How dietary landscapes impact food allergy

Duan Ni1,2,3, Alistair Senior2,5,6, Jian Tan2,3, Laurence Macia2,3,7, Ralph Nanan1,2,4,*

1 Sydney Medical School Nepean, The University of Sydney, Sydney, NSW, Australia
2 Charles Perkins Centre, The University of Sydney, Sydney, NSW, Australia
3 School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia
4 Nepean Hospital, Nepean Blue Mountains Local Health District, Sydney, NSW, Australia
5 School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW, Australia
6 Sydney Precision Data Science Centre, The University of Sydney, Sydney, NSW, Australia
7 Translational Science Hub, Global Sanofi Vaccines R&D, Brisbane, QLD, Australia

*Correspondence:
Ralph Nanan
ralph.nanan@sydney.edu.au

Abstract

Diets and environments are critical determinants for food allergy development. Harnessing unprecedented epidemiological and nutritional data, we examined the overall dietary environments for common food allergens and their intrinsic nutrient composition. We found that food and macronutrient supplies minimally impacted food allergy prevalence, but higher protein and glycine in food allergens correlated with less allergies. These findings offer new directions in food allergy research and management.
Main

Food allergies is emerging as a significant public health concern\textsuperscript{1,2}. One of the central ambitions in allergy research is to characterize the risk factors for food allergy. Previous research mostly focused on the macroenvironments (hygiene hypothesis), the individuals’ biological characteristics, and immunological aspects with gut microbiome as an intermediary\textsuperscript{1}. Factors that have been neglected, however, include the intrinsic properties of common trigger foods, like their nutrient compositions, and their extrinsic environments, like interactions with other dietary components.

Furthermore, early life exposure to allergenic foods is linked to lower risks of food allergies\textsuperscript{3-5}. Nevertheless, little is known about whether similar protective effects apply to exposures to specific food environments and nutrient environments. Moreover, studies towards what food components trigger allergy have mostly concentrated on allergenic proteins recognized by the immune system\textsuperscript{6}. However, accumulating evidence on involvements of other macronutrients, like fats and carbohydrates, in allergenic sensitization\textsuperscript{7-9}, suggested that a more holistic overview towards food components and their potential interactions is needed.

Harnessing comprehensive epidemiological and nutritional data, we systematically analyzed the extrinsic and intrinsic nutritional factors associated with the prevalence of common food allergies. We interrogated how food and nutrient environments correlate with food allergy prevalence and how the nutrient compositions of foods are linked to allergenicity. Survey of a collection of food allergy studies for the 8 most common food allergens found that extrinsic dietary environments, reflected by the food and macronutrient supplies, minimally impacted food allergy prevalence. However, a robust association was observed between the intrinsic protein and glycine content in food allergens and reduced allergy prevalence.

We harnessed a dataset investigating allergies for the 8 common food allergens (milk, egg, peanut, tree nut, wheat, soy, fish and shellfish) in Europe from 2001 to 2021\textsuperscript{10} (Figure 1A). This dataset presented a robust epidemiological landscape in a broad spatiotemporal context for these common and well-characterized food allergies. Figure 1B summarized the pooled estimates for all age self-reported food allergy lifetime and point prevalence. Milk and egg exhibited the highest lifetime and point prevalence respectively.

Prevalence data were logit transformed and the food and macronutrient data of the same year were collated for analysis. Annual gross domestic product (GDP) was included for adjustments. GDP reflects the socioeconomic status/wealth, potentially influencing allergy prevalences\textsuperscript{2}.

Early introduction of allergenic foods confers protection against food allergies later in life\textsuperscript{3-5}. To test whether exposures via dietary environments confers similar effects, food supplies were used as proxies, and modelled against corresponding food allergy prevalence. Linear models were fitted for the food allergy prevalence against the corresponding scaled food supply and GDP data, with consideration of their potential interactions. As shown in Figure 1C, there was no significant effect of the supplies of the 8 individual allergens on allergy prevalence. Only for tree nut, GDP showed a significant effect. These results suggested that exposures via allergenic food supplies
and the socioeconomic environments, were not generally associated with corresponding food allergy prevalence.

We previously showed that allergic diseases could be modulated via diets and were associated with the overall nutrient environments, another critical aspect of dietary exposures\textsuperscript{11-13}. Food allergy prevalence was thus compared against the macronutrient supplies of the corresponding countries and time points. We found no association, except for tree nut and fish exhibiting negative correlations with fat supplies (Figure 1D). This indicated that the overall nutrient environments, represented by macronutrient supplies, minimally influenced food allergy prevalence.

After extrinsic environments, we next attempted to characterize the intrinsic properties of food allergens, examining not only proteins, the main focus previously, but also other nutrient components and their potential interactions.

Macronutrient compositions of the 8 common food allergens were calculated (Figure 2A-B) and compared against the lifetime food allergy prevalence estimates in 2014\textsuperscript{14} and 2023\textsuperscript{10}. As shown in Figure 2C-D, protein compositions in foods were negatively correlated with their corresponding food allergy prevalence, while no association was found for other components (Figure S1A-C). Analysis based on individual studies yielded similar result (Figure 2E). To account for potential random confounding effects from specific foods or the countries where the studies were carried out, an array of linear mixed-effect models were fitted. Nutrient compositions of the potential allergenic foods were used as predictors. Foods and countries the data was based on were adjusted as random effects. A model considering the effects from protein content with additive random effects from foods and countries was favoured based on Akaike information criterion (Figure 2F, Table S1). It also unveiled a robust negative correlation between protein content and allergy prevalence (Figure 2G). Significantly, these associations were validated in independent American and Canadian datasets\textsuperscript{15} (Figure S1D-G).

We further interrogated whether amino acids might exert similar effects. Indeed, glycine content exhibited consistent negative correlations with allergy prevalence (Figure 2H-I).

Together, our analyses revealed that protein and glycine content, in the 8 common food allergens, negatively correlated with their allergy prevalence.

Leveraging epidemiological and nutritional data, we systematically interrogated how extrinsic and intrinsic nutritional factors are linked to food allergy prevalences. While dietary environments lacked evident effects, higher protein and glycine content in food allergens correlated with lower allergy prevalence.

It is established that early life exposure to food allergens induces tolerance against allergies\textsuperscript{3-5}. Our results suggested that exposures reflected by supplies and availabilities of allergenic foods and dietary macronutrients did not confer such protective effect. This warrants further investigations regarding the threshold and timing of exposures for allergy protection.

Instead of focusing solely on food allergenic proteins, our analyses delineated the contributions to allergenicity from different nutrient components and their plausible interactions, shifting the
paradigm for interrogating the intrinsic determinants of food allergenicity. We found that higher protein content in trigger foods was associated with lower allergy prevalence, while other nutrient components conferred negligible influences. Specifically, glycine content was linked to reduced food allergy prevalence. Our findings were based on the 8 most common food allergens, with well-characterized clinical features and well-documented epidemiological data. Further high quality epidemiological and nutritional data is needed to validate our discoveries beyond aforementioned 8 allergens. Additionally, similar findings were found for both European and North American populations. It remains to be tested if such results would pertain at a broader global scale.

It is unclear how protein content modifies food allergenicity (Figure 2J). High protein content might alter the food digestibility, interfering with allergen exposure to the host and thus tolerance induction. In this context, the amounts of proteins in allergenic foods are likely to also affect antigen uptake and processing, hence influencing allergic sensitization.

Apart from quantity, protein qualitative properties of trigger foods might be equally important. Glycine content in food allergens correlated with decreased allergy prevalence. Its small size might tweak peptide flexibility and further protein biophysical features. This might also interfere with protein digestion, and/or antigen processing and presentation. Additionally, glycine independently exhibits potent immunomodulatory capacities\(^{16}\), suppressing acute allergic responses\(^{17}\). Whether these properties could extend to high glycine foods remains to be uncovered.

Together, our study sheds unprecedented insights towards dietary factors influencing food allergy by comprehensive surveying their epidemiological and nutritional landscapes, opening novel avenues for allergy research. Our findings prompt future studies to unravel the mechanistic and pathophysiology of food allergies in a broader dietary context, which could guide food allergy managements and/or prevention.

**Methods**

Food allergy data and nutritional data were collected as described in Supplementary Information. Analyses were run in RStudio (v4.1.2) with `stats` and `lme4` packages. Detailed methods are available in Supplementary Information.

**Figure legends**

**Figure 1. A-B.** Overview of allergy studies for the 8 common food allergens (milk, egg, wheat, soy, peanut, tree nut, fish and shellfish, **A**) and their pooled results (**B**). **C.** Effect sizes of the linear modellings for scaled food allergen supplies (kg/capita/year), scaled GDP per capita (U.S. dollars) and the corresponding logit transformed food allergy prevalence. Error bars depict 95% confidence intervals. Red dot denotes statistical significance. **D.** Correlation analyses for the
logit transformed food allergy prevalence and their corresponding national protein (P, red), carbohydrate (C, cyan) and fat (F, blue) supplies (kg/capita/year) of the same year.

**Figure 2. A-B.** Overview of the nutrient compositions (g content/100g food, A) and distributions (B) of the 8 common food allergens. C-D. Correlation analyses for the protein content in food allergens and their corresponding logit transformed allergy prevalence estimates in 2023 (C) and 2014 (D). E. Correlation analysis for the protein content in food allergens and their corresponding logit transformed allergy prevalence reported in 93 studies. F. Statistical outputs for the null model and the chosen linear mixed-effect model analyzing protein content in food allergens and their corresponding logit transformed allergy prevalence. G. Estimate plot for the effect of protein content (g/g food) in food allergens and their corresponding logit transformed allergy prevalence based on the chosen linear mixed-effect model. H-I. Correlation analyses for the glycine content (mg/100g food) in food allergens and their corresponding logit transformed allergy prevalence estimates in 2023 (H) and 2014 (I). J. Model for the impacts of dietary environments and nutrient compositions on food allergy.

**Data availability**

All data used in the present study are publicly available as described in Supplementary Information.

**Acknowledgements**

This project is supported by the Norman Ernest Bequest Fund.

**Contributions**

DN participated to the study design, performed most of the analyses and wrote the manuscript. AS, JT, and LM participated in the analyses and data interpretation. RN supervised the study and wrote the manuscript.

All authors read and approved the final manuscript.

**Corresponding author**

Ralph Nanan

ralph.nanan@sydney.edu.au

Sydney Medical School Nepean, The University of Sydney.

Nepean Hospital, Level 5, South Block, Penrith NSW, 2751, Australia
Telephone: +61 2 4734 1614
Fax: +61 2 4734 1144

Ethics declarations

LM is a current employee of the Translational Science Hub Global Sanofi Vaccines R&D Brisbane, Australia. Her contribution to this work was when she was an employee of the University of Sydney. The other authors declare no competing interests.

References


The image contains a graph and a table comparing the prevalence of food allergies with the supply of various macronutrients. The top left section (A) lists the main food allergens: Milk, Egg, Wheat, Soy, Peanut, Tree nuts, Fish, and Shellfish. The top right section (B) shows a graph illustrating the prevalence of allergies against the supply of macronutrients.

The graph (B) has the x-axis labeled as "Effect size" and the y-axis as "Logit transformed allergy prevalence." The x-axis ranges from -6 to 6, and the y-axis ranges from 0.0 to 1.5. The graph includes bars for different allergens, with each bar representing the effect size of the allergy prevalence associated with the supply of macronutrients.

The table (C) compares the supply of macronutrients with the prevalence of allergies. The table includes columns for Protein supply, Carbohydrate supply, Fat supply, Milk, Egg, Wheat, Soy, Peanut, Tree nuts, Fish, and Shellfish. Each row represents the effect size (R^2) and p-value for the logistic regression model predicting allergy prevalence from the supply of macronutrients.

Finally, the bottom section (D) further details the relationship between macronutrient supply and allergy prevalence, with separate panels for each macronutrient category (Protein, Carbohydrate, Fat, Milk, Egg, Wheat, Soy, Peanut, Tree nuts, Fish, Shellfish). Each panel contains scatter plots showing the logit-transformed allergy prevalence against the macronutrient supply (kg/capita/year), with regression lines and effect sizes (R^2) indicated for each allergen.

---

**Figure 1.**

Macronutrient supply (kg/capita/year)

Macronutrient supply (kg/capita/year)
**A.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Fat</th>
<th>Carbohydrate</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>3.27</td>
<td>3.2</td>
<td>4.63</td>
</tr>
<tr>
<td>Egg</td>
<td>12.4</td>
<td>8.65</td>
<td>0.96</td>
</tr>
<tr>
<td>Wheat</td>
<td>15.1</td>
<td>2.73</td>
<td>71.2</td>
</tr>
<tr>
<td>Soy</td>
<td>38.6</td>
<td>20.7</td>
<td>27.9</td>
</tr>
<tr>
<td>Peanut</td>
<td>24.4</td>
<td>49.7</td>
<td>21.3</td>
</tr>
<tr>
<td>Tree nut</td>
<td>19.6</td>
<td>53.6</td>
<td>21</td>
</tr>
<tr>
<td>Fish</td>
<td>19.3</td>
<td>14.5</td>
<td>6.64</td>
</tr>
<tr>
<td>Shellfish</td>
<td>15.5</td>
<td>6.67</td>
<td>4.9</td>
</tr>
</tbody>
</table>

**B.**

- **R^2 = 0.4820** P value:*
- **R^2 = 0.7406** P value:*
- **R^2 = 0.2157** P value:****

**C.**

R^2 = 0.4820 P value:0.05

**D.**

R^2 = 0.7406 P value:*

**E.**

R^2 = 0.2157 P value:****

**F.**

<table>
<thead>
<tr>
<th>Estimate</th>
<th>CI (95%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prot</td>
<td>-5.55</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Carb</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Fat</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**G.**

- **R^2 = 0.5634** P value:*
- **R^2 = 0.6998** P value:*