Genet. Res., Camb. (1996), **67**, *pp.* 285–290 *With* 1 *text-figure*

SHORT PAPER

Microsatellite polymorphisms in a wild population of *Drosophila melanogaster*

PHILLIP R. ENGLAND*, DAVID A. BRISCOE AND RICHARD FRANKHAM
School of Biological Sciences, Macquarie University, N.S.W. 2109, Australia

(*Received 12 September 1995 and in revised form 1 December 1995*)

Summary

Highly variable DNA polymorphisms called microsatellites are rapidly becoming the marker of choice in population genetic studies. Until now, microsatellites have not been utilized for *Drosophila* studies. We have identified eight polymorphic microsatellite loci in *Drosophila melanogaster* and used them to characterize the genetic variation in a wild population from the Tyrrell's winery in Australia. Microsatellites were isolated from a partial genomic DNA library. All microsatellites consist of $(AC)_n$ repeats ranging from $n = 2$ to $n = 24$. Six loci were assigned to chromosomal location by genetic mapping, with three loci on chromosome II, one locus on chromosome III and two loci on the X chromosome. Up to four microsatellite loci were multiplexed in the same reaction. Microsatellite variation is substantially greater than allozyme variation in the Tyrrell's *Drosophila* population. 80% of the microsatellite loci examined are polymorphic, compared with 28% of allozymes. The mean number of alleles per polymorphic locus is 5·2 in microsatellites compared with 3·0 in allozymes. The average observed heterozygosity of polymorphic microsatellites is 47% compared with 26% for allozymes. Microsatellite variation in *Drosophila melanogaster* is similar to that reported for other insects. Higher variability commends microsatellites over allozymes for genetic studies in *Drosophila melanogaster*.

1. Introduction

Population genetics has traditionally used allozymes to estimate natural genetic variation within and among populations. However, recent advances in molecular biology indicate that variation at the DNA sequence level is more extensive, and therefore more informative. The most promising molecular marker is the microsatellite; an array of tandemly repeated nucleotide motifs (Litt & Lutty, 1989, Tautz, 1989; Weber & May, 1989). Relatively high conservation of flanking sequences allows the design of PCR primers that are used to amplify the region harbouring the microsatellite locus. Polymorphism for the number of repeats in the tandem array results in PCR products of varying lengths which are scored as allelic polymorphisms. Microsatellites are superior to other molecular genetic methods such as minisatellite fingerprinting and RAPDs because they are codominant, single-locus markers. Their high degree of polymorphism allows individuals to be genotyped with confidence and pedigree reconstruction and mapping using microsatellites are routine in humans and domestic animals (reviewed in Bruford & Wayne, 1993).

Surprisingly, microsatellites have not been used in studies of *Drosophila melanogaster*. Prior to this study, circumstantial evidence indicated that microsatellite polymorphisms would be found in *Drosophila melanogaster*. Highly polymorphic microsatellites have been found in other insects such as moths (Traut *et al.* 1992), bees (Estoup *et al.* 1993), mosquitos (Zheng *et al.* 1993), wasps (Hughes & Queller, 1993), ants (Hamaguchi *et al.* 1993; Gertsch *et al.* 1995) and butterflies (M. Bruford, pers. comm.). Before microsatellites were found to be polymorphic for repeat number, Tautz & Renz (1984) found that simple repeat sequences such as $(AC)_n$, $(TC)_n$, $(A)_n$ and $(CAG)_n$ are abundant in *Drosophila* on the basis of Southern hybridization to genomic DNA digests. Pardue *et al.* (1987) showed by *in situ* hybridization to chromosomes that $(AC)_n$ repeats are abundant and widely distributed in the *Drosophila* genome. A survey of the DNA sequence database GENBANK revealed the presence of di-, tri- and tetranucleotide repeat sequences in *Drosophila* (M. Montgomery, pers. comm.). Tautz (1989) demonstrated that a $(CAG)_n$

* Corresponding author: Phillip R. England at above address. Phone: (612) 850 8187; Fax: (612) 850 8245; E-mail: pengland@rna.bio.mq.edu.au

repeat in the *Notch* sequence of *Drosophila* is polymorphic between widely separated populations.

However, these studies provide no information about whether specific repeat sequences are polymorphic within *Drosophila* populations. We therefore undertook this study to identify and develop polymorphic microsatellites in *Drosophila melanogaster*, and to compare their variability with allozymes in a wild population.

2. Materials and methods

(i) *Microsatellite screen*

A partial *Drosophila melanogaster* genomic library was constructed and screened for the presence of $(AC)_n$ and $(TC)_n$ tandem repeats as outlined previously (Rassmann *et al.* 1991). Commercial DNA copolymers, poly(AC).poly(TG) and poly(TC).poly(AG) (Pharmacia) were used in equal proportions in probing of the colony blots.

(ii) *Primer design*

Positive clones were sequenced by the dideoxy chain termination method using either a Sequenase kit (Pharmacia) or a fluorescent dye-primer cycle sequencing kit (Applied Biosystems). Sequenced DNA was run on a polyacrylamide gel and the sequence read either manually from an autoradiograph, or automatically using an automated sequencer (Applied Biosystems). The program Primer, version 0·5 (Lincoln *et al.* 1991) was used in most cases to design PCR primers that flanked microsatellites identified by the screen. Where alternative primer pairs were suggested by the program a choice was made on the basis of both maximizing GC content of the primers and covering a range of PCR product sizes that would facilitate multiplexing and co-electrophoresis of several loci. Primers were synthesized on a Beckman Oligo 1000 DNA synthesizer.

(iii) *PCR*

PCR reactions were conducted in polycarbonate microtitre trays (Bresatec) on an MJR Thermal Cycler (MJR Research). All loci amplified with the program: 5 min at 94 °C then 30 cycles of 30 s at 94 °C, 30 s at 60 °C and 1 min at 72 °C followed by 5 min at 72 °C. 1 unit of *Taq* polymerase (Promega) was used in each 20 μl reaction with the buffer supplied by the manufacturer. Reactions were 2·5 mM for Mg^{2+}, 200 μM for unlabelled dTTP, dGTP and dCTP and 20 μM for unlabelled dATP. 1 μCi [α^{32}P] dATP (3000 mCi/mmol, Bresatec), 10 pmol of each primer and approximately 20 ng of template DNA were used per reaction. Conditions were not altered for multiplex reactions. PCR products were visualized by autoradiography after electrophoresis on 6% denaturing polyacrylamide gels. Large PCR fragments (greater than 250 bp) were electrophoresed far enough to ensure detection of polymorphism. Assignment of allele size was achieved by running the cloned allele as a standard size marker on each gel.

In order to reduce the number of reactions (and consequent costs) required to genotype individuals, all eight primer pairs were systematically trialed in all two-way, three-way and four-way combinations to identify loci that could be multiplexed in the same reaction. Products from different reactions were also combined and co-electrophoresed in the same lane where the size range of alleles at different loci did not overlap.

DNA for the library construction and PCR reactions was prepared by a high salt precipitation method (Miller *et al.* 1988) performed on whole fly homogenate. The homogenate was prepared by grinding a single fly in 100 μl buffer consisting of 10 mM Tris, 100 mM-NaCl, 10 mM EDTA and 2 μg proteinase K (Boehringer Mannheim) and incubating it overnight at 37 °C with agitation. Purified DNA was resuspended in 50 μl of H_2O and 1 μl of this was used in each PCR reaction.

(iv) *Mapping*

In order to assign the polymorphic loci to chromosomes, several highly inbred *Drosophila* lines homozygous at microsatellite loci were genotyped to identify lines which carried a different allele (+) from the marked inversion stock BLT: *In(1) sc^{S1L} sc^{8R}+S, sc^{S1} sc^{8} w^{a}*; *In(2LR) SM1, al^{2} Cy cn^{2} sp^{2}/In(2LR) bw^{V1}, ds^{33k} dp^{ov} bw^{V1}*; *In(3LR) Ubx^{130}, Ubx^{130} es/In(3LR) C, Sb; spa^{pol}* (Lindsley & Zimm, 1992). Virgin females from each inbred line were crossed to single BLT males. Heterozygous *Cy/+*, *Sb/+* male progeny from this cross were backcrossed to virgin females from the parental inbred line. Male and female parents in this backcross were genotyped along with one male and one female offspring from each of the four segregating classes; *Cy/+, Sb/+*; *Cy/+, +/+*; *Sb/+, +/+*; *+/+, +/+*.

(v) *The Tyrrell's population*

Wild flies used in this study were collected and genotyped for allozymes in 1992 from the Tyrrell's Winery near Sydney (R. K. Nurthern, pers. comm.). Offspring of the wild caught flies were stored at −80 °C until being retrieved for microsatellite analysis.

(vi) *Analysis*

Genetic data were analysed using the Biosys-1 program, release 1·7 (Swofford & Selander, 1981).

3. Results

Of 2500 clones screened, 25 showed homology with the copolymer probes. Sequencing revealed 17 clones with $(AC)_n$ repeats, 12 of which had flanking sequences

https://doi.org/10.1017/S0016672300033760 Published online by Cambridge University Press

Table 1. *Cloned microsatellite loci showing chromosome position, allele size and type, PCR primers used and extent of polymorphism in the Tyrrell's* Drosophila melanogaster *population*

Locus	Repeat type (cloned allele)	Primer sequences (5′-3′)	PCR product size (bp)	Chromosome location	Number of alleles
DmAC1	$(AC)_{12}$	tacacctgcgtgcccaac ggggctaatggagcctagtg	139	II	5
DmAC2	$(AC)_{12}$	cccattgattttgttgtgtaag gcacatgcatgtgggtgt	104	X	3
DmAC3	$(AC)_{9}$	ctccacaatccaccctcg ccatctaccacacaaccgc	196	III	3
DmAC4	$(AC)_{9+10}$	agaattcgaagtgcaagacca tgtgggcataaccaacacat	170	X	6
DmAC6	$(AC)_{13}$	ctgttcttctgccgttgtca gagctcggtacccctactcc	315		Monomorphic
DmAC7	$(AC)_{10}$	gtgattctcaacaaagcgca gcagatgtttcacccctttgt	219		5
DmAC8	$(AC)_{9}$	cccgcatcaccacaacac caagtgtcacaatgtcgatgg	95	II	6
DmAC9	$(AC)_{13}$	ctatacgcgcatcgacgtc tggctacactatccagcgc	167	II	10
DmAC10	$(AC)_{13}$	tgactacggtgccatttcaa gagcccaacgcaattaatgt	273		Monomorphic
DmAC12	$(AC)_{11}$	gtgacttccggcaatgtttt cggtacccctgctattacca	111		5

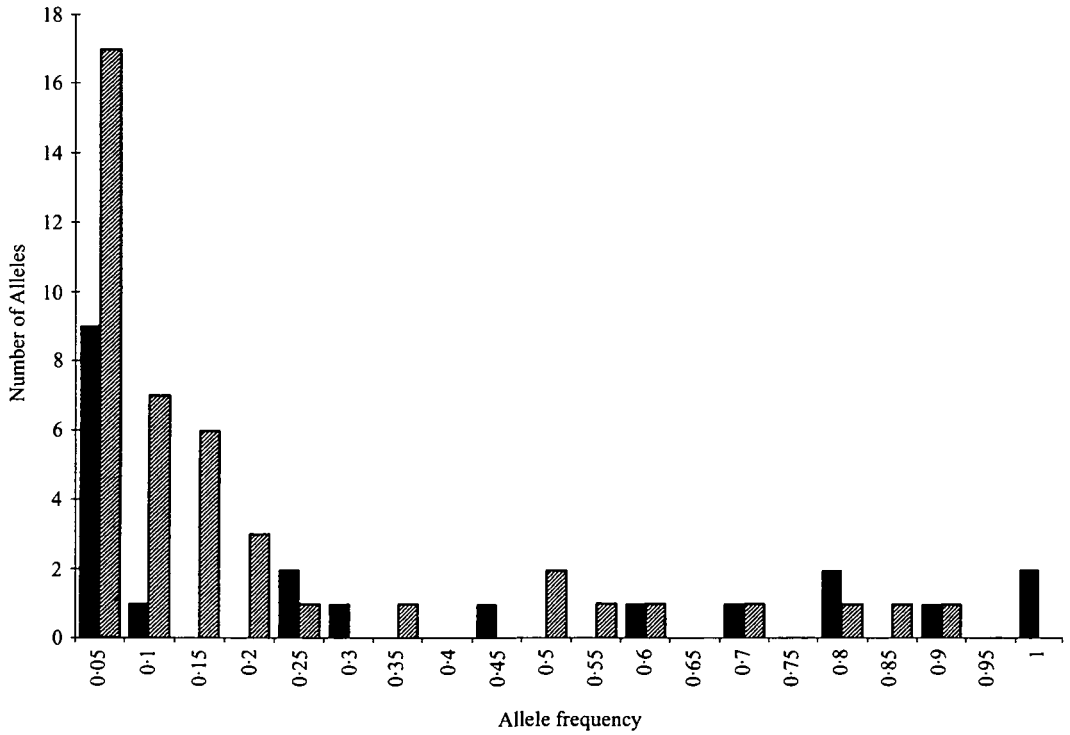

Fig. 1. Distribution of microsatellite ▨ and allozyme ■ allele frequencies in the Tyrrell's *Drosophila melanogaster* population.

suitable for designing primers. No $(TC)_n$ repeats were detected. PCR products of the expected size were amplifiable with 10 primer pairs and eight of these displayed polymorphism in the population examined, with a total of 43 alleles. Table 1 details the microsatellite loci amplified, the primers used, the number of alleles found in the Tyrrell's population and chromosome assignments of loci where they were determined. DmAC7 and DmAC12 could not be mapped because available inbred lines shared identical alleles at these loci with the BLT stock.

DmAC1, DmAC3, DmAC4 and DmAC7 were successfully multiplexed (Fig. 2) in a PCR reaction with an annealing temperature of 60 °C (see Materials and methods). DmAC2 and DmAC9 were also multiplexed using these conditions. DmAC8 and

https://doi.org/10.1017/S0016672300033760 Published online by Cambridge University Press

Table 2. *Microsatellite allele frequencies and heterozygosities in the Tyrrell's* Drosophila melanogaster *population*

Allele (repeat number	Locus							
	DmAC1	DmAC2	DmAC3	DmAC4	DmAC7	DmAC8	DmAC9	DmAC12
2					0·015			
3								
4							0·007	
5							0·116	
6					0·220*	0·024	0·493*	
7							0·007	0·007
8			0·068			0·016		
9	0·088		0·803*		0·061	0·675*	0·014	0·014
10	0·022		0·129		0·576*	0·175*	0·087*	
11	0·081	0·028			0·129	0·079	0·101	0·861*
12	0·478*	0·897*				0·032	0·043	0·007
13	0·331	0·075					0·123	
14								
15								0·111
16							0·007	
17				0·010				
18								
19				0·168				
20				0·050				
21				0·050				
22								
23				0·524*				
24				0·198				
Sample size	68	71	66	65	66	63	69	72
Observed heterozygosity	0·588	0·139	0·318	0·657	0·621	0·492	0·710	0·250
Expected H–W heterozygosity	0·647	0·176	0·334	0·666	0·600	0·506	0·714	0·246

* Denotes allele in BLT stock.

Table 3. *Microsatellite and allozyme diversity of the Tyrrell's* Drosophila melanogaster *population, either for polymorphic loci alone or with monomorphic loci included. Standard errors are in parentheses*

Marker	Mean number of alleles	Percentage of polymorphic loci	Observed Heterozygosity	Expected H–W Heterozygosity
Microsatellites (polymorphic)	5·2 (0·9)	—	0·472 (0·075)	0·490 (0·074)
Allozymes* (polymorphic)	3·0 (0·4)	—	0·258 (0·069)	0·272 (0·070)
Microsatellites (all)	4·4 (0·9)	80	0·378 (0·086)	0·392 (0·087)
Allozymes* (all)	0·84 (0·3)	28	0·072 (0·030)	0·072 (0·030)

* R. K. Nurthern, pers. comm.

DmAC12 were amplified individually with the 60 °C annealing temperature reaction.

Table 2 details the genetic diversity of the loci polymorphic in the Tyrrell's population. Polymorphic microsatellite loci have approximately twice the heterozygosity and mean number of alleles than allozymes (Table 3). Microsatellite allele distribution is unimodal except at DmAC4 and DmAC9 where distributions are bimodal (Table 2). Many allozyme and microsatellite alleles are rare (5% or less), but more microsatellite alleles occur at intermediate frequency (between 5% and 50%) (Fig. 1).

4. Discussion

We have cloned and characterized eight polymorphic microsatellites in *Drosophila melanogaster*. By comparing microsatellite and allozyme variability in the

https://doi.org/10.1017/S0016672300033760 Published online by Cambridge University Press

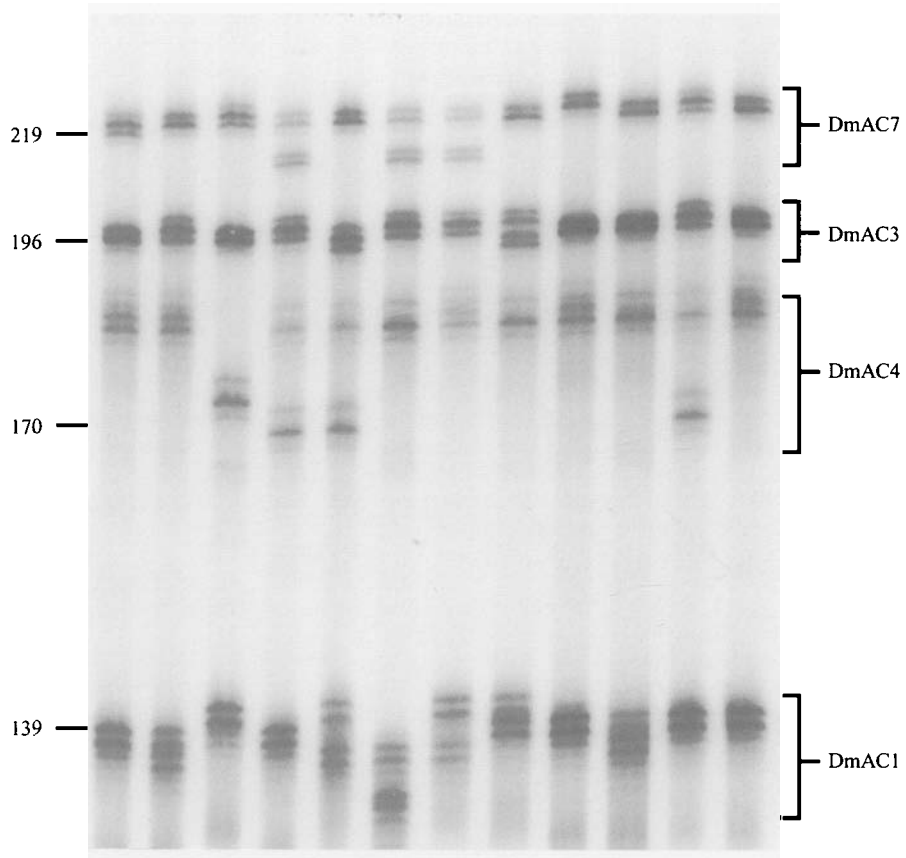

Fig. 2. Representative autoradiograph of 4 multiplexed microsatellite loci (DmAC1, DmAC3, DmAC4 and DmAC7) in 12 females from the Tyrrell's *Drosophila melanogaster* population. Numbers refer to the approximate position, in base pairs, of the cloned allele at each of the four loci.

same population we have shown that microsatellites lend greater resolution to measurements of genetic variation than do allozymes. This work represents, to our knowledge, the first published report of polymorphic microsatellites in a *Drosophila* population.

The identification of numerous microsatellite loci composed of $(AC)_n$ repeats confirms circumstantial evidence of their existence gathered by Tautz & Renz (1984), and Pardue *et al.* (1987). Pardue *et al.* (1987) detected regions of homology with $(AC)_n$ probes on chromosomes II, III and X. Consistent with these findings, we have mapped microsatellite loci to each chromosome except IV and Y.

The microsatellites we have cloned are short, with modal repeat number ranging from 8 to 13. Cloned $(AC)_n$ microsatellites in other insects also tend to be short. Choudhary *et al.* (1993) cloned six $(AC)_n$ microsatellites with repeat number ranging from 10 to 16 in the social wasp *Parachartergus colobopterus*. Estoup *et al.* (1993) found a greater abundance of short $(AC)_n$ microsatellites (less than 10 repeats) in honey bees *Apis melifera* compared to pigs. Lanzaro *et al.* (1995) found only one out of 10 $(AC)_n$ microsatellites in the mosquito *Anopheles gambiae*, with a repeat number greater than 13.

The distributions of microsatellite alleles in *Drosophila melanogaster* are unusually steep, with one or two alleles being much more common than all others (Table 2). Microsatellite alleles are usually more normally distributed (Weber, 1991). The bimodality of allele distributions obtained at two loci in our study is commonly observed in other species of invertebrates, vertebrates and plants.

For the genetic diversity measures: proportion of loci polymorphic, number of alleles per locus and observed heterozygosity, microsatellites in the Tyrrell's *Drosophila* population scored significantly higher than allozymes. Their greater resolving power should lead to the widespread adoption of microsatellite genotyping in *Drosophila* studies.

We thank Rod K. Nurthern for conducting the allozyme survey; Margaret Montgomery for searching GENBANK and for technical assistance with fly maintenance; Lynn Woodworth for fly maintenance; the Macquarie University Centre for Analytical Biotechnology for primer synthesis and Brian Neii at Westmead Hospital DNA Sequence and Synthesis Facility for automated sequencing. Andrea Taylor kindly provided competent cells and expert technical advice with the library screen. We are grateful to Emma Pharo, Andrea Taylor and Paul Sunnucks for comments on the manuscript. Our work was supported by the Australian Research Council and Macquarie University Research Grants. Publication 193 of the Research Unit for Biodiversity and Bioresources.

Note added in proof

Subsequent to the acceptance of this manuscript Goldstein and Clark published a complementary study: Goldstein,

D. B. & Clark, A. G. (1995). Microsatellite variation in North American populations of *Drosophila melanogaster*. *Nucleic Acids Research* **23**, 3882–3886.

References

Bruford, M. W. & Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Current Opinion in Genetics and Development* **3**, 939–943.

Choudhary, M., Strassmann, J. E., Solis, C. R. & Queller, D. C. (1993). Microsatellite variation in a social insect. *Biochemical Genetics* **31**, 87–96.

Estoup, A., Solignac, M., Harry, M. & Cornuet, J.-M. (1993). Characterisation of (GT)n and (CT)n microsatellites in two insect species: *Apis melifera* and *Bombus terrestris*. *Nucleic Acids Research* **21**, 1427–1431.

Gertsch, P., Pamilo, P. & Varvio, S.-L. (1995). Microsatellites reveal high genetic diversity within colonies of *Camponotus* ants. *Molecular Ecology* **4**, 257–260.

Hamaguchi, K., Ito, Y. & Takenaka, O. (1993). Dinucleotide repeat polymorphisms in a polygynous ant, *Leptothorax spinosior*, and their use for measurement of relatedness. *Naturwissenschaften* **80**, 179–181.

Hughes, C. R. & Queller, D. C. (1993). Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology* **2**, 131–137.

Lanzaro, G. C., Zheng, L., Toure, Y. T., Traore, S. F., Kafatos, F. C. & Vernick, K. D. (1995). Microsatellite DNA and isozyme variability in a West African population of *Anopheles gambiae*. *Insect Molecular Biology* **4**, 105–112.

Lincoln, S. E., Daly, M. J. & Lander, E. S. (1991). *Primer: a Computer Program for Automatically Selecting PCR Primers, Version 0.5*. Massachusetts, USA: MIT Centre for Genome Research and Whitehead Institute for Biomedical Research.

Lindsley, D. L. & Zimm, G. G. (1992). The genome of *Drosophila melanogaster*. San Diego, CA: Academic Press.

Litt, M. & Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**, 397–401.

Miller, S. A., Dykes, D. D. & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215.

Pardue, M. L., Lowenhaupt, K., Rich, A. & Nordheim, A. (1987). $(dC\text{-}dA)_n.(dG\text{-}dT)_n$ sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function. *EMBO Journal* **6**, 1781–1789.

Rassmann, K., Schlötterer, C. & Tautz, D. (1991). Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis* **12**, 113–118.

Swofford, D. L. & Selander, R. B. (1981). BIOSYS-1: A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity* **72**, 281–283.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463–6471.

Tautz, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* **10**, 4127–4138.

Traut, W., Epplen, J. T., Weichenhan, D. & Rohwedel, J. (1992). Inheritance and mutation of hypervariable $(GATA)_n$ microsatellite loci in a moth, *Ephestia kuehniella*. *Genome* **35**, 659–666.

Weber, J. L. (1991). Human DNA polymorphisms based on length variations in simple-sequence tandem repeats. In *Genome Analysis: A Practical Approach, Volume 1* (ed. K. E. Davies & S. M. Tilghman), pp. 160–181. New York: Cold Spring Harbor Laboratory Press.

Weber, J. L. & May, P. E. (1989). Abundant class of human polymorphism which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388–396.

Zheng, L., Collins, F. H., Kumar, V. & Kafatos, F. C. (1993). A detailed genetic map for the X chromosome of the malaria vector, *Anopheles gambiae*. *Science* **261**, 605–608.