Functional Diversity of *Gilliamella* Strains in Honey Bees

Chanyanuch Chaisuwan [1], Sahassawat Chailungka[2], Rujipas Yongsawas[3], Senoch Seanpong[4], Terd Dissayathanoowat[1,3] and Nuttapol Noirungsee[1,3] *

[1] Department of Biology, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand.
[2] Master of Science in Applied Microbiology (International Program), Faculty of Science, Chiang Mai University, Chiang Mai, Thailand.
[3] Research Center of Deep Technology in Beekeeping and Bee Products for Sustainable Development Goals (SMART BEE SDGs), Chiang Mai University, Chiang Mai, Thailand.
[4] Institute of Plant Sciences and Microbiology, Department of Biology, University of Hamburg, Hamburg, Germany.
* Corresponding author: nuttapol.n@cmu.ac.th

• BACKGROUND

Gilliamella, a core genus in the honey bee (*Apis* spp.) gut microbiome, plays an essential role in bee health by aiding in food digestion, promoting growth, and defending against pathogens. It has been shown to degrade pectin (a key component of pollen cell walls), detoxify harmful sugars, and protect bees from pathogens. Different strains of *Gilliamella* exhibit various metabolic functions, including the production of thiopeptides, which inhibit the growth of the bee pathogen *Melissococcus plutonius* [1]. This metabolic diversity suggests that *Gilliamella* strains confer distinct benefits to their bee hosts depending on their specific capabilities.

Although *Gilliamella* strains show high similarity in their 16S rRNA gene sequences (>99%), they display significant genetic variation across other genomic regions, with up to 40% variation in gene content. Studies using average nucleotide identity (ANI) of whole genome sequences have classified *Gilliamella* into 15 distinct species, highlighting its greater diversity beyond what is indicated by 16S rRNA alone [2, 3].

While *Gilliamella* shows some degree of host specificity, research on the functional roles of these strains is still limited. Understanding these functions could provide valuable insights into both the role of *Gilliamella* and the broader ecology of the honey bee hosts. This study seeks to explore whether *Gilliamella* groups are host-specific and to identify the genes and pathways associated with these host-specific strains. In addition, it aims to classify *Gilliamella* into functional groups, offering a deeper understanding of its role in the bee gut microbiome.

• AIMS

- To determine if groups of ***Gilliamella*** are specific to their honey bee hosts and to identify the **genes** and **pathways** with that are associated with **host specificity.**
- To classify ***Gilliamella*** strains into **functional groups** based on their genomic and metabolic capabilities.

• APPROACHES

153 *Gilliamella* genomes from *A. cerana* and *A. mellifera* from GenBank

BUSCO
97% cut-off completeness of Gammaproteobacteria markers

FastANI
Pairwise comparison of Average Nucleotide Identity (ANI)

bactaxR
Bacterial taxonomy construction and evaluation

Prokka
Genome annotation

Pangenome analysis with Panaroo

Scoary
microbial pan-GWAS
Calculation of the associations between all genes of *Gilliamella* genomes and the hosts

MinPath
Parsimony approach for biological pathway reconstructions

hclust
Hierachical cluster analysis

SIMPER
Identification of pathways most responsible for the differences between functional groups

anti SMASH
Identification of biosynthetic loci of secondary metabolites

BiGSCAPE
Construction of sequence similarity networks of biosynthetic gene clusters and classification of gene cluster families

• RESULTS

• GWAS Identifies Host-Associated Genes

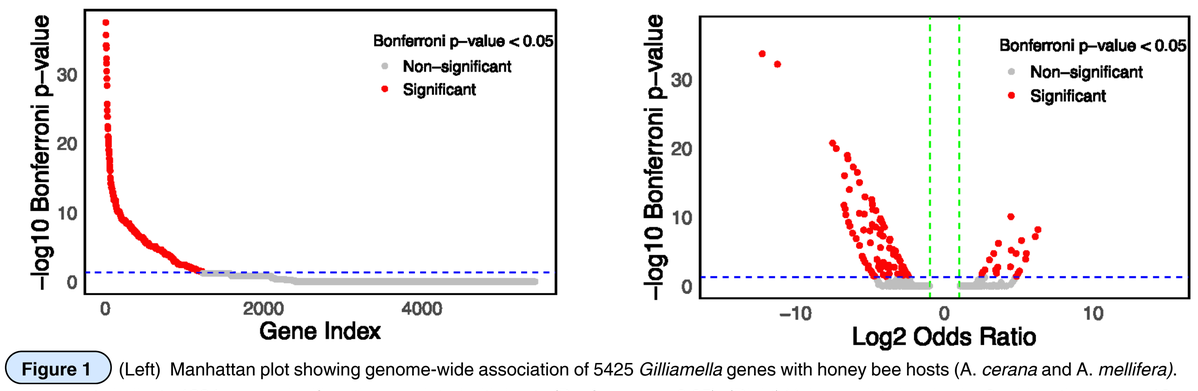


Figure 1 (Left) Manhattan plot showing genome-wide association of 5425 *Gilliamella* genes with honey bee hosts (A. *cerana* and A. *mellifera).* Red points indicate 1224 genes significantly associated with hosts (Bonferroni p < 0.05). (Right) Volcano plot showing log2 odds ratios vs. -log10 Bonferroni p-values. Positive log2 odds ratios indicate association with A. *cerana* (346 genes). Negative log2 odds ratios indicate association with A. *mellifera* (878 genes). A total of 4659 genes outside the scale range were excluded.

Host	Gene	Annotation
A. cerana	*cat*	Catalase
	levD	PTS system fructose-specific EIIA component
	yedA	putative inner membrane transporter YedA
	ytbE_2, yvgN	putative oxidoreductase YtbE, Glyoxal reductase
	vanW	Vancomycin B-type resistance protein VanW
	gntK	Thermoresistant gluconokinase
	yhbE	putative inner membrane transporter YhbE
	ytbE, ytbE_1	putative oxidoreductase YtbE
	gltC	HTH-type transcriptional regulator GltC
	uvrA	UvrABC system protein A
A. mellifera	*emrB*	Multidrug export protein EmrB
	emrA, aaeA	Multidrug export protein EmrA, p-hydroxybenzoic acid efflux pump subunit AaeA
	group_2770	putative HTH-type transcriptional regulator
	rhaR, rhaS,	HTH-type transcriptional activator RhaR, HTH-type
	pchR	transcriptional activator RhaS, Regulatory protein PchR
	resA	Thiol-disulfide oxidoreductase ResA
	manX	PTS system mannose-specific EIIAB component
	esiB	Secretory immunoglobulin A-binding protein EsiB
	group_174	Prolyl endopeptidase
	fyuA	Pesticin receptor

Table 1 Top 10 genes associated with A. *cerana* and A. *mellifera* host

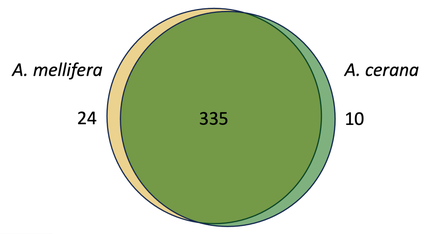


Figure 2 Venn diagram showing reconstructed pathways of *Gilliamella* from A. *cerana* and A. *mellifera* host by MinPath, mapped to the MetaCyc Metabolic Pathway database.

Host	Pathways
A. cerana	L-homocysteine biosynthesis
	L-methionine biosynthesis III
	Allantoin degradation to glyoxylate I
	Arsenate detoxification II glutaredoxin
	Arsenate detoxification III thioredoxin
A. mellifera	2, 3-dihydroxybenzoate biosynthesis
	D-erythronate degradation I
	D-erythronate degradation II
	D-threonate degradation
	L-arabinan degradation

Table 2 Top 5 Pathways associated with A. *cerana* and A. *mellifera* host

• PATHWAYS PROFILES AND BGCs

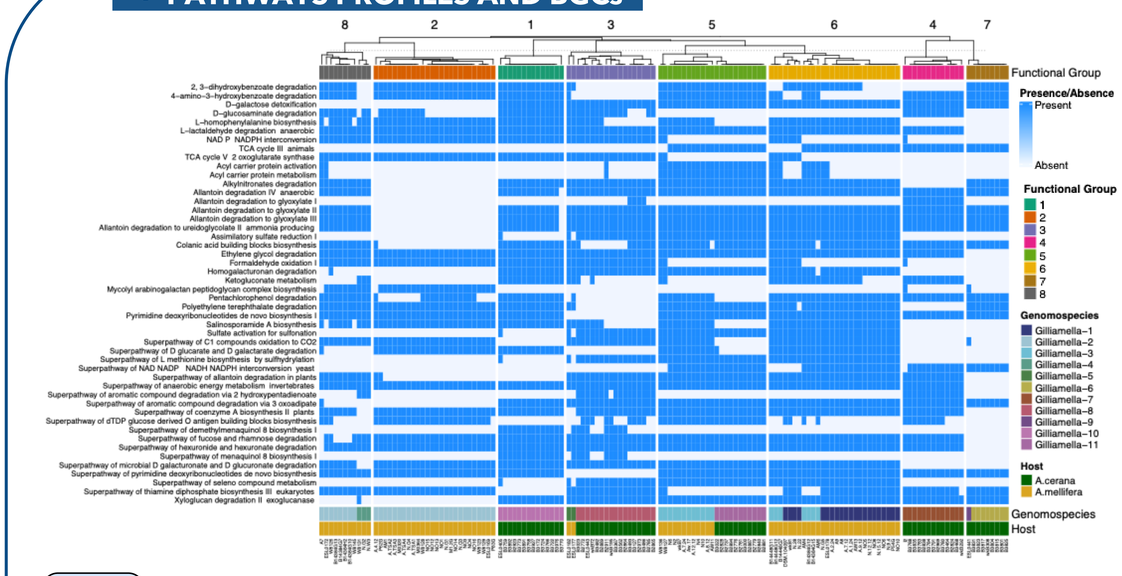


Figure 3 Heatmap displaying hierarchical clustering of *Gilliamella* strains into 8 functional groups, based on all (379) pathways using Ward's hierarchical clustering (Ward.D2) and a binary dissimilarity matrix. The heatmap shows the top 48 pathways contributing most to dissimilarity. Genomospecies classifications at the 95% ANI threshold are color-coded, and host species (A. *mellifera* and A. *cerana*) are indicated for each strain.

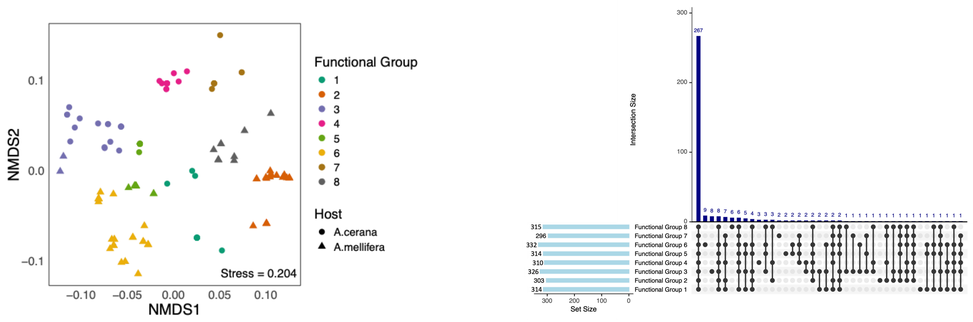
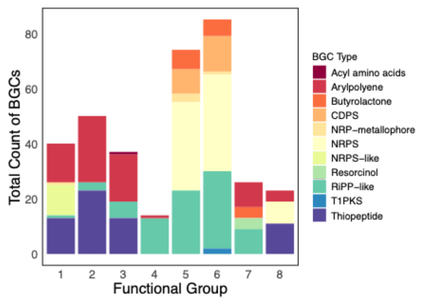


Figure 4 (Left) Non-metric Multidimensional Scaling (NMDS) plot illustrating the dissimilarity in pathways between strains of *Gilliamella*, colored by functional groups and shaped by host species (A. *mellifera* and A. *cerana*). The analysis was performed using Jaccard dissimilarity. Stress = 0.204 indicates a moderate fit. The clustering of pathways based on functional group was tested using ANOSIM (Jaccard dissimilarity), yielding an R statistic of 0.964 and a significance level of 0.001 (999 permutations). (Right) UpSet plot showing the presence and overlap of pathways across eight functional groups of *Gilliamella*, based on hierarchical clustering of pathways. Each bar represents the intersection of pathways shared among the functional groups, while the set sizes indicate the total number of pathways present in each functional group.



Network Overview	BiG-SCAPE Classes			
	NRPS	PKSother	RiPPs	Others
Number of families	15	1	6	13
Average number of BGCs per family	8	5	24	8
Max number of BGCs in a family	20	5	37	29
Families with MIBiG Reference BGCs	0	0	0	0

Figure 5 (Left) Stacked bar plot showing the distribution of Biosynthetic Gene Clusters (BGCs) across *Gilliamella* functional groups, highlighting the metabolic diversity within each group. (Right) Overview of Biosynthetic Gene Cluster (BGC) families classified by BiG-SCAPE. The lack of matches to known MIBiG references highlights the potential novelty and unexplored biosynthetic potential within *Gilliamella,* making these BGCs interesting candidates for further functional characterization and experimental validation. CDPS refers to tRNA-dependent cyclodipeptide synthases, NRP-metallophore denotes non-ribosomal peptide metallophores, NRPS stands for non-ribosomal peptide synthetase, RiPP-like represents unspecified ribosomally synthesised and post-translationally modified peptide products, and T1PKS refers to Type I Polyketide synthase.

Functional Group	Host	Function	Notable Traits
1	*A. cerana*	Pathogen defense (thiopeptide) and vitamin K producers	Lacks vitamin B1 biosynthesis pathway
2	*A. mellifera*	Pathogen defense (thiopeptide)	Lacks allantoin (purine degradation product) degradation pathways
3	Both	Pathogen defense (thiopeptide) and versatile carbohydrate degraders	Lacks pyrimidine deoxyribonucleotide biosynthesis and pentachlorophenol (pesticide) degradation pathway
4	*A. cerana*	Specialized hemicellulose and allantoin degraders	Lacks galacturonate, glucoronate, rhamnose (pectin constituent), and fucose degradation pathway
5	Both	Pathogen defense (RiPP-bacteriocin)	Lacks vitamin B1 biosynthesis pathways; members from *A. cerana* lacks pentachlorophenol (pesticide) degradation pathways
6	*A. mellifera*	Pathogen defense (RiPP-bacteriocin)	Possess Type I polyketide synthase BCG
7	*A. cerana*	Specialized hemicellulose and allantoin degraders	Lacks galacturonate, glucoronate, rhamnose (pectin constituents), and fucose degradation pathway
8	*A. mellifera*	Pathogen defense (thiopeptide)	Members of this group has been shown to inhibit *M. plutonius* via thiopeptide [1]

Table 3 *Gilliamella* functional groups classified by key functional traits. Each functional group is defined by its primary role in pathogen defense or specialized metabolism

• CONCLUSION

This study identified key genes and pathways associated with host specificity in *Gilliamella*, with a predominant focus on carbohydrate and amino acid metabolism. The pathway profiles revealed interesting functional traits, such as Functional Group 1's ability to produce vitamin K while lacking the vitamin B1 biosynthesis pathway. Additionally, functional groups associated with A. *cerana* were found to lack the pathway for pentachlorophenol degradation, possibly reflecting the population's exposure history to this pesticide. These findings provide new insights into the metabolic capabilities and ecological roles of *Gilliamella*, paving the way for further studies on its role in the bee host. Interesting future research could explore *Gilliamella*'s involvement in environmental detoxification, such as pentachlorophenol degradation, and its contribution to pathogen defense through RiPP-dependent bacteriocin production.

• ACKNOWLEDGEMENTS

- Lecturer Dr. Nuttapol Noirungsee
- Department of Microbiology, Chiang Mai University and
- Research Center of Deep Technology in Beekeeping and Bee Products for Sustainable Development Goals (SMART BEE SDGs), Chiang Mai University, Chiang Mai, Thailand.

• REFERENCES

1. Lang H, Liu Y, Duan H, Zhang W, Hu X, Zheng H. Identification of peptides from honeybee gut symbionts as potential antimicrobial agents against *Melissococcus plutonius.* Nature Communications. 2023;14(1):7650.
2. Zhang Z, Guo Y, Yang F, Li J. Pan-genome analysis reveals functional divergences in gut-restricted *Gilliamella* and *Snodgrassella.* Bioengineering. 2022;9(10):544.
3. Yang C, Hu J, Su Q, Zhang Z, Du Y, Wang J, et al. A review on recent taxonomic updates of gut bacteria associated with social bees, with a curated genomic reference database. Insect Science. 2023.