

RESEARCH ARTICLE

trnL outperforms *rbcl* as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (*Cebus capucinus*, Primates)

Elizabeth K. Mallott^{1*}, Paul A. Garber², Ripan S. Malhi^{2,3}

1 Department of Anthropology, Northwestern University, Evanston, IL, United States of America, **2** Department of Anthropology, University of Illinois at Urbana-Champaign, Urbana, IL, United States of America, **3** Carl R Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, United States of America

* elizabeth.mallott@northwestern.edu



OPEN ACCESS

Citation: Mallott EK, Garber PA, Malhi RS (2018) *trnL* outperforms *rbcl* as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (*Cebus capucinus*, Primates). PLoS ONE 13(6): e0199556. <https://doi.org/10.1371/journal.pone.0199556>

Editor: Hideyuki Doi, University of Hyogo, JAPAN

Received: January 18, 2018

Accepted: June 8, 2018

Published: June 26, 2018

Copyright: © 2018 Mallott et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequence data is available in the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) BioProject ID PRJNA299397, and the *ecotag* output for both the *rbcl* and *trnL* sequences can be found in the Supporting Information files.

Funding: This study was supported with funds from the UIUC Department of Anthropology, the Beckman Institute at UIUC, the Center for Latin American Studies at UIUC, the NSF GRFP program, and the American Philosophical Society (Lewis and

Abstract

DNA metabarcoding is a powerful tool for assessing the diets of wild animals, but there is no clear consensus on which proposed plant barcoding marker is most suitable for dietary analysis. This study compares two DNA plant barcoding markers that are commonly used for dietary analyses from degraded DNA, *rbcl* and *trnL*, to detailed dietary observations of wild white-faced capuchins (*Cebus capucinus*). Observational dietary data and fecal samples (n = 170) were collected for one year from a group of individually recognizable monkeys at La Suerte Biological Field Station, Costa Rica. DNA was extracted and portions of the *rbcl* and *trnL* chloroplast were amplified and sequenced on the Illumina MiSeq platform. Sequences were analyzed using *obitools*. Of the two barcoding markers tested, *trnL* yielded greater numbers of sequences with equal sequencing effort, higher resolution taxonomic identifications (albeit with a larger reference database), and identified a greater number of families also found in the observed diet. There was no relationship between observed capuchin feeding behavior and dietary composition based on either sequence occurrence or relative abundance of sequences using *rbcl* as a marker. However, dietary composition based on the relative abundance of *trnL* sequences was significantly positively associated with the observed percentage of feeding and foraging time capuchins' spent on each plant species. Additionally, in 35% of cases, the relative abundance of *trnL* sequences assigned to particular plant families in fecal samples was highly positively correlated with time spent consuming plants from those same families. Our results indicate that *trnL* is a more robust DNA metabarcoding marker for plant dietary analysis and may potentially be used to quantitatively assess differences in diet within or between species.

Clark Fund for Exploration and Field Research) awarded to EKM, and funds awarded to PAG and RSM by the University of Illinois Urbana-Champaign. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

DNA metabarcoding is quickly becoming a powerful tool in ecology for understanding trophic interactions and non-invasively assessing the diets of wild animals, particularly with the increasing availability of high-throughput sequencing technologies [1,2]. However, DNA barcoding marker genes must be chosen carefully to provide meaningful results [3]. The choice of marker depends on the biology of the study species, the questions being asked, the quality of the DNA sample, and completeness of reference databases [2]. Target regions must be variable enough to discriminate between closely related animal and plant species, yet flanking regions should be conserved enough to be less variable within a single species. Additionally, primer sets should readily amplify a range of species within a large taxonomic group. Finally, the target region must be long enough to yield the desired taxonomic resolution. Even though there is a consensus for some taxonomic groups as to which standardized DNA region should be used for metabarcoding, including dietary analysis (e.g., *COI* for invertebrates) [4,5], multiple barcoding target regions have been proposed for other taxonomic groups, such as plants [1,3,6,7].

An additional important consideration for dietary analysis using DNA metabarcoding is the ability of a primer to amplify often degraded and low copy number DNA from fecal samples. Mini-barcodes that target regions of DNA that are generally 150bp or smaller allow for the amplification of degraded DNA [8,9]. While these smaller barcoding regions are preferable for degraded samples [3,8,10,11], smaller barcoding regions result in less specific taxonomic identifications [12]. Additionally, in plants, using single locus barcodes can lead to less reliable taxonomic identifications due to high rates of hybridization, introgression, and polyploidy in land plants [13].

Two single-locus regions, *trnL* and *rbcL*, have been proposed for dietary analysis using DNA metabarcoding in herbivores and omnivores [1,3,14]. In particular, the *trnL*-g and *trnL*-h primer set used to amplify the P6 loop of *trnL* (UAA) [1,15–23] and the *rbcL*Z1 and *rbcL*19b mini-barcoding primers used to amplify a small region of *rbcL* [14,23–25] have been widely used. A few studies have directly compared the performance of *rbcL* and *trnL* as barcoding regions [23–26]. Results indicated that *trnL* had higher amplification success and a larger sequence divergence range in a study of wetland plants [25], while *rbcL* had a more complete publicly available reference database [26]. Both regions generally were only able to provide taxonomic resolution at the level of genus or family [9]. In addition, rates of sequence divergence, and, therefore, taxonomic resolution of mini-barcodes in these two regions may differ depending on the plant family or families of interest in a particular study [25].

Relatively less is known, however, about how dietary results from each of these barcoding regions compare to other methods of assessing diet, particularly how molecular results compare to fine-scale dietary observations over long time periods. In one study where sika deer (*Cervus nippon*) were fed a known diet, DNA metabarcoding using *trnL* was able to reliably detect all of the foods present in the diet when a local reference database was used [20]. However, taxonomic identifications were limited to the genus or family level when using a global database in the same study [20]. Similarly, in a study of red-shanked douc langurs (*Pygathrix nemaeus*) fed a known diet, DNA metabarcoding using *trnL* was only able to identify genera known to be in the diet in 31.3–37.5% of cases using a global reference database [27]. Studies comparing microhistological analysis of fecal samples or macroscopic analysis of stomach contents with DNA metabarcoding, using either *trnL* or *rbcL* as markers, found varying amounts of overlap between taxa identified by the two methods [15,18,23,28–30]. Many of these studies also showed that DNA metabarcoding provided much greater taxonomic resolution than microhistology [15,18,28,29]. Few studies have compared results from one of the DNA

metabarcoding markers with methods of direct observation of diet in a wild mammal. In a study of woodland caribou (*Rangifer tarandus caribou*), Newmaster *et al.* (2013) found that results from DNA metabarcoding using *trnL* and a global reference database and results from animal-borne video recording were highly correlated. Similarly, Hargrave (2015) found that DNA metabarcoding using *trnL* and a global reference database supplemented with local plant species sequences identified all but one taxon of plants that wild sable antelope (*Hippotragus niger*) were observed to have consumed. In a study of black and white colobus monkeys (*Colobus guereza*), Bradley *et al.* (2007) found that DNA metabarcoding using *rbcL* and a global reference database identified plants known to be part of an individual's diet during detailed behavioral observations taken during the previous two days. However, the study failed to detect some plants that an individual was observed eating and identified plants that were not recorded as part of an individual's diet but may have been consumed [24]. None of these studies have compared the precision of both *rbcL* and *trnL* with direct detailed dietary observations collected over more than six months.

To fill this gap, here we evaluate the precision and taxonomic resolution of DNA metabarcoding-based dietary analysis using both *rbcL* and *trnL*. We compare molecularly-determined diet with detailed individual-level dietary observations of a group of wild white-faced capuchins (*Cebus capucinus*) over a year-long period. We also assess whether the relative abundance of *rbcL* or *trnL* reads is correlated with the percentage of time spent feeding or foraging on food items.

Materials and methods

Study site and population

The study was conducted from January 2013 through January 2014 at La Suerte Biological Field Station in northeastern Costa Rica (10.445N, 83.784W). La Suerte Biological Field Station includes 170 ha of advanced secondary wet tropical forest and 130 ha of early secondary growth and regenerating pasture [31]. Annual rainfall during the study period was 3116 mm. A group of individually recognizable white-faced capuchins (*Cebus capucinus*) containing 21–22 individuals during the study period was observed (a detailed description of the study site and population is found in Mallott *et al.* 2017) [32]. An overview of sample collection and analysis is found in Fig 1 (Fig 1).

Observational data collection

Dietary data were collected during one-hour instantaneous focal animal samples (two-minute interval) of individually recognizable adult males, adult females, and juveniles. Data were collected between 0500 and 1600 hours, and 841 hours of dietary data were collected over 237 days. Activity budget (feeding—actively consuming food items; foraging—manipulating or searching for food items; traveling—movement between or within the crowns of trees which was not for the immediate purpose of obtaining food or engaging in social activity; resting—periods of inactivity; social—engaged in any affiliative or agonistic interaction with at least one other individual) was recorded at each two-minute interval. If an individual was feeding or foraging, diet (ripe fruit, unripe fruit, seeds, flowers, leaves, invertebrates, vertebrates, other) was recorded. Plant food items were identified to species, when possible, and identifications were aided by reference materials in the library at La Suerte Biological Field Station and the knowledge of station staff. Overall, 80% of the plant species consumed by the capuchins were identified to the species level and 86% of plant feeding and foraging records were a known species.

Fecal samples were collected from all individuals throughout the observation study period ($n = 170$) in sterile tubes, fixed in 90% ethanol, and stored at -20°C . The mean number of fecal samples collected per month was 13.1 ± 5.3 (range = 4–21). To minimize environmental

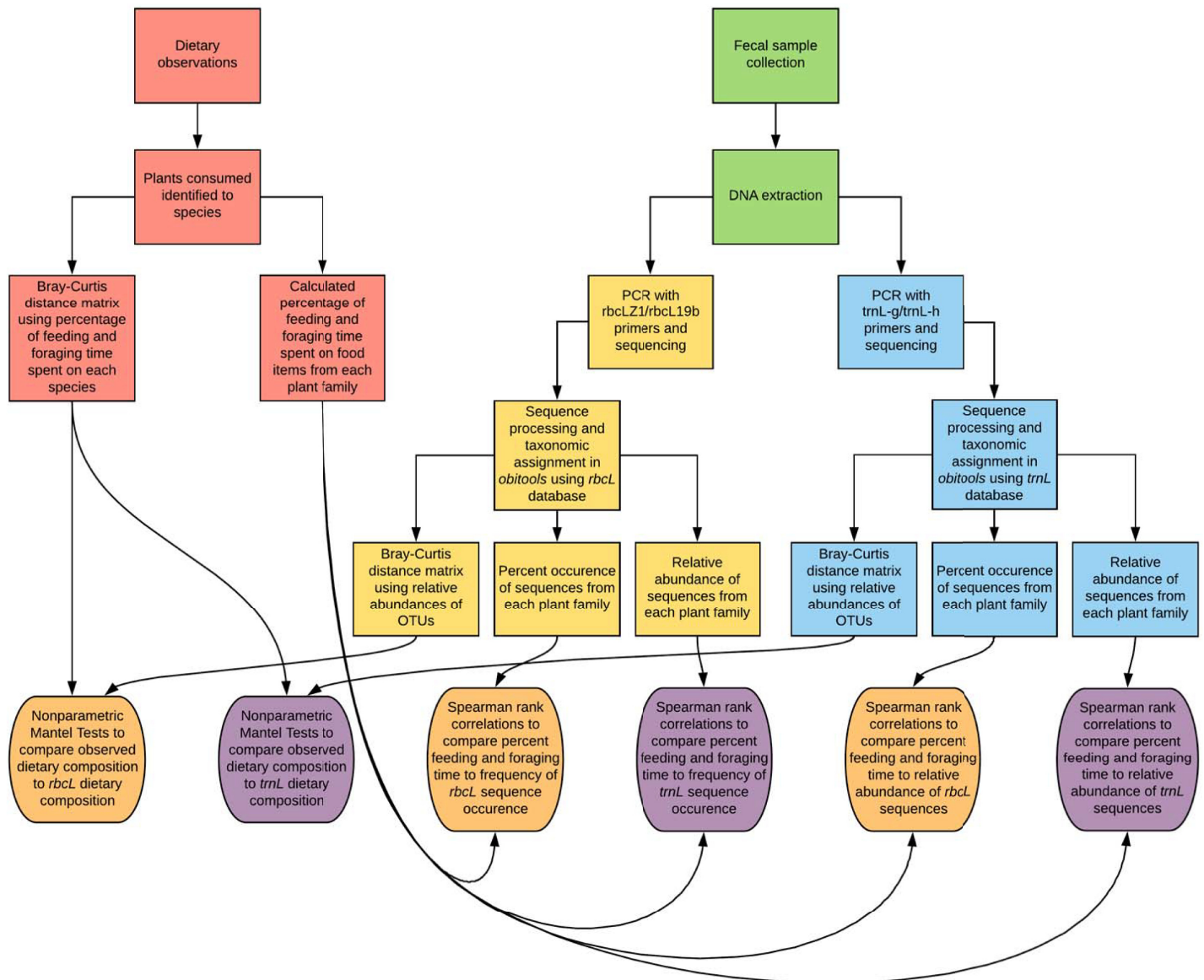


Fig 1. Summary of sample collection and analysis. Flowchart showing sample collection, sample processing for each type of sample, and statistical analyses using the resulting data from each sample type.

<https://doi.org/10.1371/journal.pone.0199556.g001>

contamination, portions of fecal material in direct contact with soil or leaf litter were not collected. CITES export permits were obtained from Sistema Nacional de Areas de Conservación in Costa Rica, and import permits for the United States were obtained from the CDC. Data collection protocols were approved by the University of Illinois IACUC, La Suerte Biological Field Station, the Ministerio de Ambiente, Energía, y Telecomunicaciones in Costa Rica, and the Comisión Nacional para la Gestión de la Biodiversidad in Costa Rica.

Molecular data collection

A QIAamp DNA Stool Mini Kit was used to extract DNA from 170 fecal samples, following the provided protocol for “Isolation of DNA from Stool for Human DNA Analysis” (using the

modifications outlined in Mallott *et al.* 2015) [33]. A 157bp segment of the *rbcL* gene was amplified using the *rbcLZ1* and *rbcL19b* primers [14,24], and a 22-143bp segment of the *trnL* (UAA) intron was amplified using the *trnL-g* and *trnL-h* primers [3]. Negative extraction and PCR controls were used to minimize the risk of contamination. Samples were individually indexed, using the same barcode for both *rbcL* and *trnL* amplicons for each sample, and the amplicons were amplified on a Fluidigm Access Array system at the Roy J. Carver Biotechnology Center at UIUC (for reaction conditions and detailed methods see Mallott *et al.* 2017). Amplicons were sequenced on the Illumina MiSeq v2 platform with paired end 2x300 bp reads. In *obitools*, raw sequences were paired with an alignment score threshold of 20, assigned to samples, de-duplicated, low-count sequences and PCR errors were removed, and the remaining sequences were assigned to taxa using the *ecotag* script and reference databases for both *rbcL* and *trnL* [34]. Reference databases for *rbcL* and *trnL* were created using *ecoPCR* [35] from release 131 of the European Nucleotide Archive. The *rbcL* database contained 1589 unique plant sequences which include 92 orders, 284 families, 881 genera, and 1054 species. The *trnL* database contained 13941 unique plant sequences which include 75 orders, 365 families, 4028 genera, and 8656 species. Tab-delimited sequence read count tables were created in *obitools*, and the resulting files were collapsed at the operational taxonomic unit, genus, and family levels. Relative abundances were then calculated in the *vegan* package [36] of R (r-project.org). Details of the bioinformatics analyses can be found in the supporting information (S1 File). Raw sequence data is available in the Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>), BioProject ID PRJNA299397, and the *ecotag* output for both the *rbcL* and *trnL* sequences can be found in the supporting information (S1 and S2 Tables).

Statistical methods

Nonparametric Mantel tests were used to compare Bray-Curtis dissimilarity matrices of monthly individual observed and molecular diets. Relative abundances of sequences were averaged across the month for individuals for whom we had multiple molecular samples in the same month. To assess relationships between diet analysis methods at the group-level, instead of individual-level, spearman rank correlations were used to test for a positive association between the percentage of the capuchin group's feeding and foraging time spent consuming plants from a specific family, and the frequency of occurrence of sequences assigned to that family (percentage of fecal samples containing sequences assigned to a family) and average relative abundance of sequences assigned to that same family within a given month for both *rbcL* and *trnL* sequencing results.

Results

Observed diet

During the study period, 47.8% of observed white-faced capuchin feeding and foraging time was spent on fruit, 49.8% on invertebrates, 1.2% on flowers, 0.5% on leaves, 0.01% on vertebrates, and 0.6% on other plant foods. *Psidium guajava* (Myrtaceae) accounted for 21.2% of feeding and foraging time on fruit items. *Hampea appendiculata* (Malvaceae) (9.5%), *Dipteryx panamensis* (Fabaceae) (8.5%), *Inga spectabilis* (Fabaceae) (6.8%), *Urera baccifera* (Urticaceae) (6.5%), and *Inga thibaudiana* (Fabaceae) (6.3%) also accounted for a substantial percentage of white-faced capuchin feeding and foraging time (Table 1). Fifty-nine species of fruit from 27 plant families were consumed during the study period (for detailed dietary results, see S3 Table).

Table 1. Observed fruit diet based on the percentage of feeding and foraging time spent on each fruit species.

Species	Family	Percentage of diet	Average monthly percentage	Months eaten
<i>Psidium guajava</i>	Myrtaceae	21.2%	19.0% ($\pm 10.8\%$)	All
<i>Hampea appendiculata</i>	Malvaceae	9.5%	13.7% ($\pm 22.9\%$)	Jan, Feb, Nov, Dec
<i>Dipteryx panamensis</i>	Fabaceae	8.5%	6.9% ($\pm 10.8\%$)	Jan—Jun
<i>Inga spectabilis</i>	Fabaceae	6.8%	4.5% ($\pm 8.6\%$)	Feb—May, Oct
<i>Urera baccifera</i>	Urticaceae	6.5%	4.1% ($\pm 12.6\%$)	Oct—Dec
<i>Inga thibaudiana</i>	Fabaceae	6.3%	4.6% ($\pm 9.6\%$)	Jan, Mar—May, Jul—Dec
<i>Ficus colubrinae</i>	Moraceae	4.1%	4.0% ($\pm 3.8\%$)	Feb, Mar, May—Aug, Oct—Dec
<i>Ficus trigonata</i>	Moraceae	3.7%	2.5% ($\pm 6.8\%$)	Apr, May, Oct, Nov
<i>Conostegia xalapensis</i>	Melastomataceae	3.4%	2.8% ($\pm 3.8\%$)	Jan, Apr—Jun, Aug, Sep, Nov, Dec
<i>Sapium grandulosum</i>	Euphorbiaceae	2.8%	1.4% ($\pm 3.0\%$)	Jul, Aug, Oct, Nov
<i>Ficus schippii</i>	Moraceae	2.2%	2.6% ($\pm 5.6\%$)	May, Jun, Nov
<i>Inga marginata</i>	Fabaceae	2.1%	2.0% ($\pm 3.2\%$)	Jan, Feb, Jun, Aug, Dec
<i>Ficus tonduzii</i>	Moraceae	1.8%	1.2% ($\pm 3.1\%$)	May, Oct
<i>Nephelium lappaceum</i>	Sapindaceae	1.7%	1.7% ($\pm 3.0\%$)	Jun—Oct
<i>Cestrum megalophyllum</i>	Solanaceae	1.7%	1.1% ($\pm 4.1\%$)	Dec
<i>Dendropanax arboreus</i>	Araliaceae	1.6%	1.6% ($\pm 5.4\%$)	Sep, Oct
<i>Piper</i> spp.	Piperaceae	1.3%	1.0% ($\pm 1.1\%$)	Feb, Apr, Nov, Dec
<i>Miconia affinis</i>	Melastomataceae	1.2%	1.1% ($\pm 2.3\%$)	Apr—Jun, Aug, Sep
<i>Casearia arborea</i>	Salicaceae	1.1%	1.1% ($\pm 2.3\%$)	Jun, Oct, Nov
<i>Miconia elata</i>	Melastomataceae	1.1%	0.5% ($\pm 0.0\%$)	Apr, May, Aug, Nov

<https://doi.org/10.1371/journal.pone.0199556.t001>

Molecular dietary analysis results

Sequencing of *rbcL* amplicons yielded 1706 unique sequences, after dereplicating and denoising the sequence dataset. Thirty-six families, 39 genera, and 36 species were identified using *rbcL* (Table 2 and S4 Table). Sequences assigned to Solanaceae were the most frequently occurring *rbcL* sequences in the fecal samples, occurring in 99.4% of samples. Sequences assigned to Melastomataceae (97.1%), Araceae (88.2%), Moraceae (78.8%), Myrtaceae (66.5%), Urticaceae (47.7%), Fabaceae (41.8%), and Convolvulaceae (40.0%) also were among the most frequently occurring *rbcL* sequences (Table 2). These families also accounted for the highest relative abundances of *rbcL* sequences, with Solanaceae accounting for 62.9% of sequences assigned to family, Melastomataceae for 17.7%, Araceae for 8.0%, Moraceae for 4.0%, Myrtaceae for 2.6%, and Urticaceae for 1.5%.

Sequencing of *trnL* amplicons yielded 4551 unique sequences. Forty-five families, 68 genera, and 69 species were identified using *trnL* (Table 2 and S5 Table). Sequences assigned to Myrtaceae were the most frequently occurring *trnL* sequences in the fecal samples, occurring in 100% of samples. Sequences assigned to Fabaceae (99.4%), Musaceae (98.8%), Solanaceae (97.7%), Piperaceae (90.0%), Malvaceae (89.4%), Urticaceae (85.9%), Euphorbiaceae (83.5%), and Melastomataceae (80.0%) were among the most frequently occurring *trnL* sequences (Table 2). Many of these families also had high relative abundances of *trnL* sequences, with Myrtaceae accounting for 50% of sequences assigned to family, Solanaceae for 16.4%, Malvaceae for 7.1%, Piperaceae for 3.3%, and Urticaceae for 3.1%.

Overall, sequencing of *trnL* amplicons resulted in 311% more sequences and 267% more unique sequences, likely partially due to the shorter length of the *trnL* amplicon. A higher percentage of *trnL* sequences were identified to both order and family, a higher percentage of observed families also were identified with *trnL*, and the *trnL* method resulted in a greater number of identified families, genera, and species when compared with *rbcL* (Table 3). This

Table 2. Percentage of observed feeding and foraging records assigned to each family of fruit and frequency of occurrence and relative abundance of *rbcL* and *trnL* sequences over the course of the entire study period.

Family	Percentage of observed fruit feeding and foraging records	Frequency of occurrence of <i>rbcL</i> sequences	Average relative abundance of <i>rbcL</i> sequences	Total relative abundance of <i>rbcL</i> sequences	Frequency of occurrence of <i>trnL</i> sequences	Average relative abundance of <i>trnL</i> sequences	Total relative abundance of <i>trnL</i> sequences
Acanthaceae	-	-	-	-	11.76%	0.02%	0.01%
Annonaceae	-	-	-	-	74.71%	2.16%	1.93%
Apiaceae	-	1.18%	0.01%	0.00%	-	-	-
Apocynaceae	0.20%	-	-	-	-	-	-
Araceae	-	88.24%	11.26%	7.99%	-	-	-
Araliaceae	1.58%	-	-	-	-	-	-
Arecaceae	0.78%	-	-	-	33.53%	0.62%	0.76%
Asteraceae	-	-	-	-	0.59%	0.00%	0.00%
Bignoniaceae	0.32%	12.35%	1.08%	0.31%	31.76%	0.22%	0.26%
Boraginaceae	0.09%	1.76%	0.01%	0.00%	-	-	-
Brassicaceae	-	0.59%	0.00%	0.00%	-	-	-
Bromeliaceae	-	-	-	-	60.00%	1.64%	2.76%
Burseraceae	-	-	-	-	18.82%	0.17%	0.13%
Calyceraceae	-	19.41%	0.96%	0.85%	-	-	-
Canellaceae	-	-	-	-	3.53%	0.00%	0.00%
Capparaceae	-	-	-	-	1.76%	0.01%	0.01%
Caprifoliaceae	-	1.18%	0.06%	0.01%	-	-	-
Caricaceae	-	-	-	-	28.24%	0.66%	2.33%
Chenopodiaceae	-	1.76%	0.00%	0.00%	-	-	-
Chrysobalanaceae	-	-	-	-	2.94%	0.01%	0.00%
Clusiaceae	0.11%	17.65%	2.09%	0.44%	24.12%	0.63%	0.01%
Convolvulaceae	0.06%	40.00%	0.24%	0.11%	18.24%	0.01%	0.00%
Cordiaceae	-	-	-	-	10.00%	0.02%	0.00%
Costaceae	-	0.59%	0.04%	0.00%	7.06%	0.13%	0.00%
Cucurbitaceae	-	8.82%	0.01%	0.00%	-	-	-
Cyclanthaceae	-	2.94%	0.01%	0.02%	-	-	-
Euphorbiaceae	2.90%	4.12%	0.22%	0.11%	83.53%	2.66%	2.03%
Fabaceae	25.09%	41.76%	4.37%	0.40%	99.41%	5.65%	2.13%
Gesneriaceae	-	2.35%	0.07%	0.02%	2.35%	0.01%	0.01%
Heliconiaceae	0.34%	-	-	-	-	-	-
Hypoxidaceae	-	-	-	-	2.35%	0.00%	0.00%
Iteaceae	-	4.71%	0.01%	0.00%	-	-	-
Lamiaceae	0.43%	17.65%	0.90%	0.47%	25.29%	0.58%	0.09%
Lauraceae	0.06%	2.35%	0.01%	0.01%	24.71%	0.12%	0.18%
Lecythidaceae	-	21.18%	0.41%	0.00%	27.65%	0.05%	0.00%
Malpighiaceae	0.06%	-	-	-	-	-	-
Malvaceae	10.05%	2.35%	0.18%	0.01%	89.41%	5.52%	7.08%
Marantaceae	-	-	-	-	2.35%	0.06%	0.00%
Melastomataceae	6.17%	97.06%	15.11%	17.73%	80.00%	2.17%	1.73%
Moraceae	16.19%	78.82%	10.14%	4.01%	60.00%	3.42%	0.72%
Musaceae	0.17%	-	-	-	98.82%	3.48%	2.55%
Myristicaceae	0.17%	-	-	-	0.59%	0.01%	0.00%
Myrtaceae	21.27%	66.47%	9.64%	2.64%	100.00%	40.69%	49.90%
Onagraceae	-	2.94%	0.00%	0.00%	-	-	-

(Continued)

Table 2. (Continued)

Family	Percentage of observed fruit feeding and foraging records	Frequency of occurrence of <i>rbcL</i> sequences	Average relative abundance of <i>rbcL</i> sequences	Total relative abundance of <i>rbcL</i> sequences	Frequency of occurrence of <i>trnL</i> sequences	Average relative abundance of <i>trnL</i> sequences	Total relative abundance of <i>trnL</i> sequences
Orchidaceae	-	2.35%	0.04%	0.00%	-	-	-
Orobanchaceae	-	0.59%	0.00%	0.00%	0.59%	0.00%	0.00%
Papaveraceae	-	7.06%	0.03%	0.00%	-	-	-
Passifloraceae	-	-	-	-	24.71%	0.08%	0.09%
Piperaceae	1.29%	-	-	-	90.00%	4.33%	3.31%
Platanaceae	-	-	-	-	23.53%	0.16%	0.02%
Poaceae	-	-	-	-	7.06%	0.09%	0.02%
Primulaceae	-	-	-	-	1.18%	0.00%	0.00%
Rhamnaceae	-	-	-	-	24.71%	1.28%	0.10%
Rosaceae	-	2.35%	0.00%	0.00%	-	-	-
Rubiaceae	1.12%	11.18%	0.67%	0.11%	47.06%	1.29%	1.26%
Salicaceae	1.15%	-	-	-	-	-	-
Sapindaceae	1.72%	-	-	-	58.24%	0.74%	0.68%
Sapotaceae	0.26%	-	-	-	22.94%	0.14%	0.28%
Simaroubaceae	0.17%	-	-	-	2.35%	0.20%	0.02%
Solanaceae	1.69%	99.41%	36.61%	62.90%	97.65%	15.24%	16.39%
Strelitziaceae	-	0.59%	0.00%	0.00%	-	-	-
Trochodendraceae	-	2.35%	0.12%	0.00%	-	-	-
Typhaceae	-	-	-	-	4.12%	0.00%	0.00%
Ulmaceae	-	5.29%	0.00%	0.00%	-	-	-
Urticaceae	6.55%	47.65%	4.97%	1.53%	85.88%	5.66%	3.06%
Velloziaceae	-	-	-	-	0.59%	0.00%	0.00%
Zingiberaceae	-	16.47%	0.75%	0.33%	20.59%	0.02%	0.12%

Average relative abundances are calculated across all fecal samples collected during the study period.

<https://doi.org/10.1371/journal.pone.0199556.t002>

result is not surprising, however, given the larger size of the *trnL* reference database used for this study. A greater percentage of *rbcL* sequences were identified at the levels of genus and

Table 3. Comparison of *rbcL* and *trnL* sequencing performance, resolution, and accuracy.

	<i>rbcL</i>	<i>trnL</i>
Sequences	2,541,027	7,910,288
Unique sequences	1706	4551
Percentage identified to order	53.8%	89.7%
Percentage identified to family	47.2%	79.2%
Percentage identified to genus	9.1%	8.2%
Percentage identified to species	7.7%	2.0%
Identified families also in observed diet	14	21
Total number of families identified	67	
Families identified with all three methods	14	
Families identified with both <i>rbcL</i> and <i>trnL</i>	19	
Families only identified with observational data	5	
Families only identified with <i>rbcL</i>	16	
Families only identified with <i>trnL</i>	20	

<https://doi.org/10.1371/journal.pone.0199556.t003>

species (Table 3), which is surprising as the *trnL* database contains four times as many genera and eight times as many species as the *rbcL* database. Of the 67 total families identified in this study, 14 were detected with all three methods and 19 were detected with both *rbcL* and *trnL* (Table 3). While the uneven distribution of fecal samples across months likely contributed to the fact that neither *rbcL* or *trnL* detected all species of food present in the observational diet, 92.5% of plant families found in the observation diet were also found using the DNA metabarcoding methodology (Table 3).

Comparison of observed and molecularly determined diet

Non-parametric Mantel tests showed a significant positive correlation between the observed dietary composition, based on feeding and foraging time in a given month, and dietary composition for each month determined with *trnL* ($\rho = 0.1025$, $p = 0.045$). These results indicate that individuals with more similar monthly diets assessed with DNA metabarcoding using *trnL* had more similar monthly diets based on feeding and foraging observations. However, the same was not true for *rbcL*. There was no relationship between observed monthly plant dietary composition and monthly dietary composition determined with *rbcL* ($\rho = 0.0653$, $p = 0.145$).

There was no relationship between the frequency of occurrence of either *rbcL* or *trnL* sequences assigned to a family in fecal samples in a given month and the percentage of feeding and foraging time the group was observed to be eating food items from that same family in that month (S6 Table). The exceptions to this were the significant positive correlations between the frequency of occurrence of *rbcL* sequences assigned to Myrtaceae and percentage of feeding and foraging time spent consuming Myrtaceae ($\rho = 0.4869$, $p = 0.046$), and the frequency of occurrence of *trnL* sequences assigned to Piperaceae and percentage of feeding and foraging time spent consuming Piperaceae ($\rho = 0.6763$, $p = 0.006$) (Fig 2).

There were several plant families for which the average relative abundance of *rbcL* or *trnL* sequences were significantly positively correlated with the percentage of feeding and foraging time the group spent consuming that same family in a given month (S6 Table). The average

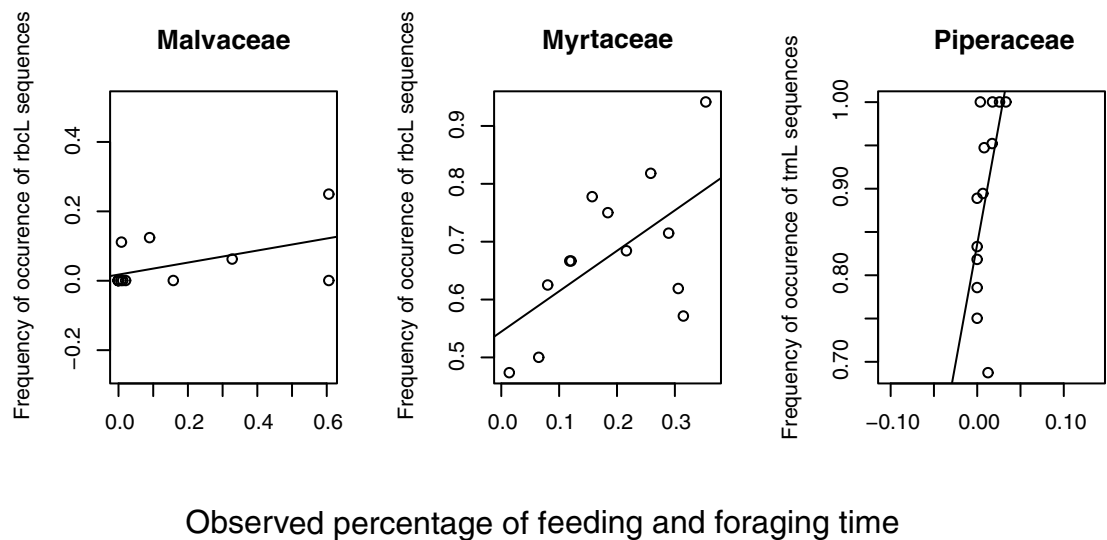


Fig 2. Relationship between feeding and foraging time and frequency of occurrence of *rbcL* and *trnL* sequences. Correlations between observed percentage of group feeding and foraging time spent on each plant family in a given month and the frequency of occurrence of *rbcL* or *trnL* sequences in the same month.

<https://doi.org/10.1371/journal.pone.0199556.g002>

relative abundance of *rbcl* sequences assigned to Fabaceae, Malvaceae, Myrtaceae, and Rubiaceae were significantly positively correlated with the percentage of monthly feeding and foraging time spent exploiting those same families ($\rho = 0.7473, p = 0.002$; $\rho = 0.5077, p = 0.038$; $\rho = 0.5840, p = 0.016$; $\rho = 0.5840, p = 0.018$) (Fig 3 and S6 Table). The average relative abundance of *trnL* sequences assigned to Malvaceae, Melastomataceae, Moraceae, Myrtaceae, Piperaceae, Rubiaceae, and Simaroubaceae were significantly positively correlated with the percentage of feeding and foraging time spent on those same families ($\rho = 0.8118, p < 0.001$; $\rho = 0.6658, p = 0.007$; $\rho = 0.5604, p = 0.025$; $\rho = 0.7418, p = 0.003$; $\rho = 0.6104, p = 0.013$; $\rho = 0.5722, p = 0.021$; $\rho = 0.5654, p = 0.022$) (Fig 4 and S6 Table). Of note, several of these families (Fabaceae, Malvaceae, Melastomataceae, Moraceae, and Piperaceae) were among the most frequently observed families consumed during the 12-month study period by this group of white-faced capuchins. There was a significant positive relationship between percentage of monthly feeding and foraging time spent consuming species from a particular family and the average relative abundance of sequences assigned to that family in four cases for *rbcl* and in seven cases for *trnL*.

Discussion

This study compared detailed observational dietary data with molecular dietary analysis using DNA metabarcoding of both *rbcl* and *trnL* chloroplast genes in a group of individually recognizable, wild white-faced capuchin monkeys. The results showed that using most metrics, *trnL* outperformed *rbcl*. Using *trnL* as a barcoding marker resulted in more sequences, more unique sequences, a greater percentage of sequences that could be identified to order or family, and results that were more comparable to those obtained from observational data. *rbcl* did have a greater percentage of sequences that could be identified to genus and species (9.1% vs. 8.2% and 7.7% vs. 2.0%), but neither marker had high rates of successful taxonomic assignments at these levels. This is not unexpected in light of previous work showing that neither *rbcl* or *trnL* have particularly high rates of species-level discrimination, particularly when compared with barcoding markers used in other organisms [1,3,6,7,13,37,38]. While these results may indicate that *trnL* has greater taxonomic resolution than *rbcl*, the difference in the number and identity of plant species present in both reference databases makes the difference

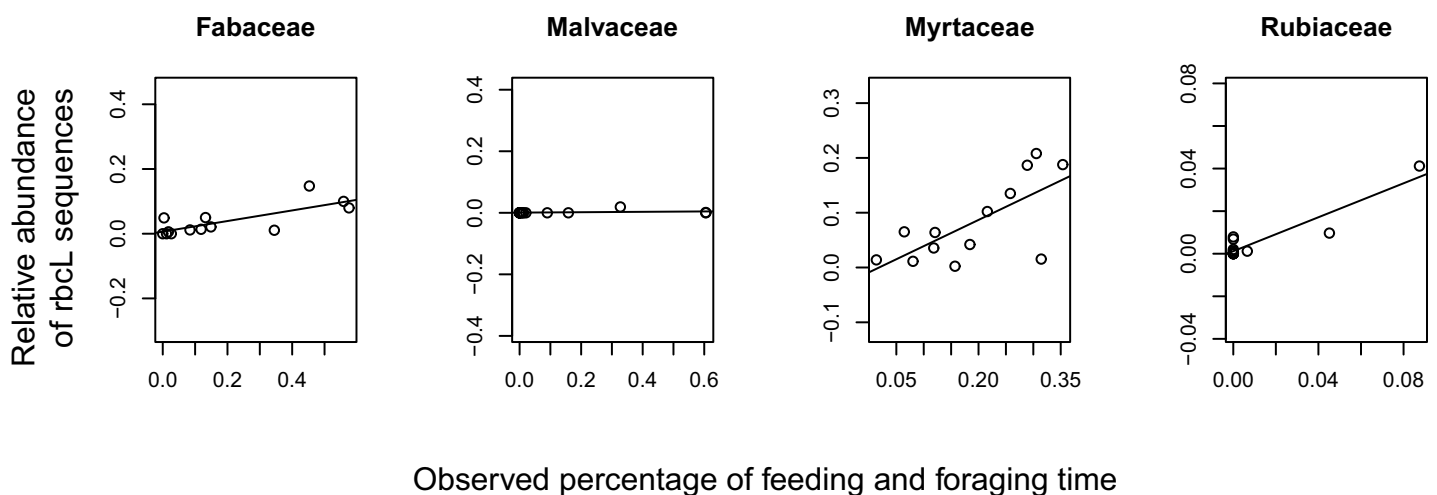


Fig 3. Relationship between feeding and foraging time and relative abundance of *rbcl* sequences. Correlations between observed percentage of group feeding and foraging time spent on each plant family in a given month and the average relative abundance of *rbcl* sequences for the same month.

<https://doi.org/10.1371/journal.pone.0199556.g003>

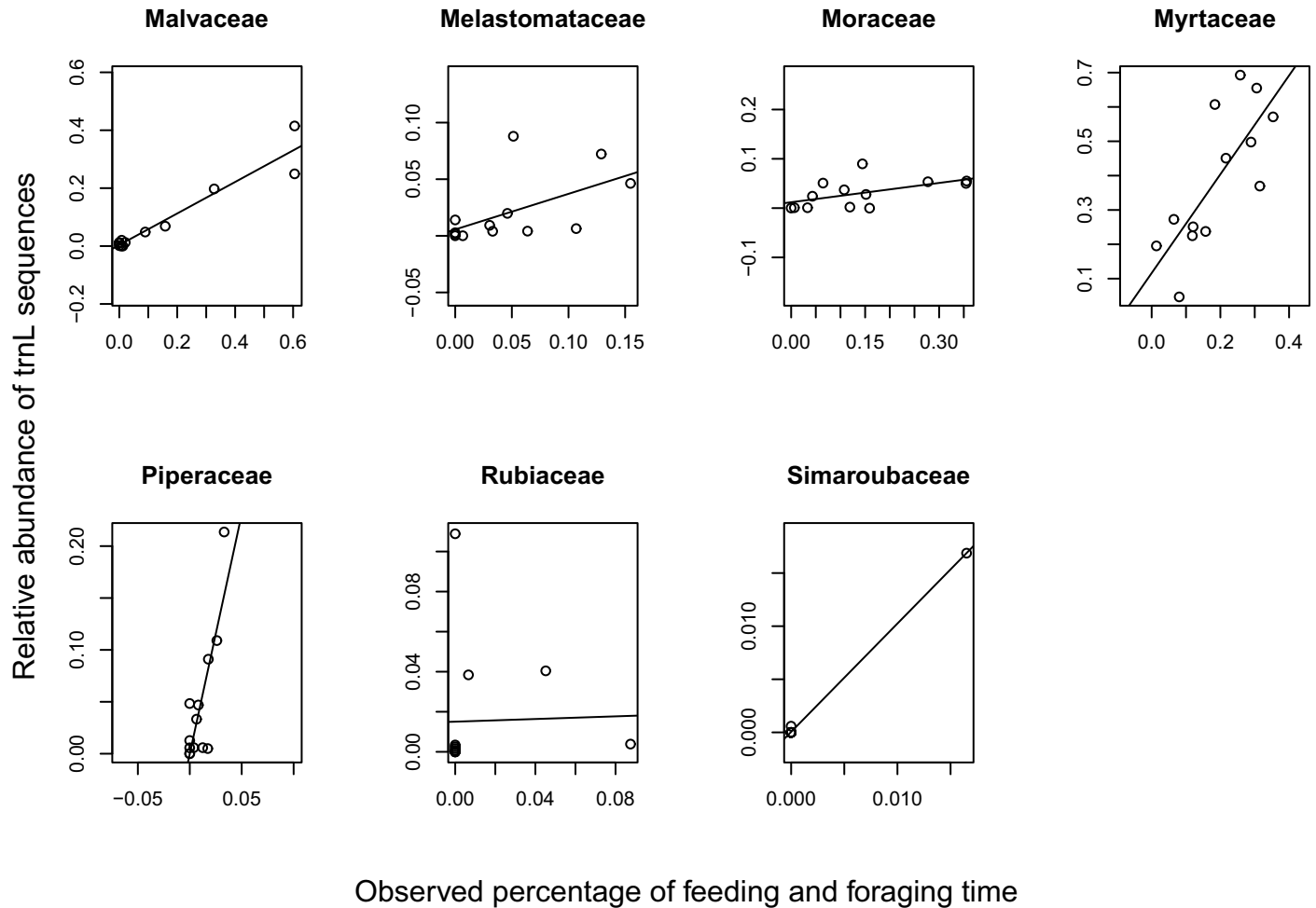


Fig 4. Relationship between feeding and foraging time and relative abundance of *trnL* sequences. Correlations between observed percentage of group feeding and foraging time spent on each plant family in a given month and the average relative abundance of *trnL* sequences for the same month.

<https://doi.org/10.1371/journal.pone.0199556.g004>

in resolution difficult to interpret. Genus- and species-level identifications would likely be improved with the use of a local reference plant DNA database for both markers. However, the lack of a good reference database is often the case in studies using DNA metabarcoding, particularly when the diet of the study species is not well characterized. Even without a local reference database, it appears that the correlation between results obtained from *trnL* sequences and the observed diet make *trnL* a good candidate for a barcoding marker that can be used to make inferences concerning broader questions about dietary differences between and within species living in the same habitats.

Some of the discrepancies between the three different identification methods are not surprising. The four families of plants identified only during behavioral observations are infrequently eaten plants—each making up less than 2% of the total white-faced capuchin diet (Table 2). Thus, it is likely that we would not have a fecal sample collected from an individual who consumed those food items within the past eight hours. Similarly, many of the families of plants that were only identified using the *trnL* method—Annonaceae, Bursaceae, Caricaceae, Cordiaceae (Boraginaceae), Passifloraceae, and Rhamnaceae—are plants that are known to be part of the white-faced capuchin diet, either at this site or other sites in Costa Rica, but

are infrequently consumed [39–41]. Other plant families identified by only the *rbcl* method (Araceae, Cyclanthaceae, and Ulmaceae) or the *trnL* method (Acanthaceae, Bromeliaceae, and Marantaceae) include plant species that are consumed by white-faced capuchins [39,40,42], but the flower or leaves are consumed, so they were not recorded as part of the fruit diet in this study. Both the *rbcl* method and the *trnL* method identified plant families that have not yet been recorded as part of the white-faced capuchin diet, but are present in this area of Costa Rica, and so may be rarely consumed by this group of white-faced capuchins. For the *rbcl* method, these include Brassicaceae, Caprifoliaceae, Curcubitaceae (either from crop raiding chayote trees from neighboring farms or consuming small undomesticated variants), Onagraceae, Orchidaceae, Papaveraceae, Rosaceae, and Strelitziaceae (while this family does not have any species native to Costa Rica, there are several ornamental varieties on the field station property). For the *trnL* method, these potentially rarely consumed families of plants include Asteraceae, Canellaceae, Capparaceae, Poaceae, Primulaceae, and Typhaceae. Both molecular methods also assigned sequences to plant families that are not found in Costa Rica (*rbcl*: Calyceraceae, Iteaceae, and Trochodendraceae; *trnL*: Platanaceae and Velloziaceae), and likely represent taxonomy assignment errors due to limitations of the reference databases.

For the majority (66.7%) of plant foods eaten by white-faced capuchins, our results indicate that frequency of occurrence of both *rbcl* and *trnL* sequences assigned to a given family in fecal samples was not correlated with the percentage of feeding and foraging time spent consuming fruits assigned to that same family. However, for plant foods such as Myrtaceae and Malvaceae, the relative abundances of *rbcl* and *trnL* sequences assigned to a give family were predictive of how much feeding and foraging time the white-faced capuchin group spent consuming fruits from that same family. This was particularly true of the relative abundance of *trnL* sequences and foods that account for at least 5% of white-faced capuchin monkey fruit consumption. Though these results suggest a predictive relationship between the relative abundance of *trnL* sequences and the relative time spent consuming fruits, the use of DNA metabarcoding to obtain quantitative dietary results is an area of active research. Whereas some studies have found no relationship between the relative abundance of *trnL* sequences and observational or microhistology data [20,30], other studies have had greater success. Srivathsan *et al.* (2015) found that the relative abundance of *trnL* sequences assigned to given genus was highly correlated with the amount of foods consumed from that genus by captive red-shanked douc langurs. In a study of large African herbivores, raw read abundance of *trnL* sequences assigned to grasses were tightly correlated with $\delta^{13}\text{C}$ values and estimates of relative C_4 -plant consumption from stable isotope data [10].

Though the results of our study suggest that *trnL* is a more robust DNA metabarcoding marker for dietary analysis and that relative abundances of *trnL* sequences may be able to be used to assess relative changes or differences in diet, the results should be interpreted cautiously. While percentage of feeding and foraging time is frequently used in primatological studies to track diet, it is not necessarily an accurate approximation of the amount of plant foods consumed [43], as individuals may consume different foods at different rates and individuals of different age, sex, or dominance status might consume different foods and in different proportions. Controlled feeding studies where the amount of a food consumed and gut passage rates are known are imperative to more precisely determine whether or not relative abundances of *trnL* can accurately be used to quantify the relative contributions of plants foods to animal diets or to compare changes in diet within or between species. Additionally, the extraction kit used likely contains DNA from plants in the Solanaceae family [1], resulting in contamination of our sequencing results. This may be responsible for the fact that Solanaceae DNA were present in a high proportion of fecal samples despite there only being observations of plants in the Solanaceae family being consumed in one month (December). Future

studies examining the impact of variation in sample storage methods, extraction methods, PCR protocols, and primer biases would be beneficial.

Comparing *trnLg/h* to other minibarcodes also would be beneficial. Several other barcodes have been proposed for plants, including *trnH-psbA* and *ITS*, and a multi-locus tiered approach using *rbcL* and *matK* has been selected as the barcode of choice by the CBOL Plant Working Group [6,7,9,13,44,45]. Several studies have used minibarcodes from these genes, or combinations of two DNA barcoding regions, to identify the plant component of animal diets [10,22,24,25,46,47], whereas other studies have used minibarcodes in different subsections of the *rbcL* gene [12,29]. Therefore, even though the *trnL* method outperformed the *rbcL* method in the context of the current study, other DNA barcoding approaches for determining the plants present in animal diets may be more accurate.

The *trnL* barcoding marker yielded results more similar to those from detailed behavioral observations when compared with the *rbcL* barcoding marker for DNA metabarcoding analyses of omnivorous and herbivorous diets. These results indicate that *trnL* is a more universal primer set with less amplification biases than *rbcL*, at least when amplifying plant DNA from fecal samples. Our data suggests that relative abundances of *trnL* sequences can be used to approximate relative time spent consuming a plant food, particularly for frequently consumed foods. However, additional studies are necessary to determine whether DNA metabarcoding using *trnL* can be used to accurately predict relative rates of consumption of plant foods.

Supporting information

S1 File. Detailed methods for the bioinformatic analyses.

(DOCX)

S1 Table. *ecotag* output for *rbcL* sequences with taxonomic annotations.

(TXT)

S2 Table. *ecotag* output for *trnL* sequences with taxonomic annotations.

(TXT)

S3 Table. Percentage of observed feeding and foraging time spent consuming plant species in each month.

(XLSX)

S4 Table. Relative abundance of *rbcL* sequences assigned specific taxa for the entire study period.

(XLSX)

S5 Table. Relative abundance of *trnL* sequences assigned specific taxa for the entire study period.

(XLSX)

S6 Table. Correlations between observed percentage of feeding and foraging time on plant families and both frequency of occurrence and relative abundances of *rbcL* and *trnL* sequences.

(XLSX)

Acknowledgments

Thank you to Maderas Rainforest Conservancy and the Molina family for allowing us to conduct research at LSBFS. Cindy Taylor, Liz Rasheed, Michaela Snead, and Elizabeth Clausing were indispensable in the field and lab. Thank you to the Roy J. Carver Biotechnology Center

at UIUC, especially Chris Wright, Mark Band, and Chris Fields. PAG wishes to thank Chrissie, Sara, and Jenni for their support. Additionally, we would like to thank several anonymous reviewers whose comments much improved the manuscript.

Author Contributions

Conceptualization: Elizabeth K. Mallott, Paul A. Garber, Ripan S. Malhi.

Formal analysis: Elizabeth K. Mallott.

Funding acquisition: Elizabeth K. Mallott, Paul A. Garber, Ripan S. Malhi.

Investigation: Elizabeth K. Mallott.

Methodology: Elizabeth K. Mallott, Ripan S. Malhi.

Resources: Ripan S. Malhi.

Supervision: Paul A. Garber, Ripan S. Malhi.

Writing – original draft: Elizabeth K. Mallott, Paul A. Garber, Ripan S. Malhi.

Writing – review & editing: Elizabeth K. Mallott, Paul A. Garber, Ripan S. Malhi.

References

1. Valentini A, Miquel C, Nawaz MA, Bellemain E, Coissac E, Pompanon F, et al. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach. *Mol Ecol Resour.* 2009; 9: 51–60. <https://doi.org/10.1111/j.1755-0998.2008.02352.x> PMID: 21564566
2. Pompanon F, Deagle BE, Symondson WOCC, Brown DS, Jarman SN, Taberlet P. Who is eating what: diet assessment using next generation sequencing. *Mol Ecol.* 2012; 21: 1931–50. <https://doi.org/10.1111/j.1365-294X.2011.05403.x> PMID: 22171763
3. Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, et al. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 2007; 35: e14. <https://doi.org/10.1093/nar/gkl938> PMID: 17169982
4. Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* 2007; 23: 167–172. <https://doi.org/10.1016/j.tig.2007.02.001> PMID: 17316886
5. Hebert PDN, Ratnasingham S, De Waard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc R Soc London B.* 2003; 270: S96–S99. <https://doi.org/10.1098/rsbl.2003.0025> PMID: 12952648
6. Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. *Proc Natl Acad Sci U S A.* 2009; 106: 12794–7. <https://doi.org/10.1073/pnas.0905845106> PMID: 19666622
7. Newmaster SGSG, Fazekas AJJ, Ragupathy S. DNA barcoding in land plants: evaluation of *rbcl* in a multigene tiered approach. *Can J Bot.* 2006; 84: 335–341. <https://doi.org/10.1139/b06-047>
8. Valentini A, Pompanon F, Taberlet P. DNA barcoding for ecologists. *Trends Ecol Evol.* 2009; 24: 110–7. <https://doi.org/10.1016/j.tree.2008.09.011> PMID: 19100655
9. Kress WJ, García-Robledo C, Uriarte M, Erickson DL. DNA barcodes for ecology, evolution, and conservation. *Trends Ecol Evol.* 2015; 30: 25–35. <https://doi.org/10.1016/j.tree.2014.10.008> PMID: 25468359
10. Kartzinel TR, Chen PA, Coverdale TC, Erickson DL, Kress WJ, Kuzmina ML, et al. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proc Natl Acad Sci.* 2015; 112: 8019–8024. <https://doi.org/10.1073/pnas.1503283112> PMID: 26034267
11. MacDonald AJ, Sarre SD. A framework for developing and validating taxon-specific primers for specimen identification from environmental DNA. *Mol Ecol Resour.* 2017; 17: 708–720. <https://doi.org/10.1111/1755-0998.12618> PMID: 27768246
12. Erickson DL, Reed E, Ramachandran P, Bourg NA, McShea WJ, Ottesen A. Reconstructing a herbivore's diet using a novel *rbcl* DNA mini-barcode for plants. *AoB Plants.* 2017; 9: plx015. <https://doi.org/10.1093/aobpla/plx015> PMID: 28533898

13. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, et al. Land plants and DNA barcodes: short-term and long-term goals. *Philos Trans R Soc B Biol Sci.* 2005; 360: 1889–1895. <https://doi.org/10.1098/rstb.2005.1720> PMID: 16214746
14. Poinar HN, Hofreiter M, Spaulding WG, Martin PS, Stankiewicz BA, Bland H, et al. Molecular coproscopy: Dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* (80-). 1998; 281: 402–406. <https://doi.org/10.1126/science.281.5375.402>
15. Soininen EM, Valentini A, Coissac E, Miquel C, Gielly L, Brochmann C, et al. Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Front Zool.* 2009; 6: 16. <https://doi.org/10.1186/1742-9994-6-16> PMID: 19695081
16. Soininen EM, Gauthier G, Bilodeau F, Berteaux D, Gielly L, Taberlet P, et al. Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLoS One.* 2015; 10: e0115335. <https://doi.org/10.1371/journal.pone.0115335> PMID: 25635852
17. Pegard A, Miquel C, Valentini A, Coissac E, Bouvier F, François D, et al. Universal DNA-based methods for assessing the diet of grazing livestock and wildlife from feces. *J Agric Food Chem.* 2009; 57: 5700–5706. <https://doi.org/10.1021/jf803680c> PMID: 19566081
18. Nichols R V., Åkesson M, Kjellander P. Diet assessment based on rumen contents: A comparison between DNA metabarcoding and macroscopy. *PLoS One.* 2016; 11: e0157977. <https://doi.org/10.1371/journal.pone.0157977> PMID: 27322387
19. Adhikari P, Park SM, Kim TW, Lee JW, Kim GR, Han SH, et al. Seasonal and altitudinal variation in roe deer (*Capreolus pygargus tianschanicus*) diet on Jeju Island, South Korea. *J Asia-Pacific Biodivers.* Elsevier Ltd; 2016; 9: 422–428. <https://doi.org/10.1016/j.japb.2016.09.001>
20. Nakahara F, Ando H, Ito H, Murakami A, Morimoto N, Yamasaki M, et al. The applicability of DNA barcoding for dietary analysis of sika deer. *DNA Barcodes.* 2015; 3: 200–206. <https://doi.org/10.1515/dna-2015-0021>
21. Bergmann GT, Craine JM, Robeson MS, Fierer N. Seasonal shifts in diet and gut microbiota of the American bison (*Bison bison*). *PLoS One.* 2015; 10: e0142409. <https://doi.org/10.1371/journal.pone.0142409> PMID: 26562019
22. De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, et al. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Mol Ecol Resour.* 2014; 14: 306–323. <https://doi.org/10.1111/1755-0998.12188> PMID: 24128180
23. Van Geel B, Protopopov A, Bull I, Duijm E, Gill F, Lammers Y, et al. Multiproxy diet analysis of the last meal of an early Holocene Yakutian bison. *J Quat Sci.* 2014; 29: 261–268. <https://doi.org/10.1002/jqs.2698>
24. Bradley BJ, Stiller M, Doran-Sheehy DM, Harris T, Chapman CA, Vigilant L, et al. Plant DNA sequences from feces: potential means for assessing diets of wild primates. *Am J Primatol.* Wiley Online Library; 2007; 69: 699–705. <https://doi.org/10.1002/ajp.20384> PMID: 17216626
25. Yang Y, Zhan A, Cao L, Meng F, Xu W. Selection of a marker gene to construct a reference library for wetland plants, and the application of metabarcoding to analyze the diet of wintering herbivorous waterbirds. *PeerJ.* 2016; 4: e2345. <https://doi.org/10.7717/peerj.2345> PMID: 27602302
26. Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, et al. Identification of amazonian trees with DNA barcodes. *PLoS One.* 2009; 4: e7483. <https://doi.org/10.1371/journal.pone.0007483> PMID: 19834612
27. Srivathsan A, Sha JCMM, Vogler AP, Meier R Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Mol Ecol Resour.* 2015; 15: 250–261. <https://doi.org/10.1111/1755-0998.12302> PMID: 25042073
28. Newmaster SG, Thompson ID, Steeves RADD, Rodgers AR, Fazekas AJ, Maloles JR, et al. Examination of two new technologies to assess the diet of woodland caribou: video recorders attached to collars and DNA barcoding. *Can J For Res.* 2013; 43: 897–900. <https://doi.org/10.1139/cjfr-2013-0108>
29. Khanam S, Howitt R, Mushtaq M, Russell JC. Diet analysis of small mammal pests: A comparison of molecular and microhistological methods. *Integr Zool.* 2016; 11: 98–110. <https://doi.org/10.1111/1749-4877.12172> PMID: 27001489
30. Hargrave A. Seasonal diets of sable antelope (*Hippotragus Niger*): Determined by microhistology and trnL sequence analysis. Texas State University. 2015. <https://doi.org/10.1017/CBO9781107415324.004>
31. Garber PA, Paciulli L. Experimental field study of spatial memory and learning in wild capuchin monkeys (*Cebus capucinus*). *Folia Primatol.* 1997; 68: 236–253. <https://doi.org/10.1159/000157250> PMID: 9360308

32. Mallott EK, Garber PA, Malhi RS. Integrating feeding behavior, ecological data, and DNA barcoding to identify developmental differences in invertebrate foraging strategies in wild white-faced capuchins (*Cebus capucinus*). *Am J Phys Anthropol*. 2017; 162: 241–254. <https://doi.org/10.1002/ajpa.23113> PMID: 27704526
33. Mallott EK, Malhi RS, Garber PA. High-throughput sequencing of fecal DNA to identify insects consumed by wild Weddell's saddleback tamarins (*Saguinus weddelli*, Cebidae, Primates) in Bolivia. *Am J Phys Anthropol*. 2015; 156: 474–481. <https://doi.org/10.1002/ajpa.22654> PMID: 25369770
34. Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E. obitools: A unix-inspired software package for DNA metabarcoding. *Mol Ecol Resour*. 2016; 16: 176–182. <https://doi.org/10.1111/1755-0998.12428> PMID: 25959493
35. Ficetola GF, Coissac E, Zundel S, Riaz T, Shehzad W, Bessi re J, et al. An in silico approach for the evaluation of DNA barcodes. *BMC Genomics*. 2010; 11: 434. <https://doi.org/10.1186/1471-2164-11-434> PMID: 20637073
36. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, et al. vegan: Community Ecology Package. 2017.
37. Hollingsworth ML, Andra Clark A, Forrest LL, Richardson J, Pennington RT, Long DG, et al. Selecting barcoding loci for plants: Evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Mol Ecol Resour*. 2009; 9: 439–457. <https://doi.org/10.1111/j.1755-0998.2008.02439.x> PMID: 21564673
38. Qu m r  E, Hibert F, Miquel C, Lhuillier E, Rasolondraibe E, Champeau J, et al. A DNA metabarcoding study of a primate dietary diversity and plasticity across its entire fragmented range. *PLoS One*. 2013; 8: e58971. <https://doi.org/10.1371/journal.pone.0058971> PMID: 23527060
39. Hogan JD, Melin AD, Mosdosy KN, Fedigan LM. Seasonal importance of flowers to Costa Rican capuchins (*Cebus capucinus imitator*): Implications for plant and primate. *Am J Phys Anthropol*. 2016; 161: 591–602. <https://doi.org/10.1002/ajpa.23059> PMID: 27492752
40. Melin AD, Hiramatsu C, Parr NA, Matsushita Y, Kawamura S, Fedigan LM. The Behavioral Ecology of Color Vision: Considering Fruit Conspicuity, Detection Distance and Dietary Importance. *Int J Primatol*. 2014; 35: 258–287. <https://doi.org/10.1007/s10764-013-9730-8>
41. Urbani B. Spatial mapping in wild white-faced capuchin monkeys (*Cebus capucinus*). University of Illinois at Urbana-Champaign. 2009.
42. Eadie EC. Feeding ecology and life history strategies of white-faced capuchin monkeys (*Cebus capucinus*). The University of New Mexico. 2012.
43. Kurland JA, Gaulin SJC. Comparability among measures of primate diets. *Primates*. 1987; 28: 71–77. <https://doi.org/10.1007/BF02382184>
44. Yao H, Song J, Liu C, Luo K, Han J, Li Y, et al. Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS One*. 2010; 5. <https://doi.org/10.1371/journal.pone.0013102> PMID: 20957043
45. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One*. 2007; 2. <https://doi.org/10.1371/journal.pone.0000508> PMID: 17551588
46. Hibert F, Sabatier D, Andrivot J, Scotti-Saintagne C, Gonzalez S, Pr vost MF, et al. Botany, genetics and ethnobotany: A crossed investigation on the elusive tapir's diet in French guiana. *PLoS One*. 2011; 6: e25850. <https://doi.org/10.1371/journal.pone.0025850> PMID: 21991372
47. Hibert F, Taberlet P, Chave J, Scotti-Saintagne C, Sabatier D, Richard-Hansen C. Unveiling the diet of elusive rainforest herbivores in next generation sequencing era? The tapir as a case study. *PLoS One*. 2013; 8: e60799. <https://doi.org/10.1371/journal.pone.0060799> PMID: 23560107