Apidologie 36 (2005) 223–229
© INRA/DIB-AGIB/ EDP Sciences, 2005
DOI: 10.1051/apido:2005010

Original article

Using drones for estimating colony number by microsatellite DNA analyses of haploid males in *Apis*[1]

F. Bernhard KRAUS[a]*, Nikolaus KOENIGER[b], Salim TINGEK[c], Robin F.A. MORITZ[a]

[a] Institut für Zoologie, Martin Luther Universität Halle-Wittenberg, Hoher Weg 4, 06099 Halle/Saale, Germany
[b] Institut für Bienenkunde, (Polytechnische Gesellschaft), Fachbereich Biologie und Informatik der J.W. Goethe – Universität Frankfurt, Karl-von-Frisch-Weg 2, 61440 Oberursel, Germany
[c] Agricultural Research Station Tenom, PO Box 197, 89908 Tenom, Sabah, Malaysia

Received 5 October 2004 – Revised 26 November 2004 – Accepted 1 December 2004

Published online 1 June 2005

Abstract – In social insects the number of colonies rather than the actual number of individuals in the population primarily determines the effective population size. Here we present a method where microsatellite data of haploid males can be used to estimate the number of male producing queens in honeybee populations. A cluster analysis based on the allelic identity by descent (AID) among male genotypes is used to group potential brother males. For each "brother cluster" the corresponding mother queen genotype is determined by Mendelian inference. We show in various simulations that although limited number of screened loci can result in slightly biased estimates, the precision improves considerably with increasing number of loci. Empirical data from microsatellite studies of the Western honeybee *Apis mellifera* and the giant Asian honeybee *Apis dorsata* are presented to illustrate the application of the procedure.

***Apis mellifera* / *Apis dorsata* / haploid male / honeybees / microsatellite / population size**

1. INTRODUCTION

Microsatellite DNA-analyses have revolutionized population genetics. The use of multiple highly variable loci allows for precise population analyses including the detection of population growth and decline, bottlenecks in the population history, migration and gene flow. Population sizes are however often difficult to determine and are usually estimated indirectly based on allele diversity. If family structured populations are studied, it is possible to infer the number of families from relatedness estimates among the sampled individuals. Although it is difficult to safely classify individuals with microsatellite loci into full and half-sisters in wild populations if no pedigree information is available (Blouin et al., 1996; Lynch and Ritland, 1999; Wang, 2002), relatedness analyses proved extremely informative in colonies of social hymenoptera where a single diploid queen is mated with one or several haploid males (Strassmann, 2001). For example microsatellite loci were instrumental in detecting the extreme degree of polyandry of honeybee queens (Estoup et al., 1994; Solignac et al., 2003) which can mate with very large numbers of males (Moritz et al., 1995, 1996; Oldroyd, 1997; Wattanachaiyingcharoen et al., 2003).

In highly eusocial monogynous insect populations it is primarily the number of colonies rather than the actual number of sexuals in the population which determines the effective population size (Kerr, 1967; Pamilo and Crozier, 1997; Pamilo et al., 1997). This is particularly true for species where reproduction operates on the colonial level by colony fission. Multiple mating only weakly enhances the effective population size and is limited by $N_e = 2.25$ times the number of colonies (Wright, 1933). The estimation of the number of colonies is

* Corresponding author: kraus@zoologie.uni-halle.de
[1] Manuscript editor: Gudrun Koeniger

therefore an important parameter for social insect population biology. A quite precise and sophisticated approach was given for honeybees by Baudry et al. (1998) where about 240 colonies were estimated to contribute to the male gene pool in a population of *Apis mellifera* L. The estimation procedure was based on the likelihood ratio of two males being brothers as opposed to being unrelated and required some elaborate computation and simulation. Here we present an alternative, simple method where we use male genotypes to determine the number of male producing colonies. A Monte Carlo simulation shows that the approach, similar to that of Li et al. (1993), is useful to estimate the number of reproductively active colonies in social insect populations. We use a fully controlled population of the Western Honeybee (*Apis mellifera*) on a North Sea island to illustrate and test the procedure under conditions where the number of contributing colonies is exactly known. As an additional empirical example, we estimate the number of colonies in a population of the giant honeybee (*Apis dorsata* Fabricius) in Borneo to further illustrate the procedure.

2. METHODS

2.1. Clustering males with similar genotypes

The goal of our analyses was to estimate the number of queens required to produce the male genotypes obtained from the DNA analyses. A simple way to group potential brother males in a cluster analysis is to determine their relatedness based on similarity at the given loci. Males from the same queen are more likely to share alleles than those of different queens. Moreover males sharing rare alleles are more likely to have received these rare alleles from the same mother queen than males sharing frequent alleles. We therefore weighed the allelic identity at the tested loci with the corresponding population wide allele frequency similar to Li et al. (1993; for diploids) to obtain an estimate for "allele identity by descent" (*AID)* between two haploid males *X* and *Y*. If an allele was shared between two males it received an identity value of $I_{xy} = 1$, if they were different $I_{xy} = 0$. If the allele was identical this could either be due to random chance or due to true relatedness. We estimated the relatedness fraction by subtracting the random probability for identical alleles.

$$AID_{xy} = 1 - p_i^2 \qquad (1)$$

where

p_i = frequency of allele i at a given locus.

In case of different alleles at a given locus, two males receive an AID_{xy} score of zero irrespective of the allele frequencies of their alleles. We can now perform this procedure for all loci and then use the arithmetic mean over all identity by descent estimates, $\overline{AID}_{xy}$, as an overall measure for allelic identity by descent between the two males *X* and *Y*:

$$\overline{AID}_{xy} = \frac{1}{n}\left(\sum_{k=1(i=j)}^{n} 1 - p_i^2 + \sum_{k=1(i\neq j)}^{n} 0 \right) \qquad (2)$$

where

n= number of loci

k = locus

i = allele in x at locus k

j = allele in y at locus k

p_i = frequency of allele i.

This identity measure can be transformed into a distance measure

$$D_{xy} = 1 - AID_{xy} \qquad (3)$$

and we obtain a distance matrix with all pairwise distances that can be analysed with standard clustering algorithms. For this analysis we first produced the distance matrix in MS EXCEL and then used the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) algorithm of the STATISTICA® software package for a cluster analysis.

2.2. Estimating number of male-producing queens

The resulting clustering tree plot was subjected to a manual inference of the least possible number of potential queens. For each cluster node we determined if the clustered male genotypes could have been potentially produced by a single diploid queen. If so, we inferred the genotype of the potential mother queen from the male genotypes in that cluster using Mendelian rules. By performing this procedure over all clusters, we received an estimate for a minimum number of male producing colonies explaining the cluster of the tree plot. Of course the minimum number of queens found in the sample relies heavily on the sample size, and has no direct relationship to the actual numbers of colonies in the population. The more males sampled, the more male producing colonies are likely to be detected. The maximum number of colonies detectable is equal to roughly half the number of males sampled, since two males can always be clustered together at random.

2.3. Testing the procedure

We tested the procedure with Monte Carlo computer simulated values and compared these with the true population sizes. We assumed variable numbers (3, 5 and 10) of loci with 10 alleles each and randomly assigned allele frequencies. Variable numbers (10, 20, 30...90) of male producing queens were generated at random and variable number of males (25, 50, and 100) sampled. The distance D_{xy} was computed (3) for all pairs of males, and the distance matrix initially stored in MS EXCEL® and then transferred to a SATISTICA® data file. A cluster analysis (UPGMA) was performed based on these distance matrices. The estimated queen mothers were assigned based on the cluster tree. The number of mismatches (males assigned to the wrong mother) as well as the overall estimate of the population size was determined. We analysed 10 replicates per parameter setting and calculated the mean and standard errors.

2.4. Empirical sample genotyping

2.4.1. The Western honeybee Apis mellifera

We tested the procedure with an experimental island population of *A. mellifera* (eight colonies) on the Island of Neuwerk (Germany) in the North Sea, which is reproductively isolated from the mainland. This closed population approach allowed for a test of the *AID* estimation procedure because our sampling and estimation results could be compared with the actual number of colonies on the island. Drones (n = 78) were caught through matings with 13 virgin queens and the fathering drone genotypes were determined from the worker offspring produced by these queens using Mendelian rules. DNA was extracted from worker brood with routine protocols (Walsh et al., 1991) and genotyped at six loci A7, A24, A28, A88, A107, A113 (Estoup et al., 1994; Solignac et al., 2003) using an ABI 310 sequencer. The distance matrices were prepared with standard spreadsheet software and population allele frequencies were estimated based on the drones alleles.

2.4.2. The Giant Asian honeybee Apis dorsata

A wild *A. dorsata* population was sampled in Sabah (Malaysia). Drones (n = 148) were caught at a drone congregation area near the Bee Research Station in Tenom. The sampled drones were stored in ethanol until DNA extraction. DNA was extracted from all 148 samples with routine protocols and genotyped at three loci A14, A76, A107 (Estoup et al., 1994; Solignac et al., 2003; Moritz et al., 1995) using the methods described for *A. mellifera* above.

3. RESULTS

3.1. Effect of the number of loci

With increasing number of loci the precision of the analysis improved. With ten variable loci the assignment of males was always correct as long as the number of colonies in the population was about half as large as the number of sampled males (Fig. 1). With increasing number of colonies the non-sampling error increases and the precision of the estimate is reduced. Interestingly, even the analyses with three and five loci yielded qualitatively useful results as long as the non-sampling error was low. Although the number of mismatches (the grouping of males which are not brothers) was considerably higher (Fig. 1), a small population could still be discriminated from a large one. The precision was however low with overestimating the size of small populations and underestimating the size of large ones. With three loci it was possible to assign 35.7% ± 14.6% (mean over all numbers of queens ± SD) of the male genotypes correctly to their mother, with 5 loci 58.1% ± 10.7% and with 10 loci 86.2% ± 10.2%.

3.2. Effect of sample size

The number of individuals sampled is more important than the number of loci tested per individual. As soon as the number of queens in a population exceeds half the sample size (males sampled) there is a systematic underestimation of the number of queens. Thus if the sample size is small, high precision analyses with large number of loci do not improve the analyses. Figure 2 shows that the analysis is primarily flawed by the non-sampling error but not by the lack of precision of the genetic analysis.

3.3. Empirical example: *A. mellifera* on Neuwerk

The six loci used were only moderately variable with 3 ± 1.1 alleles per locus on average. The distance matrix obtained by applying equations (2) and (3) to the drone genotypic data,

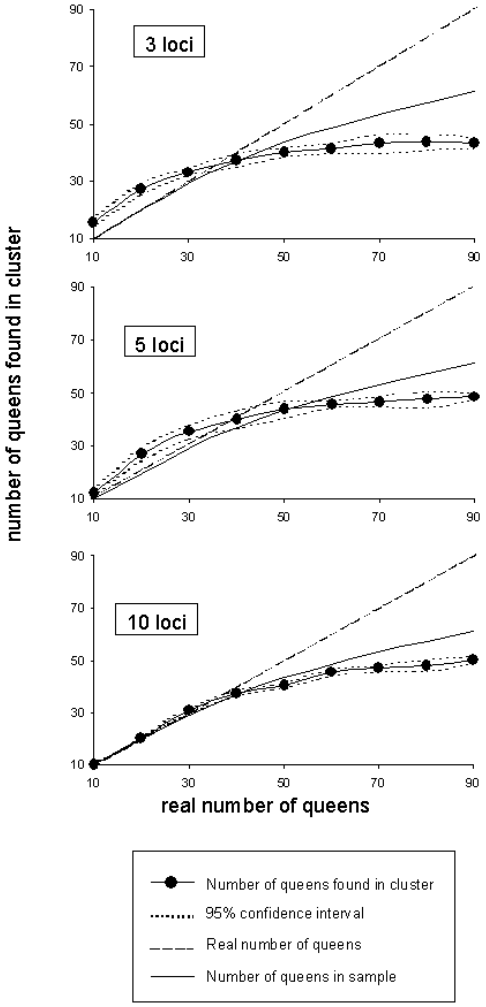

Figure 1. Effect of the number of loci on the precision of the analysis. Given are the graphs of the Monte Carlo simulations with 3, 5 and 10 loci and a random sample of 100 males. The number of queens found in the UPGMA cluster (y-axis) are plotted versus the actual number of queens initially used in the simulation. Also shown (solid line) is the number of queens which were present in the drone sample drawn from the Monte Carlo simulation (non sampling error).

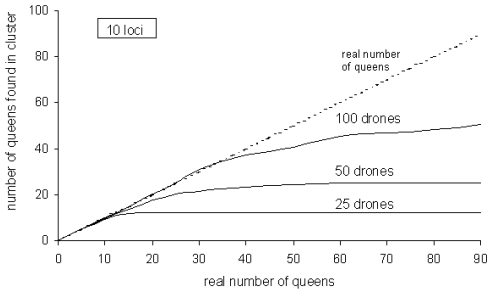

Figure 2. Effect of the number of males sampled on the preciseness of the analysis. The number of queens found in the UPGMA cluster are plotted versus the number of queens used in the Monte Carlo simulation with 10 loci. The results are show for three different sample sizes (25, 50 and 100 drones).

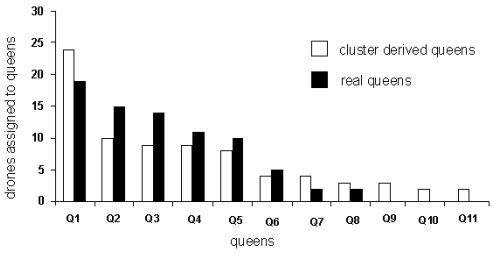

Figure 3. Distribution and number of drones assigned to queens in the cluster (white bars) and to real queens (black bars) present on the island. The two distributions do not differ significantly from each other ($P > 0.3$).

was used as basis for UPGMA analysis. We determined the putative mother queens for each cluster and obtained a total of 11 putative queens in the given sample of 78 drone genotypes. Since indeed 8 drone producing colonies were present on the island the estimates based on the UPGMA seem reasonably accurate, and as predicted by our simulations, give an overestimation of the number of colonies in a small population. The number and distribution of drones assigned to the 11 queens in the cluster (Fig. 3) did not differ significantly from the distribution to their eight real mothers ($\chi^2 = 8.23$, d.f. = 7, $P > 0.3$).

3.4. Empirical example: *A. dorsata* in Borneo

In spite of using only three loci, the genotypes of the sampled *A. dorsata* drones were extremely variable, with 15.5 ± 4.2 alleles per locus on average. Again the distance matrix was based on the drone genotypes obtained by applying equations (2) and (3) was used as basis for the UPGMA analysis. We determined the

putative mother queens for each cluster and obtained a minimum of 53 colonies in the given sample of 148 drone genotypes.

4. DISCUSSION

The present method is a simple approach to evaluate the number of male producing colonies in a population of social hymenoptera. It requires neither elaborate computer simulations nor specific population genetic software. It can be performed on the basis of user friendly, menu driven, standard spread sheet and statistics software. The goal of the procedure is not to provide precise estimates of relatedness between two haploid males (although this is also possible with this technique as with other methods). Rather, we present a colony estimate, which is therefore less prone to individual pair-wise misclassifications. So even if the number of mismatches between true and inferred parental colony is high, the overall estimate of contributing colonies is less affected. Moreover, we only search for full sib males which can be distinguished with a high probability from unrelated individuals (Blouin et al., 1996).

Although the discriminatory power of the number of loci used (3 and 6) was sufficient in this study, this may not be the case in other cases where loci are less polymorphic. Since more than 500 microsatellite loci have been developed for the honeybee (Solignac et al., 2003) a lack of diagnostic variability should not present a problem for population studies in honeybees. If drones are sampled directly on drone congregations areas (Baudry et al., 1998) it is possible to pre-screen these in pooled DNA samples for polymorphic loci (Moritz et al., 2003). In other species, however, the number of available loci is much smaller and may result in less precise estimates. Nevertheless, the simulations reveal that informative estimates are possible even with only a few loci genotyped. The method will be particularly powerful if the sample size can be adjusted to the estimated number of colonies.

Sampling errors occur at two levels, the finite sample of individuals and the finite sample of screened loci. For large populations, the limited number of sampled individuals may form the main source of error. In most cases it therefore seems more important to maximize the number of individuals rather than the number of polymorphic loci. As rule of the thumb it seems useful to increase the sample sizes until the number of estimated colonies is less than half of the sample size.

The empirical data from *A. mellifera* support the results of our simulations. The estimated number of colonies matched the actual number of colonies present in the population. By using six loci with average variability, the actual population size was only moderately overestimated. In our second empirical example, *A. dorsata* from Borneo, we only used three, but highly polymorphic loci. Here we obtained a population estimate of 53 colonies contributing to the sampled drones. Judging from our simulations, this may be a robust estimate, since it is considerably below half the total sample size (148 drones). Additionally it is possible to apply sample coverage methods to correct for the non sampling error, since the present method estimates the number of queens represented in the given sample and in some cases might therefore underestimate the number of colonies (also see Fig. 2). One approach to correct for the non sampling error was given by Cornuet and Aries (1980), but also other sample coverage methods can be used. In case of our first empirical example (*A. mellifera*) application of sample coverage methods does not lead to an increase in the number of colonies detected. This indicates that probably no colonies remained unsampled, which is indeed true since only eight colonies were present on the island. The application of sample coverage methods in our second example (*A. dorsata*) leads to an increase from 53 colonies found in the cluster to 67 colonies, indicating that some colonies remained unsampled.

The method presented here gives an estimate of the number of colonies contributing males to a given drone congregation area. Since honeybees colonies which fail to produce drones are also very likely not to reproduce via swarming the estimate can be considered as the reproductive populations size. The method can also be applied to a wider range of monogynous eusocial hymenoptera. In case of polygynous species our estimate would reflect the number of reproductively active gynes present in a population rather than the actual number of colonies. Nevertheless estimates for colony number

might still be possible in these cases, if reliable data for the average number of queens is available for the species in focus. The crucial prerequisite for our approach is that a sufficient number of males can be sampled. This is possible in many species, since the production of males is often synchronised in social insects and males are often produced in large numbers (Wilson, 1971). In the special case of species which have highly polyandrous queens, the sampling of workers from colonies and the deriving of patrilines (the males the queen has mated with) can be used to estimate the number of drone producing colonies.

ACKNOWLEDGMENTS

We would like to thank J. van Praagh for the *A. mellifera* samples from the Island of Neuwerk. Financial support was granted to RFAM by the BABE RTD network and the ALARM integrated project (EU commission DG 12).

Résumé – Utilisation des mâles pour estimer le nombre de colonies en analysant l'ADN des microsatellites des mâles haploïdes chez *Apis*. Chez les populations d'insectes sociaux, c'est principalement le nombre de colonies plutôt que le nombre réel d'individus qui détermine la taille de la population. C'est particulièrement vrai pour les espèces dont la reproduction a lieu au niveau de la colonie par fission de la colonie. Nous présentons une méthode où les données des microsatellites des mâles haploïdes peuvent être utilisées pour estimer le nombre de mâles produisant des reines présents dans une population. Une analyse de groupement basée sur l'identité génétique par descendance (AID) parmi les génotypes mâles est utilisée pour regrouper les mâles potentiellement frères. Pour chaque « groupe de frères », le génotype de la reine mère correspondante a été déterminé selon les règles mendéliennes. Dans diverses simulations nous avons montré que, bien qu'un nombre restreint de locus étudiés puisse provoquer des estimations légèrement biaisées, la précision de la méthode augmente considérablement parallèlement au nombre de locus (Fig. 1). Le nombre d'individus échantillonnés s'est révélé plus important que le nombre de locus testés par individu. Si l'échantillon est de petite taille, les analyses très précises portant sur un grand nombre de locus n'améliorent par les analyses. Les analyses sont imparfaites principalement par l'erreur de non-échantillonnage, mais non par l'absence de précision des analyses génétiques (Fig. 2).
Afin d'illustrer la méthode proposée, nous avons analysé par la méthode de l'AID les données des microsatellites d'une population d'abeilles domestiques (*Apis mellifera* L.) sur l'île de Neuwerk (Allemagne) et d'une population d'abeilles géantes asiatiques (*A. dorsata* Fabricius) à Bornéo. Dans le cas de l'île de Neuwerk, le nombre « réel » de colonies (8 colonies produisant des mâles) était connu et a pu être comparé au nombre de colonies estimé. Le nombre estimé par la méthode de l'AID a été de 11 colonies. La figure 3 donne la répartition des mâles entre les reines « réelles » et les reines estimées sur l'île.
Pour la population d'*A. dorsata*, nous avons obtenu une estimation de 53 colonies selon le groupement de l'AID sur un total de 148 mâles, dont le génotype avait été caractérisé. A en juger par nos simulations, ce peut être une estimation robuste puisqu'elle est nettement en-dessous de la moitié de la taille de l'échantillon (148 mâles).
La méthode proposée ici pour estimer la taille des populations est également utile pour une large gamme d'espèces d'Hyménoptères dont les colonies sont dirigées par une seule reine. Dans le cas d'espèces polygynes, notre estimation refléterait le nombre de reines présentes dans une population plutôt que le nombre réel de colonies. Il doit être néanmoins possible dans ces cas d'estimer le nombre de colonies, s'il existe des données fiables sur le nombre moyen de reines pour l'espèce considérée. Notre approche nécessite une autre condition préalable cruciale, à savoir qu'un nombre suffisant de mâles puisse être échantillonnés. Cela est possible chez de nombreuses espèces, puisque la production de mâles est souvent synchronisée chez les insectes sociaux et qu'ils sont souvent produits en grandes quantités.

***Apis mellifera* / *Apis dorsata* / mâle haploïde / microsatellite / taille de la population**

Zusammenfassung – Die Verwendung von Drohnen zur Abschätzung der Kolonieanzahl durch die Mikrosatelliten DNA Analyse bei haploiden Männchen von *Apis*. Bei sozialen Insekten bestimmt nicht die Anzahl der Individuen sondern die Anzahl der Kolonien die Populationsgröße. Dies trifft besonders auf Arten zu, die sich über Koloniteilung vermehren (wie z. B. die Honigbiene). Hier stellen wir eine Methode zur Abschätzung der Kolonieanzahl vor, die auf genotypischen Mikrosatelliten Daten von haploiden Männchen beruht. Bei dieser Analyse wird eine Clustermethode genutzt, die auf der genetischen Identität durch Abstammung (AID) beruht, um die Genotypen potentieller Brüder Männchen zu gruppieren. Für jedes potentielle „Brudercluster" wird die Mutter mittels der Mendelschen Regeln hergeleitet. In Simulationen konnte gezeigt werden, dass die Präzision der Methode mit zunehmender Anzahl benutzter Loci schnell ansteigt (Abb. 1). Die AID-Analyse wird weitaus stärker durch den Effekt des Stichprobenfehlers als durch die Ungenauigkeit der genetischen Analyse verzerrt (Abb. 2).

Zur Illustration der vorgestellten Methode wurden Mikrosatelliten Daten je einer Population der Westlichen Honigbiene, *A. mellifera*, von der Insel Neuwerk und einer Population der Asiatischen Riesenhonigbiene, *A. dorsata*, mittels der AID Methode analysiert. Im Falle der *A. mellifera* Population auf Neuwerk war die „wirkliche" Anzahl der Kolonien auf der Insel bekannt (insgesamt 8 Völker mit Drohnen) und konnte somit mit der geschätzten Anzahl verglichen werden. Mittels der AID Methode wurden für Neuwerk 11 Kolonien geschätzt, was nur geringfügig von den „wirklichen" vorhandenen 8 Kolonien abweicht. Abbildung 3 zeigt die Verteilung der einzelnen Drohnen auf die „wirklichen" und die geschätzten Kolonien auf Neuwerk.

Für die *A. dorsata* Population wurde die Anzahl der Kolonien auf 53 geschätzt, bei einer Stichprobengröße von 148 genotypisierten Drohnen. Ausgehend von unseren Simulationen sollte dies eine robuste Schätzung sein, da dieser Wert sich noch unterhalb der Hälfte der Stichprobe (78) befindet.

Die hier vorgestellte Methode für die Abschätzung der Kolonieanzahl kann prinzipiell auch auf andere Arten sozialer Hymenopteren angewandt werden, deren Kolonien nur eine einzige Königin besitzen. Im Falle von Arten mit mehreren Königinnen in einer Kolonie würde die Anzahl der Königinnen und nicht die Anzahl der Kolonien erfasst werden. Allerdings könnte eine Abschätzung auch in einem solchen Falle durchgeführt werden, wenn eine durchschnittliche Königinnen-Anzahl pro Kolonie bekannt ist. Eine weitere essentielle Voraussetzung für diese Methode ist natürlich, dass ausreichende Männchen vorhanden sind. Dies sollte allerdings bei vielen Arten möglich sein, weil bei sozialen Hymenopteren Männchen meist in großer Anzahl vorhanden sind und das Schwärmen der Geschlechtstiere oft synchronisiert ist.

***Apis mellifera* / *Apis dorsata* / haploide Männchen / Honigbienen / Mikrosatelliten / Populationsgröße**

REFERENCES

Baudry E., Solignac M., Garnery L., Gries M., Cornuet J.M., Koeniger N. (1998) Relatedness among honeybees (*Apis mellifera*) of a drone congregation, Proc. R. Soc. Lond. B 26, 2009–2014.

Blouin M.S., Parsons M., Lacaille V., Lotz S. (1996) Use of microsatellite loci to classify individuals by relatedness, Mol. Ecol. 5, 393–401.

Cornuet J.M., Aries F. (1980) Number of sex alleles in a sample of honeybee colonies, Apidologie 11, 87–93.

Estoup A., Solignac M., Cornuet J.M. (1994) Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies, Proc. R. Soc. Lond. B 258, 1–7.

Kerr W. (1967) Multiple alleles and genetic load in bees, J. Apic. Res. 6, 61–64.

Li C.C., Weeks D.E., Chkravarti A. (1993) Similarity of DNA fingerprints due to chance and relatedness, Hum. Hered. 43, 45–52.

Lynch M., Ritland K. (1999) Estimation of pairwise relatedness with molecular markers, Genetics 152, 1753–1766.

Moritz R.F.A., Kryger P., Koeniger G., Koeniger N., Estoup A., Tingek S. (1995) High degree of polyandry in *Apis dorsata* queens detected by DNA microsatellite variability, Behav. Ecol. Sociobiol. 37, 357–363.

Moritz R.F.A., Kryger P., Allsopp M. (1996) Competition for royalty in bees, Nature 384, 522.

Moritz R.F.A., Scharpenberg H., Lattorff H.M.G., Neumann P. (2003) A technical note for using microsatellite DNA analyses in haploid male DNA pools of social Hymenoptera, Insectes Soc. 50.

Oldroyd B.P., Clifton M.J., Wongsiri S., Rinderer T.E., Sylvester H.A., Crozier R.H. (1997) Polyandry in the genus *Apis*, particularly *Apis andreniformis*, Behav. Ecol. Sociobiol. 40, 17–26.

Pamilo P., Crozier R.H. (1997) Population Biology of Social Insect Conservation, Mem. Mus. Vict. 56, 411–419.

Pamilo P., Gertsch P., Thoren P., Seppa P. (1997) Molecular population genetics of social insects, Annu. Rev. Ecol. Syst. 28, 1–25.

Solignac M., Vautrin D., Loiseau A., Mougel F., Baudry E., Estoup A., Garnery L., Haberl M., Cornuet J.M. (2003) Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera*), Mol. Ecol. Notes 3, 307–311.

Strassmann J. (2001) The rarity of multiple mating by females in the social Hymenoptera, Insectes Soc. 48, 1–13.

Walsh P.S., Metzger D.A., Higuchi R. (1991) Chelex 100[a] as a medium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques 10, 506–513.

Wang J. (2002) An estimator for pair wise relatedness using molecular markers, Genetics 160, 1203–1215.

Wattanachaiyingcharoen W., Oldroyd B.P., Wongsiri S., Palmer K., Paar R. (2003) A scientific note on the mating frequency of *Apis dorsata*, Apidologie 34, 85–86.

Wilson E.O. (1971) The Insect Societies, Harvard University Press, Cambridge, Massachusetts.

Wright S. (1933) Inbreeding and homozygosis, Proc. Natl. Acad. Sci. USA 19, 411–420.