bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Foraging on host synthesized metabolites enables the bacterial symbiont *Snodgrassella alvi* to colonize the honey bee gut

Andrew Quinn[1,A], Yassine El Chazli[1,A], Stéphane Escrig[2], Jean Daraspe[3], Nicolas Neuschwander[1], Aoife McNally[1], Christel Genoud[3], Anders Meibom[2,4], Philipp Engel[1,*]

[1] Department of Fundamental Microbiology, University of Lausanne, Lausanne, Switzerland

[2]Laboratory of Biological Geochemistry, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

[3] Electron Microscopy Facility, University of Lausanne, Lausanne, Switzerland

[4] Center for Advanced Surface Analysis, Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland

[A] Equally contributing first authors

* Corresponding author: phillipp.engel@unil.ch

Keywords

Microbiome, microbiota, insect, symbiosis, metabolomics, NanoSIMS isotope-tracing, gnotobiotic, mass spectrometry, proteobacteria.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Summary

Dietary nutrients and microbial cross-feeding allow diverse bacteria to colonize the animal gut. Less is known about the role of host-derived nutrients in enabling gut bacterial colonization. We examined metabolic interactions within the evolutionary ancient symbiosis between the honey bee (*Apis mellifera*) and the core gut microbiota member *Snodgrassella alvi*. This Betaproteobacteria is incapable of metabolizing saccharides, yet colonizes the honey bee gut in the presence of only a sugar diet. Using comparative metabolomics, ^{13}C tracers, and Nanoscale secondary ion mass spectrometry (NanoSIMS), we show *in vivo* that *S. alvi* grows on host-derived organic acids, including citrate, glycerate and 3-hydroxy-3-methylglutarate which are actively secreted by the host into the gut lumen. *S. alvi* additionally modulates tryptophan metabolism in the gut by converting kynurenine to anthranilate. These results suggest that *S. alvi* is adapted to a specific metabolic niche in the gut that depends on host-derived nutritional resources.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

Gut bacteria and their hosts typically engage in mutualistic interactions. Metabolic exchange from the gut microbiota to the host is vital for uptake of essential nutrients, gut health and immune system function [1]. In turn, bacteria profit from a stable niche environment and frequent supply of exogenous food. The role of host secreted metabolites that benefit bacteria in the gut is less well understood. Such metabolic exchange is difficult to identify due to the overwhelming contributions of the diet and microbial products towards gut metabolites.

Deciphering the extent to which host metabolite secretion drives microbial colonization of native symbionts is aided by a simple, tractable system in which diet and microbiota-derived metabolites can be tightly controlled. The Western honey bee (*Apis mellifera*) provides such a system: Its gut microbiota is broadly stable and composed of only 8-10 genera [2,3] many of which have longstanding evolutionary associations with their host that date back to the emergence of social bees >80 mya [4-6]. Bacteria of these genera are culturable and can be inoculated individually or as defined communities into gnotobiotic bees [7]. Furthermore, while the typical honey bee diet features a rich mixture of compounds found in nectar and pollen, bees can survive for extended periods on pure sugar water diets [8].

Most members of the bee gut microbiome are primary fermenters, possessing a broad range of carbohydrate degradation enzymes that enable them to utilize hemicellulose, pectin, starch, or glycosides found in pollen [3,9]. A notable exception is the Betaproteobacteria *Snodgrassella alvi*. It colonizes the cuticular surfaces of the ileum in the hindgut and displays a markedly different metabolism [10,11]. Lacking a functional glycolysis pathway, *S. alvi* profits from acids in the gut, generating energy from an aerobic TCA cycle, and biomass through gluconeogenesis. Fermentation by other microbial members has been proposed as the primary source of short chain fatty acids (SCFAs) consumed by *S. alvi*. In particular, bacteria of the Gammaproteobacterial genus *Gilliamella* are likely mutualistic partners for metabolic crosstalk, as they co-localize with *S. alvi* within biofilms attached to the cuticular surface in the ileum and share complementary metabolic capabilities [10]. Experimental evidence from *in vitro* growth experiments bolstered this hypothesis, showing that *S. alvi* grows on the spent media of

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Gilliamella, while consuming numerous products of *Gilliamella's* metabolism, such as succinate and pyruvate [12].

Although a strong case can be made for niche exploitation through bacterial cross-feeding, this hypothesis does not explain previous results in which *S. alvi* was able to mono-colonize bees fed diets of sugar water and pollen [12]. To better understand the nutrient sources that *S. alvi* exploits, we provided bees a simple (sugar water) or complex (sugar water + pollen) diet and colonized them with *S. alvi* alone, or together with divergent strains of the genus *Gilliamella*. Surprisingly, we found that a simple sugar water diet was sufficient for *S. alvi* to colonize the honey bee gut. Subsequent metabolomics analysis indicated that host-derived carboxylic acids enable *S. alvi* colonization. We validated this hypothesis with a series of experiments to show that (i) these carboxylic acids are synthesized by the host, (ii) *S. alvi* utilizes them for growth, and (iii) the findings hold across a range of divergent *Snodgrassella* strains and species.

Results

S. alvi colonizes the honey bee gut in the presence of only sugar in the diet

We colonized microbiota-free (MF) honey bees with *S. alvi* (strain wkB2) or a mixture of *Gilliamella* strains (strains wkB1, ESL0169, ESL0182, ESL0297) or with both phylotypes together, and provided a diet composed of only sterile-filtered sugar water (sucrose) or sugar water and electron beam-sterilized polyfloral pollen for five days (**Fig. 1A**). Colonization levels of both phylotypes were assessed by qPCR and CFU plating five days post inoculation (**Fig. 1B**, **Fig. S1**). Surprisingly, *S. alvi* colonized at equivalent levels in the guts across all treatment groups, *i.e.*, sugar water in the diet was sufficient to enable *S. alvi* colonization and neither the addition of pollen to the diet nor co-colonization with *Gilliamella* significantly increased *S. alvi* loads (Wilcoxon rank sum test with *S. alvi* (sugar water) as reference group).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Importantly, qPCR analysis with universal bacterial and fungal primers showed no or very low levels of amplification in most samples, with the exception of gut samples containing pollen which resulted in relatively high background amplification with the universal bacterial primers as previously reported (**Fig. S1C, D**). Culturing of gut homogenates of MF bees on different media resulted in no microbial growth in any of the tested conditions (see methods). Hence, we can rule out that systematic contaminations with high levels of other microbes facilitated *S. alvi* colonization through cross-feeding in the mono-colonization treatment.

S. alvi depletes organic acids in the honey bee gut

To search for putative *S. alvi* growth substrates originating from the host, the pollen diet, or *Gilliamella*, we next extracted metabolites from the mid- and hindgut for a subset of the colonized bees and analyzed them via GC-MS. The presence of pollen in the gut significantly increased the abundance of nearly half of the annotated metabolites (125/233) in MF bees (**Fig. S2**). Thus, we compared results between colonized and MF bees only within each dietary treatment. When we examined mono-colonization of *S. alvi* in bees fed with sugar water, we identified multiple carboxylic acids, including citrate, 3-hydroxy-3-methylglutarate (3Hmg), and glycerate that were significantly less abundant than in the MF controls **(Fig. 1C)**. These carboxylic acids, along with others (*e.g.*, 2-ketoisocaproate, alpha-ketoglutarate and malate) were also depleted in *S. alvi* mono-colonized bees that were fed pollen **(Fig. 1C)**, despite the substantial differences in the overall metabolite profiles between the two dietary conditions. In contrast, few metabolites were significantly more abundant in *S. alvi* colonized bees relative to MF bees, with only anthranilate, a product of tryptophan metabolism, accumulating in colonized bees of both dietary treatments (**Fig. 1C**).

We then fit each metabolite with a mixed linear model to quantify how strongly metabolite changes were influenced (*i.e.*, the fixed effect) by each of the three independent experimental variables (*i.e.*, the presence of *Gilliamella, S. alvi,* or pollen) across the eight experimental conditions, This cross-conditional analysis confirmed that colonization with *S. alvi* resulted in significantly (p-value < 0.05,

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

see Supporting Data, "output.csv") decreased abundances of carboxylic acids, particularly 3Hmg, citrate, malate, fumarate, and glycerate (**Fig. 1D**). We could also infer which metabolites were co-produced, co-consumed, or cross-fed (and in which direction) between the two bacteria (**Fig. 1D**). For example, we found evidence supporting our initial hypothesis of metabolic cross-feeding from *Gilliamella* to *S. alvi* in the form of lactate, pyruvate and other unknown compounds that were more abundant with *Gilliamella* and less abundant with *S. alvi* (**Fig. 1D**, **Fig. S3**). However, we found no evidence of reverse cross-feeding from *S. alvi* to *Gilliamella*. Instead, we found that both species compete for certain metabolites and act synergistically to synthesize others. Competition, or co-consumption, centered on the carboxylic acids, as colonization with *Gilliamella* also led to depletion of citrate and glycerate, and to a lesser extent, malate, and fumarate (**Fig. 1D**, **Fig. S3**). Cooperative synthesis revolved around four metabolites, acetate, succinate, benzoate, and unknown_12.74. Of these, only acetate was also more abundant in mono-colonization (*Gilliamella*) than in MF controls (**Fig. 1D**, **Fig. S3**). Finally, nearly all the metabolites depleted with *S. alvi* colonization were positively affected by pollen, except for citrate, glycerate, acetate, and two unknown compounds, which were unaffected, and urea, whose abundance was negatively correlated with pollen. Taken together, these results show that each of the tested variables impacts nutrient availability and metabolism of *S. alvi*, although metabolic cross-feeding and a nutrient rich diet do not lead to increased colonization levels in the gut.

Host sugar catabolism provides *S. alvi* substrates *in vivo*

We next sought to measure host production of carboxylic acids in the guts of MF bees over the first six days post emergence and compare the abundances to what the bee could have extracted from the average amount of pollen consumed (24.3 ± 7.2mg/bee) as estimated based on data collected from the previous experiment (see methods, **Fig. S4** and Table S2). Moreover, in order to rule-out microbial or fungal contamination as a source of these compounds, we rigorously checked once more for live contamination by plating the guts of newly emerged (day zero) and six-day old MF bees on eight different media in three different growth environments (see methods), finding no evidence of live bacterial or fungal contamination (n=8). Even though the bees only consumed sugar water, we found that many gut

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

metabolites, including glycosylamines, sugar alcohols and carboxylic acids, were significantly more abundant six days after emergence (**Fig. S5**). Focusing specifically on the compounds depleted in the presence of *S. alvi*, we found that citrate was most abundant in the gut, with its concentration increasing from 29 ± 14 µmol/mg at Day 0 to 73 ± 34 µmol/mg at Day 6 post emergence. Substantially less citrate, 3 ± 2 µmol/mg, was extracted from pollen (**Fig. 2A**). A similar trend was also found for 3Hmg, alpha-ketoglutarate, isocitrate and kynurenine. Of *S. alvi's* putative substrates, only caffeate and lactate were more abundant in pollen than in the bee gut. The increasing abundance of many compounds over six days, as well as their generally low abundance in pollen suggests that active metabolism by the host substantially impacts the gut metabolome, i.e. easily digestible nutrients in the diet are absorbed and metabolized by the bee upstream of the hindgut, while downstream metabolic products are excreted into the hindgut and serve as substrates for *S. alvi*.

To corroborate this hypothesis, we fed MF bees for six days with 45% of the dietary sucrose replaced by 100% U-$^{13}C_6$ Glucose. Half of the bees were also provided standard pollen to assess its contribution relative to simple sugars in the synthesis of gut metabolites. We then analyzed the resulting ^{13}C isotopic enrichments in gut metabolites with a focus on those indicated as substrates for *S. alvi*. The carboxylic acids and non-essential amino acids were significantly ^{13}C enriched in the bee gut. In contrast, other compounds, such as kynurenine and essential amino acids did not show any enrichment **(Fig. 2B)**. The lack of isotope labelling indicates that these compounds were not actively synthesized from glucose by adult bees during the first six days post-emergence. Instead, they were either leftover in the gut from the larval development stage or were acquired from the larval diet, such as in the case of essential amino acids and their catabolic products. As expected, the average ^{13}C enrichment levels of most metabolites dropped in bees fed with pollen compared to those fed with sugar water only. The ^{13}C dilution can occur directly from non-labelled metabolites in pollen, as well as indirectly from host metabolism of pollen substrates. However, the average ^{13}C enrichment of carboxylic acids only dropped from 21 ± 6% to 17 ± 7%, which serves as further evidence that host metabolism of simple sugars rather than dietary consumption is the predominant source of carboxylic acids in the gut (**Fig. 2B).** We thus conclude from

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

this analysis that the carboxylic acids utilized by *S. alvi* are mostly *de novo* synthesized from sugar metabolism of the host.

NanoSIMS reveals transfer of host compounds to *S. alvi*

While our previous experiments suggested host nutrient foraging by *S. alvi*, the results did not provide direct evidence that *S. alvi* assimilates biomass from these compounds. Therefore, we probed the flow of metabolites from the bee to *S. alvi* using a "pulse-chase" ^{13}C isotope labelling experiment (ILE), and then measured the enrichment within *S. alvi* cells and surrounding host tissue using nanoscale secondary ion mass spectrometry (NanoSIMS) complemented with measurements of metabolite enrichments using GC-MS (**Fig. 3, Dataset S1**). To do so, we enriched newly emerged MF bees in ^{13}C by feeding them 100% U-$^{13}C_6$ glucose for four days after emergence. We then inoculated them with *S. alvi* and waited one day before switching their diet to naturally abundant, 98.9% ^{12}C-Glucose (**Fig. 3A**). This ensured that the largest *S. alvi* population increase, between 24- and 48-hours post-colonization, occurred without a dietary ^{13}C source, but still in a highly ^{13}C labelled environment **(Fig. 3B)**.

Accordingly, we measured rapid ^{13}C labelling turnover in *S. alvi* in 18 images coming from two bees (sampled at 48- and 72-hours post-inoculation) in which we could detect bacterial cells **(Fig. 3C, E, Dataset S2)**. The ^{13}C enrichment was substantially higher in *S. alvi* cells than in the adjacent host ileum tissue 48 hours after inoculation, but then dropped below the levels of host tissue 72 hours after inoculation (At% = start: 5.56, end: 2.25) **(Fig. 3C)**. In contrast, the ileum tissue displayed a constant ^{13}C enrichment over the course of the experiment (At% = start: 3.31, end: 3.16), whereas the cuticle lining of the gut epithelium was comparatively less enriched initially and continued dropping to almost natural ^{13}C enrichment levels after 72 hours (At% = start: 2.58, end: 1.39) **(Fig. 3C)**.

The ^{13}C enrichment of carboxylic acids and host synthesized (non-essential) amino acids in the gut dynamically shifted in response to the $^{13}C/^{12}C$ switch (**Fig. 3D, Dataset S1)**. Initially, they were highly ^{13}C enriched (38 ± 14.7%), but the labelling fell steadily to approximately half the original level (14 ± 5.8%) within 48 hours (**Fig. 3D)**. As a control, we found that essential amino acids (non-host

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

189 synthesized) were not enriched above the level of natural ^{13}C abundance (1.11%) throughout the
190 experiment. Thus, these results show that products of host glucose metabolism are actively used by *S.*
191 *alvi* to build its biomass during early colonization.

192

193 ## *In vitro* growth assays confirm active metabolism by *S. alvi*

194 We next tested whether the putative host synthesized substrates were sufficient for growth of *S. alvi* as
195 sole carbon sources in a chemically defined liquid medium "Bee9", which we derived from standard
196 M9 media (**Fig. 4A**, **Dataset S3**). We utilized Bee9, after finding that *S. alvi* consumes amino acids
197 present in standard M9 previously used to grow this bacterium [12]. Only three metabolites (citrate,
198 isocitrate, and malate) supported growth as sole carbon sources in this condition. However, the addition
199 of 3Hmg, fumarate, succinate, gamma-aminobutyrate (Gaba), kynurenine, or urea to Bee9 + citrate
200 significantly improved the maximal growth rate of *S. alvi* relative to Bee9 + citrate alone (**Fig. 4B**). In
201 particular, 3Hmg and Gaba had dramatic effects on growth rates, increasing them 3.52 ± 0.10 and 2.76
202 ± 0.04-fold, respectively.

203 Analysis of the spent media revealed that all tested metabolites, except for the glucose negative control,
204 were depleted by *S. alvi* (**Fig. S6**). This further confirmed that the identified host-derived organic acids
205 are growth substrates of *S. alvi.* Unlike the results from our colonization experiments, *S. alvi* produced
206 numerous carboxylic acids, which varied based on the carbon source. For example, acetate was a
207 metabolic substrate, but it was also produced in the presence of pyruvate or 3Hmg. Meanwhile, alpha-
208 ketoglutarate, 2-oxoisovalerate, malate and succinate were produced across most growth conditions
209 (**Fig. 4C**). Consistent with the *in vivo* metabolomics, we detected production of anthranilate when
210 kynurenine was added to the media. Additionally, we also measured the production of kynurenic acid
211 and tryptophan (**Fig. 4D).** *S. alvi* wkB2 encodes a putative kynureninase gene in its genome
212 (SALWKB2_0716, Genbank genome: CP007446 [10]) that is likely responsible for the conversion of
213 kynurenine into anthranilate. Interestingly, this gene seems to have been acquired by horizontal gene
214 transfer as indicated by a phylogenetic analysis **(Fig. S7)**: apart from other *Snodgrassella* strains, the

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

closest sequences of the kynureninase gene were not found in other Neisseriaceae, but in more distantly related Betaproteobacteria (genera *Pusilimonas* and *Alcaligenes*) and in Gammaproteobacteria (genera *Ignatzschineria* and *Acinetobacter*).

Colonization without an exogenous nutrient source is conserved across strains of the genus *Snodgrassella*

As a final step, we examined whether our results were unique to *S. alvi* wkB2 type strain, or were more generally applicable to the genus, by inoculating bees with five divergent strains of *Snodgrassella*, three native to honey bees and two native to bumble bees. We also lengthened the colonization experiment from five to ten days to account for a potential delay in colonization of the honey bee gut by non-native strains. All five strains successfully colonized bees fed only sugar water. However, the efficiency and extent of colonization varied between strains. Strains of *Snodgrassella* native to honey bees colonized consistently, while non-native strains sometimes failed to colonize (Aggregate Colonization Success: 100% vs. 70%; Fisher's Exact test: $p < 0.001$) (**Fig. 5A**). The gut metabolomic comparison between successfully colonized versus MF bees was qualitatively similar to our initial results. Carboxylic acids, such as 3Hmg, malate, fumarate, succinate, isocitrate, and glycerate were significantly less abundant in colonized guts vs MF controls across three or more of the *S. alvi* strains tested. The levels of purine and amino acid precursors hypoxanthine and urea were also depleted in the guts of colonized bees, while anthranilate again accumulated in all colonized bees (**Fig. 5B, Fig. S8**). These results show that the utilization of host-derived carboxylic acids is a conserved phenomenon in the genus *Snodgrassella* and facilitates gut colonization independent of bacterial cross-feeding or diet-derived nutrients.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Discussion

Host-derived metabolites are gaining appreciation for their importance in facilitating extracellular microbial colonization across widely disparate animal models. In some cases, bacteria graze on the chitinous lining of the murine gut or of light organ in squid [13-20]. In other instances, bacteria profit from small molecules secreted into the lumen. Commensal species can utilize lactate, 3-hydroxybutyrate and urea in the murine gut, while pathogenic *Salmonella* species utilize aspartate, malate, lactate or succinate to invade the gut [21-24]. Our study demonstrates that this phenomenon is conserved across widely disparate animal hosts and that these host-derived metabolites can represent the major carbon source for certain gut bacteria facilitating colonization and growth, independent of the diet or cross-feeding.

We took advantage of the bee as a model host to drastically minimize confounding factors, colonizing bees with a single bacterium, *S. alvi,* while restricting the host to a sole dietary substrate that is undigestible for this bacterium. We then utilized ^{13}C-glucose labelling to show that organic acids measured in the gut are synthesized by the host from dietary sugars which are then assimilated by *S. alvi* cells during gut colonization. The ^{13}C enrichment of *S. alvi* was initially higher than in the surrounding host tissue, but then dropped after the host diet was switched to ^{12}C-glucose. These dynamics are consistent with the hypothesis that *S. alvi* utilizes host metabolites that are derived from simple carbohydrate catabolism in the bee diet, rather than from grazing on components of the host epithelium lining in the gut (**Fig. 6**). Finally, we demonstrated that these organic acids sustain growth of *S. alvi in vitro*.

Our results provide new context to the substantial body of research on bacterial colonization of the bee gut. The consistent depletion of host-derived acids linked to the TCA cycle complements findings that these metabolites are depleted in bees colonized with individual microbes including *S. alvi*, *Gilliamella, Lactobacillus* Firm-5 and *Bartonella apis*, as well as when colonized with a full microbial community

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[12,25,26]. Thus, foraging on host-derived compounds, although not always essential, may be widespread in the native bee gut community. Yet, in the specific case of *S. alvi*, these host-derived nutrients seem to be key for colonization. This is supported by a previous study which found that all genes of the TCA cycle and multiple genes associated with organic acid and ketone body degradation, provide a strong fitness advantage to *S. alvi* to grow in the bee gut [27].

Strikingly, two major metabolic functions conserved across all five strains of *S. alvi* (3Hmg consumption and anthranilate production), were associated with unique features in the metabolic pathways of *S. alvi*, providing evidence for adaptation to a specialized host niche. *S. alvi* possesses a non-canonical TCA cycle (**Fig. 6**); lacking a glyoxylate shunt it is unable to grow on acetate alone, but the substitution of acetate:succinate CoA-transferase for the canonical succinyl-CoA synthetase means that acetate, rather than acetyl-CoA, is a key driver of the TCA cycle [28]. Interestingly, we found that 3Hmg consumption results in the production of acetate when *S. alvi* is grown on 3Hmg *in vitro*. Therefore, we postulate that host-derived 3Hmg enhances growth of *S. alvi* via the production of acetate, which in turn, increases the flux through the TCA cycle (**Fig. 6**). 3Hmg is an intermediate of the host's isoprenoid biosynthesis, leucine degradation and ketone body metabolism, but the reason why this metabolite is released into the gut remains elusive.

The second unique metabolic feature of *S. alvi*, the production of anthranilate, likely depends on the enzymatic activity of a kynureninase gene identified in the genome of *S. alvi* converting host-derived kynurenine into anthranilate (**Fig. 6**). Intriguingly, this is the only gene annotated in the tryptophan degradation pathway of *S. alvi*; it is highly conserved in the entire genus and has likely been acquired by horizontal gene transfer, suggesting that it encodes a conserved function that is specific to the symbiosis between *Snodgrassella* and its bee hosts.

The release of relatively valuable metabolites into the gut could be a way for the host to control community assembly and facilitate the colonization of particularly beneficial gut symbionts [29]. While the full gut microbiota has been shown to carry out several important functions for the host, the specific role of *S. alvi* has remained elusive. It is possible that its metabolic activities, such as the conversion of kynurenine into anthranilate, has a positive effect on the host, such as protection against pathogen

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

invasion [30,31]. Kynurenine is important during the larval stage of bee development, but it is also associated with neuronal defects (*e.g.*, hyperactivity and motor disfunction) in insects and vertebrates [32-34]. In contrast, anthranilate is an important precursor of the essential amino acid tryptophan and several neurotransmitters such as serotonin, tryptamine and various indole derivatives, all of which are highly beneficial to bees [35]. Intriguingly, indoles synthesized from tryptophan by *Lactobacillus* Firm-5 species in the gut were recently linked to enhanced memory and learning in honey bees [36]. Finally, we also note that the metabolism of urea by *S. alvi* could have a beneficial effect on the host. Urea constitutes a major waste product in mammals and insects. In turtle ants, ancient, specialized gut bacteria are able to recycle large amounts of nitrogen from urea into both essential and non-essential amino acids [37]. We found that urea is more abundant in the bee gut when exogenous nitrogen was absent from the diet, indicative of higher metabolic turnover and waste, which *S. alvi* could alleviate through urea fixation.

Our results and analysis fit within the context of several limitations. We examined a portion of the total gut metabolome that is amenable to analysis with gas chromatography. Further compounds of interest may be detected using complimentary analytical capabilities. While we have shown that host-derived compounds can sustain growth of *S. alvi*, the tested conditions did not reflect the natural state of the bee gut with a full microbial community and a metabolically rich diet that alters between bee bread, pollen, honey and flower nectar. Therefore, we cannot exclude that dietary or microbially-derived metabolites contribute to the growth of *S. alvi* in the native gut. In fact, we found *in vivo* evidence for cross-feeding of lactate and pyruvate from *Gilliamella*, as well as for pollen as an indirect source of key organic acids in the gut. However, these nutrient sources seem to play a minor role relative to the host, as the colonization levels of *S. alvi* did not differ significantly when pollen and/or *Gilliamella* were present. This is consistent with a previous study, which showed that the total abundance of *S. alvi*, in contrast to most other community members, does not change between nurse and forager bees that have different dietary preferences, nor does it change when pollen is removed from the diet of fully colonized bees under laboratory conditions [38].

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

How host synthesized organic acids reach the gut lumen remains elusive (**Fig. 6**). While leakage from host cells across the epithelial barrier is possible, transport via the malpighian tubules upstream of the ileum seems more likely, as they excrete nitrogenous waste and other metabolites into the gut while also regulating osmotic pressure in the hemolymph [39,40]. Future work may resolve this question through careful dissection and metabolomic analysis of the malpighian tubules and the ileum, possibly with the assistance of labeled compounds injected into the bee thorax.

Methods

Bacterial culturing

Table S1 lists all strains used in this study. Strains were isolated by directly plating gut homogenates on brain heart infusion agar (BHIA), or by first culturing them in Insectagro™ (Gibco) liquid media under microaerophilic conditions for 48 hours before colonies were streaked on plates. Species identity was confirmed by 16S rRNA gene sequencing. Strains were grown on liquid Insectagro™ media for 18 hours and then diluted in Phosphate buffer saline (PBS) and 25% glycerol at OD_{600} =1. Solutions were kept frozen at -80 °C until just before colonization, when they were thawed on ice, diluted 10-fold with PBS and mixed 1:1 with sterile sugar water (1kg/L of sucrose in water).

Experimental colonization of gnotobiotic bees

All honey bees were collected from hives at the apiary of the University of Lausanne. A different hive was selected for each colonization replicate. Microbiota-free bees were raised according to established protocols [41]. Briefly, frames containing capped brood were taken from the hive and washed quickly with a towel soaked in 3% bleach. Pupae of the appropriate age were then removed from the frames and placed into sterile emergence boxes containing 0.2 µm filter sterilized sucrose (50% w/w) sugar water. The boxes were placed in an incubator at 35 °C for 2 days, until the adult bees emerged. They

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

were then randomly transferred to separate, sterile cages assembled from plastic drinking cups and Petri dishes.

Bees were colonized by hand with 5 µL of bacterial solution, or corresponding blank control, after starving them of sugar water for 1-2 hours, stunning them on ice for 5 minutes, and transferring them with tweezers to individual microcentrifuge tubes that were cut to enable insertion of a pipette tip. Sugar water and, where specified, polyfloral pollen (Bircher Blütenprodukte) sterilized via electron beam irradiation (Studer Cables AG) were provided to cages *ad libitum* throughout each experiment. Mono- and co-colonization of bees with *S. alvi* wkB2 and the four *Gilliamella* strains were performed five times, using different hives for each treatment. Not all conditions were run in each experiment due to technical difficulties preparing the treatment groups, and the GC-MS analysis was only carried out for a subset of the samples (see **Table S2** for details**)**. The colonization experiment with the divergent *Snodgrassella* strains isolated from honey bees and bumble bees was carried out twice (see **Table S3** for details).

At the end of each experiment, bees were anesthetized with CO_2, placed on ice, and dissected with sterile forceps. Whole guts, minus the crop, were homogenized in a bead beater (FastPrep-24, MP) in tubes containing PBS and glass beads. The samples were each immediately divided into three aliquots. One was plated to count Colony Forming Units (CFUs) on BHIA as well as for contaminants on Nutrient agar (NA) and De Man, Rogosa and Sharpe agar (MRSA). Another aliquot of 500 µL was centrifuged for 15 minutes at 4 °C and 20,000 g^{-1} and the liquid supernatant snap frozen in liquid nitrogen and stored at -80 °C until metabolomic extraction. The third aliquot was frozen at -80 °C until DNA extraction.

Validation of bee sterility

To rule out contamination in the guts of newly emerged bees, two bees from each emergence box in each experiment were sacrificed and 10 µL of their homogenized guts plated on BHIA (5% CO_2), NA (aerobic) and MRSA (anaerobic) at 35ºC to check for sterility prior to colonizing the remaining bees.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

In the event of contamination detected in newly emerged bees, all bees from that emergence box were discarded from the experiment. The same contamination checks were also performed on all analyzed bees at the end of each experiment. Additional sterility checks were performed in the experiment where we compared gut metabolite levels in newly emerged and six-day old MF bees. Here we also plated eight homogenized guts on Potato dextrose agar (PDA, aerobic), Lysogeny broth agar (LBA, aerobic), Tryptic soy agar (TSA, 5% CO_2), Columbia agar + 5% sheep blood (CBA+5%SB, 5% CO_2 & anaerobic), and Tryptone yeast extract glucose agar (TYG, anaerobic).

Quantification of pollen consumption

Both whole bees and dissected bee guts were weighed to assess treatment effects and quantify pollen consumption. Whole bee and wet gut weights were both significantly higher in bees provided pollen than in those provided sugar water only, while colonization did not significantly affect weight (**Fig. S4**). (Linear mixed effects models fitted by ML with nested, random cage effects: [Bee weight, n=271, Diet: $F_{(1,262)}$=176.1, p<.001; *S. alvi*: $F_{(1,262)}$=0.1, p=.769; *Gilliamella*: $F_{(1,262)}$=1.0, p=.311; *S. alvi:Gilliamella*: $F_{(1,262)}$=0.7, p=.388]; [Gut weight, n=271, Diet: $F_{(1,262)}$=220.0, p<.001; *S. alvi*: $F_{(1,262)}$=0.3, p=.595; *Gilliamella*: $F_{(1,262)}$=0.02, p=.880; *S. alvi:Gilliamella*: $F_{(1,262)}$=2.3, p=.131]. We also weighed filled pollen troughs at the beginning and end of the experiment and calculated the amount of pollen consumed per bee. Unsurprisingly, the mean difference in gut weights between dietary groups (22.3 ± 18.3 mg/bee) closely matched the measured amount of pollen consumed by the bees across colonization treatments (24.3 ± 7.2 mg/bee) (**Table S2**). We then calculated the bee body weight (bee weight – gut weight) to determine if diet or colonization led to actual tissue weight gain in bees. While we found a slight increase in body weight from pollen (95% CI [0.69 mg, 7.51 mg]), we cannot rule out that this was in fact due to variable amounts of pollen present in the crop, which was not removed with the hindgut (Gut weight, n=271, Diet: $F_{(1,262)}$=5.2, p=.023; *S. alvi*: $F_{(1,262)}$=1.4, p=.240; *Gilliamella*: $F_{(1,262)}$=0.6, p=.442; *S. alvi:Gilliamella*: $F_{(1,262)}$=3.8, p=.052).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Isotope tracing experiments

To assess the ^{13}C metabolite enrichment in the bee gut, MF bees were provided a 45:65 ratio of 99% uniformly labelled U-$^{13}C_6$ glucose (Cambridge Isotope Laboratory) and naturally abundant sucrose (Sigma Aldrich) in the first experiment. For the NanoSIMS time-course colonization experiment, newly emerged MF bees were divided into two cages and provided pure 99% U-$^{13}C_6$ glucose (treatment) or naturally abundant (*i.e.* ~98.9% ^{12}C) -glucose (control) in water for five days (pulse) before switching both cages to a naturally abundant glucose diet (chase). On the morning of the fourth day all bees were starved for 1.5 hours and colonized with *S. alvi* wkB2. The ^{13}C glucose solution was replaced with a ^{12}C glucose solution 24 hours later. Bees were harvested at 0, 8, 24, 32, 48, 56, 72 hours after inoculation. This overlap of colonization and ^{13}C glucose feeding was chosen to both maximize the number of bacterial cells in the TEM sections and to reduce the dilution of ^{13}C label due to metabolic turnover. For each timepoint and both cages (^{13}C and ^{12}C glucose treatment), the ileum of four bees were dissected. One gut was immediately preserved for electron microscopy and NanoSIMS, while the other three guts were homogenized and aliquoted for CFU plating and metabolite extraction **(Fig. 3)**. The control samples (from the cage provided with ^{12}C glucose) were used to determine the natural ^{13}C abundance in the bee tissue and *S. alvi* cells. Image metadata and Raw NanoSIMS region of interest (ROI) values can be found on github: https://github.com/Yelchazli/Foraging-on-host-synthesized-metabolites-enables-S.-alvi-to-colonize-the-honey-bee-gut.

Quantification of bacterial loads via qPCR

DNA extraction was carried out by adding 250 µL of 2x CTAB to frozen gut homogenates. The mixture was homogenized at 6 m/second for 45 seconds in a Fast-Prep24TM5G homogenizer (MP Biomedicals) and centrifuged for 2 minutes at 2000 rpm. DNA was extracted in two-steps with Roti-Phenol and Phenol:Chloroform:Isoamylalcohol (25:24:1), then eluted in 200 µL DNAse/RNAse water and frozen. qPCR was run using target-specific primers for *Apis mellifera* actin, and specific 16S rRNA gene primers for *S. alvi* and *Gilliamella spp.* as well as 16S rRNA gene universal primers used in [12]. qPCR

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

reactions were run on a QuantStudio 3 (Applied Biosystems), and standard curves were performed using the respective amplicons on a plasmid as previously described [12].

In addition, 18S rRNA gene amplification with universal primers were run to assess fungal loads (F: TATGCCGACTAGGGATCGGG, R: CTGGACCTGGTGAGTTTCCC) [42]. qPCR reactions were run on 384-well plate QuantStudio 5 (Applied Biosystems). The corresponding standard curve for absolute quantification was performed using the 199 bp amplicon as opposed to a plasmid containing the amplicon. Data was extracted and analyzed as previously described [12].

Metabolite extractions

Frozen gut samples were divided and extracted via two established methods for short chain fatty acids (SCFAs) and soluble metabolites. Prior to soluble metabolite extraction from pure pollen grains, an equivalent amount to what was digested per bee was mixed with PBS. In order to approximate physio-chemical degradation of the pollen in the bee gut, the samples were bead beaten and incubated at 35ºC for 18 hours.

SCFAs were extracted with 750 µL diethyl ether from 75 µL of gut supernatant that was first acidified with 5 µL of 11% HCl. Isovalerate [200 µM] (Sigma Aldrich) was used as an internal standard. The solvents and resulting two-phase mixture were kept cold at -20 °C and vortexed 3x for 30 seconds. The samples were then centrifuged for 3 minutes at 4 °C and 13000 g^{-1}, and 80 µL of the ether phase was removed to a glass GC vial (Agilent). The sample was derivatized with 20 µL MTBSTFA + 1% TBDMS (Sigma Aldrich) for 1 hour at 30 °C [43].

Soluble metabolites were extracted via a modified Bligh and Dyer protocol [44]. A (5:2:1) mixture of methanol (Sigma Aldrich) and chloroform (Sigma Aldrich) and double distilled water was chilled to -20 °C and 800 µL added to 100 µL of gut homogenate. A mixture of norleucine (Sigma Aldrich), norvaline (Sigma Aldrich), and ^{13}C glucose (Cambridge Isotopes) was used as internal standards. The samples were vortexed three times for 30 seconds and extracted at -20 °C for 90 minutes. The samples were then centrifuged at 13,300 g^{-1} and 4 °C for 5 minutes. Liquid supernatant was transferred to a new

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

tube and 400 µL of cold chloroform : methanol (1:1) was added to the insoluble material. The samples were again vortexed three times, extracted for 30 minutes at -20 °C, centrifuged, and the two supernatants were combined. Phase separation was achieved by adding 200 µL of water and centrifuging at the same conditions. The upper phase was then dried overnight in a speed vacuum. The sample was derivatized with 50 µL of 20 mg/mL methoxyamine hydrochloride in pyridine (Sigma Aldrich), for 90 minutes at 33 °C followed by silylation with 50 µL of MSTFA (Sigma Aldrich) for 120 minutes at 45 °C.

GC-MS analysis

Samples were analyzed on an Agilent 8890/5977B series GC-MSD equipped with an autosampler that injected 1 µL of sample onto a VF-5MS (30m x 0.25 mm x 0.25 um) column. The SCFA samples were injected with a split ratio of 25:1, helium flow rate of 1 mL/min, and inlet temperature of 230 °C. The oven was held for 2 minutes at 50 °C, raised at 25 °C/min to 175 °C and then raised at 30 °C/minute to 280 °C and held for 3.5 minutes. The MSD was run in SIM mode with 3 target ions for each compound. The Soluble metabolite samples were injected with a split ratio of 15:1, helium flow rate of 1 mL/min and inlet temperature of 280 °C. The oven was held for 2 minutes at 125 °C, raised at 3 °C/minute to 150 °C, 5 °C/minute to 225 °C, and 15 °C/minute to 300 °C and held for 1.3 minutes. The MSD was run in scan mode from 50-500 Da at a frequency of 3.2 scan/second.

Metabolomics data analysis

Mass spectrometry features were identified by spectral matching to the NIST17 mass spectra library or to in house analytical standards. Features were then extracted using the MassHunter Quantitative Analysis software (Agilent). Data normalization and analysis was performed using custom R scripts (See Data availability). Absolute metabolite quantification was performed by fitting normalized response values to calibration curves of analytical standards. The effect of diet, and colonization with *S. alvi* and *Gilliamella* were assessed by fitting z-score normalized metabolite abundances to a mixed

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

468 linear model using the lmm2met package in R with diet, *S. alvi* and *Gilliamella* as fixed effect factors
469 and experimental replicate as a random effect [45]. The significance of each fixed effect was then adjusted
470 by the Benjamini, Hochberg, and Yekutieli method to control for false discovery. Mass isotopologues
471 were adjusted to account for natural isotope abundances following published procedures [46,47].

472

473 ## TEM section preparation and NanoSIMS analysis

474 Bee guts were dissected, the ileum was cut away from the rest of the gut and fixed in glutaraldehyde
475 solution (EMS, Hatfield, PA, US) 2.5% in Phosphate Buffer (PB 0.1M pH7.4) (Sigma, St Louis, MO,
476 US) for 2 hours at room temperature (RT) and kept at 4°C until sample preparation. Ileum pieces were
477 rinsed 3 times for 5 minutes in PB buffer and then postfixed with a fresh mixture of osmium tetroxide
478 1% (EMS, Hatfield, PA, US) with 1.5% potassium ferrocyanide (Sigma, St Louis, MO, US) in PB
479 buffer for 2 hours at RT. The samples were then washed three times in distilled water and dehydrated
480 in acetone solution (Sigma, St Louis, MO, US) at graded concentrations (30%-40min; 70%-40min;
481 100%-1h; 100%-2h). Dehydration was followed by infiltration in Epon resin (Sigma, St Louis, MO,
482 US) at graded concentrations (Epon 1/3 acetone-2 hours; Epon 3/1 acetone-2 hours, Epon 1/1-4 hours;
483 Epon 1/1-12 hours) and finally polymerized for 48 hours at 60°C in an oven. Thin sections of 100
484 nanometers (nm) were cut on a Leica Ultracut (Leica Mikrosysteme GmbH, Vienna, Austria) and
485 picked up on copper slot grid 2x1mm (EMS, Hatfield, PA, US) coated with a Polyetherimide (PEI) film
486 (Sigma, St Louis, MO, US). Sections were poststained with uranyl acetate (Sigma, St Louis, MO, US)
487 2% in H_2O for 10 minutes, rinsed several times with H_2O followed by Reynolds lead citrate for 10
488 minutes and rinsed several times with H_2O.

489 Large montages with a pixel size of 6.9 nm covering areas of around 120x120 µm were taken with a
490 transmission electron microscope Philips CM100 (Thermo Fisher Scientific, Hillsboro, USA) at an
491 acceleration voltage of 80kV with a TVIPS TemCam-F416 digital camera (TVIPS GmbH, Gauting,
492 Germany) and the alignment was performed using Blendmont command-line program from the IMOD
493 software [48].

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Four holes were made using the electron beam at the four corners of the montage to localize the area of interest for NanoSIMS imaging. Then the slot grid was deposited on 1 µL of distilled water on a 10 mm round coverslip and left to dry for 10 minutes. The slot grid was detached, letting the section attach to the coverslip.

After being coated with a 15nm gold layer to evacuate a potential build-up of a charge during the NanoSIMS analyses, the samples were analyzed by a NanoSIMS 50L (CAMECA). A primary beam of 16 keV Cs+ ions was focused to a spot size of about 100 nm on the sample surface, causing the sputtering of atoms and small molecules from the first 5-10 nm of the sample. Negative secondary ions were directed and focalized into a single beam before reaching the magnetic sector of the mass spectrometer where the ions were separated. The species of interest ($^{12}C^{14}N$- and $^{13}C^{14}N$- for this study) were then measured simultaneously using Electron Multipliers. To resolve the interference between $^{13}C^{14}N$- and $^{12}C^{15}N$-, two slits (ES and AS) were used to achieve a mass resolution higher than 10000 (Cameca definition), [49].

In the imaging mode, selected areas of 25x25 µm were implanted using a strong beam (using diaphragm D1-1 (diameter 750 µm) and a current of 16 pA) to remove the gold coating and reach stable emission conditions. After optimization of the extraction of the secondary ions, the area was imaged by scanning with a smaller beam (~100 nm, using diaphragm D1-5 (diameter 100 µm) and a current of 0.45 pA) with a 256x256 pixel resolution and a dwell time of 5µs/pixel. For each image, 10 layers were acquired. The image processing was performed using "L'image" software (Larry Nittler, Carnegie Institution of Washington). The 10 layers were aligned and stacked, 44 ms deadtime correction was applied. Regions of interests (ROI) were manually drawn around bacterial cells, gut epithelium layer, and bulk host cells. The resulting data is reported in **Dataset S1**. At% enrichment was calculated in R using the data.table package. At% enrichment formula: $13C\ At\% = {}^{14N\ 13C}/_{(14N\ 13C\ +\ 14N\ 12C)} \times 100$. **(Fig. S9).**

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

In vitro growth of *S. alvi*

S. alvi was cultured in Bee9 media, an M9-derived minimal defined media containing essential vitamins and mixtures of organic acids (see recipe in **Dataset S3** on github: https://github.com/Yelchazli/Foraging-on-host-synthesized-metabolites-enables-S.-alvi-to-colonize-the-honey-bee-gut). Growth assays were performed in 96-well plates (Corning costar) at 35 °C, 5% CO_2, with continuous orbital shaking (H1 Synergy, Biotek). OD_{600} measurements were taken every 30 minutes for 20 hours. All conditions were independently replicated three times. For metabolomics measurements, plates were immediately placed on ice at the end of each experiment, and liquid was transferred to microcentrifuge tubes. The tubes were then centrifuged at 13,300 g and 4 °C for 30 seconds, and the liquid supernatant was then snap frozen in liquid nitrogen and stored at -80 °C for metabolomic analysis.

Phylogenetic analysis of the kynureninase gene family

The kynureninase protein sequence of *S. alvi* wkB2 was retrieved from NCBI (WP 025330329.1), and blasted (NCBI, BlastP) against the "nr" database. Based on the BlastP hits, 227 divergent amino acid sequences were exported in fasta format and aligned using MUSCLE [50] with standard parameters. Tree inference and bootstrap validation were performed using IQtree with the following command: "iqtree-omp -s kyn_sequences_muscle.aln -nt 6 -m TEST -bb 1000" [51]. The tree topology was visualized using the ITOL online viewer (https://itol.embl.de/).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Data Availability

Raw data and scripts used to process data and generate figures are available on github at the following link: https://github.com/Yelchazli/Foraging-on-host-synthesized-metabolites-enables-S.-alvi-to-colonize-the-honey-bee-gut. Metabolomics data have been deposited to the EMBL-EBI MetaboLights database (DOI: 10.1093/nar/gkz1019, PMID:31691833) with the identifier MTBLS8058. https://www.ebi.ac.uk/metabolights/MTBLS8058

Acknowledgments

We would like to thank Silvia Moriano-Gutierrez, and Gonçalo Santos-Matos for providing helpful comments on the manuscript. This work was funded by the University of Lausanne, an ERC Starting Grant (MicroBeeOme), the NCCR Microbiomes, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (grant no. 180575) and a Swiss National Science Foundation project grant (grant no. 179487) to P.E.

Author Contributions

A.Q., Y.E. and P.E. conceived the project and designed the experiments. A.Q, and Y.E. performed the experiments. A.Q. analyzed the GC-MS results, while Y.E. analyzed the qPCR and NanoSIMS results. N.N. assisted with bee experiments, metabolite, and DNA extractions. A.M.N. performed preliminary in vitro growth experiments. J.D. performed the tissue sectioning and TEM imaging supervised by C.G. S.E. performed the NanoSIMS imaging and instructed on the data analysis, supervised by A.M., who also advised on the design of the NanoSIMS experiment. A.Q., Y.E. and P.E. wrote the manuscript, with contributions from S.E. and A.M. All authors reviewed the manuscript and provided feedback. As equally contributing first authors, A.Q. and Y.E. may each list their name first when referencing this work.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figures

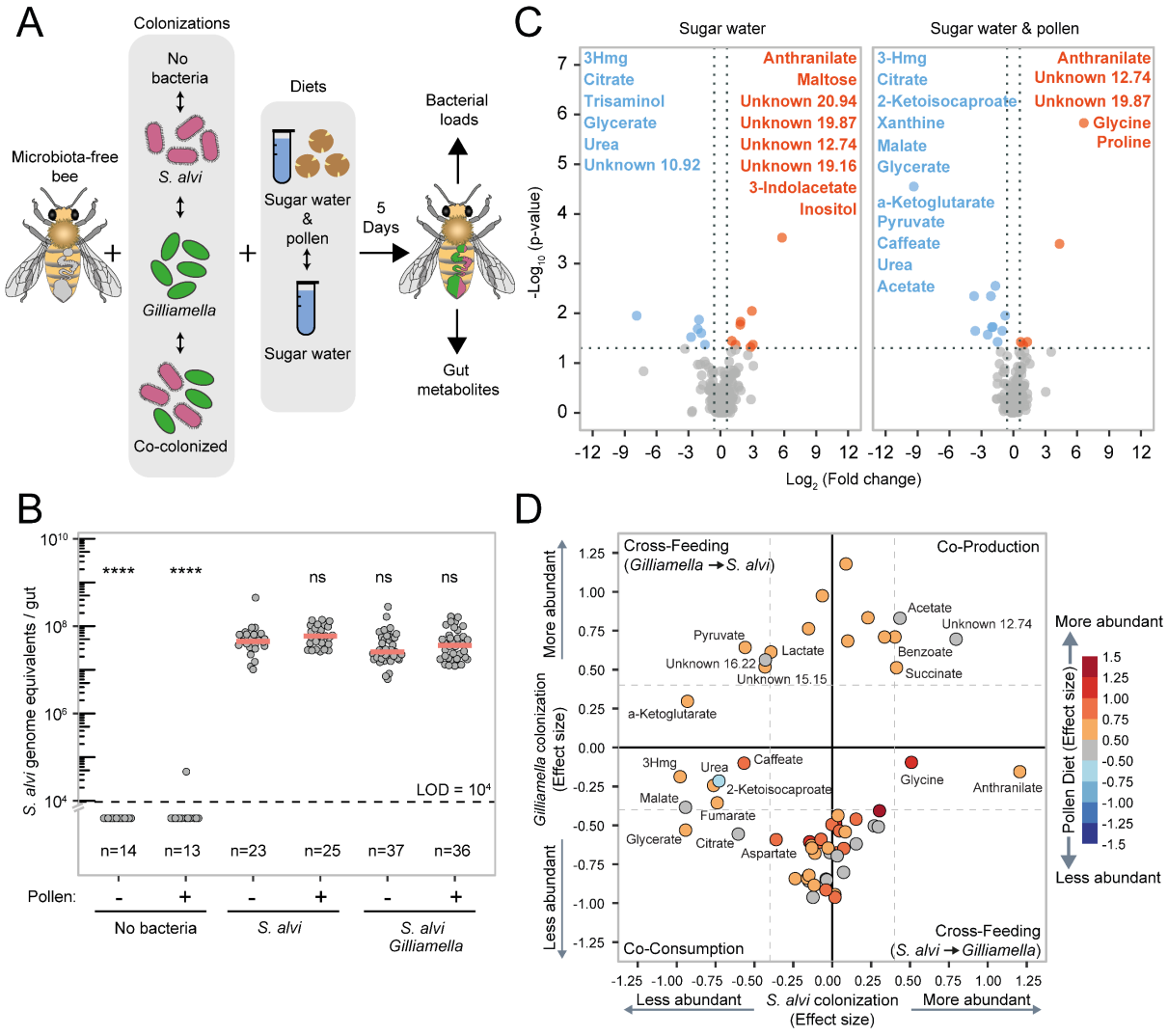


Figure 1. *S. alvi* colonizes bees fed with sugar water only and depletes host, pollen, and *Gilliamella*-derived metabolites. (A) Outline of the bee colonization experiment with four colonization treatments (no bacteria, *S. alvi*, *Gilliamella*, or co-colonized) and two dietary treatments (sterile sugar water and pollen, or sterile sugar water only). Five days after colonization gnotobiotic bees were sacrificed and the bacterial load and the metabolites in the gut quantified. **(B)** Number of *S. alvi* cells (number of bacterial genome equivalents) per bee gut determined by qPCR using *S. alvi*-specific 16S rRNA gene primers. Each value represents the number of cells in a single bee gut, red bars represent the median,

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

horizontal black line represents qPCR limit of detection (LOD = < 10^4). For each treatment, the presence/absence of pollen in the diet is indicated by +/- sign. Significance determined by Wilcoxon rank sum test with *S. alvi* (pollen -) as reference group, ns, non-significant; ****$P < 0.001$. **(C)** Significantly differentially abundant gut metabolites between *S. alvi* colonized and MF bees (i.e., 'no bacteria' treatment). Significantly depleted and produced metabolites are shown in blue and orange, respectively. Adjusted significance values were calculated using Wilcoxon rank sum test, and Benjamini-Hochberg correction, using the MF samples as a reference. **(D)** Results from mixed-linear modelling show the contributions (fixed effect sizes) of microbial colonization and diet towards z-score normalized metabolite abundances. Only metabolites significantly correlated with *S. alvi* colonization are named. Large values (+/-) indicate a strong correlation between a factor and a corresponding change in metabolite abundance. Values centered around zero (inside the dotted lines) are not significant. The effect size of pollen is indicated by the color of each metabolite, red colors indicate significantly positive correlation, while blue is a significant negative correlation. The four corners of the graph represent the four types of metabolic interactions between *S. alvi* and *Gilliamella* (from the top left: cross-feeding from *S. alvi* to *Gilliamella*, co-production by both bacteria, cross-feeding from *Gilliamella* to *S. alvi,* co-consumption). Carboxylic acids are grouped principally on the lower-left section of the plot indicating they are less abundant with *S. alvi* and *Gilliamella* colonization.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

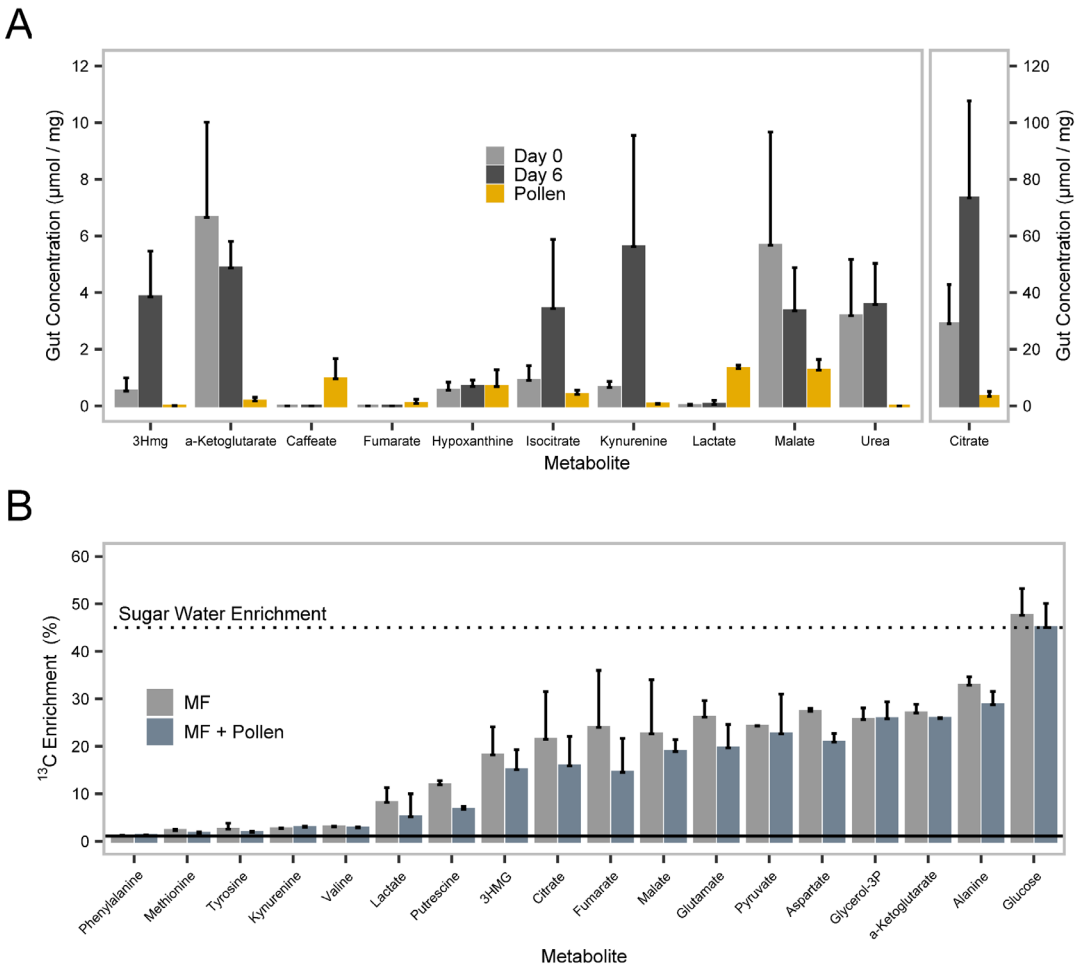


Figure 2. Gut metabolite abundances and isotope labelling reveals the host origin of *S. alvi* substrates. (A) Abundances of selected metabolites in the guts of MF bees increase from emergence (Day 0) to adult age (Day 6) (n=8). All metabolites except for caffeate and lactate were less abundant when extracted directly from an equivalent amount of pollen consumed per bee over six days (n=8). Quantitates from pollen extract were adjusted based on average total pollen consumption per bee (see methods). **(B)** The average (n=4) ^{13}C metabolite enrichment (%) in bee guts shows that carboxylic acids and non-essential amino acids become highly labelled upon feeding ^{13}C-glucose to MD bees. Dark and light bars indicate, respectively, absence and presence of pollen in the diet. The solid line denotes the natural ^{13}C enrichment (1.108%), while the dashed line denotes ^{13}C enrichment of sugar water solution fed to experimental bees.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

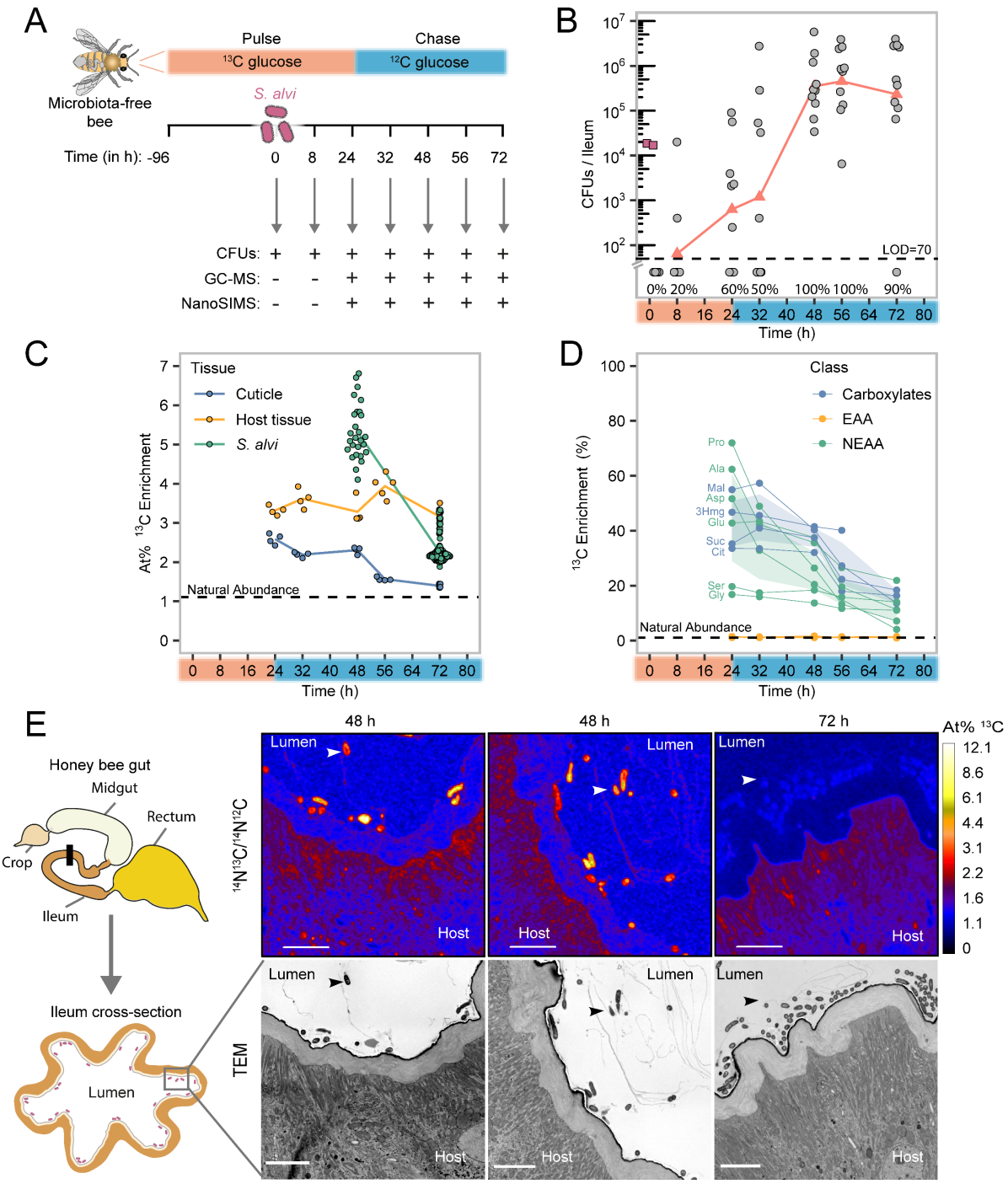


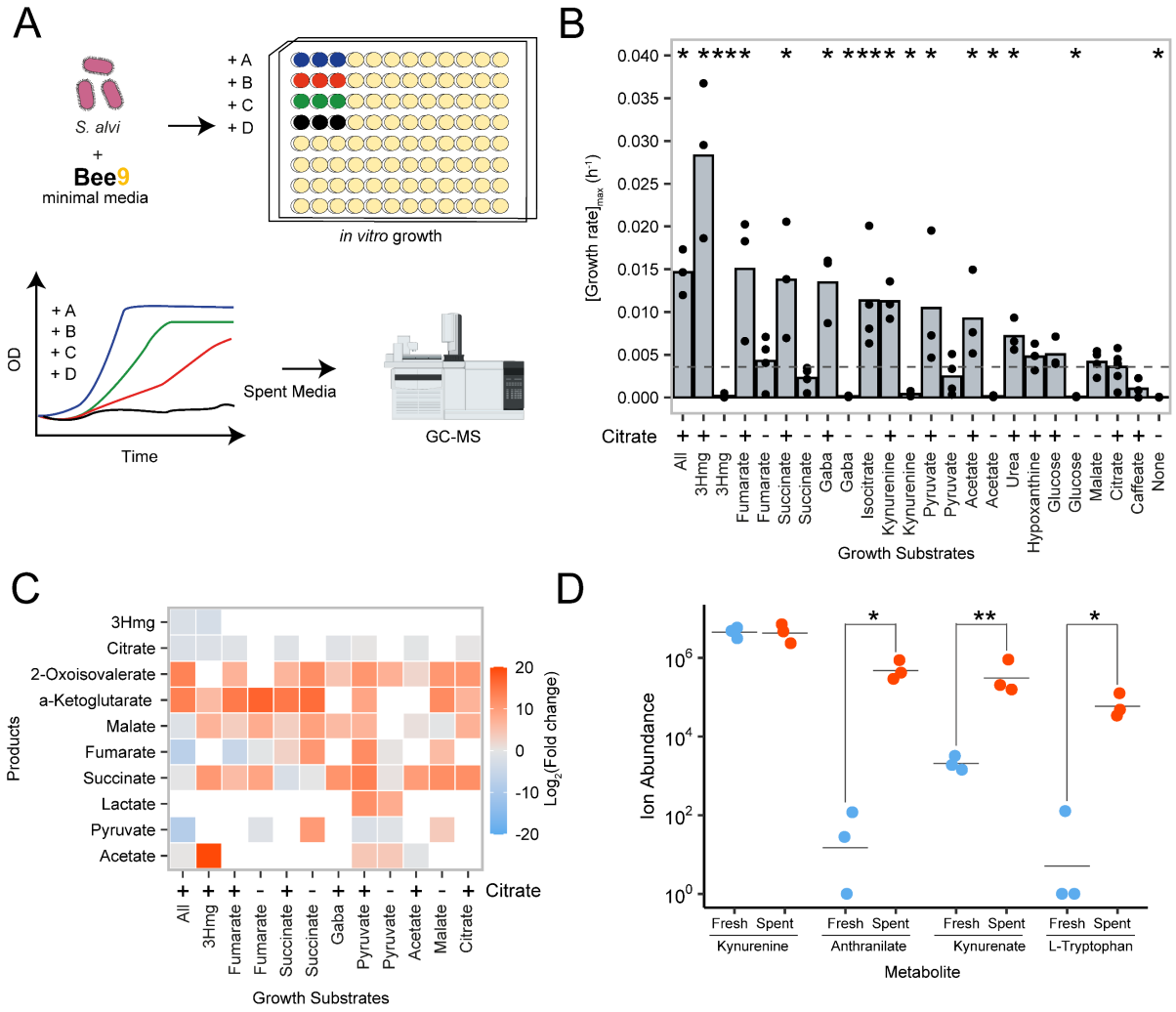
Figure 3. *S. alvi* builds biomass from substrates derived from the glucose metabolism of the host. **(A)** Experimental design of the pulse-chase isotope labelling gnotobiotic bee experiment. Newly emerged MF bees were fed 100% ^{13}C glucose for four days (i.e., 96 hours), colonized with a defined quantity of *S. alvi* and then the ^{13}C glucose was replaced by ^{12}C glucose 24 hours after colonization. The timeline indicates when bees were sampled for CFU plating, GC-MS and NanoSIMS analysis relative to the timepoint of colonization (i.e. 0h). **(B)** Colonization levels of *S. alvi* in the ileum until 72 hours

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

post inoculation show the initial strong colonization bottleneck and subsequent exponential increase in cell numbers. Colonization success is indicated as percentage of bees with detectable CFUs. Pink squares indicate the inoculum of *S. alvi* (OD= 0.1). The limit of detection (LOD) of 70 CFUs per ileum is indicated with the dashed line. **(C)** The average At% ^{13}C enrichments of *S. alvi* cells decreases rapidly, while the enrichment of host cells and epithelium remains constant over 72 hours in ileum cross-sections imaged with NanoSIMS. Each data point represents the mean of regions of interest (ROI); bacterial ROIs consist of single bacterial cells, while epithelium and host cells ROIs represent the total area encompassing epithelium or host cell tissue in each image. ROI raw counts and calculated enrichments are listed in **Dataset S1**. The dashed line at 1.1 % indicates the natural ^{13}C abundance calculated from NanoSIMS images of ^{12}C control bees. **(D)** ^{13}C enrichment of metabolites in the gut steadily decreases in the ^{12}C chase phase. Values represent the average ^{13}C enrichment of a given metabolite across the bees sampled at that timepoint. Average values and colors are shown by metabolite class: carboxylic acids (violet), non-essential amino acids (NEAA, green), essential amino acids (EAA, brown). Measured metabolites are listed in **Dataset S1**. The dotted line at 1.1 % indicates the natural ^{13}C abundance. **(E)** NanoSIMS and corresponding TEM images of 2 different ^{13}C labeled bees at 48 and 72 hours after inoculation. White arrows indicate *S. alvi* cells. At% ^{13}C represents percentage of ^{13}C atoms; the natural ^{13}C abundance is ~1.1 at%. Scale bars = 5 μM. Schematic drawing of the honey bee gut shows the region in the ileum (black bar) where cross-sections were taken and which regions of the cross-sections were imaged.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

628



629

630 **Figure 4. Host synthesized compounds enhance *S. alvi* growth *in vitro*. (A)** Schematic diagram
631 depicting growth and spent supernatant analysis. The "Bee9" media was derived from M9 base media
632 (**Dataset S3**). **(B)** The maximum growth rate across single carbon sources with (+) or without (-) citrate
633 added to the medium, reveals that growth on citrate + 3Hmg was faster than for all other conditions,
634 including "All", where all substrates were pooled together. The average value (bars) was calculated
635 from the individual biological replicates (dots). Adjusted significance values were calculated using
636 Wilcoxon rank sum test, and Benjamini-Hochberg correction, using the growth on citrate as the
637 reference state (dotted line). **(C)** Abundances of carboxylic acids in the spent supernatant, relative to
638 growth on citrate alone, varies with growth substrates. While some carboxylic acids including a-
639 ketoglutarate and malate are produced across most conditions, acetate is a growth substrate, but only

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

produced with either pyruvate or 3Hmg in the media. **(D)** Addition of kynurenine to the media leads to the production of anthranilate, kynurenic acid, and L-tryptophan. Significant changes between fresh and spent media were calculated using a paired t-test. (* P < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001).

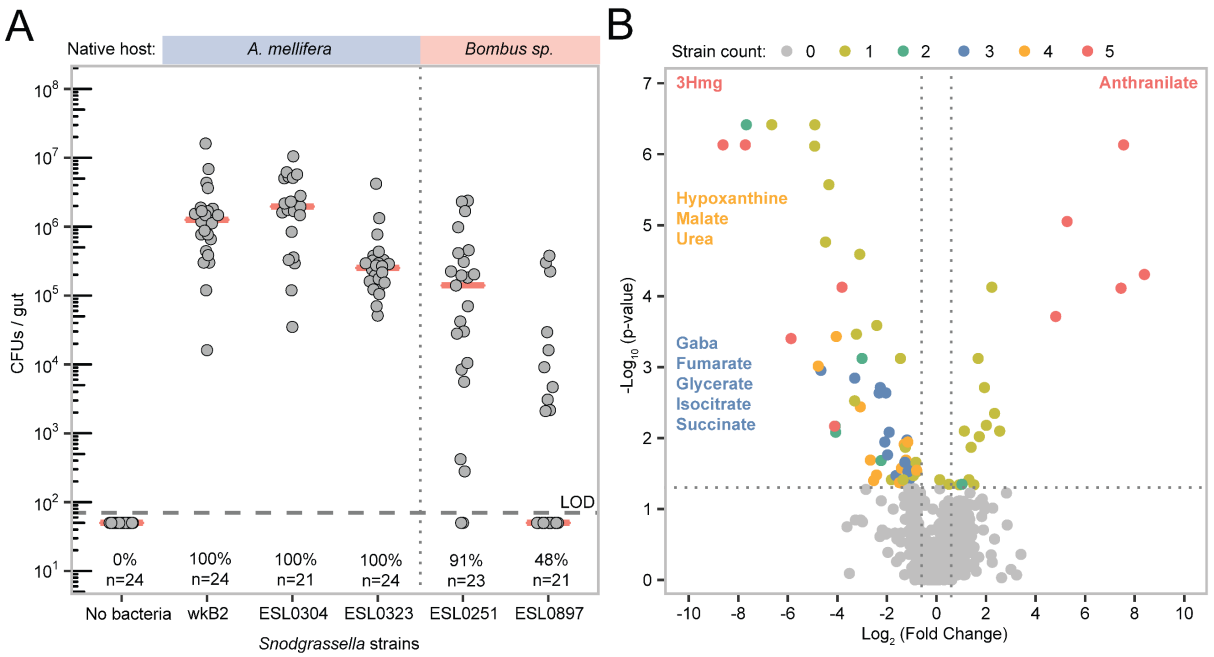
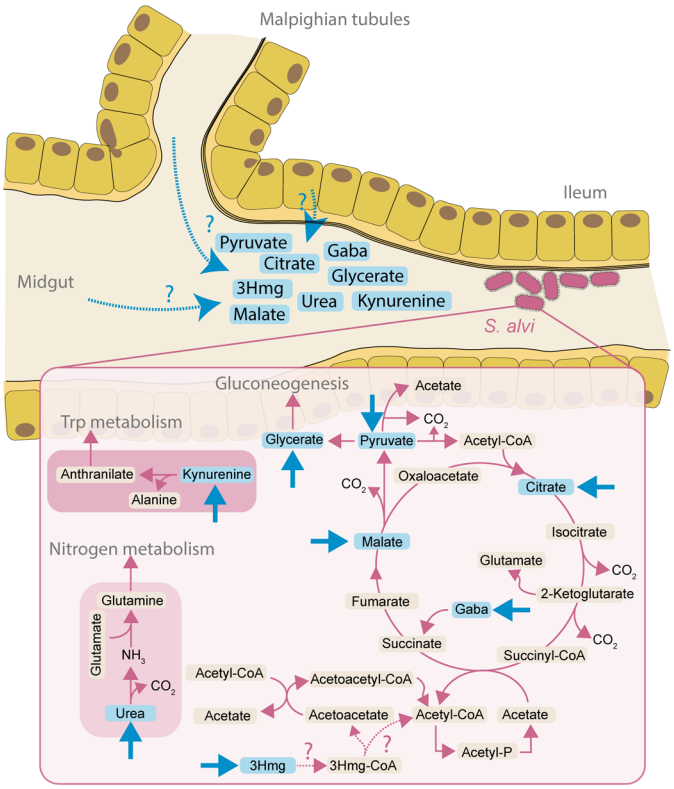


Figure 5. Gut colonization in the presence of a sugar-restricted diet and depletion of host-derived metabolites is conserved across divergent *Snodgrassella* strains. (A) Colonization levels of divergent *Snodgrassella* strains in monocolonized bees. Pink bars denote the median of CFUs/gut for each group. The LOD=70 colonies/gut is shown with a horizontal dashed line. The top bar indicates *Snodgrassella* strains native to *A. mellifera* versus *Bombus sp.* Colonization success (%) is shown for each strain, defined as guts with detectable CFU counts. (**B**) Volcano plot of gut metabolites showing the similarities in metabolite changes from MF vs colonized bees across strains. Each metabolite is plotted five times (once per *Snodgrassella* strain), and color coded by the number of strains in which it is significantly differentially abundant vs MF bees control. Only 3Hmg and anthranilate were significantly depleted or enriched by all five strains, while four of the strains depleted hypoxanthine, malate, and urea, and three of the five depleted Gaba, fumarate, glycerate, isocitrate and succinate.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



655

656 **Figure 6. Schematic model of *S. alvi* metabolism of host-derived compounds in the ileum.** Host
657 metabolites (blue) enter the ileum through the epithelial cells, or through the malpighian tubules
658 upstream of the ileum. *S. alvi* primarily utilizes the TCA cycle to generate energy, then synthesizing
659 biomass through gluconeogenic reactions. 3Hmg provides a source of acetate to fuel the TCA cycle.
660 However, the reactions synthesizing the first two steps (dotted pink arrows) have not been annotated in
661 the genome. Additionally, host synthesized urea is a source of nitrogen, and kynurenine is converted to
662 anthranilate as part of tryptophan metabolism. Blue arrows indicate the entry of substrates provided by
663 the host metabolism.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Supplementary Figures and Tables

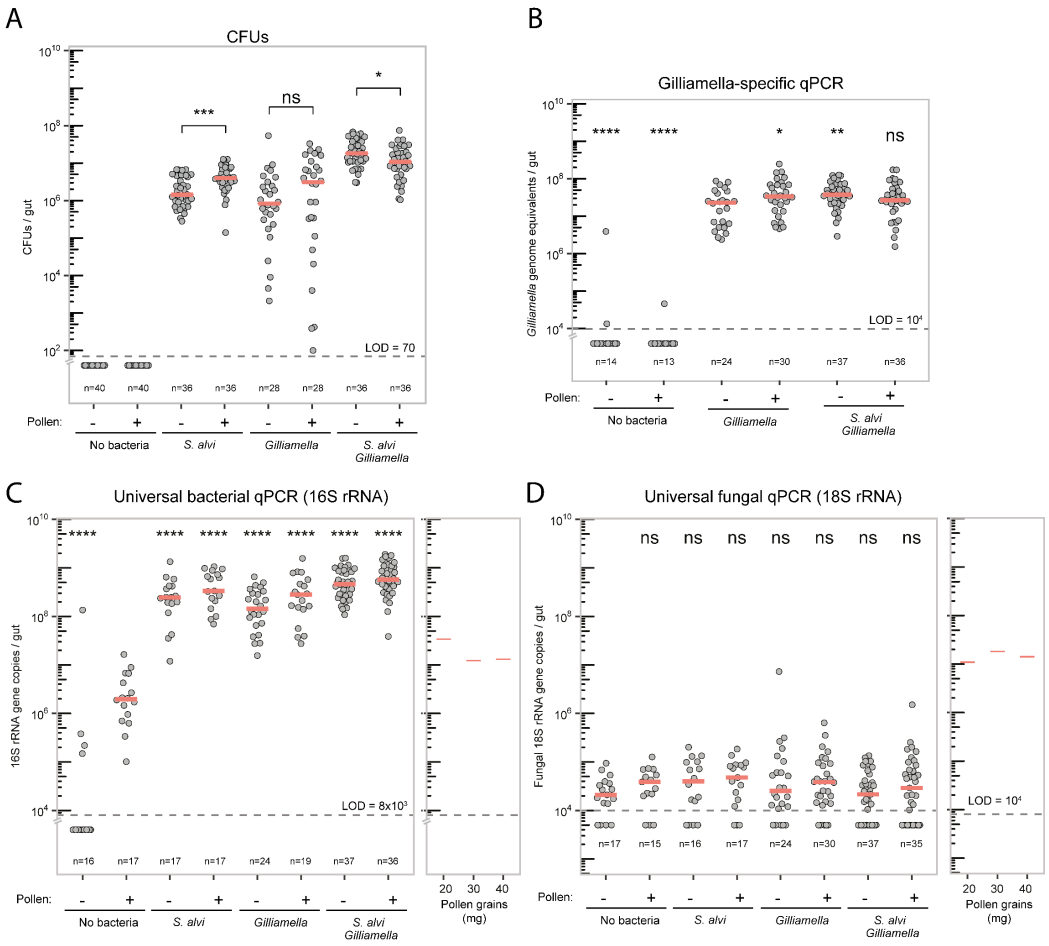


Figure S1. Total number of bacteria and fungi and colonization levels of *Gilliamella* per bee gut. **(A)** Total number of bacteria (both *S. alvi* and *Gilliamella*) was determined by plating homogenized guts on BHIA and counting CFUs. Each dot represents a single bee gut and red bars represent the median value for each group. Horizontal dotted line represents LOD = < 70. No colonies were observed in samples plotted below the LOD line. Wilcoxon rank sum test was done with *S. alvi* (Pollen -) as reference group, ns, nonsignificant; ****P < 0.001. (**B**) Bacterial cells (number of bacterial genome equivalents) per bee gut determined by qPCR using genus-specific *Gilliamella* 16S rRNA gene primers. Each dot represents the number of cells in a single bee gut, red bars represent the median, horizontal black line represents qPCR LOD =< 10^4. The presence of pollen in the diet is indicated by (P). Wilcoxon rank sum test was done with *Gilliamella* (Pollen -) as reference group, ns, nonsignificant; ****P < 0.001. (**C**) Number of 16S rRNA gene copies per bee gut and in three different quantities of

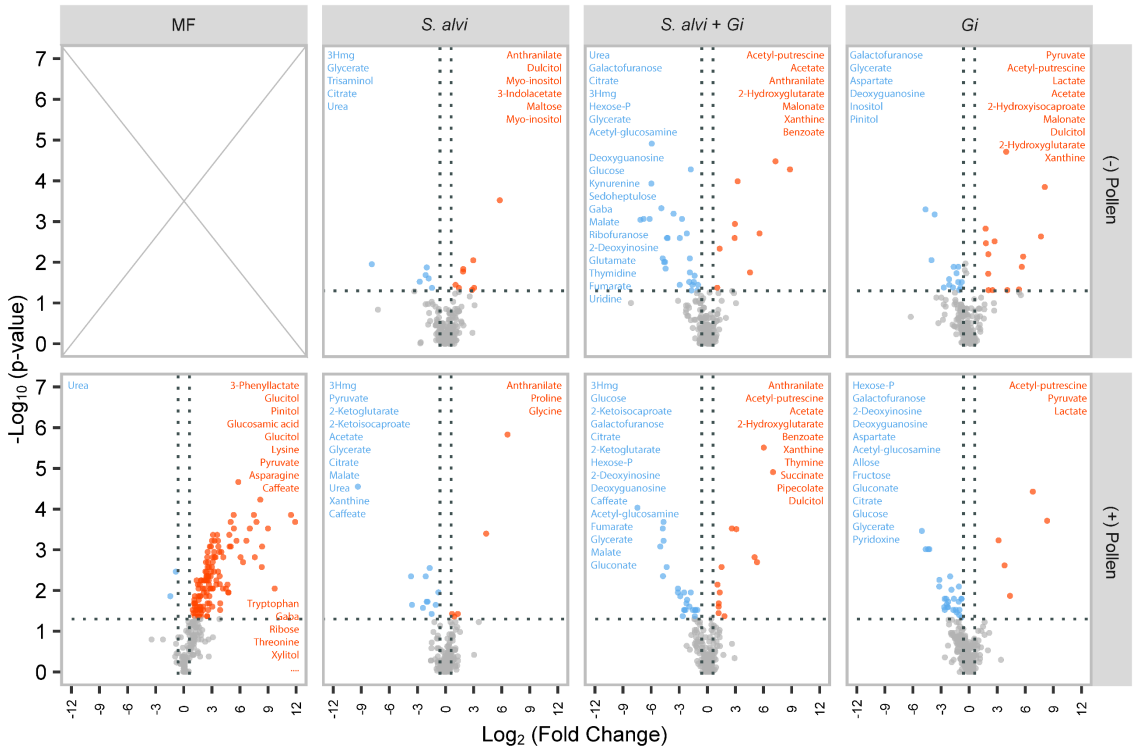
bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

677 sterilized bee pollen as determined by qPCR using universal 16S rRNA gene primers. Each dot
678 represents the 16S rRNA gene copies in a single bee gut, red bars represent the median, horizontal black
679 line represents qPCR LOD = < $8x10^3$. The presence of pollen in the diet is indicated by (P). Wilcoxon
680 rank sum test was done with MF (Pollen -) as reference group, ns, nonsignificant; ****P < 0.001. The
681 number of 16S rRNA gene copies detected in DNA extracted from different quantities of sterile pollen
682 used in the bee diet is shown in a separate plot. The high background signal in samples containing pollen
683 comes from the amplification of chloroplast DNA as shown in previous studies [52]. (**D**) Number of 18S
684 rRNA gene copies per bee gut and in three different quantities of sterilized bee pollen as determined by
685 qPCR using universal 18S rRNA gene primers. Each dot represents the 18S rRNA gene copies in a
686 single bee gut, red bars represent the median, horizontal black line represents qPCR LOD = < 10^4. The
687 presence of pollen in the diet is indicated by (P). Wilcoxon rank sum test was done with MF (Pollen -)
688 as reference group, ns, nonsignificant; ****P < 0.001.

689

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

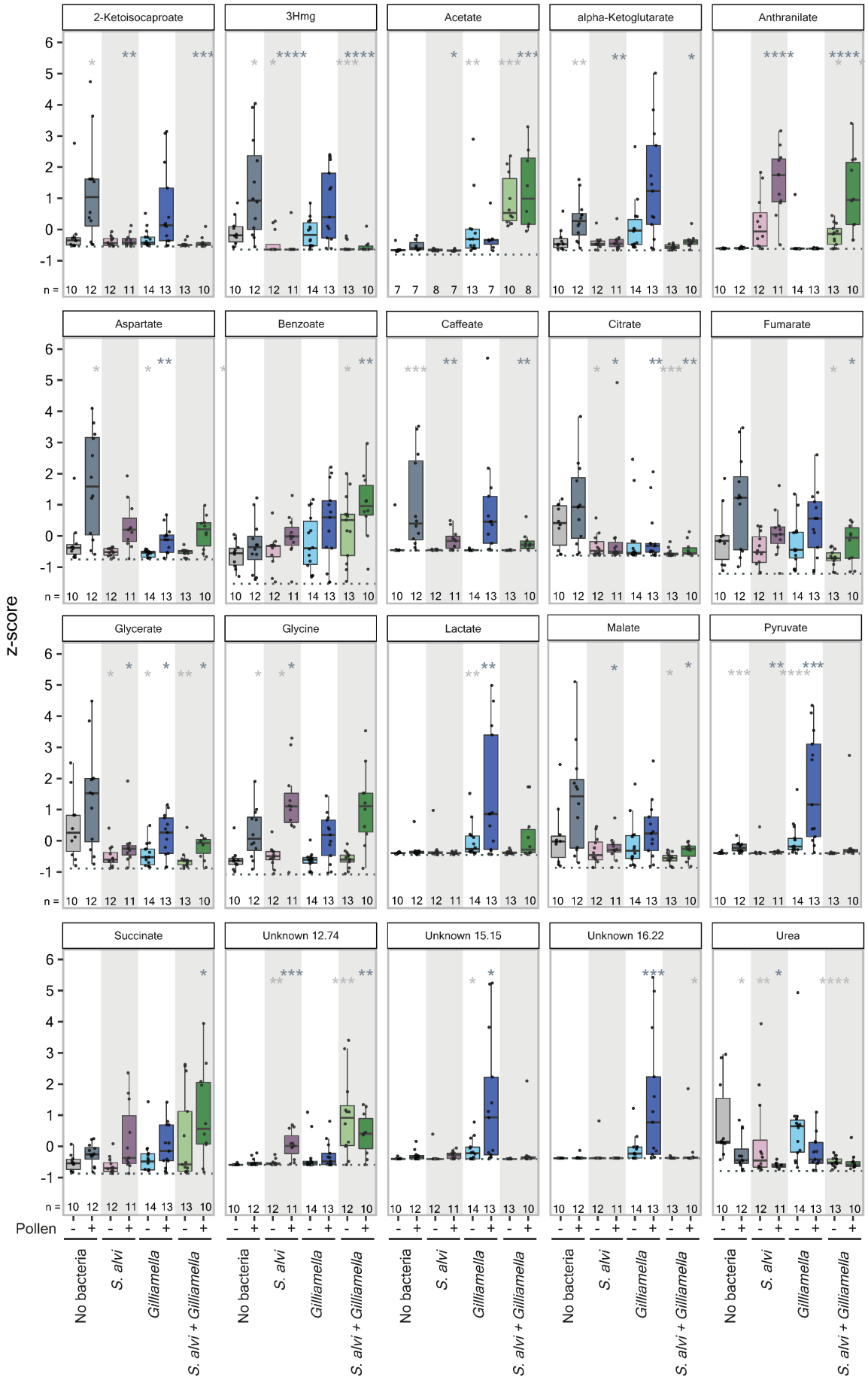
690



691

692 **Figure S2. Volcano plots highlighting gut metabolite abundance changes between colonized and**
693 **microbiota-free (MF) bees with and without pollen.** Significantly differentially abundant metabolites
694 between mono- and co-colonized *S. alvi* and *Gilliamella (Gi)* vs MF bees, as well as between MF ((+)
695 Pollen) and MF ((-) Pollen) bees. Significantly depleted and produced metabolites are shown in blue
696 and orange, respectively. Annotations are listed in order of significance for identified metabolites, with
697 isomers and other duplications removed. The full set of significantly more abundant metabolites in the
698 first panel can be found in the supporting scripts. Adjusted significance values were calculated using
699 Wilcoxon rank sum test, and Benjamini-Hochberg correction, using the MF conditions as a reference.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure S3. Z-score normalized metabolite abundances of compounds significantly affected by *S. alvi* colonization. Box plots and individual data points display values across all 8 experimental conditions. Only metabolites in which a mixed linear model showed a significant effect of *S. alvi* colonization are plotted. Multiple types of metabolic interactions are shown. For example, metabolite depletion by *S. alvi* exclusively (2-ketoisocaproate, 3Hmg, caffeate, urea) or by both species (aspartate, citrate, glycerate), metabolite production by *S. alvi* exclusively (anthranilate, unknown 12.74), cross-feeding from *Gilliamella* to *S. alvi* (pyruvate, lactate), synergistic metabolite depletion (fumarate, malate) and production (acetate, benzoate). The light and dark gray asterisks show groups significantly different from MF bees fed sugar water and pollen or only sugar water (also indicated by +/- sign), calculated using Wilcoxon rank sum test, * $P < 0.05$, ** < 0.01, *** < 0.001, **** < 0.0001. The dotted line represents the transformed LOD for each metabolite.

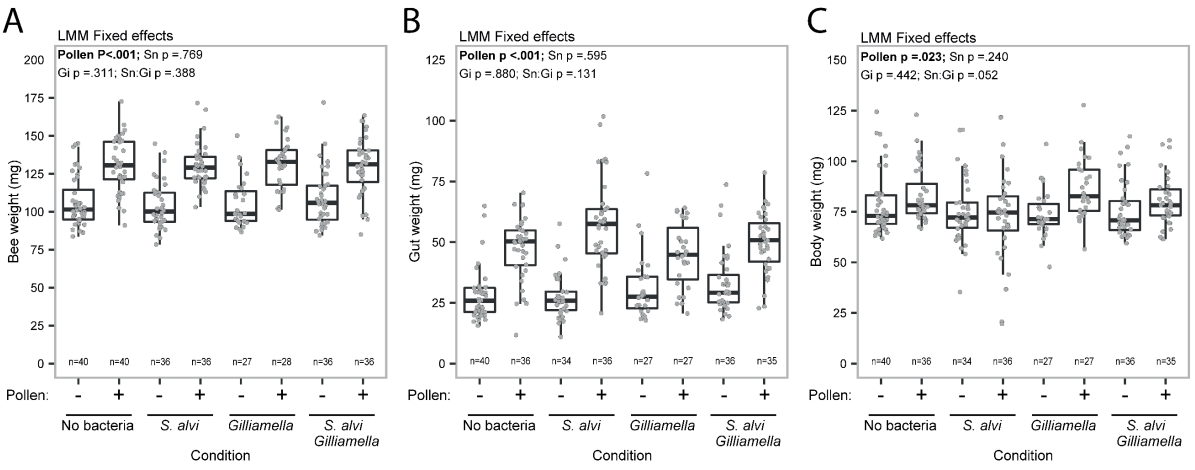
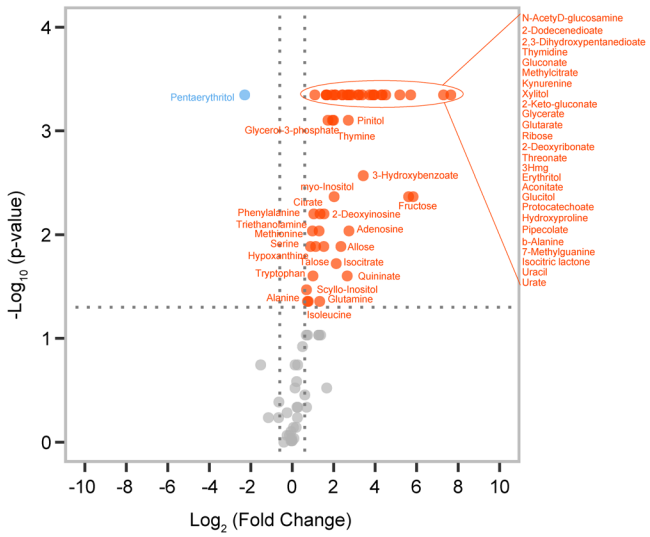


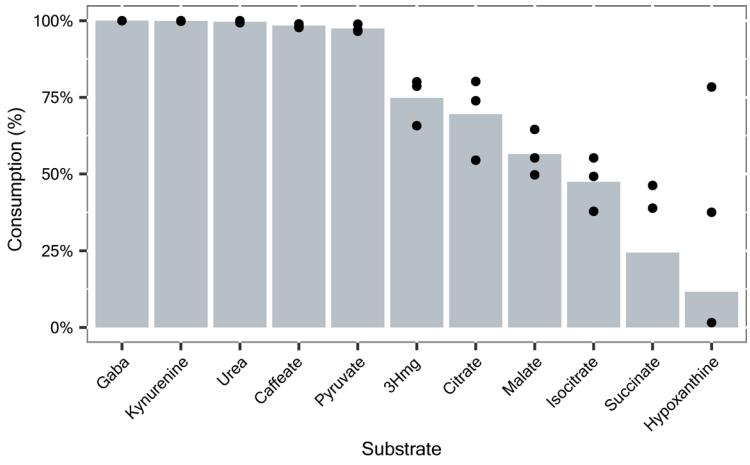
Figure S4. Pollen consumption but not colonization increases bee gut weight. (**A**) Whole bee and (**B**) dissected bee gut weights are strongly influenced by pollen in the diet, but not by colonization treatment. (**C**) The body weight, or difference between bee and gut weights, was also slightly higher across colonization groups for bees fed a pollen diet. The p-values (ANOVA) are shown for the fixed effect terms in the linear mixed models fitted by ML with nested, random cage effects.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



720

721 **Figure S5. Volcano plot highlighting gut metabolic changes between Day 0 and Day 6 after**
722 **emergence in MF bees without pollen.** Many metabolites are more abundant in the gut on Day 6 than
723 on Day 0 after emergence. Significantly less and more abundant metabolites on Day 6 are shown and
724 annotated in blue and orange, respectively, with isomers and other duplications removed. Adjusted
725 significance values were calculated using Wilcoxon rank sum test, and Benjamini-Hochberg correction,
726 using the day zero conditions as a reference (n=8).



727

728 **Figure S6. *In vitro* consumption of carbon substrates.** The percentage of each carbon source
729 consumed by the end of the *in vitro* growth experiment is shown for each tested carbon source. Data
730 was taken from the condition in which all carbon sources were added to Bee9 media in equimolar
731 amounts. Bars represent the average value from 3 biological replicates (black dots).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

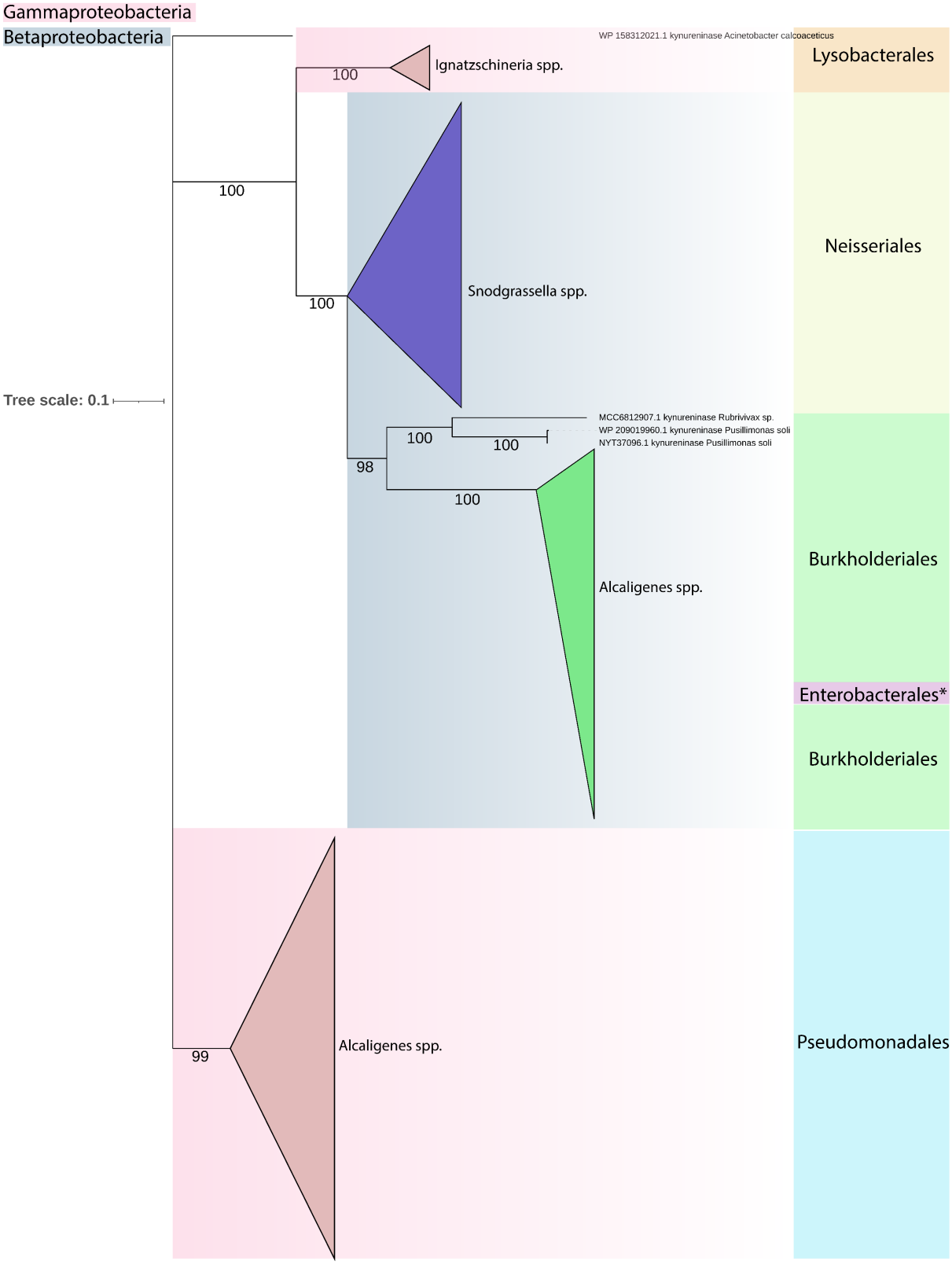


Figure S7. Phylogenetic tree of the kynureninase gene family found in *S. alvi*. The amino acid sequences of 227 homologs of the kynureninase gene identified in *S. alvi* wkb2 were retrieved from NCBI. The tree was inferred with Iqtree after aligning the sequences with MUSCLE (see methods). Tips are collapsed for clades that contain sequences of the same genus. Colored boxes on the right

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

indicate the bacterial family. Bootstrap values are indicated before every node. Gammaproteobacteria are highlighted in pink and Betaproteobacteria in blue. Asterisk indicates a sequence (MBY6345856.1) from Providencia rettgeri, a Gammaproteobacterium, which clusters together with homologs of the kynureninase of Betaproteopbacteria.

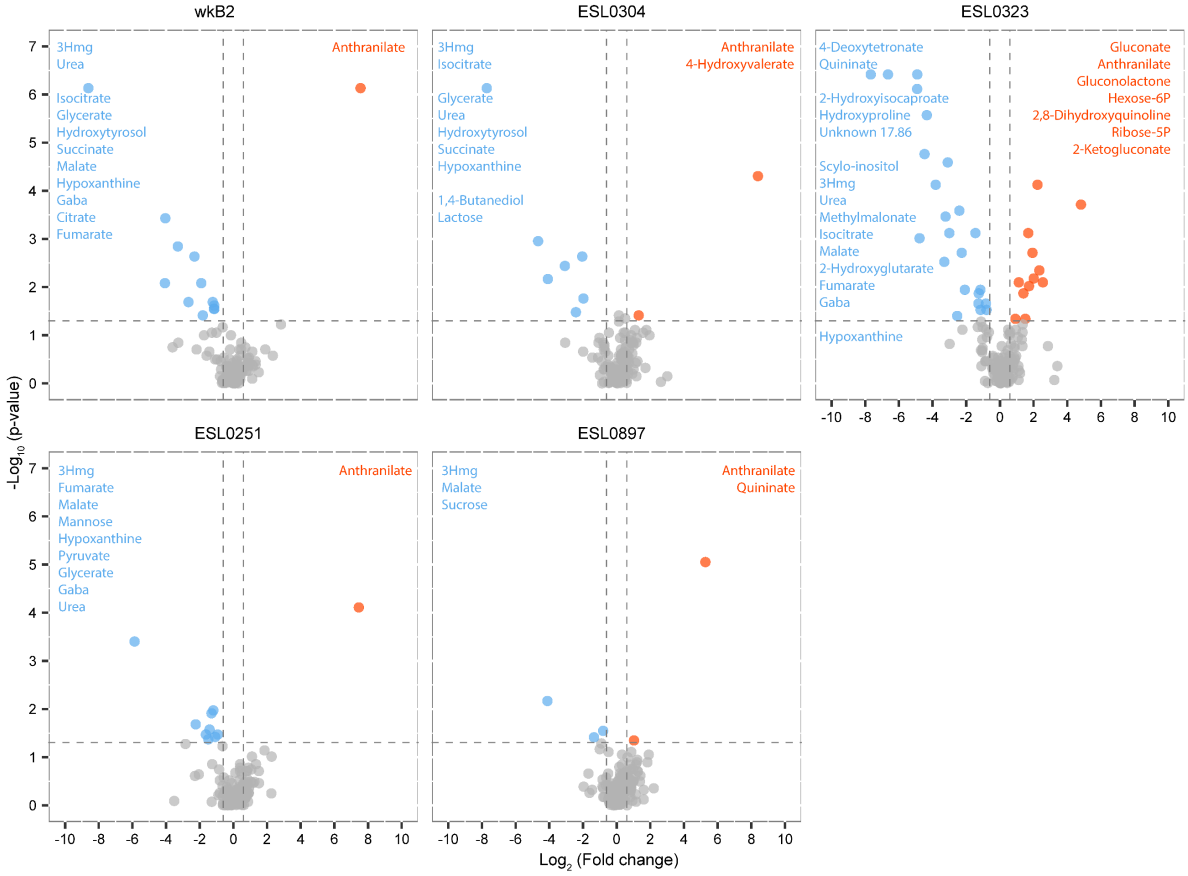


Figure S8. Gut metabolite changes in bees mono-colonized with divergent *Snodgrassella* strains relative to microbiota-free (MF) bees. Volcano plot of gut metabolites showing the similarities in metabolite changes from colonized vs MF bees across strains. Top: Strains native to *A. mellifera*; Bottom: Strains native to *Bombus sp.* Significantly depleted and produced metabolites are shown in blue and orange, respectively. Adjusted significance values were calculated using Wilcoxon rank sum test, and Benjamini-Hochberg correction, using the MF condition as a reference.

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

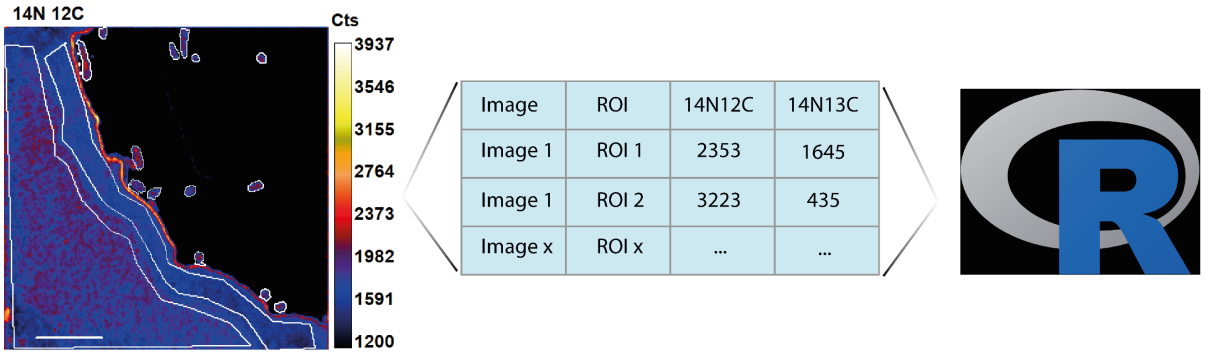


Figure S9. NanoSIMS image analysis workflow. For each image, 10 layers were acquired by bombarding the surface of the samples with Cs+ Ion beam (See methods). The image processing was performed using "L'image" software (Larry Nittler, Carnegie Institution of Washington). The 10 layers were aligned and stacked and 44 ms deadtime correction was applied. Regions of interests (ROI) were manually drawn around bacterial cells, gut epithelium layer, and bulk host cells, using a minimum threshold specified and values were extracted and reported in **Dataset S1**. At% enrichment was calculated in R using the data.table package. At% enrichment formula: $13C\ At\% = {}^{14N\ 13C}/_{(14N\ 13C\ +\ 14N\ 12C)} \times 100.$

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

759 **Table S1. Bacterial strains used in this study.** [12,53,54]

Bacterial strain	**# 16S rRNA copies**	**Culturing condition**	**Strain source**	**Place of origin**
Gilliamella apicola wkB1^T	4	BHIA, 35°C, microaerophilic	(Kwong & Moran, 2013)	New Haven, USA
Gilliamella apis ESL0169	4	BHIA, 35°C, microaerophilic	(Kešnerová et al., 2017)	Lausanne, Switzerland
Gilliamella sp. ESL0182	4	BHIA, 35°C, microaerophilic	(Ellegaard & Engel, 2018)	Lausanne, Switzerland
Gilliamella apis ESL0297	4	BHIA, 35°C, microaerophilic	This study	Lausanne, Switzerland
Snodgrassella alvi wkB2^T	4	TSA, 35°C, microaerophilic	(Kwong & Moran, 2013)	New Haven, USA
Snodgrassella alvi ESL0304	4	TSA, 35°C, microaerophilic	This study	Lausanne, Switzerland
Snodgrassella alvi ESL0323	4	TSA, 35°C, microaerophilic	This study	Lausanne, Switzerland
Snodgrassella alvi ESL0251	4	TSA, 35°C, microaerophilic	This study	Lausanne, Switzerland
Snodgrassella alvi ESL0897	4	TSA, 35°C, microaerophilic	This study	Lausanne, Switzerland

760

761

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

762 **Table S2. *S. alvi* & *Gilliamella* colonization summary counts.** Total bee numbers per experiment,
763 survival rate, number of bees used for CFU plating, qPCR and metabolomics analysis, and pollen
764 consumption are shown.

	Replicate	1	2	3	4	5	Total
Bees	MD	34	24	24	24	0	106
	MD (P)	30	23	23	25	0	101
	Sn	31	24	20	0	20	95
	Sn (P)	34	23	20	0	19	96
	Sn+Gi	32	22	24	0	16	94
	Sn+Gi (P)	29	24	20	0	22	95
	Gi	0	0	0	26	25	51
	Gi (P)	0	0	0	25	25	50
Survival (%)	MD	100%	100%	100%	96%	N/A	99%
	MD (P)	100%	100%	100%	100%	N/A	100%
	Sn	94%	100%	100%	N/A	100%	98%
	Sn (P)	100%	100%	95%	N/A	100%	99%
	Sn+Gi	97%	95%	100%	N/A	100%	98%
	Sn+Gi (P)	97%	92%	95%	N/A	100%	96%
	Gi	N/A	N/A	N/A	100%	100%	100%
	Gi (P)	N/A	N/A	N/A	96%	100%	98%
CFU	MD	10	10	10	10	N/A	40
	MD (P)	10	10	10	10	N/A	40
	Sn	10	10	10	N/A	6	36
	Sn (P)	10	10	10	N/A	6	36
	Sn+Gi	10	10	10	N/A	6	36
	Sn+Gi (P)	10	10	10	N/A	6	36
	Gi	N/A	N/A	N/A	10	18	28
	Gi (P)	N/A	N/A	N/A	10	18	28
qPCR	MD	4	3	4	6	N/A	17
	MD (P)	4	3	4	5	N/A	16
	Sn	4	9	6	N/A	6	25
	Sn (P)	4	9	8	N/A	6	27
	Sn/Gi	10	10	11	N/A	6	37
	Sn/Gi (P)	10	10	10	N/A	6	36
	Gi	N/A	N/A	N/A	10	14	24
	Gi (P)	N/A	N/A	N/A	9	21	30
Metabolomics	MD	2	1	2	5	N/A	10
	MD (P)	2	1	4	5	N/A	12
	Sn	1	1	2	N/A	6	10
	Sn (P)	1	1	3	N/A	6	11
	Sn+Gi	3	1	5	N/A	6	15
	Sn+Gi (P)	1	1	4	N/A	5	11
	Gi	N/A	N/A	N/A	5	9	14
	Gi (P)	N/A	N/A	N/A	6	7	13
Pollen Consumption (mg/bee)	MD	-	-	-	-	-	-
	MD (P)	30	33	32	13	N/A	26
	Sn	-	-	-	-	-	-
	Sn (P)	20	31	30	N/A	22	26
	Sn+Gi	-	-	-	-	-	-
	Sn+Gi (P)	19	30	33	N/A	16	24
	Gi	-	-	-	-	-	-
	Gi (P)	N/A	N/A	N/A	17	18	17

765

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

766 **Table S3. Summary of the number of bees sampled per replicate and treatment for the mono-**
767 **colonization experiment with divergent *Snodgrassella* strains.** Total number of bees per experiment,
768 survival rate, and number of bees used for CFU plating and metabolomics analysis are shown.

	Replicate	1	2	Total
Bees	MD	14	15	29
	wkB2	14	14	28
	251	16	15	31
	304	14	15	29
	323	14	15	29
	892	14	15	29
	897	15	14	29
Survival (%)	MD	93%	100%	96%
	wkB2	100%	100%	100%
	251	88%	100%	94%
	304	100%	80%	90%
	323	100%	100%	100%
	892	86%	100%	93%
	897	80%	79%	79%
CFU	MD	12	12	24
	wkB2	12	12	24
	251	11	12	23
	304	12	9	21
	323	12	12	24
	892	12	12	24
	897	12	9	21
Metabolomics	MD	6	6	12
	wkB2	6	6	12
	251	8	9	17
	304	6	6	12
	323	7	6	13
	892	6	6	12
	897	6	6	12

769

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

References

1 Tremaroli, V. & Bäckhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **489**, 242-249, doi:10.1038/nature11552 (2012).

2 Kwong, W. K. & Moran, N. A. Gut microbial communities of social bees. *Nature Reviews Microbiology* **14**, 374-384, doi:10.1038/nrmicro.2016.43 (2016).

3 Bonilla-Rosso, G. & Engel, P. Functional roles and metabolic niches in the honey bee gut microbiota. *Current Opinion in Microbiology* **43**, 69-76, doi:10.1016/j.mib.2017.12.009 (2018).

4 Kwong, W. K. *et al.* Dynamic microbiome evolution in social bees. *Science Advances* **3**, e1600513, doi:10.1126/sciadv.1600513 (2017).

5 Li, Y., Leonard, S. P., Powell, J. E. & Moran, N. A. Species divergence in gut-restricted bacteria of social bees. *Proceedings of the National Academy of Sciences* **119**, e2115013119, doi:10.1073/pnas.2115013119 (2022).

6 Wu, J. *et al.* Honey bee genetics shape the strain-level structure of gut microbiota in social transmission. *Microbiome* **9**, 225, doi:10.1186/s40168-021-01174-y (2021).

7 Engel, P. *et al.* Standard methods for research on Apis mellifera gut symbionts. *Journal of Apicultural Research* **52**, 1-24, doi:10.3896/ibra.1.52.4.07 (2013).

8 Ricigliano, V. A. & Anderson, K. E. Probing the Honey Bee Diet-Microbiota-Host Axis Using Pollen Restriction and Organic Acid Feeding. *Insects* **11**, 291, doi:10.3390/insects11050291 (2020).

9 Zheng, H., Steele, M. I., Leonard, S. P., Motta, E. V. S. & Moran, N. A. Honey bees as models for gut microbiota research. *Lab Animal* **47**, 317-325, doi:10.1038/s41684-018-0173-x (2018).

10 Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proceedings of the National Academy of Sciences* **111**, 11509-11514, doi:doi:10.1073/pnas.1405838111 (2014).

11 Martinson, V. G., Moy, J. & Moran, N. A. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl Environ Microbiol* **78**, 2830-2840, doi:10.1128/aem.07810-11 (2012).

12 Kešnerová, L. *et al.* Disentangling metabolic functions of bacteria in the honey bee gut. *PLOS Biology* **15**, e2003467, doi:10.1371/journal.pbio.2003467 (2017).

13 Schwartzman, J. A. *et al.* The chemistry of negotiation: Rhythmic, glycan-driven acidification in a symbiotic conversation. *Proceedings of the National Academy of Sciences* **112**, 566-571, doi:10.1073/pnas.1418580112 (2015).

14 Desai, M. S. *et al.* A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell* **167**, 1339-1353.e1321, doi:https://doi.org/10.1016/j.cell.2016.10.043 (2016).

15 Sonnenburg, J. L. *et al.* Glycan Foraging in Vivo by an Intestine-Adapted Bacterial Symbiont. *Science* **307**, 1955-1959, doi:10.1126/science.1109051 (2005).

16 Berry, D. *et al.* Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proceedings of the National Academy of Sciences* **110**, 4720-4725, doi:10.1073/pnas.1219247110 (2013).

17 Sicard, J.-F., Le Bihan, G., Vogeleer, P., Jacques, M. & Harel, J. Interactions of Intestinal Bacteria with Components of the Intestinal Mucus. *Frontiers in Cellular and Infection Microbiology* **7**, doi:10.3389/fcimb.2017.00387 (2017).

18 Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont. *Cell Host & Microbe* **4**, 447-457, doi:10.1016/j.chom.2008.09.007 (2008).

19 Miyashiro, T. *et al.* The N-acetyl-d-glucosamine repressor NagC of Vibrio fischeri facilitates colonization of Euprymna scolopes. *Molecular Microbiology* **82**, 894-903, doi:https://doi.org/10.1111/j.1365-2958.2011.07858.x (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

20 Wier, A. M. *et al.* Transcriptional patterns in both host and bacterium underlie a daily rhythm of anatomical and metabolic change in a beneficial symbiosis. *Proceedings of the National Academy of Sciences* **107**, 2259-2264, doi:10.1073/pnas.0909712107 (2010).

21 Zeng, X. *et al.* Gut bacterial nutrient preferences quantified *in vivo*. *Cell* **185**, 3441-3456.e3419, doi:10.1016/j.cell.2022.07.020 (2022).

22 Gillis, C. C. *et al.* Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. *Cell Host & Microbe* **23**, 54-64.e56, doi:https://doi.org/10.1016/j.chom.2017.11.006 (2018).

23 Rosenberg, G. *et al.* Host succinate is an activation signal for *Salmonella* virulence during intracellular infection. *Science* **371**, 400-405, doi:10.1126/science.aba8026 (2021).

24 Nguyen, B. D. *et al.* Import of Aspartate and Malate by DcuABC Drives H2/Fumarate Respiration to Promote Initial Salmonella Gut-Lumen Colonization in Mice. *Cell Host & Microbe* **27**, 922-936.e926, doi:10.1016/j.chom.2020.04.013 (2020).

25 Zheng, H., Powell, J. E., Steele, M. I., Dietrich, C. & Moran, N. A. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proceedings of the National Academy of Sciences* **114**, 4775-4780, doi:10.1073/pnas.1701819114 (2017).

26 Brochet, S. *et al.* Niche partitioning facilitates coexistence of closely related honey bee gut bacteria. *eLife* **10**, doi:10.7554/elife.68583 (2021).

27 Powell, J. E., Leonard, S. P., Kwong, W. K., Engel, P. & Moran, N. A. Genome-wide screen identifies host colonization determinants in a bacterial gut symbiont. *Proceedings of the National Academy of Sciences* **113**, 13887-13892, doi:10.1073/pnas.1610856113 (2016).

28 Kwong, W. K., Zheng, H. & Moran, N. A. Convergent evolution of a modified, acetate-driven TCA cycle in bacteria. *Nature Microbiology* **2**, 17067, doi:10.1038/nmicrobiol.2017.67 (2017).

29 Sharp, C. & Foster, K. R. Host control and the evolution of cooperation in host microbiomes. *Nature Communications* **13**, 3567, doi:10.1038/s41467-022-30971-8 (2022).

30 Steele, M. I., Motta, E. V. S., Gattu, T., Martinez, D. & Moran, N. A. The Gut Microbiota Protects Bees from Invasion by a Bacterial Pathogen. *Microbiology Spectrum* **9**, e00394-00321, doi:10.1128/Spectrum.00394-21 (2021).

31 Feng, Y. *et al.* Anopheline mosquitoes are protected against parasite infection by tryptophan catabolism in gut microbiota. *Nature Microbiology* **7**, 707-715, doi:10.1038/s41564-022-01099-8 (2022).

32 Cerstiaens, A. *et al.* Neurotoxic and neurobehavioral effects of kynurenines in adult insects. *Biochemical and Biophysical Research Communications* **312**, 1171-1177, doi:https://doi.org/10.1016/j.bbrc.2003.11.051 (2003).

33 Marszalek-Grabska, M. *et al.* Kynurenine emerges from the shadows – Current knowledge on its fate and function. *Pharmacology & Therapeutics* **225**, 107845, doi:https://doi.org/10.1016/j.pharmthera.2021.107845 (2021).

34 Schwarcz, R. & Stone, T. W. The kynurenine pathway and the brain: Challenges, controversies and promises. *Neuropharmacology* **112**, 237-247, doi:https://doi.org/10.1016/j.neuropharm.2016.08.003 (2017).

35 Scheiner, R., Baumann, A. & Blenau, W. Aminergic control and modulation of honeybee behaviour. *Curr Neuropharmacol* **4**, 259-276, doi:10.2174/157015906778520791 (2006).

36 Zhang, Z. *et al.* Honeybee gut Lactobacillus modulates host learning and memory behaviors via regulating tryptophan metabolism. *Nature Communications* **13**, 2037, doi:10.1038/s41467-022-29760-0 (2022).

37 Hu, Y. *et al.* Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. *Nature Communications* **9**, 964, doi:10.1038/s41467-018-03357-y (2018).

38 Kešnerová, L. *et al.* Gut microbiota structure differs between honeybees in winter and summer. *The ISME Journal* **14**, 801-814, doi:10.1038/s41396-019-0568-8 (2020).

39 Nocelli, R. C., Cintra-Socolowski, P., Roat, T. C., Silva-Zacarin, E. C. M. & Malaspina, O.

40 Yagi, S. & Ogawa, H. Effect of Tryptophan Metabolites on Fluorescent Granules in the Malpighian Tubules of Eye Color Mutants of *Drosophila melanogaster*. *Zoological Science* **13**, 97-104, 108 (1996).

bioRxiv preprint doi: https://doi.org/10.1101/2023.01.23.524906; this version posted August 20, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

41 Emery, O., Schmidt, K. & Engel, P. Immune system stimulation by the gut symbiont *Frischella perrara* in the honey bee (*Apis mellifera*). *Molecular Ecology* **26**, 2576-2590, doi:https://doi.org/10.1111/mec.14058 (2017).

42 Tauber, J. P. *et al.* Honeybee intestines retain low yeast titers, but no bacterial mutualists, at emergence. *Yeast* **39**, 95-107, doi:https://doi.org/10.1002/yea.3665 (2022).

43 den Besten, G. *et al.* The Short-Chain Fatty Acid Uptake Fluxes by Mice on a Guar Gum Supplemented Diet Associate with Amelioration of Major Biomarkers of the Metabolic Syndrome. *PLOS ONE* **9**, e107392, doi:10.1371/journal.pone.0107392 (2014).

44 Fiehn, O. *et al.* Quality control for plant metabolomics: reporting MSI-compliant studies. *The Plant Journal* **53**, 691-704, doi:10.1111/j.1365-313x.2007.03387.x (2008).

45 Wanichthanarak, K., Jeamsripong, S., Pornputtapong, N. & Khoomrung, S. Accounting for biological variation with linear mixed-effects modelling improves the quality of clinical metabolomics data. *Computational and Structural Biotechnology Journal* **17**, 611-618, doi:https://doi.org/10.1016/j.csbj.2019.04.009 (2019).

46 Nargund, S., Misra, A., Zhang, X., Coleman, G. D. & Sriram, G. Flux and reflux: metabolite reflux in plant suspension cells and its implications for isotope-assisted metabolic flux analysis. *Molecular BioSystems* **10**, 1496-1508, doi:10.1039/C3MB70348G (2014).

47 Rosenblatt, J., Chinkes, D., Wolfe, M. & Wolfe, R. R. Stable isotope tracer analysis by GC-MS, including quantification of isotopomer effects. *American Journal of Physiology-Endocrinology and Metabolism* **263**, E584-E596, doi:10.1152/ajpendo.1992.263.3.E584 (1992).

48 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology* **116**, 71-76, doi:https://doi.org/10.1006/jsbi.1996.0013 (1996).

49 Hoppe, P., Cohen, S. & Meibom, A. NanoSIMS: Technical Aspects and Applications in Cosmochemistry and Biological Geochemistry. *Geostandards and Geoanalytical Research* **37**, 111-154, doi:https://doi.org/10.1111/j.1751-908X.2013.00239.x (2013).

50 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

51 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).

52 Moran, N. A., Hansen, A. K., Powell, J. E. & Sabree, Z. L. Distinctive Gut Microbiota of Honey Bees Assessed Using Deep Sampling from Individual Worker Bees. *PLOS ONE* **7**, e36393, doi:10.1371/journal.pone.0036393 (2012).

53 Kwong, W. K. & Moran, N. A. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the. *International Journal of Systematic and Evolutionary Microbiology* **63**, 2008-2018, doi:10.1099/ijs.0.044875-0 (2013).

54 Ellegaard, K. M. & Engel, P. New Reference Genome Sequences for 17 Bacterial Strains of the Honey Bee Gut Microbiota. *Microbiol Resour Announc* **7**, doi:10.1128/mra.00834-18 (2018).