Persistence of the single lineage of transmissible ‘social cancer’ in an asexual ant

S. DOBATA,*† T. SASAKI,†‡ H. MORI,‡** E. HASEGAWA,§ M. SHIMADA* and K. TSUJI†

*Department of General Systems Studies, Graduate School of Arts and Sciences, University of Tokyo, Meguro, Tokyo 153-8902, Japan,
†Department of Agro-Environmental Sciences, Faculty of Agriculture, University of the Ryukyus, Nishihara, Okinawa 903-0213, Japan,
‡Department of Ecology and Evolutionary Biology, Graduate School of Life Sciences, University of Tohoku, Aoba, Sendai 980-8578, Japan,
§Laboratory of Animal Ecology, Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Kita-ku, Sapporo 060-8589, Japan

Abstract

How cooperation can arise and persist, given the threat of cheating phenotypes, is a central problem in evolutionary biology, but the actual significance of cheating in natural populations is still poorly understood. Theories of social evolution predict that cheater lineages are evolutionarily short-lived. However, an exception comes from obligate socially parasitic species, some of which thought to have arisen as cheaters within cooperator colonies and then diverged through sympatric speciation. This process requires the cheater lineage to persist by avoiding rapid extinction that would result from the fact that the cheaters inflict fitness cost on their host. We examined whether this prerequisite is fulfilled, by estimating the persistence time of cheaters in a field population of the parthenogenetic ant Pristomyrmex punctatus. Population genetic analysis found that the cheaters belong to one monophyletic lineage which we infer has persisted for 200–9200 generations. We show that the cheaters migrate and are thus horizontally transmitted between colonies, a trait allowing the lineage to avoid rapid extinction with its host colony. Although horizontal transmission of disruptive cheaters has the potential to induce extinction of the entire population, such collapse is likely averted when there is spatially restricted migration in a structured population, a scenario that matches the observed isolation by distance pattern that we found. We compare our result with other examples of disruptive and horizontally transmissible cheater lineages in nature.

Keywords: cheating, clonal genetic structure, Emery’s rule, evolution of cooperation, genetic caste determination, transmissible cancer

Received 15 April 2010; revision received 17 October 2010; accepted 19 October 2010

Introduction

Along with competition and exploitation, cooperation is a nearly ubiquitous feature of biological systems and has been considered to be essential in the emergence and evolution of biological complexity: from self-replicating molecules to multicellular organisms to complex societies (i.e., the major evolutionary transitions; Maynard Smith & Szathmáry 1995). A serious problem for the maintenance of cooperation is that it is highly vulnerable to any kind of cheating that originates from and imposes a fitness cost on cooperators. In natural populations, however, cooperation is widespread and intraspecific cheating normally appears to occur at low frequencies (e.g., Gilbert et al. 2007 for micro-organisms; Barron et al. 2001, Hughes & Boomsma 2008 for...
social insects). Theories of social evolution, driven mostly by inclusive fitness theory (Hamilton 1964), have achieved great success in explaining patterns of prevalence of cooperation against the threat of selfish behaviour such as worker reproduction in hymenopteran colonies (Ratnieks 1988; Ratnieks & Wenseleers 2008) and cheating in social micro-organisms (Gilbert et al. 2007). However, inclusive fitness theory predicts only whether a certain strategy is evolutionarily stable, increases in frequency in the next generation, or becomes fixed in the population with a probability higher than expected from neutrality (Hamilton 1964; Frank 1998; Roussel 2004). The underlying dynamic processes of cheater progression or elimination are often hard to capture by inclusive fitness theory itself. Among them is the persistence time of the obligate cheater lineage [see Fiegna & Velicer (2003) for a relevant laboratory experiment]. Because of their parasitic nature, the persistence time of intraspecific obligate cheater lineages is usually much shorter than that of ancestral cooperator lineages, which supports social evolution theory. For example, in queenright honeybee colonies, mutant cheater workers, which reproduce selfishly by evading policing, are evolutionarily ephemeral (Oldroyd 1994; Barron et al. 2001; Châline et al. 2002). Cancer in multicellular organisms is another example of cheating in cell subpopulations (Burt & Trivers 2006), but the cancer cells usually share the fate of their host organism when it dies. In contrast, it has been speculated that obligate social parasites of ants often originate from intraspecific cheaters (e.g. additional queens etc.) and that they speciate syntropically after which many lose the worker caste. This hypothesis is based on the phylogenetic proximity of several host and social parasite pairs of ants (Savolainen & Vepsäläinen 2003; Sumner et al. 2004; Vepsäläinen et al. 2009; see also Steiner et al. 2006), which is known as Emery’s rule sensu stricto (Emery 1909; Hölldobler & Wilson 1990). The underlying evolutionary process requires both that reproductive isolation is achieved sympatrically and that the cheater lineage persists for a sufficiently long time to adapt to its host (e.g., miniaturization of body size; Brandt et al. 2005). Previous studies mainly focused on the difficulty of sympatric speciation of social parasites (Hölldobler & Wilson 1990; Vepsäläinen et al. 2009 and references therein), but little is known about the conditions under which obligate cheater lineages can persist despite the damage inflicted on its host, once they speciated sympatrically. In this study, we examined the population genetics of a cheater lineage of the therytokously parthenogenetic ant _Pristomyrmex punctatus_ (formerly _P. pungens_), to reconstruct its evolutionary origin and persistence dynamics.

**The extraordinary social structure of _P. punctatus_**

Colonies of social insects normally provide typical examples of biological cooperation, which is characterized by a well-developed reproductive division of labour between queens and workers. However, this is not the case in _P. punctatus_. In this parthenogenetic ant, sexual reproduction and reproductive division of labour is secondarily lost: in the breeding period, all the workers are involved in reproduction first and shift to helping tasks as they age (Tsuji 1990). Previous studies have reported that unusually large individuals are sometimes found within colonies. The terminology for these large individuals and the normal workers has been rather confusing (‘large workers or ergatoid queens’ vs. ‘small workers’ in Itow et al. (1984), ‘large workers’ vs. ‘small workers’ in Tsuji (1995) and Sasaki & Tsuji (2003), ‘ergatoid queens’ vs. ‘workers’ in Peeters (1991) and Wang (2003), and ‘L-types’ vs. ‘S-types’ in Dobata et al. (2009a), respectively), partly because the extraordinary social structure of this species (Tsuji 1990) made it difficult to assess the function of these individuals. These large individuals show several characteristics typical for ergatoid (wingless and worker-like) queens in other ants, such as having distinctively larger number of ovarioles (four vs. two). According to the morphology-based definition by Molet et al. (2009) and Peeters (in press), we will hereafter use the term ‘ergatoid queens’ for the large individuals with four ovarioles and ‘workers’ for the normal (small) workers with two ovarioles, but we stress that the terminology of normal eusocial insects remains difficult to apply to _P. punctatus_, and that the above morphological definitions do not reflect their usual functions.

Unlike normal sexual ants, parthenogenetic _P. punctatus_ needs only workers for a colony to flourish (Tsuji 1990, 1994), and ergatoid queens rarely take part in cooperative tasks (Sasaki & Tsuji 2003). Instead, Tsuji (1995; fig. 4) showed that the average reproductive success of colonies containing ergatoid queens was only ca. 80% of what colonies containing only workers achieved. Recently, Dobata et al. (2009a) analysed the genetic background of the individual phenotypes over several generations using genetic markers and found that at least one genotype produced exclusively ergatoid queens. Individuals with this genotype appeared to meet the conditions for being specialized cheaters, and we will hereafter use the term ‘Cheaters’ for genetically determined ergatoid queens in _P. punctatus_. Assuming that the negative effect of ergatoid queens does not change with their frequencies and colony size, the Cheaters are expected to cause the collapse of their colony within at most 20 years when cheating would have reduced colony size to 1% of the original (0.8^{20} \approx 0.01).
In this study, we use newly developed microsatellite markers (Dobata et al. 2009b) and several population genetic tools to elucidate the detailed population genetic structure of our Pristomyrmex study population. We confirm that the dominant mode of reproduction was thelytokous parthenogenesis and that the Cheaters have a single origin in this population and form a genetically distinct lineage. Moreover, we estimate that the persistence time of this Cheater lineage should range from ca. 200–9200 generations, which is much longer than the initial expectation based on Cheaters causing the collapse of their own colony within about 20 years. We argue that two mechanisms, which were inferred from the population genetic data, enable the Cheater lineage to achieve these long persistence times: (i) horizontal transmission of Cheater individuals between colonies, so that Cheaters escape from host colonies before they collapse, and (ii) spatially restricted migration of the wingless Cheaters in a structured population, which prevents the extinction of the entire population. Finally, we compared our result with other examples of disruptive and horizontally transmissible cheater lineages in nature.

Materials and methods

Sampling

We collected 54 colonies of *P. punctatus* from an approximately 4-km$^2$ area in Kihoku, Mie Prefecture, central Japan, which was investigated in our previous study (Dobata et al. 2009a), between 11 and 13 July 2005 (Fig. 1). At this time of the year, adults belonging to parent and offspring generations coexist in the same colonies (offspring will overwinter and reproduce the following year). *P. punctatus* colonies are mostly monodomous (single-nested) and mutually hostile so that their populations are multicolonial (Tsuji & Ito 1986). Colonies were collected with their nest materials and were sorted later the same day. This procedure enabled us to collect almost all colony members, store several hundred individuals in pure ethanol and estimate colony size (the number of individuals belonging to the parental generation). Whether individuals belonged to the parent or offspring generation was determined individually, and colony sizes and individual phenotype frequencies within colonies were extrapolated from the stored subsamples of the colonies, according to the method of Tsuji (1995). Each of the collected females was dissected and classified as being an ergatoid queen or a worker, based on the number of ovarioles (see Introduction).

Genotyping

From the preserved samples, 516 females from 33 colonies (10 workers and 0–10 ergatoid queens of the parent generation per colony; five workers and 1–6 ergatoid queens of the offspring generation when present) were used for the population genetic analyses, including the ones genotyped only at four loci in our previous study (Dobata et al. 2009a). For the individuals newly added in this study, total DNA was extracted from the thorax using the Chelex method (Walsh et al. 1991) with minor modifications. Thirteen microsatellites, Pp1, Pp2 (Hasegawa et al. 2001), MYRT3 (Evans 1993), ppmb13, ppmb27, ppmb33, ppmb104, ppmb105, ppmb106, ppmb114, ppmb132, ppmb139, and ppmb204 (Dobata et al. 2009b),
and one mitochondrial marker, a downstream intron region of the 12S ribosomal RNA gene (Dobata et al. 2009a; Hasegawa et al. in press), were used in this study. Polymerase chain reactions (PCRs) were conducted following the previously described protocols (Dobata et al. 2009a,b). The amplified fluorescent PCR products were analysed using an automated sequencer (CEQ 8000; Beckman & Coulter, Fullerton, CA, USA).

Genetic analysis and analysis of clonality

To characterize the 13 microsatellites, we used GENEPOP version 4.0 (Raymond & Rousset 1995) to assess the number of alleles, the observed and expected heterozygosities, deviations from HWE at each locus and linkage disequilibria (LD) between pairs of loci. The multilocus genotype (MLG) of each individual was determined using both the genotypes of the 13 microsatellites and the mtDNA haplotype. We checked that those MLGs we observed more than once resulted from clonal redundancy rather than sexual reproduction by estimating $P_{gen}(f)$, the probability of encountering a given MLG in an individual when taking into account departure from HWE, and $P_{sex}(f)$, the probability of encountering this MLG more than once in the population as the result of distinct sexual reproductive events (Parks & Werth 1993). The calculations were conducted using GENCLONE version 2.0 (Arnaud-Haond & Belkhir 2007). Owing to software limitations, we could only use the parental generation ($n = 493$) in these $P_{gen}(f)$ and $P_{sex}(f)$ calculations. We then estimated the clonal richness index $R = (\text{the number of MLGs} - 1)/\text{(the number of analysed individuals} - 1)$ (Dorken & Eckert 2001) using all analysed individuals ($n = 516$).

Phylogenetic relationship among MLGs

To assess the phylogenetic relationship between the MLGs as implemented in the software POPULATIONS version 1.2.30 (Langella 1999), a neighbour-joining (NJ) dendrogram was constructed using chord distances (Cavalli-Sforza & Edwards 1967). For calculation of genetic distances, only microsatellite genotypes were used. An additional worker from Nakijin, Okinawa Island, southern Japan, was genotyped at all the microsatellites used in this study to determine the root of the dendrogram.

The phylogenetic relationships among MLGs was more thoroughly analyzed through Bayesian clustering of the MLG data as implemented in the software STRUCTURE version 2.2.1 (Pritchard et al. 2000; Falush et al. 2003), which assigns individuals probabilistically to a given number ($K$) of clusters. We used only the microsatellite genotype data for this analysis. The algorithm assumes that the population within each of $K$ clusters is at Hardy–Weinberg equilibrium (HWE) and at linkage equilibrium. Because $P. punctatus$ is parthenogenetic and the HWE assumption is inappropriate, the data were treated as haploid to relax the assumption of the model regarding the statistical independence of the two alleles within an individual at each locus, as recommended by Falush et al. (2003). Iteration parameters were set to a burn-in of 50 000 and Markov Carlo run lengths of 100 000 under an admixture ancestry model assuming correlated allele frequencies without prior population information. Ten independent simulations were run for $K = 1$ to $K = 15$, and the final posterior probability of $K$ was determined using the run with highest probability for each $K$.

F-statistics

We calculated $F$-statistics as defined by Weir & Cockerham (1984) using the program FSTAT version 2.9.3.2 (Goudet 1995). The 95% confidence intervals were obtained through bootstrapping over loci as implemented in the software. $F$-statistics were also calculated at each of the 13 loci. Given their monophyly (see Results), Cheaters were treated as forming distinct subpopulations even within the same colonies. In addition to the global estimate of $F$-statistics, we obtained pairwise $F_{ST}$ values between the subpopulations, which were used for analysing the mode of Cheater transmission and isolation by distance (see Results).

Results

The number of alleles and the observed and expected heterozygosities for each locus are given in Table 1. When using all individuals ($n = 516$) for the calculations, departures from HWE were highly significant after Bonferroni corrections at all but one (ppmb105) locus (Table 1). Significant LD was detected in 75 of 78 pairs of loci, which indicates clonality (De Meed & Baloux 2004). We also estimated these values using only one individual per MLG ($n = 86$), resulting in significant HWE departures remaining at nine loci and LD between 58 pairs. Using 13 nuclear microsatellite markers and one mtDNA marker, we recognized a total of 86 MLGs in the study population. The clonality indices $P_{gen}(f)$ and $P_{sex}(f)$ were very small ($\leq 0.000016$ and $\leq 0.008$, respectively), and the clonal richness index $R$ was 0.165, both strongly supporting asexuality.

Identification of the Cheater lineage

Figure 2 shows the NJ dendrogram describing the phylogenetic relationships between the 86 MLGs. The frequency of the ergatoid queens and workers revealed
that some MLGs were shared by ergatoid queens and workers (Fig. 2). As we have discussed previously (Dobata et al. 2009a), these are most likely caused by ancestral developmental plasticity for caste determination [see also Discussion]. Moreover, within the ergatoid queens, the number of ocelli, which is often correlated with reproductive capacity, varied from zero to three [see also Sasaki & Tsuji (2003) and Dobata et al. (2009a)]. Notably, one clade containing 10 MLGs (n = 86 individuals, shaded darker in Fig. 2, bootstrap support of 50%) consisted exclusively of ergatoid queens. Bayesian clustering analysis of the MLG data as

Table 1 Characteristics of 13 microsatellite loci in the study population of Pristomyrmex punctatus

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allele size range (bp)</th>
<th>NA</th>
<th>H_E (N = 516)</th>
<th>H_O (N = 516)</th>
<th>H_E (N = 86)</th>
<th>H_O (N = 86)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pp1</td>
<td>220–225</td>
<td>2</td>
<td>0.475</td>
<td>0.775*</td>
<td>0.429</td>
<td>0.616*</td>
</tr>
<tr>
<td>Pp2</td>
<td>224–238</td>
<td>3</td>
<td>0.513</td>
<td>0.787*</td>
<td>0.530</td>
<td>0.802*</td>
</tr>
<tr>
<td>MYRT3</td>
<td>177–185</td>
<td>4</td>
<td>0.688</td>
<td>0.671*</td>
<td>0.724</td>
<td>0.570*</td>
</tr>
<tr>
<td>ppm13</td>
<td>272–308</td>
<td>5</td>
<td>0.694</td>
<td>0.607*</td>
<td>0.730</td>
<td>0.733*</td>
</tr>
<tr>
<td>ppm27</td>
<td>157–160</td>
<td>2</td>
<td>0.451</td>
<td>0.682*</td>
<td>0.420</td>
<td>0.570*</td>
</tr>
<tr>
<td>ppm33</td>
<td>298–324</td>
<td>4</td>
<td>0.626</td>
<td>0.994*</td>
<td>0.632</td>
<td>0.965*</td>
</tr>
<tr>
<td>ppm104</td>
<td>160–162</td>
<td>2</td>
<td>0.466</td>
<td>0.686*</td>
<td>0.473</td>
<td>0.616</td>
</tr>
<tr>
<td>ppm105</td>
<td>190–192</td>
<td>2</td>
<td>0.188</td>
<td>0.198</td>
<td>0.305</td>
<td>0.326</td>
</tr>
<tr>
<td>ppm106</td>
<td>312–342</td>
<td>11</td>
<td>0.639</td>
<td>0.934*</td>
<td>0.670</td>
<td>0.837*</td>
</tr>
<tr>
<td>ppm114</td>
<td>308–312</td>
<td>3</td>
<td>0.367</td>
<td>0.436*</td>
<td>0.287</td>
<td>0.302</td>
</tr>
<tr>
<td>ppm132</td>
<td>268–278</td>
<td>5</td>
<td>0.390</td>
<td>0.477*</td>
<td>0.323</td>
<td>0.372</td>
</tr>
<tr>
<td>ppm139</td>
<td>142–147</td>
<td>4</td>
<td>0.528</td>
<td>0.324*</td>
<td>0.499</td>
<td>0.221*</td>
</tr>
<tr>
<td>ppm204</td>
<td>191–231</td>
<td>10</td>
<td>0.757</td>
<td>0.996*</td>
<td>0.778</td>
<td>0.977*</td>
</tr>
</tbody>
</table>

NA, number of alleles; Nind, number of genotyped individuals; H_E, expected heterozygosity; H_O, observed heterozygosity. *Significant deviation from Hardy–Weinberg equilibrium after Bonferroni corrections. The calculations were performed both using all sampled individuals (N = 516) and using only one individual per multilocus genotype (MLG) (N = 86).

Fig. 2 The rooted neighbour-joining dendrogram for the 86 multilocus genotypes (MLGs) sampled in the study population of Pristomyrmex punctatus. Numbers beside the branches give bootstrap values ≥50%. Two mtDNA haplotypes are indicated by filled (product length: 338 bp, GenBank accession no. AB556947) and open (342 bp, GenBank accession no. AB556946) squares, respectively, on the tip of the branch. Each MLG harboured only one mtDNA haplotype. The phenotype frequencies of each MLG are shown by stacked circles, the size of which reflects sample size (illustrated in box). The distinctiveness of the two mtDNA haplotypes was confirmed previously by sequencing partial fragments of the mitochondrial cytochrome oxidase I (COI) gene (Dobata et al. 2009a). Because the mtDNA haplotypes were not segregating into single clades of the dendrogram, the occurrence of some sexual reproduction and introgression of mtDNA are suspected (see Discussion).
implemented in Structure version 2.2.1 showed that the log likelihood ($\ln[P(D|K)]$) saturated at $K \geq 7$ (Fig. 3a). An additional ad hoc statistic, $\Delta K$, which provides a better predictor of the number ($K$) of clusters at the uppermost hierarchical level (Evanno et al. 2005) showed peaks at both $K = 2$ and $K = 4$ (Fig. 3b), but these $K$ clusters corresponded neither to the mitochondrial haplotypes nor the Cheaters (see Discussion). Based on this observation, we compared the clustering results from $K = 2$ to $K = 7$ (Fig. 4). All individuals belonging to the darker-shaded clade in Fig. 2 were assigned to the same cluster at all values of $K$, confirming the monophyly of these 10 MLGs (Fig. 4). Because this clade contains the Cheater genotype identified in the previous study (Dobata et al. 2009a), we concluded that this single clade corresponds to the Cheater lineage and that all extant Cheaters in the study population have a single ancestor. Hereafter, we call individuals of the other lineages ‘Noncheaters’ (note that the Noncheaters contain ergatoid queens and workers).

Almost all (85 of 86, 98.8%) of the Cheaters had three conspicuous ocelli (Fig. 2). In stark contrast, this proportion was much smaller in the ergatoid queens of the Noncheaters (11 out of 90, 12.2%). The difference was statistically highly significant (Fisher’s exact test, $\chi^2 = 133.1$, d.f. = 1, $P < 10^{-35}$).

**Mode of cheater transmission**

Next we assessed the mode of transmission of the Cheaters. We previously speculated that the Cheaters can

---

**Fig. 3** Results of multilocus genotype clustering by Structure. (a) The estimated logarithmic probability of data ($\ln[P(D|K)]$) being explained by subdivision in $K$ groups and (b) the $\Delta K$ statistic (Evanno et al. 2005) as a function of the number of clusters tested.

<table>
<thead>
<tr>
<th>mtDNA</th>
<th>Cheater</th>
<th>Noncheater</th>
</tr>
</thead>
<tbody>
<tr>
<td>338</td>
<td>342</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4** Population structure assessed by multilocus genotype clustering by Structure. Each of the individuals is represented by a thin vertical bar. Some bars are partitioned by $K$ different colors, each representing an individual’s proportionate assignment to each of given $K$ clusters, whereas other individuals have genotypes that are mostly or fully compatible with a single cluster.

© 2010 Blackwell Publishing Ltd
migrate and are horizontally transmitted among colonies as a 'transmissible social cancer', based on the fact that Cheaters with identical MLG were found in several colonies (Dobata et al. 2009a). Because of the limited availability of genetic markers, however, we could previously not rule out that the Cheaters mainly 'hitchhike' with colony fission, which would imply vertically transmission (new colonies of Noncheaters are always generated by fission; Itow et al. 1984; Tsuji 1988). Given our present evidence for monophyly, the Cheaters could now be treated as forming a distinct subpopulation from the Noncheaters even within the same colony. The pairwise genetic differentiations of Cheater and Noncheater subpopulations, measured by $F_{ST}$ values, were assessed to check whether the inferred horizontal transmission of the Cheaters was correct. If the Cheaters are vertically transmitted, pairwise genetic differentiations between Cheater subpopulations were expected to be positively correlated with those between Noncheater subpopulations, whereas no correlations were expected in the case of horizontal transmission. The partial correlation of the pairwise $F_{ST}$ values between Cheater and Noncheater subpopulations was $r = 0.152$, which was not statistically significant ($n = 45$, controlling for their geographical distances, partial Mantel test with 10 000 permutations, $P = 0.242$). We therefore conclude that horizontal transmission of Cheaters occurs frequently in addition to the vertical transmission that would automatically happen when colonies bud.

**Population genetic structure**

Table 2 shows the global estimates of $F$-statistics, treating the Cheater and Noncheater subpopulations separately. The considerably negative values of estimated $F_R$ indicate clonal reproduction (Balloux et al. 2003). Because our colony-level analysis revealed that individuals of the Cheater lineage did not coexist with ergatoid queens of the Noncheater lineages in almost all (9 of 10, 90%) Cheater-containing colonies, we could easily estimate the sizes of Cheater and Noncheater subpopulations for these colonies, according to the phenotype-based estimation method of Tsuji (1995). In the remaining single colony (K2) in which ergatoid queens of the Cheater and Noncheater lineages coexist, the proportion of ergatoid queens of the Cheater lineage in the subsample (revealed by the genotypic analysis) was extrapolated to infer the size of the Cheater subpopulation (see Materials and methods). The above calculations enabled us to obtain the average number of individuals in each subpopulation ($N$) as harmonic means of sizes of Cheater and Noncheater subpopulations in the colonies. We then compared tentative estimates of migration and mutation rates of Cheaters and

<table>
<thead>
<tr>
<th>Subpopulations</th>
<th>$F_{IS}$ (95% CI)</th>
<th>$F_{ST}$ (95% CI)</th>
<th>$F_{RT}$ (95% CI)</th>
<th>$F_{IT}$ (95% CI)</th>
<th>$F_{RT} - F_{IS}$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noncheater</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(n = 33)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheater</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(n = 10)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The $F$-statistics, group sizes ($n$), migration rates ($m$) and mutation rates ($\mu$) estimated after modifying the standard equations according to specific idiosyncratic details of the biology of *Pristomyrmex punctatus* (see Data S1, Supporting information)
Noncheaters among subpopulations per generation, by applying the classical equation for diploid organisms:

\[(m + \mu) = (1-F_{ST})/(4NF_{ST})\]

where \(m\) is the individual migration rate among subpopulations per generation and \(\mu\) is the mutation rate of a locus (\(\mu\) can be ignored when \(\mu << m\)). Unfortunately, however, this approximation often leads to biased estimations, because real populations frequently violate the underlying assumptions (Whitlock & McCauley 1999). To obtain more accurate estimates, we therefore considered several relevant effects of the idiosyncratic biology of \(P.\) punctatus (see Data S1, Supporting information) to obtain modified equations: 

\[(m + \mu) = (1-F_{ST})/(2NF_{ST})\] for haploids and \[(m + \mu) = -(1 + F_{IS})/(4NF_{IS})\] for clonal diploids (De Meeuws & Balloux 2005). The estimated values of migration and mutation rates are listed in Table 2 and allowed the conclusion that the Cheater individuals do migrate at a higher rate than the Noncheaters.

In addition to the global estimates, \(F\)-statistics were also calculated for each of the 13 loci (Table A1a) and the \((m + \mu)\) values were likewise estimated according to the above three equations (here \(\mu\) is the mutation rate of each locus). In all cases, the estimated \((m + \mu)\) values of the Noncheater subpopulations were so small that the mutation rate \(\mu\) could not be ignored. We took advantage of this property and considered the \((m + \mu)\) values as representing the upper bounds of \(\mu\) (when \(m\) was set to 0). Most of the estimates fell within \(10^{-6}\) to \(10^{-4}\) (Table A1b), which is near the lower end of the range of microsatellite mutation rates of other organisms (Ellegren 2000), and we subsequently used these for the estimation of persistence time of the Cheater lineage.

Isolation by distance

Tests of isolation by distance were conducted using ln-transformed pairwise geographical distances \((m)\) and pairwise \(F_{ST}/(1-F_{ST})\) values, as implemented in the program GENEPOP version 4.0 (Raymond & Rousset 1995), and Mantel tests were separately performed on distance submatrices of the Cheater and Noncheater subpopulations, using the package ade4 (Thioulouse et al. 1997) implemented in \(R\) version 2.9.0 (R Development Core Team 2005). We found statistically significant isolation by distance at a very local geographical scale (Fig. 5; within a radius of ca. 4 km; Cheaters, \(n = 45, r = 0.728\), Mantel test with 10,000 permutations, \(P = 0.016\); Noncheaters, \(n = 528, r = 0.487, P = 0.000099\)).

Estimating the age of the Cheater lineage

Most of the individuals (63 of 86, 73.3%) in the Cheater lineage harboured an identical MLG. Moreover, this MLG was found in all the Cheater-containing colonies. Figure 6 shows the minimum spanning network of the MLGs in the Cheater lineage, based on the chord distance (Cavalli-Sforza & Edwards 1967). The star-like topology indicates that the modal MLG is the ancestral one. Open circles illustrate how circle area reflects sample size.

Fig. 5 Isolation by distance of the study population. The pairwise \(F_{ST}/(1-F_{ST})\) values were regressed on the ln-transformed pairwise geographical distances \((m)\). The Cheater (filled circles) and Noncheater (open circles) subpopulations were regressed separately, but the pairwise \(F_{ST}\) calculations were conducted using the Cheater and Noncheater subpopulations simultaneously.

Fig. 6 Minimum spanning network among the 10 multilocus genotypes (MLGs) of the Cheater lineage. The letters in the circles correspond to the MLGs listed in Table A2. The pairwise genetic distances were based on the chord distance (Cavalli-Sforza & Edwards 1967). The star-like topology indicates that the modal MLG is the ancestral one. Open circles illustrate how circle area reflects sample size.
represents that of the ancestral Cheater, which enabled us to determine the ancestral and mutated alleles of each locus (Table A2). Ignoring back mutation, the time in generations when all the mutated alleles coalesce into the ancestral allele at allelic position \( j \) of locus \( i \), \( t_{ij} \), is

\[
t_{ij} = \log_{(1 - \frac{1}{2})} P_{ij}
\]

where \( \mu_i \) is the mutation rate per locus \( i \) per generation, and \( P_{ij} \) is the observed probability of survival of the ancestral allele, i.e., the probability that an allele at allelic position \( j \) of locus \( i \) matches the common ancestral allele (Hartl & Clark 2007). Because this method considers the extant alleles only, the estimation should be conservative and the actual persistence time might therefore be longer.

We used the values \( \mu_i \) as estimated from the \( F \)-statistics applied to each locus (Table A1b), from which the mutation rate of each allele can be calculated as \( \mu_i / 2 \). For individuals of \( P. punctatus \) having an annual life cycle, the estimated number of generations equals the number of years. The \( t_{ij} \) values were estimated at all the allelic positions (including those with no substitution) and then averaged. By considering a mutation as occurring when a heterozygous locus becomes fixed at one of the alleles, the estimated mean persistence times of the Cheater lineage were between 570 and 9209 years, depending on the estimation methods of mutation rates (Table A2). Alternatively, when we only considered novel alleles as new mutations, we obtained estimates between 236 and 4126 years. The true value should lie in between these estimates.

**Discussion**

The increased number of genetic markers, relative to our previous study (Dobata et al. 2009a), enabled us to elucidate the detailed population genetic structure of the study population of \( P. punctatus \). First, we confirmed that thelytokous parthenogenesis is the dominant mode of reproduction, based on several indices of clonality, LD and the strongly negative \( F_{IS} \) values that we obtained. Itow et al. (1984) conducted a karyological study in \( P. punctatus \) and showed the occurrence of meiosis in germ cells, indicating that \( P. punctatus \) has automictic parthenogenesis. A cross-generational genetic comparison could not find any evidence for recombination (Dobata et al. 2009a), which implies that the restoration of ploidy is most likely accomplished by central fusion. The low level of recombination in \( P. punctatus \) likely contributes to the species fitting the population genetic models assuming strict clonality.

We found several indications for the possible occurrence of sexual reproduction, as the high variance in \( F_{IS} \) estimates across loci (−0.986 to −0.290; Table A1b) is considered to be a symptom of rare sexual reproduction (Balloux et al. 2003; De Meeus et al. 2006) and the discordance between the mtDNA haplotypes and the nuclear microsatellite genotypes (Figs 2 and 4) suggests the same. Males are occasionally found in the study populations (Itow et al. 1984; Hasegawa et al. 2001; Dobata et al. 2009a), and ergatoid queens (including Cheaters) have apparently functional spermathecae to store sperm (Itow et al. 1984; Dobata et al. 2009a, A. Gotoh, J. Billen, K. Tsuji, T. Sasaki, F. Ito, in preparation). However, although it is possible for the lineages to exchange alleles, dissection of more than 1000 ergatoid queens from the study population did not produce a single case of stored sperm in a spermatheca (T. Sasaki, personal observation, Dobata et al. 2009a). Moreover, several population genetic indices strongly suggest that clonality is a good approximation. Particularly, \( P_{ex}(f) \) value, which is an attribute of each MLG and is the probability of encountering the MLG more than once resulting from distinct sexual reproductive events, was ≤0.008. Therefore, sexual reproduction is likely to be extremely rare in \( P. punctatus \).

The issue of whether reproductive isolation between incipient social parasites and their hosts can actually be fully achieved when they originate through sympatric speciation remains controversial. The best studied example is the so-called ‘microgyne morph’ of \( Myrmica rubra \) (Savolainen & Vepsäläinen 2003; Steiner et al. 2006; Vepsäläinen et al. 2009), in which genetic differentiation between social parasites and their host seems to be almost perfect in spite of some gene flow still taking place. In \( P. punctatus \), thelytokous parthenogenesis of both Cheaters and Noncheaters provides a unique situation under which reproductive isolation is nearly fully achieved between them, making this case comparable with \( M. rubra \).

**The single Cheater lineage fixed for ergatoid queen production**

Our previous study (Dobata et al. 2009a) reported that the two MLGs which we revealed in the present study to dominate the Cheater lineage (84 out of 86, 97.7%) produced only ergatoid queens over several generations. Therefore, we concluded that this biased production is genetically determined. However, we did not consider that the Cheaters may have had more than one origin in the study population. In the present study, the estimated monophyly of the Cheater MLGs ruled out the possibility that recurrent mutations for cheating contributed to the apparent prevalence of the Cheaters in our study population. Based on their morphology, the Cheater lineage can be said to represent a peculiar
case of strong genetic caste determination (sensu Anderson et al. 2008), although this does not entail the usual functions expected from their morphology.

In spite of the increased number of genetic markers used in our present study, however, it can still not be ruled out that some lineages belonging to Noncheaters also have some genetic propensity of producing ergatoid queens rather than workers, as reported in other social insects (i.e., weak genetic caste determination sensu Anderson et al. 2008), and they thus can be more selfish than other Noncheater lineages in their colony. Because caste fate in most social insects is developmentally plastic, is affected by various environmental conditions and thus varies considerably across generations, it would be premature to conclude the existence of such propensity before investigating several generations. We also reported that the proportions of ergatoid queens to workers in some MLGs vary across generations, most likely due to environmental effects (Dobata et al. 2009a). It is worth examining whether the degree of developmental plasticity would have at least some genetic basis.

Mechanisms of persistence of the Cheater lineage

The persistence time of the Cheater lineage was estimated as ranging from about 200 to 9200 years, depending on the mutation rate. This result is contrary to the initial expectation that the disruptiveness of the Cheaters should cause the collapse of their colony within about 20 years (see Introduction), and we can thus conclude that the conditions for Emery’s rule sensu stricto to apply are apparently fulfilled in *P. punctatus*.

The estimated long persistence times indicate that there are some special mechanisms contributing to the persistence of the Cheater lineage. The first is the migration ability of the Cheater individuals among colonies, as already suggested by Dobata et al. (2009a). Previous studies have shown that strict nestmate discrimination inhibits the movement of *P. punctatus* workers among colonies (Tsuji & Itô 1986; Sanada-Morimura et al. 2003), which is consistent with the low individual migration rate of the Noncheaters estimated in this study. In stark contrast, however, we also confirmed genetically that the Cheaters transmit themselves horizontally among colonies, which allows them to escape from an evolutionary dead-end, at least in the short term. However, the risk of encountering a longer-term evolutionary dead-end remains, namely the collapse of the whole population, based on the estimated higher migration rates of disruptive Cheaters compared to Noncheaters.

Generally, the migration of cheaters in a subdivided population is a potentially serious threat to cooperation. In the framework of inclusive fitness theory, an increasing level of cheater migration homogenizes the mixture of cheaters and cooperators in subpopulations, leading to lowered relatedness (*r*), instability of cooperation (*rb−c < 0*, where *b* and *c* represent the benefit and cost of cooperation), and eventually the fixation and extinction of the cheaters themselves. In this sense, migration is a double-edged sword for the cheaters to escape an evolutionary dead-end: too low or no migration results in the local extinction of cheaters, whereas too high a migration rate results in the global extinction of both cooperators and cheaters.

Based on our genetic data, we propose that spatially limited migration of the Cheaters may facilitate the persistence of their lineage despite their disruptiveness. Given that reproductive isolation is nearly achieved through thyletokous parthenogenesis, the Noncheater–Cheater microevolutionary dynamics of *P. punctatus* becomes rather similar to host–pathogen, host–parasite, or prey–predator ecological interactions. Persistence conditions have been studied more extensively in these other two-species interactions, showing that spatially restricted migration uncouples the extinction–recolonization process of one species from that of the other (Huffaker 1958, Caswell 1978; Comins et al. 1992; Bonsall et al. 2002; Briggs & Hoopes 2004). In theory, the persistence of cheaters in a spatially structured population has been implicitly demonstrated for cooperator–cheater systems (e.g., Nowak & May 1992; Koella 2000), and our present study of *P. punctatus* showed the significant isolation by distance in both the Cheaters and the Noncheaters (Fig. 5) to generate the required population viscosity. This pattern is likely caused by the total lack of wings in *P. punctatus* females, meaning that they can only disperse on foot, a constraint that also inhibits the rapid spatial spread of Cheaters. However, the detailed effects of the estimated rates of migration and the spatial structure on the persistence of the Cheaters and Noncheaters remain to be assessed further.

Comparison with other migrating cheaters

A comparison with other examples of migrating cheater lineages and their hosts in nature provides interesting comparative insights. In honeybees, the nestmate recognition system limits the migration of workers among colonies, which also inhibits the spread and persistence of cheaters (Oldroyd 1994; Barron et al. 2001; Chaline et al. 2002), whereas a mutant cheater lineage of the Cape honeybee (*Apis mellifera capensis*) that migrates among captive colonies of the neighbouring subspecies *A. m. scutellata* has caused mass extinction of invaded colonies within about 10 years (Neumann & Moritz 2002; Dietemann et al. 2007). The Cape honeybee...
cheaters have wings and can thus migrate further than *P. punctatus*. The observed rapid collapse of the host population might in part be attributed to this property.

In the cooperation of cell ‘societies’, allograft immunity usually rejects the migration of cells between multicellular organisms, which also inhibits the spread and persistence of cancer cells (Burt & Trivers 2006; see also Cremer & Sixt 2009). However, some cancer cells have evolved the capability of migration (or transmission) among organisms, and this has been confirmed recently in some mammals (Murgia et al. 2006; Pearse & Swift 2006; Siddle et al. 2007; Rebbeck et al. 2009; Murchison et al. 2010). In Tasmanian devils (*Sarcophilus harrisii*), a fatal cancer transmits through aggression among individuals (biting) and now poses a serious threat to the host. A population genetic study of Tasmanian devils found no significant isolation by distance (Jones et al. 2004), which, together with radio-tracking studies, suggests that high dispersal ability may be a fatal risk of extinction, in line with the inferences based on our data.

Although a direct comparison of persistence time (ideally in terms of generations) is difficult because studied systems have different life cycles, the aforesaid examples illustrate that migration of cheaters in natural populations, when they impose fitness costs on their host, generally causes a relatively rapid collapse of cooperation, resulting in a short persistence time of the cheater lineages themselves. A possible way for the cheater lineages to escape extinction is a rapid adaptation to their hosts by reducing their degree of disruptiveness (or ‘virulence’). Such reduced virulence of former-cheaters has evolved the capability of migration (or transmission) among organisms, and this has been confirmed recently in some mammals (Murgia et al. 2006; Pearse & Swift 2006; Siddle et al. 2007; Rebbeck et al. 2009; Murchison et al. 2010). In Tasmanian devils (*Sarcophilus harrisii*), a fatal cancer transmits through aggression among individuals (biting) and now poses a serious threat to the host. A population genetic study of Tasmanian devils found no significant isolation by distance (Jones et al. 2004), which, together with radio-tracking studies, suggests that high dispersal ability may be a fatal risk of extinction, in line with the inferences based on our data.

Although a direct comparison of persistence time (ideally in terms of generations) is difficult because studied systems have different life cycles, the aforesaid examples illustrate that migration of cheaters in natural populations, when they impose fitness costs on their host, generally causes a relatively rapid collapse of cooperation, resulting in a short persistence time of the cheater lineages themselves. A possible way for the cheater lineages to escape extinction is a rapid adaptation to their hosts by reducing their degree of disruptiveness (or ‘virulence’). Such reduced virulence of former-cheaters has been reported in an ant social parasite (Brandt et al. 2005; see also Introduction), in an endemic population of Cape honeybee (Moritz et al. 2008), and in a sexually transmitted cancer cell lineage in dogs (CTVT; Murgia et al. 2006; Rebbeck et al. 2009). However, whether such rapid evolutionary response to avoid extinction is feasible should depend on the initial degree of virulence, which is unknown in the aforesaid examples.

Of particular interest is the proximate mechanism by which the Cheaters can transmit horizontally among colonies. We hypothesize that the Cheaters might have evolved some unidentified strategies to overcome the nestmate discrimination system of host colonies, a phenomenon that is well studied in social parasite species (Brandt et al. 2005; Cervo 2006; Lambardi et al. 2007). Such adaptations of Cheaters may also lead to counter-adaptations of the hosts (i.e., the colonies of the Noncheaters), similar to the recently observed ability of Tasmanian devils to resist the invasion of transmissible cancer cells (Jones et al. 2009). Alternatively, the difference in migration rates might simply reflect a difference in reproductive capacity between Cheaters and Noncheaters when Cheaters would lay more eggs and could produce the next generation more readily than Noncheaters after immigrating into another colony. The next challenge is therefore to examine whether there are co-evolved traits in Cheaters and the Noncheaters, in spite of their relatively recent origin.

Acknowledgements

We thank the Motomi Ito laboratory at University of Tokyo for supplying reagents and analysis tools, M. Hamaguchi for field support, C. Peeters for comments on caste definition, N. Nomura, S-I. Morinaga and T. Kato for comments on genetic analyses, J. J. Boomsma for discussion, and the late R. H. Crozier and three anonymous referees for comments on the manuscript. This research was supported in part by the Japanese Ministry of Education, Culture, Sports, Science and Technology Grants-in-Aid for Scientific Research (17207003, 17657029, 18047017, 18370012, and 20033015). S.D. was supported by a Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists (18-11584).

References


© 2010 Blackwell Publishing Ltd
workers (Apis mellifera capensis Esch.). Naturwissenschaften, 95, 507–513.


Steiner FM, Schlick-Steiner BC, Konrad H et al. (2006) No sympatric speciation here: multiple data sources show that the ant Myrmica microrubra is not a separate species but an alternate reproductive morph of Myrmica rubra. Journal of Evolutionary Biology, 19, 777–787.


## Supporting information

Additional Supporting Information may be found in the online version of this article:

**Data S1** Validation of the estimation of migration and mutation rates using F-statistics and its descendents.

## Appendix

**Table A1** (a) The F-statistics of each of the 13 microsatellite loci. (b) The estimation of the upper bounds of the mutation rate per locus, based on \((m_1 + m_2)\) estimates of Noncheater subpopulations using the same setup as in Table 2.

<table>
<thead>
<tr>
<th>Subpopulations</th>
<th>Locus</th>
<th>(F_{IS})</th>
<th>(F_{IT})</th>
<th>(F_{ST}) diploid</th>
<th>(F_{ST}) haploid</th>
<th>(F_{ST}) diploid</th>
<th>(F_{IS}) based</th>
<th>(F_{ST}) haploid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noncheater</td>
<td>Pp1</td>
<td>-0.807</td>
<td>-0.568</td>
<td>0.132</td>
<td>0.612</td>
<td>9.287 (\times 10^{-5})</td>
<td>3.378 (\times 10^{-6})</td>
<td>1.791 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>Pp2</td>
<td>-0.877</td>
<td>-0.715</td>
<td>0.087</td>
<td>0.608</td>
<td>1.482 (\times 10^{-4})</td>
<td>1.981 (\times 10^{-6})</td>
<td>1.821 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>MYRT3</td>
<td>-0.290</td>
<td>0.125</td>
<td>0.322</td>
<td>0.518</td>
<td>2.974 (\times 10^{-5})</td>
<td>3.458 (\times 10^{-5})</td>
<td>2.628 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb13</td>
<td>-0.552</td>
<td>-0.108</td>
<td>0.286</td>
<td>0.636</td>
<td>3.526 (\times 10^{-5})</td>
<td>1.146 (\times 10^{-5})</td>
<td>1.617 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb27</td>
<td>-0.824</td>
<td>-0.686</td>
<td>0.076</td>
<td>0.478</td>
<td>1.717 (\times 10^{-4})</td>
<td>3.016 (\times 10^{-6})</td>
<td>3.085 (\times 10^{-6})</td>
</tr>
<tr>
<td></td>
<td>ppmb33</td>
<td>-0.986</td>
<td>-0.669</td>
<td>0.160</td>
<td>0.956</td>
<td>7.414 (\times 10^{-5})</td>
<td>2.005 (\times 10^{-7})</td>
<td>1.300 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb104</td>
<td>-0.912</td>
<td>-0.662</td>
<td>0.131</td>
<td>0.770</td>
<td>9.368 (\times 10^{-5})</td>
<td>1.363 (\times 10^{-6})</td>
<td>8.437 (\times 10^{-6})</td>
</tr>
<tr>
<td></td>
<td>ppmb105</td>
<td>-0.471</td>
<td>-0.070</td>
<td>0.273</td>
<td>0.633</td>
<td>3.761 (\times 10^{-5})</td>
<td>1.586 (\times 10^{-5})</td>
<td>1.638 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb106</td>
<td>-0.765</td>
<td>-0.514</td>
<td>0.142</td>
<td>0.587</td>
<td>8.533 (\times 10^{-5})</td>
<td>4.338 (\times 10^{-6})</td>
<td>1.987 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb114</td>
<td>-0.607</td>
<td>-0.157</td>
<td>0.280</td>
<td>0.690</td>
<td>3.632 (\times 10^{-5})</td>
<td>9.144 (\times 10^{-6})</td>
<td>1.269 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb132</td>
<td>-0.651</td>
<td>-0.214</td>
<td>0.264</td>
<td>0.673</td>
<td>3.937 (\times 10^{-5})</td>
<td>7.571 (\times 10^{-6})</td>
<td>1.372 (\times 10^{-5})</td>
</tr>
<tr>
<td></td>
<td>ppmb139</td>
<td>-0.549</td>
<td>0.175</td>
<td>0.467</td>
<td>0.746</td>
<td>1.612 (\times 10^{-5})</td>
<td>1.160 (\times 10^{-5})</td>
<td>9.617 (\times 10^{-6})</td>
</tr>
<tr>
<td></td>
<td>ppmb204</td>
<td>-0.796</td>
<td>-0.382</td>
<td>0.230</td>
<td>0.696</td>
<td>4.728 (\times 10^{-5})</td>
<td>3.619 (\times 10^{-6})</td>
<td>1.234 (\times 10^{-5})</td>
</tr>
<tr>
<td>Cheater</td>
<td>Pp1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>Pp2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MYRT3</td>
<td>-0.809</td>
<td>-0.793</td>
<td>0.009</td>
<td>0.054</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb13</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb27</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb33</td>
<td>-0.977</td>
<td>-0.977</td>
<td>0.000</td>
<td>0.034</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb104</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb105</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb106</td>
<td>-0.922</td>
<td>-0.733</td>
<td>0.099</td>
<td>0.735</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb114</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb132</td>
<td>-0.977</td>
<td>-0.977</td>
<td>0.000</td>
<td>0.034</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb139</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>ppmb204</td>
<td>-0.878</td>
<td>-0.869</td>
<td>0.005</td>
<td>0.102</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Table A2. Estimating the age of the cheater lineage

<table>
<thead>
<tr>
<th>Locus</th>
<th>Allelic position</th>
<th>MLG Case 1</th>
<th>Case 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>Pp1</td>
<td>1</td>
<td>220</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>225</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>238</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>238</td>
<td>—</td>
</tr>
<tr>
<td>MYRT3</td>
<td>1</td>
<td>179</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>181</td>
<td>—</td>
</tr>
<tr>
<td>ppm213</td>
<td>1</td>
<td>272</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>272</td>
<td>—</td>
</tr>
<tr>
<td>ppm227</td>
<td>1</td>
<td>157</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>157</td>
<td>—</td>
</tr>
<tr>
<td>ppm233</td>
<td>1</td>
<td>310</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>322</td>
<td>—</td>
</tr>
<tr>
<td>ppm204</td>
<td>1</td>
<td>314</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>334</td>
<td>336*</td>
</tr>
<tr>
<td>ppm114</td>
<td>1</td>
<td>308</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>312</td>
<td>—</td>
</tr>
<tr>
<td>ppm132</td>
<td>1</td>
<td>268</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>270</td>
<td>—</td>
</tr>
<tr>
<td>ppm133</td>
<td>1</td>
<td>142</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>142</td>
<td>—</td>
</tr>
<tr>
<td>ppm204</td>
<td>1</td>
<td>203</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>225</td>
<td>—</td>
</tr>
<tr>
<td>n</td>
<td>63</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

MLG A is that of the ancestral Cheater (numbers shown in bold), and newly mutated alleles in the other MLGs are shown with an asterisk. The probability of survival of the ancestral allele, \( P_{ij} \), was calculated as \( P_{ij} = (86 - A_{mut}) / 86 \), where \( A_{mut} \) is the number of substituted alleles in the sampled individuals. \( t_{ij} \) values were calculated at all allelic positions, using the upper bounds of the mutation rate \( \mu \) of each locus \( i \) estimated by three \( F \)-statistics based methods using the Noncheater subpopulations. Because reproduction most likely occurs by automictic parthenogenesis, some substituted alleles (shown in italic) at one allelic position could not be assigned to either recombination or mutation when they matched either the ancestral allele or the mutated allele at the other allelic position. Therefore, we calculated the estimates for both cases (Case 1 and 2, respectively). See the main text for details.

MLG, multilocus genotype.
Supporting Information

Here we validate the estimation of migration and mutation rates in *P. punctatus* based on the
*F*-statistics. Given the *F*<sub>ST</sub> values and the size of each subpopulation (*N*), one can calculate the migration rate
of individuals between subpopulations per generation (*m*) and the mutation rate per locus per generation (*μ*),
according to the classical equation *F*<sub>ST</sub> = 1/[4*N*(*m* + *μ*) + 1] (*μ* can be ignored when *μ* << *m*; Weir 1996).
Unfortunately, however, this approximation often leads to biased estimations, because real populations
frequently violate the underlying assumptions (Whitlock & McCauley 1999). To obtain more accurate
estimates, we here consider several relevant effects of the complexity of the *P. punctatus* population on the
estimation of migration rate and work them into the equation.

*Clonality.* For clonal diploids subdivided in numerous demes (which is the case in *P. punctatus*), de Meeûs
and Balloux (2005) reformulated *F*-statistics and showed that the following equations characterize the
population:

\[
F_{IS} = \frac{-F_{ST}}{1 - F_{ST}} = -\frac{1}{4*N*(m + \mu) + 1} \quad (A1),
\]

\[
F_{ST} = \frac{1}{4*N*(m + \mu) + 2} \quad (A2), \text{ and}
\]

\[
F_{IT} = 0 \quad (A3).
\]

They also showed that the *F*<sub>IS</sub> value gives a more accurate estimate of (*m* + *μ*). A comparison of their
*F*<sub>ST</sub>-based estimation with the classical one indicates that clonality will tend to reduce differentiation among
subpopulations, which means that the real migration rates are lower than those estimated by the original
equation. It should be noted that the above equations also fit surprisingly well to a stepping-stone model
(which is more appropriate for *P. punctatus*), although de Meeûs and Balloux (2005) originally assumed an
island model.

Additionally, based on the nearly strict clonality of the study population, the genotypic data can be
treated as if obtained from haploids. One can calculate corresponding *F*<sub>ST</sub> values (*F*<sub>IT</sub> and *F*<sub>IS</sub> are undefined

for haploids) and apply the following equation for haploids:

\[ F_{ST} = \frac{1}{2N_{hap}(m + \mu) + 1} \]  

(A4),

where \( N_{hap} \) is the number of clonal individuals treated as being haploids.

**Extinction and recolonization.** The life cycle of the colonies of *P. punctatus* is most likely to be approximated by the extinction–recolonization process (Slatkin 1977, Wade & McCauley 1988, Whitlock & McCauley 1990), in which subpopulations undergo frequent turnovers. To assess the effect of this process on genetic differentiation among subpopulations, it has been suggested that the distinction between colonization and migration is essential (Wade & McCauley 1988). In Noncheater subpopulations of *P. punctatus*, colonization by fission is initiated by a far larger number of individuals than those involved in migration (Itow et al. 1984, Tsuji 1988). For fission-reproducing subpopulations, genetic differentiation can be approximately measured by

\[ F_{ST} = \frac{(1 + \pi_{fis})[4N(m + \mu) + 1 + \pi_{fis}]}{4} \]  

(A5),

where \( \pi_{fis} \) is the proportion \((0 \leq \pi_{fis} \leq 1)\) of subpopulations that have arisen from random fission events (equally divided into halves) in the last generation (Whitlock 1994). The equation (A5) yields \((m + \mu) = (1 + \pi_{fis})(1/F_{ST} - 1)/4N\), so that the fission colonization of Noncheater subpopulations possibly leads to subtle underestimation on the \((m + \mu)\) at the scale of \(1/2 \leq 1/(1 + \pi_{fis}) \leq 1\).

In stark contrast, the foundation of Cheater subpopulations can be achieved by as few as one individual intruding into a colony. This indicates that Cheater subpopulations are most likely to be approximated by a migrant pool model (Slatkin 1977) and that the colonization is not distinct from migration \((K \approx Nm\) in terms of Wade & McCauley 1988). This colonization process is also known to enhance genetic differentiation among Cheater subpopulations (Wade & McCauley 1988); that is, the observed level of genetic differentiation can be achieved by more frequent individual migration than that estimated from the classical equation, which supports our working hypothesis that the cheaters have a higher migration rate than...
the noncheaters (see also Figure 5).

Nonequilibrium population. Because the Cheaters are most likely to have originated from a single mutant individual, the estimated persistence time of the cheaters suggests that Cheater subpopulations have not yet reached an equilibrium between the forces of migration and genetic drift, which is assumed in the classical equation and its descendents. Therefore, the migration rate of the cheaters might be overestimated. Denoting the generation from the ancestor \( t \) by superscripts, the recurrence equation

\[
F_{ST}^{[t]} = (1 - \mu)^2 \left[ \frac{1}{2}N + (1 - m)^2 (1 - \frac{1}{2}N)F_{ST}^{[t-1]} \right]
\]

(A6)
yields

\[
F_{ST}^{[t]} = (1 - \frac{1}{2}N) [1 + \frac{(L' - 1)(L - 1)}{L}] + F_{ST}^{[0]} L^t
\]

(A7),

where \( L = (1 - m)^2 (1 - \frac{1}{2}N) \) (Whitlock 1992). As \( F_{ST}^{[0]} = 0 \) and \( N = 463 \) for Cheater subpopulations, assuming \( \mu = 10^{-5} \) gives the time \( t \) for \( F_{ST}^{[t]} \) to reach 95% of the equilibrium value as 44.7 generations, even at the migration rate \( m = 10^{-5} \) (approximately the same as that of Noncheater subpopulations). Because higher \( m \) gives shorter \( t \), we conclude that the population of Cheater subpopulations may have already reached near equilibrium.

Conclusion. In sum, the above analyses show that the estimation based on the classical equation is likely to give approximately the true value of \( (m + \mu) \), at least for Noncheater subpopulations. More importantly, it is safe to say that the Cheaters have a higher individual migration rate than the Noncheaters.
References


