Contemporary pollen flow as a multiscale process: evidence from the insect-pollinated herb, *Pulsatilla vulgaris*

Michelle F. DiLeo¹,²*, Rolf Holderegger³, Helene H. Wagner¹

¹ – Department of Ecology & Evolution, University of Toronto, 3359 Mississauga Road, Mississauga ON, L5L 1C6

² – Metapopulation Research Centre, Faculty of Biological and Environmental Sciences, University of Helsinki, PO Box 65, 00014, Finland

³ – Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland

*corresponding author: michelle.dileo@gmail.com; +358 041 7574836

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2745.12992

This article is protected by copyright. All rights reserved.
Keywords: dispersal, functional connectivity, pollen-mediated gene flow, isolation by distance, landscape ecology, landscape genetics, landscape structure, long distance dispersal, spatial scale

Running headline: pollen flow as a multiscale process

Abstract

1. Understanding the drivers and spatial scale of gene flow is essential for the management of species living in fragmented landscapes. In plants, contemporary pollen flow is typically modelled as a single spatial process, with pollen flow declining exponentially within a short distance of mother plants. However, growing evidence suggests that many species do not conform to this pattern, often showing an excess of long-distance dispersal events or sometimes even multimodality in dispersal kernels. This suggests that a single function might be insufficient to capture the true complexity of pollination, which in reality is often achieved by multiple pollinators that vary in their foraging ranges and interactions with the landscape.

2. We reconstructed realized pollen flow and assessed pollen immigration for seven populations of the insect-pollinated herb, *Pulsatilla vulgaris*. We quantified the effects of distance, floral resources, and landscape composition over multiple spatial scales, and tested the hypotheses that within-population pollen flow is related to resources and landscape context measured locally, and that among-population pollen flow is related to features measured at larger spatial scales.

3. We found that pollen flow within populations was more likely to occur amongst near neighbours, but that among-population pollen flow was random with respect to source
populations. We further found that local floral density could explain patterns of within-population pollination distances and population-level selfing rates, whereas pollen immigration rates were best explained by the proportion of forest within a radius of 500 m around focal populations.

4. *Synthesis*. Together our results suggest that within- and among-population contemporary pollen flow may be governed by different underlying processes, possibly related to differences in the foraging range and habitat use of bee species that contribute to pollination at different scales. This highlights the critical need for researchers to take a more pollinator-eyed view of contemporary pollen flow in plants by (1) recognizing that within- and among-population gene flow by pollen may depend on different sets of pollinators that respond to features at different spatial scales (2) considering additional factors that may alter attractiveness, detectability, and accessibility of plants to pollinators beyond the effects of distance.

**Introduction**

The spatial scaling of dispersal and gene flow has far-reaching consequences for ecological and evolutionary processes including the maintenance of genetic diversity (Lopez *et al.* 2009), the dynamics of shifting ranges (Aguilee *et al.* 2016), and the structure and function of metacommunities (Thompson & Gonzalez 2017). There is thus a keen interest in understanding and generalizing patterns of contemporary gene flow across heterogeneous landscapes. In plants, gene flow is commonly modelled as a single spatial process, assuming that gene flow decays with distance. However, many plant species interact with multiple pollen vectors, each of which might contribute to gene flow over different spatial scales and interact with the environment in
different ways (Auffret et al. 2017). Thus modelling pollen-mediated gene flow with a single geographic function might fail to capture the true complexity of pollination. Indeed, mounting evidence from parentage-based reconstruction of contemporary gene flow suggests that many animal-pollinated plants do not conform to commonly used distance decay functions and often exhibit pervasive dispersal over surprisingly large distances (e.g. Robledo-Arnuncio & Gil 2005; Ismail et al. 2012; Bezemer et al. 2016).

The increased availability of molecular markers and improvement of parentage assignment methods (Wang 2004; Moran & Clark 2011) have resulted in an explosion of studies reconstructing seed and pollen flow kernels for a variety of plant species. These studies have uncovered a remarkable diversity of dispersal kernels (Ashley 2010; Bullock et al. 2017), with many species showing multimodal and long-distance dispersal events that cannot be explained by the availability and distance to suitable habitat for seed deposition or pollen donors alone (e.g. García, Jordano & Godoy 2007; Bezemer et al. 2016). The potential multiscale nature of gene flow has received a great deal of attention in the seed dispersal literature (Nathan 2006). Examples include using mechanistic and agent-based models to capture patterns of seed dispersal (Will & Tackenberg 2008), and linking multimodality in seed dispersal kernels to distinct frugivores (Spiegel & Nathan 2007), or distinct phases of frugivore behaviour (Rodríguez-Pérez, Wiegand & Santamaria 2012; Córtes & Uriarte 2013). In comparison, few studies have explicitly considered the processes that underlie multimodality in pollen dispersal kernels (but see Castilla et al. 2017; Rhodes, Fant & Skogen 2017). Notable exceptions come from pollen dye transfer studies that have quantified contributions of distinct pollinators to pollen flow over small spatial scales (e.g. Schmitt 1980; Cresswell et al. 1995; Radar et al. 2011), and genetic studies on wind-pollinated trees, which have modelled realized pollen flow as two-component processes

This article is protected by copyright. All rights reserved.
accounting for local air flows versus long-distance air drafts (e.g. Goto et al. 2006; Slavov et al. 2009; Niggemann et al. 2012). We are thus still lacking key knowledge of the drivers of multiscale pollen flow for generalist-pollinated plants at large spatial scales.

Attributes of individual plants and the surrounding landscapes can modify pollinator behaviour and thus impact mating and the scale of pollen flow of plant populations (Loveless & Hamrick 1984; Hegland 2014; Schuepp, Herzog & Entling 2014). At the level of individual plants, scent or the size of floral displays can impact spatial patterns of contemporary pollen flow beyond the effects of distance by influencing the attractiveness of individuals to pollinators (e.g. Jakobsson, Lazaro & Totland 2009; DiLeo et al. 2014). Likewise, larger plant populations may experience higher rates of pollen immigration (i.e. among-population gene flow) by attracting pollinators from further away (Van Geert, Triest & Van Rossum 2014); although the opposite can be true when pollinators must move further and more often to forage between small populations that have few floral resources (i.e. the fragmentation paradox; Kramer et al. 2008). Landscape context is also expected to play a critical role in determining both patterns of within-population and among-population pollen flow. Populations surrounded by landscapes that harbour suitable pollinator nesting habitat may be exposed to higher abundances and diversity of pollinators and may achieve higher visitation rates, thus experience lower selfing and higher seed set than populations surrounded by landscapes unsuitable for nesting (Klein, Steffan-Dewenter & Tscharntke 2003). Landscape features can additionally impact rates of pollen immigration by modifying the detectability of populations to pollinators (e.g. canopy gaps; Walters & Stiles 1996), and by providing corridors or barriers to dispersal (Kamm et al. 2010; Krewenka et al. 2011; Lander et al. 2011).
Considering the identity of pollinators and how they might respond to features of individual plants and landscape structure can improve models of contemporary pollen flow and explain substantial variation in patterns of reproductive success (Young 2002; Brunet & Holmquist 2009; Kramer, Fant & Ashley 2011; Rhodes, Fant & Skogen 2017). This is easiest to do for specialist-pollinated plants, where pollination is achieved by few species that strongly differ in their morphology and/or behaviour. For example, Schmitt (1980) found that butterflies increased the genetic neighbourhood of *Senecio* plants by bypassing near neighbours in comparison to bumblebees, which tended to forage sequentially on neighbouring plants. However, it becomes harder to generate hypotheses about the spatial scaling and drivers of gene flow when it is achieved through the movement of many pollinator species, which is true for many bee-pollinated species. In such cases, it may be possible to draw on the pollinator literature to conceptualize the potential multiscale nature of pollination. For example, it is well documented that the spatial scale at which insect pollinators move and perceive the landscape increases with body size and sociality: social and large-bodied bees have larger foraging ranges and tend to respond to the landscape at much larger spatial scales than small-bodied and solitary bees (Gathmann & Tscharntke 2002; Steffan-Dewenter et al. 2002; Westphal, Steffan-Dewenter & Tscharntke 2006; Greenleaf et al. 2007; Zurbuchen et al. 2010; Benjamin, Reilly & Winfree 2014). Small bodied bees (e.g. *Osmia*) forage on the scale of tens to hundreds of meters, whereas bumblebees and honey bees have been recorded to move up to 10 and 14 kilometres, respectively (Zurbuchen et al. 2010 and references therein). This implies that for generalist-pollinated plants, within- and among-population pollen flow may be dominated by different sets of pollinators. This is rarely considered in plant pollen flow models and may offer an explanation as to why pollen dispersal kernels are often fat-tailed or, in some cases, multimodal (Ashley 2010).
Here we test the hypothesis that landscape effects on contemporary pollen flow are scale dependent for the generalist-pollinated herb, *Pulsatilla vulgaris*. The main pollinators of *P. vulgaris* are small-bodied and solitary or semi-social members of the bee genera *Lasioglossum*, *Osmia* and *Andrena*, although flowers are also visited by large bodied and social species including multiple species of *Bombus* queens and *Apis mellifera* (Kratochwil 1988; Fay & Barlow 2014). We hypothesize that within-population pollen flow will be associated with landscape and floral features measured at small spatial scales. In contrast, we hypothesize that rates of pollen immigration into local populations of *P. vulgaris* will be related to landscape features measured at larger spatial scales. Although seed flow is an important form of landscape-scale gene flow in this species (DiLeo et al. 2017), virtually nothing is known about the scale of pollination. We therefore used paternity analysis and tracked realized pollen flow within- and among-populations, including contemporary pollen immigration rates. We asked: (1) Is pollination more likely to occur among near neighbours or does it occur randomly with respect to available fathers; does this differ within and among populations? (2) Which landscape and ecological factors explain within-population pollination distances, selfing, and pollen immigration rates?

**Materials and Methods**

**Study Species**

*Pulsatilla vulgaris* (Ranunculaceae) is a perennial, hermaphroditic herb of conservation concern and a flagship species of calcareous grasslands across central Europe. Among other calcareous grassland species in our study region in the Franconian Jura, *P. vulgaris* is the first to flower,
with flowers opening when temperatures reach above 12-15°C (usually around March-April) and lasting 4-6 weeks (Kratochwil 1988; Hensen, Oberprieler & Wesche 2005). Plants normally produce 1-3 hermaphroditic, purple flowers, although up to 16 flowers per plant have been observed in our study region. *P. vulgaris* is self-compatible, but previous work demonstrated the absence of spontaneous self-pollination in a pollinator exclusion experiment, suggesting that bees are required for successful pollination (Wells & Barling 1971). Flowers are protogynous and exhibit spatial separation of stamens and pistils (herkogamy), however separation of the male and female reproductive structures in both space and time are incomplete (Jonsson, Rosquist & Widen 1991). Thus, self-fertilization is possible through both autogamy (i.e. self-fertilization occurring within the same flower) and geitonogamy (i.e. fertilization occurring between two flowers of the same individual). *P. vulgaris* flowers are pollinated by a variety of bees; predominately by members of Hymenoptera Apoidea (Kratochwil 1988). A study in southwestern Germany found two species, *Lasioglossum lineare* and *Andrena bicolor*, to be the most important pollinators of *P. vulgaris* (Kratochwil 1988), whereas a study in the U.K. found the most important pollinator to be *Osmia bicolor* (Fay & Barlow 2014). In our study region in south-central Germany, *O. bicolor, A. bicolor*, multiple species of *Lassiglossum* (*L. calceatum, L. fulvicorne, L. pauxillum, L. morio*) and queens of *Bombus* (*B. pascuorum, B. pratorum*) have been observed visiting flowers of *P. vulgaris* (K. Weber, Landschaftspflegeverband Landkreis Bamberg e.V., personal obs.), but we have no quantitative information on the community present in the region.

**Study Area and Sampling**

The study area is a 10 x 15 km region in the southern Franconian Jura, Germany (Fig. 1). The region is characterized by a series of valleys and plateaus ranging in elevation from 410-610 m.
The plateaus contain a mix of agricultural fields, forest, grassland and settlements. Calcareous grassland habitat is found on shallow soils on the plateaus, or more typically on steep slopes at the margin of the plateaus and valleys. From April-May 2009, leaf material was collected from 1327 flowering *P. vulgaris* individuals across all 57 populations in the study region. We collected leaves from all flowering individuals in populations that had less than 40 flowering individuals and in seven focal populations where seeds were collected for paternity analysis (see below; Fig. 1). For populations that exceeded 40 individuals (*n* = 18 populations), leaf material was collected from 30-40 randomly chosen plants. For each plant where leaves were collected, we recorded the number of flowers per plant (used to calculate population size and floral density, see below), and GPS coordinates to an accuracy of 1m after post-processing with a portable handheld Trimble GeoExplorer (Trimble, Concord, ON, Canada). For paternity analysis, offspring were collected in the form of mature seeds from 6-9 mother plants from each of seven populations varying in size, isolation, and landscape context (Fig. 1). Focal populations for paternity analysis were selected to represent a variety of population sizes (15-57 flowering plants per focal population) and landscape contexts (varying in proportion of forest and seminatural habitat surrounding focal populations) across the study region. Our original sampling covered 12 populations in pairs that were spatially close but varied in population size; however, seeds from five populations were collected too early and we were not able to excise the underdeveloped embryos for genotyping. Although we did not have complete sampling of all potential fathers within the study region, we sampled all potential fathers within the seven focal populations and neighbouring populations. Thus we could confidently assign fathers within populations and we assumed that unassigned offspring were the product of pollen immigration. Additional *P. vulgaris* populations exist to the south and west of
the study region but this should not influence our results as our analysis methods allowed for the presence of unsampled populations (see below).

Genotyping

We extracted genomic DNA from leaves and embryos (carefully excised from the seed) using the QIAGEN DNeasy Plant Mini Kit following the manufacturer’s protocol (QIAGEN, Mississauga, ON, Canada). Samples were amplified and genotyped at seven species-specific and polymorphic microsatellite markers (pv2, pv7, pv27, pv33, pv56, pv65a, pv65b) using published conditions (DiLeo et al. 2015). Markers did not depart from Hardy-Weinberg equilibrium and were found to be unlinked (see DiLeo et al. 2017).

Reconstructing Contemporary Pollen Flow

We conducted a paternity analysis (including selfing) for offspring using COLONY2 (Wang 2004; Wang & Santure 2009; Jones & Wang 2010). This program reconstructs full and half-sib families using genetic information and has been found to assign a greater proportion of correct paternities than the commonly used categorical assignment program CERVUS, especially when there is incomplete sampling of potential fathers (Walling et al. 2010). COLONY2 allows for errors in genotyping, and we set stochastic error rates individually per locus based on the number of mother-offspring genotype mismatches observed in our progeny arrays (Table S1). We allowed both male and female polygamy, and used the FL-PLS mode with a medium run length and high precision. We chose FL-PLS, which combines the full likelihood method (FL) and pairwise likelihood scores (PLS), to maximize accuracy while reducing computational time as...
suggested by the program author. To test the sensitivity of COLONY2 to assign paternities based on the pool of potential fathers included, we conducted parentage analysis at three spatial scales. First, we conducted separate parentage analyses for each of the seven populations where we had offspring, including potential fathers only within each population. Second, we again conducted separate analyses for each population, but this time included potential fathers from all populations within a 1 km radius. Third, we conducted a single analysis including all potential fathers from all 57 populations within the study region. The assignments did not substantially differ between the three runs, and thus we only present results of the third analysis. To ensure that our genetic markers had sufficient information to discriminate among potential fathers, we calculated exclusion probabilities assuming a single parent is known (Wang 2007) using the program COANCESTRY (Wang 2011). This calculates the probability that an individual chosen randomly from the population is excluded as the father of an offspring.

Based on the paternity assignments, we categorized each offspring as either being selfed, outcrossed from within the population (hereinafter ‘within’), outcrossed from outside the population (i.e. pollen immigrants) or unassigned. COLONY2 gives individual probabilities of assignments, and we used 90% probability as a cut-off for assigned offspring. Offspring that had greater than 90% probability of being assigned to an unsampled father were assumed to be the result of a pollen immigration event. Some offspring were assigned to multiple potential fathers, each with probabilities less than 90%. If the multiple potential fathers were from within the population of the sampled mother and had a joint probability of greater than 90%, the offspring was categorized as “within”. Likewise, if the multiple fathers were from outside the population with a combined probability of greater than 90%, the offspring was categorized as a pollen
immigrant. If the multiple potential fathers were from a combination of within and outside, the
offspring was categorized as unassigned.

Spatial scaling of within- and among-population pollen flow

To evaluate the spatial scale of pollination within populations and to test if pollination is
more likely to occur among near neighbours or occurs randomly with respect to available fathers,
we compared observed mean outcrossed pollination distances to those expected under each
hypothesis. Here we were not interested in modelling the exact parameters of a dispersal kernel,
but interested in whether observed pollen flow conformed more to a pattern of distance decay or
randomness. Null distributions were generated for each population separately by sampling $x$
distances (with replacement) between focal mothers and all potential fathers within populations,
with either a probability of $1/distance$ in meters (distance decay hypothesis) or a uniform
probability (randomness hypothesis), where $x$ equalled the number of empirically sampled
offspring per focal mother (see Table 1 and Appendix S1). We took the mean of sampled
distances for each of 999 permutations. One-sided p-values were calculated as the proportion of
permutations where the difference in sampled pollination distance was greater than or equal to
the observed mean pollination distance (distance decay hypothesis) or less than or equal to the
observed mean pollination distance (randomness hypothesis).

To evaluate the spatial scale of among-population pollination we repeated the above
analysis but generated expected pollination distances by sampling distances between focal
populations and all potential source populations within the study region. Because we were only
able to assign a small number of among-population pollination events ($n=17$), we combined
mean distances across populations. All statistical analyses were performed in R 3.2.3 (R Core Development Team, 2015).

**Predictors of pollen flow**

We used mixed models to evaluate the role of local conspecific floral density and isolation of mother plants on selfing rates, pollination distances of within-population outcrossed offspring, and correlated paternity (Robledo-Arnuncio, Austerlitz & Smouse 2007). Local floral density was measured as the number of flowers within two meters of the mother plant, and mother isolation was measured as the mean distance of the mother plant to all other plants within the population (mean neighbour distance). A radius of two meters around mother plants was chosen as it gave the strongest correlation with selfing rates and pollination distances compared to lower (1 m) and higher (3 m) tested values. The effects of floral density and mother isolation on selfing was evaluated with a generalized mixed model with a binomial error distribution. Each offspring was entered into the model as a binary response of either selfed (1) or outcrossed (0). To control for multiple sampling of offspring per mother and per population, we included a nested random effect (1|population/mother), where offspring were nested within mothers, which were nested within populations. The effects of floral density and mother isolation on observed within-population pollination distances were evaluated with a linear mixed effects model with the same nested random effect. Both the predictors and response of the linear mixed effect model were log transformed to linearize relationships.

The effects of floral density and mother isolation on correlated paternity were also evaluated with a linear mixed effects model. Correlated paternity measures the proportion of
offspring pairs, sampled from a single mother plant, that share the same father (Ritland 1989), and was calculated using KINDIST in POLDISP 1.0 (Robledo-Arnuncio et al. 2007). Because correlated paternity represents a single value per mother, only population was included as a random effect in the model. Linear mixed models were estimated using maximum likelihood for model selection and all predictors were scaled to allow comparison of regression coefficients. For each response variable, we tested all combinations of predictors and chose the best model based on AICc. We assessed significance of the best model with likelihood ratio tests, and calculated variance explained by fixed effects (marginal $R^2$) using the protocol of Nakagawa and Schielzeth (2013). Generalized mixed effects models and linear mixed models were estimated using the lme4 (Bates et al. 2014) package in R.

At the level of populations, we used two logistic regressions with binomial-distributed error to evaluate the effect of population size, population isolation, and landscape composition on the proportions of selfed and outside sired offspring per population. Population size was counted as the number of flowering plants per population. Population isolation was measured as the mean distance of each focal population to all others in the study region (mean neighbour distance). We were most interested in two types of landscape features: seminatural habitat and forest. Here we consider seminatural habitat to include permanent grasslands, groves, mires, barren land, and orchards following the classification of Steffan-Dewenter et al. (2002). Seminatural habitat provides suitable nesting and foraging habitat for wild bees and has been shown to strongly predict abundance and diversity of bees in European landscapes (Steffan-Dewenter et al. 2002; Ockinger & Smith 2007; Le Feon et al. 2010). Thus, we used the amount of seminatural habitat as a proxy for local pollinator abundance and expected the proportion of selfed offspring to be lower, and pollen immigration to be higher, in populations surrounded by...
more seminatural habitat. Forest can alter the detectability of plant populations to pollinators and has been found to act as a barrier to dispersal for bumblebees (Kreyer et al. 2004). Therefore, we expected the proportion of selfed offspring to be higher, and the proportion of pollen immigration to be lower in populations surrounded by more forest.

We further expected the proportion of selfed offspring and pollen immigration to be related to landscape composition at different spatial scales. The main pollinators of *P. vulgaris* are small-bodied bees, which have limited dispersal ability (180-600 m for bees of similar size; Gathmann & Tscharntke 2002; Zurbuchen et al. 2010) and thus likely contribute strongly to local pollination patterns (i.e. decreasing selfing rate), but not to among-population pollen flow (i.e. increasing pollen immigration rate). In contrast, large bee visitors to *P. vulgaris* such as those in the genera *Bombus* and *Apis* are able to disperse long distances (up to 10 and 14 km, respectively; Zurbuchen et al. 2010) and likely contribute to patterns of among-population gene flow. At the time of flowering, only the queens of *Bombus* are flying, which have dispersal distances up to three times those of *Bombus* workers (Dreier et al. 2014; Carvell et al. 2017).

Small-bodied and solitary bees tend to perceive the landscape at much finer spatial scales than large-bodied and social bees with large foraging ranges (Steffan-Dewenter et al. 2002; Westphal, Steffan-Dewenter & Tscharntke 2003, 2006; Knight et al. 2009; Benjamin, Reilly & Winfree 2014). Thus, we expected that the proportion of selfed offspring would correlate most strongly to landscape measured at small spatial scales, and the proportion of pollen immigration to correlate most strongly to landscape measured at large spatial scales. To explore these potential scale-dependent effects, we quantified landscape composition at five nested spatial scales, at buffers of 50, 100, 250, 500, and 1000m around population centroids. We quantified the proportion of
forests and seminatural habitat within each buffer using digital land use maps (Tatsaechliche Nutzung ALKIS 2008-2009 database; Bayerische Vermessungsverwaltung, Munich, Germany).

Logistic regressions were implemented as generalized linear models (GLM) with binomial error distribution and logit links. Because we only have a sample size of $n = 7$ focal populations, we tested the effect of each predictor separately (population size, population isolation, proportion of forest at five spatial scales, proportion of seminatural habitat at five spatial scales) on the proportion of selfed and immigrant offspring per population to avoid overfitting the data. However, more than one predictor was found to be significantly associated with the proportion of immigrant offspring, and so we further conducted multiple regressions including a maximum of two predictors from the above analysis per model, as an exploratory measure. Pearson correlations between the predictor variables ranged from 0.008-0.95, with an average correlation of 0.45 across predictors (Table S2). The most highly correlated predictors were proportion of forest measured at 250 and 500 m radii. Population size and isolation were not highly correlated with predictors of landscape context (Table S2). We tested each model for overdispersion by dividing the residual deviance by the residual degrees of freedom. All models that included the proportion of selfed offspring as the response variable exhibited overdispersion. We thus corrected standard errors using quasi-binomial models (Bolker et al. 2009). We assessed model significance by conducting Chi-square analysis of deviance, or F-tests for models exhibiting overdispersion. We selected the best buffer radius (50m, 100m, 250m, 500m, 1000m) for each landscape predictor/response pair by choosing the model with the lowest residual deviance and lowest $AIC_c$ (proportion of immigrant) or quasi-$AIC_c$ ($QAIC_c$; proportion of selfed). $QAIC_c$ was calculated for the overdispersed proportion of selfed offspring data using the
Results

Reconstructing contemporary pollen flow

All offspring from four mothers in population A03, from one mother in A41, and from two mothers in G05a were aborted and were thus excluded from analysis. We also excluded samples that failed to amplify (i.e. could not be successfully genotyped) at more than two loci. In total, 394 offspring were genotyped and included in the paternity analysis: final sample sizes per population are shown in Table 1. Our markers were highly polymorphic with 11-33 alleles per locus and provided a multi-locus exclusion probability of 0.999, suggesting that our markers had sufficient information to discriminate among potential fathers (Table S1). COLONY2 assigned 382 offspring to a single father (sampled or unsampled) with greater than 90% probability, and twelve offspring were assigned to multiple potential fathers with individual probabilities of less than 90%. Ten of these twelve offspring were categorized as either within-population pollen flow or pollen immigration events based on the rules described in the methods. In total 201 offspring were selfed, 129 offspring were sired by fathers from within the same population, 62 were sired by fathers from outside the population of the sampled mother (i.e. pollen immigration), and two offspring were unassigned and excluded from further analysis. In 17 of 62 cases of pollen immigration, we could assign a sampled father and thus reconstruct realized among-population pollen flow.
Spatial scaling of within- and among-population pollen flow

Distributions of realized pollen flow distances are shown in Figure S1. Mean distances of realized pollen flow within populations ranged from 2.66 – 9.49 m, and mean among-population pollen flow was 6.6 km (Table 2). Permutation tests revealed that for all populations except A26, mean within-population pollination distance conformed to a pattern of distance decay (Table 2; Fig. S2). In contrast, mean within-population pollination distance in population A26 and mean among-population pollination distance conformed to patterns of randomness (Fig. 2; Fig. S2).

Predictors of pollen flow

An initial screening of the data showed that floral density and mean neighbour distance affected pollination distances in opposite directions in A25 compared to the rest of the studied populations. Removing A25 from the linear mixed effect models for mean pollination distance did not alter the outcome of model selection, but did strengthen relationships with the predictor variables. We thus present results with population A25 excluded here. The best model explaining pollination distances showed that pollination distance significantly decreased with increasing local floral density around maternal plants and increased with mean neighbour distance (marginal $R^2=0.5$; Likelihood ratio test, $\chi^2(1)=19.9, p<0.001$; Table S3, Fig. 3). Tests for the significance of individual fixed factors found that floral density had a significant effect on pollination distances (Likelihood ratio test, $\chi^2(1)=5.13, p=0.02$), but mean neighbour distance had a non-significant effect (Likelihood ratio test, $\chi^2(1)=2.4, p=0.12$). Repeating the analysis for

This article is protected by copyright. All rights reserved.
population A25 alone showed that mean neighbour distance had a significant negative relationship with pollination distance (marginal $R^2=0.56$; Likelihood ratio test, $\chi^2(1)=5.4$, $p=0.02$). There was a non-significant positive relationship between local floral density and pollination distance in A25 (marginal $R^2=0.05$; Likelihood ratio test, $\chi^2(1)=0.38$, $p=0.54$).

Neither local floral density nor mean neighbour distance explained significant variation in selfing or correlated paternity, with the null (random effects) model being selected as the most likely for both predictors (Table S3).

The proportion of selfed offspring per population showed a significant negative relationship with population size (number of flowering plants) and was not significantly associated with any other predictor variable (Table 3; Fig. 4). In contrast, the proportion of immigrants per population showed significant relationships with several predictors, including positive relationships with population size, population isolation, and seminatural habitat at 1000 m, and negative relationships with forest at 100-1000 m radii (Table 3). The best predictor of the proportion of immigrants per population (selected as the model with the lowest residual deviance and lowest $\text{AIC}_c$) was the proportion of forest at 500 m (Table 3; Fig. 4). Multiple regressions including up to two predictors per model for the proportion of immigrants were not well supported, exhibiting high $\text{AIC}_c$ values compared to single predictor models (Table S4).

**Discussion**

We assessed the multi-scale determinants of contemporary pollen flow in *P. vulgaris* across seven populations varying in size and landscape context. We found that within-population pollen
flow conformed to a pattern of near neighbour mating, but importantly this did not translate to the landscape scale – among-population pollen flow was not more likely to occur amongst neighbouring populations but rather occurred randomly with respect to available source populations. We further found that patterns of within- and among- population pollen flow were best explained by different ecological and landscape factors: within-population pollen flow distances and the proportion of selfed offspring per population were best explained by the availability of floral resources, whereas pollen immigration rates were best explained by landscape context at large spatial scales. Together these results suggest that within- and among- population patterns of contemporary pollen flow may be governed by different underlying processes and highlight the importance of using multi-scale approaches when modelling gene flow in generalist-pollinated plants.

Pollen flow as a multi-scale process

We found that patterns of near-neighbour mating at the local scale did not translate to the landscape scale (Fig. 2), suggesting that within- and among-population pollen flow are governed by different underlying processes. These scale-dependent effects match well with expectations based on the foraging ranges of *P. vulgaris* pollinators. The main pollinators of *P. vulgaris* are small-bodied bees (4-10 mm; Meyer 2007) with estimated maximum dispersal abilities within the range of 180-600 m (Gathmann & Tscharntke 2002; Zurbuchen et al. 2010), which are likely to dominate local pollination but are unable to move pollen between populations. In contrast, less frequent large-bodied visitors of *P. vulgaris* such as *Apis mellifera* and the queens of *Bombus pascuorum* and *B. pratorum* have estimated foraging ranges of several kilometres (Knight et al.
2009; Lepais et al. 2010; Zurbuchen et al. 2010; Carvell et al. 2017). Our high observed mean among-population pollen flow distance (6.6 km) suggests that although pollen immigration is rare compared to selfing and within-population outcrossing, these large-bodied bees might play a critical role in maintaining gene flow at the landscape scale. However, it should be noted that we could reconstruct among-flow pollen distance in only 17 of 62 cases and thus the true scale of pollen flow remains unknown. Despite this, very few among-population pollen flow events were assigned to neighbouring populations that were exhaustively sampled (e.g. no pollen flow between A25 and A26; Fig. 1), suggesting that pollen immigrants are arriving from further away in the landscape. This result adds to growing evidence that contemporary pollen flow can occur at surprising large spatial scales (Lander, Boshier & Harris 2010; Buehler et al. 2012; Kremer et al. 2012; Noreen et al. 2016), and that different rules need to be applied to the tail end of dispersal kernels (Nathan 2006; Slavov et al. 2009; Garcia & Borda-de-Agua 2017).

We cannot exclude the possibility that our results reflect multiscale habitat selection or resource use by the same set of pollinators. For example, a single species may use different strategies for foraging locally and for long-distance movements (Keller & Holderegger 2013), which could result in multimodal dispersal kernels (e.g. Kleyheeg et al. 2017). Another hypothesis that has been offered to explain the breakdown of distance decay in pollen dispersal kernels over larger spatial scales is that pollinators are limited by their perceptual range (Stacy et al. 1996; Dick, Etchelecu & Austerlitz 2003); it is easy for a pollinator to find the next nearest plant within a population where floral density is high, but when a pollinator leaves a patch their limited perceptual range might prevent them from finding the next nearest patch or population. It is thus possible that the large-bodied pollinators of *P. vulgaris* also contribute to patterns of within-population pollination where sequential floral visits produce steep dispersal kernels.
within but not between populations. Field observations from other parts of the range of *P. vulgaris* showed that large-bodied bees of the genera *Bombus* make up less than 3% and *Apis* less than 12% of total visits to flowers and thus should not contribute as strongly to within-population pollen flow than its main small bodied pollinators (Kratochwil 1988), however further studies are needed to specifically characterize the pollinator community in our study region.

As an important counterpoint, a recent study by Castilla *et al.* (2017) found that small-bodied bees contributed just as much as large-bodied bees to long-distance pollen flow in the tropical understory tree *Miconia affinis*. This suggests that bee body sizes do not always predict pollination distances. It is unclear how general a pattern this might be as very few studies have explicitly linked direct observations of pollinators and their traits to reconstructed pollen flow distances over large spatial scales (but see Ahmed *et al.* 2009; Rhodes, Fant & Skogen 2017).

Future work should aim to do this for other generalist-pollinated plants, using, for example, the approach of Castilla *et al.* (2017). They linked traits on an entire community of pollinators to pollen flow distances by observing single pollination events in the field, collecting the pollinator for later identification, and then excluding other pollinators while fruits developed prior to genotyping and paternity assignment.

*Beyond distance effects*

We found that geographic distance among plants and populations (i.e. mother and population isolation) were poor predictors of pollen flow in *P. vulgaris*. This result adds to a growing literature supporting a more functional approach to modelling gene flow in plant populations by
considering factors beyond distance that affect the movement and habitat use of seed and pollen vectors (Robledo-Arnuncio et al. 2014; Auffret et al. 2017; Krauss et al. 2017). Although within-population pollen flow decayed with distance for *P. vulgaris*, this pattern was best explained by local floral density. Likewise, population selfing rates were associated with population size (i.e. the number of flowering individuals). In both cases, the relationship was negative; large populations tended to have lower overall selfing rates, but mothers surrounded by more flowers within populations suffered from reduced outcrossing distances (Figs 3-4). Interestingly, we found that in the population with the most flowers (A25), pollination distance increased with local floral density (Fig. 3a). This might reflect scale-dependent optimal foraging of pollinators. For example, bee visitation rates and spillover of pollination to nearby flowers tends to increase with floral density, but at very high densities this relationship can flip due to increased competition (Goulson 2000; Veddeler, Klein & Tscharntke 2006; Hegland 2014). It should be noted that although we recorded if individuals were flowering or not, we did not assess the developmental stage of anthers and it is thus possible that our measure of local floral density overestimated the true availability of pollen donors.

Pollen immigrations rates were best explained by landscape context, showing a negative relationship with proportion of forest measured at a 500 m radius (Table 3; Fig. 4). This result is supported by evidence from the pollinator literature showing that large-bodied bee abundance (Westphal, Steffan-Dewenter & Tscharntke 2006), and pollinator services (Kremen et al. 2004; Benjamin, Reilly & Winfree 2014) are associated with landscape context and floral resources at large spatial scales. It is unclear if the negative relationship between forest cover and pollen immigration reflects an effect of landscape resistance or makes populations harder to detect for overhead bees (Kreyer et al. 2004). Kamm et al. (2010) found that open areas promoted

This article is protected by copyright. All rights reserved.
contemporary pollen flow among individuals of the insect-pollinated tree *Sorbus domestica*, suggesting that open landscape enhances movement of pollinators. Similarly, Aavik *et al.* (2017) found a negative relationship between allelic richness and forest at a 500 m radius around populations of the insect-pollinated herb *Rhinanthus osiliensis* suggesting that forest has a negative effect on among-population bee movement and subsequent gene flow. On the other hand, woody habitat and forest edges provide suitable nesting spots for some bee species (Kells & Goulson 2003; Somme, Mayer & Jacquemart 2014), and thus a minimum level of forest might be required to support populations. However, we caution that although forest was the best predictor, population size and seminatural habitat at 1000 m also had high support, with delta AIC$_c$ values less than two in model selection (Table 3). We found no support for additive effects (Table S4), but we caution that with a sample size of $n = 7$ focal populations the power to parse potential additive effects is limited. Future work sampling a wider landscape context with more focal populations will help to clarify our results.

**Conclusions**

Here we show that pollen flow in *P. vulgaris* is a multiscale process, and that differential patterns of within- and among-population pollen flow are likely driven by the foraging of distinct groups of pollinators over different spatial scales, or multiscale resource use of large-bodied bees. It remains to be seen how general this pattern is across other bee-pollinated species (e.g. see Opedal *et al.* 2017; Castilla *et al.* 2017), and other key aspects of life history and behaviour of vectors may need to be considered (e.g. grooming behaviour and optimal foraging; Holmquist, Mitchell & Karron 2012; Mitchell *et al.* 2013; Olsson *et al.* 2015). However, we show that
drawing on knowledge of the spatial scaling of habitat use of bees provides a good starting point for testing hypotheses about the drivers of contemporary pollen flow, and contributes to a more mechanistic understanding of commonly observed patterns of multimodality in pollen dispersal kernels.

**Acknowledgements**

Funding was provided by the Natural Sciences and Engineering Council of Canada (NSERC CGS-D and Michael Smith Foreign Study Supplement to M.F.D., Discovery Grant to H.H.W, and NSERC-CREATE ADVENT), an Ontario Graduate Scholarship to M.F.D., and the Government of Central Franconia, Bavaria, Germany. We thank Amaneet Lochab, Michael Liu, Qasim Muhammad, Julie Wang, and Hardikkumar Patel for assistance with DNA extraction and genotyping, Yessica Rico, Henry Lehnert, Juergen Boehmer for support in the field, and Klaus Weber for assistance with pollinator identification. We thank Marie-Josée Fortin, Marc Johnson, Rob Ness, and Nusha Keyghobadi for comments on an earlier version of this manuscript. We thank Ignasi Bartomeus and two anonymous reviewers for comments that greatly improved the paper.

**Author’s contributions**

HHW initiated the study and conducted the field work. MFD conducted the lab work, analysed the data with contributions from RH and HHW, and wrote the manuscript. All authors contributed to the editing of the manuscript and gave final approval for publication.
Data accessibility

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.3d9r8p1 (DiLeo et al., 2018)

References


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


Table 1: Sample sizes for the seven *Pulsatilla vulgaris* populations where paternity analysis was conducted. All flowering individuals were genotyped within the seven populations. Population size (number of flowering plants), and the number of offspring successfully genotyped across mothers and flowers per population are shown. The number of offspring assigned as selfed and pollen immigrants based on COLONY2 paternity analysis are given. For population locations see Figure 1.

<table>
<thead>
<tr>
<th>Population</th>
<th>Number of Flowering Plants</th>
<th>Sample Sizes</th>
<th>Number selfed</th>
<th>Number pollen immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mothers</td>
<td>Flowers</td>
<td>Genotyped offspring</td>
</tr>
<tr>
<td>A03</td>
<td>44</td>
<td>2</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>A21</td>
<td>21</td>
<td>6</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>A25</td>
<td>57</td>
<td>8</td>
<td>8</td>
<td>79</td>
</tr>
<tr>
<td>A26</td>
<td>22</td>
<td>7</td>
<td>7</td>
<td>60</td>
</tr>
<tr>
<td>A41</td>
<td>15</td>
<td>9</td>
<td>9</td>
<td>78</td>
</tr>
<tr>
<td>A45</td>
<td>22</td>
<td>6</td>
<td>6</td>
<td>55</td>
</tr>
<tr>
<td>G05a</td>
<td>46</td>
<td>6</td>
<td>6</td>
<td>31</td>
</tr>
</tbody>
</table>
Table 2: Observed mean within-population pollination distances for each population, observed mean among-population pollination distance (pooled across all populations), and expected mean pollination distances based on null hypotheses of distance decay and randomness. Standard errors (SE) of null distributions are in brackets, and p-values of permutation tests are shown.

<table>
<thead>
<tr>
<th>Population</th>
<th>observed mean pollination distance (m)</th>
<th>null: distance decay</th>
<th>expected mean pollination distance (m) (SE)</th>
<th>p-value</th>
<th>null: randomness</th>
<th>expected mean pollination distance (m) (SE)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A03</td>
<td>9.49</td>
<td></td>
<td>9.79 (0.06)</td>
<td>0.55</td>
<td></td>
<td>14.01 (0.06)</td>
<td>0.008</td>
</tr>
<tr>
<td>A21</td>
<td>7.55</td>
<td></td>
<td>10.78 (0.17)</td>
<td>0.73</td>
<td></td>
<td>21.04 (0.21)</td>
<td>0.005</td>
</tr>
<tr>
<td>A25</td>
<td>3.31</td>
<td></td>
<td>4.77 (0.02)</td>
<td>0.99</td>
<td></td>
<td>9.80 (0.04)</td>
<td>0.001</td>
</tr>
<tr>
<td>A26</td>
<td>4.86</td>
<td></td>
<td>4.24 (0.01)</td>
<td>0.05</td>
<td></td>
<td>4.97 (0.01)</td>
<td>0.41</td>
</tr>
<tr>
<td>A41</td>
<td>2.74</td>
<td></td>
<td>2.75 (0.01)</td>
<td>0.47</td>
<td></td>
<td>3.79 (0.02)</td>
<td>0.005</td>
</tr>
<tr>
<td>A45</td>
<td>4.26</td>
<td></td>
<td>6.81 (0.04)</td>
<td>0.99</td>
<td></td>
<td>12.73 (0.08)</td>
<td>0.001</td>
</tr>
<tr>
<td>G05a</td>
<td>2.66</td>
<td></td>
<td>3.33 (0.04)</td>
<td>0.68</td>
<td></td>
<td>6.34 (0.08)</td>
<td>0.009</td>
</tr>
<tr>
<td>Among</td>
<td>6655.50</td>
<td></td>
<td>2809.29 (24.22)</td>
<td>0.001</td>
<td></td>
<td>6112.23 (22.72)</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Table 3: Estimates for population-level logistic regressions, showing the relationship of the proportion of selfed offspring or proportion of pollen immigrants per population, with population size (number of flowering plants), population isolation (mean nearest neighbour distance) and landscape composition (proportion of forest and seminatural habitat) measured at increasing radii from population centroids. Estimates are scaled to allow comparison across predictors. Standard errors of model estimates are given (SE).

The best fitting model for each predictor type was selected as the model with the lowest residual deviance and is indicated in bold. Model significance was assessed with analysis of deviance with F-tests (proportion selfed) or Chi-square tests (proportion immigrant) and significant models are marked with asterisks.
### Table 3 continued

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Proportion Selfed</th>
<th>Proportion Immigrant</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (SE)</td>
<td>Deviance</td>
</tr>
<tr>
<td>population size</td>
<td>-0.63 (0.26)</td>
<td>30.1*</td>
</tr>
<tr>
<td>population isolation</td>
<td>-0.32 (0.47)</td>
<td>60.0</td>
</tr>
<tr>
<td>proportion of forest measured at:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 m</td>
<td>0.57 (0.30)</td>
<td>38.6</td>
</tr>
<tr>
<td>100 m</td>
<td>0.45 (0.29)</td>
<td>43.8</td>
</tr>
<tr>
<td>250 m</td>
<td>0.41 (0.33)</td>
<td>50.6</td>
</tr>
<tr>
<td>500 m</td>
<td>0.48 (0.37)</td>
<td>49.3</td>
</tr>
<tr>
<td>1000 m</td>
<td>0.38 (0.40)</td>
<td>55.6</td>
</tr>
<tr>
<td>proportion of seminatural habitat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>measured at:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 m</td>
<td>-0.11 (0.47)</td>
<td>64.5</td>
</tr>
<tr>
<td>100 m</td>
<td>-0.13 (0.36)</td>
<td>63.7</td>
</tr>
<tr>
<td>250 m</td>
<td>0.18 (0.36)</td>
<td>62.4</td>
</tr>
<tr>
<td>500 m</td>
<td>0.16 (0.42)</td>
<td>63.4</td>
</tr>
<tr>
<td>1000 m</td>
<td>-0.41 (0.37)</td>
<td>53.2</td>
</tr>
</tbody>
</table>

* P < 0.05; **P < 0.01
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1365-2745.12992
This article is protected by copyright. All rights reserved.
Figure 1: Study area in the Franconian Alb, Germany, showing all known populations of *Pulsatilla vulgaris* in the area and the main land use categories of forest and seminatural habitat. Focal populations where seeds were collected for paternity analysis are indicated by squares and labelled by name. Leaves were collected from all populations, and the fill of each circle indicates the sampling effort; all flowering individuals were sampled in black populations, and only a subset of 30-40 individuals in white populations. Population size (approximate number of flowering individuals) is indicated by the size of circles. See Table 1 for population sizes of focal populations.
Figure 2: Histograms showing permuted null distributions of mean pollination distances under null hypotheses of distance decay (dark grey bars) and randomness (light grey bars) and their overlap (medium grey) for an example of within-population pollen flow in population A41 (a) and among-population pollen flow (b). Observed mean pollination distances are shown with the vertical dotted lines. Null distributions and observed mean pollination distances for the other populations are shown in Figure S2 and p-values of permutation tests are given in Table 2.
Figure 3: Scatterplots showing the relationships of within-population pollination distances and number of flowers within 2 m of the mother plant (a), and mother isolation (mean neighbour distance; b). Predictors and pollination distances were log transformed to linearize relationships. Square symbols and dashed lines represent relationships for population A25, and circles and solid lines represent relationships for all other populations. See Table S3 for results of model selection on these data.
Figure 4: Scatterplot showing the best supported relationships of logistic regressions; population size (number of flowering plants) and the proportion of selfed offspring per population (a), and proportion of forest measured within 500 metres of focal populations and the proportion of pollen immigrants per population (b). The size of the points represents population size. Solid lines represent fitted values from logistic regression. See Table 3 for results of model selection on these data.