

# DietR –

## A dietary analysis tool for ASA24 and NHANES in R



Rie Sadohara<sup>1</sup>, David Jacobs<sup>2</sup>, Mark A. Pereira<sup>2</sup>, Abigail Johnson<sup>2</sup>

<sup>1</sup>Kyoto-city, Kyoto, Japan <sup>2</sup>Division of Epidemiology & Community Health, University of Minnesota

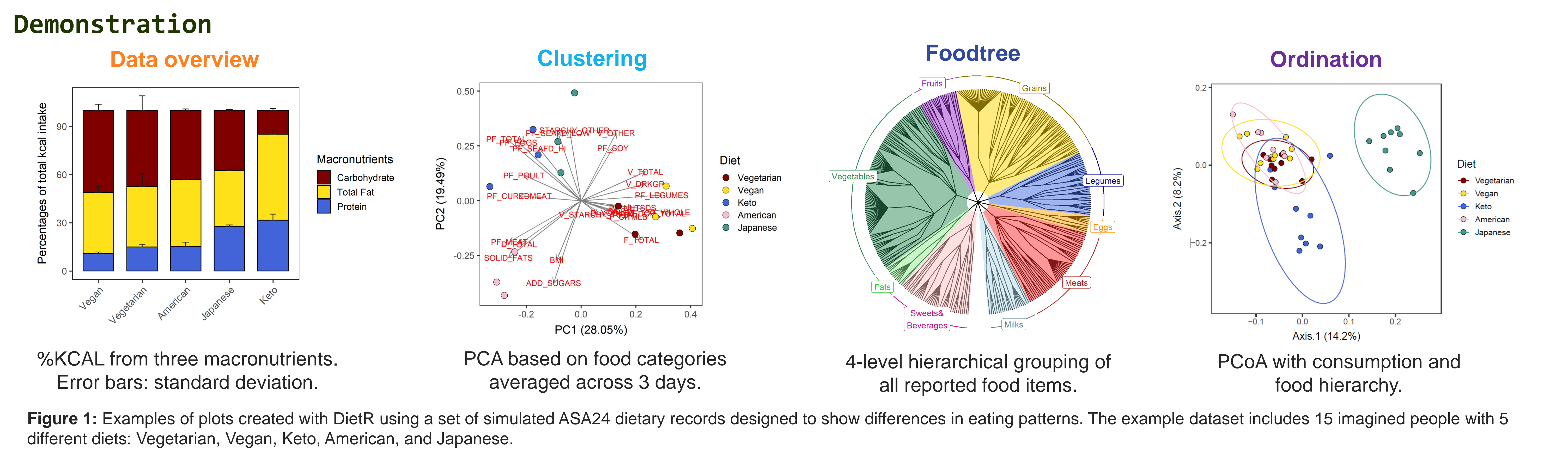


### Background

- Analysis of 24-hour recall data can be complicated and difficult.
- Many dietary datasets and dietary analysis tools are written in SAS.
- R is open-source and customizable with packages.
- We developed a package “DietR” to analyze NHANES and ASA24 data with R.

### Functionality of DietR

<b>Data preparation</b>	Load, filter, compute total food intake for each participant, compute means of food intake across days/groups, filter the total data for outliers.
<b>Data overview</b>	Data summary, % KCAL by macronutrients in barcharts.
<b>Diversity</b>	Compute $\alpha$ -diversity indices for dietary records, participants, or food groups.
<b>Clustering</b>	Principal component analysis (PCA), $k$ -means, select the optimal $k$
<b>Foodtree</b>	Build foodtrees [1] where foods in FNDDS are hierarchically grouped, visualize foodtrees, generate individual food consumption tables (“vegan” package [2]).
<b>Ordination</b>	Principal Coordinate Analysis (PCoA) based on their food consumption amount and the similarity of foods taken into account (“vegan” package [2]).



### Use case vignette: nuts/seeds/legumes diversity & body measures

#### Background and Research question

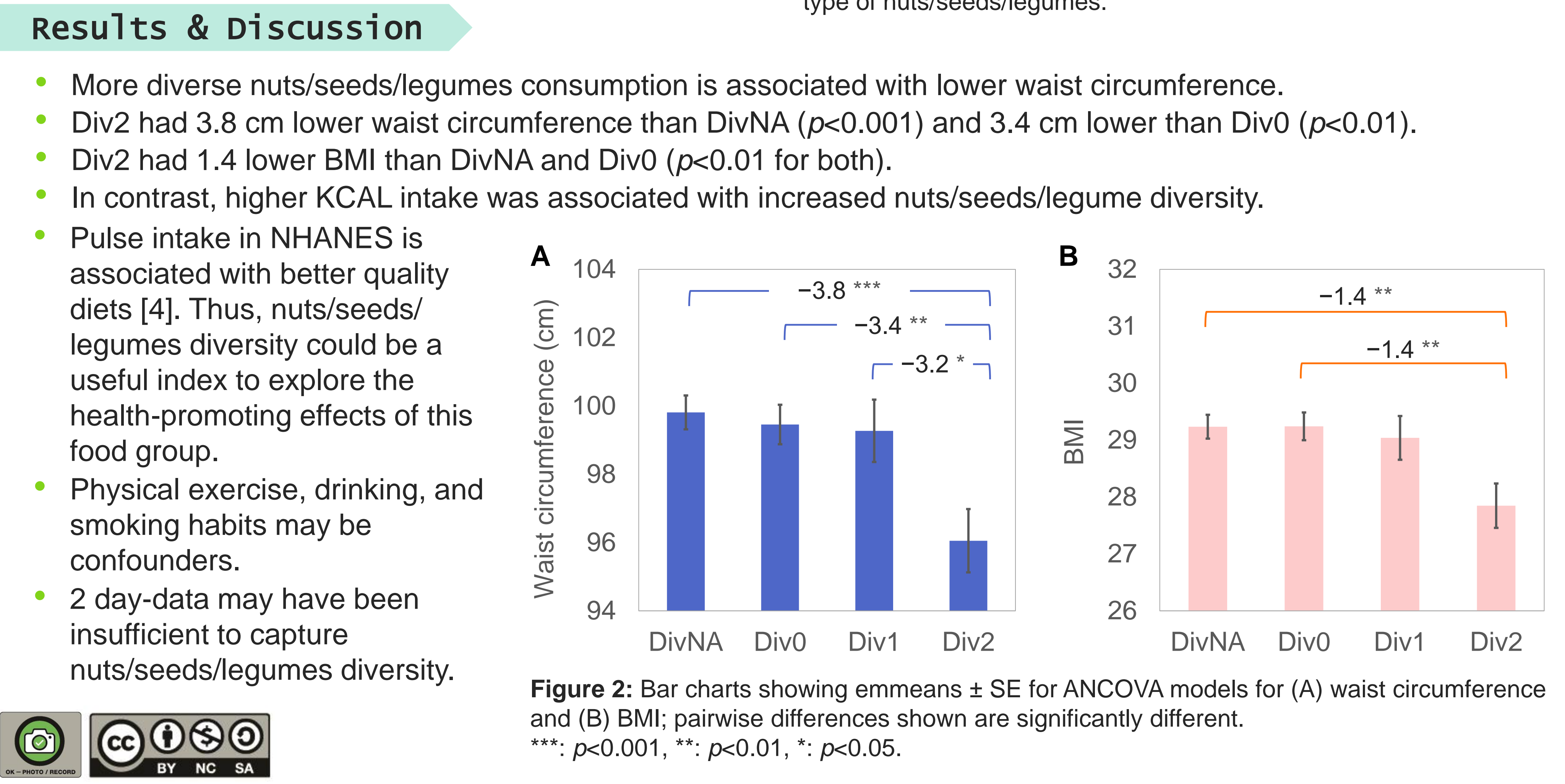
- Previous studies suggest nuts/seeds/legumes have positive impacts on health [3].
- Is diversity of nuts/seeds/legumes consumption related to body measures, e.g. waist size or BMI?

#### Methods

- NHANES 2015–16, n=3,641, 18+ yo, with waist circumference & BMI.
- Extracted reported food items with their foodcodes starting with 4 (Foodcode 4xxxxxxx: nuts/seeds/legumes) from 2 days of recalls.
- Calculated nuts/seeds/legumes  $\alpha$ -diversity.
- Defined diversity groups (**Table 1**).
- Analysis of covariance (ANCOVA) with Age, Sex, Race, Income, Education, KCAL as covariates

DivGroup	n	nuts/seeds/legumes consumed	Diversity index
DivNA	1,819	0	NA
Div0	1,105	1	0
Div1	360	>1	0.027 – 0.66
Div2	357	>1	0.66 – 1.95

**Table 1:**  $\alpha$ -diversity groups. DivNA represents no intake of nuts/seeds/legumes. Div0 are individuals who consumed 1 type of nuts/seeds/legumes. Div1 and Div2 consumed more than 1 type of nuts/seeds/legumes.



### Availability

- GitHub repo   
<https://github.com/computational-nutrition-lab/DietR>
- Website with tutorials   
<https://computational-nutrition-lab.github.io/DietR/>
- Preprint on medRxiv   
<https://doi.org/10.1101/2023.07.07.23292390>

### References

- Johnson AJ, Vangay P, Al-Ghalith GA, et al. Daily sampling reveals personalized diet-microbiome associations in humans. Cell Host Microbe. 2019;25(6):789-802.
- Simpson GL, Minchin PR, De Caceres M, et al. vegan: Community Ecology Package. 2022.
- Karlsen MC, Ellmore GS, McKeown N. Seeds—Health benefits, barriers to incorporation, and strategies for practitioners in supporting consumption among consumers. Nutr Today. 2016;51(1):50-59.
- Mitchell DC, Marinangeli CPF, Pigat S, et al. Pulse intake improves nutrient density among US adult consumers. Nutrients. 2021;13(8):2668.

### Acknowledgements

This project was supported by internal institutional start-up funds from the University of Minnesota. The authors would like to thank **Mo Hutti** for the create\_corr\_frame function which generates a correlation table with ordination axes and variables; **Pajau Vangay** for the collapse\_by\_correlation function which removes correlated variables; and **Suzie Hoops** for the matrix multiplication operation and her insights into statistical analyses.