Entomologia
Experimentalis et Applicata

DOI: 10.1111/eea.12460

In silico mining of microsatellites and analysis of genetic diversity among inter- and intra-generic aphids of the subfamily Aphidinae

Sridhar Vaddi[1]*, Venkatesan Thiruvengadam[2], Sunil Joshi[3], Sushil Kumar Jalali[2], Lakshmana Reddy Dhoranalapalli Chinnappareddy[4], Padma Nimmakayala[5], Yan Tomason[5], Gopinath Vajja[5] & Umesh K. Reddy[5]

[1]Division of Entomology and Nematology, ICAR-Indian Institute of Horticultural Research, Bengaluru 560089, India, [2]Division of Molecular Entomology, National Bureau of Agricultural Insect Resources, Bengaluru 560024, India, [3]Division of Insect Systematics, National Bureau of Agricultural Insect Resources, Bengaluru 560024, India, [4]Division of Biotechnology, ICAR-Indian Institute of Horticultural Research, Bengaluru 560089, India, and [5]Department of Biology, Gus R. Douglas Institute, West Virginia State University, WV 25112-1000, USA

Accepted: 1 March 2016

Key words: cross-amplification, phenogram, polymorphism, tribe, *Aphis gossypii*, *Acyrthosiphon pisum*, Hemiptera, Aphididae

Abstract

Nearly 5 000 aphid species damage crops, either by sucking plant sap or as disease-transmitting vectors. Microsatellites are used for understanding molecular diversity and eco-geographical relationships among aphid species. Expressed sequence tag (EST)-microsatellite motifs were identified through an in silico approach using inbuilt simple sequence repeat mining tools in aphid EST dataset. Microsatellite mining revealed one in every five aphid genes as containing a repeat motif, and out of 9 290 EST microsatellites mined from *Aphis gossypii* Glover and *Acyrthosiphon pisum* (Harris) (both Hemiptera: Aphididae), 80% were of A and/or T (AT, ATA, AAT, AATA, and ATTT) motifs, and the rest contained G and/or C motifs. All microsatellite sequences were annotated using BLAST. Primers for EST microsatellites were designed using the Primer 3.0 tool. 106 primer pairs of both dinucleotide repeats (DNRs) and trinucleotide repeats (TNRs), representing open reading frames (ORFs) and untranslated regions (UTRs), were synthesized to amplify 15 aphid species belonging to the subfamily Aphidinae, collected from diverse hosts. Four hundred forty-five polymorphic alleles were amplified. Fifty TNR and 23 DNR microsatellites amplified across the species studied. Polymorphism information content values of microsatellites ranged from 0.23 to 0.91, amplifying 2–16 alleles. Genetic similarity indices were estimated using the 'NTSYS-pc' software package. Unweighted pair group with arithmetic mean and principal component analysis resolved taxonomic relationships of the aphid species studied. The new aphid microsatellites developed will provide valuable information to researchers to study Indian aphid species diversity and genetic relationships.

Introduction

With approximately 5 000 species, aphids are the most widespread group of pests in agricultural systems (Remaudiere & Remaudiere, 1997). They cause extensive losses in field crops, horticultural crops, and forest trees, either directly or indirectly, by transmitting several viral diseases as vectors (Minks & Harrewijn, 1988; Sandstrom & Moran, 1999). Among the aphids, members of the subfamily Aphidinae are polyphagous in nature and are known to be prominent pests devastating both agricultural and horticultural crops. Aphids migrate long and short distances. Study of their dispersal ecology and genetic population structure is very important from an economic point of view, mostly owing to their nature as pests (Loxdale et al., 1993).

Molecular tools are used for characterizing biotypes, gene flow, genetic structure, diversity, nature of dispersal, and extent of mutability among species at the population

*Correspondence: Sridhar Vaddi, Division of Entomology and Nematology, ICAR-Indian Institute of Horticultural Research, Bengaluru 560089, India. E-mail: vsridhar@iihr.res.in

level. Previously, isozymes were used for studying genetic structure in aphids (Loxdale et al., 1993). Recent reports show that mitochondrial DNA (mtDNA)-specific markers are used for studying phylogenetic populations, genetics, and evolutionary aspects. Mitochondrial molecular data have helped immensely in understanding the genealogies or phylogenies and also geographical distribution to some extent (Sunnucks, 2000; Coeur d'Acier et al., 2007). Lately microsatellites have replaced mtDNA markers and isozymes, as these are polymorphic in nature, are abundant, are multi-allelic, and have high transferability. As expressed sequence tag (EST) microsatellites are developed from expressed regions, these regions are expected to be conserved across species compared to neutral microsatellites resulting in better transferability of the EST-based microsatellites.

Plieske & Struss (2001) have shown that microsatellite markers developed for one species can also be used in related species with no additional cost for primer development. Microsatellites also allow easier data integration because they are amplified by using very stringent polymerase chain reaction (PCR) conditions. Microsatellites are used for answering biological questions on gene frequencies, gene flow, genetic structure, nature of dispersal, and mutability in aggressive sexual species such as aphids, where genotypic arrays are constantly reshuffled in every generation (Sunnucks et al., 1997; Sunnucks, 2000; Caillaud et al., 2004; Jun et al., 2011). In silico mining of microsatellites using EST sequences is considered more economical compared with the higher cost involved in development of microsatellites through next-generation sequencing (Martins et al., 2012). In addition, higher transferability of microsatellites has been reported (Weng et al., 2007). EST microsatellites have some intrinsic advantages over genomic microsatellites because they are quickly obtained by mining and are present in expressed regions of the genome. The usefulness of these EST microsatellites also lies in their expected transferability because the primers are designed from the more conserved coding regions of the genome (Varshney et al., 2006). With an increasing availability of ESTs, EST microsatellites can be mined at meager costs from EST databases (Sourdille et al., 2001; Zhang et al., 2005; Yu & Li, 2007). Therefore, EST microsatellites were used to study phylogenetic relationships between species (Sunnucks et al., 1996; Simon et al., 1999a,b; Massonnet et al., 2002; Llewellyn et al., 2003), genetic diversity (Figueroa et al., 1999; Wilson et al., 2002; Simon et al., 2003), sexual reproduction (Sunnucks et al., 1997; Delmotte et al., 2002; Papura et al., 2003), or evolution (Wilson et al., 1999). However, all these studies are based on only a few publicly available genomic microsatellites, isolated from a limited number of aphid species (Sunnucks et al., 1996; Simon et al., 1999a; Vanlerberghe-Masutti et al., 1999; Sloane et al., 2001; Caillaud et al., 2004; Wilson et al., 2004). Aphid species from which microsatellites have been isolated include *Aphis glycines* Matsumura (Jun et al., 2011), *Sitobion miscanthi* (Takahashi) (Sunnucks et al., 1996; Wilson et al., 1997; Simon et al., 1999b), *Aphis gossypii* Glover (Vanlerberghe-Masutti et al., 1999), *Myzus persicae* (Sulzer) (Sloane et al., 2001), and *Macrosiphoniella tanacetaria* (Kaltenbach) (Massonnet et al., 2002). Feasibility of cross-species amplification of microsatellite loci was also reported in different genera of aphids, namely *Rhopalosiphum* and *Acyrthosiphon* (Wilson et al., 2004). However, to the best of our knowledge, very few accessible EST microsatellites have been reported, characterized, or evaluated in aphid species, with particular reference to subfamily Aphidinae.

A broad diversity of aphid species is reported from India, and most of them have been grouped into tribes within subfamilies. However, there is a lack of information on isolation and characterization of microsatellite loci of aphid species of various tribes in India. Therefore, this study was aimed at differentiating 15 aphid species (parthenogenetic, viviparous) belonging to the tribes Aphidini and Macrosipihini of the subfamily Aphidinae (Hemiptera: Aphididae), collected from southern India on various host plants. Gene-specific microsatellites were mined, and transferability by cross-species amplification was ascertained for use in estimating molecular relationships among species.

Materials and methods

Insect collections

Ten aphid species of the tribe Aphidini and five species of the tribe Macrosiphini were used in this study. They were collected from South India except *Macrosiphum euphorbiae* (Thomas), which was collected from the USA and included as an outgroup (Table 1) and maintained as homogenous populations under laboratory conditions.

DNA extraction

DNA was isolated from five parthenogenetic, viviparous specimens of each taxon using the QIAquick kit (Qiagen, Valencia, CA, USA), according to the manufacturer's instructions.

Microsatellite mining

Non-redundant EST-microsatellite sequences belonging to Aphididae [41 517 sequences from the pea aphid, *Acyrthosiphon pisum* (Harris), and 3 050 from the cotton aphid, *A. gossypii*] were downloaded in FASTA format from http://insectacentral.com and used to identify

Table 1 List of aphid species of the subfamily Aphidinae and their respective host plants

Tribe	Aphid species	Host plants
Aphidini	*Aphis affinis* del Guercio	*Mentha viridis* L. (Lamiaceae)
	Aphis craccivora Koch	*Vigna unguiculata* (L.) Walp. (Fabaceae)
	Aphis fabae Scopoli	*Solanum nigrum* L. (Asteraceae)
	Aphis nerii Boyer de Fonscolombe	*Calotropis gigantea* (L.) (Asclepiadaceae)
	Aphis gossypii Glover	*Gossypium* spec. (Malvaceae)
	Rhopalosiphum maidis (Fitch)	*Zea mays* L. (Poaceae)
	Schizaphis graminum (Rondani)	*Cynodon dactylon* (L.) Pers. (Poaceae)
	Brachysiphoniella montana (van der Goot)	Unidentified weed (Poaceae)
	Hysteroneura setariae (Thomas)	*Eleusine coracana* Gaertn. (Poaceae)
	Melanaphis sacchari (Zehntner)	*Sorghum bicolor* (L.) Moench (Poaceae)
Macrosiphini	*Brachycaudus helichrysi* (Kaltenbach)	*Chromolaena odorata* (L.) (Asteraceae)
	Brevicoryne brassicae (L.)	*Brassica campestris* L. (Brassicaceaea)
	Lipaphis erysimi (Kaltenbach)	*Raphanus sativus* L. (Brassicaceae)
	Uroleucon sonchi (L.)	*Sonchus arvensis* L. (Asteraceae)
	Macrosiphum euphorbiae (Thomas)	*Cucurbita* spec. (Cucurbitaceae) (collected from West Virginia, USA)

microsatellite motifs using a mining tool available at the Cotton Microsatellite Database (www.cottonmarker.org). We selected the following microsatellite repeat units: dinucleotide repeats (DNR) if ≥7 repeats, trinucleotide repeats (TNR) if ≥6 repeats, tetranucleotide repeats (TTNR) if ≥5 repeats, and pentanucleotide repeats (PTR) if ≥4 repeats. Randomly, DNR and TNR microsatellite regions were selected for primer design using Primer3 software (Rozen & Skaletsky, 2000) as DNRs and TNRs are abundant and well distributed in the genome.

PCR reactions

PCRs were carried out in a total volume of 10 µl containing 10 ng of the DNA template, 1× Taq buffer, 2 mM $MgCl_2$, 0.2 mM deoxynucleotide triphosphates (dNTPs), 1 U Taq DNA polymerase (Fermentas, Waltham, MA, USA), and 0.5 µM each of the forward and reverse primers. Amplification was performed in a GeneAmp PCR 9700 System Thermal Cycler (Applied Biosystems, Waltham, MA, USA) programed as follows: 94 °C for 2 min, followed by 35 cycles of 94 °C for 30 s, 50–65 °C for 30 s, 72 °C for 1 min, and a final extension step at 72 °C for 10 min; controls were maintained by adding sterile water. Amplified products were separated using 3% superfine resolution agarose gel. When scoring, stutters were avoided, and only discernible bands were taken as alleles.

Data analysis

Presence or absence of each of the fragments was scored as a binary unit character (present = 1, absent = 0). Genetic similarities based on Jaccard's coefficients were calculated using the SIMQUAL program of the Numerical Taxonomy Multivariate Analysis System (NTSYS-pc) v.2.0 software package (Rohlf, 1997). The resulting genetic similarity indices were used for generating a tree, using UPGMA, NTSYS-pc version 2 (Rohlf, 1997). Robustness of the clustering was verified by bootstrapping (Felsenstein, 1985) using PAUP v.4.0. Principal component analysis (PCA) based on genetic similarity matrices was performed using DCENTER and EIGEN algorithms of the NTSYS-pc software package. Polymorphic information content (PIC) value was calculated using the standard formula (Anderson et al., 1993; Eujayl et al., 2004).

Results

Characterization of microsatellite motifs

We report 9 290 microsatellites (2 724 from ORF, 6 266 from 5′ UTR and 3′ UTR regions) mined from Aphididae (41 517 ESTs from the pea aphid, *A. pisum*, and 3 050 from the cotton aphid, *A. gossypii*). Number of repeat units ranged 7–65 for DNRs, 6–23 for TNRs, and 5–14 for TTNRs. AT, ATA, AAT, AATA, and ATTT were 80%, and the remaining 20% were repeat classes that contained Gs and/or Cs. Of the various microsatellites isolated, DNRs constituted 24%, TNRs 46%, TTNRs 21%, and PNRs 9%. The most common motif type of DNRs was AT/TA (75%), followed by AC/TG (10%) and CA/GT (8%). Among TNRs, the highest count obtained was for AAT/TTA (72%) followed by CCA/GGT (11%). TTNRs constituted 21% of the total repeats, of which AAAT/TTTA was the most frequent (74%) followed by ATGC/TACG (7%).

Primer design

We designed 1 890 primer pairs for repeat motifs that consisted of high-quality flanking sequences, and the

expected amplicon sizes ranged from 100 to 300 bp. With this stringency, 433 primer pairs could be designed for tetra microsatellites, followed by 1 013 for TNRs and 445 for DNRs. For the rest of the SSR-containing sequences, primers could not be picked reliably as the flanking sequence quality was not suitable for picking up the ideal primers.

Amplification pattern, polymorphism, and cross-genera transferability

We tested a subset of 106 randomly picked DNR and TNR primer pairs to amplify 15 aphid species collections included in the study. Except for 15 microsatellites, all the primer sets amplified appropriately sized bands in aphid species. Out of 91 microsatellites, 73 markers were polymorphic (Table S1) and fully amplified across the species belonging to both tribes, with comparable size of the amplicons. They were considered as transferable markers. PIC was estimated for all 73 transferable markers and was found to be in the range 0.23–0.91, amplifying 2–16 alleles, respectively. Of the 73 microsatellites, 50 were TNRs, the remaining 23 were DNRs (Table S1). Interestingly, in this study, more alleles were amplified for TNRs than for DNRs.

Analysis of gene content and annotation

The microsatellites containing 73 EST sequences were annotated using the sequence-similarity search program BLAST, with 'blastx' option, so as to predict the gene and its function using a threshold E-value cutoff at 1×10^{-6} against the Swiss-Prot protein database. This analysis revealed that 45 EST sequences were from unknown genes, and the remaining 28 could be annotated to known genes (Table 2).

Molecular relationships

Using 73 microsatellite markers, in total 445 amplicons were obtained and found to be transferable across tribes in this study. Microsatellite analysis revealed that maximum

Table 2 Functional annotation of 28 cross-amplified microsatellite markers in aphids matching with genes from genomic databases of various organisms

Marker name	Present in ORF	BLAST score	Match	Gene	Organism
ap1	–	3×10^{-79}	Q05974	*Ras-related protein Rab-1A*	*Lymnaea stagnalis*
ap4	–	6×10^{-50}	Q94518	*Nascent polypeptide-assoc complex subunit α*	*Drosophila melanogaster*
ap6	–	2×10^{-17}	Q0IIM1	*E3 ubiquitin-protein ligase RNF168*	*Bos taurus*
ap12	Y	4×10^{-13}	Q9U3V5	*Protein tiptop*	*D. melanogaster*
ap15	N	3×10^{-77}	P08266	*DNA-directed RNA polymerase II subunit RPB2*	*D. melanogaster*
ap17	Y	4×10^{-10}	B1H1Z8	*Cohesin loading complex subunit SCC4 homolog*	*Xenopus tropicalis*
ap19	N	6×10^{-85}	P25867	*Ubiquitin-conjugating enzyme E2-17 kda*	*D. melanogaster*
ap23	Y	6×10^{-11}	O43290	*U4/U6.U5 tri-snrnp-associated protein 1*	*Homo sapiens*
ap28	Y	1×10^{-142}	P12261	*Elongation factor 1-gamma*	*Artemia salina*
ap31	N	4×10^{-12}	P0C6S7	*Ankyrin repeat and sterile alpha motif domain-containing protein 1B*	*Rattus norvegicus*
ap44	Y	0	P82798	*Transcriptional regulator ATRX* (Fragment)	*Macropus eugenii*
ap46	Y	0	P87268	*Putative uncharacterized protein YDR426C*	*Saccharomyces cerevisiae*
ap58	N	2×10^{-58}	Q9NFT7	*Hexokinase type 2*	*D. melanogaster*
ap59	Y	6×10^{-73}	P23380	*V-type proton atpase 16 kdaproteolipid subunit*	*D. melanogaster*
ap63	–	6×10^{-47}	P47947	*Troponin C, isoform 1*	*D. melanogaster*
ap66	N	0	Q19157	*LIM domain-containing protein pin-2*	*Caenorhabditis elegans*
ap77	N	4×10^{-7}	Q9BUN5	*Coiled-coil domain-containing protein 28B*	*H. sapiens*
ap79	N	5×10^{-12}	P26262	*Plasma kallikrein*	*Mus musculus*
ap90	Y	6×10^{-27}	P23380	*V-type proton atpase 16 kdaproteolipid subunit*	*D. melanogaster*
ap94	N	4×10^{-13}	Q9U3V5	*Protein tiptop*	*D. melanogaster*
ap95	Y	0	P04755	*Acetylcholine receptor subunit beta-like 1*	*D. melanogaster*
ap96	Y	3×10^{-52}	P40320	*S-adenosylmethionine synthase*	*D. melanogaster*
ap98	–	1.3×10^{-2}	Q8IWY8	*Zinc finger and SCAN domain-containing protein*	*H. sapiens*
ap101	Y	1×10^{-5}	Q9U3V5	*Protein tiptop*	*D. melanogaster*
ap102	Y	3×10^{-5}	Q6NRS2	*PQ-loop repeat-containing protein 1*	*X. laevis*
ap103	N	6×10^{-47}	P47947	*Troponin C, isoform 1*	*D. melanogaster*
ap104	Y	3×10^{-79}	P62925	*Eukaryotic translation initiation factor 5A*	*S. frugiperda*
ap105	–	1×10^{-13}	P91753	*Mitotic apparatus protein p62*	*Lytechinus pictus*

genetic distance (GD) was 73%, whereas the minimum was 36% among various taxa in the study. Genetic distance was larger between tribe groups than within genera and species, as predicted. Our study revealed that GD between Aphidini and Macrosiphini tribes was 68%, and within a species or genus of Aphidini and Macrosiphini, GD was 60 and 64%, respectively. An UPGMA phenogram was constructed using the entire dataset (Figure 1). Overall robustness of phenetic and putative phylogenetic topologies was evaluated by (1) bootstrap analysis with 1 000 replicates and (2) comparison with a neighbor-joining tree and PCA. Bootstrap values (BV) generally ranged from 50 to 100, indicating that the clusters supported with higher bootstrap values may be considered as robust groupings of various taxa. A monophyletic cluster of Aphidini was basally resolved with a BV support of 95. This cluster further splits into two sister clades: one with *Aphis affinis* del Guercio and *Aphis nerii* Boyer de Fonscolombe and the other with *Aphis craccivora* Koch, *Aphis fabae* Scopoli, *Brachysiphoniella montana* (van der Goot), *A. gossypii*, *Schizaphis graminum* (Rondani), and *Rhopalosiphum maidis* (Fitch). Another monophyletic cluster of the tribe Macrosiphini formed with a BV of 60, with one major exception of joining, viz. the species *Hysteroneura setariae* (Thomas), originally classified as member of the tribe Aphidini. This species figured on top of the cluster displaying affinity to the Aphidini. Two other species formed a separate lineage occupying the bottom of the tree, one belonging to the tribe Aphidini [*Melanaphis sacchari* (Zehntner)], and the other to the tribe Macrosiphini (*Ma. euphorbiae*), which was used as an independent outgroup. Scale length between *Hy. setariae* and the Macrosiphini cluster is very narrow (Figure 1). Even though *Me. sacchari* belongs to the Aphidini tribe morphologically, it genetically shows near similarity with the tribe Macrosiphini; this needs to be investigated further. The first three Eigen vectors of PCA absorbed 32.1, 20.0, and 15.5% variation, totaling 67.6% of the overall variation, thus indicating robustness of the dataset and reliability of the analysis. The pattern of clustering obtained from PCA was more robust for separating taxa of different tribes into separate clusters (Figure 2). Incongruities observed in the tree topologies pertaining to *Me. sacchari* and *Ma. euphorbiae* were not found in PCA analysis, thus indicating its importance in molecular divergence analysis of aphids.

Discussion

Aphid species are difficult to identify and categorize owing to their small size and morphological similarity. Moreover, morphological variation is known to occur in aphids due to environmental factors (Loxdale & Brookes, 1989; Cenis et al., 1993; Figueroa et al., 1999). To overcome these problems, advanced techniques like isozyme polymorphism or molecular markers have been used for assessing

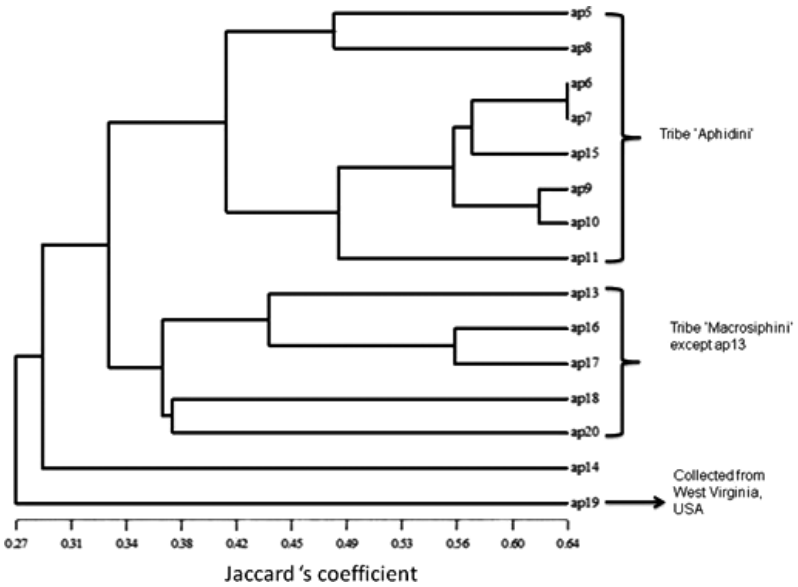

Figure 1 Phenogram showing 15 aphid species grouped under various clusters based on microsatellite amplification and polymorphic data. ap5: *Aphis affinis*; ap6: *A. craccivora*; ap7: *A. fabae*; ap8: *A. nerii*; ap9: *A. gossypii*; ap10: *Schizaphis graminum*; ap11: *Rhopalosiphum maidis*; ap13: *Hysteroneura setariae*; ap14: *Melanaphis sacchari*; ap15: *Brachysiphoniella montana*; ap16: *Uroleucon sonchi*; ap17: *Lipaphis erysimi*; ap18: *Brachycaudus helichrysi*; ap19: *Macrosiphum euphorbiae*; ap20: *Brevicoryne brassicae*.

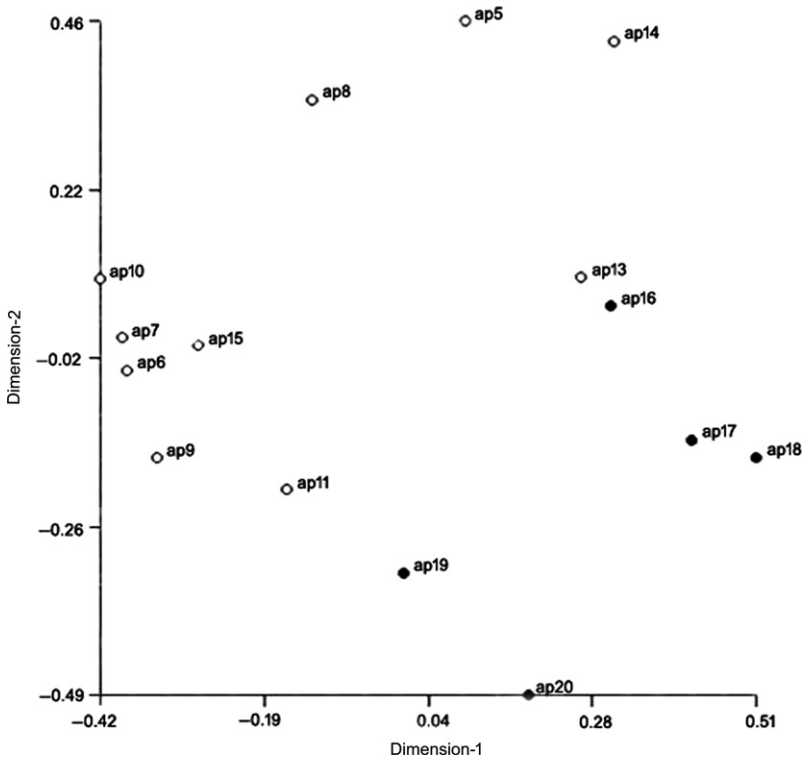

Figure 2 Principal component analysis of 15 aphid species belonging to the tribes Aphidini (open dots) and Macrosiphini (filled dots) of the subfamily Aphidinae. ap5: *Aphis affinis*; ap6: *A. craccivora*; ap7: *A. fabae*; ap8: *A. nerii*; ap9: *A. gossypii*; ap10: *Schizaphis graminum*; ap11: *Rhopalopsiphum maidis*; ap13: *Hysteroneura setariae*; ap14: *Melanaphis sacchari*; ap15: *Brachysiphoniella montana*; ap16: *Uroleucon sonchi*; ap17: *Lipaphis erysimi*; ap18: *Brachycaudus helichrysi*; ap19: *Macrosiphum euphorbiae*; ap20: *Brevicoryne brassicae*.

aphid genetic diversity (Black et al., 1992). Our major objective in this study was to assess inter- and intra-generic relationships and genetic diversity among a selection of aphid species belonging to the subfamily Aphidinae, distributed in South India.

We mined 1 890 usable microsatellite markers, one of the largest microsatellite resources of Aphididae, and tested a subset of 106 for amplification. Fifteen of the 106 markers resolved alleles in the expected size range, pointing to the usefulness of a majority of the archived markers for characterizing these important pests of field, forest, and horticulture crops. As in our study, SSR markers have been developed from genomic resources specific to *A. glycines* (Michel et al., 2009), asexual lineages of various *Sitobion* species (Sunnucks et al., 1996; Simon et al., 1999a; Wilson et al., 1999), a few microsatellites have been identified from *M. persicae* ESTs or genomic DNA sequences (Caillaud et al., 2004), and 18 microsatellite loci from *A. gossypii* and *A. fabae*.

BLAST search and annotation classified these markers as located within genes, which proved very useful for developing functional markers for a variety of genetic studies. Out of the several microsatellites isolated, DNRs accounted for 24%, whereas TNRs, TTNRs, and PNRs accounted for 46, 21, and 9%, respectively. Similarly, the distribution of perfect microsatellites spanning 10 bp or more in the pea aphid EST and genomic sequences were analyzed (Meglécz et al., 2007). The observed 0.86% genomic SSR abundance in pea aphids is similar to that in several *Drosophila* spp. (Meglécz et al., 2007). A similar feature of base composition in microsatellite motifs is also found in the human genome (Subramanian et al., 2003), whereas in some cereal genomes such as rice, *Oryza sativa* L., CCG is the most abundant of all the TNRs (Zhao & Kochert, 1993; Grover et al., 2007).

Microsatellites have been widely used in many other insects as DNA markers (Hughes & Queller, 1993; Goldstein & Clark, 1995) to understand geographical distribution and for resolving phylogenetic relationships. In this study, we evaluated aphid samples from India for polymorphisms and genetic diversity, with 106 randomly selected microsatellite primers. This is the first instance where microsatellite markers have been mined from a genomic resource and evaluated against Indian aphid species.

As in our study, the genetic population structure of *S. avenae* was revealed by microsatellites (De Barro et al., 1995). The aphid's geographical structure was relatively unusual, but microsatellite studies helped understand it (Hales et al., 1997). In aphids, microsatellites are currently the only co-dominant genetic marker tools sufficiently polymorphic to identify clones and clonal lineages (Wilson et al., 2003) and evolution by mutation in asexual lineages (Wilson et al., 1999, 2003). In this study, a robust set of 73 microsatellites was found to be transferable across the tribes of Aphidini and could be used to validate classical taxonomic classification, with minor exceptions which need to be studied further. In view of the high cost of developing genomic microsatellites, our study may aid in development of EST-microsatellite markers specific to Indian aphids belonging to the subfamily Aphidinae. Moreover, the new microsatellites developed will provide valuable information to researchers studying (Indian) aphid species, without significant molecular resources available.

In contrast to previous studies, our study using EST-microsatellite markers revealed a high PIC in species of Indian aphids. This set of markers can be used in future research to understand geographical distributions, biotype identification, and metapopulation structure in field collections, and mutational dynamics in aphids which are constantly co-evolving with resistant gene sources of Indian crop plants, in particular, and other aphid species, in general.

Acknowledgements

SV is highly grateful for the financial assistance received from the Indian Council of Agricultural Research (ICAR) as part of a HRD program under National Agricultural Innovative Project (NAIP) for carrying out this work in the USA. The authors are thankful to Dr. Leela Sahijram for critically going through and improving the article. This study was carried out as part of an Outreach Programme on Sucking Pests. Technical assistance was received from the staff of the Department of Biology, Gus R Douglas Institute, West Virginia State University, WV, USA, which is gratefully acknowledged.

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD & Sorrels ME (1993) Optimising parental selection for genetic linkages maps. Genome 36: 181–186.

Black WC, DuTeau NM & Puterka GJ (1992) Use of the random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) to detect DNA polymorphisms in aphids (Homoptera: Aphididae). Bulletin of Entomological Research 82: 151–159.

Caillaud MC, Mondor-Genson G, Levine-Wilkinson S, Miezetfrantz A, Simon JC & Coeur d'Acier A (2004) Microsatellite DNA markers for the pea aphid *Acyrthosiphon pisum*. Molecular Ecology Notes 4: 446–448.

Cenis JL, Perez P & Fereres A (1993) Identification of aphid (Homoptera: Aphididae) species and clones by random amplified polymorphic DNA. Entomological Society of America 86: 545–550.

Coeur d'Acier A, Jousseling E, Martin JF & Rasplus JY (2007) Phylogeny of the genus *Aphis* Linnaeus, 1758 (Homoptera: Aphididae) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 42: 598–611.

De Barro PJ, Sherratt TN, Brookes CP, David O & 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. Proceedings of the Royal Society of London B 262: 321–327.

Delmotte F, Leterme N, Gauthier JP, Rispe C & Simon JC (2002) Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. Molecular Ecology 11: 711–723.

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K et al. (2004) *Medicago truncatula* EST-SSRs reveal cross-species genetic markers for *Medicago* spp. Theoretical and Applied Genetics 108: 414–422.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Figueroa CC, Simon JC, Le Gallic JF & Niemeyer HM (1999) Molecular markers to differentiate two morphologically close species of the genus *Sitobion*. Entomologia Experimentalis et Applicata 92: 217–223.

Goldstein DB & Clark AG (1995) Microsatellite variation in North American populations of *Drosophila melanogaster*. Nucleic Acids Research 23: 3882–3886.

Grover A, Aishwarya V & Sharma PC (2007) Biased distribution of microsatellite motifs in the rice genome. Molecular Genetics and Genomics 277: 469–480.

Hales DF, Tomiuk J, Wöhrmann K & Sunnucks P (1997) Evolutionary and genetic aspects of aphid biology: a review. European Journal of Entomology 94: 1–55.

Hughes CR & Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. Molecular Ecology 2: 131–137.

Jun TH, Michel AP & Roufmian MA (2011) Development of soybean aphid genomic SSR markers using next generation sequencing. Genome 54: 360–367.

Llewellyn KS, Loxdale HS, Harrington R, Brookes Clark CPSJ & Sunnucks P (2003) Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. Molecular Ecology 12: 21–34.

Loxdale HD & Brookes CP (1989) Use of genetic markers (allozymes) to study the structure, overwintering and dynamics of pest aphid populations. Electrophoretic Studies on

Agricultural Pests (ed. by HD Loxdale & J den Hollander), pp. 231–270. Clarendon Press, Oxford, UK.

Loxdale HD, Hardie J, Halbert S, Foottit R, Kidd NAC & Carter CI (1993) The relative importance of short- and long-range movement of flying aphids. Biological Review 68: 291–311.

Martins RF, Zina V, Da Silva EB, Rebelo MT, Figueiredo E et al. (2012) Isolation and characterization of fifteen polymorphic microsatellite loci for the citrus mealybug, *Planococcus citri* (Hemiptera: Pseudococcidae), and cross-amplification in two other mealybug species. Journal of Genetics 91: 75–78.

Massonnet B, Simon JC & Weisse WW (2002) Metapopulation structure of the specialised herbivore *Macrosiphoniella tanacetaria* (Homoptera, Aphididae). Molecular Ecology 11: 2511–2521.

Meglécz E, Anderson SJ, Bourguet D, Butcher R, Caldas A et al. (2007) Microsatellite flanking region similarities among different loci within insect species. Insect Molecular Biology 16: 175–185.

Michel AP, Zhang W, Jung JK, Kang ST & Mian MAR (2009) Population genetic structure of *Aphis glycines*. Environmental Entomology 38: 1301–1311.

Minks AK & Harrewijn P (1988) Aphids: Their Biology, Natural Enemies, and Control. Elsevier, Amsterdam, The Netherlands.

Papura D, Simon JC, Halkett F, Delmotte F, Le Gallic JF & Dedryver CA (2003) Predominance of sexual reproduction in Romanian populations of the aphid *Sitobion avenae* inferred from phenotypic and genetic structure. Heredity 90: 397–404.

Plieske J & Struss D (2001) Microsatellite markers for genome analysis in *Brassica* I. development in *Brassica napus* and abundance in Brassicaceae species. Theoretical and Applied Genetics 102: 689–694.

Remaudiere G & Remaudiere M (1997) Catalogue des Aphididae du Monde. Institut National de la Recherche Agronomique, Paris, France.

Rohlf FJ (1997) NTSYSpc. Numerical Taxonomy and Multivariate Analysis System, v.2.2. Exeter Software, Setauket, NY, USA.

Rozen S & Skaletsky H (2000) Primer3 on the WWW for General Users and for Biologist Programmers. Methods in Molecular Biology 132: Bioinformatics Methods and Protocols (ed. by S Misener & SA Krawetz), pp. 365–386. Humana Press, Totowa, NJ, USA.

Sandstrom J & Moran N (1999) How nutritionally imbalanced is phloem sap for aphids? Entomologia Experimentalis et Applicata 91: 203–210.

Simon JC, Baumann S, Sunnucks P, Hebert PD, Pierre JS et al. (1999a) Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. Molecular Ecology 8: 531–545.

Simon JC, Leterme N & Latorre A (1999b) Molecular markers linked to breeding system differences in segregating and natural populations of the cereal aphid *Rhopalosiphum padi*. Molecular Ecology 8: 965–973.

Simon JC, Carré S, Boutin M, Prunier-Leterme N, Sabater-Munoz B & Latorre A (2003) Host-based divergence in populations of the pea aphid: insights from nuclear markers and the prevalence of facultative symbionts. Proceedings of the Royal Society of London B 270: 1703–1712.

Sloane MA, Sunnucks P, Wilson ACC & Hales DF (2001) Microsatellite isolation, linkage group identification and determination of recombination frequency in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). Genetic Research 77: 251–260.

Sourdille P, Tavaud M, Charmet G & Bernard M (2001) Transferability of wheat microsatellites to diploid Triticeae species carrying the A, B and D genomes. Theoretical and Applied Genetics 103: 346–352.

Subramanian S, Mishra RK & Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biology 4: R13.

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology and Evolution 15: 199–203.

Sunnucks P, England PR, Taylor AC & Hales DF (1996) Microsatellite and chromosome evolution of parthenogenetic *Sitobion* aphids in Australia. Genetics 144: 747–756.

Sunnucks P, De Barro PJ, Lushai G, Maclean N & Hales D (1997) Genetic structure of an aphid studied using microsatellites: cyclic parthenogenesis, differentiated lineages and host specialization. Molecular Ecology 6: 1059–1073.

Vanlerberghe-Masutti F, Chavigny P & Fuller SJ (1999) Characterization of microsatellite loci in the aphid species *Aphis gossypii* Glover. Molecular Ecology 8: 693–695.

Varshney RK, Graner A & Sorrells ME (2006) Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23: 45–48.

Weng Y, Azhaguvel P, Michels GJ & Rudd JC (2007) Cross-species transferability of microsatellite markers from six aphid (Hemiptera: Aphididae) species and their use for evaluating biotypic diversity in two cereal aphids. Insect Molecular Biology 16: 613–622.

Wilson ACC, Sunnucks P & Hales DFH (1997) A microsatellite linked to the X-chromosome of *Sitobion* aphids reveals random inheritance of the X-chromosome in males. Genetical Research 6: 233–236.

Wilson ACC, Sunnucks P & Hales DFH (1999) Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. Molecular Ecology 8: 1655–1666.

Wilson ACC, Sunnucks P, Blackman RL & Hales DFH (2002) Microsatellite variation in cyclically parthenogenetic populations of *Myzus persicae* in south-eastern Australia. Heredity 88: 258–266.

Wilson ACC, Sunnucks P & Hales DFH (2003) Heritable genetic variation and potential for adaptive evolution in asexual aphids (Aphidoidea). Biological Journal of the Linnean Society 79: 115–135.

Wilson ACC, Massonnet B, Simon JC, Prunier-Leterme N, Dolatti L et al. (2004) Cross-species amplification of microsatellite loci in aphids: assessment and application. Molecular Ecology Notes 4: 104–109.

Yu H & Li Q (2007) Development of EST-SSRs in the Mediterranean blue mussel, *Mytilus galloprovincialis*. Molecular Ecology Notes 7: 1308–1310.

Zhang LY, Bernard M, Leroy P, Feuillet C & Sourdille P (2005) High transferability of bread wheat EST-derived SSRs to other cereals. Theoretical and Applied Genetics 111: 677–687.

Zhao X & Kochert G (1993) Phylogenetic distribution and genetic mapping of a (GGC)n microsatellite from rice (*Oryza sativa* L.). Plant Molecular Biology 21: 607–614.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. List of EST microsatellites and their amplification pattern in aphids.