

# Cross-species transferability of microsatellite markers from six aphid (Hemiptera: Aphididae) species and their use for evaluating biotypic diversity in two cereal aphids

Y. Weng, P. Azhaguvel, G. J. Michels Jr and J. C. Rudd

Texas A&M University Agricultural Research and Extension Center, Amarillo, Texas, USA

## Abstract

The abundance and distribution of microsatellites, or simple sequence repeats (SSRs) were explored in the expressed sequence tag (EST) and genomic sequences of the pea aphid, *Acyrtosiphon pisum* (Harris), and the green peach aphid, *Myzus persicae* (Sulzer). A total of 108 newly developed, together with 40 published, SSR markers were investigated for their cross-species transferability among six aphid species. Genetic diversity among six greenbug, *Schizaphis graminum* (Rondani) and two Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) biotypes was further examined with 67 transferable SSRs. It was found that the pea aphid genome is abundant in SSRs with a unique frequency and distribution of SSR motifs. Cross-species transferability of EST-derived SSRs is dependent on phylogenetic closeness between SSR donor and target species, but is higher than that of genomic SSRs. Neighbor-joining analysis of SSR data revealed host-adapted genetic divergence as well as regional differentiation of greenbug biotypes. The two Russian wheat aphid biotypes are genetically as diverse as the greenbug ones although it was introduced into the USA only 20 years ago. This is the first report of large-scale development of SSR markers in aphids, which are expected to have wide applications in aphid genetic, ecological and evolutionary studies.

**Keywords:** Simple sequence repeats, biotypic diversity, aphids, *Schizaphis graminum*, *Diuraphis noxia*, *Acyrtosiphon pisum*.

Received 26 June 2007; accepted after revision 9 July 2007; first published online 22 August 2007. Correspondence: Yiqun Weng, Texas A&M University Agricultural Research and Extension Center, 6500 Amarillo Blvd. W., Amarillo, Texas 79106, USA. Tel.: +1 806 677 5600; fax: +1 806 677 5644; e-mail: yweng@ag.tamu.edu

## Introduction

Simple sequence repeats (SSRs), or microsatellites, are short elements consisting of tandem repeat units of one to six base pairs in length which are widely distributed throughout all eukaryotic genomes (Tautz, 1989). Compared with other DNA-based molecular markers, SSRs have a number of advantages, such as their abundance and dispersion throughout the entire genome, high information content, co-dominant inheritance, reproducibility and genomic specificity, and adaptation to automation (Morgante & Plivieri, 1993; Powell *et al.*, 1996). SSR markers have been used extensively for genome mapping, DNA fingerprinting and a wide range of genetic diversity, population and evolutionary studies in many eukaryotic organisms.

Two types of SSR marker can be recognized based on their origin. One is derived from genomic sequences (genomic SSRs), and the other from expressed sequence tags (ESTs) (EST-SSRs). In general, genomic SSRs are more polymorphic than EST-SSRs. However, except in species with whole or partial genome sequences available, development of genomic SSRs is time-consuming and costly. In contrast, with increasingly available ESTs, EST-SSRs can be developed at trivial cost from the EST databases. As EST-SSRs are developed from transcribed regions that are better conserved among genomes; they show a higher degree of transferability to closely related species (e.g. Sourdille *et al.*, 2001; Yu *et al.*, 2004; Zhang *et al.*, 2005) than do genomic SSRs. Therefore, genomic SSRs are more often used for genetic diversity studies within a single species, whereas EST-SSRs can be used in phylogenetically related species where SSR markers are relatively scarce or hard to develop because of lack of sequence information.

In recent years, SSR markers have also been used in a number of studies with aphids, such as work on population genetics (e.g. Sunnucks *et al.*, 1996; Simon *et al.*, 1999; Massonnet *et al.*, 2002; Llewellyn *et al.*, 2003); genetic diversity (e.g. Figueroa *et al.*, 1999; Fuller *et al.* 1999; Wilson *et al.*, 2002; Simon *et al.*, 2003), sexual reproduction (e.g. Sunnucks *et al.*, 1997; Delmotte *et al.*, 2002; Papura *et al.*, 2003) and evolution (e.g. Wilson *et al.*, 1999). However, all of these studies were based on only a few publicly available

genomic SSRs isolated from a limited number of aphid species (Sunnucks *et al.*, 1996; Simon *et al.*, 1999; Vanlerberghe-Masutti *et al.*, 1999; Sloane *et al.*, 2001; Caillaud *et al.*, 2004; summarized in Wilson *et al.*, 2004). To our knowledge, no publicly accessible EST-SSRs have been reported in any aphid species to date. The lack of highly polymorphic, easy-to-use molecular markers such as SSRs may be one important reason why there have been so few genetic mapping studies in aphids (e.g. Hawthorne & Via, 2001; Braendle *et al.*, 2005).

The greenbug, *Schizaphis graminum* (Rondani), and Russian wheat aphid (RWA, *Diuraphis noxia* (Kurdjumov)) are the two most important aphid pests of wheat (*Triticum aestivum* L.) in the Southern Great Plains of the USA. They are especially notorious as a result of periodic appearance of new prevailing biotypes that overcome host resistance. Biotypic variation is common in aphids, but the definition of biotype varies in different insects (Diehl & Bush, 1984). Greenbug and RWA biotypes were defined by their abilities to damage different plant genotypes (Shufran *et al.*, 1991; Haley *et al.*, 2004). So far, nearly 30 greenbug (Porter *et al.* 1997; Burd & Porter 2006) and five RWA biotypes (Haley *et al.*, 2004; Burd *et al.* 2006) have been identified. Independently of, and well before, the advent of modern agriculture, it was believed that the greenbug species complex was composed of host-adapted races that diverged on grass species (Porter *et al.*, 1997; Shufran *et al.*, 2000). Biotypes are comprised of genetically diverse individuals that share similar virulence genes (Anstead *et al.*, 2002). Previous investigations have suggested that virulence of the greenbug is controlled by at least two genes plus modifiers (Puterka & Peters, 1989, 1995). However, no further study was made on the genetic basis of virulence in the greenbug against host resistance in wheat or sorghum. One method of studying the genetics of virulence is genetic mapping of the virulence locus/loci with SSR markers. Unfortunately, no such markers exist for either the greenbug or the RWA. However, increasingly available genomic resources for other aphids provide opportunities to explore the use of cross-species transferable SSR markers in the two species. For example, there are many EST sequences of the pea aphid, *Acyrtosiphon pisum* (Harris) and the

green peach aphid, *Myzus persicae* (Sulzer) in GenBank (<http://www.ncbi.nih.gov>). Whole genome sequencing of the pea aphid is underway (<http://www.hgsc.bcm.tmc.edu/projects/aphid/>). Wilson *et al.* (2004) demonstrated that there is high transferability of genomic SSRs among different aphid species. A large scale analysis among different insect taxa indicated high conservation of SSR flanking sequences (Megléczy *et al.*, 2007). These genomic resources give us new opportunities for developing molecular markers in target species. The objectives of this study were to: (1) explore pea aphid and green peach aphid ESTs and genomic sequences to characterize the abundance and distribution of SSRs in the two genomes; (2) evaluate the transferability of identified SSRs among different aphid species; and (3) use transferable SSR markers to investigate genetic diversities among the greenbug and RWA biotypes.

## Results and discussion

### *Abundance and distribution of SSRs in EST and genomic sequences of the pea aphid*

We analyzed the distribution of perfect SSRs spanning 10 bp or more in the pea aphid EST and genomic sequences. For practical purposes, we only counted SSRs with motifs containing di-, tri-, tetra- and pentanucleotides. Thus, for a 10 bp SSR, one occurrence may comprise a repeat of five dinucleotides, four trinucleotides, three tetranucleotides or two penta-nucleotides. The SSR motifs include both strands of the DNA sequence. AAT, for example, also includes TTA and the reverse complements TAA and ATT. The total lengths of each repeat type were used to estimate abundance of SSRs in the genome.

The pea aphid has a genome of approximately 525 Mb (Sabater-Muñoz *et al.*, 2006). We screened 36.7 Mb of the expressed portion (ESTs) and 19.54 Mb of its genomic sequences for perfect SSRs. We also examined 11.62 Mb of EST sequences of the green peach aphid. The abundance and distribution of the different types of SSR repeats are shown in Table 1. It seems that both the genomic and EST sequences of the pea aphid genome were abundant in SSRs. For example, 12 275 SSRs were detected among 67 309 ESTs, which accounted for 0.53% of the EST

SSR source*	Sequence examined	SSR length (%)	SSR density (# per kb)	# SSRs	Distribution† (%)			
					Di-	Tri-	Tetra-	Penta-
PA-EST	36.71 Mb	0.53	0.33	12 275	17.5	54.5	10.1	17.8
PA-Genomic	19.54 Mb	0.86	0.53	10 425	26.5	39.6	11.9	22.0
GPA-EST	11.62 Mb	0.28	0.19	2 178	9.8	66.2	10.9	13.1

**Table 1.** Abundance of simple sequence repeats (SSRs) in the pea aphid and green peach aphid genomes

\*PA, pea aphid; GPA, green peach aphid.

†Di-, dinucleotide repeats; Tri-, trinucleotide repeats; Tetra-, tetranucleotide repeats; Penta-, pentanucleotide repeats.

sequences examined, and there was, on average, one SSR per 3.0 kb (0.33 SSR per kb EST). The green peach aphid ESTs contained fewer SSRs than the pea aphid ESTs (0.28 vs. 0.53%). SSRs were more abundant in the genomic sequences than in the coding regions of the pea aphid. SSRs were present in 0.86% of the genome with an average density of one SSR per 1.89 kb genomic DNA. This distribution of SSRs in ESTs and genomic sequences was consistent with earlier observations in both plants and animals (e.g., Metzgar *et al.*, 2000; Tóth *et al.*, 2000; Morgante *et al.*, 2002; Li *et al.*, 2004).

Mono- and hexanucleotide SSR repeats were not counted in this study. Therefore, it is difficult to compare the SSR abundance of the pea aphid with other insect species. However, if only the di-, tri-, tetra- and pentanucleotide SSR motifs are considered, it seems that the abundance of SSRs is taxon- and species-dependent. Megléc *et al.* (2007) analysed abundance and motif distribution of SSRs in eight dipteran (six *Drosophila* spp., *Aedes aegypti* and *Anopheles gambiae*), one lepidopteran (*Bombyx mori*) and one hymenopteran (*Apis mellifera*) species for which whole genome sequences are available. The abundance of SSRs with the four repeat types in these species varied from 0.31% in *Ae. aegypti* to 2.1% in *Ap. mellifera*. The 0.86% genomic SSR abundance in pea aphids found in the present study is similar to several *Drosophila* spp. analysed by Megléc *et al.* (2007). Accurate SSR statistics in the pea aphid should be available once its whole genome sequence is completed.

Analysis of the distribution of the four repeat types revealed that trinucleotide repeats predominate the coding regions (EST sequences) in both the pea and the green peach aphid genomes, accounting for 54.5 and 66.2%, respectively of the total SSRs identified (Table 1). In pea aphid genomic sequences, 39.6% of SSRs were triplets, and 26.5% were dinucleotides. This is in sharp contrast with results from Megléc *et al.* (2007) in which trinucleotide repeats were the third most abundant repeat type after mono- and dinucleotide repeats in all 10 insect species examined. The predominance of trinucleotide repeats in coding regions occurs in many other organisms (Metzgar *et al.*, 2000; Subramanian *et al.*, 2003; La Rota *et al.*, 2005; Peng & Lapitan, 2005), and is probably the consequence of negative selection against frameshift mutations in the coding regions (Metzgar *et al.*, 2000).

The base composition of the SSR motifs of each repeat type is biased strongly toward As and Ts in both EST and genomic SSRs of the pea aphid (Table 2), which is strikingly different from that in other insect species investigated. For example, in all dinucleotide EST-SSRs, 74.4% contain AT/TA, but only 1.77% have GC/CG. In genomic SSRs, 81.1% of trinucleotides have either AAT/TAA/TTA/ATT (42.1%) or ATA/TAT (39.0%). In the di-, tri-, tetra- and pentanucleotide motifs, the most abundant repeat types are AT, AAT, AAAT

**Table 2.** Base composition of different motifs in pea aphid genomic and expressed sequence tag simple sequence repeats (EST-SSRs)

SSR motif base composition	EST-SSR (%)				Genomic SSR (%)			
	Di-	Tri-	Tetra-	Penta-	Di-	Tri-	Tetra-	Penta-
A and T only	74.4	64.3	65.4	58.2	72.7	81.1	68.6	76.5
G and C only	1.77	5.24	0.72	0.00	1.16	3.88	0.41	0.00
Mixed*	23.8	30.5	33.9	41.8	26.3	15.1	30.1	33.5

\*Base combinations other than A and T only, or G and C only (e.g. AG, TAG, ATGC etc.).

and AAAAT, respectively (data not shown). A similar feature of base composition in SSR motifs was also found in the human genome (Subramanian *et al.*, 2003), but contrasts with that in some cereal genomes such as rice (*Oryza sativa* L.) where (CCG)<sub>n</sub> is the most abundant among all trinucleotide repeats (Zhao & Kochert, 1993; Grover *et al.*, 2007). However, Rajendrakumar *et al.* (2007) reported a high percentage of AT/TA dinucleotide repeats in rice. Why the SSR motif base composition in the pea aphid genome skews so heavily toward A and T is not known.

To summarize, database mining of the pea aphid and green peach aphid EST and genomic DNA sequences suggests that both genomes are rich in SSRs, yet the abundance and distribution of repeat motifs vary greatly as compared with other insect species.

#### *Cross-species transferability of EST- and genomic SSRs among six aphid species*

Of the 24 878 SSRs identified from the pea aphid and green peach aphid EST or genomic DNA sequences (Table 1), 48 genomic SSRs and 43 EST-SSRs of the pea aphid, as well as 13 EST-SSRs of the green peach aphid were selected. Four genomic SSRs of the greenbug were also developed through database mining (data not shown). Thus, together with 40 published genomic SSRs, a total of 148 SSRs was used in the cross-species transferability study, which were from seven aphid species: the pea aphid, the green peach aphid, the greenbug, English grain aphid, *Sitobion avenae* (F.), bird cherry-oat aphid, *Rhopalosiphum padi* (L.), *Acyrtosiphon loti* (Tobald) and *Sitobion miscanthi* (Takahashi) (see Supplementary Material Table S1 for details of all 148 markers). The six target aphid species used as DNA templates in the PCR amplification were the pea aphid, green peach aphid, bird cherry-oat aphid, RWA, cotton aphid, *Aphis gossypii* (Glover) and the greenbug. The cross-species amplification results were summarized in Table 3 and more detailed information of PCR amplification in individual aphid species was provided in Supplementary Material Table S2.

Of the 148 SSRs, 12 failed to amplify from genomic DNA templates of all six aphid species. Therefore, only data for the remaining 136 SSRs are shown in Table 3. Not all of the

SSR source	SSR #	PA	GPA	RWA1	BCOA	CA	GBE	Ave. %†
<b>EST-SSRs</b>								
<i>Acyrtosiphon pisum</i>	41	36	29	29	22	18	21	
%‡		88.1	70.7	70.7	53.7	43.9	51.2	58.5
<i>Myzus persicae</i>	12	9	8	9	9	7	10	
%		75.0	66.7	75.0	75.0	58.3	83.3	71.7
<b>Genomic SSRs</b>								
<i>A. pisum</i>	41	37	9	11	15	11	8	
%		90.2	21.9	26.8	36.6	26.8	19.5	26.3
<i>Acyrtosiphon</i> §	7	6	3	3	5	4	5	
%		85.7	42.9	42.9	71.4	57.1	71.4	57.1
<i>M. persicae</i>	8	5	6	3	2	4	4	
%		62.5	75.0	37.5	25.0	50.0	50.0	45.0
<i>Rhopalosiphum padi</i>	8	5	4	6	5	2	5	
%		62.5	50.0	75.0	62.5	25.0	62.5	56.3
<i>Schizaphis graminum</i>	4	2	2	2	3	0	4	
%		50.0	50.0	50.0	75.0	0.0	100.0	45.0
<i>Sitobion</i> §	15	9	8	8	10	5	7	
%		60.0	53.3	53.3	66.7	33.3	46.7	52.2

\*PA, pea aphid; GPA, green peach aphid; RWA1, Russian wheat aphid (biotype RWA1); BCOA, bird cherry-oat aphid; CA, cotton aphid; GBE, greenbug (biotype E).

†Average percentages of marker transferability across different species. Amplification from the same species as the SSR donor was excluded in calculation.

‡Percentage of successful amplification of SSR markers used.

§The genomic SSRs were originally developed from two different species in the genus. See Supplementary Material Table S1 for details.

SSRs were able to amplify from their own genome (SSR donor genome). For example, for the 43 newly developed EST-SSRs from the pea aphid, 88.1% (36) had successful amplification in the pea aphid, and was 75.0% for the 12 green peach aphid EST-SSRs (Table 3). Failure of amplification for these primer pairs in their respective donor genome may have been as a result of primer mismatch, and in the case of EST-SSRs, the extension of primers across a splice site or the presence of large introns in the genomic DNA fragment to be amplified. Also, in this study, only a single PCR programme was applied to all markers, and no effort was undertaken to optimize amplification conditions for unsuccessful primer pairs.

The overall rate of successful amplification of all 136 SSRs across six species was 51.9%. The success rates for the 53 EST- and 83 genomic SSRs were on average 65.1 and 46.9%, respectively, suggesting EST-SSRs are more transferable than genomic SSRs. However, it should be pointed out that, in this study, the 53 EST-SSRs underwent BLAST searches to maximize the cross-species amplification during the selection process (see Experiment Procedures below). Therefore, the high transferability of these EST-SSRs might overestimate the average for pea aphid or green peach aphid EST-SSRs.

Transferability of EST-SSRs was higher in phylogenetically closer species. Among the seven SSR donor or target aphid species, the pea aphid, green peach aphid and the RWA belong to the tribe Macrosiphini and the bird cherry-oat aphid, cotton aphid, greenbug and the English grain aphid belong to the tribe Aphidini. The transferability of pea aphid and green peach aphid EST-SSRs was on average 72.5%

in three Macrosiphini species and 60.9% in three Aphidini species. However, for the newly developed 41 pea aphid genomic SSRs, there was no difference in cross-species transferability in aphids between the two tribes (24.3 vs. 27.6%). This was also true when all 83 genomic SSRs were considered (47.5 vs. 44.3%).

One interesting observation in this study was the difference in transferability between the 41 newly developed pea aphid genomic SSRs and other 42 previously developed genomic SSRs, which was 26.3 and 51.1%, respectively (based on data in Table 3). The transferability of the 42 previously developed SSRs was similar to Wilson *et al.* (2004) (59% successful amplification in aphid species within the same tribe). The reason for this difference between the two sets of genomic SSRs is not known. As 90.2% (37/41) of the new SSRs were able to amplify their own genome (Table 3), there should have been no problems in primer binding with template DNAs during PCR reactions. One possible reason is sampling error because in the set of published SSRs, the number of markers from each of the five species tested was limited (Table 3). In addition, while all 41 new SSRs were explored from database mining and have perfect SSR motifs, the published SSRs were all developed through genomic library screening and molecular cloning, and 13 out of the 42 SSRs contain compound SSR motifs (see Supplementary Material Table S1). It would be interesting to see whether this has anything to do with the observed transferability difference between the two sets of genomic SSRs.

Eight markers used in this study (SmS10, Sm11, SmS17b, SmS49 and SmS43ii, Mpmz3, Mpmz25 and M86X) have been shown to be X chromosome-linked in different aphid

**Table 3.** Cross-species transferability of expressed sequence tag (EST)- and genomic simple sequence repeats (SSRs) in six aphid species\*

species (Wilson *et al.*, 2004). Wilson *et al.* (2004) found that X-linked loci have higher cross-species amplification rates than autosomal loci. In the present study, the average rate of successful amplification of the eight X-linked markers across six aphid species was 60.4%, whereas that of all 83 genomic SSRs was 46.9%. The X-linked marker Mpmz25 was able to amplify products in all six aphid species tested. Two markers from *Sitobion miscanthi*, SmS10 and SmS43ii, both had successful amplification in five out of the six species (see Supplementary Material Table S2). However, it is not known whether the loci detected by the eight genomic SSR markers are all located in the X chromosomes of the six aphid species examined in the present study. If this is true, X-linked loci may be more conserved in aphids than autosomal loci (Wilson *et al.*, 2004).

It should be pointed out that successful amplification by transferable SSRs across different aphid species may not necessarily suggest that the SSR motif(s) exist in the target species. While this can be verified by sequencing the PCR products, another way is to test allelic diversity in the target genome, as described below in the present study of biotypic variation in the greenbug and RWA genomes. However, nontransferable SSRs failed to amplify because of the absence of respective motifs in the target aphid genomes. One reason for this failure could be that these motifs may have arisen only in the lineage of the SSR-donor species (e.g. pea aphid). Such a hypothesis can be confirmed by testing a large number of additional aphid species to identify the divergence points for these EST-SSRs. The second reason could be that the sequences (ESTs) of the SSR donor and target genomes being compared are not true orthologues carrying the microsatellite.

In this study, we have demonstrated that EST-SSRs from the pea aphid and the green peach aphid are highly transferable even in distantly related aphid species from different tribes. As there are increasingly available genomic resources (ESTs and whole genome sequences), large number of SSR markers can be developed at trivial cost by database mining. In the past, many studies were hindered because of the lack of suitable molecular markers. We believe that this should no longer be a constraint, and we expect SSR markers to be a powerful tool in aphid research, especially in aphid pests that are important in agriculture where less effort has been expended to develop genomic resources.

#### *Genetic diversity of greenbug and Russian wheat aphid biotypes*

All SSR markers being able to amplify in either the greenbug or the RWA genome or both were employed to characterize the genetic diversity among biotypes in each aphid species. In total, 67 SSRs were used, including 25 EST-SSRs (one from the green peach aphid and 24 from the pea aphid), four genomic SSRs of the greenbug and 10 newly

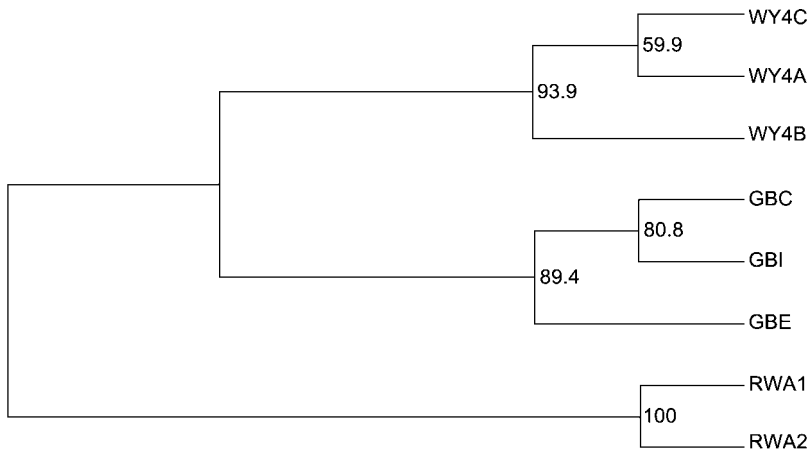
developed pea aphid genomic SSRs. The remaining 28 were published genomic SSRs. Of the 25 EST-SSRs, all are transferable to the greenbug and 22 to the RWA. Among the 42 genomic SSRs, only 28 can amplify both genomes, and seven each amplified the greenbug or the RWA genome only. Therefore, 60 and 57 SSRs had successful amplification in the greenbug and RWA genome, respectively.

Genomic DNAs from six greenbug (C, E, I, WY4A, WY4B and WY4C) and two RWA (RWA1 and RWA2) biotypes were used as templates in PCR with 67 SSR markers. In total, 219 alleles were detected by the 67 SSRs. On average, each EST and genomic SSR was able to detect 2.43 and 1.51 alleles, respectively, in both aphid species. Among the four greenbug genomic SSRs (Sgg1 to Sgg4), Sgg2 and Sgg3 were from the same genomic DNA sequence (GENBANK accession no. DQ887281), and Sgg4 was from the mitochondrial DNA sequence (GENBANK accession no. NC001658, see Supplementary Material Table S1). Sgg4 detected one monomorphic band in each of the eight DNA templates. Sgg1 and Sgg2 each detected six alleles in the greenbug, and on average 1.5 alleles in the RWA. Among the 67 SSRs, 63 were developed from aphid species other than the greenbug or the RWA. This result indicates that both genomic- and EST-SSRs from other aphid species are useful in the cereal aphids, but that EST-SSRs are able to detect more allelic variations. Genomic- and EST-SSRs detected on average 2.14 and 2.84 alleles, respectively in the greenbug, and 1.31 and 1.77, respectively in the RWA.

Polymorphism levels among different greenbug biotypes varied greatly. For example, of the 60 SSRs, 20 were polymorphic between biotypes C and E with a polymorphism level of 33.3%, and between WY4A and WY4C the level was 15.0%. The highest degree of polymorphism, 58.2% (39/67), was between biotypes E and WY4A. The genetic diversity of the six greenbug biotypes is explained well by the phylogenetic tree (cladogram) constructed from the SSR data as shown in Fig. 1. This is a consensus tree based on 1000 bootstrapping repetitions and the two RWA biotypes were used as the outgroups. The percentages supporting particular nodes among the 1000 trees are also indicated in Fig. 1. The lowest probability was 59.9% for the node branching biotypes WY4A and WY4C, suggesting the consensus tree generated from the SSR data is highly reliable.

Clustering analysis grouped the six greenbug biotypes into two clades. The first clade was composed of biotypes C, E and I, which are all 'agricultural' biotypes commonly found on sorghum and wheat. The second clade is made up of biotypes WY4A, WY4B and WY4C, which are all laboratory strains collected from the state of Wyoming.

Using a mitochondrial DNA marker, Shufran *et al.* (2000) grouped all 'agricultural' biotypes (C, E, K, I and J) of the



**Figure 1.** Cladogram of greenbug (GB) biotypes C, E, I, WY4A, WY4B and WY4C, and Russian wheat aphid (RWA) biotypes RWA1 and RWA2 based on 67 genomic and EST-SSR loci. Bootstrap sampling of alleles was carried out for 1000 repetitions. Percentages of these supporting nodes were shown. RWA was used as the outgroup in phylogenetic tree construction.

greenbug into one clade. Similarly, using amplified fragment length polymorphism fingerprinting among eight greenbug biotypes, Zhu-Salzman *et al.* (2003) found that biotypes C, E, I and K share more common polymorphisms among themselves than with other biotypes. Genetic studies and field sampling have suggested that greenbug biotypes are host-adapted races (Shufran *et al.*, 2000; Zhu-Salzman *et al.*, 2003; Burd *et al.*, 2006). Common genetic factors exist among different biotypes, enabling them to predominate and thrive in monoculture crops (Zhu-Salzman *et al.*, 2003). Host-based divergence in populations was also observed in the pea aphid (Via *et al.*, 2000; Simon *et al.*, 2003; Frantz *et al.*, 2006), RWA (Dolatti *et al.*, 2005), lettuce root aphid, *Pemphigus bursarius* (L.) (Miller *et al.*, 2005) and English grain aphid, *Sitobion avenae* (De Barro *et al.*, 1995; Sunnucks *et al.*, 1997). The molecular data from 67 SSRs in the present study clearly support the earlier notion of the host-adapted nature of the greenbug biotypes, but with higher resolution. While grouped into the same clade, biotypes C, E and I can be further separated into two subgroups – C and I in one and E in the other (Fig. 1). This is consistent with their initial host associations. Both C and I were initially identified in sorghum (Harvey & Hackerott, 1969; Harvey *et al.*, 1991), and E was first identified as overcoming the wheat resistance gene *Gb2* (Porter *et al.*, 1982). Clearly, the higher resolving power was a result of more SSRs being used in the present study, all of which are nuclear DNA markers and should be more representative of the greenbug genome.

Apart from host-adapted biotypic genetic variation, regional differentiation among greenbug biotypes is also obvious. Biotypes WY4A, WY4B and WY4C were collected in Wyoming and all of these were grouped into one clade in this study. Geographical differentiation is well known in aphid species (e.g. Martinez-Torres *et al.*, 1997; Simon *et al.*, 1999, 2002; Dedryver *et al.*, 2001; Dolatti *et al.*, 2005; Guo *et al.*, 2005). However, in 'agricultural' biotypes, host association seems to be the determining factor of

biotypic variation. In recent field surveys, biotypes E and I exhibited the greatest host range including major small grain crops and a number of grasses, and they were the only biotypes collected in all four states of the USA surveyed (Nebraska, Kansas, Oklahoma and Texas) (Burd & Porter, 2006). Nevertheless, these 'agricultural' biotypes were consistently grouped into the same clade in several studies (Shufran *et al.*, 2000; Zhu-Salzman *et al.*, 2003; the present study).

Greenbug biotypes have been defined by their virulence relationship to a selected group of plant genotypes (differentials). So far, over 120 greenbug biotypes have been identified with differentials from wheat, barley, sorghum and rye (John Burd, pers. comm.). Obviously, the ability to identify new biotypes depends on the identification of new resistance genes in host plants. All six available wheat differentials (*Gb1* to *Gb6*) have the same reactions to infestation by E and I (Burd & Porter, 2006), which may have been classified as the same biotype if no other differentials were available. This is also true for biotypes WY4A, WY4B and WY4C, which can overcome all six host resistance genes of wheat (John Burd, pers. comm.). This is the basis of the idea that greenbug biotypes are comprised of genetically diverse individuals sharing similar virulence genes (Anstead *et al.*, 2002). Molecular data in the present study clearly support the genetically heterogeneous nature of greenbug biotypes.

The RWA is a native of central Asia that was first identified in the USA in 1986 and was designated as biotype RWA1 when a new biotype, RWA2 was discovered in 2004 (Haley *et al.*, 2004). In the present study, of the 57 SSR markers tested, 19 (33.4%) were polymorphic between RWA1 and RWA2, suggesting obvious genetic differentiation between the two biotypes. This level of diversity between RWA1 and RWA2 is almost the same as that among the 'agricultural' greenbug biotypes C, E and I examined, which was 33.3% (20/60). It is well known that genetic variations exist among RWA populations from different regions of the world (e.g. Puterka *et al.*, 1992; Dolatti *et al.*, 2005), but whether

the biotypic diversity between RWA1 and RWA2 observed here preceded or followed its introduction into the USA is not clear.

## Conclusions

In this study, we explored the databases of the pea aphid and green peach aphid EST and genomic sequences and developed new SSR markers. We examined cross-species transferability of these SSR markers and transferable markers were employed for evaluating genetic diversity of eight greenbug and RWA biotypes. It is clear that the expressed portion and genomic sequences of the pea aphid genome are a rich source of SSRs. While trinucleotide repeats predominate the SSR motifs, the base composition significantly skews toward A and T nucleotides in all repeat types in the pea aphid genome.

The pea aphid and green peach aphid EST-SSRs are highly transferable to other aphid species, even to those in different tribes. These SSRs will be a valuable source of molecular markers in various studies in aphids, especially for genetic and comparative mapping, or evolutionary studies.

SSR-based phylogenetic analysis suggested significant host-associated genetic variation among the greenbug biotypes. Regional differentiation of greenbug biotypes is also evident. Significant genetic variations comparable to 'agricultural' greenbug biotypes also exist within RWA biotypes, even though it was discovered in the USA only about 20 years ago.

## Experimental procedures

### Database mining and development of SSR markers

The pea aphid and green peach aphid ESTs were acquired from dbEST at the NCBI (National Center for Biotechnology Information, Bethesda, MD) and are available at <http://www.ncbi.nlm.nih.gov>. As of 5 March 2007, there were 67 309 pea aphid ESTs corresponding to about 36.7 Mb in size, 17 567 green peach aphid ESTs (approximately 11.0 Mb) deposited in dbEST, and over four million reads of pea aphid genomic sequences deposited by the Pea Aphid Genome Project (<http://www.hgsc.bcm.tmc.edu/projects/aphid/>). All EST sequences and the first 20 000 genomic sequencing reads (total 19.5 Mb) were downloaded for SSR mining and sequence analysis. In addition, several genomic sequences of the greenbug deposited at NCBI were also downloaded for mining and development of SSR primers.

Simple perfect SSRs were screened electronically using the SSR Primer Discovery Program (<http://hornbill.csp.latrobe.edu.au/cgi-bin/pub/ssrprimer/indexssr.pl>). In this program, the minimum length of SSRs was set as 10 bp, and only SSRs with repeat units of 2–5 bp in the motif were reported. For practical use, SSRs with mononucleotide and hexanucleotide repeats were not counted or included in the final statistics. In running this program, all the default parameters were used. The input for this program is EST or genomic sequences in FASTA format, and the output includes a list of SSRs with reports of repeat type, SSR motif, SSR length, position, left and right primer sequences and their  $T_m$ , GC (%),

pair complementarity values at the 3' and 5' ends and the expected PCR product size etc. The output was downloaded into Microsoft Excel for further analysis.

Numerous SSRs were identified from the above database mining. To select EST-SSRs for evaluation of their cross-species transferability, the following steps were taken: (1) select SSRs with minimum length of 20 bp. (2) remove primers with redundant sequences. Many EST sequences in dbEST are highly redundant, including the pea aphid raw genomic sequences. This was carried out by checking the redundancy of primer sequences (use 'Sort' command in Microsoft Excel to align primer sequences). (3) To maximize cross-species transferability, each EST-SSR primer was further subjected to BLAST search against the dbEST database. SSR primers with hits in more than one aphid species were selected, whereas those having BLAST hits in only the SSR source species (pea aphid or green peach aphid) were excluded. However, for selecting pea aphid genomic SSRs, only the first two steps were taken. Finally, 108 new SSRs (43 pea aphid EST-SSRs, 48 pea aphid genomic SSRs, 13 green peach aphid EST-SSRs and four greenbug genomic SSRs) were developed and used in the present study.

For purposes of comparison, 40 additional genomic SSRs from five different aphid species published earlier were also tested in this study. These included 13 (SmS10, SmS16b, SmS17b, SmS19, SmS23, SmS24, SmS30, SmS43ii, SmS49, Sm10, Sm11, Sm12 and Sm17) from *Sitobion miscanthi* (Wilson *et al.*, 2004), four (SavS3.R, SavS3.43, SavS5.L and Sa4Σ) from the English grain aphid, *Sitobion avenae* (Simon *et al.*, 1999; Wilson *et al.*, 2004), eight (Mpmz2, Mpmz3, Mpmz9, Mpmz25, M40A, M49A, M63A and M86X) from the green peach aphid, *Myzus persicae* (Sloane *et al.*, 2001; Wilson *et al.*, 2004), eight (R1.35, R5.10, R2.73, R5.29.b, R6.3, R3.171, R5.138 and R5.50) from the bird cherry-oat aphid, *Rhopalosiphum padi* (Simon *et al.*, 2001) and seven (AIB04M, AIB07M, AIC09M, ApG10M, ApH 05M, ApH 08M and ApH 10M) from pea aphids (Caillaud *et al.*, 2004). Detailed information on all the 148 SSR markers has been listed in Supplementary Material Table S1. Information for all other SSRs from the pea aphid or green peach aphid database mining in this study is also available upon request from the corresponding author.

### Aphids, genomic DNA isolation and PCR

Six aphid species were used to evaluate the transferability of SSR markers: the pea aphid (strain LSR1.G1.AC), cotton aphid (*Aphis gossypii*), greenbug (biotype E), RWA (biotype RWA1), bird cherry-oat aphid and the green peach aphid. Single colonies of the pea aphid and RWA1 were provided by David Stern (Princeton University, Princeton, NJ) and J. P. Michaud (Kansas State University, Hays, KS), respectively. The DNA sample of the cotton aphid was kindly supplied by Gary A. Thompson (University of Arkansas, Little Rock, AR). Colonies of the green peach aphid and the bird cherry-oat aphid were each established from a single nymph from local greenhouse populations at Bushland, Texas, USA.

To investigate biotypic variation with SSR markers, six biotypes of the greenbug (C, E, I, WY4A, WY4B and WY4C) and two biotypes of the RWA (RWA1 and RWA2) were employed. C, E and I were 'agriculture' biotypes originally identified from either wheat or sorghum in the Southern Plains, whereas WY4A, WY4B and WY4C were all laboratory strains collected from wheat fields in the state of Wyoming. The reason for choosing the six biotypes was twofold: to verify the nature of host association and to examine possible geographical differentiation of greenbug biotypes. WY4A,

WY4B and WY4C were provided by John Burd (USDA-ARS, Stillwater, OK) and RWA2 was provided by Gary Puterka (USDA-ARS, Stillwater, OK, USA).

For each species, approximately 50 live aphids were collected, flash frozen and ground into fine powder in liquid nitrogen. The CTAB method (Murray & Thompson, 1980) was used to extract genomic DNAs. Each PCR consisted of 10 ng template DNA, 0.5 µM each of two primers and 1× PCR master mix (Promega Inc., Madison, WI, USA) in a total volume of 10.0 µl, which was performed in a PTC-200 thermocycler (Bio-Rad, Hercules, CA, USA). A single, touch-down PCR program was designed for all primer sets: after 3 min initial denaturation at 95 °C, six cycles were performed for 45 s at 94 °C, 5 min at 68 °C, 1 min at 72 °C with the annealing temperature being reduced by 2 °C per cycle; next, the PCR was continued by eight cycles of 45 s at 94 °C, 2 min at 58 °C, 1 min at 72 °C with the annealing temperature reduced by 1 °C per cycle, followed by 25 cycles of 45 s at 94 °C, 2 min at 50 °C and 1 min at 72 °C.

For most SSR markers, the variation in sizes of PCR products allowed the patterns to be resolved unambiguously in 4% high-resolution agarose gels stained with ethidium bromide. For SSR markers generating DNA fragments of very similar size from different DNA templates, PCR products were size-fractionated in nondenaturing polyacrylamide gels and banding patterns were visualized by SYBR Gold staining (Molecular Probes, Eugene, OR, USA). For quality control, during PCR analysis, primers that amplified null alleles were repeated at least one more time in order to rule out the possibility of the failure of PCR amplification. Detailed information for the PCR amplification products in six aphids with 136 SSR markers is listed in Supplementary Material Table S2.

#### Data analysis

The PCR products were scored in binary format, with the presence of a band being scored as 1 and its absence as 0, thus generating a binary matrix. In the cross-species transferability study, a SSR marker was considered transferable if it successfully amplified target fragment(s) in a species other than the SSR-donor. For evaluation of biotypic diversity, the binary matrix of different biotypes and transferable markers was used as input in data analysis with the software package PHYLIP 3.66 (available at <http://evolution.genetics.washington.edu/phylip.html>; Felsenstein, 1989). Specifically, 1000 bootstrapped data sets were created using SEQBOOT, after which RESTDIST was used to construct pairwise distance matrices using the multiple set option of RESTDIST and site length option of 40, then subjected to neighbour-joining (NJ) cluster analysis using the neighbor option of NEIGHBOR (Saitou & Nei, 1987) and a consensus tree was plotted in CONSENSE. The cladogram of greenbug biotypes was drawn with the computer program TreeView (version 1.6.6; <http://taxonomy.zoology.gla.ac.uk/rod/treeview/>) in which the RWA was treated as the outgroup.

#### Acknowledgements

We thank Sabina Kassymzhanova for technical help. The authors are indebted to Drs John Burd (USDA-ARS, Stillwater, OK), Gary Puterka (USDA-ARS, Stillwater, OK), David Stern (Princeton University, Princeton, NJ), Gary A. Thompson (University of Arkansas, Little Rock, AR) and J. P. Michaud (Kansas State University, Hays, KS) for

providing aphid or aphid DNA samples used in this study. We also thank two anonymous reviewers for suggestions to improve the manuscript. This research was partially supported by the Texas Cropping System Program and the United States Department of Agriculture, National Research Initiative Competitive Grant Program (CSREES 2006-35301-16892) to Y.W.

#### References

- Anstead, J.A., Burd, J.D. and Shufran, K.A. (2002) Mitochondrial DNA sequence divergence among *Schizaphis graminum* (Hemiptera: Aphididae) clones from cultivated and non-cultivated hosts: haplotype and host associations. *Bull Entomol Res* **92**: 17–24.
- Braendle, C., Caillaud, M.C. and Stern, D.L. (2005) Genetic mapping of *aphicarus*-a sex-linked locus controlling a wing polymorphism in the pea aphid (*Acyrtosiphon pisum*). *Heredity* **94**: 435–442.
- Burd, J.D., Porter, D.R., Puterka, G.J., Haley, S.D. and Peairs, F.B. (2006) Biotypic variation among north American Russian wheat aphid (Homoptera: Aphididae) populations. *J Econ Entomol* **99**: 1862–1866.
- Burd, J.D. and Porter, D.R. (2006) Biotypic diversity in greenbug (Hemiptera: Aphididae): characterizing new virulence and host associations. *J Econ Entomol* **99**: 959–965.
- Caillaud, M.C., Mondor-Genson, G., Levine-Wilkinson, S., Mieuze, L., Frantz, A., Simon, J.C., et al. (2004) Microsatellite DNA markers for the pea aphid *Acyrtosiphon pisum*. *Mol Ecol Notes* **4**: 446–448.
- De Barro, P.J., Sherratt, T.N., Brookes, C.P., David, O. and Maclean, N. (1995) Spatial and temporal genetic variation in British field populations of the grain aphid *Sitobion avenae* (F.) (Hemiptera: Aphididae) studied using RAPD-PCR. *Proc R Soc Lond B Biol Sci* **262**: 321–327.
- Dedryver, C.A., Hulle, M., Le Gallic, J.F., Caillaud, M. and Simon, J.C. (2001) Coexistence in space and time of sexual and asexual populations of the cereal aphid *Sitobion avenae*. *Oecologia* **128**: 379–388.
- Delmotte, F., Leterme, N., Gauthier, J.P., Rispe, C. and Simon, J.C. (2002) Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Mol Ecol* **11**: 711–723.
- Diehl, S.R. and Bush, G.L. (1984) An evolutionary and applied perspective of insect biotypes. *Ann Rev Entomol* **29**: 471–504.
- Dolatti, L., Ghareyazie, B., Moharrampour, S. and Noori-Daloi, M.R. (2005) Evidence for regional diversity and host adaptation in Iranian populations of the Russian wheat aphid. *Entomol Exp Appl* **114**: 171–180.
- Felsenstein, J. (1989) PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics* **5**: 164–166.
- Figueroa, C.C., Simon, J.-C., Le Gallic, J.F. and Niemeyer, H.M. (1999) Molecular markers to differentiate two morphologically-close species of the genus *Sitobion*. *Entomol Exp Appl* **92**: 217–223.
- Frantz, A., Plantegenest, M., Mieuze, L. and Simon, J.C. (2006) Ecological specialization correlates with genotypic differentiation in sympatric host-populations of the pea aphid. *J Evo Biol* **19**: 392–401.
- Fuller, S.J., Chavigny, P., Lapchin, L. and Vanlerberghe-Masutti, F. (1999) Variation in clonal diversity in glasshouse infestations of



- the aphid, *Aphis gossypii* Glover in southern France. *Mol Ecol* **8**: 1867–1877.
- Grover, A., Aishwarya, V. and Sharma, P.C. (2007) Biased distribution of microsatellite motifs in the rice genome. *Mol Genet Genomics* **277**: 469–480.
- Guo, W., Shen, Z.R., Li, Z.H. and Gao, L.W. (2005) Migration and population genetics of the grain aphid *Macrosiphum miscanti* (Takahashi) in relation to the geographic distance and gene flow. *Progress in Natural Science* **15**: 1000–1004.
- Haley, S.D., Peairs, F.B., Walker, C.B., Rudolph, J.B. and Randolph, T.L. (2004) Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Sci* **44**: 1589–1592.
- Harvey, T.L. and Hackerott, H.L. (1969) Recognition of a greenbug biotype injurious to sorghum. *J Econ Entomol* **62**: 776–779.
- Harvey, T.L., Kofoid, K.D., Martin, T.J. and Sloderbeck, P.E. (1991) A new greenbug virulent to E-biotype resistant sorghum. *Crop Sci* **31**: 1689–1691.
- Hawthorne, D.J. and Via, S. (2001) Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* **412**: 904–907.
- La Rota, M., Kantety, R.V., Yu, J.-K. and Sorrells, M.E. (2005) Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley. *BMC Genomics* **6**: 23.
- Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E. (2004) Microsatellites within genes: structure, function, and evolution. *Mol Biol Evol* **21**: 991–1007.
- Llewellyn, K.S., Loxdale, H.S., Harrington, R. Harrington, R., Brookes, C.P., Clark, S.J., *et al.* (2003) Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. *Mol Ecol* **12**: 21–34.
- Martinez-Torres, D., Moya, A., Hebert, P.D.N. and Simon, J.C. (1997) Geographic distribution and seasonal variation of mitochondrial DNA haplotypes in the aphid *Rhopalosiphum padi* (Hemiptera: Aphididae). *Bulletin Entomol Res* **87**: 161–167.
- Massonnet, B., Simon, J.-C., Weisser, W.W. (2002) Metapopulation structure of the specialized herbivore *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). *Mol Ecol* **11**: 2511–2521.
- Megléc, E., Anderson, S.J., Bourguet, D., Butcher, R., Caldas, A., Cassel-Lundhagen, A., *et al.* (2007) Microsatellite flanking region similarities among different loci within insect species. *Insect Mol Biol* **16**: 175–185.
- Metzgar, D., Bytof, J., Wills, C. (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res* **10**: 72–80.
- Miller, N.J., Kift, N.B. and Tatchell, G.M. (2005) Host-associated populations in the lettuce root aphid, *Pemphigus bursarius* (L.). *Heredity* **94**: 556–564.
- Morgante, M. and Plivieri, A.M. (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant J* **3**: 175–182.
- Morgante, M., Hanafey, M. and Powell, W. (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat Genet* **30**: 194–200.
- Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight DNA. *Nucleic Acids Res* **8**: 4321–4325.
- Papura, D., Simon, J.-C., Halkett, F., Delmotte, F., Le Gallic, J.F. and Dedryver, C.A. (2003) Predominance of sexual reproduction in Romanian populations of the aphid *Sitobion avenae* inferred from phenotypic and genetic structure. *Heredity* **90**, 397–404.
- Peng, J.H. and Lapitan, N.L. (2005) Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Funct Integr Genomics* **5**: 80–96.
- Porter, D.R., Burd, J.D. and Teetes, G. (1997) Greenbug (Homoptera, Aphididae) biotypes, selected by resistant cultivars or preadapted opportunists? *J Econ Entomol* **90**: 1055–1065.
- Porter, K.B., Peterson, G.L. and Vise, O. (1982) A new greenbug biotype. *Crop Sci* **22**: 847–850.
- Powell, W., Morgante, M., Andre, C., Henfey, M., Vogel, J., Tingy, S., *et al.* (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol Breed* **2**: 225–238.
- Puterka, G.J. and Peters, D.C. (1989) Inheritance of greenbug, *Schizaphis graminum* (Rondani), virulence to *Gb2* and *Gb3* resistance genes in wheat. *Genome* **32**: 109–114.
- Puterka, G.J. and Peters, D.C. (1995) Genetics of greenbug, *Schizaphis graminum* (Rondani) virulence to resistance genes in sorghum. *J Econ Entomol* **88**: 421–429.
- Puterka, G.J., Burd, J.D. and Burton, R.L. (1992) Biotypic variation in a worldwide collection of Russian wheat aphid (Homoptera, Aphididae). *J Econ Entomol* **85**: 1497–1506.
- Rajendrakumar, P., Biswal, A.K., Balachandran, S.M., Srinivasarao, K. and Sundaram, R.M. (2007) Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. *Bioinformatics* **23**: 1–4.
- Sabater-Muñoz, B., Legeai, F., Risper, C., Bonhomme, J., Dearden, P., Dossat, C., *et al.* (2006) Large-scale gene discovery in the pea aphid *Acyrtosiphon pisum* (Hemiptera). *Genome Biol* **7**: R21.
- Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol & Evol* **4**: 406–425.
- Shufran, K.A., Burd, J.D., Anstead, J.A. and Lushai, G. (2000) Mitochondrial DNA sequence divergence among greenbug (Homoptera : Aphididae) biotypes: evidence for host-adapted races. *Insect Mol Biol* **9**: 179–184.
- Shufran, K.A., Margolies, D.C. and Black, W.C.IV. (1991) DNA fingerprinting to study spatial and temporal distributions of an aphid *Schizaphis graminum* (Homoptera: Aphididae). *Bull Entomol Res* **82**: 407–416.
- Simon, J.-C., Baumann, S., Sunnucks, P., Hebert, P.D.N., Pierre, J.-S., Le Gallic, J.-F., *et al.* (1999) Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Mol Ecol* **8**: 531–545.
- Simon, J.-C., Leterme, N., Delmotte, F., Martin, O. and Estoup, A. (2001) Isolation and characterization of microsatellite loci in the aphid species, *Rhopalosiphum padi*. *Mol Ecol Notes* **1**: A33.
- Simon, J.-C., Risper, C. and Sunnucks, P. (2002) Ecology and evolution of sex in aphids. *Trend Ecol Evol* **17**: 34–39.
- Simon, J.-C., Carre, S., Boutin, M., Prunier-Leterme, N., Sabater-Munoz, B., Latorre, A., *et al.* (2003) Host-based divergence in populations of the pea aphid: insights from nuclear markers and the prevalence of facultative symbionts. *Proc R Soc Lond B Biol Sci* **270**: 1703–1712.
- Sloane, M.A., Sunnucks, P., Wilson, A.C.C. and Hales, D.F. (2001). Microsatellite isolation, linkage group identification and determination of recombination frequency in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). *Genet Res* **77**: 251–260.
- Sourdille, P., Tavaud, M., Charmet, G. and Bernard, M. (2001) Transferability of wheat microsatellites to diploid Triticeae

- species carrying the A, B and D genomes. *Theor Appl Genet* **103**: 346–352.
- Subramanian, S., Mishra, R.K. and Singh, L. (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol* **4**: R13.
- Sunnucks, P., De Barro, P.J., Lushai, G., Maclean, N. and Hales, D.F. (1997) Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages and host specialization. *Mol Ecol* **6**: 1059–1073.
- Sunnucks, P., England, P.R., Taylor, A.C. and Hales, D.F. (1996) Microsatellite and chromosome evolution of parthenogenetic *Sitobion* aphids in Australia. *Genetics* **144**: 747–756.
- Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* **17**: 6463–6470.
- Tóth, G., Gaspari, Z. and Jurka, J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res* **10**: 967–981.
- Vanlerberghe-Masutti, F., Chavigny, P. and Fuller, S.J. (1999) Characterization of microsatellite loci in the aphid species *Aphis gossypii* Glover. *Mol Ecol* **8**: 693–695.
- Via, S., Bouck, A.C. and Skillman, S. (2000) Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution* **54**: 1626–1637.
- Wilson, A.C.C., Massonnet, B., Simon, J.C., Prunier-Leterme, N., Dolatti, L., Llewellyn, K.S., et al. (2004) Cross-species amplification of microsatellite loci in aphids: assessment and application. *Mol Ecol Notes* **4**: 104–109.
- Wilson, A.C.C., Sunnucks, P. and Hales, D.F. (1999) Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. *Mol Ecol* **8**: 1655–1666.
- Wilson, A.C.C., Sunnucks, P., Blackman, R.L. and Hales, D.F. (2002) Microsatellite variation in cyclically parthenogenetic populations of *Myzus persicae* in south-eastern Australia. *Heredity* **88**: 258–266.
- Yu, J.K., La Rota, M., Kantety, R.V. and Sorrells, M.E. (2004) EST derived SSR markers for comparative mapping in wheat and rice. *Mol Genet Genomics* **271**: 742–751.
- Zhang, L.Y., Bernard, M., Leroy, P., Feuillet, C. and Sourdille, P. (2005) High transferability of bread wheat EST-derived SSRs to other cereals. *Theor Appl Genet* **111**: 677–687.
- Zhao, X. and Kochert, G. (1993) Phylogenetic distribution and genetic mapping of a (GGC)<sub>n</sub> microsatellite from rice (*Oryza sativa* L.). *Plant Mol Biol* **21**: 607–614.
- Zhu-Salzman, K., Li, H., Klein, P.E., Gorena, R.L. and Salzman, R.A. (2003) Using high-throughput amplified fragment length polymorphism to distinguish sorghum greenbug (Homoptera: Aphididae) biotypes. *Agri Forest Entomol* **5**: 311–315.

### Supplementary material

The following supplementary material is available for this article:

**Table S1.** Information for 148 genomic and EST-SSRs used in the present study.

**Table S2.** Cross-species amplification of 136 SSR markers among six aphid species.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/abs/10.1111/j.0307-6975.2007.00757.x>

(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.