

## Evolution and homoplasy at the *Bem6* microsatellite locus in three sweetpotato whitefly (*Bemisia tabaci*) cryptic species

Dickey *et al.*

RESEARCH ARTICLE

Open Access

# Evolution and homoplasmy at the *Bem6* microsatellite locus in three sweetpotato whitefly (*Bemisia tabaci*) cryptic species

Aaron M Dickey<sup>1,3</sup>, Paula M Hall<sup>2</sup>, Robert G Shatters Jr<sup>1</sup> and Cindy L McKenzie<sup>1\*</sup>

## Abstract

**Background:** The evolution of individual microsatellite loci is often complex and homoplasmy is common but often goes undetected. Sequencing alleles at a microsatellite locus can provide a more complete picture of the common evolutionary mechanisms occurring at that locus and can reveal cases of homoplasmy. Within species homoplasmy can lead to an underestimate of differentiation among populations and among species homoplasmy can produce a misleading interpretation regarding shared alleles and hybridization. This is especially problematic with cryptic species.

**Results:** By sequencing alleles from three cryptic species of the sweetpotato whitefly (*Bemisia tabaci*), designated MEAM1, MED, and NW, the evolution of the putatively dinucleotide *Bem6* (CA)<sub>8</sub>imp microsatellite locus is inferred as one of primarily stepwise mutation occurring at four distinct heptaucleotide tandem repeats. In two of the species this pattern yields a compound tandem repeat. Homoplasmy was detected both among species and within species.

**Conclusions:** In the absence of sequencing, size homoplasious alleles at the *Bem6* locus lead to an overestimate of alleles shared and hybridization among cryptic species of *Bemisia tabaci*. Furthermore, the compound heptanucleotide motif structure of a putative dinucleotide microsatellite has implications for the nomenclature of heptanucleotide tandem repeats with step-wise evolution.

**Keywords:** Tandem repeat, Homoplasmy, Compound microsatellite, Stepwise mutation, *Bemisia tabaci*

## Background

Satellite DNA was originally described as the bands produced by genomic DNA in a CsCl buoyant density gradient that fell outside the principal band [1]. These bands were found to be common in eukaryotes, have a GC content different from the principal band (bulk single copy DNA), and consist of tandemly repeated sequences [2]. Tandem sequence repetition has since become synonymous with the term satellite DNA [3]. Tandemly repetitive sequences have two parameters, the sequence motif, and the copy number (e.g. ATTAT TATT contains three copies of the ATT trinucleotide motif). The diversity/variability within each parameter has led to the creation of multiple classification schemes [3-7] and synonyms [8-12] within tandem-

repeat nomenclature, particularly when copy number is less than 10<sup>3</sup>.

Tandem repeat marker loci contain the repeat region together with conserved flanking sequence and are usually non-coding. Evolution inside the repeat region is relatively fast due to a mutation rate that is 10<sup>3</sup>-10<sup>5</sup> times higher than that of the genome as a whole [12-14]. Evolution of the tandem repeats occurs primarily by DNA slippage during replication (microsatellites) [15,16] or by gene conversion and crossover during meiosis (minisatellites) [17] but see Richard and Paques [18]. In the majority of cases, slippage causes an allele to change size by one repeat motif at a time in a stepwise fashion while gene conversion can cause copy number changes in larger multiples [18]. Tandem repeats tend to mutate faster with increasing copy number [19,20], and tend to expand when copy number is low and contract when copy number is high [21-23]. This observation is consistent with upper size constraints on copy number [24].

\* Correspondence: Cindy.McKenzie@ars.usda.gov

<sup>1</sup>USDA-ARS, U.S. Horticultural Research Laboratory, 2001 South Rock Rd, Fort Pierce, FL 34945, USA

Full list of author information is available at the end of the article

Mutations can also occur in the area flanking a tandem repeat [25], or into the tandem repeat itself, thereby causing an imperfect or interrupted repeat [26]. These can cause size homoplasmy; cases where alleles among individuals have the same-size fragments (identical in character state), but arose in different lineages and are thus not identical by descent.

*Bemisia tabaci* Gennadius is a cryptic species complex composed of at least 24 morphologically identical species [27,28]. Most of these species are regionally endemic, but two are globally invasive agricultural pests, infesting and feeding on hundreds of crop plants in many diverse agroecosystems [29]. These species originated in the region bordering the Mediterranean Basin (MED) and in the Middle East/Asia Minor (MEAM1) region, and have also been referred to extensively in the literature as biotype Q and biotype B, respectively [28]. The recent arrival of MED to the United States in 2004 [30] raised concerns over possible hybridization with MEAM1, established in the United States since the mid 1980's [31]. This outcome at present seems an unlikely prospect given that MEAM1 and MED are almost completely reproductively isolated [32]. Laboratory hybrids can occur, but are both rare and sterile [33], while field hybrids are also rare and do not persist [34]. Monitoring the spread of MED within the U.S. and distinguishing molecularly among MED and MEAM1 has been a principal aim of the biotype Q taskforce [35]. During this effort, a partial sequence of the mtCO1 gene [36,37] has been the gold standard for molecular identification, but two microsatellite loci, *Bem6* and *Bem23* [38] have also been used as diagnostic markers [31,34]. Because of the importance of the *Bem6* nuclear locus in *B. tabaci* cryptic species determination in North America, its diversity among economically important species and the origin of shared alleles is of interest.

The *Bem6* microsatellite was described as an imperfect (CA)<sub>8</sub> tandem repeat where CA is the dinucleotide motif and 8 is the copy number [38] (Table 1). However, it has been noted that among U.S. samples, alleles at this locus occur in multiples of 7 base pairs rather than the expected 2 base pair multiples [31]. In addition, apparent hybrids and shared alleles appeared to be present among MED, MEAM1, and endemic New World (NW) whitefly samples in the dataset [31,34]. Because of this, alleles at

this locus were sequenced in order to 1) better characterize the tandem repeat nature of this marker, and 2) evaluate the possibility of hybrids and shared alleles at this locus among cryptic species.

## Methods

Between 2006 and 2011, 63 *Bem6* fragments representing 11 genotyped alleles were sequenced from 60 whiteflies (Table 2) as a supplementary part of large state and continent-wide surveys of *Bemisia tabaci* whiteflies [31,34]. The sequencing effort included samples from lab colonies as well as samples from Spain, Columbia, Israel, and Morocco. This represented between 1-2% of the whiteflies genotyped at this locus. Alleles not sequenced were very rare. Allele names are based on the average of estimated microsatellite fragment lengths calculated by Genemapper 4.0 (Applied Biosystems, Foster City, CA) rounded to the nearest base pair (Figure 1). These estimates are based on genotyping and form the basis for potential homoplasmy, as alleles are rarely sequenced in practice [39]. The actual allele size as determined via sequencing may be different. Individuals were identified to species using MEAM1, MED, and NW specific primers based on the mitochondrial CO1 gene [36] with reference to the species-level groups identified by Dinsdale et al. [28].

Alleles were PCR amplified using the unlabeled primers described by De Barro et al. [38]. Amplified alleles were direct sequenced in most cases but were also cloned using an Invitrogen TOPO TA<sup>®</sup> Cloning Kit (Life Technologies, Carlsbad, CA) when it was necessary to sequence both alleles from a heterozygote. Sequencing reactions were run on an Applied Biosystems 3730XL DNA analyzer using a BigDye<sup>®</sup> Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). To be certain both alleles were sequenced from heterozygotes, a minimum of 16 colonies were sampled. Sequences were aligned first using the large gap setting in Sequencher 4.7 (Genecodes, Ann Arbor, MI) and then manually in Sequencher 4.7 and in Mesquite 2.74 [40].

## Results

### Homoplasmy among whitefly species

In the three species, four perfect heptanucleotide tandem repeats designated H1-H4 (Figure 2a) were found

**Table 1 Nucleotide sequence of the *Bem6* (CA)<sub>8</sub> imperfect microsatellite and flanking region [38]**

Position	DNA Sequence																																	
1	T	T	A	C	A	C	T	T	A	A	C	A	C	C	A	G	A	A	C	T	T	T	T	A	A	T	C	A	T	A	C	A	A	
34	A	T	C	A	T	T	T	C	C	T	A	A	<b>C</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>T</b>	<b>T</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>C</b>	<b>A</b>	<b>C</b>
67	<b>A</b>	C	T	A	T	A	C	A	A	A	T	T	C	A	T	A	T	C	A	A	C	C	A	T	T	A	C	A	A	C	A	C	A	A
100	C	C	T	T	T	T	C	A	T	T	A	G	T	A	T	A	G	C	A	T	T	A	C	T	C	C	T	A	A	C	A	C	A	A
133	A	G	T	T	C	A	T	A	A	T	T	A	G	T	A	T	T	A	T	A	A	C	A	T	A	A	G	C	C	A	T	C	A	A

The repeat region is in bold and imperfections in the CA motif are italicized.

**Table 2 Alleles sequenced from *B. tabaci* cryptic species**

Estimated allele size	Species	Individuals sequenced	Allele frequency*
182	?	1	‡
189	NW	3	‡‡
195	NW	1	0.07
196	MED	3	<0.01
202	NW	2	0.51
203	MED	4	0.07
209	NW	2	0.16
210	MED	20	0.91
210	MEAM1	2	<0.01
216	NW	2	0.24
216	MEAM1	17	0.96
217	MED	3	0.01
223	MEAM1	3	0.04

63 microsatellite fragments sequenced from 60 individual whiteflies (some heterozygotes). Estimated allele size is rounded to the nearest whole number but see Figure 1.

\*Estimates of North American field/greenhouse collected *Bemisia tabaci* Bem6 allele frequencies by cryptic species taken from McKenzie et al. 2012.

‡ Singleton allele, species undetermined.

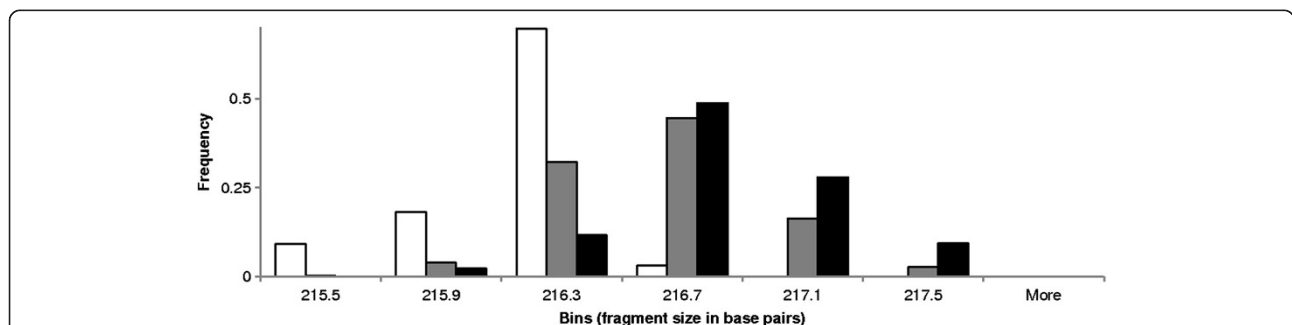
‡‡ Allele unique to lab colony.

preceding the 3' flanking region of the *Bem6* microsatellite described by DeBarro et al. [38]. From 63 fragments representing 11 genotyped alleles (based on estimated size), 16 unique alleles (based on sequencing) were found (Figure 3). The sequence of the MEAM1 allele 216 was identical in 17 individual whiteflies collected from five U.S. States and contained a compound heptanucleotide repeat of motifs H1-H3 (Figure 2). The sequence of the NW allele 216 was identical in two whiteflies, one each from Texas and Mexico and is characterized by the presence of the H4 motif (Figure 2). The sequence of the MED allele 217 was identical in two whiteflies from Florida and one from Spain and contained a compound heptanucleotide repeat of motifs H1 and H2 (Figure 2). Instances of potential homoplasy among MEAM1 allele 216, NW allele 216,

and MED allele 217 were caused by an overlapping distribution of allele size estimates (Figure 1). Additional cases of potential homoplasy ( $\leq 1$  base pair estimated length difference) among cryptic species were found at alleles 209/210 and 195/196 (Figure 3). In MED whiteflies, allele 210 is the most common in North America with a frequency of 91%. Only two MEAM1 individuals had this allele meaning there would be a very low incidence ( $\sim 1$  in 2500) of MEAM1 individuals incorrectly identified. As reported previously [34], there was no evidence of allele sharing or hybridization among cryptic species at the *Bem6* locus (Figure 3).

#### Homoplasy within whitefly species

Within species homoplasy was found for MED at allele 210 (Figure 3) and for MEAM1 at alleles 210 and 223 (Figure 3). In MED, two different 210 alleles were sequenced from 20 individuals from four countries and six U.S. States. The alleles differ from MED allele 217 by a single deletion or from MED allele 203 by a single insertion. This was the most frequent case of intraspecific homoplasy detected in the data. The allele designated "Q2" in Figure 3 was found in seven individuals collected in Guatemala and the U.S. States of Georgia and Oregon. The mitochondrial haplotypes of all seven individuals clustered with the haplotype Q2 described in McKenzie et al. [31] indicating maternal ancestry in the western Mediterranean region [41-43] (data not shown). The second allele 210 in MED was found in 13 individuals collected from Israel, Morocco, and the U.S. states of California, Pennsylvania, Florida, and Michigan. Of these 13, 11 had mitochondrial haplotypes that clustered with the haplotype Q1 described in McKenzie et al. [31] indicating maternal ancestry in the eastern Mediterranean region [41-43]. The other two individuals, a whitefly from Morocco and a whitefly from a California lab colony founded from individuals collected in Spain, had mitochondrial haplotypes consistent with western Mediterranean maternal ancestry (data not shown).



**Figure 1** Frequency histogram of allele size estimates calculated by Genemapper 4.0 based on the Rox400 standard curve for the alleles designated 216 NW (white, n = 33 fragments, median = 216.1 base pairs), 216 MEAM1 (gray, n = 4304 fragments, median = 216.4 base pairs), and 217 MED (black, n = 43 fragments, median = 216.6 base pairs). Putatively shared alleles are present among the three cryptic species.





mitochondrial barcode, but the sequence at the *Bem6* locus is most similar to other NW alleles. *Bem6* alleles in MEAM1 and MED contain compound tandem repeats of H1-H3 while NW features only H4 in tandem (Figure 3).

Within each species, the H3 motif appears to be the most stable, never showing copy number variation in tandem. This motif, AACACAC, is the most similar to the (CA<sub>8</sub>)<sub>imp</sub> originally described by De Barro et al. [38]. Singleton copies of this motif are also found in both flanking regions of the originally described microsatellite (residues 45-51 and 94-100 in Table 1) and these persist in all three species studied here. The flanking region itself also appears relatively stable though a few species specific polymorphisms are present (Table 3).

### Discussion

Generally the design of new microsatellites for each species separately is recommended [39], but this is not always possible when cryptic species are present or species level taxonomy is in flux. It should be noted that the *Bem6* microsatellite was originally isolated from the Australia species (P. De Barro, personal communication) see Dinsdale et al. [28] so it would be interesting to see if a similar pattern of heptanucleotide tandem repeat evolution is also present in this cryptic species. Determining the ancestral state of the *Bem6* locus within each cryptic species is not possible due to low allele number and paucity of home range sampling. Columbian NW are apparently ancestral to North and Central American NW [44] so allele 189 may be the ancestral state for NW. This should be investigated further with increased sampling throughout the home range of NW. It is not possible to speculate about the ancestral state of the *Bem6* locus within MED and MEAM1 due to the small number of alleles sampled in the native range. Further exploration of MED diversity in particular, including its apparently ancestral sub Saharan sub group [42], could help resolve this.

### Heptanucleotide repeats

Heptanucleotide tandem repeats have not received a lot of attention and this could be due, in part, to their exclusion from some classical definitions of both microsatellites and minisatellites e.g. [3]. While many accept some general

rules for classifying tandemly repetitive DNA sequences [16,39], several authors have argued that some of these rules may be arbitrary [7,46]. Heptanucleotide repeats have been shown to be more common than tetra- and penta-nucleotide repeats in several plant taxa [47] suggesting that defining microsatellites as motifs between 1 and 6 nucleotides in length may also be unnecessarily exclusive. The data presented here is consistent with step-wise evolution, indicating that the four heptanucleotide motifs probably evolve like microsatellites [48] rather than minisatellites [17]. In addition, the alleles described may also have arisen via length independent slippage [49,50] expected when copy number is low.

### Compound microsatellites

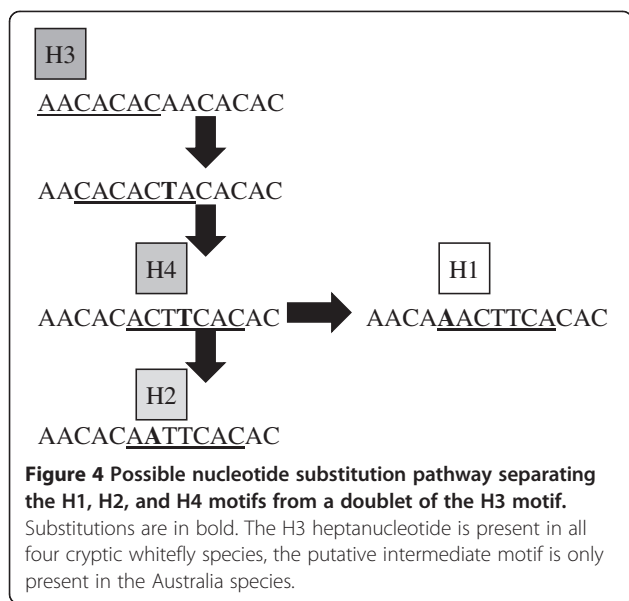
For the species in this study, a microsatellite originally described as an imperfect dinucleotide repeat is characterized as a series of perfect heptanucleotide repeats. This change seems to be a function of the particular species in question as the motifs described here never occur in tandem in the original microsatellite (Figure 2). ‘Imperfections’ or ‘interruptions’ are commonly invoked in the tandem repeat literature [16,51,52] and are characterized this way because they break up a long perfect microsatellite into two shorter microsatellites [53]. But an interruption can also lead to the creation of a new ‘proto’ microsatellite [54,55]. If the newly created sequence motif is duplicated adjacent to the original repeat, it will form a compound microsatellite. This may be a common occurrence as compound microsatellites are 15 times more abundant in genomic DNA than random microsatellite distribution expectations [55].

To our knowledge, this is the first report of a compound heptanucleotide tandem repeat. While the pattern both within and among species is consistent with step-wise evolution of tandem repeats, choosing among alternative alignments was difficult and probably exacerbated by the absence of guanine in the repeat area. In the end, the alignment selected was the only one found where within-species size-sequential alleles were only separated by one mutation step. The next best alignment also featured four different heptanucleotide motifs, but had two additional point mutations separating NW alleles 189 and 195. One possible mutation pathway among the four identified motifs is given in Figure 4 with either one or two point mutations separating each motif. Kofler et al. [55] found that the motifs comprising most compound microsatellite pairs differed by a single mutation and suggested that such mutations represent the dominant mechanism underlying the origin of compound microsatellites. The evolutionary fate of a given microsatellite ‘interruption’ will probably depend on the structural properties of the new motif [56], but may ultimately yield a compound microsatellite.

**Table 3 Species specific polymorphisms in flanking regions of the *Bem6* microsatellite**

	Position				
Species	22	23	43	69	129
MEAM1 + MED	A	A	C	A	T
Australia	T	T	.	T	C
NW	T	.	A	.	T

Dots denote identity with MEAM1 and MED alleles.



In the evolution via nucleotide substitution scenario shown in Figure 4, the H3 motif is hypothesized to be the ancestral motif since it is present in all three cryptic species studied here as well as the reference Australia species. However, the H1 motif is common to MEAM1, MED, and NW suggesting it either arose independently in each species or was found in the ancestor of the three species and gave rise to the H4 motif in NW and the H2 motif in the common ancestor of MEAM1 and MED. Untangling the phylogenetic history and demonstrating the phylogenetic utility of this compound microsatellite will require sequencing alleles from additional members of the cryptic species complex.

#### Homoplasmy and cryptic species

Mutations occurring in either the flanking region or the repeat region can lead to detectable size homoplasmy in tandem repeat markers [57,58]. In addition, back mutations can lead to undetectable size homoplasmy e.g. a 210 allele can arise either from a 203 allele or a 217 allele and distinguishing among these alternatives is impossible without parentage. The false signal given by apparently shared, but size homoplasious alleles in microsatellites is not generally considered a problem when many loci with high variability are used in population genetics [59], but see Balloux et al. [60]. Homoplasmy can overestimate the frequency of shared alleles among species [51] and this has the potential to be misleading when those species are cryptic [61]. In this study, homoplasious alleles would have led to an overestimate of alleles shared and hybridization among cryptic species if not sequenced and if not used in combination with mitochondrial DNA sequences. After sequencing, no evidence of interspecific hybridization was found at this locus.

Relative to the number of mitochondrial haplotypes found, very few *B. tabaci* haplotypes have spread globally [42]. Low genetic diversity is also apparent at the *Bem6* nuclear locus in North American samples of invasive MEAM1 and MED. This locus initially appeared diagnostic for the cryptic species MEAM1 and MED. Even after genotyping several thousand samples, 99% of MEAM1 whiteflies had the 216 and/or 223 alleles and 98% of MED samples had the 203 and/or 210 alleles (Table 2). But rare alleles and putative hybrids surfaced as sample size increased and, without sequencing, would have led to misinterpretation of allele sharing among both native and invasive cryptic species and an overestimation of hybrid frequency. These pitfalls might be even greater in the native range of MEAM1 and MED due to higher genetic diversity. Researchers should therefore use caution when using microsatellites for diagnostic purposes.

#### Conclusions

Sequencing 11 alleles at the putatively dinucleotide *Bem6* locus from three members of a cryptic species complex revealed four different heptanucleotide tandem repeat motifs. The sequencing data is consistent with step-wise evolution, suggesting the locus evolves like a microsatellite rather than a minisatellite. In addition, the alleles described may also have arisen via length independent slippage expected when copy number is low. In two of the species, a compound heptanucleotide repeat is formed and, to our knowledge, this is the first such report.

Homoplasious alleles at the *Bem6* locus would have led to an overestimate of alleles shared and hybridization among cryptic species if not sequenced. After sequencing, no evidence of interspecific hybridization remained. These results highlight the need for caution when using microsatellites for cryptic species discrimination and diagnostics.

#### Competing interest

Authors declare that they have no competing interests and are responsible for the content of this paper.

#### Authors' contributions

AMD conceived and conducted experiments, analyzed data, and drafted the manuscript. PMH conceived and conducted experiments and analyzed data. RGS conceived experiments, analyzed data, and edited the manuscript. CLM initiated and conceived experiments. All authors read and approved the manuscript.

#### Acknowledgements

The study was funded by Lance S Osborne and CLM with funding from the nursery and floriculture initiative. We thank Paul De Barro, Greg O'Corry-Crowe, and one anonymous reviewer for their suggestions on an earlier draft of this manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.

#### Author details

<sup>1</sup>USDA-ARS, U.S. Horticultural Research Laboratory, 2001 South Rock Rd, Fort Pierce, FL 34945, USA. <sup>2</sup>Mid-Florida Research & Education Center, University of Florida, 2725 Binion Rd, Apopka, FL 32703, USA. <sup>3</sup>Current address: Mid-Florida Research & Education Center, University of Florida, 2725 Binion Rd, Apopka, FL 32703, USA.

Received: 21 March 2013 Accepted: 26 June 2013

Published: 2 July 2013

#### References

- Sager R, Ishida MR: Chloroplast DNA in *Chlamydomonas*. *Proc Natl Acad Sci USA* 1963, **50**(4):725–730.
- Britten RJ, Kohne DE: Repeated sequences in DNA. *Science* 1968, **161**:529–540.
- Tautz D: Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In *DNA Fingerprinting: State of the Science*, Volume 67. Edited by Chakraborty R, Epplen JT, Jeffreys AJ. Basel, Switzerland: Birkhauser Verlag; 1993:21–28.
- Dean C, Schmidt R: Plant genomes: a current molecular description. *Annu Rev Plant Biol* 1995, **46**(1):395–418.
- Navajas-Perez R, Paterson AH: Patterns of tandem repetition in plant whole genome assemblies. *Mol Genet Genomics* 2009, **281**(6):579–590.
- Tamaki K: Minisatellite and microsatellite DNA typing analysis. In *Molecular forensics*. Edited by Rapley R, Whitehouse D. Chichester, U.K: John Wiley & Sons; 2007:71–89.
- Fondon JW III, Hammock EAD, Hannan AJ, King DG: Simple sequence repeats: genetic modulators of brain function and behavior. *Trends Neurosci* 2008, **31**(7):328–334.
- Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Krumlin E, et al: Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 1987, **235**(4796):1616–1622.
- Tautz D, Renz M: Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res* 1984, **12**(10):4127–4138.
- Litt M, Luty JA: A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* 1989, **44**(3):397–401.
- Jeffreys AJ, Wilson V, Thein SL: Hypervariable 'minisatellite' regions in human DNA. *Nature* 1985, **314**:67–73.
- Weber JL, Wong C: Mutation of human short tandem repeats. *Hum Mol Genet* 1993, **2**(8):1123–1128.
- Schug MD, Mackay TFC, Aquadro CF: Low mutation rates of microsatellite loci in *Drosophila melanogaster*. *Nat Genet* 1997, **15**(1):99–102.
- Baer CF, Miyamoto MM, Denver DR: Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat Rev Genet* 2007, **8**:619–631.
- Levinson G, Gutman GA: Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* 1987, **4**(3):203–221.
- Ellegren H: Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 2004, **5**(6):435–445.
- Jeffreys AJ, Neil DL, Neumann R: Repeat instability at human minisatellites arising from meiotic recombination. *EMBO J* 1998, **17**(14):4147–4157.
- Richard GF, Paques F: Mini- and microsatellite expansions: the recombination connection. *EMBO Rep* 2000, **1**(2):122–126.
- Brinkmann B, Klintschar M, Neuhuber F, Hohne J, Rolf B: Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet* 1998, **62**(6):1408–1415.
- Kelkar YD, Tyekucheva S, Chiaromonte F, Makova KD: The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Res* 2008, **18**(1):30–38.
- Primmer CR, Saino N, Muller AP, Ellegren H, others: Directional evolution in germline microsatellite mutations. *Nat Genet* 1996, **13**(4):391.
- Buschiazio E, Gemmill NJ: The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays* 2006, **28**(10):1040–1050.
- Huang QY, Xu FH, Shen H, Deng HY, Liu YJ, Liu YZ, Li JL, Recker RR, Deng HW: Mutation patterns at dinucleotide microsatellite loci in humans. *Am J Hum Genet* 2002, **70**(3):625–634.
- Nauta MJ, Weissing FJ: Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics* 1996, **143**(2):1021–1032.
- Grimaldi MC, Crouau-Roy B: Microsatellite allelic homoplasy due to variable flanking sequences. *J Mol Evol* 1997, **44**(3):336–340.
- Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A: Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Mol Biol Evol* 1998, **15**(10):1275–1287.
- De Barro PJ, Liu SS, Boykin LM, Dinsdale AB: *Bemisia tabaci*: a statement of species status. *Annu Rev Entomol* 2011, **56**:1–19.
- Dinsdale A, Cook L, Riginos C, Buckley YM, De Barro P, Dinsdale A: Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodidae: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. *Ann Entomol Soc Am* 2010, **103**:196–208.
- Perring TM: The *Bemisia tabaci* species complex. *Crop Prot* 2001, **20**:725–737.
- Dennehy TJ, Decain BA, Harpold VS, Zaborac M, Morin S, Fabrick JA, Nichols RL, Brown JK, Byrne FJ, Xianchun LI: Extraordinary resistance to insecticides reveals exotic Q biotype of *Bemisia tabaci* in the New World. *J Econ Entomol* 2010, **103**:2174–2185.
- McKenzie CL, Hodges G, Osborne LS, Byrne FJ, Shatters RG Jr: Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in Florida—investigating the Q invasion. *J Econ Entomol* 2009, **102**(2):670–676.
- Elbaz M, Lahav N, Morin S, Elbaz M: Evidence for pre-zygotic reproductive barrier between the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Bull Entomol Res* 2010, **100**:581–590.
- Sun DB, Xu J, Luan JB, Liu SS: Reproductive incompatibility between the B and Q biotypes of the whitefly *Bemisia tabaci* in China: genetic and behavioral evidence. *Bull Entomol Res* 2011, **101**(2):211–220.
- McKenzie CL, Bethke J, Byrne FJ, Chamberlin J, Dennehy TJ, Dickey AM, Gilrein D, Hall P, Luswig S, Oetting R, et al: Distribution of *Bemisia tabaci* (Hemiptera: Aleyrodidae) biotypes in North America after the Q invasion. *J Econ Entomol* 2012, **105**(3):753–766.
- Dalton R: Whitefly infestations: the Christmas invasion. *Nature* 2006, **443**:898–900.
- Shatters RG Jr, Powell CA, Boykin LM, Liansheng H, McKenzie CL: Improved DNA barcoding method for *Bemisia tabaci* and related Aleyrodidae: development of universal and *Bemisia tabaci* biotype-specific mitochondrial cytochrome c oxidase I polymerase chain reaction primers. *J Econ Entomol* 2009, **102**(2):750–758.
- Boykin LM, Shatters RG Jr, Rosell RC, McKenzie CL, Bagnall R, De Barro P, Frohlich DR: Global relationships of *Bemisia tabaci* (Hemiptera: Aleyrodidae) revealed using bayesian analysis of mitochondrial COI DNA sequences. *Mol Phylogenet Evol* 2007, **44**:1306–1319.
- De Barro PJ, Scott KD, Graham GC, Lange CL, Schutze MK: Isolation and characterization of microsatellite loci in *Bemisia tabaci*. *Mol Ecol Notes* 2003, **3**:40–43.
- Selkoe KA, Toonen RJ: Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett* 2006, **9**(5):615–629.
- Maddison WP, Maddison DR: *Mesquite: a modular system for evolutionary analysis*; 2001. In. 2.74 <http://mesquiteproject.org> edn.
- Tsagarakou A, Tsigenopoulos CS, Gorman K, Lagnel J, Bedford ID: Biotype status and genetic polymorphism of the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Greece: mitochondrial DNA and microsatellites. *Bull Entomol Res* 2007, **97**:29–40.
- De Barro P, Ahmed MZ: Genetic networking of the *Bemisia tabaci* cryptic species complex reveals pattern of biological invasions. *PLoS ONE* 2011, **6**(10):e25579.
- Dickey AM, Osborne LS, Shatters RG Jr, Hall PM, McKenzie CL: Population genetics of invasive *Bemisia tabaci* (Hemiptera: Aleyrodidae) cryptic species in the United States based on microsatellite markers. *J Econ Entomol* 2013, **106**(3):1355–1364.
- Alemandri V, De Barro P, Bajerman N, Arguello Caro EB, Dumon AD, Mattio MF, Rodriguez SM, Truol G: Species within the *Bemisia tabaci* (Hemiptera: Aleyrodidae) complex in soybean and bean crops in Argentina. *J Econ Entomol* 2012, **105**(1):48–53.
- Boykin LM, Armstrong KF, Kubatko L, De Barro P: Species delimitation and global biosecurity. *Evol Bioinform* 2012, **8**:1–37. 2012.
- Hannan AJ: Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for missing heritability. *Trends Genet* 2010, **26**(2):59–65.
- Mun JH, Kim DJ, Choi HK, Gish J, Debelle F, Mudge J, Denny R, Endre G, Saurat O, Duzde AM, et al: Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. *Genetics* 2006, **172**(4):2541–2555.



48. Valdes AM, Slatkin M, Freimer NB: **Allele frequencies at microsatellite loci: the stepwise mutation model revisited.** *Genetics* 1993, **133**(3):737–749.
49. Dieringer D, Schlotterer C: **Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species.** *Genome Res* 2003, **13**(10):2242–2251.
50. Zhu Y, Strassmann JE, Queller DC: **Insertions, substitutions, and the origin of microsatellites.** *Genet Res* 2000, **76**(3):227–236.
51. Angers B, Bernatchez L: **Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information.** *Mol Biol Evol* 1997, **14**(3):230–238.
52. Estoup A, Tailliez C, Cornuet JM, Solignac M: **Size homoplasmy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae).** *Mol Biol Evol* 1995, **12**(6):1074–1084.
53. Chistiakov DA, Hellems B, Volckaert FAM: **Microsatellites and their genomic distribution, evolution, function, and applications: a review with special reference to fish genetics.** *Aquaculture* 2006, **255**:1–29.
54. Dettman JR, Taylor JW: **Mutation and evolution of microsatellite loci in neurospora.** *Genetics* 2004, **168**:1231–1248.
55. Kofler R, Schlotterer C, Luschützky E, Lelley T: **Survey of microsatellite clustering in eight fully sequenced species sheds light on the origin of compound microsatellites.** *BMC Genomics* 2008, **9**:612.
56. Bacolla A, Larson JE, Collins JR, Li J, Milosavljevic A, Stenson PD, Cooper DN, Wells RD: **Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties.** *Genome Res* 2008, **18**(10):1545–1553.
57. van Oppen MJH, Rico C, Turner GF, Hewitt GM: **Extensive homoplasmy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids.** *Mol Biol Evol* 2000, **17**(4):489–498.
58. Curtu AL, Finkeldey R, Gailing O: **Comparative sequencing of a microsatellite locus reveals size homoplasmy within and between European oak species (*Quercus* spp.).** *Plant Mol Biol Rep* 2004, **22**(4):339–346.
59. Estoup A, Jarne P, Cornuet JM: **Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis.** *Mol Ecol* 2002, **11**(9):1591–1604.
60. Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J: **Microsatellites can be misleading: an empirical and simulation study.** *Evolution* 2000, **54**(4):1414–1422.
61. Skrede I, Borgen L, Brochmann C: **Genetic structuring in three closely related circumpolar plant species: ALFP versus microsatellite markers and high-arctic versus arctic-alpine distributions.** *Heredity* 2009, **102**:293–302.

doi:10.1186/1756-0500-6-249

**Cite this article as:** Dickey *et al.*: Evolution and homoplasmy at the *Bem6* microsatellite locus in three sweetpotato whitefly (*Bemisia tabaci*) cryptic species. *BMC Research Notes* 2013 **6**:249.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

