Relaxed selection underlies genome erosion in socially parasitic ant species

Schrader, Lukas; Pan, Hailin; Bollazzi, Martin; Schiøtt, Morten; Larabee, Fredrick J.; Bi, Xupeng; Deng, Yuan; Zhang, Guojie; Boomsma, Jacobus J.; Rabeling, Christian

Published in:
Nature Communications

DOI:
10.1038/s41467-021-23178-w

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Schrader, L., Pan, H., Bollazzi, M., Schiøtt, M., Larabee, F. J., Bi, X., Deng, Y., Zhang, G., Boomsma, J. J., & Rabeling, C. (2021). Relaxed selection underlies genome erosion in socially parasitic ant species. Nature Communications, 12, [2918]. https://doi.org/10.1038/s41467-021-23178-w
Relaxed selection underlies genome erosion in socially parasitic ant species

Lukas Schrader1,2✉, Hailin Pan3, Martin Bollazzi4, Morten Schiøtt1, Fredrick J. Larabee5, Xupeng Bi3, Yuan Deng3, Guojie Zhang1,3,6, Jacobus J. Boomsma1,9 & Christian Rabeling7,8,9✉

Inquiline ants are highly specialized and obligate social parasites that infiltrate and exploit colonies of closely related species. They have evolved many times convergently, are often evolutionarily young lineages, and are almost invariably rare. Focusing on the leaf-cutting ant genus *Acromyrmex*, we compared genomes of three inquiline social parasites with their free-living, closely-related hosts. The social parasite genomes show distinct signatures of erosion compared to the host lineages, as a consequence of relaxed selective constraints on traits associated with cooperative ant colony life and of inquilines having very small effective population sizes. We find parallel gene losses, particularly in olfactory receptors, consistent with inquiline species having highly reduced social behavioral repertoires. Many of the genomic changes that we uncover resemble those observed in the genomes of obligate non-social parasites and intracellular endosymbionts that branched off into highly specialized, host-dependent niches.
Radical shifts to specialized ecological niches change selection regimes often resulting in fundamental genomic and phenotypic transformation. Selection on traits adapted to the ancestral niche will relax and positive selection will favor adaptations to new ecological settings and matching life histories. In addition, niche specialization can lead to small isolated populations with limited gene flow, decreasing the efficiency of selection. Recent research revealed that consequences of relaxed selection are partly predictable, both at a general level and at the level of particular traits. The consequences of relaxed selection can be dramatic, e.g., in the convergence of eyesight in animals that shifted from a diurnal habitat to permanent darkness. Cave fish and subterranean mammals rapidly lost or reduced vision, consistent with eyes as complex organs no longer being maintained by stabilizing selection. Chloroplasts and the genetic capacity for photosynthesis degenerated in plants that transitioned to a holoparasitic lifestyle, and genomes and body plans of parasitic Myxozoa are significantly reduced compared to those of free-living cnidarians. Similar changes are also known from other obligate parasites. Finally, evolutionary transitions in bacteria from a free-living to an endosymbiotic lifestyle are almost invariably associated with convergent phenotypic and genomic reductions. In many of these cases, changes in effective population sizes due to permanent asexual reproduction or frequent bottlenecks reduce the efficacy of selection. This results in impoverished genomes with an excess of slightly deleterious alleles and only essential functions being actively maintained by selection. Genomes of recently acquired endosymbionts have eroded significantly less than genomes of ancient endosymbionts and similar trends apply to the nuclear and cytoplasmic genomes of parasitic plants and animals. In general, genomic and phenotypic changes following niche specialization appear to often evolve surprisingly fast.

It is less generally known that there are also examples of social niche shifts that have induced rapid change in morphological and behavioral phenotypes, with the most extreme transitions being found in inquiline social parasites of free-living ants. These morphologically and behaviorally highly specialized species parasitize colonies of other ants by exploiting their hosts’ social behavior such as brood care, food resources, and an established and protected nest environment. Analogous to intracellular and specialized non-social parasites, the evolution of inquilines has classically been described as a degenerative process almost invariably associated with convergent morphological and behavioral characteristics across independently evolved lineages. This so-called “social parasite syndrome” subsumes an array of regressive phenotypic changes, such as the reduction or complete loss of the worker caste, simplified mouthparts, antennae and integuments, loss of certain exocrine glands, and a nervous system of reduced complexity likely associated with a holoparasitic lifestyle. The genus Acromyrmex is well-suited for comparative studies of social parasite evolution because it harbors five recently evolved species of inquilines (Table 1). Three of these, the Brazilian Acromyrmex ameliae, the Panamanian Acromyrmex insinuator, and the Uruguayan Acromyrmex charruanus, are morphologically less specialized than the other two, the South American Pseudoatta argentina and the Brazilian Acromyrmex fowleri, both of which exhibit suits of behavioral and morphological traits that strictly adhere to the social parasite syndrome. To explore whether the dynamics of evolutionary change in inquiline social parasites follow similar patterns as documented in non-social parasites, intracellular endosymbionts, and other lineages that underwent radical niche shifts, we obtained and analyzed full genome sequences of three Acromyrmex inquilines and their respective host species.

Our analyses confirm that inquiline lineages have accumulated convergently evolved signatures of genome-wide and trait-specific genetic erosion. In particular, we find that (1) natural selection in inquiline species is relaxed overall, most likely because many ancestral social traits have become redundant and because N_c, which is expected to be low in social insects per se, is further decreased in social parasites, (2) losses and reductions have particularly affected the olfactory receptor repertoire of social parasites, (3) genomic changes are most extreme in the highly modified social parasite with the most quintessential match to the inquiline syndrome, and (4) losses/reductions at the genome-level emerged through fixation of different kinds of genomic rearrangements. We infer that relaxed selection accelerated general genome erosion in social parasites and alleviated evolutionary constraints, which facilitated rapid adaptive evolution of specific traits associated with a socially parasitic lifestyle. Our findings advance our understanding of the genomic consequences of transitioning to a novel, highly specialized life history and provide detailed insights into the molecular evolution of social parasitism in ants.

Results
To identify genetic changes associated with the shift from cooperative social colony life to social parasitism, we de novo sequenced genomes of four Acromyrmex species, including the three inquiline social parasites A. charruanus, A. insinuator, and Pseudoatta argentina, as well as A. heyeri, the host species of A. charruanus and P. argentina. The genome of A. echinatior, the host of A. insinuator, was published previously and included in our comparisons. De novo assembly produced genomes of similar quality as previously published genomes of other ants. Genome assembly completeness, assessed with BUSCO (“genome” mode), ranged between 91.4% for A. charruanus and 97.3% for A. insinuator, which is comparable to other published attine genome assemblies. Protein-coding genes were annotated for the four newly sequenced genomes and reannotated for seven previously published attine genomes. We annotated between 13,249 (P. argentina) and 16,193 (A. heyeri) protein-coding genes (median = 14,688) and retrieved 6338 single-copy orthologs. BUSCO scores for protein annotations were lower than for the genome assemblies, particularly for P. argentina (89.1% vs. 97.1%), but these differences could be accounted for in our analyses of gene family size evolution (see below). Time-calibrated phylogenetic analyses of orthologous fourfold degenerate sites inferred two independent origins of inquiline social parasitism in Acromyrmex (Fig. 1). First, a South American lineage where A. heyeri separated from the common (putatively socially parasitic) ancestor of A. charruanus and P. argentina before the two social parasites diverged and, second, a Central American speciation event when A. insinuator diverged from its host A. echinatior. Both origins of inquilineism are evolutionarily recent, estimated to be ~2.5 Ma for the divergence between A. heyeri and the common ancestor of A. charruanus and P. argentina, and ~1 Ma for the divergence between A. insinuator and A. echinatior. The two South American social parasites,
A. charruanus and P. argentina, diverged ~1.6 Ma ago from their most recent common ancestor. Our analyses largely agreed with previously published divergence time estimates in the leaf-cutting ants, i.e., the divergence between Acromyrmex and Atta (~12 Ma, compared to 10–16 Ma45,47,48), but our present study suggests a more recent divergence of At. cephalotes and At. colombica (~2 Ma) than previously estimated (~7 Ma)47.

Genomic evidence of relaxed selection in social parasites. All inquiline social parasite species are rare, patchily distributed, live in small populations, and likely to be inbred27,35. Inbreeding and small census population sizes are expected to significantly decrease the effective population sizes (Ne) of social parasites relative to their non-parasitic relatives49,50. Coalescent analyses on one host-parasite pair (A. echinator and A. insinuator) and a third, non-parasitic, closely related species (A. octospinosus)40 for which the appropriate samples (two individuals from different colonies) could be collected, confirm that Ne was indeed consistently smaller in the social parasite compared to its free-living relatives (Fig. 2A). Evolutionary rate analyses revealed genome-wide relaxation of selection on coding sequences across all socially parasitic lineages (Fig. 2B–E). Consistent with relaxed selection, codon usage was overall less biased in parasites (see Supplementary Information for details), although the signal was much weaker than at the level of non-synonymous substitution rates. Average dN/dS ratios of single-copy orthologs were significantly increased in social parasite lineages relative to their non-parasitic sister lineages (pMann–Whitney <1e−5, Fig. 2B), and selection intensity parameter log2(k) analysis51 showed this difference to be due to an overall preponderance of relaxed selection (Fig. 2C–E). We uncovered clear signatures of relaxed purifying selection (in coding regions with average dN/dS < 1, Fig. 2C) and

A. insinuator, A. charruanus and P. argentina were included in this study.
relaxed positive selection (where average \( d_\text{S}/d_\text{S} > 1 \), Fig. 2D). In both cases, average evolutionary rates in social parasites shifted towards \( d_\text{S}/d_\text{S} \) ratios of 1, which is indicative of neutral evolution. Accordingly, the selection intensity parameter \( \log_2(k) \) was weakly negative for most genes throughout these genomes, and strongly negative for a long-stretched tail end (Fig. 2E). However, there was also a majority of genes that experienced a modest intensification of selection as the peaks towards the right in Fig. 2E indicate. Our results are consistent with a prevailing pattern of genome-wide erosion across social parasite genomes punctuated by intensified selection on specific genes.

In addition to genome-wide signatures of relaxed selection, the transition from cooperative social colony life to exploitative social parasitism should particularly relax selection on genes coding for complex phenotypic traits associated with social behavior such as brood care, foraging, and nest construction. We identified 233 genes (3.34% of the tested genes) showing evidence of relaxed selection in at least one of the social parasite branches (at FDR < 0.1 and relaxation parameter \( k < 1 \)) among single-copy orthologs (see Supplementary Figs. 14–16, Supplementary Table 34). However, we also detected signatures of intensified selection (at FDR < 0.1, \( k > 1 \)) in 102 genes (1.46%, Supplementary Figs. 17–19, Supplementary Table 35), suggesting that the convergent niche shifts to social parasitism also triggered adaptive changes within an overall context of more genomic erosion.

**Gene loss in social parasites.** Gene content and synteny have generally evolved rapidly across the genomes of ants and of fungus-growing ants in particular.7,32,33 In such generally dynamic background of change, we expected to see genomic signatures of trait loss of similar or larger magnitude than what is known for non-social animal and plant parasites.20,54 To explore gene family evolution in social parasites, we modeled gene family size changes with CAFE, accounting for potential biases from...
incomplete assemblies and annotations. Models included Acromyrmex host and social parasite species, with At. colombica serving as outgroup, from which no inquilines are known. Our analysis showed that gene family evolution at three of the four social parasite nodes is indeed largely characterized by gene losses (Fig. 3A, B). We uncovered high rates of gene loss and low rates of gene gain in P. argentina, and also at the two origins of social parasitism (i.e., in A. insinuator and the ancestor of A. charruanus and P. argentina). Intriguingly, however, A. charruanus showed a different pattern, with the number of gene gains exceeding the number of gene losses, similar to the pattern observed in free-living attine species. This indicates that during the first ~1 Ma of inquiline evolution, A. insinuator and the common ancestor of A. charruanus and P. argentina followed a parallel pattern of predominantly gene loss, but that the later split between A. charruanus and P. argentina induced remarkably divergent trajectories in gene family size evolution.

We identified gene families characterized by high rates of gene loss in social parasite species compared to their hosts by functional enrichment analyses. This showed that gene loss had most significantly affected olfactory receptors (ORs), as only GO terms related to olfaction were significantly enriched (e.g., GO:0004984 “olfactory receptor activity”, FDR = 5.6e-6, Supplementary Table 45). No GO terms were significantly enriched among genes gained in social parasites or genes lost in the host species. The odorant receptor gene family is directly relevant to the behavior of social insects because colony members are crucially dependent on efficient chemical communication via olfactory cues. It was recently documented that the OR gene family experienced a significant expansion in ants55,56. Our comprehensive annotation and analysis of OR repertoires (Fig. 4, Table 2, see Supplementary Information for details) confirmed a parallel reduction in all three Acromyrmex social parasites, consistent with these species being less engaged in social interactions than their closely related host species. In A. insinuator we identified 32 fewer genes containing the characteristic 7tm6 domain of ORs than in its host A. echinatior. In the other lineage, we detected a comparable reduction of 38 fewer ORs in A. charruanus and a more massive reduction in P. argentina where only 263 OR genes remained compared to 467 OR genes in the host A. heyeri. This 44% loss implied that P. argentina approached the number of OR genes typical for non-social Hymenoptera such as Nasonia vitripennis with 225 ORs57, consistent with P. argentina having the most highly specialized socially parasitic lifestyle of all known Acromyrmex inquilines43,58 (Table 1). These results suggest that relaxed purifying selection on formerly important social traits has affected odorant receptor gene losses of social parasites in proportion to their degree of phenotypic specialization.

To gain a more detailed understanding of OR gene family evolution in fungus-growing ants, we reconstructed a family-wide gene tree, dividing genes into 321 clades belonging to 27 gene subfamilies (Fig. 4A). More than 30% of all OR clades (101 of 321) contained fewer paralogs in at least one parasite species compared to their respective hosts (orange highlights in Fig. 4A), with six clades being convergently reduced in the parasites relative to their hosts (orange number labels). In accordance, fewer large OR gene arrays (>8 genes) were retained in parasites...
Gene losses in social parasite lineages occurred significantly more often in clades that had previously experienced gene gains during the radiation of free-living Atta and Acromyrmex leaf-cutting ant species (blue blocks; cf. basal branches in Fig. 1). Reconciled gene tree of OR clade 285 belonging to the 9-exon subfamily that diversified independently and at a particularly high rate in the two leaf-cutting ant genera and subsequently underwent convergent gene losses in all social parasites. Branches with inferred gene deletion events in inquiline social parasites and their hosts are marked in orange diamonds, duplications are marked with circles colored according to the taxonomic affiliation of the ants, and speciation nodes are marked with triangles. Background shading colors highlight the independent expansions in Atta and Acromyrmex.

Table 2 Gene family sizes in Atta and Acromyrmex leaf-cutting ants for odorant receptors (ORs), elongases, gustatory receptors (GRs), major royal jelly and yellow proteins (MRJP/y), and cuticular proteins (CPRs).

|         | ORs | Elongases | GRs | MRJP/y | CPRs |
|---------|-----|-----------|-----|--------|------|
| A. echinatior | 456 (138) | 14 (2) | 62 (7) | 24 (2) | 32 (8) |
| A. insinuator | 424 (65) | 12 (3) | 61 (5) | 21 (3) | 31 (8) |
| A. heyeri | 467 (181) | 11 (4) | 61 (8) | 22 | 31 (9) |
| A. charruanus | 429 (72) | 12 (1) | 60 (9) | 22 (2) | 33 (9) |
| P. argentina | 263 (73) | 11 (3) | 27 (4) | 13 (4) | 31 (10) |
| At. colombica | 472 (72) | 14 | 62 (9) | 20 (1) | 34 (8) |
| At. cephalotes | 454 (114) | 13 (1) | 62 (5) | 17 (5) | 32 (7) |

Incomplete genes (due to misannotations, misassemblies, or pseudogenization) are given in parentheses.
The repertoire of *P. argentina* made this social parasite cluster away from the leaf-cutting ant species together with representatives of *Cyphomyrmex*, *Myctomomyrctius*, *Paratrichomyrmex*, and *Trachymyrmex* (until a recent revision, the latter three belonged to the paraphyletic grade previously referred to as *Trachymyrmex*). That the two less modified *Acromyrmex* social parasites clustered together based on their socially parasitic life history rather than their phylogenetic co-ancestry suggest that OR gene repertoires tend to become convergently reduced relatively early during the inquiline specialization process (Fig. 5A).

Previous studies have suggested that during ant evolution the expansions in size and glomeruli number of the olfactory lobes (OL) have remained correlated with the expansion of the OR gene repertoire. Using microCT scans, we explored whether the relative size of the OL was consequently smaller in inquiline social parasite species compared to their hosts, mirroring the reduced OR repertoire encoded in their genomes. As hypothesized, relative OL size was ca. 12% smaller in *P. argentina* compared to their host *A. charruanus* (Fig. 5B), and analogously OL size was reduced by ca. 38% in *A. insinuator* relative to its host *A. echinatior*. This reduction in social parasites was marginally non-significant (one-sided *t*-test, *p* = 0.059, *df* = 2, *t* = −2.65). The results should be regarded as preliminary because only a single individual per species could be brain-imaged due to the rarity of well-preserved samples of the social parasites. Notwithstanding this caveat, the preliminary analysis of brain morphology provided here suggests that reductions are proportional to the degree of specialization found in social parasites.

We also explored whether gene losses in social parasites were apparent in other gene families than in the OR family. We produced detailed annotations of four candidate gene families: (1) gustatory receptors (GRs), which are involved in chemical communication similar to odorant receptors; (2) Major Royal Jelly (MRJP)/yellow genes, which play a role in caste differentiation; (3) elongases, which contribute to the production of cuticular hydrocarbons; and (4) cuticular protein genes (CPRs), which are involved in exoskeleton formation.

To gain a more detailed understanding of the putative mutational events accountable for gene losses in social parasite genomes, we analyzed gene synten and whole-genome rearrangements. To that end, we produced detailed annotations of four candidate gene families: (1) gustatory receptors (GRs), which are involved in chemical communication similar to odorant receptors; (2) Major Royal Jelly (MRJP)/yellow genes, which play a role in caste differentiation; (3) elongases, which contribute to the production of cuticular hydrocarbons; and (4) cuticular protein genes (CPRs), which are involved in exoskeleton formation.
genome alignments (Fig. 6). Using gene orthology information, we assessed microsynteny of the Acromyrmex and P. argentina genomes relative to At. colombica within larger syntenic blocks and quantified synteny conservation by calculating gene order correlation coefficients (see Supplementary Information). Two perfectly syntenic regions have a correlation coefficient of \( \tau = 1 \), whereas microsyntenic rearrangements, such as inversions and translocations, decrease the value for \( \tau \) where microsyntenic rearrangements, such as inversions and translocations, decrease the value for \( \tau \) where

\[
\tau = \frac{f_{\text{si}} - f_{\text{trans}}} {f_{\text{si}} + f_{\text{trans}}}
\]

and quantification of orthologous genes, with \( \tau = 1 \) for perfectly conserved syntenic regions and \( \tau < 1 \) for syntenic regions containing microsyntenic changes, and evaluated for significance with one-sided Mann–Whitney tests: hosts vs parasites: \( p = 0.0037, z = 2.67, n = 610 \); A. echinatior vs A. insinuator: \( p = 0.0191, z = 2.06, n = 236 \); A. heyeri vs A. charruanus: \( p = 0.3223, z = 0.46, n = 250 \); A. heyeri vs P. argentina: \( p = 0.0091, z = 2.36, n = 246 \). Boxplot centers show the median, hinges show the first and third quartiles, and whiskers show the 1.5 x inter-quartile range. B–D Social parasite genome evolution is characterized by an excess of inversions in A. insinuator (B), transpositions in the stemgroup of A. charruanus and P. argentina (C), and deletions in P. argentina (D). Compared to A. insinuator and P. argentina, the A. charruanus genome experienced relatively few deletions, transpositions and duplications after the split from its common ancestor with P. argentina, which was characterized by a very high number of transpositions.

**Discussion**

Our comparative study revealed that the genomes of the three inquiline social parasite species P. argentina, A. charruanus, and A. insinuator have substantially diverged from their closely related Acromyrmex hosts. However, the details of the structural divergences differ, because A. insinuator evolved inquiline independently from the other two social parasites that share a common ancestor but diverged and evolved independently for the past 1.63 million years. Our data indicate that these divergences from free-living hosts are reflected in genome-wide and gene-family-wide signatures of relaxed selection, an increased abundance of various types of genomic rearrangements, and a substantial history of gene loss, contrasting previous studies. Gene losses were detected most prominently in chemosensory gene families including olfactory and GRs and they accumulated over just 1–2.5 million years. Such dramatic evolutionary changes are only expected after fundamental changes in social niche and life history and are predicted to be at least partly convergent when such changes are analogous. Our results largely match this scenario because ancestral selection pressures on traits directly related to cooperative social colony life, such as chemosensory perception for communication, synaptic plasticity for behavioral differentiation, and secretion of semiochemicals for social interactions could all be alleviated because social parasites came to exploit the foraging efforts, nursing behavior, and colony infrastructure of their hosts.

The fundamental changes in life history experienced by ant inquiline social parasites are on par with those of intracellular endosymbionts and the most specialized non-social parasites, and they likely surpass transitions made by subterranean and cave-dwelling animals where a social or symbiotic niche dimension is lacking. The consistent signatures of relaxed selection coinciding with regressions and functional losses in phenotypic traits among other social parasites and the deep divergences from their non-specialist sister clades that remained much closer to the ancestral lifestyle. The small effective population sizes and low population densities of social parasites, as well as their frequent inbreeding via sib-mating, all reduce the efficacy of natural selection and increase the prevalence of genetic drift, simultaneously enhancing the rates by which ancestral traits are lost and constraining the rates by which novel adaptive traits can emerge. Similar mechanisms have been demonstrated to underlie globally relaxed selective constraints in obligate plant parasites and intracellular endosymbionts, both of which experienced significant effects of drift and reduced gene flow due to their specialized life histories and modified population structures. Genome evolution in ant social parasites therefore strikingly resembles the dynamics known from other species that transitioned to highly specialized and unusual niches.
The ‘unalso’ here is important because the genomes of ants have evolved for more than 100 million years under a selection regime that emanated from a single major transition to a novel superorganismal level of social organizational mediated by a unique gene regulation network27 for obligate queen-worker caste segregation based on unconditional altruism2. This produced extant biodiversity of 17 subfamilies, 337 genera, and more than 13,800 extant species23. It is therefore no surprise that parallel shifts to a socially parasitic lifestyle abandoning this fundamental ancestral condition, usually based on outbreeding and larger effective populations, leave significant genomic footprints31. The results of our analyses of just three of these inquiline species confirm that ant social parasites offer important ‘loss-of-function’ study systems for identifying hallmarks of cooperative social colony life at both the phenotypic and the genomic levels31,67. Analogous evolutionary dynamics have been identified in Myxozoa, microscopic obligate parasites that evolved from multicellular cnidian ancestors, where many genes essential for complex multicellularity were secondarily lost when parasitic species regressed to a dramatically reduced body plan16. Ant genomes evolve relatively fast with rapid losses of synteny and quickly changing gene content47,52,53, potentially related to the high recombination rates of the social Hymenoptera that transitioned to superorganismality74–79. A previous comparative genomics study of fungus-growing ants further demonstrated that rates of synteny loss are significantly higher in attine ants than in ants in general17. The fact that Acromyrmex inquilines show even faster rates of synteny loss underlines that their genomes might be prone to very rapid modification, which is consistent with ant inquiline lineages being almost invariably of recent origin28. The acceleration of evolutionary change under relaxation of selective constraints likely also contributed to the emergence of novel adaptive traits in social parasites31, although at a slower rate because selection needs to be strong to overcome erosion by drift. Life as an inquiline social parasite universally requires new and highly specific adaptations to host finding, intrusion into mature host colonies, and manipulation of the hosts’ brood care behavior to favor the non-related inquiline brood. Some of these adaptations have been studied in A. insinuator77–79, A. melalia80, and a number of recent reviews have also underlined the consistency of convergent adaptations across ant inquiline social parasites28,31,81. The fact that we did not detect clear genomic signatures of adaptive change shared across the three independently evolved social parasite species may well be because convergent evolution does not necessarily rely on the same set of genes, particularly in lineages heavily affected by drift, or because changes were concentrated in regulatory, non-coding regions. It is also possible that adaptive changes are mediated epigenetically31. These aspects were beyond the scope of our present analyses and deserve further study.

We found that OR repertoires were consistently reduced across inquiline parasites, with the most dramatic reduction in the phenotypically and behaviorally most specialized inquiline P. argentina. These results need to be evaluated in light of the general expansion of the OR gene family as key evolutionary innovation associated with the emergence of complex social behavior in ants in general55,60,61,82. Inquiline ants rely on the brood care behavior and social infrastructure provided by their host colony and therefore depend to a lesser degree on the sophisticated and fine-tuned general chemical communication systems of their closely related free-living hosts, as long as they are adapted to successfully recognize and manipulate host colonies. Inquilines have in fact lost most ancestral social traits and do not engage in brood-care, colony founding, foraging, or nest-defense behaviors (e.g.77,78), so it is unsurprising that many ORs became obsolete and were consequently no longer maintained by natural selection. The similarities of these losses to the regressive evolution of olfactory repertoires in blood-sucking parasites17–19, the optic tectum in cave fish83, and the erosion of plastomes in parasitic plants14 are quite striking.

In contrast to ORs, we did not find convergent reductions in GRs, CPRs, MRJPs, and elongases, suggesting either that these gene families are maintained by natural selection in social parasites, or alternatively, that these genes are in close linkage to neighboring genes maintained by selection. OR genes are usually arranged in homogeneous and well-defined tandem arrays in the genome53. This increases the likelihood of non-allelic recombination due to their inherent repetitiveness84, driving both diversification and reduction of gene families arranged like this. Such a scenario could possibly explain both the general accumulation of OR diversification in free-living ants and the rapid secondary loss in the inquiline species studied here. Intriguingly, the only other gene family where we found substantial reduction in at least one social parasite was the GR family, which also has long tandem arrays in ant genomes.

The dramatic reduction of both ORs and GRs in P. argentina is associated with an extremely modified morphological and behavioral phenotype (Table 1, Fig. 5). It is intriguing that P. argentina is also highly divergent when compared to A. charruanus, which has not nearly the same extent of pronounced reduction in body size, OL, antennal segments, and internal mouthparts41. This implies that P. argentina and A. charruanus have evolved at very different rates, both at the genomic and phenotypic levels. Since the divergence of the two species, the rates of codon sequence evolution and gene loss decreased in A. charruanus, but increased further in P. argentina. In particular, the frequencies of genome rearrangements (transpositions) were increased in the common ancestor of A. charruanus and P. argentina and continued to increase in P. argentina, but was stalled during the later evolution of A. charruanus, so the overall rate of genome evolution came to resemble the rates observed in the free-living, non-parasitic Acromyrmex species. It is possible that A. charruanus quickly reached a stable fitness plateau after gene flow with ancestral Pseudoatta had ceased, or that new hybridization/introgression with its host somehow constrained further genome evolution, but any such considerations remain speculative with the currently available data. Pseudoatta argentina is the only social parasite species in our comparison where secondary host shifts are known because in addition to parasitizing A. heyeri, P. argentina has also been recorded to usurp colonies of A. lundii41,43,58. Consistent with secondary host acquisition, P. argentina has a significantly larger distribution range than A. insinuator and A. charruanus, inhabiting large parts of northern Argentina, Uruguay, and southern Brazil. However, just like the other inquilines, P. argentina is always exceedingly rare and has low population densities, rendering this species as difficult to study as almost all other ant inquilines41.
Methods

Detailed methods and additional information is provided in the Supplementary Information.

Sample collection, sequencing, assembly, and annotation. Colonies of Acromyrmex heyeri and its social parasites Acromyrmex charruanaus and Pseudactena argentina were collected at Plantación Cruz Roja, Uruguay. Acromyrmex insinuator was collected from colonies of its host Acromyrmex echinatior in Gamboa, Panama. All work with live ants was conducted in compliance with ethical regulations for animal testing and research. DNA was isolated from several dozens of ants collected from a single colony for each species using a CTAB-based method. RNA was isolated from pools of 10–30 males for A. insinuator, or pupae, males, and queens, for A. charruanaus and P. argentina, or workers for A. heyeri always using the RNeasy Plus Universal Mini Kit (Qiagen).

De novo genome assemblies for the genomes of A. charruanaus, A. heyeri, A. insinuator, and P. argentina were produced from ~30 Gb (A. insinuator), ~34 Gb (P. argentina), ~35 Gb (A. charruanaus), and ~40 Gb (A. heyeri) short-read sequencing data, using six different insert-size libraries (200, 500, and 800 bp paired-end libraries, and 2, 5, and 10 kb mate-pair libraries) for each genome. After filtering low-quality raw read data, reads were further corrected based on K-mer frequency. First, we indexed each read with 17-mer to build a library of 17-mer frequencies. Second, we corrected bases that were likely wrongly sequenced, based on the concept that bases wrongly sequenced typically yield low frequency k-mers. Initial assemblies were generated with SOAPdenovo based on corrected reads that passed quality control. Gap filling was done for each assembly using paired-end reads of short insert sizes. To generate comparable gene annotations, we (re-)annotated all eleven attine genome analyzed in this study using the same annotation pipeline. For annotating protein-coding genes of the four de novo sequenced genomes (A. charruanaus, A. insinuator, and seven previously published attine genomes (A. echinatior, A. cedrosi, and T. septentrionalis)) and seven previously published attine genomes (A. echinatior, A. cephalotes, A. colombica, C. costatus, M. zetekii, P. cornetzi, P. carinatus, A. insinuator, P. rubens, A. insinuator, and T. septentrionalis), we generated homology-based, de novo and transcriptome-based predictions and combined the results of three methods in gleon v1.0.1 to produce integrated gene sets for all species. Combined predictions from homology, de novo and transcriptome-based methods are summarized in Supplementary Table 21.

Whole protein sequences of 17 species (A. echinatior, A. cephalotes, A. colombica, Camponotus floridanus, C. costatus, Dinoponera quadrispecis, Harpegnathos saltator, Linepithema humile, Monomorium pharaois, M. zetekii, Oecophylla smaragdina, P. carinatus, Pogonomyrnex barbatus, Solenopsis invicta, T. septentrionalis, Vollenhovia emeryi, and Wasmannia auropunctata; see Supplementary Table 22) were used as references to perform homology-based gene predictions. We used augustus v3.2.5 with default parameters to produce de novo predictions with homology-based annotations from A. echinatior as the training set. For transcriptome-based gene predictions, RNAseq data generated from pooled individuals (see above) were aligned to the genome of each corresponding species with TopHat v2.1.0 (options: -r 20 -mat-sdev 10 -1 50000 -solex1.3-quals).

Secondly, we used Cufflinks v2.0.2 (with -l 10000) to predict gene transcripts according to the alignment results from TopHat. We screened genomes for tandem repeats with Tandem Repeats Finder and for transposable elements (TEs) using a combination of LTR_FINDER, PILER, RepeatProteinMask, Wu-BlatX, and RepeatMasker.

Phylogenetic analyses and divergence estimates. We used orthofinder to establish orthology of protein-coding genes and to define a set of single-copy orthologs between all attine species included in our study. We aligned peptide sequences of single-copy orthologs with prank and back-translated to CDS with pal2nal. After removing orthologs showing evidence of recombination (with Phi) and orthologs with putatively saturated sites, we extracted fourfold degenerate sites using custom scripts (4dSite.pl available at https://github.com/scheradler/insulineGenomics). Using a Maximum Parsimony-inferred tree as starting tree (see Supplementary Information), we reconstructed the phylogeny of the 11 species under a Maximum Likelihood framework with RaxML v8.2.12 and under a Bayesian framework using MrBayes v3.2.6 in unpartitioned analyses on the concatenated alignment of fourfold degenerate sites. We used MCMCtree (PAML v4.9 h) for inferring divergence date estimates between the different attine species and calibrated the tree using fossil evidence and estimates of genus splits published previously (see Supplementary Information).

Demographic inference. We used MSMC2 v2.1.1 (available at https://github.com/sterkel/MSMC2) to infer effective population sizes over time for A. echinatior, A. octospinus, and A. insinuator from single individual sequencing data of two individuals of each species collected in Panama (Supplementary Methods). We used short-read (100 bp), small-insert paired-end Illumina sequencing to generate 20× coverage per individual. After quality control with trimmomatic, we mapped the qualified paired-end reads to the A. echinatior reference genome with bwa and removed duplicates with picard. We generated multihaploid and mask files for each species and each individual. We restricted our analysis to larger scaffolds over 2 MB and reduced time segments to 1×2 = 15×1 = 1×2 to avoid overfitting, following recommendations by the developers of MSMC2. We ran 300 bootstrap replications for all four haplotypes per species.

Evolutionary rate analyses. Analyses of signatures of selection were performed on single-copy orthologs (SCOs) across the available leaf-cutting ant genomes (Acromyrmex and Atta). We retrieved 7750 single-copy orthogroups with orthofinder, aligned protein sequences with prank using the inferred tree species (see above) as guide tree, and generated CDS alignments with pal2nal. Orthogroups that showed significant signs of recombination (p <0.01 as tested with Phi) were removed.

We identified genes showing significant evidence for positive selection by running adaptive branch-site tests (absREL, implemented in HYPHY v2.3.14) on CDS alignments and the species tree as input. Similarly, we tested for relaxed selection in protein-coding single-copy orthologs by running RELAX (implemented in HYPHY v2.3.14). RELAX fits three dN/dS rate categories under the alternative model to the reference branch set and computes a selection intensity parameter (see Table 13). We then used short-read (100 bp), small-insert paired-end Illumina sequencing to generate 20× coverage per individual. After quality control with trimmomatic, we mapped the qualified paired-end reads to the A. echinatior reference genome with bwa and removed duplicates with picard. We generated multihaploid and mask files for each species and each individual. We restricted our analysis to larger scaffolds over 2 MB and reduced time segments to 1×2 = 15×1 = 1×2 to avoid overfitting, following recommendations by the developers of MsMC2. We ran 300 bootstrap replications for all four haplotypes per species.

Codon usage bias. We calculated GC content at the first (GC1), second (GC2), and third (GC3) codon positions, as well as the effective number of codons (ENC) for each CDS in single-copy orthologs across species of Acromyrmex and Atta. To do so, we compared ENC–GC3 plots and neutrality plots to assess the extent of codon usage bias in the different species.

Gene family size evolution. We computed gene family clusters between all available genomes from Acromyrmex species and A. colombica using blastp all-vs-all
all searches across predicted proteomes and subsequent MCL clustering (MCL v.14-137). We ran blastp v.2.6.0+ using an e-value cutoff of 1e–5. We excluded non-leaf-cutting ant species to reduce parameter estimation complexity.

Modeling. Following gene family clustering, we excluded gene families that contained proteins identified as TE proteins or TE-derived proteins, based on annotations of all proteomes done with TransprosPv5 v.08222010. Using CAFE v4.0, we modeled gene family size evolution in Acromyrmex host and social parasite genomes based on gene family clusters predicted with MCL. CAFE was run on the dataset comprising the ultrametric species divergence tree, the estimated species-specific error rates, and the gene family clustering. After fully parameterized models repeatedly failed to converge in CAFE, we ran all possible variations of 2-parameter models for each branch-specific $\lambda$ and $\mu$ values. All 2-parameter models were run independently at least eight times to produce robust parameter estimates. The best-scoring $\lambda$ and $\mu$ parameter estimates were clustered by Hartigan-Wong $k$-means clustering ($k = 2$ to $k = 8$), choosing each time 1000 random sets as initial centers (nsart = 1000). The best fit to the data was a model with six discrete rates of $\lambda$ and $\mu$. To confirm that the extreme rates of gene loss inferred for P. argentina were not due to a possible incompleteness of gene annotations, we repeated the analysis of gene family sizes with independently generated homology-based gene annotations (produced with GeMoMa) for all included species (see Supplementary Information). GO enrichment analyses with topGO were run on gene families consistently smaller in the host species compared to their hosts and on gene families consistently larger in the host species compared to the social parasites. The latter analysis did not yield any significantly enriched GO terms. Similarly, no GO terms were enriched in lineage-specific genes in each of the three social parasite species.

Whole-genome alignment and synteny. We generated whole-genome alignments for all ant species with the phylogeny-aware aligner progressiveCactus v0.1 (default settings), using the inferred species tree. We then used halnExtMarizeMutations to summarize inferred mutations at each branch of the underlying phylogeny. We calculated transposition (P), insertion (I), deletion (D), inversion (V), and duplication (U) events per million years of evolution, based on inferred divergence estimates from the phylogeny. We grouped mutation events by size (len = bases) for each branch in the phylogeny and calculated for each mutation type the relative number of events per Ma as the percent increase or decrease compared to the mean number of mutational events across all branches.

Synteny was inferred based on the conserved order of ortholog genes across At. colombica and the five Acromyrmex and Pseudomyrmex species using i-ADHoRe v3.01, using ortholog information between protein-coding genes from orthofinder v2.2.6. We processed the inferred pairwise multiplexes in R, restricting our analyses to longer regions spanning at least 25 genes in each genome. This filtering reduced the number of analyzed syntentic regions from 5425 to 1803. We calculated a Kendall’s rank correlation coefficient ($\tau$) as a quantifier for syntenic conservation for each syntenic region (see details in Supplementary Information), with perfectly syntonic regions having a coefficient of 1. We tested for a significant decrease in syntenic conservation relative to A. colombica comparing the two host genomes (A. echinatior and A. heyeri) to the three inquiline social parasite species (A. insinuator, A. charrusana, and P. argentina) using a one-sided approximate (Monte Carlo) Wilcoxon–Mann–Whitney Test.

Odorant receptor annotation. ORs were annotated in all available ant genomes. We initially retrieved manually curated OR annotations for A. echinatior. At. cephalotes, and Solenopsis invicta published by$^{81}$. Gene annotations for A. echi

natior were again manually refined using unpublished A. echinatior antennal RNAseq data (pers. comm. Bitao Qiu). Most of these manual refinements were adding short first exons that were intentionally left out in the original curated gene set (McKenzie pers. comm.). We used these manually curated annotations to manually annotate OR genes in A. insinuator. Based on the manual curated gene sets from A. echinatior, A. insinuator, and At. cephalotes, we then annotated OR genes in genomes using a gene annotation pipeline that combines evidence from the homology-based gene predictors GeMoMa and exonerate as well as blast hits with EvidenceModeler (see Supplementary Information). For each species, the final set of predicted ORs comprised models coding for an intact 7tm6 domain (predicted with pfam_scan) compared to the social parasites. The latter analysis did not yield any significant enrichment of syntenic families consistently larger in the host species compared to their hosts and on gene families consistently smaller in the hosts

Data availability. Raw sequencing data, genome assemblies, and annotations have been deposited in GenBank with BioProject accession codes PRJNA552756, PRJNA605929, and PRJNA174946 and in the CNGB Sequence Archive with accession codes CNP0000530 and CNP0000605. Further data are available in figshare with identifiers https://doi.org/10.5061/dn/figshare.14224253. Previously published ant genomes were retrieved from GenBank.

Code availability. Scripts and code used for data generation and analyses are archived on Github (https://github.com/schraderL/inquilineGenomics, https://doi.org/10.5281/zenodo.4685040) and documented in the Supplementary Information.

Received: 7 August 2020; Accepted: 6 April 2021; Published online: 18 May 2021

References

1. Hendry, A. P. Eco-evolutionary Dynamics. (Princeton University Press, 2020).
2. Sexton, J. P., Montiel, J., Shay, J. E., Stephens, M. R. & Slater, R. A. Evolution of Ecological Niche Breadth. Annu. Rev. Ecol. Evol. Syst. 48, 183–206 (2017).
3. Foote, A. D. et al. Convergent evolution of the genomes of marine mammals. Nat. Genet. 47, 272–275 (2015).
4. Zhang, G. et al. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science 339, 456–460 (2013).
5. Yin, W. et al. Evolutionary trajectories of snake genes and genomes revealed by comparative analyses of five-pacer vipers. Nat. Commun. 7, 13107 (2016).
6. Slater, R. A., Hirst, M. & Sexton, J. P. Niche breadth predicts geographical range size: a general ecological pattern. Ecol. Lett. 16, 1104–1114 (2013).
7. Hung, C. M., Drovenski, S. V. & Zink, R. M. The roles of ecology, behaviour and effective population size in the evolution of a community. Mol. Ecol. 26, 3775–3784 (2017).
8. Li, S., Jovelín, R., Yoshiga, T., Tanaka, R. & Cutler, A. D. Specialist versus generalist life histories and nucleotide diversity in Caenorhabditis nematodes. Proc. Biol. Sci. 281, 20132858 (2014).
9. Ohta, T. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23, 263–286 (1992).
10. Lahiri, D. C. et al. Selected variation in the wild. Trends Ecol. Evol. 24, 487–496 (2009).
11. Niemiller, M. L., Fitzpatrick, B. M., Shah, P., Schmitz, L. & Near, T. J. Evidence for repeated loss of selective constraint in rhodopsin of amblyopid cavefishes (Teleostei: Amblyopsidae). Evolution 67, 732–748 (2013).
12. Emerling, C. A. Retracted but not gone: patterns of vision gene loss and retention in subterranean mammals. Integr. Comp. Biol. 58, 441–451 (2018).
13. Retaux, S. & Casane, D. Evolution of eye development in the darkness of caves: adaptation, drift, or both? EvoDevo 4, 26 (2013).
14. Wicke, S. et al. Mechanistic model of evolutionary rate variation en route to a nonphototrophic lifestyle in plants. Proc. Natl Acad. Sci. USA 113, 9045–9050 (2016).
15. Gunbang, R. A. Photosynthetic evolution in parasitic plants: insight from the chloroplast genome. BioEssays 26, 235–247 (2004).
16. Chang, E. S. et al. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. Proc. Natl Acad. Sci. USA 112, 14912–14917 (2015).
17. Kirkness, E. F. et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. Proc. Natl Acad. Sci. USA 107, 1268–1273 (2010).
18. Benoit, J. B. et al. Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. Nat. Commun. 7, 10165 (2016).
19. International Glossina Genome Initiative. Genome sequence of the tsetse fly (Glossina morsitans): vector of African trypanosomiasis. Science 344, 380–386 (2014).
20. Tsai, I. J. et al. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496, 57–63 (2013).
21. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. Mol. Biol. Evol. 32, 820–832 (2015).
22. Li, H. et al. Convergent evolution of complex structures for ant-bacterial defensive symbiosis in fungus-farming ants. Proc. Natl Acad. Sci. USA 115, 10707–10712 (2018).
23. Davies, N. B., Bourke, A. F. G. & de L. Brooke, M. Cuckoo parasitoids and parasitic ants: interspecific brood parasitism as an evolutionary arms race. Trends Ecol. Evol. 4, 274–278 (1989).
24. Wilson, E. O. Social modifications related to rarerness in ant species. Evolution 17, 249–253 (1963).
25. Wernegreen, J. J. et al. Population genomics of eusocial insects: the costs of a superorganism. Mol. Biol. Evol. 5, 439–455 (2013).
26. Vogel, A. et al. Footprints of parasitism in the genome of the parasitic flowering plant Cuscuta campestris. Nat. Commun. 9, 2315 (2018).
27. Tsai, I. J. et al. The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496, 57–63 (2013).
28. Romiguier, J. et al. Population genomics of eusocial insects: the costs of a superorganism. Mol. Biol. Evol. 5, 439–455 (2013).
29. Wilson, E. O. Social modifications related to rarerness in ant species. Evolution 17, 249–253 (1963).
77. Sumner, S., Hughes, W. O. H. & Boomsma, J. J. Evidence for differential selection and potential adaptive evolution in the worker caste of an inquiline social parasite. *Behav. Ecol. Sociobiol.* **54**, 256–263 (2003).

78. Sumner, S., Nash, D. R. & Boomsma, J. J. The adaptive significance of inquiline parasitic workers. *Proc. Biol. Sci.* **270**, 1315–1322 (2003).

79. Nehring, V., Dani, F. R., Turillazzi, S., Boomsma, J. J. & d’Ettorre, P. Integration strategies of a leaf-cutting ant social parasite. *Anim. Behav.* **108**, 55–65 (2015).

80. Soares, I. M. et al. Comparative reproductive biology of the social parasite *Acromyrmex untenus* de Souza, Soares & Della Lucia and of its host *Acromyrmex subterraneus subterraneus* Forel (Hymenoptera: Formicidae). *Neotrop. Entomol.* **39**, 714–719 (2010).

81. Gruter, C., Jongepier, E. & Foitzik, S. Insect societies fight back: the evolution of defensive traits against social parasites. *Philos. Trans. R. Soc. Lond. B Biol. Sci*. **373**, 20170200–20170210 (2018).

82. McKenzie, S. K. & Kronauer, D. J. C. The genomic architecture and molecular evolution of ant odorant receptors. *Genome Res.* **28**, 1757–1765 (2018).

83. Loomis, C. et al. An adult brain atlas reveals broad neuroanatomical changes in independently evolved populations of Mexican Cavefish. *Front. Neuroanat.* **13**, 88 (2019).

84. Sumner, S., Hughes, W. O. H. & Boomsma, J. J. Evidence for differential selection and potential adaptive evolution in the worker caste of an inquiline social parasite. *Behav. Ecol. Sociobiol.* **54**, 256–263 (2003).

85. Brandt, M., Foitzik, S., Fischer-Blass, B. & Heinze, J. The coevolutionary dynamics of obligate ant social parasite systems—between prudence and antagonism. *Biol. Rev. Camb. Philos. Soc.* **80**, 251–267 (2005).

86. Gandon, S. Local adaptation and the geometry of host-parasite coevolution. *Ecol. Lett.* **5**, 246–256 (2002).

87. Schiffels, S. & Wang, K. MSMC and MSMC2: the multiple sequentially Markovian Coalescent. *Methods Mol. Biol.* **2090**, 147–166 (2020).

88. Schrader, L. Relaxed selection underlies genome erosion in socially parasitic ant species. <https://doi.org/10.5281/zenodo.4608504> (2021).

**Acknowledgements**

This study was supported by grants from the European Research Council (ERC Advanced grant 325085) (J.J.B.), the US National Science Foundation (DEB-1654829 and CAREER DEB-1943626), the University of Rochester, and Arizona State University (C.R.), as well as the Lundbeck Foundation (R190-2014-2827) (G.Z.). The Visitor’s Office of the Smithsonian Tropical Research Institute provided logistic help and facilities to work in Gamboa and the Autoridad Nacional del Ambiente and the Mar (ANAM) gave permission to sample and export ants from Panama. We gratefully acknowledge the Dirección General de Recursos Naturales Renovables for permission to conduct field research in Uruguay, and Daniel Ramírez from Cambium Forestal for facilitating access to our Uruguayan field research sites. M.B. thanks Leticia Tejera for her discovery of the *P. argentina* population.

**Author contributions**

J.J.B., C.R. and G.Z. (in alphabetic order) conceived of the study, M.B., C.R., J.J.B. and M.S. collected and extracted the samples, H.P., X.B. and Y.D. performed sequencing, assembly, and annotation, L.S. performed the phylogenomic, demographic and evolutionary rate analyses, analyzed gene family size evolution, micro synteny, and genome rearrangements. F.J.L. performed the microCT analyses. L.S. designed the figures. L.S., J.J.B. and C.R. wrote the manuscript with contributions from G.Z. and H.P. All authors read and approved the final manuscript.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-23178-w.

**Correspondence** and requests for materials should be addressed to L.S., J.J.B. or C.R.

**Peer review information** *Nature Communications* thanks Alisandra Denton, Laurent Keller, J. Roman Arguello and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2021, corrected publication 2021