OPEN ACCESS Freely available online

PLOS ONE

Identification of an HLA-A2-Restricted Epitope Peptide Derived from Hypoxia-Inducible Protein 2 (HIG2)

Sachiko Yoshimura[1,2], Takuya Tsunoda[1,2,3], Ryuji Osawa[1,2], Makiko Harada[2], Tomohisa Watanabe[2], Tetsuro Hikichi[2], Masahiro Katsuda[1], Motoki Miyazawa[1], Masaji Tani[1], Makoto Iwahashi[1], Kazuyoshi Takeda[4], Toyomasa Katagiri[3,5], Yusuke Nakamura[3,6], Hiroki Yamaue[1]*

1 Second Department of Surgery, Wakayama Medical University, Wakayama, Japan, **2** OncoTherapy Science Inc, Research and Development Division, Kanagawa, Japan, **3** Laboratory of Molecular Medicine Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **4** Department of Immunology, Juntendo University School of Medicine, Tokyo, Japan, **5** Division of Genome Medicine, Institute for Genome Research, The University of Tokushima, Tokushima, Japan, **6** Department of Medicine, University of Chicago, Chicago, Illinois, United States of America

Abstract

We herein report the identification of an HLA-A2 supertype-restricted epitope peptide derived from hypoxia-inducible protein 2 (HIG2), which is known to be a diagnostic marker and a potential therapeutic target for renal cell carcinoma. Among several candidate peptides predicted by the HLA-binding prediction algorithm, HIG2-9-4 peptide (VLNLYLLGV) was able to effectively induce peptide-specific cytotoxic T lymphocytes (CTLs). The established HIG2-9-4 peptide-specific CTL clone produced interferon-γ (IFN-γ) in response to HIG2-9-4 peptide-pulsed HLA-A*02:01-positive cells, as well as to cells in which HLA-A*02:01 and HIG2 were exogenously introduced. Moreover, the HIG2-9-4 peptide-specific CTL clone exerted cytotoxic activity against HIG2-expressing HLA-A*02:01-positive renal cancer cells, thus suggesting that the HIG2-9-4 peptide is naturally presented on HLA-A*02:01 of HIG-2-expressing cancer cells and is recognized by CTLs. Furthermore, we found that the HIG2-9-4 peptide could also induce CTLs under HLA-A*02:06 restriction. Taken together, these findings indicate that the HIG2-9-4 peptide is a novel HLA-A2 supertype-restricted epitope peptide that could be useful for peptide-based immunotherapy against cancer cells with HIG2 expression.

Citation: Yoshimura S, Tsunoda T, Osawa R, Harada M, Watanabe T, et al. (2014) Identification of an HLA-A2-Restricted Epitope Peptide Derived from Hypoxia-Inducible Protein 2 (HIG2). PLoS ONE 9(1): e85267. doi:10.1371/journal.pone.0085267

Editor: Rachel Louise Allen, University of London, St George's, United Kingdom

Received September 20, 2013; **Accepted** November 25, 2013; **Published** January 8, 2014

Funding: The authors have no support or funding to report.

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

* E-mail: yamaue-h@wakayama-med.ac.jp

Introduction

Renal cell carcinoma (RCC) comprises approximately 2–3% of all human malignancies [1]. Although patients with localized RCC can be curable by radical nephrectomy, approximately 30% of patients are observed to have metastasis at the time of diagnosis, and the median survival is only 1.5 years. Furthermore, 30% of patients experience a relapse after initial surgery, and no adjuvant treatment has yet been established [2–4]. Several molecular targeting agents, including the recently approved VEGFR tyrosine kinase inhibitor [5], were developed as novel therapeutics for RCC, but the majority of patients eventually develop treatment-resistant disease [6–13]. It is notable that RCC is one of the most immune responsive cancers. IL-2 based immunotherapy is currently the only curative treatment for metastatic RCC, but it is poorly tolerated, with significant side effects, and the efficacy has been limited to a 20% response rate, including a 5–10% complete response rate [14–17]. This limited success poses further challenges to improve the efficacy of immunotherapies for RCC. While therapeutic vaccines that induce immunity in response to tumor antigens have been under investigation for decades, the number of antigens identified in RCC and the efficacy in clinical trials have been limited [18–21].

Hypoxia-inducible protein 2 (*HIG2*) was first annotated as a novel gene induced by hypoxia and glucose deprivation [22]. A recent functional analysis revealed that HIG2 is a novel lipid droplet protein that stimulates intracellular lipid accumulation [23]. We reported HIG2 upregulation in RCC, and suggested its usefulness as a diagnostic biomarker for RCC [24]. Our findings also implied that HIG2 might be a good molecular target for the development of novel cancer treatment, because its expression was hardly detectable in normal organs except for the fetal kidney. Importantly, significant growth suppression of RCC cells occurred when endogenous *HIG2* was suppressed by *HIG2*-specific RNAi, suggesting that HIG2 has an essential role in the proliferation of RCC cells. An additional study revealed that HIG2 expression was found in 86% of human RCC tissue samples (80/93) and also correlated with the clinicopathological characteristics and survival of RCC patients [25].

In the present study, we focused on HIG2 as a novel tumor antigen, which induces antigen-specific cytotoxic T lymphocytes (CTLs) against RCC cells. We investigated the HIG2-derived epitope peptide restricted to HLA-A*02:01, the most common HLA class I type in Caucasians and the second most common type in the Japanese population [26,27], and demonstrate that this epitope peptide can also be presented by another HLA-A2 supertype allele. Thus, this epitope peptide would be applicable for peptide-based immunotherapies for RCC patients with HLA-A2.

Ethics statement

The study protocol was approved by the Institutional Review Board of OncoTherapy Science, Inc. and written informed consent was obtained from all subjects, in accordance with the guidelines of the Ethical Committee on Human Research of Wakayama Medical University, School of Medicine, OncoTherapy Science, Inc., The University of Tokyo, Juntendo University School of Medicine, The University of Tokushima and University of Chicago.

Materials and Methods

Peptides

HIG2-derived 9-mer and 10-mer peptides that have high binding affinity (binding score >10) to HLA-A:*02:01 were predicted by the binding prediction software "BIMAS" (http://www-bimas.cit.nih.gov/molbio/hla_bind), and the homologous sequences were examined by the homology search program "BLAST" (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Selected high affinity peptides and the HLA-A*02:01-restricted HIV-derived epitope peptide (ILKEPVHGV) [28] were synthesized by Sigma (Ishikari, Japan). The purity (>90%) and the sequences of the peptides were confirmed by analytical HPLC and a mass spectrometry analysis, respectively. Peptides were dissolved in dimethylsulfoxide at 20 mg/ml and stored at −80°C.

Cell lines

T2 (HLA-A*02:01, lymphoblast), Jiyoye (HLA-A32, Burkitt's lymphoma), EB-3 (HLA-A3/Aw32, Burkitt's lymphoma), *Cercopithecus aethiops*-derived COS7 and A498 (HLA-A*02:01, kidney carcinoma) cells were purchased from the American Type Culture Collection (Rockville, MD). PSCCA0922 (HLA-A*02:06/A*31:01, a B cell line) was provided by the Health Science Research Resources Bank (Osaka, Japan). Caki-1 (HLA-A*24:02/A*23:01, renal clear cell carcinoma) cells were provided by the Cell Resource Center for Biomedical Research Institute of Development, Aging and Cancer at Tohoku University. The HIG2 expression in A498 and Caki-1 cells was confirmed by a Western blotting analysis [24]. T2, Jiyoye, EB-3 and PSCCA0922 cells were maintained in RPMI1640 (Invitrogen, Carlsbad, CA), A498 and Caki-1 cells were maintained in EMEM (Invitrogen) and COS7 cells were maintained in DMEM (Invitrogen). Each medium was supplemented with 10% fetal bovine serum (GEMINI Bio-Products, West Sacramento, CA) and 1% antibiotic solution (Sigma-Aldrich, ST. Louis, MO).

Gene transfection

The plasmid encoding *HLA-A*02:01* was a generous gift from Dr. Kawakami (Keio University, Tokyo Japan). cDNA fragments encoding *HLA-A*02:06* or *HIG2* (GenBank Accession Number NM_013332) were cloned into the pcDNA3.1/myc-His vector (Invitrogen). Plasmid DNAs containing *HLA-A*02:01*, *HLA-A*02:06* and/or *HIG2* were transfected into COS7 cells using Fugene 6 (Roche Diagnostics, Indianapolis, IN) according to the manufacturer's instructions. COS7 cells were incubated with the transfection mixture at 37°C overnight prior to use as stimulator cells. The introduction of the targeted proteins was confirmed by a Western blotting analysis.

In vitro CTL induction

$CD8^+$ T cells and monocyte-derived dendritic cells (DCs) were prepared from peripheral blood of healthy volunteers (either HLA-A*02:01 or HLA-A*02:06 positive) with written informed consent. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque PLUS (GE Healthcare, Uppsala, Sweden) and $CD8^+$ T cells were harvested by positive selection with a Dynal CD8 Positive Isolation Kit (Invitrogen). Monocytes were enriched from the $CD8^-$ cell population by adherence to a tissue culture dish (Becton Dickinson, Franklin Lakes, NJ) and were cultured in AIM-V (Invitrogen) containing 2% heat-inactivated autologous serum (AS), 1,000 U/ml of GM-CSF (R&D Systems, Minneapolis, MN) and 1,000 U/ml of interleukin (IL)-4 (R&D Systems) on day 1. On day 4, 0.1 KE/ml of OK-432 (Chugai Pharmaceutical Co., Tokyo, Japan) was added in the culture to induce the maturation of DCs. On day 7, DCs were pulsed with 20 μg/ml of the respective synthesized peptides in the presence of 3 μg/ml of β2-microglobulin (Sigma-Aldrich, ST. Louis, MO) in AIM-V at 37°C for 4 h [29]. These peptide-pulsed DCs were then incubated with 30 μg/ml of mitomycin C (MMC) (Kyowa Hakko Kirin Co. Ltd., Tokyo, Japan) at 37°C for 30 min. Following washing out the residual peptide and MMC, DCs were cultured with autologous $CD8^+$ T cells on 48 well plates (Corning, Inc., Corning, NY) (each well contained 1.5×10^4 peptide-pulsed DCs, 3×10^5 $CD8^+$ T cells and 10 ng/ml of IL-7 (R&D Systems) in 0.5 ml of AIM-V/2% AS). Two days later, these cultures were supplemented with IL-2 (CHIRON, Emeryville, CA) (final concentration: 20 IU/ml). On days 14 and 21, T cells were further re-stimulated with the autologous peptide-pulsed DCs, which were freshly prepared every time. On day 28, the CTL activity against peptide-pulsed T2 or PSCCA0922 cells was examined by an interferon (IFN)- γ enzyme-linked immunospot (ELISPOT) assay.

IFN-γ enzyme-linked immunospot (ELISPOT) assay

The human IFN-γ ELISPOT kit and AEC substrate set (BD Biosciences) were used to analyze the T cell response to the respective peptides. The ELISPOT assay was performed according to the manufacturer's instructions. Briefly, T2 or PSCCA0922 cells were pulsed with 20 μg/ml of the respective peptides at 37°C for 20 h, and the residual peptide that did not bind to cells was washed out to prepare peptide-pulsed cells as the stimulator cells. After removing 500 μl of supernatant from each well of *in vitro* CTL-inducing cultures, 200 μl of cell culture suspensions were harvested from each well and distributed to two new wells (100 μl each) on Multiscreen-IP 96 well plates (Millipore, Bedford, MA). The cells were co-incubated with peptide-pulsed cells (1×10^4 cells/well) at 37°C for 20 h. HIV peptide-pulsed cells were used as a negative control. Spots were captured and analyzed by an automated ELISPOT reader, ImmunoSPOT S4 (Cellular Technology Ltd, Shaker Heights, OH) and the ImmunoSpot Professional Software package, Version 5.0 (Cellular Technology Ltd).

CTL expanding culture

The peptide-specific CTLs harvested from ELISPOT-positive wells after *in vitro* CTL induction were expanded by a modified protocol based on the previously described methods [30,31]. A total of 5×10^4 CTLs was cultured with 5×10^6 MMC-inactivated Jiyoye or EB-3 cells (30 μg/ml at 37°C for 30 min treatment) in 25 ml of AIM-V/5% AS containing 40 ng/ml of anti-CD3 monoclonal antibody (BD Biosciences, San Diego, CA) on day 0. IL-2 was added 24 h later (final concentration: 120 IU/ml), and fresh AIM-V/5% AS containing 30 IU/ml of IL-2 was provided on days 5, 8 and 11. On day 14, CTLs were harvested and the CTL activity was examined by an IFN-γ enzyme-linked immunosorbent assay (ELISA).

Establishment of CTL clones

CTL clones were established by the limiting dilution method. Briefly, CTLs were diluted to 0.3, 1 or 3 cells per well in 96 well

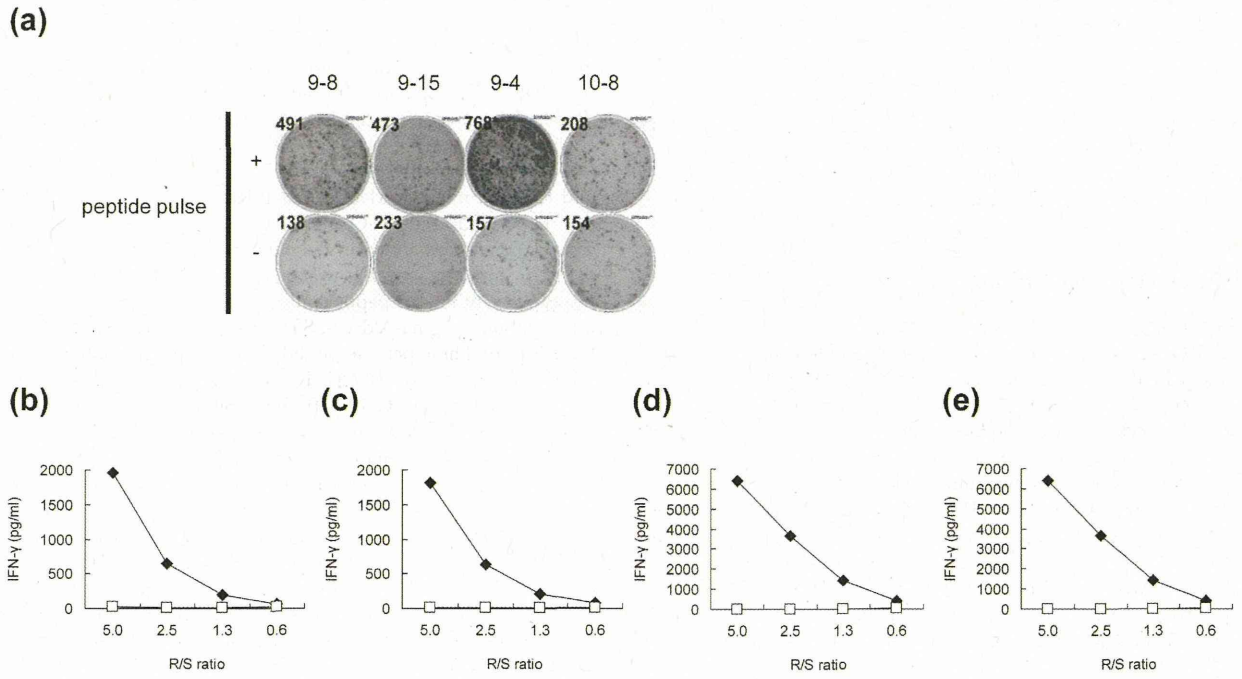

Figure 1. The IFN-γ production in response to the HIG2-9-8, HIG2-9-15, HIG2-9-4 or HIG2-10-8 peptide. (a) The IFN-γ production from cells induced by the indicated peptide-pulsed DCs was examined by an ELISPOT assay using T2 cells. "+" indicates the wells in which cells were stimulated with T2 cells pulsed with the indicated peptide and "−" indicates the wells in which cells were stimulated with HIV peptide-pulsed T2 cells. The IFN-γ production from cells induced with HIG2-9-8 (b), HIG2-9-15 (c), HIG2-9-4 (d) or HIG2-10-8 (e) peptide stimulation after CTL expanding culture was examined by ELISA. Cells were stimulated with T2 cells pulsed with the corresponding peptide (closed diamonds) or HIV peptide (open squares) at the indicated responder/stimulator ratio (R/S ratio). Similar results were obtained from three independent experiments.
doi:10.1371/journal.pone.0085267.g001

round bottom plates (Corning, Inc.), and were cultured with MMC-treated 1×10^4 Jiyoye and EB-3 cells in 125 μl AIM-V containing 5% AB serum and 30 ng/ml of an anti-CD3 monoclonal antibody on day 0. IL-2 was added to each well on day 10 (final concentration: 125 IU/ml). On day 14, an IFN-γ ELISPOT assay was performed to measure the CTL activity of each clone.

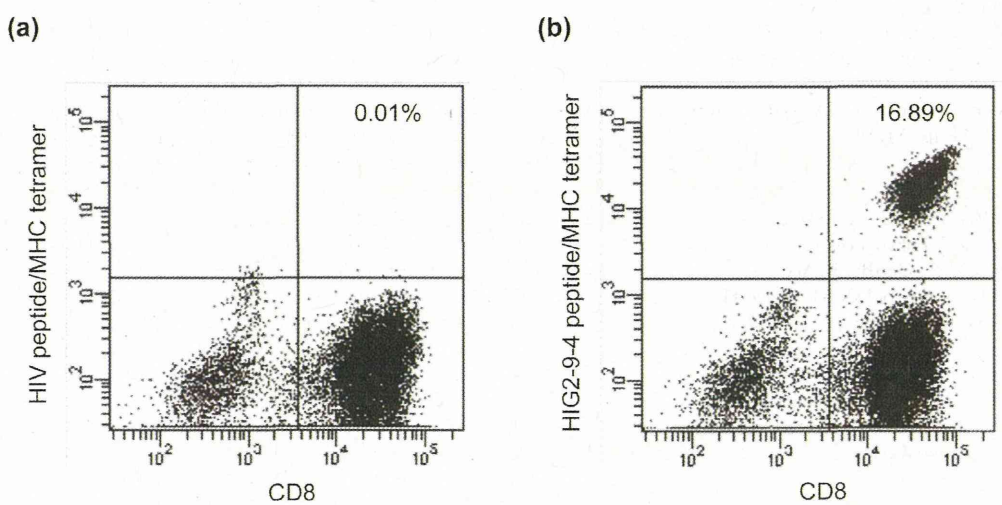

Figure 2. The expression of a HIG2-9-4 peptide-specific T cell receptor on CD8+ T cells. The expression of the HIG2-9-4 peptide-specific T cell receptor was examined on $CD3^+CD4^-$ cells following CTL expansion culture of HIG2-9-4 peptide-induced CTLs. (a) A quadrant gate was set based on the staining results with the HIV peptide/HLA-A*02: 01 tetramer. (b) $CD8^+$ T cells expressing the HIG2-9-4 peptide/HLA-A*02: 01-specific T cell receptor were detected. Similar results were obtained from three independent experiments.
doi:10.1371/journal.pone.0085267.g002

Table 1. Candidate peptides derived from HIG2 restricted with HLA-A*02:01.

Peptide name	Amino acid sequence (mer)	Binding Score
HIG2-9-8	YLLGVVLTL (9)	836.253
HIG2-9-13	VLTLLSIFV (9)	650.311
HIG2-9-15	TLLSIFVRV (9)	488.951
HIG2-9-4	VLNLYLLGV (9)	271.948
HIG2-9-9	LLGVVLTLL (9)	83.527
HIG2-9-22	RVMESLEGL (9)	31.957
HIG2-9-6	NLYLLGVVL (9)	28.027
HIG2-10-8	YLLGVVLTLL (10)	836.253
HIG2-10-29	GLLESPSPGT (10)	113.047
HIG2-10-4	VLNLYLLGVV (10)	14.495
HIG2-10-15	TLLSIFVRVM (10)	13.174
HIG2-10-18	SIFVRVMESL (10)	12.248

The binding score was obtained from the BIMAS website (http://www-bimas.cit.nih.gov/molbio/hla_bind).
doi:10.1371/journal.pone.0085267.t001

IFN-γ enzyme-linked immunosorbent assay (ELISA)

The CTL activity was examined by IFN-γ ELISA. Peptide-pulsed cells (1×10^4 cells/well) or gene-transfected cells (5×10^4 cells/well) were used to stimulate CTLs at several responder/stimulator ratios in 200 μl of AIM-V/5% AS on 96 well round bottom plates (Corning Inc.). After 24 h of incubation, cell-free supernatants were harvested, and the IFN-γ production was examined by an IFN-γ ELISA kit (BD Biosciences) according to the manufacturer's instructions.

Flow cytometry

The expression of peptide-specific T cell receptors was examined on FACS-Canto II (Becton Dickinson, San Jose, CA) using PE-conjugated peptide/MHC tetramer (Medical and Biological Laboratories, Nagoya, Japan) according to the manufacturer's instructions. Briefly, *in vitro* expanded CTLs were incubated with peptide/MHC tetramer at room temperature for 10 min, and then a FITC-conjugated anti-human CD8 mAb, APC-conjugated anti-human CD3 mAb, PE-Cy7-conjugated anti-human CD4 mAb and 7-AAD (BD Biosciences) were added and incubated at 4°C for 20 min. HIV peptide (ILKEPVHGV)/HLA-A*02: 01 tetramer was used as a negative control.

Cytotoxicity assay

The cytotoxic activity of the induced CTL clones was tested by a 4 h ^{51}Cr release assay as described previously [32]. Data are presented as the means ± SD of triplicate samples. Student's t test was used to examine the significance of the data.

Results

CTL induction with HLA-A*02:01-binding peptides derived from HIG2

We synthesized twelve 9-mer and 10-mer peptides, corresponding to parts of the HIG2 protein that had been suggested to bind to HLA-A*02:01 by the prediction with the BIMAS program (Table 1). After *in vitro* culture to induce CTLs, IFN-γ production was observed specifically when cells were stimulated with T2 cells that had been pulsed with the HIG2-9-8 peptide (YLLGVVLTL), HIG2-9-4 peptide (VLNLYLLGV), HIG2-9-15 peptide (TLLSIFVRV) or HIG2-10-8 peptide (YLLGVVLTLL) among all of the candidate peptides shown in Table 1 (Fig. S1 showing all 12 wells of one experiment and Fig. 1a showing representative wells). After CTL-expanding culture, cells still produced IFN-γ in response to the respective peptides in a responder/stimulator ratio-dependent manner, and HIG2-9-4 peptide-specific CTLs produced a higher amount of IFN-γ than CTLs stimulated with other peptides (Figs. 1b–e). In the independent experiments using PBMCs from other 2 donors, HIG2-9-4 peptide-specific CTLs produced the highest amount of IFN-γ (data not shown). We confirmed the existence of HIG2-9-4/HLA-A*02:01-specific $CD8^+$ T cells by tetramer staining. A significant population of $CD3^+CD4^-CD8^+$ cells expressed the HIG2-9-4/HLA-A*02:01-specific T cell receptor after the expansion of cells obtained by *in vitro* CTL induction (Fig. 2).

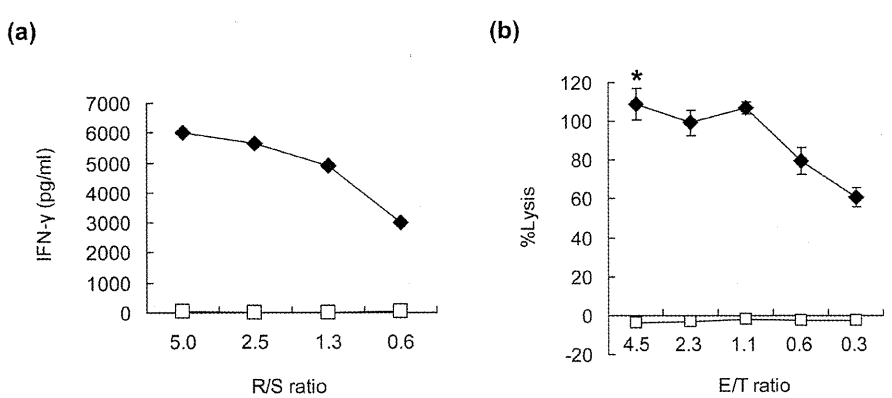

Figure 3. The IFN-γ production and cytotoxic activity of a HIG2-9-4 peptide-specific CTL clone. (a) An established CTL clone was stimulated with T2 cells pulsed with the HIG2-9-4 peptide (closed diamonds) or HIV peptide (open squares). The IFN-γ production in the culture supernatant was examined by ELISA. R/S ratio; responder/stimulator ratio. (b) The cytotoxic activity of the HIG2-9-4 peptide-specific CTL clone was examined against peptide-pulsed T2 cells (close diamond) or T2 cells pulsed with the HIV peptide (open square). E/T ratio; effector/target ratio. All experiments were performed in triplicate. The representative results from three independent experiments are shown. *$P<0.001$
doi:10.1371/journal.pone.0085267.g003

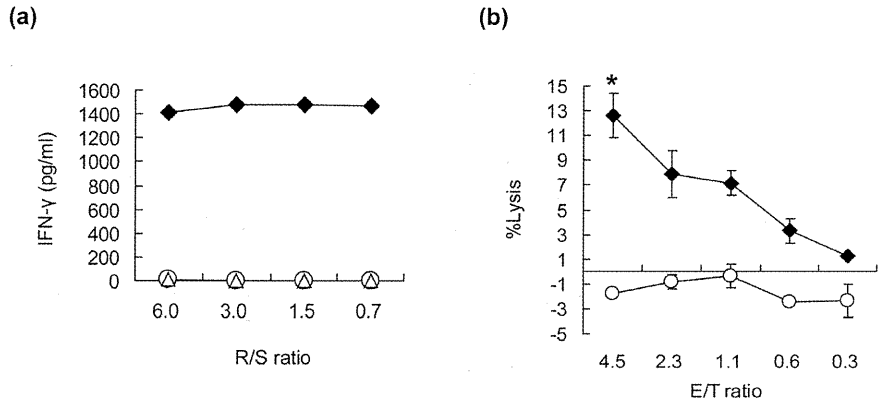

Figure 4. The recognition of HIG2 and HLA-A*02:01-expressing cells by a HIG2-9-4 peptide-specific CTL clone. (a) A HIG2-9-4 peptide-specific CTL clone was stimulated with COS7 cells expressing both HIG2 and HLA-A*02:01 (close diamond), or either HIG2 alone (open circle) or HLA-A*02:01 alone (open triangle), then the IFN-γ production was examined by ELISA. R/S ratio; responder/stimulator ratio. (b) The cytotoxic activity of the HIG2-9-4 peptide-specific CTL clone was examined against HLA-A*02:01-positive HIG2-expressing A498 cells (closed diamond) or HLA-A*02:01-negative HIG2-expressing Caki-1 cells (open circle). E/T ratio; effector/target ratio. All experiments were performed in triplicate. Representative results from three independent experiments are shown. *; $P<0.001$.
doi:10.1371/journal.pone.0085267.g004

Establishment of HIG2-9-4 peptide-specific CTL clones

We subsequently established HIG2-9-4 peptide-specific CTL clones by the limiting dilution of induced CTLs. The established HIG2-9-4 peptide-specific CTL clone produced a large amount of IFN-γ when it was stimulated with HIG2-9-4 pulsed-T2 cells, while no IFN-γ production was detected when they were stimulated with HIV-peptide-pulsed-T2 cells (Fig. 3a). Furthermore, the HIG2-9-4 peptide-specific CTL clone exerted substantial cytotoxic activity against T2 cells pulsed with the HIG2-9-4 peptide, but not those pulsed with the HIV peptide (Fig. 3b). However, we failed to establish any CTL clones that reacted with HIG2-9-8, HIG2-9-15 or HIG2-10-8 peptides, even after several attempts using multiple donors (data not shown). In addition, we found no homologous sequence to the HIG2-9-4 peptide by a homology search using the BLAST algorithm (data not shown), indicating that the HIG2-9-4 peptide is a unique epitope peptide among the candidate peptides predicted by the BIMAS program that can induce potent and stable CTLs.

Specific CTL response to HIG2 and HLA-A*02:01-expressing cells

To further verify the recognition of HIG2-expressing cells with HLA-A*02:01 by the HIG2-9-4-specific CTL clone, we prepared COS7 cells in which either or both of two plasmids designed to express the full-length of HIG2 and HLA-A*02:01 were transfected. The HIG2-9-4-specific CTL clone produced IFN-γ when the cells were exposed to the COS7 cells expressing both HIG2 and HLA-A*02:01, while no IFN-γ production was observed when they were exposed to COS7 cells expressing either HIG2 or HLA-A*02:01 (Fig. 4a). Furthermore, the HIG2-9-4 peptide-specific CTL clone demonstrated cytotoxic activity against A498 cells expressing both HLA-A*02:01 and HIG2, while no

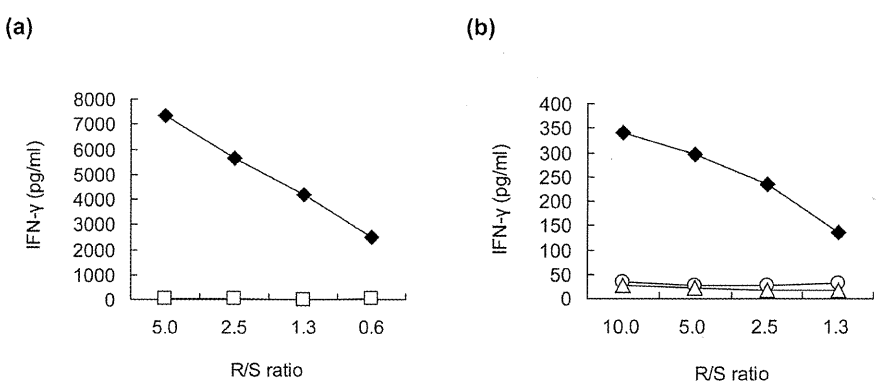

Figure 5. The HLA-A*02:06-restricted response of a HIG2-9-4 peptide-specific CTL clone. (a) A HIG2-9-4 peptide-specific CTL clone was induced from HLA-A*02:06-positive PBMCs, and stimulated with HLA-A*02:06-positive PSCCA0922 cells pulsed with the HIG2-9-4 peptide (close diamond) or HIV peptide (open square). (b) The HIG2-9-4 peptide-specific CTL clone was stimulated with COS7 cells expressing both HIG2 and HLA-A*02:06 (close diamond), or either HIG2 alone (open circle) or HLA-A*02:06 alone (open triangle). The IFN-γ production in the culture supernatant was examined by ELISA. R/S ratio; responder/stimulator ratio. The representative results from three independent experiments are shown.
doi:10.1371/journal.pone.0085267.g005

cytotoxicity was observed against HIG2-expressing Caki-1 cells without HLA-A*02:01 expression (Fig. 4b).

The HIG2-9-4 peptide cross-reacts with HLA-A*02:06

We additionally evaluated the cross-reactivity of the HIG2-9-4 peptide with HLA-A*02:06, since HLA-A*02:06 differs from HLA-A*02:01 by a single amino acid, and some reports have indicated the presentation of HLA-A*02:01-restricted peptides on HLA-A*02:06 [33,34]. Similar to the HLA-A*02:01 experiments, potent CTL clones were established from the PBMCs of HLA-A*02:06-positive donors by stimulation with the HIG2-9-4 peptide. An established CTL clone showed potent IFN-γ production when it was exposed to HIG2-9-4 peptide-pulsed HLA-A*02:06-positive PSCCA0922 cells (Fig. 5a). Furthermore, this CTL clone recognized COS7 cells that expressed both HIG2 and HLA-A*02:06 and produced IFN-γ, while no IFN-γ production was observed when stimulated with COS7 cells that expressed either HIG2 or HLA-A*02:06 (Fig. 5b). These results suggested that the HIG2-9-4 peptide is cross-reactive with HLA-A*02:06 to induce CTLs that show CTL activity against HLA-A*02:06- and HIG2-expressing cells.

Discussion

The recent FDA approvals of the cellular immunotherapy, Sipuleucel-T (Provenge), and immunomodulatory antibody, ipilimumab (Yervoy), have provided a proof of concept that the immune system can be used as a new approach to treat cancer [35,36]. Immunization with HLA-restricted epitope peptides derived from tumor antigens is a strategy that has been vigorously pursued to activate the immune system [37-40]. Unfortunately, many of the vaccine trials using epitope peptides failed to demonstrate clinical efficacy due, at least in part, to the potential immune escape mechanisms, which are attributed to the loss of tumor antigen expression by tumor cells [41-43]. Accordingly, the selection of tumor antigens which play a key role in tumor cell proliferation or survival is considered to be important to overcome immune escape. If a targeted tumor antigen is essential for tumor growth, the downregulation of this tumor antigen as a form of immune escape is expected to impair tumor progression.

Correspondingly, in the guidelines from the FDA (Guidance for Industry: Clinical Considerations for Therapeutic Cancer Vaccines), multi-antigen vaccines which contain multiple tumor antigens in order to generate multiple tumor-specific immunological responses were mentioned to effectively hinder escape mechanisms. We therefore consider that the identification of epitope peptides derived from multiple tumor antigens which are involved in tumor progression or survival can contribute to the development of multi-antigen vaccines, and can improve the efficacy of peptide vaccine therapies. We have previously identified epitope peptides derived from various tumor antigens, each of which plays a key role in tumor progression, and some of these peptides have been applied for clinical trials as multi-peptide vaccines [44–46].

In this study, we identified an HLA-A2 supertype-restricted epitope peptide derived from HIG2. HIG2 was upregulated in RCC and hardly detectable in normal organs except for the fetal kidney, and importantly, HIG2 expression was found to be directly associated with the proliferation of RCC cells [24]. Hence, RCC cells are thought to maintain HIG2 expression even under immunoselective pressure, or to otherwise exhibit tumor growth suppression resulting from the loss of HIG2 expression.

IFN-γ-producing stable CTL clones specific to the HIG2-9-4 peptide (VLNLYLLGV) were established from HLA-A2 (either A*02:01 or A*02:06)-positive PBMCs, and these clones responded specifically to COS7 cells that expressed both HIG2 and HLA-A2 (A*02:01 or A*02:06). We also revealed that HIG2-9-4-specific HLA-A*02:01-restricted CTLs exerted cytotoxic activity against RCC cells that were positive for both HIG2 and HLA-A*02:01, but not against negative cells. These results suggested that HLA-A2 (A*02:01 or A*02:06)-restricted HIG2-9-4 peptide-specific CTLs are inducible and stable, and these CTLs substantially respond to HIG2-expressing cells through the endogenous processing of the HIG2-9-4-peptide and the subsequent presentation with the HLA-A2 (A*02:01 or A*02:06) molecule on the cell surface. In addition, HIG2 is an oncofetal antigen, as described above, and no homologous sequence to the HIG2-9-4 peptide was demonstrated by a homology search using the BLAST algorithm. Thus, HIG2-9-4 peptide-specific CTLs should not induce unintended immunological responses to normal cells, such as those associated with autoimmune diseases, even if this novel and unique peptide induces strong immune responses against HIG2-expressing RCC.

HIG2 expression was found in the majority of RCC patients (86%) [25], and additionally, the HLA-A2 supertype is the most common HLA class I type in Caucasians and the second most common type in the Japanese population [26,27]. Therefore, identification of HLA-A2 supertype-restricted epitope peptides derived from HIG2 could be applicable for immunotherapies in a wide variety of RCC patients. As well as finding novel tumor antigens which are widely expressed in cancer patients, finding epitope peptides restricted to major HLA Class I types will facilitate further development of cancer immunotherapies. We are now conducting clinical trials to examine the immunogenicity and safety of a HIG2-9-4 peptide vaccine in RCC patients.

Supporting Information

Figure S1 Response to the HIG2-9-8, HIG2-9-15, HIG2-9-4 or HIG2-10-8 peptide detected by IFN-γ ELISPOT assay. The IFN-γ production from cells induced by the indicated peptide-pulsed DCs in 12 wells for each peptide was examined by an ELISPOT assay. "+" indicates the wells in which cells were stimulated with T2 cells pulsed with the indicated peptide and "−" indicates the wells in which cells were stimulated with HIV peptide-pulsed T2 cells. The wells in which the difference between peptide-pulsed cells and HIV peptide-pulsed cells were over 50 spots are indicated by squares.
(TIF)

Acknowledgments

The authors thank Dr. Kawakami (Keio University, Tokyo, Japan) for providing the expression vector, and the Cell Resource Center for Biomedical Research Institute of Development, Aging and Cancer at Tohoku University for providing the cell lines.

Author Contributions

Conceived and designed the experiments: TT RO HY. Performed the experiments: SY MH TW TH. Analyzed the data: SY MH TW TH. Wrote the paper: SY. Scientific advise: MK MM MT MI. Support to draft the manuscript: KT TK YN.

References

1. Jemal A, Siegel R, Xu J, Ward E (2010) Cancer statistics, 2010. CA Cancer J Clin 60: 277–300.
2. Cohen HT, McGovern FJ (2005) Renal-cell carcinoma. N Engl J Med 353: 2477–2490.
3. National Comprehensive Cancer Network (2012) NCCN Clinical Practice Guidelines in Oncology. Kidney Cancer. Version2.2012. Available: http://www.tri-kobe.org/nccn/guideline/urological/english/kidney.pdf
4. Patil S, Ishill N, Deluca J, Motzer RJ (2010) Stage migration and increasing proportion of favorable-prognosis metastatic renal cell carcinoma patients: implications for clinical trial design and interpretation. Cancer 116: 347–354.
5. Rini BI, Escudier B, Tomczak P, Kaprin A, Szczylik C, et al. (2011) Comparative effectiveness of axitinib versus sorafenib in advanced renal cell carcinoma (AXIS): a randomised phase 3 trial. Lancet 378: 1931–1939.
6. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, et al. (2007) Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. N Engl J Med 356: 115–124.
7. Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, et al. (2009) Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol 27: 3584–3590.
8. Escudier B, Szczylik C, Hutson TE, Demkow T, Staehler M, et al. (2009) Randomized phase II trial of first-line treatment with sorafenib versus interferon Alfa-2a in patients with metastatic renal cell carcinoma. J Clin Oncol 27: 1280–1289.
9. Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, et al. (2007) Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med 356: 2271–2281.
10. Escudier B, Pluzanska A, Koralewski P, Ravaud A, Bracarda S, et al. (2007) Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. Lancet 370: 2103–2111.
11. Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, et al. (2008) Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. Lancet 372: 449–456.
12. Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, et al. (2010) Phase 3 trial of everolimus for metastatic renal cell carcinoma: final results and analysis of prognostic factors. Cancer 116: 4256–4265.
13. Sternberg CN, Davis ID, Mardiak J, Szczylik C, Lee E, et al. (2010) Pazopanib in locally advanced or metastatic renal cell carcinoma: results of a randomized phase III trial. J Clin Oncol 28: 1061–1068.
14. McDermott DF (2009) Immunotherapy of metastatic renal cell carcinoma. Cancer 115: 2298–2305.
15. Sparano JA, Fisher RI, Sunderland M, Margolin K, Ernest ML, et al. (1993) Randomized phase III trial of treatment with high-dose interleukin-2 either alone or in combination with interferon alfa-2a in patients with advanced melanoma. J Clin Oncol 11: 1969–1977.
16. Atzpodien J, Kirchner H, Jonas U, Bergmann L, Schott H, et al. (2004) Interleukin-2- and interferon alfa-2a-based immunochemotherapy in advanced renal cell carcinoma: a Prospectively Randomized Trial of the German Cooperative Renal Carcinoma Chemoimmunotherapy Group (DGCIN). J Clin Oncol 22: 1188–1194.
17. Fyfe G, Fisher RI, Rosenberg SA, Sznol M, Parkinson DR, et al. (1995) Results of treatment of 255 patients with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. J Clin Oncol 13: 688–696.
18. Bleumer I, Tiemessen DM, Oosterwijk-Wakka JC, Völler MC, De Weijer K, et al. (2007) Preliminary analysis of patients with progressive renal cell carcinoma vaccinated with CA9-peptide-pulsed mature dendritic cells. J Immunother 30: 116–122.
19. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT (2002) Cancer/testis antigens: an expanding family of targets for cancer immunotherapy. Immunol Rev 188: 22–32.
20. Tsuji T, Matsuzaki J, Kelly MP, Ramakrishna V, Vitale L, et al. (2011) Antibody-targeted NY-ESO-1 to mannose receptor or DEC-205 in vitro elicits dual human CD8+ and CD4+ T cell responses with broad antigen specificity. J Immunol 186: 1218–1227.
21. Amato RJ, Shingler W, Goonewardena M, de Belin J, Naylor S, et al. (2009) Vaccination of renal cell cancer patients with modified vaccinia Ankara delivering the tumor antigen 5T4 (TroVax) alone or administered in combination with interferon-α (IFN-α): a Phase 2 trial. J Immunother 32: 765–772.
22. Denko N, Schindler C, Koong A, Laderoute K, Green C, et al. (2000) Epigenetic regulation of gene expression in cervical cancer cells by the tumor microenvironment. Clin Cancer Res 6: 480–487.
23. Gimm T, Wiese M, Teschemacher B, Deggerich A, Schödel J, et al. (2010) Hypoxia-inducible protein 2 is a novel lipid droplet protein and a specific target gene of hypoxia-inducible factor-1. FASEB J 24: 4443–4458.
24. Togashi A, Katagiri T, Ashida S, Fujioka T, Maruyama O, et al. (2005) Hypoxia-inducible protein 2 (HIG2), a novel diagnostic marker for renal cell carcinoma and potential target for molecular therapy. Cancer Res 65: 4817–4826.
25. Seo T, Konda R, Sugimura J, Iwasaki K, Nakamura Y, et al. (2010) Expression of hypoxia-inducible protein 2 in renal cell carcinoma: A promising candidate for molecular targeting therapy. Oncol Lett 1: 697–701.
26. Cao K, Hollenbach J, Shi X, Shi W, Chopek M, et al. (2001) Analysis of the frequency of HLA-A, B and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. Hum Immunol 62: 1009–1030.
27. Itoh Y, Mizuki N, Shimada T, Azuma F, Itakura M, et al. (2005) High-throughput DNA typing of HLA-A, -B, -C, and -DRB1 loci by a PCR-SSOP-Luminex method in the Japanese population. Immunogenetics 57: 717–729.
28. Tsomides TJ, Aldovini A, Johnson RP, Walker BD, Young RA, et al. (1994) Naturally processed viral peptides recognized by cytotoxic T lymphocytes on cells chronically infected by human immunodeficiency virus type 1. J Exp Med 180: 1283–1293.
29. Celis E, Tsai V, Crimi C, DeMars R, Wentworth PA, et al. (1994) Induction of anti-tumor cytotoxic T lymphocytes in normal humans using primary cultures and synthetic peptide epitopes. Proc Natl Acad Sci U S A 91: 2105–2109.
30. Uchida N, Tsunoda T, Wada S, Furukawa Y, Nakamura Y, et al. (2004) Ring finger protein (RNF) 43 as a New Target for Cancer Immunotherapy. Clin Can Res 10: 8577–8586.
31. Suda T, Tsunoda T, Daigo Y, Nakamura Y, Tahara H (2007) Identification of human leukocyte antigen-A24-restricted epitope peptides derived from gene products upregulated in lung and esophageal cancers as novel targets for immunotherapy. Cancer Sci 98: 1803–1808.
32. Takeda K, Yamaguchi N, Akiba H, Kojima Y, Hayakawa Y, et al. (2004) Induction of tumor-specific T cell immunity by anti-DR5 antibody therapy. J Exp Med 199: 437–448.
33. Sidney J, Southwood S, Mann DL, Fernandez-Vina MA, Newman MJ, et al. (2001) Majority of peptides binding HLA-A*0201 with high affinity crossreact with other A2-supertype molecules. Hum Immunol 62: 1200–1216.
34. Fleischhauer K, Tanzarella S, Russo V, Sensi ML, van der Bruggen P, et al. (1997) Functional heterogeneity of HLA-A*02 subtypes revealed by presentation of a MAGE-3-encoded peptide to cytotoxic T cell clones. J Immunol 159: 2513–2521.
35. Kantoff PW, Higano CS, Shore ND, Berger ER, Small EJ, et al. (2010) Sipuleucel-T immunotherapy for castration-resistant prostate cancer. N Engl J Med 363: 411–422.
36. Robert C, Thomas L, Bondarenko I, O'Day S, M D JW, et al. (2011) Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. N Engl J Med 364: 2517–2526.
37. Boon T (1993) Tumor antigens recognized by cytolytic T lymphocytes: present perspectives for specific immunotherapy. Int J Cancer 54: 177–180.
38. Rimoldi D, Rubio-Godoy V, Dutoit V, Lienard D, Salvi S, et al. (2000) Efficient simultaneous presentation of NY-ESO-1/LAGE-1 primary and nonprimary open reading frame-derived CTL epitopes in melanoma. J Immunol 165: 7253–7261.
39. Rosenberg SA, Yang JC, Schwartzentruber DJ, Hwu P, Topalian SL, et al. (2003) Recombinant fowlpox viruses encoding the anchor-modified gp100 melanoma antigen can generate antitumor immune responses in patients with metastatic melanoma. Clin Cancer Res 9: 2973–2980.
40. Pecher G, Häring A, Kaiser L, Thiel E (2002) Mucin gene (MUC1) transfected dendritic cells as vaccine: results of a phase I/II clinical trial. Cancer Immunol Immunother 51: 669–673.
41. Rosenberg SA, Yang JC, Restifo NP (2004) Cancer immunotherapy: moving beyond current vaccines. Nat Med 10: 909–915.
42. Sampson JH, Heimberger AB, Archer GE, Aldape KD, Friedman AH, et al. (2010) Immunologic escape after prolonged progression-free survival with epidermal growth factor receptor variant III peptide vaccination in patients with newly diagnosed glioblastoma. J Clin Oncol 28: 4722–4729.
43. DuPage M, Mazumdar C, Schmidt LM, Cheung AF, Jacks T (2012) Expression of tumour-specific antigens underlies cancer immunoediting. Nature 482: 405–409.
44. Okuno K, Sugiura F, Hida JI, Tokoro T, Ishimaru E, et al. (2011) Phase I clinical trial of a novel peptide vaccine in combination with UFT/LV for metastatic colorectal cancer. Exp Ther Med 2: 73–79.
45. Kono K, Iinuma H, Akutsu Y, Tanaka H, Hayashi N, et al. (2012) Multicenter, phase II clinical trial of cancer vaccination for advanced esophageal cancer with three peptides derived from novel cancer-testis antigens. J Transl Med 10: 141.
46. Obara W, Ohsawa R, Kanehira M, Takata R, Tsunoda T, et al. (2012) Cancer peptide vaccine therapy developed from oncoantigens identified through genome-wide expression profile analysis for bladder cancer. Jpn J Clin Oncol 42: 591–600.

Involvement of B3GALNT2 overexpression in the cell growth of breast cancer

TAISUKE MATSUO[1], MASATO KOMATSU[1], TETSURO YOSHIMARU[1], KAZUMA KIYOTANI[1], YASUO MIYOSHI[2], MITSUNORI SASA[3] and TOYOMASA KATAGIRI[1]

[1]Division of Genome Medicine, Institute for Genome Research, The University of Tokushima; [2]Department of Surgery, Division of Breast and Endocrine Surgery, Hyogo College of Medicine, Hyogo 663-8501; [3]Department of Surgery, Tokushima Breast Care Clinic, Tokushima 770-0052, Japan

Received August 24, 2013; Accepted October 7, 2013

DOI: 10.3892/ijo.2013.2187

Abstract. A number of glycosyltransferases have been identified and biologically characterized in cancer cells, yet their exact pathophysiological functions are largely unknown. Here, we report the critical role of β1,3-*N*-acetylgalactosaminyltransferase II (*B3GALNT2*), which transfers N-acetylgalactosamine (GalNAc) in a β1,3 linkage to N-acetylglucosamine, in the growth of breast cancer cells. Comprehensive transcriptomics, quantitative PCR and northern blot analyses indicated this molecule to be exclusively upregulated in the majority of breast cancers. Knockdown of *B3GALNT2* expression by small interfering RNA attenuated cell growth and induced apoptosis in breast cancer cells. Overexpression of B3GALNT2 in HEK293T cells prompted secretion of the gene product into the culture medium, suggesting that B3GALNT2 is potentially a secreted protein. Furthermore, we demonstrated that B3GALNT2 is N-glycosylated on both Asn-116 and Asn-174 and that this modification is necessary for its secretion in breast cancer cells. Our findings suggest that this molecule represents a promising candidate for the development of a novel therapeutic targeting drug and a potential diagnostic tumor marker for patients with breast cancer, especially TNBC.

Introduction

Breast cancer is a highly heterogeneous disease that is currently classified by the expression profiling of estrogen receptor (ER), progesterone receptor (PgR) and the human epidermal growth factor receptor 2 (HER2) (1,2). Endocrine therapies such as tamoxifen and aromatase inhibitor have produced a significant improvement in outcomes for patients with ER-positive breast cancer. The HER2-targeting therapies, such as trastuzumab and lapatinib, have significantly improved the outlook for patients with HER2-positive breast cancer. However, increased risk of endometrial cancer with long-term tamoxifen administration and of bone fracture due to osteoporosis in postmenopausal women undergoing aromatase inhibitor treatment are recognized side effects (3-5). In addition, triple negative breast cancer (TNBC), defined as tumors that are characterized by lack of ER, PgR and HER2, accounts for ~15% of all breast cancers and shows significantly poorer prognosis compared with other types of breast cancers because of a lack of clinically established targeted therapies (6,7). Due to the emergence of these side effects, endocrine-resistant and chemo-resistant breast cancers and TNBC, it is necessary to search for novel molecular targets for drugs based on well-characterized mechanisms of action.

Current 'omics' technologies, including transcriptomics, are a very useful approach for identifying novel therapeutic targets for various cancers, including breast cancer (8-10). We previously used DNA microarray to analyze the genome-wide gene expression profiles of TNBCs and normal human vital organs including heart, lung, liver and kidney (10). After comparing the expression profiles of TNBCs and normal human tissues, we focused on β1,3-*N*-acetylgalactosaminyltransferase II (*B3GALNT2*), which was significantly upregulated in TNBCs compared with normal breast ducts (10). B3GALNT2 was first identified as a novel glycosyltransferase having β1,3-glycosyltranferase motifs, which are highly conserved in β1,3-galactosyltranferase and β1,3-*N*-acetylglucosaminyltranferase families, using a BLAST search (11). The purified putative catalytic domain of this protein reportedly has *N*-acetylgalactosaminyltransferase *in vitro* activity and β1,3-linkage as determined by NMR spectroscopic analysis. However, to date, no reports have characterized the biologic function of B3GALNT2 or the significance of its transactivation in clinical breast cancer cell growth.

Glycosylation plays crucial roles in a variety of biological functions, such as cell-cell and cell-substrate interactions, differentiation and signal transduction in tumor cells (12). The glycoproteins of tumor cells are often aberrant, both in structure and in quantity, leading to abnormal biological

Correspondence to: Dr Toyomasa Katagiri, Division of Genome Medicine, Institute of Genome Research, The University of Tokushima, 3-18-15 Kuramoto-cho, Tokushima 770-8503, Japan
E-mail: tkatagi@genome.tokushima-u.ac.jp

Key words: glycosyltransferase, molecular target, breast cancer, secreted protein

functions, including cell proliferation, migration, invasion and transformation (13-15). Although these alterations are caused by the dysregulated expression or structure of specific glycosyltransferase (16), their overview remains poorly understood.

In this study, we show that *B3GALNT2* is overexpressed in breast cancers including TNBC and that downregulation of *B3GALNT2* results in a significant reduction of breast cancer cell growth due to apoptosis induction. Moreover, we demonstrate that overexpression of B3GALNT2 results in its secretion into culture medium. Our findings suggest that B3GALNT2 represents a promising candidate for the development of molecular targeting therapy and might be a suitable diagnostic marker for breast cancer.

Materials and methods

Cell lines and specimens. Human breast cancer cell lines, BT-20, HCC1143, HCC1395, HCC1599, MCF-7, MDA-MB-453, OCOB-F, T47D and ZR-75-1 and human embryonic kidney fibroblast HEK293T cells were obtained from the American Type Culture Collection (ATCC, Rockville, MD, USA). BSY-1 cell line was a kind gift from Dr Takao Yamori of the Division of Molecular Pharmacology, Cancer Chemotherapy Center, Japanese Foundation for Cancer Research. All cells were cultured under conditions recommended by the ATCC as previously described (17). We monitored the cell morphology of these cell lines by microscopy and confirmed that they had maintained their morphologic states in comparison with the original morphologic images. No mycoplasma contamination was detected in the cultures of any of these cell lines using a Mycoplasma Detection kit (Takara, Kyoto, Japan) in 2011. A total of 30 TNBCs and 13 normal mammary tissues were obtained with informed consent from patients who were treated at the Tokushima Breast Care Clinic, Tokushima, Japan, as previously described (10). This study and the use of all clinical materials described above, was approved by the Ethics Committee of The University of Tokushima.

Reverse transcription and real-time PCR. Total RNA from clinical breast cancer samples and breast cancer cell lines was isolated using a NucleoSpin RNA II (Takara) according to the manufacturer's instructions. The poly A-RNA of normal human heart, liver, kidney, lung and mammary gland (MG) (Takata Clontech) was reverse transcribed as described previously (17,18). Real-time PCR analysis was performed using Power SYBR Green PCR Master Mix (Life Technologies, Carlsbad, CA, USA) using an ABI PRISM 7500 Real-Time PCR system (Life Technologies) according to the manufacturer's instructions. Gene-specific primers used for real-time PCR were as follows: 5'-AAGACCTGTGAGACAGGAATGC-3' and 5'-GTTCTGGGTGAAAGTGCCAG-3' for *B3GALNT2* and 5'-ATTGCCGACAGGATGCAG-3' and 5'-CTCAGGAGGA GCAATGATCTT-3' for *ACTB* as a quantitative control.

Northern blot analysis. Isolation of mRNA from breast cancer cell lines was performed using mRNA Purification Kit 4 (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's instructions. The northern blot for breast cancer cell lines was prepared as described previously (17). The breast cancer blot and human multiple tissue blots (MTN and MTN II) were hybridized with [α^{32}P]-dCTP labeled PCR products of *B3GALNT2* by RT-PCR using a primer set as follows: 5'-TGATGTGGTAGTTGGCGTGT-3' and 5'-AGAACTCCCCCTCCATCATT-3'. Prehybridization, hybridization and washing were performed as described previously (17). The blots were autoradiographed with intensifying screens at -80°C for 14 days.

Gene silencing effect by siRNA treatment. BT-20, MDA-MB-453 and ZR-75-1 cells were plated onto 12-well plates ($5x10^4$, $5x10^4$ and $2.5x10^4$ cells/well, respectively). Transfection of 10 nM siRNAs was performed using Lipofectamine RNAiMax (Life Technologies) according to the manufacturer's instructions. The target sequences for three *B3GALNT2* and a control *EGFP* siRNAs were 5'-GUCAACGUGUGCUUGUG AA-3' for siB3GALNT2-1, 5'-CGAGCUUCAAUUUGUUGC U-3' for siB3GALNT2-2, 5'-CCGGAAAGUGGCAGGAG UU-3' for siB3GALNT2-3 and 5'-GCAGCACGACUUCUUC AAG-3' for siEGFP. To evaluate the knockdown effect of the siRNAs by quantitative RT-PCR, total RNA was extracted from the siRNA-transfected cells at 6 days after siRNA transfection. The gene-specific primers are described above. These experiments were performed in duplicate. To quantify cell viability, MTT assays were performed using a Cell-Counting Kit-8 according to the manufacturer's recommendations (Dojindo, Kumamoto, Japan). Absorbance at 450 nm was measured with the microplate reader Infinite 200 (Tecan, Männedorf, Switzerland). These experiments were performed in duplicate (MDA-MB-453 and ZR-75-1 cells) or triplicate (BT-20 cells).

Plasmids. To construct the B3GALNT2 expression vectors, the entire coding sequence of *B3GALNT2* cDNA was amplified by RT-PCR using KOD plus DNA polymerase (Toyobo, Osaka, Japan) and cloned into the pCAGGSn3FC expression vector in frame with Flag-tag at the COOH terminus. The primer sets of B3GALNT2-wild-type (WT) was as follows: 5'-ATAAG AATGCGGCCGCATGCGAAACTGGCTGGTGC-3' and 5'-CCGCTCGAGTCTTGCTTGACATCGACAAGG-3' (the underlined letters indicate the *Not*I or *Xho*I sites, respectively). We also performed conventional two-step mutagenesis PCR to generate mutants in which N116A and N174A were substituted to alanines, as described previously (14). The primer sets were 5'-GTAAACTACTCGCCATCACAAATC-3' and 5'-GATTT GTGATGGCGAGTAGTTTAC-3' for N116A, 5'-GTTTCCA GAGGGCCATCACTGTC-3' and 5'-GACAGTGATGGCCC TCTGGAAAC-3' for N174A (the underlined letters indicate the mutation sites, respectively). The DNA sequences of all constructs were confirmed by DNA sequencing (ABI3500XL, Life Technologies).

Fluorescence activated cell-sorting (FACS) analysis. FACS analysis was performed as previously described (10). Briefly, the cells were collected at 2, 4 and 6 days after treatment of siRNAs against *B3GALNT2* or *EGFP* as a control. The cells were fixed by 70% ethanol at room temperature for 30 min and then incubated at 37°C for 30 min with 1 mg/ml RNase A, followed by staining with 20 μg/ml propidium iodide at room temperature for 30 min in the dark. The DNA content of 10,000 cells was analyzed with a FACSCalibur flow cytometer and CellQuest software (BD Biosciences, Franklin Lakes, NJ, USA).

Western blot analysis. Western blot analysis was performed as previously described (19). Briefly, the cells were lysed in lysis buffer [50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.5% Nonidet P-40, 0.5% CHAPS] including 0.1% protease inhibitor cocktail III (Calbiochem, San Diego, CA, USA). The cell lysates were incubated on ice for 30 min and centrifuged at 15,000 rpm for 15 min to remove cell debris. Then, the proteins were mixed with SDS sample buffer [25 mM Tris-HCl (pH 6.8), 0.8% SDS, 5% glycerol] and boiled for 5 min. After SDS-PAGE, membranes blotted with proteins were incubated with anti-Flag M2 (Sigma-Aldrich, St. Louis, MO, USA, F3165) or anti-β-actin (AC-15, Sigma-Aldrich, A-5441) monoclonal antibodies diluted at 1:5,000 and PARP rabbit polyclonal antibody (Cell Signaling Technology, 9542) diluted at 1:500, respectively. Finally, the membrane was incubated with horseradish peroxidase (HRP)-conjugated secondary antibody for 1 h and protein bands were visualized by enhanced chemiluminescence detection reagents (ECL, GE Healthcare).

Immunocytochemical staining. Immunocytochemical staining was performed as previously described (10). To examine the subcellular localization of the B3GALNT2 protein, HEK293T cells were plated onto an 8-well glass slide (Thermo Fisher Scientific, Rochester, NY, USA) at a density of $1.0x10^4$ cells/well and transfected with 0.2 μg each of expression plasmids using FuGENE 6 reagent (Promega, Madison, WI, USA) according to the manufacturer's recommendations. To detect exogenous B3GALNT2-Flag, anti-Flag M2 mouse antibody was used at 1:1,000 and Alexa 488-conjugated anti-mouse antibody. The Golgi apparatus were visualized by staining with anti-Golgi-58k mouse monoclonal antibody (Sigma-Aldrich) (20). Cell morphology was analyzed by Alexa Fluor 488 phalloidin (Molecular Probes, Eugene, OR, USA) diluted at 1:1,000.

Inhibition of N-glycosylation. HEK293T cells were plated onto 6-well dishes at a density of $2.0x10^5$ cells/well and transfected with 1 μg each of expression plasmids using FuGENE 6 reagent according to the manufacturer's instructions. To validate the N-glycosylation of B3GALNT2, the cells were cultured with 10 μg/ml tunicamycin (Sigma-Aldrich), an inhibitor against N-glycosylation, for 24 h, at 4 h after transfection. Cells were then lysed with SDS sample buffer.

Preparation of secreted B3GALNT2. To detect secreted B3GALNT2 protein, we transfected B3GALNT2-expressing plasmids to HEK293T cells as described above. The medium was changed to DMEM with 0.1% FBS at 48 h after transfection. Cells were cultured for a further 48 h and then supernatants were collected. After removing debris by centrifugation at 15,000 rpm for 15 min, we performed acetone precipitation at -30°C overnight. Then, the pellets were resuspended with SDS sample buffer.

Statistical analysis. Statistical analysis was conducted using Student's t-test. A P-value of <0.05 was considered to be statistically significant. Box blot analysis was performed with StatView J 5.0 software using the microarray data of 13 normal ducts and 30 TNBCs [the Gene Expression Omnibus database accession no. GSE38959; (10)].

Results

B3GALNT2 is upregulated in breast cancer specimens and cell lines. To identify the novel therapeutic targets for TNBC therapy, we previously performed genome-wide gene expression profile analysis of TNBC and normal human tissues by DNA microarray analysis (10). In parallel with this approach, we attempted to search for the genes that encode proteins containing glycosyltransferase motifs, either based on reported information or according to prediction by the protein-motif program SMART because we previously succeeded in identifying and characterizing the cancer-specific glycosyltransferase GALNT6, that plays a critical role for mammary carcinogenesis, as a drug target (14,15). Among the glycosyltransferase genes that are upregulated in expression profiles of breast cancers, we focused on the *B3GALNT2* gene, which encodes a glycosyltransferase, as a drug target for breast cancer. Gene expression profiling analysis showed significant overexpression of *B3GALNT2* in TNBC cases, compared with normal ductal cells (Fig. 1A). Using quantitative RT-PCR, we verified that *B3GALNT2* was upregulated >2-fold in 5 of 10 TNBC samples compared with clinically normal breast tissues, whereas *B3GALNT2* was hardly expressed in the heart, lung, liver, kidney and mammary gland (Fig. 1B). Subsequent northern blot analysis of 10 breast cancer cell lines detected an approximately 4.7-kb *B3GALNT2* transcript at a high level in 8 of the 10 breast cancer cell lines that were examined (Fig. 1C).

Furthermore, multiple tissue northern blot analysis revealed that a 4.7-kb transcript of *B3GALNT2* was slightly expressed in heart, skeletal muscle, spleen, kidney, liver, small intestine, placenta, lung, prostate and ovary, while a 2.4-kb transcript was exclusively expressed in testis (Fig. 1D). According to the National Center for Biotechnology Information (NCBI) database, two representative transcripts of 4,755 nucleotides (*B3GALNT2-V1*, GenBank accession no. NM_152490) and 2,022 nucleotides (*B3GALNT2-V2*, GenBank accession no. BC029564) that share the same open reading frame encoding a 500-amino acid protein seemed to correspond to the two bands observed in multiple tissue northern blot analysis (Fig. 1E).

Effect of B3GALNT2 on cell growth. To assess whether *B3GALNT2* is essential for the growth or survival of breast cancer cells, we transfected synthetic oligonucleotide siRNAs against *B3GALNT2* into the TNBC cell line, which consisted of BT-20 cells in which *B3GALNT2* was highly expressed. The mRNA levels of *B3GALNT2* in the cells transfected with siB3-1, -2 or -3 were significantly downregulated in comparison with cells transfected with siEGFP as a control (Fig. 2A). We also observed a significant decrease in the number of viable cells measured by MTT assay (Fig. 2A). Similarly, silencing of *B3GALNT2* expression by siB3-3 markedly decreased cell proliferation of the non-TNBC cell lines MDA-MB-453 and ZR-75-1, in which *B3GALNT2* was highly expressed (Fig. 2B and C).

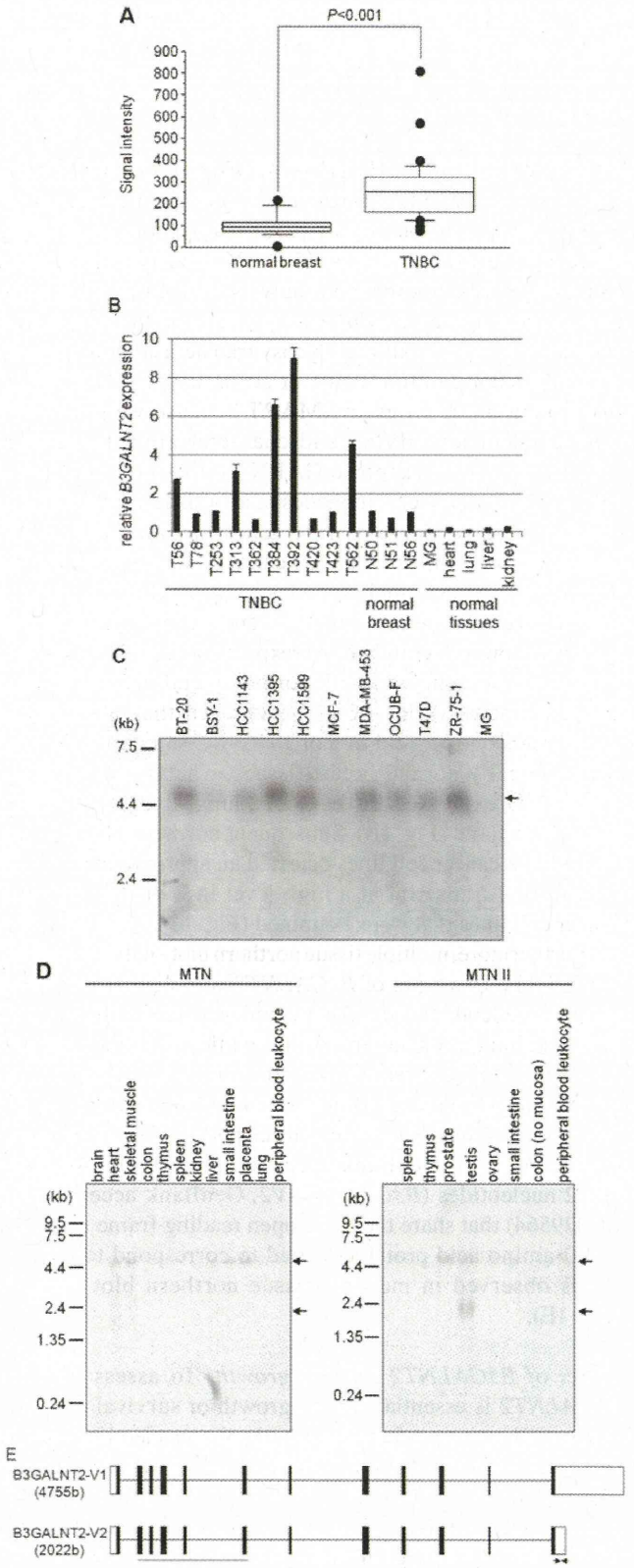

Figure 1. *B3GALNT2* overexpression in breast cancer cells. (A) Box blot analysis of *B3GALNT2* gene expression using microarray data from microdissected cells of 13 normal ductal tissues and 30 TNBC tissues. (B) Real-time RT-PCR result of *B3GALNT2* in TNBC (T), normal breast ducts (N) and normal tissues (n=2). (C) Northern blot analysis of *B3GALNT2* in 10 breast cancer cell lines and a mammary gland (MG). (D) Multiple human normal tissue northern blot analysis of *B3GALNT2*. (E) Genomic structure of *B3GALNT2* variants. Black and white boxes indicate coding and non-coding regions, respectively. The arrows indicate a primer set for real-time RT-PCR. The line shows a probe position for northern blot analysis.

To further assess the knockdown effect of *B3GALNT2*, we performed FACS analysis and found that the percentage of sub-G1 population was clearly increased in a time-dependent manner in *B3GALNT2*-depleted BT-20 cells (Fig. 3A). In addition, we observed an obvious cleaved PARP at day 4 after treatment of siRNA against *B3GALNT2* (siB3-3) (Fig. 3B). Interestingly, 4 days after transfection of siB3-3, severe disrupted cytoskeletal organization was observed by immunocytochemistry with fluorescence-labeled phalloidin (Fig. 3C), suggesting that depletion of *B3GALNT2* resulted in suppression of breast cancer cell growth due to apoptosis. These findings suggest that *B3GALNT2* is crucial for both TNBC and non-TNBC cell growth.

Secretory nature and N-glycosylation of B3GALNT2. The *B3GALNT2* gene is reported to encode a type II transmembrane enzyme, which possesses a single transmembrane domain, a stem region and a C-terminal catalytic domain for enzyme activity (21). To first investigate the subcellular localization of B3GALNT2 in mammalian cells, we transiently transfected the Flag-tagged B3GALNT2 construct (B3GALNT2-Flag) into HEK293T cells and then performed immunocytochemical staining analysis. B3GALNT2-Flag was observed to have highly intense staining in the Golgi apparatus, but it was also diffusely observed in cytoplasm (Fig. 4A). Furthermore, because many type II glycosyltransferases are found as secreted soluble enzymes through proteolytic cleavage of the stem region (22), we hypothesized that B3GALNT2 has a secretory nature. To investigate this possibility, HEK293T cells were transiently transfected with B3GALNT2-Flag and then western blot analysis with anti-Flag antibody was performed using cell lysates and culture media. We detected a band of B3GALNT2-Flag in both cell lysates and culture media, but its molecular weight in culture media was smaller than that in cell lysate, suggesting the possibility that the B3GALNT2 protein was cleaved during its secretion into culture media (Fig. 4B).

N-glycosylation on many glycosyltransferases is known to be associated with its biological functions, especially its secretion (22,23). In addition, the NetNGlyc 1.0 server (http://cbs.dtu.dk/services/NetNGlyc) and NCBI database predict that B3GALNT2 possesses two potential N-linked glycosylation at positions Asn-116 and Asn-174 that are completely conserved among other species, such as mouse, rat and *Xenopus laevis* (Fig. 4C). To confirm whether B3GALNT2 is N-glycosylated, we treated HEK293T cells transfected with B3GALNT2-wild-type (B3GALNT2-WT) with tunicamycin, which is an N-glycosylation inhibitor and then analyzed the molecular weight of B3GALNT2-WT protein by western blot analysis. As expected, tunicamycin treatment resulted in a size reduction of B3GALNT2-WT (Fig. 4D, WT). Next, to assess the N-glycosylated amino acids of B3GALNT2, we constructed the expression plasmids of N116A or N174A single mutation and N116AN174A double mutation, in which the conserved asparagine residues at position 116 or 174 were replaced by alanine residues. When HEK293T cells were transfected with these plasmids without tunicamycin treatment, the molecular weight of N116A and N174A was smaller than that of B3GALNT2-WT, respectively and the molecular weight of N116AN174A double mutant was signifi-

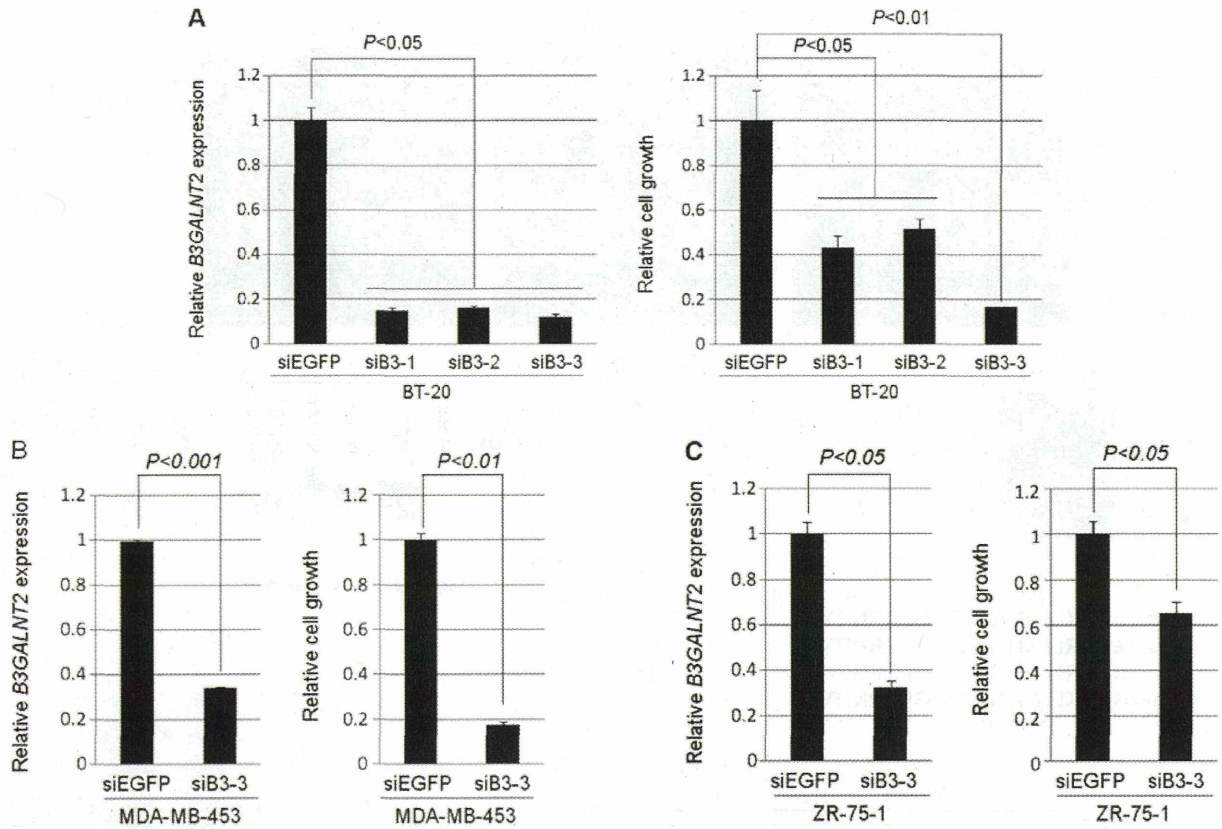

Figure 2. The knockdown effects of *B3GALNT2* on cell proliferation by siRNA in BT-20 (A), MDA-MB-453 (B) and ZR-75-1 (C) cells. Real-time PCR of *B3GALNT2* in siEGFP- or siB3GALNT2-treated cells (n=2). *ACTB* was used as a quantitative control for real-time RT-PCR. Cell proliferation was determined by MTT assay (BT-20; n=3, MDA-MB-453 and ZR-75-1; n=2).

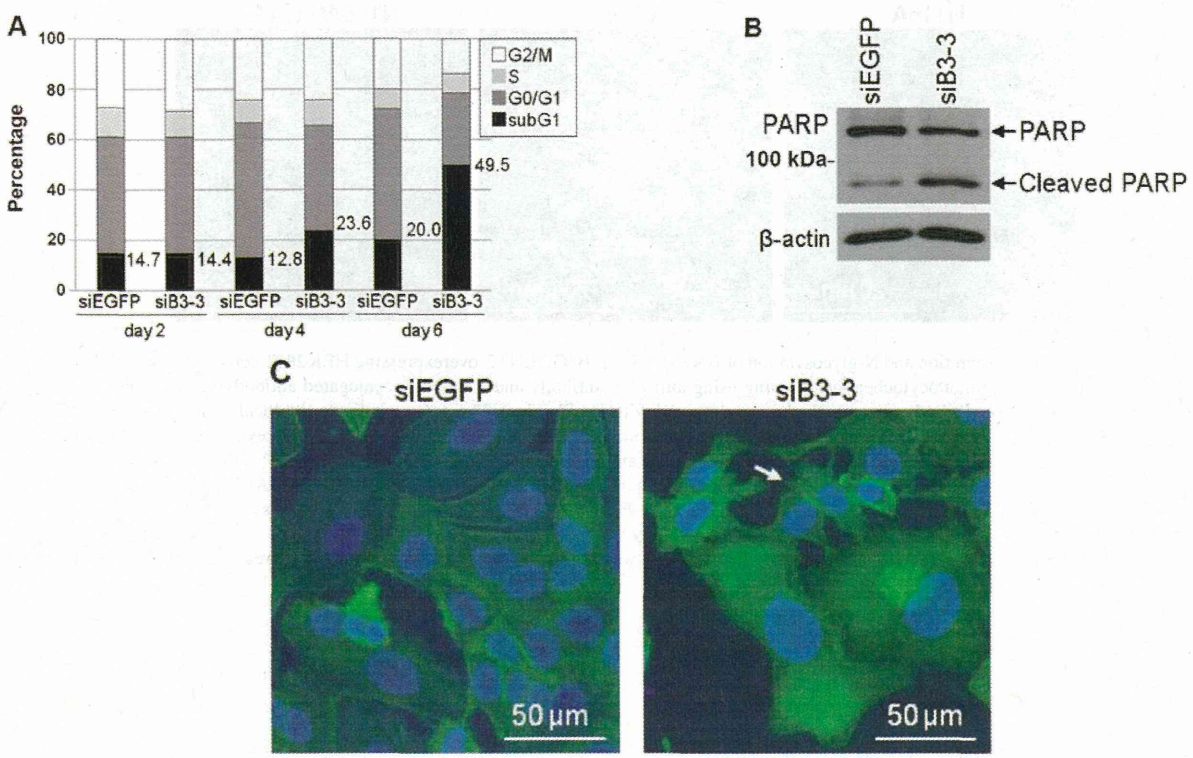

Figure 3. Knockdown of *B3GALNT2*-induced apoptosis in BT-20 TNBC cells. (A) FACS analysis of B3GALNT2-depleted BT-20 cells at each time-point. Numbers are the percentage of the sub-G1 population. (B) Cleaved PARP in BT-20 cells at 4 days after siRNA treatment was detected by western blot analysis. (C) Immunocytochemical staining in BT-20 cells at 4 days after siRNA treatment using Alexa 488-conjugated phalloidin (green) and DAPI (blue). The arrows indicate the altered cell morphology.

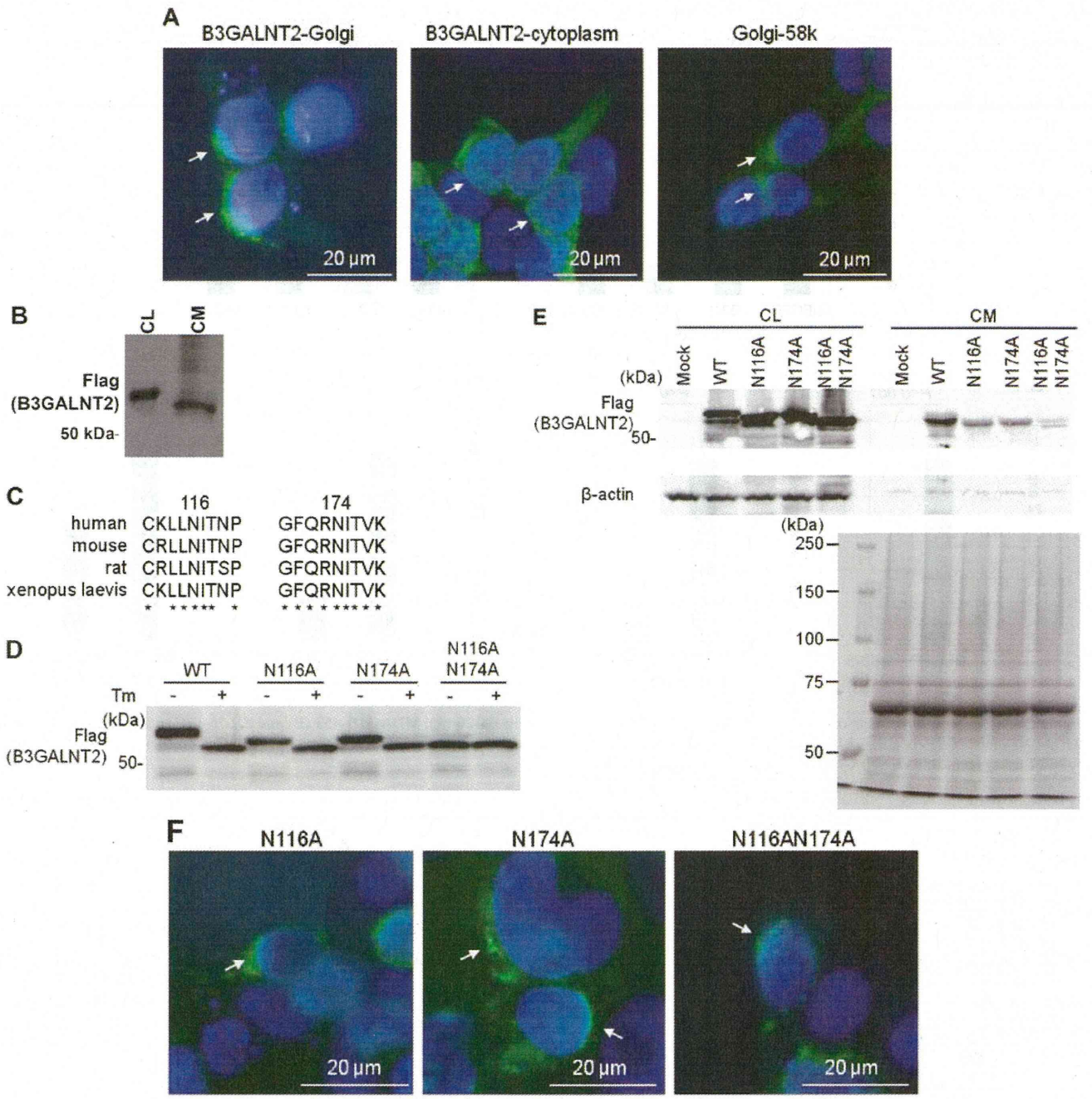

Figure 4. Subcellular localization, secretion and N-glycosylation of B3GALNT2 in B3GALNT2-overexpressing HEK293T cells. (A) Intracellular localization of B3GALNT2 was analyzed by immunocytochemical staining using anti-Flag antibody and Alexa 488-conjugated antibody (green). The Golgi apparatus were visualized by anti-Golgi-58k antibody (green). Nuclei were stained with DAPI (blue). The arrows indicate the localization of B3GALNT2-Flag or Golgi-58k. (B) B3GALNT2 of cell lysate (CL) and culture media (CM) from HEK293T cells, which were transiently expressing B3GALNT2-Flag, was detected by western blot analysis. (C) Homology search of two glycosylation sites of human B3GALNT2 with mouse (NP_848755), rat (NP_001138323) and *Xenopus laevis* (NP_001084830). Asterisks indicate the conserved amino acid residues among these species. (D) B3GALNT2-overexpressing HEK293T cells were cultured with (+) or without (-) tunicamycin (Tm). B3GALNT2-Flag was detected by western blot analysis using Flag monoclonal antibody. (E) B3GALNT2 of cell lysates (CL) and culture media (CM) was detected by western blot analysis. A comparison of the total protein amounts from culture media was performed by CBB staining. (F) Immunocytochemical staining was performed using anti-Flag antibody (green) and DAPI (blue). The arrows indicate the localization of B3GALNT2-Flag.

cantly smaller. In addition, tunicamycin treatment induced a decreased molecular mass of N116A and N174A, although the N116AN174A double mutant mass was not decreased (Fig. 4D).

To evaluate the effect of N-glycosylation on the secretion of the B3GALNT2 protein, we examined its secretion in HEK293T cells transfected with B3GALNT2-WT, -N116A, -N174A and N116AN174A, respectively. The secretion of B3GALNT2-N116A or -N174A was reduced compared with WT, while the secretion of N116AN174A mutant in culture media was drastically reduced (Fig. 4E). However, all of the mutants were mainly localized to the Golgi apparatus as well as B3GALNT2-WT

(Fig. 4F), thus suggesting no effects of these mutants on its Golgi retention. Taken together, both Asn-116 and Asn-174 amino acids of B3GALNT2 are N-glycosylation sites and their N-glycosylation is necessary for B3GALNT2 secretion.

Discussion

Glycosylation is a posttranslational modification and is associated with various physiologic events. The aberrant expression of glycosyltransferase and the immature glycan structure of proteins and lipids are observed in many cancers. These phenomena are also involved in the development and progression of cancers (13-16,24). Abnormalities of the glycan structure of glycoproteins are frequently observed in breast cancer cells (13-15). In particular, we previously identified and characterized the oncogenic roles of a cancer-specific glycosyltransferase, UDP-*N*-acetyl-α-D-galactosamine (GalNAc): polypeptide *N*-acetylgalactosaminyltransferase 6 (GALNT6) that regulated cell proliferation and cytoskeleton structure through aberrant *O*-glycosylation and stabilization of an oncoprotein mucin 1 (MUC1) (14) and fibronectin (15), which indicated that the development of GALNT6 inhibitors would be valuable for breast cancer therapy. To further elucidate the oncogenic role of aberrant glycosyltransferase expression, we attempted to identify cancer-specific glycosyltransferases that are exclusively upregulated in breast cancers through the analysis of comprehensive gene expression profiles of TNBC and normal human tissues. In this study, we focused on a breast cancer-specific glycosyltransferase, *B3GALNT2* and showed its potential as a druggable target by showing its critical roles in breast cancer cell growth.

B3GALNT2 was indicated to be the member of the β1,3-glycosyltransferase (β3GT) family by having three β3GT motifs and its function was shown by *in vitro* analyses to be a synthesis of GalNAcβ1-3GlcNAcβ1-R structure on both N-glycans and O-glycans of proteins (11). However, the biological and biochemical functions of B3GALNT2 have not been clarified in mammalian cells, including human cancer cells, primarily because the GalNAc β1-3GlcNAcβ1-R structure has been reported only in α-dystroglycan in mammalian cells (25). Recently, mutations in the *B3GALNT2* gene were identified in individuals with dystroglycanopathy by whole-exome and Sanger sequencing technologies, suggesting that α-dystroglycan is the potential substrate of B3GALNT2 (26). In contrast, the expression of α-dystroglycan has been reported to be frequently downregulated in breast cancers (27). Moreover, Stevens *et al* (26) showed that exogenous V5-tagged B3GALNT2 was mainly localized in the endoplasmic reticulum (ER) of C2C12 myoblasts, which is not concordant with our results indicating that exogenous B3GALNT2-Flag was mainly localized in the Golgi apparatus (Fig. 4A). This discrepancy may be due to differences in the experimental procedures employed, including the type of cell lines or the different expression vector constructs. Indeed, both previous findings and our results showed only exogenous expression of B3GALNT2 in mammalian cells. Hence, further study is necessary to clarify the biological roles or exact subcellular-localization of glycosylated α-dystroglycan in endogenous B3GALNT2-overexpressing breast cancer cells.

We first identified *B3GALNT2* to be upregulated in TNBCs, but demonstrated that the silencing of *B3GALNT2* expression by siRNA resulted in significant suppression of the growth of non-TNBC and TNBC cell lines (Fig. 2). Nevertheless, the overexpression of *B3GALNT2* into HEK293 or NIH3T3 cells could not enhance cell proliferation (data not shown), indicating that *B3GALNT2* is indispensable for the survival of breast cancer cells, but *B3GALNT2* alone may not be sufficient for transformation activity. Furthermore, B3GALNT2 was shown to be a secreted protein (Fig. 4), but addition of the conditioned media of the B3GALNT2-transfected HEK293T cells into the culture media of HEK293A cells could not enhance cell growth (data not shown). These results suggest that the intrinsic glycosyltransferase activity of B3GALNT2 might be critical for breast cancer cell growth. However, it has been reported that GnT-V secreted from WiDr colon cancer cells is directly involved in tumor angiogenesis in a glycosylation-independent manner, thus providing biological importance for the secretion of this glycosyltransferase (28). Therefore, further studies are needed to clarify the precise biological roles of the secreted form of B3GALNT2.

Furthermore, we demonstrated that N-glycosylation at Asn-116 or Asn-174 of B3GALNT2 is critical for its efficient secretion, but the effects of these posttranslational modifications on its biological functions, including enzyme activity, are unknown. Some studies have shown that N-glycosylation on glycosyltransferases is required for their proper-folding and/or enzymatic activities (29-31). Therefore, further study is necessary to clarify the precise pathophysiological roles of B3GALNT2 in mammary carcinogenesis through identification and characterization of its specific substrates and to screen inhibitors targeting glycosyltransferase activity of B3GALNT2 for potential breast cancer therapeutic applications.

In conclusion, we demonstrated that overexpression of cancer-specific glycosyltransferase B3GALNT2 is critical for the growth or survival in breast cancers, including TNBC and ER-positive breast cancer. Our findings showed the usefulness of B3GALNT2 as a promising diagnostic and/or therapeutic target for breast cancer.

Acknowledgements

We thank Drs Kaoru Takegawa (Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University) and Tomoya Fukawa (The University of Tokushima) for helpful discussions and Ms. Hinako Koseki for technical assistance. This study was supported by a Grant-in-Aid for Young Scientists (B) (grant nos. 23790369 and 25860240) from MEXT.

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T and Thun MJ: Cancer Statistics, 2008. CA Cancer J Clin 58: 71-96, 2008.
2. Di Cosmo S and Baselga J: Management breast cancer with targeted agents: Importance of heterogeneity. Nat Rev Clin Oncol 7: 139-147, 2010.
3. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD and Feuer EJ: Cancer Intervention and Surveillance Modeling Network (CISNET) Collaborators: Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 353: 1784-1792, 2005.
4. Early Breast Cancer Trialists' Collaborative Group (EBCTCG): Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365: 1687-1717, 2005.

5. Adamo V, Iorfida M, Montalto E, Festa V, Garipoli C, Scimone A, Zanghì M and Caristi N: Overview and new strategies in metastatic breast cancer (MBC) for treatment of tamoxifen-resistant patients. Ann Oncol 18: 53-57, 2007.
6. Foulkes WD, Smith IE and Reis-Filho JS: Triple-negative breast cancer. N Engl J Med 363: 1938-1948, 2010.
7. Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN and Pusztai L: Response to neoadjuvant therapy and longterm survival in patients with triple-negative breast cancer. J Clin Oncol 26: 1275-1281, 2008.
8. Petricoin EF III, Hackett JL, Lesko LJ, Puri RK, Gutman SI, Chumakov K, Woodcock J, Feigal DW Jr, Zoon KC and Sistare FD: Medical applications of microarray technologies: a regulatory science perspective. Nat Genet 32: 474-479, 2002.
9. Nishidate T, Katagiri T, Lin ML, Mano Y, Miki Y, Kasumi F, Yoshimoto M, Tsunoda T, Hirata K and Nakamura Y: Genome-wide gene-expression profiles of breast-cancer cells purified with laser microbeam microdissection: identification of genes associated with progression and metastasis. Int J Oncol 25: 797-819, 2004.
10. Komatsu M, Yoshimaru T, Matsuo T, Kiyotani K, Miyoshi Y, Tanahashi T, Rokutan K, Yamaguchi R, Saito A, Imoto S, Miyano S, Nakamura Y, Sasa M, Shimada M and Katagiri T: Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. Int J Oncol 42: 478-506, 2013.
11. Hiruma T, Togayachi A, Okamura K, Sato T, Kikuchi N, Kwon YD, Nakamura A, Fujimura K, Gotoh M, Tachibana K, Ishizuka Y, Noce T, Nakanishi H and Narimatsu H: A novel human beta1,3-N-acetylgalactosaminyltransferase that synthesizes a unique carbohydrate structure, GalNAcbeta1-3GlcNAc. J Biol Chem 279: 14087-14095, 2004.
12. Hart GW and Copeland RJ: Glycomics hits the big time. Cell 143: 672-676, 2010.
13. Fuster MM and Esko JD: The sweet and sour of cancer: glycans as novel therapeutic targets. Nat Rev Cancer 5: 526-542, 2005.
14. Park JH, Nishidate T, Kijima K, Ohashi T, Takegawa K, Fujikane T, Hirata K, Nakamura Y and Katagiri T: Critical roles of mucin 1 glycosylation by transactivated polypeptide N-acetylgalactosaminyltransferase 6 in mammary carcinogenesis. Cancer Res 70: 2759-2769, 2010.
15. Park JH, Katagiri T, Chung S, Kijima K and Nakamura Y: Polypeptide N-acetylgalactosaminyltransferase 6 disrupts mammary acinar morphogenesis through O-glycosylation of fibronectin. Neoplasia 13: 320-326, 2011.
16. Potapenko IO, Haakensen VD, Lüders T, Helland A, Bukholm I, Sørlie T, Kristensen VN, Lingjaerde OC and Børresen-Dale AL: Glycan gene expression signatures in normal and malignant breast tissue; possible role in diagnosis and progression. Mol Oncol 4: 98-118, 2010.
17. Park JH, Lin ML, Nishidate T, Nakamura Y and Katagiri T: PDZ-binding kinase/T-LAK cell-originated protein kinase, a putative cancer/testis antigen with an oncogenic activity in breast cancer. Cancer Res 66: 9186-9195, 2006.
18. Kim JW, Akiyama M, Park JH, Lin ML, Shimo A, Ueki T, Daigo Y, Tsunoda T, Nishidate T, Nakamura Y and Katagiri T: Activation of an estrogen/estrogen receptor signaling by BIG3 through its inhibitory effect on nuclear transport of PHB2/REA in breast cancer. Cancer Sci 100: 1468-1478, 2009.
19. Shimo A, Nishidate T, Ohta T, Fukuda M, Nakamura Y and Katagiri T: Elevated expression of protein regulator of cytokinesis 1, involved in the growth of breast cancer cells. Cancer Sci 98: 174-181, 2007.
20. Nogueira E, Fidalgo M, Molnar A, Kyriakis J, Force T, Zalvide J and Pombo CM: SOK1 translocates from the Golgi to the nucleus upon chemical anoxia and induces apoptotic cell death. J Biol Chem 283: 16248-16258, 2008.
21. Paulson JC and Colley KJ: Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. J Biol Chem 264: 17615-17618, 1989.
22. El-Battari A, Prorok M, Angata K, Mathieu S, Zerfaoui M, Ong E, Suzuki M, Lombardo D and Fukuda M: Different glycosyltransferases are differentially processed for secretion, dimerization, and autoglycosylation. Glycobiology 13: 941-953, 2003.
23. Kato T, Suzuki M, Murata T and Park EY: The effects of N-glycosylation sites and the N-terminal region on the biological function of beta1,3-N-acetylglucosaminyltransferase 2 and its secretion. Biochem Biophys Res Commun 329: 699-705, 2005.
24. Brockhausen I: Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta 1473: 67-95, 1999.
25. Yoshida-Moriguchi T, Yu L, Stalnaker SH, Davis S, Kunz S, Madson M, Oldstone MB, Schachter H, Wells L and Campbell KP: O-mannosyl phosphorylation of alpha-dystroglycan is required for laminin binding. Science 327: 88-92, 2010.
26. Stevens E, Carss KJ, Cirak S, Foley AR, Torelli S, Willer T, Tambunan DE, Yau S, Brodd L, Sewry CA, Feng L, Haliloglu G, Orhan D, Dobyns WB, Enns GM, Manning M, Krause A, Salih MA, Walsh CA, Hurles M, Campbell KP, Manzini MC; UK10K Consortium, Stemple D, Lin YY and Muntoni F: Mutations in B3GALNT2 cause congenital muscular dystrophy and hypoglycosylation of α-dystroglycan. Am J Hum Genet 92: 354-365, 2013.
27. Sgambato A, Migaldi M, Montanari M, Camerini A, Brancaccio A, Rossi G, Cangiano R, Losasso C, Capelli G, Trentini GP and Cittadini A: Dystroglycan expression is frequently reduced in human breast and colon cancers and is associated with tumor progression. Am J Pathol 162: 849-860, 2003.
28. Saito T, Miyoshi E, Sasai K, Nakano N, Eguchi H, Honke K and Taniguchi N: A secreted type of β1,6-N-acetylglucosaminyltransferase-V (GnT-V) induces tumor angiogenesis without mediation of glycosylation. A novel function of GnT-V distinct from the original glycosyltransferase activity. J Biol Chem 277: 17002-17008, 2002.
29. Toki D, Sarkar M, Yip B, Reck F, Joziasse D, Fukuda M, Schachter H and Brockhausen I: Expression of stable human O-glycan core 2 beta-1,6-N-acetylglucosaminyltransferase in Sf9 insect cells. Biochem J 325: 63-69, 1997.
30. Barbier O, Girard C, Breton R, Bélanger A and Hum DW: N-glycosylation and residue 96 are involved in the functional properties of UDP-glucuronosyltransferase enzymes. Biochemistry 39: 11540-11552, 2000.
31. Christensen LL, Bross P and Ørntoft TF: Glycosylation of the N-terminal potential N-glycosylation sites in the human alpha1,3-fucosyltransferase V and -VI (hFucTV and -VI). Glycoconj J 17: 859-865, 2000.

OPEN ACCESS Freely available online

PLOS | ONE

Genome-Wide Association Study of Breast Cancer in the Japanese Population

Siew-Kee Low[1,2]*, Atsushi Takahashi[1], Kyota Ashikawa[3], Johji Inazawa[4], Yoshio Miki[5], Michiaki Kubo[3], Yusuke Nakamura[2,6], Toyomasa Katagiri[7]

1 Laboratory for Statistical Analysis, Center for Integrative Medical Sciences, The Institute of Physical and Chemical Research (RIKEN), Yokohama, Japan, **2** Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **3** Laboratory for Genotyping Development, Center for Integrative Medical Sciences, The Institute of Physical and Chemical Research (RIKEN), Yokohama, Japan, **4** Department of Molecular Cytogenetics, Medical Research Institute and School of Biomedical Science, Tokyo Medical and Dental University, Tokyo, Japan, **5** Department of Genetic iagnosis, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan, **6** Department of Medicine and Surgery, The University of Chicago, Chicago, Illinois, United States of America, **7** Division of Genome Medicine, Institute for Genome Research, The University of Tokushima, Japan

Abstract

Breast cancer is the most common malignancy among women in worldwide including Japan. Several studies have identified common genetic variants to be associated with the risk of breast cancer. Due to the complex linkage disequilibrium structure and various environmental exposures in different populations, it is essential to identify variants associated with breast cancer in each population, which subsequently facilitate the better understanding of mammary carcinogenesis. In this study, we conducted a genome-wide association study (GWAS) as well as whole-genome imputation with 2,642 cases and 2,099 unaffected female controls. We further examined 13 suggestive loci ($P<1.0\times10^{-5}$) using an independent sample set of 2,885 cases and 3,395 controls and successfully validated two previously-reported loci, rs2981578 (combined *P*-value of 1.31×10^{-12}, OR = 1.23; 95% CI = 1.16–.30) on chromosome 10q26 (*FGFR2*), rs3803662 (combined *P*-value of 2.79×10^{-11}, OR = 1.21; 95% CI = 1.15–.28) and rs12922061 (combined *P*-value of 3.97×10^{-10}, OR = 1.23; 95% CI = 1.15–.31) on chromosome 16q12 (*TOX3-LOC643714*). Weighted genetic risk score on the basis of three significantly associated variants and two previously reported breast cancer associated loci in East Asian population revealed that individuals who carry the most risk alleles in category 5 have 2.2 times higher risk of developing breast cancer in the Japanese population than those who carry the least risk alleles in reference category 1. Although we could not identify additional loci associated with breast cancer, our study utilized one of the largest sample sizes reported to date, and provided genetic status that represent the Japanese population. Further local and international collaborative study is essential to identify additional genetic variants that could lead to a better, accurate prediction for breast cancer.

Citation: Low S-K, Takahashi A, Ashikawa K, Inazawa J, Miki Y, et al. (2013) Genome-Wide Association Study of Breast Cancer in the Japanese Population. PLoS ONE 8(10): e76463. doi:10.1371/journal.pone.0076463

Editor: Xiaoping Miao, MOE Key Laboratory of Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, China

Received July 16, 2013; **Accepted** August 29, 2013; **Published** October 15, 2013

Funding: This work was conducted as a part of Biobank Japan Project that was supported by the Ministry of Education, Culture, Sports, Sciences and Technology from the Japanese Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: amanda.sklow@src.riken.jp

Introduction

Breast cancer is the most common malignancy among women worldwide. In Japan, breast cancer comprises approximately 19% of all female cancers; it is the fifth leading cause of cancer death among women with an estimated death of 12,731 in 2011. Its incidence is about 86.0 cases/100,000 individuals/year and 56,289 were newly diagnosed to have breast cancer in 2007 (http://ganjoho.jp/data/public/statistics/backnumber/2012/files/cancer_statistics_2012.pdf). Even though the 5-year survival rate for breast cancer is relatively better compared to other malignancies, the age-adjusted incidence and mortality rate for breast cancer has revealed a significant increase since 1970s in Japan. Hence, breast cancer is one of the most important medical issues to be addressed. In particular, individual risk assessment (genetic and environmental factors), early detection by biomarkers and mammography screening are critically important to reduce breast cancer-associated death.

Although risk factors such as age, age at menarche, ethnicity, reproductive and menstrual history, oral contraceptives, hormone therapy, radiation exposure, mammographic breast density, alcohol intake, dietary folate intake, physical activity and benign breast diseases have been reported [1–7], it is well known that breast cancer is a complex polygenic disease in which genetic factors play an important role in disease etiology and pathogenesis. Individuals who have first-degree relatives with breast cancer are indicated to have approximately 2.1-fold higher risk for the disease [8]. Previous linkage analysis identified mutations in two highly-penetrant genes, *BRCA1* and *BRCA2*, as one of the major cause of inherited cancer in many families [9,10]. In addition, mutations in

ATM, *TP53*, *CHEK2*, *PTEN*, *CDH1*, *STK11* and *PALB2* genes also confer risk to breast cancer [11–17]. Nevertheless, mutations in these genes are not common in the general population and account for 5–0% of breast cancer cases. Hence, there is likely to be other genetic variants that contribute to the etiology of this cancer. Since 2007, approximately 67 common genetic variants have been identified to be associated with breast cancer through genome-wide association studies (GWAS) and international collaborative study from European and Asian descendants [18–34]. Due to the complex linkage disequilibrium, differences in allele frequencies and environmental exposure in different populations, it is of importance to identify common genetic variants associated with breast cancer in specific populations, which subsequently facilitate the development of useful prediction systems. Our group has previously reported a GWAS to identify common genetic variants associated with hormonal receptor positive breast cancer [35], but the current study has increased the sample size, which is one of the biggest sample size that represent the Japanese population, and has used a genotyping panel with better coverage aiming to identify common genetic variants associated with all types of breast cancer. Furthermore, we also evaluated the association of previously-identified loci showing the association with breast cancer in the European population and East Asian in the current dataset. Lastly, we conducted whole-genome imputation by referring to 1000G reference panel to increase the coverage of this GWAS study.

Table 1. Demographic data of patients recruited for this study.

	GWAS Set	Validation Set
Case	2642	2885
Age	56.9	59.8
Menopause status		
Postmenopausal	2088	2201
Premenopausal	99	184
Unknown	455	500
Family history		
Yes*	305	365
No	2337	2520
Estrogen Receptor		
Positive	1146	1459
Negative	663	378
Progesterone Receptor		
Positive	949	1207
Negative	824	574
HER2 Receptor		
0	372	465
+1	324	307
+2	185	151
+3	176	105
No Staining Information	354	441
	GWAS Set	**Validation Set**
Control	2099	3395
Age	56.0	44.4
Menopause status		
Postmenopausal	1511	1377
Premenopausal	13	231
Unknown	807	1787

*Family members who have breast and/or ovarian cance.
doi:10.1371/journal.pone.0076463.t001

Subject and Methods

Study population

We recruited all DNA samples from the Biobank Japan Project (http://biobankjp.org). The Biobank Japan is a bank that has collected DNAs and serum of nearly 200,000 individuals, who had been diagnosed to have one or more of 47 common diseases including various types of cancer from 66 collaborating hospitals in Japan. One of the major objectives of this project is to identify genetic variants that are associated with common diseases and to identify individuals who are at risk for various diseases. In this study, we selected a total of 5,610 breast cancer patients who had been registered in the Biobank Japan. We subsequently divided these patients into two groups of 2,725 and 2,885 cases to be used for the discovery phase and for the validation phase, respectively. For controls, we included 2,331 and 3,395 females consisting of healthy volunteers from Midosuji Rotary Club, Osaka, Japan and Health Science Research Resource Bank as well as individuals in the Biobank who had no history of cancer as controls for discovery and validation phases, respectively. The demographic data of patients recruited for this study is summarized in Table 1. All individuals who participated in this study provided written inform consent. This study was approved by the ethical committees of the Institute of Medical Sciences, the University of Tokyo and RIKEN Center for Integrative Medical Sciences.

Genotyping and quality control

For the GWAS discovery stage, we genotyped both case and control samples using Illumina OmniExpress BeadChip that contained a total of 733,202 SNPs. After a standard SNP quality control which excluded SNPs with call rate of <0.98, those that deviated from the Hardy-Weinberg equilibrium ($P \leq 1.0 \times 10^{-6}$), those on the X chromosome and non-polymorphic SNPs, a total of 550,026 SNPs were used for further analysis. The cluster plot of 100 SNPs that revealed the strongest associations were checked by visual observation to exclude SNPs with ambiguous genotypes. For sample quality control, we evaluated cryptic relatedness for each sample with identity-by-state method. To examine population stratification of this study, we performed principal component analysis (PCA) using EIGENSTRAT software v2.0 (http://genepath.med.harvard.edu/reich/Software.htm) with four reference populations from the HapMap data as reference including Europeans (represented by Caucasian from UTAH, CEU), Africans (represented by Yoruba from Ibadan, YRI) and East Asians (represented by Japanese from Tokyo, JPT, and Han Chinese from Beijing, CHB) (Figure S1a). We plotted the scatter plot by using the top two associated principal components (eigenvectors) to identify outliers who did not belong to the JPT/CHB cluster. Subsequently, we performed PCA analysis using only the genotype information of the case and control subjects to further evaluate the population substructure (Figure S1b). Quantile-quantile (Q-Q) plot was constructed using observed *P*-values against expected *P*-values and an inflation factor value (λ-value) that was calculated to assess potential population stratification of the study subjects (Figure S2a and S2b). After performing PCA, we selected 2,642 cases and 2,099 controls within the major Japanese (Hondo) cluster for subsequent analysis (Figure S1b).

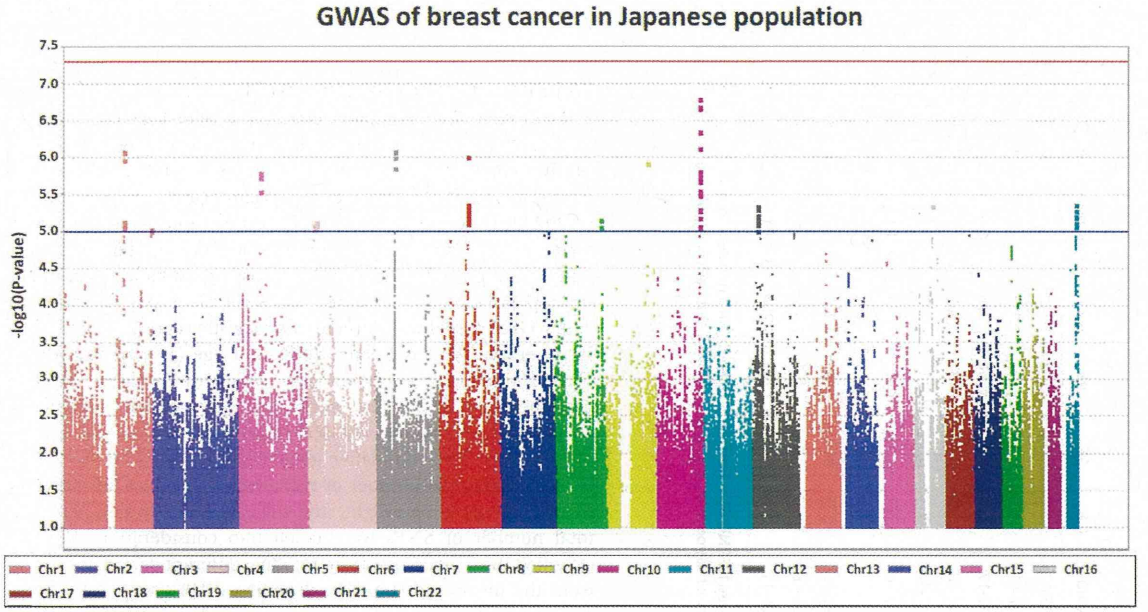

Figure 1. The Manhattan plot for GWAS of breast cancer in the Japanese population. This plot is based on $-\log_{10}$ (*P*-value) from GWAS and imputation analysis against chromosome position, each color represents different chromosome. Blue line indicate suggestive association threshold, $P=1\times10^{-5}$ while red line indicate genome-wide significant threshold $P<5\times10^{-8}$.
doi:10.1371/journal.pone.0076463.g001

Imputation analysis

To increase the power and coverage of the genome-wide association scan, we performed whole genome imputation using 1000G of East Asian population (Japanese in Tokyo JPT, Chinese in Beijing CHB and Chinese in Denver CHD) Phase I Integrated Release Version 2 dataset as reference panel to infer missing genotypes. Briefly, we prepared the input files after quality control, which excluded SNPs with genotyping rate of <98%, those that deviated from HWE (HWE $P\leq1.0\times10^{-6}$) and those with MAF of <0.01. We then confirmed that the allele frequencies of the reference allele are comparable between the GWAS dataset and the reference panel with differences of <0.15. By using MACH1.0 (http://www.sph.umich.edu/csg/abecasis/MACH/index.html), we performed haplotype phasing with the samples' genotypes referring 1000G reference panel, estimated the map crossover and error rates using 20 iterations of the Markov chain. Subsequently, we imputed the missing genotypes using Minimac (http://genome.sph.umich.edu/wiki/Minimac). We utilized stringent imputation quality control by excluding SNPs with r^2 value of <0.9.

Validation study

After evaluating the associations from GWAS and whole genome imputation, we selected a total of 13 candidate loci that showed suggestive association ($P<1.0\times10^{-5}$) with breast cancer risk for further validation by an independent set of 2,885 cases and 3,395 controls. We genotyped the cases with the multiplex-PCR Invader assay [36] and the control samples with either Illumina OmniExpress BeadChip Kits or by imputation. To verify the accuracy of the imputation analysis, we also included surrogate SNPs that showed close link ($r^2>0.8$ and D' = 1.00) to the imputed SNPs and were included in the genotype platform. Considering multiple testing at this validation stage, we applied Bonferroni significance threshold at $P<3.85\times10^{-3}$ (0.05/13 independent tests).

Evaluation of previously reported loci

To verify previously-reported loci showing the association with breast cancer in the European and East Asian populations, we evaluated 67 loci in the current Japanese GWAS dataset [18–34]. Among the 67 SNPs examined, 6 SNPs are not polymorphic in the Japanese population, 26 SNPs are same as the previously-reported SNPs, 33 and 2 SNPs are SNPs having r^2-value of more than 0.8 and 0.7 to the previously-reported SNPs, respectively (Table S4).

Statistical Analysis

The case-control associations of the GWAS discovery set and validation set were evaluated using logistic regression analysis after considering age as confounding factor from PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/). The associations of the imputed SNPs were generated with mach2dat software which utilized the output results from Minimac (dosage of the imputed SNP). To have an overview of the association of SNPs with breast cancer, a Manhattan plot of the study was plotted using Haploview 4.1. Meta-analysis for the combined analysis of the discovery and validation phase was performed using inverse-variance method and heterogeneity between the two phases was evaluated using Cochran's Q test. Regional association plots were

Table 2. Association study of SNPs on chromosome 10q26.13 and 16q12.1.

CHR	SNP	BP	Stage	RA	NRA	NCASES	NCONTROLS	RAF_Case	RAF_Ctrl	P_value	OR	SE	L95	U95	P_hetero	Gene	rel.loci
10	rs2981578	123340311	GWAS	C	T	2642	2097	0.571	0.517	2.25E-07	1.238	0.041	1.142	1.342		*FGFR2*	0
10	rs2981578	123340311	Rep	C	T	2883	3395	0.556	0.512	1.63E-06	1.213	0.040	1.121	1.313			
10	rs2981578	123340311	Combined	C	T	5525	5492	0.563	0.514	1.31E-12	1.225	0.028	1.158	1.296	7.18E-01		
16	rs12922061	52635000	GWAS	T	C	2641	2099	0.287	0.245	4.50E-06	1.244	0.048	1.133	1.365		*LOC643714*	0
16	rs12922061	52635000	Rep	T	C	2880	3395	0.278	0.239	1.41E-05	1.219	0.046	1.115	1.333			
16	rs12922061	52635000	Combined	T	C	5521	5494	0.282	0.241	3.97E-10	1.231	0.032	1.153	1.314	7.60E-01		
Another SNP on 16p12 that independently associated with breast cancer																	
16	rs3803662	52586341	GWAS	T	C	2642	2097	0.570	0.531	9.09E-05	1.178	0.042	1.085	1.279		*LOC643714*	0
16	rs3803662	52586341	Rep	T	C	2880	3392	0.572	0.517	4.69E-08	1.245	0.040	1.151	1.347			
16	rs3803662	52586341	Combined	T	C	5522	5489	0.571	0.522	2.79E-11	1.213	0.029	1.146	1.284	3.40E-01		

CHR: chromosome, SNP: single nucleotide polymorphism, BP: SNP genomic location, RA: Risk allele, NRA: Non-risk allele, NCASES: Number of cases, NCONTROLS: Number of controls, RAF: risk allele frequency, P_value: *P*-value from logistic regression analysis after age adjustment, OR: odds ratio, L95: lower 95% confidence interval, U95: upper 95% confidence interval, P_hetero: heterogeneity test with Cochran Q-test, rel.loci: distance of the SNP from the gene, GWAS: genome-wide association study, Rep: validation study.

doi:10.1371/journal.pone.0076463.t002

generated using Locus Zoom (http://csg.sph.umich.edu/locuszoom/).

Weighted genetic risk score (wGRS)

To evaluate the cumulative effects of genetic variants associated with breast cancer risk, we conducted weighted genetic risk score (wGRS) analysis on the basis of genotypes of five SNPs, three significant SNPs (rs2981578 of 10q26/*FGFR2*, rs3803662 and rs12922061 of 16q12/*TOX3*) from this study and two SNPs (rs6557161 of 6q25/*ESR1* and rs10509168 of 10q21/*ZNF365*) that were reported to be associated with breast cancer risk in East Asian population and indicated suggestive association in this study. The wGRS model was developed by logistic regression analysis by incorporating five associated-SNPs and age (as covariates) using GWAS dataset to obtain the estimates (weight) of each corresponding SNP. This model was subsequently validated in an independent samples dataset drawn from the validation phase of this study. The cumulative genetic risk scores were determined by multiplying the number of risk alleles (0/1/2) of an individual by its corresponding weight, and subsequently the sum across the total number of SNPs were taken into consideration. We then classified the genetic risk score into five different categories created from the mean and standard deviation (SD); group 1, < mean-1SD; group 2, mean-1SD to mean; group 3, mean to mean+1SD; group 4, mean+1SD to mean+2SD, group 5, > mean+2SD. Odds ratio and 95% confidence interval were calculated using group 1 as a reference.

Results

In this study, we genotyped a total of 2,725 cases and 2,311 controls with Illumina OmniExpress BeadChip Kits that contained 733,202 SNPs to identify genetic variants associated with the susceptibility to breast cancer in the Japanese population. After quality check of the SNP genotyping data, a total of 550,026 autosomal SNPs were examined for the association by logistic regression analysis. Quantile-quantile (Q-Q) plot and the genomic inflation factor (λ) of the test statistic of this GWAS based on 550,026 SNPs with all samples was 1.183 suggesting the existence of some population substructure (Figure S2a). To exclude the possibility of population substructure for our sample population, we performed principal component analysis (PCA). Although all the subjects participating in this study were clustered in the Asian population, there was a small portion of samples that were separated from the major Japanese (Hondo) cluster when PCA analysis was performed using only the genotype information of the case and control in the study (Figure S1a and S1b). We then used samples from the major Japanese (Hondo) cluster consisting of 2,642 cases and 2,099 controls, and found that the λ-value improved to 1.027 (Figure S2b). Hence, subsequent analysis was carried out using only samples from the major Japanese cluster. Whole genome imputation utilizing 1000G database as reference panel successfully estimated 7,791,127 SNPs. After stringent quality control by excluding SNPs with r^2-value of <0.9, the total number of SNPs that were taken into account was 5,335,291. The Manhattan plot, plotting $-\log_{10}$ (*P*-value) from the GWAS and imputation analysis against the chromosome position, showed that there were no genetic loci achieving genome-wide significance with the threshold *P*-value of $<5\times10^{-8}$ (Figure 1).

To identify additional susceptible loci associated with breast cancer, we conducted a validation study of 13 genetic loci showing suggestive association ($P<1\times10^{-5}$) with breast cancer after excluding SNPs that showed linkage disequilibrium (LD) coefficient (r^2) of >0.8 within each LD block by examining an

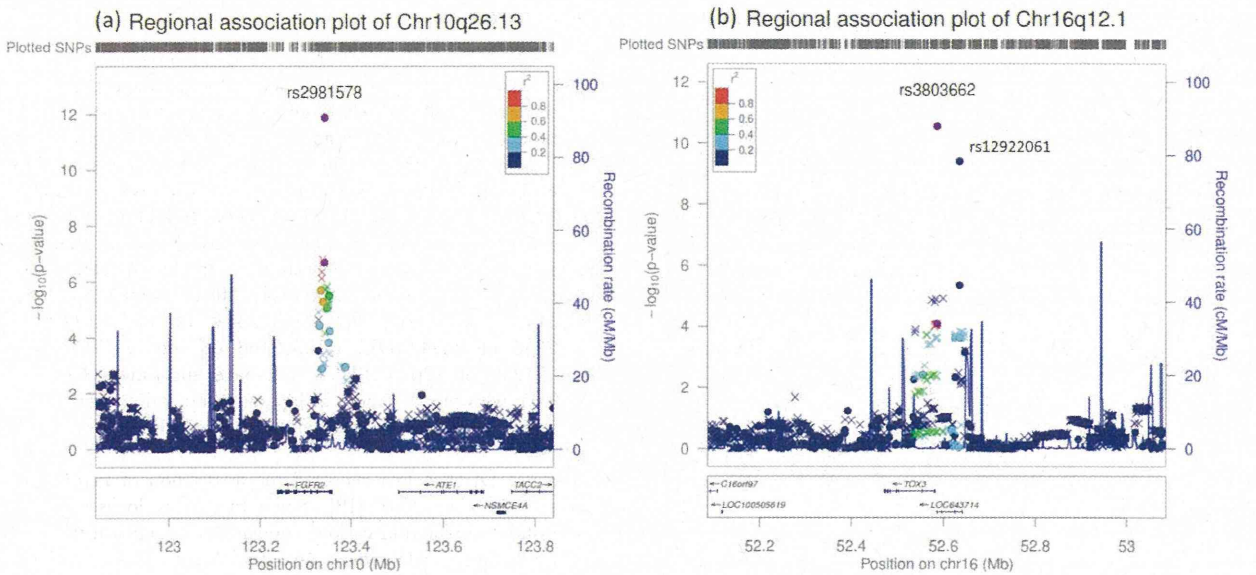

Figure 2. Regional association plots for two significantly associated loci with breast cancer in Japanese population, (a) chromosome 10q26.13 (*FGFR2*) and (b) chromosome 16q21.1 (*TOX3-LOC643714*). SNPs from the GWAS are plotted as circles; imputed SNPs are plotted as crosses. The color intensity reflects the extent of LD with the marker SNP: red, ($r^2 \geq 0.8$), orange ($0.6 \leq r^2 \leq 0.8$), green ($0.4 \leq r^2 \leq 0.6$), light blue ($0.2 \leq r^2 \leq 0.4$) and dark blue ($r^2 < 0.2$). Purplish blue lines represent local recombination rates. The SNP position is based on NCBI build 37.
doi:10.1371/journal.pone.0076463.g002

independent set of 2,885 breast cancer cases and 3,395 controls. Among the 13 loci tested, three SNPs (rs2981578 on chromosome 10q26.13 and rs3803662 along with rs12922061 on chromosome 16q12.1) were successfully validated with Bonferroni-corrected *P*-value of $<3.85\times10^{-3}$ (0.05/13 independent tests) as shown in Table 2 and Table S1. Inverse variance meta-analysis indicated that these three SNPs surpass genome-wide significance level (*P*-value$<5\times10^{-8}$) after combining the GWAS and the validation study with no significant heterogeneity (*P*-value >0.05) between the two stages (Table 2).

The most significantly associated SNP, rs2981578 (combined *P*-value of 1.31×10^{-12}, OR = 1.23; 95% CI = 1.16–.30), is located within the second intron of the *FGFR2* gene on chromosome 10q26.13 (Table 2 and Figure 2a). Variants on this gene have been the most frequently validated to be associated with breast cancer in multiple populations. For chromosome 16q21.1, we successfully validated rs12922061 (combined *P*-value of 3.97×10^{-10}, OR = 1.23; 95% CI = 1.15–.31) to be significantly associated with breast cancer (Table 2 and Figure 2b). After conditioning the effect of rs12922061, rs3803662 remained suggestively associated and was successfully validated after additional samples with a combined *P*-value of 2.79×10^{-11} (OR = 1.21; 95% CI = 1.14–.25). Two of these SNPs remained significant (*P*-value<0.0001) after performing condition analysis by using one of the SNP as covariate, suggesting the independency of association with breast cancer (Table S2). Additionally, the r^2 value between these two SNPs is only 0.17, indicating they are not closely linked with each other. Haplotype analysis of the two SNPs did not reveal stronger association than a single SNP association after 100,000 permutation analysis (Table S3). The SNP, rs3803662, is located in the last exon of *LOC643714* and near to the 5′ end of *TOX3*; whilst rs12922061 is located in the first intron of *LOC643714*.

In addition to perform GWAS for breast cancer in Japanese population, we also evaluated the association of previously-reported breast cancer risk loci in the European and East Asian populations. We evaluated a total of 61 SNPs after excluding 6 SNPs that are not polymorphic in Japanese population (Table S4). Among the 61 SNPs, eight SNPs (rs4415084 of 5p12/*MRPS30*, rs6557161 of 6q25/*ESR1*, rs7465364 of 8p21/*RPL17p33*, rs672888 of 8q24/*MYC*, rs10509168 of 10q21/*ZNF365*, rs1219648 of 10q26/*FGFR2*, rs17221259 of 12p13/*ATF7IP* and rs3803662 of 16q12/*TOX3*) showed suggestive association (*P*-value<0.05) with breast cancer in Japanese population (Table 3). All of these suggestively-associated SNPs possessed the same risk allele and showed the same direction of association that was indicated in the previous reports.

After developing wGRS model using five SNPs from the GWAS dataset, the model was subsequently validated in an independent sample set represented by the validation samples. The cumulative effect of five SNPs evaluated by the wGRS analysis indicated that odds ratio of each category increased according to the level of risk score, and individuals who are in category five carrying the most risk alleles have 2.2 times higher risk to develop breast cancer when utilizing category 1 as a reference (Table 4).