

NEWS AND VIEWS

COMMENT

A comment on García *et al.* (2005, 2007) and related papers on mating patterns and gene dispersal in *Prunus mahaleb*

C. M. HERRERA

Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas (CSIC), Avenida Américo Vespucio s/n, E-41092 Sevilla, Spain

In two studies on mating patterns and spatial components of pollen and seed dispersal of *Prunus mahaleb* based on parentage analysis, García *et al.* (2005, 2007) depicted their 196 focal trees as a spatially isolated population where all reproductive trees had been genotyped. Additional distributional data for *P. mahaleb* trees in their study area, however, revealed that García and colleagues' depiction of their study system bears little resemblance to reality. The trees these authors studied did not form a discrete, geographically isolated population. Around 300 ungenotyped reproductive trees occurred within the 1.5-km distributional gap to the nearest population proclaimed by García and colleagues. Since exhaustive sampling of potential parental genotypes is essential in parentage analyses, the occurrence of a large number of ungenotyped trees in the immediate neighbourhood of focal trees can severely affect the main conclusions of García *et al.* (2005, 2007) as well as of several related publications on gene dispersal and mating patterns of *P. mahaleb* conducted on the same trees and relying on the same false premises of spatial isolation and exhaustive sampling.

Keywords: gene flow, parentage analysis, paternity analysis, pollen dispersal, seed dispersal

Received 20 April 2009; revision received 18 May 2009; accepted 19 May 2009

Parentage analysis of progeny arrays based on highly-variable molecular markers has substantially contributed to our understanding of mating systems and gene dispersal patterns in wild populations of plants (Jones & Ardren 2003). One of the methods commonly used, known as parentage analysis, involves genotyping samples of seeds or seedlings produced by maternal parents of known or

unknown genotypes, and then probabilistically inferring the identity of paternal parents by the application of several statistical procedures (Jones & Ardren 2003; Kalinowski *et al.* 2007). For instance, paternity analysis to infer the contemporary dispersal of pollen requires that the genotypes of all or most individuals potentially siring the offspring of focal maternal parents are known, or at least a good estimate of the proportion of candidate parents sampled. As shown by recent studies, this crucial requirement is most easily met in discrete, relatively small populations that are spatially isolated from other conspecific populations. In such situations, the true fathers of the offspring from maternal families can be identified exactly (Krauss 1999; García *et al.* 2005, 2007; Robledo-Arnuncio & Gil 2005; Smouse & Robledo-Arnuncio 2005; Robledo-Arnuncio & García 2007).

García *et al.* (2005, 2007) used microsatellite markers and parentage analysis to study the mating patterns and the spatial components of pollen and seed dispersal in a group of 196 individuals of the insect-pollinated, vertebrate-dispersed, gynodioecious tree *Prunus mahaleb* (Rosaceae). García *et al.* (2005: 1822–1823) described their study trees as a 'highly isolated population where all trees have been sampled', also highlighting that 'this represents a thorough sample of the whole population and we are confident that it includes all the reproductive trees'. They also noted that the nearest *P. mahaleb* population in the study region was located at 1.5 km distance and that the area had been exhaustively sampled. García *et al.* (2007: 1948) provided a similar description of the study system, likewise depicting the stand of trees studied as an isolated population separated from the nearest other population by 1.5 km.

I found that García *et al.*'s (2005, 2007) depiction of their study system bears little resemblance to reality and that their study trees do not form a discrete, geographically isolated population. As shown in Fig. 1, the 196 *P. mahaleb* trees studied by García and colleagues, far from being a geographically isolated group, are part of a larger population extending in a southwest-northeast direction for about 2 km. Around 300 adult reproductive trees occur within the 1.5 km distributional gap proclaimed by García and colleagues (Fig. 1). In August 2008, when I found and mapped these additional trees, all of them bore fruits, empty fruit pedicels, and/or inflorescence remains, thus attesting their reproductive status. Furthermore, the heights (>3 m) and trunk diameters (mostly between 20–40 cm) of these trees were essentially the same as those of the trees studied by García *et al.* (2005, 2007), which points to their similarity in terms of age and, presumably, population history (see Appendix S1).

Did the ignoring of this large number of ungenotyped parents in the immediate neighbourhood of genotyped focal trees affect the main conclusions of García *et al.*

Correspondence: Carlos M. Herrera, Fax: + 34 954 62 11 25; E-mail: herrera@cica.es

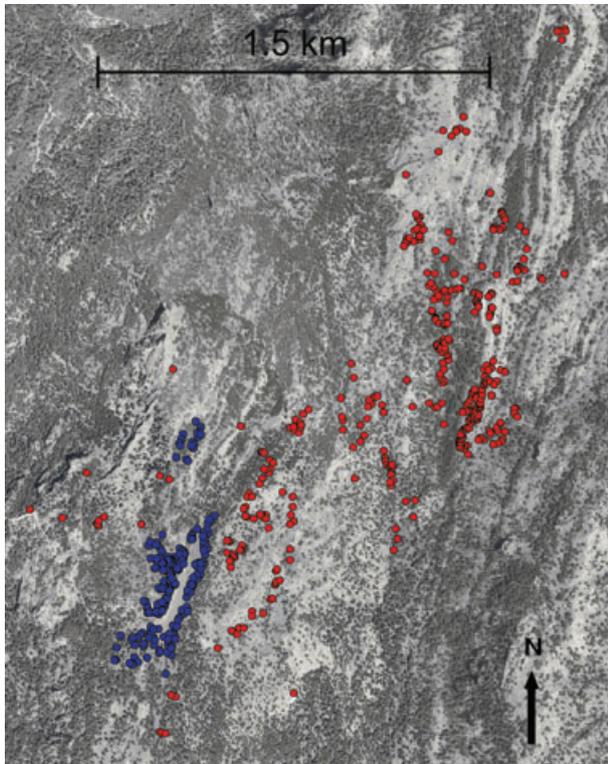


Fig. 1 Aerial photograph showing the distribution of *Prunus mahaleb* trees (dots) in the study area of García *et al.* (2005, 2007) in the Sierra de Cazorla, southeastern Spain. Blue dots represent the locations of the trees studied by the above authors ($N = 196$; taken from Fig. 1 in García *et al.* 2007); red dots are the $N = 391$ additional adult trees found in the area (geographical coordinates of these additional trees are available in Appendix S1).

(2005, 2007)? Only corresponding studies applying exhaustive sampling and genotyping of all the additional trees in the area could answer this question with certainty. Even in the absence of such an experiment, some consequences can be anticipated, for it is well known that exhaustive sampling of potential parental genotypes is essential for the reliability of paternity, maternity or parentage studies (Jones & Ardren 2003; Robledo-Arnuncio & Gil 2005; Smouse & Robledo-Arnuncio 2005; Robledo-Arnuncio & García 2007). The presence of a significant number of ungenotyped reproductive trees in the neighbourhood of focal ones can have at least the following two major implications for the results of García *et al.* (2005, 2007).

Firstly, the fraction of progeny for which true parents were assigned was likely overestimated or, to put it in another way, a certain proportion of parental assignments were probably erroneous. Such an effect could be quantitatively important, as suggested by simple simulations of paternity analyses performed using the program CERVUS 3.0.3 (Kalinowski *et al.* 2007). Using the allelic frequencies for nine microsatellite loci in the example file

furnished with the program, I simulated father assignment rates given a known mother for two situations: (a) sampling 104 (95%) of 109 trees, i.e. a situation akin to that purported in García *et al.* (2005, 2007), treating their 104 hermaphrodite trees as the only potential pollen donors; and (b) sampling 104 (33.3%) of 312 trees (i.e. 587 total trees in the area as given in Fig. 1 \times 0.531 hermaphrodite fraction). The expected proportion of unassigned progeny rose from 3% in case (a) to 63% in case (b). The rise in the uncertainty in parental assignment was even greater for estimates without information on maternal genotypes (from 4% to 88%), as is the case in parentage analysis of dispersed seeds with unknown maternal parents. A second likely consequence, stemming from both the uncertainty in parental assignments and the spatial proximity of many ungenotyped potential parents, is that rates of pollen and seed immigration into the studied population (i.e. long-distance gene immigration), and the distances over which gene flow occurred, were probably overestimated. For example, García *et al.* (2007) estimated that around 20% of dispersed seeds were derived from other populations more than 1.5 km away, because their presumed parental genotypes did not match any of the genotyped focal trees. Many of these presumed long-range immigrants, however, might have been the progeny of nearby-growing but ungenotyped parents (Fig. 1).

The potential consequences of the two above described effects should also be kept in mind in relation to several related studies likewise based on parentage analysis and conducted on the same stand of *P. mahaleb* trees and, apparently, the same seed progenies as studied by García *et al.* (2005, 2007). Godoy & Jordano (2001), Jordano *et al.* (2007), Robledo-Arnuncio & García (2007) and Fortuna *et al.* (2008) all assumed spatial isolation and exhaustive genotyping of potential parents. As these premises have been proven here to be false, the conclusions of these studies with respect to mating patterns and pollen and seed dispersal should also be interpreted with caution. For example, estimates of seed immigration rates from other populations presented by Godoy & Jordano (2001) and Jordano *et al.* (2007) are probably inflated by the large number of ungenotyped trees in the immediate neighbourhood. Likewise, the spatial organization of pairwise mating events ('spatial mating networks') reported by Fortuna *et al.*'s (2008) might be severely distorted by erroneous or uncertain parentage assignments. From all the preceding considerations it appears that a 'major remaining challenge in parentage analysis is to obtain appropriate and complete field samples' and 'that more effort should be devoted to creative sampling practices (i.e. spend more time in the field)' (Jones & Ardren 2003: 2521).

Acknowledgement

Comments from two anonymous reviewers contributed to improve the manuscript.

References

- Fortuna MA, García C, Guimaraes PR, Bascompte J (2008) Spatial mating networks in insect-pollinated plants. *Ecology Letters*, **11**, 490–498.
- García C, Arroyo JM, Godoy JA, Jordano P (2005) Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population. *Molecular Ecology*, **14**, 1821–1830.
- García C, Jordano P, Godoy JA (2007) Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: patterns in distance and direction. *Molecular Ecology*, **16**, 1947–1955.
- Godoy JA, Jordano P (2001) Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. *Molecular Ecology*, **10**, 2275–2283.
- Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. *Molecular Ecology*, **12**, 2511–2523.
- Jordano P, García C, Godoy JA, García-Castaño JL (2007) Differential contribution of frugivores to complex seed dispersal patterns. *Proceedings of the National Academy of Sciences USA*, **104**, 3278–3282.
- Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.
- Krauss SL (1999) Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). *Molecular Ecology*, **8**, 217–226.
- Robledo-Arnuncio JJ, García C (2007) Estimation of the seed dispersal kernel from exact identification of source plants. *Molecular Ecology*, **16**, 5098–5109.
- Robledo-Arnuncio JJ, Gil L (2005) Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity*, **94**, 13–22.
- Smouse PE, Robledo-Arnuncio JJ (2005) Measuring the genetic structure of the pollen pool as the probability of paternal identity. *Heredity*, **94**, 640–649.

doi: 10.1111/j.1365-294X.2009.04392.x

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Geographical coordinates and trunk diameter distribution of the *Prunus mahaleb* trees represented by red dots in Fig. 1.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.