Genet. Res., Camb. (2004), **83**, *pp.* 169–176. *With 3 figures.* © 2004 Cambridge University Press
DOI: 10.1017/S0016672304006858 Printed in the United Kingdom

Sex-specific selection on the human X chromosome?

PATRICIA BALARESQUE[1,2]*, BRUNO TOUPANCE[1], QUINTANA-MURCI LLUIS[3], BRIGITTE CROUAU-ROY[2] AND EVELYNE HEYER[1]*

[1] *Unité Eco-Anthropologie MNHN/CNRS/P7 UMR5145, Paris, France*
[2] *Laboratoire Evolution et Diversité Biologique, Toulouse, France*
[3] *CNRS URA1961, Institut Pasteur, Paris, France*

(*Received 12 August 2003 and in revised form 12 March 2004*)

Summary

Genes involved in major biological functions, such as reproductive or cognitive functions, are choice targets for natural selection. However, the extent to which these genes are affected by selective pressures remains undefined. The apparent clustering of these genes on sex chromosomes makes this genomic region an attractive model system to study the effects of evolutionary forces. In the present study, we analysed the genetic diversity of a X-linked microsatellite in 1410 X-chromosomes from 10 different human populations. Allelic frequency distributions revealed an unexpected discrepancy between the sexes. By evaluating the different scenarios that could have led to this pattern, we show that sex-specific selection on the tightly linked VCX gene could be the most likely cause of such a distortion.

1. Introduction

Natural selection is one of the main forces shaping the patterns of genetic variability in the human genome, although its role has been often neglected in most population genetics studies. Indeed, most genetic polymorphisms used in population genetic studies are assumed to be neutral and affected mainly by both mutation and genetic drift. Interestingly, the recent results from human genome sequencing have revealed that each category of repeated sequences possesses a specific dynamics in space and time. This suggests a combined and complex action of different evolutionary forces. Furthermore, these results revealed that microsatellites, which are among the most used markers in population genetics studies, displayed a non-random distribution through the human genome: there are fewer polymorphic loci on the X-chromosome compared with autosomes (International Sequencing Human Genome Consortium (ISHGC), 2001). The X-chromosome, known to harbour a number of genes involved in human fertility (Wang *et al.*, 2001; Saifi & Chandra, 1999) and in cognitive functions (Gécz & Mulley, 2000; Hurst & Randerson, 1999; Graves & Delbridge, 2001) may be potential target for natural selection. However, the precise extent to which these genes are affected by selective pressures, or a possible variation in selective pressures acting between both sexes, remains undefined. In addition, the precise extent to which these genes have an influence on surrounding sequences either through direct or background selection remains poorly studied. We investigated the allelic diversity of an X-linked dinucleotide microsatellite (DXS8175) located in the Xp22.3 region, a gene-rich region, in 10 human populations and focused on the allelic distributions. We found strikingly different allelic frequency distributions between males and females. In this study, we investigated the likely demographic and selective scenarios as the bases of these observations.

* Corresponding authors: Equipe de Génétique des populations, Musée de l'Homme, 17, place du Trocadéro, 75116 Paris, France. Tel: +33 1 44057253. Fax: +33 1 44057241. e-mail: heyer@mnhn.fr or balares@mnhn.fr

2. Materials and methods

(i) *Samples and PCR amplification*

We genotyped 951 unrelated subjects, for a total of 1410 chromosomes belonging to 10 different human populations from Africa (Akan and Yacouba from

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

Table 1. *Distribution of alleles frequencies in males (M) and females (F), $\Delta = p_f - p_m$, and statistical test of differentiation between the two sexes*

	DXS 8175[c] alleles frequencies											Exact test differentiation (Raymond & Rousset, 1995)	
Populations	10	11	12	13	14	15	16	17	18	19	20	χ^2	*P* value
European													
Corsican													
M[a] 63[b]				6	10	46	37	2				1·75	0·417
F 60				3	13	32	48	3					
$\Delta = p_f - p_m$				**−3**	**4**	**−14**	**12**	**2**					
Sardinian													
M 36			11		11	47	22	8				10·65	0·005
F 94			1		18	51	30						
$\Delta = p_f - p_m$			**−10**		**7**	**4**	**8**	**−8**					
Orcadian													
M 32						34	59	6				3·14	0·208
F 76					11	38	46	5					
$\Delta = p_f - p_m$					**11**	**4**	**−13**	**−1**					
African													
Akan													
M 59	5		8		36	17	17	14	2	2		15·05	0·000
F 166	5		3		18	16	45	11		1	1		
$\Delta = p_f - p_m$	**−0**		**−6**		**−18**	**−1**	**28**	**−2**	**−2**	**−1**	**1**		
Yacouba													
M 62			3		19	27	37	6	3	3		1·21	0·546
F 46	2		2		28	13	39	11	2	2			
$\Delta = p_f - p_m$	**2**		**−1**		**9**	**−14**	**2**	**4**	**−1**	**−1**			
Amhara													
M 31				7	26	19	32	13	3			0·35	0·838
F 56		5	5	5	20	14	36	13	2				
$\Delta = p_f - p_m$		**5**	**5**	**−1**	**−6**	**−5**	**3**	**−0**	**−1**				
Oromo													
M 31			7	3	13	13	55	10				1·97	0·373
F 80	3		6		23	21	38	6		4			
$\Delta = p_f - p_m$	**3**		**−0**		**10**	**8**	**−17**	**−4**		**4**			
Moroccan Berber													
M 38					11	47	32	5		5		0·63	0·728
F 136	2		1	2	14	43	33	5		1			
$\Delta = p_f - p_m$	**2**		**1**	**2**	**4**	**−4**	**2**	**−0**		**−4**			
Mozabit Berber													
M 85			1	1	5	34	48	9	1			2·25	0·324
F 42	2		2		14	24	50	7					
$\Delta = p_f - p_m$	**2**		**1**	**−1**	**10**	**−10**	**2**	**−2**	**−1**				
South American													
Bolivian													
M 55					9	60	29	2				0·24	0·889
F 162				1	7	57	33	1					
$\Delta = p_f - p_m$				**1**	**−2**	**−3**	**4**	**−1**					

[a] Sexes.
[b] Number of chromosomes analysed.
[c] Alleles are named according to Scozzari *et al.* (1997).

Ivory Coast, Amhara and Oromo from Ethiopia, Algerian Mozabits and Moroccan Berbers from North Africa), Europe (Sardinian, Corsican and Orcadian) and South-America (Bolivian) (see Table 1). DXS8175 microsatellite amplifications were performed according to Malaspina *et al.* (1997) and Scozzari *et al.* (1997). PCR primers were fluorescently labelled and the PCR products were run in a standard 6% denaturing gel and detected using an ABI 373A automated sequencer. GeneScan software (ABI) and Genotyper software package (ABI) were used to size the amplified alleles. In addition, we sequenced a total

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

of 50 microsatellites randomly chosen from the 10 populations to control for possible indel events in flanking sequences (Grimaldi & Crouau-Roy, 1997) and showed that no indels have caused length homoplasy. Moreover, no null alleles have been detected as previously reported by Scozzari *et al.* (1997).

(ii) *Statistical analysis*

The distributions of allele frequencies in the two sexes were compared using the exact test of population differentiation, implemented in GENEPOP software (Raymond & Rousset, 1995), well adapted for allele frequencies comparisons (Goudet *et al.*, 1996). The probability of type I error for the test was set at 0·05 and a Bonferroni correction for multiple tests used following the method of Dunn and Sidak (Ury, 1976). The analytical study of the evolution at X-linked loci under selection was performed with Mathematica 4.1 (Wolfram, 2001). Figures were obtained with R software (Ihaka & Gentleman, 1996).

3. Results and Discussion

The DXS8175 displays a total of 10 alleles, ranging from 10 to 20 repeats in all populations studied, a common feature for a dinucleotide microsatellite (Zhivotovsky *et al.*, 2003; Renwick *et al.*, 2001). Allelic frequencies in the 10 populations revealed similar distributions to those observed by Scozzari *et al.* (1997) in 30 populations from Europe, Africa, Asia and the Americas for the same marker. However, we observed a discrepancy in allele frequencies between males and females in all populations. We found that two populations (Akan and Sardinian) show a significant difference in allele frequencies (alleles 12, 14, 16, 17) between the sexes (Table 1). This difference remained significant after Bonferonni correction for multiple tests.

By evaluating the different scenarios that could have led to this pattern, it appeared that two main possible and testable hypotheses could explain the discrepancy in allele frequencies: an admixture event or a sex-specific selection acting on a gene located on the X chromosome affecting the DXS8175 marker by a hitch-hiking effect.

(i) *Admixture*

To test whether admixture is the putative cause for the observed differences, we need to take into consideration the parental populations. If a population results from the admixture of two different founder groups with significant differences in allele frequencies, discrepancies are expected in the offspring generation (F1). The sex difference in the F1 generation is equal to half the difference among parental generation (with an opposite sign). In the more extreme case, where all males and females come from two different single populations (corresponding to 100% admixture), differences in allele frequencies between sexes in the parental populations have to be twice that seen in the F1 generation. In a moderate case, and so more realistic for human populations, if admixture is less than 100%, the frequency difference between sexes in the parental populations has to be much higher than twice the F1 difference. The necessary frequencies under this hypothesis were not observed in our study.

For example, the highest discrepancy in allele frequencies was observed in the Akan population, in which allele 16 showed the highest male to female frequency divergence: $\Delta = p_f - p_m = +0{\cdot}28$. In order to explain such a result, we could consider a fictitious case (100% admixture) where the most divergent populations in allele frequencies found in our sample would represent the parental populations. Considering males of the Orcadian and females of the Sardinian populations, the maximal value which could be reached in the F1 generation would be $\Delta = p_f - p_m = 0{\cdot}30/2 = 0{\cdot}15$ (allele 16_m: 0·59 and allele 16_f: 0·29), a value much lower than 0·28.

Others pairs of parental populations would result in lower expected divergence or in male frequencies being higher than female frequencies. Furthermore, the expected frequency patterns for others alleles would never match the observed Akan pattern. Using similar calculations, it can easily be shown that the discrepancy in sex frequency for the Sardinian population (for allele 12 or 17) cannot be explained by admixture. The admixture hypothesis seems not adapted to explain the sex frequency discrepancies in the Akan or the Sardinian population. The historical and demographic parameters necessary to achieve such a situation (100% admixture) from very distant populations are very unrealistic and an admixture event leading to the observed pattern of sex-allele discrepancies is unlikely. Such a difference in allele frequencies between males and females was never observed when we compared allele frequencies for other microsatellites on other parts of the genome (data not shown).

This argument can be extended to a multiple allele approach by noticing that the distance between sexes in the parental populations, as measured by F_{ST}, has to be approximately 4 times greater than the distance between the two sexes in the F1 generation. Such a distance was never achieved (data not shown).

(ii) *Selection*

(a) *Nearness of* VCX/Y *genes*

If the DXS8175 microsatellite is in the vicinity of a gene under selective pressures, this could explain such

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

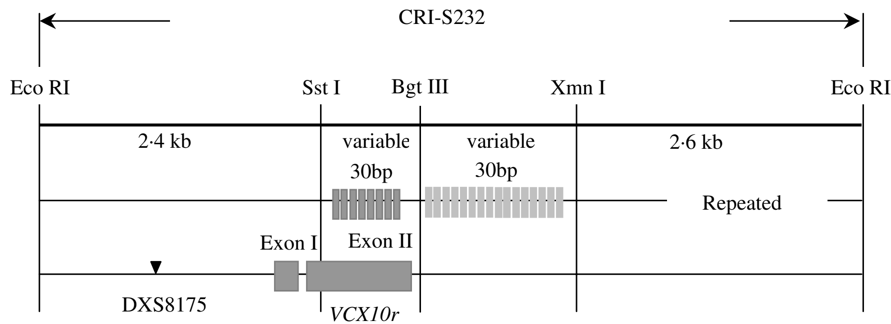

Fig. 1. Relative positions of DXS8175 microsatellite and *VCX10r* gene on the CRI-S232 genomic element (adapted from Lahn & Page, 2000; Fukami *et al.*, 2000).

a discrepancy in allele frequencies. By performing *in silico* investigation, we located the DXS8175 microsatellite ~1 kb upstream of the *VCX10r* gene (Fukami *et al.*, 2000; Lahn & Page, 2000; Balaresque *et al.*, 2003; Fig. 1) within a CRI-S232 duplicated element (Ballabio *et al.*, 1990; Li *et al.*, 1992), also called segmental duplications. The recombination fraction expected between the DXS8175 microsatellite and the *VCX10r* gene is approximately 0·001 %, confirming the association between the microsatellite and the gene.

(b) *Antagonistic allelic sex-specific selection on the X-chromosome*

Differential sex selection on the human X-chromosomes is an alternative explanation for the discrepancy in allele frequencies distributions in males and females. Population genetic theory shows that stable polymorphisms can be maintained by selection at sex-linked loci when alleles are antagonistically selected in the two sexes (see for example Crow & Kimura, 1970, p. 278). Here we will show that in such cases large differences between allele frequencies in males and females can be obtained at equilibrium. Consider a single locus on the X-chromosome with two alleles (A1 and A2) under zygotic selection with one allele (A1) being deleterious in males (heterogametic sex) but advantageous in females (homogametic sex). Selective effects of A1/A2 alleles were parameterized by a selective coefficient s reducing YA1 male fitness to $w1=1-s$ (YA2 male fitness being $w2=1$), a selective coefficient t reducing homozygous A2A2 female fitness to $w22=1-t$ (homozygous A1A1 female fitness being $w11=1$), and a dominance parameter h (i.e. heterozygous A1A2 female fitness was $w12=1-ht$). We supposed that the fitness of heterozygous females was intermediate between those of homozygous females (i.e. h varying between 0 and 1), thus excluding under- or over-dominance phenomena. This fitness model was a slight modification of the one used by Rice (1984) and was adopted because of its symmetry, all parameters (s, t and h) varying between 0 and 1.

The precise analysis of the evolution at X-linked loci under selection has been done by many authors (see for example Cannings, 1967; Crow & Kimura, 1970; Rice, 1984). The existence of a stable polymorphic equilibrium depended on the three parameters (s, t, h) as shown in Fig. 2. Globally, the region for a stable polymorphism was reduced with increase in the dominance parameter h.

Using our fitness model, equilibrium A1 frequencies in females ($\hat{p}_f$) and in males ($\hat{p}_m$) were:

$$\hat{p}_f=\frac{t[2-h(2-s)]-s}{2t[1-h(2-s)]}, \tag{1}$$

$$\hat{p}_m=\frac{(1-s)[t(2-h(2-s))-s]}{s^2+t[2(1-s)-h(2-s)^2]}. \tag{2}$$

Combining these equations (which are equivalent to equations 7 and 8 of Rice, 1984), the difference in allelic frequencies, $\Delta=\hat{p}_f-\hat{p}_m$, was:

$$\Delta=\frac{s[t(2-h(2-s))-s][s-(2-s)ht]}{2t[1-h(2-s)][s^2+t\{2(1-s)-h(2-s)^2\}]}. \tag{3}$$

This difference was strictly positive in the stable polymorphic equilibrium region, i.e. equilibrium A1 female frequency $\hat{p}_f$ was always higher than equilibrium A1 male frequency $\hat{p}_m$. The dominance parameter h had little influence on the difference between $\hat{p}_f$ and $\hat{p}_m$. As a result, we restricted our analysis to the case of a complete dominance of A1 over A2 in females (i.e. $h=0$). In this situation, the discrepancy between female and male A1 frequencies at equilibrium reduced to:

$$\Delta=\frac{s^2(2t-s)}{2t[s^2+2t(1-s)]}. \tag{4}$$

The exact value of the excess of A1 allele in females (as measured by Δ) obviously depended on the selective coefficients s and t, and can be visualized in the (s, t) parameters space as lines with equal Δ values

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

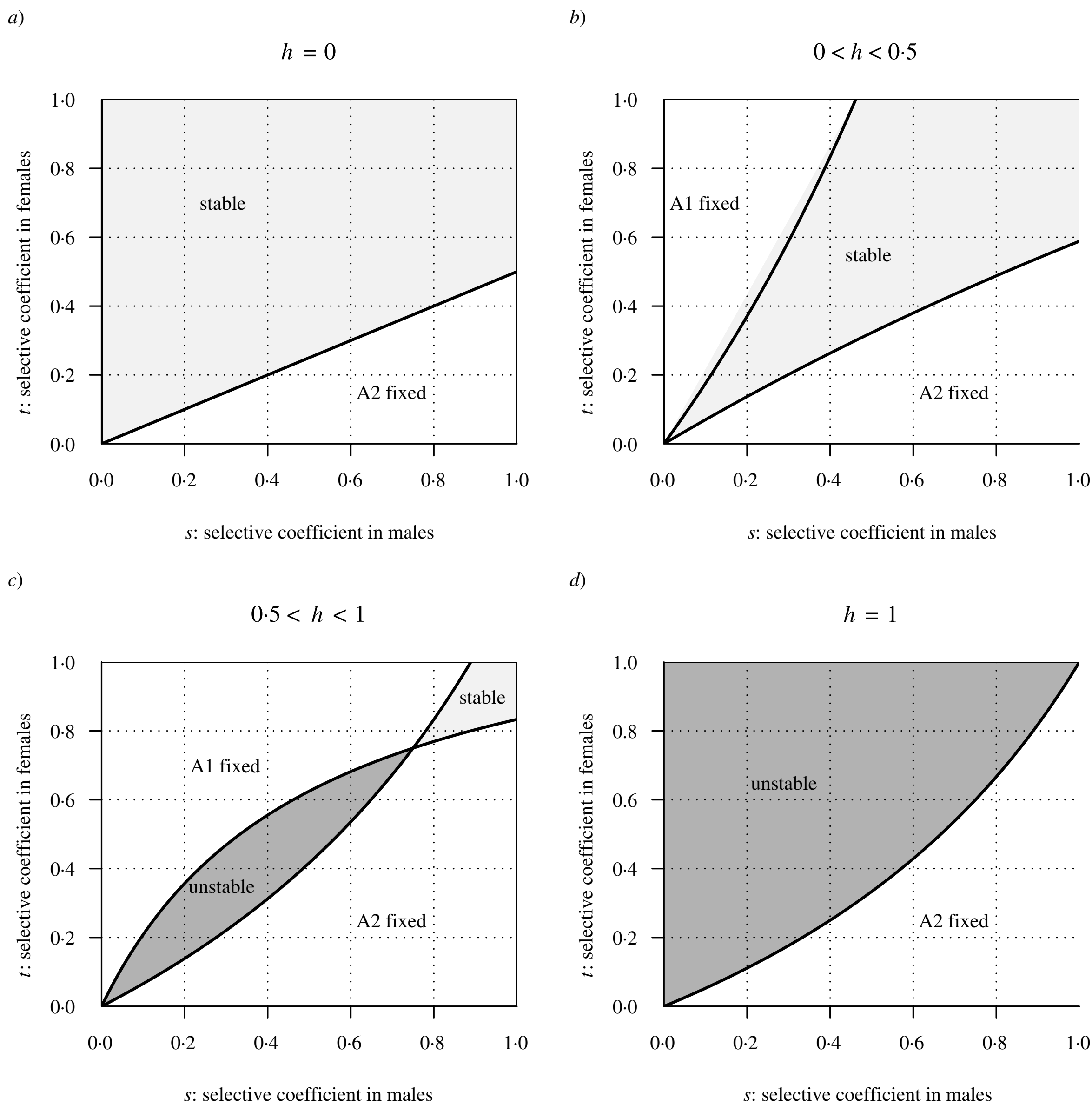

Fig. 2. Equilibrium at X-linked loci under antagonistic selection in the two sexes (allele A1 being deterious in males and advantageous in females). White areas correspond to monomorphic equilibrium (A1 fixed or A2 fixed), light grey areas correspond to stable polymorphic equilibrium, and dark grey areas correspond to unstable polymorphic equilibrium. Stability is possible when $(1-ht)(2-s)>2(1-s)$ and $(1-ht)(2-s)>2(1-t)$. (*a*) When A1 is dominant in females ($h=0$), two equilibrium states are possible: a stable polymorphic state and a monomorphic state (A2 fixed). (*b*) When A1 is partially dominant in females ($0<h<0{\cdot}5$), there are two monomorphic states (A1 fixed or A2 fixed) and a reduced area for the polymorphic state. (*c*) When A1 is partially recessive in females ($0{\cdot}5<h<1$), an unstable polymorphic state appears. (*d*) When A1 is partially recessive in females ($h=1$), polymorphism is unstable, thus A2 always becomes fixed.

(Fig. 3). Globally, Δ increased with increasing s and t (i.e. under high selection) and tended to 0·5 with males being only YA2 and females being only A1A2 heterozygotes. For more realistic selective parameters (i.e. lower s and t values), Δ was strongly decreased and would be impossible to detect for selective coefficients s and t lower than 0·2 ($\Delta<0{\cdot}05$).

With this model, we showed that sexual antagonism is sufficient to create differences in allele frequencies in males and females. The more extreme pattern we observed (Akan population) can be explained in this framework, although it requires very high selective coefficients both in males and in females ($s=0{\cdot}748$ and $t=0{\cdot}675$ for $\hat{p}_m=0{\cdot}169$ and $\Delta=0{\cdot}277$). We must, however, note (1) that the true Δ value in the Akan population might be lower than that observed in our sample, and (2) that the observed Δ value might not be the equilibrium one. Classical population genetics results show that an initial difference in allele frequencies in males and females for a neutral locus on the X-chromosome needs some generations to disappear with fluctuations around 0 from generation to

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

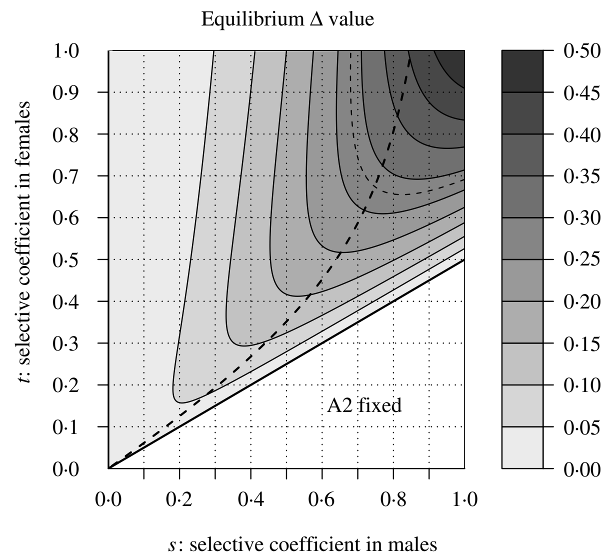

Fig. 3. Equilibrium $\Delta = \hat{p}_f - \hat{p}_m$ value in selective coefficients (s, t) space with complete dominance of A1 in females ($h=0$). The stable polymorphic equilibrium region is delimited by the thick plain line: A2 gets fixed for parameters values below the line. Thin plain lines represent (s, t) values yielding a given Δ values (shown values from 0·05 to 0·45). The thin dashed line represents $\Delta = 0{\cdot}277$. The thick dashed line represents an equilibrium A1 frequency in males $\hat{p}_m = 0{\cdot}169$. The dashed line intersect at $s = 0{\cdot}748$ and $t = 0{\cdot}675$.

generation (being successively positive and negative). Adding sexual antagonistic selection (as in our model) results in a translation of the equilibrium value from 0 (in the neutral case) to a positive value. In the non-equilibrium phase, fluctuations may well transiently increase Δ to a value much higher than its final equilibrium value. Random genetic drift or moderate admixture could be recurrent sources of displacement from the equilibrium. In some situations (especially when the discrepancy is reduced by such phenomena), selection could transiently drive the system to higher values than expected at equilibrium.

The model with an advantage in the homogametic sex could explain the pattern observed in the Akan population. Similarly, the reverse frequency pattern observed in the Sardinian population (higher frequency of allele 12 or 17 in males than in females) could be explained by the reverse model where A1 is advantageous in males and deleterious in females.

(c) VCX/Y *genes: target for selection*

We showed that a differential selection between the two sexes on a gene in the vicinity of the DXS8175 microsatellite would create such a difference in allelic frequency distribution at equilibrium in a given population. The *VCX10r* gene, a member of the *VCX/Y* gene family, is a good candidate as expression analysis showed that all copies of the *VCX/Y* gene family have a testis-specific expression, probably in the germ cells (Fukami *et al.*, 2000; Lahn & Page, 2000). Their involvement in female reproductive functions remains to be defined but, to our knowledge, no expression studies of the *VCX/Y* gene family have been performed in foetal ovary tissue and therefore a role of *VCX* members during oogenesis cannot be ruled out. A similar sex-specific selective process has been observed in *Drosophila* for sexually antagonistic genes (Rice, 1992; Chippindale *et al.*, 2001) in which some genes are advantageous in the heterogametic sex whereas they are disadvantageous in the homogametic sex. Genes located on the human X-chromosome constitute potential and interesting targets on which antagonistic selective pressures between both sexes could be acting (Gibson *et al.*, 2002).

A second class of model could involve selection acting at the gametic level rather than at the zygotic level, resulting in similar sex frequency discrepancies provided that selection acts antagonistically in the two sexes (data not shown).

A third class of model with alleles acting as a segregation distorter in males but being deleterious in females would probably also result in a difference in allele frequency at equilibrium. As we have shown, the DXS8175 microsatellite is located near the *VCX* gene within the duplicated element CRI-S232. Interestingly, in a recent study Lahn & Page (2000) and Lahn *et al.* (2001) reported that the *VCX/Y* genes could act as meiotic distorters. Their statement is mainly based on two observations: (i) the molecular characteristics of *VCX/Y* genes that recalled those of the fruitfly X-linked Stellate (Ste) and Y-linked crystal, which are meiotic drive elements in *Drosophila melanogaster* (Belloni *et al.*, 2002); (ii) recombination between CRI-S232 elements is known to cause frequent deletions in the X-chromosome short arm, resulting in steroid sulfatase deficiency (X-ichthyosis). This could be a satisfactory explanation for an old speculation of male bias among the offspring of ichthyosis-carrier females reported in some human populations (Filippi & Meera Khan, 1968; Gladstein *et al.*, 1979).

(iii) *Differences among populations*

The observed differences in allele frequencies between sexes, dependent on the allele or the population, could be due to several causes including differential demographic histories of some populations associated with variation in linkage disequilibrium (LD) levels (Ardlie *et al.*, 2002): if the mutation arrived more recently in one population, there is a higher LD between the gene under selection and the microsatellite, and therefore it can be detected through the microsatellite polymorphism in this population. In a population where the mutation arrived earlier, LD has been reduced through recombination between the selected gene

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

and the microsatellite and it is very difficult to detect such an effect. This could explain why such an association is only detectable in 2 populations in 10. Among these, the Sardinian population is known to show high levels of linkage disequilibrium (Taillon-Miller *et al.*, 2000; Angius *et al.*, 2002).

Moreover, local selection on the *VCX* gene could also explain differences between populations. The instability of this region through misalignment between duplicated elements could lead to a copy number polymorphism of *VCX* genes: these differences among populations have been documented for other gene families (Trask *et al.*, 1998), suggesting that variable selective patterns may be expected across populations.

4. Conclusions and perspectives

We have shown that discrepancies in allele frequencies between males and females are probably due to sex-specific selective pressures. The model of Rice (1984) seems well adapted to illustrate intra-locus antagonistic pressures acting on sex-specific-linked alleles. Although this model with two alleles is clearly an over-simplification of the reality, it provides an interesting framework to explain our data. The concept of antagonistic intra- or inter-locus selective pressures becomes especially relevant when the candidate loci are polymorphic and are part of a multigenic family in which different members act in synergy. Recent results on the human genome reported that segmental duplications constitute approximately 5% of the human genome and that all copies of each family share about 90–100% similarity (ISHGC, 2001; Samonte & Eichler, 2002). These large blocks of sequence similarity provide the substrate for aberrant recombination leading to variation in copy number among individuals or populations (Menashe *et al.*, 2003). This observation underlines the fact that duplicated sequences are part of an ongoing process that results in a novel form of large-scale variation in the human genome (Eichler, 2001), which may be subject to complex selective patterns.

We would like to thank the following collaborators for DNA samples: A. Chaventre, G. F. De Stefano, A. Baali, M. Cherkaoui, J. M. Dugoujon, G. Bellis, P. Kouamé, A. Pitte, M. S. Isaad, G. Larrouy and A. Sevin. Thanks go also to R. Jambou-Clerc for his technical assistance; and to P. H. Gouyon and A. Sibert for fruitful discussions. This work was supported by the Research group 'interaction génomique' (CNRS).

References

Angius, A., Bebbere, D., Petretto, E., Falchi, M., Forabosco, P., Maestrale, B., *et al.* (2002). Not all isolates are equal: linkage disequilibrium analysis on Xq13.3 reveals different patterns in Sardinian sub-populations. *Human Genetics* **111**, 9–15.

Ardlie, K. G., Kruglyak, L. & Seielstad, M. (2002). Patterns of linkage disequilibrium in the human genome. *Nature Review Genetics* **3**, 299–309.

Balaresque, P., Toupance, B., Heyer, E. & Crouau-Roy, B. (2003). Evolutionary dynamics of duplicated microsatellites shared by sex chromosomes. *Journal of Molecular Evolution* **57**, S128–S137.

Ballabio, A., Bardoni, B., Guioli, S., Basler, E. & Camerino, G. (1990). Two families of low-copy-number repeats are interspersed on Xp22.3: implications for the high frequency of deletions in this region. *Genomics* **8**, 263–270.

Belloni, M., Tritto, P., Bozzetti, M. P., Palumbo, G. & Robbins, L. G. (2002). Does Stellate cause meiotic drive in *Drosophila melanogaster*? *Genetics* **161**, 1551–1559.

Cannings, C. (1967). Equilibrium, convergence and stability at a sex-linked locus under natural selection. *Genetics* **56**, 613–618.

Chippindale, A. K., Gibson, J. R. & Rice, W. R. (2001). Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in Drosophila. *Proceedings of the National Academy of Sciences of the USA* **98**, 1671–1675.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York.

Eichler, E. E. (2001). Recent duplication, domain accretion and the dynamic mutation of the human genome. *Trends in Genetics* **17**, 661–669.

Filippi, G. & Meera Khan, P. (1968). Linkage studies on X-linked ichthyosis in Sardinia. *American Journal of Human Genetics* **20**, 564–569.

Fukami, M., Kirsch, S., Schiller, S., Richter, A., Benes, V., Franco, B., *et al.* (2000). A member of a gene family on Xp22.3, VCX-A, is deleted in patients with X-linked non-specific mental retardation. *American Journal of Human Genetics* **67**, 563–573.

Gécz, J. & Mulley, J. (2000). Genes for cognitive function: developments on the X. *Genome Research* **10**, 157–163.

Gibson, J. R., Chippindale, A. K. & Rice, W. R. (2002). The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London, Series B* **269**, 499–505.

Gladstein, K., Shapiro, L. J. & Spence, M. A. (1979). Estimating sex-ratio biases in X-linked disorders: is there an excess of males in families with X-linked ichthyosis? *American Journal of Human Genetics* **31**, 1091–1094.

Goudet, J., Raymond, M., de Meeus, T. & Rousset, F. (1996). Testing differentiation in diploid populations. *Genetics* **144**, 1933–1940.

Graves, J. A. M. & Delbridge, M. L. (2001). The X: a sexy chromosome. *Bioessays* **23**, 1091–1094.

Grimaldi, M. C. & Crouau-Roy, B. (1997). Microsatellite allelic homoplasy due to variable flanking sequences. *Journal of Molecular Evolution* **44**, 336–340.

Hurst, L. D. & Randerson, J. P. (1999). An eXceptional chromosome. *Trends in Genetics* **15**, 383–385.

Ihaka, R. & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299–314.

International Sequencing Human Genome Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

Lahn, B. T. & Page, D. C. (2000). A human sex-chromosomal gene family expressed in male germ cells and encoding variably charged proteins. *Human Molecular Genetics* **9**, 311–319.

Lahn, B. T., Pearson, N. M. & Jegalian, K. (2001). The human Y chromosome, in the light of evolution. *Nature Review Genetics* **2**, 207–216.

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press

Li, X. M., Yen, P. H. & Shapiro, L. J. (1992). Characterization of a low copy repetitive element S232 involved in the generation of frequent deletions of the distal short arm of the human X chromosome. *Nucleic Acids Research* **20**, 1117–1122.

Malaspina, P., Ciminelli, B. M., Viggiano, L., Jodice, C., Cruciani, F., Santolamazza, P., *et al*. (1997). Characterization of a small family (CAIII) of microsatellite-containing sequences with X-Y homology. *Journal of Molecular Evolution* **44**, 652–659.

Menashe, I., Man, O., Lancet, D. & Gilad, Y. (2003). Different noses for different people. *Nature Genetics* **34**, 143–144.

Raymond, M. & Rousset, F. (1995). GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–249.

Renwick, A., Davison, L., Spratt, H., King, J. P. & Kimmel, M. (2001). DNA dinucleotide evolution in humans: fitting theory to facts. *Genetics* **159**, 737–747.

Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742.

Rice, W. R. (1992). Sexually antagonistic genes: experimental evidence. *Science* **256**, 1436–1439.

Saifi, G. M. & Chandra, H. S. (1999). An apparent excess of sex- and reproduction-related genes on the human X chromosome. *Proceedings of the Royal Society of London, Series B* **266**, 203–209.

Samonte, R. V. & Eichler, E. E. (2002). Segmental duplications and the evolution of the primate genome. *Nature Reviews Genetics* **3**, 65–72.

Scozzari, R., Cruciani, F., Malaspina, P., Santolamazza, P., Ciminelli, B. M., Torroni, A., *et al*. (1997). Differential structuring of human populations for homologous X and Y microsatellite loci. *American Journal of Human Genetics* **61**, 719–733.

Taillon-Miller, P., Bauer-Sardina, I., Saccone, N. L., Putzel, J., Laitinen, T., Cao, A., *et al*. (2000). Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nature Genetics* **25**, 324–328.

Trask, B. J., Friedman, C., Martin-Gallardo, A., Rowen, L., Akinbami, C., Blankenship, J., *et al*. (1998). Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Human Molecular Genetics* **7**, 13–26.

Ury, H. K. (1976). A comparison of four procedures for multiple comparisons among means (pairwise contrasts) for arbitrary sample sizes. *Technometrics* **18**, 89–97.

Wang, P. J., McCarrey, J. R., Yang, F. & Page, D. C. (2001). An abundance of X-linked genes expressed in spermatogonia. *Nature Genetics* **27**, 422–426.

Wolfram, S. (2001). *Mathematica (version 4.1): A System for doing Mathematics by Computer*. Reading, MA: Addison-Wesley.

Zhivotovsky, L. A., Rosenberg, N. A. & Feldman, M. W. (2003). Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. *American Journal of Human Genetics* **72**, 1171–1186.

https://doi.org/10.1017/S0016672304006858 Published online by Cambridge University Press