Table 3. PSA recurrence-free survival in Cox regression analysis

	PSA recurrence-free survival univariate analysis		*P*-value	PSA recurrence-free survival multivariate analysis		*P*-value
	Hazard ratio	95% CI		Hazard ratio	95% CI	
Age (years)						
<70	1		0.66	1		0.58
≥70	1.19	0.54–2.62		0.79	0.34–1.82	
p32 expression						
Weak~moderate	1		0.001†	1		0.01†
Strong	5.43	2.04–14.41		4.10	1.39–12.07	
Preoperative PSA (ng/mL)						
<10.0	1		<0.01†	1		0.001†
≥10.0	4.58	1.99–10.55		4.56	1.86–11.20	
Gleason score						
≤6	1			1		
7	1.50	0.49–4.58	0.46	0.88	0.27–2.83	0.83
≥8	7.75	2.32–25.84	0.001†	3.22	0.87–11.89	0.08
Pathological stage						
pT2	1		0.002†	1		0.60
pT3-4	3.45	1.593–7.473		1.27	0.51–3.17	
Surgical margin						
Negative	1		0.02†	1		0.51
Positive	2.48	1.10–5.59		1.35	0.54–3.40	

†Statistically significant. CI, confidence interval; PSA, prostate-specific antigen.

the localization of p32 in prostate cancer cell lines. In the immunocytochemistry, p32 was granularly stained with antibodies in the PC3, LNCaP and RWPE-1 cells (Fig. 3B). The p32 staining was completely co-localized with the mitochondria visualized with a MitoTracker Red dye, suggesting that the expressed p32 is exclusively localized at mitochondria. Androgen deprivation did not lead nuclear translocation of p32 in these cells (data not shown). We then examined the intracellular localization of p32 in the prostate cancer cell line by western blotting. The cells were separated into nuclear extract and cytosolic fractions. B23 and TFAM were used as markers for the nuclei and cytosolic fraction, respectively. In all prostate cell lines, p32, as well as TFAM, was found mostly in the cytosolic fraction (Fig. 3C). These results suggest that p32 is localized at mitochondria but not at nuclear in prostate cancer cell lines.

Androgen receptor (AR) signaling plays a critical role in prostate cancer. We investigated p32 expression on prostate cancer cell lines after androgen deprivation. We observed reduced expression of p32 in androgen-sensitive LNCaP cells, but not insensitive PC3 and Du145 cells (Fig. 3D and Supporting Information Fig. S2). These results suggest that p32 expression might be regulated by AR signaling.

Involvement of p32 in prostate cancer cell proliferation. To examine the role of p32 in prostate cancer cell proliferation, the cell growth curve was determined by counting cells after siRNA treatment. As shown in Figure 4, a significant inhibitory effect on PC3 cell proliferation was observed in p32 siRNA-treated cells compared with the control cells that were not transfected with siRNA or cells treated with control siRNA. In contrast, siRNA-p32 had no inhibitory effect on RWPE-1 cell proliferation. The expression of p32 protein was significantly reduced in RWPE-1, PC3 and LNCaP cells by the p32 siRNA compared with the control siRNA treatments (Fig. 5A). β-actin was not affected by the p32 knockdown for 72 h in RWPE-1, PC3 and LNCaP cells (Fig. 5A). These results suggest that knockdown of p32 specifically inhibits growth of prostate cancer cells PC3 but not non-cancerous prostate epithelial cells.

Cell cycle arrest by p32 siRNA mediated knockdown. To gain insight into the molecular roles of p32, we tested whether p32 siRNA could affect cell cycle progression by using flow cytometry. p32 knockdown-induced growth suppression was associated with cell cycle arrest in PC3 cells. As shown in Figure 5B, compared with the control cells, the percentage of cells in the G1 phase (29.8%) decreased whereas the percentage in the S phase (63.3%) increased significantly after the p32-siRNA treatment in PC3 cells. On the other hand, statistically significant change was not observed in the noncancerous RWPE-1 cells. PC3 cells with downregulated p32 expression left out the S phase more slowly than the noncancerous cells. These results show that p32 depletion inhibits PC3 cell proliferation due to the blocking of progression from the G1 to S phases.

Downregulation of cyclin D1 and induction of p21 in p32 siRNA-mediated knockdown. To confirm the molecular mechanism of the G1/S phase arrest after p32 siRNA treatment, we examined specific cell cycle regulators for the G1/S phase transition: CDK4, CDK6, cyclinD1, p27Kip1 and p21Waf1. Cells treated with or without p32 siRNA for 72 h were subjected to western blot analysis. Expression of cyclin D1 was significantly (>60%) reduced or downregulated (compared with controls), while p21Waf1 expression was upregulated in p32 siRNA-treated PC3 and LNCaP cells (Fig. 5A). In noncancerous RWPE-1 cells, expression of p21Waf1 and CDK6 was observed in p32-depleted cells, while cyclin D1 expression was not changed. In the present study, we report, for the first time, that p32 depletion resulted in significant accumulation of p21, a decrease in cyclin D1 levels, a significant inhibition of cyclinD-CDK2 activity and cell cycle arrest.

Discussion

In this experiment, we have shown that p32 is important for cancer progression based on the following evidence: (i) p32 is highly expressed in prostate tumor samples and associated with the Gleason score, PSA and pathological stage; (ii) high p32 expression is adversely related to survival; (iii) p32 is overexpressed in several prostate cancer cell lines and knocking down p32 inhibits the growth of prostate cancer cells but not noncancerous cells; and (iv) p32-mediated knockdown in PC3 induces G1/S phase arrest through increased p21Waf1 and decreased cyclin D1 expression. The increased expression of p32 in

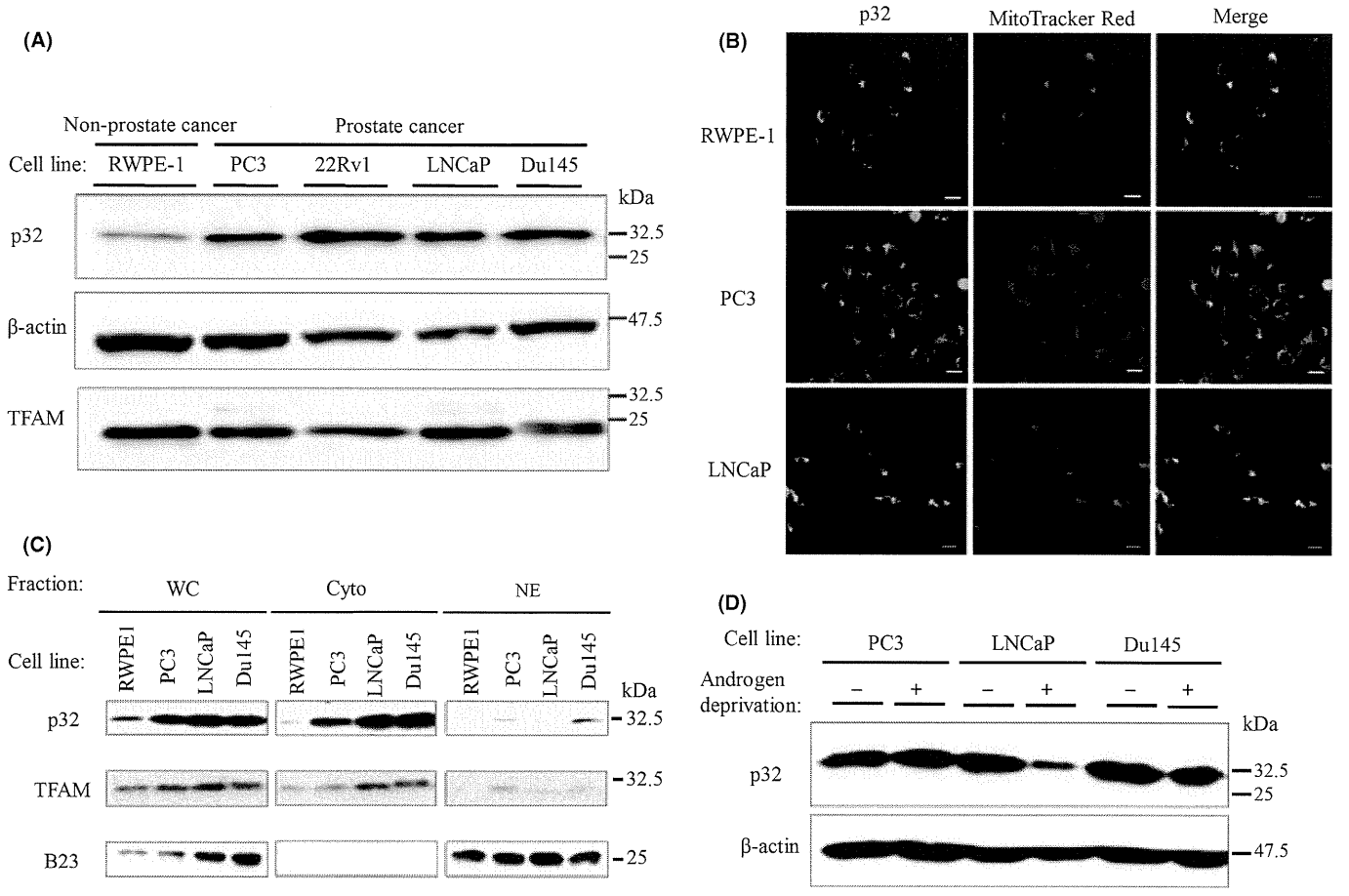

Fig. 3. Expression of p32 in prostate cancer cell lines. (A) Western blot analysis of p32 expression in prostate cancer and noncancerous cell lines. β-actin and TFAM were used as the internal control. (B) Immunocytochemistry of p32 in prostate cancer cells. Mitochondria and p32 were visualized with a mitochondria-staining dye, MitoTracker Red (middle panels), and anti-p32 antibodies (left panel), respectively. The right panels are merged. Scale bar, 20 μm. (C) Subcellular localization of p32 in prostate cancer cells. The prostate cell lysates (WC) were separated into nuclear extract (NE) and cytosolic (Cyto) fractions. The indicated proteins were detected by immunoblotting. B23 and TFAM were used as markers for nuclei and cytosolic, respectively. (D) Downregulation of p32 by androgen deprivation. Cells were treated with charcoal-treated FBS medium for 6 days. The p32 protein was analyzed by western blot.

human prostate cancer suggests that mitochondrial p32 is a key molecule of prostate tumorigenesis and p32 expression might be a potential molecular marker for the diagnosis of prostate cancer.

The current study is the first to demonstrate the relation between clinicopathological factors and patient survival with the expression of p32 in patients with prostate cancer. Chen *et al.*[28] reported that p32 might be an independent predictive factor for breast cancer prognosis and upregulation of p32 might play an important role in the metastasis of breast cancer. A possible tumor suppressor role had been previously introduced,[29,30] but the overexpression of p32 had been reported in several cancer tissue samples including breast cancer.[28,31] Ghosh *et al.*[29] have reported that p32 accumulates in inflammatory subsquamous tissue during tumor initiation and is overexpressed with progression in papillomatic and acanthotic tissues. These observations suggest that high expression of p32 might play an important role in the tumorigenesis of several cancers. It was reported that p32 overexpression was common in breast cancers but rare in some other malignant tissues, such as prostate cancer.[11,31] However, in the present study, p32 overexpression was observed in Japanese prostate cancer patients who were very strictly defined and sorted. Reasons for the discrepancy are not clear at present. Our *in vitro* studies also support that overexpression of p32 accelerates prostate tumor progression. It might be possible because of different race or different environments, such as lifestyle and nutritional factors.

The p32 knockdown cells exhibited reduced synthesis of the mitochondrial DNA-encoded OXPHOS polypeptides and were less tumorigenic *in vivo*.[32] We also observed that knockdown of p32 in MEF cells showed decreased expression of the mitochondrial-encoded component of complex IV (COX I and COX II). These observations suggest that p32 was involved in mitochondrial translation and depletion of p32 impaired the mitochondrial respiratory activity (manuscript in preparation).

In mammalian cells there are two main ways to generate energy in the form of adenosine triphosphate from glucose, oxidative phosphorylation and glycolysis. Oxidative phosphorylation occurs in mitochondria with carbon dioxide and water as end products, whereas glycolysis from glucose to lactic acid takes place in cytoplasm. In carcinogenesis, glucose will be metabolized into lactic acid instead of carbon dioxide and water, even in the presence of oxygen, which is known as the Warburg effect.[22,33–35] Recently, Fogal *et al.*[32] reported that knocking down p32 expression in human cancer cells strongly shifted their metabolism from OXPHOS to glycolysis. They suggested that tumor cells used p32 to regulate the balance between OXPHOS and glycolysis.

p32 is involved in mitochondrial respiration through the expression of mitochondrial-encoded polypeptide.[32] Knockdown

doi: 10.1111/j.1349-7006.2010.01828.x

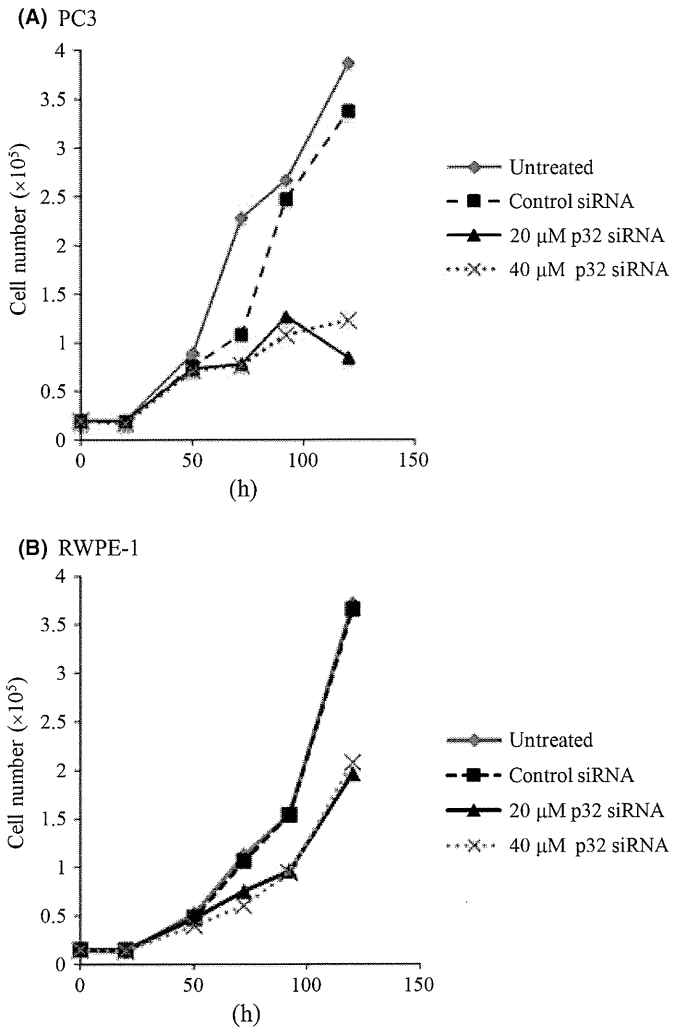

Fig. 4. RNA interference-mediated depletion of p32 inhibits PC3 cell proliferation. Cell proliferation monitored in cell lines PC3 (A) and RWPE-1 (B). The diamond, square, triangle and cross represent cells treated with nothing, 40 μM control siRNA, 20 μM p32 siRNA and 40 μM p32 siRNA, respectively. The cell number was counted at 24, 48, 72, 96 and 120 h after siRNA transfection.

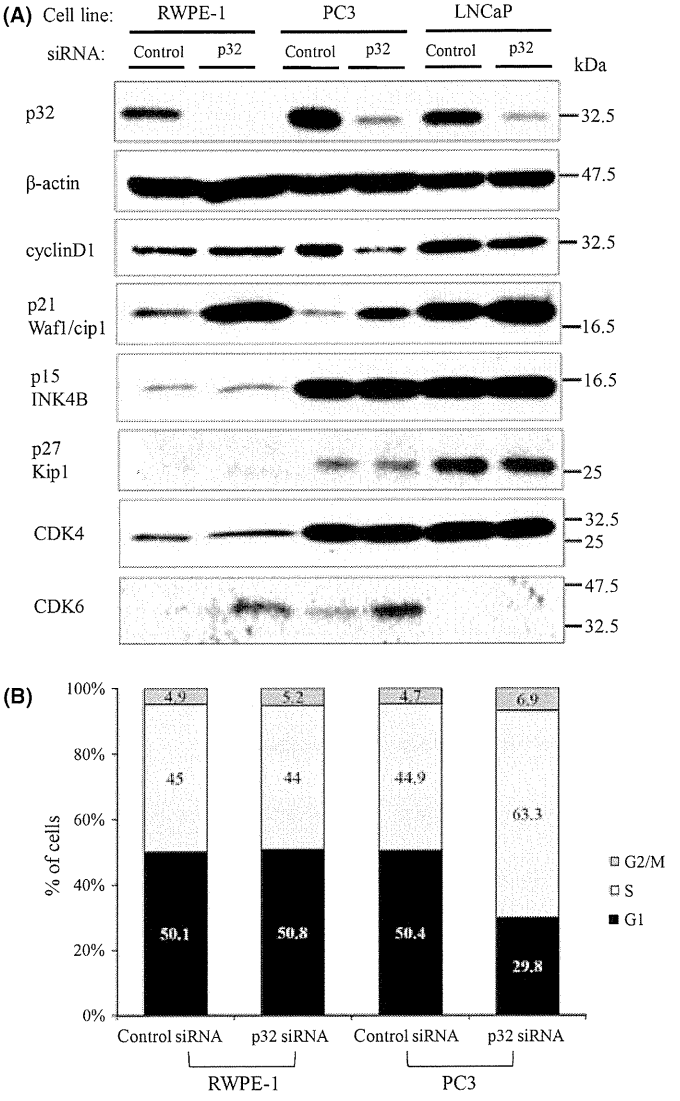

Fig. 5. Analysis of cell cycle in p32 knockdown. (A) siRNA against p32 (p32 siRNA) in PC3, LNCaP and RWPE-1. Cells were transfected with p32 or control siRNA. After subsequent culture for 72 h, the p32 protein levels were analysed by western blot. Likewise, the expressions of specific cell cycle regulators such as CDK4, CDK6, cyclinD1 and p21Waf1 were analyzed. (B) At 72 h after siRNA transfection, cells were fixed and stained with propidium iodide. Then the DNA content was measured by flow cytometry. Three separate experiments were carried out, all of which exhibited similar trends. The results of one representative experiment are shown. Cell cycle phase distributions or the number (%) of cells present at the G1, S or G2/M phases in each experimental condition were determined.

of p32 inhibits mitochondrial respiration, and this inhibition causes a collapse of the proton gradient across the mitochondrial inner membrane, thereby collapsing the mitochondrial membrane potential ($\Delta\Psi m$).[36] This inhibition can produce reactive oxygen species (ROS).[36] Excessive production of ROS gives rise to the activation of events that lead to death and survival as signaling molecules.[37] Han *et al.*[38] reported that antimycin A, an electron transport chain inhibitor in mitochondria, can produce ROS in cells and ROS might have roles in cell cycle progression via regulating cell cycle-related proteins. Therefore, it is possible that ROS levels in PC3 cells can affect the cell cycle-related proteins, resulting in G1/S phase arrest of the cell cycle. In this experiment, depletion of p32 might induce cell cycle arrest through the ROS or oxidative stress because p32 is important for the maintenance of OXPHOS.

p32 was detected in various cellular compartments other than mitochondria, such as Golgi, nucleus, cytosol and the cell surface in different cell types.[1,7,9,10,17] This evolutionarily conserved protein seems not to be expressed equally throughout the cell because it shows different localization in different cell types under different physiological conditions. It is possible to be involved in many pathological conditions, especially in cancers. We observed nuclear localization of p32 in 23 samples. Strong statistical significance was observed between p32 nuclear expression and strong p32 cytoplasmic staining ($P < 0.01$). p32 nuclear expression was significantly increased in a higher Gleason score, pathological stage and preoperative PSA (Table 2), suggesting that nuclear function of p32 might be involved in tumor progression. p32 may shed light on previous reports that it interacts with nuclear proteins like alternate mRNA splicing factor SF2,[1] lamin B receptor,[39] suggesting that nuclear p32 might be involved in the regulation of alternative splicing.

Fogal *et al.*[32] also reported that p32 knockdown breast cancer cell clones (MDA-MB-435) in nude mice was monitored.

The author demonstrated that p32 knockdown cells produced smaller tumors than the controls, suggesting that p32 is involved in tumorgenesis *in vivo*. This result suggests that p32 is an important factor for cell growth of cancer cells *in vivo*.

In the present study, we are the first to identify that p32 is essential for cell proliferation in prostate cancer, suggesting that p32 may be a novel marker of clinical progression in prostate cancer. Its unique localization in tumors and its tumor cell-specific suppression of proliferation may make p32 a useful target in the diagnosis and therapy for prostate tumors and possibly for other tumors.

Acknowledgments

The authors would like to acknowledge the technical expertise of the Support Center for Education and Research, Kyushu University. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Technology, Sports, and Culture of Japan (#22249018 and #21590337).

Disclosure Statement

No potential conflicts of interest are disclosed.

Reference

1 Muta T, Kang D, Kitajima S, Fujiwara T, Hamasaki N. p32 protein, a splicing factor 2-associated protein, is localized in mitochondrial matrix and is functionally important in maintaining oxidative phosphorylation. *J Biol Chem* 1997; **272**: 24363–70.
2 Petersen-Mahrt SK, Estmer C, Ohrmalm C, Matthews DA, Russell WC, Akusjarvi G. The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J* 1999; **18**: 1014–24.
3 Ghebrehiwet B, Silvestri L, McDevitt C. Identification of the Raji cell membrane-derived C1q inhibitor as a receptor for human C1q. Purification and immunochemical characterization. *J Exp Med* 1984; **160**: 1375–89.
4 Jiang J, Zhang Y, Krainer AR, Xu RM. Crystal structure of human p32, a doughnut-shaped acidic mitochondrial matrix protein. *Proc Natl Acad Sci U S A* 1999; **96**: 3572–7.
5 Matthews DA, Russell WC. Adenovirus core protein V interacts with p32–a protein which is associated with both the mitochondria and the nucleus. *J Gen Virol* 1998; **79**: 1677–85.
6 Dedio J, Muller-Esterl W. Kininogen binding protein p33/gC1qR is localized in the vesicular fraction of endothelial cells. *FEBS Lett* 1996; **399**: 255–8.
7 Soltys BJ, Kang D, Gupta RS. Localization of P32 protein (gC1q-R) in mitochondria and at specific extramitochondrial locations in normal tissues. *Histochem Cell Biol* 2000; **114**: 245–55.
8 Ghebrehiwet B, Lim BL, Peerschke EI, Willis AC, Reid KB. Isolation, cDNA cloning, and overexpression of a 33-kD cell surface glycoprotein that binds to the globular "heads" of C1q. *J Exp Med* 1994; **179**: 1809–21.
9 Majumdar M, Meenakshi J, Goswami SK, Datta K. Hyaluronan binding protein 1 (HABP1)/C1QBP/p32 is an endogenous substrate for MAP kinase and is translocated to the nucleus upon mitogenic stimulation. *Biochem Biophys Res Commun* 2002; **291**: 829–37.
10 Brokstad KA, Kalland KH, Russell WC, Matthews DA. Mitochondrial protein p32 can accumulate in the nucleus. *Biochem Biophys Res Commun* 2001; **281**: 1161–9.
11 Rubinstein DB, Stortchevoi A, Boosalis M *et al.* Receptor for the globular heads of C1q (gC1q-R, p33, hyaluronan-binding protein) is preferentially expressed by adenocarcinoma cells. *Int J Cancer* 2004; **110**: 741–50.
12 Storz P, Hausser A, Link G *et al.* Protein kinase C is regulated by the multifunctional chaperon protein p32. *J Biol Chem* 2000; **275**: 24601–7.
13 Kamal A, Datta K. Upregulation of hyaluronan binding protein 1 (HABP1/p32/gC1qR) is associated with Cisplatin induced apoptosis. *Apoptosis* 2006; **11**: 861–74.
14 Itahana K, Clegg HV, Zhang Y. ARF in the mitochondria: the last frontier? *Cell Cycle* 2008; **7**: 3641–6.
15 Itahana K, Zhang Y. Mitochondrial p32 is a critical mediator of ARF-induced apoptosis. *Cancer Cell* 2008; **13**: 542–53.
16 Reef S, Shifman O, Oren M, Kimchi A. The autophagic inducer smARF interacts with and is stabilized by the mitochondrial p32 protein. *Oncogene* 2007; **26**: 6677–83.
17 Sengupta A, Banerjee B, Tyagi RK, Datta K. Golgi localization and dynamics of hyaluronan binding protein 1 (HABP1/p32/C1QBP) during the cell cycle. *Cell Res* 2005; **15**: 183–6.
18 Hebert JR, Hurley TG, Olendzki BC, Teas J, Ma Y, Hampl JS. Nutritional and socioeconomic factors in relation to prostate cancer mortality: a cross-national study. *J Natl Cancer Inst* 1998; **90**: 1637–47.
19 Song Y, Oda Y, Hori M *et al.* N-myc downstream regulated gene-1/Cap43 may play an important role in malignant progression of prostate cancer, in its close association with E-cadherin. *Hum Pathol* 2010; **41**: 214–22.
20 Giovannucci E, Liu Y, Platz EA, Stampfer MJ, Willett WC. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *Int J Cancer* 2007; **121**: 1571–8.
21 Epstein JI, Allsbrook WC Jr, Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol* 2005; **29**: 1228–42.
22 Warburg O. On the origin of cancer cells. *Science* 1956; **123**: 309–14.
23 Dang CV. p32 (C1QBP) and cancer cell metabolism: is the Warburg effect a lot of hot air? *Mol Cell Biol* 2010; **30**: 1300–2.
24 Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* 2009; **9**: 153–66.
25 Polyak K, Kato JY, Solomon MJ *et al.* p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. *Genes Dev* 1994; **8**: 9–22.
26 Kanki T, Ohgaki K, Gaspari M *et al.* Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. *Mol Cell Biol* 2004; **24**: 9823–34.
27 Uchiumi T, Ohgaki K, Yagi M *et al.* ERAL1 is associated with mitochondrial ribosome and elimination of ERAL1 leads to mitochondrial dysfunction and growth retardation. *Nucleic Acids Res* 2010; **38**: 5554–68.
28 Chen YB, Jiang CT, Zhang GQ, Wang JS, Pang D. Increased expression of hyaluronic acid binding protein 1 is correlated with poor prognosis in patients with breast cancer. *J Surg Oncol* 2009; **100**: 382–6.
29 Ghosh I, Chowdhury AR, Rajeswari MR, Datta K. Differential expression of Hyaluronic Acid Binding Protein 1 (HABP1)/P32/C1QBP during progression of epidermal carcinoma. *Mol Cell Biochem* 2004; **267**: 133–9.
30 Kitazawa S, Takenaka A, Kondo T, Mizoguchi A, Kitazawa R. Protruding disordered loop of gC1qR is specifically exposed and related to antiapoptotic property in germ cell lineage. *Histochem Cell Biol* 2006; **126**: 665–77.
31 Fogal V, Zhang L, Krajewski S, Ruoslahti E. Mitochondrial/cell-surface protein p32/gC1qR as a molecular target in tumor cells and tumor stroma. *Cancer Res* 2008; **68**: 7210–18.
32 Fogal V, Richardson AD, Karmali PP, Scheffler IE, Smith JW, Ruoslahti E. Mitochondrial p32 protein is a critical regulator of tumor metabolism via maintenance of oxidative phosphorylation. *Mol Cell Biol* 2010; **30**: 1303–18.
33 Stern R, Shuster S, Neudecker BA, Formby B. Lactate stimulates fibroblast expression of hyaluronan and CD44: the Warburg effect revisited. *Exp Cell Res* 2002; **276**: 24–31.
34 Semenza GL. Tumor metabolism: cancer cells give and take lactate. *J Clin Invest* 2008; **118**: 3835–7.
35 Herrmann PC, Herrmann EC. Oxygen metabolism and a potential role for cytochrome c oxidase in the Warburg effect. *J Bioenerg Biomembr* 2007; **39**: 247–50.
36 Balaban RS, Nemoto S, Finkel T. Mitochondria, oxidants, and aging. *Cell* 2005; **120**: 483–95.
37 Boonstra J, Post JA. Molecular events associated with reactive oxygen species and cell cycle progression in mammalian cells. *Gene* 2004; **337**: 1–13.
38 Han YH, Kim SH, Kim SZ, Park WH. Antimycin A as a mitochondria damage agent induces an S phase arrest of the cell cycle in HeLa cells. *Life Sci* 2008; **83**: 346–55.
39 Marschall M, Marzi A, aus dem Siepen P *et al.* Cellular p32 recruits cytomegalovirus kinase pUL97 to redistribute the nuclear lamina. *J Biol Chem* 2005; **280**: 33357–67.

doi: 10.1111/j.1349-7006.2010.01828.x

nature genetics

Common variation in *GPC5* is associated with acquired nephrotic syndrome

Koji Okamoto[1,2], Katsushi Tokunaga[2], Kent Doi[1], Toshiro Fujita[1], Hodaka Suzuki[3], Tetsuo Katoh[3], Tsuyoshi Watanabe[3], Nao Nishida[2], Akihiko Mabuchi[2], Atsushi Takahashi[4], Michiaki Kubo[5], Shiro Maeda[6], Yusuke Nakamura[7] & Eisei Noiri[1,8]

Severe proteinuria is a defining factor of nephrotic syndrome irrespective of the etiology. Investigation of congenital nephrotic syndrome has shown that dysfunction of glomerular epithelial cells (podocytes) plays a crucial role in this disease[1]. Acquired nephrotic syndrome is also assumed to be associated with podocyte injury. Here we identify an association between variants in *GPC5*, encoding glypican-5, and acquired nephrotic syndrome through a genome-wide association study and replication analysis (*P* value under a recessive model (P_{rec}) = 6.0×10^{-11}, odds ratio = 2.54). We show that GPC5 is expressed in podocytes and that the risk genotype is associated with higher expression. We further show that podocyte-specific knockdown and systemic short interfering RNA injection confers resistance to podocyte injury in mouse models of nephrosis. This study identifies *GPC5* as a new susceptibility gene for nephrotic syndrome and implicates GPC5 as a promising therapeutic target for reducing podocyte vulnerability in glomerular disease.

Nephrotic syndrome, which is characterized by heavy proteinuria with decreased serum protein, can occur in association with various primary glomerular diseases and glomerulopathy induced by systemic diseases. Primary causes of this syndrome include minimal change disease, focal segmental glomerulosclerosis, membranous nephropathy, membranoproliferative glomerulonephritis and IgA nephropathy. Systemic diseases such as diabetes mellitus, amyloidosis and systemic lupus erythematosus also cause nephrotic syndrome. Proteinuria in nephrotic syndrome is thought to result from abnormalities in glomerular podocytes, including podocyte foot-process retraction and slit diaphragm re-organization[2–5]. Genetic abnormalities associated with congenital nephrotic syndrome have been investigated, and mutations in several podocyte-related genes, such as *NPHS1*, *NPHS2* and *ACTN4*, are known to be responsible for nephrotic syndrome[6]. However, acquired nephrotic syndrome is assumed to be a multifactorial disease caused by numerous genetic and environmental factors. Several studies have shown the contribution of podocyte injury to the pathogenesis of acquired nephrotic syndrome. Recently, genome-wide admixture scans identified a region containing *MYH9* and *APOL1* to be associated with susceptibility to focal segmental glomerulosclerosis, a subtype of adult-onset acquired nephrotic syndrome[7–9].

To identify additional susceptibility genes contributing to the common pathway of massive proteinuria in various glomerular and systemic diseases, we conducted a genome-wide association study (GWAS) using individuals with biopsy-proven nephrotic syndrome (cases). The characteristics of the cases with nephrotic syndrome are presented in **Supplementary Table 1**. We first genotyped 268,068 SNPs in 195 cases with nephrotic syndrome and 1,546 controls using a SNP genotyping array identical to that described in a previous report[10] (**Fig. 1a**). The *P* value distributions in the first test are shown in **Supplementary Figure 1**. No SNPs achieved genome-wide significance in the initial discovery screen. To overcome the type II error in the GWAS, we conducted subsequent studies. We selected 10,299 SNPs showing either an allelic or genotypic $P < 0.05$ and genotyped them in an independent sample set comprising 231 cases with nephrotic syndrome and 1,548 controls. We obtained genotype results for 7,782 SNPs using a custom-made SNP genotype assay; the *P* value distributions from the second test are shown in **Supplementary Figure 2**. Among the genotyped SNPs ($P < 0.05$; 94 SNPs), we selected the ten that showed the lowest *P* values in the combined analysis of the first and second panels (**Supplementary Table 2**). Genotyping a third independent panel of 431 cases with nephrotic syndrome and 3,371 controls by TaqMan SNP genotyping identified associations with rs16946160 in intron 2 of *GPC5* (encoding glypican-5) on chromosome 13 and with rs11086243 on chromosome 20. The combined *P* value for rs16946160 under a recessive model was $P = 6.0 \times 10^{-11}$ (**Table 1**). In contrast, the *P* value for rs11086243 under the

[1]Department of Nephrology and Endocrinology, Department of Hemodialysis and Apheresis, University Hospital, The University of Tokyo, Tokyo, Japan. [2]Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [3]Department of Nephrology, Hypertension, Diabetology, Endocrinology and Metabolism, Fukushima Medical University School of Medicine, Fukushima, Japan. [4]Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Yokohama, Japan. [5]Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Japan. [6]Laboratory for Endocrinology and Metabolism, Center for Genomic Medicine, RIKEN, Yokohama, Japan. [7]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [8]Science and Technology Research Partnership for Sustainable Development (SATREPS), Japan Science Technology (JST), Tokyo, Japan. Correspondence should be addressed to E.N. (noiri-tky@umin.ac.jp).

Received 1 November 2010; accepted 23 February 2011; published online 27 March 2011; doi:10.1038/ng.792

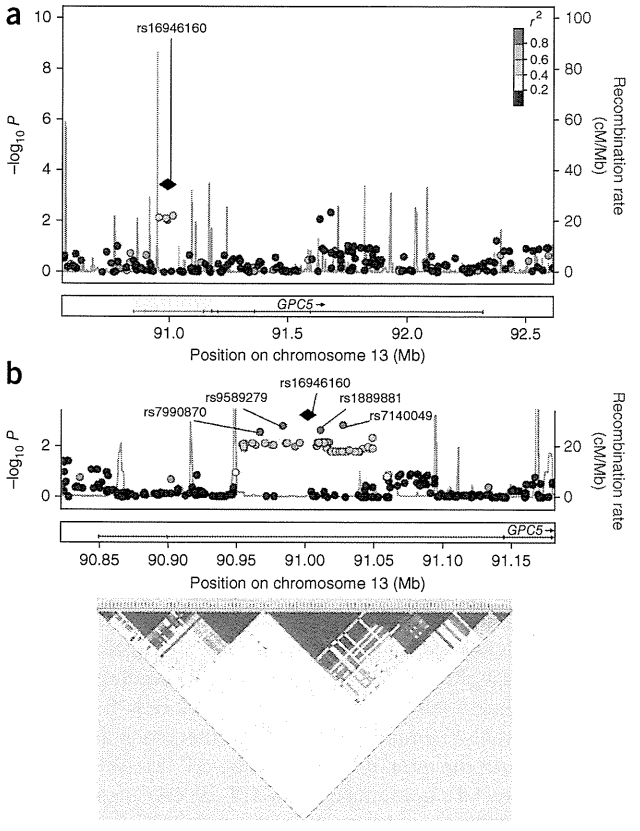

Figure 1 Association *P* value plot of the *GPC5* region in nephrotic syndrome and LD construction. (**a**) Genotyped SNPs shown with *P* values for association with nephrotic syndrome around the *GPC5* region. rs16946160 is represented as a black diamond. All other SNPs are color-coded according to the strength of LD (as measured by r^2) with rs16946160. (**b**) Evidence for association with nephrotic syndrome around the intron 2 region following imputation using HapMap II JPT and CHB reference panels. The plot includes pairwise *D'* values from the HapMap release 22 for the JPT and CHB populations.

allelic model was $P = 2.5 \times 10^{-6}$, which did not reach genome-wide significance. We performed direct sequencing of all exons and the promoter region of *GPC5* (**Supplementary Fig. 3**) as well as imputation analysis (**Fig. 1b**) and haplotype analysis in the promoter and intron 2 regions (**Supplementary Fig. 4**) in 201 cases with nephrotic syndrome and 300 healthy unrelated individuals. There was neither a single marker nor a haplotype that showed a stronger association with nephrotic syndrome than rs16946160. *GPC5* is located on the long arm of chromosome 13 (13q32)[11], and glypicans belong to the cell-surface heparan sulfate proteoglycan family. This study is the first, to our knowledge, to indicate the contribution of glypicans, a cell-surface heparan sulfate proteoglycan, to nephrotic syndrome using a genome-wide approach.

Next, we conducted functional analyses and examined the potential pathological role of GPC5 in nephrotic syndrome. We measured the mRNA level of *GPC5* extracted from peripheral blood leukocytes in 37 unrelated healthy volunteers. The *GPC5* mRNA level among individuals with the risk genotype at rs16946160 was significantly higher than those with the non-risk genotypes ($P_{rec} < 0.0001$) (**Fig. 2a**). Podocytes are located on the urinary lumen side of the glomerular basement membrane; thus, urine is a logical place to look for podocytes and for markers of podocyte injury. Podocytes and their products are detectable in urine and are altered substantially in association with disease progression in model systems and in humans[12–15]. Urinary GPC5 was presumed to be released mainly from podocytes because podocyte-specific *Gpc5* knockdown mice had minimum detectable urine Gpc5 protein levels compared to wild-type mice (**Supplementary Fig. 5a**). Therefore, we examined urinary podocyte protein levels among individuals carrying the various risk and non-risk genotypes, which could provide indirect evidence for a functional effect of the intron 2 variant of *GPC5*. Urinary GPC5 levels were higher in individuals with the risk genotype than in individuals with the non-risk genotypes ($P_{rec} = 0.0095$) (**Supplementary Fig. 5b**). We found no *GPC5* splicing variants in mRNA obtained from peripheral blood leukocytes of individuals with each genotype or in those from the renal cortex of individuals with each genotype (**Fig. 2b**). Analysis of the GWAS data, followed by imputation and linkage disequilibrium (LD) analyses, suggested that the most strongly associated SNP is rs16946160 or a closely linked SNP located in intron 2. Further analyses will be needed to determine the causal SNP influencing *GPC5* expression.

The expression of GPC5 in adult humans is high in the kidney, brain and liver[16]. We performed immunofluorescence staining, which showed that expression of GPC5 was prominent in the glomerulus and that it co-localized with podocyte markers and capillary endothelial cells in the human kidney (**Fig. 2c**). This finding is relevant to the disarrangement of podocytes and endothelial cells, which are often observed during the development of nephrotic syndrome. In podocyte-specific knockdown mice (**Supplementary Fig. 6a,b**), more than 90% of Gpc5 expression in the glomerulus disappeared (**Supplementary Fig. 6c**). These results suggest that

Table 1 Association of rs16946160 and rs11086243 genotypes with nephrotic syndrome

SNP ID	Panel	Case		Control		Allele		Genotype	
Gene locus		Freq.	GG/GA/AA	Freq.	GG/GA/AA	*P*	OR (95% CI)	GG + GA vs. AA	OR (95% CI)
	1	0.237	115/63/14	0.167	1,063/409/50	7.0×10^{-4}	1.55 (1.20–1.99)	5.8×10^{-3}	2.33 (1.25–4.35)
rs16946160	2	0.195	156/55/17	0.159	1,071/411/35	0.048	1.29 (1.00–1.65)	2.0×10^{-5}	3.44 (1.89–6.25)
GPC5	3	0.224	267/121/34	0.174	2,317/930/119	3.2×10^{-4}	1.37 (1.16–1.63)	8.7×10^{-6}	2.39 (1.61–3.55)
	Combined	0.219	538/239/65	0.168	4,451/1,750/204	2.6×10^{-7}	1.39 (1.22–1.57)	6.0×10^{-11}	2.54 (1.91–3.40)
		Freq.	CC/CT/TT	Freq.	CC/CT/TT	*P*	OR (95% CI)	CC vs. CT + TT	OR (95% CI)
	1	0.170	135/52/7	0.225	941/522/88	0.014	1.42 (1.07–1.87)	0.016	1.48 (1.07–2.05)
rs11086243	2	0.165	159/66/5	0.228	929/533/86	2.5×10^{-3}	1.49 (1.15–1.93)	8.1×10^{-3}	1.49 (1.11–2.01)
SULF2-PREX1	3	0.182	285/124/15	0.223	2,025/1,177/163	5.6×10^{-3}	1.30 (1.08–1.56)	5.1×10^{-3}	1.36 (1.10–1.68)
	Combined	0.175	579/242/27	0.225	3,895/2,232/337	2.5×10^{-6}	1.37 (1.20–1.57)	6.6×10^{-6}	1.42 (1.22–1.65)

Freq., minor allele frequency; OR, odds ratio; 95% CI, 95% confidence interval.

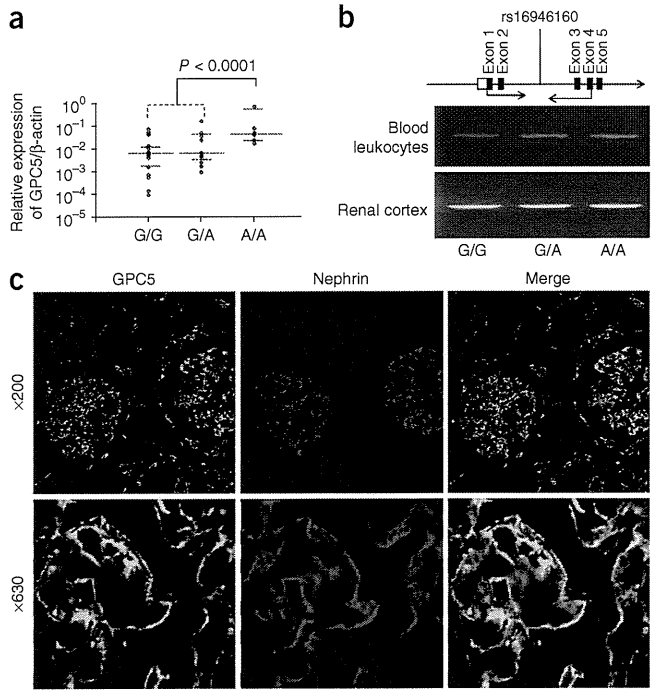

Figure 2 Association between rs16946160 genotype and *GPC5* mRNA expression and localization of expression in human kidney. (**a**) The mRNA expression level of *GPC5* in peripheral blood leukocytes from healthy individuals with different genotypes (18 G/G, 11 G/A and 8 A/A) relative to β-actin measured using real-time quantitative PCR (qPCR). Lines show quartile points. (**b**) Scheme of qPCR for screening splicing variant. No splicing variant was detected using intron-spanning RT-PCR either in blood leukocytes or the renal cortex. (**c**) Immunofluorescence staining of human kidney was visualized using confocal microscopy. GPC5 was detected as Alexa-488 fluorescence (green), and Nephrin was visualized using Alexa-633 fluorescence (blue).

GPC5 is expressed predominantly in podocytes, is localized to the cell surface membrane and is partially released according to intraglomerular circumstances. In fact, glypicans can capture and hold various endogenous and exogenous products[17]. It is possible that some unknown molecules trapped by GPC5 induce and amplify podocyte injury. Although our collection of cases with focal segmental glomerulosclerosis and IgA nephropathy involved a limited sample size, it is notable that we observed similar odds ratios for different etiologies of nephrotic syndrome, with the exception of IgA nephropathy. This latter finding suggests that different pathophysiological mechanisms of proteinuria might be involved in IgA nephropathy. Altogether, these results support the view that variation in *GPC5* is associated with a common pathway of podocyte injury and proteinuria (**Supplementary Table 3**).

The glypican family members also bind a multitude of growth factors, including fibroblast growth factor 2 (FGF2)[17]. Indeed, glypicans, which are necessary for FGF2 signal transduction[18,19], are implicated in several signal transduction cascades that regulate cell proliferation and shape[20], as studied mostly in the field of oncology. Proliferation of tumor cells is enhanced by GPC5 through FGF2 signaling[21]. In fact, FGF2–fibroblast growth factor receptor (FGFR) binding and signaling has been clearly shown for GPC1 using K562 cells[18]. Researchers from a recent study[22] similarly reported the interaction of GPC3 with FGF2 using both the human lymphoblastic cell line WI-L2 and the human colon rectal carcinoma cell line HT-29. GPC3 and GPC5 are approximately 40% homologous to each other, although GPC1 and GPC3 belong to different protein subfamilies[23]. Therefore, the interaction of glypicans with FGF2 and the FGFR pathway is universal beyond subfamilies. We examined the association between Fgf2 signal transduction and Gpc5 using rat cultured glomerular epithelial

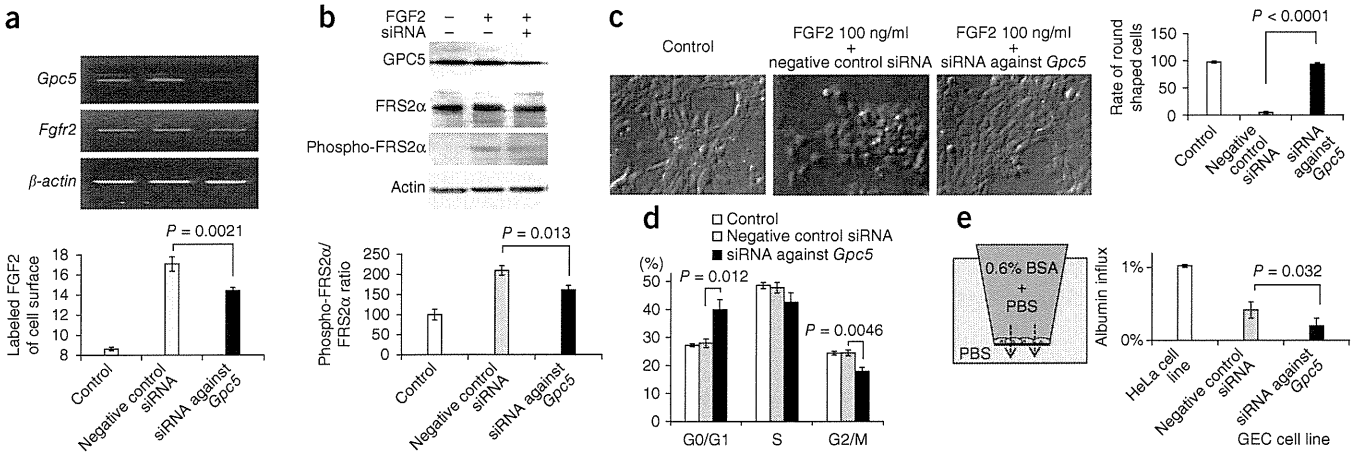

Figure 3 *In vitro* functional analysis of GPC5. (**a**) Knockdown of *Gpc5* inhibits FGF2 binding. Cells were collected and incubated with 10 nM biotinylated FGF2 ligand for 45 min at 4 °C. After staining with R-phycoerythrin-conjugated streptavidin, we counted 10,000 cells by flow cytometry ($n = 6$). Agarose gel electrophoresis of the products generated by RT-PCR using primers specific to rat *Gpc5*, *Fgfr2* and *β-actin*. Rat cultured glomerular epithelial cells (GEC) were transfected with PBS (lane 1), negative control siRNA (lane 2) and siRNA against *Gpc5* (lane 3). (**b**) Signal transduction in cells with or without knockdown in response to FGF2. We stimulated the GEC cells with 1 ng/ml FGF2 for 60 min. Cell lysates were subjected to protein blot analysis and blotted to examine the presence of Gpc5, Frs2α, phosphorylated Frs2α and actin. We calculated the phospho-Frs2α/Frs2α ratio ($n = 4$). (**c**) GEC cells were seeded in the absence (vehicle) or presence of 10 ng/ml FGF2 with or without knockdown. Morphology was analyzed using a microscope (Eclipse, Nikon) and photographed using a digital camera. The rates of round cells were determined ($n = 5$). (**d**) In the knockdown group, prevalence of the G0/G1 phase was increased and that of the G2/M phase was decreased ($n = 4$). (**e**) Schematic depiction of the paracellular permeability influx assay. Treated cells' monolayer on type 1 collagen-coated BD BioCoat cell culture inserts were incubated for 8 h. Albumin permeability across the monolayer was then determined using the protein concentration in outer chamber relative to that in the inner chamber ($n = 3$ in quadruplicate). All error bars in this figure are s.e.m.

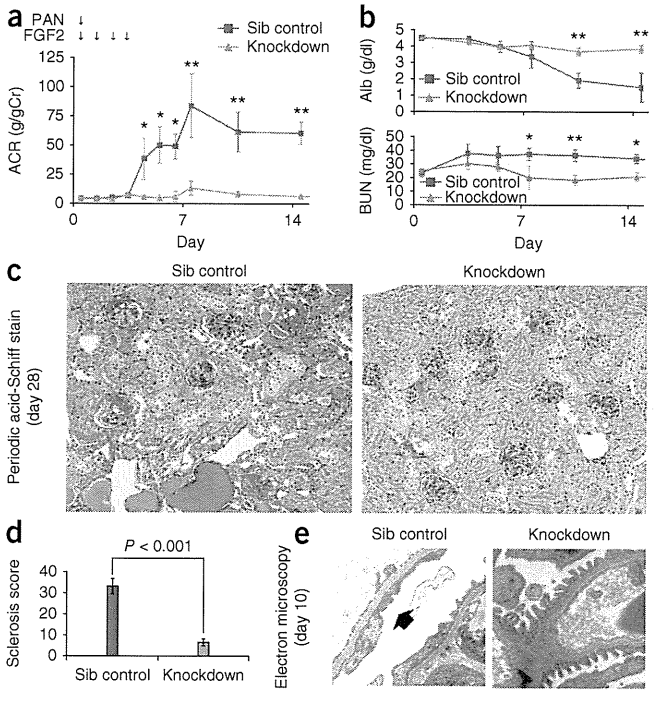

Figure 4 PAN-FGF2–induced nephrotic syndrome in wild-type sibling (sib) mice is attenuated by *Gpc5* conditional knockdown. (**a**) Albumin to creatine ratio (ACR) was significantly increased in sibling control mice after PAN-FGF2 injection after day 4 ($n = 8$; error bars are s.e.m.; $^*P < 0.05$, $^{**}P < 0.01$). (**b**) In the sibling control group, elevated blood urea nitrogen (BUN) (mg/dl) was found on days 3, 5, 7, 10 and 14, and decreased serum albumin (g/dl) was found on days 10 and 14. We found only mild changes in knockdown mice on days 7 (BUN only), 10 and 14 ($n = 8$; error bars are s.e.m.; $^*P < 0.05$, $^{**}P < 0.01$). (**c**) Periodic acid-Schiff stain of the renal cortex at day 28. Focal segmental sclerosis and cast formation by Tamm-Horsfall protein were found in the sibling control group. Almost all glomeruli were unaffected in knockdown mice that received PAN-FGF2. (**d**) The sclerosis score of glomerulus in nephrotic model on day 28. The score of knockdown mice was significantly lower than those of sib control mice ($n = 8$; error bars are s.e.m.). (**e**) Electron microscopy conducted on day 10 revealed virtually no foot-process effacement in conditional knockdown mice. Sibling control mice showed typical findings of foot-process effacement (arrow) in glomerular epithelial cells.

cells (GECs), which had previously been established by our group[24,25] (the general characteristics are presented in **Supplementary Fig. 7**). Binding of FGF2 to the cell surface was decreased in GECs in which *Gpc5* was knocked down by short interfering RNA (siRNA), although we found no difference in *Fgfr* expression (**Fig. 3a**). Protein blot analysis showed that FGF2-induced phosphorylation of Frs2α, located immediately downstream of Fgfr, was significantly reduced by *Gpc5* knockdown (**Fig. 3b**). These results suggest that GPC5 enhances FGF2 signaling in podocytes. We next examined the effect of amplified Fgf2 signaling on podocyte morphology and found that FGF2-induced dedifferentiation, as measured by reduced arborization, was improved with *Gpc5* knockdown (**Fig. 3c**). Our results show that *Gpc5* knockdown increased the prevalence of cells in the G0/G1 phase and decreased the fraction of cells in the G2/M phase in GECs (**Fig. 3d**). To evaluate the functional relevance of GPC5 in nephrotic syndrome, we examined the filtration barrier function of GECs using a paracellular permeability influx assay[26]. As shown in **Figure 3e**, *Gpc5* knockdown caused a smaller albumin influx across the cell monolayer. These *in vitro* analyses indicate that GPC5-FGF2 signaling affects podocyte differentiation and function, taking a representative role of albumin permeability.

As described in a previous report, FGF2 signaling has been presumed as a potential harmful signal in several nephrotic syndrome etiologies[27–30]. To date, pre-receptor signaling alterations of heparan sulfate proteoglycan composition, which are of crucial importance for FGF2 signaling, have not been investigated in podocytes. To further explore the role of GPC5-FGF2 signaling in nephrotic syndrome, we produced podocyte-specific *Gpc5* knockdown mice with a conditional transgenic system using a podocin-promoter–driven artificial microRNA (miRNA) against mouse *Gpc5* mRNA. The mice were fertile and developed normally up to adulthood with no major phenotypes. Because single injection of either FGF2 or puromycin aminonucleoside (PAN) to mice did not induce nephrotic-heavy proteinuria, we injected both PAN and FGF2 to induce nephrotic syndrome. Massive albuminuria occurred in BDF1-background sibling control mice and lasted for more than 10 days (**Fig. 4a**). In contrast, mild albuminuria occurred in the podocyte-specific *Gpc5* knockdown mice. In the sibling control group, we observed a decrease of serum albumin on days 10 and 14, although knockdown mice showed only mild changes. In the sibling control group, we observed an increase of blood urea nitrogen on days 7, 10 and 14 (**Fig. 4b**). Light microscopy showed that approximately 25% of glomeruli showed global sclerosis; 50% of glomeruli showed segmental sclerosis at day 28 in sibling control mice. Focal tubulointerstitial damage was also found in the sibling control mice at day 28. In contrast, 90% of glomeruli showed virtually no changes, and tubulointerstitial damage was modest in conditional knockdown mice (**Fig. 4c**). The glomerular sclerotic score in *Gpc5* knockdown mice was significantly lower than that in sibling control mice (**Fig. 4d**). Electron microscopy showed foot-process effacement in the sibling control mice at

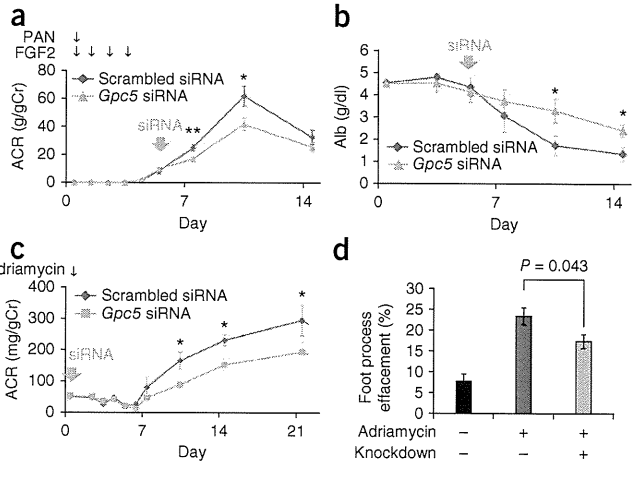

Figure 5 siRNA against *Gpc5* ameliorates PAN-FGF2–induced nephrotic syndrome and Adriamycin-induced albuminuria. (**a**) ACR (g/gCr) was significantly decreased on days 7, 10 and 14 after *Gpc5* siRNA treatment in the PAN-FGF2–induced nephrotic syndrome model (day 5) ($n = 6$; $^*P < 0.05$, $^{**}P < 0.01$). (**b**) In the scrambled-siRNA group, a decrease of serum albumin (Alb; g/dl) was found on days 7, 10 and 14. Decreased serum albumin on days 10 and 14 was significantly attenuated in *Gpc5*-siRNA–treated mice in the PAN-FGF2–induced nephrotic syndrome model (day 5) ($n = 6$; $^*P < 0.05$). (**c**) ACR (mg/gCr) was significantly decreased on days 7, 10, 14 and 21 after siRNA treatment in Adriamycin-induced albuminuria model (day 5) ($n = 6$; $^*P < 0.05$). (**d**) Electron microscopic quantification of foot processes at day 21 after Adriamycin injection in the albuminuria model ($n = 6$). All error bars in this figure are s.e.m.

day 10, although the structure of the foot process remained almost normal in conditional knockdown mice (**Fig. 4e**). These *in vivo* data strongly suggest a role for the GPC5-FGF2 pathway in nephrotic syndrome pathogenesis.

To clarify the efficacy of suppressing *Gpc5* in nephrotic syndrome as a possible therapeutic target, we conducted a systemic siRNA injection experiment after the development of proteinuria. PAN and FGF2 injection induced massive albuminuria in BDF1 mice from day 5 to day 14 (**Fig. 5a**). We injected siRNA against *Gpc5* or scrambled siRNA by the hydrodynamic method on day 5 and confirmed the efficacy of gene knockdown in the kidney on day 8 (**Supplementary Fig. 8**). Albuminuria had decreased remarkably in the *Gpc5* siRNA-treated mice when observed on days 7 and 10 (**Fig. 5a**). *Gpc5* siRNA treatment also increased serum albumin on days 10 and 14 (**Fig. 5b**). Glomerular pathological damage was significantly attenuated by *Gpc5* siRNA treatment (**Supplementary Fig. 9**). Knocking down *Gpc5* by systemic siRNA injection in mice was effective for the recovery of proteinuria and reducing pathological change, even after the development of nephrotic syndrome. One possible explanation is that the hub protein GPC5 traps FGF2 from circulation and induces FGF2 signaling in podocytes, and that the FGF2 signal results in structural alterations that cause nephrotic syndrome. In another nephrosis model induced by Adriamycin, *Gpc5* siRNA injection significantly reduced albuminuria and foot-process effacement (**Fig. 5c,d**). These data suggest that this pathway plays a crucial role in the podocyte injury observed in different glomerular diseases.

In conclusion, our GWAS, replication and functional follow-up studies identified *GPC5* as a new susceptibility gene for nephrotic syndrome that contributes to podocyte injury in glomerular diseases with different etiologies. The functional importance of GPC5 in podocyte injury has been shown *in vitro* and *in vivo*. Application of our findings could lead to new strategies for improving nephrotic syndrome management.

URLs. HapMap project, http://hapmap.ncbi.nlm.nih.gov/; NCBI RefSeq database, http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all participating doctors and staff from collaborating institutes for providing DNA samples. We also thank the technical staff of the Laboratory for Endocrinology and Metabolism at the RIKEN Center for Genomic Medicine for providing technical assistance. This study was partially supported by the BioBank Japan Project on the Implementation of Personalized Medicine, Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, to E.N. (#3023168). This study was also supported by a Grant-in-Aid for Scientific research on Priority Area 'Comprehensive Genomics' from MEXT, Japan to K.T. (#17020003).

AUTHOR CONTRIBUTIONS

K.O. performed all experimental work and data analysis and wrote the first draft of the manuscript. A.T. and M.K. obtained raw GWAS data. N.N. supported the variation screening. Y.N. planned and designed the BioBank project. K.O., S.M., K.T. and E.N. designed the experiments. T.F., S.M., A.M., K.T. and E.N. supervised the project. K.D., H.S., T.K. and T.W. collected the case data and clinical data. The manuscript was finalized by K.O. with the assistance of all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Somlo, S. & Mundel, P. Getting a foothold in nephrotic syndrome. *Nat. Genet.* **24**, 333–335 (2000).
2. Reiser, J. *et al.* Induction of B7-1 in podocytes is associated with nephrotic syndrome. *J. Clin. Invest.* **113**, 1390–1397 (2004).
3. Mathieson, P.W. Update on the podocyte. *Curr. Opin. Nephrol. Hypertens.* **18**, 206–211 (2009).
4. Johnstone, D.B. & Holzman, L.B. Clinical impact of research on the podocyte slit diaphragm. *Nat. Clin. Pract. Nephrol.* **2**, 271–282 (2006).
5. Kawachi, H. *et al.* Slit diaphragm dysfunction in proteinuric states: identification of novel therapeutic targets for nephrotic syndrome. *Clin. Exp. Nephrol.* **13**, 275–280 (2009).
6. Doi, K., Okamoto, K., Tokunaga, K., Fujita, T. & Noiri, E. Genome study of kidney disease in the age of post genome-sequencing. *Endocr. Metab. Immune Disord. Drug Targets* **8**, 173–183 (2008).
7. Kopp, J.B. *et al. MYH9* is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat. Genet.* **40**, 1175–1184 (2008).
8. Freedman, B.I. *et al.* The apolipoprotein L1 (*APOL1*) gene and nondiabetic nephropathy in African Americans. *J. Am. Soc. Nephrol.* **21**, 1422–1426 (2010).
9. Genovese, G. *et al.* Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* **329**, 841–845 (2010).
10. Unoki, H. *et al.* SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat. Genet.* **40**, 1098–1102 (2008).
11. Veugelers, M. *et al.* Characterization of glypican-5 and chromosomal localization of human *GPC5*, a new member of the glypican gene family. *Genomics* **40**, 24–30 (1997).
12. Hara, M., Yanagihara, T., Itoh, M., Matsuno, M. & Kihara, I. Immunohistochemical and urinary markers of podocyte injury. *Pediatr. Nephrol.* **12**, 43–48 (1998).
13. Hara, M. *et al.* Apical cell membranes are shed into urine from injured podocytes: a novel phenomenon of podocyte injury. *J. Am. Soc. Nephrol.* **16**, 408–416 (2005).
14. Nakamura, T. *et al.* Urinary excretion of podocytes in patients with diabetic nephropathy. *Nephrol. Dial. Transplant.* **15**, 1379–1383 (2000).
15. Wong, W. *et al.* Chronic humoral rejection of human kidney allografts is associated with MMP-2 accumulation in podocytes and its release in the urine. *Am. J. Transplant.* **10**, 2463–2471 (2010).
16. Shmueli, O. *et al.* GeneNote: whole genome expression profiles in normal human tissues. *C. R. Biol.* **326**, 1067–1072 (2003).
17. Bernfield, M. *et al.* Functions of cell surface heparan sulfate proteoglycans. *Annu. Rev. Biochem.* **68**, 729–777 (1999).
18. Steinfeld, R., Van Den Berghe, H. & David, G. Stimulation of fibroblast growth factor receptor-1 occupancy and signaling by cell surface-associated syndecans and glypican. *J. Cell Biol.* **133**, 405–416 (1996).
19. Schlessinger, J., Lax, I. & Lemmon, M. Regulation of growth factor activation by proteoglycans: what is the role of the low affinity receptors? *Cell* **83**, 357–360 (1995).
20. De Cat, B. & David, G. Developmental roles of the glypicans. *Semin. Cell Dev. Biol.* **12**, 117–125 (2001).
21. Williamson, D. *et al.* Role for amplification and expression of glypican-5 in rhabdomyosarcoma. *Cancer Res.* **67**, 57–65 (2007).
22. Song, H.H., Shi, W. & Filmus, J. OCI-5/rat glypican-3 binds to fibroblast growth factor-2 but not to insulin-like growth factor-2. *J. Biol. Chem.* **272**, 7574–7577 (1997).
23. Filmus, J., Capurro, M. & Rast, J. Glypicans. *Genome Biol.* **9**, 224 (2008).
24. Kim, B.S., Chen, J., Weinstein, T., Noiri, E. & Goligorsky, M.S. VEGF expression in hypoxia and hyperglycemia: reciprocal effect on branching angiogenesis in epithelial-endothelial co-cultures. *J. Am. Soc. Nephrol.* **13**, 2027–2036 (2002).
25. Chen, J. *et al.* VEGF-induced mobilization of caveolae and increase in permeability of endothelial cells. *Am. J. Physiol. Cell Physiol.* **282**, C1053–C1063 (2002).
26. Rico, M. *et al.* WT1-interacting protein and ZO-1 translocate into podocyte nuclei after puromycin aminonucleoside treatment. *Am. J. Physiol. Renal Physiol.* **289**, F431–F441 (2005).
27. Ray, P.E., Liu, X.H., Xu, L. & Rakusan, T. Basic fibroblast growth factor in HIV-associated hemolytic uremic syndrome. *Pediatr. Nephrol.* **13**, 586–593 (1999).
28. Ogata, S., Yorioka, N. & Kohno, N. Glucose and prednisolone alter basic fibroblast growth factor expression in peritoneal mesothelial cells and fibroblasts. *J. Am. Soc. Nephrol.* **12**, 2787–2796 (2001).
29. Floege, J. *et al.* Basic fibroblast growth factor augments podocyte injury and induces glomerulosclerosis in rats with experimental membranous nephropathy. *J. Clin. Invest.* **96**, 2809–2819 (1995).
30. Kriz, W., Hahnel, B., Rosener, S. & Elger, M. Long-term treatment of rats with FGF-2 results in focal segmental glomerulosclerosis. *Kidney Int.* **48**, 1435–1450 (1995).

ONLINE METHODS

Study participants. All participants gave written informed consent to participate in this study, which was approved in advance by the research ethics committees of Graduate School of Medicine, The University of Tokyo; BioBank Japan; and the Fukushima Medical University School of Medicine. For the GWAS, we selected case-control samples from the subjects enrolled in BioBank Japan. Nephrotic syndrome groups comprised 195 (case 1) and 231 individuals (case 2). Control groups comprised 1,546 (control 1) and 1,548 individuals (control 2) (**Supplementary Note** and **Supplementary Table 1**). No population stratification was observed in these sample sets (**Supplementary Fig. 10**).

For the validation study (third panel; see **Supplementary Note**), we obtained 431 case samples from Fukushima Medical University, The University of Tokyo and BioBank Japan following the previous criteria of discovery case 1 and case 2. We obtained 300 population-based control samples and 3,071 disease control samples from BioBank Japan. No overlap was found in cases or controls among any subject in any of the stages. All control subjects were matched with cases for ethnic origin and geographic area.

SNP genotyping and quality control. Genomic DNA from peripheral blood leukocytes was extracted using standard protocols. In the first stage, we genotyped 268,068 SNPs from autosomal chromosomes; these SNPs were selected as the tagging SNPs for the Japanese population from the JSNP21 or HapMap database 22 using high-density oligonucleotide arrays (Perlegen Sciences) as described previously[10]. For the association study, we used 205,203 SNPs having call rate >90% and no extreme departure from Hardy-Weinberg equilibrium ($P \geq 10^{-6}$). In the second screening, genotyping was performed using GeneChip SNP arrays (10,299 SNPs; Affymetrix) as reported previously[10]. In total, 7,782 SNPs having a call rate >90% (2,235 SNPs were excluded), with no extreme departure from Hardy-Weinberg equilibrium ($P \geq 0.01$ in controls; 363 SNPs were excluded), a concordance rate in the quality control samples (55 SNPs were excluded) >90% and a concordance rate in the multiplex-PCR invader assay[31] of the best 96 SNPs quality control samples (7 SNPs were excluded) >99% were used for the association study. The top ten SNPs from the first and second panels in the third panel were genotyped using the TaqMan SNP genotyping assay (Applied BioSystems). The typing results of the multiplex-PCR invader assay and the TaqMan assay were >95% and >99%, identical to that from direct sequencing.

Variation screening. We subjected 201 cases and 300 age-matched healthy controls in the third panel to direct sequencing using BigDye Terminator v3.1 Cycle Sequencing (ABI 3730 DNA analyzer, Applied Biosystems) as previously reported[32]. The average sequence success rate was 98.6%. Variations were validated with the DigiTag2 assay[33] or TaqMan assay (**Supplementary Table 4**).

Imputation analysis. Genotype imputation was performed using MACH 1.0 (ref. 34). The JPT and CHB individuals obtained from Phase II HapMap database (release 24) were used as references. We excluded the imputed SNPs with minor allele frequency < 0.01 or *R square* values <0.3.

RNA isolation and real-time PCR. We isolated total RNA from the peripheral blood leukocytes of healthy volunteers using a PAXgene Blood RNA Kit (QIAGEN) and from cultured cells using TRIzol (Invitrogen). Following DNase I treatment, we synthesized complementary DNA using oligo(dT) primers and ImProm-II reverse transcriptase (Promega). We performed real-time quantitative PCR (qPCR) using the SYBR green system (SYBR Green I PCR Master Mix, Applied BioSystems). Normalization of the relative expression of the genes of interest to the β-actin gene was performed. The qPCR primer sequences are presented in **Supplementary Table 5**. All samples were analyzed in triplicate.

Screening for alternative splicing variants. We performed RT-PCR to detect splicing variants using primers designed within exon 1 and exon 4 of *GPC5* (**Supplementary Table 5**). Blood leukocytes were obtained from healthy volunteers; renal cortexes were from diagnostic biopsy samples. Amplified products were confirmed by direct sequencing. We measured amplified products using gel electrophoresis.

Immunofluorescence staining. For human glomerular staining, we obtained kidney specimens from the Human Tissues and Biofluids Bank (Asterand). All samples were surgically dissected by tumor-related disease, not by kidney disease. We used frozen sections to perform GPC5 and Nephrin immunofluorescence double staining. Five-micrometer-thick sections were dried and fixed with 4% paraformaldehyde for 20 min at room temperature (20 ± 5 °C). We washed the slides with PBS three times and incubated them with GPC5 (R&D Systems) and Nephrin (Acris Antibodies GmbH) antibodies. For characterization of rat cultured glomerular epithelial cells (GEC)[24], antibodies of WT-1 (Santa Cruz Biotechnology), Synaptopodin (Acris Antibodies GmbH), Thy-1.1 (MyBioSource), 516 (provided by Dr. Kawachi[35]) and GSA3 (as described previously[36]) were used to characterize GEC (**Supplementary Fig. 6**). After washing the slides with PBS, we incubated them with fluorescence conjugated secondary antibody (Invitrogen) for each primary antibody for 40 min. The sections were then examined visually using confocal microscopy (LSM 510 Meta NLO imaging system, Carl Zeiss) or fluorescent microscopy (E600, Nikon).

***In vitro* knockdown.** For *in vitro* knockdown experiments, we used GECs transfected with either Stealth RNAi Negative Control Med GC (Invitrogen) or with predesigned *Gpc5* siRNA targeted against *Gpc5* (NM_001107285.1_stealth_1056, Invitrogen) using Lipofectamine 2000 (Invitrogen). BLOCK-iT Alexa Fluor Fluorescent Oligo was co-transfected and used for transfection control. We incubated the cells for 72 h to allow knockdown of *Gpc5*. The sense and antisense sequences are in **Supplementary Table 6**. We extracted cellular RNA using TRIzol (Invitrogen) and treated it with DNaseI (Roche). Single-strand complementary DNA was synthesized as described above. To determine the extent of knockdown, we quantified the mRNA (**Fig. 3a**) and protein levels (**Fig. 3b**) of Gpc5 in transfected cells.

Analysis of biotinylated FGF2 binding by flow cytometry. We biotinylated recombinant human FGF2 (generously provided by Kaken Pharmaceutical) using the Pierce EZ-Link MicroSulfo-NHS-LC biotinylation kit (Thermo Fisher Scientific). To label 25 μg of FGF2, 8 μl of a 9 mM Sulfo-NHS-LC-Biotin solution was incubated in dimethyl sulfoxide with FGF2 in a total volume of 300 μl for 1 h at room temperature. We used a G-25 column (GE Healthcare) to remove excess unbound biotin. After knockdown with siRNA against *Gpc5* by liposome system as described above, we washed the GECs twice in ice-cold PBS with 3% heat-inactivated FBS/0.02% sodium azide. Cells were incubated with 10 nM biotinylated FGF2 at 4 °C for 45 min. We then stained the cells with 5 μg/ml R-phycoerythrin–conjugated streptavidin (Jackson ImmunoResearch Laboratories) at 4 °C for 30 min in the dark. The cells were washed and resuspended in 500 μL of PBS. We measured the biotinylated FGF2 using the FACSCalibur analyzer (BD Biosciences)[37].

Signal transduction. Sub-confluent GECs were serum-starved (0.1% FBS) for 18 h and then stimulated with FGF2. For morphological analysis, cells were treated with FGF2 (100 ng/ml) for 15 min and then analyzed using an inverted microscope (Eclipse, Nikon) and photographed using a digital camera. For signal transduction analyses, cells were treated with FGF2 (1 ng/ml) for 60 min and were rinsed with ice-cold PBS and lysed in a buffer containing 0.5% Triton X-100, 50 mM Tris, pH 7.5, 150 mM sodium chloride, 1 mM phenylmethylsulfonyl fluoride, 1 mM sodium orthovanadate, 10 mM sodium fluoride, 1 μg/ml leupeptin and 1 μg/ml aprotinin. Heparitinase and chondroitinase ABC-digested lysates as previously reported[18] were boiled for 5 min and subjected to sodium dodecyl sulfate PAGE (SDS-PAGE). For protein blot analysis, proteins were transferred electrophoretically to a membrane, which were blocked in 5% skim milk. The membranes were incubated with primary antibody as follows: GPC5 antibody (R&D Systems), FRS2α antibody (Abcam), Phospho-FRS2α antibody (Cell Signaling) and Actin antibody (Santa Cruz Biotechnology). Immunoreactivity was detected using horseradish peroxidase–conjugated secondary antibody and enhanced chemiluminescence (ECL) (Amersham Pharmacia Biotech) according to the manufacturer's protocol. The signal was detected using a CCD camera system (LAS-4000mini, Fuji Photo Film).

Cell cycle assay. For cell cycle analysis, 0.5×10^5 cells were suspended in 0.5 ml of 50 μg/ml of propidium iodide solution (w/v) and 10 μg/ml RNase A (w/v) in PBS. We then analyzed cell cycle distribution using flow cytometry as

doi:10.1038/ng.792

previously described after gating the transfection control positive cells[37]. We repeated the experiment three times under identical conditions.

Albumin influx assay. We used a simple albumin influx assay to evaluate the filtration barrier function of the podocyte monolayer, following the method previously reported[26]. Briefly, we placed treated GECs into wells of a 24-well BD BioCoat cell culture insert with 0.4 μm pore (1,600,000/m^2) coated type 1 collagen (BD Co.). Cells were cultured for 18 h at 37 °C and then washed twice with PBS supplemented with 1 mM $MgCl_2$ and 1 mM $CaCl_2$ to preserve the cadherin-based junctions. We adjusted the concentration of the cultured cells so that they would be fully confluent at the time of analysis. The chambers over the monolayer of the cells on the filter were filled with RPMI 1640 (Invitrogen) containing 0.6% BSA (w/v), and the lower chambers were filled with RPMI 1640 only. After 8 h incubation in a CO_2 incubator, we measured the ratio of albumin concentration, lower chamber versus upper chamber, using the Bradford assay (Bio-Rad).

Animal experiment protocol. All experiments were conducted in accordance with the US National Institutes of Health *Guide for the Care and Use of Laboratory Animals*. In conditional knockdown experiments, we used podocyte-specific *Gpc5* knockdown mice of BDF1 background. We produced these mice using a conditional transgenic system with podocin-promoter–driven miRNA against mouse *Gpc5*. The construct of the vector contains the podocin promoter and primary miRNA sequences, which include flanking and loop sequences from an endogenous miRNA that directs the excision of the three kinds of engineered miRNA against mice *Gpc5* (Invitrogen), resulting in knockdown of *Gpc5* selectively in podocytes (12–16 weeks of age). Their wild-type siblings served as the sibling control in the conditional knockdown experiments (**Supplementary Fig. 7a**). Their genotype was determined using PCR and DNA blot; the selective knockdown of *Gpc5* in podocytes was confirmed (**Supplementary Fig. 7b** and **Supplementary Table 5**). We assigned the animals to two groups: sibling control mice and knockdown mice. To induce nephrotic syndrome in the conditional knockdown model, we injected puromycin aminonucleoside (PAN; Sigma-Aldrich) subcutaneously on day 0 (300 mg/kgBW) and injected FGF2 intravenously on days 0 and 1 (5 μg per animal).

For the siRNA treatment experiments, we used wild-type BDF1 mice 12–14 weeks of age (Japan SLC) for the PAN-FGF2–induced nephrotic syndrome model and BALB/C mice 10–12 weeks of age for the Adriamycin-induced albuminuria model. The mouse *Gpc5* siRNA and scrambled siRNA was obtained from Invitrogen as an annealed *in vivo* grade siRNA. The siRNA sequences are shown in **Supplementary Table 6**. Then, 400 μg of synthetic siRNA dissolved in 0.8 ml PBS was rapidly injected intravenously following the hydrodynamic method as reported previously[38,39]. The animals were assigned into two groups: scrambled-siRNA mice and *Gpc5*-siRNA mice. In the nephrotic model, PAN was injected subcutaneously on day 0 (500 mg/kg) and then FGF2 was injected intravenously on days 0, 1, 2 and 3 (5 μg per animal); siRNA was injected on day 5 in the siRNA-treated group (**Fig. 5a**). In the albuminuria model, Adriamycin (10 mg/kg) was injected intravenously on day 0, and siRNA was injected on day 0 in the siRNA-treated group (**Fig. 5c**).

During the experiments, urine samples were collected daily during days 1–7 and then on days 10 and 14 (and day 21 in Adriamycin model) using glass shield metabolic cages (Metabolica, Sugiyama-Gen). We killed mice from each group at the end of days 10 and 28 (and day 21 in Adriamycin model). We perfused their kidneys thoroughly with PBS and removed them for light-microscopic, electron-microscopic and immunofluorescence analysis.

Analyses of urinary albumin, urinary creatinine, serum blood urea nitrogen (BUN) and serum albumin. We determined the urinary and serum albumin levels with ELISA using a murine microalbuminuria ELISA kit (AlbuwellM, Exocell), measured the urine creatinine concentration with Nescoat VLII CRE (Alfresa Pharma) and measured the blood urea nitrogen concentration using BUN Wako B (Wako). All measurements were performed in duplicate. Albuminuria was determined as the ratio of urinary albumin to creatinine (g/gCr or mg/gCr). All procedures were performed in accordance with the manufacturer's protocols.

Renal pathology. After the overnight fixation with 10% neutral buffered formaldehyde, mice kidneys were dehydrated with alcohol and embedded into paraffin for light microscopic analysis. We counted the glomeruli showing sclerosis in periodic acid-Schiff–stained sections using the following modified scoring system, so called Sclerosis score: each was scored as 0–4 according to the percentage of glomeruli involved (0%, 0; 1–25%, 1; 26–50%, 2; 51–75%, 3; 76–100%, 4)[40]. To obtain the final score of 0–100 for the individual samples, we counted 25 glomeruli. For electron-microscopic examination, small pieces of renal cortex were fixed in 2.5% glutaraldehyde overnight, washed with 50 mM Tris-(hydroxymethyl) aminomethane-HCl buffer (pH 7.6), dehydrated in sucrose, fixed again in OsO_4 and embedded into Epon. Ultrathin sections were examined under Hitachi electron microscope (H-1000).

Statistical analysis. Subject data are presented as means ± s.d. (human) or s.e.m. (*in vitro* and animal experiments). χ^2 analysis was used to assess Hardy-Weinberg equilibrium. We analyzed the overall distribution of alleles using 2 × 2 contingency tables with statistical significance defined as $P < 0.05$. We selected the tag SNPs based on estimated haplotype construction calculated using Haploview 4.1 (Broad Institute) with Japanese genotype data from the HapMap project. Permutation-based hypothesis testing (10,000 permutations) was performed to examine the associations of estimated haplotype frequencies.

31. Ohnishi, Y. *et al.* A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46**, 471–477 (2001).
32. Okamoto, K. *et al.* Identification of *KCNJ15* as a susceptibility gene in Asian patients with type 2 diabetes mellitus. *Am. J. Hum. Genet.* **86**, 54–64 (2010).
33. Nishida, N., Tanabe, T., Takasu, M., Suyama, A. & Tokunaga, K. Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal. Biochem.* **364**, 78–85 (2007).
34. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu. Rev. Genomics Hum. Genet.* **10**, 387–406 (2009).
35. Kawachi, H., Koike, H., Kurihara, H., Sakai, T. & Shimizu, F. Cloning of rat homologue of podocin: expression in proteinuric states and in developing glomeruli. *J. Am. Soc. Nephrol.* **14**, 46–56 (2003).
36. Nosaka, K. *et al.* Permeable type I collagen membrane promotes glomerular epithelial cell growth in culture. *Kidney Int.* **43**, 470–478 (1993).
37. Lai, J.P. *et al.* Sulfatase 2 up-regulates glypican 3, promotes fibroblast growth factor signaling, and decreases survival in hepatocellular carcinoma. *Hepatology* **47**, 1211–1222 (2008).
38. Bradley, S.P. *et al.* Gene silencing in the endocrine pancreas mediated by short-interfering RNA. *Pancreas* **31**, 373–379 (2005).
39. Larson, S.D., Jackson, L.N., Chen, L.A., Rychahou, P.G. & Evers, B.M. Effectiveness of siRNA uptake in target tissues by various delivery methods. *Surgery* **142**, 262–269 (2007).
40. Saito, T., Sumithran, E., Glasgow, E.F. & Atkins, R.C. The enhancement of aminonucleoside nephrosis by the co-administration of protamine. *Kidney Int.* **32**, 691–699 (1987).

nature genetics

Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma

Vinod Kumar[1,2], Naoya Kato[3], Yuji Urabe[1], Atsushi Takahashi[2], Ryosuke Muroyama[3], Naoya Hosono[2], Motoyuki Otsuka[4], Ryosuke Tateishi[4], Masao Omata[4], Hidewaki Nakagawa[2], Kazuhiko Koike[4], Naoyuki Kamatani[2], Michiaki Kubo[2], Yusuke Nakamura[1,2] & Koichi Matsuda[1]

To identify the genetic susceptibility factor(s) for hepatitis C virus–induced hepatocellular carcinoma (HCV-induced HCC), we conducted a genome-wide association study using 432,703 autosomal SNPs in 721 individuals with HCV-induced HCC (cases) and 2,890 HCV-negative controls of Japanese origin. Eight SNPs that showed possible association ($P < 1 \times 10^{-5}$) in the genome-wide association study were further genotyped in 673 cases and 2,596 controls. We found a previously unidentified locus in the 5′ flanking region of *MICA* on 6p21.33 (rs2596542, $P_{combined} = 4.21 \times 10^{-13}$, odds ratio = 1.39) to be strongly associated with HCV-induced HCC. Subsequent analyses using individuals with chronic hepatitis C (CHC) indicated that this SNP is not associated with CHC susceptibility ($P = 0.61$) but is significantly associated with progression from CHC to HCC ($P = 3.13 \times 10^{-8}$). We also found that the risk allele of rs2596542 was associated with lower soluble MICA protein levels in individuals with HCV-induced HCC ($P = 1.38 \times 10^{-13}$).

It is estimated that more than 170 million people are infected with HCV worldwide[1]. Persistent HCV infection causes CHC and, subsequently, fatal liver diseases such as liver cirrhosis and HCC. Therefore, the treatment of HCV carriers is an issue of global importance. HCC is the third most common cause of cancer-related deaths[2], and HCV infection accounts for 30–70% of the individuals with HCC[3,4]. HCV-induced HCC is a multistep and progressive liver disease in which disease progression may be influenced by both environmental and genetic risk factors. The impact of host genetic variation on progression to CHC after HCV exposure is well documented by recent genome-wide association studies (GWAS)[5–7]. However, no comprehensive analyses have been performed to explore the genetic basis of HCV-induced HCC. Therefore, we conducted a GWAS for HCV-induced HCC.

We genotyped the DNA of 721 individuals with HCV-induced HCC and 2,890 HCV-negative controls (**Supplementary Table 1**) from BioBank Japan[8]. After the initial standard SNP quality filters, we obtained genotyping results for 432,703 SNPs for association analysis. Because progression from CHC to liver cancer is strongly affected by age and gender[3], we performed a logistic regression analysis by including age and gender as covariates at all tested loci in our analyses. The genetic inflation factor (λ) was 1.03, indicating that there is no or little population stratification (**Supplementary Fig. 1**). Although no SNPs cleared the GWAS significance threshold ($P < 5 \times 10^{-8}$) at this stage, we identified eight independent loci showing possible association ($P < 1 \times 10^{-5}$; **Supplementary Fig. 2**).

In the replication stage, 673 cases from an independent HCC cohort from the University of Tokyo and 2,596 HCV-negative controls from BioBank Japan were genotyped at these eight SNPs. We observed a significant replication of association at rs2596542 on chromosome 6p21.33 ($P = 8.62 \times 10^{-9}$, odds ratio (OR) = 1.44, 95% confidence interval (CI) 1.27–1.63; **Table 1**), whereas the remaining seven SNPs failed to replicate the association (**Supplementary Table 2**). Furthermore, the combination analysis of the GWAS and replication study data at rs2596542 revealed a highly significant association in which the frequency of the risk allele A is higher in cases ($P = 4.21 \times 10^{-13}$, OR = 1.39; **Fig. 1** and **Table 1**) after the age and gender adjustment, without any heterogeneity ($P = 0.24$) between the two stages. To further investigate the impact of rs2596542 on the complex nature of the HCV-induced HCC phenotype, we genotyped 1,730 individuals with CHC who had not developed liver cirrhosis or HCC during their recruitment. As a result, rs2596542 was found to have no association with chronic hepatitis C susceptibility ($P = 0.61$) but was significantly associated with progression from CHC to HCC ($P = 3.13 \times 10^{-8}$, OR = 1.36; **Table 2**).

Because heavy alcohol consumption (>50 g per day) as well as poor response to interferon (IFN) treatment were shown to be the major risk factors for HCC among individuals with CHC[9], we evaluated the effect of alcohol consumption as a confounding factor and found that rs2596542 remained highly significant even after adjustment for this factor (non-HCV versus HCC, OR = 1.39, $P = 1.22 \times 10^{-11}$; CHC versus HCC, OR = 1.25, $P = 2.31 \times 10^{-4}$; **Supplementary Table 3**). The major genotypes of HCV can be determined by a serotyping

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. [2]Center for Genomic Medicine, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan. [3]Unit of Disease Control Genome Medicine, The Institute of Medical Science, University of Tokyo, Tokyo, Japan. [4]Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. Correspondence should be addressed to K.M. (koichima@ims.u-tokyo.ac.jp).

Received 20 July 2010; accepted 23 March 2011; published online 17 April 2011; doi:10.1038/ng.809

Table 1 Association results of rs2596542 in the GWAS, replication stage and combined analysis

SNP	Chr. (locus)	Stage	Case RAF	Control RAF	*P*	OR (95% CI)
rs2596542 (A/G)	6 (*MICA*)	GWAS[a]	0.388	0.331	4.50×10^{-6}	1.34 (1.16–1.53)
		Replication[a]	0.413	0.331	8.62×10^{-9}	1.44 (1.27–1.63)
		Combined[a]	0.400	0.331	4.21×10^{-13}	1.39 (1.27–1.52)
				MH test	7.76×10^{-12}	1.35 (1.24–1.47)

We analyzed 1,394 cases with HCC (721 in the GWAS and 673 in the replication) and 5,486 controls (2,890 in GWAS and 2,596 in replication). Chr., chromosome; RAF, risk allele frequency (allele A); OR, odds ratio for the minor allele calculated by considering the major allele as a reference; MH, Mantel-Haenszel.
[a]*P* values and ORs are adjusted for age and gender by logistic regression analysis under an additive model.

assay that is based on the type-specific antibodies produced by the infected host[10]. A subgroup analysis for HCV serotypes or history of IFN therapy indicated that this variation is associated with HCC susceptibility independently of HCV genotypes or treatment response (**Supplementary Fig. 3**). Consistent with this result, rs1051796, which had $r^2 = 0.7$ and $D' = 0.95$ with rs2596542, was not associated with IFN response ($P = 0.89$) according to previously published data in the Japanese population[11].

rs2596542 is located within the class I major histocompatibility complex (MHC) region. The human MHC region encompasses the complex and extended linkage disequilibrium (LD) structure[12,13]. Several HLA alleles and genes within MHC region have been implicated in HCV infection or clearance or in response to treatment[14–16]. Therefore, we searched the whole 7.5-Mb extended MHC region using GWAS data to test the possibility of other associated loci. We found a moderate association peak at rs9275572 ($P = 4.99 \times 10^{-5}$), which is located between *HLA-DQA* and *HLA-DQB* loci (**Supplementary Fig. 4**). Subsequent replication and combination analyses at rs9275572 indicated a significant association with HCV-induced HCC ($P = 9.38 \times 10^{-9}$, OR = 1.30; **Supplementary Table 4**). The multiple logistic regression analysis to control for alcohol consumption along with age and gender also indicated a significant association at rs9275572 ($P = 3.21 \times 10^{-8}$, OR = 1.29; **Supplementary Table 5**). However, rs2596542 was not in high LD with rs9275572 ($D' = 0.41$, $r^2 = 0.16$), and both SNPs remained associated with HCC even after conditional analysis on each other and had small reductions in their ORs upon conditioned analysis (OR = 1.23, $P = 4.43 \times 10^{-6}$ and OR = 1.17, $P = 0.00059$, respectively; **Supplementary Table 6**). A haplotype analysis between these two markers showed four possible haplotypes, with haplotype AA showing higher risk (with OR = 1.44) compared to the major haplotype GG (**Supplementary Table 7**). However, the OR for the risk haplotype was 1.32 with $P = 2.31 \times 10^{-10}$ after comparing against all observed haplotypes in the population (**Supplementary Table 7**), which is weaker than that of rs2596542 alone (OR = 1.39, $P = 4.21 \times 10^{-13}$). Hence, the impact of rs2596542 is much stronger than the haplotype of two SNPs, suggesting that rs2596542 is a principal genetic factor in this region. We also found that rs9275572 has a moderate association with CHC susceptibility as well as progression from CHC to HCC ($P = 0.03$ and $P = 2.58 \times 10^{-5}$, OR = 1.09 and OR = 1.29, respectively; **Supplementary Table 8**). Because *HLA-DQ* and *HLA-DR* alleles were shown to be associated with viral persistence and early liver disease among Japanese individuals[16], further study will be needed to confirm whether the association at rs9275572 is because of its LD with *HLA-DQ* or *DR* alleles.

In this regard, it is interesting to note that rs9275572 had a very strong expression quantitative trait locus effect on *HLA-DQB1* ($\log_{10}$ odds (LOD) ≥ 19.48) and *HLA-DRB4* alleles (LOD ≥ 26.88)[17]. Thus, it will be important to test the functional effect of the common haplotype (AA; **Supplementary Table 7**), which tags the risk alleles at these two SNPs.

Two SNPs, rs12979860 and rs8099917, at the *IL28B* locus were reported to be associated with spontaneous clearance of HCV virus[18] and response to pegylated IFN-α and ribavirin therapy[11], respectively. However, we found no association at rs12979860 and rs8099917 in our dataset (**Supplementary Table 9**). Because we used non-HCV control subjects rather than subjects who had cleared HCV infection spontaneously, and because only about 20% of the cases with HCC had been treated with IFN, our study may not be suitable to detect associations at the *IL28B* locus. In addition, the protective C allele at rs12979860 is nearly fixed throughout east Asia, with a frequency of more than 91% in the Japanese population as compared to 67% in European Americans[6], indicating a role for other factors in spontaneous clearance.

The top associated SNP, rs2596542, is located 4.7 kb upstream of *MICA*, the MHC class I polypeptide-related sequence A gene, and 41.7 kb downstream of the *HLA-B* gene (**Supplementary Fig. 5**). The regional association plot at the rs2596542 locus, made using genotype data from the GWAS (**Fig. 1**) and imputation analysis (**Supplementary Fig. 6**), revealed that all of the modestly associated SNPs are tightly

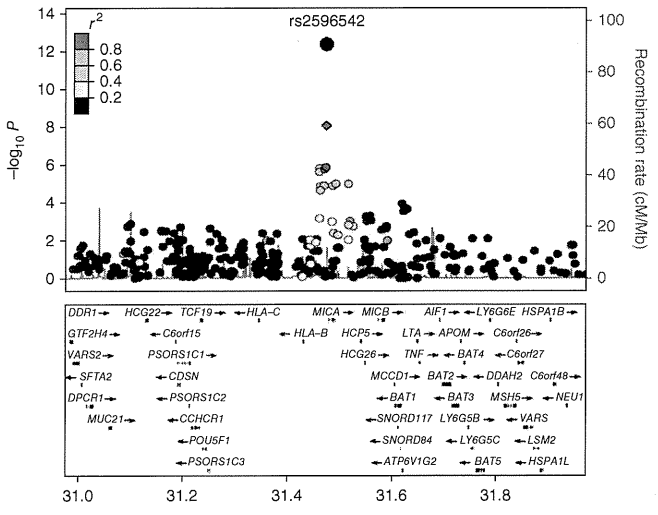

Figure 1 Regional association plot at rs2596542. Above, the *P* values of genotyped SNPs are plotted (as $-\log_{10}$ values) against their physical position on chromosome 6 (NCBI Build 36). The *P* value for rs2596542 at the GWAS stage, replication stage and combination analysis is represented by a purple diamond, circle and diamond, respectively. Estimated recombination rates from the HapMap JPT population show the local LD structure. Inset, the SNP's colors indicate LD with rs2596542 according to a scale from $r^2 = 0$ to $r^2 = 1$ based on pairwise r^2 values from HapMap JPT. Below, gene annotations from the UCSC genome browser.

Table 2 rs2596542 (A/G) is associated with progression from CHC to HCC

Subjects	RAF	(Comparison) *P*[a]	OR[a]	95% CI
Healthy	0.331			
CHC	0.333	(Healthy vs. CHC) 0.61	1.02	0.94–1.10
HCC	0.398	(CHC vs. HCC) 3.13×10^{-8}	1.36	1.22–1.51

We analyzed 5,486 controls, 1,730 CHC cases and 1,394 HCC cases. RAF, risk allele frequency (allele A); OR, odds ratio for the minor allele by considering the major allele as a reference.
[a]Calculated by logistic regression analysis, by PLINK upon age and gender adjustment under additive model.

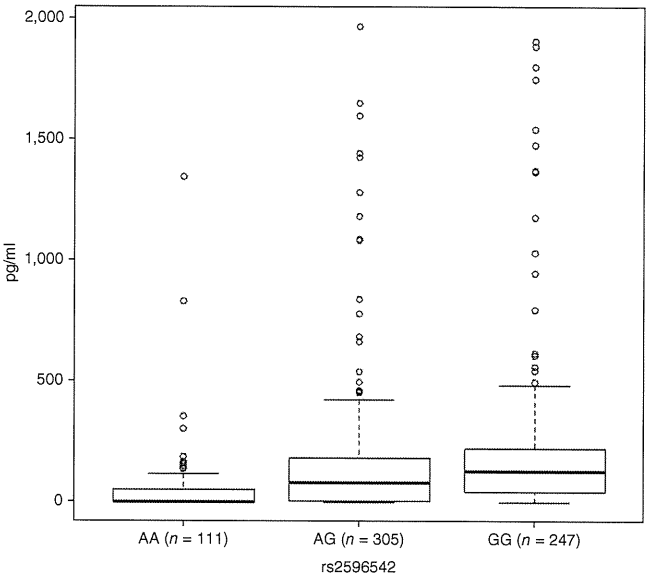

Figure 2 Correlation between soluble MICA levels and rs2596542 genotype. The *x* axis shows the genotypes at rs2596542, and the *y* axis shows the concentration of soluble MICA in pg/ml. The number of independent samples tested in each group is shown in parentheses. Each group is shown as a box plot, and the median values are shown as thick dark horizontal lines (median values of AA = 0, AG = 43.6 and GG = 77.74). The box covers the twenty-fifth to seventy-fifth percentiles, and the whiskers outside the box extend to the highest and lowest value within 1.5 times the interquartile range. Points outside the whiskers are outliers. We tested the difference in the median values among genotypes using the Kruskal-Wallis test ($P = 1.6 \times 10^{-13}$). We plotted the box plots using default settings in R (see URLs).

linked to rs2596542 ($r^2 > 0.4$) and are confined to the *MICA* gene locus. On the other hand, the imputation analysis of *HLA*-tagging SNPs did not show any evidence of linkage with rs2596542 (Online Methods and **Supplementary Table 10**), suggesting that *MICA* is a disease-associated candidate gene at this locus.

MICA is a membrane protein that acts as a ligand for NKG2D to activate anti-tumor effects through natural killer cells and $CD8^+$ T cells[19]. On the other hand, MICA is secreted into the serum by cleavage at the transmembrane domain with matrix metalloproteinases[20,21] and inhibits the anti-tumor effect of natural killer cells and $CD8^+$ T cells by blocking their action[22–24]. Elevated expression of both the membrane-bound and soluble forms of MICA (sMICA) have been reported in several cancers, including HCC[25–27]. Exon 5 of *MICA* encodes the transmembrane domain and contains a variable number of tandem repeats (VNTR) consisting of 4, 5, 6 or 9 repeats of GCT or one additional G nucleotide insertion into the 5-GCT-repeat allele (referred as A4, A5, A6, A9 and A5.1, respectively). The insertion of G (A5.1) causes a premature stop codon and subsequent loss of the transmembrane domain, leading to altered subcellular localization[28]. Therefore, we tested whether rs2596542 is in linkage with functional *MICA* VNTR alleles.

We further genotyped 673 cases with HCV-induced HCC and 890 non-HCV controls for the *MICA* VNTR locus with capillary-based electrophoresis (**Supplementary Fig. 7**). A case-control analysis revealed that the *MICA* VNTR is associated with HCV-induced HCC (global $P = 4.55 \times 10^{-7}$; **Supplementary Table 11**). Particularly, alleles A9 and A6 were associated with conferring a higher risk of HCC (OR = 1.73 and OR = 1.34, respectively), whereas the A5 and A5.1 alleles had a protective effect. Comparison of the genotypes at rs2596542 and the VNTR locus revealed that the A risk allele at rs2596542 is in LD with the A9 and A4 alleles, and the non-risk G allele is in LD with the A5 and A5.1 alleles, whereas we observed no linkage between an A6 allele and rs2596542 (**Supplementary Table 12**). We also genotyped 124 individuals with CHC; however, we observed no significant association between individuals with CHC and controls or individuals with CHC and HCC (**Supplementary Tables 13,14**).

We then tested whether the VNTR alleles, rs2596542 alleles, or VNTR-rs2596542 haplotypes had any association with MICA expression in individuals with HCV-induced HCC. We determined sMICA levels by ELISA using a total of 665 HCC serum samples (**Supplementary Table 15**). Notably, rs2596542 was significantly correlated with sMICA levels, and specifically, the risk genotype AA was associated with low levels of sMICA ($P = 1.38 \times 10^{-13}$; **Fig. 2**), whereas VNTR alleles (**Supplementary Fig. 8**) and VNTR- rs2596542 haplotypes (**Supplementary Table 16**) showed no strong association. The absence of any correlation between *MICA* VNTR alleles and sMICA suggests that sMICA levels are not regulated by post-translational processing or a premature stop codon caused by A5.1 alleles in individuals with HCC. We also examined the sMICA level in different stages of HCV-induced liver disease (in non-HCV subjects and those with CHC and HCV-induced liver cirrhosis) and found that sMICA level was elevated at the early stage of disease and was not correlated with disease progression (**Fig. 3**). Additionally, the risk allele A was also correlated with low sMICA levels in subjects with CHC (**Supplementary Fig. 9**). These findings suggest that MICA expression was induced by factors caused by chronic HCV infection,

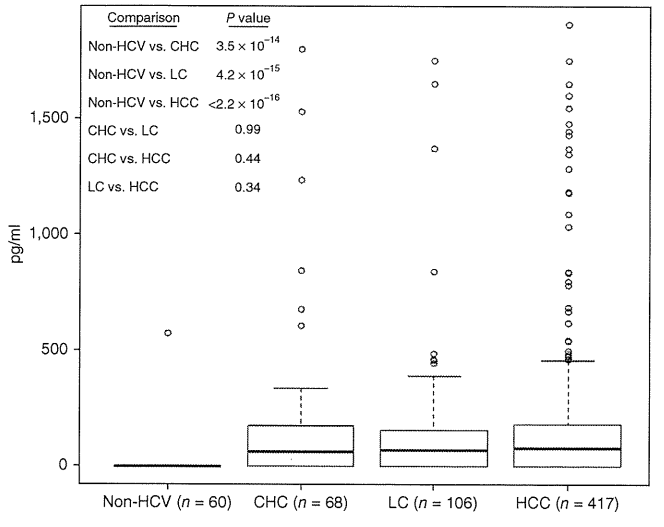

Figure 3 Correlation between soluble MICA and HCV-related diseases. The *x* axis shows the disease stages after HCV infection, and the *y* axis shows the concentration of soluble MICA in pg/ml. The number of independent samples tested in each group is shown in parentheses. Each group is shown as a box plot, and the median values are shown as thick dark horizontal lines (median values of non-HCV = 0, CHC = 64.55, LC = 72.11 and HCC = 77.98). The box covers the twenty-fifth to seventy-fifth percentiles, and the whiskers outside the box extend to the highest and lowest value within 1.5 times the interquartile range. Points outside the whiskers are outliers. We tested the difference in the median values among the disease groups using the Wilcoxon rank test. The box plots were plotted using default settings in R. Non-HCV, individuals not exposed to HCV infection; CHC, individuals with chronic hepatitis C; LC, individuals with liver cirrhosis; HCC, individuals with hepatocellular carcinoma.

similar to various types of stresses such as viral infection, inflammation and heat shock[29,30]. The levels of sMICA were shown to be directly proportional to the level of membrane-bound MICA[25], and membrane bound MICA is essential for activating natural killer cells and $CD8^+$ T cells to eliminate virus-infected cells[19]. Considering the association of the risk allele A with low levels of sMICA, our findings suggest that the individuals who carry the rs2596542 A allele would express low levels of membrane-bound MICA in response to HCV infection, which thus leads to poor or no activation of natural killer cells and $CD8^+$ T cells against virus-infected cells. Eventually, these individuals are likely to progress from CHC to HCC. Notably, several SNPs that are in absolute linkage with rs2596542 are located within the promoter or enhancer region of *MICA* and may alter the binding of stress-inducible transcriptions factors such as heat shock proteins (**Supplementary Table 17**). In this regard, it is important to analyze the factors that regulate *MICA* expression, particularly in the context of CHC. Although, the molecular mechanism whereby *MICA* polymorphisms confer the risk of disease progression should be characterized in the future, our findings reveal a crucial role of genetic variations in the host innate immune system in the development of HCV-induced HCC.

URLs. R, http://cran.r-project.org/; PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/; Primer3 v0.3.0, http://frodo.wi.mit.edu/primer3/; LocusZoom, http://csg.sph.umich.edu/locuszoom/; FastSNP, http://fastsnp.ibms.sinica.edu.tw/pages/input_CandidateGeneSearch.jsp.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We would like to thank all the subjects and the members of the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan who donated their DNA for this work. We also thank the technical staff of the Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, and the Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo. This work was conducted as a part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government.

AUTHOR CONTRIBUTIONS

K.M. and Y.N. conceived of the study; Y.N., V.K., M.K. and K.M. designed the study; V.K., Y.U., R.M. and N.H. performed genotyping; V.K., Y.N. and K.M. wrote the manuscript; A.T. and N. Kamatani performed quality control at the genome-wide phase; Y.N., K.M., H.N. and M.K. managed DNA and serum samples belonging to BioBank Japan; N. Kato, R.T., M. Otsuka, M. Omata and K.K. managed replication DNA and serum samples; V.K. analyzed the data, performed VNTR genotyping, ELISA and summarized the whole results; Y.N. obtained funding for the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Global Burden of Hepatitis C Working Group. Global burden of disease (GBD) for hepatitis C. *J. Clin. Pharmacol.* **44**, 20–29 (2004).
2. Parkin, D.M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **55**, 74–108 (2005).
3. Umemura, T., Ichijo, T., Yoshizawa, K., Tanaka, E. & Kiyosawa, K. Epidemiology of hepatocellular carcinoma in Japan. *J. Gastroenterol.* **44** (Suppl 19), 102–107 (2009).
4. Vong, S. & Bell, B.P. Chronic liver disease mortality in the United States, 1990–1998. *Hepatology* **39**, 476–483 (2004).
5. Ge, D. *et al.* Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. *Nature* **461**, 399–401 (2009).
6. Thomas, D.L. *et al.* Genetic variation in *IL28B* and spontaneous clearance of hepatitis C virus. *Nature* **461**, 798–801 (2009).
7. Rauch, A. *et al.* Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* **138**, 1338–1345, 1345.e1–7 (2010).
8. Kamatani, Y. *et al.* A genome-wide association study identifies variants in the *HLA-DP* locus associated with chronic hepatitis B in Asians. *Nat. Genet.* **41**, 591–595 (2009).
9. Schütte, K., Bornschein, J. & Malfertheiner, P. Hepatocellular carcinoma–epidemiological trends and risk factors. *Dig. Dis.* **27**, 80–92 (2009).
10. Tanaka, T. *et al.* Significance of specific antibody assay for genotyping of hepatitis C virus. *Hepatology* **19**, 1347–1353 (1994).
11. Tanaka, Y. *et al.* Genome-wide association of *IL28B* with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat. Genet.* **41**, 1105–1109 (2009).
12. Anonymous. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* **401**, 921–923 (1999).
13. de Bakker, P.I. *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat. Genet.* **38**, 1166–1172 (2006).
14. Kuniholm, M.H. *et al.* Specific human leukocyte antigen class I and II alleles associated with hepatitis C virus viremia. *Hepatology* **51**, 1514–1522 (2010).
15. Wang, J.H. *et al.* Ethnic and geographical differences in HLA associations with the outcome of hepatitis C virus infection. *Virol. J.* **6**, 46 (2009).
16. Singh, R., Kaul, R., Kaul, A. & Khan, K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. *World J. Gastroenterol.* **13**, 1770–1787 (2007).
17. Dixon, A.L. *et al.* A genome-wide association study of global gene expression. *Nat. Genet.* **39**, 1202–1207 (2007).
18. Ge, D. *et al.* Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. *Nature* **461**, 399–401 (2009).
19. Bauer, S. *et al.* Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. *Science* **285**, 727–729 (1999).
20. Salih, H.R., Rammensee, H. & Steinle, A. Cutting edge: down-regulation of MICA on human tumors by proteolytic shedding. *J. Immunol.* **169**, 4098–4102 (2002).
21. Waldhauer, I. *et al.* Tumor-associated MICA is shed by ADAM proteases. *Cancer Res.* **68**, 6368–6376 (2008).
22. Jinushi, M. *et al.* Impairment of natural killer cell and dendritic cell functions by the soluble form of MHC class I-related chain A in advanced human hepatocellular carcinomas. *J. Hepatol.* **43**, 1013–1020 (2005).
23. Groh, V., Wu, J., Yee, C. & Spies, T. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature* **419**, 734–738 (2002).
24. Doubrovina, E.S. *et al.* Evasion from NK cell immunity by MHC class I chain-related molecules expressing colon adenocarcinoma. *J. Immunol.* **171**, 6891–6899 (2003).
25. Kohga, K. *et al.* Serum levels of soluble major histocompatibility complex (MHC) class I-related chain A in patients with chronic liver diseases and changes during transcatheter arterial embolization for hepatocellular carcinoma. *Cancer Sci.* **99**, 1643–1649 (2008).
26. Jinushi, M. *et al.* Expression and role of MICA and MICB in human hepatocellular carcinomas and their regulation by retinoic acid. *Int. J. Cancer* **104**, 354–361 (2003).
27. Groh, V. *et al.* Broad tumor-associated expression and recognition by tumor-derived gamma delta T cells of MICA and MICB. *Proc. Natl. Acad. Sci. USA* **96**, 6879–6884 (1999).
28. Ota, M. *et al.* Trinucleotide repeat polymorphism within exon 5 of the MICA gene (MHC class I chain-related gene A): allele frequency data in the nine population groups Japanese, Northern Han, Hui, Uygur, Kazakhstan, Iranian, Saudi Arabian, Greek and Italian. *Tissue Antigens* **49**, 448–454 (1997).
29. Groh, V. *et al.* Costimulation of CD8αβ T cells by NKG2D via engagement by MIC induced on virus-infected cells. *Nat. Immunol.* **2**, 255–260 (2001).
30. Groh, V. *et al.* Cell stress-regulated human major histocompatibility complex class I gene expressed in gastrointestinal epithelium. *Proc. Natl. Acad. Sci. USA* **93**, 12445–12450 (1996).

ONLINE METHODS

Sample collections. We obtained DNA from 721 HCV-related HCC cases, 1,730 CHC cases and 5,486 HCV-negative controls from the BioBank Japan project[31]. For replication analysis, DNA from 673 HCV-induced HCC cases was obtained from a prospective HCC study cohort of the University of Tokyo. A diagnosis of CHC, liver cirrhosis or HCC were based on histological, clinical and laboratory findings obtained by trained physicians. Case samples with HBV co-infection were excluded from the analysis. Interferon was administrated to 20.4% of HCC cases and 70.1% of cases were not treated. The remaining 9.5% of the cases lacked information about interferon treatment. The non-HCV controls obtained from BioBank Japan contained case-mixed individuals after excluding all individuals with cancer, chronic hepatitis B, diabetes or tuberculosis. All subjects were of Japanese origin and provided written informed consent. The clinical and demographic details of the samples are summarized in **Supplementary Table 1**. We also obtained serum samples from BioBank Japan and the University of Tokyo (**Supplementary Table 12**). This research project was approved by the ethical committees of the University of Tokyo and RIKEN.

SNP genotyping and quality control. In the GWAS, 721 individuals with HCV-related liver cancer and 2,890 controls were genotyped using Illumina HumanHap610-Quad and Illumina HumanHap550v3 Genotyping BeadChip, respectively. In the replication stage, 673 cases with HCV-related disease, 1,730 cases with CHC and 2,596 controls were genotyped by the multiplex PCR-based Invader assay (Third Wave Technologies) and the Illumina HumanHap610-Quad, respectively. The common SNPs between the Illumina HumanHap550v3 and the Illumina HumanHap610-Quad arrays from all autosomal chromosomes were included for the analysis. We applied standard SNP quality control filters to exclude SNPs with low call rate (<99%), a Hardy-Weinberg equilibrium $P < 1.0 \times 10^{-6}$ for controls and minor allele frequency of <0.01. In the end, we obtained 432,703 SNPs for the analysis. In the replication analysis, the allele discrimination plots were validated by two well-trained researchers (the plots are available on request). We excluded samples with low genotyping rate (<99%) and employed principal component analysis to avoid the population stratification issue, in which individuals belonging only to Hondo cluster were included in the analysis (**Supplementary Fig. 10**)[32].

Statistical analysis. The association of SNPs with the disease phenotype in the GWAS, replication stage and combination analyses was tested using multivariate logistic regression analysis after adjusting for age at recruitment (continuous) and gender by assuming an additive model and using PLINK[33]. In the GWAS, the genetic inflation factor (λ) was derived by applying logistic regressed P values for all the tested SNPs. The quantile-quantile plot was drawn using R. The ORs were calculated by considering the major allele as a reference, unless it was stated otherwise elsewhere. The combined analysis of the GWAS and replication stage was verified by conducting the Mantel-Haenszel method. We considered $P < 5 \times 10^{-8}$ as the genome-wide significance threshold, which is the Bonferroni-corrected threshold for the number of independent SNPs genotyped in HapMap Phase 2 (ref. 34). Heterogeneity across the two stages was examined by using the Breslow-Day test[35].

For multiple logistic regression analysis at rs2596542 using the R program, we considered age at recruitment (≤60 or >60 years)[3], gender (male or female) and alcohol consumption (non-drinkers, ≤ 50 g alcohol per day or >50 g alcohol per day) as covariates from both the GWAS and replication stage cases with HCC and non-HCV controls. Association at the *MICA* VNTR locus was analyzed by Fisher's exact test, and the global P value was calculated using a χ^2 test. Statistical comparisons between genotypes and sMICA levels were performed by Kruskal-Wallis test or Wilcoxon rank test using R. We employed the R package haplo.stats to infer haplotypes and to perform haplotype association analysis. P values for association between sMICA levels and haplotype distribution were obtained by score test under an additive model by using the haplo.score function. ORs and 95% confidence intervals were calculated from the coefficients of the GLM model by considering the major haplotype as a reference. We used the haplo.cc function to calculate these statistical values.

HCV serotype. HCV serotype data was available for 531 cases with HCC from the replication stage. HCV serotype was examined by serotyping assay (SRL Laboratory) according to previously reported methods[36]. According to the Simmonds classification[37], serotype 1 corresponded to disease types 1a and 1b, whereas serotype 2 corresponded to disease types 2a and 2b.

***MICA* VNTR locus genotyping.** We followed the method suggested by Applied Biosystems. Briefly, the 5′ end of the forward primer was labeled with 6-FAM, and the 5′ end of reverse primer was labeled with the GTGTCTT non-random sequence to promote addition of As. The primer sequences were previously reported[28]. The PCR products were mixed with Hi-Di Formamide and GeneScan-600 LIZ size standard and separated using a GeneScan system on a 3730xl DNA analyzer (Applied Biosystems). GeneMapper software (Applied Biosystems) was used to assign the repeat fragment size (**Supplementary Fig. 7**).

Quantification of soluble MICA. sMICA levels were measured by sandwich enzyme-linked immunosorbent assay, as described in the manufacturer's instructions (R&D Systems).

Imputation and association analysis at HLA allele tagging SNPs. We obtained a SNP or a combination of SNPs which can tag HLA alleles in the Japanese population from a previous study[13]. The untyped genotypes of these SNPs were imputed in the GWAS samples by using a hidden Markov model programmed in MACH[38] and haplotype information from HapMap JPT samples. We applied the same SNP quality criteria as in the GWAS for selecting SNPs for the analysis. The association was tested on all SNPs that passed the quality control criteria using logistic regression analysis conditioned on age and gender.

Initially, we obtained the pair-wise LD between HLA alleles tagging SNPs and rs2596542. We performed case-control association analysis in our GWAS dataset. As shown in **Supplementary Table 9**, none of the HLA-tagging SNPs showed evidence of linkage or association except rs2844521, and rs2844521 was in absolute linkage with rs2596542 ($r^2 = 1$, $D' = 1$) and thus showed similar association. We obtained actual genotype data at rs2596501, as this SNP is included on the 550K SNP platform, and inferred the haplotype between rs2844521 and rs2596501. However, the haplotype GT (the G allele of rs2844521 and the T allele of rs2596501), which is reported to tag the *HLA-B*3501* allele ($r^2 = 1$, $D' = 1$), was not associated with HCC in our GWAS dataset ($P = 0.39$). We also performed a conditional logistic regression analysis on rs2596501 (data not shown) and found no effect on the association between rs2596542 and HCV-induced HCC. This data suggested that rs2596542 association is independent of *HLA-B*3501*. Although we observed mild association between other *HLA-B* alleles (*HLA-B*5401*, $P = 0.004$; *HLA-B*6701*, $P = 0.012$) and HCV-induced HCC, the association at rs2596542 alone was the most significant. Taken together, we found no strong evidence for linkage of HLA alleles with rs2596542.

Software. For general statistical analysis, we used R statistical environment version 2.6.1 or plink version 1.06. The Haploview software version 4.2 (ref. 39) was used to calculate LD and to draw Manhattan plots. Primer3 v0.3.0 web tool was used to design primers. We used LocusZoom for plotting regional association plots. We used FastSNP[40] web tool for functional annotation of SNPs (see URLs for all software packages).

31. Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).
32. Yamaguchi-Kabata, Y. *et al.* Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
33. 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).
34. Frazer, K.A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
35. Breslow, N. & Day, N. Statistical methods in cancer research. Volume II–The design and analysis of cohort studies. *IARC Sci. Publ.* 1–406 (1987).
36. Tsukiyama-Kohara, K. *et al.* A second group of hepatitis C viruses. *Virus Genes* **5**, 243–254 (1991).
37. Simmonds, P. *et al.* Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. *J. Gen. Virol.* **75**, 1053–1061 (1994).
38. Scott, L.J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
39. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
40. Yuan, H.Y. *et al.* FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.* **34**, W635–W641 (2006).

nature genetics

Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population

Tomomitsu Hirota[1], Atsushi Takahashi[2], Michiaki Kubo[3], Tatsuhiko Tsunoda[4], Kaori Tomita[1], Satoru Doi[5], Kimie Fujita[6], Akihiko Miyatake[6], Tadao Enomoto[7], Takehiko Miyagawa[8], Mitsuru Adachi[9], Hiroshi Tanaka[10], Akio Niimi[11], Hisako Matsumoto[11], Isao Ito[11], Hironori Masuko[12], Tohru Sakamoto[12], Nobuyuki Hizawa[12], Masami Taniguchi[13], John J Lima[14], Charles G Irvin[15], Stephen P Peters[16], Blanca E Himes[17], Augusto A Litonjua[17], Kelan G Tantisira[17], Scott T Weiss[17], Naoyuki Kamatani[18], Yusuke Nakamura[19] & Mayumi Tamari[1]

Bronchial asthma is a common inflammatory disease caused by the interaction of genetic and environmental factors[1,2]. Through a genome-wide association study and a replication study consisting of a total of 7,171 individuals with adult asthma (cases) and 27,912 controls in the Japanese population, we identified five loci associated with susceptibility to adult asthma. In addition to the major histocompatibility complex and *TSLP-WDR36* loci previously reported, we identified three additional loci: a *USP38-GAB1* locus on chromosome 4q31 (combined $P = 1.87 \times 10^{-12}$), a locus on chromosome 10p14 ($P = 1.79 \times 10^{-15}$) and a gene-rich region on chromosome 12q13 ($P = 2.33 \times 10^{-13}$). We observed the most significant association with adult asthma at rs404860 in the major histocompatiblity complex region ($P = 4.07 \times 10^{-23}$), which is close to rs2070600, a SNP previously reported for association with FEV_1/FVC in genome-wide association studies for lung function. Our findings offer a better understanding of the genetic contribution to asthma susceptibility.

The precise mechanisms underlying the development and progression of bronchial asthma have not been fully elucidated. Genome-wide association studies (GWAS) have contributed to identification of common genetic variants related to common diseases, including asthma[3]. GWAS of asthma in European and African-ancestry populations[4–8] have been conducted, and a recent large-scale, consortium-based GWAS among European populations reported ten loci strongly associated with asthma[9]. To provide further information for a better understanding of the genetic basis of asthma, GWAS using different ethnic populations are also needed.

Because there have been few large-scale analyses of asthma in Asian populations, we conducted a GWAS with 1,532 cases and 3,304 controls in the Japanese population, using Illumina HumanHap 550v3 and 610-Quad BeadChips (**Supplementary Table 1**). After principal component analysis and quality-control filtering, we excluded one outlier before the association analysis and subjected a total of 458,847 SNP loci to statistical analysis; we then generated a quantile-quantile plot using the results of a Cochran-Armitage test (**Supplementary Fig. 1a,b**). The genomic inflation factor (λ_{GC}) was 1.07, and we calculated the $\lambda_{1,000}$ adjusted by sample size to be 1.03, indicating a low possibility of false-positive associations resulting from population stratification. We observed the strongest association (at rs8192565) in the major histocompatibility complex (MHC) class III region on chromosome 6p21 (**Fig. 1**), and two variants in the *NOTCH4* and *C6orf10* loci reached the genome-wide significant threshold ($P < 5 \times 10^{-8}$; **Supplementary Fig. 1c**).

We validated the association of previously reported loci for bronchial asthma in GWAS. A recent large-scale, consortium-based GWAS identified ten loci for asthma with genome-wide significance[9]. Our GWAS replicated three SNPs reported in the previous study: rs1063355 on chromosome 6p21 ($P = 0.0010$), rs11071559 on chromosome 15q22 ($P = 0.0036$) and rs744910 on chromosome 15q22 ($P = 0.0095$) (**Supplementary Table 2a,b**). We also checked the associations of genes reported by previous GWAS for asthma[5–7,9]. The SNPs for which we observed replication of the association with $P < 0.05$ in our GWAS are shown in **Supplementary Table 2c**.

[1]Laboratory for Respiratory Diseases, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan. [2]Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Tokyo, Japan. [3]Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan. [4]Laboratory for Medical Informatics, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan. [5]Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, Habikino, Osaka, Japan. [6]Miyatake Asthma Clinic, Osaka, Japan. [7]Nonprofit Organization (NPO) Japan Health Promotion Supporting Network, Wakayama, Japan. [8]Miyagawa Clinic, Gifu, Japan. [9]First Department of Internal Medicine, Showa University School of Medicine, Shinagawa, Tokyo, Japan. [10]Third Department of Internal Medicine, Sapporo Medical University School of Medicine, Sapporo, Hokkaido, Japan. [11]Department of Respiratory Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [12]Division of Respiratory Medicine, Institute of Clinical Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan. [13]National Clinical Research Center, National Hospital Organization, Sagamihara National Hospital, Kanagawa, Japan. [14]Nemours Children's Clinic, Centers for Clinical Pediatric Pharmacology and Pharmacogenetics, Jacksonville, Florida, USA. [15]Vermont Lung Center, Department of Medicine and Physiology, University of Vermont, Burlington, Vermont, USA. [16]Center for Human Genomics, Section of Pulmonary, Critical Care, Allergy and Immunologic Diseases, Wake Forest University School of Medicine, Winston Salem, North Carolina, USA. [17]Channing Laboratory, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [18]Laboratory for International Alliance, Center for Genomic Medicine, RIKEN, Yokohama, Kanagawa, Japan. [19]Laboratory of Molecular Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. Correspondence should be addressed to M.T. (tamari@src.riken.jp).

Received 5 April; accepted 16 June; published online 31 July 2011; doi:10.1038/ng.887

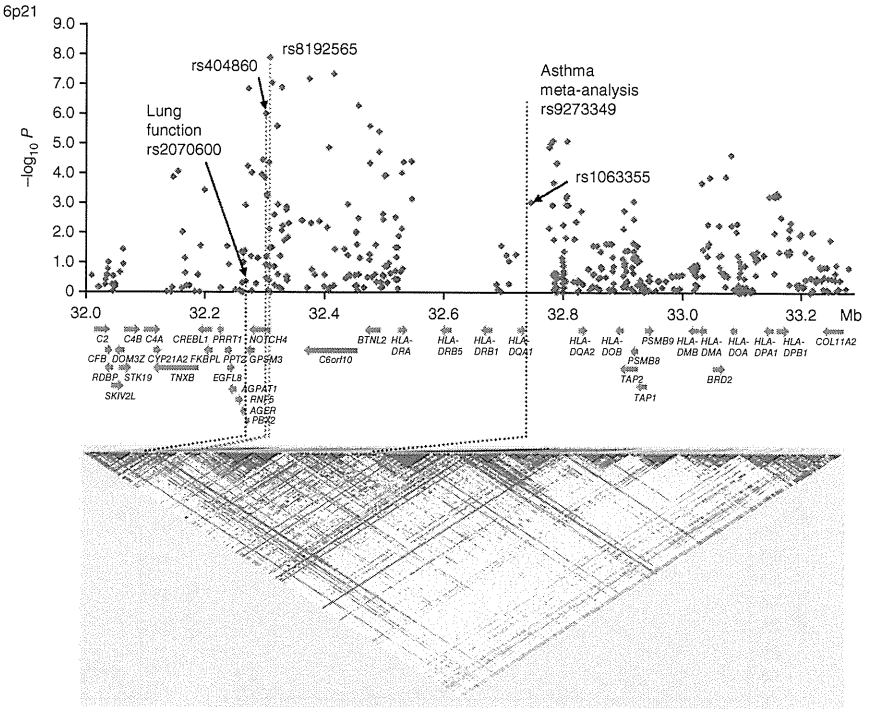

Figure 1 Case-control association results and LD map of the MHC region. The blue plots show the $-\log_{10}$ of the Cochrane-Armitage trend *P* values for the association results of the GWAS. We drew the LD maps based on the *D′* values using the genotype data of all cases and controls in the GWAS. Red dotted line, rs404860; blue dotted line, rs8192565; black dotted lines, locations of the SNPs giving significant signals for lung function (rs2070600), FEV_1/FVC, in the GWAS and for bronchial asthma in a large scale, consortium-based GWAS (rs9273349 and rs1063355). rs1063355 is absolute LD with rs9273349 ($r^2 = 1$) in the HapMap JPT and CEU populations. Blue arrows indicate the locations of genes. The LD values among the SNPs are shown in **Supplementary Table 4**.

To investigate additional susceptibility loci for adult asthma, we conducted a replication study using an independent set of 5,639 cases and 24,608 controls. We genotyped the 102 SNPs that showed Cochran-Armitage trend $P < 1 \times 10^{-4}$ in our GWAS after excluding 31 SNPs having $r^2 > 0.9$ with some of these 102 SNPs. In the replication study, 24 of the 102 SNPs replicated their significant associations with Bonferroni-corrected $P < 4.9 \times 10^{-4}$ (calculated as 0.05/102) (**Supplementary Table 3**). When we combined the data from the GWAS and replication study using the Mantel-Haenszel method, all of the 24 SNPs in five genomic loci were significantly associated with adult asthma at genome-wide significant levels ($P < 5 \times 10^{-8}$): chromosome 4q31 (rs7686660, combined $P = 1.87 \times 10^{-12}$, odds ratio (OR) = 1.16), 5q22 (rs1837253, $P = 1.24 \times 10^{-16}$, OR = 1.17), 6p21 (rs404860, $P = 4.07 \times 10^{-23}$, OR = 1.21), 10p14 (rs10508372, $P = 1.79 \times 10^{-15}$, OR = 1.16) and 12q13 (rs1701704, $P = 2.33 \times 10^{-13}$, OR = 1.19) (**Table 1**). A recent study using a candidate gene approach reported that the marker SNP rs1837253, which is located 5.7 kb upstream of the transcription initiation site of *TSLP*, was associated with asthma and airway hyper-responsiveness[10]. It was also reported that the HLA locus is associated with asthma and allergen sensitization[9,11]. Considering these previously reported results, our data identified three loci newly associated with adult asthma.

After we combined the data of the GWAS and the replication study (**Supplementary Table 3**), 18 of the 26 SNPs in the MHC region showed association beyond the genome-wide significance threshold of $P < 5 \times 10^{-8}$. We observed the strongest association in the MHC class III region (**Table 1** and **Fig. 1**). A recent large-scale European GWAS reported an association of the *HLA-DQ* locus (rs9273349) with asthma, and cases with later-onset asthma seemed to be influenced more markedly by this locus than cases with childhood-onset asthma[9]. To examine the relationship between rs9273349 and rs8192565, we conducted a logistic regression analysis using the GWAS data. Because rs9273349 was not included in our GWAS data, we used rs1063355 as a proxy SNP ($r^2 = 1$ between rs1063355 and rs9273349 in the HapMap JPT and CEU populations) (**Supplementary Table 2b**). The strong association of SNPs in the MHC class III locus remained after

Table 1 Summary of the association results of the GWAS and the replication study

			Cases		Controls				
dbSNP locus	Allele (risk allele)	Stage	Total	RAF	Total	RAF	$P^{a,b}$	OR (95% CI)[c]	P_{het}[e]
rs7686660	T/G	GWAS	1,532	0.31	3,301	0.27	1.33×10^{-4}	1.21 (1.10–1.33)	
4q31	(T)	Replication	5,623	0.30	24,606	0.27	3.33×10^{-9}	1.15 (1.10–1.20)	
		Combined[d]	7,155	0.30	27,907	0.27	1.87×10^{-12}	1.16 (1.11–1.21)	0.29
rs1837253	T/C	GWAS	1,532	0.40	3,304	0.36	3.50×10^{-5}	1.21 (1.11–1.32)	
5q22	(C)	Replication	5,629	0.39	24,601	0.35	1.02×10^{-12}	1.17 (1.12–1.22)	
		Combined[d]	7,161	0.39	27,905	0.35	1.24×10^{-16}	1.17 (1.13–1.22)	0.43
rs404860	A/G	GWAS	1,532	0.53	3,304	0.48	2.33×10^{-6}	1.24 (1.13–1.35)	
6p21	(A)	Replication	5,583	0.55	24,600	0.50	6.42×10^{-18}	1.20 (1.15–1.25)	
		Combined[d]	7,115	0.55	27,904	0.50	4.07×10^{-23}	1.21 (1.16–1.25)	0.54
rs10508372	T/C	GWAS	1,532	0.47	3,304	0.43	3.47×10^{-5}	1.21 (1.11–1.31)	
10p14	(C)	Replication	5,612	0.47	24,608	0.43	1.31×10^{-11}	1.15 (1.11–1.20)	
		Combined[d]	7,144	0.47	27,912	0.43	1.79×10^{-15}	1.16 (1.12–1.21)	0.36
rs1701704	T/G	GWAS	1,532	0.23	3,304	0.19	2.16×10^{-6}	1.30 (1.17–1.45)	
12q13	(G)	Replication	5,590	0.21	24,608	0.18	7.22×10^{-9}	1.16 (1.11–1.22)	
		Combined[d]	7,122	0.21	27,912	0.18	2.33×10^{-13}	1.19 (1.14–1.25)	0.06

RAF, risk allele frequency; OR, odds ratio; CI, confidence interval.
[a]*P* value of the Cochran-Armitage trend test for each stage. [b]*P* values in the GWAS indicate λ-corrected probability, P_{GC}. [c]We calculated odds ratios (ORs) and confidence intervals (CIs) using the non-susceptible allele as a reference. [d]We calculated the results of combined analyses using the Mantel-Haenszel method. [e]The results of Breslow-Day test.

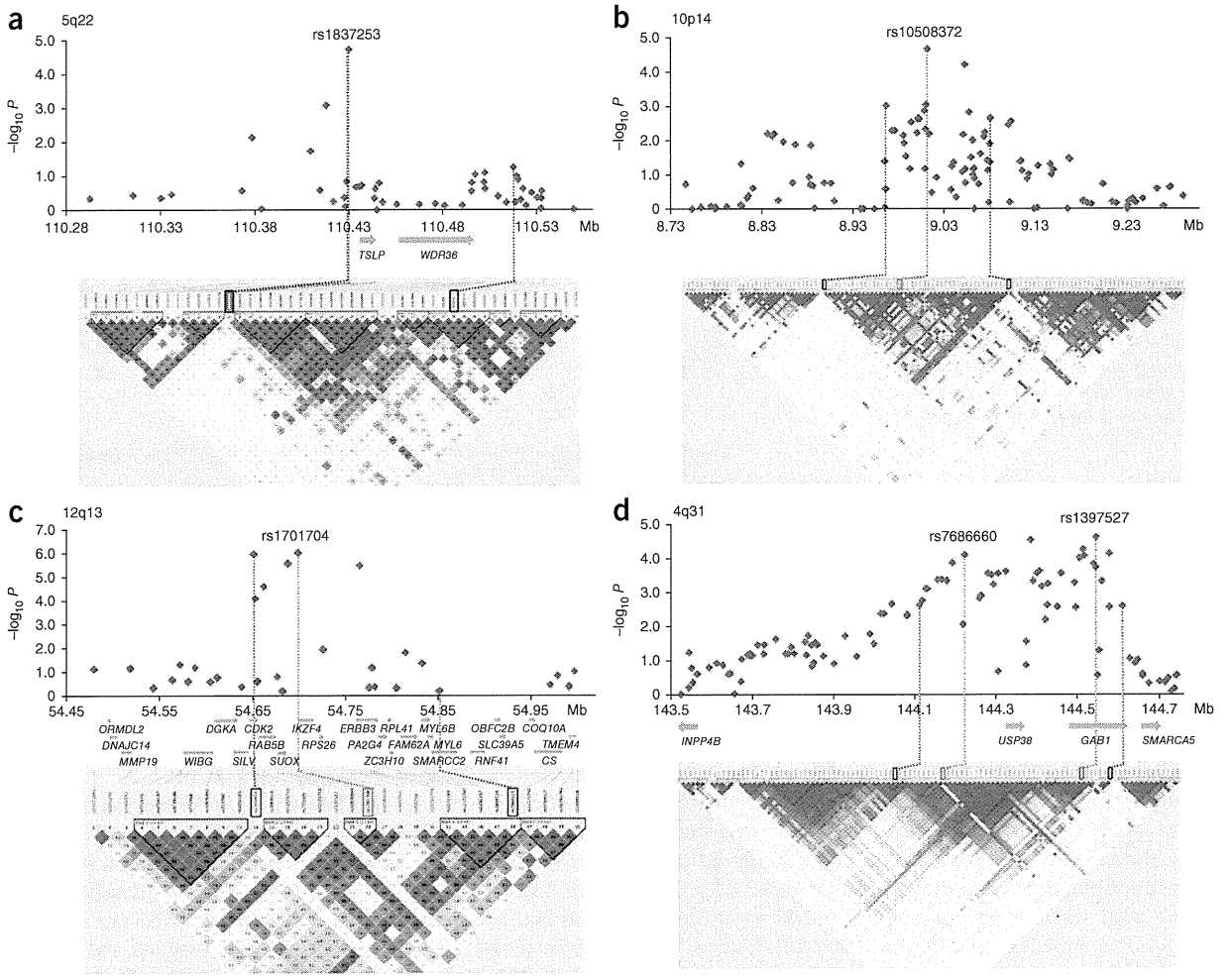

Figure 2 Case-control association results and LD maps of the four candidate regions. (**a**–**d**) *P* value plot, genomic structures and LD maps of chromosomes 5q22 (**a**), 10p14 (**b**), 12q13 (**c**) and 4q31 (**d**). The blue plots show the $-\log_{10}$ of Cochrane-Armitage trend *P* values for the association results of the GWAS. We drew the LD maps based on the *D′* values using the genotype data of all cases and controls in the GWAS. Black and red dotted lines indicate the ranges of the susceptible regions and positions of marker SNPs, respectively. Blue arrows indicate the locations of genes.

adjustment for rs1063355 (**Supplementary Fig. 2**). This result indicated that the association observed in our study was independent of rs1063355 (and hence, rs9273349) in the *HLA-DQ* locus.

Recent GWAS of natural variations in lung function have shown that a non-synonymous coding SNP in *AGER* (rs2070600) on chromosome 6p21 is associated with forced expiratory volume in one second (FEV_1)/forced vital capacity (FVC)[12,13]. Reduction of FEV_1/FVC is a characteristic of obstructive lung diseases such as asthma[14]. rs2070600 showed relatively weak linkage disequilibrium (LD) with rs404860 in the Japanese population ($r^2 = 0.115$ and $D' = 0.805$; **Supplementary Table 2b**) and did not show association with adult asthma in our study. However, rs404860 is located close to rs2070600, 32.9 kb downstream of the SNP (**Fig. 1**). Asthma is a heterogenous disease with differences in severity and natural history, and lower pulmonary function is a feature of severe asthma. Thus, further genetic studies of asthma with detailed phenotypic characterization might help to clarify the relationship between genetic determinants of lung function and susceptibility to asthma.

To further investigate the gene(s) conferring susceptibility to adult asthma mapped to four loci (excluding the MHC region), we performed detailed mapping analyses using 1,532 cases and 3,304 controls from the GWAS (**Fig. 2a–d**). We selected and genotyped tag SNPs (with minor allele frequency ≥0.05 and $r^2 < 0.8$) using the data from the HapMap phase 2 and 3 (release 27) Japanese individuals. The fine mapping analysis of the region on chromosome 5q22 using 13 tag SNPs showed that rs1837253 represented an associated LD block spanning 88 kb that included two genes, *TSLP* and *WDR36* (**Fig. 2a** and **Supplementary Table 4a**). Respiratory viruses, multiple protease allergens and inflammatory cytokines can induce TSLP expression in airway epithelial cells, and TSLP plays a critical role in sensing environmental agents and mediating T helper type 2 (T_H2) cell responses following exposure[15]. Although further investigation is required to determine the function of the related variants in asthma, current annotation suggests that *TSLP* is the most plausible susceptibility gene in this locus.

Fine mapping analysis of the locus on chromosome 10p14 with 22 tag SNPs showed that rs10508372 is located within a 112-kb LD block that contains no reported genes (**Fig. 2b** and **Supplementary Table 4b**). This region is a gene desert and is located 1 Mb downstream of *GATA3*, a master regulator of T_H2 cell differentiation[16]. As enhancers might influence the expression of the gene with Mb-long genomic distance[17], it should be verified whether the variants within the associated region affect the activity of long range *cis*-regulatory elements controlling *GATA3* expression.

The fourth locus is located in a gene-rich region on chromosome 12q13. Fine mapping analysis using 12 additional tag SNPs indicated

that rs1701704 is present within a 201-kb LD block including 13 genes (**Fig. 2c** and **Supplementary Table 4c**). Recent association studies have shown that rs1701704 at this locus is associated with type I diabetes[18] and alopecia areata[19]. rs1701704 is located 2 kb upstream from *IKZF4* (also known as *EOS*), which was shown to be involved in differentiation of regulatory T cells as a coregulator of FOXP3-directed gene silencing[20]. Regulatory T cells maintain pulmonary homeostasis and prevent harmful immune responses to innocuous inhaled antigens[21].

We finely mapped the fifth locus on chromosome 4q31 using 11 tag SNPs. The 497-kb susceptibility LD block region including rs1397527 contained *USP38-GAB1* (**Fig. 2d** and **Supplementary Table 4d**). *USP38* encodes ubiquitin-specific peptidase 38, whose function is unclear. GAB1 is a scaffolding adopter protein that plays an important role in the signaling pathway activated by cytokine receptors for IL-3, IL-6, interferon α and γ, and B-cell and T-cell receptors[22].

To examine the association of these variants with adult asthma in populations of different ancestries, we conducted an association study in 499 non-Hispanic individuals of European ancestry with mild to severe asthma from three adult asthma populations and 639 controls obtained from Illumina's iControl database. We found that rs1837253 on 5q22 showed a significant association with adult asthma ($P = 0.023$) (**Supplementary Table 5**). Hence, we conclude that the locus containing *TSLP* may be a common susceptible region for asthma across various ethnic groups.

In conclusion, we identified five loci that confer susceptibility to adult asthma in the Japanese population at genome-wide significant levels. The three new loci identified in our GWAS were not observed in a previous study with a large meta-analysis in which 65.4% of the cases were subjects with childhood-onset asthma. In contrast, we recruited subjects with adult asthma, though 19.6% of the cases in the GWAS were subjects with childhood-onset asthma. Asthma has various clinical phenotypes that show age-related variation. The heterogeneity of phenotypes between these studies might have influenced these results. The peak of association at 6p21 in this study was located close to rs2070600, which was associated with FEV_1/FVC in GWAS of lung function in healthy individuals[12,13]. We found for the first time, to our knowledge, an association of asthma with the *TSLP-WDR36* locus at a genome-wide significant level. A SNP in the *TSLP-WDR36* locus was associated with adult asthma in populations of both Japanese and non-Hispanic individuals of European ancestry. Further studies are necessary for better understanding of the clear role of genetic etiology and the pathophysiology of bronchial asthma.

URLs. The Leading Project for Personalized Medicine, http://biobankjp.org/; Haploview v4.2, http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview; PLINK statistical software v1.06, http://pngu.mgh.harvard.edu/~purcell/plink/; Illumina's IconDB resource, http://www.illumina.com/science/icontroldb.ilmn.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all the subjects for participating in the study and also thank the collaborating physicians for helping with sample collection. We are grateful to the members of BioBank Japan and the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. We thank M.T. Shimizu, H. Sekiguchi, A.I. Jodo, N. Kawaraichi and the technical staff of the Center for Genomic Medicine for providing technical assistance. This work was conducted as a part of the BioBank Japan Project supported by of Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was also partly supported by grants from the Ministry of Health, Labour and Welfare, Japan. We acknowledge the American Lung Association (ALA) and the ALA's Asthma Clinical Research Centers investigators and research teams for use of Leukotriene Modifier or Corticosteroid Salmeterol (LOCCS) and Effectiveness of Low Dose Theophylline as an Add-on Treatment in Asthma (LODO) data, with additional funding from HL071394 and HL074755 from the National Heart, Lung, and Blood Institute and the Nemours Children's Clinic. GlaxoSmithKline supported the conduct of the LOCCS Trial by an unrestricted grant to the ALA. We acknowledge Sepracor, Inc. for use of the Asthma Trial data. C.G.I. was supported by grants from the US National Institutes of Health NCRR RR022675 and RR015557 as well as the ALA, Asthma Clinical Research Center award. S.P.P. serves as an advisor to the data coordinating center of the ALA Asthma Clinical Research Centers and served as the principal investigator of the LOCCS trial.

AUTHOR CONTRIBUTIONS

T.H. and M. Tamari designed the study and drafted the manuscript. A.T. and T.T. analyzed the GWAS data. T.H., K.T. and M.K. performed the genotyping for the GWAS. S.D., K.F., A.M., T.E., T.M., M.A., H.T., A.N., H. Matsumoto, I.I., H. Masuko, T.S., N.H. and M. Taniguchi collected subjects and participated in the diagnostic evaluations. B.E.H., A.A.L., K.G.T., J.J.L., C.G.I., S.P.P. and S.T.W. conducted an association study in a non-Hispanic population of European ancestry. M. Tamari and S.T.W. wrote the manuscript. M.K., N.K. and Y.N. contributed to the overall GWAS study design.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Barnes, P.J. Immunology of asthma and chronic obstructive pulmonary disease. *Nat. Rev. Immunol.* **8**, 183–192 (2008).
2. Vercelli, D. Discovering susceptibility genes for asthma and allergy. *Nat. Rev. Immunol.* **8**, 169–182 (2008).
3. Kabesch, M. Novel asthma-associated genes from genome-wide association studies: what is their significance? *Chest* **137**, 909–915 (2010).
4. Moffatt, M.F. *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
5. Himes, B.E. *et al.* Genome-wide association analysis identifies *PDE4D* as an asthma-susceptibility gene. *Am. J. Hum. Genet.* **84**, 581–593 (2009).
6. Sleiman, P.M. *et al.* Variants of *DENND1B* associated with asthma in children. *N. Engl. J. Med.* **362**, 36–44 (2010).
7. Li, X. *et al.* Genome-wide association study of asthma identifies *RAD50-IL13* and *HLA-DR/DQ* regions. *J. Allergy Clin. Immunol.* **125**, 328–335 (2010).
8. Mathias, R.A. *et al.* A genome-wide association study on African-ancestry populations for asthma. *J. Allergy Clin. Immunol.* **125**, 336–346 (2010).
9. Moffatt, M.F. *et al.* A large-scale, consortium-based genomewide association study of asthma. *N. Engl. J. Med.* **363**, 1211–1221 (2010).
10. He, J.Q. *et al.* A thymic stromal lymphopoietin gene variant is associated with asthma and airway hyperresponsiveness. *J. Allergy Clin. Immunol.* **124**, 222–229 (2009).
11. Moffatt, M.F. *et al.* Association between quantitative traits underlying asthma and the *HLA-DRB1* locus in a family-based population sample. *Eur. J. Hum. Genet.* **9**, 341–346 (2001).
12. Repapi, E. *et al.* Genome-wide association study identifies five loci associated with lung function. *Nat. Genet.* **42**, 36–44 (2010).
13. Hancock, D.B. *et al.* Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat. Genet.* **42**, 45–52 (2010).
14. Weiss, S.T. Lung function and airway diseases. *Nat. Genet.* **42**, 14–16 (2010).
15. Ziegler, S.F. & Artis, D. Sensing the outside world: TSLP regulates barrier immunity. *Nat. Immunol.* **11**, 289–293 (2010).
16. Ho, I.C. *et al.* GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. *Nat. Rev. Immunol.* **9**, 125–135 (2009).
17. Wasserman, N.F. *et al.* An 8q24 gene desert variant associated with prostate cancer risk confers differential *in vivo* activity to a MYC enhancer. *Genome Res.* **20**, 1191–1197 (2010).
18. Hakonarson, H. *et al.* A novel susceptibility locus for type 1 diabetes on Chr12q13 identified by a genome-wide association study. *Diabetes* **57**, 1143–1146 (2008).
19. Petukhova, L. *et al.* Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature* **466**, 113–117 (2010).
20. Pan, F. *et al.* Eos mediates Foxp3-dependent gene silencing in $CD4^+$ regulatory T cells. *Science* **325**, 1142–1146 (2009).
21. Lambrecht, B.N. & Hammad, H. The role of dendritic and epithelial cells as master regulators of allergic airway inflammation. *Lancet* **376**, 835–843 (2010).
22. Sármay, G. *et al.* The multiple function of Grb2 associated binder (Gab) adaptor/scaffolding protein in immune cell signaling. *Immunol. Lett.* **104**, 76–82 (2006).