Theor Appl Genet (2004) 108:368–377
DOI 10.1007/s00122-003-1443-5

ORIGINAL PAPER

A.-C. Thuillet · T. Bataillon · P. Sourdille · J. L. David

Factors affecting polymorphism at microsatellite loci in bread wheat [*Triticum aestivum* (L.) Thell]: effects of mutation processes and physical distance from the centromere

Received: 14 June 2003 / Accepted: 12 August 2003 / Published online: 16 October 2003

Abstract The effects of factors known to influence the level of polymorphism at microsatellite loci were studied using 99 markers and seven lines of bread wheat. Mutational factors as well as indirect selective events shape diversity at these loci. Theory predicts that the selection of favorable alleles should reduce polymorphism at neutral neighboring loci in genomic areas with low recombination rates. In wheat, local recombination rate is positively correlated with physical distance from the centromere. Seventy four loci among the 99 used could be physically located on the chromosome. We studied how the following affected the diversity among a set of inbred lines: the length of the alleles, the motif (CA versus CT), the structure of the loci (perfect versus imperfect) and the chromosomal position of the loci. For each locus, we determined whether the polymorphism observed at a locus was compatible with the Stepwise Mutation Model (SMM) or the Two-Phase Model (TPM). Both the mutation rate and the compatibility with the SMM or the TPM were shown to be variable between loci. Wheat microsatellite loci were found to be more variable when segregating alleles were perfect and had long motifs (composed of many repetitions). Diversity observed at 19 loci was not compatible with the SMM. Loci located in distal regions, with presumably high recombination rates, had longer allele sizes and were more polymorphic than loci located in proximal regions. We conclude that both mutation factors and indirect selective events vary according to the local recombination rate and therefore jointly influence the level of polymorphism at microsatellite loci in wheat.

Communicated by J. Dvorak

A.-C. Thuillet · T. Bataillon · J. L. David (✉)
INRA-UMR Diversité et Génomes des Plantes Cultivées,
Domaine de Melgueil, 34130 Mauguio, France
e-mail: david@ensam.inra.fr
Fax: +33-4-67293990

P. Sourdille
Station d'Amélioration des Plantes Domaine de Crouël,
INRA, Unité Génome, 63039 Clermont-Ferrand Cedex, France

Introduction

Positive indirect selection (Maynard-Smith and Haig 1974), as well as background selection against deleterious mutations, reduce locally the effective population size (Charlesworth et al. 1993; Nordborg et al. 1996). Neutral genetic diversity is hence expected to decrease in genomic regions with low local recombination rates. Polymorphism data based on either Restriction Fragment Length Polymorphisms (RFLPs) or on nucleotide polymorphism surveys, fit the prediction of reduced polymorphism in low recombination regions in *Drosophila* (Begun and Aquadro 1992), in humans (Nachman et al. 1998; Przeworski et al. 2000), in sea beets (Kraft et al. 1998), in several species of the genus *Aegilops* (Dvorak et al. 1998a) and to a lesser extent in tomatoes (Stephan and Langley 1998; Baudry et al. 2001). An alternative explanation for the positive correlation between nucleotide diversity and local recombination rates may also be that there is increased mutation rates in regions with high recombination rates (Marais et al. 2001; Lercher and Hurst 2002). Recombination is expected to create new RFLP profiles with a significant frequency (Timmermans et al. 1996). Hence, different types of markers would reveal different patterns of polymorphism across a local recombination gradient depending on their propensity to create new variants by recombination.

Microsatellite loci exhibit, on average, high mutation rates (Jarne and Lagoda 1996), ranging from 10^{-6} up to 10^{-2} mutation events per generation (Weber and Wong 1993; Schug et al. 1997; Thuillet et al. 2002). Conflicting findings were reported on the relationship between the levels of polymorphism for microsatellite loci and the variation in local recombination rates (Michalakis and Veuille 1996; Schlötterer et al. 1997; Schug et al. 1998; Payseur and Nachman 2000; Kauer et al. 2002). Because of high mutation rates, a positive correlation between the level of neutral polymorphism and recombination rate is expected to be observed under background selection but possibly not under positive selection, as the fast regeneration of polymorphism at a neutral locus through

mutation may erase the loss of diversity caused by positive selection at adjacent loci (Wiehe 1998). Therefore, microsatellite loci could potentially be used to assess the relative importance of background selection and hitchhiking in shaping the levels of diversity in different genomic regions. However, variations in the mutation rate and differences in the model of mutation among loci, may obscure the expected correlation between polymorphism and local recombination rate. First, new alleles can be generated through different types of mutation involving the gain/loss of a variable number of repeats in a single mutation event (Di Rienzo et al. 1994). Next, mutation rates at microsatellite loci are positively correlated with allele length, measured by the number or repeats in the microsatellite motif (Wierdl et al. 1997; Kruglyak et al. 1998; Schlötterer et al. 1998; Xu et al. 2000) and the number of bases in the repeated motif plays a role as well (Ellegren 1995; Chakraborty et al. 1997). Mutation rates can furthermore be modified depending on the structure of the locus (perfect versus interrupted or compound) (Brinkmann et al. 1998). Eventually, the way the diversity is measured in microsatellite loci depends on the model of mutation assumed for these loci. Under the Stepwise Mutation Model (SMM), mutant alleles are created by the addition or deletion of one repeat (Kimura and Otha 1978). In this case, the variance in allele length σ^2 at a locus can be used to estimate the diversity θ defined as $4N_e\mu$ with N_e the effective population size and μ the mutation rate. Indeed, $2\sigma^2$ is equal to θ under SMM (Moran 1975). Under alternative models, like the Two-Phase Model (TPM), mutant alleles also originate mostly from single steps, but sometimes are created from insertion or deletion of several repeated motifs (Di Rienzo et al. 1994). If we consider two loci with an identical scaled mutation rate, the variance in allele length will be larger for a TPM locus than for a SMM locus, but both loci are expected to have roughly the same number of alleles per locus (Feldman et al. 1999). Polymerase slippage during DNA replication is supposed to be the main mechanism of mutation in microsatellites (Tachida and Iizuka 1992), but recombination and gene conversion have also been shown experimentally to be involved in micosatellite mutations both in yeast (Jankowski and Nag 2002) and in wheat (Li et al. 2002). Variations in the rate and model of mutation for microsatellite loci could therefore be caused by variations in the recombination/gene conversion rates. If this is true, it will be difficult to differentiate the footprints of indirect selective events from mere differences in mutation rates.

In wheat, it has been shown for the B genome that local recombination rates increase dramatically with the physical distance from the centromere (Lukaszewski and Curtis 1993; Akhunov et al. 2003). Approximately 80% of crossing over occurs within the distal regions, representing about 20% of the physical length of chromosomes. Dvorak et al. (1998a) showed in selfing species of the related genus *Aegilops* that levels of polymorphism revealed by RFLP markers were dramatically lower in proximal regions than in distal regions. A recent study found a positive correlation between the genetic distance from the centromere and the number of alleles per microsatellite locus in hexaploid wheat using 26 loci (Huang et al. 2002). In our study, we examine the patterns of polymorphism existing at 99 loci among seven lines inbred of bread wheat. Our main objective was to investigate the causes of variation in polymorphism at microsatellite loci including both the intrinsic properties of the locus and their position along the recombination gradient of wheat chromosomes. The factors studied were the allele length, the motif of the repeated array, the structure of the repeated array (perfect vs interrupted), the genome of origin (A, B, D), the model of mutation (SMM vs TPM) and, when possible, the physical location on the chromosome was used as a predictor of local recombination rate.

Materials and methods

Plant material and microsatellite loci description

Ninety nine microsatellite loci, defined and genetically mapped by Röder et al. (1998a), were amplified by PCR from genomic DNAs of seven inbred lines of bread wheat (*Triticum aestivum* ssp. *aestivum* L.): Opata, Chinese Spring, Courtot, Renan, Recital, Eureka and Arche. DNA extraction and amplification conditions described by Röder et al. (1998a) were used. The polymorphism was revealed by polyacrylamide gel electrophoresis followed by silver staining. We determined the length of each allele in base pairs from which we deduced the number of repeats per allele. All loci had di-nucleotide motifs: 18 were $(CA)_n$, 48 were $(CT)_n$ and 33 were imperfect (i.e. presented an interruption in their repeated sequence) (Table 1).

Chromosomal location of the loci

The physical position of each locus was determined by finding common markers between physical maps, and different genetic maps obtained with the cross Synthetic Altar84/*Ae. squarrosa* #219 × Opata M 85 (Gill, Gill and Endo 1993; Hohmann et al. 1994; Delaney et al. 1995a, b; Mickelson-Young, Endo and Gill 1995; Nelson et al. 1995a, b, c; Van Deynze et al. 1995; Gill et al 1996a, b; Marino et al. 1996; Röder et al. 1998a, b; Biagetti et al. 1999; Boyko et al. 1999; Sarma et al. 2000; Qi and Gill 2001; Varshney et al 2001; Cabrera et al. 2002; Faris and Gill 2002; Weng and Lazar 2002; Zhou et al. 2002; see also http://www.graingenes.org). In some studies, the microsatellite markers studied here had been directly assigned to chromosome intervals using deletions lines (Endo and Gill 1996). In the other cases, close flanking genetic markers, whose physical assignment were known, were used to physically locate the microsatellite loci. When the two flanking markers belonged to the same chromosome interval, the microsatellite locus was assigned to this interval. If not, the most distant distal and proximal breakpoints between the chromosome intervals of the two flanking loci were retained as the boundaries of the chromosome interval in which the microsatellite was located. When a microsatellite had only a flanking marker in a genetic proximal or distal position, the interval was built, respectively, to span the chromosome from the proximal or distal breakpoint of the interval containing the flanking marker to the distal end or proximal end, respectively, of the chromosome arm. When needed, the genetic distance between markers was adjusted between studies differing in their number of lines using linear regression on common markers in the same chromosome arm. The chromosome intervals were expressed in a relative fraction of chromosome arm length. The physical position of each microsatellite locus was then estimated as

Table 1 Genomic position and properties of the microsatellite loci used in this study. L/S: long or short chromosome arm; FL: boundaries in a relative fraction length of the chromosome arm on which the locus was located (see text). Properties: first letter p/i for perfect or imperfect structure; second criterion: motif of the repeated array CT or CA, – for imperfect loci; third criterion: model of mutation TPM (t) or SMM (s); asterisk (*) means that the locus was used to study the relationship between different polymorphism indices and the physical distance from the centromere

Chrom.	location	Xgwm	FL	Properties				Chrom.	location	Xgwm	FL	Properties			
1A	L	99	0.83–1	p	CA	s	*	4A	L	160	0.66–1	p	CT	t	*
	S/L	135	0.52S–0.45L	p	CT	t	*		S	4	0.63–1	i	–	s	*
	S	164	0.5–0.81	p	CT	s	*		S	601	0.63–1	p	CT	s	*
	S	136	0.9–1	p	CT	t	*	4B	L	495	0.1–0.41	p	CT	s	*
1B	L	153	0.69–0.87	p	CT	t	*		L	513	0.1–0.41	p	CA	s	*
	L	124	0.85–0.87	i	–	t	*		S	107	–	p	CT	s	
	L	140	0.94–1	p	CT	t	*	5A	L	126	0.76–1	p	CA	s	*
	S	413	0.56–0.9	p	CT	s	*		L	179	0.76–1	p	CA	s	*
1D	L	458	0–0.76	p	CA	s			S	415	0.67–1	i	–	s	*
	L	232	0.76–0.87	p	CT	s	*	5B	L	371	0–0.69	i	–	s	
	L	642	0.76–0.87	p	CA	s	*		L	408	0.76–1	i	–	s	*
	S	337	0–0.81	i	–	s			L	604	0.76–1	p	CT	t	*
2A	L	312	0–0.77	p	CT	t			S	213	0.56–0.71	p	CT	t	*
	L	328	0–0.77	p	CA	s			S	335	0.56–0.71	i	–	s	*
	L	294	0.77–0.85	i	–	t	*		S	540	0.71–1	i	–	s	*
	L	445	0.76–1	p	CT	s	*		S	234	0.81–1	i	–	s	*
	S/L	122	0.47S–0.78L	i	–	s	*	5D	L	174	0–0.76	p	CT	s	
	S	10	0–0.47	i	–	s	*		L	182	0–0.76	p	CT	s	
	S	275	0–0.47	p	CT	s	*		L	271	0–0.76	i	–	s	
	S	339	0–0.47	p	CT	s	*		L	583	0–0.76	p	CA	s	
	S	425	0–0.47	p	CT	s	*		L	212	0.76–1	p	CT	s	*
	S	558	0–0.77	p	CA	s			L	272	0.76–1	p	CA	s	*
	S	372	0.36–0.78	p	CT	s	*		S	190	0.78–1	p	CT	s	*
	S	614	0.78–1	i	–	s	*		S	358	0.78–1	i	–	s	*
	S	636	0.78–1	i	–	s	*	6A	L	427	0.9–1	i	–	s	*
2B	L	55.2	0–0.69	i	–	s			L	570	0.9–1	i	–	s	*
	L	120	0.65–0.69	i	–	s	*		S	334	0.65–0.99	p	CT	s	*
	L	388	0.65–0.69	i	–	s	*	6B	L	626	0–0.4	i	–	t	*
	S	374	0.4–0.56	p	CA	t	*		L	219	0.4–0.76	i	–	s	*
	S	148	0.56–0.73	p	CA	t	*		S	193	0–0.25	i	–	t	*
	S	257	0.83–0.89	p	CA	s	*		S	361	0–0.25	i	–	s	*
	S	210	0.83–1	p	CT	s	*		S	70	0–0.25	i	–	s	*
2D	L	157	0.26–0.58	p	CT	s	*		S	613	0.68–1	p	CT	t	*
	L	539	0.66–0.76	p	CT	s	*	6D	S	469	0.79–1	p	CT	s	*
	L	311	0.94–1	p	CT	s	*		S	325	0.79–1	p	CT	s	*
	S	102	0.36–0.47	p	CT	s	*								
	S	261	0.41–0.89	p	CT	t	*								
3A	L	674	-	i	–	s		7A	S/L	260	0.88S–0.8L	p	CT	s	*
	S	2	-	p	CA	s			S	130	0.73–0.84	p	CA	s	*
3B	L	533	-	p	CT	t			S	233	0.84–1	p	CT	s	*
	L	181	0.8–1	p	CT	s	*		S	60	0.84–1	p	CA	s	*
	L	247	0.8–1	p	CT	t	*		S	635	0.84–1	i	–	s	*
	L	340	0.8–1	p	CT	s	*	7B	L	333	0–0.84	p	CT	s	
	S	285	0.2–0.33	p	CT	s	*		S	297	0.27–1	i	–	s	
	S	284	0.2–0.39	p	CT	s	*		S	400	0.73–1	p	CA	s	*
	S	566	0.2–0.39	i	–	s	*		S	537	0.73–1	i	–	s	*
3D	L	52	0–0.78	i	–	s		7D	L	44	0–0.1	p	CT	s	*
	L	645	0.65–1	i	–	s	*		L	437	0–0.14	p	CT	t	*
	L	664	0.65–1	p	CT	s	*		S	295	0–0.82	p	CT	s	
	S	2d	0–0.87	p	CA	s									
	S	161	0.39–1	p	CT	s									
	S?	341	0–0.87	p	CT	s									

the middle of the interval. When analysing the effect of the physical position, we discarded any locus for which the interval was longer than half of a chromosome arm. The main characteristics of the 99 microsatellite loci are summarised in Table 1.

Measures of diversity

Nei's heterozygosity, $He = 1 - \sum_i p_1^2$, where p_i is the frequency of the i^{th} observed allele. The allele length variance σ^2 was calculated as $\hat{\sigma}^2 = \frac{1}{6}\sum_{i=1}^{7}(r_i - r_\bullet)^2$, where r_i was the number of repeats in the

allele of the *i*th wheat line, *r.* the average of r_i over the seven lines. The number of different alleles at a locus detected in the sample was noted as *k*.

Data analysis

The effects of different factors on the three diversity statistics (*He*, σ^2 and *k*) were studied using linear regression or ANOVA (SAS 1999). These factors included the mean allele length, the structure of the alleles (perfect versus imperfect), the motif of the repeated array in the case of perfect loci (CA vs CT), the model of mutation (SMM vs TPM, see below), the genome of origin (A, B, D) and, when applicable, the physical location of markers. For this latter factor, the analysis was performed on diversity indexes and allele length, as well as on residues of a multivariate regression carried out to remove the effect of the other listed factors, i.e. the mutation model, the motif, structure and genome origin.

Determination of the model of mutation

The principle was to determine the mutation model of the loci using jointly the number of alleles and the variance in allele length observed in the sample of seven inbred lines. Under a strict SMM, the values of *k* and σ^2 follow a joint distribution that is completely specified by the genetic diversity θ ($4Ne\mu$) of the sampled population and the sample size (here seven). Under a TPM, *k* can take comparatively lower values when σ^2 has high values if mutations involving large changes in allele length occur at a reasonable frequency. The joint distribution of *k* and σ^2 observed in a sample of seven lines for loci mutating under a SMM can be obtained "empirically" by using a large set of simulated independent samples drawn from populations spanning a large range of θ values. Actual loci, whose observed paired values of *k* and σ^2 would not fall within a confidence area (comprising 99% of the probability mass of the simulated distribution), are considered as non-SMM mutating loci (or accepted as TPM mutating loci). Instead of simulating populations followed by the sampling of seven lines, a coalescent procedure was used to directly simulate samples drawn within populations undergoing complete selfing at mutation-drift equilibrium, with known θ values (Donnely and Tavaré 1995). The coalescent process has been shown to describe the genealogical history of genes in finite populations and to allow efficient simulations of numerous independent population samples (Hudson 1990). To determine the range of values of θ to simulate samples compatible with our real data set, we estimated its value on the 99 loci using $\hat{\theta}$, an estimator based on the expected number of alleles, $E(k)$, observed in a sample of size *n*. $\hat{\theta}$ is the solution of $E(k) = \sum_{i=0}^{n-1} \frac{\hat{\theta}}{\hat{\theta}+i}$ (Ewens 1972). $\hat{\theta}$ estimates ranged from 1 to 20. Therefore we simulated samples with values of θ ranging from 1 to 20 by a step of 1. We verified that higher values (until $\hat{\theta} = 100$) did not substantially change the results of our test (data not shown).

One thousand independent samples of seven individuals were simulated for every parametric value of θ. In each case, two mutation models were simulated: a perfect Stepwise Mutation Model to build the confidence area of the paired value of *k* and σ^2, and a Two-Phase Model (TPM) as an illustration. For simulating this TPM, a proportion of 0.80 (*p*) of the mutations were single steps as in a perfect SMM and a proportion of 0.2 increased or reduced the allele length of *x* repeats, where *x* was drawn at random from a geometric distribution of parameter α, whose value was set to 0.15. This value for α resulted in a mean variation of about six repeats. The values used for *p* and α were consistent with values proposed by Di Rienzo et al. (1994) to fit actual polymorphism data.

The confidence area was built as follows. Under the SMM and for the range of θ values used, the simulated samples were grouped according to their observed *k* values. For each value of *k*, a confidence interval, including 99% of the simulated σ^2 values of that group, was defined. A similar confidence interval was built under the simulated TPM. We rejected the SMM assumption for any locus with a σ^2 value lying out of the 99% confidence interval (given the observed *k* at this locus).

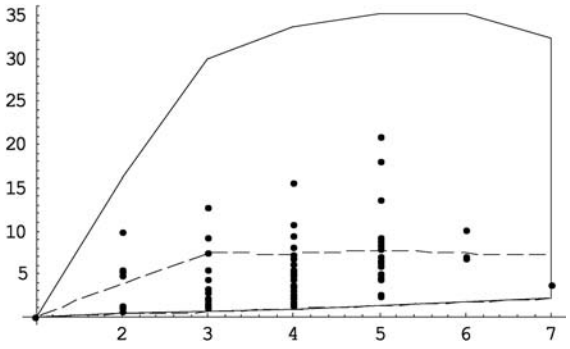

Fig. 1 Standard deviation in allele size (y axis) versus the number of alleles (x axis). Each point represents the actual data for a locus. The *dotted line* represents the space occupied by 99% of the expected values obtained by simulation under a SMM. The *full line* represents the space occupied by 99% of the expected values obtained by simulation under the TPM with specified parameters (see the text: P=0.8; α=0.15)

Results

Models of mutation

In our simulations, the allele-size variance (σ^2) increased with the number of alleles per locus (*k*), until a maximum of five alleles per locus, under both models of mutation. For each value of *k*, σ^2 was on average higher for the TPM than for the SMM and the region simulated under the TPM included the region simulated under the SMM (see Fig. 1). The standard deviation of allele sizes (σ) was used instead of the variance in allele length (σ^2) for graphical convenience in Fig. 1. Nineteen out of the 99 loci were lying outside of the 99% confidence area defined under the SMM and were located in regions compatible with the simulated TPM (Fig. 1). These 19 loci were considered to be TPM mutating loci while the remaining 80 were accepted as SMM mutating loci. The model of mutation was evenly distributed among perfect and imperfect loci (χ^2=1.59, *df*=1, P=0.20). The repartition of loci into these two classes (SMM vs TPM) explained 45% of the variation of σ^2 (ANOVA, $P<10^{-4}$), while no significant differences were observed in *k* or in *He* on the 99 loci. However, *He* and *k* values were higher in TPM perfect loci than in SMM perfect loci (*He*: 0.68 vs 0.54, P<0.03, and *k*: 4.13 vs 3.35, P<0.07). Using loci located in the physical interval spanning less than half a chromosome arm, no difference was detected in the repartition along the chromosome between the TPM and SMM loci, the average physical distance from the centromere of TPM loci being very similar to that of SMM loci (TPM 0.67 n=16, SMM 0.70 n=58).

Table 2 Effect of the motif on diversity parameters and on the allele length for 66 perfect microsatellites. *He*: Nei's heterozygosity, σ^2: variance in repeat number (see methods for definitions)

Item	*He*	Number of alleles (*k*)	σ^2	Number of repeats
CT (n=48)	0.62	3.79	47.1	23.4
CA (n=18)	0.49	2.83	25.2	20.1
R^2	0.07	0.12	0.02	0.04
P	<0.04	<0.005	0.31	0.11

Table 3 Effect of the physical distance from the centromere of microsatellite loci on their variation in mean allele length and heterozygosity level *He*, the number of alleles per locus *k* and the variance in allele length σ^2 (regression on 74 loci). The regression was carried out (1) on raw data, (2) on residual values (residuals I) of *He, k* and σ^2 obtained by ANOVA declaring as controlled factors structure of the locus, genome origin and model of mutation, and (3) on residual values of *He, k* and σ^2 obtained by an ANOVA with the same factors as "Residual I" and using allele length as a covariate

Effect of the physical distance from the centromere		*He*	*k*	σ^2	Mean length
Raw data	R^2	0.05	0.058	0.002	0.064
	P	*P*<0.06	*P*<0.04	*P*>0.65	<0.03
Residuals I	R^2	0.06	0.07	0.02	0.11
	P	<0.04	<0.02	>0.35	*P*<0.01
Residuals II	R^2	0.02	0.03	3.10^{-4}	
	P	>0.2	>0.15	>0.8	

Effect of the length of the alleles on the diversity

Among the 99 loci studied, a positive correlation was found between the mean number of repeats over the seven wheat lines (hereafter referred as allele length) and the diversity *He* (*P*<0.01, R^2=0.07), the number of alleles *k* (*P*<0.001, R^2=0.12) and the variance in allele length σ^2 (*P*<0.001, R^2=0.12). For perfect loci, allele length explained 13% and 14% of the total variation of *He* and *k*, respectively, and accounted for 31% of the variation of σ^2. Among imperfect loci, allele length was significantly correlated with only the number of alleles (*P*<0.03). Imperfect loci had longer alleles than perfect loci (ANOVA, 22.5 vs 29 repeats, *P*<0.001), but they did not have significantly higher *He*, *k* or σ^2 values than perfect loci, and these latter values showing even slightly higher values than imperfect loci.

Factors affecting the allele length

The microsatellites of the D genome were significantly shorter than those of the A and B genomes (B genome: 26.5 repeats, A: 25.6 and D: 21.1, ANOVA *P*<0.04, R^2=0.06) but had the same level of polymorphism since no differences were observed among the three genomes for *He*, *k* and σ^2 .

For the 99 loci, allele length was not significantly higher in TPM loci (28.4 repeats) than in SMM loci (23.8 repeats, *P*<0.08, using a Student test accounting for different variances in allele lengths in the two categories). However, when only perfect loci were considered, TPM loci had a longer allele length than SMM loci (27.6 repeats vs 21.1, *P*<0.02). Part of the differences in *He* and *k* between the TPM and SMM perfect loci reported above may be therefore explained by the difference in their allele length. Yet, according to regression between allele length and *k*, a difference of 6.5 repeats would account for a difference of only 0.4 alleles, i.e. approximately 50 % of the actual difference observed in *k* between TPM and SMM (0.78). Consequently, the model of mutation may cause variation in polymorphism level at microsatellite loci independently of its influence on the average allele length.

Motif affects perfect microsatellite loci diversity independently of allele length

In perfect loci, CA microsatellites had significantly fewer alleles and lower heterozygosity levels than CT microsatellites (Table 2). The relationship between allele length and locus diversity was similar for the two motifs: in CA microsatellites, the gain in *k* per added repeat was 0.077, which was similar to that of CT (0.044), both gains being significant. Hence, only a small part of the difference in *k* (0.17) between the motifs could be explained by their slight, but not significant, difference in allele length (2.8 repeats between CA and CT loci, Table 2). The actual difference in allele number between the motifs was around 0.96, i.e. 5-times higher. This suggests that the motif itself has an intrinsic effect on loci diversity.

Effect of the physical position of the locus on their mean allele length and polymorphism level

Seventy four loci were eventually located within a chromosome interval spanning less than half a chromosome (Table 1) and were kept for further analysis. On raw data, the physical distance from the centromere was significantly and positively correlated to the allele number and the allele length. The correlation was close

to significance for *He* (Table 3) and no effect of the physical distance was detected on σ^2. Yet, the variation in *He*, *k* and allele length was moderately explained by the distance from the centromere. The model of mutation, the motif, the structure and the genome origin of the allele, have been previously shown to be involved in the variation of *He*, *k*, σ^2 and allele length. Therefore we used a multifactorial analysis of variance to remove their effect. We first verified that (i) the model of mutation of the loci (TPM vs SMM), (ii) their motif and structure pooled in a unique factor (imperfect, perfect-CA, perfect-CT) and (iii) their genome origin (AB vs D) were not correlated with their physical location, thereby producing no confounding effect with physical distance from the centromere (ANOVAs, data not shown). The residual values of *He*, *k*, σ^2 and allele length obtained after the multifactorial analysis were then regressed against the physical distance from the centromere (Table 3). The correlation between physical distance from the centromere with residuals increased in comparison to raw data for all variables. Particularly, the correlation between physical distance from the centromere and allele length rose dramatically, as the determination coefficient of regression increased from R^2=0.06 to R^2=0.11. As allele length and physical distance appeared to have partly confounding effects on genetic diversity, the allele length effect on *He*, *k* and σ^2 was in turn removed before again regressing against the residual values for the physical distance from the centromere. The physical distance from the centromere no longer had any significant effect on *He* or *k*, the determination coefficient being reduced by more than 50% (Table 3). The opposite was not true as allele length still explained significant variation in *He*, *k* and σ^2 when the effects of the distance from the centromere and other factors were removed (data not shown). In brief, we conclude that the loci in more proximal regions were on average shorter, had fewer alleles and a lower heterozygosity level than loci located in more distal positions.

Discussion

In this study, a large number of microsatellite loci (99) were used to survey the polymorphism in a sample of only seven inbred lines of wheat. Diversity indexes, such as *He*, are known to be more sensitive to a low number of loci sampled than to a low number of individuals (see for instance Nei 1987, pp 179–181). Therefore, we believe that our conclusions regarding the factors influencing microsatellite diversity and allele length in wheat are likely to be robust.

Differences in the mutation model across loci

Using a simulation procedure, 19 out of 99 loci exhibited joint values of allele number and the variance in allele length that were not compatible with a Stepwise Mutation Model. Other tests allow detection of loci following a TPM, but they require independent knowledge of the demographic history of the populations sampled (Di Rienzo et al. 1994). In our case, the test assumed a mutation-drift equilibrium. The test could be inappropriate if this assumption is not true. First, wheat, considered as a population, could have experienced a bottleneck during modern breeding. In this case, the number of alleles *k* are expected to be reduced relative to both the heterozygosity level *He* and the allele length variance σ^2 (Nei et al. 1975; Garza and Williamson 2001). The SMM confidence area would have been over-estimated and some TPM loci misclassified as SMM. Alternatively, if wheat has undergone a dramatic demographic expansion through the spread of agriculture, the resulting increase in diversity would lead here to an underestimation of the SMM area and a spurious assignation of some loci to the TPM region. For a highly selected crop like bread wheat and despite a huge apparent demographic size, the number of effective reproductive individuals coming from breeding nurseries for the past hundred years should be much lower than the previous reproductive sizes achieved when wheat was grown as landraces (Huang et al. 2002), yet no dramatic reduction of diversity has been observed in elite material since the first decades of the 20th century (Donini et al. 2000; Christiansen et al. 2002). As our sample is mostly based on recent varieties, we consider that a recent genetic bottleneck is a more likely hypothesis to explain the patterns of diversity observed here. We therefore have probably detected a minimum number of loci evolving under a TPM.

Determining the mutation model of microsatellite loci allowed us to explain most of the variation in σ^2 detected among loci with similar number of alleles. If the proportion of loci evolving under a TPM found here is a good proxy for wheat and other organisms, estimators of diversity and genetic distances designed specifically for SMM microsatellites, as σ^2 is used here, should be employed carefully. Actually, the allele number per locus was the diversity measure most consistently and most strongly correlated with the controlled factors investigated compared to other measures of diversity. This was also observed in a previous study in wheat (Huang et al. 2002). This parameter should therefore deserve more attention in diversity studies.

From our data, no specific factor such as the physical distance from the centromere or the nature of the repeated array, clearly explained the distribution of TPM vs SMM loci. This suggests that high recombination regions are not prone to produce more TPM loci than low recombination regions. As perfect TPM microsatellite loci were longer and more polymorphic than perfect SMM loci, it is likely that their mutation rate is also higher.

Effect of allele length

It has been clearly shown from direct observations of mutation events in different organisms that the mutation rate of microsatellites increases with the length of the

repeated array (Wierdl et al. 1997; Schlötterer et al. 1998; Ellegren 2000a; Xu et al. 2000; Vigouroux et al. 2002). As an increase of polymorphism for loci with longer repeated arrays was observed here, this phenomenon also exists in bread wheat, our study being consistent with previous results obtained on 26 loci (Huang et al. 2002). Evidence for increased polymorphism at loci with longer alleles was found on perfect microsatellites but was less clear on imperfect ones, for which the number of alleles (but neither *He* nor σ^2) was weakly correlated with the allele size. As our sample contained only 33 imperfect loci, this could have prevented us from detecting the effect of length on their diversity level. This also suggests that the purity of the repeats plays an important role in the mutation process of microsatellites. Previous studies showed that an interruption in the repeated array of microsatellites tends to stabilise the loci (Brinkmann et al. 1998; Chigagure et al. 2000). For a given length, imperfect microsatellites are thus expected to have lower mutation rates and less polymorphism than perfect ones. In our case, imperfect loci were longer than perfect ones but the two categories had approximately the same level of polymorphism. This would mean that they effectively have a reduced mutation rate, compared to what would be expected given their length.

The microsatellite loci in the D genome carried shorter alleles on average than the loci located in the A and B genomes. No significant reduction in allele number or in Nei's heterozygosity was observed in this study, whereas in other studies the D genome had fewer alleles per locus than the A and B genomes in Huang et al. (2002) and more than the A genome in Zhang et al. (2002). The effect of allele length on mutation rate could explain part of the difference in genetic diversity observed in microsatellite loci between the three genomes of wheat. Consequently, mutation rates should be taken into account in studies using these markers to investigate the effects of bottlenecks due to the formation of hexaploid wheat through interspecific crosses between the diploid *Triticum tauschii* (D genome) and the tetraploid wheat *Triticum turgidum* (A and B genomes), as well as the effects of geneflows between *T. tauschii*, the tetraploid and hexaploid wheats (Dvorak et al 1998b).

Motifs modify the mutation rate across loci independently of the allele length

The 48 microsatellite loci with a CT motif had a higher number of alleles and level of heterozygosity than the 18 loci with a CA motif. These results are consistent with the trend reported in Huang et al. (2002) on four CA microsatellite loci. As the distribution of motifs was found to be homogenous with respect to the physical distance from the centromere, the reduction in polymorphism observed at loci with a CA motif is not likely to be due to the indirect effect of selection caused by recombination restriction. This suggests a difference in the mutation rate between the two types of microsatellites. An estimation of this difference was computed assuming that the mean effective size, *Ne*, of the two types of loci were similar within each study, even if they could vary between studies because of a much more extensive coverage of genetic diversity in Huang et al. (2002). Thus, if θ_{CA} is the estimation of genetic diversity for microsatellites with the CA motif and θ_{CT} that for the microsatellites with the CT motif, then the ratio θ_{CT}/θ_{CA} estimates the ratio of the mutation rates of the two types of loci. Using Ewens' estimator $\hat{\theta}$ based on the number of alleles in a sample of a given size (see methods), the ratio θ_{CT}/θ_{CA} was 2.1 in our study and 2.2 in the Huang et al. (2002) study. This consistent result suggests that CT microsatellites roughly mutate twice as fast as CA microsatellites in wheat. We furthermore suggest here that this difference in allele number was not likely to be explained by a difference in the length of the repeated array, since microsatellites with a CT motif were only slightly longer than microsatellites with a CA motif.

Recombination rate affects the diversity level and allele length

In wheat, the genetic distance from the centromere is not proportional to the physical distance from the centromere, since most of the recombination takes place in the distal part of the chromosomes (Lukazewski and Curtis 1993; Gill et al. 1996a, b; Akhunov 2003). Thus the local coefficient of exchange or recombination rate increases from the centromere to the distal portion of the chromosomes. In this study, microsatellite loci showed more genetic diversity in distal than in proximal regions. The number of alleles per locus was the diversity index most correlated with the physical distance from the centromere (R^2=0.07) and with the genetic distance from the centromere as well (R^2=0.23, Huang et al. 2002), whereas the heterozygosity level appeared less affected. This confirms that genetic diversity in wheat is reduced in low recombination areas compared with high recombination areas, as has been observed in closely relative *Aegilops* species using RFLP (Dvorak et al. 1998a) and other organisms (Begun and Aquadro 1992; Nachman et al. 1998; Kraft et al. 1998, Stephan and Langley 1998; Przeworski et al. 2000; Baudry et al. 2001).

As recombination is reduced in proximal regions of the chromosomes in wheat, indirect selection may lead to a reduction of the genetic diversity of the neutral loci located in this area (Nordborg et al. 1996). Nevertheless, proximal regions in wheat are considered to be less gene-rich than distal portions (Gill et al. 1996a, b), and as the frequency of selection events is likely to depend on the number of expressed genes, one could expect a decrease in the intensity of selection from the distal region to the proximal region of the chromosome. In this case, a low recombination rate would have a reduced impact on the decrease of neutral polymorphism. Actually, two arguments can be given to explain that selective pressure is

rather strong in proximal regions. First, there is only a shallow increase in the number of restriction fragments of expressed genes (ESTs) from the centromeric to the distal region (Akhunov et al. 2003). Second, in yeast, genes that are prone to give a lethal mutant, and hence submitted to strong purifying selection, are concentrated in low recombination areas (Pal and Hurst 2003). In wheat, single-copy genes are more often found in low recombination regions (Akhunov et al. 2003) than in high recombination regions, but no evidence has yet been reported that mutants of these genes would be more lethal than multi-copy genes.

An alternative but mutually non-exclusive explanation for the relationship between polymorphism and distance from the centromere may be that the mutation rate in microsatellite loci is higher in distal regions than in proximal ones. The microsatellites located in distal regions had longer alleles than loci in centromeric regions, the physical distance from the centromere still explaining 11% of the total variation of the residual allele length after accounting for the variation in allele length due to the model of mutation, structure, motif and genome origin. As longer alleles are expected to have comparatively higher mutation rates, distal regions should generate more mutant alleles at microsatellite loci, and consequently, should be more polymorphic, as observed. It is not clear if recombination by itself could play a role in lengthening the repeated array of microsatellite alleles. Genetic recombination and unequal crossing-over (Jankowski and Nag 2002) have been suggested as well as replication slippage (Tachida and Iizuka 1992) in mutation models of microsatellite alleles, but under such mechanisms, mutant alleles are not expected to become longer than the ancestral one. The repartition of the model of mutation of loci (SMM vs TPM) did not depend upon the physical distance from the centromere and thus could not account for the difference in allele length detected between proximal and distal regions.

From our data, we were not able to identify a single cause explaining the reduction of diversity in a low recombination region, differences in mutation rates being at least as likely as a selectively driven local reduction of population effective size in wheat.

We showed that diversity at microsatellite loci depends on numerous factors as diverse as the nucleotides composing the repeated array, their structure, their underlying model of mutation and their physical position in the chromosome. Some of these factors, like the motif, are clearly intrinsic properties of the locus, whereas others, like physical location, are external factors of variation. Other factors such as the allele length could be the result of the interaction of intrinsic and external factors, since allele length was found to depend on the model of mutation as well as the physical location. The integration of this knowledge in analysing patterns of diversity at microsatellite loci should improve the evolutionary interpretation of their polymorphism.

Acknowledgments We thank Eduard Akhunov, Karyn Deal and Jan Dvorak for helpful comments on the paper. This work was supported by a grant from the Bureau des Ressources Génétiques.

References

Akhunov ED, Goodyear JA, Geng S, Qi L, Echalier B, Gill BS, Lazo G, Chao S, Anderson OD, Linkiewicz AM, Dubcovsky J, La Rota M, Sorrells ME, Miftahudin, Zhang D, Nguyen HT, Gustafson PJ, Khwaja Hossain VK, Kianian SF, Peng J, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Choi DW, Close TJ, Dilbirligi M, Gill KS, Walker-Simmons KM, Steber C, McGuire PE, Qualset CO, Dvorak J (2003) The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res 13:753–763

Baudry E, Kerdelhue C, Innan H, Stephan W (2001) Species and recombination effects on DNA variability in the tomato genus. Genetics 158:1725–1735

Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Nature 356:519–520

Biagetti M, Vitellozzi F, Ceoloni C (1999) Physical mapping of wheat-*Aegilops longissima* breakpoints in mildew-resistant recombinant lines using FISH with highly repeated and low-copy DNA probes. Genome 42:1013–1019

Boyko EV, Gill KS, Mickelson-Young L, Nasuda S, Raupp WJ, Ziegle JN, Singh J, Hassawi DS, Fritz AK, Namuth D, Lapitan NLV, Gill BS (1999) A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. Theor Appl Genet 99:16–26

Brinkmann B, Klintschar M, Neuhuber F, Huhne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62:1408–1415

Cabrera A, Martin A, Barro F (2002) In-situ comparative mapping (ISCM) of *Glu-1* loci in *Triticum* and *Hordeum*. Chrom Res 10:49–54

Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri- and tetra-nucleotides microsatellite loci. Proc Natl Acad Sci USA 94:1041–1046

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303

Chigagure NN, Baxter GD, Barker SC (2000) Microsatellite loci of the cattle tick *Boophilus microplus* (Acari: Ixodidae). Exp Appl Acarology 24:951–956

Christiansen MJ, Andersen SB, Ortiz R (2002) Diversity changes in an intensively bred wheat germplasm during the 20th century. Mol Breed 9:1–11

Delaney DE, Nasuda S, Endo TR, Gill BS, Hulbert SH (1995a) Cytologically based physical maps of the group-3 chromosomes of wheat. Theor Appl Genet 91:780–782

Delaney DE, Nasuda S, Endo TR, Gill BS, Hulbert SH (1995b) Cytologically based physical maps of the group-2 chromosomes of wheat. Theor Appl Genet 91:568 573

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166–3170

Donini P, Law JR, Koebner RMD, Reeves JC, Cooke RJ (2000) Temporal trends in the diversity of UK wheat. Theor Appl Genet 100:912–917

Donnelly P, Tavaré S (1995) Coalescents and genealogical structure under neutrality. Annu Rev Genet 29:401–421

Dvorak J, Luo M, Yang Z (1998a) Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. Genetics 148:423–434

Dvorak J,·Luo MC, Yang ZL, Zhang HB (1998b) The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670

Ellegren H (1995) Mutation rates at porcine microsatellite loci. Mammal Genome 6:376–377

Ellegren H (2000a) Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genet 24:400–402

Endo TR, Gill BS (1996) The deletion stocks of common wheat. J Hered 87:295–307

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Pop Biol 3:87–112

Faris JD, Gill BS (2002) Genomic targeting and high-resolution mapping of the domestication gene Q in wheat. Genome 45:706–718

Feldman MW, Kumm J, Pritchard J (1999) Mutation and migration in models of microsatellite evolution. In: Goldstein DB, Schlötterer C (eds) Microsatellite evolution: inferences from population data. Oxford University Press, Oxford, pp 98–115

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. Mol Ecol 10:305–318

Gill KS, Gill BS, Endo TR (1993) A chromosome region-specific mapping strategy reveals gene-rich telomeric ends in wheat. Chromosoma 102:374–381

Gill KS, Gill BS, Endo TR, Boyko EV (1996a) Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. Genetics 143:1001–1012

Gill KS, Gill BS, Endo TR, Taylor T (1996b) Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. Genetics 144:1883–1891

Hohmann U, Endo TR, Gill KS, Gill BS (1994) Comparison of genetic and physical maps of group-7 chromosomes from *Triticum aestivum* L. Mol Gen Genet 245:644–653

Huang XQ, Börner A, Röder M, Ganal MW (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor Appl Genet 105:699–707

Hudson RR (1990) Gene genealogies and the coalescent process. Oxford Surveys Evolution Biol 7:1–44

Jankowski C, Nag K (2002) Most meiotic CAG repeat tract-length alterations in yeast are SPO11 dependent. Mol Gen Genet 267:64–70

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

Kauer M, Zangerl B, Dieringer D, Schlotterer C (2002) Chromosomal patterns of microsatellite variability contrast sharply in African and non-African populations of *Drosophila melanogaster*. Genetics 160:247–256

Kimura M, Otha T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc Natl Acad Sci USA 75:2868–2872

Kraft T, Sall T, Magnusson-Rading I, Nilsson NO, Hallden C (1998) Positive correlation between recombination rates and levels of genetic variation in natural populations of sea beet (*Beta vulgaris* subsp. *maritima*). Genetics 150:1239–1244

Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc Natl Acad Sci USA 95:10774–10778

Lercher MJ, Hurst LD (2002) Human SNP variability and mutation rate are higher in regions of high recombination. Trends Genet 18:337–340

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Biol Evol 11:2453–2465

Lukaszewski AJ, Curtis CA (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor Appl Genet 85:121–127

Marais G, Mouchiroud D, Duret L (2001) Does recombination improve selection on codon usage? Lessons from nematode and fly complete-genomes. Proc Natl Acad Sci USA 90:5688–5692

Marino CL, Nelson JC, Lu YH, Sorrells ME, Leroy P, Tuleen NA, Lopes CR, Hart GE (1996) Molecular genetic maps of the group-6 chromosomes of hexaploid wheat (*Triticum aestivum* L. em. Thell). Genome 3:359–366

Maynard-Smith J, Haig D (1974) The hitchicking effect of a favourable gene. Genetic Res 23:23–35

Michalakis Y, Veuille M (1996) Length variation of CAG/CAA trinucleotide repeats in natural populations of *Drosophila melanogaster* and its relation to the recombination rate. Genetics 143:1713–1725

Mickelson-Young L, Endo TR, Gill BS (1995) A cytological ladder-map of the wheat homeologous group-4 chromosomes. Theor Appl Genet 90:1007–1011

Moran PA (1975) Wandering distributions and the electrophoretic profile. Theor Pop Biol 8:318–330

Nachman MW, Bauer VL, Cromwell SL, Aquadro CF (1998) DNA variability and recombination rates at X-linked loci in humans. Genetics 150:1133–1141

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution 29:1–10

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Lu YH, Nègre S, Bernard M, Leroy P (1995a) Molecular mapping of wheat. Homoeologous group 3. Genome 38:525–533

Nelson JC, Sorrells ME, Van Deynze AE, Lu YH, Atkinson M, Bernard M, Leroy P, Faris JD, Anderson JA (1995b) Molecular mapping of wheat. Major genes and rearrangements in homoeologous groups 4, 5 and 7. Genetics 141:721–731

Nelson JC, Van Deynze AE, Autrique E, Sorrells ME, Lu YH, Merlino M, Atkinson M, Leroy P (1995c) Molecular mapping of wheat. Homoeologous group 2. Genome 38:516–524

Nordborg M, Charlesworth B, Charlesworth D (1996) The effect of recombination on background selection. Genet Res 67:159–174

Pal C, Hurst RD (2003) Evidence for co-evolution and gene order and recombination rate. Nature Genet 33:392–395

Payseur BA, Nachman MW (2000) Microsatellite variation and recombination rate in the human genome. Genetics 156:1285–1298

Przeworski M, Hudson RR, DiRienzo A (2000) Adjusting the focus on human variation. Trends Genet 16:296–302

Qi LL, Gill BS (2001) High-density physical maps reveal that the dominant male-sterile gene *Ms3* is located in a genomic region of low recombination in wheat and is not amenable to map-based cloning. Theor Appl Genet 103:998–1006

Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998a) A microsatellite map of wheat. Genetics 149:2007–2023

Röder MS, Korzun V, Gill BS, Ganal MW (1998b) The physical mapping of microsatellite markers in wheat. Genome 41:278–283

Sarma RN, Fish L, Gill BS, Snape JW (2000) Physical characterization of the homoeologous Group-5 chromosomes of wheat in terms of rice linkage blocks, and physical mapping of some important genes. Genome 43:191–198

SAS (1999) SAS/STAT user's guide, version 8.1. Cary, North California

Schlötterer C, Ritter R, Harr B, Brem G (1998) High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. Mol Biol Evol 15:1269–1274

Schlötterer C, Vogl C, Tautz D (1997) Polymorphism and locus-specific effects on polymorphism at microsatellite loci in natural *Drosophila melanogaster* populations. Genetics 146:309–320

Schug MD, Mackay TF, Aquadro CF (1997) Low mutation rates of microsatellite loci in *Drosophila melanogaster*. Nature Genet 15:99–102

Schug MD, Wetterstrand KA, Gaudette MS, Lim RH, Hutter CM, Aquadro CF (1998) The distribution and frequency of microsatellite loci in *Drosophila melanogaster*. Mol Ecol 7:57–70

Stephan W, Langley CH (1998) DNA polymorphism in *Lycopersicon* and crossing-over per physical length. Genetics 150:1585–1593

Tachida W, Iizuka M (1992) Persistence of repeated sequences that evolve by replication slippage. Genetics 131:471–478

Timmermans MCP, Prem DO, Messing J (1996) Characterization of a meiotic crossover in Maize identified by a restriction fragment length polymorphism-based method. Genetics 143:1771–1783

Thuillet AC, Bru D, David JL, Roumet P, Santoni S, Sourdille P, Bataillon T (2002) Direct estimation of mutation rate for ten microsatellite loci in durum wheat, *Triticum turgidum* (L.) Thell. ssp *durum* desf. Mol Biol Evol 19:122–125

Van Deynze AE, Dubcovsky J, Gill KS, Nelson JC, Sorrells ME, Dvorak J, Gill BS, Lagudah ES, McCouch SR, Appels R (1995) Molecular-genetic maps for group-1 chromosomes of Triticeae species and their relation to chromosomes in rice and oat. Genome 38:45–59

Varshney RK, Prasad M, Roy JK, Roder MS, Balyan HS, Gupta PK (2001) Integrated physical maps of 2DL, 6 BS and 7DL carrying loci for grain protein content and pre-harvest sprouting tolerance in bread wheat. Cereal Res Commun 29:33–40

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JS, Doebley J (2002) Rate and pattern of mutation at microsatellite loci in maize. Mol Biol Evol 19:1251–60

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Human Mol Genet 2:1123–1128

Weng Y, Lazar MD (2002) Comparison of homeologous group-6 short arm physical maps of wheat and barley reveals a similar distribution of recombinogenic and gene-rich regions. Theor Appl Genet 104:1078–1085

Wiehe T (1998) The effect of selective sweeps on the variance of the allele distribution of a linked multiallele locus: hitch-hiking of microsatellites. Theor Pop Biol 53:272–283

Wierdl M, Dominska M, Petes TD (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics 146:769–779

Xu X, Peng M, Fang Z (2000) The direction of microsatellite mutations is dependent upon allele length. Nature Genet 24:396–399

Zhang XY, Li CW,·Wang LF, Wang·HM, You GX, Dong YS (2002) An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large-scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production. Theor Appl Genet 106:112–117

Zhou WC, Kolb FL, Bai GH, Shaner G, Domier LL (2002) Genetic analysis of scab resistance QTL in wheat with microsatellite and AFLP markers. Genome 45:719–727