# Vertical inheritance and loss-driven evolution of secretion systems in the bee gut microbiota

Samuel A. Acheampong[1] & Waldan K. Kwong[1,2]

[1] Gulbenkian Institute for Molecular Medicine, Oeiras 2780-156, Portugal

[2] Faculty of Medicine, University of Lisbon, Avenida Professor Egas Moniz, Lisbon 1649-028, Portugal

Correspondence: samuel.acheampong@gimm.pt

**Abstract**

The stability of gut bacterial communities is determined by complex inter-cellular interactions such as competition, cooperation and host dynamics. A mechanism proposed to mediate these interactions is bacterial secretion systems: specialized protein complexes that secrete effector molecules into neighbouring cells or the surrounding environment to influence community stability. However, the forces driving secretion system distribution and evolution in host-associated microbiomes remain unclear. Here, we show that secretion systems in the bee gut microbiome are predominantly vertically inherited and evolve primarily through recurrent gene loss rather than horizontal acquisition. Using comparative genomic analysis, we found that bee gut symbionts mostly encode type I, V, and VI secretion systems. In contrast, the pathogen-associated type II and III systems are missing, but they retain evolutionarily related pili and flagella. We found weak association between the presence of specific secretion systems and the bee hosts, suggesting that these systems are maintained for interbacterial interactions rather than host-specific adaptation. Co-phylogenetic analyses show congruence between bacterial strain phylogenies and most secretion system phylogenies, indicating a vertical mode of transmission. Only a subtype of the type VI system in the *Orbaceae* and the type IV system in *Snodgrassella* spp. show evidence of horizontal transmission. The lack of horizontal transfers means that losses of secretion systems is a permanent evolutionary event in almost all lineages of the bee gut microbiota. Our study provides a uniquely comprehensive analysis of secretion systems across an entire gut bacterial community, giving insight into how microbiomes evolve and maintain functional interactions within host-associated environments.

## Introduction

Many animals and plants harbour host-specific microbiomes composed of bacterial communities that influence host physiology, development and ecological adaptation (McFall-Ngai et al. 2013; Christian et al. 2015; Hou et al. 2022). Within these host-associated environments, bacteria interact continuously with the host and with co-occurring microorganisms to persist. The traits that mediate such interactions, ranging from metabolic exchange to cooperation and competition, are therefore subject to long-term evolutionary pressures shaped by transmission mode, ecological stability, and opportunities for genetic exchange (Ghoul and Mitri 2016; Baishya and Wakeman 2019; Gorter et al. 2020). Yet, how these pressures influence the evolution and persistence of microbial interaction traits in stable host-associated microbiomes remains poorly understood.

One of the traits that mediate these intercellular interactions is the bacterial secretion systems. These are specialized protein complexes that transport a wide range of effector molecules (mainly proteins) across the cell envelope into neighbouring cells or the extracellular environment, facilitating functions such as host colonization, nutrient acquisition, competition & cooperation, and DNA uptake & competence (Aoki et al. 2010; Leo et al. 2010; Zhang et al. 2011; Green and Mecsas 2016; Sassone-Corsi et al. 2016; Harms et al. 2017; Lin et al. 2017; Speare et al. 2018; Luo et al. 2025; Waksman 2025). In Gram-negative bacteria, secretion systems comprise a diverse set of architectures and functions (T1SS–T9SS, Fig. 1A), and have evolutionarily related extracellular appendages such as type IV and tight adherence pili, and flagella (Büttner and He 2009; Green and Mecsas 2016; Cianciotto and White 2017; Lasica et al. 2017; Mignolet et al. 2018; Coulthurst 2019; Li et al. 2019; Meuskens et al. 2019; Spitz et al. 2019; Whitfield and Brun 2024). Although these systems and appendages are considered important determinants of microbial ecology, they have been predominantly studied in pathogenic and free-living bacteria, mostly in isolation or in simplified experimental systems (Green and Mecsas 2016; Nazir et al. 2017; Nicholson and Champion 2022; Maphosa et al. 2023; Jamali et al. 2025). They are often considered highly dynamic traits, frequently shaped by horizontal gene transfer, gene loss, duplications and rearrangements (Abby and Rocha 2012; Denise et al. 2020).

Growing evidence suggests that these secretion machineries are also important in non-pathogenic host associations (even commonly retained in genomically-reduced endosymbionts) (Okazaki et al. 2013; Steele et al. 2017; Speare et al. 2018; Schepers et al. 2021; Suria et al. 2022; Zboralski et al. 2022; Wangthaisong et al. 2023; Bontemps et al. 2024; Motta et al. 2024), a view once encapsulated by the concept of symbionts as “intelligent pathogens” (Deakin and Broughton 2009). Contrary to pathogens and free-living bacteria, bacteria belonging to long-term, host-restricted microbiomes may experience different evolutionary trajectories. These symbiotic environments are often characterized by ecological stability, repeated transmission through the host, and reduced exposure to external microbial gene pools – conditions that tend to limit horizontal gene transfer and promote genome streamlining (Moran et al. 2008; Ganesan et al. 2022; Perreau and Moran 2022). Under such constraints, complex and energetically costly traits such as secretion systems may be selectively

retained, modified, or lost depending on their long-term utility within the community. As such, secretion systems can be viewed as ecological traits whose evolutionary trajectory reflects a bacterium's interactions with its host and co-occurring microorganisms. The question remains whether, in such stable microbiomes, secretion systems are primarily acquired horizontally or inherited vertically and what the relative contributions of gene gains and losses are.

The gut microbiome of eusocial bees, including honey bees, bumble bees, and stingless bees, offers a powerful model system for studying these dynamics. Despite its taxonomic simplicity, the bee gut microbiome is remarkably stable across individual bees, colonies, and generations reflecting strong host–microbe associations and potential long-term coevolution (Kwong and Moran 2016a; Kwong, Medina, et al. 2017; Motta and Moran 2024). The bee gut microbiota supports multiple host functions, including nutrient metabolism, protection against pathogens, and modulation of host immunity (Zheng et al. 2016; Kešnerová et al. 2017; Zheng et al. 2017; Lee et al. 2018; Raymann and Moran 2018; Zheng et al. 2019; Steele et al. 2021; Z. Zhang et al. 2022). Transmission of the microbiome occurs largely through interactions among nestmates, reinforcing long-term host association and limiting exposure to external microbial reservoirs (Powell et al. 2014; Kwong and Moran 2016a). These features make the bee gut microbiome particularly well suited for testing hypotheses about the evolutionary stability, transmission mode, and diversification of bacterial interaction traits.

Previous work on the bee gut microbiome has focused primarily on taxonomic composition, metabolic capabilities, and contributions to host physiology, including nutrition, immune modulation, and protection against pathogens. However, the cellular mechanisms that mediate interactions among symbionts and the evolutionary processes shaping these mechanisms remain comparatively underexplored. While individual secretion systems such as the type VI secretion system (T6SS) have been examined in one or few lineages (Steele et al. 2017; Steele and Moran 2021; Schmidt et al. 2023; Motta et al. 2024; Jones et al. 2025), no systematic comparative analysis has addressed the secretion systems repertoires across the microbiome, how they are transmitted and maintained across the bacterial lineages, and what ecological functions they may serve within the community.

Here, we investigate the mode of transmission and evolution of bacterial secretion systems in the bee gut microbiome using comparative genomics and cophylogenetic analyses of 391 bee-associated bacterial genomes spanning the dominant Gram-negative gut lineages. We characterize secretion system repertoires, assess their phylogenetic congruence, and reconstruct patterns of gain and loss over evolutionary timescales. By integrating metagenomic datasets from bees of different host species and age groups, we further evaluate their distribution and abundance in natural communities. Together, our results reveal an evolutionary trajectory in which secretion systems are predominantly vertically inherited and driven by gene loss rather than horizontal acquisition, providing insight into how stable microbial communities evolve and maintain functional interactions within host-associated environments.

## Results

### Bee gut symbionts encode a restricted secretion system repertoire

To systematically identify and quantify secretion systems in bee-associated bacteria, we developed a customised detection pipeline integrating established computational tools TXSScan and SecReT6 (J. Li et al. 2015; Abby et al. 2016; J. Zhang et al. 2022), with system-specific thresholds calibrated to minimise false positives (Materials and Methods). This approach, coupled with manual curation, allowed for greater sensitivity compared to prior gene survey studies that relied on general-purpose annotation tools such as the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Barret et al. 2013; Ugarte-Ruiz et al. 2015; Wallner et al. 2021; Liyanapathiranage et al. 2022; Robinson et al. 2022; Robinson et al. 2023). We analysed 391 high-quality, publicly available genomes from bee gut symbionts and bee-associated environments (e.g. honey- and pollen-derived isolates), screening for the presence and absence of conserved core components of bacterial secretion systems and related appendages (Fig. S1-S5). Phylogenetically related bacteria from non-bee hosts and environmental sources were included as outgroups to enable comparative genomic analyses. Subsequent analyses focused on strains from the five dominant Gram-negative lineages of the bee gut microbiota: *Apibacter* (*Weeksellaceae*), *Bartonella* (*Bartonellaceae*), *Bombella* and *Commensalibacter* (*Acetobacteraceae*), *Frischella* and *Gilliamella* (*Orbaceae*), and *Snodgrassella* (*Neisseriaceae*). No secretion systems were detected in the Gram-positive core members (*Lactobacillus*, *Bombilactobacillus*, and *Bifidobacterium*), and these genomes were therefore excluded from further analyses.

Across the analysed genomes, the type V secretion system (T5SS) was the most prevalent, with 84.1% (329 of 391) encoding the T5aSS (classical autotransporters) subtype, 52.4% the T5bSS (two-partner secretion) subtype, and 22.5% encoding the T5cSS (trimeric autotransporter adhesins) subtype (Fig. 1B). The type I secretion system (T1SS) was the second most common, present in 59.3% of the genomes, while the type VI secretion system (T6SS) was found in 51.7%. The T9SS was the rarest, found in only 5.4% of the genomes. To assess copy number variation within genomes, candidate secretion system loci were evaluated using a completion-based criterion requiring the presence of mandatory core components, while accounting for both single-locus and multi-locus architectures defined in the system models (Materials and Methods). Most genomes encoding a given secretion system carried a single copy (Fig. 1B; Fig. S6). In contrast, T5aSS, encoded by a single gene, frequently occurred in multiple copies, with up to eight copies per genome. This pattern was not lineage-restricted, as multi-copy T5aSS were detected across multiple bee-associated lineages (Fig. S6). By contrast, *Bartonellaceae* and most *Acetobacteraceae* genomes typically encoded two copies of the T1SS, whereas *Orbaceae* and *Neisseriaceae* predominantly carried a single copy. Multi-copy T5cSS was restricted to *Neisseriaceae*, which also consistently encoded multiple T5aSS genes (Fig. S6).

### Virulence-associated secretion systems are absent, but related appendages are widespread

The type II secretion system (T2SS) and type III secretion system (T3SS) are commonly associated with pathogenicity (Cianciotto and White 2017; Maphosa et al. 2023). They are known to inject virulence factors into hosts cells, enabling tissue invasion. Strikingly, none of the bee gut or bee-associated bacteria encoded either system (Fig. 1A), suggesting the bee gut and bee-associated bacteria may be lacking host tissue invasive traits. Instead, we observed the widespread presence of the T2SS-evolutionarily related appendages, tight adherence (Tad) pili and type IV pili (T4P) (Peabody et al. 2003) in 22.5% and 74.4% of genomes, respectively and T3SS-related flagellar apparatus (Grognot and Taute 2021; Akahoshi and Bevins 2022) was encoded by 60.6% of genomes (Fig. 1B). These appendages have been implicated in mediating biofilm formation, host attachment, surface sensing, surface motility and inter-bacterial aggregation (Mignolet et al. 2018; Ligthart et al. 2020; Grognot and Taute 2021; Akahoshi and Bevins 2022; Geiger et al. 2024; Whitfield and Brun 2024). All genomes encoding Tad, T4P, or flagella carried a single copy of these appendages (Fig. 1B; Fig. S6).

**Secretion system repertoires are determined by bacterial lineage rather than host identity**

Given that secretion systems and appendages often reflect niche adaptation and are non-randomly distributed across bacterial species (Green and Mecsas 2016), we examined their occurrence across the genomes to identify potential host- or lineage-specific patterns in system repertoires. We observed marked lineage-specificity in the distribution of many systems (Fig. 2A). T1SS was highly prevalent in all genera except *Apibacter*. The T2SS-related Tad is present only in the *Orbaceae* family (*Frischella* and *Gilliamella*), while the T4P is present only in *Frischella*, *Gilliamella*, and *Snodgrassella*. T5cSS was found mostly in *Snodgrassella* and T9SS was only carried by *Apibacter*, suggesting phylogenetically constrained functions. Overall, *Gilliamella* and *Snodgrassella* carried the most diverse and variable secretion system repertoires, while *Apibacter*, *Bartonella*, *Bombella*, and *Commensalibacter* possessed highly conserved and streamlined profiles.

To quantify these differences, we analysed the alpha- and beta-diversity of the secretion system repertoires of the strains (Fig. 2B, Fig. S7). *Frischella*, *Gilliamella*, and *Snodgrassella* exhibited significantly higher richness, Shannon, and Simpson diversity than the other lineages (Kruskal-Wallis, $p < 0.01$), consistent with their broader range of encoded systems. However, evenness values within these lineages remained moderate, indicating that, although they possess a larger repertoire overall, their profiles are skewed toward a core set of dominant systems that occur in most strains, and additional systems are present only sporadically. In contrast, *Apibacter*, *Bartonella*, *Bombella*, and *Commensalibacter* exhibited both low richness and high uniformity, consistent with their streamlined repertoires. PCoA and beta-dispersion analyses using both abundance and presence/absence data showed a strong lineage effect on secretion system composition, with lineage explaining 68.6% of the variance among strains (Fig. S7).

To assess whether secretion systems show non-random associations across strains, we computed pairwise co-occurrence using Pearson correlations on presence–absence matrices (Fig. S8A). Overall,

177 the co-occurrence analysis revealed no strong positive correlations among secretion systems indicating
178 that the presence of one system provides little predictive information about the presence of another. A
179 few patterns were nevertheless detectable: for example, moderate positive associations between Tad,
180 T4P, and the flagellum, and between T4P and T5aSS, whereas T5cSS and the flagellum exhibited a
181 strong negative association. Considering that the dataset is heavily dominated by *Frischella* &
182 *Gilliamella* (n = 204) and *Snodgrassella* (n = 88), and that lineage-dependence (i.e., phylogenetic signal)
183 could influence the co-occurrence patterns, we evaluated associations while explicitly accounting for
184 phylogeny using Coinfinder (Whelan et al. 2020). This approach estimates whether co-occurrence is
185 more frequent than expected after controlling for shared evolutionary history. In *Frischella* &
186 *Gilliamella*, we found that T1SS, Tad, and T6SS tended to co-occur more often than expected
187 irrespective of phylogeny ($p < 0.05$, Fig. S8B). In *Snodgrassella*, T1SS and T5bSS showed a weak but
188 non-significant pattern of co-occurrence ($p = 0.08$, Fig. S8C).
189 Since the analysed bacterial strains originate from a variety of bee host species, and strains can
190 be highly specific to their particular bee hosts (Kwong and Moran 2016a; Powell et al. 2016; Kwong,
191 Medina, et al. 2017; Ellegaard et al. 2020), we considered if secretion systems also display host
192 specificity. When mapped to host identity, we found patterns of host-specificity for only a few systems,
193 supported by weak clustering in PCoA analysis (Fig. 2C, Fig. S9). T9SS was present only in *Apis cerana*,
194 *Apis dorsata*, and the bumble bee subgenus *Melanobombus*, with this distribution driven by the
195 specificity of *Apibacter* strains to these hosts. Intriguingly, the T4SS appeared to be enriched in bumble
196 bee derived strains compared to strains from honey bees. The T4SS is infrequently found in bee gut
197 strains overall (Fig. 1B) but is sporadically present in members of all 5 analysed bee gut lineages.
198
199 **Secretion systems are predominantly vertically inherited**
200 To assess whether secretion systems were vertically inherited or horizontally acquired, we
201 compared the phylogenies of secretion systems to their corresponding bacterial lineage phylogenomic
202 trees. Phylogenetic congruence, as a proxy of mode of inheritance, was evaluated using three distance-
203 based cophylogenetic methods, PACo, Parafit and Hommola, and one event-based method, eMPRess
204 (Legendre et al. 2002; Hommola et al. 2009; Hutchinson et al. 2017; Santichaivekin et al. 2021).
205 Significant congruence in these phylogenies would suggest co-divergence of secretion systems with the
206 bacterial strains. Across most systems and lineages, we observed significant phylogenetic congruence
207 ($p < 0.05$) in at least two of the distance-based methods and in the event-based method (Table 1),
208 consistent with vertical inheritance. To visually validate these results, we constructed species-system
209 tree tanglegrams and heatmaps of pairwise amino acid identity from the alignments used to build each
210 secretion system tree.
211 In most cases, the congruence tests and heatmap visualizations were consistent. The T1SS in
212 *Bartonella* appeared to be an exception, with congruence tests suggesting horizontal transmission.
213 However, upon closer examination, we found that all bee-associated *Bartonella* strains encoded two

214 distinct T1SS copies that formed two well-supported clades, each corresponding to one of the copies.
215 When these clades were analysed separately, both exhibited signatures of vertical inheritance (Table 1,
216 Fig. S10). Phylogenetic topology and additional sequence alignments indicated that these copies are
217 unlikely to represent recent paralogs arising from gene duplication, suggesting instead that they may
218 have independent evolutionary histories. A BLAST search revealed the closest hits of clade 1, at ≥55%
219 amino acid similarity, were from *Bartonella* strains of other hosts such as humans (*Bartonella tamiae*)
220 and bats (*Bartonella sp. HY406*), while that of the clade 2, at ≥30% amino acid similarity, were from
221 strains of distant Beta- and Gammaproteobacteria. This suggests that the clade 1 T1SS was vertically
222 inherited from an ancestral *Bartonellaceae,* while the clade 2 T1SS may have been acquired horizontally
223 by a common ancestor prior to, or during the early stages of association with bees, and transmitted
224 vertically afterwards.
225
226 **Horizontal acquisition of T4SS and T6SS in two lineages**
227 Exceptions to the pattern of strict vertical inheritance were observed for the T4SS of
228 *Snodgrassella* and the T6SS subtype i1 within the *Frischella*/*Gilliamella* lineage (Table 1). In
229 *Snodgrassella*, the T4SS phylogeny showed topological incongruence with the species tree (Fig. 3A)
230 and pairwise T4SS nucleotide similarity was lower and far more variable (mean = 70.1% ± 21.41%)
231 than the overall average nucleotide identity (ANI) among the strains (mean = 93.13% ± 6.35%) (Fig.
232 3B). Notably, *Snodgrassella sp.* B3088 harbours two distinct, highly divergent T4SS loci. BLAST
233 analysis shows the closest hit to locus1 T4SS is the conjugal T4SS from a relative within the
234 *Neisseriaceae*, *Neisseria flavescens* (72.1% identity), while locus2 matches genes from *Paralysiella*
235 *testudines* (70.8%). This suggests the copies arose from independent acquisition rather than duplication.
236 Furthermore, the two *Snodgrassella communis* strains LMG28360 and R-54236 share 100% T4SS
237 identity to one another but very low identity (<40%) to the T4SS of the other strains, including other *S.*
238 *communis* strains, suggesting another instance of independent horizontal acquisition.
239 Two of our four cophylogenetic tests revealed incongruence specifically in T6SS subtype i1
240 within the *Frischella*/*Gilliamella* lineage, which is supported by tanglegram visualization (Fig. 3C). The
241 T6SS is categorised into four subtypes (T6SSi-iv) of which the T6SSi is further subdivided into i1, i2,
242 i3, i4a, i4b, and i5 (Boyer et al. 2009). *Frischella* and *Gilliamella* generally encode a single copy of the
243 T6SS, the majority of which (123/204 strains) corresponds to subtype i2, and which appears to be
244 vertically inherited (Table 1). Subtype i1 is restricted to five *Gilliamella* strains, where subtype i1 is
245 present instead of i2, and two *Frischella* strains, which carry both subtypes. While *Frischella* strains
246 PEB0191 and ESL0167 cluster together in the species tree (98.8% ANI), their subtype i1 T6SSs are
247 highly divergent, sharing only 52.22% nucleotide similarity (Fig. 3D). Conversely, *Gilliamella* strains
248 R-53144 and Choc5-1 show the opposite pattern: despite being more divergent at the genome level
249 (94.0% ANI), their T6SS are nearly identical (99.0% nucleotide similarity). This is consistent with
250 horizontal transfer of T6SS between these strains, but while some T6SSs are known to be encoded on

mobile genetic elements (García-Bayona et al. 2021; Li et al. 2021; Jana et al. 2022; Unni et al. 2022), we found no clear evidence of transposon or plasmid-related genes near the subtype i1 locus in these bee symbionts. Nonetheless, the phylogenetic incongruence and discordant sequence similarities suggests the T6SS subtype i1 can be transmitted horizontally in *Frischella*/*Gilliamella*.

**Gene loss is the dominant force driving secretion system evolution**

To further explore the evolutionary history of these systems, we mapped the binarized distribution of each detected system onto the respective lineage-specific phylogenomic tree and examined patterns of predicted gain and loss events (Fig. 4). Overall, the reconstruction of events shows that secretion system evolution is dominated by repeated losses rather than gains. For instance, although the T1SS is highly prevalent, across both closely and distantly related species within the *Weeksellaceae*, it is entirely absent in the bee *Apibacter* strains (Fig. 4A) suggesting that its loss likely occurred prior to *Apibacter* diversification. In contrast, the T6SS is markedly enriched in bee *Apibacter* strains compared to their relatives. However, within *Apibacter*, several strains lack T6SS, indicating more recent and likely independent loss events.

Possible rare cases of secretion system gain were observed in the *Bartonellaceae*: two strains possessed the T4SS despite it appearing to have been lost early in the bee *Bartonella* lineage (Fig. 4B). These two T4SSs were most similar to those of *Brucella/Ochrobactrum* and the plant root bacteria *Mesorhizobium*, suggesting that they were recent gain events. For *Bombella* and *Gilliamella*, the scattered distribution of their T4SS (Fig. 4C, D), and the lack of clear comparisons with related taxa makes it difficult to deduce their loss or gain events; nonetheless, their closest BLAST hits were to the conjugal T4SS in bacteria such as *Pseudomonas*, *Frigidibacter*, *Acetobacter* and *Salmonella*, which may indicate horizontal acquisition events early in the evolution of these bee-associated lineages. In *Gilliamella* and *Snodgrassella*, multiple independent loss events of the T1SS (25 events in *Gilliamella*, 10 in *Snodgrassella*), T4P (7 events in *Gilliamella*, 1 in *Snodgrassella*) and T6SS (18 events in *Gilliamella*, 27 in *Snodgrassella*) appear to have occurred, but predominantly after the diversification of the bee lineages. The absence of the T1SS, Tad, and T6SS in several divergent branches suggests recent, independent loss events within these lineages. *Snodgrassella* strains typically encode two T6SS: subtype i3 at single genomic locus and subtype i2 that is often split across two separate loci. Both showed strict vertical inheritance (Table 1), in agreement with a previous analyses (Steele et al. 2017). Our phylogenetic reconstruction indicates that a post-diversification ancestor likely encoded both subtypes (Fig. 4E). Multiple independent gene loss events within the lineage subsequently produced strains harbouring either zero, one, or both subtypes.

Overall, while rare gains do occur via horizontal gene transfer, system losses are far more frequent and typically irreversible.

**Secretion system abundances vary with host age and species**

While our genome-based comparative analysis revealed the prevalent secretion systems encoded by the bee gut and bee-associated bacterial genomes, it does not capture their diversity or distribution within natural microbial communities. To address this, we analysed three publicly available metagenomic datasets derived from *Apis cerana* and *Apis mellifera* hosts of varying ages and geographic origins (Ellegaard and Engel 2019; Ellegaard et al. 2020; Wu et al. 2021). We constructed a curated database of secretion system component sequences identified from our genome analysis and used these as queries against the metagenomes under stringent thresholds and thus were able to differentiate between secretion systems encoded by bee specific bacteria and those from other low abundance or transient bacteria. Adjusted for abundance (read depth), most detected secretion systems in metagenomes belonged to members of the bee microbiome (Fig. S11-S13).

Across the metagenomes, the taxonomic distribution of secretion system components from bee bacterial lineages largely mirrored the patterns observed in the genomes (Fig. S14). For example, T9SS was exclusively associated with *Apibacter*, Tad systems with *Frischella* and *Gilliamella*, and T1SS was broadly distributed across most taxa except *Apibacter*. As expected from their dominance among Gram-negative bacteria in individual gut communities, components originating from *Gilliamella* and *Snodgrassella* were highly represented (Fig. S14, S15). Among all metagenomic reads assigned to secretion systems in the dataset from Ellegaard et al. (2019), 60.37% were assigned to *Gilliamella*, 13.04% were assigned to *Snodgrassella*, 21.25% were assigned to *Bartonella*, while fewer than 0.01% were assigned to *Apibacter*, predominantly corresponding to the T9SS. Similar patterns were observed in the datasets from Ellegaard et al. (2020) and Wu et al. (2021), in which 73.06% and 51.68% of the total assigned reads, respectively, corresponded to *Gilliamella*, while 12.19% and 19.93%, respectively, corresponded to *Snodgrassella*.

We next examined whether secretion system diversity varied as a function of host age (Fig. 5). To quantify secretion system abundance, we mapped raw reads to their respective components, normalised by gene length and sequencing depth, and expressed the results as copies per million (CPM). Overall, secretion system diversity declined significantly with host age (Fig. 5A). This pattern was not attributable to shifts in the diversity of contributing taxa, which did not differ significantly across age groups for any classified taxon. Younger bees harboured more even profiles of secretion systems, reflected in higher Shannon and Simpson diversity indices, whereas older bees were dominated by a smaller set of systems, particularly flagella and T5SS (Fig. 5A). To assess the extent to which taxonomic composition contributed to these age-associated differences, we calculated the relative abundance of each secretion system after normalising by the abundance of its contributing taxa (Fig. S16). The decline in T6SS with age occurred mainly within *Snodgrassella* and *Gilliamella*, whereas increases in T1SS, Tad, and flagella were driven largely by *Bartonella*, *Commensalibacter*, and *Frischella*. These patterns suggest that age-related changes in secretion system composition arise not only from shifts in genus-level community structure but also from differences among strains within genera. Functionally, the elevated abundance of Tad, T4P, and T6SS in young bees suggests that early gut communities are

325 primarily engaged in processes of colonization and interbacterial competition. In contrast, the increased
326 abundance of flagellar systems, T1SS, and T5SS in older bees may indicate a functional shift toward
327 motility, sensing and possibly resource exploitation, and host-associated interactions.

328 Comparisons between hosts further revealed striking differences in the diversity and abundance
329 of secretion systems. Among young bees, *A. mellifera* exhibited significantly higher secretion system
330 diversity than *A. cerana*, a difference that corresponded to the greater diversity of contributing taxa in
331 *A. mellifera* (Fig. 5B). Consistent with this, certain secretion systems, including elevated T6SS and
332 T9SS, were more abundant in *A. cerana* and were strongly influenced by the *A. cerana*-characteristic
333 taxon *Apibacter* (Fig. 5B, Fig. S14). Interestingly, these host-specific differences diminished with age.
334 In older bees, both the taxonomic composition and secretion system diversity did not differ significantly
335 between *A. mellifera* and *A. cerana* (Fig. 5C).

## Discussion

Here, we show that the secretion systems of the bee gut microbiome follow an evolutionary trajectory distinct from that described in most free-living and pathogenic bacteria. Rather than being shaped primarily by frequent horizontal gene transfer, secretion systems in bee gut microbiome are predominantly vertically inherited and shaped through recurrent gene loss. This pattern reveals secretion systems as long-term ecological traits, retained and maintained due to their advantage in stable host-associated conditions, rather than being employed in transient adaptive responses to fluctuating environments.

Across multiple bacterial lineages of the bee gut microbiome, we observe secretion systems as primarily vertically inherited indicating long-term co-divergence with their bacterial hosts. Given the prevalence in T1SS, T5SS and T6SS, we infer that the bee gut represents a highly competitive interbacterial niche (Kanonenberg et al. 2018; Coulthurst 2019; Meuskens et al. 2019). Theoretical predictions suggests that in such competitive niches, these interaction traits would be subjected to horizontal transfers (Coyte et al. 2015; Niehus et al. 2015). Consistent with these predictions, horizontal transfer of secretion systems is widely reported in other gut-associated and pathogenic bacteria. In the human gut, T6SS loci located on integrative and conjugative elements (ICEs) can spread across co-resident *Bacteroidales* species within and between hosts (Coyne et al. 2014; García-Bayona et al. 2021). Similarly, in the core mouse gut bacteria *Mucispirillum schaedleri*, a T6SS locus appears to have been acquired from distantly related taxa such as *Campylobacter* or *Helicobacter* (Loy et al. 2017). In pathogens such as *Legionella pneumophila* and *Neisseria gonorrhoeae*, the type IV pili and T4SS, both facilitating DNA uptake, are located on large mobile genetic elements and thus allow their own horizontal dissemination (Hamilton et al. 2005; Lautner et al. 2013). More broadly, systems such as T3SS and flagella are frequently horizontally acquired across diverse taxa, including environmental *Vibrio cholerae*, plant pathogens such as *Erwinia* and *Dickeya*, and animal pathogens including *Salmonella*, *Shigella*, and *Yersinia* (Naum et al. 2009; Brown and Finlay 2011; Abby and Rocha 2012; Morita et al. 2013).

The contrasting evolutionary trajectory in the bee gut microbiome likely reflects fundamental ecological differences of this environment. The bee gut represents a stable, host-restricted environment in which microbial transmission occurs primarily through social interactions (Powell et al. 2014; Kwong and Moran 2016a). In such a closed system, there is limited exposure to external microbial gene pools and, therefore, the observed vertical inheritance likely reflects strong selection for maintaining interaction traits that are consistently beneficial across host generations. Furthermore, we find that host identity exerts only a weak influence on secretion system repertoires despite the described host specificity among strains (Kwong et al. 2014; Powell et al. 2016; Ellegaard et al. 2020; Zhou et al. 2025). This contrasts with free-living and pathogenic systems, where fluctuating environments and host immune pressures drive an evolutionary arms race and promote horizontal acquisition of invasion and immune-evasion mechanisms (Bliven and Maurelli 2016). Notably, the well-described systems in this

373 context, T2SS and T3SS, also happen to be mostly acquired horizontally (Gophna et al. 2003; Brown
374 and Finlay 2011; Abby and Rocha 2012; Cianciotto and White 2017), and are completely missing in the
375 bee gut microbiome. This absence does not imply an absence of host interaction: several systems
376 retained across bee gut lineages have been described to mediate bacterial-host interactions in other
377 bacteria. Rather, these interactions likely occur within a stable, co-evolved framework, where secretion
378 systems are maintained because they provide long-term persistence advantage. Rare horizontal
379 acquisitions, such as the subtype i1 T6SS in *Frischella*/*Gilliamella* and T4SS in some *Snodgrassella*
380 strains, may represent lineage-specific exceptions, providing adaptive fine-tuning under intense
381 competitive pressure.

382 We observe gene losses as the dominant evolutionary force shaping secretion system repertoires.
383 Analysis of gain and loss events revealed repeated, independent losses occurring largely within each bee
384 gut microbiome lineage, rather than simple one-time events at the point of diversification. This pattern
385 indicates that, following evolutionary association with bee gut, different strains within the same lineage
386 have repeatedly and irreversibly lost specific secretion systems while remaining viable and persistent
387 members of the community. Thus, these losses do not only imply a uniform drive toward adaptive
388 streamlining of secretion system repertoires, but also point to multiple lineage- and strain-specific
389 strategies for persistence within the gut ecosystem. For example, while systems such as the T6SS and
390 T1SS are widespread within certain lineages, their repeated loss in distinct branches demonstrates that
391 these systems are not universally required for persistence in the bee gut, and that strains can instead
392 adopt alternative ecological strategies (perhaps niche partitioning or reliance on other interaction
393 systems). In this context, the rare lineage-restricted gain events detected, such as the apparent
394 reacquisition of T4SS variants in some *Bartonella* strains, might best be interpreted as a compensatory
395 mechanism for persistence, rather than evidence of an ongoing repertoire-shaping.

396 Within this evolutionary framework, the secretion systems retained across the bee gut
397 microbiome represent functional solutions to their ecological demands, including colonisation,
398 persistence and competition. Widely distributed across this community are the T1SS and T5SS, the latter
399 being structurally simple yet functionally versatile. In other bacteria, T5aSSs are associated with
400 enzymatic activity and host interaction, T5bSSs with antibacterial competition, and T5cSSs with
401 adhesion, autoaggregation, and biofilm formation (Linke et al. 2006; Aoki et al. 2010; Leo et al. 2010;
402 Serra et al. 2011; Casasanta et al. 2017; Michalska et al. 2017; Trunk et al. 2018; Gucinski et al. 2019;
403 Meuskens et al. 2019; Cuthbert et al. 2022; Qin et al. 2022). Likewise, the T1SS functions as a generalist
404 system mediating biofilm formation, nutrient acquisition, and interbacterial competition through the
405 secretion of adhesins, degradative enzymes, siderophores, and antimicrobials (Thomas et al. 2014;
406 Green and Mecsas 2016; Sassone-Corsi et al. 2016; De León et al. 2017; Kanonenberg et al. 2018; Smith
407 et al. 2018; Mortzfeld et al. 2022; Parker and Davies 2022; Bui et al. 2023; Karbelkar et al. 2024). The
408 high prevalence of these systems and the distinct subtype distribution of the T5SS suggest lineage-
409 specific ecological strategies within the community.

410 An example of this lineage-specific patterning is observed in *Bartonella* and *Commensalibacter*,
411 which are enriched in T1SS and either the T5aSS or T5bSS while lacking other secretion systems.
412 Although these taxa persist alongside the abundant, multi-system *Frischella*/*Gilliamella* and
413 *Snodgrassella*, they exhibit marked dominance in aged and winter bees, a characteristic shift linked to
414 dietary changes, particularly reduced pollen intake (Kešnerová et al. 2020; C. Li et al. 2022). In this
415 context, reliance on T1SS and T5aSS-associated enzymatic functions may confer a competitive
416 advantage in nutrient-limited gut environments. In *Bartonella*, the presence of two T1SS copies,
417 including one likely acquired early in its evolutionary association with bees, suggests a gain of function
418 that may further enhance persistence and competitive fitness in densely populated niches (Galán 2009;
419 Mendonça et al. 2011; Ghosh and O'Connor 2017). By contrast, in *Commensalibacter*, which encodes
420 only T1SS and T5bSS, these systems likely partition roles, with the T1SS mediating secretion of
421 degradative enzymes and adhesins and the T5bSS facilitating competitive interactions. A similar
422 ecological strategy may be observed in *Bombella* (formerly *Saccharibacter* and *Parasaccharibacter*;
423 Fig. S3), which inhabits nutrient-rich niches such as honey and pollen, where T5aSS-mediated
424 enzymatic activity likely enhances access to complex substrates (Zheng et al. 2016; Kwong, Mancenido,
425 et al. 2017; Zheng et al. 2019; Horak et al. 2020). Lastly, the T5cSS, which is mostly restricted to
426 *Snodgrassella*, has been demonstrated to be essential for exporting adhesin to initiate cell-to-cell
427 connections, autoaggregation, and eventually biofilm formation in this lineage (Lariviere et al. 2024).
428 The T6SS is a contractile nanomachine best known for mediating interbacterial competition,
429 but has also been found to function in nutrient acquisition, spatial and temporal organisation, biofilm
430 formation, and stress resistance (Zhang et al. 2011; Weber et al. 2015; Wong et al. 2016; Lin et al. 2017;
431 Speare et al. 2018; Chen et al. 2019; Custodio et al. 2021; Serapio-Palacios et al. 2022; Xiong et al.
432 2024; Liang et al. 2025; Luo et al. 2025). In the bee gut microbiome, studies have shown a role for T6SS
433 in mediating interbacterial competition and bacterial-host interactions: in *Snodgrassella*, T6SS subtype
434 i3 mediates interspecific bacterial competition and are associated with Rhs family of toxins (Steele et
435 al. 2017; Steele and Moran 2021), and the subtype i2 is involved in stimulating host immune functions
436 (Motta et al. 2024). Given the close phylogenetic relationship with the T6SS subtype i2 of *Snodgrassella*
437 (Steele et al. 2017), the T6SS subtype i2 in *Frischella* and *Gilliamella* may also be involved in bacteria-
438 host interaction (Schmidt et al. 2023). In fact, *Gilliamella* and *Snodgrassella* can form layered biofilm
439 close to each other in the ileum (Martinson et al. 2012) yet reports suggest their loads in the gut are not
440 affected by each other in a T6SS-dependent manner (Motta et al. 2024; Quinn et al. 2024). The T6SS
441 subtype i1 in *Frischella*, which is also horizontally acquired by a subset of *Gilliamella* strains lacking
442 subtype i2, is likely to mediate interbacterial competition; accordingly, its horizontal acquisition may
443 represent an adaptive mechanism promoting competitive fitness and persistence in these strains.
444 The less prevalent secretion systems of the bee microbiome further support lineage-specific
445 niche specialization. The T4SS is rare and exhibits patchy distribution across members of *Bartonella*,
446 *Bombella*, *Gilliamella*, and *Snodgrassella*. In *Bartonella*, we find an early loss of the system during the

447 diversification of the bee-associated lineage. Perhaps this loss was a prerequisite for the establishment
448 of a non-pathogenic relationship with their bee hosts, since the T4SS in other *Bartonella* spp. has been
449 shown to secrete virulence factors (Schmid et al. 2004; Schröder et al. 2011; Siamer and Dehio 2015).
450 While their apparent recent acquisition of a T4SS closely related to the pathogenic
451 *Brucella/Ochrobactrum* may contradict the hypothesis, it is possibly an adaptive mechanism to evade
452 the host immune response during the niche establishment (Dugelay et al. 2025; Wei et al. 2025). On the
453 contrary, the T4SSs either retained or independently acquired in *Bombella*, *Gilliamella*, and
454 *Snodgrassella* correspond to conjugative systems, which primarily mediate DNA transfer and are
455 therefore more likely to function in facilitating horizontal gene exchange within these lineages
456 (Waksman 2025; Zhang et al. 2025). The retention of the T2SS- and T3SS-evolutionary related
457 appendages, Tad, T4P, and flagella, further highlights the importance of adhesion, surface sensing,
458 motility, and biofilm formation in lineage-specific niche specialization (Mignolet et al. 2018; Ligthart
459 et al. 2020; Grognot and Taute 2021; Akahoshi and Bevins 2022; Geiger et al. 2024; Whitfield and Brun
460 2024). Tad and T4P are particularly enriched in *Gilliamella* and *Snodgrassella*, the dominant, co-
461 occurring taxa within the ileum, where their adhesion and biofilm functions likely support persistence
462 and facilitate coexistence through spatiotemporal regulation within a confined niche (Y. Li et al. 2022).
463 In lineages lacking Tad and T4P, such as *Bartonella* and *Bombella*, flagella-mediated motility and
464 surface sensing might provide an alternative strategy to locate and colonize available niches.
465 *Apibacter* represents an extreme case of secretion system streamlining, encoding only T6SS and
466 T9SS. The T9SS, restricted to the *Fibrobacteres–Chlorobi–Bacteroidetes* superphylum, mediates
467 secretion of proteins involved in gliding motility, biofilm formation, macromolecule degradation, metal
468 acquisition, and virulence (Lasica et al. 2017; Veith et al. 2017; Rocha et al. 2024; Niu et al. 2025).
469 Given *Apibacter*'s limited capacity for monosaccharide utilization and unlikely involvement in complex
470 polysaccharide degradation (Kwong and Moran 2016b; Praet et al. 2016; Kwong et al. 2018; W. Zhang
471 et al. 2022), the universal presence of T9SS likely support complementary functions in polysaccharide
472 processing. *Apibacter* primarily colonizes the guts of *Apis cerana*, *Apis dorsata*, and bumble bees, co-
473 occurring with *Snodgrassella* in *A. cerana* but occupying niches where *Snodgrassella* is absent in A.
474 dorsata (Kwong and Moran 2016b; Kwong, Medina, et al. 2017; Kwong et al. 2018; W. Zhang et al.
475 2022; Suzuki et al. 2025). The presence of T9SS may allow *Apibacter* to utilize host-derived glycans or
476 colonize specific surfaces inaccessible to *Snodgrassella* and other lineages. Although the role of T9SS
477 in interbacterial interactions remains poorly characterized, it has been shown to secrete antibacterial
478 toxins (Jana et al. 2020). In addition, most of the Rhs toxins in *Apibacter* appear to be associated with
479 the T9SS (Steele et al. 2017; Steele and Moran 2021) suggesting a role in bacterial competition. Perhaps,
480 while T9SS mediates nutrients acquisition and bacterial competition, T6SS further enhances fitness by
481 modulating host immunity, potentially providing a selective advantage in certain hosts.
482 Finally, our metagenomic analysis further underpins how the secretion system repertoires reflect
483 the colonization, persistence and maintenance of the bee gut microbiome as a community. Functionally,

elevated Tad, T4P, and T6SS in young bees might indicate that early gut communities are focused on colonization and interbacterial competition, consistent with the initial assembly of the bee gut microbiota, where successful colonizers rely on adhesion, niche acquisition, and competitive persistence (Powell et al. 2016; Jones et al. 2025). In contrast, flagella dominate in older bees, possibly reflecting a shift toward persistence and resource exploitation in gut environment where the hosts' dietary consumption both changes and declines (Crailsheim et al. 1992).

Taken together, our comparative genomic, evolutionary, and metagenomic analyses converge on a model in which secretion systems in bee gut symbionts are primarily vertically inherited and shaped predominantly by gene losses. At the same time, ecological pressures, particularly host age, species identity, and interbacterial competition, modulate their abundance in vivo. The result is a community that balances evolutionary conservation with ecological plasticity: stable, vertically transmitted systems underpin long-term host association, while dynamic shifts in deployment across host development reinforce functional specialization and gut microbiota stability. Our study establishes a framework for understanding how secretion systems contribute to the stability of the bee gut microbiota and provides a basis for future experimental investigations into how microbial interactions influence community structure and host physiology.

## Materials and methods

### Genomic and metagenomic data

Genomic data were obtained for the core Gram-negative bacterial lineages associated with the bee gut microbiota, including *Apibacter*, *Bartonella*, *Bombella*, *Commensalibacter*, *Gilliamella*, *Frischella*, and *Snodgrassella*. For each lineage, NCBI RefSeq GenBank genome assemblies and corresponding proteomes were downloaded in August 2023. one representative genome from both closely and distantly related bacteria species within each lineage's respective family was included. Tables S1-S5 summarize the bacteria strains, accession numbers, and geographical information associated with the genomic data used in this study. This study also incorporated raw sequencing reads from published bee gut metagenomic datasets (Ellegaard and Engel 2019; Ellegaard et al. 2020; Wu et al. 2021). The reads were retrieved from NCBI using the SRA Toolkit, based on their corresponding BioProject and SRA (SRR) accessions.

### Whole-genome phylogeny

Whole-genome phylogeny for each lineage was constructed based on the concatenated single-copy genes. Single-copy orthologues were inferred from the set of proteomes for each family – *Weeksellaceace* (*Apibacter* and outgroups, 302 orthologues), *Bartonellaceae* (*Bartonella* and outgroups, 247 orthologues), *Acetobacteraceae* (*Bombella* & *Commensalibacter* and outgroups, 24 orthologues), *Orbaceae* (*Frischella* & *Gilliamella* and outgroups, 337 orthologues), and *Neisseriaceae* (*Snodgrassella* and outgroups, 174 orthologues) - using the default parameters of OrthoFinder (Emms and Kelly 2019). The corresponding coding sequences of the orthologues were extracted and aligned based on their codons with the *-alignSeq* parameter of MACSE translational alignment (Ranwez et al. 2011; Ranwez et al. 2021). The aligned sequences were concatenated, and the maximum likelihood tree was inferred using the substitutional model finder of IQ-Tree (Minh et al. 2020). Based on the model finder module, the trees for the *Apibacter*, *Bartonella*, *Bombella* & *Commensalibacter*, and *Gilliamella* & *Frischella* were inferred based on the general time-reversible (GTR) substitutional models with empirical base frequencies, proportion of invariant sites and a free rate model of rate categories. The tree of the *Snodgrassella* lineage was inferred with the symmetric (SYM) substitutional model. All trees were inferred with 1,000 bootstrapping replicates. Phylogenetic trees were visualised on Interactive Tree of Life (iTOL) viewer (Letunic and Bork 2024).

### Detection and distribution of bacterial secretion systems in genomes

To detect the secretion systems in the bee gut microbiome, we employed an optimized approach rather than relying on the broad, general-purpose genome annotations produced by tools such as Prokka or the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) or relying solely on protein identity scores using tools such as BLAST. By targeting secretion system–specific gene sets and architectures, it achieves greater sensitivity and accuracy than both default annotations and previous functional gene

538 surveys of these systems (Barret et al. 2013; Ugarte-Ruiz et al. 2015; Wallner et al. 2021;
539 Liyanapathiranage et al. 2022; Robinson et al. 2022; Robinson et al. 2023).

540 Components of the type I, II, III, IV, V and IX secretion systems and evolutionarily related
541 appendages tight adherence (Tad), type IV pili (T4P) and flagellum were detected with Macsyfinder
542 TXSScan (Abby et al. 2016). For each genome, we generated a proteome FASTA file directly from the
543 corresponding GenBank annotation using an in-house script. This approach was preferred over the use
544 of the NCBI-provided proteome files, which include only representative proteins based on protein
545 accession numbers and may omit duplicated proteins. In contrast, the custom-generated proteomes retain
546 proteins according to their locus tags, allowing accurate identification of duplicated or paralogous
547 components. Components of the type VI secretion systems were detected with SecReT6 (J. Li et al.
548 2015; J. Zhang et al. 2022) which inputs the GenBank file and thus takes into consideration duplicates
549 and genomic locations.

550 TXSScan was used to detect components of each secretion system using HMM profile models.
551 A system was reported as present based on the maximum allowable distance between genes, the
552 minimum number of mandatory components, and the total number of components detected. Secretion
553 systems are typically composed of multiple proteins, among which a highly conserved core set is
554 essential for assembly and function; these core components are generally encoded together within an
555 operon (Green and Mecsas 2016). This detection strategy leverages the characteristic genomic
556 organisation of secretion systems. Core components, which are highly conserved and essential, form the
557 basis of the computational model and are required for a system to be reported. The less conserved or
558 functionally redundant accessory components are included to improve sensitivity while avoiding false
559 positives. The software evaluates whether detected components are co-localized, reflecting their typical
560 arrangement in operons, and applies thresholds for the minimum number of mandatory components and
561 overall system completeness. Likewise, the genomic organisation is used to further infer systems being
562 encoded in a single locus, where most components occur together, or as multi-loci with components
563 distributed across several genomic regions. The maximum number of genes between the components
564 was set to 5 for T1SS, T2SS, T4P, Tad, and T9SS; 10 for T3SS; 20 for Flagellum, and 1 for the subtypes
565 of T5SS. The minimum number out of the total mandatory components was set as 3 out of 3 for T1SS;
566 3 out of 4 for T2SS; 7 out of 9 for T3SS; 4 out of 6 for T4P; 4 out of 7 for Tad; 9 out of 11 for Flagellum;
567 7 out of 8 for T9SS; and 1 for the subtypes of T5SS. Lastly, the minimum required total number of
568 components was set to 3 out of 3 for T1SS; 6 out of 13 for T2SS; 7 out of 9 for T3SS; 5 out of 18 for
569 T4P; 6 out of 13 for Tad; 10 out of 11 for Flagellum; 7 out of 11 for T9SS; and 1 for each subtype of
570 T5SS.

571 SecReT6 was used to detect T6SS components employing a database of experimentally
572 validated protein sequences. Like the TXSScan, this method also considers the minimum number of
573 T6SS components and their genomic organisation. Moreover, phylogenetic analysis of T6SS core
574 components indicates the presence of at least, four T6SS subtypes (T6SSi-iv) across different taxa

575 (Boyer et al. 2009). The canonical T6SSi (subdivided into i1, i2, i3, i4a, i4b, and i5) occurs primarily in
576 *Proteobacteria*, T6SSii in *Francisella* pathogenicity islands, T6SSiii in *Bacteriodota*, and T6SSiv in
577 *Amoebophilus* (Boyer et al. 2009; Bröms et al. 2010; Russell et al. 2014; Bayer-Santos et al. 2019). To
578 detect T6SS in the genomes, Initial detections were considered for components with protein identity of
579 ≥ 30% to validated components, and a group of co-localized components was regarded as a locus when
580 consisting of ≥ 5 different types of components. The TssB protein, which has been shown to be sufficient
581 to classify the T6SS (J. Li et al. 2015) was extracted and used to build the phylogeny for classification
582 on the SecReT6 web server. Only T6SSi and T6SSiii existed in the datasets. Based on that, a T6SS was
583 categorized as complete/putative when it had ≥ 12 for subtype T6SSiii (disregarding the less well-
584 characterized TssQ and TssR) and ≥ 13 for subtype T6SSi (noting that TssI occurred farther away from
585 the locus in some genomes). The detected proteins were further analysed manually and visually using
586 Geneious Prime v2024.0.7 to identify possible false positives, as previously reported (Steele and Moran
587 2021) with some modifications. Briefly, a ≥ 50% identity cutoff and a minimum length of 800 amino
588 acids were applied to filter TssH hits, which mostly occurred in loci with two or more detected TssH.
589 An amino acid length of ≥300 was used to filter out TssI hits. T6SSiii lacks TssL; however, at 30%
590 amino acid identity, OmpA-like proteins (Russell et al. 2014; Zoued et al. 2014; Wang et al. 2018) were
591 detected and considered false positives, thus filtered out.
592
593 **Phylogeny of secretion systems**
594 Phylogenetic trees were constructed for secretion systems detected in each lineage using the
595 concatenated protein sequences of components that appear as single copy in a locus/region. Basically, a
596 system was considered occurring as a locus when the number of unique components as mentioned above
597 co-occurred. In most instances, most, if not all, components co-occurred in a single locus and thus
598 assigned a locus number based on their genomic coordinates. In some instances, such as the T6SS, set
599 of unique components co-occurred in 2 genomic locations but were considered as a single system as the
600 set of components on either locus constituted the total component count for that system. The
601 phylogenetic trees were constructed based on the components within the locus, either intact or combined.
602 To this end, type I secretion system (T1SS) phylogenies for *Bartonella* and *Bombella* &
603 *Commensalibacter* were inferred using concatenated ABC transporter (Abc), membrane fusion protein
604 (Mfp), and outer membrane factor (Omf) proteins, whereas T1SS phylogenies for *Frischella* &
605 *Gilliamella* and *Snodgrassella* were constructed using Mfp and Omf only. Flagellar phylogenies were
606 reconstructed using FlgB, FliE, and the shared type III secretion system components SctJNRSTUV for
607 Bartonella; FlgBC, FliE, and SctJNQRSTUV for *Bombella* & *Commensalibacter*; and FliE together
608 with SctJRSTUV for *Frischella* & *Gilliamella*. Type IV pilus (T4P) phylogenies were inferred using
609 FimT and PilBCDMOQTVW for *Frischella* & *Gilliamella*, and FimT together with
610 PilBCEMNOPQTUVWX for *Snodgrassella*. The tight adherence (Tad) system phylogeny for
611 *Snodgrassella* was reconstructed using concatenated RcpA and TadABCDFV proteins. For *Apibacter*,

612 the type IX secretion system (T9SS) phylogeny was inferred from concatenated GldJKLMN, PorQUV,
613 and SprAET proteins. Finally, type VI secretion system (T6SS) phylogenies were constructed using
614 concatenated TssBCEFGHIKNOP for *Apibacter*, TssABCDEGHJKLM for *Frischella* & *Gilliamella*,
615 and TssABCDEFGHIKLM for *Snodgrassella*.

616 T6SS phylogenies were constructed in two ways: (i) subtype-specific trees, based on the
617 phylogenetic classification of the conserved TssB protein; and (ii) a combined T6SS (all subtypes) tree,
618 generated by concatenating all single-copy T6SS loci present in each genome without separating
619 subtypes into independent trees. In this combined tree, loci belonging to different T6SS subtypes remain
620 distinct concatenated components in the alignment but are included together in a single phylogeny. The
621 locus tags of each component were used to extract their corresponding protein sequences. The sequences
622 were aligned using MUSCLE (Edgar 2022), concatenated, and phylogenetic trees were constructed
623 using IQTree (Minh et al. 2020). The trees for each lineage were constructed with the LG+F+I+G4
624 model, and each tree was constructed with 1,000 bootstrap replicates. For each system, the individual
625 component gene trees were also inferred under the same model as ultrafast booststrap distributions with
626 1,000 bootstrap replicates using IQ-Tree. Phylogenetic trees were visualized on Interactive Tree of Life
627 (iTOL) viewer (Letunic and Bork 2024).
628

629 **Test for phylogenetic congruence and reconciliation**

630 The coevolution relationships between the secretions systems and their bacterial hosts were
631 tested with three separate distance-based cophylogenetic methods, ParaFit (Legendre et al. 2002), PACo
632 (Hutchinson et al. 2017), and Hommola (Hommola et al. 2009), and one event-based method, eMPRess
633 (Santichaivekin et al. 2021). These methods were originally described to test for co-speciation between
634 parasites and their hosts by fitting their respective phylogenetic distances.

635 Briefly, these pattern-based methods use the secretion system (treated as the parasite) and
636 lineage (treated as the host) phylogenetic distances to test whether an interaction between the two
637 phylogenies is due to chance alone. Thus, using the R script described by (Nishida and Ochman 2021),
638 the phylogenetic distance matrices were computed from each phylogenetic tree, transformed by
639 principal coordinates and superimposed to assess the degree of evolutionary congruence while
640 accounting for lineage-secretion system associations. Each lineage–secretion system pair was analysed
641 through 10 simulations, with 999 permutations performed for each simulation. The tests were performed
642 as Type I error tests with the null hypothesis that the lineage and their secretion systems evolved
643 independently, i.e., a *p-value* $\leq 0.05$ signifies congruence between both phylogenies.

644 The event-based method reconciles the two phylogenies by mapping the system phylogeny onto
645 the bacterial lineage phylogeny. Reconciliations were performed with the loss cost fixed at 1, while
646 duplication and transfer costs were independently varied from 1 to 3, generating nine cost combinations
647 in total. For each cost scheme, a reconciliation using maximum parsimony was computed and statistical
648 significance was assessed for 100 sampling using the built-in randomization (p-value) module. The

649 resulting p-values from the nine analyses were averaged to obtain a single summary estimate of
650 cophylogenetic significance.
651
652 **Metagenome assembly and annotation**
653 Publicly available raw metagenome data and metadata from the studies (Ellegaard and Engel
654 2019; Ellegaard et al. 2020; Wu et al. 2021) were prefetched from the NCBI SRR repository. These are
655 metagenomes obtained from either *Apis mellifera* or *Apis cerana*. For each metagenomic sample, paired-
656 end reads were extracted, and the quality of the raw reads was checked with fastQC. Since each
657 metagenome was sequenced with different technologies, trim galore was used to trim the reads using
658 the auto default settings, which include an initial quality check before trimming detected the adaptor.
659 QC logs were manually inspected to ensure the trimmed adaptors corresponded with the reported
660 adaptors in their source studies. Contaminating host-derived reads were filtered out by first mapping the
661 raw reads against a database of the source host genomes, *A. mellifera* (PRJNA471592) or *A. cerana*
662 (PRJNA324433), using bowtie2 (Langmead and Salzberg 2012). Samtools (Danecek et al. 2021) was
663 used to generate bam-files of the filtered reads (*view -b -f 12 -F 256*), sort, and extract the paired-reads.
664 Each pair of unfiltered and filtered reads was profiled for their represented communities using Kraken2
665 (Wood et al. 2019), visualized as a community profile and manually inspected to validate the removal
666 of host-derived reads. The filtered paired-end reads were assembled using the default setting of
667 MEGAHIT assembler (D. Li et al. 2015). The resulting contigs were filtered to remove contigs shorter
668 than 1000bp. Protein coding sequences of the contigs were predicted using Prokka version 1.14
669 (Seemann 2014).
670
671 **Detection and distribution of secretion systems in metagenome data**
672 Secretion system components were identified in metagenomic assemblies based on protein
673 sequence similarity. A custom BLAST database was constructed using all predicted protein sequences
674 of secretion system components from the individual genome analysis, with each protein linked to its
675 corresponding genome metadata. Predicted protein sequences from the metagenomes were queried
676 against this database using BLASTp (e-value < 1e-20, minimum identity > 40%), retaining only the top
677 hit for each query. To distinguish bee-specific components, we examined the distribution of sequence
678 identities for each component and established thresholds based on modal distributions. A minimum
679 threshold of ≥70% identity was required for considering a component as bee specific. Coding sequences
680 corresponding to each component were further taxonomically classified using Kraken2 (Wood et al.
681 2019) and validated manually through BLASTn searches against the NCBI non-redundant nucleotide
682 database. Manual inspection confirmed that the classified genus of contigs matched the genus of the
683 corresponding reference protein, supporting their assignment as bee-associated bacterial components.
684 Potential false positives were noted for some components, particularly ATPases of the type III secretion
685 system (T3SS) and flagellar system, which occasionally occurred as singletons in taxa where no

686 complete system was detected in genome analyses. To avoid this, we required at least two unique
687 components for a secretion system to be considered present in each metagenome, removing hits based
688 solely on single components.

689 The relative abundance of each secretion system was estimated from the summed abundances
690 of its constituent components. All predicted coding sequences were mapped against quality-filtered
691 metagenomic reads using BWA-MEM2 (Vasimuddin et al. 2019). Coverage was calculated with
692 BEDTools (Quinlan and Hall 2010), and read counts were normalized by gene length. To account for
693 differences in sequencing depth across samples, abundances were further normalized as copies per
694 million reads (CPM). From the overall abundance matrix, we extracted secretion system components to
695 calculate (i) the total abundance of each secretion system per metagenome, and (ii) diversity indices of
696 secretion systems within and across samples.

## Author contributions

**Samuel A. Acheampong**: Conceptualization; Data curation; Software; Formal analysis; Validation; Investigation; Visualization; Methodology; Writing—original draft; Writing—review and editing. **Waldan K. Kwong**: Conceptualization; Resources; Supervision; Funding acquisition; Validation; Writing—original draft; Project administration; Writing—review and editing.

## Competing interests

The authors declare no competing interests.

## Acknowledgements

The authors thank members of the Microbial Genomics and Symbiosis lab for constructive discussions, and Sara Louro for assistance in curating metagenome data. This work was supported by Fundação para a Ciência e a Tecnologia (FCT) doctoral fellowship UI/BD/154902/2023 to SAA, and European Molecular Biology Organization Installation Grant 5045-2022, European Research Council Starting Grant 101042912, and FCT individual grant CEECIND/01358/2021 to WKK.

## References


Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EPC. 2016. Identification of protein secretion systems in bacterial genomes. *Sci. Rep.* 6:23080.

Abby SS, Rocha EPC. 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLOS Genet.* 8:e1002983.

Akahoshi DT, Bevins CL. 2022. Flagella at the host-microbe Interface: key functions intersect with redundant responses. *Front. Immunol.* 13:828758.

Aoki SK, Diner EJ, de Roodenbeke C t'Kint, Burgess BR, Poole SJ, Braaten BA, Jones AM, Webb JS, Hayes CS, Cotter PA, et al. 2010. A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature* 468:439–442.

Baishya J, Wakeman CA. 2019. Selective pressures during chronic infection drive microbial competition and cooperation. *Npj Biofilms Microbiomes* 5:16.

Barret M, Egan F, O'Gara F. 2013. Distribution and diversity of bacterial secretion systems across metagenomic datasets. *Environ. Microbiol. Rep.* 5:117–126.

Bayer-Santos E, Ceseti L de M, Farah CS, Alvarez-Martinez CE. 2019. Distribution, function and regulation of type 6 secretion systems of *Xanthomonadales*. *Front. Microbiol.* 10:1635.

Bliven KA, Maurelli AT. 2016. Evolution of bacterial pathogens within the human host. *Microbiol. Spectr.* 4:10.1128/microbiolspec.vmbf-0017–2015.

Bontemps Z, Paranjape K, Guy L. 2024. Host–bacteria interactions: ecological and evolutionary insights from ancient, professional endosymbionts. *FEMS Microbiol. Rev.* 48:fuae021.

Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? *BMC Genomics* 10:104.

Bröms J, Sjöstedt A, Lavander M. 2010. The role of the *Francisella Tularensis* pathogenicity island in type VI secretion, intracellular survival, and modulation of host cell signaling. *Front. Microbiol.* 1.

Brown NF, Finlay B. 2011. Potential origins and horizontal transfer of type III secretion systems and effectors. *Mob. Genet. Elem.* 1:118–121.

Bui D-C, Luo T, McBride JW. 2023. Type 1 secretion system and effectors in *Rickettsiales*. *Front. Cell. Infect. Microbiol.* 13:1175688.

Büttner D, He SY. 2009. Type III protein secretion in plant pathogenic bacteria. *Plant Physiol.* 150:1656–1664.

Casasanta MA, Yoo CC, Smith HB, Duncan AJ, Cochrane K, Varano AC, Allen-Vercoe E, Slade DJ. 2017. A chemical and biological toolbox for Type Vd secretion: characterization of the phospholipase A1 autotransporter FplA from *Fusobacterium nucleatum*. *J. Biol. Chem.* 292:20240–20254.

Chen C, Yang X, Shen X. 2019. Confirmed and potential roles of bacterial T6SSs in the intestinal ecosystem. *Front. Microbiol.* 10.

Christian N, Whitaker BK, Clay K. 2015. Microbiomes: unifying animal and plant systems through the lens of community ecology theory. *Front. Microbiol.* 6:869.

Cianciotto NP, White RC. 2017. Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infect. Immun.* 85:10.1128/iai.00014-17.

Coulthurst S. 2019. The Type VI secretion system: a versatile bacterial weapon. *Microbiology* 165:503–515.

Coyne MJ, Zitomersky NL, McGuire AM, Earl AM, Comstock LE. 2014. Evidence of extensive DNA transfer between *Bacteroidales* species within the human gut. *mBio* 5:10.1128/mbio.01305-14.

Coyte KZ, Schluter J, Foster KR. 2015. The ecology of the microbiome: networks, competition, and stability. *Science* 350:663–666.

Crailsheim K, Schneider LHW, Hrassnigg N, Bühlmann G, Brosch U, Gmeinbauer R, Schöffmann B. 1992. Pollen consumption and utilization in worker honeybees (*Apis mellifera carnica*): dependence on individual age and function. *J. Insect Physiol.* 38:409–419.

Custodio R, Ford RM, Ellison CJ, Liu G, Mickute G, Tang CM, Exley RM. 2021. Type VI secretion system killing by commensal Neisseria is influenced by expression of type four pili.Filloux A, Storz G, Filloux A, editors. *eLife* 10:e63755.

Cuthbert BJ, Hayes CS, Goulding CW. 2022. Functional and structural diversity of bacterial contact-dependent growth inhibition effectors. *Front. Mol. Biosci.* 9:866854.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* 10:giab008.

De León KB, Zane GM, Trotter VV, Krantz GP, Arkin AP, Butland GP, Walian PJ, Fields MW, Wall JD. 2017. Unintended laboratory-driven evolution reveals genetic requirements for biofilm formation by *Desulfovibrio vulgaris Hildenborough*. *mBio* 8:10.1128/mbio.01696-17.

Deakin WJ, Broughton WJ. 2009. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat. Rev. Microbiol.* 7:312–320.

Denise R, Abby SS, Rocha EPC. 2020. The evolution of protein secretion systems by co-option and tinkering of cellular machineries. *Trends Microbiol.* 28:372–386.

Dugelay C, Celli J, Terradot L. 2025. The *Brucella* type IV secretion system and effector proteins. *microLife* 6:uqaf036.

Edgar RC. 2022. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* 13:6968.

Ellegaard KM, Engel P. 2019. Genomic diversity landscape of the honey bee gut microbiota. *Nat. Commun.* 10:446.

Ellegaard KM, Suenami S, Miyazaki R, Engel P. 2020. Vast differences in strain-level diversity in the gut microbiota of two closely related honey bee species. *Curr. Biol.* 30:2520-2531.e7.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238.

Galán JE. 2009. Common themes in the design and function of bacterial effectors. *Cell Host Microbe* 5:571–579.

790 Ganesan R, Wierz JC, Kaltenpoth M, Flórez LV. 2022. How it all begins: bacterial factors mediating the
791 colonization of invertebrate hosts by beneficial symbionts. *Microbiol. Mol. Biol. Rev.*
792 86:e00126-21.

793 García-Bayona L, Coyne MJ, Comstock LE. 2021. Mobile type VI secretion system loci of the gut
794 Bacteroidales display extensive intra-ecosystem transfer, multi-species spread and geographical
795 clustering. *PLOS Genet.* 17:e1009541.

796 Geiger CJ, Wong GCL, O'Toole GA. 2024. A bacterial sense of touch: T4P retraction motor as a means
797 of surface sensing by *Pseudomonas aeruginosa* PA14. *J. Bacteriol.* 206:e00442-23.

798 Ghosh S, O'Connor TJ. 2017. Beyond paralogs: the multiple layers of redundancy in bacterial
799 pathogenesis. *Front. Cell. Infect. Microbiol.* 7:467.

800 Ghoul M, Mitri S. 2016. The ecology and evolution of microbial competition. *Trends Microbiol.* 24:833–
801 845.

802 Gophna U, Ron EZ, Graur D. 2003. Bacterial type III secretion systems are ancient and evolved by
803 multiple horizontal-transfer events. *Gene* 312:151–163.

804 Gorter FA, Manhart M, Ackermann M. 2020. Understanding the evolution of interspecies interactions
805 in microbial communities. *Philos. Trans. R. Soc. B Biol. Sci.* 375:20190256.

806 Green ER, Mecsas J. 2016. Bacterial secretion systems: an overview. *Microbiol. Spectr.*
807 4:10.1128/microbiolspec.vmbf-0012–2015.

808 Grognot M, Taute KM. 2021. More than propellers: how flagella shape bacterial motility behaviors.
809 *Curr. Opin. Microbiol.* 61:73–81.

810 Gucinski GC, Michalska K, Garza-Sánchez F, Eschenfeldt WH, Stols L, Nguyen JY, Goulding CW,
811 Joachimiak A, Hayes CS. 2019. Convergent evolution of the Barnase/EndoU/Colicin/RelE
812 (BECR) fold in antibacterial tRNase toxins. *Structure* 27:1660-1674.e5.

813 Hamilton HL, Domínguez NM, Schwartz KJ, Hackett KT, Dillard JP. 2005. *Neisseria gonorrhoeae*
814 secretes chromosomal DNA via a novel type IV secretion system. *Mol. Microbiol.* 55:1704–
815 1721.

816 Harms A, Liesch M, Körner J, Québatte M, Engel P, Dehio C. 2017. A bacterial toxin-antitoxin module
817 is the origin of inter-bacterial and inter-kingdom effectors of *Bartonella*. *PLOS Genet.*
818 13:e1007077.

819 Hommola K, Smith JE, Qiu Y, Gilks WR. 2009. A permutation test of host–parasite cospeciation. *Mol.*
820 *Biol. Evol.* 26:1457–1468.

821 Horak RD, Leonard SP, Moran NA. 2020. Symbionts shape host innate immunity in honeybees. *Proc.*
822 *R. Soc. B Biol. Sci.* 287:20201184.

823 Hou K, Wu Z-X, Chen X-Y, Wang J-Q, Zhang D, Xiao C, Zhu D, Koya JB, Wei L, Li J, et al. 2022.
824 Microbiota in health and diseases. *Signal Transduct. Target. Ther.* 7:135.

825 Hutchinson MC, Cagua EF, Balbuena JA, Stouffer DB, Poisot T. 2017. paco: implementing Procrustean
826 approach to cophylogeny in R. *Methods Ecol. Evol.* 8:932–940.

827 Jamali H, Akrami F, Layeghkhavidaki H, Bouakkaz S. 2025. Bacterial protein secretion systems:
828 mechanisms, functions, and roles in virulence. *Microb. Pathog.* 206:107790.

Jana B, Keppel K, Fridman CM, Bosis E, Salomon D. 2022. Multiple T6SSs, mobile auxiliary modules, and effectors revealed in a systematic snalysis of the *Vibrio parahaemolyticus* pan-genome. *mSystems* 7:e00723-22.

Jana B, Salomon D, Bosis E. 2020. A novel class of polymorphic toxins in *Bacteroidetes*. *Life Sci. Alliance* 3:e201900631.

Jones KR, Song Y, Rinaldi SS, Moran NA. 2025. Effects of priority on strain-level composition of the honey bee gut community. *Appl. Environ. Microbiol.* 91:e00828-25.

Kanonenberg K, Spitz O, Erenburg IN, Beer T, Schmitt L. 2018. Type I secretion system—it takes three and a substrate. *FEMS Microbiol. Lett.* 365:fny094.

Karbelkar AA, Font M, Smith TJ, Sondermann H, O'Toole GA. 2024. Reconstitution of a biofilm adhesin system from a sulfate-reducing bacterium in *Pseudomonas fluorescens*. *Proc. Natl. Acad. Sci.* 121:e2320410121.

Kešnerová L, Emery O, Troilo M, Liberti J, Erkosar B, Engel P. 2020. Gut microbiota structure differs between honeybees in winter and summer. *ISME J.* 14:801–814.

Kešnerová L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, Engel P. 2017. Disentangling metabolic functions of bacteria in the honey bee gut. *PLOS Biol.* 15:e2003467.

Kwong WK, Engel P, Koch H, Moran NA. 2014. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci.* 111:11509–11514.

Kwong WK, Mancenido AL, Moran NA. 2017. Immune system stimulation by the native gut microbiota of honey bees. *R. Soc. Open Sci.* 4:170003.

Kwong WK, Medina LA, Koch H, Sing K-W, Soh EJY, Ascher JS, Jaffé R, Moran NA. 2017. Dynamic microbiome evolution in social bees. *Sci. Adv.* 3:e1600513.

Kwong WK, Moran NA. 2016a. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* 14:374–384.

Kwong WK, Moran NA. 2016b. Apibacter adventoris gen. nov., sp. nov., a member of the phylum Bacteroidetes isolated from honey bees. *Int. J. Syst. Evol. Microbiol.* 66:1323–1329.

Kwong WK, Steele MI, Moran NA. 2018. Genome Sequences of Apibacter spp., Gut Symbionts of Asian Honey Bees. *Genome Biol. Evol.* 10:1174–1179.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357–359.

Lariviere PJ, Ashraf AHMZ, Gifford I, Tanguma SL, Barrick JE, Moran NA. 2024. Virulence-linked adhesin drives mutualist colonization of the bee gut via biofilm formation. :2024.10.14.618124. Available from: https://www.biorxiv.org/content/10.1101/2024.10.14.618124v1

Lasica AM, Ksiazek M, Madej M, Potempa J. 2017. The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front. Cell. Infect. Microbiol.* 7:215.

Lautner M, Schunder E, Herrmann V, Heuner K. 2013. Regulation, integrase-dependent excision, and horizontal transfer of genomic islands in *Legionella pneumophila*. *J. Bacteriol.* 195:1583–1597.

Lee FJ, Miller KI, McKinlay JB, Newton ILG. 2018. Differential carbohydrate utilization and organic acid production by honey bee symbionts. *FEMS Microbiol. Ecol.* 94:fiy113.

Legendre P, Desdevises Y, Bazin E. 2002. A statistical test for host–parasite coevolution. *Syst. Biol.* 51:217–234.

Leo JC, Elovaara H, Bihan D, Pugh N, Kilpinen SK, Raynal N, Skurnik M, Farndale RW, Goldman A. 2010. First analysis of a bacterial collagen-binding protein with collagen toolkits: promiscuous binding of YadA to collagens may explain how YadA interferes with host processes. *Infect. Immun.* 78:3226–3236.

Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 52:W78–W82.

Li C, Tang M, Li X, Zhou X. 2022. Community dynamics in structure and function of honey bee gut bacteria in response to winter dietary shift. *mBio* 13:e01131-22.

Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinforma. Oxf. Engl.* 31:1674–1676.

Li DY, Liu YL, Liao XJ, He TT, Sun SS, Nie P, Xie HX. 2021. Identification and characterization of EvpQ, a novel T6SS effector encoded on a mobile genetic element in *Edwardsiella piscicida*. *Front. Microbiol.* 12.

Li J, Yao Y, Xu HH, Hao L, Deng Z, Rajakumar K, Ou H-Y. 2015. SecReT6: a web-based resource for type VI secretion systems found in bacteria. *Environ. Microbiol.* 17:2196–2202.

Li Y, Leonard SP, Powell JE, Moran NA. 2022. Species divergence in gut-restricted bacteria of social bees. *Proc. Natl. Acad. Sci.* 119:e2115013119.

Li YG, Hu B, Christie PJ. 2019. Biological and structural diversity of type IV secretion systems. *Microbiol. Spectr.* 7:10.1128/microbiolspec.psib-0012–2018.

Liang C, Lin L, Wang X, Zhu W. 2025. The secretion of *Pseudomonas* unconventional peroxidase facilitates extracellular carbon acquisition from heterogeneous lignin. *Commun. Biol.* 8:1318.

Ligthart K, Belzer C, de Vos WM, Tytgat HLP. 2020. Bridging bacteria and the gut: functional aspects of type IV pili. *Trends Microbiol.* 28:340–348.

Lin J, Zhang W, Cheng J, Yang X, Zhu K, Wang Y, Wei G, Qian P-Y, Luo Z-Q, Shen X. 2017. A *Pseudomonas* T6SS effector recruits PQS-containing outer membrane vesicles for iron acquisition. *Nat. Commun.* 8:14888.

Linke D, Riess T, Autenrieth IB, Lupas A, Kempf VAJ. 2006. Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol.* 14:264–270.

Liyanapathiranage P, Wagner N, Avram O, Pupko T, Potnis N. 2022. Phylogenetic distribution and evolution of type VI secretion system in the genus *Xanthomonas*. *Front. Microbiol.* 13:840308.

Loy A, Pfann C, Steinberger M, Hanson B, Herp S, Brugiroux S, Gomes Neto JC, Boekschoten MV, Schwab C, Urich T, et al. 2017. Lifestyle and horizontal gene transfer-mediated evolution of *Mucispirillum schaedleri*, a core member of the murine gut microbiota. *mSystems* 2:10.1128/msystems.00171-16.

Luo C, Gu H, Pan D, Zhao Y, Zheng A, Zhu H, Zhang C, Li C, Zhang J, Chen C, et al. 2025. *Pseudomonas aeruginosa* T6SS secretes an oxygen-binding hemerythrin to facilitate competitive growth under microaerobic conditions. *Microbiol. Res.* 293:128052.

Maphosa S, Moleleki LN, Motaung TE. 2023. Bacterial secretion system functions: evidence of interactions and downstream implications. *Microbiology* 169:001326.

Martinson VG, Moy J, Moran NA. 2012. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl. Environ. Microbiol.* 78:2830–2840.

McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF, et al. 2013. Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci.* 110:3229–3236.

Mendonça AG, Alves RJ, Pereira-Leal JB. 2011. Loss of genetic redundancy in reductive genome evolution. *PLoS Comput. Biol.* 7:e1001082.

Meuskens I, Saragliadis A, Leo JC, Linke D. 2019. Type V secretion systems: an overview of passenger domain functions. *Front. Microbiol.* 10:1163.

Michalska K, Gucinski GC, Garza-Sánchez F, Johnson PM, Stols LM, Eschenfeldt WH, Babnigg G, Low DA, Goulding CW, Joachimiak A, et al. 2017. Structure of a novel antibacterial toxin that exploits elongation factor Tu to cleave specific transfer RNAs. *Nucleic Acids Res.* 45:10306–10320.

Mignolet J, Panis G, Viollier PH. 2018. More than a Tad: spatiotemporal control of *Caulobacter* pili. *Curr. Opin. Microbiol.* 42:79–86.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37:1530–1534.

Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42:165–190.

Morita M, Yamamoto S, Hiyoshi H, Kodama T, Okura M, Arakawa E, Alam M, Ohnishi M, Izumiya H, Watanabe H. 2013. Horizontal gene transfer of a genetic island encoding a type III secretion system distributed in *Vibrio cholerae*. *Microbiol. Immunol.* 57:334–339.

Mortzfeld BM, Palmer JD, Bhattarai SK, Dupre HL, Mercado-Lubio R, Silby MW, Bang C, McCormick BA, Bucci V. 2022. Microcin MccI47 selectively inhibits enteric bacteria and reduces carbapenem-resistant *Klebsiella pneumoniae* colonization in vivo when administered via an engineered live biotherapeutic. *Gut Microbes* 14:2127633.

Motta EVS, Lariviere PJ, Jones KR, Song Y, Moran NA. 2024. Type VI secretion systems promote intraspecific competition and host interactions in a bee gut symbiont. *Proc. Natl. Acad. Sci.* 121:e2414882121.

Motta EVS, Moran NA. 2024. The honeybee microbiota and its impact on health and disease. *Nat. Rev. Microbiol.* 22:122–137.

Naum M, Brown EW, Mason-Gamer RJ. 2009. Phylogenetic evidence for extensive horizontal gene transfer of type III secretion system genes among enterobacterial plant pathogens. *Microbiology* 155:3187–3199.

Nazir R, Mazurier S, Yang P, Lemanceau P, van Elsas JD. 2017. The ecological role of type three secretion systems in the interaction of bacteria with fungi in soil and related habitats is diverse and context-dependent. *Front. Microbiol.* 8.

947 Nicholson KR, Champion PA. 2022. Bacterial secretion systems: Networks of pathogenic regulation
948 and adaptation in mycobacteria and beyond. *PLoS Pathog.* 18:e1010610.

949 Niehus R, Mitri S, Fletcher AG, Foster KR. 2015. Migration and horizontal gene transfer divide
950 microbial genomes into multiple niches. *Nat. Commun.* 6:8924.

951 Nishida AH, Ochman H. 2021. Captivity and the co-diversification of great ape microbiomes. *Nat.*
952 *Commun.* 12:5632.

953 Niu P, Guo R, Ding C, Yu S. 2025. Advances in the type IX secretion system: an exploration focusing
954 on *Riemerella anatipestifer*. *Curr. Res. Microb. Sci.* 9:100404.

955 Okazaki S, Kaneko T, Sato S, Saeki K. 2013. Hijacking of leguminous nodulation signaling by the
956 rhizobial type III secretion system. *Proc. Natl. Acad. Sci.* 110:17131–17136.

957 Parker JK, Davies BW. 2022. Microcins reveal natural mechanisms of bacterial manipulation to inform
958 therapeutic development. *Microbiology* 168:001175.

959 Peabody CR, Chung YJ, Yen M-R, Vidal-Ingigliardi D, Pugsley AP, Saier MH. 2003. Type II protein
960 secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology*
961 149:3051–3072.

962 Perreau J, Moran NA. 2022. Genetic innovations in animal–microbe symbioses. *Nat. Rev. Genet.* 23:23–
963 39.

964 Powell E, Ratnayeke N, Moran NA. 2016. Strain diversity and host specificity in a specialized gut
965 symbiont of honeybees and bumblebees. *Mol. Ecol.* 25:4461–4471.

966 Powell JE, Martinson VG, Urban-Mead K, Moran NA. 2014. Routes of acquisition of the gut microbiota
967 of the honey bee *Apis mellifera*. *Appl. Environ. Microbiol.* 80:7378–7387.

968 Praet J, Aerts M, Brandt ED, Meeus I, Smagghe G, Vandamme P. 2016. Apibacter mensalis sp. nov.: a
969 rare member of the bumblebee gut microbiota. *Int. J. Syst. Evol. Microbiol.* 66:1645–1651.

970 Qin J, Hong Y, Morona R, Totsika M. 2022. Cysteine-dependent conformational heterogeneity of
971 *Shigella flexneri* autotransporter IcsA and implications of its function. *Microbiol. Spectr.*
972 10:e03410-22.

973 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features.
974 *Bioinformatics* 26:841–842.

975 Quinn A, El Chazli Y, Escrig S, Daraspe J, Neuschwander N, McNally A, Genoud C, Meibom A, Engel
976 P. 2024. Host-derived organic acids enable gut colonization of the honey bee symbiont
977 Snodgrassella alvi. *Nat. Microbiol.* 9:477–489.

978 Ranwez V, Chantret N, Delsuc F. 2021. Aligning protein-coding nucleotide sequences with MACSE. In:
979 Katoh K, editor. Multiple Sequence Alignment: Methods and Protocols. New York, NY:
980 Springer US. p. 51–70. Available from: https://doi.org/10.1007/978-1-0716-1036-7_4

981 Ranwez V, Harispe S, Delsuc F, Douzery EJP. 2011. MACSE: multiple alignment of coding sequences
982 accounting for frameshifts and stop codons. *PLOS ONE* 6:e22594.

983 Raymann K, Moran NA. 2018. The role of the gut microbiome in health and disease of adult honey bee
984 workers. *Curr. Opin. Insect Sci.* 26:97–104.

985 Robinson L, Liaw J, Omole Z, Corcionivoschi N, Hachani A, Gundogdu O. 2022. In silico investigation
986 of the genus *Campylobacter* type VI secretion system reveals genetic diversity in organization
987 and putative effectors. *Microb. Genomics* 8:000898.

988 Robinson LA, Collins ACZ, Murphy RA, Davies JC, Allsopp LP. 2023. Diversity and prevalence of type
989 VI secretion system effectors in clinical *Pseudomonas aeruginosa* isolates. *Front. Microbiol.*
990 13.

991 Rocha ST, Shah DD, Shrivastava A. 2024. Ecological, beneficial, and pathogenic functions of the type
992 9 secretion system. *Microb. Biotechnol.* 17:e14516.

993 Russell AB, Wexler AG, Harding BN, Whitney JC, Bohn AJ, Goo YA, Tran BQ, Barry NA, Zheng H,
994 Peterson SB, et al. 2014. A type VI secretion-related pathway in *Bacteroidetes* mediates
995 interbacterial antagonism. *Cell Host Microbe* 16:227–236.

996 Santichaivekin S, Yang Q, Liu J, Mawhorter R, Jiang J, Wesley T, Wu Y-C, Libeskind-Hadas R. 2021.
997 eMPRess: a systematic cophylogeny reconciliation tool. *Bioinformatics* 37:2481–2482.

998 Sassone-Corsi M, Nuccio S-P, Liu H, Hernandez D, Vu CT, Takahashi AA, Edwards RA, Raffatellu M.
999 2016. Microcins mediate competition among *Enterobacteriaceae* in the inflamed gut. *Nature*
1000 540:280–283.

1001 Schepers MJ, Yelland JN, Moran NA, Taylor DW. 2021. Isolation of the *Buchnera aphidicola* flagellum
1002 basal body complexes from the *Buchnera* membrane. *PLOS ONE* 16:e0245710.

1003 Schmid MC, Schulein R, Dehio M, Denecker G, Carena I, Dehio C. 2004. The VirB type IV secretion
1004 system of *Bartonella henselae* mediates invasion, proinflammatory activation and antiapoptotic
1005 protection of endothelial cells. *Mol. Microbiol.* 52:81–92.

1006 Schmidt K, Santos-Matos G, Leopold-Messer S, El Chazli Y, Emery O, Steiner T, Piel J, Engel P. 2023.
1007 Integration host factor regulates colonization factors in the bee gut symbiont Frischella
1008 perrara.Xavier KB, Garrett WS, Xavier KB, Leulier F, Florez Platino L, editors. *eLife*
1009 12:e76182.

1010 Schröder G, Schuelein R, Quebatte M, Dehio C. 2011. Conjugative DNA transfer into human cells by
1011 the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. *Proc.*
1012 *Natl. Acad. Sci.* 108:14643–14648.

1013 Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069.

1014 Serapio-Palacios A, Woodward SE, Vogt SL, Deng W, Creus-Cuadros A, Huus KE, Cirstea M, Gerrie
1015 M, Barcik W, Yu H, et al. 2022. Type VI secretion systems of pathogenic and commensal
1016 bacteria mediate niche occupancy in the gut. *Cell Rep.* [Internet] 39. Available from:
1017 https://www.cell.com/cell-reports/abstract/S2211-1247(22)00492-2

1018 Serra DO, Conover MS, Arnal L, Sloan GP, Rodriguez ME, Yantorno OM, Deora R. 2011. FHA-
1019 mediated cell-substrate and cell-cell adhesions are critical for *Bordetella pertussis* biofilm
1020 formation on abiotic surfaces and in the mouse nose and the trachea. *PLOS ONE* 6:e28811.

1021 Siamer S, Dehio C. 2015. New insights into the role of *Bartonella* effector proteins in pathogenesis.
1022 *Curr. Opin. Microbiol.* 23:80–85.

1023 Smith TJ, Font ME, Kelly CM, Sondermann H, O'Toole GA. 2018. An N-terminal retention module
1024 anchors the giant adhesin LapA of *Pseudomonas fluorescens* at the cell surface: a novel
1025 subfamily of type I secretion systems. *J. Bacteriol.* 200:e00734-17.

1026 Speare L, Cecere AG, Guckes KR, Smith S, Wollenberg MS, Mandel MJ, Miyashiro T, Septer AN. 2018.
1027 Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host.
1028 *Proc. Natl. Acad. Sci.* 115:E8528–E8537.

1029 Spitz O, Erenburg IN, Beer T, Kanonenberg K, Holland IB, Schmitt L. 2019. Type I secretion systems—
1030 one mechanism for all? *Microbiol. Spectr.* 7:10.1128/microbiolspec.psib-0003–2018.

1031 Steele MI, Kwong WK, Whiteley M, Moran NA. 2017. Diversification of Type VI Secretion System
1032 Toxins Reveals Ancient Antagonism among Bee Gut Microbes. *mBio* 8:10.1128/mbio.01630-
1033 17.

1034 Steele MI, Moran NA. 2021. Evolution of Interbacterial Antagonism in Bee Gut Microbiota Reflects
1035 Host and Symbiont Diversification. *mSystems* 6:10.1128/msystems.00063-21.

1036 Steele MI, Motta EVS, Gattu T, Martinez D, Moran NA. 2021. The Gut Microbiota Protects Bees from
1037 Invasion by a Bacterial Pathogen. *Microbiol. Spectr.* 9:e00394-21.

1038 Suria AM, Smith S, Speare L, Chen Y, Chien I, Clark EG, Krueger M, Warwick AM, Wilkins H, Septer
1039 AN. 2022. Prevalence and diversity of type VI secretion systems in a model beneficial
1040 symbiosis. *Front. Microbiol.* 13:988044.

1041 Suzuki A, Hisamoto S, Sakamoto Y. 2025. Dynamics of the hindgut microbiota of the Japanese honey
1042 bees (*Apis cerana japonica*) throughout the overwintering period. *PeerJ* 13:e20050.

1043 Thomas S, Holland IB, Schmitt L. 2014. The type 1 secretion pathway — the hemolysin system and
1044 beyond. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1843:1629–1641.

1045 Trunk T, Khalil HS, Leo JC. 2018. Bacterial autoaggregation. *AIMS Microbiol.* 4:140–164.

1046 Ugarte-Ruiz M, Stabler RA, Domínguez L, Porrero MC, Wren BW, Dorrell N, Gundogdu O. 2015.
1047 Prevalence of type VI secretion system in spanish *Campylobacter jejuni* isolates. *Zoonoses*
1048 *Public Health* 62:497–500.

1049 Unni R, Pintor KL, Diepold A, Unterweger D. 2022. Presence and absence of type VI secretion systems
1050 in bacteria. *Microbiology* 168:001151.

1051 Vasimuddin Md, Misra S, Li H, Aluru S. 2019. Efficient architecture-aware acceleration of BWA-MEM
1052 for multicore systems. In: 2019 IEEE International Parallel and Distributed Processing
1053 Symposium (IPDPS). p. 314–324. Available from:
1054 https://ieeexplore.ieee.org/document/8820962

1055 Veith PD, Glew MD, Gorasia DG, Reynolds EC. 2017. Type IX secretion: the generation of bacterial
1056 cell surface coatings involved in virulence, gliding motility and the degradation of complex
1057 biopolymers. *Mol. Microbiol.* 106:35–53.

1058 Waksman G. 2025. Molecular basis of conjugation-mediated DNA transfer by gram-negative bacteria.
1059 *Curr. Opin. Struct. Biol.* 90:102978.

1060 Wallner A, Moulin L, Busset N, Rimbault I, Béna G. 2021. Genetic diversity of type 3 secretion system
1061 in *Burkholderia* s.l. and links with plant host adaptation. *Front. Microbiol.* 12:761215.

1062 Wang X, Sun B, Xu M, Qiu S, Xu D, Ran T, He J, Wang W. 2018. Crystal structure of the periplasmic
1063 domain of TssL, a key membrane component of type VI secretion system. *Int. J. Biol. Macromol.*
1064 120:1474–1479.

1065 Wangthaisong P, Piromyou P, Songwattana P, Wongdee J, Teamtaisong K, Tittabutr P, Boonkerd N,
1066 Teaumroong N. 2023. The type IV secretion system (T4SS) mediates symbiosis between
1067 *Bradyrhizobium sp.* SUTN9-2 and legumes. *Appl. Environ. Microbiol.* 89:e00040-23.

1068 Weber BS, Ly PM, Irwin JN, Pukatzki S, Feldman MF. 2015. A multidrug resistance plasmid contains
1069 the molecular switch for type VI secretion in Acinetobacter baumannii. *Proc. Natl. Acad. Sci.*
1070 112:9442–9447.

1071 Wei Z, Zhang S, Wang X, Bai J, Wang H, Yang Y, Zhai J. 2025. Beyond survival to domination:
1072 *Brucella*'s multilayered strategies for evading host immune responses. *Front. Microbiol.*
1073 16:1608617.

1074 Whelan FJ, Rusilowicz M, McInerney JO. 2020. Coinfinder: detecting significant associations and
1075 dissociations in pangenomes. *Microb. Genomics* 6:e000338.

1076 Whitfield GB, Brun YV. 2024. The type IVc pilus: just a Tad different. *Curr. Opin. Microbiol.*
1077 79:102468.

1078 Wong MJQ, Liang X, Smart M, Tang L, Moore R, Ingalls B, Dong TG. 2016. Microbial herd protection
1079 mediated by antagonistic interaction in polymicrobial communities. *Appl. Environ. Microbiol.*
1080 82:6881–6888.

1081 Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol.*
1082 20:257.

1083 Wu J, Lang H, Mu X, Zhang Z, Su Q, Hu X, Zheng H. 2021. Honey bee genetics shape the strain-level
1084 structure of gut microbiota in social transmission. *Microbiome* 9:225.

1085 Xiong X, Wan W, Ding B, Cai M, Lu M, Liu W. 2024. Type VI secretion system drives bacterial diversity
1086 and functions in multispecies biofilms. *Microbiol. Res.* 279:127570.

1087 Zboralski A, Biessy A, Filion M. 2022. Bridging the gap: type III secretion systems in plant-beneficial
1088 bacteria. *Microorganisms* 10:187.

1089 Zhang J, Guan J, Wang M, Li G, Djordjevic M, Tai C, Wang H, Deng Z, Chen Z, Ou H-Y. 2022. SecReT6
1090 update: a comprehensive resource of bacterial type VI secretion systems. *Sci. CHINA Life Sci.*
1091 66:626–634.

1092 Zhang L, Hinz AJ, Nadeau J-P, Mah T-F. 2011. *Pseudomonas aeruginosa* tssC1 links type VI secretion
1093 and biofilm-specific antibiotic resistance. *J. Bacteriol.* 193:5510–5513.

1094 Zhang S, Wu F, Zhao H, Zhao L, Li D, Yang F, Liu L. 2025. Type IV secretion systems and conjugation
1095 in gram-negative pathogens. *FASEB J.* 39:e71116.

1096 Zhang W, Zhang X, Su Q, Tang M, Zheng H, Zhou X. 2022. Genomic features underlying the
1097 evolutionary transitions of *Apibacter* to honey bee gut symbionts. *Insect Sci.* 29:259–275.

1098 Zhang Z, Mu X, Cao Q, Shi Y, Hu X, Zheng H. 2022. Honeybee gut *Lactobacillus* modulates host
1099 learning and memory behaviors via regulating tryptophan metabolism. *Nat. Commun.* 13:2037.

1100 Zheng H, Nishida A, Kwong WK, Koch H, Engel P, Steele MI, Moran NA. 2016. Metabolism of toxic
1101 sugars by strains of the bee gut symbiont *Gilliamella apicola*. *mBio* 7:10.1128/mbio.01326-16.

1102 Zheng H, Perreau J, Powell JE, Han B, Zhang Z, Kwong WK, Tringe SG, Moran NA. 2019. Division
1103 of labor in honey bee gut microbiota for plant polysaccharide digestion. *Proc. Natl. Acad. Sci.*
1104 116:25909–25916.

Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc. Natl. Acad. Sci.* 114:4775–4780.

Zhou N, Zheng Q, Liu Y, Huang Z, Feng Y, Chen Y, Hu F, Zheng H. 2025. Strain diversity and host specificity of the gut symbiont *Gilliamella* in *Apis mellifera, Apis cerana* and *Bombus terrestris*. *Microbiol. Res.* 293:128048.

Zoued A, Brunet YR, Durand E, Aschtgen M-S, Logger L, Douzi B, Journet L, Cambillau C, Cascales E. 2014. Architecture and assembly of the Type VI secretion system. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1843:1664–1673.

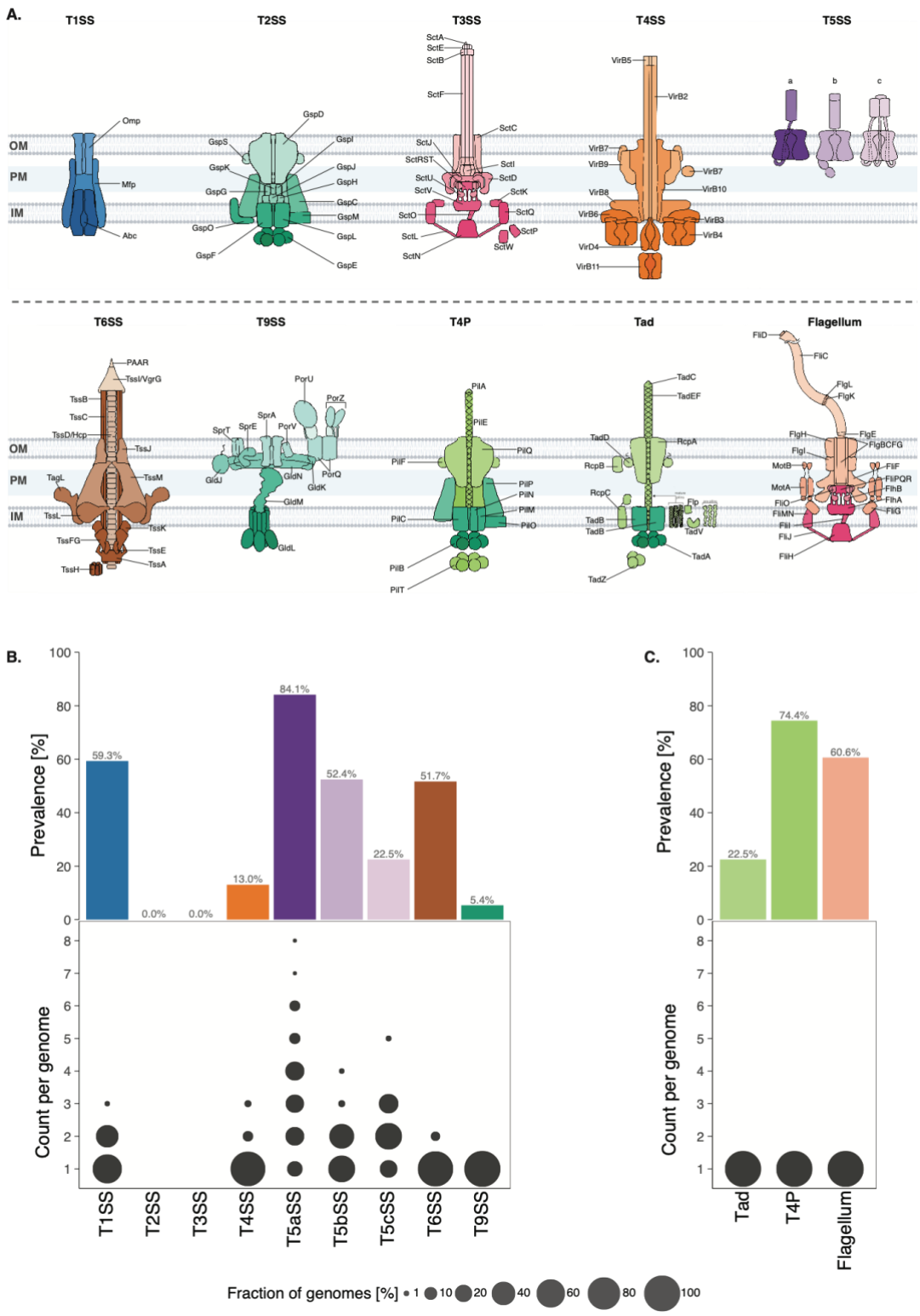


**Figure 1 Prevalence of secretion systems and appendages in the bee bacterial community**. **(A)** Illustration of secretion systems Type I (T1SS) to Type 9 (T9SS) and their evolutionary-related appendages Type IV pili (T4P), Tight-adherence pili (Tad) and Flagella, showing putative assembly of individual structural components across the inner membrane (IM), periplasm (PM) and the outer membrane (OM), of Gram-negative bacteria. Occurrence and the fraction of bacteria per count of the **(B)** secretion systems T1SS – T9SS and **(C)** their evolutionary related appendages among the Gram-negative bee bacteria.

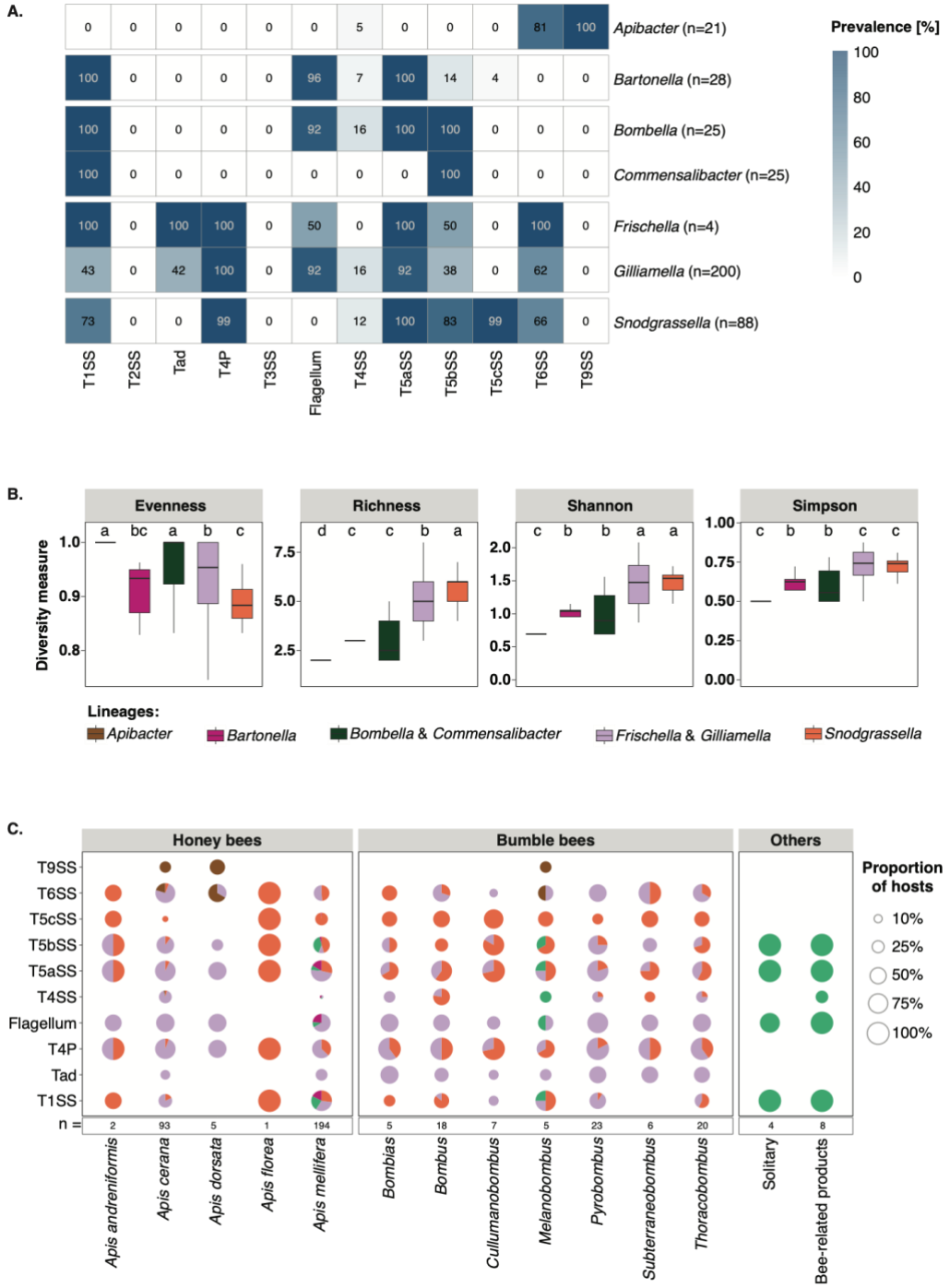


**Figure 2 Secretions systems are maintained for the bacterial community than for host adaptation.** **(A)** Prevalence of each secretion system and related appendages among each of the Gram-negative bee bacterial genera. Genus names on the y-axis with the number of strains analysed in parenthesis. Values indicate the rounded percentage. **(B)** Alpha diversity of the distribution of all secretion systems within each genus. Pairwise significance was assessed with Wilcoxon rank-sum tests (BH corrected), and letters indicate groups not significantly different (Tukey HSD). **(C)** Prevalence of the secretion systems and related appendages across different bee hosts. On the x-axis, the bee species (Honey bees/*Apis*) and subgenera (Bumble bees/*Bombus*). The n is the total number of bacteria isolated from each host used in the analysis. The pie size indicates the fraction of bacteria (i.e. n) per host and the slices show the bacterial lineage contribution. Colours of the slices correspond to the legend in (B).

**Table 1 Co-phylogeny analysis of bacterial lineages and their secretion systems.** *P*-values reported from three separate distance-based cophylogenetic methods (PACo, ParaFit, Hommola) and one one event-based method (eMPRess) between the lineage phylogeny and their respective system phylogeny. Systems phylogeny were constructed with a concatenation of single copy components: 11 (both T6SS and T9SS) proteins for *Apibacter*; 3 (T1SS) and 9 (Flagellum) proteins for *Bartonella*; 3 (T1SS) and 11 (Flagellum) proteins for *Bombella* & *Commensalibacter*; 2 (T1SS), 10 (T4P), 7 (Tad), and 12 (T6SS) proteins for *Frischella* & *Gilliamella*; and 2 (T1SS), 14 (T4P), and 12 (T6SS) proteins for *Snodgrassella*.

| | | ***p-value*** | | | |
|---|---|---|---|---|---|
| **Lineage** | **System** | **PACo** | **Parafit** | **Hommola** | **eMPRess** |
| ***Apibacter*** | T6SS | 0 | 0.001 | 0 | 0.020 |
| | T9SS | 0 | 0.002 | 0 | 0.010 |
| ***Bartonella*** | **T1SS (all clades)** | **0.999** | **1.000** | **0.870** | 0.010 |
| | T1SS clade 1 | 0 | 0.001 | 0 | 0.010 |
| | T1SS clade 2 | 0 | 0.001 | 0 | 0.010 |
| | Flagellum | 0 | 0.001 | 0 | 0.010 |
| | T5aSS | 0.004 | 0.015 | 0.010 | 0.010 |
| ***Bombella & Commensalibacter*** | T1SS | 0.001 | 0.001 | 0 | 0.010 |
| | Flagellum | 0 | 0.411 | 0 | 0.010 |
| | T5aSS | 0 | 0.915 | 0 | 0.010 |
| | T5bSS | 0 | 0.001 | 0 | 0.010 |
| ***Gilliamella & Frischella*** | T1SS | 0 | 0.001 | 0 | 0.010 |
| | T4P | 0 | 0.002 | 0 | 0.010 |
| | Tad | 0 | 0.001 | 0 | 0.010 |
| | Flagellum | 0 | 0.005 | 0 | 0.010 |
| | T4SS | 0 | 0.022 | 0 | 0.011 |
| | T5aSS | 0 | 0.001 | 0 | 0.010 |
| | T5bSS | 0 | 0.001 | 0 | 0.010 |
| | T6SS (all subtypes) | 0 | 0.001 | 0 | 0.010 |
| | **T6SS subtype i1** | 0.045 | 0.019 | **0.055** | **0.257** |
| | T6SS subtype i2 | 0 | 0.001 | 0 | 0.010 |
| ***Snodgrassella*** | T1SS | 0 | 0.001 | 0 | 0.010 |
| | T4P | 0 | 0.001 | 0 | 0.010 |
| | **T4SS** | **0.459** | **0.669** | **0.877** | **0.327** |
| | T5aSS | 0 | 0.001 | 0 | 0.010 |
| | T5bSS | 0 | 0.001 | 0 | 0.010 |
| | T6SS (all subtypes) | 0.007 | 0.068 | 0 | 0.010 |
| | T6SS subtype i2 | 0 | 0.001 | 0 | 0.010 |
| | T6SS subtype i3 | 0 | 0.001 | 0 | 0.010 |

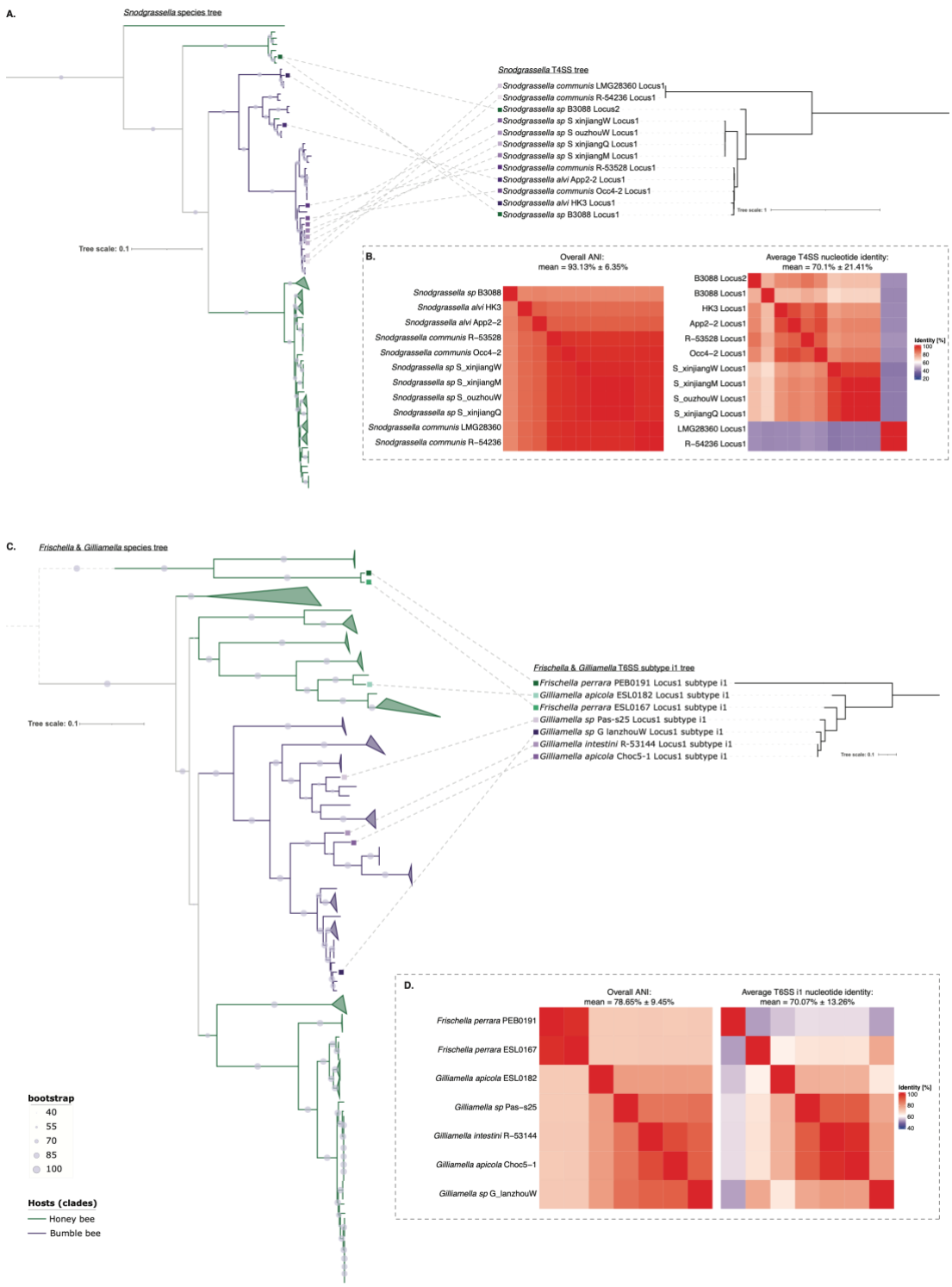


**Figure 3 Phylogenetic relationships and sequence similarity of secretion system loci across bee gut symbionts.** (A) Tanglegram between maximum-likelihood species phylogeny and the Type IV secretion system phylogeny of *Snodgrassella* shows incongruence. (B) Pairwise average nucleotide identity (ANI) among *Snodgrassella* genomes and average nucleotide identity of the T4SS loci. (C) Tanglegram between maximum-likelihood species phylogeny and the Type VI secretion system subtype i1 phylogeny of *Frischella* & *Gilliamella.* (D) Corresponding ANI and T6SS average nucleotide identity comparisons. Tree node support values are shown as bootstrap percentages. Host origin (honey bee vs. bumble bee) indicated by colour of species tree branches.

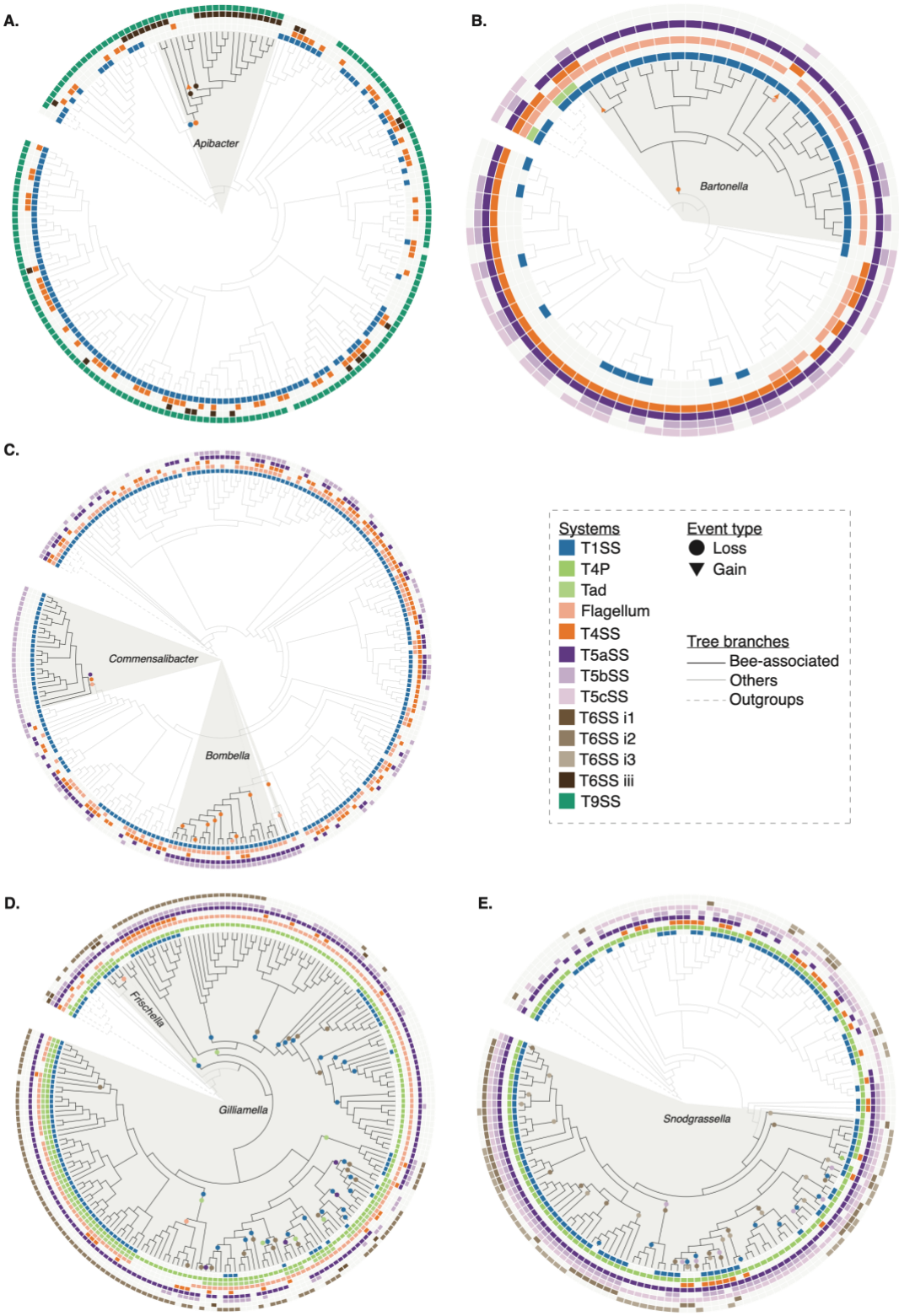


**Figure 4 Loss of secretion systems are the predominant evolutionary mechanism in bee bacteria.** Phylogenetic tree of the bee-related and both closely- and distant-related bacteria (see Supplementary Fig. S1 -5 for phylogeny) of **(A)** *Apibacter*, **(B)** *Bartonella*, **(C)** *Bombella* and *Commensalibacter*, **(D)** *Frischella* and *Gilliamella*, and **(E)** *Snodgrassella*. The bee bacteria clades are shaded in grey. The presence and absence of each secretion system is shown as the heatmap around the phylogenetic tree. The loss and gain events shown as filled shapes on the possible clade/branches of the event.

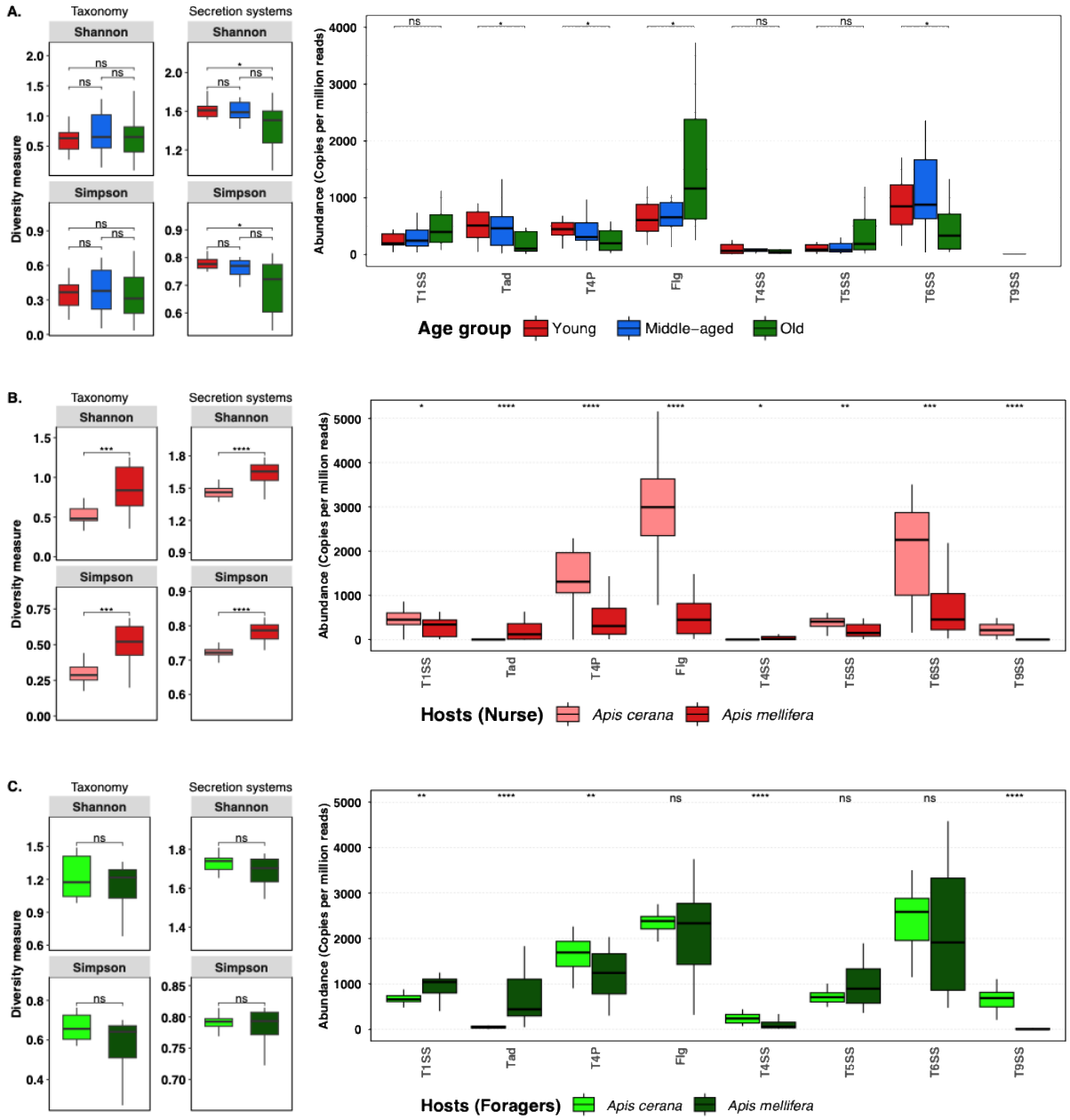


**Figure 5 Diversity of secretion systems is influenced by the age of hosts**. Shannon and Simpson diversity of only the taxa that harbours secretion systems, the diversity of the total abundance of secretion systems and the abundances of reads corresponding to each secretion system within metagenomic dataset obtained from **(A)** young (10 days old), middle-aged (22-24 days) and old (≥48 days) *Apis mellifera*; **(B)** from nurses (young) *Apis cerana* and *Apis mellifera*; and **(C)** from foraging (old) *Apis cerana* and *Apis mellifera*. Statistical significance was assessed using Wilcoxon rank-sum tests (two groups) or Kruskal–Wallis tests (≥3 groups), with Benjamini–Hochberg FDR correction where applicable. Significance levels: P < 0.05 (*); < 0.01 (**); < 0.001 (***); < 0.0001 (****); ns, not significant.