Maternity assignment and queen replacement in a social wasp

JOHN M. PETERS†, DAVID C. QUELLER§, JOAN E. STRASSMANN AND CARLOS R. SOLÍS

Department of Ecology and Evolutionary Biology, Rice University, P.O. Box 1892, Houston, Texas 77251, U.S.A.

SUMMARY

Assigning offspring to parents is important for understanding the evolution of reproductive conflicts and cooperation, particularly in the model systems represented by social insects. Molecular genetic markers are often used to exclude, and occasionally used to assign, candidate parents. However, their use in social insects has been unsatisfactory so far because candidate mothers are often highly related and candidate fathers are unknown. Here, we show that microsatellite loci can be scored from each mother's stored sperm permitting effective maternity assignment. The theoretical power of this method is huge, and we demonstrate its practical utilization in this large-scale study of the wasp, *Polistes annularis*. All 219 genotyped daughters were either assigned to a unique mother or shown to be the progeny of an uncollected dead mother. The data reveal an unexpectedly high number of changes in reproductive dominance. Maternity assignments using this method should help solve many difficult questions in social evolution.

1. INTRODUCTION

Molecular genetic markers have proven valuable in detecting the crucial elements of kinship which are difficult to observe directly. For example, the study of bird mating systems were transformed when multilocus DNA fingerprinting revealed that frequently a social father is not the genetic father (Birkhead & Møller 1992). Actual determinations of parentage are even more useful (Gibbs *et al.* 1990); they are rarer, however, because they require the exclusion of large numbers of candidates (and this often includes some which have not been genotyped).

The problem of parentage assignment has been most severe in one of the groups where it would be most useful: the social insects. Social insects provide model systems for understanding the evolution of reproductive altruism, but this understanding is hindered by incomplete knowledge of the level of direct reproduction by the members of a colony. Unfortunately, effective maternity assignments in social insects have remained beyond the reach of molecular markers, except for certain simple cases (Ross 1988; Mueller 1994). There are several reasons for this. First, the candidate mothers are often related and so their progeny have quite similar genotypes. Second, some mothers may be dead, so the methods need to be powerful enough to avoid false assignment of their progeny to surviving females. Third, for social insects other than termites, the paternal genotypes have been unavailable. The fathers mate and generally die long before their progeny are born.

Microsatellite markers (Queller *et al.* 1993) can overcome these problems, because of two important factors: (i) these single-locus codominant markers are highly variable (in the number of repeats of simple sequence motifs); and (ii) because they are amplified by the polymerase chain reaction (PCR), the DNA samples can be minute. The second advantage allows scoring of very young embryos and, more crucially, the sperm stored in each female's spermatheca (Queller 1993). Evans (1993) showed how genotyping stored sperm allows recovery of alleles of dead mates which can then be used with the maternal genotype to assign maternity. In this report, we demonstrate the theoretical power of the method and report its first large-scale application.

2. THE POWER OF MATERNITY ASSIGNMENT

The logic of maternity assignment is simple. Any candidate mother is excluded for a particular progeny if her genotype, along with those of her mate(s), could not have produced the progeny genotype at any locus. Consider an example under the haplodiploid genetic system shared by ants, bees and wasps (diploid females arise from fertilized eggs and haploid males from unfertilized eggs). An AC daughter cannot come from an AC mother mated to a B male, but she could have come from an AC mother mated to an A male. If all

† Current address: Section of Neurobiology and Behavior, Cornell University, Ithaca, New York 14853, U.S.A.
§ To whom correspondence should be addressed.

candidate mothers but one are excluded, then the remaining one can be assigned as the mother provided all candidate mothers were genotyped, or that missing candidates can be shown to be sufficiently unlikely to match the offspring. Because microsatellites typically are well-behaved codominant loci (used, for that reason, in linkage mapping, see Weissenbach *et al.* 1992), power equations can be derived to estimate the probabilities of matches.

Bearing the points made in the previous paragraph in mind, what then is the probability that the genotype of a randomly chosen diploid daughter is inconsistent with a candidate mother, related by r to the true mother, and her m haploid mates? Assume that mating is random. For a codominant locus with allele frequencies p_i ($\Sigma p_i = 1$), the probability of exclusion is:

$$E_l = 1 - \sum_i p_i^2 [1-(1-p_i)^m][1-(1-r)(1-p_i)^2]$$

$$- \sum_i \sum_{j \neq i} p_i p_j [1-(1-p_j)^m][r+(1-r)p_i^2(1-p_i)]$$

$$- \sum_i \sum_{j \neq i} p_i p_j [1-(1-p_i)^m][p_j+(1-p_j)(1-r)p_j]$$

$$+ \sum_i \sum_{j \neq i} p_i p_j [1-(1-p_i)^m-(1-p_j)^m+(1-p_i-p_j)^m]$$

$$\times [(r+(1-r)p_i)p_j+(1-r)p_i p_j]. \quad (1)$$

The first summation is for homozygous daughters (frequency p_i^2); the two brackets represent the probability of matching at least one of the candidate paternal alleles and one of the maternal alleles, respectively. The remaining three summations have the same general form, but for heterozygous offspring, with the daughter's maternal and paternal alleles indexed by i and j. The first of these is for the candidate father(s) having at least one A_j allele and the candidate mother having at least one A_i; the second is for the reverse configuration; and the final one subtracts out double counts that have occurred for A_iA_j mothers mated to multiple males collectively possessing the same two alleles. Because exclusion is based on the simple presence of alleles, the fraction of sperm contributed by different males is irrelevant. The expression is more complex if it is possible that the offspring is a full sister of the candidate mother (but in our study, the parental and offspring generations are completely distinct, so this issue does not arise).

E_l in equation (1) is the exclusion probability for a single candidate mother using a single marker locus. The overall probability of excluding f candidate mothers with L loci is

$$E_{\text{total}} = \left(1 - \prod_{l=1}^{L} (1-E_l)\right)^f.$$

This is easily solved for the number of loci needed (L) when all E's are the same; figure 1 shows some examples. It shows the number of loci required to exclude, 98% of the time ($E_{\text{total}} = 0.98$), every one of f candidate mothers who are full sisters ($r = 0.75$) of the true mother. If the genotypes of the males are unknown (equivalent to $m = \infty$; no useful information from mate genotypes), too many loci are required;

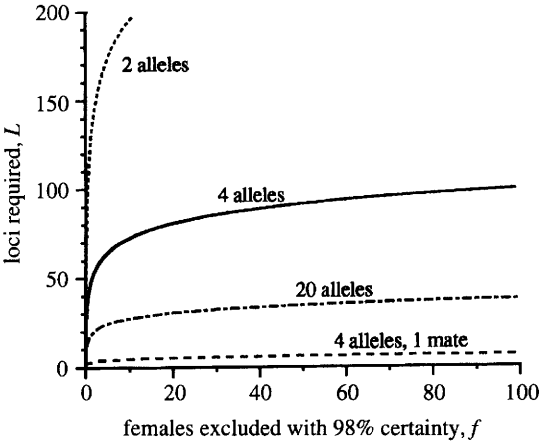

Figure 1. Number of loci required to exclude all f non-maternal females who are full sisters of the actual mother, for 98% of randomly chosen daughters. The examples are chosen to roughly represent good allozyme loci (two equally frequent alleles); moderately variable microsatellites (four equally frequent alleles); and extremely variable microsatellites or minisatellites (20 equally frequent alleles). The top three lines represent cases when only the female parental genotypes are known. The bottom line shows the added power obtained if each female mates once and her mate's genotype is also known.

exclusion of as few as five candidate mothers would require in excess of 170 good allozyme loci (two equally frequent alleles) and 60 moderately variable microsatellites (four equally frequent alleles). Even with extremely variable microsatellite or minisatellite loci (20 equally frequent alleles) more than 20 loci be would required. However, if females mate once and their mate's genotypes are known, it would suffice to score only five of the moderately variable microsatellites and ten such loci will exclude over 1000 candidate mothers (with 98% certainty).

3. MATERNITY ASSIGNMENT IN *POLISTES ANNULARIS*

We put the methodology to a practical test with colonies of the primitively eusocial wasp *Polistes annularis* collected at Lake Travis, Texas, the site of extensive previous studies (Strassmann 1979, 1981, 1983, 1989; Sullivan & Strassmann 1984; Hughes *et al.* 1987; Queller & Strassmann 1988; Queller *et al.* 1990). These colonies had been founded approximately two months before collection by groups of inseminated female foundresses, known to be close relatives (Queller *et al.* 1990). By collecting at night, we were able to capture all living foundresses (3–7 per colony), but other foundresses who died before collection may also have contributed to the brood.

Collected foundresses, their stored sperm, and samples of eggs, larvae, and pupae (no adult progeny had yet emerged) from nine colonies were amplified for seven microsatellite loci (see table 1 and figure 2). For all samples, except the sperm, microsatellites were amplified and visualized as in Hughes & Queller (1993) and Choudhary *et al.* (1993) with minor modifications. Averaged over the seven loci, genotypes were success-

Table 1. *Microsatellite loci used.*

locus name	PCR primers (5′–3′)	annealing temperature /°C	allele percentages
PAN48AAT	see Hughes & Queller (1993)	55	58, 15, 2, 24, 2
PAN69AAT	see Hughes & Queller (1993)	55	2, 63, 17, 18
PAN95TAG	TTTTACCGGGCACTTTGG CGGACATCGAACGTTGTG	55	69, 20, 6, 5
PAN111bAAT	TCGTCATATTTATCGTTGGAGTGGA TGCCTCCCTTTGACTTTTGAGAG	55	4, 17, 3, 34, 3, 5, 32, 1
PAN117AAT	see Hughes & Queller (1993)	55	9, 14, 76, 1
PBE424AAT	GGCCAATTATTATCTCCATGCAATA CGTGCATCCTTCAGAAACAATACTT	51	2, 28, 1, 10, 11, 27, 20, < 1
PBE492AAT	AACGACGAGCGTTAATAATTCATG CACGTCTGTGCATAAGAAGTACGG	55	8, 8, 9, 6, 5, 10, 18, 22, 3, 8, < 1, 1, 3

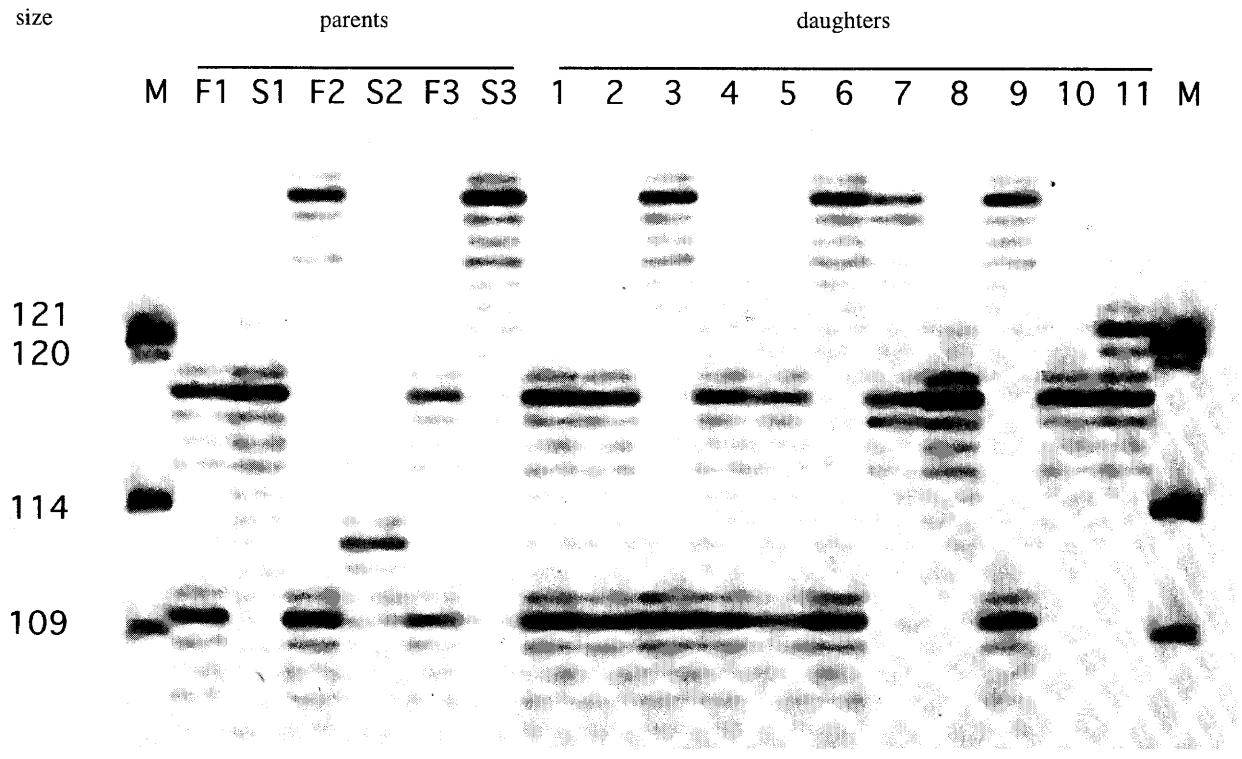

Figure 2. Microsatellite genotypes. The two outside lanes (M) are size markers (the C lane from an M13 sequence) labelled in nucleotides. The interior lanes contain amplified genotypes for locus PAN111bAAT from colony 7. Three parent sets are shown, both foundress (F) and her stored sperm (S). Each sperm lane shows a single allele (each allele shows one strong band plus several weak stutter bands characteristic of microsatellites). The remaining 11 lanes are daughters. Lanes 1, 2, 4, 5, 8 and 10 show daughters of the first foundress and lanes 3, 6, 7, and 9 show daughters of the third foundress; note that these are distinguishable only because of the sperm. Lane 11 matches none of the collected foundresses and is the daughter of an uncollected, presumably dead, foundress.

fully obtained for 99% of foundresses, 96% of pupae, 91% of larvae and 66% of eggs. The lower success rate for eggs seemed to be due to the degradation over time of initially small amounts of DNA (some loci were run before others). A few brood with fewer than three successfully scored loci were omitted from the investigation as there was insufficient information for successful maternity analysis.

Sperm samples were collected and processed as follows. The spermathecae were removed in 10% NaCl and the sperm clump dissected out from the membranes using insect pins. The sperm were then lysed by incubating at 65 °C for 10 minutes in 20 μl of 50 mM Dithiothreitol and 20 μl of 500 mM KOH. This solution was neutralized with 20 μl of 500 mM HCl and 6.66 μl of 500 mM Tris-HCl (pH 9.0): 2 μl of this solution was used in a 10 μl PCR reaction. The PCR method followed is detailed in Hughes & Queller (1993) except that the usual 35–40 cycles were preceded by 23 cycles and the addition of more Taq polymerase. The first five cycles had extended (two-minute) annealing times. Detailed protocols are available on request. 97.9% of our 280 sperm samples amplified (not always on the first attempt); this rises to

99.27% if we exclude one individual with four non-amplified loci, presumably because the DNA sample was poor. Results were replicable for a given sperm sample and locus, showing they were unaffected by any stochastic sampling effects caused by very small numbers of amplified sperm.

Estimates of relatedness were obtained using Queller & Goodnight's (1989) method, as implemented by the Macintosh computer program Relatedness 4.2b. Colonies were weighted equally, with standard errors estimated by jackknifing over colonies. Figure 2 shows an example of the microsatellite genotypes for foundresses, their stored sperm, and some of the colony's daughters. In this and other colonies, the sperm samples usually showed a single allele at each locus, indicating a single mating. Only two of 40 foundresses were mated twice (two alleles at one or more loci in their stored sperm). The males are unrelated to their mates ($r = -0.02$, s.e. $= 0.04$) and to the mates of other females in the same colony ($r = 0.02$, s.e. $= 0.04$). As expected, foundresses were closely related ($r = 0.62$, s.e. $= 0.08$). Substituting our estimated allele frequencies, $m = 1$, $r = 0.62$, and $E_{\text{total}} = 0.99$ into the power equations (1, 2) shows that we can be 99% confident that a match is unique, even if there were 30 other singly mated candidate mothers (including any that died before collection). As *P. annularis* associations rarely exceed ten foundresses (Strassmann 1981), we can be confident that a uniquely assigned mother was indeed the actual mother.

Maternity assignments were attempted for the 219 eggs, larvae, and pupae with at least three loci scored (it was assumed that all subjects were female, although two may have been haploid males because they had only one allele at each locus). At this stage of the season most of progeny function as workers, but there is fitness at stake. The majority of colonies lose all their foundresses before termination of the colony cycle (Queller & Strassmann 1988), then an old worker mates and becomes queen (Hughes *et al.* 1987).

Each of the offspring were either successfully assigned to a single foundress (141 offspring) or excluded for all collected foundresses (78 offspring). The absence of offspring consistent with the genotypes of more than one foundress-sperm pair confirms the predicted power of the method. Corroborative evidence that assignments were correct comes from ovarian dissections which identified a single foundress with well-developed ovaries in seven colonies (colonies 8 and 9 had no collected foundress with mature eggs in her ovaries, probably because the dominant queen had died very recently). These seven queens ought to be the mothers of a large proportion of at least the most recent brood of eggs, and in fact 88% of the eggs were assigned to them.

Assignment of the older brood, however, reveals much richer histories. To incorporate information from the unassigned offspring we used the discovery that foundresses are nearly always singly mated, and grouped unassigned offspring into the largest possible full siblingships, ie. into the largest groups in which all individuals shared one allele (from the putative father) and the other allele had no more than two allelomorphs (from the putative mother). The accuracy of this procedure is difficult to evaluate theoretically but it appears to work well for the level of genetic information in this dataset; when applied to the 141 offspring with assigned mothers it resulted in only three errors. This procedure showed that colonies had as many as five different dead foundresses who contributed to the brood.

The assignments and the full sibling groupings are shown in figure 3, and they reveal surprisingly complex and variable histories of reproductive dominance. Dominance appears considerably weaker early in the season. On average, 86% of the eggs, 69% of the larvae and only 55% of the pupae were daughters of the foundresses dominant at the time their age class was produced. (We do not report formal measurements of reproductive skew (Keller & Vargo 1993; Reeve & Ratnieks 1993) because the numbers of foundresses alive at the three times are unknown, though the number generally declines due to mortality.)

Even more striking are the varying individual histories of queen turnover. Colonies 2, 16, and 17 show a single continuously reigning queen. Others show replacement of an absent queen, presumably after her death (colony 25 and probably 7). In two colonies the queen must have died shortly before collection and no clear successor had yet appeared (colony 9 and probably 8; recall that these two had no foundress with well-developed ovaries). Colonies 8, 10, and 13 show supersedure, with a deposed queen remaining as a subordinate. Two colonies have multiple queen changes. In colony 8, the weakly dominant queen at the time pupae and larvae were initiated was superseded by another female who died shortly before collection, leaving the colony to a third queen, as yet undecided (the assigned eggs in this colony are the only dubious assignments in the dataset; the foundress was uniquely difficult to exclude because only three of her sperm loci amplified, and she was doubly mated). In colony 13, three foundresses still alive at the time of collection engaged in a see-saw battle for reproductive dominance. Female 1 produced most of the pupae, females 2 and 3 most of the older larvae, female 1 the younger larvae, and female 3 the eggs (the larval age distinction is not shown in figure 3). Overall, there were nine queen changes in the nine colonies, considerably more than expected from shorter-term observational studies (Strassmann 1983). Thus, the data indicate that early-season *Polistes annularis* colonies exhibit much richer reproductive histories than previously thought.

But before fully accepting this conclusion it is necessary to consider the likelihood of false exclusions, particularly given the large fraction of offspring that could not be assigned to any collected foundress. False exclusion could result from several causes. First, it could be caused by errors in scoring or data entry, but all genotypes were checked independently by scoring and entering them twice. Second, mutation to non-parental alleles could also cause false exclusions, but not very many. Edwards *et al.* (1992) estimated mutation rates for five trimeric and tetrameric human

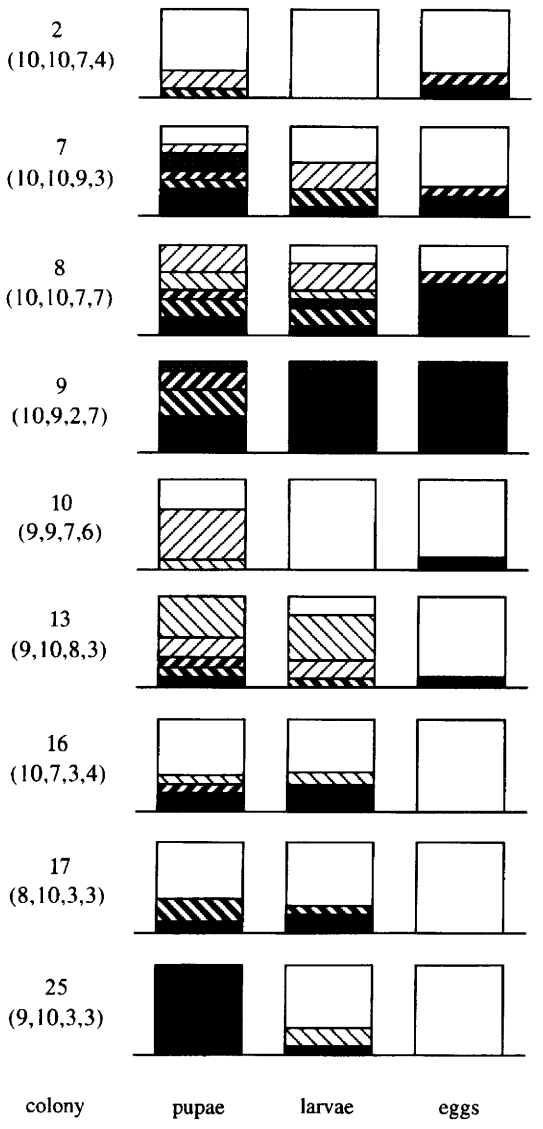

Figure 3. Reproductive success of *P. annularis* foundresses. In each colony, the fraction of brood assigned to a known foundress is shown by the height of one of three histogram segments dominated by white: □, ▨, or ▨. Offspring whose genotypes were incompatible with all known foundresses in their colony were grouped into the largest possible classes consistent with being full sisters; these sibships are shown by five darker segment patterns: ▨, ▨, ▨, ▨, ■. The patterns apply consistently within each row; for example, the large white blocks for colony 2 show that the same mother dominated during production of pupae, larvae, and eggs. The four numbers in brackets listed for each colony are the numbers of genotyped pupae, larvae, eggs, and foundresses (all collected foundresses were genotyped).

microsatellite loci, and the highest was about 10^{-4}. Taking this value for all seven of our loci, the probability that at least one mutation would occur in a given daughter–parent set is only $1-0.9999^{14} = 0.0014$ (the exponent comes from seven loci multiplied by two alleles). In addition, some mutations would not cause false exclusions.

False exclusion might also occur if there are undetected null alleles. A sequence variant at a PCR priming site, if it causes amplification to fail, could result in the scoring of a diploid individual as a homozygote because of its single non-null allele. However, a null allele would cause the haploid sperm not to amplify at all, so the high success rate of our sperm amplifications shows that nulls must be rare or absent.

If false exclusions are caused by either mutation or null alleles, the largest class of excluded offspring would be excluded for only one locus (double exclusions would require the conjunction of two rare events). However, for the consistently amplified brood classes (pupae and larvae), 63 of the 64 excluded daughters were excluded at more than one locus.

Finally, false exclusion could occur if additional males contributed sufficiently small amounts of sperm that their alleles were not detected. But this must be rare when care is taken to expose autoradiographs sufficiently. Second alleles that differ by a single repeat unit from the main sperm band may 'hide' under stutter bands (see figure 2), but this is unlikely to happen at all seven loci.

If maternity assignment uncovers surprising results in this thoroughly studied species, its value should be even greater elsewhere. It should be particularly useful when direct observation of egg-laying is impossible, when females use sneaky reproductive behaviours, when eggs or brood are eaten and when egg-layers contribute differently to workers and reproductives. In addition, parentage assignment data can be used to test theories in many of the most important and difficult areas of social insect biology including within-colony kin discrimination (Getz 1991), the genetic basis of traits (Robinson & Page 1988), the evolution of reproductive division of labour (Reeve & Ratnieks 1993) and queen versus worker control of reproduction (Ratnieks 1988).

We thank K. Barefield and C. R. Hughes for help in developing primer pairs, M. Choudhary for fieldwork and J. P. Field for discussions and comment. This work was partly funded by NSF grants BSR-90215114 and IBN-9210051, and by the Rice Scholars Program. This work is, in part, J. M. P.'s undergraduate honour's thesis.

REFERENCES

Birkhead, T. R. & Møller, A. P. 1992 *Sperm competition in birds.* London, New York: Academic Press.

Choudhary, M., Strassmann, J. E., Solís, C. R. & Queller, D. C. 1993 Microsatellite variation in a social insect. *Biochem. Genet.* **31**, 87–96.

Edwards, A., Hammond, H., Jin, L., Caskey, C. & Chakraborty, R. 1992 Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241–253.

Evans, J. D. 1993 Parentage analyses in ant colonies using simple sequence repeat loci. *Molec. Ecol.* **2**, 393–397.

Getz, W. 1991 The honey bee as a model kin recognition system. In *Kin recognition* (ed. P. Hepper), pp. 358–412. Cambridge University Press.

Gibbs, H. L., Weatherhead, P., Boag, P., White, B., Tabak L. & Hoysak, D. 1990 Realized reproductive success of polygynous red-winged blackbirds revealed by DNA markers. *Science, Wash.* **250**, 1394–1398.

Goodnight, K. F. & Queller, D. C. 1994 *Relatedness*, version 4.2b. Pascal program for Macintosh computers.

Hughes, C. R. & Queller, D. C. 1993 Detection of highly polymorphic microsatellite loci in a species with little enzyme polymorphism. *Molec. Ecol.* **2**, 131–137.

Hughes, C. R., Beck, M. O. & Strassmann, J. E. 1987 Queen succession in the social wasp, *Polistes annularis*. *Ethology* **76**, 124–132.

Keller, L. & Vargo, E. L. 1993 Reproductive structure and reproductive roles in colonies of eusocial insects. In *Queen number and sociality in insects* (ed. L. Keller), pp. 16–44. Oxford University Press.

Mueller, U. G., Eickwort, G. C. & Aquadro, C. F. 1994 DNA fingerprinting analysis of parent–offspring conflict in a bee. *Proc. natn. Acad. Sci. U.S.A.* **93**, 5143–5147.

Queller, D. C. 1993 Genetic relatedness and its components in polygynous colonies of social insects. In *Queen number and sociality in insects* (ed. L. Keller), pp. 132–152. Oxford University Press.

Queller, D. C. & Goodnight, K. F. 1989 Estimation of genetic relatedness using allozyme data. *Evolution* **43**, 258–275

Queller, D. C. & Strassmann, J. E. 1988 Reproductive success and group nesting in the paper wasp *Polistes annularis*. In *Reproductive success: studies of variation in contrasting breeding systems*, (ed. T. H. Clutton-Brock), pp. 76–96. University of Chicago Press.

Queller, D. C., Hughes, C. R. & Strassmann, J. E. 1990 Wasps fail to make distinctions. *Nature, Lond.* **344**, 388.

Queller, D. C., Strassmann, J. E. & Hughes, C. R. 1993 Microsatellites and Kinship. *Trends Ecol. Evol.* **8**, 285–288.

Ratnieks, F. L. W. 1988 Reproductive harmony via mutual policing by workers in eusocial hymenoptera. *Am. Nat.* **132**, 217–236.

Reeve, H. K. & Ratnieks, F. L. W. 1993 Queen–queen conflicts in polygynous societies: mutual tolerance and reproductive skew. In *Queen number and sociality in insects* (ed. L. Keller) pp. 45–85. Oxford University Press.

Robinson, G. E. & Page, R. E. 1988 Genetic determination of guarding and undertaking in honey bee colonies. *Nature, Lond.* **333**, 356–358.

Ross, K. R. 1988 Differential reproduction in multiple-queen colonies of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae). *Behav. Ecol. Sociobiol.* **23**, 341–355.

Strassmann, J. E. 1979 Honey caches help female paper wasps (*Polistes annularis*) survive Texas winters. *Science, Wash.* **204**, 207–209.

Strassmann, J. E. 1981 Wasp reproduction and kin selection: Reproductive competition and dominance hierarchies among *Polistes annularis* foundresses. *Flor. Entommol.* **64**, 74–88.

Strassmann, J. E. 1983 Nest fidelity and group size among *Polistes annularis* foundresses. *J. Kans. Entommol. Soc.* **56**, 621–634.

Strassmann, J. E. 1989 Group colony foundation in *Polistes annularis* (Hymenoptera: Vespidae). *Psyche* **96**, 223–236.

Sullivan, J. D. & Strassmann, J. E. 1984 Physical variability among nest foundresses in the polygynous social wasp *Polistes annularis*. *Behav. Ecol. Sociobiol.* **15**, 249–256.

Weissenbach, J., Gyapay, G., Dib, C., Morissette, J., Millasseau, P., Vaysseix, G. & Lathrop, M. 1992 A second generation linkage map of the human genome. *Nature, Lond.* **359**, 794–801.

Received 15 November 1994; accepted 7 December 1994