Familial Correlations, Segregation Analysis, and Nongenetic Correlates of Soy Isoflavone–Metabolizing Phenotypes

C. L. FRANKENFELD,*,† C. ATKINSON,* W. K. THOMAS,* E. L. GOODE,*,† A. GONZALEZ,*,‡ T. JOKELA,§ K. WÄHÄLÄ,§ S. M. SCHWARTZ,†,|| S. S. LI,* AND J. W. LAMPE*,†,‡,[1]

*Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109; †Department of Epidemiology, University of Washington, Seattle, Washington 98195; ‡Interdisciplinary Program in Nutritional Sciences, University of Washington, Seattle, Washington 98195; §Department of Chemistry, University of Helsinki, Helsinki, Finland 00014; and ||Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109

Particular intestinal bacteria metabolize the soy isoflavone daidzein to equol and *O*-desmethylangolensin (*O*-DMA), metabolites that can be identified in urine. Individuals that harbor bacteria capable of producing equol or *O*-DMA are known as equol producers (approximately 30%–50% of the population) and *O*-DMA producers (approximately 80%–90% of the population), respectively. The equol-producer phenotype has been associated with sex hormone–related outcomes in several studies. However, the bacteria responsible for these phenotypes have not yet been identified and factors that influence the manifestation of these phenotypes are not well understood. To evaluate familial clustering of and nongenetic factors associated with these phenotypes, 410 individuals from 112 families participated in phenotyping (3-day soy challenge and Day 4 spot urine collection). In segregation analyses of the equol-producer phenotype, the Mendelian dominant model provided the most parsimonious fit to the data, suggesting that the pattern of inheritance of the equol-producer phenotype is consistent with an autosomal dominant trait. This phenotype was positively associated with education (p trend = 0.01), but not with sex, smoking, or several dietary factors. Results of the segregation analyses of the *O*-DMA–producer phenotype were inconclusive; no other models provided a more parsimonious fit to the data than the general model. This phenotype was inversely associated with age in a nonlinear model (p = 0.01), positively associated with age- and sex-adjusted height (odds ratio [OR] 10-cm increase = 0.38, 95% confidence interval [CI] = 0.15, 0.95) and body mass index (kg/m^2) (OR = 0.91, 95% CI = 0.85, 0.96), but not with sex, education, smoking, or several dietary factors. These results suggest the equol-producer phenotype may be under some degree of genetic control and that there are likely other environmental factors not evaluated in the present analysis that contribute to both of these phenotypes. These results provide a foundation for further work to refine our understanding of heritable and environmental determinants of daidzein-metabolizing phenotypes. Exp Biol Med 229:902–913, 2004

Key words: daidzein; equol; *O*-desmethylangolensin; intestinal bacteria; soy; family

Introduction

Intestinal bacteria can metabolize the soy isoflavone daidzein to equol and/or *O*-desmethylangolensin (*O*-DMA) (1–4), metabolites that can be identified in host urine. To date, two distinct phenotypes of bacterial daidzein metabolism have been observed in human populations; that is, individual humans are either equol producers (approximately 30%–50% of the population; Refs. 5–10) or nonproducers and are either *O*-DMA producers (approximately 80%–90% of the population; Refs. 6, 9) or nonproducers. These phenotypes appear to be stable in individuals over time (6), suggesting that there may be some degree of genetic predisposition. In contrast to the variability observed in humans, studies with small numbers of animals observed that all rats, mice, and chimpanzees, produce equol (reviewed in Ref. 11). Although a small study of nonhuman primates observed that these mammals excrete *O*-DMA (4), variability in *O*-DMA excretion in

This work was supported by the National Institutes of Health Grants R03CA089785 and T32CA009168 (C.L.F.). Some of the results of this manuscript were obtained using the S.A.G.E. package of genetic epidemiology software, which is supported by a U.S. Public Health Service RR03655. The work was carried out partially within the EU project PHYTOHEALTH QLRT-2001-02453. This study does not necessarily reflect the views of the commission and in no way anticipates the commission's future policy in this area.

[1] To whom requests for reprints should be addressed at Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M4-B402, P.O. Box 19024, Seattle, WA 98109–1024. E-mail: jlampe@fhcrc.org

Received May 11, 2004.
Accepted June 11, 2004.

1535-3702/04/2299-0902$15.00

nonhuman mammals has not yet been evaluated systematically.

The variability in bacterial daidzein metabolism may be a determinant of human disease risk, particularly sex hormone–related conditions. The capacity to produce equol has been associated with circulating reproductive hormones in premenopausal women (12), inversely associated with mammographic density in postmenopausal women (13) and prostate cancer in Japanese men (10, 14), and positively associated with bone mineral density in response to isoflavone intervention (11). Among postmenopausal women, the capacity to produce *O*-DMA has been positively associated with urinary 2-hydroxyestrone excretion (15) and with mammographic density (13), an intermediate marker of breast cancer risk (16).

Both genetic predisposition and shared environment have been associated with overall intestinal bacteria profile (17–23) and with other markers of the presence of particular intestinal bacteria, such as the breath methane-excreter phenotype (24, 25). However, little is known about the specific factors that influence the capacity to harbor particular intestinal bacteria, including daidzein-metabolizing bacteria. We therefore conducted a family study in order to evaluate genetic and nongenetic factors associated with the equol-producer and *O*-DMA–producer phenotypes.

Materials and Methods

Study Participants. Families with at least three closely related members from at least two generations were eligible to participate; closely related included members who belonged to spousal, sibling, parental, or grandparental pairings. Participants were recruited using posted, print, radio, and Web site advertisements. After an individual was recruited into the study, study information and a consent form were mailed to other family members. Exclusion criteria for individuals were residence outside the United States, less than 10 years of age (no upper age limit was specified), allergy to soy, and chronic systemic (oral or intravenous) antibiotic use. Participants who took antibiotics within 3 months prior to the study were asked to complete the protocol at least 3 months after the completion of their antibiotic therapy. The Institutional Review Board at the Fred Hutchinson Cancer Research Center (FHCRC) approved all procedures, and written, informed consent (12 years of age or older) or assent (10 or 11 years of age) was obtained from all participants. In addition, parental consent was obtained for all participants younger than 18 years of age.

Study Procedures. When signed consent forms were received at FHCRC from at least three family members, each individual was mailed a urine collection kit (Fisher Scientific, Pittsburgh, PA), a questionnaire, and soy food items. In order to ascertain daidzein-metabolizing phenotypes, each participant supplemented his/her usual diet with a soy food item once per day for three consecutive days (soy bar, Revival Soy, ~83 mg daidzein per day or soy nuts, GeniSoy, ~10 mg daidzein per day). Because phenotype determination was based on the presence or absence of equol/*O*-DMA, the difference in daidzein dose between these two foods did not bias phenotype determination. On the morning of the fourth day, each participant collected a first-void urine sample (50–80 ml). Participants mailed urine samples and completed questionnaires to FHCRC. Urine was aliquoted and stored at –20°C until analysis.

Questionnaire Data. Questions included date of birth, race, education, general health, height, weight, nutritional supplement use, general dietary restrictions, smoking, alcohol consumption, and in relation to the soy challenge, the date of urine collection and amount of soy bar or nuts consumed each of the 3 days. For additional dietary data, participants were asked to indicate the number of servings (per day, week, or month) of particular categories of food eaten in the past month. Two separate, but similar, questionnaires were used for adults and children (individuals 10–18 years of age). The questionnaire for children excluded questions about smoking and alcohol consumption, and, specific to females, pregnancy history and hormonal contraceptive use.

Isoflavonoid Analysis. Urine samples were ether-extracted (26) and analyzed for isoflavonoids (equol, *O*-DMA, and daidzein) by gas chromatography–mass spectrometry (GC-MS) as described elsewhere (15). The mean intraassay coefficients of variation (CVs) for isoflavonoids in the quality control urine samples, measured in duplicate in each batch, were <7%. The mean interassay CVs in the quality control urine samples were <26%. Given the sensitivity of the assay, urine concentrations less than 182 nmol/l (44 ng/ml) of equol and 170 nmol/l (44 ng/ml) of *O*-DMA were considered below detectable limit. Equol and *O*-DMA producers were defined as individuals with any detectable concentration of equol and *O*-DMA, respectively. Urine samples that took more than 14 days to arrive at the laboratory after the reported date of collection were excluded from analysis ($n = 6$).

Isoflavone Stability. In preparation for measuring the isoflavonoid content of urine specimens shipped at ambient temperature, we conducted a study to determine the stability of equol and *O*-DMA in urine kept at room temperature. Four known equol producers (not from the family study sample) followed the same soy consumption and urine collection protocol as described in Study Procedures section, with the exception that urine samples were brought to FHCRC on the day of collection. One aliquot was frozen on Day 0 and additional aliquots were kept at room temperature and then frozen every 2 days for 14 days. Aliquots were subsequently analyzed for isoflavonoids.

Creatinine Analysis. To assure sufficient concentration of the urine samples, urinary creatinine concentrations were measured on aliquots from all overnight urine

samples based on a kinetic modification of the Jaffe reaction using the Roche Reagent for Creatinine (Roche Diagnostic Systems, Nutley, NJ) on a Roche Cobas Mira Plus chemistry analyzer, after first diluting the samples with distilled water 1:50. The assay was linear to 20 mg/dl and the intraassay and interassay CVs were 1%–2%.

Statistical Analysis. In order to test the stability of isoflavones left at room temperature, between-day CVs were calculated for the eight time points for each of the four urine samples.

Associations between nongenetic factors and daidzein-metabolizing phenotypes were analyzed in 410 individuals, and familial aggregation and segregation was evaluated in 366 individuals. Nine participants who reported consumption of less than half of the soy bar on any day were excluded from analysis. Forty-four individuals who completed the daidzein-metabolizing phenotyping were not included in aggregation and segregation analyses because of insufficient number of informative family members completing the study, that is, not having at least two familial relationships that were being considered in this analysis.

To evaluate familial aggregation of the daidzein-metabolizing phenotypes, uniform weight intraclass and interclass correlations (ICC) and standard errors were calculated for familial relationships between individuals using the FCOR program of S.A.G.E. (version 4.4, Case Western Reserve University, Cleveland, OH). EE software (Quantitative Genetic Epidemiology, FHCRC, Seattle, WA) was used to calculate age-adjusted ICC.

To evaluate inheritance of the equol-producer and *O*-DMA–producer phenotypes, segregation analysis was performed using the SEGREG program of S.A.G.E. under the Class A regressive logistic model (27). Mendelian inheritance was assumed to be through an autosomal locus with two alleles A and B, where the A allele was associated with the relevant phenotype. The likelihood for family data (28) was calculated as a function of the genotype-specific baseline susceptibility parameters (β_{AA}, β_{AB}, β_{BB}); an age parameter (α); the population allele frequency (qA) assuming Hardy-Weinberg equilibrium; and the probability that a parent with each genotype will transmit the A allele (τ_{AA}, τ_{AB}, τ_{BB}). Residual associations (γ) among relative pairs were allowed. Seven transmission models were fit to the data: a Mendelian dominant model, a Mendelian recessive model, a Mendelian codominant, a "no major type" model that assumes random susceptibility, an environmental model with two susceptibility groups, an environmental model with three susceptibility groups, and a general (full) model in which all parameters in the likelihood function were estimated without restriction. Parameters in each model were estimated using maximum likelihood methods, and the likelihood of each nested model was compared with that under a general model using a likelihood ratio test. Models with likelihoods that were not statistically significantly worse than the likelihood of the general model may be considered to fit the data better than models that were statistically significantly different than the general model. Akaike's information criterion (AIC), calculated as $-2(\ln L) + 2n_{parameters}$, identifies the model with the most parsimonious fit to the data.

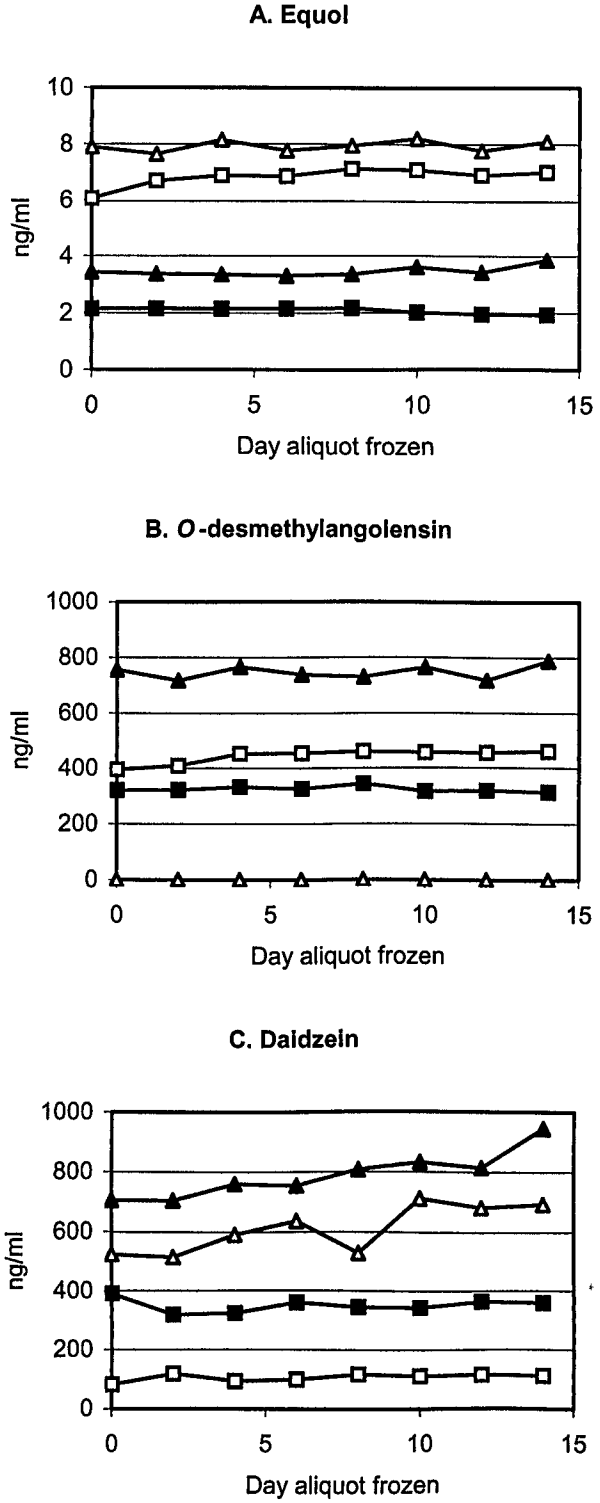

Figure 1. Concentrations of (A) equol, (B) *O*-desmethylangolensin, and (C) daidzein in urine provided from four individuals. Each individual is represented by a different symbol. One urine aliquot was frozen on Day 0, and additional aliquots were kept at room temperature and then frozen every 2 days for 14 days.

Table 1. Intraclass and Interclass Correlations (with Standard Error in Parentheses) by Familial Relationship for Equol-Producer and *O*-DMA–Producer Phenotypes in 112 Pedigrees

Relationship type	*n* pairs	Equol-producer phenotype	*O*-DMA–producer phenotype
Parent:parent (spouse:spouse)	60	0.18 (0.13)	−0.21 (0.13)
Parent:offspring	274	0.12 (0.07)	0.13 (0.07)
Father:son	50	0.25 (0.14)	0.01 (0.18)
Mother:son	68	0.16 (0.12)	0.22 (0.12)
Father:daughter	57	0.09 (0.14)	0.27 (0.13)
Mother:daughter	99	0.04 (0.11)	0.13 (0.10)
Sibling:sibling	73	0.07 (0.13)	0.22 (0.17)
Brother:brother	10	−0.23 (0.32)	−0.13 (0.43)
Sister:brother	39	0.16 (0.16)	0.22 (0.15)
Sister:sister	24	0.13 (0.22)	0.32 (0.18)
Grandparent:grandchild[a]	53	0.16 (0.15)	0.25 (0.14)
Paternal grandmother:grandson	5	0.58 (0.43)	0.38 (0.58)
Maternal grandmother:grandson	17	0.29 (0.23)	0.46 (0.19)
Maternal grandfather:granddaughter	4	−1.00 (NE[b])	−0.18 (0.22)
Maternal grandmother:granddaughter	19	−0.37 (0.24)	−0.58 (0.50)

[a] Correlations were not estimable for paternal grandfather:grandson, maternal grandfather:grandson, paternal grandmother:granddaughter, and paternal grandfather:granddaughter relationships because of small numbers of pairs (≤3 pairs)
[b] NE = Not estimable because of small numbers of pairs.

Fractional polynomial regression was also used to model the association between age and each daidzein-metabolizing phenotype (29); these models and associated plots provide a good visual representation of potentially nonlinear trends. Eight power terms (–2, –1, –0.5, 0, 0.5, 1, 2, 3) for age were evaluated, and the best fit model was selected based on deviance value. A first-degree model was chosen because no significant gains were achieved with additional terms in the models. Associations between daidzein-metabolizing phenotypes and other nongenetic factors were analyzed using generalized estimating equations (GEE). Because it was not possible to assign separate

Table 2. Parameter Estimates and Model Fit from Segregation Analysis of the Equol-Producer Phenotype in 112 Pedigrees[a]

Parameters	Mendelian dominant	Mendelian recessive	Mendelian codominant	Two-group environmental	Three-group environmental	No major type	Full model
Baseline parameter (β)							
AA	**−2.457**	**−1.497**	**−2.010**	**−0.323**	**−0.323**	**−0.323**	**−1.684**
AB	−2.457	1.367	**−2.010**	−0.323	**−0.323**	−0.323	**−1.687**
BB	**1.017**	**1.367**	**1.064**	**−0.323**	**−0.323**	−0.323	**0.724**
Age parameter (α)	**−0.016**	**−0.014**	**−0.015**	**−0.008**	**−0.008**	**−0.008**	**−0.024**
Transmission probability (τ)							
AA	1	1	1	0.708	0.778	N/A	**1.000**
AB	0.5	0.5	0.5	0.708	0.778	N/A	**0.994**
BB	0	0	0	0.708	0.778	N/A	**0.000**
Allele frequency (qA)	**0.280**	**0.778**	**0.301**	**0.708**	**0.778**	N/A	**0.184**
Residual associations (γ)							
Father:mother	0	0	0	0	0	0	0
Father:offspring	**−1.093**	**−1.276**	**−1.155**	**0.293**	**0.293**	**0.293**	**−1.151**
Mother:offspring	−1.093	−1.276	−1.155	0.293	0.293	0.293	−1.151
Sibling:sibling	−1.093	−1.276	−1.155	0.293	0.293	0.293	−1.151
Likelihood ratio statistic"−2(ln L)"	486.305	489.075	485.912	494.079	494.079	494.079	481.616
Number of parameters estimated	5	5	6	5	6	3	7
Akaike's information criterion (AIC)	496.305	499.075	497.912	504.079	506.079	500.079	495.616
Comparison to full model							
Chi-square (χ^2)	4.689	7.459	4.296	12.463	12.463	12.463	—
Degrees of freedom (df)	2	2	1	2	1	4	—
P value	0.096[b]	0.024	0.038	0.002	<0.001	0.014	—

[a] Numbers in bold are independent parameter estimates. Numbers underlined are parameters estimated on a boundary (not included in the number of parameters estimated).
[b] Model likelihood not statistically significantly worse than the likelihood of the general model.

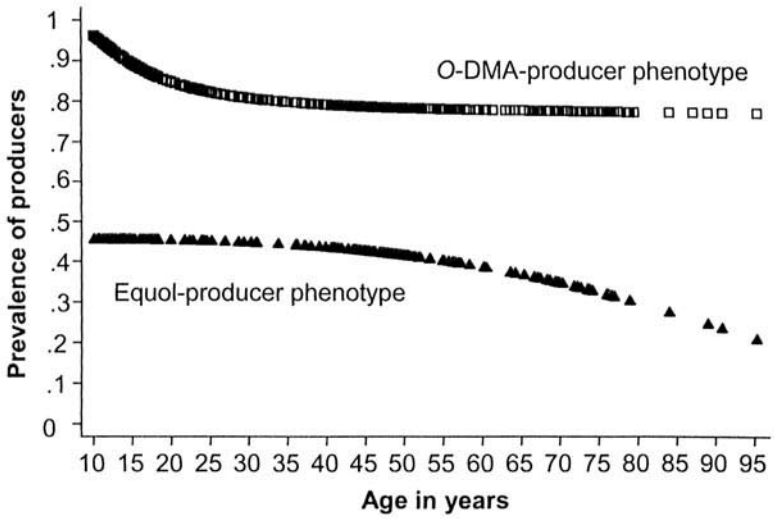

Figure 2. Predicted prevalence of producers, from a first-degree fractional polynomial regression, in relation to age for each of the equol-producer (closed triangles) and *O*-desmethylangolensin–producer (open squares) phenotypes in 410 individuals. The cube of age was selected as the best fit for the equol-producer phenotype, and the square root of age was selected as the best fit for the *O*-desmethylangolensin–producer phenotype.

correlations for each of the types of relative pairings, an exchangeable correlation across familial relationships was used to control for familial correlation. Because the hypothesis to be tested was that particular nongenetic factors give rise to daidzein-metabolizing phenotypes, equol-producer and *O*-DMA–producer phenotypes were considered as dependent variables and nongenetic factors were considered as independent variables. Effect modification by sex of several factors potentially related to sex hormone concentrations, such as smoking, anthropometry, age, and diet, was modeled for each phenotype by including an interaction term between sex and each factor using GEE as described for main effects analyses. SAS (version 8.0, SAS Institute Inc., Cary, NC) and Stata (version 8.0, Stata Corp, College Station, TX) were used.

All confidence intervals (CI) are expressed as 95% CI. Because each phenotype is a marker of the capacity to harbor particular bacteria that is distinct from the other phenotype, the factors associated with each phenotype are not necessarily expected to be the same; thus, the results and discussion are presented in relation to each phenotype rather than in relation to the exposures.

Results

Isoflavone Stability. There was no loss of equol, *O*-DMA, or daidzein in urine left at room temperature for 14 days (Fig. 1). Between-day CVs for equol concentration ranged from 2.6% to 5.4% among the four individuals tested for isoflavone stability. Among the three *O*-DMA producers, between-day CV for *O*-DMA concentration ranged from 3.4% to 5.8% among individuals. For all four individuals, between-day CV for daidzein concentration ranged from 6.7% to 13.4%. In our study, the average length of time between urine collection and processing was 4 days.

Participants and Demographics. One hundred and twelve informative families with at least two different types of relative pairs (spousal, sibling, parental, or grandparental) completed the study. Pedigrees ranged in size from three to nine family members, with a mean of four members. Participants were between 10 and 96 years of age, and the average age was 39 years. There was no difference in the prevalence of equol-producer and *O*-DMA–producer phenotypes in the 366 participants included in familial aggregation and segregation analyses compared with the overall 410 participants ($P = 0.8$, $P = 0.7$, respectively).

Daidzein-Metabolizing Phenotypes. All participants had urinary daidzein concentrations above the assay detection limit (all were $\geq$109 ng/ml) and had acceptable urinary creatinine concentrations (all were >8 mg/dl).

One hundred seventy-two (42%) individuals were equol producers, and 337 (82%) individuals were *O*-DMA producers. In our study, 202 (85%) equol nonproducers were *O*-DMA producers and 135 (78%) equol producers were *O*-DMA producers. The correlation between the equol-producer and *O*-DMA–producer phenotypes within an individual was weak (ICC = 0.09 [standard error = 0.11]), with no difference in relation to sex (ICC_{male} = 0.12 [SE = 0.09], ICC_{female} = 0.07 [SE = 0.07]).

Equol-Producer Phenotype. Weak correlations for main familial relationships were observed in relation to the equol-producer phenotype (Table 1). Correlations were stronger for father or mother with son than father or mother with daughter and positive correlations between sisters and between sister and brother and an inverse correlation between brothers were observed. Adjustment for age did not alter the overall interpretation of these results.

The fit of the Mendelian dominant model was not statistically significantly different from the general model

Table 3. Demographic and General Health Characteristics in 410 Individuals in Relation to Each Daidzein-Metabolizing Phenotype, Expressed as a Percentage of Individuals Within Each Phenotype

Characteristics	Equol				*O*-DMA			
	Producers (n = 172)	Nonproducers (n = 238)	Odds ratio	95% CI[a]	Producers (n = 337)	Nonproducers (n = 73)	Odds ratio	95% CI[a]
Demographics								
Sex								
Female	58	61	1.00	reference	61	56	1.00	reference
Male	42	39	1.07	0.73, 1.58	39	44	0.84	0.50, 1.37
Race								
Caucasian	88	88	1.00	reference	89	84	1.00	reference
Asian	5	3	1.26	0.26, 6.23	2	10	0.23	0.07, 0.74
Other	7	7	0.99	0.43, 2.30	7	7	0.94	0.25, 3.59
Education (years completed)								
≤12	8	18	1.00	reference	12	22	1.00	reference
13–15	35	20	2.01	0.95, 4.22	20	22	1.63	0.73, 3.63
16	23	23	1.86	0.90, 3.86	23	21	2.05	0.91, 4.57
>16	20	15	2.74	1.28, 5.87	16	22	1.38	0.61, 3.13
Self-reported health								
Excellent	83	83	1.00	reference	83	79	1.00	reference
Good	15	13	1.13	0.62, 2.07	13	16	0.77	0.39, 1.53
Fair/poor	1	3	0.93	0.28, 3.07	2	3	0.70	0.15, 3.30
Anthropometry[b]								
Height (10-cm increase)	—	—	0.92	0.54, 1.58	—	—	0.33	0.14, 0.78
Weight (10-kg increase)	—	—	0.99	0.93, 1.05	—	—	0.84	0.78, 0.91
Body mass index (kg/m^2)	—	—	1.01	0.95, 1.05	—	—	0.90	0.85, 0.95
Reproductive characteristics[c]								
Pregnancy history								
Never pregnant	19	18	1.00	reference	20	11	1.00	reference
Ever pregnant	81	81	0.94	0.47, 1.89	80	84	0.50	0.18, 1.44
Oral contraceptive use								
Never use	80	70	1.00	reference	74	77	1.00	reference
Current or past use	20	30	1.75	0.87, 3.53	26	23	0.87	0.37, 2.05

[a] CI = Confidence interval.
[b] Height, weight, and BMI as continuous variables, adjusted for age and sex.
[c] Women age 18 and older only (n = 81 equol producers, n = 115 equol nonproducers; n = 159 *O*-DMA producers, n = 37 *O*-DMA nonproducers).

(Table 2) and provided the most parsimonious fit to the data (AIC = 496.305) among the models compared with the general model. The fits of the Mendelian recessive and codominant, two- and three-group environmental, and no major gene models were significantly worse than the general model. Among the Mendelian models, the dominant model was not significantly different than the codominant model (P = 0.53) nor was the recessive model significantly different than the codominant model (P = 0.08). The interpretation of the other parameters (baseline parameter, age parameter, transmission probability, allele frequency, and residuals associations) does not provide meaningful information about the potential mode of transmission of these phenotypes, but these parameter estimates are presented to provide data regarding model fit.

Older individuals appeared to be less likely to be equol producers compared with younger individuals (P = 0.09) (Fig. 2). There was no association of sex, race, overall self-reported health, or anthropometric characteristics with the equol-producer phenotype (Table 3). In adults, there was a positive association of greater education and being an equol producer (P trend = 0.01). Among women, there was some suggestion that current or past use of oral contraceptives, compared with never use, was positively associated with being an equol producer.

Smoking, having companion animals, and several dietary factors were not strongly associated with the equol-producer phenotype (Table 4), but there was some suggestion that reported consumption of a low-fat diet was associated with being an equol producer.

Because of the differences in familial correlations by sex-related familial subtypes, we evaluated, on an exploratory basis, interactions between sex and several factors in relation to the daidzein-metabolizing phenotypes. In relation to the equol-producer phenotype, there were no interactions between sex and age, age-adjusted anthropometric characteristics (height, weight, and body mass index [BMI]), consumption of caffeinated beverages, or, in adults, smoking status or consumption of alcohol (P interaction > 0.3).

Table 4. Dietary Factors, Smoking, and Exposure to Animals in Relation to Each Daidzein-Metabolizing Phenotype in 410 Individuals, Expressed as a Percentage of Individuals Within Each Phenotype for Categorical Variables or as Mean (Standard Deviation) for Continuous Variables

Characteristics	Equol				*O*-DMA			
	Producers (n = 172)	Nonproducers (n = 238)	Odds ratio	95% CI[a]	Producers (n = 337)	Nonproducers (n = 73)	Odds ratio	95% CI[a]
Smoking								
Individual smoking status[b]								
Never	65	60	1.00	reference	63	60	1.00	reference
Former	27	30	0.75	0.44, 1.26	30	30	1.01	0.53, 1.90
Current	7	7	1.08	0.45, 2.60	7	9	0.73	0.26, 2.04
Household smoking								
None	77	79	1.00	reference	80	90	1.00	reference
Second-hand exposure	5	5	0.89	0.35, 2.30	8	5	0.55	0.19, 1.59
Current smoker	5	5	1.13	0.47, 2.70	10	5	0.51	0.19, 1.39
Animals								
Companion animals/livestock[c]								
Have any companion animals or livestock	70	71	0.99	0.64, 1.52	70	73	0.89	0.53, 1.49
Have dog	38	38	0.98	0.59, 1.62	30	27	0.90	0.50, 1.63
Have cat	41	44	0.95	0.58, 1.54	43	42	0.90	0.51, 1.58
Have cow or livestock	3	3	1.13	0.46, 2.81	3	3	NE[d]	NE[d]
Diet								
Special diet[c]								
Vegetarian or vegan	7	3	1.14	0.51, 2.52	6	8	0.83	0.37, 1.86
Low fat	8	1	2.28	0.93, 5.64	5	7	0.58	0.21, 1.60
Dietary supplement use[c]								
Any dietary supplements	64	68	0.92	0.61, 1.40	67	63	1.12	0.66, 1.89
Vitamin (single or multivitamin)	55	61	0.86	0.56, 1.32	60	52	1.19	0.69, 2.05
Mineral (single or multimineral)	28	34	0.79	0.46, 1.36	33	25	1.43	0.74, 2.76
Probiotic or prebiotic	4	3	1.06	0.33, 3.35	4	1	NE[d]	NE[d]
Alcohol consumption[b]								
Never	11	16	1.00	reference	13	17	1.00	reference
Former	12	15	1.23	0.53, 2.84	13	16	1.13	0.41, 3.11
Current	76	66	1.60	0.77, 3.31	71	65	1.52	0.69, 3.32
Usual preparation of meat eaten								
Only baked, broiled, or poached	42	34	1.00	reference	37	37	1.00	reference
Fried only or fried in addition to baked, broiled or poached	49	59	0.68	0.44, 1.04	56	49	1.15	0.67, 2.01
Meat not eaten	9	6	1.37	0.61, 3.10	6	10	0.73	0.28, 1.93
Usual type of grains eaten								
White	19	23	1.00	reference	21	21	1.00	reference
Whole-grain	12	9	1.60	0.74, 3.45	9	12	0.78	0.31, 2.00
Mix of white and whole-grain	68	68	1.25	0.77, 2.06	69	64	1.07	0.56, 2.01
How often fried foods eaten (weekly)								
0–2	41	35	1.00	reference	35	51	1.00	reference
3–5	42	48	0.74	0.48, 1.14	49	30	2.35	1.32, 4.22
6+	16	16	0.83	0.46, 1.48	16	16	1.37	0.66, 2.82
Weekly food groups servings								
Beans or lentils	1.8 (1.8)	1.8 (2.0)	1.01	0.92, 1.12	1.7 (2.0)	2.0 (1.9)	0.95	0.82, 1.07
Caffeinated beverages	8.6 (10)	9.5 (12)	1.00	0.98, 1.01	8.2 (10)	13 (15)	0.97	0.95, 0.99
Cereals or grains	19 (13)	19 (16)	1.00	0.98, 1.01	19 (15)	19 (14)	1.00	0.98, 1.02
Chicken or other poultry	2.4 (2.5)	2.4 (1.9)	1.01	0.92, 1.10	2.5 (2.3)	2.1 (1.7)	1.11	0.95 1.29
Dairy	12 (8.5)	11 (9.0)	1.01	0.99, 1.04	12 (8.6)	10 (9.9)	1.02	0.99, 1.05
Eggs	2.3 (2.3)	2.0 (2.2)	1.04	0.96, 1.14	2.0 (1.8)	2.6 (3.6)	0.91	0.82, 1.00
Fish	1.5 (1.8)	2.0 (4.2)	0.95	0.87, 1.03	1.8 (3.6)	1.9 (2.8)	0.99	0.93, 1.06
Fruit	13 (8.0)	12 (8.1)	1.00	0.98, 1.03	13 (8.0)	12 (8.4)	1.01	0.98, 1.04
Red meat	2.1 (2.0)	2.1 (2.4)	1.01	0.92, 1.11	2.0 (2.0)	2.5 (2.8)	0.92	0.82, 1.01
Soy	1.8 (2.8)	2.2 (5.5)	0.97	0.92, 1.02	2.1 (4.8)	1.8 (3.1)	1.02	0.95, 1.09

Table 4. (Continued)

Characteristics	Equol				*O*-DMA			
	Producers (n = 172)	Nonproducers (n = 238)	Odds ratio	95% CI[a]	Producers (n = 337)	Nonproducers (n = 73)	Odds ratio	95% CI[a]
Sweetened nondiet drinks	4.5 (6.9)	4.7 (7.1)	1.00	0.98, 1.03	4.4 (6.3)	6.0 (9.8)	0.98	0.95, 1.01
Vegetables	11 (6.6)	12 (8.0)	0.99	0.96, 1.01	11 (7.4)	12 (7.8)	0.98	0.95, 1.02

[a] CI = Confidence interval.
[b] Adults only (age 18 years or older): n = 124 equol producers, 181 equol nonproducers; n = 242 *O*-DMA producers, n = 63 *O*-DMA nonproducers.
[c] Not mutually exclusive categories. For companion animals or livestock the comparison group is no companion animals or livestock. For special diet the comparison group is no special diet. For supplement use the comparison group is no supplement use.
[d] NE: Not estimated because of small numbers.

O-DMA–Producer Phenotype. An inverse correlation in *O*-DMA–producer phenotype between parents was observed, and the correlation between siblings was stronger than the correlation between parents and children (Table 1). There was no correlation between father and son, but positive correlations for other parent:offspring subtypes were observed. Positive correlations between sisters and between sister and brother and an inverse correlation between brothers were observed. Adjustment for age did not alter the overall interpretation of these results.

In segregation analyses, the fit of the Mendelian dominant and recessive models, no major gene, and the environmental models were all statistically significantly worse than the general model (Table 5). A better fit was achieved when spousal residual associations were also estimated for these models; thus, a likelihood ratio comparison for the codominant model could not be made because the number of estimated parameters was the same as for the general model. Within the Mendelian models, the dominant model was not significantly different than the codominant model (P = 0.86), nor was the recessive model significantly different than the codominant model (p = 0.74).

There was a significant inverse association of being an *O*-DMA producer with age (P = 0.01) (Fig. 2). There was no association of sex, self-reported overall health, or education with the *O*-DMA–producer phenotype (Table 3). Asian individuals were less likely to be *O*-DMA

Table 5. Parameter Estimates and Model Fit from Segregation Analysis of the *O*-DMA–Producer Phenotype in 112 Pedigrees[a]

Parameters	Mendelian dominant	Mendelian recessive	Mendelian codominant	Two-group environmental	Three-group environmental	No major type	Full model
Baseline parameter (β)							
AA	**−0.832**	**1.870**	**−0.823**	**1.706**	**1.709**	**1.710**	**1.720**
AB	−0.832	−0.881	**−0.823**	1.706	**1.709**	1.710	**1.720**
BB	**1.729**	**−0.881**	**1.711**	**1.706**	**1.709**	1.710	**1.502**
Age parameter (α)	**−0.021**	**−0.023**	**−0.020**	**−0.021**	**−0.021**	**−0.021**	**−0.015**
Transmission probability (τ)							
AA	1	1	1	0.000	1.000	N/A	**1.000**
AB	0.5	0.5	0.5	0.000	1.000	N/A	0.000
BB	0	0	0	0.000	1.000	N/A	**0.000**
Allele frequency (qA)	**0.007**	**0.982**	**0.019**	**0.000**	**1.000**	**N/A**	**0.000**
Residual associations (γ)							
Father:mother	**−1.370**	**−1.317**	**−1.378**	**−1.379**	**−1.378**	**−1.780**	**−1.378**
Father:offspring	**1.670**	**1.706**	**1.666**	**1.667**	**1.667**	**1.662**	**1.658**
Mother:offspring	**0.981**	**0.895**	**0.990**	**0.992**	**0.990**	**0.990**	**0.966**
Sibling:sibling	**1.310**	**1.310**	**1.307**	**1.312**	**1.313**	**1.315**	**1.338**
Likelihood ratio statistic "−2(ln L)"	320.061	319.984	320.094	320.214	321.672	320.219	319.677
Number of parameters estimated	8	8	9	7	8	6	9
Akaike's information criterion (AIC)	336.061	335.984	338.094	334.214	337.672	332.219	337.677
Comparison to full model							
Chi-square (χ^2)	0.384	0.307	0.417	0.537	1.995	0.542	—
Degrees of freedom (df)	1	1	0	2	1	3	—
P value	0.535	0.580	N/E	0.765	0.158	0.910	—

[a] Numbers in bold are independent parameter estimates. Numbers underlined are parameters estimated on a boundary (not included in the number of parameters estimated).

producers, as compared with Caucasian individuals. Including adjustment for age and sex, independent and inverse associations were observed for BMI (kg/m^2) (odds ratio [OR] = 0.91, CI = 0.85, 0.96) and a 10-cm increase in height (OR = 0.38, CI = 0.15, 0.95) when these variables were included in the same model.

Smoking, having companion animals or livestock, and most dietary factors were not significantly associated with the *O*-DMA–producer phenotype (Table 4). However, lower consumption of caffeinated beverages and greater consumption of fried foods were associated with being *O*-DMA producer.

There were no interactions between sex and age, age-adjusted height or BMI, or, in adults, years of education or alcohol consumption in relation to *O*-DMA–producer phenotype (*P* interaction > 0.2). Associations were stronger among females than among males for a 10-cm increase in age-adjusted height (OR_{female} = 0.12, CI = 0.03, 0.51; OR_{male} = 0.67, CI = 0.25, 1.81; *P* interaction = 0.08), a 10-kg increase in age-adjusted weight (OR_{female} = 0.77, CI = 0.68, 0.87; OR_{male} = 0.94, CI = 0.84, 1.04; *P* interaction = 0.03), and weekly servings of caffeinated beverages (OR_{female} = 0.95, CI = 0.92, 0.99; OR_{male} = 0.99, CI = 0.96, 1.02; *P* interaction = 0.11). Compared with never-smokers, former-smoking females were less likely to be *O*-DMA producers and former-smoking men were more likely to be *O*-DMA producers (OR_{female} = 0.51, CI = 0.25, 1.06; OR_{male} = 2.10, CI = 0.64, 6.82; *P* interaction = 0.07), and current smoking was not associated with *O*-DMA–producer phenotype in relation to sex (OR_{female} = 0.97, CI = 0.96, 0.99; OR_{male} = 0.99, CI = 0.98, 1.00; *P* interaction = 0.89).

Discussion

We observed that the isoflavonoids of interest are stable in urine kept at room temperature for at least 14 days, a period that is sufficient to allow for shipping by standard mail service in the United States. These observations suggest that individuals can be recruited from a large geographic area, findings that are relevant for large-scale population studies of these metabolites and associated phenotypes.

We observed prevalences of equol producers and *O*-DMA producers similar to those observed in other studies (5–10). Our results suggest that these phenotypes are not strongly correlated within an individual. Similarly, we observed in another study, a weak, nonsignificant within-individual association; among postmenopausal women, 95% of equol producers were *O*-DMA producers, and 79% of equol nonproducers were *O*-DMA producers (15).

The results of our segregation analyses suggest that the inheritance of equol-producer phenotype may be consistent with autosomal dominant transmission, despite the fact that the familial correlations we observed were weak. These observations are not necessarily inconsistent; observing a high degree of familial correlation does not prove the existence of a genetic mechanism, and, likewise, a low degree of familial correlation does not exclude a genetic mechanism (30). Segregation analysis may be considered strong statistical evidence regarding, although not proof of, genetic control (30).

There was some suggestion that older adults were less likely to be equol producers than children or young to middle-aged adults. These results may be explained by either a cohort effect or age-related changes in intestinal bacteria populations. A cohort effect has been hypothesized for *Helicobacter pylori* (31), which colonizes in the gastrointestinal tract; lower prevalence has been observed in younger individuals and may be associated with improved hygiene over time. We would expect to observe a higher prevalence of equol producers in older individuals if changes in hygiene were associated with eradication of the equol-producing bacteria; thus, the lower prevalence we observed in older individuals is less likely to be a cohort effect. In two cross-sectional studies, bacterial diversity in healthy young adults was greater than the diversity in healthy elderly adults (32, 33), suggesting that there are age-related changes in intestinal microflora, but this has not yet been evaluated prospectively.

We observed a positive association of greater education with being an equol producer. There is no apparent biologic mechanism by which education may be associated with equol-producer phenotype. However, education may be serving as a marker for other exposures, such as geographic location, medication use, or exposure to environmental contaminants and infectious agents.

Because in other classes of mammals, all appear to be equol producers (11), we evaluated an association with having companion animals. Equol-producer phenotype was not associated with having companion animals, suggesting that shared environment with or frequent exposure to other mammals may not influence equol-producer phenotype.

Several observational studies have reported differences in dietary intake, particularly fat and fiber, in relation to the equol-producer phenotype (5, 7). However, in intervention studies of fiber and soy, no alterations in equol-producer phenotype have been observed in response to interventions (8, 34–36). We observed small or no differences in dietary factors and equol-producer phenotype. However, our evaluation of diet was not very sensitive and would have likely detected only large differences between producers and nonproducers.

The familial correlations observed for the *O*-DMA–producer phenotype are consistent with some degree of genetic predisposition. However, segregation analyses did not provide evidence for genetic or for environmental influence on the *O*-DMA–producer phenotype. We observed sex-related patterns in the correlations in relationship subtypes, suggesting a possible role for sex-linked heritability or differences in shared environmental factors or differences in response to environmental factors between

male and female parents with their children and between male and female siblings.

Age was inversely associated with the *O*-DMA–producer phenotype. Specifically, children were more likely to be *O*-DMA producers than were adults. A similar observation has been seen for intestinal colonization by *Oxalobacter formigenes*; *O. formigenes* colonization has been detected in nearly all children (37), but in only about 75%–80% of adults (38). These results suggest that factors associated with late childhood may be related to change in the intestinal microflora composition. There are several substantial changes that can occur during this time of life. Perhaps the most substantial change that occurs during this time of life is puberty. Although no studies have evaluated an association between the onset of puberty and changes in intestinal bacteria, studies of oral microflora in females in relation to the onset of puberty and during pregnancy suggest that changes in sex hormones during these times may result in alterations of the oral microbial environment (39–41).

We observed that Asian individuals were less likely to be *O*-DMA producers. Racial differences may reflect differences in genetic predisposition or in environment that influence the likelihood of being an *O*-DMA producer. However, we had few non-Caucasian participants in our study, and the difference in producer prevalence between Caucasians and Asians may have been observed by chance.

Height and BMI were significantly and independently inversely associated with *O*-DMA–producer phenotype, and these associations were stronger in females. There are no studies known to have evaluated intestinal bacteria in relation to anthropometric characteristics. In a recent hypothesis paper, Beard and Blaser (42) propose that indigenous microflora are among the many factors that influence adult height, but there is currently little evidence to support this hypothesis. Among women, because of their relationship with hormone concentrations, height and BMI may be serving as surrogates for hormone exposure (43, 44).

Smoking has been associated with oral microflora composition (45), but little is known about smoking and intestinal microflora. We observed no association between *O*-DMA–producer phenotype and smoking. However, when stratified by sex, women who were former smokers were half as likely to be *O*-DMA producers as women who were never-smokers, and men who were former smokers were twice as likely to be *O*-DMA producers as men who were never-smokers. In contrast, in another study, a significant positive association was observed between the *O*-DMA–producer phenotype and former-smoking status in postmenopausal women (13); former smokers were approximately six times as likely to be *O*-DMA producers than *O*-DMA nonproducers. Although the potential mechanism for these associations is unclear, these observations suggest that starting and stopping smoking may be associated with intestinal bacteria in a manner that is different from nonsmoking or continual smoking, and that these associations may be different in relation to sex and, among women, to menopausal status. However, no studies have yet evaluated intestinal bacteria changes in relation to smoking cessation. The existence and magnitude of an association between smoking and the capacity to produce *O*-DMA remains unclear.

There are a few considerations in the interpretation of our study. First, no correction for ascertainment bias was necessary because participants were not ascertained by phenotype. Second, a relatively simple assessment of dietary factors was used to minimize participant burden, and this may have contributed to the lack of association we observed between phenotypes and dietary factors. Third, in order to better maintain individual privacy, we did not ascertain cohabitation of family members; thus, we were unable to evaluate the degree to which cohabitation or length of cohabitation influences these phenotypes. Fourth, we did not to test for statistical significance of the ICC for equol-producer and *O*-DMA–producer phenotypes. The pairs included in the calculation of these correlations are not independent; for example, if there are three siblings, each sibling would be included twice to represent all the possible pairings of these siblings. As such, standard methods of significance testing cannot be applied. Because this is the first large-scale family study of these phenotypes and our analyses are largely hypothesis generating, we did not believe it was appropriate to apply more complex methods of significance testing. Nonetheless, the standard errors presented provide information about the precision of our estimates.

Equol-producer and *O*-DMA–producer phenotypes are markers of the presence of particular intestinal bacteria capable of metabolizing daidzein to equol and *O*-DMA, respectively. Although results from association studies have suggested potential candidate bacteria responsible for the equol-producer phenotype (46, 47) and the *O*-DMA–producer phenotype (48), the specific bacteria that are responsible for this conversion in humans have not yet been definitively identified. It is likely that there are interindividual differences in the bacteria responsible for these phenotypes (1).

An understanding of whether and how daidzein-metabolizing phenotypes can be modified may provide an avenue for prevention research in targeted populations. Identifying genetic and nongenetic factors associated with and the bacteria responsible for these daidzein-metabolizing phenotypes is a step toward such an understanding. The results of our study provide information regarding the genetic and nongenetic influences on these markers of intestinal bacteria profile. If the genetic association is confirmed by other studies, these results would provide a foundation for identifying genes associated with capacity to harbor bacteria capable of producing equol. Further evaluation is also warranted to evaluate the associations, and underlying mechanisms, we observed for several nongenetic factors with both daidzein-metabolizing pheno-

types. Overall, these results, from the largest study of these phenotypes to date, provide preliminary directions for further study of genetic and nongenetic influence on daidzein-metabolizing phenotypes.

We thank the following individuals: Dr. Lue Ping Zhao for his contribution to the study design and Kristin Woodward (FHCRC) for her help with participant recruitment.

1. Atkinson C, Berman S, Humbert O, Lampe JW. *In vitro* incubation of human feces with daidzein and antibiotics suggests interindividual differences in the bacteria responsible for equol production. J Nutr 134:596–599, 2004.
2. Chang Y-C, Nair MG. Metabolism of daidzein and genistein by intestinal bacteria. J Nat Prod 58:1892–1896, 1995.
3. Bowey E, Adlercreutz H, Rowland I. Metabolism of isoflavones and lignans by the gut microflora: a study in germ-free and human flora associated rats. Food Chem Toxicol 41:631–636, 2003.
4. Blair RM, Appt SE, Franke AA, Clarkson TB. Treatment with antibiotics reduces plasma equol concentration in cynomologus monkeys (*Macaca fascicularis*). J Nutr 133:2262–2267, 2003.
5. Rowland IR, Wisemen H, Sanders TAB, Adlercreutz H, Bowey EA. Interindividual variation in metabolism of soy isoflavones and lignans: influence of habitual diet on equol production by the gut microflora. Nutr Cancer 36:27–32, 2000.
6. Kelly GE, Nelson C, Waring MA, Joannou GE, Reeder AY. Metabolites of dietary (soya) isoflavones in human urine. Clin Chim Acta 223:9–22, 1993.
7. Lampe JW, Karr SC, Hutchins AM, Slavin JL. Urinary equol excretion with a soy challenge: influence of habitual diet. Proc Soc Exp Biol Med 217:335–339, 1998.
8. Hutchins AM, Slavin JL, Lampe JW. Urinary isoflavonoid phytoestrogen and lignan excretion after consumption of fermented and unfermented soy products. J Am Diet Assoc 95:545–551, 1995.
9. Arai Y, Uehara M, Sato Y, Kimira M, Eboshida A, Adlercreutz H, Watanabe S. Comparison of isoflavones among dietary intake, plasma concentration and urinary excretion for accurate estimation of phytoestrogen intake. J Epidemiol 10:127–135, 2000.
10. Akaza H, Miyanaga N, Takashima N, Naito S, Hirao Y, Tsukamoto T, Mori M. Is daidzein non-metabolizer a high risk for prostate cancer? A case-controlled study of serum soybean isoflavone concentration. Jpn J Clin Oncol 32:296–300, 2002.
11. Setchell KD, Brown NM, Lydeking-Olsen E. The clinical importance of the metabolite equol—a clue to the effectiveness of soy and its isoflavones. J Nutr 132:3577–3584, 2002.
12. Duncan AM, Merz-Demlow BE, Xu X, Phipps WR, Kurzer MS. Premenopausal equol excretors show plasma hormone profiles associated with lowered risk of breast cancer. Cancer Epidemiol Biomarkers Prev 9:581–586, 2000.
13. Frankenfeld CL, McTiernan A, Aiello EJ, Thomas WK, LaCroix K, Schramm J, Schwartz SM, Holt VL, Lampe JW. Mammographic density in relation to daidzein-metabolizing phenotypes in overweight, postmenopausal women. Cancer Epidemiol Biomarkers Prev 13:1156–1162. 2004.
14. Ozasa K, Nakao M, Watanabe Y, Hayashi K, Miki T, Mikami K, Mori M, Sakauchi F, Washio M, Ito Y, Suzuki K, Wakai K, Group AT. Serum phytoestrogens and prostate cancer risk in a nested case-control study among Japanese men. Cancer Sci 95:65–71, 2004.
15. Frankenfeld CL, McTiernan A, Tworoger SS, Atkinson C, Thomas WK, Stanczyk F, Marcinova S, Weigle S, Weiss NS, Holt VL, Schwartz SM, Lampe JW. Serum hormone and sex hormone binding globulin concentrations and urinary hydroxylated estrogen metabolites in postmenopausal women in relation to daidzein-metabolizing phenotypes. J Steroid Biochem Mol Biol 88:399–409. 2004.
16. Boyd NF, Lockwood GA, Martin LJ, Byng JW, Yaffe MJ, Tritchler DL. Mammographic density as a marker of susceptibility to breast cancer: a hypothesis. IARC Sci Publ 154:163–169, 2001.
17. Simhon A, Douglas JR, Drasar BS, Soothill JF. Effect of feeding on infants' faecal flora. Arch Dis Child 57:54–58, 1982.
18. Sepp E, Julge K, Vasar M, Naaber P, Bjorksten B, Mikelsaar M. Intestinal microflora of Estonian and Swedish infants. Acta Paediatr 86:956–961, 1997.
19. Edlund C, Beyer G, Hiemer-Bau M, Ziege S, Lode H, Nord CE. Comparative effects of moxifloxacin and clarithromycin on the normal intestinal microflora. Scand J Infect Dis 32:81–85, 2000.
20. Satokari RM, Vaughan EE, Akkermans AD, Saarela M, De Vos WM. Polymerase chain reaction and denaturing gradient gel electrophoresis monitoring of fecal *Bifidobacterium* populations in a prebiotic and probiotic feeding trial. Syst Appl Microbiol 24:227–231, 2001.
21. Brigidi P, Swennen E, Vitali B, Rossi M, Matteuzzi D. PCR detection of *Bifidobacterium* strains and *Streptococcus thermophilus* in feces of human subjects after oral bacteriotherapy and yogurt consumption. Int J Food Microbiol 81:203–209, 2003.
22. Zoetendal EG, Akkermans ADL, Akkermans-van Vliet WM, de Visser JAGM, de Vos WM. The host genotype affects the bacterial community in the human gastrointestinal tract. Microb Ecol Health Dis 13:129–134, 2001.
23. Malaty HM, Engstrand L, Pedersen NL, Graham DY. *Helicobacter pylori* infection: genetic and environmental influences. A study of twins. Ann Intern Med 120:982–986, 1994.
24. Florin THJ, Zhu G, Kirk KM, Martin NG. Shared and unique environmental factors determine the ecology of methanogens in the humans and rats. Am J Gastroenterol 95:2872–2879, 2000.
25. Flatz G, Czeizel A, Metneki J, Flatz SD, Kuhnau W, Jahn D. Pulmonary hydrogen and methane excretion following ingestion of an unabsorbable carbohydrate: a study of twins. J Pediatr Gastroenterol Nutr 4:936–941, 1985.
26. Heinonen S, Wähälä K, Adlercreutz H. Identification of isoflavone metabolites dihydrodaidzein, dihydrogenistein, 6′-OH-*O*-DMA, and *cis*-4-OH-equol in human urine by gas chromatography-mass spectroscopy using authentic reference compounds. Anal Biochem 274:211–219, 1999.
27. Bonney GE. Regressive logistic models for familial disease and other binary traits. Biometrics 42:611–625, 1986.
28. Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. Hum Hered 21:523–542, 1971.
29. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Appl Stat 43:429–467, 1994.
30. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of Genetic Epidemiology. Monographs in Epidemiology and Biostatistics. New York: Oxford University Press, Vol. 22: 1993.
31. Roosendaal R, Kuipers EJ, Buitenwerf J, van Uffelen C, Meuwissen SG, van Kamp GJ, Vandenbroucke-Grauls CM. *Helicobacter pylori* and the birth cohort effect: evidence of a continuous decrease of infection rates in childhood. Am J Gastroenterol 92:1480–1482, 1997.
32. Hopkins MJ, Macfarlane GT. Changes in predominant bacterial populations in human faeces with age and with *Clostridium difficile* infection. J Med Microbiol 51:448–454, 2002.
33. Hopkins MJ, Sharp R, Macfarlane GT. Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. Gut 48:198–205, 2001.
34. Lampe JW, Skor HE, Li S, Wähälä K, Howald WN, Chen C. Wheat bran and soy protein feeding do not alter urinary excretion of the isoflavan equol in premenopausal women. J Nutr 131:740–744, 2001.
35. Kirkman LM, Lampe JW, Campbell DR, Martini MC, Slavin JL.

Urinary lignan and isoflavonoid excretion in men and women consuming vegetable and soy diets. Nutr Cancer 24:1–12, 1995.
36. Karr SC, Lampe JW, Hutchins AM, Slavin JL. Urinary isoflavonoid excretion in humans is dose-dependent at low to moderate levels of soy protein consumption. Am J Clin Nutr 66:46–51, 1997.
37. Sidhu H, Enatska L, Ogden S, Williams WN, Allison MJ, Peck AB. Evaluating children in the Ukraine for colonization with the intestinal bacterium *Oxalobacter formigenes*, using a polymerase chain reaction-based detection system. Mol Diagn 2:89–97, 1997.
38. Kwak C, Jeong BC, Kim HK, Kim EC, Chox MS, Kim HH. Molecular epidemiology of fecal *Oxalobacter formigenes* in healthy adults living in Seoul, Korea. J Endourol 17:239–243, 2003.
39. Jensen J, Liljemark W, Bloomquist C. The effect of female sex hormones on subgingival plaque. J Periodontol 52:599–602, 1981.
40. Nakagawa S, Fujii H, Machida Y, Okuda K. A longitudinal study from prepuberty to puberty of gingivitis. Correlation between the occurrence of *Prevotella intermedia* and sex hormones. J Clin Periodontol 21:658–665, 1994.
41. Muramatsu M, Takaesu Y. Oral health status related to subgingival bacterial flora and sex hormones in salivia during pregnancy. Bull Tokyo Dent Coll 35:139–151, 1994.
42. Beard AS, Blaser MJ. The ecology of height: the effect of microbial transmission on human height. Perspect Biol Med 45:475–498, 2002.
43. Sowers MF, Beebe JL, McConnell D, Randolph J, Jannaush M. Testosterone concentrations in women aged 25–50 years: associations with lifestyle, body composition, and ovarian status. Am J Epidemiol 153:256–264, 2001.
44. Verkasalo PK, Appleby PN, Allen NE, Davey G, Adlercreutz H, Key TJ. Soya intake and plasma concentrations of daidzein and genistein: validity of dietary assessment among eighty British women (Oxford arm of the European Prospective Investigation into Cancer and Nutrition). Br J Nutr 86:415–421, 2001.
45. Haffajee AD, Socransky SS. Relationship of cigarette smoking to the subgingival microbiota. J Clin Periodontol 28:377–388, 2001.
46. Ueno T, Uchiyama S, Kikuchi N. The role of intestinal bacteria on biological effects of soy isoflavones in humans. J Nutr 132:594S, 2002.
47. Tsangalis D, Ashton JF, McGill AEJ, Shah NP. Enzymic transformation of isoflavone phytoestrogens in soymilk by B-glucosidase-producing *Bifidobacteria*. J Food Sci 67:3104–3113, 2002.
48. Schoefer L, Mohan R, Braune A, Birringer M, Blaut M. Anaerobic C-ring cleavage of genistein and daidzein by *Eubacterium ramulus*. FEMS Microbiol Lett 208:197–202, 2002.