Convergent evolution of a nutritional symbiosis in ants

Raphaella Jackson
Queen Mary University of London  https://orcid.org/0000-0001-5792-1837

David Monnin
Queen Mary University of London

Patapios Alexandros Patapiou
Royal Veterinary College  https://orcid.org/0000-0002-8537-2262

Gemma Golding
Queen Mary University of London

Heikki Helanterä
University of Oulu

Jan Oettler
University of Regensburg

Jürgen Heinze
University of Regensburg  https://orcid.org/0000-0002-9403-5946

Yannick Wurm
Queen Mary University of London  https://orcid.org/0000-0002-3140-2809

Chloe Economou
Queen Mary University of London

Michel Chapuisat
University of Lausanne  https://orcid.org/0000-0001-7207-199X

Lee Henry (✉ l.henry@qmul.ac.uk)
Queen Mary University of London

Article

Keywords: ants, symbioses, nutrient-supplementing microbes, endosymbiotic mutualisms

DOI: https://doi.org/10.21203/rs.3.rs-830142/v1

License: ☕️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Convergent evolution of a nutritional symbiosis in ants

Raphaella Jackson¹, David Monnin¹, Patapios A. Patapiou ¹², Gemma Golding¹, Heikki Helanterä³,
Jan Oettler⁴, Jürgen Heinze⁴, Yannick Wurm¹⁵, Chloe K. Economou¹, Michel Chapuisat⁶, Lee M.
Henry¹*

¹School of Biological and Chemical Sciences, Queen Mary University of London, London, E1 4DQ,
United Kingdom
²Department of Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, AL9 7TA,
United Kingdom
³Ecology and Genetics research unit, University of Oulu, 90014, Finland
⁴Zoology/Evolutionary Biology, University of Regensburg, Regensburg, 93040, Germany
⁵Alan Turing Institute, London NW1 2DB, UK
⁶Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

*Corresponding author. Email: l.henry@qmul.ac.uk

Abstract

Ants are among the most successful organisms on earth. It has been suggested that forming symbioses
with nutrient-supplementing microbes may have contributed to their success, by allowing ants to
invade otherwise inaccessible niches. However, it is unclear whether ants have repeatedly evolved
symbioses to overcome the same nutrient limitations. Here, we address this question by comparing the
independently evolved symbioses in Camponotus, Cardiocondyla, Formica and Plagiolepis ants. Our
analysis reveals the only metabolic function consistently retained in all of the symbiont genomes is
the capacity to synthesise tyrosine, which is essential for insect cuticles. We also reveal that in certain
multi-queen lineages, only a fraction of queens carry the symbiont, suggesting ants differ in their
colony-level reliance on symbiont-derived nutrients. Our results suggest symbioses can arise to solve
common problems, but hosts may differ in their dependence on symbionts, highlighting the
evolutionary forces influencing the persistence of long-term endosymbiotic mutualisms.
Main Text

Introduction

Ants are among the most ecologically dominant organisms in terrestrial ecosystems, and part of their success lies in their ability to occupy a wide range of habitats. It has been suggested that acquiring nutrient-provisioning symbionts may have allowed certain ant lineages to survive in nutrient imbalanced habitats. For example, gut-associated symbionts are thought to have enabled transitions to arboreal lifestyles in several ant lineages by relaxing their need for nitrogen and allowing them to feed predominantly on plant-derived resources, such as extrafloral nectaries and insect honeydew. Symbiont acquisitions may therefore represent key adaptations that have allowed ants to significantly expand the ecological niches in which they can forage and complete development, thereby contributing to their widespread ecological success.

At least four ant lineages have evolved symbioses where the microbes are housed within specialised cells called bacteriocytes that surround the midgut. Bacteriocytes are a common feature of ancient nutritional symbioses, where symbionts are essential for host development and strictly vertically transmitted through host generations. The most well-studied of these in ants is *Blochmannia*, the obligate symbiont of carpenter ants (*Camponotus*). *Blochmannia* provides its host with essential amino acids that can improve brood production, especially when proteins are scarce. *Blochmannia* is also thought to aid hosts in nitrogen recycling and synthesises the aromatic amino acid tyrosine, which is an important component of insect cuticles. The symbiont of *Cardiocondyla obscurior*, *Candidatus* Westeberhardia cardiocondylae, hereafter *Westeberhardia*, despite having a highly reduced genome, has also retained the capacity to synthesise tyrosine through a shared metabolic pathway with its ant host. Two additional ant genera, *Formica* and *Plagiolepis*, are also known to harbour symbionts within bacteriocytes surrounding the midgut, suggesting they also play a role in provisioning nutrients for their hosts. However, the functional role of the symbionts in *Formica* and *Plagiolepis* is currently unknown.

While the acquisition of nutrient provisioning symbionts has repeatedly allowed insects to invade nutrient imbalanced niches, such as plant sap and blood feeding, it is less clear why these
relationships evolve in predominantly omnivorous insects such as ants. In particular, it is unclear whether ants have repeatedly evolved symbioses to overcome the same vital nutrient limitations. This has limited our understanding of the metabolic challenges facing omnivorous insects, and how nutritional symbioses evolve to overcome them.

The aim of this study is to determine whether the four bacteriocyte-associated symbioses in ants represent ancient nutritional mutualisms that have evolved to serve similar functions for their hosts. We first characterise the genomes of the symbionts in Formica and Plagiolepis, and several new strains of Westeberhardia from phylogenetically divergent Cardiocondyla lineages. Using a comparative approach, we then ask whether the symbionts from all four ant lineages have retained metabolic pathways in their highly reduced genomes that suggest they serve similar nutrient-provisioning roles for their hosts. We then investigate the phylogenetic and intracolony distributions of symbionts in diverse Formica and Cardiocondyla species to determine the origins of each symbiosis and its prevalence across species and castes. This survey reveals that in many ant lineages that maintain multi-queen (polygynous) colonies, only a fraction of queens carry the symbiont, suggesting species differ in their dependence on symbiont-derived nutrients at the colony level. We present evidence that suggests species differences in symbiont retention are not correlated with changes in symbiont functionality and discuss how ant feeding ecology and sociality can impact dependence on nutritional symbioses.

Results and discussion

Genome characteristics of ancient obligate symbionts

We first tested the hypothesis that each of the four ant lineages hosts its own ancient strictly vertically transmitted symbionts that have co-speciated with its host. To address this aim, we compared the genomes of symbionts from 13 species of ants, representing four independently evolved symbioses. This includes symbionts from three Formica, two Plagiolepis, and an additional three Cardiocondyla species that we sequenced, in addition to four previously published genomes from Blochmannia, the obligate symbiont of Camponotus ants, and the one pre-existing Westeberhardia genome from Cardiocondyla obscurior.6,10–13
We found the gene order of single copy orthologs in symbionts is perfectly conserved in ant species belonging to the same genus (Fig. 1). This type of extreme structural stability of genomes only occurs in symbionts that have been strictly vertically transmitted within a matriline and has been documented in the obligate symbionts of whiteflies, psyllids, cockroaches and aphids. In contrast, genome structure differed substantially between symbionts from different ant genera (Fig. 1, Fig S1). In *Formica* and *Cardiocondyla* ant species, we also find that the host and symbiont phylogenies are in general concordance (Fig S2). This strongly suggests the symbioses in each of the four ant lineages are independently acquired ancient associations that have co-diversified with their hosts.

Our phylogenetic analysis demonstrates the four symbiont lineages have distinct phylogenetic origins (Fig. 2). *Formica* and *Plagiolepis* ants each harbour a *Sodalis*-allied symbiont, whereas *Westeberhardia* and *Blochmannia* belong to the large Enterobacteriaceae family. All of the symbionts have evidence of advanced genome reduction, which is characterized by reduced genome size, GC content, and number of coding sequences similar to other ancient obligate symbionts of insects. The three strains of *Westeberhardia* we analysed have extremely small (0.45-0.53Mb) GC depleted genomes (22-26%) that were similar to the figures reported for the strain in *Cardiocondyla obscurior*; confirming it has one of the smallest genomes of any known Gammaproteobacterial endosymbiont (Fig. 2). By comparison, the *Sodalis*-allied symbionts have genomes around twice the size (1.37-1.38Mb) and GC content (~41%) of that of *Westeberhardia* (Fig. 2) suggesting they are in an earlier stage of genome reduction. The *Formica* and *Plagiolepis* symbionts have a similar size, GC range and number of coding sequences as known obligate symbionts such as *Candidatus Doolittlea endobia* and several *Serratia symbiotica* lineages that are co-obligate symbionts in aphids. The degree of genome reduction in *Westeberhardia* and *Blochmannia* suggest that they are older associations than the *Sodalis* symbionts found in *Plagiolepis* and *Formica* ants.

**Bacteriocyte-associated endosymbionts**

Using fluorescent in situ hybridisation, we determine whether the *Sodalis*-allied symbionts we sequenced are localised in bacteriocytes to confirm they are the associations first observed by Lillienstern and Jungen in the early 1990’s.
Consistent with Lilienstern’s findings, we found the *Sodalis* symbiont in *Formica* ants is distributed in bacteriocytes surrounding the midgut in adult queens (Fig. 3A). The symbionts are also found in eggs and ovaries of adult queens, indicating they are vertically transmitted from queens to offspring (Fig. 3B-C). Sectioning of *F. cinerea* larvae shows the bacteriocytes to be arranged in a single layer of cells surrounding the midgut, as well as in clusters of bacteriocytes closely situated to the midgut (Fig. 3D-D’). In adult *Plagiolepis* queens, no bacteriocytes were found around the midgut, suggesting the symbionts migrate to the ovaries prior to or during metamorphosis. Apart from that, the localisation of the symbiont in *Plagiolepis* was the same as in *Formica* – symbionts in larval midgut bacteriocytes, ovaries and eggs (Fig. S3) – supporting Jungen’s cytological findings (10). Bacteriocytes are also found surrounding the midgut in *Camponotus* and *Cardiocondyla* ants, indicating the symbionts are localised in a similar manner in all four ant lineages.

**Conservation of metabolic functions in ant endosymbionts**

Despite on-going genome reduction, obligate symbionts of insects typically retain gene networks required for maintaining the symbiosis with their host, such as pathways for synthesising essential nutrients. This has resulted in the symbionts of sap- and blood-feeding insects converging on genomes that have retained the same sets of metabolic pathways – to synthesis essential nutrients missing in their hosts’ diets. Here we test the hypothesis that the four bacteriocyte-associated symbionts of ants have been acquired to perform similar functions. For this, we assess whether they have consistently retained metabolic pathways to synthesis the same key nutrients. Two major patterns stand out.

First, we find that the four divergent ant symbionts have all retained the shikimate pathway, which produces chorismate, along with most of the steps necessary to produce tyrosine from this precursor (Table 1 and Table S2). Both the symbiont of *Formica* and *Westeberhardia* lack one of the genes required to produce tyrosine. However in *Westeberhardia* it is believed the host provides the gene to complete the final step of the pathway, and we find this gene is also present in the *Formica* ant genomes (Fig S4). In addition, all symbionts except *Westeberhardia* can produce phenylalanine which is a precursor that can be converted to tyrosine by their hosts. Tyrosine is important for
insect development as it is used to produce L-DOPA, which is a key component of insect cuticles. In carpenter ants, weevils and grain beetles, removal or inhibition of their symbionts, which are thought to provision hosts with tyrosine, causes cuticle development to suffer. Recent evidence from turtle ants (Cephalotes) suggest obligate gut microbes also assist in cuticle development through the production of tyrosine and phenylalanine alongside other compounds. This suggests tyrosine provisioning by symbionts can play a crucial role in cuticle formation across diverse ant lineages. Tyrosine provisioning is also the likely function of Westeberhardia in Cardiocondyla ants, as this is one of the few nutrient pathways retained in this symbiont. Our analysis confirms the shikimate pathway, and the symbiont portions of the tyrosine pathway, have been retained in Westeberhardia from three phylogenetically diverse Cardiocondyla lineages, providing additional support for this hypothesis. In addition to tyrosine, most of the symbionts have retained the capacity to produce vitamin B9 (tetrahydrofolate) and all can perform the single step conversions necessary to produce alanine and glycine. However, our gene enrichment analysis indicates that tyrosine, and the associated chorismate biosynthetic process, are the only enriched pathways associated with nutrient provisioning that are shared by all of the symbiont genomes (Table S1). This suggests that provisioning of tyrosine, or tyrosine precursors, is of general importance across all bacteriocyte-associated symbioses of ants.

Second, our comparative analysis revealed clear differences in the pathways lost or retained across symbionts (Table 1 and Table S2). This is most evident when comparing Blochmannia with Westeberhardia, the latter of which has lost the capacity to synthesise most essential nutrients. The symbionts of Formica or Plagiolepis, in contrast, have retained the capacity to synthesise many of the same amino acids and B vitamins as Blochmannia, suggesting they may perform similar functions for their hosts. However, Blochmannia has retained more biosynthetic pathways, particularly those involved in the synthesis of essential amino acids. Experimental studies have confirmed that Blochmannia provisions hosts with essential amino acids. The absence of several core essential amino acids in the Formica and Plagiolepis symbionts may reflect differences in the dietary ecology of the different ant genera, although this would require experimental validation. The retention of the full complement of essential amino acids in the highly reduced genome of Blochmannia does however
indicate it plays a more substantive nutrient-provisioning role for its hosts than the other ant
symbionts we investigated.

Previous work on the extracellular gut symbionts of several arboreal ant lineages identified
nitrogen recycling via the urease operon as a function that may be of key importance for ant
symbioses\textsuperscript{1,2,31,32}. However, we do not find any evidence that the symbionts of \textit{Formica}, \textit{Plagiolepis},
or \textit{Cardiocondyla} play a role in nitrogen recycling via the urease operon (Table 1). This suggests
nitrogen recycling may play an important role for more strictly herbivorous ants, such as \textit{Cephalotes}.

Our results, however, indicate tyrosine may be universally required for cuticle synthesis across a
broader range of ant lineages.

The origins and losses of symbioses in \textit{Formica} and \textit{Cardiocondyla}

We investigated the presence of the symbiont in phylogenetically diverse \textit{Formica} and \textit{Cardiocondyla}
species to identify the evolutionary origins and losses of the symbiosis. Although the symbiont in
\textit{Plagiolepis} was present in \textit{P. pygmaea} and two unknown \textit{Plagiolepis} species we investigated, we did
not have sufficient phylogenetic sampling to assess the origins of the symbiosis.

In \textit{Formica}, we find the symbiont is restricted to a single clade in the paraphyletic
Serviformica group (Figure 4A). The species in this clade are socially polymorphic, forming both
multi-queen and single-queen colonies\textsuperscript{33}. Based on a previously dated phylogeny of \textit{Formica} ants, we
estimate the symbiosis originated approximately 12-22 Million years ago\textsuperscript{34}. In \textit{Cardiocondyla}, the
symbiosis is widespread throughout the genus. The prevalence of the symbiont in \textit{Cardiocondyla}, in
combination with its highly reduced genome, suggests it is a very old association that likely dates
back to the origins of the ant genus some 50-75 Million years ago\textsuperscript{35}. However, the symbiont was
absent in two clades, the argentea and palearctic groups (Figure 4B). This may represent true
evolutionary losses in these clades. It is tempting to speculate that these losses are linked to a notable
change in social structure in these two \textit{Cardiocondyla} clades, having gone from the ancestral state of
maintaining multi-queen colonies to single-queen colonies\textsuperscript{36}, however it is not clear how this could
impact the symbiosis.
Evidence of variation in colony-level dependence on symbionts

Observations from individual studies on *F. cinerea* and *F. lemani*⁸,⁹, as well as *Cardiocondyla obscurior*⁶, reported rare cases of ant queens not harbouring their symbionts in nature. This called into question the degree these insects depend on symbionts for nutrients, and whether the symbiosis may be breaking down in certain host lineages. However, given the limited number of species and populations studied, it is unclear how often colonies are maintained with uninfected queens in nature, and whether this differs across species, suggesting species may differ in their dependence on their symbiont. To answer this question, we assessed the presence of the symbionts in 838 samples from 147 colonies of phylogenetically diverse *Formica* and *Cardiocondyla* species collected across 8 countries.

Our investigation reveals the natural occurrence of uninfected queens is a widespread phenomenon in many *Formica* and *Cardiocondyla* species (Figure 4). We confirmed the absence of symbionts in queens, and that they have not been replaced with another bacterial or fungal symbiont, using diagnostic PCR, whole genome and deep-coverage amplicon sequencing (Table S3, Table S4, Fig S5 & S6). There was also clear evidence of variation across host species. In *Formica*, queens and workers of *F. fusca* always carried the symbiont, whereas queens and workers of *F. lemani, F. cinerea*, and *F. selysi* showed varying degrees of individuals not carrying the symbionts (Figure 4A, Table S5). A similar pattern can be seen in *Cardiocondyla*, where queens of several species, such as *C. obscurior*, always carry the symbiont, compared to lower incidences in other species (Figure 4B). Klein et al⁶ identified a single *C. obscurior* colony with uninfected queens in Japan, however, queens of this species nearly always carry the symbiont in nature.

The degradation and eventual loss of symbionts from bacteriocytes has been reported in males, and in sterile castes of aphids and ants ³⁷, which do not transmit symbionts to offspring. In reproductive females, bacteriocytes may degrade as a female ages; however, symbionts are typically retained at high bacterial loads in the ovaries, as this is required to maintain the symbionts within the germline ²². It is of note that all of the symbiotic ant species we investigated maintain multi-queen colonies, and the vast majority had at least one queen, often more, within a colony that carried the symbiont (Table S5). We hypothesize that species that maintain colonies with uninfected queens may
be able to retain sufficient colony-level fitness with only a fraction of queens harbouring the symbiont and receiving its nutritive benefits.

Dependence on symbionts in a socioecological context

The retention of symbionts in queens and workers of some species, but not others, suggests species either differ in their dependence on symbiont-derived nutrients, or that symbionts have lost the capacity to make nutrients in certain host lineages. Our analysis of symbiont genomes did not reveal any structural differences, such as the disruption of metabolic pathways, which could explain differences in symbiont retention between host species (Table S2). This suggests differences in the retention of symbionts may reflect differences in host ecologies.

In ants, which occupy a wide range of feeding niches, reliance on symbiont-derived nutrients will largely depend on lineage-specific feeding ecologies. For example, Camponotus ants have been shown to be predominantly herbivores. Blochmannia, in turn, has retained the capacity to synthesise key nutrients missing in plant-based diets, such as essential amino acids. Blochmannia is also always present in queens and workers \( ^{22} \), which is a testament to the importance of these nutrients for the survival of its primarily herbivorous host. In contrast, Formica and Cardiocondyla species are largely thought to be omnivores \( ^{38} \). Diet flexibility and altered foraging efforts may therefore reduce their reliance on a limited number of symbiont-derived nutrients allowing colonies of some species to persist with uninfected queens in some contexts. Silvanid beetles and grain weevils, for example, can survive in the absence of their tyrosine-provisioning symbionts \( ^{27,39,40} \) when provided nutritionally balanced diets, such in the laboratory \( ^{39} \) or in cereal grain elevators \( ^{41,42} \). Similarly, studies on Cardiocondyla and Camponotus have shown they can maintain sufficient colony health in the absence of their symbionts, if provided a balanced diet \( ^{22,43} \). It would be interesting to know whether species of Formica and Cardiocondyla that always carry the symbiont in nature, such as \( F. \) fusca and \( C. \) obscurior, have more restricted diets with less access to nutrients such as tyrosine, as this may explain their dependence on their symbiont for nutrients and tendency to harbour them in queens.

Although it is unusual for bacteriocyte-associated symbionts to be absent in reproductive females, the fact that it is simultaneous occurring in phylogenetically diverse species from many
locations suggests the symbiosis may have persisted in this manner over evolutionary time. Perhaps through diet flexibility colonies can be maintained with uninfected queens in some contexts, however we expect them to be disadvantaged in other ecological scenarios. Fluctuating environmental conditions may therefore eventually purge asymbiotic queens from lineages, allowing the symbiosis to be retained over longer periods of evolutionary time. The multiple-queen colony lifestyle present in all symbiotic *Formica* and *Cardiocondyla* species we investigated may provide an additional social buffer that limits the costs to individual queens being asymbiotic. Workers will still nourish larvae and queens without symbionts and colony fitness may be maintained through the reproductive output of nestmate queens that carry the symbiont. These two factors, diet flexibility and multi-queen systems, may result in prolonged persistence of asymbiotic individuals, which allows us to detect them, while not ultimately preventing long-term maintenance of the symbiosis.

Our data suggests that symbiotic relationships can evolve to solve common problems but also rapidly breakdown if the symbiosis is no longer required. We have identified tyrosine provisioning as a unifying function across bacteriocyte-associated symbionts of ants. But we have also shown species can vary in how much they depend on symbionts for nutrients. Our results demonstrate that ants have a unique labile symbiotic system, allowing us to better understand the evolutionary forces that influence the persistence and breakdown of long-term endosymbiotic mutualisms.

*Candidatus* Hugann liliensternia and *Candidatus* Jungenella plagiensis

We propose the names *Candidatus* Hugann liliensternia for the *Sodalis*-allied symbiont found in *Formica*. The genus name is derived from the combined first names of the first authors parents, and the species name is in honour of Margarete Lilienstern who first identified the symbiont. Similarly, we propose the name of *Candidatus* Jungenella plagiensis for the *Plagiolepis*-bound symbiont. The genus name is in honor of Hans Jungen who originally discovered the symbiont, and the species name is derived from *Plagiolepis*, the genus in which the symbiont can be found.

Materials and Methods
Detailed protocols for each of the following sections are available in the supplementary materials, under supplementary methods.

**Whole Genome Sequencing and Analysis**

Queens from 3 *Formica* species (*fusca, lemani, cinerea*), 2 *Plagiolepis* species (*pygmaea*, spp.), and 3 *Cardiocondyla* species (*minutior, mauritanica, wroughtonii*) were sequenced using the Illumina Hiseq 4000. Raw reads were trimmed, filtered, and assembled using SPAdes V3.11.1. Genomes were then annotated using Prokka V1.14.6. Pathway completeness was assessed using manual curation and the metacyc resources for *E. coli* str. K-12. Single copy orthologs were identified using Orthofinder V2.2.7. Enriched functional categories and pathways were identified using David.

**Taxonomic Analysis**

The phylogeny of ant genera used in Figure 1 is based on with additional tip placements based on. The phylogeny of endosymbiont species displayed in Figure 2 was created first using GtoTree to generate the gene alignment making use of GtoTree’s standard 172 gene set defined for gammaproteobacteria. The phylogeny was generated using a partitioned analysis in Raxml v8.2.11 with 100 rapid bootstrap inferences followed by a ML search. Detailed description of methods and phylogenies based on alternate gene sets available in supplementary methods (Fig S7).

The phylogeny of ant species used in Figure 4 is based on with additional tip placements based on a phylogeny of cytochrome B sequences based on the work of including sequences from individuals we had sequenced for *Formica* (Fig S8) and with additional tip placements based on for *Cardiocondyla*.

**FISH Microscopy**

FISH was performed on eggs, queen guts, queen ovaries (whole mount) and on larvae (cytological sections), using 16S rRNA oligonucleotide probes specifically targeting the symbionts. Samples were mounted using Vectashield hardset antifade mounting media with DAPI and visualised using a Leica DMRA2 epi-fluorescent microscope.
Symbiont Screening Procedure

We screened 838 individuals, a mixture of queens and workers, from 147 colonies across 29 species for the presence of symbionts using a combination of diagnostic PCR screening and Illumina 16S deep coverage sequencing (Table S3). Diagnostic PCRs were carried out by amplifying the symbiont 16S rRNA genes from total genomic DNA extracted from individual ants. Custom primer pairs (Table S5) were used for screening *Sodalis* and *Westeberhardia*, respectively. Positive queen diagnostic PCR results were confirmed using Sanger sequencing.

For Illumina 16S deep coverage sequencing, the 515F/806R primer pair was used to amplify the V4 region of the 16S rRNA gene in two runs of 16S sequencing in 177 *Cardiocondyla* and *Formica* samples (Table S3). Additionally we conducted a run of ITS fungal sequencing using the ITS5/5.8S_fungi primer pair (Table S3), to investigate whether any fungal symbiont replacement could be detected.

16S sequencing data was analysed using Mothur v.1.41.3, to cluster reads into OTUs were clustered at 99% similarity. ITS sequencing data was analysed using USEARCH and UPARSE to cluster reads in zero-radius OTUs (ZOTUs). Data was then processed using R to remove OTUs/ZOTUs at below 1 percent relative abundance in a sample and generate visualizations.

Data Availability

All data collected in association with this paper, alongside associated genome assemblies, are available under BioProject accession PRJNA639935.

References

1. Feldhaar, H. *et al.* Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. *BMC Biol.* 5, 48 (2007).

2. Russell, J. A. *et al.* Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21236–21241 (2009).
3. Sanders, J. G. et al. Dramatic differences in gut bacterial densities correlate with diet and habitat in rainforest ants. *Integr. Comp. Biol.* **57**, 705–722 (2017).

4. Moran, N. A., McCutcheon, J. P. & Nakabachi, A. Genomics and Evolution of Heritable Bacterial Symbionts. *Annu. Rev. Genet.* **42**, 165–190 (2008).

5. Hopkins, T. L. & Kramer, K. J. Insect Cuticle Sclerotization. *Annu. Rev. Entomol.* **37**, 273–302 (1992).

6. Klein, A. et al. A novel intracellular mutualistic bacterium in the invasive ant Cardiocondyla obscurior. *ISME J.* **10**, 376–388 (2016).

7. Buchner, P. Endosymbiosis of Animals with Plant Microorganisms. *Intersci. Publ.* 909 (1965).

doi:10.2307/3757184

8. Jungen, H. Endosymbionten bei Ameisen. *Insectes Soc.* **15**, 227–232 (1968).

9. Lilienstern, M. Beiträge zur Bakteriensymbiose der Ameisen. (1932).

10. Degnan, P. H. Genome sequence of Blochmannia pennsylvanicus indicates parallel evolutionary trends among bacterial mutualists of insects. *Genome Res.* **15**, 1023–1033 (2005).

11. Gil, R. et al. The genome sequence of Blochmannia floridanus: Comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci.* **100**, 9388–9393 (2003).

12. Williams, L. E. & Wernegreen, J. J. Unprecedented loss of ammonia assimilation capability in a urease-encoding bacterial mutualist. *BMC Genomics* **11**, 687 (2010).

13. Williams, L. E. & Wernegreen, J. J. Sequence Context of Indel Mutations and Their Effect on Protein Evolution in a Bacterial Endosymbiont. *Genome Biol. Evol.* **5**, 599–605 (2013).

14. McCutcheon, J. P. & von Dohlen, C. D. An Interdependent Metabolic Patchwork in the Nested Symbiosis of Mealybugs. *Curr. Biol.* **21**, 1366–1372 (2011).

15. Patiño-Navarrete, R., Moya, A., Latorre, A. & Peretó, J. Comparative Genomics of Blattabacterium cuenoti: The Frozen Legacy of an Ancient Endosymbiont Genome. *Genome Biol. Evol.* **5**, 351–361 (2013).

16. Sloan, D. B. & Moran, N. A. Genome Reduction and Co-evolution between the Primary and Secondary Bacterial Symbionts of Psyllids. *Mol. Biol. Evol.* **29**, 3781–3792 (2012).

17. Santos-Garcia, D., Mestre-Rincon, N., Ouvrard, D., Zchori-Fein, E. & Morin, S. Portiera Gets
Wild: Genome Instability Provides Insights into the Evolution of Both Whiteflies and Their Endosymbionts. *Genome Biol. Evol.* **12**, 2107–2124 (2020).

18. Tamas, I. 50 Million Years of Genomic Stasis in Endosymbiotic Bacteria. *Science (80-. ).* **296**, 2376–2379 (2002).

19. Husnik, F. & McCutcheon, J. P. Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc. Natl. Acad. Sci.* **113**, E5416–E5424 (2016).

20. Monnin, D. *et al.* Parallel Evolution in the Integration of a Co-obligate Aphid Symbiosis. *Curr. Biol.* **30**, 1949-1957.e6 (2020).

21. Blochmann, F. Über das Vorkommen bakterienähnlicher Gebilde in den Geweben und Eiern verschiedener Insekten. *Zentbl. Bakteriol.* **11**, 234–249 (1887).

22. Sauer, C., Dudaczek, D., Holldobler, B. & Gross, R. Tissue Localization of the Endosymbiotic Bacterium ‘Candidatus Blochmannia floridanus’ in Adults and Larvae of the Carpenter Ant Camponotus floridanus. *Appl. Environ. Microbiol.* **68**, 4187–4193 (2002).

23. Moran, N. A., Plague, G. R., Sandström, J. P. & Wilcox, J. L. A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 14543–14548 (2003).

24. Manzano-Marín, A., Oceguera-Figueroa, A., Latorre, A., Jiménez-García, L. F. & Moya, A. Solving a bloody mess: B-vitamin independentmetabolic convergence among gammaproteobacterial obligate endosymbionts from blood-feeding arthropods and the Leech haementeria officinalis. *Genome Biol. Evol.* **7**, 2871–2884 (2015).

25. Flydal, M. I. & Martinez, A. Phenylalanine hydroxylase: Function, structure, and regulation. *IUBMB Life* **65**, 341–349 (2013).

26. Simonet, P. *et al.* Disruption of phenylalanine hydroxylase reduces adult lifespan and fecundity, and impairs embryonic development in parthenogenetic pea aphids. *Sci. Rep.* **6**, 34321 (2016).

27. Anbutsu, H. *et al.* Small genome symbiont underlies cuticle hardness in beetles. *Proc. Natl. Acad. Sci.* **114**, E8382–E8391 (2017).

28. José de Souza, D., Devers, S. & Lenoir, A. Blochmannia endosymbionts and their host, the ant
Camponotus fellah: Cuticular hydrocarbons and melanization. *C. R. Biol.* **334**, 737–741 (2011).

29. Kiefer, J. S. T. *et al.* Inhibition of a nutritional endosymbiont by glyphosate abolishes mutualistic benefit on cuticle synthesis in *Oryzaephilus surinamensis*. *Commun. Biol.* **4**, (2021).

30. Duplais, C. *et al.* Gut bacteria are essential for normal cuticle development in herbivorous turtle ants. *Nat. Commun.* **12**, 1–6 (2021).

31. Bisch, G. *et al.* Genome Evolution of Bartonellaceae Symbionts of Ants at the Opposite Ends of the Trophic Scale. *Genome Biol. Evol.* **10**, 1687–1704 (2018).

32. Rubin, B. E. R., Kautz, S., Wray, B. D. & Moreau, C. S. Dietary specialization in mutualistic acacia-ants affects relative abundance but not identity of host-associated bacteria. *Mol. Ecol.* **28**, 900–916 (2019).

33. Brelsford, A. *et al.* An Ancient and Eroded Social Supergene Is Widespread across *Formica* Ants. *Curr. Biol.* **30**, 304-311.e4 (2020).

34. Borowiec, M. L., Cover, S. P. & Rabeling, C. The evolution of social parasitism in *Formica* ants revealed by a global phylogeny. *bioRxiv* (2020). doi:10.1101/2020.12.17.423324

35. Moreau, C. S. Phylogeny of the Ants: Diversification in the Age of Angiosperms. *Science* (80-. ). **312**, 101–104 (2006).

36. Heinze, J. Life-history evolution in ants: the case of Cardiocondyla. *Proc. R. Soc. B Biol. Sci.* **284**, 20161406 (2017).

37. Fukatsu, T. & Ishikawa, H. Soldier and male of an eusocial aphid *Colophina arma* lack endosymbiont: Implications for physiological and evolutionary interaction between host and symbiont. *J. Insect Physiol.* **38**, 1033–1042 (1992).

38. Seifert, B. *The ants of Central and North Europ.* (Lutra, 2018).

39. Hirota, B. *et al.* A novel, extremely elongated, and endocellular bacterial symbiont supports cuticle formation of a grain pest beetle. *MBio* **8**, (2017).

40. Kuriwada, T. *et al.* Biological role of *Nardonella* endosymbiont in its weevil host. *PLoS One* **5**, 1–7 (2010).
418  41. Huger, A. Experimentelle untersuchungen über die künstliche symbiontenelimination bei
419  verratsschädlingen: Rhizopertha Dominica F. (bostrychidae) und Oryzaephilus Surinamensis
420  L. (Cucujidae). Zeitschrift für Morphol. und Ökologie der Tiere 44, 626–701 (1956).
421  42. Mansour, K. On the microorganism-free and the infected Clandra granaria L. Bull. Soc. roy.
422  entomol. Egypt. 19, 290–306 (1935).
423  43. Ün, Ç. et al. Cytoplasmic incompatibility between Old and New World populations of a tramp
424  ant. Evolution (N. Y). 1775–1791 (2021). doi:10.1111/evo.14261
425  44. Nurk, S. et al. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. in
426  Research in Computational Molecular Biology (eds. Deng, M., Jiang, R., Sun, F. & Zhang, X.)
427  158–170 (Springer Berlin Heidelberg, 2013).
428  45. Seemann, T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069
429  (2014).
430  46. Caspi, R. et al. The MetaCyc Database of metabolic pathways and enzymes and the BioCyc
431  collection of pathway/genome databases. Nucleic Acids Res. 36, 623–631 (2008).
432  47. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome
433  comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16, 1–14
434  (2015).
435  48. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths
436  toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13
437  (2009).
438  49. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large
439  gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57 (2009).
440  50. Ward, P. S., Blaimer, B. B. & Fisher, B. L. A revised phylogenetic classification of the ant
441  subfamily Formicinae (Hymenoptera: Formicidae), with resurrection of the genera Colobopsis
442  and Dinomyrmex. Zootaxa 4072, 343–357 (2016).
443  51. Lee, M. D. GToTree: A user-friendly workflow for phylogenomics. Bioinformatics 35, 4162–
444  4164 (2019).
445  52. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large
phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

53. Romiguier, J., Rolland, J., Morandin, C. & Keller, L. Phylogenomics of palearctic Formica species suggests a single origin of temporary parasitism and gives insights to the evolutionary pathway toward slave-making behaviour. *BMC Evol. Biol.* **18**, 1–8 (2018).

54. Goropashnaya, A. V., Fedorov, V. B., Seifert, B. & Pamilo, P. Phylogenetic relationships of Palearctic Formica species (hymenoptera, Formicidae) based on mitochondrial cytchrome b sequences. *PLoS One* **7**, (2012).

55. Oettler, J., Suefuji, M. & Heinze, J. The evolution of alternative reproductive tactics in male cardiocondyla ants. *Evolution (N. Y.)* **64**, 3310–3317 (2010).

56. Heinze, J., Seifert, B. & Zieschank, V. Massacre of the innocents in Cardiocondyla thoracica: Manipulation by adult ant males incites workers to kill their immature rivals. *Entomol. Sci.* **19**, 239–244 (2016).

57. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4516–4522 (2011).

58. Epp, L. S. *et al.* New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Mol. Ecol.* **21**, 1821–1833 (2012).

59. P. D. Schloss *et al.* Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).

60. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

61. Edgar, R. C. **UPARSE**: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–8 (2013).

62. King, T., Butcher, S. & Zalewski, L. Apocrita - High Performance Computing Cluster for Queen Mary University of London. *Queen Mary Univ. London, Tech. Rep.* 1–2 (2017).

63. Rafiqi, A. M., Rajakumar, A. & Abouheif, E. Origin and elaboration of a major evolutionary transition in individuality. *Nature* (2020). doi:10.1038/s41586-020-2653-6
Acknowledgments

The authors thank Sabine Frohschammer, Sylvia Cremer, Tina Wanke, Masaki Suefuji, D. Ortius, K. Yamauchi, and Dominic Burns for providing ant samples. This research utilised Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT. This project was funded by L.M.H.’s NERC IRF (NE/M018016/1), and Marie Curie (H2020-MSCA-IF- 2017-796778-SYMOBLIGA).

Author Contributions:

Designed research: R.J., D.M., L.M.H.

Performed research: R.J., D.M., P.A.P., G.G., C.K.E.

Contributed new reagents/analytic tools/samples: H.H., J.O., J.H., Y.W., M.C.

Analysed data: R.J., D.M.

Wrote the paper: R.J., D.M., L.M.H.

Competing Interest Statement: We declare no conflict of interest.
Figure 1: (A) Ant lineages known to host bacteriocyte-associated symbionts (red font) and lineages not known to (black font), based on 63. Phylogeny root (grey font) not examined for symbionts. (B) Visualisation of symbiont genomes showing near perfect conservation of gene order in the symbionts of ant species that belong to the same genus. All genomes and annotations were generated in this study except the *Blochmannia* symbionts and the *Westeberhardia* strain from *C. obscurior* 6,10–13. *Evidence of symbionts were detected in embryos of *Anoplolepis* 63 but it is unclear if they are localised in bacteriocytes in larvae and adults.
Figure 2: A phylogeny of gammaproteobacterial endosymbionts using a partitioned analysis of 172 genes in Raxml rooted to *Xanthomonas axonopodis*. Node labels reflect support based on 100 rapid bootstraps. Bar plots represent the size (in Mbp) and GC content of symbiont genomes. Bars are colour coded to represent hypothesised relationship between symbiont and host. Species names in red in the phylogeny indicate the four bacteriocyte-associated symbionts of ants. Genomes sequenced and assembled for this paper are referenced to as ‘novel symbiont’ lineages.
Figure 3. Fluorescent in situ hybridisation (FISH) generated images showing the localisation of symbionts in *Formica* ants. A-C. Whole mount FISH of *Formica fusca*: queen gut (A, crop and proventriculus on the right, midgut in the middle, hindgut and Malpighian tubules on the left), ovaries (B) and egg (C). DAPI staining of host tissue in blue, symbiont stained in red. D-D’. FISH on transversal cytological sections of *Formica cinerea* larva midgut. DAPI staining only, showing host nuclei of bacteriocytes in a single layer surrounding the midgut (D), and a magnified region highlighting symbiont in red localised within bacteriocytes and in a bacteriome (D’).
Figure 4. Phylogenetic distributions of symbionts in queens of *Formica* (A) and *Cardiocondyla* (B) species. Pie charts represent the proportion of queens sampled that carried the symbiont (red) and those that did not (grey). Numbers represent the number of queens positive for the symbiont over total queens sampled (intracolony infection frequencies in Table S5). *Formica* phylogeny is based on 55 and *Cardiocondyla* phylogeny is based on 55, with major clades highlighted. Dashed lines indicate species added to the original source phylogeny based on additional published phylogenies (specified in the Taxonomic Analysis section of the methods). Starred names are provisional names of a recognized morphospecies to be described by B. Seifert.
Table 1: Comparison of the retention and losses of metabolic pathways for key nutrients across ant symbionts. Pathways displayed are based on those that have been shown to play important roles in other ant and insect symbiosis. Detailed breakdowns of these nutrient pathways along with analysis of other precursor, core metabolite synthesis, and transcriptional pathways, are available in Table S2.

*Tyrosine is considered a non-essential amino acid because it can be synthesised by most eukaryotic hosts from phenylalanine.
Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.docx
- SupplementaryTables.xlsx