De Vito *et al. Nutrition Journal* (2025) 24:71
https://doi.org/10.1186/s12937-025-01138-0

Nutrition Journal

RESEARCH

Open Access

Identifying and characterizing shared and ethnic background site-specific dietary patterns in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

Roberta De Vito[1†], Briana Stephenson[2†], Daniela Sotres-Alvarez[3], Anna-Maria Siega-Riz[4], Josiemer Mattei[5], Maria Parpinel[6], Brandilyn A. Peters[7], Sierra A. Bainter[8], Martha L. Daviglus[9], Linda Van Horn[10] and Valeria Edefonti[11,12*]

Abstract

Background A *posteriori* dietary patterns (DPs) are critical for capturing actual dietary behaviour. However, assessing their reproducibility across (sub)populations requires novel modelling approaches beyond descriptive statistics. Multi-study factor analysis derives DPs that are shared among all studies/subpopulations and those specific to a study or subpopulation of interest. Bayesian implementation of the multi-study factor analysis (BMSFA) is more flexible than frequentist as it imposes fewer assumptions and improves factor selection.

Methods We applied BMSFA to 24-h dietary recalls from the baseline visit (2008–2011) of the US Hispanic Community Health Study/Study of Latinos ($n = 16{,}415$). The analysis was conducted on 42 common nutrients to identify shared and subpopulation-specific DPs. Subpopulations were defined based on the cross-classification of ethnic background (Cuban, Dominican Republic, Mexican, Puerto Rican, Central and South American) and study site (Bronx, Chicago, Miami, San Diego) resulting in 12 Ethnic Background Site (EBS) categories. Regression analysis characterized DPs in terms of food groups, overall diet quality, socio-demographic/lifestyle factors, adjusting for survey design.

Results We identified four shared DPs across all EBS categories: *Plant-based foods*, *Processed foods, Dairy products*, and *Seafood*. Additionally, twelve EBS-specific DPs were identified—one for each EBS category. Most EBS-specific DPs were further grouped into overarching profiles: *Animal vs. vegetable source, Animal source only*, and *Poultry vs. dairy products*, to capture nuances within animal-based DPs. Puerto Rican background participants from Chicago expressed a strikingly different DP from all others (i.e., high on beta-carotene and low on starch/iron/thiamin).

Higher overall diet quality was observed with increasing categories of *Plant-based foods*, *Seafood*, and the "Puerto Rican background – Chicago" EBS-specific DP, whereas increasing categories of *Dairy products*, *Processed foods*, and the remaining EBS-specific DPs were related to lower diet quality. Compared to non-US-born participants, US-born individuals had significantly higher adjusted mean scores in absolute value for most DPs. Specifically, they

†Roberta De Vito and Briana Stephenson contributed equally to this work.

*Correspondence:
Valeria Edefonti
valeria.edefonti@unimi.it
Full list of author information is available at the end of the article

exhibited lower adherence to the *Plant-based foods* and *Dairy products* DPs but higher adherence to *Processed foods*, *Seafood*, and six EBS-specific DPs.

Conclusions The BMSFA successfully captured sources of dietary homogeneity and heterogeneity among US Hispanic/Latino adults across ethnic backgrounds and study sites. The study highlighted the crucial role of nativity on DPs.

Keywords Bayesian analysis, Bayesian multi-study factor analysis, Dietary patterns, Factor analysis, Hispanics/Latinos, Hispanic Community Health Study/Study of Latinos, Multi-study factor analysis, Reproducibility of dietary patterns

Background

A posteriori dietary patterns (DPs) [1] obtained by using multivariate statistical methods [i.e., principal component analysis (PCA), factor analysis (FA), and cluster analysis [2]], are critical for capturing population-specific dietary behaviors shaped by geography/climate, socioeconomic status, food supply, ethnic background, or religion [3]. However, their specificity limits generalizability compared to *a priori* DPs, which compare subjects' diet against benchmark diets based upon scientific evidence or theory for specific diseases [4]. The lack of a standardized approach for DP analysis (e.g., definition of input variables and their preprocessing, DP identification method, and DP labeling) has hindered the identification of genuinely reproducible *a posteriori* DPs, both across subpopulations of a single study and across different studies involving similar populations (i.e., *cross-study* reproducibility [5]). These challenges, compounded by inconsistencies in reporting and subjective DP labeling, have traditionally restricted meaningful comparisons of *a posteriori* DPs, thereby preventing firm conclusions about their health benefits or risks in study populations of interest [6, 7].

To our knowledge, only one article [8] has explored the reproducibility of *a posteriori* DPs within subpopulations belonging to the same study. This analysis relied on stratified PCAs across US region, sex, and race to determine the optimal number of DPs to retain for the final analysis of the overall sample. However, in a multicultural population, a comprehensive assessment of how regional and ethnic diversity contribute to unique DPs requires more efficient statistical approaches to manage and accommodate this additional complexity [9]. Ethnic backgrounds diversity poses additional challenges in the collection of dietary information, including the coverage of typical foods and traditional recipes from all different ethnic backgrounds [10].

To address the reproducibility of PCA- or FA-derived DPs across ethnically diverse subpopulations within a single study, methods used in pioneering [11, 12] and recent [13–19] research on DP reproducibility across different populations can be explored. Two major statistical approaches are commonly used [3]: 1. stratified approach, in which separate DPs are derived for each study and reproducibility is assessed via congruence coefficients or agreement measures [13, 14, 18], and 2. pooled approach, in which studies are combined into a single dataset to identify shared DPs [15, 16].

In between the two approaches, multi-study factor analysis [20] (MSFA) has been recently applied in nutritional epidemiology [17] for its ability to detect both shared and study-specific DPs across different studies. Within a large multi-cultural study, the MSFA allows to identify DPs that are consistently present across subpopulations—thus considered "reproducible"—while also capturing subpopulations-specific DPs. Bayesian MSFA (BMSFA) enhances this approach by providing greater flexibility, imposing fewer assumptions, and improving both DP identification and the selection of the optimal number of DPs to retain, a critical advantage as the number of subpopulations analyzed increases [21].

The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) [22] is an ongoing US multi-site community-based cohort studying health and risk factors of cardiovascular and pulmonary outcomes of Hispanic/Latino adults from various ethnic backgrounds and geographic regions [23]. This cohort provides the unique opportunity to define both shared and subpopulation-specific DPs by the cross-section of ethnic background and study site (EBS categories), supporting the development of targeted nutritional interventions. Previous HCHS/SOL papers have identified DPs using methods conceptually similar to MSFA but differing in scope and execution [24, 25]. One paper [24] identified food-based DPs on 24-h dietary recalls at baseline with stratified FA by ethnic background, but it ignored site-specific heterogeneity and relied on subjective judgments of pattern similarity and thus not statistically evaluated. Another paper [25] used cluster analysis by ethnic background and site but focused on frequency consumption levels from the food propensity questionnaire at follow-up rather than nutrient intakes, which measure amount of consumption.

This paper leverages BMSFA to jointly identify shared and EBS-specific DPs based on nutrient intake derived from 24-h recalls, addressing the following key research questions: 1. Are there empirically estimable DPs shared across HCHS/SOL participants? 2. Are there one or more EBS-specific DPs? 3. How do the BMSFA-derived DPs compare to those obtained from traditional approaches in nutritional epidemiology? 4. Are the identified shared and EBS-specific DPs interpretable in terms of food groups, overall diet quality, socio-demographic and lifestyle factors?

Methods

Reporting of information in the current article followed recommendations included in the Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut) statement [26].

Study design

The HCHS/SOL enrolled 16,415 adults aged 18–74 years residing in four US study sites (Bronx, Chicago, Miami, San Diego) from six Hispanic/Latino ethnic backgrounds (Cuban, Dominican Republic, Mexican, Puerto Rican, Central and South American) at baseline (2008–2011). Sampling design and cohort selection have been described previously [27].

Protocol and measurements

The baseline visit included questionnaires administered in Spanish or English depending on the participant's language preference, along with anthropometry and two 24-h dietary recalls, among other procedures. The 1 st recall was administered in person and the 2nd, preferably ≤ 30 days from the 1 st one, was done via telephone. Both recalls were conducted by trained interviewers using the Nutrition Data System for Research software, version 11 [28]. Virtually all participants (99%) provided at least one recall.

Selection of participants

From the original cohort at baseline (*n* = 16,415), participants were excluded who self-identified as belonging to other/multi/mixed ethnic backgrounds (*n* = 938), had recall data deemed unreliable by the interviewer (*n* = 65), or provided extreme (i.e., < 1st or > 99th sex-specific percentiles) energy intake (*n* = 171). Ethnic background site categories with < 200 participants—"fair" sample size to carry out FA [29]—(i.e., South American – Bronx, *n* = 184) and participants missing Hispanic/Latino background (n = 36) were also excluded. A total of 15,021 participants were included for this baseline analysis, categorized into 12 EBS combinations (Supplementary Fig. 1).

Specification of variables and data preprocessing

From the Nutrition Data System for Research software list of 139 available items [28], we selected 42 nutrients that best represent the overall diet of Hispanics/Latinos (for details, see the manual on diet and supplements at visit 1 available at: https://sites.cscc.unc.edu/hchs/node/4061); specifically, we expanded the fat profile to capture cardiovascular disease-related dietary habits [30, 31] and included only nutrient intake from foods, excluding supplements. We derived the final intake from either a single reliable recall or the average of two reliable recalls collected at baseline. Nutrient intakes were log-transformed (base e) to improve the normality of factor and error terms (details provided in Supplementary Methods).

Statistical analysis

Identification of nutrient-based dietary patterns

After reassuring factorability checks [29], we carried out BMSFA [21] on the EBS-specific log-transformed data correlation matrices to estimate unobservable shared (K) and EBS-specific (J_s) factors, known as DPs. Compared with the frequentist approach [20] [here named frequentist multi-study factor analysis (FMSFA)], BMSFA offers two advantages: 1. a better-defined loading structure via the prior distribution that makes loadings extreme, and 2. the combination of eigenvalue decomposition and 5% variance explained cut-off to choose the number of shared and EBS-specific factors. After estimating the factor loadings, we applied the varimax rotation to the shared factor-loading matrix to obtain a better-defined loading structure. We named "dominant nutrients" [29] those showing shared rotated (or EBS-specific unrotated) factor loadings ≥| 0.60| (≥| 0.30|). Factor scores estimated the extent to which each DP summarizes each participant's diet and were calculated using the Thurstone method [32, 33] (correlation with Bartlett method ≥ 0.90 for the shared factors).

Internal consistency and reproducibility of nutrient-based dietary patterns

We evaluated the internal consistency of DPs using standardized Cronbach's alphas and *alphas-when-item-deleted* [29]. We assessed their internal reproducibility by comparing BMSFA-derived DPs with principal component FA (PCFA)-derived [16] and FMSFA-derived DPs [17, 20]. Throughout the paper, the congruence coefficient (CC) was used to compare pairs of DPs, with the following cut-offs: a 0.85 ≤ CC ≤ 0.94 indicating "fair similarity" and a CC ≥ 0.95 indicating "equivalence" [13] (details provided in Supplementary Methods).

Interpreting dietary patterns in terms of food groups, overall diet quality, socio-demographic and lifestyle characteristics

We separately classified participants based on factor score quintiles (shared factors) or tertiles (EBS-specific factors) to validate DPs against selected 24-h-recall-based food groups, overall diet quality, and socio-demographic and lifestyle factors collected at baseline. For the food-group-based validation, we calculated the percentage deviation of the food-group mean intake in the top-quantile category relative to its overall (i.e., based on the total sample size) or EBS-specific mean, as based on linear regression models adjusted for age, sex, body mass index (BMI), total energy intake, as well as EBS category and the remaining DPs (shared factors only); for each food group, values exceeding 100% indicate higher-than-mean consumption—relative to the HCHS/SOL overall (shared factor) or EBS-specific (EBS-specific factor) mean—and those below 100% indicate lower-than-mean consumption. To provide a comparison between the identified DPs and the Alternative Healthy Eating Index (AHEI- 2010) (11-component diet quality score, range: 0–110, lowest to highest quality) [34], we estimated the adjusted mean AHEI- 2010 score within quantile-based categories of factor scores using a linear regression model adjusted for age, sex, BMI, total energy intake, as well as EBS category and the remaining DPs (shared factors only) (details in Supplementary Methods). For the validation of DPs against selected socio-demographic and lifestyle characteristics, we used the Pearson Chi-square test of independence (categorical characteristics) or the ANOVA (continuous characteristics), with adjustment for multiple comparisons via False Discovery Rate Method [35]. A focus on dietary acculturation [36] was further provided by estimating for each DP the adjusted mean scores within categories of nativity/years lived in mainland US (< 10 years, ≥ 10 years, and US born) using a linear regression model adjusted for age, sex, BMI, total energy intake, as well as EBS category and the remaining DPs (shared factors only).

Except for the implementation of FMSFA and BMSFA, all statistical analyses accounted for HCHS/SOL complex survey design, including survey weights, stratification and clustering [37]. All statistical tests were two-sided. Calculations were carried out using the open-source statistical computing environment R, version 3.6.2 [38], with its libraries "statmod" [39],"psych"[40], "nFactors" [41], "ggplot2" [42], "survey" [43], and "MSFA" [44].

Results

Population characteristics

Socio-demographic and lifestyle characteristics by EBS categories were included in Table 1. The largest and smallest EBS categories were participants of Mexican background from San Diego (n= 3775) and Bronx (n= 205), respectively. Across all EBS categories, most individuals were between 18 and 44 years and reported an income of less than $30,000; in most EBS categories, the majority of individuals were first-generation immigrants and non-consumers of supplements (all p-values < 0.001). Differences between EBS categories were observed for sex, years living in the United States, age of immigration, marital status, employment status, education, physical activity, BMI, and energy intake. The highest percentage of individuals born in mainland US was identified for the Cuban background – Miami category, which also had the highest percentage of not meeting the 2008 Physical Activity Guidelines for Americans and the highest mean energy intake. Individuals of Mexican background from the Bronx had the highest percentage of being 18–44 years old and married, receiving an income < $30,000, and following the 2008 Physical Activity Guidelines for Americans. They also showed almost the lowest mean energy intake. The Mexican background – Chicago individuals had the highest AHEI- 2010 mean value, whereas individuals of Puerto Rican backgrounds from Chicago and the Bronx had the lowest AHEI- 2010 scores, and the lowest mean age at arrival at mainland US. The Mexican background – San Diego category had the highest percentage of individuals in the > $30,000 income category. The South American background – Miami category had the highest percentage of highly educated individuals and individuals who used dietary supplements.

Identification of nutrient-based dietary patterns

Factorability of correlation matrices was confirmed for each of the 12 EBS-specific and overall correlation matrices (Supplementary Results and Supplementary Table 1). The BMSFA estimated four shared DPs, common to all EBS categories (explaining 62.5% of the total variance, Supplementary Table 2) and one EBS-specific DP for each of the 12 categories (10.7%– 14.4% of variance explained, Supplementary Table 3). A heatmap illustrates the factor loadings for shared and EBS-specific DPs (Fig. 1).

Shared dietary patterns

Factor 1, namely *Plant-based foods,* showed high (i.e., dark-blue color in Fig. 1, values in Supplementary Table 2) factor loadings on vegetable protein, phosphorus, magnesium, iron, zinc, copper, potassium, manganese, thiamin, niacin, pantothenic acid, vitamin B6, natural folate, soluble and insoluble dietary fiber. Factor 2, namely *Processed foods,* showed high loadings on long-chain saturated and monounsaturated fatty acids, linoleic and linolenic acid, total trans fatty acids, and natural alpha-tocopherol. Factor 3, namely *Dairy products*,

Table 1 Baseline socio-demographic and lifestyle characteristics (weighted percentages and standard error in parenthesis) by ethnic background and study site (EBS). Hispanic Community Health Study/Study of Latinos, 2008–2011

Characteristic	Total (n = 15,021)	D BX (n = 1362)	CA BX (n = 217)	CA CHI (n = 417)	CA MIA (n = 1012)	Cu MIA (n = 2241)	M BX (n = 205)	M CHI (n = 2403)	M SD (n = 3775)	PR BX (n = 1790)	PR CHI (n = 761)	SA CHI (n = 373)	SA MIA (n = 465)	p-value[a]
Age (y) (%)														
18–44	59.3 (0.8)	63.3 (2.2)	66.1 (4.2)	66.8 (3.4)	63.6 (2.3)	45.2 (1.5)	85.2 (4.1)	72.7 (1.2)	62.7 (1.6)	52.7 (2.0)	55.6 (2.8)	63.9 (4.5)	56.3 (3.2)	<.001
45–55	19.0 (0.5)	18.5 (1.4)	19.0 (3.4)	15.6 (2.3)	18.0 (1.5)	21.6 (0.9)	6.2 (1.6)	15.3 (1.0)	19.2 (1.0)	20.9 (1.3)	22.6 (2.4)	17.7 (2.5)	20.3 (2.0)	
55–74	21.6 (0.6)	18.2 (1.4)	14.9 (2.8)	17.5 (2.3)	18.4 (1.6)	33.2 (1.4)	8.6 (4.0)	12.0 (0.8)	18.1 (1.2)	26.4 (1.7)	21.8 (1.4)	18.4 (3.1)	23.4 (2.7)	
Sex (%)														
Female	52.5 (0.6)	60.1 (1.9)	52.3 (4.5)	45.0 (3.1)	54.9 (2.2)	48.0 (1.1)	53.4 (4.2)	49.6 (1.3)	54.8 (1.3)	51.2 (1.8)	48.0 (2.1)	47.1 (3.4)	60.1 (2.5)	<.001
Immigrant Generation (%)														
First	77.2 (0.8)	82.7 (2.1)	81.3 (4.2)	93.9 (2.3)	96.3 (0.8)	93.5 (0.9)	94.8 (2.1)	84.9 (1.4)	66.6 (1.6)	49.2 (1.8)	46.1 (2.6)	95.1 (2.2)	96.7 (0.9)	<.001
Nativity/Years lived in mainland United States (50 states and DC) (%)														
Born in mainland United States	28.3 (1.0)	24.1 (1.9)	25.1 (5.2)	36.7 (5.5)	41.5 (2.4)	50.2 (1.9)	34.0 (4.8)	25.4 (1.6)	21.7 (1.7)	6.6 (1.0)	3.9 (1.0)	40.3 (4.1)	46.3 (3.5)	<.001
< 10 y	50.2 (0.8)	58.4 (2.2)	56.2 (5.0)	57.7 (5.1)	55.3 (2.2)	43.4 (1.7)	60.8 (4.9)	59.7 (1.5)	48.4 (1.6)	45.6 (1.9)	43.4 (2.6)	54.9 (4.3)	50.4 (3.4)	
≥ 10 y	21.5 (0.8)	17.5 (2.1)	18.7 (4.2)	5.7 (2.3)	3.3 (0.8)	6.4 (0.9)	5.2 (2.1)	14.8 (1.4)	29.9 (1.4)	47.8 (2.0)	52.7 (2.5)	4.9 (2.2)	3.3 (0.9)	
Age of immigration[b] (y)	26.9 (0.3)	25.6 (0.5)	24.9 (1.0)	24.8 (0.7)	27.7 (0.5)	35.3 (0.5)	22.9 (1.0)	23.0 (0.3)	25.1 (0.6)	16.8 (0.8)	17.1 (0.7)	26.5 (1.0)	31.0 (0.9)	<.001
Employment (%)														
Retired and not currently employed	8.2 (0.4)	8.3 (0.9)	6.6 (1.5)	5.2 (1.2)	3.2 (0.7)	11.5 (0.9)	0.8 (0.5)	3.1 (0.4)	5.5 (0.6)	16.9 (1.6)	14.2 (1.7)	3.4 (0.8)	4.0 (1.2)	<.001
Not retired and not currently employed	41.1 (0.7)	42.6 (2.3)	34.9 (4.4)	31.5 (3.6)	38.3 (1.9)	46.7 (1.4)	37.7 (4.8)	31.5 (1.1)	41.7 (1.5)	45.2 (2.1)	38.8 (2.5)	23.1 (2.8)	31.5 (2.6)	
Part-time (≤ 35 h)	16.8 (0.5)	17.2 (1.5)	21.0 (3.5)	22.3 (2.8)	23.3 (1.6)	11.1 (0.7)	18.8 (3.3)	19.4 (1.1)	19.7 (1.1)	11.6 (1.2)	13.4 (2.2)	28.7 (2.7)	24.1 (2.1)	
Full-time (35 + hrs)	33.9 (0.7)	31.9 (2.0)	37.5 (4.5)	41.0 (2.8)	35.2 (1.7)	30.7 (1.2)	42.8 (4.6)	46.0 (1.3)	33.0 (1.5)	26.2 (1.8)	33.6 (2.6)	44.8 (2.7)	40.4 (2.9)	
Marital Status (%)														
Single	33.8 (0.7)	47.6 (1.9)	46.3 (4.4)	27.7 (3.3)	39.9 (1.8)	25.5 (1.3)	27.3 (4.7)	26.2 (1.5)	28.9 (1.4)	51.5 (1.9)	37.8 (2.7)	30.3 (4.6)	30.8 (2.7)	<.001
Married/living with partner	49.9 (0.8)	36.7 (1.8)	42.2 (4.4)	58.7 (3.9)	44.0 (1.8)	52.3 (1.5)	65.8 (4.9)	63.1 (1.5)	57.6 (1.7)	29.9 (1.7)	41.5 (2.9)	56.7 (4.3)	46.5 (3.1)	
Separated/divorced/widowed	16.3 (0.5)	15.7 (1.3)	11.5 (2.2)	13.6 (2.0)	16.1 (1.3)	22.2 (1.2)	6.9 (1.8)	10.7 (0.6)	13.4 (1.0)	18.5 (1.6)	20.7 (2.1)	13.0 (2.1)	22.7 (2.9)	

Table 1 (continued)

Characteristic	Total (n =15,021)	D BX (n =1362)	CA BX (n =217)	CA CHI (n =417)	CA MIA (n =1012)	Cu MIA (n =2241)	M BX (n =205)	M CHI (n =2403)	M SD (n =3775)	PR BX (n =1790)	PR CHI (n =761)	SA CHI (n =373)	SA MIA (n =465)	p-value[a]
Yearly household income (%)														
< $30 k	61.7 (1.0)	66.4 (1.9)	62.5 (4.8)	66.2 (4.0)	69.3 (2.4)	64.5 (1.6)	75.0 (4.1)	65.4 (1.4)	52.6 (2.5)	64.6 (1.9)	55.4 (2.5)	59.2 (4.3)	61.3 (3.2)	<.001
≥ $30 k	31.8 (1.0)	25.9 (1.6)	27.7 (4.2)	30.5 (3.3)	21.4 (2.0)	23.8 (1.4)	17.7 (4.0)	29.7 (1.3)	44.8 (2.5)	29.8 (1.9)	41.6 (2.6)	38.1 (4.4)	33.1 (3.1)	
Not reported	6.4 (0.4)	7.7 (1.0)	9.8 (2.5)	3.3 (1.2)	9.3 (1.1)	11.7 (1.1)	7.4 (1.9)	4.9 (0.6)	2.6 (0.4)	5.6 (0.7)	3.0 (0.8)	2.7 (1.1)	5.7 (1.2)	
Education status (%)														
Less than high school	33.0 (0.7)	37.2 (1.9)	36.0 (4.2)	39.6 (4.2)	39.3 (2.0)	21.6 (1.1)	48.0 (5.3)	49.8 (1.7)	29.2 (1.6)	39.4 (2.1)	31.3 (2.3)	33.0 (4.3)	15.0 (2.3)	<.001
High school or equivalent	28.4 (0.6)	23.5 (1.9)	30.0 (4.6)	19.8 (2.6)	26.3 (1.6)	30.0 (1.4)	36.1 (4.5)	30.5 (1.6)	28.5 (1.3)	26.8 (1.5)	30.4 (2.1)	27.3 (3.2)	25.3 (2.4)	
Greater than high school	38.6 (0.8)	39.3 (1.8)	34.1 (4.6)	40.6 (4.9)	34.3 (1.8)	48.5 (1.5)	16.0 (3.7)	19.8 (1.2)	42.3 (2.0)	33.7 (2.0)	38.2 (2.2)	39.7 (3.7)	60.0 (2.8)	
Met 2008 Physical Activity Guidelines for Americans (%)														
No	57.6 (0.9)	42.9 (2.1)	40.5 (2.2)	53.1 (4.2)	63.1 (2.4)	78.5 (1.5)	28.4 (5.1)	54.9 (1.3)	58.9 (1.6)	43.3 (2.0)	63.0 (3.0)	50.6 (3.3)	65.0 (2.9)	<.001
Dietary Supplement Use (%)														
No	59.3 (0.7)	63.8 (2.1)	66.5 (4.1)	56.6 (4.0)	57.5 (2.1)	57.6 (1.5)	69.6 (5.2)	67.7 (1.3)	52.2 (1.6)	65.5 (1.8)	61.4 (2.0)	62.6 (4.7)	46.7 (2.6)	<.001
Body Mass Index[c] (%)														
Underweight	1.2 (0.1)	1.4 (0.4)	1.6 (1.3)	0.1 (0.1)	1.1 (0.4)	1.5 (0.3)		0.8 (0.3)	1.4 (0.4)	0.9 (0.3)	0.4 (0.3)	2.5 (1.7)	0.6 (0.4)	<.001
Normal	21.6 (0.5)	20.9 (1.5)	21.3 (3.8)	24.2 (2.6)	22.5 (2.0)	23.2 (1.1)	18.0 (3.2)	20.6 (1.0)	21.7 (1.1)	17.5 (1.5)	22.2 (2.5)	25.0 (3.4)	32.2 (3.4)	
Overweight	37.4 (0.6)	36.5 (1.9)	36.9 (4.3)	43.4 (3.3)	38.2 (2.3)	37.4 (1.2)	39.8 (5.0)	37.9 (1.3)	39.4 (1.6)	33.3 (1.8)	31.0 (2.1)	42.7 (2.8)	38.8 (3.1)	
Obesity	39.8 (0.7)	41.2 (2.0)	40.2 (4.6)	32.3 (2.7)	38.2 (1.6)	37.9 (1.3)	42.3 (5.3)	40.7 (1.5)	37.5 (1.7)	48.3 (2.0)	46.4 (2.5)	29.8 (3.3)	28.4 (2.7)	
Energy intake[b] (kcal/day)	1910.1 (10.8)	1620.0 (28.6)	1634.6 (70.4)	1929.6 (44.9)	1910.7 (37.0)	2093.3 (19.9)	1652.1 (61.6)	2000.4 (17.5)	1937.8 (19.5)	1756.1 (24.7)	1982.8 (37.3)	1919.1 (52.5)	2028.8 (39.0)	<.001

Abbreviations: AHEI Alternative Healthy Eating Index, *BX* Bronx, *CA* Central American, *Cu* Cuban, *CHI* Chicago, *D* Dominican; *M* Mexican, *MIA* Miami, *PR* Puerto Rican, *SA* South American, *SD* San Diego

[a] P-values are from Chi-square test of independence for categorical characteristics and from ANOVA for continuous characteristics. P-values were adjusted for multiple comparisons using the False Discovery Rate method

[b] Values are survey weighted means with standard errors indicated in parenthesis

[c] Weight categories were as follows: Underweight (< 18.5 kg/m^2), Normal (18.5–24.9 kg/m^2), Overweight (25–29.9 kg/m^2), and Obesity (≥ 30 kg/m^2)

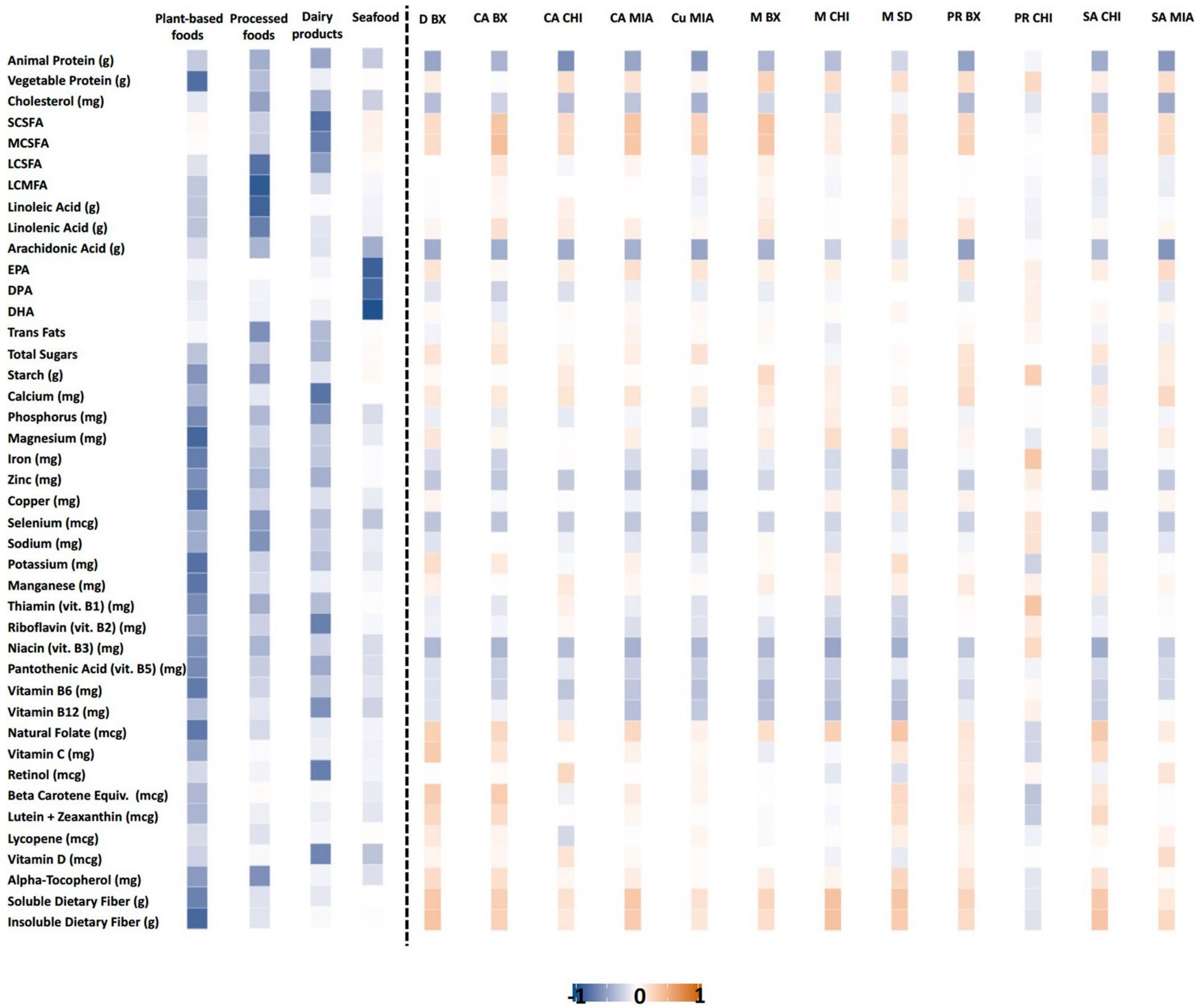

Fig. 1 Heatmap of the estimated factor-loading matrix for the shared and ethnic background site-specific dietary patterns identified with the BMFA. Dashed line indicates the division between the shared and ethnic background site-specific dietary patterns. Hispanic Community Health Study/Study of Latinos, 2008–2011. ABBREVIATIONS: BMSFA: Bayesian multi-study factor analysis; BX: Bronx; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; DHA: docosahexaenoic acid; DPA: docosapentaenoic acid; EPA: eicosapentaenoic acid; M: Mexican; MCSFA: medium-chain saturated fatty acids; LCMFA: long-chain monounsaturated fatty acids; LCSFA: long-chain saturated fatty acids; MIA: Miami; PR: Puerto Rican; SA: South American; SCSFA: short-chain saturated fatty acids; SD: San Diego

showed high loadings on short- and medium-chain saturated fatty acids, calcium, riboflavin, vitamin B12, retinol, and vitamin D. Factor 4, namely *Seafood*, showed high loadings on eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), and docosahexaenoic acid (DHA).

Ethnic background site (EBS)-specific dietary patterns

Ten of the identified EBS-specific DPs represented variants of an animal profile, as they were characterized by two or three nutrients among animal protein, arachidonic acid, and niacin (dark-blue color in Fig. 1). Based on CCs (Table 2) and visual inspection of the factor loadings (Supplementary Table 3), additional profile similarities were related to common ethnic background or site and were expressed through the following three overarching DPs (i.e., DPs shared among some EBS-specific categories):

1. *Animal vs. vegetable source*: animal protein/arachidonic acid *vs.* soluble and insoluble dietary fiber (CC =0.95, indicating equivalence, between DPs of Dominican background – Bronx and South American background – Chicago);

Table 2 Factor congruence coefficients[a,b] between pairs of ethnic background site-specific dietary patterns. Hispanic Community Health Study/Study of Latinos, 2008–2011

	D BX	CA BX	CA CHI	CA MIA	Cu MIA	M BX	M CHI	M SD	PR BX	PR CHI	SA CHI	SA MIA
D BX	1	**0.91**	0.75	**0.91**	0.83	0.76	0.78	0.82	**0.91**	0.49	***0.95***	0.82
CA BX		1	0.74	**0.91**	0.83	0.81	0.69	0.78	**0.88**	0.51	**0.88**	0.76
CA CHI			1	0.84	**0.86**	0.84	0.67	0.51	**0.90**	0.01	0.75	**0.92**
CA MIA				1	**0.93**	**0.92**	**0.87**	0.83	**0.91**	0.35	**0.93**	**0.85**
Cu MIA					1	0.83	0.75	0.64	**0.87**	0.25	**0.88**	**0.89**
M BX						1	**0.85**	0.76	**0.86**	0.14	0.78	0.79
M CHI							1	**0.86**	0.71	0.36	0.84	0.69
M SD								1	0.68	0.61	0.84	0.51
PR BX									1	0.18	**0.85**	**0.93**
PR CHI										1	0.50	0.05
SA CHI											1	0.79
SA MIA												1

Abbreviations: BX Bronx, *CA* Central American, *Cu* Cuban, *CHI* Chicago, *D* Dominican, *M* Mexican, *MIA* Miami, *PR* Puerto Rican, *SA* South American, *SD* San Diego

[a] The congruence coefficient matrix is symmetric with respect to the main diagonal (all 1's); 66 (i.e., number of ethnic background site-specific categories*(number of ethnic background site-specific categories − 1)/2) is the total number of (unique) congruence coefficients between the ethnic background site-specific dietary patterns

[b] Congruence coefficients range between 0 and 1 in absolute value. Values between 0.85 and 0.94 indicate fair similarity between corresponding dietary patterns and were shown in bold typeface in the upper triangular matrix; values ≥ 0.95 indicate equivalence of the corresponding dietary patterns and were shown in bold and italics typeface in the upper triangular matrix

Of the 24 congruence coefficients indicating fair similarity/equivalence between pairs of dietary patterns (in bold typeface in the upper triangular matrix), 11 were underlined in the upper triangular matrix to target similarities between DPs belonging to the same overarching structures

2. *Animal source only*: animal protein/arachidonic acid/cholesterol *vs.* no other dominant nutrients (0.85 ≤ CC ≤ 0.94, indicating fair similarity, for DPs of all the following pairs: Puerto Rican background – Bronx, Central American background – Chicago, Cuban background – Miami, and South American background – Miami);
3. *Poultry vs. dairy products*: less clearly identified and characterized by combinations of animal protein/arachidonic acid/niacin/vitamin B6/vitamin B12 *vs.* small- and medium-chain saturated fatty acids/soluble and insoluble fiber (0.85 ≤ CC ≤ 0.94, indicating fair similarity, for four out of six possible pairs among the following: Central American background – Bronx, Central American background – Miami, Mexican background – Bronx, Mexican background – Chicago).

In contrast, one EBS-specific category, Puerto Rican background – Chicago, was characterized by a strikingly different DP, high on beta-carotene and low on starch, iron, and thiamin (Fig. 1), which showed no fair similarity with any other EBS-specific DP (Table 2). In addition, the Mexican background – San Diego DP loaded high on iron, niacin, vitamin B6, vitamin B12, as opposed to natural folate, soluble and insoluble fiber (Fig. 1), showing fair similarity with the Mexican background – Chicago DP only (Table 2).

In total, 24 (36%) CCs indicated fair similarity/equivalence among EBS-specific DPs, of which 11 (underlined in the table) concerned pairs of DPs belonging to the same overarching structure (Table 2).

Internal reproducibility of the identified patterns

Nutrient communalities and internal consistency of DPs were satisfactory, further supporting our selection of nutrients (Supplementary Table 4 and Supplementary Results).

The internal reproducibility of BMSFA-derived DPs was satisfactory (Fig. 2, Supplementary Results and Supplementary Tables 5–8). The same number of DPs and a similar total variance explained were found for the shared DPs under BMSFA, FMSFA, and PCFA (Supplementary Table 5). The BMSFA-derived shared DPs were also equivalent to their counterparts from FMSFA (all CCs ≥ 0.97) and PCFA (all CCs ≥ 0.95) (Fig. 2). Visual inspection of the factor-loading matrix for the shared DPs suggested that BMSFA was more effective in: 1. shrinking/increasing moderately low/high loadings in absolute value towards 0 or 1, and 2. forcing potentially dominant nutrients to load on one DP instead of two (Supplementary Table 5).

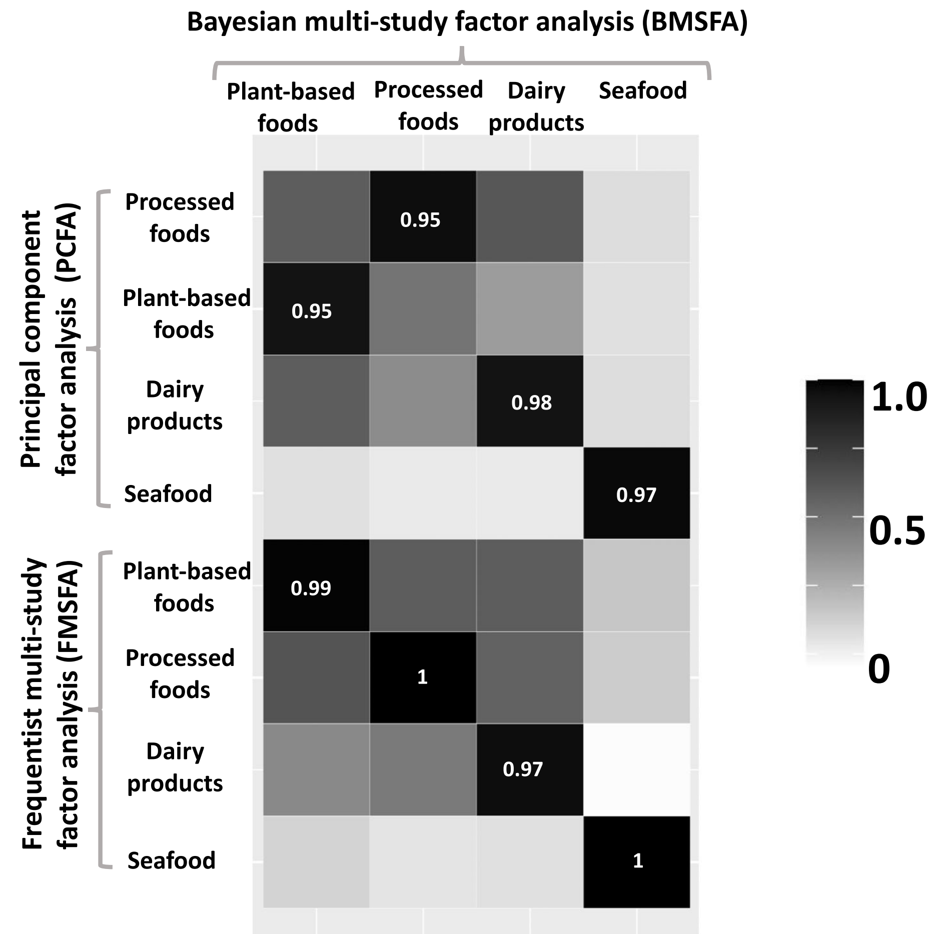

Fig. 2 Heatmap of the factor congruence coefficients[a,b] between the shared dietary patterns identified with the Bayesian and frequentist multi-study factor analysis and the overall-sample dietary patterns from principal component factor analysis. Hispanic Community Health Study/Study of Latinos, 2008–2011. [a]Congruence coefficients range between 0 and 1 (in absolute value), with values between 0.85 and 0.94 indicating fairly similarity and values ≥ 0.95 indicating equivalence of the corresponding dietary patterns. [b]Dietary patterns were ordered in terms of proportion of total variance explained in each solution (see Supplementary Table 5)

While percentages of explained variances were similar to their BMSFA-derived EBS-specific counterparts, FMSFA-derived EBS-specific DPs generally contrasted animal and vegetable food sources (animal protein/cholesterol/arachidonic acid *vs.* vegetable protein/folate/soluble and insoluble fiber in ≥ 8 DPs, Supplementary Table 6). Consequently, the FMSFA-derived EBS-specific DPs showed fewer nuances than the corresponding BMSFA-derived ones. Accordingly, when examining the percentage of pairs of EBS-specific DPs achieving fair similarity or equivalence, FMSFA showed 92%, in contrast to 36% for BMSFA (Supplementary Table 7 *vs.* Table 2). Furthermore, in the one-to-one comparison between the FMSFA- and BMSFA-derived versions of the same EBS-specific DP, only four DPs demonstrated fair similarity across the two approaches (Supplementary Table 8).

Top consumers of shared dietary patterns by ethnic background site category

Ethnic background site categories were well represented in the top-quintile category of each shared DP, with a prevalence around the expected 20% (i.e., 18–22%) for most EBS categories. Major deviations from 18–22% were observed for the *Seafood* DP, with percentages as low as 13.1% (Mexican background – Bronx) and as high as 24.3% (South American background – Chicago) (Supplementary Table 9 and Supplementary Results).

Table 3 Deviation (%) of the adjusted food-group mean for individuals in the top quintile of each shared dietary pattern, relative to the overall adjusted mean. Hispanic Community Health Study/Study of Latinos, 2008–2011.[a,b]

Food group	*Shared dietary patterns*			
	Plant-based foods	*Processed foods*	*Dairy products*	*Seafood*
Fruit—Citrus	**163**	88	120	115
Fruit—Others[c]	**159**	86	98	108
Vegetables—Dark green	**159**	90	104	**152**
Vegetables—Orange	**153**	83	89	121
Vegetables—Tomato	**149**	121	109	102
White Potatoes	130	118	100	123
Vegetables—Starchy	140	**164**	102	95
Vegetables—Beans	**223**	118	78	89
Vegetables—Others[d]	**149**	133	99	130
Refined Grain	127	**153**	122	101
Whole Grain	**149**	90	96	91
Red Meat	124	**169**	111	69
Deli Meat	107	**185**	**141**	101
Poultry	127	135	92	**151**
Fish	133	120	103	**405**
Eggs	98	**148**	127	126
Nuts and Seeds	**235**	**203**	**59**	102
Milk	135	76	**215**	100
Cheese	100	**147**	**214**	90
Yogurt	125	68	**183**	110
Milk-based Dessert	102	**169**	**246**	100
Sugar	112	96	123	99
Dessert	90	139	138	99
Sugar-sweetened Beverage	110	**158**	112	95
Diet Beverage	110	96	89	106
Alcohol	**173**	**153**	101	**152**
Snack—Overall	135	**180**	102	99
Regular fat	125	**181**	99	113
Reduced fat	118	**184**	120	116

[a] Mean intakes of food groups (servings/day) were adjusted for age, sex, body mass index, total energy intake, and ethnic background site category

[b] If the relative consumption of a food group exceeded 100%, it indicates that individuals belonging to the top-quintile category of factor score were characterized by a relatively high consumption of that food group, compared with the reference HCHS/SOL overall mean, and vice versa when the relative intake is below 100%. Percentages below 60% or above 140% were indicated in bold typeface

[c] Fruit – Others group included fruit juice (excluding citrus juice), fruit (excluding citrus fruit), avocado and similar, fried fruits, and fruit-based savory snack

[d] Vegetables – Others group included other vegetables, fried vegetables, vegetable juice, and pickled foods

Food groups associated with the identified dietary patterns

Shared dietary patterns

Table 3 shows the deviation (%) of the adjusted food-group mean for individuals in the top-quintile of each shared DP, relative to the overall adjusted mean. Individuals in the top-quintile category of the *Plant-based foods* DP were characterized by higher-than-mean (i.e., > 140%) intakes of most fruit and vegetables, whole grain, nuts and seeds, and alcohol, compared to the overall HCHS/SOL mean. Individuals in the top-quintile category of *Processed foods* DP were characterized by higher-than-mean intakes of starchy vegetables, refined grain, red meat, deli meat, eggs, nuts and seeds, cheese, milk-based desserts, sugar-sweetened beverages, alcohol, snacks, and added fats. Individuals in the top-quintile category of the *Dairy products* DP showed higher-than-mean consumptions of deli meats, milk, cheese, yogurt, and milk-based desserts; they also showed a lower-than-mean (i.e., < 60%) consumption of nuts and seeds. Finally, individuals in the top-quintile category of the *Seafood* DP presented an extremely high fish intake, and a higher-than-mean consumption of dark green vegetables, poultry, and alcohol.

Ethnic background site (EBS)-specific dietary patterns

Compared with each EBS-specific mean, participants in the top-tertile category of most EBS-specific DPs showed a higher-than-mean consumption of poultry and alcohol *vs.* a lower-than-mean consumption of nuts and seeds, thus confirming the general animal source of most EBS-specific DPs. Red meat was also highly consumed in the top-tertile category of half of them. Overarching DPs additionally showed:

1. *Animal vs. vegetable source*: lower-than-mean consumption of non-citrus fruit (Dominican background – Bronx) or beans (South American background – Chicago);
2. *Animal source only*: higher-than-mean consumption of red meat (except for South American background – Miami, which reached 132%) *vs.* no lower-than-mean consumption of any food groups;
3. *Poultry vs. dairy products*: lower-than-mean consumption of one or more among cheese, yogurt, and milk-based desserts; however, lower-than-mean consumption of dairy products was also common to individuals of Central American background from all sites and those of Cuban background from Miami, who belonged to the *Animal source only* overarching DP.

Top-consumers of the Puerto Rican background – Chicago category showed a higher-than-mean consumption of yogurt, milk-based desserts, (non-citrus) fruit, dark green and orange vegetables; this was different from any other EBS-specific DP, including that of their Bronx counterparts of Puerto Rican background. Finally, individuals of Mexican background from Chicago and San Diego showed a similar profile, including red meat, poultry, and alcohol, but fewer nuts and seeds, and beans. However, no similarities were found with the few Mexican background individuals from Bronx, whose top-consumers consumed citrus fruit and sugar-sweetened beverages (Table 4).

Alternative Healthy Eating Index associated with the identified patterns

Shared dietary patterns

Higher quintile-based categories of *Plant-based foods* and *Seafood* DPs were consistently and significantly associated with increased mean AHEI- 2010 scores up to a ~ 5-point increment (p for trend < 0.001), suggesting a higher overall diet quality for top-consumers of these DPs. *Dairy products* and *Processed foods* DPs showed the opposite trend (p for trend < 0.001): lower mean AHEI- 2010 scores (i.e., lower quality) were consistently reported for increasing quintile-based categories of these DPs, with an overall ~ 3- and ~ 5-point decrease, respectively (Fig. 3).

Ethnic background site (EBS)-specific dietary patterns

Lower mean AHEI- 2010 scores were observed across increasing tertile-based categories of 11 EBS-specific combinations (all p for trend < 0.05, except for Central American background – Bronx: $p = 0.12$). In contrast, participants of Puerto Rican background from Chicago exhibited an increase of approximately 2 points in mean AHEI- 2010 scores with the highest DP category. In addition, for the same tertile-based category, mean AHEI-2010 scores markedly differed across EBS-specific DPs, with Puerto Rican background participants from Bronx and Chicago showing the lowest mean scores [39.85 (SE: 0.20) and 42.15 (SE: 0.22), respectively] and Mexican background participants from Chicago and San Diego showing the highest mean scores [56.26 (SE: 0.15) and 54.27 (SE: 0.22), respectively] (Fig. 3).

Socio-demographic and lifestyle factors associated with the identified patterns

Shared dietary patterns

While top-consumers of all shared DPs were more likely to be younger and males, percentages of the youngest or male participants were even more extreme in top-consumers of the *Processed foods* DP. Notably, top-consumers of the *Plant-based foods* and *Seafood* DPs were less likely to have spent ≥ 10 years in the US, whereas those of the *Dairy products* and *Processed foods* DPs were more likely to have spent ≥ 10 years in the US; top-consumers of the *Seafood* DP only were more likely to be born in the United States. Top-consumers of the *Processed foods* DP showed a lower weighted mean age at immigration. While most of the overall population did not meet the 2008 Physical Activity Guidelines for Americans, top-consumers of the *Plant-based foods* and *Processed foods* DPs were more likely to be active. Use of supplements was more likely in the top-consumers of the *Plant-based foods* and *Dairy products* DPs and less likely in top-consumers of the *Processed foods* DP (Table 5, with most adjusted p-values < 0.05).

When investigating the role of nativity/years lived in mainland US in regression models, adjusted mean scores of the *Plant-based foods* and *Dairy products* DPs showed a similar trend: mean scores were negative for US born participants, but they became positive and further increased as increasing years were spent in the US. Also,

Table 4 Deviation (%) of the adjusted food-group mean for individuals in the top tertile of each ethnic background site-specific dietary pattern, relative to the ethnic background site-specific adjusted mean. Hispanic Community Health Study/Study of Latinos, 2008–2011.[a,b]

	Ethnic background site (EBS)-specific dietary patterns											
Food group	**D BX**	**CA BX**	**CA CHI**	**CA MIA**	**Cu MIA**	**M BX**	**M CHI**	**M SD**	**PR BX**	**PR CHI**	**SA CHI**	**SA MIA**
Fruit—Citrus	82	79	108	96	93	**175**	94	64	66	135	76	134
Fruit—Other[c]	**56**	60	74	83	84	117	77	71	96	**151**	77	89
Vegetables—Dark green	66	**42**	129	72	91	73	91	60	71	**205**	65	125
Vegetables—Orange	81	80	116	105	93	103	105	69	81	**147**	83	127
Vegetables—Tomato	83	80	106	105	110	124	94	89	79	110	91	97
White Potatoes	88	113	**157**	119	123	128	82	91	105	101	91	110
Vegetables—Starchy	86	106	80	82	90	72	114	91	90	126	94	104
Vegetables—Beans	100	107	82	62	81	64	**43**	**44**	85	98	**56**	**59**
Vegetables – Other[d]	81	78	112	111	106	123	101	82	88	118	81	101
Refined Grain	121	99	93	106	102	94	119	109	93	65	116	98
Whole Grain	79	84	108	93	94	**44**	73	86	81	115	81	100
Red Meat	**157**	134	**150**	**149**	**160**	138	**141**	138	**152**	94	137	132
Deli Meat	124	67	92	114	113	93	114	100	113	74	**145**	119
Poultry	**156**	**178**	**190**	**144**	**142**	**142**	**150**	129	**153**	105	**143**	**167**
Fish	80	105	106	64	77	137	86	75	**55**	72	77	63
Eggs	111	124	93	119	126	94	98	82	125	104	95	**150**
Nuts and Seeds	**18**	**16**	**20**	**53**	**47**	**5**	**55**	62	65	105	**28**	**36**
Milk	76	78	66	95	93	96	90	103	81	105	91	79
Cheese	84	**58**	**57**	62	76	70	80	69	87	84	**47**	93
Yogurt	67	**0**	90	139	80	109	69	93	**46**	**151**	**27**	85
Milk-based Dessert	65	**47**	**59**	**43**	**57**	80	**177**	67	100	**150**	75	80
Sugar	86	87	87	74	82	78	83	87	93	108	82	98
Dessert	88	67	100	**42**	68	83	135	91	67	94	64	97
Sugar-sweetened Beverage	120	113	113	111	97	**153**	127	128	106	76	114	104
Diet Beverage	103	102	95	88	96	100	107	82	90	100	100	100
Alcohol	116	90	**228**	**196**	**143**	**282**	**173**	**150**	**160**	90	**158**	106
Snack—Overall	63	71	65	65	65	**58**	87	76	64	88	71	76
Regular fat	103	82	106	94	105	89	102	76	91	97	110	101
Reduced fat	100	116	87	105	101	83	129	104	94	102	111	103

Abbreviations: BX Bronx, *CA* Central American, *Cu* Cuban, *CHI* Chicago, *D* Dominican, *M* Mexican, *MIA* Miami, *PR* Puerto Rican, *SA* South American, *SD* San Diego

[a] Mean intakes of selected food groups (servings/day) were adjusted for age, sex, body mass index, and total energy intake

[b] If the relative consumption of a food group exceeded 100%, it indicates that individuals belonging to top-tertile category of factor score were characterized by a relatively high consumption of that food group, compared with the ethnic background site-specific overall mean, and vice versa when the relative intake is below 100%. Percentages below 60% or above 140% were indicated in bold typeface

[c] Fruit – Others group included fruit juice (excluding citrus juice), fruit (excluding citrus fruit), avocado and similar, fried fruits, and fruit-based savory snack

[d] Vegetables – Others group included other vegetables, fried vegetables, vegetable juice, and pickled foods

adjusted mean scores of the *Processed Foods* and *Seafood* DPs showed similar trends: US-born participants had higher (2–threefold, positive) mean scores compared to participants with less than 10 years lived in the US; however, this behavior was less pronounced or lost (i.e., scores became negative) in those living 10 or more years in the US (Fig. 4).

Ethnic background site (EBS)-specific dietary patterns

The identified EBS-specific DPs were sparingly related to the selected socio-demographic and lifestyle factors. The two DPs expressed by Chicago and San Diego individuals of Mexican background were significantly related to most of the selected variables in the same direction: top consumers of these DPs were more likely to be: younger, first generation immigrants, who have lived less than 10 years in the US, married/living with a partner, reporting

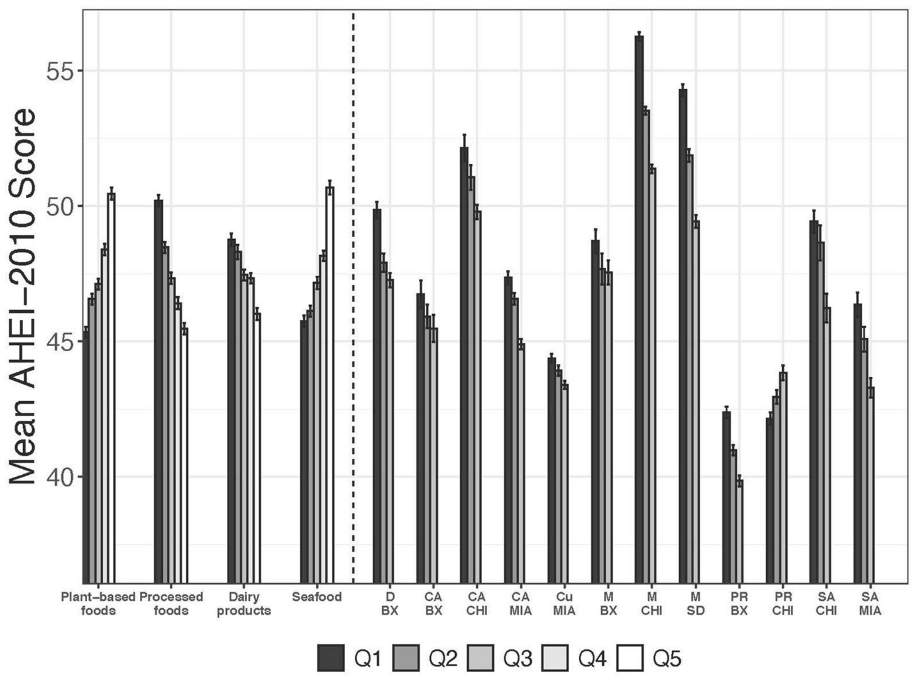

Fig. 3 Mean AHEI- 2010 scores (standard errors) by quintiles of shared dietary patterns and by tertiles of ethnic background site-specific dietary patterns[a,b,c]. Hispanic Community Health Study/Study of Latinos, 2008–2011. ABBREVIATIONS: AHEI: Alternative Healthy Eating Index; BX: Bronx; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego. [a]Weighted mean AHEI- 2010 scores were adjusted for age, sex, body mass index, and total energy intake; models for the shared dietary patterns were further adjusted for ethnic background site categories and for the other shared dietary patterns (quintiles). [b]The AHEI- 2010 measures overall diet quality in terms of adherence to the Dietary Guidelines for Americans 2010; compared to the Healthy Eating Index 2010, it also incorporates additional components that focus on foods and nutrients to predict the risk of chronic disease. The total AHEI- 2010 score ranges from 0 to 110, with higher scores indicating a healthier diet. [c]Shared patterns are shown in quintile-based categories, while ethnic background site-specific patterns are shown in tertile-based categories. The same color scale was adopted for the first three quintile-based categories and for the three tertile-based categories

an income less than $30,000, and having less than a high school education (Supplementary Table 10 and Supplementary Results).

Adjusted mean scores for nativity/years lived in mainland US were the highest (up to ~5 folds) for US-born participants in six EBS categories, including Mexican background participants from Chicago and San Diego (but not from Bronx), Dominican background participants from Bronx, Cuban background participants from Miami, as well as Central American background participants from Bronx and Miami (but not from Chicago). Among those participants, differences were found by years lived in mainland US. Specifically:

1. in Mexican background participants from Chicago (p for trend <0.001) and San Diego (p for trend $=0.037$) and Dominican background participants from Bronx (p for trend $=0.09$), mean scores were still high and positive in participants who have lived less than 10 years in the US, but were negligible or negative after 10 years from migration;
2. in Central American background participants from Bronx and Miami, mean scores switched from negative to positive for increased length of US residence (p for trend <0.05 for both), suggesting increased adaptation towards US-born-like behaviors.

Their counterparts from Chicago (i.e., Central American background – Chicago) exhibited the opposite trend, with adjusted mean scores shifting from positive (< 10 and ≥ 10 years spent in the US) to negative (US-born). Finally, Puerto Rican background participants from Chicago still relate with nativity/years lived in mainland US in a different way than any other category, including their Bronx counterparts. Their adjusted mean scores were negative across all categories of nativity/years lived in mainland US, with the most pronounced difference observed between those who had lived in US for <10 years or were US-born compared to those who had spent ≥ 10 years in the US (Fig. 4).

Table 5 Socio-demographic and lifestyle characteristics (weighted percentages and standard error in parenthesis) for participants by shared quintile-based (Q1, Q5) categories. Hispanic Community Health Study/Study of Latinos, 2008–2011

	Plant-based foods			*Processed foods*			*Dairy products*			*Seafood*		
	Q1	Q5	p-value[a]	Q1	Q5	p-value[a]	Q1	Q5	p-value[a]	Q1	Q5	p-value[a]
Age (y) (%)												
18–44	38.3 (1.9)	42.4 (2.2)	< 0.001	30.8 (1.6)	55.9 (2.3)	< 0.001	35.0 (1.8)	41.5 (2.2)	0.002	39.6 (2.1)	39.8 (1.7)	< 0.001
45–55	25.8 (1.5)	24.6 (1.3)		23.8 (1.4)	26.8 (1.7)		28.2 (1.5)	25.0 (1.6)		27.0 (1.5)	26.5 (1.5)	
55–74	36.0 (1.7)	33.0 (2.0)		45.3 (1.7)	17.2 (1.7)		36.7 (1.6)	33.5 (2.0)		33.4 (1.9)	33.7 (1.7)	
Sex (%)												
Female	69.9 (1.6)	41.9 (1.9)	< 0.001	79.5 (1.4)	24.1 (1.5)	< 0.001	56.2 (1.4)	49.1 (2.0)	< 0.001	55.8 (1.9)	47.8 (1.7)	< 0.001
Immigrant Generation (%)												
First	98.9 (0.3)	97.9 (0.7)	< 0.001	98.5 (0.4)	97.8 (0.7)	< 0.001	98.5 (0.4)	98.0 (0.7)	< 0.001	97.3 (0.8)	97.9 (0.7)	< 0.001
Nativity/Years lived in mainland United States (50 states and DC) (%)												
Born in mainland United States	28.2 (1.4)	27.8 (1.7)	< 0.001	25.7 (1.3)	25.9 (1.4)	< 0.001	28.2 (1.7)	27.5 (1.6)	0.003	25.3 (1.4)	28.6 (1.5)	< 0.001
< 10 y	45.9 (1.3)	53.6 (1.7)		61.6 (1.3)	42.1 (1.4)		53.7 (1.6)	48.4 (1.5)		46.6 (1.6)	54.5 (1.4)	
≥ 10 y	25.9 (1.4)	18.6 (1.3)		12.7 (0.9)	32.0 (1.6)		18.1 (1.3)	24.1 (1.4)		28.1 (1.5)	16.9 (1.1)	
Age of immigration[b] (y)	30.6 (1.6)	30.1 (1.6)	0.007	31.7 (0.6)	26.5 (0.6)	< 0.001	31.1 (0.6)	30.7 (0.7)	0.748	30.5 (0.6)	30.2 (0.6)	0.016
Employment (%)												
Retired and not currently employed	13.2 (1.2)	11.4 (1.2)	< 0.001	18.4 (1.3)	4.3 (0.8)	< 0.001	12.2 (1.1)	12.9 (1.3)	0.47	11.7 (1.2)	12.1 (1.2)	0.025
Not retired and not currently employed	40.9 (1.7)	35.2 (1.9)		41.8 (1.7)	29.1 (1.8)		38.4 (1.6)	36.3 (2.2)		38.6 (1.8)	33.9 (1.5)	
Part-time (≤ 35 h)	15.6 (1.4)	15.8 (1.3)		15.8 (1.1)	15.8 (1.5)		16.1 (1.1)	15.2 (1.3)		15.1 (1.4)	16.9 (1.2)	
Full-time (35 + hrs)	30.3 (1.6)	37.6 (1.9)		24.1 (1.4)	50.9 (2.3)		33.3 (1.7)	35.7 (1.9)		34.6 (1.7)	37.0 (2.0)	
Marital Status (%)												
Single	17.9 (1.3)	16.0 (1.4)	0.011	16.6 (1.4)	22.2 (1.8)	< 0.001	18.5 (1.3)	17.4 (1.5)	0.028	19.2 (1.6)	17.1 (1.2)	< 0.001
Married/living with partner	57.4 (1.8)	64.3 (1.8)		58.2 (1.7)	62.0 (2.2)		57.6 (1.7)	63.6 (2.1)		60.2 (1.9)	61.2 (1.9)	
Separated/divorced/widowed	24.6 (1.5)	19.7 (1.4)		25.1 (1.4)	15.8 (1.5)		23.9 (1.3)	18.9 (1.4)		20.6 (1.4)	21.7 (1.5)	
Yearly household income (%)												
< $30 k	71.8 (1.8)	64.8 (2.1)	< 0.001	69.6 (1.8)	62.9 (2.1)	< 0.001	67.9 (1.8)	63.4 (2.6)	0.012	60.7 (2.0)	64.1 (2.1)	< 0.001
≥ $30 k	19.9 (1.6)	31.8 (2.1)		24.4 (1.6)	32.9 (2.0)		25.6 (1.8)	31.0 (2.7)		32.6 (1.9)	31.7 (2.3)	
Missing	8.3 (1.0)	3.3 (0.7)		6.0 (0.7)	4.2 (0.8)		6.5 (0.9)	5.6 (1.1)		6.6 (1.0)	4.1 (0.7)	
Education status (%)												
Less than high school education	42.2 (1.7)	36.1 (2.0)	0.027	41.6 (1.8)	36.3 (2.3)	0.003	44.9 (1.7)	35.0 (2.1)	< 0.001	39.6 (1.9)	36.3 (1.8)	0.783
	25.4 (1.5)	25.6 (1.5)		21.7 (1.2)	27.4 (1.9)		22.9 (1.3)	26.1 (1.9)		27.4 (2.0)	25.9 (1.8)	
High school education or equivalent												
Greater than high school education	32.4 (1.7)	38.3 (1.9)		36.6 (1.8)	36.3 (2.2)		32.2 (1.6)	38.9 (2.2)		33.0 (1.8)	37.8 (2.1)	
Met 2008 Physical Activity Guidelines for Americans (%)												
No	69.6 (1.6)	56.9 (2.1)	< 0.001	66.8 (1.7)	52.4 (2.2)	< 0.001	64.8 (1.7)	64.1 (2.1)	0.07	66.0 (1.9)	61.2 (2.1)	0.706

Table 5 (continued)

	Plant-based foods			*Processed foods*			*Dairy products*			*Seafood*		
	Q1	**Q5**	**p-value[a]**	**Q1**	**Q5**	**p-value[a]**	**Q1**	**Q5**	**p-value[a]**	**Q1**	**Q5**	**p-value[a]**
Supplemental Use (%)												
No	59.3 (1.8)	47.6 (2.0)	< 0.001	45.0 (1.6)	61.6 (2.1)	< 0.001	57.5 (1.7)	46.6 (2.1)	< 0.001	50.2 (2.0)	50.3 (1.6)	0.053
Body Mass Index[c] (%)												
Underweight	1.2 (0.3)	1.1 (0.3)	0.086	0.4 (0.2)	1.9 (0.5)	0.002	1.1 (0.2)	1.3 (0.4)	0.349	1.6 (0.4)	0.9 (0.4)	0.079
Normal	20.0 (1.2)	23.0 (1.2)		18.6 (1.0)	23.6 (1.1)		19.9 (1.0)	24.6 (1.3)		23.0 (1.2)	18.8 (1.0)	
Overweight	35.6 (1.3)	39.5 (1.4)		37.6 (1.3)	34.9 (1.4)		37.9 (1.4)	35.6 (1.4)		35.1 (1.4)	39.0 (1.5)	
Obesity	43.3 (1.6)	36.6 (1.5)		43.3 (1.4)	39.6 (1.5)		41.0 (1.4)	38.5 (1.6)		40.3 (1.4)	41.3 (1.5)	
Energy intake[b] (kcal/day)	1385 (15.2)	2443 (26.7)	< 0.001	1286 (13.3)	2621 (24.5)	< 0.001	1615 (19.9)	2259 (22.6)	< 0.001	1951 (21.5)	1941 (23.0)	0.047

[a] *P*-values are from Chi-square test of independence for categorical characteristics and from ANOVA for continuous characteristics. *P*-values were adjusted for multiple comparisons using the False Discovery Rate method. The adjustment was carried out considering all dietary patterns simultaneously

[b] Values are weighted means with standard errors indicated in parenthesis

[c] Weight categories were as follows: Underweight (< 18.5 kg/m^2), Normal (18.5–24.9 kg/m^2), Overweight (25–29.9 kg/m^2), and Obesity (≥ 30 kg/m^2)

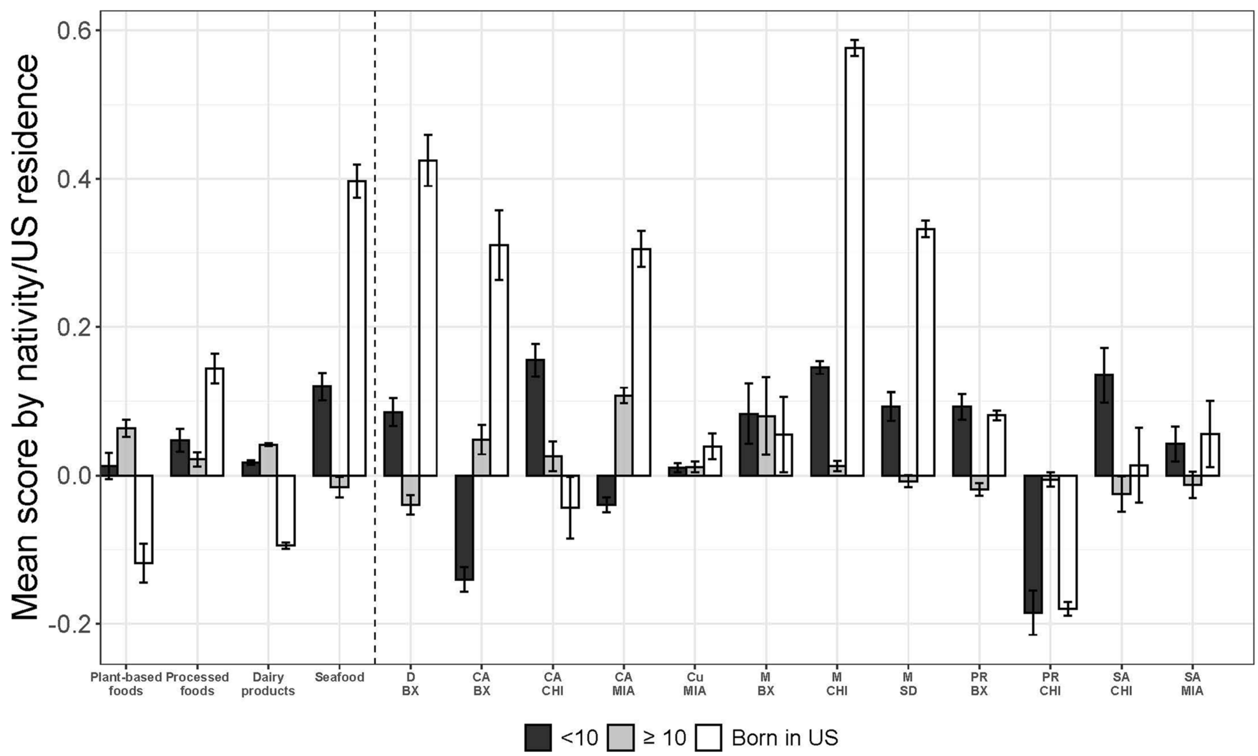

Fig. 4 Mean dietary pattern scores (standard errors) for shared and ethnic background site-specific dietary patterns by nativity/years lived in mainland US in categories[a,b]. Hispanic Community Health Study/Study of Latinos, 2008–2011. ABBREVIATIONS: BX: Bronx; CA: Central American; Cu: Cuban; CHI: Chicago; D: Dominican; M: Mexican; MIA: Miami; PR: Puerto Rican; SA: South American; SD: San Diego. [a]Weighted mean dietary pattern scores were adjusted for age, sex, education, and total energy intake; models for the shared dietary patterns were further adjusted for ethnic background site categories and for the other shared dietary patterns (continuous). [b]The same color scale was adopted for the first three quintile-based categories and for the three tertile-based categories

Discussion

With the application of BMSFA, this paper identified four shared and reproducible *a posteriori* DPs (*Plant-based foods*, *Processed foods, Dairy products*, and *Seafood*) across the 12 EBS categories available in HCHS/SOL. Most of the additional 12 EBS-specific DPs—one for each category—represented variants of an animal profile associated with poultry consumption and were further grouped into three overarching DPs: *Animal vs. vegetable source*, *Animal source only*, and *Poultry vs. dairy products*. The BMSFA-derived shared DPs were equivalent to their counterparts under PCFA and FMSFA; however, the BMSFA-derived EBS-specific DPs showed more nuances than the FMSFA-derived ones. The identified DPs were associated with selected food groups, the overall diet quality as measured with the AHEI- 2010, and socio-demographic and lifestyle factors, including acculturation. A higher overall diet quality was observed for increasing score categories of the *Plant-based foods* and *Seafood* DPs. The *Dairy products*, *Processed foods*, and 11 EBS-specific DPs showed the opposite trend. When evaluating nativity/years lived in mainland US, the *Plant-based foods* and *Dairy products* DPs showed the highest negative adjusted means in US-born vs non-US-born participants, whereas the *Processed foods*, *Seafood*, and six EBS-specific DPs showed the highest positive adjusted means in US-born participants. Puerto Rican background participants from Chicago expressed a strikingly different DP, as their EBS-specific DPs was characterized by yogurt, milk-based desserts, and selected fruits and vegetables. Consistently, this DP showed opposite trends compared to all other subpopulations. It was directly related with overall diet quality and showed negative adjusted means across all categories of nativity/years lived in mainland US.

The shared factors reflect common DPs among Hispanic/Latino adults in the US, regardless of ethnic background and study site. These DPs closely align with general Hispanic/Latino culinary traditions [45], while also reflecting acculturation to more US-American diets [46]. Although they predominantly targeted vegetable and animal sources of foods, and their fat components, they excluded grains (represented by starch, total sugars, and/or fiber) or meat (represented by animal protein and/or cholesterol). This may seem unexpected, given the central role of cereals and meat in Hispanic/Latino diets [47–49]. However, our analysis indicates that none of the nutrient combinations targeting grains or meat, except for the *Processed foods* DP, were consumed strongly enough to emerge as an additional shared DP.

While primarily summarizing animal-based profiles, the 12 subpopulation-specific DPs reflect a blend of background-specific (e.g., culinary traditions, cultures, beliefs) [45] and site-specific (e.g., food access and environment) [50] factors. Within overarching DPs, EBS-specific categories vary by ethnic background and/or study site. Even outside any overarching structure, individuals of Puerto Rican background from Bronx and Chicago showed distinct DPs, and individuals of Mexican background from San Diego and Chicago shared similar DPs, but not with those from the Bronx. These findings highlight the importance of analyzing data at the cross-section of ethnic background and site and using *a posteriori* DPs, which revealed site-specific differences by background—insights that were not captured by previous AHEI- 2010-based analyses [51].

A few papers presented *a posteriori* DPs derived with factor or cluster analyses on Hispanics/Latinos living in the US [24, 25, 52–56]. Earlier than HCHS/SOL, two of these studies [52, 54] derived DPs on Hispanics/Latinos alone. They proposed a cluster analysis on Mexican American adults from the US National Health and Nutrition Examination Survey (NHANES) 2001/2002 [52] and a FA on Puerto Rican adults from the Boston area [54]. In the former study [52], the *Traditional Mexican* and the *Meat* DPs present grains (tacos/tortillas) on two separate DPs, consumed in combination with either legumes or (red) meat. Their *Poultry and alcohol* DP resembles most of our EBS-specific DPs. Dairy products loaded high on a separate animal-source *Milk and baked products* DP, which also showed similarities with our *Processed foods* DP due to the presence of cakes, cookies, and pizza. In the other study [54], the *Meat, processed meat, and French fries* and the *Sweets, sugary beverages, and dairy desserts* DPs share similarities with our *Processed foods* DP. However, while our DP reflects acculturation to typical dietary behaviors seen in US adults, the previous DPs likely capture modern industrialized diets related to nutrition transition, as they were shown not to be related to acculturation [57]. While revealing the major role of oils, the *Rice, beans, and oils* DP also confirms the rice and legumes' combination typically found in a Hispanic/Latino diet.

Within HCHS/SOL, two papers derived *a posteriori* DPs at either ethnic background-specific [24] or EBS-specific [25] levels. Maldonado et al. [24] described one "fully" (i.e., shared among all ethnic backgrounds) and four "partially" reproducible DPs derived using FA on 34 food groups from 24-h recalls stratifying by six ethnic background-specific categories. Their fully reproducible *Burgers, Fries, & Soft drinks* DP is fairly similar to our shared *Processed foods* DP. Consistent with our *Seafood* DP, Maldonado's *Fish* DP [24] loaded high on fish and to less extent poultry; similarly to our EBS-specific DP for Dominican background – Bronx, fish was also opposed to poultry in their Dominican participants. Our *Dairy products* DP shares similarities with Maldonado's *Egg & Cheese* DP [24], although the additional presence of starchy vegetables and processed meats also suggested overlapping with our *Processed foods* DP for Dominican and Puerto Rican backgrounds. Maldonado's *White Rice, Beans, & Red Meats* partially overlaps with our *Plant-based foods* DP; however, ours included 7 additional vegetable and fruit groups and was more focused on whole rather than refined grains. Whereas Maldonado et al. [24] derived DPs separately by ethnic background using standard FA, Stephenson et al. [25] used robust profile clustering to jointly classify individuals and 129 food propensity questionnaire items into global or local food patterns, based on 9 of our 12 EBS categories. Stephenson's *Global Profile 1* favored a more frequent consumption of fruits, vegetables, poultry, and fish and therefore contained a mix of elements from our *Plant-based foods* and *Seafood* DPs. Their *Global Profile 2* favored a more frequent consumption of foods with oils, added sugars, and eggs, and contained many elements of our *Processed foods* DP. Like our BMSFA results, poultry was commonly consumed across all EBS categories but with different frequencies [25].

Finally, a recent article [56] assessed reproducibility of nutrient-based DPs of Mexican or other Hispanic adult participants from 2007–2012 NHANES and 2008–2011 HCHS/SOL by applying separate standard FAs for each study (NHANES and HCHS/SOL) but not further by ethnicity or geography. Among the five DPs identified as similar across studies [56], the *Dairy* and *Fats/oils* DPs share similarities with our *Dairy products* and *Processed foods* DPs, respectively. The joint presence of most minerals, group B vitamins, and fiber suggests some similarities exist between their *Grains/legumes* [56] and our *Plant-based foods* DP, although starch was only 0.57 in our analysis (Supplementary Table 2). While vitamin C plus carotenoids and animal protein plus cholesterol did not load high on any of our DPs, Varela et al. [56] also identified a *Fruits/veggies* and a *Meats* DPs loading high on most carotenoids and total protein, respectively. In contrast, by separating EPA, DHA, and DPA from total PUFA, we captured the *Seafood* DP not identified in Varela et al. [56].

Unadjusted analyses on single DPs and age at immigration or nativity/years lived in mainland US suggest that dietary practice outcomes worsen with years lived in mainland US, in line with others [36, 58]. When multiple regression models on nativity/years lived in mainland US were adopted, results aligned more with those suggesting nativity is per se a strong predictor (e.g. [59, 60]) or a stronger predictor of dietary acculturation compared with years spent in the US (e.g. [55]). Currently, there is still little agreement on how acculturation should be measured [57, 61, 62], and further research is needed to consider the complex interplay between socio-demographic factors, DPs, acculturation, and/or health outcomes [63], likely within a mediation analysis approach [64].

The current analysis has strengths and limitations. The HCHS/SOL is to date the largest and most comprehensively characterized cohort of adults of diverse Hispanic and Latino background in the United States. The sampling design implemented at each study site—where the sample distributions mirrored the Hispanic/Latino background concentrations in that area by age, sex, and country of origin—is superior to the convenience samples typically used in epidemiologic cohort studies [23]. Methodologically, MSFA identified shared and EBS-specific DPs in a single step within one statistical model by improving over the two extreme and unrealistic scenarios of no sharing or complete sharing of DPs across subpopulations. Compared to FMSFA, the introduction of prior distributions (which act like rotations) in BMSFA has provided a clearer interpretation of the EBS-specific DPs observed, while maintaining equivalence with the shared DPs from FMSFA or PCFA. Among limitations, although the study design allows inference to the target areas surrounding the study sites, results are not generalizable to the total Hispanic/Latino adult population in the US [23]. However, the four study sites are located among cities with highest concentrations of Hispanics/Latinos. Based on 24-h recall data, our identified DPs may not represent the participant's usual diet and may have failed to capture episodically consumed foods well. In general, self-reported dietary assessment tools are prone to measurement error, and 24-h recalls have been shown to underestimate total dietary intake [65]. Systematic under-reporting of energy and protein intake was observed in a biomarker calibration study in HCHS/SOL that varied by ethnic background [66], which may explain some of the lower mean energy values in certain EBS categories. Finally, although we introduced percentage deviations and stringent cut-offs to enhance the interpretation and comparison of DPs, we acknowledge that the more comprehensive analysis provided by the proposed advanced method has increased the complexity of interpreting the identified DPs.

Conclusions

This application of BMSFA within HCHS/SOL reveals shared and EBS-specific DPs that clearly document the dietary habits of Hispanic/Latino adults residing in the US. The shared DPs were equivalent to those identified by traditional (PCFA) and novel (FMSFA) approaches and closely align with general Hispanic/Latino culinary traditions, while also reflecting acculturation to more US-American diets. While primarily summarizing animal-based profiles, the EBS-specific DPs reflect a blend of influence based on a participant's ethnicity and/or geographic residency. Although our results were reasonable, implications for potential preventive strategies require further assessment of more complex patterns including acculturation, socio-demographic factors, and/or health outcomes.

Abbreviations

BMI	Body mass index
BMSFA	Bayesian multi-study factor analysis
DP	Dietary pattern
EBS	Ethnic background and site
FA	Factor analysis
FMSFA	Frequentist multi-study factor analysis
HCHS/SOL	Hispanic Community Health Study/Study of Latinos
MSFA	Multi-study factor analysis
PCA	Principal component analysis
PCFA	Principal component factor analysis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12937-025-01138-0.

Supplementary Material 1.

Acknowledgements

The authors thank the staff and participants of HCHS/SOL for their important contributions. Investigators website—http://www.cscc.unc.edu/hchs/.

Authors' contributions

RDV, BJS, and VE designed overall research plan; DS-A: assisted with the research plan and writing of the manuscript; AMS-R, JM, and MP provided nutritional expertise on results of the study and a critical review of the manuscript; RDV performed most of the statistical analyses, with contributions from BS; RDV prepared all tables and figures except for Online Resource Fig. 1; RDV assisted in writing the Methods section and provided critical review of the manuscript; BS assisted in writing the Results section and provided critical review of the manuscript; BP, SAB, MLD, LVH assisted with interpretation of results and provided critical review of the manuscript; VE wrote the paper; VE and RDV had primary responsibility for final content. AMS-R, MLD, and LVH were involved in obtaining the cohort funding; AMS-R, MLD, LVH, and DS-A were involved in conducting the study including dietary data collection. All authors reviewed and approved the final manuscript.

Funding

This work was supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I/N01-HC- 65233), University of Miami (HHSN268201300004I/N01-HC- 65234), Albert Einstein College of Medicine

(HHSN268201300002I/N01-HC- 65235), University of Illinois at Chicago (HHSN268201300003I/N01- HC- 65236 Northwestern Univ), and San Diego State University (HHSN268201300005I/N01-HC- 65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, National Institutes of Health (NIH) Institution-Office of Dietary Supplements. Roberta De Vito was supported by US NIH grant NIGMS/NIH COBRE CBHD P20GM109035. Sierra Bainter was supported by US NIH grant K01MH122805. Valeria Edefonti was supported by the Young Investigator Grant 2020, from Università degli Studi di Milano and from the Italian Ministry of University and Research (grant numbers PRIN 20227YCB5P, MUR Progetto Eccellenza to DISCCO).

Data availability

The HCHS/SOL fully supports data sharing for HCHS/SOL–approved manuscript proposals with outside investigators. All data sharing is conducted in accordance with HCHS/SOL study and NIH policies and governed by a Data and Materials Distribution Agreement (DMDA) between UNC and the external institution, ensuring the confidentiality and privacy of HCHS/SOL participants and their families. Alternatively, de-identified HCHS/SOL data are publicly available at BioLINCC and dbGaP for the subset of the study cohort who authorized general use of their data at the time of informed consent. Supporting code to implement the methods performed in this study is publicly and freely available. In detail, we finalized an R package, named MSFA, to implement the methods presented in De Vito et al. [De Vito R, Bellio R, Trippa L, and Parmigiani G, "Multi-study factor analysis" and "Bayesian multi-study factor analysis for high-throughput biological data"] and their application in nutritional epidemiology. The package is available on GitHub: https://github.com/rdevito/MSFA.

Declarations

Ethics approval and consent to participate

This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving research study participants were approved by the Institutional Review Boards from each study site and the University of North Carolina at Chapel Hill (Coordinating Center). Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Biostatistics and Data Science Institute, Brown University, Providence, RI, USA. [2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [3]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [4]Departments of Nutrition and Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA, USA. [5]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [6]Department of Medicine - DMED, Università degli Studi di Udine, Udine, Italy. [7]Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY, USA. [8]Department of Psychology, University of Miami, Coral Gables, FL, USA. [9]Institute for Minority Health Research, University of Illinois at Chicago, Chicago, IL, USA. [10]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [11]Branch of Medical Statistics, Biometry and Epidemiology "G. A. Maccacaro", Department of Clinical Sciences and Community Health - Dipartimento Di Eccellenza 2023 - 2027, Università degli Studi di Milano, Via Celoria 22, 20133 Milan, Italy. [12]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.

Received: 25 September 2024 Accepted: 14 April 2025
Published online: 04 May 2025

References

1. Zhao J, Li Z, Gao Q, Zhao H, Chen S, Huang L, Wang W, Wang T. A review of statistical methods for dietary pattern analysis. Nutr J. 2021;20:37.
2. Gleason PM, Boushey CJ, Harris JE, Zoellner J. Publishing nutrition research: a review of multivariate techniques–part 3: data reduction methods. J Acad Nutr Diet. 2015;115:1072–82.
3. Jannasch F, Riordan F, Andersen LF, Schulze MB. Exploratory dietary patterns: a systematic review of methods applied in pan-European studies and of validation studies. Br J Nutr. 2018;120:601–11.
4. Schulze MB, Martinez-Gonzalez MA, Fung TT, Lichtenstein AH, Forouhi NG. Food based dietary patterns and chronic disease prevention. BMJ. 2018;361:k2396.
5. Edefonti V, De Vito R, Salvatori A, Bravi F, Patel L, Dalmartello M, Ferraroni M. Reproducibility of A Posteriori Dietary Patterns across Time and Studies: A Scoping Review. Adv Nutr. 2020;11:1255–81.
6. Krebs-Smith SM, Subar AF, Reedy J. Examining Dietary Patterns in Relation to Chronic Disease: Matching Measures and Methods to Questions of Interest. Circulation. 2015;132:790–3.
7. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhouser ML, Boushey CJ, Schap TE, Reedy J. The Dietary Patterns Methods Project: synthesis of findings across cohorts and relevance to dietary guidance. J Nutr. 2015;145:393–402.
8. Judd SE, Letter AJ, Shikany JM, Roth DL, Newby PK. Dietary Patterns Derived Using Exploratory and Confirmatory Factor Analysis are Stable and Generalizable Across Race, Region, and Gender Subgroups in the REGARDS Study. Front Nutr. 2014;1:29.
9. Tucker KL. Dietary patterns, approaches, and multicultural perspective. Appl Physiol Nutr Metab. 2010;35:211–8.
10. de Souza RJ, Zulyniak MA, Desai D, Shaikh MR, Campbell NC, Lefebvre DL, Gupta M, Wilson J, Wahi G, Atkinson SA, et al. Harmonization of Food-Frequency Questionnaires and Dietary Pattern Analysis in 4 Ethnically Diverse Birth Cohorts. J Nutr. 2016;146:2343–50.
11. Balder HF, Virtanen M, Brants HA, Krogh V, Dixon LB, Tan F, Mannisto S, Bellocco R, Pietinen P, Wolk A, et al. Common and country-specific dietary patterns in four European cohort studies. J Nutr. 2003;133:4246–51.
12. Mannisto S, Dixon LB, Balder HF, Virtanen MJ, Krogh V, Khani BR, Berrino F, van den Brandt PA, Hartman AM, Pietinen P, et al. Dietary patterns and breast cancer risk: results from three cohort studies in the DIETSCAN project. Cancer Causes Control. 2005;16:725–33.
13. Castello A, Buijsse B, Martin M, Ruiz A, Casas AM, Baena-Canada JM, Pastor-Barriuso R, Antolin S, Ramos M, Munoz M, et al. Evaluating the Applicability of Data-Driven Dietary Patterns to Independent Samples with a Focus on Measurement Tools for Pattern Similarity. J Acad Nutr Diet. 2016;116(1914–1924):e1916.
14. Castello A, Lope V, Vioque J, Santamarina C, Pedraz-Pingarron C, Abad S, Ederra M, Salas-Trejo D, Vidal C, Sanchez-Contador C, et al. Reproducibility of data-driven dietary patterns in two groups of adult Spanish women from different studies. Br J Nutr. 2016;116:734–42.
15. Moskal A, Pisa PT, Ferrari P, Byrnes G, Freisling H, Boutron-Ruault MC, Cadeau C, Nailler L, Wendt A, Kuhn T, et al. Nutrient patterns and their food sources in an International Study Setting: report from the EPIC study. PLoS ONE. 2014;9:e98647.
16. Edefonti V, Hashibe M, Ambrogi F, Parpinel M, Bravi F, Talamini R, Levi F, Yu G, Morgenstern H, Kelsey K, et al. Nutrient-based dietary patterns and the risk of head and neck cancer: a pooled analysis in the International Head and Neck Cancer Epidemiology consortium. Ann Oncol. 2012;23:1869–80.
17. De Vito R, Lee YCA, Parpinel M, Serraino D, Olshan AF, Zevallos JP, Levi F, Zhang ZF, Morgenstern H, Garavello W, et al. Shared and Study-specific Dietary Patterns and Head and Neck Cancer Risk in an International Consortium. Epidemiology. 2019;30:93–102.
18. Murakami K, Shinozaki N, Fujiwara A, Yuan X, Hashimoto A, Fujihashi H, Wang HC, Livingstone MBE, Sasaki S. A Systematic Review of Principal Component Analysis-Derived Dietary Patterns in Japanese Adults: Are Major Dietary Patterns Reproducible Within a Country? Adv Nutr. 2019;10:237–49.
19. Bakolis I, Hooper R, Bachert C, Lange B, Haahtela T, Keil T, Hofmaier S, Fokkens W, Rymarczyk B, Janson C, et al. Dietary patterns and respiratory health in adults from nine European countries-Evidence from the GA(2)LEN study. Clin Exp Allergy. 2018;48:1474–82.

20. De Vito R, Bellio R, Trippa L, Parmigiani G. Multi-study factor analysis. Biometrics. 2019;75:337–46.
21. De Vito R, Bellio R, Trippa L, Parmigiani G. Bayesian multi-study factor analysis for high-throughput biological data. The Annals of Applied Statistics. 2021;15:1723–41.
22. Sorlie PD, Aviles-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Schneiderman N, Raij L, Talavera G, Allison M, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol. 2010;20:629–41.
23. Pirzada A, Cai J, Heiss G, Sotres-Alvarez D, Gallo LC, Youngblood ME, Aviles-Santa ML, Gonzalez HM, Isasi CR, Kaplan R, et al. Evolving Science on Cardiovascular Disease Among Hispanic/Latino Adults: JACC International. J Am Coll Cardiol. 2023;81:1505–20.
24. Maldonado LE, Adair LS, Sotres-Alvarez D, Mattei J, Mossavar-Rahmani Y, Perreira KM, Daviglus ML, Van Horn LV, Gallo LC, Isasi CR, Albrecht SS. Dietary Patterns and Years Living in the United States by Hispanic/Latino Heritage in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). J Nutr. 2021;151:2749–59.
25. Stephenson BJK, Sotres-Alvarez D, Siega-Riz AM, Mossavar-Rahmani Y, Daviglus ML, Van Horn L, Herring AH, Cai J. Empirically Derived Dietary Patterns Using Robust Profile Clustering in the Hispanic Community Health Study/Study of Latinos. J Nutr. 2020;150:2825–34.
26. Lachat C, Hawwash D, Ocke MC, Berg C, Forsum E, Hornell A, Larsson C, Sonestedt E, Wirfalt E, Akesson A, et al. Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. PLoS Med. 2016;13:e1002036.
27. Lavange LM, Kalsbeek WD, Sorlie PD, Aviles-Santa LM, Kaplan RC, Barnhart J, Liu K, Giachello A, Lee DJ, Ryan J, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. Ann Epidemiol. 2010;20:642–9.
28. Nutrition Data System for Research (NDS-R) version. NDS- R software, 1998–1999 Regents of the University of Minnesota. Minneapolis (MN): Nutrition Coordinating Center, University of Minnesota; 2011.
29. Pett MA, Lackey NR, Sullivan JJ. Making sense of factor analysis: the use of factor analysis for instrument development in health care research. Thousand Oaks, CA: Sage; 2003.
30. Islam MA, Amin MN, Siddiqui SA, Hossain MP, Sultana F, Kabir MR. Trans fatty acids and lipid profile: A serious risk factor to cardiovascular disease, cancer and diabetes. Diabetes Metab Syndr. 2019;13:1643–7.
31. Lichtenstein AH. Dietary fat and cardiovascular disease risk: quantity or quality? J Womens Health (Larchmt). 2003;12:109–14.
32. Johnson RA, Wichern DW. Applied multivariate statistical analysis. 5th ed. Upper Saddler River, NJ: Prentice Hall; 2002.
33. DiStefano C, Zhu M, Mindrila D. Understanding and using factor scores: Considerations for the applied researcher. Pract Assess Res Eval. 2009;14:1–11.
34. Chiuve SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, Stampfer MJ, Willett WC. Alternative dietary indices both strongly predict risk of chronic disease. J Nutr. 2012;142:1009–18.
35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57:289–300.
36. Satia JA. Dietary acculturation and the nutrition transition: an overview. Appl Physiol Nutr Metab. 2010;35:219–23.
37. Lumley T. Complex surveys: a guide to analysis using R. Hoboken: Wiley; 2011.
38. R Development Core Team. R: A Language and Environment for Statistical Computing 2024 Vienna, Austria: R Foundation for Statistical Computing, Available at: http://www.R-project.org (23 Sept 2024, date last accessed).
39. Giner G, Smyth GK: Statmod: probability calculations for the Inverse Gaussian distribution. arXiv preprint arXiv:160306687 2016, Available at: https://arxiv.org/abs/1603.06687 (23 Sept 2024, date last accessed).
40. Revelle W: Psych: procedures for psychological, psychometric, and personality research. Evanston, IL: Northwestern University. R package version 2.4.6, 2024, Available at: https://CRAN.R-project.org/package=psych (23 Sept 2024, date last accessed).
41. Raîche G, Walls TA, Magis D, Riopel M, Blais JG: Non-graphical solutions for Cattell's scree test. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences. 2013;9:23.
42. Wickham H. Ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
43. Lumley T. Survey: analysis of complex survey samples. R package version 4.4.1, 2024, Available at: https://cran.r-project.org/web/packages/survey/index.html (23 Sept 2024, date last accessed).
44. De Vito R, Bellio R: MSFA. GitHub repository. Available at: https://githubcom/rdevito/MSFA (23 Sept 2024, date last accessed).
45. Ellithorpe ME, Takahashi B, Alumit Zeldes G, Dorrance-Hall E, Chavez M, Plasencia J. Family and Cultural Perceptions About Meat Consumption among Hispanic/Latino and White Adults in the United States. Ecol Food Nutr. 2022;61(3):353–66.
46. Haytowitz DB, Pehrsson PR. USDA's National Food and Nutrient Analysis Program (NFNAP) produces high-quality data for USDA food composition databases: Two decades of collaboration. Food Chem. 2018;238:134–8.
47. Dixon LB, Sundquist J, Winkleby M. Differences in energy, nutrient, and food intakes in a US sample of Mexican-American women and men: findings from the Third National Health and Nutrition Examination Survey, 1988–1994. Am J Epidemiol. 2000;152:548–57.
48. Siega-Riz AM, Sotres-Alvarez D, Ayala GX, Ginsberg M, Himes JH, Liu K, Loria CM, Mossavar-Rahmani Y, Rock CL, Rodriguez B, et al. Food-group and nutrient-density intakes by Hispanic and Latino backgrounds in the Hispanic Community Health Study/Study of Latinos. Am J Clin Nutr. 2014;99:1487–98.
49. Bermudez OI, Falcon LM, Tucker KL. Intake and food sources of macronutrients among older Hispanic adults: association with ethnicity, acculturation, and length of residence in the United States. J Am Diet Assoc. 2000;100:665–73.
50. Sanchez-Flack JC, Anderson CAM, Arredondo EM, Belch G, Martinez ME, Ayala GX. Fruit and Vegetable Intake of US Hispanics by Food Store Type: Findings from NHANES. J Racial Ethn Health Disparities. 2019;6:220–9.
51. Mattei J, Sotres-Alvarez D, Daviglus ML, Gallo LC, Gellman M, Hu FB, Tucker KL, Willett WC, Siega-Riz AM, Van Horn L, Kaplan RC. Diet Quality and Its Association with Cardiometabolic Risk Factors Vary by Hispanic and Latino Ethnic Background in the Hispanic Community Health Study/Study of Latinos. J Nutr. 2016;146:2035–44.
52. Carrera PM, Gao X, Tucker KL. A study of dietary patterns in the Mexican-American population and their association with obesity. J Am Diet Assoc. 2007;107:1735–42.
53. Lin H, Bermudez OI, Tucker KL. Dietary patterns of Hispanic elders are associated with acculturation and obesity. J Nutr. 2003;133:3651–7.
54. Noel SE, Newby PK, Ordovas JM, Tucker KL. A traditional rice and beans pattern is associated with metabolic syndrome in Puerto Rican older adults. J Nutr. 2009;139:1360–7.
55. Sofianou A, Fung TT, Tucker KL. Differences in diet pattern adherence by nativity and duration of US residence in the Mexican-American population. J Am Diet Assoc. 2011;111(1563–1569):e1562.
56. Varela JJ, Mattei J, Sotres-Alvarez D, Mossavar-Rahmani Y, McClain AC, Maldonado LE, Daviglus ML, Stephenson BJ. Examining Generalizability across Different Surveys: Comparing Nutrient-Based Food Patterns and Their Cross-Sectional Associations with Cardiometabolic Health in the United States Hispanic/Latino Adults. Curr Dev Nutr. 2024;8:103797.
57. Mattei J, McClain AC, Falcon LM, Noel SE, Tucker KL. Dietary Acculturation among Puerto Rican Adults Varies by Acculturation Construct and Dietary Measure. J Nutr. 2018;148:1804–13.
58. Perez-Escamilla R. Acculturation, nutrition, and health disparities in Latinos. Am J Clin Nutr. 2011;93:1163S-1167S.
59. Duffey KJ, Gordon-Larsen P, Ayala GX, Popkin BM. Birthplace is associated with more adverse dietary profiles for US-born than for foreign-born Latino adults. J Nutr. 2008;138:2428–35.
60. Siega-Riz AM, Pace ND, Butera NM, Van Horn L, Daviglus ML, Harnack L, Mossavar-Rahmani Y, Rock CL, Pereira RI, Sotres-Alvarez D. How Well Do U.S. Hispanics Adhere to the Dietary Guidelines for Americans? Results from the Hispanic Community Health Study/Study of Latinos. Health Equity. 2019;3:319–27.
61. Arandia G, Nalty C, Sharkey JR, Dean WR. Diet and acculturation among Hispanic/Latino older adults in the United States: a review of literature and recommendations. J Nutr Gerontol Geriatr. 2012;31:16–37.
62. Ayala GX, Baquero B, Klinger S. A systematic review of the relationship between acculturation and diet among Latinos in the United States: implications for future research. J Am Diet Assoc. 2008;108:1330–44.
63. Barcenas CH, Wilkinson AV, Strom SS, Cao Y, Saunders KC, Mahabir S, Hernandez-Valero MA, Forman MR, Spitz MR, Bondy ML. Birthplace, years of residence in the United States, and obesity among Mexican-American adults. Obesity (Silver Spring). 2007;15:1043–52.

64. Liu J, Fulda KG, Tao MH. Association between acculturation and metabolic syndrome in Hispanic adults mediated by fruits intake. Public Health Nutr. 2021;24:6472–6.
65. Jonnalagadda SS, Mitchell DC, Smiciklas-Wright H, Meaker KB, Van Heel N, Karmally W, Ershow AG, Kris-Etherton PM. Accuracy of energy intake data estimated by a multiple-pass, 24-hour dietary recall technique. J Am Diet Assoc. 2000;100:303–8 quiz 309–311.
66. Mossavar-Rahmani Y, Shaw PA, Wong WW, Sotres-Alvarez D, Gellman MD, Van Horn L, Stoutenberg M, Daviglus ML, Wylie-Rosett J, Siega-Riz AM, et al. Applying Recovery Biomarkers to Calibrate Self-Report Measures of Energy and Protein in the Hispanic Community Health Study/Study of Latinos. Am J Epidemiol. 2015;181:996–1007.