Theor Appl Genet (2000) 101:1029–1038

Y.C. Li · T. Fahima · J.H. Peng · M.S. Röder
V.M. Kirzhner · A. Beiles · A.B. Korol · E. Nevo

Edaphic microsatellite DNA divergence in wild emmer wheat, *Triticum dicoccoides*, at a microsite: Tabigha, Israel

Received: 4 February 2000 / Accepted: 31 March 2000

Abstract Twenty eight microsatellite markers were used to analyze genetic divergence in tandem dinucleotide repeated DNA regions between two edaphic subpopulations of *Triticum dicoccoides* growing on the contrasting terra rossa and basalt soils from a microsite at Tabigha, north of the Sea of Galilee, Israel. The terra rossa soil niche was drier and more stressful than the basalt throughout the growing season (November to May). Significant microsatellite divergence in allele distribution, repeat length, genetic diversity, and linkage disequilibria were found between emmer wheat from the two soil types over two short transects of 100 m each. Soil-*specific and -unique* alleles and linkage disequilibria were observed in the terra rossa and basalt subpopulations. A permutation test showed that the effects of random genetic drift were very low for the significant genetic diversity at microsatellite loci between the two subpopulations, suggesting that an adaptive molecular pattern derived by edaphic selection may act upon variation of the microsatellites.

Key words Genetic diversity · Edaphic selection · Microsatellite · Microsite divergence · Wild emmer wheat · *Triticum dicoccoides*

Introduction

Macro- and micro-geographic variations in genetic polymorphisms of plant and animal populations have been extensively studied (reviewed in Nevo 1988, 1998 Linhart and Grant 1996). Wild emmer wheat, *Triticum dicoccoides*, is the allotetraploid progenitor of all bread wheats (Feldman 1976). Previous allozyme (reviewed in Nevo 1988, 1998) and RAPD-DNA (Nevo 1998) microgeographic studies for *T. dicoccoides* and other cereals have shown significant nonrandom adaptive molecular genetic differentiation at single and multilocus structures in either protein-coding regions or randomly amplified polymorphic DNAs among microecological environments in contrasting soil types or topography.

Microsatellite DNAs, tandem repeats of short oligonucleotides, have been widely used for population genetics and ecology (Saghai-Maroof et al. 1994; Terauchi and Konuma 1994; Awadalla and Ritland 1997; Innan et al. 1997; Van Treuren et al. 1997) because of their abundant high polymorphism across genomes, in particular the noncoding regions of genomes. Some authors have sought to explain the ubiquitous occurrence of microsatellites in terms of a functional significance (e.g. Stallings et al. 1991; Kashi et al. 1997). By contrast, most models of simple repeat evolution, in the absence of close linkage to a selected locus (''genetic hitchhiking'', Slatkin 1995), assume selective neutrality (Tachida and Iizuka 1992; Shriver et al. 1993; Valdes et al. 1993; Di Rienzo et al. 1994). In fact, allelic sizes are tightly constrained (Garza et al. 1995; Nauta and Weissing 1996). The levels of variation varied in different microsatellite loci and different populations (e.g. Saghai-Maroof et al. 1994; Innan et al. 1997). Mutational biases or selective constraints on allele size may truncate and converge allelic distributions in divergent populations or species (Garza et al. 1995; Slatkin 1995; Amos et al. 1996; Nauta and Weissing 1996; Primmer et al. 1996). Some authors suggested that natural selection controls the level of microsatellite variation (e.g. Harding et al. 1992; Epplen et al. 1993; Stephan and Cho 1994; Garza et al. 1995; Innan et al. 1997). These previous studies did not prove whether microsatellite loci are a selected target or just a convenient molecular marker. However, contradictory interpretations for microsatellite pattern suggest the need for deeper research on microsatellite variation in natural populations on a microgeographical scale.

Communicated by H.F. Linskens

Y.C. Li · T. Fahima · J.H. Peng · V.M. Kirzhner · A. Beiles
A.B. Korol · E. Nevo (✉)
Institute of Evolution, University of Haifa, Haifa 31905
e-mail: nevo@research.haifa.ac.il, Fax: +972-48246554

M. Röder
Institute for Plant Genetics and Crop Plant Research,
Corrensstrasse 3, 06466 Gatersleben, Germany

This study analyzed genetic diversity in DNA dinucleotide microsatellites of two edaphic subpopulations of *T. dicoccoides* from two contrasting soil types, terra rossa and basalt, at the Tabigha microsite in northern Israel, demonstrating microgeographic molecular-genetic divergence caused by ecological stress and edaphic selection.

Materials and methods

Habitat of wild emmer wheat

T. dicoccoides (genomic constitution AABB) is distributed all over the Fertile Crescent, but its center of distribution is in the Upper Jordan Valley and its surroundings (Nevo and Beiles 1989). It grows mainly on *basaltic* and *terra rossa* soil types and often grows together with wild barley and wild oats. The studied microgeographical site is at Tabigha near the Sea of Galilee, at the Mediterranean Sea level (Nevo et al. 1988). The red terra rossa and brown basalt soils were weathered from Middle-Eocene hard limestone and Upper-Pleistocene basalt flows, respectively (Atlas of Israel 1970). The brown basalt soil is subject to extremes of high and low water content due to its montmorinolitic clays. Hence, its water regime is more variable in terms of water content than that of the kaolinitic terra rossa soil, thereby representing a wider ecological niche than terra rossa. The terra rossa soil layer is also shallower, and numerous bare rocks predominate over the terrain with pockets of soil varying in depth. The topography of the terra rossa is more hilly than that of the flat, deeper soil of the southwestern dipping basalt. The terra rossa soil niche appeared to be clearly drier, narrower, and more stressful than the basalt in terms of the availability of water and moisture throughout the growing season (November to May). The differences were directly reflected in the different plant formations on the two soil types, which display long-term biotic differentiation (see Nevo et al. 1988). On terra rossa, the vegetation, including wild emmer wheat, was sparser, lower in height, and earlier in both ripening and drying than vegetation on the basalt.

Sampling

Sampling for this study was conducted in 1985 in two transects 200 m apart (Nevo et al. 1988). Each transect was 100 m long and equally subdivided into 50 m of red terra rossa and 50 m of brown basalt separated by a sharp geological boundary. Seeds were collected from 50 plants, about 1-m apart in each half of a transect. In the present study, 155 individuals from the two transects (76 and 79 from the terra rossa and basalt, respectively) were analyzed for their microsatellite variation.

PCR assay of microsatellites

Genomic DNA was extracted from seedlings using the method of Junghans and Metzlaff (1990). Primer pairs of the 28 dinucleotide microsatellites used in this study were described by Röder et al. (1995, 1998). Table 1 presents the repetitive motifs and chromosomal locations of these microsatellites. The procedure used to detect microsatellite polymorphism followed Fahima et al. (1998). Fragment sizes were calculated using the Fragment Manager (Pharmacia) computer program by comparison with internal size standards, which were added to each lane in the loading buffer.

Data analysis

The core sequence repeat number for each allele was calculated according to the fragment sizes and the number of repeat units at the corresponding loci in Chinese Spring (Table 1). Gene diversity (*He*; Nei 1973) of emmer wheat in the terra rossa and basalt soils was estimated for each locus. The genetic distance between the terra rossa and basalt subpopulations was measured by the allele-sharing distance (D_{AS}; Chakraborty and Jin 1993) defined for hypervariable markers. It has been used for microsatellite data, for example, in humans (Bowcock et al. 1994) and honeybees (Estoup et al. 1995). The programs *Microsat* (Goldstein et al. 1995), *POPGENE* (Yeh et al. 1997), and *STATISTICA* (Statsoft 1993) were used to perform statistical analyses. The χ^2-test was used to test the homogeneity of allele frequencies at loci between the two subpopulations and to test the significance of linkage disequilibria. Since repeat numbers at microsatellite loci are not distributed normally in our studied population, a nonparametric statistic, the Kruskal-Wallis (K-W) test, was used to test the difference in repeat number between subpopulations. Hartley's *F*-ratio was used to test the homogeneity of variances in repeat number at each locus between the two subpopulations. The log-linear model test was used to test the genetic and soil effects on linkage disequilibria.

Table 1 Microsatellite loci screened, core sequence, and chromosomal locations of microsatellite loci in Chinese Spring. A, B = genome A, B; S = short arm; L = long arm; imp = imperfect

Marker	Repeat type	Chromosome location
GWM136	$(CT)_{58}$	1AS
GWM99	$(AC)_{21}$	1AL
GWM95	$(AC)_{16}$	2AS
GWM294	$(GA)_9TA(GA)_{15}$	2AL
GWM218	$(CT)_{18}$	3AS
GWM162	$(CA)_{14}AA(CA)_4$	3AL
GWM601	$(CT)_{17}$	4AS
GWM637	$(CA)_{18}$	4AL
GWM415	$(GA)_{25}$imp	5AS
GWM186	$(GA)_{26}$	5AL
GWM459	$(GA)_{>28}$	6AS
GWM169	$(GA)_{23}$	6AL
GWM60	$(CA)_{30}$	7AS
GWM332	$(GA)_{36}$	7AL
GWM18	$(CA)_{17}GA(TA)_4$	1BS
GWM124	$(CT)_{27}(GT)_{18}$imp	1BL
GWM429	$(CT)_{25}$	2BS
GWM120	$(CT)_{11}(CA)_{18}$	2BL
GWM389	$(CT)_{14}(GT)_{16}$	3BS
GWM340	$(GA)_{26}$	3BL
GWM368	$(AT)_{25}$	4BS
GWM251	$(CA)_{28}$	4BL
GWM540	$(CT)_3(CC)(CT)_{16}$	5BS
GWM408	$(CA)_{>22}(TA)(CA)_7(TA)_9$	5BL
GWM361	$(GA)_{20}$imp	6BS
GWM219	$(GA)_{35}$imp	6BL
GWM537	$(CA)_{18}(TA)_{13}$	7BS
GWM577	$(CA)_{14}(TA)_6$	7BL

Permutation test

A permutation test was performed for the observed soil-*specificity* and soil-*uniqueness* of alleles at each locus and for linkage disequilibria (LD) between loci. For this analysis, rare alleles (with five copies in the whole Tabigha population) were excluded from consideration. The permutation test was conducted to determine the effects of random sampling. For that purpose, genotypes within each pair of soil subpopulations were randomly shuffled 5000-times while maintaining the initial sample sizes.

Results

Distribution of allelic diversity at polymorphic microsatellite loci

All individuals of *T. dicoccoides* were considered homozygous due to selfing. In this study, all 28 microsatellite primer pairs could produce DNA amplifications except for one marker, GWM637. Of the 27 markers, each of 26 amplified a single fragment in each individual, regarded as 26 loci; one marker, GWM332, amplified two fragments in many individuals, and we regarded the two fragments as two alleles at two loci (GWM332a and 332b). For the total 28 loci detected in this study, only GWM601 was monomorphic for emmer wheat in both soils. With the exception of the locus GWM332b (due to a ''null allele(s)''), the total alleles amplified over 27 loci were 170 and 177 in the terra rossa (TR) and basalt (BA) subpopulations, respectively; 142 alleles were shared between the two subpopulations (Table 2). The allele distributions at microsatellite loci were not equal between the two edaphic subpopulations. For example, Fig. 1 illustrates different allele distributions at five microsatellite loci between the closest sampling points in the terra rossa and basalt soils. There were distinct alleles and a higher allelic diversity in the basalt section than in the terra rossa section. In particular, GWM332a and GWM332b were completely linked on the long arm of chromosome 7A in the population of *T. dicoccoides* × *Triticum durum* (Peng et al., unpublished data). At GWM332a, similar alleles were detected in all individuals of wild emmer in both subpopulations. At GWM332b, however, soil-*unique* alleles (including a ''null'' allele) were observed (Fig. 2). This locus was not included in other genetic analyses because it is hard to estimate the repeat number of the null allele. These facts suggest that the locus GWM332b was a duplication of the common locus GWM332a. According to the allele distribution in the terra rossa and basalt subpopulations, we could expect that some soil-*specific* mutation occurred in the tandem repeat region and also in the flanking region of GWM332b, particularly in the subpopulation dwelling on basalt soil.

Permutation test

We assumed, as a critical measure, that soil-niche did not affect microsatellites, and that allele frequencies should

Table 2 Number of alleles, genetic diversity (*He*; Nei 1973), and allele-sharing distance (D_{AS}) at microsatellite loci of *T. dicoccoides* between terra rossa (TR) and basalt (BA) soil types at the Tabigha microsite, Upper Galilee, Israel

Locus	Number of alleles[a]				ARN		K-W test H(1, *N*)	Variance		Hartley's *F* (df_1, df_2>70)	*He*		D_{AS}
	Total	TR	BA	Shared	TR	BA		TR	BA		TR	BA	
GWM018	7	1	2	4	32.9	33.2	16.078***	28.55	2.53	11.3****	0.731	0.789	0.519
GWM060	9	2	2	5	21.2	22.7	1.522	16.36	37.86	2.3***	0.713	0.762	0.349
GWM095	6	0	1	5	15.5	17.6	32.766****	3.88	3.13	1.2	0.587	0.775	0.650
GWM099	3	0	1	2	7.7	8.4	4.265*	0.22	8.21	37.3 ****	0.441	0.369	0.138
GWM120	10	0	3	7	32.7	33.9	6.612**	4.26	7.83	1.8*	0.681	0.722	0.298
GWM124	12	3	1	8	51.6	53.7	7.825**	20.30	27.78	1.9*	0.827	0.865	0.470
GWM136	11	3	1	7	62.0	58.2	0.813	61.45	76.76	0.3	0.179	0.675	0.449
GWM162	6	0	0	6	23.6	22.4	21.252****	3.08	2.42	1.1	0.700	0.769	0.523
GWM169	8	1	0	7	23.1	22.0	0.381	6.03	2.19	2.7***	0.777	0.760	0.142
GWM186	8	3	1	4	23.9	22.7	1.139	36.36	9.24	3.9***	0.714	0.780	0.400
GWM218	12	2	1	9	21.6	23.2	0.191	56.34	26.75	2.1**	0.813	0.874	0.487
GWM219	6	1	2	3	22.1	26.1	12.244***	5.68	36.22	6.4****	0.644	0.751	0.336
GWM251	8	2	3	3	23.4	18.6	33.664****	41.46	34.91	1.2	0.670	0.587	0.533
GWM294	9	2	0	7	20.7	19.3	9.383**	15.48	8.41	1.8*	0.800	0.819	0.584
GWM332	5	0	1	4	20.4	20.3	0.091	1.45	2.74	1.9*	0.653	0.692	0.232
GWM340	12	0	5	7	27.0	29.1	9.491**	32.39	38.47	1.2	0.659	0.868	0.559
GWM361	6	1	1	4	22.0	21.6	11.343***	1.56	0.66	2.4***	0.643	0.562	0.465
GWM368	7	1	0	6	23.0	22.9	0.412	9.88	12.33	1.2	0.715	0.780	0.264
GWM389	8	1	0	7	32.6	31.7	0.694	11.73	3.37	3.5***	0.794	0.760	0.252
GWM408	6	0	1	5	28.2	25.0	25.306****	9.39	9.05	0.1	0.719	0.647	0.557
GWM415	2	0	0	2	19.6	19.1	10.565**	2.16	0.20	10.6****	0.266	0.025	0.145
GWM429	6	1	2	3	15.6	15.1	4.735*	19.43	26.04	1.3	0.606	0.609	0.187
GWM459	9	0	1	8	25.3	26.2	2.701	16.95	9.39	1.8*	0.834	0.845	0.299
GWM537	16	4	6	6	33.9	36.3	0.470	11.28	50.22	4.5****	0.804	0.885	0.726
GWM540	4	0	0	4	18.1	17.4	10.711***	2.13	0.81	2.6***	0.645	0.632	0.177
GWM577	8	0	0	8	28.7	34.4	20.093****	31.53	66.18	2.1**	0.665	0.834	0.537
GWM601	1	0	0	1							0.000	0.000	0.000
Total	205	28	35	142									
Mean	7.1	1.0	1.3	5.3							0.661	0.683	0.421

*, **, ***, **** = $P < 0.05, 0.01, 0.001, 0.0001$

[a] TR = terra rossa-unique, BA = basalt-unique

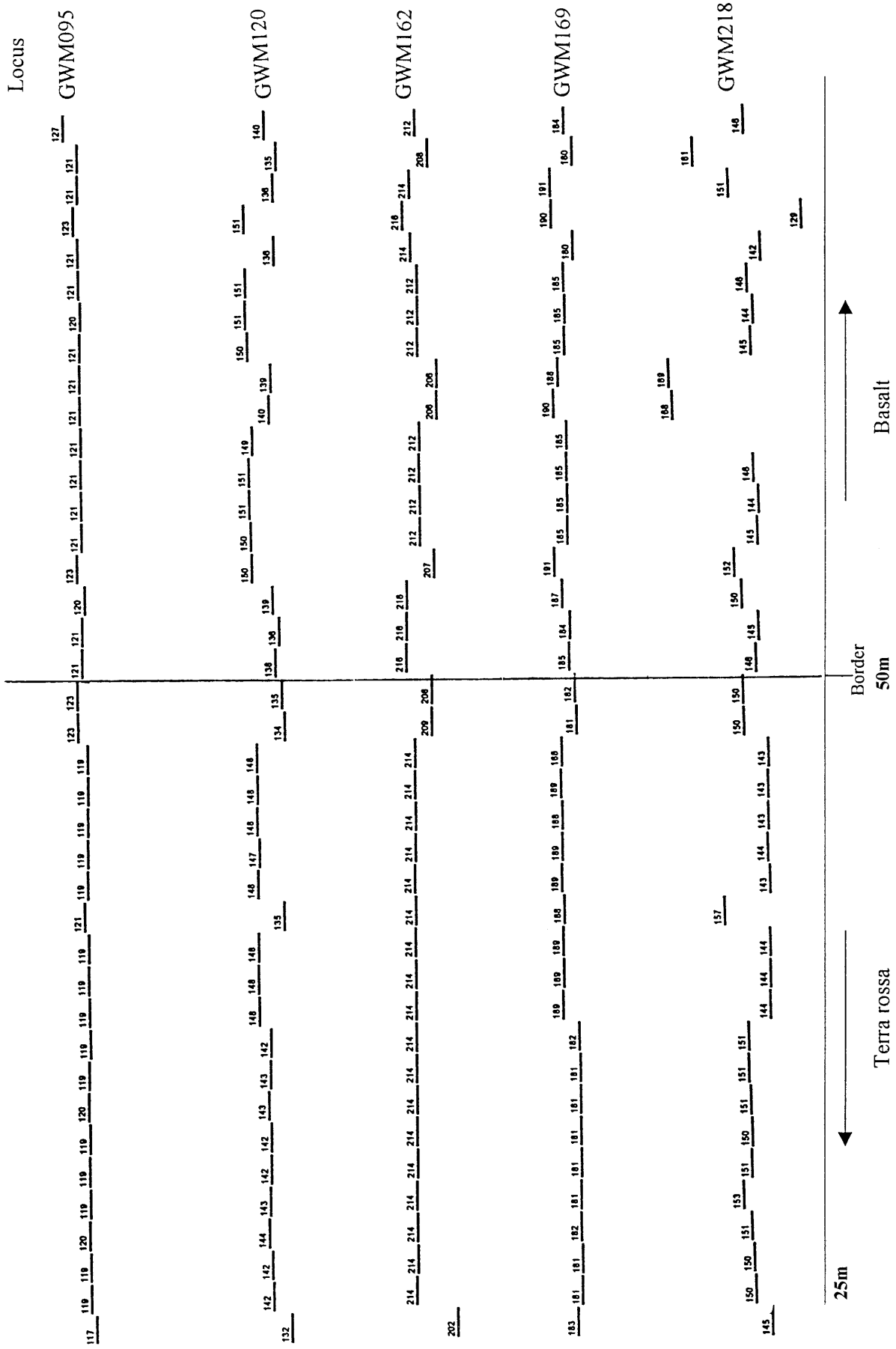

Fig. 1 Allele distribution at the GWM095, GWM120, GWM162, GWM169 and GWM218 microsatellite loci on the spatially close parts of the terra rossa and basalt soils for *T. dicoccoides*. Allele sizes are numbered by base pairs including tandem-repeated and flanking regions. The first lane on the left is for Chinese Spring

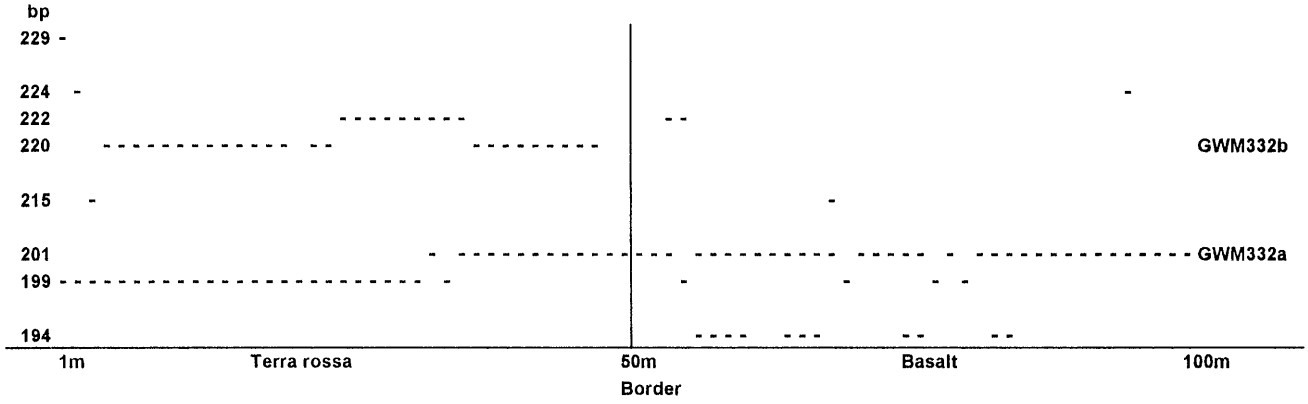

Fig. 2 Wild emmer microsatellite DNA allele distribution detected by the GWM332 primer pairs on the terra rossa and basalt soils. Allele sizes are marked by base pairs including tandem-repeated and flanking regions. The first lane on the left is for Chinese Spring

Table 3 Summary of the permutation test for the observed allele-specificity, excluding the rare alleles[a], based on 5000 random shuffling of samples from the terra rossa and basalt subpopulations at Tabigha, Israel

Subpopulation	Significance of χ^2 test for allele frequencies[b]	Number of real observed niche-specific alleles (n_o)	Cases of n_p[c] $\geq n_o$ in 5000 permutations	Significance of allele niche-specificity
Terra rossa	0.05	39	0	0.0002
	0.01	30	0	0.0002
Basalt	0.05	42	0	0.0002
	0.01	25	0	0.0002

[a] An allele was excluded if the observation of an allele was ≤ 5
[b] χ^2 test for homogeneity of allele frequencies between the two subpopulations
[c] n_p = the number of niche-specific alleles observed in the permutations

be equal between the terra rossa and basalt subpopulations. A χ^2 test was used to compare the difference in each allele distribution between the two edaphic subpopulations. After excluding rare alleles (observed in ≤ 5 individuals in the whole population), 39 and 30 terra rossa-*specific* and 42 and 25 basalt-*specific* alleles were found at $P < 0.05$ and 0.01 significance levels, respectively (Table 3). Among these alleles, 15 and 14 were terra rossa- and basalt-*unique*, respectively.

Out of 5000 random permutations of sampled individuals, no cases showed more soil-*specific* alleles than those observed in the real data of both subpopulations. The results indicated that the possibility of obtaining the observed soil-*specific* alleles by chance was < 0.0002. The significance of allele soil-*specificity* was $P < 0.0002$ in both edaphic subpopulations (Table 3).

Genetic diversity on the terra rossa and basalt soils

Although the difference in overall gene diversities (*He*; Nei 1973) was not large between the basalt (0.683) and the terra rossa (0.661) subpopulations, gene diversities were obviously higher in the basalt than in the terra rossa at 16 microsatellite loci (Table 2). In particular, the gene diversities were 0.188, 0.209, and 0.169 higher in the basalt than in the terra rossa subpopulations at loci GWM095, GWM340 and GWM577, respectively. However, an opposite pattern of *He* estimates was observed at another nine loci between the basalt and terra rossa subpopulations. At locus GWM415, the difference of *He* between the two soil types was quite large (0.241).

The genetic distance between the two subpopulations in the terra rossa and basalt soils was estimated using the allele-sharing distance index (D_{AS} = 1- proportion of sharing alleles, ranged from 0 to 1 (Chakraborty and Jin 1993) for each microsatellite locus (Table 2). The largest D_{AS} was 0.726 at GWM537, followed by another nine loci ($D_{AS} > 0.510$), then by another 12 loci, their D_{AS} ranging from 0.232 to 0.487. At the remaining five polymorphic loci, the genetic distances between the terra rossa and basalt subpopulations ranged from 0.138 to 0.187. The overall D_{AS} between the terra rossa and basalt subpopulations was also high (0.421). If the distance (200 m) between transect I and II was taken into account, the genetic distances between the terra rossa and basalt subpopulations were very similar in transect I ($D_{AS} = 0.480$) and transect II ($D_{AS} = 0.501$), which was a double check, as the genetic distance between the terra rossa and basalt subpopulations combined the two transects.

Diversity of repeat numbers between the terra rossa and basalt subpopulations

The average repeat numbers (ARNs) of microsatellite core sequences were significantly (Kruskal-Wallis test, $P < 0.05–0.00005$) larger at eight loci in the basalt subpopulation than in the terra rossa subpopulation; but the ARNs were significantly ($P < 0.05–0.00005$) smaller at another nine loci in the basalt than in the terra rossa subpopulation (Table 2). The difference of repeat numbers between the two subpopulations did not seem to relate to motif types. For example, among five loci with the motif $(CA)_n$ or $(AC)_n$, ARNs at two loci (GWM095 and GWM099) were significantly ($P < 0.00005$ and 0.05) larger in the basalt subpopulation than in the terra rossa one; at loci GWM162 and GWM251, however, the ARNs were significantly ($P < 0.00005$) smaller in the basalt than in the terra rossa subpopulation. Among nine loci with the motif $(GA)_n$, the ARNs at three loci (GWM294, GWM361, and GWM415) were significantly ($P < 0.01–0.0005$) larger in the terra rossa than in the basalt subpopulation, but, at loci GWM219 and GWM340, the ARNs were significantly ($P < 0.01–0.0005$) larger in the basalt than in the terra rossa subpopulation. Notably, most loci with compound or imperfect motifs, such as $(CT)_n(GT)_k$, $(CA)_n(TA)_k$, $(CA)_nAA(CA)_k$, etc., showed different repeat numbers in the terra rossa and basalt subpopulations (Table 2).

Variances (σ^2) in repeat number were significantly (Hartley's *F*-test, $P < 0.05–0.00005$) larger at ten loci and smaller ($P < 0.01–0.00005$) at other seven loci in the terra rossa than in the basalt subpopulations. The results suggest that distributions of repeat numbers (alleles) were not equal between the two subpopulations.

Discriminant analysis

Stepwise discriminant analysis was conducted in an attempt to classify individuals in the terra rossa and basalt subpopulations by microsatellite variation. Data from five loci (GWM099, GWM136, GWM169, GWM368 and GWM577) were excluded from the analysis in order to minimize the reduction in sample size due to missing data. At each step the variable (locus) with the largest *F* to-enter value was chosen for inclusion in the model. The stepping was terminated when no other variable had an *F* to-enter value greater than that specified under the option *F* to-enter. The classification results are presented in Table 4. At the third step, three selected loci (GWM095, GWM408, and GWM120) could correctly classify 75% of the 129 individuals into their original subpopulations. In this case, the difference among the subpopulation centroids was also significant ($F_{(3,125)} = 31.08$, $P < 0.00005$). At the fifth step, five loci (GWM095, GWM408, GWM120, GWM124, and GWM429) chosen from the remaining 21 polymorphic loci were sufficient to correctly classify most individuals into their original soil niches. The overall multivariate *F* ratio for the test of differences among the subpopulation centroids was significant ($F_{(5,123)} = 17.92$, $P < 0.00005$). According to the five loci, 83% of 129 sampled individuals were correctly classified in their original subpopulations. With the maximum selected variables (nine loci) 88% of the 129 sample individuals could be correctly classified by the stepwise procedure. The difference among the subpopulation centroids was significant ($F_{(9,119)} = 17.54$, $P < 0.00005$).

The standard method was also used to choose all variables simultaneously in the model. The entire data set (the 21 polymorphic loci) could offer a higher discriminating power: according to the 21 loci, 89% of the 129 individuals were correctly classified into their original subpopulations; the difference among the population centroids was significant ($F_{(21,107)} = 7.63$, $P < 0.00005$). These results demonstrate that genetic differentiation is substantial and significant between the two subpopulations in the terra rossa and basalt soils.

Table 4 Classification matrix[a] of discriminant analysis for 129 individuals sampled from the terra rossa (TR) and basalt (BA) subpopulations of *T. dicoccoides* at the Tabigha microsite

Selected loci	Subpopulation	Correct classification (%)	Subpopulation		*F* test for difference of centroids
			TR	BA	
3	TR	73.7	56	20	$F_{(3,125)} = 31.08$****
	BA	76.9	18	60	
	Total	75.3	74	80	
5	TR	84.0	63	12	$F_{(5,123)} = 26.17$****
	BA	82.4	13	61	
	Total	83.2	76	73	
9	TR	89.2	66	8	$F_{(9,119)} = 17.54$****
	BA	86.5	10	64	
	Total	87.8	76	72	
21	TR	93.8	60	4	$F_{(21,107)} = 7.63$****
	BA	84.5	10	55	
	Total	89.1	70	59	

****, $P < 0.00005$

[a] Rows: observed classifications; columns: predicted classification

Table 5 Niche-*specificity*[a] of linkage disequilibria at microsatellite loci of *T. dicoccoides* in the terra rossa and basalt subpopulations at the Tabigha microsite, Israel

Subpopulation	Significance level (χ^2-test)	Total LDs	Proportion of niche-*specific* LDs		No. cases $P_o \geq P_e$	Significance of niche-*specificity* or *uniqueness*[c]
			In real data (P_o)	In 5000 permutations (P_e)		
With rare alleles:						
Terra rossa	0.001	902	0.904	0.475–0.886	0	0.0002
	0.01	1612	0.849	0.500–0.798	0	0.0002
	0.05	2665	0.794	0.488–0.795	0	0.0002
	Total	5179	0.830			
Basalt	0.001	316	0.747	0.158–0.760	1	0.0002
	0.01	750	0.676	0.226–0.655	0	0.0002
	0.05	1516	0.648	0.248–0.601	0	0.0002
	Total	2582	0.660			
Without rare alleles ($n_o \leq 5$)[b]						
Terra rossa	< 0.001	881	0.904	0.519–0.912	0	0.0002
	< 0.01	1523	0.848	0.549–0.834	0	0.0002
	< 0.05	2490	0.797	0.528–0.772	0	0.0002
	Total	4894	0.832			
Basalt	< 0.001	521	0.661	0.239–0.834	28	0.005
	< 0.01	597	0.613	0.310–0.729	65	0.013
	< 0.05	1228	0.589	0.336–0.679	80	0.016
	Total	2348	0.650			

[a] Random test performed by 5000 permutation experiments
[b] Alleles observed in ≤ 5 individuals in the whole population were considered as rare alleles
[c] Significance based on 5000 permutations

Linkage disequilibrium (LD) between microsatellite loci

Non-random association of alleles between loci is supposed to be in linkage disequilibrium (LD), and this may arise from selfing, selection, etc. (Hartl and Clark 1997). The values of LDs were computed using the methods of Weir (1979). With all alleles, 5179 and 2582 LDs ($P < 0.05$–0.001) were observed between loci in the terra rossa and basalt subpopulations, respectively (Table 5). The higher the significance of LDs, the larger the proportions of niche-*specific* LDs in both subpopulations. In the terra rossa subpopulation, the proportions of niche-*specific* LDs were larger (0.794–0.904) than those (0.648–0.747) in the basalt at $P < 0.05$, 0.01, and 0.001.

Permutation test

Out of 5000 permutations of the random shuffling of sampled individuals, only one case showed a higher LD niche-*specificity* in the basalt subpopulation at the $P < 0.001$ level. No case was found with a higher LD niche-*specificity* in the basalt subpopulation at the $P < 0.05$ and 0.01 levels, and in the terra rossa at all levels (Table 5). Permutation-test results indicated that the significance of LD niche-*specificity* was very high ($P \leq 0.0002$) in both subpopulations. The proportions of niche-*specific* LDs in the terra rossa subpopulation did not change at the $P < 0.05$, 0.01, and 0.001 levels due to excluding the rare alleles observed in ≤ 5 individuals in the whole population, and the LD niche-*specificity* values were still highly significant ($P < 0.0002$) based on 5000 permutations. In the basalt subpopulation, although the proportions of niche-*specific* LDs decreased after the rare alleles were excluded, the niche-*specificity* for LDs was also significant ($P < 0.016$, 0.013, 0.005) based on 5000 random permutations at the 0.05, 0.01, and 0.001 levels, respectively (Table 5). These results suggest that the possibility of obtaining LD niche-*specificity* by stochastic effects was highly improbable.

The log-linear model test indicated that genome and chromosomes did not affect the LDs, but the soil effect was strong ($\chi^2 = 336.52$, df = 1, $P < 0.00005$) for the significance of the LDs. Sample size could not be the cause of a significant difference between the two subpopulations in the percentages of the significant LDs since the sample sizes were similar between the terra rossa (76) and basalt (79) soils.

Discussion

Microsatellite allele soil-specificity or uniqueness

At the Tabigha microsite, plant associations and plant species and diversity, i.e. the biota, differ on the contrasting terra rossa and basalt soils (Nevo et al. 1988). The Tabigha population under study involves thousands of *T. dicoccoides* plants. The population density of *T. dicoccoides* on basalt soil is higher than that on terra rossa, and morphological (spike color) polymorphism also varies between the two soil types as well as between the plant phenotypes, slender versus robust plants on the terra rossa and basalt, respectively. These differences were associated with soil type, but not with genetic drift. The previous allozyme (protein-coding region) analysis suggested that allozyme polymorphisms in wild emmer at the Tabigha microsite are partly adaptive, and that differentiation in the allozymes for both single and multilocus

structures derives primarily from edaphic stress (Nevo et al. 1988).

The present study demonstrated that allele distributions were associated with soil type at most microsatellite loci screened in the terra rossa and basalt wild emmer wheat subpopulations at Tabigha. The significant differences in average repeat numbers and the variances in repeat number also indicated that the two subpopulations were not equal. Discriminant analysis suggested that the observed microsatellite divergence between the two subpopulations was substantial, and sufficient to make the subpopulations niche-characteristic. Niche-*specific* or -*unique* alleles were observed in both subpopulations dwelling in the terra rossa and basalt soils. A random test based on 5000 permutations proved that the observed niche-*specificity* or *uniqueness* of alleles did not arise from a random drift effect, but may arise from edaphic selection by soil-*specific* environments. Microsatellite loci may be the selected target or may be located within the DNA regions subjected to natural selection.

Soil-specific linkage disequilibria

In this study, significant linkage disequilibria caused by allelic associations (Hartl and Clark 1997) in the terra rossa subpopulation were double those observed in the basalt. Most of the observed linkage disequilibria (LDs, $P < 0.05$, 0.01, 0.001) were niche-*specific* or niche-*unique* in both subpopulations, particularly in the terra rossa subpopulation. A permutation test based on 5000 random shufflings of samples suggested that the possibility of obtaining niche-*specific* LDs by random factors was very low in both subpopulations, particularly in the terra rossa where it was highly improbable. Soil-*specificity* was the main factor causing the divergence in linkage disequilibria between the terra rossa and basalt types, overriding the expected results due to the high selfing rate in wheat (99%, Golenberg 1986) that may cause linkage disequilibria by hitchhiking (Korol et al. 1994). The high LD *niche-specificity* and *–uniqueness* may be an adaptation to the microenvironment on the terra rossa and basalt soils.

Natural selection on microsatellite variation

Many of the uses of microsatellite loci as molecular markers in population biology depend on the assumption that they are neutral (e.g. Shriver et al. 1993; Valdes et al. 1993; Di Rienzo et al. 1994; Awadalla and Ritland 1997). However, some previous studies suggest that natural selection may control the level of microsatellite variation (e.g. Harding et al. 1992; Epplen et al. 1993; Stephan and Cho 1994; Garza et al. 1995; Innan et al. 1997). In a study of the selfing plant *Arabidopsis thaliana*, Innan et al. (1997) suggested that there is no reason to reject some form of balancing selection on microsatellites, in particular on microsatellites involved in an important function (e.g. gene regulation). Although a general function for microsatellite loci is not yet known, certain effects have been observed for some microsatellites that could indicate their regulatory functional role in selection at particular loci (reviewed in Kashi et al. 1997). Previous experiments showed that $(TC)_n$, $(GA)_n$, and $(TA)_n$ control the transcriptional activity of *Ultrabithorax*, *hsp26*, and *actin5C* genes in *Drosophila*. The $(TG)_n$ repeat also modulates the transcription activity of the *prolactin* gene of rat (reviewed in Kashi et al. 1997). This regulatory activity was independent of location and is consistent with the observation that microsatellites seem to be evenly spread throughout the genome and not concentrated in traditional areas of transcription regulators (Stallings et al. 1991).

Another interpretation results from the fact that *T. dicoccoides* is a highly selfing plant. Correspondingly, differential selection on fitness-related target loci will affect many other sites in the genome, either linked or unlinked to the selected loci (Korol et al. 1994). This may generate multiple niche-specific linkage disequilibria, and it is not an easy task to discriminate such a scenario from direct selection.

Studies of natural and experimental plant populations have revealed a high degree of genetic organization in predominantly selfing species, such as *Avena barbata* (Allard 1975), *T. dicoccoides* (Nevo et al. 1988; Nevo and Beiles 1989), and *Aegilops peregrina* (Nevo et al. 1994). The genetic organization is manifested in intense correlations in allelic state over loci and also in striking microgeographical heterogeneity, as described in the present study. Both aspects of this structure are facilitated by self-fertilization. The development of non-random associations of alleles within populations is facilitated because inbreeding reduces heterozygosity and thus also the randomizing effect of recombination. However, not all of the genetic variability is locked into multilocus complexes, and the recombinational potential remains substantial in predominantly selfing plants, such as *A. barbata* (Allard 1975). As a result of these factors, considerable free genetic variability remains in the populations and is available for long-term response to natural selection (Allard 1975). Similarly, in the predominantly selfing *T. dicoccoides*, there is also considerable free genetic variability in natural populations responsible for natural selection. Spatial differentiation is facilitated because self-fertilization retards gene flow from population to population. Although the low outcrossing rate in *T. dicoccoides* retards gene flow by pollen dispersal, potential seed dispersal by ants or small mammals well surpasses the pollen dispersal, since ants may move seeds of wild emmer well over 10 m (Golenberg 1986; Nevo et al. 1988). Gene flow will quickly modify patterns of geographic differentiation. This edaphic divergence of alleles and LDs is unexpected, without selection, along such short transects in dense and old population stands, particularly since some of the seeds are dispersed across transects by ants and mice (Golenberg 1986; Nevo et al. 1988). The previous study of the same materials suggest-

ed that edaphic selection operates on allozymic (protein-coding) diversity (Nevo et al. 1988). Our present results may suggest that edaphic selection influences the divergence of non-coding microsatellite DNA. Hence, natural edaphic selection may operate not only on protein-coding regions, but also upon the presumably regulatory non-coding tandem repeated regions of DNA. The observed microsatellite divergence may contribute to fitness as an evolutionary adaptive strategy. The enormous diversity of repeat numbers at selective microsatellite loci would provide a molecular source of new genetic diversity for rapid evolutionary adaptation of a population to novel ecological challenges (Kashi et al. 1997). Therefore, the present microgeographic differentiation in the two edaphic subpopulations may not be caused only by self-fertilization of wild emmer. The interaction between natural selection and the restriction of recombination associated with the mating system of predominant selfing may be the primary force responsible for the genetic organization of microsatellites in wild emmer wheat.

Acknlowledgments We thank Dr. E. Hallerman of the Virginia Polytechnic Institute and State University in the USA for his instructive comments on this paper. We also thank Dr. V. Korzun and Ms. K. Wendehake of the Institute for Plant Genetics and Crop Plant Research in Germany for their excellent technical instructions and the Graduate School of the University of Haifa for its partial financial aid. This research is supported by the Israel Discount Bank Chair of Evolutionary Biology and the Ancell-Teicher Research Foundation for Genetics and Molecular Evolution.

References

Allard RW (1975) The mating system and microevolution. Genetics 79: 115–126

Amos W, Sawcer SJ, Feaker RW, Rubinstein DC (1996) Microsatellites show mutational bias and heterozygote instability. Nature Genet 13: 390–391

Atlas of Israel (1970) Survey of Israel. Ministry of Labour, Jerusalem. Elsevier/North Holland, Amsterdam

Awadalla P, Ritland K (1997) Microsatellite variation and evolution in the *Mimulus guttatus* species complex with contrasting mating systems. Mol Biol Evol 14: 1023–1034

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368: 455–457

Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distance. In: Pena DJ, Chakraborty R, Epplen JT, Jeffrey AJ (eds) DNA fingerprinting: state of the science. Birkhäuser Verlag, Basel, pp 153–175

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

Epplen C, Melmer G, Siedlaczck I, Schwaiger FW, Mäueler W, Epplen JT (1993) On the essence of ''meaningless'' simple repetitive DNA in eukaryote genomes. In: Pena DJ, Chakraborty R, Epplen JT, Jeffrey AJ (eds) DNA fingerprinting: state of the science. Birkhäuser Verlag, Basel, pp 153–175

Estoup A, Garnery L, Solignac M, Cornuet JM (1995) Microsatellite variation in honey bee (*Apis mellifera L.)* populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140: 679–695

Fahima T, Röder MS, Grama A, Nevo E (1998) Microsatellite DNA polymorphism and divergence in *Triticum dicoccoides* accessions highly resistant to yellow rust. Theor Appl Genet 96: 187–195

Feldman M (1976) Wheats. In: Simmonds NW (ed) *Evolution of crop plants*. Longman, London, pp 120–128

Garza JC, Slatkin M, Freimer NB (1995) Microsatellite allele frequencies in human and chimpanzees, with implications for constraints on allele size. Mol Biol Evol 12: 594–603

Goldstein DB, Ruiz Linares A, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics 139: 463–471

Golenberg EM (1986) Multilocus structure in plant populations: population and genetic dynamics of *Triticum dicoccoides*. PhD thesis, State University of New York

Harding RM, Boyce AJ, Clegg JB (1992) The evolution of tandemly repetitive DNA: recombination rules. Genetics 132: 847–859

Hartl DL, Clark AG (1997) Principles of population genetics. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146: 1441–1452

Junghans H, Metzlaff M (1990) A simple and rapid method for the preparation of total DNA. *Biotechniques* 8: 176

Kashi Y, D King and M Soller (1997) Simple sequence repeats as a source of quantitative genetic variation. Trends Genet 13: 74–78

Korol AB, Preygel IA, Preygel SI (1994) Recombination variability and evolution. Chapman and Hall, London

Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. Annu Rev Ecol Syst 27: 237–277

Nauta MJ, Weissing FJ (1996) Constraints on allele size at microsatellite loci: implications for genetic differentiation. Genetics 143: 1021–1032

Nei M (1973) Analysis of gene diversity in subdivided population. Proc Natl Acad Sci USA 70: 3321–3323

Nevo E (1988) Genetic diversity in nature: patterns and theory. Evol Biol 23: 217–246

Nevo E (1998) Molecular evolution and ecological stress at global, regional and local scales: the Israeli perspective. J Exp Zool 282: 95–119

Nevo E, Beiles A (1989) Genetic diversity of wild emmer wheat in Israel and Turkey: structure, evolution, and application in breeding. Theor Appl Genet 77: 421–455

Nevo E, Beiles A, Krugman T (1988) Natural selection of allozyme polymorphisms: a microgeographical differentiation by edaphic, topographical, and temporal factors in wild emmer wheat (*Triticum dicoccoides*). Theor Appl Genet 76: 737–752

Nevo E, Beiles A, Krugman T (1994) Edaphic natural selection of allozyme polymorphisms in *Aegilops peregrina* at a Galilee microsite in Israel. Heredity 72: 109–112

Primmer CR, Ellegren H, Saino N, Moller AP (1996) Directional evolution in germline microsatellite mutations. Nature Genet 13: 391–393

Röder, MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246: 327—333

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

Saghai-Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allad W (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91: 5466–5470

Shriver MD, Jin L, Chakraborty R, Boerwinkle E (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics 134: 983–993

Slatkin M (1995) Hitchhiking and associative overdominance at a microsatellite locus. Mol Biol Evol 12: 473–480

Statsoft (1993) *Statistica* for windows, Release 4.5F. Ramon Science Center, USA

Stallings RL, Ford AF, Nelson D, Torney DC, Hildebrand CE, Morzis RK (1991) Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. Genomics 10: 807–815

Stephan W, Cho S (1994) Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. Genetics 136: 333–341

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

Terauchi R, Konuma A (1994) Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. Genome 37: 794–801

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revisited. Genetics 133: 737–749

Van Treuren R, Kuittinen H, Kärkkäinene K, Baena-Gonzalez E, Savolainen O (1997) Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*. Mol Biol Evol 14: 220–229

Weir BS (1979) Inferences about linkage disequilibrium. Biometrics 35: 235–254

Yeh FC, Yang RC, Boyle T, Ye ZH, Mao JX (1997) *POPGENE*, a user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Center, University of Alberta, Canada