EXPLORING SPECIES DIVERSITY AND FUNCTIONAL GENES IN NATURAL HONEY FROM SON LA, VIETNAM USING METAGENOMIC APPROACH

Nguyen Hoang Nam[1*], Pham Minh Duc[1], Kim Thi Phuong Oanh[2], Tran Thi Thuy[1]
[1]Hanoi National University of Education, [2]Institute of Biology - Vietnam Academy of Science and Technology

ARTICLE INFO

Received: 10/01/2025

Revised: 31/3/2025

Published: 31/3/2025

KEYWORDS

Honey
Metagenomics
Species composition
Gene function
Vietnam

ABSTRACT

Honey is a widely consumed food and a habitat for diverse microorganisms. In this paper, we employed the metagenomics approach to investigate the species composition and genetic make-up of natural honey from Moc Chau, Yen Chau and Quynh Nhai districts, Son La province in Vietnam, where is favorable for honey production. The results revealed that honey contains DNA from various microorganisms, notably the bacterium *Apilactobacillus kunkeei* and the yeast *Zygosaccharomyces rouxii*. Observable differences in microbial composition among 06 honey samples have been noted, suggesting that metagenomics could be applied to trace honey's geographic and biological origins. The gene function profiles of microorganisms in honey indicate diverse metabolic activities and biological processes, with similar functional gene compositions across samples. The functional group "Defense Mechanisms" accounts for an average of 2.94±0.13%, reflecting microbial adaptations to the harsh environment of honey through efflux pump systems. Genes associated with "Secondary compound metabolism" average 1.52±0.43%, demonstrating that honey is a source of bioactive compounds. This study provides a honey metagenomic data resource for exploring novel microorganisms and genes in Vietnamese honey.

KHÁM PHÁ ĐA DẠNG LOÀI VÀ CHỨC NĂNG GENE TRONG MẬT ONG RỪNG TẠI SƠN LA, VIỆT NAM BẰNG METAGENOMICS

Nguyễn Hoàng Nam[1*], Phạm Minh Đức[1], Kim Thị Phương Oanh[2], Trần Thị Thúy[1]
[1]Trường Đại học Sư phạm Hà Nội, [2]Viện Sinh học - Viện Hàn lâm Khoa học và Công nghệ Việt Nam

THÔNG TIN BÀI BÁO

Ngày nhận bài: 10/01/2025

Ngày hoàn thiện: 31/3/2025

Ngày đăng: 31/3/2025

TỪ KHÓA

Mật ong
Metagenomics
Thành phần loài
Chức năng gene
Việt Nam

TÓM TẮT

Mật ong là thực phẩm phổ biến của con người và môi trường sống của nhiều vi sinh vật. Tỉnh Sơn La, Việt Nam có điều kiện tự nhiên thuận lợi cho sự phát triển của ong mật. Trong bài báo này, chúng tôi đã sử dụng phương pháp metagenomics để khảo sát thành phần sinh vật và các gene trong mật ong tự nhiên thu tại Mộc Châu, Yên Châu và Quỳnh Nhai, tỉnh Sơn La. Kết quả cho thấy mật ong chứa DNA của nhiều loài sinh vật, nổi bật là vi khuẩn *Apilactobacillus kunkeei* và nấm *Zygosaccharomyces rouxii*. Sự khác biệt về thành phần sinh vật giữa các mẫu mật ong cho thấy tiềm năng của metagenomics trong việc truy xuất nguồn gốc mật ong. Hồ sơ chức năng gene tiết lộ sự đa dạng về trao đổi chất và quá trình sinh học, với thành phần chức năng gene tương đồng giữa các mẫu. Nhóm "Cơ chế phòng vệ" chiếm 2,94 ± 0,13%, phản ánh khả năng thích nghi của vi sinh vật với môi trường mật ong thông qua các protein bơm đào thải. Các gene liên quan đến chuyển hóa hợp chất thứ cấp chiếm 1,52 ± 0,43%, chứng tỏ rằng mật ong là nguồn các hợp chất có hoạt tính sinh học. Nghiên cứu này cung cấp nguồn dữ liệu metagnome nhằm khai thác thông tin vi sinh vật và gene mới từ mật ong tại Việt Nam.

DOI: https://doi.org/10.34238/tnu-jst.11854

[*] Corresponding author. *Email: nam7.hnue@gmail.com*

1. Introduction

Honey is renowned for its uniquely sweet taste and complex chemical composition, contributing to its nutritional and therapeutic properties [1]. Given its probiotic potential and natural antibacterial properties, honey is a valuable resource in medical and wellness applications [2], [3]. The microbiota found in honey consists of a variety of bacteria, fungi, and other microorganisms that originate from multiple sources, including the bee's digestive tract, pollen, and the surrounding environment [4]-[7]. Despite honey's extreme conditions, including high osmotic pressure, low pH, and antibacterial properties, these microorganisms have uniquely adapted and thrived. Researching the biodiversity of honey is crucial as it offers insights into ecosystem health, particularly the well-being of honeybee populations [8]. By analyzing the pollen content of honey, researchers can identify key plants of bee foraging, which can inform conservation strategies. Furthermore, studying the microorganisms in honey can enhance food safety by identifying both beneficial and potentially harmful species [9].

Son La, a mountainous province in northern Vietnam, is known for its favorable conditions for honey production [10]. The rich biodiversity of Son La's plants contributes to its honey's distinctiveness and high quality. Metagenomics has become a key tool to gain a deeper understanding of the microbial composition in Son La's honey. Metagenomic methods allow researchers to analyze the genetic material of all microorganisms in a honey sample, offering a comprehensive picture of the microbial biodiversity without the need to culture individual species [7]. Via this approach, researchers can identify microbial species and their functional roles, shedding light on their contributions to honey's flavor, preservation, and health-promoting properties [11].

While there has been some research on the chemical composition of honey in Vietnam [12], [13], studies focusing on the biodiversity of microorganisms in honey remain limited. To fully understand the biological characteristics of honey produced in Vietnam, in this study, we used a metagenomics approach to explore the species composition and genetic make-up of six natural honey samples collected in Son La province, Vietnam. This knowledge will enhance the quality control of honey products and open new avenues for exploring their potential across industries, from food to medicine.

2. Materials and methods

All chemicals used for DNA extraction and purification are analytical grade and purchased from Merck and/or Sigma Aldrich.

2.1. Sampling and processing

Table 1. *The location of the beehives in this study*

Sample	District	Coordinates
Q1	Quynh Nhai	21°45'57.5"N 103°41'21.3"E
Q2		21°45'57.3"N 103°41'21.7"E
Y1	Yen Chau	20°59'17.0"N 104°14'25.0"E
Y2		20°59'15.0"N 104°14'26.0"E
C1	Moc Chau	20°48'35.0"N 104°42'27.0"E
C2		20°48'32.0"N 104°42'10.0"E

In this study, six natural honey samples were collected from various locations (Table 1) within Son La province, Vietnam, from May to June 2022. The honey samples were collected directly from natural hives in the wild.

Samples were preserved at cool temperatures and transported to the laboratory to prepare for DNA extraction. From each sample, 62.5 mL of honey was divided into five 50 mL Falcon tubes, and filled with milli-Q water to reach 50 mL. The tubes were heated to 65 °C for 15 minutes, cooled, and centrifuged at 4,620 × g for 30 minutes. The supernatant was discarded, and the pellet was resuspended in 5 mL of milli-Q water. After combining all the suspensions into a single tube and diluting with milli-Q water to a total volume of 50 mL, the pellet was resuspended in 1.5 mL of milli-Q water and centrifuged at 4,620 × g for 15 minutes. The supernatant was discarded again, and the remaining suspension was transferred to a fresh 2 mL Eppendorf tube for centrifugation at 16,000 × g for 2 minutes to obtain a pellet. The DNA extraction followed a protocol from a previous study [5]. Briefly, the pellet was mixed with 900 µL of TE buffer and four 2 mm glass beads, then vortexed for 1 minute. After adding 100 µL of Lysozyme solution and incubating at 37 °C for 50 minutes, the tube was centrifuged, and the supernatant was removed. Then, 40 µL of proteinase K and 960 µL of 2.5% CTAB lysis solution were added, vortexed, and incubated at 65 °C for 90 minutes. Chloroform/isoamyl alcohol (CI) was added, and the mixture was centrifuged at 7,600 × g for 10 minutes. 800 µL of the supernatant was transferred to a fresh tube, followed by the addition of NaCl and CTAB solutions. After additional centrifugation with CI, the DNA was precipitated with ethanol, washed with 70% ethanol, and treated with RNase A in TE buffer. Finally, the DNA was dissolved in 100 µL of TE buffer and stored at 4 °C.

2.2. DNA quality checking, library preparation, and sequencing

These steps were performed by NovogeneAIT Genomics Singapore Pte Ltd. DNA concentration, degradation degree, purity, and contamination were checked following their criteria. Samples with an OD value between 1.8 and 2.0 and DNA content above 1 µg were used for library preparation. Following the manufacturer's recommendations, sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit for Illumina (NEB, USA). The libraries were sequenced on an Illumina NovaSeq 6000 system (Illumina, Inc.), and paired-end (PE) reads (2 x 150 bp) were generated.

2.3. Data processing

The quality of reads was checked using FastQC (v0.11.9) [14]. Raw data were processed using Trimmomatic (v0.39) [15]. The remaining data were mapped by Bowtie2 (v2.4.5) [16] with the human reference genome GRCh38.p14 (GCF_000001405.40) downloaded from the National Center for Biotechnology Information (NCBI). Unmapped reads were used for taxonomic classification by Kraken2 (v2.1.2) [17] and metagenomic assembly by MEGAHIT (v1.2.9) [18]. Bracken (v2.7) [19] was used to estimate the relative abundance from Kraken2 reports. QUAST (v5.2.0) [20] was used to check the quality of assembled contigs. The protein-coding genes in the contig sets and their corresponding protein sequences were predicted using the Prodigal software (v2.6.3) [21] with the *-p meta* option. The functions of the protein sequences generated by Prodigal were annotated using the reCOGnizer software (v1.11.1) [22]. The reCOGnizer software utilizes the Conserved Domains Database (CDD) to annotate functions based on protein domains.

2.4. Data availability

The sequencing raw data in this study was deposited and available in the NCBI under the Bioproject ID PRJNA1023004 with the Sequence Read Archive (SRA) number from SRR26265546 to SRR26265551.

3. Results and Discussion

3.1. Data filtering

Overall, the raw sequencing data for all samples demonstrated good quality. The percentage of data discarded is not high (less than 10% after Trimmomatic processing), indicating the good quality and high purity of the DNA samples. The amount of sequencing data discarded after Bowtie2 mapping ranged from 0.01% to 0.16%. Figure 1 shows the amount of data before and after the filtering steps.

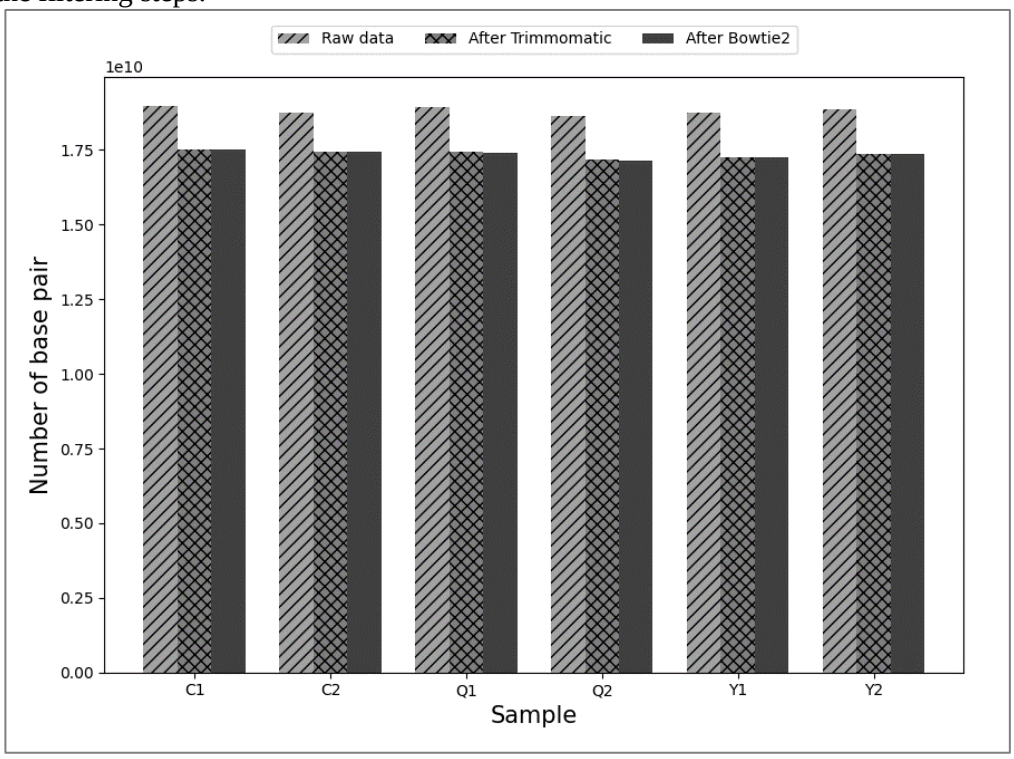

Figure 1. *Bar plots showing numbers of base pairs before and after cleaning steps*

3.2. Taxonomic profiles

Using the Bracken for six metagenome samples in honey that were collected at three different districts showed that the eukaryotic group accounted for a high proportion in samples C1, Q1, Q2, Y1, and Y2 (ranging from 58 to 93%) (Figure 2). Exclusively, in sample C2, the eukaryotic group only accounted for 21%, much lower than that of bacteria (79%). The proportions of protozoa, archaea, and viruses in this study were negligible, ranging up to 0.6%. Overall, the findings of this study align with previous research, suggesting that the DNA composition in honey samples can vary greatly and does not follow a specific pattern [6], [23]. Some studies have indicated that the majority of DNA in honey samples belongs to viruses, accounting for as much as 40% [4] to over 98% [6], followed by bacteria and eukaryotes. In contrast, specific honey samples from Greece showed a high proportion of bacteria (~53%), followed by plants (~31%), other eukaryotes (~14%), and viruses (~2%) [7]. Meanwhile, an analysis of honey samples from China showed that the DNA pool in a given sample could be dominated by bacteria, fungi, plants, animals, or viruses [23]. While current evidence in this study remains insufficient to elucidate the relationship between microbial composition and environmental factors fully, many studies have demonstrated that the distinct metagenomic profiles observed in

honey samples from different regions may help in tracing the geographic origins of honey products available on the market [23] - [25].

In the eukaryotic group, honeybee DNA accounted for the highest proportion in samples C1, Y1, and Y2, whereas fungi dominated in samples Q1 and Q2, comprising 80% – 84% of the total eukaryotic reads. Among fungi, *Zygosaccharomyces rouxii* was the most prominent species, accounting for 76% – 85% of the phylum Ascomycota. Other fungal species, such as *Zygotorulaspora mrakki* (4% – 6%) and members of the genera *Saccharomyces* and *Kluyveromyces*, appeared in smaller proportions. Within the genus *Apis*, *A. cerana* consistently represented over 90% of honeybee DNA across all six samples, while other species were trivial, contributing only 3% – 4%. The small number of reads assigned to the other bees can be caused by two reasons: (1) the DNA of these bees contaminated the honey of *A. cerana* through their contact in the wild, and (2) the similarity between the genomic sequences of these species leads to false-positive classification results [6].

The plant composition accounted for only 2% – 6% of the eukaryotic group. In this study, the phylum Streptophyta was evenly represented, with the class Magnoliopsida dominant. Orders such as Asterales, Malpighiales, Poales, and Fagales collectively accounted for over 70% of the plant reads. Other species were less abundant but demonstrated significant diversity, highlighting the ecological richness around the hives. Plant composition at lower taxonomic ranks, such as genus and species, can vary depending on geographical regions. Our findings revealed a different taxonomic distribution compared to a previous study [25], which reported *Schleichera* and *Syzygium* as predominant genera. These variations likely reflect regional differences in floral resources available to the bees, which influence pollen collection and, consequently, the botanical composition of honey.

The relative abundance of bacteria in samples C1, C2, Q1, Q2, Y1, and Y2 were 37%, 79%, 28%, 7%, 15%, and 41%, respectively, reflecting significant dissimilarity. Among bacterial phyla, Proteobacteria dominated only in samples Q1 and Q2, with proportions of 60% and 90%, respectively, followed by Firmicutes (6% – 29%), Actinobacteria (1% – 6%), Bacteroidetes (2% – 3%), and other phyla in smaller percentages. In contrast, in samples C1, C2, Y1, and Y2, the Firmicutes phylum dominated, accounting for 85% (Y2) to 99% (C1).

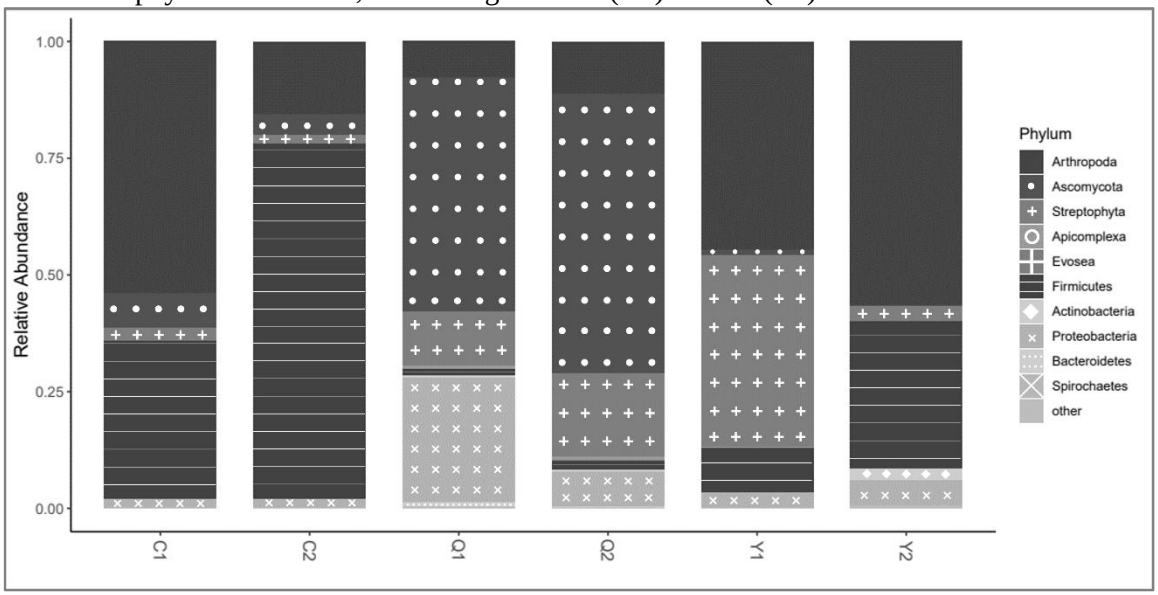

Figure 2. *Bar plot illustrates the relative abundance of phylum in each sample. Only the top 10 phyla are shown.*

For *Apilactobacillus*, a significant bacterial genus reported in other bee-related studies [26], remarkable differences were observed between samples from distinct regions in our study. In C1, C2, and Y1, *Apilactobacillus* constituted 34% – 77% of the bacterial population, with over 96% identified as *A. kunkeei*. In samples Q1 and Q2, however, *Apilactobacillus* was only 1%, with the three species, *Apilactobacillus* sp. *SG5_A10*, *A. bombintestini*, and *A. kunkeei* present in approximately equal proportions. *A. kunkeei* ferments sugars like glucose and fructose, producing lactic acid and other organic acids that contribute to honey's preservation and flavor [27]. Thriving in honey's acidic pH (3.5 – 4.5) [1], *A. kunkeei* is pivotal in enhancing honey's sensory attributes and shelf-life. Its fermentation creates an environment unfavorable to spoilage microorganisms, highlighting its significance in food preservation and potential probiotic applications [28]. In addition, we also detected *Enterococcus faecium*, *Lactobacillus panisapium*, *Escherichia coli*, and *Paraburkholderia fungorum* at varying levels, reflecting the diversity of bacteria in the honey samples. While some species are commonly associated with bee gut microbiota or environmental sources, *E. coli* may indicate potential contamination and should be further investigated.

As mentioned, some previous studies have shown that most of the DNA in honey samples is viral, ranging from 40% [4] to more than 98% [29]. The viral group often dominates in other studies of the honey metagenome. However, in our samples, the percentage of viral DNA was very low; the highest was in sample C1, with 0.6%. There were differences in the viral species composition of the samples in different regions. While the read percentage of Polybotosvirus accounted for the majority in samples Q1 and Q2, the group of viruses *Apis mellifera* filamentous virus (AmFV) was dominant in the other four samples. In addition to its primary host, *A. mellifera*, AmFV has also been detected in populations of *A. cerana* in China [30]. Studies on this virus indicated that it is relatively harmless to bees [31].

3.3. Metagenomic assembly

Table 2 presents the results of metagenome assembly. The samples' contiguity is different, expressed through the average length of contigs (auN), the index N50, L50, and the number of contigs in each sample. Sample C1 has a high N50 index but a low L50 index and the lowest number of contigs, indicating that this sample has the best continuity, while Q2 has the worst continuity. The contig sets were used for gene prediction and functional annotation.

Table 2. *Reference-free metrics for metagenomic assembly result of each sample*

Sample	No. contigs	Total length	auN	N50	L50
C1	167,377	323,769,217	28,018	11,464	5,425
C2	162,804	278,111,171	13,782	3,411	18,216
Q1	342,323	395,716,891	52,599	1,465	35,177
Q2	400,709	298,592,043	9,188	704	117,939
Y1	1,386,043	1,424,228,826	5,514	1,244	250,196
Y2	336,418	415,887,259	14,656	2,459	19,301

3.3. Functional profiles

To our knowledge, limited research has been conducted on the functional gene composition of metagenomes in honey. The gene function profiles (Figure 3) in this study highlight the diverse metabolic activities and biological processes the microbial community performs in honey. Notably, the functional gene composition across honey samples is mainly consistent. Prominent functional groups include carbohydrate metabolism and transport, cell membrane/wall/capsule formation, replication, transcription, translation, and ion metabolism and transport – essential processes required for cellular survival. Compared to the gut microbiota of honeybees [32], the honey microbiota also exhibits a high abundance of genes related to carbohydrate metabolism

and transport, underscoring the distinctive characteristic of microorganisms thriving in honey that adapt to a high-sugar environment.

Notably, the "Defense mechanisms" functional group accounts for an average proportion of 2.94 ± 0.13%. Honey is an environment characterized by high sugar concentration, low water activity, low pH, and the presence of antimicrobial compounds. Microorganisms thriving under such conditions require defense mechanisms to adapt to the environment. Most genes in this group in our honey samples encode efflux pump proteins, particularly multidrug efflux pumps, indicating that the microorganisms in honey may possess the ability to resist multiple types of antibiotics. Genes related to the transport and metabolism of secondary compounds were also present in all four samples, with an average proportion of 1.52 ± 0.43%. Although their proportion is not high, their presence could still be valuable for researching and developing of highly applicable compounds such as antimicrobial agents and antioxidants.

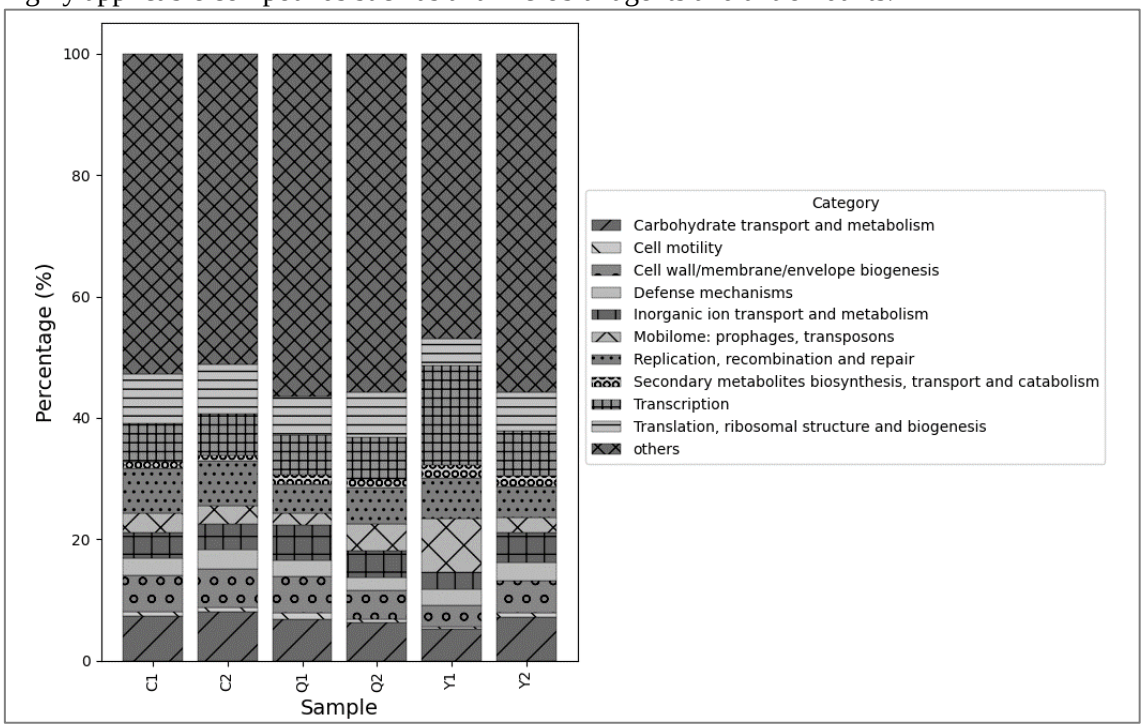

Figure 3. *Bar plot showing the abundance of gene function in each sample*

4. Conclusions

This study provides a foundational analysis of the microbial and functional gene composition in honey samples from Son La province. The microbial composition of honey samples varies significantly and is influenced by environmental factors. We identified the presence of efflux pump genes in metagenomic sequences, demonstrating microbial adaptation to the extreme conditions in honey. Our findings highlight the diverse microbial ecosystem of honey and its potential for developing biologically active compounds.

Acknowledgement

Nguyen Hoang Nam was funded by Vingroup JSC and supported by the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Institute of Big Data, code VINIF.2021.TS.127.

REFERENCES

[1] P. M. Da Silva, C. Gauche, L. V. Gonzaga, A. C. O. Costa, and R. Fett, "Honey: Chemical composition, stability and authenticity," *Food Chemistry*, vol. 196, pp. 309-323, 2016.

[2] H. Abuelgasim, C. Albury, and J. Lee, "Effectiveness of honey for symptomatic relief in upper respiratory tract infections: a systematic review and meta-analysis," *BMJ Evidence-Based Medicine*, vol. 26, no. 2, pp. 57-64, 2021.

[3] Z. N. François, M. F. G. Perin, and K. P. Marie, "Antimicrobial activity of probiotic strain *Lactobacillus plantarum* isolated from 'sha'a' and assessment of its viability in local honey," *Journal of Microbiology, Biotechnology and Food Sciences*, vol. 2021, no. 10, pp. 226-231, 2021.

[4] H. K. Wirta, M. Bahram, K. Miller, T. Roslin, and E. Vesterinen, "Reconstructing the ecosystem context of a species: Honey-borne DNA reveals the roles of the honeybee," *PLOS One*, vol. 17, no. 7, 2022, Art. no. e0268250.

[5] S. Bovo, V. J. Utzeri, A. Ribani, R. Cabbri, and L. Fontanesi, "Shotgun sequencing of honey DNA can describe honey bee derived environmental signatures and the honey bee hologenome complexity," *Sci. Rep.*, vol. 10, no. 1, 2020, Art. no. 1.

[6] S. Bovo, A. Ribani, V. J. Utzeri, G. Schiavo, F. Bertolini, and L. Fontanesi, "Shotgun metagenomics of honey DNA: Evaluation of a methodological approach to describe a multi-kingdom honey bee derived environmental DNA signature," *PLOS One*, vol. 13, no. 10, 2018, Art. no. e0205575.

[7] A. Galanis, P. Vardakas, M. Reczko, V. Harokopos, P. Hatzis, E. M. C. Skoulakis, G. A. Pavlopoulos, and S. Patalano, "Bee foraging preferences, microbiota and pathogens revealed by direct shotgun metagenomics of honey," *Molecular Ecology Resources*, vol. 22, no. 7, pp. 2506-2523, 2022.

[8] T. L. Pham, T. T. D. Bui, T. T. Ha, T. H. Nguyen, M. S. Yoo, Y. S. Cho, and V. Q. Dong, "The gut microbiota at different developmental stages of *Apis cerana* reveals potential probiotic bacteria for improving honeybee health," *Microorganisms*, vol. 10, no. 10, 2022, Art. no. 10.

[9] C. Billington, J. M. Kingsbury, and L. Rivas, "Metagenomics approaches for improving food safety: A review," *Journal of Food Protection*, vol. 85, no. 3, pp. 448-464, 2021.

[10] M. H. Bui, T. T. Tran, H. N. Nguyen, and T. T. H. Le, "Species composition and nest distribution of natural honey bee (Hymenoptera: Apidae) in the mountainous area of Northwest, Vietnam," *TNU Journal of Science and Technology*, vol. 229, no. 01, pp. 198-203, 2023.

[11] F. E. Buytaers, A. Saltykova, W. Mattheus, B. Verhaegen, N. H. C. Roosens, K. Vanneste, V. Laisnez, N. Hammami, B. Pochet, V. Cantaert, K. Marchal, S. Denayer, and S. C. J. De Keersmaecker, "Application of a strain-level shotgun metagenomics approach on food samples: resolution of the source of a *Salmonella* food-borne outbreak," *Microb. Genom.*, vol. 7, no. 4, 2021, Art. no. 000547.

[12] T. N. T. Nguyen, N. T. Nguyen, D. T. Tran, P.-C. Kuo, B. T. Nguyen, N. T. Le, H. T. Le, H. D. T. Nguyen, D. C. Vu, T. L. Ho, N. A. Le, and T. T. T. Nguyen, "Chemical composition analysis and antioxidant activity of *Coffea robusta* monofloral honeys from Vietnam," *Foods*, vol. 11, no. 3, 2022, Art. no. 3.

[13] N. Q. Pham, H. B. Luu, T. T. Vu, and T. T. Q. Cung, "Comparison of antibacterial activities of some kinds of honey in Vietnam," *Vietnam Journal of Food Control*, vol. 5, no. 1, pp. 77-88, 2022.

[14] S. Andrews, "FastQC: A quality control tool for high throughput sequence data," 2022. [Online]. Available: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [Accessed: August 24, 2023]

[15] A. M. Bolger, M. Lohse, and B. Usadel, "Trimmomatic: a flexible trimmer for Illumina sequence data," *Bioinformatics*, vol. 30, no. 15, pp. 2114-2120, 2014.

[16] B. Langmead and S. L. Salzberg, "Fast gapped-read alignment with Bowtie 2," *Nat Methods*, vol. 9, no. 4, 2012, Art. no. 4.

[17] D. E. Wood, J. Lu, and B. Langmead, "Improved metagenomic analysis with Kraken 2," *Genome Biol*, vol. 20, no. 1, Art. no. 257, 2019.

[18] D. Li, C. M. Liu, R. Luo, K. Sadakane, and T. W. Lam, "MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph," *Bioinformatics*, vol. 31, no. 10, pp. 1674-1676, 2015.

[19] J. Lu, F. P. Breitwieser, P. Thielen, and S. L. Salzberg, "Bracken: estimating species abundance in metagenomics data," *PeerJ Computer Science*, vol. 3, 2017, Art. no. e104.

[20] A. Mikheenko, V. Saveliev, and A. Gurevich, "MetaQUAST: evaluation of metagenome assemblies," *Bioinformatics*, vol. 32, no. 7, pp. 1088-1090, 2016.

[21] D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, and L. J. Hauser, "Prodigal: prokaryotic gene recognition and translation initiation site identification," *BMC Bioinformatics*, vol. 11, no. 1, 2010, Art. no. 119.

[22] J. C. Sequeira, M. Rocha, M. M. Alves, and A. F. Salvador, "UPIMAPI, reCOGnizer and KEGGCharter: Bioinformatics tools for functional annotation and visualization of (meta)-omics datasets," *Computational and Structural Biotechnology Journal*, vol. 20, pp. 1798-1810, 2022.

[23] S. Liu, D. Lang, G. Meng, J. Hu, M. Tang, and X. Zhou, "Tracing the origin of honey products based on metagenomics and machine learning," *Food Chemistry*, vol. 371, 2022, Art. no. 131066.

[24] H. K. Wirta, N. Abrego, K. Miller, T. Roslin, and E. Vesterinen, "DNA traces the origin of honey by identifying plants, bacteria and fungi," *Sci. Rep.*, vol. 11, no. 1, 2021, Art. no. 1.

[25] S. Ullah, F. Huyop, R. A. Wahab, I. G. A. Sujana, N. S. Antara, and I. B. W. Gunam, "Using pollen DNA metabarcoding to trace the geographical and botanical origin of honey from Karangasem, Indonesia," *Heliyon*, vol. 10, no. 12, 2024, Art. no. e33094.

[26] J. Crovadore, R. Chablais, F. Raffini, B. Cochard, M. Hänzi, F. Gérard, K. K. Jensen, and F. Lefort, "Draft genome sequences of 3 strains of *Apilactobacillus kunkeei* isolated from the bee gut microbial community," *Microbiology Resource Announcements*, vol. 10, no. 13, 2021, Art. no. e00088-21.

[27] A. Endo, S. Maeno, Y. Tanizawa, W. Kneifel, M. Arita, L. Dicks, and S. Salminen, "Fructophilic lactic acid bacteria, a unique group of fructose-fermenting microbes," *Appl. Environ. Microbiol.*, vol. 84, no. 19, 2018, Art. no. e01290-18.

[28] T. C. Olofsson and A. Vásquez, "Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee *Apis mellifera*," *Curr. Microbiol.*, vol. 57, no. 4, pp. 356-363, 2008.

[29] S. Bovo, A. Ribani, V. J. Utzeri, G. Schiavo, F. Bertolini, and L. Fontanesi, "Shotgun metagenomics of honey *DNA*: Evaluation of a methodological approach to describe a multi-kingdom honey bee derived environmental DNA signature," *PLOS One*, vol. 13, no. 10, 2018, Art. no. e0205575.

[30] C. Hou, B. Li, Y. Luo, S. Deng, and Q. Diao, "First detection of *Apis mellifera* filamentous virus in *Apis cerana* cerana in China," *Journal of Invertebrate Pathology*, vol. 138, pp. 112-115, 2016.

[31] U. Hartmann, E. Forsgren, J.-D. Charrière, P. Neumann, and L. Gauthier, "Dynamics of *Apis mellifera* filamentous virus (AmFV) infections in honey bees and relationships with other parasites," *Viruses*, vol. 7, no. 5, 2015, Art. no. 5.

[32] P. Engel and N. A. Moran, "Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis," *Gut. Microbes*, vol. 4, no. 1, pp. 60-65, 2013.