Dietary Patterns Derived by Reduced Rank Regression Are Inversely Associated with Type 2 Diabetes Risk across 5 Ethnic Groups in the Multiethnic Cohort[1,2]

Simone Jacobs,[3,4] Janine Kroeger,[5,6] Matthias B Schulze,[5,6] Laura K Frank,[7] Adrian A Franke,[3] Iona Cheng,[8] Kristine R Monroe,[9] Christopher A Haiman,[9] Laurence N Kolonel,[3] Lynne R Wilkens,[3] Loic Le Marchand,[3] Carol J Boushey,[3] and Gertraud Maskarinec[3]*

[3]Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI; [4]Epidemiology and Biostatistics Unit, Institute of Public Health, Heidelberg University, Heidelberg, Germany; [5]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany; [6]German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany; [7]Unit of Social Determinants of Health, Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin, Germany; [8]Cancer Prevention Institute of California, Fremont, CA; and [9]Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA

Abstract

Background: Reduced rank regression (RRR) is an approach to identify dietary patterns associated with biochemical markers and risk of type 2 diabetes (T2D).

Objective: We aimed to derive dietary patterns associated with adiponectin, leptin, C-reactive protein (CRP), and triglycerides (TGs) and to examine the prospective associations of these patterns with T2D risk in 5 ethnic/racial groups with differences in T2D rates.

Methods: The Multiethnic Cohort (MEC) included 215,831 African-American, Japanese-American, Latino, Native Hawaiian, and white adults living in Hawaii and California who completed a validated quantitative food-frequency questionnaire in 1993–1996. T2D status was based on self-report with confirmation by administrative data. Serum CRP and TGs and plasma adiponectin and leptin were measured ~10 y after baseline in a subset (n = 10,008) of participants. RRR was applied to dietary data and biomarker information of 10,008 MEC participants in the combined population and in each ethnic/racial group. RRR-derived dietary patterns, simplified by removal of foods that were not found to be important, were subsequently evaluated for association with T2D risk in 155,316 cohort members (8687 incident T2D cases diagnosed by 2010) by using Cox proportional hazards regression.

Results: Combining ethnic/racial groups, we identified a dietary pattern low in processed and red meat, sugar-sweetened beverages, diet soft drinks, and white rice and high in whole grains, fruit, yellow-orange vegetables, green vegetables, and low-fat dairy that was inversely associated with CRP, TGs, and leptin and positively related to adiponectin. Comparing extreme tertiles, the dietary pattern predicted a 16–28% significantly lower T2D risk in the combined study population and also separately in African Americans, Japanese Americans, Latinos, Native Hawaiians, and whites. Ethnicity-specific derived patterns varied only modestly from the overall pattern and resulted in comparable associations with T2D.

Conclusion: This identified dietary pattern may lower T2D risk through its impact on adipokines, by lowering chronic inflammation and dyslipidemia across 5 ethnic/racial groups. *Curr Dev Nutr* 2017;1:e000620.

Keywords: adipokines, adiponectin, biomarker, C-reactive protein, dietary pattern, leptin, Multiethnic Cohort, reduced rank regression, triglycerides, type 2 diabetes

Manuscript received February 13, 2017. Initial review completed February 22, 2017. Revision accepted April 17, 2017. First published April 17, 2017.

[1] SJ was supported by a postdoctoral fellowship from the German Research Foundation (DFG JA 2564/1-1). The Multiethnic Cohort study was supported by NIH grants R37 CA54281 and U01 CA164973. The diabetes data collection was funded by R21 DK073816 (principal investigator: GM). The work of the University of Hawaii Cancer Center Nutritional Support Shared Resource was supported in part by P30 CA071789.

[2] Author disclosures: S Jacobs, J Kroeger, MB Schulze, LK Frank, AA Franke, I Cheng, KR Monroe, CA Haiman, LN Kolonel, LR Wilkens, L Le Marchand, CJ Boushey, and G Maskarinec, no conflicts of interest.

*To whom correspondence should be addressed. E-mail: gertraud@cc.hawaii.edu.

[10] Abbreviations used: AHEI, Alternative Healthy Eating Index; aMED, Alternate Mediterranean Diet Score; CRP, C-reactive protein; DASH, Dietary Approaches to Stop Hypertension; HbA1c, glycated hemoglobin; HEI, Healthy Eating Index; MEC, Multiethnic Cohort; QFFQ, quantitative FFQ; RRR, reduced rank regression; $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; $RRRDS_{ethni}$, dietary pattern scores derived ethnicity-specifically; SSB, sugar-sweetened beverage; T2D, type 2 diabetes.

Introduction

More than 415 million people worldwide have type 2 diabetes (T2D)[10]; by 2040, the number of people with T2D is predicted to reach 642 million (1). In the United States, higher T2D prevalence rates are reported in Asian Americans, Pacific Islanders (2, 3), African Americans,

and Hispanics (4) than in whites. Excess body weight, genetic predisposition, and metabolic factors, as reflected in different T2D biomarker profiles across ethnic groups (5), along with modifiable lifestyle factors such as diet and exercise, may contribute to ethnic/racial differences in T2D.

Dietary patterns may capture the synergistic effects of multiple influential aspects of diet. Two approaches to characterizing overall diet are commonly distinguished: a priori indexes and a posteriori patterns. A priori indexes evaluate dietary quality and are constructed from adherence to dietary recommendations made on the basis of existing scientific evidence relating diet to chronic diseases. A posteriori–derived dietary patterns are identified through exploratory data-driven approaches, such as factor analysis, which is used to identify common underlying patterns of food consumption (6). Dietary patterns have been associated with T2D; in the Hawaii component of the Multiethnic Cohort (MEC), the associations between the a priori indexes Healthy Eating Index (HEI)–2010 and the alternative HEI (AHEI)–2010 and the alternate Mediterranean diet (aMED) and Dietary Approaches to Stop Hypertension (DASH) scores and T2D risk were strongest in whites compared with Native Hawaiians and Japanese Americans (7), whereas factor analysis–derived dietary patterns were not significantly associated with T2D in Native Hawaiians and showed inconsistent associations in whites and Japanese Americans (8). These inconsistent findings may be due to the a priori methods not taking into account ethnicity-specific consumption patterns—for example, the concentration on one starch, such as rice—and a posteriori methods not being optimally suited to identify dietary patterns predictive of disease risk. A method that combines a priori and a posteriori approaches is reduced rank regression (RRR) (9). In contrast to a priori and a posteriori methods, RRR-derived dietary patterns use a data-driven approach, identify factors that explain covariation in food consumption, are associated with intermediate disease markers, and incorporate ethnicity-specific differences in diet and biomarkers, and thus are likely to include foods that are related to disease risk. Previous studies identified dietary patterns predictive of T2D by using different intermediate markers as response variables, including inflammatory markers (10–13), HOMA-IR (14, 15), glycated hemoglobin (HbA1c) (10, 15), fasting glucose (15), adiponectin (10, 16), HDL cholesterol (10, 16), and TGs (16).

Because previous studies (10–16) did not investigate RRR-derived dietary patterns and T2D risk in different ethnic groups in the United States, the first aim was to compare dietary patterns associated with T2D-related biomarkers (17, 18) [adiponectin, leptin, TGs, and C-reactive protein (CRP)] derived in the combined study population and each ethnic/racial group. The second aim was to evaluate the associations of these dietary patterns with T2D risk in African-American, Japanese-American, Latino, Native Hawaiian, and white MEC participants.

Methods

Study population

The MEC is a prospective cohort study primarily established to study lifestyle and genetic factors and cancer among different ethnic/racial groups in Hawaii and California (19). The cohort includes 215,831 men and women, aged 45–75 y at recruitment (1993–1996), living in Hawaii and in Los Angeles, California, with the following ethnic/racial distribution: African American (16.3%), Japanese American (26.4%), Latino (22.0%), Native Hawaiian (6.5%), white (22.9%), and other ancestry (5.8%) (19). From 1993 to 1996, participants enrolled in the cohort by completing a self-administered mail questionnaire on diet, demographic characteristics, medical conditions, lifestyle factors, and anthropometric measures (19). Biological specimens (mainly blood and urine samples) were collected from a subset of the cohort (n = 68,740) primarily during 2001–2006. Of these 68,740 participants, a panel of biochemical markers was assessed in a biomarker subcohort of 12,578 predominantly fasting individuals, who were selected from controls in case-control studies within the MEC.

Dietary assessment

Dietary data were collected at baseline by a validated and calibrated self-administered quantitative FFQ (QFFQ) with >180 food items (20) specifically designed for use in this multiethnic population (19). Food mixtures were disaggregated into their components, and each ingredient was assigned to the relevant food item. Individual food items and foods from mixed dishes were classified into 41 food groups on the basis of nutrient profiles and culinary uses.

Ascertainment of T2D

On the basis of the information from 3 questionnaires [1993–1996 (baseline), 1999–2002, and 2003–2007] and 3 sources of administrative data [i.e., Medicare claims (21), California hospital discharge diagnoses (22), and a Hawaii health plan linkage (2)], we developed a strict definition of T2D. Only participants with ≥1 self-reported T2D diagnosis on one of the questionnaires and confirmation by ≥1 administrative data source were considered incident cases. The first report of a T2D diagnosis was considered as the year of discovery because exact dates of diagnosis were unavailable.

Laboratory procedures

Adiponectin and leptin from plasma were measured by using ELISA kits (catalog no. DRP300; R&D Systems). Insulin from serum was measured by using an ELISA kit (catalog no. EZHI-14K; EMD Millipore). All ELISA protocols were followed in accordance with the manufacturers' instructions. A Cobas Mira Plus chemistry autoanalyzer (Roche Diagnostics) was used to measure serum glucose (kit from Randox), CRP, and TGs (kits from Pointe Scientific, Inc.) per the manufacturers' instructions. HOMA-IR was calculated as follows: [fasted insulin (microunits/liter) × fasted glucose (milligrams/deciliter)]/405 (23).

For the current analysis, the 4 biomarkers—TGs, leptin, CRP, and adiponectin—were chosen as the intermediate markers of T2D serving as response variables for RRR for the following reasons: they are affected by diet (24–27), they showed a cross-sectional association with HOMA-IR as a marker of a prediabetic stage in nondiabetics from all ethnic groups in the current analysis (r across ethnic groups: −0.28 to −0.47 for adiponectin, 0.34–0.47 for leptin, 0.14–0.25 for CRP, and 0.25–0.40 for TGs), and previous studies indicate that the 4 biomarkers are related to the pathophysiology

TABLE 1 Food group intake medians by tertile of the original (nonsimplified) $RRRDS_{comb}$ and correlations of the food groups with the $RRRDS_{comb}$ in the biomarker subcohort[1]

	All ethnic groups combined			African American			Japanese American			Latino			Native Hawaiian			White		
	T1	T3	*r*	T1	T3	*r*	T1	T3	*r*	T1	T3	*r*	T1	T3	*r*	T1	T3	*r*
n	3334	3340		645	676		800	689		1003	1173		705	593		181	209	
$RRRDS_{comb}$ score	−10.2 to −0.34	0.35 to 6.05		−5.91 to −0.34	0.35 to 4.95		−3.73 to −0.34	0.35 to 3.65		−6.38 to −0.34	0.36 to 5.39		−10.2 to −0.34	0.36 to 6.05		−4.36 to −0.34	0.36–5.66	
range																		
Food groups, g/d																		
Processed meat	22.1	9.5	−0.38*	20.0	8.4	−0.40*	21.4	9.5	−0.43*	20.4	9.6	−0.35*	27.6	12.3	−0.38*	17.6	7.1	−0.42*
Red meat	53.7	22.5	−0.57*	41.1	16.9	−0.56*	47.8	21.0	−0.54*	69.2	26.2	−0.64*	60.5	27.8	−0.53*	45.6	16.3	−0.56*
SSBs	50.9	11.8	−0.31*	63.9	11.8	−0.34*	50.9	11.8	−0.27*	59.2	11.8	−0.29*	59.2	11.8	−0.39*	11.8	0.0	−0.23*
Diet drinks	0.0	0.0	−0.30*	0.0	0.0	−0.23*	11.8	0.0	−0.29*	0.0	0.0	−0.24*	11.8	0.0	−0.37*	11.8	0.0	−0.42*
White rice	71.1	28.3	−0.30*	16.4	16.4	−0.14*	400	200	−0.45*	13.2	16.4	−0.06*	286	142	−0.42*	56.6	16.4	−0.35*
Whole grains	20.1	91.6	0.40*	33.9	94.1	0.39*	15.9	121	0.46*	17.8	85.3	0.39*	22.1	88.8	0.39*	23.5	74.4	0.34*
Fruit	129	342	0.40*	120	339	0.40*	112	336	0.51*	176	387	0.39*	126	293	0.38*	106	340	0.43*
Yellow-orange vegetables	15.7	34.4	0.37*	13.6	32.7	0.40*	16.5	39.1	0.45*	15.5	33.4	0.34*	17.9	31.1	0.34*	14.4	40.2	0.42*
Green vegetables	87.3	138	0.33*	77.3	140	0.35*	95.5	156	0.38*	76.3	123	0.29*	101	137	0.29*	85.5	167	0.46*
Low-fat dairy	19.7	134	0.32*	15.7	120	0.30*	17.0	120	0.32*	12.8	172	0.35*	23.7	134	0.27*	42.6	199	0.30*

[1] Values are medians of food group intakes and Pearson correlation coefficients (*r*) between the food group intakes and $RRRDS_{comb}$; n = 10,008. Only food groups with factor loadings ≥0.2 in the combined analysis are shown. *$P < 0.05$. $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; SSB, sugar-sweetened beverage; T, tertile.

of T2D (17, 18). Because all biomarkers should refer to the same stage of the pathway from exposure to disease, our aim was to focus on upstream markers linking diet to T2D risk. HOMA-IR and HbA1c may refer to a later stage of the pathogenesis of diabetes than TGs, leptin, CRP, and adiponectin and were therefore not included as response variables in the current analysis. The use of HDL cholesterol instead of leptin as response in the RRR method yielded results similar to the current ones.

Statistical analysis

RRR. Of the 12,578 participants in the biomarker subcohort, we excluded T2D cases with diagnosis before or at blood draw, ethnicities/races other than the 5 major groups, incomplete main confounder information, and incomplete or implausible (HDL cholesterol > total cholesterol) biomarker information, resulting in a final study population of 10,008 biomarker subcohort members, which we used for identification of dietary patterns by RRR. Characteristics between the biomarker subcohort and the full MEC cohort were similar overall, except that proportionally fewer whites, slightly fewer Japanese Americans, and more African Americans, Native Hawaiians, and Latinos are present in the biomarker subcohort (data not shown).

RRR is a type of multivariate regression that aims to reduce the dimensionality of complex data. RRR was applied by using the partial least-squares procedure in SAS to identify linear functions of predictors (i.e., food group intake in grams) that explain as much variation in the response variables (here, T2D-related biomarkers) as possible (9). A more detailed description of the method has been previously published (9, 28). All statistical analyses were performed by using SAS release 9.4 (SAS Institute).

RRR of dietary data was applied to biomarker subcohort members combining ethnic groups (n = 10,008). Ethnicity-specific derivation of RRR patterns was conducted in a separate analysis given ethnic/racial differences in previously reported dietary pattern and dietary index analyses in the MEC (7, 8). Sex-specific RRR results were of similar magnitude; therefore, men and women were combined.

RRR does not allow for direct adjustment for potential confounders. Therefore, values for the predictors and response variables were adjusted before entry into the RRR model by using the residual method (29). The median of the variable of interest was added to the residuals of a linear regression of that variable on the potential confounders in order to maintain the proper range. Food groups were adjusted for energy intake and race/ethnicity, all biomarkers for race/ethnicity, and CRP additionally for season of blood draw to account for seasonal fluctuations.

Association of dietary patterns with biomarkers. The longitudinal associations between the dietary pattern score and T2D-related biomarkers were examined in a regression model with the score as exposure and biomarkers as outcome and the calculation of geometric means of biomarker concentration across dietary pattern tertiles, adjusted for confounder, which were selected a priori, namely BMI (continuous), total energy intake (continuous, kilojoules per day), sex, age (continuous), education (≤12 y, 13–15 y, or ≥16 y), race/ethnicity (white, African American, Japanese

TABLE 2 Food group intake medians and factor loadings by tertile of original (nonsimplified) $RRRDS_{ethni}$ in the biomarker subcohort[1]

	African American			Japanese American			Latino			Native Hawaiian			White		
	T1	T3	Factor loadings	T1	T3	Factor loadings	T1	T3	Factor loadings	T1	T3	Factor loadings	T1	T3	Factor loadings
n	683	684		742	743		1080	1079		631	631		198	197	
Food groups, g/d															
Processed meat	17.1	9.1	−0.21	23.8	7.8	−0.32	21.0	8.6	−0.25	24.1	16.5	−0.18	16.0	7.3	−0.12
Red meat	39.7	17.1	−0.38	49.2	18.9	−0.26	66.7	24.1	−0.37	57.6	31.4	−0.37	45.5	19.0	−0.29
Poultry	63.9	38.9	−0.26	37.4	25.8	−0.04	49.7	39.1	−0.06	41.7	29.1	−0.21	38.3	27.8	−0.12
Shellfish	3.8	0.9	−0.33	4.3	2.4	−0.04	2.5	1.3	−0.05	4.4	2.7	−0.07	1.3	2.3	0.09
Eggs	11.4	8.7	−0.11	14.8	6.5	−0.24	12.8	8.9	−0.08	12.6	12.1	0.06	9.9	12.1	0.06
Other tubers and potatoes	22.1	13.9	−0.21	24.0	16.1	<0.01	22.7	16.4	−0.03	49.8	30.7	−0.28	24.6	16.5	−0.08
French fries	3.3	0.0	−0.15	5.8	0.0	−0.13	5.8	2.3	−0.11	4.6	2.3	−0.27	5.8	2.3	−0.11
SSBs	29.6	14.8	−0.07	50.9	5.9	−0.15	128	5.9	−0.26	50.9	11.8	−0.15	11.8	0.0	−0.37
Diet soft drinks	0.0	0.0	−0.19	0.0	0.0	−0.14	0.0	0.0	−0.16	25.5	0.0	−0.30	0.0	0.0	−0.09
White rice	16.4	28.3	0.18	400	200	−0.26	8.2	16.4	0.05	286	143	−0.21	46.1	16.4	−0.30
Fish	11.1	11.9	0.06	20.8	16.1	−0.04	6.1	9.6	0.21	23.2	21.3	0.03	13.6	14.8	0.10
Nuts	3.1	3.6	0.08	3.1	3.5	0.17	1.7	2.2	0.22	3.4	5.0	0.29	3.8	3.0	−0.09
Whole grains	37.5	86.1	0.28	17.8	90.5	0.27	17.8	91.7	0.32	32.9	47.4	0.04	34.8	49.5	0.07
Fruit	160	270	0.11	120	293	0.31	188	358	0.23	152	216	0.18	142	295	0.28
Yellow-orange vegetables	15.6	27.5	0.21	17.7	36.7	0.34	16.9	33.0	0.24	23.4	20.8	−0.05	15.6	29.0	0.15
Green vegetables	83.7	129	0.25	104	138	0.21	78.6	126	0.24	119	111	−0.09	99.7	137	0.27
Cruciferous vegetables	33.6	51.3	0.20	39.6	52.5	0.19	21.8	38.3	0.18	47.2	37.4	−0.14	35.0	46.7	0.23
Tomatoes	5.8	24.9	0.21	5.8	10.3	0.10	10.3	24.9	0.11	9.9	24.9	0.19	12.4	24.9	0.16
Other vegetables	0.0	2.5	0.19	2.5	4.8	0.12	0.0	2.5	0.10	2.5	2.5	−0.02	1.3	4.8	0.24
Low-fat dairy	19.7	103	0.23	19.7	89.6	0.20	14.5	180	0.24	38.0	85.3	0.05	97.2	130	0.06
Legumes	19.4	19.5	0.05	11.2	11.2	0.27	69.8	40.8	−0.15	13.9	12.4	0.11	13.1	18.0	0.30
Cottage cheese	0.0	0.0	0.09	0.0	0.0	0.13	1.8	1.8	−0.06	0.0	0.0	0.24	0.0	1.8	0.12
Coffee	171	240	0.06	248	240	−0.11	338	240	−0.10	154	248	0.22	260	260	−0.01
Alcohol, drinks/d	0.1	0.0	−0.13	0.0	0.0	−0.06	0.1	0.1	0.15	0.0	0.1	0.23	0.2	0.6	−0.01

[1] Values are medians of the food group intakes and factor loadings; *n* = 10,008. Only food groups with factor loadings ≥0.2 in ≥1 of the ethnic groups are shown. For African Americans, Native Hawaiians, Japanese Americans, and Latinos, the factor loadings signs were reversed to enable better comparison across ethnic groups. $RRRDS_{ethni}$, dietary pattern scores derived ethnicity-specifically; SSB, sugar-sweetened beverage; T, tertile.

American, Latino, or Native Hawaiian; only in combined analysis), physical activity (<30 or or ≥30 min/d), and smoking status (never smoker, past smoker, or current smoker). For missing values of smoking status and physical activity, a missing category was created for each variable (~1% missing data). The significance of linear trends across dietary pattern tertiles was tested by assigning each participant the median value for his or her tertile and modeling this value as a continuous variable.

Association of dietary patterns with T2D. For the application of the RRR-derived dietary patterns in the entire MEC (*n* = 215,831), we first excluded the biomarker subcohort participants (*n* = 12,578) to obtain independent samples for the identification of dietary pattern and to assess the association of the pattern with T2D risk. Second, we excluded (with some overlap) prevalent diabetes cases at cohort entry (*n* = 28,153), ethnic groups other than the major ethnic groups (*n* = 13,994), and those with missing information on essential covariates (*n* = 11,940), resulting in a final number of 155,316 participants representing an independent sample without biomarker measurements.

To make the results more generalizable and easier to interpret, and to reduce the possibility of overfitting, we simplified the dietary pattern score by calculating the unweighted sum of the *z*-standardized intakes of food groups with factor loadings ≥0.2 of the absolute value (9). We subsequently applied the simplified scores in the full cohort and estimated HRs for risk of T2D by using Cox proportional hazards regression for the continuous dietary pattern scores and across dietary pattern tertiles. Observation started at the time of cohort entry and ended at the T2D time of discovery, death, or closure date for follow-up (31 December 2010). The model was adjusted for age as a strata variable and for the same confounders as used in the biomarker analysis in the log-linear component of the model.

In sensitivity analyses, we first excluded all participants who provided blood (*n* = 68,740). We then investigated the impact of excluding participants with lipid-lowering and anti-inflammatory medication use on the results of the RRR analysis. We also excluded nonfasting participants in the RRR analysis, as well as participants with acute inflammation (indicated by CRP concentrations >10 mg/L), and participants with extreme energy intakes that fell outside the recommended cutoffs (<500 and >3500 kcal/d) (30). Finally, because body fat may confound associations between food intake and biomarkers, we additionally adjusted all biomarker values for BMI before their use as response variables in the RRR in a separate analysis.

TABLE 3 Baseline characteristics of the biomarker subcohort by tertile of original (nonsimplified) $RRRDS_{comb}$[1]

	All ethnic groups combined			African American			Japanese American			Latino			Native Hawaiian			White		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
$RRRDS_{comb}$ score	−10.2 to	−0.34 to	0.35 to	−5.91 to	−0.34 to	0.35 to	−3.73 to	−0.34 to	0.35 to	−6.38 to	−0.34 to	0.36 to	−10.2 to	−0.34 to	0.36 to	−4.36 to	−0.34 to	0.36 to
range	−0.34	0.35	6.05	−0.34	0.35	4.95	−0.34	0.35	3.65	−0.34	0.35	5.39	−0.34	0.35	6.05	−0.34	0.35	5.66
n	3334	3334	3340	645	730	676	800	739	689	1003	1065	1173	705	594	593	181	206	209
Sex, % male	58.0	42.4	38.7	40.0	33.4	29.1	66.0	42.4	40.5	59.3	47.1	43.6	62.7	44.6	36.9	60.8	44.2	40.7
Age, y	56.1 ±	58.4 ±	60.6 ±	56.4 ±	58.8 ±	61.4 ±	56.3 ±	58.9 ±	62.1 ±	57.7 ±	59.4 ±	60.9 ±	53.0 ±	54.4 ±	56.6 ±	57.7 ±	61.0 ±	63.3 ±
	7.77	8.17	8.12	8.3	8.88	8.72	8.01	8.2	8.04	7.06	7.12	6.96	6.96	7.52	8.3	8.00	8.73	8.12
BMI, kg/m^2	27.5 ±	26.3 ±	25.6 ±	28.7 ±	27.5 ±	26.8 ±	25.0 ±	24.0 ±	23.2 ±	27.6 ±	26.7 ±	26.1 ±	29.1 ±	27.3 ±	26.5 ±	26.9 ±	25.2 ±	24.9 ±
	4.99	4.55	4.15	5.83	4.7	4.39	3.42	3.38	3.13	4.34	4.15	3.74	5.38	5.26	4.56	5.05	4.17	4.02
Education, %																		
≤12 y	41.1	39.1	37.5	31.2	27.8	22.9	25.1	29.4	31.5	62.8	59.2	54.7	40.4	37.1	34.1	28.7	16.5	17.7
13–15 y	33.1	32.8	32.3	42.0	41.6	40.5	33.4	30.3	30.0	26.0	28.3	27.2	37.3	33.7	37.4	23.2	30.1	26.8
≥16 y	25.9	28.1	30.2	26.8	30.6	36.5	41.5	40.3	38.5	11.2	12.5	18.1	22.4	29.2	28.5	48.1	53.4	55.5
Smoking status, %																		
Never	40.6	49.9	52.5	40.5	47.8	48.1	39.9	57.1	60.1	40.4	47.6	55.5	42.9	49.2	47.9	37.6	44.7	37.8
Past	41.5	37.7	37.6	37.4	38.5	41.7	46.8	35.5	35.1	40.9	37.7	33.0	38.7	37.6	40.5	48.1	43.7	50.2
Current	16.9	11.0	8.5	21.4	12.9	9.5	13.0	7.0	4.2	16.9	11.7	8.4	18.0	12.8	11.5	14.4	10.7	12.0
Physical activity, %																		
<30 min/d	35.1	35.3	30.1	42.3	41.6	34.3	29.9	30.7	26.1	44.0	42.1	37.6	23.2	25.3	20.1	29.3	22.8	16.3
≥30 min/d	63.7	63.6	68.4	55.7	56.4	63.8	70.1	69.3	73.2	54.1	56.0	59.8	75.6	74.0	79.6	70.7	76.7	83.7
Energy intake, kJ/d	10122 ±	8073 ±	9978 ±	8725 ±	7318 ±	9348 ±	9181 ±	7550 ±	8880 ±	10,966 ±	8627 ±	10,914 ±	11,532 ±	8723 ±	10,430 ±	9070 ±	7895 ±	9090 ±
	5117	3716	4530	4768	3460	4444	3393	2822	2905	6104	4191	5038	5123	3974	5000	4305	3266	3391

[1]Values are means ± SDs unless otherwise indicated; *n* = 10,008. Missing data: Smoking status, *n* = 123; Physical activity, *n* = 130. $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; T, tertile.

Results

The RRR method applied to all biomarker subcohort participants resulted in 4 scores representing a linear combination of the *z*-standardized intakes of 41 food groups, with each group multiplied by an individual weight. The first score explained 2.1% of the total variation in biomarkers but was determined mainly by leptin and not by the other biomarkers, and the third and the fourth score explained only a low variation in biomarkers (≦0.2%). Therefore, we only considered the second RRR dietary pattern score [dietary pattern score obtained by combining ethnic groups ($RRRDS_{comb}$)] in the subsequent analyses, which explained 4.8% of the total variation in food groups, 1.0% of the total variation in all 4 biomarkers, 2.8% of adiponectin variation, 0.05% of leptin variation, 0.3% of CRP variation, and 0.7% of TG variation. When applying RRR in each ethnic/racial group separately, the second set of RRR scores [dietary pattern scores derived ethnicity-specifically ($RRRDS_{ethni}$)] explained 3.0–5.3% of the total variation in foods and 1.2–3.6% of the total variation in all 4 biomarkers across ethnic groups.

The $RRRDS_{comb}$ was characterized by low consumption of processed and red meat, white rice, and sugar-sweetened beverages (SSBs) and high consumption of whole grains, fruit, yellow-orange and green vegetables, and low-fat dairy (**Table 1**). These food groups showed factor loadings ≥0.2 combining ethnic groups and were significantly correlated with $RRRDS_{comb}$ across ethnic groups, with $r > 0.2$, except for white rice in African Americans and Latinos.

The comparison of $RRRDS_{ethni}$ (**Table 2**) and $RRRDS_{comb}$ (Table 1) showed similarities between the 2 scores. With regard to ethnicity-specific characteristics, the pattern for African Americans was additionally characterized by low poultry and shellfish intakes, the pattern for Japanese Americans by low egg consumption, the pattern for Latinos by high nut and fish consumption, the pattern for Native Hawaiians by low French fries and poultry consumption and high consumption of nuts and coffee, and the pattern for whites by high legume consumption. Individuals in the upper tertile of the $RRRDS_{comb}$ were more likely to be older, female, physically active, a never smoker, with a low BMI and a higher education (with the exception of Japanese Americans) (**Table 3**).

Association of dietary pattern with biomarkers measured ~10 y later

The $RRRDS_{comb}$ was positively correlated with adiponectin and inversely with leptin, CRP, and TGs after adjustment (**Table 4**); and ≥1 of the biomarkers showed significant correlations with the food groups. Multivariable-adjusted adiponectin means significantly increased across $RRRDS_{comb}$ tertiles across most groups, whereas mean TG, CRP, and leptin significantly decreased across most groups (**Figure 1**). Except for a lack of association of the $RRRDS_{ethni}$ with leptin in African Americans and with CRP in Native Hawaiians, the $RRRDS_{ethni}$ correlated positively with adiponectin and negatively with leptin, CRP, and TGs in all ethnic/racial groups (data not shown).

TABLE 4 Pearson correlations between the original (nonsimplified) $RRRDS_{comb}$ and food groups with biomarkers in the biomarker subcohort[1]

	Adiponectin		Leptin		CRP		TGs	
	r	Partial *r*	*r*	Partial *r*	*R*	Partial *r*	*r*	Partial *r*
Dietary pattern score	0.17*	0.05*	−0.02*	−0.06*	−0.05*	−0.03*	−0.08*	−0.05*
Food groups								
Processed meat	−0.07*	<0.01	−0.03*	0.02*	0.02*	0.02	0.02*	<0.01
Red meat	−0.09*	−0.02*	0.01	0.02*	0.03*	0.02*	0.05*	0.03*
SSBs	−0.06*	−0.01	−0.01	<0.01	0.02	0.01	0.02*	0.01
Diet drink	−0.04*	−0.01	0.10*	0.04*	0.04*	0.01	0.01	−0.01
White rice	−0.07*	<0.01	−0.06*	0.01	−0.02*	−0.01	0.02*	0.02
Whole grains	0.05*	0.01	−0.05*	−0.04*	−0.05*	−0.03*	−0.04*	−0.02*
Fruit	0.09*	−0.01	0.07*	<0.01	<0.01	−0.01	−0.03*	−0.01
Yellow-orange vegetables	0.08*	<0.01	0.05*	−0.02*	<0.01	−0.01	−0.03*	−0.02
Green vegetables	0.07*	<0.01	0.07*	<0.01	−0.02	−0.03*	−0.02	−0.01
Low-fat dairy	0.06*	0.01	0.04*	<0.01	−0.02	−0.02*	−0.02*	−0.01

[1] n = 10,008. Values are Pearson correlation coefficients and partial Pearson correlation coefficients adjusted for age (continuous), BMI (continuous), total energy intake (continuous; kilojoules per day), sex, education (≤12 y, 13–15 y, or ≥16 y), race/ethnicity (white, African American, Native Hawaiian, Japanese American, or Latino), physical activity (<30 or ≥30 min/d), and smoking status (never smoker, past smoker, or current smoker) between the original $RRRDS_{comb}$ and food group intakes with adiponectin, leptin, CRP, and TGs. The $\log_e$-transformed biomarker concentrations were also adjusted for ethnicity, CRP additionally for season of blood draw to account for seasonal fluctuations, and food groups for energy intake and ethnicity by using the residual method. Only food groups with factor loadings ≥0.2 are shown. *$P < 0.05$. CRP, C-reactive protein; $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; SSB, sugar-sweetened beverage.

Simplification of pattern score and application in the full cohort

The simplified (see Methods) $RRRDS_{comb}$ (food variables, n = 10) had a correlation coefficient of 0.82 with the original $RRRDS_{comb}$ (food variables, n = 41) in the biomarker subcohort, and the simplified $RRRDS_{ethni}$ (food variables, n = 8–11) had correlation coefficients ranging from 0.71 to 0.82 with the original $RRRDS_{ethni}$ (food variables, n = 41) across ethnic groups indicating a certain score similarity.

The simplified RRR-derived pattern showed moderate correlations with the previously published a priori index scores (7) (r = 0.65, 0.51, 0.42, and 0.67 for HEI-2010, AHEI-2010, aMED, and DASH, respectively) and factor analysis–derived patterns (8) (r = −0.43, 0.44, and 0.49 for the "fat and meat," "vegetables," and "fruit and milk" factors, respectively) in the full cohort.

Dietary pattern and incident T2D in the full cohort

We identified 8687 incident T2D cases between cohort entry and 2010 among 155,316 participants of the full MEC (mean follow-up time = 14.8 y). Comparing extreme tertiles, the $RRRDS_{comb}$ was significantly related to a 16–28% T2D risk reduction in the combined analysis and across ethnic groups (**Table 5**). RRR scores 1, 3 and 4 were not related to T2D risk (data not shown). When comparing extreme tertiles of the $RRRDS_{ethni}$, there was a significant 16–31% reduction in T2D risk across ethnic groups in the multivariable-adjusted model (**Table 6**), with apparently stronger associations in Native Hawaiians.

Sensitivity analyses

After additional adjustment of all biomarker values for BMI before their use as response variables in the RRR, diet soft drinks were no longer an important component of the second RRR pattern, whereas the other foods identified as important in the main analysis still showed factor loadings >0.15; and a similar association of the pattern with biomarkers and slightly weaker associations with T2D as in the main analysis was observed. The exclusion of participants with lipid-lowering and anti-inflammatory medication use, nonfasting participants, participants with acute inflammation, and participants with extreme energy intake yielded similar results in the RRR analysis; and all participants who provided blood yielded similar results compared with the main analysis in the Cox regression (data not shown).

Discussion

With the use of the RRR method, we derived a dietary pattern low in processed and red meat, SSBs, diet soft drinks, and white rice and high in whole grains, fruit, yellow-orange and green vegetables, and low-fat dairy. This dietary pattern was longitudinally associated with T2D-related biomarkers of inflammation, dyslipidemia, and adipokines. Comparing extreme tertiles, the pattern was associated with a 16–28% lower T2D incidence in 5 ethnic groups. Ethnicity-specific derived dietary patterns showed similar characteristics with small ethnic differences and yielded a similar magnitude of a 16–31% risk reduction for T2D across ethnic groups comparing extreme tertiles, with an apparently stronger risk reduction in Native Hawaiians.

The simplified RRR-derived pattern showed moderate correlations with the previously published a priori index scores (7) and factor analysis–derived patterns (8) in the full cohort. The RRR-derived dietary pattern was significantly inversely associated with T2D risk in 5 ethnic groups, whereas the associations of the a priori indexes HEI-2010, AHEI-2010, aMED, and DASH had larger effect estimates with T2D in whites than in Native Hawaiians and Japanese Americans (7); and factor analysis–derived dietary patterns were not significantly associated with T2D in Native Hawaiians (8) in previous analyses within the Hawaii component of the MEC. These discrepancies in findings may be due to the fact that the RRR method provides an advantage over classic

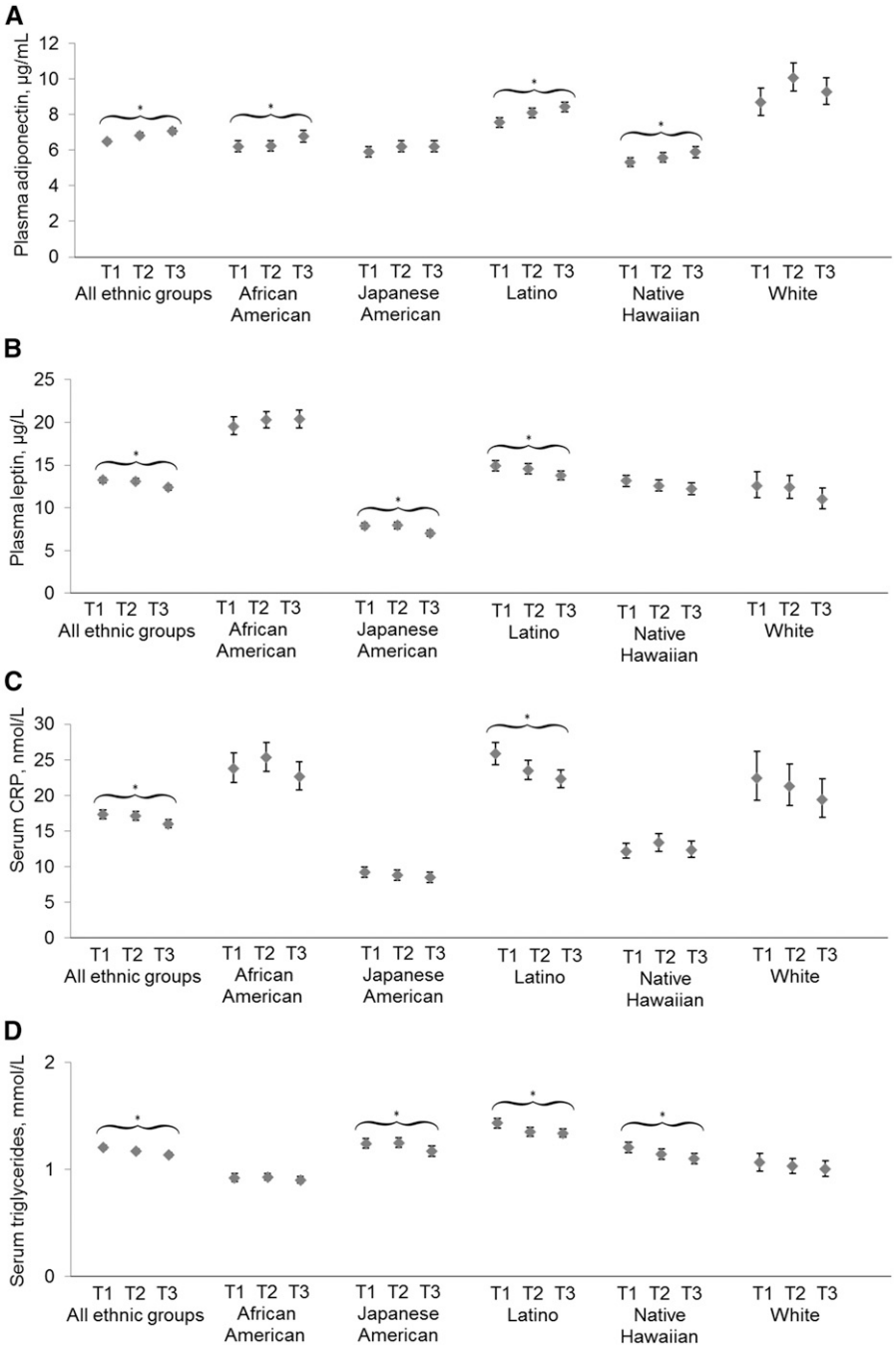

FIGURE 1 Multivariable-adjusted geometric means and 95% CIs of $\log_e$-transformed adiponectin (A), leptin (B), CRP (C), and TGs (D) across original $RRRDS_{comb}$ tertiles stratified by ethnic group (n = 10,008). Adjusted for age (continuous), BMI (continuous), total energy intake (continuous; kilojoules per day), sex, education (≤12 y, 13–15 y, or ≥16 y), physical activity (<30 or ≥30 min/d), smoking status (never smoker, past smoker, or current smoker), and race/ethnicity (for combined analysis). *P-trend < 0.05. CRP, C-reactive protein; $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; T, tertile.

data-driven methods by incorporating previous knowledge of the relation between diet and biomarkers as intermediate measures of disease risk and, therefore, is more likely to identify a disease-related dietary pattern (9). A priori dietary indexes were originally created and tested in individuals of European and African-American (for DASH) heritage and therefore food consumption patterns of other ethnic groups are not as well represented in the a priori indexes. In contrast, the RRR method identifies dietary patterns that typically exist in the study population, including foods consumed by different ethnic groups.

Whereas some of the important food components of the current RRR-derived pattern were also components of some of the previously (7) investigated a priori indexes (e.g., processed and red meat, whole grains, SSBs, fruit), the food components "white

TABLE 5 Ethnicity-specific HRs (95% CIs) of T2D by tertile of the simplified $RRRDS_{comb}$ in the full MEC[1]

	Tertile of simplified dietary pattern score			Continuous simplified dietary pattern score[2]
	1	2	3	
Ethnic groups combined				
$RRRDS_{comb}$ score range	−7.94 to −0.37	−0.37 to 0.30	0.30 to 13.3	
Diabetes cases/population at risk, *n*/*n*	3654/52,526	2695/51,566	2338/51,224	8687/155,316
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.87 (0.83, 0.92)	0.79 (0.75, 0.84)	0.91 (0.89, 0.93)
African American				
$RRRDS_{comb}$ score range	−7.28 to −0.37	−0.37 to 0.30	0.30 to 10.7	
Diabetes cases/population at risk, *n*/*n*	496/7712	503/9086	370/7588	1369/24,386
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.92 (0.81, 1.05)	0.81 (0.70, 0.94)	0.93 (0.87, 0.99)
Japanese American				
$RRRDS_{comb}$ score range	−6.85 to −0.37	−0.37 to 0.30	0.30 to 6.78	
Diabetes cases/population at risk, *n*/*n*	1268/16,591	796/14,372	748/14,778	2812/45,741
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.88 (0.80, 0.96)	0.84 (0.76, 0.93)	0.93 (0.89, 0.97)
Latino				
$RRRDS_{comb}$ score range	−7.94 to −0.37	−0.37 to 0.30	0.30 to 10.9	
Diabetes cases/population at risk, *n*/*n*	706/10,495	639/10,950	529/10,674	1874/32,119
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.94 (0.84, 1.05)	0.81 (0.72, 0.92)	0.89 (0.85, 0.93)
Native Hawaiian				
$RRRDS_{comb}$ score range	−7.57 to −0.37	−0.37 to 0.30	0.30 to 13.3	
Diabetes cases/population at risk, *n*/*n*	422/3761	216/2545	257/3145	895/9451
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.84 (0.71, 1.00)	0.83 (0.71, 0.99)	0.95 (0.89, 1.00)
White				
$RRRDS_{comb}$ score range	−6.41 to −0.37	−0.37 to 0.30	0.30 to 8.62	
Diabetes cases/population at risk, *n*/*n*	762/13,967	541/14,613	434/15,039	1737/43,619
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.83 (0.74, 0.93)	0.72 (0.63, 0.81)	0.87 (0.83, 0.92)

[1] *n* = 155,316. Participants from the biomarker subcohort were excluded in the full cohort for independent samples. The simplified dietary pattern score was calculated as the sum of unweighted standardized intakes of food items, based on the RRR results derived by combining ethnic groups (fruit + low-fat dairy + green vegetables + yellow-orange vegetables + whole grains – processed meat – red meats – white rice – sugar-sweetened beverages – diet soft drinks). Stratified by age (continuous), adjusted for BMI (continuous), total energy intake (continuous; kilojoules per day), sex, education (≤12 y, 13–15 y, or ≥16 y), race/ethnicity (white, African American, Native Hawaiian, Japanese American, or Latino; only in combined analysis), physical activity (<30 or ≥30 min/d), and smoking status (never smoker, past smoker, or current smoker). MEC, Multiethnic Cohort; ref, reference; RRR, reduced rank regression; $RRRDS_{comb}$, dietary pattern score obtained by combining ethnic groups; T2D, type 2 diabetes.

[2] *z*-Standardized (mean = 0, SD = 1).

rice," "green vegetables," "yellow-orange vegetables," and "low-fat dairy" are newly identified foods that had not been part of the a priori indexes, with the exception of "low-fat dairy" in the DASH index. A meta-analysis reported that white rice consumption was consistently directly associated with T2D risk, particularly in Asian populations (31), which may be due to its contribution to dietary glycemic load (e.g., white rice explained 58.5% of the dietary glycemic load in Japanese women) (32). Moreover, in comparison with minimally processed whole grains, white rice is poor in nutrients, including insoluble fiber and magnesium, that have been associated with lower T2D risk in the MEC (33).

Recent meta-analyses of cohort studies also described inverse associations of low-fat dairy (34) and green leafy and yellow vegetables (35) with T2D risk. A high consumption of low-fat dairy products may reduce T2D risk via weight reduction and lower inflammation (36), consistent with the weak inverse correlation between low-fat dairy products and CRP seen in the current analysis. Green vegetables are rich sources of fiber, polyphenols, vitamin C, and other bioactive compounds that contain anti-inflammatory properties (37), consistent with a weak inverse correlation of green vegetable intake with CRP in the current analysis. Processed meat (38), red meat (38, 39), and SSBs (40) were also directly related to T2D risk; and fruit intake (35) was inversely related to T2D risk in meta-analyses. Independent of BMI, red meat correlated inversely with adiponectin and directly with CRP, leptin, and TGs in the current study, whereas previous studies described no association of red meat with adiponectin (41, 42) and TGs (42) and a direct association of red meat with CRP, which was not independent of BMI (41, 42). With the use of the BMI-adjusted biomarker concentrations as responses in RRR, only diet soft drinks were no longer an important component of the pattern. Confounding or reverse causation might therefore explain the identification of diet soft drinks as a pattern component.

An increasing number of studies have used the RRR approach to identify dietary patterns predictive of T2D by using different intermediate markers as response variables, including inflammatory markers (10–13), HOMA-IR (14, 15), HbA1c (10, 15), fasting glucose (15), adiponectin (10, 16), HDL cholesterol (10, 16), and TGs (16). The explained variation in biomarkers in these studies ranged from 3.9% to 8% (10–16), which is slightly higher than in our findings (1.2–3.6% across ethnic groups). This may be due to the assessment of biomarkers in the follow-up examination after ~10 y in the current analysis, whereas previous studies mainly used cross-sectional assessments of diet and biomarkers.

Because previous analyses (10, 16, 43) suggested that a single RRR dietary pattern is unlikely to explain several different and independent pathways, the dietary pattern in the present analysis

TABLE 6 Ethnicity-specific HRs (95% CIs) for T2D by tertile of the simplified $RRRDS_{ethni}$ in the full MEC cohort[1]

	Tertile of simplified dietary pattern score			Continuous simplified dietary pattern score[2]
	1	2	3	
African American				
Diabetes cases/population at risk, *n*/*n*	494/8165	475/8147	400/8074	1369/24,386
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.99 (0.87, 1.13)	0.84 (0.73, 0.96)	0.93 (0.88, 0.99)
Japanese American				
Diabetes cases/population at risk, *n*/*n*	1201/15,509	864/15,173	747/15,059	2812/45,741
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.88 (0.80, 0.96)	0.82 (0.74, 0.91)	0.92 (0.88, 0.96)
Latino				
Diabetes cases/population at risk, *n*/*n*	722/10,802	609/10,690	543/10,627	1874/32,119
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.89 (0.80, 1.00)	0.83 (0.74, 0.94)	0.92 (0.88, 0.97)
Native Hawaiian				
Diabetes cases/population at risk, *n*/*n*	394/3245	303/3156	198/3050	895/9451
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.90 (0.77, 1.05)	0.69 (0.57, 0.82)	0.89 (0.84, 0.95)
White				
Diabetes cases/population at risk, *n*/*n*	766/14,725	529/14,488	442/14,406	1737/43,619
Multivariable-adjusted HR (95% CI)	1.00 (ref)	0.81 (0.72, 0.91)	0.76 (0.67, 0.86)	0.89 (0.85, 0.93)

[1] n = 155,316. Participants from the biomarker subcohort were excluded in the full cohort for independent samples. The simplified dietary pattern score was calculated as the sum of unweighted standardized intakes of food groups, based on RRR pattern analysis derived separately for African Americans (yellow-orange vegetables + cruciferous vegetables + green vegetables + tomatoes + low-fat dairy + whole grains − processed meat − red meat − poultry − shellfish − other potatoes and tubers), Japanese Americans (green vegetables + yellow-orange vegetables + legumes + fruit + low-fat dairy + whole grains − processed meat − red meat − eggs − white rice), Latinos (fish + green vegetables + yellow-orange vegetables + fruit + nuts + low-fat dairy + whole grains − processed meat − red meat − sugar-sweetened beverages), Native Hawaiians (coffee + alcohol + nuts + cottage cheese − red meat − poultry − diet soft drinks − other potatoes and tubers − French-fried potatoes − white rice), and whites (legumes + cruciferous vegetables + green vegetables + other vegetables + fruit − red meat − white rice − sugar-sweetened beverages). For African Americans, Native Hawaiians, Japanese Americans, and Latinos, the factor loadings signs were reversed to enable better comparison across ethnic groups. Stratified by age (continuous), adjusted for BMI (continuous), total energy intake (continuous; kilojoules per day), sex, education (≤12 y, 13–15 y, or ≥16 y), physical activity (<30 or ≥30 min/d), and smoking status (never smoker, past smoker, or current smoker). MEC, Multiethnic Cohort; ref, reference; RRR, reduced rank regression; $RRRDS_{ethni}$, dietary pattern scores derived ethnicity-specifically; T2D, type 2 diabetes.

[2] z-Standardized (mean = 0, SD = 1).

seems to mainly explain variations in adiponectin and TGs and, to a lesser extent, in CRP and scarcely in leptin. We detected some similarities in important food groups in the present analysis compared with the patterns identified in largely white study populations that used biomarkers of similar pathways (10–13) (e.g., processed and red meat, SSBs, diet soft drinks, and fruit). Unique to the present pattern, "white rice" was also an important pattern score component in whites, Native Hawaiians, and Japanese Americans. This may be due to the Hawaiian part of the MEC population who reported a higher intake of white rice, with mean values of 56, 242, and 275 g/d in whites, Native Hawaiians, and Japanese Americans compared with 37 and 37 g/d in African Americans and Latinos, respectively, in the MEC. Only 1–2% of participants in the Nurses' Health Study reported consumption of ≥107 g white rice/d (44).

The pattern derived in the combined population compared with the patterns derived in specific ethnic/racial groups showed differences in strength of associations with the biomarkers, with the least significant associations observed in whites and African Americans. This may be due to the smaller sample size of whites in the biomarker subcohort and ethnic differences in the biomarker profile (5). In all groups, the ethnicity-specific derived patterns were characterized by a high contribution of processed meat, red meat, SSBs, and fruit; and the relation of the patterns with T2D risk was mainly similar to that of the combined pattern. These findings indicate a common ground of protective and harmful foods across ethnic groups. Ethnic differences were observed with regard to the importance of specific food groups: for example, white rice was an important contributor in whites, Native Hawaiians, and Japanese Americans. These ethnic groups primarily reside in Hawaii where food intake often combines elements of Eastern and Western diets. Further examples of ethnic differences include a high contribution of legumes to the pattern associated with whites, poultry with African Americans and Native Hawaiians, eggs with Japanese Americans, and nuts and fish with Latinos. This may be due to the consumption of different foods within food groups across ethnic groups.

As a major advantage in the present study, information on dietary intake from cohort entry, assessed biomarker information ~10 y later, and long-term follow-up for T2D risk with a validated diagnosis (i.e., self-report confirmed by administrative data) was available. The use of a QFFQ designed for the relevant ethnic populations allowed us to study a heterogeneous population with wide variations in dietary habits, which may contribute to the differences of strength of association in the dietary pattern–T2D association that the authors found. Multiple data sources for T2D status were available, making it possible to create a robust definition of diagnosis that provides high specificity and avoids misclassification.

RRR shares a number of limitations with the data-driven approaches, including that the identified food intake patterns are specific to the population under study. This can partially be addressed by validation efforts in differing populations, such as were performed in 2 previous studies (45, 46) and across ethnic groups in the present study. The consistency of the results across 5 different ethnic populations in the current analysis adds considerably to the validity of the findings. In this analysis, we reduced the data dependency of the pattern variables by constructing

simplified dietary patterns, and we applied a split-sample approach to address this issue, although generalizability to other populations remains a concern. Although the validation of the QFFQ with 24-h recalls indicated acceptable results (20), the 1-time dietary assessment by self-reported QFFQ was a limitation.

In conclusion, the results of the current analysis in a prospective cohort with 5 ethnic groups suggest that a diet high in fruit, low-fat dairy, green and yellow-orange vegetables, and whole grains and low in processed and red meat, white rice, SSBs, and diet soft drinks may lower the risk of developing T2D, possibly by influencing adipokine concentrations, inflammation, and dyslipidemia. The findings of the current analysis highlight the importance of studying relations between nutrition, T2D-related biomarkers, and T2D risk across ethnic groups with different food-consumption habits and varying T2D rates. These findings need to be validated in other nonwhite cohorts before translating the results into specific recommendations, including ethnicity-specific food recommendations for high-risk groups.

Acknowledgments

The authors' responsibilities were as follows—SJ and GM: formulated the research question, conducted the statistical analysis, interpreted the results, and finalized the manuscript; SJ: wrote the first draft of the manuscript; LRW: contributed to the statistical analysis and interpretation of the results; JK, AAF, MBS, IC, LKF, KRM, LNK, CAH, LRW, LLM, and CJB: critically reviewed the draft manuscript and contributed to the revised draft of the manuscript; IC, LNK, LLM, CAH, KRM, and LRW: designed the overall cohort study and were responsible for the data and biospecimen collections; GM: had primary responsibility for final content and was responsible for the integrity of the work as a whole; and all authors: read and approved the final manuscript.

References

1. International Diabetes Federation. IDF Diabetes Atlas [Internet]. 7th ed. Brussels (Belgium): International Diabetes Federation; 2015 [cited 2016 Oct 1]. Available from: http://www.diabetesatlas.org.
2. Maskarinec G, Erber E, Grandinetti A, Verheus M, Oum R, Hopping BN, Schmidt MM, Uchida A, Juarez DT, Juarez K, et al. Diabetes incidence based on linkages with health plans: the multiethnic cohort. Diabetes 2009;58:1732–8.
3. Grandinetti A, Kaholokula JK, Theriault AG, Mor JM, Chang HK, Waslien C. Prevalence of diabetes and glucose intolerance in an ethnically diverse rural community of Hawaii. Ethn Dis 2007;17:250–5.
4. Joseph JJ, Echouffo-Tcheugui JB, Carnethon MR, Bertoni AG, Shay CM, Ahmed HM, Blumenthal RS, Cushman M, Golden SH. The association of ideal cardiovascular health with incident type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis. Diabetologia 2016;59:1893–903.
5. Morimoto Y, Conroy SM, Ollberding NJ, Kim Y, Lim U, Cooney RV, Franke AA, Wilkens LR, Hernandez BY, Goodman MT, et al. Ethnic differences in serum adipokine and C-reactive protein levels: the Multiethnic Cohort. Int J Obes (Lond) 2014;38:1416–22.
6. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13:3–9.
7. Jacobs S, Harmon BE, Boushey CJ, Morimoto Y, Wilkens LR, Le Marchand L, Kröger J, Schulze MB, Kolonel LN, Maskarinec G. A priori-defined diet quality indexes and risk of type 2 diabetes: the Multiethnic Cohort. Diabetologia 2015;58:98–112.
8. Erber E, Hopping BN, Grandinetti A, Park SY, Kolonel LN, Maskarinec G. Dietary patterns and risk for diabetes: the Multiethnic Cohort. Diabetes Care 2010;33:532–8.
9. Hoffmann K, Schulze MB, Schienkiewitz A, Nothlings U, Boeing H. Application of a new statistical method to derive dietary patterns in nutritional epidemiology. Am J Epidemiol 2004;159:935–44.
10. Heidemann C, Hoffmann K, Spranger J, Klipstein-Grobusch K, Mohlig M, Pfeiffer AF, Boeing H; European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study Cohort. A dietary pattern protective against type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study cohort. Diabetologia 2005;48:1126–34.
11. Schulze MB, Hoffmann K, Manson JE, Willett WC, Meigs JB, Weikert C, Heidemann C, Colditz GA, Hu FB. Dietary pattern, inflammation, and incidence of type 2 diabetes in women. Am J Clin Nutr 2005;82:675–84; quiz: 714–5.
12. Liese AD, Weis KE, Schulz M, Tooze JA. Food intake patterns associated with incident type 2 diabetes: the Insulin Resistance Atherosclerosis Study. Diabetes Care 2009;32:263–8.
13. McGeoghegan L, Muirhead CR, Almoosawi S. Association between an anti-inflammatory and anti-oxidant dietary pattern and diabetes in British adults: results from the national diet and nutrition survey rolling programme years 1–4. Int J Food Sci Nutr 2015;67:553–61.
14. McNaughton SA, Mishra GD, Brunner EJ. Dietary patterns, insulin resistance, and incidence of type 2 diabetes in the Whitehall II Study. Diabetes Care 2008;31:1343–8.
15. Batis C, Mendez MA, Gordon-Larsen P, Sotres-Alvarez D, Adair L, Popkin B. Using both principal component analysis and reduced rank regression to study dietary patterns and diabetes in Chinese adults. Public Health Nutr 2016;19:195–203.
16. Frank LK, Jannasch F, Kroger J, Bedu-Addo G, Mockenhaupt FP, Schulze MB, Danquah I. A dietary pattern derived by reduced rank regression is associated with type 2 diabetes in an urban Ghanaian population. Nutrients 2015;7:5497–514.
17. Sattar N, Wannamethee SG, Forouhi NG. Novel biochemical risk factors for type 2 diabetes: pathogenic insights or prediction possibilities? Diabetologia 2008;51:926–40.
18. Chen GC, Qin LQ, Ye JK. Leptin levels and risk of type 2 diabetes: gender-specific meta-analysis. Obes Rev 2014;15:134–42.
19. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, Pike MC, Stram DO, Monroe KR, Earle ME, Nagamine FS. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000;151:346–57.
20. Stram DO, Hankin JH, Wilkens LR, Pike MC, Monroe KR, Park S, Henderson BE, Nomura AM, Earle ME, Nagamine FS, et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol 2000;151:358–70.
21. Setiawan VW, Virnig BA, Porcel J, Henderson BE, Le Marchand L, Wilkens LR, Monroe KR. Linking data from the Multiethnic Cohort Study to Medicare data: linkage results and application to chronic disease research. Am J Epidemiol 2015;181:917–9.
22. State of California Office of Statewide Health Planning and Development. Health Care Information Division—patient discharge data [Internet]. 2015 [cited 2015 Sep 7]. Available from: http://www.oshpd.ca.gov/HID/Products/PatDischargeData/PublicDataSet/index.html.
23. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412–9.
24. Lopez-Jaramillo P. The role of adiponectin in cardiometabolic diseases: effects of nutritional interventions. J Nutr 2016;146(Suppl):422S–6S.
25. Galán I, Valencia-Martn JL, Guallar-Castillón P, Rodriguez-Artalejo F. Alcohol drinking patterns and biomarkers of coronary risk in the Spanish population. Nutr Metab Cardiovasc Dis 2014;24:189–97.

26. Oliveira A, Rodriguez-Artalejo F, Lopes C. Alcohol intake and systemic markers of inflammation—shape of the association according to sex and body mass index. Alcohol Alcohol 2010;45:119–25.
27. Sluik D, Brouwer-Brolsma EM, de Vries JH, Geelen A, Feskens EJ. Associations of alcoholic beverage preference with cardiometabolic and lifestyle factors: the NQplus Study. BMJ Open 2016;6:e010437.
28. Weikert C, Schulze MB. Evaluating dietary patterns: the role of reduced rank regression. Curr Opin Clin Nutr Metab Care 2016;19:341–6.
29. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65(4 Suppl):1220S–8S; discussion: 1229S–31S.
30. Willett WC. Nutritional epidemiology. New York: Oxford University Press; 1998.
31. Hu EA, Pan A, Malik V, Sun Q. White rice consumption and risk of type 2 diabetes: meta-analysis and systematic review. BMJ 2012;344:e1454.
32. Murakami K, Sasaki S, Takahashi Y, Okubo H, Hosoi Y, Horiguchi H, Oguma E, Kayama F. Dietary glycemic index and load in relation to metabolic risk factors in Japanese female farmers with traditional dietary habits. Am J Clin Nutr 2006;83:1161–9.
33. Hopping BN, Erber E, Grandinetti A, Verheus M, Kolonel LN, Maskarinec G. Dietary fiber, magnesium, and glycemic load alter risk of type 2 diabetes in a multiethnic cohort in Hawaii. J Nutr 2010;140:68–74.
34. Gao D, Ning N, Wang C, Wang Y, Li Q, Meng Z, Liu Y, Li Q. Dairy products consumption and risk of type 2 diabetes: systematic review and dose-response meta-analysis. PLoS One 2013;8:e73965.
35. Wang PY, Fang JC, Gao ZH, Zhang C, Xie SY. Higher intake of fruits, vegetables or their fiber reduces the risk of type 2 diabetes: a meta-analysis. J Diabetes Investig 2016;7:56–69.
36. Bordoni A, Danesi F, Dardevet D, Dupont D, Fernandez AS, Gille D, Dos Santos CN, Pinto P, Re R, Rémond D, et al. Dairy products and inflammation: a review of the clinical evidence. Crit Rev Food Sci Nutr 2017;57:2497–525.
37. Brighenti F, Valtuena S, Pellegrini N, Ardigo D, Del Rio D, Salvatore S, Piatti P, Serafini M, Zavaroni I. Total antioxidant capacity of the diet is inversely and independently related to plasma concentration of high-sensitivity C-reactive protein in adult Italian subjects. Br J Nutr 2005; 93:619–25.
38. Micha R, Michas G, Mozaffarian D. Unprocessed red and processed meats and risk of coronary artery disease and type 2 diabetes—an updated review of the evidence. Curr Atheroscler Rep 2012;14:515–24.
39. Pan A, Sun Q, Bernstein AM, Schulze MB, Manson JE, Willett WC, Hu FB. Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. Am J Clin Nutr 2011;94: 1088–96.
40. Imamura F, O'Connor L, Ye Z, Mursu J, Hayashino Y, Bhupathiraju SN, Forouhi NG. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. Br J Sports Med 2016;50:496–504.
41. Ley SH, Sun Q, Willett WC, Eliassen AH, Wu K, Pan A, Grodstein F, Hu FB. Associations between red meat intake and biomarkers of inflammation and glucose metabolism in women. Am J Clin Nutr 2014; 99:352–60.
42. Montonen J, Boeing H, Fritsche A, Schleicher E, Joost HG, Schulze MB, Steffen A, Pischon T. Consumption of red meat and whole-grain bread in relation to biomarkers of obesity, inflammation, glucose metabolism and oxidative stress. Eur J Nutr 2013;52:337–45.
43. Hoffmann K, Zyriax BC, Boeing H, Windler E. A dietary pattern derived to explain biomarker variation is strongly associated with the risk of coronary artery disease. Am J Clin Nutr 2004;80: 633–40.
44. Sun Q, Spiegelman D, van Dam RM, Holmes MD, Malik VS, Willett WC, Hu FB. White rice, brown rice, and risk of type 2 diabetes in US men and women. Arch Intern Med 2010;170:961–9.
45. InterAct Consortium. Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct Study. Diabetologia 2014;57:321–33.
46. Imamura F, Lichtenstein AH, Dallal GE, Meigs JB, Jacques PF. Generalizability of dietary patterns associated with incidence of type 2 diabetes mellitus. Am J Clin Nutr 2009;90:1075–83.