Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
International Journal for Parasitology 34 (2004) 943–950
INTERNATIONAL Journal for PARASITOLOGY
www.parasitology-online.com

Non-Mendelian transmission of alleles at microsatellite loci: an example in *Ixodes ricinus*, the vector of Lyme disease☆

Thierry de Meeûs[a,*], Pierre-François Humair[b], Christoph Grunau[c], Christelle Delaye[a], François Renaud[a]

[a]*Génétique et Evolution des Maladies Infectieuses, Equipe Evolution des Systèmes Symbiotiques, UMR 2724 CNRS-IRD, BP 64501, 911 Av. Agropolis, 34394 Montpellier Cedex 5, France*
[b]*Institut de Zoologie, Rue Emile-Argand 11, 2007 Neuchâtel, Switzerland*
[c]*Institut de Génétique Humaine (IGH), CNRS UPR 1142, 141 rue de la Cardonille, 34396 Montpellier, France*

Received 5 February 2004; received in revised form 15 April 2004; accepted 18 April 2004

Abstract

Microsatellite loci are usually considered to be neutral co-dominant and Mendelian markers. We undertook to study the inheritance of five microsatellite loci in the European Lyme disease vector, the tick *Ixodes ricinus*. Only two loci appeared fully Mendelian while the three others displayed non-Mendelian patterns that highly frequent null alleles could not fully explain. At one locus, IR27, some phenomenon seems to hinder the PCR amplification of one allele, depending on its origin (maternal imprinting) and/or its size (short allele dominance). DNA methylation, which appeared to be a possible explanation of this amplification bias, was rejected by a specific test comparing the amplification efficiency that did not differ between unmethylated and experimentally methylated DNA. The role of allele size in heterozygous individuals was then revealed from the data available on field collected ticks and consistent with the results of a theoretical approach. These observations highlight the need for prudence while inferring reproductive systems (selfing rates), parentage or even allelic frequencies from microsatellite markers, in particular for parasitic organisms for which molecular approaches often represent the only way for population biology inferences.
© 2004 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: Microsatellite loci; Inheritance; Null alleles; Parental imprinting; Methylation; Allele size; Short allele dominance

1. Introduction

Microsatellite loci have had a considerable popularity in population genetics studies during the past decade because these highly polymorphic and relatively easy to use molecular markers are considered neutral, codominant and Mendelian (Lehmann et al., 1996; Jarne and Lagoda, 1996). Nevertheless, some technical problems are often met, such as null alleles (Paetkau and Strobeck, 1995; Pemberton et al., 1995; Brookfield, 1996), misscoring of allele size (Delmotte et al., 2001) and short allele dominance (Wattier et al., 1998).

During a previous study at the population level of *Ixodes ricinus*, a hard tick known as the main vector of Lyme disease in Europe (Balmelli and Piffaretti, 1996; Crump, 2000; Humair and Gern, 2000), microsatellites displayed unusual patterns. In particular, these microsatellites showed huge and highly variable heterozygote deficit (Wright's, 1965, F_{is}) that null alleles could not totally explain (De Meeûs et al., 2002a). We therefore undertook further studies in order to better understand the causes of such phenomena and their consequences for population genetics inferences.

In this paper, we present and discuss the results obtained during the analysis of eight families of *I. ricinus* ticks that were genotyped at the five polymorphic microsatellite loci described by Delaye et al. (1998).

As PCR amplification problems were met for maternal alleles at one locus (locus IR27), just as if parental imprinting was occurring at that locus, we designed an experiment to test for the existence of such a phenomena. Imprinting may result from methylation

☆ Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijpara.2004.04.006.

* Corresponding author. Tel.: +33-467-41-63-10; fax: +33-467-41-62-99.
E-mail address: demeeus@mpl.ird.fr (T. de Meeûs).

0020-7519/$30.00 © 2004 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijpara.2004.04.006

processes (Chakraborty, 1989; Barlow, 1993; LaSalle and Lalande, 1996; Jaenisch, 1997; Bestor, 2000) that may disturb DNA amplification (Gérard Cuny, personal communication). DNA from many species can contain considerable amounts of the fifth base 5-methylcytosine (5mC). 5mC is chemically different from cytosine and has an effect on the structure of DNA (Behe and Felsenfeld, 1981; Banyay and Graslund, 2002; Nathan and Crothers, 2002). For instance, T_m (melting temperature: the temperature at which half the DNA strands are single-stranded and half are double-stranded) for 5mC containing polynucleotides were found to be higher than for the unmethylated polynucleotides (Klump and Loffler, 1985). We wondered therefore whether methylation of template DNA has an effect on PCR amplification efficiencies that is measurable in a quantitative PCR approach. An experimental molecular manipulation of tick DNA was then designed to specifically test the possible role of methylation of tick DNA in the parental imprinting-like pedigrees noticed for locus IR27.

Finally, and because the parental imprinting hypothesis was not confirmed by the methylation experiment, we re-analysed the available field data at this locus, and tried to adjust them to a theoretical model where DNA amplification of alleles of heterozygous individuals is biased toward the shortest alleles (short allele dominance) as it was already observed for a red alga by Wattier et al. (1998).

2. Materials and methods

2.1. Pedigrees

To make sure of females' virginity, 105 nymphs were collected in a woodland near Neuchâtel named 'Bois de l'Hôpital' (47° North latitude, 6.75° East longitude, North Western Switzerland, see De Meeûs et al., 2002a) in spring 1996 by the flag technique (e.g. Aesclimann, 1972) and were allowed to feed on Swiss white mice in the Neuchâtel laboratory. Engorged nymphs were individually maintained at relative humidity close to saturation (>95%) and at room temperature until moult was completed. When adult ticks emerged from the nymphal exuviae, surviving female ($n = 53$) and male ($n = 50$) ticks were placed and maintained in separate tubes. One month after moult, 48 pairs could be formed and each pair was maintained in a tube to allow mating. One month later, females were allowed to feed on the ears of New Zealand white rabbits. Engorged females were individually maintained, as previously described for engorged nymphs, until they laid eggs. Ticks were preserved in 70% ethanol: males after mating, females after egg laying and larvae after hatching (6–7 months after nymph sampling). Only 16 parent pairs and their offspring (380 larvae of which were genotyped) could be obtained (other pairs failed to produce offspring).

2.2. Genotyping parents and their offspring

Genotyping of the families was undertaken during the year 1998. We used the five polymorphic microsatellite loci described by Delaye et al. (1998), IR8, IR25, IR27, IR32 and IR39, with the same amplification methods and allele scoring (see also De Meeûs et al., 2002a). Among the 16 parent pairs that were genotyped, only eight could be exploited, since in the eight other pairs both parents were homozygous for the same allele at all loci. Four pairs could be used for locus IR8 (pairs 1, 4, 6 and 8), two pairs for locus IR25 (pairs 6 and 8), six pairs for locus IR27 (pairs 1, 2, 4, 6, 7 and 8), two pairs for locus IR32 (pairs 3 and 5) and three pairs for locus IR39 (pairs 3, 4 and 8). Larval individuals could be analysed for only one locus (not enough DNA).

The goodness of fit between the genotypic proportions observed in the offspring to those expected under Mendelian expectations was tested using the multinomial test (e.g. Weir, 1990).

When required (i.e. non-Mendelian pedigrees of locus IR27) PCR amplification was repeated. This was then followed by an asymmetric PCR (one primer 10 times in excess), in order to reveal a potentially hidden allele as successfully used for locus IR39 in De Meeûs et al. (2002a) (asymmetric PCR may compensate for the decrease in affinity of one primer with the DNA sample). The genotype resulting from the mixture of an equal amount of DNA of the mother and the father of the problematic pedigrees was also compared to the pattern observed in the offspring. This was made to test if the presence of one particular allele may modify the amplification of another allele when mixed in the same PCR reaction.

2.3. Testing the possible role of DNA methylation

PCR amplification can be described by the formula $Y_n = Y_{n-1}(1 + E)$ with E being efficiency of each amplification step between 0 and 1, n the step number and Y the number of amplifiable DNA molecules. Considering an efficiency Em for the methylated template Ym, an efficiency Eu for the unmethylated DNA template Yu, and considering that methylation cannot be transmitted to amplified DNA, the above equation becomes $Y_n = Yu_{n-1}(1 + Eu) + Ym_{n-1}(1 + Em)$ with $Y_0 = Ym_0$. If Em is smaller than Eu, a considerable underestimation of the amount of sample DNA could occur (e.g. 10.2% after 25 PCR rounds with hypothetical $Em = 0.88$ and $Eu = 0.98$). The whole sequence corresponding to locus IR27, a $(CA)_n$ microsatellite, contains 20 CpG pairs (among which 11 pairs are in the amplified region and three pairs are in the two primer binding sites) (see Delaye et al., 1998) (CpG means a guanidin residue G followed by a cytosine residue C in 5′ to 3′ direction, connected covalently by a phosphate ester p). Genomic DNA from two male (M1, M2) and two female *I. ricinus* individuals (F2, F3), sampled in Switzerland during 1998, was extracted and PCR amplified for locus IR27.

PCR products were cloned into pGEM-T-Easy (Promega), and the inserts were re-amplified with universal primers T7 and SP6. PCR products were purified and quantified spectrophotometrically. PCR products were precipitated with PEG solution (Rosenthal et al., 1993) resolved in 10 mM Tris–HCl (pH 8) and quantified by measuring the optical density at 260 nm (Eppendorf BioPhotometer). In 1 μg PCR product, cytosines in CpG pairs were methylated with 4 U M.SssI (NEB). The CpG specific cytosine methylase M.SssI (New England Biolabs) methylates all cytosine residues (C5) within the double-stranded dinucleotide recognition sequence $5'\cdots CG\cdots 3'$. In an identical reaction mix, M.SssI was replaced by BSA (Bovine Serum Albumin, mock treatment). Methylation was confirmed with antibodies against 5-methylcytosine (Megabase Research) on dot-blots of 300 ng PCR product. 0.1, 1, 10 and 100 pg M.SssI-methylated and mock treated PCR products were amplified using IR27 primers. Increasing amounts of double-strand PCR products were detected by incorporation of the fluorescence dye SYBR Green I (LightCycler FastStart DNA kit, Roche Diagnostics). Real-time quantification of fluorescence intensity was performed in a LightCycler (Roche). Threshold cycles (crossing points) obtained with the real-time PCR were plotted against the logarithm of the template concentration. Straight lines were drawn and slopes were determined with the Excel (Microsoft) trend line function. Amplification efficiencies were calculated with $E = [10^{(-1/\text{slope})}] - 1$ (e.g. Pfaffl, 2001; Arezi et al., 2003). Amplification efficiencies for each methylated and unmethylated template sample were then compared by a Wilcoxon signed rank test for paired samples, computed by S-Plus 2000 Professional release 2 (MathSoft, Inc.).

2.4. Re-analysis of field data for locus IR27

Wright's (1965) F_{is} is a measure of the heterozygote deficit that can be found in a population relative to what is expected under Hardy–Weinberg expectations (see Hartl and Clark, 1989, page 294). It is a measure of the impact of non-random mating on genotypic frequencies found at each allele of each locus. Short allele dominance can be recognised by regressing the F_{is} of each allele against allele sizes (see Wattier et al., 1998), we thus attempted to find such a pattern in the available data (De Meeûs et al., 2002a). The raw data are available at http://cepm/cepm/sitewebess/gb/demeeus/tdemeeusgb.html/rawdatameeus.html (see also De Meeûs et al., 2002a). To test for the influence of allele size on the whole data set, but controlling for other factors, we used a generalised linear modelling approach with S-Plus 2000 professional release 2 (Mathsoft Inc.). The initial model was of the form f_i = poly(AllelSize, 2) + Site + Year + Sex + Site:Year + poly(AllelSize, 2): Sex + Site:Sex + Year:Sex + Site:Year:Sex + Constant, with f_i being the heterozygote deficit for allele i at locus IR27 estimated with Weir and Cockerham's (1984) method, poly(AllelSize, 2) being a quadratic function of the allele size, Site being the sample location, Year the year of sampling, Sex the sex of the tick and : meaning interaction. A quadratic expression was preferred for allele size because it gave better results than the monotonous relationship. Nevertheless, we checked that the monotonous case leads to similar conclusions. Because f_i significance depends on sample size (N) and allelic frequencies (p_i), we weighted the contribution of each observation by the product $Np_i(1 - p_i)$. In this way, the bigger and more polymorphic a sample is, the more weight is given to f_i. Following the principle of parsimony we have looked for the simplest possible model (see Crawley, 2002, page 221). We retained the minimum model after dropping terms in turn, using the Akaike Information Criterion (AIC) (S-PLUS 2000 Guide to Statistics, Volume1, 1999. Data Analysis Products Division, MathSoft, Seattle, WA).

2.5. Modelling DNA amplification bias in heterozygous individuals

The basic assumption of the model is that in a heterozygous individual, the shortest allele will be better (or faster) amplified by a factor proportional to the size difference between the two alleles in concern. Let α (between 0 and 1) be the effectiveness of the bias. Then, the heterozygosity expected, for the microsatellite locus under consideration in a panmictic population, between alleles i and j of frequencies p_i and p_j and sizes s_i and s_j, respectively, will be:

$$2p_ip_j\left[1 - \alpha\frac{s_i - s_j}{s_n - s_1}\right] \text{ if } s_i > s_j$$

and

$$2p_ip_j\left[1 - \alpha\frac{s_j - s_i}{s_n - s_1}\right] \text{ if } s_i < s_j$$

with n the number of alleles, $s_i < s_{i+1}$ and thus s_n the size of the longest allele.

Thus, the expected observed heterozygosity of allele i with all other alleles will be:

$$H_i = 2p_i\left\{(1-p_i) - \frac{\alpha}{s_n - s_1}\left[\sum_{j=1}^{j=i-1}(s_i - s_j)p_j + \sum_{j=i+1}^{j=n}(s_j - s_i)p_j\right]\right\}$$

which represents the expected heterozygosity under the panmictic assumption minus the proportion of those genotypes seen as homozygous for the shortest alleles. The expected observed homozygosity of allele i will thus be what is expected under panmixia plus the proportion of individuals carrying i that were heterozygous with a longer allele and seen homozygous for i, which is:

$$F_i = p_i\left\{p_i + 2\frac{\alpha}{s_n - s_1}\sum_{j=i+1}^{j=n}(s_j - s_i)p_j\right\}$$

Consequently the estimation of the frequency of allele i will be erroneously observed as $p_i' = F_i + \frac{1}{2}H_i$.

We can thus derive the expected observed heterozygote deficit at allele i defined by Wright (1965) as:

$$F_{is_i} = 1 - \frac{H_i}{2p_i'(1-p_i')}$$

We have applied this model to female *I. ricinus* from our Switzerland samples (De Meeûs et al., 2002a) because we expect female F_{is} to be less affected by Wahlund effects, i.e. heterozygote deficits due to the mixing of individuals from different populations (e.g. Hartl and Clark, 1989) than males. Indeed, *I. ricinus* females do not migrate much as compared to males (De Meeûs et al., 2002a). The expected F_{is} was computed for each allele in each sample by adjusting the allelic frequencies in order to get p_i's similar to the p_is estimated from the Swiss samples. This means that for each allele in each Swiss sample, assuming $\alpha = 1$, we searched for the allelic frequency that led to an estimate equal to the observed one. The correlation between these expected and the observed F_{is} values were then tested by a Pearson correlation test under S-Plus 2000 professional release 2 (Mathsoft Inc.).

3. Results

3.1. Pedigrees

In this section, N will always refer to sample size (number of genotyped offspring) and P to the multinomial exact probability to obtain by chance a genotypic distribution as different or more different than the one observed in the offspring under the null hypothesis of Mendelian inheritance (exact P-value).

3.1.1. Locus IR8

Taking into account that this locus is X-linked (De Meeûs et al., 2002a), the four couples examined provided a descent that conforms to Mendelian expectations (pair 1: $N = 23$, $P = 0.63$; pair 4: $N = 10$, $P = 0.11$; pair 6: $N = 42$, $P = 0.27$ and pair 8: $N = 17$, $P = 0.63$).

3.1.2. Locus IR39

Each of the three pairs analysed also provided a descent that conform to Mendelian expectations (pair 3: $N = 36$, $P = 0.08$; pair 4: $N = 34$, $P = 0.64$ and pair 8: $N = 18$, $P = 0.44$).

3.1.3. Locus IR32

For pair 3 the mother and the father displayed the homozygous phenotypes 247/247 and 233/233, respectively. The expected progeny composition was thus 100% 233/247 heterozygous phenotypes. However, we obtained five 233/247, eight 233/233, four 247/247 and seven bandless phenotypes (blanks). For pair 5, a similar observation was made. The parents 235/235 (mother) and 233/233 (father) produced six 233/235, five 233/233, five 235/235 and eight bandless phenotypes in their offspring. These unexpected results may be explained assuming that each member of the two parental pairs carried a null allele at a heterozygous state (e.g. Null/247 and Null/233 for pair 3). With this assumption, the observed proportions fit Mendelian expectations ($P = 0.7$ and 0.9 for pairs 3 and 5, respectively). Taking into account the estimated frequency of null alleles found in Neuchâtel in 1996 for males and females (0.69 and 0.54, respectively), using Brookfield's (1996) method on the raw data (see Section 2), the probability of drawing at random a pair that is heterozygous for a null allele is 0.21. Thus, the null allele explanation stands up.

3.1.4. Locus IR25

Two pairs were studied (pairs 6 and 8). For pair 8, the mother and the father were 147/147 and 154/154, respectively, and produced five 147/154, one 147/147, three 154/154 and three bandless phenotypes in their offspring. This unexpected result can be interpreted as for those observed at locus IR32. Indeed, if both parents are assumed heterozygous with a null allele, the observed proportions fit to Mendelian expectations ($P = 0.5$). Here estimates of null allele frequencies in Neuchâtel in 1996 (0.27 and 0.47 in females and males, respectively, unpublished data) gives a probability of 0.2 for such an event to occur. For pair 6, the mother and the father were 128/147 and 134/148, respectively, and provided 11 128/128, 13 134/147/148, nine 128/134/148 and 10 147/147, i.e. only unexpected or even impossible genotypes. The most parsimonious explanation we found for this surprising result is that the father was heterozygous with a duplicated gene on one chromosome (134–148) and a null allele on the other chromosome. If so the observed proportions fit expected ones ($P = 0.9$). Three and four banded individuals were occasionally observed in some populations for this locus in field collected individuals (unpublished data). However, only one male out of 48 displayed three bands in Neuchâtel in 1996 for that locus. When combined with the frequency of null alleles estimated for this site in males the event of having drawn by chance a male heterozygous for one duplication and one null allele appears relatively unlikely ($P = 0.01$).

3.1.5. Locus IR27

Inheritance in six families (pairs 1, 2, 4, 6, 7 and 8) was tested. Only one (pair 8, female 117/119, male 123/125) presented genotypic distributions in agreement with Mendelian expectations ($P = 0.515, N = 13$). Three pairs were also in agreement with Mendelian proportions assuming one null allele to be in the father (pair 4, 34 offspring, $P = 0.345$) or in each parent (pairs 1 and 6, 20 and seven offspring, $P = 0.436$ and 0.231, respectively). Note that for Neuchâtel in 1996, null alleles were estimated to be in frequencies 0.26 and 0.17

in females and males, respectively, at this locus (unpublished data), which makes such interpretations not unlikely ($P = 0.28$ for pair 4 and $P = 0.11$ for pairs 1 and 6). For pair 2 (female 113/123, male 119/119, nine offspring) null alleles cannot be invoked. Indeed, maternal alleles always appeared poorly amplified in two 113/119 offspring and five 119/123 offspring or were absent (two 119 offspring). For pair 7 (female 117/125, male 119/119, eight offspring), three offspring displayed genotype 119/125 (poorly amplified maternal 125) and five displayed only allele 119. Thus for pairs 2 and 7, maternal alleles seem to experience some difficulties in being amplified in the offspring. The same patterns held when the PCR was repeated, even with the asymmetric design. The DNA mixture of each pair of parents provided expected results (the bands from each parent could be seen). If one accepts the interpretation that maternal alleles of pairs 2 and 7 suffer from an amplification bias, the proportions observed actually correspond to Mendelian expectations ($P > 0.45$).

3.2. Methylation experiment

Mean efficiencies values were $Em = 0.99$ and $Eu = 0.97$ for the methylated and the un-methylated DNA, respectively. These values are not significantly different (Wilcoxon signed rank test for paired samples, $P = 0.63$). Clearly, there is no evidence that methylated and unmethylated templates are amplified with different efficiency.

3.3. Re-analysis of field data for locus IR27

The minimal model obtained was $f_i = \text{poly(AllelSize, 2)} + \text{Site} + \text{Constant}$ ($R^2 = 0.36$) (Table 1). Both factors (allele size and site) are highly significant. The shape of the dependency of heterozygote deficit on allele size can be seen in Fig. 1. Whatever the exact shape of the relationship, it is clear that the longer an allele the smaller the F_{is} estimate. No such results could be observed at any of the four other loci.

3.4. Modelling DNA amplification bias in heterozygous individuals

The results of this approach are summarised in Fig. 2 for a 100% effectiveness of amplification bias ($\alpha = 1$)

Table 1
Results of the generalised linear model for allelic heterozygote deficit in Locus IR27

	Df	Dev.	Res. Df	Res. dev	F	Pr(F)
Null			133	25.95		
Poly(AllelSize, 2)	2	4.12	131	21.84	15.21	0.0001
Site	8	5.18	123	16.65	4.79	0.0001

Df, degrees of freedom; Dev., deviance; Res., residual; F, the Fisher value of the analysis of variance; Pr(F), the P-value of the test; Null, the null model.

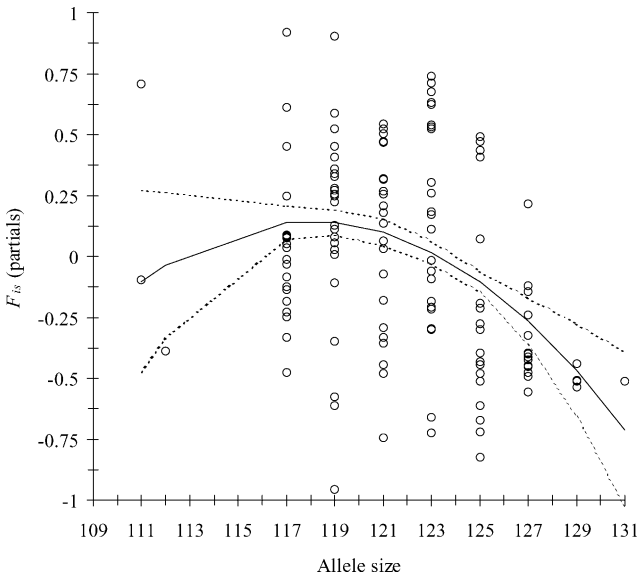

Fig. 1. Shape of the relationship between partials (corrected for the effect of site) of heterozygote deficits (F_{is}) (circles) and allele size extracted from the generalised linear modelling (see Table 1 and text for details) at locus IR27. The fit of the model (plain line) and the 95% confidence intervals (dashed lines) are represented. Note that removing alleles 111 and 112 from the data set did not change the shape of the curve and resulted in a better fit ($R^2 = 0.39$).

(lower values simply render the expected F_{is}s lower) and different frequency distributions of alleles (Table 2). There is a general tendency for high F_{is} values for the shortest alleles and a decrease as alleles get longer. The frequency distribution of alleles is critical on the expected F_{is} and may even reverse the trend of the curve (see Fig. 2).

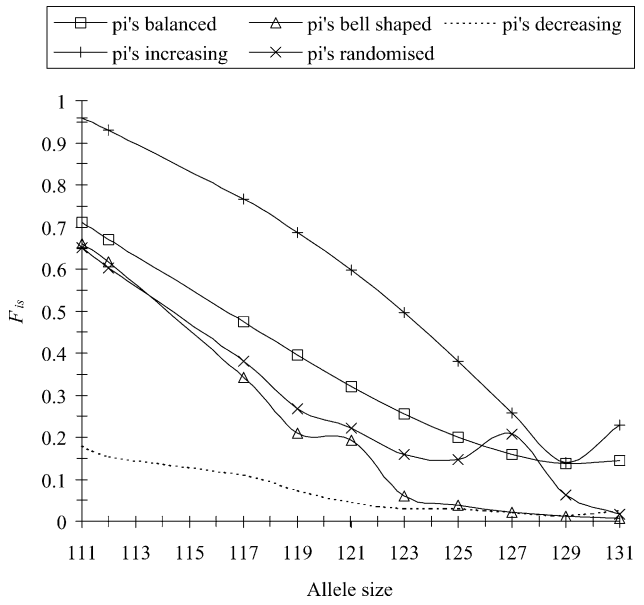

Fig. 2. Heterozygote deficit (F_{is}) expected values under the hypothesis of a more efficient amplification of the shortest allele in heterozygous DNA's during PCR, for different frequency distributions (p_i's) (see Table 2 for details) of the different alleles observed at locus IR27 in *Ixodes ricinus* populations. The effectiveness of the amplification bias was set to $\alpha = 1$.

Table 2
Theoretical distributions of allelic frequencies used to explore the model of influence of allele size on heterozygote deficits of locus IR27

Alleles	Balanced	Bell shaped	Decreasing	Increasing	Randomised
111	0.1	0.025	0.6000	0.0025	0.0100
112	0.1	0.055	0.2000	0.0025	0.0025
117	0.1	0.100	0.1000	0.0050	0.6000
119	0.1	0.160	0.0500	0.0100	0.0025
121	0.1	0.300	0.0200	0.0100	0.0200
123	0.1	0.170	0.0100	0.0200	0.0050
125	0.1	0.100	0.0100	0.0500	0.1000
127	0.1	0.050	0.0050	0.1000	0.2000
129	0.1	0.027	0.00250	0.2000	0.0500
131	0.1	0.013	0.00250	0.6000	0.0100

Under the hypothesis of 100% effectiveness of amplification bias ($\alpha = 1$) in favour of the shortest alleles, the correlation between the allelic F_{is}s observed in *I. ricinus* females from Switzerland and the expected one is highly significant ($R = 0.38$, $P = 0.007$). However, observed F_{is} are most of the time higher than the expected ones, especially so for the intermediate sized alleles (Fig. 3). Intermediate sized alleles are generally the most frequent ones, and thus the most often associated with null alleles that are known to be rather frequent at this locus (De Meeûs et al., 2002a).

4. Discussion

First, the low number of successful pairs obtained and the high proportion of genetically identical parents observed for the five loci (highly polymorphic, see De Meeûs et al., 2002a) may suggest that the nymphs collected to build the pairs were highly related.

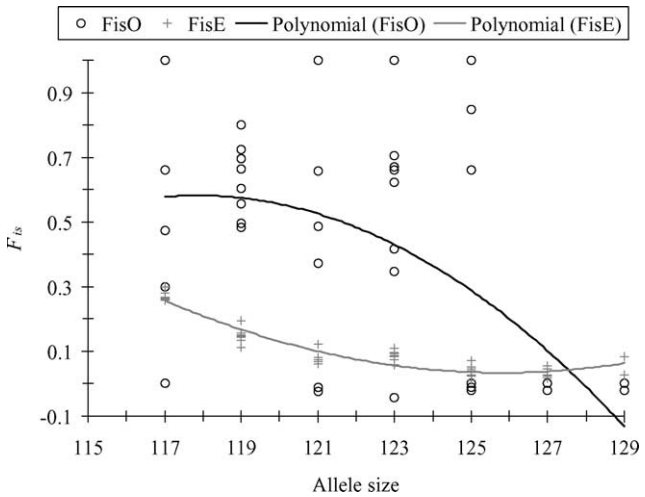

Fig. 3. Comparison between expected F_{is} (FisE) under the hypothesis of 100% effectiveness of amplification bias ($\alpha = 1$) in favour of the shortest alleles and the F_{is} observed in *Ixodes ricinus* females in Switzerland (FisO) at locus IR27. Expected values were estimated for each allele in each sample after allelic frequencies adjustments to the allelic frequencies observed in Switzerland female ticks (see text for more explanation).

Microsatellites are often considered to be Mendelian codominant markers (Jarne and Lagoda, 1996). Here, for most of the pedigrees that did not fit Mendelian expectations, null alleles and locus duplication seemed to explain our results. However, null alleles did not satisfactorily explain all the heterozygote deficits found in wild ticks (De Meeûs et al., 2002a). Thus, the null alleles and gene duplication interpretation does not exclude other phenomena that remain to be found.

One locus (IR27) displayed a 'parental imprinting like' behaviour in two pedigrees. Parental imprinting is a widespread phenomenon in living organisms (Chakraborty, 1989). Imprinting may result from methylation processes (Chakraborty, 1989; Barlow, 1993; LaSalle and Lalande, 1996; Jaenisch, 1997; Bestor, 2000) that may disturb DNA amplification (Gérard Cuny, personal communication). Our results show that if differences in the amplification efficiencies of unknown samples with supposedly identical sequences are observed, the reason for this does not lie in DNA methylation. However, we could not compare the interaction between allele size and DNA methylation in heterozygous individuals of our samples and pedigrees. To our knowledge, segregation distortions (Slettan et al., 1997; Chavarriaga-Aguire et al., 1998) and null alleles (Paetkau and Strobeck, 1995; Gullberg et al., 1997) were evidenced in pedigree studies on microsatellite markers of various organisms, but a parental imprinting-like behaviour was not reported until now, except for some human disease associated microsatellite loci (Petronis and Gottesman, 2000; Liu et al., 2001).

Another explanation may be that the flanking sequence of both alleles of both females of pedigrees 2 and 7 had some mutations that made them less competitive than allele 119 for amplification (almost null alleles). The alleles involved in these two parental imprinting-like pedigrees were observed to behave normally in the other pedigrees. Furthermore, asymmetric PCR (one primer in excess) did not reveal more heterozygous ticks from the wild at locus IR27 (De Meeûs et al., 2002a). Thus, this phenomenon does not seem to be allele-specific.

The hypothesis of amplification bias in favour of the shortest alleles (short allele dominance) appears as the best explanation for locus IR27 behaviour. To our knowledge, this phenomenon was only reported in one instance for one microsatellite locus of a red alga (Wattier et al., 1998).

The regression analysis and the model we built show that, in case of short allele dominance, the relationship between F_{is} and allele size is not expected to be linear, that it depends on allelic frequency distribution and that allele frequency estimation is highly biased.

The corresponding bias in F_{is} estimate will thus be very difficult to assess unless one can re-amplify the DNA with an optimisation protocol as in Wattier et al. (1998) (not possible here because all the DNA of our samples has

already been used). In our case, the mechanism seem more complicated than a simple competition during the PCR, as the mixture of the parents involved in the problematic pedigrees did not reveal any amplification problem for the longest allele as in Wattier et al. (1998). It is also worth noticing that in our case, the phenomenon seems very strong as it concerns a much narrower allele size range than in Wattier et al. (1998).

In pedigrees 2 and 7 for locus IR27, if we interpret the observed homozygous individuals as heterozygous for the longest allele (i.e. two 119/125 for pedigree 2 and five 119/125 for pedigree 7) the goodness of fit tests give $P = 0.18$ and 0.01, respectively. Given that more than 10 tests were undertaken for the pedigree experiments, $P = 0.01$ may have been obtained by chance and the short allele dominance hypothesis appears consistent with our result. Moreover, together with the important frequency of real null alleles (blanks) observed at locus IR27 (De Meeûs et al., 2002a), this phenomenon rather well explains the heterozygote deficits observed for that locus in the field.

The proximal (i.e. molecular) cause of this amplification bias, if methylation cannot be invoked, remains fairly mysterious and deserves further studies at the molecular level. This phenomenon, depending on its effectiveness, leads to erroneous F_{is} estimates and, which is worse, to erroneous allelic frequency estimates for the loci concerned as shown by our theoretical approach. This confirms there is a need to cautiously interpret heterozygote deficits found in the wild before parentage or reproductive systems inferences can be made. More generally, each time an unexpected heterozygote deficit is observed, a regression test between allele sizes and heterozygote deficits, correcting for the effect of other relevant factors (as the sampling location), should always be done to make sure no such phenomenon exists in the microsatellites in use. This conclusion is particularly relevant for parasites and infectious diseases for which the use of molecular markers often represents the only opportunity to assess their population biology and epidemiology (e.g. De Meeûs et al., 2002a,b; Prugnolle et al., 2002). Indeed, 'molecular tools are increasingly being used to address questions about parasite epidemiology' Constantine (2003).

Acknowledgements

An international program of scientific collaboration of the Centre National de la Recherche Scientifique (France, PICS No. 290) and Fonds National Suisse de la Recherche Scientifique (Switzerland, request No. 31.42919.95) financed this work. We would like to thank Lorenza Beati, Jerôme Goudet, Patrick Durand, Bruno Betchart and André Aeschlimann for their help, Michel Raymond who kindly provided the program 'Multinom' for multinomial tests, Christine Chevillon and Philippe Jarne who corrected an early version of this manuscript and two anonymous referees who helped to improve our manuscript. Thierry de Meeûs and François Renaud are supported by the CNRS.

References

Aesclimann, A., 1972. *Ixodes ricinus*, Linné, 1758 (Ixodoidea, Ixodidae). Essai préliminaire de synthèse sur la biologie de cette espèce en Suisse. Acta Tropica 29, 321–340.

Arezi, B., Xing, W., Sorge, J.A., Hogrefe, H.H., 2003. Amplification efficiency of thermostable DNA polymerases. Anal. Biochem. 321, 226–235.

Balmelli, T., Piffaretti, J.C., 1996. Analysis of the genetic polymorphism of *Borrelia burgdorferi* sensu lato by multilocus enzyme electrophoresis. Int. J. Syst. Bacteriol. 46, 167–172.

Banyay, M., Graslund, A., 2002. Structural effects of cytosine methylation on DNA sugar pucker studied by FTIR. J. Mol. Biol. 324, 667–676.

Barlow, D.P., 1993. Methylation and imprinting: from host defence to gene regulation. Science 260, 309–310.

Behe, M., Felsenfeld, G., 1981. Effects of methylation on a synthetic polynucleotide: the B–Z transition in poly(dG-m5dC).poly(dG-m5dC). Proc. Natl. Acad. Sci. USA 78, 1619–1623.

Bestor, T.H., 2000. The DNA methyltransferases of mammals. Hum. Mol. Genet. 9, 2395–2402.

Brookfield, J.F.Y., 1996. A simple method for estimating null allele frequency from heterozygote deficiency. Mol. Ecol. 5, 453–455.

Chakraborty, R., 1989. Can molecular imprinting explain heterozygote deficiency and hybrid vigor? Genetics 122, 713–717.

Chavarriaga-Aguire, P., Maya, M.M., Bonierbale, M.W., Kresovich, S., Fregene, M.A., Tohme, J., Kochert, G., 1998. Microsatellites in cassava (Manihot esculenta Crantz): discovery, inheritance and variability. Theor. Appl. Genet. 97, 493–501.

Constantine, C.C., 2003. Importance and pitfalls of molecular analysis to parasite epidemiology. Trends Parasitol. 19, 346–348.

Crawley, M.J., 2002. Statistical Computing. An Introduction to Data Analysis using S-Plus, Wiley, Chichester.

Crump, J.A., 2000. Lyme disease pitfalls and diagnosis. N. C. Med. J 61, 318–320.

Delaye, C., Aeschlimann, A., Renaud, F., Rosenthal, B., De Meeûs, T., 1998. Isolation and characterisation of microsatellite markers in the *Ixodes ricinus* complex (Acari: Ixodidae). Mol. Ecol. 7, 360–361.

Delmotte, F., Leterne, N., Simon, J.C., 2001. Microsatellite allele sizing: difference between automated capillary electrophoresis and manual technique. Biotechniques 31, 810–818.

De Meeûs, T., Beati, L., Delaye, C., Aeschlimann, A., Renaud, F., 2002a. Sex biased dispersal in the vector of Lyme disease, *Ixodes ricinus*. Evolution 59, 1802–1807.

De Meeûs, T., Renaud, F., Mouveroux, E., Reynes, J., Galeazzi, G., Mallié, M., Bastide, J.M., 2002b. The genetic structure of *Candida glabrata* populations in AIDS and non-AIDS patients. J. Clin. Microbiol. 40, 2199–2206.

Gullberg, A., Tegelström, H., Olsson, M., 1997. Microsatellites in the sand lizard (*Lacerta gracilis*): description, variation, inheritance, and applicability. Biochem. Genet. 35, 281–295.

Hartl, D.L., Clark, A.G., 1989. Principles of Population Genetics, 2nd ed., Sinauer Associates, Sunderland, MA.

Humair, P.F., Gern, L., 2000. The wild hidden face of lyme borreliosis in Europe. Microbes Infect. 2, 915–922.

Jaenisch, R., 1997. DNA methylation and imprinting: why bother? Trends Genet. 13, 323–329.

Jarne, P., Lagoda, J.L., 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11, 424–429.

Klump, H.H., Loffler, R., 1985. Reversible helix/coil transitions of left-handed Z-DNA structures. Comparison of the thermodynamic properties of poly(dG).poly(dC), poly[d(G–C)].poly[d(G–C)],

and poly(dG–m5dC).poly(dG–m5dC). Biol. Chem. Hoppe Seyler 366, 345–353.

LaSalle, J.M., Lalande, M., 1996. Homologous association of oppositely imprinted chromosomal domains. Science 272, 725–728.

Lehmann, T., Hawley, W.A., Kamau, L., Fontenille, D., Simard, F., Collins, F.H., 1996. Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: comparison of microsatellites and allozyme loci. Heredity 77, 192–208.

Liu, J., Nyholt, D.R., Magnussen, P., Parano, E., Pavone, P., Geschwind, D., Lord, C., Iversen, P., Hoh, J., The Autism Genetic Resource Exchange Consortium, Ott, J., Gilliam, C., 2001. A genomewide screen for autism susceptibility loci. Am. J. Hum. Genet. 69, 327–340.

Nathan, D., Crothers, D.M., 2002. Bending and flexibility of methylated and unmethylated EcoRI DNA. J. Mol. Biol. 316, 7–17.

Paetkau, D., Strobeck, C., 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. Mol. Ecol. 4, 519–520.

Pemberton, J.M., Slate, J., Bancroft, D.R., Barret, J.A., 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. Mol. Ecol. 4, 249–252.

Petronis, A., Gottesman, I.I., 2000. Psychiatric epigenetics: a new focus for the new century. Mol. Psychiatry 5, 342–346.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real time RT-PCR. Nucleic Acids Res. 29, 2002–2007.

Prugnolle, F., De Meeûs, T., Durand, P., Sire, C., Théron, A., 2002. Genetic evidence for sex-biased dispersal in a parasitic organism (*Schistosoma mansoni*). Mol. Ecol. 11, 1231–1238.

Rosenthal, A., Coutelle, O., Craxton, M., 1993. Large-scale production of DNA sequencing templates by microtitre format PCR. Nucleic Acids Res. 21, 173–174.

Slettan, A., Olsaker, I., Øystein, L., 1997. Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. Heredity 78, 620–627.

Wattier, R., Engel, C.R., Saumitou-Laprade, P., Valero, M., 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). Mol. Ecol. 7, 1569–1573.

Weir, B.S., 1990. Genetic Data Analysis, Sinauer Associates, Sunderland, MA.

Weir, B.S., Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38, 1358–1370.

Wright, S., 1965. The interpretation of population structure by F-statistics with special regard to system of mating. Evolution 19, 395–420.