Insect Biochemistry and Molecular Biology 113 (2019) 103228

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Analysis of the extent of synteny and conservation in the gene order in aphids: A first glimpse from the *Aphis glycines* genome

Mauro Mandrioli[a,*], Giulia Melchiori[a], Michela Panini[b], Olga Chiesa[b], Rosanna Giordano[c,d], Emanuele Mazzoni[b], Gian Carlo Manicardi[a]

[a] Dipartimento di Scienze della Vita, Università di Modena e Reggio Emilia, Via Campi 213/D, 41125, Modena, Italy
[b] Dipartimento di Scienze delle produzioni vegetali sostenibili, Università Cattolica del Sacro Cuore, Piacenza, Italy
[c] Puerto Rico Science, Technology & Research Trust, San Juan, PR, USA
[d] Know Your Bee, Inc., San Juan, PR, USA

ARTICLE INFO

Keywords:
Gene mapping
Synteny
Chromosomal rearrangements
Holocentric chromosomes
Aphids

ABSTRACT

In the last decade several insect genomes have been sequenced, but for most the chromosomal mapping of the identified scaffolds/annotated genes is not available. The lack of this information makes it difficult to analyse various genetic aspects, including the presence of genome rearrangements and the extent of synteny within and across species. We mapped five multigenic DNA families (major and minor rDNAs, histone gene cluster, esterases and carotenoid desaturases) and seven scaffolds corresponding to 9 Mb of the soybean aphid, *Aphis glycines,* genome and identified *loci* spanning the four soybean aphid chromosomes. A comparative analysis of the localization of the annotated *A. glycines* genes with respect to the peach potato aphid, *Myzus persicae,* and the fly, *Drosophila melanogaster,* evidenced a lower degree of synteny between the two aphid species than in the aphid-fly comparison. Only 1.4 genes per syntenic block were observed in aphids in contrast to 2.3 genes per block in flies. This higher chromosomal rearrangement rate in aphids could be explained considering that they possess holocentric chromosomes that can favour the stabilization and inheritance of chromosomal rearrangements. Lastly, our experiments did not detect the presence of chimeric assemblies in the newly available *A. glycines* biotype 1 genome, differently from what reported in assembled genome of other aphid species, suggesting that chromosomal mapping can be used to ascertain the quality of assembled genomes.

1. Introduction

In the last decade, insect genomics has had highly exciting developments. From a first phase where most of the sequencing effort was focused on the genus *Drosophila*, other insect genomes have followed, such as that of the honey bee *Apis mellifera* (Honeybee Genome Sequencing Consortium, 2006), the red flour beetle *Tribolium castaneum* (Tribolium Genome Sequencing Consortium, 2008), the yellow fever mosquito *Aedes aegypti* (Dudchenko et al., 2017), the mosquito *Culex quinquefasciatus* (Arensburger et al., 2010) and the blood-sucking bug *Rhodnius prolixus* (Mesquita et al., 2015).

The currently available insect genomes belong to both major evolutionary lineages of insects (holometabola and hemimetabola) and to six different insect orders (i.e. Honeybee Genome Sequencing Consortium, 2006; Grimmelikhuijzen et al., 2007; Tribolium Genome Sequencing Consortium, 2008; Arensburger et al., 2010; Brucker et al., 2012; Chilana et al., 2012; Mesquita et al., 2015). These achievements have greatly increased our knowledge of the biology of insects and provided a better understanding of applied aspects, such as insecticide resistance that poses an increasing problem for pest control (Grimmelikhuijzen et al., 2007). For example, it has been possible to identify new insecticide targets in addition to the few previously known, such as acetylcholinesterase, sodium channels and GABA receptors (Grimmelikhuijzen et al., 2007). Furthermore, it has been possible to gain a better understanding of the olfactory signalling cascades, responsible for host-seeking behaviour (plant/insect interactions), allowing for the development of more efficient trapping/capture strategies and tools (Grimmelikhuijzen et al., 2007).

Investigations in the above applied research fields is particularly needed in aphids, since this *taxon* includes significant plant pests, which, by virtue of being phloem feeders, not only impart direct damage to crops, but also serve as vector of debilitating plant diseases (Van Emden and Harrington, 2007). Genomic resources in the Aphididae are, therefore, critically needed to advance both basic and applied

* Corresponding author. Department of Life Sciences, University of Modena and Reggio Emilia, Via Campi 213/D, 41125 Modena, Italy.
E-mail address: mauro.mandrioli@unimore.it (M. Mandrioli).

https://doi.org/10.1016/j.ibmb.2019.103228
Received 15 January 2019; Received in revised form 3 July 2019; Accepted 21 August 2019
Available online 22 August 2019
0965-1748/

Table 1
List of primers for the amplification of FISH probes on scaffolds.

SBAphidCtg1000	1000F1	5′- TTTTATCCAAAAATCACTTGTCCAT
	1000R1	5′- CTATTCGTTGTTTGCAATTACACTG
	1000F2	5′- TATTTATACGGGTTATGCGGAGATA
	1000R2	5′- GCTTAGCTAAAGATTTATGCGACTG
SBAphidCtg1002	1002F1	5′- ATTGTACAGCTGTCTAGGTGGTTTC
	1002R1	5′- AAACTGAGCACTAGGAACTTGAAGA
	1002F2	5′- TATGCGGGTAGTGACTTAGTGTGTA
	1002R2	5′- TTAGTGGGGTAGTCAGTCGTTTTAG
SBAphidCtg1004	1004F1	5′- GTAGCGATTGTGTAAGTGCAATATG
	1004R1	5′- TTTTTGGGTCACCAATTTTAATAGA
	1004F2	5′- TCTAGATAAGCGGTTTCATGATTTC
	1004R2	5′- CAATAGCTTAGCCTTGACGAAGTAG
SBAphidCtg1016	1016F1	5′- ATTGTCGCGCGTTACATACTTAGCTG
	1016R1	5′- ATAGCCCTAGCTCTAGAGCTCTGAGA
	1016F2	5′- TTTGGCCTATAAATCTCGCTCTAGG
	1016R2	5′- CTCCGATATTTAACTCTATATCTCT
SBAphidCtg1039	1039F1	5′- GTACCACCGTCTATCTGTCTCAAGT
	1039R1	5′- ACGATTTTTAACGATTTATGGATGA
	1039F2	5′- GTCGAAATGAGGAAATAGAGTTGAA
	1039R2	5′- ATCGTACTCTTTTCACAAGAACCAC
SBAphidCtg6374	6374F1	5′- GTAGCATCTCCTCCAGATGAAATTA
	6374R1	5′- AATTGTAGTTTCGTCGCTTATTGAC
	6374F2	5′- GTACCTACACGTTTGAGGAGAAAAA
	6374R2	5′- ACTATACGTGCGATAATGCCTCTAC
SBAphidCtg6376	6376F1	5′- CTAGCAACTGCTCCAGATGTTTTAT
	6376R1	5′- TTAAGTAGTTTCCTCGCTTTAAGTG
	6376F2	5′- CTACCTACACGAAAGACCAGTTTTT
	6376R2	5′- TTTAGACGTGCACTAATGCGAACTA

research (Mandrioli et al., 2017).

Current resources in aphid genomics include the genome assemblies for the following aphid species: the pea aphid, *Acyrthosiphon pisum* (International Aphid Genomics Consortium, 2010), the Russian wheat aphid, *Diuraphis noxia* (Nicholson et al., 2015), a draft of the genome of the soybean aphid, *Aphis glycines* (from a composite of multiple North American populations) (Wenger et al., 2017), the cotton aphid *Aphis gossypii* (Quan et al., 2019), as well as four additional species for which the complete genome papers are in different stages of publications, but whose genome sequences are available on AphidBase: *Aphis glycines* biotype 1 (Giordano et al., in preparation), the black cherry aphid, *Myzus cerasi*; the peach potato aphid, *Myzus persicae;* and the bird cherry-oat aphid, *Rhopalosiphum padi.*

Despite the availability of sequenced genomes, few studies have focused on chromosomal mapping of annotated genes in aphids and these studies are only specific to the pea aphid (Brisson and Davis, 2008; Manicardi et al., 2015a, b; Mandrioli et al., 2017). The first pea aphid genetic map was conducted by Hawthorne and Via (2001) with the aim to study the pea aphid host plant specialization. In particular, they developed a linkage map of 173 dominant amplified fragment length polymorphism (AFLP) markers grouped into four linkage groups. This set of markers was successively enriched with seven additional AFLP markers on the X chromosome (Braendle et al., 2005).

Using a different approach, Mandrioli et al. (2017) assessed the localization of eight *A. pisum* scaffolds (covering 5 Mb and 83 genes) identifying *loci* spanning all the four *A. pisum* chromosomes and in this manner favoured the identification of chromosomal markers useful for the study of chromosomal rearrangements. The availability of mapping data in aphids could be beneficial in several ways and in particular to: i) elucidate the evolution of aphid karyotypes; ii) provide genetic evidence for the origin of virulent biotypes that have evolved in response to plants with various resistance levels and genes (Cooper et al., 2015).

In the present work we assessed the localization of seven *A. glycines* scaffolds (spanning 9 Mb and 53 genes), as well as the localization of five multigenic DNA families (major and minor rDNAs, histone gene cluster, esterases and carotenoid desaturases). This approach allowed the identification of *loci* spanning all four *A. glycines* chromosomes. In order to evaluate the impact of holocentric chromosomes on the occurrence of genome rearrangements in aphids in comparison to the typical monocentric chromosomes, we performed an analysis of the synteny between *A. glycines* and the peach potato aphid, *M. persicae,* as well an analysis of the former with the fruit fly, *D. melanogaster.*

2. Materials and methods

The specimens of soybean aphid *A. glycines* used in the present research were obtained from a laboratory strain belonging to the *A. glycines* Biotype 1 collected in Urbana, Illinois (USA) in 2001. Aphids were maintained asexually on weekly-changed soybean leaves at 20 °C at a light-dark regime of 16 h light: 8 h darkness in the insectary of the Università Cattolica del Sacro Cuore in Piacenza (Italy), according to the letter of authority issued under Directive 2008/61/EC by the Italian Central Phytosanitary Service (prot. n. 0020273, dated August 03, 2016) allowing the rearing of the aphids in controlled conditions, in order to avoid a possible spread of the pest in Italy. Live aphids used in this research were never managed outside the dedicated insectary and at the end of the experiments the colony was eliminated by deep freezing and subsequent incineration.

Chromosome preparations were obtained within the insectary premises, from live parthenogenetic females, by the spreading of embryo cells as reported by Monti et al. (2011), whereas chromomycin A_3 (CMA_3) staining was done as described by Mandrioli et al. (2017).

DNA extraction was performed using the Wizard® SV Genomic DNA Purification System (Promega), according to the manufacturer's instructions using *A. glycines* adults collected in the insectary, immediately preserved in 95% ethanol and stored at −20 °C until DNA extraction. DNA samples were quantified by spectrophotometric absorbance measurements at 260 nm using a Nanodrop ND 1000 Spectrophotometer (NanoDrop Technologies). The Long PCR Enzyme Mix (Fermentas), combined to a digoxigenin (DIG)- and biotin-labelling of the probe with the PCR DIG labelling kit (Roche), was used to amplify and label two contiguous 20 Kbp long probes for each scaffold. The use of two probes for each scaffold was preferred as it allows longer labelled chromosomal portions and consequently stronger fluorescent signals on chromosomes, as already assessed in Mandrioli et al. (2017). Oligonucleotide primers (Table 1) were specifically designed on the *A. glycines* Biotype 1 scaffold sequences (available in Aphidbase) using the on-line software Primer 3 (http://bioinfo.ut.ee/primer3/).

In order to amplify the complete cluster of genes coding for histones, the primers HIS-CLUST-F (5′-CGAAACCGTAAAGGGTACGA) and HIS-CLUST-R (5′-GGCGGCTTTGACTTT ATTGA) were designed on the basis of the *A. pisum* unplaced genomic scaffold 368 (NW_003383857.1, from base 259987 to 272662). Primers were used to amplify a 7379 bp fragment with a Hybaid thermal-cycler using the Long PCR Enzyme Mix (Fermentas) with a 68 °C annealing temperature and extension for 8 min (25 cycles), according to the manufacturer's instructions.

The 5S rDNA repeat unit of *A. pisum* was amplified by PCR using the primers, 5S–F (5′-TGCACGTAGTGTTCCCAAGC) and 5S-R (5′-ACGACC ATACCACGTTGAATAC), designed according to the 5S sequences of aphid *Acyrthosiphon magnoliae* available in GenBank (ID: X01518). The above two primers were designed to prevent primer cross-hybridization with other pol III-controlled genes (Geiduschek and Tochini-Valentini,

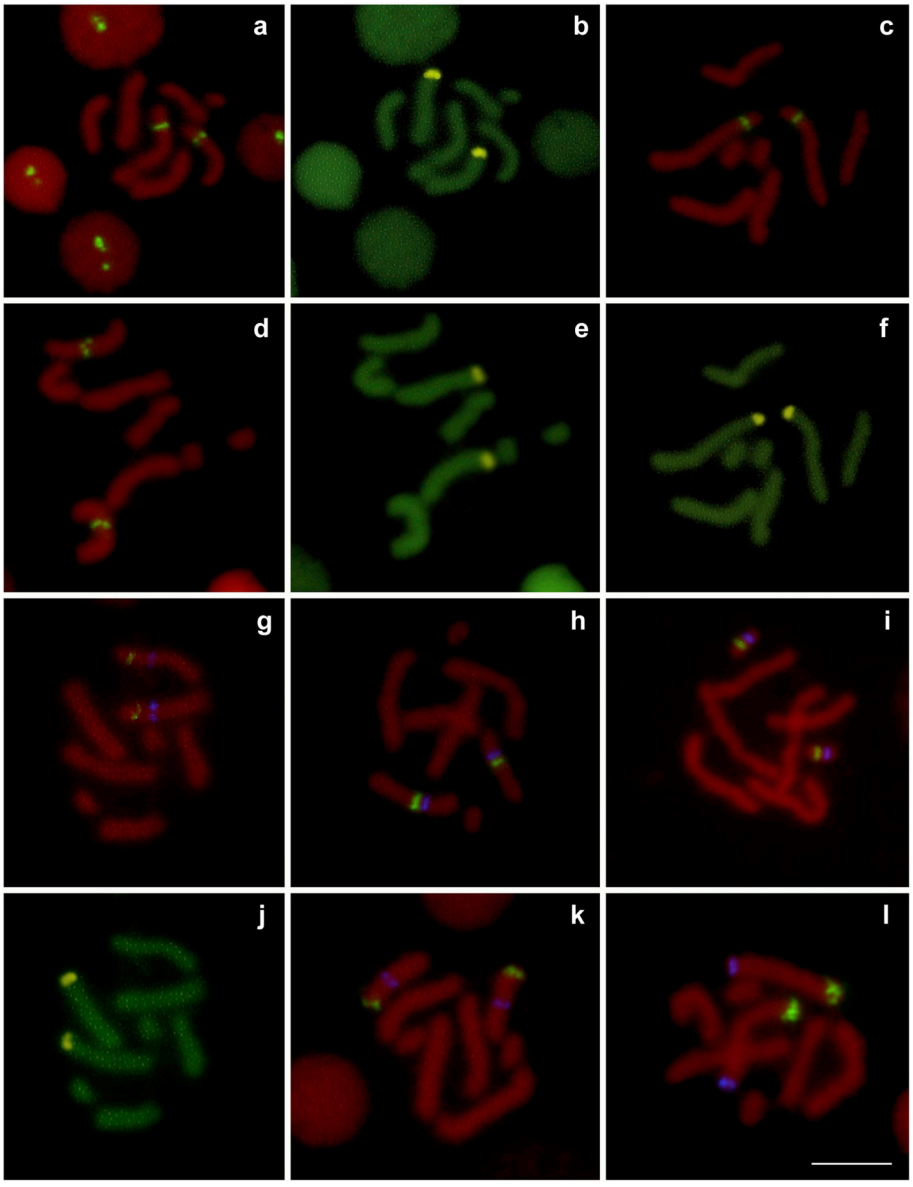

Fig. 1. *A. glycines* chromosomes stained with propidium iodide after fluorescent *in situ* hybridization (**a, c-d, g-i, k-l**) allowed to determine the localization of genes coding for carotenoid desaturase (**a**), 5S rDNA (**d**), esterase (**h**), 28S rDNA (**l**), together with the localization of the scaffolds 1039 (**c**), 1002 (**g**), 1000 (**h, k**), 1004 (**i**), 6374 (**i**), 6376 (**k**) and 1016 (**l**). In order to define the reciprocal position on chromosomes, double FISH experiments with FITC-labelled probes (in green) for histone cluster (**g**), scaffold 1000 (**h**) scaffold 1004 (**i**), scaffold 6376 (k) and 28S rDNA (**l**) and AMCA-labelled probes (in blue) 1002 (**g**), esterase (**h**), scaffold 6374 (**i**), scaffold 1000 (**k**) and scaffold 1016 (**l**) were also performed. CMA_3 staining (**b, e-f, j**) labelled a telomeric portion of the two X chromosomes thus facilitating their identification. Bar corresponds to 100 μm. **1039**: scaffold SBAphidCtg1039. **1016**: scaffold SBAphidCtg1016. **1000**: scaffold SBAphidCtg1000. **6376**: scaffold SBAphidCtg6376. **6374**: scaffold SBAphidCtg6374. **1002**: scaffold SBAphidCtg1002. **1004**: scaffold SBAphidCtg1004. **tor**: carotenoid desaturase. **hist**: histone gene cluster. **est**: esterase. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

1988). The amplification was performed with a thermocycler Hybaid at an annealing temperature of 59 °C for 30 s and an extension temperature of 72 °C for 45 s.

An internal portion of the 28S rDNA gene was amplified using the primers 28S–F (5′-AACAAACAACCGATACGTTCCG-3′) and 28S-R (5′-CTCTGTCCGTTTACAACCGAGC-3′), designed using the *A. pisum* 28S rDNA sequences available in GenBank (ID: KY558389.1). Amplification was performed using a Hybaid thermal cycler at an annealing temperature of 60 °C for 1 min with an extension temperature of 72 °C for 1 min.

The desaturase genes (*tor*) was amplified using the primers F (5′-ACTGGACACATTTTGACATCCT) and R (5′-TCAATGTCGGGCGTAA ATTACT), designed on the sequence of the *A. pisum* carotene desaturase gene (*tor*) available in GenBank (ID: 100169245).

A portion of the esterase gene was amplified using the primers esterase_F (5′-AAATCATATTTCCCGGGTTC) and esterase_R (5′–AGGTT CACTAAGATTACTCA), as previously reported in Rivi et al. (2013).

The amplification of the probes for 28S, 5S, esterase and *tor* were performed with an amplification mix containing 100 ng of genomic DNA as template, 1 μM of each primer, 200 μM dNTPs, and 2 U of DyNAZyme II polymerase (Finnzymes Oy, Vantaa, Finland). The results of the PCR amplification were evaluated in a 1.2% agarose gel electrophoresis.

Random priming probe biotin-labelling was performed with the Biotin High Prime (Roche), whereas the PCR digoxigenin labelling were performed using the Dig High Prime (Roche). Both labelling were performed according to the Roche protocols.

Fluorescent *in situ* hybridization (FISH) was performed as described by Monti et al. (2011) using fluorescein isothiocyanate (FITC)-conjugated anti-DIG antibodies (Roche) for the DIG labelled probes and

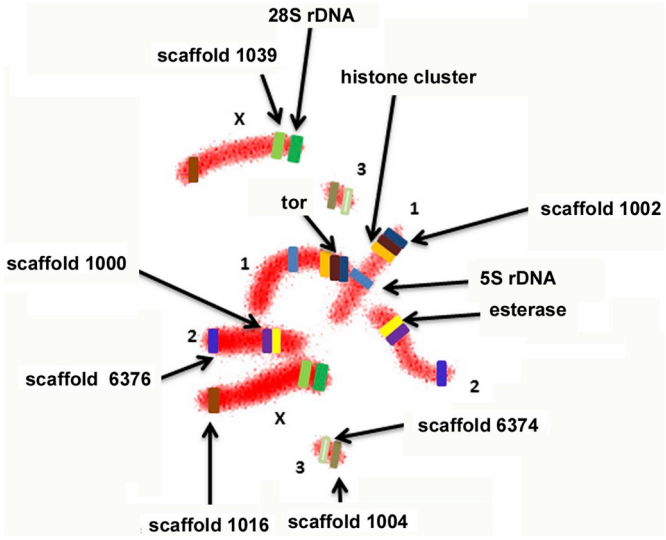

Fig. 2. Schematic summary showing the chromosomal localization of the mapped scaffolds and multigene families. **1039**: scaffold SBAphidCtg1039. **1016**: scaffold SBAphidCtg1016. **1000**: scaffold SBAphidCtg1000. **6376**: scaffold SBAphidCtg6376. **6374**: scaffold SBAphidCtg6374. **1002**: scaffold SBAphidCtg1002. **1004**: scaffold SBAphidCtg1004. **tor**: carotenoid desaturase. **hist**: histone gene cluster. **est**: esterase.

aminomethylcoumarin acetate (AMCA) conjugated-avidin for the biotin labelled probes. FISH slides were observed using a Zeiss Axioplan epifluorescence microscope. Photographs of the fluorescent images were taken using a CCD camera (Spot, Digital Instrument, Madison, USA) and the Spot software supplied with the camera. Images were processed using Adobe Photoshop (Adobe Systems, Mountain View, CA).

Search for orthologues of the genes mapped in *A. glycines*, was done in *M. persicae* and *D. melanogaster* by BLAST alignments in GenBank (http://blast.ncbi.nlm.nih.gov/Blast.cgi) both at the DNA and protein level, according to Moreno-Hagelsieb and Latimer (2008). A further confirmation of the identification, limited to aphids, was performed by BLAST alignments of aphid genomes using AphidBase (http://www.aphidbase.com). *M. persicae* genomic data were downloaded from Aphidbase (http://bipaa.genouest.org/is/aphidbase/myzus_persicae/). Successively, synteny was manually evaluated by comparing the distribution of genes in the aphid scaffolds with respect to the *D. melanogaster* genome assembly using Genome Browser (https://www.ncbi.nlm.nih.gov/genome).

3. Results

In order to conduct a first mapping of the *A. glycines* genome, a FISH approach based on the use of PCR-made long probes was conducted. The mapped scaffolds were: SBAphidCtg1000 (1,67 Mb), SBAphidCtg1002 (1,37 Mb), SBAphidCtg1004 (1,3 Mb), SBAphidCtg1016 (0,9 Mb), SBAphidCtg1039 (0,67 Mb), SBAphidCtg6374 (1,29 Mb) and SBAphidCtg6376 (1 Mb).

The above set of scaffolds, together with the localization of genes belonging to the multigene family 28S rDNA, 5s rDNA, esterase, *tor* and histone cluster, allowed the identification of at least two chromosomal markers for each *A. glycines* chromosomes (Figs. 1 and 2).

The desaturase gene (*tor*) (Fig. 1a), the gene of the histone cluster and the scaffold SBAphidCtg1002 mapped in close proximity near a terminal region of autosome 1 (Fig. 1g), whereas the 5S rDNA gene was located in an interstitial position of the same autosome (Fig. 1d).

The 28S rDNA genes were mapped on a telomeric portion of the two X chromosomes (Fig. 1l), whereas the scaffold SBAphidCtg1016 mapped at the opposite telomere (Fig. 1l). The scaffold SBAphidCtg1039 was localized in a sub-terminal position in the 28S rDNA-bearing telomere of the X chromosomes (Fig. 1c).

The scaffold SBAphidCtg6376 (Fig. 1k) was mapped in a telomeric position of autosomes 2, whereas scaffold SBAphidCtg1000 (Fig. 1h) and the esterase genes (Fig. 1h) were found located in a intercalary position closer to the opposite telomere of the same chromosomes.

Scaffolds SBAphidCtg1004 (Fig. 1i) and SBAphidCtg6374 (Fig. 1i) were mapped near the opposite termini of autosomes 3.

The localization of the seven *A. glycines* scaffolds allowed the comparison of the syntenic extent of the 53 localized genes in the soybean aphids as compared to *D. melanogaster* and *M. persicae* and thus permitted the evaluation of the chromosome rearrangement rate in two species of aphids in comparison to *D. melanogaster* (Figs. 3–9). The synteny analysis evidenced an average of 1.4 genes per syntenic block in aphids, suggesting a low degree of synteny between the two aphid species, in contrast to the 2.3 genes per syntenic block in the *A. glycines* - *D. melanogaster* comparison.

4. Discussion

With rare exceptions, the structure of aphid chromosomes and their evolution have received little attention. The few available published papers have been focused mainly on the X chromosomes in view of their involvement in sex determination and their enrichment in satellite DNA and heterochromatin (Bizzaro et al., 1999; Mandrioli and Borsatti, 2007; Monti et al., 2011; Jaquiéry et al., 2011, 2018; Mandrioli and Manicardi, 2012, 2014; Wilson et al., 2014; Muggli et al., 2015; Mandrioli et al., 2017). Actually, cytogenetics can serve as an independent and important mode of inquiry to validate currently available aphid genome assemblies. The need for this validation is strongly supported by recent data that points to the occurrence of a very high rate of misassembly in the *A. pisum* genome (Jaquiéry et al., 2018). Indeed, according to Jaquiéry et al. (2018), the pea aphid genome assembly Acyr 2.0 contains widespread errors and more than half of the scaffolds larger than 150 kb are chimeras of X and autosomes. These results confirmed a previous conclusion, based on a genetic map of 305 microsatellite markers (Jaquiéry et al., 2011), that high rates of scaffold misassembly were present in the pea aphid genome.

Moreover, Jaquiéry et al. (2018) suggested that assembly errors involving fragments of less than 10 kb could be present, but not detected given the low resolution of their analyses. Their conclusion was that the genome of the pea aphid presents considerable assembly problems to a degree that goes beyond the assembly results obtained with current assembly pipelines (e.g. Salzberg et al., 2004; Muggli et al., 2015).

As shown by results reported in Mandrioli et al. (2017), cytogenetics is a valuable tool not only to study aphid karyotype evolution, but also to furnish additional data useful for the assembly of genomes in aphids (including the reassembly of the pea aphid genome). Indeed, Mandrioli et al. (2017) developed an approach, based on 10 kbp long probes, which showed that cytogenetic data, can lead to genome assemblies of higher quality for aphids and other species.

In order to proceed with a first analysis of the extent of synteny and conservation in the gene order of aphids, we mapped seven *A. glycines*

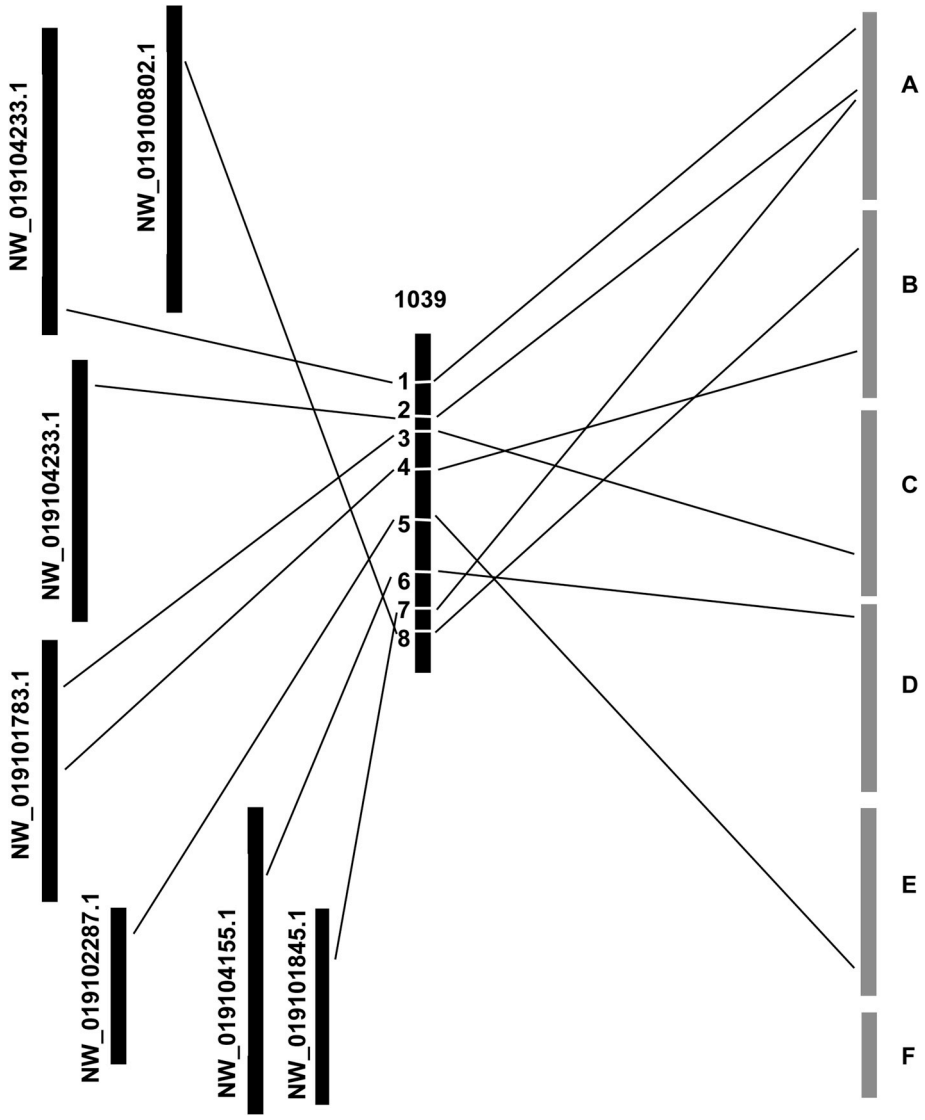

Fig. 3. Comparison of the gene map of *A. glycines* scaffold 1039 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: armadillo. **2**: vacuolar protein sorting 26. **3**: arginine N-methyltransferase. **4**: glycil-tRNA synthase. **5**: ribosomal protein L3. **6**: ribosomal biosynthesis regulatory protein. **7**: ubiquinone biosynthesis protein Coq7. **8**: COP9 signalosome subunit 8.

scaffolds (spanning 9 Mb and 53 genes) and five multigenic DNA families (major and minor rDNAs, histone gene cluster, esterases and carotenoid desaturases) identifying loci spanning all the four soybean aphid chromosomes, without evidencing any hybridization signal which could suggest the presence of chimeric assemblies in the biotype 1 *A. glycines* genome assembly. Our cytogenetic study allowed us to identify at least two chromosomal markers for each chromosome, located at the opposite terminal portions. *A. glycines* is therefore the first aphid species whose karyotype can be compared in different biotypes to determine if they are associated with chromosomal rearrangements.

Aphids possess holocentric chromosomes with a kinetic activity that is diffused along the entire chromosome which favours the stabilization of chromosomal rearrangements (Manicardi et al., 2015a, 2015b). Moreover, aphid reproduction mainly consists of apomictic parthenogenesis that fosters the inheritance of rearranged karyotypes (Manicardi et al., 2015a, 2015b). In view of these unique biological properties, we evaluated not only the synteny between aphids (*A. glycines* vs *M. persicae)*, but also the synteny extant between *A. glycines* and *Drosophila* to evaluate the impact of holocentrism on the evolution rate of aphid genomes.

The synteny analysis performed, comparing *A. glycines* gene maps with those of the peach potato aphid *M. persicae*, evidenced a low degree of synteny between the two aphid species in comparison to the aphid-fly comparison. An average of only 1.4 genes per syntenic block was observed when comparing the two aphid species while 2.3 genes/block were noted in the comparison *A. glycines - Drosophila*. Our results strongly mirror those found in Lepidoptera by d'Alençon et al. (2010), who compared *Bombyx mori* and the noctuid species *Spodoptera frugiperda* and *Helicoverpa armigera*, evidencing high rates of local genome rearrangements that resulted in small conserved syntenic blocks of

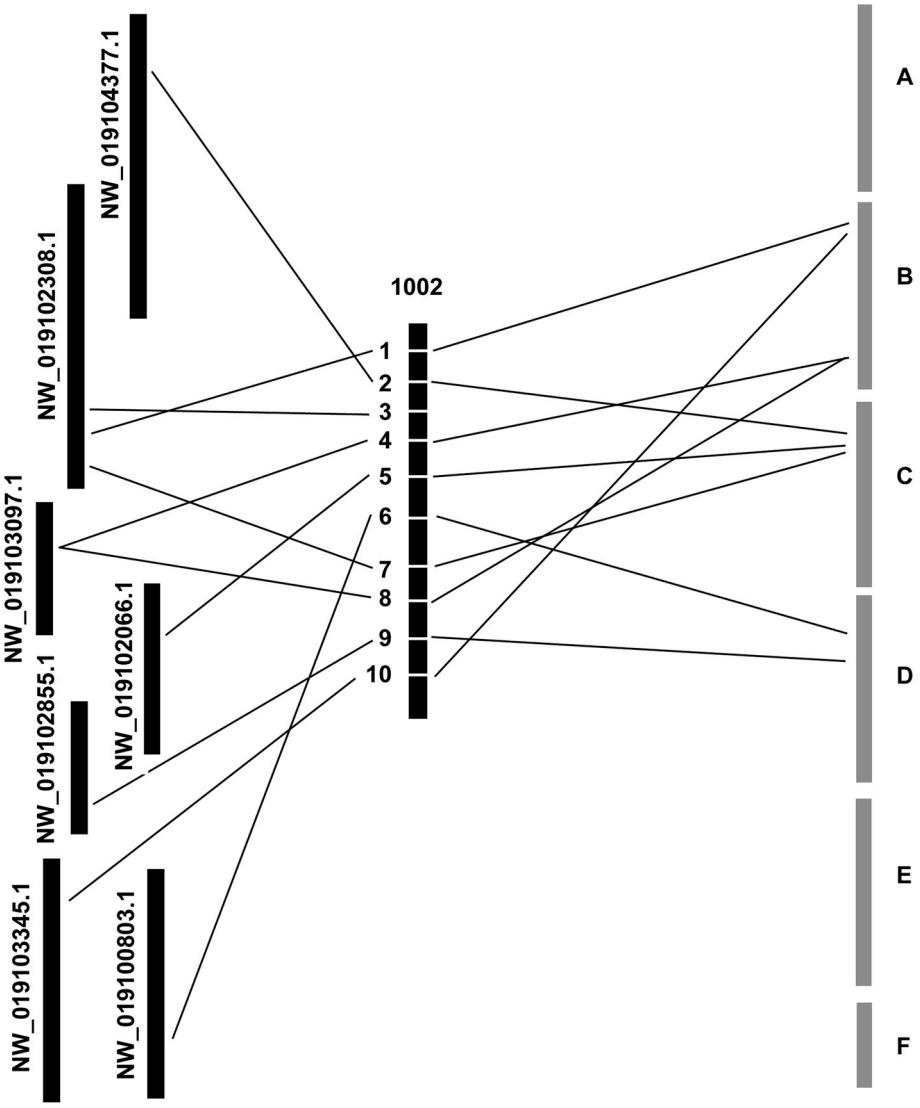

Fig. 4. Comparison of the gene map of *A. glycines* scaffold 1002 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: vacuolar fusion protein MON1. **2**: glutaredoxin. **3**: ATPase ASNA1. **4**: ribosomal protein L24. **5**: pre-mRNA splicing factor 38. **6**: ubiquitin-conjugating enzyme E2. **7**: transcription factor jun. **8**: ribosomal protein L24e. **9**: prefoldin. **10**: mitochondrial ribosomal protein L48.

genes approximately containing 1.3 genes per block between *B. mori* and the two noctuid species and 2.0 genes per block between *S. frugiperda* and *H. armigera*. This corresponds to approximately two chromosome breaks per Mb of DNA per million years pointing to an evolutionary rate that is much higher than that found among species in the genus *Drosophila*, supporting the hypothesis that holocentric chromosomes can greatly increase the number of local chromosomal rearrangements in respect to monocentric ones.

The analysis of the localization of the different studied gene families allowed a comparison with other aphid species, where these genes have been mapped. In particular, in aphids major rDNA genes are generally arranged as tandemly repeated clusters that have been localized at one telomere of the X chromosomes by *in situ* hybridization with 28S rDNA probes (Blackman et al., 2000; Mandrioli et al., 2011; Manicardi et al., 2015a, b). Exceptions include the interstitial position of rDNA genes in *Amphorophora idaei* (Fenton et al., 1994) and the autosomal localization of NORs in *Schoutedenia lutea* (Hales, 1989). According to the work presented herein, the major rDNA genes are likewise present at a telomere of the X chromosomes in *A. glycines* as well, confirming that the 28S rDNA genes are a *bona fide* marker for the sex chromosomes in aphids.

In contrast to the conserved position of the major rDNA genes, the localization of the 5S rDNA on aphid chromosomes varies between species. FISH experiments evidenced a single cluster located in an interstitial position in the autosomes 1 of *Aphis nerii* (Mandrioli et al., 2011) and two interstitial clusters on X chromosomes, together with a 5S rDNA cluster in an intercalary region of autosome 1, in *A. pisum* (Bizzaro et al., 2000). Using fluorescent *in situ* hybridization (FISH) experiments it has been clearly assessed that the 5S rDNA genes are located in an interstitial position of autosomes 1 in *A. glycines*, in

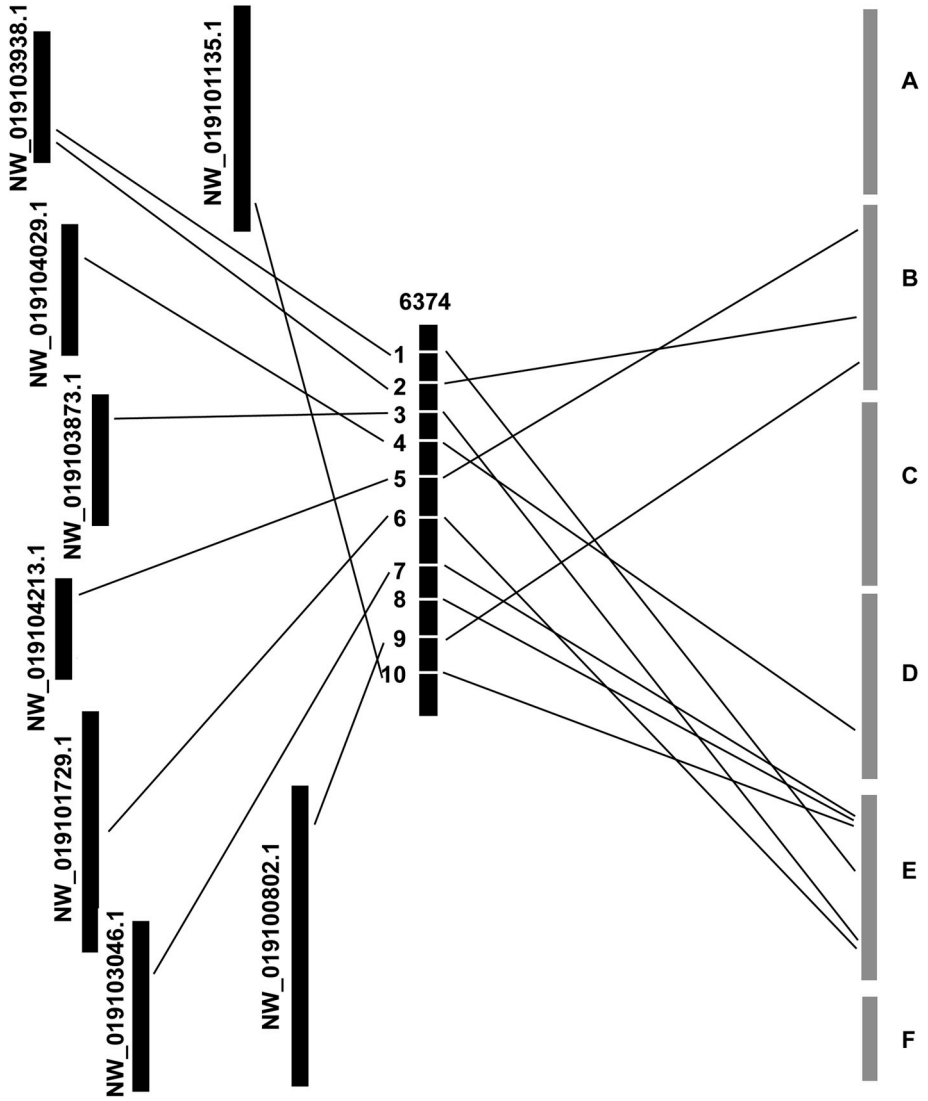

Fig. 5. Comparison of the gene map of *A. glycines* scaffold 6374 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: Ofd1 protein. **2**: nuclear pore complex protein nup107. **3**: glutaminyl-tRNA synthase. **4**: integrator complex subunit 9. **5**: ACC transporter ABCE. **6**: DUF2040. **7**: RagA G protein. **8**: proteasome suunit beta. **9**: acin1. **10**: ribosomal protein L37ae.

accordance to data previously reported for *Aphis nerii* (Mandrioli et al., 2011).

Histone coding genes have been studied in four aphid species, with some differences in their reported localization. In the wheat aphid, *Diuraphis noxia,* the histone H3 and H4 coding genes were co-localized at the 28S rDNA-bearing ends of the two X chromosomes (Novotná et al., 2011). As previously reported, this co-localization is not a unique feature of aphids and it suggests that histone genes could be inserted into the intergenic spacer region, as reported in Coleoptera (Roehrdanz et al., 2010). This hypothesis is further supported by FISH experiments with the H3 probe showing that the histone gene cluster is, like the rDNA genes, present in the silver nitrate positive bridge connecting the two X chromosomes at prometaphase, during male determination (Novotná et al., 2011). In contrast to these results, for *A. pisum*, *M. persicae* and *D. noxia,* the histone genes have been mapped in a single cluster located in an interstitial position of autosome 1 (Mandrioli and Manicardi, 2013). The results presented in this work for *A. glycines* confirmed that aphid histone genes are present in a single chromosome in the aphid karyotype and that the localization on autosome 1 appears to be the most conserved and likely most ancient position.

Most of the genes/scaffold identified in this work mapped towards the terminal portion of the *A. glycines* chromosomes. According to previous literature, due to the holocentric nature of aphid chromosomes, most of the crossing over events occur at the termini due to the fact that a centrally positioned chiasma cannot be resolved at anaphase I (Nokkala et al., 2004) so that the localization of most genes towards the chromosomal termini is evolutionary favoured. This suggestion is supported by what has been observed in the holocentric chromosomes of the hemipteran *Psylla foersteri* (Psylloidea), where the condensation of holocentric bivalents result in spatial reorganization of the chiasma

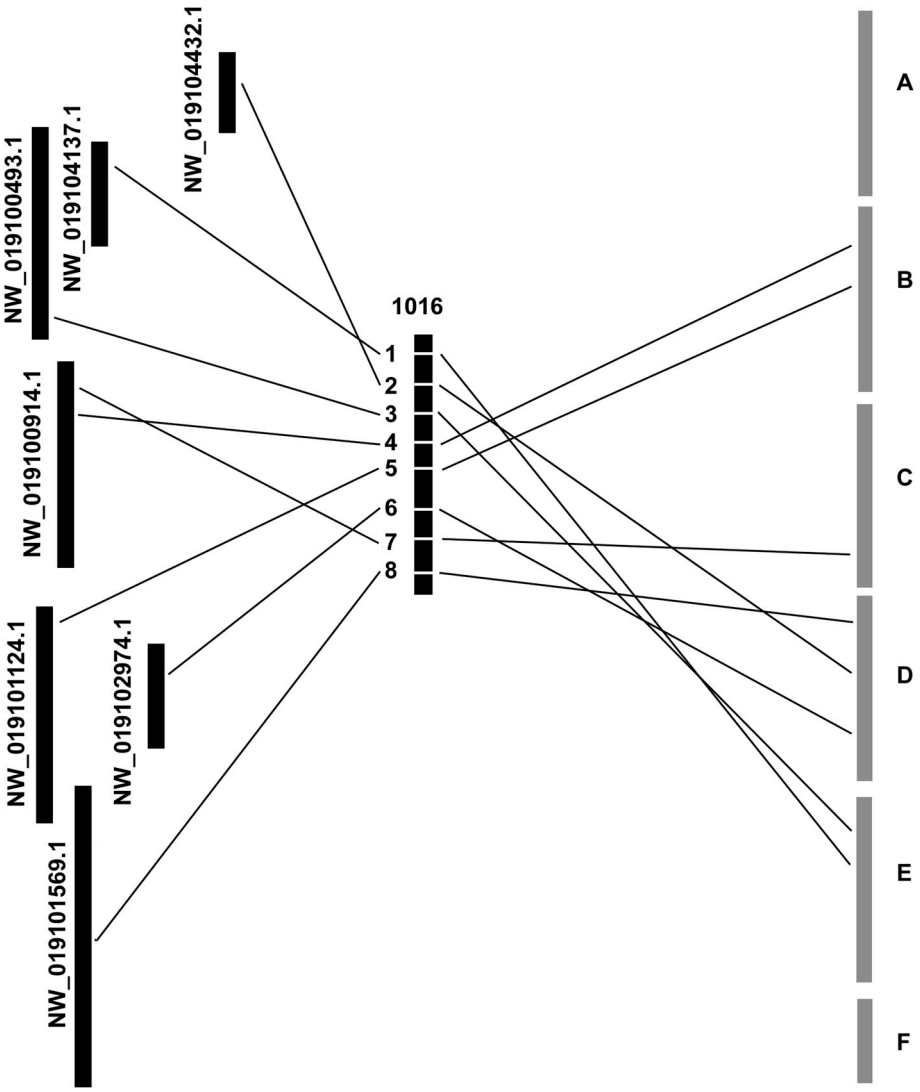

Fig. 6. Comparison of the gene map of *A. glycines* scaffold 1016 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: ataxin-2. **2**: tyrosine phosphatase. **3**: kibra. **4**: menin. **5**: transcription factor TFIIB. **6**: elongator complex protein 4. **7**: eukaryotic translation initiation factor 3 subunit K. **8**: UBA.

loops (Nokkala et al., 2004). In particular, it has been suggested that the holocentric structure of chromosomes act indirectly as restrictive factors upon the number and position of chiasmata that may be formed (Nokkala et al., 2004). Similarly, it has been observed that strong limitations are present in the number of crossing over in the holocentric chromosomes of the nematode *Caenorhabditis elegans* since each chromosome may form a single chiasma per bivalent at one of the termini (Riddle et al., 1997; Hillers and Villeneuve, 2003).

Our cytogenetic data provided valuable results that can be used to better disentangle problems related to the evolutionary trends of the aphid genome organization and karyotype structure. Moreover, our experiments did not detect the presence of chimeric assemblies in the newly available Biotype 1 *A. glycines* assembled genome. The integration of cytogenetics data into the standardized work flow in the genome assembly pipeline of future aphid genome projects can provide validation for the commonly used automated assembly pipelines. Indeed this revised approach could be useful to avoid the problems reported in the current pea aphid genome assembly.

Acknowledgements

The authors would like to thank A. Nardelli for his help in the preparation of the first set of chromosome slides. The *M. persicae* clone G006 genomic DNA sequence data were downloaded from AphidBase (http://www.aphidbase.com), made available in advance of publication. Funding for *M. persicae* clone G006 genomic sequencing was provided by USDA-NIFA Award 2010-65105-20558. Funding for the sequencing of the genome of *Aphis glycines* Biotype 1 was provided by the United Soybean Board (USB), Illinois Soybean Association (ISA), and the North Central Soybean Research Program (NCSRP) to R. Giordano.

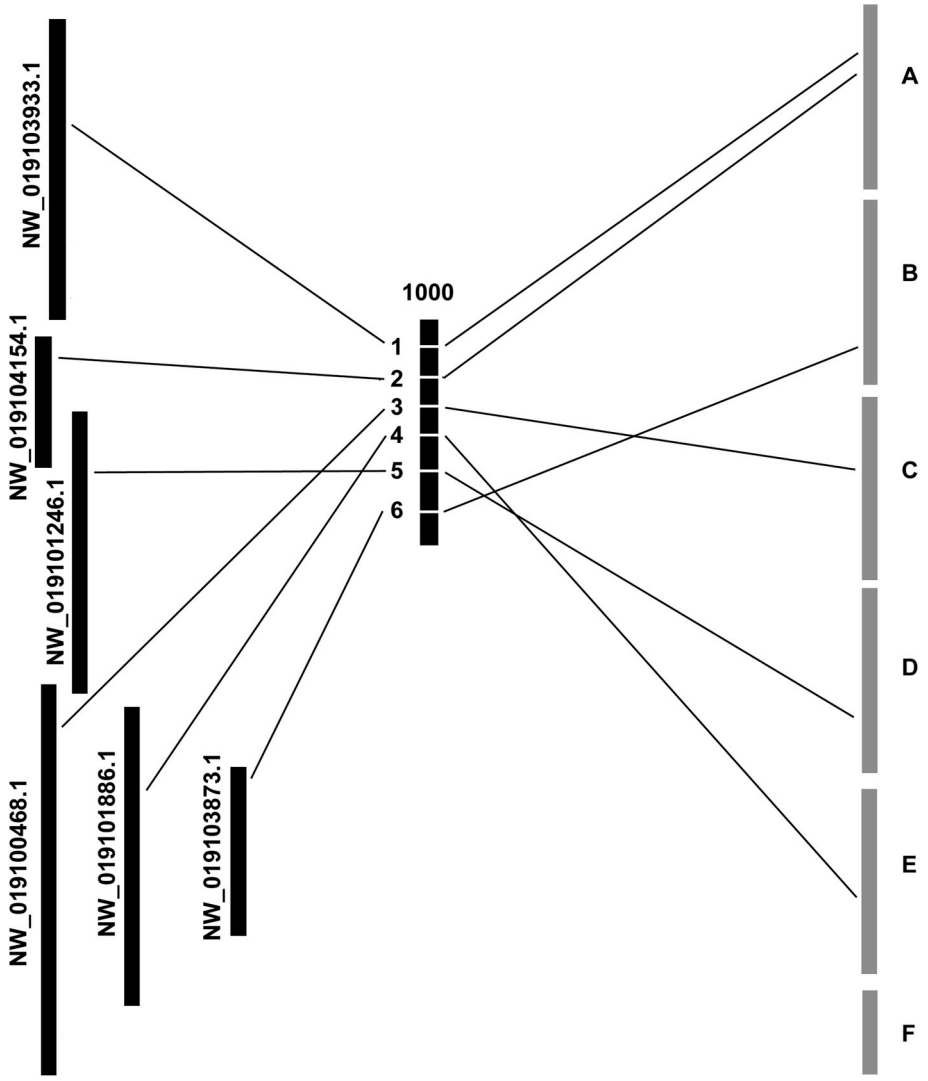

Fig. 7. Comparison of the gene map of *A. glycines* scaffold 1000 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: wings apart-like. **2**: GPI transamidase component PIG-T. **3**: rad4. **4**: mitochondrial matrix Mmp37. **5**: defective in cullin neddylation protein. **6**: dynactin subunit 6.

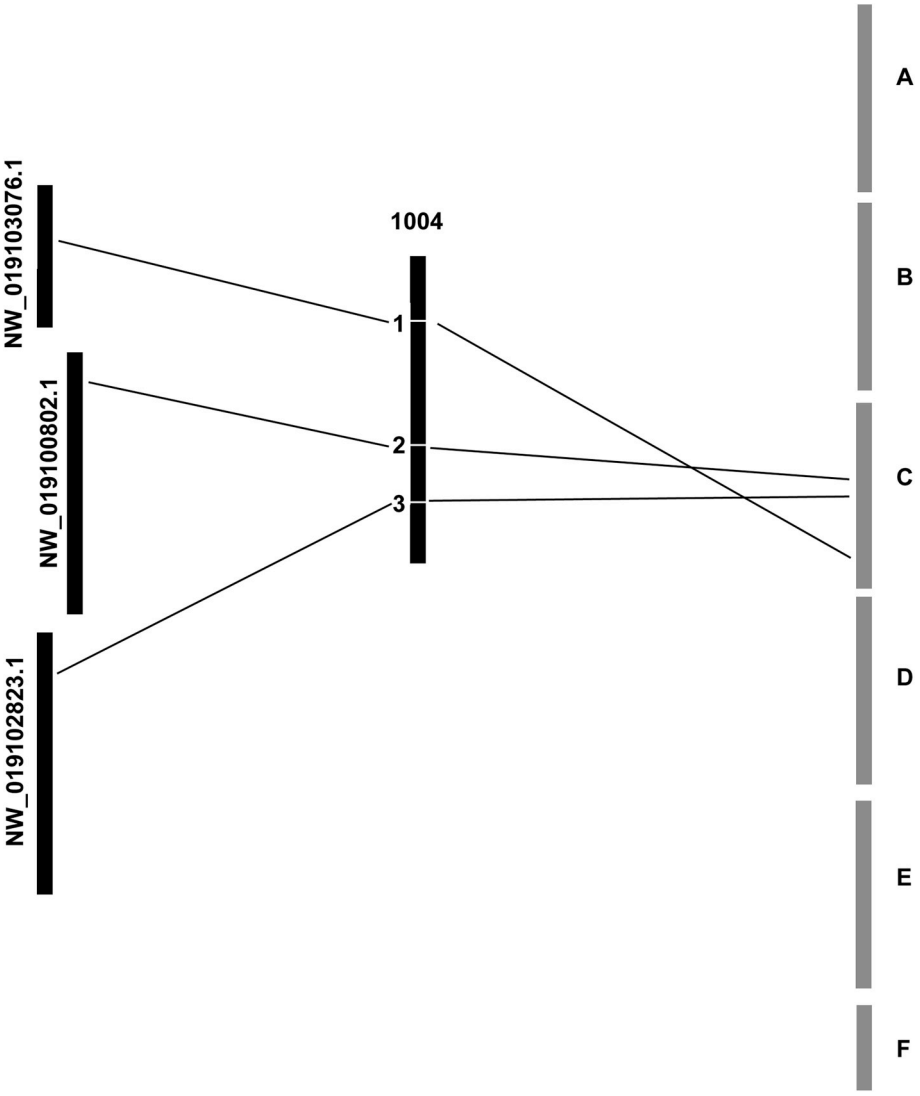

Fig. 8. Comparison of the gene map of *A. glycines* scaffold 1004 with *D. melanogaster* Muller elements A-F (grey bars at right) and *M. persicae* scaffolds (black bars at left). **1**: af4/fmr2 family member 4. **2**: FAD/NAD-binding protein. **3**: ribosome-binding factor A.

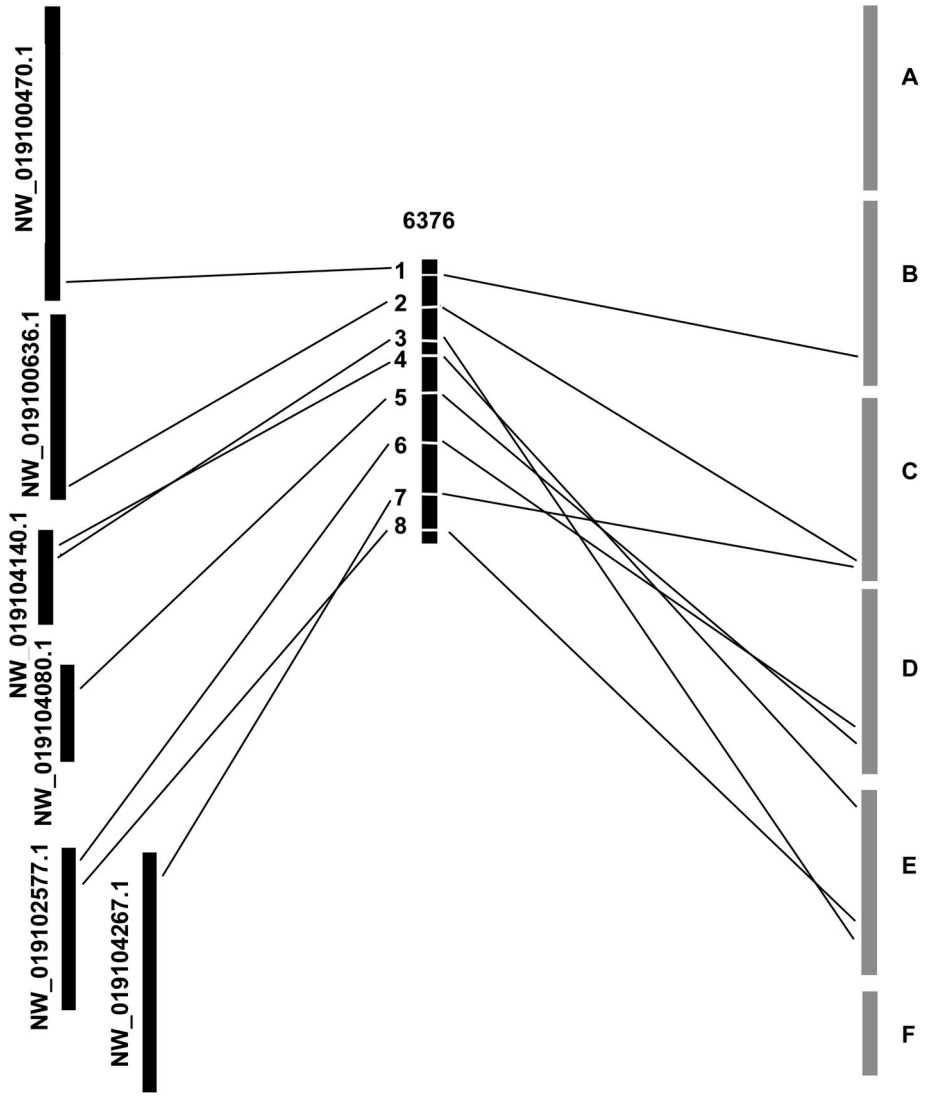

Fig. 9. Comparison of the gene map of *A. glycines* scaffold 6376 with the *D. melanogaster* Muller elements A-F (grey bars at right) and the *M. persicae* scaffolds (black bars at left). **1**: cullin. **2**: thiolase. **3**: DNA polymerase epsilon. **4**: transcription initiation factor IIF, beta subunit. **5**: mitochondrial ribosomal protein S26. **6**: transcription factor AP-2. **7**: mitochondrial F1–F0 ATP synthase subunit F **8**: INO80 complex subunit E.

References

Arensburger, P., Megy, K., Waterhouse, R.M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C.L., Campbell, K.S., Casola, C., Castro, M.T., Chandramouliswaran, I., Chapman, S.B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., Fraser-Liggett, C., Guigo, R., Haas, B., Hammond, M., Hansson, B.S., Hemingway, J., Hill, S.R., Howarth, C., Ignell, R., Kennedy, R.C., Kodira, C.D., Lobo, N.F., Mao, C., Mayhew, G., Michel, K., Mori, A., Liu, N., Naveira, H., Nene, V., Nguyen, N., Pearson, M.D., Pritham, E.J., Puiu, D., Qi, Y., Ranson, H., Ribeiro, J.M., Roberston, H.M., Severson, D.W., Shumway, M., Stanke, M., Strausberg, L., Sun, C., Sutton, G., Tu, Z.J., Tubio, J.M., Unger, M.F., Vanlandingham, D.L., Vilella, A.J., White, O., White, J.R., Wondji, C.S., Wortman, J., Zdobnov, E.M., Birren, B., Christensen, B.M., Collins, F.H., Cornel, A., Dimopoulos, G., Hannick, L.I., Higgs, S., Lanzaro, G.C., Lawson, D., Lee, N.H., Muskavitch, M.A., Raikhel, A.S., Atkinson, P.W., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330, 86–88.

Bizzaro, D., Barbolini, E., Mandrioli, M., Manicardi, G.C., Mazzoni, E., 1999. Cytogenetic characterization of the holocentric chromosomes in the aphids *Myzus varians* and *Myzus cerasi*. Caryologia 52, 81–85.

Bizzaro, D., Mandrioli, M., Zanotti, M., Giusti, M., Manicardi, G.C., 2000. Chromosome analysis and molecular characterization of highly repeated DNAs in the aphid *Acyrthosiphon pisum* (Aphididae, Hemiptera). Genetica 108, 197–202.

Blackman, R.L., Spence, J.M., Normark, B.B., 2000. High diversity of structurally heterozygous karyotypes and rDNA arrays in parthenogenetic aphids of the genus *Trama*. Heredity 84, 254–260.

Braendle, C., Caillaud, M.C., Stern, D.L., 2005. Genetic mapping of aphicarus: a sex-linked locus controlling a wing polymorphism in the pea aphid (*Acyrthosiphon pisum*). Heredity 94, 435–442.

Brisson, J.A., Davis, G.K., 2008. Pea aphid. In: Hunter, W. (Ed.), In: Kole, C. (Ed.), Genome Mapping and Genomics in Arthropods, vol. 1. Springer-Verlag: Berlin & Heidelberg, pp. 59–67.

Brucker, R.M., Funkhouser, L.J., Setia, S., Pauly, R., Bordenstein, S.R., 2012. Insect innate immunity database (IIID): an annotation tool for identifying immune genes in insect genomes. PLoS One 7, e45125.

Chilana, P., Sharma, A., Rai, A., 2012. Insect genomic resources: status, availability and future. Curr. Sci. 102, 571–580.

Cooper, S.G., Concibido, V.C., Estes, R., Hunt, D.W.A., Jiang, G., Krupke, C.H., McCornack, B., Mian, R., O'Neal, M., Poysa, V., Voldseth, D.P., Ragsdale, D., Tinsley, N., Wang, D., 2015. Geographic distribution of soybean aphid biotypes in the United States and Canada during 2008–2010. Crop Sci. 55, 2598–2608.

d'Alençon, E., Sezutsu, H., Legeai, F., Permal, E., Bernard-Samain, S., Gimenez, S., Gagneur, C., Cousserans, F., Shimomura, M., Brun-Barale, A., Flutre, T., Couloux, A., East, P., Gordon, K., Mita, K., Quesneville, H., Fournier, P., Feyereisen, R., 2010. Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. Proc. Natl. Acad. Sci. U.S.A. 107, 7680–7685.

Dudchenko, O., Batr, a S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., Aiden, E.L., 2017. *De novo* assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science 356, 92–95.

Fenton, B., Birch, A.N.E., Malloch, G., Woodford, J.A.T., Gonzalez, C., 1994. Molecular analysis of ribosomal DNA from the aphid *Amphorophora idaei* and an associated fungal organism. Insect Mol. Biol. 3, 183–189.
Geiduschek, E.P., Tocchini-Valentini, G.P., 1988. Transcription by RNA polymerase III. Annu. Rev. Biochem. 57, 873–914.
Grimmelikhuijzen, C.J., Cazzamali, G., Williamson, M., Hauser, F., 2007. The promise of insect genomics. Pest Manag. Sci. 63, 413–416.
Hales, D.F., 1989. The chromosomes of *Schoutedenia lutea* (Homoptera, Aphididae, Greenideinae) with an account of meiosis in male. Chromosoma 98, 295–300.
Hawthorne, D.J., Via, S., 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412, 904–907.
Hillers, K.J., Villeneuve, A.M., 2003. Chromosome-wide control of meiotic crossing over in *C. elegans*. Curr. Biol. 13, 1641–1647.
Honeybee Genome Sequencing Consortium, 2006. Insights into social insects from the genome of the honey bee. Nature 443, 931–949.
International Aphid Genomics Consortium, 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8, e1000313.
Jaquiéry, J., Stoeckel, S., Rispe, C., Mieuzet, L., Legeai, F., Simon, J.C., 2011. Accelerated evolution of sex chromosomes in aphids, an X0 system. Mol. Biol. Evol. 29, 837–847.
Jaquiéry, J., Peccoud, J., Ouisse, T., Legeai, F., Prunier-Leterme, N., Gouin, A., Nouhaud, P., Brisson, J.A., Bickel, R., Purandare, S., Poulain, J., Battail, C., Lemaitre, C., Mieuzet, L., Le Trionnaire, G., Simon, J.C., Rispe, C., 2018. Disentangling the causes for faster-X evolution in aphids. Genome Biol. Evol. 10, 507–520.
Mandrioli, M., Borsatti, F., 2007. Analysis of heterochromatin epigenetic markers in the holocentric chromosomes of the aphid *Acyrthosiphon pisum*. Chromosome Res. 15, 1015–1022.
Mandrioli, M., Manicardi, G.C., 2012. Unlocking holocentric chromosomes: new perspectives from comparative genomics? Curr. Genom. 13, 343–349.
Mandrioli, M., Manicardi, G.C., 2013. Chromosomal mapping reveals a dynamic organization of the histone genes in aphids. Entomologia 1, 7. https://doi.org/10.4081/entomologia.2013.e2.
Mandrioli, M., Manicardi, G.C., 2014. Mapping the aphid genome: the cytogenetic dimension of a pest crop insect. In: Sharakhov, I.V. (Ed.), Protocols for Cytogenetic Mapping of Arthropod Genome. Taylor and Francis Group, pp. 325–348.
Mandrioli, M., Azzoni, P., Lombardo, G., Manicardi, G.C., 2011. Composition and epigenetic markers of heterochromatin in the aphid *Aphis nerii* (Hemiptera: Aphididae). Cytogenet. Genome Res. 133, 67–77.
Mandrioli, M., Zambonini, G., Manicardi, G.C., 2017. Comparative gene mapping as a tool to understand the evolution of pest crop insect chromosomes. Int. J. Mol. Sci. 18, 1919–1933.
Manicardi, G.C., Mandrioli, M., Blackman, R.L., 2015a. The cytogenetic architecture of the aphid genome. Biol. Rev. 90, 112–125.
Manicardi, G.C., Nardelli, A., Mandrioli, M., 2015b. Fast chromosomal evolution and karyotype instability: occurrence of recurrent chromosomal rearrangements in the peach potato aphid *Myzus persicae* (Hemiptera, Aphididae). Biol. J. Linn. Soc. 116, 519–529.
Mesquita, R.D., Vionette-Amaral, R.J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F.A., Minx, P., Spieth, J., Carvalho, A.B., Panzera, F., Lawson, D., Torres, A.Q., Ribeiro, J.M., Sorgine, M.H., Waterhouse, R.M., Montague, M.J., Abad-Franch, F., Alves-Bezerra, M., Amaral, L.R., Araujo, H.M., Araujo, R.N., Aravind, L., Atella, G.C., Azambuja, P., Berni, M., Bittencourt-Cunha, P.R., Braz, G.R., Calderón-Fernández, G., Carareto, C.M., Christensen, M.B., Costa, I.R., Costa, S.G., Dansa, M., Daumas-Filho, C.R., De-Paula, I.F., Dias, F.A., Dimopoulos, G., Emrich, S.J., Esponda-Behrens, N., Fampa, P., Fernandez-Medina, R.D., da Fonseca, R.N., Fontenele, M., Fronick, C., Fulton, L.A., Gandara, A.C., Garcia, E.S., Genta, F.A., Giraldo-Calderón, G.I., Gomes, B., Gondim, K.C., Granzotto, A., Guarneri, A.A., Guigó, R., Harry, M., Hughes, D.S., Jablonka, W., Jacquin-Joly, E., Juárez, M.P., Koerich, L.B., Lange, A.B., Latorre-Estivalis, J.M., Lavore, A., Lawrence, G.G., Lazoski, C., Lazzar, i CR., Lopes, R.R., Lorenzo, M.G., Lugon, M.D., Majerowicz, D., Marcet, P.L., Mariotti, M., Masuda, H., Megy, K., Melo, A.C., Missirlis, F., Mota, T., Noriega, F.G., Nouzova, M., Nunes, R.D., Oliveira, R.L., Oliveira-Silveira, G., Ons, S., Orchard, I., Pagola, L., Paiva-Silva, G.O., Pascual, A., Pavan, M.G., Pedrini, N., Peixoto, A.A., Pereira, M.H., Pike, A., Polycarpo, C., Prosdocimi, F., Ribeiro-Rodrigues, R., Robertson, H.M., Salerno, A.P., Salmon, D., Santesmasses, D., Schama, R., Seabra-Junior, E.S., Silva-Cardoso, L., Silva-Neto, M.A., Souza-Gomes, M., Sterkel, M., Taracena, M.L., Tojo, M., Tu, Z.J., Tubio, J.M., Ursic-Bedoya, R., Venancio, T.M., Walter-Nuno, A.B., Wilson, D., Warren, W.C., Wilson, R.K., Huebner, E., Dotson, E.M., Oliveira, P.L., 2015. Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. Proc. Natl. Acad. Sci. U.S.A. 112, 14936–14941.
Monti, V., Giusti, M., Bizzaro, D., Manicardi, G.C., Mandrioli, M., 2011. Presence of a functional $(TTAGG)_n$ telomere-telomerase system in aphids. Chromosome Res. 19, 625–633.
Moreno-Hagelsieb, G., Latimer, K., 2008. Choosing BLAST options for better detection of orthologs as reciprocal best hits. Bioinformatics 24, 319–324.
Muggli, M.D., Puglisi, S.J., Ronen, R., Boucher, C., 2015. Misassembly detection using paired-end sequence reads and optical mapping data. Bioinformatics 31, i80–i88.
Nicholson, S.J., Nickerson, M.L., Dean, M., Song, Y., Hoyt, P.R., Rhee, H., Kim, C., Puterka, G.J., 2015. The genome of *Diuraphis noxia*, a global aphid pest of small grains. BMC Genomics 16, 429.
Nokkala, S., Kuznetsova, V.G., Maryanska-Nadachowska, A., Nokkala, C., 2004. Holocentric chromosomes in meiosis. I. Restriction of the number of chiasmata in bivalents. Chromosome Res. 12, 733–739.
Novotná, J., Havelka, J., Stary, P., Koutecky, P., Vıtkova, M., 2011. Karyotype analysis of the Russian wheat aphid, *Diuraphis noxia* (Kurdjumov) (Hemiptera: Aphididae), reveals a large X chromosome with rRNA and histone gene families. Genetica 139, 281–289.
Quan, Q., Hu, X., Pan, B., Zeng, B., Wu, N., Fang, G., Cao, Y., Chen, X., Huang, Y., Zhan, S., 2019. Draft genome of the cotton aphid *Aphis gossypii*. Insect Biochem. Mol. Biol (in press).
Riddle, D.L., Blumenthal, T., Meyer, B.J., Priess, J.R., Albertson, D.G., Rose, A.M., Villeneuve, A.M., 1997. Chromosome organization, mitosis, and meiosis. In: Riddle, D.L., Blumenthal, T., Meyer, B.J., Priess, J.R. (Eds.), *C. elegans* II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 47–48.
Rivi, M., Monti, V., Mazzoni, E., Cassanelli, S., Panini, M., Anaclerio, M., Cigolini, M., Corradetti, B., Bizzaro, D., Mandrioli, M., Manicardi, G.C., 2013. A1-3 chromosomal translocations in Italian populations of the peach potato aphid *Myzus persicae* (Sulzer) not linked to esterase-based insecticide resistance. Bull. Entomol. Res. 103, 278–285.
Roehrdanz, R., Heilmann, L., Senechal, P., Sears, S., Evenson, P., 2010. Histone and ribosomal RNA repetitive gene clusters of the boll weevil are linked in a tandem array. Insect Mol. Biol. 19, 463–471.
Salzberg, S.L., Church, D., Di Cuccio, M., Yaschenko, E., Ostell, J., 2004. The genome assembly archive: a new public resource. PLoS Biol. 2, e285.
Tribolium Genome Sequencing Consortium, 2008. The genome of the model beetle and pest *Tribolium castaneum*. Nature 452, 949–955.
Van Emden, H.F., Harrington, R., 2007. Aphids as Crop Pests. CABI Publishing, Oxfordshire, UK.
Wenger, J.A., Cassone, B.J., Legeai, F., Johnston, J.S., Bansal, R., Yates, A.D., Coates, B.S., Pavinato, V.A., Michel, A., 2017. Whole genome sequence of the soybean aphid, *Aphis glycines*. Insect Biochem. Mol. Biol. (17), 30005–X. https://doi.org/10.1016/j.ibmb.2017.01.005. pii: S0965-1748.
Wilson, A.C.C., Delgado, R.N., Vorburger, C., 2014. Biased transmission of sex chromosomes in the aphid *Myzus persicae* is not associated with reproductive mode. PLoS One 9, e116348.