Genome-wide developed microsatellites reveal a weak population differentiation in the hoverfly *Eupeodes corollae* (Diptera: Syrphidae) across China

Meng-Jia Liu¹,²,³#, Xiao-qiang Wang¹#, Ling Ma², Li-Jun Cao², Hong-Ling Liu³, De-Qiang Pu¹,³*, Shu-Jun Wei²*

1. Industrial Crop Research Institute, Sichuan Academy of Agricultural Sciences, 159 Huajin Road, Chengdu 610300, China
2. Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
3. Institute of Plant Protection, Sichuan Academy of Agricultural Sciences, 20 Jingjusi Road, Chengdu 610066, China

# These authors contributed equally.

Corresponding authors: Shu-Jun Wei, De-Qiang Pu

Shu-Jun Wei, Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China, Email: shujun268@163.com; Tel.: +86-010-51503439

De-Qiang Pu, Institute of Plant Protection, Sichuan Academy of Agricultural Sciences, 20 Jingjusi Road, Chengdu 610066, China, Email: pdqpudeqiang@163.com.

Running title: *E. corollae* genetic structure
Abstract: The hoverfly, *Eupeodes corollae*, is a worldwide natural enemy of aphids and a plant pollinator. To provide insights into the biology of this species, we examined its population genetic structure by obtaining 1.15-GB random genomic sequences using next-generation sequencing and developing genome-wide microsatellite markers. A total of 79,138 microsatellite loci were initially isolated from the genomic sequences; after strict selection and further testing of 40 primer pairs in eight individuals, 24 polymorphic microsatellites with high amplification rates were developed. These microsatellites were used to examine the population genetic structure of 96 individuals from four field populations collected across southern to northern China. The number of alleles per locus ranged from 5 to 13 with an average of 8.75; the observed and expected heterozygosity varied from 0.235 to 0.768 and from 0.333 to 0.785, respectively. Population genetic structure analysis showed weak genetic differentiation among the four geographical populations of *E. corollae*, suggesting a high rate of gene flow reflecting likely widespread migration of *E. corollae* in China.

Key words: *Eupeodes corollae*, microsatellite, genome, population genetic structure, migration
**Introduction**

*Eupeodes corollae* is one of the most common hoverflies with a worldwide distribution [1, 2]. The larval stage of this species is mostly insectivorous, feeding mainly on aphids [3-5] while adults are pollinators [6-8]. Many hoverfly species are important biological control agents of aphids due to their rapid dispersal and absence of summer diapause compared with other aphidophaga [9]. Understanding the biology and behavior of hoverflies can help in assessing their potential as biological control agents of aphids.

Hoverflies migrate seasonally as revealed by radar monitoring [10] and isotopic tools [11]. Population genetic analysis is also frequently employed to reveal the migration of species as a complementary approach to traditional methods [12-15]. In populations of the hoverflies *Cheilosia longula* [16], *Blera fallax* [17], *Sphaerophoria scripta* and *Episyrphus balteatus* [18], population genetic differentiation has not been found between some regions, suggesting migratory movements of these hoverflies between regions including southern and northern regions of Europe [18, 19]. However, some hoverflies, such as *E. balteatus* and *Scaeva selenitica*, are only partially migratory [20].

Previous studies reported that *E. corollae* is a highly migratory species in Europe [21-23], but its migratory behavior of *E. corollae* remains unclear in other areas. *E. corollae* is commonly found across China, but the ecology and biology of this species has rarely been studied [8]. In this study, we conducted a preliminary examination of the population genetic structure of *E. corollae* in China. First, we obtained random genomic sequences of *E. corollae* using next-generation sequencing and developed an effective and informative set of microsatellite markers of *E. corollae*. We used this novel set of microsatellite markers to investigate the genetic structure of four *E. corollae* populations collected from four representative regions across China.

**Materials and methods**

**Sample collection and DNA extraction**

A male adult from a laboratory (Sichuan Academy of Agriculture Sciences)-reared line of *E. corollae* was used for generating genome sequences. Four field populations of *E. corollae* were collected from China in March to July 2017 (Table 1, Figure 1a). To avoid the sampling of siblings, adults in a site were collected using insect net with individuals sampled separated by about 20 meters. A total of eight individuals from field collections were used for initial testing of selected primers. Twenty-four individuals from each of the four populations were then used for a population level survey. All samples were stored in absolute ethanol, frozen at −80 °C and stored at the Integrated Pest Management Laboratory of the Beijing Academy of Agriculture and Forestry Sciences. The thorax from each individual *E. corollae* was used for genomic DNA extraction using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany).

**Genome sequencing and assembly**

The extracted genomic DNA from a laboratory-reared individual was used in constructing a high-throughput sequencing library with 500-bp insert size using the Illumina TruSeq DNA
PCR-Free HT Library Prep Kit (Illumina, San Diego, CA, USA). The prepared library was sequenced on an Illumina Hiseq4000 Sequencer using the Hiseq Reagent Kit v3 (Illumina, San Diego, CA, USA) by Beijing BerryGenomics Co., Ltd. The raw data were trimmed by removing the low quality reads using Trimmomatic 0.36 [24] and then the sequences were evaluated by FastQC v 0.11.5 [25]. IDBA was used to assemble the generated genomic sequences with K-mer from 20 to 140 [26].

**Genome-wide microsatellite survey and primer design**

MSDB was used to search all potential microsatellite loci (repeat units of 2, 3, 4, 5, and 6 corresponding to the minimum number of repeats of 7, 5, 4, 4, and 4, respectively) from the assembled genomic sequences of *E. corollae* [27]. QDD was used to isolate microsatellites and design primers [28]. The outputs of primer pairs from QDD were further filtered by the following criteria [29, 30]: (i) the corresponding microsatellites were pure and specific; (ii) the design strategy of ‘A’ was used to avoid primer secondary structure and repeats; (iii) the minimum distance between the 3’ end of a primer and its target region should be longer than 10 bp; (iv) the annealing temperature for each primer pairs was set between 58 °C and 62 °C to avoid large differences among primers; (v) the estimated PCR product size of the primer pairs was from 100 to 350 bp.

**Polymorphic microsatellite isolation**

After screening primers from the QDD program, a universal primer (CAGGACCAGGCTACCGTG) was added to the 5’ end of each selected forward primer to allow efficient combining with the fluorescent label [31]. Amplifications were performed using the GoTaq Green Master Mix (Promega, USA) in a final volume of 10 μl system with 0.5 μl of template DNA (5–20 ng/μl), 5 μl of Master Mix (Promega, Madison, WI, USA), 3.94 μl of ddH2O, 0.08 μl forward primer, 0.16 μl reverse primer and 0.32 μl universal primer labeled with fluorescence (FAM, HEX, and ROX sequencing dyes). The PCR protocol was set as: 5 min for 95 °C, 35 cycles of amplification with 95 °C for 30s, 56 °C for 40s, and 72 °C for 40s. Final extension was with 72 °C for 15 min. PCR products were analyzed on an ABI 3730xl DNA Analyzer (Applied Biosystems, USA) using the GeneScan 500 LIZ size standard (Applied Biosystems, USA). Genotyping was conducted by GENEMAPPER 4.0 (Applied Biosystems, USA). Those primer pairs with amplification efficiency lower than 75%, showing monomorphism in eight individuals, or producing more than two peaks (non-specific amplification) were discarded.

**Genetic diversity and population genetic structure analyses**

GENEPOP version 4.0.11 [32] was used to test the likelihood of deviation from Hardy-Weinberg equilibrium (HWE) and the linkage disequilibrium (LD) at each microsatellite locus, the inbreeding coefficient (FIS) and pairwise population differentiation (FST). Allele frequencies, expected heterozygosity (HE) and observed heterozygosity (HO) were calculated with the macros Microsatellite Tools [33].

Population genetic structure was analyzed by STRUCTURE version 2.3.4 [34]. The clustering test was replicated 30 times for each K value ranging from 1 to 5 with a burn-in of
100,000 iterations followed by 200,000 Markov Chain Monte Carlo iterations. The Delta (K) method was used to estimate optimal K values by submitting the STRUCTURE output to Structure Harvester Web 0.6.94 [35]. Visualization of the results was handled by CLUMPP version 1.1.2 [36] and DISTRUCT version 1.1 [37]. Additional, BAPS version 6.0 software (Bayesian analysis of population structure) was used to incorporate spatial information into clustering of individuals.

Results and discussion

Genomic sequences of E. corollae

A total of 51.53 Gb paired-end (PE) sequences (184,394,506 reads each with a length of 150 bp) was obtained and the genomic size of E. corollae was estimated to be 12315 Mb. Trimmed reads were assembled into 2563327 scaffolds with a total length of 1.15 Gb ranging from 100 bp to 437.63 KB, with an N50 of 1510 bp. These contigs were used for microsatellite discovery.

Microsatellite characteristics of E. corollae

79,138 microsatellite loci were isolated from the randomly sequenced genome sequences of E. corollae with 5000 (6.32%) dinucleotide repeat (DNR) sites, 29221 (36.92%) trinucleotide repeat (TNR) sites, 30988 (39.16%) tetranucleotide repeat (TTNR) sites, 6635 (8.38%) pentanucleotide repeats (PNR) sites and 7294 (9.22%) hexanucleotide repeat (HNR) sites. The frequency of dinucleotide repeats in E. corollae is unusually low when compared with other insect species such as Grapholita molesta [30] (Lepidoptera), Aphis glycines (Hemiptera) [38] and Obolodiplosis robiniae (Diptera) [39], which shows the distribution of microsatellites to vary among species [40, 41].

Development of variable microsatellite markers

The QDD program initially generated 18114 primer pairs; we selected those corresponding to tri- and tetra-nucleotide microsatellites for further filtering under criteria listed in the methods and obtained 40 primer pairs. These primer pairs were validated in eight individuals of E. corollae; six pairs with no polymorphism and ten pairs with low amplification efficiency (< 75%) were discarded. The remained 24 primer pairs that generated polymorphic genotypes were used for population-level examination. Development of an appropriate set of markers is often the first step in population genetic and evolutionary studies. The recent development of genomic sequencing technology has made it relatively easy to isolate powerful microsatellites from large numbers of candidates at a genome-wide scale [42]. This method has been used in population structure analyses in many species, such as Grapholita molesta [30], Frankliniella occidentalis [43] and Carposina sasakii [29]. In our study, the 24 microsatellites developed are highly efficient in terms of amplification and polymorphism, enabling us to assess the population genetic structure of E. corollae.

Population genetic diversity

A total of 96 individuals with 24 individuals from each of the four populations was used for the genetic diversity study. The number of alleles per locus for all individuals ranged from 5
to 13 with an average of 8.75, which showed the level of polymorphism of the selected loci. The observed ($H_O$) and expected ($H_E$) heterozygosity values ranged from 0.235 to 0.768 and from 0.333 to 0.785, respectively. Four loci (S01, S07, S24, S39) showed a significant gap between observed and expected values, while the inbreeding coefficient ($F_{IS}=(H_E-H_O)/H_E$) calculated by GENEPOP for these loci was relatively high.

Significant deviations from HWE after sequential Bonferroni correction [44] ($P < 0.05$) were detected in 9 of 24 loci (S01, S02, S07, S11, S14, S24, S25, S33&S39), and 3 of the 24 loci (S07, S24 & S39) deviated in all populations. None of the loci were in linkage disequilibrium (LD) in the four populations.

**Population genetic structure**

Pairwise Fst analysis showed no significant differentiation between each pair of populations with $F_{ST}$ values ranging from -0.007 to 0.001 (Table 4). BAPS analysis showed all populations clustered into one group (Figure 1a) while STRUCTURE analysis showed an optimal value of $K=3$. All populations were evenly spread across the three clusters, indicating a lack of genetic differentiation among populations (Figure 1b). This pattern of genetic structure is congruent with an estimated pairwise Fst values among populations. The geographically related pairs of populations had relatively small Fst value while the distantly related pairs of populations had relatively larger Fst values (Table 4).

A lack of population differentiation is common in hoverflies. For example, a previous study on the hoverfly *Cheilosia naruska* from Finland showed that the species lacks differentiation at both the genetic and phenotypic levels [45]. Another study of two hoverfly species (*Episyrphus balteatus* and *Sphaerophoria scripta*) in Europe using 12 species-specific microsatellite markers also revealed a lack of genetic differentiation within species [18]. High levels of genetic diversity associated with a lack of structuring at a large spatial scale may indicate a high tolerance to environmental variability and a high migration rate [46]. Our study indicated that *E. corollae* in China may be highly mobile. The geographically related pattern of population structure may indicate that migration is restricted by geographical barriers. Our study provides preliminary insight into the biology and ecology of *E. corollae*. Further denser sampling is required to assess the population genetic structure of this species as well as other approaches to investigate its migration pattern.

Microsatellite markers are popular and powerful DNA markers because they are cost-effective and with a high diversity[41]. With the development of next-generation sequencing, genome-wide single nucleotide polymorphisms (SNPs) are becoming more powerful to screen genome-wide polymorphisms in a rapid and cost-effective manner [47]. Incorporating high-density SNPs in population genetic analysis may provide information on biology and ecology, such migration routes, of *E. corollae*, and help to understand adaptive evolution in this species [48].

**Conclusions**

We developed 24 microsatellite markers in *E. corollae* at a genome-wide scale which provides genetic markers for population genetic analyses of this species. Our preliminary
examination of four geographical populations of *E. corollae* across China suggested weak
but geographically lined population differentiation. The results provide insight into
migration of *E. corollae* in China.

**Acknowledgments**

We thank Prof. Ary Hoffmann from The University of Melbourne for his revisions and
suggestions on the manuscript, Xu-Bo Wang and Hua-Yan Chen for their help on the
collection of specimens, Yan-Jie Lv for her help on the molecular works. This research was
supported by the Applied Foundation of Science & Technology Department of Sichuan
Province (2018YYJC0468), the National Natural Science Foundation (31472025), the Natural
Science Foundation of Beijing Municipality (6162010), the Beijing Key Laboratory of
Environmentally Friendly Pest Management on Northern Fruits (BZ0432), all of China.

**Author Contributions**

Conceiving and design, Shu-Jun Wei and De-Qiang Pu; Data curation, L-Jun Cao, Ling Ma;
Formal analysis, Meng-Jia Liu, Ling Ma and Li-Jun Cao; Methodology, Li-Jun Cao and Shu-Jun
Wei; Resources, De-Qiang Pu, Ya-Jun Gong; Supervision, Shu-Jun Wei; Validation, Ya-Jun
Gong and Li-Jun Cao; Writing – original random, Meng-Jia Liu, Ling Ma; Writing – review &
editing, Meng-Jia Liu and Shu-Jun Wei.

**Conflicts of Interest**

The authors declare no conflict of interest.

**References**

1. Van Veen, M.P.; Moore, S.J., *Hoverflies of Northwest Europe: identification keys to
   the Syrphidae*. KNNV Publishing Utrecht: 2004.
2. Rojo, S.; Isidro, P.; Perez-Bañón, M.; Marcos-García, M. Revision of the hoverflies
   (Diptera: Syrphidae) from the Azores archipelago with notes on Macaronesian
   Syrphid fauna. *Life and Marine Sciences* **1997**, *15*, 65-82.
3. Rojo, S.; Hopper, K.R.; Marcos-García, M.A. Fitness of the hover flies *Episyrphus
   balteatus* and *Eupeodes corollae* faced with limited larval prey. *Entomologia
   Experimentalis et Applicata* **1996**, *81*, 53-59.
4. Mengual, X.; Stähls, G.; Rojo, S. Molecular phylogeny of Allograpta (Diptera,
   Syrphidae) reveals diversity of lineages and non-monophyly of phytophagous taxa.
   *Molecular Phylogenetics and Evolution* **2008**, *49*, 715-727.
5. Scott, S.; Barlow, C. Effect of hunger on the allocation of time among pea plants by
   the larvae of an aphidophagous hover fly, *Eupeodes corollae* [Dipt: Syrphidae].
   *Entomophaga* **1990**, *35*, 163-172.
6. Barbir, J.; Dorado, J.; Fernández-Quintanilla, C.; Blanusa, T.; Maksimovic, C.;
   Badenes-Pérez, F.R. Wild rocket – effect of water deficit on growth, flowering, and
   attractiveness to pollinators. *Acta Agriculturae Scandinavica, Section B — Soil &
   Plant Science* **2014**, *64*, 482-492.
7. Jauker, F.; Wolters, V. Hover flies are efficient pollinators of oilseed rape. *Oecologia
   2008*, *156*, 819-23.
8. Pu, D.q.; Shi, M.; Wu, Q.; Gao, M.q.; Liu, J.F.; Ren, S.p.; Yang, F.; Tang, P.; Ye, G.y.; Shen, Z.c. Flower--visiting insects and their potential impact on transgene flow in rice. *Journal of Applied Ecology* 2014, 51, 1357-1365.

9. Putra, N.S.; Yasuda, H. Effects of prey species and its density on larval performance of two species of hoverfly larvae, Episyrphus balteatus de Geer and Eupeodes corollae Fabricius (Diptera: Syrphidae). *Applied Entomology and Zoology* 2006, 41, 389-397.

10. Hu, G.; Lim, K.S.; Horvitz, N.; Clark, S.J.; Reynolds, D.R.; Sapir, N.; Chapman, J.W. Mass seasonal bioflows of high-flying insect migrants. *Science* 2016, 354, 1584-1587.

11. Raymond, L.; Vialatte, A.; Plantegenest, M. Combination of morphometric and isotopic tools for studying spring migration dynamics in *Episyrphus balteatus*. *Ecosphere* 2014, 5, 1-16.

12. Wei, S.J.; Shi, B.C.; Gong, Y.J.; Jin, G.H.; Chen, X.X.; Meng, X.F. Genetic structure and demographic history reveal migration of the diamondback moth *Plutella xylostella* (Lepidoptera: Plutellidae) from the southern to northern regions of China. *PLoS ONE* 2013, 8, e59654.

13. Liedvogel, M.; Akesson, S.; Bensch, S. The genetics of migration on the move. *Trends Ecol Evol* 2011, 26, 561-9.

14. Zhan, S.; Merlin, C.; Boore, J.L.; Reppert, S.M. The monarch butterfly genome yields insights into long-distance migration. *Cell* 2011, 147, 1171-85.

15. Zhan, S.; Zhang, W.; Niitepold, K.; Hsu, J.; Haeger, J.F.; Zalucki, M.P.; Altizer, S.; De Roode, J.C.; Reppert, S.M.; Kronforst, M.R. The genetics of monarch butterfly migration and warning coloration. *Nature* 2014, 514, 317.

16. Milankov, V.; Francuski, L.; Ludoski, J.; Ståhls, G.; Vujic, A. Genetic structure and phenotypic diversity of two northern populations of Cheilosia aff. longula (Diptera: Syrphidae) has implications for evolution and conservation. *European Journal of Entomology* 2010, 2013, 107, 305.

17. Rotheray, E.; Lepais, O.; Nater, A.; Krützen, M.; Greminger, M.; Goulson, D.; Bussière, L. Genetic variation and population decline of an endangered hoverfly Blera fallax (Diptera: Syrphidae). *Conservation Genetics* 2012, 13, 1283-1291.

18. Raymond, L.; Plantegenest, M.; Vialatte, A. Migration and dispersal may drive to high genetic variation and significant genetic mixing: the case of two agriculturally important, continental hoverflies (*Episyrphus balteatus* and *Sphaerophoria scripta*). *Molecular ecology* 2013, 22, 5329-5339.

19. Raymond, L.; Plantegenest, M.; Gauffre, B.; Sarthou, J.P.; Vialatte, A. Lack of genetic differentiation between contrasted overwintering strategies of a major pest predator Episyrphus balteatus (Diptera: Syrphidae): implications for biocontrol. *PLoS one* 2013, 8, e72997.

20. Odermatt, J.; Frommen, J.G.; Menz, M.H. Consistent behavioural differences between migratory and resident hoverflies. *Animal Behaviour* 2017, 127, 187-195.

21. Stubbs, A.E.; Falk, S.J., *British hoverflies: An illustrated identification guide*. British Entomological and Natural History Society; 2002.

22. Svensson, B.G.; Janzon, L.A. Why does the hoverfly *Metasyrphus corollae* migrate? *Ecological Entomology* 1984, 9, 329-335.

23. Speight, M.C. A mass migration of Episyrphus balteatus and Eupeodes corollae arriving in the south-west and remarks on other migrant hoverflies (Diptera: Syrphidae) in Ireland. *Irish Naturalists' Journal* 1996, 25, 182-183.
Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114-2120.

Andrews, S. FastQC: a quality control tool for high throughput sequence data. **2010**.

Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. In *IDBA – A practical iterative de bruijn graph De Novo assembler*, Berlin, Heidelberg, 2010; Springer Berlin Heidelberg: Berlin, Heidelberg, 2010; pp 426-440.

Du, L.; Li, Y.; Zhang, X.; Yue, B. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. *Journal of Heredity* **2013**, *104*, 154-157.

Meglécz, E.; Costedoat, C.; Dubut, V.; Gilles, A.; Malausa, T.; Pech, N.; Martin, J.-F. QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* **2010**, *26*, 403-404.

Wang, Y.Z.; Cao, L.J.; Zhu, J.Y.; Wei, S.J. Development and characterization of novel microsatellite markers for the peach fruit moth Carposina sasakii (Lepidoptera: Carposinidae) using next-generation sequencing. *International Journal of Molecular Sciences* **2016**, *17*, 362.

Song, W.; Cao, L.J.; Wang, Y.Z.; Li, B.Y.; Wei, S.J. Novel microsatellite markers for the oriental fruit moth Grapholita molesta (Lepidoptera: Tortricidae) and effects of null alleles on population genetics analyses. *Bulletin of Entomological Research* **2017**, *107*, 349-358.

Blacket, M.J.; Robin, C.; Good, R.T.; Lee, S.F.; Miller, A.D. Universal primers for fluorescent labelling of PCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. *Molecular ecology resources* **2012**, *12*, 456-463.

Raymond, M. GENEPOP : population genetics software for exact tests and ecumenism. Vers. 1.2. *J. Hered.* **1995**, *86*, 248-249.

Cao, L.J.; Wang, Z.H.; Gong, Y.J.; Zhu, L.; Hoffmann, A.A.; Wei, S.J. Low genetic diversity but strong population structure reflects multiple introductions of western flower thrips (Thysanoptera: Thripidae) into China followed by human-mediated spread. *Evolutionary Applications* **2017**, *10*, 391-401.

Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **2000**, *155*, 945-959.

Earl, D.A.; vonHoldt, B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **2012**, *4*, 359-361.

Jakobsson, M.; Rosenberg, N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801-1806.

Rosenberg, N.A. DISTRACT: a program for the graphical display of population structure. *Molecular Ecology Notes* **2004**, *4*, 137-138.

Bai, X.; Zhang, W.; Orantes, L.; Jun, T.-H.; Mittapalli, O.; Mian, M.A.R.; Michel, A.P. Combining next-generation sequencing strategies for rapid molecular resource development from an invasive aphid species, Aphis glycines. *PloS one* **2010**, *5*, e11370.

Yao, Y.; Zhao, W.; Shang, X. Development of polymorphic microsatellite markers of Obolodiplosis robiniae (Haldeman) (Diptera: Cecidomyiidae), a North American pest invading asia. *Journal of Insect Science* **2015**, *15*, 127-127.
40. Ellegren, H. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* **2004**, *5*, 435.

41. Selkoe, K.A.; Toonen, R.J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* **2006**, *9*, 615-629.

42. Queirós, J.; Godinho, R.; Lopes, S.; Gortazar, C.; De la Fuente, J.; Alves, P. Effect of microsatellite selection on individual and population genetic inferences: an empirical study using cross-specific and species-specific amplifications. *Molecular Ecology resources* **2015**, *15*, 747-760.

43. Cao, L.J.; Li, Z.M.; Wang, Z.H.; Zhu, L.; Gong, Y.J.; Chen, M.; Wei, S.J. Bulk development and stringent selection of microsatellite markers in the western flower thrips *Frankliniella occidentalis*. *Scientific Reports* **2016**, *6*, 26512.

44. Rice, W.R. Analyzing tables of statistical tests. *Evolution; International Journal of Organic Evolution* **1989**, *43*, 223-225.

45. Milankov, V.; Francuski, L.; Ludoški, J.; Ståhls, G.; Vujić, A. Estimating genetic and phenotypic diversity in a northern hoverfly reveals lack of heterozygosity correlated with significant fluctuating asymmetry of wing traits. *Journal of Insect Conservation* **2010**, *14*, 77-88.

46. Verhoeven, K.J.; Macel, M.; Wolfe, L.M.; Biere, A. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society of London B: Biological Sciences* **2011**, *278*, 2-8.

47. Behura, S.K. Molecular marker systems in insects: current trends and future avenues. *Molecular Ecology* **2006**, *15*, 3087-3113.

48. Ball, A.D.; Stapley, J.; Dawson, D.A.; Birkhead, T.R.; Burke, T.; Slate, J. A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (Taeniopygia guttata). *BMC Genomics* **2010**, *11*, 218.
### Tables

**Table 1** Collection information of *Eupeodes corollae* for microsatellite development and population genetic structure analysis

| Code  | Collection location                  | Collection location | Longitude (°E),   | Latitude(°N)   | Crop field | Collection date | Number |
|-------|-------------------------------------|---------------------|-------------------|---------------|------------|-----------------|--------|
| HNHK  | Haikou, Hainan Province             | Haikou, Hainan      | 110°27’20.034”    | 20°1’42.3444” | Rape       | 01/03/2017      | 24     |
| BJFS  | Fangshan, Beijing                   | Fangshan, Beijing   | 115°51’46.2096”   | 39°43’36.3108”| Weeds      | 08/06/2017      | 24     |
| YNYX  | Yuxi, Yunnan Province               | Yuxi, Yunnan        | 102°32’42.2448”   | 24°22’13.6092”| Rape       | 20/06/2017      | 24     |
| HLHB  | Harbin, Heilongjiang Province       | Harbin, Heilongjiang| 126°40’25.4136”   | 45°38’9.8952” | Watermelon | 28/07/2017      | 24     |
### Table 2 Twenty-four microsatellite loci developed for *Eupeodes corollae*

| Locus     | Motif | Forward primer | Reverse primer | Size(bp) | FL |
|-----------|-------|----------------|----------------|----------|----|
| EC7-S01   | (ACG)7| CCTATACATAACGGGCGG | CCCAGCGAAGGATGTCTCC | 103      | HEX |
| EC7-S02   | (ACG)7| CCCTCAACAGCCATTCGGAT | ACCACGCTGACCATGTTGAA | 115      | HEX |
| EC7-S03   | (AGC)8| GCTTCTAGAGCCTACTTGTT | CTCAGTATCTGGCGCTTTCC | 116      | HEX |
| EC7-S06   | (AGC)7| AGCTTCCAGTTCCAAAGCC | CCAGCGAACCAAAACACAG | 127      | HEX |
| EC7-S07   | (ATC)10| TACGCTCTGCTCCTGCTC | ACGGGAATCGACAGCAGCT | 130      | HEX |
| EC7-S08   | (ATC)10| TCAGTACGTACGGAAGG | GTGGTCCTGGGAAGCTGTC | 131      | HEX |
| EC7-S09   | (ATC)10| GCTGCCTATCACTTGGCCT | GTGGTCCTCACTGAGCTCG | 133      | HEX |
| EC7-S11   | (AAG)11| AGCGAAGAAACATGCACCG | GAAGGCTCTGGATGAGCGG | 150      | HEX |
| EC10-S13  | (AAG)8| CACAGAAGTTTTGCTGGA | GGTAAAGGTGATGTGGGCG | 158      | FAM |
| EC7-S14   | (ATC)9| AACACCGAACTCCAAACCG | TTTCAACATTCGGCTGCTG | 161      | FAM |
| EC10-S16  | (AAC)7| TGGAGGCAGCTGGATTGCCAGC | TTGGATGACAGCTGCTG | 180      | FAM |
| EC7-S17   | (AAC)12| CATTGGAAAGGCTGACAG | TGGAACTCCATGCGATCC | 186      | FAM |
| EC7-S18   | (AAC)7| TGCTTGACGATTACACGT | GTGTTGACGAGGATGGACG | 187      | FAM |
| EC7-S21   | (ACG)7| TGCAATGCTGACACCGAC | CGGAGTCCCAACTCTATGAC | 200      | FAM |
| EC7-S22   | (CCG)7| TGGTGTTGGAGGTGAAAAGT | GTTGTTGACATCGTGACACG | 203      | FAM |
| EC11-S23  | (ACG)7| CGTGGGTCTTGGCTTCTG | TGGACTTTGCTGCTACACG | 204      | FAM |
| EC7-S24   | (ACC)7| GTGGTGCTCATGCTACAGG | TGGAGTCATTCCGAGCG | 212      | FAM |
| EC7-S25   | (ATC)7| CGCAGACATCATCATTG | TGGAGTGTGACATCCGAC | 215      | ROX |
| EC12-S26  | (AGC)7| GTTGGTGGTGTTGGAGG | GTGGTGGTGGAGGA | 220      | ROX |
| EC10-S29  | (ACG)11| CATGAAACCATCGCGTCTC | ATACCCCTGATCCAGCGG | 225      | ROX |
| EC33-S31  | (ATC)33| TAACCTGGTGCGCTAGTCTT | GTGGTGGCAGACATGCTG | 259      | ROX |
| EC13-S33  | (AAAG)13| AGGGCACTTTGAATCCACCG | TGACTCCGAATGTGGCCTAG | 285      | ROX |
| EC7-S36   | (AGAT)24| TGGGCTCAAGTGTGAAACACGA | AACAGCTTTGGCCCTACCGA | 310      | ROX |
| EC20-S39  | (ATC)8| CCATCGCGAACTGTTTCCTCT | TGCTGCTATGTGCTCGTC | 324      | ROX |
| Locus    | Allele | F_{IS} | HLHB | HNHK | YNX | BJFS | HLHB | HNHK | YNX | BJFS | HLHB | HNHK | YNX | BJFS | HLHB | HNHK | YNX |
|----------|--------|--------|------|------|-----|------|------|------|-----|------|------|------|-----|------|------|------|-----|
| EC7-S01  | 9      | 0.64   | 0.49 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 | 0.30 |
| EC7-S02  | 6      | 0.12   | 0.29 | 0.52 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 | 0.21 |
| EC7-S03  | 10     | -0.01  | -0.08 | -0.06 | -0.16 | 0.10 | 0.79 | 0.71 | 0.82 | 0.47 | 0.54 | 0.51 | 0.61 | 0.48 | 0.58 | 0.54 | 0.71 |
| EC7-S06  | 5      | 0.01   | -0.21 | -0.06 | 0.14 | 0.65 | 0.75 | 0.84 | 0.49 | 0.62 | 0.55 | 0.59 | 0.53 | 0.61 | 0.67 | 0.63 | 0.46 |
| EC7-S07  | 7      | 0.47   | 0.42 | 0.45 | 0.40 | 0.01 | 0.01 | 0.01 | 0.00 | 0.62 | 0.64 | 0.68 | 0.62 | 0.33 | 0.38 | 0.38 | 0.38 |
| EC7-S08  | 5      | -0.08  | 0.15 | 0.02 | 0.09 | 0.90 | 0.34 | 1.00 | 0.08 | 0.54 | 0.49 | 0.38 | 0.55 | 0.58 | 0.42 | 0.38 | 0.50 |
| EC7-S09  | 8      | 0.02   | -0.11 | 0.16 | 0.09 | 0.35 | 0.86 | 0.25 | 0.80 | 0.73 | 0.67 | 0.69 | 0.64 | 0.71 | 0.75 | 0.58 | 0.58 |
| EC7-S11  | 6      | 0.47   | 0.52 | -0.06 | -0.05 | 0.01 | 0.00 | 1.00 | 1.00 | 0.47 | 0.51 | 0.24 | 0.18 | 0.25 | 0.25 | 0.25 | 0.19 |
| EC10-S13 | 10     | 0.16   | 0.08 | 0.19 | -0.13 | 0.29 | 0.17 | 0.13 | 0.98 | 0.79 | 0.77 | 0.77 | 0.81 | 0.67 | 0.71 | 0.63 | 0.91 |
| EC7-S14  | 6      | 0.38   | -0.02 | 0.12 | -0.09 | 0.01 | 0.19 | 0.04 | 0.84 | 0.34 | 0.45 | 0.56 | 0.42 | 0.21 | 0.46 | 0.50 | 0.46 |
| EC10-S16 | 10     | -0.03  | -0.10 | 0.06 | -0.05 | 0.35 | 0.68 | 0.15 | 0.72 | 0.72 | 0.83 | 0.62 | 0.79 | 0.74 | 0.92 | 0.58 | 0.83 |
| EC7-S17  | 7      | 0.00   | -0.18 | 0.15 | -0.03 | 0.48 | 0.78 | 0.55 | 0.96 | 0.58 | 0.64 | 0.64 | 0.59 | 0.58 | 0.75 | 0.54 | 0.61 |
| EC7-S18  | 6      | -0.18  | -0.13 | 0.12 | -0.12 | 1.00 | 1.00 | 0.61 | 1.00 | 0.41 | 0.26 | 0.33 | 0.34 | 0.48 | 0.29 | 0.29 | 0.38 |
| EC7-S21  | 11     | 0.07   | -0.01 | 0.08 | 0.06 | 0.58 | 0.01 | 0.44 | 0.14 | 0.67 | 0.74 | 0.77 | 0.70 | 0.63 | 0.75 | 0.71 | 0.67 |
| EC7-S22  | 13     | 0.05   | -0.03 | 0.03 | -0.02 | 0.91 | 0.59 | 0.76 | 0.44 | 0.79 | 0.81 | 0.73 | 0.70 | 0.75 | 0.83 | 0.71 | 0.71 |
| EC11-S23 | 8      | 0.13   | 0.18 | 0.01 | 0.02 | 0.58 | 0.11 | 0.63 | 0.92 | 0.76 | 0.76 | 0.72 | 0.77 | 0.67 | 0.63 | 0.71 | 0.75 |
| EC7-S24  | 13     | 0.34   | 0.37 | 0.62 | 0.33 | 0.04 | 0.00 | 0.00 | 0.00 | 0.63 | 0.72 | 0.66 | 0.73 | 0.42 | 0.46 | 0.25 | 0.50 |
| EC7-S25  | 8      | 0.05   | 0.16 | 0.28 | 0.08 | 0.88 | 0.03 | 0.03 | 0.74 | 0.66 | 0.69 | 0.64 | 0.67 | 0.63 | 0.58 | 0.46 | 0.62 |
| EC12-S26 | 8      | 0.27   | 0.15 | 0.08 | 0.10 | 0.16 | 0.18 | 0.43 | 0.95 | 0.79 | 0.73 | 0.72 | 0.83 | 0.58 | 0.63 | 0.67 | 0.75 |
| EC10-S29 | 6      | -0.14  | 0.26 | -0.10 | 0.10 | 1.00 | 0.11 | 0.54 | 1.00 | 0.37 | 0.34 | 0.53 | 0.28 | 0.42 | 0.25 | 0.58 | 0.30 |
| EC33-S31 | 5      | 0.02   | 0.25 | 0.02 | 0.05 | 0.53 | 0.07 | 0.43 | 0.84 | 0.43 | 0.33 | 0.42 | 0.44 | 0.42 | 0.25 | 0.42 | 0.42 |
| EC13-S33 | 8      | 0.14   | -0.01 | 0.05 | -0.11 | 0.60 | 0.22 | 0.42 | 0.04 | 0.53 | 0.43 | 0.44 | 0.60 | 0.46 | 0.43 | 0.42 | 0.67 |
| EC7-S36  | 10     | 0.00   | 0.18 | 0.13 | 0.06 | 0.30 | 0.56 | 0.82 | 0.17 | 0.75 | 0.76 | 0.71 | 0.74 | 0.75 | 0.63 | 0.63 | 0.70 |
| EC20-S39 | 14     | 0.52   | 0.40 | 0.66 | 0.60 | 0.00 | 0.02 | 0.00 | 0.00 | 0.60 | 0.63 | 0.54 | 0.52 | 0.29 | 0.38 | 0.19 | 0.21 |
| All      | 8.75   | 0.14   | 0.12 | 0.17 | 0.06 | 0.59 | 0.60 | 0.58 | 0.58 | 0.50 | 0.53 | 0.48 | 0.55 |
Table 4 Pairwise Fst of 4 *Eupeodes corollae* populations based on 24 microsatellites

| Population | BJFS | HLHB | HNHK | YNYX |
|------------|------|------|------|------|
| BJFS       | —    | 0.901| 0.306| 0.892|
| HLHB       | -0.004 | —    | 0.838| 0.973|
| HNHK       | 0.005 | -0.001| —    | 0.468|
| YNYX       | -0.004 | -0.007| 0.001| —    |

The bottom triangle shows the pairwise Fst values, while the upper triangle shows the corresponding P values.
**Figures and figure legends**

**Figure 1** Collection sites of *Eupeodes corollae* (a) and population genetic structure analysis of four geographical populations using BAPS (a) and STRUCTURE (b). BAPS analysis showed that all population are clustered into one cluster (blue color in figure a). STRUCTURE analysis showed that the optimal delta K was three and all populations were composed of the three clusters. Codes for the population are shown in Table 1.
