu solucionat la vida quan m'oblidava de demanar medi... I al nostre equip de secretàries, que ens informen de tots els tràmits. També he d'agrair molt l'ajuda de la Conxita...sense els teus coneixements d'estadística alguns anàlisis haguessin estat un calvari! Per últim, gràcies a aquest

doctorat he conegut a tot l'equip de biologia evolutiva de la UAB: a l'Antonio, el Mauro, la Pili, el Tomàs, la Montse, el Francisco, la Blanca, la Luz, la Doris... De veritat, moltes gràcies a tots!

Un doctorat no serveix només per aprendre a fer ciència, implica moltes més coses! A mi m'ha permès anar a viure amb el Chus, conèixer moltíssima gent, aprendre anglès, anar de congressos (no m'enrotllaré més perquè ja vaig per la tercera pàgina però es podria fer tota una dissertació)...i viatjar! Decididament les estades de doctorat són un gran invent... Vaig fer una estada curteta a Viena, però molt intensa. Tot i que els experiments no van sortir tot lo bé que esperàvem, el prof. C. Schlötterer, la Viola Nolte i la resta de l'equip van ser molt acollidors. Però les meves dues grans estades han estat a Ithaca... no la grega, la nordamericana. Ithaca és una petita ciutat situada a la zona dels Finger Lakes, a l'estat de Nova York. No se si és perquè es troba envoltada de deus, cascades, natura i animals, o si perquè molt propet d'allí l'any 1969 s'hi va celebrar el festival de Woodstock, però a Ithaca s'hi respira un ambient especial. I com no podia ser d'una altra manera, allí vaig conèixer a una gent meravellosa: la Nadia, el Nathan, la Naty, la Heather, el Carl, el Zhen, la Vanessa, el Frank, el Mike, la Martha, l'Andy, la Wendy i sobretot el Chip, una gran persona i un boníssim divulgador i investigador.

Hi ha molta més gent amb la que he passat bons moments i que encara no he citat: el Mac i les seves paraules d'ànim, el Pere i les captures de crancs, la Mar i l'Alhambra de Granada, la Inés excompanya de CAP... Moltes gràcies també al Ramon, l'Amaia, l'Agustí i la Berta dels Serveis Científicotècnics, per la vostra feina tan ràpida i ben feta. I finalment gràcies senyor Agaur per totes les beques que m'heu atorgat! En especial a en Bernat Xancó i la Montserrat Boada, que sempre m'han atès molt bé. A més de la beca predoctoral de la Generalitat de Catalunya 2009FIC-00096, aquesta tesi també s'ha finançat amb els projectes del Ministerio de Ciencia y Tecnología CGL2006-13423-C02-02 i de la Generalitat de Catalunya SGR2009-636.

En aquesta recta final, he d'agrair els ànims i les correccions del Francesc i la Marta, indispensables per la finalització d'aquesta tesi. I sobretot moltes gràcies a tu, Chus, pel teu suport moral i logístic que tant he necessitat en aquesta última etapa.

Moltes gràcies a tots!

Nothing in Biology Makes Sense Except in the Light of Evolution

Theodosius Dobzhansky, 1973

It is natural selection that gives direction to changes, orients chance, and slowly, progressively produces more complex structures, new organs, and new species. Novelties come from previously unseen association of old material. To create is to recombine.

François Jacob, 1977

Per a vatros

CONTINGUTS

Agraïments

Continguts

1. Introducció general	1
1. 1. La recombinació és una avantatge evolutiu	3
1. 2. La recombinació i els mecanismes que l'originen	4
1. 3. Els mapes genètics o de lligament	6
1. 4. La distribució de la recombinació al llarg del cromosoma	7
1. 5. Les inversions cromosòmiques com a model genètic	8
1. 6. Les inversions cromosòmiques i la recombinació	10
1. 7. El manteniment de les inversions a les poblacions naturals	13
1. 8. <i>Drosophila subobscura</i> com a organisme model	15
1. 9. Primeres observacions i estudis de la recombinació a <i>D. subobscura</i>	18
1. 10. El polimorfisme per inversions en <i>D. subobscura</i> és adaptatiu	21
1. 11. Estudi de les inversions de <i>D. subobscura</i> a nivell molecular	22
2. 1. Objectius	27
2. 2. Estructura de la tesi	29
3. Publicacions	
3. 1. Determinació de la recombinació al llarg del cromosoma O de <i>Drosophila subobscura</i> i elaboració d'un mapa genètic.....	31
<u>Publication 1</u> : A recombination survey using microsatellites: the O chromosome of <i>Drosophila subobscura</i>	33
3. 2. Estudi de la influència de les inversions en la formació d'entrecreuaments al llarg del cromosoma.	55
<u>Publication 2</u> : Recombination and selection in the maintenance of the adaptive value of inversions	57
3. 3. Diferenciació genètica entre poblacions naturals de <i>D.</i> <i>subobscura</i>.	79
<u>Publication 3</u> : Gene flow and gene flux shape evolutionary patterns of variation in <i>D. subobscura</i>	81
<u>Publication 4</u> : Recent origin of a chromosomal inversion revealed by its association with nucleotide haplotypes	107

3. 4. Flux gènic entre ordenaments de <i>D. subobscura</i>.	115
<u>Publication 5</u> : The effect of recombination on medium-sized inversions: selection and gene flux.	117
4. Discussió	149
4. 1. Mapa genètic del cromosoma O de <i>D. subobscura</i>	151
4. 2. La recombinació al cromosoma O de <i>Drosophila subobscura</i> , a nivell general i regional.	154
4. 3. Les inversions regulen la formació d'entrecreuaments al llarg del cromosoma.	155
4. 4. Patrons selectius relacionats amb inversions.	157
4. 5. Estudi del contingut genètic de les inversions en diferents poblacions. ..	160
4. 6. Estudi del flux gènic entre ordenaments.	161
4. 7. Variabilitat genètica vs distància al punt de trencament.	165
4. 8. Patrons de selecció	166
4. 9. Datació dels ordenaments cromosòmics.	167
5. Conclusions	169
6. Bibliografia	173
Informe dels directors de la tesi	187
I. Supplementary figures publication 1	189
II. Supplementary figures publication 3	190
III. Supplementary figures publication 5	201

1. Introducció general

1. INTRODUCCIÓ

La recombinació és un mecanisme biològic que permet l'obtenció de noves combinacions genètiques. L'avantatge principal que representa és l'increment de la variabilitat genètica dels organismes, conferint noves possibilitats d'adaptació a diferents ambients. S'ha de destacar la universalitat d'aquest mecanisme, que s'ha descrit en virus, bacteries i organismes superiors. En els eucariotes la recombinació genètica normalment es produeix durant la meiosi, com a resultat de l'entrecruament de cromosomes homòlegs aparellats. En organismes sense reproducció sexual s'ha descrit la recombinació mitòtica, on la cèl·lula diploide resultant presenta combinacions al·lèliques diferents de la parental. En organismes amb reproducció sexual també s'ha observat de manera esporàdica la recombinació mitòtica en cèl·lules somàtiques. La recombinació genètica es va començar a estudiar a principis del segle XX (STURTEVANT 1913), però no va ser fins els anys 90 del segle passat quan es van realitzar els primers estudis on es relacionava la taxa de recombinació amb la diversitat nucleotídica (BEGUN and AQUADRO 1992). Es va poder observar que la recombinació és un factor clau modulant els nivells de polimorfisme nucleotídic, donat que explica aproximadament un quart de la variabilitat observada entre gens (NAVARRO *et al.* 1997).

1. 1. La recombinació és un avantatge evolutiu

La recombinació meiòtica és un mecanisme biològic present en la reproducció sexual de tots els organismes eucariòtics a causa dels seus efectes positius en l'evolució dels organismes. A principis del segle XX es va observar que la recombinació incrementa la taxa d'adaptació, permetent que mutacions beneficioses que inicialment es trobaven en diferents cromosomes, es poguessin combinar en un mateix cromosoma (MORGAN 1913; FISHER 1930; MULLER 1932). Principalment, la recombinació actua trencant el desequilibri de lligament creat per la deriva genètica. A partir d'aquesta premissa, s'han estudiat una sèrie de models on s'observa que el desequilibri de lligament disminueix la variabilitat i, en conseqüència, la taxa d'adaptació (revisat en FELSENSTEIN 1974). Segons el Trinquet de Muller (*Muller's ratchet*) les mutacions deletèries s'acumulen més fàcilment en una població sense recombinació, com a conseqüència de la deriva i el

l·ligament (MULLER 1964). Per altra banda, l'efecte de Hill-Robertson descriu que, en poblacions finites, el desequilibri de l·ligament entre llocs seleccionats reduirà l'efectivitat global de la selecció (HILL and ROBERTSON 1966). La recombinació trenca el desequilibri, permetent a la selecció actuar de manera independentment en diversos *loci*. Com les mutacions fortament seleccionades (tant positiva com negativament) no és probable que segreguin durant molt de temps en les poblacions naturals, ja que es fixaran o es perdran ràpidament, l'efecte Hill-Robertson s'espera que tingui un paper important quan la selecció és feble o moderada. Un altre model, el de selecció de fons (*Background selection*), explica com la selecció negativa en mutacions fortament deletèries elimina de les poblacions les variants que es troben lligades, siguin neutres o feblement seleccionades (CHARLESWORTH *et al.* 1993). Finalment, el model de interferència selectiva (*interference selection*), explica com el lligament creat per la selecció feble pot alterar la probabilitat de fixació d'altres mutacions (COMERON and KREITMAN 2002).

1. 2. La recombinació i els mecanismes que l'originen

En organismes superiors, la transferència d'informació d'un cromosoma al seu homòleg durant la meiosi es deu a dos mecanismes: els entrecreuaments, on es produeix l'intercanvi recíproc d'informació entre cromàtides no germanes, i la conversió gènica, que és la transferència no recíproca d'un curt interval d'informació entre cromàtides homòlogues no germanes (CHOVNICK 1973). El mecanisme de la recombinació s'inicia per uns talls en la doble cadena d'ADN entre dues cromàtides no germanes de dos cromosomes homòlegs, que són realitzats per una endonucleasa específica de la meiosi (Figura 1b). Els extrems lliures de les cadenes formen ponts d'hidrogen entre les cromàtides no complementàries (Figura 1c), donant lloc a la formació del quiasma, que conté regions d'ADN heterodúplex. Segons el model de Holliday, que és el més senzill, es forma un punt d'intersecció del quiasma que pot migrar al llarg de l'heterodúplex (Figura 1d). La resolució de l'heterodúplex es pot realitzar de dues formes equiprobables (Figura 1e), però només amb una de les dues formes de resolució té lloc l'entrecruament dels marcadors situats fora de la regió d'ADN heterodúplex (Figura 1f). Els aparellaments erronis de l'ADN heterodúplex es reparen mitjançant un sistema de correcció enzimàtic, que escindeix les bases mal aparellades d'una de les dues

cadena d'ADN de cada cromàtida i omple el buit amb la base complementària (Figura 1h). El resultat serà que algunes cromàtides presentaran fragments de conversió gènica.

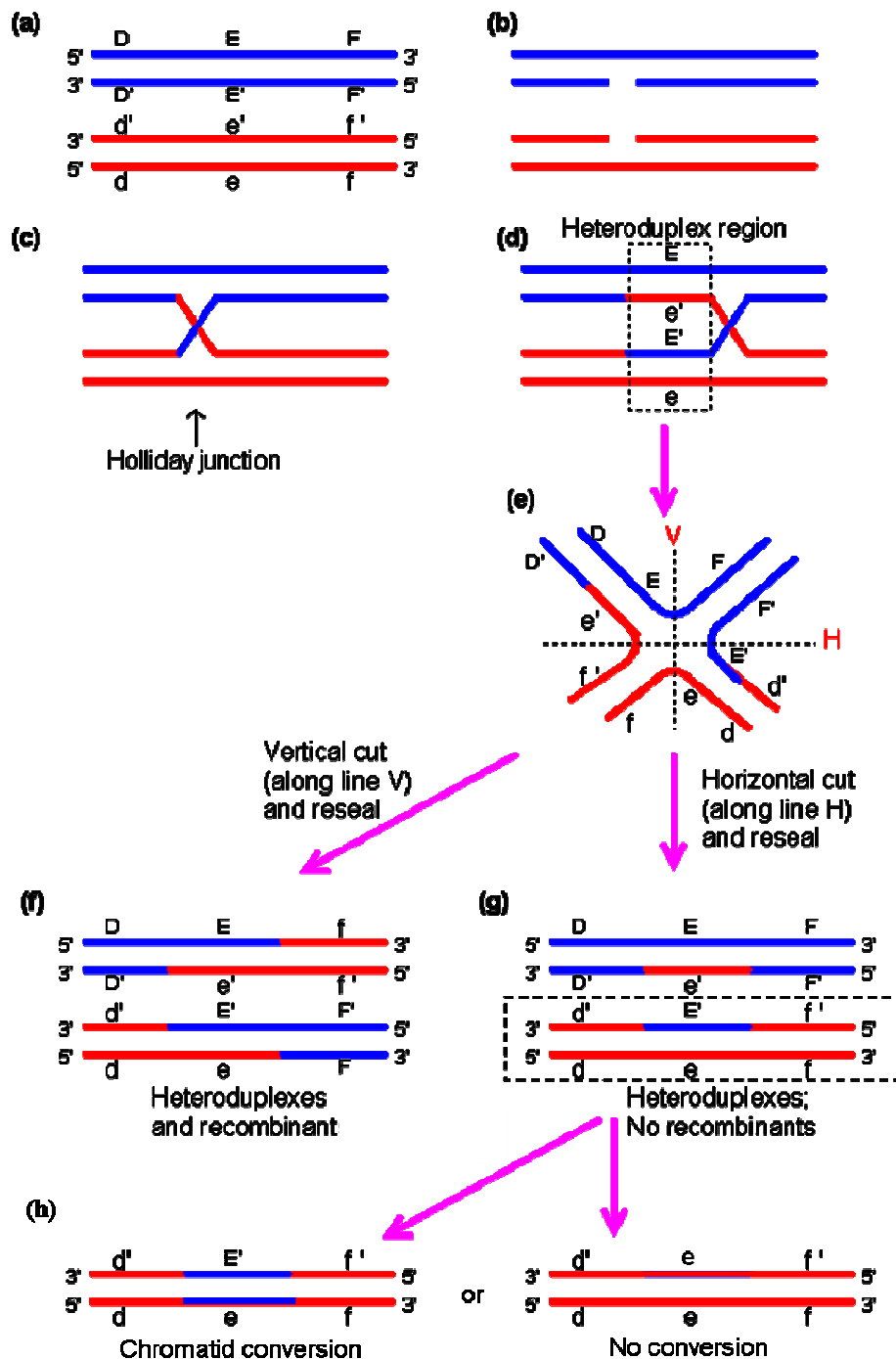


Figura 1: El model de Holliday. Adaptació de Molecular Biology Web Book, de Frank Lee (2009). <http://www.web-books.com/MoBio/>. a) els cromosomes homòlegs es troben alineats, b) s'originen talls en la doble cadena d'ADN entre cromàtides no germanes, c) unió dels extrems lliures amb ponts d'hidrogen, d) migració del punt d'intersecció, e) resolució de l'heterodúplex, f) el tall vertical produeix cromosomes recombinants, g) el tall horitzontal només origina regions heterodúplex, h) la correcció del heterodúplex pot produir fragments de conversió gènica.

1. 3. Els mapes genètics o de lligament

Alfred H. Sturtevant va realitzar el primer mapa genètic o de lligament quan estava realitzant la tesis doctoral al laboratori de Thomas H. Morgan (STURTEVANT 1913). Fent encreuaments, va observar que la freqüència de recombinació variava entre parelles de marcadors diferents. Sturtevant va descobrir que podia usar el percentatge de cromosomes recombinants per quantificar la distància entre dos marcadors i així construir un mapa genètic lineal, i va definir que una unitat de mapa o centimorgan (cM) equival a una freqüència de recombinació de l'1%, és a dir, 1 de cada 100 productes meiótics és recombinant. A més, va raonar que la probabilitat que es produeixin entrecruaments entre gens que es troben situats molt a prop és mínima, però com més separats es trobin dos gens, més gran serà la probabilitat que es produeixin entrecruaments entre ells. Més tard es va concloure que els entrecruaments d'ordre parell que es produeixin entre dos marcadors consecutius no poden ser detectats, i que com més separats es trobin aquests dos marcadors més gran serà la probabilitat que es produeixin i per tant que es subestimi la freqüència de recombinació. Per aquesta raó, es van desenvolupar funcions matemàtiques que estableixen una relació entre la freqüència de recombinació i la distància, és a dir, les funcions de mapa. La funció de mapa de John B. S. Haldane assumeix que els entrecruaments al llarg d'un cromosoma es distribueixen seguint una distribució de Poisson, sense tenir en compte que hi pugui haver interferència entre entrecruaments (HALDANE 1919). La interferència prediu que quan es produeix un entrecruament, es redueix la probabilitat que se'n produeixi un altre en un lloc adjacent. Es pot quantificar mitjançant el coeficient de coincidència, que és el quocient entre els dobles recombinants observats i esperats, donat que la suma dels coeficients d'interferència i de coincidència és 1. S'ha observat que el grau d'interferència i el número d'entrecruaments varia entre organismes (STAM 1978). Anys més tard, Damodar D. Kosambi va deduir una segona funció de mapa més realista, assumint que la distribució dels entrecruaments al llarg del cromosoma no és a l'atzar, sinó que depèn de la interferència (KOSAMBI 1944). Després, l'any 1955, Newton E. Morton va proposar l'estimador de la puntuació *Lod* per detectar lligament entre dos *loci* (MORTON 1955). *Lod* és un acrònim de l'expressió anglesa *Log of Odds* (logaritme de les probabilitats). Es calcula com el logaritme del quocient de probabilitats entre la hipòtesis que, donada una certa freqüència de recombinació, els dos marcadors es trobin lligats i la hipòtesis nul·la (que segreguin independentment, és a

dir, que la freqüència de recombinació sigui del 50%). Les puntuacions de 3 o superiors indiquen que els dos marcadors es troben lligats i puntuacions de -2 o inferiors indiquen l'absència de lligament.

1. 4. La distribució de la recombinació al llarg del cromosoma

La freqüència i la distribució dels entrecreuaments al llarg del cromosoma sembla que és molt variable (SINGH *et al.* 2009). Aquests paràmetres varien entre espècies properes, com ara dins del gènere *Drosophila* (TRUE *et al.* 1995) o bé quan es comparen humans i ximpanzés (PTAK *et al.* 2005). En humans, s'ha observat que la taxa de recombinació varia entre poblacions (FEARNHEAD and SMITH 2005) i també entre individus (COOP *et al.* 2008). En *Drosophila*, on els mascles no presenten recombinació, s'ha descrit que la taxa de recombinació pot variar entre línies (BROOKS and MARKS 1986). A més, els punts calents recombinacionals (*recombinational hotspot*) s'han descrit en un ampli ventall d'organismes, incloent plantes, llevats i animals (revisat a SINGH *et al.* (2009)).

La taxa de recombinació al llarg del cromosoma es pot estudiar a nivell local, o bé a nivell global. Els estudis a nivell local fets en *D. melanogaster* i *D. pseudoobscura* mostren que la taxa de recombinació és molt heterogènia (CIRULLI *et al.* 2007; SINGH *et al.* 2009). En canvi, quan la taxa de recombinació s'estudia a nivell global, el patró que s'obté varia entre espècies, com ara entre *D. melanogaster* i *D. pseudoobscura* (TRUE *et al.* 1995). En *D. pseudoobscura* i *D. mauritiana* s'ha observat que la recombinació en sentit ampli és uniforme al llarg del cromosoma (TRUE *et al.* 1995; HAMBLIN and AQUADRO 1999). En canvi, en *D. melanogaster* s'observa heterogeneïtat al llarg del cromosoma amb regions on la recombinació es troba fortament disminuïda, com ara vora el centròmer (LINDSLEY and ZIMM 1992; TRUE *et al.* 1995) o en regions heterocromàtiques (ROBERTS 1965).

1. 5. Les inversions cromosòmiques com a model genètic

Les inversions cromosòmiques són anomalies estructurals que consisteixen en que un segment del cromosoma es fragmenta i es torna a unir però orientat en sentit contrari. S'han descrit dos tipus d'inversions: les pericèntriques, que són les menys comunes i inclouen el centròmer en la regió invertida, i les paracèntriques, que són les més abundants i només afecten un braç cromosòmic. Aquestes mutacions solen ser força freqüents; per exemple, aproximadament tres quartes parts de les espècies del gènere *Drosophila* són polimòrfiques per inversions paracèntriques (KRIMBAS 1992). Estudis recents de comparació de genomes han mostrat la presència de nombroses regions invertides específiques d'espècie, suggerint que els reordenaments juguen un paper important en l'evolució dels genomes de les espècies (AYALA and COLUZZI 2005). Per exemple, la inversió O_3 s'observa en *D. madeirensis* però en canvi s'ha perdut en *D. subobscura*, tot i que és una espècie propera (MUNTÉ *et al.* 2005). La reducció de la recombinació a l'interior de les inversions pot facilitar el procés d'especiació. Per exemple, les inversions poden facilitar l'acumulació de gens que contribueixen a l'aïllament reproductiu entre poblacions connectades per flux gènic (NOOR *et al.* 2001). A nivell intraespecífic és freqüent trobar polimorfisme per inversions, on els cromosomes invertits es troben segregant amb els seus homòlegs no invertits (HOFFMANN and RIESEBERG 2008). En *Drosophila*, les inversions solen ser grans i cobreixen una part significativa del cromosoma (KRIMBAS and POWELL 1992). A més, el nombre d'inversions que presenten diferents espècies varia molt dintre del gènere *Drosophila* (HOFFMANN and RIESEBERG 2008).

L'estudi de les inversions cromosòmiques es va iniciar al principi del segle XX, quan Sturtevant i col·laboradors van observar per primer cop que els mapes gènics de *D. melanogaster* i *D. simulans* es trobaven invertits (STURTEVANT 1921; STURTEVANT 1926). Els insectes dípters són especialment adequats per a l'estudi de les inversions cromosòmiques a nivell citològic gràcies al gran tamany dels seus cromosomes politènics, observables al microscopi òptic. Aquests cromosomes gegants són el resultat de múltiples replicacions cromosòmiques sense divisió nuclear. En preparacions mitòtiques i meiòtiques d'individus heterozigots per inversions, s'observa la formació d'una nansa que és necessària per l'aparellament dels dos cromosomes homòlegs. La tinció amb orceïna posa de manifest bandes que divideixen el cromosoma en seccions,

permetent així dibuixar mapes citològics, com ara el de *D. subobscura* (KUNZE-MÜHL and MÜLLER 1958). Les preparacions de cromosomes politènics permeten detectar la freqüència de les inversions en les poblacions naturals, localitzar els punts de trencament de les inversions, detectar altres aberracions cromosòmiques (com ara delecions, duplicacions o translocacions) i localitzar marcadors genètics per hibridació *in situ* (Figura 2). Aquesta tècnica és encara la manera més precisa de localitzar marcadors en les espècies on no s'ha seqüenciat el genoma (com ara *D. subobscura*), ja que la comparació amb espècies properes no sol ser adequada perquè normalment no hi ha sintènia i l'ordre dels gens varia (SANTOS *et al.* 2010). Les tècniques moleculars han permès la identificació de noves inversions, per exemple mitjançant la detecció de patrons de lligament inusuals o per la comparació de l'ordre dels marcadors moleculars entre espècies. Ja sigui amb tècniques clàssiques o modernes, s'han identificat inversions en un ampli ventall d'espècies de microorganismes, plantes, insectes i animals (HOFFMANN *et al.* 2004). Per últim, s'espera que les noves tècniques d'ultraseqüenciació permetin obtenir nous genomes i una identificació més ràpida i acurada de les inversions originades a diferents llinatges (revisat en HOFFMANN and RIESEBERG (2008).



Figura 2: Preparació de cromosomes politènics d'un individu heterocariotípic de *D. subobscura* amb els ordenaments O_{3+4+1}/O_{3+4} , on s'observa la nansa formada per la inversió O_1 .

Està àmpliament acceptat que l'origen de les inversions és únic (POWELL 1997). Tanmateix, encara no hi ha consens sobre quin és el mecanisme que les origina. En algunes espècies s'ha observat que la formació de les inversions està relacionada amb la recombinació ectòpica produïda pels elements transposables (CÁCERES *et al.* 1999). Així, per la formació d'una inversió es requereixen dues seqüències repetides que es trobin al mateix cromosoma en orientació invertida. La manca de restes d'elements transposables prop dels punts de trencament d'algunes inversions ha fet que recentment

es proposin nous mecanismes, com ara el *staggered breaks model* (CASALS and NAVARRO 2007). Aquest model proposa que el primer pas és la formació de talls escalonats a les cadenes d'ADN, concretament en dos punts diferents del mateix cromosoma. L'inversió es produirà perquè els mecanismes reparadors lligaran els extrems 5' amb els extrems 3' del punt de trencament contrari. Per últim, els espais resultants són reomplerts, amb la conseqüent formació de duplicacions invertides a cada extrem de la inversió. Depenent de la longitud dels talls escalonats, les regions duplicades seran o no detectables. Finalment, es poden induir inversions de manera artificial mitjançant raigs X (SPERLICH *et al.* 1977).

1. 6. Les inversions cromosòmiques i la recombinació

A principis del segle XX, tal i com hem vist abans, Sturtevant i col·laboradors van observar que la formació d'entrecreuaments es trobava reduïda en femelles heterozigotes per inversions però no en les homozigotes (STURTEVANT and PLUNKETT 1926). Aquests investigadors van suposar que si la inversió era suficientment petita, s'impediria la formació d'entrecreuaments a causa de la dificultat en l'aparellament dels cromosomes homòlegs. En canvi, si la inversió fóra suficientment llarga perquè es donés un entrecreuament, el resultat seria la formació de cromosomes acèntrics i dicèntrics amb la conseqüent formació de gàmetes desequilibrats i zigots inviàbles. Tanmateix, Beadle i Sturtevant no van detectar la proporció d'òvuls inviàbles que esperaven trobar i, per tant, van proposar que els productes meiòtics en els òvuls de *Drosophila* estan ordenats linealment, i només un dels dos nuclis més externs dona lloc al nucli funcional de l'òvul (BEADLE and STURTEVANT 1935). Per tant, els dos interns (que contindrien els gàmetes desequilibrats) i un dels externs formaria els corpuscles polars (Figura 3). CARSON (1946) va demostrar citològicament aquest mecanisme treballant amb la mosca *Sciara*. En altres espècies, com ara el blat, els productes gamètics desequilibrats no s'eliminen, causant així importants problemes d'esterilitat (STRICKBERGER 1976). Per últim, la inversió podia ser prou gran perquè es produïssin un nombre parell d'entrecreuaments i, en aquest cas, sempre es formarien gàmetes viàbles.

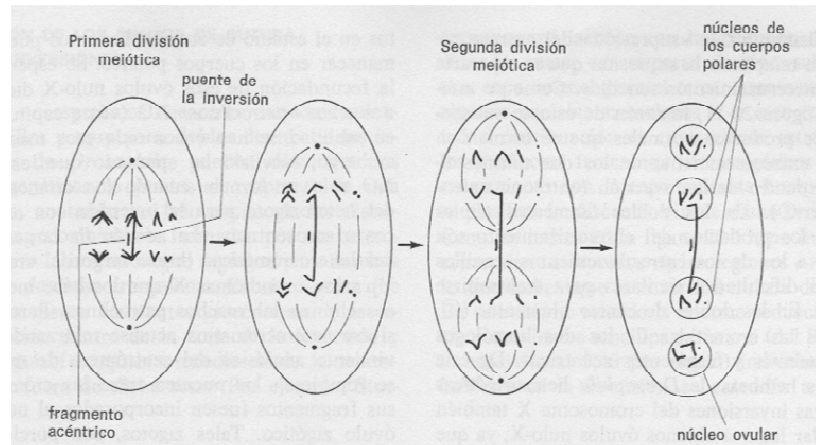


Figura 3: Esquema de la distribució lineal dels productes meiótics en els òvuls de *Drosophila* segons Beadle i Sturtevant. Figura 22.11 de STRICKBERGER (1976).

A finals dels anys 70, Paul A. Roberts, en una revisió sobre la genètica de les aberracions cromosòmiques, va explicar que les inversions cromosòmiques redueixen la recombinació per dues raons. Per una banda, la nansa d'inversió dels heterocariotips és un impediment físic que inhibeix parcialment la formació de quiasmes, sobretot vora els punts de trencament i, per altra banda, els entrecreuaments d'ordre imparell donen lloc a la formació de gàmetes inviàbles (ROBERTS 1976). En aquesta revisió va resumir que l'efectivitat d'una inversió reduint la formació d'entrecreuaments depèn de la seva longitud, de la seva posició al braç, de la sensibilitat del braç a patir reordenacions i de la presència d'inversions en altres cromosomes, que pot comportar l'augment de la recombinació. Així doncs, les inversions petites o de mida mitjana, amb els punts de trencament localitzats en la regió mitja-distal, serien les que inhibirien més eficaçment els entrecreuaments. Per últim, també va observar que la reducció dels entrecreuaments no es limita a la regió invertida, sinó que afecta a gran part del braç degut a la disrupció de la sinapsi. En general, es troba fortament reduïda vora el punt de trencament i augmenta gradualment al allunyar-se d'ell. Malgrat això, dintre de la inversió les restriccions són més fortes ja que només un nombre parell d'entrecreuaments produirà gàmetes balancejats, mentre que fora de les inversions qualsevol entrecreuament produirà descendència recombinant. A més, la reducció de la recombinació fora de la inversió no és simètrica als dos punts de trencament, sinó que varia en funció de l'espècie. En algunes espècies s'ha observat que la reducció és major a l'extrem distal de la inversió (ROBERTS 1962), mentre que en altres espècies la reducció a l'extrem proximal és superior (CARSON 1953; STUMP *et al.* 2007).

Per una banda, s'ha descrit que la presència d'inversions augmenta la taxa de recombinació en altres regions del genoma (WHITE and MORLEY 1955; KRIMBAS and POWELL 1992). Per exemple, la taxa de recombinació en *D. virilis* incrementa en les zones més allunyades de la inversió (KOMAI and TATAKU 1940). Per altra banda, l'efecte Schultz-Redfield implica que la presència d'inversions en un determinat cromosoma fa augmentar la taxa de recombinació en altres cromosomes (SCHULTZ and REDFIELD 1951). Aquest efecte s'ha observat en *D. melanogaster* (STEINBERG 1936), en *D. virilis* (KOMAI and TAKAKU 1942) i en *D. robusta* (CARSON 1953).

Arcadi Navarro i col·laboradors, van realitzar l'any 1997 un estudi teòric per veure l'efecte que produeixen les inversions sobre el flux gènic, ja sigui produït per entrecreuaments d'ordre parell o bé per conversió gènica (NAVARRO *et al.* 1997). Van fer servir dos models, però aquell que tenia en compte el fenomen de la interferència era el que s'ajustava millor a les dades empíriques. L'estudi el van resumir en tres conclusions principals: primer, la recombinació per entrecreuaments es troba totalment inhibida vora els punts de trencament però en canvi no ho està la conversió gènica. Segon, a la part central de les inversions s'espera que la taxa de flux gènic sigui més elevada, principalment deguda a recombinació per entrecreuaments. Tercer, en inversions curtes el flux gènic es troba uniformement distribuït ja que la conversió gènica és la força principal, mentre que en inversions llargues s'espera més flux gènic a les zones centrals de la inversió que a la vora dels punts de trencament. Aquesta conclusió té relació amb l'observació que la interferència juga un paper més important en inversions petites (CHOVNICK 1973). En canvi, en zones no afectades per inversions, l'efecte dels entrecreuaments seria més important que la conversió gènica a l'hora de determinar els patrons de variabilitat nucleotídica (NAVARRO *et al.* 1997). Altres autors, com Ary A. Hoffmann suggereix que la conversió gènica pot estar reduïda al voltant dels punts de trencament, on l'aparellament meiótic es troba alterat (HOFFMANN *et al.* 2004; HOFFMANN and RIESEBERG 2008). Donat que tant la conversió gènica com els entrecreuaments necessiten la formació d'un quiasma, sembla difícil pensar que vora els punts de trencament només es vegi fortament inhibida la recombinació per entrecreuament.

L'estudi pioner en la caracterització de la conversió gènica en *Drosophila* va ser el realitzat per Arthur Chovnick l'any 1973 (CHOVNICK 1973). Planteja que l'estimació de la conversió gènica és imprescindible per tal de caracteritzar adequadament el paper de les inversions com a barreres d'intercanvi en poblacions naturals. Partia de la base que en individus heterozigots, la conversió genètica seria important en aquelles regions on els cromosomes homòlegs estiguessin correctament aparellats permetent l'intercanvi. Estudiant el gen *rossy* de *D. melanogaster*, que es troba localitzat en la zona central de la inversió *In(3LR)*, va arribar a la conclusió que la conversió gènica en homozigots i en heterozigots presentava el mateix ordre de magnitud. Ell mateix apunta que la localització d'aquest gen en una zona central de la inversió pot ser la clau de l'obtenció d'aquest resultat. També raona que la interferència afecta als entrecreuaments, però no a la conversió gènica. Conclou que la formació de dobles entrecreuaments i la conversió gènica, que es produeixen amb taxes més elevades que la mutació, redueixen la diferenciació genètica entre ordenaments.

1. 7. El manteniment de les inversions a les poblacions naturals

Tot i que en general la recombinació és beneficiosa pels organismes vius, tal i com s'ha argumentat anteriorment, la recombinació també podria trencar les associacions entre gens que interactuen, produint una pèrdua d'eficàcia biològica i comportant una càrrega recombinacional (SANTOS 2009). La reducció de la recombinació és un dels factors claus per explicar el manteniment de les inversions a les poblacions naturals, ja que està estretament relacionat amb el seu valor adaptatiu. Dos articles recents han resumit els principals mecanismes que explicarien el manteniment de les inversions (HOFFMANN and RIESEBERG 2008; SANTOS 2009), que s'exposen a continuació:

Theodosius Dobzhansky va publicar l'any 1950 el primer mecanisme proposat, la hipòtesi de la coadaptació dels complexos poligènics (DOBZHANSKY 1950). Aquest mecanisme es descompon en dues components. En primer lloc, la sobredominància o heterosi dels heterocariotips s'explicaria pel manteniment de conjunts d'al·lels que interaccionen entre ells dins les inversions. Aquestes interaccions epistàtiques determinarien que l'eficàcia biològica dels heterocariotips sigui superior a l'esperada per les contribucions dels *loci* individuals. Aquest model implica que hi hagi com a

mínim dos *loci* lligats a una inversió que presentin epístasi positiva. El segon component és que aquestes interaccions són específiques de cada població. Per tant, l'intercanvi genètic entre cromosomes de diferents poblacions trencaria els complexos gènics coadaptats produint descendència amb una menor eficàcia biològica. Dobzhansky partia de la base que la selecció natural és la responsable de l'origen de l'heterosi i també del seu manteniment, preservant les variants genètics heteròtiques fins i tot quan el valor adaptatiu dels individus homozigots és molt petit. Si el polimorfisme per inversions d'una població és elevat, serà molt probable la formació d'heterozigots amb una elevada eficàcia biològica, i molt menor la probabilitat de formació homozigots. Per tant, a les espècies sense autofecundació, el polimorfisme per inversions promou que hi hagi una eficàcia biològica mitjana elevada i la preservació d'una elevada plasticitat evolutiva.

Marvin Wasserman, l'any 1968 va escriure un article en el que també explicava que el polimorfisme per inversions es pot mantenir a les poblacions si els heterocariotips presenten major eficàcia biològica que els dos homozigots (amb o sense inversió) gràcies a les interaccions epistàtiques de les combinacions de supergens o complexos gènics coadaptats (WASSERMAN 1968). A més, Wasserman va tenir en compte que l'homozigositat estructural no implica homozigositat gènica (CARSON 1958). S'ha de parlar d'homocariotips quan ens referim a homozigots pel mateix ordenament, que en poblacions naturals presentaran diferències en el contingut genètic a no ser que presentin una relació de parentiu directa. D'aquesta manera, si existeix més d'una combinació de supergens per un mateix ordenament cromosòmic en una població, els homocariotips presentaran menor eficàcia biològica degut a la disrupció de les combinacions epistàtiques favorables per recombinació, que entre heterocariotips no es veurà afectada. Aquesta recombinació produirà una disminució de la eficàcia biològica global de les combinacions afavorides.

Una hipòtesi més recent és la de l'adaptació local, proposada per Kirkpatrick i Barton l'any 2006 (KIRKPATRICK and BARTON 2006). Segons aquests autors, les inversions es troben afavorides a les poblacions naturals ja que la reducció de la recombinació en els heterocariotips manté una sèrie d'al·lels lligats a la inversió i adaptats localment, on les relacions epistàtiques entre ells no són necessàries. Sense la inversió, les combinacions favorables es perdrien per recombinació. Com l'epístasi

entre al·lels no és un requisit, aquesta hipòtesi és de caràcter més general. El destí d'aquesta inversió és gairebé la fixació, i el polimorfisme s'explica per migració o per al·lels deleteris capturats dins de la inversió.

També s'ha proposat que la inversió en sí mateixa és la diana de la selecció, gràcies a la mutació que es genera just al punt de trencament. Per exemple, en humans s'ha observat que l'expressió d'un determinat gen queda truncada per efectes de posició o la disrupció directa del gen (TADIN-STRAPPS *et al.* 2004). El manteniment d'aquestes inversions dependrà de l'eficàcia biològica de la descendència.

La subdominància es dona quan els heterocariotips presenten una eficàcia biològica menor que els homocariotips (amb o sense inversió). La disminució de l'eficàcia biològica es pot deure a diferents factors, com ara que hi hagi dificultats estructurals en l'aparellament dels heterozigots o bé per incompatibilitats genètiques (KIRKPATRICK and BARTON 2006). En diverses espècies de plantes s'ha observat que la freqüència d'entrecreuaments senzills a l'interior de les inversions és molt elevada, i en conseqüència es produeixen gàmetes desequilibrats que donen lloc a problemes d'esterilitat (RIESEBERG 2001). Aquest procés no manté el polimorfisme per inversions a la població, sinó més aviat produirà la fixació o la pèrdua de la inversió.

Per últim, les inversions també poden ser neutres. En aquest cas, la probabilitat de dispersió o pèrdua a les poblacions naturals depèn de la seva mida efectiva, de la deriva genètica i de la migració. S'ha defensat aquesta hipòtesi en els casos en què no s'han observat desviacions significatives de l'equilibri Hardy-Weinberg en el polimorfisme per inversions (ARADOTTIR and ANGUS 2004).

1. 8. *Drosophila subobscura* com a organisme model

L'espècie *Drosophila subobscura* va ser descrita als anys 30 del segle passat per J. E. Collin (Figura 4). És una espècie de dípter del grup *obscura*, que es distribueix pràcticament per tota Europa, excepte a la part central i septentrional d'Escandinàvia. A més s'ha observat al nord d'Àfrica fins el Sàhara, i al pròxim orient fins Iran, però no es coneix amb precisió la seva distribució per l'antiga URSS. Per últim, també es troba a

les illes Canàries, Açores i Madeira (KRIMBAS 1993). L'any 1978, aquesta espècie va colonitzar les costes de Xile i posteriorment, l'any 1982, es va detectar a la costa oest Nordamericana (BECKENBACH and PREVOSTI 1986). Les dues colonitzacions no han estat independents sinó seqüencials (PASCUAL *et al.* 2007), de manera que la colonització de *D. subobscura* representa un experiment natural amb dues rèpliques, sent un cas molt interessant per estudiar qüestions de tipus evolutiu (BRNCIC and BUDNIK 1980; AYALA *et al.* 1989). *Drosophila subobscura* presenta dos pics d'expansió al llarg de l'any, un més important a la primavera i un altre de menor intensitat a la tardor (SERRA *et al.* 1987; PASCUAL *et al.* 1993; ARGEMÍ *et al.* 2003). És una espècie amb una gran flexibilitat adaptativa, ja que s'ha capturat en regions amb diferents condicions climàtiques i d'hàbitat (MONCLÚS 1964; CODINA and PEREZ 1980). L'elevada abundància, que facilita la seva captura, i el seu ric polimorfisme cromosòmic per inversions, fan que sigui una espècie de referència en l'àmbit de la genètica evolutiva, com ho és la seva homòloga americana *D. pseudoobscura*. A nivell ecològic *D. subobscura* és una espècie molt desconeguda, però en canvi es desenvolupa fàcilment al laboratori. El seu cicle biològic és d'aproximadament 20 dies en condicions òptimes de densitat larvària i de temperatura (18°C) (ORENGO and PREVOSTI 1994). Per tant, totes aquestes característiques fan que sigui una espècie molt usada en estudis genètics, evolutius, ecològics i de biologia en general (KRIMBAS 1993; POWELL 1997).

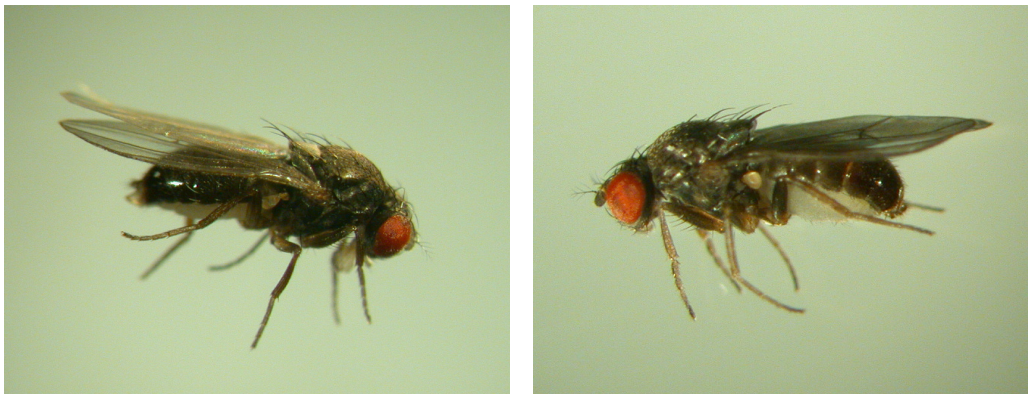


Figura 4: Imatges de *D. subobscura*, a l'esquerra s'observa un mascle i a la dreta una femella.

El cariotip de *D. subobscura* es considera que és primitiu dins del seu gènere, ja que presenta cinc parells de cromosomes acrocèntrics (A (cromosoma X), E, J, O i U) i una parella de cromosomes puntiformes (dot) (MAINX *et al.* 1953). La preparació dels cromosomes politènics que es mostra a la figura 5 és excepcional, ja que es pot veure el cariotip complet de *D. subobscura* amb tots els cromosomes units pel centròmer i la

presència d'algunes inversions. El polimorfisme per inversions és molt elevat en tots els seus cromosomes acrocèntrics (KRIMBAS and LOUKAS 1980; KRIMBAS 1993). Els primers estudis sobre polimorfisme cromosòmic per inversions daten de mitjans dels anys 50 del segle passat (STUMM-ZOLLINGER 1953). L'existència de la soca homocariotípica *chcu* (O_{3+4} , J_{ST} , U_{ST} , E_{ST} i A_{ST}) permet analitzar fàcilment el polimorfisme cromosòmic en aquesta espècie (LANKINEN and PINSKER 1977). S'ha observat que la majoria d'inversions en *D. subobscura* no es troben soles a la natura, si no que formen complexes amb altres inversions. Aquests complexos poden estar formats per inversions imbricades, com ara O_{3+4+8} o O_{3+4+23} , o bé per inversions no imbricades com O_{3+4+1} O_{3+4+7} . La línia contínua sota els números indica que les inversions es troben imbricades. L'any 1958 Kunze-Mühl i Müller van publicar un mapa citològic de *D. Subobscura*, dividit en 100 seccions i 405 subseccions, i que també mostrava els punts de trencament de les inversions (KUNZE-MÜHL and MÜLLER 1958). C. B. Krimbas va fer una revisió d'aquest mapa afegint els punts de trencament de noves inversions descrites amb posterioritat (KRIMBAS 1993).

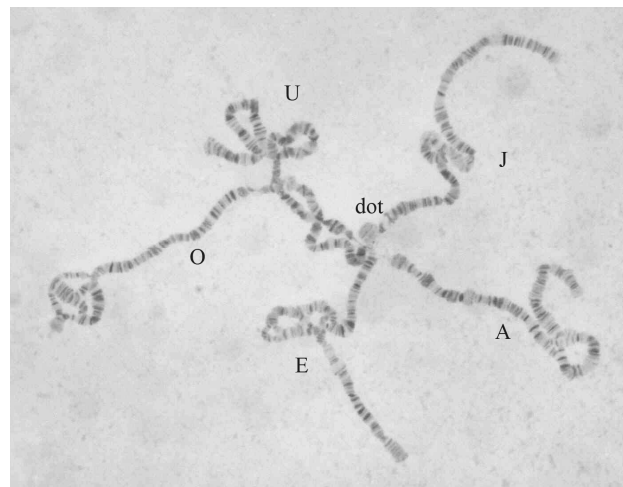


Figura 5: El cariotip de *D. subobscura* observat al microscopi òptic en una preparació de cromosomes politènics. S'observen les inversions A_2/A_{ST} , $E_{1+2+9+12}/E_{ST}$, J_1/J_{ST} , O_{ST}/O_{3+4} i U_{1+2}/U_{ST} .

La present tesi doctoral es centrarà només en el cromosoma O per una sèrie de motius. Per una banda, és un dels cromosomes que presenta major polimorfisme per inversions, ja que se n'han descrit fins a 26 inversions diferents en aquest cromosoma (BALANYÀ *et al.* 2003). A més, existeix una soca de gens letals equilibrats *Va/Ba* (*Varicose / Bare*) que permet obtenir línies homocariotípiques i heterocariotípiques letals pel cromosoma O seguint un patró d'encreuaments adequat (MESTRES *et al.* 1990; MESTRES *et al.* 1998). D'aquesta manera, no cal recórrer als dificultosos encreuaments consanguinis,

imprescindibles per obtenir línies homozigòtiques de qualsevol dels altres cromosomes. La soca *Va/Ba* va ser obtinguda per Diether Sperlich (SPERLICH *et al.* 1977) i les seves principals característiques es poden trobar a MESTRES and SERRA (2008). Per últim, el cromosoma O s'ha dividit clàssicament en dues regions: el segment distal SI, on es troben les inversions imbricades O_{3+4} , i el segment proximal SII, que ocupa les restants dues tercers parts de la longitud total del cromosoma O.

1. 9. Primeres observacions i estudis de la recombinació a *D. subobscura*

El primer grup de recerca que va abordar el tema de la recombinació a *D. subobscura* va ser l'encapçalat per John B. S. Haldane durant la Segona Guerra Mundial (PHILIP *et al.* 1944). La primera característica que va cridar l'atenció a aquest grup d'investigadors va ser que l'espècie contenia un número elevat d'inversions en el seu cariotip i, a més a més, existia polimorfisme a les poblacions. Per tant, era un patró que es desviava clarament de l'espècie model *D. melanogaster*. També observaren que els individus heterocariotípics eren més viables i fèrtils que els homocariotípics, i ho van relacionar amb la reducció de la recombinació per part de les inversions. Un altre fet els va sorprendre, i és que quan es realitzaven encreuaments al laboratori, en comptes d'obtenir igual proporció de larves homo- i heterocariotípiques per un cromosoma concret, normalment es produïa una reducció significativa dels homocariotips. Per tant, van concloure que sobre aquestes inversions actuava la selecció natural, tot i que es desconeixia el motiu. Quan es realitzaven encreuaments consanguinis del tipus germà-germana durant més de deu generacions, a la darrera generació encara detectaven individus heterocariotípics, confirmant de nou l'efecte de la selecció. També, i fent servir l'àmplia col·lecció de mutants de la que disposaven, van començar a estudiar les relacions de lligament entre ells, observant que la reducció dels entrecreuaments i la seva particular distribució era deguda a les inversions. Per exemple, van trobar interferències negatives, com ara que la presència d'un entrecreuament en la regió proximal d'una inversió incrementava molt la freqüència d'entrecreuaments a la regió just fora de la inversió però a la zona distal de la mateixa.

El primer mapa genètic del cromosoma A (X) de *D. subobscura* va ser realitzat per A. L. M. Christie l'any 1939, i contenia 120 unitats de longitud (CHRISTIE 1939). Es tracta d'un mapa aproximat, ja que es treballava amb mutants induïts per raigs X sense tenir en compte les inversions cromosòmiques que poguessin presentar. L'any 1945, Helen Spurway va elaborar un segon mapa comprovant prèviament que les soques mutants no presentessin inversions (SPURWAY 1945). Aquest mapa de lligament del cromosoma X es va realitzar amb 14 *loci* i presentava 150 unitats de longitud. Aquest nou mapa es va comparar amb el del cromosoma X de *D. virilis*, que era el més llarg d'aquella època amb 170 unitats de longitud. Posteriorment, l'any 1947 M. J. Bird va estimar la longitud del mapa genètic del cromosoma X de *D. subobscura* en 250 unitats (BIRD 1947).

El primer treball de recombinació amb el cromosoma O (llavors anomenat cinquè grup de lligament) va ser dut a terme per KOSKE and MAYNARD SMITH (1954). Usant la seva col·lecció de mutants per aquest cromosoma, van veure l'homologia entre el cromosoma O de *D. subobscura* i el braç 3R de *D. melanogaster*. L'elaboració del mapa genètic va ser difícil, degut a que alguns dels marcadors utilitzats, com ara *Ba* (*Bare*) i *DI* (*Delta*, actualment conegut pel nom de *Varicose*), eren letals recessius. A més, les soques mutants presentaven ordenaments cromosòmics diferents, amb el problema addicional que la nomenclatura d'aquests ordenaments no era homogènia entre els diferents laboratoris europeus que treballaven amb *D. subobscura* (el grup anglès ja esmentat, el suís de Stumm-Zollinger i l'austríac de Mainx). Per aquesta raó Th. Koske i J. Maynard Smith van optar per fer diferents mapes genètics en funció dels ordenaments de les soques mutants disponibles. Així van detectar que la recombinació també es trobava reduïda fora de les inversions, justificant que podia ser conseqüència d'una absència de sinapsis dels cromosomes homòlegs en aquestes regions, com també havien argumentat anteriorment Th. Dobzhansky i C. Epling a *D. pseudoobscura* (DOBZHANSKY and EPLING 1948). KOSKE and MAYNARD SMITH (1954) també van observar que la taxa de recombinació incrementava en les regions més distants de les inversions dins del mateix cromosoma. Finalment, i aquests autors van descriure la presència d'un gen dominant supressor dels entrecreuaments i que van anomenar C5 que no es trobava associat a cap inversió.

Costa B. Krimbas i Eleftherios Zouros l'any 1969 van publicar un article en el que usant les inversions com a marcadors, observaven que la recombinació es trobava suprimida entre inversions no imbricades del mateix ordenament (KRIMBAS and ZOUROS 1969). Concretament, a l'analitzar la descendència d'encreuaments entre individus heterocariotípics $O_{\underline{3+4+1}}/O_{ST}$, no van trobar cap individu recombinant. Anàlogament, tampoc van trobar recombinants en la descendència d'individus $O_{\underline{3+4+7}}/O_{ST}$. En el llibre que va publicar l'any 1993, C. B. Krimbas defensa que el fort desequilibri de lligament entre inversions no imbricades de *D. subobscura* es podia deure tant a l'absència total de recombinació en femelles heterozigòtiques, com a una forta pressió selectiva per tal de mantenir interaccions epistàtiques entre inversions. L'escola grega també va ser la primera en publicar un mapa genètic del cromosoma O de *D. subobscura*, concretament l'any 1979, per M. Loukas i col·laboradors (LOUKAS *et al.* 1979). Aquest mapa va ser realitzat amb 16 marcadors genètics, i usant la funció de mapa de Haldane, que no té en compte el fenomen de la interferència. Així, van determinar que la longitud del cromosoma O seria de 228,3 cM.

Anys després es va tornar a abordar la qüestió de les inversions lligades no imbricades de *D. subobscura* (SPERLICH and FEUERBACH-MRAVLAK 1974). Es van analitzar els cromosomes politènics de la descendència d'individus heterocariotípics $O_{\underline{3+4+7}}/O_{ST}$ i el resultat va tornar a ser que no es va trobar ni un sol recombinant dels 366 descendents analitzats, tot i que les inversions $O_{\underline{3+4}}$ i O_7 es troben molt separades, per aproximadament unes 72 bandes citològiques. SPERLICH and FEUERBACH-MRAVLAK (1974) van argumentar que els cromosomes recombinants eren eliminats per selecció i que per tant les relacions epistàtiques eren importants per mantenir el desequilibri de lligament entre inversions no imbricades de *D. subobscura*.

Als anys 90, els estudis sobre la recombinació en *D. subobscura* es van reprendre a Barcelona (MESTRES *et al.* 1998). També mitjançant cromosomes politènics, i després d'analitzar 415 individus de la descendència, es va observar que la recombinació entre individus heterocariotípics $O_{\underline{3+4+8+7}}/O_{\underline{3+4}}$, que també presentaven inversions lligades no imbricades, es trobava inhibida. De la mateixa manera, tampoc es van trobar recombinants dels 437 individus analitzats de la descendència d'heterocariotips $O_5/O_{\underline{3+4}}$, amb la diferència que aquesta vegada les inversions no imbricades no estaven lligades, sinó que es trobaven en cromosomes diferents

(MESTRES *et al.* 1998). En tots aquests estudis no es va poder descartar la presència de cromosomes recombinants com a resultat de dobles entrecreuaments dins de les inversions.

1. 10. El polimorfisme per inversions en *D. subobscura* és adaptatiu

El polimorfisme per inversions en *D. subobscura* ha estat extensament estudiat en més d'un centenar de poblacions naturals a la regió paleàrtica (KRIMBAS and LOUKAS 1980; SPERLICH and PFRIEM 1986; KRIMBAS 1992; MENOZZI and KRIMBAS 1992; KRIMBAS 1993; BALANYÀ *et al.* 2006). La clàssica discussió de si la distribució en clines latitudinals de certes inversions en *D. subobscura* era conseqüència de factors històrics o selectius es va poder respondre definitivament gràcies a la colonització del continent americà. Tot i que a les costes americanes sols va arribar un reduït nombre de colonitzadors (PASCUAL *et al.* 2001; MESTRES *et al.* 2005), la seva dispersió va suposar la formació de clines latitudinals per les inversions cromosòmiques de les que eren portadors els individus colonitzadors. El fet que les clines de Nord i Sudamèrica presenten el mateix sentit que a la regió paleàrtica confirma que les clines són adaptatives (PREVOSTI *et al.* 1988; BALANYÀ *et al.* 2003;). A més a més, s'ha observat que les clines latitudinals en les freqüències d'inversions han anat variant al llarg del temps (SOLÉ *et al.* 2002; BALANYÀ *et al.* 2003; BALANYÀ *et al.* 2004). De fet, s'ha trobat una correlació positiva entre l'augment de la temperatura i l'augment dels ordenaments típics de regions càlides (BALANYÀ *et al.* 2006). Per tant, sembla que el polimorfisme per inversions de *D. subobscura* s'està adaptant a l'escalfament global del nostre planeta. Finalment, s'ha descrit que els ordenaments $O_{\underline{3+4+7}}$ i O_5 presenten un efecte heteròtic en poblacions americanes (MESTRES *et al.* 2001).

Tal i com s'ha comentat, sembla que les clines latitudinals de *D. subobscura* responen a canvis de temperatura. En aquests sentit, es va fer un experiment per tal d'obtenir gens candidats a l'adaptació termal (LAAYOUNI *et al.* 2007). Mitjançant *microarrays*, es va analitzar el transcriptoma de tres estocs mantinguts durant tres anys a diferents règims de temperatura (fred: 13°C, òptim: 18°C, càlid: 22°C). El resultat va ser que es van identificar un elevat número de gens candidats a la regulació termal a tots

els cromosomes, i alguns d'ells es trobaven localitzats a l'interior d'inversions cromosòmiques. Per tant, és interessant poder determinar quins d'aquests gens poden tenir un paper fonamental en l'adaptació termal de *D. subobscura*, analitzant-los a nivell nucleotídic.

1. 11. Estudi de les inversions de *D. subobscura* a nivell molecular

L'efecte de les inversions en la variabilitat nucleotídica depèn principalment de dos factors, el flux gènic i l'edat de les inversions (NAVARRO *et al.* 2000). A la vegada, el flux gènic es pot veure influenciat pel valor adaptatiu de les inversions (HOFFMANN and RIESEBERG 2008) i també per la mida de les inversions (NAVARRO *et al.* 1997). En *D. subobscura* s'han realitzat estudis moleculars per mirar de resoldre aquestes qüestions; la gran majoria d'aquests s'han centrat en el segment SI del cromosoma O i en el cromosoma sexual (A).

L'any 1994, es va publicar un article en què s'analitzava per primera vegada la conversió gènica en *D. subobscura* (ROZAS and AGUADÉ 1994). Per fer-ho, van analitzar el gen nuclear *rp49*, que es troba situat a la vora d'un dels punts de trencament de la inversió O₃, en els ordenaments O_{ST} i O₃₊₄, on les inversions imbricades 3+4 cobreixen la totalitat del segment SI del cromosoma O. ROZAS and AGUADÉ (1994) van trobar una elevada diferenciació genètica entre els dos ordenaments, fet que va permetre determinar les regions convertides. Anys després, es va tornar a analitzar la variabilitat en el gen *rp49*, però aquest cop afegint dos ordenaments més que també es troben situats a la regió SI del cromosoma O: O₃₊₄₊₈ i O₃₊₄₊₂₃ (ROZAS *et al.* 1999). Tot i que es van trobar diferències significatives en totes les comparacions entre ordenaments, les més elevades es trobaven en les comparacions que incloïen l'ordenament O_{ST}. També es va quantificar la conversió gènica usant el mètode desenvolupat per Esther Betrán i col·laboradors (BETRÁN *et al.* 1997). Es va trobar que el flux gènic no estava totalment suprimit dintre les inversions, i que la seva magnitud depenia de la distància del marcador al punt de trencament de la inversió. Al gen *Acph1* (situat al segment SI) en els ordenaments O_{ST} i O₃₊₄, també s'ha pogut observar una elevada diferenciació genètica (NAVARRO-SABATÉ *et al.* 1999). En una anàlisi conjunta de vuit gens diferents localitzats als ordenaments O_{ST} i O₃₊₄, es va observar que els nivells de variabilitat eren

similars independentment de la localització dels gens utilitzats, amb una elevada diferenciació entre els ordenaments i considerables nivells de desequilibri de lligament intra i interlocus (MUNTÉ *et al.* 2005). La gran diferenciació genètica, fins i tot en les parts centrals de la inversió, s'atribuïren a l'absència d'entrecreuaments d'ordre parell o bé a que els cromosomes recombinants eren eliminats per la selecció natural per mantenir els complexos gènics coadaptats. Per últim, també s'han realitzat estudis amb el gen *Odh* de *D. subobscura*, situat al segment SII dintre de les inversions O_1 , O_5 i O_7 i fora de la inversió O_2 però a la vora del seu punt de trencament, analitzant un fragment de 800 nucleòtids. Amb aquest marcador molecular es pretenia estudiar nous aspectes de la colonització americana per part de *D. subobscura* (MESTRES *et al.* 2004; GÓMEZ-BALDÓ *et al.* 2008). A les poblacions del nou continent es van trobar només 11 haplotips diferents, segurament com a resultat de l'efecte fundador, i alguns clarament associats a certs ordenaments cromosòmics (O_5 , O_{3+4+2} y O_{3+4+7}). En canvi, en les poblacions paleàrtiques analitzades (Barcelona, Espanya i Mt. Parnes, Grècia) es van observar molts haplotips diferents (ARAÚZ *et al.* 2008), i només es va detectar associacions parcials en el casos de l'ordenament O_{3+4+1} i la inversió O_5 . Aquesta última inversió sols presentava dos haplotips, i un d'ells es troba completament associat a aquesta inversió en les poblacions americanes (MESTRES *et al.* 2004; GÓMEZ-BALDÓ *et al.* 2008).

En el primer estudi realitzat al cromosoma sexual de *D. subobscura*, es va analitzar per als ordenaments A_{ST} , A_1 i A_2 una regió de 5,4 Mb als ordenaments A_{ST} , A_1 i A_2 que incloïa el gen *yellow* (MUNTÉ *et al.* 2000). El fragment analitzat es troba molt a la vora del punt de trencament de la inversió A_1 , i fora de la inversió A_2 . La localització de la regió es correspon amb la diferenciació genètica detectada, ja que només va ser significativa en les comparacions amb l'ordenament A_1 . Recentment es va realitzar un altre estudi al cromosoma sexual, concretament en els ordenaments A_{ST} i A_2 , però aquest cop usant la informació de cinc gens nuclears (NÓBREGA *et al.* 2008). La inversió A_2 és senzilla i de mida mitjana (41,3 cM), amb una longitud genètica força més gran que el mínim teòric per tal que es formin entrecreuaments d'ordre parell (20 cM). Tot i que es va detectar diferenciació genètica significativa en cinc gens, aquesta depenia de la localització del gen, i la màxima diferenciació es va trobar en el gen més proper del punt de trencament de la inversió. Per tant, tot i la presència de flux gènic, sobretot en

zones centrals de la inversió, aquest no era suficient per homogeneïtzar el contingut genètic de les dues inversions.

Per altra banda, també s'han realitzat estudis poblacionals en *D. subobscura* a nivell molecular. Mitjançant mapes de restricció del gen *rp49*, es va fer l'anàlisi de quatre poblacions pels ordenaments O_{ST} i O₃₊₄ (ROZAS *et al.* 1995). Tres d'aquestes poblacions es trobaven al continent europeu i la quarta a l'illa de Tenerife. El resultat va ser que no es van trobar diferències entre les poblacions del continent, i la única diferenciació significativa era deguda a la població de Tenerife. Aquests resultats suggerien que hi havia un elevat flux gènic entre poblacions, contrari a les observacions prèvies de Dobzhansky (DOBZHANSKY 1950). L'any 2001 es va publicar un altre estudi poblacional a nivell molecular; aquest cop amb 10 *loci* microsatèl·lits, no necessàriament relacionats amb inversions, i comparant cinc poblacions europees i dues nord americanes (PASCUAL *et al.* 2001). El resultat va ser que no es va trobar diferenciació genètica entre les poblacions europees, fet que també suggeria un elevat flux gènic entre poblacions. En canvi, a les poblacions americanes es va observar una disminució en el nombre d'al·lels i l'heterozigositat; aquests fets es van associar a un fort efecte fundador amb una posterior expansió.

Per últim, també s'han realitzat datacions d'alguns ordenaments cromosòmics de *D. subobscura*, usant el contingut nucleotídic d'alguns dels gens que s'han seqüenciat. Les primeres datacions que es van obtenir van ser les dels ordenaments O_{ST} i O₃₊₄. Les estimes es van obtenir utilitzant la variabilitat silenciosa (és a dir, de posicions sinònimes i no codificants) de les seqüències més divergents (extraient aquelles que presentaven conversió gènica) i usant *D. pseudoobscura* com a outgroup (ROZAS and AGUADÉ 1994). Les estimes que es van obtenir van ser de 0,58 milions d'anys (Ma) en l'ordenament O_{ST} i de 0,73 Ma pel O₃₊₄. Posteriorment, usant una metodologia diferent, es van tornar a datar els ordenaments O_{ST} i O₃₊₄ i els O₃₊₄₊₈ i O₃₊₄₊₂₃. Aquest cop es va tenir en compte la variabilitat silenciosa mitjana de totes les seqüències, extraient aquelles que presentaven conversió gènica (ROZAS *et al.* 1999). A més, es van usar *D. guanche* i *D. madeirensis* com a grups externs. Les edats obtingudes per aquest mètode van ser inferiors a les anteriors (O_{ST}: 0,24 Ma; O₃₊₄: 0,33 Ma; O₃₊₄₊₈: 0,34 Ma; O₃₊₄₊₂₃: 0,28 Ma). L'edat dels ordenaments O_{ST} i O₃₊₄ es va tornar a datar usant el gen *Acph1* també situat al segment SI (NAVARRO-SABATÉ *et al.* 1999). L'edat es va estimar mitjançant el mètode descrit en ROZAS *et al.* (1999), i la datació obtinguda va ser força

similar (O_{ST} : 0,26 Ma; O_{3+4} : 0,31 Ma). També s'han obtingut algunes datacions del cromosoma A fent servir el mètode descrit per J. Rozas (ROZAS *et al.* 1999). L'edat estimada per l'ordenament A_1 va ser de 0,13 Ma (MUNTÉ *et al.* 2000) i la de la inversió A_2 en 0,16 milions d'anys (NÓBREGA *et al.* 2008).

2. 1. OBJECTIUS

El principal objectiu d'aquesta tesi doctoral és estudiar l'efecte de la recombinació en el cromosoma O de *Drosophila subobscura* respecte a la variabilitat genètica i la selecció en inversions. El treball s'ha realitzat usant dos tipus de marcadors moleculars. Per una banda, els *loci* microsatèl·lits han de permetre obtenir un mapa genètic del cromosoma O, i també determinar com les inversions modifiquen els patrons i taxes de recombinació. Per altra banda, els gens nuclears ens han de permetre estudiar els patrons de variabilitat nucleotídica determinats per les inversions a les poblacions naturals. Els objectius particulars que es van establir, van ser els següents:

1. Realitzar un mapa genètic del cromosoma O per dos tipus d'homocariotips: O_{ST} i O_{3+4} , mitjançant 14 marcadors microsatèl·lits, dels quals es coneix la seva localització citològica.
2. Analitzar el patró recombinacional al llarg del cromosoma O, correlacionant la distància genètica amb la distància física.
3. Estudiar l'efecte de les inversions inhibint la recombinació en individus heterocariotípics que només diferien per una inversió (O_{3+4+6}/O_{3+4} , O_{3+4+7}/O_{3+4}), així com heterocariotips que diferien per dues inversions no imbricades (O_{3+4+7}/O_{3+4+8}).
4. Analitzar eficàcia biològica dels homocariotips i heterocariotips en funció de la recombinació.
5. Inferir l'efecte del flux gènic sobre la diferenciació genètica entre poblacions. Per una banda, s'analitzaran 6 gens nuclears (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* i *Fmr1*) en els ordenaments O_{ST} i O_{3+4} en dues poblacions Europees: Barcelona (Espanya) i Mt. Parnes (Grècia). Per l'altra banda, també s'estudiaran les seqüències nucleotídiques del gen *Odh* en l'ordenament O_5 en diverses poblacions.
6. Detectar l'empremta de la selecció associada a les inversions, mitjançant l'anàlisi del seu contingut genètic usant gens candidats a l'adaptació tèrmica (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* i *Fmr1*) en relació als ordenaments cromosòmics O_{3+4} , O_{3+4+1} i O_{3+4+7} .
7. Finalment, s'estimarà l'edat dels ordenaments O_{3+4+1} , O_{3+4+7} , O_{3+4} i O_{ST} , usant gens situats dintre de les inversions i a la vora dels punts de trencament.

2. 1. OBJECTIVES

The main goal of this PhD Thesis is to study the effects of recombination on genetic variability and selection in the O chromosome of *D. subobscura*. The study has been approached using two types of genetic molecular markers. Microsatellite markers allow us to obtain a genetic map of the O chromosome and to describe how inversions modulate patterns and recombination rates. Then, nuclear genes enable us to study the nucleotide variability patterns caused by the influence of chromosomal inversions in natural populations. The specific objectives established are detailed below:

1. To carry out a genetic map of the O chromosome for homokaryotypes O_{ST} and O_{3+4} , using 14 microsatellite markers with a known cytological location.
2. To analyze the recombinational pattern across the O chromosome correlating genetic and physical distances.
3. To study the recombination reduction mediated by inversions in heterokaryotypes differing by one inversion (O_{3+4+6}/O_{3+4} , O_{3+4+7}/O_{3+4}) and a heterokaryotype differing by two non-overlapping inversions (O_{3+4+7}/O_{3+4+8}).
4. To analyze fitness of homo and heterokaryotypes, taking into account recombination.
5. To infer the influence of gene flow on genetic differentiation between populations. On one hand, six nuclear genes will be analyzed (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* and *Fmr1*) in O_{ST} and O_{3+4} arrangements from two European populations: Barcelona (Spain) and Mt. Parnes (Grece). On the other hand, nucleotide sequences of the *Odh* gene from the O_5 arrangement will be studied in different populations.
6. To detect the footprint of selection associated with inversions, analyzing their genetic content using candidate genes (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* and *Fmr1*) in chromosomal arrangements O_{3+4} , O_{3+4+1} and O_{3+4+7} .
7. Finally, to estimate the age of O_{3+4+1} , O_{3+4+7} , O_{3+4} and O_{ST} arrangements using genes located inside and close to the inversion breakpoint.

2. 2. ESTRUCTURA DE LA TESI

I. Determinació de la recombinació al llarg del cromosoma O de *Drosophila subobscura* i elaboració d'un mapa genètic.

Publicació 1: A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*

II. Estudi de la influència de les inversions en la formació d'entrecreuaments al llarg del cromosoma.

Publicació 2: Recombination and selection in the maintenance of the adaptive value of inversions

III. Diferenciació genètica entre poblacions naturals de *D. subobscura*.

Publicació 3: Gene flow and gene flux shape evolutionary patterns of variation in *D. subobscura*

Publicació 4: Recent origin of a chromosomal inversion revealed by its association with nucleotide haplotypes

IV. Flux genètic entre ordenaments de *D. subobscura*.

Publicació 5: The effect of recombination on medium sized inversions: selection and gene flux

I. Determinació de la recombinació al llarg del cromosoma O de *Drosophila subobscura* i elaboració d'un mapa genètic.

3. PUBLICACIONS

3. 1. Determinació de la recombinació al llarg del cromosoma O de *Drosophila subobscura* i elaboració d'un mapa genètic.

Publication 1: A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*

Publicació 1: Anàlisi de la recombinació usant microsatèl·lits: el cromosoma O de *Drosophila subobscura*

RESUM

La recombinació juga un paper important en l'adaptació de les espècies, ja que és una força evolutiva que influeix en els patrons d'organització genòmica. No obstant, en algunes ocasions la recombinació pot ser desavantatjosa, ja que pot trencar combinacions adaptatives de gens, com ara els complexos gènics coadaptats. Els mapes genètic i citològics permeten analitzar la recombinació al llarg del cromosoma. En aquest estudi es va comparar la recombinació en línies O_{ST} i O_{3+4} de *D. subobscura*, usant 14 *loci* microsatèl·lits. Els mapes genètics que es van obtenir presentaven longituds similars: respectivament 184 cM i 196 cM pels cromosomes O_{ST} i O_{3+4} . Per la majoria de parelles de marcadors analitzats, sembla que una mida mostral de 150 individus és suficient per obtenir estimes de recombinació correctes, amb l'excepció d'aquells marcadors que van hibridar a la mateixa banda citològica. La taxa de recombinació sembla que és bastant uniforme al llarg del cromosoma O, tot i que s'han observat algunes diferències a nivell regional. S'han detectat alguns *hot* i *coldspots* recombinacionals, i el seu nombre varia entre els dos tipus de línies homocariotípiques (O_{ST} i O_{3+4}). Aquesta variabilitat pot ser atribuïda a diferències en el contingut genètic de cada ordenament, o bé a diferències entre línies.

A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*.

CINTA PEGUEROLES, PEDRO A. ARAÚZ, MARTA PASCUAL AND FRANCESC MESTRES

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avd. Diagonal 645, 08028, Barcelona, Spain.

*Corresponding author: cpegueroles@ub.edu

Abstract

Recombination plays an important role in species adaptation since it acts as an evolutionary force that can influence genome pattern organization. However, recombination can be detrimental in some situations, causing the breakdown of some adaptive gene combinations such as coadapted gene complexes. Genetic and cytological chromosome maps allow recombination throughout the genome to be analyzed. In this study we compare the recombination rate of two types of homokaryotypic lines of *D. subobscura* (O_{ST} and O_{3+4}) using a set of at least 13 microsatellite *loci*. The genetic maps obtained present similar lengths: 184 cM and 196 cM for O_{ST} and O_{3+4} chromosomes, respectively. For most pairs of markers analyzed, a sample size of about 150 individuals appeared sufficient to obtain appropriate recombination values, with the exception of markers located in the same cytological band. Recombination rates seemed to be fairly uniform along the O chromosome, but some regional differences were observed. Several recombination hot and coldspots were detected, and their numbers were different in the homokaryotypic line types (O_{ST} and O_{3+4}). This variability could be attributed to differences between the genetic content of the two arrangements or to differences between the lines.

Key words: *D. subobscura*, recombination, genetic map, microsatellite *loci*, hotspots, coldspots

Introduction

Meiotic recombination is a biological mechanism which is widespread in the sexual reproduction of eukaryotic organisms due to its advantageous effects. Recombination is universal to organisms at all levels of the evolutionary scale, from viruses to vertebrates. From an evolutionary point of view, recombination plays an important adaptive role in species (Felsenstein and Yokoyama 1974, Otto and Barton 1997). Recombination can be advantageous because it increases the genetic variability of offspring. Furthermore, it is an evolutionary force in itself, capable of influencing patterns of genome organization and sequence variability. For instance, recombination can lead to the generation of new genes by exon shuffling, and to intron duplication by non-symmetrical recombination (Patthy 1999, Kolkman and Stemmer 2001). However, recombination can be detrimental in some situations, breaking down some adaptive gene combinations, such as coadapted gene complexes (Dobzhansky 1950). Inversions are chromosomal mutations which suppress recombination in heterokaryotypic individuals since crossovers usually give rise to unbalanced meiotic products. For this reason, coadapted gene complexes in inverted regions are more easily maintained over time. Genetic maps, which are estimated from the recombination frequency between genetic markers, allow us to understand the behavior of the recombination rate across the chromosome, as long as cytological maps are available.

The *Drosophila* genus is an appropriate model for carrying out recombination studies. Species belonging to this genus have a short life cycle and are easily manipulated in the laboratory. Furthermore, 12 *Drosophila* species have been fully sequenced (Clark et al. 2007). Broad-scale recombination studies with the *Drosophila* species have produced contrasting results. In these studies, the recombination rate across the chromosome is represented as the slope of the correlation between genetic and physical lengths, which is proportional to the recombination rate. In *D. mauritiana* (True et al. 1996) and *D. pseudoobscura* (Hamblin and Aquadro 1999) the recombination rate was found to be fairly uniform across the chromosome. In contrast, studies in *D. melanogaster* (Lindsley and Zimm 1992, True et al. 1996) showed that recombination rates varied between regions, generally reporting lower recombination rates near the centromere. However, these previous studies inferred recombination rate in only one chromosomal arrangement while in the present study we have compared the

recombination behavior along the chromosome in two different chromosomal arrangements.

For the *Drosophila* species whose genomes have not been sequenced, molecular markers such as microsatellite *loci* may have been isolated. These markers, if cytologically localized as in *D. subobscura* (Santos et al. 2010), can be used to study the recombination process. *Drosophila subobscura* has a primitive karyotype, composed of 5 pairs of acrocentric chromosomes (A = X, J, U, E and O, Krimbas 1993) and one dot chromosome. High inversion polymorphism in all acrocentric chromosomes has been observed in *D. subobscura* (Krimbas and Loukas 1980, Krimbas 1993). Homokaryotypic lines of the O chromosome can be obtained by carrying out the appropriate crosses with *chcu* (*cherry curled*, a homokaryotypic strain) and *Va/Ba* (*Varicose/Bare*, a balanced lethal strain) (Mestres et al. 1990, Mestres et al. 1998). This chromosome is usually divided into two segments: the proximal segment is known as SII and extends about two thirds along the whole length, and the distal segment is known as SI (Figure 1). For instance, the O_{ST} and O_{3+4} chromosomes differ by a pair of overlapping inversions (O_3 and O_4) in the SI segment (Krimbas and Loukas 1980, Krimbas 1993). The overlapping inversions cover most of this chromosomal segment (Figure 1).

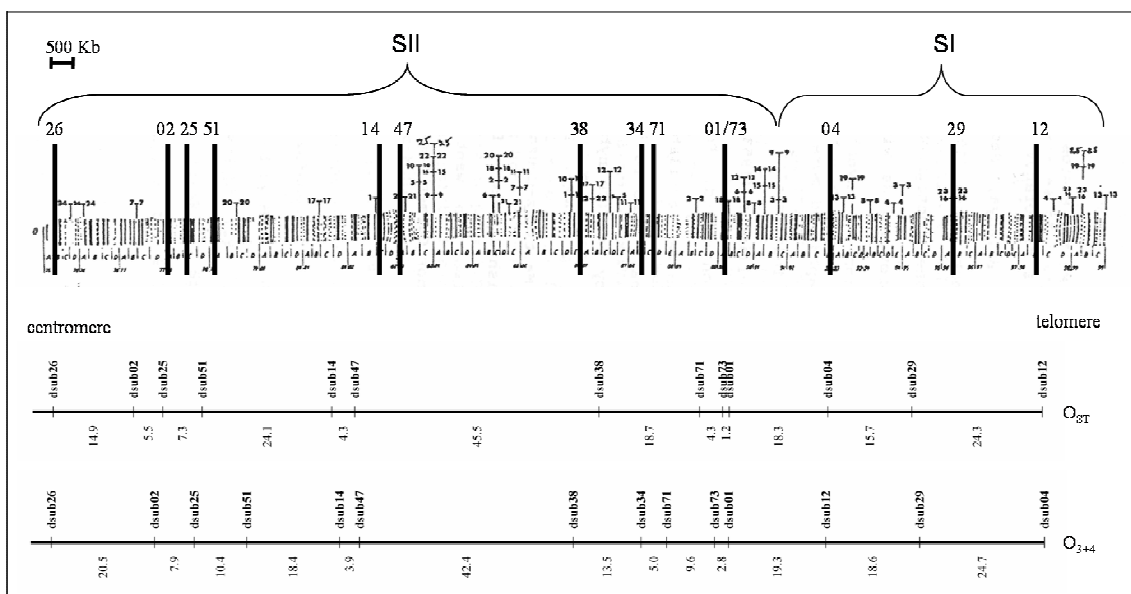


FIGURE 1: The O_{ST} arrangement cytological map presenting the microsatellite marker locations adapted from Santos et al. (2010). Below, genetic maps of O_{ST} and O_{3+4} expressed in cM.

Our main objective was to determine the recombination rates across chromosome O of *D. subobscura*. For this purpose, we analyzed two different types of homokaryotypic lines, O_{ST} and O₃₊₄, using a set of highly polymorphic microsatellite *loci* cytologically localized (Santos et al. 2010) and distributed along this chromosome. More specifically, we analyzed whether there were regional recombination rate differences. Finally, we examined whether there were differences in the recombination rates in regions presenting high density inversion breakpoints.

Materials and methods

Fly samples and DNA genotyping

Isochromosomal lines were established from different natural populations of *D. subobscura* (Sunne in Sweden, Drøbak in Norway and Mount Parnes in Greece) to obtain lines presenting genetic variability for the markers used (Ruiz-Martin 2006, Araúz et al. 2009). In all these populations, the chromosomes were made homozygous by crossing flies with *chcu* (*cherry curled*, homokaryotypic strain) and *Va/Ba* (*Varicose/Bare*, a balanced lethal strain), as previously described (Mestres et al. 1990, Mestres et al. 1998). Inversions were identified by analyzing the polytene chromosomes of third instar larvae. A total of 14 microsatellite markers (Pascual et al. 2000) previously located in the O chromosome (Santos et al. 2010, Figure 1) were used to characterize the different lines. Seven O_{ST} and 24 homokaryotypic O₃₊₄ chromosomal lines were genotyped (Supplementary material) and two lines for each arrangement (S07BBB and S01HC1 for O_{ST}; MP13 and MP16 for O₃₊₄), differing in all or most microsatellites, were selected (Table 1). The two O_{ST} lines, S07BBB and S01HC1, were crossed with each other, and the F₁ females were backcrossed to line S01HC1 to produce the recombinant backcross (BC) offspring. The S01HC1 line was selected to generate the BC offspring due to the presence of a null allele in *locus* *dsub29*. The same protocol was followed for O₃₊₄ lines MP13 and MP16, with the F₁ being backcrossed to line MP16, which was randomly selected. In order to analyze recombination, we amplified 13 markers in 165 O_{ST} BC individuals and 14 markers in 180 O₃₊₄ BC individuals. *Locus* *dsub34* was not amplified in O_{ST} crosses because it was not informative since both lines presented the same allele. Due to technical difficulties, only 117 O_{ST} and 156 O₃₊₄ individuals were genotyped for *locus* *dsub26*.

Table 1. Molecular markers across the O chromosome of *D. subobscura*.

Marker	Primers	Annealing temp.	Cytological location ^b	Multiplex ^c	O _{ST}		O ₃₊₄	
					S07BBB	S01HC1	MP13	MP16
dsub26	F: 6-Fam-GCAGGTCGGTACCTGTCAAT R: ATTTGCAGTGGCTACCCCTA	54°C	75B	A1	465	494	459	486
dsub02 ^a	F: 6-Fam-CCAGGTACGTGTACACACAGG R: TGACAAAAAGGACAACCTTCG	54°C	78A	A2	227	221	223	221
dsub25	F: Hex-TCGCGTCTTGTGTCTAAGGA R: GGCAGCAGAGGAACTACCTG	54°C	78C	A1	277	254	278	266
dsub51	F: 6-Fam-GTAAGAATGGGCAAGCGAAT R: TTGAAGCAGGTCAGACGATG	50°C	79A	B2	392	401	393	405
dsub14 ^a	F: Hex-AGCTGCCCTCCCAATGAT R: TTCCTCTGGTCAACCCTTTG	54°C	82C	A1	122	120	116	120
dsub47	F: Hex-AGTGGCCCCAAAAATAGTCA R: GCATTTTATGTTGCCCCATT	54°C	82D-83A	B1	152	137	137	150
dsub38	F: Hex-ATCCCAATACCCCTGGTAGC R: TTCCATTAAGGCCCACTCAC	50°C	86E	B2	409	417	417	416
dsub34	F: 6-Fam-AGTCCCTCGGGTATGACTGA R: CGAGCAGTGTACGCAGGTTA	54°C	88B	A2	107	107	113	107
dsub71	F: Ned-TAGCTCAATGGCCAGCTCTT R: CCAAGGCGCTCAAGATAGAC	54°C	88C-D	A2	263	261	263	273
dsub73	F: Hex-CAAAACGCGCTTTTCTTTTC R: CACACACAGTCCCCTCTCT	54°C	90A	A2	273	260	273	271
dsub01 ^a	F: Hex-CCAGAGCACTCGTGTAAGCA R: ACGTTTTGTCTTTTCGCTGGT	54°C	90A	B1	300	303	295	301
dsub04 ^a	F: 6-Fam-GCACTTGAAGTCTTGTGGCA R: TTGACGACTTCATGCTCAGG	54°C	92D	B1	195	198	206	185
dsub29	F: Ned-TTTGTTTCGCTTTCCATTC R: TTCATGCGGTTTTTCATGGTA	54°C	98B	B1	251	N	259	253
dsub12	F: 6-Fam-TCTCTGTGCTCCTGCCACTA R: CCCAAGCATTATGCAATTT	50°C	96A-B	B2	268	269	269	273

^a previously designed primers (Pascual et al. 2000, 2001); ^b Santos et al. 2010; ^c Markers amplified together; N: refers to null alleles.

Genomic DNA was extracted using the Gloor protocol (Gloor et al. 1993). Some primers were already assayed (Pascual et al. 2000, 2001) and others were designed in the present work using the *primer 3* program (Rozen et al. 2000) (Table 1); forward primers were fluorescently labeled with 6-Fam, Hex and Ned (Table 1). PCR multiplex amplifications were performed on groups of three or four microsatellites (Table 1) with 2.5 µl buffer 10x, 3 µl dNTPs (1 mM), 0.3 µl *Taq* polymerase (Amersham), 5 µl primer mix (1 µM each), 13.1 µl water and 1 µl DNA. The PCR thermal profile was 15 min at 95°C for initial denaturation, followed by 30 cycles of 30 sec at 94°C, 90 sec at the required annealing temperature (Table 1), 60 sec at 72°C, and a final 30 min extension at 60°C. *Locus* dsub26 was amplified alone due to technical difficulties with the PCR multiplex amplification, using 2.5 µl buffer 10x, 3 µl dNTPs, 0.2 µl *Taq* polymerase, 0.5

μl for each primer (10 mM), 16.3 μl water and 1 μl DNA, and the same PCR conditions. PCR fragment analyses were carried out with a 3730 Analyzer (Applied Biosystems) in the Serveis Científico-Tècnics at the Universitat de Barcelona. Finally, allele sizes were assigned using the GeneMapper 3.7 program (Applied Biosystems) by reference to the internal standard Rox 70-500 (BioVentures).

Data analysis

We carried out pairwise analyses with MapMaker software (Lander et al. 1987), to obtain genetic maps. Genetic distances were expressed in centimorgans (cM), and estimated using Kosambi's map function (Kosambi et al. 1944). This function assumes that the likelihood of a crossover is not uniform across the chromosome because of the presence of interference. The LOD score value is used as a statistical estimate of whether two *loci* are likely to be located near each other on a chromosome. LOD score values of three or greater generally indicates linkage between *loci*.

The recombination rate across the chromosome is expressed in centimorgans per kb (cM/kb). We estimated the sequence length assuming that in average the genetic content of all cytological bands is the same as considered in Munté et al. (2005). We obtained the number of cytological bands from the Kunze-Mühl und Müller map (1958). It is known that chromosome O of *D. subobscura*, 2 of *D. pseudoobscura* and 3R of *D. melanogaster* are homologous and correspond to Muller's E element. Nevertheless, there are high levels of rearrangement within them although the genetic content expressed as the number of genes is the same (Santos et al 2010). Furthermore, differences in length in Muller's E element have been detected among the 12 *Drosophila* genomes (Clark et al. 2007). We expect that the length of this chromosome in *D. subobscura* would be similar to *D. pseudoobscura* (30.7 Mb) being the closest relative with its whole genome sequenced (Clark et al. 2007). Assuming this length we have obtained a rough average estimation of the sequence length per band. Finally, markers dsub01 and dsub73 hybridized in the same cytological band and to be able to estimate recombination rate between them, the analysis was carried out as if they were separated by one cytological band. Thus, this assumption could put a lower bound on the recombination intensity between these two markers.

We performed statistical analyses with the Statistica v6 software (StatSoft). First of all, since all variables follow a normal distribution, we performed regressions between the genetic and physical distances among consecutive markers in O_{ST} and O_{3+4} BC individuals; these regression lines were used to determine whether the two arrangements were the same, analyzing variance homogeneity and implementing a test based on Fisher's F statistics to compare them. Finally, we plotted a representation of the recombination rates between markers (with their confidence intervals) according to the physical distance between them, for both O_{ST} and O_{3+4} individuals.

Results

Genetic maps

We detected high recombination levels in the current study: 82% of BC O_{ST} individuals and 89% of BC O_{3+4} individuals presented at least one crossover event. The maximum number of crossovers in an individual was five, and the most likely number of crossovers for BC O_{ST} individuals was one, compared to two for BC O_{3+4} individuals. A χ^2 test was used to compare the observed number of crossovers per chromosome and the expected number considering that they followed a Poisson distribution. Data obtained from BC O_{ST} individuals followed a Poisson distribution ($\chi^2 = 2.49$, p-value = 0.646) with a weighted mean (λ) of 1.61. By contrast, data from BC O_{3+4} individuals did not follow this distribution ($\chi^2 = 15.29$, p-value = 0.004), and 1.77 was its weighted mean.

Genetic maps (Figure 1) were constructed using all pairwise recombination frequencies. The order of the markers in both types of arrangements was consistent with their cytological location, further resolving the location of markers (dsub73 and dsub01) which were previously localized in the same chromosomal band by fluorescent in situ hybridization (Santos et al. 2010). The recombination frequencies between consecutive markers were always smaller than 0.5 and in most cases LOD scores were higher than three except between markers dsub47-dsub38 for O_{ST} crosses (Table 2). These two markers were the most distant ones, and therefore presented the highest recombination. To summarize, it would appear that we used a sufficient number of markers across the chromosome to obtain good estimations of the recombination frequencies between consecutive markers (with the exception of markers dsub47-dsub38). Furthermore, we inferred the minimum sample size to obtain an accurate estimation in our recombination studies. For this purpose, we computed the recombination frequency between all

consecutive markers for samples of individuals increasing by 25 each time. We then calculated the differences between consecutive samples for each pair of *loci* and estimated their joint variance. Finally, we plotted the variance for samples of individuals increasing by increments of 25 (Figure 2). We detected a decrease in variance with increasing sample size. We studied whether our data followed a negative exponential function. Using a non-linear regression model, our data fit the function $y = 0.0022e^{-0.018x}$ and $y = 0.0045e^{-0.023x}$ for O_{ST} and O_{3+4} individuals respectively, where the estimation of coefficients was obtained through the Marquardt method and x represents the sample size and y the standard deviation. For O_{ST} and O_{3+4} , p-values are 0.00000006 and 0.0000 respectively, showing the goodness of the fitted models. Both functions are asymptotically close to 0, and the variation of standard deviations for a sample size of 150 individuals is negligible.

Table 2. Recombination frequency (RF) and LOD scores between consecutive markers.

O_{ST}				O_{3+4}			
Markers	R. F.	LOD	kb	Markers	R. F.	LOD	kb
dsub26 / dsub02	0.145	14.16	3362	dsub26 / dsub02	0.194	14.42	3362
dsub02 / dsub25	0.055	34.22	820	dsub02 / dsub25	0.078	32.02	820
dsub25 / dsub51	0.073	30.72	902	dsub25 / dsub51	0.103	26.99	902
dsub51 / dsub14	0.224	11.53	4510	dsub51 / dsub14	0.176	17.19	4510
dsub14 / dsub47	0.043	36.52	492	dsub14 / dsub47	0.039	40.77	492
dsub47 / dsub38	0.361	2.77	5494	dsub47 / dsub38	0.345	3.72	5494
dsub38 / dsub71	0.179	15.92	1968	dsub38 / dsub34	0.132	16.55	1722
				dsub34 / dsub71	0.050	38.67	246
dsub71 / dsub73	0.042	33.95	2214	dsub71 / dsub73	0.094	19.67	2214
dsub73 / dsub01	0.012	44.38	82	dsub73 / dsub01	0.028	44.26	82
dsub01 / dsub04	0.176	16.14	3116	dsub01 / dsub12	0.184	17.38	2296
dsub04 / dsub29	0.152	18.96	3526	dsub12 / dsub29	0.178	17.38	2132
dsub29 / dsub12	0.226	11.34	2132	dsub29 / dsub04	0.229	12.05	2870

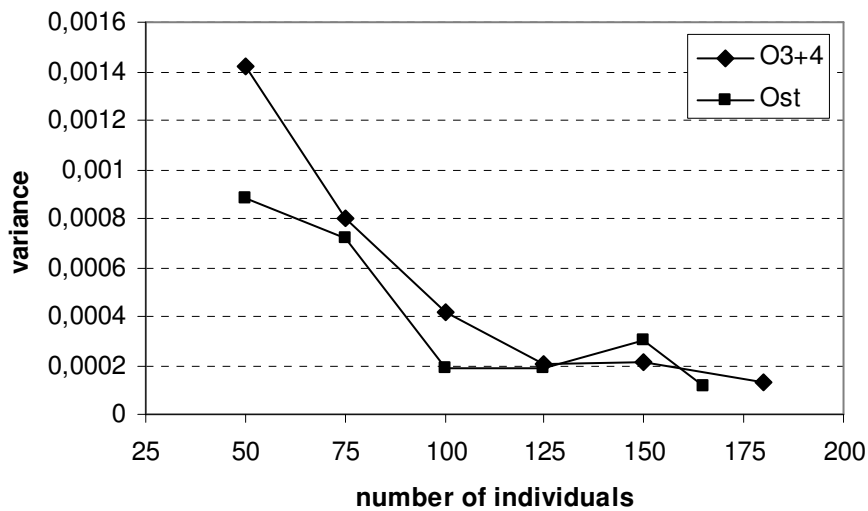


FIGURE 2: Joint variance of the differences between consecutive samples for each pair of *loci* (variance) vs samples increasing by 25 individuals each time (number of individuals) in O_{ST} and O_{3+4} .

The number of recombinant individuals found between consecutive markers ranged from two to 59 in the O_{ST} arrangement and from five to 58 in the O_{3+4} arrangement. We were able to compute the correlation between the number of recombinants and physical distances (kb), since both variables follow a normal distribution (p-values = 0.3326 and 0.2083 for O_{ST} and p-values = 0.0607 and 0.4186 for O_{3+4}). In fact, there is a significant correlation between the number of recombinant individuals and the physical distance between markers (for the O_{ST} $R^2 = 0.6946$, p-value = 0.0008; for the O_{3+4} $R^2 = 0.8084$, p-value = 0.0). Thus, in order to obtain sufficient recombinant individuals to carry out proper statistics, it would be advisable to increase the sample size between markers separated approximately 100 kb, which is approximately the physical distance between markers *dsub73-dsub01* estimated from *D. pseudoobscura* (Santos et al. 2010).

Recombination rate distribution along the chromosome

The accumulative genetic distance for each microsatellite locus in relation to locus *dsub26*, which is the nearest marker to the centromere, was plotted against the corresponding physical distance (Figure 3). Interestingly, there was a linear relationship between genetic and physical distances in both crosses (O_{ST} $R^2 = 0.9927$, p-value = 0.0 and O_{3+4} $R^2 = 0.9932$, p-value = 0.0). This suggests that on average, the recombination rate is fairly constant across the chromosome. The line slope (b) in Figure 3 was

proportional to the recombination rate, and allowed us to compute the uniform recombination rate for O_{ST} ($b= 6.627$ cM/Mb) and O_{3+4} ($b= 7.194$ cM/Mb).

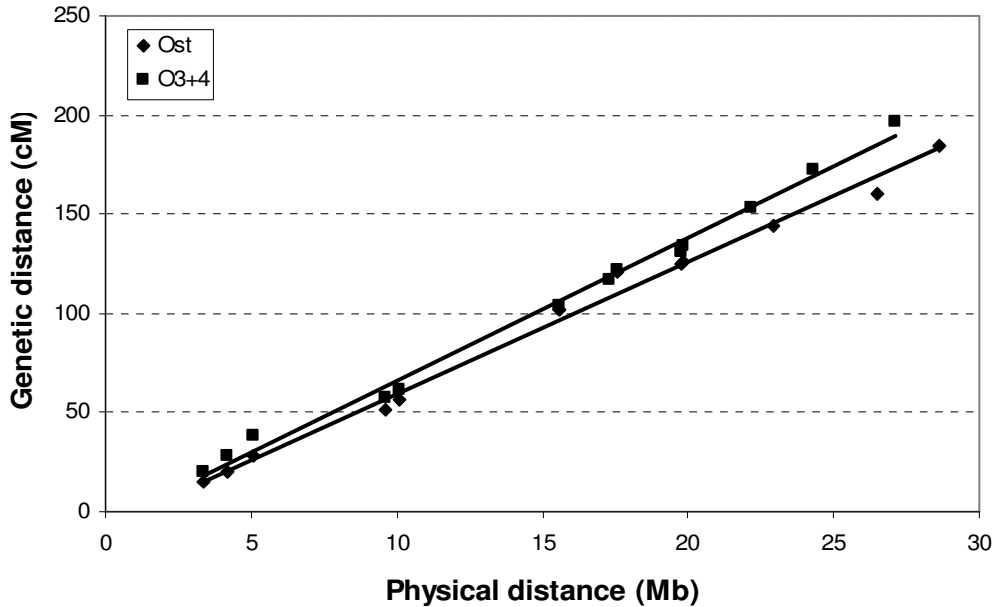


FIGURE 3: Physical (Mb) vs genetic distances (cM) in O_{ST} and O_{3+4} BC individuals with reference to the first *locus* studied (dsub26).

Furthermore, to assess whether there were differences between each type of karyotype (O_{ST} and O_{3+4}), we compared only the SII region of the O chromosome because it is equivalent in both arrangements, while in region SI the markers are rearranged (Figure 1). There was also a significant correlation between genetic and physical distances between markers in both arrangements for region SII, O_{ST} ($R^2 = 0.803$, p -value = 0.001) and O_{3+4} individuals ($R^2 = 0.798$, p -value = 0.001). The variance homogeneity test indicated that the variances of the two groups can be considered homogenous ($F = 1.322$, p -value = 0.351) and consequently we were able to carry out a regression line comparison test. This comparison indicated that there were significant differences between the two regression lines ($F = 4.945$, p -value = 0.024).

Regional differences in the recombination rate

For both O_{ST} and O_{3+4} individuals, the recombination rates between markers and their confidence intervals were plotted according to the physical distance between markers. It appeared that the recombination rate was not homogeneous at regional level, and in some cases confidence intervals between some markers did not overlap. In the O_{ST} arrangement, we identified a coldspot between markers dsub71-dsub73 (Figure 4).

In the O_{3+4} arrangement, we identified two hotspots between markers dsub73-dsub01 and dsub34-dsub71 (Figure 4).

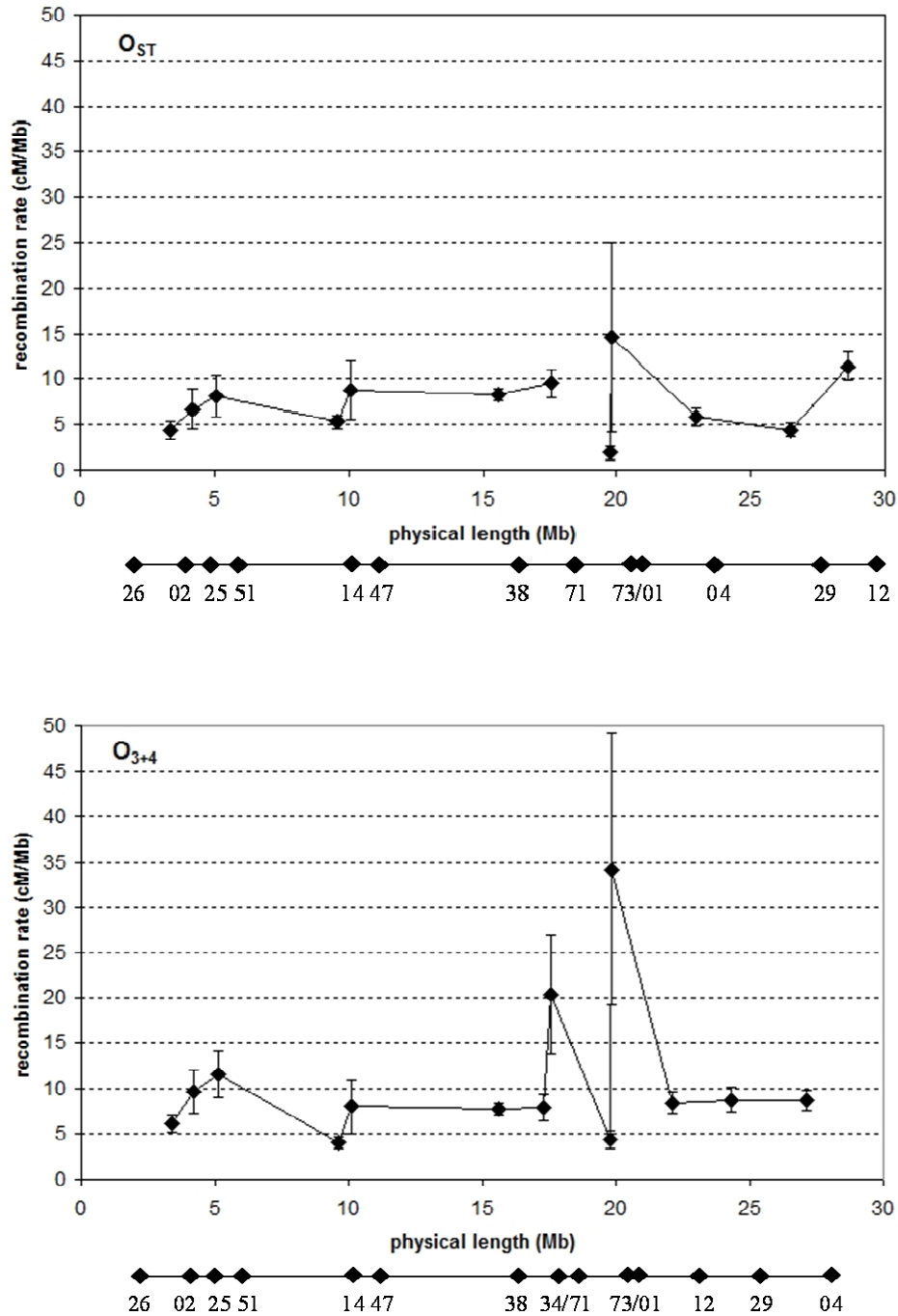


FIGURE 4: Recombination rates between markers and their confidence intervals plotted according to the physical distance between them, for O_{ST} and O_{3+4} BC individuals.

Furthermore, we compared recombination rates in regions with high density inversion breakpoints with the remaining regions in BC individuals. For O_{3+4} crosses, regions with high density were those between markers dsub47-dsub38, dsub38-dsub34 and dsub01-dsub12 (Figure 1). For O_{ST} crosses, regions with high density breakpoints were those between markers dsub47-dsub38, dsub38-dsub71, dsub01-dsub04 and dsub04-dsub29. Assuming that the number of crossovers along the O chromosome follows a Poisson distribution, we compared the λ parameter of the regions with high density of inversion breakpoints with the λ parameter of the regions with low density. There were no significant differences between these parameters for either of the crosses (in a maximum likelihood test with Wilks approximation, the p-values are always > 0.95).

Finally, we detected the number of double recombinants between three consecutive markers using MapMaker software (Table 3). Most double recombinants were found in regions with a recombination frequency higher than 20%. However, we identified one double recombinant between markers dsub71-dsub01 in O_{3+4} BC individuals despite the recombination frequency being smaller than 15%. Most surprisingly, three double recombinants were detected between markers dsub02-dsub51 in O_{ST} BC individuals, despite them presenting a recombination frequency smaller than 10%.

Table 3. Recombination frequency (RF) and double recombinants (DR) among three consecutive markers.

O_{ST}	RF	DR	O_{3+4}	RF	DR
dsub26 / dsub25	0.20	0	dsub26 / dsub25	0.28	0
dsub02 / dsub51	0.09	3	dsub02 / dsub51	0.18	0
dsub25 / dsub14	0.26	3	dsub25 / dsub14	0.25	3
dsub51 / dsub47	0.26	1	dsub51 / dsub47	0.21	1
dsub14 / dsub38	0.39	2	dsub14 / dsub38	0.35	2
dsub47 / dsub71	0.40	0	dsub47 / dsub34	0.34	3
dsub38 / dsub73	0.21	1	dsub38 / dsub71	0.18	0
			dsub34 / dsub73	0.14	0
dsub71 / dsub01	0.06	0	dsub71 / dsub01	0.11	1
dsub73 / dsub04	0.18	0	dsub73 / dsub12	0.18	3
dsub01 / dsub29	0.31	1	dsub01 / dsub29	0.35	2
dsub04 / dsub12	0.30	6	dsub12 / dsub04	0.36	4

*Discussion***Genetic maps**

The genetic map of the O chromosome of *Drosophila subobscura* completely matched its cytological map, obtained by the localization of the same microsatellite markers by fluorescent in situ-hybridizations (Santos et al. 2010). Furthermore, it allowed refining the location order of markers mapping in the same cytological band. The number of crossovers per chromosome in O_{ST} recombinants followed a Poisson distribution. This was not the case for O₃₊₄ BC individuals, which could be a consequence of the higher proportion of extreme values and recombination coldspots observed. Despite these discrepancies, the weighted means of the Poisson distributions were similar (1.61 and 1.77 in O_{ST} and O₃₊₄ BC individuals, respectively) although the higher value obtained for the O₃₊₄ arrangement was caused by an excess of two crossover classes per chromosome. Most of the individuals presented one or two crossovers, consistent with previous values described in different organisms (Ford and Hamerton 1956). These values can arise from meiotic regulation since it has been suggested that having at least one crossover per bivalent is fundamental for correct chromosomal meiotic disjunction (Nicklas 1974) and multiple crossovers cannot be randomly distributed due to interference (Carpenter 1988).

We estimated the genetic length of the O chromosome of *D. subobscura* as 190.7 cM, which is the average length between the O_{ST} (184.4 cM) and O₃₊₄ arrangements (197.0 cM). The length obtained in the present study was smaller than the 228.3 cM obtained in a previous work (Loukas et al. 1979). The previous genetic map was obtained using 16 markers located in the O chromosome but with unknown cytological location, and using the Haldane map function, which does not consider interference. Interference reduces the probability of a crossover forming close to an existing one and it is expected to be present between markers with recombination frequencies higher than 15%. Given that recombination frequencies between markers were higher than 15%, the genetic length of the O chromosome in the map obtained by Loukas et al. (1979) could be somewhat overestimated because they used the Haldane function (Lynch and Walsh, 1998).

Genetic maps of chromosome 2 of *D. pseudoobscura* (homologous to the O chromosome of *D. subobscura*) have also been obtained in different studies. The total

genetic length estimated varies from 101 cM (Hamblin and Aquadro 1999) to 204 cM (Ortiz-Barrientos et al. 2006). These discrepancies were attributed to the different number of microsatellite markers used in each study (10 genes and 18 microsatellite markers, respectively), and it was concluded that the best estimate was approximately 200 cM, this being the genetic length obtained using the higher number of markers (Ortiz-Barrientos et al. 2006). Thus, the chromosomal lengths of *D. subobscura* and *D. pseudoobscura* seem to be similar. On the other hand, genetic lengths of chromosome 3R of *D. melanogaster*, *D. simulans* and *D. mauritiana* (which are homologous to the O chromosome of *D. subobscura*) are smaller, approximately 56 cM, 75 cM and 125 cM, respectively (True et al. 1996). Recombination frequency estimates were obtained using a different number of markers in each species, the number of markers used in *D. melanogaster* being much higher. Thus the different number of markers or their recombination behavior could be responsible for the differences in length between these species, since chromosome 3R of *D. melanogaster* and *D. simulans* are 27.9 Mb and 27.5 Mb, respectively (Clark et al. 2007). On the other hand, chromosome 2 of *D. pseudoobscura* is 30.7 Mb (Clark et al. 2007); although this chromosome has a slightly larger number of base pairs compared to the homologous chromosome in the species of the *melanogaster* group, the size alone cannot explain the differences in genetic length. The recombination rate in *D. subobscura* ranges from 2 to 11 cM/Mb with a weighted mean of 7 cM/Mb. These values are very similar to those estimated for *D. pseudoobscura* (from 4 to 11 cM/Mb with a weighted mean of 7 cM/Mb) using data from Ortiz-Barrientos et al. (2006). Recombination rates in *D. pseudoobscura* are high compared to many other taxa, including its close relative *D. persimilis*, and almost four times higher than those observed in *D. melanogaster* (Ortiz-Barrientos et al. 2006). Therefore, the fact that *D. pseudoobscura* and *D. subobscura* have greater genetic lengths than the three species of the *melanogaster* group is probably due to a combination of a greater number of base pairs and a higher recombination rate.

Differences between arrangements

In our study, we observed some differences in the genetic distance between consecutive markers when comparing the two types of homokaryotypic lines. The average of these differences was 3 cM, and the greatest difference detected was 5.7 cM. In *D. melanogaster*, Ashburner (1989) reported that some factors, such as temperature or the hatching time of the eggs, could affect recombination rates. However, in the present

study, we maintained all chromosomal lines and performed all the crosses at a constant temperature (18°C), and we randomly genotyped some descendents of the total offspring. Therefore it does not appear that these mediated environmental factors can explain the differences in the recombination rate. Furthermore, the recombination frequency of a chromosome could increase if other chromosomes of the karyotype are heterozygotes for inversions (Schultz-Redfield 1951). This could affect our estimates, because the O_{3+4} and O_{ST} lines came from Mt. Parnes in the South and Sunne in the North respectively, and Southern populations are expected to have higher inversion polymorphism than Northern ones (Krimbas 1993). Despite only knowing the arrangement of the O chromosome, higher recombination rates detected in O_{3+4} BC individuals could be due to inversions in other chromosomes in the lines. The homokaryotypic lines were obtained after three generations of crosses with laboratory strains, and therefore the genetic background for all chromosomes except for the O chromosome could be highly homogeneous. Nevertheless, the starting variability of the wild individuals may have affected the final chromosomal content. Individuals from the Sunne population presented a negligible variability of chromosomal inversions for all chromosomes without taking the O chromosome into account (mean heterozygosity 0.23, computed using data from Ruiz-Martín 2006), most individuals being homokaryotypes for the ST arrangements, as are the laboratory strains. Therefore our O_{ST} lines most probably do not present any inversion in the other chromosomes of the karyotype. However, the individuals from the Mt. Parnes population presented a larger inversion polymorphism for the same chromosomes (mean heterozygosity 0.55, computed using data from Ruiz-Martín 2006), thus the presence of inversions in chromosomes other than the O chromosome cannot be completely ruled out. Furthermore, we cannot dismiss the effect the differences in the genetic content of the O chromosomes between lines had on the recombination rate, since the lines used had a completely different set of alleles in most microsatellites (Table 1).

In recombination studies of *Drosophila*, between 180 and 275 individuals are usually analyzed (Hamblin and Aquadro 1999, Staten et al. 2004 and Ortiz-Barrientos et al. 2006). In our work, we detected stabilization of the recombination frequency when we used around 150 individuals for most pairwise comparisons between consecutive markers (Figure 2), indicating that this number seemed to be sufficient to obtain correct estimates. Thus, despite the number of O_{ST}/O_{ST} individuals analyzed being smaller than

the O_{3+4}/O_{3+4} ones, both estimates seem accurate and comparable. However, the number of individuals required to obtain an accurate estimate depends on the physical distance between markers; it is advisable to increase the sample size for markers closer than approximately 100 kb. For this reason, fine-scale recombination studies (Cirulli et al. 2007, Singh et al. 2009) and recombination studies using markers separated by a small physical distance (Kulathinal et al. 2008), require the analysis of a high number of individuals.

Recombination rate distribution across the chromosome

The broad-scale recombination rate seems to be fairly uniform across the O chromosome of *D. subobscura*. Similarly, uniformity in recombination rates has also been reported in the equivalent of the O chromosome in *D. pseudoobscura* (Hamblin and Aquadro 1999) and *D. mauritiana* (True et al. 1996). However, Ortiz-Barrientos et al. (2006) in *D. pseudoobscura* and True et al. (1996) in *D. melanogaster* and *D. simulans* also detected lower recombination rates near the pericentromeric region. Nevertheless, fine-scale recombination studies, that is when studying physical distances smaller than 2 Mb or when substantially increasing the number of markers across the chromosome, have reported that the recombination rate can be quite heterogeneous (Cirulli et al. 2007, Kulathinal et al. 2008, Singh et al. 2009). Similarly, we also detected differences in the regional recombination rate by detecting recombination hot and coldspots. The presence of recombination hotspots has been observed in a wide range of organisms, including several *Drosophila* species (reviewed in Singh et al. 2009). For instance, in *D. pseudoobscura*, recombination hotspot analysis was carried out in a region of approximately 2000 kb on the X chromosome (Cirulli et al. 2007). They observed the presence of recombination hotspots, the hottest region being 38-fold than the coldest, and 15-fold than the background recombination rate. Moreover, hotspots were also detected when analyzing the recombination rate of the whole X chromosome of *D. pseudoobscura* using 50 markers separated by an average of 466 kb (Kulathinal et al. 2008). They estimated that the 95% confidence interval of the recombination rate spanned from 0.9 cM/Mb to 22 cM/Mb, which corresponds to ~24-fold variation. In the present study, the recombination rates in the O chromosome of *D. subobscura* spanned from 1.9 cM/Mb to 14.6 cM/Mb in O_{ST} and from 4.1 cM/Mb to 34.1 cM/Mb in O_{3+4} . Considering the confidence intervals as in Kulathinal et al. (2008), variation was ~21 and ~15-fold in O_{ST} and O_{3+4} , respectively. Thus, we can conclude

that the recombination rate at the regional level in *D. subobscura* is heterogeneous since in the O_{ST} arrangement, we identified a coldspot between markers dsub71-dsub73 (Figure 4). Moreover, in the O_{3+4} arrangement, we identified two hotspots between markers dsub73-dsub01 and dsub34-dsub71 (Figure 4). However, the study which analyzed the local recombination rate in *D. melanogaster* failed to find hotspots, since in a region of 1.2 Mb in the X chromosome, the recombination rate varied ~ 3.5 -fold (Singh et al. 2009). Finally, recombination rate estimations depend on the physical distance between markers, which was obtained assuming that in average cytological bands contain approximately the same sequence length. Although deviations of this assumption could affect the hot and coldspot regions estimations, the regions that we have detected as recombination outliers do not seem to deviate from the general banding pattern of the chromosome.

It is well known that recombination rates and Kosambi's map function estimates are equivalent for values below a recombination rate of 15%, because double crossover effects due to interference are negligible in these regions. However, markers dsub71-dsub01 and dsub02-dsub51, whose recombination frequencies are 0.11 and 0.09, present double recombinants (Table 3). Mutation can be discarded since the probability of a new allele with the same size arising from the parental line is very small given that the alleles from the two parental lines were quite different (Table 1). Recombination could arise from double crossovers or gene conversion affecting the *locus* under study, although with our approximation it is not possible to distinguish between both types of events.

Double crossovers, which are undetectable between two markers, can affect recombination rate estimates. However, despite being undetectable, we can make a rough estimation of double crossovers, considering that recombination frequency is uniform for markers separated by approximately the same distance. Nevertheless, the estimation of the distances needs to be considered as an approximation, since genetic content of cytological bands could not be uniform in all chromosomal regions. In the O_{ST} homokaryotypes one double recombinant was detected between markers dsub38-dsub73. Since the physical distance between these markers (4182 kb) is similar to the distance between markers dsub51-dsub14 (4510 kb), we can extrapolate that approximately one double crossover occurred between them. Using the same rationale,

markers separated by the highest physical distance (dsub47-dsub38, 5494 kb) could be affected by approximately six double crossovers, which is the number of double recombinants estimated between dsub04-dsub12 with an estimated physical distance of 5658 kb. Thus, genetic distance between markers separated by a physical distance higher than 4.5 Mb requires the use of Kosambi's map function in order to obtain a more realistic genetic distance.

We can conclude that despite having detected recombination rate homogeneity across the chromosome, we observed significant regional differences. Several recombination hot and coldspots were detected, and their numbers were different in the homokaryotypic line types (O_{ST} and O_{3+4}). This variability could be attributed to differences between the genetic content of the two arrangements or to differences between the lines.

Acknowledgements

We would like to thank Prof. Joan Balanyà (Dept. of Genetics, Universitat de Barcelona) for the analyses of the polytene chromosomes, Prof. Concepció Arenas (Dept. of Statistics Universitat de Barcelona) for her statistical advice and Mr. Robin Rycroft (Language Advisory Service, Universitat de Barcelona) for the English grammar corrections. Furthermore, we would also like to thank Prof. Charles F. Aquadro, Prof. Martha Hamblin (Cornell University), Prof. Luís Serra, Prof. Ferran Palero and Prof. Enrique Blanco (Universitat de Barcelona) for their valuable comments on the manuscript. This work was supported by a pre-doctoral fellowship to CP (2009FIC-00096) from the Generalitat de Catalunya (Spain). The research was funded by projects CGL2006-13423-C02-02 from the Ministerio de Ciencia y Tecnología (MCYT, Spain) and SGR2009-636 from the Generalitat de Catalunya (Spain).

Literature cited

- Araúz PA, Mestres F, Pegueroles C, Arenas C, Tzannidakis G, Krimbas CB and Serra, L, 2009. Tracking the origin of the American colonization by *Drosophila subobscura*: genetic comparison between Eastern and Western Mediterranean populations. *J Zool Syst Evol Res* 47:25-34.
- Ashburner M, 1989. *Drosophila*, a laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

- Carpenter ATC, 1988. Thoughts on recombination nodules, meiotic recombination, and chiasmata. In: Kucherlapati R, Smith GR, editors. Genetic recombination. American Society of Microbiology, Washington, DC pp 529-548.
- Cirulli ET, Kliman RM, Noor MAF, 2007. Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. J Mol Evol 64:129-135.
- Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al., 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450:203-218.
- Dobzhansky Th, 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. Genetics 35:288-302.
- Felsenstein J, Yokoyama S, 1974. The evolutionary advantage of recombination. Genetics 78:737-756.
- Ford CF, Hamerton JL, 1956. The chromosomes of man. Nature 178:1020-1023.
- Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK, Robertson HM, Engels WR, 1993. Type I repressors of P element mobility. Genetics 135:81-95.
- Hamblin MT, Aquadro CF, 1999. DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: a study of the second chromosome. Genetics 153:859-869.
- Kolkman JA, Stemmer WPC, 2001. Directed evolution of proteins by exon shuffling. Nat Biotechnol 19:423-428.
- Kosambi DD, 1944. The estimation of map distances from recombination values. Ann Eugen 12:172-175.
- Krimbas CB, 1993. *Drosophila subobscura*: Biology, Genetics and Inversion Polymorphism. Verlag Dr. Kovac, Hamburg.
- Krimbas CB, Loukas M, 1980. The inversion polymorphism of *Drosophila subobscura*. Evol Biol 12:163-234.
- Kulathinal RJ, Bennett SM, Fitzpatrick CL, Noor MAF, 2008. Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. Proc Natl Acad Sci USA 105:10051-10056.
- Kunze-Mühl E, Müller E, 1958. Weitere Untersuchungen über die chromosomale Struktur und die natürlichen Strukturtypen von *Drosophila subobscura* Coll. Chromosoma 9:559-570.

- Lander ES, Green P, Abrahamsom J, Barlow A, Daly MJ, Lincoln SE, Newburg L, 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.
- Lindsley DL, Zimm GG, 1992. *The Genome of Drosophila melanogaster*. Academic Press, San Diego.
- Loukas M, Krimbas CB, Mavragani-Tsipidou P, Kastritsis CD, 1979. Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. *J Heredity* 70:17-26.
- Lynch M, Walsh B, 1998. *Genetics and analysis of quantitative traits*. Sinauer Associates, Inc. Sunderland, MA.
- Mestres F, Sanz J, Serra L, 1998. Chromosomal structure and recombination between inversions in *Drosophila subobscura*. *Hereditas* 128:105-113.
- Mestres F, Pegueroles G, Prevosti A, Serra L, 1990. Colonization of America by *Drosophila subobscura*: lethal genes and the problem of the O₅ inversion. *Evolution* 44:1823-1836.
- Munté A, Rozas J, Aguadé M and Segarra C, 2005. Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* 169:1573-1581.
- Nicklas RB, 1974. Chromosome segregation mechanisms. *Genetics* 78:205-213.
- Otto SP, Barton NH, 1997. The evolution of recombination: removing the limits to natural selection. *Genetics* 147:879-906.
- Ortiz-Barrientos D, Chang AS, Noor MAF, 2006. A recombinational portrait of the *Drosophila pseudoobscura* genome. *Genet Res* 87:23-31.
- Patthy L, 1999. Genome evolution and exon-shuffling – a review. *Gene* 238:103-114.
- Pascual M, Schug MD, Aquadro CF, 2000. High density of long dinucleotide microsatellites in *Drosophila subobscura*. *Mol Biol Evol* 17:1259-1267.
- Pascual M, Aquadro C, Soto V, Serra L, 2001. Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. *Mol Biol Evol* 18:731-740.
- Rozen S, Skaletsky HJ, 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

- Ruiz-Martín H, 2006. Estudi de les associacions entre al·lels de *loci* microsatèl·lits i inversions cromosòmiques en poblacions europees i colonitzadores de *D. subobscura*. Diploma d'Estudis Avançats. Dept. Genètica, Universitat de Barcelona, Spain.
- Santos J, Serra L, Solé E, Pascual M. 2010. FISH mapping of microsatellite *loci* from *Drosophila subobscura* its comparison to related species. Chrom Res, in press.
- Schultz J, Redfield H, 1951. Interchromosomal effects on crossing-over in *Drosophila*. Cold Spring Harb Symp Quant Biol 16:175-197.
- Singh ND, Aquadro CF, Clark AG, 2009. Estimation of fine-scale recombination intensity variation in the white-echinus interval of *D. melanogaster*. J Mol Evol 69:42-53.
- Staten R, Schully SD, Noor MAF, 2004. A microsatellite linkage map of *Drosophila mojavensis*. BMC Genetics 5:12.
- True JR, Mercer JM, Laurie CC, 1996. Differences in crossover frequency and distribution among three sibling species of *Drosophila*. Genetics 142:507-523.

**II. Estudi de la influència de les
inversions en la formació
d'entrecreuaments al llarg del
cromosoma.**

3. 2. Estudi de la influència de les inversions en la formació d'entrecreuaments al llarg del cromosoma.

Publication 2: Recombination and selection in the maintenance of the adaptive value of inversions

Publicació 2: La recombinació i la selecció en el manteniment del valor adaptatiu de les inversions

RESUM

Una gran quantitat de dades semblen confirmar el valor adaptatiu de les inversions en el gènere *Drosophila*. En els heterocariotips, la inhibició de la recombinació produïda per les inversions sembla fonamental per la seva adaptació. Aquest estudi mostra que la recombinació en *D. subobscura* es troba fortament inhibida per les inversions cromosòmiques, no tan sols al seu interior sinó també fora d'elles. A més, sembla que la regió fora de la inversió on la recombinació és troba reduïda és asimètrica i independent de la longitud de la inversió. Tot i la dificultat que els entrecreuaments es formin a la vora dels punts de trencament, els dos únics esdeveniments recombinacionals que s'han detectat dintre de les inversions es troben localitzats vora el punt de trencament. Per tant, la selecció ha de jugar un paper important reduint la recombinació i mantenint combinacions d'al·lels adaptatives dins de la regió inhibida. En la descendència de tots els encreuaments es va trobar un major nombre d'individus heterocariotípics que homocariotípics, independentment que fossin cromosomes recombinants o no recombinants, a pesar que les línies provenien de localitats molt distants. Per tant, la hipòtesi de la coadaptació en sentit estricte s'hauria de descartar, ja que sembla l'eficàcia biològica dels individus heterocariotípics és superior, independentment de la seva població d'origen.

Recombination and selection in the maintenance of the adaptive value of inversions

CINTA PEGUEROLES^{*}, VÍCTOR ORDÓÑEZ, FRANCESC MESTRES AND MARTA PASCUAL

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avd. Diagonal 645, 08028, Barcelona, Spain.

*Corresponding author: cpegueroles@ub.edu

Abstract

A huge amount of data seems to confirm the adaptive value of inversions in *Drosophila*. The inhibition of recombination in heterokaryotypes mediated by inversions seems fundamental in maintaining their adaptive role. The present study shows that recombination is highly suppressed in *D. subobscura* due to chromosomal inversions, not only inside the inversions but also outside them. It seems that the region outside the inversion where recombination is inhibited is asymmetrical and independent of the inversion length. Despite the difficulty of crossovers taking place near inversion breakpoints, the only two recombination events detected inside inversions were located very close to the breakpoint. Thus, selection could be largely responsible for the recombination reduction maintaining sets of adaptive alleles inside the inverted region. Heterokaryotype descendants were always in higher frequency than homokaryotypes, whether recombinant or non-recombinant, even though the strains came from very distant populations. Consequently, the strict coadaptation hypothesis can be discarded although genes in different arrangements seem to confer higher fitness on heterozygote individuals, regardless of the origin of the chromosome.

Key words: *Drosophila subobscura*, microsatellites, inversions, heterokaryotypes, crossover, heterosis

Introduction

Chromosomal inversions are defined as a type of aberration in which a segment of the chromosome is reversed. Inversions have been extensively observed when genomes of closely related species are compared, suggesting that rearrangements play an important role in genome evolution (Clark et al., 2007, Hoffmann & Rieseberg, 2008). Although chromosomal inversions are generally thought to be formed by unique non-recurrent events, about 60% of *Drosophila* species are polymorphic for paracentric inversions (Powell, 1997). Several hypotheses have been put forward to explain the fixation or maintenance of different gene arrangements in natural populations (reviewed in Hoffmann & Rieseberg, 2008). The genome comparison of different species revealed high number of species specific inversions suggesting that chromosomal inversions could be a key factor in speciation processes (Ayala & Coluzzi, 2005). For instance, inversions could enhance the accumulation of genes causing reproductive isolation (Noor et al., 2001). Furthermore, inversion polymorphism has been shown to be adaptive in response to environmental changes (Prevosti et al., 1988, Balanyà et al., 2006) conferred by the presence of sets of adaptive genes included within the inversion (Dobzhansky, 1950, Kirkpatrick & Barton, 2006). Inversions would play an important role in maintaining adaptive gene content, probably through recombination reduction (Rozas et al., 1999, Schaeffer et al., 2003, Munté et al., 2005).

In early studies, it was detected that chromosomal inversions were responsible for crossover suppression (Sturtevant, 1926, Sturtevant & Plunkett, 1926). In paracentric inversions, which are the most common ones in natural populations, odd crossovers inside the inverted segment in heterokaryotypes lead to unbalanced chromosomes with deletions and duplications. Since, in *Drosophila*, these abnormal chromatids are selectively eliminated during the formation of the polar body nuclei, its functional egg nucleus receives only normal chromosomes (Sperlich & Pfried, 1986). However, even crossovers inside inversions produce balanced chromosomes, although there could be constraints on their formation. For instance, in heterokaryotypes, the non-correct synapsis near breakpoints seems responsible for crossover reduction, while higher levels are expected in central parts of inversions (Coyne et al., 1993, Navarro & Ruiz, 1997). Thus, crossover formation could not be uniform along the inverted segment, contrary to gene conversion (Navarro et al., 1997). The length of the chromosomal segment with reduced recombination has been reported to vary with

inversion length and location, whether in the proximal or distal region of the inversion (Roberts, 1976). The presence of inversions can also modify recombination elsewhere in the genome, whether in the same chromosome or even in others (White & Morley, 1955, Krimbas, 1992). For instance, in *Drosophila* heterokaryotypes it was observed that recombination increases in the inverted chromosome in regions far away from the inversion (Komai & Tatak, 1940). Moreover, the Schultz and Redfield effect describes how in the presence of inversion heterokaryotypes, recombination could increase in the other chromosomes of the karyotype (Schultz & Redfield, 1951).

Given its rich chromosomal inversion polymorphism, *Drosophila subobscura* is a valuable model species for analyzing the effect of inversions on recombination (Krimbas, 1993). The O chromosome is particularly rich in inversion polymorphism, with 26 inversions identified. Early studies in *D. subobscura* using chromosome inversions as markers suggested that recombination between distant inversions was suppressed. For instance, no recombinants were obtained in heterokaryotypes O_{3+4+1}/O_{ST} (Krimbas & Zouros, 1969) or O_{3+4}/O_5 (Mestres *et al.*, 1998), separated by approximately 50 cytological bands. Furthermore, no recombination was observed between linked non-overlapping inversions in heterokaryotypes O_{3+4+7}/O_{st} (Krimbas & Zouros, 1969, Sperlich & Feuerbach-Mravlag, 1974) or $O_{3+4}/O_{3+4+8+7}$ (Mestres *et al.*, 1998) in spite of the larger physical distance between the non-overlapping inversions (around 72 cytological bands). This lack of recombinant chromosomes could be due to either the complete absence of crossovers throughout the chromosome, the low resolution of the markers used or selection acting against recombinants. The lack of recombinants between inversions observed in those laboratory studies, is not indicative of recombination prevented on the whole O chromosome, as reported by studies detecting recombination between inversions and allozyme loci located close to its breakpoint (Loukas *et al.*, 1980, Mestres *et al.*, 1990).

The present research aims to assess by a molecular chromosome wide approach the effect of inversions on reducing recombination. More specifically, we analyse whether recombination is inhibited inside and outside inversions in three types of *D. subobscura* heterokaryotypes by using 12 microsatellite markers distributed along the O chromosome. We estimate the length of recombination reduction outside inversions and relate it to the length of inversions by using heterokaryotypes differing in both length

and number of inversions. Finally, we discuss the evolutionary implications of inversions through their effect on recombination in natural populations and their role on adaptation.

Material and methods

Fly samples and DNA genotyping

Isochromosomal lines from natural populations of *D. subobscura* were made homozygous by crossing flies with the *chcu* (*cherry curled*, homokaryotypic strain) and the *Va/Ba* (*Varicose/Bare*, a balanced lethal strain), as previously described (Mestres *et al.*, 1990, Mestres *et al.*, 1998). In some cases, due to the presence of lethal genes in the O chromosome, homokaryotypic lines could not be obtained and, consequently, wild chromosomes were maintained in heterokaryotypic condition with the *Va* balancer chromosome. In this study, recombination dynamics in three types of heterokaryotypes (O_{3+4+7}/O_{3+4+8} , O_{3+4+7}/O_{3+4} and O_{3+4+6}/O_{3+4}) (Figure 1) was analyzed, using 12 microsatellite markers (Table 1) located in the O chromosome (Pascual *et al.*, 2000) (Santos *et al.*, 2010). Lines O_{3+4} and O_{3+4+8} were collected from Mt. Parnes in Greece, lines O_{3+4+7} from Barcelona in Spain (Araúz *et al.*, 2009) and line O_{3+4+6} was collected from Sunne in Sweden (Ruiz-Martín, 2006). The chromosomal lines were selected for their different microsatellite allele sizes, which enabled the genotype of the offspring to be distinguished. Recombinant F₂ offspring were obtained by backcrossing virgin F₁ O_{3+4+6}/O_{3+4} females with males from the MP13 line, which was randomly selected from both parental lines. For the other two crosses, the virgin F₁ O_{3+4+7}/O_{3+4+8} and O_{3+4+7}/O_{3+4} females were crossed with males from the MP101 line, which was selected to detect the null alleles in the lethal O_{3+4+7} lines (Table 1). We genotyped the microsatellite loci in 209 F₂ descendants of F₁ O_{3+4+7}/O_{3+4+8} females, 183 F₂ descendants of F₁ O_{3+4+7}/O_{3+4} females and 167 F₂ descendants of F₁ O_{3+4+6}/O_{3+4} females.

Genomic DNA from each individual was extracted using the protocol of (Gloor *et al.*, 1993). The different primers used and PCR conditions for multiplex amplification are described elsewhere (Pascual *et al.*, 2000, Pegueroles *et al.*, 2010). PCR fragment analyses were carried out with a 3730 Analyzer (Applied Biosystems), at the Serveis Científic-Tècnics of the Universitat de Barcelona. Finally, allele sizes were assigned using GeneMapper 3.7 (Applied Biosystems) by reference to the internal standard Rox 70-500 (BioVentures).

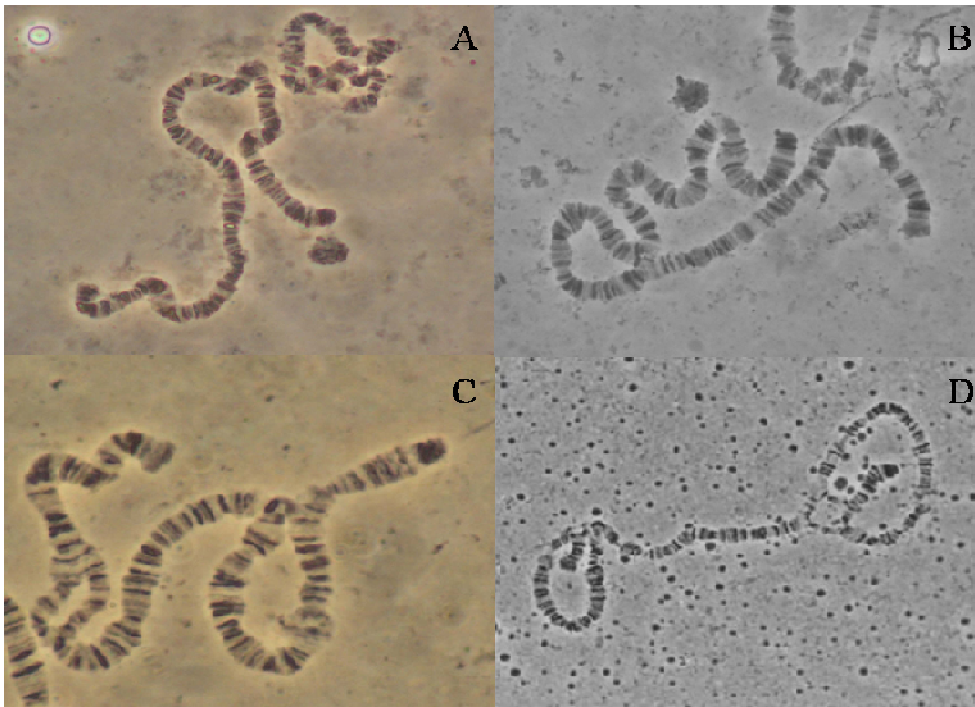


Figure 1. Polytene chromosomes of third instar larvae preparations of O_{3+4}/O_{3+4} homokaryotype (A) and the three types of heterokaryotypes used: O_{3+4+6}/O_{3+4} (B), O_{3+4+7}/O_{3+4} (C) and $O_{3+4+8+7}/O_{3+4}$ (D).

Table 1. Chromosomal arrangements and allele sizes (bp) of the twelve microsatellite loci for the lines used in the different crosses.

	BC36 × MP27	S07ABB × MP13	BC50 × MP58	MP101
	O_{3+4+7}/Va	O_{3+4+6}	O_{3+4+7}/Va	O_{3+4}
dsub26	475	441	450	477
dsub02	N	226	N	231
dsub25	252	252	252	243
dsub51	N	400	391	398
dsub14	N	120	N	N
dsub47	146	146	137	144
dsub38	419	415	438	413
dsub71	263	257	263	255
dsub01	300	298	300	263
dsub12	274	264	276	260
dsub29	248	253	257	257
dsub04	196	204	196	194

N = null allele. MP101 males were used to obtain F_2 offspring for some crosses (see Material and Methods)

Data analysis

Recombination frequencies between molecular markers were computed for each heterokaryotype cross by MapMaker software (Lander *et al.*, 1987). Moreover, for comparison, values previously obtained for homokaryotype O_{3+4}/O_{3+4} (Pegueroles *et al.*, 2010) were recalculated, excluding the two markers not used in the analyses of the heterokaryotypes. Physical distances between markers (Mb) were calculated under the assumption that all cytological bands of the O chromosome of *D. subobscura* contain the same DNA content (Schäfer *et al.*, 2010) and that this chromosome is equivalent in length to chromosome 2 of *D. pseudoobscura* (Munté *et al.*, 2005, Pegueroles *et al.*, 2010). Chi-square tests were carried out with Statistica v6 software.

Results

We covered 27 Mb of the 31 Mb of the O chromosome, which is the estimated physical distance between the most proximal (dsub26) and the most distal (dsub04) microsatellite markers. Recombination frequencies of the three types of heterokaryotypes analyzed are shown in Table 2. Recombination behavior differs substantially between the homokaryotype and the three types of heterokaryotypes (Table 2), since recombination was clearly affected by inversions in heterokaryotypes. When comparing the number of recombinant individuals between consecutive markers in heterokaryotypes and homokaryotypes with a χ^2 test, we distinguished three regions: regions where recombination is inhibited ($P < 0.03$), regions not affected by inversions ($P > 0.17$) and regions with reduced recombination (Table 2). Despite showing a significant reduction compared to homokaryotypes (Figure 1 and Table 2), the region between markers dsub29-dsub04 for O_{3+4+6}/O_{3+4} was thought to have reduced recombination, since a fair number of recombinants were found.

In O_{3+4+7}/O_{3+4+8} heterokaryotypes, recombination was suppressed across the whole chromosome, with the exception of one double recombinant out of 208 individuals near the proximal breakpoint of the O_8 inversion (Table 2). For this individual, one crossover took place between markers dsub01-dsub12 and a second crossover between dsub12-dsub29. Thus, to give a balanced chromosome and explain the viability of the recombinant, the first recombination event had to occur between the proximal breakpoint and locus dsub12 (Figure 2). Recombination in O_{3+4+6}/O_{3+4} and

O_{3+4+7}/O_{3+4} heterokaryotypes was significantly reduced in 63% and 65% of the O chromosome, respectively (Table 2). Recombinants inside O_6 inversion were not found in O_{3+4+6}/O_{3+4} heterokaryotypes. However, one double recombinant out of 391 individuals was found inside the O_7 inversion in O_{3+4+7}/O_{3+4} heterokaryotypes (Table 2). For this recombinant, one crossover occurred between markers dsub14-dsub47 and the second crossover was detected between markers dsub47-dsub38. It is worth noting that dsub38 is located outside the inversion. Consequently, to explain the viability of this individual, the crossover had to occur between locus dsub47 and the distal breakpoint of inversion O_7 . In spite of the low number of recombinants inside the inversions detected in heterokaryotypes, the even number of recombinants in the same regions in O_{3+4} homokaryotypes was much higher. In O_{3+4} homokaryotypes, 40 recombinants with an even number of crossovers were inferred out of 305 homokaryotypic individuals for the O_7 region, and 18 individuals with an even number of crossovers were inferred out of 327 individuals in the O_8 region. These results are consistent with the former inversion including a larger chromosomal segment. The inferred values can be considered a good estimate of the expected number of recombinants with even number of crossovers within each inversion. The comparison of the observed and the expected values showed significant differences in both the O_7 ($\chi^2 = 44.91$, $P = 0$) and the O_8 region ($\chi^2 = 8.83$, $P = 0.003$).

Table 2. Recombination frequency between consecutive markers in the homokaryotype and heterokaryotypes.

	O_{3+4}/O_{3+4}	O_{3+4+6}/O_{3+4}	O_{3+4+7}/O_{3+4}	O_{3+4+7}/O_{3+4+8}
	MP16/MP13	S07ABB/MP13	BC50/MP58	BC36/MP27
dsub26 / dsub02	0.194	0.174	0.000*	0.000*
dsub02 / dsub25	0.078	0.072	0.000*	0.000*
dsub25 / dsub51	0.103	0.090	0.000*	0.000*
dsub51 / dsub14	0.176	0.174	0.000*	0.000*
dsub14 / dsub47	0.039	0.018	0.005*	0.000*
dsub47 / dsub38	0.345	0.000*	0.005*	0.000*
dsub38 / dsub71	0.178	0.000*	0.000*	0.000*
dsub71 / dsub01	0.111	0.000*	0.071	0.000*
dsub01 / dsub12	0.184	0.000*	0.186	0.005*
dsub12 / dsub29	0.178	0.013*	0.158	0.005*
dsub29 / dsub04	0.229	0.089*	0.171	0.000*

Inverted regions appear in grey. All χ^2 comparison tests are done in relation to the homokaryotypes (* $P < 0.05$).

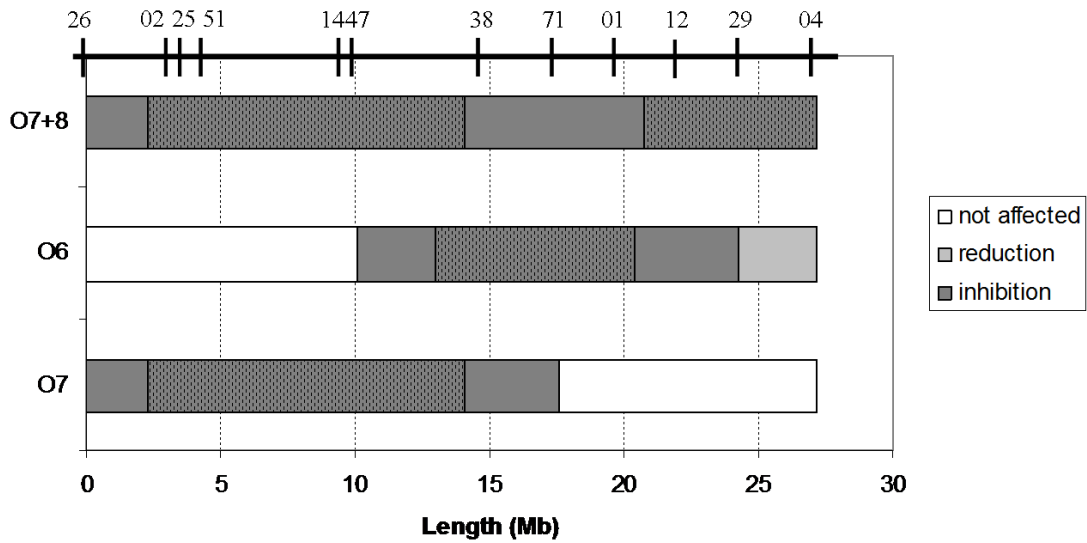


Figure 2. Diagram showing regions where recombination frequency has been inhibited, reduced or is not affected in the three types of heterokaryotypes relative to the chromosomal length. At the top, microsatellite locations are shown. Stripped regions identify inversions.

Outside O_6 inversion that is centrally located, an asymmetrical recombination reduction, larger in the distal region, was detected. Asymmetry could not be assessed in the O_7 inversions, since it is located more closely to the centromere. Although both inversions differed in length ($O_6 = 7.5$ Mb and $O_7 = 11.7$ Mb), the distal regions outside inversions in which recombination was inhibited were quite similar (Figure 2). Thus, the ratio between the distal region with inhibited recombination and inversion length was larger in inversion O_6 (52%) than inversion O_7 (30%).

The proportion of chromosomal arrangements in the offspring of F_1 heterokaryotypic females crossed with homokaryotypic O_{3+4} males was evaluated to assess whether each arrangement had the expected 50% frequency. Descendants of O_{3+4+7}/O_{3+4+8} females had no significant deviations, since 48% of individuals had the O_{3+4+8} arrangement and 52% had the O_{3+4+7} ($\chi^2 = 0.09$, $P = 0.77$ with Yates correction). Interestingly, non-equivalent segregation was found in the other two crosses, since significantly fewer descendants carrying the O_{3+4} arrangement were found, and these individuals were homozygous, since heterokaryotypic females had been crossed with homokaryotypic O_{3+4} males. In O_{3+4+6}/O_{3+4} heterokaryotypes, only 36% of the offspring carried the O_{3+4} arrangement ($\chi^2 = 4.17$, $P < 0.04$); and in O_{3+4+7}/O_{3+4} heterokaryotypes 38% of them did ($\chi^2 = 6.71$, $P < 0.01$ with Yates correction). Furthermore, reduction of O_{3+4} homozygous individuals in the offspring was observed in both recombinant and

non-recombinant chromosomes, with the same proportion in both the crosses involving the O_{3+4+6} ($\chi^2 = 0.99$, $P = 0.32$) and the O_{3+4+7} arrangement ($\chi^2 = 0$, $P = 1$).

Discussion

Recombination reduction mediated by inversions

Our results support the hypothesis that inversions effectively reduce recombination along the chromosome in *D. subobscura*, since recombination was almost completely inhibited in heterokaryotypes not only within inversions but also outside them. However, total inhibition of recombination was not observed in *Anopheles gambiae*, although a significant reduction of the recombination rate was reported both inside and outside inversions (Stump *et al.*, 2007). In an early review of the effects of heterozygosity for paracentric inversions, White & Morley (1955) indicated that suppression or drastic reduction of crossing-over outside the inverted region was a well-established pattern in *Drosophila*. Different factors might cause the recombination reduction mediated by inversions. First, it is known that only even numbers of crossovers inside inversions between the same pair of chromatids produce normal meiotic products, while otherwise recombinant chromosomes show duplications and deletions (Roberts, 1976). Second, it is sometimes possible to observe that synapsis between two homologous chromosomes is not perfect close to inversion breakpoints in polytene chromosome preparations (Figure 1D), which might interfere with their ability to recombine. Finally, selection favoring specific gene combinations within inversions could eliminate recombinant chromosomes rendering them maladaptive (Dobzhansky, 1950, Kirkpatrick & Barton, 2006, Hoffmann & Rieseberg, 2008).

According to (Krimbas, 1992), short and medium-sized inversions located from the medial to the distal part of the chromosome are the most effective recombination suppressors. This suggestion is consistent with the observed results, since O_6 , shorter and more centrally located than O_7 has proportionally more recombination inhibition outside the inversion. However, in both inversions, recombination was inhibited in approximately the same length in the region outside the distal breakpoint (3.5Mb and 3.9Mb for O_7 and O_6 inversions, respectively), despite both inversions having different lengths. Thus, it could be considered that the region inhibited outside the inversion is independent of the inversion length and just a by-product of the difficulty of

heterochromosomes to synapse correctly near the breakpoint, as previously suggested (Roberts, 1976).

Regions outside the breakpoints seem to behave in a different way with higher reduction of recombination in the distal than in the proximal region. This asymmetry had already been reported in other species although with contrasting results. In some *Drosophila* species (e.g. *D. pseudoobscura*), reduction was higher in the distal segment (Dobzhansky & Epling, 1948, Roberts, 1962), while in others (e.g. *D. robusta*) it was higher in the proximal segment (Carson, 1953). In *Anopheles gambiae*, recombination reduction was more drastic in the proximal region in one of the autosomes and the sex chromosome, and this reduction is claimed to be involved in the speciation process (Stump *et al.*, 2005, Stump *et al.*, 2007). Asymmetry was also detected in *Trimerotropis* grasshoppers with higher recombination reduction in the proximal region (White & Morley, 1955). Thus, it can be concluded that asymmetrical reduction of recombination outside inversions is a general pattern in insects, although the region suffering greater reduction can be species-specific.

The presence of inversion heterokaryotypes in *Drosophila* has been shown to increase recombination in chromosomal regions far away from the inversion (Komai & Tatak, 1940, Grell, 1962, Lucchesi & Suzuki, 1968, Krimbas, 1992). However, in our study recombination frequency between consecutive markers in regions not affected by the inversions was not significantly higher in heterokaryotypes than in homokaryotypes (Table 2). Pegueroles *et al.* (2010) reported significant differences in recombination frequency between homokaryotypes for different inversions that could be caused by differences among arrangements, differences between lines or by the Schultz and Redfield effect (i.e. recombination increase due to an interchromosomal effect). The Schultz and Redfield effect hypothesizes that the recombination frequency of a chromosome increases if other chromosomes of the karyotype are heterozygotes for inversions. Lines coming from different populations could have different levels of heterozygosity for inversions (Krimbas, 1993). In the present study, recombination in the region not affected by inversions in outbred crosses (i. e. using lines from different populations) was not significantly different compared to recombination in homokaryotypic lines, which is consistent with the similar expected heterozygosity values in F_1 females. In the homokaryotypic crosses, the expected heterozygosity for the

F₁ females is 0.55 (Mt. Parnes/Mt. Parnes), similar to the heterozygosity values expected for the F₁ females from the heterokaryotypic crosses (Mt. Parnes-Sunne = 0.58, Mt. Parnes-Barcelona = 0.54). However, in Pegueroles *et al.* (2010) the O_{ST} homokaryotypic lines coming from the North had lower recombination, which could be in agreement with the Schultz and Redfield effect with northern lines having lower mean expected heterozygosity (Sunne = 0.23) than southern lines (Mt. Parnes = 0.55). Thus, our results lead to the suggestion that observed differences between homokaryotypes are not due to differences between lines, as reported for *D. pseudoobscura* (Levine & Levine, 1955), but could rather be caused by the type of chromosomal arrangement or the Schultz and Redfield effect.

Maintenance of inversions adaptive value

Previous studies suggest that at least one crossover per bivalent could be required to have a correct chromosomal meiotic disjunction (Nicklas, 1974). However, other studies thought that chromosomal exchange was not needed for normal segregation in *Drosophila* (Sturtevant & Beadle, 1936, Hawley *et al.*, 1993). According to our data, the possibility of correct chromosome segregation without chiasma formation seems a very plausible explanation in heterokaryotypes since 43.4% and 99.9% non-recombinant chromosomes were detected between chromosomes differing by one or two inversions respectively. It might be that formation of an inversion loop is enough to ensure correct segregation of chromosomes.

However, the classic hypothesis that chiasmata have occurred but then recombinant chromosomes were preferentially segregated to the polar bodies, as suggested by Beadle & Sturtevant (1935), cannot be ruled out. Previous studies reported the absence of recombination in *D. subobscura* heterokaryotypes (Krimbas & Zouros, 1969, Sperlich & Feuerbach-Mravlag, 1974, Mestres *et al.*, 1998). By direct observation of chromosomes through dissection of third instar larvae, no recombination was detected in 139 O₃₊₄₊₁/O_{ST} flies, 49 O₃₊₄₊₁₂/O_{ST} ones, 102 O₃₊₄₊₇/O_{ST} ones (Krimbas & Zouros, 1969), 366 individuals for the same heterokaryotypes (O₃₊₄₊₇/O_{ST}) (Sperlich & Feuerbach-Mravlag, 1974), 415 O₃₊₄₊₈₊₇/O₃₊₄ flies and 437 O₅/O₃₊₄ ones (Mestres *et al.*, 1998). These studies suggest that breaking an inversion combination by recombination is a very infrequent event, but recombinant arrangements have been observed in nature, though not often. For instance, O₃₊₄₊₈₊₇ chromosomes were found at very low

frequencies (0.001% to 0.763%) in three European populations (Fontdevila *et al.*, 1983, Mestres *et al.*, 1990, Orengo & Prevosti, 1996). Another infrequent arrangement that can be seldom detected in nature is $O_{\underline{3+4+5}}$ (Krimbas, 1964) (Zivanovic & Marinkovic, 2003). Analogously, $O_{\underline{3+4+6}}$ is found in very low frequency in northern (Balanyà *et al.*, 2003, Ruiz-Martín, 2006) and some Mediterranean populations (Sperlich & Feuerbach, 1969, Zivanovic *et al.*, 2002). The formation of O_7 inversion by recombination from heterokaryotypes $O_{\underline{3+4+7}}/O_{ST}$ has been unambiguously proved in colonizing American populations of *D. subobscura*. The American O_7 inversion had the same lethal gene and the same *Odh* gene haplotype found in some $O_{\underline{3+4+7}}$ chromosomes (Gómez-Baldó *et al.*, 2008, Mestres *et al.*, 2009). Furthermore, in American populations microsatellites typically in linkage disequilibrium with the $O_{\underline{3+4+7}}$ and the O_{ST} arrangements (Simões *et al.*, 2009) are found in the O_7 chromosomal line confirming its origin by recombination (data not shown).

New inversions are hardly ever found in *D. subobscura* natural populations (Balanyà *et al.*, 2003), and their fate, due to their low frequency when formed is mainly determined by genetic drift (Powell, 1997). However, new gene arrangements could spread further in populations due to selection by different mechanisms, such as coadapted gene complexes (Dobzhansky, 1950) and local adaptation (Kirkpatrick & Barton, 2006, Hoffmann & Rieseberg, 2008). The most likely explanation for the low frequency of arrangements generated by recombination in natural populations is that they are non-adaptive because some form of epistasis is broken or synthetic lethality is generated. Sperlich and Feuerbach-Mravlag (1974) considered that epistatic interactions are relevant in maintaining the linkage disequilibrium between non-overlapping inversions in *D. subobscura*. Other authors have suggested that this strong linkage disequilibrium could be attributed to the presence of strong selective forces producing epistatic interactions or to the absence of recombination in heterozygote females (Krimbas & Zouros, 1969, Krimbas, 1993). Our results seem to reject the later hypothesis since at least one double crossover was detected in descendants of $O_{\underline{3+4+7}}/O_{\underline{3+4+8}}$ females.

Theoretical studies point out that recombination will be more efficiently inhibited near inversion breakpoints, where gene flux would be primarily due to gene conversion, whereas in the central part of the inversion both gene conversion and

crossovers could contribute to gene flux (Navarro *et al.*, 1997). In heterokaryotypes, the reduction in the observed number of double recombinants within the inversion in relation to their expected number in homokaryotypes could be due to a lower frequency of chiasmata formation inside inverted regions or to purging of recombinant chromosomes, i.e. by differential viability or meiotic drive, with non-recombinants segregating preferentially to the oocyte. The only two recombinant events detected in the proximities of the inverted region could be explained by gene conversion or double crossovers. However, since gene conversion affects a reduced region (i. e. 122 bp tract length on average in *D. subobscura*, (Betrán *et al.*, 1997)), the probability of converting this microsatellite region seems very unlikely. Thus, two double crossovers could have taken place close to inversion breakpoints suggesting that inhibition in those regions may not be as relevant as previously thought.

In the present study, a significant deviation from random segregation expectations was detected in two out of three crosses with a deficiency of homozygous descendants whether recombinant or not. This deviation was previously observed in *D. subobscura* and three hypotheses were postulated: a greater fitness of individuals carrying a certain inversion, meiotic drive and the heterotic effect of the wild chromosome in combination with the chromosome from a laboratory inbred strain (Mestres *et al.*, 1998). The meiotic drive hypothesis was recently ruled out in *D. pseudoobscura* with experiments that used five different arrangements involving reciprocal crosses (Meisel & Schaeffer, 2007). Their observations were consistent with selection maintaining arrangement polymorphism by fitness benefits from coadapted gene complexes, which would be maintained by suppressed recombination between arrangements. A decrease in fitness in recombinants might be explained by breaking down sets of coadapted gene complexes (Dobzhansky, 1950) or sets of genes individually adapted to local conditions (Kirkpatrick & Barton, 2006). Under the coadaptive hypothesis, inversion heterozygotes are adaptively superior to the corresponding homozygotes only if the chromosome pair has been drawn from the same or neighboring localities (Dobzhansky & Epling, 1948). This is not the case in our experiment since heterokaryotype descendants are always more frequent regardless of the origin of the lines used: they come from very distant populations, such as Greece, Spain or Sweden. For instance, no fitness differences were detected in the offspring of the O_{3+4+7}/O_{3+4+8} crossed to an O_{3+4} homokaryotypic line from Mt. Parnes, despite

O₃₊₄₊₇ chromosome came from Barcelona and O₃₊₄₊₈ chromosome came from Mt. Parnes. Finally, two types of offspring can be distinguished, those obtained by backcrossing with an O₃₊₄ parental line and those obtained from crosses with a different O₃₊₄ line. Reduction in the number of homokaryotypes in the offspring is similar in both cases, which indicates that the reduction is independent of individuals being inbred or outbred for the O chromosome. The overdominance observed could be produced by the presence of recessive deleterious mutations, which decreased fitness in homozygous individuals (Kirkpatrick & Barton, 2006). However, O₃₊₄ homokaryotypic stocks maintained in the laboratory do not show reduced viability. Thus, our results can be better explained by a heterotic effect with greater viability of heterozygotes independently of the origin of the chromosome.

In conclusion, our studies with the O chromosome of *D. subobscura* support both the local adaptation hypothesis (Kirkpatrick & Barton, 2006) and, in part the coadaptation hypothesis (Dobzhansky, 1950). Genes maintained in different arrangements seem to confer higher fitness to heterozygous individuals regardless of the origin of the chromosome. Thus, different arrangements, though they come from different populations and have a different set of alleles for neutral markers, would harbor the same adaptive alleles (captured in the inversion), granting adaptation success to a given environment.

Acknowledgements

We are grateful Joan Balanyà for the analyses of the polytene chromosomes and his comments and suggestions, Pedro A. Araúz and Héctor Ruiz-Martín for collecting the flies and Concepció Arenas for her help in the statistical analyses. We also thank Pedro Simões and Ferran Palero for their valuable comments and suggestions on the manuscript. This study was supported by a pre-doctoral fellowship to CP (2009FIC-00096) from the Generalitat de Catalunya (Spain). Research was funded by projects CGL2006-13423-C02-02 of the Ministerio de Ciencia y Tecnología (MCYT, Spain) and 2009 SGR 636 of the Generalitat de Catalunya (Spain).

Literature cited

Araúz, P. A., Mestres, F., Pegueroles, C., Arenas, C., Tzannidakis, G., Krimbas, C. B. & Serra, L. 2009. Tracking the origin of the American colonization by *Drosophila*

- subobscura*: genetic comparison between Eastern and Western Mediterranean populations. *Journal of Zoological Systematics and Evolutionary Research* **47**: 25-34.
- Ayala, F. J. & Coluzzi, M. 2005. Chromosome speciation: Humans, *Drosophila*, and mosquitoes. *Proceedings of the National Academy of Sciences of the USA* **102**: 6535-6542.
- Balanyà, J., Oller, J. M., Huey, R. B., Gilchrist, G. W. & Serra, L. 2006. Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* **313**: 1773-5.
- Balanyà, J., Serra, L., Gilchrist, G. W., Huey, R. B., Pascual, M., Mestres, F. & Sole, E. 2003. Evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. *Evolution* **57**: 1837-45.
- Beadle, G. W. & Sturtevant, A. H. 1935. X-chromosome inversions and meiosis in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the USA* **21**: 384-390.
- Betrán, E., Rozas, J., Navarro, A. & Barbadilla, A. 1997. The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* **146**: 89-99.
- Carson, H. L. 1953. The effect of inversions on crossing over in *Drosophila robusta*. *Genetics* **38**: 168-186.
- Clark, A. G., Eisen, M. B., Smith, D., *et al.* 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**: 203-218.
- Coyne, J. A., Meyers, W., Crittenden, A. P. & Sniegowski, P. 1993. The fertility effects of pericentric inversions in *D. melanogaster*. *Genetics* **134**: 487-496.
- Dobzhansky, T. 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35**: 288-302.
- Dobzhansky, T. & Epling, C. 1948. The suppression of crossing-over in inversion heterozygotes of *Drosophila pseudoobscura*. *Proceedings of the National Academy of Sciences of the USA* **34**: 137-141.
- Fontdevila, A., Zapata, C., Alvarez, G., Sanchez, L., Mendez, J. & Enriquez, I. 1983. Genetic coadaptation in the chromosomal polymorphism of *Drosophila*

- subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. *Genetics* **105**: 935-955.
- Gloor, G. B., Preston, C. R., Johnson-Schlitz, D. M., Nassif, N. A., Phillis, R. W., Benz, W. K., Robertson, H. M. & Engels, W. R. 1993. Type I repressors of P element mobility. *Genetics* **135**: 81-95.
- Gómez-Baldó, L., Latorre, A., Serra, L. & Mestres, F. 2008. Molecular variation in the *Odh* gene in Chilean natural populations of *Drosophila subobscura*. *Hereditas* **145**: 154-162.
- Grell, R. F. 1962. A new model for secondary nondisjunction: the role of distributive pairing. *Genetics* **47**: 1737-1754.
- Hawley, R. S., McKim, K. S. & Arbel, T. 1993. Meiotic segregation in *Drosophila melanogaster* females: molecules, mechanisms, and myths. *Annual Review of Genetics* **27**: 281-317.
- Hoffmann, A. A. & Rieseberg, L. H. 2008. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics* **39**: 21-42.
- Kirkpatrick, M. & Barton, N. 2006. Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419-434.
- Komai, T. & Tatak, T. 1940. Two independent inversions in the X-chromosome of *D. virilis* and their effects on crossing over and disjunction. *Cytologia* **11**: 245-260.
- Krimbas, C. B. 1964. The genetics of *Drosophila subobscura* populations. I. Inversion polymorphism in populations of southern Greece. *Evolution* **18**: 541-552.
- Krimbas, C. B. (1992) The inversion polymorphism of *Drosophila subobscura*. In: *Drosophila Inversion Polymorphism* (Krimbas, C. B. & Powell, J. R., eds.). pp. 127-220. CRC Press, Inc., Boca Raton, Florida.
- Krimbas, C. B. 1993. *Drosophila subobscura: Biology, Genetics and Inversion polymorphism*. Verlag Dr. Kovac, Hamburg.
- Krimbas, C. B. & Zouros, E. 1969. Crossing over suppression between linked but non overlapping inversions in *Drosophila subobscura*. *Drosophila Information Service* **44**: 71-72.
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E. & Newburg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**: 174-181.

- Levine, R. P. & Levine, E. E. 1955. Variable crossing over arising in different strains of *Drosophila pseudoobscura*. *Genetics* **40**: 399-405.
- Loukas, M., Krimbas, C. B. & Sourdis, J. 1980. The genetics of *Drosophila subobscura* populations. XIII. A study of lethal allelism. *Genetica* **54**: 197-207.
- Lucchesi, J. C. & Suzuki, D. T. 1968. The interchromosomal control of recombination. *Annual Review of Genetics* **2**.
- Meisel, R. P. & Schaeffer, S. W. 2007. Meiotic transmission of *Drosophila pseudoobscura* chromosomal arrangements. *PLoS One* **2**: e530.
- Mestres, F., Sanz, J. & Serra, L. 1998. Chromosomal structure and recombination between inversions in *Drosophila subobscura*. *Hereditas* **128**: 105-13.
- Mestres, F., Pegueroles, G., Prevosti, A. & Serra, L. 1990. Colonization of America by *Drosophila subobscura*: lethal genes and the problem of the O₅ inversion. *Evolution* **44**: 1823-1836.
- Mestres, F., Balanya, J., Pascual, M., Arenas, C., Gilchrist, G. W., Huey, R. B. & Serra, L. 2009. Evolution of Chilean colonizing populations of *Drosophila subobscura*: lethal genes and chromosomal arrangements. *Genetica* **136**: 37-48.
- Munté, A., Rozas, J., Aguadé, M. & Segarra, C. 2005. Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* **169**: 1573-81.
- Navarro, A. & Ruiz, A. 1997. On the fertility effects of pericentric inversions. *Genetics* **147**: 931-3.
- Navarro, A., Betran, E., Barbadilla, A. & Ruiz, A. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695-709.
- Nicklas, R. B. 1974. Chromosome segregation mechanisms. *Genetics* **78**: 205-213.
- Noor, M. A. F., Grams, K. L., Bertucci, L. A. & Reiland, J. 2001. Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the USA* **98**: 12084-12088.
- Orengo, D. J. & Prevosti, A. 1996. Temporal changes in chromosomal polymorphism of *Drosophila subobscura* related to climatic changes. *Evolution* **50**: 1346-1350.
- Pascual, M., Schug, M. D. & Aquadro, C. F. 2000. High density of long dinucleotide microsatellites in *Drosophila subobscura*. *Molecular Biology and Evolution* **17**: 1259-67.

- Pegueroles, C., Araúz, P. A., Pascual, M. & Mestres, F. 2010. A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*. *Genetica*, in press.
- Prevosti, A., Ribó, G., Serra, L., Agudé, M., Balanyà, J., Monclús, M. & Mestres, F. 1988. Colonization of America by *Drosophila subobscura*: Experiment in natural populations that supports the adaptive role of chromosomal inversion polymorphism. *Proceedings of the National Academy of Sciences USA* **85**: 5597-5600.
- Powell, J. R. 1997. Progress and prospects in evolutionary biology: The *Drosophila* model. *Oxford University Press, New York*.
- Roberts, P. A. 1962. Interchromosomal effects and the relation between crossing over and nondisjunction. *Genetics* **47**: 1691-1709.
- Roberts, P. A. (1976) The genetics of chromosome aberration. In: *The genetics and biology of Drosophila*, Vol. 1a. pp. 67-184. Academic Press, London, UK.
- Rozas, J., Segarra, C., Ribó, G. & Agudé, M. 1999. Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *Drosophila subobscura*. *Genetics* **151**: 189-202.
- Ruiz-Martín, H. (2006) Estudi de les associacions entre al·lels de loci microsatèl·lits i inversions cromosòmiques en poblacions europees i colonitzadores de *D. subobscura*. In: *Diploma d'Estudis Avançats. Dept. Genètica*, Universitat de Barcelona, Spain
- Santos, J., Serra, L., Solé, E. & Pascual, M. 2010. FISH mapping of microsatellite loci from *Drosophila subobscura* and its comparison to related species *Chrosomal Research*, in press.
- Schaeffer, S. W., Bernhardt, M. J. & Anderson, W. W. 2003. Evolutionary rearrangement of the amylase genomic regions between *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Journal of Heredity* **94**: 464-71.
- Schäfer, M. A., Mazzi, D., Klappert, K., Kauranen, H., Vieira, J., Hoikkala, A., Ritchie, M. G. & Schlötterer, C. 2010. A microsatellite linkage map for *Drosophila montana* shows large variation in recombination rates, and a courtship song trait maps to an area of low recombination. *Journal of Evolutionary Biology* **23**: 518-527.

- Schultz, J. & Redfield, H. 1951. Interchromosomal effects on crossing-over in *Drosophila*. *Cold Spring Harbor Symposia on Quantitative Biology* **16**: 175–197.
- Simões, P., Calabria, G., Picão-Osório, J., Naas, I., Serra, L., Balanyà, J. & Pascual, M. (2009) Inversion polymorphisms and linkage disequilibrium patterns in ancestral and invasive populations of *Drosophila subobscura*. 12th ESEB Congress, Turin (Italy).
- Sperlich, D. & Feuerbach-Mravlag, H. 1969. Austausch-ungleichgewicht zwischen unabhängigen inversionen in natürlichen populationen von *Drosophila subobscura*. [Recombination disequilibrium of independent inversions in natural populations of *Drosophila subobscura*.]. *Theoretical and Applied Genetics* **39**: 104-112.
- Sperlich, D. & Feuerbach-Mravlag, H. 1974. Epistatic gene interaction, crossing over, and linked and unlinked inversions in *Drosophila subobscura*. *Evolution* **28**: 67-75.
- Sperlich, D. & Pfriem, P. (1986) Chromosomal polymorphism in natural and experimental populations. In: *The genetics and biology of Drosophila* (eds. Ashburner, M., Carson, H. L. and Thompson Jr., J. N.), Vol. 3e. pp. 257-309. Academic Press, N. Y.
- Stump, A. D., Pombi, M., Goeddel, L., Ribeiro, J. M. C., Wilder, J. A., Torre, A. D. & Besansky, N. J. 2007. Genetic exchange in 2La inversion heterokaryotypes of *Anopheles gambiae*. *Insect Molecular Biology* **16**: 703-709.
- STUMP, A. D., M. C. FITZPATRICK, N. F. LOBO, S. TRAORE, N. F. SAGNON *et al.*, 2005 Centromere-proximal differentiation and speciation in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences USA* **102**: 15930-15935.
- Sturtevant, A. H. 1926. A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biologisches Zentralblatt* **46**: 697-702.
- Sturtevant, A. H. & Beadle, G. W. 1936. The relations of inversions in the X chromosomes of *Drosophila melanogaster* to crossing-over and disjunction. *Genetics* **28**: 554-604.
- Sturtevant, A. H. & Plunkett, C. R. 1926. Sequence of corresponding third-chromosome genes in *Drosophila melanogaster* and *D. simulans*. *Biological Bulletin* **50**: 56-60.

- White, M. J. D. & Morley, F. H. W. 1955. Effects of pericentric rearrangements on recombination in grasshopper chromosomes. *Genetics* **40**: 604–619.
- Zivanovic, G. & Marinkovic, D. 2003. Viabilities of originally natural O-chromosomal inversion homo- and heterokaryotypes in *D. subobscura*. *Hereditas* **139**: 128-142.
- Zivanovic, G., Andjelkovic, M. & Marinkovic, D. 2002. Chromosomal inversion polymorphism of *D. subobscura* from south-eastern part of Europe. *Journal of Zoological Systematics and Evolutionary Research* **40**: 201-204.

III. Diferenciació genètica entre poblacions naturals de *D. subobscura*.

3. 3. Diferenciació genètica entre poblacions naturals de *D. subobscura*.

Publication 3: Gene flow and gene flux shape evolutionary patterns of variation in *D. subobscura*

Publicació 3: El flux gènic determina els patrons evolutius de variació en *D. subobscura*

RESUM

El flux gènic és un factor que incrementa la variabilitat genètica dins de les inversions (*gene flux*) i disminueix la diferenciació genètica entre poblacions (*gene flow*). Les inversions cromosòmiques estan influenciades per la selecció, i el seu paper reduïnt la recombinació assegura el manteniment del seu contingut genètic. No obstant, segons la hipòtesi de la coadaptació de Dobzhansky, s'esperaria trobar diferenciació genètica entre inversions i poblacions. En el present estudi, hem comparat la diversitat nucleotídica entre dues poblacions de *Drosophila subobscura*, provinents dels dos extrems de la regió Mediterrània, Mount Parnes (Grècia) i Barcelona (Espanya), seleccionant sis gens localitzats al cromosoma O d'aquesta espècie: *Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* i *Fmr1*. En ambdues poblacions, l'anàlisi s'ha centrat en ordenaments que difereixen per dues inversions solapades, O_{ST} i O₃₊₄. El gen *Fmr1*, que és l'únic situat dintre de la regió invertida, presenta elevada diferenciació a nivell nucleotídic en els dos ordenaments. Per altra banda, no s'ha detectat diferenciació genètica significativa entre les poblacions de Mt. Parnes i Barcelona per cap dels sis gens, indicant alts nivells de flux gènic. Finalment, els ordenaments O_{ST} i O₃₊₄ s'haurien pogut formar aproximadament en el mateix període.

Gene flow and gene flux shape evolutionary patterns of variation in *D. subobscura*

CINTA PEGUEROLES^{1*}, ALBERT FERRÉS¹, CHARLES F. AQUADRO², FRANCESC MESTRES¹ AND MARTA PASCUAL¹

¹Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avd. Diagonal 645, 08028, Barcelona, Spain.

²Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

*Corresponding author: cpegueroles@ub.edu

Abstract

Gene flux and gene flow are important factors increasing genetic variability within inversions and decreasing genetic differentiation between populations. Chromosomal inversions are under strong selection and its role in recombination reduction ensures maintenance of its genetic content. Nevertheless, according to Dobzhansky's coadaptation hypothesis, genetic differentiation should be expected between different inversions as well as between populations for the same inversion. In order to get further insight into this subject, two *Drosophila subobscura* populations from both sides of the Mediterranean region, Mt. Parnes (Greece) and Barcelona (Spain) were analysed. In both populations, analyses were focused on two different arrangements differing by a couple of overlapping inversions, O_{ST} and O₃₊₄. Nucleotide diversity levels were obtained for six nuclear genes (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* and *Fmr1*) located across the O chromosome. Interestingly, the only gene located inside the inverted region (*Fmr1*), showed high nucleotide differentiation between both arrangements. However, no significant genetic differences were detected within arrangements between Mt. Parnes and Barcelona populations for none of the six genes, indicating high levels of gene flow between populations. Finally, molecular dating techniques indicate that the O_{ST} and O₃₊₄ arrangements could have originated in a similar time period.

Key words: *D. subobscura*, populations, nuclear genes, gene flow, inversions dating

Introduction

The Natural Selection theory, proposed by Charles Darwin in the XIX century, is one of the most important and more studied topics in biology. However, due to its complexity, the exact mechanisms involved in the origin of species are still unclear. Chromosomal inversions are clearly involved in natural selection, since their evolutionary significance has been studied using many groups of animals and plants. (reviewed in HOFFMANN and RIESEBERG 2008)). Genomic comparative studies have revealed the existence of numerous inverted regions between species suggesting an important role in the speciation process (AYALA and COLUZZI 2005; SCHAEFFER *et al.* 2008; SANTOS *et al.* 2010). Selection can maintain genetic differences between arrangements which will be enhanced by the recombination reduction inside inverted regions (NAVARRO *et al.* 1997; STUMP *et al.* 2007). Inversions polymorphism has been described in several *Drosophila* species (KRIMBAS and POWELL 1992), which vary seasonally, with altitude and latitude (DOBZHANSKY 1943; KRIMBAS 1993; KENNINGTON *et al.* 2006). For instance, *Drosophila subobscura*, which harbours a rich polymorphism in paracentric inversions (KRIMBAS and POWELL 1992; KRIMBAS 1993), has been crucial to prove the adaptive role of inversions, as shown by the latitudinal clines on inversion frequencies found in Palearctic and also in colonizing populations (KRIMBAS and LOUKAS 1980; PREVOSTI *et al.* 1988; BALANYÀ *et al.* 2003). Moreover, changes in chromosomal polymorphism matching global warming (SOLÉ *et al.* 2002; BALANYA *et al.* 2004; BALANYA *et al.* 2006) also suggest the adaptive value of inversions in *D. subobscura*. Despite the adaptive role of inversions seems undoubtful, the genetic basis of their adaptation is unknown. Several hypotheses have been suggested to explain the maintenance of inversion polymorphism in populations (HOFFMANN and RIESEBERG 2008). The coadaptation hypothesis (DOBZHANSKY 1950) suggests that natural selection maintains favourable combinations of alleles that interact epistatically within and between arrangements in a certain population. The local adaptation hypothesis explains that inversions will spread in a population at migration-selection balance when they capture at least two advantageous alleles that are individually adapted to local conditions (KIRKPATRICK and BARTON 2006). Thus, epistasis is not required in this later hypothesis.

The low gene flux observed between some arrangements in some *Drosophila* species for genes located inside the inverted region could support both hypotheses (LAAYOUNI *et al.* 2003; SCHAEFFER *et al.* 2003; MUNTÉ *et al.* 2005; HOFFMANN and RIESEBERG 2008). Lower fitness of heterozygous individuals from different populations of *D. pseudoobscura* was detected by Dobzhansky in laboratory experiments and allowed him to propose the coadaptation hypothesis (DOBZHANSKY 1950). However in that same species molecular data rejected the strict coadaptation model since no genetic differentiation was found among populations within inversions (SCHAEFFER *et al.* 2003). In *D. subobscura* high genetic differentiation between European populations was detected when chromosomal arrangements were used as markers, since their frequency widely varies between populations (KRIMBAS 1993). Since chromosomal arrangements are under strong selection, gene flow between populations could be underestimated using them as markers. However low levels of genetic differentiation between European populations were obtained using other markers such as RFLPs (ROZAS *et al.* 1995) or microsatellite *loci* (PASCUAL *et al.* 2001). This discrepancy in population genetics differentiation using different markers could be overcome by analyzing the gene content of the same inversion in different populations.

Our aim in the present research was to analyze genetic differentiation between two populations of *D. subobscura* from the western and eastern Mediterranean: Barcelona (Spain) and Mt. Parnes (Greece) which despite being located at approximately the same latitude, their inversion polymorphism differs substantially (KRIMBAS and LOUKAS 1980; KRIMBAS 1993; ARAÚZ *et al.* 2009). Six nuclear genes have been sequenced in homokaryotypic individuals carrying arrangements O_{ST} and O_{3+4} , since they were the more common arrangements in both populations. The six nuclear genes used have been chosen based on their differential expression between populations maintained at different temperature regimes (LAAYOUNI *et al.* 2007). Thus, these genes could be under selection in relation to chromosomal inversions. We compare the genetic content within and between arrangements for the two populations, and discuss if the Dobzhansky's coadaptation model, which requires genetic differentiation between populations, is observed in *D. subobscura*.

Material and methods

Fly samples, DNA isolation and sequencing

Drosophila subobscura isochromosomal lines were established from two distant European populations (Barcelona, Spain and Mt. Parnes, Greece) following the pattern of genetic crosses described in (MESTRES *et al.* 1998). Chromosomal arrangements for each line were determined by analysis of polytene chromosomes of third instar larvae (MESTRES *et al.* 1998). A total of 25 O₃₊₄ lines (13 from Barcelona and 12 from Mt. Parnes) and 12 O_{ST} lines (5 from Barcelona and 7 from Mt. Parnes) were non-lethal and sequenced in the present study. Furthermore, *D. madeirensis* and *D. pseudoobscura* were used for interspecific analyses. *D. madeirensis* sequences were obtained from a laboratory strain while *D. pseudoobscura* sequences were downloaded from Flybase (<http://flybase.org>).

Six genes located on the O chromosome were selected: *Pif1A* (PFTAIRE-interacting factor 1A), *Abi* (Abelson interacting protein), *Sqd* (Squid), *Yrt* (Yurt), *Atpa* (Na pump α subunit), *Fmr1* (Fragile X mental retardation) since they showed differential expression between laboratory populations maintained at different temperature regimes (LAAYOUNI *et al.* 2007). Although their cytological location had been previously reported, new *in situ* hybridizations have been carried out in the present study, following standard protocols (LAAYOUNI *et al.* 2000). After thoroughly inspection, four out of six genes proved to be not correctly localized previously and the correct locations are given in Table 1. Primers used for amplification and sequencing reaction are listed in Table 1. Some primers were reported in a previous study (LAAYOUNI *et al.* 2007) and others were designed in the present study using Primer Designer v1.01 (1990 Scientific and Educational Software). Single fly genomic DNA was extracted using Genra Puregene Cell Kit (Qiagen).

PCR amplifications were carried out in a 25 μ l total volume with 3.5 μ l buffer 10 \times , 3 μ l dNTPs (1 mM), 0.2 μ l Taq DNA Polymerase (Qiagen), 0.5 μ l forward primer (10 μ M), 0.5 μ l reverse primer (10 μ M), 16.3 μ l water and 1 μ l DNA. Amplification conditions were 4 min at 94°C of initial denaturation, and 35 cycles with 30 sec at 94°C, 30 sec at the required annealing temperature and 3 min at 72°C, and 5 min at 72°C of final extension. PCR amplification products were purified with ExoI-SAP (BioLabs-Promega). Sequencing reactions were obtained with the ABI Prism BigDye Terminators

3.1 cycle sequencing kit (Applied Biosystems) on an ABI 3730 sequencer at the Biotechnology Resource Center of Cornell University. In some cases, due to technical difficulties, the PCR amplification reactions were modified in a 15 μ l total volume with 3 μ l buffer 10x, 3 μ l dNTPs (1 mM), 0.45 μ l DMSO, 0.3 μ l Phire Hot Start DNA Polymerase (Finnzymes), 0.75 μ l forward primer (10 μ M), 0.75 μ l reverse primer (10 μ M), 5.75 μ l water and 1 μ l DNA. Amplification conditions were 30 sec at 98°C of initial denaturation, and 35 cycles with 5 sec at 98°C denaturation, 5 sec at the required annealing temperature and 90 sec at 72°C, and 7 min at 72°C for final extension. PCR amplification products were purified with ExoI-SAP (TebuBio-Promega). Sequencing reactions were obtained with the ABI Prism BigDye Terminators 3.1 cycle sequencing kit (Applied Biosystems) using a 3730 Analyzer (Applied Biosystems), from the Serveis Científic-Tècnics at Universitat de Barcelona. Sequences were assembled with SeqMan II (DNASTAR) and aligned with Clustal W (THOMPSON *et al.* 1994) implemented in BioEdit v7 (HALL 1999).

Data analysis

The six analyzed genes can be divided into two groups according to their cytological location (Figure 1). *Fmr1* gene is located within inversion O₄ in region SI, which includes the chromosome segment covered by the O₃₊₄ arrangement. The other genes (*Pif1A*, *Abi*, *Sqd*, *Yrt* and *Atp α*) are located in region SII (which includes the chromosome segment that can probably freely recombine between arrangements O_{ST} and O₃₊₄). Distance in basepairs of each gene to the nearest inversion breakpoints was estimated assuming that all cytological bands contain the same genetic content and that the length of the O chromosome of *D. subobscura* is equivalent to chromosome 2 of *D. pseudoobscura* (PEGUEROLES *et al.* 2010). The exonic and intronic regions were finally determined by blasting against the genome of *D. pseudoobscura* and the resulting expected proteins were compared to the corresponding *D. pseudoobscura* proteins to ensure the correct assignment of the exons. Exonic and intronic composition of each gene is shown in Figure 1.

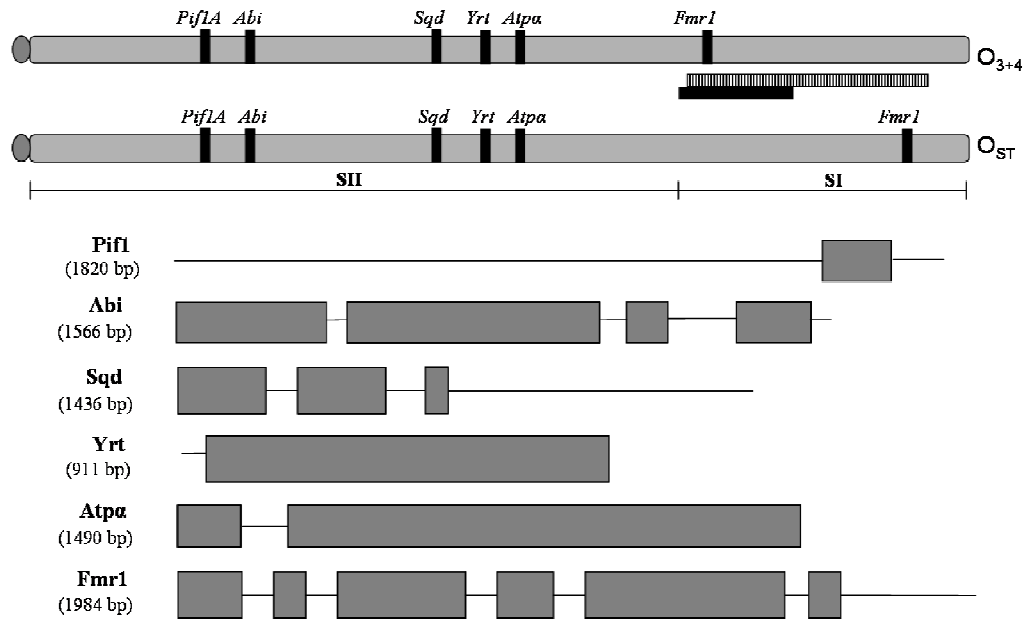


Figure 1: Top: location of the 6 gene regions studied in the O_{3+4} and O_{ST} arrangements. O_3 inversions is represented with a black bar, and O_4 inversion with a striped bar. Bottom: exonic regions are highlighted using grey boxes, and intronic regions are represented by lines. For each gene, the size of the sequenced fragment is given in parentheses.

Table 1: Location of the 6 genes carried out by *in situ* hybridization and primers used in the amplification and sequencing

Gene Symbol ^d	Location	Length (bp)	Primers
<i>PiflA</i> (CG42599)	79D	1820	5'-AATGTATCACAAGGAGAACG-3' ^{a, b, c} 5'-CTCCTGGTAGTACTGCAGAT-3' ^{a, b, c} 5'-GTCCAACACTATGGTTGATGCC-3' ^b 5'-ACGACTTTACAGCTAGCTGG-3' ^b
<i>Abi</i> (CG9749)	81A	1566	5'-CCTTGTCCGCATAGTAGTCA-3' ^{a, b} 5'-GTGAGTAGTAATTAGGTTTCG-3' ^{a, b} 5'-ACTCACGTGGCATACTGCTC-3' ^b
<i>Sqd</i> (CG16901)	85D	1436	5'-AACTAACCTTGTTCCTCTCC-3' ^{a, c} 5'-TTACACACGCTTCGTCAGTT-3' ^{a, b, c} 5'-CATTATAGTAGCCACCAGGA-3' ^b 5'-TCTTGGTGGTTGTTGCTGTG-3' ^b
<i>Yrt</i> (CG9764)	86E	911	5'-CTGGACATCATCGAGAAGGA-3' ^{a, c} 5'-ACATTGGCCAGCTTCACTTG-3' ^{a, c} 5'-AACTGGAGTCTGGCTGGCGT-3' ^b 5'-TACGCCGGATCATCGTTGCT-3' ^b
<i>Atpa</i> (CG5670)	87C	1490	5'-TCATAAGATCTCTCCTGAGG-3' ^{a, b, c} 5'-GCAATATCCTCAACGGTCTC-3' ^{a, b, c} 5'-CATCATTGAGGCGCGCAGCT-3' ^b 5'-CCAGGGCTAGTTCTATCGTA-3' ^b
<i>Fmr1</i> (CG6203)	98A	1984	5'-ACAGCCAAGTCGTTCTACCA-3' ^{a, b, c} 5'-CCATTCACCAGACCTTCCTT-3' ^{a, b, c} 5'-GGCAAGGTAATCGGCAAGAA-3' ^b

^a primers used in the amplifications; ^b primers used for sequencing; ^c primers previously designed (LAAYOUNI *et al.* 2007); ^d Gene symbol of the homologous gene in *D. melanogaster*.

Analyses were carried out separately for each gene and also for the concatenated data set. *Abi* gene was not included in the final concatenate data set due to the insufficient number of individuals. For this gene, sequencing of some individuals failed due to the presence of a gene duplication (data not shown). Only individuals with sequences for the five remaining genes were used in the concatenation. Genes concatenation was carried out using Concatenator v1 software (PINA-MARTINS and PAULO 2008). DnaSP v5 (LIBRADO and ROZAS 2009) was used to estimate the standard parameters for each population and arrangement. We estimated the number of haplotypes (h), number of polymorphic sites (S), the number of singletons, the nucleotide diversity (π) (NEI 1987), nucleotide diversity in synonymous sites and non-coding positions (π_{sil}) (NEI and GOJOBORI 1986) and heterozygosity in silent sites (θ_{sil}) (WATTERSON 1975) were also estimated. Divergence per silent site (K_{sil}) between *D. subobscura* and *D. madeirensis* or *D. pseudoobscura* (NEI and GOJOBORI 1986) was also obtained. Genetic differentiation was assessed using D_{xy} (NEI 1987), F_{ST} (HUDSON *et al.* 1992) and S_{nn} (HUDSON 2000) and estimating its significance with the p-value after 10000 replicates. These analyses were carried out excluding insertion-deletion events (indels). However, due to the huge number of indels observed in the *Pif1A* gene, analyses were also carried out considering indels only with the tetrallelic model implemented in DnaSP v5 (LIBRADO and ROZAS 2009). Tajima's D (TAJIMA 1989) and Fu and Li's D (FU and LI 1993) were used to test for neutrality. Fu and Li's D (FU and LI 1993) was computed both without outgroup and using *D. madeirensis* and *D. pseudoobscura* as outgroups. Population size changes were determined by the statistic R_2 (RAMOS-ONSINS and ROZAS 2002) and its significance was assessed by coalescent simulations using 1000 replicates considering either no recombination, free recombination, or level of recombination (ρ) determined from the sequences directly. The population recombination rate (ρ) was obtained using the composite likelihood method of Hudson (HUDSON 2001) implemented in LDhat v2.1 (<http://www.stats.ox.ac.uk/~mcvean/LDhat>). Gene genealogies were constructed using the Neighbour-Joining method with the maximum composite likelihood model implemented in the Mega v4 program (TAMURA *et al.* 2007). Bootstrap values were determined after 1000 replicates.

Age of inversions was estimated with the *Fmr1* gene, since it is located inside the inversions following two different approaches: using silent nucleotide diversity of the two most divergent individuals and using silent nucleotide diversity of all individuals except recombinants. Number of substitutions per site per year was estimated using the divergence per silent site between *D. subobscura* and *D. pseudoobscura* and assuming that the divergence of the two species occurred 8 My ago (RAMOS-ONSINS *et al.* 1998). Finally, Wilcoxon matched pairs tests were carried out using Statistica v6.

Results

Nucleotide variation

Estimates of genetic variability levels of the six genes for each arrangement (O_{3+4} and O_{ST}) and for each population are presented in Table 2. In all cases, the number of haplotypes detected approached the number of sequences analyzed. Nucleotide diversity (π : $Z = 0.58$, $P = 0.56$) as well as number of polymorphic sites (S : $Z = 0.04$, $P = 0.97$) were similar when comparing both populations. No significant differences when comparing arrangements were detected for π ($Z = 1.33$, $P = 0.18$), although significant differences were obtained for S values ($Z = 2.9$, $P = 0.004$), probably due to the smaller sample size of O_{ST} arrangements. The number of polymorphic sites within each arrangement was similar for all genes, except for *Pif1A*. This gene presented the highest nucleotide diversity and was detected as an outlier in a box-plot representation (data not shown), probably due to its larger proportion of intronic regions. To compensate for this deviation, nucleotide diversity, heterozygosity per site and divergence were analyzed using only synonymous sites and non-coding positions. For each gene, π_{sil} and θ_{sil} values were quite similar, with θ_{sil} values being slightly higher (Table 2). The patterns of gene divergence changed substantially depending on which outgroup was used. For instance, *Abi* was the less divergent gene when using *D. madeirensis*, while *Sqd* was the less divergent gene when using *D. pseudoobscura* as outgroup. The highest π_{sil} and K_{sil} values were both found in the *Yrt* gene, which suggests that this gene might have a higher mutation rate.

Table 2: Nucleotide variation and divergence for each gene and arrangement (O_{3+4} and O_{ST}) in Mt. Parnes (MP) and Barcelona (BC) populations.

	<i>Pif1A</i>		<i>Abi</i>		<i>Sqd</i>		<i>Yrt</i>		<i>Atpa</i>		<i>Fmr1</i>	
	MP	BC	MP	BC	MP	BC	MP	BC	MP	BC	MP	BC
O_{3+4}												
<i>n</i>	12	12	12	7	12	10	12	12	12	12	11	10
<i>h</i>	12	12	12	7	12	10	12	12	12	11	11	10
<i>S</i>	120	114	32	24	25	21	33	35	22	27	38	28
singletons	55	42	20	15	16	14	18	21	9	16	25	18
π	0.021	0.022	0.005	0.006	0.005	0.005	0.010	0.010	0.005	0.005	0.005	0.004
π_{sil}	0.023	0.023	0.012	0.014	0.006	0.006	0.034	0.035	0.015	0.016	0.009	0.008
θ_{sil}	0.026	0.024	0.017	0.015	0.008	0.007	0.042	0.041	0.013	0.019	0.012	0.010
K_{sil}^a	0.034	0.032	0.013	0.013	0.030	0.028	0.078	0.078	0.018	0.018	0.025	0.025
K_{sil}^b	0.191	0.190	0.173	0.173	0.116	0.115	0.327	0.326	0.208	0.209	0.151	0.151
O_{ST}												
<i>n</i>	5	7	2	3	5	6	5	6	5	6	4	6
<i>h</i>	5	7	2	3	5	5	5	6	4	6	4	6
<i>S</i>	82	95	11	16	11	9	23	27	17	12	26	34
singletons	57	49	11	16	8	5	20	22	7	7	16	18
π	0.022	0.023	0.007	0.007	0.004	0.003	0.011	0.011	0.006	0.003	0.007	0.007
π_{sil}	0.023	0.024	0.017	0.017	0.005	0.004	0.037	0.037	0.014	0.011	0.014	0.014
θ_{sil}	0.024	0.025	0.017	0.017	0.005	0.004	0.040	0.043	0.013	0.011	0.014	0.014
K_{sil}^a	0.032	0.038	0.012	0.013	0.031	0.030	0.077	0.078	0.018	0.015	0.020	0.023
K_{sil}^b	0.190	0.194	0.173	0.173	0.116	0.115	0.322	0.326	0.208	0.207	0.149	0.149

n, sample size; *h*, number of haplotypes; *S*, number of polymorphic sites; π , nucleotide diversity including all sites; π_{sil} , nucleotide diversity in synonymous sites and non-coding positions; θ_{sil} , heterozygosity in silent sites; K_{sil}^a , divergence per silent site between *D. subobscura* and *D. madeirensis*; K_{sil}^b , divergence per silent site between *D. subobscura* and *D. pseudoobscura*.

Genetic differentiation between populations and arrangements

When excluding *Fmr1* gene, no fixed mutations differentiating both population were found for none of the two arrangements (O_{3+4} and O_{ST}) and the percentage of shared mutations for the concatenated data set between Mt. Parnes and Barcelona was 44% between O_{3+4} chromosomes and 38% between O_{ST} chromosomes. When comparing populations for the same arrangement, F_{ST} values were small in all cases and negative in 75% of the comparisons (Table 3). Negative F_{ST} values imply that the mean number of differences within population is higher than between populations. Between arrangements, values were small and negative only in 43% of the cases but significant differences were only found for *Fmr1*, which is consistent with its cytological location within inversion O_4 . Moreover, there were significant differences between arrangements for the concatenated data set when considering *Fmr1* but the analysis of the concatenated data set excluding *Fmr1* showed small and non significant values. The highest D_{xy} values were detected for *Pif1A*, (Table 3) which could be explained by its

large portion of intronic content giving higher variability levels. Genetic differentiation with this gene was also estimated using insertions and deletions only, due to their high frequency. It is found that D_{xy} values for indels in *Pif1A* are one order of magnitude higher than in the other genes (estimates obtained excluding gaps). In contrast, F_{ST} estimates from indel events were similar to the other F_{ST} values.

Table 3: Genetic differentiation between O_{3+4} and O_{ST} arrangements from Mt. Parnes (MP) and Barcelona (BC)

		<i>Pif1A</i>	<i>Pif1A</i> ^a	<i>Sqd</i>	<i>Yrt</i>	<i>Atpa</i>	<i>Fmr1</i>	Conc. ^b	Conc. ^c
O_{3+4} MP vs O_{3+4} BC	D_{xy}	0.021	0.049	0.004	0.010	0.005	0.004	0.010	0.009
	F_{st}	-0.017	-0.012	-0.050	-0.008	-0.044	0.028	-0.0214	-0.015
	$P(Snn)$	-	-	-	-	-	-	-	-
O_{ST} MP vs O_{ST} BC	D_{xy}	0.022	0.127	0.003	0.01	0.005	0.007	0.011	0.010
	F_{st}	-0.014	-0.079	-0.048	-0.117	0.079	-0.100	0.038	0.013
	$P(Snn)$	-	-	-	-	-	-	-	-
O_{3+4} MP vs O_{ST} MP	D_{xy}	0.020	0.154	0.004	0.010	0.005	0.011	0.010	0.010
	F_{st}	-0.053	-0.067	-0.031	-0.028	0.035	0.466	-0.054	0.094
	$P(Snn)$	-	-	-	-	-	***	-	**
O_{3+4} BC vs O_{ST} BC	D_{xy}	0.022	0.046	0.004	0.010	0.004	0.011	0.010	0.011
	F_{st}	0.018	0.011*	0.011	-0.070	-0.048	0.497	0.022	0.155
	$P(Snn)$	-	-	-	-	-	***	-	**
O_{3+4} MP vs O_{ST} BC	D_{xy}	0.022	0.139	0.004	0.010	0.004	0.012	0.010	0.011
	F_{st}	0.006	-0.007	-0.021	-0.06	-0.037	0.487	0.021	0.153
	$P(Snn)$	-	-	-	-	-	***	-	*
O_{3+4} BC vs O_{ST} MP	D_{xy}	0.021	0.051	0.004	0.010	0.006	0.011	0.010	0.010
	F_{st}	-0.031	0.017	0.025	-0.047	0.089	0.487	-0.026	0.116
	$P(Snn)$	-	-	-	-	-	**	-	*

Statistical significance was assessed using the p-value of the Snn statistic (-, not significant; *, $0.01 < P < 0.05$; **, $0.001 < P < 0.01$; ***, $P < 0.001$); ^a analysis using InDel events; ^b concatenated data set of genes *Pif1A*, *Sqd*, *Yrt* and *Atpa*; ^c concatenated data set with the addition of *Fmr1* gene

Genealogy reconstruction for *Fmr1* and the concatenated data set were carried out separately using *D. madeirensis* and *D. pseudoobscura* as outgroups (Figure 2). In both phylogenetic trees obtained, sequences from different populations were randomly clustered. However, in the *Fmr1* genealogy, two groups corresponding to the O_{3+4} and the O_{ST} arrangements were observed, despite only sequences with O_{3+4} arrangements were grouped with a high bootstrap value. By visual inspection (Supplementary data) it was detected that four out of ten sequences for the O_{ST} arrangement were recombinants (asterisks show the recombinant individuals in Figure 2A), but none was detected for the O_{3+4} arrangement. The region affected by recombination was approximately 161 bp for MP19 and FBC58, and 10 bp for MP28 and BC55 individuals. When excluding these recombinant individuals, nine fixed differences were obtained between both arrangements combining lines from both populations. Furthermore, after excluding

these recombinants, the phylogenetic reconstruction for the two groups of arrangements showed reciprocal monophyly with more than 89 % bootstrap support (data not shown). However, in the gene genealogy of the concatenated data set, both populations and arrangements were randomly distributed (Figure 2B) and, although some lines with the same arrangements were grouped together bootstrap values were below 30%.

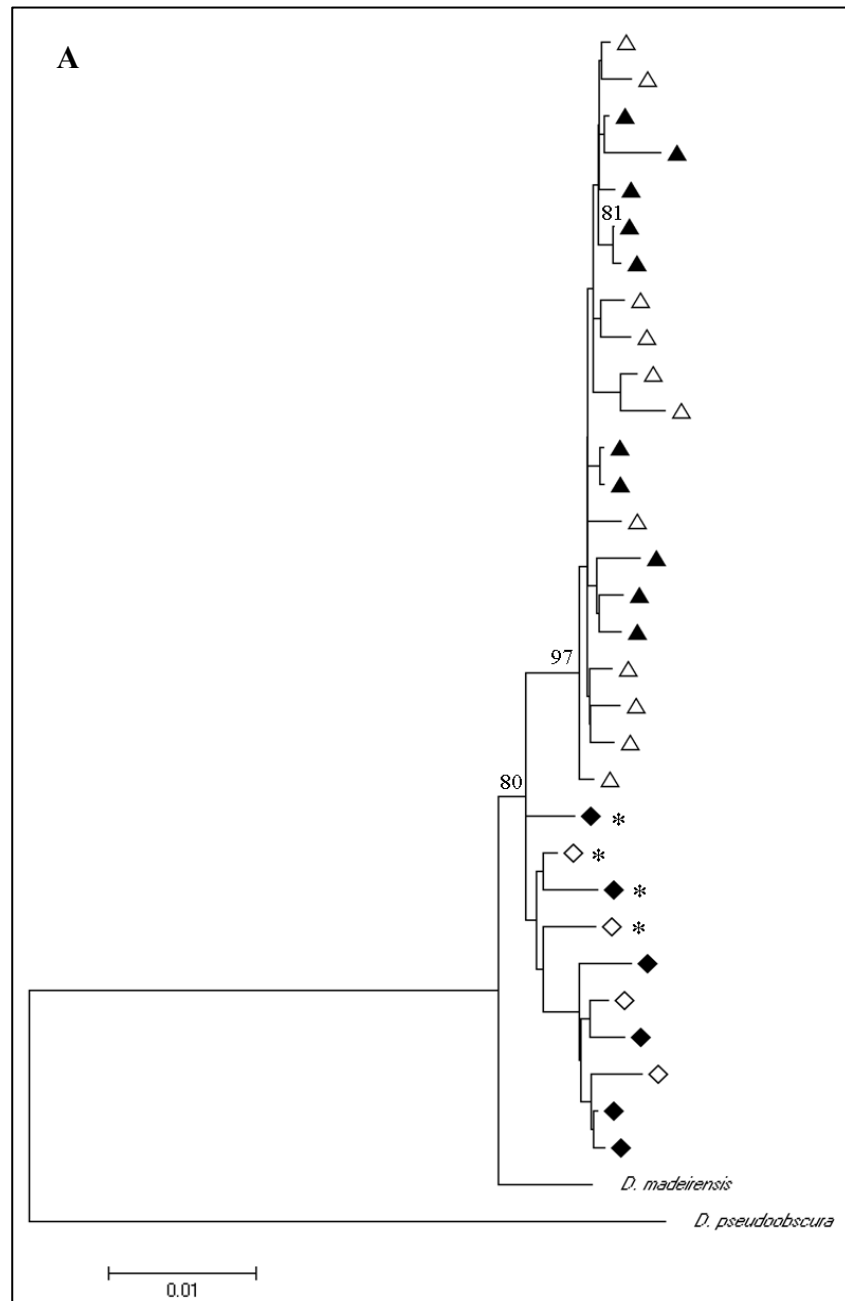


Figure 2A: Neighbor-joining gene genealogy based on total nucleotide variation in the *Fmr1* gene, using *D. madeirensis* and *D. pseudoobscura* as outgroups. Only bootstrap values above 70% are included. (Δ) O_{3+4} from Mt. Parnes, (\blacktriangle) O_{3+4} from Barcelona, (\diamond) O_{ST} from Mt. Parnes and (\blacklozenge) O_{ST} from Barcelona. Asterisks indicate recombinant individuals.

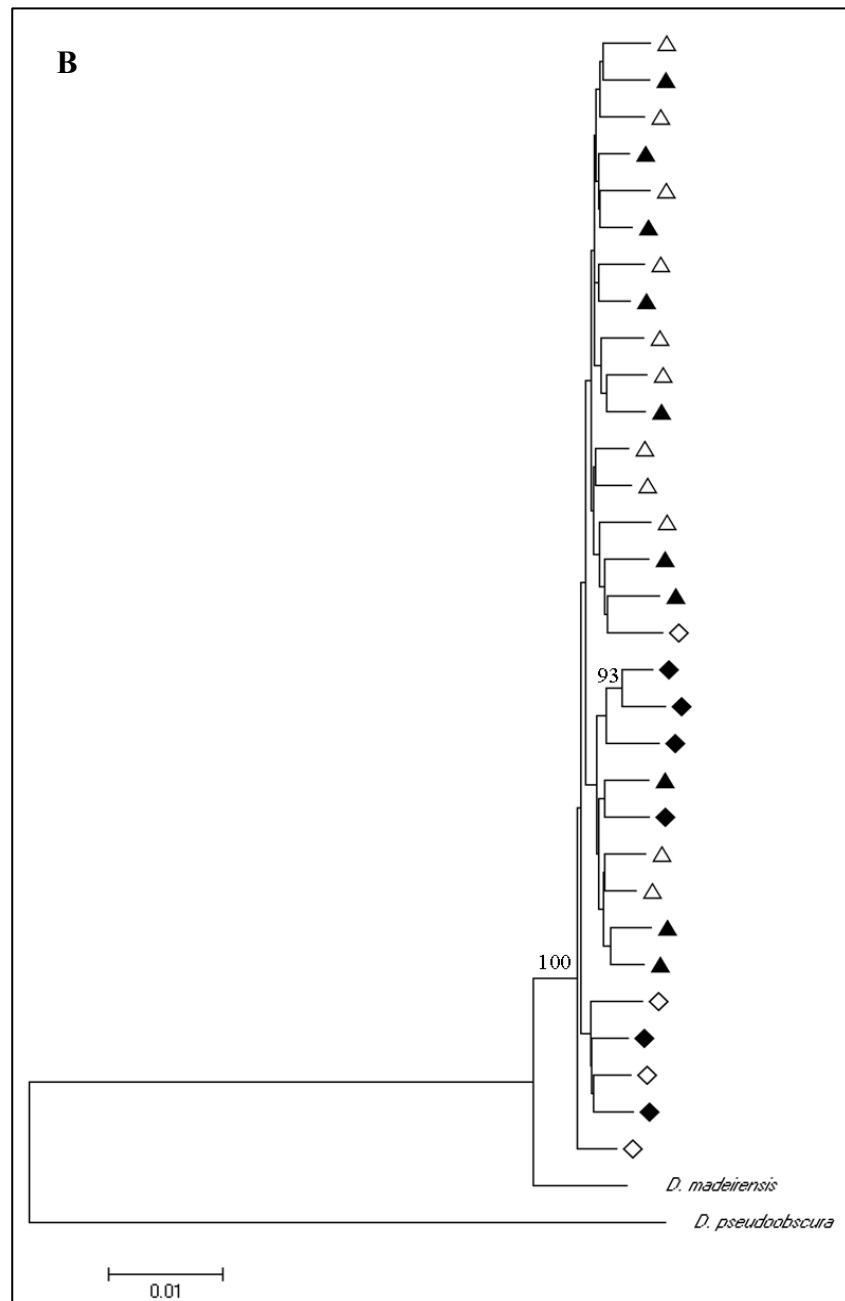


Figure 2B: Neighbor-joining gene genealogy based on total nucleotide variation in the concatenated data set of genes *Pif1A*, *Sqd*, *Yrt*, *Atpa* and *Fmr1*, using *D. madeirensis* and *D. pseudoobscura* as outgroups. Only bootstrap values above 70% are shown. (Δ) O_{3+4} from Mt. Parnes, (\blacktriangle) O_{3+4} from Barcelona, (\diamond) O_{ST} from Mt. Parnes and (\blacklozenge) O_{ST} from Barcelona.

Population demography and selection

Neutrality tests were non significant and showed negative values for the concatenated data set and for each gene separately, except for the *Atpa* gene in Mt. Parnes (Table 4). Negative D values correspond to an excess of low frequency polymorphism, which could be due to either demography or selection. To further elucidate this issue, the R_2

statistic (RAMOS-ONSINS and ROZAS 2002) was estimated, since it is more suitable for small sample sizes presenting recombination. R_2 estimates ranged between 0.09 and 0.18 and significance was assayed considering three alternative scenarios, since rejection of the null hypothesis of constant population size depends on the level of recombination implemented (Table 4). When no recombination was included in the coalescent simulations, most of the comparisons were non significant. However, when recombination (ρ) was taken into account in the coalescent simulations, most of the comparisons were significant. Finally, the same set of genes presented significant p-values both using ρ or free recombination in the simulations.

Table 4: Neutrality tests for each gene and population, Rho estimated using the LDhat program and test of population expansion using Ramos-Onsins and Rozas' R_2 statistic.

Mt. Parnes	Tajima's D	Fu and Li's D ^a	R_2	ρ	p-value ^b	p-value ^c	p-value ^d
<i>Pif1A</i>	-0.723	-1.142	0.101	100	0.129	0.000	0.000
<i>Abi</i>	-1.099	-0.603	0.091	69	0.034	0.010	0.008
<i>Sqd</i>	-1.017	-1.299	0.092	81	0.062	0.035	0.028
<i>Yrt</i>	-0.871	-0.109	0.096	100	0.079	0.019	0.019
<i>Atpa</i>	0.942	1.104	0.176	13	0.881	0.945	0.990
<i>Fmr1</i>	-1.100	-0.753	0.088	27	0.033	0.008	0.000
Concatenate ^e	-0.799	-0.670	0.100	100	0.070	0.000	0.000
Barcelona							
<i>Pif1A</i>	-0.326	-0.475	0.116	100	0.340	0.101	0.101
<i>Abi</i>	-0.908	-1.569	0.087	68	0.003	0.003	0.002
<i>Sqd</i>	-0.776	-1.335	0.104	62	0.113	0.093	0.079
<i>Yrt</i>	-0.833	-0.418	0.099	100	0.137	0.033	0.033
<i>Atpa</i>	-0.648	-1.033	0.102	79	0.151	0.091	0.069
<i>Fmr1</i>	-0.619	-0.896	0.109	17	0.179	0.133	0.051
Concatenate ^e	-0.518	-0.810	0.110	100	0.197	0.000	0.000

^a using *D. pseudoobscura* as an outgroup; ^b R_2 p-value estimated by coalescent simulations without recombination; ^c R_2 p-value estimated by coalescent simulations with estimated levels of recombination; ^d R_2 p-value estimated by coalescent simulations with free recombination; ^e genes *Pif1A*, *Sqd*, *Yrt*, *Atpa* and *Fmr1* included in the concatenate data set. $P < 0.05$ are in bold.

Age of inversions

Sequences from *Fmr1* were used for estimating the age of the inversions, since it is the only gene analysed which is located in section SI, in which both arrangements differ. Recombinant individuals were excluded because only variation originated by mutation is useful for this analysis (Figure 2). The ages of inversions were estimated for each population separately and combining them since their origin is unique (Table 5). Older estimates were obtained using the information from the two more distant individuals, in relation to the values obtained using the mean silent nucleotide diversity. Nevertheless,

the age estimates for O_{3+4} and O_{ST} were quite similar with both approaches, although older coalescent time was detected for O_{3+4} when using the most divergent individuals, while older coalescent time was detected for O_{ST} with the mean nucleotide diversity.

Table 5: silent nucleotide diversity and age of the different gene arrangements using *Fmr1* gene

	π_{sil}^a	Time	π_{sil}^b	Time
O_{3+4} MP	0.014	0.77	0.009	0.45
O_{3+4} BC	0.016	0.87	0.008	0.41
O_{ST} BC	0.014	0.72	0.009	0.48
O_{3+4}^c	0.019	1.01	0.008	0.44
O_{ST}^c	0.015	0.77	0.009	0.49

^a π_{sil} between the most divergent individuals; ^b mean π_{sil} after excluding recombinant individuals; ^c estimations obtained combining individuals from both populations

Discussion

Nucleotide variation

The present study focuses on the analysis of six genes in the context of two different arrangements: the O_{ST} and the O_{3+4} . Five of these genes are located outside the inverted region and only one (*Fmr1*) was located inside the O_4 inversion and approximately 1.3 Mb from the nearest breakpoint. In terms of nucleotide variability, lower nucleotide diversity values are expected in markers located approximately 1Mb from the inversion breakpoint (ANDOLFATTO *et al.* 2001). However, despite *Fmr1* gene could be included in this category, it does not present the lowest levels of π_{sil} . In a previous study, no reduction of the variation was observed in genes located close to the inversion breakpoints when compared with genes inside the inverted region (MUNTÉ *et al.* 2005). However, in other *Drosophila* species a correlation between nucleotide diversity and distance was observed, with lower levels of nucleotide diversity in genes located near inversion breakpoints (LAAYOUNI *et al.* 2003).

In the present study, *Yrt* showed both the highest π_{sil} and K_{sil} , suggesting that this gene could present a high neutral mutation rate, as previously observed in the *Acph1* gene (MUNTÉ *et al.* 2005). Excluding *Yrt*, since it could act as an outlier for the nucleotide variability, estimates of π_{sil} ranged from 0.004 to 0.024 for the O_{ST} arrangement and from 0.006 to 0.023 for O_{3+4} in the region where genes can freely recombine between both arrangements. For genes located within the inversion (MUNTÉ *et al.* 2005), π_{sil} ranged from 0.004 to 0.013 for the O_{ST} arrangement, and from 0.008 to

0.013 for the O_{3+4} arrangement after excluding the *Acph1* gene. Moreover, the *Fmr1* gene estimated π_{sil} was 0.014 for the O_{ST} and 0.009 for O_{3+4} . Consequently, for genes located inside and outside the inverted regions, the ranges of nucleotide variability mostly overlap although the highest value is found outside the inversions. Thus, mutation and recombination could both have contributed to recover variability levels within the inverted regions. In fact, the variability for *Fmr1* gene in O_{ST} was partly due to the presence of recombinants, since the values for this arrangement were much reduced when they were excluded (Table 2 and Table 5). Most probably these recombinant individuals were generated by gene conversion, since estimated recombination tracts affected 161 bp and 10 bp. It has been previously suggested that recombination should be higher in central parts of the inversion because genes located in this region could be influenced by double crossovers and gene conversion, while genes close to inversions breakpoints should be mainly influenced by gene conversion (NAVARRO *et al.* 1997). Thus, recombination reduction mediated by inversions could play an important role in the maintenance of genetic differentiation between O_{ST} and O_{3+4} arrangements.

Genetic differentiation between populations

Genetic differentiation was assessed using two independent statistics, D_{xy} (NEI 1987) and F_{ST} (HUDSON *et al.* 1992). D_{xy} is an estimate of the average number of nucleotide substitutions per site between arrangements. However, its value is influenced by levels of nucleotide diversity, and they can be affected by factors such as neutral mutation rate and recombination. Interestingly, the gene presenting the highest D_{xy} values was *Pif1A*, which at the same time has the highest π value due to the high proportion of intronic regions that were amplified. On the other hand, F_{ST} estimates do not seem to be influenced by nucleotide diversity levels, being more suitable to compare levels of differentiation between different genes.

One of the aims of the present work was to study the genetic differentiation between two distant populations at the molecular level. Differences in chromosomal inversions between these two populations had been previously observed (ARAÚZ *et al.* 2009). For instance, the O_{3+4} arrangement represents the 28% of the Barcelona population, but its frequency increases to 52% in Mt. Parnes. Furthermore, inversions that did not appear or had a very low frequency in one population, showed a significant

frequency in the other (15% O_{3+4+7} , 24% O_{3+4+8} and 4% O_{3+4+1} in Barcelona vs 0% O_{3+4+7} , 3% O_{3+4+8} and 14% O_{3+4+1} in Mt. Parnes). These discrepancies could be generated by these two regions being considered pleistocenic refugia (TABERLET *et al.* 1998; HEWITT 1999) and maintained by low gene flow or selection to local environments (KOVACEVIC and SCHAEFFER 2000). However, no significant genetic differentiation was detected between both populations within arrangements, suggesting no differences neither due to their population of origin nor to gene location. In other *Drosophila* species, such as *D. pseudoobscura* (SCHAEFFER *et al.* 2003), no genetic differentiation was found between populations within arrangements using genes located in the inverted region. No significant genetic differentiation was found among three European populations of *D. subobscura* (one from Holland and two from Spain) when analysing restriction length polymorphism and comparing O_{ST} and O_{3+4} arrangements (ROZAS *et al.* 1995). These results suggest that gene flow between populations is high, and agree with several studies detecting non-negligible gene flow between natural European populations of *D. subobscura* (LATORRE *et al.* 1992; PASCUAL *et al.* 2001; ZIVANOVIC *et al.* 2007). Gene flow between distant populations could be attributed to migration both by passive transportation associated to human activities (PASCUAL *et al.* 2007) and to the active dispersal capabilities of *D. subobscura* (SERRA *et al.* 1987).

Genetic differentiation between arrangements

No significant genetic differentiation between chromosomal arrangements was observed in the present study for genes located outside the inversions. However, the gene placed inside O_{ST} and O_{3+4} (*Fmr1*) showed significant genetic differentiation. For this gene, F_{ST} values were similar to those obtained in previous studies using other genes located within the same inverted region (ROZAS *et al.* 1999; MUNTÉ *et al.* 2005). High gene flux between arrangements was also found in *D. pseudoobscura* (SCHAEFFER and ANDERSON 2005) and *D. buzzatti* (LAAYOUNI *et al.* 2007). The differentiation between arrangements could be due to the recent origin of an inversion, to recombination reduction mediated by inversions, and to selection binding together adaptive genes captured within inversions. Many observations support the adaptive value of inversions in *D. subobscura* such as the latitudinal clines for some chromosomal arrangements (PREVOSTI *et al.* 1988). In particular, O_{ST} and O_{3+4} present opposite latitudinal clines, with O_{ST} being more frequent in northern Europe and O_{3+4} being more frequent in the south (SOLÉ *et al.* 2002; BALANYÀ *et al.* 2004). However,

none of the nine different genes located across the O_{3+4} arrangement that have been studied so far seem to be directly under selection (MUNTÉ *et al.* 2005) and present study). Nevertheless, as mentioned above genetic differentiation could be also maintained by lower rate of exchange inside the inverted regions. This effect could be reinforced by the presence of overlapped inversions since the O_{3+4} arrangement which is formed by two overlapped inversions reduces in a greater extent genetic exchange in comparison to single inversions (NÓBREGA *et al.* 2008).

Population demography and selection

The hypothesis that the excess of rare alleles detected by the Tajima's D test may be the result of a recent population expansion was further confirmed by the R_2 statistic. Thus, the observed pattern of nucleotide variability could reflect an expansion of Barcelona and Mt. Parnes populations. Negative Tajima's D values obtained for genes located inside the inverted regions in previous studies (ROZAS *et al.* 1999; MUNTÉ *et al.* 2005), could also be due to an expansion of the species or hitchhiking related to the inversion. However, the duplicated *Acp70A* gene, which is located outside inverted regions, showed positive but non significant Tajima's D values (CIRERA and AGUADÉ 1998), led to discard the hypothesis of an expansion of the whole species (ROZAS *et al.* 1999). Our analyses of the five genes not associated with inversions suggest that *D. subobscura* has passed through a population expansion, although we also detected a positive D value for one gene in Mt. Parnes. More R_2 values are significant in Mt. Parnes than in Barcelona, suggesting that Mt. Parnes had a higher expansion, in accordance with its larger effective population size (ARAÚZ *et al.* 2009).

Age of inversions

The age of inversions can be estimated considering their unique origin (POWELL 1997). These estimates can be inferred from the time of coalescence of the sequences, taking into account that variability accumulated in the sequences is proportional to its origin. As previously mentioned, divergence was initially estimated using two species of the *obscura* group as outgroups (*D. madeirensis* and *D. pseudoobscura*). On one hand, *D. madeirensis* is closely related to *D. subobscura*, since their divergence time was estimated at about 0.63 Mya by RAMOS-ONSINS *et al.* (1998). This species is endemic of Madeira Island and has a small effective population size. However, its nucleotide diversity is similar to *D. subobscura*, and therefore not showing the lower levels of

variation typically found in an island (RAMOS-ONSINS *et al.* 1998; KHADEM *et al.* 2001) also noted the presence of shared polymorphisms between these two species due to their common ancestry. On the other hand, *D. pseudoobscura* divergence time with respect to *D. subobscura* was estimated to be about 8 Mya (RAMOS-ONSINS *et al.* 1998). *D. pseudoobscura* is a native species from North America and its effective size could be quite similar to *D. subobscura* (PASCUAL *et al.* 2000). For *Abi* and *Atpa* genes, values of silent nucleotide diversity were quite similar to the divergence between *D. subobscura* and *D. madeirensis*. Due to this fact, it was compulsory to use a more divergent outgroup such as *D. pseudoobscura* to estimate the age of the inversions. Furthermore, individuals detected as recombinants (whenever due to crossing over or to gene conversion) were excluded from the analysis since variability originated by recombination causes an overestimation of the ages (ROZAS *et al.* 1999). In the present study, recombinants were observed only among the O_{ST} lines. This bias could arise from the frequencies of the different arrangements in these two populations. All arrangements O_{3+4+X}, where X includes non overlapped inversions with 3+4, can freely recombine in SI. In Barcelona and Mt. Parnes, as shown by ARAÚZ *et al.* (2009), it is more likely to find O_{3+4+X} homozygous (35% and 74% respectively) than O_{ST}/O_{3+4+X} heterozygous (13% and 18% respectively) or O_{ST} homozygous (1.3% and 1.1% respectively). Therefore, the probability of sampling a recombining O_{ST} chromosome is higher than to find a recombinant O_{3+4+X}.

The age of the inversion can be estimated using the silent nucleotide diversity per site between the two most divergent individuals (ROZAS and AGUADÉ 1994) or the mean silent nucleotide diversity per site of all individuals except recombinants (NÓBREGA *et al.* 2008; ROZAS *et al.* 1999). As expected, when using the most divergent individuals our estimations of the inversion age were older than when using the average. The estimated age of O₃₊₄ arrangement when mixing populations and using the deepest π_{sil} values (1.01 Myr) was older than the estimates obtained separately for Barcelona and Mt. Parnes populations (0.77 and 0.87 Myr). However, when using the mean π_{sil} values of all individuals except recombinants, the coalescent time (0.44 Myr) was approximately the same than for Barcelona and Mt. Parnes populations (0.41 Myr and 0.45 Myr respectively). Age estimates of O_{ST} and O₃₊₄ arrangements were quite similar in previous studies, differing in 0.15 Myr in (ROZAS and AGUADÉ 1994), 0.09 Myr in (ROZAS *et al.* 1999) and 0.05 Myr in (NAVARRO-SABATÉ *et al.* 1999). In all these

studies O_{ST} was slightly younger than O_{3+4} . In the present study, we also detected similar age estimates for the two arrangements, differing in 0.07 and 0.15 Myr, with the slightly older arrangement changing upon the methodology employed. These discrepancies could be due to differences in sample size, undetected recombination events which would affect differentially both methods and genes localization relative to the inversion breakpoints.

Acknowledgements

We thank Ferran Palero for his valuable comments on the manuscript. We also thank Joan Balanyà for the analyses of the polytene chromosomes, Pedro A. Araúz for the collection of the flies and the collaboration with the crosses to obtain the homokaryotypic lines, Gemma Calabria and Olga Dolgova for their collaboration in the *in situ* hybridizations. Furthermore, we want to thank Vanessa Bauer-DuMont for her technical support. This work was supported by a pre-doctoral fellowship to CP (2009FIC-00096) from the Generalitat de Catalunya (Spain). Research was founded by projects CGL2006-13423-C02-02 from the Ministerio de Ciencia y Tecnología (MCYT, Spain) and 2009 SGR 636 from the Generalitat de Catalunya (Spain).

Literature cited

- ANDOLFATTO, P., F. DEPAULIS and A. NAVARRO, 2001 Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetics Research* **77**: 1-8.
- ARAÚZ, P. A., F. MESTRES, C. PEGUEROLES, C. ARENAS, G. TZANNIDAKIS *et al.*, 2009 Tracking the origin of the American colonization by *Drosophila subobscura*: genetic comparison between Eastern and Western Mediterranean populations. *Journal of Zoological Systematics and Evolutionary Research* **47**: 25-34.
- AYALA, F. J., and M. COLUZZI, 2005 Chromosome speciation: Humans, *Drosophila*, and mosquitoes. *Proceedings of the National Academy of Sciences of the USA* **102**: 6535-6542.
- BALANYÀ, J., J. M. OLLER, R. B. HUEY, G. W. GILCHRIST and L. SERRA, 2006 Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* **313**: 1773-1775.
- BALANYÀ, J., L. SERRA, G. W. GILCHRIST, R. B. HUEY, M. PASCUAL *et al.*, 2003 Evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. *Evolution* **57**: 1837-1845.

- BALANYÀ, J., E. SOLÉ, J. OLLER, D. SPERLICH and L. SERRA, 2004 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. II. European populations. *Journal of Zoological Systematics and Evolutionary Research* **42**: 191-201.
- CIRERA, S., and M. AGUADÉ, 1998 Molecular evolution of a duplication: The sex-peptide (Acp70A) gene region of *Drosophila subobscura* and *Drosophila madeirensis* *Molecular Biology and Evolution* **15**: 988-996.
- DOBZHANSKY, T., 1943 Temporal changes in the composition of populations of *D. pseudoobscura*. *Genetics* **28**: 162-186.
- DOBZHANSKY, T., 1950 Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35**: 288-302.
- FU, Y. X., and W. H. LI, 1993 Statistical tests of neutrality of mutations. *Genetics* **133**: 693-709.
- HALL, T. A., 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98.
- HEWITT, G. M., 1999 Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**: 87-112.
- HOFFMANN, A. A., and L. H. RIESEBERG, 2008 Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics* **39**: 21-42.
- HUDSON, R. R., 2000 A new statistic for detecting genetic differentiation. *Genetics* **155**: 2011-2014.
- HUDSON, R. R., 2001 Two-locus sampling distributions and their application. *Genetics* **159**: 1805-1817.
- HUDSON, R. R., M. SLATKIN and W. P. MADDISON, 1992 Estimation of levels of gene flow from DNA-sequence data. *Genetics* **132**: 583-589.
- KENNINGTON, W. J., L. PARTRIDGE and A. A. HOFFMANN, 2006 Patterns of diversity and linkage disequilibrium within the cosmopolitan inversion *In(3R)Payne* in *Drosophila melanogaster* are indicative of coadaptation. *Genetics* **172**: 1655-1663.
- KHADEM, M., J. ROZAS, C. SEGARRA and M. AGUADÉ, 2001 DNA variation at the *rp49* gene region in *Drosophila madeirensis* and *Drosophila subobscura* from

- Madeira: inferences about the origin of an insular endemic species. *Journal of Evolutionary Biology* **14**: 379-387.
- KIRKPATRICK, M., and N. BARTON, 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419-434.
- KOVACEVIC, M., and S. W. SCHAEFFER, 2000 Molecular population genetics of X-linked genes in *Drosophila pseudoobscura*. *Genetics* **156**: 155-172.
- KRIMBAS, C. B., 1993 *Drosophila subobscura: Biology, Genetics and Inversion polymorphism*, Verlag Dr. Kovac, Hamburg.
- KRIMBAS, C. B., and M. LOUKAS, 1980 The inversion polymorphism of *Drosophila subobscura*. *Evolutionary Biology* **12**: 163-234.
- KRIMBAS, C. B., and J. F. POWELL, 1992 *Drosophila Inversion Polymorphism*. CRC Press, Boca Raton, Florida.
- LAAYOUNI, H., F. GARCIA-FRANCO, B. E. CHAVEZ-SANDOVAL, V. TROTTA, S. BELTRAN *et al.*, 2007 Thermal evolution of gene expression profiles in *Drosophila subobscura*. *BMC Evolutionary Biology* **7**: 42.
- LAAYOUNI, H., E. HASSON, M. SANTOS and A. FONTDEVILA, 2003 The evolutionary history of *Drosophila buzzatii*. XXXV. Inversion polymorphism and nucleotide variability in different regions of the second chromosome. *Molecular Biology and Evolution* **20**: 931-944.
- LAAYOUNI, H., M. SANTOS and A. FONTDEVILA, 2000 Toward a physical map of *Drosophila buzzatii*. Use of randomly amplified polymorphic dna polymorphisms and sequence-tagged site landmarks. *Genetics* **156**: 1797-1816.
- LATORRE, A., C. HERNANDEZ, D. MARTINEZ, J. A. CASTRO, M. RAMON *et al.*, 1992 Population structure and mitochondrial DNA gene flow in Old World populations of *Drosophila subobscura*. *Heredity* **68**: 15-24.
- LIBRADO, P., and J. ROZAS, 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.
- MESTRES, F., J. SANZ and L. SERRA, 1998 Chromosomal structure and recombination between inversions in *Drosophila subobscura*. *Hereditas* **128**: 105-113.
- MUNTÉ, A., J. ROZAS, M. AGUADÉ and C. SEGARRA, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* **169**: 1573-1581.

- NAVARRO-SABATÉ, A., M. AGUADÉ and C. SEGARRA, 1999 The relationship between allozyme and chromosomal polymorphism inferred from nucleotide variation at the *AcpH-1* gene region of *Drosophila subobscura*. *Genetics* **153**: 871-889.
- NAVARRO, A., E. BETRAN, A. BARBADILLA and A. RUIZ, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695-709.
- NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- NEI, M., and T. GOJOBORI, 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**: 418-426.
- NÓBREGA, C., M. KHADEM, M. AGUADÉ and C. SEGARRA, 2008 Genetic exchange versus genetic differentiation in a medium-sized inversion of *Drosophila*: the A_2/A_{ST} arrangements of *Drosophila subobscura*. *Molecular Biology and Evolution* **25**: 1534-1543.
- PASCUAL, M., C. F. AQUADRO, V. SOTO and L. SERRA, 2001 Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. *Molecular Biology and Evolution* **18**: 731-740.
- PASCUAL, M., M. P. CHAPUIS, F. MESTRES, J. BALANYÀ, R. B. HUEY *et al.*, 2007 Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**: 3069-3083.
- PASCUAL, M., M. D. SCHUG and C. F. AQUADRO, 2000 High density of long dinucleotide microsatellites in *Drosophila subobscura*. *Molecular Biology and Evolution* **17**: 1259-1267.
- PEGUEROLES, C., P. A. ARAÚZ, M. PASCUAL and F. MESTRES, 2010 A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*. *Genetica*, in press.
- PINA-MARTINS, F., and O. S. PAULO, 2008 CONCATENATOR: Sequence data matrices handling made easy. *Molecular Ecology Resources* **8**: 1254-1255.
- POWELL, J. R., 1997 Progress and prospects in evolutionary biology: The *Drosophila* model. Oxford University Press, New York.
- PREVOSTI, A., G. RIBÓ, L. SERRA, M. AGUADÉ, J. BALANYÀ *et al.*, 1988 Colonization of America by *Drosophila subobscura*: Experiment in natural populations that

- supports the adaptive role of chromosomal inversion polymorphism. *Proceedings of the National Academy of Sciences USA* **85**: 5597-5600.
- RAMOS-ONSINS, S. E., and J. ROZAS, 2002 Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**: 2092-2100.
- RAMOS-ONSINS, S. E., C. SEGARRA, J. ROZAS and M. AGUADE, 1998 Molecular and chromosomal phylogeny in the *obscura* group of *Drosophila* inferred from sequences of the *rp49* gene region. *Molecular Phylogenetics and Evolution* **9**: 33-41.
- ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. *Proceedings of the National Academy of Sciences USA* **91**: 11517-11521.
- ROZAS, J., C. SEGARRA, G. RIBÓ and M. AGUADÉ, 1999 Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *Drosophila subobscura*. *Genetics* **151**: 189-202.
- ROZAS, J., C. SEGARRA, C. ZAPATA, G. ALVAREZ and M. AGUADE, 1995 Nucleotide polymorphism at the *rp49* region of *Drosophila subobscura*: Lack of geographic subdivision within chromosomal arrangements in Europe. *Journal of Evolutionary Biology* **8**: 355-367.
- SANTOS, J., L. SERRA, E. SOLÉ and M. PASCUAL, 2010 FISH mapping of microsatellite loci from *Drosophila subobscura* and its comparison to related species *Chrosomal Research*, in press.
- SCHAEFFER, S. W., and W. W. ANDERSON, 2005 Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura*. *Genetics* **171**: 1729-1739.
- SCHAEFFER, S. W., A. BHUTKAR, B. F. MCALLISTER, M. MATSUDA, L. M. MATZKIN *et al.*, 2008 Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* **179**: 1601-1655.
- SCHAEFFER, S. W., M. P. GOETTING-MINESKY, M. KOVACEVIC, J. R. PEOPLES, J. L. GRAYBILL *et al.*, 2003 Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. *Proceedings of the National Academy of Sciences USA* **100**: 8319-8324.
- SERRA, L., G. PEGUEROLES and F. MESTRES, 1987 Capacity of dispersal of a colonizing species: *Drosophila subobscura*. *Genetica* **73**: 223-235.

- SOLÉ, E., J. BALANYÀ, D. SPERLICH and L. SERRA, 2002 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. I. Mediterranean populations from southwestern Europe. *Evolution* **56**: 830-835.
- STUMP, A. D., M. POMBI, L. GOEDEL, J. M. C. RIBEIRO, J. A. WILDER *et al.*, 2007 Genetic exchange in 2La inversion heterokaryotypes of *Anopheles gambiae*. *Insect Molecular Biology* **16**: 703-709.
- TABERLET, P., L. FUMAGALLI, A. G. WUST-SAUCY and J. F. COSSON, 1998 Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**: 453-464.
- TAJIMA, F., 1989 Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.
- TAMURA, K., J. DUDLEY, M. NEI and S. KUMAR, 2007 MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.
- THOMPSON, J. D., D. G. HIGGINS and T. J. GIBSON, 1994 CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.
- WATTERSON, G. A., 1975 Number of segregating sites in genetic models without recombination. *Theoretical Population Biology* **7**: 256-276.
- ZIVANOVIC, G., C. ARENAS and F. MESTRES, 2007 The genetic structure of Balkan populations of *Drosophila subobscura*. *Hereditas* **144**: 120-128.

3. 4. Flux gènic entre ordenaments de *D. subobscura*.

Publication 4: Recent origin of a chromosomal inversion revealed by its association with nucleotide haplotypes

Publicació 4: Deteminació de l'origen recent d'una inversió cromosòmica gràcies a la seva associació amb haplotips nucleotídics

RESUM

En aquest estudi s'ha obtingut la seqüència nucleotídica d'un fragment de 800 parells de bases del gen *Odh* (Octanol deshidrogenasa) localitzat en el cromosoma O de *Drosophila subobscura* en línies cromosòmiques O₅ originaries de Mar del Plata (Argentina), Sunne (Suècia), Zanjic (Montenegro) i Drobak (Noruega). Algunes d'aquestes línies eren letals en combinació homocarítipica (Mar del Plata, Sunne i Zanjic), mentre que la procedent de Drobak presentava viabilitat normal. Només es van obtenir dos haplotips diferents: un, detectat a les línies O₅ de Mar del Plata, Sunne i Zanjic, és el que ja s'havia observat completament associat a la inversió O₅ en les poblacions colonitzadores d'Amèrica i també en altres línies europees amb aquesta inversió (Taulé, França i Lilla Edet, Suècia). L'altre, obtingut a la línia procedent de Drobak, ja havia estat detectat en associació amb una línia O₅ originaria de Lovain-la-Neuve (Bèlgica). El fet de trobar un mateix haplotip en localitats geogràficament tant allunyades (Escandinàvia i els Balcans) dóna idea del gran flux gènic entre les poblacions de *D. subobscura*. Finalment, la explicació més probable de l'obtenció de només dos haplotips associats a la inversió O₅ és el seu origen relativament recent respecte a la escala del temps evolutiu.

Recent origin of a chromosomal inversion revealed by its association with nucleotide haplotypes

PEDRO A. ARAÚZ¹, CINTA PEGUEROLES¹, GORAN ZIVANOVIC², PEDRO J. FERNÁNDEZ-IRIARTE³, HÉCTOR RUIZ-MARTÍN¹, JOAN BALANYÀ¹, LLUÍS SERRA¹ AND FRANCESC MESTRES¹

¹Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avd. Diagonal 645, 08028, Barcelona, Spain.

²Department of Genetics, Institute for Biological Research ‘Sinisa Stankovic’, University of Belgrade, Serbia.

³Departamento de Biología, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, 7600 Mar del Plata, Argentina.

*Corresponding author: fmestres@ub.edu

Published in 2009 in Hereditas 146(5): 257-259.

The colonization of America by *Drosophila subobscura* can be considered a natural experiment with two replicates (North and South America) and offers a great opportunity to study several evolutionary phenomena. This process has been analyzed from its early stages (BRNCIC et al. 1981) and some valuable information has been obtained. For instance, the magnitude of the founding event has been estimated using different genetic markers, the non-independence of both colonizations (North and South America) has been established and subsequent evolution of American colonizing populations has been examined (for a revision see BALANYÀ et al. 2003, 2006; MESTRES et al. 2005).

However, one point still remains obscure: the Palearctic population from which the colonization started. Although some evidence indicates a possible origin from the Mediterranean region (PASCUAL et al. 2001, 2007; GILCHRIST et al. 2001), it is not clear from which population the colonization started. A key element in this puzzle is the O₅ inversion. Although it has a rather erratic distribution in the Palearctic region and always is detected in moderate (Scandinavia) or low (other regions) frequencies (for a revision see KRIMBAS 1992, 1993; ZIVANOVIC and MESTRES 2000), it is frequent in American populations (AYALA et al. 1989; PREVOSTI et al. 1989; MESTRES et al. 2005). Not only can this inversion be detected in most American populations, but it also presents a clinal distribution with latitude in both colonized areas (PREVOSTI et al. 1988; BALANYÀ et al. 2003). Another particular feature of this inversion is its complete association with a lethal gene in America. All American O₅ inversions carry the same lethal gene, and this gene is always detected in association with the O₅ inversion (for a revision see MESTRES et al. 2005, 2009). Most interesting is the proven heterotic effect of this inversion in American populations, in spite of carrying a lethal gene (MESTRES et al. 2001). To characterize the O₅ inversion at the molecular level, the *Odh* gene was selected because it is located inside the inversion close to its proximal breakpoint (MESTRES et al. 2004). In American populations, the O₅ inversions are always associated to a particular haplotype of the *Odh* gene (Haplotype 1, GeneBank/EMBL accession number AJ496667). This inversion is difficult to obtain in the Old World and until now only three European O₅ inversions could be sequenced for their *Odh* gene. Two of them carried Haplotype 1, whereas the other carried a different one (Haplotype 2, GeneBank/EMBL accession number AJ496717), which differed in only four positions (MESTRES et al. 2004).

The aim of the present research has been to obtain additional O₅ inversions from the Palearctic region to gather more information on the origin of the American colonization by checking their viability and sequencing their *Odh* gene. We have analyzed three O₅ chromosomal lines from Zanjic (Montenegro), Sunne (Sweden) and Drobak (Norway), respectively. Furthermore, by sequencing one O₅ lethal chromosomal line obtained from the colonizing population of Mar del Plata (Argentina) we wanted to confirm that the *D. subobscura* populations of the Atlantic coast of Argentina originated from a secondary colonization (probably from the Chilean populations established around 1978), and thus discard the possibility of a new and independent colonization from the Palearctic region.

The O₅ chromosomal line from Zanjic (collected in June 1997) was obtained following the procedure presented in ZIVANOVIC et al. (2000). Samples from Scandinavia (Sunne and Drobak) and Argentina (Mar del Plata) were collected in August 2005 and November 2005, respectively. The Scandinavian and Argentinean O₅ chromosomal lines were obtained using the pattern of genetic crosses described in MESTRES et al. (1998). The chromosomal line from Zanjic had been previously analyzed for its viability and was proved to carry a lethal gene not allelic with that found in the American O₅ inversions (ZIVANOVIC and MESTRES 2000). The chromosomal line from Sunne was lethal and not allelic with those found in America, whereas that from Drobak presented normal viability. The O₅ from Mar del Plata carried the lethal gene completely associated with this inversion in America (FERNÁNDEZ-IRIARTE et al. 2009). The *Odh* gene of all these O₅ lines was sequenced using the procedure described in MESTRES et al. (2004) and GÓMEZ-BALDÓ et al. (2008). The nucleotide haplotypes obtained are shown in Figure 1 together with the three sequences previously obtained from European O₅ lines (MESTRES et al. 2004). The chromosomal lines from Mar del Plata, Sunne and Zanjic presented the same haplotype (Haplotype 1) found in all O₅ inversions from America and also detected in Taulé (France) and Lilla-Edet (Sweden). However, the O₅ from Drobak presented the other haplotype (Haplotype 2) also found in Europe (in Louvain-la-Neuve, Belgium) (MESTRES et al. 2004). As the O₅ chromosomal line from Mar del Plata carries Haplotype 1, this result supports the hypothesis that the colonization of the Atlantic coast of Argentina is a secondary event from the original colonization of Chile (around 1978), in opposition to an independent invasion from the Old World (FERNANDEZ-IRIARTE et al. 2009). Until now, all

evidence indicates that only one O_5 (carrying the lethal gene and the *Odh* haplotype completely associated with this inversion in America) reached the American continent (for a revision see MESTRES et al. 2005).

Table 1: Nucleotide haplotypes for the *Odh* gene obtained from the O_5 chromosomal lines of Mar del Plata (Argentina), Sunne (Sweden), Zanjic (Montenegro) and Drobak (Norway). The polymorphic positions shown are those present in the American populations of *D. subobscura* and the *Odh* haplotype of the *Va* balancer chromosome is shown as a reference. For the aim of comparison, the haplotypes of the O_5 chromosomal lines from Taulé (France), Lilla-Edet (Sweden) and Louvain-la-Neuve (Belgium) have also been included (MESTRES et al. 2004).

	20	87	102	103	119	167	213	221	254	272	305	311	329	374	455	479	491	509	510	557	572	575	602	604	609	610	614	616	628	644	655	680	683	686
Va	T	C	G	T	T	C	G	C	C	C	T	G	C	C	C	G	G	G	A	G	C	A	G	T	C	G	T	C	A	G	C	A	T	A
Mar del Plata					T					A							C		T	C				A	T	A	G		C	A		T	C	T
Sunne					T					A							C		T	C				A	T	A	G		C	A		T	C	T
Zanjic					T					A							C		T	C				A	T	A	G		C	A		T	C	T
Drobak					T												C		C						T	A	G		C	A		T	T	
Taulé					T					A							C		T	C				A	T	A	G		C	A		T	C	T
Lilla-Edet					T					A							C		T	C				A	T	A	G		C	A		T	C	T
Louvain-la-Neuve					T												C		C						T	A	A		C	A		T	T	

In the Palearctic populations only two haplotypes are detected. These haplotypes differ only in four positions and are in complete association with the inversion O_5 , i. e., they are not detected in other chromosomal arrangements. Moreover, after analyzing a total of 132 sequences from Barcelona and Mt. Parnes, no association between chromosomal arrangements and *Odh* haplotypes has been detected (ARAÚZ et al. 2008). The complete association between Haplotype 1 and the O_5 inversion in America can be easily explained as a product of the founder event. However, the associations between both *Odh* haplotypes and the O_5 inversions in the Palearctic region require other explanations. Because the *Odh* gene is located close to the proximal breakpoint of this inversion, its linkage disequilibrium with the inversion is expected to be high. One possible explanation for this association could be that the O_5 was a relatively recent inversion with regard to evolutionary time. Most probably, this inversion appeared only once in a Palearctic population (carrying one *Odh* nucleotide haplotype), originating a strong linkage disequilibrium difficult to break because the *Odh* gene is located close to the proximal breakpoint of the inversion. From its point of origin, the inversion had to spread to a huge area as indicated by the presence of Haplotype 1 in Montenegro, France and Scandinavia. This expansion of the O_5 inversion is compatible with the levels of gene flow observed in this species in the European and American regions

(LATORRE et al. 1992; PASCUAL et al. 2001; ZIVANOVIC et al. 2007; MESTRES et al. 2009). A sporadic genetic exchange (crossing-over or gene conversion) could explain the presence of Haplotype 2 associated to the O₅ inversion. An alternative explanation could be a selective advantage of this association. Most likely the inversion would contain several adaptive genes, a phenomenon well documented in the literature (for a recent revision see HOFFMANN and RIESEBERG 2008) and the possibility could exist that the *Odh* gene was included in a co-adapted complex. If this inversion was relatively ancient, more *Odh* nucleotide haplotypes (arisen by genetic exchange in any of its forms) with equivalent or rather similar fitness would be expected to be found. Thus, the paucity of haplotypes detected points to a recent origin of the O₅ inversion. Finally, a third possibility consists in the recent appearance of an advantageous allele close to the *Odh* gene that was able to spread to fixation within the O₅ gene pool. However, the advantageous effect of this allele would be weak, because the inversion O₅ is found at low frequencies in the Palearctic populations where it is detected. For all these reasons, it seems reasonable to assume a recent origin of the inversion.

Acknowledgements

We thank Mr. R. Rycroft (S.A.L. Universitat de Barcelona) for corrections to the English manuscript. This work was supported by grants CGL2006-13423-C02-02 of M. E. C. (Spain), 2009 SGR 636 from the Generalitat de Catalunya (Spain), 143040 Ministry of Science and Technological Development (Serbia) and CONICET (PIP 6357/05-06) grant awarded to PFI (member of the Scientific Researcher Program of CONICET, Argentina).

Literature cited

- Araúz, P. A., Mestres, F., Peris-Bondia, F., Latorre, A. and Serra, L. 2008. Molecular analysis of a colonization: *Odh* gene variability in Mediterranean and American populations of *D. subobscura*. — Annual Meeting of the Society of Molecular Biology and Evolution. Barcelona (Spain).
- Ayala, F. J., Serra, L. and Prevosti, A. 1989. A grand experiment in evolution: the *D. subobscura* colonization of the Americas. — *Genome* 31: 246–255.
- Balanyà, J., Oller, J. M., Huey, R. B., Gilchrist, G. W. and Serra, L. 2006. Global genetic change tracks global climate warming in *D. subobscura*. — *Science* 313: 1773–1775.

- Balanyà, J., Serra, L., Gilchrist, G. W., Huey, R. B., Pascual, M., Mestres, F. and Solé, E. 2003. Evolutionary pace of chromosomal polymorphism in colonizing populations of *D. subobscura*: An evolutionary time series. — *Evolution* 57: 1837–1845.
- Brcic, D., Prevosti, A., Budnik, M., Monclús, M. and Ocaña, J. 1981. Colonization of *D. subobscura* in Chile. I. First population and cytogenetic studies. — *Genetica* 56: 3–9.
- Fernández-Iriarte, P. J., Balanyà, J., Pascual, M., Mestres, F., Hasson, E., Fontdevila, A. and Serra, L. (2009). Tracking the origin of an invasive species: *D. subobscura* in Argentina. — *J. Evol. Biol.* 22: 650–658.
- Gilchrist, G. W., Huey, R. B. and Serra, L. 2001. Rapid evolution of wing size clines in *D. subobscura*. — *Genetica* 112–113: 273–286.
- Gómez-Baldó, L., Latorre, A., Serra, L. and Mestres, F. 2008. Molecular variation in the *Odh* gene in Chilean natural populations of *D. subobscura*. — *Hereditas* 145: 154–162.
- Hoffmann, A. A. and Rieseberg, L. H. 2008. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? — *Annu. Rev. Ecol. Evol. Syst.* 39: 21–42.
- Krimbas, C. B. 1992. The inversion polymorphism of *D. subobscura*. — In: C. B. Krimbas, C. B. and Powell, J. R. (eds). *Drosophila* inversion polymorphism. CRC Press. Boca Raton, Florida, USA. p. 127–220.
- Krimbas, C. B. 1993. *D. subobscura*: Biology, Genetics and inversion polymorphism. Verlag Dr. Kovac. Hamburg.
- Latorre, A., Hernández, C., Martínez, D., Castro, J. A., Ramón, C. and Moya, A. 1992. Population structure and mitochondrial DNA gene flow in Old World populations of *D. subobscura*. — *Heredity* 68: 15–24.
- Mestres, F., Sanz, J. and Serra, L. 1998. Chromosomal structure and recombination between inversions in *D. subobscura*. *Hereditas* 128: 105–113.
- Mestres, F., Balanyà, J., Arenas, C., Solé, E. and Serra, L. 2001. Colonization of America by *D. subobscura*: Heterotic effect of chromosomal arrangements revealed by the persistence of lethal genes. — *Proc. Natl. Acad. Sci. USA* 98: 9167–9170.

- Mestres, F., Abad, L., Sabater-Muñoz, B., Latorre, A. and Serra, L. 2004. Colonization of America by *D. subobscura*: Association between *Odh* gene haplotypes, lethal genes and chromosomal arrangements. — *Genes Genet. Syst.* 79: 233–244.
- Mestres, F., Balanyà, J., Pascual, M., Arenas, C., Solé, E. and Serra, L. 2005. Lethal genes and the colonization of America by *D. subobscura*. — *Current Topics in Genetics* 1: 31–57.
- Mestres, F., Balanyà, J., Pascual, M., Arenas, C., Gilchrist, G. W., Huey, R. B. and Serra, L. 2009. Evolution of Chilean colonizing populations of *D. subobscura*: lethal genes and chromosomal arrangements. — *Genetica* 136: 37–48.
- Pascual M., Aquadro C. F., Soto V. and Serra, L. 2001. Microsatellite variation in colonizing and Palearctic populations of *D. subobscura*. — *Mol. Biol. Evol.* 18: 731–740.
- Pascual, M., Chapuis, M. P., Mestres, F., Balanyà, J., Huey, R. B., Gilchrist, G. W. Serra, L. and Estoup, A. 2007. Introduction history of *D. subobscura* in the New World: a microsatellite-based survey using ABC methods. — *Mol. Ecol.* 16: 3069–3083.
- Prevosti, A., Serra, L., Aguadé, M., Ribó, G., Mestres, F., Balañà, J. and Moclús, M. 1989. Colonization and establishment of the Palearctic species *D. subobscura* in North and South America. — In: Fontdevila, A. (ed.), *Evolutionary biology of transient unstable populations*. Springer-Verlag, Berlin, p. 114–129.
- Prevosti, A., Ribó, G., Serra, L., Aguadé, M., Balanyà, J., Monclús, M. and Mestres, F. 1988. Colonization of America by *D. subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. — *Proc. Natl. Acad. Sci. USA* 85: 5597–5600.
- Zivanovic, G. and Mestres, F. 2000. Lethal genes in O₅ chromosomes of *D. subobscura* from Europe and America. — *J. Zool. Syst. Evol. Res.* 38: 123–126.
- Zivanovic, G., Andjelkovic, M., and Marinkovic, D. 2000. Genetic load and coadaptation of chromosomal inversion. II. O-chromosomes in *D. subobscura* populations. — *Hereditas* 133: 105–113.
- Zivanovic, G., Arenas, C., Mestres, F. 2007. The genetic structure of Balkan populations of *D. subobscura*. — *Hereditas* 144: 120–128.

**IV. Flux gènic entre ordenaments de
*D. subobscura.***

3. 4. Flux gènic entre ordenaments de *D. subobscura*.

Publication 5: The effect of recombination on medium sized inversions: selection and gene flux

Publicació 5: L'efecte de la recombinació sobre les inversions de mida mitjana: selecció i flux gènic

RESUM

Els patrons d'organització genòmica poden estar influenciats per la recombinació, que és una força evolutiva important. En general, la recombinació és avantatjosa, com demostra l'evolució del sexe en la majoria dels organismes. No obstant, la recombinació pot ser desavantatjosa en algunes situacions, com ara quan es produeix el trencament de combinacions de gens adaptatives. Aquest treball s'ha realitzat amb individus provinents de dues poblacions (Barcelona, Espanya i Mount Parnes, Grècia) i es basa en l'estudi de sis gens (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa*, *Fmr1*) localitzats al cromosoma O de *D. subobscura* en tres ordenaments considerats adaptatius (O_{3+4+1} , O_{3+4+7} and O_{3+4}). Tot i que s'ha detectat diferenciació genètica significativa en els gens localitzats a l'interior dels ordenaments, sembla que les inversions de mida mitjana (O_1 i O_7) reduïrien menys la recombinació que no pas les inversions imbricades (O_{3+4}). La diversitat nucleotídica estimada va ser similar en tots els casos, fins i tot en els gens situats a la vora dels punts de trencament. La recombinació, ja sigui produïda per entrecreuament o conversió gènica, sembla que no està totalment suprimida a l'interior de les inversions. El gen *Atpa*, que es troba localitzat entre inversions no imbricades, podria estar influenciat per la selecció associada a les inversions O_1 i O_7 , i l'escombrat selectiu (*selective sweep*) sembla que es va produir fa 0,04 i 0,06 Ma respectivament. Finalment, l'edat dels ordenaments O_{3+4+1} i O_{3+4+7} es va estimar en 0,23 Ma i 0,21 Ma respectivament.

The effect of recombination on medium-sized inversions: selection and gene flux

CINTA PEGUEROLES^{1*}, CHARLES F. AQUADRO², FRANCESC MESTRES¹ AND MARTA PASCUAL¹

¹Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Avd. Diagonal 645, 08028, Barcelona, Spain.

²Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853

*Corresponding author: cpegueroles@ub.edu

Abstract

The patterns of genome organization can be influenced by recombination, which is an important evolutionary force. In general, the presence of recombination is advantageous, as shown by the evolution of sex in most complex organisms. However, recombination can be disadvantageous in some situations, breaking some adaptive combinations of genes. The present study, which was carried out using individuals from Barcelona (Spain) and Mount Parnes (Greece), focus on the study of six genes (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa*, *Fmr1*) located across the O chromosome of *D. subobscura*, in relation to three potential adaptive arrangements (O_{3+4+1} , O_{3+4+7} and O_{3+4}). Despite significant genetic differentiation using F_{ST} was found between genes located within arrangements, single medium-size inversions (O_1 and O_7) seem to present a more relaxed recombination reduction than overlapped inversions (O_{3+4}). Nucleotide diversity was similar for all genes, even for those located close to inversions breakpoints. Recombination, which can be due to crossover and gene conversion, seems not to be totally suppressed inside inversions. The *Atpa* gene, which is located between non-overlapped inversions, seems to be influenced by selection relative to O_1 and O_7 arrangements with non synonymous fixed differences between them. The selective sweeps could be dated in 0.04 Myr for O_1 and 0.06 Myr for O_7 . Finally, the age of arrangements O_{3+4+1} and O_{3+4+7} was estimated in 0.23 Myr and 0.21 Myr respectively.

Key words: *D. subobscura*, nuclear genes, genetic differentiation, recombination, gene conversion, selection

Introduction

Inversions are chromosomal mutations in which a segment of the chromosome breaks off and attaches in the reverse direction being encountered in many organism such as bacteria, plants, insects and mammals (HOFFMANN *et al.* 2004). Inversions play an important role modulating patterns of recombination in the whole genome. It is known that for heterokaryotypic individuals only even crossovers inside the inverted regions give rise to balanced meiotic products (STURTEVANT 1926). Crossovers are highly reduced inside as well as outside the inverted regions (PEGUEROLES *et al.* (a); STUMP *et al.* 2007; WHITE and MORLEY 1955). Some studies showed that recombination increases in the region of the chromosome non covered by inversions (KOMAI and TATAKU 1940), despite this is not a general pattern (PEGUEROLES *et al.* (a)). Furthermore inversions can increase recombination in other chromosomes (SCHULTZ and REDFIELD 1951).

The formation of an inversion seems to be a unique event (POWELL 1997), whether mediated by ectopic recombination or staggered breaks models (CASALS and NAVARRO 2007). Thus, genetic differentiation among arrangements should be high when a new inversion arises. This differentiation would be eroded through time by even crossovers and gene conversion, and to a lesser extend by mutation. However, genetic differentiation is expected to be maintained in areas around the inversion breakpoints where meiotic pairing is almost blocked (NAVARRO *et al.* 1997) and in areas where associations with alleles are maintained by selection (HOFFMANN and RIESEBERG 2008). The adaptive value of inversions, has been shown in several species of dipterans, where they can be more easily studied due to the presence of giant chromosomes in some tissues (PREVOSTI *et al.* 1988; STUMP *et al.* 2005; HOFFMANN and RIESEBERG 2008). Several hypotheses have been proposed to explain their adaptive value (revised in HOFFMANN and RIESEBERG 2008). Under the coadaptation hypothesis, the adaptive value of inversions is determined by a certain set of alleles in a gene arrangement which confer higher fitness when combined with another gene arrangement from the same population (DOBZHANSKY 1950). Thus, this model is divided into two components; the first, the presence of epistatic interactions among genes within the inverted region and among genes of different arrangements from the same population and the second, different coadapted genes could be maintained for the same chromosomal arrangements in different populations. Alternatively, it has been postulated that adaptive inversions can be spread in natural populations due to locally adaptive alleles where no epistatic

interactions neither differences among populations are required (KIRKPATRICK and BARTON 2006). The high levels of linkage disequilibrium among non adjacent genes were interpreted as an evidence for epistasis within inversions, although linkage disequilibrium patterns were independent from the population of origin (SCHAEFFER *et al.* 2003). In several *Drosophila* species, some gene arrangements are formed by complexes of non-overlapped inversions, and their maintenance in natural populations was related to epistatic interactions among them (SPERLICH and FEUERBACH-MRAVLAG 1974; KUMAR and GUPTA 1991). However, it has also been suggested that this strong linkage disequilibrium could be attributed to the absence of recombination in heterozygote females (KRIMBAS and ZOUROS 1969; KRIMBAS 1993). Nevertheless, in *D. subobscura*, it was shown that the O₇ inversion, which is seldom found in natural populations, was formed by recombination between the O_{ST} and O₃₊₄₊₇ arrangements ruling out this later explanation (GÓMEZ-BALDÓ *et al.* 2008). *D. subobscura* is a model organism that has been largely studied due to its rich inversion polymorphism. Its chromosomal inversion frequency presents latitudinal clines as a consequence of their adaptive value and not of historic events, as shown by the same clines evolving in colonizing populations from both American hemispheres (PREVOSTI *et al.* 1988).

The aim of the present work is to track the footprint of selection associated to thermal adaptation mediated by chromosomal arrangements. The present research focus on three arrangements of the O chromosome of *D. subobscura*: O₃₊₄₊₁, O₃₊₄₊₇ and O₃₊₄, which in Europe and America present significant clines negatively correlated with latitude (PREVOSTI *et al.* 1988; BALANYÀ *et al.* 2003; BALANYÀ *et al.* 2004), thus suggesting that selection could be shaping their genetic patterns favoring them in warm environments. The O₃₊₄ arrangement is widespread in the Palearctic region (KRIMBAS and LOUKAS 1980; KRIMBAS 1992; BALANYÀ *et al.* 2004), while the O₃₊₄₊₁ arrangement is mainly restricted to the Balkan Peninsula (KRIMBAS and LOUKAS 1980; KRIMBAS 1992; ARAÚZ *et al.* 2009) and the O₃₊₄₊₇ arrangement presents higher frequency in the Iberian and the Anatolia Peninsulas (KRIMBAS and LOUKAS 1980; KRIMBAS 1992; ARAÚZ *et al.* 2009). Their distribution reflects the origin of the chromosomal lines used in the present study (O₃₊₄₊₁ from Mount Parnes, O₃₊₄₊₇ from Barcelona and O₃₊₄ from both populations). We have selected six candidate genes suggested to be involved in thermal regulation, since they presented differences of expression among laboratory populations maintained at different temperatures (LAAYOUNI *et al.* 2007). Thus, some of

these genes could be influenced by selection on these adaptive arrangements. Its location inside and outside the inverted regions allowed us to estimate the influence of a single inversion in modeling the patterns of genetic variation across the chromosome.

Material and methods

Fly samples and DNA sequencing

Forty six isochromosomal lines for the O chromosome of *D. subobscura* from Barcelona (Spain) and Mount Parnes (Greece) were used (ARAÚZ *et al.* 2009). Specifically, three different arrangements were selected: 11 O_{3+4+1} and 12 O_{3+4} lines from Mt. Parnes and 10 O_{3+4+7} and 12 O_{3+4} lines from Barcelona. The six genes selected were *Pif1A* (PFTAIRE-interacting factor 1A), *Abi* (Abelson interacting protein), *Sqd* (Squid), *Yrt* (Yurt), *Atpa* (Na pump α subunit) and *Fmr1* (Fragile X mental retardation). Their location on the O chromosome was determined by *in situ* hybridization (Figure 1A) and their intronic and exonic content was previously determined (PEGUEROLES *et al.* (b)) (Figure 1B). Genomic DNA extraction, DNA amplification and sequencing reactions were carried out as reported in (PEGUEROLES *et al.* (b)). PCR fragment analyses were performed with a 3730 Analyzer (Applied Biosystems), in the Serveis Científico-Tècnics de la Universitat de Barcelona. Partial sequences were assembled with SeqMan II (DNASTAR) and complete sequences were multiply aligned with Clustal W (THOMPSON *et al.* 1994) implemented in BioEdit v7 (HALL 1999). *D. pseudoobscura* sequences were downloaded from Flybase (<http://flybase.org>) and were used as outgroups.

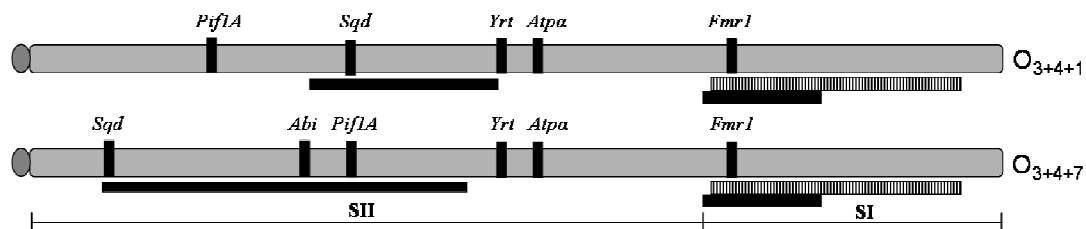


Figure 1A: Location of the 6 gene regions studied in the O_{3+4+1} and O_{3+4+7} arrangements. Black bars in section SII indicate the location of the O_1 or O_7 inversions and the black and striped bars in section SI the location of O_{3+4} inversions.

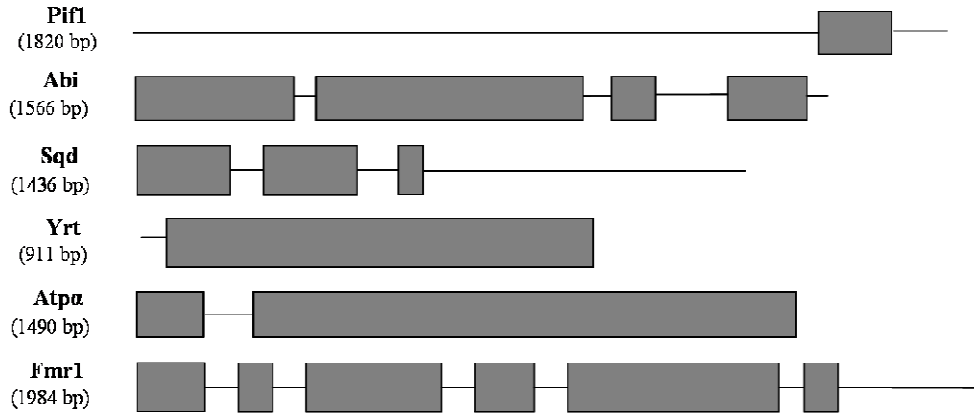


Figure 1B: Gene regions analyzed are represented, exons with grey boxes, and introns regions with bars.

Data analysis

The distance of each gene to the nearest inversion breakpoints in bp was estimated assuming that all cytological bands contain the same genetic content and the length of the O chromosome of *D. subobscura* is equivalent to the chromosome 2 of *D. pseudoobscura* (PEGUEROLES *et al.* 2010). Analyses were carried out for each region separately and for the concatenated data set. Due to the presence of duplications in the *Abi* gene (data not shown), very few individuals could be sequenced for the O_{3+4+1} arrangement. Concatenate data set does not contain *Abi* gene since only the chromosomal lines with all genes sequenced were included. Eleven O_{3+4} and nine O_{3+4+1} chromosomal lines for Mt. Parnes population and ten O_{3+4} and eight O_{3+4+7} chromosomal lines for Barcelona population were concatenated. Genes concatenation was carried out with Concatenator v1 (PINA-MARTINS and PAULO 2008). Nucleotide polymorphism, genetic differentiation, linkage disequilibrium, gene conversion, neutrality tests and selection tests were estimated with DnaSP v5 (LIBRADO and ROZAS 2009). The estimated standard parameters of nucleotide polymorphism were: the number of haplotypes (h), the number of polymorphic sites (S) and the number of singletons. It was also estimated the nucleotide diversity (π) (NEI 1987), nucleotide diversity in synonymous sites and non-coding positions (π_{sil}) (NEI and GOJOBORI 1986), the heterozygosity in silent sites (θ_{sil}) (WATTERSON 1975) and the divergence per silent site between *D. subobscura* and *D. pseudoobscura* (K_{sil}) (NEI and GOJOBORI 1986). The confidence intervals for the π_{sil} estimations were computed from coalescent simulations using recombination levels estimated from our nucleotide sequences. Genetic distance

was computed by F_{ST} (HUDSON *et al.* 1992) and its genetic differentiation significance was estimated as S_{nn} with 10000 replicates (HUDSON 2000). In all cases, analyses were carried out excluding sites with gaps for all chromosomal lines. Due to the high amount of insertion-deletion events (indels) observed in the *Pif1A* gene, analysis were also performed considering only indels with the tetraallelic model implemented in DnaSP v5 (LIBRADO and ROZAS 2009). Tajima's D (TAJIMA 1989) and Fu and Li's D (FU and LI 1993) neutrality tests were estimated using *D. pseudoobscura* as an outgroup. Finally, HKA test (HUDSON *et al.* 1987) and McDonald and Kreitman test (MCDONALD and KREITMAN 1991) were carried out to detect the footprint of selection.

For each locus and also for the concatenated data set, we estimated the percentage of pairwise comparisons between informative sites presenting significant linkage disequilibrium (LD). Their statistical significance was analyzed with the Chi-square test after Bonferroni correction (WEIR 1996). The overall linkage disequilibrium statistic ZnS was also calculated (KELLY 1997). Recombination estimations were obtained using different approaches, thus, the population recombination rate (ρ) was estimated using the composite likelihood method of Hudson (HUDSON 2001) that was computed with the LDhat v2.1 program (<http://www.stats.ox.ac.uk/~mcvean/LDhat>). Recombination networks were constructed using the SplitsTree4 program (HUSON and BRYANT 2006). Finally, gene conversion tracts (GCT) were obtained using the (BETRÁN *et al.* 1997) method implemented in DnaSP v5 (LIBRADO and ROZAS 2009).

The age of the inversions were computed for all genes located inside the inverted regions (Figure 1). It was estimated using the average silent nucleotide diversity of the individuals analyzed for each arrangement except those carrying gene conversion tracts (ROZAS *et al.* 1999; NÓBREGA *et al.* 2008). The number of substitutions per site per year was estimated using the divergence per silent site between *D. subobscura* and *D. pseudoobscura*, assuming that the two species diverged 8 Mya (RAMOS-ONSINS *et al.* 1998).

Results

Nucleotide variation

The concatenated data set for all genes included 7743 sites. After excluding gaps, 7722 sites remained for O_{3+4+1} and O_{3+4} arrangements from Mt. Parnes and 7725 for O_{3+4+7} and O_{3+4} from Barcelona. All indel polymorphisms were detected in non-coding regions. A total of 281 polymorphic sites (48% singletons), which corresponded to at least 292 mutations were detected in Mt. Parnes, and 266 polymorphic sites (45% singletons) in Barcelona, corresponding to a minimum of 269 mutations. For both populations, the nucleotide variation and divergence within chromosomal arrangements for the six studied gene regions is presented in Table 1. Due to the low number of sequences available for the *Abi* gene in the O_{3+4+1} arrangement, these sequences were not considered for further analyses. In both populations, the number of haplotypes was approximately the same as the number of sequences analyzed, except for *Sqd* and *Atpa* genes in O_{3+4+1} arrangement, in agreement with the lower number of polymorphic sites presented in these genes for that arrangement. The *PiflA* gene presented the highest number of polymorphic sites and nucleotide diversity (π) due to the highest proportion of intronic region amplified for this gene (Figure 1). In order to reduce the bias caused by the different proportion of intronic-exonic regions amplified for each gene, genetic variability was estimated only in silent positions (both synonymous and non-coding sites) (Table 1). No significant differences per gene were detected when comparing silent nucleotide variability (π_{sil}) among chromosomal arrangements (Figure 2), with the exception of the *Atpa* gene that presented significant lower values for O_{3+4+1} arrangement in comparison to O_{3+4} arrangement. Thus, a reduction of the variability associated with the inversions was only detected in this gene despite being located outside O_1 and O_7 inversions. However, significant differences in the levels of variability were detected among some genes (Figure 2). For instance, *Yrt* and *PiflA* genes presented larger variability than *Sqd* and *Atpa* genes. Moreover, the *Yrt* gene showed not only the highest value for π_{sil} but also for θ_{sil} and K_{sil} , suggesting a high neutral substitution rate. In summary, values obtained for each gene were quite similar independently of the arrangement and the population of origin.

Table 1: Nucleotide variation and divergence within chromosomal arrangement for the six studied genes in both populations

	<i>Pif1A</i>		<i>Abi</i>		<i>Sqd</i>		<i>Yrt</i>		<i>Atpa</i>		<i>Fmr1</i>	
Mt.Parnes	O ₃₊₄₊₁	O ₃₊₄	O ₃₊₄₊₁	O ₃₊₄	O ₃₊₄₊₁	O ₃₊₄	O ₃₊₄₊₁	O ₃₊₄	O ₃₊₄₊₁	O ₃₊₄	O ₃₊₄₊₁	O ₃₊₄
<i>n</i>	9	12	3	12	10	12	11	12	9	12	10	11
<i>h</i>	9	12	3	12	5	12	11	12	4	12	10	11
<i>S</i>	104	120	8	32	12	25	28	33	6	22	30	38
Singletons	61	55	8	20	8	16	14	18	6	9	24	25
π	0.021	0.021	0.003	0.005	0.002	0.005	0.009	0.010	0.001	0.005	0.004	0.005
π_{sil}	0.021	0.023	0.008	0.014	0.005	0.006	0.031	0.034	0.003	0.015	0.008	0.009
θ_{sil}	0.027	0.026	0.008	0.025	0.008	0.008	0.033	0.042	0.005	0.013	0.013	0.012
K_{sil}^a	0.188	0.191	0.173	0.174	0.115	0.115	0.327	0.327	0.204	0.208	0.150	0.151
Barcelona	O ₃₊₄₊₇	O ₃₊₄	O ₃₊₄₊₇	O ₃₊₄	O ₃₊₄₊₇	O ₃₊₄	O ₃₊₄₊₇	O ₃₊₄	O ₃₊₄₊₇	O ₃₊₄	O ₃₊₄₊₇	O ₃₊₄
<i>n</i>	9	12	10	7	9	10	9	12	9	12	8	10
<i>h</i>	9	12	9	7	9	10	9	12	8	11	7	10
<i>S</i>	89	114	33	24	16	21	31	35	22	27	19	28
Singletons	54	42	21	15	11	14	18	21	17	16	12	18
π	0.017	0.022	0.006	0.006	0.003	0.005	0.010	0.010	0.004	0.005	0.003	0.004
π_{sil}	0.018	0.023	0.015	0.014	0.004	0.006	0.035	0.035	0.006	0.016	0.006	0.008
θ_{sil}	0.021	0.024	0.018	0.015	0.006	0.007	0.039	0.041	0.010	0.019	0.007	0.010
K_{sil}^a	0.187	0.190	0.172	0.173	0.115	0.115	0.325	0.326	0.204	0.209	0.150	0.151

n, sample size; *h*, number of haplotypes; *S*, number of polymorphic sites; π , nucleotide diversity in all sites; π_{sil} , nucleotide diversity in synonymous sites and non-coding positions; θ_{sil} , heterozygosity in silent sites; K_{sil}^a , divergence per silent site between *D. subobscura* and *D. pseudoobscura*.

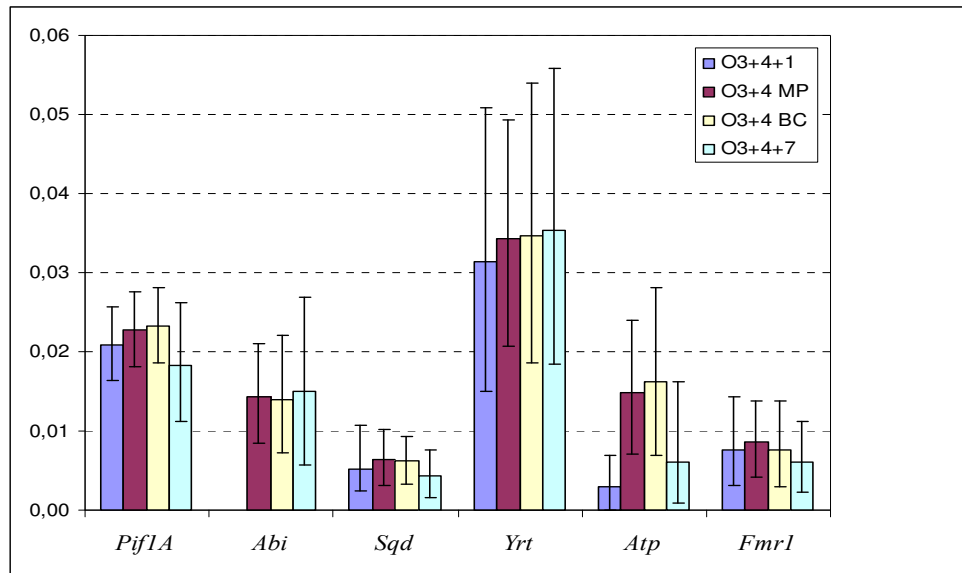


Figure 2: Silent nucleotide diversity estimations for each gene and arrangement, and their confidence interval estimated by coalescent simulations.

Neutrality tests

Tajima's *D* and Fu and Li's *D* tests were carried out to assess whether variability pattern within arrangements differ from those expected under neutrality (Table 2). In most of the comparisons (93%) values obtained were negative although only significant in the

O_{3+4+1} arrangement for *Atpa* gene. The footprint of selection was not detected for any of the genes studied as assessed by Mc Donald and Kreitman test, since the ratios of polymorphism and divergence in synonymous and nonsynonymous sites were not significantly different. Finally, HKA tests were in all cases not significant, suggesting that there are not significant differences in polymorphic variability relative to divergence among different regions.

Table 2: Tajima's D and Fu and Li's (with and without an outgroup) neutrality tests within each chromosomal arrangement

		Tajima's D	Fu and Li's D ^c
<i>Pif1A</i>	O_{3+4}^a	-0.580	-0.800
	O_{3+4}^b	-0.104	0.081
	O_{3+4+1}	-0.654	-0.972
	O_{3+4+7}	-0.357	-0.850
<i>Abi</i>	O_{3+4}^a	-1.144	-0.962
	O_{3+4}^b	-0.542	-1.001
	O_{3+4+7}	-0.899	-0.794
<i>Sqd</i>	O_{3+4}^a	-0.992	-1.656
	O_{3+4}^b	-0.810	-1.036
	O_{3+4+1}	-0.731	-1.598
	O_{3+4+7}	-1.234	-1.561
<i>Yrt</i>	O_{3+4}^a	-0.810	-0.224
	O_{3+4}^b	-1.024	-0.888
	O_{3+4+1}	-0.684	0.118
	O_{3+4+7}	-0.888	-0.789
<i>Atpa</i>	O_{3+4}^a	0.158	-0.612
	O_{3+4}^b	-0.721	-0.941
	O_{3+4+1}	-1.728*	-2.236*
	O_{3+4+7}	-1.347	-0.309
<i>Fmr1</i>	O_{3+4}^a	-1.401	-1.227
	O_{3+4}^b	-0.982	-1.225
	O_{3+4+1}	-1.620	-1.956
	O_{3+4+7}	-0.637	-1.432

^a O_{3+4} from Mt. Parnes population; ^b O_{3+4} from Barcelona population; ^c Fu and Li's D using *D. pseudoobscura* as an outgroup; * P<0.05

Genetic differentiation between arrangements

The estimates of genetic differentiation between $O_{3+4} - O_{3+4+1}$, $O_{3+4} - O_{3+4+7}$, $O_{3+4+1} - O_{3+4+7}$ and $O_{3+4} - O_{3+4}$ arrangements using F_{ST} are shown in Table 3. When possible, the comparison of arrangements pairs coming from the same or different populations was carried out. No fixed mutations were detected between $O_{3+4} - O_{3+4+1}$, $O_{3+4} - O_{3+4+7}$ and $O_{3+4+1} - O_{3+4+7}$ arrangements and the percentage of shared mutations never was superior to 50%. Similar results were obtained whichever was the population of origin of the O_{3+4} arrangement. Significant genetic differentiation was detected when analyzing the concatenate data set between $O_{3+4} - O_{3+4+1}$, $O_{3+4} - O_{3+4+7}$ and $O_{3+4+1} - O_{3+4+7}$. However, different behavior patterns were observed when analyzing genes separately. When comparing O_{3+4} arrangements from both populations, most of the F_{ST} values obtained were negative, indicating more differences within arrangements than between them. However, comparing different arrangements, most of the values were positive. Significant differences were detected for those genes located inside the inverted regions. Nevertheless, the *Atpa* gene, which is located outside the inversions and between the non-overlapped inversions, presented the largest F_{ST} values (Table 3). No correlation was found between F_{ST} values and the distance to the nearest inversion breakpoint (Figure 3). Most of the genes, except *Atpa*, presented quite similar values of F_{ST} independently of being located inside or outside the inversions.

Indels in the intronic region of *PiflA* gene were informative for genetic differentiation (supplementary material). For instance, S_{nn} values for $O_{3+4} - O_{3+4+1}$ arrangements were not significant when excluding gaps, but significant when considering only indel information. However, there were no informative indels for genetic differentiation in the $O_{3+4} - O_{3+4+7}$ arrangements. For the $O_{3+4+1} - O_{3+4+7}$ arrangements, the S_{nn} statistic detected significant differentiation when analyzing only indels, but K_s^* statistic was also significant when excluding indel events ($K_s^* = 3.391$, $K_{st}^* = 0.009$), in accordance to the slightly higher power of this statistic when levels of variation are very high (HUDSON 2000).

Table 3: Genetic differentiation between chromosomal arrangements estimated as F_{ST} and the statistical significance of S_{nn} , a measure of how often the nearest neighbours of sequences are found in the same population. Shaded genes are located inside inversions

	O_{3+4} BC	O_{3+4} MP	O_{3+4} BC	O_{3+4} BC	O_{3+4} MP	O_{3+4+1} MP
	O_{3+4} MP	O_{3+4+1} MP	O_{3+4+1} MP	O_{3+4+7} BC	O_{3+4+7} BC	O_{3+4+7} BC
<i>Pif1A</i>	-0.017	-0.028 ns	-0.020 ns	0.146 **	0.126 **	0.076 ns
<i>Pif1A^a</i>	-0.012	0.420 **	0.383 **	0.107 *	0.037 ns	0.371 **
<i>Abi</i>	-0.037	-	-	0.065 *	0.042 *	-
<i>Sqd</i>	-0.050	0.082*	0.071 *	0.040 *	0.016 *	0.071 **
<i>Yrt</i>	-0.008	0.017 ns	-0.033 ns	-0.001 ns	0.013 ns	0.024 ns
<i>Atp</i>	-0.044	0.614 ***	0.656 ***	0.525 **	0.482 **	0.367 ***
<i>Fmr1</i>	0.028	-0.033 ns	0.008 ns	0.004 ns	0.010 ns	-0.005 ns
Concatenated	-0.015	0.099 *	0.121 *	0.182 **	0.156 **	0.092 *

ns, not significant; *, 0.01<P<0.05; **, 0.001<P<0.01; ***, P<0.001; ^a analysis using only the total number of InDel events.

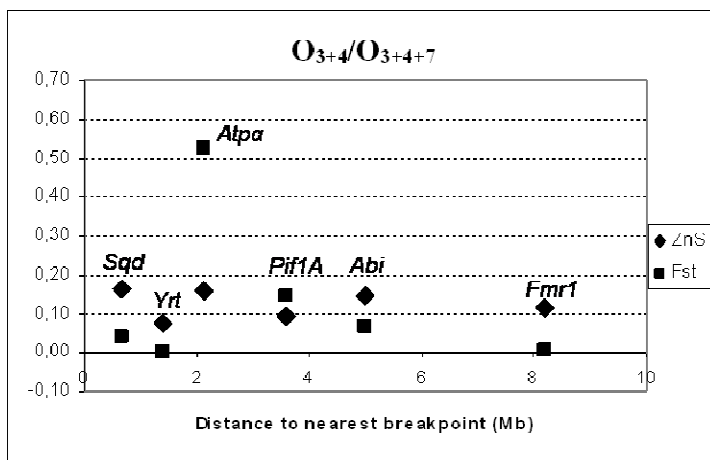
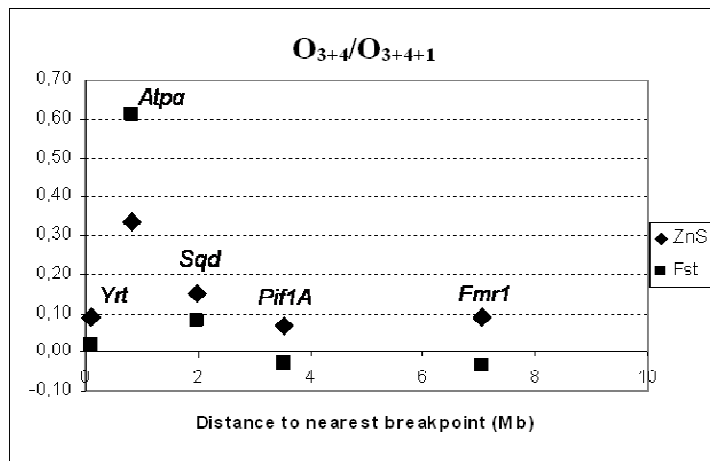


Figure 3: Genetic differentiation estimated as F_{ST} and LD using ZnS between arrangements O_{3+4} and O_{3+4+1} in Mt. Parnes (top) and between arrangements O_{3+4} and O_{3+4+7} in Barcelona population (bottom) vs their distance to the nearest inversion breakpoint.

Linkage disequilibrium

The number of informative sites for each locus and chromosomal arrangement are given in Table 4. For *Atpa* gene in O_{3+4+1} , no informative sites were observed since all variability within this arrangement is due to singletons (Table 1). The percentage of pairwise comparisons between informative sites presenting significant linkage disequilibrium (LD) is given in Table 4 with and without Bonferroni correction. After applying the Bonferroni procedure, as expected, values obtained for each arrangement separately were lower than when grouping them. The *Sqd* gene in Barcelona was the only exception to this observation. However, without correction, in general lower levels of linkage disequilibrium were obtained for the O_{3+4} arrangement. Only *Atpa* (in both populations) and *Sqd* (in Barcelona) presented higher LD levels when grouping sequences of different arrangements, as expected. Similarly, higher values were obtained for the *Atpa* gene when grouping arrangements for the ZnS statistic.

To test the relationship between LD and gene location, ZnS values obtained grouping sequences were plotted against the distance to the nearest inversion breakpoint (Figure 3). No relationship was found, since all genes presented ZnS values ranging from 0.07 to 0.15, except for *Atpa* gene with 0.34 and 0.24 for Mt. Parnes and Barcelona, respectively. Moreover, all R^2 values were plotted against the nucleotide distance (bp) for the concatenated data set (Figure 4). Despite the number of significant pairwise comparisons decreased with distance, some of them involved genes separated by 16 Mb in the O chromosome. The number of significant R^2 values was higher for the $O_{3+4} - O_{3+4+7}$ than for the $O_{3+4} - O_{3+4+1}$ (866 and 660 respectively), which could be influenced by the number of genes located inside the inverted region (three inside O_7 and one inside O_1). Between $O_{3+4} - O_{3+4+7}$ arrangements, 464 significant comparisons took place between intronic positions (16.4% between different genes), 287 between intronic and exonic regions (76% between different genes) and only 115 between exonic regions (39.1% between different genes). Similarly, between $O_{3+4} - O_{3+4+1}$ arrangements, 348 significant comparisons took place between intronic positions (42.8% between different genes), 224 between intronic and exonic regions (75.4% between different genes) and only 88 between exonic regions (33% between different genes). For the *Atpa* gene, the 63% and the 52% of the significant comparisons for the $O_{3+4} - O_{3+4+7}$ and the $O_{3+4} - O_{3+4+1}$ respectively, included nucleotides related to aminoacid changes.

Table 4: Association between informative polymorphic nucleotide sites, ZnS statistic values and the recombination parameter (Rho) per bp in Mt. Parnes and Barcelona populations. Shaded genes are located inside inversions.

	Mt. Parnes	Informative sites	% LD ^a	% LD ^b	ZnS ^c	Rho/length
<i>PiflA</i>	<u>O₃₊₄₊₁</u>	40	8.5	0	0.172	0.055*
	<u>O₃₊₄</u>	61	7.7	0	0.123	0.055*
	<u>O₃₊₄₊₁</u> & <u>O₃₊₄</u>	68	6.7	1	0.066	0.055*
<i>Sqd</i>	<u>O₃₊₄₊₁</u>	4	100	3	0.714	0.000
	<u>O₃₊₄</u>	8	10.7	3	0.212	0.055
	<u>O₃₊₄₊₁</u> & <u>O₃₊₄</u>	12	15.2	6	0.149	0.030
<i>Yrt</i>	<u>O₃₊₄₊₁</u>	13	12.8	0	0.195	0.046
	<u>O₃₊₄</u>	15	7.6	0	0.124	0.110*
	<u>O₃₊₄₊₁</u> & <u>O₃₊₄</u>	21	11.4	9	0.090	0.110*
<i>Atpa</i>	<u>O₃₊₄₊₁</u>	0				0.009
	<u>O₃₊₄</u>	13	21.8	4	0.214	0.038
	<u>O₃₊₄₊₁</u> & <u>O₃₊₄</u>	19	71.3	25	0.335	0.007
<i>Fmr1</i>	<u>O₃₊₄₊₁</u>	5	0	0	0.089	0.004
	<u>O₃₊₄</u>	13	9	0	0.167	0.021
	<u>O₃₊₄₊₁</u> & <u>O₃₊₄</u>	22	9.1	7	0.089	0.033
Barcelona						
<i>PiflA</i>	<u>O₃₊₄₊₇</u>	35	9.58	0	0.178	0.030
	<u>O₃₊₄</u>	71	12.72	0	0.147	0.055*
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	79	11	1.9	0.094	0.055*
<i>Abi</i>	<u>O₃₊₄₊₇</u>	12	32	0	0.327	0.008
	<u>O₃₊₄</u>	9	14	0	0.270	0.033
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	17	19.85	5	0.144	0.013
<i>Sqd</i>	<u>O₃₊₄₊₇</u>	5	10	1	0.206	0.021
	<u>O₃₊₄</u>	6	20.00	20	0.262	0.070*
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	10	22.22	8.9	0.165	0.070*
<i>Yrt</i>	<u>O₃₊₄₊₇</u>	13	11.54	0	0.182	0.067
	<u>O₃₊₄</u>	14	6.59	2.2	0.117	0.089
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	28	9	2.4	0.078	0.110*
<i>Atpa</i>	<u>O₃₊₄₊₇</u>	5	10	1	0.200	0.000
	<u>O₃₊₄</u>	11	14.55	5.5	0.204	0.018
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	24	42.75	17	0.262	0.003
<i>Fmr1</i>	<u>O₃₊₄₊₇</u>	7	33.33	0	0.365	0.006
	<u>O₃₊₄</u>	10	15.56	0	0.187	0.007
	<u>O₃₊₄₊₇</u> & <u>O₃₊₄</u>	13	16.67	2.6	0.114	0.012

^a percentage of pairwise comparisons with significant LD by the χ^2 test ($P < 0.005$)

^b percentage of pairwise comparisons with significant LD after Bonferroni correction (WEIR 1996)

^c ZnS over all pairwise comparisons using only informative sites

* Rho per gene > 100

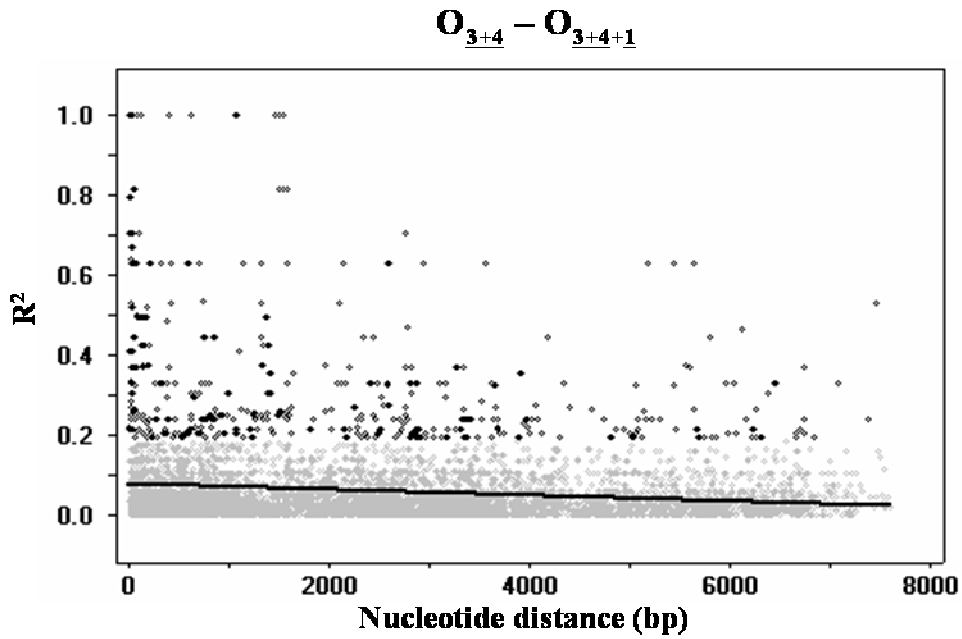


Figure 4A

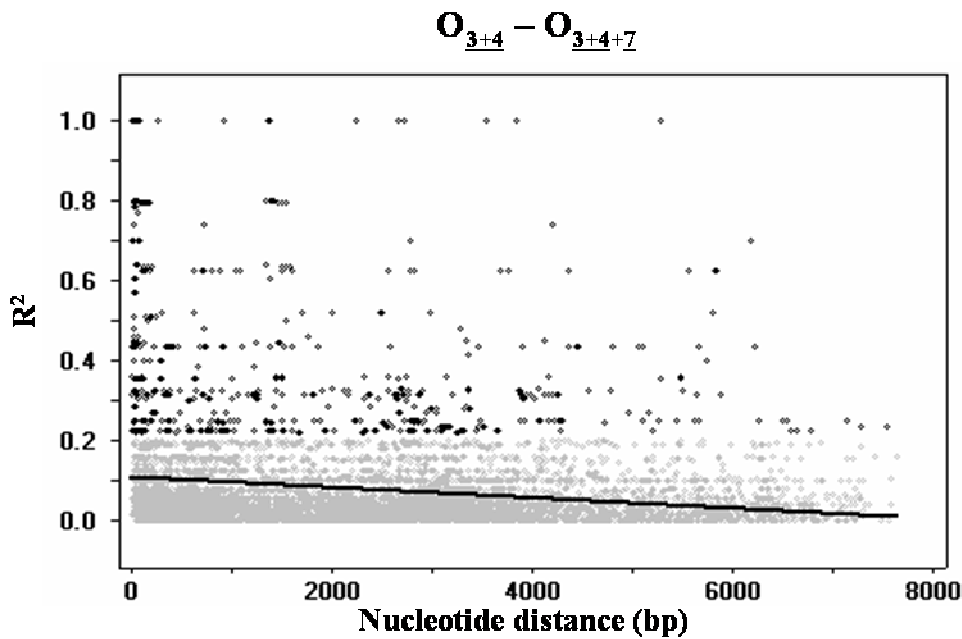


Figure 4B

Figure 4: Pairwise comparisons between informative sites for linkage disequilibrium that were significant after χ^2 test when grouping $O_{3+4+1} - O_{3+4}$ (A) and $O_{3+4+7} - O_{3+4}$ (B).

Gene flux

In each population, recombination per gene was computed for each arrangement separately and combining them (Table 4). In some cases it was not possible to estimate the exact amount of recombination, since Rho values greater than 100 are treated as 100 by the LDhat program. In all cases, recombination estimated within O_{3+4+1} or O_{3+4+7} arrangements were lower than for the O_{3+4} arrangement. When grouping arrangements, recombination was similar or higher than in the O_{3+4} arrangement for most genes with the exception of *Sqd* and *Atpa* genes in Mt. Parnes, and *Abi* and *Atpa* genes in Barcelona. Recombination networks were also constructed, where recombination events are graphically represented using blue lines. Due to high levels of variability in *Pif1A* gene and the concatenate data set, recombination networks could not be performed. For most genes, the reticulated network showing connections between arrangements suggest high levels of recombination among them (Figure 5). Only in the networks for the *Atpa* gene it was possible to distinguish two clades, corresponding to the two analyzed arrangements in each population (Figure 5). In this later case, it is also possible to observe that low number of recombination connections were detected within O_{3+4+1} and O_{3+4+7} chromosomal lines in front of the large number of connections in the O_{3+4} arrangement. Moreover, some chromosomal lines appeared in the incorrect group, suggesting that the entire sequence amplified recombined.

Gene conversion tracts (GCT) were estimated using (BETRÁN *et al.* 1997) approach (Table 5). Despite the ψ values were small, ranging from 10^{-3} to 10^{-4} in all cases, some GCT were detected. A larger number of tracts were observed between $O_{3+4} - O_{3+4+7}$ (8) than between $O_{3+4} - O_{3+4+1}$ (6). The length of the tracts was quite variable, from 7 to 1573 bp, but larger tracts were always found in *Atpa* gene; since GCT are expected to be small, large tracts observed might be due to crossover events.

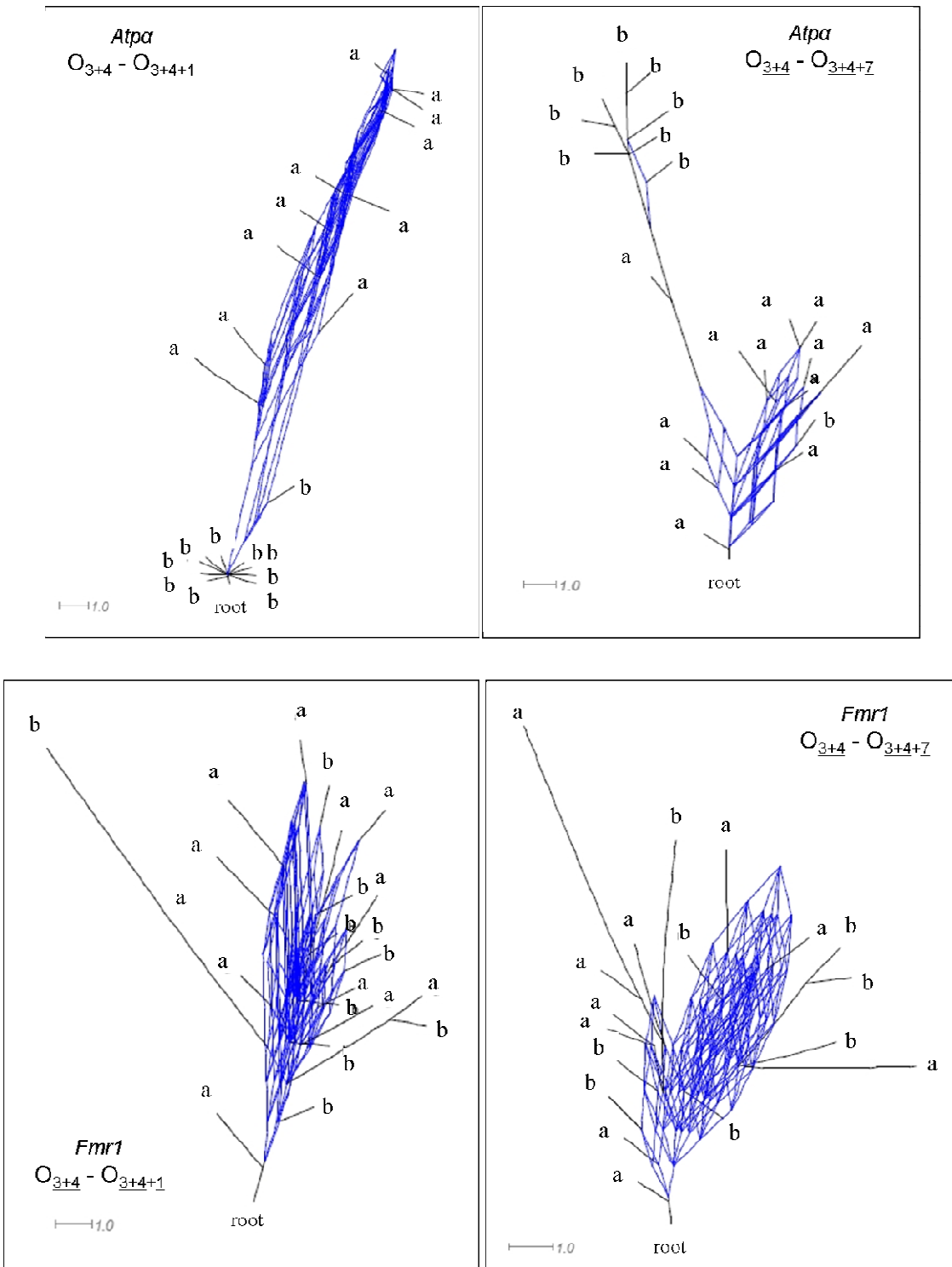


Figure 5: Recombination networks for *Atpa* and *Fmr1* genes for $O_{3+4+1} - O_{3+4}$ and $O_{3+4+7} - O_{3+4}$ comparisons. Individuals designated with letter **a** present the arrangement O_{3+4} and individuals designated with letter **b** present the arrangement O_{3+4+1} or O_{3+4+7} .

Table 5: Gene conversion tracts detected. Shadowed genes are located inside inversions

O_{3+4}/O_{3+4+1}	Converted ^a	Informative sites	ψ ^b	Length (bp)
<i>Pif1A</i>		1	0.00017	
<i>Sqd</i>		3	0.00041	
<i>Yrt</i>	O_{3+4}	4	0.00063	111
<i>Atpa</i>	O_{3+4}	17	0.00541	1573
<i>Atpa</i>	O_{3+4+1}	17	0.00541	52
<i>Fmr1</i>		0	0	
O_{3+4}/O_{3+4+7}				
<i>Pif1A</i>	O_{3+4}	32	0.00464	466
<i>Pif1A</i>	O_{3+4+7}	32	0.00464	103
<i>Pif1A</i>	O_{3+4+7}	32	0.00464	208
<i>Abi</i>		1	0.00023	
<i>Sqd</i>	O_{3+4+7}	3	0.00043	8
<i>Yrt</i>	O_{3+4+7}	2	0.00037	7
<i>Atpa</i>	O_{3+4}	18	0.00614	1422
<i>Atpa</i>	O_{3+4+7}	18	0.00614	1573
<i>Fmr1</i>		0	0	

^a converted arrangement; ^b probability of a site being informative of a gene conversion event.

Age of inversions

The neutral mutation rate was estimated from the mean silent nucleotide diversity since it seems to be more appropriate than to use silent nucleotide diversity from the two more divergent chromosomal lines (ROZAS *et al.* 1999; PEGUEROLES *et al.* 2010). Since only variability due to mutation is useful to make correct estimations, individuals presenting GCT were excluded from the analysis. The age of the inversions were estimated using only genes located inside inversions. However, genes located inside the inversion but in a central position are susceptible to present double crossovers and thus, the age of the inversion could be overestimated. For this reason, *Sqd* gene, which is the closest to O_7 inversion breakpoints, seems to be the more appropriated to carry out the estimation of the age of the O_{3+4+7} arrangement. If the age of the O_{3+4+7} arrangement is estimated using *Pif1A* and *Abi* genes the values obtained were 0.63 Myr and 0.70 Myr respectively. Nevertheless, when using *Sqd* gene, the age estimated was 0.21 Myr. The age of O_{3+4+1} arrangements was 0.23 Myr when using *Sqd* gene, which is the only analyzed gene located inside the O_1 inversion.

*Discussion***Genetic differentiation: single vs overlapped inversions**

Since the unique origin of most inversions has been widely accepted (POWELL 1997), during the first stages genetic variability within the inversion should be low and genetic differentiation among arrangements should be high. Through time, variability can be recovered by two non excluding mechanisms, mutation and recombination and the latter, which includes crossovers and gene conversion, seems to occur at a higher rate than mutation (SCHAEFFER and ANDERSON 2005). However, selective pressure can maintain genetic differentiation among arrangements (HOFFMANN *et al.* 2004). Non selective factors, such as the length and the age of inversions, can also contribute in the maintenance genetic differentiation among arrangements. A theoretical study suggested that the minimum genetic length necessary for double crossovers formation inside inverted regions is 20 cM (NAVARRO *et al.* 1997). Thus, the higher the length of the inversion, the higher would be the probability of double crossovers occurrence. Considering that the genetic length of the O chromosome of *D. subobscura* is 190.7 cM (PEGUEROLES *et al.* 2010), the genetic length of O₁ and O₇ inversions was estimated in 37.7 cM and 72.9 cM respectively, being higher than the minimum genetic length required for even crossovers formation (20cM). Thus, despite the probability of even crossovers could be larger in O₇ than in O₁, nucleotide diversity for both arrangements in the concatenated data set was quite similar (0.0072 vs 0.0073, respectively). It was suggested that the effect of inversions on nucleotide variability depends mainly on two factors: gene flux and the age of inversions (NAVARRO *et al.* 2000). In the present study, only data from *Sqd* gene (the closest to the inversion breakpoint) was used to estimate the age of O₃₊₄₊₇ arrangement, since other genes located inside inversions but in a more central position could be influenced by recombination. The age estimated for both arrangements is around 0.20 Myr, being younger than previous estimations of O₃₊₄ arrangement (NAVARRO-SABATÉ *et al.* 1999; ROZAS *et al.* 1999; PEGUEROLES *et al.* (b)), consistent with the derived status of O₃₊₄₊₁ and O₃₊₄₊₇ arrangements. Thus, since the ages of O₁ and O₇ inversions are quite similar, we can conclude that similar levels of gene flux are occurring in both inversions independently of their length. It seems that the length of the inversions does not seem to affect levels of recombination after a given threshold.

The study of eight genes in the SI region, located within the O_3 inversion, revealed high levels of genetic differentiation between the O_{3+4} and the O_{ST} arrangements, as well as lower genetic variability in the latter (MUNTÉ *et al.* 2005). They estimated the genetic length of O_3 inversion in 27.4 cM, considering that the genetic length of the O chromosome of *D. subobscura* is 228.3 cM (LOUKAS *et al.* 1979). However, with the newer estimation of its genetic length (190.7 cM) (PEGUEROLES *et al.* 2010), the length of O_3 inversion was estimated in 22.9 cM, being very close to the minimum size necessary to allow double crossovers formation. Furthermore, it is worth pointing out that O_3 inversions although ancestral is never found alone in natural populations, and it is always encountered forming a complex of two overlapped inversions. Thus both genetic differentiation and variability could be influenced not only by the small length of the inversion but also by the presence of overlapped inversions.

In the present study, F_{ST} values obtained for arrangements differing by single inversions were significant but smaller than compared to overlapped inversions (O_{ST} vs O_{3+4+X} , where X means another inversion) (ROZAS *et al.* 1999; MUNTE *et al.* 2005; PEGUEROLES *et al.* (b)). Similarly, genetic differentiation between A_2 and A_{ST} arrangements (also differing by a single inversion), was significant but F_{ST} values were also smaller than when comparing with the overlapped inversions (NÓBREGA *et al.* 2008). Finally, one of the genes located inside SI region (*rp49*) was also sequenced in O_{3+4+8} and O_{3+4+23} arrangements, and again F_{ST} values obtained comparing these arrangements with the O_{3+4} (also differing only by a single inversion) were significant but lower than when comparing overlapping inversions (ROZAS *et al.* 1999). The same pattern was observed in *D. pseudoobscura*, since high levels of genetic differentiation were found when comparing arrangements differing by overlapped inversions (SCHAEFFER *et al.* 2003). Thus, high levels of genetic differentiation were only found in those comparisons involving multiple overlapped inversions implying the formation of at least two inversions loops during meiosis (O_{ST} vs O_{3+4+X}).

The rule of higher differentiation involving overlapping inversions seem not to be accomplished when comparing O_{3+4+8} and O_{3+4+23} arrangements (O_8 and O_{23} overlapped) (ROZAS *et al.* 1999) or O_{3+4+1} and O_{3+4+7} (O_1 and O_7 also overlapped) (present research). This lack of high differentiation could be explained by

recombination mediated by other arrangements differing by a single non-overlapped inversion. Most of the inversions described in the O chromosome of *D. subobscura* form complex arrangements with O_{3+4} inversion (ARAÚZ *et al.* 2009), which presents high frequency in southern European populations, the area where most of the analyzed samples come from (KRIMBAS 1992; KRIMBAS 1993). Thus, O_{3+4+8} , O_{3+4+23} , O_{3+4+1} and O_{3+4+7} chromosomes could recombine with one of the several arrangements differing only by one inversion. On the contrary, during meiosis, heterokaryotypes carrying the O_{ST} arrangement and another arrangement in general will form two inversion loops. Thus, recombination rate in heterokaryotypes for the O_{ST} arrangement in the SI region should be highly reduced. The O_{ST} could more easily maintain a cluster of adaptive genes conferring a selective advantage and explaining the significant latitudinal clines in Europe and America (PREVOSTI *et al.* 1988; BALANYÀ *et al.* 2003).

Nucleotide variability and gene distribution across the inversion

The recovery of variability is not expected to be uniform across the inverted region, since recombination is expected to be more reduced near breakpoints due to chromosomal mispairing (NAVARRO *et al.* 1997). Thus, the location of genes inside the inverted regions can be a key factor to take into account when studying levels of nucleotide diversity. We have detected significant differences in the silent nucleotide levels, despite no relation with its location inside the inversion was detected. Moreover, we have not detected inferior variability in our genes located fairly close to breakpoints, as observed in previous experimental studies of *D. subobscura* (MUNTÉ *et al.* 2005; NÓBREGA *et al.* 2008) despite in other species such as *D. buzzatti* the reduction was detected (LAAYOUNI *et al.* 2003). Furthermore, no high F_{ST} values were found for *Yrt* (outside the inversion) and *Sqd* genes (inside the inversion) despite being located very close to O_1 and O_7 breakpoints, respectively (Figure 2). Nevertheless, *rp49* and *P236* genes presented higher levels of genetic differentiation when located near inversion breakpoints (ROZAS and AGUADÉ 1994; NÓBREGA *et al.* 2008) than when located in a more central position (ROZAS *et al.* 1999). Thus, genetic differentiation close to inversion breakpoints can also be eroded through time at a gene specific rate.

Estimates of linkage disequilibrium and recombination

The observation of high strong genetic differentiation for genes located across an inverted region was interpreted in two possibilities: the lack of double crossovers or selection acting against recombinant individuals (MUNTÉ *et al.* 2005; SCHAEFFER and ANDERSON 2005). However, the low and similar F_{ST} values that we detect for genes located within the O_{3+4+7} arrangement, independently of their position relative to the breakpoint, are compatible with recombination taking place inside the inversions whether due to crossovers formation or gene conversion.

If recombination in heterokaryotypes is reduced and knowing that recombination in homokaryotypes is free, we would expect lower levels of LD within arrangement than when grouping sequences from two arrangements. However, this pattern was only clearly observed in *Atpa* gene despite being located outside the O_1 and O_7 inversions. The lower ZnS values when grouping sequences from different arrangements suggest that recombination between them would not be negligible. Moreover, the low genetic differentiation detected for most of the genes, in spite of their location and the genetic size of the inversions also suggest that recombination between arrangements is more frequent than previously thought (NAVARRO *et al.* 1997). We estimated recombination for each gene and arrangement using the composite likelihood method of Hudson (HUDSON 2001), since it is a fast methodology that performs well in terms of variance in comparison to other estimators (MCVEAN *et al.* 2002). Thus, we obtained global recombination estimations, since this methodology does not distinguish among crossovers and gene conversion. For *Pif1A*, *Yrt* and *Sqd* genes it was not possible to determine the exact value of recombination since Rho was higher than 100, which is the maximum value allowed by the program (MCVEAN *et al.* 2002). It seems that high mutation rate can cause homoplasy that mimics the effects of recombination (MCVEAN *et al.* 2002). The *Pif1A* and *Yrt* genes showed the highest levels of variability in the O_{3+4} arrangement, in the first gene probably due to large proportion of intronic content amplified and the second gene seems to present a high mutation rate. However, for *Sqd* gene Rho was also higher than 100 for the O_{3+4} arrangement in Barcelona despite the estimated low value of π , suggesting true high recombination rate.

As expected, some amount of recombination was detected for all analyzed genes. First, new variants can appear *de novo* by mutation and can be spread in

populations by recombination between homokaryotypes. Second, recombination in heterokaryotypes carrying a single inversion could not be totally suppressed, thus increasing variability in the derived inversions. For all studied genes recombination in O_{3+4} was higher than in O_{3+4+1} and O_{3+4+7} arrangements. On one hand, the O_{3+4} arrangement is older than the other two that can be considered derived arrangements. On the other hand, the frequency of O_{3+4} arrangement is higher than the two others (ARAÚZ *et al.* 2009) suggesting a higher effective size. Thus, assuming Hardy-Weinberg equilibrium, the expected frequency of O_{3+4} homokaryotypes in Mt. Parnes and Barcelona (0.365 and 0.104, respectively) would be larger than the expected for O_{3+4+1} and O_{3+4+7} homokaryotypes (0.026 and 0.030, respectively). Moreover, for all studied genes, larger recombination values were obtained for the O_{3+4} arrangement in Mt. Parnes, in agreement with the higher abundance of this arrangement in that population (ARAÚZ *et al.* 2009). Recombination estimates are expected to be highly reduced in heterokaryotypes by the presence of the inversion (NAVARRO *et al.* 1997). This influence was observed when grouping sequences for genes *Sqd* and *Abi*, which are located inside the inverted region, and also in *Atpa* gene, located outside the inversion. Thus, this is in agreement with the observation that recombination is affected inside but also outside the inverted regions (WHITE and MORLEY 1955; PEGUEROLES *et al.* (a)).

To estimate gene conversion we selected the (BETRÁN *et al.* 1997) method which depends mainly in two parameters: ψ , that it is the probability of a site being informative for a gene conversion event and ϕ , that is the parameter of the gene conversion tract length distribution. The method also detects the observed number of GCT and its length, which are inferior to the true number, especially for ψ values lower than 0.01 (BETRÁN *et al.* 1997). The ψ parameter is related to the extent of genetic differentiation among sequences. High genetic differentiation leads to high ψ values, and consequently to a better estimation of the number and length of GCT. In the present work the low ψ values obtained, with the exception of *Atpa* gene (Table 5) can be related to the low genetic differentiation between arrangements for most genes independently of their location relative to inversion breakpoints. In previous studies carried out with *D. subobscura*, ψ values obtained when comparing different arrangements ranged from 10^{-2} to 10^{-3} , and consequently the number of GCT detected was superior (ROZAS *et al.* 1999) to the present study. Similarly, studies carried out with other species (*D. buzzatii* and *D. pseudoobscura*) and comparing sequences from

different arrangements detected similar ψ values (LAAYOUNI *et al.* 2003; SCHAEFFER and ANDERSON 2005). In the present study, the *Atpa* gene presented the highest ψ values, which are in the same range than those obtained in the above mentioned studies, despite being located outside the inversion and the largest number of GCT. However, three of the four GCT detected in *Atpa* gene affected long fragments (1422 bp or higher), suggesting that this fragments could correspond to double crossovers. The individuals carrying the GCT for *Atpa* gene correspond to those detected as recombinants in the recombination network (Figure 5). The FBC48 chromosomal line was also detected as recombinant in the network in spite of no GCT tracts were detected. Visual inspection of the sequence showed that this particular chromosomal line presented five polymorphic sites not shared (supplementary material) and this higher variability can be misled to recombination. Then, due to the low genetic differentiation observed in our data, gene conversion could be underestimated. Thus, overlapped inversions, which are susceptible to present more genetic differentiation than single inversions, seem a better model to estimate gene conversion rate.

Patterns of selection

The six genes studied were selected from a list of candidate genes to be involved in thermal adaptation (LAAYOUNI *et al.* 2007). When grouping arrangements, significant pairwise comparisons for LD were obtained for genes separated by a high physical distance (16 Mb). These LD observations could be influenced by the maintenance of coadapted gene complexes (HOFFMANN and RIESEBERG 2008). However, no significant differences in the pattern obtained for Tajima's D and Fu and Li's D tests (with and without an outgroup) were found. In nearly all comparisons values were negative, suggesting an excess of polymorphisms at low frequency. In most of the cases values were not significant, with the exception of *Atpa* gene for the O_{3+4+1} arrangement. However, we can not discard that extreme values were assessed as significant as a consequence of the bias to negative values. Since the same pattern was observed for genes located inside or outside the inverted regions, the excess of rare alleles could not be necessarily related with the sweep associated to the release of a new inversion. In a previous study carried out with O_{3+4} and O_{ST} arrangements, this pattern was interpreted as an expansion of Mt. Parnes and Barcelona populations (PEGUEROLES *et al.* (b)).

In our study, several observations lead to the proposal that *Atpa* gene could be influenced by selection. The *Atpa* gene, which molecular function is sodium : potassium exchanging ATPase and cation transmembrane transporter, was the only of the six genes selected with previous evidences to respond to temperature stimulus (FENG *et al.* 1997). We observed that this gene, in spite of being located outside O_1 and O_7 inversions, presents high F_{ST} and LD values and low recombination between O_{3+4} - O_{3+4+1} and O_{3+4} - O_{3+4+7} arrangements, while this pattern is not observed in other genes also located outside the inverted regions. Moreover, when excluding recombinant individuals, fixed differences are found between arrangements, which corresponded to nonsynonymous changes. Thus, all these observations suggest the existence of a recent selective sweeps for the O_1 and O_7 inversions. According to the low variability detected for O_{3+4+1} and O_{3+4+7} arrangements the selective sweep can be dated in 0.04 and 0.06 Myr respectively, being younger than both arrangements (around 0.20 Myr). However, Tajima and Fu and Li's D were significant for O_{3+4+1} arrangement, but McDonald and Kreitman and HKA tests failed to find the footprint of selection, probably due to the lack of power that it is consequence of the low number of informative sites. Further analyses are required to further elucidate the presence of selection. The observation of linkage disequilibrium between non-overlapping inversions, as the O_{3+4+1} and O_{3+4+7} studied arrangements, could be consequence of the presence of epistatic interactions between them (KRIMBAS 1993; SPERLICH and FEUERBACH-MRAVLAG 1974). However, from the analysis of those pairwise comparisons showing significant LD do not showed the presence of epistatic interactions. Moreover, no significant differentiation was found for this gene when comparing the O_{3+4} and the O_{ST} arrangements (PEGUEROLES *et al.* (b)), thus suggesting that the selective value of the *Atpa* gene is related to O_1 and O_7 inversions.

Acknowledgments

First of all, we want to thank Ferran Palero for his comments, suggestions and support. Many thanks to Pedro A. Araúz for the collection of the flies and the collaboration carrying out the crosses, Joan Balanyà for the analyses of the polytene chromosomes and Gemma Calabria and Olga Dolgova for their collaboration in the *in situ* hybridizations. Thanks to Albert Ferrés for his help obtaining some sequences. Finally, we want to acknowledge the entire Aquadro lab for their friendship and valuable comments. This work was supported by a pre-doctoral fellowship to CP (2009FIC-

00096) from the Generalitat de Catalunya (Spain). Research was founded by projects CGL2006-13423-C02-02 from the Ministerio de Ciencia y Tecnología (MCYT, Spain) and 2009 SGR 636 from the Generalitat de Catalunya (Spain).

Literature

- ARAÚZ, P. A., F. MESTRES, C. PEGUEROLES, C. ARENAS, G. TZANNIDAKIS *et al.*, 2009 Tracking the origin of the American colonization by *Drosophila subobscura*: genetic comparison between Eastern and Western Mediterranean populations. *Journal of Zoological Systematics and Evolutionary Research* **47**: 25-34.
- BALANYÀ, J., L. SERRA, G. W. GILCHRIST, R. B. HUEY, M. PASCUAL *et al.*, 2003 Evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. *Evolution* **57**: 1837-1845.
- BALANYÀ, J., E. SOLÉ, J. OLLER, D. SPERLICH and L. SERRA, 2004 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. II. European populations. *Journal of Zoological Systematics and Evolutionary Research* **42**: 191-201.
- BETRÁN, E., J. ROZAS, A. NAVARRO and A. BARBADILLA, 1997 The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* **146**: 89-99.
- CASALS, F., and A. NAVARRO, 2007 Chromosomal evolution: inversions: the chicken or the egg? *Heredity* **99**: 479-480.
- DOBZHANSKY, T., 1950 Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35**: 288-302.
- FENG, Y., L. HUYNH, K. TAKEYASU and D. M. FAMBROUGH, 1997 The *Drosophila* Na,K-ATPase -subunit gene: gene structure, promoter function and analysis of a cold-sensitive recessive-lethal mutation. *Genes and Function* **1**: 99-117.
- FU, Y. X., and W. H. LI, 1993 Statistical tests of neutrality of mutations. *Genetics* **133**: 693-709.
- GÓMEZ-BALDÓ, L., A. LATORRE, L. SERRA and F. MESTRES, 2008 Molecular variation in the *Odh* gene in Chilean natural populations of *Drosophila subobscura*. *Hereditas* **145**: 154-162.

- HALL, T. A., 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98.
- HOFFMANN, A. A., C. M. SGRO and A. R. WEEKS, 2004 Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution* **19**: 482-488.
- HOFFMANN, A. A., and L. H. RIESEBERG, 2008 Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics* **39**: 21-42.
- HUDSON, R. R., M. KREITMAN and M. AGUADÉ, 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153-159.
- HUDSON, R. R., M. SLATKIN and W. P. MADDISON, 1992 Estimation of levels of gene flow from DNA-sequence data. *Genetics* **132**: 583-589.
- HUDSON, R. R., 2000 A new statistic for detecting genetic differentiation. *Genetics* **155**: 2011-2014.
- HUDSON, R. R., 2001 Two-locus sampling distributions and their application. *Genetics* **159**: 1805-1817.
- HUSON, D. H., and D. BRYANT, 2006 Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254-267.
- KELLY, J. K., 1997 A test of neutrality based on interlocus associations. *Genetics* **146**: 1197-1206.
- KIRKPATRICK, M., and N. BARTON, 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419-434.
- KOMAI, T., and T. TATAKU, 1940 Two independent inversions in the X-chromosome of *D. virilis* and their effects on crossing over and disjunction. *Cytologia* **11**: 245-260.
- KRIMBAS, C. B., and E. ZOUROS, 1969 Crossing over suppression between linked but non overlapping inversions in *Drosophila subobscura*. *Drosophila Information Service* **44**: 71-72.
- KRIMBAS, C. B., and M. LOUKAS, 1980 The inversion polymorphism of *Drosophila subobscura*. *Evolutionary Biology* **12**: 163-234.
- KRIMBAS, C. B., 1992 The inversion polymorphism of *Drosophila subobscura*, pp. 127-220 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Inc., Boca Raton, Florida.

- KRIMBAS, C. B., 1993 *Drosophila subobscura: Biology, Genetics and Inversion polymorphism*, Verlag Dr. Kovac, Hamburg.
- KUMAR, A., and J. P. GUPTA, 1991 Heterosis and the lack of coadaptation in *Drosophila nasuta*. *Heredity* **67**: 275-279.
- LAAYOUNI, H., E. HASSON, M. SANTOS and A. FONTDEVILA, 2003 The evolutionary history of *Drosophila buzzatii*. XXXV. Inversion polymorphism and nucleotide variability in different regions of the second chromosome. *Molecular Biology and Evolution* **20**: 931-944.
- LAAYOUNI, H., F. GARCIA-FRANCO, B. E. CHAVEZ-SANDOVAL, V. TROTTA, S. BELTRAN *et al.*, 2007 Thermal evolution of gene expression profiles in *Drosophila subobscura*. *BMC Evolutionary Biology* **7**: 42.
- LIBRADO, P., and J. ROZAS, 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.
- LOUKAS, M., C. B. KRIMBAS, P. MAVRAGANI-TSIPIDOU and C. D. KASTRITSIS, 1979 Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. *Journal of Heredity* **70**: 17-26.
- MCDONALD, J. H., and M. KREITMAN, 1991 Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652-654.
- MCVEAN, G., P. AWADALLA and P. FEARNHEAD, 2002 A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics* **160**: 1231-1241.
- MENOZZI, P., and C. B. KRIMBAS, 1992 The inversion polymorphism of *Drosophila subobscura* revisited: synthetic maps of gene arrangement frequencies and their interpretation. *Journal of Evolutionary Biology* **5**: 625-641.
- MUNTÉ, A., J. ROZAS, M. AGUADÉ and C. SEGARRA, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* **169**: 1573-1581.
- NAVARRO, A., E. BETRAN, A. BARBADILLA and A. RUIZ, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695-709.
- NAVARRO, A., A. BARBADILLA and A. RUIZ, 2000 Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in *Drosophila*. *Genetics* **155**: 685-698.

- NAVARRO-SABATÉ, A., M. AGUADÉ and C. SEGARRA, 1999 The relationship between allozyme and chromosomal polymorphism inferred from nucleotide variation at the *AcpH-1* gene region of *Drosophila subobscura*. *Genetics* **153**: 871-889.
- NEI, M., and T. GOJOBORI, 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**: 418-426.
- NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- NÓBREGA, C., M. KHADEM, M. AGUADÉ and C. SEGARRA, 2008 Genetic exchange versus genetic differentiation in a medium-sized inversion of *Drosophila*: the A_2/A_{ST} arrangements of *Drosophila subobscura*. *Molecular Biology and Evolution* **25**: 1534-1543.
- PEGUEROLAS, C., P. A. ARAÚZ, M. PASCUAL and F. MESTRES, 2010 A recombination survey using microsatellites: the O chromosome of *Drosophila subobscura*. Submitted.
- PEGUEROLAS, C., V. ORDÓÑEZ, M. PASCUAL and F. MESTRES, Recombination and selection in the maintenance of the adaptive value of inversions. In preparation (a).
- PEGUEROLAS, C., A. FERRÉS, C. F. AQUADRO, F. MESTRES and M. PASCUAL, Gene flow and gene flux shape evolutionary patterns of variation in *D. subobscura*. In preparation (b).
- PINA-MARTINS, F., and O. S. PAULO, 2008 CONCATENATOR: Sequence data matrices handling made easy. *Molecular Ecology Resources* **8**: 1254-1255.
- POWELL, J. R., 1997 Progress and prospects in evolutionary biology: The *Drosophila* model. Oxford University Press, New York.
- PREVOSTI, A., G. RIBÓ, L. SERRA, M. AGUADÉ, J. BALANYÀ *et al.*, 1988 Colonization of America by *Drosophila subobscura*: Experiment in natural populations that supports the adaptive role of chromosomal inversion polymorphism. *Proceedings of the National Academy of Sciences USA* **85**: 5597-5600.
- RAMOS-ONSINS, S. E., C. SEGARRA, J. ROZAS and M. AGUADE, 1998 Molecular and chromosomal phylogeny in the *obscura* group of *Drosophila* inferred from sequences of the *rp49* gene region. *Molecular Phylogenetics and Evolution* **9**: 33-41.

- ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. Proceedings of the National Academy of Sciences USA **91**: 11517-11521.
- ROZAS, J., C. SEGARRA, G. RIBÓ and M. AGUADÉ, 1999 Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *Drosophila subobscura*. Genetics **151**: 189-202.
- SCHAEFFER, S. W., M. P. GOETTING-MINESKY, M. KOVACEVIC, J. R. PEOPLES, J. L. GRAYBILL *et al.*, 2003 Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. Proceedings of the National Academy of Sciences USA **100**: 8319-8324.
- SCHAEFFER, S. W., and W. W. ANDERSON, 2005 Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura*. Genetics **171**: 1729-1739.
- SCHULTZ, J., and H. REDFIELD, 1951 Interchromosomal effects on crossing-over in *Drosophila*. Cold Spring Harbor Symposia on Quantitative Biology **16**: 175-197.
- SPERLICH, D., and H. FEUERBACH-MRAVLAG, 1974 Epistatic gene interaction, crossing over, and linked and unlinked inversions in *Drosophila subobscura* Evolution **28**: 67-75.
- STUMP, A. D., M. C. FITZPATRICK, N. F. LOBO, S. TRAORE, N. F. SAGNON *et al.*, 2005 Centromere-proximal differentiation and speciation in *Anopheles gambiae*. Proceedings of the National Academy of Sciences USA **102**: 15930-15935.
- STUMP, A. D., M. POMBI, L. GOEDEL, J. M. C. RIBEIRO, J. A. WILDER *et al.*, 2007 Genetic exchange in 2La inversion heterokaryotypes of *Anopheles gambiae*. Insect Molecular Biology **16**: 703-709.
- STURTEVANT, A. H., 1926 A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. Biologisches Zentralblatt **46**: 697-702.
- TAJIMA, F., 1989 Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics **123**: 585-595.
- THOMPSON, J. D., D. G. HIGGINS and T. J. GIBSON, 1994 CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research **22**: 4673-4680.

- WATTERSON, G. A., 1975 Number of segregating sites in genetic models without recombination. *Theoretical Population Biology* **7**: 256-276.
- WEIR, B. S., 1996 *Genetic data analysis II: Methods for discrete population genetic data*. Sinauer Associates, Sunderland, MA.
- WHITE, M. J. D., and F. H. W. MORLEY, 1955 Effects of pericentric rearrangements on recombination in grasshopper chromosomes. *Genetics* **40**: 604–619.

4. Discussió general

DISCUSSION

En la present tesi doctoral s'ha estudiat la recombinació al cromosoma O de *D. subobscura* a diferents nivells. Està generalment acceptat que les inversions redueixen la recombinació tant al seu interior com a les regions veïnes (STURTEVANT and BEADLE 1936), permetent el manteniment de combinacions genètiques adaptatives (DOBZHANSKY 1950; KIRKPATRICK and BARTON 2006). Degut a l'elevat flux gènic aquestes combinacions genètiques poden ser les mateixes per un mateix ordenament en diferents poblacions (ROZAS *et al.* 1995). A l'hora, els ordenaments poden diferir en el seu contingut genètic, tal i com s'ha vist en estudis previs (ROZAS *et al.* 1999; SCHAEFFER *et al.* 2003). Per tant, el flux gènic es podria trobar restringit a causa de les inversions.

4. 1. Mapa genètic del cromosoma O de *Drosophila subobscura*.

Un dels objectius de la present tesi doctoral era l'obtenció de dos mapes genètics del cromosoma O. Aquests mapes genètics es van elaborar amb 14 marcadors microsatèl·lits, i es van obtenir per dos tipus d'homocariotips: O_{ST} i O₃₊₄. Tots dos presenten una longitud genètica força semblant (O_{ST}: 184 cM i O₃₊₄: 196 cM), però malgrat això s'observen diferències entre els dos ordenaments. No sembla probable que aquestes diferències estiguin causades per factors externs, ja que els dos encreuaments es van realitzar simultàniament i van estar sotmesos a les mateixes condicions de temperatura i medi de cultiu. El temps de desenvolupament tampoc sembla un factor clau, ja que els individus adults genotipats es van agafar a l'atzar d'un conjunt d'individus que portaven una setmana emergint. La diferència en la mida mostral tampoc sembla que presenti una gran influència en l'obtenció d'estimes de recombinació, ja que a partir dels 150 individus analitzats les freqüències de recombinació entre marcadors no presenten grans oscil·lacions.

En canvi, hi ha altres factors que podrien haver modulats les taxes de recombinació obtingudes per cada ordenament. S'ha observat que la freqüència de recombinació d'un determinat cromosoma pot augmentar per la presència d'inversions en heterozigosi en altres cromosomes (SCHULTZ and REDFIELD 1951). Tot i que sols coneixiem el cariotip pel cromosoma O de les línies que vam usar, el protocol que

permet la seva obtenció tendeix a augmentar la homozigosis en tot el cariotip. També sabem que el polimorfisme per inversions s'espera que sigui major en poblacions del sud d'Europa que no pas en poblacions nòrdiques (SOLÉ *et al.* 2002; BALANYÀ *et al.* 2004). De fet, la heterozigositat mitjana en la població nòrdica de la qual provenien les línies O_{ST} era inferior a l'obtingut per la població del sud d'on provenien les O₃₊₄. Així, la probabilitat que els cromosomes de la resta del cariotip fossin heterozigòtics per inversions és major per l'ordenament O₃₊₄ que pel O_{ST}, fet que es correspondria amb la major longitud genètica obtinguda dels homocariotips O₃₊₄. De totes maneres, no es pot descartar que les diferències en la taxa de recombinació estiguin determinades pel contingut genètic de les diferents línies, efecte que s'ha descrit en *Drosophila* (BROOKS and MARKS 1986).

Si estimem la longitud del cromosoma O de *D. subobscura* com la mitjana dels valors obtinguts per cada ordenament, concloem que la longitud genètica d'aquest cromosoma, estimada amb la funció de mapa de Kosambi (KOSAMBI 1944), és d'aproximadament 190,7 cM. Aquesta estima és inferior a l'obtinguda en un mapa genètic realitzat anteriorment, on s'havia estimat en 228,3 cM (LOUKAS *et al.* 1979). Aquest mapa s'havia obtingut fent servir 16 marcadors dels que es desconeixia la seva localització citològica, i usant la funció de mapa de Haldane (HALDANE 1919). Aquesta funció, a diferència de la descrita per Kosambi, no té en compte el fenomen de la interferència, és a dir, que la formació d'un entrecreuament redueix la probabilitat de formació d'un altre entrecreuament a prop seu. Les estimes obtingudes per aquestes funcions de mapa es comencen a diferenciar quan les freqüències de recombinació entre marcadors són del 15% o superiors, donat que com més separats es trobin dos marcadors, més gran serà la probabilitat de formació de múltiples entrecreuaments i aquesta probabilitat estarà influenciada per la interferència (LYNCH and WALSH 1998). En ambdós mapes trobem marcadors separats per freqüències de recombinació superiors al 15%, i en conseqüència l'ús de la funció de mapa de Kosambi és molt recomanable. Per tant, no podem descartar que la longitud genètica del mapa de (LOUKAS *et al.* 1979) estigui sobreestimada a causa de la interferència.

Curiosament, en el nostre estudi es van detectar dobles entrecreuaments entre marcadors microsatèl·lits separats per una freqüència de recombinació inferior al 15%. Aquesta observació es podria explicar per mutació o per recombinació. Donat que en un

dels casos els al·lels de les dues línies parentals difereixen per un elevat nombre de còpies (per 11 repeticions), sembla molt poc probable que la mutació expliqui els dobles recombinants detectats. Per tant, només la recombinació dels marcadors, atribuïble a entrecreuaments o a conversió gènica permetria explicar aquesta observació.

Els dobles entrecreuaments formats entre dos marcadors consecutius no es poden detectar, per això és necessari corregir les freqüències de recombinació obtingudes experimentalment (que podrien estar subestimades) usant una funció de mapa. En *D. pseudoobscura* s'ha observat que en regions de 2 Mb de longitud no es formen dobles entrecreuaments (KULATHINAL *et al.* 2008). Donat que en el nostre estudi hi ha marcadors separats per una longitud superior, és important tenir una idea del número de dobles entrecreuaments que es podrien formar entre els marcadors. Per exemple, es pot quantificar el número de dobles entrecreuaments que es produeixen entre tres marcadors consecutius, i assumir que el número de dobles recombinants que es formaran entre dos marcadors separats per una distància física semblant serà aproximadament el mateix. D'aquesta manera es va poder estimar que el nombre aproximat d'individus amb dobles entrecreuaments entre els marcadors que es troben separats una distància física més gran en el nostre estudi (dsub51-14: 4510 Kb i dsub47-38: 5494 Kb) serien 1 i 6 respectivament, que representa respectivament el 3% i el 10% dels individus recombinants entre aquests marcadors. Per tant, la distància genètica entre marcadors separats més de 4,5 Mb estaria infravalorada si no es corretgeixen les freqüències de recombinació obtingudes amb la funció de mapa.

A la literatura trobem dos mapes genètics del cromosoma 2 de *D. pseudoobscura*, equivalent a l'element E de Muller, que és homòleg al cromosoma O de *D. subobscura* (HAMBLIN and AQUADRO 1999; ORTIZ-BARRIENTOS *et al.* 2006). En l'obtenció d'ambdós mapes es va fer servir la funció de mapa de Kosambi, però la distància genètica obtinguda pels dos estudis varia substancialment, presumiblement com a conseqüència del diferent nombre de marcadors usat (ORTIZ-BARRIENTOS *et al.* 2006). El mapa més antic, de 101 cM, es va estimar fent servir solament 10 marcadors (HAMBLIN and AQUADRO 1999), mentre que el posterior, de 204 cM, es va calcular amb 18 marcadors (ORTIZ-BARRIENTOS *et al.* 2006). També s'han estimat les longituds de l'element E de Muller en altres espècies del subgènere *Sophophora*: 56 cM el cromosoma 3R de *D. melanogaster*, 75 cM el de *D. simulans* i 125 cM el de *D.*

mauritiana (TRUE *et al.* 1995). Segons aquestes estimes, sembla que la longitud genètica d'aquest element a les dues espècies del grup *obscura* és superior al de les espècies del grup *melanogaster*. La disponibilitat d'alguns genomes seqüenciats, permet veure el número de Mb dels genomes del grup *melanogaster* és inferior al del grup *obscura* (27.9 Mb el cromosoma 3R de *D. melanogaster*, 27.5 Mb el de *D. simulans* i 30.7 Mb el de *D. pseudoobscura*). Malgrat això, la reducció no és proporcional a la distància genètica, que s'hauria de trobar també afectada per les diferents taxes de recombinació entre espècies. De fet, ja s'ha observat que la taxa de recombinació de *D. pseudoobscura* és superior a la d'altres taxons, sent quatre vegades superior que l'estimada per *D. melanogaster* (ORTIZ-BARRIENTOS *et al.* 2006). Comparant els mapes genètics obtinguts al nostre estudi i el de *D. pseudoobscura* d'Ortiz-Barrientos, s'observa que les taxes de recombinació obtingudes entre marcadors són molt similars, amb una mitjana ponderada de 7 cM/Mb en els dos casos.

4. 2. La recombinació al cromosoma O de *Drosophila subobscura*, a nivell general i regional.

Hi ha autors que defensen que la recombinació meiòtica podria estar regulada per tal d'assegurar la formació com a mínim d'un quiasma per bivalent, per facilitar la correcta segregació dels cromosomes homòlegs (NICKLAS 1974). En el nostre estudi s'ha observat que la freqüència de cromosomes recombinants és molt alta, ja que la descendència F_2 analitzada mostra que gairebé tots els individus són recombinants. La funció de Poisson descriu la freqüència de classes (enteses com el nombre d'entrecreuaments per cromosoma) assumint que els entrecreuaments es formen de manera independent els uns amb els altres. La mitjana ponderada de la distribució de Poisson és similar en els dos casos (O_{ST} : 1,61 i O_{3+4} : 1,77), indicant que la majoria dels cromosomes presentaven entre un i dos entrecreuaments. Malgrat aquestes semblances, en la descendència dels homocariotips O_{ST} , el número d'entrecreuaments per cromosoma s'ajusta a una distribució de Poisson, però en canvi no és així pels homocariotips O_{3+4} , a causa d'un excés de la classe dos entrecreuaments per cromosoma. En general, l'ordenament O_{3+4} presenta major freqüència de recombinació que l' O_{ST} .

És important tenir en compte que la taxa de recombinació al llarg del cromosoma es pot estudiar a nivell local i a nivell global. Del nostre estudi podem concloure que, la taxa de recombinació al cromosoma O de *D. subobscura* a nivell global és força uniforme, tal i com s'havia detectat en el cromosoma 2 de *D. pseudoobscura* (HAMBLIN and AQUADRO 1999) o en el 3R de *D. mauritiana* (TRUE *et al.* 1995). En canvi, en un estudi posterior del cromosoma 2 de *D. pseudoobscura* (ORTIZ-BARRIENTOS *et al.* 2006) es va observar que la taxa de recombinació és menor en les regions pericentromèriques, així com als cromosomes 3R de *D. melanogaster* i *D. simulans* (TRUE *et al.* 1995). Per tant, la uniformitat de la recombinació al llarg del cromosoma a nivell global és una característica que varia entre espècies. En canvi, a nivell regional hem observat que la taxa de recombinació no és homogènia, ja que es detecta la presència de *hot* i *coldspots* recombinacionals. En altres drosofílids s'han realitzat estudis de recombinació a nivell local (en regions d'unes poques Mb), on es mostra que la recombinació és altament heterogènia (CIRULLI *et al.* 2007; SINGH *et al.* 2009). A més, l'existència de *hotspots* recombinacionals s'ha descrit en un elevat nombre d'organismes (revisat a SINGH *et al.* (2009)). En resum, la taxa de recombinació en el cromosoma O de *D. subobscura* és força uniforme a nivell global, tot i que es detecta variació a nivell regional.

4. 3. Les inversions regulen la formació d'entrecreuaments al llarg del cromosoma.

La reducció de la recombinació per part de les inversions es pot produir per diferents motius. Per una banda, a l'interior de la nansa d'inversió que es forma en individus heterocariotípics, només els entrecreuaments d'ordre parell produiran gàmetes viables (STURTEVANT and BEADLE 1936; ROBERTS 1976; NAVARRO *et al.* 1997). A més, tal i com es sol observar en les preparacions de cromosomes politènics, la sinapsis no és perfecta a la vora dels punts de trencament de la nansa d'inversió (ROBERTS 1976). Finalment, en el cas que la recombinació trenqués combinacions beneficioses d'al·lels, la selecció podria actuar eliminant els cromosomes amb poca eficàcia biològica (KIRKPATRICK and BARTON 2006; HOFFMANN and RIESEBERG 2008). A *D. subobscura*, les inversions cromosòmiques són mecanismes reductors de la recombinació molt eficaços, donat que en heterocariotips s'ha observat que la recombinació es troba gairebé inhibida a l'interior de les inversions i en zones adjacents. En canvi, aquest

patró no és comú a tots els organismes; per exemple, en el dípter *Anopheles gambiae*, s'ha observat que hi ha una reducció significativa de la recombinació, però en cap cas aquesta està inhibida (STUMP *et al.* 2007).

Els heterocariotips O_{3+4+6}/O_{3+4} i O_{3+4+7}/O_{3+4} , que només presenten en heterozigosi una inversió per cromosoma (O_6 i O_7), ens van permetre comparar com es veu influenciada la recombinació per la mida i la localització de les inversions. La inversió O_6 (7,5 Mb) és més petita i presenta una posició central al cromosoma O, mentre que la inversió O_7 (11,8 Mb) es localitza més prop del centròmer (KUNZE-MÜHL and MÜLLER 1958). En els dos casos hi ha pràcticament inhibició de la recombinació al seu interior, donat que només s'ha detectat un doble recombinant en els aproximadament 200 individus analitzats per l'encreuament O_{3+4+7}/O_{3+4} . A més, la regió distal fora de la inversió amb la recombinació inhibida té una longitud similar en els dos casos, O_6 : 3,9 Mb i O_7 : 3,5 Mb. Per tant, o bé el fragment afectat fora de la inversió és independent de la longitud d'aquesta, o bé la reducció per part de la inversió O_6 és proporcionalment major en relació a la seva mida. Altres estudis mostren que les inversions mitjanes o petites situades en un regió mitjana-distal, com és el cas de la O_6 , podrien suprimir la recombinació de manera més efectiva (KRIMBAS and POWELL 1992). Per altra banda, la inversió O_6 permet observar que la reducció de la recombinació es desigual als dos extrems de la inversió, sent major a la part distal. El segment que presenta reducció de la recombinació fora de la inversió varia en funció de l'espècie. En algunes espècies és major el segment de la regió distal (DOBZHANSKY and EPLING 1948; WHITE and MORLEY 1955; ROBERTS 1962), mentre que en altres és superior a l'extrem proximal (CARSON 1953; STUMP *et al.* 2007).

S'ha descrit en diferents espècies d'insectes que la recombinació augmenta en els individus heterocariotípics en les regions del cromosoma més allunyades de la inversió (KOMAI and TATAKU 1940; GRELL 1962; LUCCHESI and SUZUKI 1968; KRIMBAS and POWELL 1992). Aquesta observació no sembla que afecti a *D. subobscura* perquè no s'han observat diferències significatives en la freqüència de recombinació a les zones no afectades per la inversió dels individus heterocariotípics i en la mateixa regió pels homocariotípics. També s'havia descrit un efecte intercromosòmic pel qual la recombinació en un cromosoma determinat pot augmentar per la presència d'inversions en heterozigosi a altres cromosomes (SCHULTZ and REDFIELD 1951), que podria ser

rellevant tenint en compte la diferent heterosigositat esperada en les poblacions d'origen de les nostres línies. Podem descartar que hi hagi diferències entre línies, tal i com s'havia observat anteriorment en *Drosophila* (BROOKS and MARKS 1986), ja que no s'han detectat canvis significatius en la freqüència de recombinació en les zones no afectades per la inversió, independentment de quina fos la línia usada.

Alguns estudis proposen que la formació d'un quiasma per bivalent seria necessari per assegurar una disjunció meiòtica correcta (NICKLAS 1974), en canvi altres defensen que no es necessària la formació del quiasma per a la correcta segregació (STURTEVANT and BEADLE 1936; HAWLEY *et al.* 1993). Dels aproximadament 200 individus amb l'heterocariotip O_{3+4+7}/O_{3+4+8} analitzats, només n'hi havia un portador d'un cromosoma recombinant. Per tant, la segregació de cromosomes sense la formació de quiasma sembla plausible. Per altra banda, no podem descartar que els cromosomes recombinants hagin segregat de manera preferent als cossos polars. Usant les inversions cromosòmiques com a marcadors genètics, en *D. subobscura* ja s'havia observat l'absència de cromosomes recombinants entre inversions no imbricades (KRIMBAS and ZOUROS 1969; SPERLICH and FEUERBACH-MRAVLAK 1974; MESTRES *et al.* 1998). En poblacions naturals, la formació d'ordenaments cromosòmics per recombinació entre inversions no imbricades, com ara la de l'ordenament O_7 a partir d'heterozigots O_{3+4+7}/O_{ST} , s'ha detectat en molt poques ocasions (GÓMEZ-BALDÓ *et al.* 2008; MESTRES *et al.* 2009). Per tant, aquesta inhibició de la recombinació entre inversions no imbricades sembla ser un fenomen freqüent.

4. 4. Patrons selectius relacionats amb inversions.

Els entrecreuaments dins de les inversions han de ser d'ordre parell per tal que els cromosomes recombinants siguin viables i no presentin duplicacions i delecions. En heterocariotips, sols es va detectar un doble recombinant a les inversions O_7 i O_8 , i cap per la inversió O_6 , que presenta una mida més petita que les anteriors. A partir dels encreuaments entre individus homocariotípics, es va observar que a les regions on es troba la inversió O_7 s'havien produït un 13.1 % d'entrecreuaments d'ordre parell del total d'individus analitzats, per tant la reducció en la formació de dobles entrecreuaments en els heterocariotips és molt significativa. Per una banda, aquest fet

podria ser degut a l'impediment físic associat a la formació de la nansa d'inversió (NAVARRO *et al.* 1997), tot i que els únics dobles entrecreuaments que s'han detectat estaven localitzats al voltant del punt de trencament, a la regió on sembla que l'impediment físic hauria de ser major. Per altra banda, els cromosomes recombinants podrien presentar menor eficàcia biològica o segregar de manera preferencial als corpuscles polars.

En la descendència dels heterocariotips O_{3+4+7}/O_{3+4+8} amb una línia homocariotípica O_{3+4} (amb els dos cromosomes homòlegs idèntics per descendència) provinent de Mt. Parnes, es va observar la segregació esperada del 50% per cada cromosoma. Donat que el cromosoma O_{3+4+7} provenia de Barcelona i el O_{3+4+8} de Mt. Parnes, sembla que el postulat de Th. Dobzhansky no es compliria (DOBZHANSKY 1950). En canvi, en els heterocariotips O_{3+4+6}/O_{3+4} i O_{3+4+7}/O_{3+4} , es va observar un dèficit significatiu d'homozigots O_{3+4} . Per obtenir la progènie recombinant, els heterocariotips O_{3+4+6}/O_{3+4} es van retroencreuar amb la línia O_{3+4} parental, mentre que en el O_{3+4+7}/O_{3+4} , es va usar una línia O_{3+4} diferent. Per tant, la reducció és independent del fet que els individus siguin o no consanguinis pel cromosoma O. Aquesta reducció d'individus homozigots afectava a cromosomes recombinants (on la recombinació tenia lloc a la regió del cromosoma més allunyada de la inversió) i no recombinants. La menor eficàcia biològica dels homocariotips O_{3+4} a causa d'al·lels recessius deleteris no sembla probable, ja que aquest ordenament es troba freqüentment en homozigosi en poblacions sudeuropees (KRIMBAS 1993; SOLE *et al.* 2002; ARAÚZ *et al.* 2009). A més, aquestes línies O_{3+4} es mantenen al laboratori sense que s'hagin detectat problemes de viabilitat. En un estudi previ, s'havien detectat desviacions significatives de la segregació a l'atzar en cromosomes no recombinants (MESTRES *et al.* 1998). S'ha suggerit que aquesta observació es podia explicar mitjançant tres hipòtesis: la major eficàcia biològica dels individus portadors d'una certa inversió, *meiotic drive* (durant la meiosi, segregació preferencial d'un cromosoma enfront del seu homòleg) o un efecte heteròtic del cromosoma del camp en combinació amb un cromosoma d'una línia de laboratori consanguínia. En el nostre cas, una major eficàcia de l'ordenament O_{3+4+6} no sembla una hipòtesi plausible, ja que aquest ordenament segurament és format recurrentment per recombinació però no incrementa en freqüència a les poblacions naturals estudiades (BALANYÀ *et al.* 2004). La hipòtesi del *meiotic drive* es va descartar en encreuaments recíprocs realitzats en *D. pseudoobscura*, on es va concloure que

l'excés d'heterozigots s'explicaria per selecció (MEISEL and SCHAEFFER 2007). Pel que fa a l'efecte heterotíc s'han formulat principalment dues hipòtesis per explicar el valor adaptatiu de les inversions: el model de coadaptació de (DOBZHANSKY 1950) i el model de l'adaptació local de (KIRKPATRICK and BARTON 2006). La hipòtesi de Dobzhansky explica que els heterozigots són adaptativament superiors degut a les relacions epistàtiques entre al·lels d'ordenaments cromosòmics de la mateixa localitat. En el nostre estudi s'ha observat un dèficit d'individus O_{3+4}/O_{3+4} independentment que fossin individus homozigòtics (tinguessin exactament el mateix contingut genètic) o només homocariotípics (tinguessin el mateix ordenament però diferent contingut genètic). Per tant, tot i que no podem afirmar la presència o absència de relacions epistàtiques entre gens, podem concloure que hi ha un efecte heterotíc de les inversions ja que els individus heterozigòtics presenten un major eficàcia biològica, provinguin o no de la mateixa població d'origen.

El destí d'una inversió nova està determinat principalment per la deriva genètica, atès que a l'inici la seva freqüència és molt baixa. En *D. subobscura* s'han descrit inversions que només s'han observat una vegada en poblacions naturals, confirmant aquestes hipòtesis (BALANYÀ *et al.* 2003). Però si les noves inversions són beneficioses es poden dispersar a les poblacions per mecanismes selectius, principalment mitjançant els dos mecanismes que s'han descrit a l'apartat anterior (DOBZHANSKY 1950; KIRKPATRICK and BARTON 2006). En *D. subobscura* s'ha observat recurrentment la formació d'inversions soles per recombinació entre inversions no imbricades (KRIMBAS 1993; BALANYÀ *et al.* 2003; MESTRES *et al.* 2009). El fet que aquestes inversions generalment desapareixen de les poblacions, podria indicar que presenten menor eficàcia biològica. Hi ha autors que consideren que la presència de relacions epistàtiques entre inversions és la responsable de l'elevat desequilibri de lligaments entre inversions no imbricades en *D. subobscura* (SPERLICH and FEUERBACH-MRAVLAK 1974). No obstant, altres autors suggerien que aquest desequilibri de lligament també podia estar produït per l'absència de recombinació en femelles heterocariotípiques (KRIMBAS and ZOUROS 1969; KRIMBAS 1993). Aquesta hipòtesi l'hem de descartar ja que em detectat un individu recombinant dels heterocariotips O_{3+4+7}/O_{3+4+8} analitzats i per tant la hipòtesi selectiva sembla més plausible.

4. 5. Estudi del contingut genètic de les inversions en diferents poblacions.

Les qüestions relacionades amb el flux genètic entre poblacions, es van abordar amb dues aproximacions. Per una banda, es van analitzar 6 gens nuclears (*Pif1A*, *Abi*, *Sqd*, *Yrt*, *Atpa* i *Fmr1*) en individus que presentaven els ordenaments O_{ST} i O_{3+4} i que provenien de poblacions molts distants: Barcelona (Espanya) i Mt. Parnes (Grècia). Cinc d'aquests gens es localitzen a la regió II i per tant fora de les inversions, i només un, el *Fmr1*, a la regió I i per tant l'interior de la inversió. Per l'altra banda, es van estudiar les seqüències nucleotídiques del gen *Odh* en individus que presentaven l'ordenament O_5 i que provenien de diverses poblacions (Montenegro, Suècia, Noruega, França i Bèlgica). Aquesta inversió és molt poc freqüent a les poblacions naturals (ZIVANOVIC and MESTRES 2000), i per això només disposem d'un individu seqüenciat per cadascuna de les poblacions estudiades.

La diferenciació genètica entre poblacions es va estimar de dues maneres, amb l'estadístic D_{xy} , que és una estima del número mitjà de substitucions nucleotídiques per lloc entre ordenaments (NEI 1987), i amb el F_{ST} , que és el número mitjà de diferències de les seqüències d'una població en relació al número mitjà entre les dues poblacions (HUDSON *et al.* 1992). En la caracterització del polimorfisme per inversions de les poblacions de Barcelona i Mt. Parnes utilitzades en aquesta tesi, es van observar diferències en les freqüències dels ordenaments de les dues poblacions (ARAÚZ *et al.* 2009). En canvi, a nivell genètic, no es van detectar diferències significatives entre poblacions per una mateixa inversió per cap dels sis gens estudiats, independentment de la seva localització relativa a la inversió. En estudis previs a *D. subobscura* on s'estudiava la diferenciació genètica entre poblacions del continent europeu, però usant mapes de restricció o marcadors microsatèl·lits, tampoc s'havia detectat diferenciació genètica significativa (ROZAS *et al.* 1995; PASCUAL *et al.* 2001). També s'ha observat un flux genètic important entre poblacions utilitzant els gens letals com a marcadors tant a poblacions paleàrtiques com colonitzadores (ZIVANOVIC *et al.* 2007; MESTRES *et al.* 2009). De la mateixa manera, tampoc s'havia observat diferenciació genètica significativa entre poblacions Nord-americanes de *D. pseudoobscura* (SCHAEFFER *et al.* 2003). Aquests resultats suggereixen que el flux genètic entre poblacions distants pot ser elevat. En estudis realitzats a *D. melanogaster*, usant marcadors microsatèl·lits no

necessàriament relacionats amb inversions, sí que es va observar diferenciació significativa entre poblacions africanes, d'on és originària aquesta espècie (CARACRISTI and SCHLÖTTERER 2003; POOL and AQUADRO 2006). Per tant, la diferenciació entre poblacions podria dependre del seu aïllament que estaria afectat per la seva capacitat dispersiva (SERRA *et al.* 1987) o per la migració facilitada pel transport passiu (PASCUAL *et al.* 2007). En el cas del gen *Odh* de *D. subobscura* només s'han observat dos haplotips en totes les poblacions analitzades. L'observació del mateix haplotip entre poblacions europees molt llunyanes suggereix que l'ordenament té un origen recent i que en aquesta espècie el flux gènic entre poblacions és elevat. Per últim, donat que la diferenciació entre poblacions no es detecta per gens localitzats a l'interior de les inversions, la teoria de la coadaptació entesa tal i com la va formular Th. Dobzhansky, que implica la diferenciació del contingut genètic de les inversions entre diferents poblacions (DOBZHANSKY 1950), sembla que no s'aplica en *D. subobscura* com demostren els nostres estudis ni en *D. pseudoobscura* (SCHAEFFER *et al.* 2003).

Per altra banda, els tests de neutralitat realitzats per cadascun dels gens en cada població van ser en general negatius i no significatius. Aquest excés d'al·lels a baixa freqüència per tots els gens, independentment de la seva localització en relació a la inversió, i els resultats obtinguts amb l'estadístic R_2 , suggereixen que ambdues poblacions es trobarien en expansió (RAMOS-ONSINS and ROZAS 2002). A més, els valors de R_2 mostren que l'expansió de Mt. Parnes podria ser major que la de Barcelona, que es podria relacionar amb la major grandària efectiva estimada per aquesta població (ARAÚZ *et al.* 2009).

4. 6. Estudi del flux gènic entre ordenaments.

Està àmpliament acceptat que cada inversió cromosòmica s'ha produït per un esdeveniment únic (POWELL 1997). Per tant, durant les primeres etapes la variabilitat genètica a dintre de les inversions seria molt baixa i també hi hauria elevada diferenciació genètica i desequilibri de lligament entre ordenaments. En el cas que el contingut genètic de les inversions fos adaptatiu, la selecció podria mantenir en menor o major grau la diferenciació genètica entre ordenaments en funció de la pressió selectiva (HOFFMANN *et al.* 2004). Tot i això, i sobretot si no actua la selecció, s'espera que a

mesura que passi el temps es vagi recuperant la variabilitat, principalment per recombinació (entesa com entrecreuaments i conversió gènica) i, en un menor grau, per mutació neutra (SCHAEFFER and ANDERSON 2005). A *D. subobscura* s'ha determinat que la taxa de conversió genètica per bp per generació ($2,73 \cdot 10^{-7}$) és un ordre de magnitud més gran que la de recombinació ($7,06 \cdot 10^{-8}$, estimada en aquest treball) i dos ordres de magnitud més gran que la de mutació ($2,16 \cdot 10^{-9}$) (BETRÁN *et al.* 1997). La recombinació també dependrà de característiques intrínseques de la inversió, com ara la seva longitud. Un estudi teòric va estimar que la longitud genètica mínima necessària perquè es produeixin dobles entrecreuaments a l'interior d'una inversió és de 20 cM (NAVARRO *et al.* 1997). A partir d'aquí, com més gran sigui la inversió, més elevada serà la probabilitat que es produeixin dobles entrecreuaments. Considerant que la longitud del cromosoma O de *D. subobscura* es de 190,7 cM (el present treball), la longitud aproximada de les inversions O_1 i O_7 seria major que el mínim teòric requerit per la formació de dobles entrecreuaments (37,7 cM i 72,9 cM respectivament). Per altra banda, l'edat que hem estimat pels dos ordenaments és força similar (aproximadament 0,20 Ma). En general, entre els dos ordenaments no s'observen diferències significatives per la diversitat nucleotídica silenciosa. Per tant, pot ser que la longitud de les inversions sigui un factor clau limitador del flux gènic quan aquestes són petites (al volant de 20 cM o longituds menors), però un cop superat el llindar de la mida, i amb el pas del temps, no s'observin diferències en la variabilitat.

En tots es estudis amb gens nuclears que s'han realitzat al segment SI, s'ha observat que la diferenciació genètica entre els ordenaments O_{ST} i O_{3+4} és elevada i significativa (ROZAS and AGUADÉ 1994; NAVARRO-SABATÉ *et al.* 1999; MUNTÉ *et al.* 2005 i també en el present treball). Els alts nivells de diferenciació genètica detectats entre aquests ordenaments poden estar influenciats per la inhibició de la recombinació degut al complex que formen les dues inversions imbricades. Els anàlisis que s'han realitzat en gens nuclears localitzats a l'interior d'inversions senzilles de mida mitjana, també mostren diferenciació genètica significativa (ROZAS *et al.* 1999; NÓBREGA *et al.* 2008 i en el present estudi). En canvi, els valors de diferenciació genètica són menors en inversions senzilles que en inversions imbricades (NAVARRO-SABATÉ *et al.* 1999; ROZAS *et al.* 1999; MUNTÉ *et al.* 2005 i en el present treball). Els estudis realitzats en inversions imbricades al mateix cromosoma en *D. pseudoobscura* també mostren alts

nivells de diferenciació, comparables als obtinguts en *D. subobscura* (SCHAEFFER *et al.* 2003).

Per altra banda, al cromosoma O de *D. subobscura* la majoria d'inversions cromosòmiques es localitzen en zones d'alta densitat de punts de trencament d'inversions (KRIMBAS 1993), i per tant a les poblacions naturals podem trobar individus heterocariotípics per inversions solapades però localitzades en diferents cromosomes. En aquestes sistemes s'esperaria trobar una elevada diferenciació genètica, però entre O_{3+4+8} - O_{3+4+23} (ROZAS *et al.* 1999) i O_{3+4+1} - O_{3+4+7} (el present treball) s'observa diferenciació genètica significativa però menor que l'obtinguda al comparar O_{3+4} i O_{ST} . A més, la moderada diferenciació entre O_{3+4+1} i O_{3+4+7} pel gens situats dintre de les inversions és sorprenent donat que aquests ordenaments presenten diferents distribucions i rarament coincidiran en una mateixa població (KRIMBAS 1993; ARAÚZ *et al.* 2009). La manca d'alts nivells de diferenciació genètica es podria explicar per la recombinació amb altres ordenaments amb els quals només difereixin per una inversió. Per exemple, la majoria de les inversions descrites a *D. subobscura* formen complexes amb l'ordenament O_{3+4} , que és el més freqüent en poblacions del sud d'Europa (KRIMBAS 1993; SOLÉ *et al.* 2002; ARAÚZ *et al.* 2009). Per això, a les poblacions naturals s'espera trobar un elevat nombre d'individus que seran heterozigots per O_{3+4} i un altre ordenament (com ara O_{3+4+8} , O_{3+4+23} , O_{3+4+1} i O_{3+4+7}), de manera que els dos cromosomes dels individus sols difereixin per una inversió i poden presentar major freqüència de recombinació. En canvi, el flux gènic entre O_{ST} i altres ordenaments s'espera que es trobi molt reduït, donat que la seva presència en individus heterozigòtics implicarà la formació de dues nanses d'inversió, amb l'excepció de quan l'heterozigot es formi amb inversions senzilles com ara O_6 , O_5 i O_{11} , que es solen trobar a baixa freqüència i restringides al nord d'Europa, o bé amb nous ordenaments formats per recombinació, que també es troben a molt baixa freqüència.

Els alts nivells constants de desequilibri de lligament observats a l'interior d'inversions imbricades van ser interpretats per manca de formació de dobles entrecreuaments, o bé per una forta pressió selectiva (MUNTÉ *et al.* 2005; SCHAEFFER and ANDERSON 2005). Per la majoria de gens del nostre estudi, encara que es trobin localitzats a l'interior de la inversió o a la vora del punt de trencament, s'observa que el desequilibri de lligament entre seqüències d'un tipus d'ordenament és major que quan es barregen seqüències d'ordenaments diferents. Aquesta observació, sumada als baixos

nivells de diferenciació detectats, suggereixen elevada recombinació. Per això hem obtingut estimes directes de la recombinació, usant el mètode de Hudson (HUDSON 2001), implementat al programa LDhat (MCVEAN *et al.* 2002). Amb aquest mètode s'obtenen estimes globals de recombinació, ja que no distingeix entre entrecreuaments o conversió gènica. La recombinació que es detecta principalment és deguda a l'intercanvi entre ordenaments diferents, tot i que la recombinació entre homocariotips pot comportar la dispersió de variants noves que s'han originat per mutació. En el nostre estudi hem observat que la recombinació de l'ordenament O_{3+4} és major que la dels O_{3+4+1} i O_{3+4+7} , que pot ser conseqüència de la major edat i mida efectiva per l'ordenament O_{3+4} . A més, la recombinació estimada per l'ordenament O_{3+4} a Mt. Parnes és superior a la de Barcelona, i podria tenir relació amb la major abundància d'aquest ordenament a la població de Mt. Parnes (KRIMBAS 1993; ARAÚZ *et al.* 2009). Per altra banda, sembla que la recombinació entre ordenaments és menor per alguns gens situats a l'interior de la inversió i també pel gen *Atpa*, malgrat que es troba situat a fora de la regió invertida. Per últim, també es van realitzar xarxes de recombinació per cadascun dels gens amb el programa SplitsTree (HUSON and BRYANT 2006). El resultat va ser que gairebé per tots els gens els nivells de recombinació eren molt elevats, amb l'excepció pel gen *Atpa*, on s'observaven dos grups clarament diferents a cada població, tant a Barcelona (les seqüències amb l'ordenament O_{3+4+7} i O_{3+4}) com a Mt. Parnes (les seqüències amb l'ordenament O_{3+4+1} i O_{3+4}).

Per altra banda, també es van realitzar estimes de fragments de conversió gènica (GCT, acrònim de l'anglès *gene conversions tracts*) mitjançant el mètode de BETRÁN *et al.* (1997). Aquest es basa en dos paràmetres, ϕ , que és un paràmetre de la distribució de la longitud dels fragments de conversió gènica i ψ , que és probabilitat d'una posició de ser informativa per detectar la conversió gènica. Una major diferenciació genètica comporta l'obtenció de valors de ψ més elevats i en conseqüència una millor estima dels GCT. En general, els valors obtinguts del paràmetre ψ han estat baixos al nostre estudi en comparació amb estimes prèvies obtingudes pel sistema O_{3+4}/O_{ST} (BETRÁN *et al.* 1997), o en altres espècies com *D. buzzatti* (LAAYOUNI *et al.* 2003) o *D. pseudobscura* (SCHAEFFER and ANDERSON 2005). Per tant, com a conseqüència de la baixa diferenciació genètica detectada en els gens estudiats (tot i que en la majoria de casos era significativa), amb l'excepció del gen *Atpa*, les estimes de conversió gènica a partir de les nostres seqüències estarien subestimades. En el gen *Atpa*, on les estimes de

conversió gènica no haurien d'estar subestimades per manca de diferenciació, només s'han detectat quatre fragments, i tres d'ells estarien produïts per entrecreuaments ja que la seva longitud és superior a 1400bp.

4. 7. Variabilitat genètica vs distància al punt de trencament.

La recuperació de la variabilitat a l'interior de les inversions sembla que no és uniforme. Estudis teòrics van preveure que la variabilitat nucleotídica hauria de ser menor a la vora dels punts de trencament que en les parts més centrals de les inversions (NAVARRO *et al.* 1997). Posteriorment, estudis experimentals van mostrar que la variabilitat nucleotídica és menor en marcadors separats per una distància d'aproximadament 1 Mb o inferior del punt de trencament (ANDOLFATTO *et al.* 2001). A més, s'ha de tenir en compte si els gens es troben dintre o fora de les inversions, ja que és a l'interior de les inversions on la formació d'entrecreuaments presentarà més restriccions. En els nostres estudis, el gen *Sqd* es situa dintre de la inversió i a 0,66 Mb de la inversió O_{3+4+7} , el gen *Odh* es troba dintre de la inversió i a 0,74 Mb de la inversió O_5 , el *Fmr1* es troba dintre de la inversió i a 1,3 Mb del punt de trencament de la inversió O_{3+4} , l'*Yrt* es troba fora de la inversió i a 0,08 Mb de la inversió O_{3+4+1} , l'*Atpa* es localitza fora de la inversió i a 0,82 Mb de la inversió O_{3+4+1} , i finalment. De tots aquests gens, només s'observa l'esperada disminució de la variabilitat en dos casos: al gen *Atpa*, tot i que en aquest cas es podria deure a un procés selectiu més que a la seva localització respecte el punt de trencament de la inversió, i al gen *Odh*, on només s'han identificat 2 haplotips en totes les línies analitzades fins al moment, però pot ser conseqüència de l'origen recent de la inversió O_5 . Els estudis realitzats en *D. subobscura* mostren que en general no hi ha una disminució de la variabilitat associada a la proximitat dels punts trencament. Per exemple, en set gens situats a menys de 1Mb dels punts de trencament de les inversions O_{ST} i O_{3+4} no es va observar disminució de la variabilitat, i fins i tot el gen *Acph1*, que està situats a menys de 0,2 Mb, presentava major variabilitat que la resta (MUNTÉ *et al.* 2005). Els alts valors de variabilitat i divergència observats en els gens *Acph1* (MUNTÉ *et al.* 2005) i *Yrt* (el present treball) suggereixen que podrien presentar una taxa de mutació anormalment alta. En un altre estudi en *D. subobscura*, però en aquest cas entre els ordenaments A_2 i A_{ST} tampoc es va observar reducció en la variabilitat per l'únic gen localitzat a menys d'una Mb del

punt de trencament (NÓBREGA *et al.* 2008). En canvi, en altres espècies com *D. buzzatti*, sí que s'observa la reducció esperada (LAAYOUNI *et al.* 2003). Per tant, les dades experimentals no mostren sempre aquesta reducció en la variabilitat, que sembla estar fortament influenciada per l'espècie. Factors com el flux gènic, la mutació i l'edat de les inversions poden influenciar en gran part els resultats obtinguts.

4. 8. Patrons de selecció

Diverses observacions porten a la conclusió que les inversions de *D. subobscura* són adaptatives, però potser la més important és l'existència de clines latitudinals del mateix sentit per la freqüència de les inversions, en poblacions europees i colonitzadores (PREVOSTI *et al.* 1988; BALANYÀ *et al.* 2003). En concret els ordenaments O_{ST} i O_{3+4} , que són els que s'han estudiat més exhaustivament en *D. subobscura*, presenten distribucions oposades; el primer és més freqüent al nord d'Europa i presenta una correlació positiva amb la latitud, mentre que l'ordenament O_{3+4} presenta una correlació negativa (KRIMBAS 1993; MESTRES *et al.* 1994; SOLÉ *et al.* 2002; BALANYÀ *et al.* 2004). Cap dels 9 gens nuclears analitzats a l'interior de la inversió O_{3+4} sembla que estigui directament influenciat per la selecció (MUNTÉ *et al.* 2005 i també el present treball). Per tant, la detecció de patrons selectius no sembla fàcil ni tan sols en sistemes que es troben altament diferenciats. Els sis gens nuclears utilitzats en el present estudi van estar triats d'un llistat de gens candidats a estar involucrats en la adaptació termal (LAAYOUNI *et al.* 2007). Tot i això, dels quatre ordenaments estudiats (O_{3+4} , O_{3+4+1} , O_{3+4+7} i O_{ST}) només es van trobar desviacions significatives del model neutralista pel gen *Atpa* a l'ordenament O_{3+4+1} .

El gen *Atpa* codifica per la subunitat α de la proteïna ATPasa en *Drosophila*, i les seves funcions moleculars són les de bomba Na-K i transport de cations transmembrana, sent una molècula indispensable pel correcte funcionament del metabolisme cel·lular. A la literatura trobem informació de la seva relació amb respostes a estímuls relacionats amb la temperatura (FENG *et al.* 1997), ja que a *Drosophila* el fenotip descrit sensible a la temperatura està relacionat amb mutacions causades als canals de ions o bé en proteïnes que regulen la funció dels canals. Diverses observacions indiquen que el gen *Atpa* podria estar influenciat per la selecció

relacionada amb els ordenaments O_{3+4+1} i O_{3+4+7} . Aquest gen es troba entremig de dues inversions no solapes ($O_1 - O_7$ i O_{3+4}) però malgrat això, als ordenaments O_{3+4+1} i O_{3+4+7} s'ha detectat baixa variabilitat (accentuada en l'ordenament O_{3+4+1}), alta diferenciació genètica, elevat desequilibri de lligament i menys recombinació entre seqüències de diferents ordenaments que entre seqüències del mateix ordenament, malgrat que s'han detectat clarament individus recombinants. Quan s'exclouen aquests individus recombinants, s'observa que les diferències fixades entre ordenaments donen lloc a canvis aminoacídics. En canvi, no s'observen diferències quan es comparen els ordenaments O_{ST} i O_{3+4} i per tant, el procés selectiu només estaria relacionat amb els ordenaments O_{3+4+1} i O_{3+4+7} . La baixa variabilitat sinònima observada pel gen *Atpa* ens permet datar "l'escombrat selectiu" en aproximadament 0,05 Ma. Per tant, aquest és més recent que l'edat dels ordenaments O_{3+4+1} i O_{3+4+7} , estimada en aproximadament 0,20 Ma. Malgrat aquestes observacions, el test de McDonald i Kreitman no refuta la hipòtesi neutralista, tot i que podria ser conseqüència de les poques posicions polimòrfiques de les que es disposa per realitzar el test. L'elevat desequilibri de lligament entre inversions no imbricades podria explicar-se per la presència d'interaccions epistàtiques entre inversions (SPERLICH and FEUERBACH-MRAVLJAG 1974; KRIMBAS 1993).

4. 9. Datació dels ordenaments cromosòmics.

L'edat de les inversions es pot estimar considerant que cadascuna d'elles s'ha format per un esdeveniment únic (POWELL 1997). L'estima de l'edat es realitza inferint el temps de coalescència de les seqüències, donat que la variabilitat acumulada per mutació neutra és proporcional al temps que fa que es va originar la inversió. Per aquesta raó s'han d'excloure els individus recombinants de l'anàlisi, ja que el seu ús comportaria la sobreestima de l'edat de les inversions. L'edat es pot estimar usant la diversitat nucleotídica silenciosa mitjana de tots els individus excepte dels recombinants (ROZAS *et al.* 1999), o bé usant la diversitat nucleotídica silenciosa entre els dos individus més divergents, que comportarà l'obtenció d'estimes més antigues (ROZAS and AGUADÉ 1994). En el nostre estudi hem observat que l'ús de les seqüències més divergents presenta més oscil·lacions, probablement donat que no tota la variabilitat acumulada ha aparegut per mutació sinó també per recombinació.

Per altra banda, es van obtenir seqüències de dues espècies per usar-les com a grup extern, concretament de *D. madeirensis* i *D. pseudoobscura*. Tot i que les dues espècies pertanyen al grup *obscura*, el temps de divergència amb *D. subobscura* es força diferent. *D. subobscura* i *D. madeirensis* van divergir fa aproximadament 0,63 Ma, mentre que amb *D. pseudoobscura* van divergir abans, fa aproximadament 8 Ma (RAMOS-ONSINS *et al.* 1998). Tot i que *D. madeirensis* és una espècie endèmica de l'illa de Madeira amb una grandària efectiva poblacional petita, els seus nivells de diversitat nucleotídica són similars als de *D. subobscura*, i no s'observa la disminució esperada en illes (KHADEM *et al.* 2001). Els mateixos autors observen que *D. subobscura* i *D. madeirensis* presenten polimorfismes compartits heretats de l'ancestre comú. Usant *D. madeirensis* com a grup extern vam observar que els valors de diversitat nucleotídica i divergència eren del mateix ordre de magnitud i que, per alguns gens, les edats dels ordenaments eren superiors al temps de divergència entre *D. subobscura* i *D. madeirensis*. Per totes aquestes raons, finalment varem escollir *D. pseudoobscura* com a grup extern de *D. subobscura*.

Dels estudis dels gens localitzats a l'interior de les inversions O_{3+4} , usant la diversitat nucleotídica silenciosa mitjana de tots els individus excepte dels recombinants, concloem que l'edat dels ordenaments O_{ST} i O_{3+4} és força similar, sent respectivament 0,49 i 0,44 Ma. En estudis realitzats anteriorment usant aquesta metodologia, les datacions van donar valors semblants entre ordenaments però inferiors als del nostre treball, sent 0,24 Ma per O_{ST} i 0,33 Ma per O_{3+4} (ROZAS *et al.* 1999), i 0,26 Ma per O_{ST} i 0,31 Ma per O_{3+4} (NAVARRO-SABATÉ *et al.* 1999). L'edat dels ordenaments O_{3+4+1} i O_{3+4+7} també s'ha estimat amb la diversitat nucleotídica silenciosa mitjana de tots els individus excepte dels recombinants, amb la informació del gen que es troba a l'interior de la inversió i més a la vora del punt de trencament (*Sqd*). Es va evitar l'ús de gens localitzats en zones centrals, perquè en aquestes zones la probabilitat de recombinar és major (NAVARRO *et al.* 1997) i per tant l'origen de les inversions que s'obtidria estaria sobreestimat. Les estimes són semblant als dos ordenaments, on l'edat del O_{3+4+1} és de 0,23 Ma, i la del O_{3+4+7} és de 0,21 Ma. Per tant, tal i com s'esperava, l'edat d'aquests ordenaments derivats és inferior a la del O_{3+4} .

5. CONCLUSIONS

1. S'han obtingut dos mapes genètics del cromosoma O: la longitud genètica del cromosoma O_{ST} és de 184 cM, mentre que és de 196 cM en el cromosoma O_{3+4} .
2. L'anàlisi d'un mínim de 150 individus sembla suficient per obtenir bones estimes de les freqüències de recombinació entre marcadors separats fins a 5,5 Mb.
3. La taxa de recombinació en el cromosoma O de *D. subobscura* és força uniforme a nivell global. Malgrat això, es detecten variacions a nivell regional, amb la presència de *hot* i *coldspots* recombinacionals.
4. Les inversions redueixen fortament la recombinació en els heterocariotips. En l'anàlisi de la descendència d'individus homocariotípics s'observa que el 85,7 % de cromosomes són recombinants. En canvi, en heterocariotips per una inversió de mida mitjana només el 50% de la descendència és recombinant i només per marcadors localitzats fora de la inversió. Finalment, en heterocariotips per dues inversions de mida mitjana sols el 0,5% és recombinant.
5. La reducció de la recombinació per part de les inversions s'observa a l'interior d'aquestes i també fora però vora del punt de trencament. A fora de les inversions, la reducció és asimètrica, sent major a la part distal.
6. L'excés d'individus heterocariotípics a la descendència dels encreuaments es podria explicar per un efecte heteròtic degut al seu avantatge selectiu. Aquesta diferència s'observa independentment de la població d'origen dels ordenaments.
7. Per als 6 gens estudiats al llarg del cromosoma O no s'han detectat diferències significatives entre poblacions pel mateix ordenament. Per tant, el flux genètic entre poblacions distants de *D. subobscura* sembla ser molt elevat.
8. S'ha trobat diferenciació genètica significativa entre les inversions O_{3+4} i O_{ST} deguda al gen *Fmr1*, que es troba localitzat en la regió invertida. A més, aquesta diferenciació és independent de la població d'origen de les inversions.

9. Els ordenaments que difereixen per inversions senzilles de mida mitjana (O_{3+4+1} - O_{3+4} i O_{3+4+7} - O_{3+4}) presenten menor diferenciació genètica que els que difereixen per ordenaments complexos amb inversions imbricades en un mateix cromosoma (O_{3+4} - O_{ST}).

10. Les estimes de recombinació suggereixen que hi ha flux gènic entre inversions, i a més no s'observa una reducció de la variabilitat a les proximitats del punt de trencament de les inversions.

11. El gen *Atpa*, que es troba situat fora de les inversions, sembla que presenta un escombrat selectiu recent en els ordenaments O_{3+4+1} i O_{3+4+7} , datat respectivament en 0,04 i 0,06 Ma.

12. Les estimes de conversió gènica obtingudes per tots els gens estudiats excepte *Atpa* segurament estan subestimades a causa de la baixa diferenciació gènica entre ordenaments. En el gen *Atpa* s'han detectat 4 fragments, dels quals el 75% són fragments superiors de 1400 bp, probablement originats per entrecreuaments.

13. L'edat dels ordenaments O_{3+4+1} i O_{3+4+7} és força semblant (aproximadament 0,20 Ma) i inferior a la dels ordenaments O_{ST} i O_{3+4} (respectivament 0,48 i 0,41 Ma). La baixa variabilitat detectada a l'inversió O_5 suggereix un origen més recent.

5. CONCLUSIONS

1. Two genetic maps of the O chromosome were obtained: the genetic length obtained for O_{ST} chromosome was 184 cM, while it was 196 cM in O₃₊₄.
2. The analysis of 150 individuals seems sufficient to calculate accurate recombination frequencies for markers separated by less than 5.5 Mb.
3. Overall, the recombination rate across the O chromosome was fairly uniform. However, at the local level recombination was heterogeneous and some hot and coldspots regions were detected.
4. Inversions highly reduce recombination in heterokaryotypes. In homokaryotypic individuals 85.7% of their offspring was recombinant. However, for heterokaryotypic individuals carrying a medium-size inversion only 50% of recombinants were detected, and affecting only markers located outside the inversion. Finally, for heterokaryotypic individuals carrying two medium-size inversions only 0.5% of recombinant offspring was observed.
5. Recombination reduction mediated by inversions was detected inside as well as outside them but close to the inversion breakpoints. Reduction outside inversions was asymmetrical and more severe in the distal region.
6. The excess of heterokaryotypic individuals in the offspring could be explained by a heterotic effect. This result was attained independently of the population of origin of the different arrangements.
7. For the 6 genes studied across the O chromosome, no significant genetic differentiation was detected within arrangement between populations. Thus, gene flow between distant populations of *D. subobscura* could be high.
8. Significant genetic differentiation associated to *Fmr1* gene was detected between inversions O₃₊₄ and O_{ST}, which was independent of the population of origin.

9. Genetic differentiation between arrangements differing by medium-size inversion ($O_{\underline{3+4+1}}-O_{\underline{3+4}}$ and $O_{\underline{3+4+7}}-O_{\underline{3+4}}$) is lower than between arrangements differing by overlapped inversions on the same chromosome ($O_{\underline{3+4}}-O_{ST}$).

10. Recombination estimates revealed that gene flux between inversions could take place. Moreover, no reduction of the nucleotide variability near the inversion breakpoints was detected.

11. The *Atpa* gene, despite being located outside inversions, presents a recent selective sweep in arrangements $O_{\underline{3+4+1}}$ and $O_{\underline{3+4+7}}$, which was dated at 0.04 and 0.06 My respectively.

12. Gene conversion estimates obtained for all genes except *Atpa* may have been underestimated due to the low genetic differentiation between arrangements. For the *Atpa* gene four tracts were detected and 75% of them were larger than 1400 bp, suggesting that the recombinant chromosomes probably originated by crossovers.

13. The age of arrangements $O_{\underline{3+4+1}}$ and $O_{\underline{3+4+7}}$ was quite similar (approximately 0.20 My) and is inferior to those estimated for arrangements O_{ST} and $O_{\underline{3+4}}$ (0.48 and 0.41 My respectively). The low variability observed in O_5 inversion suggests a more recent origin.

6. BIBLIOGRAFIA

- ANDOLFATTO, P., F. DEPAULIS and A. NAVARRO, 2001 Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetics Research* **77**: 1-8.
- ARADOTTIR, G. I., and R. B. ANGUS, 2004 A chromosomal analysis of some water beetle species recently transferred from *Agabus* Leach to *Ilybius* Erichson, with particular reference to the variation in chromosome number shown by *I. montanus* Stephens (Coleoptera: Dytiscidae). *Hereditas* **140**: 185-192.
- ARAÚZ, P. A., F. MESTRES, F. PERIS-BONDIA, A. LATORRE and L. SERRA, 2008 Molecular analysis of a colonization: *Odh* gene variability in Mediterranean and American populations of *D. subobscura*. Annual meeting of the Society of Molecular Biology and Evolution. Barcelona, Spain.
- ARAÚZ, P. A., F. MESTRES, C. PEGUEROLES, C. ARENAS, G. TZANNIDAKIS, KRIMBAS, C. B. and L. SERRA, 2009 Tracking the origin of the American colonization by *Drosophila subobscura*: genetic comparison between Eastern and Western Mediterranean populations. *Journal of Zoological Systematics and Evolutionary Research* **47**: 25-34.
- ARGEMÍ, M., F. MESTRES, A. PREVOSTI and L. SERRA, 2003 Microevolutionary dynamics of a community of Drosophilids. *Journal of Zoological Systematics and Evolutionary Research* **41**: 57-63.
- AYALA, F. J., and M. COLUZZI, 2005 Chromosome speciation: Humans, *Drosophila*, and mosquitoes. *Proceedings of the National Academy of Sciences USA* **102**: 6535-6542.
- AYALA, F. J., L. SERRA and A. PREVOSTI, 1989 A grand experiment in evolution: the *Drosophila subobscura* colonization of the Americas. *Genome* **31**: 246-255.
- BALANYÀ, J., L. SERRA, G. W. GILCHRIST, R. B. HUEY, M. PASCUAL, F. MESTRES and E. SOLÉ, 2003 Evolutionary pace of chromosomal polymorphism in colonizing populations of *Drosophila subobscura*: an evolutionary time series. *Evolution* **57**: 1837-1845.
- BALANYÀ, J., E. SOLÉ, J. OLLER, D. SPERLICH and L. SERRA, 2004 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. II.

- European populations. *Journal of Zoological Systematics and Evolutionary Research* **42**: 191-201.
- BALANYÀ, J., J. M. OLLER, R. B. HUEY, G. W. GILCHRIST and L. SERRA, 2006 Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* **313**: 1773-1775.
- BEADLE, G. W., and A. H. STURTEVANT, 1935 X-chromosome inversions and meiosis in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences USA* **21**: 384-390.
- BECKENBACH, A. T., and A. PREVOSTI, 1986 Colonization of North America by the European species, *Drosophila subobscura* and *D. ambigua*. *American Midland Naturalist* **115**: 10-18.
- BEGON, M., C. B. KRIMBAS and M. LOUKAS, 1980 The genetics of *Drosophila subobscura* populations. XV. Effective size of a natural population estimated by three independent methods. *Heredity* **45**: 335-350.
- BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature* **356**: 519-520.
- BETRÁN, E., J. ROZAS, A. NAVARRO and A. BARBADILLA, 1997 The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* **146**: 89-99.
- BIRD, M. J., 1947 The genetics and cytology of *Drosophila subobscura*. PhD Thesis, University College, London.
- BRNCIC, D., and M. BUDNIK, 1980 Colonization of *Drosophila subobscura* Collin in Chile. *Drosophila Information Service* **55**: 20.
- BROOKS, L. D., and R. W. MARKS, 1986 The organization of genetic variation for recombination in *Drosophila melanogaster*. *Genetics* **114**: 525-547.
- BURLA, H., H. JUNGEN and G. BACHLI, 1986 Population structure of *Drosophila subobscura*: non-random microdispersion of inversion polymorphism on a mountain slope. *Genetica* **70**: 9-15.
- CÁCERES, M., J. M. RANZ, A. BARBADILLA, M. LONG and A. RUIZ, 1999 Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**: 415-418.

- CARACRISTI, G., and C. SCHLÖTTERER, 2003 Genetic differentiation between American and European *Drosophila melanogaster* populations could be attributed to admixture of African alleles. *Molecular Biology and Evolution* **20**: 792-799.
- CARSON, H. L., 1946 The selective elimination of inversion dicentric chromatids during meiosis in the egg of *Sciara impatiens*. *Genetics* **31**: 95-113.
- CARSON, H. L., 1953 The effect of inversions on crossing over in *Drosophila robusta*. *Genetics* **38**: 168-186.
- CARSON, H. L., 1958 Response to selection under different conditions of recombination in *Drosophila*. *Cold Spring Harbor Symposia on Quantitative Biology* **23**: 291-306.
- CASALS, F., and A. NAVARRO, 2007 Chromosomal evolution: inversions: the chicken or the egg? *Heredity* **99**: 479-480.
- CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289-1303.
- CHOVNICK, A., 1973 Gene conversion and transfer of genetic information within the inverted regions of inversion heterozygotes. *Genetics* **75**: 123-131.
- CHRISTIE, A. L. M., 1939 The effect of X-rays on sex in *Drosophila subobscura* and an account of some sex-linked characters. *Journal of Genetics* **39**: 47-60.
- CIRULLI, E. T., R. M. KLIMAN and M. A. F. NOOR, 2007 Fine-scale crossover rate heterogeneity in *Drosophila pseudoobscura*. *Journal of Molecular Evolution* **64**: 129-135.
- CODINA, M., and M. M. PEREZ, 1980 Primeros datos sobre la distribución estacional de las especies de *Drosophila* en Cataluña. *Boletín de la Real Sociedad Española de Historia Natural (Biol.)* **78**: 143-154.
- COMERON, J. M., and M. KREITMAN, 2002 Population, evolutionary and genomic consequences of interference selection. *Genetics* **161**: 389-410.
- COOP, G., X. Q. WEN, C. OBER, J. K. PRITCHARD and M. PRZEWORSKI, 2008 High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science* **319**: 1395-1398.
- DOBZHANSKY, Th., 1950 Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics* **35**: 288-302.

- DOBZHANSKY, Th., and C. EPLING, 1948 The suppression of crossing-over in inversion heterozygotes of *Drosophila pseudoobscura*. Proceedings of the National Academy of Sciences USA **34**: 137-141.
- FEARNHEAD, P., and N. G. C. SMITH, 2005 A novel method with improved power to detect recombination hotspots from polymorphism data reveals multiple hotspots in human genes. American Journal of Human Genetics **77**: 781-794.
- FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. Genetics **78**: 737-756.
- FENG, Y., L. HUYNH, K. TAKEYASU and D. M. FAMBROUGH, 1997 The *Drosophila* Na,K-ATPase -subunit gene: gene structure, promoter function and analysis of a cold-sensitive recessive-lethal mutation. Genes and Function **1**: 99-117.
- FISHER, R. A., 1930 *The genetical theory of natural selection*. Oxford University Press, Oxford, UK.
- FONTDEVILA, A., C. ZAPATA, G. ALVAREZ, L. SANCHEZ, J. MENDEZ and I. ENRIQUEZ 1983 Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. Genetics **105**: 935-955.
- GÓMEZ-BALDÓ, L., A. LATORRE, L. SERRA and F. MESTRES, 2008 Molecular variation in the *Odh* gene in Chilean natural populations of *Drosophila subobscura*. Hereditas **145**: 154-162.
- GRELL, R. F., 1962 A new model for secondary nondisjunction: the role of distributive pairing. Genetics **47**: 1737-1754.
- HALDANE, J. B. S., 1919 The combination of linkage values and the calculation of distances between the loci of linked factors. Journal of Genetics **8**: 299-309.
- HAMBLIN, M. T., and C. F. AQUADRO, 1999 DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: a study of the second chromosome. Genetics **153**: 859-869.
- HAWLEY, R. S., K. S. MCKIM and T. ARBEL, 1993 Meiotic segregation in *Drosophila melanogaster* females: molecules, mechanisms, and myths. Annual Review of Genetics **27**: 281-317.
- HEWITT, G. M., 1999 Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society **68**: 87-112.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. Genetical Research **8**: 269-294.

- HOFFMANN, A. A., and L. H. RIESEBERG, 2008 Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology, Evolution, and Systematics* **39**: 21-42.
- HOFFMANN, A. A., C. M. SGRO and A. R. WEEKS, 2004 Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution* **19**: 482-488.
- HUDSON, R. R., 2001 Two-locus sampling distributions and their application. *Genetics* **159**: 1805-1817.
- HUDSON, R. R., M. SLATKIN and W. P. MADDISON, 1992 Estimation of levels of gene flow from DNA-sequence data. *Genetics* **132**: 583-589.
- HUSON, D. H., and D. BRYANT, 2006 Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254-267.
- KHADEM, M., J. ROZAS, C. SEGARRA and M. AGUADÉ, 2001 DNA variation at the *rp49* gene region in *Drosophila madeirensis* and *Drosophila subobscura* from Madeira: inferences about the origin of an insular endemic species. *Journal of Evolutionary Biology* **14**: 379-387.
- KIRKPATRICK, M., and N. BARTON, 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**: 419-434.
- KOMAI, T., and T. TATAKU, 1940 Two independent inversions in the X-chromosome of *D. virilis* and their effects on crossing over and disjunction. *Cytologia* **11**: 245-260.
- KOMAI, T., and T. TAKAKU, 1942 On the effect of the X-chromosome inversions on crossing-over in *Drosophila virilis*. *Cytologia* **12**: 357-365.
- KOSAMBI, D. D., 1944 The estimation of map distances from recombination values. *Annual of Eugenics* **12**: 172-175.
- KOSKE, T., and J. MAYNARD SMITH, 1954 Genetics and cytology of *Drosophila subobscura*. X. The fifth linkage group. *Journal of Genetics* **52**: 521-541.
- KRIMBAS, C. B., 1992 The inversion polymorphism of *Drosophila subobscura*, pp. 127-220 in *Drosophila Inversion Polymorphism*, edited by C. B. KRIMBAS and J. R. POWELL. CRC Press, Inc., Boca Raton, Florida.
- KRIMBAS, C. B., 1993 *Drosophila subobscura: Biology, Genetics and Inversion polymorphism*. Verlag Dr. Kovac, Hamburg.
- KRIMBAS, C. B., and M. LOUKAS, 1980 The inversion polymorphism of *Drosophila subobscura*. *Evolutionary Biology* **12**: 163-234.

- KRIMBAS, C. B., and J. F. POWELL, 1992 *Drosophila Inversion Polymorphism*. CRC Press, Boca Raton, Florida.
- KRIMBAS, C. B., and E. ZOUROS, 1969 Crossing over suppression between linked but non overlapping inversions in *Drosophila subobscura*. *Drosophila Information Service* **44**: 71-72.
- KULATHINAL, R. J., S. M. BENNETT, C. L. FITZPATRICK and M. A. F. NOOR, 2008 Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. *Proceedings of the National Academy of Sciences USA* **105**: 10051-10056.
- KUNZE-MÜHL, E., and E. MÜLLER, 1958 Weitere Untersuchungen über die chromosomale Struktur und die natürlichen Strukturtypen von *Drosophila subobscura* Coll. *Chromosoma* **9**: 559-570.
- LAAYOUNI, H., E. HASSON, M. SANTOS and A. FONTDEVILA, 2003 The evolutionary history of *Drosophila buzzatii*. XXXV. Inversion polymorphism and nucleotide variability in different regions of the second chromosome. *Molecular Biology and Evolution* **20**: 931-944.
- LAAYOUNI, H., F. GARCIA-FRANCO, B. E. CHAVEZ-SANDOVAL, V. TROTTA, S. BELTRAN, M. COROMINAS and M. SANTOS 2007 Thermal evolution of gene expression profiles in *Drosophila subobscura*. *BMC Evolutionary Biology* **7**: 42.
- LANKINEN, P., and W. PINSKER, 1977 Allozyme constitution of two standard strains of *Drosophila subobscura*. *Experientia* **33**: 1301-1302.
- LINDSLEY, D. L., and G. G. ZIMM, 1992 *The genome of Drosophila melanogaster*. Academic Press, San Diego, CA.
- LOUKAS, M., C. B. KRIMBAS, P. MAVRAGANI-TSIPIDOU and C. D. KASTRITSIS, 1979 Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. *Journal of Heredity* **70**: 17-26.
- LUCCHESI, J. C., and D. T. SUZUKI, 1968 The interchromosomal control of recombination. *Annual Review of Genetics* **2**: 53-86.
- LYNCH, M., and B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc., Sunderland, MA.
- MAINX, F., T. KOSKE and E. SMITAL, 1953 Untersuchungen über die chromosomale Struktur europische Vertreter der *Drosophila obscura* Gruppe. *Zeitschrift für induktive Abstammungs- und Vererbungslehre* **85**: 354-372.

- MCVEAN, G., P. AWADALLA and P. FEARNHEAD, 2002 A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics* **160**: 1231-1241.
- MEISEL, R. P., and S. W. SCHAEFFER, 2007 Meiotic transmission of *Drosophila pseudoobscura* chromosomal arrangements. *PLoS One* **2**: e530.
- MENOZZI, P., and C. B. KRIMBAS, 1992 The inversion polymorphism of *Drosophila subobscura* revisited: synthetic maps of gene arrangement frequencies and their interpretation. *Journal of Evolutionary Biology* **5**: 625-641.
- MESTRES, F., and L. SERRA, 2008 The Va/Ba balanced lethal strain: thirty years of research in *Drosophila subobscura* (1977–2007). *Journal of Zoological Systematics and Evolutionary Research* **46**: 190-191.
- MESTRES, F., J. SANZ and L. SERRA, 1998 Chromosomal structure and recombination between inversions in *Drosophila subobscura*. *Hereditas* **128**: 105-113.
- MESTRES, F., G. PEGUEROLES, A. PREVOSTI and L. SERRA, 1990 Colonization of America by *Drosophila subobscura*: lethal genes and the problem of the O₅ inversion. *Evolution* **44**: 1823-1836.
- MESTRES, F., L. ABAD, B. SABATER-MUNOZ, A. LATORRE and L. SERRA, 2004 Colonization of America by *Drosophila subobscura*: association between *Odh* gene haplotypes, lethal genes and chromosomal arrangements. *Genes and Genetic Systems* **79**: 233-244.
- MESTRES, F., J. BALANYÀ, C. ARENAS, E. SOLE and L. SERRA, 2001 Colonization of America by *Drosophila subobscura*: heterotic effect of chromosomal arrangements revealed by the persistence of lethal genes. *Proceedings of the National Academy of Sciences USA* **98**: 9167-9170.
- MESTRES, F., J. BALANYÀ, C. SEGARRA, A. PREVOSTI and L. SERRA, 1994 O chromosome inversion polymorphism in Northern and Atlantic Europe and its implications in the American colonization by *Drosophila subobscura*. *Zeitschrift für Zoologische Systematik und Evolutionsforschung* **32**: 108-116.
- MESTRES, F., J. BALANYÀ, M. PASCUAL, C. ARENAS, E. SOLÉ, and L. SERRA, 2005 Lethal genes and the colonization of America by *D. subobscura*. *Topics in Current Genetics* **1**: 31-55.
- MESTRES, F., J. BALANYÀ, M. PASCUAL, C. ARENAS, G. W. GILCHRIST, R. B. HUEY and L. SERRA, 2009 Evolution of Chilean colonizing populations of *Drosophila subobscura*: lethal genes and chromosomal arrangements. *Genetica* **136**: 37-48.

- MONCLÚS, M., 1964 Distribución y ecología de Drosophilidos en España. I. Especies de *Drosophila* en la región catalana. *Genética Ibérica* **16**: 143-165.
- MORGAN, T. H., 1913 *Heredity and sex*. Columbia University Press, New York.
- MORTON, N. E., 1955 Sequential tests for the detection of linkage. *The American Journal of Human Genetics* **7**: 277-318.
- MULLER, H. J., 1932 Some genetic aspects of sex. *American Naturalist* **66**: 118-138.
- MULLER, H. J., 1964 The relation of recombination to mutational advance. *Mutation Research* **1**: 2-9.
- MUNTÉ, A., M. AGUADÉ and C. SEGARRA, 2000 Nucleotide variation at the *yellow* gene region is not reduced in *Drosophila subobscura*: A study in relation to chromosomal polymorphism. *Molecular Biology and Evolution* **17**: 1942-1955.
- MUNTÉ, A., J. ROZAS, M. AGUADÉ and C. SEGARRA, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. *Genetics* **169**: 1573-1581.
- NAVARRO, A., A. BARBADILLA and A. RUIZ, 2000 Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in *Drosophila*. *Genetics* **155**: 685-698.
- NAVARRO, A., E. BETRAN, A. BARBADILLA and A. RUIZ, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**: 695-709.
- NAVARRO-SABATÉ, A., M. AGUADÉ and C. SEGARRA, 1999 The relationship between allozyme and chromosomal polymorphism inferred from nucleotide variation at the *Acp-1* gene region of *Drosophila subobscura*. *Genetics* **153**: 871-889.
- NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- NICKLAS, R. B., 1974 Chromosome segregation mechanisms. *Genetics* **78**: 205-213.
- NÓBREGA, C., M. KHADEM, M. AGUADÉ and C. SEGARRA, 2008 Genetic exchange versus genetic differentiation in a medium-sized inversion of *Drosophila*: the A_2/A_{ST} arrangements of *Drosophila subobscura*. *Molecular Biology and Evolution* **25**: 1534-1543.
- NOOR, M. A. F., K. L. GRAMS, L. A. BERTUCCI and J. REILAND, 2001 Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences USA* **98**: 12084-12088.

- ORENGO, D. J., and A. PREVOSTI, 1994 Preadult competition between *Drosophila subobscura* and *Drosophila pseudoobscura*. *Zeitschrift für Zoologische Systematik und Evolutionsforschung* **32**: 44-50.
- ORTIZ-BARRIENTOS, D., A. S. CHANG and M. A. F. NOOR, 2006 A recombinational portrait of the *Drosophila pseudoobscura* genome. *Genetics Research* **87**: 23-31.
- PASCUAL, M., F. J. AYALA, A. PREVOSTI and L. SERRA, 1993 Colonization of North America by *Drosophila subobscura*: Ecological analysis of three communities of drosophilids in California. *Zeitschrift für Zoologische Systematik und Evolutionsforschung* **31**: 216-226.
- PASCUAL, M., M. D. SCHUG and C. F. AQUADRO, 2000 High density of long dinucleotide microsatellites in *Drosophila subobscura*. *Molecular Biology and Evolution* **17**: 1259-1267.
- PASCUAL, M., C. F. AQUADRO, V. SOTO and L. SERRA, 2001 Microsatellite variation in colonizing and palearctic populations of *Drosophila subobscura*. *Molecular Biology and Evolution* **18**: 731-740.
- PASCUAL, M., M. P. CHAPUIS, F. MESTRES, J. BALANYÀ, R. B. HUEY, G. W. GILCHRIST, L. SERRA and A. ESTOUP, 2007 Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology* **16**: 3069-3083.
- PHILIP, U., J. M. RENDEL, H. SPURWAY and J. B. S. HALDANE, 1944 Genetics and karyology of *Drosophila subobscura*. *Nature* **154**: 260-262.
- POOL, J. E., and C. F. AQUADRO, 2006 History and structure of Sub-Saharan populations of *Drosophila melanogaster*. *Genetics* **174**: 915-929.
- POWELL, J. R., 1997 *Progress and prospects in Evolutionary Biology: The Drosophila model*. Oxford University Press, New York.
- PREVOSTI, A., G. RIBÓ, L. SERRA, M. AGUADÉ, J. BALANYÀ, M. MONCLÚS AND F. MESTRES, 1988 Colonization of America by *Drosophila subobscura*: Experiment in natural populations that supports the adaptive role of chromosomal inversion polymorphism. *Proceedings of the National Academy of Sciences USA* **85**: 5597-5600.
- PTAK, S. E., D. A. HINDS, K. KOEHLER, B. NICKEL, N. PATIL, D. G. BALLINGER, M. PRZEWORSKI, K. A. FRAZER AND S. PAABO, 2005 Fine-scale recombination patterns differ between chimpanzees and humans. *Nature Genetics* **37**: 429-434.

- RAMOS-ONSINS, S. E., and J. ROZAS, 2002 Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**: 2092-2100.
- RAMOS-ONSINS, S. E., C. SEGARRA, J. ROZAS and M. AGUADE, 1998 Molecular and chromosomal phylogeny in the *obscura* group of *Drosophila* inferred from sequences of the *rp49* gene region. *Molecular Phylogenetics and Evolution* **9**: 33-41.
- RIESEBERG, L. H., 2001 Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* **16**: 351-358.
- ROBERTS, P. A., 1962 Interchromosomal effects and the relation between crossing over and nondisjunction. *Genetics* **47**: 1691-1709.
- ROBERTS, P. A., 1965 Difference in behaviour of eu- and heterochromatin: crossing over. *Nature* **205**: 725-726.
- ROBERTS, P. A., 1976 The genetics of chromosome aberration, pp. 67-184 in *The Genetics and Biology of Drosophila*. Academic Press, London.
- RODRIGUEZ-TRELLES, F., 2003 Seasonal cycles of allozyme by chromosomal inversion gametic disequilibrium in *Drosophila subobscura*. *Evolution* **57**: 839-848.
- RODRIGUEZ-TRELLES, F., G. ALVAREZ and C. ZAPATA, 1996 Time series analysis of seasonal changes of the O inversion polymorphism of *Drosophila subobscura*. *Genetics* **142**: 179-187.
- ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. *Proceedings of the National Academy of Sciences USA* **91**: 11517-11521.
- ROZAS, J., C. SEGARRA, G. RIBÓ and M. AGUADÉ, 1999 Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *Drosophila subobscura*. *Genetics* **151**: 189-202.
- ROZAS, J., C. SEGARRA, C. ZAPATA, G. ALVAREZ and M. AGUADE, 1995 Nucleotide polymorphism at the *rp49* region of *Drosophila subobscura*: Lack of geographic subdivision within chromosomal arrangements in Europe. *Journal of Evolutionary Biology* **8**: 355-367.
- SANTOS, J., L. SERRA, E. SOLÉ and M. PASCUAL, 2010 FISH mapping of microsatellite loci from *Drosophila subobscura* and its comparison to related species *Chromosomal Research*, in press.
- SANTOS, M., 2009 Recombination load in a chromosomal inversion polymorphism of *Drosophila subobscura*. *Genetics* **181**: 803-809.

- SCHAEFFER, S. W., and W. W. ANDERSON, 2005 Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura*. *Genetics* **171**: 1729-1739.
- SCHAEFFER, S. W., M. J. BERNHARDT and W. W. ANDERSON, 2003 Evolutionary rearrangement of the amylase genomic regions between *Drosophila melanogaster* and *Drosophila pseudoobscura*. *Journal of Heredity* **94**: 464-471.
- SCHULTZ, J., and H. REDFIELD, 1951 Interchromosomal effects on crossing-over in *Drosophila*. *Cold Spring Harbor Symposia on Quantitative Biology* **16**: 175-197.
- SERRA, L., G. PEGUEROLES and F. MESTRES, 1987 Capacity of dispersal of a colonizing species: *Drosophila subobscura*. *Genetica* **73**: 223-235.
- SINGH, N. D., C. F. AQUADRO and A. G. CLARK, 2009 Estimation of fine-scale recombination intensity variation in the *white-echinus* interval of *D. melanogaster*. *Journal of Molecular Evolution* **69**: 42-53.
- SOLÉ, E., J. BALANYÀ, D. SPERLICH and L. SERRA, 2002 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. I. Mediterranean populations from southwestern Europe. *Evolution* **56**: 830-835.
- SPERLICH, D., and H. FEUERBACH-MRAVLAVAG, 1974 Epistatic gene interaction, crossing over, and linked and unlinked inversions in *Drosophila subobscura*. *Evolution* **28**: 67-75.
- SPERLICH, D., and P. PFRIEM, 1986 Chromosomal polymorphism in natural and experimental populations, pp. 257-309 in *The Genetics and Biology of Drosophila* (eds. Ashburner, M., Carson, H. L. and Thompson Jr., J. N.). Academic Press, N. Y.
- SPERLICH, D., H. FEUERBACH-MRAVLAVAG, P. LANGE, A. MICHAELIDIS and A. PENTZOS-DAPONTE, 1977 Genetic load and viability distribution in central and marginal populations of *Drosophila subobscura*. *Genetics* **86**: 835-848.
- SPURWAY, H., 1945 The genetics and cytology of *Drosophila subobscura*. I. Element A. Sex-linked mutants and their standard order. *Journal of Genetics* **46**: 278-286.
- STAM, P., 1978 Interference in genetic crossing over and chromosome mapping *Genetics* **92**: 573-594.
- STEINBERG, A. G., 1936 The effect of autosomal inversions on crossing over in the X-chromosome of *D. melanogaster*. *Genetics* **21**: 615-624.

- STRICKBERGER, M. W., 1976 *Genética, 2ª edición*. Omega, Barcelona.
- STUMM-ZOLLINGER, E., 1953 Vergleichende Untersuchung über die Inversionshäufigkeit bei *Drosophila subobscura* in Populationen der Schweiz und Südwesteuropas. *Zeitschrift für induktive Abstammungs- und Vererbungslehre* **85**: 382-407.
- STUMP, A. D., M. POMBI, L. GOEDDEL, J. M. C. RIBEIRO, J. A. WILDER, A. D. TORRE AND N. J. BESANSKY, 2007 Genetic exchange in 2La inversion heterokaryotypes of *Anopheles gambiae*. *Insect Molecular Biology* **16**: 703-709.
- STURTEVANT, A. H., 1913 The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology* **14**: 43-59.
- STURTEVANT, A. H., 1921 A case of rearrangement of genes in *Drosophila*. *Proceedings of the National Academy of Sciences USA* **7**: 235-237.
- STURTEVANT, A. H., 1926 A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biologisches Zentralblatt* **46**: 697-702.
- STURTEVANT, A. H., and G. W. BEADLE, 1936 The relations of inversions in the X chromosomes of *Drosophila melanogaster* to crossing-over and disjunction. *Genetics* **28**: 554-604.
- STURTEVANT, A. H., and C. R. PLUNKETT, 1926 Sequence of corresponding third-chromosome genes in *Drosophila melanogaster* and *D. simulans*. *Biology Bulletin* **50**: 56-60.
- TABERLET, P., L. FUMAGALLI, A. G. WUST-SAUCY and J. F. COSSON, 1998 Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**: 453-464.
- TADIN-STRAPPS, M., D. WARBURTON, F. A. M. BAUMEISTER, S. G. FISCHER, J. YONAN, T. C. GILLIAM AND A. M. CHRISTIANO, 2004 Cloning of the breakpoints of a *de novo* inversion of chromosome 8, inv (8)(p11.2q23.1) in a patient with Ambras syndrome. *Cytogenetic and Genome Research* **107**: 68-76.
- TRUE, J. R., J. M. MERCER and C. C. LAURIE, 1995 Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* **142**: 507-523.
- WASSERMAN, M., 1968 Recombination-induced chromosomal heterosis *Genetics* **58**: 125-139

- WHITE, M. J. D., and F. H. W. MORLEY, 1955 Effects of pericentric rearrangements on recombination in grasshopper chromosomes. *Genetics* **40**: 604–619.
- ZIVANOVIC, G., and F. MESTRES, 2000 Lethal genes in O₅ chromosomes of *Drosophila subobscura* from Europe and America. *Journal of Zoological Systematics and Evolutionary Research* **38**: 123-126.
- ZIVANOVIC, G., C. ARENAS and F. MESTRES, 2007 The genetic structure of Balkan populations of *Drosophila subobscura*. *Hereditas* **144**: 120-128.



Facultat de Biologia

Departament de Genètica

Av. Diagonal, 645
Edifici Annex
08028 – BARCELONA
Tel. 93 402 14 98
93 403 49 89
Fax 93 403 44 20

INFORME DELS DIRECTORS DE LA TESI

Publicació 1. Respecte al primer article, Cinta Pegueroles va dur a terme la majoria del treball experimental (obtenir les línies cromosòmiques, genotipar-les, fer els encreuaments, analitzar els resultats, etc.). Va realitzar els càlculs necessaris i participà activament en la discussió. El primer esborrany del treball el va redactar ella mateixa. Aquest article està actualment en segona revisió a la revista GENETICA (IF = 1,98).

Publicació 2. En el segon article, la doctorada també va ser la persona que va portar el pes de la part experimental. Va realitzar els càlculs estadístics i va participar activament en la discussió del resultat. Ella va redactar el primer manuscrit de l'article. Aquest article està enviat a la revista JOURNAL OF EVOLUTIONARY BIOLOGY (IF=3,92).

Publicació 3. En el tercer article, Cinta Pegueroles es va encarregar de posar a punt el patró de seqüenciació dels gens escollits i va col·laborar activament en la obtenció de les línies cromosòmiques de Barcelona i Mt. Parnes. Posteriorment, ella es va encarregar de la seqüenciació de les línies i de l'anàlisi dels resultats. Va tenir una gran participació en la discussió dels resultat i preparà un primer esborrany de l'article. Aquest article es pensa enviar a la revista GENETICS (IF = 4.00).

Publicació 4. Respecte al quart article, cronològicament fora el seu primer treball de tesi, entroncada amb la línia de recerca principal del grup. Ella va obtenir les línies cromosòmiques O₅ de Suecia i Noruega, va participar en la seva seqüenciació, en l'anàlisi dels resultats i en la discussió del treball. Aquest article està publicat a la revista HEREDITAS (IF = 1,18).

Publicació 5. En el darrer article, Cinta Pegueroles va col·laborar en l'obtenció de les línies cromosòmiques i les va seqüenciar pels gens escollits. Amb les dades obtingudes realitzà tots els càlculs necessaris i va participar activament en la discussió dels seus resultats. També va ser responsable de redactar el primer esborrany del treball. Aquest article es pretén enviar a la revista MOLECULAR BIOLOGY AND EVOLUTION (IF= 7,28).

Signat:

Marta Pascual Berniola

Francesc Mestres Naval

Barcelona, 8 de Març de 2010

Supplementary table publication 1: O₃₊₄ and O_{ST} homokaryotypic chromosomal lines genotyped for the present study. Lines highlighted in grey are those selected for the crosses, since they differed for most of the markers.

	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub	dsub
O₃₊₄	26	02	25	51	14	47	38	34	71	01	73	04	29	12
MP13	459	223	278	393	116	137	417	113	263	273	295	206	259	269
MP16	486	221	266	405	120	150	416	107	273	271	301	185	253	273
FMP16	454	221	272	387	120	142	409	111	270	263	295	200	251	271
FMP17	465	211	266	415	120	152	416	109	257	265	312	212	253	275
FMP20	?	229	279	398	120	148	411	107	263	273	295	196	253	264
FMP29	456	230	277	395	120	137	395	107	263	271	295	218	257	269
FMP6	436	221	272	391	120	146	406	107	260	265	296	193	253	268
MP10	438	228	265	?	?	142	404	107	263	260	291	212	253	275
MP101	477	231	242	398	N	144	412	107	255	263	297	193	253	260
MP17	438	229	278	413	N	146	403	109	270	273	295	191	265	269
MP24	473	230	265	442	120	142	417	109	265	271	297	200	261	269
MP3	454	230	271	382	120	137	416	106	263	277	289	195	255	275
MP4	449	225	262	395	120	146	416	116	255	269	295	200	251	269
MP41	485	231	279	401	120	137	404	111	255	260	295	204	?	284
MP49	467	223	?	391	120	137	412	109	257	277	?	195	273	268
MP51	440	231	254	398	120	137	398	107	260	267	312	195	257	272
MP54	?	229	?	411	N	142	412	107	262	273	294	206	255	273
MP58	475	225	267	401	120	148	415	109	257	267	297	204	248	268
MP6	449	243	258	400	120	137	417	111	265	?	314	204	273	268
MP67	459	231	256	407	122	156	413	107	263	267	301	200	257	272
MP71	463	229	275	393	N	152	413	107	258	265	295	204	253	261
MP72	440	231	258	393	120	137	?	111	250	260	295	196	253	268
MP73	454	230	270	389	120	142	416	111	271	260	298	195	253	268
MP89	?	239	251	387	120	148	416	107	261	277	297	196	247	277
O_{ST}														
S07BBB	465	227	277	392	122	152	409	107	263	273	300	195	251	268
S01HC1	494	221	254	401	120	137	417	107	261	260	303	198	N	269
D01ACA	454	229	263	393	108	135	423	107	263	262	?	206	251	262
MP19	473	223	257	?	120	142	412	107	257	269	300	191	251	269
MP28	445	221	273	397	120	137	417	107	255	260	297	195	251	275
MP84	438	221	260	397	120	142	407	111	264	263	289	195	251	269
MP65	465	219	277	401	126	141	401	111	262	265	300	212	257	269

MP and FMP = Mt. Parnes; S = Sunne; D = Drøbak; ? = unknown allele

Supplementary tables publication 3

Tables show the nucleotide polymorphic sites for each gene in the *D. subobscura* lines. Different lines are grouped according to their gene arrangement. In the superior part there are lines with the O₃₊₄ arrangement, and in the inferior part the O_{ST} arrangement. Dots indicate nucleotides identical to the reference sequence while dashes indicate gaps. Exonic regions appear shaded. The position of each polymorphic site belongs to the alignment of *D. subobscura*, *D. madeirensis* and *D. pseudoobscura* sequences.

Yrt	8	21	23	32	110	140	143	161	216	263	269	287	305	320	356	362	368	383	389	413	452	461	485	488	515	524	530	533	536	542	545	567	615	617	638	650	665	671	681	683	695	731	771	827	838	863	872	893	899	908					
FMP2	A	T	A	G	G	C	T	C	C	A	G	G	G	G	G	G	G	G	T	A	C	C	C	G	G	C	T	C	G	G	A	G	C	T	C	A	C	G	C	G	G	G	T	G	G	C	G	3+4							
FMP3	.	.	.	C	.	.	G	.	.	T	A	T	T	C	.	.	.	C	T	G	T	C	C	.	.	3+4						
FMP4	.	.	.	C	.	.	G	A	.	C	C	C	.	.	3+4					
FMP6	.	C	.	C	.	T	G	T	A	.	C	C	.	C	.	.	3+4				
FMP10	.	.	.	C	.	.	G	T	T	.	T	T	A	.	C	C	.	.	3+4					
FMP16	.	.	.	C	.	T	G	T	T	C	C	C	.	.	3+4					
MP1	.	.	.	C	.	T	G	.	.	T	A	.	C	A	C	.	C	C	.	.	3+4			
MP2	.	C	.	C	.	T	G	T	T	G	C	C	.	.	3+4				
MP3	-	.	.	C	.	.	G	.	.	T	A	T	T	G	C	.	C	C	.	.	3+4		
MP4	.	.	.	C	.	.	G	.	.	.	C	A	A	T	.	.	3+4				
MP6	.	C	.	C	.	.	G	.	.	T	A	G	C	C	.	.	3+4			
MP10	.	C	.	C	.	T	G	A	.	.	C	C	.	.	3+4
FBC5	G	.	T	C	.	T	G	T	T	C	C	C	.	.	3+4				
FBC15	.	.	.	C	.	.	G	.	T	T	A	G	C	.	.	3+4			
FBC29	.	.	.	C	.	.	G	A	.	C	.	T	T	.	.	.	C	C	T	T	C	C	.	.	3+4			
FBC30	.	C	.	C	.	.	G	G	C	C	.	.	3+4			
FBC37	.	.	.	C	.	.	G	A	T	T	A	.	T	G	.	.	.	T	3+4					
FBC48	.	.	.	C	.	.	G	.	.	.	C	A	A	T	.	.	3+4			
FBC49	.	.	.	C	.	.	G	.	T	T	A	T	G	C	.	3+4			
BC23	.	.	.	C	.	T	G	T	.	.	C	C	.	C	.	3+4			
BC31	.	.	.	C	.	.	G	.	T	T	A	T	.	3+4				
BC34	.	.	.	C	C	.	G	T	G	.	.	T	T	.	3+4				
BC45	.	.	.	C	.	T	G	C	.	T	3+4			
BC56	.	.	.	C	.	.	G	T	.	T	A	C	.	3+4			
MP19	.	C	.	C	.	T	G	C	C	T	T	T	C	.	.	ST				
MP22	.	.	.	C	.	.	G	C	C	.	ST			
MP23	-	.	.	C	.	T	G	C	C	.	ST			
MP28	.	.	.	C	.	.	G	T	A	.	C	C	C	ST					
MP84	.	.	.	C	.	.	G	A	T	T	A	A	.	T	G	ST					
FBC58	.	.	.	C	.	.	G	.	.	.	C	A	A	T	ST				
FBC63	.	.	.	C	.	.	G	C	.	C	ST			
FBC92	.	.	.	C	.	.	G	T	C	.	C	ST			
FBC100	.	.	.	C	.	T	G	T	T	C	C	C	ST					
BC39	.	.	.	C	.	T	G	C	C	T	T	.	A	C	ST						
BC55	.	C	.	C	.	.	G	.	T	A	C	T	T	T	G	T	A	ST				

<i>Atpα</i>	22	52	91	172	173	186	189	191	207	212	224	238	319	436	742	746	868	940	1177	1243	1249	1339	1351	1375	1381	1420	1444	1487	1547	1549	1555	1564	1594			
FMP2	G	G	C	C	A	A	C	T	A	C	C	T	T	C	C	C	G	A	G	C	C	C	C	C	C	G	C	C	G	C	T	C	T	3+4		
FMP3	A	T	A	.	G	.	.	.	C	.	A	G	G	A	G	3+4
FMP4	A	T	A	.	G	.	.	.	C	C	G	A	G	T	.	.	A	T	.	.	A	3+4	
FMP6	A	T	A	T	G	C	A	C	.	A	G	C	T	.	.	A	T	C	.	A	3+4		
FMP10	A	T	A	T	.	T	T	.	A	T	.	.	A	3+4		
FMP16	A	T	A	.	G	.	.	.	C	.	G	T	.	.	A	T	C	.	A	3+4		
MP1	A	T	A	C	.	.	.	G	A	T	.	.	A	3+4		
MP2	A	T	A	T	G	C	A	C	.	A	G	C	C	.	.	C	A	T	C	.	A	3+4		
MP3	A	T	A	.	G	.	.	.	C	.	A	G	C	T	.	.	A	T	.	.	A	3+4		
MP4	A	T	A	.	G	.	.	.	C	.	A	G	A	T	3+4	
MP6	A	T	A	T	G	C	A	C	.	A	G	C	G	T	.	.	A	T	C	.	A	3+4		
MP10	A	T	A	T	G	C	A	C	.	A	G	C	C	.	.	.	G	A	T	.	.	A	3+4		
FBC5	A	T	A	.	G	.	.	.	C	.	A	G	C	C	T	.	.	A	T	.	.	A	3+4		
FBC12	A	T	A	.	G	.	.	.	C	.	G	G	A	T	.	.	A	T	C	.	A	3+4		
FBC15	A	T	A	.	G	.	.	.	C	.	A	G	C	.	.	.	G	T	.	.	A	T	C	.	A	3+4		
FBC29	A	T	A	T	G	C	A	C	.	A	G	C	.	T	A	T	.	.	A	T	.	.	A	3+4			
FBC37	A	T	A	T	G	C	A	C	.	A	G	C	.	.	.	C	A	T	.	.	A	3+4		
FBC48	A	T	A	.	G	.	.	.	C	.	A	G	C	.	T	.	T	T	.	T	.	.	A	T	C	T	A	3+4		
FBC49	A	T	A	A	3+4	
BC23	A	T	A	.	G	.	.	.	C	C	.	G	G	A	T	C	.	A	3+4		
BC31	A	T	A	.	G	.	.	.	C	.	A	T	.	T	A	T	.	.	A	3+4		
BC34	A	T	A	.	G	.	.	.	C	.	A	A	T	.	.	A	T	.	.	.	3+4		
BC45	A	T	A	T	G	C	A	C	.	A	G	C	G	A	T	.	.	A	3+4		
BC56	A	T	A	T	G	C	A	C	.	A	G	C	G	A	T	.	.	A	3+4		
MP19	A	T	A	.	G	.	.	.	C	.	A	A	ST		
MP23	A	T	A	.	G	.	.	.	C	.	A	G	C	A	T	.	.	A	ST		
MP22	ST	
MP28	ST	
MP84	A	T	A	T	G	C	A	C	.	A	G	C	C	.	.	C	A	T	C	.	A	ST		
FBC58	A	T	A	.	G	.	.	.	C	.	A	A	T	ST	
FBC63	A	T	A	.	G	.	.	.	C	.	A	G	C	T	.	.	A	T	C	.	A	ST	
FBC92	A	T	A	.	G	.	.	.	C	C	G	A	T	C	.	A	ST		
FBC100	A	T	A	.	G	.	.	.	C	.	A	G	C	A	T	.	.	A	ST	
BC39	A	T	A	T	G	C	A	C	.	A	G	C	C	A	T	C	.	A	ST		
BC55	A	T	A	.	G	.	.	.	C	C	G	A	T	.	.	A	ST		

Supplementary tables publication 5

Tables show the nucleotide polymorphic sites for each gene in the *D. subobscura* lines. Different lines are grouped according to their gene arrangement, which is indicated in the right side. Dots indicate nucleotides identical to the reference sequence while dashes indicate gaps. Exonic regions appear shaded.

