IGF-I and *IGFBP-3* Polymorphisms in Relation to Circulating Levels among African American and Caucasian Women

Aimee A. D'Aloisio,[1] Jane C. Schroeder,[4] Kari E. North,[4] Charles Poole,[4] Suzanne L. West,[3] Gregory S. Travlos,[2] and Donna D. Baird[1]

[1]Epidemiology Branch and [2]Cellular and Molecular Pathology Branch, National Institute of Environmental Health Sciences; [3]Health, Social, and Economics Research, RTI International, Research Triangle Park, North Carolina and [4]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Abstract

Circulating insulin-like growth factor-I (IGF-I) and IGF-binding protein-3 (IGFBP-3) levels have been associated with common diseases. Although family-based studies suggest that genetic variation contributes to circulating IGF-I and IGFBP-3 levels, analyses of associations with multiple *IGF-I* and *IGFBP-3* single nucleotide polymorphisms (SNP) have been limited, especially among African Americans. We evaluated 30 *IGF-I* and 15 *IGFBP-3* SNPs and estimated diplotypes in association with plasma IGF-I and IGFBP-3 among 984 premenopausal African American and Caucasian women. In both races, *IGFBP-3* rs2854746 (Ala^{32}Gly) was positively associated with plasma IGFBP-3 (CC versus GG mean difference among Caucasians, 631 ng/mL; 95% confidence interval, 398-864; African Americans, 897 ng/mL; 95% confidence interval, 656-1,138), and *IGFBP-3* diplotypes with the rs2854746 GG genotype had lower mean IGFBP-3 levels than reference diplotypes with the CG genotype, whereas *IGFBP-3* diplotypes with the CC genotype had higher mean IGFBP-3 levels. *IGFBP-3* rs2854744 (-202 A/C) was in strong linkage disequilibrium with rs2854746 in Caucasians only, but was associated with plasma IGFBP-3 in both races. Eight additional *IGFBP-3* SNPs were associated with ≥5% differences in mean IGFBP-3 levels, with generally consistent associations between races. Twelve *IGF-I* SNPs were associated with ≥10% differences in mean IGF-I levels, but associations were generally discordant between races. Diplotype associations with plasma IGF-I did not parallel *IGF-I* SNP associations. Our study supports that common *IGFBP-3* SNPs, especially rs2854746, influence plasma IGFBP-3 levels among African Americans and Caucasians but provides less evidence that *IGF-I* SNPs affect plasma IGF-I levels. (Cancer Epidemiol Biomarkers Prev 2009;18(3):954–66)

Introduction

Insulin-like growth factor-I (IGF-I), a peptide with structural similarities to insulin, has been implicated in many biological processes, including cell cycle regulation, differentiation, proliferation, hormone secretion, and apoptosis. IGF-binding proteins (IGFBP) help regulate the activity of IGFs by influencing their bioavailability and degradation and may also have independent effects through interactions with cell surface molecules (1, 2). IGFBP-3 binds ~90% of circulating IGF-I (3) and has also been reported to inhibit growth and promote apoptosis (4, 5) independent of its effects on IGF-I.

Circulating IGF-I levels, and to a lesser extent IGFBP-3 levels, have been studied in association with cardiovascular disease, diabetes, and cancer (6-8). Estimates from twin- or family-based studies suggest that genetic factors may account for up to 50% of the interindividual variation in plasma IGF-I levels (9, 10) and up to 60% of the variation in plasma IGFBP-3 levels (9, 11). In adults, age is the nongenetic factor most consistently associated with IGF-I blood levels, with lower levels associated with advancing age (3, 12-21). Women have lower circulating IGF-I (13, 16, 17, 19, 21, 22) but higher IGFBP-3 levels (13, 16, 17, 21, 22) than men, and two studies have suggested that African American women have higher circulating IGF-I levels than Caucasian women (17, 23).

Our research goal was to investigate relations between *IGF-I* and *IGFBP-3* polymorphisms and their circulating protein levels among African American and Caucasian women. Prior analysis of dense single nucleotide polymorphisms (SNP) and IGF-I and IGFBP-3 levels among African Americans has been limited to the Multiethnic Cohort Study, which included a random sample of ~150 African Americans in their IGF SNP analyses (24). We selected 45 SNPs in *IGF-I* and *IGFBP-3* and examined whether these SNPs and their estimated diplotypes (paired haplotypes) were associated with plasma IGF-I and IGFBP-3 levels in premenopausal African American and Caucasian women who participated in the National Institute of Environmental Health Sciences (NIEHS) Uterine Fibroid Study (UFS).

Materials and Methods

Study Population. The current study population consisted of 984 premenopausal women (582 African

Received 9/12/08; revised 12/18/08; accepted 12/29/08; published OnlineFirst 2/24/09.

Grant support: Intramural Research Program of the NIH, NIEHS and NIH Office of Research on Minority Health; NIEHS grant P30ES10126 (C. Poole). Genotyping was supported by the University of North Carolina-GlaxoSmithKline Center of Excellence in Pharmacoepidemiology and Public Health.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Requests for reprints: Aimee D'Aloisio, Epidemiology Branch, National Institute of Environmental Health Sciences, P.O. Box 12233, Mail Drop A3-05, Research Triangle Park, NC 27709. Phone: 919-541-3225; Fax: 919-541-2511. E-mail: daloisio@niehs.nih.gov

doi:10.1158/1055-9965.EPI-08-0856

Americans and 402 Caucasians) who participated in the NIEHS UFS and had available DNA samples. The UFS was designed to estimate the prevalence of uterine leiomyomata (fibroids) among African American and Caucasian women and to evaluate potential etiologic factors for fibroids. Details of the parent study were described previously (25, 26). Briefly, a random sample of 2,384 George Washington University female health plan members, ages 35 to 49 years, was identified for potential enrollment into the parent study (25, 26). The study was approved by Institutional Review Boards at NIEHS and George Washington University, and the consent form specified use of specimens for genetic polymorphism analyses.

UFS eligibility criteria were met by 1,786 of the 2,102 women who consented to eligibility screening. Most ineligible women were excluded because they no longer attended the health plan clinic where the parent study was based (71%) or they had been misidentified as a 35- to 49-year-old female (16%). Enrollment occurred from 1996 to 1999. Approximately 20% of eligible women refused participation, resulting in a total of 1,430 participants in the parent study (26). Demographic characteristics, reproductive history, smoking status, and alcohol use were assessed from telephone interviews and self-administered questionnaires. Body weight was measured at the clinic visit.

We restricted the current study to women who self-identified as African American or Caucasian (n = 1,323) to facilitate race-specific analyses and excluded postmenopausal women because they did not attend the UFS clinic visit for ultrasound screening and blood collection (n = 178). Race and menopausal status criteria for the current study were met by 1,145 women, and DNA was extracted for 984 of the 1,003 women with blood samples.

Sample Collection and Assays. Fasting blood samples were collected by venipuncture, and plasma was stored at −80°C. Plasma IGF-I was measured at NIEHS using a double-antibody RIA by extraction method (Nichols Institute Diagnostics), with a reported detection limit of 0.06 ng/mL. Plasma IGFBP-3 was measured at NIEHS by a double-antibody immunoradiometric assay (Diagnostic Systems Laboratories), with a reported detection limit of 0.05 ng/mL. The mean interassay coefficient of variation on replicate quality-control samples was 8.8% for IGF-I and 4.2% for IGFBP-3.

Genomic DNA was extracted from whole blood using a phenol/chloroform procedure or a modified salt precipitation protocol (GenQuik Protocol; Orochem Technologies).

Genetic Polymorphisms. Race-specific tag SNPs in *IGF-I* and *IGFBP-3* were selected using Genome Variation Server software.[5] We used the Seattle SNPs database as the African American and Caucasian reference population for *IGF-I* and the HapMap database for *IGFBP-3* because it has not been evaluated by Seattle SNPs. We expanded coverage to include 5 kb outside the 5′ and 3′ ends of each gene, specified a pairwise correlation coefficient (r^2) of 0.8 to identify tag SNPs that capture variation across each gene, and selected tag SNPs with a minor allele frequency (MAF) of at least 5% among the African American or Caucasian reference populations. The Genome Variation Server software identified 29 tag SNPs for *IGF-I* and 12 tag SNPs for *IGFBP-3* that met these criteria. In addition, we selected five SNPs *a priori* including one nonsynonymous *IGF-I* SNP (rs17884626), one synonymous *IGF-I* SNP (rs3729846), two nonsynonymous *IGFBP-3* SNPs (rs2854746 and rs9282734), and an *IGFBP-3* promoter SNP (rs2854744, -202 A/C) evaluated previously in association with circulating IGFBP-3 levels and health outcomes (24, 27-35).

Genotyping was done using the TaqMan genotyping approach (36-38) at the The University of North Carolina at Chapel Hill. Allele-specific oligonucleotide probes for 39 selected SNPs were purchased from Applied Biosystems (ABI) ''TaqMan Validated and Coding SNP or Pre-Designed SNP Genotyping Assays.'' ABI attempted to develop custom assays for the 6 remaining SNPs through their ''Custom TaqMan SNP Genotyping Assays'' service. Two *IGF-I* tag SNPs were dropped from analyses, including one for which a custom assay could not be developed and one with a predesigned assay that did not meet ABI technical specifications. In addition, we genotyped an alternate *IGF-I* tag SNP to replace one that was inconsistent with Hardy-Weinberg equilibrium in our African American population. We genotyped 30 *IGF-I* and 15 *IGFBP-3* SNPs (Supplementary Figs. S1 and S2).

PCR amplification was done on an ABI GeneAmp PCR System 9700 thermal cycler with dual 384-well blocks, and endpoint plates were read using the ABI 7900HT system. Fluorescent VIC and 6-FAM reporter dyes differentiated wild-type and variant alleles. The Sequence Detection System 2.3 software automatically called alleles; experienced operators reviewed the software output. The samples' DNA concentrations were validated using a NanoDrop ND-1000 spectrophotometer before dilution to 5 ng/μL using DNA-grade sterile water. Samples were placed into 96-well microtiter plates with two blank and two known DNA standard (Control DNA CEPH Individual 1347-02; ABI) samples and were subsequently aliquoted into 384-well PCR plates. Quality-control measures included blinded genotyping of 28 duplicate samples representing 22 women, which produced concordant results for all samples. The overall call rate was 98.8%, and only 5 women had <50% of complete allele calls for the 45 SNPs assayed. We confirmed that SNP genotype frequencies were consistent with Hardy-Weinberg equilibrium within each racial group using the exact test statistic with 1 df (α = 0.01; ref. 39).

Diplotype Estimation. To simultaneously evaluate associations between linked polymorphic loci and their plasma protein levels, we estimated race-specific *IGF-I* and *IGFBP-3* diplotypes. SNPs were excluded from race-specific diplotype analyses if their MAF in our study population was <5% for tag SNPs or <3% for *a priori* SNPs within the racial group being evaluated. Women missing genotype data for >50% of the SNPs considered for diplotype analyses within a gene were excluded from diplotype estimation for that gene (one Caucasian and three African Americans for *IGFBP-3* and three Caucasians and two African Americans for *IGF-I*). We

[5] Genome Variation Server version 1.04. Seattle (WA): Seattle SNPs Program for Genomic Applications. Available from: http://gvs.gs.washington.edu/GVS/ [updated 2006 Jun 16; cited 2006 Jul 9].

Table 1. Characteristics of premenopausal Caucasian and African American women with genotype information from NIEHS UFS

Characteristics	Caucasians (n = 402), n (%)	African Americans (n = 582), n (%)
Age (y)		
35-39	137 (34.1)	219 (37.6)
40-44	135 (33.6)	205 (35.2)
≥45	130 (32.3)	158 (27.1)
Education		
High school	12 (3.0)	121 (20.8)*
Some postsecondary	33 (8.2)	265 (45.5)
College degree	133 (33.1)	123 (21.1)
Graduate degree	217 (54.0)	68 (11.7)
Missing	7 (1.7)	5 (0.9)
Current oral contraceptives †	37 (9.2)	28 (4.8)
Parity		
0	236 (58.7)	120 (20.6)
1	54 (13.4)	134 (23.0)
2	91 (22.6)	194 (33.3)
≥3	21 (5.2)	134 (23.0)
Body mass index		
Underweight-normal weight (<25)	236 (58.7)	147 (25.3)
Overweight (25 to <30)	95 (23.6)	178 (30.6)
Obese (≥30)	71 (17.7)	257 (44.2)
Smoking status		
Current smoker	31 (7.7)	172 (29.6)
Former smoker	139 (34.6)	133 (22.9)
Never smoked	232 (57.7)	277 (47.6)
Alcohol intake (past year)		
<0.5 drinks/wk	60 (14.9)	300 (51.5)
≥0.5 drinks/wk	314 (78.1)	245 (42.1)
Missing	28 (7.0)	37 (6.4)
	Mean (SD)	Mean (SD)
Plasma IGF-I (ng/mL) †	172 (58.7)	165 (74.0)
Plasma IGFBP-3 (ng/mL) †	4,524 (822.3)	4,085 (859.7)

*Includes 11 women with less than high school education.
† n missing: oral contraceptives: 3 Caucasians and 4 African Americans; Plasma IGF-I and IGFBP-3: 1 Caucasian and 8 African Americans.

examined race-specific linkage disequilibrium (LD) patterns using Haploview software (40) to identify SNPs in each gene that could be combined for estimating diplotypes. First, we identified LD blocks consisting of individual SNPs (with MAF of at least 5%) in strong LD [95% of pairwise SNP comparisons with one-sided 95% confidence intervals (95% CI) for D' within 0.70-0.98; ref. 41]. Next, we used the Tagger approach (42) in Haploview to identify pairs of redundant SNPs in strong LD (pairwise r^2 values of at least 0.8) and excluded one member of each redundant pair from diplotype estimation unless both SNPs were selected *a priori*. Race-specific pairwise r^2 values are available for *IGFBP-3* and *IGF-I* SNPs in Supplementary Tables S1 to S4.

Race-specific diplotypes representing defined groups of SNPs in each gene were estimated using PHASE version 2.1 (43, 44), which allocates the most likely diplotype to each subject, with the prior assumption that frequently observed haplotypes with less ambiguity due to homozygosity are more probable. PHASE also provides a posterior probability estimate that expresses the uncertainty associated with each diplotype assignment. To reduce the number of race-specific diplotype groups for analyses, we combined single (unlinked) SNPs with an adjacent LD block and adjacent LD blocks with each other if the larger groups resulted in diplotypes estimated with at least 90% certainty (posterior probability) for at least 90% of observations. Otherwise, diplotype groups were composed of individual SNPs or LD blocks. We assigned women to their most probable diplotype for each group; however, if their most probable diplotype had a posterior probability of <90%, we classified their diplotype group as missing.

Statistical Analyses. Primary statistical analyses were stratified by race and conducted using SAS version 9.1 (SAS Institute). We used linear regression to estimate associations between *IGF-I* and *IGFBP-3* SNP variants and their respective circulating protein levels. Estimated associations are unadjusted because there are no known factors other than race that would predict both plasma levels and SNP variants.

For individual SNP analyses, we generally used codominant inheritance models that estimated associations separately for heterozygous and homozygous variants relative to the reference genotype, which was defined as the most common race-specific homozygous genotype in our study population. However, when there were ≤10 women with the homozygous variant genotype, we used a dominant model that combined heterozygous and homozygous variants for comparison with the reference genotype. We evaluated the concordance of estimated associations between races (the difference in mean differences between races; Supplementary Tables S5 and S6) by combining the data for both races and applying linear regression models that included multiplicative genotype-by-race interaction terms with separate variables for race and genotypes.

Table 2. Estimated mean differences in plasma IGFBP-3 levels associated with *IGFBP-3* SNP variants relative to estimated mean IGFBP-3 levels for reference genotypes (italicized) based on linear regression models for premenopausal Caucasian and African American women

IGFBP-3 SNP*	*IGFBP-3* location	Caucasians						African Americans					
		LD block†	Diplotype group (position)†	MAF (%)‡	Genotype	*n*§	Mean differences or means, ng/mL (95% CI)‖	LD block†	Diplotype group (position)†	MAF (%)‡	Genotype	*n*§	Mean differences or means, ng/mL (95% CI)‖
rs903889	Upstream	1	1 (1)	22.0	GG	22	−398 (−754 to −42)	Outside	1 (1)	9.4	GT, GG	97	−131 (−318 to 55)
					GT	131	−67 (−240 to 106)						
					TT	*246*	*4,572 (4,470-4,674)*				*TT*	*471*	*4,108 (4,031-4,185)*
rs924140	Upstream	1	—	46.8	TT	82	581 (352-810)	1	—	38.9	CC	95	−560 (−761 to −360)
					CT	210	375 (190-560)				CT	251	−276 (−427 to −125)
					CC	*107*	*4,210 (4,059-4,361)*				*TT*	*222*	*4,304 (4,194-4,414)*
rs2854744	Upstream	1	1 (2)	46.3	AA	79	592 (364-819)	1	1 (2)	42.2	CC	107	−534 (−728 to −339)
					AC	204	353 (170-536)				AC	258	−258 (−412 to −105)
					CC	*107*	*4,221 (4,073-4,370)*				*AA*	*195*	*4,305 (4,190-4,421)*
rs2854746	Exon 1	1	1 (3)	41.7	CC	65	631 (398-864)	1	1 (3)	32.1	CC	52	897 (656-1,138)
					CG	203	379 (206-551)				CG	259	458 (318-598)
					GG	*130*	*4,226 (4,091-4,360)*				*GG*	*256*	*3,798 (3,699-3,897)*
rs2471551	Intron 1	1	1 (4)	19.4	CC	17	−345 (−745 to 55)	1	1 (4)	20.3	CC	27	−465 (−796 to −133)
					CG	119	−119 (−296 to 57)				CG	173	−258 (−412 to −105)
					GG	*260*	*4,583 (4,484-4,682)*				*GG*	*364*	*4,193 (4,106-4,280)*
rs9282734	Exon 2	—	—	0.4	AC	3	698 (−233 to 1,629)	—	1 (5)	3.3	AC, CC	36	24 (−263 to 311)
					AA	*393*	*4,520 (4,439-4,601)*				*AA*	*532*	*4,079 (4,007-4,151)*
rs2453837	Intron 3	—	—	0	*GG*	*397*	*4,524 (4,444-4,605)*	—	—	0	*GG*	*568*	*4,083 (4,012-4,154)*
rs6953668	Intron 3	—	—	0.6	AG	5	394 (−330 to 1,119)	1	1 (6)	5.0	AG, AA	55	−124 (−362 to 115)
					GG	*394*	*4,522 (4,441-4,603)*				*GG*	*513*	*4,094 (4,020-4,169)*
rs3110697	Intron 3	1	1 (5)	41.4	AA	62	−425 (−670 to −181)	Outside	1 (7)	36.1	AA	78	−551 (−764 to −338)
					AG	203	−188 (−366 to −10)				AG	251	−365 (−513 to −217)
					GG	*131*	*4,693 (4,554-4,832)*				*GG*	*238*	*4,325 (4,219-4,431)*
rs2453840	Intron 4	2	1 (6)	18.3	AA	12	−535 (−1,008 to −62)	2	2 (1)	9.2	AC, AA	98	209 (24-394)
					AC	121	−45 (−221 to 131)						
					CC	*263*	*4,560 (4,461-4,658)*				*CC*	*470*	*4,045 (3,969-4,122)*
rs2453839	Intron 4	2	—	20.1	CC	14	−325 (−760 to 111)	2	2 (2)	40.9	CC	93	−167 (−380 to 45)
					CT	131	−114 (−285 to 57)				CT	279	−28 (−186 to 129)
					TT	*250*	*4,574 (4,474-4,675)*				*TT*	*194*	*4,130 (4,009-4,251)*
rs6670	Exon 5; 3′ untranslated	Outside	2 (1)	21.2	AA	12	74 (−402 to 551)	3	3 (1)	12.6	AT, AA	133	−44 (−210 to 121)
					AT	145	79 (−90 to 249)						
					TT	*240*	*4,494 (4,390-4,598)*				*TT*	*434*	*4,100 (4,020-4,180)*
rs13223993	Downstream	Outside	2 (2)	21.4	AA	19	39 (−343 to 421)	3	3 (2)	43.6	AA	116	−17 (−216 to 182)
					AG	131	42 (−131 to 216)				AG	264	−147 (−308 to 14)
					GG	*244*	*4,522 (4,419-4,624)*				*GG*	*184*	*4,163 (4,039-4,286)*
rs2270628	Downstream	3	2 (3)	21.0	TT	15	−316 (−743 to 112)	4	3 (3)	36.2	TT	83	84 (−131 to 300)
					CT	137	−15 (−186 to 156)				CT	247	−64 (−218 to 90)
					CC	*247*	*4,544 (4,441-4,646)*				*CC*	*236*	*4,101 (3,991-4,211)*
rs12671457	Downstream	3	2 (4)	16.5	AC, CC	118	−119 (−296 to 59)	4	3 (4)	5.1	AC, CC	55	−218 (−456 to 21)
					AA	*267*	*4,558 (4,460-4,657)*				*AA*	*510*	*4,106 (4,032-4,181)*

*Represent tag SNPs except for 3 SNPs that were selected *a priori*: rs2854744 (-202 A/C), rs2854746 (Ala^{32}Gly), and rs9282734 (His^{158}Pro).

† LD blocks consist of SNPs with MAF ≥ 5% in strong LD (95% of pairwise SNP comparisons with one-sided 95% CIs for D' within 0.70-0.98). Group refers to SNP combination for diplotype estimation; position refers to SNP order in each group. SNPs were excluded from diplotype estimation based on the Tagger algorithm ($r^2 \geq 0.8$) or MAF < 5% for tag SNPs or MAF < 3% for SNPs selected *a priori*.

‡ Based on total n = 402 (Caucasians) and total n = 582 (African Americans).

§ Caucasians: n = 401, excluded 1 missing plasma IGFBP-3. African Americans: n = 574, excluded 8 missing plasma IGFBP-3.

‖ Mean plasma IGFBP-3 levels for reference genotypes estimated from linear regression intercepts are italicized in the last row for each SNP. Mean differences in plasma IGFBP-3 levels for heterozygous and homozygous variants relative to reference genotypes were estimated from β values in linear regression models.

To estimate diplotype associations with plasma IGF-I or IGFBP-3 levels, we used separate race-specific models for each diplotype group, with the most common diplotype as the reference category. Diplotypes assigned to ≤5 women were combined into one ''rare diplotype'' category. We used an empirical-Bayes method of information-weighted averaging to enhance the validity and precision of regression estimates (45). Specifically, we assumed a prior mean of 0, because we had no prior information to group diplotypes according to the anticipated direction of associations with plasma levels. We specified a prior variance corresponding to ±1 SD of the mean plasma levels $(2 * SD / 3.92)^2$ of IGF-I (prior variances: African Americans 1,419 and Caucasians 901) and IGFBP-3 (prior variances: African Americans 186,819 and Caucasians 174,161). This method shrinks regression estimates toward the prior mean such that imprecise estimates based on smaller numbers of observations are shifted further toward the prior mean than more precise estimates. We applied the shrinkage estimator for each diplotype and report posterior medians (50th percentile of the posterior probability distribution) and 95% posterior limits (2.5th and 97.5th percentiles of the posterior probability distribution). Regression estimates and 95% CIs estimated by linear regression are available in Supplementary Tables S7 to S10.

Results

Participant Characteristics. Race-specific mean plasma IGF-I and IGFBP-3 levels and other characteristics of the study population are displayed in Table 1. African Americans were less likely than Caucasians to have a

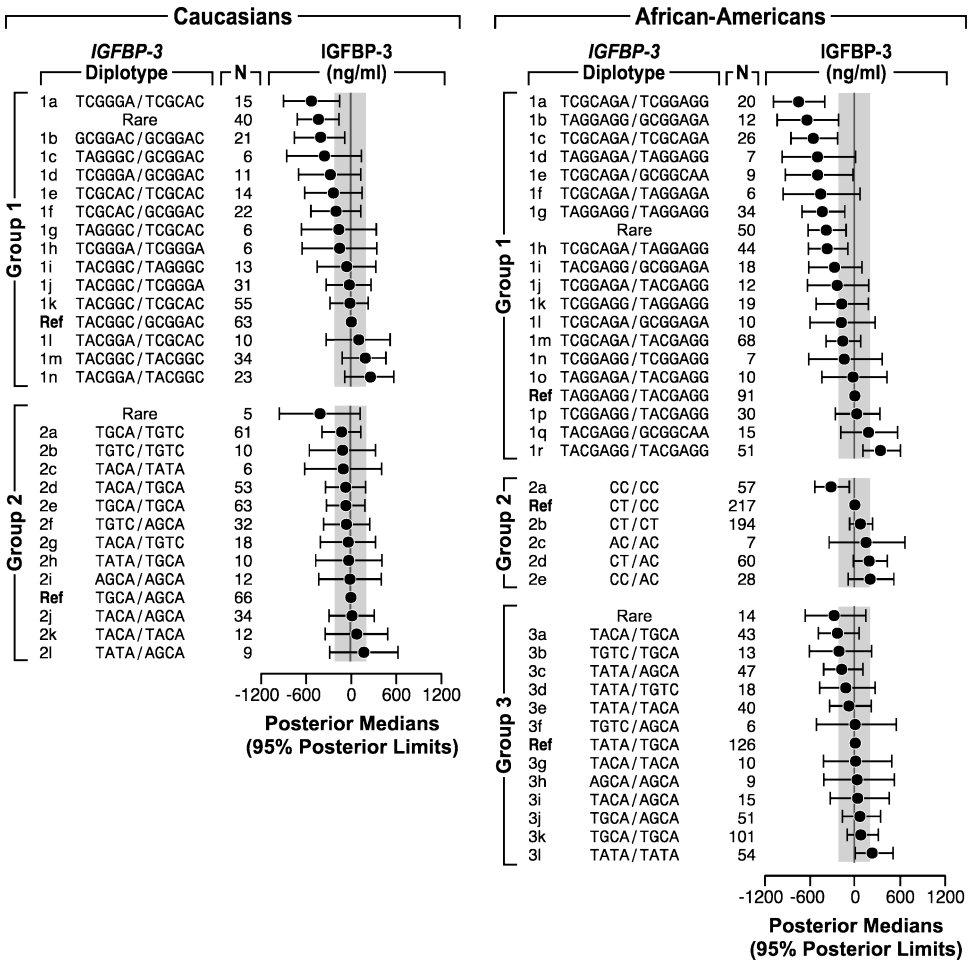

Figure 1. Estimated differences from mean plasma IGFBP-3 levels associated with *IGFBP-3* diplotypes for premenopausal Caucasian and African American women based on race-specific linear regression models of each *IGFBP-3* diplotype group on plasma IGFBP-3 levels. Posterior medians (95% posterior limits) were estimated using empirical-Bayes method of information-weighted averaging by assuming a prior mean of 0 and a prior variance (Caucasians 174,161 and African Americans 186,819) corresponding to 1 SD of mean plasma IGFBP-3 $(2 * SD / 3.92)^2$. For Caucasian models, group 1 includes $n = 370$ and group 2 includes $n = 391$. For African American models, group 1 includes $n = 539$, group 2 includes $n = 563$, and group 3 includes $n = 547$. *Shaded area,* ±200 ng/mL, which is ~5% of the mean plasma IGFBP-3 levels (Caucasians group 1: 4,661 ng/mL and group 2: 4,587 ng/mL; African Americans group 1: 4,327 ng/mL, group 2: 4,060 ng/mL, and group 3: 4,106 ng/mL) for each reference diplotype.

Table 3. Estimated mean differences in plasma IGF-I levels associated with *IGF-I* SNP variants relative to estimated mean IGF-I levels for reference genotypes (italicized) based on linear regression models for premenopausal Caucasian and African American women

IGF-I SNP*	*IGF-I* location	Caucasians						African Americans					
		LD block †	Diplotype group (position) †	MAF (%) ‡	Genotype	*n* §	Mean differences or means, ng/mL (95% CI)‖	LD block †	Diplotype group (position) †	MAF (%) ‡	Genotype	*n* §	Mean differences or means, ng/mL (95% CI)‖
rs35767	Upstream	Outside	1 (1)	16.0	AG, AA	116	5 (−7 to 18)	1	1 (1)	40.8	AA	99	−1 (−19 to 17)
											AG	264	−2 (−16 to 11)
					GG	*280*	*171 (164-177)*				*GG*	*205*	*166 (156-176)*
rs5742612	Upstream	—	—	3.9	AG, GG	30	21 (−1 to 42)	—	—	3.3	AG, GG	35	5 (−20 to 31)
					AA	*368*	*171 (165-177)*				*AA*	*536*	*164 (158-171)*
rs5742614	Intron 1	—	—	1.8	CG	14	18 (−14 to 49)	1	1 (2)	11.5	CG, CC	124	8 (−6 to 23)
					GG	*385*	*172 (166-178)*				*GG*	*444*	*163 (156-170)*
rs3729846	Exon 2	—	—	0				—	—	0.8	CT	9	33 (−16 to 82)
					CC	*399*	*172 (166-178)*				*CC*	*557*	*164 (158-170)*
rs12821878	Intron 2	1	2 (1)	22.7	AA	21	0 (−26 to 27)	Outside	1 (3)	5.8	AG, AA	64	−15 (−34 to 5)
					AG	139	5 (−8 to 17)						
					GG	*237*	*171 (163-178)*				*GG*	*505*	*167 (160-173)*
rs10860869	Intron 2	1	—	29.1	TT	31	−5 (−28 to 17)	Outside	2 (1)	32.3	TT	59	11 (−10 to 32)
					AT	170	6 (−6 to 18)				AT	249	12 (−1 to 25)
					AA	*196*	*170 (162-179)*				*AA*	*262*	*158 (149-167)*
rs1019731	Intron 2	1	2 (2)	12.3	AC, AA	93	10 (−4 to 23)	—	—	3.4	AC	39	−10 (−34 to 14)
					CC	*304*	*170 (164-177)*				*CC*	*528*	*165 (159-172)*
rs7956547	Intron 2	1	2 (3)	27.3	CC	29	−5 (−28 to 18)	2	2 (2)	26.2	CC	43	8 (−16 to 32)
					CT	160	2 (−10 to 14)				CT	210	10 (−3 to 23)
					TT	*209*	*172 (164-180)*				*TT*	*314*	*160 (152-168)*
rs5742626	Intron 2	—	—	0.5	CT	4	−3 (−61 to 54)	2	2 (3)	7.2	CT, CC	78	−8 (−26 to 9)
					TT	*394*	*172 (166-178)*				*TT*	*491*	*166 (159-173)*
rs17880975	Intron 2	—	—	0.1	AG	1	−78 (−193 to 37)	—	—	2.3	AG	26	8 (−21 to 38)
					GG	*395*	*173 (167-178)*				*GG*	*540*	*165 (158-171)*
rs2033178	Intron 2	2	3 (1)	6.2	AG, AA	47	5 (−13 to 22)	—	—	4.7	AG, AA	50	−24 (−46 to −3)
					GG	*349*	*172 (166-178)*				*GG*	*513*	*167 (161-174)*
rs17884646	Intron 2	—	—	0	*TT*	*397*	*172 (166-178)*	—	—	0	*TT*	*570*	*165 (159-171)*
rs5742657	Intron 2	—	—	2.2	AG, GG	17	18 (−11 to 46)	2	—	12.1	AG, GG	128	3 (−11 to 18)
					AA	*383*	*171 (166-177)*				*AA*	*440*	*164 (157-170)*
rs5742663	Intron 2	—	—	0.1	GT	1	12 (−103 to 127)	2	—	11.6	GT, GG	123	8 (−6 to 23)
					TT	*396*	*172 (167-178)*				*TT*	*444*	*163 (156-170)*
rs11829586	Intron 2	—	—	2.3	AG	18	16 (−12 to 43)	2	—	11.5	AG, AA	121	7 (−8 to 22)
					GG	*380*	*172 (166-178)*				*GG*	*445*	*163 (156-170)*
rs4764884	Intron 2	2	—	25.7	TT	25	−16 (−40 to 8)	2	2 (4)	23.7	TT	34	10 (−16 to 36)
					CT	153	0 (−12 to 12)				CT	200	5 (−8 to 18)
					CC	*218*	*173 (165-181)*				*CC*	*333*	*162 (154-170)*
rs5742683	Intron 2	—	—	0.1	AG	1	12 (−102 to 127)	2	—	10.6	AG, GG	113	5 (−11 to 20)
					AA	*396*	*172 (166-178)*				*AA*	*454*	*164 (157-171)*
rs17884626	Exon 3	—	—	0				—	—	1.0	CT	11	13 (−31 to 57)
					CC	*399*	*172 (167-178)*				*CC*	*559*	*164 (158-171)*
rs5009837	Intron 3	2	—	30.3	TT	36	−6 (−27 to 15)	Outside	2 (5)	41.3	TT	105	8 (−9 to 26)
					CT	167	6 (−6 to 18)				CT	260	10 (−3 to 24)
					CC	*192*	*170 (162-178)*				*CC*	*203*	*158 (148-168)*
rs17727841	Intron 3	2	3 (2)	19.1	CC	15	−4 (−35 to 27)	—	—	4.7	CG, CC	49	23 (1-44)
					CG	123	−4 (−17 to 8)						

(*Continued on the following page*)

Table 3. Estimated mean differences in plasma IGF-I levels associated with *IGF-I* SNP variants relative to estimated mean IGF-I levels for reference genotypes (italicized) based on linear regression models for premenopausal Caucasian and African American women (Cont'd)

IGF-I SNP*	*IGF-I* location	Caucasians						African Americans					
		LD block †	Diplotype group (position) †	MAF (%) ‡	Genotype	*n* §	Mean differences or means, ng/mL (95% CI)‖	LD block †	Diplotype group (position) †	MAF (%) ‡	Genotype	*n* §	Mean differences or means, ng/mL (95% CI)‖
					GG	*261*	*174 (167-181)*				*GG*	*520*	*163 (156-169)*
rs4764883	Intron 3	2	—	30.9	CC	37	−8 (−29 to 12)	3	—	45.8	TT	132	−12 (−29 to 5)
					CT	167	4 (−8 to 17)				CT	248	−11 (−25 to 3)
					TT	*186*	*171 (163-179)*				*CC*	*180*	*173 (162-183)*
rs9308315	Intron 3	2	—	28.2	AA	28	−16 (−39 to 7)	3	3 (1)	49.8	TT	150	−12 (−28 to 5)
					AT	166	3 (−9 to 15)				AT	262	−11 (−26 to 4)
					TT	*201*	*172 (164-180)*				*AA*	*152*	*173 (161-185)*
rs978458	Intron 3	2	3 (3)	28.0	TT	28	−13 (−37 to 10)	3	—	37.5	TT	83	9 (−10 to 28)
					CT	164	4 (−8 to 16)				CT	259	7 (−6 to 20)
					CC	*202*	*172 (163-180)*				*CC*	*224*	*160 (151-170)*
rs5742692	Intron 3	—	—	2.1				3	3 (2)	11.7	GG	11	14 (−31 to 58)
					AG	17	18 (−11 to 46)				AG	111	4 (−12 to 19)
					AA	*382*	*171 (166-177)*				*AA*	*444*	*164 (157-170)*
rs11111262	Intron 3	2	—	9.9	AG, AA	76	−4 (−19 to 10)	—	—	2.7	AG, AA	27	40 (11-68)
					GG	*321*	*173 (167-180)*				*GG*	*536*	*163 (157-169)*
rs1520220	Intron 3	2	3 (4)	19.1	GG	13	−33 (−65 to 0)	3	3 (3)	34.7	GG	73	10 (−9 to 29)
					CG	124	7 (−6 to 19)				CG	250	6 (−7 to 19)
					CC	*258*	*171 (164-178)*				*CC*	*245*	*161 (151-170)*
rs3730204	Exon 4; 3′ untranslated	—	—	2.3	CT	18	−24 (−52 to 3)	—	—	0.3	CT	3	6 (−79 to 90)
					TT	*378*	*173 (168-179)*				*TT*	*562*	*165 (159-171)*
rs6214	Exon 4; 3′ untranslated	Outside	3 (5)	39.8	TT	67	17 (0-34)	Outside	3 (4)	45.2	CC	124	−3 (−20 to 14)
					CT	183	−2 (−14 to 11)				CT	263	−4 (−18 to 10)
					CC	*148*	*170 (161-180)*				*TT*	*178*	*167 (156-178)*
rs6219	Exon 4; 3′ untranslated	3	4 (1)	10.0	CT, TT	78	−6 (−20 to 9)	4	4 (1)	9.1	CT, TT	94	16 (−1 to 32)
					CC	*321*	*174 (167-180)*				*CC*	*475*	*162 (155-169)*
rs2946834	Downstream	3	4 (2)	33.0	AA	39	−7 (−28 to 13)	4	4 (2)	49.3	AA	142	16 (0-33)
					AG	181	1 (−11 to 13)				AG	279	2 (−12 to 17)
					GG	*175*	*172 (164-181)*				*GG*	*147*	*159 (147-171)*

*Represent tag SNPs except for 2 SNPs that were selected *a priori*: rs3729846 ($Thr^{52}Thr$) and rs17884626 ($Ala^{115}Thr$).

† LD blocks consist of SNPs with MAF ≥ 5% in strong LD (95% of pairwise SNP comparisons with one-sided 95% CIs for D' within 0.70-0.98). Group refers to SNP combination for diplotype estimation; position refers to SNP order in each group. SNPs were excluded from diplotype estimation based on the Tagger algorithm ($r^2 \geq 0.8$) or MAF < 5% for tag SNPs or MAF < 3% for SNPs selected *a priori*. For African Americans, rs9308315 was the substitute for rs4764883 (Hardy-Weinberg $p < 0.01$).

‡ Based on total $n = 402$ (Caucasians) and total $n = 582$ (African Americans).

§ Caucasians: $n = 401$, excluded 1 missing plasma IGF-I. African Americans: $n = 574$, excluded 8 missing plasma IGF-I.

‖ Mean plasma IGF-I levels for reference genotypes estimated from linear regression intercepts are italicized in the last row for each SNP. Mean differences in plasma IGF-I levels for heterozygous and homozygous variants relative to reference genotypes were estimated from β values in linear regression models.

college or graduate degree (33% versus 87%), to report regular alcohol consumption (42% versus 78%), or to be nulliparous at UFS enrollment (21% versus 59%). African Americans were more likely than Caucasians to be overweight or obese (75% versus 41%) and to report current smoking (30% versus 8%). Few women were currently taking oral contraceptives.

IGFBP-3. For plasma IGFBP-3, we focused on estimated differences in mean levels of at least 200 ng/mL for index genotypes or diplotypes relative to the reference (~±5% of the estimated mean level for the reference genotype or diplotype, which ranged from 3,798 to 4,693 ng/mL). We disregarded imprecise associations with rare SNPs having ≤10 observations with heterozygous and homozygous variants.

***IGFBP-3* SNP Analyses.** The variants for 10 *IGFBP-3* SNPs (rs903889, rs924140, rs2854744, rs2854746, rs2471551, rs3110697, rs2453840, rs2453839, rs2270628, and rs12671457) were associated with differences of ≥200 ng/mL in estimated mean IGFBP-3 levels in at least one racial group when compared with reference genotypes (Table 2). In both races, variants for rs924140, rs2854744, and rs2854746 were associated with differences in plasma IGFBP-3 levels of ~500 to 900 ng/mL for homozygous variants and ~300 to 500 ng/mL for heterozygotes relative to the estimated mean levels for reference genotypes. Pairwise r^2 values for all three SNPs were at least 0.8 among Caucasians, indicating strong LD in our study population. Among African Americans, rs924140 and rs2854744 were also in strong LD ($r^2 = 0.82$); however, neither SNP was in LD with rs2854746 (r^2 = 0.30-0.34). In both races, rs3110697 was also in moderate LD with rs924140 and rs2854744 (r^2 = 0.55-0.65), and its variants were inversely associated with plasma IGFBP-3 (~430 ng/mL lower in Caucasians and ~550 ng/mL lower in African Americans with the AA genotype, with smaller differences estimated for the AG genotype relative to those with the GG genotype). Plasma IGFBP-3 was also inversely associated with rs2471551 variants among African Americans (~460 ng/mL lower for the CC genotype and ~260 ng/mL lower for the CG genotype relative to the GG genotype).

Inverse associations between homozygous variants for five SNPs (rs903889, rs2471551, rs2453840, rs2453839, and rs2270628) and plasma IGFBP-3 were noted among Caucasians, but estimates were relatively imprecise because they were based on <25 observations. Two of these SNPS, rs2453840 and rs2453839, were in strong LD ($r^2 = 0.87$) among Caucasians. Among African Americans, combined homozygous and heterozygous variants for rs2453840 and rs12671457 were positively and inversely associated with plasma IGFBP-3 respectively, with differences of ~200 ng/mL relative to reference genotypes.

SNP associations with plasma IGFBP-3 showed little evidence of discordance by race based on estimated African American versus Caucasian differences in these associations (Supplementary Table S5). Three possible exceptions were rs2854746 (CC versus GG genotype difference between races, 266; 95% CI, −70 to 602), rs2453840 (AA/AC versus CC genotype difference between races, 298; 95% CI, 45-551), and rs2270628 (TT versus CC genotype difference between races, 400; 95% CI, −88 to 888). However, estimated differences in mean differences between races, especially for homozygous variants, were fairly imprecise.

***IGFBP-3* Diplotype Analyses.** Among Caucasians, three LD blocks accounted for 10 of 12 *IGFBP-3* SNPs included in diplotype analyses, with 2 SNPs outside LD blocks. After excluding two redundant *IGFBP-3* SNPs ($r^2 \geq 0.8$), we created two diplotype groups, as described previously (Table 2), and completed diplotype estimation in each group for 93% to 98% of Caucasians. Overall, we estimated 48 unique diplotypes, including 20 classified as rare based on assignment to ≤5 women.

Among African Americans, four LD blocks accounted for 11 of 14 *IGFBP-3* SNPs included in diplotype analyses, with 3 SNPs outside LD blocks. After excluding one redundant *IGFBP-3* SNP ($r^2 \geq 0.8$), we created three diplotype groups (Table 2) and completed diplotype estimation in each group for 94% to 99% of African Americans. Overall, we estimated 71 unique diplotypes, which included 33 rare diplotypes.

Six Caucasian *IGFBP-3* group 1 diplotypes (1a-1f) were associated with decreases of ≥200 ng/mL in estimated mean IGFBP-3 levels relative to the reference diplotype (Fig. 1). All six diplotypes (1a-1f) included the GG genotype for rs2854746 (third diplotype position), and five diplotypes (1a, 1b, and 1d-1f) included the CC genotype for rs2854744 (second diplotype position). The only two Caucasian *IGFBP-3* group 1 diplotypes (1m and 1n) with the CC genotype for rs2854746 were positively associated with plasma IGFBP-3.

Ten African American *IGFBP-3* group 1 diplotypes (1a-1j) were inversely associated with plasma IGFBP-3 relative to the reference diplotype, including eight (1a-1h) with the GG genotype for rs2854746 (third diplotype position), three (1a, 1c, and 1e) with the CC genotype for rs2854744 (second position), and the only one (1c) with the CC genotype for rs2471551 (fourth diplotype position; Fig. 1). In addition, the only African American *IGFBP-3* group 1 diplotype (1r) with the CC genotype for rs2854746 was positively associated with plasma IGFBP-3. One African American *IGFBP-3* group 2 diplotype (2a) and two group 3 diplotypes (3a and 3b) were associated with lower mean IGFBP-3 levels relative to the reference diplotype, whereas one group 3 diplotype (3l) was associated with a higher mean level.

IGF-I. For plasma IGF-I, we focused on estimated differences in mean levels of at least 16 to 18 ng/mL for index genotypes or diplotypes relative to the reference (~±10% of the estimated mean level for the reference genotype or diplotype, which ranged from 158 to 176 ng/mL). We did not consider imprecise associations with rare SNPs having ≤10 observations with heterozygous and homozygous variants.

***IGF-I* SNP Analyses.** Among Caucasians, homozygous variants for two common *IGF-I* SNPs (rs1520220 and rs6214) and variants for five rare (MAF < 5%) *IGF-I* SNPs (rs5742612, rs5742614, rs5742657, rs5742692, and rs3730204) were associated with ≥10% differences in estimated mean IGF-I levels relative to reference genotypes (Table 3). However, with the exception of the positive association with the rs6214 TT genotype, estimates were relatively imprecise due to ≤30 observations with variants. None of the *IGF-I* SNPs noted above

were in LD ($r^2 < 0.4$), except for strong LD between rs5742657 and rs5742692 ($r^2 = 0.94$).

Among African Americans, combined homozygous and heterozygous variants for three rare (MAF < 5%) *IGF-I* SNPs (rs2033178, rs17727841, and rs11111262) were associated with differences of ~15% to 25% relative to estimated mean IGF-I levels for the reference genotypes, although the association with the rs11111262 variants was based on <30 observations (Table 3). In addition, variants for two more common *IGF-I* SNPs (rs6219 and rs2946834) were associated with 10% increases in estimated mean IGF-I levels relative to reference genotypes. Of the five *IGF-I* SNPs noted above, only rs17727841 and rs11111262 were in moderate LD ($r^2 = 0.56$).

Based on models that included interaction terms with race, estimated mean differences in IGF-I levels

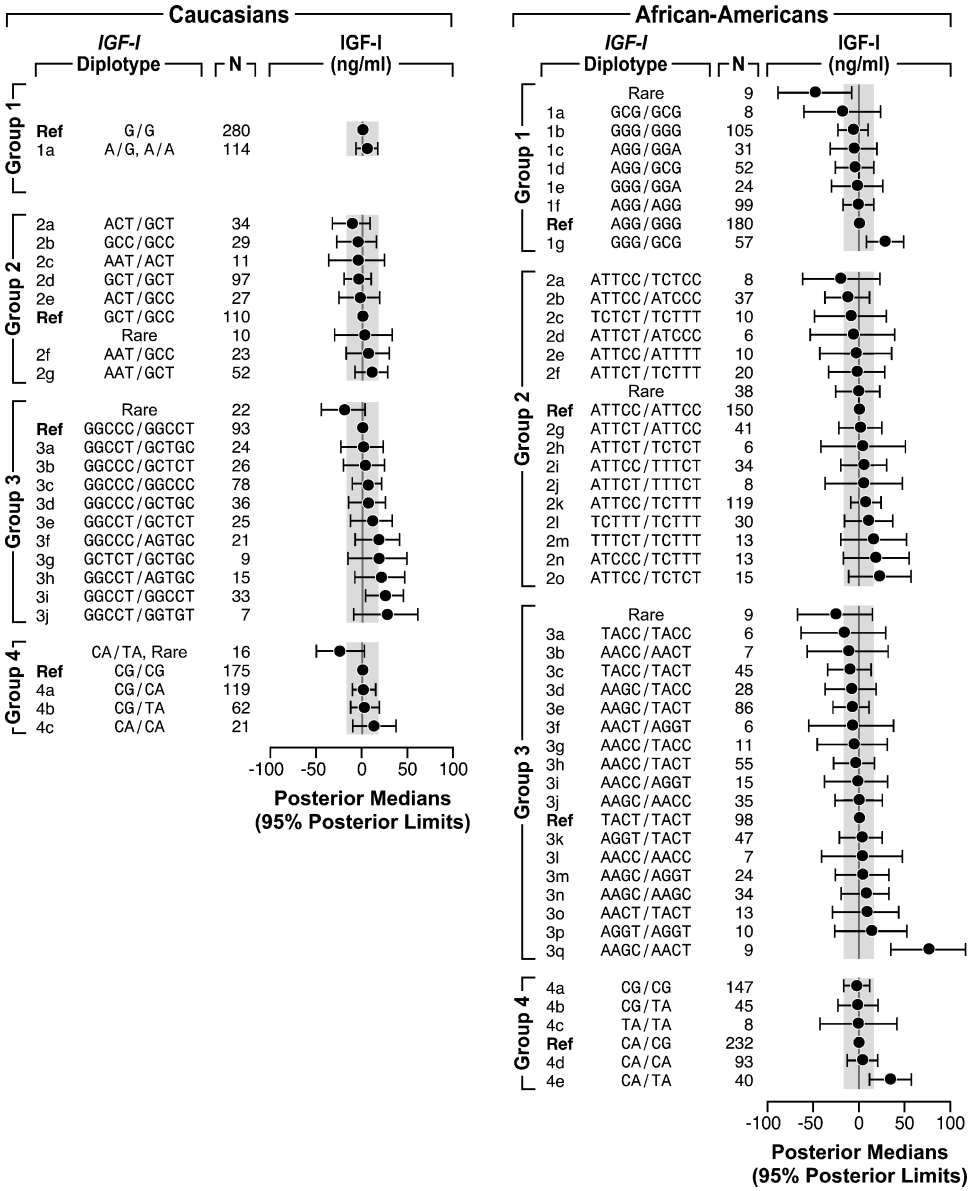

Figure 2. Estimated differences from mean plasma IGF-I levels associated with *IGF-I* diplotypes for premenopausal Caucasian and African American women based on race-specific linear regression models of each *IGF-I* diplotype group on plasma IGF-I levels. Posterior medians (95% posterior limits) were estimated using empirical-Bayes method of information-weighted averaging by assuming a prior mean of 0 and a prior variance (Caucasians 901 and African Americans 1,419) corresponding to 1 SD of mean plasma IGF-I $(2 * SD / 3.92)^2$. For Caucasian models, group 1 includes $n = 394$, group 2 includes $n = 393$, group 3 includes $n = 389$, and group 4 includes $n = 393$. For African American models, group 1 includes $n = 565$, group 2 includes $n = 558$, group 3 includes $n = 545$, and group 4 includes $n = 565$. *Shaded area,* ±10% of the mean plasma IGF-I levels (Caucasians group 1: 171 ng/mL, group 2: 173 ng/mL, group 3: 165 ng/mL, and group 4: 172 ng/mL; African Americans group 1: 166 ng/mL, group 2: 162 ng/mL, group 3: 166 ng/mL, and group 4: 162 ng/mL) for each reference diplotype.

for 13 *IGF-I* SNP variants varied by at least 16 ng/mL in African Americans versus Caucasians (Supplementary Table S6). These included 4 SNPs within or near exon 4, which was where plasma IGF-I associations were predominantly noted: rs11111262 (AG/AA versus GG genotype difference between races, 44; 95% CI, 13-75), rs1520220 (GG versus CC genotype difference between races, 43; 95% CI, 1-85), rs6219 (CT/TT versus CC genotype difference between races, 21; 95% CI, −1 to 44), and rs2946834 (AA versus GG genotype difference between races, 24; 95% CI, −4 to 52). However, estimated differences in mean differences between races were relatively imprecise.

***IGF-I* Diplotype Analyses.** Among Caucasians, three LD blocks accounted for 15 of 17 *IGF-I* SNPs included in diplotype analyses, with 2 SNPs outside LD blocks. After excluding six redundant *IGF-I* SNPs ($r^2 \geq 0.8$), we created four diplotype groups (Table 3) and completed diplotype estimation in each group for 97% to 99% of Caucasians. Overall, we estimated 40 unique diplotypes, including 13 classified as rare based on assignment to ≤5 women.

Among African Americans, four LD blocks accounted for 16 of 20 *IGF-I* SNPs included in diplotype analyses, with 4 SNPs outside LD blocks. After excluding six redundant *IGF-I* SNPs ($r^2 \geq 0.8$), we created four diplotype groups (Table 3) and completed diplotype estimation in each group for 95% to 99% of African Americans. Overall, we estimated 73 unique diplotypes, which included 25 rare diplotypes.

Five Caucasian *IGF-I* group 3 diplotypes (3f-3j) were associated with ≥10% increases in estimated mean IGF-I levels relative to the reference diplotype, although two estimates (3g and 3j) were relatively imprecise due to few observations with either diplotype (Fig. 2). One African American *IGF-I* group 1 diplotype (1g) and one group 4 diplotype (4e) were associated with increases of 15% to 20% in mean IGF-I levels relative to the reference diplotypes. In addition, one African American *IGF-I* group 1 diplotype (1a), four group 2 diplotypes (2a and 2m-2o), and two group 3 diplotypes (3a and 3q) were associated with ≥10% differences in mean IGF-I levels compared with the reference diplotypes, although estimates were fairly imprecise due to small numbers of observations. *IGF-I* diplotype associations with plasma IGF-I did not parallel SNP associations unlike associations between *IGFBP-3* diplotypes and plasma IGFBP-3.

Discussion

Evidence of a causal association was strongest for the nonsynonymous *IGFBP-3* SNP, rs2854746, with plasma IGFBP-3 levels. In both races, the rs2854746 CC genotype was associated with higher mean IGFBP-3 levels than were estimated for the GG genotype, whereas mean levels for the CG genotype were intermediate. In addition, *IGFBP-3* diplotypes with the rs2854746 GG genotype had consistently lower mean IGFBP-3 levels than those estimated for reference diplotypes with the CG genotype in both races, whereas *IGFBP-3* diplotypes with the CC genotype had higher mean IGFBP-3 levels.

Biological evidence supports a causal relation of rs2854746 with plasma IGFBP-3, because this SNP results in an amino acid change from alanine to glycine, and protein sequence analysis suggests that the amino acid coded by rs2854746 is within the region of IGFBP-3 responsible for IGF-I binding (46). The Multiethnic Cohort Study, a large study of Caucasian women from the Breast and Prostate Cancer Cohort Consortium, and a small Caucasian study reported associations between rs2854746 and plasma IGFBP-3 that were consistent with our findings (24, 35, 47). Therefore, individual rs2854746 associations and correspondence with *IGFBP-3* diplotype findings in both races, in addition to biological evidence, support a causal association between rs2854746 and plasma IGFBP-3.

Other genetic studies of circulating IGFBP-3 have not evaluated rs2854746, but several studies have examined the *IGFBP-3* promoter SNP, rs2854744 (-202 A/C), predominantly among Caucasians. Several studies reported higher mean IGFBP-3 levels among individuals with the AA genotype compared with the CC genotype and intermediate levels among those with the AC genotype (24, 27-35). We also noted increases in mean IGFBP-3 levels for rs2854744 AA versus CC genotypes in both races, although plasma IGFBP-3 associations with diplotypes that included rs2854744 variants were not as consistent as those with diplotypes that included rs2854746 variants, especially among African Americans. Consistent with the Multiethnic Cohort Study (24), we noted strong LD between rs2854744 and rs2854746 ($r^2 = 0.82$) in Caucasians that may partly explain associations between rs2854744 variants and plasma IGFBP-3, although these two SNPs were not in LD among African Americans ($r^2 = 0.34$). Deal et al. (28) reported that promoter activity was increased *in vitro* in association with the rs2854744 A allele, which suggests that rs2854744 may influence circulating IGFBP-3 levels independent of its association with rs2854746. Similar to our study, four studies reported decreased IGFBP-3 levels in association with rs3110697 variants relative to the reference genotype (24, 33-35), and the Multiethnic Cohort Study also reported that rs3110697 was not in strong LD with rs2854744 or rs2854746 among both races (24).

We also reported consistent inverse associations between *IGFBP-3* rs2471551 variants and plasma IGFBP-3 among both races. This SNP has potential functional relevance as it is located near a splice site (<20 bp from the 5′ side of exon 2). Canzian et al. (27) and Diorio et al. (34) reported that the rs2471551 CC genotype was inversely associated with circulating IGFBP-3 relative to the GG genotype in Caucasian women.

Many epidemiologic studies of *IGF-I* have focused on the dinucleotide CA repeat polymorphism (position −969) located in the promoter ~1 kb upstream of the transcription site; however, associations between CA repeat polymorphisms and circulating IGF-I levels have been inconsistent (23, 30, 31, 48-55). Methodologic differences in the categorization of repeat genotypes and the potential for substantial misclassification during genotyping make it difficult to compare results across studies (56). Although we did not evaluate this repeat polymorphism, we evaluated three *IGF-I* SNPs within 5 kb of the 5′ and 3′ ends of the gene and found that two were associated with ≥10% differences in mean IGF-I levels (rs5742612 among Caucasians and rs2946834 among African Americans). However, Diorio et al. (34)

reported no association between rs5742612 variants and plasma IGF-I within their study of Caucasian women. There have been no reports of LD between the *IGF-I* repeat polymorphism and any of the SNPs in our study, with the exception of rs5742612 in a Chinese population (57). Due to rs5742612 MAF differences for Chinese versus African Americans or Caucasians, it is unlikely that rs5742612 would be in LD with the repeat polymorphism in our study population.

We estimated higher mean IGF-I levels in association with the rs6214 TT versus CC genotypes among Caucasians. In contrast, Al-Zahrani et al. (29) and Canzian et al. (27) reported no association between rs6214 and circulating IGF-I within predominantly Caucasian study populations. Consistent with the Multiethnic Cohort Study (24), we found no association between rs35767 and plasma IGF-I in either racial group. However, Canzian et al. (27) and Patel et al. (35) noted associations between rs35767 and circulating IGF-I in large studies of Caucasian women. We also noted an inverse association between plasma IGF-I and the rs1520220 GG versus CC genotypes among Caucasians and a positive association with rs2946834 AA versus GG genotypes among African Americans, although estimated differences in race-specific associations may not be meaningful given their imprecision. Al-Zahrani et al. (29) and Patel et al. (35) reported higher mean IGF-I levels in association with rs1520220 and rs2946834 variants relative to reference genotypes in Caucasian women, although Al-Zahrani et al. (29) reported that only the association with rs1520220 variants remained after adjustment for rs2946834. However, rs1520220 and rs2946834 variants were not associated with plasma IGF-I in the Multiethnic Cohort Study (24), and Diorio et al. (34) reported no association with rs1520220 variants. Comparison of race-specific *IGF-I* SNP and diplotype associations with plasma IGF-I suggests that an untyped functional polymorphism may lie near or within the untranslated region of exon 4, but we could not identify this polymorphism from evaluation of the literature.

A strength of this study is that participants were randomly selected from health plan membership roles, with response rates of ~80% for both races. However, selection bias could exist if eligible women excluded from our analysis differed from the women who were included with respect to their plasma IGF-I or IGFBP-3 levels or genotypes. In particular, 161 (14%) eligible women lacked DNA for genotyping primarily because of no available blood samples, which included a slightly greater proportion of missing African Americans (16%) than Caucasians (12%).

The use of diplotype analysis strengthened our study as it provided support for detecting which SNPs may be causally associated with circulating protein levels, and it assisted with identifying regions where untyped SNPs that influence circulating protein levels may reside. Diplotype associations were unlikely to be biased by the exclusion of women who had diplotypes estimated with low certainty (posterior probability < 90%), because only 1% to 7% of women were excluded from analyses of each diplotype group based on this criterion. Rare diplotypes were assigned with lower certainty, as the PHASE software assumes that frequently observed haplotypes with less ambiguity are more probable. We combined rare diplotypes assigned to ≤5 women into a single category, but we did not interpret associations with these categories due to their heterogeneity.

Despite restricting our study to Caucasians and African Americans and stratifying analyses by race, population stratification within each racial group is a potential limitation of our study. Population stratification is more likely to bias results within African Americans due to their inherently greater admixture than Caucasians. However, the degree of bias depends on the number of ethnicities and the range of genotype frequencies within the racial group in addition to the true magnitude of genotype association with the outcome (58-60). Concordance between races for *IGFBP-3* SNP findings suggests that population stratification was less likely to bias these results, although population stratification within races may still be present. However, population stratification within races may have more strongly influenced the *IGF-I* SNP findings, as there were notable differences between races.

Our use of information-weighted averaging intentionally biased estimates of associations with race-specific diplotypes toward the null because we assumed a null value for the prior mean. However, this approach increased the precision of estimates, particularly for diplotypes assigned to small numbers of women. Despite the increase in bias with estimating posterior medians, a reduction in the overall mean square error based on a greater decrease in variance of estimates has been shown with simulation studies and an occupational cohort study (61). Although we did not interpret our results based on hypothesis tests, this approach also reduces the likelihood of type I error with multiple comparisons (62-64).

The parent study obtained only one measurement of plasma IGF-I and IGFBP-3 from study participants; however, the Nurses' Health Study reported high correlations (>0.8) for plasma IGF-I and IGFBP-3 measurements repeated in premenopausal women over time (65). Age and sex are strong predictors of circulating IGF-I and IGFBP-3; however, these factors were unlikely to influence our results because our study population was restricted to premenopausal women within a 15-year age range (35-49 years), and adjusting for age did not affect results (data not shown).

A major strength of our study was the large number of African American participants, because previous research has mostly focused on relations between *IGF-I* and *IGFBP-3* SNPs and their circulating protein levels in Caucasians. Because African Americans have more genetic heterogeneity than Caucasians, the frequency of etiologically relevant SNPs may differ and may at least partly explain racial disparities in the burden of cancer and cardiovascular disease. Therefore, assessing *IGF-I* and *IGFBP-3* SNPs that predict circulating IGF-I and IGFBP-3 levels will improve our understanding of the biological role of IGF-I and IGFBP-3 in the etiology of common diseases.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be

hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The epidemiologic study was managed by Glenn Heartwell, and the clinical coordinator was Dr. Joel Schechtman. Dr. Mary Watson prepared DNA samples for genotyping. Dr. Jason Luo performed genotyping within the Mammalian Genotyping Core of The University of North Carolina at Chapel Hill. Dr. Sue Edelstein prepared the figures. Drs. Abee Boyles and Stephanie London reviewed an earlier version of the article.

References

1. Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: biological actions. Endocr Rev 1995;16:3–34.
2. Hwa V, Oh Y, Rosenfeld RG. The insulin-like growth factor-binding protein (IGFBP) superfamily. Endocr Rev 1999;20:761–87.
3. Juul A, Main K, Blum WF, Lindholm J, Ranke MB, Skakkebaek NE. The ratio between serum levels of insulin-like growth factor (IGF)-I and the IGF binding proteins (IGFBP-1, 2 and 3) decreases with age in healthy adults and is increased in acromegalic patients. Clin Endocrinol (Oxf) 1994;41:85–93.
4. Gill ZP, Perks CM, Newcomb PV, Holly JM. Insulin-like growth factor-binding protein (IGFBP-3) predisposes breast cancer cells to programmed cell death in a non-IGF-dependent manner. J Biol Chem 1997;272:25602–7.
5. Schedlich LJ, Graham LD. Role of insulin-like growth factor binding protein-3 in breast cancer cell growth. Microsc Res Tech 2002;59:12–22.
6. Giovannucci E. Insulin-like growth factor-I and binding protein-3 and risk of cancer. Horm Res 1999;51:34–41.
7. Renehan AG, Zwahlen M, Minder C, O'Dwyer ST, Shalet SM, Egger M. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. Lancet 2004;363:1346–53.
8. Sandhu MS. Insulin-like growth factor-I and risk of type 2 diabetes and coronary heart disease: molecular epidemiology. Endocr Dev 2005;9:44–54.
9. Harrela M, Koistinen H, Kaprio J, et al. Genetic and environmental components of interindividual variation in circulating levels of IGF-I, IGF-II, IGFBP-1, and IGFBP-3. J Clin Invest 1996;98:2612–5.
10. Pantsulaia I, Trofimov S, Kobyliansky E, Livshits G. Genetic regulation of the variation of circulating insulin-like growth factors and leptin in human pedigrees. Metabolism 2005;54:975–81.
11. Pantsulaia I, Trofimov S, Kobyliansky E, Livshits G. Pedigree-based quantitative genetic analysis of interindividual variation in circulating levels of IGFBP-3. J Bone Miner Metab 2002;20:156–63.
12. LeRoith D. Insulin-like growth factors. N Engl J Med 1997;336:633–40.
13. Kaklamani VG, Linos A, Kaklamani E, Markaki I, Mantzoros C. Age, sex, and smoking are predictors of circulating insulin-like growth factor 1 and insulin-like growth factor-binding protein 3. J Clin Oncol 1999;17:813–7.
14. Holmes MD, Pollak MN, Hankinson SE. Lifestyle correlates of plasma insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations. Cancer Epidemiol Biomarkers Prev 2002;11:862–7.
15. Lukanova A, Toniolo P, Akhmedkhanov A, et al. A cross-sectional study of IGF-I determinants in women. Eur J Cancer Prev 2001;10:443–52.
16. Probst-Hensch NM, Wang H, Goh VH, Seow A, Lee HP, Yu MC. Determinants of circulating insulin-like growth factor I and insulin-like growth factor binding protein 3 concentrations in a cohort of Singapore men and women. Cancer Epidemiol Biomarkers Prev 2003;12:739–46.
17. DeLellis K, Rinaldi S, Kaaks RJ, Kolonel LN, Henderson B, Le Marchand L. Dietary and lifestyle correlates of plasma insulin-like growth factor-I (IGF-I) and IGF binding protein-3 (IGFBP-3): the multiethnic cohort. Cancer Epidemiol Biomarkers Prev 2004;13:1444–51.
18. Harris TB, Kiel D, Roubenoff R, et al. Association of insulin-like growth factor-I with body composition, weight history, and past health behaviors in the very old: the Framingham Heart Study. J Am Geriatr Soc 1997;45:133–9.
19. Goodman-Gruen D, Barrett-Connor E. Epidemiology of insulin-like growth factor-I in elderly men and women. The Rancho Bernardo Study. Am J Epidemiol 1997;145:970–6.
20. Johansson H, Baglietto L, Guerrieri-Gonzaga A, et al. Factors associated with circulating levels of insulin-like growth factor-I and insulin-like growth factor binding protein-3 in 740 women at risk for breast cancer. Breast Cancer Res Treat 2004;88:63–73.
21. Morimoto LM, Newcomb PA, White E, Bigler J, Potter JD. Variation in plasma insulin-like growth factor-1 and insulin-like growth factor binding protein-3: personal and lifestyle factors (United States). Cancer Causes Control 2005;16:917–27.
22. Chang S, Wu X, Yu H, Spitz MR. Plasma concentrations of insulin-like growth factors among healthy adult men and postmenopausal women: associations with body composition, lifestyle, and reproductive factors. Cancer Epidemiol Biomarkers Prev 2002;11:758–66.
23. Jernstrom H, Chu W, Vesprini D, et al. Genetic factors related to racial variation in plasma levels of insulin-like growth factor-1: implications for premenopausal breast cancer risk. Mol Genet Metab 2001;72:144–54.
24. Cheng I, Henderson KD, Haiman CA, et al. Genetic determinants of circulating IGF-I, IGFBP-1 and IGFBP-3 levels in a multiethnic population. J Clin Endocrinol Metab 2007;92:3660–6.
25. Day Baird D, Dunson DB, Hill MC, Cousins D, Schectman JM. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. Am J Obstet Gynecol 2003;188:100–7.
26. Baird DD, Dunson DB, Hill MC, Cousins D, Schectman JM. Association of physical activity with development of uterine leiomyoma. Am J Epidemiol 2007;165:157–63.
27. Canzian F, McKay JD, Cleveland RJ, et al. Polymorphisms of genes coding for insulin-like growth factor 1 and its major binding proteins, circulating levels of IGF-I and IGFBP-3 and breast cancer risk: results from the EPIC study. Br J Cancer 2006;94:299–307.
28. Deal C, Ma J, Wilkin F, et al. Novel promoter polymorphism in insulin-like growth factor-binding protein-3: correlation with serum levels and interaction with known regulators. J Clin Endocrinol Metab 2001;86:1274–80.
29. Al-Zahrani A, Sandhu MS, Luben RN, et al. IGF1 and IGFBP3 tagging polymorphisms are associated with circulating levels of IGF1, IGFBP3 and risk of breast cancer. Hum Mol Genet 2006;15:1–10.
30. Jernstrom H, Deal C, Wilkin F, et al. Genetic and nongenetic factors associated with variation of plasma levels of insulin-like growth factor-I and insulin-like growth factor-binding protein-3 in healthy premenopausal women. Cancer Epidemiol Biomarkers Prev 2001;10:377–84.
31. Lai JH, Vesprini D, Zhang W, Yaffe MJ, Pollak M, Narod SA. A polymorphic locus in the promoter region of the IGFBP3 gene is related to mammographic breast density. Cancer Epidemiol Biomarkers Prev 2004;13:573–82.
32. Schernhammer ES, Hankinson SE, Hunter DJ, Blouin MJ, Pollak MN. Polymorphic variation at the -202 locus in IGFBP3: influence on serum levels of insulin-like growth factors, interaction with plasma retinol and vitamin D and breast cancer risk. Int J Cancer 2003;107:60–4.
33. Ren Z, Cai Q, Shu XO, et al. Genetic polymorphisms in the IGFBP3 gene: association with breast cancer risk and blood IGFBP-3 protein levels among Chinese women. Cancer Epidemiol Biomarkers Prev 2004;13:1290–5.
34. Diorio C, Brisson J, Berube S, Pollak M. Genetic polymorphisms involved in insulin-like growth factor (IGF) pathway in relation to mammographic breast density and IGF levels. Cancer Epidemiol Biomarkers Prev 2008;17:880–8.
35. Patel AV, Cheng I, Canzian F, et al. IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3). PLoS ONE 2008;3:e2578.
36. Behrens M, Lange R. A highly reproducible and economically competitive SNP analysis of several well characterized human mutations. Clin Lab 2004;50:305–16.
37. Livak KJ. Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. Genet Anal 1999;14:143–9.
38. McGuigan FE, Ralston SH. Single nucleotide polymorphism detection: allelic discrimination using TaqMan. Psychiatr Genet 2002;12:133–6.
39. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76:887–93.
40. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5.
41. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225–9.
42. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005;37:1217–23.
43. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162–9.
44. Stephens M, Smith NJ, Donnelly P. A new statistical method for

haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
45. Greenland S. Bayesian perspectives for epidemiological research. I. Foundations and basic methods. Int J Epidemiol 2006;35:765–75.
46. Bairoch A, Apweiler R, Wu CH, et al. The Universal Protein Resource (UniProt). Nucleic Acids Res 2005;33:D154–9.
47. Morimoto LM, Newcomb PA, White E, Bigler J, Potter JD. Variation in plasma insulin-like growth factor-1 and insulin-like growth factor binding protein-3: genetic factors. Cancer Epidemiol Biomarkers Prev 2005;14:1394–401.
48. Rosen CJ, Kurland ES, Vereault D, et al. Association between serum insulin growth factor-I (IGF-I) and a simple sequence repeat in IGF-I gene: implications for genetic studies of bone mineral density. J Clin Endocrinol Metab 1998;83:2286–90.
49. Jernstrom H, Sandberg T, Bageman E, Borg A, Olsson H. Insulin-like growth factor-1 (IGF1) genotype predicts breast volume after pregnancy and hormonal contraception and is associated with circulating IGF-1 levels: implications for risk of early-onset breast cancer in young women from hereditary breast cancer families. Br J Cancer 2005;92:857–66.
50. Rietveld I, Janssen JA, Hofman A, Pols HA, van Duijn CM, Lamberts SW. A polymorphism in the IGF-I gene influences the age-related decline in circulating total IGF-I levels. Eur J Endocrinol 2003;148:171–5.
51. Missmer SA, Haiman CA, Hunter DJ, et al. A sequence repeat in the insulin-like growth factor-1 gene and risk of breast cancer. Int J Cancer 2002;100:332–6.
52. Vaessen N, Heutink P, Janssen JA, et al. A polymorphism in the gene for IGF-I: functional properties and risk for type 2 diabetes and myocardial infarction. Diabetes 2001;50:637–42.
53. DeLellis K, Ingles S, Kolonel L, et al. IGF1 genotype, mean plasma level and breast cancer risk in the Hawaii/Los Angeles multiethnic cohort. Br J Cancer 2003;88:277–82.
54. Wen W, Gao YT, Shu XO, et al. Insulin-like growth factor-I gene polymorphism and breast cancer risk in Chinese women. Int J Cancer 2005;113:307–11.
55. Giovannucci E, Haiman CA, Platz EA, Hankinson SE, Pollak MN, Hunter DJ. Dinucleotide repeat in the insulin-like growth factor-I gene is not related to risk of colorectal adenoma. Cancer Epidemiol Biomarkers Prev 2002;11:1509–10.
56. Tran N, Bharaj BS, Diamandis EP, Smith M, Li BD, Yu H. Short tandem repeat polymorphism and cancer risk: influence of laboratory analysis on epidemiologic findings. Cancer Epidemiol Biomarkers Prev 2004;13:2133–40.
57. Wong HL, Delellis K, Probst-Hensch N, et al. A new single nucleotide polymorphism in the insulin-like growth factor I regulatory region associates with colorectal cancer risk in Singapore Chinese. Cancer Epidemiol Biomarkers Prev 2005;14:144–51.
58. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. Genet Epidemiol 2004;27:14–20.
59. Ardlie KG, Lunetta KL, Seielstad M. Testing for population subdivision and association in four case-control studies. Am J Hum Genet 2002;71:304–11.
60. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. J Natl Cancer Inst 2000;92:1151–8.
61. Greenland S, Poole C. Empirical-Bayes and semi-Bayes approaches to occupational and environmental hazard surveillance. Arch Environ Health 1994;49:9–16.
62. Poole C. Multiple comparisons? No problem! Epidemiology 1991;2:241–3.
63. Greenland S, Robins JM. Empirical-Bayes adjustments for multiple comparisons are sometimes useful. Epidemiology 1991;2:244–51.
64. Greenland S. Multiple comparisons and association selection in general epidemiology. Int J Epidemiol 2008;37:430–4.
65. Missmer SA, Spiegelman D, Bertone-Johnson ER, Barbieri RL, Pollak MN, Hankinson SE. Reproducibility of plasma steroid hormones, prolactin, and insulin-like growth factor levels among premenopausal women over a 2- to 3-year period. Cancer Epidemiol Biomarkers Prev 2006;15:972–8.

IGF-I and *IGFBP-3* Polymorphisms in Relation to Circulating Levels among African American and Caucasian Women

Aimee A. D'Aloisio, Jane C. Schroeder, Kari E. North, et al.

Cancer Epidemiol Biomarkers Prev 2009;18:954-966.