

Opinion

# Open Data for Open Questions in Comparative Nutrition

Juliano Morimoto <sup>1,\*</sup>  and Mathieu Lihoreau <sup>2</sup> 

<sup>1</sup> School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Ave, Aberdeen AB24 2TZ, UK

<sup>2</sup> Research Center on Animal Cognition (CRCA), Center for Integrative Biology (CBI); CNRS, University Paul Sabatier, 31062 Toulouse, France; mathieu.lihoreau@univ-tlse3.fr

\* Correspondence: juliano.morimoto@abdn.ac.uk

Received: 16 March 2020; Accepted: 7 April 2020; Published: 9 April 2020



**Abstract:** Achieving a better understanding of the consequences of nutrition to animal fitness and human health is a major challenge of our century. Nutritional ecology studies increasingly use nutritional landscapes to map the complex interacting effects of nutrient intake on animal performances, in a wide range of species and ecological contexts. Here, we argue that opening access to these hard-to-obtain, yet considerably insightful, data is fundamental to develop a comparative framework for nutrition research and offer new quantitative means to address open questions about the ecology and evolution of nutritional processes.

**Keywords:** nutritional ecology; geometric framework for nutrition; performance landscape; open data

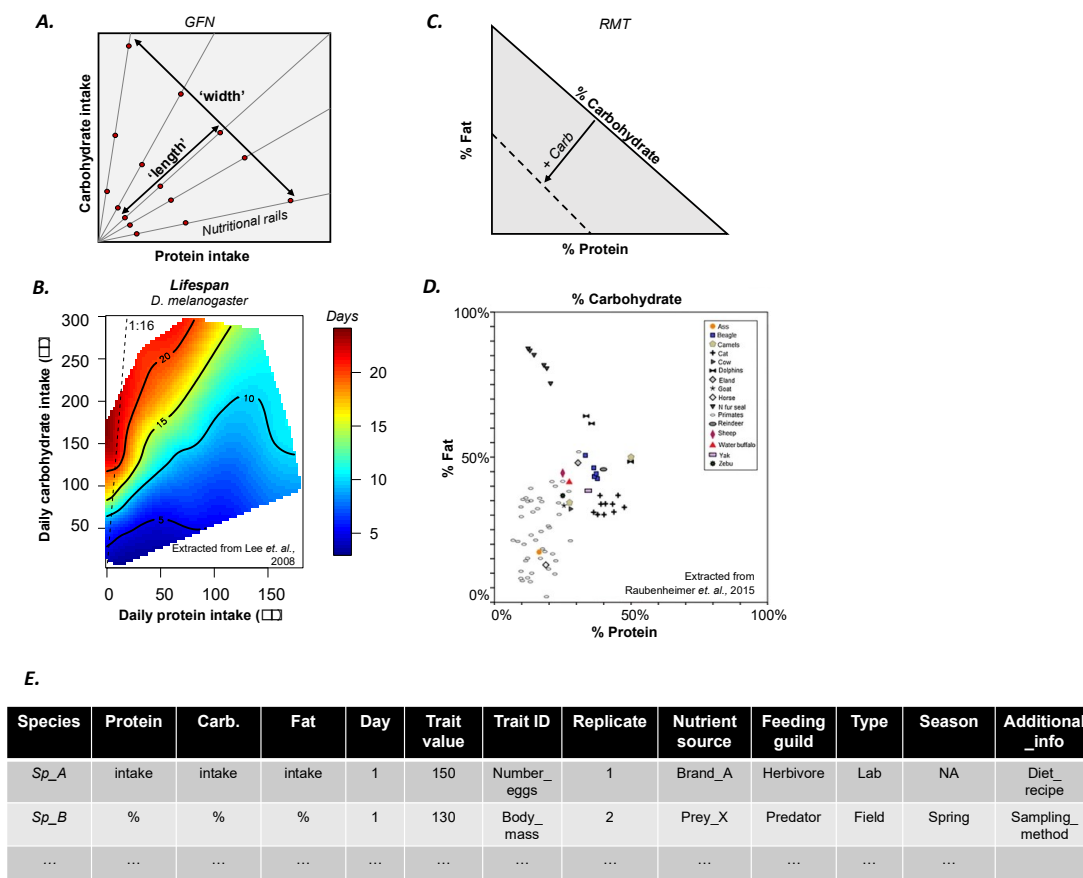
## 1. The Multidimensional Nature of Nutrition

The World Health Organisation (WHO) estimates that ca. 2.4 billion adults experience health-related problems due to malnutrition (<https://www.who.int/en/news-room/fact-sheets/detail/malnutrition>). Such alarming numbers have prompted scientists, clinicians, and policy makers to develop a more integrative nutrition science, by focusing their research on the effects of nutrition across a wide variety of contexts, spanning from fundamental science involving animal nutrition and the consequences to animal behaviour and evolution to human disease [1,2].

In recent decades, the conceptual framework known as the geometric framework for nutrition (GFN; Figure 1) [3] has revolutionized research in nutritional ecology by taking into account how variation in the quantity and balance of nutrients in foods can affect the development, metabolic health, reproduction and ageing of animals [2]. The GFN allows for direct experimental assessment of the effects of multiple nutrients and their calorie intakes on performance traits and ultimately fitness expression (e.g., Figure 1A,B) [4]. This has brought key insights into the nutritional factors of a wide variety of physiological and behavioural phenomena across feeding guilds (e.g., herbivores, carnivores, and omnivores) and taxonomic groups (e.g., slime moulds, insects, fish, birds, and mammals) [2,5]. These progresses in nutrition research also have far reaching implications in various areas of biology, ecology and evolution, including to our understanding of contemporary diseases affecting human societies [6]. Here, we argue that making GFN data open access has considerable benefits to advancing the broad field of nutrition research and tackling open questions on the ecology and evolution of nutritional processes through quantitative comparative approaches of complex nutritional systems across ecological contexts and scales of organisations. To do so, we provide an up-to-date summary of available data and a guideline for the standardized sharing of future data.

## 2. The Geometric Framework for Nutrition in a Nutshell

The GFN [2] is a state-space modelling approach that can be used to study the simultaneous contribution of two or more nutrients to performance traits (e.g., ratios of proteins to carbohydrates [7], lipids to proteins [8], proteins to lipids and carbohydrates [9,10], and amino acids [11]). Diets with different nutrient ratios (i.e., ‘nutritional rails’, Figure 1A) determine the ‘width’ of the performance landscape in the nutritional plane (i.e., ‘nutritional space’, Figure 1A). Diets with the same nutrient ratio can have different nutrient concentrations, which introduces variance in individual nutrient intake and determines the ‘length’ of the performance landscape (Figure 1A). Experimentally, the higher the number of different nutrient ratios, nutrient concentrations and replicates, the higher the resolution of the landscape (see [2] for a review).



**Figure 1.** Overview of the geometric framework for nutrition (GFN). (A) Schematic representation of the performance landscape plotted in a two-dimensional nutrient space (here defined by protein (P) and carbohydrate (C) intake). Nutritional rails are diets with varying PC ratios. Red dots represent the concentrations of the different diets. The ‘width’ of the landscape is determined by the nutritional rails. The ‘length’ of the landscape is determined both by the concentrations within each rail and the individual variation in nutrient intake within each rail. (B) Empirical performance landscape (i.e., lifespan) from a GFN study in *Drosophila melanogaster* (extracted from [7]). (C) Schematic representation of the Right-Angle Mixture Triangle (RMT) used to assess diet mixtures primarily in field studies. Note that the % of carbohydrate in the mixture increases as the axis moves towards the origin (solid arrow) so that the mixture is constrained within the limits delimited by the dashed isocline. (D) Empirical Right-Angle Mixture Triangle (i.e., milk composition) from a field study (extracted from [12]). (E) Template for storing and sharing GFN data. An Excel (.csv) version is provided in Appendix A.

Higher resolution, and thus confidence in the results, is obtained by a higher coverage of the nutritional space with more nutritional rails which are also in closer proximity, following:

$$\frac{90^\circ}{N_{rails}} = \varnothing$$

where  $N_{rails}$  is the number of nutritional rails in the experimental design and  $\varnothing$  is the average distance between rails. However, an increase in the number of nutrient ratios in arithmetic scale also results in an increase in sample size in the geometric scale, following:

$$N_{final} = NR \times C \times R$$

where  $N_{final}$  is the final sample size of the experiment,  $NR$  is the number of nutrient ratios,  $C$  is the number of nutrient concentrations, and  $R$  is the number of replicates per concentration per ratio. This relationship constrains experimental designs that can grow rapidly and compromises their feasibility. For example, GFN experiments tend to adopt a design of at least five replicates of five nutrient ratios with three concentrations each (75 replicates) [2]. For an experiment with no more than six nutrient ratios, the sample size increases from 75 to 90 replicates ( $6 \times 3 \times 5$ ) for an increase of  $\sim 17\%$  in the landscape resolution ( $15^\circ$  between nutritional rails). Data collection is therefore a bottleneck for the quality and resolution of fitness landscape and represents a trade-off between feasibility (i.e., the rapid increase in sample size) and resolution (i.e., higher coverage of the nutritional space).

For field studies, when this approach can be tedious and all required information may not be available, a derivative of the GFN known as the Right-Angle Mixture Triangle (RMT) [13] can be used to transform diet contents into relative percentages of each specific nutrient of interest. The percentage of each nutrient in the diet is represented in the axes of RMT (Figure 1C), which depicts the nutrient mixture in percentage of the diet. For example, a recent study analysed the macronutrient composition of mammalian milk, revealing that, overall, primate milk is lower in protein concentration relative to the milk of other mammals (Figure 1D) [12]. Since nutrient mixtures are represented by percentages, the concentration of the nutrients is not incorporated into the graphical representation which instead offers a graphical overview of the diet mixture (see Figure 1A). Nonetheless, in theory, RMT can be used to derive multidimensional performance landscapes and compare broader nutritional ecology patterns between populations in different environments or between species occupying the same environment [13]. Here, data collection is also a bottleneck for data resolution and statistical inferences in comparative studies, meaning that this approach also experiences a trade-off on quality (and costs) and resolution.

For the remainder of this paper, we will use the GFN to refer to both the more general GFN (described above) and the derivative RMT. All concepts proposed here are therefore applicable to and will benefit both frameworks.

### 3. Raw Data Availability: Why Does It Matter?

Performance landscapes are often drawn using extrapolation methods such as thin-plate splines (TPSs) [14]. TPSs uses information from the collected data in each nutrient ratio and concentration to extrapolate the values for the areas of the nutritional space where experimental data is lacking. This allows for the construction of a continuous landscape from discrete experimental data. This approach is widespread to study morphological, behavioural, physiological and reproductive adaptations in species, spanning from insects to mammals [6,7], including humans [15]. However, while the extrapolation method solves the data collection conundrum, it generates a new problem—TPSs often mask the information from the raw (empirical) data, making it virtually impossible to infer and/or extract these data (see, e.g., Figure 1B). As a result, each GFN study becomes an ‘island’ of information inaccessible to researchers. This limitation strongly limits the possibility to perform meta-analyses,

and discover broad-scale patterns through comparative studies, which is essentially developing the GFN from a descriptive to an analytical method.

#### 4. More Data for A Bigger Picture

GFN studies are increasingly published with nutritional response landscapes in a variety of organisms (see summary in Data S1) [2]. The data for each population and species are time consuming (and often costly) to obtain but extremely insightful for the community. For example, Lee et al. [7] ran experiments with 1008 fruit flies (28 diets, 36 replicates) to build three landscapes showing for the first time that nutrient balance, and not caloric restriction, was responsible for extending lifespan in insects (Figure 1B). Similar conclusions were later obtained in mammals by Solon-Biet et al. [9], who ran experiments with 858 mice (25 diets). This fast-growing amount of empirical data thus offers the new fascinating possibility to compare traits between populations and species. To date, however, performance surface data have not been provided in more than 60% of the GFN studies (see meta-analysis in Data S1). We believe that making such raw GFN data open access and published alongside studies (irrespective of the journal's policy) will allow for the development of a truly integrated and collaborative community of nutritional ecologists with accessible datasets from which quantitative analyses can be performed and new methods for data analysis can be designed and tested [16,17].

#### 5. How to Share GFN Data?

Although the value of open data has been widely recognized to explore broader patterns and processes across species, space and time, few fields of biological and ecological sciences have fully embraced this practice, and so the landscape of open data remains largely scattered and complex to navigate [18]. To facilitate access to and reuse of GFN data by the community, we encourage authors to adopt a simple common guideline.

1. Provide raw data used to construct landscapes in a standardized and accessible format (see example in Figure 1E; see template file (.csv) in Appendix A).
2. If publishing in an open-access journal, authors are likely (although not always) to be required to deposit the raw data in a public data repository (e.g., Dryad). If publishing in a subscription journal, we encourage authors to make raw data available as Supplementary Material. In both cases, it is important that the data format is also standardised (Figure 1C) and any additional information needed for the understanding of the dataset is also provided (e.g., README file). These steps will benefit the visibility and citations of the original paper.
3. Include 'geometric framework for nutrition' and 'performance landscapes' as keywords in publications to increase exposure on search engines.
4. If the study is already published but the data are not available, we encourage authors to deposit the raw data in a public data repository.

#### 6. Towards Quantitative and Comparative Nutrition Research

Open data have the potential to transform biological and ecological sciences with a new depth of information that can facilitate advances across disciplines and explore broader-scale patterns [19]. Nutritional ecology is not an exception and will assuredly greatly benefit from this practice, since good quality data are particularly hard to obtain. Just like the adoption of performance landscapes has brought new fundamental insights into the ecology and evolution of nutrition [4,6,7,17], formally comparing large quantities of landscapes based on data sharing will allow for new kinds of quantitative analyses characterizing the responses of individuals to nutritional conditions from different life-stages, populations and species. Recent studies have already demonstrated the potential for new theoretical and comparative avenues of research with open data in nutritional ecology [16,17]. For instance, we developed a quantitative approach to compare GFN performance landscapes within and between

species, and revealed major differences in the reproductive responses of *Drosophila melanogaster* (data from [7]) and the tephritid *Bactrocera tryoni* (data from [20]), whereby the species-specific reproductive responses are driven by differences in protein (but not carbohydrate) intake [17]. Likewise, comparative studies on the relative composition of diets using the RMT have provided key insights into the nutritional ecology of higher vertebrates in the wild including marine predators [21] as well as primates (including humans [22,23]; see above). Beyond nutrition research, strictly speaking, these data could be used to study a broad range of biological interactions mediated by food and quantitatively address general problems in ecology and evolution. Are there nutritional adaptations required for sociality? How do nutritional interactions mediate host–parasite evolution? To what extent do nutritional adaptations shape species assemblages? How do nutritional constraints associated with trophic levels determine evolutionary trade-offs? Many of these questions could be addressed through meta-analyses of open nutritional data.

## 7. Conclusions

Nutritional ecology studies increasingly use nutritional landscapes to map the complex interacting effects of nutrient intake on animal performances, in a wide range of species and ecological contexts. Here we encourage authors to adopt a common guideline to share their raw data. Opening access to these hard-to-obtain, yet considerably insightful data is fundamental to develop a comparative framework for nutrition research and offer new quantitative means to address open questions about the ecology and evolution of nutritional processes.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2075-4450/11/4/236/s1>, Data S1: List of publications using the GFN. The list was collated using two search engines (Web of Science and Google Scholar; search on 15 January 2020) with queries for ‘geometric framework’ and ‘nutrition’. In total, 68 studies fitted the criteria (i.e., data articles displaying multidimensional performance landscapes)—of which, 27 provided the raw data along with the publication or in a data repository (~39%).

**Author Contributions:** M.L. conceived the ideas; J.M. collected the data; M.L. and J.M. analysed the data; M.L. and J.M. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication. All authors have read and agreed to the published version of the manuscript.

**Funding:** M.L. was funded by CNRS and a research grant from the Agence Nationale de la Recherche (ANR-16-CE02-0002-01).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Excel (.csv) template for sharing GFN raw data.

Species	Protein	Carb.	Fat	Day	Trait Value	Trait_ID	Replicate	Nutrient Source	Feeding Guild	Type of Experiment	Season	Additional_Info
<i>Sp_A</i>	intake	intake	intake	1	150	Number_eggs	1	Brand_A	Herbivores	Lab	NA	Diet_recipe
<i>Sp_B</i>	%	%	%	1	130	Body_mass	2	Prey_X	Predator	Field	Spring	Sampling_method

## References

1. Branca, F.; DeMaio, A.; Udomkesmalee, E.; Baker, P.; Aguayo, V.M.; Barquera, S.; Dain, K.; Keir, L.; Lartey, A.; Mugambi, G.; et al. A new nutrition manifesto for a new nutrition reality. *Lancet* **2019**, *395*, 8–10. [CrossRef]
2. Simpson, S.J.; Raubenheimer, D. *The Nature of Nutrition: A Unifying Framework from Animal Adaptation to Human Obesity*, 1st ed.; Princeton University Press: Princeton, NJ, USA, 2012.
3. Simpson, S.J.; Raubenheimer, D. A multi-level analysis of feeding behaviour: The geometry of nutritional decisions. *Philos. Trans. R. Soc. B Boil. Sci.* **1993**, *342*, 381–402.
4. Simpson, S.J.; Sibly, R.M.; Lee, K.P.; Behmer, S.; Raubenheimer, D. Optimal foraging when regulating intake of multiple nutrients. *Anim. Behav.* **2004**, *68*, 1299–1311. [CrossRef]
5. Raubenheimer, D.; Simpson, S. *Eat Like the Animals: What Nature Teaches Us about the Science of Healthy Eating*; Houghton Mifflin: London, UK, 2020.

6. Raubenheimer, D.; Simpson, S.J. Nutritional Ecology and Human Health. *Annu. Rev. Nutr.* **2016**, *36*, 603–626. [[CrossRef](#)] [[PubMed](#)]
7. Lee, K.P.; Simpson, S.J.; Clissold, F.J.; Brooks, R.C.; Ballard, J.W.O.; Taylor, P.; Soran, N.; Raubenheimer, D. Lifespan and reproduction in *Drosophila*: New insights from nutritional geometry. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2498–2503. [[CrossRef](#)] [[PubMed](#)]
8. Mayntz, D.; Behrens, M.M.; Ali, S.S.; Dao, D.N.; Lucero, J.D.; Shekhtman, G.; Quick, K.L.; Dugan, L.L. Nutrient-Specific Foraging in Invertebrate Predators. *Science* **2005**, *307*, 111–113. [[CrossRef](#)] [[PubMed](#)]
9. Solon-Biet, S.M.; McMahon, A.C.; Ballard, J.W.O.; Ruohonen, K.; Wu, L.E.; Cogger, V.C.; Warren, A.; Huang, X.; Pichaud, N.; Melvin, R.G.; et al. The ratio of macronutrients, not caloric intake, dictates cardiometabolic health, aging, and longevity in ad libitum-fed mice. *Cell Metab.* **2014**, *19*, 418–430. [[CrossRef](#)] [[PubMed](#)]
10. Morimoto, J.; Senior, A.; Ruiz, K.; Wali, J.A.; Pulpitel, T.; Solon-Biet, S.M.; Cogger, V.C.; Raubenheimer, D.; Le Couteur, D.G.; Simpson, S.J.; et al. Sucrose and starch intake contribute to reduced alveolar bone height in a rodent model of naturally occurring periodontitis. *PLoS ONE* **2019**, *14*, e0212796. [[CrossRef](#)] [[PubMed](#)]
11. Arganda, S.; Bouchebti, S.; Bazazi, S.; Le Hesran, S.; Puga, C.; Latil, G.; Simpson, S.J.; Dussutour, A. Parsing the life-shortening effects of dietary protein: Effects of individual amino acids. *Proc. R. Soc. B Boil. Sci.* **2017**, *284*, 20162052. [[CrossRef](#)] [[PubMed](#)]
12. Raubenheimer, D.; Machovsky-Capuska, G.E.; Chapman, C.A.; Rothman, J.M. Geometry of nutrition in field studies: An illustration using wild primates. *Oecologia* **2014**, *177*, 223–234. [[CrossRef](#)] [[PubMed](#)]
13. Raubenheimer, D. Toward a quantitative nutritional ecology: The right-angled mixture triangle. *Ecol. Monogr.* **2011**, *81*, 407–427. [[CrossRef](#)]
14. Nychka, D.; Reinhard, F.; Paige, J.; Sain, S. *Fields: Tools for Spatial Data*; R Packag. 9.9; University Corporation for Atmospheric Research: Boulder, CO, USA, 2017.
15. Rosas, A.; Bastir, M. Thin-plate spline analysis of allometry and sexual dimorphism in the human craniofacial complex. *Am. J. Phys. Anthropol. Off. Publ. Am. Assoc. Phys. Anthropol.* **2002**, *117*, 236–245. [[CrossRef](#)] [[PubMed](#)]
16. Rapkin, J.; Jensen, K.; Archer, C.R.; House, C.M.; Sakaluk, S.K.; Del Castillo, E.; Hunt, J. The Geometry of Nutrient Space-Based Life-History Trade-Offs: Sex-Specific Effects of Macronutrient Intake on the Trade-Off between Encapsulation Ability and Reproductive Effort in Decorated Crickets. *Am. Nat.* **2018**, *191*, 452–474. [[CrossRef](#)]
17. Morimoto, J.; Lihoreau, M. Quantifying Nutritional Trade-Offs across Multidimensional Performance Landscapes. *Am. Nat.* **2019**, *193*, E168–E181. [[CrossRef](#)]
18. Culina, A.; Baglioni, M.; Crowther, T.W.; Visser, M.E.; Woutersen-Windhouwer, S.; Manghi, P. Navigating the unfolding open data landscape in ecology and evolution. *Nat. Ecol. Evol.* **2018**, *2*, 420–426. [[CrossRef](#)] [[PubMed](#)]
19. Reichman, O.J.; Jones, M.; Schildhauer, M. Challenges and Opportunities of Open Data in Ecology. *Science* **2011**, *331*, 703–705. [[CrossRef](#)] [[PubMed](#)]
20. Fanson, B.G.; Weldon, C.W.; Simpson, S.J.; Taylor, P.; Perez-Staples, D. Nutrients, not caloric restriction, extend lifespan in Queensland fruit flies (*Bactrocera tryoni*). *Aging Cell* **2009**, *8*, 514–523. [[CrossRef](#)] [[PubMed](#)]
21. Machovsky-Capuska, G.E.; Raubenheimer, D. The Nutritional Ecology of Marine Apex Predators. *Annu. Rev. Mar. Sci.* **2020**, *12*, 361–387. [[CrossRef](#)] [[PubMed](#)]
22. Cui, Z.; Wang, Z.; Zhang, S.; Wang, B.; Lu, J.; Raubenheimer, D. Living near the limits: Effects of interannual variation in food availability on diet and reproduction in a temperate primate, the Taihangshan macaque (*Macaca mulatta tcheliensis*). *Am. J. Primatol.* **2019**, *82*, e23080. [[CrossRef](#)] [[PubMed](#)]
23. Raubenheimer, D.; Rothman, J.M. Nutritional Ecology of Entomophagy in Humans and Other Primates. *Annu. Rev. Entomol.* **2013**, *58*, 141–160. [[CrossRef](#)] [[PubMed](#)]

