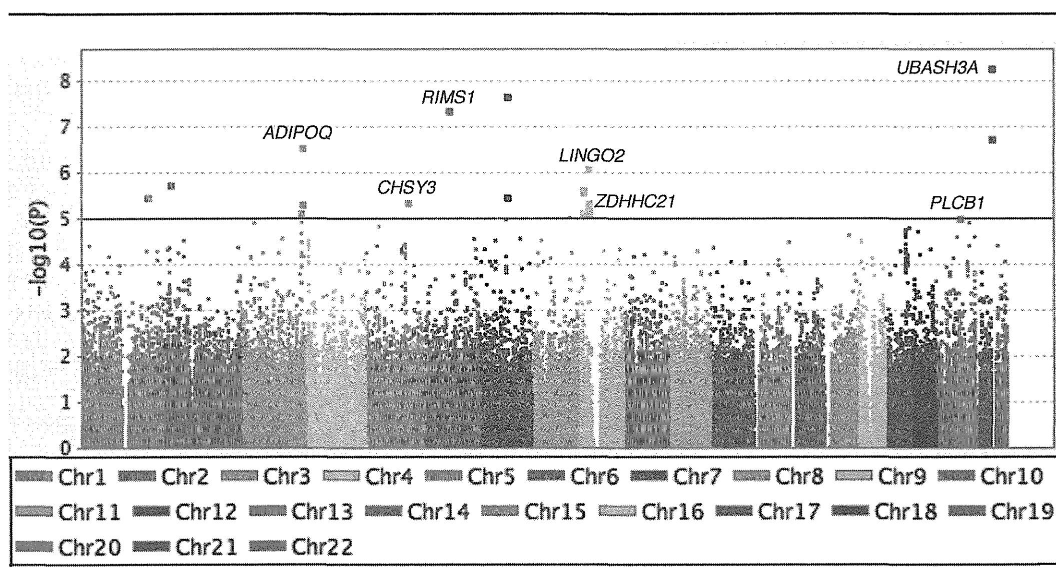

FIGURE 1 Genome-wide $-\log_{10}$ *P* value plot for stage 1 discovery analysis of confectionery-intake score. Blue line indicates the criteria for stage 2 replication phase ($P = 1 \times 10^{-5}$). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

from the expected *P* value under the null hypothesis (Figure 2). The genomic inflation factor (λ) was 1.003.

Of the 22 SNPs, five were excluded because they were located within the same LD block of the selected SNPs with the lowest *P* value in each block. An additional five SNPs were omitted as they were not within or near to a gene (<50 kb) and their MAF was <0.05. Eventually, we selected 12 SNPs (Table 2) from HERPACC-II participants for the stage 2 follow-up based on predefined criteria. The genotype distributions of these 12 SNPs were in Hardy–Weinberg equilibrium ($P > 0.05$) in both studies, with the exceptions of rs12351510 in stage 1 ($P = 0.014$) and rs2839519 in stage 2 ($P = 0.020$). The deviation from the expected genotype distribution, however, was relatively small, at less than 1% in both cases.

Among the 12 selected SNPs, the association of rs822396 with the confectionery-intake score was replicated in stage 2 of the J-MICC population ($P = 0.049$; Table 2). In the pooled analysis of stages 1 and 2 data, the smallest *P* was found for rs822396 ($P = 4.2 \times 10^{-5}$), followed by rs17042603, rs13356198, and rs1147522. The polymorphism rs822396 was shown to be an SNP in intron 1 of the *ADIPOQ* gene (IVS1-3971A>G). Further adjustment for BMI did not substantially alter the results. The *P* values for rs822396 in stages 1, 2, and 1+2 were 4.2×10^{-7}, 0.049, and 4.3×10^{-5}, respectively. The *P* values for other selected SNPs were also similar to those without BMI adjustment (data not shown), although those in stage 1 for rs10810211 and rs6039211 were 1.1×10^{-5}.

The association of the rs822396 polymorphism with the confectionery-intake score was more dominant for Japanese-style than Western-style confectionery: the respective *P* values in stages 1, 2, and 1+2 were 1.9×10^{-8}, 0.013, and 4.8×10^{-6} for Japanese-style confectionery, and 5.1×10^{-3}, 0.73, and 0.083 for Western-style confectionary.

We compared the mean confectionery-intake score by the rs822396 genotype (major homozygotes versus heterozygotes + minor homozygotes) and the background characteristics of participants in the pooled dataset (Table 3). The score was higher among participants with at least one minor allele than among those without. Moreover,

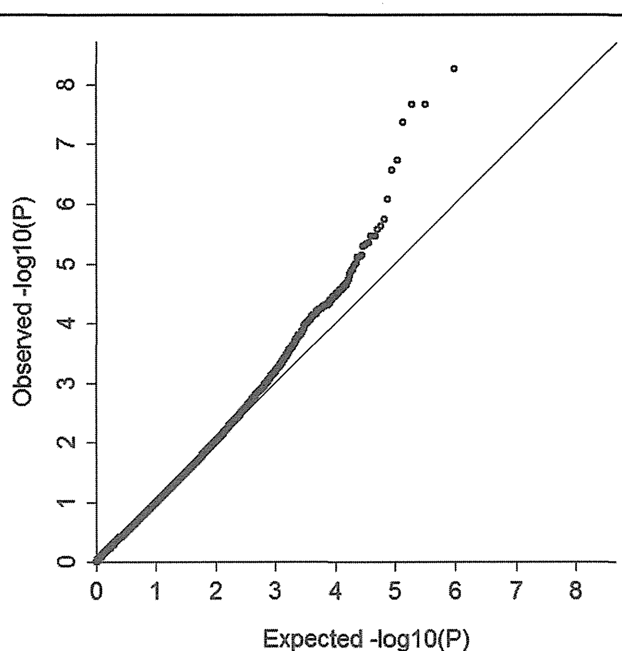

FIGURE 2 Q–Q plot for stage 1 discovery analysis of confectionery intake score. $\lambda = 1.003$.

TABLE 2 SNPs identified in GWAS analysis for confectionery intake score (n = 939 for stage 1 and 4,491 for stage 2)

SNP ID	Chromosome	Position (NCBI Build 36.3)	Major allele[a]	Minor allele[a]	Nearby gene	Minor allele frequency			β for additive model			*P* for additive model		
						Stage 1	Stage 2	Stage 1+2	Stage 1	Stage 2	Stage 1+2	Stage 1	Stage 2	Stage 1+2
rs17042603	2	21651923	A	G	No gene	0.052	0.050	0.050	0.1568	0.0155	0.0410	1.8×10^{-6}	0.23	7.0×10^{-4}
rs822396	3	188049571	A	G	*ADIPOQ*	0.055	0.060	0.060	0.1643	0.0234	0.0461	2.8×10^{-7}	0.049	4.2×10^{-5}
rs13356198	5	129284991	A	G	*CHSY3*	0.055	0.064	0.062	0.1432	0.0172	0.0362	4.7×10^{-6}	0.14	9.1×10^{-4}
rs1147522	6	72640321	C	T	*RIMS1*	0.017	0.017	0.017	0.3017	0.0043	0.0580	4.3×10^{-8}	0.84	4.5×10^{-3}
rs3897749	7	85972133	G	T	No gene	0.065	0.060	0.061	0.1379	−0.0104	0.0168	3.4×10^{-6}	0.38	0.13
rs6474850	9	14652459	G	A	*ZDHHC21*	0.182	0.167	0.170	0.0911	−0.0044	0.0129	2.4×10^{-6}	0.56	0.069
rs2890992	9	14662267	T	C	*ZDHHC21*	0.141	0.122	0.126	0.1001	−0.0006	0.0177	2.7×10^{-6}	0.95	0.026
rs10810211	9	14666870	A	G	*ZDHHC21*	0.128	0.118	0.120	0.0976	−0.0010	0.0168	7.8×10^{-6}	0.90	0.037
rs12351510	9	28641898	T	G	*LINGO2*	0.014	0.015	0.015	0.2834	−0.0039	0.0494	8.5×10^{-7}	0.86	0.022
rs6039211	20	8564588	A	G	*PLCB1*	0.479	0.535	0.525	0.0652	−0.0035	0.0072	9.97×10^{-6}	0.53	0.18
rs2839519	21	42740913	G	A	*UBASH3A*	0.177	0.170	0.171	0.0998	−0.0118	0.0077	1.9×10^{-7}	0.11	0.27
rs2839525	21	42750006	T	G	*UBASH3A*	0.181	0.174	0.175	0.1103	−0.0084	0.0123	5.5×10^{-9}	0.25	0.075

[a]Alleles are indexed to the forward strand of Center for Biotechnology Information (NCBI) Build 36.3.

this difference was consistently observed across the strata of gender, age, smoking and drinking habits, and BMI. It was particularly large (0.083, $P = 5.1 \times 10^{-6}$) in younger participants aged <55 years.

Discussion

In this study, we identified 12 candidate SNPs that were potentially related to confectionery intake in a GWAS. Among them, the association was replicated in an independent population for one SNP (rs822396 or IVS1-3971A>G) in intron 1 of the *ADIPOQ* gene.

The *ADIPOQ* gene on chromosome 3q27 encodes adiponectin, which is an adipokine that is extensively expressed in adipose tissue, and is a highly abundant plasma protein with circulating levels that are, in part, genetically controlled (27,28). Adiponectin has excited intense interest because of the robust negative correlations of its circulating levels with indices of insulin resistance and the risk of type 2 diabetes, as well as their consistent inverse associations with fat mass (28). Because central nervous insulin action might be related to inhibition of eating behavior, and might be negatively correlated with peripheral insulin resistance and obesity (29), insulin resistance associated with hypoadiponectinemia could be involved in increases in food intake.

In addition to its potential role as an insulin-sensitizing adipokine, it is hypothesized that adiponectin plays an important role in the regulation of energy homeostasis, including appetite stimulation (27,30), although its effects on food intake show considerable diversity across studies (30). Collectively, however, previous studies show that adiponectin directly and/or indirectly affects eating activities, which might partly explain our current findings.

Physiologically, adiponectin has been shown to stimulate food intake by activating AMP-activated protein kinase (AMPK) in the arcuate hypothalamus via its receptor AdipoR1 (27,31). The putative downstream pathways for food-intake regulation in response to hypothalamic AMPK are acetyl-coenzyme A carboxylase/malonyl-coenzyme A/carnitine palmitoyltransferase-1/fatty-acid oxidation and mammalian target of rapamycin signaling (31). However, it remains to be investigated whether adiponectin specifically affects consumption of energy-dense foods such as confectionery. Our relatively broad screening approach with a cutoff of $P < 1 \times 10^{-5}$ in stage 1 might have resulted in the observed low replication rate of only one of the 12 candidate SNPs identified in stage 1. Nevertheless, this replicated SNP appears to be biologically plausible as an indicator of a gene involved in eating behaviors.

The *ADIPOQ* gene spans 17 kb, contains three exons, and its translation starts at exon 2 and ends at exon 3 (32,33). SNPs throughout the gene or nearby have recently been related to the circulating levels of adiponectin in GWASs (34-36) or in studies genotyping tag SNPs (33,37). SNPs representing the most significant associations, however, vary considerably among studies. They are distributed throughout or nearby the *ADIPOQ* gene from upstream (e.g., rs864265), through the promoter region (e.g., rs822387, rs17300539), intron 1 (e.g., rs16861210, rs17366568), exon 2 (e.g., rs2241766), and intron 2 (e.g., rs3774261), to the 3′ untranslated region (UTR; e.g., rs6773957, rs2082940). These SNPs are frequently in LD with one another, so that researchers cannot easily focus on the genetic polymorphisms that are responsible for the adiponectin levels.

The rs822396 polymorphism associated with the confectionery-intake score in the present study was also in LD, albeit weak-to-

TABLE 3 Mean confectionery intake score by rs822396 genotype and background characteristics of participants in the pooled dataset (stages 1 and 2 studies)

	Major homozygotes			Heterozygotes + minor homozygotes			
Characteristics	***n***	**Mean**	**(SE)[a]**	***n***	**Mean**	**(SE)[a]**	***P***
All participants	4,800	0.278	(0.004)	628	0.326	(0.011)	4.6×10^{-5}
Gender							
Male	2,288	0.221	(0.005)	302	0.275	(0.014)	3.4×10^{-4}
Female	2,512	0.331	(0.006)	326	0.373	(0.017)	0.018
Age (years)							
<55	2,085	0.271	(0.006)	254	0.354	(0.017)	5.1×10^{-6}
≥55	2,715	0.284	(0.005)	374	0.308	(0.014)	0.12
Smoking							
Never smokers	2,784	0.318	(0.005)	380	0.362	(0.015)	4.7×10^{-3}
Current or ex-smokers	2,009	0.222	(0.006)	248	0.273	(0.017)	3.6×10^{-3}
Alcohol drinking							
Never drinkers	1,985	0.325	(0.007)	266	0.366	(0.018)	0.034
Current or ex-drinkers	2,810	0.245	(0.005)	361	0.296	(0.014)	5.9×10^{-4}
Body mass index (kg/m^2)							
<25	3,638	0.278	(0.004)	494	0.330	(0.012)	5.8×10^{-5}
≥25	1,118	0.279	(0.009)	128	0.310	(0.026)	0.26

[a]Adjusted for gender and age.

moderate, with some of the previously mentioned polymorphisms, including rs864265, rs822387, rs3774261, rs6773957, and rs2082940 in a Japanese population within the International Haplotype Map (HapMap) project (http://hapmap.ncbi.nlm.nih.gov/). Therefore, even if the rs822396 polymorphism is not directly linked with circulating adiponectin levels, genetic polymorphisms of *ADIPOQ* around the rs822396 SNP might control blood adiponectin levels, and could be associated with a propensity to favor foods of high-energy density such as confectionery. Additionally, alternative splicing sites of the *ADIPOQ* gene have been found near this SNP (within 4 kb upstream and downstream of the SNP; http://www.ensembl.org/Homo_sapiens/Gene/Splice?db=core;g=ENSG00000181092;r=3:186560479-186576252;t=ENST00000444204). Thus, the rs822396 SNP might affect the expression of *ADIPOQ* through alternative splicing.

Interestingly, the Québec Family Study by Choquette et al. (38) involving genome-wide linkage analysis found linkage on chromosome 3q27.3 with intakes of energy, lipid, and carbohydrate. As the 3q27 region harbors the *ADIPOQ* gene, this study might corroborate these earlier findings, suggesting that a variation of this gene is associated with the intake of high-calorie foods.

The association with the confectionery-intake score showed genome-wide significance for SNP rs2839525 in the stage 1 study (Table 2, $P = 5.5 \times 10^{-9}$). This was, however, not replicated in stage 2. Although we could not identify the precise reason for this discrepancy, highly significant associations found in GWASs have often failed to be replicated (39).

Although we identified and replicated an association of the rs822396 polymorphism with the confectionery-intake score, it did not reach genome-wide significance ($P < 10^{-8}$) either in stage 1 alone or in the pooled analysis of stages 1 and 2, and the association was comparatively weaker in stage 2. This might have been partly due to the relatively small number of participants ($n = 939$) in stage 1, or the simplistic self-reporting method used to assess the intake frequencies of Japanese-style and Western-style confectionery. Moreover, the difference in background characteristics of participants between the studies might have partly accounted for the weaker association in stage 2 than stage 1. The two populations differed notably in age: the average age for the stage 2 group was 8 years higher than that for stage 1 (Table 1). The association of the rs822396 polymorphism with the confectionery-intake score was much stronger in the younger group (<55 years; Table 3). When we analyzed data only from the stage 2 study by age stratum (<55 and ≥55 years), as in Table 2, the association of SNP rs822396 was more dominant in the younger age strata ($n = 1,772$; β for additive model = 0.0487; $P = 0.008$) than in the older one ($n = 2,717$; $\beta = 0.0090$; $P = 0.56$). The older age distribution of the population in stage 2 might therefore have attenuated the SNP association compared with that in stage 1; the association might have been more replicable if the stage 2 population had been more similar in age distribution to that in stage 1.

Although a more detailed questionnaire including questions on portion sizes might have provided more conclusive findings, informative data were obtained in a previous familial study based on simple questions about the intake frequencies of sweet foods (9). Furthermore, the association of an SNP in the *ADIPOQ* gene with the intake of high-energy foods such as confectionery is biologically plausible and supports the findings of a previous analysis (38).

In summary, we found that an SNP in the *ADIPOQ* gene was correlated with a preference for confectionery through a two-stage GWAS with discovery and replication phases. Given the biological plausibility and

relevant previous findings, this association warrants further follow-up and provides a good working hypothesis for experimental testing.O

Acknowledgments

We thank Miki Kokubo at the Center for Genomic Medicine of Kyoto University Graduate School of Medicine for technical assistance with the stage 1 GWAS, Kyota Ashikawa, Tomomi Aoi, and other members of the Laboratory for Genotyping Development at the Center for Genomic Medicine of RIKEN for support with genotyping in the stage 2 study, Yoko Mitsuda, Keiko Shibata, and Etsuko Kimura at the Department of Preventive Medicine of Nagoya University Graduate School of Medicine, Miki Watanabe and Isao Oze at the Division of Epidemiology and Prevention of the Aichi Cancer Center Research Institute, Fusako Katsurada at the Department of Health Science of Shiga University of Medical Science, and Mitsuhiko Matsushita and Yasunobu Sagara at the Tokushima Prefecture Health Examination Center for their cooperation, technical assistance, and valuable comments. We also thank Shinkan Tokudome at the National Institute of Health and Nutrition (formerly Nagoya City University), Chiho Goto at Nagoya Bunri University, Nahomi Imaeda at Nagoya Women's University, Yuko Tokudome at Nagoya University of Arts and Sciences, Masato Ikeda at the University of Occupational and Environmental Health, and Shinzo Maki at the Aichi Prefectural Dietetic Association for providing the food frequency questionnaire.

The discovery phase (stage 1) of this study was supported by a Grant for the CREST program of the Japan Science and Technology Agency, and a Grant for the Third Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor and Welfare of Japan (H21-3rdCancer-G003). The replication phase (stage 2) was supported by Grants-in-Aid for Scientific Research on Priority Areas (No. 17015018) and Innovative Areas (No. 221S0001) from the Japanese Ministry of Education, Culture, Sports, Science, and Technology.

© 2013 The Obesity Society

References

1. Kant AK, Graubard BI. Energy density of diets reported by American adults: association with food group intake, nutrient intake, and body weight. *Int J Obes (Lond)* 2005;29:950-956.
2. Mendoza JA, Drewnowski A, Christakis DA. Dietary energy density is associated with obesity and the metabolic syndrome in US adults. *Diabetes Care* 2007;30: 974-979.
3. Eny KM, Wolever TM, Corey PN, El-Sohemy A. Genetic variation in *TAS1R2* (Ile191Val) is associated with consumption of sugars in overweight and obese individuals in 2 distinct populations. *Am J Clin Nutr* 2010;92:1501-1510.
4. Agurs-Collins T, Fuemmeler BF. Dopamine polymorphisms and depressive symptoms predict foods intake. Results from a nationally representative sample. *Appetite* 2011;57:339-348.
5. Mizuta E, Kokubo Y, Yamanaka I, et al. Leptin gene and leptin receptor gene polymorphisms are associated with sweet preference and obesity. *Hypertens Res* 2008; 31:1069-1077.
6. Avena NM, Rada P, Hoebel BG. Sugar and fat bingeing have notable differences in addictive-like behavior. *J Nutr* 2009;139:623-628.
7. Pelchat ML. Food addiction in humans. *J Nutr* 2009;139:620-622.
8. Collaku A, Rankinen T, Rice T, et al. A genome-wide linkage scan for dietary energy and nutrient intakes: the Health, Risk Factors, Exercise Training, and Genetics (HERITAGE) Family Study. *Am J Clin Nutr* 2004;79:881-886.
9. Keskitalo K, Knaapila A, Kallela M, et al. Sweet taste preferences are partly genetically determined: identification of a trait locus on chromosome 16. *Am J Clin Nutr* 2007;86:55-63.
10. Hebebrand J, Volckmar AL, Knoll N, Hinney A. Chipping away the 'missing heritability': GIANT steps forward in the molecular elucidation of obesity – but still lots to go. *Obes Facts* 2010;3:294-303.
11. Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann N Y Acad Sci* 2010;1212:59-77.
12. Tajima K, Hirose K, Inoue M, Takezaki T, Hamajima N, Kuroishi T. A model of practical cancer prevention for out-patients visiting a hospital: the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC). *Asian Pac J Cancer Prev* 2000;1:35-47.
13. Hamajima N, Matsuo K, Saito T, et al. 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* 2001;2:99-107.
14. Inoue M, Tajima K, Hirose K, et al. Epidemiological features of first-visit outpatients in Japan: comparison with general population and variation by sex, age, and season. *J Clin Epidemiol* 1997;50:69-77.
15. Wakai K, Hamajima N, Okada R, et al. Profile of participants and genotype distributions of 108 polymorphisms in a cross-sectional study of associations of genotypes with lifestyle and clinical factors: a project in the Japan Multi-Institutional Collaborative Cohort (J-MICC) Study. *J Epidemiol* 2011;21:223-235.
16. Hamajima N. The Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study) to detect gene-environment interactions for cancer. *Asian Pac J Cancer Prev* 2007;8:317-323.
17. Tokudome S, Goto C, Imaeda N, Tokudome Y, Ikeda M, Maki S. Development of a data-based short food frequency questionnaire for assessing nutrient intake by middle-aged Japanese. *Asian Pac J Cancer Prev* 2004;5:40-43.
18. Tokudome Y, Goto C, Imaeda N, et al. Relative validity of a short food frequency questionnaire for assessing nutrient intake versus three-day weighed diet records in middle-aged Japanese. *J Epidemiol* 2005;15:135-145.
19. Goto C, Tokudome Y, Imaeda N, et al. Validation study of fatty acid consumption assessed with a short food frequency questionnaire against plasma concentration in middle-aged Japanese people. *Scand J Nutr* 2006;50:77-82.
20. Imaeda N, Goto C, Tokudome Y, Hirose K, Tajima K, Tokudome S. Reproducibility of a short food frequency questionnaire for Japanese general population. *J Epidemiol* 2007;17:100-107.
21. Lee KY, Uchida K, Shirota T, Kono S. Validity of a self-administered food frequency questionnaire against 7-day dietary records in four seasons. *J Nutr Sci Vitaminol (Tokyo)* 2002;48:467-476.
22. Date C, Fukui M, Yamamoto A, et al. Reproducibility and validity of a self-administered food frequency questionnaire used in the JACC study. *J Epidemiol* 2005; 15(Suppl 1):S9-S23.
23. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81: 559-575.
24. Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y. A high-throughput SNP typing system for genome-wide association studies. *J Hum Genet* 2001;46:471-477.
25. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-265.
26. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005;76:887-893.
27. Kadowaki T, Yamauchi T, Kubota N. The physiological and pathophysiological role of adiponectin and adiponectin receptors in the peripheral tissues and CNS. *FEBS Lett* 2008;582:74-80.
28. Cook JR, Semple RK. Hypoadiponectinemia – cause or consequence of human "insulin resistance"? *J Clin Endocrinol Metab* 2010;95:1544-1554.
29. Schloegl H, Percik R, Horstmann A, Villringer A, Stumvoll M. Peptide hormones regulating appetite – -focus on neuroimaging studies in humans. *Diabetes Metab Res Rev* 2011;27:104-112.
30. Stofkova A. Leptin and adiponectin: from energy and metabolic dysbalance to inflammation and autoimmunity. *Endocr Regul* 2009;43:157-168.
31. Minokoshi Y, Shiuchi T, Lee S, Suzuki A, Okamoto S. Role of hypothalamic AMP-kinase in food intake regulation. *Nutrition* 2008;24:786-790.
32. Takahashi M, Arita Y, Yamagata K, et al. Genomic structure and mutations in adipose-specific gene, adiponectin. *Int J Obes Relat Metab Disord* 2000;24:861-868.
33. Siitonen N, Pulkkinen L, Lindstrom J, et al. Association of ADIPOQ gene variants with body weight, type 2 diabetes and serum adiponectin concentrations: the Finnish Diabetes Prevention Study. *BMC Med Genet* 2011;12:5.
34. Ling H, Waterworth DM, Stirnadel HA, et al. Genome-wide linkage and association analyses to identify genes influencing adiponectin levels: the GEMS Study. *Obesity (Silver Spring)* 2009;17:737-744.
35. Heid IM, Henneman P, Hicks A, et al. Clear detection of *ADIPOQ* locus as the major gene for plasma adiponectin: results of genome-wide association analyses including 4659 European individuals. *Atherosclerosis* 2010;208:412-420.
36. Wu Y, Li Y, Lange EM, et al. Genome-wide association study for adiponectin levels in Filipino women identifies *CDH13* and a novel uncommon haplotype at *KNG1-ADIPOQ*. *Hum Mol Genet* 2010;19:4955-4964.
37. Hivert MF, Manning AK, McAteer JB, et al. Common variants in the adiponectin gene (*ADIPOQ*) associated with plasma adiponectin levels, type 2 diabetes, and diabetes-related quantitative traits: the Framingham Offspring Study. *Diabetes* 2008; 57:3353-3359.
38. Choquette AC, Lemieux S, Tremblay A, et al. Evidence of a quantitative trait locus for energy and macronutrient intakes on chromosome 3q27.3: the Québec Family Study. *Am J Clin Nutr* 2008;88:1142-1148.
39. Ioannidis JP. Non-replication and inconsistency in the genome-wide association setting. *Hum Hered* 2007;64:203-213.

INTERNATIONAL JOURNAL OF UROLOGY

International Journal of Urology (2013) 20, 911–916 doi: 10.1111/iju.12074

Original Article: Clinical Investigation

Prevalence of postmicturition urinary incontinence in Japanese men: Comparison with other types of incontinence

Koji Yoshimura,[1] Takeo Nakayama,[3] Akihiro Sekine,[4] Fumihiko Matsuda,[2] Shinji Kosugi,[5] Yoshio Sugino,[1] Kenichi Yoshimura,[6] Osamu Ogawa[1] and the Nagahama Cohort Research Group

[1]Department of Urology and [2]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Departments of [3]Health Informatics, [4]Genome Informatics and [5]Medical Ethics, Kyoto University School of Public Health, and [6]Translational Research Center, Kyoto University Hospital, Kyoto, Japan

Abbreviations & Acronyms
BMI = body mass index
ICIQ-SF = International Consultation Society Incontinence Questionnaire Short Form
MUI = mixed urinary incontinence
OR = odds ratio
PUI = postmicturition urinary incontinence
SUI = stress urinary incontinence
UI = urinary incontinence
UUI = urge urinary incontinence

Correspondence: Koji Yoshimura M.D., Ph.D., Department of Urology, Kyoto University Graduate School of Medicine, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 610-8507, Japan. Email: ky7527@kuhp.kyoto-u.ac.jp

Received 25 September 2012; accepted 9 December 2012. Online publication 10 January 2013

Objective: To investigate the prevalence and correlates of postmicturition urinary incontinence in Japanese men, and to compare with those of other types of urinary incontinence.

Methods: A total of 3224 male participants in a community-based survey were investigated. Three types of urinary incontinence were assessed; that is, postmicturition urinary incontinence, stress urinary incontinence and urge urinary incontinence. Age, body mass index, alcohol intake, cigarette smoking, and medical history of 18 diseases and conditions were the dependent variables for candidate correlates of the three types of incontinence.

Results: Unlike stress urinary incontinence and urge urinary incontinence, the prevalence of postmicturition urinary incontinence was constant throughout all generations (6.5% for the 30 s, 6.6% for the 40 s, 6.0% for the 50 s, 6.3% for the 60 s and 5.1% for the 70 s). The independent correlates for postmicturition urinary incontinence were asthma ($P < 0.001$; odds ratio 3.01), prostatic disease ($P < 0.001$; odds ratio 2.38), rhinosinusitis ($P = 0.001$; odds ratio 1.92), low back pain ($P = 0.003$; odds ratio 1.58), sleeplessness ($P = 0.013$; odds ratio 1.86), depression ($P = 0.024$; odds ratio 3.41) and body mass index ($P = 0.025$; odds ratio 0.73).

Conclusions: Postmicturition urinary incontinence has different characteristics from those of stress urinary incontinence and urge urinary incontinence. Unlike stress urinary incontinence and urge urinary incontinence, postmicturition urinary incontinence is not age-dependent. Several diseases related to an allergic status, such as asthma and rhinosinusitis, are correlates for postmicturition urinary incontinence.

Key words: epidemiology, male, urinary incontinence.

Introduction

UI affects substantial proportions of men. The estimated prevalence of UI increases with age, from 4.8% in men aged 19–44 years to 21.1% in men older than 65 years.[1] As most previous epidemiological studies assessing male UI focused on UUI, SUI and MUI, the prevalences and risk factors for these types of UI have been well investigated. These types of UI in men, especially UUI, result in considerable distress and deterioration of the quality of life.[2,3]

Postmicturition dribble is another type of UI, and is classified as a postmicturition symptom. This symptom has been defined by the International Continence Society in 2002 as involuntary loss of urine immediately after the individual has finished passing urine, usually after leaving the toilet in men,[4] and is distinct from terminal dribbling.[5] Several studies have found that the prevalence of postmicturition dribble is relatively high in men,[2,6–9] and the degree of distress caused by this symptom is substantial.[2] However, more

© 2013 The Japanese Urological Association

detailed information about this symptom, such as its age-stratified prevalence and correlates, is scarce.[8] In the present study, we investigated the prevalence and correlates of male postmicturition dribble using data from a Japanese cohort population, in comparison with those of other types of UI. We refer to postmicturition dribble as PUI hereafter, because we want to emphasize an aspect of urinary incontinence with this symptom for comparisons with other types of UI.

Methods

The Nagahama cohort project is a prospective study composed of a questionnaire survey, anthropometric measures, physiological measures, biochemical measures from blood samples and genomic information.[10] The baseline data for the study population were obtained from August 2008 to November 2010. All the protocols and informed consent procedures were approved by the Kyoto University Graduate School and Faculty of Medicine Ethics Committee, the ad hoc Review Board of the Nagahama Cohort Project, and the Nagahama Municipal Review Board of Personal Information Protection.

Although the questionnaire was designed to be self-reporting, the interviewers confirmed the appropriate answer by interview if no answer was written to minimize the lack of data. Self-reported information on the medical history, major comorbidities, current medication use, lifestyle and psychosocial factors were also collected. The three types of UI; that is, SUI, UUI and PUI, were assessed using the ICIQ-SF, the Japanese version of which has been validated for use.[11] When assessing the prevalence of UI, we examined all severities of UI; that is, "about once a week or less often" or more in frequency and "a small amount" or more in volume. When assessing the severity of UI, we classified it into the following three levels: mild, once or less per week; moderate, two or three times per week; severe, once or more per day.

The independent variables used in the analyses for correlates of the three types of incontinence were age; BMI; alcohol intake; cigarette smoking; and 18 medical history diseases and conditions, comprising low back pain, hypertension, rhinosinusitis, dyslipidemia, prostatic disease, diabetes mellitus, gout, sleeplessness, hepatic disease, coronary artery disease, anemia, malignant disease, atopic dermatitis, asthma, depression, reflux esophagitis, renal disease and stroke. Several diseases or conditions included in the questionnaire were omitted from the analyses because of their low prevalence or relation to injury.

Statistics

Statistical analyses were carried out using a logistic regression model. Univariate analyses were first carried out to confirm the basic relationships between each independent variable and the three types of UI. If the univariate analysis produced a P-value of less than 0.25, the variable was applied to a multivariate analysis. All P-values were two-sided, and values of $P < 0.05$ were considered significant. SPSS version 13.0 (SPSS, Chicago, IL, USA) was used for all calculations.

Table 1 Characteristics of the participants

Age (years)	
30–39	673 (20.9%)
40–49	381 (11.8%)
50–59	481 (14.9%)
60–69	1201 (37.3%)
70–	488 (15.1%)
BMI (kg/m^2)	
<18.5	108 (3.3%)
18.5–24.9	2243 (69.6%)
25.0–29.9	776 (24.1%)
30≤	97 (3.0%)
Alcohol (yes)	2617 (81.2%)
Smoking (yes)	991 (30.7%)
Medical conditions	
Low back pain (yes)	1743 (54.1%)
Hypertension (yes)	791 (24.5%)
Rhinosinusitis (yes)	385 (11.9%)
Dyslipidemia (yes)	366 (11.3%)
Prostatic disease (yes)	299 (9.3%)
Diabetes (yes)	295 (9.2%)
Gout (yes)	237 (7.4%)
Sleeplessness (yes)	211 (6.5%)
Hepatic disease (yes)	184 (5.7%)
Coronary artery disease (yes)	176 (5.5%)
Anemia (yes)	159 (4.9%)
Malignant disease (yes)	144 (4.5%)
Atopic dermatitis (yes)	136 (4.2%)
Asthma (yes)	127 (3.9%)
Depression (yes)	115 (3.6%)
Reflux esophagitis (yes)	104 (3.2%)
Renal disease (yes)	87 (2.7%)
Stroke (yes)	38 (1.2%)

Results

A total of 3228 male residents participated in the Nagahama project. Of these, four residents were excluded from the study because of insufficient data. The data for the remaining 3224 (99.9%) residents were evaluated in the study. The characteristics and distributions of the participants are shown in Table 1.

Overall, 441 residents (13.7%) had some type of UI. Of these, 334 (10.4%), 46 (1.4%) and 61 (1.9%) residents had mild, moderate and severe UI, respectively. The prevalences of mild UI and moderate to severe UI increased with aging

© 2013 The Japanese Urological Association

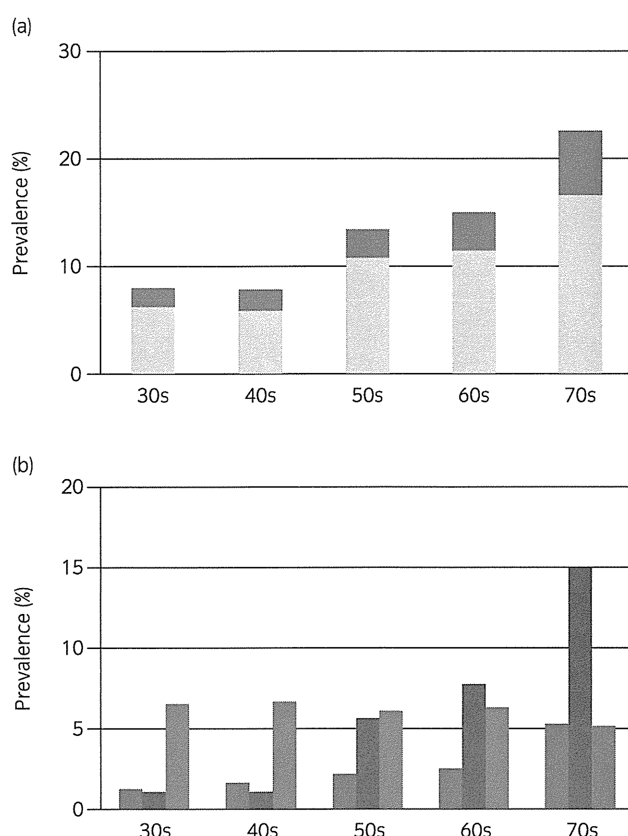

Fig. 1 (a) Prevalences of mild UI and moderate to severe UI stratified by age. (b) Prevalences of the three types of UI stratified by age. ■, moderate to severe UI; ■, mild UI; ■, SUI; ■, UUI; ■, PUI.

(Fig. 1a). Regarding the types of UI, 79 residents (2.5%) had SUI, 204 (6.3%) had UUI and 199 (6.2%) had PUI. A total of 15 residents (0.5%) had MUI. As shown in Figure 1b, the prevalences of SUI and UUI increased with aging, whereas the prevalence of PUI did not change.

The results of the analyses for correlates of each type of UI are shown in Tables 2 and 3. The univariate analyses showed that prostatic disease, age, malignant disease, anemia and hypertension were associated with SUI (Table 2). Among these variables, prostatic disease (OR 2.14), age (OR 1.29), anemia (OR 2.27) and malignant disease (OR 2.22) were independent correlates for SUI in the multivariate analyses (Table 3). Although both prostatic disease and malignant disease were independently correlated with SUI, the number of participants with prostate cancer was 22, which accounted for 10% of all participants with prostatic disease and 15% of all participants with malignant disease. Age, low back pain, diabetes, prostatic disease, hypertension, coronary artery disease, renal disease, malignant disease, sleeplessness, hepatic disease, dyslipidemia, BMI and alcohol intake were associated with UUI in the univariate analyses (Table 2). Among these variables, diabetes (OR 1.98), age (OR 1.92), low back pain (OR 1.74), alcohol intake (OR 1.84), renal disease (OR: 1.91) and prostatic disease (OR 1.48) remained statistically significant in the multivariate analyses (Table 3).

Unlike SUI and UUI, age was not associated with PUI. In contrast, asthma, prostatic disease, rhinosinusitis, sleeplessness, reflux esophagitis, anemia and BMI were associated with PUI in the univariate analyses (Table 2). Among these variables, asthma (OR 3.01), prostatic disease (OR 2.38), rhinosinusitis (OR 1.92), low back pain (OR 1.58) and BMI (OR 0.73) remained statistically significant in the multivariate analyses. Depression was another independent correlate for PUI (OR 3.41). Reflux esophagitis was marginally associated with PUI ($P = 0.09$) (Table 3). The lack of association between age and PUI was also observed after adjustment for the other independent correlates.

Discussion

The development of the three types of UI in men is deemed to be caused by different mechanisms. UUI, which is a component symptom of overactive bladder syndrome, is usually attributed to detrusor overactivity. SUI is induced by dysfunction of the urethral sphincter, which typically results from iatrogenic causes, such as prostatectomy. Meanwhile, the mechanism for the development of PUI is not well recognized. An earlier study by Stephenson *et al.* showed that most patients with PUI had normal urodynamic bladder function.[5] That study also found that half of the patients could not occlude their urethra voluntarily or "milk back" residual urine in the urethra into the bladder, suggesting contraction failure of the bulbocavernosus muscle. However, another study found that patients with PUI had normal reflex and activity of the bulbocavernosus muscle, as assessed by electromyography.[12] Despite the lack of an obvious mechanism, residual urine in the bulbar urethra has been attributed to PUI.[5]

The different mechanisms naturally suggest different prevalences and correlates of the three types of UI, and this was confirmed in the present study. Among the independent correlates of UUI, age, history of diabetes and low back pain had especially strong associations with UUI. Aging and diabetes have frequently been reported as risk factors for UI. Diabetes can induce various types of bladder dysfunctions, including detrusor overactivity,[13] and therefore the association of diabetes with UUI is reasonable. Low back pain is a novel correlate of UUI. Although this symptom can arise from various problems, age-related degenerative processes in the intervertebral disks and facet joints, spinal stenosis, and disk herniation are considered to be common causes.[14] These structural deformations of the spine might induce compression of the spinal cord or spinal roots, and potentially result in neurogenic bladder dysfunction, such as detrusor overactivity. Conversely, one longitudinal study

© 2013 The Japanese Urological Association

Table 2 Univariate analyses for correlates of SUI, UUI and PUI

	SUI		UUI		PUI	
	P-value	OR	*P*-value	OR	*P*-value	OR
Age	<0.001	1.44	<0.001	2.03	0.43	–
BMI	0.40	–	0.034	0.75	0.019	0.72
Alcohol	0.51	–	0.042	0.65	0.36	–
Smoking	0.40	–	0.66	–	0.58	–
Low back pain	0.39	–	<0.001	1.82	<0.001	1.74
Hypertension	0.044	1.62	<0.001	1.88	0.57	–
Rhinosinusitis	0.88	–	0.090	–	<0.001	2.16
Dyslipidemia	0.99	–	0.014	1.62	0.93	–
Prostatic disease	<0.001	3.25	<0.001	2.80	<0.001	2.80
Diabetes	0.48	–	<0.001	2.95	0.77	–
Gout	0.168	–	0.58	–	0.65	–
Sleeplessness	0.40	–	0.005	1.93	0.001	2.09
Hepatic disease	0.80	–	0.010	1.89	0.91	–
Coronary artery disease	0.40	–	0.001		0.050	1.68
Anemia	0.002	2.91	0.52	–	0.017	1.90
Malignant disease	<0.001	3.70	0.002	2.24	0.26	–
Atopic dermatitis	0.85	–	0.56	–	0.222	–
Asthma	0.28	–	0.47	–	<0.001	3.05
Depression	0.29	–	0.78	–	0.229	–
Reflux esophagitis	0.35	–	0.32	–	0.008	2.25
Renal disease	0.44	–	0.001	2.71	0.106	–
Stroke	0.27	–	0.79	–	0.82	–

showed that incontinence increased the risk of future back pain.[15] In any case, low back pain has a close relationship with UUI. PUI, as well as UUI, had an independent association with low back pain. The common symptom of low back pain would be better noted as an associated factor for various lower urinary tract symptoms including UUI and PUI. Similar to the case for UUI, aging was also a strong correlate of SUI. The other independent correlates of SUI were history of prostatic disease, malignant disease and anemia. Some previous studies have reported that men with prostatic disease have higher rates of UI.[16,17] Although it is well known that SUI is one of the most important complications after prostatectomy for prostatic diseases, especially prostate cancer, the present study population had a low prevalence of prostate cancer. Therefore, the association between SUI and prostatic disease and/or malignant disease in the present study would not be mainly attributable to radical prostatectomy.

The most striking observation in the present study was the lack of correlation between aging and PUI. The prevalence of PUI was found to remain constant at 5–6% throughout the generations, and the severity and frequency of this symptom were also not age dependent (data not shown). These observations suggest that this symptom does not arise through mechanical or functional dysfunction induced by aging. Furthermore, the independent correlates of PUI included uncommon parameters, such as asthma and rhinosinusitis; along with relatively common parameters, such as prostatic disease, sleeplessness and depression. There was also a marginal association between reflux esophagitis and PUI. Although some recent studies have suggested a relationship between overactive bladder and asthma,[18,19] asthma had no association with UUI, but was associated with PUI in the present study. It is recognized that asthma and rhinitis/rhinosinusitis have a strong correlation with each other.[20] The most accepted explanation for the interaction between the upper and lower airways is increased oral breathing and systemic response. In other words, impaired filtering in the nose of patients with rhinitis leads to mouth breathing, which can result in increased exposure of the lower airways to allergens. It is also known that asthmatic patients have a much greater risk of gastro-esophageal reflux disease-related symptoms and vice versa.[20] Gastro-esophageal reflux disease, including reflux esophagitis, could worsen asthma either by direct effects on airway responsiveness or through aspiration-induced inflammation, and the bronchoconstriction observed in asthma, as well as asthma medications, could induce gastro-esophageal reflux. Although these three diseases, asthma, rhinosinusitis and reflux esophagitis, have associations with one another and with allergy, the present

© 2013 The Japanese Urological Association

Table 3 Multivariate analyses for correlates of SUI, UUI and PUI

	SUI			UUI			PUI		
	P-value	OR	95% CI	*P*-value	OR	95% CI	*P*-value	OR	95% CI
Age	0.013	1.29	1.06–1.57	<0.001	1.92	1.64–2.24	–		
BMI	–			0.10			0.025	0.73	0.55–0.96
Alcohol	–			0.005	1.84	1.20–2.82	–		
Smoking	–			–			–		
Low back pain	–			<0.001	1.74	1.29–2.34	0.003	1.58	1.17–2.12
Hypertension	0.40			0.59			–		
Rhinosinusitis	–			0.12			0.001	1.92	1.33–2.77
Dyslipidemia	–			0.36			–		
Prostatic disease	0.009	2.14	1.21–3.77	0.045	1.48	1.01–2.18	<0.001	2.38	1.63–3.48
Diabetes	–			<0.001	1.98	1.36–2.90	–		
Gout	0.28			–			–		
Sleeplessness	–			0.33			0.013	1.86	1.14–3.04
Hepatic disease	–			0.24			–		
Coronary artery disease	–			0.42			0.33		
Anemia	0.022	2.27	1.13–4.57	–			0.26		
Malignant disease	0.025	2.22	1.10–4.45	0.53			–		
Atopic dermatitis	–			–			0.10		
Asthma	–			–			<0.001	3.01	1.79–5.06
Depression	–			–			0.024	3.41	1.17–9.90
Reflux esophagitis	–			–			0.09		
Renal disease	–			0.045	1.91	1.01–3.60	0.36		
Stroke	–			–			–		

study showed that each of these diseases was independently associated with PUI. As the mechanism for the development of PUI is completely unknown, as aforementioned, we cannot even speculate about the reasons why these diseases have relationships with PUI. Further investigations in this area are warranted in the future.

There have been a few epidemiological studies investigating the prevalence of PUI.[6,8,9] The previously reported results for the prevalence of PUI were quite variable ranging from 5.5 to 38.3%. Including the present study, the highly variable percentages might result from different questioning methods, interviews or self-reporting, and different races of the study populations. However, the most likely reason for the difference would be subtle differences in the formulation of questions indicating PUI. We used the ICIQ-SF, whereas others used a questionnaire based on the International Continence Society definition. Every question used in studies on the prevalence of PUI is validated for use, and the results obtained are significant. However, it is noteworthy that subtle differences in the formulation of questions can influence the outcomes.

The present study had several limitations. One is the ambiguity in the names of several diseases/conditions. For example, prostatic disease includes benign prostatic hyperplasia, prostate cancer and prostatitis, and we cannot distinguish concrete disease entities. Low back pain is also an ambiguous name for a symptom that can result from various diseases as described earlier. However, in actuality, many of the participants, who are amateurs in the medical field, did not know the exact disease causing their symptom of low back pain. Another limitation was the lack of data on bladder or bowel diseases, which can affect the function of the lower urinary tract. Despite these limitations, we believe that our observations provide some clues for the future solution of the mechanism of PUI.

Conflict of interest

None declared.

References

1 Shamliyan TA, Wyman JF, Ping R, Wilt TJ, Kane RL. Male urinary incontinence: prevalence, risk factors, and preventive interventions. *Rev. Urol.* 2009; **11**: 145–65.

2 Engstrom G, Engstrom MLW, Henningsohn L, Loof L, Leppert J. Prevalence of distress and symptoms severity from the lower urinary tract in men: a population-based study with the DAN-PSS questionnaire. *Fam. Pract.* 2003; **21**: 617–22.

© 2013 The Japanese Urological Association

3 Coyne KS, Kvasz M, Ireland AM, Milsom I, Kopp ZS, Chapple CR. Urinary incontinence and its relationship to mental health and health-related quality of life in men and women in Sweden, the United Kingdom, and the United States. *Eur. Urol.* 2012; **61**: 88–95.
4 Abrams P, Cardozo L, Fall M *et al.* The standardisation of terminology of lower urinary tract function: report from the Standardisation Sub-committee of the International Continence Society. *Neurourol. Urodyn.* 2002; **21**: 167–78.
5 Stephenson TP, Farrar DJ. Urodynamic study of 15 patients with postmicturition dribble. *Urology* 1977; **9**: 404–6.
6 Malmsten UG, Molander U, Peeker R, Irwin DE, Milsom I. Urinary incontinence, overactive bladder, and other lower urinary tract symptoms: a longitudinal population-based survey in men aged 45–103 years. *Eur. Urol.* 2010; **58**: 149–56.
7 Hrisanfow E, Hagglund D. The prevalence of urinary incontinence among women and men with chronic obstructive pulmonary disease in Sweden. *J. Clin. Nurs.* 2011; **20**: 1895–905.
8 Irwin DE, Milsom I, Hunskaar S *et al.* Population-based survey of urinary incontinence, overactive bladder, and other lower urinary tract symptoms in five countries: results of the EPIC study. *Eur. Urol.* 2006; **50**: 1306–15.
9 Coyne KS, Sexton CC, Thompson CL *et al.* The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: results from the Epidemiology of LUTS (EpiLUTS) study. *BJU Int.* 2009; **104**: 352–60.
10 Nakayama T. Implementation of a genomic epidemiology and public health activities in a community: establishing a "rule" for the protection and utilization of individual information. *Jpn. J. Genet. Counsel.* 2010; **31**: 117–21.
11 Gotoh M, Donovan J, Corcos J *et al.* Scored ICIQ-SF (International Consultation on Incontinence Questionnaire-Short Form) for symptoms and QOL assessment in patients with urinary incontinence. *Nippon Hainyokino Gakkaishi* 2001; **12**: 227–31 (in Japanese).
12 Paquin JM, D'Amour M, Perreault JP, Faucher R, Mauffette F. Electromyographic study of bulbocavernosus muscle in 15 patients with postmicturition dribbling. *Urology* 1985; **26**: 603–5.
13 Daneshgari F, Liu G, Birder L, Mitchell ATH, Chacko S. Diabetic bladder dysfunction: current translational knowledge. *J. Urol.* 2009; **182**: S18–26.
14 Deyo RA, Weinstein JN. Low back pain. *N. Engl. J. Med.* 2001; **344**: 363–70.
15 Smith MD, Russell A, Hodges PW. Do incontinence, breathing difficulties, and gastrointestinal symptoms increase the risk of future back pain? *J. Pain* 2009; **10**: 876–86.
16 Van Oyen H, van Oyen P. Urinary incontinence in Belgium: prevalence, correlates and psychosocial consequences. *Acta Clin. Belg.* 2002; **57**: 207–18.
17 Markland AD, Goode PS, Burgio KL *et al.* Correlates of urinary, fecal, and dual incontinence in older African-American and white men and women. *J. Am. Geriatr. Soc.* 2008; **56**: 285–90.
18 Coyne KS, Sexton CC, Irwin DE, Kopp ZS, Kelleher CJ, Milsom I. The impact of overactive bladder, incontinence and other lower urinary tract symptoms on quality of life, work productivity, sexuality and emotional well-being in men and women: results from the EPIC study. *BJU Int.* 2008; **101**: 1388–95.
19 Correia S, Dinis P, Rolo F, Lunet N. Prevalence, treatment and known risk factors of urinary incontinence and overactive bladder in the non-institutionalized Portuguese population. *Int. Urogynecol. J.* 2009; **20**: 1481–9.
20 Boulet LP, Boulay ME. Asthma-related comorbidities. *Expert Rev. Respir. Med.* 2011; **5**: 377–93.

Appendix I

The Nagahama Cohort Research Group

Yuri Dowa, Department of Medical-ethics/Medical Genetics, School of Public Health, Kyoto University.

Tasuro Ishizaki, Department of Health Informatics, School of Public Health, Kyoto University.

Hiroshi Kadotani, Center for Genomic Medicine, Graduate School of Medicine, Kyoto University.

Takahisa Kawaguchi, Center for Genomic Medicine, Graduate School of Medicine, Kyoto University.

Michiaki Mishima, Department of Respiratory Medicine, Graduate School of Medicine, Kyoto University.

Koichi Miyaki, Department of Health Informatics, School of Public Health, Kyoto University.

Seiho Mizusawa, Center for Genomic Medicine, Graduate School of Medicine, Kyoto University.

Yoshimitsu Takahashi, Department of Health Informatics, School of Public Health, Kyoto University.

© 2013 The Japanese Urological Association

nature genetics

Genome-wide association analyses in east Asians identify new susceptibility loci for colorectal cancer

Wei-Hua Jia[1,16], Ben Zhang[2,16], Keitaro Matsuo[3], Aesun Shin[4], Yong-Bing Xiang[5], Sun Ha Jee[6], Dong-Hyun Kim[7], Zefang Ren[1], Qiuyin Cai[2], Jirong Long[2], Jiajun Shi[2], Wanqing Wen[2], Gong Yang[2], Ryan J Delahanty[2], Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO)[8], Colon Cancer Family Registry (CCFR)[8], Bu-Tian Ji[9], Zhi-Zhong Pan[1], Fumihiko Matsuda[10], Yu-Tang Gao[5], Jae Hwan Oh[11], Yoon-Ok Ahn[12], Eun Jung Park[6], Hong-Lan Li[5], Ji Won Park[11], Jaeseong Jo[6], Jin-Young Jeong[7], Satoyo Hosono[3], Graham Casey[13], Ulrike Peters[14,15], Xiao-Ou Shu[2], Yi-Xin Zeng[1,17] & Wei Zheng[2,17]

© 2013 Nature America, Inc. All rights reserved.

To identify new genetic factors for colorectal cancer (CRC), we conducted a genome-wide association study in east Asians. By analyzing genome-wide data in 2,098 cases and 5,749 controls, we selected 64 promising SNPs for replication in an independent set of samples, including up to 5,358 cases and 5,922 controls. We identified four SNPs with association *P* values of 8.58×10^{-7} to 3.77×10^{-10} in the combined analysis of all east Asian samples. Three of the four were replicated in a study conducted in 26,060 individuals of European descent, with combined *P* values of 1.22×10^{-10} for rs647161 (5q31.1), 6.64×10^{-9} for rs2423279 (20p12.3) and 3.06×10^{-8} for rs10774214 (12p13.32 near the *CCND2* gene), derived from meta-analysis of data from both east Asian and European-ancestry populations. This study identified three new CRC susceptibility loci and provides additional insight into the genetics and biology of CRC.

CRC is one of the most commonly diagnosed malignancies in east Asia and many other parts of the world[1]. Genetic factors have an important role in the etiology of both sporadic and familial CRC[2]. However, less than 6% of CRC cases can be explained by rare, high-penetrance variants in the CRC susceptibility genes identified to date, such as the *APC, SMAD4, AXIN2, BMPR1A, POLD1, STK11, MUTYH* and DNA mismatch repair genes[2]. Over the past two decades, many candidate gene studies have evaluated common genetic risk factors for CRC; only a few of these have been replicated in subsequent studies[3]. Recent genome-wide association studies (GWAS) have identified approximately 15 common genetic susceptibility loci for CRC[4–12]. However, these newly identified genetic factors, along with known high-penetrance variations in CRC susceptibility genes, explain less than 15% of the heritability for this common malignancy[10,11]. Furthermore, with the exception of a small study conducted in Japan[12], all other GWAS have been conducted in populations of European ancestry, which differ from other populations in certain features of genetic architecture. Many of the variants discovered in populations of European ancestry show only weak or no association with CRC in other ancestry groups[13]. Therefore, additional GWAS are needed, particularly in populations not of European ancestry, to fully uncover the genetic basis for CRC susceptibility.

In 2009, we initiated the Asia Colorectal Cancer Consortium (ACCC), a GWAS in east Asians, to search for previously unknown genetic risk factors for CRC. The discovery stage (stage 1) consisted of five GWAS conducted in China, Korea and Japan, including 2,293 CRC cases and 5,780 controls (**Supplementary Table 1**). Cases and controls were genotyped using several SNP arrays, including the Affymetrix Genome-Wide Human SNP Array 6.0 (906,602 SNPs), the Affymetrix Genome-Wide Human SNP Array 5.0 (443,104 SNPs), the Illumina Infinium HumanHap610 BeadChip (592,044 SNPs), the Illumina Human610-Quad BeadChip (620,901 SNPs) and the Illumina HumanOmniExpress BeadChip (729,462 SNPs) (**Supplementary Table 1**). After quality control exclusions as described previously[14–17], 2,098 cases and 5,749 controls remained for this study (**Supplementary Tables 1** and **2**). Also excluded from the analyses were SNPs with call rate of <95%, genotype concordance rate of <95%

[1]State Key Laboratory of Oncology in South China, Cancer Center, Sun Yat-sen University, Guangzhou, China. [2]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [3]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan. [4]Molecular Epidemiology Branch, National Cancer Center, Goyang-si, Korea. [5]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China. [6]Institute for Health Promotion, Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University, Seoul, Korea. [7]Department of Social and Preventive Medicine, Hallym University College of Medicine, Okcheon-dong, Korea. [8]A complete list of members is provided in the Acknowledgements. [9]Division of Cancer Epidemiology & Genetics, National Cancer Institute, Bethesda, Maryland, USA. [10]Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [11]Center for Colorectal Cancer, National Cancer Center, Goyang-si, Korea. [12]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea. [13]Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA. [14]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [15]Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA. [16]These authors contributed equally to this work. [17]These authors jointly directed this work. Correspondence should be addressed to W.Z. (wei.zheng@vanderbilt.edu).

Received 29 March; accepted 29 November; published online 23 December 2012; doi:10.1038/ng.2505

Table 1 Association of CRC risk with the top four risk variants identified in east Asian samples

						Cases		Controls		Per-allele association		Heterogeneity	
SNP	Alleles[a]	Chr.	Gene[b]	Location (bp)[c]	Stage	Sample size	MAF	Sample size	MAF	OR (95% CI)[d]	P_{trend}	P^e	I^2
rs10774214	T/C	12p13.32	*CCND2*	4238613	GWAS	2,098	0.373	5,749	0.348	1.20 (1.09–1.32)	2.03×10^{-4}		
					Replication	5,197	0.381	5,797	0.355	1.16 (1.09–1.23)	5.80×10^{-7}		
					Overall	7,295	0.379	11,546	0.352	1.17 (1.11–1.23)	5.48×10^{-10}	0.615	0%
rs647161	A/C	5q31.1	*PITX1*	134526991	GWAS	2,098	0.353	5,749	0.308	1.22 (1.12–1.33)	3.29×10^{-6}		
					Replication	5,217	0.344	5,815	0.319	1.14 (1.07–1.21)	1.15×10^{-5}		
					Overall	7,315	0.347	11,564	0.313	1.17 (1.11–1.22)	3.77×10^{-10}	0.444	0%
rs2423279	C/T	20p12.3	*HAO1*	7760350	GWAS	2,098	0.339	5,749	0.307	1.16 (1.07–1.26)	4.96×10^{-4}		
					Replication	5,227	0.315	5,811	0.297	1.13 (1.06–1.19)	1.22×10^{-4}		
					Overall	7,325	0.322	11,560	0.302	1.14 (1.08–1.19)	2.29×10^{-7}	0.331	12%
rs1665650	T/C	10q26.12	*HSPA12A*	118477090	GWAS	2,098	0.346	5,749	0.310	1.20 (1.10–1.31)	3.88×10^{-5}		
					Replication	5,192	0.328	5,808	0.320	1.10 (1.04–1.17)	0.0018		
					Overall	7,290	0.333	11,557	0.315	1.13 (1.08–1.19)	8.58×10^{-7}	0.404	4%

Chr., chromosome; OR, odds ratio; CI, confidence interval.
[a]Minor/major allele for east Asians. OR was estimated for the minor allele. [b]Closest gene. [c]Location based on NCBI Human Genome Build 36.3. [d]Adjusted for age, sex, the first ten principal components (stage 1) and study site. [e]*P* for heterogeneity across studies in GWAS and replication was calculated using Cochran's *Q* test.

© 2013 Nature America, Inc. All rights reserved.

between positive control samples, minor allele frequency (MAF) of <5% or *P* value for Hardy-Weinberg equilibrium of $<1.0 \times 10^{-5}$ in controls for each study. Imputation was conducted for each study following the MaCH algorithm[18] using phased HapMap 2 Han Chinese in Beijing, China (CHB) and Japanese in Tokyo, Japan (JPT) samples as the reference. No apparent genetic admixture was detected, except for one sample from KCPS-II (**Supplementary Fig. 1**). Associations between CRC risk and each of the genotyped and imputed SNPs were evaluated using logistic regression within each study after adjusting for age, sex and the first ten principal components using mach2dat[18]. Meta-analyses were conducted under a fixed-effects model using the METAL program[19]. There was little evidence for inflation in the association test statistics for any of the five studies (genomic inflation factor (λ) range of 1.02 to 1.04) or for all studies combined ($\lambda = 1.01$) (**Supplementary Fig. 2** and **Supplementary Table 1**). The observed number of SNPs with small *P* values was slightly larger than that expected by chance (**Supplementary Fig. 2**).

Multiple genomic locations were found that were potentially related to CRC risk (**Supplementary Fig. 3**). Nine SNPs identified from published GWAS conducted in populations of European ancestry showed associations with CRC risk at $P < 0.05$ in stage 1 (data not shown). To improve the statistical power for evaluating these SNPs, we genotyped 6,476 additional samples to bring the total sample size to 5,252 cases and 9,071 controls. Except for the 2 SNPs that are monomorphic in east Asians (rs6691170 and rs16892766), all 16 of the other SNPs identified from published GWAS conducted in European-ancestry populations showed association with CRC risk in the same direction as reported previously (**Supplementary Table 3**). A significant association with CRC risk at $P < 0.05$ was found for 13 SNPs, including rs6687758, rs10936599, rs10505477, rs6983267, rs7014346, rs10795668, rs3802842, rs4444235, rs4779584, rs9929218, rs4939827, rs10411210 and rs961523. Except for two SNPs (rs6983267 and rs4779584), no statistically significant heterogeneity at $P < 0.05$ was observed between east Asian and European-ancestry populations (**Supplementary Table 3**).

To identify new genetic factors for CRC, we selected 64 SNPs for replication in an independent set of 5,358 cases and 5,922 controls recruited in 5 studies conducted in China, Korea and Japan (**Supplementary Table 2**). SNPs were selected from among those that (i) had MAF of >5%; (ii) showed no heterogeneity across studies ($P_{het} > 0.05$ and $I^2 < 25\%$); (iii) were not in linkage disequilibrium (LD; $r^2 < 0.2$) with any known CRC risk variant reported from previous GWAS; (iv) had high imputation quality in each of the five studies (RSQ > 0.5); and (v) were associated at $P < 0.01$ in the combined analysis of all five studies included in stage 1. These criteria were used to prioritize SNPs for replication.

Of the 64 SNPs evaluated in stage 2, 7 showed association with CRC risk at $P < 0.05$ with a direction of association consistent with that observed in stage 1 (**Table 1** and **Supplementary Table 4**). In the combined analysis of data from stages 1 and 2, *P* values for associations with two SNPs (rs647161 at 5q31.1, odds ratio (OR) = 1.17, $P = 3.77 \times 10^{-10}$, and rs10774214 at 12p13.32, OR = 1.17, $P = 5.48 \times 10^{-10}$) were lower than the conventional genome-wide significance level of 5.0×10^{-8}, providing convincing evidence for an association of these SNPs with CRC risk (**Table 1**). An additional SNP, rs2423279, showed a significant association in stage 2 after Bonferroni correction (corrected $P < 7.8 \times 10^{-4}$) but did not reach the conventional GWAS significance level for association with CRC risk in the combined analysis of all samples (OR = 1.14, $P = 2.29 \times 10^{-7}$). The association between CRC risk and each of these three SNPs was consistent across most studies (**Fig. 1**). Results for the other four SNPs that replicated in stage 2 at $P < 0.05$ (rs1665650, rs2850966, rs1580743 and rs4503064) are also presented (**Supplementary Table 4**), including one SNP (rs1665650) with an association *P* value of 8.58×10^{-7} in the combined analysis of all data from both stages (**Table 1**).

We next evaluated these top four SNPs (**Table 1**) using data from GWAS in the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO) and the Colon Cancer Family Registry (CCFR), which together include 11,870 cases and 14,190 controls of European ancestry[4,20,21]. Three of the four SNPs were replicated in the GECCO and CCFR sample, although the strength of the associations was weaker than in east Asians (**Table 2**). These results provide independent support of our findings in the east Asian population. Meta-analyses of data from both east Asian and European-ancestry populations provided strong evidence for associations of CRC risk with three SNPs, with *P* values all below the genome-wide significance threshold of 5×10^{-8} (**Table 2**). The weaker associations observed in European-ancestry populations could be explained, in part, by differences in LD patterns at these loci for east

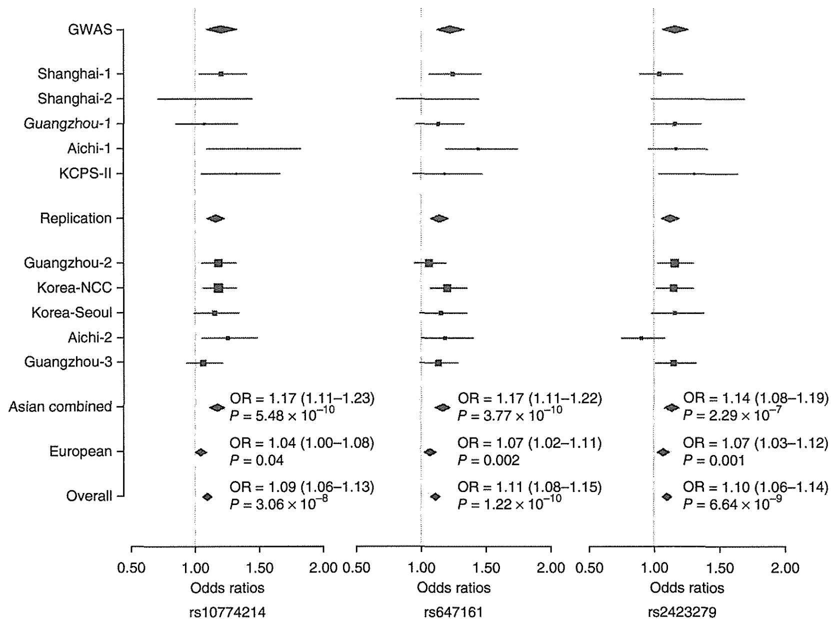

Figure 1 Forest plots for the three SNPs showing evidence of an association with CRC risk. Per-allele ORs are presented, with the area of each box proportional to the inverse variance weight of the estimate. Horizontal lines represent 95% confidence intervals.

© 2013 Nature America, Inc. All rights reserved.

Asians and Europeans (**Supplementary Fig. 4**). It is possible that causal variants in these regions are tagged by different SNPs in these two populations or that there is allelic heterogeneity, in which different underlying causal variants exist in populations of Asian and European ancestry. The difference in LD structure between Asian and European descendants and possible allelic heterogeneity in these two populations might explain, in part, why these loci were not discovered in previous studies conducted in individuals of European ancestry. The fourth SNP evaluated in the GECCO and CCFR sample, rs1665650, however, was not replicated in individuals with European ancestry (OR = 0.96, $P = 0.05$).

Stratification analyses showed that the association of CRC risk with each of the three replicated SNPs was generally consistent in Chinese, Korean and Japanese individuals ($P_{het} > 0.05$), although the association with rs2423279 was not statistically significant in Japanese, perhaps owing to a small sample size (**Supplementary Table 5**). Associations of these three SNPs with CRC risk were similar for men and women ($P_{het} > 0.05$) (**Supplementary Table 6**).

The rs10774214 SNP is located just 15 kb upstream of *CCND2*, the gene encoding cyclin D2 (**Fig. 2a**), a member of the D-type cyclin family, which also includes cyclins D1 and D3. These cyclins have a critical role in cell cycle control (from G1 to S phase) through activation of cyclin-dependent kinases (CDKs), primarily CDK4 and CDK6 (ref. 22). *CCND2* is closely related to *CCND1*, a well-established human oncogene[22,23]. Although *CCND2* has been less well studied than *CCND1*, several studies, including The Cancer Genome Atlas (TCGA), have shown that *CCND2* is overexpressed in a substantial proportion of human colorectal tumors[22–25]. Overexpression of this cyclin may be an independent predictor of survival in individuals with CRC[24]. Several other genes, including *PARP11*, *FGF23*, *FGF6*, *C12orf5* and *RAD51AP1*, are also in close proximity to the SNP identified in our study, of which both *C12orf5* (also known as *TIGAR*, encoding TP53-induced glycolysis and apoptosis regulator) and *RAD51AP1* were found to be overexpressed in CRC tissue included in TCGA[25]. rs10774214 is in strong LD with several SNPs that are located in potential transcription factor–binding sites, as determined using the TRANSFAC database[26]. Additional research may be warranted regarding possible mechanisms by which this SNP is related to CRC risk.

The rs647161 SNP is located on chromosome 5q31.1, where a cluster of SNPs were associated with CRC risk (**Fig. 2b**). Of the genes in this region (including *PITX1*, *CATSPER3*, *PCBD2*, *MIR4461* and *H2AFY*), *PITX1* is the closest to rs647161 (approximately 129 kb upstream). The *PITX1* gene (encoding paired-like homeodomain 1) has been described as a tumor suppressor gene and may be involved in the tumorigenesis of multiple human cancers[27–31], including CRC[27,32]. *PITX1* has been reported to suppress tumorigenicity by downregulating the RAS pathway, which is frequently altered in colorectal tumors[27]. Inhibition of *PITX1* induces the RAS pathway and tumorigenicity, and restoring *PITX1* in colon cancer cells inhibits tumorigenicity[27]. It also has been reported that *PITX1* may activate *TP53* (ref. 33) and regulate telomerase activity[34]. Consistent with its possible function as a tumor suppressor gene, *PITX1* has been found to be downregulated in human cancer tissue samples and cell lines[27–30,32]. CRC tissue expressing wild-type KRAS showed significantly lower expression of *PITX1* than tissue with mutant KRAS[32]. Most recently, low *PITX1* expression was found to be associated with poor survival in individuals with CRC[35]. In addition, rs6596201, which is in moderate LD with rs647161 ($r^2 = 0.25$), is an expression quantitative trait locus (eQTL) ($P = 2.42 \times 10^{-28}$) for the *PITX1* gene[36]. Several other genes at this locus, including *C5orf24*, *H2AFY* and *NEUROG1*, were also found to be highly expressed in colorectal tumors included in TCGA ($P < 0.001$)[25]. Additional studies are warranted to explore a possible role for these genes in the etiology of CRC.

Table 2 Association of CRC risk with the top three risk variants in European descendants and east Asian and European descendants combined

		MAF[b]		European-ancestry populations[c]				East Asian and European-ancestry populations combined[c]			
SNP	Alleles[a]	Cases	Controls	Cases	Controls	OR (95% CI)	P_{meta}	Cases	Controls	OR (95% CI)	P_{meta}
rs10774214	T/C	0.385	0.379	11,870	14,190	1.04 (1.00–1.09)	0.040	19,165	25,736	1.09 (1.06–1.13)	3.06×10^{-8}
rs647161	A/C	0.680	0.667	11,870	14,190	1.07 (1.02–1.11)	0.002	19,185	25,754	1.11 (1.08–1.15)	1.22×10^{-10}
rs2423279	C/T	0.263	0.252	11,870	14,190	1.07 (1.03–1.12)	0.001	19,195	25,750	1.10 (1.06–1.14)	6.64×10^{-9}

[a]Alleles (minor/major) for east Asians. [b]MAF in European-ancestry populations. [c]Summary statistics were generated using inverse variance–weighted fixed-effects meta-analysis.

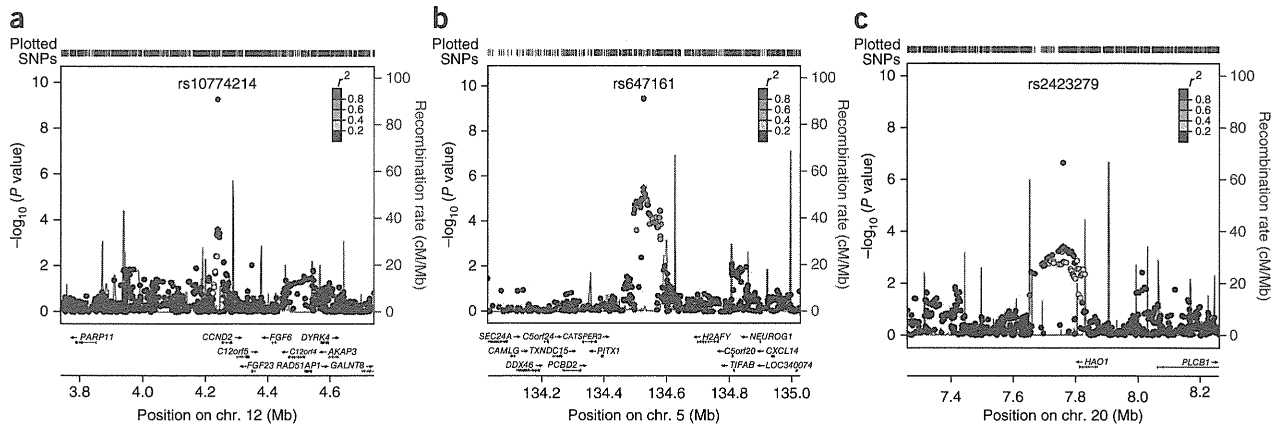

Figure 2 Regional plots of association results and recombination rates for the three SNPs showing evidence of association with CRC risk. Genotyped and imputed data from GWAS samples are plotted on the basis of their chromosomal position in NCBI Human Genome Build 36.3. For each region, the SNP selected for stage 2 replication is denoted with a diamond, and the *P* value from the combined analysis of stage 1 and 2 data is provided. (**a**–**c**) Data are shown for rs10774214 (**a**), rs647161 (**b**) and rs2423279 (**c**).

The rs2423279 SNP is located on chromosome 20p12.3, close to the *HAO1* and *PLCB1* genes (**Fig. 2c**). *HAO1* encodes hydroxyacid oxidase, which oxidizes 2-hydroxyacid. *PLCB1* encodes phospholipase C-β1, which has an important role in the intracellular transduction of many extracellular signals. Overexpression of the *PLCB1* gene has been observed in CRC tissue[25]. Possible mechanisms by which these genes are involved in CRC carcinogenesis are unknown. The rs2423279 SNP is 1,408,069 bp downstream of rs961253, a SNP previously identified in a European GWAS as being associated with CRC risk[10]. However, these two SNPs are not correlated in east Asians ($r^2 = 0$) or in Europeans ($r^2 = 0$). Adjustment for rs961253 did not change the results for rs2423279 (data not shown).

To our knowledge, this is the largest GWAS performed for CRC in east Asians, a population that differs from populations of European ancestry in CRC risk and certain aspects of genetic architecture. Results from our study, along with data from a large study conducted in a population of European ancestry, provide convincing evidence of associations with CRC risk for three new independent susceptibility loci at 5q31.1, 12p13.32 and 20p12.3. Results from this study provide new insights into the genetics and biology of CRC.

URLs. Cancer Genetic Markers of Susceptibility (CGEMS), http://cgems.cancer.gov/; Database of Genotypes and Phenotypes (dbGaP), http://www.ncbi.nlm.nih.gov/gap; EIGENSTRAT, http://genepath.med.harvard.edu/~reich/EIGENSTRAT.htm; eqtl.uchicago.edu, http://eqtl.uchicago.edu/Home.html; GTEx eQTL Browser, http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi; Haploview, http://www.broad.mit.edu/mpg/haploview/; HapMap Project, http://hapmap.ncbi.nlm.nih.gov/; IntOGen, http://www.intogen.org/home; LocusZoom, http://csg.sph.umich.edu/locuszoom/; MaCH 1.0, http://www.sph.umich.edu/csg/abecasis/MACH/; mach2dat, http://genome.sph.umich.edu/wiki/Mach2dat:_Association_with_MACH_output; METAL, http://www.sph.umich.edu/csg/abecasis/Metal/; PLINK version 1.07, http://pngu.mgh.harvard.edu/~purcell/plink/; R version 2.13.0, http://www.r-project.org/; SAS version 9.2, http://www.sas.com/; SNAP, http://www.broadinstitute.org/mpg/snap/; TRANSFAC, http://www.gene-regulation.com/pub/databases.html; UCSC Genome Browser, http://genome.ucsc.edu/; WHI investigators, https://cleo.whi.org/researchers/SitePages/Write%20a%20Paper.aspx.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. The authors wish to thank the study participants and research staff for their contributions and commitment to this project, R. Courtney for DNA preparation, J. He for data processing and analyses, and M.J. Daly for clerical support in manuscript preparation. This research was supported in part by US National Institutes of Health (NIH) grants R37CA070867, R01CA082729, R01CA124558, R01CA148667 and R01CA122364, as well as by Ingram Professorship and Research Reward funds from the Vanderbilt University School of Medicine. Participating studies (grant support) in the consortium are as follows: Shanghai Women's Health Study (US NIH, R37CA070867), Shanghai Men's Health Study (US NIH, R01CA082729), Shanghai Breast and Endometrial Cancer Studies (US NIH, R01CA064277 and R01CA092585; contributing only controls), Guangzhou Colorectal Cancer Study (National Key Scientific and Technological Project, 2011ZX09307-001-04, and the National Basic Research Program, 2011CB504303, contributing only controls; the Natural Science Foundation of China, 81072383, contributing only controls), Aichi Colorectal Cancer Study (Grant-in-Aid for Cancer Research, the Grant for the Third Term Comprehensive Control Research for Cancer and Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology, 17015018 and 221S0001), Korea–National Cancer Center Colorectal Cancer Study (Basic Science Research Program through the National Research Foundation of Korea, 2010-0010276; National Cancer Center Korea, 0910220), Korea-Seoul Colorectal Cancer Study (none reported) and KCPS-II colorectal Cancer Study (National R&D Program for Cancer Control, 0920330; Seoul R&D Program, 10526).

We wish to thank all participants, staff and investigators of GECCO and CCFR for making it possible to present the results in individuals of European ancestry for new CRC susceptibility loci identified in east Asians. Investigators (institution and location) from GECCO and CCFR who provided support for this project include (in alphabetical order) Aaron K. Aragaki (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), John A. Baron (Division of Gastroenterology and Hepatology, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA), Sonja I. Berndt (Division of Cancer Epidemiology and Genetics, National Cancer Institute, US NIH, Bethesda, Maryland, USA), Stéphane Bezieau (Service de Génétique Médicale, Centre Hospitalier Universitaire (CHU) Nantes, Nantes, France), Hermann Brenner, Katja Butterbach (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany), Bette J. Caan (Division of Research, Kaiser Permanente Medical Care Program, Oakland, California, USA), Christopher S. Carlson (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and School of Public Health,

© 2013 Nature America, Inc. All rights reserved.
npg

University of Washington, Seattle, Washington, USA), Graham Casey (Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA), Andrew T. Chan (Division of Gastroenterology, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts, USA, and Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA), Jenny Chang-Claude (Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany), Stephen J. Chanock (Division of Cancer Epidemiology and Genetics, National Cancer Institute, US NIH, Bethesda, Maryland, USA), Lin S. Chen (Department of Health Studies, University of Chicago, Chicago, Illinois, USA), Gerhard A. Coetzee (Keck School of Medicine, University of Southern California, Los Angeles, California, USA), Simon G. Coetzee (Keck School of Medicine, University of Southern California, Los Angeles, California, USA), David V. Conti (Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA), Keith Curtis (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), David Duggan (Translational Genomics Research Institute, Phoenix, Arizona, USA), Todd L. Edwards (Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA), Charles S. Fuchs (Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA, and Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA), Steven Gallinger (Department of Surgery, Mount Sinai Hospital, Toronto, Ontario, Canada, and Samuel Lunenfeld Research Institute, Toronto, Ontario, Canada), Edward L. Giovannucci (Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA, and Departments of Epidemiology and Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA), Stephanie M. Gogarten (School of Public Health, University of Washington, Seattle, Washington, USA), Stephen B. Gruber (Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California, USA), Robert W. Haile (Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA), Tabitha A. Harrison (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Richard B. Hayes (Division of Epidemiology, Department of Environmental Medicine, New York University School of Medicine, New York, New York, USA), Michael Hoffmeister (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany), John L. Hopper (Melbourne School of Population Health, The University of Melbourne, Melbourne, Victoria, Australia), Li Hsu (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and Department of Biostatistics, University of Washington, Seattle, Washington, USA), Thomas J. Hudson (Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada), David J. Hunter (Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA), Carolyn M. Hutter (Division of Cancer Control and Population Sciences, National Cancer Institute, US NIH, Bethesda, Maryland, USA), Rebecca D. Jackson (Division of Endocrinology, Diabetes, and Metabolism, Ohio State University, Columbus, Ohio, USA), Mark A. Jenkins (Melbourne School of Population Health, The University of Melbourne, Melbourne, Victoria, Australia), Shuo Jiao (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Charles Kooperberg (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Sébastien Küry (Service de Génétique Médicale, CHU Nantes, Nantes, France), Andrea Z. LaCroix (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Cathy C. Laurie (Department of Biostatistics, University of Washington, Seattle, Washington, USA), Cecelia A. Laurie (Department of Biostatistics, University of Washington, Seattle, Washington, USA), Loic Le Marchand (Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii, USA), Mathieu Lemire (Ontario Institute for Cancer Research, Toronto, Ontario, Canada), David Levine (School of Public Health, University of Washington, Seattle, Washington, USA), Noralane M. Lindor (Department of Health Sciences Research, Mayo Clinic, Scottsdale, Arizona, USA), Yan Liu (Stephens and Associates, Carrollton, Texas, USA), Jing Ma (Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA), Karen W. Makar (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Polly A. Newcomb (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA), Ulrike Peters (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA), John D. Potter (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA, and Centre for Public Health Research, Massey University, Palmerston North, New Zealand), Ross L. Prentice (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Conghui Qu (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA), Thomas Rohan (Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Yeshiva University, Bronx, New York, USA), Robert E. Schoen (Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA), Fredrick R. Schumacher (Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA), Daniela Seminara (Division of Cancer Control and Population Sciences, National Cancer Institute, US NIH, Bethesda, Maryland, USA), Martha L. Slattery (Department of Internal Medicine, University of Utah Health Sciences Center, Salt Lake City, Utah, USA), Darin Taverna (Translational Genomics Research Institute, Phoenix, Arizona, USA), Stephen N. Thibodeau (Department of Laboratory Medicine, Mayo Clinic, Rochester, Minnesota, USA, and Department of Pathology and Laboratory Genetics, Mayo Clinic, Rochester, Minnesota, USA), Cornelia M. Ulrich (Division of Preventive Oncology, German Cancer Research Center, Heidelberg, Germany, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA), Raakhee Vijayaraghavan (Genetic Basis of Human Disease Division, Translational Genomics Research Institute, Phoenix, Arizona, USA), Bruce Weir (Department of Biostatistics, University of Washington, Seattle, Washington, USA), Emily White (Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, and Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA) and Brent W. Zanke (Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada).

We also thank B. Buecher of ASTERISK; U. Handte-Daub, M. Celik, R. Hettler-Jensen, U. Benscheid and U. Eilber of DACHS; P. Soule, H. Ranu, I. Devivo, D. Hunter, Q. Guo, L. Zhu and H. Zhang of HPFS, NHS and PHS; C. Berg and P. Prorok of PLCO; T. Riley of Information Management Services Inc.; B. O'Brien of Westat Inc; B. Kopp and W. Shao of SAIC-Frederick; investigators from the Women's Health Initiative (WHI; see URLs) and the GECCO Coordinating Center. Participating studies (grant support) in the GECCO and CCFR GWAS meta-analysis are as follows: GECCO (US NIH, U01 CA137088 and R01 CA059045), DALS (US NIH, R01 CA048998), Colo2&3 (US NIH, R01 CA060987), DACHS (German Federal Ministry of Education and Research, BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, 01KH0404 and 01ER0814), HPFS (US NIH, P01 CA055075, UM1 CA167552, R01 137178 and P50 CA127003), MEC (US NIH, R37 CA054281, P01 CA033619 and R01 CA063464), NHS (US NIH, R01 137178, P50 CA127003 and P01 CA087969), OFCCR (US NIH, U01 CA074783), PMH (US NIH, R01 CA076366), PHS (US NIH, CA042182), VITAL (US NIH, K05 CA154337), WHI (US NIH, HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, HHSN271201100004C and 268200764316C) and PLCO (US NIH, Z01 CP 010200, U01 HG004446 and U01 HG 004438). CCFR is supported by the National Cancer Institute, US NIH, under RFA CA-95-011 and through cooperative agreements with members of the Colon Cancer Family Registry and principal investigators of the Australasian Colorectal Cancer Family Registry (U01 CA097735), the Familial Colorectal Neoplasia Collaborative Group (U01 CA074799; USC), the Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (U01 CA074800), the Ontario Registry for Studies of Familial Colorectal Cancer (U01 CA074783), the Seattle Colorectal Cancer Family Registry (U01 CA074794) and the University of Hawaii Colorectal Cancer Family Registry (U01 CA074806). The GWAS work was supported by a National Cancer Institute grant (U01CA122839). OFCCR was supported by a GL2 grant from the Ontario Research Fund, the Canadian Institutes of Health Research and the Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society Research Institute. B.Z. is a recipient of Senior Investigator Awards from the Ontario Institute for Cancer Research, through support from the Ontario Ministry of Economic Development and Innovation. ASTERISK was funded by a Regional Hospital Clinical Research Program (PHRC) and supported by the Regional Council of Pays de la Loire, Groupement des Entreprises Françaises dans la Lutte contre le Cancer (GEFLUC), Association Anne de Bretagne Génétique and Ligue Régionale Contre le Cancer (LRCC). PLCO data sets were accessed with approval through dbGaP (Cancer Genetic Markers of Susceptibility (CGEMS) prostate cancer scan, phs000207.v1.p1 and GWAS of Lung Cancer and Smoking, phs000093.v2.p2).

AUTHOR CONTRIBUTIONS

W.Z. conceived and directed ACCC as well as the Shanghai-Vanderbilt Colorectal Cancer Genetics Project. W.-H.J., Y.-X.Z., K.M., A.S., Y.-B.X., S.H.J., D.-H.K., U.P.

© 2013 Nature America, Inc. All rights reserved.

and G.C. directed CRC projects at Guangzhou, Aichi, Korea-NCC, Shanghai, KCPS-II, Korea-Seoul, GECCO and CCFR, respectively. B.Z., Q.C. and W.W. coordinated the project. Q.C. directed laboratory operations. J.S. performed genotyping experiments. B.Z., J.L. and W.W. performed statistical analyses. W.Z. wrote the manuscript with substantial contributions from B.Z., Q.C., J.L., X.-O.S. and R.J.D. Z.R., G.Y., B.-T.J., Z.-Z.P., F.M., Y.-T.G., J.H.O., Y.-O.A., E.J.P., H.-L.L., J.W.P., J.J., J.-Y.J. and S.H. contributed to data and biological sample collection in the original studies included in ACCC and contributed to manuscript revision. Members of GECCO and CCFR contributed to data and biological sample collection in studies included in these consortia.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2505.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Jemal, A. *et al.* Global cancer statistics. *CA Cancer J. Clin.* **61**, 69–90 (2011).
2. de la Chapelle, A. Genetic predisposition to colorectal cancer. *Nat. Rev. Cancer* **4**, 769–780 (2004).
3. Dong, L.M. *et al.* Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. *J. Am. Med. Assoc.* **299**, 2423–2436 (2008).
4. Zanke, B.W. *et al.* Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat. Genet.* **39**, 989–994 (2007).
5. Tomlinson, I. *et al.* A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat. Genet.* **39**, 984–988 (2007).
6. Broderick, P. *et al.* A genome-wide association study shows that common alleles of *SMAD7* influence colorectal cancer risk. *Nat. Genet.* **39**, 1315–1317 (2007).
7. Jaeger, E. *et al.* Common genetic variants at the *CRAC1* (*HMPS*) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat. Genet.* **40**, 26–28 (2008).
8. Tenesa, A. *et al.* Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.* **40**, 631–637 (2008).
9. Tomlinson, I.P. *et al.* A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat. Genet.* **40**, 623–630 (2008).
10. Houlston, R.S. *et al.* Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat. Genet.* **40**, 1426–1435 (2008).
11. Houlston, R.S. *et al.* Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat. Genet.* **42**, 973–977 (2010).
12. Cui, R. *et al.* Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population. *Gut* **60**, 799–805 (2011).
13. He, J. *et al.* Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. *Cancer Epidemiol. Biomarkers Prev.* **20**, 70–81 (2011).
14. Zheng, W. *et al.* Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.* **41**, 324–328 (2009).
15. Bei, J.X. *et al.* A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat. Genet.* **42**, 599–603 (2010).
16. Jee, S.H. *et al.* Adiponectin concentrations: a genome-wide association study. *Am. J. Hum. Genet.* **87**, 545–552 (2010).
17. Nakata, I. *et al.* Association between the *SERPING1* gene and age-related macular degeneration and polypoidal choroidal vasculopathy in Japanese. *PLoS ONE* **6**, e19108 (2011).
18. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
19. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
20. Peters, U. *et al.* Meta-analysis of new genome-wide association studies of colorectal cancer risk. *Hum. Genet.* **131**, 217–234 (2012).
21. Figueiredo, J.C. *et al.* Genotype-environment interactions in microsatellite stable/microsatellite instability–low colorectal cancer: results from a genome-wide association study. *Cancer Epidemiol. Biomarkers Prev.* **20**, 758–766 (2011).
22. Musgrove, E.A., Caldon, C.E., Barraclough, J., Stone, A. & Sutherland, R.L. Cyclin D as a therapeutic target in cancer. *Nat. Rev. Cancer* **11**, 558–572 (2011).
23. Mermelshtein, A. *et al.* Expression of D-type cyclins in colon cancer and in cell lines from colon carcinomas. *Br. J. Cancer* **93**, 338–345 (2005).
24. Sarkar, R. *et al.* Expression of cyclin D2 is an independent predictor of the development of hepatic metastasis in colorectal cancer. *Colorectal Dis.* **12**, 316–323 (2010).
25. Gundem, G. *et al.* IntOGen: integration and data mining of multidimensional oncogenomic data. *Nat. Methods* **7**, 92–93 (2010).
26. Matys, V. *et al.* TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**, D108–D110 (2006).
27. Kolfschoten, I.G. *et al.* A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* **121**, 849–858 (2005).
28. Chen, Y. *et al.* Decreased *PITX1* homeobox gene expression in human lung cancer. *Lung Cancer* **55**, 287–294 (2007).
29. Chen, Y.N., Chen, H., Xu, Y., Zhang, X. & Luo, Y. Expression of pituitary homeobox 1 gene in human gastric carcinogenesis and its clinicopathological significance. *World J. Gastroenterol.* **14**, 292–297 (2008).
30. Lord, R.V. *et al.* Increased *CDX2* and decreased *PITX1* homeobox gene expression in Barrett's esophagus and Barrett's-associated adenocarcinoma. *Surgery* **138**, 924–931 (2005).
31. Nagel, S. *et al.* Activation of paired-homeobox gene *PITX1* by del(5)(q31) in T-cell acute lymphoblastic leukemia. *Leuk. Lymphoma* **52**, 1348–1359 (2011).
32. Watanabe, T. *et al.* Differential gene expression signatures between colorectal cancers with and without *KRAS* mutations: crosstalk between the KRAS pathway and other signalling pathways. *Eur. J. Cancer* **47**, 1946–1954 (2011).
33. Liu, D.X. & Lobie, P.E. Transcriptional activation of p53 by Pitx1. *Cell Death Differ.* **14**, 1893–1907 (2007).
34. Qi, D.L. *et al.* Identification of *PITX1* as a *TERT* suppressor gene located on human chromosome 5. *Mol. Cell Biol.* **31**, 1624–1636 (2011).
35. Knösel, T. *et al.* Loss of desmocollin 1-3 and homeobox genes *PITX1* and *CDX2* are associated with tumor progression and survival in colorectal carcinoma. *Int. J. Colorectal Dis.* **27**, 1391–1399 (2012).
36. Zeller, T. *et al.* Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5**, e10693 (2010).

© 2013 Nature America, Inc. All rights reserved.

ONLINE METHODS

Study populations. After quality control filtering, 7,456 cases and 11,671 controls from 10 studies were included in the consortium (**Supplementary Table 2**). Detailed descriptions of participating studies and demographic characteristics of study participants are provided in the **Supplementary Note**. Briefly, the consortium included 10,730 Chinese participants, 5,544 Korean participants and 2,853 Japanese participants. Chinese participants were from five studies: the Shanghai Study 1 (Shanghai-1, n = 3,102), the Shanghai Study 2 (Shanghai-2, n = 485), the Guangzhou Study 1 (Guangzhou-1, n = 1,613), the Guangzhou Study 2 (Guangzhou-2, n = 2,892) and the Guangzhou Study 3 (Guangzhou-3, n = 2,638). Korean participants were from three studies: the Korean Cancer Prevention Study-II (KCPS-II, n = 1,301), the Seoul Study (n = 1,522) and the Korea–National Cancer Center (Korea-NCC) Study (n = 2,721). Japanese participants were from two studies: the Aichi Study 1 (Aichi-1, n = 1,346) and the Aichi Study 2 (Aichi-2, n = 1,507). We also evaluated associations for the top 4 SNPs using data from 11,870 CRC cases and 14,190 controls of European ancestry included in GECCO and CCFR, which included 14 studies from the United States, Europe, Canada and Australia[4,20,21]. Approval was granted from the relevant institutional review boards at all study sites, and all included participants gave informed consent.

Genotyping and quality control procedures. Detailed descriptions of genotyping and quality control procedures as well as design of plates and control samples are given in the **Supplementary Note**. Briefly, in stage 1, 481 cases and 2,632 controls from Shanghai-1 were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0 as described previously[14]. The average concordance percentage of quality control samples was 99.7%, with a median value of 100% in Shanghai-1 (refs. 14,37,38). Stage 1 genotyping for 296 cases and 257 controls in Shanghai-2 was performed using Illumina HumanOmniExpress BeadChips. The same method was used to genotype cases from the Guangzhou-1 (n= 694) and Aichi-1 (n = 497) studies in stage 1. The positive quality control samples in these studies had an average concordance percentage of 99.41% and a median value of 99.97%. Cases and controls in KCPS-II were genotyped using the Affymetrix Genome-Wide Human SNP Array 5.0 (ref. 16). Controls for the Guangzhou-1 and Aichi-1 studies were genotyped previously using the Illumina Human610-Quad BeadChip[15] and Illumina Infinium HumanHap610 BeadChip[17] platforms, respectively. Details of quality control procedures for these samples have been described previously[15–17]. We excluded from the analysis samples that were genetically identical or duplicated, had a genotype-determined sex that was inconsistent with self-reported data, had unclear population structure, had close relatives with a PI-HAT estimate greater than 0.25 or had a call rate of <95%. Within each study, SNPs were excluded if (i) MAF was <5%, (ii) the call rate was <95%; (iii) the genotyping concordance percentage was <95% in quality control samples; (iv) the P value for Hardy-Weinberg equilibrium was $<1.0 \times 10^{-5}$ in controls; or (v) SNPs were not on the 22 autosomes. The final numbers of cases, controls and SNPs remaining for analysis in each participating study are presented in **Supplementary Table 1**.

Genotyping for stage 2 was completed using the iPLEX Sequenom MassARRAY platform as described previously[14,39]. With the exception of samples from the Guangzhou-3 study, which were genotyped at Fudan University (Shanghai), all other samples were genotyped at the Vanderbilt Molecular Epidemiology Laboratory. The average concordance percentage of the genotyping data for positive control samples was >99% with a median value of 100% for each of the five studies. SNPs were excluded from the analysis if (i) the call rate was <95%, (ii) the genotyping concordance percentage was <95% in control samples, (iii) there was an unclear genotype call or (iv) the P value for Hardy-Weinberg equilibrium was $<7.8 \times 10^{-4}$. The numbers of SNPs remaining for analysis in each participating study in stage 2 are presented in the **Supplementary Note**.

Genotyping for samples included in the GECCO and CCFR GWAS was conducted using Illumina BeadChip arrays, with the exception of the Ontario Familial Colorectal Cancer Registry study, for which Affymetrix arrays were used[4,20,21]. Details of the quality control procedures for these samples are presented in the **Supplementary Note**.

SNP selection for replication. SNPs were selected for stage 2 replication if (i) data were available in each of the five stage 1 studies; (ii) MAF was >5% in each stage 1 study; (iii) no heterogeneity was detected across the five studies included in stage 1 (P_{het} > 0.05 and I^2 < 25%); (iv) there was no LD (r^2 < 0.2) with any known risk variant reported from previous GWAS; (v) there was no LD (r^2 < 0.2) with the other SNPs identified in this study; (vi) there was high imputation quality in each of the five studies (RSQ > 0.5); and (vii) P < 0.01 in combined analysis of all stage 1 studies.

Evaluation of population structure. We evaluated population structure in each of the five participating studies included in stage 1 by using principal-components analysis (PCA). Genotyping data for uncorrelated genome-wide SNPs were pooled with data from HapMap to generate the first ten principal components using EIGENSTRAT software[40] (see URLs). The first two principal components for each sample were plotted using R (see URLs). We identified and excluded one participant of KCPS-II who was more than 6 s.d. away from the means of principal components 1 and 2 (**Supplementary Fig. 1**). The remaining 7,847 samples showed clear east Asian origin, and these samples were included in the final genome-wide association analysis. Cases and controls in each of the five studies were in the same cluster as HapMap Asian samples. The estimated inflation factor λ ranged from 1.02 to 1.04 in these studies after adjusting for age, sex and the first ten principal components, with a λ of 1.01 for combined stage 1 data (**Supplementary Fig. 2** and **Supplementary Table 1**).

Imputation. We used the MaCH 1.0 program[18] (see URLs) to impute genotypes for autosomal SNPs that were present in HapMap Phase 2 release 22 separately for each of the five studies included in stage 1. Genotype data from the 90 Asian subjects from HapMap were used as the reference. For Guangzhou-1 and Aichi-1, cases and controls were genotyped using different platforms. To improve imputation quality[41], we identified SNPs for which data were available in both cases and controls (250,612 SNPs in Guangzhou-1 and 232,426 SNPs in Aichi-1) and used them to impute genotyping data. A total of 1,636,380 genotyped SNPs or imputed SNPs with high imputation quality (RSQ > 0.50) in all five studies were tested for association with CRC. To directly evaluate the imputation quality for the top four SNPs identified in our study, we genotyped them in approximately 2,500 samples included in stage 1. The agreement of genotype calls derived from direct genotyping and imputation was very high, with mean concordance rates of 98.05%, 95.61%, 99.84% and 97.90% for rs647161, rs10774214, rs2423279 and rs1665650, respectively (**Supplementary Table 7**).

Statistical analyses. Dosage data for genotyped and imputed SNPs for participants in each stage 1 study were analyzed using the program mach2dat[18] (see URLs). We coded 0, 1 or 2 copies of the effect allele as the dosage for genotyped SNPs, and, for imputed SNPs, we used the expected number of copies of the effect allele as the dosage score. This approach has been shown to give unbiased estimates in meta-analyses[42]. Associations between SNPs and CRC risk were assessed using ORs and 95% CIs derived from logistic regression models. ORs were estimated on the basis of the log-additive model and adjusted for age, sex and the first ten principal components. PLINK version 1.07 (see URLs) also was used to analyze genotype data[43] and yielded results virtually identical to those derived from dosage data using mach2dat[18]. Meta-analyses were performed using the inverse-variance method, assuming a fixed-effects model, and calculations were implemented in the METAL package[19] (see URLs).

Similar to stage 1, we used logistic regression models to derive ORs and 95% CIs for the 64 selected SNPs in stage 2, assuming a log-additive model with adjustment for age and sex. We performed joint analyses to generate summary results for combined samples from all studies, with additional adjustment for study site. We also conducted stratification analysis for the top four SNPs by population ancestry (Chinese, Korean or Japanese) and by sex. We used Cochran's Q statistic to test for heterogeneity[44] and the I^2 statistic to quantify heterogeneity[45] across studies as described elsewhere in detail[46]. Analyses for stage 2, as well as combined stage 1 and 2 data, were conducted using SAS, version 9.2 (see URLs), with the use of two-tailed tests. P values of $<5\times10^{-8}$ in the combined analysis was considered statistically significant.

We used Haploview version 4.2 (see URLs; ref. 47) to generate a genome-wide Manhattan plot for results from the stage 1 meta-analysis. Forest plots

© 2013 Nature America, Inc. All rights reserved.

and quantile-quantile plots were drawn using R. We drew regional association plots using the website-based tool LocusZoom, version 1.1 (see URLs; ref. 48). LD plots were generated using Haploview[47] and the UCSC Genome Browser (see URLs).

37. Long, J. *et al.* Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer. *PLoS Genet.* **8**, e1002532 (2012).
38. Shu, X.O. *et al.* Identification of new genetic risk variants for type 2 diabetes. *PLoS Genet.* **6**, pii: e1001127 (2010).
39. Zheng, W. *et al.* Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. *J. Natl. Cancer Inst.* **102**, 972–981 (2010).
40. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
41. Sinnott, J.A. & Kraft, P. Artifact due to differential error when cases and controls are imputed from different platforms. *Hum. Genet.* **131**, 111–119 (2012).
42. Jiao, S., Hsu, L., Hutter, C.M. & Peters, U. The use of imputed values in the meta-analysis of genome-wide association studies. *Genet. Epidemiol.* **35**, 597–605 (2011).
43. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
44. Lau, J., Ioannidis, J.P. & Schmid, C.H. Quantitative synthesis in systematic reviews. *Ann. Intern. Med.* **127**, 820–826 (1997).
45. Higgins, J.P. & Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
46. Zhang, B., Beeghly-Fadiel, A., Long, J. & Zheng, W. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncol.* **12**, 477–488 (2011).
47. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
48. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).

© 2013 Nature America, Inc. All rights reserved.

doi:10.1038/ng.2505

ARTHRITIS & RHEUMATISM
Vol. 65, No. 2, February 2013, pp 472–480
DOI 10.1002/art.37777
© 2013, American College of Rheumatology

PLD4 as a Novel Susceptibility Gene for Systemic Sclerosis in a Japanese Population

Chikashi Terao,[1] Koichiro Ohmura,[1] Yasushi Kawaguchi,[2] Tetsuya Nishimoto,[3] Aya Kawasaki,[4] Kazuhiko Takehara,[5] Hiroshi Furukawa,[6] Yuta Kochi,[7] Yuko Ota,[2] Katsunori Ikari,[2] Shinichi Sato,[8] Shigeto Tohma,[6] Ryo Yamada,[1] Kazuhiko Yamamoto,[8] Michiaki Kubo,[7] Hisashi Yamanaka,[2] Masataka Kuwana,[3] Naoyuki Tsuchiya,[4] Fumihiko Matsuda,[1] and Tsuneyo Mimori[1]

Objective. **Systemic sclerosis (SSc) is an autoimmune disease for which multiple susceptibility genes have been reported. Genome-wide association studies have shown that large numbers of susceptibility genes are shared among autoimmune diseases. Recently, our group identified 9 novel susceptibility genes associated with rheumatoid arthritis (RA) in a Japanese population. The aim of this study was to elucidate whether the 18 genes that displayed associations or suggestive associations for RA in our previous study are associated with SSc in Japanese.**

Methods. **We performed an association study that included 415 patients with SSc and 16,891 control subjects, followed by a replication study that included 315 patients and 21,054 control subjects. The 18 markers reported to display association with RA were analyzed for their associations with SSc in the first study, and 5 markers were further analyzed in the replication study. The inverse variance method was used to evaluate the associations of these markers with SSc in a combined study.**

Results. **In the phospholipase D4 gene (*PLD4*), rs2841277 displayed a significant association with SSc in Japanese patients ($P = 0.00017$). We observed that rs2841280 in exon 2 of *PLD4* was in strong linkage disequilibrium with rs2841277 and introduced an amino acid alteration. We also observed associations between SSc and rs6932056 in *TNFAIP3* and rs2280381 in *IRF8* ($P = 0.0000095$ and $P = 0.0030$, respectively), both of which displayed associations with SSc in a European population.**

Conclusion. **We determined that *PLD4* is a novel susceptibility gene for SSc in Japanese, thus confirming the involvement of *PLD4* in autoimmunity. Associations between SSc and *TNFAIP3* or *IRF8* were also detected in our Japanese population. SSc and RA appear to share relatively large proportions of their genetic backgrounds.**

Systemic sclerosis (SSc) is a connective tissue disease that affects 7–489 individuals per million worldwide and is characterized by the excess production of extracellular matrix molecules and fibrosis (1). Patients with SSc display skin sclerosis, obliterative microvasculopathy such as Raynaud's phenomenon, and multiorgan involvement. Severe complications of SSc sometimes develop, including interstitial lung disease, pulmonary hypertension, and renal crisis. These severe symptoms

Supported by grants-in-aid from the Ministry of Health, Labor, and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, by research grants from the Japan Rheumatism Foundation, the Waksman Foundation, and the Mitsubishi Pharma Research Foundation, and by the Genetics and Allied Research in Rheumatic Diseases Networking consortium.

[1]Chikashi Terao, MD, PhD, Koichiro Ohmura, MD, PhD, Ryo Yamada, MD, PhD, Fumihiko Matsuda, PhD, Tsuneyo Mimori, MD, PhD: Kyoto University, Kyoto, Japan; [2]Yasushi Kawaguchi, MD, PhD, Yuko Ota, MD, Katsunori Ikari, MD, PhD, Hisashi Yamanaka, MD, PhD: Tokyo Women's Medical University, Tokyo, Japan; [3]Tetsuya Nishimoto, PhD, Masataka Kuwana, MD, PhD: Keio University, Tokyo, Japan; [4]Aya Kawasaki, PhD, Naoyuki Tsuchiya, MD, PhD: University of Tsukuba, Tsukuba, Japan; [5]Kazuhiko Takehara, MD, PhD: Kanazawa University, Kanazawa, Japan; [6]Hiroshi Furukawa, MD, PhD, Shigeto Tohma, MD: Sagamihara National Hospital, National Hospital Organization, Sagamihara, Japan; [7]Yuta Kochi, MD, PhD, Michiaki Kubo, MD, PhD: RIKEN, Yokohama, Japan; [8]Shinichi Sato, MD, PhD, Kazuhiko Yamamoto, MD, PhD: University of Tokyo, Tokyo, Japan.

Address correspondence to Chikashi Terao, MD, PhD, or Koichiro Ohmura, MD, PhD, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan. E-mail: a0001101@kuhp.kyoto-u.ac.jp or ohmurako@kuhp.kyoto-u.ac.jp.

Submitted for publication June 4, 2012; accepted in revised form October 23, 2012.

and complications of SSc result in a poor prognosis and a shortened lifespan (2,3). No effective method for preventing or curing SSc has been established (4).

It is well known that SSc has genetic components (5); for example, a US study revealed that the incidence of SSc was much higher among the families of patients with SSc compared with the general population (6). Recent technologic developments enabled the use of genome-wide association studies (GWAS) to identify novel susceptibility loci for autoimmune diseases (7). GWAS of European patients with SSc revealed that *CD247* (8), HLA (8), *TNIP1*, *PSORS1C1*, and *RHOB* (9) are susceptibility loci for SSc. In addition, another GWAS identified associations between *IRF8*, *GRB10*, and *SOX5* and limited cutaneous SSc (lcSSc) in a European population (10). Furthermore, studies adopting a candidate gene approach based on subjecting genes to functional inference analysis led to the identification of *STAT4* (11), *IRF5* (12), *TBX21* (13), *NLRP1* (14), *TNFSF4* (15), *CD226* (16), *BLK* (17), and *TNFAIP3* (18) as novel susceptibility genes for SSc in Europeans. SSc association studies in Japanese populations confirmed that *STAT4* (19), *IRF5* (20), and *BLK* (21) are associated with SSc and identified *UBE2L3* as a susceptibility gene for diffuse cutaneous SSc (dcSSc) (22). An association between HLA and SSc was also detected in Asians (23). These findings suggest a clear overlap in the genetic background of SSc between different populations.

It is well known that susceptibility genes are shared by various autoimmune diseases (24). In fact, HLA (25), *STAT4* (26), and *TNFAIP3* (27,28), which are susceptibility genes for SSc, have also been reported to be associated with rheumatoid arthritis (RA). In addition, *PTPN22*, which was shown to be strongly associated with RA in a European population (29), showed a suggestive association with SSc in Europeans (30). The sharing of these susceptibility genes between RA and SSc raises the possibility that newly identified susceptibility genes for RA could also be susceptibility genes for SSc. Recently, a large Japanese consortium, the Genetic and Allied research in Rheumatic diseases Networking consortium, identified 9 novel susceptibility genes and 6 candidate susceptibility genes for RA using a meta-analysis of GWAS and replication studies (31). Four other genes, namely, HLA, *PADI4*, *CCR6*, and *TNFAIP3*, were also confirmed to display associations with RA. Here, we performed a 2-stage association study of Japanese patients with SSc, in which we genotyped these genes as candidate susceptibility loci.

PATIENTS AND METHODS

Study subjects. DNA samples were obtained from 415 patients with SSc at Kyoto University Hospital and Tokyo Women's Medical University; these samples comprised the first set. Independent DNA samples were obtained from 315 patients with SSc at Keio University Hospital, Sagamihara National Hospital, and Kanazawa University Hospital; these samples were used as the replication set. All patients were Japanese, all had a diagnosis of SSc as determined by a rheumatologist, and all fulfilled the 1980 American College of Rheumatology classification criteria for SSc (32). The patients with SSc for whom clinical information was available were classified as having lcSSc or dcSSc, according to the definitions developed by LeRoy et al (33). The control samples were described in detail in our previous study (31). The current study was approved by the local ethics committees at each institution, and written informed consent was obtained from all subjects. The basic characteristics of the study subjects are shown in Table 1.

Genotyping. The 9 novel susceptibility markers, 6 potentially associated markers, and 4 confirmed markers of RA that were identified in our previous study in a Japanese population (31) were chosen as candidate susceptibility markers for SSc in Japanese. Eighteen of the 19 markers (HLA was excluded; see Results), none of which had previously been reported to be associated with SSc in Japanese individuals, were genotyped in the current study. The 5 candidate markers in the first set that showed associations with P values less than 0.1 were further genotyped in the replication study. Single-nucleotide polymorphisms (SNPs) rs2841280 and rs894037 were chosen as candidate causative variants in the phospholipase D4 gene (*PLD4*) region. Because rs894037 was shown to be monomorphic in Japanese, rs2841280 was genotyped in 334 control subjects, in addition to all patients, for imputation reference. The patients in the first and replication studies were genotyped at Kyoto University or Tokyo Women's Medical University and at Keio University or University of Tsukuba, respectively, using TaqMan assays (Applied Biosystems). The genotyping methods in control subjects were described in detail in our previous study (31).

Briefly, control genotypes in the first set were imputed based on the genome-scanning data, using mach2dat software with HapMap Phase II East Asian Populations as reference. The control genotypes for the replication study were extracted from genome-scanning data for the markers included on Illumina HumanHap610 Quad BeadChips. The genotypes for rs6932056 (which is not included in the array) were imputed based on the genome-scanning data, using mach2dat software with HapMap Phase II East Asian Populations as reference, and were used as control data for the replication set. The genotypes for rs2841280 (which is not included in the HapMap data or the array) were also imputed in control subjects, based on the genome-scanning data, using mach2dat software. Genotyping data for the 334 control subjects as determined by TaqMan assay in combination with genome-scanning data were used as reference.

Statistical analysis. The associations between the genotyped markers and SSc were analyzed using a Cochran-Armitage trend test in both the first and replication studies. Subanalyses were performed by comparing the genotypes of