Genome evolution following an ecological shift in nectar-dwelling *Acinetobacter*

Vivianna A. Sanchez¹, Tanya Renner², Lydia J. Baker¹*, Tory A. Hendry°#

¹ Department of Microbiology, Cornell University, Ithaca, New York, USA
² Department of Entomology, The Pennsylvania State University, University Park, USA
* Current address: Department of Marine Biology and Ecology, University of Miami, Florida, USA

Running Title: Genome evolution in nectar-dwelling *Acinetobacter*

#Address correspondence to Tory A. Hendry, th572@cornell.edu
Abstract

The bacterial genus *Acinetobacter* includes species found in environmental habitats like soil and water, as well as species adapted to be host-associated or pathogenic. High genetic diversity may allow for this habitat flexibility, but the specific genes underlying switches between habitats are poorly understood. One lineage of *Acinetobacter* has undergone a substantial habitat change by evolving from a presumed soil-dwelling ancestral state to thrive in floral nectar. Here we compared the genomes of floral-dwelling and pollinator-associated *Acinetobacter*, including newly described species, with genomes from relatives found in other environments to determine the genomic changes associated with this ecological shift. Following one evolutionary origin of floral nectar adaptation, nectar-dwelling *Acinetobacter* species have undergone reduction in genome size compared to relatives and have experienced dynamic gene gains and losses as they diversified. We found changes in gene content underlying carbohydrate metabolism and nitrogen metabolism, which we predict to be beneficial in nectar environments. Gene losses follow a pattern consistent with genome streamlining, whereas gains appear to result from both evolutionary divergence and horizontal gene transfer. Most notably, nectar-dwelling *Acinetobacter* acquired the ability to degrade pectin from plant pathogens and the genes underlying this ability have duplicated and are under selection within the clade. We hypothesize that this ability was a key trait for adaptation to floral nectar, as it could improve access to nutrients in the nutritional unbalanced habitat of nectar. These results identify the genomic changes and traits coinciding with a dramatic habitat switch from soil to floral nectar.
Introduction

The gammaproteobacteria genus *Acinetobacter* is diverse and includes (1–3). These taxa inhabit a broad range of environments, including soil and water (1–3). Some *Acinetobacter* lineages have also evolved to be host-associated or animal pathogens, with a notable example being the recently emerged human pathogen *Acinetobacter baumannii* (4). Strains in the genus are phenotypically and genetically diverse and frequently adapt to new ecological niches (5, 6). However, few direct connections have been made between specific genomic changes and ecological transitions within *Acinetobacter*, or in bacteria more broadly (7). One poorly characterized habitat transition within the genus *Acinetobacter* is adaptation for growth in floral nectar. Several *Acinetobacter* species found in floral nectar appear to be most closely related to soil-dwelling relatives (4, 8). Nectar represents a significant environmental shift compared to soil habitats, likely with different selective pressures. Genomic comparisons between *Acinetobacter* adapted to floral nectar versus other habitats can uncover how bacteria evolve to new environments and which genetic traits facilitate major ecological switches.

The high genetic diversity and genomic plasticity within *Acinetobacter* may be driven by mechanisms facilitating horizontal gene transfer (HGT), including competence for natural transformation (9–12), conjugative abilities, and prevalent mobile elements such as plasmids, prophage, and insertion sequences (8, 13, 14). Horizontally acquired genomic islands are common observed throughout the *Acinetobacter* genus (15, 16) and can contain genes conferring beneficial phenotypes like antibiotic resistance and plasmid mobilization (17, 18). HGT is a source of evolutionary novelty in bacteria (19), but other sources of genetic diversity can also be important, such as error-prone polymerases in *A. baumannii* (4, 8, 20), or gene duplication followed by divergence (21). Gene duplication can also potentially lead to increased gene
expression, allowing for enhanced nutrient acquisition, temperature stress tolerance, and overall resistance to antibiotics and pesticides (22), but it is unclear how broadly important this mechanism is for bacterial adaptation.

Genetic novelty in bacteria can allow for the evolution of new traits and subsequent exploitation of new niches (23–25). In some cases, specific genetic elements have been linked to habitat-specific fitness. For instance, *A. baumannii* has resistance islands with mobile genetic elements and antibiotic resistant genes, allowing for persistence in hospital settings (6, 26).

Antibiotic resistance is a common example of a novel trait resulting from a specific environmental selective pressure because it is easily observable and important in well studied pathogen systems. Other traits that are relevant in natural systems have occasionally been connected to ecological changes in bacteria (27, 28), but such connections can be difficult to infer. In other cases, traits that are linked to success in a specific environment may be known, but not their genetic basis. For instance, the ability to access nutrients from pollen is a unique and potentially beneficial trait in floral nectar-dwelling *Acinetobacter*, but how this trait was gained is unknown (29).

Floral nectar is a nutritional reward produced by flowers to attract pollinating animals. It is high in carbohydrates; the sugar content in floral nectar can reach 90% of nectar dry weight (30, 31) and it is a resource for microbes as well (32, 33). However, floral nectar habitats create several stresses for microbes, including limitation of nutrients other than sugar (34–38). Although nectar contains amino acids and lipids (39), it can contain limiting amounts of nitrogen for some microbes (40, 41). These factors make nectar a selective environment and can lead to strong priority effects where early arriving microbes prevent subsequent colonization of flowers (40, 42).
Culture dependent and independent methods have revealed diverse microbes that thrive in these conditions (43–49). The genus *Acinetobacter* makes up a high proportion of bacterial taxa in floral nectar and is prevalent and readily cultured from nectar environments (44, 50). *Acinetobacter* is also frequently found associated with floral visitors. For example, *Acinetobacter apis* was isolated from the gut of the western honey bee, *Apis mellifera*, and bee pollen provisions and nests sometimes include *Acinetobacter* (51–54). However, it is unknown whether *Acinetobacter* found with pollinators are nectar-dwelling species being dispersed by floral visitors, or if they are specific associates of pollinators. For ease here, we refer to strains isolated from both nectar and floral visitors as nectar-dwelling strains.

Previous phylogenomic analysis of *Acinetobacter* isolates from nectar and bees found that they were closely related to soil-dwelling species (50). This previous work suggested one evolutionary origin of nectar-dwelling/bee association but did not assess evolutionary patterns within this lineage. We do so here, with and additional 15 genomes from nectar-dwelling isolates. These include the genomes of three previously described species, *A. apis*, *A. boissieri*, and *A. nectaris* (45, 51), newly sequenced *A. nectaris* isolates, and newly sequenced genomes of three recently described species, *A. pollinis*, *A. rathckeae*, and *A. baretiae* (50). For comparison, we included genomes from *A. brisouii*, which is isolated from soil and water and was previously found to be the closest relative of *A. nectaris*, as well as those from eight other environmental *Acinetobacter* species. We hypothesized that the switch to floral nectar from soil would drastically change the selective pressures experienced by this *Acinetobacter* lineage, leading to changes in gene content. We used comparative genomics to understand which genes may become redundant or beneficial for bacteria in floral nectar, and to identify acquired genes that may have facilitated this environmental switch.
Results and Discussion

Phylogeny and genome characteristics of nectar-dwelling Acinetobacter

To understand the evolutionary history of nectar-dwelling Acinetobacter, we constructed a phylogenomic tree using genomes of Acinetobacter isolates from floral nectar and floral visitors. The isolates collected from floral nectar and pollinators form a clade, with a bootstrap support of 100, separate from soil, water, and animal dwelling Acinetobacter species (Fig. 1). This confirms that there is one known evolutionary origin of a nectar-dwelling within Acinetobacter and that this group evolved from a presumed soil-dwelling ancestor. The six species in the nectar clade appear to not be isolated from environments outside of floral nectar or pollinators, based on 16S rRNA sequence comparisons to GenBank databases (50). Multiple of these species, A. nectaris, A. boissieri, and A. pollinis, are abundant and common in floral nectar from locations worldwide, and our isolates came from both North America and Europe (45, 50, 55, 56). This suggests that the clade is specialized for growth in floral nectar and/or associated with pollinators and is widely found in these habitats (44, 57, 58).

We used genomic comparisons between nectar-dwelling Acinetobacter and relatives living in distinct environments to uncover genomic patterns associated with nectar specialization. Relative to species found in other environments, species in the nectar-specialized Acinetobacter clade have smaller genomes and lower numbers of protein coding genes (Table 1). There is a significant difference between genome sizes within the nectar-specialist clade, 2.38-2.75 Mb, and the environmental clade, 2.90-4.88 Mb (p-value=0.00023). Across genomes of nectar-dwelling strains vs. environmental strains, nectar specialists have 243-977 fewer protein coding genes, a 10-30% reduction in proteins. Genomic reduction can occur for various reasons.
Genome streamlining is common for species living in stable, nutrient poor conditions such as some soil and marine habitats (59–62), and is thought to be driven by selection and facilitated when bacteria have large effective population sizes (59, 63–65). Some environmental stresses may also promote genome streamlining due to selection (66, 67). Gene loss can also be degenerative and result from genetic drift, with extreme examples occurring in bacteria that are host-restricted and experience frequent population bottlenecks (68–71).

Although nectar-dwelling bacteria may experience population bottlenecks due to the transient nature of the floral environment, we do not find strong evidence for genetic drift as is seen in host-restricted bacteria; such species often show high evolutionary rates, high rates of pseudogenes, and low genome GC content (70, 72). Evolutionary rate tests in PAML found no support for a faster rate in nectar-dwelling *Acinetobacter*, compared to the null hypothesis of a global clock across the nectar and environmental *Acinetobacter* phylogeny (69). Similarly, we did not find evidence of genomic degeneration in the form of pseudogenes, as the number of pseudogenes detected in nectar-dwelling *Acinetobacter* ranged from 125-294, while environmental *Acinetobacter* had a similar range of 187-283 and other *Acinetobacter* species fall within this range as well (Table 1) (73, 74). The only pattern found that is potentially supportive of genetic drift is that nectar isolates have slightly lower percent GC compositions relative to soil-dwelling species (Table 1). We therefore hypothesize that the habitat switch from soil to nectar altered selective pressures on gene content in nectar dwelling species and allowed for the loss of some functions through genome streamlining. To further investigate this, we sought to define the gene content and functional capacities of nectar-dwelling species compared to soil-dwelling relatives.
Gene content evolution with the switch to nectar

To determine content of predicted proteins among *Acinetobacter* clades, we performed an ortholog clustering analysis to identify shared orthologs, recent paralogs, and unique genes (75) (Table S1). This analysis resulted in 7,334 orthologs in total identified across all genomes, 1,076 of which were present across all genomes. Environmental strains had 5,558 orthologs in total, with 2,921 orthologs only found in environmental species, and 111 environmental specific orthologs shared across these species. The nectar-dwelling clade had 4,413 total orthologs, with 1,776 orthologs only found in nectar-dwelling species, and 53 of these nectar specific orthologs found in all nectar clade strains.

To trace gene gain and loss events within the nectar-specialist clade, we performed a maximum likelihood ancestral state reconstruction analysis. Overall, there have been dynamic gene gains across the evolution of the group. Substantial ortholog gain events occurred at the ancestral node for the nectar-dwelling clade, the ancestral nodes to the species *A. pollinis*, *A. rathckeae*, and *A. baretiae*, and at additional ancestral nodes (Fig.1, nodes 1, 2, 5, 6, 8 and 9).

Gene gains were significantly higher than gene losses throughout the nectar-dwelling clade, with gains of 209-345 orthologs at tips and nodes (Fig. 1, Fig. S1). Numbers of gene loss events generally increased closer to the tips of the tree (Fig. 1, Fig. S1), suggesting that genome streamlining may be a relatively recent process in the clade. Given that the nectar-dwelling genomes are smaller in total size than the genomes of relatives (Table 1), these patterns suggests that some gene gains identified in this analysis are the result of divergence leading to novel orthologs rather than horizontal acquisition of new genes alone. Selective pressures from environmental changes can increase processes like mutation rate, leading to diversification in
orthologs (76). Here, adaptation to the novel nectar environment may have led to increased gene divergence.

Gene gains and losses occurred across diverse functional categories in the nectar-dwelling clade (Fig. 2). We investigated specific changes in gene content in functional categories, with the largest differences between nectar-dwelling species and environmental relatives occurring in the categories of metabolism of aromatic compounds, nitrogen metabolism, and carbohydrate metabolism.

The number of genes involved in metabolism of aromatic compounds (Fig. 2, Fig. S2A) varied between nectar and environmental strains. Although the difference in total genes was not significant, 48 out of 67 genes in this category were present in at least one environmental *Acinetobacter* species, but completely absent in the nectar clade (Table S1). All but one of the remaining genes were present in both clades, but rarely common across all nectar-dwelling strains. Several of the genes for aromatic compound metabolism that are present in nectar-dwelling *Acinetobacter* strains are involved in benzoate metabolism, an aromatic compound that has been shown to be released by plants (77). The decrease in genes in this category suggests that nectar-dwelling *Acinetobacter* encounter a limited diversity of aromatic compounds compared to species in other environments.

Gene content differences in nectar-dwelling strains compared to environmental relatives also suggest that shifts in nitrogen and amino acid metabolism strategies accompanied the shift to nectar dwelling. Genes involved in amino acid metabolism were significantly higher in nectar-dwelling strains (p-value=0.003). This difference was mainly due to an excess of genes for the transport of certain nitrogen sources, including putrescine, proline, glycine, serine, alanine, arginine, ornithine, methionine, leucine, glutamate, and aspartate (Table S1). In contrast,
nitrogen metabolism genes were reduced by about half in nectar dwelling Acinetobacter compared to environmental species, although the difference was not significant (Table S1). Specifically, nectar-dwelling species genomes are missing glutamine synthases, glutamate synthases and ammonium transporters. Many of these are genes for which multiple orthologs are present in environmental Acinetobacter and only one has been retained in nectar-dwelling strains. Floral nectar is low in nitrogen relative to carbon (31, 78), and the ability to assimilate nitrogen sources has been linked to competition and growth in floral nectar in both yeasts (79) and Acinetobacter (80, 81). A shift towards more transport systems for nitrogen sources could be driven by selection for nitrogen scavenging.

Genes involved in carbohydrate metabolism showed an overall decrease in nectar-dwelling strains (Fig.2), however this overall decrease was not statistically significant (p-value=0.089). The loss in carbohydrate genes was driven by the subcategories of central carbohydrate metabolism and significantly by genes in the category of organic acid metabolism, (p-value=2.2e-16) (Fig. S2B). Other subcategories, notably monosaccharide metabolism, showed a significant gain in genes (p-value=3.022e-16) (Fig. S2B). Compared to environmental habitats, floral nectar provides an abundance of simple carbohydrates, including fructose, sucrose, and glucose (78, 82). Many of the nectar-dwelling species can assimilate fructose, and some can assimilate glucose and sucrose (50, 80). In comparison, non-nectar dwelling species such as A. baylyi, are often unable to utilize fructose, sucrose, or glucose as sole carbon sources, but can metabolize other diverse carbon sources (83–85). In support of these observations, we found that nectar-dwelling species tend to have more genes for metabolizing and transporting simple sugars (Table S1). For example, all nectar dwelling species have the gene coding for 1-phosphofructokinase (EC 2.7.1.56), which is part of the fructose and mannose metabolism
pathway, whereas the genomes of environmental strains were less likely to have this gene.

Nectar-dwelling strain genomes were also more likely to have the gene encoding xylanase, which is responsible for the breakdown of the common plant polysaccharide xylan to xylose (86, 87). These genes may be more beneficial for nectar-dwellers than soil-dwelling Acinetobacter.

In keeping with overall changes in carbohydrate metabolism in nectar, we found that phosphotransferase system (PTS) genes specific to nectar sugars are more common in nectar-dwelling strains compared to environmental strains (Table 2). PTS genes are a common method for bacteria to transport sugars into cells via a phosphorylation cascade (88, 89). PTS can also be involved in sensing and regulation of physiological processes related to sugar, such as carbohydrate active enzymes, chemotaxis, and biofilm formation (90). These multicomponent systems are specific to distinct molecules including fructose, mannitol, sucrose, and glucose. The sucrose-specific PTS enzyme complex (EIIABC) is present within A. apis, A. boissieri, A. rathkeae, and A. baretiae strains, but absent from all environmental strains, as well as A. nectaris and A. pollinis. Fructose-specific EIIABC complexes were found in all nectar-dwelling strains and only three environmental species, A. lwoffii, A. baylyii, and A. brisouii. This pattern suggests that PTS enzymes may have been beneficial for making the ecological switch to high sugar environments. Overall, these findings show a shift in carbohydrate metabolism in keeping with growth in nectar, although we found mainly subtle differences in gene content.

Because of the difference in sugars available to environmental versus nectar-dwelling Acinetobacter, we further investigated their carbohydrate metabolic capabilities by comparing carbohydrate active enzymes identified using the CAZY database (91, 92). Genomes of environmental species contain significantly more genes in the glycosyl transferase (GT) family (p-value = 0.001), with an average of 23 genes per environmental species genome and 15 per
nectar dweller species genome (Fig. S3). This difference is mainly driven by enzymes in the GT2 and GT4 families. The role of these specific genes within *Acinetobacter* species is unclear, but generally enzymes in these families are involved in the synthesis of cell wall, capsular, and extracellular biofilm polysaccharides (93–96), suggesting that some of these functions may be different in nectar-dwelling *Acinetobacter*. We note, however, that several biofilm formation genes are found in the nectar-dwelling strains, including *pgaABCD* genes of the poly-β-1,6-N-acetyl-D-glucosamine (PGA) operon responsible for the maintenance of biofilm stability (97). Of the 15 nectar-dwelling strains, nine have the complete PGA operon, while four out of the 11 environmental strains have the full operon. Of the individual genes in the operon *A. pollinis* strains have 2-6 times the copies of *pgaB* genes, which contain binding domains critical for export of PGA (98), suggesting biofilm formation as an important trait in nectar-dwelling environments.

In contrast to the decrease in GT enzymes, nectar-dwellers contain significantly more genes in glycoside hydrolase (GH) families (p-value=8.152e-8), averaging 15 enzymes per species compared to approximately 13 genes per environmental species (Fig. S3). This pattern was mainly driven by genes in the GH 28 family, which are involved in the breakdown of the polygalacturonic acid backbone of pectin (99). Pectin is a major component of plant cells walls, and we hypothesize, as discussed below, that the ability to degrade this polysaccharide may be beneficial in floral nectar.

**Novel functions in nectar**

We investigated which genes had been likely gained by HGT in nectar-dwelling *Acinetobacter*, with the hypothesis that such genes may confer novel functions. We found
genomic islands within all members of the nectar-dwelling clade and some strains also contained intact prophages within their genomes. Gene count from genomic islands ranged from 111 – 352 genes with approximately 50% annotated as “hypothetical” and the remainder involved in plasmid or transposon mobilization, phage replication, or Type1 secretion components. Mobilization genes were present in genomic islands within the nectar dwelling clade suggesting that movement of genomic material is facilitated by mobile elements like plasmids (Table S2).

Among known functions within genomic islands, we found genes coding for pectin degradation enzymes, specifically PL1 family pectin lyases and GH28 family polygalacturonases. All species in the nectar-dwelling clade contain at least one of these genes, with several species possessing multiple copies of genes associated with the degradation of pectin (Table S3). Pectin is a recalcitrant polysaccharide that provides structural stability in plant cell walls and the outer layers of pollen grains (100). Among bacteria, enzymes for degrading pectin are commonly found in necrotrophic plant pathogens, which use them to digest plant tissue (101). These enzymes are notably absent among all *Acinetobacter* genomes in GenBank, with the exception of the orthologs found in the nectar clade (Table S3). Sequences in GenBank with the highest similarity to nectar-clade orthologs of PL1 and GH28 genes are found outside of the genus in plant pathogens such as *Pectobacterium, Erwinia* and *Dickeya* (Fig. 3A-B). This pattern supports that these genes were acquired by nectar-dwelling *Acinetobacter* by HGT, likely from a necrotrophic plant pathogen in the Enterobacteriales.

Tracing the pattern of gains in pectin degradation genes onto the phylogeny of nectar-dwelling *Acinetobacter* suggests that at least one copy each of the pectin lyase and polygalacturonase genes were present in the common ancestor of the nectar-clade *Acinetobacter* species analyzed here (Fig. 3A-B). In several strains these two orthologs are located next to each
other on the chromosome, so they may have been gained together in one event. As seen in the
gene trees for the pectin lyase and polygalacturonase orthologs, we determined that these genes
experienced multiple duplication events with paralogs sister to each other (Fig. 3A-B). These
duplications occurred within the species *Acinetobacter pollinis*, which contains six copies of
polygalacturonase and three copies of pectin lyase (within the strain SCC477 as an example).
Additional horizontal transfers, losses, or duplication events may have occurred within the nectar
clade, as some species have multiple copies within a gene tree (*A. boissieri*, Fig. 3A) or
differences in topology between gene trees and species trees (*A. apis*, Fig. 3B). Some of these
copies are on contigs that are likely from plasmids (based on increased read depth and the
presence of plasmid replication genes), which may have facilitated duplication and transfers of
these genes. Duplication was more common for the polygalacturonase than the pectin lyase
genes, and the polygalacturonase genes were also the only example of multiple copies outside of
the *A. pollinis* (Fig. 3A).

The fact that these genes have been maintained, and even duplicated, within the nectar-
dwelling clade suggests that they may serve an important ecological role for these bacteria. In
support of this, amino acid substitutions in several of pectin degrading enzyme protein sequences
show signatures of positive selection (Table S4). Positive selection was detected at the nodes and
tips of the polygalacturonase gene tree (Fig. 3), particularly for *A. pollinis* (nine sites) and *A. apis*
(seven sites) orthologs (Table S4). The high number of duplication events of these genes in *A.
pollinis*, together with signatures of positive selection, suggests that pectin degrading enzymes
may be functionally diversifying in this species. Both of these enzymes cleave linkages in the
polygalacturonic acid backbone of pectin (102). Necrotrophic plant pathogens typically have
diverse copies of these enzymes, with slight variations in catalytic ability, in order to effectively degrade pectin (102–104). Potentially, this pattern is convergently evolving in *A. pollinis*.

To investigate the potential function of the amino acids under selection in *Acinetobacter* polygalacturonase enzymes, we generated predicted protein structures of representative orthologs from *A. pollinis* strain FNA3 and the *A. apis* type strain (105). The polygalacturonase binding site contains four highly conserved motifs NTD, G/QDD, G/SHG and RIK. NTD and RIK are highly conserved catalytic regions while G/QDD and G/SHG are substrate binding regions (106).

In nectar strain orthologs all conserved regions were present (NTD, GDD, GHG, RIK) but no sites under selection were near these active sites (Fig. 4). It is therefore unclear from their location on the protein structure how sites under selection might impact protein function.

It is not known how much nectar-dwelling *Acinetobacter* interact with major sources of pectin in plant tissue, but to our knowledge they have not been observed to infect plants. However, microbes in nectar are regularly interact with pollen grains, which are introduced into nectar by pollinator activity (107). In fact, some *Acinetobacter* strains can cause pollen grains to burst open or pseudogerminate (29). This ability is beneficial for *Acinetobacter*, as it is associated with increased growth in nectar when pollen is present (29). Floral nectar has been shown to be nitrogen limiting for both yeasts and bacteria (80, 81), so the ability to access nitrogen from pollen within nectar could increase microbial fitness. However, pollen is protected by a resistant exine layer and is difficult to degrade (108). Pectin is an essential component of pollen cell walls and pollen tubes (109), and pectin degrading enzymes have been hypothesized to be involved in pollen breakdown by bacterial gut symbionts of honey bees (110). We hypothesize that the pectin degradation enzymes in *Acinetobacter* could be involved in accessing nutrients from pollen, which could explain the apparent importance of genes coding for such
enzymes in the clade. In support of this, we find that most of the GH28 and PL1 proteins in
Acinetobacter have secretion signals (Fig. 4B-D, Fig. S4B-D), similar to secreted pectin
degrading enzymes in Pectobacterium (102), suggesting that they should act extracellularly.
Furthermore, we find the most selection on these genes within A. pollinis and A. apis. The former
shows strong impacts on pollen bursting and pseudogermination (29) and the latter was isolated
from honey bees and is likely to encounter pollen regularly. We speculate that the ability to
degrade pectin could be a key trait allowing Acinetobacter to thrive in nectar and in association
with pollinators.

Conclusions

In this study we show that the ecological change from soil-dwelling to nectar-dwelling
led to genomic reduction, possibly due to genome streamlining, a phenomenon common among
bacteria living in oligotrophic habitats (66). Even with this streamlining, nectar-dwelling species
had higher levels of gene gain events compared to gene losses. Many of the gain gains occur at
early branching nodes within the nectar clade, and the nectar clade had an increase in the number
of genes involved in monosacharride metabolism and sugar transport, likely due to the high
sugar environment of nectar (78). We also found changes in nitrogen and amino acids
metabolism genes suggesting a switch towards nitrogen scavenging in comparison to
environmental Acinetobacter, consistent with nitrogen limitation in floral nectar (55). Nectar-
dwelling Acinetobacter species have also acquired pectin degrading enzymes, presumably
through HGT, from plant pathogens. We found duplication, diversification, and positive
selection within pectin degrading genes, supporting our hypothesis that these genes may provide
an important ecological function. Overall, we find that genome evolution from gene loss,
diversification, and HGT may have all contributed to the *Acinetobacter* habitat switch to floral nectar.

**Materials and Methods**

**Phylogenetic analyses**

A phylogenomic species tree was inferred using 26 *Acinetobacter* genomes and *Pseudomonas syringae pv. tomato* strain DC3000 as an outgroup (Table 1). These included sequences from nine environmental *Acinetobacter* species, and three previously published nectar-dwelling and pollinator associated *Acinetobacter* genomes (45, 51) obtained from GenBank. We did not include strains from the clades containing the animal pathogens *A. baumannii* or *A. parvus*, as previous work found these groups to have undergone distinct evolutionary changes compared to soil-dwelling relatives (4). Our recently published nectar-dwelling *Acinetobacter* genomes (see (50) for isolation information) were assembled using Discovar de novo (111), checked for completeness using CheckM (112), and annotated using the RAST Server (112) Nucleotide sequences from previously published *Acinetobacter* genomes were downloaded from GenBank and annotated in RAST for consistency. Protein sequences were used in the PhyloPhlAn 3.0 pipeline (114) to determine conserved proteins within *Acinetobacter* genomes. PhyloPhlAn identified 399 conserved proteins and their nucleotide sequences were used for phylogenetic tree reconstruction. Nucleotide sequences for the identified protein sequences were aligned using MAFFT (115). Maximum likelihood trees were reconstructed using IQ-Tree (116) with bootstrapping set to 1000 and a symmetric substitution model (117). Welch’s t-test was used to determine the significance between genome size from the nectar clade and environmental clade.
Gene trees were reconstructed using polygalacturonase and pectin lyase genes from the
Acinetobacter genomes. Outgroups were selected by using nectar-dwelling Acinetobacter spp.
pectin lyase and polygalacturonase genes as BLAST queries in GenBank. We found that all of
the best BLAST hits for these genes were from necrotrophic plant pathogens so we included the
most similar representative sequences in the analyses. Acinetobacter polygalacturonase and
pectin lyase genes were identified in our newly sequenced genomes by RAST annotation,
BLAST of the genomes using plant pathogen orthologs as queries, and comparison with the
CAZy database to confirm that we had identified all orthologs. Genes were aligned using
MAFFT (115) and were used for maximum likelihood phylogenetic inference in IQ-Tree using
the TIM3 substitution model which was selected using model finder in IQ-Tree.

Ortholog analyses

Orthologous protein sequence clustering was conducted using OrthoMCL (75) using the
following parameters: mode = 1, inflation = 2, pi cutoff = 50. To determine the ancestral state of
orthologs across the Acinetobacter species tree, the software package Count was used (118)
implementing Wagner Parsimony. This analysis determines the genes gained and lost at each
node of the Acinetobacter phylogenomic tree minimizing state changes and assuming all
character states are reversible (119). To estimate the number of pseudogenes present within the
genomes, we used the program Pseudofinder (120). The algorithm identifies pseudogenes from
Genbank files by analyzing average coding sequence (CDS) length, fragmented CDS, and
intergenic pseudogenes and alignment lengths are compared against homologs identified by
blastp hits from a reference protein database (121). The following parameters were used to
predict potential intergenic, fragmented, truncated, and long pseudogenes: intergenic length = 30,
Plasmids were assembled using unfiltered paired end 2 x 250 Illumina sequencing reads (50) and the plasmidSPAdes assembly tool (122). Plasmids less than 1000 bp were excluded from final plasmid count. Fisher’s exact test was used to calculate all significance test between orthologs from the nectar groups and the environmental group.

Selection and protein analyses

dN/dS (ω) values were estimated for the polygalacturonase and pectin lyase coding region alignments and IQ gene tree phylogenies using codeml in the PAML v4.4 package, with gaps included (123). For each branch and node in the phylogenies, a likelihood-ratio test for positive selection was performed to compare nested branch-site models (Model A_\text{null} versus Model A) (124) (Table S4).

The AI protein prediction software, Alphafold, was used to predict protein structure of pectin degradation enzymes. The Alphafold algorithm is a neural network that generates a multiple sequence alignment (MSA) from the query protein sequence provided and extracts evolutionary information to generate protein predictions (105). A web version of Alphafold was used for these predictions (125). In order to determine if amino acid sites under selection were functionally important, we predicted the structure of polygalacturonase genes identified to be under positive selection from the PAML analysis, specifically genes from *Acinetobacter pollinis* strain FNA3 (GenBank locus tags I2F29_RS12745 and I2F29_RS12925), *Acinetobacter apis* (CFY84_RS01715), and *Phaseolibacter flectens* (L871_RS0110380). Three-dimensional protein predictions were edited using ChimeraX to highlight sites under selection (126).
Data availability

All data is available in the supplemental materials or NCBI databases. Genomic assemblies are available in GenBank under accessions VTDP00000000, VTDO00000000, VTDM00000000, VTDL00000000, JAEQDL000000000, JAERJC000000000, JAERJB000000000, JAEQDM000000000, VTDS00000000, VTDT00000000, VTDR00000000, and VTDQ00000000. Sequence reads are available in the SRA under accessions SRR26518995-SRR26519002.

Acknowledgements

We thank the Cornell Biotechnology Resource Center for sequencing and Heather Feaga for consultation on analyses. We are grateful to Rachel Vannette, Tadashi Fukami, and Bart Lievens for sharing isolates for sequencing.

Support was provided by the SUNY Graduate Diversity Fellowship (V.A.S.) and the Cornell Atkinson Center for Sustainability Sustainable Biodiversity Fund (V.A.S.). V.A.S. and T.A.H. conceived of the study and V.A.S., T.A.H., L.J.B., and T.R. performed analyses. Writing was collaborative between V.A.S., T.A.H., and T.R., with comments from L.J.B.
References


Figure 1. *Acinetobacter* species maximum likelihood phylogenomic tree based on 399 conserved protein sequences. All nodes have a bootstrap value of 100. Ancestral state reconstruction of ortholog gains and losses at nodes within the nectar-dwelling clade are shown. Nodes are labelled numerically and a pie chart at each node shows the proportion of ortholog gains (yellow) versus losses (blue). Pie size is scaled by total number of gain and loss events.
Figure 2: Heatmap showing number of orthologs across functional categories from nectar-dwelling and environmental *Acinetobacter* strains. Category assignments are based on RAST annotations (FA = fatty acid metabolism, TE = transposable elements), and a phylogenomic species tree (Figure 1) is shown. Each row is colored independently based on the variance of gene content within the category. A color scale with yellow indicating the highest number of genes and deep blue indicating the lowest number of genes is used. Genes of unknown function or categories with <6 mean orthologs per genome are excluded. The mean number of orthologs within a genome for each category is given.
Figure 3. Maximum likelihood phylogenetic trees of A) the polygalacturonase gene, and B) the pectin lyase gene from nectar-dwelling *Acinetobacter* strains and the most closely related orthologs from plant pathogens. Red bars indicate likely duplication events and blue bars indicated horizontal gene transfer events. Bootstrap values below 80 are displayed at nodes. Asterisks show nodes and tips with significant positive selection. Orthologs from outside of *Acinetobacter* are from plant pathogens (*Erwinia pyrifoliae* (GenBank accession CP023567), *Pseudomonas flectans* (JAEE01000003), *Pectobacterium carotovorum* (CP001657), *Pectobacterium odoriferum* (MTAN01000008), *Pectobacterium wasabiae* (CP015750), *Dickeya zeae* (CP006929), and *Pseudomonas fluorescens* (OV986001)). Outgroup orthologs represent the best BLAST hits (>50% amino acid sequence identity, one ortholog per species) in GenBank databases.
Table 1. Genome characteristics of nectar-dwelling Acinetobacter strains and comparison environmental strains. Type strains are designated.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Accession</th>
<th>Pseudo-genes</th>
<th>Genome Size (Mb)</th>
<th>G+C Content (%)</th>
<th>Protein Coding Genes</th>
<th>Proteins assigned to COGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. apis ANC 5114&lt;sup&gt;+&lt;/sup&gt;</td>
<td>FZLN00000000</td>
<td>125</td>
<td>2.41</td>
<td>38.3</td>
<td>2232</td>
<td>1985</td>
</tr>
<tr>
<td>A. boissieri ANC 4422&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_FMYL000000000</td>
<td>182</td>
<td>2.69</td>
<td>38.0</td>
<td>2554</td>
<td>2378</td>
</tr>
<tr>
<td>A. rathkeae EC115</td>
<td>VTDP000000000</td>
<td>192</td>
<td>2.62</td>
<td>39.2</td>
<td>2438</td>
<td>2388</td>
</tr>
<tr>
<td>A. rathkeae EC24</td>
<td>VTDO000000000</td>
<td>214</td>
<td>2.75</td>
<td>39.3</td>
<td>2565</td>
<td>2531</td>
</tr>
<tr>
<td>A. baretiae B10A&lt;sup&gt;+&lt;/sup&gt;</td>
<td>VTDM000000000</td>
<td>234</td>
<td>2.59</td>
<td>37.5</td>
<td>2531</td>
<td>2402</td>
</tr>
<tr>
<td>A. baretiae B5B</td>
<td>VTDL000000000</td>
<td>253</td>
<td>2.70</td>
<td>37.5</td>
<td>2629</td>
<td>2486</td>
</tr>
<tr>
<td>A. nectaris CIP 110549&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_AYER000000000</td>
<td>189</td>
<td>2.67</td>
<td>36.7</td>
<td>2569</td>
<td>2458</td>
</tr>
<tr>
<td>A. nectaris BB362</td>
<td>JAEQDL000000000</td>
<td>249</td>
<td>2.38</td>
<td>36.9</td>
<td>2269</td>
<td>2173</td>
</tr>
<tr>
<td>A. nectaris EC31</td>
<td>JAERJC000000000</td>
<td>200</td>
<td>2.42</td>
<td>36.7</td>
<td>2312</td>
<td>2303</td>
</tr>
<tr>
<td>A. nectaris EC34</td>
<td>JAERJB000000000</td>
<td>212</td>
<td>2.39</td>
<td>36.8</td>
<td>2288</td>
<td>2284</td>
</tr>
<tr>
<td>A. nectaris BB226</td>
<td>JAEQDM000000000</td>
<td>255</td>
<td>2.58</td>
<td>36.7</td>
<td>2421</td>
<td>2312</td>
</tr>
<tr>
<td>A. pollinis FNA11</td>
<td>VTDS000000000</td>
<td>245</td>
<td>2.67</td>
<td>36.6</td>
<td>2621</td>
<td>2599</td>
</tr>
<tr>
<td>A. pollinis FNA3</td>
<td>VTDT000000000</td>
<td>248</td>
<td>2.67</td>
<td>36.6</td>
<td>2615</td>
<td>2575</td>
</tr>
<tr>
<td>A. pollinis SCC474</td>
<td>VTDRO000000000</td>
<td>294</td>
<td>2.75</td>
<td>36.7</td>
<td>2623</td>
<td>2605</td>
</tr>
<tr>
<td>A. pollinis SCC477&lt;sup&gt;+&lt;/sup&gt;</td>
<td>VTDQQ000000000</td>
<td>262</td>
<td>2.75</td>
<td>36.7</td>
<td>2634</td>
<td>2609</td>
</tr>
<tr>
<td>A. brisouii ANC4419</td>
<td>NZ_APPR000000000</td>
<td>187</td>
<td>3.11</td>
<td>41.7</td>
<td>3036</td>
<td>3021</td>
</tr>
<tr>
<td>A. brisouii CIP 110357&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_AYEU000000000</td>
<td>184</td>
<td>3.09</td>
<td>41.6</td>
<td>3025</td>
<td>3003</td>
</tr>
<tr>
<td>A. brisouii SM1</td>
<td>NZ_JZRE000000000</td>
<td>189</td>
<td>2.99</td>
<td>41.8</td>
<td>2913</td>
<td>2651</td>
</tr>
<tr>
<td>A. baylyi ADP1</td>
<td>222</td>
<td>3.6</td>
<td>40.4</td>
<td>3376</td>
<td>2897</td>
<td></td>
</tr>
<tr>
<td>A. ursingii NIPH 706</td>
<td>NZ_APQB000000000</td>
<td>234</td>
<td>3.49</td>
<td>40</td>
<td>3468</td>
<td>2920</td>
</tr>
<tr>
<td>A. bouveti DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_APQD000000000</td>
<td>198</td>
<td>3.37</td>
<td>40</td>
<td>3468</td>
<td>2657</td>
</tr>
<tr>
<td>A. indicus ANC 4215</td>
<td>ATGH01000001</td>
<td>238</td>
<td>3.17</td>
<td>45.4</td>
<td>3064</td>
<td>2632</td>
</tr>
<tr>
<td>A. ursingii DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_APQD000000000</td>
<td>198</td>
<td>3.37</td>
<td>40</td>
<td>3468</td>
<td>2657</td>
</tr>
<tr>
<td>A. indicus ANC 4215</td>
<td>ATGH01000001</td>
<td>238</td>
<td>3.17</td>
<td>45.4</td>
<td>3064</td>
<td>2632</td>
</tr>
<tr>
<td>A. ursingii DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_APQD000000000</td>
<td>198</td>
<td>3.37</td>
<td>40</td>
<td>3468</td>
<td>2657</td>
</tr>
<tr>
<td>A. indicus ANC 4215</td>
<td>ATGH01000001</td>
<td>238</td>
<td>3.17</td>
<td>45.4</td>
<td>3064</td>
<td>2632</td>
</tr>
<tr>
<td>A. ursingii DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>NZ_APQD000000000</td>
<td>198</td>
<td>3.37</td>
<td>40</td>
<td>3468</td>
<td>2657</td>
</tr>
</tbody>
</table>

Table 2. Presence and absence of phosphotransferase system (PTS) genes within *Acinetobacter* strains and the isolation environment for each strain (citations provided where available).

<table>
<thead>
<tr>
<th>Strain</th>
<th>None</th>
<th>PTS fructose-specific EIABC complexes</th>
<th>PTS sucrose-specific EIIBC complexes</th>
<th>Strain environment (source citation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. apis ANC 5114&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. boissieri ANC 4422&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. rathkeae EC115</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. rathkeae EC24</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. baretiae B10A&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. baretiae B5B</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. nectaris CIP 110549&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. nectaris BB362</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. nectaris EC31</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. nectaris EC34</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. nectaris BB226</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. pollinis FNA11</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. pollinis FNA3</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. pollinis SCC474</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. pollinis SCC477&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. brisouii ANC4419</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. brisouii CIP 110357&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. brisouii SM1</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. baylyi ADP1</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. ursingii NIPH 706</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. bouveti DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. indicus ANC 4215</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. ursingii DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. indicus ANC 4215</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. ursingii DSM 14964&lt;sup&gt;+&lt;/sup&gt;</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Accession</td>
<td>Location</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td><em>A. apis</em> ANC5114</td>
<td>1</td>
<td>1</td>
<td>Honey bee gut (51)</td>
<td></td>
</tr>
<tr>
<td><em>A. boissieri</em> ANC4422</td>
<td>1</td>
<td>1</td>
<td>Floral nectar (45)</td>
<td></td>
</tr>
<tr>
<td>EC115</td>
<td>1</td>
<td>1</td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>EC24</td>
<td>1</td>
<td>1</td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>B10A</td>
<td>1</td>
<td>1</td>
<td>Honey bee gut (50)</td>
<td></td>
</tr>
<tr>
<td>B5B</td>
<td>1</td>
<td></td>
<td>Honey bee mouth (50)</td>
<td></td>
</tr>
<tr>
<td><em>A. nectaris</em> CIP110549</td>
<td>1</td>
<td></td>
<td>Floral nectar (45)</td>
<td></td>
</tr>
<tr>
<td>BB362</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>EC31</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>EC34</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>BB226</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>FNA11</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>FNA3</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>SCC474</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td>SCC477</td>
<td>1</td>
<td></td>
<td>Floral nectar (50)</td>
<td></td>
</tr>
<tr>
<td><em>A. brisouii</em> ANC4119</td>
<td>1</td>
<td></td>
<td>Peat (127)</td>
<td></td>
</tr>
<tr>
<td><em>A. brisouii</em> DSM189516=CIP 110357</td>
<td>1</td>
<td></td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td><em>A. brisouii</em> ANC3789=SM1</td>
<td>1</td>
<td></td>
<td>River in Endau-Rompin National Park</td>
<td></td>
</tr>
<tr>
<td><em>A. baylyi</em> CIP107474</td>
<td>1</td>
<td></td>
<td>Activated sludge plant (12)</td>
<td></td>
</tr>
<tr>
<td><em>A. ursingii</em> NIPH706</td>
<td>1</td>
<td></td>
<td>Human blood</td>
<td></td>
</tr>
<tr>
<td><em>A. bouveti</em> CIP110357</td>
<td>1</td>
<td></td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td><em>A. indicus</em> ANC4215</td>
<td>1</td>
<td></td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td><em>A. lwoffii</em> CIP64.10</td>
<td>1</td>
<td></td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td><em>A. guillouiae</em> NIPH991</td>
<td>1</td>
<td></td>
<td>Human ear swab</td>
<td></td>
</tr>
<tr>
<td><em>A. larvae</em> BRTC1</td>
<td>1</td>
<td></td>
<td>Moth larvae (128)</td>
<td></td>
</tr>
<tr>
<td><em>A. rudis</em> CIP110305</td>
<td>1</td>
<td></td>
<td>Raw milk and wastewater (129)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4: AlphaFold protein predictions of polygalacturonase genes from *Acinetobacter apis* (A), *Acinetobacter pollinis* FNA3 (B-C), and *Phaseolibacter flectens* (D). Conserved motifs and sites under selection are highlighted blue and yellow respectively. Secretion signal tags are present in B-D. These extend off the protein to the left in C-D, but backwards in B. B is rotated in Fig. S2 to show the secretion signal.