Comparison of Metabarcoding Techniques for Dietary Assessment in Herbivores and Omnivores

Nur Hartini Sariyati[1], *Hidayah* Haris[1], *Farah* Farhana Ramli[1], *Suliadi* Firdaus Sufahani[1] *Badrul* Munir Md-Zain[2], *Salmah* Yaakop[2], *Abd Rahman* Mohd-Ridwan[3] *Nor* Rahman Aifat[4], *Mohd Sanusi* Mohamed[5], *Ibnu* Maryanto[6] *Muhammad* Abu Bakar Abdul-Latiff[1] *

1Environmental Management and Conservation Research Unit (eNCORe), Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (Pagoh Campus), KM1 Jalan Panchor, 84600, Johor, Malaysia
2Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Bangi, 43600, Malaysia
3Centre for Pre-University Studies, Universiti Malaysia Sarawak, Sarawak, Kota Samarahan, 94300, Malaysia
4Faculty of Tropical Forestry, Universiti Malaysia Sabah (UMS), Jalan UMS, Sabah, Kota Kinabalu, 88400, Malaysia
5Copenhagen Zoo, Masreca 19,Persiaran Rimba Permai,Cyber 10,Selangor, Sepang, 63000, Malaysia
6Museum Zoologicum Bogoriense, Widyasatwaloka Building, Research Centre in Biosystematic and Evolution (BRIN), Jl. Raya Cibinong KM 46 Cibinong, Indonesia

Abstract. Dietary assessment plays a crucial role in comprehending the ecological dynamics and nutritional needs of herbivores and omnivores. The metabarcoding technique has emerged as a potent tool for exploring the dietary composition of these animals. However, various metabarcoding techniques have been developed, each with its own advantages and limitations. This study aims to compare the performance of different metabarcoding techniques in herbivores and omnivores diet. We systematically reviewed 159 published manuscripts in Scopus and Google Scholar, and thematic analysis was conducted across several categories, including the marker, platform, and database utilized. Preliminary findings reveal significant variations among metabarcoding techniques across these two animal groups. The trnL gene exhibited higher taxonomic resolution for herbivorous species, whereas the combination of the 'trnL + 16s rRNA' gene exhibited superior performance for omnivorous species. The Illumina platform emerged as the most commonly used method for analyzing the diets of both herbivores and omnivores, with the primary reference database being the National Centre for Biotechnology Information (NCBI). This study offers valuable insights into the strengths and limitations of different metabarcoding techniques for dietary assessment in herbivores and omnivores and optimizing metabarcoding protocols, facilitating more precise and reliable diet analyses within these ecological groups.

1 Introduction

* Corresponding author: latiff@uthm.edu.my

DNA metabarcoding involves identifying a multitude of species within a single sample, which may consist of complete organisms or environmental specimens like water, soil, and degraded DNA [1]. This innovative approach is rapidly emerging as a practical substitute for traditional morphology-based species identification, particularly in large-scale investigations conducted in environments where logistical or financial constraints pose challenges [2]. Metabarcoding is a standard tool in numerous recent studies focusing on trophic interactions, including dietary analysis [3]. In the context of diet analysis, the diet is inferred from DNA traces found in fecal samples or the gut contents of consumers [4]. This molecular dietary research, including sequencing prey identification, can expose exact food webs throughout ecosystems using a reference library [5]. Forister *et al.* [6] proposed that diets frequently display imbalanced distributions, prioritizing a few primary resources while incorporating numerous rare ones.

Conventional methods for diet studies, such as direct observation, videotaping, or fecal microscopy, have various drawbacks and limitations [7]. Observing wild animals consuming unexpected foods that defy conventional knowledge is a common occurrence, as highlighted by Mehrkam [8]. While large mammalian herbivores possess the ability to consume a variety of food plants, they tend to exhibit preferences for specific options while avoiding others [9; 10]. Certain mammals exhibit dietary specialization, depending on a specific food source, whereas others are dietary generalists capable of consuming diverse foods [11]. Additionally, the food supplies that are accessible in a mammal's habitat have a significant impact on what it consumes [12]. Mammals have three main dietary groups: carnivores, omnivores and herbivores. Herbivores are mammals that own a primary diet based on plant material. They are further categorized into more specialized diets, such as folivores, frugivores, or granivores, based on their specific nutritional preferences and adaptations [10; 13]. In contrast, omnivores have a more diverse diet that encompasses both plant and animal matter, allowing them to adapt to fluctuations in food availability.

Symondson [14] is the first researcher who used DNA barcoding to assess the diets of wild animals, focusing on invertebrates. These investigations identified the predator's diet even after the food had been digested. Even though DNA metabarcoding has been chosen as the most practical tool for studying diet since it has enhanced the process of identifying foods, assessing dietary diversity, and measuring the relative abundance of taxa in the diets of wild animals [15, 16]. However, this method has occasionally come under scrutiny due to concerns about the accuracy of DNA identification, which affects both the choice of barcode markers and the technique used to analyze the results [17, 18]. Achieving precise taxonomic information at the species or genus level can be challenging through DNA metabarcoding, especially when reference databases are incomplete [19]. Additionally, implementing and conducting DNA metabarcoding experiments can incur substantial expenses and demand specialized equipment and expertise [20]. In this review, we compiled published research that employed metabarcoding techniques for the analysis of herbivores' and omnivores' diets to identify the barcode region, platforms and database utilized in previous studies. The strengths and weaknesses of various metabarcoding approaches for nutrition evaluation in herbivores and omnivores will be discussed to improve metabarcoding procedures, enabling more precise and trustworthy diet assessments in these ecological populations.

2 Methodology

Bibliographic searches were employed to acquire data from previous studies conducted on diet metabarcoding. Peer-reviewed articles were searched in both the Scopus database and Google Scholar, focusing on indexed titles, abstracts, keywords, and topics using the keyword "METABARCODING." Relevant studies, including theses and reports related to dietary analysis based on metabarcoding approaches, were included in the search. Figure 1

illustrates the steps taken to retrieve the publications for review. In the first step, 5950 articles were generated. However, only 5540 manuscripts were accessible, and publications were screened, and only publications related to the theme of diet metabarcoding were selected for review. Any publications that failed to meet the previously defined inclusion criteria were excluded. Subsequently, 346 duplicates were removed, and the publications were downloaded for filtering purposes, resulting in 329 articles. These manuscripts were then further categorized into different types of diets (carnivore, omnivore, and herbivore), resulting in a final selection of 159 manuscripts for herbivore and omnivore metabarcoding dietary studies. Review on diet metabarcoding studies were also excluded in this review. These publications were subsequently analysed based on the marker, platform, and database utilised in the diet metabarcoding technique. All of this information was used to create the infographics.

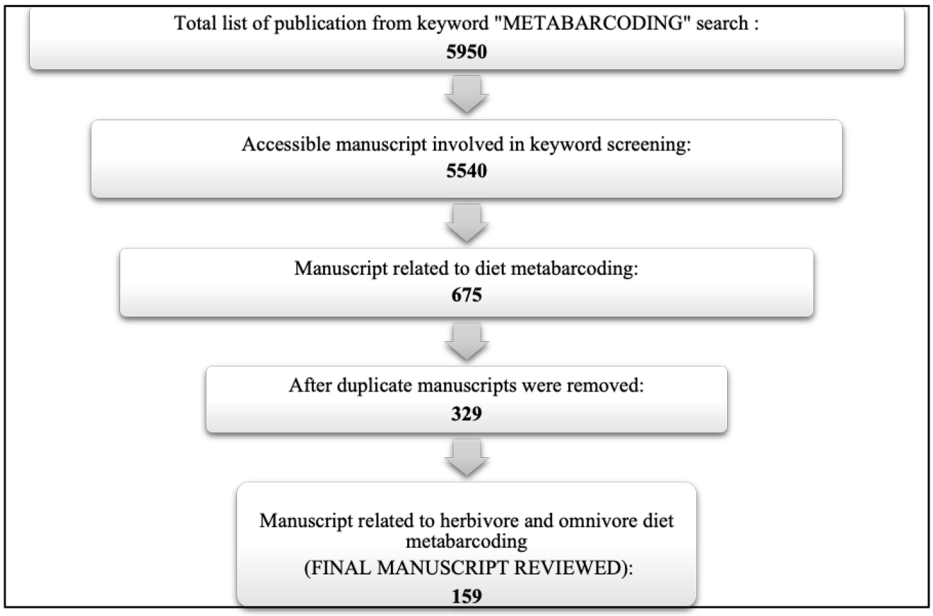

Fig. 1. Flowchart of the review process on publications related to metabarcoding

3 Results and Discussions

We identified 329 manuscripts related to diet metabarcoding from the systematic review, but only 47% of them focused on herbivore diets (94 manuscripts) and omnivore diets (65 manuscripts).

3.1 Results

3.1.1 Herbivore

There are certain articles in which multiple markers were combined for the analysis. In such cases, each article was counted separately based on the markers utilized, with multimarker disregarded as one category. This approach aimed to identify the most effective marker among those used in previous diet metabarcoding studies. Among herbivores, twelve DNA regions were utilized in metabarcoding techniques (Figure 2a). The results show that the

Transfer RNA for leucine (trnL) region was the most commonly targeted marker for plant DNA amplification in herbivores (64 manuscripts), followed by the Internal Transcribed Spacer region (ITS) nuclear region (26 manuscripts) and the Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) region (21 manuscripts). The Maturase K (matK) and NADH-plastoquinone oxidoreductase subunit J (ndhJ) regions were reported in 3 and 2 manuscripts, respectively. Other regions, such as 12S ribosomal RNA (12S rRNA), 16S rRNA, 18S rRNA, psbA-trnH, rpoC1, and atpF/H, were each mentioned in a single manuscript. In terms of platforms used, the Illumina platform was the most widely employed for diet metabarcoding, accounting for 72 out of 94 articles across all 74 (Figure 2b). Examples of Illumina platforms utilized in previous studies include Miseq, Hiseq, iSeq, and Novaseq. Other platforms utilized included Roche: 454 GS FLX (6 manuscripts), Ion Torrent PGMTM system, and ABI: 3130, 3730 (each with 3 manuscripts), while the remaining manuscripts did not specify the platforms used. Furthermore, six different database sources were employed for plant identification in herbivore diet metabarcoding (Figure 2c), either individually or as a compilation of sources. The National Center for Biotechnology Information (NCBI) had the highest number of manuscripts with 48, followed by EMBL (19 manuscripts), The Barcode of Life Data System (BOLD) (7 manuscripts), Consortium for The Barcode of Life (CBOL), and United States Department of Agriculture (USDA) each recorded in 2 manuscripts, and Consortium of Pacific Northwest Herbaria (CPNWH) with a single manuscript. However, 18 manuscripts reported customizing their own databases through plant DNA sequencing, and 4 manuscripts did not specify the database source.

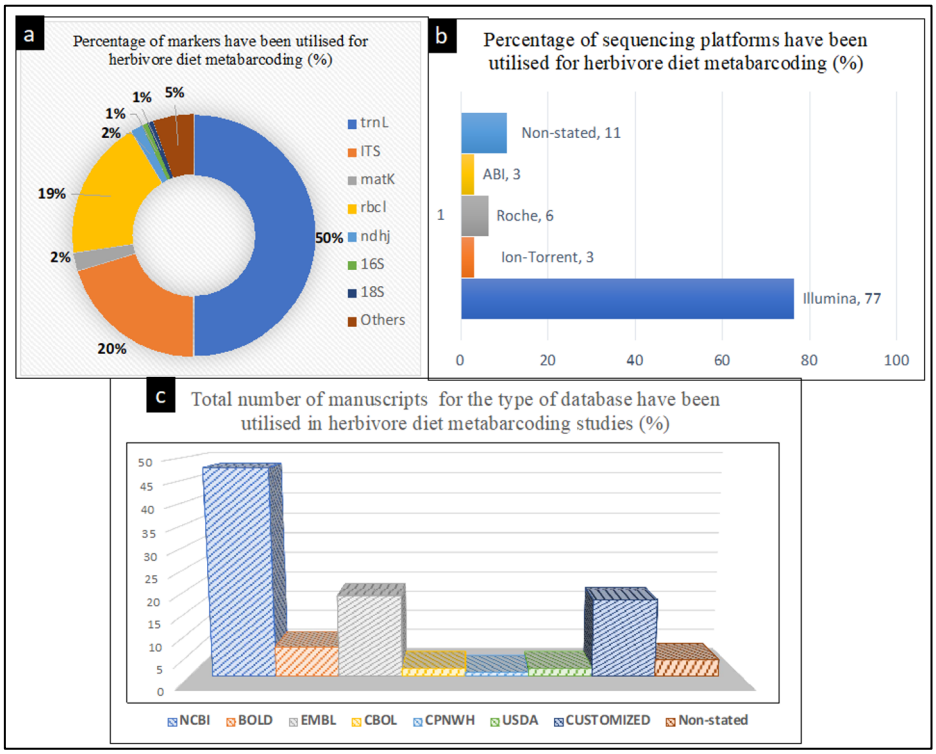

Fig. 2. a) Percentage of markers utilised in herbivore diet metabarcoding (%), **b)** Percentage of sequencing platforms utilised in herbivore diet metabarcoding (%), **c)** Total number of markers utilised in herbivore diet metabarcoding

3.1.2 Omnivore

In the case of omnivore metabarcoding techniques, there were eight primary targeted DNA loci (Figure 3a). There are few combinations of markers to amplify both plant and animal materials in an omnivore diet. The heatmap in Figure 4 shows the frequency of usage for the combination of a few DNA markers in a previous study for discovering the diet of omnivores. The combination of "trnL+ 16S rRNA" was commonly used in previous omnivore metabarcoding diets. Overall, the most widely targeted marker for plant DNA amplification in omnivores was the trnL region (41 manuscripts) specifically for plant detection and the Cytochrome c Oxidase Subunit I (COI) region with 35 manuscripts for animal detection. Then, followed by the usage of rbcL (28 manuscripts), 16S rRNA (25 manuscripts), ITS (12 manuscripts), 18S rRNA(12 manuscripts), 12S rRNA (7 manuscripts), matK (2 manuscripts) and psbA-trnH (with the single manuscript). The Illumina platform also dominates the platform usage for omnivore diet metabarcoding, accounting for 59 out of 66 publications (Figure 3b). Other platforms used included Roche: GS FLX, GS Junior (utilised in 2 studies), ABI: 3730XL and Ion-Torrent (each platform listed in a single manuscript respectively), with the remaining manuscripts not specifying the platforms used. Furthermore, either individually or as a compilation of sources, seven separate database sources were used for the identification in herbivores (Figure 3c). With 41 manuscripts, NCBI has the most, followed by BOLD (10 manuscripts) and EMBL (10 manuscripts), CBOL (2 manuscripts), SILVA: V123, V132 (2 manuscripts), USDA and CPNWH (each with a single manuscript). However, four manuscripts stated that they customized their databases using DNA sequencing, while five manuscripts did not mention the database source.

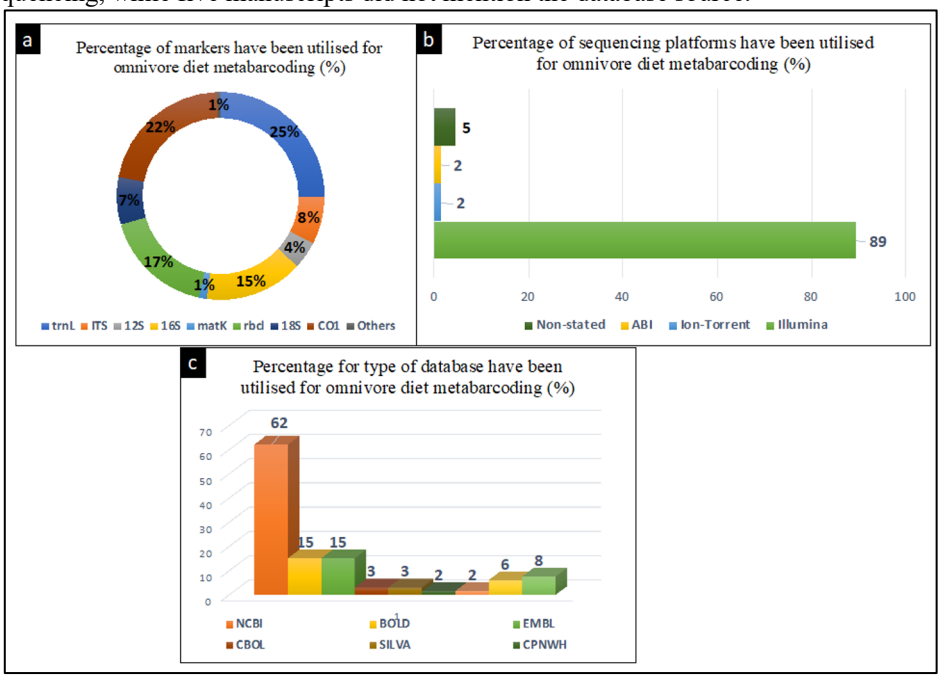

Fig. 3. a) Percentage of markers utilised in omnivore diet metabarcoding, **b)** Percentage of sequencing platforms utilised in omnivore diet metabarcoding, **c)** Total number of markers utilised in omnivore diet metabarcoding

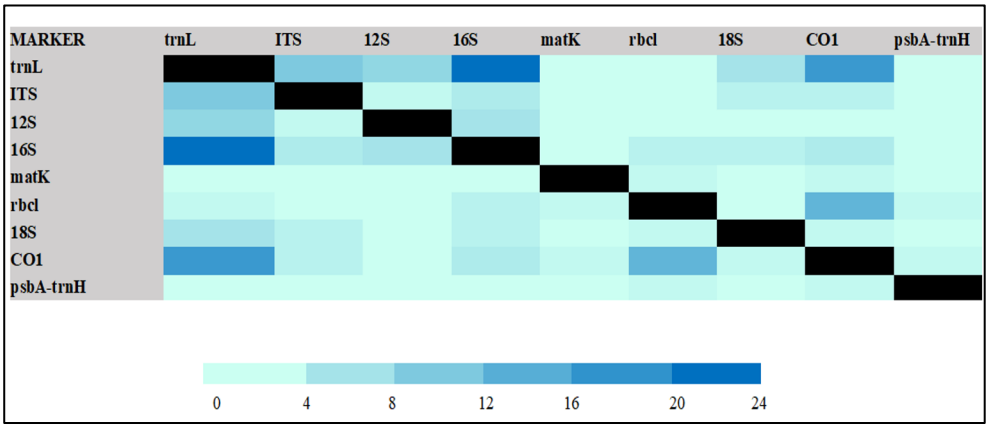

Fig. 4. Heatmap showing the frequency for employment of markers combination in omnivore's diet metabarcoding. The darker shade of blue shows a higher frequency of utilisation.

3.2 DISCUSSION

3.2.1 Finding the Optimal Barcoding Marker

The choice of markers depends on several factors, including the target species, the type of DNA samples collected, and the availability of reference databases [21]. In the context of herbivores, three specific barcode markers find extensive application in metabarcoding dietary analysis. The trnl and rbcL regions have become effective plant barcoding markers due to their ability to provide taxonomic resolution at the genus or family level [22-26]. Nevertheless, the trnL marker has been reported to be surpassed by rbcL due to rbcL's larger fragment size, which comes at the cost of reduced discriminatory power at the species level [22]. ITS nuclear region has been introduced as an alternative barcode, the second most commonly used marker for herbivores in this study. According to Chen et al. [27], this novel region can accurately identify taxa at the species and genus levels with an accuracy of approximately 91.5% and 99.8%, respectively. Yet, ITS sequences may be less represented in public databases when compared to trnL [28], which possesses a more extensive presence of sequence data, such as in GenBank. This availability of sequence data makes it easier to identify plant species or taxonomic units within herbivore food compositions, thereby improving the reliability of dietary analysis results.

The mitochondrial COI region has a high degree of conservation across the animal kingdom [29]. It has an extensive reference database for DNA metabarcoding [30], rendering it a universal marker for animal DNA barcoding. However, omnivores with diverse diets encompassing a wide range of plant and animal species can pose challenges in accurately identifying food consumed. This limitation becomes evident in COI barcoding, particularly for plant species. Ribosomal RNAs (12S, 16S, and 18S) are recommended alternatives for omnivore diet metabarcoding due to their high conservation and slower evolutionary rates compared to the entire mitochondrial genome. They play a vital role in offering species-specific signatures [31, 32]. In omnivore dietary analysis, 16S rRNA is increasingly favored over COI as it is a more conserved locus across taxonomic ranks. The employed primers can be more universally applicable while maintaining similar taxonomic resolution [33]. However, the 12S region may struggle to differentiate closely related animal species due to its highly conserved region, making it less flexible for detecting closely related species [34]. As for the 18S region, the database's reference shortage still hinders the taxonomic

identification for many species [35]. The trnL and the rbcL regions remain as dominant plant DNA markers in detecting the plant-based diet in omnivores.

Current researchers favour universal barcodes that are highly conserved DNA regions for amplifying a broad range of taxa [36]. These universal primers' adaptability is useful when working with unknown or diverse samples. Recognising that no universal primer can capture all conceivable genomic variants inside the targeted region is critical. Modifications may be required in some cases to amplify specific groups of organisms. This review shows that a limited set of primer pairs has predominantly featured for the common regions used in diet metabarcoding as seen in Table 1. Some previous studies had revealed the usage of multimakers to increase the efficiency of exploring these animals' diet by capturing broader range of dietary items and reducing the potential bias and limitation when using a single marker [30, 33]. However, this approach was not recommended as using different sets of markers is time and cost-ineffective.

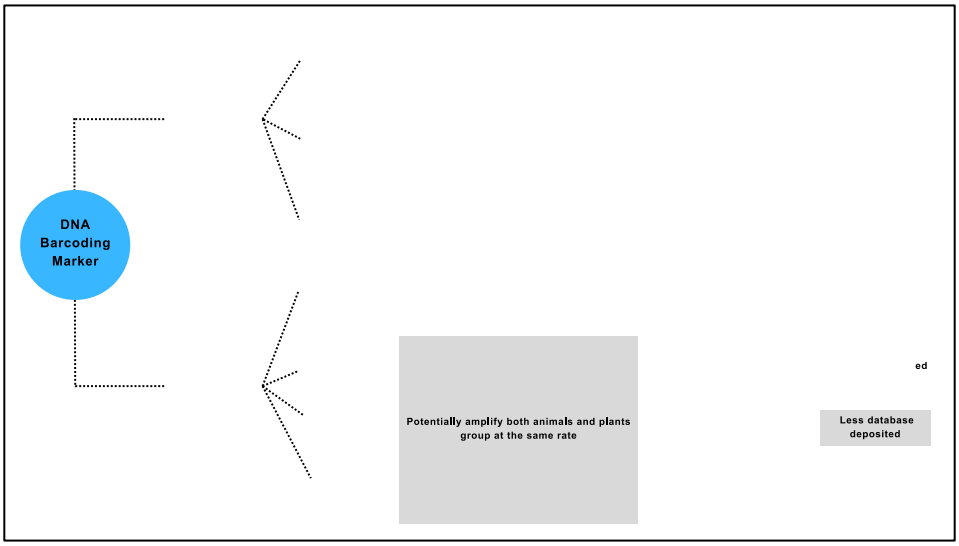

Fig. 5. Summary of criteria for targeted DNA barcoding region in diet metabarcoding studies

Table 1. List of primer pairs that were cited the most in the diet metabarcoding research

Region	Primer name	Sequence (5'-3')	Reference
trnL	*trn*L(UAA)g	GGGCAATCCTGAGCCAA	[37]
	*trn*L(UAA)h	CCATTGAGTCTCTGCACCTATC	[37]
ITS	ITS1-F	CTTGGTCATTTAGAGGAAGTAA	[38]
	ITS4-B	CAGGAGACTTGTACACGGTCCAG	[38]
rbcL	rbcL-aF	ATGTCACCACAAACAGAGACTAAAGC	[39]
	rbcL-aR	GTAAAATCAAGTCCACCGCG	[40]
16S rRNA	515F	GTGCCAGCMGCCGCGGTAA	[41]
	806R	GGACTACHVGGGTWTCTAAT	[41]
COI	LCOI490	GGTCAACAAATCATAAAGATATTGG	[42]
	HCO218	TAAACTTCAGGGTGACCAAAAAATCA	[42]

3.2.2 Selecting the Sequencing Platform

Illumina stands out as a top choice among sequencing platforms in both herbivore and omnivore diet metabarcoding. Illumina is widely recognized as a leading next-generation sequencing (NGS) platform in genomics and molecular biology [43]. It has a reputation for its early adoption of NGS technology, offering high accuracy and minimal errors crucial for precise and dependable data. Other platforms gaining attention in diet metabarcoding include Roche, Ion Torrent, and ABI. These three platforms applied different sequencing mechanisms (Illumina: sequencing by synthesis/ Roche: pyrosequencing/ ABI: sequencing by ligation), and each platform has a different level of read length, producing varied size, output, and level of accuracy. The comparison conducted by [44] found that Roche and ABI sequencing platforms outperformed Illumina in terms of accuracy, achieving accuracy rates of 99.9% and 99.4%, respectively, while Illumina had an accuracy rate of 98%. However, Illumina excelled in generating larger output data, producing approximately 600 Gb of data, and being cost-effective at a rate of 0.07 USD per million bases. In contrast, Roche and ABI had more limited data outputs of 0.7 Gb and 12.0 Gb, respectively, and incurred higher costs at 10 USD per million bases for Roche and 0.13 USD for ABI. Among few Illumina sequencing systems, the Miseq system dominated the platform used in dietary studies, followed by Hiseq, Miniseq, Novaseq, and iSeq (Figure 6), each with a varied output data size (Table 2).

Besides, Illumina sequencers are compatible with a broad range of sample preparation and library construction methods well-established in the research community [45]. The Ion Torrent sequencing technology operates by detecting the release of hydrogen ions while incorporating new nucleotides that generate a read length of around 200-600 bp [46], which is used to fill gaps in the assembly produced by other technologies. This feature makes it a cost-effective choice for focused genetic studies but has a lower data output (up to 2 Mb), providing limited results.

Other platforms that have not yet been utilized for herbivore and omnivore diet metabarcoding are Nanopore and Pac Bio. Nanopore works on the principle of minute changes in electric current across the nanopore was massively used in e-DNA metabarcoding as it generates very long reads(>200 Kb) and can produce up to 30 Gb data output [46, 47], which are useful for building new genomes, distinguishing between closely related genetic variants. However, Nanopore is less common in diet metabarcoding because Nanopore sequencing has been seen to face the problem of base-calling accuracy compared to other platforms [48]. The technology may have concerns about increasing accuracy and throughput at this stage, which must be considered for its large-scale commercial use. Besides, this platform is relatively new compared to other pioneer platforms. The protocol for specific dietary analysis has not been developed, and data analysis is more complex, requiring specialized bioinformatics skills such as using pipelines. PacBio sequencing also produces very long reads, around 1000-3000 bp [46, 49], but it's sometimes less favored because it can be more expensive per unit of genetic information [50]. Besides, it can only produce 5-8 GB data per SMRT cell (www.pacific- biosciences.com), which is much less when compared to the other existing technologies. Each sequencing platform has unique characteristics like read length, data output, and error rates. It's essential to consider the bioinformatics skills when choosing a platform since some may require more expertise for data analysis. Additionally, it's wise to check if the platform is stable or if upcoming updates could affect the research.

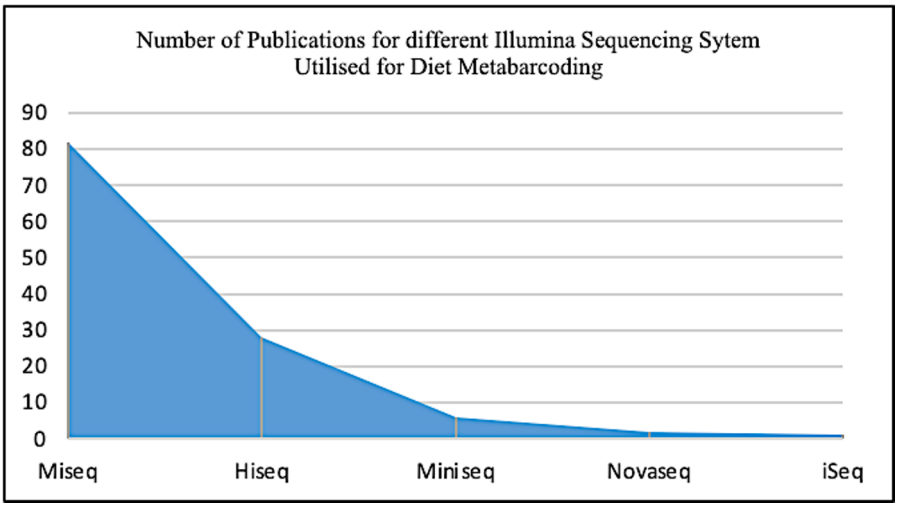

Fig. 6. The number of publications recorded for utilisation of different Illumina Sequencing System

Table 2. The length of reads and output data size for different Illumina Sequencing Systems (Illumina Inc., USA)

Illumina sequencing system	Read Length	Output data size
Miseq	typically ranging from 2 x 150 up to 2 x 300 bp paired-end reads.	15 GB per run
Hiseq	Ranging from shorter reads to long reads (2 x 50 bp to 2 x 250 bp or more, depending on the specific model).	Range from hundreds of GB to several TB per run, depending on the model and configuration
Miniseq	2 x 150 bp paired-end reads	7.5 Gb per run
Novaseq	Extensive flexibility in read length, including options for very long reads (2 x 150 bp to 2 x 300 bp, 2 x 250 bp and more).	Extremely high, ranging from several Tb to over 6 Tb per run depending on the model and configuration
iSeq	Typically provides shorter reads with options like 2 x 150 bp paired-end reads.	Up to 2 x 150 bp paired-end reads, suitable for smaller-scale applications

3.2.3 Choosing Database for Identification

The analysis indicates that the Genbank database deposited in NCBI plays a dominant role in dietary metabarcoding research for both herbivores and omnivores. NCBI's comprehensive database encompasses a wide array of biological information, including genomic sequences, gene data, protein information, and an extensive collection of scientific articles. This resource offers access to diverse DNA sequences from different barcodes of nuclear, mitochondrial and chloroplast DNA. According to Schoch et al. [51], the NCBI taxonomy curators are responsible for detailing major taxonomic groups, elucidating distinctive terms within NCBI taxonomy, recognizing external resources, and recording revisions to tools and additional resources that prove the validity of data sequence provided by this reference database. However, a study has reported incorrect species identification from NCBI after phylogenetic analysis was conducted, which involved Hypophthalmichthys sp [52], and other database sources were used to re-confirm the species identification. Other databases, such as The

Barcode of Life Data System (BOLD) and The European Molecular Biology Laboratory (EMBL), are valuable dietary analysis references. BOLD is particularly known for its utility in exploring species diversity and conservation and its relevance to taxonomic and ecological studies. It serves as a repository for DNA barcode sequences, complete with associated metadata from various species. BOLD rigorously enforces standardized data collection and documentation protocols, ensuring uniformity and data quality across barcode data [53]. Yet, the NCBI and BOLD databases are widely used for plant DNA identification compared to EMBL since they have similar performance levels for plants, achieving approximately 81% and 57%, respectively [54, 55].

EMBL seems to be the second top database used in herbivore dietary analysis. EMBL, part of the International Nucleotide Sequence Database Collaboration (INSDC), offers an extensive repository for nucleotide sequences, genome assemblies and a broad spectrum of molecular biology data, making it valuable for studying plant genomics [56]. It is a useful resource for researchers working with DNA and RNA sequences, genome assembly, functional genomics, and those needing a comprehensive sequence database [57]. Besides, the SILVA database is commonly used in detecting microbes [58, 59] yet is still valuable for detecting other eukaryotes, which makes it less utilized in diet metabarcoding studies. The choice of a database depends on the specific goals of the research. Researchers often use a combination of these databases, depending on the nature of their research questions and the specific data they require.

Nevertheless, the most effective database is often a customized database crafted by carefully selecting and compiling reference sequences from various organisms likely to be part of the diet of the target species [60, 61]. These reference sequences can encompass DNA barcodes or markers specific to different taxonomic groups such as plants, animals, and fungi. By tailoring the database to the particular ecosystem, region, or species under study, researchers enhance the accuracy and precision of their dietary analyses, especially for local species.

4 Conclusion

This study reveals that the trnL gene exhibited greater taxonomic precision when it came to herbivorous species. In contrast, the combination between trnL and 16S rRNA markers was mostly recorded in omnivorous dietary analysis, even though the COI region dominates for animal-based diet identification in the omnivore group. The choice of DNA markers for dietary analysis depends on several crucial factors, including the target species and the availability of reference databases. For platform, the Illumina platform, especially the Miseq system, stands as the most commonly utilised method in both groups of animal's diet metabarcoding, followed by Roche, Ion-Torrent, and ABI platforms. Illumina outperforms other platforms due to cost-effective and is known for its high-throughput sequencing. Yet, the choice of sequencing platform should align with research goals, and the unique characteristics of each platform, such as read length and data output, need to be considered. Besides, the key reference database being utilised in diet metabarcoding for herbivores and omnivore is the National Centre for Biotechnology Information (NCBI). However, a customized database is proposed as the best reference source for diet metabarcoding, especially for local species. Researchers often combine multiple databases or create a customised database to enhance the accuracy of dietary analysis, especially for local species or specific ecosystems. Overall, the field of dietary metabarcoding continues to evolve with a growing emphasis on choosing the right combination of markers, sequencing platforms, and databases to ensure accurate and reliable results for dietary analysis.

This project is funded by the Ministry of Natural Resources and Environmental Sustainability (NRES) under the National Conservation Trust Fund for Natural Resources (NCTF) (KeTSA (S) 600-2/1/48/6 Jld. 2 (11))-(UTHM-K449) and UTHM-REGG-Q194-2022 postgraduate grant by Universiti Tun Hussein Onn Malaysia. We are deeply indebted to Faculty of Applied Science and Technology (FAST), Universiti Tun Hussein Onn Malaysia, and Department of Wildlife and National Parks (DWNP) for providing the necessary funding, facilities and assistance.

References

1. K. Fernandes, M. van der Heyde, M. Bunce, K. Dixon, R.J. Harris, G. Wardell-Johnson, P.G. Nevill, Restor. Ecol., **26**, 6 (2018)
2. M. Liu, L.J. Clarke, S.C. Baker, G.J. Jordan, C.P. Burridge, Ecol. Entomol., **45**, 3 (2020)
3. J.M. Nielsen, E.L. Clare, B. Hayden, M.T. Brett, P. Kratina, Methods Ecol. Evol., **9,** 2 (2018)
4. H. Ando, H. Mukai, T. Komura, T. Dewi, M. Ando, Y. Isagi, Environ. DNA, **2**, 4 (2020)
5. B. Eitzinger, B.C. Rall, M. Traugott, S. Scheu, Oikos, **127**, 7 (2018)
6. M.L. Forister, V. Novotny, A.K. Panorska, L. Baje, Y. Basset, P.T. Butterill, L. Cizek, P.D. Coley, F. Dem, I.R. Diniz, P. Drozd, M.S. Fox, A.E. Glassmire, R.F. Hazen, J. Hrček, J.P. Jahner, O. Kaman, T.J. Kozubowski, T.A. Kursar, O.T. Lewis, J. Lill, R.J. Marquis, S.E. Miller, H.C. Morais, M. Murakami, H. Nickel, N.A. Pardikes, R.E, Ricklefs, M.S. Singer, A.M. Smilanich, J.O. Stireman, S. Villamarin-Cortez, S. Vodka, M. Volf, D.L. Wagner, T. Walla, G.D. Weiblen, L.A. Dyer, PNAS, **112**, 2 (2015)
7. S. Rytkönen, E.J. Vesterinen, C. Westerduin, T. Leviäkangas, E. Vatka, M. Mutanen, P. Valimaki, M. Hukkanen, M. Suokas, M. Orell, web. Ecol. Evol., **9**, 1 (2019)
8. L.R. Mehrkam, The cognitive abilities of wild animals, (Zoo Animal Learning and Training, United States, 2020)
9. N. Owen-Smith, P. Novellie, Am. Nat, **119**, 2 (1982)
10. D.E. Burkepile, C.E. Burns, C.J. Tambling, E. Amendola, G.M. Buis, N. Govender, V. Nelson, D.I. Thompson, A.D. Zinn, M.D. Smith, Ecosphere, **4**, 11 (2013)
11. L.D. V. Arregoitia, G. D'Elía, Mamm. Rev., **51**, 1 (2021)
12. K. E. Stoner, R. M. Timm, Seasonally dry tropical forest mammals: adaptations and seasonal patterns (Island Press, Washington, 2011)
13. C.H.A. Van Leeuwen, N. Villar, I.M. Sagrera, A.J. Green, E.S. Bakker, M.B. Soons, M. Galetti, P.A. Jansen, B.A. Nolet, L. SantamaríA, **9**, e09254 (2022)
14. W.O.C. Symondson, Mol. Ecol., **11**, 4 (2002)
15. L.P. da Silva, V.A. Mata, P.B. Lopes, P. Pereira, S.N. Jarman, R.J. Lopes, P. Beja, Mol. Ecol. Resour., **19**, 6 (2019)
16. A.T. Reese, T.R. Kartzinel, B.L. Petrone, P.J. Turnbaugh, R.M. Pringle, L.A. David, MSystems, **4**, 5 (2019)
17. M. Staats, A.J. Arulandhu, B. Gravendeel, A. Holst-Jensen, I. Scholtens, T. Peelen, T.W. Prins, E. Kok, Anal. Bioanal. Chem., **408**, 17 (2016)
18. B.E. Deagle, A.C. Thomas, J.C. McInnes, L.J. Clarke, E.J. Vesterinen, E.L. Clare, T.R. Kartzinel, J.P. Eveson, Mol. Ecol., **28,** 2 (2019)
19. P. Somervuo, D.W. Yu, C.C. Xu, Y. Ji, J. Hultman, H. Wirta, O. Ovaskainen, Methods Ecol. Evol., **8**, 4 (2017)

20. G. Yao, W. Ma, X. Huang, Q. Jia, J. Shen, Y. Chang, H. Ouyang, J. He, J Anal. Methods. Chem., (2020)
21. Q.L. Liu, J. Li, M.J. Wingfield, T.A. Duong, B.D. Wingfield, P.W. Crous, S.F. Chen, Stud. Mycol., **97**, 1 (2020)
22. E.K. Mallott, P.A. Garber, R.S. Malhi, PLoS One, **13**, 6 (2018)
23. M. Ismail, A. Ahmad, M. Nadeem, M.A. Javed, S.H. Khan, I. Khawaish, A.A. Sthanadar, S.H. Qari, S.M. Alghanem, K.A. Khan, M.F. Khan, S. Qamer, Saudi. J. Biol. Sci., **27**, 12 (2020)
24. N.A.F. Abdullah-Fauzi, K.V. Karuppannan, N.H.S. Mohd-Radzi, M. Gani, A.R. Mohd-Ridwan, N. Othman, H. Haris, N.H. Sariyati, N.R. Aifat, M.A.B. Abdul-Latiff, M.F.A. Abdul-Razak, B.M. Md-Zain, Zool. Stud., **61** (2022)
25. N.H.S. Mohd-Radzi, K.V. Karuppannan, N.A.F. Abdullah-Fauzi, A.R. Mohd-Ridwan, N. Othman, M.A.B. Abdul-Latiff, M. Gani, M.F.A. Abdul-Razak, B.M. Md-Zain, Biodivers. Data, J **10** (2022)
26. N.A. Osman, M.A.B. Abdul-Latiff, A.R. Mohd-Ridwan, S. Yaakop, S.M. Nor, B.M. Md-Zain, Animals, **10**, 12 (2020)
27. S. Chen, H. Yao, J. Han, C. Liu, J. Song, L. Shi, Y. Zhu, X. Ma, T. Gao, X. Pang, K. Luo, Y. Li, X. Jia, Y. Lin, C. Leon, PLoS One, **5**, 1 (2010)
28. N.A. Fahner, S. Shokralla, D.J. Baird, PLoS One, **11**, 6 (2016)
29. M. Guo, C. Yuan, L. Tao, Y. Cai, W. Zhang, Conserv. Genet. Resour., **14**, 4 (2022)
30. R.A. Collins, J. Bakker, O.S. Wangensteen, A.Z. Soto, L. Corrigan, D.W. Sims, M.J. Genner, S. Mariani, Methods. Ecol. Evol., **10**, 11 (2019)
31. S. Bandiera, S. Rüberg, M. Girard, N. Cagnard, S. Hanein, D. Chrétien, A. Munnich, S. Lyonnet, A. Henrion-Caude, PLoS One, **6**, 6 (2011)
32. M. Vences, M.L. Lyra, R.G.B. Perl, M.C. Bletz, D. Stanković, C.M. Lopes, M. Jarek, S. Bhuju, R. Geffers, C.F.B. Haddad, S. Steinfartz, Conserv. Genet. Resour., **8**, (2016)
33. E. Valsecchi, J. Bylemans, S.J. Goodman, R. Lombardi, I. Carr, L. Castellano, A. Galimberti, P. Galli, Environ. DNA, **2**, 4 (2020)
34. Valentini, P. Taberlet, C. Miaud, R. Civade, J. Herder, P.F. Thomsen, E. Bellemain, A. Besnard, E. Coissac, F. Boyer, C. Gaboriaud, P. Jean, N. Poulet, N. Roset, G.H. Copp, P. geniez, D. Pont, C. Argillier, J. Baudoin, T. Peroux, A.J. Crivelli, A. Olivier, M. Acqueberge, M.L. Brun, P.R. Moller, E. Willerslev, T. Dejean, Mol. Ecol., **25**, 4 (2016)
35. A.S. Tanabe, S. Nagai, K. Hida, M. Yasuike, A. Fujiwara, Y. Nakamura, Y. Takano, S. Katakura, Mol. Ecol. Resour., **16**, 2 (2016)
36. V.K. Gupta, V. Jindal. J. Integr. Pest. Manag., (2013)
37. P. Taberlet, E. Coissac, F. Pompanon, L. Gielly, C. Miquel, A. Valentini, T. Vermat, G. Corthier, C. Brochmann, E. Willerslev, Nucleic. Acids. Res., **35**, 3 (2007)
38. M. Gardes, T.D. Bruns, Mol. Ecol., **2**, 2 (1993)
39. R.A. Levin, W.L. Wagner, P.C. Hoch, M. Nepokroeff, J.C. Pires, E.A. Zimmer, K.J. Sytsma, Am. J. Bot., **90**, 1 (2003)
40. B. Fofana, L. Harvengt, J.P. Baudoin, P. Du Jardin, Belgian J. Bot., **129**, 2 (1997)
41. J.G. Caporaso, C.L. Lauber, W.A. Walters, D. Berg-Lyons, C.A. Lazupone, P.J. Turnbaugh, N. Fierer, R. Knight, Proc. Natl. Acad. Sci. USA, **108**, 1 (2011)
42. O. Folmer, M. Black, W. Hoeh, R. Lutz, R. Vrijenhoek, Mol. Mar. Biol. Biotechnol., **3**, 5 (1994)

43. B. Meera Krishna, M. A. Khan, S. T. Khan, Next-generation sequencing (NGS) platforms: an exciting era of genome sequence analysis, (Microbial Genomics in Sustainable Agroecosystems, Springer, 2019)
44. J. Hodzic, L. Gurbeta, E. Omanovic-Miklicanin, A. Badnjevic, Med. Arch., **71**, 4 (2017)
45. M.K. Scheible, E.K. Timpano, L.M. Boggs, K.A. Meiklejohn, Int. J. Legal. Med., **135**, 5 (2021)
46. A. K. Gupta, U. D. Gupta, Next generation sequencing and its applications (Animal biotechnology, Academic Press, 2020)
47. K. Yahara, M. Suzuki, A. Hirabayashi, W. Suda, M. Hattori, Y. Suzuki, Nat. Commun., **12**, 1 (2021)
48. L.M. Petersen, I.W. Martin, W.E. Moschetti, C.M. Kershaw, G.J. Tsongalis, J. Clin. Mirobiol., **58**, 1 (2019)
49. A. D. Bailey, Modification detection using Nanopore sequencing, (University of California, Santa Cruz 2015)
50. S. Koren, A.M. Philippy, Curr. Opin. Microbiol., **23** (2015)
51. C.L. Schoch, S. Ciufo, M. Domrachev, C.L. Hotton, S. Kannan, R. Khovanskaya, D. Leipe, R. Mcveigh, K. O'Neill, B. Robbertse, S. Sharma, V. Soussov, J.P. Sullivan, L. Sun, S. Turrner, I. Karsch-Mizrachi, Database, **2020**, baa062 (2020)
52. H. Liu, S. Fu, S. Zhang, M. Ding, A. Wang, Aquaculture, **528**, 736673 (2020)
53. L.F. Clark, Food Policy, **54**, (2015)
54. D. Shadrin, O. Valuyskikh, V. Kanev, Acta. Biol. Sib., **6** (2020)
55. K.A. Meiklejohn, N. Damaso, J.M. Robertson, PLoS One, **14**, 6 e0217084 (2019)
56. H. Susanto, Appl. Chem. Eng., **4**, 181 (2017)
57. G. Cantelli, A. Bateman, C. Brooksbank, A.I. Petrov, R.S. Malik-Sheriff, M. Ide-Smith, H. Hermjakob, P. Flicek, R. Apweiler, E. Birney, J. McEntyre, Nucleic. Acids. Res., **50**, D11-D19 (2022)
58. J. Lu, S.L. Salzberg. Microbiome, **8**, 1 (2020)
59. F.O. Glöckner, Biodivers. Inf. Sci. Stand, **3**, e36125 (2019)
60. N. Othman, H. Haris, Z. Fatin, M. F. Najmuddin, N. H. Sariyati, B. M. Md-Zain. M. A. B. Abdul-Latiff, IOP Publishing, **736**(1), 012054 (2021)
61. N. Othman, K. Munian, H. Haris, F.F. Ramli, N.H. Sariyati, M.F. Najmuddin, M.A.B. Abdul-Latiff, Sains Malays., **52**,1 (2023)