Interspecific interactions influence contrasting spatial genetic structures in two closely related damselfly species

Kahilainen, Aapo; Keränen, Inka; Kuitunen, Katja; Kotiaho, Janne Sakari; Knott, Emily

2014

Please cite the original version:

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
Interspecific interactions influence contrasting spatial genetic structures in two closely related damselfly species

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Molecular Ecology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>MEC-14-0691.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Original Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Kahilainen, Aapo; University of Jyväskylä, Department of Biological and Environmental Science  
  Keränen, Inka; University of Jyväskylä, Department of Biological and Environmental Science  
  Kuitunen, Katja; University of Jyväskylä, Department of Biological and Environmental Science; University of Jyväskylä, Department of Mathematical Information Technology  
  Kotiaho, Janne; University of Jyväskylä, Department of Biological and Environmental Science; University of Jyväskylä, Natural History Museum  
  Knott, Karelyn; University of Jyväskylä, Department of Biological and Environmental Science |
| Keywords:         | Calopteryx splendens, Calopteryx virgo, population genetics, Landscape Genetics, microsatellite |
Title: Interspecific interactions influence contrasting spatial genetic structures in two closely related damselfly species

Authors: Aapo Kahilainen¹, Inka Keränen¹, Katja Kuitunen¹,², Janne S. Kotiaho¹,³ & K. Emily Knott¹

Affiliations: ¹University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014, University of Jyväskylä, Finland

²University of Jyväskylä, Department of Mathematical Information Technology, P.O. Box 35, FI-40014 University of Jyväskylä, Finland.

³University of Jyväskylä, Natural History Museum, P.O. Box 35 FI-40014, University of Jyväskylä, Finland

Keywords: Calopteryx splendens, Calopteryx virgo, population genetics, landscape genetics, microsatellite

Correspondence: Aapo Kahilainen
University of Jyväskylä, Department of Biological and Environmental Science,
P.O. Box 35, FI-40014, University of Jyväskylä, Finland
Tel: +358 40 8053875; Fax: +358 14 617 239
E-mail: aapo.kahilainen@jyu.fi

Running title: Contrasting genetic structures in Calopteryx
Spatial genetic structure (SGS) is largely determined by colonization history, landscape and ecological characteristics of the species. Therefore, sympatric and ecologically similar species are expected to exhibit similar SGSs, potentially enabling prediction of the SGS of one species from that of another. On the other hand, due to interspecific interactions, ecologically similar species could have different SGSs. We explored the SGSs of the closely related *Calopteryx splendens* and *C. virgo* within Finland and related the genetic patterns to characteristics of the sampling localities. We observed different SGSs for the two species. Genetic differentiation even within short distances in *C. splendens* suggests genetic drift as an important driver. However, we also observed indication of previous gene flow (revealed by a negative relationship between genetic differentiation and increasing potential connectivity of the landscape). Interestingly, genetic diversity of *C. splendens* was negatively related to density of *C. virgo*, suggesting that interspecific interactions influence the SGS of *C. splendens*. In contrast, genetic differentiation between *C. virgo* sub-populations was low and only exhibited relationships with latitude, pointing to high gene flow, colonization history and range margin effects as the drivers of SGS. The different SGSs of the two ecologically similar species cautions indirect inferences of SGS based on ecologically similar surrogate species.
**INTRODUCTION**

The spatial distribution of genetic diversity is rarely homogenous, but exhibits pronounced geographic variation (Eckert *et al.* 2008). The distribution of genetic diversity within and between populations, i.e. spatial genetic structure (SGS), is driven by historic contingencies such as range contractions and expansions due to glacial cycles (Hewitt 2000) and contemporary processes of random genetic drift (Kimura & Crow 1964), gene flow between populations (Slatkin 1985; Bohonak 1999), and selection (Ranta *et al.* 2009; Funk *et al.* 2011). Since these drivers of SGS are strongly influenced by the ecology of the species (e.g. habitat preferences, dispersal abilities, mating system and life-history characteristics) and by environmental characteristics of the landscape (e.g. the amount of habitat suitable for population persistence and dispersal), sympatric and ecologically similar species can be expected to exhibit similar SGSs (Whiteley *et al.* 2006; Fortuna *et al.* 2009; Dawson 2012; Hughes *et al.* 2013).

On the other hand, increasing ecological similarity of species can lead to increasing interspecific interactions, and thus affect their SGSs. For example, in a simulation, increasing the strength of interspecific competition and the number of competing species decreased both effective population sizes and dispersal between localities, leading to increased genetic drift and differentiation between populations (Ranta *et al.* 2009). Similarly, increasing the number of competitor species led to decreased population sizes and increased genetic drift of experimental perennial ryegrass (*Lolium perenne*) populations (Nestmann *et al.* 2011).

Given that the SGS of species is important for determining management units and predicting the outcomes of land use on the viability and evolutionary potential of populations (Frankham
the generalizability of SGS across sympatric and ecologically similar species is important also for applied conservation projects. If SGSs can be generalized across species, SGSs can be predicted for other sympatric and ecologically similar species based on that of only one (or a few) representative species. This implies that conservation strategies designed for safeguarding the viability and evolutionary potential of a single species could be effective in conserving other species as well (Whiteley et al. 2006). On the other hand, if interspecific interactions between the focal species lead to differences in their SGSs, conservation strategies targeted for single species systems might not be adequate. In these cases the SGSs of each species of conservation concern would need to be determined separately.

We studied the SGSs of *Calopteryx splendens* [HARRIS, 1782] and *C. virgo* [LINNAEUS, 1758] in Finland (Fig. 1). The two damselfly species provide a suitable system for studying the generalizability of SGS across species with largely similar ecologies, since they are considerably similar in their appearance, behavior, life-history and habitat choice (Askew 2004; Wellenreuther et al. 2012; Karjalainen & Hämäläinen 2013; but see Sternberg & Buchwald 1999). To fully describe the SGSs of both species we determined how genetic variation is distributed within (allelic richness and heterozygosity) and between sub-populations (isolation-by-distance, population specific $F_{ST}$, pairwise $F_{ST}$s and $D_{est}$s, and number of genetic clusters). From each sub-population, we also recorded locality characteristics (latitude and longitude, density of conspecific males and connectivity of the landscape), which we considered to be associated with colonization history, random genetic drift and dispersal between localities. To estimate the frequency of interspecific interactions experienced, we recorded the density of heterospecific *Calopteryx* males at each sub-population. Our aims were to see (1) if the SGSs of the two *Calopteryx* species are
concordant, (2) if the SGSs of the species have similar relationships with the characteristics of the localities, and (3) if the SGSs of the two species are affected by interspecific interactions.

METHODS

Calopteryx species

The Eurasian *C. splendens* and *C. virgo* have largely overlapping biogeographic ranges with both species inhabiting small rivers and streams across most of Europe (Askew 2004; Karjalainen & Hämäläinen 2013). Within their sympatric range in southern Fennoscandia, the environmental niches of the *Calopteryx* species also overlap, with the niche of *C. splendens* being nested within that of *C. virgo* (Wellenreuther et al. 2012). However, some habitat partitioning between the two species has been reported in Central Europe, where *C. splendens* is most often found in slightly warmer (18-24°C) lowland rivers and *C. virgo* is found in cooler and more shaded rivers (13-18°C) at higher elevations (Sternberg & Buchwald 1999; Schütte & Schrimpf 2002). The different temperature affinities of the two species are reflected in differences in their range limits in the north: the range of *C. splendens* extends to roughly 64° Lat., whereas the range of *C. virgo* extends to 68° Lat. (Valle 1952; Karjalainen 2010; Karjalainen & Hämäläinen 2013). However, although *C. virgo* seems to occupy a wider range of water temperatures, we found no habitat partitioning with respect to water temperature in the two *Calopteryx* species at our study sites (Supplementary material A). Also, since the wings of *C. virgo* are more melanised than those of *C. splendens*, it has been suggested that *C. virgo* has better thermoregulation abilities allowing them to adapt to colder environments (Outomuro & Ocharan 2011). However, recent work concluded that differences
in wing melanisation between the Calopteryx species is primarily driven by sexual selection, and that it is less important for adaptation to colder environments (Svensson & Waller 2013). Capture-mark-recapture studies indicate that sexually mature individuals of both Calopteryx species rarely disperse more than 300 meters from the site of first observation (Stettmer 1996; Schutte et al. 1997; Ward & Mill 2007). The longest observed dispersal distances of adult Calopteryx individuals have been 1.7 km and 4 km for C. splendens and C. virgo, respectively (Stettmer 1996). We are not aware of studies explicitly focusing on dispersal prior to sexual maturation (i.e. natal dispersal), but we expect it to be limited since significant population differentiation has been reported for C. splendens (overall $F_{ST}$ range from 0.05 to 0.14 in different studies; Svensson et al. 2004; Chaput-Bardy et al. 2008; Viitaniemi 2009). However, we are not aware of any study on the SGS of C. virgo.

In Finland, C. splendens sub-populations are nearly exclusively sympatric with C. virgo, although C. virgo sub-populations can be locally allopatric within the distribution range of C. splendens (Tynkkynen et al. 2004; but see Ilvonen et al. 2011). The species frequently interact in locally sympatric sub-populations and occasionally hybridize (Svensson et al. 2007; Keränen et al. 2013). Both species are territorial, and the males exhibit interspecific aggression when competing for breeding territories (De Marchi 1990; Tynkkynen et al. 2004, 2006). C. virgo seems to be the stronger competitor (Tynkkynen et al. 2004, 2006), and interspecific interactions with C. virgo are known to influence the secondary sexual characters of C. splendens (Tynkkynen et al. 2005; Honkavaara et al. 2011; Kuitunen et al. 2011). Since the mating success of Calopteryx males is related to their territory holding ability (Plaistow & Siva-Jothy 1996), interspecific competition could reduce the effective population size of the weaker competitor. Indeed, a previous study suggested that the genetic diversity of C.
**splendens** declines as the proportion of *C. virgo* individuals in sympatric sub-populations increases (Viitaniemi 2009).

**Sampling localities and collecting *Calopteryx* individuals**

Sampling was conducted between 27\textsuperscript{th} of June and 29\textsuperscript{th} of July, and between 23\textsuperscript{nd} of June and 28\textsuperscript{th} of July, in 2008 and 2009, respectively. Nineteen of 40 sampling localities contained both *C. splendens* and *C. virgo* (i.e. sympatric sub-populations) and the others contained only *C. virgo* (i.e. allopatric sub-populations), nine of which were located in Northern Finland, outside the range of *C. splendens* (Fig. 1). Considering the relatively poor dispersal capabilities of *Calopteryx* damselflies (see above), we decided to sample sub-populations from different rivers separated by at least one lake and a minimum distance of ten kilometers, an exception being *C. virgo* sub-populations AJ and AK (Fig. 1; Supplementary material B, Table B1), which were only 4.9 km apart.

*Calopteryx* individuals were collected using butterfly nets. Either whole individuals or a single leg per individual (if population sizes seemed low) were stored in 95% ethanol (EtOH) at 6-8°C until DNA extraction. Removing a leg from damselflies does not impact fitness (Thompson *et al.* 2011). Since only two *C. splendens* individuals were collected from sub-population BG, these individuals were not genotyped or included in the analyses. Therefore, the sample size for *C. splendens* sub-populations was 18. For most sub-populations, genotyped individuals were collected in either 2008 or 2009, but for populations AA, AC, AD and AK (Supplementary material B, Table B1.) we genotyped individuals collected in both years to test for temporal population structure (Supplementary material C).
Recording locality characteristics

Geographical coordinates were recorded on site with a hand-held GPS device (Garmin eTrex Legend HCx or Magellan Triton 300). Latitude and longitude were used in our analyses as UTM coordinates according to grid ETRS-TM35FIN. *Calopteryx* densities were estimated by counting all individuals within a 300-meter section of river, which was defined as 150 meters both up- and downstream from the site of the perceived highest density of *Calopteryx* individuals. Although both male and female density were estimated, only male densities were included in the analyses, since females frequently move away from rivers in response to changing weather, making estimates of their density less reliable (K. Kuitunen, pers. obs.). All counts were recorded between 9 AM and 4 PM, when the damselflies are most active (Corbet 1999; Karjalainen & Hämäläinen 2013). Densities at most sites were estimated at least twice during the season and some were also estimated in both years (Supplementary material B, Table B1). Repeatability of the density estimates was high (*C. splendens*: repeatability=0.64, F=5.745, df$_1$=16, df$_2$=29, n$_0$=2.63; *C. virgo*: repeatability=0.62, F=5.477, df$_1$=26, df$_2$=47, n$_0$=2.77; Lessells & Boag 1987). We used logarithm-transformed (log$_{10}$) averages of the multiple estimates in our analyses. In the case of *C. splendens*, we added one to each density estimate before transformation, since several sites did not contain a single *C. splendens* individual. For simplicity, the logarithm-transformed density estimates are hereafter referred to simply as density estimates.

To obtain a locality specific measure for potential connectivity we quantified the total length of river habitat within a five-kilometer radius buffer zone from each sampling site. Width of river habitat was not included because breeding sites of *Calopteryx* are mainly found along the shoreline. Measurements were made from a combination of two map layers (“River
Network” & “Water formations, EU Water Framework Directive”, provided by The Finnish Environment Institute & Centres for Economic Development, Transport and the Environment) using ArcGIS 10.1 (ESRI 2012. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute). Our measure reflects potential connectivity because it does not consider whether the habitat contains Calopteryx individuals (Calabrese & Fagan 2004; Hughes et al. 2013). Our measure neglects the potentially important branching structure of the rivers (Chaput-Bardy et al. 2008), however, it more accurately reflects our sampling scheme, given that we sampled sub-populations residing in different rivers rather than different tributaries of the same river. The five kilometer radius buffer should sufficiently represent potential connectivity since this distance is at the higher range of dispersal for both C. virgo and C. splendens (Stettmer 1996; Schutte et al. 1997).

Prior to incorporating any of the variables into the analyses we checked the data for collinearity and spatial autocorrelation (Moran’s I). Collinearity was tested by calculating variance inflation factors (VIF) for all variables in R (R Core Team 2012), and spatial autocorrelation was tested across the entire 40 site dataset using SAM 4.0 (Rangel et al. 2010). VIFs were low for all variables (all below 2), suggesting no collinearity. However, a negative spatial autocorrelation (Moran’s I = -0.321; P=0.005) was observed in the density of C. splendens males in the longest distance class (center at c.a. 607 km). This occurred because nine of the sampled localities (C. virgo sub-populations) were outside the distribution range of C. splendens, resulting in spatial aggregation of localities with zero density of C. splendens (Fig. 1). Thus, we analyzed a separate subset of the C. virgo data containing only the sympatric populations (see below). There was no spatial autocorrelation in any of the other variables.
Population Genetic Methods

DNA was extracted from the legs of *Calopteryx* individuals and used for genotyping of microsatellite loci as described in Supplementary material C and Molecular Ecology Resources Primer Development Consortium et al. (2011). The microsatellite loci were assessed for reliable scoring and potential deviations from equilibrium and neutrality assumptions (Supplementary material C). Finally, the *C. splendens* dataset included 567 individuals genotyped at 12 loci and the *C. virgo* dataset included 1401 individuals genotyped at six loci (Supplementary material C). For each locus and subpopulation, unbiased expected heterozygosity (*uH*E) estimated using GenAlEx 6.5 (Peakall & Smouse 2012) and allelic richness (*AR*) adjusted by rarefaction was estimated using FSTAT 2.9.3 (updated from Goudet 1995; See Supplementary material C, Tables C3 and C4).

Statistical Methods

We studied how environmental characteristics (see above) and density of heterospecific males are related to intrapopulation genetic diversity as measured by *uH*E and *AR*, using a beta regression and a linear multiple regression, respectively. Analyses were conducted using R and the package “betareg” (Cribari-Neto & Zeleis 2010) was used for the beta regression. All independent variables were initially entered into the model and then removed on a stepwise basis by comparing nested models with a log-likelihood ratio test.

To study how genetic diversity was distributed among populations (i.e. population differentiation), we tested for isolation-by-distance (IBD), pairwise genetic differentiation between sub-populations (both *F*ST and *D*est, Weir & Cockerham 1984; Jost 2008), and
conducted individual-based spatial Bayesian clustering analyses. IBD was tested with a
permutation test of the relationship between the Euclidean geographic distance matrix (both
untransformed distances in kilometers and logarithm transformed distances) and the pairwise
relatedness matrix \((F_{ij}; \text{Ritland }1996)\) with 5000 randomizations using SPAGeDi 1.4 (Hardy
& Vekemans 2002). To visualize the relationship, the pairwise \(F_{ij}\)s were divided into distance
classes with approximately equal numbers of comparisons and averaged to acquire a distance
class specific \(F(d)\). We estimated the pairwise differentiation coefficients and their 95%
confidence intervals using the R package “diveRsity” (Keenan et al. 2013) with 5000
bootstraps for both estimates and for both species.

The individual based Bayesian clustering was conducted in the program TESS 2.3.1 (Chen et
al. 2007; Durand et al. 2009). Because all individuals from the same sub-population shared
the same coordinates, we randomized geographic coordinates for each individual in TESS
2.3.1. Specifically, latitude and longitude were allowed to vary a maximum of one kilometer
from the sampling site coordinates, with a standard deviation of 300 meters. Given what is
known about the dispersal of \(Calopteryx\) individuals (Stettmer 1996; Schutte et al. 1997;
Ward & Mill 2007), we considered these parameters for the randomization to be biologically
realistic. Assuming a conditional autoregressive Gaussian admixture model with a quadratic
trend surface degree, we ran TESS 2.3.1 for 50 000 iterations (preceded by a 50 000 iteration
burn-in) 20 times for each maximum number of clusters \((K_{\text{max}})\), which ranged from two to the
number of sampling localities \((K_{\text{max}}= 2\ldots18, \text{ and } K_{\text{max}}= 2\ldots40, \text{ for } C. \text{ splendens and } C. \text{ virgo,}
respectively). We used deviance information criterion (\(DIC\)) values and stabilization of the \(Q-
matrix of posterior probabilities (i.e. when the number of observed clusters did not increase
with an increase in \(K_{\text{max}}\)) as metrics for choosing the most appropriate \(K_{\text{max}}\) for the data. An
admixtured model was assumed since such models are expected to be more robust against overestimation of $K_{max}$ when there are genetic clines (Guillot 2009; François & Durand 2010). Once $K_{max}$ was deduced, 80 additional replicate runs were conducted to yield a total of 100 replicate runs for $K_{max}$. From these, 20 runs with the lowest DIC values were used to calculate average individual admixture proportions with CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007). These were visualized as a bar plot using DISTRUCT 1.1 (Rosenberg 2004).

To study how genetic differentiation is related to the locality characteristics and density of heterospecific *Calopteryx* males, we used a Bayesian population based analysis of genetic differentiation in the program GESTE2 (Foll & Gaggiotti 2006). GESTE2 incorporates non-genetic information in the prior distribution of population specific $F_{ST}$-values via a generalized linear model (Balding 2003; Foll & Gaggiotti 2006). GESTE2 analyses were run for a total of 100 000 iterations with a 50 000 iteration burn-in period.

Since the sizes of the two datasets differed both in the amount of genetic information derived from each individual (12 loci vs. 6 loci for *C. splendens* and *C. virgo*, respectively) and in their geographic scope (18 sub-populations spanning c.a. 270km vs. 40 sub-populations spanning c.a. 700km for *C. splendens* and *C. virgo*, respectively), subsets of the data were analyzed to confirm the results obtained from the total dataset. To do this, we conducted clustering in TESS for five subsets of the *C. splendens* dataset: three subsets each with six randomly selected loci, one subset with the same loci that were used in the *C. virgo* dataset, and one subset excluding the loci that showed some deviations from the assumptions (see above; Supplementary material F, Fig. F1) In addition, a subset of the *C. virgo* dataset including only the sympatric populations was analyzed. This was done for two reasons: (1) to check if possible differences in the SGSs of the two *Calopteryx* species are associated with
the different geographic scales of sampling and different numbers of sampled sub-
populations, and (2) to account for the negative spatial autocorrelation in the density of *C.
* *splendens* males (see above).

RESULTS

The genetic diversities of the two species were related to different locality characteristics. The 
*uH*$_E$ of *C. splendens* sub-populations was related to potential connectivity, with increasing 
potential connectivity being coupled with increasing *uH*$_E$ (Table 1). Interestingly, *C.
* *splendens* *uH*$_E$ was differentially related to the densities of conspecific versus heterospecific 
males: increasing density of conspecific males was coupled with increasing *uH*$_E$, whereas the 
increasing density of heterospecific males was coupled with decreasing *uH*$_E$. Surprisingly, 
none of the locality characteristics were related to the *AR* of *C. splendens*. The genetic 
diversity of *C. virgo* (both *uH*$_E$ and *AR*) was related to latitude and longitude, with genetic 
diversity increasing with decreasing latitude and increasing longitude (Table 1).

There was no statistically significant IBD pattern in pairwise individual kinship coefficients 
(*F*$_{ij}$) of *C. splendens* (Fig 2a; jackknifed $\beta=-6.32*10^{-5}$, S.E.=1.34*10^{-5}; 95% CI of permuted 
null distribution = [-6.67*10^{-5}, -3.55*10^{-5}]; $P=0.129$), although there was a distinct decrease 
in the kinship coefficients after the first distance class. On the other hand, for *C. virgo*, 
kinship coefficients decreased gradually from the first distance class onwards, and a 
statistically significant IBD pattern was observed (Fig 2b; $\beta=-7.89*10^{-6}$, S.E.=3.16*10^{-6}; 95%
CI of permuted null distribution = [-2.31*10^{-6}, -1.62*10^{-7}]; $P<0.001$). However, although a 
statistically significant IBD is observed, it must be noted that the *F*$_{ij}$ values for *C. virgo* are 
very low.
Furthermore, different levels of genetic differentiation were observed in the two species, both in comparisons of pairwise genetic differentiation and in individual based clustering analyses. For most of the pairwise comparisons between the *C. splendens* sub-populations, the lower boundary of the 95% confidence interval was well above 0.01 for both $F_{ST}$ and $D_{est}$ (Supplementary material D, Tables D1 and D2), whereas for majority of the pairwise comparisons between *C. virgo* sub-populations, the 95% confidence intervals either included, or were very close to zero at the lower boundaries (Supplementary material D, Tables D3 and D4). Moreover, the individual based genetic clustering in TESS resulted in a different $K_{max}$ for the two species. For *C. splendens* six genetic clusters were detected, whereas for *C. virgo* only two clusters were detected (Fig. 3). Five of the six *C. splendens* clusters each had a high affinity with a single sub-population, and the remaining cluster had a high affinity with five sub-populations in central Finland (Fig. 3a). In *C. virgo* the two genetic clusters seemed to represent two ends of a cline, with one cluster having a high affinity with nine sub-populations from southern and eastern Finland and the other with ten sub-populations in northern Finland. Sub-populations geographically intermediate to those in the two clusters showed different levels of admixture (Fig. 3b).

In addition to differences in the number of genetic clusters observed for the two *Calopteryx* species, differentiation of the sub-populations was related to different locality characteristics (Tables 1, 2 & 3; Fig. 4). In the GESTE2 analysis of *C. splendens* population specific $F_{STR}$, the models with potential connectivity were assigned higher posterior probabilities than most other models (Table 2) resulting in a summed posterior probability of 0.605 for potential connectivity. Of all the models, a model with constant and potential connectivity received the highest posterior probability, indicating that population specific $F_{STR}$ decreases with increasing
potential connectivity (Table 3; Fig. 4a). For *C. virgo*, however, latitude received the highest summed posterior probability (0.994) and a model with constant and latitude had the highest posterior probability, indicating that $F_{ST}$ increases with increasing latitude (Table 3; Fig. 4b). The population specific $F_{ST}$s of the *C. virgo* sub-populations were also lower than those of *C. splendens* sub-populations (Fig. 4).

The difference between the two species in the number of observed genetic clusters was most likely not an artifact of a difference in the number of microsatellite loci analyzed, since all of the *C. splendens* subsets also detected more than two clusters (observed $K_{max}$ was 3, 3, 5, 4 and 4 in the five respective subsets; Supplementary material F, Fig. F1). Also, the IBD pattern observed in the SGS of *C. virgo* remained statistically significant even when the dataset was reduced to a subset including only the sub-populations that are sympatric with *C. splendens* ($\beta$=7.16*10^{-6}, S.E.=3.49*10^{-6}; 95% CI of permuted null distribution = [-5.65*10^{-6}, 3.98*10^{-6}]; P=0.004). However, other patterns observed from the *C. virgo* dataset disappeared when the subset of sympatric sub-populations was analyzed: null models were selected for population specific $F_{ST}$, $uH_E$ and $AR$ (Supplementary material F, Tables F1 and F2).
DISCUSSION

Despite the ecological similarities of the sympatric _C. splendens_ and _C. virgo_ in Finland, they exhibited markedly different spatial genetic structures (SGS). Rather than supporting the hypothesis that sympatric species with similar ecologies exhibit similar SGS (Whiteley _et al._ 2006; Dawson 2012; Hughes _et al._ 2013), our results highlight the possibility that ecological similarity can, in fact, create differences in the SGS of interacting species.

_C. splendens_

The positive relationship between unbiased expected heterozygosity (\(uH_E\)) and density of conspecific males (Table 1) together with multiple genetic clusters, most corresponding to individual sub-populations (Fig. 3), indicate predominant effects of drift and low gene flow on SGS in _C. splendens_. Nonetheless, some legacy of previous gene flow is apparent in the positive relationship between potential connectivity (i.e. total length of streams within a 5 km buffer zone) and \(uH_E\) and the negative relationship between potential connectivity and population specific \(F_{ST}\) (Tables 1 & 3; Fig. 4a). _C. splendens_ apparently rarely disperse more than few kilometers and prefer dispersal along streams (Stettmer 1996; Schutte _et al._ 1997; Ward & Mill 2007; Chaput-Bardy _et al._ 2008). Previous studies have also found dispersing _C. splendens_ to have lower mating success than residents, which can further inhibit gene flow and strengthen differentiation between _C. splendens_ sub-populations (Svensson _et al._ 2006; Wellenreuther _et al._ 2010).

In contrast with previous population genetic studies on this species (Svensson _et al._ 2004; Chaput-Bardy _et al._ 2008; Viitaniemi 2009) we did not observe IBD (Fig. 2). The absence of
IBD could be attributed to genetic disequilibrium due to recent colonization followed by rapid range expansion (e.g. as seen in *Erythromma viridulum*, Watts *et al.* 2010; and *Coenagrion sticum*, Swaegers *et al.* 2013) or strong overall differentiation due to range-margin effects (Eckert *et al.* 2008). Regarding recent colonization, no major shifts in the species range in Finland have been recorded in recent decades. Although post-glacial colonization can still be reflected in the SGS, it remains unclear if short-term effects (i.e. within the range of decades; Watts *et al.* 2010; Swaegers *et al.* 2013) influence the SGS of *C. splendens*. Regarding range-margin effects, the overall genetic differentiation of *C. splendens* observed here did not differ greatly from that observed in different regions (overall $F_{ST}=0.11$ vs. 0.05 in southern Sweden and 0.14 in western France; Svensson *et al.* 2004; Chaput-Bardy *et al.* 2008). Although our results cannot be directly compared to those of previous studies because different types of markers with different levels of polymorphism were used (microsatellites vs. AFLPs; Jakobsson *et al.* 2013), similar overall differentiation between sub-populations in the Finnish *C. splendens* population were also found using AFLPs ($\Phi_{PT}=0.08$; Viitaniemi 2009). Most likely, the lack of IBD is due to strong genetic differentiation between some neighboring sub-populations (e.g. pairwise $F_{ST}=0.20$ between sub-populations AK & CJ, Fig. 1; Supplementary material D, Table D1), as a result of our sampling, which focused on sub-populations at different rivers separated by land and/or lakes.

An interesting finding in our analysis is that interspecific interactions could contribute to the strength of drift experienced by *C. splendens* sub-populations (Table 1): when *C. virgo* males are abundant, sympatric *C. splendens* sub-populations are less genetically diverse. Alternatively, such a relationship could arise due to habitat partitioning and local adaptation of the two species to different habitats. In Central Europe, the two *Calopteryx* species do
show different preferences to water temperature (Sternberg & Buchwald 1999; Schütte & Schrimpf 2002). If locally adapted, the population sizes of the two species would be expected to show opposing trends with water temperature. However, our data show that in Finland the preferred water temperatures of the two species overlap. Although we have not quantified habitat partitioning with respect to other environmental variables, densities of the two species in sympatric sub-populations do not correlate negatively, suggesting that possible habitat partitioning between the species is not strong (Supplementary material A).

Our results are concordant with a study by Viitaniemi (2009), who suggested that interspecific interactions explained a decline in the heterozygosity of *C. splendens* with increasing proportion of *C. virgo* individuals in sympatric sub-populations. We believe that the density of *C. virgo*, rather than its relative abundance in sympatric populations, better describes how frequently heterospecific individuals are encountered and is a more appropriate measure of the frequency of interspecific interactions. Unfortunately, with our data it is not possible to determine how interspecific interactions shape the effective population sizes of sub-populations directly, or reduce gene flow, or both (e.g. as suggested by Ranta *et al.* 2009). It is known that increasing aggression of *C. virgo* males towards *C. splendens* males decreases the proportion of *C. splendens* males with a territorial mating strategy in Finnish sub-populations (Tynkkynen *et al.* 2004, 2006). Since territorial males have better mating success than non-territorial males (Plaistow & Siva-Jothy 1996), it is likely that *C. virgo* influences the effective population size by accentuating the reproductive skew of *C. splendens*. However, it is also possible that interspecific aggression limits gene flow by reducing either the immigration of *C. splendens* (i.e. males immigrate to sites with fewer *C. virgo*), or the probability that immigrants obtain matings (i.e. immigrant males are not able to hold
territories at sites with high *C. virgo* abundance). Future work should examine the role of interspecific aggression as a likely explanation for the observed negative relationship between *C. splendens* genetic diversity and *C. virgo* density.

Although $uH_E$ of *C. splendens* was negatively related to *C. virgo* density, the effect of *C. virgo* density was not identified in the analysis of population specific $F_{ST}$s of *C. splendens*, which could be due to the low statistical power of GESTE2 (Balkenhol *et al.* 2009). It is more puzzling, however, that we did not observe similar locality-specific patterns with $AR$ as we did with $uH_E$ (Table 1, Supplementary material E, Table E1), since allelic richness and heterozygosity often exhibit concordant patterns even though allelic richness is more sensitive to changes in effective population size (Luikart *et al.* 1998; Eckert *et al.* 2008). A reasonable explanation for this counterintuitive pattern could be that allelic richness of the Finnish *C. splendens* population is already low, and most variation in $AR$ between sub-populations is in the relative frequencies of common alleles (Supplementary material C, Table C3). Indeed, the average number of alleles per locus observed across the whole *C. splendens* dataset is 3.38 (S.D. = 1.59) and the average frequency of the most common allele is 0.72 (S.D. = 0.20), leading to a situation in which the relative amount of variation is higher in $uH_E$ than in $AR$ (Coefficient of variation = 0.130 and 0.086 respectively). This could be a consequence of the markers, since variability of microsatellite loci in many dragonflies and damselflies is generally low (Watts 2009), or alternatively reflect the fact that populations close to range-margins generally have low genetic diversities (Eckert *et al.* 2008; Watts *et al.* 2010).
C. virgo

To our knowledge the SGS of C. virgo has not been described before, and to our surprise, we observed an SGS that was very different from that observed in C. splendens. Latitude was consistently related to all aspects of the SGS of C. virgo (uHe, AR, IBD, clustering and population specific $F_{ST}$; Tables 1, 2 & 3; Figures 2b, 3b, and 4b). In addition to latitude, the genetic diversity of C. virgo was also related to longitude, with increasing longitude being coupled with increasing genetic diversity (Table 1). The relationship of C. virgo’s SGS with latitude and longitude probably reflects a persistent signature of colonization after the last glacial maximum, and latitudinal gradients in genetic diversity at geographic scales similar to ours have been reported before (Hewitt 2000; Schmitt & Seitz 2001). Alternatively, the latitudinal trends in diversity and differentiation can result from range margin effects, which are difficult to differentiate from the effects of post glacial colonization (Eckert et al. 2008).

On the other hand, low genetic differentiation between sub-populations and a latitudinal cline in genetic diversity could be due genetic disequilibrium caused by recent colonization or range expansion (Watts et al. 2010; Swaegers et al. 2013); but, similarly to the case of C. splendens, we cannot determine this since we do not have adequate temporal data on changes in SGS or know the timing of colonization of Finland by C. virgo. However, although two genetic clusters are discovered and some pairwise genetic differences between southern and northern C. virgo sub-populations are significant (Supplementary material D, Tables D3 and D4) the genetic differentiation is low, which likely reflects gene flow between populations. Indeed, this is supported by the cline of admixture between the genetic clusters and low population specific $F_{ST}$ (Figures 3b & 4b).
Why is there a difference between the SGSs of the two Calopteryx species?

Although it is not unprecedented that closely related species show different SGSs (Hodges et al. 2007; Lehrian et al. 2009; Fortuna et al. 2009; Johansson et al. 2013), our results are surprising considering the similarities in the life-history characteristics of the two species and extent of habitat overlap in Fennoscandia. The different SGSs of the two Calopteryx species are unlikely to be explained by statistical artifacts, such as differences in the information content of the loci (Supplementary material C, Tables C2, C3 and C4; Jakobsson et al. 2013), different numbers of sampled loci, or different geographic ranges of sampling (Supplementary material F). However, since latitudinal and longitudinal variation were not significantly related to the SGS of C. virgo in a subset consisting only of the sympatric sub-populations, we cannot rule out the possibility that the lack of relationship between latitude and longitude with C. splendens SGS is partly due to a difference in scale.

Our study shows that the SGS of C. splendens is influenced more by genetic drift than is the SGS of C. virgo. Furthermore, our results suggest that interspecific interactions between C. splendens and C. virgo contribute to this difference by strengthening the magnitude of genetic drift experienced by C. splendens sub-populations while the SGS of C. virgo is not affected. A likely explanation for this discrepancy is that, in Finland, C. virgo is the stronger competitor for the best breeding habitats (Tynkkynen et al. 2004, 2005, 2006). The competitive advantage of C. virgo could be due to its ability to occupy a wider thermal niche in Fennoscandia (Wellenreuther et al. 2012; Supplementary material A). Indeed, C. virgo is more common in Finland than C. splendens and this seems to be also reflected in its SGS: the abundance of sub-populations offer more stepping stones for dispersal and gene flow, leading to less genetic differentiation between C. virgo sub-populations.
Conservation implications

Since characterizing SGS of multiple species for landscape level conservation planning is often unfeasible, surrogate approaches are in high demand from practicing conservationists to aid prediction of the effects of different management and land-use actions (Margules & Pressey 2000; Rodrigues & Brooks 2007). Although neither *C. splendens* nor *C. virgo* is endangered, the two species system can function as a test of the surrogate approach. Our results highlight the fact that use of a surrogate requires careful consideration. Interspecific interactions are an important part of the community that can have a role in shaping the SGS of natural populations, and should not be overlooked in conservation planning.

Acknowledgements

We are grateful to four anonymous referees, Matti Häkkilä, Reetta Hänninen, Sami Karjalainen, Matti Koivula, Maria Marjeta, Adriano Mazziotta, Kaisa Mustola, Mikael Puurtinen, Jenni Tikka, Janne Valkonen and the members of MCC for constructive feedback and helping out with various aspects of the manuscript.

This research was funded by the Academy of Finland (via Biological Interactions Graduate School, The Center of Excellence for Evolutionary Research, and the postdoctoral project funding of Katja Kuitunen), **Kone Foundation**, The Finnish Cultural Fund, Kuopion Luonnonystäväin Yhdistys, and Societas Biologica Fennica Vanamo.
Literature


R Core Team (2012) R: A Language and Environment for Statistical Computing.


DATA ACCESSIBILITY

Genotypes for both *Calopteryx splendens* and *Calopteryx virgo*, accompanied with sub-population specific environmental data uploaded to Dryad (doi:10.5061/dryad.v1ss7).
AUTHOR CONTRIBUTIONS

All the authors (AK, IK, KEK, KK and JSK) contributed to the design of the research; AK, IK and KK conducted the field work; AK, IK and KEK designed the laboratory protocol and conducted the laboratory work; AK and KEK conducted the analyses; AK led the writing process of the manuscript, to which all authors contributed.
Table 1. The relationships between locality characteristics and genetic diversity of *Calopteryx* species. The results considering $uH_E$ and $AR$ represent beta regression and multiple linear regression models, respectively. For *C. splendens* only a beta regression model for the relationship between $uH_E$ and locality characteristics is displayed, as $AR$ showed no statistically significant relationship with any of the locality characteristics. Descriptions of model selection can be found in Supplementary material E.

<table>
<thead>
<tr>
<th>Species</th>
<th>Dependent</th>
<th>Independent</th>
<th>Estimate</th>
<th>S.E.</th>
<th>$z\ (uH_E); t\ (AR)$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. splendens</em></td>
<td>$uH_E$</td>
<td>intercept</td>
<td>-0.884</td>
<td>0.139</td>
<td>-6.345</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\log_{10}(CsD)$</td>
<td>0.156</td>
<td>0.073</td>
<td>2.131</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PC</td>
<td>0.011</td>
<td>0.005</td>
<td>2.126</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\log_{10}(CvD)$</td>
<td>-0.183</td>
<td>0.052</td>
<td>-3.479</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><em>C. virgo</em></td>
<td>$uH_E$</td>
<td>intercept</td>
<td>1.565</td>
<td>0.560</td>
<td>2.794</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>-3.432*10^{-7}</td>
<td>8.436*10^{-8}</td>
<td>-4.068</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>3.474*10^{-7}</td>
<td>1.593*10^{-7}</td>
<td>2.181</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>$AR$</td>
<td>intercept</td>
<td>6.365</td>
<td>0.698</td>
<td>9.120</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>-5.720*10^{-7}</td>
<td>1.051*10^{-7}</td>
<td>-5.436</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E</td>
<td>4.451*10^{-7}</td>
<td>2.012*10^{-7}</td>
<td>2.213</td>
<td>0.033</td>
</tr>
</tbody>
</table>

Abbreviations of independent variables of the models: $\log_{10}(CsD)$ = logarithm transformed density estimates of *C. splendens* males; PC = potential connectivity (in kilometers; see text for description); $\log_{10}(CvD)$ = logarithm transformed density estimates of *C. virgo* males; N = latitude in UTM coordinates; E = longitude in UTM coordinates.

Table 2. The highest posterior probability models for the relationship between population specific $F_{ST}$ and locality characteristics obtained using GESTE2.

<table>
<thead>
<tr>
<th>Model #</th>
<th>Probability</th>
<th>Locality characteristics included</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. splendens</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><strong>0.455</strong></td>
<td>PC</td>
</tr>
<tr>
<td>1</td>
<td>0.282</td>
<td>Null</td>
</tr>
<tr>
<td>4</td>
<td>0.062</td>
<td>$\log_{10}(CsD)$, PC</td>
</tr>
<tr>
<td><em>C. virgo</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td><strong>0.522</strong></td>
<td>N</td>
</tr>
<tr>
<td>25</td>
<td>0.204</td>
<td>N, E</td>
</tr>
<tr>
<td>13</td>
<td>0.100</td>
<td>N, $\log_{10}(CsD)$</td>
</tr>
</tbody>
</table>
Abbreviations of independent variables of the models: PC = potential connectivity (see text for description); N = latitude in UTM coordinates; $\log_{10}(C_sD) = \log_{10}$ transformed estimates of density of $C$. splendens males; $E = \text{longitude}$ in UTM coordinates. Null stands for a null-model including only the intercept but none of the explanatory variables.

Table 3. Coefficients of the models with the highest posterior probabilities for population specific $F_{ST}$ (bold in Table 1).

<table>
<thead>
<tr>
<th>Species</th>
<th>Estimates</th>
<th>Regression coefficient</th>
<th>HDPI 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Mode</td>
</tr>
<tr>
<td>$C$. splendens</td>
<td>Constant ($\alpha_0$)</td>
<td>-2.14</td>
<td>-2.19</td>
</tr>
<tr>
<td></td>
<td>PC ($\alpha_2$)</td>
<td>-0.393</td>
<td>-0.400</td>
</tr>
<tr>
<td></td>
<td>Variance ($\sigma^2$)</td>
<td>0.336</td>
<td>0.257</td>
</tr>
<tr>
<td>$C$. virgo</td>
<td>Constant ($\alpha_0$)</td>
<td>-4.26</td>
<td>-4.21</td>
</tr>
<tr>
<td></td>
<td>E ($\alpha_4$)</td>
<td>0.628</td>
<td>0.622</td>
</tr>
<tr>
<td></td>
<td>Variance ($\sigma^2$)</td>
<td>0.347</td>
<td>0.262</td>
</tr>
</tbody>
</table>

Abbreviations of independent variables of the models: PC = potential connectivity (see text for description); N = latitude in UTM coordinates.
**FIGURE CAPTIONS**

Figure 1. Map of *Calopteryx* sampling localities within Finland. Open symbols represent sites with both *C. virgo* and *C. splendens* in sympatry and solid symbols represent sites with *C. virgo* in allopatry. Letters indicate the locality codes (see Supplementary material B, Table B1 for more information on the localities). The dashed and dotted lines represent the range limit of *C. splendens* and *C. virgo*, respectively.

Figure 2. A correlogram of average pairwise kinship coefficients (*F*(d)) between individuals of (a) *C. splendens* and (b) *C. virgo* in different distance classes. The error bars represent ± S.E. for each *F*(d). Note the different scales on the x-axis of the two panels.

Figure 3. The admixture proportions of (a) *C. splendens* and (b) *C. virgo* individuals. Different colors represent different genetic clusters with *K*=6 and *K*=2 for *C. splendens* and *C. virgo*, respectively. The samples in the two panels are arranged according to the geographic coordinate axis along which the range of variation is the largest. Thus, the samples of (a) *C. splendens* are arranged according to longitude with the westernmost sampling localities on the left hand side and the samples of (b) *C. virgo* are arranged according to latitude, with sampling localities at low latitudes on the left hand side. Letters codes indicate the sub-population (see Supplementary material B, Table B1 for more information on the localities).

Figure 4. A graphical representation of the GESTE2 models in Table 3. (a) The relationship between *C. splendens* population specific *F*<sub>ST</sub> and potential connectivity (i.e. length of rivers within a five kilometer buffer zone), and (b) the relationship between *C. virgo* population specific *F*<sub>ST</sub> and latitude (according to UTM grid ETRS-TM35FIN).
SUPPORTING INFORMATION

- Supplementary material A: Overlap in water temperature preferences of *Calopteryx* species in Finland.

- Supplementary material B: Characteristics of *Calopteryx* sub-populations and locality information.

- Supplementary material C: Genotyping protocols, characteristics of the microsatellite loci and genetic characteristics of the sub-populations.

- Supplementary material D: Pairwise genetic differentiation of the *Calopteryx* sub-populations.

- Supplementary material E: Model selection for beta regression and linear regression models on the relationships between locality characteristics and genetic diversity.

- Supplementary material F: Analyses with subsets of the *Calopteryx* datasets.
a) Molecular Ecology Page 38 of 38

Population specific $F_{ST}$ vs. Potential connectivity (km)

b) Population specific $F_{ST}$ vs. Latitude

68000000 70000000 72000000 74000000