의학석사 학위논문

Genetic variation in obesity-related genes and risk for breast cancer

비만관련 유전자의 유전적 변이와 유방암 위험도 간의 연관성 연구

2014 년 2 월

서울대학교 대학원

의과학과 의과학전공

정 서 강

A thesis of the Master's degree

비만관련 유전자의 유전적 변이와 유방암 위험도 간의 연관성 연구

Genetic variation in obesity-related genes and risk for breast cancer

February 2014

Department of Biomedical Sciences,

Seoul National University

Graduate School

Seokang Chung

비만관련 유전자의 유전적 변이와 유방암 위험도 간의 연관성 연구

지도교수 최 지 엽

이 논문을 의학석사 학위논문으로 제출함

2013 년 10 월

서울대학교 대학원

의과학과 의과학전공

정 서 강

정서강의 의학석사 학위논문을 인준함

2014 년 1 월

위 원 장 (인)

부위원장 (인)

위 원 (인)

Genetic variation in obesity-related genes and risk for breast cancer

by

Seokang Chung

A thesis submitted to the Department of Biomedical Sciences in partial fulfillment of the requirements for the Degree of Master of Science in Medicine at Seoul National University College of Medicine

January 2014

Approved by Thesis Committee:

Professor _____Chairman

Professor _____Vice chairman

Professor _____

Abstract

Introduction: Recent increasing trends in overweight and obesity lead to great interest of their effect on complex diseases including breast cancer. Although many previous association studies have been addressed genetic variations in obesity related genes on breast cancer risk, they have been focusing on single SNP analysis as well as limited consideration of genes and gene-environmental interaction. We investigated the associations of known obesity-related genes on breast cancer risk by comprehensive assessment based on individual single-nucleotide polymorphism (SNP) analysis and pathway-analysis and also the modified effects of genetic factors on the associations between obesity and breast cancer risk.

Methods: We first carried out genome-wide association study (GWAS) to identify genetic loci with body mass index (BMI) in 1,786 breast cancer cases and 1,789 healthy controls from the multicenter case-control study, based on previously known genes from GWAS catalog (http://www.genome.gov/gwastudies) and published literatures associated with obesity phenotype. Associations of breast cancer risk and SNP with BMI were assessed by logistic regression models under an additive genetic model and also the gene-based pathway was assessed by the adaptive rank-truncated product (ARTP) method with 1,000 permutations. The analyses were conducted in overall, and in the strata of menopausal status, obesity level and hormone receptor status.

Results: Among available 227,197 SNPs from candidate genes and additional significant SNPs from these participants, 9,985 SNPs in 712 genes were found

to be associated with BMI at p<0.05: 408 genes (45.9%) from GWAS catalog and 221 genes (38.1%) from the literatures for obesity were identified, respectively. Rs17804012 in *RBFOX1* (16p13.3) and rs2014791 in *LINC00317* (21q21.1) showed the most significant association with BMI (p<5E-6), which observed null associations with breast cancer risk. In contrast, the gene-based pathway analysis including less significantly associated genes with BMI (N=644, p<0.05) showed significant associations with breast cancer risk (the pathway p=3E-3). Especially, the effect of *THRB* gene on breast cancer risk had highly significant value (p=9E-4) and the association was confined to premenopausal women with BMI above 23.2 (p=5E-3). The effect of rs6550865, the representative SNP in *THRB*, was mainly associated with either estrogen receptor (ER) or progesterone receptor (PR) positive breast cancer risk when stratified analysis by hormone receptor subtypes but the difference between ER/PR subtypes did not reach significant ($P_{heterogeneity}$>0.05).

Conclusions: The group of BMI-related genes including *THRB* was significantly associated with breast cancer risk in the pathway analysis and the associations were different depending on the menopausal status and/or BMI levels. Our results suggest that consideration of the complex genetic pathway for obesity might reveal additional breast cancer susceptibility loci.

Keywords: Breast cancer, obesity-related genes, pathway analysis, polymorphism

Student number: 2012-21811

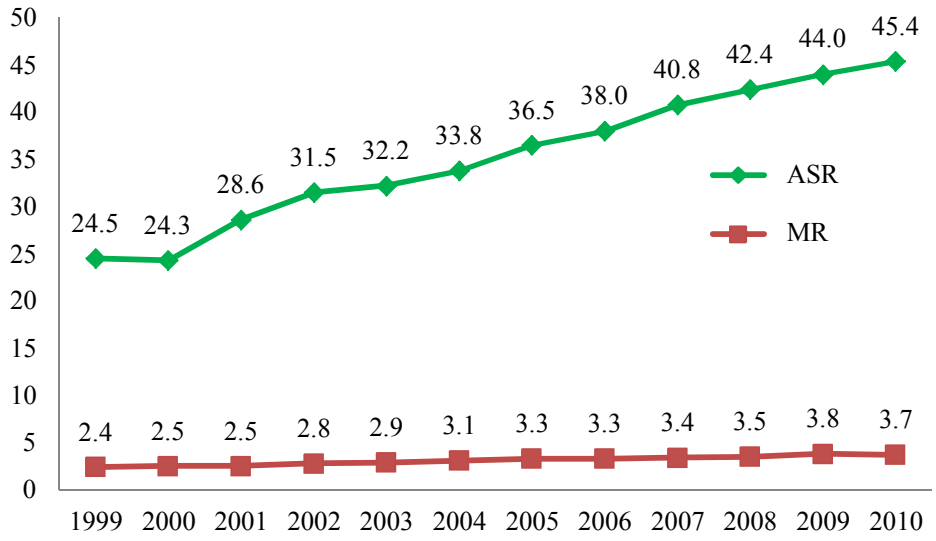
Contents

List of Tables

List of Figures

Introduction

Breast cancer is one of the rapidly increasing cancers worldwide including Korea; 14,208 new Korean female cases of breast cancer developed and 1,868 women died of this disease in 2010 [1-3]. In Korean women, the incidence of breast cancer has steadily increased since 1999, and the age-standardized incidence rates were estimated at 24.5/100,000 in 1999 and 45.4/100,000 in 2010 (Figure 1) [4]. Although the incidence rate for Korean female is lower than those in Western countries, the average annual change of cumulative probability for breast cancer in Korea has been increasing sharply at an alarming rate at 4%, which is 4-fold higher than the global cumulative probability [5]. Furthermore, the morality rates from breast cancer has still been increasing in Korea while chronological reduction of those in many Western countries was observed (Figure 1) [6].



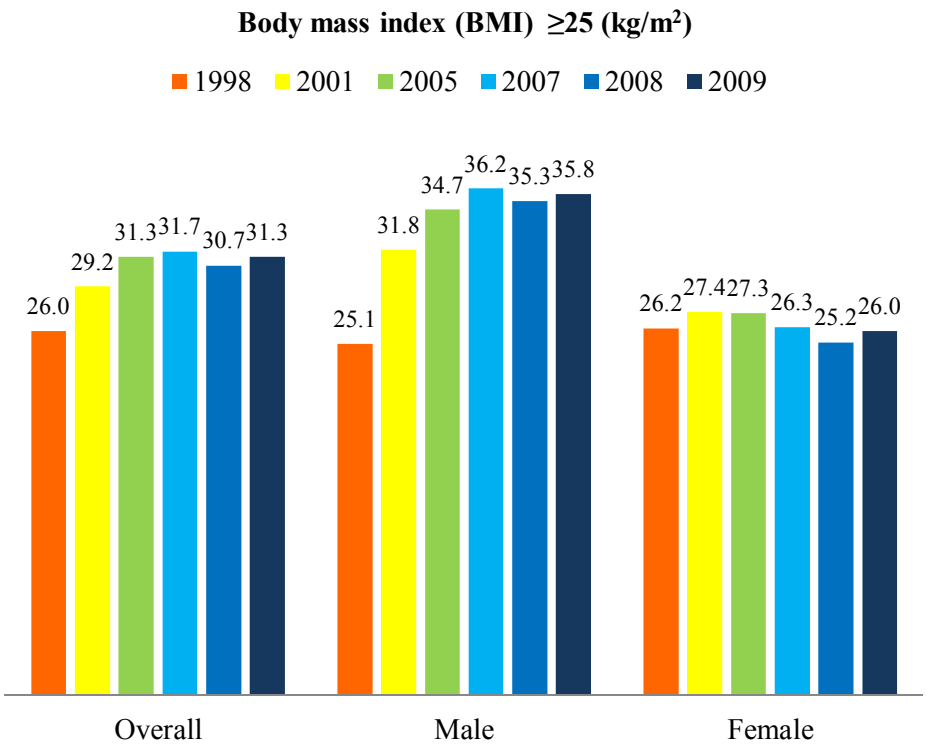
*Modified from annual report of cancer statistics in Korea in 2010, Korea Central Cancer Registry [4] and report on the statistics for National Cancer Incidence in 2009, National Cancer Center [6]

ASR, Age-standardized incidence rate; MR, mortality rate

Figure 1. Trends in age-standardized incidence rate and mortality rate of breast cancer among Korean women

1. Increased prevalence of obesity

The prevalence of overweight and obesity has been increasing worldwide; recent data in the US showed that almost 40% and 30% of adult men and women fall into the overweight category ($25 \leq BMI < 30$ (kg/m^2)) [7]. Several reports based on Korea National Health Examination and Nutrition Survey data have been addressed on obesity trends in Korea; the prevalence rates of obesity defined as equal more than 25 of BMI (kg/m^2) in Korean adults increased from 26.0% in 1998 to 31.3 in 2009 (Figure 2), which change of those of BMI was confined to men [8-11]. Khang et al. revealed the shifts to the right in the distribution of waist circumference between 1998 and 2007 among Korean adults regardless of gender [10]. This major public health challenge results in many GWAS studies with obesity (Table 1) and great interest of effect of the obesity on the trait or diseases related to obesity as well [12-16].



*Data from Korea National Health and National Examination Survey, 2013 [11]

(Age-standardized rates based on 2005 Korean population)

Figure 2. Trends of obesity prevalence in Korean adults

Table 1. Variants associated with BMI in previous genome-wide association studies

Nearest gene	SNP	Chr	Position	EA	BMI results		Ethnicity	N	Comparison category	Ref
					Effect	*P*				
NEGR1	rs3101336	1	72523773	G	4.56	1.1.E-7	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs25698958	1	72537704	A	4.66	9.9.E-8	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs2815752	1	72524461	A	0.10*	9.3.E-6	European	32,387	Per-allele change in BMI	Willer et al, 2016
SEC16B, RASAL2	rs10913469	1	176180142	C	4.29	4.2.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
TMEM18	rs2867125	2	612827	G	6.83	1.1.E-10	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs4854344	2	628144	T	6.33	2.9.E-10	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs7561317	2	634953	G	6.42	2.4.E-10	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs6548238	2	624905	C	0.26*	1.2.E-6	European	32,256	Per-allele change in BMI	Willer et al, 2010
SFRS10, ETV5, DGKG	rs7647305	3	187316984	C	4.66	3.1.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
APBB2	rs7697609	4	40651331	T	1.63	4.7.E-6	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
APBB3	rs6857327	4	40649029	T	1.60	7.8.E-6	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
GNPDA2	rs10938397	4	45023455	G	0.19*	1.0.E-5	European	32,387	Per-allele change in BMI	Willer et al, 2012
TMEM161B	rs247916	5	87566009	C	0.63	9.2.E-6	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
LGR4, LIN7C, BDJF	rs4074134	11	27603861	G	5.77	1.8.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs4923461	11	27613486	A	5.89	1.1.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008

	rs925946	11	27623778	T	4.90	2.0.E-7	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs10501087	11	27626684	T	5.68	4.2.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs6265	11	27636492	G	5.67	7.2.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
MTCH2	rs10838738	11	47619625	G	0.07*	7.1.E-6	European	32,387	Per-allele change in BMI	Willer et al, 2014
BCDIN3D, FAIM2	rs7138803	12	48533735	A	4.89	9.6.E-7	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
FTO	rs3751812	16	52375961	T	1.64	2.0.E-8	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs8050136	16	52373776	C	1.63	3.0.E-8	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs9941349	16	52382989	T	1.61	6.5.E-8	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs9930333	16	52357478	T	1.60	7.9.E-8	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs10852521	16	52362466	T	0.65	6.3.E-7	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs7190492	16	52386253	G	0.64	1.5.E-6	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs8044769	16	52396636	T	0.66	2.4.E-6	Caucasian	1,060	BMI >35 vs BMI<25	Wang et al, 2011
	rs9939609	16	52378028	A	0.33*	6.3.E-17	European	32,329	Per-allele change in BMI	Willer et al, 2009
RPGPRIP1L, FTO	rs6499640	16	52327178	A	5.55	6.0.E-8	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs8050136	16	52373776	A	8.28	4.4.E-24	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs3751812	16	52375961	T	8.92	3.3.E-24	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs7190492	16	52386253	G	6.85	2.0.E-12	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs8044769	16	52396636	C	6.85	7.9.E-16	Mixed	337,238	BMI ≥30 vs	Gudmar et al,

									18.5≤BMI≤25	2008
SH2B1, ATP2A1	rs8049439	16	28745016	C	3.92	6.0.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs478102	16	28780899	A	4.03	3.5.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs7498665	16	28790742	G	4.25	1.7.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs7498665	16	28790742	G	0.15*	5.4.E-6	European	32,361	Per-allele change in BMI	Willer et al, 2013
MC4R	rs12970134	18	56035730	A	4.23	2.6.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
	rs17782313	18	560020077	C	0.20*	3.9.E-7	European	32,385	Per-allele change in BMI	Willer et al, 2011
CHS78M, KCTD15	rs29941	19	39001372	C	4.25	5.6.E-6	Mixed	337,238	BMI ≥30 vs 18.5≤BMI≤25	Gudmar et al, 2008
KCTD15	rs11084753	19	39013977	G	0.06*	2.6.E-7	European	32,335	Per-allele change in BMI	Willer et al, 2015

* Beta estimates

2. Established risk factors for breast cancer development

Lifestyles including reproductive factors and obesity have been recognized as known risk factors for breast cancer from previous epidemiological studies (Table 2) [17-19]; longer duration of exposure to endogenous sex hormone determined by early age at menarche, late age at first full-term pregnancy and/or late age at menopause have been studied well in relation to increased risk of breast cancer. Also, those factors changed into bad side, i.e. early age at menarche, late age at first delivery and never having children caused by hormonal influence, which might be contributed to the rising trends in breast cancer incidences in Korea [20-24].

On the other hand, the effect of obesity on breast cancer by menopausal status is moderate, although many previous studies of the association between obesity and risk of breast cancer were reported by menopausal status [25-27].

Table 2. Summary of risk factors of breast cancer

Breast cancer risk factors	Magnitude of risk
Well established risk factors of breast cancer	
Increasing age	↑↑
Family history of breast cancer among 1st degree relatives	↑↑
Mutations in BRCA1 and BRCA2 genes	↑↑
History of benign breast diseases	↑↑
Late age at menopause	↑↑
Early age at menarche	↑↑
Nulliparity and older age at first birth	↑↑
High mammographic density of breast	↑↑
Hormonal replacement therapy	↑
Recent use of oral contraceptives	↑
Obesity in postmenopausal women	↑
Alcohol consumption	↑
Well established protective factors of breast cancer	
Early age at first birth	↓↓
Higher number of parity	↓↓
Longer duration of breastfeeding	↓↓
Obesity in premenopausal women	↓
Physical activity	↓
Intake of fruits and vegetables	↓

*Modified from Duitrescu et al, 2005 [17], Kelsey et al, 1993 [18], and McPherson et al, 2000 [19]

(↑↑, moderate to high increase in risk; ↑, low to moderate increase in risk; ↓↓, moderate to high decrease in risk; ↓, low to moderate decrease in risk)

3. Associations between obesity and risk of breast cancer

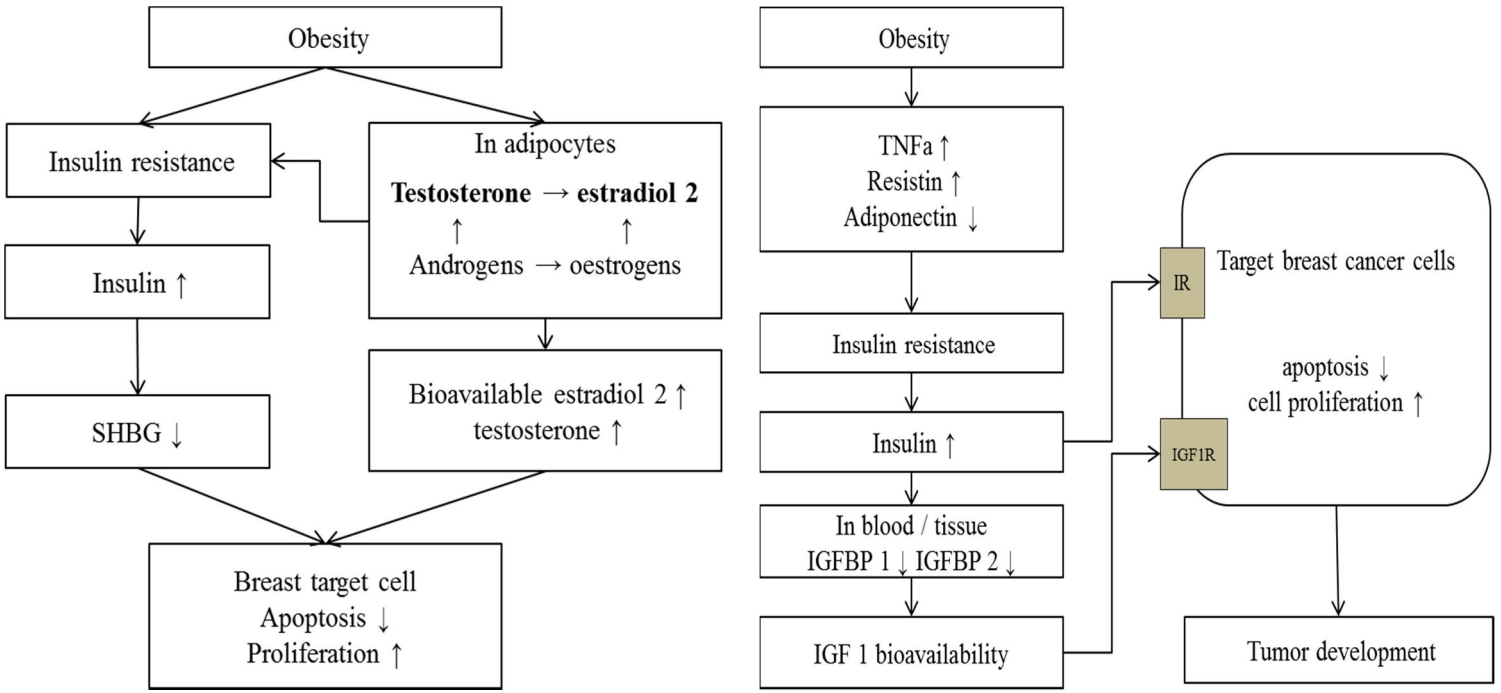
In terms of the associations between obesity and breast cancer, the opposite direction of the associations by menopausal status was observed in previous studies (Table 3). A controversial association with BMI on breast cancer risk was mainly suggested among premenopausal women, while the increased risk of breast cancer in obese women compared with normal weight women was mainly reported among postmenopausal women, although no relationship between BMI and breast cancer risk was observed in recent case-control studies [20,21,26,28-37].

Two major mechanisms is mainly suggested for the association of obesity with breast cancer risk; hormonal and non-hormonal mechanisms, respectively (Figure 3) [38]. In the hormonal mechanism, the effect of obesity on risk of breast cancer was explained by increasing available estrogen level with reduced level of sex hormone binding globulin (SHBG) and the influence is different by menopausal status. This excessive stimulation among postmenopausal women as a result of enhanced oestrogen synthesis by adipose tissue was observed while plasma estrogen concentrations are not influenced directly by body fat mass among premenopausal women since most of the estrogens produced in the ovaries and the circulating levels are under homeostatic regulation [39]. The non-hormonal mechanisms include effects of obesity on tumorigenesis by increased levels of insulin-like growth factors1 (IGF1) and insulin through growth-factor production [38].

Table 3. Previous association studies between BMI and breast cancer risk by menopausal status

	Ethnicity	N	BMI group, kg/m^2	Source of BMI	Premenopausal women			Postmenopausal women		
					N	RR (OR)	95% CI	N	RR (OR)	95% CI
Prospective cohort studies		(No.count)								
Cecchini et al, 2012 [28]	Mixed	31,731	≥30 vs <25	Measured	38 / 43	1.70	1.10-2.63	331 / 194	1.14	0.94-1.38
White et al, 2012 [29]	Mixed	82,971	≥30 vs 20-24.9	Self-reported	-	-	-	329 / 316	1.60	1.36-1.87
Harris et al, 2011 [30]	Mixed	45,799	≥27.5 vs <20.5	Self-reported	96 / 132	0.75	0.57-0.99	-	-	-
Reeves et al, 2007 [31]	UK	1,222,630	≥30 vs 22.5-24.9	Self-reported	166 /271	0.79	0.68-0.92	1,274 / 1,336	1.29	1.22-1.36
Ahn et al, 2007 [32]	Mixed	99,039	30.0-34.9 vs 18.5-22.4	Self-reported	-	-	-	175 / 134	1.55	1.22-1.96
Palmer et al, 2007 [33]	Black	59,000	≥35 vs <25	Self-reported	71 / 157	0.72	0.54-0.96	38 / 32	0.94	0.58-1.54
Lahmann et al, 2004 [26]	European	176,886	≥28.8 vs <21.6	Measured	68 / 132	0.82	0.59-1.14	239 / 98	1.36	1.06-1.75
Morimoto et al, 2002 [21]	Mixed	85,917	≥31.1 vs ≤22.6	Measured	.	.	.	103 / 37	2.52	1.62-3.93
Van den et al, 2000 [20]	Mixed	337,819	29-30.9 vs <21	Self-reported	32 / 158	0.96	0.60-1.52	224 / 363	1.21	1.01-1.46
Case-control studies		(Case/ control)								
John et al, 2001 [34]	Mixed	672 / 808	≥30 vs <25	Self-reported	179 / 298	0.60	0.45-0.79	-	-	-
Berstad et al, 2010 [35]	Mixed	3,997 / 4,041	30-34.9 vs <25	Self-reported	168 /1,342	0.86	0.68-1.09	254 / 918	0.95	0.77-1.16

Ogundiran et al, 2010 [36]	Nigerian women	1,223 / 1,101	≥28 vs <21.0	Measured	187 / 153	0.70	0.50-0.98	151 / 100	0.76	0.48-1.21
Boyd et al, 2006 [37]	Mixed	1,114 / 1,114	>27.64 vs ≤21.79	Self-reported	51 / 86	0.76	0.50-1.30	180 / 159	1.17	0.90-1.60



SHBG, Sex hormone binding globulin; TNFa, tumor necrosis factor-alpha; IGFBP, insulin-like growth factor-binding protein; IGF, insulin-like growth factor

*Source from Calle et al, 2004 [38]

Figure 3. Proposed mechanisms of obesity on the risk of breast cancer

4. Previous association studies on obesity related genetic factors and breast cancer risk

Because of the importance of maintaining genomic integrity in carcinogenesis, many association studies also have assessed genetic variations in obesity related genes, such as leptin (*LEP*) and fat mass and obesity (*FTO*) on risk for breasts cancer [40-43] (Table 4). Among them, the *FTO* gene, one of the representative genes associated with obesity, is also considered powerful classify to predict breast cancer risk in a case-control study by Kalklamani et al; the epistatic model including *FTO* genotypes further improved the prediction accuracy with higher area under the receiver operating characteristics curve (AUC) than nonepistatic model (AUC=0.68) [40].

However, previous studies focused on a limited number of genes and interaction between genetic polymorphisms of various and environmental exposure which may contribute to breast cancer development. Most association studies also have identified just a few SNPs at high level of significance using approaches of individual SNP-analysis, and they often overlooked combined associations by gene-level analysis [43,44]. In the view of gene x gene interaction and gene x environment interaction, it is expected to find the noble association between genes related to obesity and breast cancer risk by comprehensive assessment.

Table 4. Previous association studies on obesity-related genetic factors and breast cancer risk

Gene	SNP	Chr	Position	Study design	Ethnicity	N (case/ control)	EA	Breast cancer risk		Effect model	Ref
								OR	95 % CI, *P*		
FTO	rs1121980	16	53809247	Case-control	Mixed	100 / 148	C	0.60	0.037	Additive	[42]
	rs9969609	16	1355082	Case-control	Mixed	100 / 148	T	Null association			[42]
	rs1477196	16	53808258	Case-control	Mixed	354 / 364	G	2.61	1.56-4.37	Recessive	[40]
	rs9939609	16	1355082	Case-control	Mixed	355 / 364	T	Null association			[40]
	rs7206790	16	53797908	Case-control	Mixed	356 / 364	G	Null association			[40]
	rs8047395	16	53798523	Case-control	Mixed	357 / 364	G	Null association			[40]
HSD11B1	rs932335	1	209905734	Nested case-control	Caucasian	648 / 649	C	1.59	1.25-2.02	Recessive	[43]
	rs4393158	1	209851897	Nested case-control	Caucasian	648 / 649	A	Null association			[43]
	rs17317033	1	209856626	Nested case-control	Caucasian	648 / 649	C	Null association			[43]
	rs2235543	1	209860668	Nested case-control	Caucasian	648 / 649	T	Null association			[43]
	rs11807619	1	209881373	Nested case-control	Caucasian	648 / 649	T	1.41	1.10-1.80		[43]
IRS2	rs2289046	13	110407906	Nested case-control	Caucasian	648 / 649	G	0.74	0.58-0.94		[43]
	rs754204	13	110411568	Nested case-control	Caucasian	648 / 649	T	1.32	1.03-1.70		[43]
	rs12584136	13	110419354	Nested case-control	Caucasian	648 / 649	A	Null association			[43]
LEP	rs7799039	7	127878783	Meta-analysis	Mixed	2,003 / 1,967	A	Null association		Recessive	[41]
LEPR	rs1137101	1	66058513	Meta-analysis	East-Asian	285 / 447	G	0.50	0.36-0.70		[41]

	rs1137101	1	66058513	Meta-analysis	Mixed	4,627 / 5,476	G	Null association			[41]
	rs1137100	1	66036441	Meta-analysis	Mixed	4,627 / 5,476	G	Null association			[41]
MC4R	rs17782313	18	57851097	Case-control	Mixed	100 / 148	T	1.86	0.0035	Additive	[42]
PON1	rs854560	7	94946084	Meta-analysis	Mixed	1,517 / 1,379	M	2.16	1.76-2.66	Recessive	[41]
	rs662	7	94937446	Meta-analysis	Mixed	1,517 / 1,379	R	Null association			[41]

EA, Effect allele

Objectives

To investigate the associations of obesity-related genes on breast cancer risk by

1) Individual SNP analysis between breast cancer risk and SNPs related to BMI
2) gene-based pathway-analysis using SNPs related to BMI on breast cancer risk
3) additional stratification analyses by BMI, menopausal status, and hormone receptor (HR) status for 1) and 2) as above

Thus, complex mechanisms of obesity in the development of breast cancer would be explained.

Materials and methods

1. Study population

In this multicenter case-control study, breast cancer patients were drawn from the Seoul Breast Cancer Study (SEBCS) and healthy controls from a large urban population cohort that is a sub-cohort of the Korea Genome Epidemiology Study (KoGES). Final subjects consisted of 1,786 cases and 1,789 controls (Figure 4). A questionnaire was given by the trained interviewers to collect information on demographic, reproductive and other lifestyle factors from every subject. All participants in the study provided written informed consent when the questionnaire survey was conducted, and the study design was approved by the Committee on Human Research of Seoul National University Hospital (IRB No. H-0503-144-004).

1-1. The Seoul Breast Cancer Study (SEBCS)

The characteristics of the study participants have been previously described in detail [24,45]. Briefly, a total of 4,040 histologically confirmed incident breast cancer patients and 1,894 controls without cancer history were recruited from Seoul National University Hospital (SNUH) and Asan Medical Center (AMC) from 2002 to 2007. 175 cases with previous history of cancer, and no incident breast cancer (N=18), and also diagnosed as benign breast diseases (N=70) were excluded for inclusion criteria. For genome-wide scan (stage I), 2,342 cases were remained and selected from

those with sufficient DNA samples. Among them, average genome-wide IBS between each pair of individuals was calculated to select relationships with first-degree relatives or in relationships with more distant relatives whose clusters were tightly linked to the first-degree relationship. The 17 cases with call rate less than 95%, excessive heterozygosity (N=12), known IBS relatives (N=14) was additionally excluded and 2,165 cases confirmed in SEBCS. Finally 1,786 cases were included in present analyses by matching for 10-years age group since the women from community-based controls were over the age of 40 as described below. For all cases, a retrospective chart review was used to collect the medical and pathology records including information on the ER and PR status measured by an immunohistochemistry (IHC) assay conducted at the hospitals in which the patients were originally diagnosed. The ER and PR results were interpreted as positive when more than 10% of the tumor cells showed positive nuclear staining.

1-2. Korea Genome Epidemiology Study (KoGES)

Controls were selected from the health examinee cohort (HEXA) from large urban population, which is a sub-project of the KoGES to investigate major genetic and environmental factors for common diseases in the Korean population. Of 32,887 women subjects aged 40-69 years recruited between 2006 and 2007, randomly selected 2,392 healthy controls with sufficient DNA were analyzed in genome wide association (GWA) scan. Excluding subjects with known gender inconsistency (N=1), call rate less than 95% (N=252), excessive heterozygosity (N=15), known IBS relatives (N=16),

MDS outliers (N=11), known previous cancer history (N=45), 2,052 controls were identified. Final 1,789 controls were included in this study by matching to the case by 10-year age group (Figure 4).

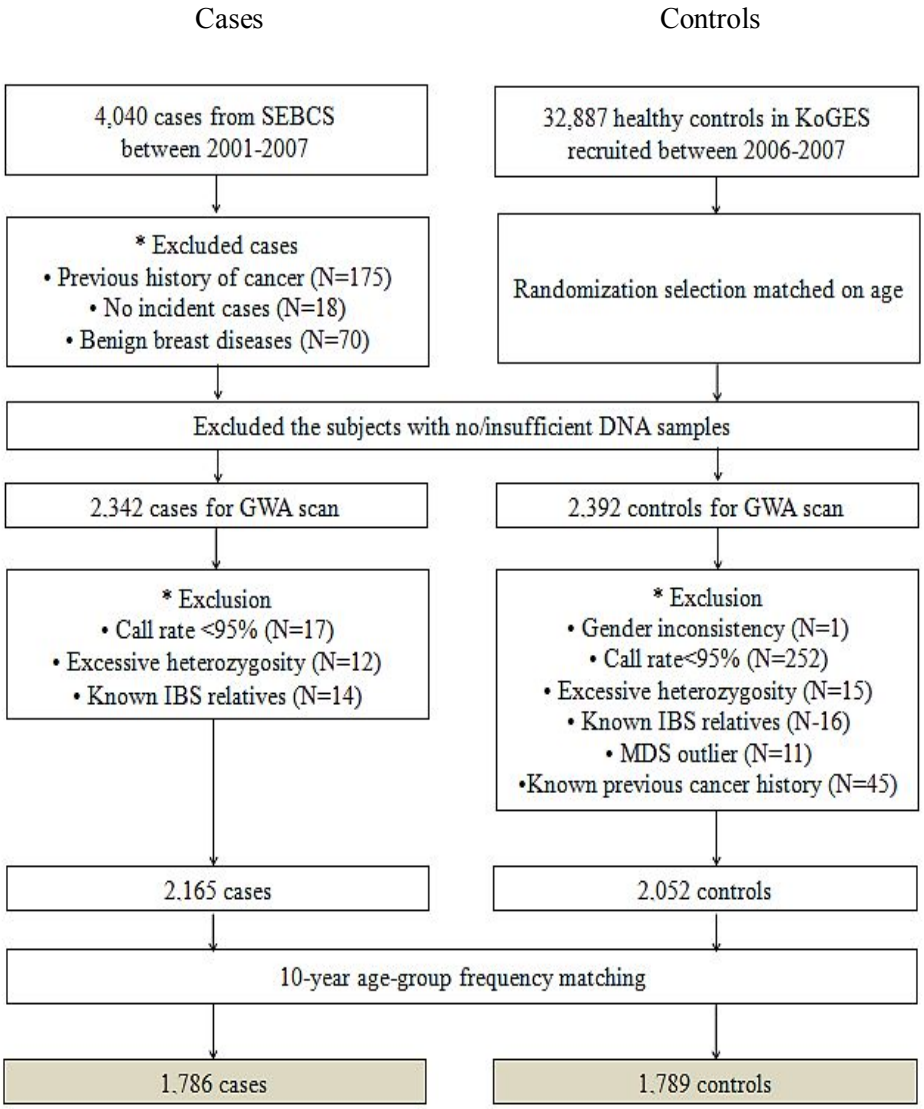


Figure 4. Selection of participants

2. Genotyping and quality control of DNA

2-1. Genome-wide SNP array

Genomic DNA was isolated from peripheral whole blood samples collected at the time of enrollment of subjects using the QUIAZEN DNA Blood Mini Kit based on their instruction (QIAGEN, Valencia, CA, USA). A GWA scan was conducted by the Affymetrix Genome-Wide Human SNP array 6.0 platform (Affymetrix Inc, Santa Clara, CA, USA).

A total of 30 samples were genotyped through SNPstream® UHT (12-plex, SNP-IT assay) for quality control of genotyping and the 99.8% average concordance rate between the samples was observed. Samples of subjects that had a genotype call rate below 95%, an incorrectly imputed gender or a high heterozygosity rate were filtered. For the quality control of marker (SNPs), the following quality control criteria were established and SNPs were excluded according to the criteria : (1) deviation from the *p*-value for Hardy-Weinberg equilibrium of $<10^{-6}$, (2) a genotyping calling of <95%, (3) a minor allele frequency of <1% (4) a poor cluster plot for either cases or controls, (5) differential missingness between cases and controls ($p<10^{-4}$), and (6) multiple positioning and/or mitochondrial SNPs (Table 5) [46].

2-2. Imputation

SNP imputation was carried out to infer the genotype of SNPs that were not observed in the Affymetrix 6.0 by using the hidden Markov model as implemented in MACH 1.0. Imputation was conducted on 555,525 autosomal

SNPs (Table 5) that were genotyped in GWA scan and passed the quality control procedure. Thereafter, the optimized model parameters were used to impute the SNPs on over 2.4 million markers using the phased CHB + JPT data from HapMap Phase II (release 22) as the reference panel. In total, 2,210,823 SNPs which showed an imputation quality score (r^2) of at least 0.30 were confirmed.

Table 5. Procedure for selection of markers

	Case	Control
SNPs at the start of quality control	909,622	
Exclusion of SNPs by		
P-value for HWE <10E-6	-	17,697
Call rate <95%	50,561[a]	94,694[a]
MAF <1E-2	210,952[a]	182,960[a]
Common SNPs	631,882	
Exclusion of SNPs by		
Differential missingness <10E-4	62,534	
Multiple position & mitochondrial SNPs	13,823	
Final SNPs in 2,165 cases and 2,052 controls	555,525	

HWE, Hardy-Weinberg equilibrium; MAF, minor allele frequency; [a]Not mutually exclusive

*Source from Kim et al, 2012 [46]

3. Gene and SNP selection for obesity and genotyping

The selection of obesity-related genetic factors involved two sources (Table 6 and Figure 5). GWAS catalog (http://www.genome.gov/gwastudies) data was used to selects 995 SNPs for obesity related phenotype (i.e., BMI, fat mass, obesity, obesity-related traits, etc.; p<1E-4) and a total of 891 genes from those SNPs were identified (Table 6). The candidate genes for obesity was selected based on previous literature including 540 obesity-related genes through five lines of evidence [47]. These evidences include reports on animal models, mendelian forms of obesity, location near linkage peaks or quantitative trait loci, human genetic association studies and other source of evidence including expression trait. In this report, the following search term in the same way as the previous study was used: 'obesity in mice' or 'obesity in animals'; 'monogenic obesity' or 'mendelian obesity syndrome'; 'genes for obesity' or 'obesity pathway genes' or 'obesity candidate genes'; 'obesity linkage studies', 'obesity expression analysis'; and 'obesity genetic association studies' or 'obesity polymorphism' or 'obesity variant'. Because the paper by Vimaleswaran et al. had searched until June, 2011, additional search between June, 2011 and July, 2013 were conducted using the same search terms and phenotype (lines) in the Pubmed database. In total, 633 candidate genes for obesity were identified of those genes searched (Figure 5). Genomic regions studied were defined to include the candidate genes ±10 kb of flanking sequence around each gene. The available total number of 185,557 SNPs located in 888 genes and 54,350 SNPs in 580 genes from GWAS

Catalog and candidate genes for obesity, respectively, were included in this analysis (Table 7). Additional 449 SNPs in 146 Genes which independently associated with BMI in our participants were added to the present study (as described in Table 7).

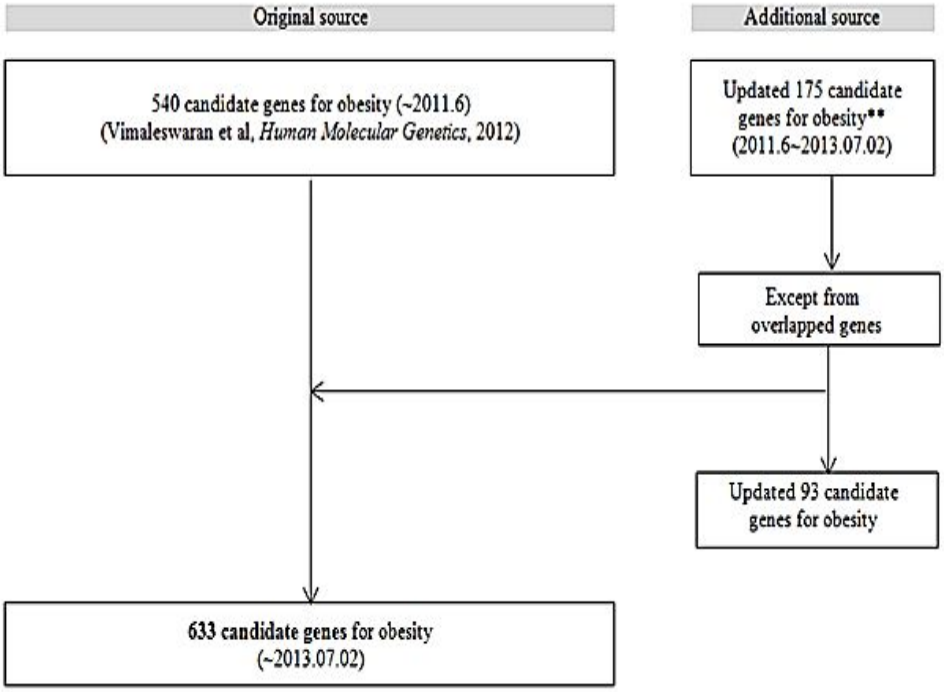
Table 6. Identification of obesity related genetic factors from GWAS catalog

- Obesity-related 994 SNPs (~2013.07.05)

Obesity related phenotype (GWAS catalog)	No. Studies	No. SNPs* (p<1E-4)
Body mass index	14	102
Body mass index (interaction)	1	12
Body mass index and cholesterol	1	2
Body mass index and fat mass	0	0
Obesity	5	25
Obesity (early onset extreme)	1	1
Obesity (extreme)	3	17
Obesity and blood pressure	1	3
Obesity and osteoporosis	1	2
Obesity-related traits (total body weight, waist circumference, waist-hip ratio)	2	856
Total SNPs		995
Total Genes		891 (Intergenic 61)

* Not mutually exclusive

Figure 5. Identification of obesity related genetic factors from previous literature



** Pubmed search term (for updated genes)
Obesity[all field] AND (mice[all field] OR animal [all field] OR monogenic[all field] OR mendelian [all field] OR pathway[all field] OR candidate [all field] OR "linkage" [all field] OR expression[all field] OR association[all field]) AND gene[all field] AND (polymorphism[all field] OR variant[all field])

Table 7. Available number of genes and SNPs in the thesis

	Genes	SNPs
GWAS catalog	888	185,557
Previous literatures	580	54,350
Additional genetic factors in our participants	146	449
Total	1,516	227,197

4. Statistical analyses

4-1. Analyses of demographic, reproductive factors and BMI on breast cancer risk

To define BMI groups for the analysis, the BMI was categorized as <18.5, 18.5-22.9, 23-24.9, and ≥25 using the questionnaire data. Breast cancer patients were grouped into either ER positive or PR positive or both ER- and PR-negative groups.

The differences in continuous and categorical variables between breast cancer cases and controls were assessed by student's *t* test and Pearson's Chi square test, respectively. Binary and polytomous unconditional logistic regression models were used to estimate the odds ratio (ORs) and corresponding 95% confidence interval (95% CI) for associations between BMI and breast cancer risk overall or subtype-specific breast cancer risk stratified by menopausal status.

The models were adjusted simultaneously for age at reference date (defined as the date of diagnosis for cases and date of enrollment for controls), family history of breast cancer (first and second degree relatives), pregnancy history (ever, never), age at first full-term pregnancy (FFTP) and age at menarche which showed the significance difference of distribution between cases and controls. Wald test was used to evaluate the heterogeneity of the associations across breast cancer subtypes defined by the ER/PR status. All statistical tests were based-on two-sided probability with a significance level of 0.05 and done with SAS, version 9.3 (SAS Institute Inc., Cary, North

Carolina).

4-2. Analyses of association between obesity-related genes/SNPs and BMI

Associations of SNPs for BMI were tested in multiple linear regression analyses using an additive genetic model. BMI levels analyzed as a quantitative trait with adjustment for age and cancer status. Statistical software used for association with BMI comprised Plink v1.01 (http://pngu.mgh.harvard.edu/~purcell/plink/anal.shtml) using the additive genetic model and Mach2qtl (http://www.sph.umich.edu/csg/abecasis/MACH/download) with allelic dosage in a linear regression.

4-3. Analyses of association between obesity-related genes/SNPs and breast cancer risk

To investigate SNPs in obesity-related genes and risk of breast cancer, individual SNP analysis and gene-based pathway analyses for breast cancer were carried out (Figure 6).

1) Individual SNP analysis

Among SNPs of candidate genes and SNPs from GWAS in the participants as described in Table 7, BMI-related 2,366 SNPs were identified (p<0.05) after applying for additional quality control; 1) selected SNPs defined by linkage disequilibrium with the pairwise $r^2 \geq 0.4$, 2) minor allele frequency (MAF) of 0.05, 3) identified imputation quality ($R^2 \geq 0.3$) in Asian populations and 4) p-value for Hardy-Weinberg test among controls

(P_{hwe}>10E-5). Unconditional logistic regression was used to estimate the per-allele odds ratio (OR) for breast cancer risk in relation to SNPs with BMI (p<50E-6) adjusting for age and family history of breast cancer in 1st and 2nd relatives using Plink v1.01 for typed data and mach2dat (http://www.sph.umich.edu/csg/abecasis/MACH/download) with allelic dosage for imputation data. The associations on breast cancer were also assessed by strata of BMI and menopausal status.

2) Gene-based pathway analysis

We conducted a gene-based pathway analysis to evaluate the associations between obesity related genes and breast cancer risk. Total of 2,366 SNPs located in 644 genes which relate to BMI were included. The gene-based pathway analysis was based on the adaptive rank truncated product (ARTP) methods and was implemented in the R package 'Adajoint' [48].

a. ARTP method

The ARTP is a one of effective approaches of combining p-value across SNPs within a biological pathway to combine SNP-level p-values [49]. As the ARTP is modified statistic of the rank truncated product statistic (RTP), the concept of RTP is described as below.

Individual L SNPs within the considered pathway can performed to test. The product of the RTP statistic could calculated to test for a null hypothesis like as following, assuming individual L SNPs within the

considered pathway using the standard Cochran-Armitage trend test and ordered statistic of those individual p -value by $p_{(1)} \leq \ldots \leq p_{(L)}$ was performed.

No. of formula

$$① \ W(K) = \prod_{i=1}^{K} p_{(i)}$$

W(K), the product of the K smallest p-values

K, predetermined truncation point, $1 \leq K \leq L$

$p_{(i)}$, resulting ith p-value in a pathway consisting of L SNPs

Because of difficulty to make a sensible choice of K, the ARTP method provides the minimum empirical p-value in final test statics by optimizing the selection of the truncation point among a set of candidates based on following minimum p-values statistics like as following.

$$② \ MinP = min_{1 \leq j \leq J} \hat{S}(K_j)$$

$\hat{S}(K_j)$, estimated p-value for $W(K_j)$, $1 \leq j \leq J$

J, candidate truncation point, $K_1 \leq \ldots \leq K_j$

However, this type of pathway, called as the SNP-based pathway analysis, has the potential drawbacks which are ignoring the structure of the underlying candidate genes that define the biologic pathway. To

overcome this weakness, a gene-based pathway analysis using the ARTP methods was used for this thesis.

b. Gene-based pathway using the ARTP method

This gene-based pathway approach based on the ARTP method is comprised of two steps: 1) to obtain the summary statistics, gene-level *p*-value, for the association between a gene within the pathway and the outcome and 2) to utilize the ARTP method once again to combine these gene-level *p*-values for observed and permuted for the pathway-level association. The ARTP method also used a permutation procedure to estimate its *p*-value by using a common set of referent statistics to avoid computationally expensive permutation algorithms. A summary procedure of the gene-based pathway using the ARTP methods was provided as following.

Assuming that the considered pathway was comprised of L genes, with the l^{th} consisting of n_1SNPs, $1 \leq l \leq L$

$p_{l,i}^{(0)}$, the *p*-value of the association test on the i^{th} SNPs of the l^{th} gene in observed dataset

$p_{l,i}^{(b)}$, the *p*-value of the association test on the i^{th} SNPs of the l^{th} gene in generated dataset based on permutation procedure, $1 \leq b \leq B$

1) 1^{st} applying to the ARTP method for summary gene level *p*-value by

combining SNP-level evidence of association within a gene

$p_l^{*(0)}$, the gene-level *p*-value for the observed data using Min P algorithm given $p_{l,i}^{(b)}$, $1 \leq i \leq n_1$, $0 \leq b \leq B$

$p_l^{*(b)}$, the gene-level *p*-value for the permutated dataset using Min P algorithm given $p_{l,i}^{(b)}$, $1 \leq i \leq n_1$, $0 \leq b \leq B$

2) 2nd applying to the ARTP method (Min *P* algorithm) for adjusted pathway level *p*-value by combining the observed and permuted gene-level *p*-values

In this analysis, the *p*-value for the gene-based pathway analysis was estimated by 1,000 parametric permutation steps adjusted for age. The results were also stratified by menopausal status and BMI levels. *P*-value of 0.05 was used as the significant threshold for gene-based analysis to account for testing 644 obesity related genes.

The most significant SNPs in gene-based pathway approach on BMI and breast cancer risk was also investigated through individual SNP analysis with consideration of BMI, menopausal status and hormone receptor subtype. To examine the potential interaction of genetic effects of the SNP on breast cancer risk, product term of menopausal status, BMI level and the SNP of interest in regression model was included. Wald test for heterogeneity in association of the SNP across ER/PR subtypes was conducted as well.

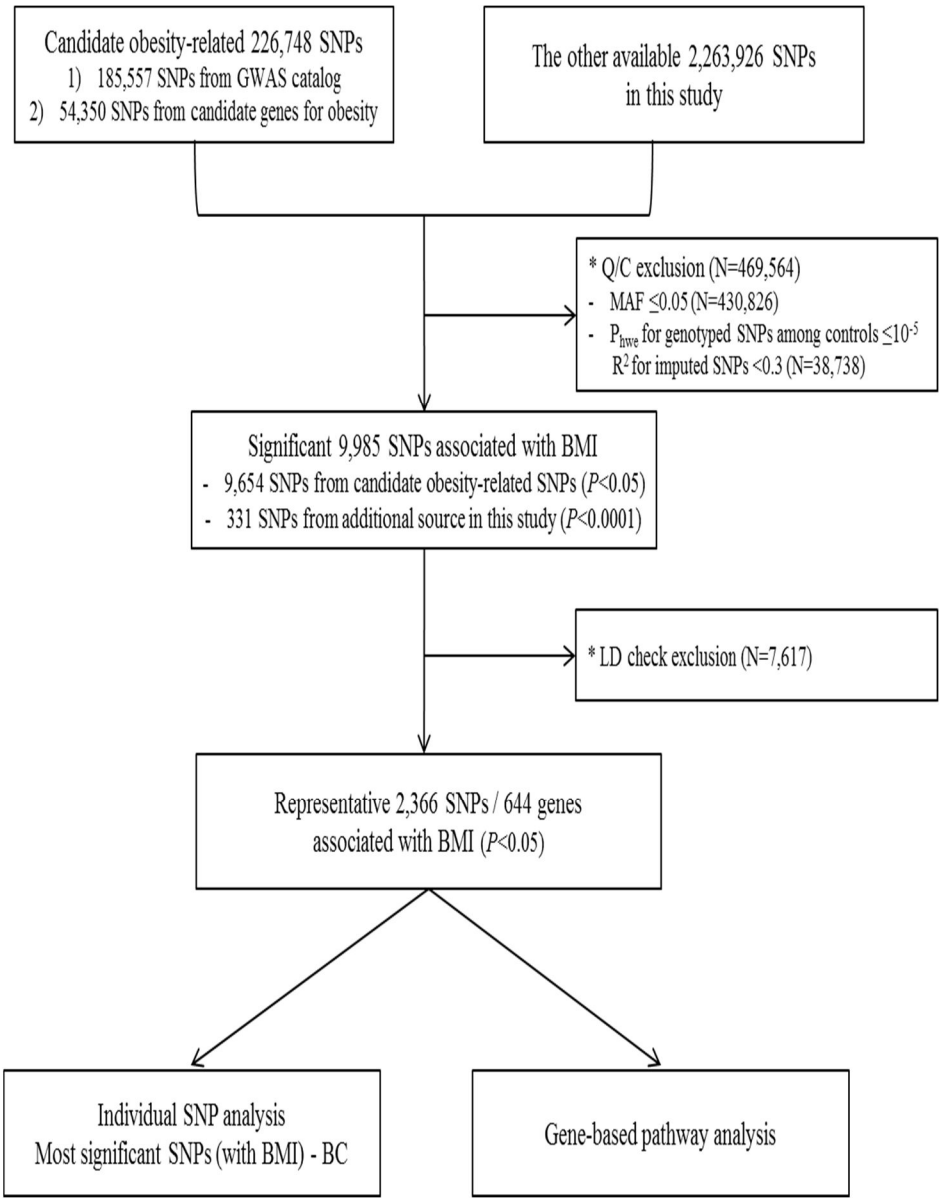


Figure 6. Procedure of the analyses

Results

1. Characteristics of participants

The baseline characteristics of the breast cancer cases and controls were shown in Table 8. Cases were more likely to have a family history of breast cancer, early age at menarche and late age at FFTP and be premenopausal women than controls (all p<1E-4), while no other significant difference was observed on the distribution of age, BMI and history of pregnancy.

Table 9 showed the association between BMI and breast cancer risk by menopausal status and hormone receptor status; no association for BMI with breast cancer risk was observed irrespective of menopausal status, but the protective effect of ER/PR-negative breast cancer was significantly suggested in postmenopausal women with above a BMI of 25 unit (OR=0.64, 95% CI=0.44-0.91; Table 9).

Table 8. The selected characteristics of participants

	Case (N=1,786)		Control (N=1,789)		*P**
	N	(%)	N	(%)	
Age (years), mean (sd)	50.8	(8.0)	50.8	(7.7)	0.8354
40-49	953	(53.4)	954	(53.3)	0.9996
50-59	555	(31.1)	556	(31.1)	
60+	278	(15.6)	279	(15.6)	
Family history					<0.0001
yes	90	(5.0)	31	(1.7)	
no	1696	(95.0)	1758	(98.3)	
Menarche age (years), mean (sd)	14.9	(1.7)	15.2	(1.9)	<0.0001
≤13	350	(19.6)	276	(15.4)	<0.0001
14-15	833	(46.6)	720	(40.3)	
≥16	575	(32.2)	688	(38.5)	
Unknown	28	(1.6)	105	(5.9)	
Ever pregnancy					0.1436
yes	1678	(94.0)	1659	(92.7)	
no	108	(6.1)	130	(7.3)	
Age at FFTP (years)[a], mean (sd)	25.7	(3.6)	25.2	(3.4)	<0.0001
<25	616	(36.7)	707	(42.6)	<0.0001
25-29	868	(51.7)	785	(47.3)	
≥30	186	(11.1)	119	(7.2)	
Unknown	8	(0.5)	48	(2.9)	
No. of children[a], mean (sd)	2.2	(0.9)	2.3	(0.9)	0.4206
1	232	(13.8)	206	(12.4)	0.2318
2	1003	(59.8)	976	(58.8)	
≥3	431	(25.7)	463	(27.9)	
Unknown	12	(0.7)	14	(0.8)	
Menopausal status					<0.0001
premenopausal women	1007	(56.4)	850	(47.5)	
postmenopausal women	770	(43.1)	883	(49.4)	
Unknown	9	(0.5)	56	(3.1)	
Age at menopause[b], mean (sd)	48.6	(5.3)	49.2	(4.6)	0.0188

[a]among parous women
[b]among postmenopausal women
**P*-value by chi-square test

Table 9. Associations of BMI on breast cancer risk by menopausal status and hormone receptor status

BMI (kg/m^2)	Control		Case		OR[a]	(95% CI[a])
	N	(%)	N	(%)		
All	N=1,789		N=1,786			
mean (sd)	23.6	(3.0)	23.4	(2.9)		
median (range)	23.2	(22.0)	23.1	(21.4)		
<18.5	40	(2.2)	38	(2.1)	0.83	0.51-1.32
18.5-22.9	768	(42.9)	812	(45.5)	Ref	
23-24.9	427	(23.9)	476	(26.7)	1.05	0.88-1.24
≥25	494	(27.6)	443	(24.8)	0.85	0.71-1.01
Unknown	60	(3.4)	17	(1.0)		
Premenopausal women	N=850		N=1,007			
<18.5	21	(2.5)	24	(2.4)	0.86	0.47-1.58
18.5-22.9	427	(50.2)	526	(52.2)	Ref	
23-24.9	199	(23.4)	241	(23.9)	0.93	0.73-1.18
≥25	187	(22.0)	211	(21.0)	0.88	0.69-1.13
Unknown	16	(1.9)	5	(0.5)		
Postmenopausal women	N=883		N=770			
<18.5	17	(1.9)	12	(1.6)	0.90	0.40-2.00
18.5-22.9	333	(37.7)	284	(36.9)	Ref	
23-24.9	223	(25.3)	232	(30.1)	1.20	0.93-1.55
≥25	305	(34.5)	232	(30.1)	0.82	0.64-1.05
Unknown	5	(0.6)	10	(1.3)		
ER/PR(+)						
All	N=1,789		N=1,261			
<18.5	40	(2.2)	31	(2.5)	1.03	0.64-1.67
18.5-22.9	768	(42.9)	574	(45.5)	Ref	
23-24.9	427	(23.9)	337	(26.7)	1.07	0.89-1.28
≥25	494	(27.6)	308	(24.4)	0.83	0.69-1.00
Unknown	60	(3.4)	11	(0.9)		
Premenopausal women	N=850		N=749			
<18.5	21	(2.5)	21	(2.8)	1.13	0.61-2.11
18.5-22.9	427	(50.2)	392	(52.3)	Ref	
23-24.9	199	(23.4)	184	(24.6)	0.98	0.76-1.25
≥25	187	(22.0)	148	(19.8)	0.82	0.63-1.06
Unknown	16	(1.9)	4	(0.5)		
Postmenopausal women	N=883		N=505			
<18.5	17	(1.9)	9	(1.8)	0.98	0.42-2.28
18.5-22.9	333	(37.7)	181	(35.8)	Ref	
23-24.9	223	(25.3)	150	(29.7)	1.20	0.91-1.59
≥25	305	(34.5)	160	(31.7)	0.86	0.65-1.12
Unknown	5	(0.6)	5	(1.0)		
ER/PR(-)						
All	N=1,789		N=465			
<18.5	40	(2.2)	7	(1.5)	0.64	0.28-1.44
18.5-22.9	768	(42.9)	213	(45.8)	Ref	
23-24.9	427	(23.9)	124	(26.7)	1.01	0.79-1.31

≥25	494	(27.6)	115	(24.7)	0.78	0.60-1.01
Unknown	60	(3.4)	6	(1.3)		
Premenopausal women		N=850		N=222		
<18.5	21	(2.5)	3	(1.4)	0.55	0.46-1.86
18.5-22.9	427	(50.2)	115	(51.8)	Ref	
23-24.9	199	(23.4)	51	(23.0)	0.91	0.63-1.32
≥25	187	(22.0)	52	(23.4)	0.97	0.67-1.41
Unknown	16	(1.9)	1	(0.5)		
Postmenopausal women		N=883		N=241		
<18.5	17	(1.9)	3	(1.2)	0.61	0.17-2.15
18.5-22.9	333	(37.7)	97	(40.3)	Ref	
23-24.9	223	(25.3)	73	(30.3)	1.09	0.77-1.55
≥25	305	(34.5)	63	(26.1)	0.63	0.44-0.91
Unknown	5	(0.6)	5	(2.1)		

[a]ORs and 95% CI analyzed by logistic regression adjusting for age, family history of breast cancer in 1st and 2nd relatives, age at menarche, age at FFTP and ever pregnancy
*All *P* for heterogeneity of association with BMI on breast cancer by subtype was not shown because none of associations were significant

2. Findings of obesity-related SNP and genes with BMI

Among available candidate obesity-related 227,197 SNPs from previous literature and additional significant SNPs in this multicenter case-control data, we found 9,985 SNPs to be associated with BMI at $p<0.05$ distributed in 712 gene loci (Table 10); 408 genes (45.9%) from GWAS catalog and 221 genes (38.1%) from previous literatures of obesity were significantly associated with BMI, respectively. The manhattan plot of the associations for BMI in these SNPs was shown in Figure 7 and their results of association on breast cancer risk among 9,985 SNPs related to BMI ($p<0.05$) were suggested in Figure 8. The genomic context and the associations with BMI for the most significant SNPs in each source (literature review including GWAS catalog and GWAS analysis in the participants) were presented in Table 11. For the most associated SNPs rs17804012 in *RBFOX1* from candidate source, BMI levels were higher by 39% (SE=0.1, p=2.5E-6) per minor allele. For the other SNP rs2014791 in *LINC00317* identified from this data, per minor-allele were higher by 34% (SE=0.1, p=2.8E-6; Table 11).

Table 10. Number of SNPs associated with BMI

	GWAS catalog		Previous literatures		Additional SNPs		Total	
Available in our case-control data								
Gene	888		580		146		1,561	
SNP	185,557		54,350		449		227,197	
Quality control check								
SNP	151,353	(81.6)	43,023	(79.2)	331	(73.7)	184,061	(81.0)
Significant SNPs/genes associated with BMI among total subjects, $p<0.05$								
Gene	408	(45.9)	221	(38.1)	112	(76.7)	712	(45.6)
SNP	7,758	(4.2)	2,513	(4.6)	331	(73.7)	9,985	(4.4)
Significant SNPs associated with BMI among total subjects according to *p*-value								
$p<10E-3$	241	(0.1)	46	(0.1)	331	(73.7)	591	(0.3)
$p<10E-4$	17	(0.0)	1	(0.0)	331	(73.7)	350	(0.2)
$p<10E-5$	5	(0.0)	0	(0.0)	228	(50.8)	233	(0.1)
Number of SNPs associated with breast cancer risk in 9,985 SNPs by *p*-value								
$p<50E-2$	341	(4.4)	207	(8.2)	15	(4.5)	532	(5.3)
$P<10E-3$	15	(0.2)	13	(0.5)	0	(0.0)	28	(0.3)
$p<10E-4$	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)

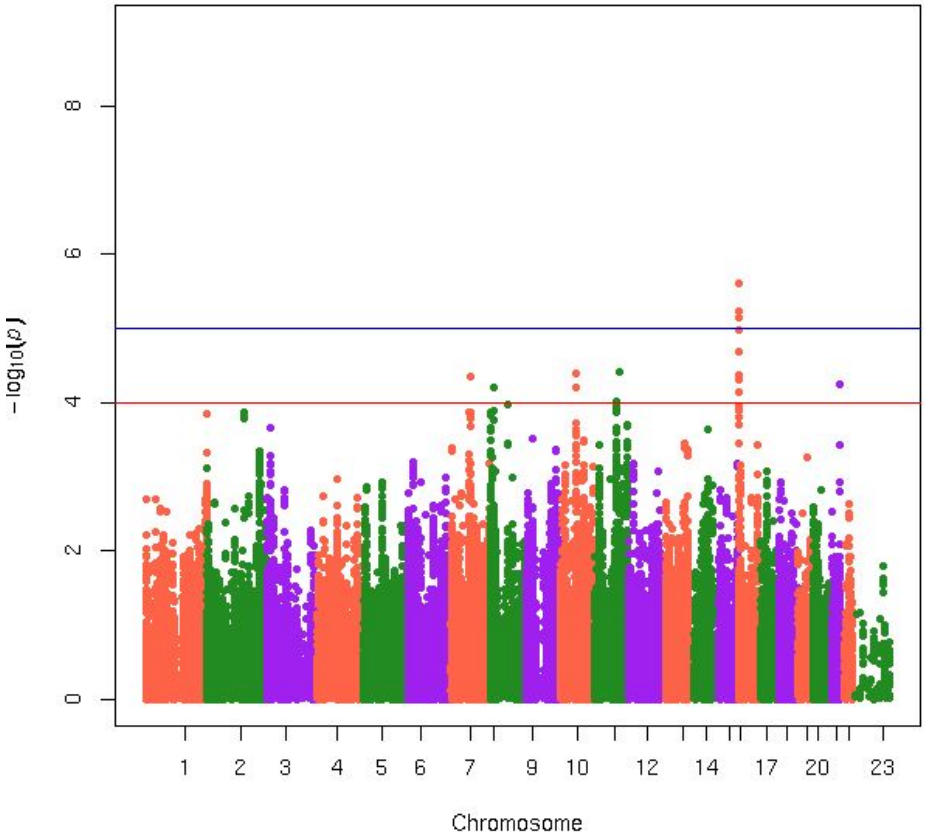


Figure 7. Manhattan plot of association between 184,061 SNPs and BMI

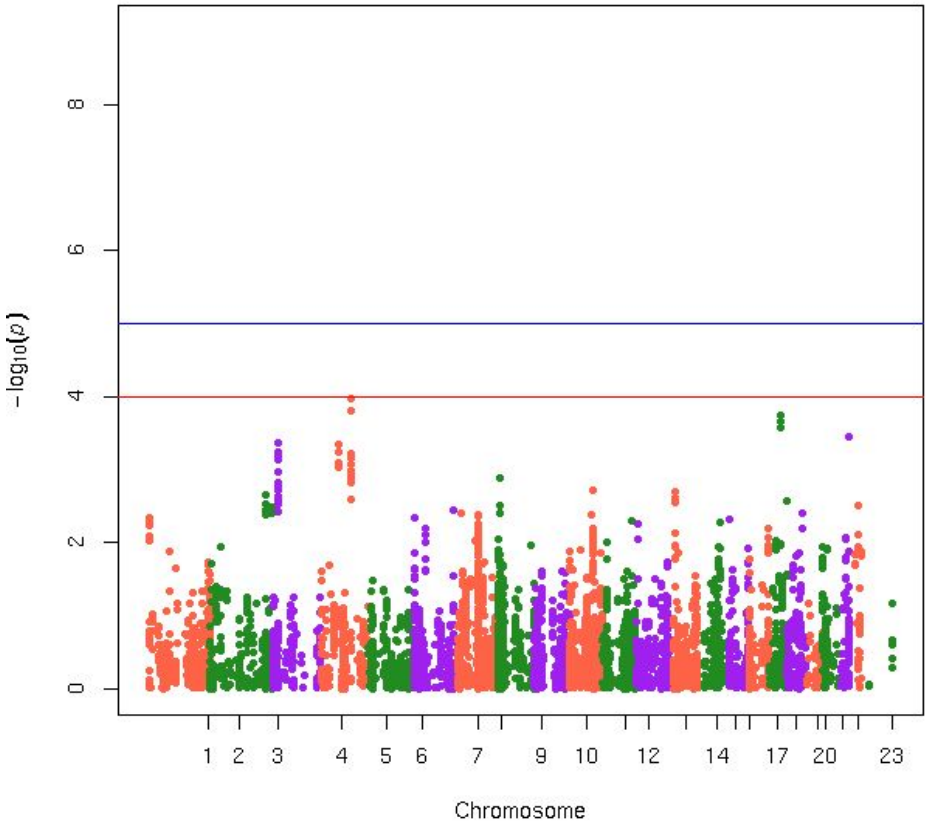


Figure 8. Manhattan plot of association between 9,985 SNPs related to BMI (p<0.05) and breast cancer risk

Table 11. Associations of SNPs on BMI (*P*<5E-6)

	Single-Nucleotide Polymorphism	
	rs17804012	**rs2014791**
Source	GWAS catalog	Noble
Genomic context		
Chromosome	16	21
Position	614867	22067821
Locus	*RBFOX1*	*LINC00317*
Gene name	RNA binding protein, fox-1 homolog 1	Long intergenic non-protein coding RNA 317
Alleles		
Reference	A	C
Minor	C	T
Minor allele frequency	0.2813	0.3824
R^2/P_{hwe}	0.8773	0.0887
Association with BMI[a] among total subjects		
Beta (se)	0.39 (0.1)	0.34 (0.1)
Association *P*	2.46.E-6	2.77.E-6
Association with BMI[a] among premenopausal subjects		
Beta (se)	0.34 (0.1)	0.11 (0.1)
Association *P*	<0.0001	0.2252
Association with BMI[a] among postmenopausal subjects		
Beta (se)	0.22 (0.1)	0.58 (0.1)
Association *P*	0.0673	<1E-4

[a]Results by linear regression adjusting for age and cancer status (case/control)

3. Findings of individual genetic variants with BMI on breast cancer risk

We confirmed null association with overall breast cancer risk for 2 SNPs which were most strongly associated with BMI: distributions of allele frequencies rs17804012 in *RBFOX1* and rs2014791 in *LINCOO317* were not significantly different between breast cancer cases and controls. These results were not different after stratified analyses by menopausal status and BMI levels (Table 12).

Table 12. Association between SNPs related to BMI and breast cancer risk

SNP	Gene	Menopausal status	BMI results[a]			BMI group	Breast cancer risk[b]		
			Beta	SE	Association *P*		OR	95% CI	Association *P*
rs17804012	*RBFOX1*	All	0.39	0.1	2.46E-06	All	1.03	0.92-1.15	0.6281
						<23.2	0.97	0.83-1.14	0.6935
						≥23.2	1.08	0.92-1.26	0.3534
		Premenopausal	0.34	0.1	<1E-4	All	1	0.85-1.17	0.9665
						<23.2	0.94	0.76-1.16	0.5808
						≥23.2	1.06	0.83-1.35	0.6342
		Postmenopausal	0.22	0.1	0.0673	All	1.08	0.92-1.27	0.3238
						<23.2	1.07	0.83-1.37	0.6229
						≥23.2	1.07	0.86-1.32	0.5438
rs2014791	*LINC00317*	All	0.34	0.1	2.77E-06	All	1	0.91-1.10	0.9789
						<23.2	0.98	0.86-1.13	0.7868
						≥23.2	1.01	0.87-1.16	0.9468
		Premenopausal	0.11	0.1	0.2252	All	0.96	0.84-1.10	0.5434
						<23.2	0.96	0.80-1.14	0.6408
						≥23.2	0.93	0.75-1.15	0.5013
		Postmenopausal	0.58	0.1	<1E-4	All	1.04	0.90-1.20	0.6101
						<23.2	1	0.80-1.25	0.9831
						≥23.2	1.07	0.88-1.29	0.5191

[a]Results by linear regression adjusting for age and cancer status; [b]Results by logistic regression adjusting for age and family history of breast cancer in 1st and 2nd relatives

4. Findings of gene-based pathway analysis related to BMI on breast cancer risk

Pathway-based analysis of all BMI-related 644 genes involved in obesity signaling was significantly associated with risk of breast cancer (pathway p=3E-3, data not shown). The gene-based analyses identified 40 genes associated with overall risk of breast cancer by 1,000 permutations (all gene p<0.05) and the selected pathway based on those 40 genes showed higher significance with breast cancer risk (p=2E-3, Table 13). Especially, the effect of *THRB* (thyroid hormone receptor, beta) gene on breast cancer risk has highly significant value (p=9E-4).

The *THRB* effects on breast cancer risk were different in subgroup analyses by menopausal status and BMI levels (Table 14). The association of pathway between *THRB* genes and breast cancer risk was only significant among women with above BMI median value 23.2 (p=3.9E-3) and premenopausal subgroups (p=4E-3). When the pathway analysis is applied to participants which considered both BMI and menopausal status, the significant association between *THRB* and breast cancer risk was confined to premenopausal women with above BMI 23.2 (p=5E-3). The association with BMI has also same directional of the results of breast cancer risk with BMI level lower per minor allele, when the effect of the representative SNP rs6550865 in *THRB* genes on BMI and breast cancer was identified by individual analysis (Table 15). In stratified analysis of hormone receptor-defined subtype, the significant association was predominately observed in

ER/PR (+) breast cancer, although there were no significant differences by subtypes and no significant interactions effect between the rs6550865, BMI and menopausal status (interaction p=0.2421, data not shown) (Table 16).

Table 13. Associations of genes and pathway related to BMI on breast cancer risk with 2,366 SNPs

Pathway *p*	Gene Abb	Gene	Cytogenic locus	No. SNPs	most associated SNP	Gene *P**
2.0E-3	***THRB***	**Thyroid hormone receptor, beta**	**20p13-p12.2**	**5**	**rs6550865**	**9.0.E-4**
	ABCC8	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	11p15.1	4	rs2237967	1.0.E-3
	CD36	CD36 molecule (thrombospondin receptor)	7q11.2	5	rs10499858	1.0.E-3
	LOC644649	Apolipoprotein O pseudogene 5	16q21	1	rs2289825	1.0.E-3
	PMAIP1	Phorbol-12-myristate-13-acetate-induced protein 1	18q21.32	1	rs9960133	1.0.E-3
	MIAT	Myocardial infarction associated transcript (non-protein coding)	4p13	1	rs6005130	1.3.E-3
	TOM1L1	Target of myb1 (chicken)-like 1	3p24.2	2	rs12936080	2.5.E-3
	FAS	Fas cell surface death receptor	10q24.1	1	rs4934433	5.0.E-3
	OSBPL6	Oxysterol binding protein-like 6	2q32.1	1	rs334616	5.0.E-3
	MRPS18A	Mitochondrial ribosomal protein S18A	6p21.3	2	rs833053	9.0.E-3
	ADRA1B	Adrenoceptor alpha 1B	5q33.3	1	rs2222308	9.4.E-3
	PVALB	Parvalbumin	11q24.2	1	rs2022068	1.2.E-2
	PRDM5	PR domain containing 5	4q25-q26	2	rs2597540	1.3.E-2
	MTHFD1L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	6q25.1	2	rs752543	1.3.E-2
	TTC8	Tetratricopeptide repeat domain 8	17q23.2	2	rs1287662	1.3.E-2
	C6orf195	Chromosome 6 open reading frame 195	6p25.2	1	rs1150656	1.8.E-2
	RAB6A	RAB6A, member RAS oncogene family	11q13.3	1	rs17132305	1.9.E-2
	NDN	Necdin, melanoma antigen (MAGE) family member	15q11.2-q12	2	rs699329	2.0.E-2
	ZNF259	Zinc finger protein 259	11q23.3	2	rs1729410	2.1.E-2
	GAA	Glucosidase, alpha; acid	17q25.2-q25.3	1	rs12600845	2.3.E-2
	AGMO	Alkylglycerol monooxygenase	7p21.2	2	rs12667785	2.3.E-2
	Intergenic	Rs1379996	20p11.22	1	rs1379996	2.3.E-2
	GRXCR1	Glutaredoxin, cysteine rich 1	4q21	1	rs2345759	2.4.E-2
	IGF1	Insulin-like growth factor 1 (somatomedin C)	12q23.2	1	rs2195240	2.4.E-2
	PARP4	Poly (ADP-ribose) polymerase family, member 4	13q11	2	rs17080735	2.5.E-2
	FHOD3	Formin homology 2 domain containing 3	18q12	4	rs3744903	2.9.E-2

TTLL7	Tubulin tyrosine ligase-like family, member 7	1p31.1	3	rs12042443	3.4.E-2
NBEA	Neurobeachin	13q13	7	rs9544394	3.6.E-2
PPYR1	Neuropeptide Y receptor Y4	5q11.2-q12	1	rs1343514	3.6.E-2
RIN2	Ras and Rab interactor 2	22q13.1	1	rs4813379	3.7.E-2
GDF3	Growth differentiation factor 3	12p13.1	1	rs7297638	3.7.E-2
GPR74	Neuropeptide FF receptor 2	22q12.1	4	rs11733404	3.7.E-2
PRDM16	PR domain containing 16	1p36.23-p33	11	rs2651898	3.8.E-2
FSIP1	Fibrous sheath interacting protein 1	15q14	1	rs7179013	3.8.E-2
ASIC2	Acid-sensing (proton-gated) ion channel 2	17q12	11	rs7226110	4.2.E-2
RBP5	Retinol binding protein 5, cellular	12p13.31	2	rs7297327	4.2.E-2
TRB3	Tribbles pseudokinase 3	14q31.3	1	rs6037460	4.5.E-2
ARAP2	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 2	4p14	1	rs11935245	4.6.E-2
COL1A1	Collagen, type I, alpha 1	17q21.33	1	rs2269336	4.7.E-2
PKNOX2	PBX/knotted 1 homeobox 2	10q11.2	1	rs840026	5.0.E-2

* adjusted for age
Results by 1,000 permutations

Table 14. Association of *THRB* on breast cancer risk stratified by subgroups

Gene Abb	Gene	Cytogenic locus	Subgroup		Gene *P**
THRB	Thyroid hormone receptor, beta	20p13-p12.2	Overall	Total women	0.0009
			BMI level	Women with under BMI 23.2	0.2607
				Women with above BMI 23.2	0.0039
			Menopausal status	Premenopausal women	0.0040
				Postmenopausal women	0.5025
			BMI level & menopausal status	Premenopausal women with under BMI 23.2	Excluded
				Premenopausal women with above BMI 23.2	0.0050
				Postmenopausal women with under BMI 23.2	0.8241
				Postmenopausal women with above BMI 23.2	0.1358

* adjusted for age

Results by 1,000 permutations

Table 15. Association between rs6550865 in *THRB* and breast cancer risk

SNP	Gene	EA	RA	EAF	P_{hwe}/R^2	Menopausal status	BMI results[a]			BMI group	Breast cancer risk[b]		
							Beta	SE	*P*		OR	95% CI	*P*
rs6550865	*THRB*	T	C	0.4964	0.0882	All	-0.19	0.07	0.0072	All	0.85	0.77-0.93	0.0004
										<23.2	0.90	0.79-1.02	0.1066
										≥23.2	0.78	0.68-0.90	0.0005
						Pre-menopausal	-0.30	0.09	0.0013	All	0.80	0.71-0.92	0.0014
										<23.2	0.88	0.74-1.04	0.1402
										≥23.2	0.72	0.59-0.89	0.0021
						Post-menopausal	-0.09	0.11	0.4189	All	0.89	0.78-1.03	0.1154
										<23.2	0.91	0.73-1.12	0.3707
										≥23.2	0.88	0.72-1.06	0.1699

EA, effect allele; RA, reference allele; EAF, effect allele frequency

[a]Results by linear regression adjusting for age and cancer status (case/control); [b]Results by logistic regression adjusting for age and family history of breast cancer in 1st and 2nd relatives

Table 16. Association between rs6550865 in *THRB* and breast cancer risk by hormone receptor status

rs6550865	Controls	ER/PR(+)			ER/PR(-)			$P_{heterogenetiy}$
genotype	N	N	OR*	95% CI*	N	OR*	95% CI*	
Total women								
CC	398	361	Ref		119	Ref		
TC	929	611	0.73	0.62-0.88	236	0.85	0.66-1.09	0.2658
TT	460	287	0.69	0.56-0.85	110	0.79	0.59-1.06	0.3684
Unknown	2	2			0			
Women with above BMI 23.2								
CC	190	191	Ref		58	Ref		
TC	452	285	0.63	0.49-0.81	118	0.86	0.60-1.23	0.1027
TT	217	132	0.59	0.44-0.80	45	0.65	0.42-1.01	0.6688
Unknown	1	1			0			
Premenopausal women								
CC	193	223	Ref		70	Ref		
TC	437	359	0.73	0.57-0.92	105	0.67	0.47-0.94	0.6392
TT	219	166	0.67	0.51-0.89	47	0.60	0.39-0.91	0.6066
Unknown	1	1			0			
Premenopausal women with above BMI 23.2								
CC	81	106	Ref		37	Ref		
TC	188	144	0.63	0.44-0.91	44	0.55	0.33-0.92	0.5907
TT	90	64	0.57	0.37-0.88	13	0.33	0.17-0.68	0.1376
Unknown	0	0			0			

* adjusting for age and family history of breast cancer in 1st and 2nd relatives

Discussion

In the present genome-wide association study among Korean women, we identified the association between previous known obesity-related candidate genes and BMI level. About 42% of candidate genes for previous research (45.9% from GWAS catalog; 38.1% from candidate genes for obesity) were shown to be associated with BMI at $p<0.05$. The variants included in most 2 significant SNPs in our study are associated with approximately 30% higher BMI level and the significant associations between polymorphisms of these 2 SNPs and breast cancer risk in individual SNP analysis were not confirmed. But, we did also observe that the effect estimate based on BMI related gene-based pathway analysis was highly associated with breast cancer risk (pathway $p=0.003$). Specifically, the effect of *THRB* gene showed stronger among premenopausal women with BMI above cutoff of 23.2 ($p<0.01$).

The null association between BMI and breast cancer risk was observed using this case-control data and the results were same in stratified analyses by menopausal status. Many previous studies including meta-analysis have reported the positive association of BMI with breast cancer risk, especially among postmenopausal women while conflicting effect persist between premenopausal BMI and risk for breast cancer [21,50-52]. However, no associations of postmenopausal BMI on breast cancer risk were also observed with large sample size using data from international collaborative

study as well as case-control studies consisted of Korean women [53-56]. The appearance of controversy may arise in part because studies of body mass index and breast cancer risk have used a wide variety of BMI categories and varying reference categories. Also, associations between BMI and breast cancer risk in postmenopausal women differed depending on subtype. According to previous studies, postmenopausal BMI was negatively associated with ER/PR-negative breast cancer, which is consistent with our significant results among subtype analysis (OR=0.63, 95% CI=0.44-0.91, BMI 25+ vs 18.5-23), while positive association with ER/PR-positive breast cancer was suggested [35,57-59]. Especially, the previous results by Berstad et al. showed the protective effect of BMI was greater on ER/PR tumors among white women, but not African-American women [35]. These differences of frequency of specific subtypes might be observed across ethnic group, such as more frequent ER/PR-negative breast cancer in African-American than other ethnic groups, could be contributed to different effect of BMI on breast cancer by race [60].

Among many BMI-related SNPs in present study, rs17804012 in *RBFOX1* locus selected from GWAS catalog data was most strongly associated with BMI (p<2.46e-06). Although the effect of rs17804012 on obesity-related phenotype was not understood yet, the associations between variants in *A2BP1* with both percent body fat and BMI in GWAS by Ma and colleagues [61]. The additional evidence for association of these variants with longitudinally measured BMI in a large population-based sample of full-heritage Pima Indians, along with the obese phenol-types observed with

A2BP1 knockout mice and high expression levels of *A2BP1* in the hypothalamus, leads us to speculate that *A2BP1* has a role in body weight regulation [62,63]. Our findings include a novel SNP rs2014791 in *LINC00317* (also known as *C21orf117*, Long intergenic non-protein coding RNA 317), which was not included in previous literature, of a plausible biological function in obesity pathogenesis, especially associated with BMI (p=2.77e-06). Further studies will be needed to confirm the variants in intergenic non-protein coding region including *LINC00317* as new susceptibility loci for obesity.

Previous animal studies using knockout and knockin mouse models have shown that thyroid hormone receptor-beta (*THRβ, TRβ*) is involved in mediating the Tri-iodothy-ronine (T3) effects on liver metabolism such as reducing plasma cholesterol and triglycerides [64,65] and the protective effect is primarily significant among more obese women (BMI 23.2+) in our gene-based pathway analysis. A recent cell line study by Park and colleagues also provided additional in vivo evidence to support the idea that *TRβ* could act as a tumor suppressor in breast cancer development and progression, by regulating the pathways by which *TRβ* blocked the estradiol 2-stimulated cell proliferation based on attenuating the JAK-STAT signaling [66]. These findings suggest possibility of association with *THRB* genes on breast cancer, which showed high degree of significance in our gene-based pathway analyses by 1,000 permutations (all p<0.0005).

Furthermore, we found evidence that genetic variation in the *THRB* genes significantly contributed to breast cancer risk among premenopausal

women. The results by Saraiva supported that serum concentration of thyroid-stimulating hormone (TSH) was significantly higher in the group of premenopausal breast cancer patients than in controls, while lower TSH levels were observed among the group of postmenopausal breast cancer patients compared to those of controls; higher TSH level was more linked to sex hormone-sensitive HR-positive breast cancer through stimulating *THRB* genes and regulating hormone stimulated cell proliferation [67].

Several potential limitations of this study must be considered. Cases and controls were selected from different study and recruitment period, in words, cases and were recruited from hospital-based SEBCS study between 2001 and 2007 and controls were recruited from community-based KoGES study between 2006 and 2007, respectively, which might arise in our results through selection bias different by different characteristics between two groups.

The bias might influence toward protective effect of the BMI-related genetic factors on breast cancer risk among overweight women as well as inverse relationship between BMI and risk for breast cancer although no statistically significant associations with BMI on breast cancer risk were observed; hospital-based recruited cases might be less healthy than controls in community based study from general population and overweight or obese women could be less included especially in the way that BMI was calculated by post-diagnostic weight among cases.

Since hormone receptor data for the patients was collected primarily from medical records without additional validation between hospitals, it is

possible that misclassification of hormone receptor status was introduced. Nevertheless, the medical records of hormone receptor are expected to be relatively exact information since data on hormone receptor is important factor for clinical decision of treatment. Furthermore, there were no significant differences when we compared the distributions of hormone receptor status between the cases diagnosed by two hospitals (SNUH and AMC, data not shown).

Although the information of exposure (including weight and height) was collected primarily from self-reported questionnaire which might have a chance of being misclassified, but this bias is unlikely to have distorted our present analysis. Another limitation is using of low significance level of 0.05 threshold for identification of markers associated with BMI and breast cancer, but not consideration of multiple comparison. However, we identified the significant association between BMI-related genes and breast cancer risk through 1,000 permutation tests based on min *P* algorithm as a means of controlling for multiple comparisons.

Our analysis focused not only individual SNP analysis which often identifies only just few of the most significant SNPs based on high level of significance but also gene-based pathway analyses that compensate those drawbacks of single-SNP analysis. Our findings from this studies was needed to confirm in further studies as the absence of replication studies, despite the present study was comprehensively conducted compared to previous associations studies in terms of embracive consideration of obesity-related genes and interactions between genetic and environmental factors.

To our knowledge, it is the first case-control association study to evaluate the association between genetic variants in obesity-related genes and BMI/breast cancer in Korean population with large sample size. Additionally, the current study is considered a knowledge based analysis approach as inclusion of all SNP/gene that associated with obesity from collection of literatures. Most previous reports did evaluate the association in limited obesity related genes and risk for breast cancer. In addition, our examination of a large number of obesity-related SNPs/genes associated with breast cancer using gene-based analysis. Gene and pathway-based analysis as a new paradigm for GWAS has been proposed to compensate the drawbacks of single-SNP analysis which often identifies only a few of the most significant SNPs that account for a small proportion of the genetic variants in complex diseases. Thus, the present study is considered to represent one of the most comprehensively conducted assessments of genetic variation in obesity-related genes and breast cancer risk.

In summary, we clarified that the associations of previous known obesity-related genes including *RBFOX1* with BMI and novel intergenic region which may linked to obesity was confirmed. Also, the BMI-related genes were significantly associated with breast cancer risk in pathway analysis and the association was different by menopausal status and BMI levels, although the association was need for identification through additional validation study. Further longitudinal studies are needed to better understand how affect obesity related genes are associated with breast cancer with environmental factors.

Reference

1. Lee JH, Yim SH, Won YJ, Jung KW, Son BH, et al. (2007) Population-based breast cancer statistics in Korea during 1993-2002: incidence, mortality, and survival. J Korean Med Sci 22 Suppl: S11-16.
2. Bray F, McCarron P, Parkin DM (2004) The changing global patterns of female breast cancer incidence and mortality. Breast Cancer Res 6: 229-239.
3. Jung KW, Won YJ, Kong HJ, Oh CM, Seo HG, et al. (2013) Cancer statistics in Korea: incidence, mortality, survival and prevalence in 2010. Cancer Res Treat 45: 1-14.
4. Korean Central Cancer Registry (2010) Annual report of cancer statistics in Korea. Korean National Cancer Registiry.
5. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, et al. (2011) Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. Lancet 378: 1461-1484.
6. National Cancer Center (2011) Report on the statistics for National Cancer Incidence 2009.
7. Flegal KM, Carroll MD, Kit BK, Ogden CL (2012) Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 307: 491-497.
8. Kim DM, Ahn CW, Nam SY (2005) Prevalence of obesity in Korea. Obes Rev 6: 117-121.

9. Park HS, Park CY, Oh SW, Yoo HJ (2008) Prevalence of obesity and metabolic syndrome in Korean adults. Obes Rev 9: 104-107.

10. Khang YH, Yun SC (2010) Trends in general and abdominal obesity among Korean adults: findings from 1998, 2001, 2005, and 2007 Korea National Health and Nutrition Examination Surveys. J Korean Med Sci 25: 1582-1588.

11. Survey KNHaNE (2013) Prevalence of obesity in Korean adults.

12. Kopelman P (2007) Health risks associated with overweight and obesity. Obes Rev 8 Suppl 1: 13-17.

13. Khandekar MJ, Cohen P, Spiegelman BM (2011) Molecular mechanisms of cancer development in obesity. Nat Rev Cancer 11: 886-895.

14. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, et al. (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41: 25-34.

15. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, et al. (2009) Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet 41: 18-24.

16. Wang K, Li WD, Zhang CK, Wang Z, Glessner JT, et al. (2011) A genome-wide association study on obesity and obesity-related traits. PLoS One 6: e18939.

17. Dumitrescu RG, Cotarla I (2005) Understanding breast cancer risk -- where do we stand in 2005? J Cell Mol Med 9: 208-221.

18. Kelsey JL, Gammon MD, John EM (1993) Reproductive factors and breast cancer. Epidemiol Rev 15: 36-47.

19. McPherson K, Steel CM, Dixon JM (2000) ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. BMJ 321: 624-628.

20. van den Brandt PA, Spiegelman D, Yaun SS, Adami HO, Beeson L, et al. (2000) Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol 152: 514-527.

21. Morimoto LM, White E, Chen Z, Chlebowski RT, Hays J, et al. (2002) Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative (United States). Cancer Causes Control 13: 741-751.

22. Ahn SH, Yoo KY (2006) Chronological changes of clinical characteristics in 31,115 new breast cancer patients among Koreans during 1996-2004. Breast Cancer Res Treat 99: 209-214.

23. Yoo KY, Kim Y, Park SK, Kang D (2006) Lifestyle, genetic susceptibility and future trends of breast cancer in Korea. Asian Pac J Cancer Prev 7: 679-682.

24. Chung S, Park SK, Sung H, Song N, Han W, et al. (2013) Association between chronological change of reproductive factors and breast cancer risk defined by hormone receptor status: results from the Seoul Breast Cancer Study. Breast Cancer Res Treat 140: 557-565.

25. Huang Z, Hankinson SE, Colditz GA, Stampfer MJ, Hunter DJ, et al. (1997) Dual effects of weight and weight gain on breast cancer risk.

JAMA 278: 1407-1411.

26. Lahmann PH, Hoffmann K, Allen N, van Gils CH, Khaw KT, et al. (2004) Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). Int J Cancer 111: 762-771.

27. McTiernan A, Rajan KB, Tworoger SS, Irwin M, Bernstein L, et al. (2003) Adiposity and sex hormones in postmenopausal breast cancer survivors. J Clin Oncol 21: 1961-1966.

28. Cecchini RS, Costantino JP, Cauley JA, Cronin WM, Wickerham DL, et al. (2012) Body mass index and the risk for developing invasive breast cancer among high-risk women in NSABP P-1 and STAR breast cancer prevention trials. Cancer Prev Res (Phila) 5: 583-592.

29. White KK, Park SY, Kolonel LN, Henderson BE, Wilkens LR (2012) Body size and breast cancer risk: the Multiethnic Cohort. Int J Cancer 131: E705-716.

30. Harris HR, Willett WC, Terry KL, Michels KB (2011) Body fat distribution and risk of premenopausal breast cancer in the Nurses' Health Study II. J Natl Cancer Inst 103: 273-278.

31. Reeves GK, Pirie K, Beral V, Green J, Spencer E, et al. (2007) Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. BMJ 335: 1134.

32. Ahn J, Schatzkin A, Lacey JV, Jr., Albanes D, Ballard-Barbash R, et al. (2007) Adiposity, adult weight change, and postmenopausal breast

cancer risk. Arch Intern Med 167: 2091-2102.

33. Palmer JR, Adams-Campbell LL, Boggs DA, Wise LA, Rosenberg L (2007) A prospective study of body size and breast cancer in black women. Cancer Epidemiol Biomarkers Prev 16: 1795-1802.

34. John EM, Sangaramoorthy M, Phipps AI, Koo J, Horn-Ross PL (2011) Adult body size, hormone receptor status, and premenopausal breast cancer risk in a multiethnic population: the San Francisco Bay Area breast cancer study. Am J Epidemiol 173: 201-216.

35. Berstad P, Coates RJ, Bernstein L, Folger SG, Malone KE, et al. (2010) A case-control study of body mass index and breast cancer risk in white and African-American women. Cancer Epidemiol Biomarkers Prev 19: 1532-1544.

36. Ogundiran TO, Huo D, Adenipekun A, Campbell O, Oyesegun R, et al. (2010) Case-control study of body size and breast cancer risk in Nigerian women. Am J Epidemiol 172: 682-690.

37. Boyd NF, Martin LJ, Sun L, Guo H, Chiarelli A, et al. (2006) Body size, mammographic density, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 15: 2086-2092.

38. Calle EE, Kaaks R (2004) Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. Nat Rev Cancer 4: 579-591.

39. Rose DP, Vona-Davis L (2010) Interaction between menopausal status and obesity in affecting breast cancer risk. Maturitas 66: 33-38.

40. Kaklamani V, Yi N, Sadim M, Siziopikou K, Zhang K, et al. (2011) The role of the fat mass and obesity associated gene (FTO) in breast cancer risk. BMC Med Genet 12: 52.

41. Liu C, Liu L (2011) Polymorphisms in three obesity-related genes (LEP, LEPR, and PON1) and breast cancer risk: a meta-analysis. Tumour Biol 32: 1233-1240.

42. da Cunha PA, de Carlos Back LK, Sereia AF, Kubelka C, Ribeiro MC, et al. (2013) Interaction between obesity-related genes, FTO and MC4R, associated to an increase of breast cancer risk. Mol Biol Rep 40: 6657-6664.

43. Feigelson HS, Teras LR, Diver WR, Tang W, Patel AV, et al. (2008) Genetic variation in candidate obesity genes ADRB2, ADRB3, GHRL, HSD11B1, IRS1, IRS2, and SHC1 and risk for breast cancer in the Cancer Prevention Study II. Breast Cancer Res 10: R57.

44. Fletcher O, Gibson L, Johnson N, Altmann DR, Holly JM, et al. (2005) Polymorphisms and circulating levels in the insulin-like growth factor system and risk of breast cancer: a systematic review. Cancer Epidemiol Biomarkers Prev 14: 2-19.

45. Choi JY, Lee KM, Park SK, Noh DY, Ahn SH, et al. (2005) Association of paternal age at birth and the risk of breast cancer in offspring: a case control study. BMC Cancer 5: 143.

46. Kim HC, Lee JY, Sung H, Choi JY, Park SK, et al. (2012) A genome-wide association study identifies a breast cancer risk variant in ERBB4 at

2q34: results from the Seoul Breast Cancer Study. Breast Cancer Res 14: R56.

47. Vimaleswaran KS, Tachmazidou I, Zhao JH, Hirschhorn JN, Dudbridge F, et al. (2012) Candidate genes for obesity-susceptibility show enriched association within a large genome-wide association study for BMI. Hum Mol Genet 21: 4537-4542.

48. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, et al. (2009) Pathway analysis by adaptive combination of P-values. Genet Epidemiol 33: 700-709.

49. Dudbridge F, Koeleman BP (2004) Efficient computation of significance levels for multiple associations in large studies of correlated data, including genomewide association studies. Am J Hum Genet 75: 424-435.

50. Connolly BS, Barnett C, Vogt KN, Li T, Stone J, et al. (2002) A meta-analysis of published literature on waist-to-hip ratio and risk of breast cancer. Nutr Cancer 44: 127-138.

51. Suzuki S, Kojima M, Tokudome S, Mori M, Sakauchi F, et al. (2013) Obesity/weight gain and breast cancer risk: findings from the Japan collaborative cohort study for the evaluation of cancer risk. J Epidemiol 23: 139-145.

52. Cheraghi Z, Poorolajal J, Hashem T, Esmailnasab N, Doosti Irani A (2012) Effect of body mass index on breast cancer during premenopausal and postmenopausal periods: a meta-analysis. PLoS One 7: e51446.

53. Nickels S, Truong T, Hein R, Stevens K, Buck K, et al. (2013) Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. PLoS Genet 9: e1003284.

54. Choi JY, Abel J, Neuhaus T, Ko Y, Harth V, et al. (2003) Role of alcohol and genetic polymorphisms of CYP2E1 and ALDH2 in breast cancer development. Pharmacogenetics 13: 67-72.

55. Chun NR (2004) A Study on Dietary Intake and Antioxdiant Mineral Status of Korean Women with Breast Cancer.

56. Do MH, Lee, S.S., (2003) Relation of breast cancer risk with alcohol consumption and physical activity: a case–control study. Korean Journal of Nutrition 36.

57. Iwasaki M, Otani T, Inoue M, Sasazuki S, Tsugane S (2007) Body size and risk for breast cancer in relation to estrogen and progesterone receptor status in Japan. Ann Epidemiol 17: 304-312.

58. Enger SM, Ross RK, Paganini-Hill A, Carpenter CL, Bernstein L (2000) Body size, physical activity, and breast cancer hormone receptor status: results from two case-control studies. Cancer Epidemiol Biomarkers Prev 9: 681-687.

59. Suzuki R, Rylander-Rudqvist T, Ye W, Saji S, Wolk A (2006) Body weight and postmenopausal breast cancer risk defined by estrogen and progesterone receptor status among Swedish women: A prospective cohort study. Int J Cancer 119: 1683-1689.

60. Joslyn SA (2002) Hormone receptors in breast cancer: racial differences in distribution and survival. Breast Cancer Res Treat 73: 45-59.

61. Ma L, Hanson RL, Traurig MT, Muller YL, Kaur BP, et al. (2010) Evaluation of A2BP1 as an obesity gene. Diabetes 59: 2837-2845.

62. Kiehl TR, Nechiporuk A, Figueroa KP, Keating MT, Huynh DP, et al. (2006) Generation and characterization of Sca2 (ataxin-2) knockout mice. Biochem Biophys Res Commun 339: 17-24.

63. Lastres-Becker I, Brodesser S, Lutjohann D, Azizov M, Buchmann J, et al. (2008) Insulin receptor and lipid metabolism pathology in ataxin-2 knock-out mice. Hum Mol Genet 17: 1465-1481.

64. Sadow PM, Chassande O, Gauthier K, Samarut J, Xu J, et al. (2003) Specificity of thyroid hormone receptor subtype and steroid receptor coactivator-1 on thyroid hormone action. Am J Physiol Endocrinol Metab 284: E36-46.

65. Shin DJ, Plateroti M, Samarut J, Osborne TF (2006) Two uniquely arranged thyroid hormone response elements in the far upstream 5' flanking region confer direct thyroid hormone regulation to the murine cholesterol 7alpha hydroxylase gene. Nucleic Acids Res 34: 3853-3861.

66. Park JW, Zhao L, Cheng SY (2013) Inhibition of estrogen-dependent tumorigenesis by the thyroid hormone receptor beta in xenograft models. Am J Cancer Res 3: 302-311.

67. Saraiva PP, Figueiredo NB, Padovani CR, Brentani MM, Nogueira CR

(2005) Profile of thyroid hormones in breast cancer patients. Braz J Med Biol Res 38: 761-765.

국문 초록

서론: 세계적으로 과체중 및 비만의 유병률이 증가하면서, 유방암을 포함한 복합질환에 미치는 비만의 영향에 대한 관심도 증대되고 있다. 기존에 비만 관련 유전자의 유전적 변이와 유방암 위험도 간의 연관성에 관하여 조사된 바 있으나, 유전적 인자 및 유전적-환경적 요인의 상호작용에 관한 제한적인 고려 및 주로 단일 SNP 분석에 초점을 맞추었다는 점에서 한계점을 지닌다. 이에 본 연구에서는 포괄적인 리뷰를 통한 비만관련 유전자를 선택하고 그것들의 개별 SNP 분석과 유전자 기반 경로 분석을 통해서 비만 관련 유전자의 유전적 변이와 유방암 위험도간의 연관성에 관해 조사하고자 하며, 비만과 유방암의 연관성에 영향을 미치는 유전적 요인의 조절 효과에 관해서도 살펴보고자 한다.

방법: 다기관 환자 대조군 연구에 포함된 1,786 명의 유방암환자와 1,789 명의 대조군을 대상으로 하여 기존에 비만의 표현형과 관련하여 문헌에 게재되거나 GWAS catalog (http://www.genome.gov/gwastudies)에 의해 알려진 비만 관련 유전자의 유전적 변이와 체질량지수 간에 전장유전체 연관분석을 우선적으로 수행하였다. BMI 와 관련 있는 SNP 과 유방암 위험도간의 연관성을 조사하고자 상가적유전자모델을 이용한 로지스틱 회귀

분석을 실시하였으며, 유전자 기반의 경로분석을 살펴보기 위해 1,000 순열과정을 통한 ARTP 방법을 사용하였다. 분석은 전체 대상자를 대상으로 평가되었으며, 폐경 상태, 비만 정도와 호르몬 수용체 상태에 따른 층화 분석도 수행되었다.

결과: 후보유전자 및 BMI 에 대한 GWAS 결과를 포함하여 이용 가능한 227,197 SNP 으로부터 유의수준 0.05 미만에서 BMI 와 연관성을 가지는 712 개의 유전자에 위치한 9,985 개의 SNP 을 확인되었으며, 그 중 GWAS catalog 의 408 개의 유전자(45.9%)와 문헌조사를 통한 후보유전자의 221 개의 유전자(38.7%)가 각각 확인되었다. *RBFOX1* 유전자에 있는 rs17804012 와 *LINC00317* 에 위치한 rs2014791 SNP 이 BMI 와 가장 강한 연관성을 가졌으며 (p<5E−6), 이와 유방암 위험도를 살펴보았을 때는 유의한 연관성이 관찰되지 않았다. 이에 반해 BMI 와 관련성 가진 644 개의 유전자 (p<0.05)를 포함한 유전자 기반경로분석을 수행했을 시 유방암 위험도와 유의한 연관성이 관찰되었다 (유전자 기반 경로분석의 유의수준 p=3E−3). 특히, *THRB* 유전자가 유방암 위험도에 미치는 영향이 가장 강하게 나타났으며 (p=9E−4) 이러한 연관성은 BMI 가 23.2 이상인 폐경 전 여성에서만 국한되어 나타났다 (p=5E−3). 호르몬 수용체 아형에 따른 층화 분석 시, *THRB* 유전자의 대표 SNP 인 rs6550865 는 주로 에스트로겐 수용체 양성 혹은 프로게

스테론 수용체 양성 유방암 위험도와 연관성이 나타났지만 아형 별 간에 연관성차이는 유의하지 않았다 ($P_{heterogeneity}$>0.05).

결론: *THRB* 유전자를 포함한 BMI 관련 유전자그룹이 유전자기반 경로분석에서 유방암과 유의한 연관성을 가짐을 확인하였으며 그러한 연관성은 폐경 상태와 BMI 에 따라 달라질 수 있음을 확인하였다. 본 결과는 환경적 요인을 고려한 비만의 복잡한 유전자 경로분석이 추가적인 유방암의 민감성 인자를 찾을 수 있음을 제시한다.

주요어: 유방암 · 비만 관련 유전자 · 경로 분석 · 다형성

학번: 2012-21811