Published OnlineFirst April 26, 2012; DOI:10.1158/1055-9965.EPI-12-0118

31. Gauderman W, Morrison J. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. Available from: http://hydra.usc.edu/gxe/. 2006.
32. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, et al. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res 1998;26:362–7.
33. Dentice M, Luongo C, Elefante A, Romino R, Ambrosio R, Vitale M, et al. Transcription factor Nkx-2.5 induces sodium/iodide symporter gene expression and participates in retinoic acid- and lactation-induced transcription in mammary cells. Mol Cell Biol 2004;24:7863–77.
34. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res 2002;12:996–1006.
35. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007;447:799–816.
36. Rayet B, Gelinas C. Aberrant rel/nfkb genes and activity in human cancer. Oncogene 1999;18:6938–47.
37. Lee CH, Jeon YT, Kim SH, Song YS. NF-kappaB as a potential molecular target for cancer therapy. Biofactors 2007;29:19–35.
38. Tschan MP, Reddy VA, Ress A, Arvidsson G, Fey MF, Torbett BE. PU.1 binding to the p53 family of tumor suppressors impairs their transcriptional activity. Oncogene 2008;27:3489–93.

Seaweed consumption and the risk of thyroid cancer in women: the Japan Public Health Center-based Prospective Study

Takehiro Michikawa[a,b], Manami Inoue[a], Taichi Shimazu[a], Norie Sawada[a], Motoki Iwasaki[a], Shizuka Sasazuki[a], Taiki Yamaji[a], Shoichiro Tsugane[a] and for the Japan Public Health Center-based Prospective Study Group

Iodine is a suspected risk factor for thyroid cancer. Seaweed accounts for about 80% of Japanese people's iodine intake. We examined the association between seaweed consumption and the risk of thyroid cancer in Japanese women. Women participating in the Japan Public Health Center-based Prospective Study ($n=52\,679$; age: 40–69 years) were followed up for a mean of 14.5 years; 134 new thyroid cancer cases, including 113 papillary carcinoma cases, were identified. Seaweed consumption was assessed using a food-frequency questionnaire and divided into three categories: 2 days/week or less (reference); 3–4 days/week; and almost daily. The Cox proportional hazards model was applied to estimate hazard ratios (HRs) and 95% confidence intervals (CIs). Seaweed consumption was clearly associated with an increased risk of papillary carcinoma (HR for almost daily consumption compared with 2 days/week or less = 1.71; 95% CI: 1.01–2.90; trend $P=0.04$). After stratification for menopausal status, an increased risk was observed in postmenopausal women (papillary carcinoma HR for almost daily consumption compared with 2 days/week or less = 3.81, 95% CI: 1.67–8.68; trend $P<0.01$), but not in premenopausal women (HR = 0.91, 95% CI: 0.44–1.91; trend $P=0.76$). This study identified a positive association between seaweed consumption and the risk of thyroid cancer (especially for papillary carcinoma) in postmenopausal women. *European Journal of Cancer Prevention* 21:254–260 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins.

European Journal of Cancer Prevention 2012, 21:254–260

Keywords: cohort study, Japanese, papillary carcinoma, seaweed, thyroid cancer

[a]Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo and [b]Environmental Epidemiology Section, Center for Environmental Health Sciences, National Institute for Environmental Studies, Tsukuba, Japan

Correspondence to Manami Inoue, MD, PhD, Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, 5-1-1 Tsukiji Chuo-ku, Tokyo 104 0045, Japan
Tel: +81 335 475 201 x3389; fax: +81 335 478 578;
e-mail: mnminoue@ncc.go.jp

Received 30 June 2011 Accepted 1 July 2011

Introduction

Few risk factors for thyroid cancer have been confirmed other than childhood exposure to ionizing radiation (World Cancer Research Fund/American Institute for Cancer Research, 2007; Dal Maso *et al.*, 2009); thus, investigations of other possible risk factors, including food and nutrition, in association with the risk of thyroid cancer have been carried out. Earlier researchers were particularly interested in the association between foods rich in iodine, an essential component of the thyroid hormones, and the risk of thyroid cancer (Bosetti *et al.*, 2001; Dal Maso *et al.*, 2009), because iodine intake is suspected to play a role in the development of thyroid cancer (Franceschi, 1998; World Cancer Research Fund/American Institute for Cancer Research, 2007). Several studies on this topic have been conducted over the years, but they had a predominantly case–control design (Dal Maso *et al.*, 2009), and the epidemiological evidence is too limited to draw any conclusions.

Seaweed is a common source of iodine. In Japan, seaweed is readily available, is considered healthy (Kim *et al.*, 2004), and is a part of the diet of most Japanese individuals [mean seaweed consumption = 13.5 g/day (Ministry of Health, Labour and Welfare, 2008)]; seaweed is the major source of iodine intake [cumulative percent contribution of seaweed to iodine intake ~80%; the cumulative contribution of other food groups such as fish and shellfish each makes up only a small percentage (Nagataki, 2008; Yasunaga *et al.*, 2009; Ministry of Health, Labour and Welfare, 2009)]. As high iodine intake is suspected to be a risk factor for thyroid cancer, high seaweed consumption may increase the risk of thyroid cancer. To date, two case–control studies have examined seaweed consumption in relation to the risk of thyroid cancer (Kolonel *et al.*, 1990; Horn-Ross *et al.*, 2001), but neither study showed an association. However, seaweed consumption was lower in the two study populations than it is in the Japanese population. Being iodine-sufficient (Maruchi *et al.*, 1971; Kikuchi *et al.*, 2008), the Japanese population, with its high rate of seaweed consumption (Iso *et al.*, 2005), represents a suitable platform for examining the effects of high seaweed consumption on the risk of thyroid cancer.

0959-8278 © 2012 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/CEJ.0b013e32834a8042

We hypothesized that seaweed consumption would be positively associated with the risk of thyroid cancer, particularly papillary carcinoma, in Japan. Papillary carcinoma is the predominant type of thyroid cancer in iodine-sufficient areas (Maruchi *et al.*, 1971; Williams *et al.*, 1977; Pettersson *et al.*, 1996); thus, high iodine intake is considered to increase the risk of papillary carcinoma. To investigate our hypothesis, we analyzed data from the Japan Public Health Center-based Prospective (JPHC) Study. Because the number of thyroid cancer cases in men (26 cases) was too small for meaningful analyses with sufficient statistical power, we included only women in this study.

Methods

Study population

The JPHC Study has been described in detail elsewhere (Tsugane and Sobue, 2001). In brief, we distributed a self-administered baseline questionnaire to elicit information on anthropometry data, medical history, health screening, lifestyle, dietary habits, and menstrual and reproductive history (for women) to all residents aged 40–59 years of five public health center areas in 1990 (Cohort I). In addition, a similar questionnaire was also distributed to all residents aged 40–69 years of six public health center areas between 1993 and 1994 (Cohort II). In this study, we excluded individuals from one area, because they provided no data on cancer incidence. Among eligible participants, 55 874 women (83%) returned the questionnaires. Those in whom cancer of any type had been diagnosed before the start of the study, those who provided incomplete responses on seaweed consumption, and those who reported extremes of calorie intake (upper and lower 1.0 percentiles) were excluded, leaving a total of 52 679 women for inclusion in the analysis (21 290 women in Cohort I and 31 389 women in Cohort II). The study was approved by the Institutional Review Board of the National Cancer Center (Tokyo, Japan).

Seaweed consumption

The questions on diet included in the questionnaire have been reported elsewhere (Tsugane *et al.*, 2001). The response options for the question on the average frequency of seaweed consumption during the previous month differed slightly between Cohort I and Cohort II; in Cohort I, there were four response options (rarely; 1–2 days/week; 3–4 days/week; and almost daily), whereas in Cohort II, there were five (never; occasionally; 1–2 days/week; 3–4 days/week; and almost daily). For statistical analysis, we combined the data from both cohorts, classifying the participants into three groups in terms of the frequency of seaweed consumption: 2 days/week or less; 3–4 days/week; and almost daily. This was justified because the number of participants in the lower consumption categories was small. We documented the validity of seaweed consumption data in subsamples, using 28-day dietary records. The Spearman rank coefficients between the frequency of seaweed consumption according to the questionnaire data and according to the dietary records were 0.33 for Cohort I (Tsubono *et al.*, 2003) and 0.40 for Cohort II (unpublished data). We could not collect information on the consumption of kelp supplements, but we assumed that the participants rarely used such supplements (Hirayama *et al.*, 2008), because they were not popular in Japan at the time the questionnaires were administered.

Follow-up

Participants were followed up from the baseline survey until 31 December 2007. Changes in residence status, including survival, were confirmed annually by the residential registry in each area. Deaths, certified according to the requirements of the Ministry of Health, Labour and Welfare, were ascertained from local public health centers (residence and death registration are required by law in Japan). Only 0.4% ($n = 192$) of the participants were lost to follow-up during the study period.

Incident cases of thyroid cancer were identified through a specific cancer registry system for the JPHC Study, which was established to collect cancer incidence data on the individuals living within the study area through a continuous surveillance of hospital records and population-based cancer registries. Death certificates were used as a source of supplementary information. The site and histological features of each case were coded according to the International Classification of Diseases for Oncology, third edition (ICD-O-3, code: C73.9) (World Health Organization, 2000). During the 766 327 person-years of follow-up (mean follow-up period: 14.5 years), we identified 134 new thyroid cancer cases. Of these, 84.3% (113 cases) were papillary carcinoma (ICD-O-3; morphology code: 8050, 8260, 8340, and 8350).

Statistical analysis

Person-years of observation were defined as the time from the date of response to the questionnaire to the date of diagnosis of thyroid cancer, the date of emigration from the study area, the date of death, or the end of the study period, whichever occurred first. Participants who were lost to follow-up were censored on the last confirmed date of their presence in the study area.

The Cox proportional hazards model was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the association between seaweed consumption and the total thyroid cancer risk as well as papillary carcinoma risk. For seaweed consumption categories, dummy variables were created, and the category of 2 days/week or less was regarded as the reference. After age-adjusted (continuous) and area-adjusted (10 public health centers) HRs with 95% CIs had been calculated, we adjusted for smoking history (never/past or current), [because

smoking influences the secretion of a thyroid-stimulating hormone associated with the risk of thyroid cancer (Mack *et al.*, 2003)], body mass index (BMI) (< 21.0, 21.0–22.9, 23.0–24.9, ≥ 25 kg/m^2) because of the strongly suspected risk factor of thyroid cancer (Kitahara *et al.*, 2011), green tea consumption (< 1, 1–2, 3–4, ≥ 5 cups/day) [because it had previously been associated with the risk of thyroid cancer in this population (Michikawa *et al.*, 2011)], total calorie intake (quartiles), health screening in the previous year (yes, no) (which might have increased the detection rate of thyroid cancer), and menopausal status [premenopausal, postmenopausal (natural or induced)]. We further added consumption of green vegetables (≤ 2, 3–4 days/week, almost daily) as a potential confounding factor because this variable changed the risk estimates by more than 10%. We also considered it necessary to adjust for alcohol consumption and consumption of other foods [fish and goitrogenic foods such as potatoes, pickled vegetables, and soy products, which inhibit the use of iodine in the thyroid gland (Gaitan, 1988; Doerge and Sheehan, 2002)] because seaweed consumption might be associated with these factors. However, adjustment for these factors did not substantially alter the results. Furthermore, to minimize the effects of different lifestyles and dietary behaviors, we stratified the participants according to smoking history (never/past or current), alcohol consumption (none or occasional/regular), BMI (< 25.0 or ≥ 25 kg/m^2), consumption of fish, green vegetables, potatoes, pickled vegetables, and soy products (≤ 2 or ≥ 3 days/week), and green tea consumption (< 1 or ≥ 1 cup/day). We then repeated the analyses after excluding cases of thyroid cancer occurring within the first 2 years of follow-up to confirm the temporal association. Trend tests were performed by allocating scores of 1, 2, and 3 for each seaweed consumption category.

Because there is an etiologic hypothesis that female hormones play a role in thyroid tumorigenesis (Chen *et al.*, 2008; Dal Maso *et al.*, 2009), it is possible that the risk factors for thyroid cancer vary between premenopausal and postmenopausal women. Therefore, we conducted stratified analyses with respect to menopausal status at baseline. We evaluated the potential interaction between seaweed consumption and menopausal status using a multiplicative interaction term with a likelihood ratio test.

All data were statistically analyzed using Stata version 11 (Stata Corp., College Station, Texas, USA).

Results

Table 1 provides background information on the women at baseline in each seaweed consumption category. Those who consumed seaweed almost daily had the highest mean age. The participants who consumed seaweed frequently were also likely to consume fish, green vegetables, potatoes, pickled vegetables, soy products, and green tea; they were also likely to have undergone health screening in the previous year.

Table 2 shows the HRs for thyroid cancer according to seaweed consumption. Seaweed consumption tended to be associated with an increased risk of thyroid cancer. When analysis was restricted to cases of papillary carcinoma, a statistically significant association was observed between seaweed consumption and the risk of cancer (HR = 1.71; 95% CI: 1.01–2.90; trend $P = 0.04$). This increased risk was found in all groups stratified according to lifestyles and dietary behaviors. The results were not substantially affected after we excluded participants with thyroid cancer diagnosed in the first 2 years of follow-up. In addition, as health screening and medical examinations might increase the likelihood of subclinical cancer detection, we further analyzed this association after excluding women with a history of chronic diseases including hypertension and diabetes mellitus (because such participants would probably be under close medical observation) and excluding cases of cancer confined within the thyroid gland. Even after this additional analysis, the positive association of seaweed consumption with the risk of papillary carcinoma persisted (trend $P = 0.04$).

A stratified analysis according to menopausal status showed a significant positive trend between seaweed consumption and the risk of thyroid cancer in postmenopausal women, but not in premenopausal women (Table 3). Compared with participants who ate seaweed 2 days/week or less, the multivariable HRs of papillary carcinoma risk for those who ate seaweed almost daily were 0.91 (95% CI: 0.44–1.91; trend $P = 0.76$) in premenopausal women and 3.81 (95% CI: 1.67–8.68; trend $P < 0.01$) in postmenopausal women. There was moderate evidence of an interaction between seaweed consumption and menopausal status for the risk of papillary carcinoma (P for interaction = 0.09). When we excluded women with a history of chronic diseases and patients with thyroid cancer confined within the thyroid gland, seaweed consumption was also positively associated with the risk of papillary carcinoma in postmenopausal women (17 cases; trend $P < 0.01$), but was not associated with the risk of thyroid cancer in premenopausal women (31 cases; trend $P = 0.49$).

Discussion

To our knowledge, this is the first prospective cohort study to report a positive association between seaweed consumption and the risk of thyroid cancer, or more specifically, the risk of papillary carcinoma. This positive association was clearly observed in postmenopausal women, but not in premenopausal women.

With regard to the observed association between seaweed consumption and the risk of thyroid cancer, we carefully

Table 1 Background information at baseline according to seaweed consumption

	Seaweed consumption		
	≤ 2 days/week (n=24 500)	3–4 days/week (n=18 105)	Almost daily (n=10 074)
Age, mean (SD)	51.8 (8.4)	51.7 (7.7)	52.4 (7.6)
Body mass index, ≥ 25 kg/m^2 (%)[a]	28.5	27.7	27.0
Current smokers (%)[a]	7.7	5.1	4.3
Regular drinkers (%)[a]	12.4	11.1	11.2
Fish consumption, almost daily (%)[a]	11.0	17.8	29.8
Green vegetable consumption, almost daily (%)[a]	18.9	33.3	56.0
Potatoes' consumption, almost daily (%)[a]	3.7	7.5	19.3
Pickled vegetables' consumption, almost daily (%)[a]	32.4	43.8	50.5
Soy products consumption, almost daily (%)[a]	22.6	38.7	64.7
Green tea consumption, ≥ 5 cups/day (%)[a]	25.9	28.7	29.9
Health screening in the previous year, yes (%)[a]	77.2	82.5	82.6
Menopausal status, premenopausal (%)[a]	34.8	33.1	32.7

[a]Age-standardized proportion.

Table 2 Hazard ratios and 95% confidence intervals for thyroid cancer according to seaweed consumption

	Seaweed consumption			
Category of consumption	≤ 2 days/week	3–4 days/week	Almost daily	Trend P
Total thyroid cancer				
Number of participants	24 500	18 105	10 074	
Number of cases	44	55	35	
Person-years of follow-up	344 385	270 006	151 935	
Age-adjusted and area-adjusted HR (95% CI)[a]	Reference	1.37 (0.91–2.06)	1.52 (0.96–2.40)	0.07
Multivariable HR1 (95% CI)[b]	Reference	1.22 (0.78–1.87)	1.41 (0.86–2.32)	0.17
Multivariable HR2 (95% CI)[b,c]	Reference	1.32 (0.81–2.15)	1.58 (0.91–2.73)	0.10
Papillary carcinoma				
Number of cases	34	45	34	
Age-adjusted and area-adjusted HR (95% CI)[a]	Reference	1.43 (0.90–2.25)	1.88 (1.15–3.07)	0.01
Multivariable HR1 (95% CI)[b]	Reference	1.20 (0.74–1.93)	1.71 (1.01–2.90)	0.04
Multivariable HR2 (95% CI)[b,c]	Reference	1.36 (0.80–2.32)	1.86 (1.03–3.34)	0.04

CI, confidence interval; HR, hazard ratio.
[a]Adjusted for age (continuous) and area (10 public health centers).
[b]Additionally adjusted for smoking habit (never/past, current), body mass index (<21.0, 21.0–22.9, 23.0–24.9, ≥ 25.0 kg/m^2), green vegetable consumption (≤ 2, 3–4 days/week, almost daily), green tea consumption (<1, 1–2, 3–4, ≥ 5 cups/day), total calorie intake (quartiles), health screening in the previous year (yes, no), and menopausal status [premenopausal, postmenopausal (natural or induced)].
[c]Women who developed thyroid cancer within the first 2 years of follow-up were excluded.

checked the possibility of detection bias. Socioeconomic status may have an effect on the frequency with which participants receive health screening and medical examinations. Advances in diagnostic tools have led to increased detection of subclinical thyroid cancer (Davies and Welch, 2006); thus, health screening and medical examinations might increase the likelihood of subclinical cancer detection. As information on socioeconomic status was not obtained from all of the participants, we adjusted for health screening in our multivariable model. Moreover, we performed an additional analysis excluding the factors possibly associated with the detection of thyroid cancer. Despite this, seaweed consumption was positively associated with papillary carcinoma. Therefore, detection bias was less likely to explain our findings.

The observed association of high seaweed consumption with an increased risk of thyroid cancer may be explained by the iodine content of seaweed. Seaweed is the major source of iodine intake in Japan [the cumulative percent contribution of seaweed to iodine intake ~80% (Nagataki, 2008; Yasunaga *et al.*, 2009; Ministry of Health, Labour and Welfare, 2009)]. Some experimental studies have related high iodine intake to the development of thyroid cancer (Kanno *et al.*, 1992; Takegawa *et al.*, 2000). In addition, there is indirect evidence that high iodine intake leads to autoimmune thyroiditis (Bagchi *et al.*, 1985; Zhu *et al.*, 1995; Teng *et al.*, 2006), which is associated with thyroid cancer, including papillary carcinoma (Ott *et al.*, 1987; Okayasu *et al.*, 1995; Antonaci *et al.*, 2009). However, we may not have fully assessed the daily intake of dietary iodine among our participants because our questionnaire did not ask about all food items containing iodine. The possibility that iodine is a mediator in the association between seaweed consumption and the risk of thyroid cancer warrants further prospective studies. In other words, other components contained in seaweed might be responsible for the association observed between seaweed consumption and the risk of thyroid cancer. For instance, seaweed contains inorganic arsenic (Nakamura *et al.*, 2008), which is associated with a risk of skin, lung, urinary bladder, kidney, and liver cancer (Tapio and Grosche,

Table 3 The association of seaweed consumption with the risk of thyroid cancer according to menopausal status[a]

	Seaweed consumption			
Category of consumption	≤ 2 days/week	3–4 days/week	Almost daily	Trend *P*
Total thyroid cancer				
Premenopausal women				
Number of participants	10402	7227	3624	
Number of cases	27	28	13	
Person-years of follow-up	146819	108561	54997	
Multivariable HR (95% CI)[b]	Reference	0.98 (0.56–1.73)	0.85 (0.41–1.75)	0.68
Postmenopausal women				
Number of participants	13275	10344	6163	
Number of cases	17	25	22	
Person-years of follow-up	186322	153632	92770	
Multivariable HR (95% CI)[b]	Reference	1.65 (0.84–3.25)	2.43 (1.18–4.98)	0.02
Papillary carcinoma				
Premenopausal women				
Number of cases	24	23	13	
Multivariable HR (95% CI)[b]	Reference	0.85 (0.46–1.57)	0.91 (0.44–1.91)	0.76
Postmenopausal women				
Number of cases	10	20	21	
Multivariable HR (95% CI)[b]	Reference	2.08 (0.93–4.63)	3.81 (1.67–8.68)	<0.01

CI, confidence interval; HR, hazard ratio.
[a]Analysis was restricted to those with complete information on menopausal status.
[b]Adjusted for age (continuous), area (10 public health centers), smoking habit (never/past, current), body mass index (<21.0, 21.0–22.9, 23.0–24.9, ≥ 25.0 kg/m^2), green vegetable consumption (≤ 2, 3–4 days/week, almost daily), green tea consumption (<1, 1–2, 3–4, ≥ 5 cups/day), total calorie intake (quartiles), and health screening in the previous year (yes, no).

2006); however, no epidemiologic evidence has been obtained so far indicating that inorganic arsenic is associated with the risk of thyroid cancer.

To date, two case–control studies have examined the association between seaweed consumption and the risk of thyroid cancer. One study carried out in Hawaii did not show any association between seaweed consumption (mean daily intake ~2.0 g/day) and the risk of thyroid cancer (Kolonel *et al.*, 1990). In another study of women in San Francisco, California, an inverse association was found between seaweed consumption and the risk of papillary carcinoma (odds ratio for consumption of ≥ 0.33 g/day vs. nonconsumers = 0.61; 95% CI: 0.44–0.84) (Horn-Ross *et al.*, 2001). Residents of Hawaii and California consume less seaweed than those of Japan [mean daily intake = 13.5 g/day (Ministry of Health, Labour and Welfare, 2008)]. Given the disparity in seaweed consumption, the cumulative percent contribution of seaweed to iodine intake must be lower in these populations than it is in the Japanese population. Thus, seaweed consumption did not appear to be an adequate proxy marker of iodine intake in either study. However, iodine intake as estimated by means of a food-frequency questionnaire in another study was not found to be significantly associated with the risk of thyroid cancer (Kolonel *et al.*, 1990; Truong *et al.*, 2010). Horn-Ross *et al.* (2001) reported no association between iodine exposure measured using toenail clippings as a biomarker and papillary carcinoma, but they found an inverse association when iodine intake was assessed through questionnaires. However, they also report that their measurements of dietary iodine intake may not have been optimal because of the limitations of dietary iodine databases. Therefore, the biomarker findings seem more relevant than the questionnaire results. The levels of iodine intake in the three populations used in the above studies were lower than the mean iodine intake in Japanese individuals [1.5 mg/day (Ministry of Health, Labour and Welfare, 2009)]. Considering all these points, it is not unreasonable to conclude that seaweed consumption, as a surrogate marker of iodine intake, may indeed be a risk factor for thyroid cancer in Japanese women.

Our study suggested a positive association between seaweed consumption and the risk of thyroid cancer in postmenopausal women, but not in premenopausal women. The mean seaweed consumption (g/day) tended to increase with age (Ministry of Health, Labour and Welfare, 2008); thus, it is possible that the difference observed in the risk of thyroid cancer simply reflects different levels of seaweed consumption. However, there is also a biological difference related to female hormones in the effects of seaweed on premenopausal and postmenopausal women (Chen *et al.*, 2008; Dal Maso *et al.*, 2009). Estrogens appear to promote the proliferation and growth of thyroid cancer cells through estrogen receptor α (ERα) (Chen *et al.*, 2008). In papillary carcinoma, the expression of ERα is higher in premenopausal women than that in postmenopausal women (Kawabata *et al.*, 2003). A recent experimental study revealed the antiestrogenic bioactivity of seaweed extract: administration of seaweed led to a reduction in circulating 17β-estradiol levels in rats, and the binding of estradiol to ERα was inhibited by a seaweed extract *in vitro* (Skibola *et al.*, 2005). This implies that antiestrogenic bioactive compounds, including seaweed, are protectively associated with premenopausal thyroid cancer risk, but not with postmenopausal thyroid cancer risk because of

the relatively low levels of estrogens and ERα in postmenopausal women. Therefore, the antiestrogenic bioactive compounds in seaweed may counteract the adverse effects of the iodine in seaweed on the risk of thyroid cancer in premenopausal women, meaning that only postmenopausal women have an increased risk of thyroid cancer related to iodine intake. Although it is not impossible that the difference we observed resulted from chance alone, because of the relatively small number of events in each stratified category by menopausal status, the difference is nevertheless noteworthy and merits further investigation.

The strengths of the present study include its prospective design, negligible loss to follow-up (0.4%), and the large sample of the Japanese general population enrolled. The fact that the participants were from an iodine-sufficient population with a high mean value of seaweed consumption also lends importance to this study. However, our study has some limitations. On the question of seaweed consumption, the correlation coefficients between the questionnaire results and dietary records were reasonable (0.33 for Cohort I; 0.40 for Cohort II). Misclassification in assessing seaweed consumption was, however, inevitable, but it was likely to have been random with respect to outcome. Hence, such a misclassification would have tended to attenuate rather than exaggerate the true association between seaweed consumption and the risk of thyroid cancer. In addition, as with all observational studies, unmeasured confounding factors (e.g. external radiation) and unknown factors could exist. We did not evaluate benign thyroid diseases (Dal Maso *et al.*, 2009) as a potential confounding factor, because only one patient reported such a history.

Conclusion

We observed an increased risk of thyroid cancer, especially papillary carcinoma, in relation to high seaweed consumption in Japanese women, and our results indicate that this association may differ according to menopausal status. However, the underlying mechanism of the risk of thyroid cancer because of seaweed consumption remains unknown, and further studies on this topic are required.

Acknowledgements

The authors thank all staff members in each study area and in the central offices for their cooperation and technical assistance. We are indebted to the Aomori, Iwate, Ibaraki, Niigata, Osaka, Kochi, Nagasaki, and Okinawa Cancer Registries for providing their incidence data.

This work was supported by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare, Japan (until 2009), and by Management Expenses Grants from the Government to the National Cancer Center (since 2010). It was also supported by a Grant-in-Aid for the Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labour and Welfare, Japan, and by a Grant-in-Aid from Keio Medical Association.

Members of the Japan Public Health Center-based Prospective Study (JPHC Study, principal investigator: S. Tsugane) Group are: S. Tsugane, M. Inoue, T. Sobue, and T. Hanaoka (National Cancer Center, Tokyo); J. Ogata, S. Baba, T. Mannami, A. Okayama, and Y. Kokubo (National Cardiovascular Center, Osaka); K. Miyakawa, F. Saito, A. Koizumi, Y. Sano, I. Hashimoto, T. Ikuta, and Y. Tanaba (Iwate Prefectural Ninohe Public Health Center, Iwate); Y. Miyajima, N. Suzuki, S. Nagasawa, Y. Furusugi, and N. Nagai (Akita Prefectural Yokote Public Health Center, Akita); H. Sanada, Y. Hatayama, F. Kobayashi, H. Uchino, Y. Shirai, T. Kondo, R. Sasaki, Y. Watanabe, Y. Miyagawa, Y. Kobayashi, and M. Machida (Nagano Prefectural Saku Public Health Center, Nagano); Y. Kishimoto, E. Takara, T. Fukuyama, M. Kinjo, M. Irei, and H. Sakiyama (Okinawa Prefectural Chubu Public Health Center, Okinawa); K. Imoto, H. Yazawa, T. Seo, A. Seiko, F. Ito, F. Shoji, and R. Saito (Katsushika Public Health Center, Tokyo); A. Murata, K. Minato, K. Motegi, T. Fujieda, and S. Yamato (Ibaraki Prefectural Mito Public Health Center, Ibaraki); K. Matsui, T. Abe, M. Katagiri, and M. Suzuki (Niigata Prefectural Kashiwazaki and Nagaoka Public Health Center, Niigata); M. Doi, A. Terao, Y. Ishikawa, and T. Tagami (Kochi Prefectural Chuo-higashi Public Health Center, Kochi); H. Sueta, H. Doi, M. Urata, N. Okamoto, F. Ide, and H. Goto (Nagasaki Prefectural Kamigoto Public Health Center, Nagasaki); H. Sakiyama, N. Onga, H. Takaesu, and M. Uehara (Okinawa Prefectural Miyako Public Health Center, Okinawa); F. Horii, I. Asano, H. Yamaguchi, K. Aoki, S. Maruyama, M. Ichii, and M. Takano (Osaka Prefectural Suita Public Health Center, Osaka); Y. Tsubono (Tohoku University, Miyagi); K. Suzuki (Research Institute for Brain and Blood Vessels Akita, Akita); Y. Honda, K. Yamagishi, S. Sakurai, and N. Tsuchiya (Tsukuba University, Ibaraki); M. Kabuto (National Institute for Environmental Studies, Ibaraki); M. Yamaguchi, Y. Matsumura, S. Sasaki, and S. Watanabe (National Institute of Health and Nutrition, Tokyo); M. Akabane (Tokyo University of Agriculture, Tokyo); T. Kadowaki (Tokyo University, Tokyo); M. Noda and T. Mizoue (International Medical Center of Japan, Tokyo); Y. Kawaguchi (Tokyo Medical and Dental University, Tokyo); Y. Takashima and Y. Yoshida (Kyorin University, Tokyo); K. Nakamura (Niigata University, Niigata); S. Matsushima and S. Natsukawa (Saku General Hospital, Nagano); H. Shimizu (Sakihae Institute, Gifu); H. Sugimura (Hamamatsu University, Shizuoka); S. Tominaga (Aichi Cancer Center Research Institute, Aichi); H. Iso (Osaka University, Osaka); M. Iida, W. Ajiki, and A. Ioka (Osaka Medical Center for Cancer and Cardiovascular Disease, Osaka); S. Sato (Chiba

Prefectural Institute of Public Health, Chiba); E. Maruyama (Kobe University, Hyogo); M. Konishi, K. Okada, and I. Saito (Ehime University, Ehime); N. Yasuda (Kochi University, Kochi); S. Kono (Kyushu University, Fukuoka).

Conflicts of interest

There are no conflicts of interest.

References

Antonaci A, Consorti F, Mardente S, Giovannone G (2009). Clinical and biological relationship between chronic lymphocytic thyroiditis and papillary thyroid carcinoma. *Oncol Res* **17**:495–503.

Bagchi N, Brown TR, Urdanivia E, Sundick RS (1985). Induction of autoimmune thyroiditis in chickens by dietary iodine. *Science* **230**:325–327.

Bosetti C, Kolonel L, Negri E, Ron E, Franceschi S, Dal Maso L, *et al.* (2001). A pooled analysis of case-control studies of thyroid cancer. VI. Fish and shellfish consumption. *Cancer Causes Control* **12**:375–382.

Chen GG, Vlantis AC, Zeng Q, van Hasselt CA (2008). Regulation of cell growth by estrogen signaling and potential targets in thyroid cancer. *Curr Cancer Drug Targets* **8**:367–377.

Dal Maso L, Bosetti C, La Vecchia C, Franceschi S (2009). Risk factors for thyroid cancer: an epidemiological review focused on nutritional factors. *Cancer Causes Control* **20**:75–86.

Davies L, Welch HG (2006). Increasing incidence of thyroid cancer in the United States, 1973–2002. *JAMA* **295**:2164–2167.

Doerge DR, Sheehan DM (2002). Goitrogenic and estrogenic activity of soy isoflavones. *Environ Health Perspect* **110 (Suppl 3)**:349–353.

Franceschi S (1998). Iodine intake and thyroid carcinoma – a potential risk factor. *Exp Clin Endocrinol Diabetes* **106 (Suppl 3)**:S38–S44.

Gaitan E (1988). Goitrogens. *Baillieres Clin Endocrinol Metab* **2**:683–702.

Hirayama F, Lee AH, Binns CW, Watanabe F, Ogawa T (2008). Dietary supplementation by older adults in Japan. *Asia Pac J Clin Nutr* **17**:280–284.

Horn-Ross PL, Morris JS, Lee M, West DW, Whittemore AS, McDougall IR, *et al.* (2001). Iodine and thyroid cancer risk among women in a multiethnic population: the Bay Area Thyroid Cancer Study. *Cancer Epidemiol Biomarkers Prev* **10**:979–985.

Iso H, Date C, Noda H, Yoshimura T, Tamakoshi A (2005). Frequency of food intake and estimated nutrient intake among men and women: the JACC Study. *J Epidemiol* **15 (Suppl 1)**:S24–S42.

Kanno J, Onodera H, Furuta K, Maekawa A, Kasuga T, Hayashi Y (1992). Tumor-promoting effects of both iodine deficiency and iodine excess in the rat thyroid. *Toxicol Pathol* **20**:226–235.

Kawabata W, Suzuki T, Moriya T, Fujimori K, Naganuma H, Inoue S, *et al.* (2003). Estrogen receptors (alpha and beta) and 17beta-hydroxysteroid dehydrogenase type 1 and 2 in thyroid disorders: possible in situ estrogen synthesis and actions. *Mod Pathol* **16**:437–444.

Kikuchi Y, Takebayashi T, Sasaki S (2008). Iodine concentration in current Japanese foods and beverages [in Japanese with English abstract]. *Nippon Eiseigaku Zasshi* **63**:724–734.

Kim MK, Sasaki S, Sasazuki S, Tsugane S (2004). Prospective study of three major dietary patterns and risk of gastric cancer in Japan. *Int J Cancer* **110**:435–442.

Kitahara CM, Platz EA, Freeman LE, Hsing AW, Linet MS, Park Y, *et al.* (2011). Obesity and thyroid cancer risk among US men and women: a pooled analysis of five prospective studies. *Cancer Epidemiol Biomarkers Prev* **20**:464–472.

Kolonel LN, Hankin JH, Wilkens LR, Fukunaga FH, Hinds MW (1990). An epidemiologic study of thyroid cancer in Hawaii. *Cancer Causes Control* **1**:223–234.

Mack WJ, Preston-Martin S, Dal Maso L, Galanti R, Xiang M, Franceschi S, *et al.* (2003). A pooled analysis of case-control studies of thyroid cancer: cigarette smoking and consumption of alcohol, coffee, and tea. *Cancer Causes Control* **14**:773–785.

Maruchi N, Furihata R, Makiuchi M (1971). Population surveys on the prevalence of thyroid cancer in a non-endemic region, Nagano, Japan. *Int J Cancer* **17**: 575–583.

Michikawa T, Inoue M, Shimazu T, Sasazuki S, Iwasaki M, Sawada N, *et al.* (2011). Green tea and coffee consumption and its association with thyroid cancer risk: a population-based cohort study in Japan. *Cancer Causes Control* **22**:985–993.

Ministry of Health, Labour and Welfare (2008). The National Health and Nutrition Survey 2006 [in Japanese]. Available at *http://www.mhlw.go.jp/bunya/kenkou/eiyou06/01.html.* [Accessed 10 June 2011]

Ministry of Health, Labour and Welfare (2009). Dietary reference intakes for Japanese, 2010 (in Japanese). Available at *http://www.mhlw.go.jp/shingi/2009/05/s0529-4.html.* [Accessed 10 June 2011]

Nagataki S (2008). The average of dietary iodine intake due to the ingestion of seaweeds is 1.2 mg/day in Japan. *Thyroid* **18**:667–668.

Nakamura Y, Narukawa T, Yoshinaga J (2008). Cancer risk to Japanese population from the consumption of inorganic arsenic in cooked hijiki. *J Agric Food Chem* **56**:2536–2540.

Okayasu I, Fujiwara M, Hara Y, Tanaka Y, Rose NR (1995). Association of chronic lymphocytic thyroiditis and thyroid papillary carcinoma. A study of surgical cases among Japanese, and white and African Americans. *Cancer* **76**: 2312–2318.

Ott RA, McCall AR, McHenry C, Jarosz H, Amin A, Lawrence AM, *et al.* (1987). The incidence of thyroid carcinoma in Hashimoto's thyroiditis. *Am Surg* **53**: 442–445.

Pettersson B, Coleman MP, Ron E, Adami HO (1996). Iodine supplementation in Sweden and regional trends in thyroid cancer incidence by histopathologic type. *Int J Cancer* **65**:13–19.

Skibola CF, Curry JD, VandeVoort C, Conley A, Smith MT (2005). Brown kelp modulates endocrine hormones in female Sprague-Dawley rats and in human luteinized granulosa cells. *J Nutr* **135**:296–300.

Takegawa K, Mitsumori K, Onodera H, Shimo T, Kitaura K, Yasuhara K, *et al.* (2000). Studies on the carcinogenicity of potassium iodide in F344 rats. *Food Chem Toxicol* **38**:773–781.

Tapio S, Grosche B (2006). Arsenic in the aetiology of cancer. *Mutat Res* **612**:215–246.

Teng W, Shan Z, Teng X, Guan H, Li Y, Teng D, *et al.* (2006). Effect of iodine intake on thyroid diseases in China. *N Engl J Med* **354**:2783–2793.

Truong T, Baron-Dubourdieu D, Rougier Y, Guenel P (2010). Role of dietary iodine and cruciferous vegetables in thyroid cancer: a countrywide case-control study in New Caledonia. *Cancer Causes Control* **21**: 1183–1192.

Tsubono Y, Kobayashi M, Sasaki S, Tsugane S (2003). Validity and reproducibility of a self-administered food frequency questionnaire used in the baseline survey of the JPHC Study Cohort I. *J Epidemiol* **13 (Suppl 1)**:S125–S133.

Tsugane S, Sobue T (2001). Baseline survey of JPHC study – design and participation rate. Japan Public Health Center-based Prospective Study on Cancer and Cardiovascular Diseases. *J Epidemiol* **11 (Suppl 6)**: S24–S29.

Tsugane S, Sasaki S, Kobayashi M, Tsubono Y, Sobue T (2001). Dietary habits among the JPHC study participants at baseline survey. Japan Public Health Center-based Prospective Study on Cancer and Cardiovascular Diseases. *J Epidemiol* **11 (Suppl 6)**:S30–S43.

Williams ED, Doniach I, Bjarnason O, Michie W (1977). Thyroid cancer in an iodide rich area: a histopathological study. *Cancer* **39**:215–222.

World Cancer Research Fund/American Institute for Cancer Research (2007). *Food, nutrition, physical activity, and the prevention of cancer: a global perspective.* Washington DC: AICR.

World Health Organization (2000). *International classification of diseases for oncology.* 3rd ed. Geneva, Switzerland: World Health Organization.

Yasunaga M, Ishikawa J, Ujike A, Nishioka C (2009). Study on daily intake of iodine from daily foods in Kagawa prefecture [in Japanese]. Kagawaken Environmental Health Research Center Report; **8**:75-78.

Zhu YP, Bilous M, Boyages SC (1995). Excess iodine induces the expression of thyroid solid cell nests in lymphocytic thyroiditis-prone BB/W rats. *Autoimmunity* **20**:201–206.

Genome-Wide Association Study in East Asians Identifies Novel Susceptibility Loci for Breast Cancer

**Jirong Long[1☯], Qiuyin Cai[1☯], Hyuna Sung[2], Jiajun Shi[1], Ben Zhang[1], Ji-Yeob Choi[2], Wanqing Wen[1], Ryan J. Delahanty[1], Wei Lu[3], Yu-Tang Gao[4], Hongbing Shen[5], Sue K. Park[2,6,7], Kexin Chen[8], Chen-Yang Shen[9,10], Zefang Ren[11], Christopher A. Haiman[12], Keitaro Matsuo[13], Mi Kyung Kim[14,15], Ui Soon Khoo[16], Motoki Iwasaki[17], Ying Zheng[3], Yong-Bing Xiang[4], Kai Gu[3], Nathaniel Rothman[18], Wenjing Wang[3], Zhibin Hu[5], Yao Liu[5], Keun-Young Yoo[6], Dong-Young Noh[7,19], Bok-Ghee Han[20], Min Hyuk Lee[21], Hong Zheng[8], Lina Zhang[8], Pei-Ei Wu[9], Ya-Lan Shieh[9], Sum Yin Chan[22], Shenming Wang[23], Xiaoming Xie[24], Sung-Won Kim[7,19], Brian E. Henderson[12], Loic Le Marchand[25], Hidemi Ito[13], Yoshio Kasuga[26], Sei-Hyun Ahn[14,15], Han Sung Kang[14,15], Kelvin Y. K. Chan[16,27], Hiroji Iwata[28], Shoichiro Tsugane[17], Chun Li[29], Xiao-Ou Shu[1], Dae-Hee Kang[2,6,7], Wei Zheng[1]*

1 Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America, **2** Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea, **3** Shanghai Center for Disease Control and Prevention, Shanghai, China, **4** Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China, **5** Department of Epidemiology and Biostatistics, Nanjing Medical University, Nanjing, China, **6** Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea, **7** Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea, **8** Department of Epidemiology and Biostatistics, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China, **9** Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, **10** College of Public Health, China Medical University, Taichung, Taiwan, **11** School of Public Health, Sun Yat-sen University, Guangzhou, China, **12** Department of Preventive Medicine, Keck School of Medicine, University of Southern California and Norris Comprehensive Cancer Center, Los Angeles, California, United States of America, **13** Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan, **14** Cancer Epidemiology Division, Breast Cancer Center, National Cancer Center, Goyang, Korea, **15** Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Ulsan, Korea, **16** Department of Pathology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China, **17** Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan, **18** Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, United States of America, **19** Department of Surgery, Seoul National University College of Medicine, Seoul, Korea, **20** Center for Genome Science, Korea National Institute of Health, Seoul, Korea, **21** Department of Surgery, Soonchunhyang University Hospital, Bucheon, Korea, **22** Department of Clinical Oncology, Queen Mary Hospital, Hong Kong, China, **23** First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, **24** Cancer Center, Sun Yat-sen University, Guangzhou, China, **25** Epidemiology Program, Cancer Research Center, University of Hawaii, Honolulu, Hawaii, United States of America, **26** Department of Surgery, Nagano Matsushiro General Hospital, Nagano, Japan, **27** Department of Obstetrics and Gynecology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China, **28** Department of Breast Oncology, Aichi Cancer Center Central Hospital, Nagoya, Japan, **29** Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee, United States of America

Abstract

Genetic factors play an important role in the etiology of both sporadic and familial breast cancer. We aimed to discover novel genetic susceptibility loci for breast cancer. We conducted a four-stage genome-wide association study (GWAS) in 19,091 cases and 20,606 controls of East-Asian descent including Chinese, Korean, and Japanese women. After analyzing 690,947 SNPs in 2,918 cases and 2,324 controls, we evaluated 5,365 SNPs for replication in 3,972 cases and 3,852 controls. Ninety-four SNPs were further evaluated in 5,203 cases and 5,138 controls, and finally the top 22 SNPs were investigated in up to 17,423 additional subjects (7,489 cases and 9,934 controls). SNP rs9485372, near the TGF-β activated kinase (*TAB2*) gene in chromosome 6q25.1, showed a consistent association with breast cancer risk across all four stages, with a *P*-value of 3.8×10^{-12} in the combined analysis of all samples. Adjusted odds ratios (95% confidence intervals) were 0.89 (0.85–0.94) and 0.80 (0.75–0.86) for the A/G and A/A genotypes, respectively, compared with the genotype G/G. SNP rs9383951 ($P=1.9\times10^{-6}$ from the combined analysis of all samples), located in intron 5 of the *ESR1* gene, and SNP rs7107217 ($P=4.6\times10^{-7}$), located at 11q24.3, also showed a consistent association in each of the four stages. This study provides strong evidence for a novel breast cancer susceptibility locus represented by rs9485372, near the *TAB2* gene (6q25.1), and identifies two possible susceptibility loci located in the *ESR1* gene and 11q24.3, respectively.

Citation: Long J, Cai Q, Sung H, Shi J, Zhang B, et al. (2012) Genome-Wide Association Study in East Asians Identifies Novel Susceptibility Loci for Breast Cancer. PLoS Genet 8(2): e1002532. doi:10.1371/journal.pgen.1002532

Editor: Mark I. McCarthy, University of Oxford, United Kingdom

Received July 21, 2011; **Accepted** December 23, 2011; **Published** February 23, 2012

Funding: This research was supported in part by U.S. National Institutes of Health grants R01CA124558, R01CA148667, R01CA64277, and R37CA070867, as well as Ingram Professorship and Research Reward funds to W Zheng; R01CA118229, R01CA092585, and Department of Defense (DOD) Idea Award BC011118 to X-O Shu; R01CA122756 and DOD Idea Award BC050791 to Q Cai; R01 CA137013 to J Long. Sample preparation and genotyping assays at Vanderbilt were conducted at the Survey and Biospecimen Shared Resources and Vanderbilt Microarray Shared Resource, which are supported in part by Vanderbilt-Ingram Cancer Center (P30 CA68485). The SeBCS was supported by BRL (Basic Research Laboratory) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2011-0001564). The KOHBRA/KOGES was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health, Welfare, and Family Affairs, Republic of Korea (#1020350). Participating studies (Principal Investigator, grant support) in the consortium are as

follows: the Shanghai Breast Cancer Study (W Zheng, R01CA64277), the Shanghai Women's Health Study (W Zheng, R37CA070867), the Shanghai Breast Cancer Survival Study (X-O Shu, R01CA118229), the Shanghai Endometrial Cancer Study (X-O Shu, R01CA092585, contributing only controls to the consortium), the Seoul Breast Cancer Study (D-H Kang, the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea, 0620410-1), the Nanjing Study (H Shen, 09KJA330001, Jiangsu, China), the Tianjin Study (K Chen, the National Natural Science Foundation of China Grant No. 30771844), the Taiwan Biobank Study (C-Y Shen, DOH97-01), the Hong Kong Study (US Khoo, Research Grant Council, Hong Kong SAR, China, HKU 7520/05M and 76730M), the Guangzhou Breast Cancer Study (Z Ren, the National Natural Science Foundation of China Grant No. 81072383), the Multiethnic Cohort Study (BE Henderson, CA63464; LN Kolonel, CA54281; and CA Haiman, CA132839), the Nagano Breast Cancer Study (S Tsugane, Grants-in-Aid for the Third Term Comprehensive Ten-Year Strategy for Cancer Control from the Ministry of Health, Labor, and Welfare of Japan, and for Scientific Research on Priority Areas, 17015049, and for Scientific Research on Innovative Areas, 221S0001, from the Ministry of Education, Culture, Sports, Science, and Technology of Japan), and the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (K Tajima, Grants-in-Aid for Scientific Research on Priority Areas, 17015052, from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; H Tanaka, Grants-in-Aid for the Third Term Comprehensive Ten-Year Strategy for Cancer Control from the Ministry of Health, Labor, and Welfare of Japan, H20-002). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agents. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: wei.zheng@vanderbilt.edu

☯ These authors contributed equally to this work.

Introduction

Breast cancer is one of the most common malignancies diagnosed among women worldwide, including those living in East Asian countries. Genetic factors play an important role in the etiology of both sporadic and familial breast cancer [1]. In the past two decades, more than 1,000 reports have been published addressing the association between variants in candidate genes and breast cancer risk. However, only a few genetic risk factors have been confirmed for this common malignancy [2]. Recent genome-wide association studies (GWAS) have identified approximately 20 common genetic susceptibility loci for breast cancer [3–14]. However, these newly-identified genetic factors, along with known high-penetrance breast cancer susceptibility genes explain less than 30% of the heritability for this cancer [2,15]. Furthermore, most GWAS were conducted among women of European ancestry, and many of the variants discovered in European-ancestry populations showed only a weak or no association with breast cancer in other ethnic groups [16,17]. For example, only 8 of 12 breast cancer risk SNPs identified in women of European ancestry were directly replicated in Chinese population [18]. Therefore, GWAS conducted in non-European women are needed to fully uncover the genetic basis for breast cancer susceptibility. Herein, we report results from a large GWAS of breast cancer conducted in East Asian women.

Results

A total of 19,091 female breast cancer cases and 20,606 female controls—including 23,891 Chinese, 11,907 Korean and 3,809 Japanese women—were included in the present study (Table 1). In Stage I, we analyzed 690,947 SNPs in 2,918 breast cancer cases and 2,324 community controls recruited from studies conducted in Shanghai, China (Figure 1, Text S1). Top 5,365 SNPs were investigated in Stage IIa including 1,613 Chinese cases and 1,800 Chinese controls recruited from studies conducted in Shanghai, China. Of the SNPs evaluated, 68 SNPs showed an association with breast cancer risk at $P \leq 0.05$ with the same direction as observed in Stage I. We performed a meta-analysis for the remaining 4,913 SNPs with data available from both Stage IIa and Stage IIb (2,359 Korean cases and 2,052 Korean controls). Twenty-six SNPs showed an association with breast cancer risk with $P_{meta} \leq 0.05$ and the association was consistent among Stages I, IIa and IIb. These SNPs, along with the 68 SNPs mentioned above, were selected for Stage III replication in 4,712 cases and 4,496 controls. Finally, based on the results of the first three stages, 22 top SNPs were selected for Stage IV evaluation in 7,489 cases and 9,934 controls.

SNP rs9485372 showed a statistically significant association with breast cancer risk in each of the four stages (Table 2). The OR (95% CI) per A allele was 0.88 (0.81–0.95), 0.86 (0.81–0.92), 0.94 (0.88–1.00) and 0.90 (0.85–0.94), respectively, for stages I to IV. The association with this SNP was remarkably consistent across all but one small study (Figure 2A). Pooled analysis of samples from all studies produced OR (95% CI) of 0.90 (0.87–0.92) and *P*-value of 3.8×10^{-12}, which is substantially lower than the conventional genome-wide significance level of 5×10^{-8} based on conservative Bonferroni adjustment of multiple comparisons at $\alpha = 0.05$, providing strong evidence for an association of this SNP with breast cancer risk.

Two other SNPs, rs9383951 and rs7107217, were also consistently replicated in each of the three replication sets. The C allele of rs9383951 was associated with decreased risk with OR (95% CI) of 0.82 (0.73–0.93), 0.90 (0.81–1.00), 0.91 (0.82–1.00), and 0.88 (0.81–0.96), respectively, for stages I to IV (Table 2). The P-value reached 1.9×10^{-6} in the pooled analysis of samples from all four stages. For SNP rs7107217, the ORs (95% CI) per C allele were 1.13 (1.04–1.23), 1.11 (1.04–1.18), 1.07 (1.00–1.14) and 1.05 (1.01–1.10), respectively, for stages I to IV, respectively (Table 2). Analyses with all subjects combined showed OR (95% CI) of 1.08 (1.05–1.11) and *P* value of 4.6×10^{-7}. Again, the association of breast cancer risk with these two SNPs was very consistent across the vast majority of participating studies (Figure 2B and 2C).

Stratified analyses showed that the associations with these three SNPs were consistent in all three East Asian populations, although the association for SNPs rs9485372 and rs7107217 was not significant for Japanese subjects, probably due to a small sample size (Table 3). Associations of these three SNPs with breast cancer risk were similar when stratified by menopausal or estrogen receptor status and none of the heterogeneity tests was statistically significant (Table S1). No significant interaction was observed with other risk factors (Table S1). After adjusted for the top 5 or 10 principal components, the results did not change significantly (Table S2).

Both SNPs rs9485372 and rs9383951 are located at chromosome 6q25.1, approximately 2.34 Mb and 350 kb from the SNP rs2046210 that we previously reported for breast cancer risk [8]. None of these three SNPs, however, are in LD ($r^2<0.1$) in any of the three populations (Asian, European and Africans) as determined using data generated in the HapMap or any of the study populations included in the current study (Table S3 and Figure S1). In an analysis including all 30,153 subjects who were genotyped for three SNPs in 6q25.1, all three SNPs remained strongly associated with breast cancer risk after mutual adjustment of the other 2 SNPs with P values of 1.4×10^{-12}, 1.3×10^{-4}, and 6.0×10^{-39} for SNPs

Author Summary

Breast cancer is one of the most common malignancies among women worldwide. Genetic factors play an important role in the etiology of breast cancer. To identify common genetic susceptibility alleles for breast cancer, we performed a four-stage genome-wide association study in 19,091 cases and 20,606 controls among East-Asian women. Single nucleotide polymorphism (SNP) rs9485372, near the TGF-beta activated kinase 1 (*TAB2*) gene at chromosome 6q25.1, was associated with breast cancer risk ($P=3.8\times10^{-12}$). SNPs rs9383951, located in intron 5 of the estrogen receptor 1 (*ESR1*) gene, and rs7107217, located at 11q24.3, were also consistently associated with breast cancer risk in all four stages with a combined *P* of 1.9×10^{-6} and 4.6×10^{-7}, respectively. This study provides strong evidence for a novel breast cancer susceptibility locus represented by rs9485372, near the *TAB2* gene (6q25.1), and identifies two possible susceptibility loci located in the *ESR1* gene and 11q24.3, respectively.

rs9485372, rs9383951 and rs2046210, respectively (Table S4). No significant interaction was observed for these three SNPs (Table S5). We also created a genetic risk score (GRS) to evaluate the combined effect of three SNPs located in 6q25.1 (Table S6). Compared with women carrying 0–1 risk variants, women carrying 6 variants had over two-fold increased risk with an OR (95% CI) of 2.36 (1.89–2.96) and a P value of 1.3×10^{-47}.

A total of 376 SNPs were successfully imputed in the LD blocks including rs2046210 and rs9485372 and the whole *ESR1* gene with RSQ≥0.3 and minor allele frequency (MAF)≥0.05. Among them, 27 SNPs showed an association with breast cancer risk with P≤0.05 after adjusted for age, rs9485372, rs9383951 and rs2046210 (Table S7). With the exception of rs4591859 and rs7776340 in the locus of rs2046210 and rs7768330 in the locus of rs9383921, all other SNPs are in the same LD block within the *ESR1* gene (Figure S2). No additional SNP in the rs9485372 locus showed an association with breast cancer risk at $p<0.05$ after adjusted for rs9485372, rs2046210, and rs9383921.

Discussion

In this large GWAS conducted in East-Asian women including 19,091 cases and 20,606 controls, we provided strong evidence for a novel breast cancer susceptibility locus represented by rs9485372 and suggestive evidence for two other loci, represented by SNPs rs9383951 and rs7107217.

We previously reported a genetic susceptibility locus at 6q25.1, represented by rs2046210, for breast cancer risk [8]. The newly identified SNPs, rs9485372 and rs9383951, also are located at chromosome 6q25.1. However, these three SNPs are not in LD and are thus representing independent breast cancer susceptibility loci. All of them were associated with breast cancer risk after mutual adjustment of the other two SNPs. SNP rs9485372 is approximately 31 Kb upstream of the TGF-β activated kinase 1/MAP3K7 binding protein 2 (*TAB2*) gene (Figure 3). The protein encoded by this gene is an activator of MAP3K7/TAK1, which is required for the IL-1 induced activation of NF-κB and MAPK8/JNK. The TGF-β pathway plays a major role in breast cancer development and progression [19]. The MAP kinases pathway is critical in regulating cell growth and cell death [20] and may

Table 1. Selected characteristics of studies participating in the Asia Breast Cancer Consortium.

Study Stage[a]	Ethnicity	No. of cases	No. of controls	age[b]	Menopause (%)[c]	ER+ (%)
Stage I						
Shanghai-I	Chinese	2,918	2,324	51.7/50.3[d]	42.9/41.7	65.3
Stage II						
Shanghai-II (IIa)	Chinese	1,613	1,800	53.2/53.4	50.2/55.1	62.5
SeBCS-I (IIb)	Korean	2,359	2,052	48.1/51.7	37.9/52.0	61.9
Stage III						
Shanghai-III	Chinese	2,601	2,386	53.8/55.1[d]	50.3/52.6	64.9
Taiwan	Chinese	1,066	1,065	51.5/47.5[d]	52.3/39.9	66.1
Nagoya	Japanese	644	644	51.4/51.1	48.5/48.5	72.8
Nagano	Japanese	401	401	53.8/54.0	54.9/65.3	74.6
Stage IV						
Nanjing	Chinese	1,786	1,837	50.6/50.2	51.3/47.6	55.7
Tianjin	Chinese	1,297	1,585	51.9/51.9	51.9/55.5	44.2
Guangzhou	Chinese	838	865	49.0/49.2	41.8/51.9	71.6
NCC	Korean	505	504	49.0/49.1	49.5/45.3	65.0
SeBCS-II	Korean	777	1,104	47.5/47.7	36.3/37.3	63.0
KOHBRA/KoGES	Korean	1,397	3,209	40.5/50.3[d]	23.3/	62.8
MEC	Japanese	889	830	66.5/66.5		85.3
Total		19,091	20,606			

[a]See the methods section for the full names of participating studies.
[b]Mean value for cases/controls.
[c]Percentage for cases/controls.
[d]Significant at $\alpha=0.01$ level.
doi:10.1371/journal.pgen.1002532.t001

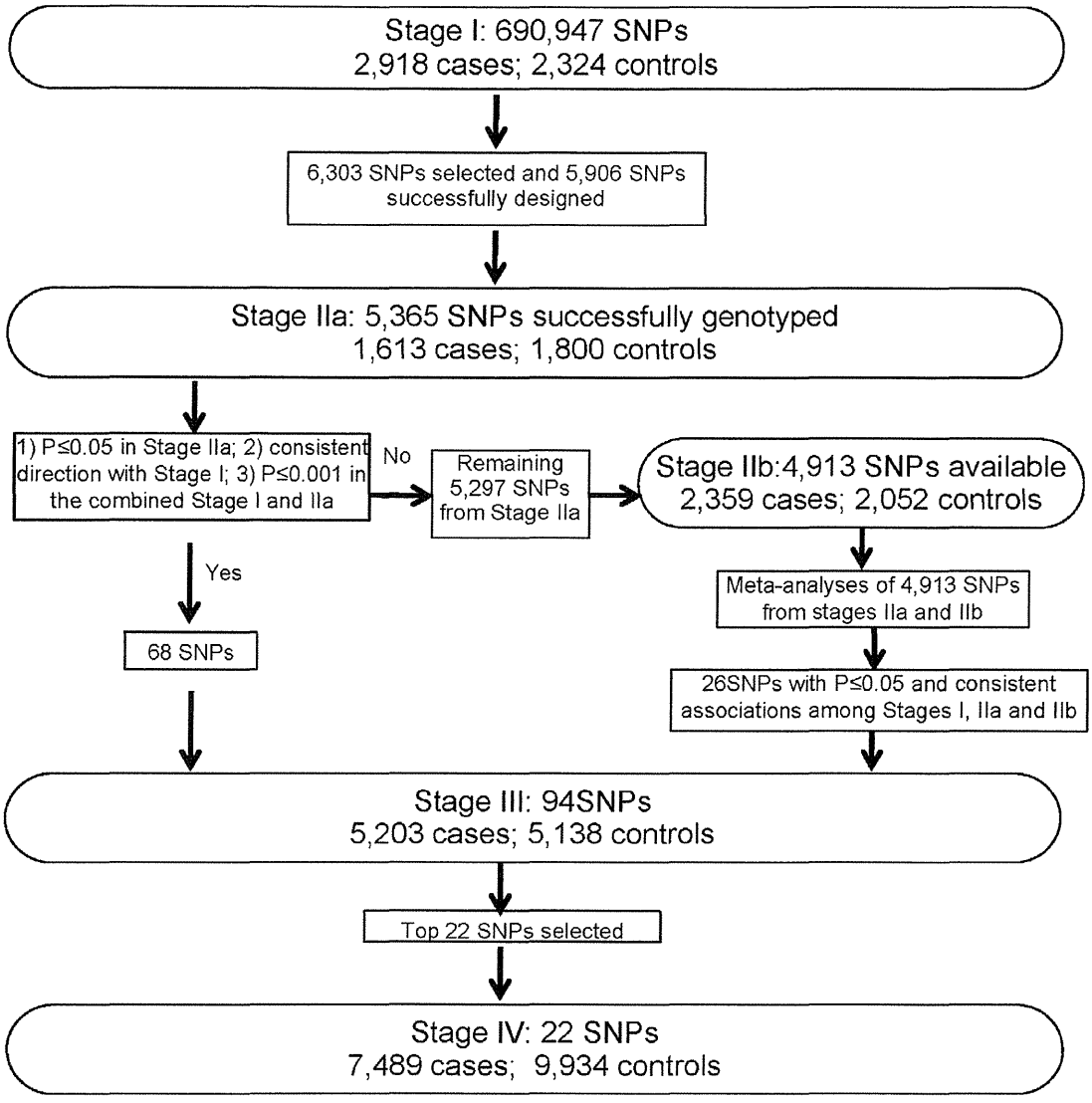

Figure 1. Overview of the study design.
doi:10.1371/journal.pgen.1002532.g001

contribute to the development of cancer [20]. Furthermore, the TAB2 protein is required for DNA damage-induced TAK1 activation, suggesting that TAB2 may play a role in DNA damage repair [21]. Other genes in the region identified in the study included *SUMO4*, *LATS1*, *PPIL4*, and *UST*. However, given the proximity of the *TAB2* gene with rs9485372 and the important role of this gene in breast carcinogenesis, it is possible that the association between rs9485372 and breast cancer risk may be mediated through the *TAB2* gene. It is also possible that the association may be mediated through regulating the *ESR1* gene, located approximately 2.5 Mb from rs9485372. This possibility was highlighted by a recent study showing that several open reading frames in the 6q25.1 regions co-expressed with *ESR1* [22]. Further research is warranted to clarify the mechanism of the association identified in the study.

SNP rs9383951 is located in intron 5 of the *ESR1* gene, an important gene that has been documented to play a key role in breast cancer development and progression. Previous candidate gene studies have extensively evaluated two SNPs, rs2234693 (PvuII) and rs9340799 (XbaI), in the *ESR1* gene in relation to breast cancer risk; the results, however, have been inconsistent [2]. Neither rs2234693 nor rs9340799 are in LD ($r^2<0.01$) with the SNPs discovered in the present study. To follow-up the lead from our previous study reporting a susceptibility locus at 6q25.1 for breast cancer [8], two recent studies conducted among women of European descent identified rs3757318 and rs9397435 in relation to breast cancer risk [11,23]. These two SNPs are in strong LD ($r^2>0.6$ in Asians) with the SNP (rs2046210) we previously reported at 6q25.1 in East Asians but not in other populations. Again, these two SNPs are not in LD ($r^2<0.01$ in Asian, European and African populations) with rs9383951 and rs9485372 identified in this study. Although the association with rs9383951 did not reach the conventional genome-wide significance, the fact that this SNP is located in the *ESR1* gene strongly suggests a true association of this SNP with breast cancer risk.

Table 2. Summary of results for the three SNPs showing a statistically or marginally significant association in all four stages with breast cancer risk, the Asia Breast Cancer Consortium.

SNP[a]	Position[b]	Study	No. of Cases/Controls	EAF (%)[c]	Per allele OR (95%CI) d	P value[d]
rs9485372 (A/G)	149650567 (6q25.1)					
		Stage I	2,770/2,175	43.5	0.88(0.81–0.95)	1.4×10^{-3}
		Stage II	3,930/3,818	47.1	0.86(0.81–0.92)	6.3×10^{-6}
		Stage III	4,081/4,074	43.2	0.94(0.88–1.00)	0.05
		Stage IV	5,186/7,440	46.2	0.90(0.85–0.94)	4.2×10^{-5}
		All stages	15,967/17,507	45.4	0.90(0.87–0.92)	3.8×10^{-12}
rs9383951 (C/G)	152337306 (6q25.1)					
		Stage I	2,916/2,319	11.4	0.82(0.73–0.93)	2.4×10^{-3}
		Stage II	3,948/3,836	10.1	0.90(0.81–1.00)	0.06
		Stage III	4,581/4,433	9.7	0.91(0.82–1.00)	0.06
		Stage IV	6,117/8,296	9.6	0.88(0.81–0.96)	3.3×10^{-3}
		All stages	17,562/18,884	10.0	0.88(0.84–0.93)	1.9×10^{-6}
rs7107217 (C/A)	128978900 (11q24.3)					
		Stage I	2,916/2,319	31.4	1.13(1.04–1.23)	3.6×10^{-3}
		Stage II	3,929/3,839	34.8	1.11(1.04–1.18)	2.1×10^{-3}
		Stage III	4,606/4,424	35.2	1.07(1.00–1.14)	0.04
		Stage IV	7,348/9,831	37.4	1.05(1.01–1.10)	0.02
		All stages	18,799/20,413	35.8	1.08(1.05–1.11)	4.6×10^{-7}

[a]Effect/reference alleles based on forward strand.
[b]From NCBI genome build 36.
[c]Effect allele frequency in controls.
[d]Adjusted for age and study sites.
doi:10.1371/journal.pgen.1002532.t002

SNP rs7107217 also showed a consistent association in all four stages, although the pooled *P*-value did not reach the conventional genome-wide significance level. This SNP is located at 11q24.3, 152 Kb downstream of the *BARX2* gene and 212 Kb upstream of the *TMEM45B* gene (Figure S3). *BARX2* is a homeobox gene for which the mouse ortholog has been shown to influence cellular processes that control cell adhesion and cytoskeleton remodeling. It has been shown, BARX2 and estrogen receptor-alpha (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion [24]. BARX2 also acts in a tumor suppressor and loss of heterozygosity of this gene, lead to poorer survival in patients with ovarian cancer [25].

It could be ideal to increase the sample size in the discovery stage and simplify the replication stages of the study. However, like many other consortium projects, financial constraints and some logistical issues prevented us for achieving the maximum statistical power. Nevertheless, with approximately 40,000 cases and controls, our study represents the largest breast cancer genetic association study in East Asian women. This consortium will continue to provide valuable resources to identify additional novel susceptibility loci for breast cancer.

In summary, in this large GWAS conducted in East Asia women, we provided convincing evidence for an association with a novel independent susceptibility locus located at 6q25.1, near the *TAB2* gene. Our study also suggests that genetic variants in the *ESR1* gene and chromosome 11q24.3 may be related to breast cancer risk. Given that multiple independent breast cancer susceptibility loci have identified in our studies and studies conducted by others in 6q25.1 that harbors the *ESR1* gene, it is possible that 6q25.1 may represent an important region for breast cancer susceptibility.

Methods

Study populations

Included in this consortium project were 19,091 cases and 20,606 controls from 14 studies (Table 1). Detailed descriptions of these participating studies and demographic characteristics of study participants are provided in Text S1. Briefly, the consortium included 23,981 Chinese women, 11,907 Korean women, 3,809 Japanese women. The Chinese women were from 8 studies: Shanghai [n = 13,642, Shanghai Breast Cancer Study, Shanghai Breast Cancer Survival Study (SBCSS), Shanghai Endometrial Cancer Study (SECS), Shanghai Women Health Study (SWHS)] [8,26], Nanjing (n = 3,623) [27], Tianjin (n = 2,882) [28], Taiwan (n = 2,131) [29], and Guangzhou (n = 1,703). The Korean women were from four studies [Seoul Breast Cancer Study (SeBCS) (n = 6,292) [30], Korea NCC (n = 1,009), KoGES (n = 3,209) [31], and KOHBRA (n = 1,397) [32]]. The Japanese women were from three studies conducted in Hawaii and Los Angeles [n = 1,719; Multiethnic Cohort Study (MEC) [33]], Nagoya (n = 1,288) [34], and Nagano (n = 802) [35] (Table 1). Approval was granted from relevant institutional review boards in all study sites; all included subjects gave informed consent.

Genotyping methods

The Genotyping protocol for Stage I has been described previously [8]. Briefly, the initial 300 subjects were genotyped using the Affymetrix GeneChip Mapping 500K Array Set. The remaining 4,985 subjects were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0. We included one negative control and at least three positive quality control (QC) samples from the Coriell Cell Repositories (http://ccr.coriell.org/) in each

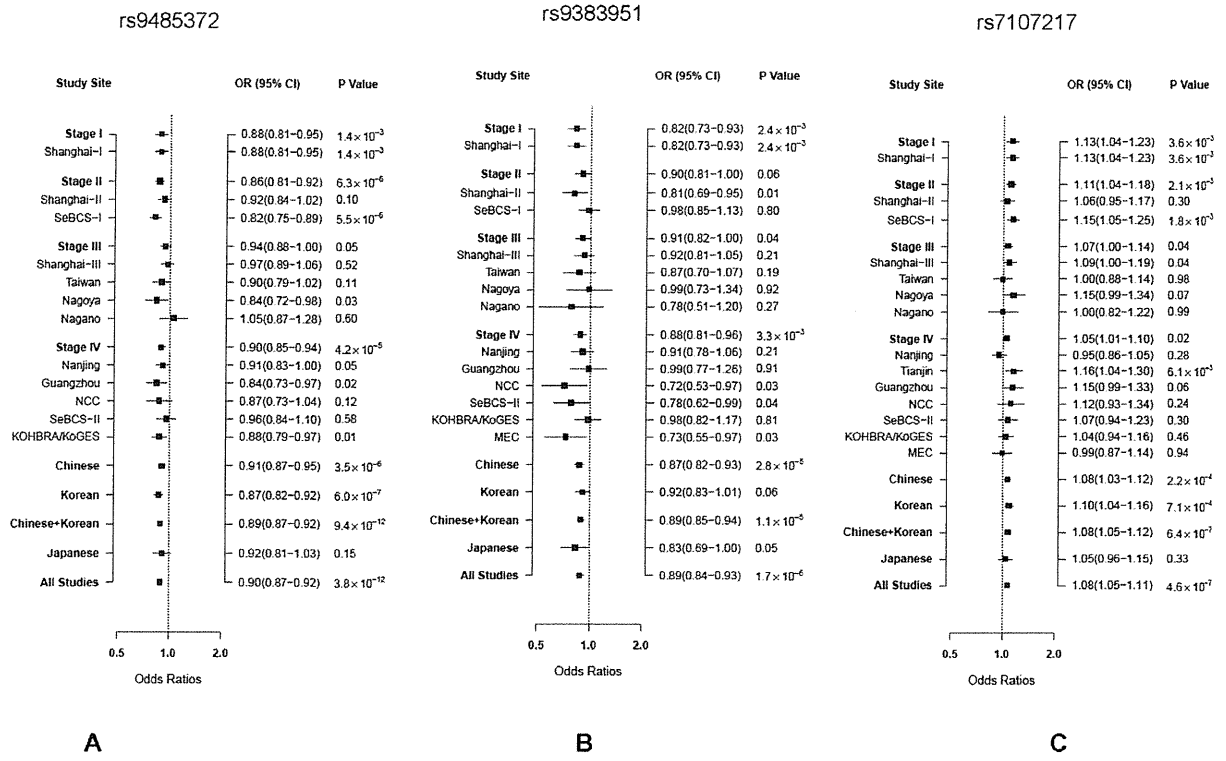

Figure 2. ORs per risk allele and 95% CIs for breast cancer associated with three SNPs by study site and ethnicity. A: rs9485372, B: rs9383951; and C: rs7107217.
doi:10.1371/journal.pgen.1002532.g002

Table 3. Association of SNPs with breast cancer risk by ethnic groups, the Asia Breast Cancer Consortium.

SNP	Study	No. of Cases/Controls	EAF (%)[a]	OR (95% CI)[b]		P value[b]
				Heterozygote	Homozygote	
rs9485372						
	Chinese	9,922/9,644	43.2	0.90(0.84–0.96)	0.83(0.76–0.90)	3.5×10^{-6}
	Korean	5,006/6,825	48.2	0.87(0.79–0.95)	0.76(0.68–0.85)	6.0×10^{-7}
	Chinese+Korean	14,928/16,469	45.2	0.89(0.85–0.94)	0.80(0.75–0.85)	9.4×10^{-12}
	Japanese	1,039/1,038	47.5	0.93(0.76–1.13)	0.84(0.66–1.07)	0.15
	All studies	15,967/17,507	45.4	0.89(0.85–0.94)	0.80(0.75–0.86)	3.8×10^{-12}
rs9383951						
	Chinese	10,625/10,180	10.7	0.86(0.80–0.92)	0.87(0.67–1.13)	3.4×10^{-5}
	Korean	5,011/6,833	9.7	0.92(0.83–1.02)	0.79(0.52–1.19)	0.06
	Chinese+Korean	15,636/17,013	10.3	0.88(0.83–0.93)	0.86(0.69–1.07)	1.3×10^{-5}
	Japanese	1,926/1,871	6.8	0.86(0.71–1.05)	0.40(0.14–1.13)	0.05
	All studies	17,562/18,884	10.0	0.88(0.83–0.93)	0.83(0.67–1.03)	1.9×10^{-6}
rs7107217						
	Chinese	11,887/11,719	32.3	1.09(1.03–1.15)	1.14(1.05–1.25)	2.2×10^{-4}
	Korean	4,987/6,824	38.7	1.13(1.04–1.23)	1.19(1.06–1.34)	7.1×10^{-4}
	Chinese+Korean	16,874/18,543	34.6	1.10(1.05–1.15)	1.16(1.08–1.24)	6.4×10^{-7}
	Japanese	1,925/1,870	47.3	1.09(0.94–1.27)	1.09(0.91–1.31)	0.33
	All studies	18,799/20,413	35.8	1.10(1.05–1.15)	1.15(1.08–1.22)	4.6×10^{-7}

[a]Effect allele frequency in controls.
[b]Adjusted for age and study sites.
doi:10.1371/journal.pgen.1002532.t003

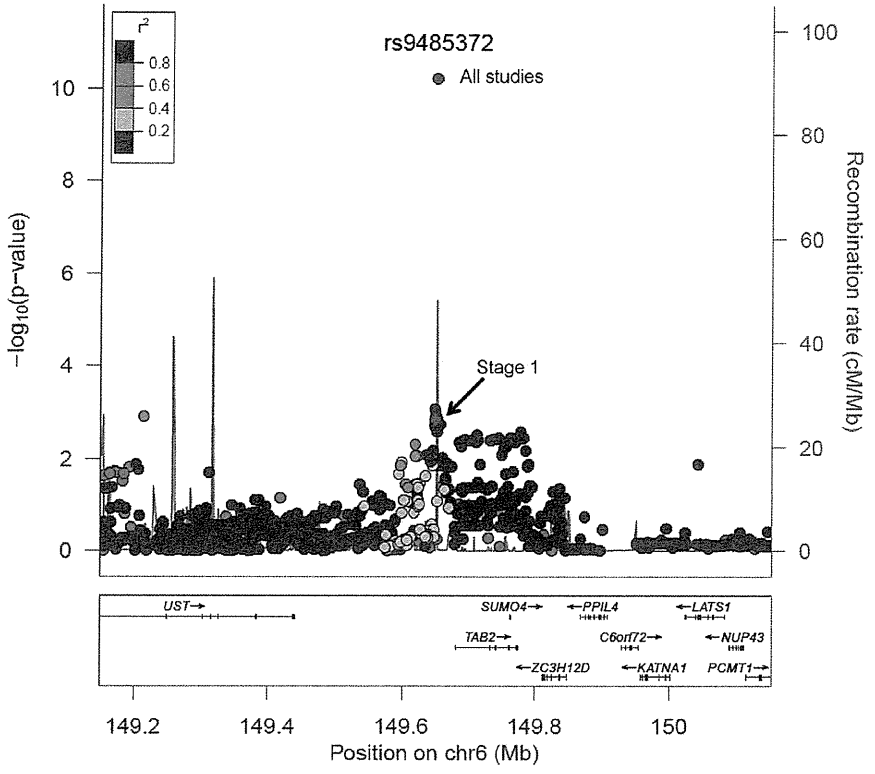

Figure 3. A regional plot of the $-\log_{10}$P-values for SNPs at 6q25.1. The LD is estimated using data from HapMap Asian population. Also shown are the SNP Build 36 coordinates in kilobases (Kb), recombination rates in centimorgans (cM) per megabase (Mb) and genes in the region (below) based on the March 2006 UCSC genome browser assembly.
doi:10.1371/journal.pgen.1002532.g003

of the 96-well plates for Affymetrix SNP Array 6.0 genotyping. A total of 273 positive QC samples were successfully genotyped, and the average concordance rate was 99.9% with a median value of 100%. The sex of all study samples was confirmed to be female. Genetically identical, unexpected duplicated samples were excluded, as were close relatives with a pair-wise proportion of identify-by-descent (IBD) estimate greater than 0.25. All samples with a call rate<95% were excluded. The SNPs were excluded if: (i) MAF<1%, (ii) call rate<95%, or (iii) genotyping concordance rate<95% in quality control samples. The final dataset included 2,918 cases and 2,324 controls for 690,947 markers. There are 21,223 SNPs that were on Affymetrix 500K Array Set but not on the Affymetrix SNP Array 6.0. These SNPs were excluded. SNPs on the Affymetrix 6.0 array but not on the Affymetrix 500k array were treated as missing data for those samples genotyped on using the Affymetrix 500k array. Similar results were obtained after excluding women genotyped by Affymetrix 500K Array Set from the analyses.

Genotyping for Stage IIa was completed using the Illumina iSelect platform. To compare the consistency between the Affymetrix and Illumina iSelect platforms, we also included 43 samples from Stage I that were genotyped by Affymetrix SNP 6.0. Similar to the QC procedures used in Stage I, the following criteria were used to exclude samples: (i) call rate<95%; or (ii) unexpected duplicated samples based on IBD estimate. SNPs were excluded if: (i) call rate<95%, or (ii) genotyping concordance rate<95% in quality control samples when compared with Affymetrix 6.0 data. After QC, the mean concordance rate was 99.85% between Illumina iSelect and Affymetrix 6.0 genotyping.

Data for the SNPs analyzed in Stage IIb were extracted from the Korean GWAS genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 chip. A total of 30 QC samples were successfully genotyped, and the concordance rate was 99.83%. The sex of all samples was confirmed to be female. The SNPs were excluded if: (1) genotype call rate<95%, (2) MAF<1% in either the cases or controls, (3) deviation from HWE at P-value$<10^{-6}$, and (4) poor cluster plot in either the cases or controls.

Genotyping for Stage III and all samples from Koreans in Stage IV was completed using the iPLEX Sequenom MassArray platform in the Vanderbilt Molecular Epidemiology Laboratory. Included in each 96-well plate as QC samples were one negative control (water), two blinded duplicates, and two samples from the HapMap project. To compare the consistency between the Affymetrix and Sequenom platforms, we also genotyped 45 samples included in Stage I. The mean concordance rate was 99.67% for the blind duplicates, 98.88% for HapMap samples, and 99.52% between Sequenom and Affymetrix 6.0 genotyping. Data quality from the Hong Kong study was low and thus data from the study were excluded for the current analysis. Genotyping for two Chinese studies (Nanjing and Guangzhou) in Stage IV was completed using the iPLEX Sequenom MassArray platform at the Fudan University, Shanghai, China. Blind duplicate QC samples were included and the mean concordance rate was 98.70%. Genotyping for the Tianjin study in Stage IV was performed using TaqMan assays. Genotyping assay protocols were developed and validated at the Vanderbilt Molecular Epidemiology Laboratory, and TaqMan genotyping assay reagents were provided to investigators of the Tianjin study (Tianjin Cancer Institute and

Hospital). For the MEC study, data for the three SNPs presented in this study were extracted from the GWA scan data generated using Illumina 660W. For SNPs not included on the chip, imputed data using HapMap as reference were extracted. Genotype frequencies for SNP rs9485372 deviated from HWE in controls (P = 0.004), therefore, this SNP was excluded in data analyses. Not all SNPs for Stage IV were genotyped in all studies included in Stage IV due to genotyping failure or the use of different genotyping platforms (Table S8).

SNP selection for replication

SNP selection for Stage II replication: Promising SNPs were selected for replication in Stage II based on the following criteria: 1) minor allele frequency (MAF)≥5%; 2) $P<0.02$ in Stage I; 3) Hardy-Weinberg equilibrium (HWE) test $P>1.0\times10^{-6}$ in controls; 4) not in strong linkage disequilibrium (LD) ($r^2<0.5$) with any of the previously confirmed breast cancer genetic risk variants or SNPs evaluated in our previous studies [8,12]; and 5) high genotyping quality as indicated by very clear genotyping clusters checked manually. When multiple SNPs are in LD with $r^2\geq0.5$, one SNP with the lowest *P*-value was selected. In total, 6,303 SNPs were selected for replication. A total of 5,906 SNPs (93.7%) were successfully designed by Illumina and included in the iSelect array. After stringent QC procedures, data from 5,365 SNPs were considered high quality for association analyses in Stage IIa, which include 1,613 breast cancer patients and 1,800 controls recruited from Shanghai studies.

SNP selection for Stage III replication: Among the 5,365 SNPs successfully genotyped in Stage IIa, 68 SNPs were selected for Stage III replication in an independent set of 5,203 cases and 5,138 controls recruited from Shanghai and several other East Asian populations (Table 1 and Text S1). The selection criteria are: 1) an association with breast cancer risk in Stage IIa with P≤0.05; 2) the direction of the association consistent in both stages; and 3) P≤0.001 in the merged data of Stage I and IIa. During the course of Stage III genotyping, genome-wide association scan data from 2,359 cases and 2,052 controls were obtained from the Seoul Breast Cancer GWAS (Stage IIb). Therefore, we performed a meta-analysis of Stage IIa and IIb data. Of the 5,297 SNPs which were not selected initially for Stage III replication based on Stage IIa data alone, data were available for 4,913 SNPs in Stage IIb. Meta-analyses of these 4,913 SNPs from Stage IIa and IIb yielded 26 additional SNPs that showed an association at $P\leq0.05$ and in the same direction among stages I, IIa, and IIb. These 26 SNPs were then added to the list of SNPs to be genotyped in Stage III.

SNP selection for Stage IV replication: Based on the results of the first three stages, 22 top SNPs were selected for Stage IV evaluation and genotyped in up to 17,423 additional subjects (7,489 cases and 9,934 controls) (Table 1 and Text S1).

Statistical analyses

Case-control differences in selected demographic characteristics and major risk factors were evaluated using t-tests (for continuous variables) and Chi-square tests (for categorical variables). Associations between SNPs and breast cancer risk were assessed using odds ratios (ORs) and 95% confidence intervals (CIs) derived from logistic regression models. ORs were estimated for heterozygote and homozygote for the variant allele compared with homozygotes for the common allele. ORs were also estimated for the variant allele based on a log-additive model and adjusted for age, and study site, when appropriate. Stratified analyses by ethnicity, menopausal status, and estrogen receptor (ER) status were carried out. PLINK version 1.06 was used to analyze genome-wide data obtained in Stage I and the replication data in Stage IIa. Results from Stage IIb were also obtained from PLINK version 1.06. Meta-analyses of Stage IIa and Stage IIb were performed using a weighted z-statistics method, where weights were proportional to the square root of the number of individuals in each sample and standardized such that the weights added up to one. The z-statistic

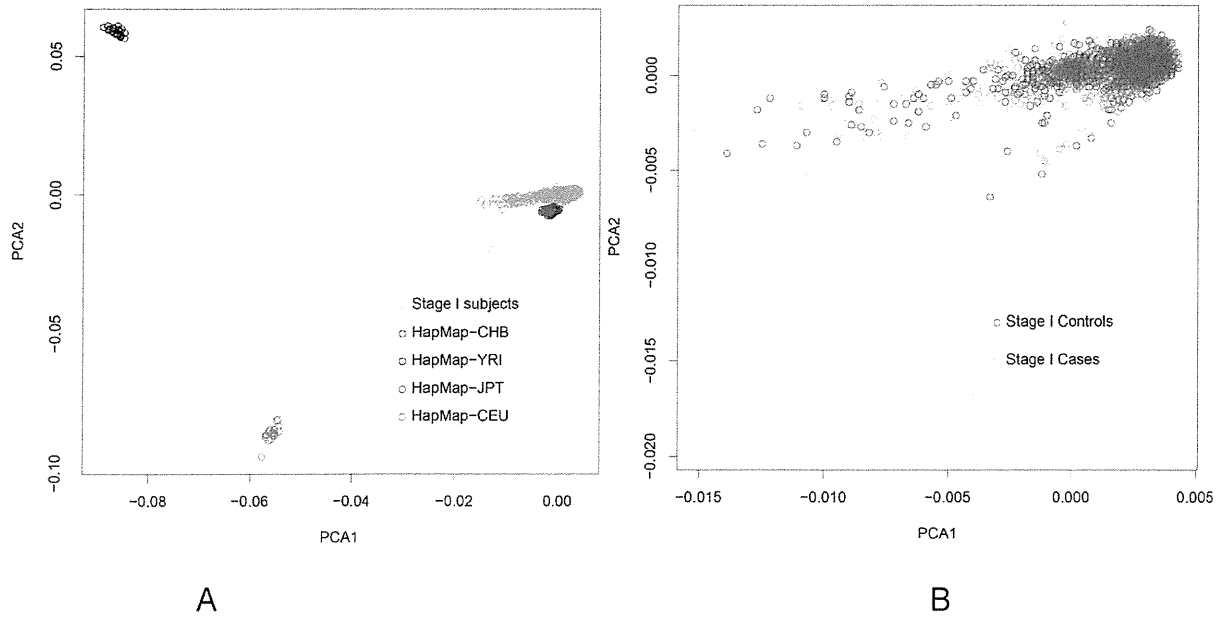

Figure 4. Principal Component Analysis (PCA) based on the first two eigenvectors obtained by PCA. A: all individuals from Stage I and HapMap; B: breast cancer cases and controls from Stage I.
doi:10.1371/journal.pgen.1002532.g004

summarizes the magnitude and direction of the effect relative to the reference allele. An overall z-statistic and p value were then calculated from the weighted average of the individual statistics. Calculations were implemented in the METAL package (http://www.sph.umich.edu/csg/abecasis/Metal). Individual data were obtained from each study for Stage IV SNPs for a pooled analysis, which were conducted using SAS, version 9.2, with the use of two-tailed tests.

We first investigated the population structure by estimating inflation factor λ using all 690,947 SNPs SNPs that passed the QC. The inflation factor λ was estimated to be 1.042, suggesting that any population substructure, if present, should not have any appreciable effect on the results. Among the final 690,947 SNPs obtained in Stage I after QC, we generated a list of 196,471 SNPs with pairwise LD<0.2 by using plink (http://pngu.mgh.harvard.edu/~purcell/plink/). Then, principal components were estimated based on these 196,471 SNPs using EIGENSTRAT [36]. We then drew a plot for all Stage I and HapMap II subjects based on the first two principal components (Figure 4). All study participants in Stage I were clustered very closely with HapMap Asians. The first 5 or 10 principal components were adjusted in the logistic regression analyses for evaluating associations of SNPs and breast cancer risk.

To evaluate the combined effect of SNPs located in chromosome 6q25.1 on breast cancer risk, we created a genetic risk score (GRS) by summing the number (0–2) of risk alleles that each woman carried for each of the three SNPs, including rs9383951, rs9485372, rs2046210. The GRS was constructed among those who had complete data for all three SNPs. We also did imputation using MACH (http://www.sph.umich.edu/csg/abecasis/MACH/index.html) with HapMap II Asian data as reference. LD structure was estimated from the flanking 100 kb of these three SNPs and the *ESR1* gene using data from HapMap II Asians (Figure S1). All SNPs in the LD blocks including rs9485372, rs2046210 and rs9383951 and SNPs inside the *ESR1* gene were analyzed in relation to breast cancer risk with age, rs9485372, rs9383951 and rs2046210 adjusted.

Supporting Information

Figure S1 Estimates of pairwise LD (r^2) for common SNPs from HapMap II Asians for the SNPs located in 6q25.1. A: LD plot for the flanking 100 kb of SNP rs9485372. B: LD plot for the upstream 100 kb of SNP rs2046210 and the ESR1 gene.
(TIF)

Figure S2 Estimates of pairwise LD (r^2) from HapMap II Asian for the SNPs showing significant associations after adjusted for rs9485372, rs9383951 and rs2046210.
(TIF)

Figure S3 A regional plot of the $-\log_{10}$P-values for SNPs at 11q24.3. The LD is estimated using data from HapMap Asian population. Also shown are the SNP Build 36 coordinates in kilobases (Kb), recombination rates in centimorgans (cM) per megabase (Mb) and genes in the region (below) based on the March 2006 UCSC genome browser assembly.
(TIF)

Table S1 Association of SNPs with breast cancer risk by menopause and ER status.
(DOCX)

Table S2 Association results adjusted for the top principal components in Stage I.
(DOCX)

Table S3 LD between the 3 SNPs that are associated with breast cancer and are located in 6q25.1.
(DOCX)

Table S4 Conditional analyses for SNPs located on 6q25.1.
(DOCX)

Table S5 Association results of SNP-SNP interaction.
(DOCX)

Table S6 Associations of breast cancer risk with the genetic risk score for the three SNPs located in chromosome 6q25.1, the Asia Breast Cancer Consortium.
(DOCX)

Table S7 SNPs in 6q25.1 showed association after adjusted for rs9485372, rs9383951 and rs2046210.
(DOCX)

Table S8 Sample size for the SNPs included in Stage IV.
(DOCX)

Text S1 Supplementary Methods.
(DOCX)

Acknowledgments

The authors wish to thank the study participants and research staff for their contributions and commitment to this project, Regina Courtney for DNA preparation, Jing He for data processing and analyses, and Mary Jo Daly for clerical support in the preparation of this manuscript.

Author Contributions

Conceived and designed the overall study: W Zheng. Performed genotyping experiments: J Shi, H Zheng. Wrote the manuscript: J Long, W Zheng, Q Cai, X-O Shu. Significantly contributed to writing the manuscript: C Li, W Wen, RJ Delahanty. Coordinated genotyping assays: Q Cai, J Long. Managed genotyping data: J Long, B Zhang. Performed statistical analyses: J Long, C Li, W Wen. Directed lab operations: Q Cai. Directed the GWAS in Korea: D-H Kang. Assisted the GWAS in Korea: H Sung, J-Y Choi. Contributed to data and biological collection of the original studies: H Shen, J-Y Choi, W Lu, Y-T Gao, H Shen, SK Park, K Chen, C-Y Shen, Z Ren, CA Haiman, K Matsuo, MK Kim, US Khoo, M Iwasaki, Y Zheng, Y-B Xiang, K Gu, N Rothman, W Wang, Z Hu, Y Liu, K-Y Yoo, D-Y Noh, B-G Han, MH Lee, H Zheng, L Zhang, P-E Wu, Y-L Shieh, SY Chan, S Wang, X Xie, S-W Kim, BE Henderson, L Le Marchand, H Ito, Y Kasuga, S-H Ahn, HS Kang, KYK Chan, H Iwata, S Tsugane, D-H Kang, X-O Shu, W Zheng.

References

1. Nathanson KL, Wooster R, Weber BL (2001) Breast cancer genetics: what we know and what we need. Nat Med 7: 552–556.
2. Zhang B, Beeghly-Fadiel A, Long J, Zheng W (2011) Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. Lancet Oncol 12: 477–488.
3. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, et al. (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447: 1087–1093.
4. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al. (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39: 870–874.
5. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, et al. (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 39: 865–869.

6. Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, et al. (2008) Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A 105: 4340–4345.
7. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, et al. (2008) Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 40: 703–706.
8. Zheng W, Long J, Gao YT, Li C, Zheng Y, et al. (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 41: 324–328.
9. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, et al. (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 41: 579–584.
10. Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, et al. (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 41: 585–590.
11. Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, et al. (2010) Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 42: 504–507.
12. Long J, Cai Q, Shu XO, Qu S, Li C, et al. (2010) Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. PLoS Genet 6: e1001002. doi:10.1371/journal.pgen.1001002.
13. Antoniou AC, Wang X, Frederickson ZS, McGuffog L, Tarrell R, et al. (2010) A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet % 19.
14. Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, et al. (2011) Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. J Natl Cancer Inst 103: 425–435.
15. Fletcher O, Houlston RS (2010) Architecture of inherited susceptibility to common cancer. Nat Rev Cancer 10: 353–361.
16. Zheng W, Cai Q, Signorello LB, Long J, Hargreaves MK, et al. (2009) Evaluation of 11 breast cancer susceptibility loci in African-American women. Cancer Epidemiol Biomarkers Prev 18: 2761–2764.
17. Zheng W, Wen W, Gao YT, Shyr Y, Zheng Y, et al. (2010) Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. J Natl Cancer Inst 102: 972–981.
18. Long J, Shu XO, Cai Q, Gao YT, Zheng Y, et al. (2010) Evaluation of breast cancer susceptibility loci in Chinese women. Cancer Epidemiol Biomarkers Prev 19: 2357–2365.
19. Benson JR (2004) Role of transforming growth factor beta in breast carcinogenesis. Lancet Oncol 5: 229–239.
20. Davis RJ (2000) Signal transduction by the JNK group of MAP kinases. Cell 103: 239–252.
21. Hinz M, Stilmann M, Arslan SC, Khanna KK, Dittmar G, et al. (2010) A cytoplasmic ATM-TRAF6-cIAP1 module links nuclear DNA damage signaling to ubiquitin-mediated NF-kappaB activation. Mol Cell 40: 63–74.
22. Dunbier AK, Anderson H, Ghazoui Z, Lopez-Knowles E, Pancholi S, et al. (2011) ESR1 Is Co-Expressed with Closely Adjacent Uncharacterised Genes Spanning a Breast Cancer Susceptibility Locus at 6q25.1. PLoS Genet 7: e1001382. doi:10.1371/journal.pgen.1001382.
23. Stacey SN, Sulem P, Zanon C, Gudjonsson SA, Thorleifsson G, et al. (2010) Ancestry-shift refinement mapping of the C6orf97-ESR1 breast cancer susceptibility locus. PLoS Genet 6: e1001029. doi:10.1371/journal.pgen.1001029.
24. Stevens TA, Meech R (2006) BARX2 and estrogen receptor-alpha (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion. Oncogene 25: 5426–5435.
25. Sellar GC, Li L, Watt KP, Nelkin BD, Rabiasz GJ, et al. (2001) BARX2 induces cadherin 6 expression and is a functional suppressor of ovarian cancer progression. Cancer Res 61: 6977–6981.
26. Gao YT, Shu XO, Dai Q, Potter JD, Brinton LA, et al. (2000) Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. Int J Cancer 87: 295–300.
27. Liang J, Chen P, Hu Z, Zhou X, Chen L, et al. (2008) Genetic variants in fibroblast growth factor receptor 2 (FGFR2) contribute to susceptibility of breast cancer in Chinese women. Carcinogenesis 29: 2341–2346.
28. Zhang L, Gu L, Qian B, Hao X, Zhang W, et al. (2009) Association of genetic polymorphisms of ER-alpha and the estradiol-synthesizing enzyme genes CYP17 and CYP19 with breast cancer risk in Chinese women. Breast Cancer Res Treat 114: 327–338.
29. Ding SL, Yu JC, Chen ST, Hsu GC, Kuo SJ, et al. (2009) Genetic variants of BLM interact with RAD51 to increase breast cancer susceptibility. Carcinogenesis 30: 43–49.
30. 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.
31. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, et al. (2009) A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. Nat Genet 41: 527–534.
32. Han SA, Park SK, Hyun AS, Hyuk LM, Noh DY, et al. (2011) The Korean Hereditary Breast Cancer (KOHBRA) Study: Protocols and Interim Report. Clin Oncol (R Coll Radiol).
33. Kolonel LN, Henderson BE, Hankin JH, Nomura AM, Wilkens LR, et al. (2000) A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 151: 346–357.
34. Hamajima N, Matsuo K, Saito T, Hirose K, Inoue M, et al. (2001) Gene-environment Interactions and Polymorphism Studies of Cancer Risk in the Hospital-based Epidemiologic Research Program at Aichi Cancer Center II (HERPACC-II). Asian Pac J Cancer Prev 2: 99–107.
35. Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, et al. (2009) Serum organochlorines and breast cancer risk in Japanese women: a case-control study. Cancer Causes Control 20: 567–580.
36. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904–909.

GASTROENTEROLOGY 2012;142:1468-1475

CLINICAL—LIVER

Consumption of n-3 Fatty Acids and Fish Reduces Risk of Hepatocellular Carcinoma

NORIE SAWADA,* MANAMI INOUE,* MOTOKI IWASAKI,* SHIZUKA SASAZUKI,* TAICHI SHIMAZU,* TAIKI YAMAJI,* RIBEKA TAKACHI,‡ YASUHITO TANAKA,§ MASASHI MIZOKAMI,‖ SHOICHIRO TSUGANE,* and the Japan Public Health Center–Based Prospective Study Group

Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tsukiji, Chuo-ku Tokyo, Japan; ‡Department of Community Preventive Medicine, Division of Social and Environmental Medicine, Niigata University Graduate School of Medical and Dental Sciences, Asahimachi-dori, Chuo-ku, Niigata, Japan; §Department of Virology and Liver Unit, Nagoya City University Graduate School of Medical Sciences, Kawasumi, Mizuho, Nagoya, Japan; ‖The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Japan

See Covering the Cover synopsis on page 1399; see editorial on page 1411.

BACKGROUND & AIMS: Fish is a rich source of n-3 polyunsaturated fatty acids (PUFAs), such as eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), and docosahexaenoic acid (DHA). Although consumption of fish and n-3 PUFA has been reported to protect against the development of some types of cancer, little is known about its association with hepatocellular carcinoma (HCC). **METHODS:** We investigated the association between fish and n-3 PUFA consumption and HCC incidence (n = 398) in a population-based prospective cohort study of 90,296 Japanese subjects (aged, 45–74 y). Hazard ratios and 95% confidence intervals (CIs) for the highest vs the lowest quintile were estimated from multivariable adjusted Cox proportional hazards regression models. We also conducted subanalyses of subjects with known hepatitis B virus (HBV) or hepatitis C virus (HCV) status, and of subjects who were anti-HCV and/or hepatitis B surface antigen positive. All tests of statistical significance were 2-sided. **RESULTS:** Among all subjects, consumption of n-3 PUFA-rich fish and individual n-3 PUFAs was associated inversely with HCC, in a dose-dependent manner. Hazard ratios for the highest vs lowest quintiles were 0.64 (95% CI, 0.42–0.96) for n-3 PUFA-rich fish, 0.56 (95% CI, 0.36–0.85) for EPA, 0.64 (95% CI, 0.41–0.98) for DPA, and 0.56 (95% CI, 0.35–0.87) for DHA. These inverse associations were similar irrespective of HCV or HBV status. **CONCLUSIONS: Consumption of n-3 PUFA-rich fish or n-3 PUFAs, particularly EPA, DPA, and DHA, appears to protect against the development of HCC, even among subjects with HBV and/or HCV infection.**

Keywords: Diet; Liver Cancer; Cancer Prevention; Omega-3 Fatty Acid.

The most important risk factor in the development of hepatocellular carcinoma (HCC) in human beings is chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV).[1] The markedly poor prognosis of HCC, with a 5-year survival rate in Japan of less than 20%,[2] emphasizes the need for effective preventive measures, particularly in hepatitis virus carriers. Although dietary factors also might be risk factors, the role of diet in the etiology of HCC remains unclear, except with regard to alcohol consumption and aflatoxin contamination.[3]

A recent prospective study showed an inverse association between white meat, including fish, and liver cancer.[4] Inverse associations with the consumption of white meat or fish were observed in some studies,[5–8] but were not confirmed in others.[9–11] Moreover, except for 2 case-control studies,[5,7] most previous epidemiologic studies of white meat or fish and HCC did not consider HCV or HBV infection status.

Fish is a rich source of n-3 polyunsaturated fatty acids (PUFAs), such as eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA), and docosahexaenoic acid (DHA), and several studies have documented a protective effect of dietary n-3 PUFA on the development of several cancers.[12,13] However, less is known about the influence of n-3 PUFA on HCC.

Here, we investigated the presence of an association between fish and n-3 PUFA consumption and HCC in a large-scale, population-based, cohort study in Japan, with consideration for HCV and HBV infection status.

Materials and Methods

Study Population

The Japan Public Health Center–based prospective study was launched in 1990. The study design has been described in detail previously.[14] The study population was defined as all residents of 11 public health center (PHC) areas across Japan who were aged 40–69 years at the start of the respective baseline survey (n = 140,420). In the present analysis, we excluded one PHC area (Tokyo) because data on cancer incidence were not available, as well as some subjects from a second PHC

Abbreviations used in this paper: ALA, alpha-linolenic acid; ALT, alanine aminotransferase; CI, confidence interval; DHA, docosahexaenoic acid; DPA, docosapentaenoic acid; EPA, eicosapentaenoic acid; FFQ, food frequency questionnaire; HBsAg, hepatitis B virus antigen; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; HR, hazard ratio; PHC, public health center; PUFAs, polyunsaturated fatty acids.

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2012.02.018

(Osaka) area for whom different definitions were used (n = 16,841). The study was approved by the Institutional Review Board of the National Cancer Center (Tokyo, Japan).

Baseline Survey

Cohort participants responded to a self-administered questionnaire at baseline in 1990 (cohort I) and 1993–1994 (cohort II). A 5-year follow-up survey was conducted in 1995 (cohort I) and 1998 (cohort II). The 5-year follow-up survey included more comprehensive information on food intake frequency than the baseline survey, and accordingly was used as baseline for the present study. We initially identified 113,378 participants as the study population at the baseline survey. The questionnaire also included information on personal medical history, smoking and drinking habits, diet, and other lifestyle factors. After exclusion of 205 participants who were found to be ineligible because of non-Japanese nationality (n = 44), late report of emigration that occurred before the start of the follow-up period (n = 155), incorrect birth date (n = 3), and duplicate registration (n = 3), the remaining 113,171 participants were considered eligible for the present study. Completed questionnaires were received from 94,999 subjects (response rate, 84%). Further, subjects who had been diagnosed with cancer before the starting point were excluded from analysis (n = 3022).

Food Frequency Questionnaire

The food frequency questionnaire (FFQ) asked subjects about their usual intake of 138 food items in standard portions/units during the previous year, including 19 fish questions. The questionnaire contained 9 frequency categories (never, 1–3 times/mo, 1–2 times/wk, 3–4 times/wk, 5–6 times/wk, once/d, 2–3 times/d, 4–6 times/d, and ≥7 times/d). Nineteen items inquired about fish and shellfish intake, including salted fish, dried fish, canned tuna, salmon or trout, bonito or tuna, cod or flat fish, seabream, horse mackerel or sardine, mackerel pike or mackerel, dried small fish, salted roe, eel, squid, octopus, prawn, short-necked clam or crab shell, vivipara, *chikuwa* (fish paste product), and *kamaboko* (fish paste product). Standard portion sizes were specified for each food item in the 3 amount choices of small (50% smaller), medium (same as standard), and large (50% larger). Fish consumption in g/day was calculated by multiplying frequency by standard portion size for each food item. In our FFQ, dishes in which food was just a constituent were not included. We calculated the daily intake of all n-3 PUFAs combined and of individual PUFAs, namely α-linolenic acid (ALA), EPA, DPA, and DHA, using a fatty acid composition table of Japanese foods.[15] Furthermore, based on the value of n-3 PUFA per 100 g edible portion of fish, we also calculated the consumption of n-3 PUFA-rich fish (salmon or trout, sea bream, horse mackerel or sardine, mackerel pike or mackerel, and eel).[15] Intake of food and nutrients was log-transformed and adjusted for total energy intake by the residual model.[16] We also used the nutrient density method and obtained similar results.

We documented the validity of the FFQ in the assessment of fish, ALA, EPA, DPA, and DHA consumption in subsamples using 14- or 28-day dietary records. Based on 102 men and 113 women in cohort I, the Spearman correlation coefficients between energy-adjusted intake of fish,[17] n-3 PUFA, ALA, EPA, DPA, and DHA[18] from the questionnaire and from dietary records were 0.37, 0.21, 0.27, 0.38, 0.32, and 0.34 for men, and 0.32, 0.34, 0.25, 0.45, 0.39, and 0.37 for women, respectively. The percentage differences between the dietary records and the FFQ for fish were −16% for men, and −1% for women.[19] Thus, validities for fish and n-3 PUFAs were considered moderate.

Among the 91,977 subjects who responded to the questionnaire and had no past history of cancer, subjects who reported extreme total energy intake (upper or lower 1.0%) were excluded, leaving 90,296 subjects for analysis.

Blood Collection and Laboratory Assays

Subjects were asked to voluntarily provided 10 mL of blood during health checkups in 1993–1995, at which time plasma alanine aminotransferase (ALT) level was determined. Samples were divided into the plasma and buffy layers, and preserved at −80°C until analysis. Among subjects who provided blood (n = 33,329), plasma samples from a portion of the subjects (n = 17,497) were screened for anti-HCV using a third-generation immunoassay (Lumipulse II Ortho HCV; Ortho-Clinical Diagnostics K.K., Tokyo, Japan)[20] and for hepatitis B virus antigen (HBsAg) by reversed passive hemagglutination with a commercial kit (Institute of Immunology Co, Ltd, Tokyo, Japan).

Follow-up and Identification of Hepatocellular Carcinoma

Subjects were followed up from the baseline survey until December 31, 2008. Changes in residence status, including survival, were identified annually through the residential registry in the respective public health center area. Among study subjects, 2775 (3.1%) moved out of their study area and 318 (0.4%) were lost to follow-up evaluation during the study period.

Incidence data on HCC were identified by active patient notification from major local hospitals in the study area and data linkage with population-based cancer registries, with permission from the local governments responsible for the registries. Death certificates were used as a supplementary information source, with 10.6% of cases in our cancer registry system based on death certificate only. Cases were coded using the International Classification of Diseases for Oncology, 3rd ed (code C22.0).[21] During an average follow-up period of 11.2 years (1,008,595 person-years), a total of 398 cases of HCC were newly diagnosed among 90,296 subjects who had returned the baseline questionnaire. In one subgroup, a total of 74 cases of HCC were newly diagnosed among 17,497 subjects who had data on anti-HCV and HBsAg status and ALT level.

Statistical Analysis

Person-years of follow-up evaluation were calculated for each subject from the date of completion of the baseline questionnaire to the date of HCC diagnosis, date of emigration from the study area, or date of death, whichever occurred first; or if none of these occurred, follow-up evaluation was through the end of the study period (December 31, 2008). Subjects who were lost to follow-up evaluation were censored at the last confirmed date of presence in the study area. Hazard ratios (HRs) of HCC were calculated by quintiles of consumption of the respective food items or nutrients, with the lowest consumption category as the reference. HRs and 95% confidence intervals (CIs) were calculated by the Cox proportional hazards model, and adjusted for age at baseline survey (continuous), sex, and study area (10 PHC areas) according to the SAS PHREG procedure (version 9.1; SAS Institute, Inc, Cary, NC). For further adjustment, additional possible confounders were incorporated into the model, namely smoking status (never, former, current); alcohol intake (almost never, 1–3 times/mo, ≥1 times/wk); body mass index (continuous); past history of diabetes mellitus (yes or no); intake of coffee (almost never, 1–4 d/wk, ≥1 cups/d); and soy foods, vegetables, vegetable oil, protein, and iron (continuous). Because of a high correlation coefficient between