1

2

3   **Supplementary Appendix**

4

5   Genetic Influences on Dietary Variety: Results from a Twin Study

6

7

8

9   Benjamin Scheibehenne, Peter M. Todd, Stéphanie M. van den Berg, Peter K. Hatemi, Lindon

10   J. Eaves, & Christian Vogler

11

12

13 List of 99 Foods from the FFQ (Twin study)

Food Name	Category	Item #
apples, applesauce, pears	fruits and vegetables	1
bananas	fruits and vegetables	2
peaches, apricots (canned, frozen or dried, whole year)	fruits and vegetables	3
peaches, apricots, nectarines (fresh, in season)	fruits and vegetables	4
canataloupe	fruits and vegetables	5
watermelon (in season)	fruits and vegetables	6
strawberries (fresh, in season)	fruits and vegetables	7
oranges	fruits and vegetables	8
orange or grapefruit juice	fruits and vegetables	9
grapefruit	fruits and vegetables	10
tang, start breakfast drinks	fruits and vegetables	11
other fruit juices, fortified fruit drinks	fruits and vegetables	12
any other fruit, including berries, fruit cocktail	fruits and vegetables	13
string beans, green beans	fruits and vegetables	14
peas	fruits and vegetables	15
chili with beans	fruits and vegetables	16
other beans such as baked beans, pintos, kidney beans, limas	fruits and vegetables	17
corn	fruits and vegetables	18
winter squash, baked squash	fruits and vegetables	19
tomatoes, tomato juice	fruits and vegetables	20
red chili sauce, taco sauce, salsa picante	fruits and vegetables	21
broccoli	fruits and vegetables	22
cauliflower or brussel sprouts	fruits and vegetables	23
spinach (raw)	fruits and vegetables	24
spinach (cooked)	fruits and vegetables	25
mustard greens, turnip greens, collards	fruits and vegetables	26
cole slaw, cabbage sauerkraut	fruits and vegetables	27
carrot, or mixed vegetables containing carrots	fruits and vegetables	28
sweet potatoes, yams	fruits and vegetables	31
other potatoes, inc. boiled, baked and potato salad	fruits and vegetables	32
rice	fruits and vegetables	33
other vegetables, inc cooked onions, summer squash	fruits and vegetables	34
biscuits, muffins, burger tolls (inc. fast food)	salty or sugary snacks	55
white bread (inc. sandwiches), bagels, crackers etc.	salty or sugary snacks	56
salty snacks (like chips, popcorn)	salty or sugary snacks	59
peanuts, peanut butter	salty or sugary snacks	60
ice cream	salty or sugary snacks	72
frozen yogurt	salty or sugary snacks	73
doughnuts, cookies, cakes, pastries	salty or sugary snacks	74
pumpkin pie, sweet potato pie	salty or sugary snacks	75
other pies	salty or sugary snacks	76
chocolate candy	salty or sugary snacks	77
other candy, jelly, honey, brown sugar	salty or sugary snacks	78
hamburgers, cheeseburgers, meat loaf	meat	36
beef - steaks, roasts	meat	37
beef stew, pot pie with carrots, other vegetables	meat	38
liver, inc. chicken livers	meat	39
pork, inc. chops	meat	40
fried chicken	meat	41
chicken or turkey, roasted, stewed or boiled	meat	42

14   List of 99 Foods from the FFQ (continued)

Food Name	Category	Item #
whole milk and drinks with whole milk (not on cereal)	liquid foods	82
2% milk and drinks with 2% milk (not on cereal)	liquid foods	83
skimmed milk, 1% milk or buttermilk (not on cereal)	liquid foods	84
regular soft drinks	liquid foods	85
diet soft drinks	liquid foods	86
beer	liquid foods	87
wine	liquid foods	88
liquor	liquid foods	89
decaffeinated coffee	liquid foods	90
coffee, not decaf	liquid foods	91
tea (hot or iced)	liquid foods	92
lemon in tea	liquid foods	93
non-dairy creamer in coffee or tea	liquid foods	94
milk in coffee or tea	liquid foods	95
cream (real) or half-and-half in coffee or tea	liquid foods	96
sugar in coffee or tea	liquid foods	97
artificial sweetener in coffee or tea	liquid foods	98
glasses of water, not counting in coffee or tea	liquid foods	99
salad dressing, mayonnaise (inc. on sandwiches)	other	29
french fries and fried potatoes	other	30
butter, margarine or other fat on vegetables, potatoes etc.	other	35
fried fish or fish sandwich	other	43
tuna fish, tuna salad, tuna casserole	other	44
shellfish (shrimp, lobster, crab, oysters, etc.)	other	45
other fish, broiled, baked	other	46
spaghetti, lasagna, other pasta with tomato sauce	other	47
pizza	other	48
mixed dishes with cheese (like macaroni & cheese)	other	49
liverwurst	other	50
hot dogs	other	51
ham, lunch meats	other	52
vegetable soup, veg. Beef, minestrone, tomato soup	other	53
other soups	other	54
dark bread, inc. rye, whole wheat, pumpernickel	other	57
corn bread, corn muffins, corn tortillas	other	58
butter on bread or rolls	other	61
margarine on bread or rolls	other	62
gravies made with meat drippings, or white sauce	other	63
high fiber, bran or granola cereals, shredded wheat	other	64
highly fortified cereals, like total, product 19, or most	other	65
other cold cereals, like corn flakes, rice krispies	other	66
cooked cereals	other	67
sugar added to cereal	other	68
eggs	other	69
bacon	other	70
sausage	other	71
cottage cheese	other	79
other cheeses and cheese spreads	other	80
flavored yogurt	other	81

15

16   **Bayesian Model Estimate**

17   Following Van den Berg et al. (2006), the variety score (i.e. the phenotype) p_{ij} of a

18   single MZ twin i from family j was modeled as

19

20    $p(MZ)_{ij} = \mu + A_j + C_j + E_{ij}$ (1)

21

22   where μ indicates a fixed effect that represents the mean of the phenotype. For DZ

23   twins, the model was specified as

24

25    $p(DZ)_{ij} = \mu + \frac{1}{2}A_j + \frac{1}{2}A'_{ij} + C_j + E_{ij}$ (2)

26

27   where A' indicates the non-shared genetic variance that is unique for each DZ twin.

28   For singletons, who neither share a common environment nor genetic variance with anyone

29   else in our sample, the model was specified as

30

31    $p(single)_i = \mu + A_i + C_i + E_i$ (3)

32

33   All parameters were assumed to be normally distributed with mean zero such that $A_j \sim$

34    $N(0, \sigma_A^2)$, $C_j \sim N(0, \sigma_C^2)$, $C_i \sim N(0, \sigma_C^2)$, $E_{ij} \sim N(0, \sigma_E^2)$, $E_i \sim N(0, \sigma_E^2)$, $A'_{ij} \sim N(0, \frac{1}{2}\sigma_A^2)$, $A'_{ij} \sim$

35    $N(0, \frac{1}{2}\sigma_A^2)$ and $A_i \sim N(0, \sigma_A^2)$. The variances itself were modelled as precisions (i.e. $\tau = 1/\sigma^2$).

36   The phenotype was standardized to mean zero and standard deviation one prior to entering the

37   JAGS model (i.e. $\mu_p = 0$, $\sigma^2_p = 1$). The model also includes the mean effect of age as a

38   covariate (that was also standardized outside JAGS). The ACE estimates did not critically

39   depend on the inclusion or exclusion of this covariate though.

40

41

42   **Prior Distributions and Sampling**

43   For all τ's we used gamma prior distributions with shape and rate parameters set to 1

44   and 0.1 respectively. The prior on μ and on the mean effect of age were normally distributed

45   with mean zero and precision = 0.1 (i.e., a variance of 10). Given that the phenotype was

46   scaled to a mean of zero and a standard deviation of 1, these priors are fairly uninformative

47   and do not overly constrain the results a priori. The prior for the genetic correlation for DZ

48   twins with opposite sex was set to uniform with range 0 to 0.5.

49   After an initial burn-in of 5,000 iterations and some thinning of the chains, we

50   obtained a total of 60,000 relatively independent samples from the joint posterior distribution

51   from three independent Markov chains. The thinning was necessary to reduce the auto-

52   correlation within the chains that was quite strong. The samples from the three independent

53   and thinned chains mixed well as indicated by a visual inspection and the Gelman-Rubin

54   statistic, suggesting that the obtained samples represented the full range of the posterior

55   distribution.

56   **JAGS code to estimate the ACE model**

```
57    model {
58
59        # MZM
60        for (i in 1:Nmzm)
61        {
62            C.MZM[i] ~ dnorm(mu.m , tau.c.m)
63            A.MZM[i] ~ dnorm(C.MZM[i] + beta * MZM[i,3], tau.a.m)
64            MZM[i,1] ~ dnorm(A.MZM[i] , tau.e.m)
65            MZM[i,2] ~ dnorm(A.MZM[i] , tau.e.m)
66        }
67
68        # MZF
69        for (i in 1:Nmzf)
70        {
71            C.MZF[i] ~ dnorm(mu.f, tau.c.f)
72            A.MZF[i] ~ dnorm(C.MZF[i] + beta * MZF[i,3], tau.a.f)
73            MZF[i,1] ~ dnorm(A.MZF[i] , tau.e.f)
74            MZF[i,2] ~ dnorm(A.MZF[i] , tau.e.f)
75        }
76
77        #DZM
78        for (i in 1:Ndzm)
79        {
80            C.DZM[i] ~ dnorm(mu.m, tau.c.m)
81            A1.DZM[i] ~ dnorm(C.DZM[i] + beta * DZM[i,3], double.tau.a.m)
82            A2.DZM[i,1] ~ dnorm(A1.DZM[i], double.tau.a.m)
83            A2.DZM[i,2] ~ dnorm(A1.DZM[i], double.tau.a.m)
84            DZM[i,1] ~ dnorm(A2.DZM[i,1] , tau.e.m)
85            DZM[i,2] ~ dnorm(A2.DZM[i,2] , tau.e.m)
86
87        }
88
89
90        #DZF
91        for (i in 1:Ndzf)
92        {
93            C.DZF[i] ~ dnorm(mu.f, tau.c.f)
```

```
94        A1.DZF[i] ~ dnorm(C.DZF[i] + beta * DZF[i,3], double.tau.a.f)
95        A2.DZF[i,1] ~ dnorm(A1.DZF[i], double.tau.a.f)
96        A2.DZF[i,2] ~ dnorm(A1.DZF[i], double.tau.a.f)
97        DZF[i,1] ~ dnorm(A2.DZF[i,1], tau.e.f)
98        DZF[i,2] ~ dnorm(A2.DZF[i,2], tau.e.f)
99      }
100
101
102     # OS
103     for (i in 1:Ndos)
104     {
105        C.DOS[i] ~ dnorm(0, tau.c.dos)
106        A1.DOS[i] ~ dnorm(C.DOS[i] + beta * DOS[i,3], tau.a1.dos)
107        A2.DOS[i,1] ~ dnorm(A1.DOS[i], tau.a2.dos.m)   # first twin is MALE
108        A2.DOS[i,2] ~ dnorm(A1.DOS[i], tau.a2.dos.f)   # second twin FEMALE
109        DOS[i,1] ~ dnorm(A2.DOS[i,1] + mu.m,  tau.e.m) # first twin is MALE
110        DOS[i,2] ~ dnorm(A2.DOS[i,2] + mu.f , tau.e.f) # second twin FEMALE
111     }
112
113
114     # separate variance components for males and females, with gamma priors
115     #Prior for hhIndex
116     tau.a.m ~ dgamma(1,.1)
117     tau.a.f ~ dgamma(1,.1)
118     tau.c.m ~ dgamma(1,.1)
119     tau.c.f ~ dgamma(1,.1)
120     tau.e.m ~ dgamma(1,.1)
121     tau.e.f ~ dgamma(1,.1)
122
123
124     double.tau.a.m <- 2*tau.a.m
125     double.tau.a.f <- 2*tau.a.f
126
127     tau.c.dos <- sqrt(tau.c.m*tau.c.f)
128     tau.a1.dos <- sqrt(tau.a.m*tau.a.f)/ dos.genetic.correlation
129     tau.a2.dos.m <- 1/ ( 1/tau.a.m - 1/tau.a1.dos)
130     tau.a2.dos.f <- 1/  ( 1/tau.a.f - 1/tau.a1.dos)
131
132     dos.genetic.correlation ~ dunif(0,0.5) #genetic correlation for dzos
133
134     mu.m ~ dnorm(0,.1)
135     mu.f ~ dnorm(0,.1)
136
137     beta ~ dnorm(0,.1) #effect of age
138 }
139
140
```

Model Input

Nmzm – number of monozygotic male twin pairs

Nmzf – number of monozygotic female twin pairs

Ndzm – number of dizygotic male twin pairs

Ndzf – number of dizygotic female twin pairs

Ndos – number of dizygotic opposite sex twin pairs

MZM – Nmzm x 3 matrix containing the phenotype of all monozygotic male twins

MZF – Nmzf x 3 matrix containing the phenotype of all monozygotic female twins

DZM – Ndzm x 3 matrix containing the phenotype of all dizygotic male twins

DZF – Ndzf x 3 matrix containing the phenotype of all dizygotic female twins

153   DOS – Ndos x 3 matrix containing the phenotype of all dizygotic twins with opposite sex

154

155   Note: Each row contains the phenotype of one twin pair. The third column indicates the age.

156   Singletons were appended to the respective matrices (one singleton twin per row)
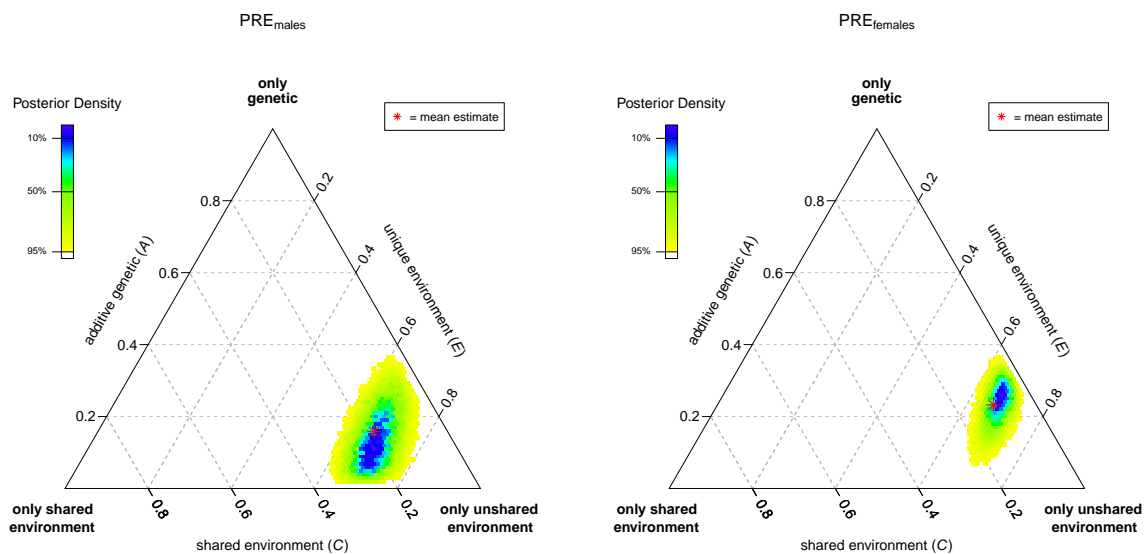
157

158

159

160   **Joint posterior distributions of the ACE model for all three variety measures**
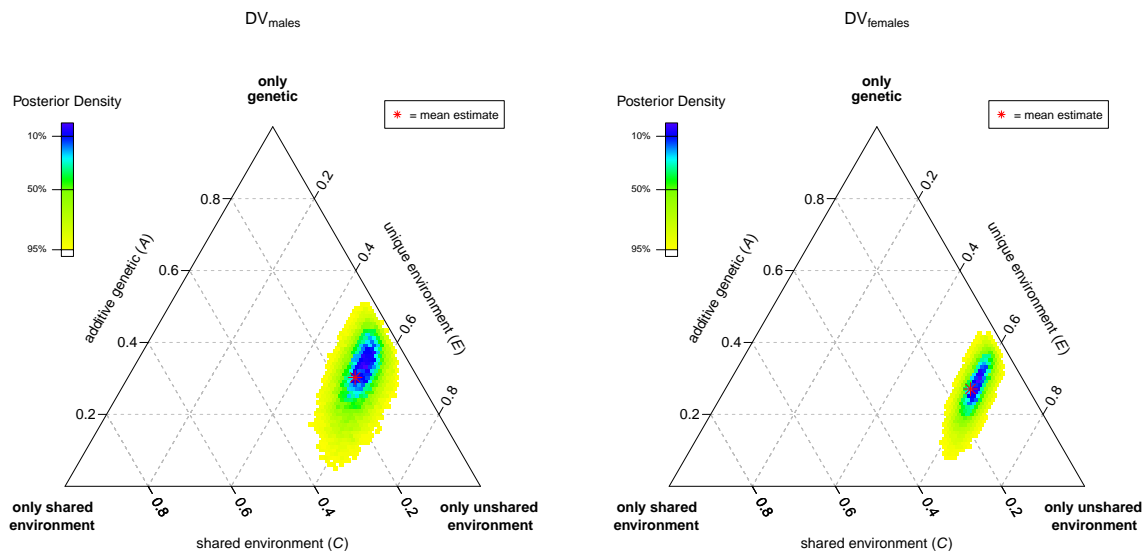
161
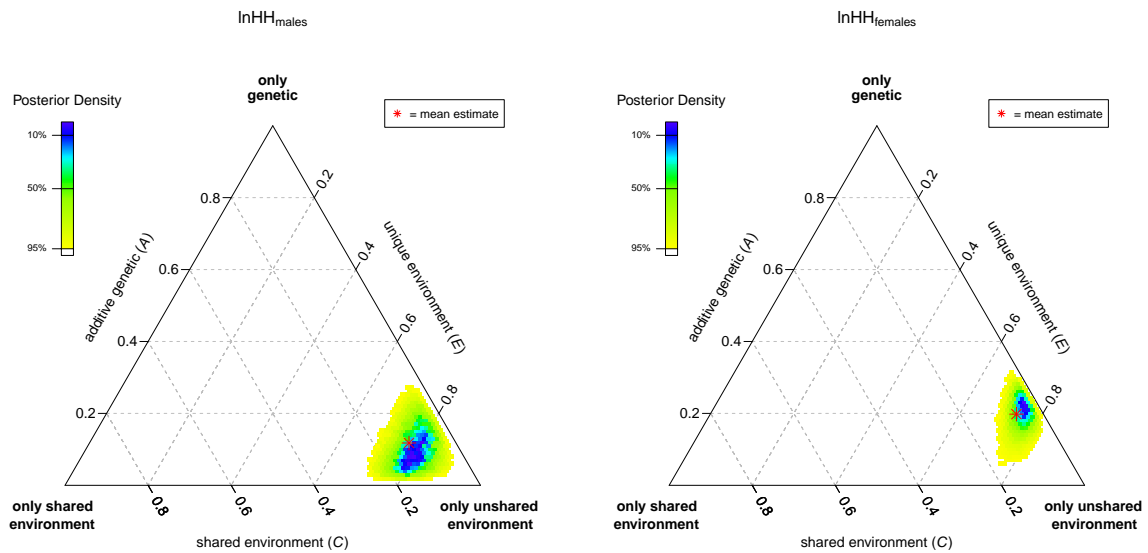
162

163

164

165 **Details of the Linear Regression Analysis**
166

Predictor	Estimate	Std.Error	t	p
(Intercept)	0.63930	0.03120	20.49	<.001
sex_male	0.01505	0.00356	4.22	<.001
age	0.00217	0.00069	3.16	.002
age^2	-0.00003	0.00001	-4.01	<.001
family income	0.00064	0.00012	5.29	<.001
education (years)	0.00308	0.00197	1.57	.117
religion_protestant	0.01485	0.00419	3.55	<.001
religion_jewish	-0.04739	0.00766	-6.19	<.001
religion_other	-0.01448	0.00785	-1.84	.065
religion_none	0.00016	0.00948	0.02	.987
family size	0.00350	0.00165	2.12	.034

167 **Table A1**: Linear Regression with DV as dependent variable.

168 R2 = .038; F(10; 5466) = 21.63, p < .001

169

170

Predictor	Estimate	Std.Error	t	p
(Intercept)	-2.97500	0.08969	-33.18	<.001
sex_male	-0.06949	0.00997	-6.97	<.001
age	-0.00735	0.00194	-3.78	<.001
age^2	0.00010	0.00002	5.22	<.001
family income	-0.00087	0.00034	-2.57	.010
education (years)	-0.00351	0.00570	-0.62	.538
religion_protestant	-0.02654	0.01170	-2.27	.023
religion_jewish	-0.00196	0.02134	-0.09	.927
religion_other	0.01890	0.02191	0.86	.388
religion_none	0.00213	0.02639	0.08	.936
family size	-0.01782	0.00460	-3.88	<.001

171 **Table A2**: Linear Regression with lnHH as dependent variable.

172 R^2 = .041; F(10; 5285) = 22.73, p < .001

173

174

Predictor	Estimate	Std.Error	t	p
(Intercept)	0.28360	0.02681	10.58	<.001
sex_male	0.02613	0.00298	8.77	<.001
age	0.00240	0.00058	4.13	<.001
age^2	-0.00003	0.00001	-6.36	<.001
family income	0.00049	0.00010	4.88	<.001
education (years)	0.00108	0.00170	0.63	.527
religion_protestant	0.01248	0.00350	3.57	<.001
religion_jewish	-0.01019	0.00638	-1.60	.110
religion_other	-0.00567	0.00655	-0.87	.387
religion_none	-0.00050	0.00789	-0.06	.950
family size	0.00590	0.00138	4.29	<.001

175 **Table A3:** Linear Regression with PRE as dependent variable.

176 R^2 = .079; $F(10; 5285)$ = 45.05, $p < .001$

177

178 **References**

179 Van Den Berg, S. M., Beem, L., & Boomsma, D. I. (2006). Fitting genetic models using

180   Markov chain Monte Carlo algorithms with BUGS. *Twin Research and Human*

181   *Genetics, 9(3)*, 334-342.