Application-Driven Honey Selection via Principal Component Analysis-Aided Optimisation

Lip Siang Yeo[a], Zhang Ru Bong[a], Karen Gah Hie Kong[a], Irine Runnie Henry Ginjom[a], Sin Yong Teng[b], Bing Shen How[a,*]

[a]Research Centre for Sustainable Technologies, Faculty of Engineering, Computing and Science, Swinburne University of Technology, Jalan Simpang Tiga, 93350, Kuching, Sarawak, Malaysia
[b]Radboud University, Institute for Molecules and Materials, P.O. Box 9010, 6500 GL, Nijmegen, the Netherlands
bshow@swinburne.edu.my

Honey, a natural sweetener produced by bees from floral nectar, is known for its diverse physicochemical properties found in various applications (e.g., medical, dietary). The clinical utilisation of honey highlights the importance of antimicrobial and antioxidant benefits in various medical treatments. The primary factor contributing to the consumer purchasing behaviour of honey for dietary needs is mainly influenced by taste, texture, and colour. This challenged the honey selection process, ensuring the suitability and quality of honey in diverse contexts. This work focuses on addressing the multi-varied honey selection problem utilising Multi-Criteria Decision-Making (MCDM) analysis to identify key factors contributing to the selection of honey attributes based on the intended use. This provides a systematic and objective means of selecting honey varieties for clinical and dietary applications. The model favours the New Zealand *Apis* Manuka honey (Manuka2) for consumer dietary attributed to its high sugar concentration (33.24 wt%) and low acidity (4.04 mEq/kg) contributing to the desirable sweet and non-acidic flavour. The raw stingless bee honey (SBHR5) from Sarawak local bee farm exhibits impressive antimicrobial properties such as high free acidity (19.19 mEq/kg), DPPH assay (908.21 mg/mL), and total phenolic content (384.27 mg GAE/kg dwb) desirable for clinical use. This work explores the potential of utilising Principal Component Analysis (PCA)-aided optimisation for honey selection and its significance in identifying honey varieties with desired characteristics.

1. Introduction

The physicochemical properties of honey are influenced by several factors (e.g. climatic conditions, nectar source, and postharvest handling practices) (Tigistu et al., 2021). These variations are expected to result in differences in honey quality (e.g., appearance, structure, and texture) and stability (e.g., water activity), across different sources, including pharmacological effects (e.g., anti-diabetic, anti-inflammatory, and antioxidant). This has subsequently boosted the health interest in honey applications in various fields (e.g., tissue engineering (Yupanqui Mieles et al., 2022), herpes treatment (Semprini et al., 2019), and manufactured food products such as powdered honey (Osés et al., 2022) and yoghurt (Merrzlov et al., 2018)). Clinical utilisation of honey emphasises the importance of antimicrobial activities of honey for treatment which inhibits bacteria during the cell recovery process (Yupanqui Mieles et al., 2022). Consumer preferences play a vital role in honey selection, with taste, aroma, colour, and texture being important factors in purchasing behaviour (Roman et al., 2013). Processing honey is challenged by the viscous properties of honey, which prioritise the selection of honey with lower moisture content (Osés et al., 2022).

The vast decision criteria from the diverse honey properties with varied correlations to one another challenged decision-makers in selecting the optimal honey source to satisfy the desired requirements. Pentoś et al. (2015) utilizes the advanced neural network model to identify the influence of various chemical properties on the electrical properties of honey. This enables the systematic assessment of honey quality based on the electrical parameters, which is useful in practice instead of the complex assessment of chemical parameters in honey. Contrary to honey properties, the deployment of systematic evaluation techniques (e.g., Analytical Hierarchy

Process (AHP), weighted sum approach) in beekeeping locations provide beekeepers with crucial insights to optimize honey production and quality (Shaharudin et al., 2023). This enables the assessment of multicriteria problems addressing suitable apiary locations for optimal honey production. Despite the diverse application of honey, there is lacking contribution towards a systematic evaluation technique in the honey literature, especially the optimal selection of honey sources to satisfy the desired requirements.

This paper proposed the use of Principal Component Analysis (PCA) as a statistical approach to solve the Multi-Criterial Decision-Making (MCDM) problem of honey selection. PCA evaluates statistically significant variables based on the contribution and achieves dimension reduction. This enables the translation of high-variable problems into manageable sizes with minimal information loss. PCA-aided optimisation has been pioneered for process optimisation (Teng et al., 2019) and technology selection (How and Lam, 2018), showcasing its effectiveness in evaluating the optimal solution. This paper explores the capability of PCA-aided optimisation in synthesising the optimal honey variety considering the diverse properties to satisfy the desirable characteristics.

2. Problem statement

The model is formulated to determine the optimal honey sources dependent on its intended application. Figure 1 describes the various source of honey (*Apis* Manuka, *Apis* Acacia, processed stingless bee, raw stingless bee) with its diverse physicochemical properties (e.g., free acidity, total sugar, phenolic, etc.) influenced by the desired characteristics based on applications in clinical or consumer dietary. Besides, the honey price and distance from the source (lower distance indicating lower carbon footprint potential) are considered as indicators for economic and environmental goals. Table 1 tabulates the preferential optimise direction (minimise = min; neutral (no preference) = neu; maximise = max) of the desired honey properties in clinical and dietary applications.
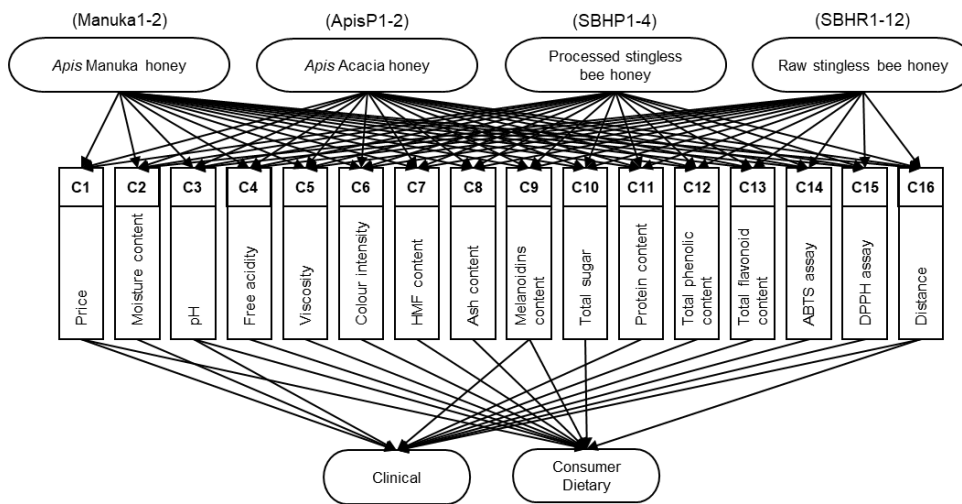

Figure 1: Superstructure of application-driven honey selection problem

Table 1: Preferential criterion optimizing direction for various honey application

Criterion	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
Clinical	Min	Neu	Min	Max	Neu	Neu	Neu	Min	Max	Neu	Max	Max	Max	Max	Max	Min
Dietary	Min	Min	Max	Min	Max	Max	Max	Min	Max	Max	Max	Max	Max	Max	Max	Min

3. Methodology

Figure 2 illustrates the proposed PCA optimisation model for honey selection. The honey database documents 20 honey sources (e.g., *Apis* Manuka, Sarawak stingless bees, etc.) with diverse properties (e.g., moisture, pH, free acidity, viscosity, phenolic, flavonoid, etc.) identified through various lab analyses. This serves as the decision criteria for the various intended honey application (e.g., clinical, consumer dietary, food production, etc.). The formulated PCA optimisation model reduces the complexity of the problem through dimension reduction by translating the largely correlated variables (honey properties) into smaller uncorrelated variables (Principal Components, PCs).
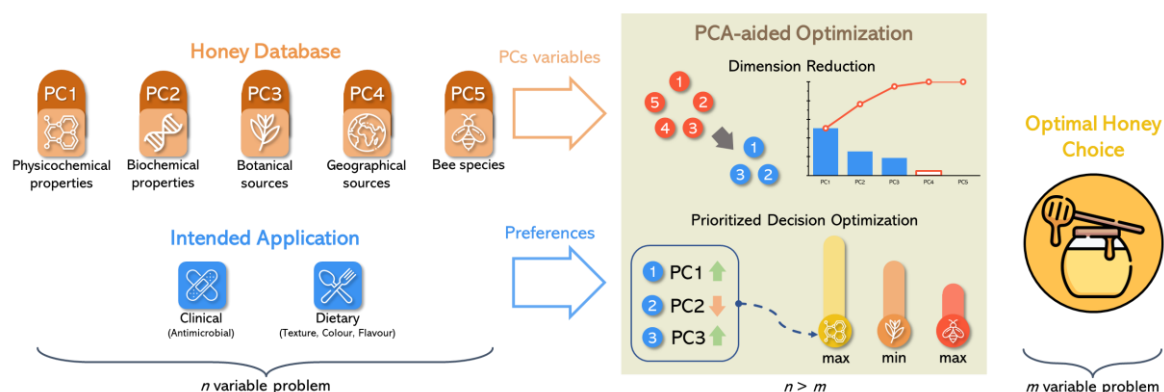
Figure 2: PCA-Aided optimization model for honey selection

3.1 Honey properties analysis

Table 2 documents the various experimental methodology utilised to analyse the various honey properties in this work. The analysis results represent the decision-making data for the honey selection problem.

Table 2: Summary of the methodology used for the selected physicochemical properties of honey samples

Methodology	Physicochemical properties	References
Refractometric method	Moisture content	(Malaysian Standard, 2017)
Spectrophotometric method	Colour intensity, Melanoidins content, HMF content, ABTS assay, DPPH assay	(Imtara et al., 2018; Malaysian Standard, 2017; Trinh et al., 2022; Pontis et al., 2014)
Titration method	Free acidity, Total sugar	(Sereia et al., 2017)
Aluminium chloride colourimetric method	Total flavonoid content	(Sereia et al., 2017)
Bradford assay	Protein content	(Ramon-Sierra et al., 2015)
Folin-Ciocalteu assay	Total phenolic content	(Kek et al., 2017)
Conductivity meter	Ash content	(IHC, 2009)
pH meter	pH	(Malaysian Standard, 2017)
Viscometer with water bath	Viscosity	(James et al., 2009)

3.2 PCA-aided decision-making model

This work follows the extension of a novel PCA application as multi-objective decision-making (MODM) analysis to evaluate the various solutions optimality based on the PCs score (How and Lam, 2018). The PCs of the scaled data (Z^{std}) is computed as the correlation matrix (R) using singular value decomposition (SVD) where U (left singular matrix), V (right singular matrix), and Σ (diagonal matrix) (see Eq.(1)).

$$R = U\Sigma V^T \tag{1}$$

The decomposed components is represented as the eigenvectors (V) and eigenvalues ($\lambda = \Sigma^2$) following the eigen decomposition of R (Bro and Smilde, 2014) (see Eq(2)).

$$R^T \cdot R = V\Sigma^2 V^T = V\lambda V^T \tag{2}$$

The variance of respective PCs is determined from the eigenvalues (λ) which, the featured PCs (m) is identified with a cumulative variance of 90 % threshold ensuring minimal information lost. Eq(3) denotes the featured PCs score (PC_m) computed as the dot product of the scaled data (Z^{std}) and the featured PCs loadings (V_m). The respective PCs optimizing direction is dependent on the preferential direction (maximize or minimize) of each criterion influenced by the intended honey application. Details of evaluating the PCs optimizing direction is documented in prior work by How and Lam (2018).

$$PC_m = Z^{std} \cdot V_m \tag{3}$$

The objective function of the model is formulated as the maximisation of the total PCs score denoted as the sum of all featured PCs score (PC_m) and the respective weightage (W_m) determined from the variance (see Eq(4)).

This enables the proposition of optimal selection with prioritised criteria from the various honey properties depending on the desired honey application.

$$Objective = \max \sum_m (PC_m \times W_m) \tag{4}$$

4. Results and discussion

Figure 3(a) illustrates that the 4 featured PCs are sufficient to describe the problem with a cumulative variance ≥ 90 %. The respective criterion contribution towards the 4 featured PCs is documented in Table 3, where the decision preference (from Table 1) will influence the PCs optimising direction. The 20 honey sources are projected onto the 4 featured PCs (scatter points) with the loadings of 15 criteria (vectors) (see Figure 3(b-c)).
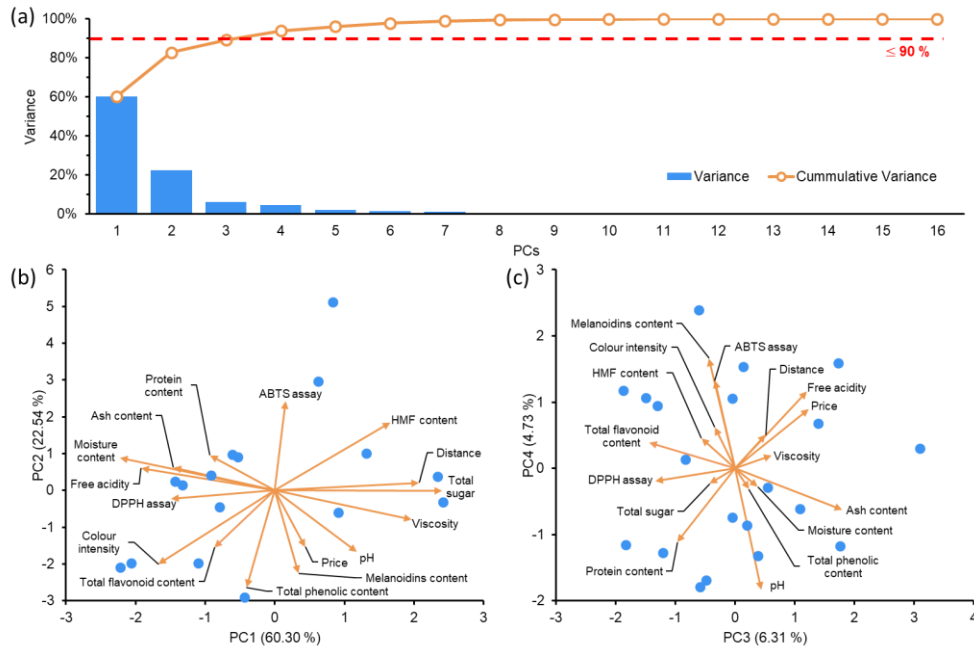

Figure 3: (a) Pareto Analysis (Dimension Reduction); (b) PCA plot PC1 vs PC2; (c) PCA plot PC3 vs PC4

Table 3: Various criteria contribution % (based on PCs loading) towards the respective featured PCs

Criteria	PC1	PC2	PC3	PC4
Price	0.53	-6.51	11.74	6.16
Moisture content	-13.76	2.20	1.12	-0.59
pH	3.85	-7.77	1.46	-25.97
Free acidity	-10.27	1.03	11.18	10.37
Viscosity	10.78	-1.75	2.88	0.29
Colour intensity	-7.77	-11.13	-0.90	2.97
HMF	7.65	9.46	-2.44	1.59
Ash content	-5.89	1.08	25.00	-3.04
Melanoidins	0.32	-14.08	-1.46	21.27
Total sugar	16.08	0.00	-1.40	-0.41
Protein content	-2.41	2.54	-7.20	-9.55
Total phenolic content	-0.46	-19.22	0.42	-0.75
Total flavonoid content	-2.08	-6.64	-16.02	1.16
ABTS assay	0.06	16.34	-0.88	13.61
DPPH assay	-6.19	-0.13	-13.87	-0.27
Distance	11.92	0.12	2.04	1.99

Figure 4 illustrates the performance score (total PCs score) of the respective honey sources for both clinical and dietary preferences. The model favours Sarawak local raw stingless bee honey (SBHR5, PCs score = 86.45 %)

for clinical application, while New Zealand *Apis* Manuka honey (Manuka2, PCs score = 89.17 %) has opted for consumer dietary. The clinical choice of SBHR5 is attributed by higher scores in dominant PCs (PC1 = 55.65 %, PC2 = 21.54 %) as compared to Manuka2 (PC1 = 0 %, PC2 = 20.58 %), which performed poorly. The dominant PC1 prioritises higher free acidity and DPPH assay of honey, indicating better antimicrobial activities of SBHR5 (free acidity = 19.19 mEq/kg, DPPH = 908.21 mg/mL) over Manuka2 (free acidity = 4.04 mq/kg, DPPH assay = 349.82 mg/mL). Interestingly, SBHR3 performed better in PC1 (62.00 %), offering free acidity and DPPH assay value at 3.49 % and 67.84 % higher than SBHR5. The pitfall of SBHR3 is attributed to the lower total phenolic content (13.68 %) and Melanoidins content (42.59 %), contributing to poorer performance (PC2 = 15.88 %) compared to SBHR5 (PC2 = 21.54 %). For consumer dietary perspective, Manuka2 (PC1 = 64.23 %, PC2 = 20.58 %) is favoured by the model over SBHR5 (PC1 = 8.58 %, PC2 = 21.54 %) attributed by the better performance in PC1. This contributes to the selection of honey with higher total sugar and lower moisture content, such as Manuka2 (total sugar = 33.24 wt%, moisture = 15.35 wt%) over SBHR5 (total sugar = 20.46 wt%, 31.42 wt%) as preferable flavour and texture for consumer. In comparison to the environmental goals, Manuka2 is imported from New Zealand (8,082 km) and contributes to higher emission potential considering longer travel distance as compared to locally sourced honey SBHR5 (25 km).
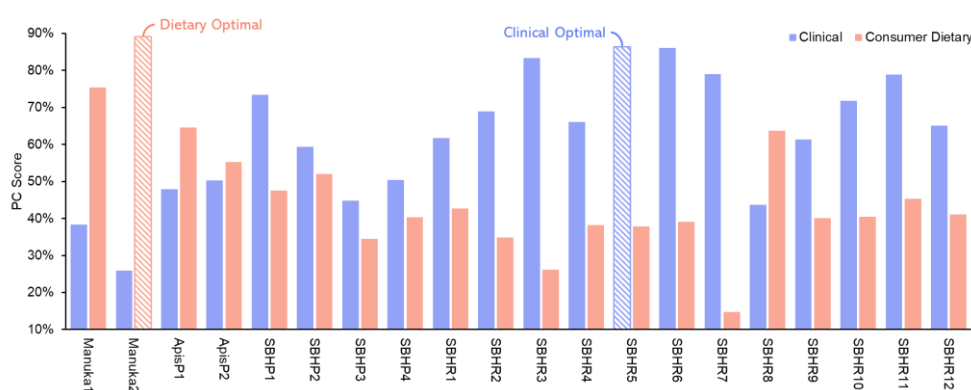

Figure 4: PCs score of various honey varieties for clinical (blue) and consumer dietary (red)

5. Conclusion

The presented model in this work demonstrated the effectiveness of utilising PCA-aided optimisation in selecting various optimal honey sources for clinical (antimicrobial properties) and consumer dietary preferences (flavour and texture). One of the *Apis* Manuka honey (Manuka2) stands out with exceptional potential for dietary among all the honey samples with PCs score of 89.17 %. This is attributed by the high sugar concentration (33.24 wt%) and lower free acidity (4.04 mEq/kg) of the *Apis* Manuka honey originated from New Zealand. It offers a sweeter and less sour flavour profile enhancing its appeal for consumer consumption. Manuka2 also exhibit more viscous texture with the highest viscosity (79.63 Pa-s) among other honey samples highlighting the preferred choice for the consumer. SBHR5, one of the raw stingless bee honey samples, exhibits promising characteristics for clinical research with a PCs score of 86.45 %. The promising antimicrobial properties of SBHR5 are favoured by the high free acidity (19.19 mEq/kg) and DPPH assay (908.21 mg/mL), indicating its capability to inhibit the growth and proliferation of pathogenic bacteria. SBHR5 also demonstrates the highest total phenolic content (384.27 mg GAE/kg dwb), outperforming other honey samples. This highlights the honey's remarkable ability to scavenge free radicals and protect against oxidative stress indicating its valuable candidate for various clinical applications. The PCA-aided optimisation model approach can guide decision-making processes in apiculture by identifying the nutritional composition of honey and prioritising factors, traits and variables that influence trait optimisation (e.g., nutrient profile, stability and texture). The expansion towards a more diverse honey properties (e.g., agar well diffusion assay and broth dilution method) and larger database considering various parameters (e.g., geological source, supply-chain value, productivity) can enhance the robustness of the PCA-aided optimisation model to consider various economic and environmental indicators for optimal honey selection.

References

Bro R., Smilde A.K., 2014, Principal component analysis. Analytical Methods, 6, 9, 2813-2831.

How B.S., Lam H.L., 2018, Sustainability evaluation for biomass supply chain synthesis: novel principal component analysis (PCA) aided optimisation approach. Journal of Cleaner Production, 189, 941-961.

Imtara H., Elamine Y., Lyoussi B., 2018, Physicochemical characterization and antioxidant activity of Palestinian honey samples. Food Science & Nutrition, 6(8), 2056-2065.

IHC, 2009, Harmonised Methods of The International Honey Commission, International Honey Commission. <ihc-platform.net/ihcmethods2009.pdf>, accessed 10/06/2023.

James O., Mesubi M., Usman L., Yeye S., Oluseyi A., Olurotimi O., Ajani O., Siyanbola T., 2009, Physical characterisation of some honey samples from North-Central Nigeria. International Journal of Physical Sciences, 4, 464-470.

Kek S.P., Chin N.L., Yusof Y.A., Tan S.W., Chua L.S., 2017, Classification of entomological origin of honey based on its physicochemical and antioxidant properties. International Journal of Food Properties, 20(sup3), S2723-S2738.

Malaysian Standard, 2017, MS2683:2017: Kelulut (stingless bee) honey - specification: quality requirements, Department of Standard Malaysia, Department of Standard Malaysia. <brizy.cloud/customfile/7512f96ecef15ae5d649ccb4015712b9.pdf>, accessed 10 June 2023.

Merrzlov S., Lomova N., Narizhniy S., Rudakova T., Snizhko O., Viktor Voroshchuk V., 2018, Managing quality and safety during the production of yoghurt with honey products. Nauka ta Innovacii, 14(6), 24-37.

Osés S.M., Cantero L., Puertas G., Fernández-Muino M. Á., Sancho M.T., 2022, Antioxidant, antimicrobial and anti-inflammatory activities of ling-heather honey powder obtained by different methods with several carriers. LWT, 159, 113235.

Pentoś K., Łuczycka D., Kapłon T., 2015, The identification of relationships between selected honey parameters by extracting the contribution of independent variables in a neural network model. European Food Research and Technology, 241(6), 793–801.

Pontis J.A., Costa L.A.M.A.D., Silva S.J.R.D., Flach A., 2014, Color, phenolic and flavonoid content, and antioxidant activity of honey from Roraima, Brazil. Food Science and Technology, 34, 69-73.

Ramon-Sierra J., Ruiz-Ruiz J., Ortiz-Vazquez E., 2015, Electrophoresis characterisation of protein as a method to establish the entomological origin of stingless bee honeys. Food Chemistry, 183, 43-48.

Roman A., Popiela-Pleban E., Kozak M., 2013, Factors influencing consumer behavior relating to the purchasing of honey part 1. The buying process and the level of consumption. Journal of Apicultural Science, 57(2), 159-172.

Semprini A., Singer J., Braithwaite I., Shortt N., Thayabaran D., McConnell M., Weatherall M., Beasley R., 2019, Kanuka honey versus aciclovir for the tropical treatment of herpes simplex labialis: a randomised controlled trial. BMJ Open, 9(5), e026201.

Sereia M.J., Março P.H., Perdoncini M.R.G., Parpinelli R.S., De Lima E.G., Anjo F.A., 2017, Techniques for the Evaluation of Physicochemical Quality and Bioactive Compounds in Honey. In: Toledo, V.D.A.A.D. (Ed.), Honey Analysis. InTech Open, DOI: 10.5772/66839.

Shaharudin N.A., Rahman S.A., Benjamin A.M., Omar M.F., Man M., 2023, On the use of multi-criteria decision making model for selecting the important criteria in meliponiculture. Universal Journal of Agricultural Research, 11(2), 336–343.

Teng S.Y., How B.S., Leong W.D., Teoh J.H., Cheah A.C.S., Motavasel Z., Lam H.L., 2019, Principal component analysis-aided statistical process optimisation (PASPO) for process improvement in industrial refineries. Journal of Cleaner Production, 225, 359-375.

Tigistu T., Worku Z., Mohammed A., 2021, Evaluation of the physicochemical properties of honey produced in Doyogena and Kachabira Districts of Kembata Tamboro zone, Southern Ethiopia. Heliyon, 7(4), e06803.

Trinh N.T.N., Tuan N.N., Thang T.D., Kuo P.C., Thanh N.B., Tam L.N., Tuoi L.H., Nguyen T.H.D., Vu D.C., Ho T.L., Anh L.N., Thuy N.T.T., 2022, Chemical composition analysis and antioxidant activity of coffea robusta monoflorals honeys from Vietnam. Foods, 11(3), 388.

Yupanqi Mieles J., Vyas C., Aslan E., Humphreys G., Diver C., Bartolo P., 2022, Honey: An advanced antimicrobial and wound healing biomaterial for tissue engineering applications. Pharmaceutics, 14(8), 1663.