Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food science

Philosophiae Doctor (PhD)
Thesis 2018:26

Spread and persistence of antibiotic resistance genes in the honeybee gut microbiota

Spredning og persistens av antibiotika resistens gener i tarmfloraen hos honningbier

Jane Ludvigsen

Spread and persistence of antibiotic resistance genes in the honeybee gut microbiota

Spredning og persistens av antibiotika resistens gener i tarmfloraen hos honningbier

Philosophiae Doctor (PhD) Thesis

Jane Ludvigsen

Norwegian University of Life Sciences
Faculty of Chemistry, Biotechnology and Food science

Ås, 2018

Thesis number 2018:26
ISSN 1894-6402
ISBN: 978-82-575-1504-1

Acknowledgements

The work in this thesis were carried out at the Faculty of Chemistry, Biotechnology and Food Science at the Norwegian University of Life Sciences, with financial support by the University.

First of all, I would like to thank my main supervisor Prof. Knut Rudi for taking on this PhD-journey with me; diving head first into the exciting world of honeybee gut microbiota. It has been an adventure with excellent supervision; pulling me up when I had my head stuck in the rectum; pushing me forward with just the right freedom to reach the ileum; guiding me, with tasks just difficult enough, all the way to the midgut; and even though I was crawling around almost drowning in nectar, you encouraged me to pull true and make it out. Thank you for these amazing years in the MiDiv lab! I have learned a lot ☺

I would also like to thank my co-supervisors Trine M. L'Abée-Lund, and Gro V. Amdam for all help along the way. Gro, you reminded and inspired me to believe in myself ☺.

My sincere gratitude goes to you - MiDiv girls - for your loving kindness, and to Davide, Claus, Daniel, Heidi and all the current and former group members of the MiDiv Lab, Environmental microbiology group and everyone I know at KMB: you have made my every work day! – NICE. Thank you for always being there: to share experiences, to come to for help, to enjoy work, to practice patience acceptance on, and for motivating me.
I am forever grateful ☺

A want to give a special note to my spiritual guide Gen Kelsang Tubchen and my Kadampa friends – you evoke peace and happiness in my mind and make each day easier with your profound advice on how to solve my PhD problems.

My dear friends and family – thank you for your endless love and for always believing in me – I love you.
Now - Let's party!

Jane

Table of Contents

Summary

Bacteria residing in the gut – the gut microbiota (GM) – are important for host health. Several parallels between the human and honeybee GM exist: i) the GM is host specific (core microbiota), ii) age dependent development of the GM microbiota, iii) the GM composition is gut part dependent, and iv) the GMs ability to provide the host with additional dietary benefits. Due to their simple GM composition, honeybees have emerged as a model to understand host-bacteria interactions. In addition, the honeybee GM has been used to study its associations with perturbations like antibiotic exposure. Although short-term perturbations by antibiotic treatment have been extensively studied, we have limited knowledge about the long-term effects. The human and animal GM is a reservoir for different antibiotic resistant genes (ARGs) that can transfer to pathogenic bacteria, a scenario that is already a global and serious threat to human infection management. In this thesis, we addressed several intriguing questions regarding this scenario: e.g. in what way do long-term antibiotic treatment affect the GM composition, how do ARG spread to and within the GM, and how are they persistent within the GM?

We used honeybees from different antibiotic treatment regimes as a model to address if antibiotic exposure will select for antibiotic resistant bacterial strains and/or if ARGs will transfer horizontally within the core microbiota. We used an experimental set up of two honeybee populations: one from Arizona, USA and one from Ås, Norway. In the USA, tetracycline is widely used in agriculture as well as to treat honeybee infections, while in Norway it is not. This set up in combination with the use of a low complex model system, allowed us to identify spread of ARG within the GM population at the bacterial strain level, and associate it with antibiotic exposure. We used a combination of techniques to investigate the honeybee GM composition and the prevalence of ARG: e.g. Bacteria culturing, quantitative PCR, Illumina whole genome shotgun sequencing, phenotypical testing and microscopy.

We focused on two bacteria important for honeybee health: *Gilliamella apicola* and *Snodgrassella alvi*. To investigate the phylogeny composition in our dataset, we compared genes found in all bacteria (of the same species) and found that strains of *G. apicola* separated into three subgroups found in bees from both Norway and Arizona. This showed that strain diversity is maintained despite long-term antibiotic exposure to the Arizonan bee

population. We also found that antibiotic exposure has an effect on the horizontal spread of transposon associated ARG within the Arizonan honeybee GM, wherein these ARGs were detected in all subgroups of *G. apicola* as well as its transfer to *S. alvi*.

Moreover, our results showed that unique tetracycline resistance genes associated differently with unique bacterial subgroups. One subgroup differed substantially both phenotypically and genotypically from the type strain of *G. apicola* and therefore it was characterized, described and proposed as a new species: *G. apis* sp. nov.

Overall, these findings show that ARG are prevalent in the core microbiota of honeybees and that long-term antibiotic exposure influences the spread of ARG within the honeybee core microbiota population rather than selecting for a few antibiotic resistant strains. This suggests that persistence of ARGs in the GM is sustained by host selection of core bacteria harboring ARGs, and that antibiotic exposure maintains the GM as a potent reservoir for ARGs. These results highlight the need to reduce unnecessary antibiotic usage to prevent spread of ARGs and demonstrate the suitability of honeybees as a model for investigating ARGs spread in bacterial populations.

Sammendrag

Bakteriene som lever i tarmen – tarmfloraen – er viktig for vertens helse. Det kan dras mange paralleller mellom tarmfloraen til mennesker og honningbier: i) tarmfloraen er verts spesifikk (kjerneflora), ii) begge utvikles med alder, iii) sammensetningen av tarmfloraen er avhengig hvor i tarmen det er, og iv) tarmfloraen kan tilføre verten energi ved nedbryting av næringsstoffer. På grunn av at honningbier har en enkelt sammensatt tarmflora, har de begynt å bli brukt som modell i studier om bakterie-vert interaksjoner. I tillegg har honningbier blitt brukt i studier hvor det er sett på hvordan tarmfloraen er assosiert med ytre påvirkninger. Selv om forandringer som skyldes kortvarige antibiotika behandlinger har blitt grundig studert, har vi heller liten kunnskap om langtids effekter. Tarmfloraen i mennesker og dyr er et reservoar for antibiotika resistens gener (ARG) som kan overføres til patogene bakterier, et senario som allerede er en global og alvorlig trussel for behandling av infeksjonssykdommer. I dette doktorgradsarbeidet, adresserte vi flere spennende spørsmål relatert til dette senarioet, som f.eks.: på hvilken måte påvirker lang-tids eksponering med antibiotika tarmfloraen sammensetning, hvordan spes ARG til og i tarmfloraen, og på hvilken måte kan ARG persistere i tarmfloraen?

Vi brukte honningbier fra områder med ulik bruk av antibiotika som en modell for å adressere om antibiotika eksponering vil kunne selektere for antibiotika resistente bakterier og/ eller om ARG vil kunne overføres mellom bakteriemedlemmene i kjernefloraen. Vi brukte et eksperimentelt oppsett med to honningbie populasjoner: en fra Arizona, USA og en fra Ås, Norge. I USA blir tetrasyklin bruk i landbruksindustrien og likeså som til behandling av infiserte bikuber, mens i Norge blir tetrasyklin ikke brukt slik. Dette oppsettet i kombinasjon med en modell som har en enkel tarmflorasammensetning, gjorde slik at vi kunne identifisere spredning av ARG innad in tarmfloraen på bakteriestamme nivå, og assosiere dette med antibiotika eksponering. Vi brukte flere ulike metoder for å undersøke honningbienes tarmflora sammensetning og prevalens av ARG der i: dyrkning av bakterier, kvantitativ PCR, Illumina hel-genom sekvensering, phenotypiske tester og mikroskopering.

Vi fokuserte på to bakterier som er viktige for honningbie helse: *Gilliamella apicola* og *Snodgrassella alvi*. For å kunne undersøke den fylogenetiske sammensetningen i vårt datasett, så sammenlignet vi genene som finnes i alle bakteriene (innenfor en bakterie spesies), og fant at ulike stammer av *G. apicola* grupperte seg i tre sub-grupper, og disse var

tilstede i bier både fra Norge og Arizona. Dette viste at mangfoldet av stammer beholdes selv under langvarig antibiotika eksponering. Vi fant også at antibiotika eksponering har en effekt på overføring av transposon-assosierte ARG i tarmfloraen hos honningbier fra Arizona, hvor i disse ARG kunne detekteres i alle sub-grupper av *G. apicola* og også i *S. alvi*.

I tillegg viser våre resultater at ulike tetrasyklinresistens gener assosierer seg ulike med ulike bakterie sub-grupper. En av disse sub-gruppene var så ulik både phenotypisk og genotypisk type-stammen *G. apicola,* at den derfor ble karakterisert, beskrevet og foreslått til å være en ny spesies: *G. apis* sp. nov.

Sett i sammenheng så viser disse funnene at ARG er prevalente i kjernefloraen hos honningbier og at langtids eksponering med antibiotika påvirker i større grad spredningen av ARG i tarmfloraen enn at den selekterer for noe få antibiotika resistente stammer. Fra dette kan det tenkes at persistens av ARG i tarmfloraen opprettholdes på grunn av verts seleksjon av kjerne floraen, og at antibiotika eksponering understøtter at tarmfloraen forblir et reservoar for ARG. Disse resultatene påpeker at det er viktig å redusere unødvendig bruk av antibiotika for å forebygge spredning av ARG og de demonstrerer at honningbier er nyttige som modell til å undersøke hvordan ARG sprer seg i bakteriepopulasjoner.

List of Papers

Paper 1

Jane Ludvigsen, Anbjørg Rangberg, Ekaterina Avershina, Monika Sekelja, Claus Kreibich, Gro Amdam, and Knut Rudi. **Shifts in the Midgut/Pyloric Microbiota Composition within a Honey Bee Apiary throughout a Season.** *Microbes and Environments* (2015), Vol. 30 No. 3 p. 235-244, doi:10.1264/jsme2.ME15019

Paper 2

Jane Ludvigsen, Davide Porcellato, Trine M. L'Abée-Lund, Gro V. Amdam, Knut Rudi. **Geographically widespread honeybee gut symbiont subgroups show locally distinct antibiotic resistance pattern.** *Molecular Ecology* (2017), Vol. 26 (23):6590-6607, doi:10.1111/mec.14392

Paper 3

Jane Ludvigsen, Davide Porcellato, Knut Rudi. **Resolving the diversity of the honeybee gut symbiont *Gilliamella*: description of *Gilliamella apis* sp. nov., isolated from the gut of honeybees (*Apis mellifera*).** *In revision; IJSEM*

Paper 4

Jane Ludvigsen, Gro V. Amdam, Knut Rudi, Trine M. L'Abée-Lund. **Linking streptomycin resistance genes (*strA-strB*) in a honeybee gut symbiont to environmental antibiotic exposure.** *Submitted; NOTE in Microbial Ecology*

Additional papers

Herstad, Kristin; Gajardo, Karina; Bakke, Anne Marie; Moe, Lars; **Ludvigsen, Jane**; Rudi, Knut; Rud, Ida; Sekelja, Monika; Skancke, Ellen (2017). A diet change from dry food to beef induces reversible changes on the faecal microbiota in healthy, adult, client-owned dogs. *BMC Veterinary Research*; Vol. 13(1).

Rudi, Knut; **Ludvigsen, Jane**; Dirven, Hubert; Steffensen, Inger-Lise (2017). Genetically and dietary induced obesity associate differently with gut microbiota in a murine intestinal tumorigenesis model. *Environmental Disease*; Vol. 2. p 45-54

Ludvigsen, Jane; Svihus, Birger; Rudi, Knut (2016). Rearing Room Affects the Non-dominant Chicken Cecum Microbiota, While Diet Affects the Dominant Microbiota. *Frontiers in Veterinary Science*; Vol. 3.

Ravi, Anuradha; Avershina, Ekaterina; Foley, Steven L.; **Ludvigsen, Jane**; Storrø, Ola; Øien, Torbjørn; Johnsen, Roar; McCartney, Anne; L'Abee-Lund, Trine Marie; Rudi, Knut (2015). The commensal infant gut meta-mobilome as a potential reservoir for persistent multidrug resistance integrons. *Scientific Reports*; Vol. 5:15317.

Ravi, Anuradha; Avershina, Ekaterina; Rudi, Knut; **Ludvigsen, Jane**; L'Abée-Lund, Trine (2014). Integrons in the Intestinal Microbiota as Reservoirs for Transmission of Antibiotic Resistance Genes. *Pathogens*; Vol. 3. p 238-248

Baruzzi F, **Ludvigsen J**, Rudi K, de Candia S (2017). Shift in raw bovine milk microbiota after spontaneous fermentation as revealed by 16S rDNA metagenomic analysis. *Draft manuscript*.

Abbreviations

ANI	Average nucleotide identity
AR	Antibiotic resistance
ARG(s)	Antibiotic resistance gene(s)
DDH	DNA-DNA hybridization
GGD	Genome to genome distance
GI-tract	Gastro intestinal tract
GM	Gut microbiota
HGT	Horizontal gene transfer
IS-element	Insertion element
MGEs	Mobile genetic elements
MIC	Minimum inhibitory concentration
ML	Maximum likelihood
nt	Nucleotide
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
SNP	single nucleotide polymorphism
Sm^r	Streptomycin resistance
SV	Sequence variants
Tc^r	Tetracycline resistance
WGS	Whole genome sequencing
WT	Wild type

~ Everything starts with a wish ~

1 Introduction

1.1 Gut microbiota

The gastro intestinal tract (GI-tract) is a stretched tube spanning from mouth to anus that is located in the center of the body (Figure 1). The gut spans form the stomach to the anus and is an organ like the heart or the brain, but with a distinction: it contains a vast number of microorganisms. These are bacteria, eukaryotes, viruses, and archaea [1], which are collectively termed the gut microbiota (GM). Bacteria make up the predominant part of the GM, but which bacteria that are present in which proportions (microbiota composition) varies among different animal species [2]. The bacteria living in the gut have collectively more genes than its host does, and thus they can perform a vast variety of functions [3, 4]. Therefore, the bacterial gut microbiota can provide its host beneficial capacities [1, 4, 5]. One main benefit is that they help to break down food substances that the host cannot, which contributes to the hosts' energy and nutrient uptake [6]. Diet is a main driver for the gut microbiota composition [7], because different bacteria can utilize and break down different substances depending on their functional gene repertoire [3]. A clear distinction between the GM composition, its complexity and host species develops due to gut structure and dietary preferences [4]. In general, a more complex GM reflect a more diverse diet [4, 8]. However, this complexity varies both, in bacterial numbers and in the microbiota composition, in dependence with the location in the gut [2, 6, 9]. Generally, there are lower numbers of bacteria closer to the mouth with increasing numbers towards the anus (Figure 1).

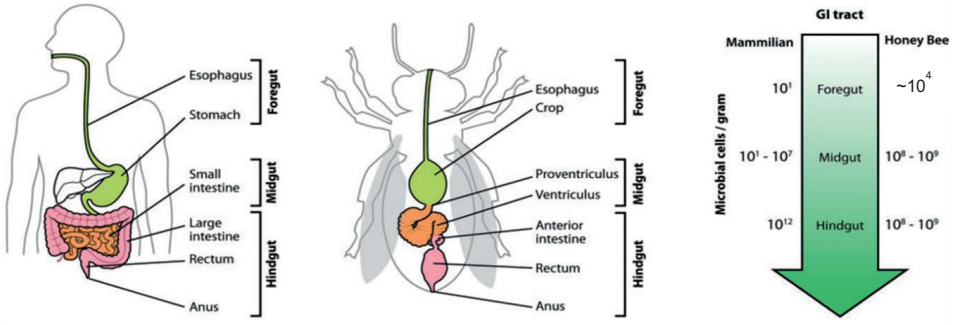

Figure 1: The digestive systems of humans and honeybees. The coloring is comparable in respect to different parts of the GI-tract system. The right side figure compare bacteria/cell numbers of different gut parts between humans and honeybees. (Adapted from [10]).

In recent years, connection are made between the GM composition and gastro intestinal diseases [11], how it impacts on host immune system [12, 13], its correlation with different host metabolic diseases [14-16], and several physical/mental conditions of the host [17, 18]. It is an intricate web of bacterial – host interactions [19], and although the knowledge is increasing about the importance of the GM and how it affects host health [20], large parts of this puzzle are still missing [21].

1.2 The honeybee gut microbiota

1.2.1 Complexity

The honeybee GM displays no exceptions from the above characteristics, and partly because of their strictly plant based and sugar rich diet, their GM complexity is low. The honeybee GI-tract can be divided into four main parts with distinct functions; crop (nectar and water is stored temporarily); midgut (digestion of food and nutrient absorption); ileum (absorption of digested food and bacterial breakdown of undigested food passed on from the midgut); rectum (feces accumulation and water absorption) [22, 23]. These parts are roughly similar to those found in humans (Figure 1). The GM complexity within these four parts differs and is reflected largely in the numbers of bacteria present: Crop (~10^4), midgut (~10^6-10^7), ileum (~10^7-10^8), and rectum (10^8-10^9) [9, 24]. These numbers are based on detected 16S rRNA gene copies per gut part, and taken together, bacteria in the crop amount for ~ 0.007 to 1%, the midgut ~1 to 4%, the ileum ~ 4 to 10%, and the rectum ~ 90-95%, of the total bacteria found in the honeybee GI-tract [24]. When compared to humans wherein the ileum harbor up to 10^7 cells/ml and the colon/rectum harbor 10^{11} - 10^{12} cells per gram [6, 25, 26], the lower complexity of the GM of honeybees is apparent (Figure 1).

1.2.2 Seasonality

The European honeybee (*Apis mellifera*) is a social insect that are used for large scale pollination in the food industry, and preserving healthy honeybees is essential for efficient crop yield [27]. They are very flexible in pollination preferences, and are managed by beekeepers in most parts of the world, thus subjected to different living conditions [28]. In summer the bees forage on different plants collecting pollen, nectar and water, while during winter in the temperate zone the bees overwinter inside the hives on a sugar diet fed them by

the beekeepers. How this change in living conditions and diet affects the honeybee GM is not known and needs to be elucidated.

The bees that emerge in spring have a much shorter lifespan (5-6 weeks) compared to the bees that emerge in the fall and will survive the winter (6 months) [22], and thus the social structure of the honeybee hives is temperature dependent [29]. The seasonal shifts impacts the honeybee life cycle that normally starts with the queen laying eggs in the early spring that hatches into larvae, and then develop into young female worker bees [22]. As these bees age, they undertake different roles within the hives, which are not strict [30-32], but mostly follow a certain chronological pattern: Nurse bees (cleans cells and feed the larvae), in-hive bees (receives and store food), guard bees (guards the entrance), and foragers (collect and deliver pollen and nectar for the hive population) [22].

1.2.3 Development

The development of the honeybee gut microbiota is a dynamic process that resembles that of a human child, which predominantly starts at birth and bacteria from the child's environment starts colonizing the gut [33, 34]. When the honeybee larvae emerges as a young bee, the inner skeleton of the larvae is broken down and the bee GI-tract is formed [22]. As a newly emerged bee the GI-tract is almost sterile and is colonized by both environmental and bee specific bacteria within the first one to two days (ileum and rectum, respectively) [9, 35, 36]. This colonization process happens due to contact with hive material and nest mates, and after two days the bacterial numbers are still low (~10^4-10^5 in both ileum and rectum) [9]. Then after three days, the bacteria become more numerous, as well as gut part and host specific, and stabilizes at day six and remains more or less stable while the bee is aging [9]. The same pattern can be described for the human child, when in the first months, the GM composition is highly variable, and then, at three years of age, the child microbiota becomes comparable to an adult's, both in composition and in number of bacteria [34, 37, 38]. The colonization of bee specific bacteria in the ileum is dependent on direct contact with nest mates, their fecal material, and food sources [9], which are also factors driving the human colonization process (mode of delivery, contact with its mother and other humans, its surroundings, and food preferences) [39-41]. In addition to diet being a main driver for shifts in the GM composition, factors like host selection (some bacteria are

favored by the host) and social behavior, are likely to be as significant, which ultimately results in a highly adapted host specific GM [2, 8, 23, 42-44].

1.2.4 Host specific composition

Comparisons of the GM composition of *Apis mellifera* from diverse geographical regions show that it is highly consistent with nine bacterial genera clusters found in almost all bees [44]. These nine clusters make up more than 95% of the total bacteria in the GI-tract [45, 46]. These bacterial clusters were designated as bacterial phylotypes based on a >97% identity cutoff of the 16SrRNA gene, and named: Gamma-1; Gamma-2 and Beta (Gamma- and Beta- *Proteobacteria*, respectively); Firm-4 and Firm-5 (*Firmucutes*); Bifido (*Actinobacteria*); Alpha-1, Alpha 2.1 and Alpha 2.2 (Alpha-*proteobacteria*) [47]. Bacteria from Gamma-1, Beta, Firm-4, Firm-5 and Bifido are present in almost all individual honeybees (*Apis*) and except for Firm-4, in most social bumblebees (*Bombus*) species [42, 47]. Collectively they are known as the honeybee GI-tract core microbiota [23]. These five phylotypes dominate the microbiota [24] and the presence of the remaining four among individual bees are more erratic [45], with Alpha-1 and Gamma-2 being *Apis* specific [42]. The described GM composition is associated with a healthy host state and changes may lead to disease [27].

In recent years, the nine phylotypes have been taxonomically determined and subsequently renamed after characterization of cultured bacteria. The bacterial species that today represent the different phylotypes are listed in table 1, of which the Alpha-2.1 phylotype is not yet eluded [23].

In general, the crop consists of mostly *Lactobacillus* sp., *P. apium* and other environmental bacteria, and the number of bacteria might fluctuate as the bee age and is exposed to different diets [48]. Contrary, the hindgut is found to be more or less stable, regarding who is there, after it has established (at day seven) [49], which is comparable to what is observed for humans [50]. The midgut contains low numbers *S. alvi* and *G. apicola*, as do ileum with *S. alvi* growing along the length of the gut wall with *G. apicola* as a layer on top of *S. alvi* extending into the lumen [24]. In younger bees, *F. perrara* is found to colonize the ileum at a greater extend [9]. The ileum also contains *Lactobacillus* sp. from both Firm-4 and Firm-5, but these are more abundant in the rectum where they dominate together with

B. asteroides [23]. Overall, the honeybee GM mainly consists of the three phyla *Proteobacteria*, *Firmicutes*, and *Actinobacteria* (Table 1). In comparison, the human gut largely contains bacteria from the phylum *Firmicutes* and *Bacteriodetes*, with sparse communities of *Proteobacteria*, *Actinobacteria* and *Verrucomicrobia* and the majority are believed to be anaerobic bacteria [51, 52]. Due to the high complexity in species diversity, and the highly adaptable GM composition, in response to different lifestyles, diets, and living environments [37, 53], only two bacterial groups are found to exist across the human population (core species) [54].

Even though the first culturing of bacteria from the GI-tract of bees was initiated by Martha Gilliam in the 1970s [55], the more comprehensive study of the bee GM is a relatively new field of research that started slightly more than a decade ago [56]. The identification of the GM bacteria has mostly used the 16SrRNA gene similarity and a 97% species identity cut off for species designation. Using this approach, sequences identified as one bee phylotype might actually consists of several bacteria species (see section 1.3: Methods for microbiota studies). This is apparent for the different *Lactobacillus* species mentioned above, which show high similarity in 16S rRNA gene identity but exhibit high genomic divergence [57], which then separates them as different species. This might hold for the other phylotypes as well, and especially for *Gilliamella apicola* (Gamma-1 phylotype) for which previous studies have reported high strain diversity [45, 58, 59]. Investigation of the Gamma-1 16S rRNA gene phylogeny across honeybees (*Apis*) and bumblebees (*Bombus*) has shown that the bumblebee strains cluster independently [44, 58, 60]. Based on their genome divergence towards *G. apicola* four new *Gilliamella* sp isolated from bumblebees have been described: *G. intestini*, *G. bombicola*, *G. bombi*, and *G. mensalis* [61]. Comparisons across bee species performed in previous studies, based on a portion of the16S rRNA gene, might thus have lacked the resolution needed to separate bacteria species. Therefore, these are exciting times in the study of honeybee GM as new and efficient methods for strain identification now exists that can be used to further resolve the honeybee GM composition.

Table 1: Bacterial species that today have replaced the old phylotype designation and their taxonomic description.

Bacteria species	Old phylotype	Bacterial class	Bacterial phyla
Gilliamella apicola [59]	Gamma-1	Gammaproteobacteria	Proteobacteria
Frischella perrara [62]	Gamma-2	Gammaproteobacteria	Proteobacteria
Snodgrassella alvi [59]	Beta	Betaproteobacteria	Proteobacteria
Bifidobacterium asteroids, Bifidobacterium coryneforme/indicium	Bifido*	Actinobacteria	Actinobacteria
Lactobacillus mellis and *Lactobacillus mellifer* [63]	Firm-4	Bacilli	Firmicutes
Lactobacillus helingborgensis, *Lactobacillus melliventris*, *Lactobacillus kimbladii*, *Lactobacillus kullabergensis* [63], *and Lactobacillus apis* [64]	Firm-5	Bacilli	Firmicutes
Bartonella apis [65]	Alpha-1	Alphaproteobacteria	Proteobacteria
Parasaccharibacter apium [66]	Alpha-2.2	Alphaproteobacteria	Proteobacteria
Alpha-2.1	Alpha-2.1	Alphaproteobacteria	Proteobacteria

*proposed split into Bifido-1 and Bifido-2 [57].

1.2.5 Ecology and function

Bacteria living in the gut are considered commensals or symbionts to their host. The former means that they coexist without harming or benefiting one another and the latter implies that one of the two (or both) benefits from the coexistence without harming the other [19]. The ecology of the GI-tract of humans and animals coincides with the varying gut microbiota composition found in its different parts [1, 6]. This is mainly so because bacteria have different ideal conditions for cell growth and adapt to the changing environment along and inside different niches of the GI-tract: e.g. variation in pH, aerobic vs anaerobic patches, the gut epithelia vs the gut lumen, gastric movement vs more static environment, and nutrient rich vs nutrient specific niches [1, 6]. The members of honeybee core GM exhibit the role as symbionts by individually contributing to distinct metabolic functions in the gut [23, 67].

Nutrient availability is among the prerequisites for cell growth and rapidly shapes the human GM: e.g. different dietary intake, which changes the availability of bacterial nutrients permanently or temporarily [68]. For honeybees it has been shown that aged diets (not fresh) have a negative effect on host health and correlates with a dysbiotic (diseased state) GM composition [69], which suggests that diet can influence the GM composition

also in honeybees. In general, the main function of the GI-tract is to aid in the nutrient uptake of the host, which influences host health [1]. In congruence with this view, members of the core GM in honeybees probably aid in the breakdown of food sources that the host cannot digest or process. This includes breakdown of pollen walls and metabolism of toxic sugars by *G. apicola* strains [46, 60]. The broad diversity of *G. apicola* strains is reflected in that only some strains are capable of pectin degradation (a main constituent of pollen walls) [46] and the metabolic capacity of sugar fermentation is strain dependent [60]. These results highlight the need for more in-depth strain characterization to deepen our understanding of the GM composition and its role in honeybee health [14, 23].

An additional beneficial function of the gut microbiota is to be a barrier against and fight off gastro-enteric pathogens [1]. Pathogens are microorganisms that infect a host and can result in host disease [70]. Obligate human gut-pathogens do not reside in the gut microbiota but are transient bacteria introduced through food or other environmental factors. Contrary, opportunistic human pathogens can be commensals of the GM and cause disease if introduced elsewhere on the body, for instance after GI-tract surgery or in urinary tract infections (*E. coli*). In honeybees, *F. perrara* shows characteristics of an opportunistic pathogen although its role in the GM is not yet determined, but it is associated with higher prevalence in honeybees within a diseased state [69]. *F. perrara* is found sporadically across the whole GI-tract, but sometimes in high numbers, when located in the pyloric region, which is located between the midgut and the ileum, where it is associated with morphological changes in host epithelial cells and honeybee immune system modulation [71]. Contrary, elevated numbers of *S. alvi* is associated with a healthy host state [69] and is found to have antagonistic effects on bee pathogens [72]. While individual bacterial species exhibit defensive properties, the GM composition should probably be “balanced”, and that of a healthy state, to be able to protect against pathogens and disease [73], as attempts to use single species as probiotics [10] has not yet been successful in honeybees [74].

In the same way as can be seen in humans and other animals, the honeybee GM can also stimulate hormone signaling [75] and their intrinsic immune system [76]. All these different functions of the GM underline what a complex environment this is, wherein microorganisms interacts, both in competition and in cooperation, with each other and its host maintaining homeostasis. Understanding the intricate web of interactions is what constitutes gut microbiota studies.

1.3 Methods for gut microbiota studies

1.3.1 Culture dependent methods

In human gut microbiota studies, the most studied part of the gut is the large intestine (colon), made possible through studies of fecal samples. Although not a perfect match, fecal samples are thought to best represent the colon gut composition without having to perform invasive surgery. Traditionally, bacterial studies refer to culturing bacteria in nutrient broths or on agar plates, and supplying them with the individual conditions needed for cell growth and proliferation [77]. Traditional culturing selects for the bacteria that grow well under rich nutrient conditions, bacteria that sustain their own metabolism (not depending on other bacteria/fungi), and bacteria not needing complex nutrients or environments for growth [78]. This bias introduced during culturing is known as "The great plate count anomaly" implying that the bacteria, which necessary conditions are not met do not grow in the laboratory [79]. Through culturing, both aerobic and anaerobic bacteria that make up large parts of the GM have been identified, but the knowledge about the aerobic proportion is in advance, due to their less complex growth requirements. In contrast, from the anaerobic and microaerophilic (needs enhanced CO_2) bacteria population culturing information is still lacking, but new methods are on the rise. New ways of mimicking growth in natural systems and thus co-dependence [77, 80-82] and new more complex methods for culturing "all" bacteria in a sample [78], including spore forming anaerobic bacteria [83] are all promising and provide new much needed knowledge.

To characterize cultured bacteria further, beyond their colony morphology visible to the naked eye, different microscopy techniques can be used to describe their cell shape, cell wall constituents and their growth patterns. Most bacteria are between ~0.4–3 μm in length [84] and colonies on an agar plate consist of thousands of bacterial cells. Light microscopy, be used to visualize bacterial movement in liquid culture at 400× magnification. With magnification at 1000× it is possible to determine bacterial shape (rods or cocci), detailed growth morphology (single, clusters, diplo, or chains) and when used in combination with Gram staining it provides knowledge of cell wall structure (Gram negative and Gram positive) [85, 86]. Even more detailed description of the cell surface can be obtained using a Scanning Electron Microscope, which sends electrons onto the cell surface that are reflected back to the detector and an image is created that can depicture the cell wall on nm scale [86].

Properties derived from culturing and microscopy are used for classification and taxonomy of bacteria and are thus important for identification [87]. After the discovery, identification, and acknowledgement of that the nt sequence in the DNA contained all the information about the organism, additional criteria for classification was applied based on the bacterial genome DNA sequence [88], hence the DNA-DNA hybridization (DDH) method . The theory behind states that genomes with more than 70% similar DNA sequences are more likely related, and today this method is still a reference criterion for distinguishing between bacteria species [89].

1.3.2 Sequencing methodology for GM and gut microbiome characterization

Because only a fraction of the microbiota can be studied using culturing methods, additional methods is needed to give more comprehensive insight into the GM composition and its functions. All information about a bacterium is stored in its DNA as genes, each with a unique function. The bacterial ribosomal small subunit gene – the 16S rRNA gene – is the most used gene for bacterial identification and phylogeny. This gene can be amplified by polymerase chain reaction (PCR) from pure culture DNA and then sequenced using the traditional dideoxynucleotide (ddNTP) Sanger sequencing technique [90]. This gene includes both highly variable and conserved regions, spanning about 1500 base pairs (bp) of which the whole or parts (variable region) can be used for taxonomic classification [91].

With advances in new sequencing technology, the possibility to detect uncultivable bacteria has emerged [92]. Several next generation sequencing technologies: SMRT® (Pacific Biosciences), Ion Torrent (Thermo Fisher Scientific), and Illumina sequencing platforms (Illumina) [93]. MinION (Oxford Nanopore Technology) technology shows great promise for longer read-length sequencing and can be applied for either for genome sequencing or community analysis [94, 95]. Each methodology has its benefits or drawback, depending on the sequencing target [96]. The most used method today is Illumina sequencing, which can be used for targeted gene sequencing of the 16S rRNA gene, in complex samples [97]. With the use of indexed universal primers, all 16S rRNA genes are amplified and then sequenced, which results in a sample vise compositional bacterial profile. This methodology includes the short read (150 – 300bp) sequencing by synthesis from Illumina, which when used as 300bp paired-end reads; sequencing reads can span about 500bp of the 16S rRNA gene. This high throughput sequencing method makes it feasible to characterize the complex GM

compositions and by including duel primer indexing large datasets can be analyzed [98]. This method can be used to compare the GM composition between different groups of samples, and provides valuable information about what factors that might influence the GM composition.

There is not always congruence between which bacteria are present and what functions the GM can perform. Therefore, additional information is needed to link which bacterium to which function/genetic ability [39]. For this purpose, Illumina shotgun metagenomics can be used, wherein transposon guided cuts are made across all DNA from a gut sample, then these fragments are amplified and sequenced. By linking the fragments together using computer algorithms (see section: 1.3.3 Data analysis) bacterial genomic content can be assembled and reveal bacterial function/genetic content for different bacteria species, and the collected functionality of the GM (microbiome) can be analyzed in response to environmental or host factors [99]. The same method (Whole genome sequencing – WGS) can be used on individual cultured bacterial cells, in which the complete genomic content of individual bacterial strains is revealed. This makes it feasible to track bacterial strains in epidemiologic studies, by using genome-wide association studies, linking phenotypic traits like antibiotic resistance, with genotypic content [100, 101]

1.3.3 Data analysis

Handling the amount of data that result from high throughput Illumina sequencing requires some computer programming skills. Fortunately, the open web source Bioconductor [102], distributes highly user-friendly and cost-free analysis solutions in the form of R-packages [103], which renders the statistically less experienced user the possibility for performing complex statistical analysis on big datasets with limited programming knowledge. High throughput sequencing data analysis is a fast moving field, thus several new R-packages exists for bacteria community analysis, and their use largely depends on study design and sample properties, since they largely reflect different statistical/mathematical approaches. A much used method for taxonomic assignment of the millions of 16S rRNA sequences resulting from Illumina sequencing, is the clustering of similar sequences based on >97% identity into operational taxonomic units (OTUs) [104] and then assigning taxonomy based on similarity within bacterial databases (Greengenes, SILVA, or RDP) [39]. Due to this cut off, this clustering method may falsely join sequences of different species together [105].

New statistical approaches have been implemented into high throughput sequencing analysis pipelines, that corrects sequencing reads and separate sequences into sequence variants (SV DADA2) or zero-radius OTUs (ZOTUS, UNOISE2), which then more likely represent bacterial species [105, 106]. This type of approach has shown to give more discriminative power between samples than the traditional OTU method [107] and are comparative between studies[108], but the use of only a part of the 16S rRNA gene will never reflect the true richness of microbial communities [92].

While the 16S rRNA gene can be used for bacterial species identification it is believed not to reflect the true bacterial evolution at the species and strain level [109]. This is because the the ribosomal proteins are very important fo preserving cell survival, it evolves more slow than proteins involved in adaptation. Therefore, it possesses limited resolution regarding intra-species diversity and additional marker genes should be used for strain level phylogeny [57]. To elucidate the more rapid dynamic evolution of GM members, one has to compare a larger part of the functional/evolving part within the bacteria genome [110]. Bacteria strain within a species can harbor different functional genes and the total gene pool of that species is called the pan-genome whereas the genes found within all strains of that species is called the core-genome [110]. By comparing/aligning some or all core genes, a phylogeny can be inferred that takes into account the different evolutionary rates of each core gene and groups together bacteria strains that are more evolutionary alike or have co-adapted to a certain environment [110, 111]. Several method exists for calculation evolutionary distances, but commonly they are based on detecting mutations, that is variation of single nucleotide differences/polymorphisms (SNPs) within a genes DNA, and then calculating distance according to the level of SNP variation between genes. For protein coding genes the amino acid sequence is used for initial alignment [111], which is then translated back to nucleotides.

Various methods for phylogeny analysis can be applied, and traditionally a neighbor joining (NJ) tree is built based on a sequence alignment of the genes of interest. NJ is a clustering method that joins together sequences that have the least SNP variations/evolutionary distance. With the increasing use of WGS, and comparisons of core genes, methods like maximum likelihood and Bayesian statistic have become popular and these are though to best reflect evolution as long as *a priory* assumptions about the data are met [112]. The maximum likelihood method calculates likelihoods for a possible tree

topology/arrangement, across several tree topologies, given the data and the model of choice, and the tree topology with the highest overall likelihood is chosen [112, 113].

The distance obtained by comparing the total number of SNPs found across the core genes of two bacterial strains is known as the average nucleotide identity (ANI), which gives information about evolutionary relationship at the species level [89, 114]. An ANI value of ~94-96% can compare to a DDH value of 70% and this method has become an invaluable tool for identification and separation of bacteria species [115]. Data analysis of WGS sequencing normally constitutes three parts; i) Quality filtering and error correction of raw fastq reads; ii) assembly of remaining reads into contigs (additionally into scaffolds); and iii) annotation of contigs to find out which genes are present within the genome. For each part, several software programs exist, that perform more or less equally well and are free of charge: e.g. i) Trimmomatic [116] and Quake [117], ii) Velvet [118] and Spadez [119], iii) Prokka [120], RAST [121] and NCBI Prokaryotic Genome Annotation Pipeline [122]. For downstream analysis, several free-of-charge R-packages are available, depending on preference, as well as commercially guided user-friendly interface (GUI) softwares (Geneious) that are excellent for annotation visualizations but also includes several software packages for sequence analysis [123].

Even though much information can be retrieved from WGS data and metagenomics it is still nessecary to actually grow bacteria and identify physiologically and biochemical properties, because some genes might not be expressed just because the gene is present [124]. This is especially important for identification of new bacterial species, but also to be able to elucidate community composition in regards to metabolic capacities of co-occurring as well and individual bacterial strains [67, 78].

1.4 Model organisms for gut microbiota studies

1.4.1 Animal models

To not inflict harm or disease to research objects, most human GM studies use fecal samples, just to prevent invasive procedures. New treatments or severe manipulations of the GM are first tried out in model animals before they are approved for human studies. Ethical guidelines exists also for animals models, but they are less stringent allowing numerous

possibilities for study design. There are no such thing as an ideal model, although different models can have several benefits, but also various drawbacks depending on the research focus. Summarized, there are three main arguments for why animal models are useful in terms of GM studies: i) The human gut microbiota is very complex and many bacteria are still uncultivable, thus we still do not know who is there on a species/strain level, ii) only the GM in the colon part of the GI-tract can be investigated through fecal samples, which leaves the function of the GM in rest of the GI-tract unexplored, and iii) there are big individual variations, which are reflected in numerous variables influencing the human GM, that are not possible to control over time or across many individuals, all of which can be overcome with the use of animal models [125]. To fully understand intrinsic host-microbiota interactions, we need model animals, and two main categories of model animals can be presumed: laboratory models and wild models [126]. Laboratory model can especially be used for manipulations and targeted studies while wild models are often used for exploring the GM microbiota and its responses in a natural setting [126].

Mice models have been used extensively as laboratory models because they are easily kept and their phylum level GM composition resembles that of humans [125]. Also, through years of genetic engineering, mice that develop diseases that mimic several human diseases exists, which can help in understanding the role of GM in disease, although carful study design is needed [125]. In recent years, it has become clear that laboratory mice exhibit a reduced GM compared to free living mice, due to their rearing within the laboratory, and thus they might not reflect a natural GM response or function [127]. Mice are also used as germ-free or gnotobiotic models, which either are without bacteria or inoculated with known bacteria, respectively, and are useful models for understanding GM colonization processes and host effects of single metabolites [128]. Other models that are used for manipulations by diet and different treatments are pigs, cattle, chicken and fish [125, 126], the latter permits inclusion of more individuals, which can be a drawback with the use of larger animals. In addition, insects have emerged as alternative cost efficient models. Insects, in general, show less complex GM compositions [4] with only 2-20 bacterial genera, of which most are cultivable [129]. The fruit fly (*Drosophila melanogaster*) is the most used species and has been used for a variety of studies [125], including gut-behavioral studies [130]. One drawback of using fruit flies for GM studies is that its GM is largely diet and environmentally dependent [129] and does not harbor a consistent core microbiota [131].

1.4.2 Honeybees as models for GM studies

There are advantages for using the honeybees as a model for GM studies. It has the a low complex core microbiota that is not found elsewhere outside the honeybee and bumblebee gut, thus the bacteria-host interactions are solid and reflected in the social transmission of the GM [44]. In contrast, only two bacterial families have been identified that are present in humans spanning different ethnicity and geographic origins and thereby might be considered a part of our core microbiota [54]. Compared to humans, all members of the core microbiota are cultivable, and they show high strain diversity [45, 57]. Even though the GM composition varies between humans and honeybees (host specific, mostly anaerobic vs grow best in enhanced CO_2 atmosphere [67]), high strain diversity is also detected in humans [3] and thus honeybee gut bacteria can be used to elucidate the role of this diversity within the GM. An obvious advantage of animal models is that investigation of the whole GI-tract is feasible. Because the human and honeybee GM is both gut part dependent, investigation of distinct parts of the honeybee GI-tract microbiota can be used to unravel gut part functionality and further elucidate the role of each bacterial member though their metabolic profiling of [67]. Honeybees harbor only an intrinsic immune system, but this simplicity might be the key to extract knowledge from an otherwise very complex web of interactions [13]. Honeybees are easily manageable in the lab and the GM is possible to manipulate, which has resulted in both germ-free and gnotobiotic models that can be used for bacterial-host interaction studies [43, 67]. The honeybee genome is described in detail [132], and all honeybee workers are sisters, which makes for a limited diversity in their genetic content. They are also numerous, thus experiments with high numbers of individuals can be performed. Finally yet importantly, honeybees are managed by humans but live their life in the wild and are thus a mix between domestic and wild animals. This makes honeybees a unique model in investigations on how both natural and highly controlled environmental settings can influence GM composition.

The use of honeybees as GM models started relatively recent, but due to their sociality and human brain similarities, they are also used for physiological, learning and aging experiments [30, 133-135]. GM, physiology and behavior is tightly connected, thus honeybees can be excellent model for complex individualistic studies, that might reveal information about the more complex human system. Importantly, it is not only beneficial to use honeybees as models for understanding human GM functionality, but it is also important

to study honeybee GM per se, to preserve a healthy and sustainable honeybee population, which is crucial for efficient food production.

1.5 Bacterial antibiotic resistance

1.5.1 Antibiotics and antimicrobial resistance

Antibiotics or anti-bacterial compounds are used as medicine to prevent or treat bacterial infection. Already in the late 19th century antibiotic effects were observed as a natural ability of the fungi *Penicillium* to inhibit bacterial growth, but Alexander Flemming was the first to publish the accidental discovery of the antibiotic Penicillin in 1928, a compound that was later purified and named Penicillin G and used to treat several infections [136]. This significant discovery initiated the search for additional natural compounds that could kill bacteria but not affect human cells. Diverse fungi and *Streptomyces* sp. living in soil are most studied and a large variety of antimicrobial compounds has been identified from these. Today several types of chemically synthesized antibiotics exist that are mimicking the original natural-occurring compounds [137], but some compound are still used in their natural form [138, 139]. In medical settings, the different antibiotics existing today are classified according to their mode of action, and thus five main bacterial target sites exists. These are the cell wall synthesis, cell membrane, the protein synthesis, and the DNA and RNA synthesis [139] (Figure 2).

Some antibiotics kill off bacteria when used above a certain concentration (bactericidal), but some only inhibit them (bacteriostatic) and depend on the host immune system for complete bacterial elimination [139]. Due to our advances within medical surgery, organ transplant technology, cancer and disease therapies more people are immunocompromised today than before and thus the need for efficient antibiotics is increasing. Pharmacokinetics, the concentration in different host tissues, the ability to penetrate a bacterial cell, and the half-life stability, are all factors influencing the effectiveness of the antibiotic compound during antibiotic treatment. The latter two are especially important in an environmental setting determining the persistence of the antibiotic in the environment. Agricultural and other environmental settings are considered low-antibiotic concentration settings [140]. If the antibiotic concentration is too low or do not kill the bacteria, chances are that the bacteria will survive and develop resistance mechanisms.

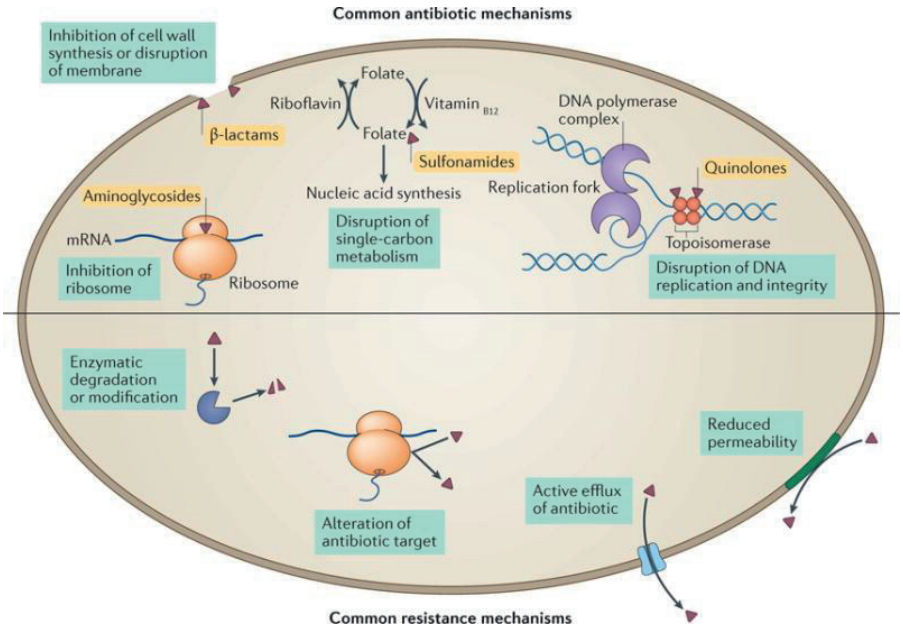

Figure 2: Main target sites of antibiotics and bacterial resistance mechanisms. Reprinted from [141].

The resistance mechanism that develops depends on the antibiotic target site and for some antibiotics, multiple mechanisms are found. The main resistance mechanisms known today are production of enzymes that inactivate the antibiotic, efflux pumps, changed target proteins, altered metabolic pathways, and production of immunity proteins (Figure 2) [141, 142]. Antibiotic resistance can be either intrinsic or acquired. Intrinsic resistance is due to cell wall impermeability or lack of the specific antibiotic target site [143], while acquired AR can be caused by mutations within antibiotic target genes [144] or by horizontal transfer of AR genes (ARGs) between bacteria that invoke the resistance mechanisms mentioned above [142]. Acquired resistance can develop following selection pressure by antibiotic exposure as a bacterial Darwinian survival mechanism either within a host or in the environment.

Although production of anti-bacterial compounds by environmental bacteria can to some extent, select for antibiotic resistance among competing neighboring bacteria, the main driver of antibiotic resistance is the extensive use of manmade antibiotics in medicine, food

industry and agriculture [145-147]. The emerge of bacteria that are resistant to three or more antibiotics (multidrug resistant) [148] is already a global health problem recognized by the WHO [149]. These multidrug resistant bacteria are worldwide causing everyday infections some not possible to treat with common antibiotics [148]. Several disease causing multidrug resistant bacteria belong to the class *γ-proteobactera*, many of which are commensal or symbionts of the human and animal GI-tract or free-living environmental opportunistic pathogens.

1.5.2 Spread and persistence of bacterial antibiotic resistance genes

The use of antibiotics in medicine has been the main driver of AR within human pathogens and the use of antibiotics as growth promoters in farming has led to an increase in AR animal symbionts [150] some of which also cause human infections [151]. Spread of AR bacteria (ARB) between farm animals and humans has been documented [152] and insects has been shown to be a vector for ARB spread within farm areas and in the food industry [153]. There are not many studies that describe ARGs in insects, in which most bacteria studied, are *E.coli* and *Enterococcus* sp. [153], which are rather human pathogens and not intrinsic insect specific. Phenotypic resistance in bacteria isolated from insects digestive tracts has been studied but these bacteria were also mostly environmental and no detection of ARGs were performed [154]. Hence, more information about ARGs in insects' core bacteria is needed to address the role of insects in spread of ARGs within the environment.

Despite numerous studies on ARGs found in humans and in natural environments, the direct link between antibiotic usage and ARGs development, spread and persistence in natural bacterial populations is not very straightforward because reducing the antibiotic exposure does not always limit the AR prevalence and persistence [155]. The development of AR due to antibiotic exposure is best demonstrated for AR that is caused by genome mutations, where persistence is maintained through vertically transmission across generation [144]. It is documented that the development of AR-causing mutations leads to fitness costs for the bacteria. However, new findings support that additional mutations arise that counteracts the fitness cost caused by the original mutation [155]. In addition, these mutations persist without any antibiotic exposure, which indicate that they do not generate a high fitness cost in the mutated bacteria [144]. Even more complex interaction and mechanism are in play for

the spread and persistence of ARGs associated with transfer elements, wherein the role of antibiotic exposure is not fully understood [156].

ARGs can transfer horizontally (horizontal gene transfer (HGT)) between bacteria by three main mechanisms: Transduction (transfer by phages as vectors), natural transformation (bacteria take up free DNA), and conjugation (direct transfer through bacterial pili of transferable elements like plasmids) (Figure 3) [157]. In human pathogens many ARGs are found within plasmids and thus plasmid transfer is extensively studied and believed to be the main cause of spread of multidrug resistant bacteria [157]. Plasmids can exist within bacterial cells separate from the bacterial genome or like small plasmids incorporated into the bacterial genome [158]. A link between bacteria living in different environments is created, since many plasmids have a broad bacterial host-range that can include all Gram negative bacteria [157]. Plasmids transfer mainly by conjugation, and close bacterial contact is thus needed for transfer to take place, which can be an obstacle for transfer in sparse bacterial communities. Large conjugative plasmids can initiate their own transfer and smaller mobilizable plasmids (in the host genome) can be co-transferred [157, 158]. Plasmids can serve as a vector for other AR-containing mobile elements like transposons. Transposons can move between plasmids and bacterial genomes by their own transfer mechanism (conjugative transposition), which is regulated and expressed by the transposons themselves [159]. Several types of transposons exist [159], but the DNA transposon normally contains one or more genes that causes a phenotypic change in the host [158] such as ARGs. Several ARGs are associated with various transposons and these transposons can be found either within the bacterial genome or within different plasmids, which makes transposons extremely potent transfer elements of ARGs.

Some acquired ARGs found in human pathogens are believed to have an environmental origin [157, 160], and it is acknowledged that the environment is a reservoir for ARGs, although there are many confounding factors to the understanding of how this transfer occurs [156]. One example is that ARGs can express different functions in environmental bacteria than the same ARGs do in a pathogen [160]. This theory is proposed for efflux pumps that pump out antibiotics in a pathogen, but is probable survival mechanisms to evade exposure of heavy metals in an environmental bacterium [147, 161]. Hence, there are many factors exerting a selection pressure in an environmental setting, causing bacteria to evolve. Apart from the obvious selection pressure of high concentrations of antibiotics on

the development of ARB, low concentrations or sub-inhibitory concentrations (below the minimum inhibitory concentration – MIC) of certain antibiotics have shown to trigger ARGs development as well as bacterial stress responses which subsequently influence the intrinsic capability of bacteria for HGT [140].

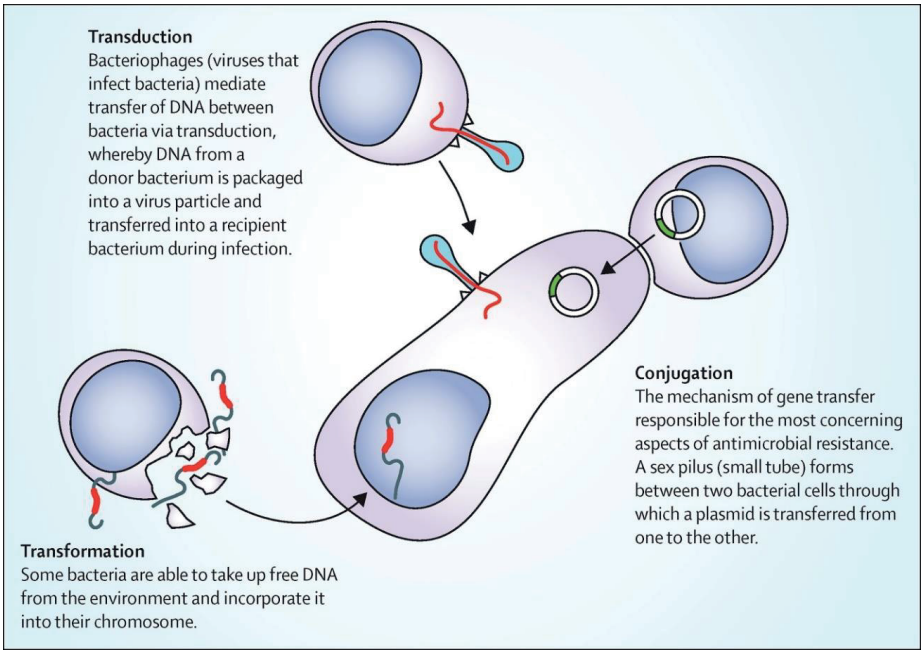

Figure 3: Three main ways of which ARG can spread between bacteria. Reprinted with permission from [147].

Continuous exposure of antibiotics within an environment will in addition to contributing to transfer of ARG, facilitate persistence of ARB and ARG in that environment. Other factors that might influence persistence of ARG are: co-selection of plasmid borne genes that imposes benefits for the host; e.g. resistance to metals or increased fitness without selection pressure, lack of or reduced fitness cost due to acquired ARG; e.g. transfer of transposon mediated resistance into the host genome [155], and high transfer rate (conjugation efficiency) of certain plasmids during absence of antibiotic selection [162]. Environments where many of these factors that facilitate transfer and persistence might come together, are considered hot spots for ARG transfer and one such environment is the GI-tract.

1.5.3 The gut microbiota as reservoir for bacterial antibiotic resistance

The gut microbiota is a densely populated bacterial environment that spans a diversity of several microbial classes [26]. The close contact between these bacteria makes the gut a perfect place for HGT between distantly related Gram positive- and Gram negative bacteria [163]. The high proportion of *proteobacteria* in the gut of many animals, including insects, facilitates the spread of broad host range plasmids by conjugation within this phylum [4, 164]. Additional transfer by transformation and transduction between GM members also happens [26] and both mechanisms can facilitate the transfer of whole plasmids [165]. HGT is shown to be more efficient within ecologically similar niches [166], which indicate that bacteria occupying the same space will more likely be able to exchange genes. The GM interacts with bacteria from different environment, since the host ingests live bacteria through feed and then defecating live bacteria. Therefore, fecal droppings are a source of ARGs spread to various environments. Human gut metagenome studies have also paved way for the belief that the GM is a reservoir for ARGs [167], and not only a hot spot for transfer, since more ARGs have been identified from gut samples than from any other environment [168, 169]. This reservoir contains more ARGs against antibiotics that have been used for a long time or been approved for use in farming and agriculture [170]. This exemplifies the impact that non-human use of antibiotics can have on human health and thus the commensal/symbiotic gut bacteria [165]. In humans, ARB and ARGs can persist up-to four years post antibiotic treatment [171]. This underlines one further factor influencing the persistence of both ARB and ARGs within the GM, which is host selection.

The GM being a reservoir for ARGs has also been documented in honeybees [172], and as seen for humans and animals, the ARGs are located within the GM symbiotic population [173]. During antibiotic treatment, the GM is exposed to both high- and low concentrations and often short-term perturbation of the GM composition occurs, where either ARB survive and proliferate or new ARB evolve [171]. These perturbations seems to be inflicted even at bacterial strain level [171, 174, 175]. Perturbations can impact GM functions and thus host health [176] (Figure 4), as seen in honeybees, when bees are treated with tetracycline and lose their ability for pathogen defense [174, 177]. The honeybee gut symbionts harbor several ARG against the antibiotic tetracycline, and the number of different gene correlate with the use of a derivative of tetracycline (Oxytetracycline) [178] to treat and prevent honeybee infections [172]. The prevalence of ARGs depends on geography, or more

accurately: on the country-specific antibiotic usage strategies. Countries that use much antibiotics in general have higher prevalence of ARGs in both GM populations and in the environment [168, 172]. AR is found in bacteria from environments not believed to have been exposed to antibiotics, like wild animals [179]. This reflects that when ARGs first are introduced into a GM environment they will most likely persist.

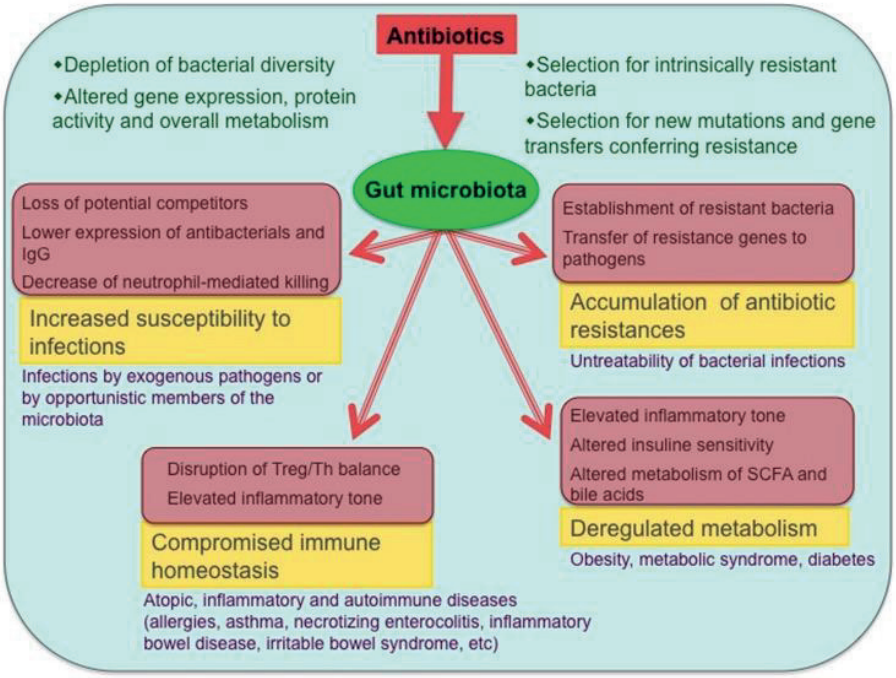

Figure 4: Ways that antibiotic treatment can influence host health. Reprinted from [176].

1.5.4 Tetracycline and Streptomycin

Tetracycline and streptomycin, although by slightly different mechanism, both function by inhibition of the 30S bacterial protein synthesis. They were discovered in the late 1940s and are used to treat bacterial infections caused by aerobic bacteria both Gram-negative and Gram-positive [180, 181]. They are still in use as therapeutic options for human infections but due to their widespread use in clinical and veterinary medicine and agriculture, broad resistance among a variety of bacteria exists [181]. This is exemplified in that bacteria from the early 1950s show very low levels of tetracycline resistance (Tc^r), but today Tc^r is found in almost all environments in both pathogenic and commensal bacteria [181-183]. Tc^r is due

to acquired *tet* genes, that are associated with mobile genetic elements (MGEs) such as plasmids and/or transposons [181]. Thirty-eight different genes are described which infer resistance in three main ways: as efflux pump proteins (23 genes), as ribosomal protection proteins (11 genes), or as enzymatic inactivation proteins (3 genes), and one gene remains unknown (Table 2) [184, 185]. Several *tet* genes are found within multiple bacterial taxa but some are also more bacterial specific, which might reflect their MGE association [182, 184]. In later years, more bacteria that harbor two different Tc^r genes conferring the same resistance mechanism are described, but if this reflects increased antibiotic usage remains inconclusive [184]. Tetracycline resistance genes are named *tet*(A), *tet*(B), *tet*(C) etc throughout the alphabet and the newest identified genes have numbers [184, 186]. The efflux pump proteins genes consists of the gene coding for the pump itself (e.g. *tet*(A)) and a gene coding for a regulator (*tetR* for that class) that represses the expression of the pump protein by binding to the pump gene promotor. When tetracycline is present, the repressor protein binds to it and the resistance protein is expressed.

Table 2: Mechanisms for known tetracycline resistance genes, as of [185].

Efflux (n = 30)	Ribosomal protection (n = 12)	Enzymatic (n = 3)	Unknown
tet(A), *tet*(B), *tet*(C), *tet*(D), *tet*(E), *tet*(G), *tet*(H), *tet*(J), *tet*(V), *tet*(Y), *tet*(Z), *tet*(30), *tet*(31), *tet*(33), *tet*(47), *tet*(35), *tet*(39), *tet*(41), *tet*(K), *tet*(L), *tet*(38), *tet*(45), *tetA*(P), *tet*(40), *otr*(B), *otr*(C), tcr3, *tet*(42), *tet*(43), tetAB(46)	*tet*(M), *tet*(O), *tet*(S), *tet*(W), *tet*(32), *tet*(Q), *tet*(T), *tet*(36), *otr*(A), tetB(P), *tet*(44)	*tet*(X), *tet*(34), *tet*(37)	*tet*(U)

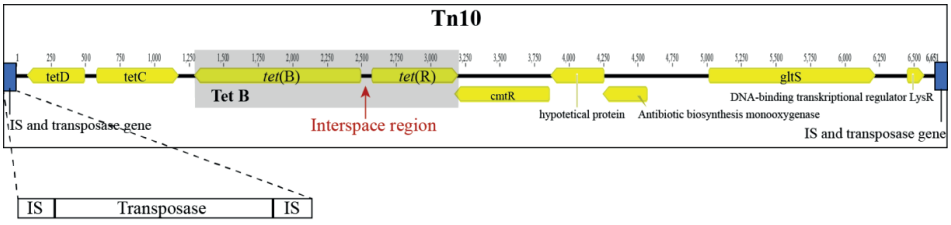

Figure 5: The Tn10 transposon and its associated genes. The Tet B determinant is shown, which constitutes of the *tet*(B) gene, that codes for the efflux pump, and its regulator *tet*(R). The transposase genes are located at either ends; shown as blue boxes and zoomed to show that there are always two inverted repeat IS elements on each side of the transposase gene, which codes for the cut and paste mechanism of the transposon.

The efflux pump type resistance genes are the most studied and among these, *tet*(B), normally associated with the transposon Tn10 (Figure5), is one most widely distributed among Gram negative bacteria [184, 185, 187]. Tetracycline is one of the antibiotics that have been shown to increase HGT and transposition even at low concentrations [188, 189].

Like seen for tetracycline resistance, streptomycin resistance (Sm^r) mostly results from acquired streptomycin resistance genes [190, 191]. Sm^r can occur due to alterations in the ribosomal binding site, reduced cell membrane permeability, and production of aminoglycoside modifying enzymes [180, 190]. Two such enzymes are the product of the linked *strA-strB* genes, which are phosphotransferases that modify the 6-hydroxy group within the streptomycin molecule and originate from *Streptomyces* sp. [180, 190]. These genes are often associated with a particular transposon, Tn*5393* [192], identified in several plasmids in human and animal commensals and pathogens [191, 193], and in environmental bacteria such as in soil [194] and plant pathogens [192]. Depending on the bacterial host, the transposon may include an insertion element (IS-element), either IS1133 or IS6100, in addition to the transposon genes (Figure 6), which are found to increase expression of the *strA-strB* and thus elevate the MIC for that particular bacterium [195].

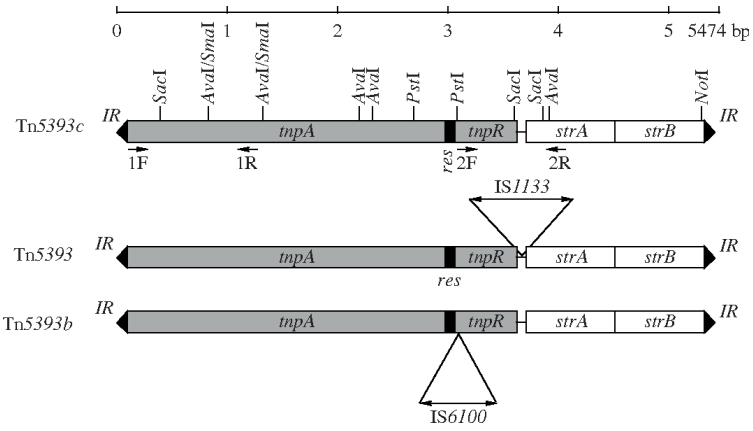

Figure 6: The Tn*5393* transposon containing the *strA-strB* resistance genes. The transposase gene and the regulator gene are shown and named *tnpA* and *tnpR*, respectively. Variations of this transposon exists depending on the inclusion of the IS elements IS*1133* or IS*6100*. Reprinted from [196].

1.5.5 Methods for identification of bacterial antibiotic resistance

Identification of phenotypic antibiotic resistance is performed by MIC determination of clinical isolates. Cultured bacteria are spread across an agar plate and paper discs inoculated with certain concentrations, are placed on top and the plate is incubated over-night. Then zone diameters across the discs are measured, which corresponds to a certain MIC concentration. Alternately a strip inoculated with a range of concentrations can be used to directly measure MIC. Pathogenic bacteria with a MIC above a certain threshold are considered resistant and the antibiotic compound tested cannot be used to treat infections caused by these bacteria. MIC thresholds are experimentally set and determining MIC should follow strict methodological guidelines (EUCAST guidelines, http://www.eucast.org). A universal detection approach is essential to be able to compare across groups of bacteria from different environments [179]. For environmental bacteria, a MIC threshold is set, by calculating the wild-type (WT, bacteria that do not harbor ARGs towards the antibiotic tested) population's concentration at which the antibiotic starts inhibiting bacterial growth, and the above concentration will be the MIC [156]. Interestingly, different ARGs towards the same antibiotic compound express different MIC, thus the phenotypical genetic background must be identified [197].

Identification of ARGs from complex samples can be done using gene specific PCR and qPCR for the quantification of ARGs. Genetic variation can be identified within samples by using targeted amplicon sequencing in the same way as the 16S rRNA gene is used for GM composition characterization. Metagenomics, where the whole genetic pool is sequenced and the resulting DNA reads are screened against a database containing known ARG is often used [141]. Different databases exists for this purpose; e.g. Antibiotic Resistance Gene Database (ARDB) [198], ResFinder [199], and PATRIC [200]. The same approach can be applied for WGS data from bacterial strains. WGS of bacterial strains has become a tool not only for ARG identification, but also for elucidating how ARGs develop and spread. Comparative genomics of: (i) core genes (see section: 1.3.3 Data analysis) can be used in tracing epidemic ARG strains, (ii) single ARG genes can unravel ARG origin and spread within a community, and (iii) whole genome comparisons can unravel acquisition of ARG and associated MGE and new ARG mechanisms. Despite these new methods, we still lack knowledge on how ARB and ARGs spread and persist within the GI-tract (*in vivo*) mainly due to the complexity of the human and animal GM, which causes difficulties when

monitoring ARG and assembling of genomes from metagenomics data [141]. To investigate transfer events *in vivo*, insects can be used as less complex models.

2 Aim and sub goals of the PhD thesis

While the commensal and symbiotic bacteria of the gut microbiota (GM) are known to harbor antibiotic resistance genes (ARGs), these genes can spread to transient pathogenic bacteria making the gut microbiota a reservoir for ARG dispersal. Therefore, more knowledge about how these ARGs establish, spread and persist within the GM is needed. The overall aim of this thesis was to investigate how ARGs spread to and within a gut bacterial community and address factors influencing the persistence of ARGs in the GM. To overcome the challenge of human GM complexity, we used honeybees as a less complex model system, to address this.

The work conducted in this thesis was divided into five sub goals:

- Establish baseline knowledge for honeybee GM composition and identify factors important for variation in the GM of Norwegian honeybees (Paper 1).
- Compare the prevalence of ARGs in two honeybee populations with different antibiotic treatment regimes (Paper 2 & Paper 4).
- Investigate strain diversity, of gut symbionts, from two populations with different antibiotic treatment regimes (Paper 2).
- Identify associations between ARGs, bacterial strains and mobile genetic elements (Paper 2 &4)

In addition, we performed a comparison of the strain diversity within the gut symbiont *Gilliamella* and characterized a new species; *Gilliamella apis* sp. nov (Paper 3), to confirm the findings in Paper 2.

3 Main results

Paper 1:

Shifts in the midgut/pyloric microbiota composition within a honeybee apiary throughout a season.

Jane Ludvigsen, Anbjørg Rangberg, Ekaterina Avershina, Monika Sekelja, Claus Kreibich, Gro V. Amdam, and Knut Rudi (2015). ***Microbes Environ*** **30, 235–244**.

To identify bacteria associated with the Norwegian honeybee population, we cultured bacteria from the midgut/pyloric region of honeybee guts. We collected ten bees from three different hives to comprise possible between-hive variation. Additionally, these bees were collected from three different location within each hive to cover age differences in the bee population. The midgut/pyloric region of the guts were aseptically dissected out and pooled, before plating and incubation in an enhanced CO_2 atmosphere at 37 °C for 2 days. In accordance to previous studies, we identified eight main bacterial groups (Figure 1) that make up the majority of the honeybee gut microbiota.

The study of honeybee gut microbiota was in its infancy at the time when we started this work, so no information was available for Norwegian honeybees. Therefore, this study was used to determine a baseline for a normal gut microbiota in Norwegian honeybees. We investigated the midgut microbiota over the course of a foraging season and under a stable diet to determine influential factors to consider when performing comparisons of gut microbiota across time and/or geography.

For this analysis we collected 30 bees (as described above) each month over a foraging season (Mai-October) and bees from two time points during winter (November and February). The low-complex honeybee gut population made it possible to use a mixed Sanger-sequencing approach of total 16S rRNA genes from each sample to investigate the most dominant bacteria in the gut population over time. We identified four bacterial components, whose relative abundance differed between months, with *Gilliamella* changing the most drastically from high abundance in Mai to an absolute low in October (Figure 2). The alpha-diversity changed from an absolute low (low diversity) in May to a high (more diverse population) in August. We also investigated the total bacteria to bee ration using

qPCR. This showed that there were more bacteria present in Mai and that the number of bacteria decreased until August with an absolute low in October.
In contrast to the changing bacterial composition throughout the foraging season, the gut microbiota composition was stable across a period of four months (November – February) when the bees were not foraging but fed a stable sugar diet inside the hives (Figure 3).

At the time (2013), the Illumina sequencing methodology was still a relatively new method for bacterial composition analysis. Nevertheless, we tried out an in house sequencing set-up for Illumina MiSeq of the total midgut 16S rRNA genes, using a control bee gut sample. With this technique we could identify the main bacteria groups as found by previous studies using 454 pyrosequencing, in addition to confirming the low complexity of the honeybee midgut (Figure 1, Supplementary Figure S2). The Illumina 16S rRNA sequencing results of the honeybee midgut/pyloric microbiota from this study were the first ever described and created a foundation for further use (Paper 5).

Paper 2:

Geographically widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant patterns.

Jane Ludvigsen, Davide Porcellato, Trine M. L'Abée-Lund, Gro V. Amdam, and Knut Rudi (2017). ***Mol Ecol***

We used a random selection of the dataset from Paper 1 (90 bees from Ås, Norway) and a dataset of 90 bees collected in Arizona, USA to investigate the prevalence of four of the most frequently identified tetracycline resistance genes in a previous study [172] using qPCR. These two datasets represent differences in antibiotic usage policies in two different geographic regions. Norway do not use antibiotics in their beekeeping while in the USA, antibiotics has been used for decades to treat beehives with honeybee specific infections as well as it is widespread used in agriculture.

We found that there was a higher prevalence of tetracycline resistance genes in the Arizonan population (100%) compared to the Norwegian population (<30%) (Figure 2a). In addition, there was a higher relative load of *tet*(B) in each bee in the Arizonan compared to the Norwegian population but the *tet*(H) relative load was not significantly different (Figure 2b).

Using culturing techniques established in Paper 1, we cultured bacteria from the midgut/pyloric region of the honeybee gut of three pooled samples (each containing 10 midguts) from both Norway and Arizona. Bacteria were selected based on colony morphology resembling *Gilliamella* and *Snodgrassella*. We genome sequenced 48 *Gilliamella* and 22 *Snodgrassella*, which showed that some Norwegian strains harbored tetracycline genes (*tet*(B) or *tet*(H)) and some not (Table 1), while all stains from the Arizonan population harbored tetracycline genes (even both genes).

We phylogenetically compared strains of *Gilliamella* and *Snodgrassella* based on nt differences within their core genes and found that *Gilliamella* cluster into four subgroups belonging to two main groups that were about 80% different (Figure 3a) and that showed different functional capabilities; e.g. ability to degrade pectin (Supplementary Figure S1). Most strains in subgroups of *Gilliamella* were found both in the Norwegian and in the Arizonan population indicating that no specific antibiotic resistant strains were favored by tetracycline treatment. Instead the *tet*(B) genes were found on the Tn10 transposon in all *Gilliamella* and were distributed among all three subgroups through HGT (Figure 3a). In contrast the *tet*(H) genes were subgroup/linage specific in both *Gilliamella* and *Snodgassella* and found within the genome and not on transferrable elements. These results show that ARG are prevalent and are spreading within the gut core microbiota of honeybees.

Paper 3:

Resolving the diversity of the honeybee gut symbiont *Gilliamella*: description of *Gilliamella apis* sp. nov., isolated from the gut of honeybees (*Apis mellifera*).

Jane Ludvigsen, Davide Porcellato, Knut Rudi (2017). *In revision: IJSEM.*

In Paper 2, we found that the two main groups of *Gilliamella* had an ANIb value of about 80%, which suggests that strains belonging to these two groups are different *Gilliamella* species. These two groups also showed differences in genome size, G+C %, and 16S rRNA gene identity towards *G. apicola* wkB1 (type strain for this species isolated from *Apis mellifera* - The Western honeybee) (Paper 2; Figure 3b, Table 2). All results summarized pointed towards Group 2 being a different species. In this paper, we then further characterized four strains – spanning Group 2: A7 and A-TSA-1 (Arizona), and N-G2 and

NO3 (Norway). These stains were characterized towards known *Gilliamella* species, by comparing draft genome distances (*in-silico* DDH – ANI and GGD), genetic content, and growth- and metabolic characteristics (Table 4). Based on differential characteristics towards other *Gilliamella*s but similarity in main respiratory quinone and predominant fatty acids, strain NO3^T was proposed type strain for a new *Gilliamella* species; *Gilliamella apis* sp. nov.

The diversity of the *Gilliamella* genus is reported to be high in several studies, which are based only on 16S rRNA gene- or core gene comparisons, but no study has conducted a thorough comparison of all strains spanning this diversity and connected between previous literatures. Therefore, we also included core genes and draft genome comparisons of 52 *Gilliamella* strains, genome sequenced in previous studies (including strains both from *Bombus* and *Apis*), which spanned the diversity reported to exist within the *Gilliamella* genus (Table 1). A phylogenetic ML tree of core genes (Figure 1) in combination with ANI and GGD calculations (Table 2 & 3) clearly showed that some strains should be renamed according to our new proposed species and that additional new species could be proposed in the future after further phenotypic characterizations.

Paper 4:

Linking streptomycin resistance genes (*strA-strB*) in a honeybee gut symbiont to environmental antibiotic exposure.

Jane Ludvigsen, Gro V. Amdam, Knut Rudi, Trine M. L'Abée-Lund (2017). *Submitted – Note for Microbial Ecology*

The antibiotics tetracycline and streptomycin are frequently used in the USA in agriculture but are not used in Norway at present. This, in combination with the use of tetracycline to treat honeybee infections in the USA have influenced and probably caused the high difference in tetracycline resistance prevalence between these two locations that we detected in Paper 2. The prevalence of streptomycin resistance in the honeybee GM has not yet been investigated. Based on our- and previous findings of tetracycline resistance genes in the honeybee GM, we hypothesized that low-level streptomycin exposure could invoke honeybee gut bacteria to acquire and maintain streptomycin resistance genes from environmental streptomycin resistant bacteria. We used the datasets from Paper 2 of 180

midguts collected from Arizona and Norway. We screened the samples using qPCR for the linked streptomycin resistance genes *strA-strB*, and we found a higher prevalence in the Arizonan population than in the Norwegian: with 17 out of 90 bees being positive in contrast to one out of 90, respectively. Interestingly, by metagenome analysis, performed using shotgun sequencing (Nextera XT) on the Illumina MiSeq platform, we identified a contig in the Arizonan population, containing the *strA-strB* genes, which belonged to *Snodrassella alvi*.

After culturing bacteria on plates with 4- and 12 mg/l streptomycin selection, using the same samples from Arizona and Norway as used in Paper 1 & 2, we could isolate one strain (E1) from the Arizonan population that contained the *strA-strB* genes. Sanger sequencing of the 16S rRNA gene verified this strain as *S. alvi*. Comparison of streptomycin MIC was performed between the *S.alvi*_E1 stain and *S. alvi* strains from both Arizona and Norway that did not posses *strA- strB* genes (WT), which showed an increased MIC of the *S. alvi*_E1 strain from 0.75 mg/l to 48 mg/l.

To assess if the *strA-strB* genes are associated with any transfer element, we genome sequenced the *S. alvi*_E1 strain. The *strA-strB* are normally found within the transposon Tn*5393*, but this was not apparent in our genome assembly. The association of the *strA-strB* genes with the transposon was verified by amplification of a PCR fragments spanning the *strA* gene and the transposon gene *tnpA* and the *IS*1133 element, that is often found within the Tn*5393*, was also identified (Figure 1). This study represent the first report of Tn*5393* associated *strA-strB* genes identified in a honeybee gut symbiont.

4 Discussion

Correlation of antibiotic exposure with ARG prevalence

In general, there is a strong correlation between a country's antibiotic usage policies and the prevalence of ARB in this country [149, 170], but the underlying driving factors for this observation are not well established. The long-term use of tetracycline in honeybee management in the USA have resulted in overall higher prevalence of tetracycline resistance genes (*tet* genes) in the honeybee GM from the USA compared to countries that do not use tetracycline in their beekeeping [172]. To further explore this observation, we compared the GM of honeybees from Norway and Arizona (USA), for the presence of four *tet* genes identified as the most prevalent in countries that do not use tetracycline in honeybee management (Paper 2). Less than 30% of the Norwegian bees were positive for *tet* genes, while all of the tested Arizonan bees harbored *tet* genes and the majority of them harbored more than one gene (Paper 2, Figure 2a). Summarized, this underlines that the above-described correlation holds true also for our study population. Although some *tet*-free honeybee symbionts, of *G. apicola* and *S. alvi*, have been previously isolated from honeybees in the USA [172], all USA strains isolated in our study (Paper 2) contained *tet* genes (Paper 2, Figure 2b). ARB and susceptible bacteria have been shown to co-occur in bacterial populations [201], such as in the Norwegian honeybee symbiotic population, but it seems like long-term antibiotic treatment in the USA might have changed these dynamics. In addition, none of *G. apicola* isolated from Norway, contained more than one *tet* gene, whereas 23 % of Arizonan *G. apicola* isolates harbored two variants (Paper 2, Figure 2c). This suggests that direct antibiotic exposure correlates not only with higher prevalence of ARGs but also with the presence of multiple ARGs within each bacterium.

In addition to direct antibiotic exposure of honeybee hives, honeybees can also encounter antibiotics in varying concentrations in the environment, especially if antibiotics are used in agriculture [140]. For example, in the USA, streptomycin is directly applied onto fruit orchards for fighting plant pathogens [202]. Interestingly, higher prevalence of acquired streptomycin resistance genes, *strA-strB*, was detected in bees from Arizona compared to Norway with 17/90 and 1/90, respectively (Paper 4). No cross-resistance is reported for tetracycline and streptomycin in environmental bacteria [194], therefore the most likely explanation for the increase in prevalence of ARGs in the Arizonan honeybee GM is the indirect exposure of honeybees to streptomycin in the environment.

Effect of long-term antibiotic exposure on strain diversity

The core microbiota members of the honeybee gut show high strain diversity [45, 57]. Short term tetracycline treatment (antibiotic exposure for approximately one week) with high tetracycline concentration (450 µg/ml) has been previously shown to alter the honeybee GM composition both in their relative abundance and at the strain level of core GM members [174, 175]. We investigated whether long-term tetracycline (oxytetracycline) exposure would change the strain diversity of the gut symbiont *G. apicola* (Paper 2). As in bacterial infections, wherein one pathogenic strain overtakes the population, the same can be argued to occur for ARB under antibiotic selective pressure (Figure 7, step VI – higher part). The resistant bacteria is favored due to its selective advantage of surviving the high antibiotic concentration during antibiotic treatment [203], but a selection for resistant bacteria has been also observed during sub inhibitory concentrations of tetracycline [204]. By comparing cultured strains from Norway and Arizona using core gene phylogenetic analysis, we found that the strain diversity was as high in Arizonan strains as in the Norwegian strains, thus no strain selection had occurred due to long-term tetracycline treatment (Paper 2, Figure 3). In *G. apicola*, this diversity is linked to differences in gene repertoires, thus different capabilities of nutrient breakdown: e.g. pectin and various sugars [46, 60]. This means that different strains may contribute to the host metabolism by production of nutrition elements that the host can utilize but also by detoxification of sugar metabolites. Hence, this diversity is important for maintaining host health and the strong host selection of core honeybee symbionts [43], and thus bacterial strains, might counteract outer low-impact selection pressures.

Spread and persistence of ARG in the GM

An alternative outcome to that only one strain overtakes the population is possible, which is that the ARB shares its ARG with the rest of the population, thus additionally contributing to preserving strain diversity (Figure 7, step VI – lower part). The *tet*(B) gene found in *G. apicola* was located on the Tn10 transposon in all strains (Paper2), which confirms that they are acquired and transferable [142]. Transposons can only spread form one genomic entity to another: e.g. from plasmid to the bacterial genome and vice versa, but we did not identify any plasmids in our cultured strains (Paper 2 & Paper 4), although plasmids containing *tet* genes have been identified in a metagenomics study (study of the microbiome) in the honeybee GM [172]. This indicate that there are plasmids circulating within this bacterial

population. The *tet*(B) gene is very similar across several bacterial genera and thus the origin of this gene is not easily identified, and only one aminoacid difference in the *tet*(B) between Norwegian and Arizonan *G. apicola* strains was identified in our data (Paper 2, Table S6). In addition, the nucleotide (nt) sequence of the Tn10 transposon was also conserved in al strains. However, there are variation in several nucleotides within the regulator area in the Tet B determinant (Paper 2, Figure 7), and we therefore used this region (Figure 5) to address spread of Tet B resistance within the *G. apicola* population. We found that the same Tet B variant could be identified in all strains across the *G. apicola* strain diversity (Paper 2, Figure 7), thus the alternative outcome is observed (Figure 7, step VI – lower part). The honeybee GM is a densely populated bacterial community that might well support HGT of ARGs through conjugation. HGT between members of this community has been documented for none-ARGs such as *rsh*-genes [43]. This spread was only detected in the Arizonan strains and since tetracycline exposure even at low concentrations can enhance transposition by the upregulation of the transposase genes [188, 189], this might explain our observation.

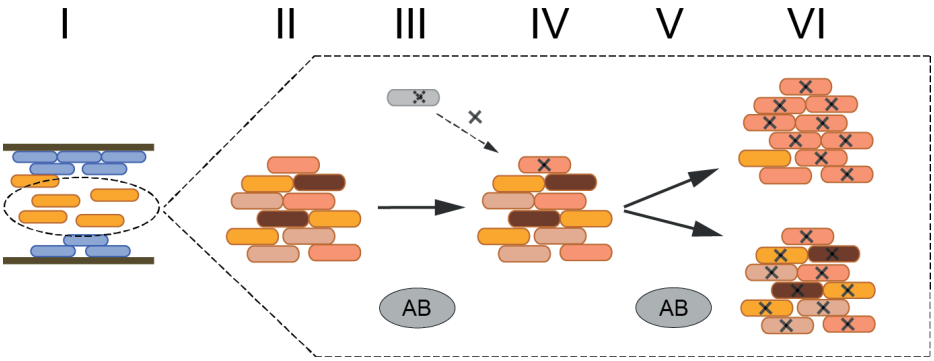

Figure 7: Main hypothesis (Paper 2) I: bacteria community within honeybee-gut intestine. II: by genome sequencing, several strains from one bacterium species, subgroups within this species can be identified (subgroups are shown with different colors). III & IV: antibiotic treatment enhances antibiotic-resistant genes (ARG) uptake into the population. Bacteria that harbour ARGs will survive and can sustain and proliferate. V: continuous exposure to antibiotics (selective pressure). VI: either the strain/subgroup with the ARG proliferates and dominates the population, or the ARG will be transferred to other subgroups by HGT. AB = antibiotic treatment, X = ARG.

The nt variation within the Tet B regulator region was county specific and in the Norwegian strains one variant was similar to others found in environmental and pathogenic *Proteobacteria*, while variants identified in the Arizonan strains where all unique for *G. apicola* (Paper 2, Figure 7, Table S10). The similarity towards environmental bacteria, could suggests a recent transfer event, while the uniqueness in the Arizonan strains might indicate an earlier acquisition. Taken together, these results suggests a model describing ARG spread to and within the GM of honeybees, wherein ARGs are acquired from environmental bacteria by honeybee gut symbionts in a geographically restricted manner and further shared between members of the symbiotic GM population. In this way, the functional diversity that different strains represent regarding nutrient breakdown is sustained, but the core microbiota is also more equipped, during antibiotic exposure, to compete against pathogens that might harbor ARGs [205]. The identification of a Tn*5393* transposon, with the *strA-strB* genes within *S. alvi*, that was identical to a transposon found in *E. coli*, and of which *strA-strB* genes were identical to those found in plant pathogens (Paper 4), further supports this model.

In our study, we identified only two *tet* genes in the Norwegian population, with *tet*(B) as the most prevalent and in congruence with previous findings, but with *tet*(H) also present at low prevalence (Paper 2, Figure 2a), which deviate from countries with similar exposure history, wherein no *tet*(H) has been reported in honeybees [172]. Interestingly and in contrast to *tet*(B) found on transposons, *tet*(H) was consistently identified within the genome of both *G. apicola* and *S. alvi* strains. The *tet*(H) gene has previously been identified in various Gram negative bacteria but in association with MGEs [185]. Therefore our finding of *tet*(H) within the genome only flanked by host genes, might represent a novel discovery. In addition, *tet*(H) also occurred to be present only within certain linages within *G. apicola* and *S. alvi*, thus consistent with a vertical transfer within the core microbiota population. Our qPCR data showed similar number of *tet*(H) genes identified in Norway and Arizona (Paper2, Figure 2b), which implies that tetracycline exposure does not influence the selection or spread of this resistance gene to the same extent as it does *tet*(B).

The occurrence of ARGs within MGEs in bacteria can infer a fitness cost due to extra energy that is needed to maintain and proliferate these extra genes during growth. Factors influencing the persistence of ARGs within a bacterium without selective pressure are low fitness costs and the high replicability of associated plasmids [162]. Also, if a long-term

association between the ARG and the MGE exists, such as for *tet*(B) and Tn10, this fitness costs can lessen due to the bacterial selection of low fitness cost MGEs [156]. Another way to reduce this cost is for the ARGs to become incorporated within the bacterial genome [156]. We identified *tet*(H) within the genome, which probably then explains why this gene can be maintained also in strains with additional *tet* genes (*tet*(B)). Thus, if a low fitness cost ARG gene is introduced it might persist in the population even without the antibiotic selection, as demostrated in the Norwegian honeybee population. Nevertheless, the fact that ARGs are found within the core microbiota, which undergoes strong host selection and is constantly shared between nest mates, is probably the main contribution to the persistence of these ARGs within the honeybee GM.

ARGs in insects

Only few studies have investigated the role of insects as a reservoir for ARGs and ARB [153] and consequently this represent a poorly understood ecosystem. However, one study identified ARB within the gut of houseflies and tracked the bacterium to fecal samples of pigs, demonstrating that insects indeed represent a link between different environments and that they can transfer ARB from one environment to another [153]. Honeybees acquire ARGs from the environment and these persist within the core microbiota (Paper 2). Honeybees have been shown to carry plant pathogens on their body, thus contributing to the spread of plant diseases [206, 207], which is also a likely scenario for ARB. Honeybees constantly interact with the environmental and human settings. Moreover, they are often transported large distances for pollination purposes, which makes honeybees a potent reservoir for large-scale ARG spread.

Other insect-related studies have mostly identified ARGs in environmental bacteria and not within the core microbiota as we have done in our studies (Paper 2 and 4). Many insects do not have as a stable core microbiota as honeybees do and their GM is highly diet dependent [129]. Although the honeybee core microbiota was also diet dependent (Paper 1, Figure 2a & Figure 3), we found the core members to be constitutively present regardless of diet or season, their relative abundance fluctuated depending on diet or season. Diet dependent fluctuations were also observed in other studies [69], and this variation might influence comparative honeybee GM studies. Moreover, ARGs, although some are found in a broad spectra of bacteria genera, some ARGs can be highly genera specific [185], which is also

acknowledged for *tet* genes found in the honeybee GM [172]. Other factors influencing honeybee GM composition is age and disease [9, 24, 27], and hence precise experimental design is needed when performing comparative honeybee GM studies, whether investigating GM composition or the presence of ARGs [185].

Taxonomy of *Gilliamella* species

The reported high strain diversity of *G. apicola* [45, 46, 60], was identified also in our study by comparative phylogeny of core genes (Paper 2, Figure 3). We further compared our strains to those isolated and genome sequenced in previous studies, which also comprised strains from different *Apis* and *Bombus* species (Paper 3). Our results were in reference with the phylogeny that was previously described for some of the strains [60], but additionally we compared the phylogeny across the existing diversity (Paper 3, Figure 1). This diversity might include more strains in the future as only parts of the *Apis* and *Bombus* species have been investigated and thus only a small part of all *Gilliamella* strains have been so far isolated and genome sequenced (~100 in NCBI as of 2017). Despite lacking this information, clear clusters of strains based on whether or not the strains come from *Apis* or *Bombus* species have emerged [60]. Within each of these two clusters, there are patterns showing bee species-specific strains, although not always consistent.

The 16S rRNA gene is widely used for identification of bacteria and inferring relations between them due to its unique properties as being both evolutionally conserved but at the same time reflecting evolutionary variation at the genus level. Though debated, for some genera, this gene do not have the resolution to separate species, based on the 97% similarity criteria and thus additional genotypic and phenotypic comparisons should be applied before determining new species [87]. One such genotypic measure is ANI, which is today widely used, and takes into the account differences in nucleotide sequence across all orthologue genes and sets the species limit to 94-96% similarity [89]. Based on this method we identified one group of *Gilliamella* strains that showed ~80% similarity to *G. apicola* (Paper 2, Figure 3). This was in line with being a new species, as the general interspecies similarity is 82% [208], a criterion today applied by NCBI for new bacteria submission. In line with being a new species, differential phenotypical traits were identified, and interestingly within the *Gilliamella* genus, the *tet*(H) gene was only associated with this new species. The *tet*(H) genes was also found within *S. alvi*, so it is not strictly species specific within the honeybee

GM, but the specificity observed within the *Gilliamella* genus calls for more in depth exploration. For the genus *Gilliamella*, different species of *Gilliamella* isolated from *Bombus* and *Apis* are as close as 98.7 % based on the 16S rRNA gene [61], and from our data (Paper 3, Table S1) we found that some strains of the newly proposed *G. apis* sp. nov., show 99.3 % 16S rRNA gene similarity towards the type strain *G. apicola* wkB1. This indicates that based even on a 16S rRNA gene cutoff of 99%, different *Gilliamella* species cannot be separated.

5 Concluding remarks and future perspectives

In this work, we investigated the prevalence, spread and persistence of acquired ARGs within the GM using honeybees as a model. We used WGS of bacterial isolates to address ARG associations with both MGEs and different bacteria species. Acquired Tc^r and Sm^r are found within the honeybee core microbiota species, either as transposon associated and/or within the bacteria genome. The honeybee core microbiotas relative composition fluctuates depending on foraging diet, while it remains stable for long time when on a steady diet, and this should be taken into consideration when performing detailed GM comparison studies in the future.

In this thesis, we have continued to unravel the diversity of the *Gilliamella* genus, which was first done in bumblebees [61], and showed that the claimed *G. apicola* diversity found within honeybees might not be as diverse after all. Phylogeny of core genes within the *Gilliamella* genus, revealed a cluster of strains with only ~80% similarity towards *G. apicola* and these strains were determined to belong to a new species with the proposed name *G. apis.* sp. nov., after additional phylogenetic analysis was performed. Based on the 16S rRNA gene the *Gilliamella* species cannot be separated even at the 99% cut-off level and this must be accounted for when and if one predicts/retrieves functional data based on metagenomics studies or 16S rRNA gene information., since these different *Gilliamella* species exhibit different functional capabilities. These differences in functional capacities are probably important for the honeybee GM dynamics and the health of the host. An intriguing question arises about whether these species share the same niche or occupy different niches that might be separated in space/ different gut parts. This question could not be addressed in this thesis since both the midgut and the pyloric region were analyzed together. Other studies of honeybee GM that have isolated and genome-sequenced strains that could be used for comparison also do not separate between different gut parts [58, 60]. Hence, this question remains for future exploration.

The complex interactions of factors contributing to spread and persistence of ARB and ARGs in a bacterial population makes it hard to certainly determine cause and effect in these matters even with the use of a low complex system like the honeybee GM. Nevertheless, our results have shed light on both spread and persistence of ARGs in the GM, were ARGs are persistent within the core microbiota by low cost maintenance in

bacterial genomes, and that ARB are not just selected for but that transposon associated ARGs are actually transferred between core microbiota species when under long-term tetracycline exposure. The potential role of honeybees to be a reservoir for ARG spread between different environments are highlighted in that honeybees interact with environmental bacteria, the honeybee gut can be a hotspot for HGT, and thus honeybees can spread ARB and ARG when transported large distances for pollination purposes.

We propose a model wherein ARGs are acquired by core microbiota members from environmental bacteria in a geographically restricted manner, and then shared within the GM, largely due to antibiotic exposure. Hence, reducing the unnecessary use of antibiotic in agricultural setting and in human medicine, where possible, is important as one means to hinder further spread of ARGs.

To explore this scenario further, an *in vivo* experimental set up can be initiated, where an Arizonan AR bacterium is introduced under low selection pressure to Norwegian honeybees and then investigate whether the bacteria will establish and proliferate or if the ARGs will be shared within the population. Colonization studies have shown that establishment of introduced strains is successful in germ free honeybees [67] and bumblebees [43], but not consistently successful when introduced into a normal GM [209]. Based on WGS information and that Tet B varies in sequence between Norwegian and Arizonan strains, both bacterial specific and Tet B specific primers can be designed, and the faith of introduced bacteria and ARGs can be monitored over time by qPCR. Thus excluding the need for biased bacterial culturing of trans conjugants. Moreover, use of sequences variants instead of OTUs might enable detection of colonization events by 16S rRNA gene sequencing, which would also unravel any compositional changes during antibiotic treatment, as seen in other honeybee GM studies [174, 175].

The honeybee GM as a model for studying GM dynamics and its ARG associations at the strain level seems promising, resulting in important information [210], which with good luck will enable us to fight bacterial AR, spread and persistence.

6 References

1. Engel, P. and N.A. Moran, *The gut microbiota of insects - diversity in structure and function.* FEMS Microbiol Rev, 2013. **37**(5): p. 699-735.
2. Kohl, K.D., M.D. Dearing, and S.R. Bordenstein, *Microbial communities exhibit host species distinguishability and phylosymbiosis along the length of the gastrointestinal tract.* Molecular Ecology: p. n/a-n/a.
3. Greenblum, S., R. Carr, and E. Borenstein, *Extensive strain-level copy-number variation across human gut microbiome species.* Cell, 2015. **160**(4): p. 583-94.
4. Dillon, R.J. and V.M. Dillon, *The gut bacteria of insects: nonpathogenic interactions.* Annu Rev Entomol, 2004. **49**: p. 71-92.
5. Koppel, N., V. Maini Rekdal, and E.P. Balskus, *Chemical transformation of xenobiotics by the human gut microbiota.* Science, 2017. **356**(6344).
6. Walter, J. and R. Ley, *The human gut microbiome: ecology and recent evolutionary changes.* Annu Rev Microbiol, 2011. **65**: p. 411-29.
7. De Filippis, F., et al., *High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome.* Gut, 2016. **65**(11): p. 1812-1821.
8. Nishida, A.H. and H. Ochman, *Rates of Gut Microbiome Divergence in Mammals.* Molecular Ecology: p. n/a-n/a.
9. Powell, J.E., et al., *Routes of acquisition of the gut microbiota of Apis mellifera.* Appl Environ Microbiol, 2014.
10. Rangberg, A., et al., *Paratransgenesis: an approach to improve colony health and molecular insight in honey bees (Apis mellifera)?* Integr Comp Biol, 2012. **52**(1): p. 89-99.
11. Matsuoka, K. and T. Kanai, *The gut microbiota and inflammatory bowel disease.* Seminars in Immunopathology, 2015. **37**: p. 47-55.
12. Liu, H., et al., *Heat Shock Proteins: Intestinal Gatekeepers that Are Influenced by Dietary Components and the Gut Microbiota.* Pathogens, 2014. **3**(1): p. 187-210.
13. Brown, E.M., M. Sadarangani, and B.B. Finlay, *The role of the immune system in governing host-microbe interactions in the intestine.* Nature Immunology, 2013. **14**: p. 660.
14. Zhang, C. and L. Zhao, *Strain-level dissection of the contribution of the gut microbiome to human metabolic disease.* Genome Med, 2016. **8**(1): p. 41.
15. Ley, R.E., et al., *Human gut microbes associated with obesity.* Nature, 2006. **444**: p. 1022.
16. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins.* Nature, 2009. **457**(7228): p. 480-4.
17. Forsythe, P., W. Kunze, and J. Bienenstock, *Moody microbes or fecal phrenology: what do we know about the microbiota-gut-brain axis?* BMC Med, 2016. **14**: p. 58.
18. Szczesniak, O., et al., *Isovaleric acid in stool correlates with human depression.* Nutr Neurosci, 2016. **19**(7): p. 279-83.
19. Hooper, L.V., *Commensal Host-Bacterial Relationships in the Gut.* Science, 2001. **292**(5519): p. 1115-1118.
20. Goodrich, J.K., et al., *Genetic Determinants of the Gut Microbiome in UK Twins.* Cell Host Microbe, 2016. **19**(5): p. 731-43.
21. Katsnelson, A., *Microbiome: The puzzle in a bee´s gut.* Nature, 2015. **521**: p. S56.
22. Davis, F.C., *The honey bee inside out*. 2004, Great Britain: Bee Craft Limited.

23. Kwong, W.K. and N.A. Moran, *Gut microbial communities of social bees.* Nat Rev Microbiol, 2016. **14**(6): p. 374-84.
24. Martinson, V.G., J. Moy, and N.A. Moran, *Establishment of characteristic gut bacteria during development of the honeybee worker.* Appl Environ Microbiol, 2012. **78**(8): p. 2830-40.
25. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body.* PLoS Biol, 2016. **14**(8): p. e1002533.
26. van Schaik, W., *The human gut resistome.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1670): p. 20140087.
27. Hamdi, C., et al., *Gut microbiome dysbiosis and honeybee health.* Journal of Applied Entomology, 2011. **135**(7): p. 524-533.
28. Vanengelsdorp, D. and M.D. Meixner, *A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them.* J Invertebr Pathol, 2010. **103 Suppl 1**: p. S80-95.
29. Hatjina, F., et al., *Population dynamics of European honey bee genotypes under different environmental conditions.* Journal of Apicultural Research, 2014. **53**(2): p. 233-247.
30. Münch, D., et al., *Obtaining Specimens with Slowed, Accelerated and Reversed Aging in the Honey Bee Model.* J Vis Exp, 2013(78).
31. Nelson, C.M., et al., *The Gene vitellogenin Has Multiple Coordinating Effects on Social Organization.* PLOS Biology, 2007. **5**(3): p. e62.
32. Toth, A.L., et al., *Nutritional status influences socially regulated foraging ontogeny in honey bees.* Journal of Experimental Biology, 2005. **208**: p. 4641-4649.
33. Koenig, J.E., et al., *Succession of microbial consortia in the developing infant gut microbiome.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4578-85.
34. Avershina, E., et al., *Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children.* FEMS Microbiol Ecol, 2014. **87**(1): p. 280-90.
35. Alvarez-Perez, S., C.M. Herrera, and C. de Vega, *Zooming-in on floral nectar: a first exploration of nectar-associated bacteria in wild plant communities.* FEMS Microbiol Ecol, 2012. **80**(3): p. 591-602.
36. Rokop, Z.P., M.A. Horton, and I.L. Newton, *Interactions between Cooccurring Lactic Acid Bacteria in Honey Bee Hives.* Appl Environ Microbiol, 2015. **81**(20): p. 7261-70.
37. Yatsunenko, T., et al., *Human gut microbiome viewed across age and geography.* Nature, 2012. **486**(7402): p. 222-7.
38. Palmer, C., et al., *Development of the Human Infant Intestinal Microbiota.* PLOS Biology, 2007. **5**(7): p. e177.
39. Sankar, S.A., et al., *The human gut microbiome, a taxonomic conundrum.* Syst Appl Microbiol, 2015. **38**(4): p. 276-86.
40. Avershina, E., et al., *Transition from infant- to adult-like gut microbiota.* Environ Microbiol, 2016. **18**(7): p. 2226-36.
41. Valles, Y., et al., *Microbial succession in the gut: directional trends of taxonomic and functional change in a birth cohort of Spanish infants.* PLoS Genet, 2014. **10**(6): p. e1004406.
42. Kwong, W.K., et al., *Dynamic microbiome evolution in social bees.* Sci Adv, 2017. **3**(3): p. e1600513.
43. Kwong, W.K., et al., *Genomics and host specialization of honey bee and bumble bee gut symbionts.* Proc Natl Acad Sci U S A, 2014. **111**(31): p. 11509-14.

44. Koch, H., et al., *Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees.* Mol Ecol, 2013.
45. Moran, N.A., et al., *Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees.* PLoS One, 2012. **7**(4): p. e36393.
46. Engel, P., V.G. Martinson, and N.A. Moran, *Functional diversity within the simple gut microbiota of the honey bee.* Proc Natl Acad Sci U S A, 2012. **109**(27): p. 11002-7.
47. Martinson, V.G., et al., *A simple and distinctive microbiota associated with honey bees and bumble bees.* Mol Ecol, 2011. **20**(3): p. 619-28.
48. Corby-Harris, V., P. Maes, and K.E. Anderson, *The Bacterial Communities Associated with Honey Bee (Apis mellifera) Foragers.* PLoS One, 2014. **9**(4).
49. Anderson, K.E., et al., *Ecological Succession in the Honey Bee Gut: Shift in Lactobacillus Strain Dominance During Early Adult Development.* Microb Ecol, 2016. **71**(4): p. 1008-19.
50. Faith, J.J., et al., *The long-term stability of the human gut microbiota.* Science, 2013. **341**(6141): p. 1237439.
51. Ley, R.E., D.A. Peterson, and J.I. Gordon, *Ecological and evolutionary forces shaping microbial diversity in the human intestine.* Cell, 2006. **124**(4): p. 837-48.
52. Lozupone, C.A., et al., *Diversity, stability and resilience of the human gut microbiota.* Nature, 2012. **489**(7415): p. 220-30.
53. De Filippo, C., et al., *Diet, Environments, and Gut Microbiota. A Preliminary Investigation in Children Living in Rural and Urban Burkina Faso and Italy.* Front Microbiol, 2017. **8**: p. 1979.
54. Sekelja, M., et al., *Unveiling an abundant core microbiota in the human adult colon by a phylogroup-independent searching approach.* ISME J, 2011. **5**(3): p. 519-31.
55. Gilliam, M. and D.K. Valentine, *Enterobacteriaceae isolated from foraging worker honey bees, Apis mellifera.* Journal of Invertebrate Pathology, 1974. **23**(1): p. 38-41.
56. Jeyaprakash, A., M.A. Hoy, and M.H. Allsopp, *Bacterial diversity in worker adults of Apis mellifera capensis and Apis mellifera scutellata (Insecta: Hymenoptera) assessed using 16S rRNA sequences.* J Invertebr Pathol, 2003. **84**(2): p. 96-103.
57. Ellegaard, K.M., et al., *Extensive intra-phylotype diversity in lactobacilli and bifidobacteria from the honeybee gut.* BMC Genomics, 2015. **16**: p. 284.
58. Engel, P., R. Stepanauskas, and N.A. Moran, *Hidden diversity in honey bee gut symbionts detected by single-cell genomics.* PLoS Genet, 2014. **10**(9): p. e1004596.
59. Kwong, W.K. and N.A. Moran *Cultivation and characterization of the gut symbionts of honey bees and bumble bees: Snodgrassella alvi gen. nov., sp. nov., a member of the Neisseriaceae family of the Betaproteobacteria; and Gilliamella apicola gen. nov., sp. nov., a member of Orbaceae fam. nov., Orbales ord. nov., a sister taxon to the Enterobacteriales order of the Gammaproteobacteria*. Int J Syst Evol Microbiol, 2012. DOI: 10.1099/ijs.0.044875-0.
60. Zheng, H., et al., *Metabolism of Toxic Sugars by Strains of the Bee Gut Symbiont Gilliamella apicola.* MBio, 2016. **7**(6).
61. Praet, J., et al., *Gilliamella intestini sp. nov., Gilliamella bombicola sp. nov., Gilliamella bombi sp. nov. and Gilliamella mensalis sp. nov.: Four novel Gilliamella species isolated from the bumblebee gut.* Syst Appl Microbiol, 2017. **40**(4): p. 199-204.
62. Engel, P., W.K. Kwong, and N.A. Moran, *Frischella perrara gen. nov., sp. nov., a gammaproteobacterium isolated from the gut of the honey bee, Apis mellifera.* Int J Syst Evol Microbiol, 2013.

63. Olofsson, T.C., et al., *Lactobacillus apinorum sp. nov., Lactobacillus mellifer sp. nov., Lactobacillus mellis sp. nov., Lactobacillus melliventris sp. nov., Lactobacillus kimbladii sp. nov., Lactobacillus helsingborgensis sp. nov. and Lactobacillus kullabergensis sp. nov., isolated from the honey stomach of the honeybee Apis mellifera.* Int J Syst Evol Microbiol, 2014. **64**(Pt 9): p. 3109-19.
64. Killer, J., et al., *Lactobacillus apis sp. nov., from the stomach of honeybees (Apis mellifera), having an in vitro inhibitory effect on the causative agents of American and European foulbrood.* Int J Syst Evol Microbiol, 2014. **64**(Pt 1): p. 152-7.
65. Kesnerova, L., R. Moritz, and P. Engel, *Bartonella apis sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria.* Int J Syst Evol Microbiol, 2016. **66**(1): p. 414-21.
66. Corby-Harris, V., et al., *Origin and effect of Alpha 2.2 Acetobacteraceae in honey bee larvae and description of Parasaccharibacter apium gen. nov., sp. nov.* Appl Environ Microbiol, 2014. **80**(24): p. 7460-72.
67. Kesnerova, L., et al., *Disentangling metabolic functions of bacteria in the honey bee gut.* PLoS Biol, 2017. **15**(12): p. e2003467.
68. David, L.A., et al., *Diet rapidly and reproducibly alters the human gut microbiome.* Nature, 2014. **505**(7484): p. 559-63.
69. Maes, P.W., et al., *Diet-related gut bacterial dysbiosis correlates with impaired development, increased mortality and Nosema disease in the honeybee (Apis mellifera).* Mol Ecol, 2016. **25**(21): p. 5439-5450.
70. Hornef, M., *Pathogens, Commensal Symbionts, and Pathobionts: Discovery and Functional Effects on the Host.* ILAR J, 2015. **56**(2): p. 159-62.
71. Engel, P., K.D. Bartlett, and N.A. Moran, *The Bacterium Frischella perrara Causes Scab Formation in the Gut of its Honeybee Host.* MBio, 2015. **6**(3): p. e00193-15.
72. Koch, H. and P. Schmid-Hempel, *Socially transmitted gut microbiota protect bumble bees against an intestinal parasite.* Proc Natl Acad Sci U S A, 2011. **108**(48): p. 19288-92.
73. Hubert, J., et al., *Changes in the Bacteriome of Honey Bees Associated with the Parasite Varroa destructor, and Pathogens Nosema and Lotmaria passim.* Microb Ecol, 2017. **73**(3): p. 685-698.
74. Schwarz, R.S., N.A. Moran, and J.D. Evans, *Early gut colonizers shape parasite susceptibility and microbiota composition in honey bee workers.* Proceedings of the National Academy of Sciences, 2016. **113**(33): p. 9345-9350.
75. Zheng, H., et al., *Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling.* Proc Natl Acad Sci U S A, 2017. **114**(18): p. 4775-4780.
76. Kwong, W.K., A.L. Mancenido, and N.A. Moran, *Immune system stimulation by the native gut microbiota of honey bees.* R Soc Open Sci, 2017. **4**(2): p. 170003.
77. Epstein, S.S., *The phenomenon of microbial uncultivability.* Curr Opin Microbiol, 2013. **16**(5): p. 636-42.
78. Allen-Vercoe, E., *Bringing the gut microbiota into focus through microbial culture: recent progress and future perspective.* Curr Opin Microbiol, 2013. **16**(5): p. 625-9.
79. Staley, J.T. and A. Konopka, *Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats.* Annu Rev Microbiol, 1985. **39**: p. 321-46.
80. Fenn, K., et al., *Quinones are growth factors for the human gut microbiota.* Microbiome, 2017. **5**(1): p. 161.
81. D'Onofrio, A., et al., *Siderophores from Neighboring Organisms Promote the Growth of Uncultured Bacteria.* Chemistry & Biology, 2010. **17**(3): p. 254-264.

82. Berdy, B., et al., *In situ cultivation of previously uncultivable microorganisms using the ichip.* Nature Protocols, 2017. **12**(10): p. 2232-2242.
83. Browne, H.P., et al., *Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation.* Nature, 2016. **533**: p. 543.
84. Levin, P.A. and E.R. Angert, *Small but Mighty: Cell Size and Bacteria.* Cold Spring Harb Perspect Biol, 2015. **7**(7): p. a019216.
85. Coico, R., *Gram Staining*, in *Current Protocols in Microbiology*. 2005, John Wiley & Sons, Inc.
86. Golding, C.G., et al., *The scanning electron microscope in microbiology and diagnosis of infectious disease.* Scientific Reports, 2016. **6**: p. 26516.
87. Rosselló-Mora, R. and R. Amann, *The species concept for prokaryotes.* FEMS Microbiology Reviews, 2001. **25**(1): p. 39-67.
88. Wayne, L.G., *International Committee on Systematic Bacteriology: announcement of the report of the ad hoc Committee on Reconciliation of Approaches to Bacterial Systematics.* Zentralbl Bakteriol Mikrobiol Hyg A, 1988. **268**(4): p. 433-4.
89. Goris, J., et al., *DNA-DNA hybridization values and their relationship to whole-genome sequence similarities.* Int J Syst Evol Microbiol, 2007. **57**.
90. Heather, J.M. and B. Chain, *The sequence of sequencers: The history of sequencing DNA.* Genomics, 2016. **107**(1): p. 1-8.
91. Woese, C.R., *Bacterial evolution.* Microbiological Reviews, 1987. **51**(2): p. 221-271.
92. Yarza, P., et al., *Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences.* Nat Rev Microbiol, 2014. **12**(9): p. 635-45.
93. Cao, Y., et al., *A Review on the Applications of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies.* Frontiers in Microbiology, 2017. **8**: p. 1829.
94. Lu, H., F. Giordano, and Z. Ning, *Oxford Nanopore MinION Sequencing and Genome Assembly.* Genomics, Proteomics & Bioinformatics, 2016. **14**(5): p. 265-279.
95. Wick, R.R., et al., *Completing bacterial genome assemblies with multiplex MinION sequencing.* Microb Genom, 2017. **3**(10): p. e000132.
96. Quail, M.A., et al., *A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.* Bmc Genomics, 2012. **13**: p. 341.
97. Sinclair, L., et al., *Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform.* PLoS One, 2015. **10**(2): p. e0116955.
98. Fadrosh, D.W., et al., *An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform.* Microbiome, 2014. **2**(1): p. 6.
99. Lloyd-Price, J., et al., *Strains, functions and dynamics in the expanded Human Microbiome Project.* Nature, 2017. **550**(7674): p. 61-66.
100. Mobegi, F.M., et al., *Deciphering the distance to antibiotic resistance for the pneumococcus using genome sequencing data.* Sci Rep, 2017. **7**: p. 42808.
101. Read, T.D. and R.C. Massey, *Characterizing the genetic basis of bacterial phenotypes using genome-wide association studies: a new direction for bacteriology.* Genome Medicine, 2014. **6**(11): p. 109.
102. Huber, W., et al., *Orchestrating high-throughput genomic analysis with Bioconductor.* Nature methods, 2015. **12**(2): p. 115-121.

103. Fox, J. and A. Leanage, *R and the Journal of Statistical Software.* Journal of Statistical Software, 2016. **73**(2).
104. Edgar, R.C., *UPARSE: highly accurate OTU sequences from microbial amplicon reads.* Nat Methods, 2013. **10**(10): p. 996-8.
105. Edgar, R.C., *UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing.* bioRxiv, 2016.
106. Callahan, B.J., et al., *DADA2: High-resolution sample inference from Illumina amplicon data.* Nature Methods, 2016. **13**: p. 581.
107. Thompson, L.R., et al., *A communal catalogue reveals Earth's multiscale microbial diversity.* Nature, 2017. **551**(7681): p. 457-463.
108. Callahan, B.J., P.J. McMurdie, and S.P. Holmes, *Exact sequence variants should replace operational taxonomic units in marker-gene data analysis.* The Isme Journal, 2017. **11**: p. 2639.
109. Ellegaard, K.M. and P. Engel, *Beyond 16S rRNA Community Profiling: Intra-Species Diversity in the Gut Microbiota.* Frontiers in Microbiology, 2016. **7**: p. 1475.
110. Medini, D., et al., *The microbial pan-genome.* Current Opinion in Genetics & Development, 2005. **15**(6): p. 589-594.
111. Baldauf, S.L., *Phylogeny for the faint of heart: a tutorial.* Trends Genet, 2003. **19**(6): p. 345-51.
112. Daniel, R.B., et al., *Quantitative Phylogenetic Analysis in the 21st Century Análisis Filogenéticos Cuantitativos en el siglo XXI.* Revista Mexicana de Biodiversidad, 2007. **78**(2): p. 225-252.
113. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.* Syst Biol, 2010. **59**(3): p. 307-21.
114. Konstantinidis, K.T. and J.M. Tiedje, *Genomic insights that advance the species definition for prokaryotes.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(7): p. 2567-2572.
115. Richter, M., et al., *JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison.* Bioinformatics, 2016. **32**(6): p. 929-31.
116. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
117. Kelley, D.R., M.C. Schatz, and S.L. Salzberg, *Quake: quality-aware detection and correction of sequencing errors.* Genome Biology, 2010. **11**(11): p. R116.
118. Zerbino, D.R. and E. Birney, *Velvet: algorithms for de novo short read assembly using de Bruijn graphs.* Genome Res, 2008. **18**(5): p. 821-9.
119. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* J Comput Biol, 2012. **19**.
120. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-9.
121. Aziz, R.K., et al., *The RAST Server: rapid annotations using subsystems technology.* BMC Genomics, 2008. **9**: p. 75.
122. Angiuoli, S.V., et al., *Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation.* Omics, 2008. **12**(2): p. 137-41.
123. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.
124. Piddock, L.J., *Assess drug-resistance phenotypes, not just genotypes.* Nat Microbiol, 2016. **1**(8): p. 16120.

125. Kostic, A.D., M.R. Howitt, and W.S. Garrett, *Exploring host-microbiota interactions in animal models and humans.* Genes Dev, 2013. **27**(7): p. 701-18.
126. Pascoe, E.L., et al., *Network analysis of gut microbiota literature: an overview of the research landscape in non-human animal studies.* ISME J, 2017. **11**(12): p. 2644-2651.
127. Rosshart, S.P., et al., *Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance.* Cell, 2017. **171**(5): p. 1015-1028 e13.
128. Seedorf, H., et al., *Bacteria from diverse habitats colonize and compete in the mouse gut.* Cell, 2014. **159**(2): p. 253-66.
129. Pernice, M., S.J. Simpson, and F. Ponton, *Towards an integrated understanding of gut microbiota using insects as model systems.* J Insect Physiol, 2014. **69**: p. 12-8.
130. Wong, A.C., et al., *Gut Microbiota Modifies Olfactory-Guided Microbial Preferences and Foraging Decisions in Drosophila.* Curr Biol, 2017. **27**(15): p. 2397-2404 e4.
131. Wong, A.C., J.M. Chaston, and A.E. Douglas, *The inconstant gut microbiota of Drosophila species revealed by 16S rRNA gene analysis.* Isme j, 2013. **7**(10): p. 1922-32.
132. The Honeybee Genome Sequencing, C., *Insights into social insects from the genome of the honeybee Apis mellifera.* Nature, 2006. **443**: p. 931.
133. Amdam, G.V., et al., *Altered physiology in worker honey bees (Hymenoptera: Apidae) infested with the mite Varroa destructor (Acari: Varroidae): a factor in colony loss during overwintering?* J Econ Entomol, 2004. **97**(3): p. 741-7.
134. Amdam, G.V., et al., *Hormonal control of the yolk precursor vitellogenin regulates immune function and longevity in honeybees.* Exp Gerontol, 2004. **39**(5): p. 767-73.
135. Munch, D., et al., *In the laboratory and during free-flight: old honey bees reveal learning and extinction deficits that mirror mammalian functional decline.* PLoS One, 2010. **5**(10): p. e13504.
136. Tan, S.Y. and Y. Tatsumura, *Alexander Fleming (1881-1955): Discoverer of penicillin.* Singapore Med J, 2015. **56**(7): p. 366-7.
137. von Nussbaum, F., et al., *Antibacterial Natural Products in Medicinal Chemistry—Exodus or Revival?* Angewandte Chemie International Edition, 2006. **45**(31): p. 5072-5129.
138. Mingeot-Leclercq, M.P., Y. Glupczynski, and P.M. Tulkens, *Aminoglycosides: activity and resistance.* Antimicrob Agents Chemother, 1999. **43**(4): p. 727-37.
139. Kohanski, M.A., D.J. Dwyer, and J.J. Collins, *How antibiotics kill bacteria: from targets to networks.* Nat Rev Microbiol, 2010. **8**(6): p. 423-35.
140. Andersson, D.I. and D. Hughes, *Microbiological effects of sublethal levels of antibiotics.* Nat Rev Microbiol, 2014. **12**(7): p. 465-78.
141. Crofts, T.S., A.J. Gasparrini, and G. Dantas, *Next-generation approaches to understand and combat the antibiotic resistome.* Nature reviews. Microbiology, 2017. **15**(7): p. 422-434.
142. van Hoek, A.H.A.M., et al., *Acquired Antibiotic Resistance Genes: An Overview.* Frontiers in Microbiology, 2011. **2**: p. 203.
143. Blair, J.M.A., et al., *Molecular mechanisms of antibiotic resistance.* Nature Reviews Microbiology, 2014. **13**: p. 42.
144. Handel, N., et al., *Interaction between mutations and regulation of gene expression during development of de novo antibiotic resistance.* Antimicrob Agents Chemother, 2014. **58**(8): p. 4371-9.
145. Aminov, R.I. and R.I. Mackie, *Evolution and ecology of antibiotic resistance genes.* FEMS Microbiol Lett, 2007. **271**(2): p. 147-61.

146. Marshall, B.M. and S.B. Levy, *Food animals and antimicrobials: impacts on human health.* Clin Microbiol Rev, 2011. **24**(4): p. 718-33.
147. Holmes, A.H., et al., *Understanding the mechanisms and drivers of antimicrobial resistance.* The Lancet, 2016. **387**(10014): p. 176-187.
148. Magiorakos, A.P., et al., *Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance.* Clin Microbiol Infect, 2012. **18**(3): p. 268-81.
149. Organization, T.W.H., *Antimicrobial resistance: global report on surveillance*. 2016.
150. Aarestrup, F.M., et al., *Associations between the use of antimicrobial agents for growth promotion and the occurrence of resistance among Enterococcus faecium from broilers and pigs in Denmark, Finland, and Norway.* Microb Drug Resist, 2000. **6**(1): p. 63-70.
151. Labro, M.T. and J.M. Bryskier, *Antibacterial resistance: an emerging 'zoonosis'?* Expert Rev Anti Infect Ther, 2014. **12**(12): p. 1441-61.
152. Katsunuma, Y., et al., *Associations between the use of antimicrobial agents for growth promotion and the occurrence of antimicrobial-resistant Escherichia coli and enterococci in the feces of livestock and livestock farmers in Japan.* J Gen Appl Microbiol, 2007. **53**(5): p. 273-9.
153. Zurek, L. and A. Ghosh, *Insects represent a link between food animal farms and the urban environment for antibiotic resistance traits.* Appl Environ Microbiol, 2014. **80**(12): p. 3562-7.
154. Ignasiak, K. and A. Maxwell, *Antibiotic-resistant bacteria in the guts of insects feeding on plants: prospects for discovering plant-derived antibiotics.* BMC Microbiology, 2017. **17**(1): p. 223.
155. Andersson, D.I. and D. Hughes, *Persistence of antibiotic resistance in bacterial populations.* FEMS Microbiol Rev, 2011. **35**(5): p. 901-11.
156. Bengtsson-Palme, J., E. Kristiansson, and D.G.J. Larsson, *Environmental factors influencing the development and spread of antibiotic resistance.* FEMS Microbiol Rev, 2018. **42**(1).
157. von Wintersdorff, C.J.H., et al., *Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer.* Frontiers in Microbiology, 2016. **7**: p. 173.
158. Bennett, P.M., *Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria.* British Journal of Pharmacology, 2008. **153**(Suppl 1): p. S347-S357.
159. Muñoz-López, M. and J.L. García-Pérez, *DNA Transposons: Nature and Applications in Genomics.* Current Genomics, 2010. **11**(2): p. 115-128.
160. Martínez Jé, L., *Natural Antibiotic Resistance and Contamination by Antibiotic Resistance Determinants: The Two Ages in the Evolution of Resistance to Antimicrobials.* Front Microbiol, 2012. **3**.
161. Martinez, J.L., et al., *Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems.* FEMS Microbiol Rev, 2009. **33**(2): p. 430-49.
162. Lopatkin, A.J., et al., *Persistence and reversal of plasmid-mediated antibiotic resistance.* Nature Communications, 2017. **8**(1): p. 1689.
163. Salyers, A.A., A. Gupta, and Y. Wang, *Human intestinal bacteria as reservoirs for antibiotic resistance genes.* Trends Microbiol, 2004. **12**(9): p. 412-6.
164. Davison, J., *Genetic Exchange between Bacteria in the Environment.* Plasmid, 1999. **42**(2): p. 73-91.

165. Huddleston, J.R., *Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes.* Infect Drug Resist, 2014. **7**: p. 167-76.
166. Smillie, C.S., et al., *Ecology drives a global network of gene exchange connecting the human microbiome.* Nature, 2011. **480**(7376): p. 241-4.
167. Ravi, A., et al., *The commensal infant gut meta-mobilome as a potential reservoir for persistent multidrug resistance integrons.* Scientific Reports, 2015. **5**: p. 15317.
168. Hu, Y., et al., *Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota.* Nat Commun, 2013. **4**: p. 2151.
169. Sommer, M.O., G. Dantas, and G.M. Church, *Functional characterization of the antibiotic resistance reservoir in the human microflora.* Science, 2009. **325**(5944): p. 1128-31.
170. Forslund, K., et al., *Country-specific antibiotic use practices impact the human gut resistome.* Genome Res, 2013. **23**(7): p. 1163-9.
171. Jernberg, C., et al., *Long-term impacts of antibiotic exposure on the human intestinal microbiota.* Microbiology, 2010. **156**(Pt 11): p. 3216-23.
172. Tian, B., et al., *Long-term exposure to antibiotics has caused accumulation of resistance determinants in the gut microbiota of honeybees.* MBio, 2012. **3**(6).
173. Levy, S.B. and B.M. Marshall, *Honeybees and tetracycline resistance.* MBio, 2013. **4**(1): p. e00045-13.
174. Raymann, K., Z. Shaffer, and N.A. Moran, *Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees.* PLOS Biology, 2017. **15**(3): p. e2001861.
175. Raymann, K., L.M. Bobay, and N.A. Moran, *Antibiotics reduce genetic diversity of core species in the honeybee gut microbiome.* Mol Ecol, 2017.
176. Francino, M.P., *Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances.* Frontiers in Microbiology, 2015. **6**: p. 1543.
177. Li, J.H., et al., *New evidence showing that the destruction of gut bacteria by antibiotic treatment could increase the honey bee's vulnerability to Nosema infection.* PLoS One, 2017. **12**(11): p. e0187505.
178. Lodesani, M. and C. Costa, *Limits of chemotherapy in beekeeping: development of resistance and the problem of residues.* Bee World, 2005. **86**(4): p. 102-109.
179. Allen, H.K., et al., *Call of the wild: antibiotic resistance genes in natural environments.* Nature Reviews Microbiology, 2010. **8**: p. 251.
180. Ramirez, M.S. and M.E. Tolmasky, *Aminoglycoside modifying enzymes.* Drug Resist Updat, 2010. **13**(6): p. 151-71.
181. Chopra, I. and M. Roberts, *Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance.* Microbiol Mol Biol Rev, 2001. **65**(2): p. 232-60 ; second page, table of contents.
182. Wang, N., et al., *Analysis of newly detected tetracycline resistance genes and their flanking sequences in human intestinal bifidobacteria.* Scientific Reports, 2017. **7**(1): p. 6267.
183. Wilkerson, C., et al., *Antibiotic Resistance and Distribution of Tetracycline Resistance Genes in Escherichia coli O157:H7 Isolates from Humans and Bovines.* Antimicrob Agents Chemother, 2004. **48**(3): p. 1066-7.
184. Roberts, M.C., *Update on acquired tetracycline resistance genes.* FEMS Microbiol Lett, 2005. **245**(2): p. 195-203.
185. Roberts, M.C. and S. Schwarz, *Tetracycline and Phenicol Resistance Genes and Mechanisms: Importance for Agriculture, the Environment, and Humans.* J Environ Qual, 2016. **45**(2): p. 576-92.

186. Levy, S.B., et al., *Nomenclature for New Tetracycline Resistance Determinants.* Antimicrob Agents Chemother, 1999. **43**(6): p. 1523-4.
187. Nesme, J., et al., *Large-scale metagenomic-based study of antibiotic resistance in the environment.* Curr Biol, 2014. **24**(10): p. 1096-100.
188. Whittle, G., N.B. Shoemaker, and A.A. Salyers, *The role of Bacteroides conjugative transposons in the dissemination of antibiotic resistance genes.* Cell Mol Life Sci, 2002. **59**(12): p. 2044-54.
189. Doucet-Populaire, F., et al., *Inducible transfer of conjugative transposon Tn1545 from Enterococcus faecalis to Listeria monocytogenes in the digestive tracts of gnotobiotic mice.* Antimicrob Agents Chemother, 1991. **35**(1): p. 185-7.
190. Shaw, K.J., et al., *Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes.* Microbiol Rev, 1993. **57**(1): p. 138-63.
191. Pezzella, C., et al., *Tetracycline and Streptomycin Resistance Genes, Transposons, and Plasmids in Salmonella enterica Isolates from Animals in Italy.* Antimicrobial Agents and Chemotherapy, 2004. **48**(3): p. 903-908.
192. Sundin, G.W. and C.L. Bender, *Dissemination of the strA-strB streptomycin-resistance genes among commensal and pathogenic bacteria from humans, animals, and plants.* Mol Ecol, 1996. **5**(1): p. 133-43.
193. L'Abée-Lund, T.M. and H. Sørum, *Functional Tn5393-Like Transposon in the R Plasmid pRAS2 from the Fish Pathogen Aeromonas salmonicida subspecies salmonicida Isolated in Norway.* Applied and Environmental Microbiology, 2000. **66**(12): p. 5533-5535.
194. Sundin, G.W., D.E. Monks, and C.L. Bender, *Distribution of the streptomycin-resistance transposon Tn5393 among phylloplane and soil bacteria from managed agricultural habitats.* Can J Microbiol, 1995. **41**(9): p. 792-9.
195. Sundin, G.W. and C.L. Bender, *Expression of the strA-strB streptomycin resistance genes in Pseudomonas syringae and Xanthomonas campestris and characterization of IS6100 in X. campestris.* Appl Environ Microbiol, 1995. **61**(8): p. 2891-7.
196. Petrova, M.A., et al., *Association of the strA-strB genes with plasmids and transposons in the present-day bacteria and in bacterial strains from permafrost.* Russian Journal of Genetics, 2008. **44**(9): p. 1116-1120.
197. Sunde, M. and M. Norstrom, *The genetic background for streptomycin resistance in Escherichia coli influences the distribution of MICs.* J Antimicrob Chemother, 2005. **56**(1): p. 87-90.
198. Liu, B. and M. Pop, *ARDB—Antibiotic Resistance Genes Database.* Nucleic Acids Research, 2009. **37**(Database issue): p. D443-D447.
199. Zankari, E., et al., *Identification of acquired antimicrobial resistance genes.* Journal of Antimicrobial Chemotherapy, 2012. **67**(11): p. 2640-2644.
200. Wattam, A.R., et al., *Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center.* Nucleic Acids Research, 2017. **45**(Database issue): p. D535-D542.
201. Lehtinen, S., et al., *Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage.* Proceedings of the National Academy of Sciences of the United States of America, 2017. **114**(5): p. 1075-1080.
202. McManus, P.S., et al., *Antibiotic use in plant agriculture.* Annu Rev Phytopathol, 2002. **40**: p. 443-65.
203. Kirkup, B.C., *Bacterial Strain Diversity Within Wounds.* Advances in Wound Care, 2015. **4**(1): p. 12-23.

204. McVicker, G., et al., *Clonal Expansion during Staphylococcus aureus Infection Dynamics Reveals the Effect of Antibiotic Intervention.* PLoS Pathogens, 2014. **10**(2): p. e1003959.
205. Levy, S.B. and B.M. Marshall *Honeybees and tetracycline resistance*. MBio, 2013. **4**, DOI: 10.1128/mBio.00045-13.
206. Alexandrova, M., et al., *Erwinia amylovora longevity in beehives, beehive products and honeybees.* Acta Horticult 2002. **590**: p. 201–205.
207. Pattemore, D.E., et al., *Evidence of the role of honey bees (Apis mellifera) as vectors of the bacterial plant pathogen Pseudomonas syringae.* Australasian Plant Pathology, 2014. **43**(5): p. 571-575.
208. Qin, Q.-L., et al., *A Proposed Genus Boundary for the Prokaryotes Based on Genomic Insights.* Journal of Bacteriology, 2014. **196**(12): p. 2210-2215.
209. Billiet, A., et al., *Effect of oral administration of lactic acid bacteria on colony performance and gut microbiota in indoor-reared bumblebees (Bombus terrestris).* Apidologie, 2016. **48**(1): p. 41-50.
210. Mather, A.E., et al., *Distinguishable epidemics of multidrug-resistant Salmonella Typhimurium DT104 in different hosts.* Science, 2013. **341**(6153): p. 1514-7.

Microbes Environ. Vol. 30, No. 3, 235-244, 2015
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME15019

MICROBES AND ENVIRONMENTS

Shifts in the Midgut/Pyloric Microbiota Composition within a Honey Bee Apiary throughout a Season

Jane Ludvigsen[1]*, Anbjørg Rangberg[1], Ekaterina Avershina[1], Monika Sekelja[1,2], Claus Kreibich[1], Gro Amdam[1,3], and Knut Rudi[1]

[1]*Norwegian University of Life Sciences, Chemistry, Biotechnology and Food science department, Campus Ås, Ås 1432, Norway;* [2]*Genetic Analysis P.O. Box 4239, Nydalen, 0401 Oslo, Norway; and* [3]*School of Life Sciences, Arizona State University, P.O. Box 874501, 427 East Tyler Mall, Tempe, AZ 85287, USA*

(Received February 9, 2015—Accepted June 22, 2015—Published online September 1, 2015)

Honey bees (*Apis mellifera*) are prominent crop pollinators and are, thus, important for effective food production. The honey bee gut microbiota is mainly host specific, with only a few species being shared with other insects. It currently remains unclear how environmental/dietary conditions affect the microbiota within a honey bee population over time. Therefore, the aim of the present study was to characterize the composition of the midgut/pyloric microbiota of a honey bee apiary throughout a season. The rationale for investigating the midgut/pyloric microbiota is its dynamic nature. Monthly sampling of a demographic homogenous population of bees was performed between May and October, with concordant recording of the honey bee diet. Mixed Sanger-and Illumina 16S rRNA gene sequencing in combination with a quantitative PCR analysis were used to determine the bacterial composition. A marked increase in α-diversity was detected between May and June. Furthermore, we found that four distinct phylotypes belonging to the *Proteobacteria* dominated the microbiota, and these displayed major shifts throughout the season. *Gilliamella apicola* dominated the composition early on, and *Snodgrassella alvi* began to dominate when the other bacteria declined to an absolute low in October. *In vitro* co-culturing revealed that *G. apicola* suppressed *S. alvi*. No shift was detected in the composition of the microbiota under stable environment/dietary conditions between November and February. Therefore, environmental/dietary changes may trigger the shifts observed in the honey bee midgut/pyloric microbiota throughout a season.

Key words: honey bee workers, midgut/pyloric microbiota, seasonal changes, *Gilliamella apicola*, *Snodgrassella alvi*

Honey bees (*Apis mellifera*) are important crop pollinators and are widely used around the world in agriculture and food production (51). The honey bee is a social insect that harbors a core gut microbiota of eight abundant phylotypes, which accounts for ~95% of all gut bacteria (34, 37). The distinct and relatively simple gut microbiota is transmitted from adult to newly hatched bees through feeding and secretion inside the colony (29). The honey bee gut microbiota is distributed throughout the entire digestive tract, in which the midgut harbors approximately 1–4% and the ileum/rectum over 90% of the most dominant bacteria found in honey bees (35). Several studies have shown the importance of gut symbionts in bee health and disease (20, 22, 28, 52, 53). In honey bees, the digestion of nutrients takes place in the midgut (10, 11, 13), and is aided by gut associated microbes. A recent study reported that the honey bee gut symbiont *G. apicola* degraded pectin (19), a main component of pollen, which is difficult to break down by the host (11).

Worker honey bees are functionally sterile helpers that perform within-nest tasks and forage. Foraging workers encounter different plants and flowers during the season, and bring back nectar and pollen to the colony. In Norway, worker honey bees actively forage between April and September. Colony food intake is a function of local environmental conditions, including temperature and precipitation. Foragers visit the plants that are available at that time of the year, and the resulting pollen and nectar is consumed or stored by the colony (13, 23, 40, 43). In September, beekeepers feed the bees a sugar mix as a replacement for the honey that is harvested. Bees cluster inside the hive between September and April due to the cold weather conditions in Norway (43) (www.norges-birokterlag.no). Honey bee colonies are active in thermo-regulation throughout the Nordic winter, and maintain core temperatures of approximately 20°C, in contrast to an in-hive temperature of approximately 30°C during the foraging season (www.norges-birokterlag.no and www.stadevægt.dk). Energy for heat production comes from the consumption of stored sugars as their winter diet. In the spring, the colony has a high demand for protein due to increased levels of brood rearing. Pollen is the primary amino acid and lipid source for honey bees, and pollen foraging is required for rebuilding colony strength through the production of new bees during the springtime (13). In addition to amino acids and lipids, pollen provides vitamins and minerals, whereas nectar is the primary carbohydrate source for the colony.

Throughout the foraging season, honey bees acquire a transient set of gut bacteria that are horizontally transmitted from the environment surrounding the colony (2). Previous studies have suggested differences in the gut microbiota composition between colonies at different locations, between colonies at the same location, and between individual bees within a colony (16, 17, 26, 34, 37). Another recent study by Corby-Harris, *et al.* (9), which performed repeated sampling of foragers, did not find any significant differences in the core

* Corresponding author. E-mail: jane.ludvigsen@nmbu.no;
Tel: +47–6723–2517.

gut microbiota composition at two different points in time (fall and spring) at one specific location. Therefore, some findings indicate a highly variable gut microbiota in honey bees, whereas others suggest a more stable gut microbiota. This potential disagreement warrants further studies.

It currently remains unclear how the microbiota of the honey bee gut responds to environmental fluctuations and dietary changes throughout a season. The gut is also spatially complex. Previous studies established that the honey bee ileum contains a stable microbiota (35, 41), while the midgut and pylorus is highly dynamic due to its peritrophic membrane and refluxes from the ileum (48). We hypothesized that environmental/dietary changes have a considerable impact on the microbiota in this dynamic part of the honey bee gut. As a first pass to address this hypothesis, we examined microbial changes in samples of the midgut/pylorus obtained from bees living in commercial hives during a season. At our location, commercial hives are subject to marked changes in diet and environment. Samples from the active foraging season were compared to the microbiota of clustering bees living in a stable winter environment at the same location. We used high throughput 16S rRNA gene sequencing (*i.e.*, both mixed Sanger- and Illumina sequencing) in combination with quantitative PCR (qPCR) to determine the microbiota composition of the samples. In addition, we performed culturing and *in vitro* competition experiments in order to address questions regarding the biology of select midgut/pyloric bacteria. The results obtained suggest that diet/environment is important for shaping the midgut/pyloric microbiota composition, and contribute to its dynamic nature.

Materials and Methods

Bee sampling

In order to perform the bee midgut/pyloric microbiota analysis throughout a foraging season, bees were sampled between May and October 2012 at the Norwegian University of Life Sciences, Ås, Norway. Information regarding available nutrition for foraging bees is summarized in Table 1. Ten worker bees from three separate colonies, for a total of 30 bees per month, were collected. The bees were picked at random with tweezers from three of the ten removable combs in the brood chamber box of each colony (one comb in the front, one in the middle, and one at the back). This sampling method had the following benefits: i) it was easy to repeat between colonies, and ii) it ensured that the bees collected were unlikely to fall into the same behavioral group. Therefore, we obtained samples that were similar between colonies and represented the diverse worker populations of those hives (13). The ten bees from the three different colonies were sampled together, randomized, and then analyzed, and the results were averaged across hives, thereby giving data that represented the biological gut bacterial composition in a population at one location. This was performed for all sampling times, except September, in which samples from different colonies were analyzed separately to obtain a snapshot of possible colony differences.

Table 1. Diet of honey bees throughout a foraging season

Sample time point	Diet
May	Dandelion, maple, and fruit trees
June	Raspberries
July	Raspberries & canola
August	Heather & honeydew
September	Sugar mix[1] (1 week[2])
October	Sugar mix[1] (1 month[3])

[1] Sugar mix composition: 37% sucrose, 19% glucose, 19% fructose, and 25% water.
[2] Bees had been fed this sugar mix for one week by the sampling time point.
[3] Bees had been fed this sugar mix for one month by the sampling time point.

An additional 30 bees were collected (ten from each of the three colonies) to facilitate the culturing of midgut/pyloric bacteria. Moreover, ten bees (three bees from two colonies and four from one colony) were separately collected and used as an average sample (hereafter referred to as the average July sample) in Illumina MiSeq sequencing, analyzing the 16S rRNA microbiota composition, as well as a control for the DNA extraction procedure's technical variation. All bees were collected in July 2012 from the same three colonies as those used in the microbiota analysis seasonal study.

In order to analyze midgut/pyloric microbiota under stable nutritional conditions, worker bees were collected from two colonies in November 2011 (24 bees) and February 2012 (30 bees), which had been fed a commercial sugar mix (37% sucrose, 19% glucose, 19% fructose and 25% water, Nordic Sugar A/S, Denmark) for one month and four months, respectively. These colonies were restricted to their hive due to cold weather, and, thus, their environment was stable/constant throughout the sampling period.

Midgut isolation and DNA extraction

Bees were anesthetized on ice directly after sampling (1), and washed in 50% ethanol before dissection. The whole gut was dissected out using a sterile dissecting forceps, and the stinger was pulled out as the bee was held by the head. By pulling the stinger, the intact digestive tract followed, separating the midgut from the crop, which remained in the bee. An illustration of the dissecting procedure is shown in Fig. S1. The midgut/pylorus was collected in micro tubes (Sarstedt, Germany) containing 0.2 g <106 μm acid-washed glass beads (Sigma-Aldrich, USA) and 500 μL S.T.A.R buffer (Roche, Switzerland) by making a cut with a sterile dissecting scissor within the pylorus part of the digestive tract (Fig. S1).

Mechanical lysis was performed using MagNA Lyzer (Roche); 6,500 rpm for 20 s for 2×; 1 min cooling at 4°C between runs. An automated DNA magnetic bead-based extraction method was used on all samples (developed by Genetic Analysis; http://www.genet-analysis.com). A Quant-iT PicoGreen dsDNA assay (Life Technologies, USA) was used for quantification of the extracted DNA (45).

Microbiota analyzes

Mixed sequencing

PCR was used to amplify approximately 450 bases of bacterial 16S rRNA genes in each midgut/pylorus sample using universal 16S rRNA primers; Forward-F11 5′-TCCTACGGGAGGCAGCAGT-3′, Reverse-A01 5′-GGACTACCAGGGTATCTAATCCTGTT-3′ as previously described (38). PCR was performed with HOT FIREpol DNA polymerase (Solis BioDyne, Estonia) in a final volume of 25 μL. Cycling conditions for PCR: Activation 95°C for 15 min, and 30 cycles of 95°C for 30 s, 60°C for 30 s, 72°C for 60 s, Final elongation 72°C for 7 min. A Quant-iT PicoGreen dsDNA assay (Life Technologies) was used to confirm successful PCR amplification, and the correct amplicon size was confirmed on 1% agarose gel.

Sanger sequencing was performed using the BigDye Terminator v1.1 Cycle Sequencing Kit (Life Technologies) on Exo1-treated PCR products in a total volume of 10 μL. Agencourt CleanSEQ Dye-terminator Removal (Beckman Coulter, USA) was used to purify the end-labeled sequences, and all Sanger sequencing was performed at Hedmark University College on a 3130 xl Genetic analyzer (Life Technologies).

The mixed Sanger sequencing method was first applied by Trosvik, *et al.* (49), and, when used in combination with multivariate curve resolution with an alternating least squares analysis (MCR-ALS) (56), mixed bacteria communities can be analyzed (3, 44). The method first aligned all sequence spectra. The co-occurrence of

the different parts of the spectra were then determined with the co-occurring nucleotides that were displayed as the component sequence. During the MCR-ALS analysis, a quality filtering of sequence data was performed, and sequences with low quality were removed. The bacterial components were subsequently base-called, and taxonomy assigned using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov). The relative ratio of the dominant bacterial components in each sample was calculated in the MCR-ALS analysis without assuming closure (not adding up to 100% due to residual noise in each sample). Matlab (MathWorks, USA) was used to determine the mean α-diversity (modified Simpson index). The α-diversity, as explained previously (3), was calculated from the raw aligned spectra from each individual sample and based on the ratio of nucleotide mixing at each position in the sequences. The α-diversity calculation method used in this study was unique to the mixed sequencing approach, in which the values could not be directly compared across different methods.

Quantitative PCR

The relative quantity of 16S rRNA genes (bacterial load) was determined through ratio calculations between universal 16S rRNA genes and the honey bee specific vitellogenin gene, both retrieved from qPCR. The vitellogenin primers amplified 150 bp of the vitellogenin gene (DNA), which encodes a yolk precursor protein abundant in blood (1). Both qPCR reactions were performed on the LightCycler 480 II (Roche), and the raw data were imported into the LinReg PCR program for CT values and PCR efficiency calculations (42). The log relative bacteria/bee DNA ratio was calculated (30, 46) using the following formula: =(log[efficiency Bee]×CT Bee)–(log[efficiency Bacteria]×CT Bacteria).

EvaGreen (25, 33), as a fluorescent marker, was used for the vitellogenin qPCR with 5× HOT FIREPol EvaGreen qPCR Mix Plus (Solis Bio Dyne) at a final concentration of 1×. One microliter of diluted gDNA (1:2) was added to the working solution for a final volume of 20 µL. The determination of the dilution factor for qPCR was based on a dilution series experiment conducted prior to the analysis. Cycling conditions: Activation 95°C for 15 min, 40 cycles of; 95°C for 30 s, 54°C for 45 s, 72°C for 30 s. Primers for vitellogenin: Forward 5′-GTTGGAGAGCAACATGCAGA-3′, Reverse 5′-TCGATCCATTCCTTGATGGT-3′ were used (1). A high-resolution melting (HRM) analysis was performed to confirm the specificity of the PCR primers, and a positive control (Bee brain-DNA) and negative control (Nuclease-free water) were added to each run. DNA from the honey bee brain was extracted from one sterile dissected bee brain using the DNeasy Blood and Tissue kit (Qiagen, USA), and its concentration was measured using an ND-1000 spectrophotometer (Thermo Scientific, USA). Universal 16S rRNA qPCR was performed using the TaqMan probe as a fluorescent marker (38) and a positive control (*Lactobacillus kunkeii*) and negative control (Nuclease-free water) were added to each run. The same primer pair as that used for mixed sequencing was applied, and we used 5× HOT FIREPol Probe qPCR Mix Plus (Solis BioDyne) in a 1× concentration, with 1 µL of diluted gDNA (1:2) at a final volume of 20 µL, with the following cycling: Activation 95°C for 15 min and 40 cycles of 95°C for 30 s, 60°C for 60 s.

Illumina sequencing

The sample used for this analysis was collected in July 2012 (hereafter referred to as the average July sample), and originated from the same three colonies as those used in the seasonal study. This sample consisted of ten bees, the midgut/pylorus of which were sampled and mixed together into one tube, crushed, and diluted with S.T.A.R buffer. This mixture was then made into aliquots to equal the amount of one sample and represented an average measurement in July 2012. The analysis of this sample was performed in triplicate with tagged-universal PRK primers targeting the 16S rRNA gene (PRK314F and PRK806R) (55), as described in (39), for initial PCR and then pooled before sequencing, after quantification with a Quant-iT PicoGreen dsDNA assay and normalization. The amplicon size was approximately 590 bp and indexing corresponding to the Illumina TruSeq LT set-up was used. Samples were sent to the University of Oslo for 250 bp paired-end MiSeq sequencing (Illumina, USA). The retrieved data were analyzed using the Quantitative Insights Into Microbial Ecology (QIIME) pipeline (6). Regarding OTU classification, the forward reads were quality filtered and clustered at a 99% homology level using a closed-reference *uclust* search against the Greengenes database (15). Additionally, the random selection of 1,000 paired-end reads was assigned taxonomic nomenclature using BLAST with a cut-off of 95% identity.

Technical controls

A positive extraction control was included in each plate run to address potential extraction procedure bias. We used the average July sample for this purpose. Because the bees throughout the season were collected at different time points, DNA was extracted in three turns on a 96-well plate. DNA from the average July sample was extracted in duplicate for each DNA extraction/plate, and this plate setup was maintained throughout the experiment and then used in the subsequent analysis of mixed sequencing and qPCR analysis. The three extraction plates each included the extraction control sample in two replicates, which then resulted in six measurements for this sample. In addition, a non-template control (elution buffer) was included in each run. Both controls followed the same workflow as the experimental samples, and this procedure ensured the detection of potential methodology bias.

Bacterium isolation and taxonomy assignment

We sampled a separate set of 30 bees in July 2012 from the same three colonies as those for the seasonal dataset and Illumina sequencing. Ten midguts/pylori were pooled in one micro tube (Sarstedt), and prepared with 500 µL 1× phosphate buffered saline (PBS) and 15% glycerol. The gut parts were frozen at –80°C before culturing. Frozen gut parts from one micro tube were homogenized and spread in parallel on Tryptic soy agar plates (TSA) (Merck KGaA, Germany) with 5% horse blood (hereafter referred to as blood agar plates; ThermoFisher Scientific, USA). The blood agar plates were incubated for two d at 37°C in a CO_2- enriched atmosphere (GasPack EZ CO_2 container system; Becton Dickinson [BD], USA) (31). Controls for sterility and the correct atmosphere were included. Colonies were randomly picked and discriminated by different morphologies and then repeatedly streaked on new blood agar plates to ensure pure cultures.

DNA extraction from the bacterial isolates was performed as previously described in this study. Universal bacteria CoverAll primers (developed by Genetic Analysis and publically available through purchase) were used for the 16S rRNA gene amplification (amplicon about 1,200 bp, targeting V3–V9) of the bacteria isolates. We used the HOT FIREpol DNA polymerase as previously described with the following cycling conditions: Activation 95°C for 15 min and 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 1 min and 20 s. Sequencing was performed as previously described. The Sanger sequences were processed with the use of CLC Main work bench 6 (CLCbio, Qiagen). Sequencing was performed using both the forward and reverse CoverAll primers, and consensus sequences were assembled from the two complimentary sequences derived from each bacterium and matched in BLAST. Sequence taxonomy was assigned with hits of more than 99% matches. Nearest identity BLAST hits were chosen as reference sequences for all isolates and the mixed sequencing components, and reference sequences for the Illumina sequencing BLAST search were also added. The sequences were then aligned and manually curated and a neighbor-joining phylogenetic tree (jukes cantor algorithm) with bootstrapping (100) was created in CLC main workbench 6.

Competition experiment

Two of the isolates (*Gilliamella apicola* and *Snodgrassella alvi*) were selected for a competition experiment to determine whether

they exhibited symbiotic, mutualistic, or competitive characteristics when grown together. The two strains were selected after screening the isolates with specific primers to yield positive amplification. The primer pair Gamma1-459-qtF 5′-GTATCTAATAGGTGCATCAA TT-3′ and Gamma1-648-qtR 5′-TCCTCTACAATACTCTAGTT-3′ was used to detect *G. apicola*, while the primer pair Beta-1009-qtF 5′-CTTAGAGATAGGAGAGTG-3′ and Beta-1115-qtR 5′-TAAT GATGGCAACTAATGACAA-3′ was used to detect *S. alvi*, as described previously (35). Both isolates were grown alone for one d in tryptic soy broth (TSB) (Merck KGaA) (31) in 1.5-mL Eppendorf tubes prepared with sterile 0.2 g <106 μm acid-washed glass beads with a starting amount of 1 μL swabbed from a blood agar plate dissolved in 100 μL TSB. All 100 μL were mixed with the respective bacterium and they were then grown together and alone under the same condition as described above for two more d with additional TSB in a total volume of 1.5 mL. The experiment was performed in triplicate, and a negative control (only TSB) was added. The bacteria were lysed, and DNA was extracted for qPCR quantification of the bacterial load in each sample on LightCycler 480 II. qPCR was performed on all samples in one run, in duplicate, using 5× HOT FIREPol EvaGreen qPCR Mix Plus at a final concentration of 1× with the following cycling conditions: Activation 95°C for 15 min and 40 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s. The products (amplicon size; *G. apicola* 210 bp, *S. alvi* 128 bp) were verified by a HRM analysis and on a 1.5% agarose gel. qPCR raw data were processed using the LinReg PCR program as previously described, and the ratio between the same bacterium grown alone and together with the other bacterium was calculated.

Statistical analysis

Each point in time consisted of data measurements from single bees that were added and averaged. The standard deviation (SD) and standard error of the mean (SEM) were determined for all points in time. We performed a one-way ANOVA to test the complete seasonal trend for each bacterium. The same statistical method was applied for the qPCR results and α-diversity analysis. We used the Tukey HSD test to test the significant difference among different points in time. In addition, we performed two-sided Student's *t*-tests for comparisons between bee colonies and to address technical variations. Corrections for multiple testing (Bonferroni) were performed by dividing the selected *p*-value on the numbers of *t*-tests performed in each analysis.

Accession numbers

Sanger-sequences were deposited in GenBank under accession numbers KM454389–KM454422.

Results

Overall microbiota composition

We used Illumina sequencing, mixed sequencing, and culturing to determine the overall composition of the microbiota in our dataset.

Illumina sequencing of the average July sample gave a total of 188,189 reads after quality filtering. QIIME analyses showed low diversity (Fig. S2, Table S1). Unfortunately, most likely due to a lack of bee-associated bacteria in the Greengenes database, we could not obtain accurate taxonomic assignments by QIIME. Therefore, we performed an in-depth BLAST search of 1,000 random Illumina sequences. These analyses showed that the five most abundant bacteria identified were: *Tatumella* sp. 40% (which gave an equal number of hits on both *Tatumella ptyseos* and *Tatumella terrea*), *G. apicola* 23%, *S. alvi* 16%, *Frischella perrara* 12%, and *Lactobacillus kalixensis* 6%.

The cultivation of midgut/pylorus samples resulted in 34 bacteria isolates, from five different bacterial phyla, which clustered with previously identified gut bacteria from honey bees (Fig. 1).

Seven main bacterial components were identified by mixed sequencing (Table 2 and Fig. S3). These bacterial components were matched with both Illumina sequencing and cultured isolate sequences, representing most of the characterized diversity (Fig. 1).

Four of these bacterial components were identified in the seasonal dataset and were taxonomically assigned as *F. perrara*, *G. apicola*, *S. alvi*, and one component, which had the closest % identity to the *Enterobacteriaceae* family. Three additional components were identified in the stable environment dataset: *Acetobacteraceae*, *Rhizobiales*, and *Lactobacillus*.

Seasonal trends in population composition

Major changes in the midgut/pyloric bacterial relative abundance were evident between May and October, and the calculated Tukey HSD test *p*-values gave significance on the 95% and 99% levels (Fig. 2a, Table S2). *G. apicola* showed high dominance early on in the season, but its relative abundance declined between May and August, with the lowest point being reached in October. A significant difference was observed in the relative abundance of *F. perrara* between August and September with a dominance peak in August, and similar results were obtained for *Enterobacteriaceae* between August and October, but with a dominance peak in September. The relative abundance of *S. alvi* was low for the first three months, then declined in August to its lowest point, but increased at the end of the season. An approximately eight-fold increase in the abundance of *S. alvi* was detected between September and October. Calculations on α-diversity gave a markedly lower diversity in May, which significantly increased until July, and then remained fairly stable throughout October (Fig. 2b). The one-way ANOVA for both the four main bacteria and α-diversity displayed significant *p*-values of $p<0.01$ with respect to the temporal trends between May and October (Table S2).

The relative quantity of 16S rRNA genes (bacterial load) was the highest in May and peaked again in September, and the lowest point being reached in October (Fig. 2c). Tukey HSD significant differences between May and June, August and September, and September and October measurements were determined at the 95% and 99% levels (Fig. 2c, Table S2). The one-way ANOVA for the whole dataset revealed significance of $p<0.01$ between May and October (Table S2).

Colony variation

Colony variations were examined among the three colonies in September, and no significant differences were observed in bacterial relative abundance after a *t*-test Bonferroni correction (Fig. S4A). However, one colony (colony 3) had a significantly higher relative bacterial load ($p<0.01$) than the two other colonies (Fig. S4B).

Population composition in a stable environment

The two time point analysis of the two colonies feeding on stored sugars through the winter showed no significant differ-

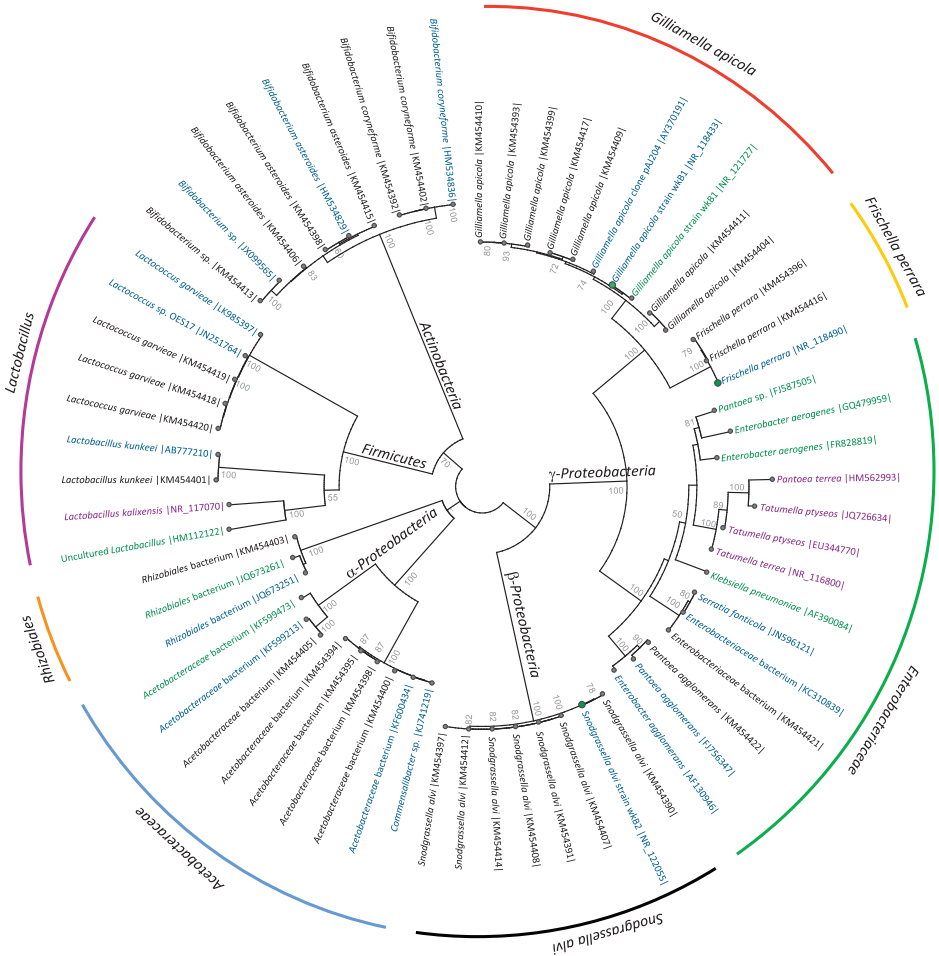

Fig. 1. Neighbor-joining phylogenetic tree of bacterial isolates from the honey bee midgut/pylorus. All 34 bacteria cultured and isolated from the honeybee midgut/pylorus collected in July 2012 were included (black text), and their reference sequences (>99% identity BLAST hits) have blue colored text. Reference sequences for Illumina BLAST hits are marked with purple text and blue text with a green node. In addition, the best percent identity BLAST hit sequences for the four bacterial components retrieved from mixed sequencing are included with green colored text. The tree was made using CLC Main workbench 6 and bootstrap values over 50% are shown. The colored circle outside the main tree shows the BLAST based taxonomic assignments for the bacterial components from mixed Sanger sequencing. The following color codes were used; going in a clockwise direction: red; *Gilliamella apicola*, yellow; *Frischella perrara*, green; *Enterobacteriaceae*; black; *Snodgrassella alvi*, blue; *Acetobacteraceae*, orange; *Rhizobiales* bacterium, purple; *Lactobacillus*.

Table 2. Assigned taxonomy by BLAST hits for bacterial components derived from mixed sequencing

Component	Assigned taxonomy	Accession No. GenBank	Dataset	E-value	% identity
Frischella perrara	*Frischella perrara*	NR_118490	Seasonal	1e-73	96
Gilliamella apicola	*Gilliamella apicola*	NR_121727	Seasonal/Stable environment	1e-83	99
Snodgrassella alvi	*Snodgrassella alvi*	NR_122055	Seasonal/Stable environment	1e-73	95
Enterobacteriaceae	*Klebsiella pneumonia*	AF390084	Seasonal	3e-60	91
	Pantoea sp.	FJ587505		3e-59	90
	Enterobacter aerogenes	FR828819		3e-59	90
Acetobacteraceae	*Acetobacteraceae* bacterium	KF599473	Stable environment	5e-53	89
Rhizobiales	*Rhizobiales* bacterium	JQ673261	Stable environment	5e-83	99
Lactobacillus	Uncultured *Lactobacillus* sp.	HM112122	Stable environment	2e-47	87

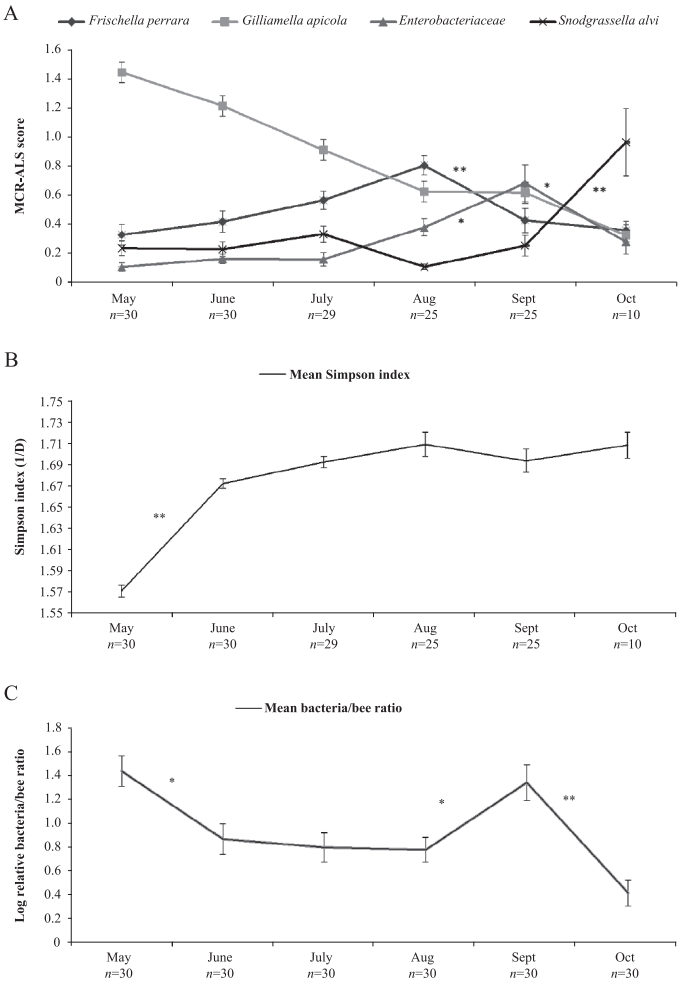

Fig. 2. Seasonal changes in the honey bee midgut/pyloric microbiota. A) Mean microbiota composition throughout a foraging season between May and October for the four dominating components found in the honey bee midgut/pylorus by mixed sequencing. The MCR-ALS score, determined by mixed sequencing, represents an approximately relative bacterial composition in the honey bee midgut/pylorus without assuming closure of the system. B) α-diversity between May and October calculated from the raw spectra of mixed sequencing. C) Mean bacterial load for each month between May and October. The calculated relative ratio between 16S rRNA genes and vitellogenin genes (bacteria/bee), in the midgut/pylorus determined by quantitative PCR, is shown. Significant differences were observed between May and October (one-way ANOVA p=<0.01) in the three analyses, and n; number of bees included in the final analyses each month. Markings show the error bars of the calculated SEM (2α=68.2% CI) for each month and Tukey HSD significant difference values, from pair-wise comparisons of the neighborly time points of the monthly average values of n bees, is shown; *=p<0.05, **=p<0.01.

ence after a *t*-test Bonferroni correction either in bacterial relative abundance between November and February (Fig. 3) or in relative bacterial load (bacteria/bee ratio), which was 0.15±0.08 and 0.25±0.1 for November and February, respectively.

Competition experiment

The results obtained in the seasonal study suggested a negative interaction between the two bacterial species *S. alvi* and *G. apicola*. Therefore, these bacteria were selected for an *in vitro* competition experiment. When grown alone, *G. apicola* and *S. alvi* both showed a steady state bacterial load; however, when grown together, *S. alvi* showed significantly less growth (Fig. 4). In contrast, *S. alvi* did not influence the growth of *G. apicola*; no significant differences were observed from *G. apicola* grown alone.

Technical validation

The calculated average value of two replicates on each plate indicated minor plate variations in the bacterial relative abundance analysis (Fig. S5). *G. apicola* in this average July sample showed a mean difference of 0.3 units in the MCR-

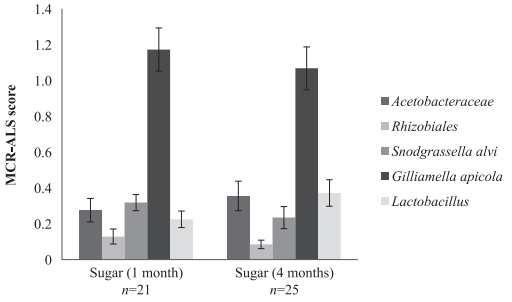

Fig. 3. The honey bee midgut/pyloric microbiota composition in the stable environment dataset. The mean microbiota composition at two different time points (November and February) under stable dietary conditions (sugar-mix) for the five dominating components found in the honey bee midgut/pylorus by mixed sequencing. At the two time points, November and February, the bees had been fed the commercial sugar mix for one month and four months, respectively. The MCR-ALS score, determined by mixed sequencing, represents an approximately relative bacterial composition in the honey bee midgut/pylorus without assuming closure of the system. Markings show the error bars of the calculated SEM (2α=68.2% CI) and *n*; number of bees included in the final analyses each month.

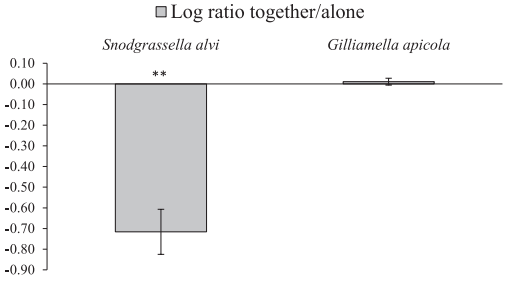

Fig. 4. *In vitro* competition experiment with *Gilliamella apicola* and *Snodgrassella alvi*. Significant differences were observed in the relative ratio between *S. alvi* when grow alone and when grown together with *G. apicola*. Markings show the error bars of the calculated SEM (2α=68.2% CI), *t*-test: ***p*=<0.01. The calculated mean values from three independent experiments are shown.

ALS score between May and August, which was the largest difference observed. The qPCR analysis showed a bacterial to bee ratio of 0.59±0.114 with respect to plate-to-plate variations.

Of the 30 midguts/pylori analyzed each month, the resulting number of samples after quality filtering during the MCR-ALS analysis was: May 100%, June 100%, July 97%, August 83%, September 83%, and October 33%. The reduced number of sequences passing quality filtering in October may have been due to the low amount of bacterial DNA that month (as determined by qPCR). Of the 54 bees (24 in November+30 in February) collected for the stable environment dataset, 87% and 83%, respectively, passed the quality filter.

Discussion

Our approach enabled the detection of distinct shifts in honey bee midgut/pyloric bacteria throughout the season. This was in contrast to Corby-Harris, *et al.* (9), who found no significant difference between the bacterial composition in spring and fall when characterizing the total gut community (crop, midgut, ileum and rectum) of foraging bees. Due to the high amount of bacteria in the hindgut (the lower part of ileum and rectum) (35), the microbiota composition in reference to this part may have been overrepresented if the intestine (midgut, ileum, and rectum) was analyzed as a whole. Therefore, our results support the midgut/pyloric microbiota being more dynamic than the hindgut microbiota. The bacterial composition of the midgut/pylorus may shift because of local metabolic processes, and/or the repeated shedding of the peritrophic membrane and reflux from the ileum (48). At the same time, the highly structured seasonal shifts observed suggest that external forces may play a significant role in shaping the midgut/pyloric microbiota. We favored environmental exposure and dietary changes as the main drivers for seasonal trends because of the strong relationship that exists between gut bacterial composition and the host diet in humans, vertebrate animals, and insects (7, 12, 14, 24). In support of this view, our measurements of the midgut/pylorus microbiota composition from hives fed the same diet over a period of four months showed a stable microbiota composition.

The marked increase observed in α-diversity between May, June, and July may, in some extent, be explained by shifts in the dominating bacteria. Comparisons between mixed sequencing and α-diversity calculations indicated that the decrease in *G. apicola* alone can not completely describe the increase in α-diversity between May and June. Therefore, we also considered this increase to be influenced by the colony being more exposed to various environmental bacteria as well as additional dietary compounds when bees start to forage than when the colony does not forage during the winter (2, 9). Although bees were foraging in May (starting in April in 2012), there was a limited food supply in Ås, Norway at that time of the year, and the available foraging plants provided pollen as a main nutrient. In June, more flowers and plants emerged and persisted, and various nectar and some pollen sources were available. This richer foraging context continued throughout the summer with more nectar-bearing flowers becoming available. We speculated that the high α-diversity that was still prevalent in October indicated that bacteria from the peak foraging season persisted in stored food reserves inside the hive for some time (2).

The peaks in bacterial load appeared to correlate with the two main dietary changes in May and September. The bacterial peak in May most likely reflected the *G. apicola* component, which we found to be highly dominant at this point. Previous studies reported that *G. apicola* was the most abundant in the ileum (35), and is the sole bacterium in the honey bee gut that is able to degrade pectin (19). Pectin is a main constituent of pollen (47), and its degradation is known to occur in the midgut (27). These findings indicate that bees foraging early on in the season mostly acquire pollen as a nutrient; hence, bacteria able to utilize pollen will proliferate and dominate in the midgut/pylorus. The peak in September likely reflected the proliferation of *Enterobacteriaceae* because this component dominates the bacterial composition in September. The start of sugar feeding in September and nectar foraging during August may both have influenced this proliferation. The latter may have had a stronger influence because elevated levels of *Enterobacteriaceae* were already

detected in July/August when the bees were still foraging. BLAST hits of our *Enterobacteriaceae* component gave the best percent identity to different bacteria genera previously isolated from honey bee guts and from plants (2, 4, 54). Furthermore, our results were consistent with previous findings by Corby-Harris, *et al.* (9) who detected *Enterobacteriaceae* only in the gut samples of forager bees in fall. Comparisons of this component to both Illumina and culturing BLAST hits gave a span of four different bacterial genera (*Pantoea*, *Enterobacter*, *Tatumella*, and *Serratia*). Reclassification within these genera in recent years, exemplified by *Enterobacter agglomerans* being transferred to *Pantoea agglomerans* by Gavini, *et al.* (21), and *Pantoea* sp. being assigned to *Tatumella* sp. by Brady, *et al.* (5), appeared to connect these results together.

Colony demography has been suggested to play a role in measuring the total bacterial load and bacterial relative abundance throughout a season. The production of new bees by the bee colony is seasonal; one-d-old bees harbor at least three orders of magnitude fewer bacteria than older workers (35). A recent study by Powell, *et al.* (41) showed differences in the microbiota composition between newly hatched bees and 16-d-old bees. We collected ten bees randomly from three different combs resulting in 30 bees per time point. This method ensured a broad representation of ages and task groups; however, since we did not sample by age, we cannot rule out some age-related influence on the gut microbiota. However, Martinson, *et al.* (35) reported that the bacterial load in young workers increased rapidly (within nine d) to that in older bees. Due to this rapid colonization, it is unlikely that our dynamic results were solely driven by changes in colony demography. The result obtained in October, with the high prevalence of *S. alvi*, may have been influenced by age-related differences. Worker bees become more similar in age as the colony prepares for winter because the production of new bees slows down and stops and the oldest bees die out (36). This compression of age in the worker caste may explain the abrupt change observed in the bacterial composition in October because the prevalence of *S. alvi* was high in the ileum of young bees (41).

A separate factor that also needs to be considered is temperature changes, which may have a major impact on bacterial communities. However, although ambient temperatures markedly change throughout the year in Norway, the within-colony environment of honey bees is more stable, with minor changes in core temperatures and only 10–20°C variations at the periphery (8, 40, 43), (www.norges-birokterlag.no, www.stadevægt.dk). In contrast to the changing ambient environment, the midgut/pyloric microbiota remained stable between November and February in our colonies. This stability suggests that the shifts observed in the relative abundance of bacteria were not driven by temperature; however, some influences of temperature cannot be ruled out and require further study.

When we compared our bacterial load results with those from the bacterial relative abundance analysis, we found a two-fold decrease in the total bacterial load in October, corresponding to an approximately four-fold increase in *S. alvi* within the bacterial composition. This, as an additional explanation for the October result, indicated *S. alvi* outgrowth only when there were few other bacteria present. Suppression was confirmed by one strain of *G. apicola* in our *in vitro* competition experiment, which showed that, when competing for the same nutrients and grown under set conditions, *G. apicola* suppressed *S. alvi* outgrowth. Although we cannot generalize from two strains, recent findings support these two species occupying different niches: Martinson *et al.* (35) conducted FISH staining of the honeybee gut and found that *S. alvi* adhered to the midgut/ileum wall, forming a bacterial layer, whereas *G. apicola* habited the luminal niche. *G. apicola* and *S. alvi* occupying different niches is in line with recent evidence for genome complementarity between these two bacteria (32). Furthermore, these two bacteria have been shown to exhibit different growth properties when grown in broth. *S. alvi* without flagella (17) grew in the bottom of a tube as a bacterial layer (31), whereas *G. apicola* with its flagella (17) was suited for competition for nutrients throughout the whole broth (31). We detected similar growth properties for our strains. These findings together with the present results indicate that, *in vivo*, luminal *G. apicola* may prevent *S. alvi* from entering the luminal niche.

The mixed Sanger sequencing used in this study was originally applied and validated for a time series analysis of mixed bacteria communities (50). The limitations of this method lie in the detection of low abundance species in high richness communities. The bacteria, which we discovered dominating the midgut/pyloric microbiota, in our dataset were previously identified as major constituents in the honey bee gut (18, 31, 34, 37), thereby supporting the suitability of the analyses. The relatively simple and defined honey bee gut composition makes mixed Sanger sequencing a better choice than high throughput sequencing because it is cheaper and less computer intensive. In addition, our OTU classification by QIIME could not distinguish between *Frischella* and *Gilliamella*, but rather classified them as *Pasteurellales*. Both bacteria were identified using Sanger sequencing, which indicated the higher sensitivity of Sanger-sequencing reads when identifying highly similar sequences. Mixed sequencing revealed that our *Enterobacteriaceae* component did not dominate the bacterial composition in July, which is in contrast to the results obtained by Illumina sequencing. Therefore, Illumina sequencing may have inferred a bias in these sequences because *Enterobacteriaceae* did not dominate amongst our bacteria isolates (only 2 out of 24 isolates) or in the gut microbiota in other studies.

Conclusion

The results of the present study exemplify major changes in the honey bee midgut/pyloric microbiota composition throughout a foraging season, whereas a stable microbiota composition was maintained under stable environmental conditions during winter. We emphasize the need for longitudinal studies to investigate and understand the gut microbiota in honey bees.

Acknowledgements

We wish to thank Ashish K. Shah for helping with bee collection and bee brain dissection, and Felix Nwosu for the Illumina MiSeq analysis setup. We also extend our thanks to the staff at the local hospital in Fredrikstad, Norway, Department of Microbiology, for

supplying the horse blood and protocols used in culturing the midgut bacteria. G.V.A. was supported by the PEW Charitable Trust and the Research Council of Norway (award# 213976 and 191699).

References

1. Amdam, G.V., K. Norberg, M.K. Fondrk, and R.E. Page Jr. 2004. Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. Proc. Natl. Acad. Sci. U.S.A. 101:11350–11355.
2. Anderson, K.E., T.H. Sheehan, B.M. Mott, P. Maes, L. Snyder, M.R. Schwan, A. Walton, B.M. Jones, and V. Corby-Harris. 2013. Microbial ecology of the hive and pollination landscape: bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). PLoS One 8:e83125.
3. Avershina, E., O. Storro, T. Oien, R. Johnsen, R. Wilson, T. Egeland, and K. Rudi. 2013. Bifidobacterial succession and correlation networks in a large unselected cohort of mothers and their children. Appl. Environ. Microbiol. 79:497–507.
4. Brady, C., I. Cleenwerck, S. Venter, M. Vancanneyt, J. Swings, and T. Coutinho. 2008. Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). Syst. Appl. Microbiol. 31: 447–460.
5. Brady, C.L., S.N. Venter, I. Cleenwerck, K. Vandemeulebroecke, P. De Vos, and T.A. Coutinho. 2010. Transfer of *Pantoea citrea*, *Pantoea punctata* and *Pantoea terrea* to the genus *Tatumella emend.* as *Tatumella citrea* comb. nov., *Tatumella punctata* comb. nov. and *Tatumella terrea* comb. nov. and description of *Tatumella morbirosei* sp. nov. Int. J. Syst. Evol. Microbiol. 60:484–494.
6. Caporaso, J.G., J. Kuczynski, J. Stombaugh, *et al.* 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7:335–336.
7. Chandler, J.A., J.M. Lang, S. Bhatnagar, J.A. Eisen, and A. Kopp. 2011. Bacterial communities of diverse Drosophila species: ecological context of a host-microbe model system. PLoS Genet. 7:e1002272.
8. Cooper, D.P., M.W. Schaffer, and L.S. Buchmann. 1985. Temperature regulationof honey bees (*Apis mellifera*) foraging in the sonoran desert. J. Exp. Biol.:1–15.
9. Corby-Harris, V., P. Maes, and K.E. Anderson. 2014. The bacterial communities associated with honey bee (*Apis mellifera*) foragers. PLoS One 9:e95056.
10. Crailsheim, K. 1988. Intestinal transport of sugars in the honeybee (*Apis mellifera* L.). J. Insect Physiol. 34:839–845.
11. Crailsheim, K. 1990. The protein balance of the honey bee worker. Apidologie 21:417–429.
12. David, L.A., C.F. Maurice, R.N. Carmody, *et al.* 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505:559–563.
13. Davis, F.C. 2004. Inside and outside influences, p. 145–157. *In* F.C. David (ed.), The Honey Bee Inside Out. Bee Craft Limited, Warwickshire.
14. De Filippo, C., D. Cavalieri, M. Di Paola, M. Ramazzotti, J.B. Poullet, S. Massart, S. Collini, G. Pieraccini, and P. Lionetti. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc. Natl. Acad. Sci. U.S.A. 107:14691–14696.
15. DeSantis, T.Z. P Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G.L. Andersen. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 72:5069–5072.
16. Disayathanoowat, T., J.P. Young, T. Helgason, and P. Chantawannakul. 2012. T-RFLP analysis of bacterial communities in the midguts of *Apis mellifera* and *Apis cerana* honey bees in Thailand. FEMS Microbiol. Ecol. 79:273–281.
17. Engel, P., V.G. Martinson, and N.A. Moran. 2012. Functional diversity within the simple gut microbiota of the honey bee. Proc. Natl. Acad. Sci. U.S.A. 109:11002–11007.
18. Engel, P., W.K. Kwong, and N.A. Moran. 2013. *Frischella perrara* gen. nov., sp. nov., a gammaproteobacterium isolated from the gut of the honey bee, *Apis mellifera*. Int. J. Syst. Evol. Microbiol. 63:3646–3651.
19. Engel, P., and N.A. Moran. 2013. Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. Gut Microbes 4:60–65.
20. Evans, J.D., and T.N. Armstrong. 2006. Antagonistic interactions between honey bee bacterial symbionts and implications for disease. BMC Ecol. 6:4.
21. Gavini, F., J. Mergaert, A. Beji, C. Mielcarek, D. Izard, K. Kersters, and J. De Ley. 1989. Transfer of *Enterobacter agglomerans* (Beijerinck 1888) Ewing and Fife 1972 to *Pantoea* gen. nov. as *Pantoea agglornerans* comb. nov. and description of *Pantoea dispersa* sp. nov. Int. J. Syst. Evol. Mircobiol. 337–345.
22. Hamdi, C., A. Balloi, J. Essanaa, *et al.* 2011. Gut microbiome dysbiosis and honeybee health. J. Appl. Entomol. 135:524–533.
23. Hatjina, F., C. Costa, R. Büchler, *et al.* 2014. Population dynamics of European honey bee genotypes under different environmental conditions. J. Apicult. Res. 53:233–247.
24. Hildebrandt, M.A., C. Hoffmann, S.A. Sherrill-Mix, S.A. Keilbaugh, M. Hamady, Y.Y. Chen, R. Knight, R.S. Ahima, F. Bushman, and G.D. Wu. 2009. High-fat diet determines the composition of the murine gut microbiome independently of obesity. Gastroenterology 137:1716–1724. e1711–1712.
25. Ihrig, J., R. Lill, and U. Mühlenhoff. 2006. Application of the DNA-specific dye EvaGreen for the routine quantification of DNA in microplates. Anal. Biochem. 359:265–267.
26. Jeyaprakash, A., M.A. Hoy, and M.H. Allsopp. 2003. Bacterial diversity in worker adults of *Apis mellifera capensis* and *Apis mellifera scutellata* (Insecta: *Hymenoptera*) assessed using 16S rRNA sequences. J. Invertebr. Pathol. 84:96–103.
27. Klungness, L.M.P.Y. 1984. A histochemical study of pollen digestion in the alimentary canal of honeybees (*Apis mellifera* L.). J. insect. Physiol. 30:511–521.
28. Koch, H., and P. Schmid-Hempel. 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. Proc. Natl. Acad. Sci. U.S.A. 108:19288–19292.
29. Koch, H., D.P. Abrol, J. Li, and P. Schmid-Hempel. 2013. Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. Mol. Ecol. 22:2028–2044.
30. Kubista, M., J.M. Andrade, M. Bengtsson, *et al.* 2006. The real-time polymerase chain reaction. Mol. Aspects. Med. 27:95–125.
31. Kwong, W.K., and N.A. Moran. 2013. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: *Snodgrassella alvi* gen. nov., sp. nov., a member of the *Neisseriaceae* family of the *Betaproteobacteria*; and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., *Orbales* ord. nov., a sister taxon to the *Enterobacteriales* order of the *Gammaproteobacteria*. Int. J. Syst. Evol. Microbiol. 63:2008–2018.
32. Kwong, W.K., P. Engel, H. Koch, and N.A. Moran. 2014. Genomics and host specialization of honey bee and bumble bee gut symbionts. Proc. Natl. Acad. Sci. U.S.A. 111:11509–11514.
33. Mao, F., W.Y. Leung, and X. Xin. 2007. Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. BMC Biotechnol. 7:76.
34. Martinson, V.G., B.N. Danforth, R.L. Minckley, O. Rueppell, S. Tingek, and N.A. Moran. 2011. A simple and distinctive microbiota associated with honey bees and bumble bees. Mol. Ecol. 20:619–628.
35. Martinson, V.G., J. Moy, and N.A. Moran. 2012. Establishment of characteristic gut bacteria during development of the honeybee worker. Appl. Environ. Microbiol. 78:2830–2840.
36. Mattila, H.R., J. Harris, and G.W. Otis. 2001. Timing production winterbees in honey bee (*Apis mellifera*) colonies. Insectes Soc. 48:88–93.
37. Moran, N.A., A.K. Hansen, J.E. Powell, and Z.L. Sabree. 2012. Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. PLoS One 7:e36393.
38. Nadkarni, M.A., F.E. Martin, N.A. Jacques, and N. Hunter. 2002. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology 148:257–266.
39. Naseribafrouei, A., K. Hestad, E. Avershina, M. Sekelja, A. Linlokken, R. Wilson, and K. Rudi. 2014. Correlation between the human fecal microbiota and depression. Neurogastroenterol. Motil. 26:1155–1162.
40. Omholt, S.W. 1987. Thermoregulation in the winter cluster of the honeybee, *Apis Mellifera*. J. Theor. Biol. 128:219–231.

41. Powell, J.E., V.G. Martinson, K. Urban-Mead, and N.A. Moran. 2014. Routes of acquisition of the gut microbiota of *Apis mellifera*. Appl. Environ. Microbiol. 80:7378–7387.
42. Ruijter, J.M., C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff, and A.F. Moorman. 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic. Acids. Res. 37:e45.
43. Seeley, D.T. 1995. The foraging abilities of a colony, p. 295–307. *In* The Wisdom of the Hive, the Social and Physiology of the Honey Bee Colonies, First ed. Harvard University Press, Cambridge, Massachusetts London, England.
44. Sekelja, M., I. Rud, S.H. Knutsen, V. Denstadli, B. Westereng, T. Naes, and K. Rudi. 2012. Abrupt temporal fluctuations in the chicken fecal microbiota are explained by its gastrointestinal origin. Appl. Environ. Microbiol. 78:2941–2948.
45. Singer, V.L., L.J. Jones, S.T. Yue, and R.P. Haugland. 1997. Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. Anal. Biochem. 249:228–238.
46. Ståhlberg, A., N. Zoric, P. Åman, and M. Kubista. 2005. Quantitative real-time PCR for cancer detection: the lymphoma case. Expert Rev. Mol. Diagn. 5:221–230.
47. Taylor, L.P.H.P. 1997. Pollen germination and tube growth. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:461–491.
48. Tellam, R.L. 1996. The peritrophic matrix., p. 86–114, *In* M.J.Lehane and P.F. Billingsley (ed.), Biology of the Insect Midgut. Chapman & Hall, London.
49. Trosvik, P., B. Skånseng, K.S. Jakobsen, N.C. Stenseth, T. Næs, and K. Rudi. 2007. Multivariate analysis of complex DNA sequence electropherograms for high-throughput quantitative analysis of mixed microbial populations. Appl. Environ. Microbiol. 73:4975–4983.
50. Trosvik, P., K. Rudi, T. Naes, A. Kohler, K.S. Chan, K.S. Jakobsen, and N.C. Stenseth. 2008. Characterizing mixed microbial population dynamics using time-series analysis. ISME J. 2:707–715.
51. vanEngelsdorp, D., and M.D. Meixner. 2010. A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. J. Invertebr. Pathol. 103:S80–S95.
52. Vasquez, A., E. Forsgren, I. Fries, R.J. Paxton, E. Flaberg, L. Szekely, and T.C. Olofsson. 2012. Symbionts as major modulators of insect health: lactic acid bacteria and honeybees. PLoS One 7:e33188.
53. Wu, M., Y. Sugimura, N. Takaya, D. Takamatsu, M. Kobayashi, D. Taylor, and M. Yoshiyama. 2013. Characterization of bifidobacteria in the digestive tract of the Japanese honeybee, *Apis cerana japonica*. J. Invertebr. Pathol. 112:88–93.
54. Yoshiyama, M., and K. Kimura. 2009. Bacteria in the gut of Japanese honeybee, *Apis cerana japonica*, and their antagonistic effect against *Paenibacillus larvae*, the causal agent of American foulbrood. J. Invertebr. Pathol. 102:91–96.
55. Yu, Y., C. Lee, J. Kim, and S. Hwang. 2005. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. Biotechnol. Bioeng. 89:670–679.
56. Zimonja, M., K. Rudi, P. Trosvik, and T. Næs. 2008. Multivariate curve resolution of mixed bacterial DNA sequence spectra: identification and quantification of bacteria in undefined mixture samples. J. Chemometr. 22:309–322.

Received: 7 September 2016 | Revised: 16 September 2017 | Accepted: 6 October 2017

DOI: 10.1111/mec.14392

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Geographically widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant patterns

Jane Ludvigsen[1] | Davide Porcellato[1] | Trine M. L'Abée-Lund[2] | Gro V. Amdam[3,4] | Knut Rudi[1]

[1]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway

[2]Department of Food Safety and Infection Biology, Norwegian University of Life Sciences, Oslo, Norway

[3]Faculty of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

[4]School of Life Sciences, Arizona State University, Tempe, AZ, USA

Correspondence
Knut Rudi, Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway.
Email: knut.rudi@nmbu.no

Abstract

How long-term antibiotic treatment affects host bacterial associations is still largely unknown. The honeybee-gut microbiota has a simple composition, so we used this gut community to investigate how long-term antibiotic treatment affects host-associated microbiota. We investigated the phylogenetic relatedness, genomic content (GC percentage, genome size, number of genes and CRISPR) and antibiotic-resistant genes (ARG) for strains from two abundant members of the honeybee core gut microbiota (*Gilliamella apicola* and *Snodgrassella alvi*). Domesticated honeybees are subjected to geographically different management policies, so we used two research apiaries, representing different antibiotic treatment regimens in their apiculture: low antibiotic usage (Norway) and high antibiotic usage (Arizona, USA). We applied whole-genome shotgun sequencing on 48 *G. apicola* and 22 *S. alvi*. We identified three predominating subgroups of *G. apicola* in honeybees from both Norway and Arizona. For *G. apicola*, genetic content substantially varied between subgroups and distance similarity calculations showed similarity discrepancy between subgroups. Functional differences between subgroups, such as pectin-degrading enzymes (*G. apicola*), were also identified. In addition, we identified horizontal gene transfer (HGT) of transposon (Tn10)-associated tetracycline resistance (Tet B) across the *G. apicola* subgroups in the Arizonan honeybees, using interspace polymorphisms in the Tet B determinant. Our results support that honeybee-gut symbiont subgroups can resist long-term antibiotic treatment and maintain functionality through acquisition of geographically distinct antibiotic-resistant genes by HGT.

KEYWORDS
Apis mellifera, gut symbionts, population ecology, tetracycline resistance, WGS

1 | INTRODUCTION

Bacterial antibiotic resistance is an emerging problem throughout the world. Due to continuous, massive horizontal gene transfer (HGT) between bacteria, antibiotic-resistant genes (ARG) are now found in almost all environments (Akhtar, Hirt, & Zurek, 2009; Bryan, Shapir, & Sadowsky, 2004; Davies & Davies, 2010; Nesme et al., 2014). However, there are large geographical differences with respect to resistance patterns. Countries with extensive use of antibiotics have generally higher levels of environmental antibiotic resistance (Appelbaum, 1992), and this predicament can also be observed in host-associated bacteria (Kumar et al., 2013; Tian, Fadhil, Powell, Kwong, & Moran, 2012). Knowledge on how resistance is maintained among symbiotic gut bacteria is still lacking. If the antibiotic treatment selects for specific antibiotic-resistant bacteria, strains within a bacterial community still remain largely unknown. A major challenge in addressing ARG spread in naturally occurring bacterial populations is their complexity. Whether existing in nature, as human or animal

host-associated microbiota, these bacterial populations comprise each at least more than 100 different genera (Sankar, Lagier, Pontarotti, Raoult, & Fournier, 2015). Contrary, honeybees are less complex models, and can be used to study not only gut-community dynamics, but also ARG and HGT within host populations (Koch & Schmid-Hempel, 2011a; Kwong & Moran, 2016; Ludvigsen et al., 2015; Martinson et al., 2011; Moran, Hansen, Powell, & Sabree, 2012; Tian et al., 2012). As for humans in different geographic regions, domesticated honeybees are subjected to geographically different antibiotic treatment regimens (Levy & Marshall, 2013). In the United States, oxytetracycline is frequently used to both prevent and treat American foulbrood (infection with the bacterium *Paenibacillus larvae*) (Genersch, 2010; Reybroeck, Daeseleire, De Brabander, & Herman, 2012; Spivak, 2000; Thompson et al., 2005). Due to this, several different tetracycline-resistant (Tc^r) genes have been identified in the honeybee gut (Tian et al., 2012). In contrast, countries that do not use oxytetracycline for beekeeping detect fewer Tc^r genes, which are also less abundant (Tian et al., 2012). In Norway, oxytetracycline is not used in commercial beekeeping (http://www.norges-birokterlag.no/). Partially due to past extensive global use of tetracycline, Tc^r genes are found worldwide within natural bacterial populations. The efflux pump type of tetracycline (Tet B), associated with the Tn10 transposon, is one of the most frequently identified determinants (Bryan et al., 2004; Martinez et al., 2009; Tian et al., 2012; Wilkerson, Samadpour, van Kirk, & Roberts, 2004). Occurrence of ARG in natural environments raises concern about the possible spread to human pathogens, as well as uncertainty about how these genes influence functionality of host symbiotic bacteria (Davies & Davies, 2010; Sommer, Dantas, & Church, 2009).

Two of the most abundant bacteria within the honeybee-gut microbiota, *Gilliamella apicola* (γ-proteobacteria) and *Snodgrassella alvi* (β-proteobacteria), play a part in maintaining its host health (Engel, Martinson, & Moran, 2012; Koch & Schmid-Hempel, 2011b). *Gilliamella apicola* and *S. alvi* live in close proximity, but occupy distinct ecological niches within the gut community. *Snodgrassella alvi* grows as a biofilm along the gut wall, while *G. apicola* exists on top of this biofilm and extends into the lumen (Kwong, Engel, Koch, & Moran, 2014; Martinson, Moy, & Moran, 2012). *Gilliamella apicola* exhibits phenotypic diversity with respect to the capacity to degrade pectin, but it has not yet been determined whether this property is linked to specific linages (Engel, Stepanauskas, & Moran, 2014; Jayani, Saxena, & Gupta, 2005; Kwong & Moran, 2015; Moran et al., 2012).

The aim of this work was to investigate strain diversity of *G. apicola* and *S. alvi*, and address whether antibiotic usage leads to a few geographically restricted, bacterial strains, or whether important functional traits are preserved through HGT of ARG among honeybee-associated bacterial strains (Figure 1). We used whole-genome shotgun sequencing to address how the phylogenetic relatedness of strains of *G. apicola* and *S. alvi* associates with genomic content (GC percentage, genome size, number of CDS and CRISPR) and Tc^r genes. Both bacteria were isolated from two different research apiaries. Norway represented low antibiotic usage, while Arizona represented high antibiotic usage. We also performed minimum inhibitory concentration (MIC) testing on cultured bacterial strains.

This work presents evidence for differences in genome content between subgroups in *G. apicola*, and HGT of Tc^r genes among the Arizonan *G. apicola* strains. Our results give insight into how different subgroups persist in the gut through ARG transfer under long-term antibiotic exposure.

2 | MATERIALS AND METHODS

2.1 | Bee sampling

Bees were sampled from research apiaries at the Norwegian University of Life Sciences and Arizona State University. The two research apiaries were not treated with oxytetracycline, so any antibiotic exposure must have occurred in the several commercial apiaries

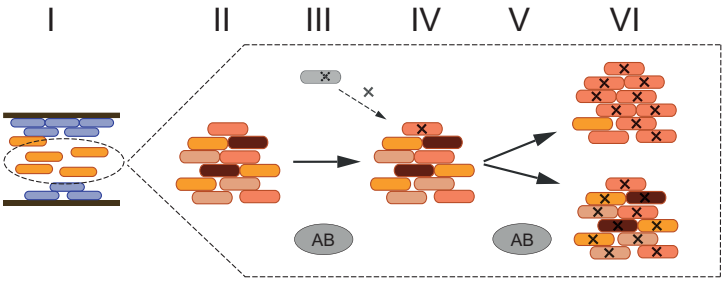

FIGURE 1 Main hypothesis. I: bacteria community within honeybee-gut intestine. II: by genome sequencing, several strains from one bacterium species, subgroups within this species can be identified (subgroups are shown with different colours). III & IV: antibiotic treatment enhances antibiotic-resistant genes (ARG) uptake into the population. Bacteria that harbour ARGs will survive and can sustain and proliferate. V: continuous exposure to antibiotics (selective pressure). VI: either the strain/subgroup with the ARG proliferates and dominates the population, or the ARG will be transferred to other subgroups by HGT. AB = antibiotic treatment, X = ARG. In the honeybee intestines, some bacteria form biofilms adhering to the gut wall (*Snodgrassella alvi*), while others live within the lumen cavity (*Gilliamella apicola*), both continuously interacting with each other. The different bacteria are shown with different colours (I)

from which the hives originated. Although specific numbers are not available on how often beehives are treated, the practice is ongoing and extensive in the USA, while antibiotics are not used in beekeeping in Norway. These research apiaries contained collections of honeybee colonies from different commercial apiaries in each country. Therefore, our samples represented a broader sampling distribution than just the two research apiaries where the colonies were housed for the experiment. We sampled three different colonies from the Norwegian apiary and nine different colonies (July 2011) from the one in Arizona for quantitative, real-time PCR (qPCR) screens. For bacterial culturing, the same three Norwegian colonies were used (July 2011), while three different Arizonian colonies were used, due to different sampling time points (October 2013).

2.2 | DNA extraction for real-time PCR screening of midguts

Ninety bees of different ages from Norway and Arizona were randomly picked from three different frames inside the hives, for a total of 180 bees that represented the selected colonies. Midgut/pyloric samples were aseptically dissected and collected from individual bees and transferred into plastic tubes with added S.T.A.R buffer (500 μl for DNA preservation) and ~0.20 g <106-μm glass beads. The samples were stored at −20°C until DNA extraction was performed. The individual honeybee guts were lysed using the MagnaLyser at 2 × 20 s at speed 6,500 rpm, with 1 min rest at 4°C between runs. An automated DNA magnetic bead-based extraction method was used on all samples (developed by Genetic Analysis; http://www.genet-analysis.com/) and used in combination with the KingFisher Flex robot. Quant-iT PicoGreen® dsDNA assay (Life Technologies™, USA) was used to quantify the extracted DNA.

The midguts/pylorus regions were all screened for four Tc^r genes (*tet*(B), *tet*(C), *tet*(W), and *tet*(H)) by qPCR on the LightCycler 480 II (Roche, Germany). These genes were selected, based on results from Tian et al. (2012), due to being the most prevalent Tc^r genes found in bees from countries that do not use oxytetracycline in beekeeping. Primer pairs for these four genes were used at a final concentration of 0.2 μM, in combination with 5×HOT FIREPol® EvaGreen qPCR Mix Plus (Solis BioDyne, Estonia) at final cons of 1.25 U/μl, with added nuclease-free water at a total of 20 μl per reaction. Cycling conditions were as follows: 95°C for 15 min; 40 cycles of 95°C for 30 s; annealing temp for 30 s; and 72°C for 30 s. Annealing temp for *tet*(B), *tet*(C), *tet*(W) and *tet*(H) were as Tian et al. (2012) described. Positive and negative controls were included for each gene analysis. The criterion for positive sample was one Cq value below the negative control Cq value. A melting point curve analysis was added at the end of each qPCR run, and the PCR amplicons of *tet*(B) and *tet*(H) were also Sanger-sequenced for verification of the correct PCR product. Additionally, we performed qPCR on the 16S rRNA gene for the same samples to account for possible differences in size of the midgut/pylorus region between samples. Primers and PCR conditions were as previously described in Ludvigsen et al. (2015).

2.3 | Bacteria culturing and DNA extraction

Thirty bees of different ages from three different colonies (10 bees from each) were randomly picked from three different frames inside the hives, representing the selected colonies from both research apiaries (Norway and Arizona) for a total of 60 bees. Bee midguts/pyloric regions were aseptically dissected and pooled together (keeping the Norwegian and Arizonan samples separate), and stored in glycerol stocks at −80°C until plating was performed. The gut parts were homogenized in PBS, using a mixing pestle. A 10-fold dilution was made using PBS to avoid overgrowth when plating. A total of 100 μl of the homogenate was plate spread on TSA with 5% horse blood, with and without tetracycline, and incubated for 2 days at 37°C in an enhanced CO_2 atmosphere (GasPak EZ CO_2 container system; Becton Dickinson, USA). Different dilutions were plate spread, depending on whether or not plates contained tetracycline. The plates contained 0, 4, 12 or 24 μg/ml of tetracycline (chosen according to previous literature) (Tian et al., 2012). The isolated bacteria were randomly picked, and we emphasized on colonies with the morphology of *G. apicola* and *S. alvi*. These were then used in qPCR screening for identification of *G. apicola* and *S. alvi* as described by Martinson et al. (2012) then later for screening on Tc^r genes (method described for midgut/pyloric samples) and then subsequently genome sequenced.

DNA extraction was performed by homogenizing ~20 μl of bacteria culture in plastic tubes with 500 μl S.T.A.R buffer and ~0.20 g <106-μm glass beads using the MagnaLyser at 2 × 20 s at speed 6,500 rpm with 1-min rest, at 4°C between runs. The MagLGC™ Total Nucleic Isolation kit (LGC Genomic, Germany) for blood samples was then used, in combination with the KingFisher Flex robot (Thermo Scientific, USA), for an automated DNA extraction protocol. Qubit dsDNA HS Assay Kit was used to measure the DNA concentration, which was normalized to 0.2 ng/μl.

2.4 | Genome sequencing; sequence assembly, annotation and analysis

Whole-genome shotgun sequencing was performed using Nextera XT protocol (Illumina, USA) and a v3 paired-end sequencing kit (2× 300 bp), following the manufacturer's guidelines for the Illumina MiSeq. The 300-bp paired-end reads were quality-filtered using Trimmomatic (Bolger, Lohse, & Usadel, 2014) and assembled using SPAdes version 3.0 (using default parameters and -k 99,127 and "–careful") (Bankevich et al., 2012). Contigs with length less than 500 bp were removed. Coding DNA sequences and ribosomal RNAs were predicted using PROKKA version 1.0 (Seemann, 2014). The GenBank files were imported into GENEIOUS software version 8.1 (USA) for subsequent analyses (Kearse et al., 2012).

RESFINDER version 2.1 (Zankari et al., 2012) was used to identify resistance genes in all genomes with a threshold of 90% and a minimum length set to 80%. In addition, each identified contig containing a Tc^r gene was run in BLAST searches using the none-redundant (nr) database at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cg).

CRISPR recognition tool (CRT) version 1.2 as GENEIOUS plug-in (Bland et al., 2007) was used to search for genome repeats for all

G. apicola and *S. alvi* (settings; minimum three repeats). The CRISPRFinder online tool (http://crispr.i2bc.paris-saclay.fr) was used to extract spacer sequences (Grissa, Vergnaud, & Pourcel, 2007) for subsequent local Blast searches against PhageDB (NCBI) in GENEIOUS version 8.1. Spacers were aligned, and those occurring in more than one genome were used for phage cross-infection mapping against phylogenetically similar genomes.

Alignment of assembled contigs harbouring Tc^r genes was performed using the Mauve plug-in for GENEIOUS version 8.1, which addressed the location of Tc^r genes within whole contigs (Darling, Mau, & Perna, 2010). Both 16S rRNA genes and pectin-degrading genes were extracted from annotated gbk files in GENEIOUS version 8.1. Maximum-likelihood (ML) trees of 16S rRNA genes and pectin-degrading enzymes were made using PhylML (Guindon et al., 2010) plug-in for GENEIOUS version 8.1 from MUSCLE (Edgar, 2004) alignments of aa sequences. Both trees were made from protein alignments with a WAG model and 100 bootstraps with optimization for topology (NNI), length and rate.

2.5 | Subgroup identification and genome-content comparisons

We performed a pan-core genome analysis on all the genomes sequenced. The protein coding sequences of all 48 *G. apicola* and 22 *S. alvi* were compared separately with an all-against-all approach, using blastp (Camacho et al., 2009). All the distances between each gene in each genome against all the genes in all the other genomes were used to construct a panmatrix using the R package MICROPAN (Snipen & Liland, 2015) (http://cran.r-project.org/). CDSs were grouped in clusters, using a threshold of 0.75 and complete linkage.

Single orthologue genes (defined as genes present in only one copy per genome and obtained from the panmatrix), present in all the genomes were used to construct the phylogenetic relationship. For all the gene clusters containing single orthologue genes present in all the genomes, the nucleotide sequences were translated to amino acids, aligned using DECIPHER r-package (Wright, 2015), and back-translated to nucleotide sequences. All the alignments were then concatenated into a single file containing all the aligned, single-copy, orthologue genes. ML tree was made using CLC genomic workbench with Jukes–Cantor distance, four substitution rate categories and empirically determined gamma substitution parameter with a bootstrap of 100. *Snodgrassella alvi* wkB2 and *G. apicola* wkB1, obtained from NCBI, were used as reference genomes.

Distances between genomes and subgroups were calculated as per cent nucleotide identity between pairs of sequences from the MUSCLE (Edgar, 2004) alignment used for the ML phylogeny, using GENEIOUS version 8.1.

A Venn diagram analysis was performed separately for both *G. apicola* and *S. alvi* to find subgroup-specific gene content. We divided the genomes of *G. apicola* into the four subgroups identified in its phylogeny analysis (Figure 3a), and by the presence of ARG (Table S3). For *S. alvi*, we used the presence of ARG to divide the groups (Table S3). Venn diagrams were computed, using R statistics and the limma package (Ritchie et al., 2015) (http://cran.r-project.org/). Additionally, associations between genome content and subgroups, and specific tetracycline genes, were identified for *G. apicola* using a genomewide association analysis (GWAS) approach. The same groups as for the Venn diagram analysis were used (GWAS results). The panmatrix output from ROARY (Page et al., 2015), with default settings, was applied as input for the GWAS analysis, using SCOARY (Brynildsrud, Bohlin, Scheffer, & Eldholm, 2016).

BLAST Ring Image Generator (BRIG) (Alikhan, Petty, Ben Zakour, & Beatson, 2011) was used to visualize genome comparisons of whole genomes. Comparisons were made against reference genomes *G. apicola* (wkB1, CP007445) and *S. alvi* (wKB2, CP007446), with upper and lower identity thresholds of 90% and 70%, respectively. A BLAST *E*-value of 0.05 was used for all comparisons. Mauve Contig Mover (Rissman et al., 2009) was applied to order the genome (TSA1) towards the reference genomes of *G. apicola* (wkB1, CP007445). The ordered concatenated contigs were used as a reference genome for comparing the antibiotic-resistant gene-mapping visualization of *G. apicola* subgroups.

2.6 | Tetracycline MIC

All *G. apicola* and *S. alvi* strains selected for genome sequencing were tested for MIC on tetracycline using MIC Test Strips (Montebello Diagnostics AS, Norway). The strains were inoculated (EUCAST guidelines, http://www.eucast.org/) on Müller Hinton agar with 5% horse blood, and incubated for 2 days at 37°C in an enhanced CO_2 atmosphere (GasPak EZ CO_2 container system; Becton Dickinson, USA).

2.7 | Statistical tests

Differences between the prevalence of specific Tc^r genes in Arizonan and Norwegian samples were determined using binominal testing and Student's *t* test. Cq values for *tet*(B) and *tet*(H) were normalized against Cq values of the 16S rRNA gene by ratio calculations as previously described in Ludvigsen, taking into account differences in PCR amplification efficiency using the LinRegPCR software (Ruijter et al., 2009). Differences in gene prevalence between Arizonan and Norwegian samples for *tet*(B) and *tet*(H) were determined using Student's *t* test in R. Variation in MIC, related to Tc^r genes, was analysed with ASCA ANOVA, using PLS Toolbox in Matlab. Cluster analysis by MIC values of all genome-sequenced *G. apicola* was performed using *K*-means clustering and ANOVA, in combination with *t* test, to determine differences in mean of MIC values within different Tet B-interspaced mutations, and differences in genome size and CG content. Both analyses were performed within STATISTICA version 12 (Dell Statistica, USA) and Excel.

2.8 | Calculation of ANIb and Tetra

We calculated the ANIb (average nucleotide identity—Blast) and Tetra (tetranucleotide signature) parameters implemented in the online tool JSpeciesWS (Richter, Rossello-Mora, Oliver Glockner, &

Peplies, 2016) between selected genomes in each subgroup. ANIb is a distance measure of the average nucleotide similarity, based on pairwise alignment of orthologue genes between two draft genomes. This method correlates well with DNA–DNA hybridization (DDH) results and can calculate species-distance delineation (Konstantinidis & Tiedje, 2005; Richter & Rosselló-Móra, 2009). A DDH of ~60%–70% represents an ANIb identity of 95%–96%, and an ANIb of <83% supports a new species. Tetra is a correlation measure of co-occurrence of tetranucleotides in two genomes and distinguishes between bacteria strains (Burall, Grim, Mammel, & Datta, 2016). The Tetra measure correlates with ANIb at the species delineation (ANIb = 95%–96%, Tetra ≥ 0.999) and can, together with ANIb, support if two genomes are the same species.

3 | RESULTS

3.1 | Prevalence of tetracycline resistance in Arizonan and Norwegian honeybees

We screened midguts from 90 bees from each of the two apiaries for *tet*(B), *tet*(C), *tet*(W) and *tet*(H) by quantitative real-time PCR. In the Norwegian samples, only *tet*(B) and *tet*(H) were identified, while *tet*(C) and *tet*(W) were also detected in the Arizonan bees (Figure 2a). For both locations, *tet*(B) was the most prevalent gene, at 22% for Norwegian bees and 100% for Arizonan bees ($p < 10^{-11}$, *t* test). There were also major differences in prevalence for *tet*(H), at 7% for Norwegian bees and 73% for Arizonan bees ($p < 10^{-11}$, *t* test). Moreover, the Arizonan samples had a higher prevalence of *tet* (B) in each bee compared to Norwegian bees ($p < 10^{-7}$, *t* test) (Figure 2b). In contrast, *tet*(H) prevalence was not significantly different between the two research apiaries but showed a relatively low quantity within each bee for both locations ($p = .12$, *t* test).

From 130 cultured bacterial strains, 94 strains were identified as *G. apicola* or *S. alvi* by a species-specific PCR, as described in Martinson et al. (2012). Selective bacterial culturing from plates, with and without tetracycline, confirmed the high prevalence of Tc^r in the Arizonan bees compared to the Norwegian bees (Figure 2c). In fact, all *Gilliamella* and *Snodgrassella* ($n = 51$) isolates from Arizonan bees had tetracycline-resistant genes, despite some isolates being selected from plates without tetracycline (Table S1). *Tet*(B) was the main determinant in Arizonan *Gilliamella* strains, and 23% was positive for both *tet*(B) and *tet*(H). Moreover, *tet*(H) and *tet*(B) genes always co-occurred. This was not true for the Norwegian strains, which had either *tet*(B) or *tet*(H). *Tet*(H) was the dominating determinant for the Arizonan *S. alvi* strains, and no strains harboured both genes. No *tet* (B) or *tet*(H) genes were identified in the Norwegian *S. alvi* strains.

3.2 | Genome sequencing, assembly and annotation of strain isolates

QPCR screening results of bacterial strains were used to select bacteria for genome sequencing (Table S1). A total of 48 *Gilliamella* and 22 *S. alvi* were sequenced (Table 1). To cover strain diversity (based on Tc^r gene content and phenotype), 27 Norwegian and 21 Arizonan

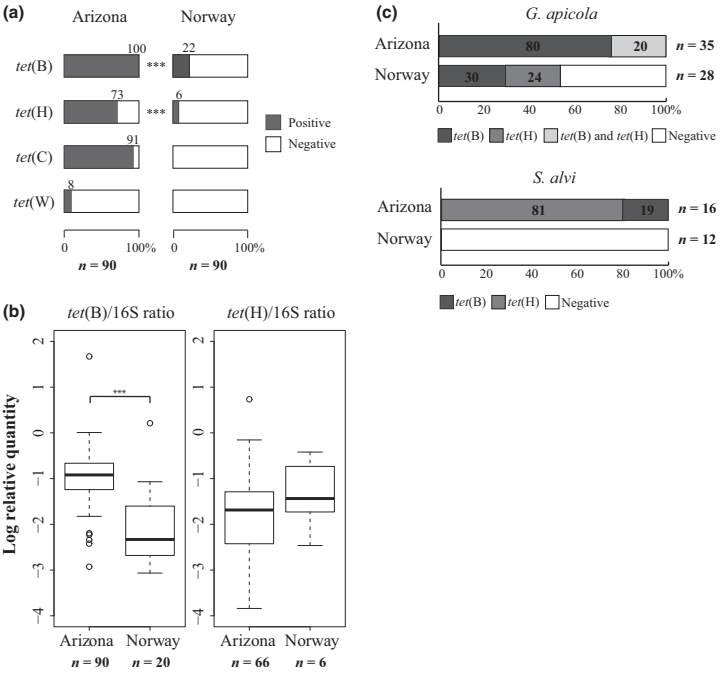

FIGURE 2 qPCR results. (a) Screening of honeybee midguts/pylori. Per cent positive samples detected for four different tetracycline genes is shown with numbers above each box. (b) Comparison of tetracycline gene prevalence between Arizona and Norway. Boxplot of log relative quantity of the *tet*(B) and *tet*(H) genes of all positive samples (midgut/pylori) shown in (a). Tetracycline gene prevalence was normalized against the total 16S rRNA gene to account for variation in midgut/pyloric size during sampling. ***p value < .01. (c) Fraction of *tet*(B) and *tet*(H) resistance among total isolated *Gilliamella apicola* and *Snodgrassella alvi* isolates. Detailed description of isolated strains is shown in Table S1

TABLE 1 Selected strains for whole-genome shotgun sequencing

	Total number of isolates	Number of isolates with resistant genes	Number of isolates with *tet*(B)	Number of isolates with *tet*(H)	Number of isolates with *tet*(B) & *tet*(H)	Number of isolates without resistant genes
Norwegian *Gilliamella apicola*	27	19	11	8	–	7
Arizonan *G. apicola*	21	21	16	0	5	–
Norwegian *Snodgrassella alvi*	9	–	–	–	–	9
Arizonan *S. alvi*	13	13	3	10	–	–

Gilliamella strains, and nine Norwegian and 13 Arizonan *S. alvi* strains were selected. Assemblies of *Gilliamella* and *S. alvi* genomes varied between 54 and 655 contigs per genome (Table 2). From the annotated contigs, bacterial 16S rRNA gene sequences for all bacteria strains were used to assign taxonomy by BLAST searches (NCBI, nr database). All *Gilliamella* strains showed 99% or more identity to *G. apicola* wkB1 (CP007445) (Table 2), and all *Snodgrassella* strains showed 99.9% or more identity to *S. alvi* wkB2 (CP007446) (Table 3). WkB1 and wkB2 were used as references in subsequent analysis of our data.

3.3 | Genome comparison of *G. apicola*

The phylogenetic relatedness between 48 strains of *G. apicola* was inferred from ML phylogeny of 1,041 single orthologue genes. These 48 strains could be divided into two main groups. Each of these could then be split into two, making four distinct subgroups (Groups 1.1, 1.2, 2.1 and 2.2) (Figure 3a). We detected Norwegian and Arizonan strains in each subgroup except one (Group 2.2), in which there were only two Norwegian strains. Therefore, three predominating subgroups were maintained across geographical locations.

The average similarity between the two main groups was 80% (measured as average per cent nt identity), which is significantly different from the within-group similarity of 96.8% ± 2.5% and 96.7% ± 2.5% (p value < 10^{-19}, t test) for Group 1 and Group 2, respectively (Figure 3a, Table S9).

Moreover, in-depth investigation of the genomes within the different subgroups showed that the average genome size and GC content were significantly different between Group 1 and Group 2 (Table 2). The average genome sizes, for Group 1 and Group 2, were 3.02M bp ± 81K bp and 2.53M bp ± 84K bp (p value < 10^{-23}, t test), respectively, resulting in a difference of genome size between of approximately 400K bp. The GC contents were 33.8% ± 0.2 and 34.7% ± 0.15 (p value < 10^{-22}, t test) for Group 1 and Group 2, respectively. Across all subgroups, genome size and GC content displayed an apparent correlation with phylogeny (Figure 3b) (p value < 10^{-29}, ANOVA), and all subgroups were significantly different from each other (largest p value all groups against all <10^{-6}, t test Bonferroni corrected).

Due to the low similarity between the two main groups in our data set, we cross-checked the distance between these two groups against a previously published algorithm and online tool JSpeciesWS (Richter et al., 2016), and calculated the ANIb (average nucleotide identity—Blast) and Tetra (tetranucleotide signature). The Tetra measurement correlates with ANIb at the species delineation (ANIb = 95%–96%, Tetra ≥ 0.999), so can be used with ANIb to indicate if two genomes are the same species. In our data, the average ANIb and Tetra between Group 1 and Group 2 were 79.6% ± 0.2% and 0.964 ± 0.002, respectively (Table S9). This suggests that they belong to different *Gilliamella* species.

We also applied ML phylogeny for 1,445 bp of the 16S rRNA gene for comparison, as the 16S rRNA gene was the choice of bacterial identification for many years. The 16S rRNA phylogeny also showed the same strains divided into the same two main groups, but no clear subgroups were detected (Fig. S1).

Comparison of the total gene content of all 48 *G. apicola* genomes resulted in 4,408 pan-genome genes and 1,480 core genes (Fig. S2a). The pan-genome rarefaction curve showed a tendency for gene saturation, indicating that the gene diversity within *G. apicola* is represented in our data set. From this gene present–absent phylogeny of *G. apicola* genomes identified by pan-genome analysis, the clustering into four subgroups was supported. Approximately 1,000 genes differed between Group 1 and Group 2, and 600–800 genes differed between the four subgroups (Fig. S3). Group 1 had approximately 400 more CDSs than Group 2, which corresponded to the smaller, average genome size also found in Group 2 (Table 2). Using BRIG, a visualization of whole-genome comparisons between all draft genomes within the subgroups revealed distinct missing parts within the draft genomes compared to the reference wkB1 *G. apicola* (Figure 4). As BRIG only compared genes that are present in the reference genome, we also performed a Venn diagram analysis to identify subgroup-specific genes. Group 1 had 214 unique genes (including unique genes in Subgroups 1.1 and 1.2), and Group 2 had 462 unique genes (including unique genes in Subgroups 2.1 and 2.2) (Table S2) that were not annotated as hypothetical.

We also included a GWAS, using SCOARY to investigate the statistical association of genes towards the subgroups. The enzymes pectate A lyase (PL1) and pectate disaccharide lyase (PL9), previous detected in honeybee *G. apicola* and the basis for discriminating this genera into different functional groups (Engel et al., 2012), were only associated with Subgroups 1.1 and 1.2 (Table S4, Supplement GWAS_Tet_genes & GWAS_Groups). Using the Venn diagram analysis of subgroup-specific gene content, we identified one more pectin-degrading enzyme, endo-polygalacturonase (GH28), which was also only found, in Group 1 (Fig. S4a, Table S2). We also detected a fourth enzyme, exo-polygalacturonase (also a GH28), in all genomes

TABLE 2 16S rRNA gene identity blast results, genomic content and assembly info for each *Gilliamella apicola* genome within the four different subgroups

	16S rRNA gene		Genome info			Assembly info				Group
Name	% ID[a]	Length (bp)	CDS	Size (bp)	GC %	Contigs[b]	N50	L50	tRNA	1
NO5	99.5	1,534	2,709	3,063,250	33.6	143	55,544	15	49	1.1
NO6	99.5	1,534	2,710	3,064,360	33.6	129	61,286	14	49	1.1
NO10	99.5	1,534	2,664	3,051,496	33.7	398	15,301	55	49	1.1
NO8	99.5	1,534	2,647	3,037,207	33.7	529	10,743	79	47	1.1
N-9-4	99.5	1,534	2,713	3,069,680	33.6	118	60,516	14	49	1.1
N-12-12	99.5	1,534	2,718	3,076,431	33.6	147	55,906	18	49	1.1
N-15-12	99.5	1,534	2,682	3,049,463	33.7	264	25,498	36	39	1.1
N-G5	99.8	1,534	2,635	3,063,691	33.5	99	89,376	12	40	1.1
A-7-12	99.9	1,534	2,805	3,143,865	33.6	91	74,031	14	46	1.1
A-1-24	99.9	1,534	2,790	3,144,522	33.6	157	48,470	22	46	1.1
A-9-12	99.9	1,534	2,794	3,145,014	33.6	189	33,789	25	45	1.1
A-2-24	99.9	1,534	2,791	3,136,687	33.6	129	50,570	21	46	1.1
AW13	99.9	1,534	2,747	3,128,857	33.6	371	16,368	57	45	1.1
A9	99.9	1,534	2,587	3,028,920	33.5	362	16,586	50	45	1.1
A8	99.9	1,534	2,622	3,046,518	33.5	226	30,769	29	43	1.1
N-22	99.4	1,534	2,610	2,979,016	33.7	304	75,752	11	47	1.1
N-28	99.4	1,534	2,600	2,974,046	33.7	92	20,593	41	47	1.1
A-8-12	99.5	1,467	2,580	2,929,550	34.0	110	70,012	13	48	1.2
Aw-17	99.5	1,534	2,580	2,928,228	34.0	130	52,252	17	50	1.2
A-7-24	99.5	1,534	2,607	2,943,177	34.0	107	67,283	13	50	1.2
A-12-12	99.5	1,467	2,610	2,942,700	34.0	102	74,588	11	50	1.2
AW11	99.5	1,534	2,580	2,942,846	34.0	328	19,809	42	49	1.2
N4	99.5	1,534	2,559	2,910,729	33.9	127	50,596	14	48	1.2
N2	99.5	1,534	2,557	2,908,650	34.0	88	76,937	12	48	1.2
N6	99.5	1,534	2,561	2,905,171	33.9	82	87,042	9	40	1.2
N10	99.5	1,534	2,550	2,898,993	34.0	202	27,039	29	48	1.2
AM4	99.4	1,442	2,618	3,063,854	33.9	166	44,082	19	38	1.2
AM6	99.4	1,442	2,513	3,003,378	34.2	655	8,403	103	49	1.2
										2
NO3	99.0	1,467	2,217	2,523,863	34.7	54	98,475	8	44	2.1
NO16	99.2	1,444	2,218	2,526,241	34.7	57	94,347	8	47	2.1
NO15	99.0	1,448	2,215	2,520,720	34.7	50	110,209	7	46	2.1
NO14	99.2	1,444	2,214	2,520,079	34.7	75	64,969	12	46	2.1
NO12	99.2	1,448	2,212	2,527,217	34.7	68	72,003	11	46	2.1
NO1	99.0	1,534	2,207	2,516,229	34.7	156	31,497	22	46	2.1
NO13	99.2	1,448	2,226	2,533,114	34.7	60	88,091	8	47	2.1
NO4	99.2	1,444	2,204	2,519,429	34.7	214	19,667	34	45	2.1
A-4-12	99.3	1,442	2,355	2,626,243	34.7	104	72,898	11	35	2.1
N-G3	99.3	1,543	2,181	2,482,892	34.7	83	66,116	11	45	2.1
N-G1	99.3	1,534	2,178	2,478,810	34.7	65	99,129	7	44	2.1
N-G4	99.1	1,534	2,169	2,436,247	34.6	142	36,046	20	46	2.1
A-TSA4	99.0	1,467	2,176	2,460,968	34.8	86	68,662	9	47	2.1
A-TSA2	99.0	1,467	2,179	2,462,160	34.8	72	80,313	9	47	2.1
A-TSA1	99.0	1,467	2,176	2,467,442	34.8	79	72,217	11	47	2.1

(Continues)

TABLE 2 (Continued)

Name	16S rRNA gene		Genome info			Assembly info				Group 1
	% ID[a]	Length (bp)	CDS	Size (bp)	GC %	Contigs[b]	N50	L50	tRNA	
A-TSA3	99.0	1,467	2,178	2,467,212	34.8	77	79,540	10	47	2.1
AM1	99.0	1,534	2,155	2,438,283	34.6	120	37,727	17	45	2.1
A7	99.0	1,534	2,389	2,710,452	34.2	125	47,965	16	45	2.1
N-W3	99.3	1,534	2,484	2,674,958	34.4	90	62,957	13	44	2.2
N-G2	99.3	1,534	2,531	2,709,192	34.4	77	92,081	9	44	2.2

[a]Against *G. apicola* wkB1 (CP007445).
[b]Number of contigs in final draft after filtering.

TABLE 3 16S rRNA gene identity blast results and ANIb values, genome info and genome assembly info for each *Snodgrassella alvi* genome

Name	16S rRNA gene		Genome info				Assembly info			
	% ID[a]	Length (bp)	ANIb[b]	CDS	Size (bp)	GC %	Contigs[c]	N50	L50	tRNA
N-S2	99.9	1,534	97.32	2,066	2,421,229	41.2	73	61,427	13	39
N-S4	99.9	1,534	97.32	2,070	2,421,486	41.2	38	133,936	6	49
N-23	99.9	1,534	97.32	2,062	2,421,693	41.2	128	38,667	21	37
N-S5	99.9	1,534	97.32	2,074	2,417,615	41.2	77	56,721	13	43
N-W4	99.9	1,534	97.33	2,070	2,421,251	41.2	75	55,544	13	41
N-S1	99.9	1,534	97.33	2,059	2,420,873	41.2	98	50,074	18	52
A-9-24	99.9	1,534	97.04	2,207	2,501,107	41.2	62	83,544	10	46
A-2-12	99.9	1,534	97.04	2,204	2,502,682	41.2	75	64,388	14	50
A-5-24	99.9	1,534	97.08	2,188	2,490,743	41.3	172	28,499	30	41
A-1-12	99.9	1,534	97.04	2,205	2,502,286	41.2	58	91,493	10	46
A-11-12	99.9	1,534	97.04	2,202	2,500,985	41.2	90	53,637	12	49
A-10-12	99.9	1,534	97.04	2,205	2,501,655	41.2	63	65,004	12	44
A2	99.9	1,534	97.08	2,112	2,425,186	41.2	84	56,386	12	42
A5	99.9	1,534	97.06	2,115	2,430,376	41.2	120	40,275	20	45
A11	99.9	1,534	97.08	2,108	2,430,265	41.2	122	42,222	17	47
A3	99.9	1,534	97.08	2,102	2,428,731	41.2	109	44,108	18	53
A12	99.9	1,534	98.03	2,059	2,399,919	41.3	214	19,825	35	43
N-S3	100	1,535	97.06	2,164	2,463,518	41.3	79	70,149	11	47
N-W7	99.9	1,534	97.05	2,116	2,423,318	41.3	62	70,113	10	50
N9	99.9	1,535	97.02	2,062	2,403,335	41.2	129	42,232	23	48
Aw-20	100	1,535	96.94	2,188	2,498,497	41	65	75,805	10	53
Aw-18	100	1,535	96.96	2,190	2,497,111	41	88	55,786	12	52

[a]Against *S. alvi* wkB2 (CP007446).
[b]Calculated against *S. alvi* wkB2 (CP007446).
[c]Number of contigs in final draft after filtering.

except for one strain (A7) in Subgroup 2.1 (Table S4). For the PL1 and PL9 enzymes, ML phylogenies of protein-aligned sequences showed that the evolution of these enzymes correlated with the specific lineages identified within Group 1 of *G. apicola* (Fig. S5). The GH28 enzyme only found in Group 1 did not strictly follow the evolution within the subgroups, and the GH28 (exo-polygalacturonase) found in both groups was only evolutionally similar to the orthologue phylogeny within Group 1.

Our 16S rRNA phylogeny included previously identified strains, with and without the capacity to phenotypically break down pectin, which also harboured the PL1 and PL6 enzymes (Engel et al., 2012). We found the strains to be separated into the two main groups identified in this work: the pectin-degrading strain in Group 1 and the non-pectin-degrading strain in Group 2 (Fig. S1).

Further elaborating on the GWAS analysis, we used the association statistic of the pectin-degrading enzymes (100% sensitivity or specificity, and FDR-corrected p value < 0.01) as the cut-off value for searches of other significantly associated subgroup-specific genes. We then compared these to the Venn diagram analysis results. Both analyses showed that Subgroup 2.2 harboured all

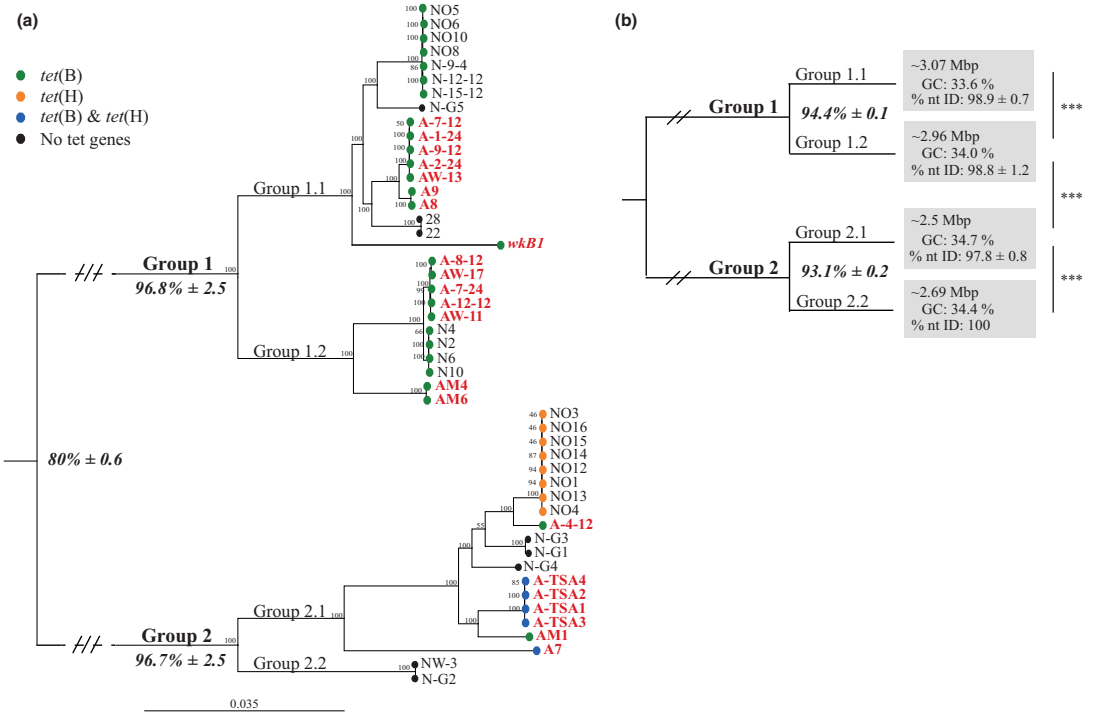

FIGURE 3 (a) Maximum-likelihood (ML) tree made from 1,041 single orthologue genes of 48 *G. apicola* strains. Bootstrap values >50 are shown at each node. Main groups and subgroups are in text above each node division. Arizonan *Gilliamella apicola* are shown in red, and Norwegian in black. Per cent nt similarity distance between the two main groups are shown at the dividing node in italics, and within-group distances are shown below the group name in italics. Distances are cut short and shown as interrupted lines to fit figure window. Colour coding at strain nodes shows tetracycline gene presence and is explained in the upper left corner. Reference *G. apicola* wkB1 is shown in red italics. Bootstrap values are given in per cent of 100 iterations. (b) A graphical representation of ML tree with detailed information about the four subgroups. Distance similarity between subgroups is shown at node division, and average genome size, GC content, and within-distance similarity for each subgroup are shown within the shaded boxes. ***p value < .01

enzymes needed for urea breakdown (urease-alpha, -beta, and -gamma, urease-accessory proteins; UreD UreE, UreF, and UreG, and urease operon transcriptional activator) (Table S2, Fig. S4a). We were not able to extract genes belonging to specific pathways for the other subgroups.

We then used the number of CRISPR arrays within the genomes and the corresponding number of incorporated spacers to address phage exposure and investigate whether this information could further divide the strains beyond the four subgroups. We could distinguish between strains at the leaf nodes (last common node) within the phylogeny (Table S5). On average, Group 2 had more arrays within each genome, but not more spacers than Group 1 (p value arrays = $<10^{-5}$, p value spacers >0.05, t test).

When comparing the spacer sequences, no overlap between spacers was detected when exceeding the last common node. In general, there was high diversity between the strains with regard to type of CRISPR spacer, but no systematic pattern could be detected beyond the last common node (data not shown). We also aligned spacer sequences to detect phage cross-infection patterns within node-specific strains. Within Subgroup 1.2, identical spacers could be identified in both Arizonan and Norwegian strains, indicating a common niche with similar exposure events at both locations for this subgroup. As the CRISPR arrays are hypervariable and strain-specific regions, different spacers could be used to identify whether we had sequenced the same strain twice, by aligning spacers from genomes with the same number of arrays and spacers incorporated. Genomes N-22 and N28 harboured completely identical spacers, as did NO5 and NO6, N-9-4 and N-12-12, and NW-3 and N-G2.

3.4 | Genome comparison of *S. alvi*

The phylogenetic relatedness between the 22 strains of *S. alvi* was inferred from ML analysis of 1,588 single orthologue genes. Those strains isolated from bees collected in Norway were mixed with the strains isolated from bees collected in Arizona (Figure 5). There seemed to be several linages among these isolates, but no clear

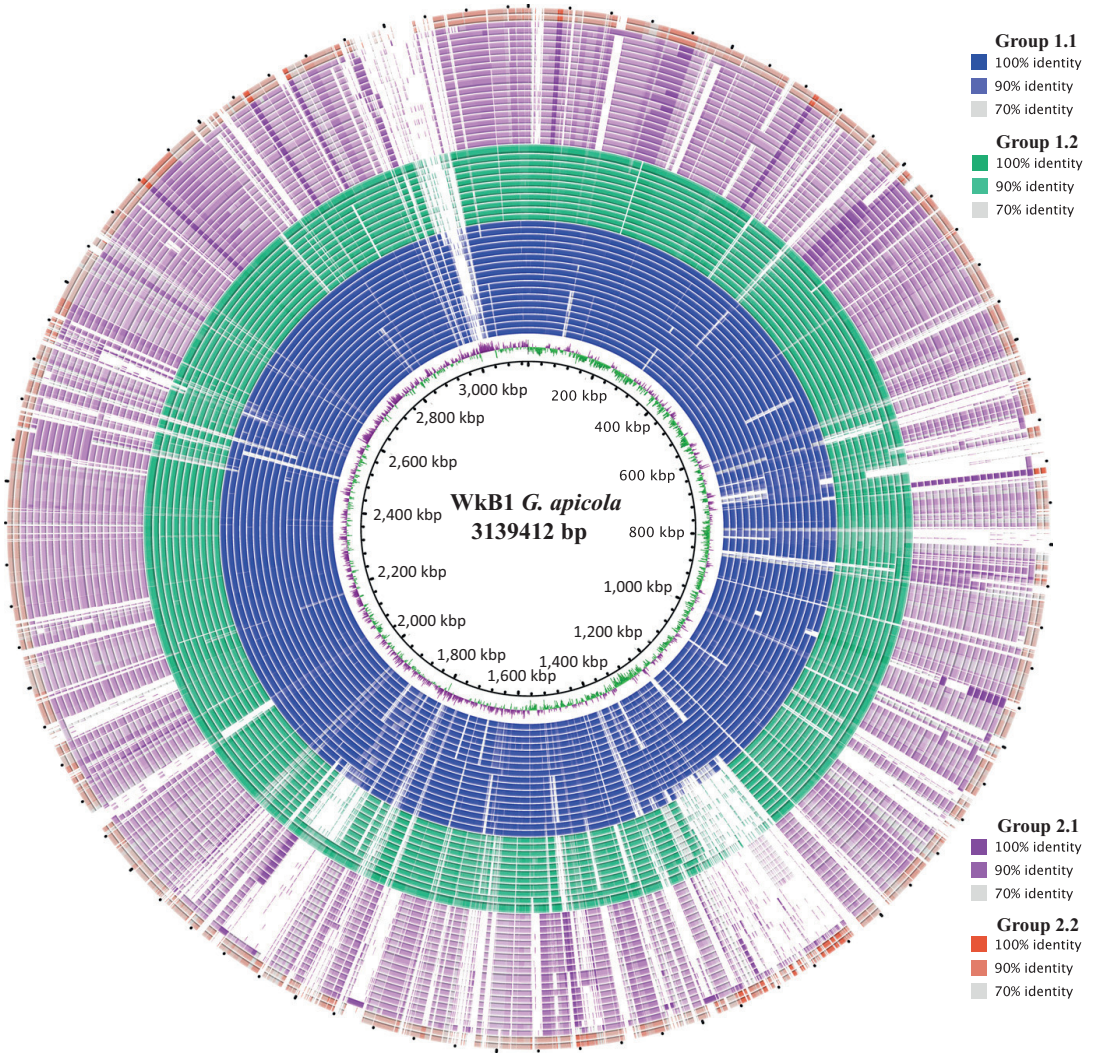

FIGURE 4 BLAST Ring Image Generator visualization from genome comparison of 48 draft genomes within the four *Gilliamella* subgroups, against the reference *G. apicola* wkB1. Genomes size of *G. apicola* wkB1 is shown in the middle, and the similarity threshold with colour key is explained in the top right and bottom right part of the figure for group 1 and group 2, respectively. The colour coding is continuous from 100% to 70%, and the 90% identity colour is for colour reference only. GC skew is shown as the inner-most ring with alternating purple and green colours: Purple = GC skew (−), Green = GC skew (+). From inner to outer rings, the order follows the order from Figure 3a—from top to bottom

subgroups could be identified as the average per cent nt similarity between the two linages separated at the first dividing node was 96.8% ± 0.7%. The within-group distances for these two linages were 97.8% ± 1.2% and 96.9 ± 1.4% for the upper and lower, respectively. Genome size and GC content were almost identical for both linages, with an average of 2.45M bp and 41.2% (Table 3).

Comparison of total gene content between all 22 *S. alvi* genomes resulted in 2,762 pan-genome genes and 1,692 core genes. The pan-genome rarefaction curve showed a tendency for gene saturation, indicating that the gene diversity within *Snodgrassella* is represented in our data set (Fig. S2b). The gene present–absent phylogeny of *S. alvi* (Fig. S6) showed differences in gene content between the linages. The BRIG visualization against the reference wkB2 detected a distinct lineage-specific evolution (Figure 6), but no significant difference between the genome linages and the number of CDSs was detected (Table 3). For the CRISPR arrays, we identified one array

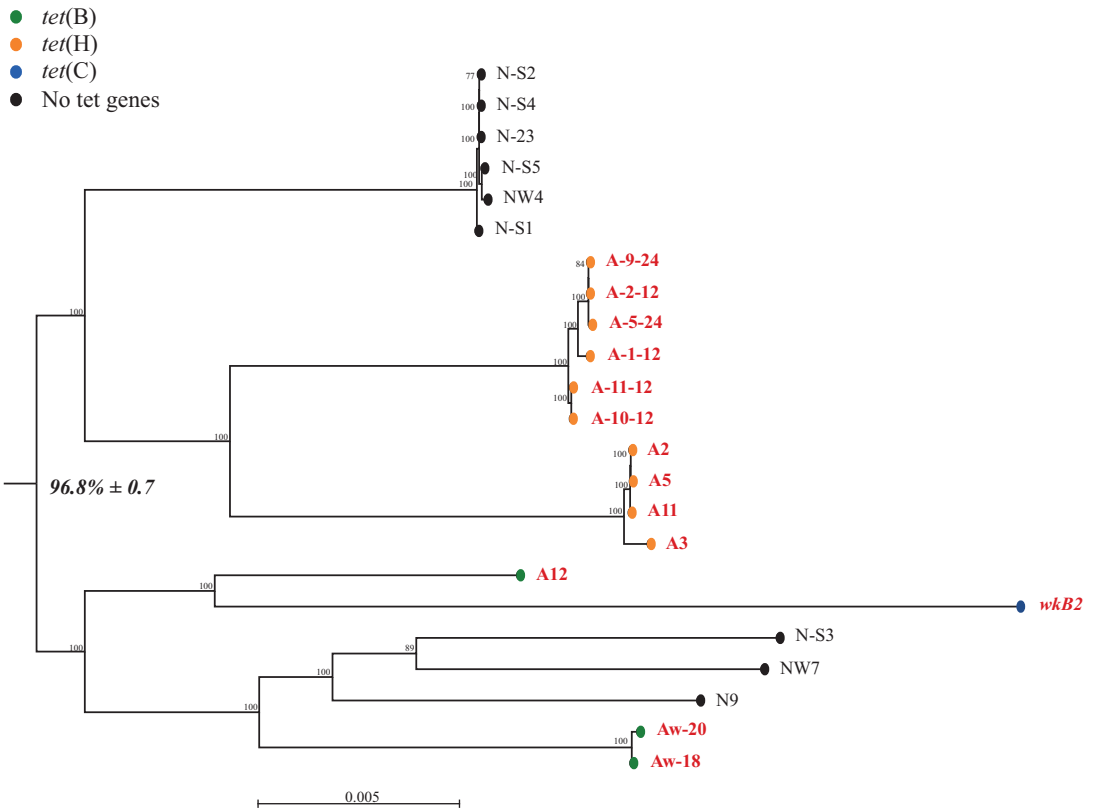

FIGURE 5 Maximum-likelihood tree made from 1,588 single orthologue genes of 22 *Snodgrassella alvi* strains. Reference wkB2 is included and shown in red italics. Arizonan *S. alvi* are shown in red, and Norwegian in black. Per cent nt similarity distance between the two linages separated by the first common node are shown at the dividing node in italics. Presence of tetracycline genes within strains has different colour-coded nodes and is explained in the upper left corner. Bootstrap values are given in per cent of 100 iterations

and three spacers in five of the nine Norwegian *S. alvi* strains (Table S5).

3.5 | Antibiotic resistance within *G. apicola*

We mapped the Tc^r gene information onto the phylogeny (Figure 3a) to connect the Tc^r gene content with the detected strain variation. Distribution of the different Tc^r genes displayed subgroup-specific clustering. *Tet*(H) genes from both Norwegian and Arizonan strains were only identified in Subgroup 2.1. The Norwegian strains without Tc^r were intermixed with the Arizonan ones within the two main subgroups.

Searching for ARG with RESFINDER version 2.1, we identified the two Tc^r genes *tet*(B) and *tet*(H) in our *Gilliamella* strains. The *tet*(B) gene was identified with 100% identity towards a reference (AP000342, *Shigella flexneri*), although this gene also showed variants within the Arizonan population among the *G. apicola* strains, with only 99.92% identity (Table S3). The reference gene was identified as a part of the transposon Tn10, which contains *tet*(B) genes (Fig. S7). (Chalmers, Sewitz, Lipkow, & Crellin, 2000; Hillen & Berens, 1994). We also identified the Tn10 transposon in our strains and found that all *tet*(B) in *G. apicola* were located within Tn10 transposons. By aligning the *tet*(B) gene, we found a difference of one nt (explaining the 99.92% identity from RESFINDER) in position 455, corresponding to an amino acid (aa) shift in position 152 in the aa sequence from Alanine (A) to Valine (V) in 11 of 24 strains. All the Norwegian strains had the A amino acid (Table S6).

Due to lack of variation in the *tet*(B) gene, we compared the whole Tet B determinant (*tet*(B) gene, *tetR*(B) and the interspace region) (Fig. S7) across all strains of *G. apicola* to identify possible HGT events. We found mutations in five nucleotides (nt numbers 16, 20, 34, 41 and 56) within the interspaced region (operator and promoter region) between *tet*(B) and *tetR*(B), making up six combinatorial variants (Figure 7a, Table S6). To verify the specificity of these mutations within the *Gilliamella* genus, we compared the Tet B

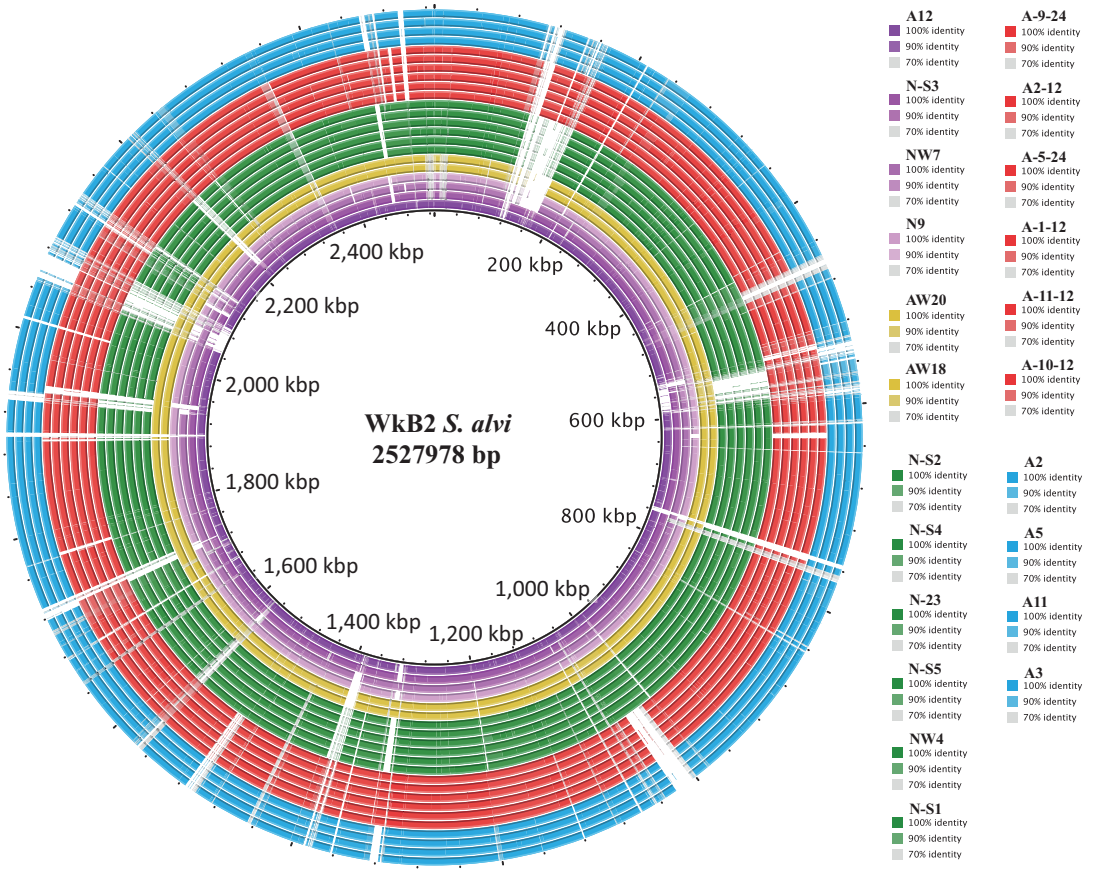

FIGURE 6 BLAST Ring Image Generator visualization from genome comparison of 22 draft genomes of *Snodgrassella alvi* against the reference wkB2 *S. alvi*. Genomes size of wkB2 *S. alvi* is shown in the middle. The colour coding is continuous from 100% to 70%, and colour keys show similarity threshold for each genome from upper left (A12) to lower right (A3) as the ring order from inner to outer. Genomes found within the same linage (compared to Figure 5) are coloured with similar colours

sequence of our isolated *G. apicola* strains to 13 sequences of Tet B from closely related bacterial genera (Table S10). These 13 sequences were all identical when aligned. Only one variant (number 3), identified in some Norwegian isolates, was identical to these sequences. However, all the Arizonan variants were unique to *G. apicola*. More distant genera inferred additional mutations within the promotor/operator region that were different from the ones found in our isolates (Figure 7a). Variants 1 and 5, only found in Arizonan isolates, were both detected within two subgroups among the *Gilliamella* population, while the rest corresponded to a specific subgroup clustering (Figure 7b).

All *tet*(H) genes in our *G. apicoal* strains were identical when aligned. Although the *tet*(H) was identical to a reference gene originally found on a transposon (Tn5706, Y11510, *Pasteurella multocida*) (Table S3), it was not associated with transfer elements in our strains, but was located within the beta-glucosidase operon in the genomic DNA. The Tet H determinant (both *tet*(H) and *tetR*(H) genes) was inserted between the aryl-phospho-beta-D-glucosidase (*BglC*) and inorganic polyphosphate/ATP-NAD kinase (*ppnK*) genes in both Norwegian and Arizonan strains.

The *sul2* gene (sulphonamide resistance) was also found in *G. apicola* strains from both Norway and Arizona (Table S3). We included this information when we searched for orthologues, which were associated with the different ARG in our data set (Tet B, Tet H, and *sul2*), using the GWAS approach and Venn diagram analysis. This analysis showed that all genes normally found in Tn10 (Fig. S7) were significantly associated with Tet B, but no additional genes could be identified. 22 orthologues were associated with the *sul2* gene (p value $< 10^{-8}$), which were all found on the same contig as *sul2* in all our *sul2*-positive strains. Blast results of *sul2* containing contigs had hits towards *sul2*-containing plasmids (IncQ1-like plasmid element) normally present in various *Enterobacteriaceae* genera, with

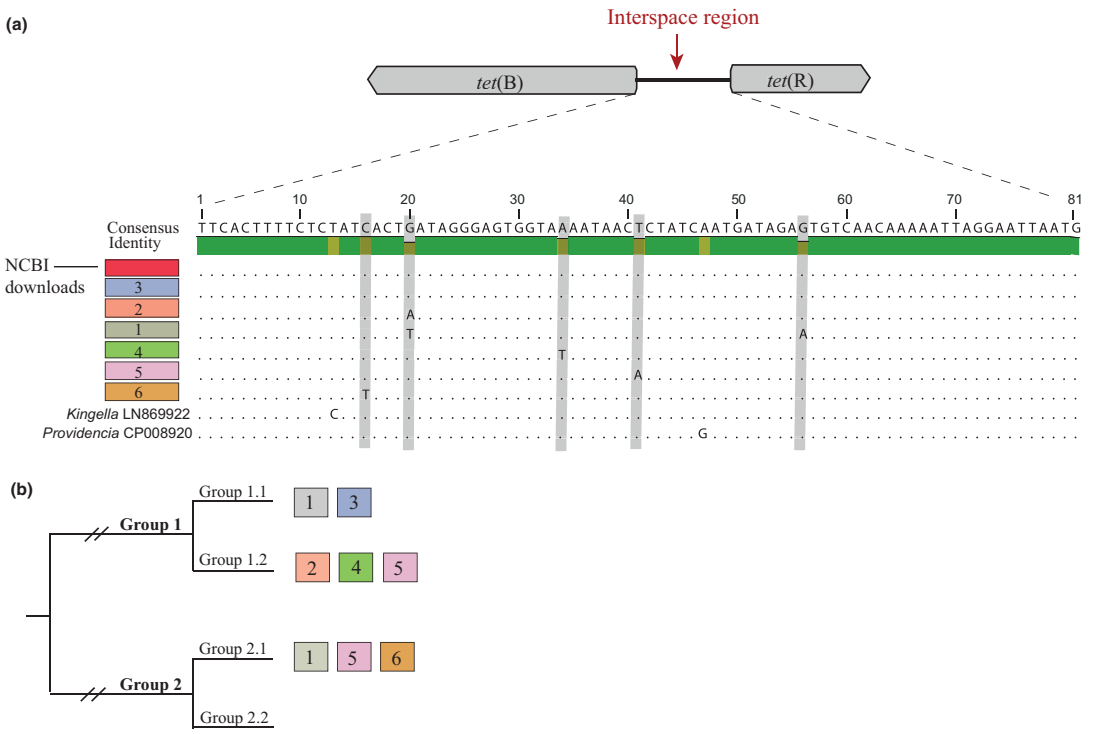

FIGURE 7 Interspace mutational variants found in *Gilliamella apicola*. (a) Dots represent consensus sequence across all compared sequences. Only one sequence is shown for each variant, coded by different colours and different numbers. Mutation positions found in *G. apicola* are highlighted in grey. Table S5 shows NCBI download sequence names and includes *Escherichia* sp., *Serratia* sp., *Salmonella* sp., *Haemophilus* sp., *Shigella* sp., *Citrobacter* sp. and *Klebsiella* sp. (b) A graphical representation of ML tree (Figure 3a) showing how the different variants distribute across the four subgroups of *G. apicola*. Colours and numbers correspond to variants coding shown in (a)

78.5% identity (*E*-value $<10^{-10}$). Despite this, these genes were intergraded in the genomes of *G. apicola*, based on core genes as neighbours (such as GMP synthase [guaA]) and were within contigs of different lengths. Mapping ARG with their insertion sites, according to the concatenated reference genome TSA1 (which harbours Tet B, Tet H, and Sul2) using BRIG, showed that Tet B and Tet H were located far apart and that the *sul2* insertion site was identical for all genomes that harboured *sul2* resistance (Fig. S8).

We performed susceptibility testing for all isolates of *G. apicola*, and found for *G. apicola* with Tet B that the MIC representing Norwegian and Arizonan strains varied between 3 and 16 mg/L, and 12 and 16 mg/L, respectively (Table S11). The MIC range for the Norwegian strains were wider, and *K*-means cluster analysis divided the Norwegian strains into two groups, based on MIC: one with lower and one with higher MIC (3 and 6 mg/L, and 12 and 16 mg/L, respectively, Table S7). These two MIC groups also resided within two subgroups, so showed correspondingly different Tet B interspace mutational variants (p value $< 10^{-11}$, one-way ANOVA) (Table S6). In addition, MIC showed a strong association with the Tc^r gene (p value; *tet*(B) 0.0001, *tet*(H) 0.04, ANCA ANOVA) and a significant difference between MIC of Tet B and Tet H (p value $< 10^{-10}$, t test).

3.6 | Tetracycline resistance within *S. alvi*

We mapped the Tc^r gene information onto the ML phylogeny (Figure 5). Strains with *tet*(B) and *tet*(H) were identified within separate linages, and the Norwegian strains without Tc^r were intermixed with the Arizonan strains.

The *tet*(B) gene in *S. alvi* showed 100% identity towards the reference by RESFINDER, which also included 100% identity towards *tet*(B) in some of the *Gilliamella* strains (Table S3). Contrary to *G. apicola*, in which all *tet*(B) genes were on Tn10 transposons, we only identified the whole Tn10 transposon within *S. alvi* in one of three strains with *tet*(B). In the other two strains, we identified the *tet*(B) gene and its regulatory gene, *tetR*(B), in addition to a truncated tetC gene originating from Tn10 transposons (Fig. S9). These genes were located next to the genomic anguibactin system regulator protein (*angR*), with a tRNA-Phe site in the vicinity. The two *S. alvi* strains with Tet B inserted in

the genomic DNA (strain Aw-18 & Aw-20) differed from the *S. alvi* strain, with the Tet B determinant located on a transposon containing contig (strain A12) by 6 aa in the terminal end of the *tetR*(B) sequence. In comparison, the transposon-associated *tetR*(B) was 4 aa shorter.

The Tet H determinant in *S. alvi* was located within the genome not associated with any apparent transfer elements, although it was identical to a reference (AJ245947, *Pasteurella aerogenes*) previously found on a pPAT1 plasmid (Table S3). It was located near the two core genes: an EamA-like transporter family protein; and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (*dhbA*), a protein involved in siderophore biosynthesis (Fig. S9). The insertion site was the same for all 10 strains with Tet H and located in close proximity to the Adhesin YadA precursor protein (*yadA*), which is a known virulence factor and cell-adhesion protein. The insertion site for Tet B was only located ~25,000 bp apart from the insertion site of Tet H (Fig. S9). GWAS analysis was not performed on the *S. alvi* genomes due to few strains, and the Venn diagram analysis did not yield any conclusive results regarding specific functional differences between the different tetracycline-associated strains (Fig. S4b, Table S8).

Minimum inhibitory concentration variation within *S. alvi* strains showed a strong association with different Tc^r genes (p value < .0001, MANOVA), and a clear difference between MIC of *tet*(B) and *tet*(H) was found (p value < .006, t test) (Table S11).

3.7 | Comparison of tetracycline-resistant genes in *G. apicola* and *S. alvi*

As Tet B was identified within both *G. apicola* and *S. alvi*, we wanted to investigate whether there could be possible HGT of Tet B between the two species. We compared the transposon-associated *tet*(B) and the related *tetR*(B) genes of *S. alvi* to the same two genes in *G. apicola* and found them to be identical. Therefore, we included the interspaced region of Tet B in the two bacterial species and identified the same mutational variant in *S. alvi* as in *G. apicola* (strain A7) (Fig. S10). Additionally, the whole Tn10 sequence was identical to the one identified in *G. apicola* A7.

We also compared the *tet*(H) of *G. apicola* and *S. alvi*, which showed 100% identity towards different reference genes (Table S3). An alignment of the aa sequence of these two genes from *G. apicola* and *S. alvi* showed that *tet*(H) in *S. alvi* was 3 aa longer and that they differed by 6 aa at the terminal end (Fig. S11). Despite this difference, *tet*(H) in *S. alvi*, was as similar to *tet*(H) in *G. apicola* as to the best blast hit (NCBI, nr database), which was 1,918 bp (out of 2,013 bp; whole *tet*(H)) with 100% identity to *Mannheimia haemolytica* (CP005383).

4 | DISCUSSION

4.1 | Strain selection from tetracycline treatment

We addressed whether there are strain differences within the gut bacterial symbionts of honeybees at two geographical locations under very different antibiotic treatment regimes. Using both phylogenetic analysis of orthologous genes and pan-genome analysis, we found three subgroups that exist across geography within the honeybee-gut symbiont *G. apicola*; thus, bacterial strains from both Norway and Arizona were present (Figures 3a and S3). Due to the complexity of natural environments, prevalence in combination with comparative and controlled studies with model organisms have mostly been used to investigate presence of ARG and population dynamics related to HGT (Auerbach, Seyfried, & McMahon, 2007; Di Cesare et al., 2015; Hannan et al., 2010; Shoemaker, Vlamakis, Hayes, & Salyers, 2001). We used a natural, but low complex environment, to address strain selection, and we were able to extract information spanning the strain variation within *G. apicola* from both populations included in our study. Within the three predominating subgroups, the tetracycline-resistant determinant Tet B was in all subgroups. Furthermore, the strains containing ARG are intermixed with non-ARG strains. Therefore, it is unlikely that a unique clonal subgroup was selected by antibiotic usage in the USA.

4.2 | Functional differences between subgroups of *G. apicola*

The four identified subgroups of *G. apicola* displayed functional differences, with two subgroups (1.1 and 1.2) having full genetic capacity for pectin degradation (Fig. S4). Although previous work showed differences in pectin-degradation capacity by *G. apicola* strains, and 16S rRNA gene analysis and single-cell sequencing have shown genetic variations within *G. apicola*, no specific subgroups have been fully identified (Engel et al., 2012, 2014; Ludvigsen et al., 2015; Moran et al., 2012). The pectin-degrading capacity of *G. apicola* may be a health benefit for the honeybee, and this functional trait has co-evolved with *G. apicola*, as the pectin-gene phylogenies correlates with the genomic phylogeny (Fig. S5). This correlation reflects a long-term evolutionary association (Kwong & Moran, 2015). On the other hand, the endo- and exo-polygalacturonases (GH28) show a more diversified selection, as subgroup phylogenies are not identical to the overall phylogeny.

The possession of polygalacturonases has been linked to diversification of ecological strategies in fungi (Sprockett, Piontkivska, & Blackwood, 2011). The wider distribution of exo-polygalacturonase (Groups 1 and 2, Fig. S5, Table S4) compared to the other three pectin-degrading enzymes (Group 1) could indicate niche-specific evolution in the *G. apicola*. Our 16S rRNA phylogeny result included two external 16S sequences of *G. apicola* (downloaded from previous work of Engel et al., 2012) (Fig. S1). One of these strains is phenotypically capable of degrading pectin and is in Group 1, the other, not phenotypically capable of degrading pectin, is in Group 2, and this also indicates that the exo-polygalacturonase in Group 2 is probably not sufficient to fully degrade pectin. The fact that some strains have only one polygalacturonase, while some have two, has been documented in other organisms, such as fungi, in which the number of polygalacturonases each fungi possesses correlates with genomes size (Sprockett et al., 2011). This holds true for *G. apicola* as well, as there was a significant difference in genome size between Group 1 and 2, with Group 1 having

a higher average genome size, while possessing both enzymes (Figure 3b, Table 2).

We detected a large difference in per cent nucleotide identity (80%) between Group 1 and Group 2 in our data set. The selected genomes were all identified with *G. apicola*-specific primers, as described by Martinson et al. (2012). These primers target the Gamma-1 phylotype previously identified from 16S rRNA deep sequencing (Cox-Foster et al., 2007; Martinson et al., 2011), and the target was renamed *G. apicola*. From 16S rRNA gene-similarity comparisons, all genomes showed 99% or more similarity to the *G. apicola* wkB1 (CP007445) reference strain (Table 2). Bacteria strains with similarities above 97% are normally considered the same species. Despite this high 16S rRNA gene similarity, we compared the average nucleotide per cent identity of all orthologue genes (ANIb) (Table S9), and found this to be much lower than the species delineation of at least 85% (Konstantinidis & Tiedje, 2005; Richter & Rosselló-Móra, 2009). The smaller genome size, fewer CDSs and lack of capacity to degrade pectin of Group 2 and that they harbour unique genes compared to Group 1 suggest that these strains have functionally diverged.

4.3 | HGT of Tet B

As the Norwegian and Arizonan honeybees underwent very different selection pressures due to no tetracycline-vs.-tetracycline treatment, we also addressed whether there was evidence for HGT of Tc^r genes among our isolates. We used the interspace region between *tet*(B) and *tetR*(B) to address Tc^r gene spread within the community. The six interspace mutational Tet B variants are distributed between different lineages within the subgroups (Table S6), implying prior independent uptake of these variant types. Only two of these variants were identified in the strains isolated from bees collected in Norway, while all six were found in the strains isolated from bees collected in Arizona. Figure 7a shows that one Norwegian variant was similar to sequences in other environmental and gut-related bacteria, such as *Escherichia coli* and *Salmonella* sp. (Table S10). The Arizonan variants were all unique for *G. apicola*, which made it possible to use this region to predict HGT between *G. apicola* strains. Therefore, two interspace variants (numbers 1 and 5) (Figure 7b, Table S6) among the Arizonan strains were identified within three predominant subgroups and found on the same transposon (Tn10). This shows that Tet B spreads between subgroups of *G. apicola* in the honeybee-gut community through HGT. Tet B is located on the Tn10 transposon, and Whittle, Shoemaker, and Salyers (2002) and Twiss, Coros, Tavakoli, and Derbyshire (2005) both state that tetracycline exposure among other stressful environments accelerates their spread. This could explain the high prevalence of *tet*(B) in the Arizonan population (Figure 2), as previously stated by Tian et al. (2012). As our data cannot account for other environmental factors that influence HGT, another possible explanation for higher prevalence of Tc^r genes in the Arizonan population is that HGT occurs randomly in the population and that there is a selection to sustain the strains that harbour Tc^r genes, due to tetracycline exposure. We also identified a Tn10 containing Tet B in *S. alvi* with identical DNA sequence to Tet B in *G. apicola*. The fact that the Tn10 transposon has a wide host range (Chopra & Roberts, 2001; Roberts, 2005), and the fact that these two bacteria coexist in the intestine and have the necessary contact for HGT (Huddleston, 2014; Martinson et al., 2012) both support our findings. Additionally, Kwong et al. (2014) previously detected HGT of non-ARG (RHS genes) between *G. apicola* and *S. alvi* underlining that HGT between these two species occurs.

4.4 | qPCR and Tet H result comparison to previous findings

Our sampled data set represents honeybees from countries with very different apicultural antibiotic-use levels and provides a general idea about the impact of antibiotic use on commercial beekeeping in Norway vs. the USA. Despite the limited data set, our qPCR screening showed similar results regarding prevalence of Tc^r genes, as previously found on average in the USA (Tian et al., 2012). However, we identified Tet H in our *S. alvi* isolates next to core genomic genes, rather than on a composite plasmid, as Tian et al. (2012) found in honeybees from Maryland. This difference is probably due to the different populations in the two studies, although the *tet*(H) gene from the two data sets is identical. Contrary, the *tet*(H) genes identified within *G. apicola* and *S. alvi* were different, and the presence of *tet*(H) within these two bacteria genera was limited to only one subgroup/lineage each (Figures 3 and 5), indicating a more clonal spread within the community, which matches that the *tet*(H) genes are found within the genome.

The *tet*(H) gene identified within *S. alvi* using RESFINDER was identical to a reference sequence, based on pairwise comparison of the nucleotide sequence. When we aligned the aa sequence, we unexpectedly found the reference sequence to be 11 aa shorter and that they differed by 2 aa in the terminal end (Fig. S11). Therefore, the origin of this gene identified remains unknown.

4.5 | The potential for niche adaptation within *G. apicola*

Smillie et al. (2011) investigated HGT within different human body-site communities and found that HGT is more frequent within bacteria that inhabit the same niche. From our data, different subgroups could, based upon their different metabolic capacity, occupy different niches. It is likely that these niches coexist within the same space within the honeybee-gut cavity, as indicated by the detected HGT within the *G. apicola* and between *G. apicola* and *S. alvi*, which are known to form biofilms within the gut (Martinson et al., 2012). Although close proximity is more likely, we cannot determine whether the strains isolated in this study inhabit the midgut or the pyloric region, as both parts were included in our samples. Other studies using single-cell sequencing and phylogenetic analysis (Engel et al., 2014; Kwong & Moran, 2015) have not differentiated between these two gut compartments, which means they are also unable to yield additional information regarding these two possible gut niches. It would be preferable for further studies on *G. apicola* to

evaluate whether the genomic differentiation observed could be explained by gut compartment segregation.

4.6 | Host selection important for HGT

Across various environments, Tc^r genes are one of the most prevalent ARGs. Findings are consistent that higher antibiotic usage produces a higher prevalence of ARG (Bryan et al., 2004; Jiang et al., 2013; Knapp, Dolfing, Ehlert, & Graham, 2010; Seyfried, Newton, Rubert, Pedersen, & McMahon, 2010; Tian et al., 2012; Wilkerson et al., 2004). However, there are conflicting opinions as to whether or not antibiotic pressure (even below inhibitory concentrations) increases HGT, or whether only the dynamics of the population are affected (e.g., selection of antibiotic-resistant strains) (Baharoglu & Mazel, 2011; Chait, Palmer, Yelin, & Kishony, 2016; Lopatkin et al., 2016). Our results suggest that HGT within honeybee-gut populations contributes to the geographical differences detected in resistance patterns within this natural environment. In addition, they help illustrate the adaptive nature of bacteria selected by the host to survive environmental shifts, which drives the gut community to sustain the long-term co-evolution of these symbionts with their host (Kwong & Moran, 2015), despite antibiotic exposure.

5 | CONCLUSION

In conclusion, our results support a model in which evolutionarily selected bacterial subgroups adapt to antibiotic treatment through independent uptake of ARG from the environment and exchange these genes within the population by HGT, in a geographically restricted manner.

ACKNOWLEDGEMENTS

We would like to thank Nicholas A. Baker and Dr. Ying Wang, who sampled bees at Arizona State University for qPCR screening and culturing. Thanks also to Claus Kreibich at the Norwegian University of Life Sciences for helping with bee collecting. Thank to Felix Nwosu for performing part of the qPCR screenings, and Prof. Nancy Moran for providing us with positive control and technical details for the qPCR screening of Tc^r genes. Thanks to Cyril Alexander Frantzen for helping with UNIX computer program installations and Anuradha Ravi for input on figure layouts. Thanks to Nicholas A. Baker and Dalial Freitak, University of Helsinki, for proofreading the scientific content of the manuscript.

DATA ACCESSIBILITY

All genomes are submitted to NCBI with Accession nos.: *Gilliamella* (48). *Gilliamella apicola*: Accession nos. NARN01-NARZ01, NAHR01-NAHX01, MZNE01-MZNI01. *Gilliamella* spp.: Accession no. NASA01-NASW01. This whole-genome shotgun bioproject was deposited at GenBank under the Accession no. PRJNA339422. *Snodgrassella* (22)—Accession nos. MVDP01, NAGW01-NAGZ01, NAHA01-NAHQ01. This whole-genome shotgun bioproject was deposited at GenBank under the Accession no. PRJNA339423. Spades annotated .gbk files available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.qj925.

CONFLICT OF INTERESTS

All authors confirm that there are no conflict of interests to declare.

AUTHORS CONTRIBUTIONS

J.L. performed laboratory work, selected data set, analysed data, made figures and wrote the manuscript. D.P. did the bioinformatics analysis on the whole-genome shotgun sequence data, performed the GWAS analysis and wrote the M&M bioinformatics part. T.M.L.L. gave input on data analysis, and read and commented on the manuscript. G.A. provided bee collection, and read and commented on the manuscript. K.R. helped with study design, ran ML analysis, and read and commented on the manuscript. All authors read and approved the manuscript.

ORCID

Jane Ludvigsen http://orcid.org/0000-0003-4916-4738

REFERENCES

Akhtar, M., Hirt, H., & Zurek, L. (2009). Horizontal transfer of the tetracycline resistance gene tetM mediated by pCF10 among *Enterococcus faecalis* in the house fly (*Musca domestica* L.) alimentary canal. *Microbial Ecology*, *58*, 509–518. https://doi.org/10.1007/s00248-009-9533-9

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics*, *12*, 402. https://doi.org/10.1186/1471-2164-12-402

Appelbaum, P. C. (1992). Antimicrobial resistance in *Streptococcus pneumoniae*: An overview. *Clinical Infectious Diseases*, *15*, 77–83. https://doi.org/10.1093/clinids/15.1.77

Auerbach, E. A., Seyfried, E. E., & McMahon, K. D. (2007). Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Research*, *41*, 1143–1151. https://doi.org/10.1016/j.watres.2006.11.045

Baharoglu, Z., & Mazel, D. (2011). Vibrio cholerae triggers SOS and mutagenesis in response to a wide range of antibiotics: A route towards multiresistance. *Antimicrobial Agents and Chemotherapy*, *55*, 2438–2441. https://doi.org/10.1128/AAC.01549-10

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*, 455–477. https://doi.org/10.1089/cmb.2012.0021

Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N. C., & Hugenholtz, P. (2007). CRISPR recognition tool (CRT): A tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics*, *8*, 209. https://doi.org/10.1186/1471-2105-8-209

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bryan, A., Shapir, N., & Sadowsky, M. J. (2004). Frequency and distribution of tetracycline resistance genes in genetically diverse, nonselected, and nonclinical *Escherichia coli* strains isolated from diverse human and animal sources. *Applied and Environment Microbiology*, *70*, 2503–2507. https://doi.org/10.1128/AEM.70.4.2503-2507.2004

Brynildsrud, O., Bohlin, J., Scheffer, L., & Eldholm, V. (2016). Rapid scoring of genes in microbial pan-genome-wide association studies with SCOARY. *Genome Biology*, *17*, 238. https://doi.org/10.1186/s13059-016-1108-8

Burall, L. S., Grim, C. J., Mammel, M. K., & Datta, A. R. (2016). Whole genome sequence analysis using JSpecies tool establishes clonal relationships between *Listeria monocytogenes* strains from epidemiologically unrelated Listeriosis outbreaks. *PLoS ONE*, *11*, e0150797. https://doi.org/10.1371/journal.pone.0150797

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST plus: Architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

Chait, R., Palmer, A. C., Yelin, I., & Kishony, R. (2016). Pervasive selection for and against antibiotic resistance in inhomogeneous multistress environments. *Nature Communications*, *7*, 10333. https://doi.org/10.1038/ncomms10333

Chalmers, R., Sewitz, S., Lipkow, K., & Crellin, P. (2000). Complete nucleotide sequence of Tn10. *Journal of Bacteriology*, *182*, 2970–2972. https://doi.org/10.1128/JB.182.10.2970-2972.2000

Chopra, I., & Roberts, M. (2001). Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and Molecular Biology Reviews*, *65*, 232–260. https://doi.org/10.1128/MMBR.65.2.232-260.2001

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., … Lipkin, W. I. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, *318*, 283–287. https://doi.org/10.1126/science.1146498

Darling, A. E., Mau, B., & Perna, N. T. (2010). ProgressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*, *5*, e11147. https://doi.org/10.1371/journal.pone.0011147

Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, *74*, 417–433. https://doi.org/10.1128/MMBR.00016-10

Di Cesare, A., Eckert, E. M., Teruggi, A., Fontaneto, D., Bertoni, R., Callieri, C., & Corno, G. (2015). Constitutive presence of antibiotic resistance genes within the bacterial community of a large subalpine lake. *Molecular Ecology*, *24*, 3888–3900. https://doi.org/10.1111/mec.2015.24.issue-15

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Engel, P., Martinson, V. G., & Moran, N. A. (2012). Functional diversity within the simple gut microbiota of the honey bee. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11002–11007. https://doi.org/10.1073/pnas.1202970109

Engel, P., Stepanauskas, R., & Moran, N. A. (2014). Hidden diversity in honey bee gut symbionts detected by single-cell genomics. *PLoS Genetics*, *10*, e1004596. https://doi.org/10.1371/journal.pgen.1004596

Genersch, E. (2010). American Foulbrood in honeybees and its causative agent, *Paenibacillus larvae*. *Journal of Invertebrate Pathology*, *103*(Suppl 1), S10–S19. https://doi.org/10.1016/j.jip.2009.06.015

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, *35*, W52–W57. https://doi.org/10.1093/nar/gkm360

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, *59*, 307–321. https://doi.org/10.1093/sysbio/syq010

Hannan, S., Ready, D., Jasni, A. S., Rogers, M., Pratten, J., & Roberts, A. P. (2010). Transfer of antibiotic resistance by transformation with eDNA within oral biofilms. *FEMS Immunology and Medical Microbiology*, *59*, 345–349. https://doi.org/10.1111/j.1574-695X.2010.00661.x

Hillen, W., & Berens, C. (1994). Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annual Review of Microbiology*, *48*, 345–369. https://doi.org/10.1146/annurev.mi.48.100194.002021

Huddleston, J. R. (2014). Horizontal gene transfer in the human gastrointestinal tract: Potential spread of antibiotic resistance genes. *Infection and Drug Resistance*, *7*, 167–176. https://doi.org/10.2147/IDR

Jayani, R. S., Saxena, S., & Gupta, R. (2005). Microbial pectinolytic enzymes: A review. *Process Biochemistry*, *40*, 2931–2944. https://doi.org/10.1016/j.procbio.2005.03.026

Jiang, L., Hu, X., Xu, T., Zhang, H., Sheng, D., & Yin, D. (2013). Prevalence of antibiotic resistance genes and their relationship with antibiotics in the Huangpu River and the drinking water sources, Shanghai, China. *Science of the Total Environment*, *458–460*, 267–272. https://doi.org/10.1016/j.scitotenv.2013.04.038

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). GENEIOUS Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Knapp, C. W., Dolfing, J., Ehlert, P. A., & Graham, D. W. (2010). Evidence of increasing antibiotic resistance gene abundances in archived soils since 1940. *Environmental Science and Technology*, *44*, 580–587. https://doi.org/10.1021/es901221x

Koch, H., & Schmid-Hempel, P. (2011a). Bacterial communities in central European bumblebees: Low diversity and high specificity. *Microbial Ecology*, *62*, 121–133. https://doi.org/10.1007/s00248-011-9854-3

Koch, H., & Schmid-Hempel, P. (2011b). Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19288–19292. https://doi.org/10.1073/pnas.1110474108

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 2567–2572. https://doi.org/10.1073/pnas.0409727102

Kumar, S. G., Adithan, C., Harish, B. N., Sujatha, S., Roy, G., & Malini, A. (2013). Antimicrobial resistance in India: A review. *Journal of Natural Science, Biology, and Medicine*, *4*, 286–291. https://doi.org/10.4103/0976-9668.116970

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

Kwong, W. K., & Moran, N. A. (2015). Evolution of host specialization in gut microbes: The bee gut as a model. *Gut Microbes*, *6*, 214–220. https://doi.org/10.1080/19490976.2015.1047129

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

Levy, S. B., & Marshall, B. M. (2013). Honeybees and tetracycline resistance. *MBio*, *4*, e00045-13.

Lopatkin, A. J., Huang, S., Smith, R. P., Srimani, J. K., Sysoeva, T. A., Bewick, S., … You, L. (2016). Antibiotics as a selective driver for conjugation dynamics. *Nature Microbiology*, *1*, 16044. https://doi.org/10.1038/nmicrobiol.2016.44

Ludvigsen, J., Rangberg, A., Avershina, E., Sekelja, M., Kreibich, C., Amdam, G., & Rudi, K. (2015). Shifts in the midgut/pyloric microbiota composition

within a honey bee apiary throughout a season. *Microbes and Environments*, *30*, 235–244. https://doi.org/10.1264/jsme2.ME15019

Martinez, J. L., Sanchez, M. B., Martinez-Solano, L., Hernandez, A., Garmendia, L., Fajardo, A., & Alvarez-Ortega, C. (2009). Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiology Reviews*, *33*, 430–449. https://doi.org/10.1111/j.1574-6976.2008.00157.x

Martinson, V. G., Danforth, B. N., Minckley, R. L., Rueppell, O., Tingek, S., & Moran, N. A. (2011). A simple and distinctive microbiota associated with honey bees and bumble bees. *Molecular Ecology*, *20*, 619–628. https://doi.org/10.1111/mec.2011.20.issue-3

Martinson, V. G., Moy, J., & Moran, N. A. (2012). Establishment of characteristic gut bacteria during development of the honeybee worker. *Applied and Environment Microbiology*, *78*, 2830–2840. https://doi.org/10.1128/AEM.07810-11

Moran, N. A., Hansen, A. K., Powell, J. E., & Sabree, Z. L. (2012). Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. *PLoS ONE*, *7*, e36393. https://doi.org/10.1371/journal.pone.0036393

Nesme, J., Cecillon, S., Delmont, T. O., Monier, J. M., Vogel, T. M., & Simonet, P. (2014). Large-scale metagenomic-based study of antibiotic resistance in the environment. *Current Biology*, *24*, 1096–1100. https://doi.org/10.1016/j.cub.2014.03.036

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., … Parkhill, J. (2015). ROARY: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, *31*, 3691–3693. https://doi.org/10.1093/bioinformatics/btv421

Reybroeck, W., Daeseleire, E., De Brabander, H. F., & Herman, L. (2012). Antimicrobials in beekeeping. *Veterinary Microbiology*, *158*, 1–11. https://doi.org/10.1016/j.vetmic.2012.01.012

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 19126–19131. https://doi.org/10.1073/pnas.0906412106

Richter, M., Rossello-Mora, R., Oliver Glockner, F., & Peplies, J. (2016). JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*, *32*, 929–931. https://doi.org/10.1093/bioinformatics/btv681

Rissman, A. I., Mau, B., Biehl, B. S., Darling, A. E., Glasner, J. D., & Perna, N. T. (2009). Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics*, *25*, 2071–2073. https://doi.org/10.1093/bioinformatics/btp356

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*, e47. https://doi.org/10.1093/nar/gkv007

Roberts, M. C. (2005). Update on acquired tetracycline resistance genes. *FEMS Microbiology Letters*, *245*, 195–203. https://doi.org/10.1016/j.femsle.2005.02.034

Ruijter, J. M., Ramakers, C., Hoogaars, W. M., Karlen, Y., Bakker, O., van den Hoff, M. J., & Moorman, A. F. (2009). Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research*, *37*, e45. https://doi.org/10.1093/nar/gkp045

Sankar, S. A., Lagier, J. C., Pontarotti, P., Raoult, D., & Fournier, P. E. (2015). The human gut microbiome, a taxonomic conundrum. *Systematic and Applied Microbiology*, *38*, 276–286. https://doi.org/10.1016/j.syapm.2015.03.004

Seemann, T. (2014). PROKKA: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Seyfried, E. E., Newton, R. J., Rubert, K. F., Pedersen, J. A., & McMahon, K. D. (2010). Occurrence of tetracycline resistance genes in aquaculture facilities with varying use of oxytetracycline. *Microbial Ecology*, *59*, 799–807. https://doi.org/10.1007/s00248-009-9624-7

Shoemaker, N. B., Vlamakis, H., Hayes, K., & Salyers, A. A. (2001). Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Applied and Environment Microbiology*, *67*, 561–568. https://doi.org/10.1128/AEM.67.2.561-568.2001

Smillie, C. S., Smith, M. B., Friedman, J., Cordero, O. X., David, L. A., & Alm, E. J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature*, *480*, 241–244. https://doi.org/10.1038/nature10571

Snipen, L., & Liland, K. H. (2015). MICROPAN: An R-package for microbial pan-genomics. *BMC Bioinformatics*, *16*, 1–8.

Sommer, M. O., Dantas, G., & Church, G. M. (2009). Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science*, *325*, 1128–1131. https://doi.org/10.1126/science.1176950

Spivak, M. (2000). Preventative antibiotic treatments for honey bee colonies. *American Bee Journal*, *140*, 867–868.

Sprockett, D. D., Piontkivska, H., & Blackwood, C. B. (2011). Evolutionary analysis of glycosyl hydrolase family 28 (GH28) suggests lineage-specific expansions in necrotrophic fungal pathogens. *Gene*, *479*, 29–36. https://doi.org/10.1016/j.gene.2011.02.009

Thompson, H. M., Waite, R. J., Wilkins, S., Brown, M. A., Bigwood, T., Shaw, M., … Sharman, M. (2005). Effects of European foulbrood treatment regime on oxytetracycline levels in honey extracted from treated honeybee (*Apis mellifera*) colonies and toxicity to brood. *Food Additives & Contaminants*, *22*, 573–578. https://doi.org/10.1080/02652030500089986

Tian, B., Fadhil, N. H., Powell, J. E., Kwong, W. K., & Moran, N. A. (2012). Long-term exposure to antibiotics has caused accumulation of resistance determinants in the gut microbiota of honeybees. *MBio*, *3*, pii: e00377-12.

Twiss, E., Coros, A. M., Tavakoli, N. P., & Derbyshire, K. M. (2005). Transposition is modulated by a diverse set of host factors in *Escherichia coli* and is stimulated by nutritional stress. *Molecular Microbiology*, *57*, 1593–1607. https://doi.org/10.1111/j.1365-2958.2005.04794.x

Whittle, G., Shoemaker, N. B., & Salyers, A. A. (2002). The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cellular and Molecular Life Sciences*, *59*, 2044–2054. https://doi.org/10.1007/s000180200004

Wilkerson, C., Samadpour, M., van Kirk, N., & Roberts, M. C. (2004). Antibiotic resistance and distribution of tetracycline resistance genes in *Escherichia coli* O157:H7 isolates from humans and bovines. *Antimicrobial Agents and Chemotherapy*, *48*, 1066–1067. https://doi.org/10.1128/AAC.48.3.1066-1067.2004

Wright, E. S. (2015). DECIPHER: Harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, *16*, 322. https://doi.org/10.1186/s12859-015-0749-z

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., … Larsen, M. V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, *67*, 2640–2644. https://doi.org/10.1093/jac/dks261

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ludvigsen J, Porcellato D, L'Abée-Lund TM, Amdam GV, Rudi K. Geographically widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant patterns. *Mol Ecol*. 2017;00:1–18. https://doi.org/10.1111/mec.14392

Title

Resolving the diversity of the honeybee gut symbiont *Gilliamella*: description of *Gilliamella apis* sp. nov[1]., isolated from the gut of honeybees (*Apis mellifera*).

Authors

Jane Ludvigsen[1], Davide Porcellato[1], Gro V. Amdam[2,3], Knut Rudi[1].

[1]Faculty of Chemistry, Biotechnology, and Food Science, Norwegian University of Life Sciences (NMBU), Chr. M. Fahlsensvei 1, 1433 Ås, Norway. [2]Norwegian University of Life Sciences, Faculty of Ecology and Natural Resource Management, Høgskoleveien 12, 1430 Ås, Norway. [3]School of Life Sciences, Arizona State University, P.O. Box 874501, 427 East Tyler Mall, Tempe, AZ 85287, USA

[1] The GeneBank accession number for the draft genome assembly of *G. apis* NO3^T is NASD00000000

Abstract

Guts of honey bees (*Apis*) and bumble bees (*Bombus*) are colonized with the symbiotic bacterial genus *Gilliamella*. This genus shows a high degree of diversity in functional capacities and in genome similarity and separates into distinct lineages. *Gilliamella apicola* wkB1^T, isolated from *Apis*, was the first species to be described, and recently four new species, isolated from *Bombus*, have been identified. In this paper, we compare several genomes from previous studies spanning this diversity in an attempt to unravel the phylogenetic relationship among the different *Gilliamella* species. We show that one lineage, isolated only from *Apis*, is different from other *Gilliamellas* described previously, based on average nucleotide identity (ANI) and genome-to-genome distance (GGD) calculations. Here, we propose the new species name for this lineage: *Gilliamella apis* sp. nov. We present the characterization of type strain NO3^T (=DSM 105629, =LMG 30293), a strain isolated from the Western honey bee *Apis mellifera*, which cluster within this lineage. Cells of NO3^T grow best in microaerophilic atmosphere with enhanced CO_2 levels at 36 °C in pH 7.0-7.5. Cells also grow well in anaerobic conditions, but not in aerobic conditions. The cells are approximately 1µm in length and rod-shaped and the genome G+C is 34.7 mol%. Differential characteristics between NO3^T and the different type strains of *Gilliamella* were revealed from API kit tests and genomic content comparisons. Main respiratory quinone for NO3^T is ubiquinone-8, and the predominant fatty acids are $C_{18:1}$ w7c/ $C_{18:1\ w6c}$, $C_{16:0,}$ consistent with the *Gilliamella* genus.

Main text

The genus *Gillamella* is considered a symbiont in both the honey bee (*Apis*) and bumble bee (*Bombus*) guts and has an impact on the health of both species [1-3]. The genus *Gilliamella*, which was first described by Kwong and Moran [4], has the following taxonomic classification: *Proteobacteria* (Phylum), *Gammaproteobacteria* (Class), *Orbales* (Order), and *Orbaceae* (Familiy), *Gilliamella* (genus). Kwong and Moran [4] established the type strain *Gilliamella apicola* wkB1^{T} (NZ_CP007445), which is isolated from the western honey bee (*Apis mellifera*). The classification of this type strain was based on the former identification of the Gamma-1 phylotype and primers designed by Jeyaprakash, Hoy [5] and Martinson, Danforth [6] based on 16S rRNA gene sequences. Taxonomies of new isolates have been based on the similarity towards the type strain 16S rRNA sequence and a diverse range have been recognized within this genus that do not fully correspond to phylogeny based on multiple genes or pan-genome analysis [7, 8]. Since June 2017, two new studies have each genome-sequenced more than 40 strains within the *Gilliamella* genus [8, 9], and over 100 strains of this beneficial symbiont, originating from different honeybees and bumble bee species, with diverse geographic origin, are currently available from NCBI databases. This large number of genomes can help elucidate the recognized diversity spanning the *Gilliamella* genus and the delineation of new species. New bioinformatic methods that use a large part of the genome sequence information to separate strains, such as the *in-silico* DNA-DNA hybridization method (Genome-to-genome-distance-calculator, GGDC) [10, 11] and average nucleotide identity (ANI) calculation using orthologue genes [12], have more or less taken over from the more traditional DDH wet lab method for deciding new species. Both methods corresponds well to DDH methodology, and proposed species delimitations are for GGD DDH >70 % genome similarity and >95% for ANI. In addition, a G+C difference of >1 has been proposed as a criterion for distinguishing species [13]. Phylogeny based purely

on 16S rRNA gene sequencing information has proven difficult since this genus show more than 99% similarity among different species [8, 14]. ANI calculations has already been used to deepen our understanding of the diversity within this genus [8, 14], and the genus *Gilliamella* currently comprises five species, as Praet, Cnockaert [14] recently determined four new species, isolated from bumblebees,: *G. bombi* (NZ_FMWS01000014), *G. bombicola* (FMAQ01000001), *G. intestine* (FMBA01000001), and *G. mensalis* (NZ_FMWR01000014).

In the present study we have implemented a maximum-likelihood (ML) phylogenetic comparison of whole genomes of 52 *Gilliamella* strains, isolated and genome-sequenced in previous studies (***Table 1***), which cover the above-mentioned diversity. This clearly shows each closest neighboring *Gilliamella* species and, in combination with comparisons of ANI and GGD calculations for all these strains against *G.apicola* wkB1^T, we unravel any taxonomic uncertainty from strains previously identified as *G. apicola* by 16S rRNA gene similarity. In a former study of 48 *Gilliamella* strains, it was suggested that one subgroups/lineage, which showed ANI similarities ~ 80% towards wkB1^T, could be a new *Gilliamella* species [8]. Through extensive comparisons of different *Gilliamella* strains we now show that this specific lineage (20 strains), isolated at present date solely from *Apis mellifera*, is genotypically and phenotypically different from other *Gilliamella* species, and therefore should be described as a new species for which we propose the name *Gilliamella apis* sp. nov. Here we report the phenotypic and genotypic characterizations of the type strain NO3^T in addition to three other strains from the same lineage.

Strain NO3^T was isolated from homogenized honeybee guts living in Ås, Norway. As part of a study to isolate tetracycline resistant bacteria from the midgut/pyloric region [8], the strains

NO3^T and A7 were isolated after 48h incubation at 37°C, in an enhanced CO_2 atmosphere (BD GasPak™ EZ CO_2 Container system, Becton Dickinson, USA [$CO_2 \geq 2.5\%$]), on tryptic soy agar (TSA) supplemented with 5% horse blood added 12 µg/ml tetracycline. From the same study, strains A-TSA1 and N-G2 were isolated on TSA + 5% horse blood without tetracycline under the above-described culturing conditions. Sub-cultivation was performed on TSA + 5% horse blood with the same culturing conditions to ensure purity before the strains were genome-sequenced using the Nextera XT DNA library preparation from Illumina® on the MiSeq system as described in [8]. The strains were mixed with heart infusion broth (HIB) with 11% glycerol and stored at -80°C until further characterization was performed.

All *Gilliamella* genomes were downloaded from NCBI, and single-copy core genes from selected *Gilliamella* strains (***Table 1***) were extracted from the pan-genome table created with the R package "micropan" [15]. 189 single-copy core gene clusters were translated to amino acid sequences, aligned, and back-translated to nucleotide sequences using the R package "Decipher" [16]. All of the back-translated alignments were combined together. An ML analysis was conducted on the combined alignment with the R package "phanghorn" [17], using a GTR model and optimized for gamma rate, variable size, and topology with the rearrangement parameter set for stochastic and a 100× bootstrap analysis. The ML tree was rooted on the *Bombus* branch since previous studies have shown *Gilliamella* strains isolated from *Bombus* to be distinct from *Gilliamella* strains isolated from *Apis* [9]. The ML analysis clearly separates the NO3^T from the other *Gilliamella* species (***Figure 1***). GGD and ANI calculations between NO3^T and wkB1^T, *G. intestini* LMG 28358^T, *G. bombicola* LMG 28359^T, *G. bombi* LMG 29879^T, *G. mensalis* LMG 29880^T support the claim that NO3^T is a new species (Table 2). Strains: SAG P17, M1-2G and P62G, described by Engel, Stepanauskas [7] and Zheng, Nishida [9], cluster together with NO3^T (***Figure 1***; red

highlighted), and both ANI (>97%) and GGD (>70%) support the argument that these strains are the same species (Table 2; dark grey highlighted). The same ANI and GGD values are apparent for strain wkB30 (**Table 3**; dark grey highlighted), which clusters with *G. bombi* (***Figure 1***; red highlighted). While it remains to be determined whether strain A7, N-G2, NW-3 and SAG I20 should be separate from *G. apis* NO3^T and whether the cluster containing wkB7 and SAG B02 should be separate from *G. apicola* wkB1^T, both ANI and GGD values are below the species delimitation (Table 2; light gray highlighted). The same observation was done for Imp1-6 and Choc5-1, which cluster with *G. bombi* and *G. intestine*, respectively (**Table 3**; light grey highlighted). All strains in question are highlighted in blue-colored boxes in ***Figure 1***. For consideration, if the ANI and GGD values between *Gilliamella* type strains are takes as a reference [14], the similarity of the remaining strains toward their closest type strain is as low (**Table 2** & **Table 3**), and there is reason for further characterizations and possible collaborate decisions on which additional new species should be proposed.

For additional comparison with previous literature, we included 16S rRNA gene BLAST searches towards all five type strains (Local BLAST in Geneious v8). NO3^T was found to be closest related with *G. apicola* wkB1^T with 98.9% similarity. Table S1 summarizes 16S rRNA gene identities towards: wkB1^T, *G. intestini* LMG 28358^T, *G. bombicola* LMG 28359^T, *G. bombi* LMG 29879^T, *G. mensalis* LMG 29880^T and NO3^T, for selected genomes from ***Table 1*** and four closely related genera (*Orbus hercynius* CN3^T (FJ612598), *Orbus sasakiae* C7^T (JN561614), *Frischella perrara* PEB0191^T (JX878306), *Schmidhempelia bombi* Bimp (AWGA01000037). It is known from previous literature that this cluster shows a maximum of only 1.1% difference to wkB1^T at the 16S rRNA gene level and that some strains are 99.3% similar, indicating the species delimitation for *G. apis* to be at 0.7% dissimilarity towards wkB1^T [8]. If we define the *Gilliamella apis* cluster (***Figure 1***: dark

yellow box) based on GGD and ANI species cut-off values (**Table 2**), the within species divergence is 0.6% based on blast hit of > 1400bp of the 16S rRNA gene towards the type strain NO3^T (Table S1). We made a 16S rRNA gene phylogeny from a Clustal W alignment of 1195bp with bootstrap100× (Geneious v8), of the strains in Table S1, which places NO3^T in the *Orbaceae*/ *Gilliamella* lineage, but in a separate cluster from any of the other four type strains (Figure S1).

The cultivation characteristics of NO3^T, A-TSA1, A7 and N-G2 are described in table S2. We tested growth on TSA, HIA, and Müller Hinton agar (MH), both with and without 5% horse blood added, as well as growth on these agars at three different temperatures: 30°C, 36°C, and 40°C. We streaked one 48h-old colony on new agar plates and incubated in microaerophilic atmosphere. All strains grew after 24h incubation with pinpoint large, white/grey colonies on TSA + 5% horse blood. Strain A7 and N-G2 showed slightly α-hemolytic colonies, which was visualized by a green/brown zone underneath the colony lawn. After 48h, NO3^T and A-TSA1 colonies were <1mm large, while A7 and N-G2 colonies were 1mm, all white/grey/opaque, slightly α-hemolytic and smooth. After 72h, NO3^T and A-TSA1 colonies were approximately 1.5–2 mm, while A7 and N-G2 colonies were 2.5 mm. On HIA, colonies were non-hemolytic, translucent, and colorless with the same size as detected on TSA + 5% horse blood.

All strains grew at 30°C and 36°C on TSA and HIA with and without 5% horse blood and no difference in colony size could be detected when horse blood was added. The strains grew better on MH agar + 5% horse blood than on MH agar without blood at both temperatures, and strain N-G2 grew on MH without blood at 36°C. At 42°C growth was weak, on TSA and HIA with and without horse blood, and was only visible where bacteria density was high. No growth was detected on MH agar or MH agar with 5 % horse blood at 42°C.

We also tested aerobic and anaerobic growth (AnaeroGen™ pouch, Thermo scientific, USA [O_2 < 0.1% and CO_2 7-15%]) at 36°C. None of the strains grew in aerobic conditions, but all strains grew in anaerobic conditions, indeed as well or better (on MH) as in microaerophilic conditions (Table S2).

For strain NO3^T Growth at pH 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0 was tested by incubating bacteria in trypic soy broth (TSB) at 36°C in enhanced CO_2 atmosphere for 48h. TSB (Sigma Aldrich, Germany) was buffered to the above-mentioned concentrations with 1.5M NaCl and 1.0M NaOH and autoclaved for 15min at 121°C. The pH was measured after autoclaving in control tubes, which showed a decreased in pH of 0.1 unit. We added 50µl of an overnight culture of NO3^T to 5ml TSB and measured both pH and absorbance before and after 48h incubation. After incubation in enhanced CO_2 atmosphere, pH was lowered 0.5 units in tubes without bacteria and the pH was 1.0 unit lower in tubes due to bacterial growth. Approximate cell density was measured using McFarland turbidity measurement and found to be highest at pH 6.5-7.5, with maximum cell density of 1.06×10^9 CFU/ml at pH 7.0, with no growth at pH 5.5. Cells grew evenly within the broth and did not clump together as described for wkB1^T [4].

Investigation of phenotypic fermentation capacities and enzyme production was done using the API kits (bioMérieux, France) on 24h-old colonies following the manufacturer's recommendations. Previous studies using the API 20NE kit have shown lack of growth in the supplied medium [4, 14], so we used API NH, which only needs 2h for detection of fermentation reactions, and we also tested API 20E. We included *G. apicola* wkB1^T (= DSMZ 104097) as a control, to enable comparison. Catalase and oxidase activity was tested using ID color Catalase (ID-ASE) (bioMérieux® sa, France) and Bactident® Oxidase reagent (MERC, Germany), respectively following manufacturers' recommendations. All five strains (NO3, A-TSA1, A7, N-G2 and wkB1^T) were negative for both catalase and

196 oxidase activity and showed similar sugar fermentation profiles: D-glucose, D-Fructose and
197 D-Sucrose positive and D-maltose negative. Using the API NH, we achieved a positive result
198 on glucose fermentation for wkB1^T, which is in congruence with [18], but in contrast to
199 Kwong and Moran [4] and Praet, Cnockaert [14], who characterized wkb1^T as glucose
200 negative. Strain A7 and wkB1^T differed from the others in being Gamma-glutamyl
201 transferase-positive. In addition, wkB1^T was positive for β-Galactosidase and γ-glutamyl
202 transferase, and A7 was lacking amygdalin fermentation, while NO3^T, A-TSA1, and N-G2
203 were all positive. WkB1^T did not yield any positive reactions using the API 20E kit, while the
204 other strains did (Table S3). Additionally, NO3^T, A7, A-TSA1 and N-G2 were positive for
205 the Vouges-Proskauer test (acetoin production), which indicate that they convert glucose by
206 the butanediol fermentation pathway [19]. Zheng, Nishida [9] showed that there is a high
207 correlation between the genotype and the phenotypic sugar fermentation capacities of
208 *Gilliamella* strains. Therefore, using the genome sequenced data we also determined the
209 fermentation capacities of xylose, arabinose, mannose, and rhamnose, from the presence of
210 nine sugar fermenting genes as previously described by Zheng, Nishida [9]. Strain NO3^T
211 possesses the genes xylA and xylB and rhaA, rhaB, and rhaD for xylene and rhamnone
212 catabolism, respectively (Table S4). Despite the presence of the rhamnose utilizing genes, the
213 API 20E test did not yield a positive result for L-Rhamnose. This discrepancy might indicate
214 that API 20E is not a suitable test for enzyme production in *Gilliamella,* since positive tests
215 other than D-glucose were only truly positive after 48h incubation. Differential phenotypic
216 and genotypic characteristics of *Gilliamella* type strains are summarized in ***Table 4***.
217
218 We morphologically characterized strain NO3^T by scanning-electron microscopy (SEM) and
219 light microscopy (phase contrast and Gram staining). For SEM imaging, bacterial cells from
220 an overnight culture (TSB in enhanced CO_2) were spun down and fixated in 0.1M PIPES-

buffer with 1.25% glutaraldehyde and 2% paraformaldehyde for 2h at room temperature. The cells were washed three times in 0.1M PIPES buffer and left overnight in 0.1M. Cells were left to attach to a glass slide coated with poly-l-lysine for one hour and then a dehydration protocol (50% → 100% ethanol) was used before the cells were dried in an automated critical point dryer. The cells were coated with 80%/20% gold and palladium particles before visualized with Zeiss EVO 50. Cell were rod-shaped and approximately 1.0× 0.25 μm in size (***Figure 2***a). Single cells, cells in pairs and in star-like clusters as well as long filamentous shapes were detected. Gram staining and phase contrast microscopy were performed on the same overnight culture but on flame fixed cells and live bacteria, respectively (***Figure 2***b and ***Figure 2***c). Cells were found to be Gram-negative, rod-shaped, single and in pairs, and in both samples we detected filamentous forms and also here star shapes aggregates are suspected. Phase contrast microscopy of live cells showed that they were vigorously moving around, which is consistent with cells growing evenly in broth. From the SEM picture, there seems to be some kind of filaments or attachment point on the short end of each rod, which could be flagella for movement, but exactly what this is needs to be confirmed.

We tested the susceptibility of NO3^T, A-TSA1, A7, G2 and wkB1^T to three different antibiotics – tetracycline, ampicillin, and streptomycin – using the MIC-strip (Montebello, Norway) and following EUCAST guidelines. Bacteria were dissolved in 0.9% NaCl to 1.0 McFarland concentration and streaked in three directions on a MH plate with 5% horse blood, and incubated for 24h in enhanced CO_2 atmosphere before MIC was read. All five strains had similar MIC range for ampicillin: 0.19–0.38 mg/l, but varied for the other two antibiotics (Table S5). MIC was lower in N-G2 (0.19 mg/l), which do not have resistance genes towards tetracycline (Ludvigsen) and wkB1^T had the highest MIC (12.0 mg/l). NO3^T

has the following MICs: tetracycline; 3.0 mg/L, ampicillin; 0.25 mg/L, and streptomycin; 3.0 mg/L. Differential results of $NO3^T$ and $wkB1^T$ are summarized in ***Table 4***.

After all sub-cultivations were performed, purity of strain $NO3^T$ was tested by sequencing 998 bp of the 16S rRNA gene and compared to the 16S rRNA gene sequence extracted from the original genome sequenced isolate. Only one nt differed between the two sequences, which possibly originate from sequencing error. Forty-eight-hour-old bacterial cells cultured in TSB were centrifuged and washed in phosphate-buffered saline (PBS) and then centrifuged again and supernatant removed. Cells were freeze-dried and sent to DSMZ (Braunschweig, Germany) for the identification of polar lipids and respiratory quinones, while live cell culture was used for fatty acid characterization. All three analyses were carried out by the Identification Service of the DSMZ. Characterization of fatty acids in $NO3^T$ was done by fatty acids methyl ester (FAME) analysis, with the Sherlock MIS (MIDI Inc, Newark, USA) system. Similar to other *Gilliamella* type trains [4, 14], the main fatty acids for $NO3^T$ were $C_{18:1\ w7c}$ and/or $C_{18:1\ w6c}$, $C_{16:0}$, and $C_{14:0}$ 3OH and/or $C_{16:1}$ iso I.

Respiratory lipoquinones were extracted using the two-stage method described by Tindall [20], and Tindall [21]. Main respiratory quinone found was ubiquinone-8, which is consistent with *G. apicola* $wkB1^T$ and other members of Orbeaceae [4, 18], but Praet, Cnockaert [14] failed to report any quinones for the other Gilliamella type strains as did they for $wkB1^T$ when tested in the same analysis. “When polar lipids from the same biomass have previously been subjected to respiratory lipoquinone analysis the two-stage method described by Tindall [20], and Tindall [21] is used to first extract the respiratory lipoquinones followed by the polar lipids” (https://www.dsmz.de/services/services-microorganisms/identification/analysis-of-polar-lipids.html). Phosphoaminoglycolipid, diphosphatidylglycerol, phosphatidylethanolamine, and phosphatidylglycerol are the main polar lipids (Figure S2).

The data presented here show that the genus *Gilliamella* consists of several lineages, which might potentially be new species and should be further characterized and possibly renamed. Comparison of genomic content and phenotypic characteristics both support that strains $NO3^T$ and A-TSA1 clearly cluster within a separate lineage, with ANI and GGD values that support this lineage to be a new species, for which we propose the name *Gilliamella apis* sp. nov.

Description of *Gilliamella apis* sp. nov.

Gilliamella apis (a'pis. L. gen. fem. n. *apis* from a honey bee, the genus name of the honey bee *Apis mellifera*, which refers to the insect host of this species).

Cells are Gram-negative, rod-shaped, and ~1.0×0.25 μm large. Optimal growth is observed on TSA in microaerophilic atmosphere enriched with CO_2 at 36°C, after 48–72h at pH 7.0–7.5. After 48h colonies are < 1mm in diameter, round and smooth with grey/white color and slightly α-hemolytic on TSA. After 72h, colonies are < 1.5mm large. On HIA colonies are translucent, none-hemolytic but with the same size as grown on TSA. Similar growth is observed in anaerobe conditions at 36°C, but it does not grow under aerobe conditions. Strain $NO3^T$ is negative for catalase and oxidase, ferments D-glukose, D-fructose, D-saccharose (sucrose), and harbors the genes xylA, xylB, rhaA, rhaB, and rhaD for xylene and rhamnose catabolism. It produces acetoin from glucose and possesses the enzyme alkaline phosphatase. It is negative for D-mannose, D-mannitol, β-Galactosidase, γ-Glutamyl transferase, urease, and indole production. It harbors the Tet H determinant that confers resistance towards tetracycline. Polar lipids are: Phosphoaminoglycolipid, diphosphatidylglycerol, phosphatidylethanolamine, and phosphatidylglycerol, and the main fatty acids are: $C_{18:1\ w7c}$ and/or $C_{18:1\ w6c}$, $C_{16:0}$, $C_{14:0}$ 3OH and/or $C_{16:1}$ iso I, and $C_{12:0}$ (in sinking order). Genome size is 2.52 Mbp and G+C is 34.7 mol%, and the genome assembly is deposited at NCBI under

accession number NASD00000000. The type strain, NO3^T (=DSM 105629, =LMG 30293) was isolated from the gut of the honey bee *A. mellifera* from Ås, Norway.

Acknowledgement

We want to thank Philipp Engel for providing us full length 16S rRNA sequences for the SAG strains, Hilde Raanaas Kolstad and Åsmund Andersen for help with SEM microscopy preparation and visualization, and Ahmed Abdelghani for assistance with freeze-drying bacteria.

Conflict of interest and ethical statement

We the authors declare no conflict of interest and that research was done in accordance with Norwegian rules for beekeeping and sample collection.

References

1. **Kwong WK, Moran NA**. Gut microbial communities of social bees. *Nat Rev Microbiol* 2016;14(6):374-384.
2. **Maes PW, Rodrigues PA, Oliver R, Mott BM, Anderson KE**. Diet-related gut bacterial dysbiosis correlates with impaired development, increased mortality and Nosema disease in the honeybee (Apis mellifera). *Mol Ecol* 2016;25(21):5439-5450.
3. **Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA**. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc Natl Acad Sci U S A* 2017;114(18):4775-4780.
4. **Kwong WK, Moran NA**. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: Snodgrassella alvi gen. nov., sp. nov., a member of the Neisseriaceae family of the Betaproteobacteria; and Gilliamella apicola gen. nov., sp. nov., a member of Orbaceae fam. nov., Orbales ord. nov., a sister taxon to the Enterobacteriales order of the Gammaproteobacteria. *Int J Syst Evol Microbiol* doi: 10.1099/ijs.0.044875-0.
5. **Jeyaprakash A, Hoy MA, Allsopp MH**. Bacterial diversity in worker adults of Apis mellifera capensis and Apis mellifera scutellata (Insecta: Hymenoptera) assessed using 16S rRNA sequences. *J Invertebr Pathol* 2003;84(2):96-103.
6. **Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S et al.** A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol* 2011;20(3):619-628.
7. **Engel P, Stepanauskas R, Moran NA**. Hidden diversity in honey bee gut symbionts detected by single-cell genomics. *PLoS Genet* 2014;10(9):e1004596.
8. **Ludvigsen J, Porcellato D, L'Abée-Lund TM, Amdam GV, Rudi K**. Geographically Widespread Honeybee-Gut Symbiont Subgroups Show Locally Distinct Antibiotic-Resistant Patterns. *Molecular Ecology*:*in Press*.
9. **Zheng H, Nishida A, Kwong WK, Koch H, Engel P et al.** Metabolism of Toxic Sugars by Strains of the Bee Gut Symbiont Gilliamella apicola. *MBio* 2016;7(6).
10. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, journal article 2013;14(1):60.
11. **Auch AF, von Jan M, Klenk HP, Göker M**. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2.
12. **Richter M, Rosselló-Móra R**. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences* 2009;106(45):19126-19131.
13. **Meier-Kolthoff JP, Klenk H-P, Göker M**. Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *International Journal of Systematic and Evolutionary Microbiology* 2014;64(2):352-356.
14. **Praet J, Cnockaert M, Meeus I, Smagghe G, Vandamme P**. Gilliamella intestini sp. nov., Gilliamella bombicola sp. nov., Gilliamella bombi sp. nov. and Gilliamella mensalis sp. nov.: Four novel Gilliamella species isolated from the bumblebee gut. *Syst Appl Microbiol* 2017;40(4):199-204.
15. **Snipen L, Liland KH**. micropan: an R-package for microbial pan-genomics. *BMC Bioinformatics*, journal article 2015;16(1):1-8.

16. **Wright ES**. DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*, journal article 2015;16(1):322.
17. **Schliep KP**. phangorn: phylogenetic analysis in R. *Bioinformatics (Oxford, England)* 2011;27(4):592-593.
18. **Engel P, Kwong WK, Moran NA**. Frischella perrara gen. nov., sp. nov., a gammaproteobacterium isolated from the gut of the honey bee, Apis mellifera. *Int J Syst Evol Microbiol* 2013.
19. **O'Meara RAQ**. The Mechanism of the Voges-Proskauer Reaction and the Diacetyl Reaction for Proteins. *British journal of experimental pathology* 1931;12(5):346-356.
20. **Tindall BJ**. A Comparative Study of the Lipid Composition of Halobacterium saccharovorum from Various Sources. *Systematic and Applied Microbiology* 1990;13(2):128-130.
21. **Tindall BJ**. Lipid composition of Halobacterium lacusprofundi. *FEMS Microbiology Letters* 1990;66(1):199-202.
22. **Kwong WK, Engel P, Koch H, Moran NA**. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc Natl Acad Sci U S A* 2014;111(31):11509-11514.

405 # Tables

406 ***Table 1:*** Strains included in this study to produce ***Figure 1***. All genomes are downloaded from NCBI from the
407 listed accession number as well as G+C mol% information. Type strains are highlighted.

Strain	Taxonomy	Source	G+C mol%	Reference	Accession
wkB11	*G. apicola*	*Bombus*	34.4	[22]	JFON00000000
Fer2-1	*G. apicola*	*Bombus*	35.4	[9]	LZGY00000000
Nev5-1	*G. apicola*	*Bombus*	35.4	[9]	LZHO00000000
App4-10	*G. apicola*	*Bombus*	35.6	[9]	LZGS00000000
Choc5-1	*G. apicola*	*Bombus*	34.9	[9]	LZHH00000000
Imp1-6	*G. apicola*	*Bombus*	34.9	[9]	LZHL00000000
wkB30	*G. apicola*	*Bombus*	34.6	[22]	JFZX00000000
LMG 28358^T	*G. intestini*	*Bombus*	34.6	[14]	FMBA00000000
LMG 28359^T	*G. bombicola*	*Bombus*	35.9	[14]	FMAQ00000000
LMG 29879^T	*G. bombi*	*Bombus*	34.6	[14]	FMWS00000000
LMG 29880^T	*G. mensalis*	*Bombus*	35.5	[14]	FMWR00000000
wkB1^T	*G. apicola*	*Apis*	33.6	[22]	NZ_CP007445
wkB108	*G. apicola*	*Apis*	34.6	[9]	LZGM00000000
wkB112	*G. apicola*	*Apis*	34.4	[9]	LZGL00000000
wkB308	*G. apicola*	*Apis*	35.6	[9]	LZGN00000000
wkB292	*G. apicola*	*Apis*	35.1	[9]	LZGO00000000
AM4	*G. apicola*	*Apis*	33.9	[8]	NARY00000000
SAG B02	*G. apicola*	*Apis*	33.7	[7]	JAIM00000000
N2	*G. apicola*	*Apis*	34.0	[8]	NARW00000000
A-8-12	*G. apicola*	*Apis*	34.0	[8]	NARR00000000
wkB7	*G. apicola*	*Apis*	34.0	[9]	NZ_CM004509
N-G5	*G. apicola*	*Apis*	33.5	[8]	NASA00000000
P54G	*G. apicola*	*Apis*	33.8	[9]	LZGJ00000000
N-9-4	*G. apicola*	*Apis*	33.6	[8]	NAHW00000000
NO5	*G. apicola*	*Apis*	33.6	[8]	NAHV00000000
A-1-24	*G. apicola*	*Apis*	33.6	[8]	MZNE00000000
A8	*G. apicola*	*Apis*	33.5	[8]	MZNG00000000
N-22	*G. apicola*	*Apis*	33.7	[8]	NASB00000000
SAG I20	*G. apicola*	*Apis*	35.0	[7]	JAIN00000000
M1-2G	*G. apicola*	*Apis*	34.6	[9]	LZGQ00000000
SAG P17	*G. apicola*	*Apis*	34.4	[7]	JAIO00000000
P62G	*G. apicola*	*Apis*	34.7	[9]	LZGI00000000
NO3^T	*G. apis*	*Apis*	34.7	[8]	NASD00000000
N-G2	*G. apis*	*Apis*	34.4	[8]	NAST00000000
NW-3	*G. apis*	*Apis*	34.4	[8]	NASW00000000
A7	*G. apis*	*Apis*	34.2	[8]	NASN00000000
N-G4	*G. apis*	*Apis*	34.6	[8]	NASV00000000
N-G1	*G. apis*	*Apis*	34.7	[8]	NASS00000000
N-G3	*G. apis*	*Apis*	34.7	[8]	NASU00000000
AM1	*G. apis*	*Apis*	34.6	[8]	NASL00000000
A-TSA1	*G. apis*	*Apis*	34.8	[8]	NASO00000000
A-TSA2	*G. apis*	*Apis*	34.8	[8]	NASP00000000
A-TSA3	*G. apis*	*Apis*	34.8	[8]	NASQ00000000
A-TSA4	*G. apis*	*Apis*	34.8	[8]	NASR00000000
A-4-12	*G. apis*	*Apis*	34.7	[8]	NASK00000000
NO1	*G. apis*	*Apis*	34.7	[8]	NASI00000000
NO4	*G. apis*	*Apis*	34.7	[8]	NASM00000000
NO12	*G. apis*	*Apis*	34.7	[8]	NASH00000000
NO13	*G. apis*	*Apis*	34.7	[8]	NASG00000000
NO14	*G. apis*	*Apis*	34.7	[8]	NASF00000000
NO16	*G. apis*	*Apis*	34.7	[8]	NASE00000000
NO15	*G. apis*	*Apis*	34.7	[8]	NASJ00000000

408

Table 2: Similarity values in % and G+C difference between selected genomes spanning the lineages found in Figure 1 towards Apis type strains G. apicola wkB1T and G. apis NO3T. Dark grey highlighted = above species cut-off values, Light grey highlighted = below species cut-off values but above similarity values between already described type strains.

	wkB1^T					**NO3^T**				
Strains	ANIb	ANIm	TETRA	GGD	G+C difference	ANIb	ANIm	TETRA	GGD	G+C difference
NO3^T	79.81	85.57	0.9625	23.2	1.05	-	-	-	100	0
LMG 8358^T	80.11	85.68	0.9705	29.2	2.02	79.33	85.58	0.9481	29.2	0.98
LMG 8359^T	80.38	85.54	0.9648	29.0	3.24	79.17	85.88	0.9494	28.9	2.2
LMG 9879^T	79.38	84.27	0.9720	22.7	0.93	78.35	84.44	0.9531	21.9	0.11
LMG 9880^T	79.47	84.26	0.9709	22.8	1.89	78.53	84.25	0.9603	21.8	0.85
A-TSA1	79.75	85.55	0.9609	23.2	1.17	97.26	97.78	0.9987	78.5	0.12
A7	80.02	85.88	0.9623	23.7	0.54	93.04	93.84	0.9971	52.5	0.51
N-G2	79.94	85.63	0.9705	23.3	0.76	91.97	93.22	0.9966	48.4	0.29
wkB11	79.09	84.26	0.9762	22.8	0.78	78.34	84.36	0.9615	21.7	0.27
wkB30	79.84	84.40	0.9727	23.1	1.03	78.58	84.54	0.9548	22.3	0.02
wkB308	77.69	84.33	0.9398	21.6	1.97	77.64	84.70	0.9378	21.4	0.92
wkB292	78.98	84.32	0.9664	22.5	1.52	78.45	84.48	0.9679	22.0	0.48
wkB7	93.18	93.56	0.9945	51.4	0.36	79.59	85.27	0.9675	22.9	0.68
M1-2G	79.83	85.36	0.9608	23.0	1.01	97.36	97.78	0.9989	78.6	0.04
P62G	79.82	85.69	0.9604	23.4	1.07	97.13	97.74	0.9992	78.1	0.03
SAG I20	80.17	86.23	0.9516	24.8	1.41	85.69	87.85	0.9849	31.3	0.37
SAG B02	92.95	93.71	0.9902	51.8	0.07	79.94	86.11	0.9645	24.2	0.97
SAG P17	79.96	86.02	0.9608	23.8	0.8	97.55	97.83	0.9964	79.9	0.25

Table 3*:* Similarity values in % and G+C difference between selected Bombus genomes towards the Bombus type strains G. intestine LMG 28358T, G. bombicola LMG 28359T, G. bombi LMG 29879T, G. mensalis LMG 29880T. Dark grey highlighted = above species cut-off values, Light grey highlighted = below species cut-off values but above similarity values between already described type strains.

	***G. intestini* LMG 28358^T**					***G. bombicola* LMG 28359^T**				
Strains	ANIb	ANIm	TETRA	GGD	G+C	ANIb	ANIm	TETRA	GGD	G+C
LMG 28358^T	-	-	-	100	0	86.70	87.95	0.9879	34.5	1.22
LMG 28359^T	86.03	87.95	0.9879	34.5	1.22	-	-	-	100	0
LMG 29879^T	89.00	91.59	0.9845	44.8	1.09	85.85	87.83	0.9743	34.3	2.31
LMG 29880^T	85.00	87.56	0.9787	33.6	0.13	90.11	92.14	0.9840	47.4	1.35
wkB11	83.98	89.92	0.9811	39.7	1.24	82.29	87.00	0.9725	31.6	2.46
Fer2-1	83.93	86.57	0.9753	31.6	0.26	83.88	86.82	0.9702	32.2	1.48
Nev5-1	84.29	87.06	0.9741	33.4	0.24	84.43	87.70	0.9699	34.2	1.46
App4-10	85.63	87.67	0.9783	33.9	0.03	87.80	89.94	0.9836	39.7	1.25
Choc5-1	94.21	94.98	0.9894	59.6	0.7	86.30	87.87	0.9785	34.8	1.92
Imp1-6	89.10	91.65	0.9814	45.7	0.77	86.00	87.73	0.9732	34.1	1.99
wkB30	89.01	91.38	0.9825	44.2	1.0	85.55	87.73	0.9732	34.1	2.22
	***G. bombi* LMG 29879^T**					***G. mensalis* LMG 29880^T**				
Strains	ANIb	ANIm	TETRA	GGD	G+C	ANIb	ANIm	TETRA	GGD	G+C
LMG 28358^T	89.30	91.57	0.9845	44.8	1.09	85.83	87.54	0.9787	33.6	0.13
LMG 28359^T	85.71	87.79	0.9743	34.3	2.31	89.65	92.13	0.9840	47.4	1.35
LMG 29879^T	-	-	-	100	0	85.94	87.07	0.9741	30.6	0.96
LMG 29880^T	85.91	87.06	0.9741	30.6	0.96	-	-	-	100	0
wkB11	83.59	88.74	0.9869	28.8	0.15	81.73	85.58	0.9824	25.4	1.11
Fer2-1	83.55	85.59	0.9789	27.1	0.83	84.06	85.73	0.9759	27.6	0.13
Nev5-1	84.31	86.21	0.9790	28.1	0.85	84.63	86-36	0.9756	28.8	0.11
App4-10	86.18	87.45	0.9743	31.5	1.06	89.23	89.79	0.9942	37.7	0.1
Choc5-1	88.54	90.45	0.9882	37.2	0.39	85.05	86.28	0.9774	28.9	0.57
Imp1-6	95.01	95.92	0.9943	63.3	0.32	85.85	86.85	0.9740	30.5	0.64
wkB30	96.85	97.27	0.9982	72.8	0.09	86.05	86.93	0.9730	30.8	0.87

Table 4: Differential characteristics of type strains: *G. apicola* wkB1^T, *G. apis* NO3^T, *G. intestini* LMG 28358^T, *G. bombicola* LMG 28359^T, *G. bombi* LMG 29879^T, and *G. mensalis* LMG 29880^T. Some results are summarized from reference literature listed in **Table 1**.

		G. apis **NO3^T**	***G. apicola*** **wkB1^T**	***G. intestini*** **LMG 28358^T**	***G. bombicola*** **LMG 28359^T**	***G. bombi*** **LMG 29879^T**	***G. mensalis*** **LMG 29880^T**
Phenotypic							
Growth at 42°C		w[1]	-	w[1]	-	w[1]	+
Growth on MH agar		-	+	NT[2]	NT[2]	NT[2]	NT[2]
pH growth range		7.0 - 7.5	6.0 - 6.5	NT[2]	NT[2]	NT[2]	NT[2]
D-Glucose		+	+	-	-	-	-
γ-Glutamyl transferase		-	+	NT[2]	NT[2]	NT[2]	NT[2]
β-Galactosidase		-	+	+	-	-	-
Catalase		-	-	-	+	+	-
Tetracycline MIC		3.0 mg/l	12.0 mg/l	NT[2]	NT[2]	NT[2]	NT[2]
Streptomycin MIC		3.0 mg/l	12.0 mg/l	NT[2]	NT[2]	NT[2]	NT[2]
Genotypic							
Xylose	xylA	+	+	-	-	-	-
	xylB	+	+	-	-	-	-
Arabinose	araA	-	+	-	-	-	+
	araB	-	+	-	-	-	+
	araD	-	+	-	-	-	+
Mannose	manA	-	+	-	+	-	-
Rhamnose	rhaA	+	+	-	-	-	-
	rhaB	+	+	-	-	-	-
	rhaD	+	+	-	-	-	
Pectin	PL1	-	+	-	-	-	-
Tetracycline gene		Tet H	Tet B				
G+C mol%		34.7	33.5	34.6	35.9	34.6	35.5
Fatty acids		$C_{16:0}$, $C_{14:0}$ 3OH and/or $C_{16:1}$ iso I, $C_{12:0}$	$C_{16:0}$, $C_{18:1\ w7c}$ or $C_{18:1\ w6c}$	$C_{16:0}$, $C_{18:1\ w7c}$, $C_{12:0}$, $C_{18:1\ w9c}$	$C_{16:0}$, $C_{18:1\ w7c}$	$C_{16:0}$, $C_{18:1\ w7c}$	$C_{16:0}$, $C_{18:1\ w7c}$, $C_{14:0}$ 3OH and/or $C_{16:1}$ iso I
Respiratory quinone		Ubiquinone 8	Ubiquinone 8	-	-	-	-

[1] w = weak growth
[2] NT= Not tested

Figures

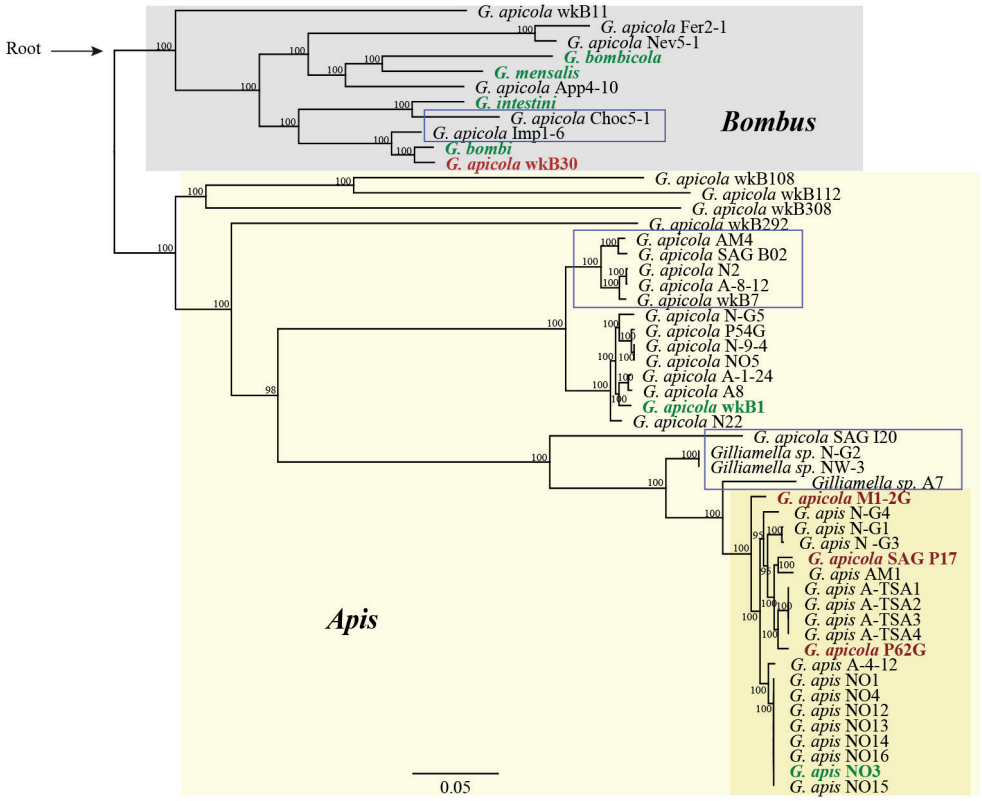

Figure 1*:* ML phylogeny of 189 single-copy core gene clusters from 52 *Gilliamella* genomes previously isolated from different *Bombus* and *Apis* species [4, 8, 9, 14, 22]. The tree is rooted by the *Bombus* lineage and names as they appear in NCBI are used and bootstrap values are shown in percent. The five type strains of *Gilliamella* are highlighted in green and clusters separately: Grey box = *Bombus* and light yellow box = *Apis*. The *G. apis* cluster is highlighted in dark yellow and strains that were previously identified as *G. apicola* but should be renamed are shown in red. Strains that fall below species delimitation cut-off of ANI and GGD (Table 2 and 3) but above the similarity values between already-described type strains are shown in blue boxes.

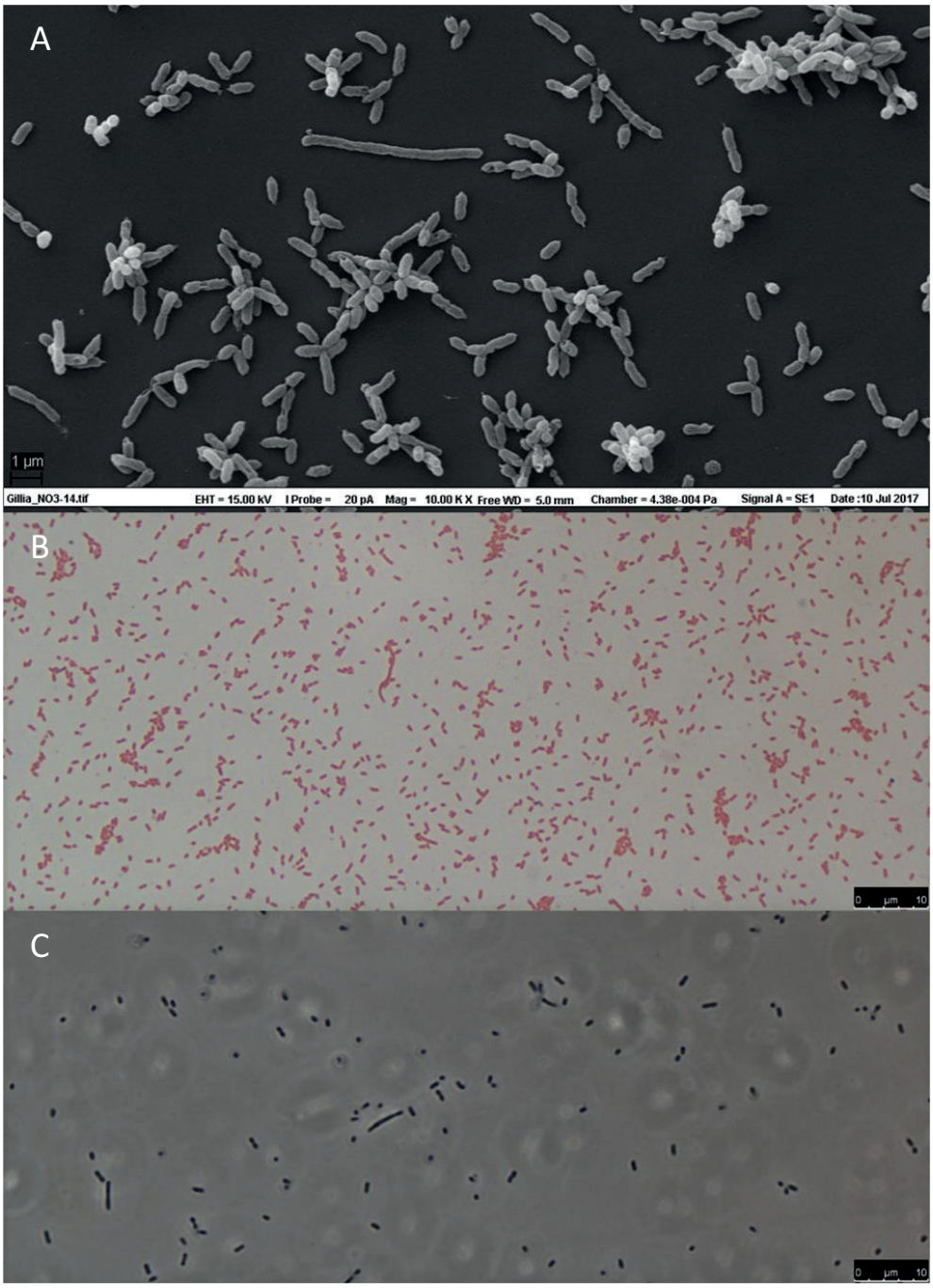

Figure 2: Microscopy of NO3^T by SEM a), light microscopy - Gram stained b), and light microscopy - phase contrast c). All pictures show rod-shaped cells that are approximately 1µm long, but filamentous cells and star shaped arrangements can be seen. Scale bar is shown in left (a) and in right lower corner (b and c).

Supplementary

Table S1: 16S rRNA gene similarity in % of strains included for 16S comparisons (Fig S1) towards the type strains: *G. apicola* wkB1^T, *G. apis* NO3^T, *G. intestini* LMG 28358^T, *G. bombicola* LMG 28359^T, *G. bombi* LMG 29879^T, *G. mensalis* LMG 29880^T.

Strain	***G. apicola***[1] **wkB1^T** CP007445	***G. apis*** **NO3^T** NASD01	***G. intestini*** **LMG 28358^T** LT631740.1	***G. bombicola*** **LMG 28359^T** LT631739.1	***G. bombi*** **LMG 29879^T** LT631738.1	***G. mensalis*** **LMG 29880^T** LT631737.1
wkB11	97.1	97.0	96.9	96.5	97.2	96.5
Fer2-1	99.0	98.0	98.6	98.6	98.5	98.6
Nev5-1	98.9	98.0	98.5	98.6	98.4	98.5
App4-10	98.1	97.3	98.1	98.4	97.9	98.4
Choc5-1	97.2	97.0	97.0	96.5	97.2	96.4
Imp1-6	98.8	98.1	98.9	98.4	99.8	98.3
wkB30	98.6	98.2	98.8	98.4	99.7	98.4
LMG 28358^T	98.5	97.8	-	98.6	99.0	98.5
LMG 28359^T	98.4	97.6	98.6	-	98.6	99.9
LMG 29879^T	98.7	97.9	99.0	98.6	-	98.5
LMG 29880^T	98.4	97.8	98.5	99.9	98.5	-
wkB108	98.1	97.7	97.5	97.0	97.5	97.0
wkB308	97.2	97.4	96.4	96.0	96.7	96.1
wkB292	99.0	98.7	98.1	98.0	98.2	98.2
AM4	99.4	99.0	98.1	98.4	98.4	98.3
SAG B02	99.2	98.9	98.0	98.0	98.1	98.1
N2	99.5	98.7	98.3	98.2	98.5	98.3
A-8-12	99.5	98.7	98.3	98.3	98.5	98.3
wkB7	99.6	98.8	98.4	98.4	98.6	98.4
N-G5	99.8	99.0	98.4	98.1	98.5	98.2
N-9-4	99.5	98.8	98.1	98.3	98.5	98.4
NO5	99.5	98.8	98.1	98.3	98.5	98.4
A-1-24	99.9	98.8	98.5	98.2	98.6	98.3
A8	99.9	98.8	98.5	98.2	98.6	98.3
wkB1^T	-	98.9	98.4	98.2	98.5	98.2
N22	99.4	98.4	97.9	98.1	98.6	98.1
SAG I20	99.1	99.5	97.9	97.8	98.1	98.0
M1-2G	99.0	99.6	97.7	97.7	97.8	97.9
SAG P17	98.7	99.5	97.6	97.5	97.8	97.7
P62G	99.3	99.7	98.0	98.0	98.2	98.2
N-G2	99.3	99.3	98.0	98.0	98.0	98.1
A7	99.0	99.3	98.0	98.0	98.0	98.0
N-G4	99.1	99.5	97.7	97.5	97.8	97.8
N-G1	99.3	99.4	98.0	97.9	98.2	98.1
AM1	99.0	99.4	97.9	97.5	97.8	97.8
A-TSA1	99.0	99.5	97.7	97.6	97.9	97.8
A-4-12	99.3	99.8	97.9	98.0	98.1	98.2
NO14	99.2	99.9	98.1	97.8	98.2	98.0
NO3^T	98.9	-	97.8	97.6	97.9	97.8
O. hercynius CN3^T	93.9	94.4	93.9	94.3	94.4	94.3
O. sasakiae C7^T	94.0	94.2	93.6	94.2	94.2	94.1
F. perrara PEB0191^T	95.3	95.6	94.7	95.0	95.0	95.2
S. bombi Bimp	96.2	96.1	95.9	96.5	95.6	96.5

[1] Values are taken from previous literature as described in Table 1.

Table S2: Growth on different agars and at different culturing conditions for four strains of *G. apis*.

Temp[2,3]	**30°C**				**36°**				**42°C**			
Strain	**NO3**	**TSA1**	**A7**	**G2**	**NO3**	**TSA1**	**A7**	**G2**	**NO3**	**TSA1**	**A7**	**G2**
CO_2												
TSA	+	+	+	+	+	+	+	+	(+)	(+)	(+)	(+)
TSA+B	+	+	+	+	+	+	+	+	(+)	(+)	(+)	(+)
HIA	+	+	+	+	+	+	+	+	(+)	(+)	(+)	(+)
HIA+B	+	+	+	+	+	+	+	+	(+)	(+)	(+)	(+)
MH	-	-	-	(+)	-	-	(+)	+	-	-	-	-
MH+B	+	(+)	(+)	+	+	+	+	+	-	-	-	-
Aerob												
TSA					-	-	-	-				
TSA+B					-	-	-	-				
HIA					-	-	-	-				
HIA+B					-	-	-	-				
MH					NT	NT	NT	NT				
MH+B					NT	NT	NT	NT				
Anaerobe												
TSA					+	+	+	+				
TSA+B					+	+	+	+				
HIA					+	+	+	+				
HIA+B					+	+	+	+				
MH					(+)	(+)	-	+				
MH+B					+	+	+	+				

[2] (+) = weak growth

[3] NT= Not tested

Table S3: Phenotypic characteristics from API NH and API 20E

Active ingrediens	Reactions/Enzymes	NO3[T]	A-TSA1	A7	N-G2	wkB1[T,4]
API NH						+
D-glukose	Acidification	+	+	+	+	+
D-fructose		+	+	+	+	-
D-maltose		-	-	-	-	+
D-saccharose (sucrose)		+	+	+	+	-
L-ornithine	Ornithine DeCarboxylase	-	-	-	-	-
Urea	Urease	-	-	-	-	-
5-bromo-3-indoxyl-caprate	Lipase	-	-	-	-	+
4-nitrophenyl-phosphate 2CHA	Alkaline Phosphatase	+	+	+	+	+
4-nitrophenyl-βD-galactopyranoside	βGalactosidase	-	-	-	-	-
Proline-4-methoxy-β-naphtylamide	Proline Arylamidase	-	-	-	-	+
Γ-glutamyl-4-methoxy- β-naphtylamide	Gamma Glutamyl Transferase	-	-	-	+	-
L-tryptophane	Indole	-	-	-	-	-
API 20E						
Trisodium citrate	INDole production	-	-	-	(+)[5]	-
Sodium pyruvate	Acetoin production (Vouges-Proskauer)	+	(+)[5]	+	(+)[5]	-
D-glucose	Fermentation	+	+	-	+	-
D-mannitol	Fermentation	-	-	-	(+)[5]	-
D-sucrose	Fermentation	(+)[5]	-	-	+	-
Amygdalin	Fermentation	+	+	-	+	-

[4] wkB1[T] (=DSMZ 104097)
[5] (+) = weak positive reaction after 48h

Table S4: Presence of sugar fermenting genes in NO3, A-TSA1, A7, N-G2

	NO3	**A-TSA1**	**A7**	**N-G2**
xylA	+	+	-	-
xylB	+	+	+	-
araA	-	-	-	-
araB	-	-	-	-
araD	-	-	-	-
manA	-	-	+	-
rhaA	+	+	-	+
rhaB	+	+	-	+
rhaD	+	+	-	+
PL1	-	-	-	-

Table S5: MICs for NO3^{T}, A-TSA1, A7, N-G2 and wkB1^{T}.

	NO3^{T}	**A-TSA1**	**A7**	**N-G2**	**wkB1^{T}** (DSMZ 104097)
Tetracycline	3.0 mg/l	3.0 mg/l	6.0 mg/l	0.19 mg/l	12.0 mg/l
Ampicillin	0.25 mg/l	0.19 mg/l	0.19 mg/l	0.25 mg/l	0.38 mg/l
Streptomycin	3.0 mg/l	0.5 mg/l	0.75 mg/l	8.0 mg/l	12.0 mg/l

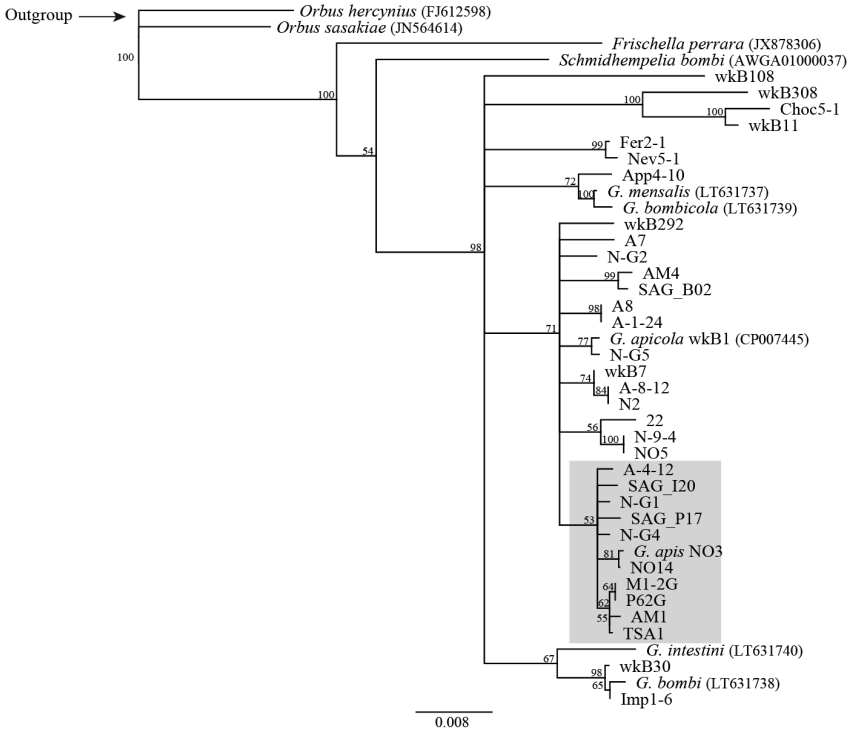

Figure S1: Neighbor-joining phylogeny made from 1195bp of the 16S rRNA gene of strains in table S1. Names of type strains are written out and the cluster of *G. apis* is highlighted in grey. Not all *Bombus* and *Apis* strains cluster separately as seen in Figure 1. Bootstrap values are shown in percent. Strain references are listed in Table 1.

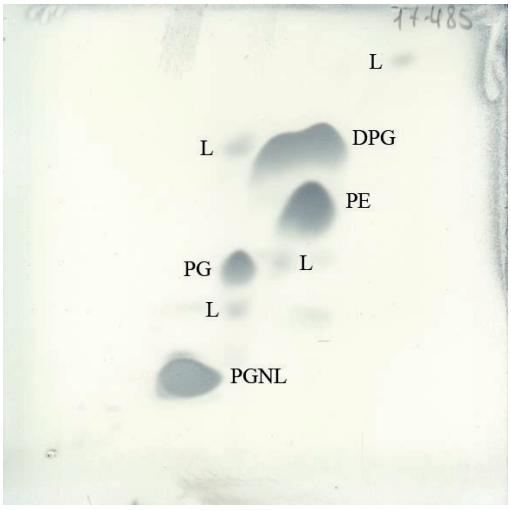

Figure S2: Polar lipid composition of NO3^T, identified by two-dimensional silica gel tin layer chromatography (DSMZ, Baunschweig, Germany). PG = Phosphatidylglycerol, PE = Phosphatidylethanolamine, DPG = Diphosphatidylglycerol, PGNL= Phosphoaminoglycolipid, L = Lipid

Leibniz-Institut
DSMZ-Deutsche Sammlung von
Mikroorganismen und Zellkulturen GmbH

21.11. 2017

Confirmation of the deposit and availability of a strain

The following information is confidential and serves only to allow the International Journal of Systematic and Evolutionary Microbiology to confirm that a strain has been deposited and will be available from the DSMZ in accordance with the Rules of the Bacteriological Code (1990 revision) as revised by the ICSP at the plenary sessions in Sydney and Paris.

The strain *Gilliamella* sp. NO3^{T} has been deposited in the DSMZ under the number **DSM 105629^{T}.**

This strain is available in the publically accessible section of the DSMZ and restrictions have not been placed on access to information concerning the presence of this strain in the DSMZ. It will be included in published and online catalogues after publication of this number by the authors.

This strain has been checked for viability in the DSMZ and is stored using one of the standard methods used in the DSMZ.

The depositor of this strain has also carried out a "depositor's check" and confirmed the identity of the strain held under this DSM number.

The DSMZ is not responsible for differences between the properties of the strain deposited in the DSMZ and properties given in the literature/databases.

E. Lang

Dr. Elke Lang
Curator responsible for the strain

GREAT AT SMALL THINGS

CERTIFICATE OF DEPOSIT

This is to certify that the following microorganism has been deposited into the BCCM/LMG Bacteria Collection and is available to the public without restriction:

LMG number: **LMG 30293**

Speciesname: **Gilliamella apis**

Depositor: **Ludvigsen Jane, Norwegian University of Life Sciences, Department of Chemistry, Biotechnology and Food Science**

Depositor no: **NO3**

Gent, 04 October 2017

ir. Claudine Vereecke
Public Collection Curator
BCCM/LMG Bacteria Collection

Laboratorium voor Microbiologie - Universiteit Gent (UGent)
K.L. Ledeganckstraat 35 - B-9000 Gent - Belgium
T +32 (0)9 264 51 08 - bccm.lmg@ugent.be

UNIVERSITEIT GENT
Belgian Science Policy Office
.be

Title

Linking streptomycin resistance genes (*strA-strB*) in a honeybee gut symbiont to environmental antibiotic exposure.

Jane Ludvigsen[1], Gro V. Amdam[2,3], Knut Rudi[1], Trine M. L'Abée-Lund[4]

[1]Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Chr. M. Falsens vei 1, 1430 Ås, Norway

[2]Faculty of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Høgskoleveien 12, 1430 Ås, Norway

[3]School of Life Sciences, Arizona State University, P.O. Box 874501, 427 East Tyler Mall, Tempe, AZ 85287, USA

[4]Department of Food Safety and Infection Biology, Norwegian University of Life Sciences, Campus Adamstuen, 0454 Oslo, Norway

Corresponding author

Jane Ludvigsen

E-mail: jane.ludvigsen@nmbu.no

Phone: +47 67232517

Running title

Streptomycin resistance genes in honeybees

Keywords

Honeybee gut symbiont, *Snodgrassella alvi*, streptomycin resistance, *strA-strB*

Abstract

Use of antibiotics in medicine and farming contributes to increasing numbers of antibiotic resistant bacteria in diverse environments. The ability of antibiotic resistance genes (ARG) to transfer between bacteria genera contributes to this spread. It is difficult to directly link antibiotic exposure to spread of ARG in a natural environment where environmental settings and study populations cannot be fully controlled. We used managed honeybees in environments with contrasting streptomycin exposure (USA: high exposure, Norway: low exposure) and mapped the prevalence and spread of transferrable streptomycin resistance genes. We found high prevalence of *strA-strB* genes in the USA compared to Norway with 17/90 and 1/90 positive samples, respectively ($p<0.00007$). We identified *strA-strB* genes on a transferrable transposon Tn*5393* in the honeybee gut symbiont *Snodgrassella alvi*. Such transfer of resistance genes increases risk of spread to new environments as honeybees are moved to new pollination sites.

Main

Antibiotic resistance genes (ARG) can spread between pathogenic and symbiotic bacteria through horizontal gene transfer (HGT), and further between host organisms: e.g. humans, farm animals, and insects [1]. Insect hosts can effectively spread antibiotic resistant bacteria (ARB) between environments [2]. Prevalence of ARG is positively correlated with use and dispersal of antibiotics within an environment [3,4], and even low levels of antibiotics can cause resistance [5]. Studies on how ARG spread in natural settings suffer from lack of standardized study populations and appropriate control over environmental factors. We take advantage of managed honeybees, which live in homeostatically regulated, highly structured societies. We study prevalence and spread of acquired streptomycin resistance in the gut microbiota of honeybees located in areas with contrasting streptomycin exposure.

We used a previous sampled dataset of 180 bees from two research apiaries: (i) Norwegian University of Life Sciences and (ii) Arizona State University [6]. Antibiotics are not applied in Norwegian agriculture, while the antibiotic streptomycin is used in the USA, e.g. prophylactic by spraying apples and pear trees repeatedly during tree blossoming [4,7]. Honeybees are important pollinators of fruit trees and applied extensively at the time of tree blooming, leading to possible exposure to varying concentrations of streptomycin left on the fruit flowers [4]. We hypothesized that honeybee gut bacteria can acquire and maintain streptomycin resistance genes from environmental streptomycin resistant bacteria. The development of ARB in honeybees is of considerable concern since these insects are transported large distances for pollination purposes, thus representing a real source of ARG spread through fecal droppings and interaction with environmental bacteria.

Resistance to streptomycin is due to gene mutations or acquired resistance genes such as *strA-strB* [8,9]. The latter is of special concern since *strA-strB* genes are mostly associated with transferrable elements, like small non-conjugative broad host range plasmids or self-transferrable (conjugative) plasmids and the transposon Tn5393 [10,11]. These transferrable elements are found in a range of environmental and pathogenic bacteria [11,10,12]. We screened for the prevalence of the linked *strA-strB* resistance genes using quantitative real-time PCR (qPCR), on extracted DNA from midgut samples (Supplementary). We found that 17 out of 90 samples were positive from Arizona, but only one of 90 samples was positive from Norway ($p < 0.00007$, Chi-squared=15.8). This first report of *strA-strB* genes in honeybees suggests that the use of streptomycin in US agriculture is associated with higher *strA-strB* gene prevalence in honeybee gut microbiota.

To explore which bacterium that harbor *strA-strB* genes, we used shotgun metagenome sequencing (Nextera XT, Illumina) (Supplementary). In samples from Arizona, we detected *strA-strB* genes by local blast search, with flanking gene sequences identified as the honeybee gut symbiont *Snodgrassella alvi* with 96% nt identity (CP007446).

Next, we isolated streptomycin resistant gut bacteria from the Arizonian samples for subsequent genome sequencing (Nextera XT, Illumina) (Supplementary). By qPCR, we identified one strain (E1) that was *strA-strB* positive. Sanger sequencing of the 16S rRNA gene (1163bp) confirmed this strain to be *S. alvi* (99% similarity, CP007446). This finding corroborated that *S. alvi* harbors streptomycin resistance genes. The assembled *S. alvi*_E1 genome (supplementary) was screened for known ARGs using ResFinder V3.0 [13], and a 100% match for *strA-strB* genes was detected towards accession number M96392 (NCBI), which is located within the Tn*5393* transposon in the plasmid pEa34 identified in *Erwinia amylovora* [11]. *E. amylovora* is a plant pathogen of apples and pears. By aligning the annotated assembled genome towards the 6705bp long Tn*5393* transposon (M96392), we resolved that our transposon was divided on two gene fragments with flanking genomic genes on each side. The Tn*5393* association with the *strA-strB* genes was verified by successfully amplifying a 2112 bp long PCR fragment that spanned the *strA* gene and the *tnpR* gene of the transposon (Fig 1)(supplementary). The gene repertoire of the Tn*5393* transposon vary slightly, and is restricted to five nt mutations within the *StrA* gene and the presence of insertion elements (IS1133 and IS6100) [14,15]. We identified IS1133 in Tn*5393* in *S. alvi*_E1 by amplifying a 799bp PCR fragment spanning the *strA* gene and IS1133. We used the Sanger sequenced PCR product to connect the two gene fragments to produce Figure 1.

Summarized, our results describe a 6764bp long Tn*5393* transposon that contains IS1133. An identical Tn*5393* was previously identified in *Escherichia coli* plasmid pVI-W9608 (EF108308). This illustrates the close link to human pathogens and plasmids with Tn*5393* can transfer between distantly related bacteria, e.g. the fish pathogen *Aeromonas salmonicida* and the plant pathogen *E. amylovora* to *Enterobacteriaceae* [11,12]. Both *E. coli* and *E. amylovora* can be found in honeybee gut as transient environmental bacteria, and *E. amylovora* is transported around by honeybees while pollinating fruit flowers and can survive in the bee gut for 36 hours [16]. This scenario plays out for other environmental bacteria as well [17], making the honeybee gut microbiota a possible hotspot for HGT and spread of ARG [18].

We identified Tn*5393* within the genome of the *S. alvi*_E1 strain and not on any plasmid. Although not frequent, this has been confirmed in *Salmonella* strains of animal origin and specific *E. amylovara* strains [19,20]. In *S. alvi*, some tetracycline resistance genes have been identified within the genome [6] indicating that ARG can become stable genomic genes that are maintained across generations.

In addition, we compared the streptomycin MIC of *S. alvi*_E1 with 13 *S. alvi* without *strA-strB* genes (WT) from both Norway and Arizona (supplementary). The MIC range of WT isolates was between 0.38-0.75 µg/ml and MIC of *S. alvi*_E1 was 12.0 µg/ml (Table S1). MIC of streptomycin resistance bacteria varies with bacteria origin and is highly influenced by the genetic background [9] and elevated expression when in associated with IS elements [15]. In our *S. alvi* isolate, MIC is low compared to other bacteria harboring IS1133 [19].

In conclusion, we are the first to report horizontal gene transferred (HGT) streptomycin resistance genes in a honeybee gut symbiont. Our data are consistent with a direct link between the use of streptomycin in agriculture and dispersal of streptomycin resistant genes. The identification of streptomycin resistance on a transferrable element in the honeybee gut microbiota contributes to the knowledge of ongoing spread of ARG between environmental bacteria and human pathogens, which underlines the importance of reducing the use of antibiotics in medicine and agriculture whenever possible.

Data availability

The assembled genome is deposited at NCBI with accession number NXEN00000000.
The Tn*5393* sequence is deposited at NCBI with the accession number: MG704836.
The metagenome data are available on request.

Acknowledgements

We thank Davide Porcellato for performing the genome sequencing of the *S. alvi*_E1 strain and Inga Leena Angell for performing the metagenome sequencing of the Norwegian sample.

Conflict of Interest: The authors declare that they have no conflict of interest.

148 References

149

150 1. Djordjevic SP, Stokes HW, Roy Chowdhury P (2013) Mobile elements, zoonotic
151 pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans,
152 production animals and soil microbiota. Frontiers in Microbiology 4:86.
153 doi:10.3389/fmicb.2013.00086
154 2. Zurek L, Ghosh A (2014) Insects represent a link between food animal farms and the urban
155 environment for antibiotic resistance traits. Appl Environ Microbiol 80 (12):3562-3567.
156 doi:10.1128/AEM.00600-14
157 3. Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A, Bork P (2013)
158 Country-specific antibiotic use practices impact the human gut resistome. Genome Res 23
159 (7):1163-1169. doi:10.1101/gr.155465.113
160 4. McManus PS, Stockwell VO, Sundin GW, Jones AL (2002) Antibiotic use in plant
161 agriculture. Annu Rev Phytopathol 40:443-465.
162 doi:10.1146/annurev.phyto.40.120301.093927
163 5. Andersson DI, Hughes D (2014) Microbiological effects of sublethal levels of antibiotics.
164 Nat Rev Microbiol 12 (7):465-478. doi:10.1038/nrmicro3270
165 6. Ludvigsen J, Porcellato D, L'Abée-Lund TM, Amdam GV, Rudi K (2017) Geographically
166 widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant
167 patterns. Molecular Ecology 26 (23):6590-6607. doi:10.1111/mec.14392
168 7. Bonaterra A, Badosa E, Cabrefiga J, Frances J, Montesinos E (2012) Prospects and
169 limitations of microbial pesticides for control of bacterial and fungal pomefruit tree diseases.
170 Trees (Berl West) 26 (1):215-226. doi:10.1007/s00468-011-0626-y
171 8. Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, Pan Y, Li J, Zhu L, Wang X, Meng Z, Zhao
172 F, Liu D, Ma J, Qin N, Xiang C, Xiao Y, Li L, Yang H, Wang J, Yang R, Gao GF, Wang J,
173 Zhu B (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of
174 human gut microbiota. Nat Commun 4:2151. doi:10.1038/ncomms3151
175 9. Sunde M, Norstrom M (2005) The genetic background for streptomycin resistance in
176 Escherichia coli influences the distribution of MICs. The Journal of antimicrobial
177 chemotherapy 56 (1):87-90. doi:10.1093/jac/dki150
178 10. Sundin GW, Bender CL (1996) Dissemination of the strA-strB streptomycin-resistance
179 genes among commensal and pathogenic bacteria from humans, animals, and plants. Mol
180 Ecol 5 (1):133-143
181 11. Chiou CS, Jones AL (1993) Nucleotide sequence analysis of a transposon (Tn5393)
182 carrying streptomycin resistance genes in Erwinia amylovora and other gram-negative
183 bacteria. Journal of bacteriology 175 (3):732-740
184 12. L'Abée-Lund TM, Sørum H (2000) Functional Tn5393-Like Transposon in the R Plasmid
185 pRAS2 from the Fish Pathogen Aeromonas salmonicida subspecies salmonicida Isolated in
186 Norway. Applied and Environmental Microbiology 66 (12):5533-5535
187 13. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup
188 FM, Larsen MV (2012) Identification of acquired antimicrobial resistance genes. Journal of
189 Antimicrobial Chemotherapy 67 (11):2640-2644. doi:10.1093/jac/dks261
190 14. Sundin GW (2000) Examination of base pair variants of the strA-strB streptomycin
191 resistance genes from bacterial pathogens of humans, animals and plants. The Journal of
192 antimicrobial chemotherapy 46 (5):848-849
193 15. Sundin GW, Bender CL (1995) Expression of the strA-strB streptomycin resistance genes
194 in Pseudomonas syringae and Xanthomonas campestris and characterization of IS6100 in X.
195 campestris. Appl Environ Microbiol 61 (8):2891-2897

16. Alexandrova M, Porrini C, Bazzi C, Carpana E, Bigliardi M, Sabatini AG (2002) Erwinia amylovora longevity in beehives, beehive products and honeybees. Acta Horticult 590:201–205
17. McArt SH, Koch H, Irwin RE, Adler LS (2014) Arranging the bouquet of disease: floral traits and the transmission of plant and animal pathogens. Ecol Lett 17 (5):624-636. doi:10.1111/ele.12257
18. Huddleston JR (2014) Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. Infect Drug Resist 7:167-176. doi:10.2147/IDR.S48820
19. Forster H, McGhee GC, Sundin GW, Adaskaveg JE (2015) Characterization of Streptomycin Resistance in Isolates of Erwinia amylovora in California. Phytopathology 105 (10):1302-1310. doi:10.1094/PHYTO-03-15-0078-R
20. Pezzella C, Ricci A, DiGiannatale E, Luzzi I, Carattoli A (2004) Tetracycline and Streptomycin Resistance Genes, Transposons, and Plasmids in Salmonella enterica Isolates from Animals in Italy. Antimicrobial Agents and Chemotherapy 48 (3):903-908. doi:10.1128/aac.48.3.903-908.2004

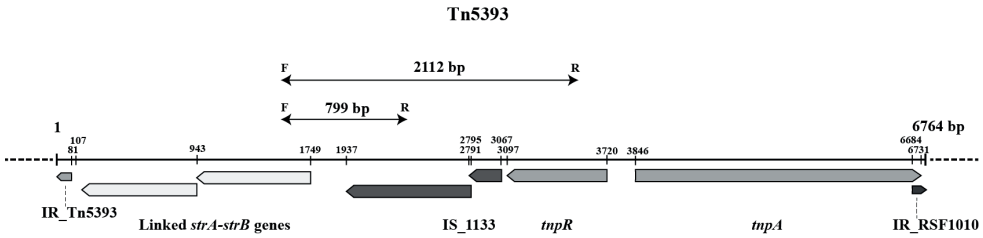

Figure 1: Drawn visualization of the 6764bp long Tn*5393* transposon identified in *S. alvi*_E1 (GenBank MG704836). Length of associated genes are shown as numbers above on each side of respective genes, and genes are visualized with different coloring and names below genes. PCR fragments used to produce the Tn*5393* consensus of *S. alvi*_E1 are shown as arrows spanning the different genes targeted. The genomic region of *S. alvi*_E1 is shown as the dotted lines on each side of the Tn*5393* transposon.

Supplementary data

Tables

Table S1: Overview of *strA-strB* gene presence and corresponding streptomycin MIC in *S. alvi* isolates. Isolates with the detonation N and A represent *S. alvi* strains from Norway and Arizona, respectively.

Strain	*strA-strB**	Streptomycin µg/ml
N-S2	-	0.38
N-S3	-	0.75
N-S4	-	0.75
N-S5	-	0.38
NW-7	-	0.75
AW-18	-	0.5
A-2-12	-	0.75
A-9-24	-	0.75
A-10-12	-	0.75
A-11-12	-	0.75
A-5-24	-	0.75
A3	-	0.5
A5	-	0.5
A-E1	+	**12**

*Results from qPCR screens = not detected (-) and detected (+).

Materials and Methods

For *strA-strB* qPCR screens, we used primers as described in Sunde et al. [21] with an end concentrations of 0.2µM in an 1×concentration of EvaGreen mix (Solis biodyne, Estonia) on Light Cycler 480II (Roche, Switzerland), with the following term cycling condition: 95°C for 15min, then 40 cycles of 95°C for 30sec, 60°C for 30sec, and 72°C for 60sec. A negative (nuclease-free water) and positive DNA control (E.coli DH5α with resistance plasmid pRAS2 with *strA-strB* genes [12]) were included in each run. Samples with one cq value less than negative control, were selected as positive. We verified the 538bp s*trA-strB* PCR fragment using Sanger sequencing and the LightRun service of GATC (Germany) and subsequently as

query in BLAST searches at NCBI (nt database) in where we got 100% match, over the whole fragment length, against *strA-strB* genes origination from several bacteria genera.

For shotgun metagenome sequencing, we used Nextera XT DNA preparation kit, Illumina and selected three samples with low Cq values from the Arizonan dataset and the positive sample from the Norwegian dataset. The prepared samples were run on the MiSeq platform (Illumina, USA) following manufacturers' protocol and the 300bp paired-end v3 kit. Honeybee-specific sequences were removed by mapping against the *Apis mellifera* genome using Geneious V8.1 [22] and subsecuently 1,325401 paired reads from all three Arizonan sample were assembled into 74,755 contigs with N50 of 893bp (min length 35bp and max length 33,954bp) using Geneious v8.1. Sanger sequenced PCR fragments of *strA-strB* was queried against the metagenome assembly and we identified the *strA-strB* genes in the Arizonan samples, in one contig of length 8135bp. No antibiotic resistance gene transfer element was identified in the metagenome assembled contig containing the *strA-strB* genes, but the flanking region of the *strA-strB* genes were queried at nt database, NCBI and we got a hit towards the honeybee gut symbiont *Snodgrassella alvi* with 96% identity. The Norwegian sample did not assemble well due to low coverage and we therefore were not able to identify which bacterium that contained the *strA-strB* genes.

We verified our finding by culturing streptomycin resistant bacteria from the honeybee gut and screened them for the *strA-strB* genes by qPCR as described before. For isolation of bacteria, we used 30 bees from each location, sampled in a previous dataset [6] and 100µl of 10^{-2} and 10^{-3} dilution of pooled homogenized (prepared as described in Ludvigsen et al. [6]) gut samples was spread on tryptic soy agar with added 5% horse blood and 4µg/ml and 12µg/ml streptomycin (Sigma Aldrich, Germany). We selected 26 strains with different morphology from Norway and Arizona from all plates. DNA was extracted as described previously in Ludvigsen et al. [6] and qPCR screen for *strA-strB* was performed as described previously form midgut samples. PCR for Sanger sequencing (GATC, Germany) of the 16S rRNA gene of the positive *strA-strB* strain was performed with CoverAll primers as described in Ludvigsen et al. [23].

We genome sequenced the isolated *S. alvi*_E1 strain using the Nextera XT DNA preparation kit, Illumina, following manufactures recommendations, as described for metagenome sequencing of midtgut samples. Fastq files were trimmed for quality below Q20 and we

assembled the genome using velvet plug in for Geneious V8.1 [24]. Contigs with length less than 200bp were removed and the final assembly consisted of 156 contigs (L50= 47,816bp). The assembled genome was annotated using RAST [25].

A positive PCR fragment, that spanned the *strA*- and the *tnpR* gene of the transposon, confirmed the Tn5393 association with the *strA-strB* genes. We used the forward primer 2F and reverse primer smAR as described in Petrova et al. [26] and Pezzella et al. [20], respectively at a 0.2µM end concentration in an 1×concentration of ready to load mix (RTL Polymerase mix, Solis biodyne, Estonia). The following term cycling conditions were applied: 95°C for 15min, then 35 cycles of 95°C for 30sec, gradient span from 48-60°C for 30sec, and 72°C for 60sec. The PCR product was visualized on 1,5% agarose gel.
For identification of IS1133 we used the primers IS1133F and smAR as previously described by Pezzella et al. [20], at a 0.2µM end concentration in an 1×concentration of ready to load mix (RTL Polymerase mix Solis biodyne, Estonia). The following term cycling conditions were applied: 95°C for 15min, then 35 cycles of 95°C for 30sec, 54°C for 30sec, and 72°C for 60sec. The PCR product was visualized on a 2% agarose gel and Sanger sequenced using the LightRun service of GATC (Germany). The two identified contigs from the genome assembly harboring the *strA-strB* genes and the transposon were mapped against the Sanger sequenced PCR fragment in Geneious v 8.1 to produce a consensus sequence of Tn5393 in *S. alvi*_E1 (Figure 1).
Streptomycin MIC was performed by using MIC Test Strips (Montebello Diagnostics AS, Norway). Inoculum of 1.0 McFarland of each strain was spread (EUCAST guidelines, http://www.eucast.org/) on hart infusion agar with 5% horse blood, and incubated for three days at 37 °C in an enhanced CO_2 atmosphere (GasPack EZ CO2 container system; Becton Dickinson, USA).

344 References

345

346 1. Djordjevic SP, Stokes HW, Roy Chowdhury P (2013) Mobile elements, zoonotic
347 pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans,
348 production animals and soil microbiota. Frontiers in Microbiology 4:86.
349 doi:10.3389/fmicb.2013.00086
350 2. Zurek L, Ghosh A (2014) Insects represent a link between food animal farms and the urban
351 environment for antibiotic resistance traits. Appl Environ Microbiol 80 (12):3562-3567.
352 doi:10.1128/AEM.00600-14
353 3. Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A, Bork P (2013)
354 Country-specific antibiotic use practices impact the human gut resistome. Genome Res 23
355 (7):1163-1169. doi:10.1101/gr.155465.113
356 4. McManus PS, Stockwell VO, Sundin GW, Jones AL (2002) Antibiotic use in plant
357 agriculture. Annu Rev Phytopathol 40:443-465.
358 doi:10.1146/annurev.phyto.40.120301.093927
359 5. Andersson DI, Hughes D (2014) Microbiological effects of sublethal levels of antibiotics.
360 Nat Rev Microbiol 12 (7):465-478. doi:10.1038/nrmicro3270
361 6. Ludvigsen J, Porcellato D, L'Abée-Lund TM, Amdam GV, Rudi K (2017) Geographically
362 widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant
363 patterns. Molecular Ecology 26 (23):6590-6607. doi:10.1111/mec.14392
364 7. Bonaterra A, Badosa E, Cabrefiga J, Frances J, Montesinos E (2012) Prospects and
365 limitations of microbial pesticides for control of bacterial and fungal pomefruit tree diseases.
366 Trees (Berl West) 26 (1):215-226. doi:10.1007/s00468-011-0626-y
367 8. Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, Pan Y, Li J, Zhu L, Wang X, Meng Z, Zhao
368 F, Liu D, Ma J, Qin N, Xiang C, Xiao Y, Li L, Yang H, Wang J, Yang R, Gao GF, Wang J,
369 Zhu B (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of
370 human gut microbiota. Nat Commun 4:2151. doi:10.1038/ncomms3151
371 9. Sunde M, Norstrom M (2005) The genetic background for streptomycin resistance in
372 Escherichia coli influences the distribution of MICs. The Journal of antimicrobial
373 chemotherapy 56 (1):87-90. doi:10.1093/jac/dki150
374 10. Sundin GW, Bender CL (1996) Dissemination of the strA-strB streptomycin-resistance
375 genes among commensal and pathogenic bacteria from humans, animals, and plants. Mol
376 Ecol 5 (1):133-143
377 11. Chiou CS, Jones AL (1993) Nucleotide sequence analysis of a transposon (Tn5393)
378 carrying streptomycin resistance genes in Erwinia amylovora and other gram-negative
379 bacteria. Journal of bacteriology 175 (3):732-740
380 12. L'Abée-Lund TM, Sørum H (2000) Functional Tn5393-Like Transposon in the R Plasmid
381 pRAS2 from the Fish Pathogen Aeromonas salmonicida subspecies salmonicida Isolated in
382 Norway. Applied and Environmental Microbiology 66 (12):5533-5535
383 13. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup
384 FM, Larsen MV (2012) Identification of acquired antimicrobial resistance genes. Journal of
385 Antimicrobial Chemotherapy 67 (11):2640-2644. doi:10.1093/jac/dks261
386 14. Sundin GW (2000) Examination of base pair variants of the strA-strB streptomycin
387 resistance genes from bacterial pathogens of humans, animals and plants. The Journal of
388 antimicrobial chemotherapy 46 (5):848-849
389 15. Sundin GW, Bender CL (1995) Expression of the strA-strB streptomycin resistance genes
390 in Pseudomonas syringae and Xanthomonas campestris and characterization of IS6100 in X.
391 campestris. Appl Environ Microbiol 61 (8):2891-2897

16. Alexandrova M, Porrini C, Bazzi C, Carpana E, Bigliardi M, Sabatini AG (2002) Erwinia amylovora longevity in beehives, beehive products and honeybees. Acta Horticult 590:201–205
17. McArt SH, Koch H, Irwin RE, Adler LS (2014) Arranging the bouquet of disease: floral traits and the transmission of plant and animal pathogens. Ecol Lett 17 (5):624-636. doi:10.1111/ele.12257
18. Huddleston JR (2014) Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. Infect Drug Resist 7:167-176. doi:10.2147/IDR.S48820
19. Forster H, McGhee GC, Sundin GW, Adaskaveg JE (2015) Characterization of Streptomycin Resistance in Isolates of Erwinia amylovora in California. Phytopathology 105 (10):1302-1310. doi:10.1094/PHYTO-03-15-0078-R
20. Pezzella C, Ricci A, DiGiannatale E, Luzzi I, Carattoli A (2004) Tetracycline and Streptomycin Resistance Genes, Transposons, and Plasmids in Salmonella enterica Isolates from Animals in Italy. Antimicrobial Agents and Chemotherapy 48 (3):903-908. doi:10.1128/aac.48.3.903-908.2004
21. Sunde M, Fossum K, Solberg A, Sorum H (1998) Antibiotic resistance in Escherichia coli of the normal intestinal flora of swine. Microbial drug resistance (Larchmont, NY) 4 (4):289-299. doi:10.1089/mdr.1998.4.289
22. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England) 28 (12):1647-1649. doi:10.1093/bioinformatics/bts199
23. Ludvigsen J, Rangberg A, Avershina E, Sekelja M, Kreibich C, Amdam G, Rudi K (2015) Shifts in the Midgut/Pyloric Microbiota Composition within a Honey Bee Apiary throughout a Season. Microbes Environ 30 (3):235-244. doi:10.1264/jsme2.ME15019
24. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18 (5):821-829. doi:10.1101/gr.074492.107
25. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75. doi:10.1186/1471-2164-9-75
26. Petrova MA, Gorlenko ZM, Soina VS, Mindlin SZ (2008) Association of the strA-strB genes with plasmids and transposons in the present-day bacteria and in bacterial strains from permafrost. Russian Journal of Genetics 44 (9):1116-1120. doi:10.1134/s1022795408090160

07.04.2017

SKJEMA 4.7 Errata

Retting av formelle feil i avhandlingen (jf. §15.3-2 i ph.d.-forskriften)

Ph.d.-kandidaten kan, etter innlevering, søke fakultetet om tillatelse til å rette formelle feil i avhandlingen. En fullstendig oversikt over de feil (errata) som ønskes rettet skal skrives inn i dette skjemaet og leveres fakultetet senest 4 uker før planlagt disputas. Det kan søkes kun én gang.

Avhandlingens tittel:	Spread and persistence of antibiotic resistance genes in the honeybee gut microbiota

Side nr	Avsnitt	Endret fra	Endret til
19	Table 1	*Snodgrassella* alvi	*Snodgrassella alvi*
19	Table 1	*Frishella perrara*	*Frischella perrara*
23	2	bacteria is present	bacteria are present
23	2	link which bacteria to which	link which bacterium to which
30	2	but these were also mostly environmental	but these bacteria were mostly environmental
48	2	in environmental bacteria (ref),	in environmental bacteria [194],
52	3	honeybee core microbiota is also diet dependent	honeybee core microbiota was also diet dependent
52	3	diet or season, but their relative	diet or season, their relative

Dette skjemaet signeres av ph.d.-kandidat og hovedveileder og oversendes fakultetet for godkjenning. Godkjent errata arkiveres i ph.d.-kandidatens doktorgradsmappe, og legges ved den endelige avhandlingens trykk-versjon som siste side.

Dato og signatur:

Ph.d.-kandidat (forfatter):	19/2-18 Jane Ludvigsen

Hovedveileder:	18/2-18 Kn + Ruf.

Errata godkjent av fakultetet: Ja ☒ Nei ☐

For fakultetet:	26.02.18 Åsmund Asheim

Norges miljø- og biovitenskapelige universitet

ISBN: 978-82-575-1504-1
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

97528 / ANDVORDGRAFISK.NO