PLOS ONE

RESEARCH ARTICLE

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

Mengjia Liu[1,2☯], Xiaoqiang Wang[1☯], Ling Ma[2], Lijun Cao[2], Hongling Liu[3], Deqiang Pu[1,3]*, Shujun Wei[2]*

1 Industrial Crop Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu, China, **2** Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, **3** Institute of Plant Protection, Sichuan Academy of Agricultural Sciences, Chengdu, China

☯ These authors contributed equally to this work.
* pdqpudeqiang@163.com (DP); shujun268@163.com (SW)

Citation: Liu M, Wang X, Ma L, Cao L, Liu H, Pu D, et al. (2019) Genome-wide developed microsatellites reveal a weak population differentiation in the hoverfly *Eupeodes corollae* (Diptera: Syrphidae) across China. PLoS ONE 14 (9): e0215888. https://doi.org/10.1371/journal.pone.0215888

Editor: Maohua Chen, Northwest A&F University, CHINA

Received: April 9, 2019

Accepted: August 22, 2019

Published: September 26, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: Financial support was provided by the Applied Foundation of Science & Technology Department of Sichuan Province (2018YYJC0468), the Natural Science Foundation of Beijing Municipality (6184037), the Beijing Key Laboratory of Environmentally Friendly Pest Management on Northern Fruits (BZ0432) and BAAFS-MOU Pest

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.

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

Control Research Joint Laboratory. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

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. *Eupeodes 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, Fig 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 paired-end 150 bp 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]. The genome size of P. solenopsis was estimated by JELLYFISH v2.2.6 software with a K-mer method [26]. IDBA was used to assemble the generated genomic sequences with K-mer from 20 to 140 [27].

Genome-wide microsatellite survey and primer design

Microsatellite markers were developed from genome sequences as in previous publications [28–30]. 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* [31]. QDD was used to isolate microsatellites and design primers [32]. The outputs of primer pairs from QDD were further filtered by the following criteria [33, 34]: (i) the corresponding microsatellites were pure and specific; (ii) the

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

Code	Collection location	Longitude (˚E),	Latitude(˚N)	Crop field	Collection date	Number
HNHK	Haikou, Hainan Province	110˚27′20.034″	20˚1′42.3444″	Rape	01/03/2017	24
BJFS	Fangshan, Beijing	115˚51′46.2096″	39˚43′36.3108″	Weeds	08/06/2017	24
YNYX	Yuxi, Yunnan Province	102˚32′42.2448″	24˚22′13.6092″	Rape	20/06/2017	24
HLHB	Harbin, Heilongjiang Province	126˚40′25.4136″	45˚38′9.8952″	Watermelon	28/07/2017	24

https://doi.org/10.1371/journal.pone.0215888.t001

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 (CAGGACCAGGCTAC CGTG) was added to the 5′ end of each selected forward primer to allow efficient combining with the fluorescent label [35]. Amplifications were performed using the GoTaq Green Master

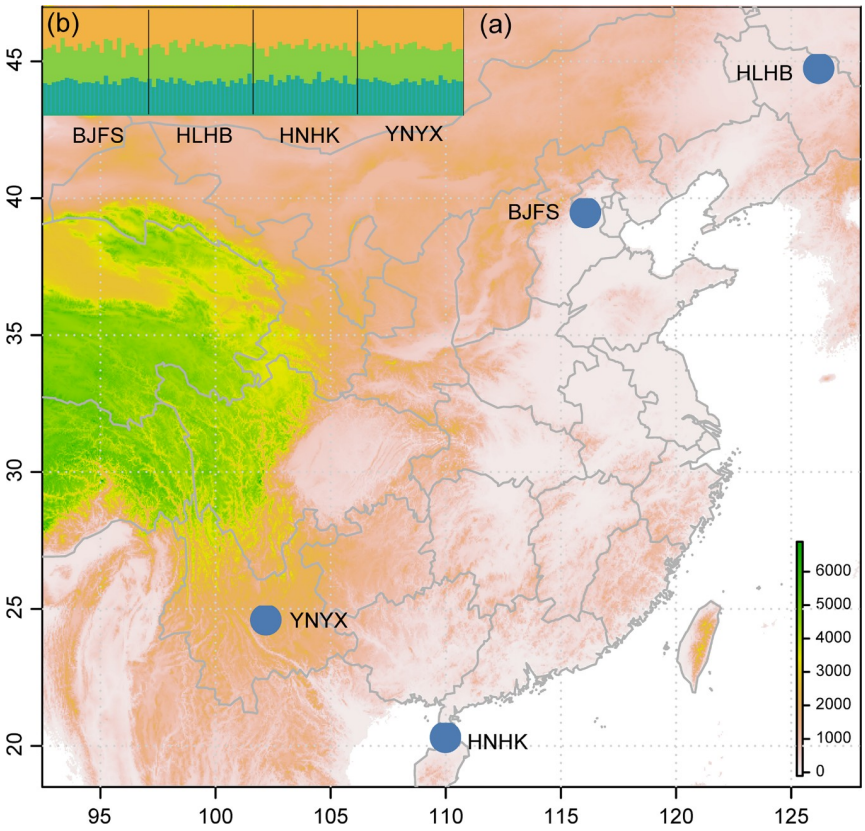

Fig 1. Collection sites of *Eupeodes corollae* (a) and population genetic structure analysis of four geographical populations using BAPS (a) and STRUCTURE (b). The map was drawn in R function *map_data*. BAPS analysis showed that all population are clustered into one cluster (blue color in figure a). 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.

https://doi.org/10.1371/journal.pone.0215888.g001

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 ddH_2O, 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 [36] 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 (F_{IS}) and pairwise population differentiation (F_{ST}). Allele frequencies, expected heterozygosity (H_E) and observed heterozygosity (H_O) were calculated with the macros Microsatellite Tools [37].

Population genetic structure was analyzed by STRUCTURE version 2.3.4 [38]. 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 [39]. Visualization of the results was handled by CLUMPP version 1.1.2 [40] and DISTRUCT version 1.1 [41]. 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*

The genomic size of *E. corollae* was estimated to be 12,315 Mb. A total of 51.53 Gb paired-end (PE) sequences (184,394,506 reads each with a length of 150 bp) was obtained. Trimmed reads were assembled into 2,563,327 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*

In total 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* [34] (Lepidoptera), *Aphis glycines* (Hemiptera) [42] and *Obolodiplosis robiniae* (Diptera) [43], which shows the distribution of microsatellites to vary among species [44, 45].

Development of variable microsatellite markers

The QDD program initially generated 18114 primer pairs (S1 Table); we selected those corresponding to tri- and tetra-nucleotide microsatellites for further filtering under criteria listed in the methods and obtained 40 primer pairs (S2 Table). These primer pairs were validated in

Table 2. Twenty-four microsatellite loci developed for *Eupeodes corollae*.

Locus	Motif	Forward primer	Reverse primer	Size(bp)	FL
EC7-S01	(ACG)7	CCTATACATAACGGGCCGGG	CCCAGCGAAGGATGTTCTCC	103	HEX
EC7-S02	(ACG)7	CCCTCAACAGCCATTCCGAT	ACCAGCGTGACCATGTTGAA	115	HEX
EC7-S03	(AGC)8	GCCTTGCAGAGCCTACTGTT	CTCAGTAGTCTGGCGCTTCC	116	HEX
EC7-S06	(AGC)7	AGCTTCCCAGTTCCAAAGCC	CCAGCGAACCAACAAACCAG	127	HEX
EC7-S07	(ATC)10	TACGCCTCTGTCTTTGCCTC	AACGGGAATCGACAAGCACT	130	HEX
EC7-S08	(ATC)10	TCAGTAACGTCACGAAGGGC	GTGGTCCTGGAAGCTGTCTC	131	HEX
EC7-S09	(ATC)10	GCTGCCTTATCACTTGCCCT	TGTGGTCCAACTGAGTGTCG	133	HEX
EC7-S11	(AAG)11	AGCGAAAGAACAATGCCACG	GAAGGTCTCTGGATGGACGG	150	HEX
EC10-S13	(AAG)8	CACACGAACTTCTGGCTGGA	GGGTAAGGTGTAGTGTGGGC	158	FAM
EC7-S14	(ATC)9	AACACCCGAACTCCAAACCG	TTTCAACATTCGCGTCGCTG	161	FAM
EC10-S16	(AAC)7	TGGAGCGAGCTGGATTGATC	TTCGAGTGATGAGCCTGTGG	180	FAM
EC7-S17	(AAC)12	CATTGGAAAGGCTGCAACGG	TGGAACTCCATGGCATTCCG	186	FAM
EC7-S18	(AAC)7	TGCCTTGACGATTACCACGT	GATGGTGACGGATTGCGACT	187	FAM
EC7-S21	(ACG)7	TGCATGGATGGACACCAGAC	GCGATGCCAACCTCATGTAC	200	FAM
EC7-S22	(CCG)7	TGGTGTGGAGGGTGGAAATG	GTTTGTGCATCCGTGAACGA	203	FAM
EC11-S23	(ACG)7	CTGAGGGCTTGCTTCATGTG	TGGACTTTCGTGTACCAGCC	204	FAM
EC7-S24	(ACC)7	GTCGTCCTCATCGTCACAGG	TCATTGATTCGGCAGCAGGT	212	FAM
EC7-S25	(ATC)7	CGCACAGCATCACATCCATG	TAAGTGCGAGTACGGGCATT	215	ROX
EC12-S26	(AGC)7	GGTAGTGGCATCAGTGGAGG	GTTGGTGGTTGGGATGCAAA	220	ROX
EC10-S29	(ACG)11	CATGAACCCATCAGCGTCCT	ATACCCTGATCCAGCCCGAT	225	ROX
EC33-S31	(ATC)33	TAACTGGGTGGCATCGGTTC	GTTTGTGCGACTTGTGAGCT	259	ROX
EC13-S33	(AAAG)13	AGGGCAGCTATTGAATCCCG	TGACTCCGAATGTGCTCAGG	285	ROX
EC7-S36	(AGAT)24	TGGGCTCAAGTGTAAACGGA	AACAGCTTTGCCCTACCGAA	310	ROX
EC20-S39	(ATC)8	CCATCGCGAACTGTTCCTCT	TGCTGCTATGTCTCCGTGTT	324	ROX

FL, fluorescent label.

https://doi.org/10.1371/journal.pone.0215888.t002

eight individuals of *E. corollae*; six pairs with no polymorphism (S12, S15, S30, S32, S34, S40) and ten pairs (S30, S35, S32, S04, S10, S19, S20, S27, S37, S38) 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 [46]. This method has been used in population structure analyses in many species, such as *Grapholita molesta* [34], *Frankliniella occidentalis* [47] and *Carposina sasakii* [33]. 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* (Table 2).

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 five 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

Table 3. Summary statistics of 24 microsatellite markers for *Eupeodes corollae* validated in four populations. F_{IS}, inbreeding coefficient; He, expected heterozygosity; Ho, observed heterozygosity; HWE, average P-value of Hardy–Weinberg equilibrium.

Locus	Allele	F_{IS}				HWE				H_E				H_O			
		BJFS	HLHB	HNHK	YNYX	BJFS	HLHB	HNHK	YNYX	BJFS	HLHB	HNHK	YNYX	BJFS	HLHB	HNHK	YNYX
EC7-S01	9	0.64	0.49	0.51	0.30	0.00	0.00	0.00	0.09	0.46	0.72	0.61	0.55	0.17	0.38	0.30	0.39
EC7-S02	6	0.12	0.29	0.52	-0.21	0.22	0.23	0.01	0.63	0.33	0.35	0.34	0.34	0.29	0.25	0.17	0.42
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	0.02	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

https://doi.org/10.1371/journal.pone.0215888.t003

observed and expected values, while the inbreeding coefficient ($F_{IS} = (H_E\text{-}H_O)/H_E$) calculated by GENEPOP for these loci was relatively high (Table 3).

Significant deviations from HWE after sequential Bonferroni correction [48] ($P < 0.05$) were detected in 9 of 24 loci (S01, S02, S07, S11, S14, S24, S25, S33&S39), and three 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 F_{ST} 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

Table 4. Pairwise F_{ST} 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 F_{ST} values, while the upper triangle shows the corresponding P values. See Table 1 for population code.

https://doi.org/10.1371/journal.pone.0215888.t004

clustered into one group (Fig 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 (Fig 1B). This pattern of genetic structure is congruent with an estimated pairwise F_{ST} values among populations. Pairs of nearby populations had relatively small F_{ST} value while pairs of populations with large geographical distance had relatively larger F_{ST} 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 [49]. 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 [50]. 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 [45]. 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 [51]. 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 [52].

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.

Supporting information

S1 Table. All primers pairs design in the study.
(XLSX)

S2 Table. Forty primer pairs used for initial evaluation.
(XLSX)

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.

Author Contributions

Conceptualization: Deqiang Pu, Shujun Wei.

Data curation: Ling Ma, Lijun Cao.

Formal analysis: Mengjia Liu, Xiaoqiang Wang, Ling Ma, Lijun Cao.

Funding acquisition: Deqiang Pu, Shujun Wei.

Investigation: Xiaoqiang Wang, Ling Ma, Lijun Cao, Hongling Liu, Deqiang Pu.

Methodology: Lijun Cao, Shujun Wei.

Project administration: Deqiang Pu.

Resources: Xiaoqiang Wang, Deqiang Pu.

Software: Lijun Cao.

Supervision: Shujun Wei.

Validation: Lijun Cao.

Writing – original draft: Mengjia Liu, Xiaoqiang Wang, Ling Ma.

Writing – review & editing: Mengjia Liu, Shujun Wei.

References

1. Van Veen MP, Moore SJ. 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. ARQUIPÉLAGO Ciências Biológicas e Marinhas = Life and Marine Sciences. 1997; 15:65–82.
3. Rojo S, Hopper KR, Marcos-García MA. Fitness of the hover flies Episyrphus balteatus and Eupeodes corollae faced with limited larval prey. Entomologia Experimentalis et Applicata. 1996; 81(1):53–9.
4. Mengual X, Ståhls 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 (3):715–27. https://doi.org/10.1016/j.ympev.2008.09.011 PMID: 18848633
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(2):163–72.
6. Barbir J, Dorado J, Fernández-Quintanilla C, Blanusa T, Maksimovic C, Badenes-Pérez FR. Wild rocket–effect of water deficit on growth, flowering, and attractiveness to pollinators. Acta Agriculturae Scandinavica, Section B—Soil & Plant Science. 2014; 64(6):482–92. https://doi.org/10.1080/09064710.2014.925575
7. Jauker F, Wolters V. Hover flies are efficient pollinators of oilseed rape. Oecologia. 2008; 156(4):819–23. https://doi.org/10.1007/s00442-008-1034-x PMID: 18438687.
8. Dq Pu, Shi M, Wu Q, Gao Mq, Liu JF, Ren Sp, et al. Flower-visiting insects and their potential impact on transgene flow in rice. Journal of Applied Ecology. 2014; 51(5):1357–65.
9. Putra NS, 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(3):389–97.
10. Hu G, Lim KS, Horvitz N, Clark SJ, Reynolds DR, Sapir N, et al. Mass seasonal bioflows of high-flying insect migrants. Science. 2016; 354(6319):1584–7. https://doi.org/10.1126/science.aah4379 PMID: 28008067
11. Raymond L, Vialatte A, Plantegenest M. Combination of morphometric and isotopic tools for studying spring migration dynamics in Episyrphus balteatus. Ecosphere. 2014; 5(7):1–16.
12. Wei SJ, Shi BC, Gong YJ, Jin GH, Chen XX, Meng XF. 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(4):e59654. https://doi.org/10.1371/journal.pone.0059654 PMID: 23565158; PubMed Central PMCID: PMC3614937.
13. Liedvogel M, Akesson S, Bensch S. The genetics of migration on the move. Trends Ecol Evol. 2011; 26 (11):561–9. https://doi.org/10.1016/j.tree.2011.07.009 PMID: 21862171.
14. Zhan S, Merlin C, Boore JL, Reppert SM. The monarch butterfly genome yields insights into long-distance migration. Cell. 2011; 147(5):1171–85. https://doi.org/10.1016/j.cell.2011.09.052 PMID: 22118469; PubMed Central PMCID: PMC3225893.

15. Zhan S, Zhang W, Niitepold K, Hsu J, Haeger JF, Zalucki MP, et al. The genetics of monarch butterfly migration and warning coloration. Nature. 2014; 514(7522):317. https://doi.org/10.1038/nature13812 PMID: 25274300
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; 107(3):305.
17. Rotheray E, Lepais O, Nater A, Krützen M, Greminger M, Goulson D, et al. Genetic variation and population decline of an endangered hoverfly Blera fallax (Diptera: Syrphidae). Conservation genetics. 2012; 13(5):1283–91.
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 (E pisyrphus balteatus and S phaerophoria scripta). Molecular ecology. 2013; 22(21):5329–39. https://doi.org/10.1111/mec.12483 PMID: 24138027
19. Raymond L, Plantegenest M, Gauffre B, Sarthou JP, 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(9):e72997. https://doi.org/10.1371/journal.pone.0072997 PMID: 24023799; PubMed Central PMCID: PMC3759392.
20. Odermatt J, Frommen JG, Menz MH. Consistent behavioural differences between migratory and resident hoverflies. Animal Behaviour. 2017; 127:187–95.
21. Stubbs AE, Falk SJ. British hoverflies: An illustrated identification guide: British Entomological and Natural History Society; 2002.
22. Svensson BG, Janzon LÅ. Why does the hoverfly Metasyrphus corollae migrate? Ecological entomology. 1984; 9(3):329–35.
23. Speight MC. A mass migration of Episyrphus balteatus and Eupeodes corollae arriving in the southwest and remarks on other migrant hoverflies (Diptera: Syrphidae) in Ireland. Irish Naturalists' Journal. 1996; 25(5):182–3.
24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404
25. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.
26. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011; 27(6):764. https://doi.org/10.1093/bioinformatics/btr011 PMID: 21217122
27. Peng Y, Leung HCM, Yiu SM, Chin FYL, editors. IDBA–A practical iterative de bruijn graph De Novo assembler2010; Berlin, Heidelberg: Springer Berlin Heidelberg.
28. Castoe TA, Poole AW, Gu W, Jason de Koning A, Daza JM, Smith EN, et al. Rapid identification of thousands of copperhead snake (Agkistrodon contortrix) microsatellite loci from modest amounts of 454 shotgun genome sequence. Molecular Ecology Resources. 2010; 10(2):341–7. https://doi.org/10.1111/j.1755-0998.2009.02750.x PMID: 21565030
29. Gardner MG, Fitch AJ, Bertozzi T, Lowe AJ. Rise of the machines–recommendations for ecologists when using next generation sequencing for microsatellite development. Molecular Ecology Resources. 2011; 11(6):1093–101. https://doi.org/10.1111/j.1755-0998.2011.03037.x PMID: 21679314
30. Abdelkrim J, Robertson BC, Stanton J-AL, Gemmell NJ. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. BioTechniques. 2009; 46(3):185–92. https://doi.org/10.2144/000113084 PMID: 19317661
31. 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(1):154–7. https://doi.org/10.1093/jhered/ess082 PMID: 23144492
32. Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, et al. QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics. 2010; 26(3):403–4. https://doi.org/10.1093/bioinformatics/btp670 PMID: 20007741
33. Wang YZ, Cao LJ, Zhu JY, Wei SJ. 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(3):362. https://doi.org/10.3390/ijms17030362 PMID: 26999103
34. Song W, Cao LJ, Wang YZ, Li BY, Wei SJ. 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(3):349–58. Epub 2016/11/08. https://doi.org/10.1017/S0007485316000936 PMID: 27819214.

35. Blacket MJ, Robin C, Good RT, Lee SF, Miller AD. Universal primers for fluorescent labelling of PCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. Molecular ecology resources. 2012; 12(3):456–63. https://doi.org/10.1111/j.1755-0998.2011.03104.x PMID: 22268566

36. Raymond M. GENEPOP: population genetics software for exact tests and ecumenism. Vers. 1.2. J Hered. 1995; 86:248–9.

37. Cao LJ, Wang ZH, Gong YJ, Zhu L, Hoffmann AA, Wei SJ. 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(4):391–401. https://doi.org/10.1111/eva.12461 PMID: 28352298

38. Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 2000; 155(2):945–59. PMID: 10835412

39. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 2012; 4 (2):359–61. https://doi.org/10.1007/s12686-011-9548-7

40. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007; 23 (14):1801–6. https://doi.org/10.1093/bioinformatics/btm233 PMID: 17485429

41. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes. 2004; 4(1):137–8. https://doi.org/10.1046/j.1471-8286.2003.00566.x

42. Bai X, Zhang W, Orantes L, Jun T-H, Mittapalli O, Mian MAR, et al. Combining next-generation sequencing strategies for rapid molecular resource development from an invasive aphid species, Aphis glycines. PloS one. 2010; 5(6):e11370. https://doi.org/10.1371/journal.pone.0011370 PMID: 20614011

43. 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(1):127-. https://doi.org/10.1093/jisesa/iev104 PMID: 26386040

44. Ellegren H. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics. 2004; 5:435. https://doi.org/10.1038/nrg1348 PMID: 15153996

45. Selkoe KA, Toonen RJ. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology letters. 2006; 9(5):615–29. https://doi.org/10.1111/j.1461-0248.2006.00889.x PMID: 16643306

46. 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(4):747–60. https://doi.org/10.1111/1755-0998.12349 PMID: 25403329

47. Cao LJ, Li ZM, Wang ZH, Zhu L, Gong YJ, Chen M, et al. Bulk development and stringent selection of microsatellite markers in the western flower thrips *Frankliniella occidentalis*. Scientific Reports. 2016; 6:26512. https://doi.org/10.1038/srep26512 PMID: 27197749

48. Rice WR. Analyzing tables of statistical tests. Evolution; international journal of organic evolution. 1989; 43(1):223–5. https://doi.org/10.1111/j.1558-5646.1989.tb04220.x PMID: 28568501

49. 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(1):77–88. https://doi.org/10.1007/s10841-009-9226-1

50. Verhoeven KJ, Macel M, Wolfe LM, 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(1702):2–8.

51. Behura SK. Molecular marker systems in insects: current trends and future avenues. Molecular ecology. 2006; 15(11):3087–113. https://doi.org/10.1111/j.1365-294X.2006.03014.x PMID: 16968257

52. Ball AD, Stapley J, Dawson DA, Birkhead TR, 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(1):218.