Background

Honeybees are highly social animals. They rely on each other to ensure the greater good of their society/hive they are a member of. Honeybees in a hive can be broadly categorized into three groups that are nurses, foragers, and winters bees. The nurses stay in the hive and nurse the larvae, the foragers go out to collect pollen and nectar to feed the larvae for hive proliferation, and lastly, the winter bees are the ones whose purpose is to live through the winter to ensure the survival of the hive. The gut microbiome composition of a honeybee is specific although relatively simple. There have been studies centered around the compositional changes in the gut microbiota, yet few have focused on the changes in the absolute bacterial abundance. Furthermore, not enough information is present about the gut microbiota of winter bees and as mentioned above, winter bees are essential for the survival of the hive into the following foraging season.

The members of the gut microbiome phylotypes studied here are *Gilliamela, Snodgrassellla, Bifidobacterium, Lactobacillus* Firm-4*, Lactobacillus* Firm-5 which are core members and present in almost every honeybee. The two additional non-cores that are also studied are *Bartonella and Frischella*. Quantification of these 7 core members was done in a single colony for a period of 2 years. The data and analysis showcased that the longer-living winter bees had greater bacterial load and lower species richness. The majority was taken up by *Commensalibacter* and *Bartonella*. Overall, the data analysis revealed that there lies a considerable difference in the gut microbiome of winter and foraging bees and greater attention should be paid to the role it plays in the ability of the winter bees to survive the winter.

Studies on gut microbiota contents of bees have been carried out before, aimed at observing the composition of the gut microbiota of different bee types or during different seasons. They all drew the same conclusion that the community composition remains stable. However, a caveat of these previous studies is that they were mostly based on comparative analyses of relative abundances by using 16S rRNA gene amplicon sequencing. Such analyses cannot provide insights into the extent or directionality of abundance changes, especially if microbial loads vary substantially between samples.

Also, change in the total abundance of the microbiota could by itself be an important characteristic of different bee types, seasons, or environmental exposure. This is crucial because winter bees are critical for colony health and survival during the cold season when resources are limited, and most colony losses occur.qPCR and 16S rRNA gene amplicon sequencing was used to assess differences in the gut microbiota of nurses, foragers, and winter bees. Bacterial loads of major community members in 566 individual worker bees sampled from a single hive over 2 years were used. 24 adult worker bees were sampled of the kind *Apis mellifera* each month from a single hive located on the Dorigny campus of the University of Lausanne, Switzerland. Sample collection during the foraging season was collected from foragers returning to the hive entrance with pollen on their legs. During the winter months, sampling of the winter bees was taken from the top of the frames from inside the hive. Each sampling time point took place at the middle of each month (±3 days) between April 2015 and April 2017 (Kešnerová, Trolio, Liberti, Erkosar, & Engel, 2019).

Data Analysis

All the data analysis was done on the data contained in the file `02_df_mm_calc.csv`.The relevant data for the analysis performed was majorly comprised of the microbe targets, the sample number, sampling date, the bee type, and the normalized cell numbers. The microbes are referred to by a shorthand in the data and will be done so similarly in this report. They are *Gilliamela (Ga), Snodgrassella (Sa), Bifidobacterium (Bi), Lactobacillus* Firm-4 (F4)*, Lactobacillus* Firm-5 (F5)*, Bartonella (Ba),* and *Frischella (Fp).*
The first analysis as shown in Figure 1A. was to find the monthly changes in the absolute abundance assessed by qPCR, as determined by the number of genome equivalents per sample, of seven phylotypes

monitored monthly, depicted as mean values (±SE) of the analyzed bees. The first steps for this were to extract the data and store it in a pandas dataframe. From there the processed data was further cleaned up by removing the redundant indexing and the data for the universal primers. The rows where the target microbe was not detected were also removed. To achieve the desired statistical parameters for plotting later a custom function was defined called `summarySE`. This function took in the cleaned-up dataframe and the desired confidence interval as the function inputs. The function returns a dataframe called `stat`. This dataframe is indexed by the date and the corresponding microbe target microbe along with the calculated stats which were 'N' (number of occurrences), the mean of the cell numbers, the standard deviation of the mean, standard error, and the confidence interval of 0.95. Most of the statistics were calculated using predefined python functions. The confidence interval was calculated using by first calculating the confidence interval multiplier `ciMulti` with the ppf (point percent function) function of the `scipy.stats` library and then multiplying with the standard error.

The above-created dataframe was then organized into a dictionary to later make use of the `plot_date()` function of the `matplotlib` library. This dictionary contained the target microbe as the key and the date, mean cell number and the standard errors as the corresponding values. The data was plotted with the date on the x-axis and the log of the mean cell numbers for the corresponding date on the y-axis. Separate lines were used to represent the mean abundance pattern of each different target microbe, with the core members marked with solid lines and the two non-cores with dotted lines along with bars representing the standard error.

The second analysis as shown in Figure 1B is the Relative community composition of the gut microbiota of bees sampled in each month as based on the seven monitored phylotypes. To plot this graph, the original data had to be reorganized. The reorganized data was converted to a dataframe where the indexing was done with sample IDs and the columns were the 7 phylotypes being studied and the values in the columns were the cell numbers for each phylotype for that corresponding sample ID in the indexing. In other words, the cell numbers were stored according to the sample IDs along the rows for all the 7 phylotypes. This reformatting of the data was done using nested-for-loops. Now to find the relative abundance. The first statistic required was the row-wise sum of the cell numbers for all the phylotypes for every sample ID. This was done using the `pandas` libraries `sum()` function and a new column called 'sum' was added into the dataframe that was created above. To find the relative abundance for each of the microbe for a particular sample ID, the cell numbers for each microbe had to be divided by the sum of cell numbers in for that sample ID stored om the last column. This process was applied to every sample ID to achieve a dataframe with the relative abundance as the datapoints with each microbe as the column and indexing done with the sample ID. The relative abundance dataframe was then plotted as a 100% stacked bar graph, with the sample IDs on the x-axis and relative abundance as % on the y-axis.

For the third analysis which was the Projection of the abundances of monitored phylotypes into the first and second principal components in all analysed bees, together with correlation vectors representing variables driving the separation on both axes as shown in Figure 2. For this analysis the first step involved repurposing the code block used in the above analysis for creating a data frame where the columns were the 7 phylotypes and the cell number were the values in each column. The 0s or missing values were then substituted by a one and the log of the whole dataframe was stored in another dataframe. This is the dataframe for whom the principal components were calculated .Using functions from the package `sklearn`, the data was scaled and then the principal components were calculated and stored in a new dataframe. The `np.round()` function with the `pca.explained_variance_ratio_* 100` was also used to calculate the variance captured by each principal component. A column with the corresponding bee type was added to the PCA dataframe to colour the datapoints during plotting accordingly . The principal components 1 and 2 were plotted on the x and y axis respectively with each point was coloured according to the type of bee it belonged to either winter or forager. The projections or eigenvectors for each phylotype were plotted, they have been multiplied by factor of 2.5 so as to make it easier to visualize the difference in magnitude amongst the phylotypes, the projections are not a true representation of the eigenvectors and values but the relative relation amongst the phylotypes is maintained and easier to realize this way.

Conclusion

The conclusions and inferences that can be drawn from the data analysis performed on the 16s rRNA data of honeybee gut microbiome will be elucidated in this section.

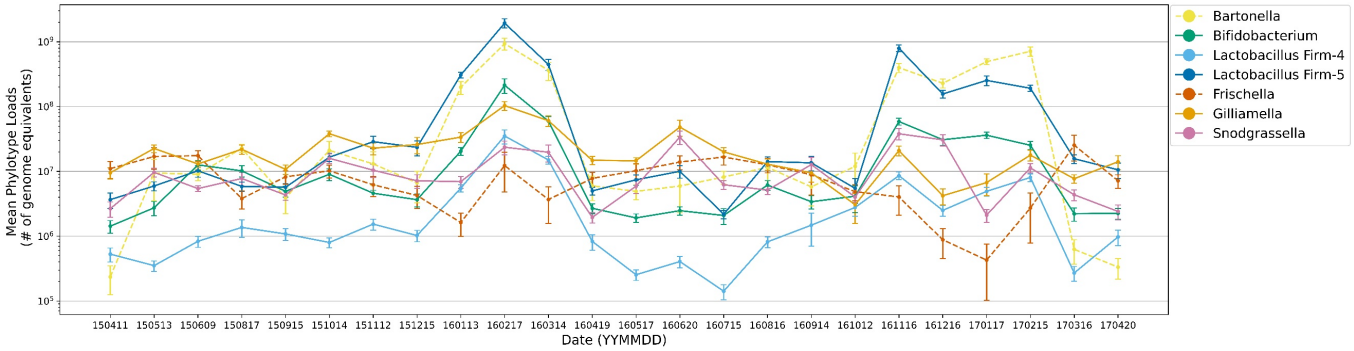

Figure 1A

In Figure 1A the absolute abundance of monitored phylotypes can be clearly seen, the variation amongst the phylotypes was not very significant as can be seen by the coincidence of the lines on the graph. However clear changes in the values of absolute abundance can be seen. More specifically the changes in absolute abundance of gut microbiome between winter bees and forging bees is extremely evident. The two clear crests observed on the figure are between time points 13th January 2016 (160113 on the graph) to 19th April 2016 (160419) and 16th November 2016 (161116) to 15th February 2017 (170215). Both time periods match up exactly with the time period of winter bees. Interestingly, al the phylotypes observed the same pattern of seasonal dip and rise albeit in different magnitudes. Justifying the motivation to study the gut microbiota of winter bees and what enables them to survive the winter period.

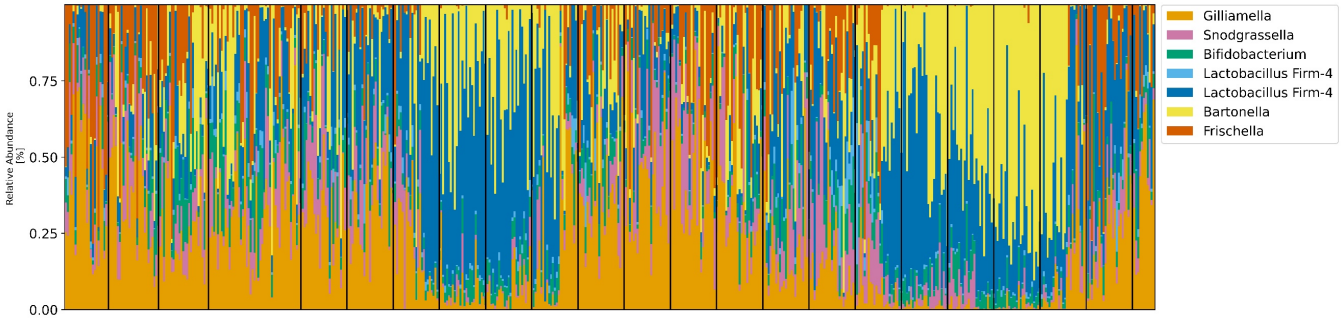

Figure 1B

After observing the absolute abundance of the phylotypes, it was necessary to study the relative abundance of the 7 phylotypes to each other. The relative abundance analysis can be observed in the Figure 1B. Here we can see that relative abundance of the microbiomes is uniform throughout the year apart from *Gilliamella* comprising a little less than the majority in the foraging months. But, come wintertime the relative abundance curve completely changes. There can be observed a clear increase in the share made up by the *Lactobacillus* Firm-5 and *Bartonella* as shown by the blue and yellow sections on the graph parallel to the winter months.

The principal component analysis seen in Figure 2 further revealed a clear separation among the winter and forager bees (Fig. 2) along principal component 1 (PC1). The Separation was driven mainly by 4 phylotypes *Lactobacillus* Firm-4*, Lactobacillus* Firm-5*, Bartonella,* and *Bifidobacterium* as can be seen by the eigenvectors in the figure above. Also, corroboratively, these 4 phylotypes are the same phylotypes that observed the maximum changes in absolute abundance between seasons represented the by crests and troughs created by these four phylotypes.

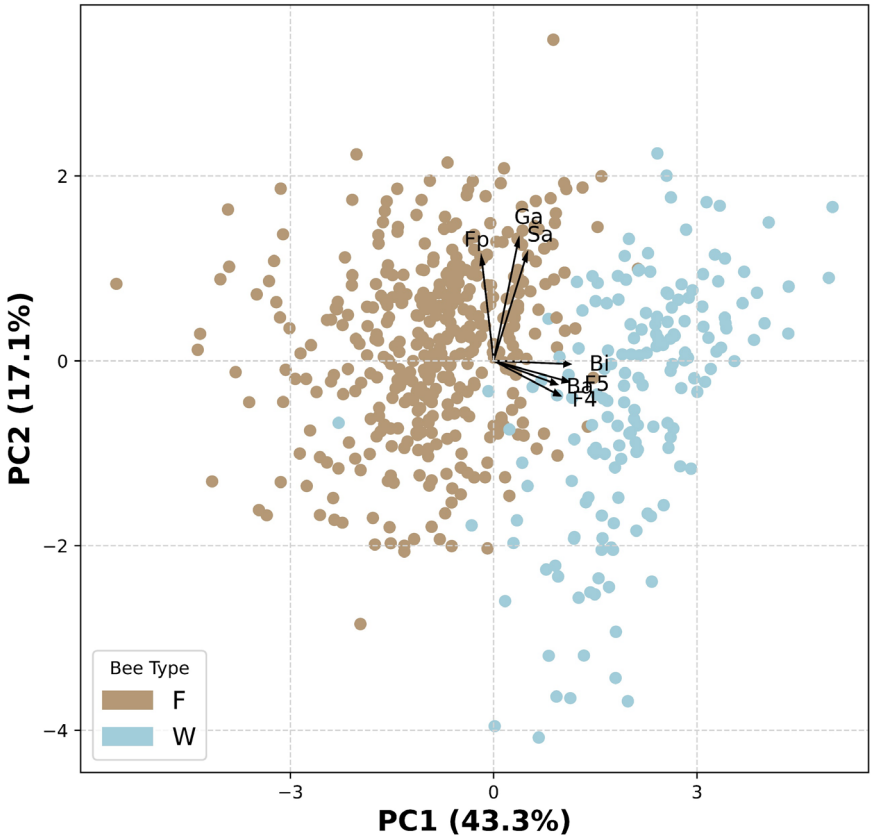

Figure 2

From the above three analysis it can be concluded that the two usually overlooked aspects when studying honeybee gut microbiome which are the absolute bacterial content and winter bees are as equally important as bacterial composition and foraging bees. The absolute bacterial content and PCA show a new a trend amongst bees of different kinds (foraging and winter) which had not been unearthed before.

References

Kešnerová, L., Trolio, M., Liberti, J., Erkosar, B., & Engel, P. (2019). Gut microbiota structure differs between honeybees in winter and summer. *The ISME Journal*, 801-814.