Title
Sex Hormones Genotypes and Phenotypes and Determinants of Sex Hormone Binding Globulin in relation to Type 2 Diabetes Risk

Author
GOTO, ATSUSHI

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Sex Hormones Genotypes and Phenotypes and Determinants of Sex Hormone Binding Globulin in relation to Type 2 Diabetes Risk

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Epidemiology

by

Atsushi Goto

2012

ABSTRACT OF THE DISSERTATION

Sex Hormones Genotypes and Phenotypes and

Determinants of Sex Hormone Binding Globulin in relation to

Type 2 Diabetes Risk

by

Atsushi Goto

Doctor of Philosophy in Epidemiology

University of California, Los Angeles, 2012

Professor Simin Liu, Chair

Type 2 diabetes (T2D) is a common and complex metabolic disease that is associated with serious complications and premature mortality. Several prospective studies and clinical experiments support the role of sex hormones in the development of T2D.

This dissertation examines the roles of sex hormones and sex hormone-binding globulin (SHBG) in the development of T2D as well as the determinants of SHBG levels. Chapter 1 introduces general background information regarding sex hormones and SHBG in relation to the pathogenesis of T2D as well as determinants of plasma SHBG levels. Chapter 2 explores whether the levels of sex hormones and SHBG may play a role in the inverse relation between coffee intake and the risk of T2D in the Women's Health Study. Chapter 3 examines the potential determinants of circulating levels of SHBG in a pooled analysis of studies included in the Women's Health Initiative Randomized Trial and Observational Study. Chapter 4 investigates roles of sex hormone pathway genes in the development of T2D and examines

consistency across African-American women, Hispanic women, and European-American based on data from the Women's Health Initiative SNP Health Association Resource (WHI-SHARe) and the WHI Genomics and Randomized Trials Network (WHI-GARNET).

Our findings indicated that the observed inverse associations of caffeinated-coffee and caffeine intakes with the risk of T2D were substantially attenuated after taking into account of SHBG levels. Findings from a pooled analysis of studies included in the Women's Health Initiative suggested that age, usage of exogenous estrogen, BMI, and lifestyle factors may be associated with plasma levels of SHBG. In a genetic study examining associations between sex hormone pathway genes and T2D risk, possible associations between genes involved in sex hormones synthesis and metabolism, estrogen receptors, and androgen receptor and the risk of T2D were seen in African Americans, and possible relations between genes involved in the metabolism of androgen and progesterone receptor and the risk of T2D risk were seen in Hispanics. The directions of effect estimates were generally consistent across ethnicity.

These findings support the notion that sex hormones and SHBG may play roles in the pathogenesis of T2D. Also, the findings suggest that age, usage of exogenous estrogen, physical activity, regular coffee intake, and adiposity may play roles in the regulation of SHBG metabolism. Further studies in replicative and experimental settings are warranted to confirm the findings.

The dissertation of Atsushi Goto is approved.

Jørn Olsen

Frank Sorvillo

Qing Zhou

Simin Liu, Committee Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

VITA

1998-2004	Medical Doctor School of Medicine Yokohama City University, Kanagawa, Japan
2003 Summer	Clinical Clerkship School of Medicine Harvard University, Cambridge, Massachusetts
2004-2005	Residency Training Program Fujisawa City Hospital, Kanagawa, Japan
2005 Fall	Internship Kanazawa Public Health Center of Yokohama, Kanagawa, Japan
2005-2006	Residency Training Program Yokohama City University Hospital, Kanagawa, Japan
2006-2008	Endocrinology and Diabetes Training Program International Medical Center of Japan, Tokyo, Japan
2008-2010	Master of Public Health in Epidemiology School of Public Health University of California, Los Angeles, California
2009	University Fellowship in Epidemiology
2010	Wayne Soohoo Memorial Scholarship.
2010-2012	Endocrinology and Diabetes Training Program National Center for Global Health and Medicine, Tokyo, Japan

PUBLICATIONS AND PRESENTATIONS

Goto A, Takahashi Y, Kishimoto M, Tanabe A, Kajio H, Hasuo K, and Noda M. Hypopituitarism caused by bilateral internal carotid artery aneurysms with a carotid-cavernous fistula. Intern Med. 47(8): 815-816, 2008.

Goto A, Takahashi Y, Kishimoto M, Nakajima Y, Nakanishi K, Kajio H, and Noda M. A Case of Fulminant Type 1 Diabetes Associated With Significant Elevation of Mumps Titers. Endocr. J. 55(3): 561-564, 2008.

Goto A, Takaichi M, Kishimoto M, Takahashi Y, Kajio H, Shimbo T, and Noda M. Assessment of Combined Parameters with C-Peptide Levels as Predictors for Future Needs of Insulin Treatment in Japanese Type 2 Diabetic Subjects." The American Diabetes Association (ADA) 68th Scientific Sessions, San Francisco, June 6-10, 2008. Poster presentation

Goto A, Takahashi Y, Kishimoto M, Minowada S, Aibe H, Hasuo K, Kajio H, Noda M. Primary aldosteronism associated with severe rhabdomyolysis due to profound hypokalemia. Intern Med. 48(4):219-23, 2009.

Goto A, Takaichi M, Kishimoto M, Takahashi Y, Kajio H, Shimbo T, and Noda M. Body mass index, fasting plasma glucose levels, and C-peptide levels as predictors of the future insulin use in Japanese type 2 diabetic patients. Endocr J. 57(3):237-44, 2010.

Goto A, Song Y, Chen BH, Manson JE, Buring JE, Liu S. Coffee and Caffeine Consumption in Relation to Plasma Sex Hormone-Binding Globulin and the Risk of Type 2 Diabetes. The American Heart Association (AHA) 50th Cardiovascular Disease Epidemiology and Prevention Annual Conference, San Francisco, March 5th, 2010. Oral presentation.

Goto A, Takaichi M, Kishimoto M, Takahashi Y, Kajio H, Shimbo T, and Noda M. Duration of Hospitalization and Duration of Type 2 Diabetes as Predictors of Change in Glycemic Control before and after Hospitalizations among Japanese Type 2 Diabetic Patients. Society for Epidemiologic Research the 43rd Annual Meeting, Seattle, June 23-26, 2010. Poster Presentation.

Goto A, Song Y, Chen BH, Manson JE, Buring JE, Liu S. Coffee and caffeine consumption in relation to sex hormone-binding globulin and risk of type 2 diabetes in postmenopausal women. Diabetes. 60(1):269-75, 2011.

Goto A, Liu S. Is caffeine protective against Type 2 diabetes? Diabetes Manage. 1(4), 351–353, 2011 (Editorial).

Chapter 1 Background

1.1 Introduction

Type 2 diabetes (T2D) is a common and complex metabolic disease, associated with serious complications and premature mortality. An accelerating increase in the prevalence of T2D is expected worldwide and the increase in developing countries is estimated to be substantial.[1-3] The fundamental basis of T2D pathogenesis involves the impairment of insulin action and secretion. Risk factors for T2D include obesity, diet (high fat and high glycemic index/load), physical inactivity, a family history of T2D, age, race/ethnicity, smoking, a history of gestational diabetes, and sex.[4-6] Importantly, increasing physical activity and restricting the total energy intake has been shown to substantially decrease the incidence of T2D.[7] However, the prevalence of T2D is increasing, even in developed countries. This finding suggests the need for translating these findings into the community and may also indicate an incomplete understanding of the pathogenesis of T2D. To date, little is known about the exact mechanisms of obesity in the development of T2D. Epidemiologic evidence suggests that obesity is a stronger risk factor for diabetes among women than men.[8] Moreover, several prospective studies and clinical experiments support the role of sex hormones in the development of T2D.[9-11]

Similar to obesity, the reported inverse association between coffee consumption and the risk of T2D is stronger among women, suggesting the existence of sex differences.[12] Importantly, laboratory experiments have shown that caffeine and other major components of coffee (both cafestol and kahweol) alter the expressions and activities of liver enzymes.[13] Furthermore, previous studies have associated coffee consumption with circulating levels of sex hormones and

sex hormone-binding globulin (SHBG).[14,15] Coffee consumption has been suggested to influence the metabolism of sex hormones.[16]

Although increasing evidence suggests important roles of sex hormones and SHBG in glucose metabolism, the evidence in human is insufficient to support this notion. In this dissertation, we conducted three studies to investigate the roles of sex hormones and SHBG in the development of T2D as well as determinants of SHBG levels to better understand the pathophysiology linking sex hormones and SHBG with T2D. We first examined whether the levels of sex hormones and SHBG may play a role in the inverse relation between coffee intake and T2D risk in the Women's Health Study. Then, we examined potential determinants of the circulating levels of SHBG in a pooled analysis of studies included in the Women's Health Initiative Randomized Trial and Observational Study. Finally, we investigated the roles of sex hormone pathway genes in the development of T2D based on data from the Women's Health Initiative SNP Health Association Resource (SHARe). This chapter will provide general background information on sex hormones, SHBG, determinants of SHBG levels, and their relation to the risk of T2D.

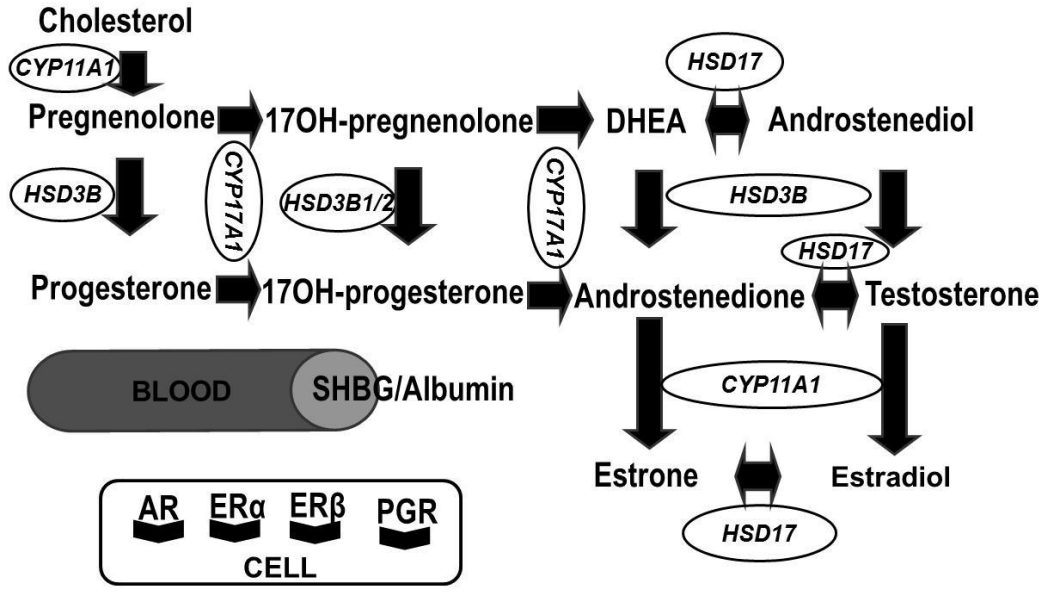
1.2 Background

Sex Hormones and Sex hormone-binding globulin

The primary sex hormones are testosterone in men, and estradiol and progesterone in the women. Circulating testosterone and estradiol levels are regulated by SHBG, which restricts the passage of these hormones across the plasma membrane of target cells. Classically, free estradiol (unbound form) and free testosterone pass across the cell membrane and bind to intracellular sex

hormone receptors; the hormone-receptor complex then binds to DNA in the nucleus, facilitating the transcription of various genes. In addition to functioning as a regulator of the free concentration of steroid hormones, SHBG was recently shown to play a central role in permitting certain steroid hormones to act without entering the cell. [17] These actions result in rate changes in the synthesis of proteins coded by the genes being transcribed, resulting in change in the concentrations of these proteins in specific target cells or the enhancement of cell responses to the hormone. Similar to estradiol and testosterone, 80 % of the circulating progesterone is bound to albumin, 18% is bound to corticosteroid-binding globulin, and 2 % is not bound to a carrier protein (i.e., free progesterone). Transcription occurs after progesterone binds to progesterone receptors, resulting in the formation of messenger RNA. [18]

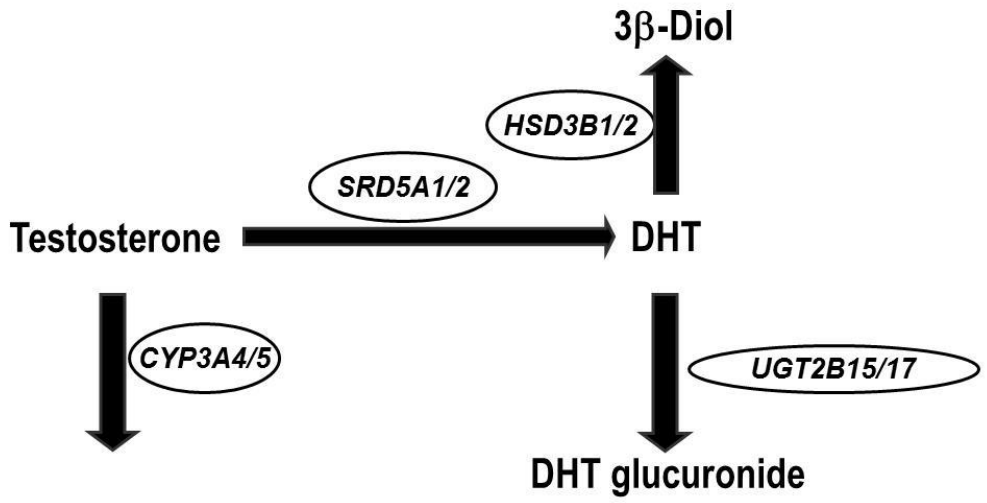
Figure 1. Synthesis, transport, and receptors of sex hormones



Androgen: Synthesis and Metabolism and Mechanisms of Actions

The primary androgens are testosterone and its active metabolite, dihydrotestosterone (DHT). Androgens are synthesized and secreted in the testes, ovary, and adrenals. The synthesis and secretion of testosterone is regulated by cholesterol desmolase (CYP11A1), which allows the conversion of cholesterol into pregnenolone. The subsequent conversion of pregnenolone to testosterone involves multiple pathways in which cytochrome P-450 17A1 (CYP17A1), 17β-hydroxysteroid dehydrogenase 1/2/3(HSD17B1/2/3), and 3β-hydroxysteroid dehydrogenase 1 and 2 (HSD3B1/2) play important roles (Figure 1). After testosterone has been taken up by target tissues, testosterone can be further reduced by 5α-reductase enzymes (SRDA5A1/2) to yield a more active androgen metabolite, dihydrotestosterone (DHT). Testosterone and DHT bind to the same intracellular receptor (androgen receptor, AR) (Figure 2). This hormone-receptor complex then binds to DNA in the nucleus of a target cell, altering the rate of formation of multiple protein transcripts. Testosterone and androstenedione can also be converted into estradiol through a process termed aromatization. Aromatization is performed by a cytochrome P450 aromatase enzyme (CYP19), which is expressed in the granulose cells of the ovary, the Leydig cells of the testes, and many other tissues including the placenta, brain, pituitary, liver, and adipose tissue. Indeed, many of the effects of circulating testosterone result from its conversion into either 5α-DHT or estradiol within target tissues. Testosterone is oxidized by cytochrome P450 3A4 (CYP3A4) and cytochrome P450 3A5 (CYP3A5) into hydroxyl-testosterone. 3-Beta-hydroxysteroid dehydrogenases (HSD3B1/2) catalyze the oxidation of DHT into 3β-diol, and the uridine diphosphate glycosyltransferase 2 family, member B15/17 (UGT2B15/17) catalyzes the glucuronidation of DHT into DHT glucuronide (Figure 2).

Figure 2. Metabolism of Androgen



Relation between Androgens and the Risk of Type 2 Diabetes

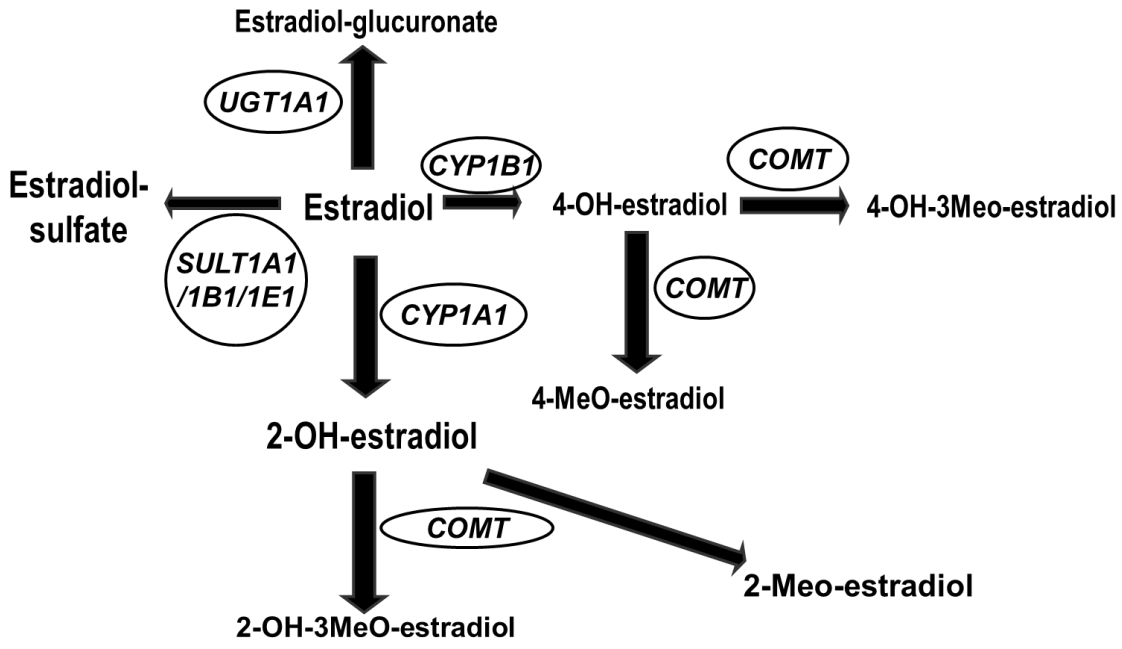
A recent systematic review has reported that in cross-sectional studies, women with T2D had higher levels of testosterone, compared with control subjects, whereas men with T2D had lower levels of plasma total testosterone. [9] A prospective nested case-control study also associated higher testosterone levels with T2D risk among postmenopausal women.[10] In experiments performed using female rats treated with testosterone after an oophorectomy, the rats became insulin resistant,[11] and androgen receptor knockout mice exhibited insulin resistance and impaired glucose tolerance,[19] providing further evidence of a direct association between testosterone and T2D risk. Furthermore, testosterone therapy in postmenopausal women has been shown to increase visceral adiposity,[20] a strong risk factor for diabetes, while anti-androgen therapy decreases visceral adiposity in obese women with polycystic ovarian syndrome.[21]

Estrogen: Synthesis and Metabolism and Mechanisms of Actions

The naturally occurring estrogens are estradiol, estrone, and estriol. These estrogens are secreted primarily by the granulose cells of the ovarian follicles, the corpus luteum, and the placenta in premenopausal women. Importantly, ovarian estrogen synthesis is negligible in postmenopausal women, in whom estrogens are synthesized in peripheral tissues, such as subcutaneous fat. The biosynthesis of these estrogens depends on the cytochrome P450 aromatase enzyme (CYP19), which converts testosterone to estradiol and androstenedione to estrone. This reaction occurs in the ovaries, adipose, liver, muscle, and brain. Estrogen is synthesized by the conversion of testosterone and androstenedione by cytochrome P450, family 11, subfamily A, polypeptide 1 (CYP11A1). Thus, the synthesis of estrogen is limited by the circulating levels of testosterone and androstenedione, which are synthesized using 17α-hydroxylase (CYP17). Among the three major forms of estrogen, estradiol exerts the most potent estrogenic effect. There are two principal types of nuclear estrogen receptors to which estrogen binds: estrogen receptor α (ERα) and estrogen receptor β (ERβ). After binding estrogen, these receptors form homodimers and bind to DNA, altering its transcription. Most tissues contain only one type of receptor, but overlapping does occurs, with some tissues containing both ERα and ERβ. ERα is found primarily in the uterus, kidneys, liver, and heart, whereas ERβ is found primarily in the ovaries, prostate, lungs,gastrointestinal tract, hemopoietic system, and central nervous system (CNS). These receptors also form heterodimers with ERα binding to ERβ. Estradiol is metabolized into its metabolites by several enzymes. Sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1(SULT1A1); sulfotransferase family 1E, estrogen-preferring, member 1(SULT1E1); and sulfotransferase family 1B, member 1(SULT1B1)

catalyze the sulfate conjugation of estradiol into estradiol-sulfate. In addition, estradiol is conjugated by uridine diphosphate glycosyltransferase 1 family, polypeptide A1 (UGT1A1) into estradiol glucuronide. Estradiol is also catalyzed into hydroxyl-estradiol by cytochrome P450 lAl (CYP1A1) and 1B1 (CYP1B1). Then, catechol-O-methyltransferase (COMT) catalyzes the methylation of hydroxyl-estradiol (Figure 3).

Figure 3. Metabolism of Estrogen



Relation between Estrogens and the Risk of Type 2 Diabetes

Epidemiologic evidence suggests that higher levels of estradiol at baseline are associated with the future development of T2D in both women and men.[9,10] This finding, however, is inconsistent with the results from clinical trials of oral estrogen therapy, in which estrogen treatment lowered the risk of T2D among postmenopausal women.[22-24] Exogenous

estrogen has been suggested to have different physiological effects depending on the dose, duration, and mode of administration, possibly explaining these conflicting results.[10] Moreover, female ERα knockout mice exhibited insulin resistance and impaired glucose tolerance,[25] suggesting that estrogen may have beneficial effects on glucose metabolism. Further studies are needed to elucidate the biological mechanisms through which estrogens affect glucose utilization.

SHBG: Synthesis and Metabolism and Mechanisms of Action

SHBG is synthesized primarily in the liver and binds to androgens and estrogens, thereby regulating the biologically active fraction of sex hormones.[26] Thus, the circulating SHBG level is a major determinant of the metabolic clearance of sex hormones, modulating the access of such sex hormones to their target tissues. Moreover, the plasma membranes of various kinds of cells have been recently shown to be capable of binding specifically and with a high affinity to SHBG, mediating sex hormone signaling at the cell membrane through SHBG receptors.[17,27]

Relation between SHBG and the Risk of Type 2 Diabetes

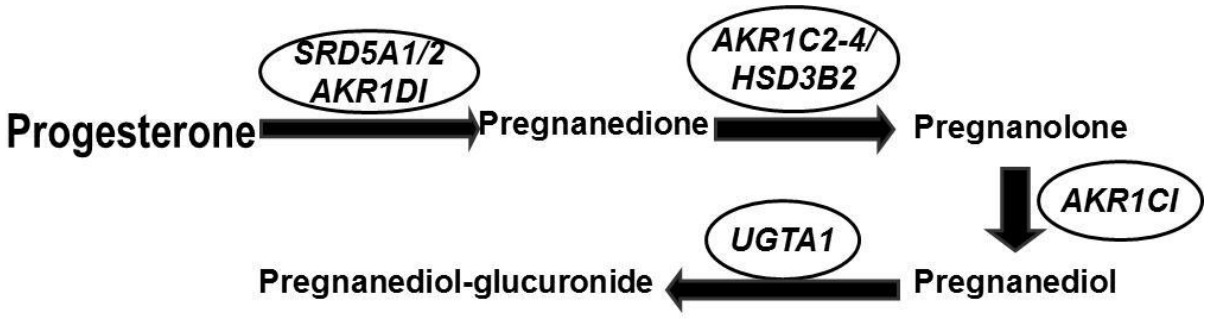
Emerging data indicate that higher SHBG levels are associated with a lower risk of T2D in both women and men.[9] Importantly, recent molecular epidemiologic studies have shown that genetically determined levels of SHBG were inversely associated with the risk of T2D based on the results of a Mendelian randomization (MR) analysis,[28,29] and the direction and magnitude of the association of the genetically determined levels of SHBG and the T2D risk were consistent with the multivariable-adjusted association observed in earlier studies. MR analysis is based on the principle that the random assortment of genotypes during meiosis is unrelated to confounding factors.[30] Therefore, these molecular epidemiologic works support a protective role of the plasma

SHBG levels in T2D. Further, a recent meta-analysis found that higher SHBG levels were associated with a lower risk of metabolic syndrome in both men and women[31] and one recent prospective study showed that the lower SHBG levels, but not the testosterone levels, were associated with the incident of metabolic syndrome in men.[6]

Progesterone: Synthesis and Metabolism and Mechanisms of Actions

Progesterone is a sex-steroid hormone secreted by the corpus luteum, the placenta, and the follicle. Small amounts also enter the circulation from tissues that secrete steroid hormones, such as the testes and adrenal cortex. Similar to testosterone and estradiol, progesterone binds to an intracellular receptor (progesterone receptor: PGR). The progesterone receptor is bound to a heat shock protein in the absence of progesterone, and the binding of progesterone releases the heat shock protein, exposing the DNA-binding domain of the receptor. Transcription occurs after progesterone binds to progesterone receptors, resulting in formation of messenger RNA.[18] Progesterone is converted to pregnanedione by 5α-reductase (SRDA5A1/2) and 5β-reductase(AKR1D1). Pregnanedione is then catalyzed into pregnanolone by aldo-keto reductase family 1, member C2-4 (AKR1C2-4) and HSD3B2, and aldo-keto reductase family 1, member C1 (AKR1C1) converts pregnanolone into pregnanediol. Then, pregnanediol is conjugated to glucuronic acid by UGT1A1 and excreted in the urine. [18] The main target organs of progesterone are the uterus, the breasts, and the brain. Progesterone plays important roles in the progestational changes that occur in the endometrium and in the cyclic changes that occur in the cervix and vagina. It also has an antiestrogenic effect on myometrial cells. In addition, progesterone decreases the number of estrogen receptors in the endometrium and accelerates estrogen metabolism.

Figure 4. Metabolism of Progesterone



Relation between Progesterone and the Risk of Type 2 Diabetes

Blood progesterone levels rise continuously throughout pregnancy,[32] and the onset of gestational diabetes typically occurs during the second trimester of pregnancy, when the circulating levels of progesterone are relatively high.[33] These observations have led to the hypothesis that increased levels of progesterone might be responsible for the development of gestational diabetes. In a study of 2,081 pregnant women, women who received a prophylactic intramuscular injection of 17α-hydroxyprogesterone caproate had a 2.9-fold increased risk of gestational diabetes, compared with women who did not.[34] However, in a double-blind randomized controlled trial of 1091 women, the weekly administration of 17α-hydroxyprogesterone caproate did not result in a significant difference in the incidence of gestational diabetes.[35] Experimental evidence in rodents supports the roles of progesterone in glucose metabolism. The administration of high-dose progesterone to female diabetes-prone *db/db* mice accelerated the onset of diabetes.[36] Furthermore, mice lacking a progesterone receptor manifested with increased insulin secretion as a result of pancreatic β cell proliferation.[36] These data suggest that progesterone might play roles in glucose metabolism and possibly in the development of T2D.

Relation between Coffee Consumption and the Risk of Type 2 Diabetes

Coffee is among the most widely consumed beverages in the world. Knowledge of the health effects of coffee allows individuals to make informed decisions about coffee consumption. Coffee consumption is also considered to be a potentially modifiable preventive/causative factor that might contribute to disease prevention. For these reasons, coffee consumption has been extensively studied in relation to various diseases, including T2D.[12]

Cumulative evidence from previous prospective studies has documented an association between greater coffee consumption and a lower risk of T2D.[12,37] Several possible explanations have been put forth to explain the potential protective effect of coffee consumption on T2D, including the effects coffee components such as magnesium, potassium, chlorogenic acid, and caffeine on insulin sensitivity and pancreatic β cell function by coffee components.[12] Though caffeine consumption can acutely lower insulin sensitivity,[38] the chronic intake of caffeine may contribute to the protective effect of coffee consumption on T2D. To date, however, little is known about the exact mechanisms responsible for the potential protective effect of coffee on the development of type 2 diabetes. Importantly, evidence from a systematic review suggests that sex differences exist in the inverse association between coffee consumption and the risk of T2D.[12] These results suggest that sex hormone signaling may play an important role in the potential protective effect of coffee against T2D.

Determinants of Circulating SHBG levels

Given the clinical importance of SHBG, the determinants of the plasma SHBG levels are of great interest. The estrogen/androgen balance is generally assumed to regulate SHBG

synthesis.[26] However, SHBG synthesis in the liver is considered to be regulated by hormonal factors (sex hormones,[39-41] thyroid hormone,[40,41] prolactin,[41] insulin-like growth factor I,[42] and insulin[41,42]), other environmental factors (adiposity,[41] physical activity level,[43] and dietary factors[41]), age,[44] and genetic factors.[45] To date, however, the literature remains inconclusive as to whether these factors play direct roles in the regulation of SHBG levels in humans.

1.3 Specific Aims

The aims of this dissertation are to examine the roles of sex hormones and SHBG in the development of T2D as well as to identify determinants of the SHBG levels. In the following chapters, we will discuss the results with regard to our specific aims: 1) to explore whether the levels of sex hormones and SHBG may account for the inverse relation between coffee intake and the risk of T2D in the Women's Health Study; 2) to examine potential determinants of the circulating levels of SHBG in a pooled analysis of studies included in the Women's Health Initiative Randomized Trial and Observational Study; and 3) to investigate the associations between sex hormone pathway genes and the risk of T2D using data from the Women's Health Initiative SNP Health Association Resource (SHARe) and the WHI Genomics and Randomized Trials Network (WHI-GARNET).

.

1.5 References

1. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract.* Nov 5 2009.

2. Roglic G, Unwin N. Mortality attributable to diabetes: Estimates for the year 2010. *Diabetes Res Clin Pract.* Nov 13 2009.

3. Thomas MC, Hardoon SL, Papacosta AO, et al. Evidence of an accelerating increase in prevalence of diagnosed Type 2 diabetes in British men, 1978-2005. *Diabet Med.* Aug 2009;26(8):766-772.

4. Standards of medical care in diabetes--2012. *Diabetes care.* Jan 2012;35 Suppl 1:S11-63.

5. Hu FB, Manson JE, Stampfer MJ, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *The New England journal of medicine.* Sep 13 2001;345(11):790-797.

6. Joslin EP, Kahn CR. *Joslin's diabetes mellitus*. 14th ed. Philadelphia, Pa.: Lippincott Williams & Willkins; 2005.

7. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med.* Feb 7 2002;346(6):393-403.

8. Willett WC, Dietz WH, Colditz GA. Guidelines for healthy weight. *N Engl J Med.* Aug 5 1999;341(6):427-434.

9. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA.* Mar 15 2006;295(11):1288-1299.

10. Ding EL, Song Y, Manson JE, Rifai N, Buring JE, Liu S. Plasma sex steroid hormones and risk of developing type 2 diabetes in women: a prospective study. *Diabetologia.* Oct 2007;50(10):2076-2084.

11. Rincon J, Holmang A, Wahlstrom EO, et al. Mechanisms behind insulin resistance in rat skeletal muscle after oophorectomy and additional testosterone treatment. *Diabetes.* May 1996;45(5):615-621.

12. van Dam RM, Hu FB. Coffee consumption and risk of type 2 diabetes: a systematic review. *JAMA.* Jul 6 2005;294(1):97-104.

13. Park BK, Kitteringham NR. Assessment of enzyme induction and enzyme inhibition in humans: toxicological implications. *Xenobiotica.* Nov 1990;20(11):1171-1185.

14. Kotsopoulos J, Eliassen AH, Missmer SA, Hankinson SE, Tworoger SS. Relationship between caffeine intake and plasma sex hormone concentrations in premenopausal and postmenopausal women. *Cancer.* Jun 15 2009;115(12):2765-2774.

15. Nagata C, Kabuto M, Shimizu H. Association of coffee, green tea, and caffeine intakes with serum concentrations of estradiol and sex hormone-binding globulin in premenopausal Japanese women. *Nutrition and cancer.* 1998;30(1):21-24.

16. Lucero J, Harlow BL, Barbieri RL, Sluss P, Cramer DW. Early follicular phase hormone levels in relation to patterns of alcohol, tobacco, and coffee use. *Fertil Steril.* Oct 2001;76(4):723-729.

17. Rosner W, Hryb DJ, Khan MS, Nakhla AM, Romas NA. Sex hormone-binding globulin mediates steroid hormone signal transduction at the plasma membrane. *The Journal of steroid biochemistry and molecular biology.* Apr-Jun 1999;69(1-6):481-485.

18. Ganong's review of medical physiology. New York: McGraw-Hill Medical; 2010.

19. Lin HY, Xu Q, Yeh S, Wang RS, Sparks JD, Chang C. Insulin and leptin resistance with hyperleptinemia in mice lacking androgen receptor. *Diabetes.* Jun 2005;54(6):1717-1725.

20. Lovejoy JC, Bray GA, Greeson CS, et al. Oral anabolic steroid treatment, but not parenteral androgen treatment, decreases abdominal fat in obese, older men. *Int J Obes Relat Metab Disord.* Sep 1995;19(9):614-624.

21. Gambineri A, Pagotto U, Tschop M, et al. Anti-androgen treatment increases circulating ghrelin levels in obese women with polycystic ovary syndrome. *Journal of endocrinological investigation.* Jul 2003;26(7):629-634.

22. Kanaya N, Murray PA, Damron DS. The differential effects of midazolam and diazepam on intracellular Ca2+ transients and contraction in adult rat ventricular myocytes. *Anesth Analg.* Dec 2002;95(6):1637-1644, table of contents.

23. Margolis KL, Bonds DE, Rodabough RJ, et al. Effect of oestrogen plus progestin on the incidence of diabetes in postmenopausal women: results from the Women's Health Initiative Hormone Trial. *Diabetologia.* Jul 2004;47(7):1175-1187.

24. Bonds DE, Lasser N, Qi L, et al. The effect of conjugated equine oestrogen on diabetes incidence: the Women's Health Initiative randomised trial. *Diabetologia.* Mar 2006;49(3):459-468.

25. Ribas V, Nguyen MT, Henstridge DC, et al. Impaired Oxidative Metabolism and Inflammation are Associated with Insulin Resistance in ER{alpha} Deficient Mice. *Am J Physiol Endocrinol Metab.* Nov 17 2009;298(2):E304-E319.

26. Anderson DC. Sex-hormone-binding globulin. *Clinical endocrinology.* Jan 1974;3(1):69-96.

27. Kahn SM, Hryb DJ, Nakhla AM, Romas NA, Rosner W. Sex hormone-binding globulin is synthesized in target cells. *The Journal of endocrinology.* Oct 2002;175(1):113-120.

28. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N Engl J Med.* Sep 17 2009;361(12):1152-1163.

29. Perry JR, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Human molecular genetics.* Feb 1 2010;19(3):535-544.

30. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomisation and causal inference in observational epidemiology. *PLoS medicine.* Aug 26 2008;5(8):e177.

31. Joslin EP, Marble A. *Joslin's Diabetes mellitus.* 11th ed. Philadelphia,: Lea & Febiger; 1971.

32. Jackson JM, Defor TA, Crain AL, et al. Self-reported diabetes is a valid outcome in pragmatic clinical trials and observational studies. *Journal of clinical epidemiology.* May 5 2012.

33. Joslin EP. *Joslin's Diabetes mellitus.* 12th ed. Philadelphia: Lea & Febiger; 1985.

34. Rebarber A, Istwan NB, Russo-Stieglitz K, et al. Increased incidence of gestational diabetes in women receiving prophylactic 17alpha-hydroxyprogesterone caproate for prevention of recurrent preterm delivery. *Diabetes care.* Sep 2007;30(9):2277-2280.

35. Margolis KL, Lihong Q, Brzyski R, et al. Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements. *Clin Trials.* 2008;5(3):240-247.

36. Picard F, Wanatabe M, Schoonjans K, Lydon J, O'Malley BW, Auwerx J. Progesterone receptor knockout mice have an improved glucose homeostasis secondary to beta -cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America.* Nov 26 2002;99(24):15644-15648.

37. Huxley R, Lee CM, Barzi F, et al. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: A systematic review with meta-analysis. *Arch Intern Med.* Dec 14 2009;169(22):2053-2063.

38. Greer F, Hudson R, Ross R, Graham T. Caffeine ingestion decreases glucose disposal during a hyperinsulinemic-euglycemic clamp in sedentary humans. *Diabetes.* Oct 2001;50(10):2349-2354.

39. Lee IR, Dawson SA, Wetherall JD, Hahnel R. Sex hormone-binding globulin secretion by human hepatocarcinoma cells is increased by both estrogens and androgens. *The Journal of clinical endocrinology and metabolism.* Apr 1987;64(4):825-831.

40. Rosner W, Aden DP, Khan MS. Hormonal influences on the secretion of steroid-binding proteins by a human hepatoma-derived cell line. *The Journal of clinical endocrinology and metabolism.* Oct 1984;59(4):806-808.

41. Plymate SR, Matej LA, Jones RE, Friedl KE. Inhibition of sex hormone-binding globulin production in the human hepatoma (Hep G2) cell line by insulin and prolactin. *The Journal of clinical endocrinology and metabolism.* Sep 1988;67(3):460-464.

42. Crave JC, Lejeune H, Brebant C, Baret C, Pugeat M. Differential effects of insulin and insulin-like growth factor I on the production of plasma steroid-binding globulins by human hepatoblastoma-derived (Hep G2) cells. *The Journal of clinical endocrinology and metabolism.* Apr 1995;80(4):1283-1289.

43. Hawkins VN, Foster-Schubert K, Chubak J, et al. Effect of exercise on serum sex hormones in men: a 12-month randomized clinical trial. *Med Sci Sports Exerc.* Feb 2008;40(2):223-233.

44. Muller M, den Tonkelaar I, Thijssen JH, Grobbee DE, van der Schouw YT. Endogenous sex hormones in men aged 40-80 years. *European journal of endocrinology / European Federation of Endocrine Societies.* Dec 2003;149(6):583-589.

45. Xita N, Tsatsoulis A. Genetic variants of sex hormone-binding globulin and their biological consequences. *Molecular and cellular endocrinology.* Mar 5 2010;316(1):60-65.

Chapter 2

Coffee and Caffeine Consumption in Relation to Sex Hormone–Binding Globulin and Risk of Type 2 Diabetes in Postmenopausal Women: the Women's Health Study

2.1 Introduction

Previous prospective studies have documented an inverse association between coffee consumption and the risk of type 2 diabetes (T2D)[1,2] especially among women.[2] Coffee intake may improve glucose tolerance via the activation of energy metabolism and the enhancement of insulin sensitivity and β-cell function,[2,3] although much of the molecular mechanism remains unknown. Previous cross-sectional studies have associated coffee intake with the plasma levels of sex hormones or sex hormone-binding globulin (SHBG).[4,5] In addition, a large body of observational and experimental data has suggested the importance of sex hormones in the development of T2D.[6-8] Notably, recent experiments indicate that SHBG not only regulates the biologically active fraction of sex hormones but may bind to its own receptors at the plasma membranes of a variety of cells, directly mediating the intracellular signalling of sex hormones.[9] More recently, prospective studies of men and women incorporating both genetic and phenotypic assessment of SHBG have revealed a strong inverse association between the SHBG levels and T2D risk.[10] However, no studies have comprehensively evaluated the interrelationships of coffee consumption in relation to sex hormones and SHBG with respect to T2D risk. To examine whether and to what extent sex hormones or SHBG may account for the potential protective effect of coffee intake against the development of T2D, we analyzed data from a prospective case-control study of women. In particular, we evaluated the associations of coffee consumption

with plasma levels of sex hormones and SHBG, as well as the direct association between coffee consumption and T2D risk during a 10-year follow-up period. Further, we investigated whether the association of coffee consumption with T2D risk was attenuated by further adjusting for plasma sex hormones or SHBG. Finally, we examined whether coffee intake may interact with specific SHBG genotypes in affecting T2D risk.

2.2 Research Design and Methods

The Women's Health Study (WHS) is a randomized, double-blind, placebo-controlled trial originally designed to evaluate the balance of benefits and risks of low-dose aspirin and vitamin E in the primary prevention of cardiovascular disease and cancer.[11] Of the 39,876 participants aged 45 years and older, 98% of participants completed a 131-item semiquantitative food frequency questionnaire (SFFQ). At baseline, participants were asked if they were willing to provide blood samples by mail. Women who responded affirmatively and were eligible to be enrolled into the run-in phase were mailed a blood collection kit. 28,345 (71%) provided baseline blood samples, of which we restricted our study to 6,574 postmenopausal women who were not using hormone replacement therapy (HRT) at the time of blood collection. By February 2005, 366 of these initially healthy women reported developing incident type 2 diabetes. Controls were matched in 1:1 ratio to cases by age (within 1 year), duration of follow-up (within 1 month), race, and fasting status at time of blood draw (82% provided fasting blood samples, defined as $\geq$ 10 h since the last meal). Based on these eligibility criteria, 359 cases and 359 controls were included in our analyses. Written informed consent was obtained from all participants. This study was

approved by the Institutional Review Boards of Brigham and Women's Hospital, Harvard Medical School and the University of California at Los Angeles (UCLA).

Assessment of dietary intake

In the SFFQ, participants were asked how often on average during the previous year they had consumed caffeinated and decaffeinated-coffee ("one cup"), tea ("one cup or glass"), different types of caffeinated soft drinks ("one glass, bottle, or can"), and chocolate products (e.g., "bar or packet"). Participants could choose from nine responses (never or less than one per month, one to three per month, one per week, two to four per week, five to six per week, one per day, two to three per day, four to five per day, and six or more per day). Using U.S. Department of Agriculture food composition data supplemented with other sources, we estimated that the caffeine contents were 137 mg per cup of coffee, 47 mg per cup of tea, 46 mg per bottle or can of cola beverage, and 7 mg per serving of chocolate candy.[12] A validation study from a similar cohort of women reported high correlations between intake of coffee and other caffeinated beverages assessed with SFFQ and with four 1-week diet records (coffee, $r = 0.78$; tea, $r = 0.93$; and caffeinated sodas, $r = 0.85$).[13]

Ascertainment of incident type 2 diabetes

Details regarding ascertainment of incident type 2 diabetes in our cohorts have been reported previously.[14] After excluding those with diabetes at baseline, all participants were asked annually whether and when they had a diagnosis of diabetes since baseline. Using the diagnostic criteria of the American Diabetes Association,[15] all self-reported cases of T2D were confirmed by a supplemental questionnaire. The self-reported diabetes in the WHS was validated against

physician-led telephone interviews, supplementary questionnaires, and medical record reviews, all yielding positive predictive values >91%.[16]

Laboratory procedures

A mailed blood collection kit contained instructions, three 10 mL EDTA Vacutainer tubes, three 4.5 mL sodium citrate tubes, the supplies needed to draw a sample of blood, a completed overnight courier air bill, and a gel-filled freezer pack. The gel-filled freezer pack was frozen overnight to serve as a coolant during the mailing. The women were asked to provide a fasting blood sample in the morning drawn into two EDTA and two citrate tubes, and to return the completed blood kit via overnight courier. All the samples arrived in our laboratory within 24-30 hours of venipuncture. Upon receipt, the samples were kept chilled until processing. After centrifugation for 20 minutes (2500 rpm, 4 ºC) each sample was pipetted into a 2-mL Nunc vial. The samples were stored in liquid nitrogen tanks until the time of the laboratory analyses. Laboratory personnel were blinded to the case-control status, and the matched case-control pairs were handled identically and assayed in a random order during the same analytical run. The plasma concentrations of the sex hormones and SHBG were measured using chemiluminescent immunoassays (Elecsys autoanalyzer 2010, Roche Diagnostics, Indianapolis, IN, USA), which have been validated for the measurement of plasma sex hormones and SHBG levels.[17-19] For the hormone levels used in this study, the coefficients of variation from blinded quality control samples were 5.2% for estradiol, 7.4% for testosterone, 2.8% for dehydroepiandrosterone sulfate (DHEAS), and 2.8% for SHBG. The detailed methods for SNP selection and the genotyping of the SHBG single nucleotide polymorphisms (SNPs) have been described previously.[10] Two informative SNPs associated with plasma SHBG levels were included in our study: rs6259 in

exon 8, which encodes an amino acid substitution of asparagine for aspartic acid and may lead to the reduced clearance of SHBG; and rs6257 in intron 1.

Statistical analysis

Following the conventional practice used in previous studies, we categorized caffeinated-coffee, decaffeinated-coffee, and tea consumption by the aggregating nine possible responses for caffeinated-coffee, decaffeinated-coffee, and tea on the SFFQ into four categories (no cups per day, less than one cup to one cup per day, 2 to 3 cups per day, and 4 or more cups per day). We also categorized caffeine consumption into four categories (≤50, 51-250, 251-500, and >500 mg/day).

Baseline characteristics were compared between case patients and controls using the paired t-test for continuous variables and McNemar's test for categorical variables. To assess the association of caffeine-related beverage consumption (caffeinated-coffee, decaffeinated-coffee, tea, and caffeine) with sex hormones (estradiol, testosterone, and DHEAS) and SHBG levels, we calculated the geometric means of sex hormones and SHBG plasma levels according to the four categories of caffeine-related beverage consumption. We used multiple linear regression models to adjust for matching factors (age, race, duration of follow-up, and time of blood draw), smoking status (never, past, and current smokers), physical activity level (rarely/never, <1, 1-3, and ≥4 times/week), alcohol use (rarely/never, 1-3 drinks/months, 1-6 drinks/week, ≥1 drinks/day), total calories (≤1500, 1501-2000, 2001-2500, and >2500 kcal/day), and body mass index (BMI) (continuous). To test for a linear trend across increasing categories of caffeine-related beverage consumption, we computed the median value for each category and included this as a continuous variable in the multiple linear regression models.

To assess the relations of caffeine-related beverage consumption with T2D risk, we used conditional logistic regression models to adjust for matched pairs (Match-adjusted model). We further adjusted for smoking status (never, past, and current smokers), physical activity (rarely/never, <1, 1-3, and ≥4 times/week), family history of diabetes (yes or no), alcohol use (rarely/never, 1-3 drinks/months, 1-6 drinks/week, ≥1 drinks/day), total calories (≤1500, 1501-2000, 2001-2500, and >2500 kcal/day), and BMI (continuous) (Categorical model). To test for a linear trend across increasing categories of caffeine-related beverage consumption, we computed the median value for each category and included this as a continuous variable in the conditional logistic regression models. Because caffeine and caffeinated-coffee consumption were associated with plasma SHBG levels but not with sex hormones levels, in subsequent analyses, we further included the plasma SHBG levels (≤20.0, 20.1-25.0, 25.1-30.0, and >30.0 nmol/L) in the models (Categorical model + SHBG).

To further provide a visual representation of the dose-response curve, we fitted quadratic spline models by including transformed variables of caffeine-related beverage consumption to multiple regression models using a single knot at the middle category boundary used in each Categorical model [20]. Finally, we examined the potential effect modification by two informative *SHBG* SNPs, rs6259 and rs6257, using the dominant genetic model. We calculated the adjusted plasma SHBG levels and odds ratios for T2D for combinations of *SHBG* genotypes and caffeinated-coffee intake levels (2 cups/day or more vs. less than 2 cups/day). Wald tests were used to test for statistical interaction by entering product terms to the regression models. We performed the χ^2 test to evaluate Hardy-Weinberg equilibrium for rs6259 and rs6257 among the

controls. All the statistical analyses were conducted using SAS (version 9.2; SAS Institute, Cary, NC).

2.2 Results

Compared with control participants, T2D cases had a greater proportion of traditional risk factors at baseline (Table 1). T2D cases had higher levels of plasma total estradiol and lower levels of plasma SHBG, but plasma total testosterone and DHEAS appeared to be similar between cases and controls.

Caffeinated-coffee and caffeine intakes were positively associated with plasma SHBG levels but not with sex hormones (Table 2 and Table 3). For caffeinated-coffee, the multivariate-adjusted geometric mean levels of plasma SHBG were 26.6 nmol/L (95% confidence interval [CI], 18.9-37.4) in women consuming ≥4 cups/day and 23.0 nmol/L (95% CI, 16.5-32.0) in non-drinkers (*P* for trend = 0.01). For caffeine, the multivariate-adjusted geometric mean levels of plasma SHBG were 26.6 nmol/L (95% CI, 19.0-37.4) in women consuming >500 mg/day and 22.9 nmol/L (95% CI, 16.5-32.0) in women consuming ≤50 mg/day (*P* for trend = 0.02) (Table 2). We found similar results using quadratic spline models that imposed smooth dose-response relations (Table 2 and Figure 1A). The spline plots indicated that heavy drinkers of caffeinated-coffee (>2 cups/day) were associated with higher levels of plasma SHBG (Figure 1A). In contrast, decaffeinated-coffee and tea intakes were not associated with plasma SHBG levels and sex hormone levels (Table 3).

Caffeinated-coffee and caffeine intakes were also inversely associated with risk of T2D (Table 4). The multivariate-adjusted odds ratios (ORs) of T2D were 0.47 (95% CI: 0.23 – 0.94; *P*

for trend = 0.047) for ≥4 cups/day of caffeinated-coffee compared with non-drinkers, and 0.56 (95% CI: 0.27 – 1.15; *P* for trend = 0.18) for >500 mg/day compared with ≤50 mg/day of caffeine. Little or no association of decaffeinated-coffee and tea consumption with type 2 diabetes was observed. After further adjusting for plasma SHBG levels, the inverse associations of caffeinated-coffee and caffeine with T2D risk were attenuated (Categorical model + SHBG). Compared with non-drinkers, the ORs for ≥4 cups/day of caffeinated-coffee changed from 0.47 to 0.71 (95% CI, 0.31 – 1.61; *P* for trend = 0.47). Similarly, compared with <50 mg/day of caffeine, the ORs for ≥500 mg/day were changed from 0.56 to 0.89 (95% CI, 0.38 – 2.10; *P* for trend = 0.91). In contrast, further adjustment for plasma sex hormones instead of SHBG did not change the association. Similar results were shown in our fitted spline logistic regression models (Table 4, Spline model). In the spline plots, before adjusting for plasma SHBG, an inverse trend between caffeinated-coffee and risk of type 2 diabetes was observed above 2 cups/day of caffeinated-coffee (Figure 1B). This trend disappeared after further adjusting for plasma SHBG (Figure 1C).

Finally, we estimated the multivariable-adjusted geometric mean levels of plasma SHBG and multivariable-adjusted ORs of T2D for combinations of SHBG genotypes and caffeinated-coffee intake levels (≥2 cups/day vs. <2 cups/day) (Table 5). We detected no departure from Hardy-Weinberg equilibrium for rs6259 and rs6257 SNPs among the controls (*P*=0.24 and *P*=0.06, respectively). Carriers of the rs6259 minor allele and non-carriers of the rs6257 minor allele who consumed high caffeinated-coffee had a lower risk of T2D in directions corresponding to their associated plasma SHBG levels: as compared with low-drinkers (<2 cups/day) without the rs6259 minor allele, high-drinkers (≥2 cups/day) with the minor allele had

20% higher plasma SHBG levels (27.8 vs. 23.2 nmol/L) and were associated with lower risk of type 2 diabetes (OR =0.54; 95% CI, 0.26-1.11). Similarly, as compared with low-drinkers (< 2 cups/day) with the rs6257 minor allele, high-drinkers (≥2 cups/day) without the minor allele had 24% higher plasma SHBG levels (25.2 vs. 20.3 nmol/L) and had a lower risk T2D (OR =0.38; 95% CI, 0.18-0.83).

2.3 Discussion

In this prospective study of postmenopausal women, caffeinated-coffee and caffeine intakes were positively associated with plasma SHBG levels. Also, we observed an inverse association between intake of caffeinated-coffee and caffeine and risk of T2D. The associations were largely attenuated after adjustment for SHBG levels. Finally, carriers of the rs6259 minor allele and non-carriers of the rs6257 minor allele who consumed high caffeinated-coffee had a lower risk of T2D in directions corresponding to their associated plasma SHBG levels. These findings suggest that SHBG may account for the inverse association between caffeinated-coffee and T2D risk.

The inverse associations of caffeinated-coffee and caffeine intake with T2D risk observed in our study are consistent with findings from previous studies.[1,2] Several possible explanations have been put forth to explain the protective effect of coffee consumption on T2D risk, including the effects on insulin sensitivity and β cell function by varying coffee components such as magnesium, potassium, chlorogenic acid, and caffeine.[2] To date, however, little is known about the underlying mechanisms. Evidence from a systematic review suggests the sex differences in

the inverse association between coffee and the risk of T2D.[2] Moreover, both observational and experimental data indicate the important roles of sex hormones in the development of T2D.[6-8] SHBG is synthesized primarily in the liver and binds androgens with high affinity and estrogens with low affinity, thereby regulating the biologically active fraction of sex hormones.[21] Recently, it has been shown that the plasma membranes of a variety of cells are able to bind SHBG specifically and with high affinity, and SHBG mediates sex hormones signaling at the cell membrane through the SHBG receptors.[9] This discovery of the function of SHBG as a mediator of a steroid-signaling system has drawn much interest to biologic effects of SHBG. We first reported that genetically determined lower levels of SHBG were associated with T2D risk using Mendelian randomization analyses[10], findings of which have been replicated by a large consortium of case-control studies.[22] Taken together, we hypothesized that caffeinated-coffee consumption may lower the risk of T2D possibly via altering SHBG metabolism.

We found that caffeine and caffeinated-coffee intakes were positively associated with plasma SHBG levels, which is consistent with earlier studies.[4,5,23-25] Little or no association between decaffeinated-coffee and plasma SHBG levels suggest that caffeine may be a key component of coffee responsible for determining plasma SHBG levels. Moreover, our findings of little or no relations between caffeine-related beverage consumption and sex hormones suggest that caffeine may increase the level of plasma SHBG without directly altering sex hormones levels. Caffeine and other major components of coffee (cafestol and kahweol) alter the expressions and activities of liver enzymes.[26-29] Because SHBG is synthesized and metabolized primarily in the liver,[21] coffee intake may affect SHBG metabolism in the liver, influencing the plasma SHBG levels.[5]

Coffee may increase plasma SHBG levels, resulting not only in affecting the biologic actions of sex hormones by binding to circulating androgens and estrogens but also in exerting direct metabolic effects.[9] Our findings thus provide a new explanation for the potential protective effect of coffee consumption on the T2D risk. Notably, we found that carriers of the rs6259 minor allele and non-carriers of the rs6257 minor allele who consumed high caffeinated-coffee had a lower risk of T2D in directions corresponding to their associated plasma SHBG levels. These findings may further support the notion that SHBG may account for the potential protective effect of caffeinated-coffee on T2D. In contrast, the role of specific sex-steroids in relation to the coffee-diabetes relation remains to be determined.

The strengths of our study include its prospective study design with 10-year follow-up with comprehensive assessment of baseline variables, blood samples, and SHBG genotypes. Nevertheless, our study has several limitations. First, cross-sectional analyses of coffee consumption and plasma SHBG may be a concern, although it is not likely that endogenous sex hormones or SHBG would influence the consumption. Second, we cannot exclude the possibilities of residual confounding from unmeasured or incompletely measured covariates even though we have adjusted for many major risk factors for type 2 diabetes. Third, misclassifications of dietary intakes and biomarker measures-are inevitable. For example, there may be measurement errors of plasma sex hormones and SHBG due to the limitations of stored samples. However, as cases were identified prospectively and case-control pair were matched and handled in an identical fashion in the same analytical run, any potential misclassifications should affect case and control equally. Therefore, such misclassifications were likely to be non-differential, which would lead to an underestimation of the associations. Fourth, there is a

concern about the possibility of residual confounding from unmeasured time-dependent confounders when a standard method is performed to adjust for both an exposure and a measured intermediate variable. Finally, our study only included postmenopausal women, which may limit the generalizability of our findings to premenopausal women or men.

In conclusion, our results suggest that SHBG levels may account for the potential protective effect of habitual coffee consumption against the risk of T2D among postmenopausal women. A better understanding of the underlying mechanisms requires further investigation in both observational and experimental settings.

2.5 Acknowledgements

Presented was published in the Diabetes journal on October 28th, 2010.[30]

This work was presented in part at the 50th Cardiovascular Disease Epidemiology and Prevention Annual Conference 2010, March 5th, 2010, San Francisco, California.

Table 1. Baseline Characteristics between Participants with Incident Cases of Type 2 Diabetes and Control Participants among 718 Women

Characteristics	Cases	Controls	P-values *
No.	359	359	
Age (yr)	60.3 ± 6.1	60.3 ± 6.1	
Caucasian (%)	93.5	93.5	
BMI (kg/m^2)	30.9 ± 6.1	26.0 ± 5.0	<0.001
Alcohol (g/day)	2.62 ± 7.4	4.19 ± 8.3	0.008
Smoking (% current)	14.5	13.7	0.74
Physical activity (% ≥ once/wk)	30.7	38.7	0.02
Family history of diabetes (%)	48.5	24.0	<0.001
Past postmenopausal hormone use (%)	34.0	29.3	0.17
Ever oral contraceptive use (%)	50.4	48.0	0.57
Age at menopause	48.0 ± 6.2	48.0 ± 5.8	0.79
Years since menopause	12.2 ± 8.2	12.2 ± 8.0	0.77
Age at menarche <12 (%)	25.4	21.7	0.23
Age at first pregnancy of ≥6 months, <25 (%)	63.4	57.2	0.37
Pregnancies ≥5 (%)	18.7	19.9	0.69
Marital status (% currently married)	65.7	68.2	0.28
Caffeine-related beverages			
Caffeinated-coffee (% ≥ 4 cups/day)	13.8	20.9	0.01
Decaffeinated-coffee (% ≥ 4 cups/day)	2.3	4.3	0.20
Tea (% ≥ 4 cups/day)	5.2	2.9	0.13
Caffeine (% ≥ 500 mg/day)	14.3	21.1	0.02
Sex hormones			
SHBG (nmol/l)	22.3 ± 13.8	36.9 ± 17.4	<0.001
Estradiol (pg/mL)	24.6 ± 15.9	20.5 ± 11.3	<0.001
Testosterone (ng/dL)	29.8 ± 19.1	28.9 ± 19.1	0.49
DHEAS (µg/dL)	91.0 ± 61.3	92.6 ± 53.7	0.67

Abbreviation: SHBG, sex hormone-binding globulin; DHEAS, dehydroepiandrosterone.
Data are means ± SD.

* Baseline characteristics were compared between case patients and controls using the paired t-test for continuous variables and the McNemar's test for categorical variables.

Table 2. Geometric Mean Levels of Sex Hormone-binding Globulin according to Caffeinated-Coffee, Decaffeinated-Coffee, Tea, and Caffeine Consumption

Plasma SHBG levels (nmol/L)	Categories of intake				*P* for trend *
	No cups per day	Less than one cup or one cup per day	Two to three cups per day	Four or more cups per day	
Caffeinated-coffee, median (n)	0 (185)	1.0 (187)	2.5 (212)	4.5 (122)	
Match-adjusted model	23.1 (16.1-33.1)	22.7 (15.8-32.7)	24.7 (17.2-35.5)	27.6 (19.1-39.9)	0.002
Categorical model †	23.0 (16.5-32.0)	22.8 (16.3-31.8)	23.6 (16.9-33.0)	26.6 (18.9-37.4)	0.01
Spline model ‡	22.9 (16.5-31.9)	22.7 (16.2-31.8)	24.1 (17.2-33.6)	26.0 (18.5-36.5)	
Decaffeinated-coffee, median (n)	0 (401)	0.4 (188)	2.5 (84)	4.5 (23)	
Match-adjusted model	24.5 (17.1-35.2)	24.3 (16.8-35.4)	25.2 (17.1-37.3)	26.7 (17.4-40.9)	0.44
Categorical model †	23.7 (17.0-33.1)	22.3 (15.8-31.3)	23.2 (16.2-33.3)	22.9 (15.5-33.9)	0.75
Spline model ‡	23.6 (16.9-32.9)	23.5 (16.8-33.0)	23.6 (16.5-33.6)	23.5 (16.2-34.3)	
Tea, median (n)	0 (242)	0.4 (351)	2.5 (76)	4.5 (28)	
Match-adjusted model	25.6 (17.9-36.7)	23.6 (16.4-33.8)	24.6 (16.9-35.8)	21.6 (14.3-32.5)	0.27
Categorical model †	24.5 (17.6-34.1)	23.3 (16.7-32.6)	23.3 (16.5-32.8)	21.4 (14.7-31.3)	0.24
Spline model ‡	24.4 (17.5-33.9)	23.7 (17.0-33.1)	22.5 (16.1-31.5)	22.7 (15.9-32.3)	
Caffeine category (mg/day)	≤50	51 – 250	251 – 500	>500	
Caffeine intake, Median (mg/day) (n)	13 (131)	140 (209)	366 (230)	656 (123)	
Match-adjusted model	22.9 (15.8-33.0)	22.5 (15.6-32.5)	23.6 (16.4-34.0)	26.9 (18.5-39.0)	0.008
Categorical model †	22.9 (16.5-32.0)	23.2 (16.7-32.3)	23.0 (16.5-32.1)	26.6 (19.0-37.4)	0.02
Spline model ‡	23.0 (16.5-32.1)	23.0 (16.6-32.0)	23.7 (17.0-33.0)	25.6 (18.3-35.8)	

Data are expressed as geometric means (95% CI). Abbreviation: SHBG, sex hormone-binding globulin; CI, confidence interval.

* *P* for trend are based on median values in categories of the participants.

** Match-adjusted model: adjusted for age, race, duration of follow-up, and time of blood draw.

† Categorical model: adjusted for matching factors, smoking status, physical activity, alcohol use, total calories, and BMI.

‡ Spline model: estimates at category medians from quadratic spline regression models with one knot at the middle category boundaries, adjusted for covariates used in categorical model.

Table 3. Geometric Mean Levels of Total Estradiol, Total Testosterone, and DHEAS according to Caffeinated-Coffee, Decaffeinated-Coffee, Tea, and Caffeine Consumption

	Categories of intake				*P* for trend *
	No cups per day	Less than one cup or one cup per day	Two to three cups per day	Four or more cups per day	
Caffeinated-coffee, median (n)	0 (185)	1.0 (187)	2.5 (212)	4.5 (122)	
Total estradiol (pg/mL)	23.6 (17.9-31.1)	21.2 (16.1-28.1)	23.9 (18.0-31.6)	22.0 (16.5-29.2)	0.76
Total testosterone (ng/dL)	20.7 (13.4-32.0)	21.5 (13.8-33.3)	22.8 (14.6-35.4)	22.3 (14.3-35.0)	0.23
DHEAS (µg/dL)	87.8 (57.6-133.7)	89.0 (58.2-136.1)	94.4 (61.6-144.6)	93.3 (60.4-144.0)	0.28
Decaffeinated-coffee, median (n)	0 (401)	0.4 (188)	2.5 (84)	4.5 (23)	
Total estradiol (pg/mL)	23.0 (17.4-30.5)	21.6 (16.2-28.8)	22.7 (16.8-30.8)	21.8 (15.7-30.3)	0.67
Total testosterone (ng/dL)	21.0 (13.6-32.4)	18.7 (11.9-29.2)	19.7 (12.3-31.6)	19.9 (11.9-33.3)	0.55
DHEAS (µg/dL)	86.3 (56.5-131.8)	83.1 (53.8-128.5)	84.5 (53.5-133.6)	88.5 (53.7-145.6)	0.99
Tea, median (n)	0 (242)	0.4 (351)	2.5 (76)	4.5 (28)	
Total estradiol (pg/mL)	22.9 (17.3-30.2)	23.5 (17.8-31.1)	24.6 (18.5-32.7)	22.4 (16.3-30.8)	0.60
Total testosterone (ng/dL)	21.7 (14.1-33.6)	23.0 (14.8-35.6)`	21.9 (14.0-34.3)	19.7 (11.9-32.4)	0.49
DHEAS (µg/dL)	87.8 (57.7-133.5)	94.9 (62.2-144.8)	88.2 (57.2-136.1)	78.1 (48.3-126.2)	0.36
Caffeine category (mg/day)	< 50	50 – 249	250 – 499	≥ 500	
Caffeine intake, Median (mg/day) (n)	13 (131)	140 (209)	366 (230)	656 (123)	
Total estradiol (pg/mL)	23.0 (17.4-30.4)	22.4 (16.9-29.5)	24.2 (18.3-32.0)	22.4 (16.9–29.8)	0.79
Total testosterone (ng/dL)	19.4 (12.5-30.0)	21.9 (14.2-33.8)	22.1 (14.2-34.2)	22.1 (14.2-34.5)	0.21
DHEAS (µg/dL)	84.2 (55.1-128.7)	90.2 (59.1-137.6)	92.3 (60.3-141.3)	90.5 (58.7-139.5)	0.42

Data are expressed as geometric mean (95% CI). Abbreviation: DHEAS, dehydroepiandrosterone sulfate; CI, confidence interval.

* *P* for trend are based on median values in categories of the participants.

All models were adjusted for matching factors (age, race, duration of follow-up, and time of blood draw), smoking status, physical activity, alcohol use, total calories, and BMI.

Table 4. Odds Ratios for Type 2 Diabetes according to Caffeinated-Coffee, Decaffeinated-Coffee, Tea, and Caffeine Consumption

	Categories of intake				
	No cups per day	**Less than one cup to one cup per day**	**Two to three cups per day**	**Four or more cups per day**	***P* for trend** *
Caffeinated-coffee (cases/controls)	99/86	103/84	105/107	49/73	
Median (cups per day)	0	1.0	2.5	4.5	
Match-adjusted model†	1.00	1.04 (0.68-1.60)	0.82 (0.55-1.23)	0.55 (0.34-0.90)	0.008
Categorical model‡	1.00	0.95 (0.52-1.74)	0.94 (0.53-1.67)	0.47 (0.23-0.94)	0.047
Categorical model + SHBG §	1.00	0.92 (0.46-1.84)	0.96 (0.48-1.94)	0.71 (0.31-1.61)	0.47
Spline model ‖	1.00	1.11 (0.63-1.97)	0.92 (0.55-1.54)	0.61 (0.34-1.11)	
Spline model + SHBG ¶	1.00	0.87 (0.45-1.68)	0.86 (0.47-1.58)	0.80 (0.39-1.65)	
Decaffeinated-coffee (cases/controls)	211/190	94/94	37/47	8/15	
Median (cups per day)	0	0.4	2.5	4.5	
Match-adjusted model†	1.00	0.87 (0.60-1.26)	0.63 (0.38-1.03)	0.49 (0.20-1.20)	0.03
Categorical model ‡	1.00	1.16 (0.70-1.95)	0.77 (0.38-1.55)	0.72 (0.19-2.69)	0.39
Categorical model + SHBG §	1.00	1.03 (0.56-1.90)	1.11 (0.46-2.71)	1.33 (0.27-6.49)	0.71
Spline model ‖	1.00	1.30 (0.90-1.88)	0.61 (0.33-1.14)	0.47 (0.17-1.29)	
Spline model + SHBG ¶	1.00	1.31 (0.84-2.03)	0.91 (0.42-2.00)	0.92 (0.26-3.27)	
Tea (cases/controls)	114/128	180/171	37/39	18/10	
Median (cups per day)	0	0.4	2.5	4.5	
Match-adjusted model†	1.00	1.21 (0.87-1.69)	1.06 (0.64-1.74)	2.03 (0.91-4.54)	0.23
Categorical model ‡	1.00	1.12 (0.70-1.77)	1.22 (0.59-2.50)	1.53 (0.43-5.47)	0.46
Categorical model + SHBG §	1.00	0.97 (0.55-1.72)	1.14 (0.48-2.72)	1.74 (0.44-6.93)	0.42
Spline model ‖	1.00	1.39 (0.99-1.93)	1.52 (0.83-2.77)	1.32 (0.51-3.38)	
Spline model + SHBG ¶	1.00	1.37 (0.93-2.03)	1.37 (0.67-2.78)	1.57 (0.54-4.58)	
Caffeine category (mg/day)	≤50	51 – 250	251 - 500	>500	
Caffeine intake, Median (mg/day)	13	140	366	656	
(cases/controls)	67/64	109/100	118/112	49/74	
Match-adjusted model †	1.00	1.01 (0.65-1.56)	0.99 (0.63-1.54)	0.62 (0.37-1.04)	0.06
Categorical model ‡	1.00	1.00 (0.56-1.82)	1.26 (0.68-2.32)	0.56 (0.27-1.15)	0.18
Categorical model + SHBG §	1.00	0.94 (0.47-1.88)	1.44 (0.68-3.04)	0.89 (0.38-2.10)	0.91
Spline model ‖	1.00	1.56 (0.80-3.04)	1.56 (0.81-3.01)	0.92 (0.48-1.76)	
Spline model + SHBG ¶	1.00	1.53 (0.71-3.30)	1.70 (0.79-3.62)	1.32 (0.61-2.82)	

Data are ORs (95% CI). Abbreviation: SHBG, sex hormone-binding globulin; ORs, odds ratios; CI, confidence interval

* *P*-values for trend are based on median levels in categories. † Match-adjusted model: stratified on matched pairs using conditional logistic regression models. ‡ Categorical model: further adjusted for smoking status, physical activity, family history of diabetes, alcohol use, total calories, and BMI. § Categorical model + SHBG: further adjusted for plasma SHBG. ‖ Spline model: odds ratios comparing odds at category medians from quadratic logistic spline models with one knot at the middle category boundaries adjusted for covariates used in categorical model. ¶ Spline model plus plasma SHBG: further adjusted for plasma SHBG.

Table 5. Caffeinated-Coffee Consumption in Relation to Plasma SHBG and Type 2 Diabetes Stratified by SHBG SNPs

	SHBG Genotype							
		rs6259 GG (wild type)		AG or AA (variant)		rs6257 CT or CC (variant)		TT (wild type)
Plasma SHBG Levels (nmol/L) *	No.		No.		No.		No.	
Caffeinated-Coffee Intake								
Low (< 2 cups per day)	277	23.2 (16.6-32.5)	70	23.3 (16.4-32.9)	69	20.3 (14.2-29.0)	285	23.2 (16.6-32.4)
High (≥ 2 cups per day)	246	24.4 (17.4-34.3)	76	27.8 (19.4-39.7)	72	21.6 (15.2-30.8)	244	25.2 (17.9-35.4)
***P* for interaction** †	0.18				0.84			
Odds Ratios ‡	**(cases/ controls)**		**(cases/ controls)**		**(cases/ controls)**		**(cases/ controls)**	
Caffeinated-Coffee Intake								
Low (< 2 cups per day)	152/125	1.00 (reference)	40/30	0.90 (0.40-2.00)	42/27	1.00 (reference)	149/136	0.41 (0.19-0.88)
High (≥ 2 cups per day)	122/124	0.70 (0.42-1.16)	26/50	0.54 (0.26-1.11)	34/38	0.40 (0.16-1.02)	109/135	0.38 (0.18-0.83)
***P* for interaction** †	0.79				0.13			

* Multivariate-adjusted geometric mean SHBG levels with 95% confidence intervals for combinations of SHBG genotypes and caffeinated-coffee intake levels (2 cups/day or more vs. less than 2 cups/day) adjusted for matching factors, smoking status, physical activity, family history of diabetes, alcohol use, total calories, and BMI.

† Wald tests were used to test for statistical interaction by entering product terms into the regression models.

‡ Multivariable adjusted odds ratios and 95% confidence intervals of type 2 diabetes risk for combinations of SHBG genotypes and caffeinated-coffee intake levels (2 cups/day or more vs. less than 2 cups/day) adjusted for matching factors, smoking status, physical activity, family history of diabetes, alcohol use, total calories, and BMI.

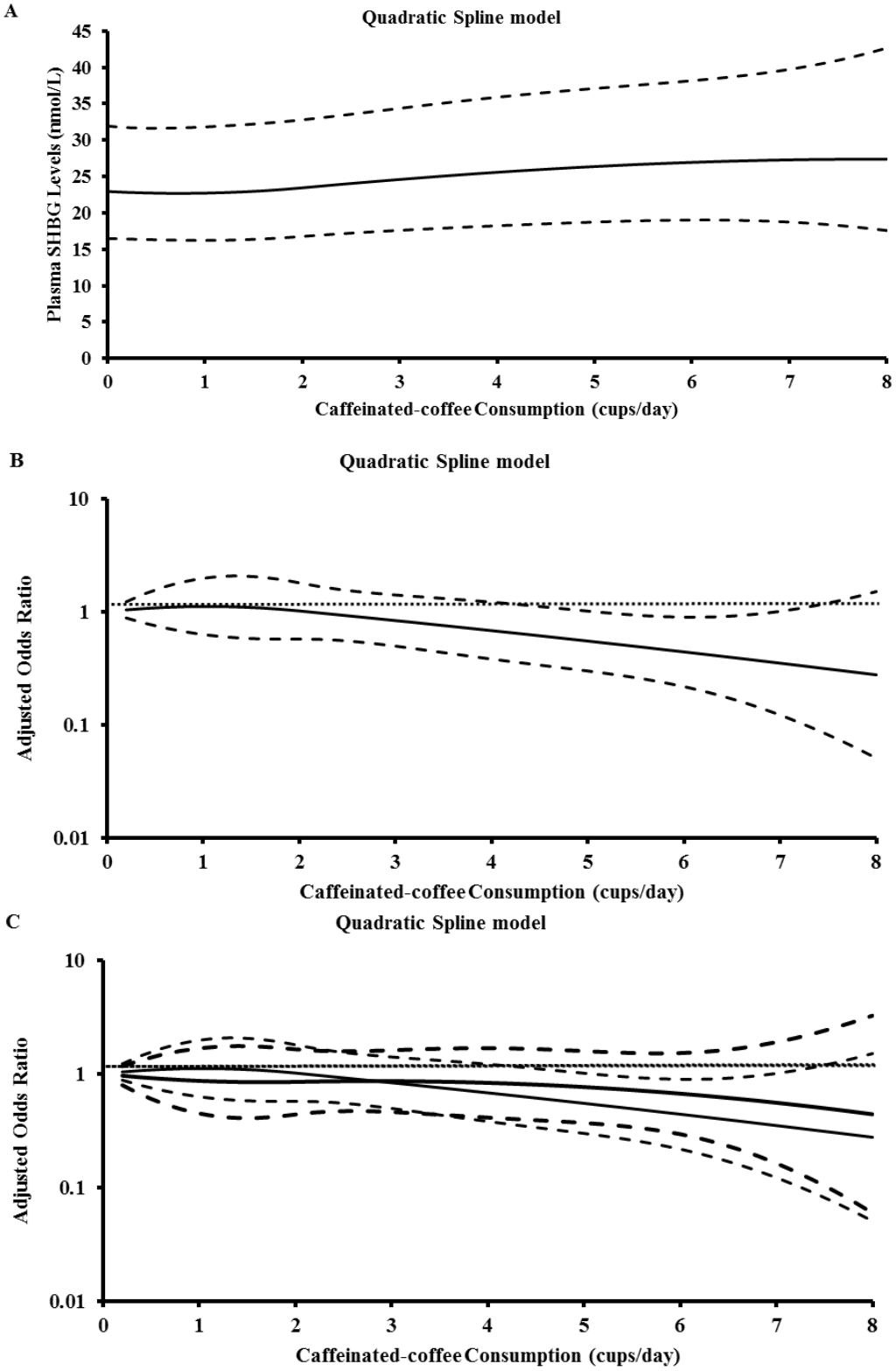
Figure legends

Panel A shows the geometric mean SHBG levels adjusted for matching factors, smoking status, physical activity, family history of diabetes, alcohol use, total calories, and BMI from quadratic spline model (solid curve) with pointwise 95% confidence limits (dashed curves).

Panel B shows the odds ratio of type 2 diabetes from quadratic conditional logistic spline model adjusted for matching factors, smoking status, physical activity, family history of diabetes, alcohol use, total calories, and BMI (solid curve) with pointwise 95 % confidence limits (dashed curves).

Panel C shows the multivariate-adjusted odds ratio of type 2 diabetes from quadratic conditional logistic spline model with further adjustment for plasma SHBG.

Figure 1. Estimated Plasma SHBG and Risk of Type 2 Diabetes According to Caffeinated-Coffee Consumption



2.7 References

1. Huxley R, Lee CM, Barzi F, et al. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: A systematic review with meta-analysis. *Arch Intern Med.* Dec 14 2009;169(22):2053-2063.

2. van Dam RM, Hu FB. Coffee consumption and risk of type 2 diabetes: a systematic review. *JAMA.* Jul 6 2005;294(1):97-104.

3. Isogawa A, Noda M, Takahashi Y, Kadowaki T, Tsugane S. Coffee consumption and risk of type 2 diabetes mellitus. *Lancet.* Feb 22 2003;361(9358):703-704.

4. London S, Willett W, Longcope C, McKinlay S. Alcohol and other dietary factors in relation to serum hormone concentrations in women at climacteric. *The American journal of clinical nutrition.* Jan 1991;53(1):166-171.

5. Kotsopoulos J, Eliassen AH, Missmer SA, Hankinson SE, Tworoger SS. Relationship between caffeine intake and plasma sex hormone concentrations in premenopausal and postmenopausal women. *Cancer.* Jun 15 2009;115(12):2765-2774.

6. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA.* Mar 15 2006;295(11):1288-1299.

7. Ding EL, Song Y, Manson JE, Rifai N, Buring JE, Liu S. Plasma sex steroid hormones and risk of developing type 2 diabetes in women: a prospective study. *Diabetologia.* Oct 2007;50(10):2076-2084.

8. Rincon J, Holmang A, Wahlstrom EO, et al. Mechanisms behind insulin resistance in rat skeletal muscle after oophorectomy and additional testosterone treatment. *Diabetes.* May 1996;45(5):615-621.

9. Rosner W, Hryb DJ, Khan MS, Nakhla AM, Romas NA. Sex hormone-binding globulin mediates steroid hormone signal transduction at the plasma membrane. *The Journal of steroid biochemistry and molecular biology.* Apr-Jun 1999;69(1-6):481-485.

10. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N Engl J Med.* Sep 17 2009;361(12):1152-1163.

11. Buring JE, Hennekens CH. The Women's Health Study: summary of the study design. *J Myocard. Ischemia.* 1992;4:27-29.

12. Watt BK, Merrill AL. *Composition of Foods: Raw, Processed, Prepared, 1963-1992: Agriculture Handbook no. 8.*: Washington, DC, U.S. Department of Agriculture; 1993.

13. Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *International journal of epidemiology.* Dec 1989;18(4):858-867.

14. Song Y, Manson JE, Buring JE, Liu S. A prospective study of red meat consumption and type 2 diabetes in middle-aged and elderly women: the women's health study. *Diabetes care.* Sep 2004;27(9):2108-2115.

15. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes care.* Jul 1997;20(7):1183-1197.

16. Liu S, Lee IM, Song Y, et al. Vitamin E and risk of type 2 diabetes in the women's health study randomized controlled trial. *Diabetes.* Oct 2006;55(10):2856-2862.

17. Reynders M, Anckaert E, Schiettecatte J, Smitz J. Evaluation of a new automated electrochemiluminescent sex hormone-binding globulin (SHBG) immunoassay. *Clin Chem Lab Med.* 2005;43(1):86-89.

18. Sanchez-Carbayo M, Mauri M, Alfayate R, Miralles C, Soria F. Elecsys testosterone assay evaluated. *Clin Chem.* Aug 1998;44(8 Pt 1):1744-1746.

19. Yang DT, Owen WE, Ramsay CS, Xie H, Roberts WL. Performance characteristics of eight estradiol immunoassays. *Am J Clin Pathol.* Sep 2004;122(3):332-337.

20. Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology (Cambridge, Mass.* Jul 1995;6(4):356-365.

21. Anderson DC. Sex-hormone-binding globulin. *Clinical endocrinology.* Jan 1974;3(1):69-96.

22. Perry JR, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Hum Mol Genet.* Feb 1;19(3):535-544.

23. Lucero J, Harlow BL, Barbieri RL, Sluss P, Cramer DW. Early follicular phase hormone levels in relation to patterns of alcohol, tobacco, and coffee use. *Fertil Steril.* Oct 2001;76(4):723-729.

24. Nagata C, Kabuto M, Shimizu H. Association of coffee, green tea, and caffeine intakes with serum concentrations of estradiol and sex hormone-binding globulin in premenopausal Japanese women. *Nutrition and cancer.* 1998;30(1):21-24.

25. Ferrini RL, Barrett-Connor E. Caffeine intake and endogenous sex steroid levels in postmenopausal women. The Rancho Bernardo Study. *American journal of epidemiology.* Oct 1 1996;144(7):642-644.

26. Park BK, Kitteringham NR. Assessment of enzyme induction and enzyme inhibition in humans: toxicological implications. *Xenobiotica.* Nov 1990;20(11):1171-1185.

27. Higgins LG, Cavin C, Itoh K, Yamamoto M, Hayes JD. Induction of cancer chemopreventive enzymes by coffee is mediated by transcription factor Nrf2. Evidence that the coffee-specific diterpenes cafestol and kahweol confer protection against acrolein. *Toxicol Appl Pharmacol.* Feb 1 2008;226(3):328-337.

28. Huber WW, Rossmanith W, Grusch M, et al. Effects of coffee and its chemopreventive components kahweol and cafestol on cytochrome P450 and sulfotransferase in rat liver. *Food Chem Toxicol.* Apr 2008;46(4):1230-1238.

29. Huber WW, Scharf G, Rossmanith W, et al. The coffee components kahweol and cafestol induce gamma-glutamylcysteine synthetase, the rate limiting enzyme of chemoprotective glutathione synthesis, in several organs of the rat. *Arch Toxicol.* Jan 2002;75(11-12):685-694.

30. Goto A, Song Y, Chen BH, Manson JE, Buring JE, Liu S. Coffee and caffeine consumption in relation to sex hormone-binding globulin and risk of type 2 diabetes in postmenopausal women. *Diabetes.* Jan 2011;60(1):269-275.

Chapter 3

Age, Body Mass Index, Reproductive History, Usage of exogenous estrogen, and Lifestyle Factors in Relation to Circulating Sex Hormone-binding Globulin Levels in Multiethnic Postmenopausal Women: A Pooled Analysis of Studies in The Women's Health Initiative Randomized Trial and Observational Study

3.1 Introduction

Sex hormone-binding globulin (SHBG) is synthesized primarily in the liver and binds to androgens with a high affinity and estrogens with a low affinity.[1] The primary role of SHBG was thought to regulate the biologically active fraction of sex hormones and their accessibility to target cells.[1] Remarkably, the plasma membranes of various cells are able to bind to SHBG specifically and with a high affinity, and SHBG mediates sex hormone signaling at the cell membrane through SHBG receptors.[2,3] The discovery of a function for SHBG as a mediator of a novel steroid-signaling system has drawn much attention to the biologic effects of SHBG. Emerging evidence indicates that SHBG may play important roles in the development of hormone-dependent cancers,[4,5] cardiovascular diseases,[6] type 2 diabetes (T2D),[7] and hip fracture.[8] Importantly, recent molecular epidemiologic studies have shown that genetically determined levels of SHBG were inversely associated with T2D risk using Mendelian randomization (MR) analysis,[9,10] and the direction and magnitude of the association between the genetically determined levels of SHBG and T2D risk were consistent with the multivariable-adjusted association observed in earlier studies. Further, a recent meta-analysis found that lower SHBG levels were associated with a higher risk of metabolic syndrome in both men and

women.[11] Given the clinical importance of SHBG, an investigation of the determinants of the plasma SHBG levels is of great interest. Both genetic and environmental factors are hypothesized to determine the plasma levels of SHBG.[12] However, no studies have comprehensively examined the potential non-genetic determinants of circulating SHBG levels in a large-scale, multiethnic population-based study. Therefore, we investigated the relation of age, usage of exogenous estrogen, reproductive history, adiposity, and lifestyle factors with the circulating SHBG levels among participants of the Women's Health Initiative Clinical Trial and Observational Study (WHI-CT and OS).

3.2 Study Design and Methods

Study Population

The Women's Health Initiative (WHI) is a long-term national health study that has focused on strategies for preventing heart disease, breast and colorectal cancer, and osteoporotic fractures in postmenopausal women. The original WHI study included 161,808 postmenopausal women enrolled between 1993 and 1998. The WHI has two major parts: a partial factorial randomized clinical trial (CT) and an observational study (OS); both were conducted at 40 Clinical Centers nationwide. The CT enrolled 68,132 postmenopausal women between the ages of 50-79 into trials testing three prevention strategies. The Observational Study (OS) was designed to examine the relationship between lifestyle, environmental, medical and molecular risk factors and specific measures of health or disease outcomes. This component involved tracking the medical history and health habits of 93,676 women not participating in the CT. The Institutional Review Boards

at all the participating institutions approved study protocols and procedures. Each of the participants signed informed consent to participate in the study. We included all 13,547 women in the WHI-CT and WHI-OS whose blood SHBG levels had been measured. Because multiple measurements of the SHBG levels were available for some participants, we randomly selected one observation for each participant. Because this is a cross-sectional study at baseline, all participants with SHBG measurements were included regardless of hormone therapy use in the CT and usage of hormone therapy in the CT was not used as a covariate in this study. Descriptions of the 11 studies that were included in the current study are provided in the Supplementary Table.

Measurements of Plasma Sex Hormone Binding Globulin Levels

For each study participant, blood was collected at the baseline visit after at least a 12-h fast and then stored at −80 °C to -70 °C. Samples used for the hormone measurements were taken from these baseline specimens. The plasma SHBG levels were measured using an electrochemiluminescence immunoassay (ECL), a solid-phase, two-site chemiluminescent immunoassay (solid-phase, two-site CIA), or an immunoradiometric assay (IRMA) (Supplementary Table). The inter-assay coefficients of variation ranged between 3.7% and 17.7%.

Measurements of Demographic factors, Reproductive History, and Lifestyle Factors

Height was measured using a wall-mounted stadiometer, and weight was measured while the participants were wearing light clothing. The body mass index (BMI) was calculated as the weight in kilograms divided by the height in meters, squared. Baseline information on medical history, marital status, and reproductive history was collected via several self-administered

questionnaires. To assess usage of exogenous estrogen and oral contraceptives, a detailed interview was conducted at the baseline visit.

The methods of data collection and validation of dietary factors have been reported previously.[13,14] Participants completed a standardized, self-administered FFQ developed for the WHI to estimate average daily dietary intake over the previous 3-month period. The FFQ was based on instruments used in the WHI feasibility studies[15,16] and the original National Cancer Institute/Block FFQ.[17] The three sections of WHI FFQ included 19 adjustment questions for the calculation of the nutrient content of specific food items, 122 composite and single food line items asking about frequency of consumption and portion size, and 4 summary questions that asked about the usual intake of fruits and vegetables and added fats for comparison with information gathered from the line items. The dietary database, linked to the University of Minnesota Nutrition Coordinating Center Nutrition Data System for Research (Nutrition Coordinating Center, Minneapolis, MN, USA), is based on the U.S. Department of Agriculture standard reference releases and manufacturer information.[18] The construction of FFQs used at the Fred Hutchinson Cancer Research Center has been described in detail elsewhere.[16] This FFQ has demonstrated reasonably good validity as a measurement of dietary intake compared with 24-h dietary recall interviews and food records.[19] The WHI alcohol consumption questionnaires asked whether they ever consumed at least 12 alcoholic drinks of any kind, and those who answered yes were asked whether they still drank alcohol to distinguish never, former, and current drinkers from each other. Ever drinkers were also asked how many alcoholic beverages they consumed each day, week, or month over different ages in their lives. Alcohol intake (grams/day) was calculated from alcohol from beer, wine, and liquor, as well as from foods. The

WHI physical activity questionnaire collected usual frequency, duration, and pace of recreational walking, frequency and duration of other recreational activities or exercises (mild, moderate and strenuous), household, and yard activities.[20] Summary continuous variables estimating weekly energy expenditure (metabolic equivalent [MET]-hours per week) from each type of activity (mild, moderate, strenuous, walking, household, and yard) were calculated from the questionnaire. Summary variables were created by combining frequency, duration and MET-estimated intensity in the following equation: [(Frequency of activity per week × Minutes per session × MET for that activity) / (60 min/hour)].[20] These summary variables in "MET-hours" quantify the total kilocalories expended per kilogram per week. The questionnaire has been reported to have moderate to high test-retest reliability in a diverse sample of post-menopausal women.[20]

Statistical Analysis

Biomarker levels were log-transformed in all the analyses to enhance normality. First, we compared baseline characteristics across quintiles of the plasma SHBG levels with the use of analysis of variance (ANOVA) for continuous variables or chi-squared test for categorical variables (Table 1). Then, we examined the association of each lifestyle factor separately with SHBG levels (Table 2-5). The lifestyle factors included in the analyses were physical activity, alcohol intake, cigarette smoking, intakes of coffee, decaffeinated-coffee, and tea, and categorized food groups. Dietary factors examined were red meat (beef, pork or lamb as a main dish), fish, eggs, soy, fruits, vegetables, whole grains, refined grains, potatoes (French fries, fried potatoes, sweet potatoes, yams, cassava, and yucca), and sweets (chocolate candy, candy bars, ice cream, pudding, custard, flan, low-fat or non-fat frozen desserts, doughnuts, cake, pastry,

Pop-Tarts, pan dulce, cookies, pumpkin pie, and sweet potato pie). After categorizing lifestyle factors into five groups (quintiles for physical activity levels and alcohol consumption and categories for others), we calculated geometric means and standard errors of plasma SHBG levels according to categories of lifestyle factors. *P*-values for linear trend were computed by assigning category score for each category and including the variable as a continuous variable in the models. We also computed *P*-values for quadratic trend by further including a quadratic term in each linear trend model. In model 1, the study-specific geometric means and standard errors of plasma SHBG levels were estimated using a generalized linear model with adjustment for age (continuous), ethnicity (categorical), body mass index (BMI) (continuous), cigarette smoking (never, past, or current), alcohol consumption (Non/past drinkers, <1 drink/week, or >=1 drink/week), physical activity (MET-h/wk) (continuous), and total energy intake (continuous). Then, we pooled estimates using the DerSimonian and Laird random-effects model.[21] In Model 2, we further adjusted for reproductive history: hormone use, oral contraceptive use, age of menarche (categorical, ≥ 13 years old or not), age of first pregnancy (categorical, ≥ 25 years old or not), parity (categorical, ≥ 5 or not), and marital status (categorical, currently married or not).

To jointly examine the association between age, BMI, and lifestyle factors and SHBG levels, we further included these factors in the model simultaneously (Full model) (Table 6). To examine the relative importance of each factor, we estimated study- and ethnicity-specific means and standard errors of log-transformed SHBG levels per 5-year in age, BMI unit (kg/m^2), 7.5 MET-hours/week for physical activity (e.g., brisk walking for 30 minutes, five days a week), 10 cigarettes smoked/day, 10 grams alcohol ingested/day, and one serving/day for each dietary factor with adjustment for total energy intake (continuous), and reproductive history (usage of

exogenous estrogen, usage of oral contraceptives, age of menarche [<13 vs. ≥ 13 years old], age of first pregnancy [<25 vs. ≥ 25 years old], parity [<5 vs. ≥ 5 livebirths], and marital status [currently married vs. not married]). Then, we pooled the results using the DerSimonian and Laird random-effects model. To examine whether the associations between these factors and SHBG levels differed by ethnicity, we also examined the associations between those factors and SHBG levels according to ethnicity. Heterogeneity between studies and ethnicity was evaluated using chi-squared tests and the *I*-squared statistic.[22] To examine sources of heterogeneity between studies included in this study, we further stratified studies according to SHBG assay types (ECL, solid-phase, two-site CIA, or IRMA). We also stratified studies according to study outcomes (cardiovascular diseases or type 2 diabetes, cancer, fractures, and others). To provide a visual representation of the dose-response curve, we fitted unrestricted quadratic spline models by including transformed variables of each factor to the regression models (with knots at the 20th, 40th, 60th, and 80th percentiles).[23] Also, because SHBG synthesis in the liver is also considered to be regulated by sex hormones[24-26], we further stratified by usage of exogenous estrogen to examine whether the associations still remain among never users of exogenous estrogen. Finally, to examine the associations between exogenous estrogen and reproductive history and SHBG levels, we estimated pooled means and standard errors of log-transformed SHBG levels according to usage of exogenous estrogen, usage of oral contraceptives, years since menopause (per 5-year), menarche age (≥13 vs. <13), number of parities (≥ 5 vs. <5), marital status (currently married vs. not married) using the same model (Full model) (Table 7). Because the results in Native American women were quite imprecise due to small number, we did not present

the results for Native American women. We conducted statistical analyses using SAS (version 9.3; SAS institute, Cary, NC) and STATA (version 11.2; StataCorp, College Station, TX).

3.3 Results

The baseline characteristics according to quintiles of plasma SHBG levels are presented in Table1. Compared to women with lower levels of plasma SHBG, women with higher levels of plasma SHBG tended to be older, African-Americans or Asian/Pacific Islanders, have low adiposity measures, be current smokers, engage in physical activity, and have higher education levels, but less likely to have a history of diabetes or hypertension. Also, they tended to be past users of postmenopausal hormone and never users of oral contraceptives, had more years since menopause, and underwent menarche at earlier age.

Table 2 shows the relations of physical activity levels, alcohol consumption, and cigarette smoking with plasma SHBG levels. After multivariable adjustment for age, ethnicity, BMI, cigarette smoking, alcohol consumption, and total energy intake (Model 1), physical activity levels were positively associated with plasma levels of SHBG. Further adjustment for reproductive history (Model 2) or restriction to the control subjects only (data not shown) resulted in similar findings. In Model 2, compared with women in the lowest quintile of physical activity levels, women in the highest quintile had 5.3 % higher SHBG levels (the adjusted geometric mean levels of SHBG [95% confidence interval]: 46.6(41.8-52.0) in the lowest quintile; 49.1(45.1-53.5) in the highest quintile; *P* for linear trend < 0.001; *P* for quadratic trend =0.95) (Table 2). After multivariable adjustment (Model 1-2), we did not observe a clear linear association between alcohol consumption and SHBG levels (*P* for linear trend = 0.27), although

a suggestion of a nonlinear association was seen (*P* for quadratic trend = 0.04) (Table 2). While we did not observe association between years of smoking and SHBG levels, there appeared to be a weak positive association between cigarettes per day and SHBG levels (Table 2). After multivariable adjustment (Model 2), compared to never or past smokers, women who smoked 35 or more cigarettes per day had 2.4 % higher SHBG levels (*P* for linear trend = 0.002; *P* for quadratic trend = 0.002) (Table 2). Restricting the analysis to the controls only resulted in similar findings (data not shown).

There was a positive association between regular coffee intake and SHBG levels, whereas decaffeinated-coffee and tea intakes were not associated with SHBG levels (Table 3). After multivariable adjustment for age, ethnicity, BMI, physical activity, cigarette smoking, alcohol consumption, total energy, and reproductive history (Model 2), compared with women who drank no cups of coffee per day, women who drank 6 or more cups per day had 7.4 % higher SHBG levels (*P* for linear trend <0.001; *P* for quadratic trend = 0.002) (Table 3). Further restricting the analysis to the controls only did not materially change the results (data not shown).

Associations of red meat, fish, egg, dairy, and soy intakes with SHBG levels are presented in Table 4. After multivariable adjustment for age, ethnicity, BMI, physical activity, cigarette smoking, alcohol consumption, total energy, and reproductive history (Model 2), a weak inverse association between red meat intake and SHBG levels (Table 4). Compared to non-consumers of red meat, consumers of one or more per day of red meat had 3.1 % lower levels of SHBG (Table 4). Further restriction to the control subjects only resulted in similar results. After multivariable adjustment (Model 2), consumers of one or more servings per day of egg had 5.1 % lower levels of SHBG than non-consumers (Table 4). Similar results were seen when we further restriction to

control subjects only (data not shown). Compared with non-consumers, consumers of one or more servings per day of soy had 2.3 % higher levels of SHBG after multivariable adjustment (Model 2) (Table 4). Further restriction to the control subjects only resulted in similar findings. Conversely, fish consumption and dairy consumption were not associated with SHBG levels (Table 4).

We did not observe an association between intakes of fruits, vegetable, whole grains, refined grains, and potatoes, and SHBG levels after multivariable adjustment for age, ethnicity, BMI, physical activity, cigarette smoking, alcohol consumption, total energy, and reproductive history (Table 5). In contrast, a moderate positive association was seen between sweets and SHBG levels (Table 5). Women who consumed three or more servings per day of sweets had 10.6 % higher levels of SHBG than women who consumed less than one serving per week. Similar results were seen when the analysis was further restricted to only the control subjects (data not shown).

To jointly examine the association between lifestyle factors and SHBG levels, dietary factors and other factors were simultaneously included in the models (Full model) (Table 6). Age, physical activity, and regular coffee intake were positively associated with SHBG levels, whereas BMI was inversely associated with SHBG levels. Age was moderately and positively associated with SHBG levels; the pooled adjusted mean difference in log-SHBG levels per 5-year increment in age was 0.037 (= 0.066 log-transformed standard deviation [SD]) (95% confidence interval in log-SHBG levels = 0.026 to 0.047). Stratification according to the usage of exogenous estrogen indicated that age was positively associated with SHBG levels among both ever users and never users of exogenous estrogen (mean difference in log-SHBG levels

among ever users= 0.024 [= 0.043 SD], 95% confidence interval = 0.003 to 0.045; mean difference among never users=0.045 [= 0.080 SD], 95% confidence interval = 0.029 to 0.060) (data not shown). No substantial heterogeneity was seen across the studies (Supplementary Figure 1A), nor was heterogeneity seen across ethnicities (*P*-value for ethnic difference=0.16) (data not shown). Spline curves indicated that an age of 60 years or older was associated with higher SHBG levels (Figure 1A). BMI was inversely associated with SHBG levels (the pooled adjusted mean difference in log-SHBG levels per one unit increase in BMI=-0.028 [= -0.050 SD], 95% confidence interval= -0.033 to -0.024). Similar patterns were observed for association of waist circumference and SHBG levels (data not shown). BMI was inversely associated with SHBG levels among both ever users and never users of exogenous estrogen (mean difference in log-SHBG levels among ever users= -0.035 [= -0.062 SD], 95% confidence interval= -0.042 to -0.028; mean difference among never users = -0.027 [= -0.048 SD], 95% confidence interval=-0.032 to -0.022) (data not shown). Although the direction of the association was largely consistent across studies, with exception of W9 and AS110 in African-Americans, heterogeneity was suggested across the studies (Supplementary Figure 1B) and across ethnicities (*I*-squared for ethnic difference=77.7%; *P*-value for ethnic difference =0.004) (data not shown). Stratification according to the SHBG assay types or according to the study outcomes did not explain the heterogeneity across the studies. Spline curves indicated an inverse association of BMI with SHBG levels especially at BMI below 30 (Figure 1B). Physical activity levels were positively associated with SHBG levels (the pooled adjusted mean difference in log-SHBG levels per increment of 7.5 MET-hours/week = 0.017 [= 0.030 SD]; 95% confidence interval= 0.008 to 0.027). Stratification according to the usage of exogenous estrogen also indicated that physical

activity levels were positively associated with SHBG levels (mean difference in log-SHBG levels among ever users = 0.032 [=0.057 SD], 95% confidence interval=0.013 to 0.050; mean difference among never users=0.019 [=0.034 SD]; 95% confidence interval=0.003 to 0.035) (data not shown). No suggestion of ethnic differences was seen (*P*-value for ethnic difference =0.72) (data not shown), but heterogeneity across the studies, especially among African-Americans, was suggested (Supplementary Figure 1C). Stratification according to the study outcome or the SHBG assay types did not explain the heterogeneity across studies. Spline curves showed a linear relation between physical activity levels measured in MET-h/week and SHBG levels (Figure 1C). Regular coffee intake was moderately and positively associated with SHBG levels (the pooled adjusted mean difference in log-SHBG levels per cup/day=0.027 [= 0.052 SD], 95% confidence interval= 0.013 to 0.046). Further stratification according to the usage of exogenous estrogen showed similar patterns among both ever users and never users of exogenous estrogen (mean difference in log-SHBG levels among ever users= 0.024 [=0.043 SD], 95% confidence interval=0.010 to 0.037; mean difference among never users= 0.027 [=0.048 SD], 95% confidence interval=0.014 to 0.039) (data not shown). We did not observe any heterogeneity across the studies (Supplementary Figure 1D) or ethnicities (*I*-squared for ethnic difference=0.0%; *P*-value for ethnic difference =0.59) (data not shown). Spline curves suggested that that heavy drinkers of regular coffee (>2 cups/day) had higher levels of plasma SHBG (Figure 1D). Although Model 2 suggested associations of number of cigarettes smoked per day and intakes of red meat, eggs, soy, and sweets with SHBG levels, the Full model did not indicate any apparent associations between these factors and SHBG levels (Table 6). Spline curves for cigarettes smoking showed no association between number of cigarettes smoked per day and

SHBG levels for less than 40 cigarettes per day, and an increasing trend observed for more than 60 cigarettes per day is not a statistically stable feature (Supplementary Figure 2A). Also, Models 2 suggested a possible nonlinear association between alcohol intake and SHBG levels. The Full model, however, did not indicate a clear linear or nonlinear relation between alcohol intake and SHBG levels (Table 6 and Supplementary Figure 2B).

Finally, we examined the associations between exogenous estrogen and reproductive history and SHBG levels with adjustment for covariates included in the Full model (Table 7). The usage of exogenous estrogen and marital status were positively associated with SHBG levels, whilst other factors were not apparently associated with SHBG levels after mutual adjustment for age, BMI, lifestyle factors, and total energy intake. Current estrogen users had higher SHBG levels than past or never users of exogenous estrogen (mean difference in log-SHBG, 0.225 [= 0.401 SD]; 95% confidence interval, 0.043 to 0.408). Heterogeneity was suggested across the studies (Supplementary Figure 1E) and ethnicities (P-value for ethnic difference <0.001) (data not shown), but stratification according to SHBG assay types or according to study outcomes did not explain the heterogeneity across studies. Also, compared with never users of exogenous estrogen, ever users had higher SHBG levels (mean difference in log-SHBG= 0.083 [= 0.149 SD], 95% CI= 0.012 to 0.153). In addition, compared with women who never or past married, women who were currently married had lower levels of SHBG (mean difference in log-SHBG, -0.033 [= -0.059 SD], 95% CI= -0.062 to -0.003). In contrast, we did not find any associations between usage of oral contraceptives, years since menopause, parities, and age of menarche with SHBG levels (Table 7).

3.4 Discussion

In this large-scale multiethnic study among postmenopausal women consisting of 11 studies nested in the Women's Health Initiative, positive associations with SHBG levels were seen for age, use of exogenous estrogen, physical activity, and regular coffee intake, whereas BMI was inversely associated with SHBG levels. Our comprehensive analyses in the large-scale multiethnic study support the notion that in addition to sex hormones, age, adiposity, physical activity, and regular coffee intake could also conceivably play roles in the regulation of SHBG synthesis.

The positive association between age and SHBG levels observed in this study is in line with earlier studies.[27,28] Several mechanisms have been proposed to explain this positive association of age with SHBG levels. One explanation is that SHBG levels are elevated with aging secondary to the decline in sex hormone production.[29] However, because oral estrogen is known to increase SHBG levels, the age-related decline in estrogen levels in women might not result in an elevation of the SHBG levels. The second explanation is that the age-related decrease in GH and IGF-I levels might lead to the increased SHBG levels.[30] Nevertheless, whether the decline in the GH and IGF- I levels alone are responsible for the apparent age-associated increase in SHBG levels remains unclear. The third explanation may be that the age-related decline in insulin secretion or insulin sensitivity[31] may result in the elevation of SHBG levels. However, plasma SHBG was recently shown to be significantly increased, rather than decreased, after insulin treatment in diabetic patients.[32] Therefore, no clear conclusion can be reached regarding the mechanisms underlying the age-associated increase in SHBG levels.

As expected, our findings indicated that usage of exogenous sex hormones was positively associated with SHBG levels. This observation agrees well with the findings from *in-vitro* experiments that SHBG synthesis in the liver is partly regulated by sex hormones.[24-26] Previous cross-sectional studies have also reported that the use of oral estrogen among postmenopausal women was associated with 2-fold greater levels of SHBG.[33] Further, the usage of oral contraceptives has been reported to be associated with a 2-fold increase in SHBG levels among premenopausal women.[34] Our results, however, did not indicate a relation between the usage of oral contraceptives and SHBG levels among postmenopausal women. This finding may indicate that the past use of oral contraceptives might have little impact on SHBG levels, whereas the use of exogenous estrogen perhaps influence the regulation of SHBG in the liver among postmenopausal women. In addition, our findings indicated that currently married women had lower SHBG levels than single women, whereas an earlier study reported that married nulliparous postmenopausal women had higher SHBG levels than single nulliparous postmenopausal women.[35] It was difficult to examine the association among nulliparous women in the current study because there were few nulliparous women. Thus, future work is needed to better understand the relationship. Also, we did not find any association between number of parities or age of menarche and SHBG levels. These findings are consistent with previous studies.[35] Number of parities and age of menarche were associated with SHBG levels among premenopausal but not among postmenopausal women. Although previous research reported that the number of years since menopause was positively associated with the SHBG levels,[35] there was no association between years since menopause and SHBG levels after multivariable adjustment for potential confounding factors, including age. However, this finding is not

particularly surprising in light of the fact that age and years since menopause are highly correlated (r=0.71 in our study). In fact, we observed a positive association between age and SHBG levels even after adjustment for the number of years since menopause (data not shown). This finding may suggest that age, but not years since menopause, might be an important predictor of SHBG levels, although further studies are needed to determine the relative importance of these two factors.

Physical activity levels were positively associated with higher levels of SHBG and the positive association in all ethnic groups. In a cross-sectional study of 400 men, physical activity levels were positively associated with plasma SHBG levels after adjustment for age, BMI, smoking status, and alcohol intake.[28] A randomized clinical trial of 102 men also showed a greater increase in SHBG levels among exercisers, compared with a control groups (increase at 3 months from SHBG levels at baseline=14.3% among exercisers vs. 5.7% among control subjects, P-value=0.04).[36] However, little is known about the mechanisms underlying the possible effect of exercise on SHBG metabolism. One explanation is that improved insulin sensitivity induced by physical activity might influence SHBG production in the liver. Another explanation is that loss of adiposity might mediate the potential effect of physical activity on SHBG levels.

We observed an inverse association between BMI and SHBG levels and similar patterns were observed in the association between waist circumference and SHBG levels. The inverse association observed in this study is consistent with previous studies.[37] The underlying mechanisms responsible for the inverse association of BMI with SHBG levels are not well understood. Because adiposity is associated with higher blood insulin levels and it has been reported that insulin inhibits SHBG production in hepatoma cells,[26] hyperinsulinemia in women

with higher BMI levels could lead to lower SHBG levels. Meanwhile, as mentioned earlier, a recent study showed that insulin treatment in human resulted in higher SHBG levels.[32] Therefore, another explanation is that hyperinsulinemia might merely represent a compensatory increase in insulin levels secondary to insulin resistance in obese people, and insulin resistance itself might play a major role in the regulation of SHBG levels.

Regular coffee intake was positively associated with SHBG levels. In contrast, we observed no associations between decaffeinated-coffee and tea intakes with SHBG levels. These findings are in line with earlier studies examining association between coffee intake and SHBG levels.[38,39] One recent study found that caffeinated-coffee and caffeine intakes, but not decaffeinated-coffee intake, were positively associated with circulating levels of SHBG.[39] Because there were few decaffeinated coffee or tea drinkers in our study, we lacked statistical power to detect an association with SHBG levels. Molecular mechanisms responsible for the relation between regular coffee consumption and SHBG levels are largely unknown. Caffeine and other major components of coffee (cafestol and kahweol) alter the expressions and activities of liver enzymes.[40-43] Because SHBG is synthesized and metabolized primarily in the liver,[1] coffee intake may affect SHBG metabolism in the liver and may influence the plasma levels of SHBG.[38] Importantly, several studies have shown that caffeine increases the resting metabolic rate.[44] Also, SHBG production is regulated by the metabolic state of the liver via hepatocyte nuclear factor-4α (HNF-4α).[45] Taken together, caffeine might influence SHBG production by altering the HNF-4α levels in response to changes in the metabolic rate in the liver.

No apparent association was seen between soy intake and SHBG levels in this study (Full Model) (Table 6). Our results are consistent with findings from a recent meta-analysis[46] that

included 47 randomized or residential crossover trials of soy or isoflavones for 4 or more weeks. The summary mean differences in SHBG levels (nmol/L) (soy or isoflavones groups vs. control groups) were -2.19 (95% confidence interval= -6.37 to 1.99) in premenopausal women and -0.87 (95% confidence interval= -3.52 to 1.78) in postmenopausal women.[46] Therefore, these findings suggest that soy or isoflavones intake might not play a major role in the metabolism of SHBG.

The strengths of our study include its large sample size and multiethnic study design with detailed measures of variables. Several issues regarding our study design need to be recognized. First, the major limitation of this study is its inability to establish a temporal relationship between exposure (e.g., lifestyle factors) and SHBG levels, although it is unlikely that SHBG levels alter the exposure. Second, the studies pooled in the current study differed in the SHBG measurement methods, the laboratories where the measurements were performed, and the study outcomes that were selected, which might have introduced sources of heterogeneity. Third, we cannot exclude the possibility of residual confounding arising from incompletely measured or unmeasured covariates. Fourth, biases due to measurement errors in exposure variables, covariates, and SHBG levels may exist and may vary across the studies. However, because such biases are likely to be non-differential, the estimates would tend to attenuate toward the null. Finally, although our study population was well-characterized and nested within the WHI, the possibility of selection bias cannot be ruled out.

In conclusion, this large-scale multiethnic study among postmenopausal women consisting of 11 studies include in the Women's Health Initiative found that age, usage of exogenous estrogen, physical activity, and regular coffee intake were positively associated with SHBG levels, whereas BMI was inversely associated with SHBG levels. Because these factors are risk

factors (or protective factors) for cardiovascular diseases, type 2 diabetes, cancers, and fractures, the findings may have important implications for various diseases. Further studies are needed in both experimental and observational settings.

Table 1. Baseline characteristics according to quintiles of plasma sex hormone-binding globulin levels

	Q1 (lowest)	Q2	Q3	Q4	Q5(Highest)	*P*-values*
	(n=2756)	(n=2668)	(n=2711)	(n=2707)	(n=2705)	
Plasma SHBG levels (nmol/L)	22.9 (18.7-26.0)	34.4 (31.8-36.9)	45.1 (42.1-48.2)	60.0 (55.8-65.4)	93.2 (80.0-122.0)	
Age (years)	63.3 ± 7.0	65.1 ± 7.3	65.7 ± 7.3	66.5 ± 7.4	65.9 ± 7.8	< 0.001
Race (% European-Americans)	67.0	68.9	70.2	71.4	61.1	
Race (% African-Americans)	19.3	18.3	17.7	16.5	22.5	
Race (% Hispanics)	8.0	7.3	6.5	7.0	8.3	< 0.001
Race (% Asian/Pacific Islanders)	3.8	3.7	3.7	3.7	7.0	
Race (% Native Americans)	1.3	1.0	1.3	0.7	0.5	
Body Mass Index (kg/m^2)	32.3 ± 6.1	30.4 ± 5.9	28.9 ± 5.9	27.1 ± 5.5	25.8 ± 5.5	< 0.001
Waist circumference(cm)	97.7 ± 12.5	93.0 ± 12.9	89.0 ± 12.6	84.1 ± 12.1	80.7 ± 12.2	< 0.001
Alcohol consumption (grams/day)	0.06 (0.01-1.98)	0.12 (0.01-3.52)	0.50 (0.01-4.60)	0.51 (0.01-6.00)	0.12 (0.01-2.76)	< 0.001
Current smoking (%)	6.9	8.0	8.7	9.7	10.7	< 0.001
Physical activity (total MET-hours/week)	4.5 (0.5-12.1)	5.3 (1.3-14.0)	7.5 (1.5-16.3)	8.3 (2.5-17.8)	8.8 (2.5-19.2)	< 0.001
Past postmenopausal hormone use (%)	27.0	30.2	29.4	31.3	40.5	< 0.001
Ever oral contraceptive use (%)	40.1	34.9	32.2	31.5	30.6	< 0.001
Age at menopause	46.6 ± 6.8	46.9 ± 6.7	47.0 ± 6.7	47.6 ± 6.5	47.0 ± 6.5	0.009
Years since menopause	16.9 ± 9.0	18.5 ± 9.3	19.0 ± 9.4	19.2 ± 9.1	19.1 ± 9.4	< 0.001
Age at menarche, <12 (%)	50.6	49.9	51.9	54.4	56.2	< 0.001
Age at first pregnancy of ≥6 months, <20 (%)	24.7	23.8	21.3	20.3	22.3	0.003
History of hypertension (%)	49.0	42.1	39.1	34.8	33.1	< 0.001
Pregnancies, ≥5 (%)	20.9	20.8	19.6	17.5	16.9	< 0.001
Marital status (% currently married)	56.5	55.1	51.8	52.7	51.3	< 0.001
Education (% college graduate)	30.5	30.7	31.2	35.6	36.0	< 0.001
Total estradiol (pg/mL)	12.6 (8.8-18.1) (n=2730)	11.4 (8.0-16.5) (n=2625)	10.3 (7.3-15.0) (n=2666)	9.4 (6.8-14.0) (n=2641)	10.1 (6.8-21.0) (n=2584)	< 0.001
Total testosterone (ng/dL)	18.3 (12.1-25.5) (n=1464)	20.2 (13.1-28.4) (n=1414)	20.5 (13.2-29.8) (n=1530)	21.2 (13.4-31.5) (n=1575)	16.0 (0.1-28.0) (n=1803)	< 0.001

Abbreviations: SHBG, sex hormone-binding globulin; MET, metabolic equivalent.
Unless otherwise indicated, values are means ± SD or median (interquartile range)
*Baseline characteristics were compared between groups using analysis of variance (ANOVA) for continuous variables or chi-squared test for categorical variables.

Table 2. Adjusted geometric mean plasma levels of SHBG according to categories of physical activity, alcohol, and smoking categories

Plasma SHBG levels (nmol/L)	Categories (Lowest)				(Highest)	P for liner trend*	P for quadratic trend*
Physical activity	Q1	Q2	Q3	Q4	Q5		
Median (MET-h/w) (n)	0 (2544)	2.5 (2540)	7.0 (2665)	13.7 (2409)	27.8 (2538)		
Range (MET-h/w)	0-0.5	0.8-4.2	4.2-10.0	10.0-18.8	18.8-134.2		
Model 1[1]	46.7(45.5-47.9)	45.3(44.1-46.5)	46.9(47.0-49.5)	48.2(47.0-49.5)	50.3(49.0-51.6)	<0.001	0.89
Model 2[2]	46.6(41.8-52.0)	44.9(41.6-48.4)	46.7(42.4-51.5)	47.7(43.1-52.6)	49.1(45.1-53.5)	<0.001	0.95
Alcohol	Q1	Q2	Q3	Q4	Q5		
Median(g/day) (n)	0.00 (2708)	0.02 (2699)	0.13 (2701)	1.99 (2703)	13.9 (2702)		
Range	0.00-0.01	0.01-0.03	0.03-0.99	0.99-6.49	6.49-191.9		
Model 1[1]	45.8(41.8-50.2)	46.5(42.2-51.3)	47.3(43.1-51.8)	48.5(44.2-53.3)	46.5(43.0-50.3)	0.33	0.04
Model 2[2]	45.8(41.5-50.6)	46.4(41.9-51.3)	47.2(42.9-52.0)	48.3(44.0-53.1)	46.4(43.0-50.0)	0.27	0.04
Years of smoking	Never	<10 yrs	10-29 yrs	30-49 yrs	50- yrs		
(n)	(6997)	(1301)	(2545)	(2048)	(301)		
Model 1[1]	47.0(42.9-51.5)	47.8(44.3-51.5)	47.5(43.2-52.2)	46.4(42.6-50.5)	45.6(41.5-50.2)	0.29	0.48
Model 2[2]	47.0(42.8-51.6)	48.0(44.1-52.1)	47.1(42.7-52.0)	46.5(42.6-50.7)	46.3(41.6-51.6)	0.35	0.67
Cigarettes per day	Never or Past	<5	5-24	25-34	35-		
(n)	(4918)	(196)	(786)	(127)	(63)		
Model 1[1]	46.5(42.6-50.8)	44.8(41.6-48.3)	49.1(45.2-53.3)	50.8(45.7-56.6)	41.8(37.1-47.2)	0.002	0.001
Model 2[2]	46.5(42.5-50.8)	45.5(42.2-49.0)	49.6(45.1-54.4)	50.4(46.2-55.0)	47.6(45.3-50.0)	0.002	0.002

Data are expressed as adjusted geometric mean (95% confidence interval).
Abbreviation: SHBG, sex hormone-binding globulin; m, month; w, week; d, day; yrs, years; MET-h/w, metabolic equivalent (MET)-hours per week; Q1, quintile 1; Q2; quintile 2; Q3; quintile 3; Q4, quintile 4; Q5, quintile 5.
* Based on median values in categories of the participants with exception for smoking. For increasing categories of years of smoking, 0, 5, 20, 40, and 60 years of smoking were assigned, and for increasing categories of cigarettes per day, 0, 1, 15, 30, and 45 cigarettes per day were assigned to compute *P*-values for trend.
[1] Model 1: Geometric mean and confidence interval of plasma SHBG levels, and *P*-values for linear trend and quadratic trend were estimated using random-effects model with adjustment for age (continuous), ethnicity (categorical), body mass index (BMI) (continuous), cigarette smoking (never, past, or current), alcohol consumption (none/past drinkers, <1 drink/week, or >=1 drink/week), physical activity (MET-h/wk) (continuous), and total energy intake (continuous) that were not examined as the primary independent variable.
[2] Model 2: Further adjusted for reproductive history: usage of exogenous estrogen, usage of oral contraceptives, age of menarche (<13 vs. ≥ 13 years old), age of first pregnancy (<25 vs. ≥ 25 years old), parity (<5 vs. ≥ 5 livebirths), and marital status (currently married vs. not married).

Table 3. Adjusted geometric mean plasma levels of SHBG according to categories of regular coffee, decaffeinated-coffee, and tea intakes

Plasma SHBG levels (nmol/L)	Categories (Lowest)				(Highest)	*P* for linear Trend*	*P* for quadratic trend*
Regular coffee (n)	(1646)	(2135)	(4175)	(1245)	(393)		
Range (cups/d)	0	1	2-3	4-5	≥6		
Model 1[1]	47.7(43.3-52.6)	46.2(41.9-50.9)	47.2(43.7-51.0)	50.4(46.3-54.8)	52.8(47.7-58.5)	<0.001	0.002
Model 2[2]	47.4(43.0-52.3)	45.9(41.5-50.8)	47.3(43.6-51.3)	50.9(46.6-55.5)	53.1(47.8-58.9)	<0.001	0.002
Decaffeinated coffee (n)	(2215)	(859)	(645)	(121)	(22)		
Range (cups/d)	0	1	2-3	4-5	≥6		
Model 1[1]	51.5(44.3-60.0)	50.9(43.2-60.1)	52.0(45.2-59.9)	49.4(43.4-56.3)	51.6(41.4-64.3)	0.62	0.82
Model 2[2]	51.5(44.0-60.4)	50.8(42.5-60.6)	51.8(44.9-59.7)	49.3(43.4-55.9)	53.2(42.4-66.7)	0.64	0.57
Tea (n)	(4444)	(706)	(581)	(98)	(28)		
Range (cups/d)	0	1	2-3	4-5	≥6		
Model 1[1]	51.1(43.5-59.9)	49.5(43.9-55.8)	50.2(44.6-56.5)	52.6(42.7-64.8)	51.1(41.2-63.3)	0.70	0.47
Model 2[2]	50.9(43.2-60.0)	49.9(43.5-57.2)	50.2(44.2-56.9)	54.6(42.6-70.1)	51.5(41.8-63.3)	0.46	0.44

Data are expressed as adjusted geometric mean (95% confidence interval).
Abbreviation: SHBG, sex hormone-binding globulin; d, day; MET-h/w, metabolic equivalent (MET)-hours per week.
* For increasing categories, 0, 1, 2.5, 4.5, and 6 cups/day were assigned to compute *P*-values for trend.
1 Model 1: Geometric mean and confidence interval of plasma SHBG levels, and *P*-values for linear trend and quadratic trend were estimated using random-effects model with adjustment for age (continuous), ethnicity (categorical), body mass index (BMI) (continuous), cigarette smoking (never, past, or current), alcohol consumption (none/past drinkers, <1 drink/week, or >=1 drink/week), physical activity (MET-h/wk) (continuous), and total energy intake (continuous).
2 Model 2: Further adjusted for reproductive history: usage of exogenous estrogen, usage of oral contraceptives, age of menarche (<13 vs. ≥ 13 years old), age of first pregnancy (<25 vs. ≥ 25 years old), parity (<5 vs. ≥ 5 livebirths), and marital status (currently married vs. not married).

Table 4. Adjusted geometric mean plasma levels of SHBG according to categories of red meat, fish, egg, dairy, and soy intakes

Plasma SHBG levels (nmol/L)	Categories (Lowest)				(Highest)	*P* for linear trend*	*P* for quadratic trend*
Red meat							
Median(s/w)(n)	0 (1974)	0.6 (7588)	2.0 (2027)	3.4 (1718)	7.6 (206)		
Range (s/w)	0	<1	1-2	2-7	≥ 7		
Model 2[1]	48.1(44.6-53.4)	47.4(43.3-51.9)	50.4(46.2-42.4)	44.7(40.3-49.6)	46.6(39.3-55.2)	0.04	0.03
Fish							
Median(s/w) (n)	0 (725)	0.6 (4748)	1.5 (3825)	3.0 (3925)	8.5 (290)		
Range (s/w)	0	<1	1-2	2-7	≥ 7		
Model 2[1]	46.9(42.3-51.8)	47.1(43.2-51.4)	47.2(42.6-52.2)	46.9(42.8-51.3)	48.4(43.1-54.5)	0.44	0.14
Egg							
Median(s/w) (n)	0 (2368)	0.5 (7919)	2.0 (2042)	3.4 (1074)	7.6(110)		
Range (s/w)	0	<1	1-2	2-7	≥ 7		
Model 2[1]	47.3(42.7-52.3)	47.2(43.3-51.5)	47.0(42.2-52.4)	45.9(41.5-50.7)	44.9(38.9-51.8)	0.03	0.57
Dairy							
Median(s/d) (n)	0.6 (419)	1.5 (590)	4.6 (3750)	12.0(6830)	26.7(1951)		
Range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	51.2(44.3-59.2)	48.2(43.0-53.9)	47.1(42.6-52.2)	47.3(43.5-51.3)	45.5(41.3-50.0)	0.11	0.56
Soy							
Median(s/w) (n)	0 (11397)	0.2 (1526)	2.0 (247)	3.4 (268)	7.6 (75)		
Range (s/w)	0	<1	1-2	2-7	≥ 7		
Model 2[1]	46.8(43.0-50.9)	47.4(42.7-52.6)	49.4(43.0-56.8)	53.4(46.2-61.7)	47.9(39.2-58.5)	0.007	0.18

Data are expressed as adjusted geometric mean (95% confidence interval).

Abbreviation: SHBG, sex hormone-binding globulin; s/w, serving/week; *P*, *P*-value.

* Based on median values in categories of the participants.

[1] Model 2: Geometric mean and confidence interval of plasma SHBG levels, and *P*-values for linear trend and quadratic trend were estimated using random-effects model with adjustment for age (continuous), ethnicity (categorical), body mass index (BMI) (continuous), cigarette smoking (never, past, or current), alcohol consumption (none/past drinkers, <1 drink/week, or >=1 drink/week), physical activity (metabolic equivalent (MET)-hours per week) (continuous), total energy intake (continuous), and reproductive history: usage of exogenous estrogen, usage of oral contraceptives, age of menarche (<13 vs. ≥ 13 years old), age of first pregnancy (<25 vs. ≥ 25 years old), parity (<5 vs. ≥ 5 livebirths), and marital status (currently married vs. not married).

Table 5. Adjusted geometric mean plasma levels of SHBG according to categories of fruits, vegetables, whole grains, refined grains, potatoes, and sweets intakes

Plasma SHBG levels (nmol/L)	Categories (Lowest)				(Highest)	*P* for linear trend*	*P* for quadratic trend*
Fruits,							
median(s/w) (n)	0.2 (407)	1.7 (483)	4.8 (2801)	12.2 (7105)	24.5 (2717)		
range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	46.0(41.1-51.5)	47.5(41.9-53.9)	46.7(42.4-51.4)	47.0(43.1-51.3)	47.6(43.1-52.5)	0.39	0.51
Vegetables,							
median(s/w) (n)	0.6 (119)	1.5 (111)	4.8 (2461)	12.5 (7982)	26.3 (2840)		
range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	48.8(42.0-56.7)	51.1(41.6-62.8)	46.8(42.3-51.9)	46.8(42.8-51.2)	47.8(43.8-52.2)	0.09	0.10
Whole grains							
median(s/w) (n)	0.6 (565)	1.5 (925)	4.5 (5469)	10.5 (6011)	24.9 (543)		
range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	46.3(42.3-50.8)	47.7(42.8-53.2)	46.7(42.6-51.2)	47.1(43.0-51.7)	47.3(43.0-52.1)	0.41	0.90
Refined grains							
median(s/w) (n)	0.6 (36)	1.6 (61)	5.1 (1046)	14.4 (6240)	30.3 (6130)		
range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	44.2(36.9-53.0)	46.0(39.9-53.0)	48.3(42.3-55.1)	47.4(43.3-51.9)	46.6(42.8-50.7)	0.58	0.80
Potatoes,							
median(s/w) (n)	0 (643)	0.6 (4032)	1.5 (3048)	3.6 (5168)	8.3 (622)		
range (s/w)	0	<1	1-2	2-7	≥ 7		
Model 2[1]	46.8(41.3-52.9)	47.3(43.1-51.9)	46.7(42.5-51.3)	47.0(43.0-51.4)	47.4(43.1-52.0)	0.57	0.23
Sweets,							
median(s/w) (n)	0.5 (1902)	1.5 (1628)	4.1 (5123)	10.5 (4248)	26.5 (612)		
range (s/w)	<1	1-2	2-7	7-21	≥ 21		
Model 2[1]	45.0(40.5-50.0)	46.0(41.3-51.1)	46.3(42.7-50.3)	48.5(44.0-53.5)	49.8(44.8-55.4)	0.001	0.02

Data are expressed as adjusted geometric mean (95% confidence interval).
Abbreviation: SHBG, sex hormone-binding globulin; s/w, serving/week; *P*, *P*-value.
* Based on median values in categories of the participants.
[1] Model 2: Geometric mean and confidence interval of plasma SHBG levels, and *P*-values for linear trend and quadratic trend were estimated using random-effects model with adjustment for age (continuous), ethnicity (categorical), body mass index (BMI) (continuous), cigarette smoking (never, past, or current), alcohol consumption (none/past drinkers, <1 drink/week, or >=1 drink/week), physical activity (metabolic equivalent (MET)-hours per week) (continuous), total energy intake (continuous), and reproductive history: usage of exogenous estrogen, usage of oral contraceptives, age of menarche (<13 vs. ≥ 13 years old), age of first pregnancy (<25 vs. ≥ 25 years old), parity (<5 vs. ≥ 5 livebirths), and marital status (currently married vs. not married).

Table 6. Adjusted mean difference in log-transformed plasma SHBG levels in relation to age, BMI, and lifestyle factors.

log-SHBG levels*	Pooled (n=8712)
Age (per 5-year)	0.037(0.026 to 0.047)
BMI (per unit)	-0.028(-0.033 to -0.024)
Physical activity (per 7.5 MET-h/w)	0.017(0.008 to 0.027)
Alcohol (per 10 grams/d)	-0.012(-0.022 to -0.002)
Cigarettes (per 10/d)	-0.011(-0.023 to -0.000)
Regular Coffee (per cup/d)	0.029 (0.013 to 0.046)
Red meat (per s/d)	-0.201(-0.592 to 0.190)
Fish (per s/d)	0.028 (-0.020 to 0.076)
Eggs (per s/d)	-0.060(-0.130 to 0.010)
Dairy (per s/d)	-0.017(-0.028 to -0.006)
Soy (per s/d)	-0.017(-0.112 to 0.078)
Fruits (per s/d)	-0.002(-0.017 to 0.012)
Vegetables (per s/d)	0.002(-0.010 to 0.014)
Whole grains (per s/d)	0.002(-0.014 to 0.017)
Refined grains (per s/d)	-0.009(-0.017 to -0.000)
Potatoes (per s/d)	-0.015(-0.053 to 0.024)
Sweets (per s/d)	0.013(-0.002 to 0.027)

Abbreviation: SHBG, sex hormone-binding globulin; log-SHBG levels, log-transformed SHBG; s/d, serving/day; cup/d, cup/day; MET-h/w, metabolic equivalent [MET]-hours per week.
Data are expressed as adjusted mean (95% confidence interval).
Adjusted means and confidence intervals of log-transformed SHBG levels per 5-year in age, BMI unit, 7.5 MET-hours/week for physical activity, 10 cigarettes smoked/day, 10 grams of alcohol ingested/day, and one serving/day for dietary factors were computed using the random-effects model with adjustment for total energy intake (continuous), and reproductive history (usage of exogenous estrogen, usage of oral contraceptives, age of menarche (<13 vs. ≥ 13 years old), age of first pregnancy (<25 vs. ≥ 25 years old), parity (<5 vs. ≥ 5 livebirths), and marital status (currently married vs. not married).

Table 7. Adjusted mean difference in log-transformed plasma SHBG levels in relation to reproductive history and usage of exogenous estrogen

log-SHBG levels*	Pooled (n=8712)
Oral contraceptive use	-0.013 (-0.037 to 0.011)
(ever vs. never)	
Exogenous estrogen use	
(current vs. past/never)	0.225 (0.043 to 0.408)
(ever vs. never)	0.083 (0.012 to 0.153)
Years since menopause	0.001 (-0.007 to 0.010)
(per 5-year)	
Menarche age	-0.000 (-0.021 to 0.021)
(≥13 vs. <13)	
Number of parities	-0.028 (-0.062 to 0.006)
(≥ 5 vs. <5)	
Marital status	-0.033 (-0.062 to -0.003)
(currently married vs. not married)	

Abbreviation: SHBG, sex hormone-binding globulin; log-SHBG levels, log-transformed SHBG; *P*, *P*-value.
Data are expressed as adjusted mean (95% confidence interval) of log-transformed SHBG levels.
Adjusted mean (95% confidence interval) of log-transformed SHBG levels according to exogenous estrogen usage, oral contraceptives usage, and reproductive history (years since menopause, menarche age, number of parities, and marital status) were estimated using random-effects model with adjustment for age (continuous), BMI (continuous), physical activity levels (MET-hours/week [continuous]), cigarettes smoking per day (continuous), alcohol ingestion per day (in gram [continuous]), total energy intake (continuous), all dietary factors included in Table 6, and reproductive history that were not examined as the primary independent variable.

Supplementary Table

Number of individuals, study design, outcome of interests, and methods used to measure SHBG levels and their CVs for studies included in the pooled analysis

Study	Full name	Study design	Outcome	Number of Individuals	Method	CV
AS090	WHI sex hormone and genetic risk factors for hip fracture	Nested case-control study	Hip fracture	400 cases 400 controls	Solid-phase two-site CIA	intraassay CV : 4.1–7.7%, interassay CV : 5.8–13%.
AS110	Sex steroid hormones and risk of coronary heart disease	Nested case-control study	Coronary heart disease	311 cases 297 controls	Solid-phase two-site CIA	interassay CV : 8.0%
AS167	Sex Hormones, Risk Factors and Risk of ER-Positive and ER-Negative Breast Cancers	Case-cohort study	Breast cancer	306cases 581 controls	Solid-phase two-site CIA	intraassay CV : 2.5 - 5.3%. interassay CV : 5.2 - 6.6%.
AS238	Biochemical predictors of type 2 DM in women	Nested case-control study	Type 2 Diabetes	497 cases 1,007 controls	ECL	interassay CV : 5.4%
BA7	Endogenous estradiol and the effects of estrogen therapy on major outcomes of WHI	Case-cohort study	Stroke, Coronary heart disease, Venous thromboembolism, and Dementia/ MCI.	2,837 cases 373 controls	Solid-phase two-site CIA	intraassay CV :4.0 - 7.5%. interassay CV : 8.0 - 13%.
BA9	Biochemical antecedents of fracture in minority women	Nested case-control study	Fracture	1,057 cases 1,101 controls	Solid-phase two-site CIA	interassay CV : 9.1%.
BA21	Understanding the role of sex hormones in colorectal cancer	Nested case-control study	Colorectal cancer	353 cases 735 controls	Solid-phase two-site CIA	interassay CV : 3.7%.
W5	Correlates of endogenous sex hormone concentrations in WHI	Cross-sectional study	not applicable	300 random samples	IRMA	intraassay CV :5.7%. interassay CV : 17.7%.

W9	Biological markers of the effect of HT on risk of fractures in the Women's Health Initiative Clinical Trial	Nested case-control study	Hip fracture	682 cases 641 controls	Solid-phase two-site CIA	intraassay CV : 4.1–7.7%, interassay CV : 5.8–13%.
W10	Biological markers of the effect of HT on risk of breast cancer in the Women's Health Initiative Clinical Trial	Nested case-control study	Breast cancer	682 cases 641 controls	Solid-phase two-site CIA	intraassay CV : 4.0–7.5%, interassay CV : 8.0–13%.
W18	HT hormone pretest	Cross-sectional study	Not applicable	193 random samples	Solid-phase two-site CIA	intraassay CV : 4.1–7.7%, interassay CV : 5.8–13%.

Abbreviation: CV; coefficient of variation; SHBG, sex hormone-binding globulin; Solid-phase two-site CIA, Solid-phase, two-site chemiluminescent immunoassay using the Immulite Analyzer; IRMA, immunoradiometric assay .

Figure Legends

Figure 1A-D and Supplementary Figure 2A-B: Estimated plasma SHBG levels using spline analyses

Unrestricted quadratic spline models with the inclusion of transformed variables for each factor in the regression models (with knots at the 20th, 40th, 60th, and 80th percentiles) with adjustment for age (continuous), BMI (continuous), physical activity levels (MET-hours/week [continuous]), cigarettes smoking per day (continuous), alcohol ingestion per day (in gram [continuous]), total energy intake (continuous), all dietary factors included in Table 6, and reproductive history that were not examined as the primary independent variable were used to plot curves (solid curve=point estimates; dashed curves=pointwise 95% confidence limits).

Supplementary Figure 2A-E: Pooled mean difference in log-SHBG levels according to ethnicity and studies

The overall estimates were obtained using a random-effects model. The dots indicate the point estimates for study- and ethnicity-specific log-transformed SHBG levels and the horizontal lines represent 95% confidence intervals (CIs) per 5-year in age, BMI unit, 7.5 MET-hours/week for physical activity (e.g., brisk walking for 30 minutes, five days a week), 10 cigarettes smoked/day, 10 grams of alcohol ingested/day, and one serving/day for each dietary factor were estimated using generalized linear models with adjustment for age (continuous), BMI (continuous), physical activity levels (MET-hours/week [continuous]), cigarettes smoking per day (continuous), alcohol ingestion per day (in gram [continuous]), total energy intake (continuous), all dietary factors included in Table 6, and reproductive history that were not examined as the primary independent variable. Then, the results were pooled using the DerSimonian and Laird random-effects model.

Abbreviations: ES, estimate; 95%CI, 95% confidence interval.

Figure 1 Estimated plasma SHBG levels using spline analyses

Figure 1A. Estimated plasma SHBG using spline analyses according to age

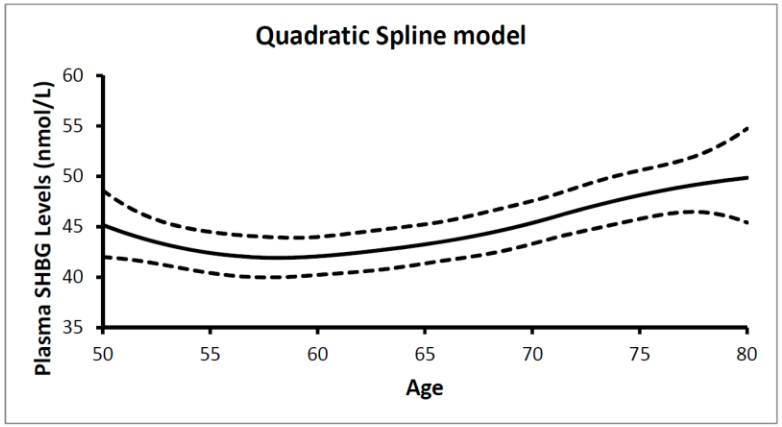


Figure 1B. Estimated plasma SHBG using spline analyses according to body mass index (BMI)

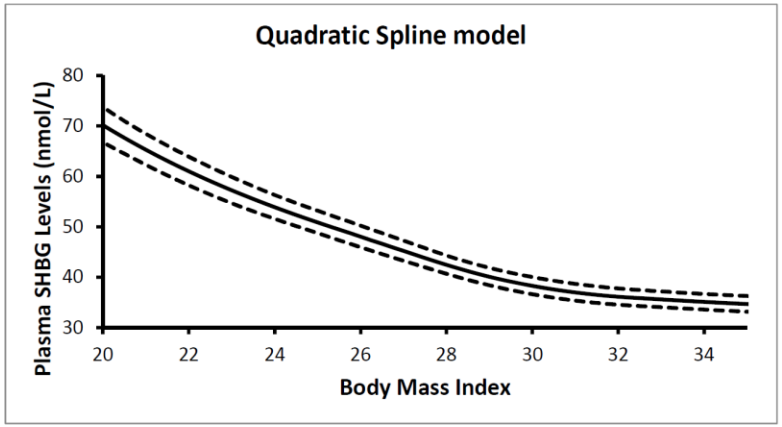


Figure 1C. Estimated plasma SHBG using spline analyses according to physical activity

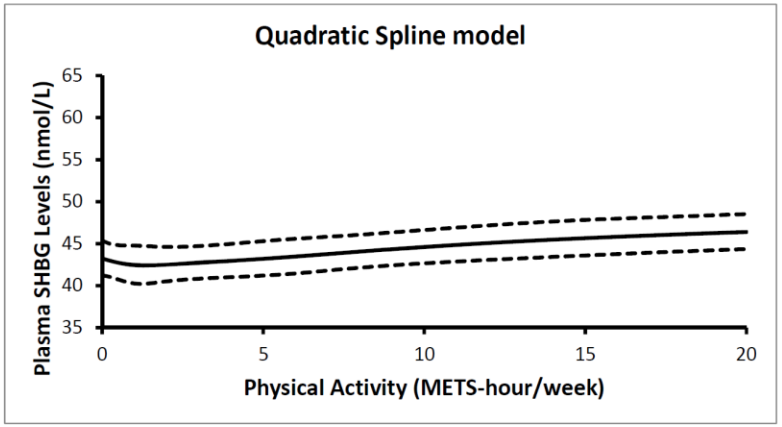
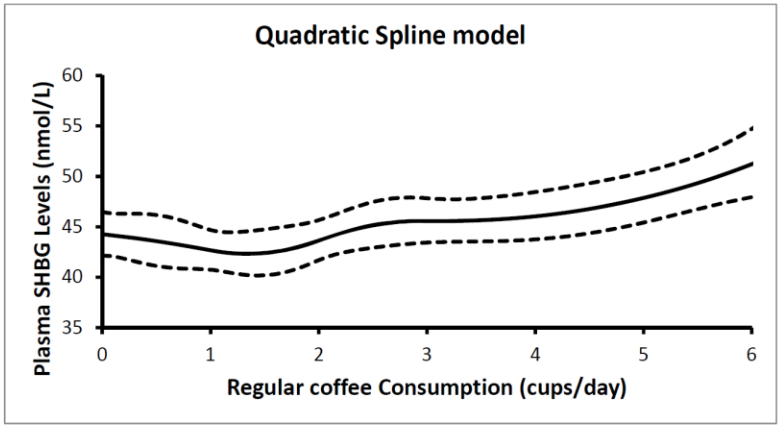
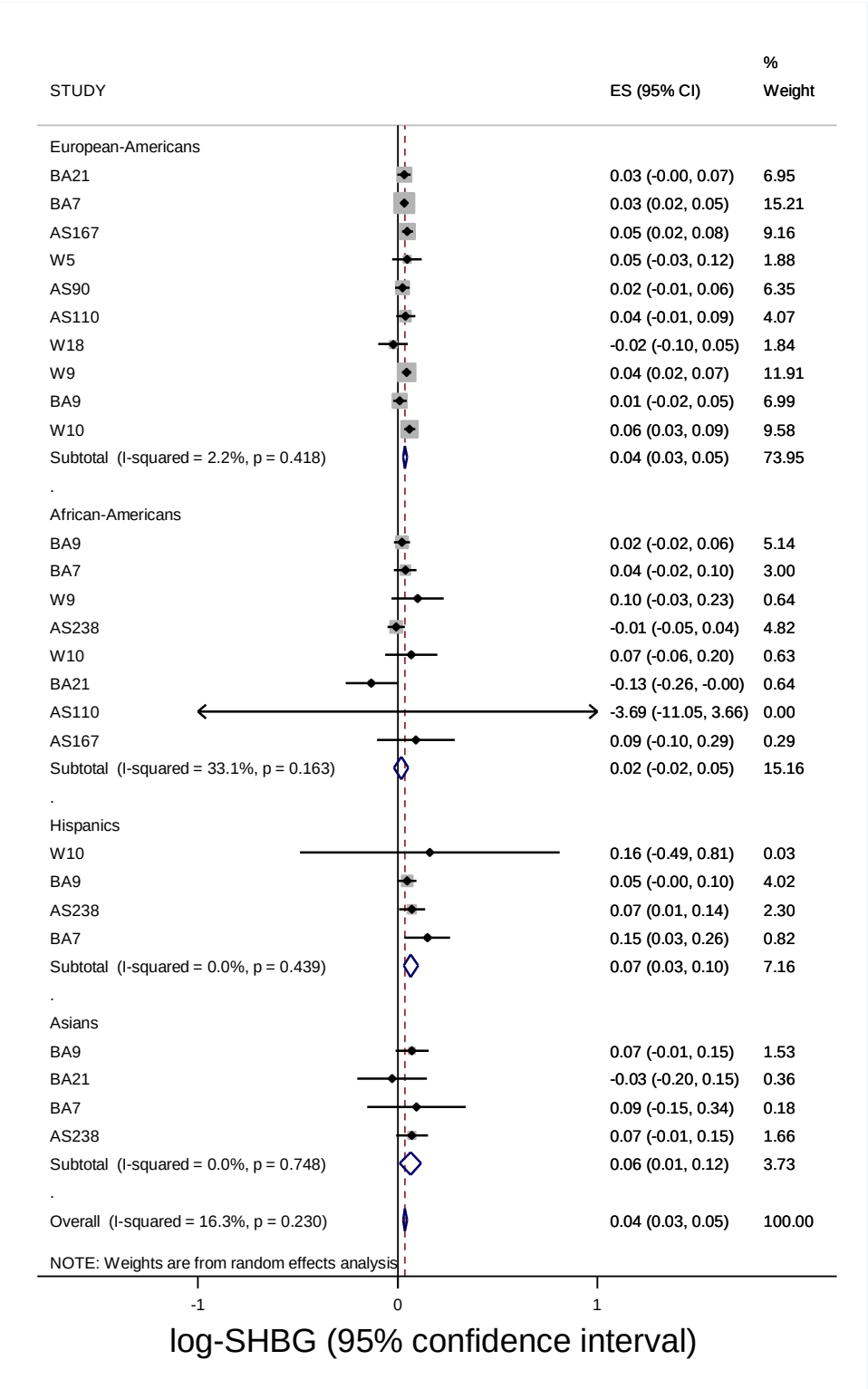


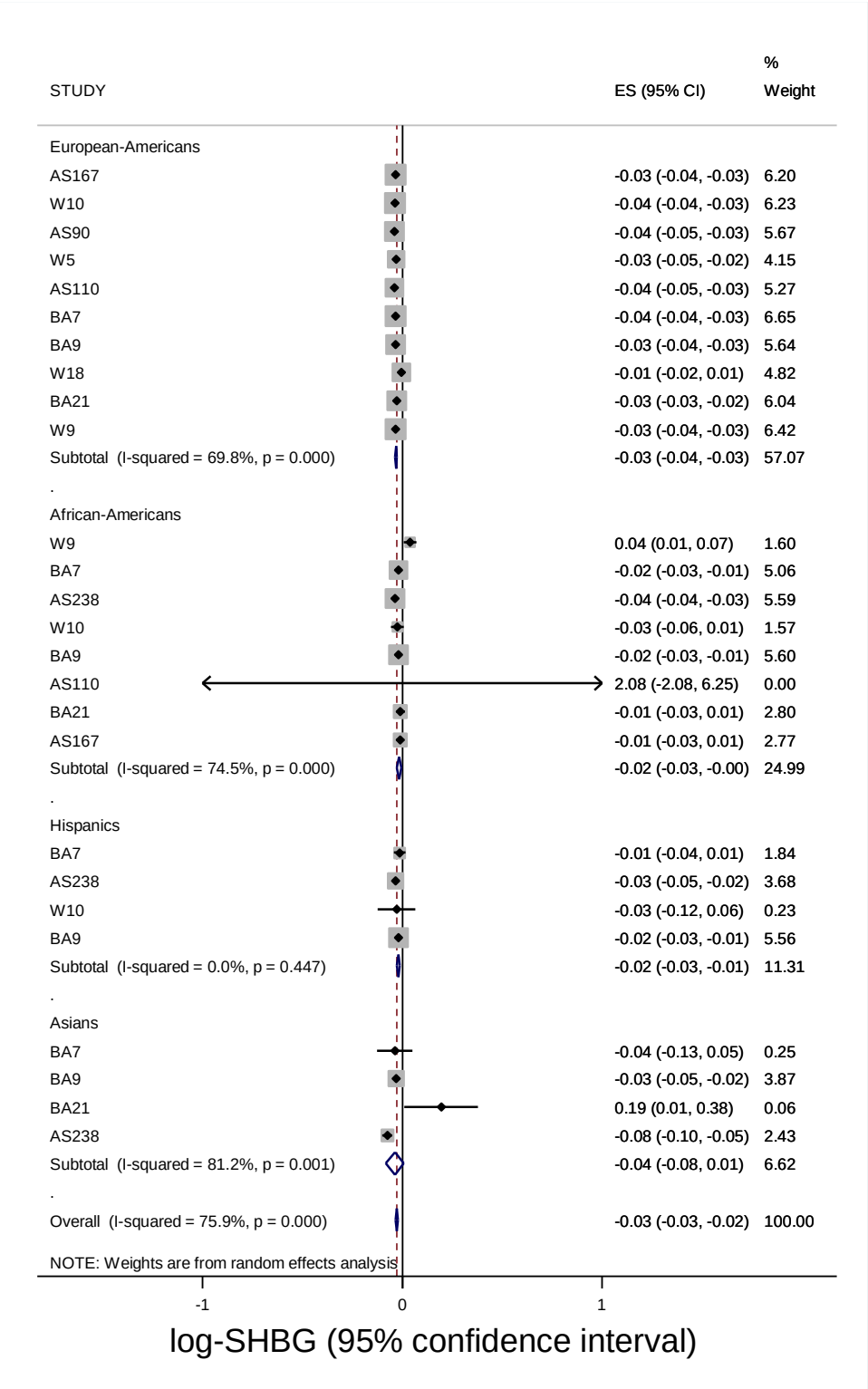
Figure 1D Estimated plasma SHBG using spline analyses according to regular coffee intake



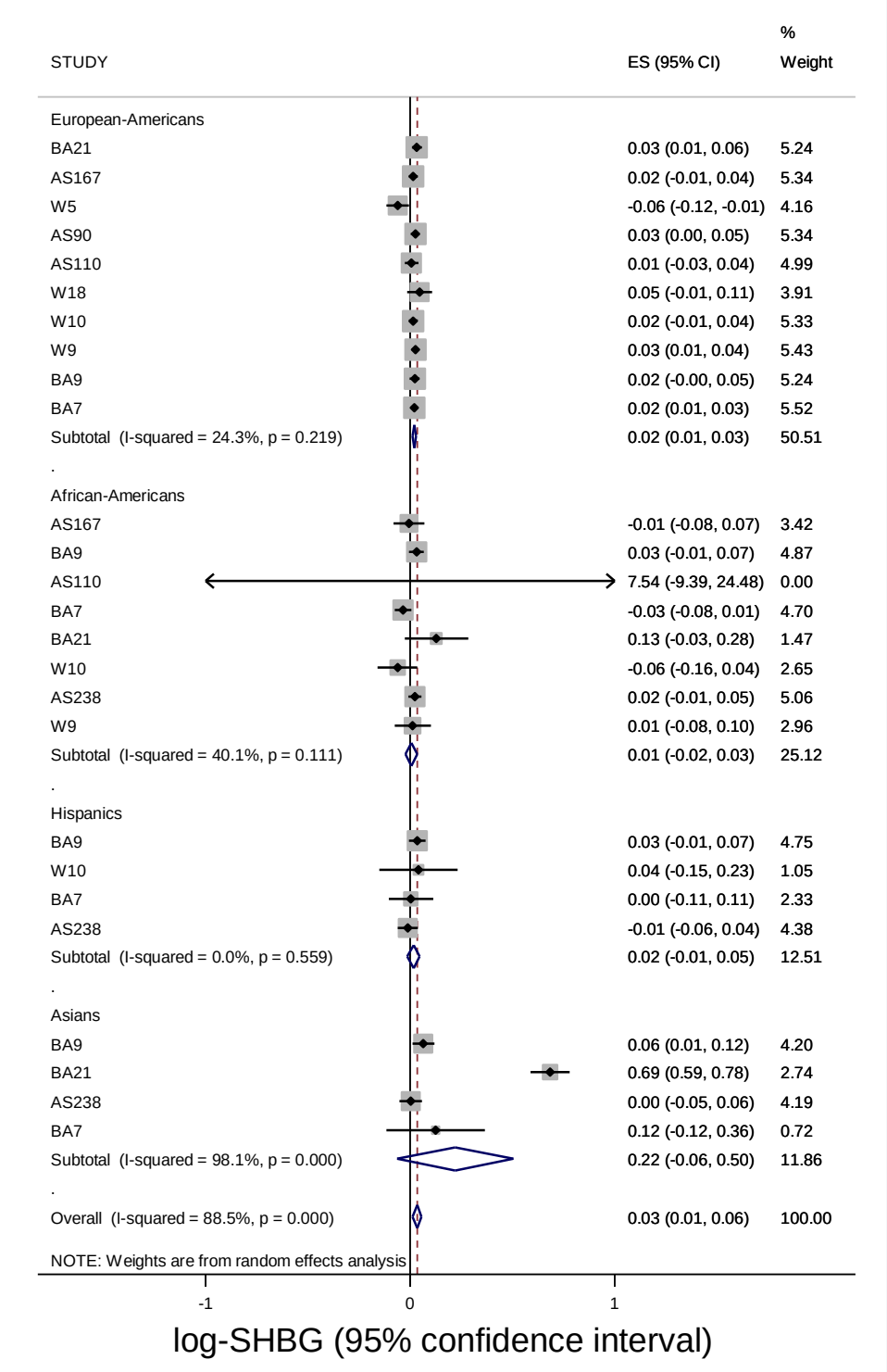
Supplementary Figure 1A Mean difference in log-SHBG levels per 5-year increment in age



Supplementary Figure 1B Mean difference in log-SHBG levels per BMI unit

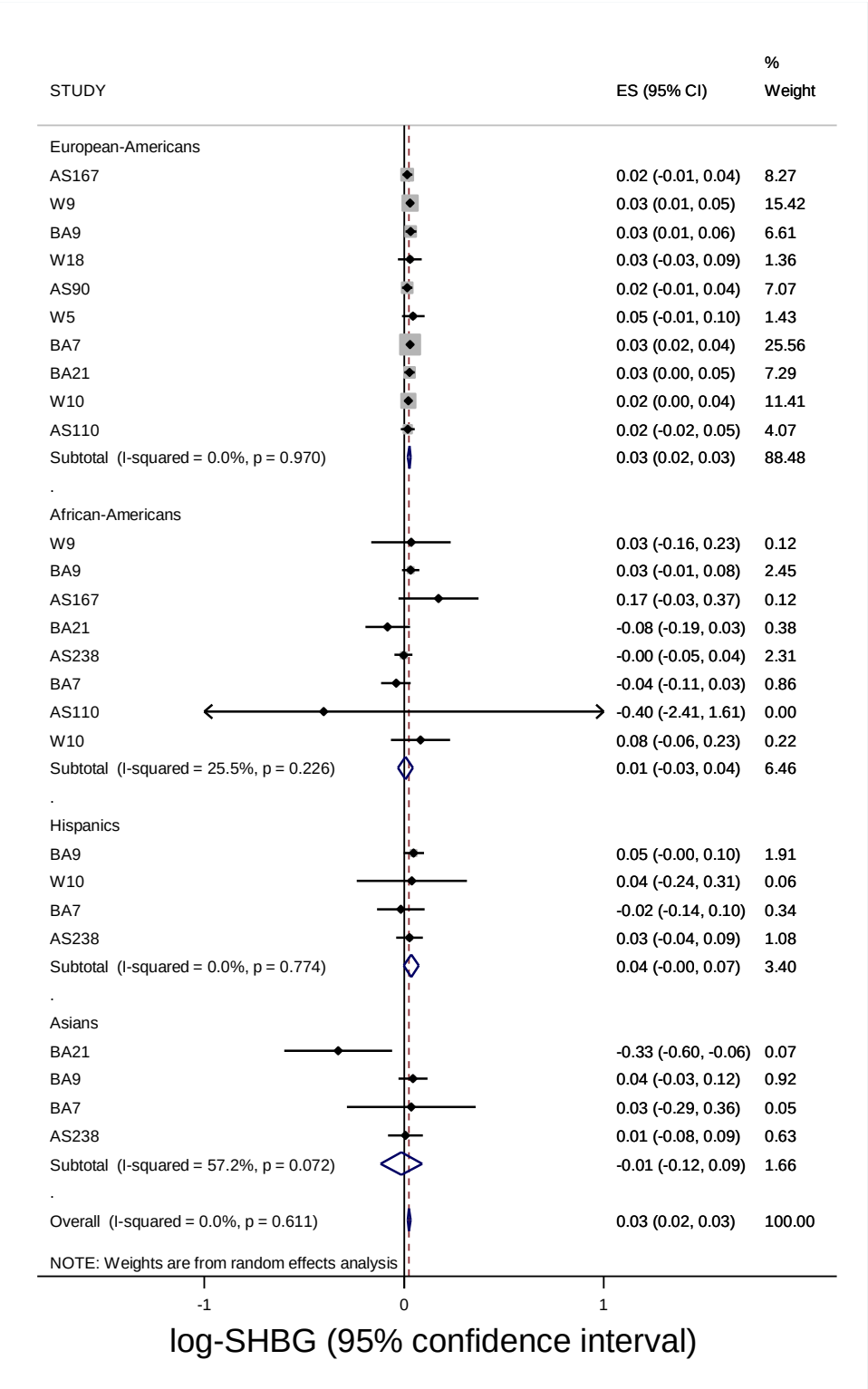


Supplementary Figure 1C Mean difference in log-SHBG levels per increment of 7.5 MET-h/w of physical activity

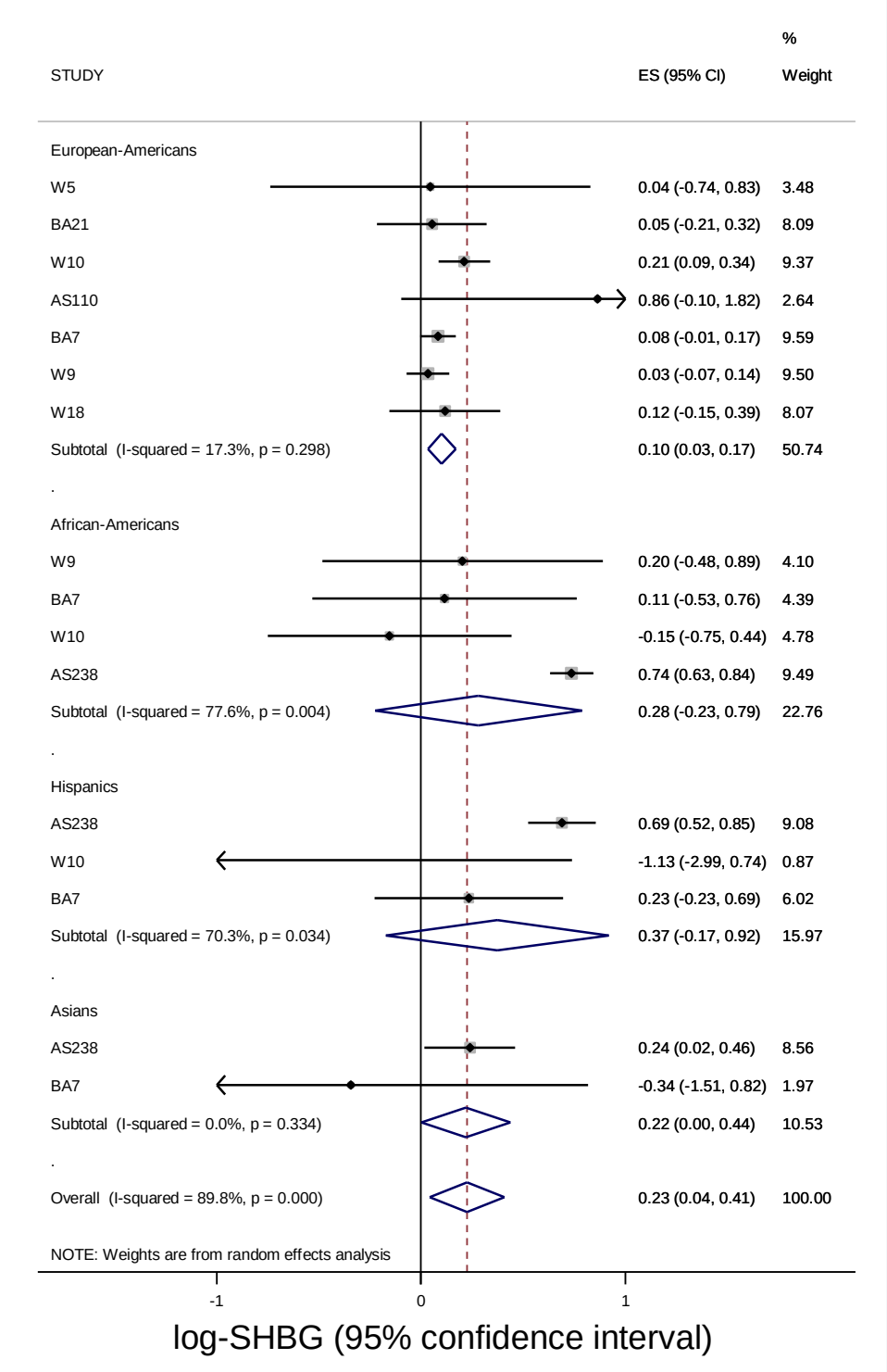


Abbreviation: MET-h/w, metabolic equivalent-hours per week

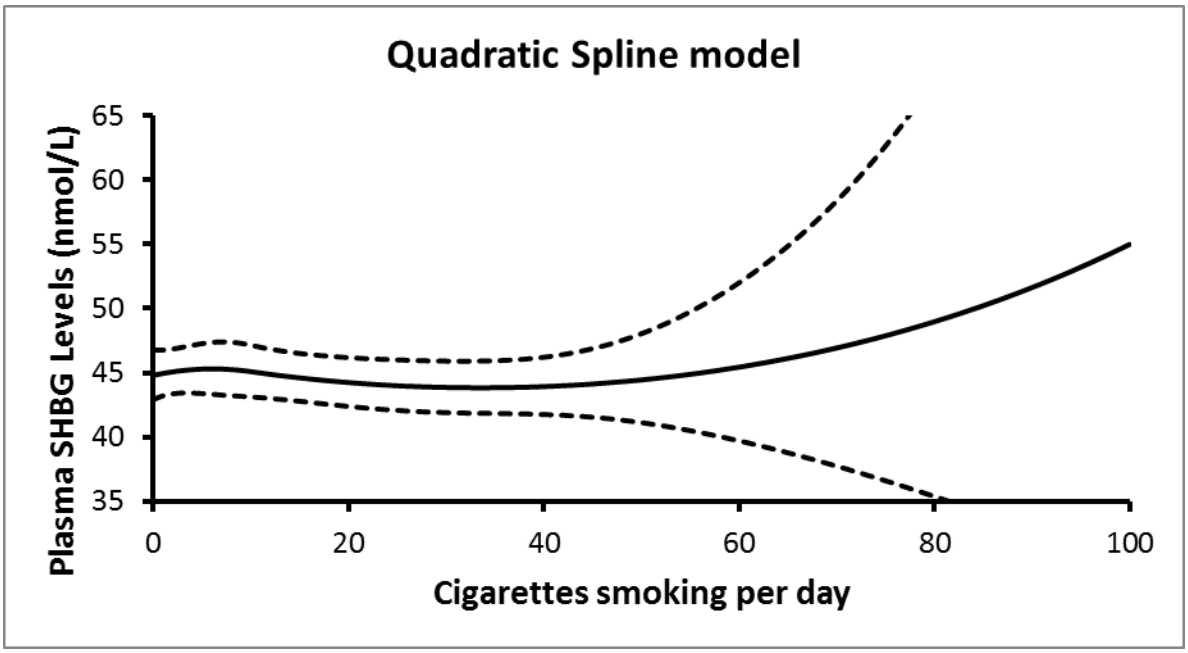
Supplementary Figure 1D Mean difference in log-SHBG levels per one cup/day of regular coffee intake



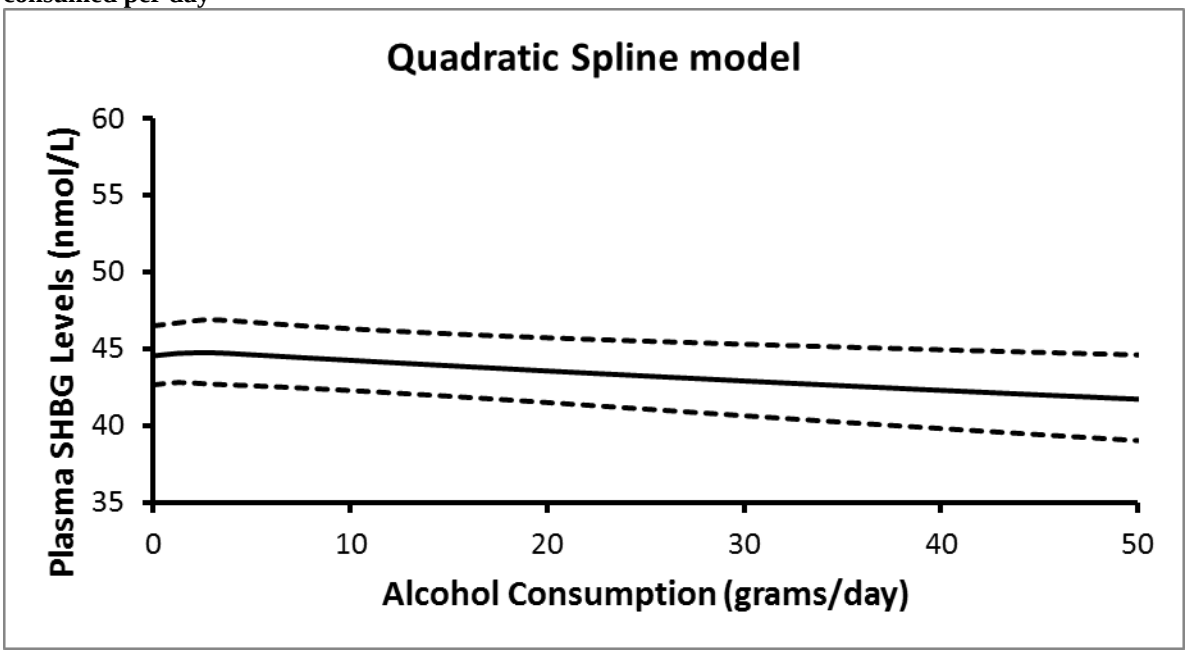
Supplementary Figure 1E Mean difference in log-SHBG levels by usage of exogenous estrogen (Current vs. Past or Never)



Supplementary Figure 2A. Estimated plasma SHBG levels using spline analyses according to number of cigarettes smoked per day



Supplementary Figure 2B. Estimated plasma SHBG levels using spline analyses according to alcohol consumed per day



3.6. References

1. Anderson DC. Sex-hormone-binding globulin. *Clinical endocrinology.* Jan 1974;3(1):69-96.

2. Kahn SM, Hryb DJ, Nakhla AM, Romas NA, Rosner W. Sex hormone-binding globulin is synthesized in target cells. *The Journal of endocrinology.* Oct 2002;175(1):113-120.

3. Rosner W, Hryb DJ, Khan MS, Nakhla AM, Romas NA. Sex hormone-binding globulin mediates steroid hormone signal transduction at the plasma membrane. *The Journal of steroid biochemistry and molecular biology.* Apr-Jun 1999;69(1-6):481-485.

4. Key T, Appleby P, Barnes I, Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *Journal of the National Cancer Institute.* Apr 17 2002;94(8):606-616.

5. Allen NE, Key TJ, Dossus L, et al. Endogenous sex hormones and endometrial cancer risk in women in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Endocr Relat Cancer.* Jun 2008;15(2):485-497.

6. Haffner SM, Moss SE, Klein BE, Klein R. Sex hormones and DHEA-SO4 in relation to ischemic heart disease mortality in diabetic subjects. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. *Diabetes care.* Oct 1996;19(10):1045-1050.

7. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA.* Mar 15 2006;295(11):1288-1299.

8. Lee JS, LaCroix AZ, Wu L, et al. Associations of serum sex hormone-binding globulin and sex hormone concentrations with hip fracture risk in postmenopausal women. *The Journal of clinical endocrinology and metabolism.* May 2008;93(5):1796-1803.

9. Perry JR, Weedon MN, Langenberg C, et al. Genetic evidence that raised sex hormone binding globulin (SHBG) levels reduce the risk of type 2 diabetes. *Human molecular genetics.* Feb 1 2010;19(3):535-544.

10. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N Engl J Med.* Sep 17 2009;361(12):1152-1163.

11. Brand JS, van der Tweel I, Grobbee DE, Emmelot-Vonk MH, van der Schouw YT. Testosterone, sex hormone-binding globulin and the metabolic syndrome: a systematic review and meta-analysis of observational studies. *Int J Epidemiol.* Feb 2011;40(1):189-207.

12. Xita N, Tsatsoulis A. Genetic variants of sex hormone-binding globulin and their biological consequences. *Molecular and cellular endocrinology.* Mar 5 2010;316(1):60-65.

13. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Control Clin Trials.* Feb 1998;19(1):61-109.

14. Anderson GL, Manson J, Wallace R, et al. Implementation of the Women's Health Initiative study design. *Ann Epidemiol.* Oct 2003;13(9 Suppl):S5-17.

15. Henderson MM, Kushi LH, Thompson DJ, et al. Feasibility of a randomized trial of a low-fat diet for the prevention of breast cancer: dietary compliance in the Women's Health Trial Vanguard Study. *Prev Med.* Mar 1990;19(2):115-133.

16. Kristal AR, Feng Z, Coates RJ, Oberman A, George V. Associations of race/ethnicity, education, and dietary intervention with the validity and reliability of a food frequency questionnaire: the Women's Health Trial Feasibility Study in Minority Populations. *American journal of epidemiology.* Nov 15 1997;146(10):856-869.

17. Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol.* 1990;43(12):1327-1335.

18. Schakel SF, Sievert YA, Buzzard IM. Sources of data for developing and maintaining a nutrient database. *J Am Diet Assoc.* Oct 1988;88(10):1268-1271.

19. Patterson RE, Kristal AR, Tinker LF, Carter RA, Bolton MP, Agurs-Collins T. Measurement characteristics of the Women's Health Initiative food frequency questionnaire. *Ann Epidemiol.* Apr 1999;9(3):178-187.

20. Meyer AM, Evenson KR, Morimoto L, Siscovick D, White E. Test-retest reliability of the Women's Health Initiative physical activity questionnaire. *Med Sci Sports Exerc.* Mar 2009;41(3):530-538.

21. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled clinical trials.* Sep 1986;7(3):177-188.

22. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in medicine.* Jun 15 2002;21(11):1539-1558.

23. Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology (Cambridge, Mass.* Jul 1995;6(4):356-365.

24. Lee IR, Dawson SA, Wetherall JD, Hahnel R. Sex hormone-binding globulin secretion by human hepatocarcinoma cells is increased by both estrogens and androgens. *The Journal of clinical endocrinology and metabolism.* Apr 1987;64(4):825-831.

25. Rosner W, Aden DP, Khan MS. Hormonal influences on the secretion of steroid-binding proteins by a human hepatoma-derived cell line. *The Journal of clinical endocrinology and metabolism.* Oct 1984;59(4):806-808.

26. Plymate SR, Matej LA, Jones RE, Friedl KE. Inhibition of sex hormone-binding globulin production in the human hepatoma (Hep G2) cell line by insulin and prolactin. *The Journal of clinical endocrinology and metabolism.* Sep 1988;67(3):460-464.

27. Dechaud H, Denuziere A, Rinaldi S, Bocquet J, Lejeune H, Pugeat M. Age-associated discrepancy between measured and calculated bioavailable testosterone in men. *Clinical chemistry.* Apr 2007;53(4):723-728.

28. Muller M, den Tonkelaar I, Thijssen JH, Grobbee DE, van der Schouw YT. Endogenous sex hormones in men aged 40-80 years. *European journal of endocrinology / European Federation of Endocrine Societies.* Dec 2003;149(6):583-589.

29. Caldwell JD, Jirikowski GF. Sex hormone binding globulin and aging. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme.* Mar 2009;41(3):173-182.

30. Vermeulen A, Kaufman JM, Giagulli VA. Influence of some biological indexes on sex hormone-binding globulin and androgen levels in aging or obese males. *The Journal of clinical endocrinology and metabolism.* May 1996;81(5):1821-1826.

31. Muzumdar R, Ma X, Atzmon G, Vuguin P, Yang X, Barzilai N. Decrease in glucose-stimulated insulin secretion with aging is independent of insulin action. *Diabetes.* Feb 2004;53(2):441-446.

32. Simo R, Barbosa-Desongles A, Lecube A, Hernandez C, Selva DM. Potential Role of Tumor Necrosis Factor-alpha in Downregulating Sex Hormone-Binding Globulin. *Diabetes.* Feb 2012;61(2):372-382.

33. Gower BA, Nyman L. Associations among oral estrogen use, free testosterone concentration, and lean body mass among postmenopausal women. *The Journal of clinical endocrinology and metabolism.* Dec 2000;85(12):4476-4480.

34. Kuhnz W, Staks T, Jutting G. Pharmacokinetics of levonorgestrel and ethinylestradiol in 14 women during three months of treatment with a tri-step combination oral contraceptive: serum protein binding of levonorgestrel and influence of treatment on free and total testosterone levels in the serum. *Contraception.* Dec 1994;50(6):563-579.

35. Moore JW, Key TJ, Bulbrook RD, et al. Sex hormone binding globulin and risk factors for breast cancer in a population of normal women who had never used exogenous sex hormones. *Br J Cancer.* Nov 1987;56(5):661-666.

36. Hawkins VN, Foster-Schubert K, Chubak J, et al. Effect of exercise on serum sex hormones in men: a 12-month randomized clinical trial. *Med Sci Sports Exerc.* Feb 2008;40(2):223-233.

37. Haffner SM, Katz MS, Dunn JF. Increased upper body and overall adiposity is associated with decreased sex hormone binding globulin in postmenopausal women. *Int J Obes.* Jul 1991;15(7):471-478.

38. Kotsopoulos J, Eliassen AH, Missmer SA, Hankinson SE, Tworoger SS. Relationship between caffeine intake and plasma sex hormone concentrations in premenopausal and postmenopausal women. *Cancer.* Jun 15 2009;115(12):2765-2774.

39. Goto A, Song Y, Chen BH, Manson JE, Buring JE, Liu S. Coffee and caffeine consumption in relation to sex hormone-binding globulin and risk of type 2 diabetes in postmenopausal women. *Diabetes.* Jan 2011;60(1):269-275.

40. Park BK, Kitteringham NR. Assessment of enzyme induction and enzyme inhibition in humans: toxicological implications. *Xenobiotica.* Nov 1990;20(11):1171-1185.

41. Higgins LG, Cavin C, Itoh K, Yamamoto M, Hayes JD. Induction of cancer chemopreventive enzymes by coffee is mediated by transcription factor Nrf2. Evidence that the coffee-specific diterpenes cafestol and kahweol confer protection against acrolein. *Toxicol Appl Pharmacol.* Feb 1 2008;226(3):328-337.

42. Huber WW, Rossmanith W, Grusch M, et al. Effects of coffee and its chemopreventive components kahweol and cafestol on cytochrome P450 and sulfotransferase in rat liver. *Food Chem Toxicol.* Apr 2008;46(4):1230-1238.

43. Huber WW, Scharf G, Rossmanith W, et al. The coffee components kahweol and cafestol induce gamma-glutamylcysteine synthetase, the rate limiting enzyme of chemoprotective glutathione synthesis, in several organs of the rat. *Arch Toxicol.* Jan 2002;75(11-12):685-694.

44. Astrup A, Toubro S, Cannon S, Hein P, Breum L, Madsen J. Caffeine: a double-blind, placebo-controlled study of its thermogenic, metabolic, and cardiovascular effects in healthy volunteers. *The American journal of clinical nutrition.* May 1990;51(5):759-767.

45. Selva DM, Hammond GL. Thyroid hormones act indirectly to increase sex hormone-binding globulin production by liver via hepatocyte nuclear factor-4alpha. *Journal of molecular endocrinology*. Jul 2009;43(1):19-27.

46. Hooper L, Ryder JJ, Kurzer MS, et al. Effects of soy protein and isoflavones on circulating hormone concentrations in pre- and post-menopausal women: a systematic review and meta-analysis. *Human reproduction update*. Jul-Aug 2009;15(4):423-440.

Chapter 4

Genes in Sex Hormone Pathways and Diabetes Risk among African-American, Hispanic, and European-American Postmenopausal Women: the Women's Health Initiative SNP Health Association Resource (WHI-SHARe) and Genomics and Randomized Trials Network (WHI-GARNET)

4.1 Introduction

An increasing body of evidence suggests that androgens, estrogens, and sex hormone binding globulin (SHBG) may play important roles in insulin resistance, impaired glucose tolerance, and the development of type 2 diabetes (T2D) and gestational diabetes.[1-3] However, observational studies directly investigating the associations between sex hormones levels and T2D risk have been inconclusive. Genetic studies have reported that single nucleotide polymorphisms (SNPs) in sex hormone pathways were associated with circulating levels of sex hormones and SHBG.[4] Therefore, linking polymorphisms in sex hormone pathways to T2D risk would provide further insight in determining whether sex-steroids and SHBG play a causal role in T2D development. While previous genome wide association studies (GWAS) have not identified top signals in the sex-steroid pathway for T2D risk, it is possible that because the large number of SNPs tested, real but relatively modest associations are likely to be missed. This may be especially true after conservative adjustment for multiple comparisons using the Bonferonni correction. As an alternative, a number of approaches including genome-wide pathway analysis and candidate gene approaches (or candidate pathway analysis) using SNPs on genome-wide arrays have been proposed to further address this limitation of GWAS.[5-8] The genome-wide

pathway-based approaches are useful for investigating enrichment for known pathways in a genome-wide scan by taking advantage of gene sets based on prior biological knowledge. In an attempt to specifically examine the role of genes in the sex-steroid pathways for T2D risk, we conducted a comprehensive analysis of sex steroid pathways utilizing already genotyped SNPs on genome-wide arrays among a large population of African-American and Hispanic postmenopausal women who participated in the WHI-SHARe cohort (n=11,678).

4.2 Study Design and Methods

Study Population

The Women's Health Initiative (WHI) is a long-term national health study that has focused on strategies for preventing heart disease, breast and colorectal cancer, and osteoporotic fractures in postmenopausal women. The original WHI study included 161,808 postmenopausal women enrolled between 1993 and 1998. The WHI has two major parts: a partial factorial randomized Clinical Trial (CT) and an Observational Study (OS); both were conducted at 40 Clinical Centers nationwide. The CT enrolled 68,132 postmenopausal women between the ages of 50-79 years into trials testing three prevention strategies. The OS examines the relationship between lifestyle, environmental, medical and molecular risk factors and specific measures of health or disease outcomes. This component involves tracking the medical history and health habits of 93,676 women not participating in the CT.

The WHI SNP Health Association Resource (SHARe) was conducted among 12,008 African-American and Hispanic participants in the WHI CT or OS with archived DNA specimens who agreed to the Supplemental Consent, which allowed for data sharing. To evaluate

the consistency across multiple populations, we further assessed whether the SNPs that were significantly associated among African-Americans or Hispanics in WHI-SHARe were also associated among the European-American population using data from the WHI-GARNET (the Genomics and Randomized Trials Network). WHI-GARNET participants were women enrolled in the WHI Hormone Therapy (HT) Trial who met eligibility requirements for this study and eligibility for submission to dbGaP, and who provided DNA samples. Of the approximately 27,000 women who participated in the HT trial, incident diabetes cases and matched controls who were free of prevalent or incident diabetes, coronary heart disease, stroke, and venous thrombosis were included. Matching criteria were age (±5 years), race/ethnicity, hysterectomy status, enrollment date (±1.5 years), and length of follow-up (±48 months). Controls were also prioritized on the basis of availability of plasma CVD biomarker availability (glucose, high density lipoprotein cholesterol, low density lipoprotein cholesterol, total cholesterol, insulin, triglycerides, C-reactive protein, and fibrinogen). The WHI-GARNET participants belonged to the following self-identified ethnic groups: European-American (87%), African American (5%), Hispanic (3%), Asian/Pacific Islander (1.8%), and Native American (0.7%).

Genotyping Methods

DNA samples were extracted from the buffy coats of blood samples collected at baseline. For WHI-SHARe the Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix, Inc., Santa Clara, CA) was used to genotype all samples. Lab technicians were blinded to the disease status during the genotyping processes. We excluded subjects on the basis of genotyping failure and quality control (n=149), relatedness (n=56), discordant race (n=56), and missing phenotype information (n=40). For the related individuals, the relative with the highest call rate was retained, while the other family members were excluded. After applying all exclusions, a

total of 8,180 African American women and 3,498 Hispanic women were included in the analysis. For WHI-GARNET, the Illumina HumanOmni1-Quad SNP platform (Illumina inc., San Diego, CA) was used to genotype all the samples at the Broad Institute (Cambridge, MA). Sample and SNP level processing quality control (QC) were performed in a standardized fashion at the GARNET Data Coordinating Center at University of Washington. After applying all exclusions, a total of 3,203 women (2,006 controls and 1,197 cases) were included in the WHI-GARNET study.

Choice of markers

Genes involved in sex-hormone pathways include 1) genes involved in sex hormone synthesis and transport: *CYP 11A1/17A1/19*(cytochrome P-450 11A1/17A1/19 genes), *HSD3B1/B2* (3β-hydroxysteroid dehydrogenase 1/2/3 genes), *HSD17B1/B2/B3* (17β-hydroxysteroid dehydrogenase 1/2/3 genes), and *SHBG* (sex hormone binding globulin gene); 2) genes involved in androgen metabolism and a gene encoding androgen receptor: *SRD5A1/2* (steroid 5-*α* reductase type I/II genes), *CYP3A4/5* (cytochrome P-450 3A4/5 genes), *UGT2B15/17* (uridine diphosphate-glucuronosyltransferase 2B15/17 genes), *HSD3B1/B2* (3β-hydroxysteroid dehydrogenase 1/2 gene), and *AR* (androgen receptor gene); 3) genes involved in estrogen metabolism and genes encoding estrogen receptors: *CYP1A1* (cytochrome P-450 1A1 gene), *CYP1B1* (cytochrome P-450 1B1 gene), *COMT* (catechol-O-methyltransferase gene), *SULT1A1/1E1/2A1* (sulfotransferase 1A1/1E1/2A1 genes), *UGT1A1*(uridine diphosphate-glucuronosyltransferase 1A1 gene), and *ERα/β* (estrogen receptor *α/β* genes); and 4)genes involved in progesterone metabolism and a gene encoding progesterone receptor: *AKR1C1-4/D1* (aldo-keto reductase C1-4/D1 genes), and *PGR* (progesterone receptor gene).[9] Using the online databases NetAffx (http://www.affymetrix.com) and NCBI(http://www.ncbi.nlm.nih.gov), we

searched for SNPs within 30kb upstream or downstream of those genes in the sex hormone pathways, which are covered by the Affymetrix Genome-Wide Human SNP Array 6.0. In total, we identified 1,388 SNPs within these genes. Numbers of SNPs per gene are listed in the Supplementary Table 1.

Ascertainment of Clinical Diabetes

In the WHI-SHARe, we included 2,582 participants (1,964 African-American and 618 Hispanic women) as clinical diabetes cases who self-reported at baseline ever having been diagnosed with diabetes when they were not pregnant (951 incident cases among African-American and 394 Hispanic women) or during follow-up receiving treatment for diabetes (1,013 prevalent cases among African-American and 224 prevalent cases among Hispanic women). Remaining women were non-cases. In WHI-GARNET , all incident cases of T2D self-reported during follow-up in the WHI HT trial were selected as potential cases from the August 14, 2009 database, using the cutoff date of July 7, 2002 (for the estrogen plus progestin trial) or February 29, 2004 (for the estrogen-alone trial). In total, we included 1,197 incident cases. Self-reported diabetes in the WHI has been found to be a reliable indicator of diagnosed diabetes based on medication inventories, fasting glucose levels, and medical record review.[10,11]

Statistical Analyses

The χ^2 test tests and Student *t* tests were used for comparisons of means and proportions between case participants and noncase participants (Table 1). To assess the relation between each SNP and diabetes risk, we fitted unconditional logistic regression models to estimate odds ratios (ORs) under additive genetic models separately for each ethnic group with adjustment for age, geographic regions, and potential population stratification by including the top 3 principal components from the principal component analysis (PCA)-based method (Table 2 and Table 3).

[12] To correct for multiple testing, we used false discovery rate (FDR) *q* values.[13] To display the -log10 (P) and LD with the index SNP by chromosomal location for regions of interest, we also generated regional association plots using SNAP (Supplementary Figure 1).[14] In addition, a SNPs-set enrichment analysis (SSEA) was performed using the PLINK set-test (Supplementary Table 2). In the SSEA, a gene-level score was directly computed from averaging SNP-level test statistics using unconditional logistic regression analysis with adjustment of age, geographic regions, and the top 3 principal components from the PCA-based method. After pruning SNPs in high LD (R^2>0.5), up to the top 5 SNPs with P-value smaller than 0.05 were selected. The same analysis was performed with 1,000 of simulated SNP sets in which the phenotype status of the individuals was permuted. An empirical *P*-value for the SNP set was computed by calculating the number of times the test statistic of the simulated SNP sets exceeded that of the original SNP set.[15] FDR *Q* values < 0.05 for SNPs level associations and empirical *P*-values < 0.05 for gene level associations from were considered significant. Finally, we assessed whether significant SNPs in WHI-SHARe were also associated with T2D risk in 3,203 European-Americans in the WHI-GARNET using genotyped SNPs and imputed SNPs. Genotype imputations were conducted in BEAGLE software[16] using the 1000 Genomes data (December 2010 release). SNPs with minor allele frequency >5% and high imputation quality (r^2 >0.5) were included in the analysis. Similar to WHI-SHARe, unconditional logistic regression models were fitted with adjustment of matching factors and ancestry estimates by including the top 3 principal components for global ancestry. Statistical analyses were conducted using PLINK (PLINK (V.1.07; http://pngu.mgh.harvard.edu/purcell/plink/), R (version 2.14; R Development Core Team 2008) and STATA (version 11.2; StataCorp, College Station, TX).

4.3 Results

In both African-American and Hispanic women, case participants had greater BMI and were more likely to have a family history of diabetes and use postmenopausal hormone. Also, case participants tended to exercise less than controls (Table 1).

Table 2 shows associations between SNPs in sex hormone pathway and diabetes risk in African-American women. Among African-American women, 2 SNPs at *ERα*, 5 SNPs at *AR*, 15 SNPs at *HSD3B1*, 1 SNP at *HSD17B3*, 1 SNP at *SULT1A1*, and 6 SNPs at *UGT2B15* were significantly associated with diabetes risk. For *ERα*, rs9479126 located at intron 4 of *ERα* was the most significant SNP associated with diabetes risk. Odds ratios for T2D were 1.20 (95% confidence interval, 1.09-1.32; *Q*-value= 0.045) for rs9479126. For *AR*, the most significant SNP (rs7472268) was located at upstream region of *AR* and the corresponding odds ratio was 1.22 (95% confidence interval, 1.11-1.34; *Q*=0.004). Among 15 SNPs at *HSD3B1* significantly associated with diabetes risk, rs2362811 had the most significant *P*-values (odds ratio= 0.83; 95% confidence interval, 0.74-0.92; *Q*=0.004). This SNP is located upstream of *HSD3B1* and downstream of *HSD3B2*. Some other SNPs located around this SNP and SNPs within *HSD3B1* were also significantly associated with diabetes risk (Supplementary Figure 1A). Also, rs7019471 at the downstream region of *HSD17B3* (odds ratio=1.20; 95% confidence interval, 1.09-1.33; *Q* =0.03) and rs2411453 at intron 1 of *SULT1A1* (odds ratio=1.15; 95% confidence interval, 1.06-1.25; *Q*=0.002) were significantly associated with diabetes risk. Six SNPs upstream of *UGT2B15* were significantly associated with diabetes risk. Rs294754 was the most significant SNP among these SNPs (odds ratio=0.86; 95% confidence interval, 0.79-0.94; *Q*=0.02) and other significant SNPs were in moderately high linkage disequilibrium with

rs29754 (Supplementary Figure 1B). Five genes with significant SNPs were also significant in SSEA in African-Americans (empirical *P*-values[*P*]: ERα, 0.02; AR, 0.04; HSD3B1, 0.004; SULT1A1, 0.03; and UGT2B15, 0.03) (Supplementary Table 2).

Table 3 shows associations between SNPs in sex hormone pathway and diabetes risk in Hispanic women. Among Hispanic women, 2 SNPs at *CYP3A4* and 7 SNPs at *PGR* were significantly associated diabetes risk. A SNP, rs473706, at the upstream region of *CYP3A4* was the most significant SNP in the *CYP3A4* region (odds ratio=1.67; 95% confidence interval, 1.17-2.39; *Q*=0.05). Another SNP, rs2737418, located upstream of *CYP3A4* was also significantly associated with diabetes risk in Hispanic women. These two SNPs were in high linkage disequilibrium (R^2=1.00; D'= 1.00) (Supplementary Figure 1C). Of the 7 SNPs at *PGR* that were significantly associated with diabetes risk among Hispanic women, rs948516 at the upstream region of *PGR* was the most significant one (odds ratio=0.66; 95% confidence interval, 0.56-0.78; *Q*=0.00004). Other significant SNPs in high linkage disequilibrium with rs948516 were also significantly associated with diabetes risk in Hispanic women (Supplementary Figure 1D). PGR was also found to be significant in SSEA (*P* = 0.002), but CYP3A4 was not (*P*=0.07). SSEA also indicated *HSD17B2* (*P*=0.048), and *HSD3B2* (*P*=0.04) with significant associations (Supplementary Table 2).

We have further examined whether the directions of effect estimates for the most significant SNP at each gene among either African-Americans or Hispanics were consistent across ethnic groups (Tables 2-3). Due to low minor allele frequencies, ORs and 95% CIs for rs7472268 at *AR* and rs7019471 at *HSD17B3* displayed wide variations comparing Hispanics and African-Americans. With one exception (for rs2362811 at *HSD3B1)*, however, the directions of ORs for all significant variants (at *ERα, UGT2B15*, *SULT1A1, CYP3A4*, and *PGR*) were

generally consistent across ethnic groups. This was true even for SNPs that displayed evidence of significant heterogeneity across ethnicity.

Finally, we examined consistency across populations. We were able to examine 21 SNPs at *ESRα*, *HSD3B1*, *PGR*, *SULT1A1*, and *UGT2B15* among three ethnic groups: African-Americans in WHI-SHARe; Hipanics in WHI-SHARe; and European-Americans in WHI-GARNET (Table 4 and Supplementary Figure 2). Two SNPs at *PGR* and 1 SNP at *SULT1A1* showed the same directions in all three ethnic groups, and 15 SNPs (= 71% of the 21 SNPs) showed the same directions in at least two ethnic groups. Of note, SNPs in *HSD3B1* were significantly replicated in WHI-GARNET (OR for the most significant SNP [rs2011191], 1.15; 95% CI 1.04-1.28; $P=9.3\times10^{-3}$).

4.4 Discussion

In this large population of African-American and Hispanic minority participants in the WHI-SHARe cohort, SNPs in *ERα*, *AR*, *HSD3B1*, *HSD17B3*, *SULT1A*, and *UGT2B15* were significantly associated with diabetes risk among African-American women and SNPs in CYP3A4 and PGR were significantly associated with diabetes risk among Hispanic women. The directions for the significant variants were generally consistent across ethnic groups.

Previous genome wide association studies (GWAS) identified a number of top genetic loci for T2D risk.[17] These include *FTO*, *PPARG*, *KCNJ11*, *TCF7L2*, *HHEX*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8*, and *WFS1*. Although these top signals are not attributable to the sex hormone pathways, accumulating evidence suggests that sex hormones and SHBG play a significant role in T2D development. Previous GWAS might have missed variants in sex

hormone pathways, because GWAS involve large numbers of statistical tests, which reduce statistical power to detect more modest associations, such as those in sex hormone pathway genes. Hypothesis-driven targeted genotyping may be better suited to assess these more subtle associations.

Several variants in *AR* were significantly associated with diabetes risk among African-American women. Testosterone exerts its effects through binding to androgen receptors. Thus, if testosterone influences glucose metabolism, polymorphisms in *AR* might be associated with T2D risk. Previous studies have shown that CAG repeat polymorphism at exon 1 of *AR* was associated with HbA1c levels[18] and HOMA-IR, the homeostasis model assessment of insulin resistance.[19] Also, androgen receptor knockout mice manifested insulin resistance and impaired glucose tolerance[20], and mice with selective knockdown of androgen receptor in adipocytes manifested visceral obesity and hyperglycemia.[21] Further, a recent systematic review has reported that in cross-sectional studies, women with type 2 diabetes had higher levels of testosterone compared to controls, whereas men with type 2 diabetes had lower levels of plasma total testosterone[1], supporting active roles of androgen in the development of type 2 diabetes.

Variants in estrogen receptor genes might also play roles in the development of type 2 diabetes. We observed that in African-American women SNPs in intron 4 and intron 6 of *ERα* were moderately associated with diabetes risk. Earlier studies reported that polymorphisms in intron 1, 2, 3, and 4 of the *ERα* have been found to be associated with T2D risk in Chinese, Hungarian, French, Swedish, and African-American and European-American cohorts.[22] A cohort study of 1928 French participants reported that SNPs in introns 3 and 4 of the *ESR1* gene were associated significantly with T2D risk (rs3020314, rs985694, P = 0.0009-0.001).[22] Evidence from clinical trials in human also supports the role of estrogen in the development of diabetes, in

which oral estrogen therapy lowered the risk of T2D among postmenopausal women.[23-25] Moreover, female *ERα* knockout mice developed insulin resistance and impaired glucose tolerance.[26] Further, epidemiologic evidence suggests that higher levels of estradiol at baseline are associated with the future development of T2D in both women and men.[1,27] Although we did not observe clear associations between variants in *ERβ* and diabetes risk in African-American and Hispanic women, mice lacking *ERβ* developed fasting hypoglycemia, increased levels of skeletal muscle glucose transporter 4 (GLUT4), and pancreatic islet hypertrophy.[28] Thus, *ERα* and *ERβ* might have distinct functions and we need further studies to understand the roles of these receptors in human.

SNPs in *HSD3B1* were associated with diabetes risk among African-American women and replicated in WHI-GARNET. *HSD3B1* encodes 3β-hydroxysteroid dehydrogenases that catalyze several reactions in the androgen pathway, resulting in production of testosterone. One case-control study of 561 participants reported that rs6203 and rs1047303 at exon 4 in the HSD3B1 gene were associated with essential hypertension risk.[29] No previous studies, however, reported association between variants at *HSD3B1* and T2D risk. *HSD3B1* catalyzes the formation of all classes of steroid hormones, including aldosterone, androstenedione, and testosterone. Thus, the significant findings observed in our study might not necessarily indicate the important role of sex hormone pathways.

A SNP at *HSD17B3* was associated with diabetes risk in African-American in WHI-SHARe. *HSD17B3* encodes 17-beta-hydroxysteroid dehydrogenase that converts androstenedione to testosterone. No studies have reported associations between variants at *HSD17B3* and diabetes risk. Also, in African-American, a SNP at *SULT1A1* was significantly associated with diabetes risk. *SULT1A1* encodes sulfotransferase which converts estradiol to

estradiol-sulfate. However, no previous studies showed that SNPs at *SULT1A1* was associated with diabetes risk. Thus, further studies are needed to confirm the findings. We observed that 6 SNPs in the upstream region of *UGT2B15* were also significantly associated with diabetes risk among African-American women in WHI-SHARe. *UGT2B15* encodes UDP-glucuronyltransferase, family 2, beta-15 which is highly expressed in adipose tissue[30] and glucuronidates androgens and their metabolites. There are, however, little data on the role of *UGT2B15* on T2D risk or glucose metabolism.

Among Hispanic women in WHI-SHARe, 2 SNPs at *CYP3A4* were associated with diabetes risk. *CYP3A4* encodes cytochrome P450, subfamily IIIA, polypeptide 4, which catalyzes testosterone to hydroxy-testosterone. A previous study of 5,259 unrelated Japanese individual reported that a polymorphism, rs55951658, at exon 4 of *CYP3A4* was significantly associated with prevalence of type 2 diabetes.[31] In addition to the function of CYP3A4 to catalyze testosterone, CYP3A4 involves in oxidation of a number of drug, steroids, fatty acids, and xenobiotics. Therefore, the significant association of SNPs in *CYP3A4* with T2D might not be through metabolism of sex hormones.

SNPs at the upstream region of *PGR* were also significantly associated diabetes risk among Hispanic women in WHI-SHARe. These SNPs were in high linkage disequilibrium. Blood progesterone levels rise continuously throughout pregnancy,[11] and the onset of gestational diabetes typically occurs during the second trimester of pregnancy, when the circulating levels of progesterone are high.[32] These observations have led to the hypothesis that increased levels of progesterone might be responsible for the development of gestational diabetes. In a study of 2,081 pregnant women, women who received a prophylactic intramuscular injection of 17α-hydroxyprogesterone caproate had a 2.9-fold increased risk of gestational diabetes, compared

with women who did not.[3] In addition, female mice lacking progesterone receptor showed low fasting glucose levels and high insulin levels on glucose injection, suggesting an important role of the progesterone pathway in glucose metabolism and beta-cell function.[33]

Despite the fact that there was some inconsistency of significant SNPs among African-American women and Hispanic women in WHI-SHARe and European-American women in WHI-GARNET, it is important to note that the directions of the effect estimates were generally consistent. Minor allele frequencies of significant SNPs among Hispanics were relatively higher than other ethnic groups. Thus, we might have had insufficient power to detect significant associations among African-Americans and European-Americans.

Of note, SNPs at genes encoding the major rate-limiting enzymes involved in sex hormones synthesis (*CYP17A1* and *CYP19A1*) were not significantly associated with diabetes risk. SNPs at these genes have been reported to be associated with sex hormone levels and cancer risks such as endometrial cancer[34], suggesting distinct roles of these enzymes. Further, although previous studies have reported several significant associations between variants in *SHBG* and diabetes risk[2], no significant SHBG –T2D risk were observed in the current study. The SNPs with the lowest *P*-value in the *SHBG* region were rs1641528 (OR=0.97, 95% CI=0.90 to 1.04, *P*=0.39) among African-Americans and rs9898876 (OR=1.15, 95% CI=0.97 to 1.37, *P*=0.12) among Hispanics. This may be due to the fact that none of the 5 SNPs in the *SHBG* gene included in the Affymetrix array were functional, which may not have captured the SHBG-T2D relation. In addition, the estimates were too imprecise to compare the directions of associations across ethnic groups. Taken together, future work needs to be performed to elucidate the role of *SHBG* gene in the development of T2D.

The major strength of this study is that this study uses the prior knowledge of sex hormone pathways and its potential association with type 2 diabetes, which may increase power to detect associations compared to broad, agnostic approach used in most genome wide association (GWA) studies. Other strengths include the use of a large, diverse population with a well-characterized assessment of epidemiologic profiles and genotyping among population of African-Americans, Hispanics, and European-Americans. However, several limitations deserve further consideration. First, although we observed significant genes/SNPs in WHI-SHARe among African-Americans or Hispanics, none of the significant SNPs were significantly associated with diabetes risk among both African-Americans and Hispanics. Further, only SNPs at HSD3B1 were replicated in WHI-GARNET. Potential explanations for the heterogeneity among ethnic groups include differences in allele frequencies, statistical power, environmental modifiers, allelic and genetic heterogeneity, and linkage disequilibrium patterns. Second, our sample size in Hispanics provided limited power to detect associations, especially for rarer variants. Third, population stratification may be a concern when all individuals in this study self-identified their ethnicity. However, we adjusted for 3 principal components for global ancestry to control for potential population stratification.[35] Finally, potential biases due to the misclassification of diabetes due to undiagnosed diabetes must be considered. However, because such biases are likely to be nondifferential with respect to genotypes, the observed estimates are likely be attenuated if at all.[36]

In conclusion, in this large population of African-American and Hispanic minority participants in the WHI-SHARe cohort, we found that SNPs in sex hormone pathways were associated with diabetes risk in African American and Hispanic women. Our study provides additional support for the roles of sex hormones in the development of T2D. Further studies

including fine-mapping with emphasis on functional loci in genes from the sex hormone pathways are warranted.

Table 1. Characteristics for cases and noncases included in WHI-SHARe by ethnic groups

Characteristic	African American cases (n=1,964)	African American noncases (n=6,216)	*P*-value	Hispanic cases (n=618)	Hispanic noncases (n=2,880)	*P*-value
Age (years)	62.2(6.8)	61.4 (7.1)	<0.001	60.4(6.7)	60.2(6.7)	0.52
BMI (kg/m^2)	33.1(6.5)	30.3(6.2)	<0.001	31.5(5.8)	28.3(5.3)	<0.001
Family history of diabetes, %	63.0	41.3	<0.001	62.4	38.6	<0.001
Current smoking, %	11.4	11.5	0.55	7.4	6.7	0.83
Alcohol intake, servings/w	0.7(5.0)	1.2(4.3)	<0.001	0.6(4.6)	1.4(3.6)	<0.001
Physical activity*	8.0(11.2)	10.2(13.1)	<0.001	8.5(12.3)	11.2(14.1)	<0.001
Ever PMH users, %	49.0	55.0	<0.001	52.2	63.1	<0.001

Data are mean (standard deviation) or %.
Chi squared tests and Student *t* tests were used for comparisons of means and proportions between cases and noncases.
Abbreviations: BMI= body mass index; PMH = postmenopausal hormone.
* Total energy expenditure of energy from recreational physical activity including walking, mild, moderate and strenuous physical activity levels (MET-hours/week)

Table 2. Associations of the significant SNPs among African-American women with clinical diabetes risk, Comparing results for African-American women (n=8,180) and Hispanic women (n=3,498).

Gene(s)	SNP ID	Minor allele	MAF in AA	OR[1] in AA	95% CIs in AA	*P*-value in AA	*Q*-value in AA	MAF in His	OR[1] in His	95% CIs in His	*P*-value in His
ERα	rs9479126	G	0.16	1.20	1.09-1.32	2.6×10^{-4}	0.045	0.13	1.05	0.87-1.25	0.63
	rs17081685	C	0.20	1.17	1.07-1.28	5.2×10^{-4}	0.045	0.24	1.01	0.88-1.17	0.86
AR	rs1458814	A	0.17	1.19	1.09-1.31	2.2×10^{-4}	0.004	0.02	0.59	0.32-1.07	0.08
	rs1458816	G	0.17	1.20	1.09-1.32	1.3×10^{-4}	0.004	0.02	0.63	0.36-1.10	0.11
	rs12353907	A	0.18	1.19	1.08-1.31	2.4×10^{-4}	0.004	0.02	0.64	0.39-1.05	0.08
	rs7472268	T	0.17	1.22	1.11-1.34	4.7×10^{-5}	0.004	0.03	0.66	0.41-1.06	0.09
	rs1331253	G	0.22	1.14	1.05-1.25	2.1×10^{-3}	0.031	0.02	0.70	0.43-1.14	0.15
HSD3B1	rs17023809	C	0.11	1.15	1.03-1.29	1.5×10^{-2}	0.047	0.02	0.81	0.48-1.36	0.42
	rs2362811	T	0.16	0.83	0.74-0.92	2.9×10^{-4}	0.011	0.44	1.02	0.90-1.16	0.73
	rs1885001	T	0.47	1.12	1.04-1.20	2.7×10^{-3}	0.014	0.31	0.95	0.83-1.09	0.47
	rs10754407	T	0.46	1.12	1.04-1.20	3.0×10^{-4}	0.014	0.31	0.96	0.84-1.10	0.58
	rs6685275	A	0.29	1.12	1.03-1.21	5.4×10^{-3}	0.023	0.07	0.74	0.56-0.97	0.03
	rs2011191	A	0.40	0.91	0.84-0.98	1.3×10^{-2}	0.043	0.32	1.08	0.94-1.24	0.27
	rs951693	G	0.47	1.12	1.04-1.20	2.8×10^{-3}	0.014	0.31	0.94	0.82-1.8	0.41
	rs6676283	C	0.37	1.12	1.04-1.21	2.5×10^{-3}	0.014	0.07	0.80	0.62-1.03	0.08
	rs6686779	C	0.24	0.89	0.81-0.97	8.0×10^{-3}	0.030	0.34	1.04	0.91-1.19	0.57
	rs6205	G	0.37	1.12	1.04-1.21	2.5×10^{-3}	0.014	0.08	0.82	0.64-1.05	0.12
	rs2362810	T	0.29	0.88	0.81-0.96	2.5×10^{-3}	0.014	0.49	0.95	0.83-1.08	0.41
	rs4659181	C	0.41	1.10	1.02-1.19	1.0×10^{-2}	0.036	0.29	0.95	0.82-1.09	0.43
	rs7553527	G	0.11	0.79	0.70-0.90	4.8×10^{-4}	0.011	0.47	0.99	0.87-1.12	0.84
	rs2208382	C	0.50	1.12	1.05-1.21	1.7×10^{-3}	0.014	0.30	0.94	0.81-1.07	0.34
	rs6668763	A	0.50	1.12	1.05-1.21	1.6×10^{-3}	0.014	0.30	0.93	0.81-1.06	0.27
HSD17B3	rs7019471	T	0.15	1.20	1.09-1.33	2.8×10^{-4}	0.027	0.02	0.92	0.54-1.55	0.74
SULT1A1	rs2411453	C	0.28	1.15	1.06-1.25	1.1×10^{-3}	0.002	0.46	1.06	0.93-1.21	0.37
UGT2B15	rs17146785	G	0.28	0.88	0.81-0.96	2.4×10^{-3}	0.022	0.04	1.00	0.72-1.39	0.99
	rs983347	A	0.28	0.88	0.81-0.96	3.7×10^{-3}	0.022	0.04	0.99	0.71-1.37	0.94
	rs983346	A	0.28	0.88	0.81-0.96	3.4×10^{-3}	0.022	0.04	1.00	0.72-1.39	0.99
	rs294754	A	0.22	0.86	0.79-0.94	8.3×10^{-4}	0.020	0.02	1.13	0.74-1.74	0.57
	rs1513559	G	0.46	0.91	0.85-0.98	1.4×10^{-2}	0.049	0.18	0.99	0.84-1.17	0.92
	rs1849536	T	0.37	1.10	1.02-1.19	1.1×10^{-2}	0.045	0.45	1.07	0.94-1.21	0.34

Abbreviations: SNP= single nucleotide polymorphism; OR=odds ratio; CIs = confidence intervals; *Q*-value, false discovery rate (FDR) *Q*-value; MAF=minor allele frequency; AA=African Americans; His=Hispanics.[1]Unconditional logistic regression models were used to estimate ORs (additive genetic model) adjusted for age (continuous), region, and 3 principal components for global ancestry.

Table 3. Associations of the significant SNPs among Hispanic women with clinical diabetes risk, Comparing results for Hispanic women (n=3,498) and African-American women (n=8,180).

Gene(s)	SNP ID	Minor allele	MAF in His	OR[1] in His	95% CI in His	*P*-value in His	*Q*-value in His	MAF in AA	OR[1] in AA	95% CI in AA	*P*-value in AA
CYP3A4	rs2737418	T	0.03	1.58	1.14-2.18	6.1×10^{-3}	0.0488	0.29	1.02	0.94-1.10	0.66
	rs473706	T	0.03	1.67	1.17-2.39	5.1×10^{-3}	0.0488	0.20	1.02	0.93-1.11	0.75
PGR	rs4754732	G	0.23	0.67	0.57-0.78	8.5×10^{-7}	<0.0001	0.10	0.97	0.86-1.10	0.68
	rs948516	G	0.23	0.66	0.56-0.78	6.5×10^{-7}	<0.0001	0.11	0.97	0.86-1.10	0.63
	rs10501972	T	0.23	0.67	0.57-0.79	1.3×10^{-6}	<0.0001	0.07	1.04	0.90-1.20	0.60
	rs632542	T	0.25	0.73	0.63-0.85	6.6×10^{-5}	0.0010	0.10	1.03	0.91-1.17	0.62
	rs1942845	C	0.23	0.69	0.58-0.81	5.8×10^{-5}	0.0002	0.15	1.09	0.99-1.20	0.09
	rs522819	A	0.26	0.71	0.61-0.84	2.8×10^{-5}	0.0006	0.10	1.04	0.91-1.18	0.56
	rs485283	G	0.25	0.73	0.63-0.85	6.9×10^{-5}	0.0010	0.12	1.00	0.89-1.12	0.98

Abbreviations: SNP= single nucleotide polymorphism; OR=odds ratio; CI = confidence intervals; FDR=false discovery rate; MAF=minor allele frequency; AA=African Americans; His=Hispanics.

[1]Unconditional logistic regression models were used to estimate ORs (additive genetic model) adjusted for age (continuous), region, and 3 principal components for global ancestry.

Table 4. Associations of the significant SNPs among African-American women or Hispanic women with clinical diabetes risk, Comparing results for European-American women (n=3,203), African-American women (n=8,180), and Hispanic women (n=3,498).

Gene(s)	SNP ID	I/G	Coded allele	CAF in EA	OR[1] in EA	95% CI in EA	CAF in AA	OR[1] in AA	95% CI in AA	CAF in His	OR[1] in His	95% CI in His
ESRα	rs17081685	G	C	0.07	0.93	0.76-1.13	0.20	1.17	1.07-1.28	0.24	1.01	0.88-1.17
HSD3B1	rs2362811	I	T	0.43	1.02	0.92-1.13	0.16	0.83	0.74-0.92	0.44	1.02	0.90-1.16
	rs1885001	I	T	0.68	1.14	1.02-1.27	0.47	1.12	1.04-1.20	0.31	0.95	0.83-1.09
	rs10754407	I	T	0.32	0.88	0.79-0.98	0.46	1.12	1.04-1.20	0.31	0.96	0.84-1.10
	rs2011191	G	A	0.35	0.87	0.78-0.97	0.40	0.91	0.84-0.98	0.32	1.08	0.94-1.24
	rs951693	I	G	0.32	0.88	0.79-0.98	0.47	1.12	1.04-1.20	0.31	0.94	0.82-1.8
	rs6686779	I	C	0.65	1.15	1.03-1.28	0.24	0.89	0.81-0.97	0.34	1.04	0.91-1.19
	rs2362810	I	T	0.43	1.02	0.92-1.13	0.29	0.88	0.81-0.96	0.49	0.95	0.83-1.08
	rs4659181	I	C	0.66	1.11	0.99-1.23	0.41	1.10	1.02-1.19	0.29	0.95	0.82-1.09
	rs7553527	G	G	0.45	1.05	0.94-1.16	0.11	0.79	0.70-0.90	0.47	0.99	0.87-1.12
	rs2208382	I	C	0.32	0.88	0.79-0.98	0.50	1.12	1.05-1.21	0.30	0.94	0.81-1.07
	rs6668763	I	A	0.32	0.88	0.78-0.98	0.50	1.12	1.05-1.21	0.30	0.93	0.81-1.06
PGR	rs4754732	I	G	0.32	0.93	0.83-1.04	0.10	0.97	0.86-1.10	0.23	0.67	0.57-0.78
	rs948516	G	G	0.32	0.93	0.83-1.04	0.11	0.97	0.86-1.10	0.23	0.66	0.56-0.78
	rs10501972	I	T	0.31	0.93	0.83-1.04	0.07	1.04	0.90-1.20	0.23	0.67	0.57-0.79
	rs632542	I	T	0.35	0.94	0.84-1.05	0.10	1.03	0.91-1.17	0.25	0.73	0.63-0.85
	rs1942845	G	C	0.29	0.91	0.82-1.02	0.15	1.09	0.99-1.20	0.23	0.69	0.58-0.81
	rs485283	I	G	0.36	0.93	0.84-1.04	0.12	1.00	0.89-1.12	0.25	0.73	0.63-0.85
SULT1A1	rs2411453	I	C	0.61	1.12	0.95-1.32	0.28	1.15	1.06-1.25	0.46	1.06	0.93-1.21
UGT2B15	rs1513559	I	G	0.10	1.03	0.86-1.23	0.46	0.91	0.85-0.98	0.18	0.99	0.84-1.17
	rs1849536	I	T	0.78	0.93	0.82-1.06	0.37	1.1	1.02-1.19	0.45	1.07	0.94-1.21

Abbreviations: SNP= single nucleotide polymorphism; OR=odds ratio; CI = confidence interval; FDR=false discovery rate; CAF=coded allele frequency; EA= European-American; AA=African Americans; His=Hispanics.

[1]Unconditional logistic regression models were used to estimate ORs (additive genetic model) adjusted for age (continuous), region, and 3 principal components for global ancestry in WHI-SHARe (AA and His) and matching factors and 3 principal components in WHI-GARNET (EA).

[2]Imputed (I) or genotyped (G).

Supplementary Table 1. Number of SNPs covered by the Affymetrix 6.0 array

Gene(s)	Number of SNPs on Affy6.0
AKR1C1	26
AKR1C2	31
AKR1C3	50
AKR1C4	112
AKR1D1	110
AR	89
COMT	17
CYP1A1	6
CYP1B1	40
CYP3A4	16
CYP3A5	9
CYP11A1	14
CYP17A1	11
CYP19A1	86
ESR1	171
ESR2	28
HSD3B1	46
HSD3B2	40
HSD17B1	1
HSD17B2	104
HSD17B3	99
PGR	102
SHBG	5
SRD5A1	37
SRD5A2	41
SULT1A1	2
SULT1E1	50
SULT2A1	13
UGT1A1	7
UGT2B15	24
UGT2B17	1

Abbreviations: Affy6.0= Affymetrix 6.0 array

Supplementary Table2 . SNPs-set enrichment analysis: Associations between sex hormone pathways genes and diabetes risk in African American and Hispanic women

Gene	Empirical *P*-values in AA	Empirical *P*-values in His
AKR1C1	>0.99	>0.99
AKR1C2	0.23	>0.99
AKR1C3	0.71	0.22
AKR1C4	0.76	0.37
AKR1D1	0.62	0.13
AR	0.04	0.06
COMT	>0.99	>0.99
CYP1A1	>0.99	>0.99
CYP1B1	0.10	>0.99
CYP3A4	>0.99	0.07
CYP3A5	>0.99	>0.99
CYP11A1	>0.99	0.10
CYP17A1	0.13	>0.99
CYP19A1	0.73	>0.99
ERα	0.02	0.51
ERβ	0.12	0.37
HSD3B1	0.004	0.07
HSD3B2	0.09	0.04
HSD17B1	>0.99	>0.99
HSD17B2	0.73	0.048
HSD17B3	0.04	0.71
PGR	0.30	0.002
SHBG	>0.99	>0.99
SRD5A1	0.08	0.49
SRD5A2	0.51	>0.99
SULT1A1	0.002	>0.99
SULT1E1	0.13	0.54
SULT2A1	>0.99	>0.99
UGT1A1	>0.99	0.07
UGT2B15	0.03	>0.99
UGT2B17	>0.99	>0.99

Abbreviations: OR=odds ratio; AA=African Americans; His=Hispanics.
A gene-level score was directly computed from averaging SNP-level test statistics using unconditional logistic regression analysis with adjustment of age, geographic regions, and the top 3 principal components from the PCA-based method. After pruning SNPs in high LD (R^2>0.5), up to the top 5 SNPs with P-value smaller than 0.05 were selected. The same analysis was performed with 1,000 simulated SNP sets in which the phenotype statuses of the individuals were permuted. An empirical P-value for the SNP set was computed by calculating the number of times the test statistic of the simulated SNP sets exceeded that of the original SNP set.

Figure legends

Supplementary Figure 1A and B. Regional association plots.
Regional association plots for *HSD3B1* and *UGT2B15* among African-Americans.

Supplementary Figure 1C and D. Regional association plots.
Regional association plots for *CYP3A4* and *PGR* among Hispanics.

Supplementary Figure 2A-E. Forest plots for significant single nucleotide polymorphisms by population.
The overall estimates were obtained using a random-effects model. The dots indicate the odds ratios for each population and the horizontal lines represent 95% confidence intervals (CIs). Abbreviations: EA, European Americans; AA, African Americans; His, Hispanics; ALL, random-effects meta-analysis of the three populations.

Supplementary Figure 1A-D. Regional association plots

Figure 1A. Regional association plots for *HSD3B1* among African-Americans

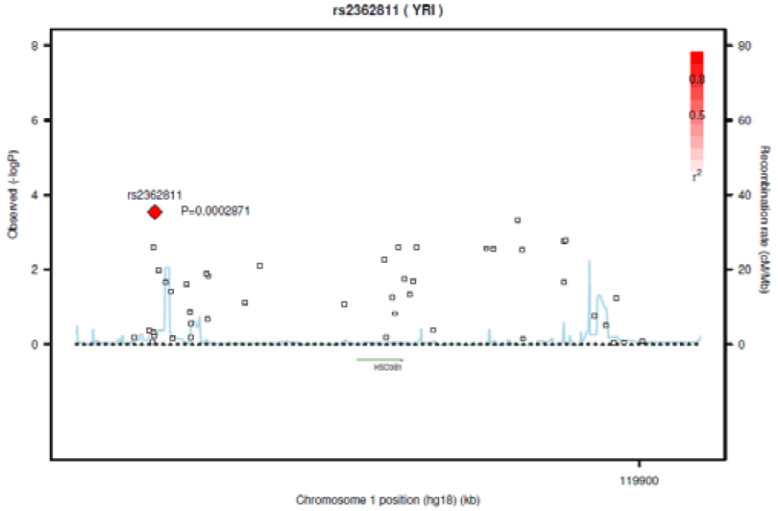


Figure 1B. Regional association plots for *UGT2B15* among African-Americans

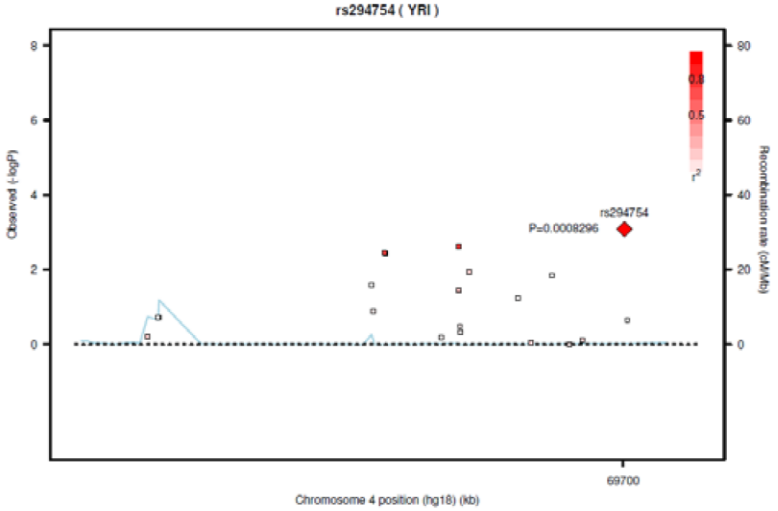


Figure 1C. Regional association plots for *CYP3A4* among Hispanics

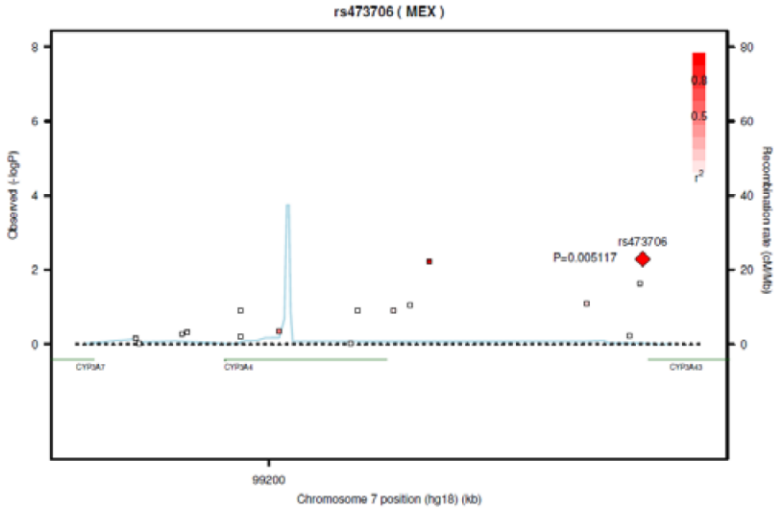
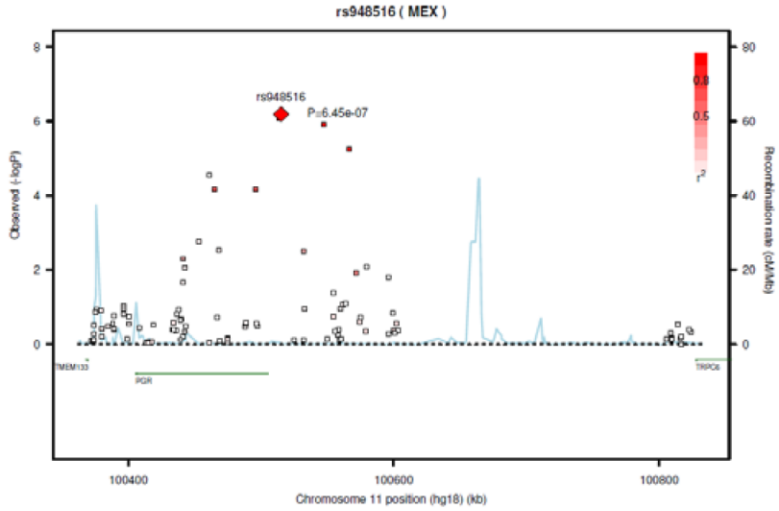
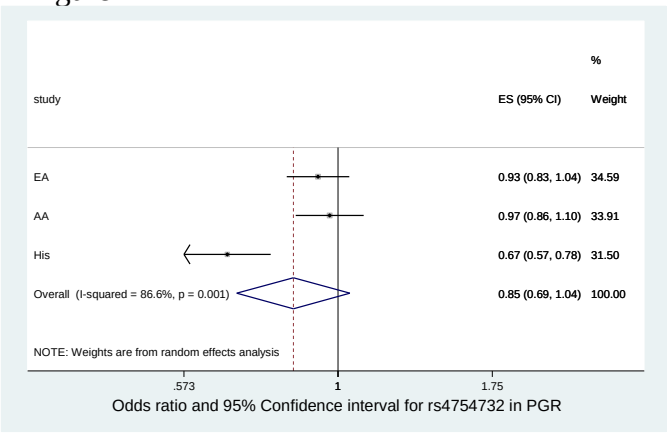


Figure 1D. Regional association plots for *PGR* among Hispanics



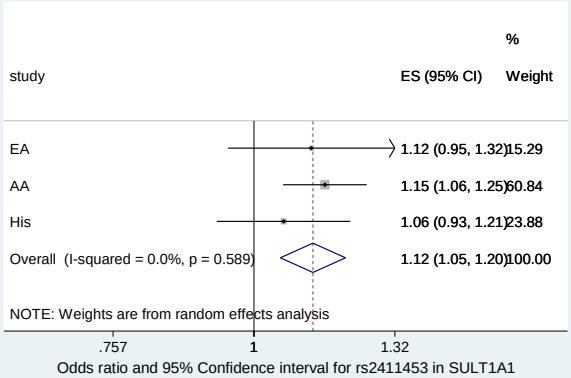
Supplementary Figure 2A-E Forest plots for significant single nucleotide polymorphisms by population

Figure 2A



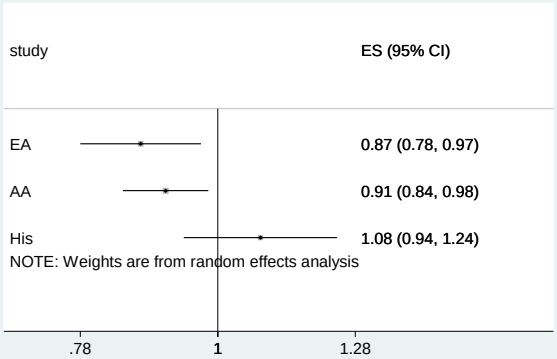
Odds ratio and 95% Confidence interval for rs4754732 in PGR

Figure 2B



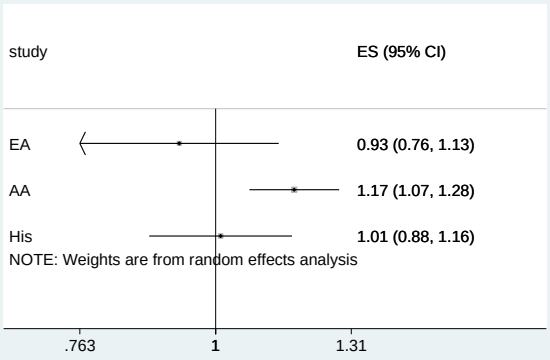
Odds ratio and 95% Confidence interval for rs2411453 in SULT1A1

Figure 2C



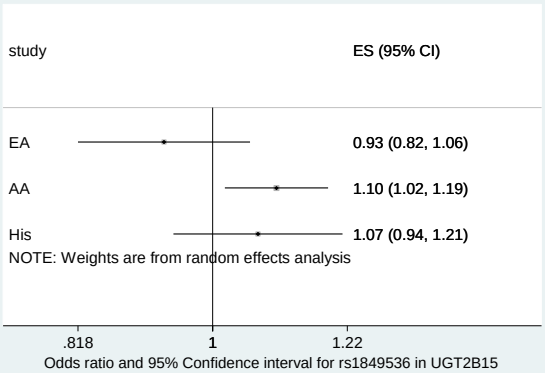
Odds ratio and 95% Confidence interval for rs2011191 in HSD3B1

Figure 2D



Odds ratio and 95% Confidence interval for rs17081685 in ERa

Figure 2E



Odds ratio and 95% Confidence interval for rs1849536 in UGT2B15

4.6 References

1. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA.* Mar 15 2006;295(11):1288-1299.

2. Ding EL, Song Y, Manson JE, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N Engl J Med.* Sep 17 2009;361(12):1152-1163.

3. Rebarber A, Istwan NB, Russo-Stieglitz K, et al. Increased incidence of gestational diabetes in women receiving prophylactic 17alpha-hydroxyprogesterone caproate for prevention of recurrent preterm delivery. *Diabetes care.* Sep 2007;30(9):2277-2280.

4. Ahn J, Schumacher FR, Berndt SI, et al. Quantitative trait loci predicting circulating sex steroid hormones in men from the NCI-Breast and Prostate Cancer Cohort Consortium (BPC3). *Hum Mol Genet.* Oct 1 2009;18(19):3749-3757.

5. Sober S, Org E, Kepp K, et al. Targeting 160 candidate genes for blood pressure regulation with a genome-wide genotyping array. *PLoS One.* 2009;4(6):e6034.

6. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: A review of statistical methods and recommendations for their application. *American journal of human genetics.* Jan 2010;86(1):6-22.

7. Wang K, Li M, Bucan M. Pathway-Based Approaches for Analysis of Genomewide Association Studies. *Am J Hum Genet.* Oct 26 2007;81(6).

8. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods and prospects for future development. *Trends Genet.* Apr 3 2012.

9. Henderson BE, Ponder BAJ, Ross RK. *Hormones, genes, and cancer*. Oxford ; New York: Oxford University Press; 2003.

10. Margolis KL, Lihong Q, Brzyski R, et al. Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements. *Clin Trials.* 2008;5(3):240-247.

11. Jackson JM, Defor TA, Crain AL, et al. Self-reported diabetes is a valid outcome in pragmatic clinical trials and observational studies. *Journal of clinical epidemiology.* May 5 2012.

12. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* Aug 2006;38(8):904-909.

13. Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* Aug 5 2003;100(16):9440-9445.

14. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics.* Dec 15 2008;24(24):2938-2939.

15. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics.* Sep 2007;81(3):559-575.

16. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *American journal of human genetics.* Feb 2009;84(2):210-223.

17. Frayling TM. Genome-wide association studies provide new insights into type 2 diabetes aetiology. *Nature reviews. Genetics.* Sep 2007;8(9):657-662.

18. Skjaerpe PA, Giwercman YL, Giwercman A, Svartberg J. Androgen receptor gene polymorphism and the metabolic syndrome in 60-80 years old Norwegian men. *International journal of andrology.* Jun 1 2010;33(3):500-506.

19. Rodriguez-Gonzalez G, Ramirez-Moreno R, Perez P, et al. The GGN and CAG repeat polymorphisms in the exon-1 of the androgen receptor gene are, respectively, associated with insulin resistance in men and with dyslipidemia in women. *The Journal of steroid biochemistry and molecular biology.* Feb 2009;113(3-5):202-208.

20. Lin HY, Xu Q, Yeh S, Wang RS, Sparks JD, Chang C. Insulin and leptin resistance with hyperleptinemia in mice lacking androgen receptor. *Diabetes.* Jun 2005;54(6):1717-1725.

21. McInnes KJ, Smith LB, Hunger NI, Saunders PT, Andrew R, Walker BR. Deletion of the Androgen Receptor in Adipose Tissue in Male Mice Elevates Retinol Binding Protein 4 and Reveals Independent Effects on Visceral Fat Mass and on Glucose Homeostasis. *Diabetes.* Mar 13 2012.

22. Dahlman I, Vaxillaire M, Nilsson M, et al. Estrogen receptor alpha gene variants associate with type 2 diabetes and fasting plasma glucose. *Pharmacogenet Genomics.* Nov 2008;18(11):967-975.

23. Kanaya N, Murray PA, Damron DS. The differential effects of midazolam and diazepam on intracellular Ca2+ transients and contraction in adult rat ventricular myocytes. *Anesth Analg.* Dec 2002;95(6):1637-1644, table of contents.

24. Margolis KL, Bonds DE, Rodabough RJ, et al. Effect of oestrogen plus progestin on the incidence of diabetes in postmenopausal women: results from the Women's Health Initiative Hormone Trial. *Diabetologia.* Jul 2004;47(7):1175-1187.

25. Bonds DE, Lasser N, Qi L, et al. The effect of conjugated equine oestrogen on diabetes incidence: the Women's Health Initiative randomised trial. *Diabetologia.* Mar 2006;49(3):459-468.

26. Ribas V, Nguyen MT, Henstridge DC, et al. Impaired Oxidative Metabolism and Inflammation are Associated with Insulin Resistance in ER{alpha} Deficient Mice. *Am J Physiol Endocrinol Metab.* Nov 17 2009;298(2):E304-E319.

27. Ding EL, Song Y, Manson JE, Rifai N, Buring JE, Liu S. Plasma sex steroid hormones and risk of developing type 2 diabetes in women: a prospective study. *Diabetologia.* Oct 2007;50(10):2076-2084.

28. Barros RP, Gabbi C, Morani A, Warner M, Gustafsson JA. Participation of ERalpha and ERbeta in glucose homeostasis in skeletal muscle and white adipose tissue. *American journal of physiology. Endocrinology and metabolism.* Jul 2009;297(1):E124-133.

29. Shimodaira M, Nakayama T, Sato N, et al. Association of HSD3B1 and HSD3B2 gene polymorphisms with essential hypertension, aldosterone level, and left ventricular structure. *European journal of endocrinology / European Federation of Endocrine Societies.* Oct 2010;163(4):671-680.

30. Tchernof A, Levesque E, Beaulieu M, et al. Expression of the androgen metabolizing enzyme UGT2B15 in adipose tissue and relative expression measurement using a competitive RT-PCR method. *Clinical endocrinology.* May 1999;50(5):637-642.

31. Yamada Y, Matsuo H, Watanabe S, et al. Association of a polymorphism of CYP3A4 with type 2 diabetes mellitus. *Int J Mol Med.* Nov 2007;20(5):703-707.

32. Joslin EP. *Joslin's Diabetes mellitus*. 12th ed. Philadelphia: Lea & Febiger; 1985.

33. Picard F, Wanatabe M, Schoonjans K, Lydon J, O'Malley BW, Auwerx J. Progesterone receptor knockout mice have an improved glucose homeostasis secondary to beta -cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America.* Nov 26 2002;99(24):15644-15648.

34. Olson SH, Bandera EV, Orlow I. Variants in estrogen biosynthesis genes, sex steroid hormone levels, and endometrial cancer: a HuGE review. *American journal of epidemiology.* Feb 1 2007;165(3):235-245.

35. Yu K, Wang Z, Li Q, et al. Population substructure and control selection in genome-wide association studies. *PLoS ONE.* 2008;3(7):e2551.

36. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology*. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.

Chapter 5

Conclusions

5.1 Conclusions and Future Research Directions

Previous epidemiologic and experimental studies have identified obesity, diet (high fat and high glycemic index/load), physical inactivity, family history of type 2 diabetes (T2D), age, race/ethnicity, smoking, history of gestational diabetes, and sex as risk factors for T2D.[1-3] However, there is an accelerating increase in the prevalence of T2D worldwide and the increase in developing countries is estimated to be substantial.[4-6] Moreover, the prevalence of T2D is increasing even in developed countries. This trend suggests the need to translate these findings into the community and may also indicate an incomplete understanding of the pathogenesis of T2D. Thus, the pathogenesis of T2D must be better understood to curb the escalating diabetes epidemic. Understanding the roles of sex hormones in the development of T2D and determinants of sex hormone-binding globulin (SHBG) are vital in terms of the preventing of T2D because mechanisms underlying the relations between sex hormones or the SHBG levels, and the risk of T2D may possibly become therapeutic targets, and determinants of the circulating SHBG levels could potentially be important factors in preventing T2D in the future. This dissertation examined the roles of sex hormones and SHBG in the development of T2D as well as the potential determinants of the SHBG levels.

Chapter 1 introduced general background information on sex hormones and SHBG in relation to the pathogenesis of T2D. Chapter 2 explored whether the levels of sex hormones and SHBG may play a role in the inverse relation between coffee intake and the risk of T2D in the

Women's Health Study. Chapter 3 examined potential determinants of the circulating levels of SHBG in a pooled analysis of studies included in the Women's Health Initiative Randomized Trial and Observational Study. Chapter 4 investigated the roles of sex hormone pathway genes in the development of T2D using data from the Women's Health Initiative SNP Health Association Resource (SHARe).

Our findings in Chapter 2 indicated that the observed inverse associations of caffeinated-coffee and caffeine intakes with the risk of T2D were substantially attenuated after taking into account of SHBG levels. These findings suggest that SHBG levels may account for the potential protective effect of habitual coffee consumption against the risk of T2D among postmenopausal women. However, whether this relationship was causal remain to be determined due to the nature of observational study. Clearly, further mechanistic and intervention studies are needed to determine the long-term effect of caffeinated-coffee and caffeine intakes on T2D risk. In addition, it is important to note that caffeinated coffee and caffeine intakes may have adverse effects, such as arrhythmias, dyslipidemia, pregnancy complications, and drug interactions.[7] Meanwhile, current public health recommendations on coffee consumption should not be changed.

The results from a pooled analysis of studies in the Women's Health Initiative described in Chapter 3 suggested that age, BMI, exogenous estrogen use, and lifestyle factors may be associated with the plasma levels of SHBG. If these factors determine the circulating levels of SHBG, substantial clinical implications would exist. Importantly, the SHBG levels are thought be involved in the development of various diseases, such as hormone dependent cancers, cardiovascular diseases, T2D, and fractures. Because BMI and lifestyle factors are potentially

modifiable, these findings might have important implications for the prevention of various diseases. However, more studies, including randomized controlled trials to examine the effects of lifestyle factors on SHBG levels and in *vivo* and in *vitro* experiments are clearly needed to confirm the notions.

In Chapter 4, we examined the associations between genes in sex hormone pathways and T2D risk. The findings suggested that genes in sex hormones synthesis and metabolism, estrogen receptors, and androgen receptor may be associated with T2D risk in African-Americans and genes encoding androgen metabolism and progesterone receptor may be associated with T2D risk in Hispanics. Importantly, the directions for the most significant single nucleotide polymorphisms (SNPs) were generally consistent across ethnic groups, suggesting that these loci may be contributed to the risk of T2D across different ethnic groups. These findings provide additional support for the roles of sex hormones in the development of T2D. To confirm that these variants are truly associated with T2D risk, it is necessary to replicate the findings in independent cohorts. Future work may also include fine-mapping of these loci, integrating gene expression analysis into analyses of SNPs in sex hormone pathways,[8] evaluating gene-gene and gene-environment interactions, and analyzing epigenetics of sex hormone pathways.

Although further studies are necessary to confirm the present findings, our studies provide additional supports for the roles of sex hormones and SHBG in the development of T2D, and the roles of several non-genetic factors as potential determinants of SHBG.

5.2 References

1. Joslin EP, Kahn CR. *Joslin's diabetes mellitus.* 14th ed. Philadelphia, Pa.: Lippincott Williams & Willkins; 2005.

2. Standards of medical care in diabetes--2012. *Diabetes care.* Jan 2012;35 Suppl 1:S11-63.

3. Hu FB, Manson JE, Stampfer MJ, et al. Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *The New England journal of medicine.* Sep 13 2001;345(11):790-797.

4. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res Clin Pract.* Nov 5 2009.

5. Roglic G, Unwin N. Mortality attributable to diabetes: Estimates for the year 2010. *Diabetes Res Clin Pract.* Nov 13 2009.

6. Thomas MC, Hardoon SL, Papacosta AO, et al. Evidence of an accelerating increase in prevalence of diagnosed Type 2 diabetes in British men, 1978-2005. *Diabet Med.* Aug 2009;26(8):766-772.

7. Higdon JV, Frei B. Coffee and health: a review of recent human research. *Crit Rev Food Sci Nutr.* 2006;46(2):101-123.

8. Zhong H, Yang X, Kaplan LM, Molony C, Schadt EE. Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *American journal of human genetics.* Apr 9 2010;86(4):581-591.