Honeybee Genetics Shape the Strain-Level Structure of Gut Microbiota in Social Transmission

Jiaqiang Wu  
China Agricultural University

Haoyu Lang  
China Agricultural University

Xiaohuan Mu  
China Agricultural University

Zijing Zhang  
China Agricultural University

Qinzhi Su  
China Agricultural University

Xiaosong Hu  
China Agricultural University

Hao Zheng (✉ hao.zheng@cau.edu.cn)  
China Agricultural University  https://orcid.org/0000-0003-3313-1675

Research

Keywords: Apis mellifera, gut microbiota, GWAS, Bifidobacterium, host specificity, type IV pili, polysaccharide utilization loci

Posted Date: December 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-129095/v1

License: ☇️  This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

**Background**: Honeybee gut microbiota transmitted via social interactions are beneficial to the host health. Although the microbial community is relatively stable, individual variations and high strain-level diversity have been detected across honeybees. Although the bee gut microbiota structure is influenced by environmental factors, the heritability of the gut members and the contribution of the host genetics remains elusive. Considering bees within a colony are not readily genetically identical due to the polyandry of queen, we hypothesize that the microbiota structure can be shaped by host genetics.

**Results**: We used shotgun metagenomics to simultaneously profile the microbiota and host genotypes of individuals from hives of four different subspecies. Gut composition is more distant between genetically different bees at both phylotype- and “sequence-discrete population”-level. We then performed a successive passaging experiment within colonies of hybrid bees generated by artificial insemination, which revealed that the microbial composition dramatically shifts across batches of bees during the social transmission. Specifically, different strains from the phylotype of Snodgrassella alvi are preferentially selected by genetically varied hosts, and strains from different hosts show a remarkably biased distribution of single-nucleotide polymorphism in the Type IV pili loci. A genome-wide association analysis identified that the relative abundance of a cluster of Bifidobacterium strains is associated with the host glutamate receptor gene that is specifically expressed in the bee brain. Finally, mono-colonization of Bifidobacterium with a specific polysaccharide utilization locus impacts the expression and alternative splicing of the gluR-B gene, which is associated with an altered circulating metabolomic profile.

**Conclusions**: Our results indicated that host genetics influence the bee gut composition, and suggest a gut-brain connection implicated in the gut bacterial strain preference. Honeybees have been used extensively as a model organism for social behaviors, genetics, and gut microbiome. Further identification of host genetic function as shaping force of microbial structure will advance our understanding of the host-microbe interactions.

**Full Text**

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the latest manuscript can be downloaded and accessed as a PDF.

**Figures**
Figure 1

Gut microbiota compositions differ across genetically varied honeybees at both phylotype- and SDP-level.

a, Neighbor-joining tree constructed from allele-sharing distances between honeybee subspecies. Nodes with 100% support are marked with dots. The scale bar represents raw genetic distance per variable site. b, ADMIXTURE analysis showing clustering of bee samples into four groups (K=4). One colony each of OH (n=29) and AF bees (n=8), and two colonies each of YF (n=10) and SK (n=10) bees were sampled for the metagenomic analysis. Each bar represents one bee individual. c, Relative abundance of phylotypes in the guts of bee individuals from different subspecies. (d–g) Relative abundance of sequence-discrete populations (SDPs) for four core gut members: Gilliamella (d), Bifidobacterium (e), Lactobacillus Firm-5 (f), and Bartonella (g). Dissimilarity of phylotype- (h–k) and SDP-level (l, m) gut composition profiles between paired samples as measured by Bray-Curtis distance. The distances are separated by the category of comparison: comparison between bees from different subspecies (SP), between bees from different hives of the same subspecies (HV), between individuals from the same colony (ID) (NS, not significant; *, p<0.05; **, p<0.01, ***, p<0.001; Wilcoxon test).
Gut microbiota compositions differ across genetically varied honeybees at both phylotype- and SDP-level. 

a. Neighbor-joining tree constructed from allele-sharing distances between honeybee subspecies. Nodes with 100% support are marked with dots. The scale bar represents raw genetic distance per variable site.
b. ADMIXTURE analysis showing clustering of bee samples into four groups (K=4). One colony each of OH (n=29) and AF bees (n=8), and two colonies each of YF (n=10) and SK (n=10) bees were sampled for the metagenomic analysis. Each bar represents one bee individual. 
c. Relative abundance of phylotypes in the guts of bee individuals from different subspecies. 

(d–g) Relative abundance of sequence-discrete populations (SDPs) for four core gut members: Gilliamella (d), Bifidobacterium (e), Lactobacillus Firm-5 (f), and Bartonella (g). Dissimilarity of phylotype- (h–k) and SDP-level (l, m) gut composition profiles between paired samples as measured by Bray-Curtis distance. The distances are separated by the category of comparison: comparison between bees from different subspecies (SP), between bees from different hives of the same subspecies (HV), between individuals from the same colony (ID) (NS, not significant; *, p<0.05; **, p<0.01, ***, p<0.001; Wilcoxon test).
Figure 2

Gut community shifts during the successive passage within the colony. (a) Study design for serial transmission of gut microbiota in two lines of hybrid bees generated by artificially insemination. We sampled guts from three batches of hybrid bees (B1–B3) during the passage and from the founding workers that initiate the colony (O-A’, O-Y’). (b) Bray-Curtis distance of the gut communities between bees from different batches at the OTU-level of 16S rRNA sequences. (c) Bray-Curtis distance of the gut communities at the SDP-level by metagenomic sequencing. (d) Relative abundance profiles of SDPs from Bifidobacterium and Lactobacillus Firm-5 for the founding workers (O-A’, O-Y’) and the B3 batch of individuals (O-A, O-Y). (e–l) Comparison of the relative abundances of the SDPs of Bifidobacterium Bifido-1.2 (e, g), Bifido-1.4 (f, h), Lactobacillus Firm5-1 (i, k), and Firm 5-4 (j, l) between the founding workers and the B3 individuals. Wilcoxon test was used to compare the average of Bray-Curtis distance or the relative abundance of each microbiota taxon between each two bee groups, and P value <0.05 indicating statistical significance. (NS, not significant; **, p < 0.01; ***, p < 0.001; Wilcoxon test).
Figure 2

Gut community shifts during the successive passage within the colony. (a) Study design for serial transmission of gut microbiota in two lines of hybrid bees generated by artificially insemination. We sampled guts from three batches of hybrid bees (B1–B3) during the passage and from the founding workers that initiate the colony (O-A’, O-Y’). (b) Bray-Curtis distance of the gut communities between bees from different batches at the OTU-level of 16S rRNA sequences. (c) Bray-Curtis distance of the gut communities at the SDP-level by metagenomic sequencing. (d) Relative abundance profiles of SDPs from Bifidobacterium and Lactobacillus Firm-5 for the founding workers (O-A’, O-Y’) and the B3 batch of individuals (O-A, O-Y’). (e–l) Comparison of the relative abundances of the SDPs of Bifidobacterium Bifo-1.2 (e, g), Bifo-1.4 (f, h), Lactobacillus Firm5-1 (i, k), and Firm 5-4 (j, l) between the founding workers and the B3 individuals. Wilcoxon test was used to compare the average of Bray-Curtis distance or the relative abundance of each microbiota taxon between each two bee groups, and P value <0.05 indicating statistical significance. (NS, not significant; **, p < 0.01; ***, p < 0.001; Wilcoxon test).
SNPs of Type IV pili component genes of Snodgrassella alvi are differentially distributed between the founding workers and the B3 bees in the colony. (a) A heatmap showing the minor allele frequency for missense SNPs that are significantly different in founding workers (O-A') and the B3 batch of bees (O-A). Each row represents one bee metagenomic sample, each column is one site in the T4P genes. The tree on the right illustrates a dendrogram of clustering (Ward’s method). (b) Whole-genome phylogenetic tree of isolated Snodgrassella strains (Additional file 1: Dataset S1) using the maximum-likelihood algorithm based on the concatenation of core protein sequences. The lines aligned to tree leaves represent corresponding gene sequences with missense (red dot) and silent (grey dot) SNPs indicated.
**Figure 3**

SNPs of Type IV pili component genes of *Snodgrassella alvi* are differentially distributed between the founding workers and the B3 bees in the colony. (a) A heatmap showing the minor allele frequency for missense SNPs that are significantly different in founding workers (O-A') and the B3 batch of bees (O-A). Each row represents one bee metagenomic sample, each column is one site in the T4P genes. The tree on the right illustrates a dendrogram of clustering (Ward’s method). (b) Whole-genome phylogenetic tree of isolated *Snodgrassella* strains (Additional file 1: Dataset S1) using the maximum-likelihood algorithm based on the concatenation of core protein sequences. The lines aligned to tree leaves represent corresponding gene sequences with missense (red dot) and silent (grey dot) SNPs indicated.
Relative abundances of the heritable Bifidobacterium SDP Bifido-1.4 are associated with genetic variants in the genomic locus containing the metatropic glutamate receptor gene. (a) Genome-wide Manhattan plot: each dot represents the $-\log$ of the p value for the association of the SNP with the relative abundance of the Bifido-1.4 SDP. The red box highlights the associated locus on chromosome 13 containing the gene gluR-B. The threshold for genome-wide significance was corrected for multiple testing with a weighted Bonferroni adjustment. The threshold value was set at $-\log(p) > 2e-8$. We used both General Linear Model (GLM) and Mixed Linear Model (MLM) for the association tests (b) Quantile-quantile plot showing deviation from the expected distribution of p values. The 1204diagonal (red) line represents the expected distribution. (c) Close-up plots of ~0.4 Mb window around the SNPs with the highest associations. The coloring of each circle is proportional to the significance. The exon intron architecture of the gluR-B gene and two neighboring genes (LOC408435, LOC422843) are shown at the bottom. (d) Relative abundance of the Bifido-1.4 SDP in bees with different genotypes at the gluR-B-associated SNP (chr13-7527228; p < 0.001, Wilcoxon test).
Relative abundances of the heritable Bifidobacterium SDP Bifido-1.4 are associated with genetic variants in the genomic locus containing the metatropic glutamate receptor gene. (a) Genome-wide Manhattan plot: each dot represents the $-\log$ of the p value for the association of the SNP with the relative abundance of the Bifido-1.4 SDP. The red box highlights the associated locus on chromosome 13 containing the gene gluR-B. The threshold for genome-wide significance was corrected for multiple testing with a weighted Bonferroni adjustment. The threshold value was set at $-\log(p) > 2e-8$. We used both General Linear Model (GLM) and Mixed Linear Model (MLM) for the association tests (b) Quantile-quantile plot showing deviation from the expected distribution of p values. The 1204diagonal (red) line represents the expected distribution. (c) Close-up plots of ~0.4 Mb window around the SNPs with the highest associations. The coloring of each circle is proportional to the significance. The exon intron architecture of the gluR-B gene and two neighboring genes (LOC408435, LOC422843) are shown at the bottom. (d) Relative abundance of the Bifido-1.4 SDP in bees with different genotypes at the gluR-B-associated SNP (chr13-7527228; p < 0.001, Wilcoxon test).
Figure 5

Bifidobacterium impacts gene expression and GABA level in the brain, and affects the hemolymph metabolome of honeybees. (a) Targeted metabolomics indicates that the GABA concentrations are increased in the brain of Bifido-1.4 mono-inoculated honeybees than the microbiota-free (MF) bees. (b) Relative expression levels of gluR-B gene in brains of MF and Bifido-1.4 inoculated bees. (c) Differential splicing events of gluR-B gene in the brains of MF and Bifido-1.4 inoculated bees. Benjamini-Hochberg corrected p values are shown. (d) Heatmaps of differentially abundant metabolites identified by a quasi-targeted metabolomics of hemolymph. Colors indicate the Z-score based on the normalized relative concentration of each metabolite. The tree on the left illustrates a dendrogram of clustering (Ward’s method). (e) Partial least squares discriminant analysis based on 328 metabolites (Additional file 3: Dataset S3) detected from the hemolymph of MF (n = 6) and Bifido-1.4 inoculated bees (n = 6) showing different clustering groups (95% confidence regions). (f) Syntenic loci of the PUL in B. asteroides strains
from different SDPs. Homologous genes are connected by gray bars. (g) Boxplots of the read counts for GH43-24 specific to Bifido-1.4 in each genotype at the gluR-B-associated SNP (chr13-7527228).

**Figure 5**

Bifidobacterium impacts gene expression and GABA level in the brain, and affects the hemolymph metabolome of honeybees. (a) Targeted metabolomics indicates that the GABA concentrations are increased in the brain of Bifido-1.4 mono-inoculated honeybees than the microbiota-free (MF) bees. (b) Relative expression levels of gluR-B gene in brains of MF and Bifido-1.4 inoculated bees. (c) Differential splicing events of gluR-B gene in the brains of MF and Bifido-1.4 inoculated bees. Benjamini-Hochberg corrected p values are shown. (d) Heatmaps of differentially abundant metabolites identified by a quasi-targeted metabolomics of hemolymph. Colors indicate the Z-score based on the normalized relative concentration of each metabolite. The tree on the left illustrates a dendrogram of clustering (Ward’s method). (e) Partial least squares discriminant analysis based on 328 metabolites (Additional file 3: Dataset S3) detected from the hemolymph of MF (n = 6) and Bifido-1.4 inoculated bees (n = 6) showing
different clustering groups (95% confidence regions). (f) Syntenic loci of the PUL in B. asteroides strains from different SDPs. Homologous genes are connected by gray bars. (g) Boxplots of the read counts for GH43-24 specific to Bifo-1.4 in each genotype at the gluR-B-associated SNP (chr13-7527228).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DatasetS1Genomes.xlsx
- DatasetS2SNP.xlsx
- DatasetS3Transcriptome.xlsx
- DatasetS4Metabolome.xlsx