UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

Inversion polymorphisms: an insight into its molecular genetic content

João Guilherme Patrício Picão de Almeida Osório

MESTRADO EM BIOLOGIA EVOLUTIVA E DO DESENVOLVIMENTO

2009

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA ANIMAL

Inversion polymorphisms: an insight into its molecular genetic content

João Guilherme Patrício Picão de Almeida Osório

Dissertação de mestrado orientada por:

Professora Doutora Marta Pascual – Dept. Genètica, Facultat de Biologia, Universitat de Barcelona

Professora Doutora Margarida Matos – Dept. Biologia Animal, Faculdade de Ciências, Universidade de Lisboa

MESTRADO EM BIOLOGIA EVOLUTIVA E DO DESENVOLVIMENTO

2009

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I want to thank the many people involved in this thesis directly or on the back stage:

First of all, I would like to thank my advisor Professor Marta Pascual for all the support and friendship throughout this research year. I want to stand out how Marta received me, even though with no previous announcement. Our scientific interaction was extremely and contributed greatly to my formation.

To the great pos-doc Pedro Simões (a.k.a. "El Presidente", "Dream team Evolution", Pedrinho or simply as Simões *et al.*) who became a great friend and a great scientific colleague.

All the people of the Evolutionary Genetics Lab or connect to it: my thanks to Víctor Hugo, Gemma, Cinta, Chus,Victor Ordoñez, Celia, Xavi, Albert, Salvatore and Champi; and to the Professors Francesc, Joan and Lluís. It was a great year at the personal and scientific level.

Professor Margarida Matos, for the passionate way of teaching evolution and for all the effort made for thesis to be conducted in Barcelona.

To Professor Élio Sucena for teaching science.

All my friend inside and/or outside science!!!! Many Thanks…(in alphabetic order) A.B., Action, Alex Maria, Aninhas, Cocas, Emília, Fred Cabeças, Francisca, Hany Isadora, J.P., Lima, Machado, Marialva, Moscoso, Patraquim, Pedro Simões, Rita, Sara, VH, Tona, Zuca to all. (In case that I forgot someone, sorry but Alex have to print my thesis)

Isadora to this fantastic year that I hope to be many more.!!

Also, Thanks Aninhas, Pedro Simões and Soraia for this year in Barcelona,

but for some different reasons as above.

Thanks to the all great music, I would not be able to work or do anything else.

I would like to thank the fantastic Erasmus Grant that gave me financial support (€200/month)

To Claudio Alonso for allowing me finish the present thesis during the first weeks of my staying in his lab.

Finally, I am deeply thankful to my family, especially to my great Mother!!

1. RESUMO

As inversões cromossómicas são uma característica inerente ao genoma do género *Drosophila*. As inversões paracêntricas são as mais comuns estando presentes em 60% das suas espécies. Este tipo de inversões não envolvem o centrómero no segmento invertido. Indivíduos heterozigóticos para inversões, isto é heterocariótipicos, apresentam uma redução da recombinação na parte invertida do cromossoma como resultado de problemas mecânicos no emparelhamento dos cromossomas homólogos. Para além do problema físico no emparelhamento, um evento de "crossing-over" em número ímpar durante a meiose originaria gâmetas não balanceados. A reduzida troca de informação genética entre inversões leva a que sejam herdadas como unidades Mendelianas.

Drosophila subobscura é constituída por cinco cromossomas acrocêntricos e um puntiforme: cromossoma A (cromossoma sexual), J, U, E and O (autossómicos); e tem um rico polimorfismo de inversões em todos os seus cromossomas de forma simples (única inversão, ex. A_1) ou composta em ordenamento (múltiplas inversões, ex. O_{3+4}). Esta espécie é nativa da região Paleárctica onde apresenta clines latitudinais para muitos ordenamentos cromossómicos (simples e compostos), ou seja, a frequência dos ordenamentos varia com a latitude. A colonização do continente Americano veio reforçar a importância da selecção natural como principal mecanismo responsável pela formação dos clines. No final da década de setenta *D.subobscura* colonizou a América do Sul e posteriormente a América do Norte, e rapidamente foram estabelecidos clines latitudinais paralelos aos do Velho Mundo. Apesar da evidencia do valor adaptativo das inversões, os mecanismos envolvidos na manutenção e evolução destes polimorfismos ainda estão em debate.

Dois cenários selectivos não mutuamente exclusivos foram propostos para esta problemática: a hipótese de coadaptação *sensu* Dobzhansky e a hipótese de adaptação local de Kirpatrick e Barton. A primeira tem duas características: dentro de um ordenamento os alelos de diferentes *loci* interagem epistaticamente dentro de uma população, e interacções entre diferentes ordenamentos da mesma população promovem a vantagem dos heterocariótipos. Como consequência destes efeitos epistáticos positivos dentro de populações, troca genética entre populações iria romper as associações de alelos coadaptados e levaria uma depressão na eficácia biológica desses indivíduos. Ou seja, os ordenamentos cromossómicos têm propriedades específicas de cada população.

A hipótese de adaptação local não necessita que hajam interacções epistáticas para que a inversão seja favorecida. Quando uma inversão surge, irá capturar conjuntos de alelos que estão independentemente adaptados às condições locais. Os alelos serão mantidos na inversão porque a recombinação é reduzida nos heterocariótipos, e consequentemente a frequência da inversão irá aumentar por ter vantagem sobre os susceptíveis de recombinação. Deste modo, o modelo prevê que a composição genética das inversões seja mantida nas diferentes populações.

Com o intuito de estudar a evolução do conteúdo genético das inversões e inferir qual destes mecanismos selectivos propostos para a evolução e manutenção dos polimorfismos das inversões melhor explica os padrões de variação clinal em *D. subobscura*, estudámos 31 microssatélites que cobrem todo o genoma de *Drosophila subobscura* e sua associação com as inversões cromossómicas em quatro populações Europeias ancestrais e uma Norte Americana recentemente colonizada.

Inicialmente, estudámos padrões gerais de variabilidade e diferenciação tanto a nível de microsatélites como de polimorfismo cromossómico. Para testar as hipóteses de mecanismos selectivos anteriormente expostas, analisámos a existência de diferenciação genética através da estatística *F* (F_{ST}) em diferentes ordenamentos na mesma e em diferentes populações Europeias; e também do mesmo ordenamento entre populações Europeias. De seguida, estudámos os níveis de associações não aleatória, "linkage disequilibrium" (LD), entre os alelos dos microssatélites e os ordenamentos utilizando *loci* localizados dentro e fora dos ordenamentos. Os padrões de desequilíbrio em populações ancestrais e colonizadoras foram comparados para estudar o impacto ao nível genético de um efeito de gargalo associado à colonização da América do Norte.

Tendo esta informação em conta, discutimos se os padrões destas associações são devidas a processos históricos associados com a formação da inversão ou a selecção na formação dos clines latitudinais destes ordenamentos cromossómicos adaptativos. Também interpretámos os padrões obtidos à luz do recente trabalho teórico de Kirpatrick e Barton (2006).

Inesperadamente, os resultados obtidos revelaram histórias distintas para os diferentes cromossomas nas populações Europeias. O cromossoma A não apresentou diferenciação na composição genética entre diferentes ordenamentos na mesma população bem como em distintas populações tanto nos marcadores localizados dentro como fora dos ordenamentos. O mesmo padrão foi observado para o mesmo ordenamento entre populações. Nós propusemos que a falta de diferenciação poderia ser uma consequência de homoplasia. A alta variabilidade do cromossoma sexual nas populações ancestrais, provavelmente resultante de *background selection*, levaria a homoplosia dos marcadores neutros. Apesar desta variabilidade genética e ausência

de diferenciação entre ordenamentos, apenas os *loci* dentro das inversões apresentavam padrões de desequilíbrio, e os ordenamentos variaram na sua associação com os alelos de microssatélites. Algumas destes padrões de LD mantiveram-se entre populações sugerindo a acção da selecção em regiões adjacentes a estes *loci*.

O conteúdo genético dentro dos ordenamentos dos cromossomas J, E e O manteve-se constante entre populações Europeias, ou seja, a identidade de cada ordenamento manteve-se ao longo da Europa. Pelo contrário, na mesma população a composição dentro dos distintos ordenamentos era significativamente diferente. Estes padrões de composição génica dos ordenamentos estão de acordo com a hipótese de adaptação local. Em estudos anteriores foi reportado a existência de fluxo genético entre populações e entre diferentes ordenamentos, no entanto os padrões genéticos mantêm-se. O que sugere que forte selecção estaria a contrabalançar o efeito homogeneizador da migração e da troca genética entre ordenamentos. Assim, o nosso argumento que a adaptação local seria o mecanismo responsável pelo valor adaptativo dos cromossomas J, E e O ganha robustez.

Os ordenamentos do cromossoma U apresentavam composições específicas para cada população Europeia, ou seja, o mesmo ordenamento era diferente na sua composição entre populações. Adicionalmente, os diferentes ordenamentos eram distintos tanto na mesma população como entre populações. Os padrões de LD também eram específicos de cada população, ou seja, as associações não ao acaso entre alelos de microssatélites e ordenamentos diferiam entre populações. Mesmo tendo em conta a existência de migração e troca genética entre ordenamentos, as características especificas de cada ordenamento em cada população são mantidas. Consequentemente, forte pressão selectiva dentro das inversões mantém o conteúdo

genético especifico de cada população no cromossoma U, o que está de acordo com a hipótese de coadaptação *sensu* Dobzhansky.

A população Norte Americana apresentava variabilidade reduzida, e tanto a composição genética como os padrões de associação eram distintos dos Europeus. O efeito gargalo foi determinante na composição genética desta população que apresentava grande LD ao longo do genoma e tendo os ordenamentos dos diferentes cromossomas em associação com vários alelos. Não podemos fazer nenhuma ilação se os processos no continente Americano são os mesmo que no Velho Continente. Contudo, segundo o trabalho teórico de Kirkpatrick e Barton são necessários dois *loci* para o cenário da adaptação local se processar. Assim, nesta recente população os *loci* que estão a dar o valor selectivo à inversão poderão ser poucos e não coincidirem com os ancestrais devido ao efeito fundador.

Os dados obtidos e discutidos nesta dissertação indicam que diferentes mecanismos selectivos podem estar a afectar o genoma de *Drosophila subobscura*. Estudos clássicos de associação entre aloenzimas e inversões também apresentavam um padrão de não diferenciação do mesmo ordenamento entre populações nos cromossomas E e O. O modelo de adaptação local parece estar amplamente distribuído no género *Drosophila*, uma vez que encontrámos os mesmo padrões em estudos anteriores em diversas espécies de *Drosophila*.

PALAVRAS-CHAVE: Polimorfismo das Inversões Cromossómicas, Adaptação Local, Coadaptação, *Drosophila subobscura*, "Linkage Disequilibrium", Populações Naturais

2. ABSTRACT

There is compelling evidence supporting an adaptive explanation for the evolution of inversion polymorphism in *D.subobscura*. Nevertheless, the specific mechanisms that underlie the maintenance and evolution of these polymorphisms are still in debate. The analysis of the associations between chromosomal inversions and microsatellite alleles is extremely useful to address the evolution of the genic content of inversions. Here we present a survey of 31 microsatellite *loci* covering the whole genome of *D.subobscura* and their association with chromosomal inversions on four European (Barcelona, Mount Parnes, Drøbak and Sunne) and one North American populations (Bellingham). We observed high linkage disequilibrium (LD) between most microsatellite loci and chromosomal arrangements in the colonizing population as a result of the founder event. Among European populations we have observed: i) no genetic differentiation between arrangements and within the same arrangements between populations for the sex-chromosome, ii) higher genetic differentiation between arrangements within populations than within the same arrangements between populations for chromosomes J, E and O, and iii) the arrangements of chromosome U had population-specific properties. The LD patterns between autosomal arrangements and microsatellite *loci* were in accordance with the genetic differentiation of the arrangements. Furthermore, we discuss whether this LD patterns in most localities are due to historical processes associated with the inversion formation or to selection. Our results provide strong support for distinct selective mechanisms on the maintenance of inversion polymorphisms, namely: local adaptation hypothesis on chromosome J, E and O, and coadaptation *sensu* Dobzhansky on U chromosome.

KEYWORDS: Chromosome Inversion Polymorphisms, Local adaptation, Coadaptation, *Drosophila subobscura*, Linkage Disequilibrium, Natural Populations

3. INTRODUCTION

3.1. Chromosomal Inversions

During the evolution of the genus *Drosophila*, it appears that chromosomal inversion polymorphism represents an inherent attribute of its genome (GREGORY 2004; KRIMBAS and POWELL 1992), with about 60% of its species presenting paracentric inversions (POWELL 1997). This kind of inversions are characterized by an inverted segment of the chromosome where the breakpoints occur on one side of the centromere (HOFFMANN *et al.* 2004), *i.e.*, the inversion does not involve the centromere. Individuals heterokaryotipic for inversions present reduced recombination (POWELL 1997). The products of single or odd crossover events within the inversion are unviable by originating unbalanced gametes (HOFFMANN *et al.* 2004; POWELL 1997). Additionally, inversion heterozygotes promote mechanical pairing problems in the inversion breakpoints (GRIFFITHS *et al.* 2004). These factors contribute to inversions becoming "*inherited intact as single simple Mendelian units"* for the majority of the chromosome (POWELL 1997).

3.2. Origin of Chromosomal Inversions

Surprisingly, although the study of chromosome rearrangements has been developed since the 30's, the evolution and origin of such structures remains largely unclear (CASALS and NAVARRO 2007). The broadest idea comes from Finnegan in 1989 (FINNEGAN 1989) – the ectopic recombination (or illegitimate recombination or non-allelic homologous recombination; (CASALS and NAVARRO 2007). This idea is supported by the existence, in some *Drosophila* lineages, of transposable elements in inversion breakpoints. If these segments are placed on the chromosome in an inverted

orientation, the result of a recombination event will produce a chromosome inversion (CASALS and NAVARRO 2007; FINNEGAN 1989). A recent work by Ranz and collaborators raised another possible scenario (RANZ *et al.* 2007). The authors found a high frequency of association between the breakpoint regions of the inversions and inverted duplications of genes or other nonrepetitive sequences. Furthermore, a very low association between the breakpoint regions and inverted repetitive sequences was found in the *melanogaster* species group. In this new circumstance, the duplications at the breakpoints are not the cause of the inversion, but a by-product of staggered breaks caused by the occurrence of an inversion - staggered breaks model (CASALS and NAVARRO 2007; RANZ *et al.* 2007). This novel data shed a fresh light onto the mechanisms behind chromosomal rearrangements and stand, together with the previous model, as the two most predominant explanations for this process.

3.3. Adaptive Role of Chromosomal Inversions

The adaptation paradigm is one of the most debated topics in evolutionary biology (LEWONTIN 1974; ORR 2005). Despite this, much is still unknown about the genetic basis underlying adaptation. Furthermore, its impact in the shaping of genomes relative to other evolutionary forces/processes is sometimes not easy to access (e.g. (HAHN and RAUSHER 2008). Namely, historical contingency could have a confounding effect or bias some interpretations in this problematic (GOULD and LEWONTIN 1979; TRAVISANO *et al.* 1995). Chromosomal inversion polymorphisms were the first genetic markers in which the action of natural selection on natural populations was empirically shown, through the pioneering work of Dobzhansky and colleagues (DOBZHANSKY 1970). The seasonal frequency changes (DOBZHANSKY 1943; FONTDEVILA *et al.* 1983; RODRIGUEZ-TRELLES *et al.* 1996) as well as their

clinal variation at a broad geographical scale (BALANYA *et al.* 2003; BALANYA *et al.* 2004; KRIMBAS and POWELL 1992) points to an adaptive role for these inversions polymorphisms.

Latitudinal clines for most chromosome arrangements are well establish in Europe and have been maintained for many years in *Drosophila subobscura* populations (BALANYA *et al.* 2004; KRIMBAS and POWELL 1992). Furthermore, the repeatable clinal patterns for chromosomal inversions found in both North and South America few years after its colonization (PREVOSTI *et al.* 1988), and the maintenance of these patterns reinforce the idea that the clines evolved by natural selection (BALANYA *et al.* 2003). More recently it was shown that the frequency of these arrangements tracks global warming at a worldwide scale (BALANYA *et al.* 2006). However, much is still unknown about the mechanisms that underlie the maintenance and evolution of these polymorphisms, *i.e.*, the genetic basis of this adaptation (see (HOFFMANN and RIESEBERG 2008).

Two main selective scenarios, not mutually exclusive, have been suggested to explain the evolution and maintenance of these polymorphisms: the coadaptation hypothesis *sensu* Dobzhansky (DOBZHANSKY 1950) and the local adaptation hypothesis (KIRKPATRICK and BARTON 2006). The former concept entails two features; alleles at different *loci* within gene arrangements interact epistatically within local populations, and interactions between different gene arrangements from the same population promote selective advantage in heterokaryotypes (DOBZHANSKY 1950; PRAKASH and LEWONTIN 1968). As a consequence of these positive epistatic interactions within local populations, genetic exchange among chromosomal arrangements from different populations will lead to an outbreeding depression due to the disruption of the sets of coadapted alleles (DOBZHANSKY and EPLING 1948). Thus,

chromosomal arrangements have population-specific properties.

Under the local adaptation hypothesis (KIRKPATRICK and BARTON 2006) no epistasis is required for inversions to be favored. When an inversion arises it may capture together modules of alleles that are independently adapted to local conditions. These alleles will be maintained in the inversion due to recombination reduction in heterokaryotypes. Consequently, they will increase in frequency since they will have higher fitness relative to the ones susceptible to recombination. Therefore, the alleles harbored within the inversion will spread, and as a result the inversion will spread with them. One expected output of this model would be no genetic differentiation within the same gene arrangements along different populations, that is, gene arrangements would have their own identity. On the other hand, under the coadaptation model, genetic differentiation is expected both between similar gene arrangements across populations and between different gene arrangements in the same population. It is important to state that although this local adaptation model does not require epistasis, it also does not neglect it, as the authors propose that epistasis might in fact enhance the evolution of chromosome inversions.

3.4. Molecular Markers and Inversions

The association studies between chromosomal inversions and molecular markers are extremely useful to address the evolution of the genic content of inversions. In the last decades of the XX century several studies were done trying to establish an association between allozymes and inversions (LOUKAS and KRIMBAS 1975; LOUKAS *et al.* 1979; PRAKASH and LEWONTIN 1968, 1971; ZOUROS and KRIMBAS 1973). The general conclusion was a significant non-random association

between some allozymes with inversions, with different populations presenting the same association patterns. Generally, these studies observed similar patterns of considerable genetic differentiation between arrangements and no differentiation on the same arrangement for most populations analyzed. Molecular studies conducted in *D.subobscura* reveal the same patterns of significant genetic differentiation between gene arrangements, even though genetic exchange can occur within inversions between gene arrangements by gene conversion and/or by double crossovers (MUNTE *et al.* 2005; NOBREGA *et al.* 2008; ROZAS and AGUADÉ 1990,1994; ROZAS *et al.* 1999). Recent molecular studies have claimed that coadaptation is an important mechanism for the evolution of chromosomal inversion in *Drosophila* (KENNINGTON *et al.* 2006; SCHAEFFER *et al.* 2003). In both works the authors found significant genetic differentiation between chromosomal arrangements, which is in accordance with the coadaptation model *sensu* Dobzhansky, although no genetic differentiation was found for the same chromosomal arrangements in different populations. Also, they report evidence for statistical epistatic selection in *loci* inside the inversion based on the detection of regions with high linkage disequilibrium (LD) interspersed by regions of low LD.

3.5. Objectives

The majority of the molecular studies mentioned above were performed in few chromosomes and/or populations. So, a more global picture of the molecular genetic content of these gene arrangements both at the whole genome level and also at a broad geographical scale is still missing. Here, we present a survey of 31 microsatellites covering the whole genome of *Drosophila subobscura* and their

association with chromosomal inversions on four European ancestral populations and one North American recently colonized population of this species.

We will thus be able to address the evolution of the genic content of inversions and infer which of the two selective mechanisms proposed for the evolution and maintenance of inversion polymorphisms - the coadaptation and the local adaptation models – better fits our data. Specifically we aim to:

- test the existence of genetic differentiation in different arrangements and/or in the same chromosomal arrangements across different populations.

- study the levels of linkage disequilibria (LD) between microsatellite alleles and inversions, using several *loci* that are mapped inside and outside these chromosomal arrangements.

- analyze LD patterns in ancestral and colonizing *D. subobscura* populations. This will allow us to determine to what extent do the populations of the Old World that span a broad latitudinal and longitudinal range present the same patterns of LD as well as compare those with the ones presented by the colonizing population. We will also address the impact of the bottleneck associated with the colonization of North America at the genetic level.

- Taking all this information in account, we will further discuss whether this linkage disequilibrium patterns in most localities are due to historical processes associated with the inversion formation or to selection driving the latitudinal clines of these adaptive chromosomal arrangements.

4. MATERIAL AND METHODS

4.1. Geographic Samples

Wild *Drosophila subobscura* samples were collected from four European populations and one North America population in different years and seasons (Figure 1). The Barcelona (Spain) population was sampled in October 2007 (95 individuals), Mount Parnes (Greece) in May of 2006 (103 individuals), Drøbak (Norway) and Sunne (Sweden) in August 2005 (80 and 63 individuals, respectively) and Bellingham (USA) in October 2006 (94 individuals).

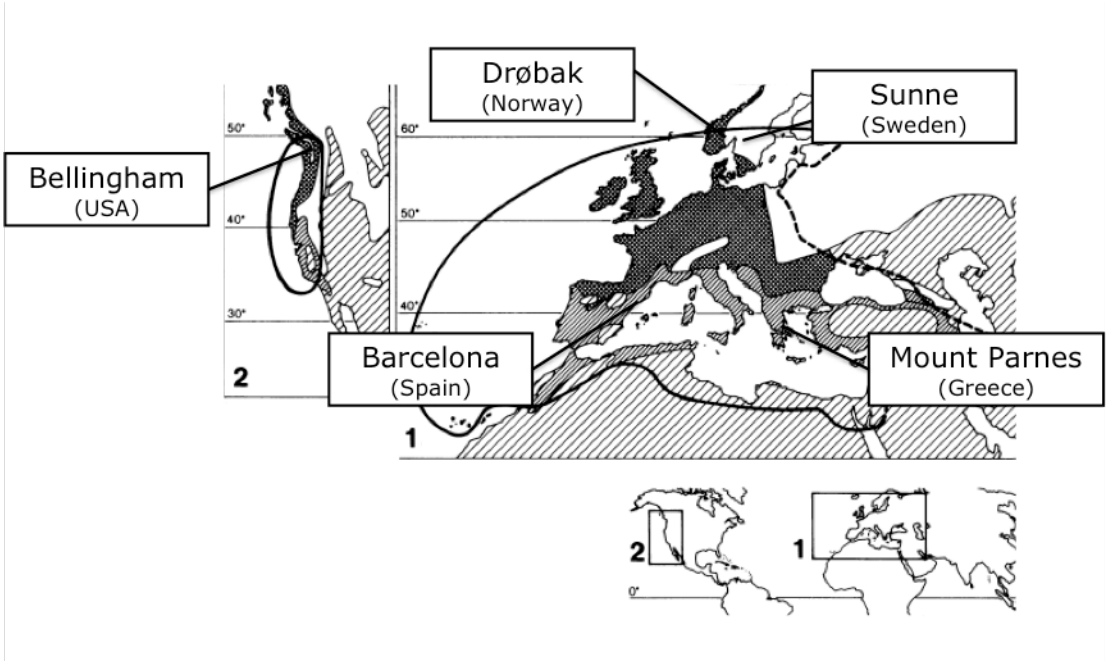

FIGURE 1. - Map of Europe (1) and North America (2) and the sampled sites. Adapted from PREVOSTI *et al.* 1988

4.2. Chromosomal Inversions

The karyotype of *D. subobscura* consists of five acrocentric chromosomes and a dot chromosome. (MAINX *et al.* 1953) named as A (=X, the sex chromosome), J (chromosomal element D of Mueller/Sturtevant/Novitski and homologous to arm 3L

in *Drosophila melanogaster*), U (chromosomal element B and homologous to arm 2L), E (chromosomal element C and homologous to arm 2R), and O (chromosomal element E and homologous to arm 3R).

In order to score the chromosomal arrangements, wild-caught males or males descendant from isofemale lines were individually crossed with virgin females of the *chcu* strain. The *chcu* strain carries the recessive markers cherry eyes and curled wings and its genetic background is highly homogeneous and homokaryotypic for the chromosomal arrangements A_{ST}, J_{ST}, U_{ST}, E_{ST} and O_{3+4} (KOSKE and MAYNARD SMITH 1954). One female third-instar larva from each cross was dissected and then examined for its polytene chromosomes to ascertain the arrangements of one set of the chromosomes from the wild. The remaining of the larva was preserved in ethanol at -80ºC for posterior DNA extraction. Salivary glands were stained with 2% orcein in 60% acetic acid mixed 50: 50 with lactic acid. The chromosomal arrangements were designated according to (KUNZE-MÜHL and MÜLLER 1958). All crosses were kept at 18ºC.

4.3. DNA Extraction and Microsatellite Amplification

Genomic DNA was individually extracted from each larva of the crosses (also used to score inversion polymorphism). The samples were hydrated with TE 0,1x (10mM TrisHCL, 1mM EDTA, pH 8) overnight at 4ºC prior to extraction. Each larva was homogenized in a microtube with 160µl of cold solution I (10mM Tris, 60mM NaCl, 5% Sucrose, 10mM EDTA, pH 7,8). Then, 200µl of solution II (300mM Tris, 1,25% SDS, 5% Sucrose, 10mM EDTA, pH 8) were added, the homogenate was mixed by inversion of the tube and the tube was placed in a water bath at 65ºC for 30min. A total of 60µl of 3M Sodium Acetate was added and the tube was cooled at -

20ºC for 20min. After centrifugation for l5min (12000 r.p.m.) an equal volume of Isopropanol (~400µl) was added to the aqueous phase and the sample was maintained at room temperature for 5 min. After centrifugation for 10min the pellet was rinsed with 500µl of 70% ethanol, vacuum dried, and resuspended in 50µl of double deionized sterile water. The primers are described in Table 1.

The thirty-one microsatellite *loci* analyzed in this study were: dsub11, dsub37, dsub76, dsub05, dsub21, dsub39, dsub70, dsub19, dsub18, dsub59, dsub69, dsub74, dsub62, dsub27, dsub10, dsub03, dsub42, dsub64, dsub15, dsub46, dsub68, dsub79, dsub20, dsub53, dsub28, dsub02, dsub14, dsub38, dsub01, dsub04, dsub12 (Table 1). These markers had been previously identified and characterized in *D. subobscura* (PASCUAL et al. 2000), and were chosen due to their localization on the chromosomes in respect to their inversions (Figure 2; chromosomal localization given by Santos personal communication).

The markers were amplified using six different multiplex PCR reactions. The *loci* of each chromosome were amplified together on a multiplex PCR reaction using the Qiagen Multiplex Amplification Kit. The *loci* dub64, dsub46 and dsub20 were amplified apart on two PCR reactions: dsub64 with the Amersham Taq polymerase and dsub46+dsub20 with the Qiagen Kit. The amplification reactions with the Qiagen Kit were performed for a total volume of 15 µl with 7,4 µl of the Kit's Master Mix, 1,5µl of primer mix (2 µM each primer) and 1µl of DNA. The multiplex amplification reactions with the Amersham Taq polymerase were carried out in a 15µl reaction containing 1,5µl of 10x Buffer, 2,4µl dNTP (at 1mM each), 0,12µL of Taq polymerase, 2,1µl of $MgCL_2$ (25mM), 1,2µL of primer mix (2 µM each primer), 0,15 µl of BSA(10mg/ml) and 1µl of DNA.

TABLE 1 - Primers and characteristics of the 31 microstallite *loci*.

Chromosome	Microsatellite *locus* *	Localization	Foward / Reverse (5' - 3')	Dye	Allele Range (bp)	*ch-cu* Allele (bp)
A	dsub11	03B	TTATGCTTATGAGCGCGATG GGACAACGTTGCTCAGTTCA	HEX	228 - 256	232
	dsub37	06C	TTGTTGTCCAACTGCTGACC TAAACGACGAACCCAAGAGG	NED	204 - 267	242
	dsub76	08B	TGTCGGGGTTAGGCTGAG GCTGGGATTAGGCTTGCAT	HEX	293 - 341	316
	dsub05	08E	ACTGCAAATGCAGCAAACAG AAACGCGTACAGCATGAGTG	NED	133 - 180	152
	dsub21	11A	GTGCCGTTCCTGGAAATAAA ACCCATTCAAGTATGCCGAG	HEX	166 - 208	185
	dsub39	12B	GATACGTTCCTCGGTGCATT CGGAATGGAGATACGATTGG	6-FAM	262 - 294	273
	dsub70	13A	CAGGAAAGTACTTGGAAAGTCG GTGCGTGGAAAAATGGAACT	NED	287 - 342	311
	dsub19	15B-C	AGGAACACCATAGCCACAGC TTGCGATGACAGTAAGGCAG	6-FAM	173 - 209	181
J	dsub18	21C	ATAAAACGTTTCTGCGGCAT AGTCAAGCGGTAGTCAGGGA	NED	193 - 237	216
	dsub59	23D	GGGAGCCCCCTGAGTAGATA CCGGCTCTAATCAAGCACTC	NED	212 - 314	258
	dsub69	24D	GGCTCTACGGTTCAGACTGC TTTTATTGCCACACCAGACG	HEX	127 - 168	143
	dsub74	25B	GCGCCTACCTTAACCCAATTA GTACCCGCTCTTTGCTATGC	6-FAM	149 - 198	168
	dsub62	29C-D	GGCATCTCGGCATTTACATA GATTACGCGGCAATTTGTTT	6-FAM	267 - 337	313
	dsub27	33A	TATGGCTGTTTGTCTTGGCA TTGTTGCCCCTGTATCTTCC	HEX	209 - 258	224
U	dsub10	36A	GATTGCTTGATTGGCGATTT GGGGCATCGATTATGAACAC	NED	193 - 327	258,260,262,264
	dsub03	37B	CTCCAAACTGCTGCTCATCC TTTCTCAAACCACAGCAGGA	6-FAM	131 - 162	145
	dsub42	44B	CAACTTCCTTTGCGTGTCAA TGGAGCAACACCAGCTACAG	NED	107 - 140	120
	dsub64	48B	TGAAGATGCTGATGGTTGGT GGGATCCGCAACTTGTTGTT	HEX	148 - 194	163
	dsub15	52A	AAAGACTTTGACGCGACGAT CCATGTCTGGCCAACTATCA	6-FAM	225 - 269	255
E	dsub46	54D-E	CGCCAATGCTCGTAATACAC CCGAGCTCTCAGCTCTCAGT	NED	317 - 347	327
	dsub68	54D-E	AATTGAATTCTTCGCCCAAC AGCCAGCAGCAAATGTAGGA	HEX	202 - 240	206
	dsub79	58A	ATTTCCACATGCCACAACAC CCTGTCAATGCGTGACATTT	6-FAM	123 - 177	149
	dsub20	62B	CCACCCTAAGTTTTGCCTCA CAGCCGCAGAACAGAAAAAT	NED	233 - 307	290,292
	dsub53	70A-B	AAACATTTGGGCAAGGACAC TTCCTGGTTTCGTTCGTTTC	6-FAM	278 - 333	308
	dsub28	73C	GCTAGTCGCTAGTCGCTCGT ATTTAAGCCAGGCCGAGAGT	HEX	159 - 197	177
O	dsub02	78A	CCAGGTACGTGTCACACAGG TGACAAAAAGGACAACCTTCG	NED	210 - 254	232
	dsub14	82C	AGCTGCCCTCCCAATGAT TTCCTCTGGTCAACCCTTTG	HEX	107 - 129	121
	dsub38	86E	ATCCCAATACCCCTGGTAGC TTCCATTAAGGCCCACTCAC	HEX	392 - 461	417
	dsub01	90A	CCAGAGCACTCGTGTAAGCA ACGTTTTGTCTTTCGCTGGT	HEX	255 - 281	263
	dsub04	92D	GCACTTGAAGTCTTGTGGCA TTGACGACTTCATGCTCAGG	6-FAM	164 - 242	197
	dsub12	98B	TCTCTGTGCTCCTGCCACTA CCCCAAGCATTATGCAATTT	6-FAM	216 - 308	272

All reactions were performed on an AB GeneAmp PCR System 2700 machine using the following steps: 15 min at 94°C, then

30 cycles of 30s at 94°C, 30s at 50°C and 30s at 72°C followed by 30min at 60°C. After amplification, the products were visualized in an agarose gel and then loaded on

an ABI PRISM 3700 automatic sequencer from the Scientific and Technical Services of the University of Barcelona, with CST ROX 70–500 (BioVentures, Inc.) used as internal molecular ladder, and allele sizes were assigned with GeneMapper™ ID version 3.7 (Applied Biosystems, Inc.)

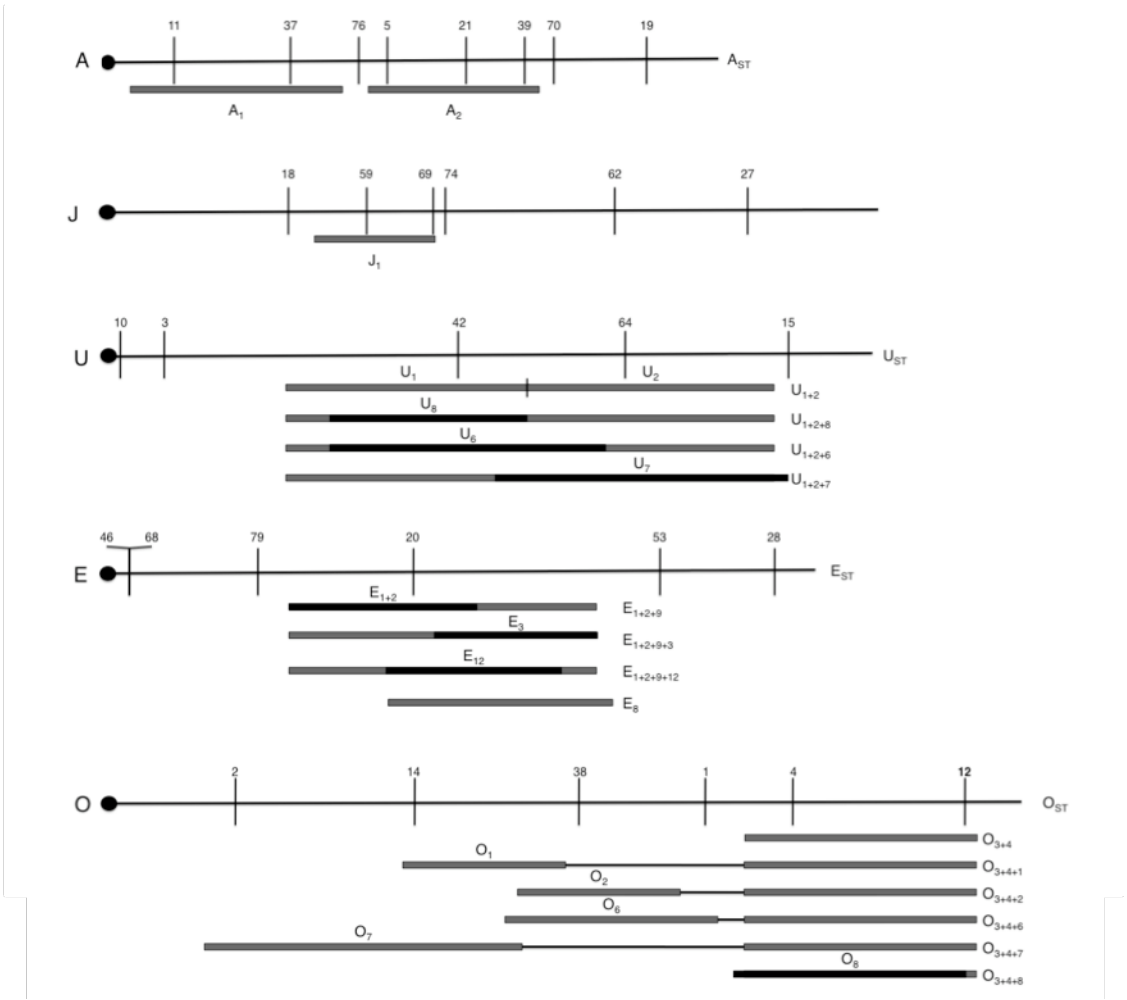

FIGURE 2.-Schematic representation of the location of the microsatellite *loci* in chromosomes A, J, U, E and O of *Drosophila subobscura*, , with inversion positions also marked (for O chromosome are only represented the most common arrangements). The centromere is placed on the left (black circle) and the telomere on the right. The linear order of micrsosatellite is their localization in the standard arrangements

4.4. Statistical Methods

Molecular Genetic variability was measured using both allelic richness and expected heterozigosity (*He*, or gene diversity) with FSAT software package (GOUDET 2001). The variability of each population (Figure 3) was an average of the variability of all microsatellite *loci* of all chromosomes. Patterns of inversion variability were obtained considering each chromosome as a *locus* and each arrangement as an allele. The allelic richness and expected heterozygosity were compared among populations with the Wilcoxon Matched Pairs Test using Statistica 9 software.

Genetic differentiation was calculated with *F*-statistics (WEIR and COCKERHAM 1984) and their significance were determined using FSTAT (GOUDET 2001) with sequential Bonferroni correction. To determine the genetic differentiation of chromosomal arrangements we treated each arrangement of a European population as a "subpopulation". Then, we compared the genetic differentiation (F_{ST}) of all pairs of comparisons between arrangements of the same chromosome and grouped in three categories: i) the same chromosome arrangement in different populations, ii) different chromosome arrangements in the same population, and iii) different chromosome arrangements in different populations. The F_{ST} in the *loci* that mapped inside and outside the inversions was calculated by pair of comparisons of the inversions represented on the majority of the populations: A_{ST}, A_1 and A_2; J_{ST} and J_1; E_{ST}, E_{1+2}, E_{1+2+9} and $E_{1+2+9+12}$; U_{ST}, U_{1+2} and U_{1+2+8}; O_{ST}, O_{3+4} and O_{3+4+8}. We used the Mann-Whitney *U* test of Statistica 9 software to compare the groups of *loci* inside and outside for each category, and to analyze the differentiation of all *loci* of each chromosome in the different categories.

Linkage disequilibrium between genetic markers and inversions for each

chromosome in each population was quantified with the multiallelic version of Lewontin's *D'*-statistic, $D'm=\Sigma_{ij}\ p_i q_i|D'_{ij}|$ (LEWONTIN 1964) using the software PowerMarker version 3.25 (LIU and MUSE 2005). Statistical significance was evaluated using the Fisher exact test implemented in PowerMarker. The *P*-values were obtained with 10000 permutations (Markov chain Monte Carlo gave similar results) and adjusted for multiple comparisons.

The specific LD patterns between microsatellite alleles and chromosomal arrangements was calculated with an interallelic disequilibrium measure (*D'* statistic) between multiallelic markers implemented in MIDAS (GAUNT *et al.* 2006). The significance was measured using a Chi square (χ^2) with Yates correction. Only the significant associations after the correction were represented. The chromosomal arrangements with a frequency of less than 5% in a population were not analyzed to avoid the detection of singletons. In the same way, for a significant association between microsatellite alleles and inversions, only the alleles that appeared more than three times were considered to avoid spurious LD patterns.

5. RESULTS

5.1. Population Variability

The populations differ in their genetic diversity according to their geographical localization on both levels of biological organization: inversion polymorphisms and microsatellite *loci* (Figure 3). The two Southern European populations, Barcelona and Mount Parnes, were the most variable for microsatellite *loci* (Table 2) both in allelic richness (mean value of 16.51) and gene diversity (mean value of 0.83) (see Figure 3B and 3D, respectively). For these molecular markers, although the two Northern European populations - Drøbak and Sunne - had significantly lower values of allelic richness (12.08 and 11.09, respectively) relative to both Southern populations (Wilcoxon Match Pairs Test, $P < 0.0003$ for all population comparions), their gene diversity (or expected heterozygosity) was relatively similar (mean $He = 0.8$ for the Northern populations *vs.* mean $He = 0.83$ for the Southern populations). The colonizing population, Bellingham, was the less variable for the microsatellite *loci* and presented both a significantly lower allelic richness (mean value of 5.52) (Wilcoxon Match Pairs Test, $P < 0.000002$ for all population comparisons) and gene diversity ($He = 0.71$; Wilcoxon Match Pairs Test, $P < 0.0002$ for all population comparisons) relative to Europe.

The patterns of inversion variability were measured considering each chromosome as a *locus* and each arrangement as an allele (Figure 3). The populations of Barcelona and Mount Parnes had the highest values of allelic richness (averages of 4.89 and 4.74, respectively; here allelic richness refers to number of arrangements per chromosome per population standardized to the same sample size in all populations)

and *He* for their inversion polymorphism. However differences in allelic richness between Northern and Southern locations were not statistically significant (Wilcoxon Matched Pairs Test (W), $P > 0.05$). However, differences were statistically significant when comparing Barcelona *vs.* Northern European populations for *He* (W, $P < 0.05$).

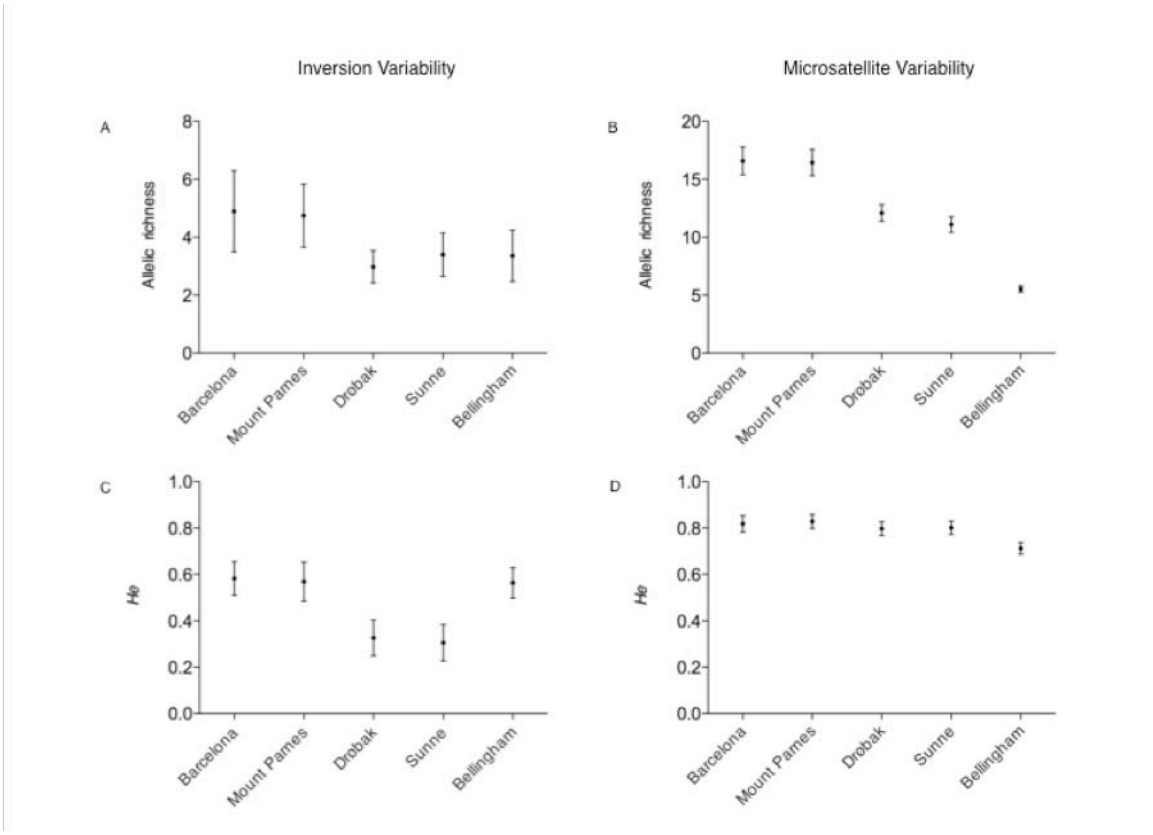

FIGURE 3.- Variability of the inversion polymorphism (A and C) and the microsatellite *loci* (B and D) in each population. The variability is measured on their allelic richness and expected heterozygosity (*He*).

Bellingham also presented lower, although not significant, allelic richness in comparison with the Southern European populations i.e. its values of allelic richness were identical to those of the two North European populations (average of 3.36; Figure 3A). The Bellingham population, on the other hand, had similar expected heterozygotes for their inversions as the two South European populations (*He* = 0.56; Figure 3C) and significantly higher values of *He* relative to the two Northern European populations (W, $P < 0.05$ for both population comparisons). That is, in spite

of Bellingham having roughly the same number of different chromosomal arrangements as the North European populations, their frequencies were intermediate as happens in the South European populations.

Analysis of the genetic diversity patterns for each chromosome - that is, the genetic diversity for the microsatellite *loci* that mapped in each chromosome – showed that the genetic composition of all European chromosomes were significantly different from the colonizing in their allelic richness (W, P <0.05) and in the expected heterozygosity only for the A and O chromosomes (W, $P < 0.05$) – see Table 2. There were also significant differences in allelic richness for the A chromosome between the Southern populations (Barcelona and Mount Parnes) and the Northern populations (Drøbak and Sunne) (W, $P < 0.05$), and on the O chromosome for both Southern populations and Sunne (W, $P < 0.05$).

The levels of microsatellite *loci* variability also differed between the sex (A) and autosomal (J, U, E and O) chromosomes in Europe and North American populations (Table 3). While A chromosomes were more variable than autosomes in European populations (particularly for *He*, Mann-Whitney U Test, $P < 0.05$ for all populations), the North American population had slightly higher variability for the autosomal chromosomes, although differences were not statistically significant (Mann-Whitney U Test, $P = 0.11$ for both allelic richness and *He*).

TABLE 2 – Genetic variability of the 31 microsatellite *loci* in all populations

Chromosome	Microsatellite *locus*	Population									
		Barcelona		Mount Parnes		Drøbak		Sunne		Bellingham	
		Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*
A	dsub11	13,183	0,883	12,451	0,882	9,830	0,881	10,000	0,901	6,524	0,805
	dsub37	27,817	0,951	31,442	0,952	19,599	0,929	22,000	0,909	4,000	0,722
	dsub76	13,866	0,841	16,518	0,898	13,630	0,864	11,995	0,869	5,762	0,728
	dsub05	17,699	0,907	16,157	0,918	10,922	0,884	11,994	0,864	5,524	0,712
	dsub21	17,126	0,893	17,333	0,915	7,854	0,847	9,991	0,853	4,000	0,560
	dsub39	26,043	0,943	21,654	0,913	14,626	0,808	12,980	0,847	3,762	0,633
	dsub70	16,656	0,916	20,180	0,923	19,648	0,928	16,993	0,923	4,000	0,725
	dsub19	15,023	0,895	14,055	0,898	14,600	0,885	10,000	0,861	5,945	0,729
J	dsub18	16,998	0,885	12,991	0,885	7,998	0,820	8,960	0,810	7,000	0,600
	dsub59	24,996	0,766	20,955	0,865	14,412	0,614	8,980	0,520	4,000	0,678
	dsub69	14,998	0,855	10,988	0,831	10,673	0,790	9,960	0,792	4,000	0,746
	dsub74	16,999	0,886	10,973	0,828	13,717	0,862	11,960	0,859	5,000	0,797
	dsub62	18,999	0,935	24,928	0,943	13,899	0,908	15,000	0,910	6,000	0,770
	dsub27	15,998	0,872	16,937	0,872	-	-	-	-	5,000	0,764
U	dsub10	30,540	0,950	25,195	0,944	13,814	0,906	12,986	0,898	9,558	0,856
	dsub03	8,347	0,769	12,588	0,784	9,793	0,833	8,999	0,866	4,999	0,695
	dsub42	8,943	0,818	8,684	0,754	8,784	0,766	8,986	0,709	6,000	0,789
	dsub64	19,641	0,923	25,465	0,916	12,508	0,839	11,986	0,795	6,794	0,811
	dsub15	13,268	0,857	13,338	0,871	7,906	0,800	7,000	0,793	6,000	0,774
E	dsub46	8,914	0,404	10,287	0,469	7,833	0,649	6,000	0,740	5,808	0,746
	dsub68	14,783	0,844	14,577	0,836	12,754	0,805	11,000	0,755	6,797	0,781
	dsub79	3,786	0,266	4,489	0,248	5,982	0,380	5,000	0,441	2,792	0,367
	dsub20	24,909	0,939	27,008	0,951	11,992	0,918	11,000	0,919	6,790	0,790
	dsub53	13,421	0,834	14,558	0,819	12,764	0,821	12,987	0,829	6,750	0,789
	dsub28	12,824	0,682	10,969	0,717	11,530	0,558	9,999	0,684	4,000	0,724
O	dsub02	15,909	0,867	19,851	0,906	16,350	0,886	14,975	0,884	5,000	0,747
	dsub14	2,815	0,139	5,329	0,350	3,918	0,192	3,000	0,207	2,970	0,143
	dsub38	19,707	0,883	18,469	0,881	8,997	0,796	7,998	0,824	5,815	0,754
	dsub01	12,253	0,881	12,438	0,879	12,859	0,888	11,983	0,843	5,995	0,801
	dsub12	27,916	0,945	20,404	0,921	21,493	0,936	16,979	0,914	7,815	0,825
	dsub04	19,413	0,916	18,429	0,910	11,889	0,854	10,983	0,823	6,815	0,720

Note: we have no results for dsub27 in Drøbak and Sunne

5.2. Population Differentiation

Significant genetic differentiation (measured through F_{ST} values) between the European and the colonizing population was found for microsatellite *loci* (Figure 4B). Likewise, genetic differentiation was also found between Northern and Southern European populations for most chromosomes, with the loci in the U chromosome showing the largest genetic differentiation between them. In contrast, we found almost absence of genetic differentiation between the two Northern or the two Southern populations (Figure 4B).

TABLE 3 - Differences of sex and autosomal chromosomal microsatellite variability in Europe and North American populations

Chromosome	Population									
	Barcelona		Mount Parnes		Drøbak		Sunne		Bellingham	
	Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*	Allelic richness	*He*
Sex	18,43 ± 1,94	0,90 ± 0,01	18,72 ± 2,10	0,91 ± 0,01	13,84 ± 1,52	0,88 ± 0,01	13,24 ± 1,50	0,88 ± 0,01	4,94 ± 0,39	0,70 ± 0,03
Autosomal	15,93 ± 1,46	0,79 ± 0,05	15,65 ± 1,33	0,8 ± 0,04	11,45 ± 0,79	0,76 ± 0,04	10,31 ± 0,70	0,76 ± 0,04	5,73 ± 0,33	0,72 ± 0,03
Difference S:A	n.s	*	n.s	*	n.s	*	n.s	*	n.s	n.s

Mean values of variability ± Standard error. Mann-Whitney *U* test; ns $P>0,05$, * $P<0,05$

At the chromosomal polymorphism level there were also differences between populations (Figure 4A and Table 4). The differences in frequency of all chromosomes arrangements were higher between the Southern and the Northern populations of Europe than with Bellingham (Figure 4A). In fact, while the chromosomal polymorphism in the Northern Europe populations was almost limited to "standard" chromosomes; in contrast, the Southern populations and the colonizing one had a more balanced polymorphism for all chromosomes (Table 4).

5.3. Genetic content of the chromosomal arrangements

The genetic composition of the different chromosomal arrangements between the European populations was measured through genetic differentiation values (F_{ST}) between them. The genetic composition of all chromosomal arrangements of Bellingham was very different from the composition of chromosomal arrangements of Europe for all pairwise comparisons of populations due to the founder effect. Since our question was if there were differences in the genetic content of the chromosomal arrangements between populations possibly at mutation-drift equilibrium, Bellingham was excluded from this analysis to avoid any bias in the results.

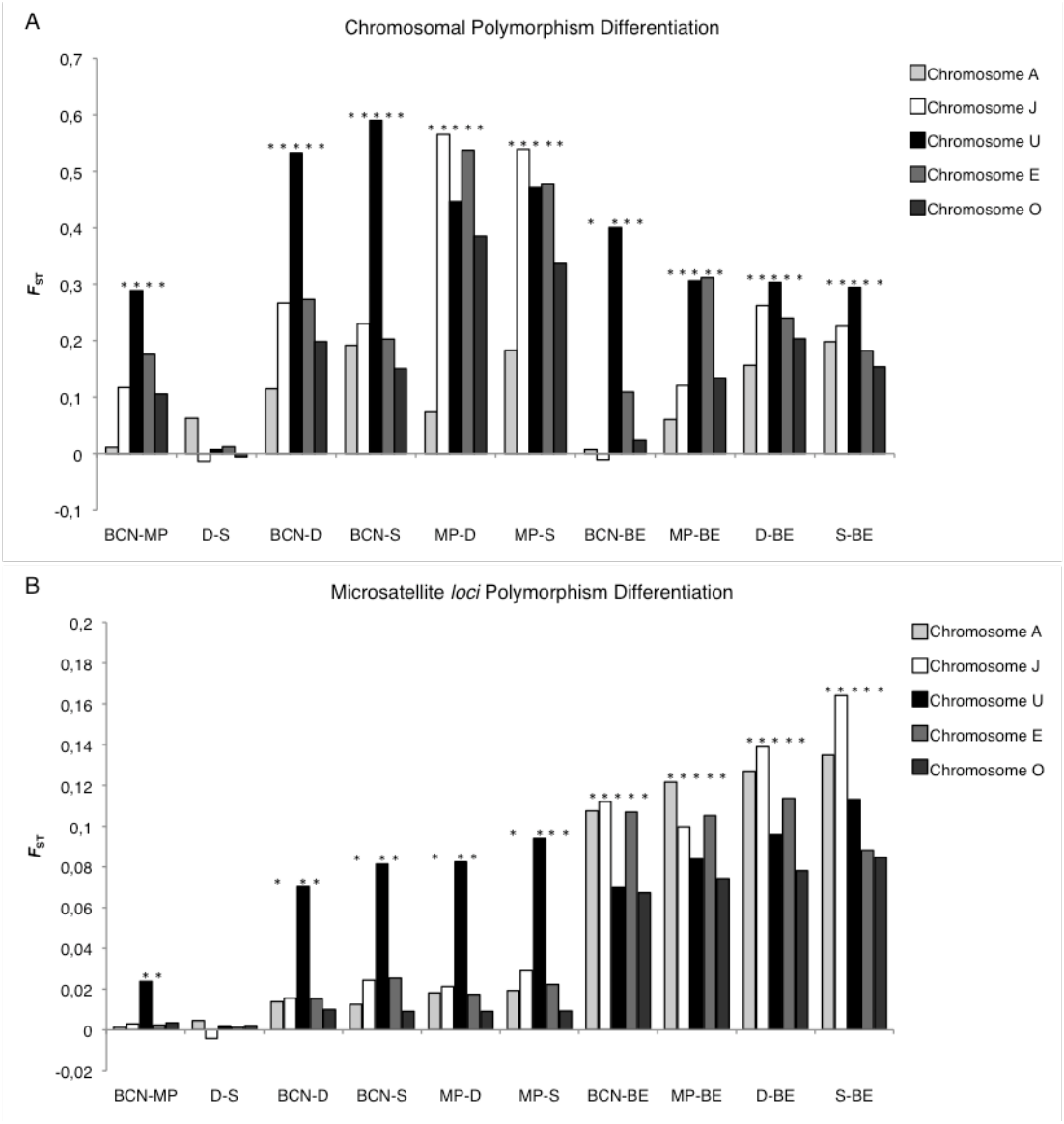

FIGURE 4.- Genetic differentiation (F_{ST}) of the chromosomal polymorphism (A) and microsatellite *loci* (B) for each chromosome by pair comparisons of populations. The populations are named as BCN (Barcelona), MP (Mount Parnes), D (Drøbak), S (Sunne) and B (Bellingham). The asterisks (*) represent the comparisons that were significantly different after Bonferroni correction

The analyses of genetic differentiation (F_{ST}) were done for all pairs of comparisons between arrangements of the same chromosome and grouped in threecategories: i) the same chromosome arrangement in different populations, ii) different chromosome arrangements in the same population, and iii) different

TABLE 4 – Chromosomal arrangement frequency in four European and one North American populations

Chromosome	Arrangement	Population				
		Barcelona	Mount Parnes	Drøbak	Sunne	Bellingham
A	A_{ST}	0,558	0,515	0,735	0,902	0,638
	A_1	0,105	0,233	0,235	0,078	-
	A_2	0,337	0,252	0,029	0,02	0,362
	N	95	103	68	51	94
J	J_{ST}	0,358	0,136	0,757	0,729	0,362
	J_1	0,642	0,864	0,243	0,271	0,638
	N	95	103	74	59	94
U	U_{ST}	0,074	-	0,789	0,839	0,394
	U_1	0,011	0,01	0,026	0,018	-
	U_{1+2}	0,758	0,291	0,184	0,089	0,085
	U_{1+2+6}	-	0,466	-	-	-
	U_{1+2+7}	-	0,107	-	-	-
	U_{1+2+8}	0,158	0,126		0,054	0,521
	N	95	103	76	56	94
E	E_{ST}	0,505	0,165	0,986	0,945	0,66
	E_8	0,011	0,126	-	-	-
	E_{1+2}	0,211	0,058	0,014	0,055	-
	E_{1+2+9}	0,116	0,505	-	-	0,064
	$E_{1+2+9+3}$	0,021	0,01	-	-	0,277
	$E_{1+2+9+12}$	0,137	0,136	-	-	-
	N	95	103	74	55	94
O	O_{ST}	0,263	0,107	0,703	0,655	0,245
	O_5	-	-	0,041	0,086	0,096
	O_6	-	0,01	0,176	0,138	-
	O_7	0,011	-	-	-	-
	O_{3+4}	0,263	0,592	0,041	0,069	0,234
	O_{3+4+1}	0,053	0,136	-	-	-
	O_{3+4+2}	0,032	-	-	-	0,223
	O_{3+4+6}	0,011	-	-	0,017	-
	O_{3+4+7}	0,116	0,01	-	-	0,043
	O_{3+4+8}	0,221	0,058	-	0,034	0,16
	O_{3+4+10}	-	0,029	-	-	-
	O_{3+4+13}	0,011	-	-	-	-
	O_{3+4+17}	-	0,01	-	-	-
	O_{3+4+22}	0,011	0,049	-	-	-
	$O_{3+4+23+2}$	0,011	-	0,041	-	-
	N	95	103	74	58	94

chromosome arrangements in different populations (Figure 5). These three categories were subdivided according to the microsatellite *loci* that mapped inside the inversion (white), outside the inversion (black) and all the *loci* localized in the chromosome (gray). The A chromosome presented low genetic differentiation in the three categories (Figure 5A) and no significant differentiation between the *loci* inside and outside the inversions was observed (Mann-Whitney U Test, $P > 0.05$). Chromosomes J, E and O had the same pattern of genetic differentiation (Figure 5B, 5D and 5E). There was very low genetic differentiation in the same chromosomal arrangement

along the different European populations, both in the *loci* inside and outside, and no significant differentiation between them (P >0.05). On the other hand, there was high genetic differentiation between the different inversions in the same population for the *loci* inside the arrangements of the J, E and O chromosomes. In contrast, the microsatellite *loci* outside the arrangements presented low differentiation and were significant different from the *loci* inside (P <0.05 for J and O chromosomes; P <0.001 for the E chromosome). The pattern of genetic composition in different arrangements for different populations was the same as before: high differentiation for the *loci* located inside the considered inversions, low differentiation for the *loci* outside, and significant differentiation between the two groups of *loci* (P <0.001 for J chromosome and E; P <0.01 for the O chromosome). So, for the J, E and O chromosomes the same chromosome arrangements in different populations were more similar between them than different arrangements in the same population for the microsatellite *loci* that mapped inside the arrangements. Nevertheless, the microsatellite *loci* that mapped outside the arrangements always presented low genetic differentiation. The U chromosome showed an opposite scenario (see Figure 5C): the *loci* inside the arrangement had high genetic differentiation between the same arrangement in different populations and these *loci* were significant different from the *loci* located outside the arrangements (P <0.001) which presented low differentiation. Moreover, in the U chromosome, the *loci* inside the different arrangements were genetic differentiated in the same population, however were not significantly different from the *loci* outside (P >0.05). As for the different arrangements in different populations, *loci* located inside the inversion presented significant differentiation and there was significant genetic differentiation between the *loci* inside and outside the inversions (P <0.001).

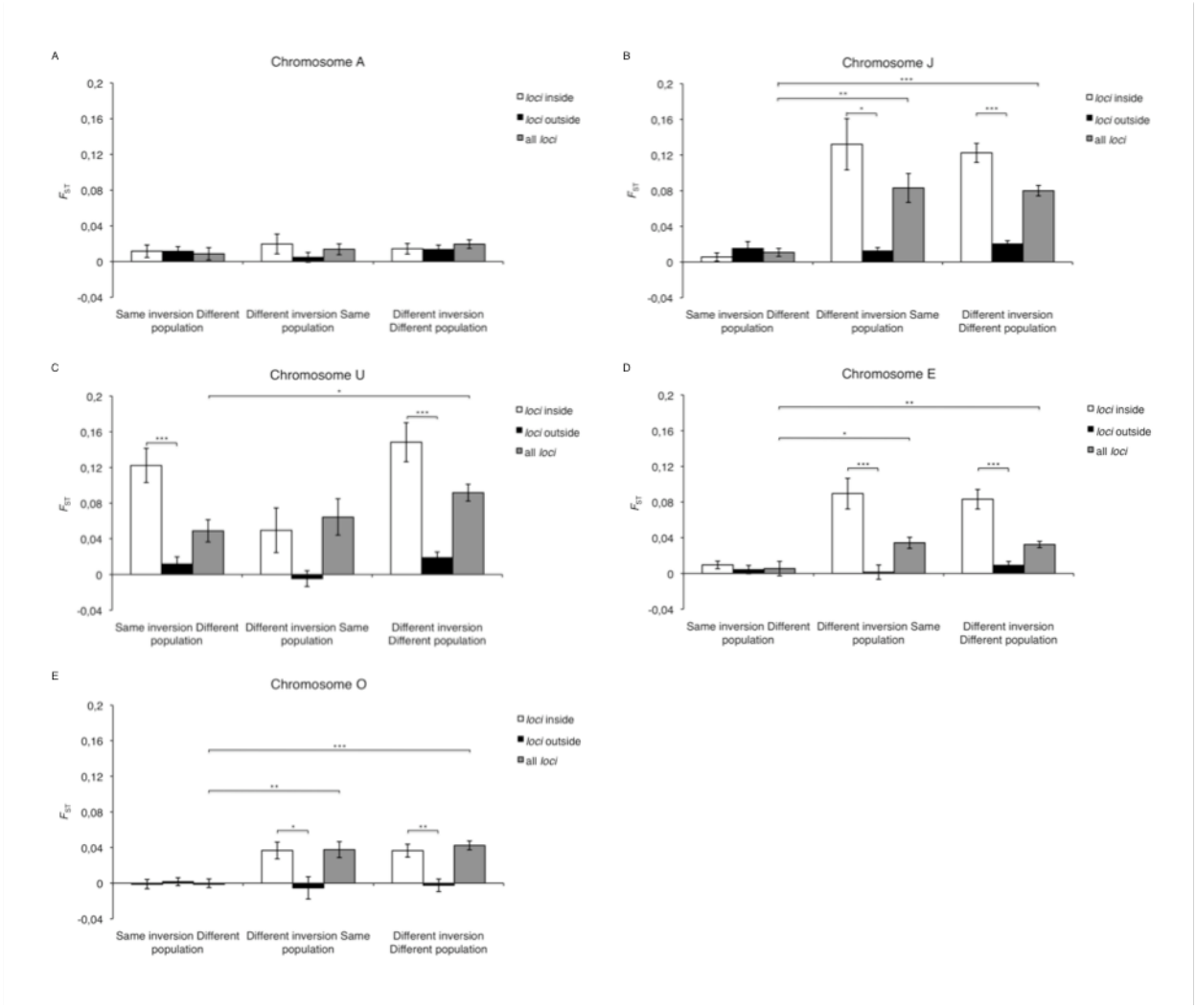

FIGURE 5 .- Genetic differentiation (F_{ST}) of the arrangements of each chromosome. The pairs of comparisons between arrangements of the same chromosome and grouped in three categories: i) the same chromosome arrangement in different populations, ii) different chromosome arrangements in the same population, and iii) different chromosome arrangements in different populations. These three categories were subdivided according to the microsatellite *loci* that mapped inside the inversion (white), outside the inversion (black) and all the *loci* localized in the chromosome (gray). The error bars represent the standard error. Asterisks (*) represent the significance difference between the *loci* inside and outside of each category, and the differentiation of all *loci* of each chromosome in the different categories (Mann-Whitney *U* test; * P<0,05, ** P<0,01, *** P<0,001).

5.4. Linkage Disequilibrium

Analyses of linkage disequilibrium between the neutral molecular markers and the chromosomal arrangements were conducted to increase the knowledge of the genetic content of the different chromosomal arrangements in the different

populations of *D.subobscura* and the evolutionary forces behind it. Non-random association between the microsatellite *loci* and chromosomal inversions was analyzed through the multiallelic version of Lewontin's *D'* – statistic (LEWONTIN 1964) (see Table 5). The general pattern for the five chromosomes was significant linkage disequilibrium (LD) for the markers localized inside the arrangements in the case of the Old World populations; and significant LD along all chromosomes for the New World population (Table 5 and Figure 2). For the A chromosome, Barcelona and Mount Parnes had significant LD in all markers located within the inversion A_2 with the exception of dsub21 in Barcelona (Fisher's Exact tests, $P < 0.05$; Table 5). As for the two Northern populations, no LD was detected for chromosome A in Sunne, probably due to the high frequency of Ast arrangement in that population (Table 4). In Drøbak, presenting also a significant proportion of A1 (Table 4) significant linkage disequilibrium (*D*'m) was observed for the *locus* most proximally located within this inversion (Table 5).

For chromosome J in the populations of Barcelona, Drøbak and Sunne, very significant *D'm* (Table 5) were observed for *loci* inside the inversion J_1 (dsub59 and dsub69), and non-significant LD for the *loci* outside the inversion, even in the ones nearby the inversion (dsub18 and dsub74). Nevertheless, no significant LD was observed in those *loci* for Mount Parnes, and only *locus* dsub27 presented significant LD although it is located towards the telomere and not within any inversion. Similarly, the marker located inside the arrangements U_{1+2} and U_{1+2+8} (dsub42) had significant LD with the chromosome U for all European populations. For chromosome E, *locus* dsub20, located inside the different arrangements, presented significant non-random association with the inversions (Table 5) in the South Europe

TABLE 5 - Linkage disequilibrium between microsatellite *loci* and chromosomes - *D*'m

Chromosome	Microsatellite *locus*	Population				
		Barcelona	Mount Parnes	Drøbak	Sunne	Bellingham
A	dsub11	0,210 n.s.	0,221 n.s.	**0,596** **	0,464 n.s.	**0,415** **
	dsub37	**0,588** **	0,494 n.s.	0,463 n.s.	0,976 n.s.	0,322 *
	dsub76	0,417 n.s.	0,424 n.s.	0,427 n.s.	0,706 n.s.	**1,000** ***
	dsub05	0,470 *	**0,414** **	0,368 n.s.	0,623 n.s.	**0,706** ***
	dsub21	0,324 n.s.	**0,449** ***	0,292 n.s.	0,372 n.s.	**0,983** ***
	dsub39	**0,595** **	**0,522** ***	0,459 n.s.	0,637 n.s.	**0,568** ***
	dsub70	0,422 n.s.	0,470 n.s.	0,564 *	0,701 n.s.	**0,983** ***
	dsub19	0,333 n.s.	0,290 n.s.	0,491 n.s.	0,514 n.s.	0,189 n.s.
J	dsub18	0,395 n.s.	0,488 n.s.	0,296 n.s.	0,353 n.s.	**0,854** ***
	dsub59	**0,749** **	0,942 n.s.	**0,724** ***	**0,631** ***	**0,921** ***
	dsub69	**0,568** ***	0,414 n.s.	**0,908** ***	**0,780** ***	**0,650** ***
	dsub74	0,384 n.s.	0,259 n.s.	0,364 n.s.	0,520 n.s.	**0,891** ***
	dsub62	0,440 n.s.	0,569 n.s.	0,399 n.s.	0,538 n.s.	0,104 n.s.
	dsub27	0,375 n.s.	0,528 *	- -	- -	0,109 n.s.
U	dsub10	0,491 n.s.	0,491 n.s.	0,418 n.s.	0,470 n.s.	**0,733** ***
	dsub03	0,267 n.s.	0,306 n.s.	0,486 n.s.	0,450 n.s.	**0,617** ***
	dsub42	0,381 *	**0,816** ***	**0,434** ***	**0,700** ***	**0,775** ***
	dsub64	0,460 n.s.	0,450 n.s.	0,463 n.s.	0,683 *	**0,761** ***
	dsub15	0,423 n.s.	0,318 *	0,361 n.s.	0,309 n.s.	**0,542** ***
E	dsub46	0,281 n.s.	0,321 n.s.	1,000 n.s.	0,685 n.s.	**0,645** ***
	dsub68	0,297 n.s.	0,324 n.s.	1,000 n.s.	0,631 n.s.	**0,716** ***
	dsub79	0,233 n.s.	**0,412** **	1,000 n.s.	1,000 n.s.	**0,866** **
	dsub20	**0,703** ***	**0,716** ***	1,000 n.s.	0,901 n.s.	**0,669** ***
	dsub53	0,335 n.s.	0,396 n.s.	1,000 n.s.	0,599 n.s.	**0,578** ***
	dsub28	0,403 n.s.	0,265 n.s.	1,000 n.s.	0,591 n.s.	0,072 n.s.
O	dsub02	0,487 *	0,463 n.s.	0,406 n.s.	0,402 n.s.	**0,291** **
	dsub14	0,504 n.s.	0,441 *	0,462 n.s.	0,366 n.s.	0,421 n.s.
	dsub38	0,480 n.s.	0,394 n.s.	0,367 n.s.	0,233 n.s.	**0,544** ***
	dsub01	0,406 n.s.	0,383 n.s.	**0,676** ***	**0,663** ***	**0,765** ***
	dsub04	**0,551** **	0,490 n.s.	0,426 n.s.	0,457 n.s.	**0,534** ***
	dsub12	0,613 n.s.	0,462 n.s.	0,571 n.s.	0,487 n.s.	**0,641** ***

Fisher's Exact test * P<0,05, ** P<0,01, *** P<0,001. The significant values after the Bonferroni are in bold.

populations. The pattern of LD for O chromosome was not as clear as described above for the other chromosomes probably because this chromosome has multiple arrangements (Table 4). The two Southern populations had significant LD (P <0.05; Table 5) in markers located inside the inversion O_7, dsub02 for Barcelona and dsub14 for Mount Parnes. The Northern had strong LD on dsub01 that is positioned inside the inversion O_6 (P <0.001; Table 5)

The former analysis gave a general picture of LD patterns between the microsatellite markers and the chromosomes. Nonetheless, the specific chromosomal arrangements that were in LD with those markers, and the microsatellite alleles

involved on it remained unanswered. For that, it was applied an interallelic disequilibrium measure between microsatellite alleles and chromosomal arrangements (Table 6).

TABLE 6 - Interallelic disequilibrium (*D'*) between microsatellite alleles and chromosomal arrangements

Chromosome	Arrangement	Microsatellite *locus*	Population				
			Barcelona	Mount Parnes	Drøbak	Sunne	Bellingham
A	A_{ST}	dsub11			242		242
		dsub76					306 / 308
		dsub05	151				160
		dsub21					181
		dsub39					282
		dsub70					311/ 315
	A_1	dsub11			246		
		dsub37			218		
		dsub76			306		
		dsub21		169			
		dsub39		276			
		dsub70		311			
		dsub19		193			
	A_2	dsub37	220 / 242				216
		dsub76		310 / 320			303/ 318 / 320
		dsub05	152	156			145 / 166
		dsub21					179 / 185 / 195
		dsub39	277	277			275
		dsub70		309			301 / 305
		dsub19		201			
J	J_{ST}	dsub18		212			210
		dsub59	245	245	245	245	245 / 251
		dsub69	143 / 150		143	143	147
		dsub74					160 / 166
	J_1	dsub18			218		208 / 212
		dsub59			262	285	260 / 266
		dsub69	141		145 / 147	145 / 147	149 / 151
		dsub74					151 / 162 / 164
		dsub62			297		
		dsub27					
U	U_{ST}	dsub10					222 / 283
		dsub03	147				145
		dsub42				125	109 / 120
		dsub64					161 / 188
		dsub15					241 / 253
	U_{1+2}	dsub10		237			
		dsub42		115 / 122 / 128	115		
	U_{1+2+8}	dsub10		226			208 / 212 / 220
		dsub03		147			158
		dsub42	130	124 / 126			126
		dsub64	153				163 / 164
		dsub15	225				247 / 251

Note: Only microsatellite alleles (in bp) presenting significant *D'* coefficients with inversions are shown ($\chi 2$ Yates correction, $P<0.05$).

TABLE 6 - Continued

Chromosome	Arrangement	Microsatellite *locus*	Population				
			Barcelona	Mount Parnes	Drøbak	Sunne	Bellingham
E	E_{ST}	dsub46					327 / 339
		dsub68					214 / 221
		dsub79					148
		dsub20	254 / 255	251 / 255			243 / 248 / 255
		dsub53					297
		dsub28					
	E_{1+2}	dsub20	243	243			
	E_{1+2+9}	dsub79		149			
		dsub20					257
		dsub28	177				
	$E_{1+2+9+3}$	dsub46					335
		dsub68					223
		dsub79					149
		dsub20					269
		dsub53					299
	$E_{1+2+9+12}$	dsub79		151			
		dsub20		277			
	E_{8}	dsub20		254			
O	O_{ST}	dsub02		222			
		dsub01		263	263	263	269
		dsub04	196 / 198	196 / 198			198
		dsub12	252				252
	O_{3+4}	dsub02	230				
		dsub38	413				
		dsub01					261 / 277
		dsub04					197
		dsub12					267 / 271
	O_{3+4+1}	dsub14		123			
		dsub38		405			
	O_{3+4+8}	dsub01					265
		dsub04					194 / 242
		dsub12					262 / 268
	O_{3+4+7}	dsub02	234				252
		dsub38	411				413
		dsub01					261
		dsub12	268				
	O_{3+4+2}	dsub02					231
		dsub38					424
		dsub01					259
		dsub04					196
		dsub12					269
	O_{6}	dsub38					424
		dsub01			261 / 265	196	
		dsub04			196		

Note: Only microsatellite alleles (in bp) presenting significant *D'* coefficients with inversions are shown ($\chi 2$ Yates correction, $P<0.05$).

The Southern European populations had significant LD between inversion A_2 and allele 277 (bp) of *locus* dsub39 (*D'* =1 in Barcelona and *D'* =0.65 in Mount Parnes, with both *P* <0,001) that is located near its distal breakpoint (Figure 2). In the

proximal breakpoint of inversion A_2 there was also LD in Barcelona and Mount Parnes although different alleles were implicated (*locus* dsub05; see Table 6). Interestingly, in Drøbak all *loci* inside inversion A_1 presented significant LD with this arrangement (Table 6), which had more than 20% frequency in the population (Table 4). Bellingham presented significant LD for almost all *loci* analyzed with the arrangements A_{ST} and A_1 sometimes with several alleles of the same *locus* in disequilibrium.

All populations presented significant LD patterns between J_{ST} and two *loci* (dsub59 and dsub69) located within the J_1 inversion (Table 6). Remarkably, the microsatellite alleles in disequilibrium with J_{ST} were the same across populations for *locus* dsub59 (allele 245 bp), and along Barcelona, Drøbak and Sunne for *locus* dsub69 (allele 143 bp). In the same way, E and O chromosomes presented significant LD values between markers located inside the inversions and their "standard" chromosomes for the same allele on different populations (Table 6). Barcelona, Mount Parnes and Bellingham had allele 255bp of dsub20 in LD with E_{ST}; Barcelona and Mount Parnes had alleles 196 and 198bp of dsub04 (mapped inside the arrangement O_{3+4}) in LD with O_{ST}; and finally, Mount Parnes and the two Northern populations had allele 263bp of dsub01 (mapped inside the inversion O_6) in LD with O_{ST}. Like the other chromosomes, U had higher levels of LD in a marker located inside the arrangements (dsub42), however the alleles presenting significant LD with different arrangements differed between populations (Table 6).

It is curious to note that the alleles responsible for the LD patterns with all the chromosome's inversions (except for chromosome J) in Europe and the North American populations were, in the majority of cases, not the same ones. Furthermore,

for that colonizing population the association between alleles and inversions was almost complete for the majority of the comparisons, depending on the chromosome.

6. DISCUSSION

The native Palearctic fly *Drosophila subobscura* spans more than 30º latitude in the Old World: from North Africa to Scandinavia, where well-documented, stable latitudinal clines in inversion frequencies have been described (KRIMBAS and POWELL 1992). Shortly after the two non-independent colonization events of North and South America (PASCUAL *et al.* 2007) the species developed parallel latitudinal clines reinforcing the idea of an adaptive role of the chromosomal inversion polymorphism on the establishment of the clines (AYALA *et al.* 1989; PREVOSTI *et al.* 1988). Furthermore, it has been shown that changes in frequency of these polymorphisms are responding to global climate warming (BALANYA *et al.* 2006). However, the selective mechanisms involved in the maintenance of chromosomal polymorphisms are still a matter of dispute, being the two main hypothesis the coadaptation *sensu* Dobzhansky (DOBZHANSKY 1950) and the local adaptation by Kirkpatrick and Barton (KIRKPATRICK and BARTON 2006).

6.1 Patterns of Genetic Variability and Differentiation

The European populations differ in their microsatellite *loci* variability according to their latitudinal localization (Figure 3). The Southern European populations were more variable in their allelic richness than the Northern European ones, but did not differ in their expected heterozigosity. This difference between allelic richness and *He* is a classic pattern of a population bottleneck (NEI *et al.* 1975). After a bottleneck it takes more time to a population to recover their ancestral number of alleles than their number of heterozigotes for a neutral marker. Both parameters

depend on the severity of the bottleneck and the subsequent intrinsic rate of growth, but the number of alleles depends on the mutation rate to generate new alleles. The decrease of variability towards North was probably due to a postglaciar expansion, where the Southern peninsulas acted as "variability refuges" (HEWITT 2000). Furthermore, the smaller effective population size in the Northern localities, due to their marginal location relative to the distribution area of the species, can also contribute to this reduced genetic variability. Significant molecular genetic differentiation between the Southern and Northern European populations was found for almost all chromosomes, being particularly high for chromosome U (Figure 4B). This result could be a byproduct of the genetic composition of the chromosomal arrangements in different populations (see below). At the inversion polymorphism level, the Northern European populations consistently presented lower level of variability when compared to the Southern ones, particularly for *He* (Figure 3A and 3C). Furthermore, the frequencies of arrangements between South and North of Europe were very distinct (Figure 4A). This is caused by the well-defined cline of inversion frequencies in Europe (KRIMBAS and POWELL 1992).

Patterns of molecular genetic variability in the Bellingham population clearly show the effects of the colonization event of the North American continent (Figure 3). Due to this event, both allelic richness and *He* were reduced relative to the ancestral populations at the microsatellite level but only reduction in allelic richness for inversions. Despite Bellingham being localized in the extreme north of the cline in North America, it did not present the same pattern of reduction in chromosomal variability as the North European populations. This is explained by the fact that the North America cline is not so steep as the European one (BALANYA *et al.* 2003). Bellingham presents a low number of chromosomal arrangements with intermediate

frequencies (low allelic richness and high *He*, respectively Figure 3A and 3C), and as a result was less different in the arrangement frequency from the Southern European populations than were the Northern European populations (Figure 4A). Furthermore, this colonizing population presented significant genetic differentiation at the microsatellite level relative to all other ancestral populations, as a consequence of the founder event associated with colonization of North America (Table 2 and Figure 4B), in agreement with previous studies (PASCUAL *et al.* 2001, 2007). Interestingly, the genetic differentiation of Bellingham was larger with the Northern than the Southern European populations both at the microsatellite and inversion level. This is in agreement with the previously suggested Mediterranean origin of the colonizers (BRNCIC *et al.* 1981; PASCUAL *et al.* 2007). Furthermore the higher similarity with the Barcelona population points to a more probable western Mediterranean origin of the colonizers (Figure 4).

The lower genetic variability for the sex-linked *loci* relative to autosomal ones in Bellingham (Table 3) are also a consequence of the bottleneck. The sex chromosome has a smaller effective population size than the autosomes (males have only one copy; N_{eX}=3/4 N_{eA}), and consequently after a bottleneck it takes more time to recover its genetic variability. However, the ancestral populations were more variable in the sex-linked *loci*, which is consistent with the background selection model (CHARLESWORTH *et al.* 1993), in which the neutral markers linked to the deleterious alleles are purged from the population. In a heterogametic system recessive deleterious mutations are removed more efficiently from the sex chromosome, because there is less chance for recombination to combine different neutral alleles with the deleterious mutation. Thus, as expected under the background selection

model, the ancestral populations have more neutral variation on the sex chromosome, as also observed in other *Drosophila* species (e.g. (KAUER *et al.* 2002).

6.2. Genetic Composition of Chromosomal Arrangements: Different Histories for Different Chromosomes

Two main hypotheses were proposed to address the selective mechanisms underlying inversion polymorphisms: the coadaptation *sensu* Dobzhansky (DOBZHANSKY 1950) and the local adaptation model (KIRKPATRICK and BARTON 2006). The coadaptation model has two outputs about genetic variation on the chromosomal arrangements: genetic differentiation among different chromosomal arrangements and genetic differentiation within a chromosomal arrangement among populations. Instead, the local adaptation model does not have chromosomal population-specific properties. So, the genetic content would remain constant in the same arrangement across populations. We tested these predictions in our dataset.

Unexpectedly, we did not observe a general pattern of genetic composition of the arrangements in all chromosomes that could be explained by one of these hypotheses. In fact, our data revealed distinct evolutionary mechanisms for the patterns of European chromosomes in respect to the maintenance of their polymorphisms. This observation in itself contradicts previous indications that the O chromosome is representative of all chromosomes in *D. subobscura* (SANTOS 2007).

Chromosome A was the only chromosome that did not present a pattern of arrangements genetic composition predicted by any of the aforementioned hypothesis (Figure 5A). The different arrangements were not significantly different from each

other in the same or in different populations, independently of the *loci* being inside or outside the inversion. Different non-exclusive processes can cause the lack of genetic differentiation between arrangements. First, this chromosome had high variability on the ancestral populations (Table 2) probably due to background selection, as discussed above. Consequently, this high variability might generate homoplasy in the allele distribution and thus obscure possible genetic differences. Alternatively, genetic exchange between arrangements can contribute to reduce genetic differentiation as reported between the A_2/A_{ST} arrangements (NOBREGA *et al.* 2008). Despite the genetic exchange found by Nóbrega and collaborators they also detected strong genetic differentiation between arrangements in markers inside the A_2 inversion. However, this observation does not contradict our results in the sense that this study used a colonizing population and isolated from the continent, from the island of Madeira (KHADEM *et al.* 1998)). Similarly, in our data there was high genetic differentiation between A_2/A_{ST} in the colonizing American population (Bellingham, F_{ST}=0,33), that in a certain way could be compared with the result mentioned above: the bottleneck effect may possibly increase the genetic differentiation between arrangements.

In spite of the arrangements of chromosome A did not differ genetically between them as well as the *loci* inside and outside the inversions, the LD patterns relative to the inversions presented significant differences across *loci* (Table 5 and 6). Exclusively the markers inside the inversions presented non-random association with chromosome A (Table 5). Markers located outside of the inversion but near its breakpoints (dsub76 and dsub70) did not present significant LD, making the inversion's origin as the cause for the LD ("founding" disequilibrium, see below) less likely. The arrangements also differed on their associations with the microsatellite

alleles (*D'*, Table 6). The Southern European populations had clear indications of LD in inversion A_2, especially in Mount Parnes. The same allele of *locus* dsub39, which is located near its distal breakpoint, presented LD with A_2 on both Southern populations. The *locus* located near the proximal breakpoint of inversion A_2 also presented LD in Barcelona and Mont Parnes although different alleles are involved. Moreover, a *locus* located in the center of the A_2 inversion (dsub21) also presents significant LD in Mount Parnes. The inversion A_2 is typical of the Southern populations, so the strong LD observed could be caused by selection acting on genes nearby those *loci*, especially in the distal region. Inversion A_1 is common in Northern latitudes and their frequencies increased in North Atlantic European populations (BALANYA *et al.* 2004). Therefore, the strong LD observed in Drøbak between inversion A_1 and the *loci* localized inside it might be a selective sweep on inversion A_1 in the northern populations.

The patterns presented by chromosomes J, E and O were in accordance with the local adaptation hypothesis (see Figure 5B, 5D and 5E). The genetic content of each chromosomal arrangement remained constant at a wide range of geographical distribution and for all *loci* whether inside or outside the inversions. Previous studies also observed no evidence for geographic differentiation for a given chromosome O arrangement (ROZAS and AGUADÉ 1990; ROZAS *et al.* 1995, 1999). Moreover, in the same population *loci* located inside the inversion were highly differentiated between arrangements. Therefore the gene arrangements had their own "identity" across Europe. A neutral hypothesis might have the same pattern of non-differentiation of the same arrangement, in the sense that each inversion had a monophyletic origin (POWELL 1997) and will originate gene-inversion disequilibria for all polymorphic *loci* included in the inversion or near it ("founding" disequilibrium). Nevertheless,

there is evidence for genetic exchange (gene flux) between different chromosomal arrangements through gene conversion and/or double crossovers (MUNTE *et al.* 2005; NOBREGA *et al.* 2008; ROZAS and AGUADÉ 1990,1994; ROZAS *et al.* 1999), and gene flow between populations (PASCUAL *et al.* 2001). Both gene flux and gene flow were supported by our data. In the first case, there was no genetic differentiation in the *loci* located outside the inversion between different arrangements in the same population. Secondly, in the same line of reasoning, the *loci* outside the inversions were not genetically differentiated between different populations whether it was in the same or in different arrangements (Figure 5). In spite of all gene flux and flow the chromosomal arrangements "identity" was kept. So, strong selection must be invoked in order to counterbalance these aforementioned homogenizing processes and thus contributing to the maintenance of the existing latitudinal clines in these chromosomal inversions

The LD analyses for chromosomes J, E and O were in accordance with the local adaptation scenario. Non-random association between chromosomal arrangements and microsatellites was almost exclusive for the *loci* inside the inversions, whether those were near the breakpoint or in the center of the inversion. In general, the *loci* localized outside the inversions did not have significant LD with the arrangements, even those near the inversion breakpoint – the most striking being dsub74 on chromosome J – ruling out, once again, the "founding" disequilibrium effect. The same LD pattern across populations, that is involving the same microsatellite alleles with the same chromosome arrangement such as for *locus* dsub59 and the J_{ST} inversion (Table 6), suggests that the genetic composition of the arrangements was maintained by selection. As (LEWONTIN 1974) said: "*The observation of significant linkage disequilibrium that is consistent between*

populations is a very sensitive detector of natural selection". Therefore, our hypothesis that selection through a local adaptation process has been holding the genetic content of the chromosomal arrangements of J, E and O became more robust. Remarkably, the classical association studies between allozymes and inversions in *D.subobscura* made on the 70's also presented a consistency of the LD patterns in different populations for chromosomes E and O (LOUKAS and KRIMBAS 1975; LOUKAS *et al.* 1979; ZOUROS and KRIMBAS 1973); it is interesting that some of this allozymes mapped nearby the microsatellite *loci* that were in LD: *Est-9* is close to dsub20 on chromosome E and, *Lap* to dsub01 and *Acph* to dsub04 on chromosome O.

In fact, it seems that the local adaptation mechanism is widespread through the genus *Drosophila*: *D.pseudoobscura* and *D.persimilis* (PRAKASH and LEWONTIN 1968, 1971), *D.pseudoobscura* (SCHAEFFER et al. 2003) and *D.melanogaster* (KENNINGTON *et al.* 2006) presented similar genetic composition for the same arrangement irrespective of the geographic localization. Nonetheless, these studies claimed evidence for coadaptation since they observed epistatic selection within inversions (on the statistical sense, view (PHILLIPS 2008)) and genetic differentiation between different arrangements. Yet, epistatic selection and genetic differentiation between gene arrangements do not rule out the local adaptation mechanism (KIRKPATRICK and BARTON 2006; SANTOS 2009).

The genetic pattern of European U chromosome was distinct from the other chromosomes (Figure 5C). The genetic composition inside the inversion of the same arrangement was largely dissimilar across Europe populations. The arrangements of chromosome U had population-specific properties (genetic compositions) in the Old World. Even with genetic exchange and migration, as before, detected by no differentiation in the *loci* outside the inversions. Therefore, strong selection pressure

inside the inversions has been maintaining their population-specific gene content in chromosome U, which is in agreement with the coadaptation hypothesis *sensu* Dobzhansky (DOBZHANSKY 1950). The different chromosomal arrangements also were distinct to each other in the same population (Figure 5C). However, the markers inside the inversion were not significantly different from the outside *loci*, because the variance for the *loci* inside was high as a result of the different comparisons. Linkage disequilibrium analyzes were also in accordance with the coadaptation model. The microsatellite *locus* in LD with the chromosome U (Table 5) was located inside the inversion (dsub42).

6.3. Genetic Composition of Chromosomal Arrangements and the Founder Effect

The Bellingham population experienced a huge bottleneck thirty years ago at the time of the colonization of North America (PASCUAL *et al.* 2007). As discussed above this has lead to major changes at the molecular genetic level in comparison to the European populations. Few years after the colonization parallel latitudinal clines to those in Europe were developed (PREVOSTI *et al.* 1988). We analyzed the LD in Bellingham and compared with the ancestral populations (Table 5 and 6). The founder event associated with the colonization of North America caused a "demographic" disequilibrium. Bellingham presented extremely high levels of LD for almost all *loci* analyzed, sometimes with several alleles of the same *locus* in disequilibrium with the inversion. In the majority of cases, the alleles presenting LD with the arrangements in Europe and North America were not the same. This dissimilarity could be attributed to the bottleneck effect, since the LD in the assayed putatively neutral markers in Europe was not complete. This discrepancy in patterns raises some questions on the formation of the European *vs.* North American clinal patterns: Is the genetic

composition of chromosomal arrangements in America different from Europe? Does the discrepancy in neutral markers reflect differences in the relevant adaptive genes determining the clinal pattern of inversions in the two continents? And finally, are the processes behind the clines different between the continents? To tackle this set of questions we should have other North American populations sampled at different latitudes. Nonetheless, according to (KIRKPATRICK and BARTON 2006) model it is only required two *loci* to the local adaptation scenario to establish. Then, in this recent population the *loci* that are promoting the adaptive value of the inversion could be few and undetectable at this scale since LD due to the founder effect in most *loci* could prevent the detection of selection.

7. CONCLUDING REMARKS

We will now briefly highlight some of the most relevant points of this study:

- Southern European *D. subobscura* populations present larger genetic variability than Northern ones in both microsatellite and chromosomal inversion markers.

- Genetic differentiation between northern and southern populations is higher for chromosomal arrangements than for microsatellites due to the selection acting on the former, generating the latitudinal clines in Europe.

- Linkage disequilibrium patterns in molecular markers within inversions reveal selection acting on some chromosomal arrangements in the European populations.

- Our results provide strong support for distinct selective mechanisms on the maintenance of inversion polymorphisms, namely: local adaptation hypothesis on chromosome J, E and O, and coadaptation *sensu* Dobzhansky on U chromosome.

- Linkage disequilibrium patterns in molecular markers in colonizing populations spans along the chromosomes due to the founder effect.

- Different alleles are responsible for the LD patterns in the colonizing *vs*. de ancestral populations due to the bottleneck during the colonization process.

- Chromosomal inversions are an important selective feature of *Drosophila* genomes and should be included in the studies of adaptation processes, since sets of genes could be maintained together an act as a constrainer or enhancer to the process.

8. REFERENCES

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.

BALANYA, 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.

BALANYA, 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.

BALANYA, J., E. SOLE, J. OLLER, D. SPERLICH and L. SERRA, 2004 Long-term changes in the chromosomal inversion polymorphism of *Drosophila subobscura*. II. European populations. J. Zool. Syst. Evol. Res. **42:** 191-201.

BRNCIC, D., A. PREVOSTI, M. BUDNIK, M. MONCLÚS and J. OCAÑA, 1981 Colonization of *D. subobscura* in Chile I. First population and cytogenetic studies. Genetica **56:** 3–9.

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.

DOBZHANSKY, T., 1943 Genetics of natural populations. IX. Temporal changes in the composition of populations of *Drosophila 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.

DOBZHANSKY, T., 1970 *Genetics of the Evolutionary Process*. Columbia Univ. Press, New York.

DOBZHANSKY, T., and C. EPLING, 1948 The supression of crossing over in inversion heterozygotes of *Drosophila pseudoobscura*. Genetics **34:** 137-141.

FINNEGAN, D. J., 1989 Eukaryotic transposable elements and genome evolution. Trends Genet. **5:** 103-107.

FONTDEVILA, A., C. ZAPATA, G. ALVAREZ, L. SANCHEZ, J. MENDEZ *et al.*, 1983 Genetic coadaptation in the chromosomal polymorphism of *Drosophila subobscura*. I. Seasonal changes of gametic disequilibrium in a natural population. Genetics **105:** 935-955.

GAUNT, T. R., S. RODRIGUEZ, C. ZAPATA and I. N. DAY, 2006 MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. BMC Bioinformatics **7:** 227.

GOUDET, J., 2001 FSTAT, a program to estimate and test gene diversities and fixation indices, http:/www.unil.ch/izea/softwares/fstat.html.

GOULD, S. J., and R. C. LEWONTIN, 1979 Spandrels of San-Marco and the panglossian parafigm - A critique of the adaptationist program. Proc. R. Soc. Lond. Ser. B **205:** 581-598.

GREGORY, T. R., 2004 *The Evolution of the Genome*. Elsevier Academic Press.

GRIFFITHS, A., S. WESSLER, R. C. LEWONTIN, W. GELBART and D. SUZUKI, 2004 *Introduction to Genetic Analysis*. W. H. Freeman.

HAHN, M. W., and M. RAUSHER, 2008 Toward a selection theory of molecular evolution. Evolution **62:** 255-265.

HEWITT, G. M., 2000 The genetic legacy of the Quaternary ice ages. Nature **405:** 907-913.

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? Annu. Rev. Ecol. Evol. Syst. **39:** 21-42.

HOFFMANN, A. A., C. M. SGRO and A. R. WEEKS, 2004 Chromosomal inversion polymorphisms and adaptation. Trends Ecol. Evol. **19:** 482-488.

KAUER, M., B. ZANGERL, D. DIERINGER and C. SCHLOTTERER, 2002 Chromosomal patterns of microsatellite variability contrast sharply in African and Non-African populations of *Drosophila melanogaste*r. Genetics **160:** 247–256.

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 A. BREHM, 1998 Tracing the colonization of Madeira and the Canary Islands by *Drosophila subobscura* through the study of the *rp49*gene region. J. Evol. Biol. **11:** 439-452.

KIRKPATRICK, M., and N. BARTON, 2006 Chromosome inversions, local adaptation and speciation. Genetics **173:** 419-434.

KOSKE, T., and J. MAYNARD SMITH, 1954 Genetics and cytology of *Drosophila subobscura*. X. The fith linkage group. J. Gen. **52:** 521-541.

KRIMBAS, C. B., and J. R. POWELL, 1992 *Drosophila Inversion Polymorphism*. CRC Press.

KUNZE-MÜHL, E., and E. MÜLLER, 1958 Weitere Untersuchungen über die chromosomale Struktur und die naturlichen Strukturtypen von Drosophila subobscura coll. Chromosoma **9:** 559-570.

LEWONTIN, R. C., 1964 The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49:** 49-67.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York

LIU, K., and S. V. MUSE, 2005 PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics **21:** 2128-2129.

LOUKAS, M., and C. B. KRIMBAS, 1975 The genetics of *Drosophila subobscura* populations. V. A study of linkage disequilibrium in natural populations between genes and inversions of the E chromosome. Genetics **80:** 331-347.

LOUKAS, M., C. B. KRIMBAS and Y. VERGINI, 1979 The genetics of *Drosophila subobscura* populations. IX. Studies on linkage disequilibrium in four natural populations. Genetics **93:** 497-523.

MAINX, F., T. KOSKE and E. SMITAL, 1953 Untersuchungen über die chromosomale Struktur europāischer Vertreter der *Drosophila obscura* Gruppe

. Zeitschrift für Inducktive Abstammungs- und Vererbungslehre **85:** 354–372.

MUNTE, A., J. ROZAS, M. AGUADE and C. SEGARRA, 2005 Chromosomal inversion polymorphism leads to extensive genetic structure: a multilocus survey in *Drosophila subobscura*. Genetics **169:** 1573-1581.

NEI, M., T. MARUYAMA and R. CHAKRABORTY, 1975 The bottleneck effect and genetic variability in populations. Evolution **29:** 1-10.

NOBREGA, C., M. KHADEM, M. AGUADE and C. SEGARRA, 2008 Genetic exchange versus genetic differentiation in a medium-sized inversion of *Drosophila*: the

A2/Ast arrangements of *Drosophila subobscura*. Mol. Biol. Evol. **25:** 1534-1543.

ORR, H. A., 2005 The genetic theory of adaptation: a brief history. Nat. Rev. Genet. **6:** 119-127.

PASCUAL, M., C. F. AQUADRO, V. SOTO and L. SERRA, 2001 Microsatellite variation in colonizing and Palearctic populations of *Drosophila subobscura*. Mol. Biol. Evol. **18:** 731-740.

PASCUAL, M., M. P. CHAPUIS, F. MESTRES, J. BALANYA, R. B. HUEY *et al.*, 2007 Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. Mol. Ecol. **16:** 3069-3083.

PASCUAL, M., M. D. SCHUG and C. F. AQUADRO, 2000 High density of long dinucleotide microsatellites in *Drosophila subobscura*. Mol Biol Evol **17:** 1259-1267.

PHILLIPS, P. C., 2008 Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet **9:** 855-867.

POWELL, J. R., 1997 *Progress and Prospects in Evolutionary Biology : The Drosophila Model*. Oxford University Press, New York.

PRAKASH, S., and R. C. LEWONTIN, 1968 A molecular approach to the study of genic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements of *Drosophila*. Proc. Natl. Acad. Sci. U S A **59:** 398-405.

PRAKASH, S., and R. C. LEWONTIN, 1971 A molecular approach to the study of genic heterozygosity in natural populations. V. Further direct evidence of coadaptation in inversions of *Drosophila*. Genetics **69:** 405-408.

PREVOSTI, A., G. RIBO, L. SERRA, M. AGUADE, J. BALANA *et al.*, 1988 Colonization of America by *Drosophila subobscura*: experiment in natural populations that

supports the adaptive role of chromosomal-inversion polymorphism. Proc. Natl. Acad. Sci. U S A **85:** 5597-5600.

RANZ, J. M., D. MAURIN, Y. S. CHAN, M. VON GROTTHUSS, L. W. HILLIER *et al.*, 2007 Principles of genome evolution in the *Drosophila melanogaster* species group. Plos Biology **5:** 1366-1381.

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, 1990 Evidence of extensive genetic exchange in the *rp49* region among polymorphic chromosome inversions in *Drosophila subobscura*. Genetics **126:** 417-426.

ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. Proc. Natl. Acad. Sci. U S A **91:** 11517-11521.

ROZAS, J., C. SEGARRA, G. RIBO and M. AGUADE, 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. AGUADÉ, 1995 Nucleotide Polymorphism at the *rp49* region of *Drosophila subobscura*: lack of geographic subdivision within chromosomal arrangements in Europe. J. Evol. Biol. **8:** 355-367.

SANTOS, M., 2007 Evolution of total net fitness in thermal lines: *Drosophila subobscura* likes it 'warm'. J. Evol. Biol. **20:** 2361-2370.

SANTOS, M., 2009 Recombination Load in a Chromosomal Inversion Polymorphism of *Drosophila subobscura*. Genetics **181:** 803-809.

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. Proc. Natl. Acad. Sci. U S A **100:** 8319-8324.

TRAVISANO, M., J. A. MONGOLD, A. F. BENNETT and R. E. LENSKI, 1995 Experimental tests of the roles of adaptation, chance, and history in evolution. Science **267:** 87-90.

WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358–1370.

ZOUROS, E., and C. B. KRIMBAS, 1973 Evidence for linkage disequilibrium maintained by selection in two natural populations of *Drosophila subobscura*. Genetics **73:** 659-674.