'Paternity analysis of pollen-mediated gene flow for Fraxinus excelsior L. in a chronically fragmented landscape'

By C F E Bacles and R A Ennos.

1 **ORIGINAL ARTICLE**

2 **Paternity analysis of pollen-mediated gene flow**

3 **for *Fraxinus excelsior* L. in a chronically**

4 **fragmented landscape**

5

6 **CFE Bacles[1,2] and RA Ennos[1,3]**

7

8 [1]*Institute of Evolutionary Biology, Ashworth Laboratories, School of Biological*

9 *Sciences, The University of Edinburgh, Edinburgh, EH9 3JT, UK*

10 [2]*Present address: INRA UMR BIOGECO 69 Route d'Arcachon 33612 Cestas*

11 *cedex France*

12 [3]*Correspondence: Dr. Richard A. Ennos. Institute of Evolutionary Biology,*

13 *Ashworth Laboratories, School of Biological Sciences, The University of Edinburgh,*

14 *Edinburgh, EH9 3JT, UK. Email: rennos@ed.ac.uk; Phone :+44 131 650 5411 ;*

15 *Fax : +44 131 662 0478*

16

17 **Keywords**: *Fraxinus excelsior* L., habitat fragmentation, landscape genetics,

18 genetic connectivity, microsatellites, paternity assignment

19

20 **Running title**: Genetic connectivity of ash forest remnants

21

22 **Word count** (excluding references, tables and figures): 7610 words

23

24

25 Paternity analysis based on microsatellite marker genotyping was used to infer

26 contemporary genetic connectivity by pollen of three population remnants of the

27 wind-pollinated, wind-dispersed tree *Fraxinus excelsior,* in a deforested Scottish

28 landscape. By deterministically accounting for genotyping error and comparing a

29 range of assignment methods, individual-based paternity assignments were used

30 to derive population-level estimates of gene flow. Pollen immigration into a 300ha

31 landscape represents between 43% and 68% of effective pollination, mostly

32 depending on assignment method. Individual male reproductive success is

33 unequal, with 31 of 48 trees fertilising one seed or more, but only three trees

34 fertilising more than ten seeds. Spatial analysis suggests a fat-tailed pollen

35 dispersal curve with 85% of detected pollination occurring within 100m, and 15%

36 spreading between 300m and 1900m from the source. Identification of immigrating

37 pollen sourced from two neighbouring remnants indicates further effective dispersal

38 at 2900m. Pollen exchange among remnants is driven by population size rather

39 than geographic distance, with larger remnants acting predominantly as pollen

40 donors, and smaller remnants as pollen recipients. Enhanced wind dispersal of

41 pollen in a barren landscape ensures that the seed produced within the catchment

42 includes genetic material from a wide geographic area. However, gene flow

43 estimates based on analysis of non-dispersed seeds were shown to underestimate

44 realised gene immigration into the remnants by a factor of two suggesting that

45 predictive landscape conservation requires integrated estimates of post-recruitment

46 gene flow occurring via both pollen and seed.

47 **Introduction**

48 The accuracy of our prediction of the response of forest trees to deforestation and

49 population fragmentation relies on an understanding of how pollen and seed

50 movement is modified as a consequence of changes in the landscape (Sork and

51 Smouse, 2006). Trees, which are characterised by their individual longevity, high

52 intra-population genetic diversity, and often substantial potential for gene flow via

53 pollen and seed, may be particularly well-equipped to withstand habitat disturbance

54 (Hamrick, 2004). Although theoretical predictions of reduced genetic diversity and

55 elevated inbreeding following habitat fragmentation (Young *et al.*, 1996) are upheld

56 for a number of wind-pollinated temperate tree species (Sork e*t al.*, 2002; Jump

57 and Penuelas, 2006), a recent review of empirical studies conducted in neotropical

58 tree species suggests that fragmentation generally has more complex effects

59 (Lowe *et al.*, 2005).

60
61 An emerging picture is of an increase in pollen and seed-mediated gene flow

62 across deforested landscapes (Aldrich and Hamrick, 1998; Dick, 2001; White *et al.*,

63 2002; Bittencourt and Stebbenn, 2007) However this enhanced gene flow does not

64 necessarily lead to an increase in genetic diversity or reduction in inbreeding if a

65 limited number of pollen and seed sources contribute to the gene pool (Aldrich and

66 Hamrick, 1998; O'Connell *et al.*, 2006; Sork *et al.*, 2002). Moreover the effect of

67 fragmentation is often not uniform over the landscape. Smaller fragments tend to

68 receive proportionally more pollen immigration than larger fragments because of a

69 paucity of local pollen donors (Sork and Smouse, 2006). It is clear from these

70 considerations that to understand the genetic connectivity of tree species living in

71 fragmented habitats requires an appreciation of the contemporary processes of

72 dispersal and establishment and an analysis of how they are affected by the spatial

73 scale of fragmentation and the heterogeneity of the landscape in which the

74 fragmentation occurs (Sork and Smouse, 2006).

75
76 The combined development of highly polymorphic microsatellite markers and

77 statistical analysis of parentage assignment (reviewed in Jones and Ardren, 2003)

78 have made it possible to gather empirical evidence of contemporary gene

79 movement within various landscapes for wild animals (e.g. Hazlitt *et al.*, 2006) and

80 plants (e.g. Bittencourt and Stebbenn, 2007). For measuring contemporary gene

81 flow these methods have many advantages over previous approaches that relied

82 on inferences from genetic structure and yielded estimates of historical, average

83 values of effective migration rate (Sork *et al.*, 1999; Whitlock and McCauley, 1999).

84 However a number of problems are beginning to emerge with adopting parentage

85 assignment approaches for measuring gene flow in practice.

86
87 The first issue is that the microsatellite methodology widely used for genotyping

88 has significant assay limitations that may call the accuracy of the pedigree

89 inference into question (Dakin and Avise, 2004; Hoffman and Amos, 2005; Slavov

90 *et al.*, 2005). Two main types of genotyping error can be distinguished, allele

91 dropouts (Dakin and Avise, 2004) and erroneous calling of allele size (Amos *et al.*,

92 2007) both of which can be either of a systematic or stochastic nature. While some

93 studies suggest that it is best to discard affected loci from parentage analyses

94 where occurrence of non-amplifying (null) alleles is suspected (Dakin and Avise,

95 2004), Wagner *et al.* (2006) argue that when the number of loci is low,

96 discriminatory power may decrease dramatically as a result. They suggest that a

97 better alternative to either removing loci or ignoring the presence of null alleles is to

98 accommodate them within the analyses. Indeed the use of many, even moderately

99 variable, loci rather than fewer hypervariable ones, reduces the impact of error at

100   any particular locus on parentage assignment (Hoffman and Amos, 2005; Slavov *et*

101   *al.*, 2005). Despite the profusion of recent publications establishing that even a low

102   genotyping error rate had non-trivial consequences for parentage and relatedness

103   studies, quantification and publication of error rates are not yet routinely performed

104   (Hoffman and Amos, 2005).

105

106   The second issue is that conclusions drawn from the analysis depend on the

107   method of paternity assignment adopted and assumptions about the size and

108   genetic composition of the population of potential paternal parents that have not

109   been sampled (Oddou-Muratorio *et al.*, 2003; Burczyck and Chybicki, 2004). While

110   simple exclusion is a useful starting point for paternity inference, refined statistical

111   approaches are necessary to assess the confidence in paternity assignment

112   (Marshall *et al.,* 1998). For example in natural tree populations it is virtually

113   impossible to sample all trees contributing to the reproductive pollen pool. It is

114   therefore necessary to assess the risk of excluding a candidate pollen parent on

115   the sole grounds that it has not been sampled. Recent methods based on either

116   Likelihood or Bayesian approximation allow us to estimate the statistical precision

117   of a paternity assignment for a given sample of a reproductive population (Marshall

118   *et al.*, 1998; Nielsen *et al.*, 2001; Gerber *et al.*, 2003; Araki and Blouin, 2005;

119   Hadfield *et al.*, 2006). Overall, it is preferable to use more than one of these

120   approaches to estimate genetic exchange among populations (Oddou-Muratorio *et*

121   *al.*, 2003).

122

123   In plants a further complication with measuring interpopulation gene flow using

124   parentage assignment arises because gene flow is effected by two asynchronous

125   dispersal processes, the first involving pollen and the second involving seeds. The

126   genetic material transferred between populations via pollen and incorporated into

127   seed present on a mother tree only brings about gene flow if that seed is recruited

128   into the local population. Parentage assignment of recruited seedlings is difficult to

129   track with current analytical tools (Sork and Smouse, 2006). Therefore

130   contemporary gene flow among fragmented tree populations has often been

131   estimated by measuring pollen movement to pre-recruitment seeds. These

132   estimates of gene flow are reasonable if seed-mediated gene flow among existing

133   populations is rare, and seed dispersal is primarily important for recolonisation and

134   range expansion. However in tree species with a high potential for long-distance

135   seed dispersal and subsequent recruitment this assumption may be invalid

136   (Smouse and Sork, 2004) and realised gene flow following seed dispersal may

137   differ significantly from gene flow measured in pre-recruitment seed within a

138   population. The extent of the discrepancy between gene flow measured from

139   parentage analysis before and after seed recruitment has still to be properly

140   documented.

141

142   The overall objective of the current study is to describe the genetic connectivity

143   among population fragments of common ash (*Fraxinus excelsior*) in a chronically

144   deforested landscape in the Southern Uplands of Scotland. Previous work on these

145   fragments has shown that they maintain high levels of genetic diversity and weak

146   inter-fragment differentiation ($\Theta = 0.080$), indicating that historical gene flow has

147   not been limited ($Nm = 3.48$). We also found from an analysis of seed families,

148   using a mixed-mating model approach (Ritland 2002), that contemporary matings

149   are, on average, predominantly outcrossed ($t_m = 0.971 \pm 0.028$) and using a

150   neighbourhood model approach (Burczyk *et al.*, 2002) that contemporary effective

151   pollen dispersal distance within the landscape averages 328 m (Bacles *et al.*,

152   2005). Both seeds collected from forest fragments and newly recruited seedlings

153   were found to harbour high levels of genetic diversity comparable to that of the

154   adult population suggesting an essential contribution of long distance dispersal to

155   genetic diversity in this wind-pollinated wind-dispersed species (Bacles *et al.*, 2005;

156   2006). The present paper complements these studies by quantifying the genetic

157   exchange among the individual fragments brought about by pollen flow and relates

158   this to the size and landscape context of the fragments.

159

160   To achieve this we estimate pollen-mediated gene flow and male reproductive

161   success of local *F. excelsior* trees from a paternity analysis of non-dispersed seeds

162   genotyped at hypervariable microsatellite markers. This is the best methodological

163   approach currently available to address this question. Nonetheless, in full

164   awareness of the potential limitations of the methodology, we take a number of

165   steps in order to ensure that our estimation describes true biological phenomena.

166   Firstly, we quantify genotyping error at marker loci and set out to minimise error

167   due to possible mis-scoring or null alleles using a simple deterministic approach.

168   Secondly, we use a range of paternity assignment methods to obtain a confidence

169   interval rather than a point estimate of pollen-mediated gene flow. Lastly, we

170   compare such estimates derived from non-dispersed seeds with those derived from

171   seedlings establishing in the same *F. excelsior* remnants estimated by means of

172   parent-pair analysis (Bacles *et al.*, 2006), to assess how they relate to absolute

173   levels of genetic exchange among remnant populations.

Material and methods

Study species

Fraxinus excelsior, common ash, is a post-pioneer tree species widespread in temperate Europe and native throughout the British Isles. The phylogeography of the species is now well described (Ferrazzini *et al.*, 2007; Heuertz *et al.*, 2004; Morand *et al.*, 2002). *F. excelsior* displays a complex, polygamous sexual system (FRAXIGEN, 2005) in which individuals may be classified phenotypically across a continuum from purely male to purely female with a whole range of hermaphroditic intermediates. Hermaphrodite individuals are self-fertile and levels of seed sets are similar in hermaphrodite and female trees, but in natural populations, *F. excelsior* is preferentially outcrossed and male fertility of hermaphrodite trees appears to be much lower than that of male trees (FRAXIGEN, 2005). Fruits are dry and winged, adapted to wind-dispersal. Regular fruit bearing begins around 20 years of age but fruiting phenology will vary depending on latitude, altitude, temperature and between years with great variation from no seeding to masting (FRAXIGEN, 2005).

Sampling and data collection

The study site is a highly deforested catchment of 900ha (Moffat Dale) located 80km south of Edinburgh, Scotland (N 55° 23' 51" W 3° 19' 50"), which forms part of a glacially derived landscape in which steep sided valleys have been carved by ice. Many native tree species including *F. excelsior* are confined to steep and narrow stream sides situated at the bottom of steep valleys inaccessible to grazing herbivores. Populations of *F. excelsior* tend to be very small, comprising ten to 30 mature individuals, with no natural regeneration in grazed areas. Remnant stands are typically separated by hundreds of metres although some can be isolated by more than one kilometre. In this catchment, *F. excelsior* is present in only five

Figure 1

200   forest remnants, two of them within the Carrifran Burn and three others in its

201   immediate surroundings (Figure 1, see also Bacles *et al.* 2005).

202

203   Two remnants in the bare open landscape of the Carrifran Burn, CDa and CMa,

204   and one remnant confined to a higher altitude dense conifer plantation upstream of

205   Carrifran in Swine Cleuchs, SCa, chosen for their heterogeneity in size, density

206   and landscape features, were exhaustively sampled for adult trees and family

207   arrays. Two neighbouring remnants, in Spoon Burn (SBa), the adjacent valley to

208   Moffat Dale nearest to Carrifran, downstream and in Whitewells (Wa) located at the

209   bottom of the Moffat Dale where the Carrifran streams drain into Moffat Water

210   (Figure 1), were partially sampled for adult trees to gather an indication of potential

211   pollen immigration because they are the only two other known local pollen sources

212   within ten km.

213

214   In 2000 and 2001, leaf material was collected from all mature trees in CDa, CMa

215   and SCa (comprising 30, four and 12 individuals respectively, Figure 1). Leaf

216   material was also collected from two trees (A, B) isolated from the nearest

217   remnants by a distance of 250 m and from a sample of 20 mature trees in SBa and

218   Wa (Figure 1). Such sampling represents no less than 40% of the composition of

219   these two remnants. 30 fruits, or all fruits if the seed crop was less, were collected

220   from all 19 trees producing fruits in 2000, a non-masting year, in each of remnants

221   CDa, CMa and SCa. In total, we sampled 88 trees and 483 seeds from 19 families.

222

223   The complete sample of 88 trees and 483 seeds were genotyped for five

224   microsatellite markers previously developed for *F. excelsior*, namely, M2-30B, 1.19

225   and 3.1 (Brachet *et al.*, 1999), and FEMSATL2 and FEMSATL5 (Lefort *et al.*, 1999).

226   DNA isolation, amplification by polymerase chain reaction (PCR) and

227   electrophoretic separation of PCR products were carried out as described

228   elsewhere (Bacles *et al.*, 2005).

229

230   **Evaluating microsatellite scoring and accounting for mistyping**

231   Out of 483 seeds genotyped, 61 presented a mismatch with their mother at one or

232   more loci. Furthermore, genotypes observed at loci 3.1 and FEMSATL 5 suggest

233   departure from Mendelian segregation (Bacles *et al.*, 2005) and occurrence of null

234   alleles which has also been discussed in other *F. excelsior* studies (e.g. Heuertz *et*

235   *al.*, 2001; Morand *et al.*, 2002).

236

237   In the rare instances where correction for genotyping error is applied in empirical

238   studies, it is generally introduced as a global stochastic error rate (Gerber *et al.*,

239   2000; Marshall *et al.*, 1998). A major drawback of such practice is that the benefits

240   of accounting for error are often outweighed by costs in precision of paternity

241   assignment which becomes uninformative (Oddou-Muratorio *et al.*, 2003; Morissey

242   and Wilson, 2005). Therefore, here we chose to deterministically account for allele

243   dropouts and size miscalls by performing two successive transformations to the

244   raw multilocus genotypes (referred to hereafter as RAW).

245

246   Erroneous allele sizing is most likely to occur between alleles of similar size and

247   when alleles are rare. Therefore, we applied an initial transformation to account for

248   size miscalls in the form of a binning procedure. At each locus, rare alleles were

249   binned with common alleles of the nearest size. Alleles were deemed rare when

250   they occurred at a frequency of less than 0.01 for the entire dataset. The procedure

251   was applied strictly to loci M2.30B and FEMSATL2. For loci 1.19, 3.1 and FEMSATL5,

252   more common alleles were also binned in order to reflect difficulties in gel scoring

253   for 1.19 and difficulties in respect of Mendelian segregation for 3.1 and FEMSATL5

254 (Bacles *et al.* 2005). The procedure reduced the number of alleles observed from

255 54 to 29, from 35 to 20, from 10 to 7, from 46 to 12 and from 17 to 8 at loci

256 FEMSATL2, M2-30B, 1.19, FEMSATL5 and 3.1 respectively in the binned dataset

257 (referred to hereafter as BIN).

258

259 A second transformation was then performed to account for the possible

260 occurrence of allele dropouts. A one-allele dropout model was applied to each

261 locus by introducing a new (i.e. non-observed) allele, by rescoring every individual

262 with a non-amplifying genotype as homozygote null, and every observed

263 homozygote as heterozygote null in the transformed dataset (hereafter referred to

264 as BINNULL).

265

266 For each dataset, genotyping error rates were quantified by means of direct

267 comparison of offspring-mother genotype at each locus and averaged over loci in

268 CERVUS 2.0 (Marshall *et al.*, 1998). Loci were retained for subsequent analyses if

269 the estimated genotyping error was less than 5%. In order to assess the

270 discriminatory power of each dataset, a paternity exclusion probability (PEP) was

271 computed for each locus and accumulated over loci in FAMOZ (version released on

272 17.04.2007, Gerber *et al.*, 2003).

273

274 **Estimating contemporary pollen-mediated gene flow at the landscape scale**

275 Paternity analyses were undertaken to identify the pollen parent of the 422 seeds

276 that shared a compatible multilocus genotype with their putative mother only and

277 excluding the 61 seeds presenting at least one mismatching allele with their

278 mother. Pollen parents were considered either among the 48 trees sampled in

279 CDa, CMa and SCa, including the possibility for mother trees to self, or as pollen-

280 mediated gene flow from outside the landscape covered by the three completely

281 censed remnants (Figure 1).

282

283 For each of the RAW, BIN and BINNULL datasets, paternity was assigned using both

284 a simple exclusion (SE) and a maximum-likelihood (ML) approach in FAMOZ (Gerber

285 *et al.*, 2003). In each case, outcomes of paternity assignment may be, for each

286 individual seed, either that one unique individual among the 48 trees of Carrifran

287 and Swine Cleuchs is assigned as its pollen parent, or that its paternity is

288 unresolved with more than one possible pollen parent among the 48 trees, or

289 finally that all 48 trees are excluded as potential pollen parents and its paternity is

290 assigned to immigrant pollen. A range of values for apparent pollen-mediated gene

291 flow into the landscape is subsequently obtained as the percentage of seeds in the

292 sample for which paternity was assigned to immigrant pollen.

293

294 In FAMOZ, confidence in paternity assignment is estimated using a simulation

295 procedure for hypothesis testing (Gerber *et al.*, 2000). The paternity of the 422

296 seeds sampled from 19 mother trees in three *F. excelsior* remnants of Moffat Dale

297 was assigned to the most-likely fathers detected by means of 'log of the odds'

298 ratios (LOD scores) based on pollen pool gene frequencies estimated from progeny

299 arrays in MLTR (Ritland, 2002). We chose to approximate the (non-observed) allele

300 frequencies of the reproductive population by using the observed pollen pool

301 frequencies instead of the frequencies observed for the small sample of 88 mature

302 trees because the latter, which is sampled *a priori* based on spatial location, may

303 be a poor estimate of the actual reproductive population if gene flow is extensive.

304 No significant genotypic association was detected among any pair of loci (Bacles *et*

305 *al.*, 2005). LOD scores over all loci were therefore obtained by adding LOD scores

306 calculated for each locus.

307

308   Confidence in paternity assignment was then determined in FAMOZ by comparing

309   the distribution of the LOD scores of the most-likely fathers of 50 000 randomly

310   generated seeds with their father randomly chosen among the 48 trees to the

311   distribution of LOD scores of the most-likely fathers of 50 000 seeds whose paternal

312   genotype was randomly generated according to pollen pool allele frequencies. The

313   test threshold for rejecting a candidate as a true father (TF), was chosen at the

314   intersection of the two distributions of LOD scores to minimise both type I error,

315   wrongly considering as resulting from pollen immigration a seed sired by a sampled

316   father, and type II error, wrongly assigning true pollen immigration to a sampled

317   father (Gerber *et al.*, 2000). For paternity assignment by simple exclusion (SE), all

318   candidate males with a positive LOD score (i.e. test threshold TRUE FATHER=0) were

319   not excluded from being the true father.

320

321   Global results of paternity assignment obtained for each of RAW, BIN and BINNULL

322   datasets with both SE and ML methods are discussed in respect of estimated error

323   rates, confidence levels in assignments and estimated pollen-mediated gene flow

324   at the landscape scale. The dataset/method combination found to minimise

325   genotyping error rates while maximising confidence in paternity assignments was

326   retained for subsequent detailed description of individual male reproductive

327   success. In particular, results were summarised in order to identify the number of

328   sires among the sample trees and the number of seeds they sired among the

329   sampled seeds. The pollen dispersal curve was estimated by plotting the distance

330   between mother trees and pollen donors for each most-likely assignment. When

331   more than one likely father was identified (unresolved assignment), a fraction of the

332   seed was assigned to all likely fathers evenly and proportionally to the number of

333   likely fathers found.

334 **Estimating the fractional pollen contribution of forest remnants and**

335 **identifying local sources of pollen immigration**

336 In order to estimate landscape connectivity and pollen-mediated genetic exchange

337 among forest remnants, it may be most relevant to assess the relative pollen

338 contribution of forest remnants to the seed crop rather than to define individual

339 paternity *per se.* It has been argued that such (meta)population-scale phenomena

340 may be better addressed with fractional-likelihood assignment methodology that

341 will assign a fraction of the paternity of a given seed to all male candidates with a

342 positive LOD score in proportion to their likelihood probability (Nielsen *et al.*, 2001).

343 In order to estimate potential pollen immigration into CDa, CMa, and SCa from

344 other known *F. excelsior* remnants of Moffat Dale, SBa and Wa (Figure 1), we

345 estimated the posterior expectation of the number of sampled offspring in each of

346 five remnants by means of fractional-likelihood assignment in PATRI (Signorovitch

347 and Nielsen, 2002).

348

349 The approach in PATRI also allows us to make prior assumptions about the

350 proportion of the pollen parents that have not been sampled (Nielsen *et al.*, 2001).

351 While the actual effective male population size is unknown, we do have some

352 expectations of the number of trees occurring in the landscape and likely to

353 contribute to the pollen pool. Therefore, we tested the sensitivity of the fractional-

354 likelihood assignment to assumptions made on the population size (N) by

355 successively repeating analyses for an N of 88, the number of trees sampled; an N

356 of 150, the approximate number of *F. excelsior* trees occurring in the catchment

357 and N modeled as a uniform function varying between 100 and 500. We compared

358 results with those of a maximum-likelihood assignment performed in FAMOZ when

359 considering all 88 trees sampled in Moffat Dale.

360

361 **Comparing potential to realised pollen-mediated gene flow among forest**

362 **remnants.**

363 To assess whether estimates of pollen-mediated gene flow from seeds that have

364 not yet dispersed reflect estimates of pollen-mediated gene flow seen after seed

365 dispersal and establishment, for each of remnant CDa, CMa and SCa, we used

366 most-likely fathers to attribute the origin of the pollen grain to either local pollen,

367 foreign pollen of known source (in other identified remnants) or of unknown source.

368 We compared these figures with previously published results derived from a ML

369 parent-pair analysis performed on seedlings establishing in the same three

370 remnants (Bacles *et al.*, 2006).

371

372 In addition we estimated total gene flow into fragments using genotypic data

373 generated both from progeny arrays (T_p) and from established seedlings (T'_s). Note

374 that pollen grains only carry one gene copy while diploid seeds carry two. Let *A*

375 and *A'* represents the number of local seeds fertilised by immigrant pollen in

376 progeny arrays and establishing seedlings respectively. Let B' be the number of

377 immigrating seeds, and C and C' the total number of seeds sampled in progeny

378 arrays and establishing seedlings respectively, then:

379

380 $Tp = (A / 2C) \times 100$ **Equation 1**

381 and

382 $T's = (A' + 2B') / 2C' \times 100$ **Equation 2**

383

384 In equation 1, seeds are sampled on known mother trees and are all local. *Tp*

385 therefore assumes that pollen is the main vector of gene flow among populations

386 and that seed dispersal is mostly local. Results are discussed in terms of

387   comparison of potential (i.e. *ante* dispersal) and realised (i.e. *post* dispersal and

388   establishment) gene flow in the three heterogeneous forest remnants.

389 **Results**

390 **Genotyping error and choice of dataset**

391 Genotyping error rates estimated by means of offspring-mother genotype

392 comparison are reported for each locus and each genotype transformation in Table

393 1. They were found to be highest for loci 3.1 (error (RAW) = 0.2847) and FEMSATL 5

394 (error (RAW) = 0.2237). Drastically reducing the number of alleles, from 17 to 8 and

395 from 46 to 12 at locus 3.1 and FEMSATL5 respectively, decreases the error rate only

Table 1

396 slightly (error (3.1, BIN) = 0.2679; error (FEMSATL5, BIN) = 0.2133) but additional

397 inclusion of a null allele decreases the error significantly (error (3.1, BINNULL) =

398 0.1499; error (FEMSATL5, BINNULL) = 0.0762) suggesting that null alleles may be

399 responsible for the non-Mendelian segregation observed in progeny arrays at

400 these loci. In contrast, genotyping error rates estimated at loci 1.19, M2.30B and

401 FEMSATL2 were under 5% (Table 1), and lowest at loci M2.30B and FEMSATL2 after

402 binning of alleles (error (M2.30B, BIN) = 0.0347; error (FEMSATL2, BIN) = 0.0443) and

403 with inclusion of a null allele at locus 1.19 (error (1.19, BINNULL) = 0.0289). Overall,

404 inclusion of loci 3.1 and FEMSATL5 in estimates increases mean error across loci

405 dramatically, up to 3 times (Table 1). When these loci were excluded, estimates of

406 mean error across loci were consistently under 5%.

407

408 Paternity exclusion probabilities (PEP) estimated per locus for RAW vary between

409 0.594 at locus 1.19 and 0.864 at locus FEMSATL2 (Table 1) reflecting variation in

410 level and evenness of polymorphism among loci (Bacles *et al.*, 2005). As expected,

411 reducing the number of alleles at each locus systematically lowers discriminatory

412 power among genotypes at each locus, albeit moderately, with PEP estimated per

413 locus from BIN varying from 0.504 at locus 1.19 to 0.857 at locus FEMSATL2.

414 Conversely, additional inclusion of a null allele results in contrasting effects on

415 single locus estimates of PEP (Table 1). However, cumulative estimates of PEP,

416 including all five loci, are consistently very high across datasets, reaching values

417 upward of 99.9%. Excluding loci 3.1 and FEMSATL5 decreased cumulative PEP only

418 slightly (Table 1).

419

420 On the basis both that including loci 3.1 and FEMSATL5 increases genotyping error

421 to rates well above 5% and that excluding them hardly affects multilocus genotype

422 discrimination of individuals, they were not retained for subsequent analyses.

423 Meanwhile, considering loci 1.19, M2.30B and FEMSATL2 only, mean genotyping

424 error rate is lowest when multilocus genotypes are transformed with binning at loci

425 M2.30B and FEMSATL2 and with binning and inclusion of a null allele at locus 1.19

426 (dataset hereafter referred to as BIN3NULL1; error (MEAN, BIN3NULL1) = 0.0360). No

427 identical multilocus genotype was found among the 88 trees sampled in Moffat

428 Dale and cumulative PEP is estimated at 0.991 which is nearly as high as for RAW

429 data (Table 1).

430

431 **Contemporary pollen-mediated gene flow at the landscape scale**

432 Results of SE and ML paternity analysis for RAW, BIN, BIN3NULL1 datasets and for

433 assignment of pollen parents to the 422 seeds that do not show any mismatch with

434 their mother on raw data, are given in table 2. 43% to 68% of the 422 seeds

Table 2 435 analysed were found to have been fertilized with pollen dispersed from trees

436 located outside the landscape covered by three *F. excelsior* remnants of Carrifran

437 and Swine Cleuchs. Highest estimates were obtained using a ML method (Table 2).

438 Differences between SE and ML estimates of apparent pollen flow are due to a

439 number of seeds that were not assigned a father among the 48 sampled trees with

440 the ML method because their LOD was positive but below the given threshold for

441 assignment. However they were assigned one (or several potential) fathers with

442 the SE method, most frequently for the bin dataset ($N_{assigned}$ (BIN, SE) = 239, Table

443   2). False rejection of true sampled fathers was lowest (Type I error < 0.05 for TF =

444   2.90) for fewer seeds ($N_{assigned}$ = 141) with the BIN3NULL1 transformation (Table 2).

445

446   **Individual male reproductive success and distance of pollen dispersal**

447   On the basis that the BIN3NULL1 dataset is characterised by the lowest estimates of

448   genotyping error and type I error in ML assignment, the most-likely pollen parents

449   identified for 141 seeds under these conditions were retained for subsequent

450   description of male reproductive success and spatial patterns of pollen dispersal at

451   the landscape scale.

452

453   In total, 31 of the 48 trees sampled in CDa, CMa and SCa were found to sire one

454   or more of the sampled seeds (Supplementary Table S1). Of these, 14 trees

455   fertilised three seeds or less, while three trees fertilised more than ten seeds each.

456   In the latter case, all fertilized seeds were sampled from neighbouring trees within

457   the same remnant as the sire.

458

459   This pattern in individual male reproductive success is reflected in the L-shape of

460   the pollen dispersal curve constructed by plotting the distribution of spatial

461   distances between the 141 assigned seeds (i.e. based on spatial location of

462   mother trees) and their most-likely pollen-parent (Figure 2a). Over 80% of effective

463   pollen dispersal is confined to less than 100m from the source, corresponding to

Figure 2   464   local pollen movement within remnant. The observed proportion of these local

465   pollinations is significantly higher than expected under random dispersal (48%,

466   Wilcoxon two-sided signed rank test p-value<0.05, Figure 2a). However, a number

467   of rarer events, each representing less than 10% of total effective pollen dispersal

468   (which is up to 25% less than expected from random dispersal), were identified

469   between 200m and 600 m, and between 1600m and 1800m corresponding to inter-

470   remnant long-distance dispersal.

471

472   Further analysis that included trees in partially sampled remnants SBa and Wa

473   allowed us to refine the shape of the pollen dispersal distribution (Figure 2b). ML

474   paternity assignment of 163 seeds (TF=2.92, Type I error <0.05; Type II error<0.28)

475   attributed their paternity either to trees sampled in CDa, CMa and SCa or to trees

476   sampled in SBa and Wa (Supplementary Table S2). The inclusion of these nearby

477   pollen sources allowed the detection of a small number of effective pollination

478   events at distances between 1100m and 1600m, between 1800m and 200m and

479   as great as 2900m (Figure 2b).

480

481   **Pollen-mediated genetic connectivity of forest remnants**

482   Pollen contribution of the five *F. excelsior* population remnants sampled for adult

483   trees was estimated from fractional likelihood (FL) paternity analysis of a subset of

484   404 seeds and 83 male candidates with no missing value in their multilocus

485   genotype at 1.19, M2.30B and FEMSATL2 because PATRI computes complete

486   genotypes only (Signorovitch and Nielsen, 2002). Estimates of the relative

487   contribution of the five *F. excelsior* remnants of Moffat Dale to effective pollination

Table 3 488   in remnants CDa, CMa and SCa are similar when a FL or ML approach is applied to

489   all 88 genotyped trees (Table 3). However, the absolute fractional contribution of

490   remnants to paternity of sampled seeds is highly sensitive to the sampled fraction

491   of reproductive adults (Nielsen *et al.* 2001) as illustrated by decreasing posterior

492   expectation of the number of sampled offspring fathered in each remnant with

493   increasing prior *N* (Table 3). Overall however, the relative pollen contribution of the

494   five forest remnants remains unchanged. CDa contributes most to paternity of the

495   sampled seeds. SCa, SBa and Wa also contribute in decreasing proportion (Table

496   3), while remnant CMa is a poor contributor (Table 3). The analysis clearly

497   demonstrates that neighbouring remnants may act as sources of immigrant pollen.

498

499   Estimates of pollen-mediated genetic exchange among remnants CDa, CMa and

500   SCa derived from ML paternity analysis confirm that the largest remnant CDa acts

501   as a pollen donor, siring 26% of the seeds sampled in remnant CMa, located 600m

502   away, and 7% of the seeds sampled in the most spatially isolated remnant, SCa,

Figure 3   503   located 1700m away (Figure 3). Conversely, CMa, the smallest remnant (N_{trees} =

504   4), only sired 3% of its local seeds, and 1% the seeds within CDa (Figure 3).

505

506   **Potential and realised pollen-mediated gene flow**

507   How such pollen-mediated genetic exchange will impact on genetic structure

508   depends on dispersal and establishment of the seeds. Estimates of potential

509   pollen-mediated gene flow into remnants CDa, CMa, and SCa from ML paternity

Table 4   510   analysis of 422 seeds collected on mother trees before their dispersal described

511   above (65% to 94%, Table 4) are comparable to those of potential pollen-mediated

512   gene flow from a ML parent-pair analysis of 60 seedlings that were establishing in

513   the same three remnants the following year (70% to 100%, Table 4). However,

514   such comparison also show that pollen-mediated gene flow realised after seed

515   dispersal and seedling establishment is much lower, ranging from 12.5% in

516   remnant CMa to 17.5% in remnants CDa and SCa. Total gene flow estimates from

517   progeny arrays are much lower (T_p ranging between 32.5% and 47%) than from

518   establishing seedlings (*T's* ranging between 67.5% and 87.5%, Table 4).

519 **Discussion**

520 Despite a number of significant concerns over genotyping error and uncertainties

521 associated with statistical modeling, the application of paternity assignment

522 analysis in the fragmented populations of *F. excelsior* has significantly enhanced

523 our understanding of their genetic behaviour. It is clear that the population

524 fragments within a single valley receive about half their pollen from outside the

525 valley. Remnants within the valley are genetically connected via pollen flow, but the

526 patterns of pollen flow among fragments are not symmetrical; pollen is

527 preferentially transferred from large to small fragments. The analysis has also

528 demonstrated that the effective pollen dispersal curve is fat-tailed. While the

529 majority of detected pollen movement occurs over short distances (within 100m),

530 there is still substantial pollen flow occurring over distances greater than 1km.

531 Although these general conclusions are important for guiding the management of

532 fragmented tree populations, this study has also highlighted the practical difficulties

533 associated with obtaining quantitative assessments of gene flow from large scale

534 studies that rely on parentage analysis.

535

536 A predictive understanding of the genetic connectivity of fragmented populations

537 requires reliable estimation of contemporary gene dispersal across heterogeneous

538 landscapes (Sork and Smouse, 2006). While development of both molecular

539 techniques and statistical tools has greatly improved prospects for accuracy, the

540 application of parentage analyses to natural populations remains an evolving area

541 of research leading to regular reanalysis of empirical data within new statistical

542 frameworks (e.g. Hadfield *et al.*, 2006; Slate *et al.*, 2000). At the centre of the

543 debate, lies the question of sensitivity of parentage analyses to partial sampling of

544 the reproductive population and to genotyping error at marker loci (Nielsen *et al.*,

545 2001; Oddou-Muratorio *et al.*, 2003; Slavov *et al.*, 2005). In order to obtain reliable

546   population-level inference of gene flow from a collection of individual-level paternity

547   assignments, we chose to address these concerns by applying a range of paternity

548   analysis methods to *F. excelsior* population remnants of the chronically deforested

549   catchment of Moffat Dale (Table 2 and 3).

550

551   Critically, application of parentage analyses to estimating gene movement in

552   natural populations relies on a conundrum: accuracy in estimation of the proportion

553   of the reproductive population that has not been sampled (i.e. immigrant gene flow)

554   strongly increases as the proportion of the -yet unknown- reproductive population

555   that has not been sampled decreases (Oddou-Muratorio *et al.*, 2003). Approaching

556   true reproductive population size seems particularly important to analyses

557   performed when using a fractional likelihood approach in PATRI because estimating

558   the absolute contribution of *F. excelsior* trees to paternity of sampled seeds is

559   sensitive to input prior information on the effective male population size (Table 3).

560   The advantage of the hypothesis testing-based simulation approach to determining

561   assignment confidence in FAMOZ is that it does not require any assumption on the

562   size of the true reproductive population. Comparison of SE and ML methods for a

563   range of transformed multilocus genotypes accounting for genotyping error *sensu*

564   *lato* at microsatellite markers suggest an immigration of at least 43% (SE, BIN) and

565   up to 68% (ML, RAW) of the pollen fertilising seeds from 19 trees of three forest

566   remnants of the Moffat Dale catchment.

567

568   The range in gene flow rates seems mostly affected by the choice of paternity

569   assignment method rather than by dataset transformation. Indeed, while

570   transformation of raw data allowed us both to reduce mean genotyping error (for

571   BIN3NULL1 at 3.60%), and to minimise false rejection of true fathers that were

572   sampled (for BIN3NULL1 Type I< 5%) to acceptable levels, estimates of gene flow

573   between RAW, BIN and BIN3NULL1 vary by up to 15% or a given paternity

574   assignment method. Variation in gene flow estimates between the SE and ML

575   methods can be attributed to the fact that, under ML, between 10% and 16% of

576   seeds are not assigned a father among the sampled trees. This is because the LOD

577   score of candidates is positive but below the given threshold for assignment

578   (TF=2.90, Table 2). Although paternity of these seeds cannot be assigned at the

579   chosen confidence threshold, it is arguable that their paternity should necessarily

580   be attributed to immigrant pollen. Indeed, it has been demonstrated that

581   assignment error may be much higher than random on unobserved (i.e. immigrant)

582   events (Slate *et al.*, 2000) which suggests that estimates of pollen-mediated gene

583   flow from ML method (here inclusive of seeds that were not assigned a father

584   because genotypically compatible candidates had a low LOD score) should be seen

585   as upper limits. Conversely, ML analysis suggests that even for a strict LOD score

586   threshold, Type II error of wrongly assigning immigrant pollen to an unrelated

587   sampled tree is high (up to 27% for BIN and TF=2.50, Table 2) indicating substantial

588   cryptic gene flow. Therefore, apparent pollen flow estimated with relaxed

589   assignment (equivalent to TF>0) by SE method, and those obtained with BIN data

590   because allele binning results in lower discrimination of multilocus genotypes, are

591   most conservative, with increased risk of cryptic gene flow and therefore represent

592   lower limits of effective pollen immigration into forest remnants of Moffat Dale.

593

594   We justify our deterministic transformation of genotypic data not as substitution of

595   raw datasets for transformed ones that may be more biased but rather as a simple

596   way of minimising genotyping error and its possible influence on paternity

597   assignment. The transformed dataset still includes a mean genotyping error rate of

598   about 3.6% per locus which may still have an impact on the conclusions drawn

599   from this study. Nonetheless, we deliberately chose not to include this global rate in

600 paternity analyses because there is evidence that including global genotyping error

601 rates inflates errors in paternity assignments (Oddou-Muratorio *et al.*, 2003; Slavov

602 *et al.,* 2005). Given the limitations of the dataset, which are here clearly quantified,

603 a range of estimates from several paternity analyses provides an ecologically

604 meaningful interpretation of pollen-mediated gene flow at the landscape scale.

605

606 Extensive contemporary pollen-mediated gene flow averaging 60% has already

607 been reported in plots located within continuous stands of *F. excelsior* (Hebel *et al.*,

608 2007) and of other wind-pollinated temperate tree species, for instance for *Quercus*

609 (Dow and Ashley, 1998) covering only small areas. Contemporary pollen-mediated

610 gene flow estimates of 43% to 68% for *F. excelsior* in the fragmented landscape of

611 Moffat Dale are comparatively higher because all *F. excelsior* trees were sampled

612 in an area of 300ha suggesting that *F. excelsior* maintains extensive pollen

613 exchange across a landscape heavily deforested not only locally but also at the

614 wider regional scale of the Southern Uplands of Scotland (>50km). Although trees

615 standing solitarily in grazed pastures, spatially isolated from congeners following

616 deforestation, were once described as *living-dead* (Janzen, 1986), there is now

617 plethora of evidence of reproductive activity of isolated pasture trees, mainly in

618 tropical species (Aldrich and Hamrick, 1998; Dick, 2001; White *et al.*, 2002)

619 corroborating our findings of enhanced pollen-mediated gene flow following

620 anthropogenic disturbance. Nonetheless, of the two *F. excelsior* trees of Moffat

621 Dale that are isolated from others by a distance of at least 250m (Figure 1), neither

622 produced a seed crop nor did they contribute to effective pollination of seeding

623 trees for the sampled reproductive season. On the basis of this observation, we

624 cannot reject the hypothesis that such isolated pasture trees are *living-dead.*

625 Similarly, only three of the 48 trees sampled locally have a high male reproductive

626 success and 26 of them contribute to effective pollination of fewer than two seeds

627   to none (Supplementary Table S1). However, whereas the evidence suggests that

628   most sampled trees have a low individual male reproductive success locally, we

629   cannot reject either the hypothesis that pollen from such trees effectively emigrated

630   to other forest remnants outside the sample area, as the presence of a large

631   component of immigrant pollination of either unknown origin or originated from

632   identified neighbouring sources would suggest (Figure 3).

633

634   Several ecological factors may have contributed to confer such an advantage to

635   long-distance pollination. Firstly, comparison of temporal variation of effective

636   pollen movement between mast seeding and non-masting years showed that in a

637   non-masting year, as is the case in the present study pollen-mediated gene flow

638   was favoured in a *F. excelsior* stand in southern England (FRAXIGEN 2005).

639   Furthermore, in such a situation, the small seed crops that were produced by a

640   number of trees (in particular, trees SCa34 and SCa38 displayed only one seed

641   branch with fewer than 10 seeds, Supplementary Table S1) may create a sampling

642   effect. Secondly, temporal variation in individual flowering phenology may greatly

643   affect mate availability. Indeed, Gerard *et al.* (2006) not only found that co-

644   flowering individuals were patchily distributed in space in a *F. angustifolia* and *F.*

645   *excelsior* hybrid zone, they also detected an asymmetry in male reproductive

646   success with early flowering trees participating more as pollen donors than late

647   flowering ones. A scenario where immigrant pollen would be preferentially available

648   during the period of stigma receptivity of seeding trees of Moffat Dale would also

649   favour long-distance pollination.

650

651   However, high levels of gene immigration are not necessarily sufficient to prevent

652   assortative mating and selfing (Gerard *et al.*, 2006). For gene flow to become an

653   efficient force counteracting the deleterious genetic effects of habitat

26

654 fragmentation, gene flow must not only be sustained at high levels across seasons,

655 but must also be qualitatively diverse. Here we find that efficient pollen immigration

656 allows for new and diverse genetic material to establish in the seed generation

657 (Bacles *et al.*, 2005). Such genetically diverse pollen pool composition may be

658 explained by the type of decay of pollen dispersal. Indeed, Klein *et al.* (2006)

659 demonstrated by means of simulation that fat-tailed dispersal kernels lead

660 asymptotically to a diverse propagule pool containing a balance of mixing of the

661 propagules of two sources and therefore that the diversity of the pollen pool of a

662 mother plant should increase with increased spatial isolation. Pollen dispersal

663 patterns observed for *F. excelsior* in Moffat Dale seem to corroborate such

664 theoretical findings. The majority of detected pollen dispersal was found between

665 near-neighbours at distances under 100m (Figure 2). However, not only were a

666 number of rare events detected among forest remnants at distances up to 2900m,

667 in proportions departing from random dispersal (Wilcoxon two-sided signed ranked

668 test, $p<0.05$), but undetected events were also in the majority and may have

669 originated at much greater distances, suggesting a L-shaped pollen dispersal

670 kernel with a tail spreading over several kilometres and underlining the difficulty of

671 detecting long distance dispersal (Nathan, 2006).

672

673 Such pollen dispersal effective over long distances can be linked to landscape

674 features resulting from habitat disturbance. Indeed, in the Southern Uplands of

675 Scotland, deforestation and land use for pasture have greatly opened the barren

676 landscape that is regularly battered by strong winds. It is therefore likely that wind-

677 mediated pollen movement for *F. excelsior* has been facilitated by the modification

678 of the landscape in Moffat Dale. In particular, genetic connectivity of forest

679 remnants seems favoured by landscape openess and remnant size rather than by

680 geographic proximity. Indeed, although no seeding trees were sampled in two of

681   five remnants (SBa and Wa), fractional-likelihood paternity assignment shows that

682   their contribution to effective pollination of seeding trees of the other three extant

683   remnants (CDa, CMa and SCa) is higher than that of remnant CMa which is a

684   much smaller remnant of only 4 trees located in the bed of the river running

685   through an exposed and barren pasture in the Carrifran valley. Within remnant

686   pollination for CMa is indeed much lower than for other remnants (Figure 3),

687   highlighting the fact that remnants are smaller, but are spatially well connected to

688   neighbouring forest remnants tend to receive proportionally higher gene flow simply

689   because there are fewer potential local pollen donors (Sork and Smouse, 2006).

690

691   An important practical consequence of the high rate of pollination by immigrant

692   pollen is that the locally produced seed in Carrifran will contain genes sampled

693   from a wide geographic area around the valley How the high rate of success of

694   immigrant pollen in the production of local seeds will ultimately affect the genetic

695   structure of *F. excelsior* remnants depends on how much of the pollen pool genetic

696   diversity is effectively carried into successive generations by established seedlings

697   that reach maturity. Natural regeneration in Moffat Dale has been severely limited

698   by continuous grazing pressure. Colonisation of mountain grasslands by *F.*

699   *excelsior* seedlings has been found to be connected to grazing activities, with

700   seedlings found preferentially in high layers of vegetation, in shaded and ungrazed

701   areas (Julien *et al.*, 2006)*.* Pasture habitats in Moffat Dale may therefore be

702   unfavourable to seedling establishment. Thus, actual gene flow may be recruitment

703   limited rather than dispersal limited (Imbert and Lefèvre, 2003). In fact, comparison

704   of total gene flow estimated here from non-dispersed seeds with total gene flow

705   estimated from newly established seedlings in three *F. excelsior* remnants shows

706   that actual recruitment of genes carried by immigrant pollen is limited (Table 4).

707   Note that the ratio of potential to realised pollen-mediated gene flow is low, not

708   because there seems to be an advantage conferred to recruitment of local seeds

709   fertilised with local pollen but because the majority of establishing seedlings have

710   immigrated into the remnants (Bacles *et al.*, 2006). This indicates that, in cases

711   when seed dispersal is an important vector of long-distance dispersal, estimating

712   seed-mediated gene flow is essential to predicting landscape connectivity (Sork

713   and Smouse 2006).

714

715   In the Southern Uplands of Scotland and in other severely deforested landscapes,

716   conservation management aimed at sustainable forest restoration without human

717   intervention must move away from *conservation gardening* (Hobbs, 2007).

718   Predictive conservation necessitate better understanding of the evolution of

719   dispersal in a changing environment (Kokko and Lopez-Sepulcre, 2006) and an

720   appreciation of how population genetic processes operate in ecological space and

721   time. Bringing knowledge of contemporary gene flow among population remnants

722   generated from this study and others into conservation will ensure that the

723   evolutionary processes maintaining genetic connectivity and evolutionary potential

724   are restored at the landscape scale (Meagher, 2007).

Acknowledgements

742 **Literature cited**

743 Aldrich PR, Hamrick JL (1998) Reproductive dominance of pasture trees in a

744 fragmented tropical forest mosaic. *Science* **28**: 103-105.

745 Amos W, Hoffman JI, Frodsham A*, et al.* (2007) Automated binning of

746 microsatellite alleles: problems and solutions. *Mol. Ecol. Notes* **7**: 10-14.

747 Araki H, Blouin MS (2005) Unbiased estimation of relative reproductive success of

748 different groups: evaluation and correction of bias caused by parentage

749 assignment errors. *Mol. Ecol.* **14**: 4097-4109.

750 Bacles CFE, Burczyk J, Lowe AJ, Ennos RA (2005) Historical and contemporary

751 mating patterns in remnant populations of the forest tree *Fraxinus excelsior* L.

752 *Evolution* **59**: 979-990.

753 Bacles CFE, Lowe AJ, Ennos RA (2006) Effective seed dispersal across a

754 fragmented landscape. *Science* **311**: 628.

755 Brachet S, Jubier MF, Richard M, Jung-Muller B, Frascaria-Lacoste N (1999) Rapid

756 identification of microsatellite loci using 5' anchored PCR in the common ash

757 *Fraxinus excelsior. Mol. Ecol.* **8**: 157-168.

758 Bittencourt JVM, Sebbenn AM (2007) Patterns of pollen and seed dispersal in a

759 small fragmented population of the wind-pollinated tree *Araucaria angustifolia* in

760 southern Brazil. *Heredity* **99**: 580-591.

761 Burczyk J, Adams WT, Moran GF, Griffin AR (2002) Complex patterns of mating

762 revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the

763 neighbourhood model. *Mol. Ecol.* **11**: 2379-2391.

764 Burczyk J, Chybicki IJ (2004) Cautions on direct gene flow estimation in plant

765 populations. Evolution **58**: 956-963.

766 Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis.

767 *Heredity* **93**: 504-509.

768   Dick CW (2001) Genetic rescue of remnant tropical trees by an alien pollinator.

769   *Proc. Roy. Soc. London B* **268**: 2391-2396.

770   Dow BD, Ashley MV (1998) High levels of gene flow revealed by paternity analysis

771   using microsatellites. **89**: 62-70

772   Ferrazzini D, Monteleone I, Belletti P (2007) Genetic variability and divergence

773   among italian populations of common ash (*Fraxinus excelsior* L.). *Ann. For. Sci.* 64:

774   159-168.

775   FRAXIGEN (2005) *Ash species in Europe: biological characteristics and practical*

776   *guidelines for sustainable use* Oxfod Forestry Institute, University of Oxford, U.K.

777   128 pp.

778   Gerard PR, Klein E, Austerlitz F, Fernandez-Manjarres JF, Frascaria-Lacoste N

779   (2006) Assortative mating and differential male mating success in an ash hybrid

780   zone population. *BMC Evol. Biol.* **6**:

781   Gerber S, Chabrier P, Kremer A (2003) FAMOZ: a software for parentage analysis

782   using dominant, codominant and uniparentally inherited markers. *Mol. Ecol. Notes*

783   **3**: 479-481.

784   Gerber S, Mariette S, Streiff R, Bodenes C, Kremer A (2000) Comparison of

785   microsatellites and amplified fragment length polymorphism markers for parentage

786   analysis. *Mol. Ecol.* **9**: 1037-1048.

787   Hadfield JD, Richardson DS, Burke T (2006) Towards unbiased parentage

788   assignment: combining genetic behavioural and spatial data in a Bayesian

789   framework. *Mol. Ecol.* **15**: 3715-3730.

790   Hamrick J (2004) Response of forest trees to global environmental changes. *For.*

791   *Ecol. Manage.* **197**: 323-335.

792   Hazlitt SL, Sigg DP, Eldridge MDB, Goldizen AW (2006) Restricted mating

793   dispersal and strong breeding group structure in a mid-sized marsupial mammal

794   (*Petrogale penicillata*). *Mol. Ecol.* **15**: 2997-3007.

795 Hebel I, Aldinger E, Haas R, *et al.* (2007) Pollen dispersal and pollen contamination

796 in an ash seed orchard (*Fraxinus excelsior* L.). *Allgemeine forst und jagdzeitung*

797 **178**: 44-49.

798 Heuertz M, Hausman F, Hardy OJ, Vendramin GG, Frascaria-Lacoste N,

799 Vekemans X (2004) Nuclear microsatellites reveal contrasting patterns of genetic

800 structure between western and southeastern European populations of the common

801 ash (*Fraxinus excelsior* L.). *Evolution* **58**: 976-988.

802 Heuertz M, Hausman F, Tsvetkov I, Frascaria-Lacoste N, Vekemans X (2001)

803 Assessment of genetic structure within and among Bulgarian populations of the

804 common ash (*Fraxinus excelsior* L.). *Mol. Ecol.* **10**: 1615-1623.

805 Hobbs R (2007) Managing plant populations in fragmented landscapes: restoration

806 or gardening? *Aus. J. Bot.* **55**: 371-374.

807 Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection

808 approaches, common sources and consequences for paternity exclusion. *Mol.*

809 *Ecol.* **14**: 599-612.

810 Imbert E, Lefèvre F (2003) Dispersal and gene flow of *Populus nigra* (Salicaceae)

811 along a dynamic river system. *J. Ecol.* **91**: 447-456.

812 Janzen DH (1986) Blurry catastrophes. *Oikos* **47**: 1-2.

813 Jones AG, Ardren WR (2003) Methods of parentage analyses in natural

814 populations. *Mol. Ecol.* **12**: 2511-2523.

815 Julien M-P, Alard D, Balent G (2006) Patterns of ash (*Fraxinus excelsior* L.)

816 colonization in mountain grasslands: the importance of management practices.

817 *Plant Ecology* **183**: 177-189.

818 Jump AS and Penuelas J (2006) Genetic effects of chronic habitat fragmentation in

819 a wind-pollinated tree. *Proc. Natl. Acad. Sci. USA* **103**: 8096-8100.

820 Klein E, Lavigne C, Gouyon P-H (2006) Mixing of propagules from discrete sources

821 at long distances: comparing a dispersal tail to an exponential. *BMC Ecol.* **6**: 3.

822 Kokko H, Lopez-Sepulcre A (2006) From individual dispersal to species range:

823 perspectives in a changing world. *Science* **313**: 789-790.

824 Lefort F, Brachet S, Frascaria-Lacoste N, Edwards KJ, Douglas GC (1999)

825 Identification and characterization of microsatellite loci in ash (*Fraxinus excelsior*

826 L.) and their conservation in the olive family (Oleaceae). *Mol. Ecol.* **8**: 1075-1092.

827 Lowe AJ, Boshier DH, Ward M, Bacles CFE, Navarro C (2005) Genetic resource

828 impacts of habitat loss and degradation: reconciling empirical evidence and

829 predicted theory for neotropical trees. *Heredity* **95**: 255-273.

830 Marshall TC, Slate J, Kruuk LEB, Pemberton J (1998) Statistical confidence for

831 likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7**: 639-655.

832 Meagher TR (2007) Paternity analysis in a fragmented landscape. *Heredity* **99**:

833 563-564.

834 Morand ME, Brachet S, Rossignol P, Dufour J, Frascaria-Lacoste N (2002) A

835 generalized heterozygote deficiency assessed with microsatellites in French

836 common ash populations. *Mol. Ecol.* **11**: 377-385.

837 Morissey MB, Wilson AJ (2005) The potential costs of accounting for genotypic

838 errors in molecular parentage analyses. *Mol. Ecol.* **14**: 4111-4121.

839 Nathan R (2006) Long-distance dispersal of plants. *Science* **313**: 786-788.

840 Nielsen R, Mattila DK, Clapham PJ, Palsbøll PJ (2001) Statistical approaches to

841 paternity analysis in natural populations and applications to the north Atlantic

842 Humpback whale. *Genetics* **157**: 1673-1682.

843 O'Connell LM, Mosseler A, Rajora OP (2006) Impacts of forest fragmentation on

844 the mating system and genetic diversity of white spruce (*Picea glauca*) at the

845 landscape scale. *Heredity* **97**: 418-426.

846 Oddou-Muratorio S, Houot M-L, Demesure-Musch B, Austerlitz F (2003) Pollen

847 flow in the wild service tree, *Sorbus torminalis* (L.) Crantz. I. Evaluating the

848 paternity analysis procedure in continuous populations. *Mol. Ecol.* **12**: 3427-3439.

849 Ritland K (2002) Extensions of models for the estimation of mating systems using n

850 independent loci. *Heredity* **88**: 221-228.

851 Signorovitch J, Nielsen R (2002) PATRI-paternity inference using genetic data.

852 *Bioinformatics* **18**: 341-342.

853 Slate J, Marshall T, Pemberton J (2000) A retrospective assessment of the

854 accuracy of the paternity inference program CERVUS. *Mol. Ecol.* **9**: 801-808.

855 Slavov GT, Howe GT, Gyaourova AV, Birkes DS, Adams WT (2005) Estimating

856 pollen flow using SSR markers and paternity exclusion: accounting for mistyping.

857 *Mol. Ecol.* **14**: 3109-3121.

858 Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees. *For. Ecol.*

859 *Manage.* **197**:21-38.

860 Sork VL, Davis FW, Smouse PE, Apsit PE, Dyer RJ, Fernandez, JF *et al.* (2002)

861 Pollen movement in declining populations of California Valley oak, *Quercus lobata:*

862 where have all the fathers gone? *Mol. Ecol.* **11**: 1657-1668.

863 Sork VL, Nason JD, Campbell DR, Fernandez JF (1999) Landscape approaches to
864 historical and contemporary gene flow in plants. *Trends Ecol. Evol.* **14**: 219-224.

865 Sork VL, Smouse PE (2006) Genetic analysis of landscape connectivity in tree

866 populations. *Landsc. Ecol.* **21**: 821-836.

867 Wagner AP, Creel S, Kalinowski ST (2006) Estimating relatedness and

868 relationships using microsatellite loci with null alleles. *Heredity* **97**: 336-345

869 White GM, Boshier DH, Powell W (2002) Increased pollen flow counteracts

870 fragmentation in a tropical dry forest: An example from *Swietenia humilis* Zuccarini.

871 *Proc Natl. Acad. Sci. USA* **99**: 2038-2042.

872 Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration:

873 F_{st} different 1/(4Nm+1). *Heredity* **82**: 117-125.

874 Young A, Boyle T, Brown AHD (1996) The population genetic consequences of

875 habitat fragmentation for plants. *Trends. Ecol. Evol.* **11**: 413-418.

876 **Titles and legends of figures**
877

878 **Figure 1** Distribution of *Fraxinus excelsior* mature trees in Moffat Dale remnants.

879 Each dot represents a tree. A) Mature trees grow in small forest remnants,

880 confined to steep slopes (elevations are given in meters) along streams

881 (highlighted in dark lines). An exhaustive sampling and mapping was performed in

882 remnants CMa (*N*=4), CDa (*N*=30) and SCa (*N*=12) while in two larger remnants

883 SBa and Wa which include approximately 50 mature trees each, 20 individuals

884 were sampled throughout each of them as potential sources of immigrant pollen

885 flowing into remnants CMa, CDa and SCa. Two lone trees of the Carrifran Burn

886 (Labelled A and B) were also sampled. B to F) Close up of spatial distribution of

887 individuals sampled in remnant CDa, CMa, SCa, SBa and Wa respectively. In CDa,

888 CMa, and Sca, all individuals producing fruits in 2000 are represented by a star: 30

889 seeds were collected throughout the tree canopy from 11, 2 and 6 trees in CDa,

890 CMa and SCa respectively. The background map is a section of Ordnance Survey

891 product Land-line.Plus-nt11 © Crown copyright Ordnance Survey. An EDINA

892 digimap / JISC supplied service. © Figure 1 was originally published in Bacles *et al.*

893 (2005) reprinted with kind permission from *Evolution* (Blackwell Publishing)..

894

895

896 **Figure 2** Comparison of the frequency distribution of possible (white) and detected

897 (black) effective pollen dispersal events for *Fraxinus excelsior* in relation to the

898 location of pollen donors within the fragmented landscape of Moffat Dale. 2a)

899 Detection within three censed remnants. Possible dispersal distances were

900 estimated from Euclidian distances between 141 assignable seeds, spatially

901 located on 19 mother trees, and all 48 *F. excelsior* candidate pollen parents

902 sampled in remnants CDa, CMa and SCa of Moffat Dale (Figure 1). Detected

903   pollen dispersal distances were estimated from Euclidian distances between the

904   141 assignable seeds and their likely father when identified by means of maximum-

905   likelihood paternity analysis in FAMOZ (Gerber et al. 2003, Table 2). 2b) Detection

906   including neighbouring sources of immigrant pollen. Possible dispersal distances

907   were estimated from Euclidian distances between 163 assignable seeds, spatially

908   located on 19 mother trees, and all 88 *F. excelsior* candidate pollen parents

909   sampled in all five remnants of Moffat Dale (Figure 1). Detected pollen dispersal

910   distances were estimated from Euclidian distances between the 163 assignable

911   seeds and their likely father when identified by means of maximum-likelihood

912   paternity analysis in FAMOZ (Gerber et al. 2003, Table 3). In both situations, when

913   more than one likely father was identified (unresolved assignment), a fraction of the

914   seed was assigned to all likely fathers evenly and proportionally to the number of

915   likely fathers (Supplementary Table S1 and S2). Distance distributions of detected

916   pollen dispersal were found to differ significantly from random dispersal (Wilcoxon

917   two-sided signed rank test, *p*-value <0.05 for both $n =141$ and $n =163$).

918

919   **Figure 3** Schematic map of pollen-mediated genetic exchange among three

920   *Fraxinus excelsior* forest remnants varying in their population size, density and

921   degree of spatial isolation to other forest remnants in the mosaic landscape of

922   Moffat Dale. Estimates of gene movement within remnants (continuous white

923   arrows), of gene flow among remnants (continuous black arrows) and of gene

924   immigration from external sources (dashed white arrows) are based on results of

925   ML paternity analysis performed in FAMOZ (Gerber et al. 2003) 422 seeds sampled

926   from all 19 seeding trees in remnants CDa (N_{trees}=30), CMa (N_{trees}=4) and SCa

927   (N_{trees}=12), considering all 48 trees occurring within the landscape, including two

928   isolated trees (A and B, Figure 1) as potential pollen donors (Table 2). Relevant

929   potential geographic barriers to gene flow among remnants are highlighted:

930 Remnants CDa and CMa are located in close proximity (600 m) at the bottom of

931 the bare and open valley while remnant SCa is most isolated over a ridge (dashed

932 white rectangle) located about 1700 m away and surrounded by a dense closed

933 conifer plantation (continuous black square).

934   FIGURE 1

935

936　FIGURE 2

2a) 48 candidate fathers

2b) 88 candidate fathers

937　Inter-distance classes between seeds and most-likely pollen parents (m)
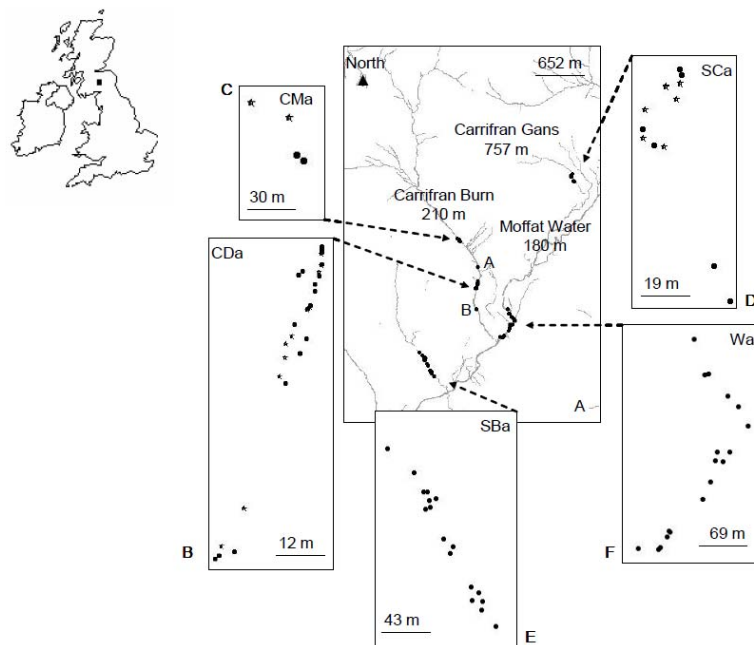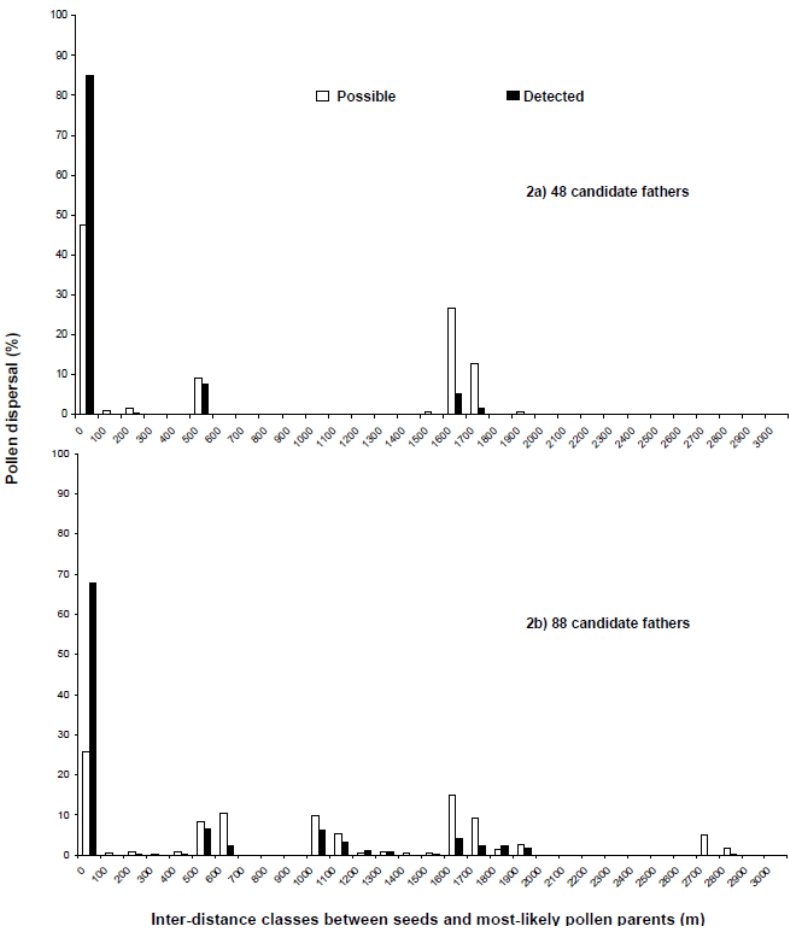
938   FIGURE3

939

940 **Table 1.** Estimates of genotyping error and paternity exclusion probabilities (PEP) at each of five *Fraxinus excelsior* microsatellite loci and overall,

941 computed in CERVUS 2.0 (Marshall et al. 1998) and based on genotyping of 483 seeds sampled from 19 mothers and of 88 trees with no transformation

942 of genotypes (RAW), with binning of rare alleles (BIN), with binning of rare alleles and inclusion of a generalised null allele (NULL). Detailed procedures for

943 allele binning and null allele inclusion are given in material and methods section. Genotyping error estimates based on comparison of the genotype of

944 the 483 seeds with that of their mother. The number of mismatching seed genotypes is given ($N_{mismatch}$) but the number of comparisons may be under

945 483 when values are missing ($N_{comparison}$). Lowest estimates of genotyping error for each locus are highlighted in bold.

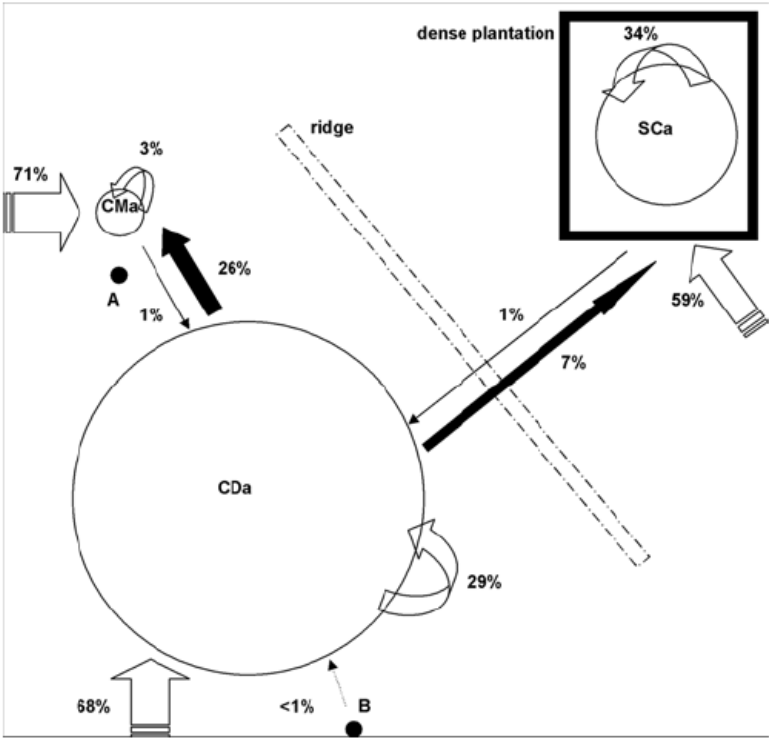| Dataset	RAW			BIN			BINNULL		
Locus	$N_{mismatch}/N_{comparison}$	error	PEP	$N_{mismatch}/N_{comparison}$	error	PEP	$N_{mismatch}/N_{comparison}$	error	PEP
3.1	129/431	0.2847	0.692	118/431	0.2679	0.680	69/483	**0.1499**	0.653
FEMSATL5	132/434	0.2237	0.808	102/434	0.2133	0.713	39/483	**0.0762**	0.685
1.19	13/469	0.0332	0.594	9/469	0.0292	0.504	12/483	**0.0289**	0.606
M2.30B	26/460	0.0378	0.855	23/460	**0.0347**	0.837	45/483	0.0641	0.841
FEMSATL2	34/481	0.0463	0.864	32/481	**0.0443**	0.857	34/483	0.0457	0.870
Overall (excl. 3.1 and 5)	-	0.0391	0.992	-	0.0361	0.988	-	0.0462	0.992
Overall (all)	-	0.1251	0.999	-	0.1179	0.999	-	0.0732	0.999

42

946 **Table 2.** Comparison of global results of paternity analysis of 422 *Fraxinus excelsior* seeds sampled from 19 mother trees and 48 candidate fathers and
947 of their translation into percentage of apparent pollen-mediated gene flow into three forest remnants of the Moffat Dale catchment for a range of
948 paternity assignment methods and of microsatellite genotype transformations. Simple exclusion (SE) and maximum-likelihood (ML) paternity analyses
949 were conducted in FAMOZ (Gerber *et al.,* 2003). Lod of the odds score (LOD) thresholds (TF) for ML paternity assignment and Type I error of false
950 rejection and Type II error of false assignment were determined by means of 50 000 simulations in FAMOZ. Both SE and ML methods were applied to
951 seeds genotyped at microsatellites FEMSATL2 (Brachet *et al.*, 1999), M2.30B and 1.19 (Lefort *et al.*, 1999) that displayed a multilocus genotype
952 compatible with the multilocus genotype of their mother, with no further transformation (RAW), after allele binning at all three loci (BIN), after additional
953 inclusion of a generalised null allele at locus 1.19 (BIN3NULL1). Results highlighted in bold were retained for detailed analysis of genetic connectivity
954 among remnants.

Dataset	Paternity assignment method	TF	Type I error	Type II error	N_{seed} Excluded	N_{seed} Not assigned	N_{seed} Assigned, resolved	N_{seed} Assigned, unresolved	Apparent pollen flow
					LOD=0	0<LOD<TF	LOD>TF, unique	LOD>TF, multiple	LOD≤TF
RAW		0	-	-	219	-	120	83	52%
BIN	SE	0	-	-	183	-	125	114	43%
BIN3NULL1		0	-	-	224	-	119	79	53%
RAW		3.20	<0.10	<0.13	219	68	102	33	68%
BIN	ML	2.50	<0.07	<0.28	183	41	135	63	53%
BIN3NULL1		**2.90**	**<0.05**	**<0.22**	**224**	**57**	**92**	**49**	**67%**

955 **Table 3.** Comparison of fractional and maximum likelihood estimation of the contribution of the five *Fraxinus excelsior* forest remnants of Moffat Dale to
956 effective pollination of seeds sampled in three of them. Results are based on the 422 seeds sampled in remnants CDa, CMa and SCa which displayed
957 a genotype compatible with that of their mother, transformed to minimise genotyping error. [†]20 trees were genotyped in each of SBa and Wa as an
958 indication of potential local sources of immigrant pollen. The most-likely number of seeds sired in each remnant was determined by means of maximum-
959 likelihood paternity analysis in FAMOZ (Gerber *et al.,* 2003) considering all 88 trees genotyped and assigning paternity at a threshold TF of 2.92 (Type I
960 error <0.05, Type II error <0..28). The fractional-likelihood paternity analysis was computed in PATRI (Signorovitch and Nielsen, 2002). The posterior
961 expectation of the number of seeds fertilised by trees from each of five forest remnants was estimated for a given population size (*N*) of 88 (the number
962 of trees genotyped), of 150 (approximating the total number of trees occurring in Moffat Dale) and modelled as a uniform function of 100-500
963 individuals.

		Fractional likelihood assignment in PATRI			Maximum likelihood assignment in FAMOZ
		Posterior expectations of the number of sampled seeds			Most likely number of seed sired
	Prior N	88	150	Uniform [100-500]	88
	CDa	88	56	23	91
Remnant	CMa	3	2	1	4
	SCa	32	28	19	35
Gene flow[†]	SBa	15	11	6	18
	Wa	8	7	4	15
All		146	104	53	163

964 **Table 4.** Comparison of pollen and total gene flow estimates from maximum-likelihood (ML) paternity analysis of non-dispersed seeds with estimates

965 from ML parent-pair analysis of established seedlings in three *Fraxinus excelsior* forest remnants of Moffat Dale. Forest remnants CDa, CMa and SCa

966 were exhaustively sampled for 48 adult trees which were all considered as potential pollen donors and seeds were sampled from all trees producing

967 fruits, respectively, 11, 2 and 6 trees in CDa, CMa and SCa. [#]sample size given as number of seed families/total number of seeds. [*]Results of ML

968 parent-pair analysis of seedlings establishing in *F. excelsior* remnants CDa, CMa and SCa were previously reported in Bacles *et al.* (2006). [§]See

969 Equation 1 estimating total gene flow from progeny arrays (T_p) and where A is the number of local seeds fertilised by immigrant pollen and C is the total

970 number of seeds sampled [†]See Equation 2 estimating total gene flow from established seedlings (T'_s) and where A' is the number of seeds fertilised by

971 immigrant pollen, B' is the number of immigrating seeds and C' is the total number of established seedlings sampled.

	progeny arrays: ML paternity analysis *ante* seed dispersal					established seedlings: ML parent-pair analysis[*] *post* establishment						
	origin of pollen (N_{seed})			gene flow estimates		origin of pollen and seed (N_{seed})				gene flow estimates		
	sample size[#]	local seed		pollen flow	total gene flow via pollen	sample size	local seed		foreign seed	potential pollen flow	realised pollen flow	total gene flow
		local pollen	foreign pollen				local pollen	foreign pollen				
remnant	$C^{§}$		$A^{§}$		$T_p{}^{§}$	$C'^{†}$		$A'^{†}$	$B'^{†}$	*local seed*		$T'_s{}^{†}$
CDa	11/282	81	201	**71%**	**35%**	20	3	7	10	**70%**	**17.5%**	**67.5%**
CMa	2/32	1	31	**97%**	**48%**	20	0	5	15	**100%**	**12.5%**	**87.5%**
SCa	6/108	37	71	**66%**	**33%**	20	2	7	11	**78%**	**17.5%**	**72.5%**

972

973 **Supplementary Table S1.** Individual contribution of *Fraxinus excelsior* trees to effective pollination of non-dispersed seeds sampled in three forest
974 remnants of the Moffat Dale catchment. Results are based on the 422 seeds sampled in remnants CDa, CMa and SCa which displayed a genotype
975 compatible with that of their mother, transformed to minimise genotyping error, considering all 48 sires growing in the three remnants. The most-likely
976 number of seeds sired by individual trees was determined by means of maximum-likelihood paternity analysis in FAMOZ (Gerber et al. 2003) considering
977 48 candidates genotyped and assigning paternity at a threshold TF of 2.90 (Type I error <0.05, Type II error <0.22). When more than one likely father
978 was identified (unresolved assignment), a fraction of the seed was assigned to all likely fathers evenly and proportionally to the number of likely fathers.
979

						MOTHER TREE																
				CDa									CMa			SCa							All
SIRE	102	105	108	111	112	118	123	125	126	129	130	CDa all	24	26	CMa all	33	34	35	36	38	41	SCa all	
n	22	29	23	29	24	21	27	26	23	30	29	**283**	20	11	**31**	22	2	28	24	3	29	**108**	**422**
19					0.33	1.00		9.33	8.00	1.00		19.67											**19.67**
130	1.00	1.00					4.00			1.00		7.00											**7.00**
107			1.00	2.00						3.00		6.00											**6.00**
109		0.33	1.33		2.00	1.33	1.00					6.00											**6.00**
114		1.00		1.00		1.00				2.33		5.33							0.50			0.50	**5.83**
110		3.00		0.75	1.00				1.00			5.75											**5.75**
113								1.00		0.50		1.50	0.50		0.50	1.00		1.00	0.67			2.67	**4.67**
116		1.00		1.00							1.00	3.00	1.50		1.50								**4.50**
119	0.33	1.00				1.00				2.00		4.33											**4.33**
115	0.33			0.33	1.00			0.83		0.50		3.00	1.00		1.00								**4.00**
125		0.33	0.33	1.00		1.33	1.00					4.00											**4.00**
120													1.00	2.00	3.00								**3.00**
101				0.25								0.25		0.50	0.50			1.00	1.00			2.00	**2.75**
111				1.50							1.00	2.50											**2.50**
129		0.39		1.00	1.00							2.39											**2.39**
104		0.25								0.83		1.08	0.50		0.50				0.50			0.50	**2.08**
118		0.14										0.14	0.50		0.50				0.17		1.00	1.17	**1.81**

	ID												Sum			Sum							Sum	Total
	127						0.50		0.50		0.50		1.50											**1.50**
	128				0.50	1.00							1.50											**1.50**
	102		0.14		1.25								1.39											**1.39**
	112	0.33	1.00										1.33											**1.33**
	105		0.48	0.33			0.33						1.14							0.17			0.17	**1.31**
	126		1.00		0.25								1.25											**1.25**
	124		1.00										1.00											**1.00**
	121		0.14										0.14	0.50		0.50				0.17			0.17	**0.81**
	122		0.14										0.14							0.17			0.17	**0.31**
	103																							
	108																							
	117																							
	123																							
	CDa all	**2.00**	**12.36**	**3.00**	**7.50**	**7.67**	**6.50**	**8.00**	**11.67**	**9.00**	**11.67**	**2.00**	**81.36**	**5.50**	**2.50**	**8.00**	**1.00**		**2.00**	**3.33**		**1.00**	**7.33**	**96.69**
CMa	26				0.50	1.33			1.33				3.17											**3.17**
CMa	24													0.50	0.50	1.00								**1.00**
CMa	27																							
CMa	28																							
	CMa all				**0.50**	**1.33**			**1.33**				**3.17**	**0.50**	**0.50**	**1.00**								**4.17**
SCa	32																3.50	1.00	4.00	5.00			13.50	**13.50**
SCa	37																0.50		7.00	4.00		1.00	12.50	**12.50**
SCa	42																3.00			1.00			4.00	**4.00**
SCa	33																3.00			0.50			3.50	**3.50**
SCa	38		0.14									1.00	1.14					1.00		0.17		1.00	2.17	**3.31**
SCa	43											0.50	0.50											**0.50**
SCa	41																			0.50			0.50	**0.50**
SCa	40						0.50						0.50											**0.50**

36																	0.50					0.50	**0.50**
31										0.50	0.50												**0.50**
34		0.25									0.25												**0.25**
35																							
SCa all		0.39				0.50					2.00	2.89				10.00	2.00	11.00	11.67	0.00	2.00	36.67	**39.56**
B		0.25								0.33		0.58											**0.58**
A																							
All sires	**2**	**13**	**3**	**8**	**9**	**7**	**8**	**13**	**9**	**12**	**4**	**88**	**6**	**3**	**9**	**11**	**2**	**13**	**15**	**0**	**3**	**44**	**141**

980

48

981 **Supplementary Table S2.** Individual contribution of *Fraxinus excelsior* trees to effective pollination of non-dispersed seeds sampled in three forest
982 remnants of the Moffat Dale catchment. Results are based on the 422 seeds sampled in remnants CDa, CMa and SCa which displayed a genotype
983 compatible with that of their mother, transformed to minimise genotyping error, considering 88 sires. 20 trees were genotyped in each of remnant SBa
984 and Wa as an indication of potential local sources of immigrant pollen. The most-likely number of seeds sired by individual trees was determined by
985 means of maximum-likelihood paternity analysis in FAMOZ (Gerber et al. 2003) considering all 88 trees genotyped and assigning paternity at a threshold
986 TF of 2.92 (Type I error <0.05, Type II error <0..28). When more than one likely father was identified (unresolved assignment), a fraction of the seed was
987 assigned to all likely fathers evenly and proportionally to the number of likely fathers.
988

		MOTHER TREE																						
		CDa												CMa			SCa							All
	SIRE	102	105	108	111	112	118	123	125	126	129	130	**CDa all**	24	26	**CMa all**	33	34	35	36	38	41	**SCa all**	
	n	*22*	*29*	*23*	*29*	*24*	*21*	*27*	*26*	*23*	*30*	*29*	***283***	*20*	*11*	***31***	*22*	*2*	*28*	*24*	*3*	*29*	***108***	***422***
CDa	19					0.25	1.00		9.33	8.00	1.00		19.58											**19.58**
	130	1.00	1.00					4.00			1.00		7.00											**7.00**
	109		0.33	1.33		2.00	1.33	1.00					6.00											**6.00**
	110		3.00		0.70	1.00				1.00			5.70											**5.70**
	107			1.00	1.50						2.50		5.00											**5.00**
	114		0.50			0.50		1.00			2.25		4.25							0.50			0.50	**4.75**
	113								1.00		0.50		1.50	0.50		0.50	1.00		1.00	0.50			2.50	**4.50**
	119	0.33	1.00					1.00			2.00		4.33											**4.33**
	125		0.33	0.33	1.00		1.33	1.00					4.00											**4.00**
	115	0.33				0.25	1.00		0.83		0.50		2.92	1.00		1.00								**3.92**
	120													1.00	2.00	3.00								**3.00**
	116		0.67			0.50						1.00	2.17	0.75		0.75								**2.92**
	101			0.20									0.20		0.50	0.50			1.00	1.00			2.00	**2.70**
	111			1.50								1.00	2.50											**2.50**
	118		0.14										0.14	0.50		0.50				0.17		1.00	1.17	**1.81**
	129		0.24			1.00	0.50						1.74											**1.74**

	ID												Sub			Sub						Sub	Total
	104		0.10								0.75		0.85	0.50		0.50			0.33			0.33	**1.68**
	128				0.50	1.00							1.50										**1.50**
	102		0.15		1.20								1.35										**1.35**
	112	0.33	1.00										1.33										**1.33**
	127						0.33		0.50		0.50		1.33										**1.33**
	105		0.48	0.33			0.33						1.14						0.17			0.17	**1.31**
	126		1.00		0.20								1.20										**1.20**
	121		0.14										0.14	0.50		0.50			0.17			0.17	**0.81**
	124		0.67										0.67										**0.67**
	122		0.14										0.14						0.17			0.17	**0.31**
	103																						
	108																						
	117																						
	123																						
	CDa all	2.00	10.90	3.00	6.80	6.50	5.83	8.00	11.67	9.00	11.00	2.00	76.70	4.75	2.50	7.25	1.00	2.00	3.00	1.00		7.00	90.95
CMa	26				0.50	1.25			1.33				3.08										**3.08**
	24													0.25	0.50	0.75							**0.75**
	27																						
	28																						
	CMa all				0.50	1.25			1.33				3.08	0.25	0.50	0.75							**3.83**
SCa	32																3.50	1.00	4.00	5.00		13.50	**13.50**
	37																0.50		7.00	4.00	1.00	12.50	**12.50**
	42																1.75			1.00		2.75	**2.75**
	33																1.75			0.50		2.25	**2.25**
	38		0.14								0.50		0.64					0.50		0.17	0.50	1.17	**1.81**
	41																			0.50		0.50	**0.50**
	36																			0.50		0.50	**0.50**

	2	11	3	7	8	6	8	13	9	11	3	82	5	3	8	9	2	13	15	0	3	40	130
43										0.33		0.33											**0.33**
40						0.33						0.33											**0.33**
31										0.33		0.33											**0.33**
34		0.10										0.10											**0.10**
35																							
SCa all		0.24				0.33				1.16		1.73	7.50	1.50	11.00	11.67					1.50	33.17	**34.90**
B		0.10								0.25		0.35											**0.35**
A																							
CDa+CMa+SCa	2	11	3	7	8	6	8	13	9	11	3	82	5	3	8	9	2	13	15	0	3	40	**130**
20		2.67		0.50	0.50							3.67	0.50		0.50								**4.17**
1			3.00				1.00					4.00											**4.00**
6	0.33	0.50		0.70	0.50					0.50		2.53											**2.53**
4		0.10								0.25		0.35		1.00	1.00								**1.35**
7	0.33			0.50	0.50							1.33											**1.33**
2									1.00			1.00	0.25		0.25								**1.25**
18	1.00											1.00											**1.00**
9					0.50							0.50											**0.50**
10				0.50								0.50											**0.50**
12					0.50							0.50											**0.50**
16																0.50						0.50	**0.50**
11													0.25		0.25								**0.25**
15					0.25							0.25											**0.25**
5		0.10										0.10											**0.10**
17		0.10										0.10											**0.10**
3																							

(Row group label: SBa)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Sum	Total
	8																			
	13																			
	14																			
	19																			
SBa all		1.67	3.47	3.00	2.20	1.25	1.50	1.00	1.00	0.75	15.83	1.00	1.00	2.00	0.50			0.50		**18.33**
	11															0.50	1.00	0.50	2.00	**2.00**
	15		1.00								1.00						1.00		1.00	**2.00**
	16	0.33		0.50	0.50						1.33				0.50				0.50	**1.83**
	13									0.33	0.33						1.00		1.00	**1.33**
	12														1.25				1.25	**1.25**
Wa	19														1.25				1.25	**1.25**
	4																1.00		1.00	**1.00**
	6									1.00	1.00									**1.00**
	14												1.00	1.00						**1.00**
	18											1.00		1.00						**1.00**
	1									0.50	0.50									**0.50**
	7																0.33		0.33	**0.33**
	3		0.10								0.10									**0.10**
	8		0.10								0.10									**0.10**
	17		0.10								0.10									**0.10**
	2																			
	5																			
	9																			
	10																			
	20																			

Wa all	0.33	1.30		0.50		0.50						1.83	4.47	1.00	1.00	2.00	3.00	0.50		3.33		1.50	8.33	**14.80**
SBa+Wa	2	5	3	3	1	2	1	0	1	1	2	20	2	2	4	4	1	0	3	0	2	9	**33**
All sires	**4**	**16**	**6**	**10**	**9**	**8**	**9**	**13**	**10**	**12**	**5**	**102**	**7**	**5**	**12**	**12**	**2**	**13**	**18**	**0**	**4**	**49**	**163**

989
990

53