Nuclei were visualized by 4,6-diamidino-2-phenylindole counterstaining.

Patients and tumor tissues

This study was carried out after obtaining approval from the internal review board. Frozen samples of invasive breast carcinoma tissue were obtained from a case series of 52 patients who had undergone surgery at the National Defense Medical College Hospital, Tokorozawa, Japan, between 2005 and 2007. These frozen tumor samples were used to evaluate levels of SDF-1 mRNA by quantitative RT-PCR. From formalin-fixed, paraffin-embedded samples of the corresponding 52 tumors, 4-μm-thick sections were prepared and subjected to immunohistochemistry for SDF-1 protein expression.

From among 247 consecutive patients who had undergone mastectomy or breast-conserving surgery for unilateral invasive breast carcinoma at the National Defense Medical College Hospital between 1995 and 1999, complete medical records and appropriate tissue samples for constructing tissue microarrays were available for 223 patients. The tumors resected from these 223 patients were included in another retrospective immunohistochemical study of SDF-1 expression.

The 223 patients had been followed up for a median period of 74 months (range, 1–151 months), during which there were 58 relapses and 30 deaths. After surgical therapy, the patients with ER- and/or PR-positive breast cancer had received endocrine therapy for 2 years or more, and the patients with large tumors and/or nodal metastasis had received adjuvant chemotherapy. Fifteen patients with locally advanced breast cancer had received preoperative chemotherapy, for example, two or more courses of the cyclophosphamide-epirubicin-5-fluorouracil (CEF) regimen. Two-hundred and twenty-one patients were females and two were males. Additional patient characteristics are summarized in Table 1. For these 223 tumors, histological types and nuclear grade were re-examined for the present study by two observers (T. K. and H. T.). Other clinicopathological data were collected from the medical records and pathology reports.

Tissue microarray construction

We constructed tissue microarray blocks as described previously [11]. Briefly, double tissue cores 2 mm in diameter were taken from each donor block, and these core specimens were transferred to a recipient block using a Tissue Microarrayer (Beecher Instruments, Silver Spring, MD, USA). One Tissue microarray (TMA) block contained a maximum of 26 tumor samples, and 13 TMA sets were prepared for the present study.

Immunohistochemistry

Immunohistochemistry was performed on whole tissue sections from the 52 tumors, which were also subjected to quantitative RT-PCR, and on the TMA sections from the 223 tumors. Antibodies used were mouse monoclonal anti-human SDF-1 (R&D Systems), mouse monoclonal anti-human CXCR4 (MAB173, R&D systems), mouse monoclonal anti-human ER (clone 1D5, Dako, Glostrup, Denmark), rabbit polyclonal anti-HER2 antibody included in a HercepTest kit (Dako), and murine IgG isotype control antibody (R&D Systems). After deparaffinization of 4-μm-thick tissue sections, antigens were retrieved by microwaving in 10 mM sodium citrate (pH 6.0) for SDF-1, CXCR4, and murine isotype antibody, or by autoclaving in 10 mM Tris–HCl (pH 9.0) for ER. Non-specific binding was blocked by incubation in 1% normal swine serum (Dako) in PBS. The slides were incubated with the primary antibodies at 4°C overnight and then reacted with a dextran polymer reagent combined with secondary antibodies and peroxidase (Envision Plus; Dako) for 1 h at room temperature. Specific antigen–antibody reactions were visualized with 0.2% diaminobenzidine tetrahydrochloride and hydrogen peroxide. Counterstaining was performed using Mayer's hematoxylin.

The expression of SDF-1 was assessed according to the proportion of the stained area (<10% as negative, ≥10% as positive). Positive staining was further classified into two patterns according to the localization of SDF-1 immunoreactivity: cytoplasmic-dominant staining (C-pattern), and membrane-dominant staining (M-pattern). For CXCR4, nuclear staining in ≥10% and <10% of carcinoma cells was defined as positive and negative expression, respectively. A HER2 score was assigned according to the standard procedure, and a score of 3+ was classified as positive. ER was defined as positive if nuclear staining was seen in ≥10% of carcinoma cells. The results of immunohistochemistry were evaluated by two observers (T.K. and H.T.) independently, and cases with discrepant judgments were re-evaluated by discussion until consensus was obtained.

Statistical analysis

Comparisons between groups were evaluated using chi-squared test, Fisher's exact test, or Mann–Whitney *U*-test. Multiple comparisons (post hoc test) were carried out by the Bonferroni method. Survival curves of patients were drawn using the Kaplan–Meier method. Cox's univariate and multivariate proportional hazards models were used to explore the association of variables with disease-free and overall survival. For all the tests, differences at $P < 0.05$ were considered to be statistically significant. All analyzes were performed using the JMP 6.0 software package for Windows (SAS Institute Inc., Cary, NC, USA).

Table 1 Clinicopathological implication of SDF-1 expression in surgically resected breast cancers

	Total (n = 223)	SDF-1		
		"High expression" (n = 158) [n (%)]	"Low expression" (n = 65) [n (%)]	P value
Age				
Median y (range)		52 (30 ~ 82y)		
≤52	111	87 (79)	24 (37)	
>52	112	71 (45)	41 (63)	0.018
ER status				
Negative	90	49 (31)	41 (63)	
Positive	133	109 (69)	24 (37)	<0.0001
HER2 status				
Score 0–2+	202	148 (94)	54 (83)	
Score 3+	21	10 (6)	11 (17)	0.021
Nuclear grade				
1	45	38 (24)	7 (11)	
2	95	78 (49)	17 (26)	
3	83	42 (27)	41 (63)	<0.0001
Tumor size				
<5.0 cm	176	128 (81)	48 (74)	
≥5.0 cm	42	25 (16)	17 (26)	0.13
Unknown	5	5 (3)	0 (0)	
Lymph node metastasis				
(−)	117	86 (55)	31 (48)	
(+)	101	67 (42)	34 (52)	0.29
Unknown	4	4 (3)	0 (0)	
Distant metastasis				
(−)	211	151 (96)	60 (92)	
(+)	8	3 (2)	5 (8)	0.052
Unknown	4	4 (2)	0 (0)	
Stage				
1 or 2	181	131 (83)	50 (77)	
3 or 4	37	22 (14)	15 (23)	0.16
Unknown	5	5 (3)	0 (0)	
Tumor histology				
IDC	194	136 (86)	58 (89)	
ILC	10	10 (6)	0 (0)	
Special[a]	19	12 (8)	7 (11)	0.095
CXCR4				
Negative	62	35 (22)	27 (42)	
Positive	161	123 (78)	38 (58)	0.0033

Abbreviation: *IDC* invasive ductal carcinoma, *ILC* invasive lobular carcinoma

[a] Mucinous carcinoma (6), Medullary carcinoma (3), Spindle cell carcinoma (1), Apocrine carcinoma (2), Tubular carcinoma (5), Micropapillary (1), Endocrine carcinoma (1)

Results

SDF-1 expression is induced by estradiol in ER-positive breast cancer cell lines

To determine whether breast cancer cells produce SDF-1, we subjected breast cancer cell lines to quantitative RT-PCR, ELISA, and immunofluorescence analysis. Quantitative RT-PCR detected a low level of SDF-1 mRNA in the ER-positive cell lines MCF-7 and T47D (clear columns in Fig. 1a), but the mRNA level increased about 7-fold and 21-fold, respectively, in response to estradiol treatment (filled columns in Fig. 1a). In contrast, the level of SDF-1 mRNA was very low in the ER-negative cell lines MDA-MB-231, MDA-MB435, and MDA-MB-436, and did not change after estradiol treatment (Fig. 1a).

The results of ELISA were similar to those of quantitative RT-PCR. Mean levels of secreted SDF-1 protein in

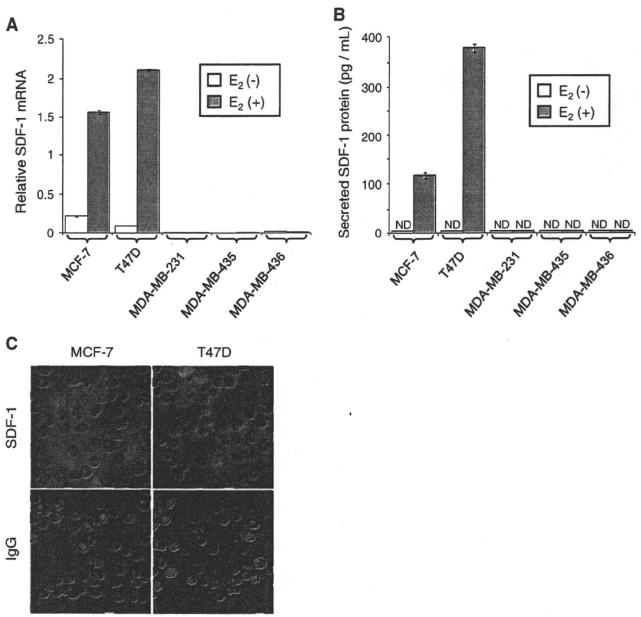

Fig. 1 SDF-1 expression in breast cancer cell lines. **a** Effect of estradiol on SDF-1 transcriptional activity in the five cell lines evaluated by quantitative RT-PCR. After treatment with 10 nM estradiol (*filled columns*) or ethanol vehicle (*clear columns*) for 48 h, total RNA was extracted. The relative SDF-1 mRNA values were calculated by estimating the ratio of SDF-1 copies to 18s-rRNA copies. *Columns*, mean values of three samples; *bars*, standard deviation; E2, estradiol. **b** Effect of estradiol on SDF-1 protein secretion by the five cell lines evaluated by ELISA. After treatment with 10 nM estradiol (*filled columns*) or ethanol vehicle (*clear columns*) for 72 h, the level of SDF-1 protein in the supernatant of each cell line was measured. *Columns*, mean values of three samples; *bars*, standard deviation; E2, estradiol; ND, not detectable. **c** Expression of SDF-1 protein detected by immunofluorescence in the ER-positive cell lines, MCF-7 and T47D. *Upper*: Cells were used after treatment with estradiol for 72 h. SDF-1 immunoreactivity was visualized in both cell lines. *Lower*: Immunofluorescence using murine IgG isotype control antibody showed no immunoreaction

the supernatant of MCF-7 and T47D cells after estradiol treatment were 117 and 378 pg/ml, respectively (filled columns in Fig. 1b). In contrast, SDF-1 protein was undetectable in the supernatant of these two ER-positive cell lines without estradiol treatment, or in the supernatant of the three ER-negative cell lines, irrespective of estradiol treatment (Fig. 1b).

Immunofluorescence analysis revealed diffuse and strong cytoplasmic SDF-1 immunoreactivity in almost all cells of the MCF-7 and T47D cell lines after estradiol treatment (Fig. 1c), whereas no, or only faint SDF-1 immunoreactivity was observed in the three ER-negative cell lines, regardless of estradiol treatment (data not shown). Taken together, these findings indicated that ER-positive breast cancer cell lines expressed SDF-1 in response to estradiol, whereas ER-negative breast cancer cell lines did not.

Correlation between protein expression pattern and mRNA level of SDF-1 in surgically resected breast cancers

Next, we conducted an immunohistochemical study to compare the expression levels of SDF-1 mRNA and protein in 52 samples of surgically resected invasive breast carcinoma and 13 samples of normal mammary gland tissue located adjacent to the tumors. Unremarkable mammary duct epithelia were found to have strong cytoplasmic immunoreactivity for SDF-1 (Fig. 2a). Among the 52 samples of breast cancer, 31 (60%), 18 (34%), and 3 (6%) showed C-, M-, and negative SDF-1 immunoreaction patterns, respectively (Fig. 2b–d). Cancer-associated fibroblasts (CAFs) sometimes showed weak SDF-1 staining (arrows in Fig. 2b–d), but the intensity and area of the

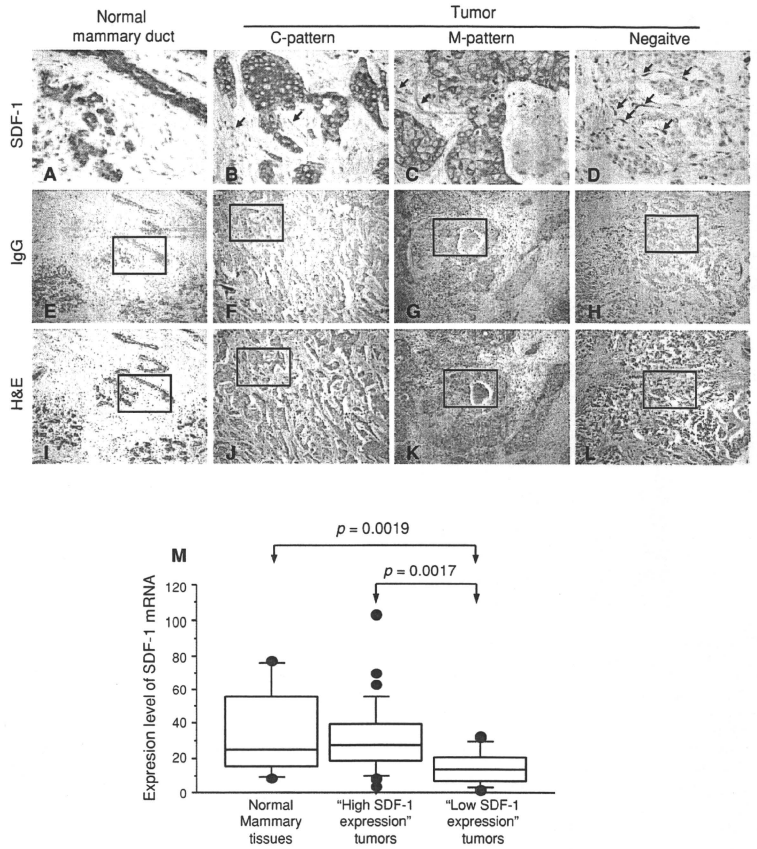

Fig. 2 SDF-1 expression in samples of human breast cancer. **a–d** SDF-1 immunoreaction pattern in surgically resected breast cancer tissues and normal mammary tissues. **a** Unremarkable mammary duct epithelia show strong cytoplasmic SDF-1 immunoreactivity. **b** C-pattern. Strong cytoplasmic-dominant immunoreactivity is evident in all constituent cells. **c** M-pattern. Moderate membrane-dominant immunoreactivity is evident in almost all constituent cells. **d** Negative immunoreaction pattern. **e–h** Immunohistochemistry using murine IgG control antibody for **a–d**. Murine IgG isotype control antibody showed no immunoreactivity. **i–l**, Hematoxylin and eosin (H&E) sections for **a–d**. H&E and IgG control images are shown at ×100 magnification with *boxed areas* indicating the ×400 magnification images shown for SDF-1 staining. Cancer-associated fibroblasts showed a weak immunoreaction for SDF-1 (*arrows* in Fig. b–d). **m** Correlation between results of immunohistochemistry and quantitative RT-PCR. The median expression levels of SDF-1 mRNA in normal mammary gland tissues ($n = 13$), in the "high SDF-1 expression" group comprising tumors with the C-pattern ($n = 31$), and in the "low SDF-1 expression" group comprising tumors with the M- and negative patterns ($n = 21$), were 25.0, 27.4 and 13.6, respectively. Bonferroni-adjusted comparisons between groups showed that mRNA levels differed significantly between the tumors with "high" and "low SDF-1 expression" ($P = 0.0017$), and between the normal mammary tissues and the tumors with "low SDF-1 expression" ($P = 0.0019$). The *Y*-axis represents the relative level of SDF-1 mRNA expression normalized against 18s-rRNA, calculated relative to the sample with the lowest expression, which was assigned a value of 1.0. *Boxes and whiskers*, 25th to 75th and 10th to 90th percentiles, respectively; the median is the central line in each box; circles, outliers

SDF-1 immunoreactivity on tumor cells were much stronger and wider than on CAFs. Immunohistochemistry using mouse IgG control antibody showed no immunoreactivity (Fig. 2e–h). RT-PCR showed that the median levels of SDF-1 mRNA expression in the groups of tumors with C-, M-, and negative patterns were 27.4 (interquartile range, 18.2–39.8), 14.2 (interquartile range, 8.5–21.7), and 6.0 (interquartile range, 4.5–20.4), respectively. The median levels of SDF-1 mRNA differed significantly between the C- and M-pattern groups and between the C- and negative pattern groups ($P = 0.0011$ and 0.031, respectively, Mann–Whitney U-test).

When the 52 tumor samples were classified into two groups, i.e., "high SDF-1 expression" tumors comprising the C-pattern group, and "low SDF-1 expression" tumors comprising the negative plus M-pattern groups, the median levels of SDF-1 mRNA were 27.4 (interquartile range, 18.2–39.8) and 13.6 (interquartile range, 6.5–20.5) in the tumors showing "high" and "low SDF-1 expression", respectively. The median level of SDF-1 mRNA in samples of normal mammary gland tissue was 25.0 (interquartile range, 15.2–58.4). The levels of SDF-1 mRNA differed significantly between the tumors showing "high" and "low SDF-1 expression" ($P = 0.0017$), and between normal mammary gland tissue and the tumors showing "low SDF-1 expression" ($P = 0.0019$, post hoc Bonferroni test; Fig. 2m). Taken together, the data indicated that the pattern of immunoreactivity for SDF-1 protein was significantly correlated with the level of its mRNA.

Clinicopathological and prognostic implications of SDF-1 immunoreactivity for the entire patient cohort

We next evaluated the patterns of SDF-1 protein expression using another large cohort of patients with invasive breast cancer for whom long-term follow-up data were available, in order to clarify the clinicopathological and prognostic significance of tumor-derived SDF-1. One-hundred and fifty-eight patients (71%) had tumors showing "high SDF-1 expression" and 65 (29%) had tumors showing "low SDF-1 expression", the latter comprising 55 (24%) tumors with an M-pattern and 10 (5%) with a negative pattern. "High SDF-1 expression" showed a more significant association with younger patient age ($\leq$52 years, $P = 0.018$), HER2 negativity ($P = 0.021$), and lower nuclear grade ($P < 0.0001$) than "low SDF-1 expression". In particular, positive ER staining was detected more frequently in tumors with "high SDF-1 expression" (109 of 158, 69%) than in those with "low SDF-1 expression" (24 of 65, 37%) ($P < 0.0001$, Table 1). Tumor size, lymph node status, or clinical stage was not significantly correlated with SDF-1 status. Tumor histology also showed no significant correlation with SDF-1 status, but all 10 invasive lobular carcinomas showed "high SDF-1 expression". Nuclear CXCR4 staining was positive in 78% (123 of 158) of tumors with "high SDF-1 expression," and this proportion was significantly higher than that among tumors with "low SDF-1 expression" (58%, 38 of 65, $P = 0.0033$, Table 1).

Among the 223 patients examined, there was a significant difference in the disease-free and overall survival curves between the "high" and "low SDF-1 expression" groups ($P = 0.0058$ and 0.017, respectively; Fig. 3a, b). Univariate analysis also showed that clinical stage, nuclear grade, ER status, and HER2 status were significant or almost significant indicators of clinical outcome (Table 2). Multivariate analysis using the Cox proportional hazard model including these indicators showed that SDF-1 status was a factor significantly predictive of disease-free survival ($P = 0.036$), along with clinical stage ($P < 0.0001$) and nuclear grade ($P = 0.0012$). However, only clinical stage was selected as a significant prognostic factor for overall survival ($P < 0.0001$), and the impact of SDF-1 status was marginal ($P = 0.099$, Table 2). Univariate analysis showed that CXCR4 was of no prognostic significance (data not shown).

Clinicopathological and prognostic implications of SDF-1 immunoreactivity among ER-positive patients

As our data clearly indicated a marked correlation between the expressions of ER and SDF-1 in vitro and in vivo, we carried out further analysis to clarify the significance of SDF-1 expression in the subgroup of patients with ER-positive tumors. Among the 133 cases of ER-positive breast cancer, 109 (82%) showed "high SDF-1 expression." SDF-1 status was not significantly correlated with nuclear grade, tumor size, lymph node status, stage, histological type, or nuclear CXCR4 expression, but "high SDF-1 expression" was correlated with younger patient age ($P = 0.0025$; Table 3). A HER2 score of 3+ was detected in only two cases, both of which showed "low SDF-1 expression". Additionally, among only five cases that had distant metastasis at the time of diagnosis, two showed "high SDF-1 expression" and three showed "low SDF-1 expression".

In the 133 ER-positive patients, there was a significant difference in the disease-free and overall survival curves between the groups showing "high" and "low SDF-1 expression" ($P = 0.0047$ and 0.036, respectively; Fig. 3c, d). Furthermore, in these cases, multivariate analysis showed that SDF-1 status was an independent prognostic factor not only for disease-free survival, but also for overall survival ($P = 0.015$ and 0.046, respectively; Table 4). Nuclear grade and clinical stage also had a significant impact on overall survival (Table 4). Among ER-negative cases, there was no significant difference in the survival curves between the two groups (Fig. 3e, f).

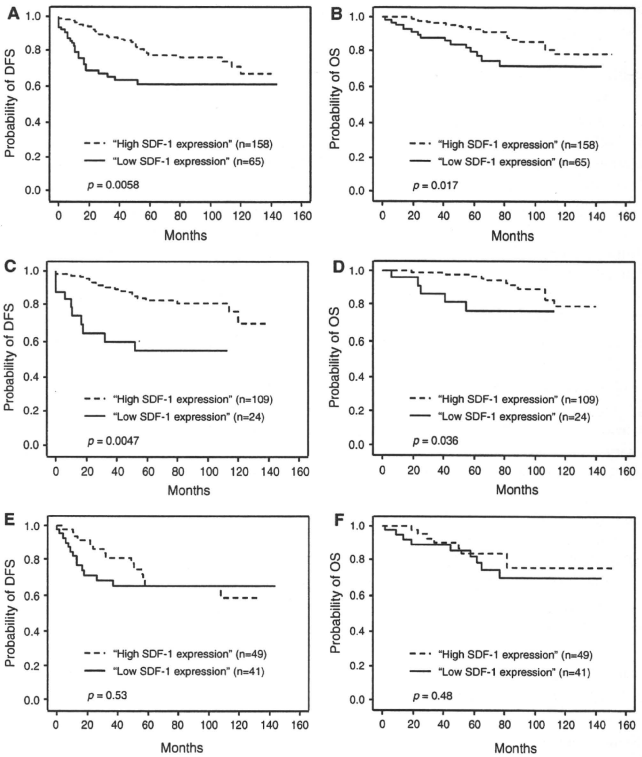

Fig. 3 Prognostic impact of SDF-1 status detected by immunohistochemistry in the patients with primary breast cancer. **a–b** all 223 patients; **c–d** 133 patients with ER-positive tumors; **e–f** 90 patients with ER-negative tumors. **a** Disease-free survival curves for the 158 patients whose tumors showed "high SDF-1 expression" and the 65 patients whose tumors showed "low SDF-1 expression." The two curves differ significantly ($P = 0.0058$). **b** Overall survival curves for the 158 patients whose tumors showed "high SDF-1 expression" and the 65 patients whose tumors showed "low SDF-1 expression." The two curves differ significantly ($P = 0.017$). **c** Disease-free survival curves for the 109 patients whose tumors showed "high SDF-1 expression" and the 24 patients whose tumors showed "low SDF-1 expression". The two curves differ significantly ($P = 0.0047$). **d** Overall survival curves for the 109 patients whose tumors showed "high SDF-1 expression" and the 24 patients whose tumors showed "low SDF-1 expression". The two curves differ significantly ($P = 0.036$). **e** Disease-free survival curves for the 49 patients whose tumors showed "high SDF-1 expression" and the 41 patients whose tumors showed "low SDF-1 expression". The two curves do not differ significantly ($P = 0.53$) **f** Overall survival curves for the 49 patients whose tumors showed "high SDF-1 expression" and the 41 patients whose tumors showed "low SDF-1 expression." The two curves do not differ significantly ($P = 0.48$). All P values were calculated using the Cox proportional hazards model

Discussion

In the present study, we showed that the expression of SDF-1 was regulated by estradiol in the ER-positive cell lines, MCF-7, and T47D, and found that SDF-1 status was significantly correlated with several clinically important factors, especially ER status, in samples of human breast cancer. In addition, we revealed that SDF-1 status was a statistically significant prognostic factor among cases of ER-positive breast cancer. To our knowledge, this is the first study to have demonstrated the importance of tumor-derived SDF-1 in ER-positive breast cancers.

Table 2 Prognostic indicators detected by Cox's univariate and multivariate analyses in patients with primary breast cancer

Variable	Univariate			Multivariate		
	HR	(95%CI)	*P* value	HR	(95%CI)	*P* value
Disease-free survival						
SDF-1						
"High expression"	1		0.0058	1		0.036
"Low expression"	1.45	(1.11–1.89)		1.38	(1.02–1.86)	
Clinical stage						
1 or 2	1		<0.0001	1		<0.0001
3 or 4	2.83	(2.16–3.72)		2.67	(2.01–3.54)	
Nuclear grade						
1	1		<0.0001	1		0.0012
2	1.30	(0.78–2.52)		1.51	(0.90–2.93)	
3	3.10	(1.93–5.90)		2.34	(1.40–4.55)	
ER						
Positive	1		0.10	1		0.39
Negative	1.24	(0.96–1.61)		0.88	(0.65–1.19)	
HER2						
Negative	1		0.025	1		0.21
Positive	1.57	(1.06–2.19)		1.29	(0.86–1.86)	
Overall survival						
SDF-1						
"High expression"	1		0.017	1		0.099
"Low expression"	1.56	(1.09–2.21)		1.39	(0.94–2.07)	
Clinical stage						
1 or 2	1		<0.0001	1		<0.0001
3 or 4	3.34	(2.34–4.88)		3.23	(2.23–4.78)	
Nuclear grade						
1	1		0.0002	1		0.16
2	1.22	(0.60–3.38)		1.32	(0.65–3.68)	
3	3.36	(1.79–9.13)		1.96	(0.96–5.49)	
ER						
Positive	1		0.047	1		0.58
Negative	1.42	(1.00–2.03)		1.23	(0.74–1.73)	
HER2						
Negative	1		0.0068	1		0.11
Positive	1.94	(1.22–2.87)		1.49	(0.90–2.35)	

Abbreviation: *95%CI* 95% confidence interval

Many researchers have investigated the role of the CXCR4/SDF-1 axis in tumor biology, and demonstrated the importance of CXCR4 expression for tumor progression in vitro [12–20] and in vivo [21–26]. In addition, some types of cancer cell lines have been shown to produce both SDF-1 and its receptor, CXCR4, and to use this signaling system in an autocrine manner [6, 27–29]. A number of clinical studies of glioma [30], colorectal cancer [31], gastric cancer [32], and oral squamous cell carcinoma [33] have concluded that higher SDF-1 expression is correlated with lymph node and/or distant metastasis and poorer clinical outcome.

In contrast, using a mouse model, Wendt et al. [34, 35] showed that breast and colon cancer cell lines with forced expression of SDF-1 established fewer secondary tumors than those with null SDF-1 expression. In patients with testicular germ cell tumors, high SDF-1 expression was correlated with longer relapse-free survival [36]. These differences in the clinical impact of tumor-derived SDF-1 might depend on the type of cancer.

Table 3 Clinicopathological implication of SDF-1 expression in surgically resected ER-positive breast cancers

	Total (n = 133)	SDF-1		
		"High expression" (n = 109) [n (%)]	"Low expression" (n = 24) [n (%)]	P value
Age				
≤52	67	60 (55)	7 (29)	
>52	66	49 (45)	17 (71)	0.025
Nuclear grade				
1	38	32 (29)	6 (24)	
2	65	56 (51)	9 (38)	
3	30	21 (20)	9 (38)	0.15
Tumor size				
<5.0 cm	109	90 (82)	19 (79)	
≥5.0 cm	20	15 (14)	5 (21)	0.53
Unknown	4	4 (4)	0 (0)	
Lymph node metastasis				
(−)	72	60 (55)	12 (50)	
(+)	57	45 (41)	12 (50)	0.65
Unknown	4	4 (4)	0 (0)	
Stage				
1 or 2	109	91 (83)	18 (75)	
3 or 4	20	14 (13)	6 (25)	0.21
Unknown	4	4 (4)	0 (0)	
Tumor histology				
IDC	116	94 (86)	22 (92)	
ILC	6	6 (6)	0 (0)	
Special[a]	11	9 (8)	2 (8)	0.49
CXCR4				
Negative	32	24 (22)	8 (33)	
Positive	101	85 (78)	16 (67)	0.29

Abbreviation: *IDC* invasive ductal carcinoma, *ILC* invasive lobular carcinoma

[a] Mucinous carcinoma (5), Tubular carcinoma (5), Micropapillary (1)

Several previous reports and our results have indicated that SDF-1 is one of the ERGs [6, 37–39]. Hall and Korach [6] showed that SDF-1 induced proliferation of ER-positive breast and ovarian cancer cell lines, and proposed that SDF-1 might be a strong mediator of estrogen-induced cell proliferation. However, our clinical data demonstrated that tumors with high SDF-1 expression showed more indolent characteristics, i.e., ER positivity, a lower tumor grade, and lack of HER2 overexpression. At present, the function of SDF-1 as an ERG in ER-positive tumors in vivo remains unclear.

We conducted subset analysis of the prognostic significance of SDF-1 immunoreactivity in groups of patients with ER-positive and ER-negative tumors. In a subset of 133 ER-positive tumors, multivariate analysis showed that "high SDF-1 expression," as well as lower nuclear grade and earlier clinical stage, were independent indicators of better prognosis. On the other hand, in a subset of 90 ER-negative tumors, the "high" and "low SDF-1 expression" groups had similar clinical outcomes. In general, ERG expression has been thought to reflect the activity of ER signaling pathways in breast cancers, and it has been proved that several ERGs, e.g., PR, are indicators of responsiveness to endocrine therapy [4, 40]. Therefore, it is possible that SDF-1 immunoreactivity might be a new and powerful predictor of endocrine responsiveness in breast cancer cells.

In our present study, 54% (49 of 90) of ER-negative breast cancers had "high SDF-1 expression." It has been shown that SDF-1 expression can be induced by hypoxia-induced factor-1 (HIF-1) in endothelial cells [41], by fibroblast growth factor-2 (FGF-2) in stromal cells [42], and by bone morphogenetic protein (BMP) in cancer-associated fibroblasts (CAFs) [43]. Thus, multiple alternative pathways mediated by these molecules may be involved in the upregulation of SDF-1 in ER-negative breast cancers.

Interestingly, the intracellular distribution pattern of SDF-1 immunoreactivity in surgically resected breast cancers was a distinctive characteristic, and significantly

Table 4 Prognostic indicators detected by Cox's univariate and multivariate analyses in patients with ER-positive primary breast cancer

Variable	Univariate			Multivariate		
	HR	(95%CI)	*P* value	HR	(95%CI)	*P* value
Disease-free survival						
SDF-1						
"High expression"	1		0.0047	1		0.015
"Low expression"	1.87	(1.22–2.67)		1.70	(1.12–2.52)	
Clinical stage						
1 or 2	1		<0.0001	1		<0.0001
3 or 4	2.87	(1.96–4.19)		2.53	(1.70–3.75)	
Nuclear grade[a]						
1, 2	1		<0.0001	1		0.0041
3	2.28	(1.59–3.30)		1.77	(1.20–2.62)	
Overall survival						
SDF-1						
"High expression"	1		0.036	1		0.046
"Low expression"	1.89	(1.05–3.21)		1.86	(1.01–3.26)	
Clinical stage						
1 or 2	1		0.0009	1		0.0040
3 or 4	2.57	(1.50–4.39)		2.28	(1.32–3.94)	
Nuclear grade[a]						
1, 2	1		0.0008	1		0.019
3	2.42	(1.45–4.19)		1.92	(1.12–3.37)	

Abbreviation: *95%CI* 95% confidence interval

[a] There was no relapse and death among the cases with ER-positive and nuclear grade 1, so we combined the category of "nuclear grade 1" with "2" and performed subset analyses of ER-positive patients

correlated with the amount of SDF-1 mRNA. Tumors with cytoplasmic-dominant immunoreactivity had a higher level of SDF-1 mRNA than those with negative and membrane-dominant immunoreactivity. Therefore, we evaluated SDF-1 protein expression immunohistochemically in the present cohort of 223 invasive breast cancers based on criteria emphasizing its intracellular distribution. In previous studies, SDF-1 immunoreactivity was detected on the cell membrane of gastric and ovarian tumor cells [32, 44], and in the cytoplasm of colorectal tumor cells [31]. No previous report has described both membranous and cytoplasmic immunoreactivity in a single type of cancer. Further studies are needed to validate the reproducibility of the present criteria for judgment of immunohistochemical data.

We showed that "high SDF-1 expression" was significantly correlated with nuclear expression of CXCR4 in all 223 breast cancers. This finding was concordant with previously reported results indicating that SDF-1 stimulation induced rapid nuclear internalization of CXCR4 [45, 46], and confirmed that the CXCR4/SDF-1 axis plays an important role in the progression of breast cancer.

In summary, the present study revealed that breast cancers showing "high SDF-1 expression" have a higher frequency of ER positivity, HER2 negativity, lower nuclear grade, and better patient outcome, not only overall, but also in patients with ER-positive tumors. Examination of SDF-1 expression in ER-positive invasive breast cancers might be useful for identification of patients with a potentially better clinical outcome, and could help avoid the prescription of unnecessary chemotherapy for them.

Acknowledgments We thank Dr. Yumi Miyazaki and Ms. Kozue Suzuki for technical assistance and Dr. Keiichi Iwaya for helpful discussion. This work was supported by grants from the Ministry of Defense, Japan, the Ministry of Health, Labor, and Welfare, Japan, the Princess Takamatsu Cancer Research Fund, and the Foundation for Promotion of Defense Medicine.

References

1. Marugame T, Matsuda T, Kamo K, Katanoda K, Ajiki W, Sobue T (2007) Cancer incidence and incidence rates in Japan in 2001 based on the data from 10 population-based cancer registries. Jpn J Clin Oncol 37(11):884–891
2. Yoshimoto M, Tada K, Hori H et al (2004) Improvement in the prognosis of Japanese breast cancer patients from 1946 to 2001—an institutional review. Jpn J Clin Oncol 34(8):457–462
3. EBCTCG (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365(9472):1687–1717

4. Ciocca DR, Elledge R (2000) Molecular markers for predicting response to tamoxifen in breast cancer patients. Endocrine 13(1):1–10
5. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn HJ (2007) Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer. Ann Oncol 18(7):1133–1144
6. Hall JM, Korach KS (2003) Stromal cell-derived factor 1, a novel target of estrogen receptor action, mediates the mitogenic effects of estradiol in ovarian and breast cancer cells. Mol Endocrinol 17(5):792–803
7. Tashiro K, Tada H, Heilker R, Shirozu M, Nakano T, Honjo T (1993) Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins. Science 261(5121):600–603
8. Nagasawa T, Kikutani H, Kishimoto T (1994) Molecular cloning and structure of a pre-B-cell growth-stimulating factor. Proc Natl Acad Sci USA 91(6):2305–2309
9. Burger JA, Kipps TJ (2006) CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. Blood 107(5):1761–1767
10. Kulbe H, Levinson NR, Balkwill F, Wilson JL (2004) The chemokine network in cancer—much more than directing cell movement. Int J Dev Biol 48(5–6):489–496
11. Shinto E, Tsuda H, Ueno H et al (2005) Prognostic implication of laminin-5 gamma 2 chain expression in the invasive front of colorectal cancers, disclosed by area-specific four-point tissue microarrays. Lab Invest 85(2):257–266
12. Muller A, Homey B, Soto H et al (2001) Involvement of chemokine receptors in breast cancer metastasis. Nature 410(6824):50–56
13. Barbero S, Bonavia R, Bajetto A et al (2003) Stromal cell-derived factor 1alpha stimulates human glioblastoma cell growth through the activation of both extracellular signal-regulated kinases 1/2 and Akt. Cancer Res 63(8):1969–1974
14. Kijima T, Maulik G, Ma PC et al (2002) Regulation of cellular proliferation, cytoskeletal function, and signal transduction through CXCR4 and c-Kit in small cell lung cancer cells. Cancer Res 62(21):6304–6311
15. Yoon Y, Liang Z, Zhang X et al (2007) CXC chemokine receptor-4 antagonist blocks both growth of primary tumor and metastasis of head and neck cancer in xenograft mouse models. Cancer Res 67(15):7518–7524
16. Chinni SR, Sivalogan S, Dong Z et al (2006) CXCL12/CXCR4 signaling activates Akt-1 and MMP-9 expression in prostate cancer cells: the role of bone microenvironment-associated CXCL12. Prostate 66(1):32–48
17. Ehtesham M, Winston JA, Kabos P, Thompson RC (2006) CXCR4 expression mediates glioma cell invasiveness. Oncogene 25(19):2801–2806
18. Fernandis AZ, Prasad A, Band H, Klosel R, Ganju RK (2004) Regulation of CXCR4-mediated chemotaxis and chemoinvasion of breast cancer cells. Oncogene 23(1):157–167
19. Retz MM, Sidhu SS, Blaveri E et al (2005) CXCR4 expression reflects tumor progression and regulates motility of bladder cancer cells. Int J Cancer 114(2):182–189
20. Orimo A, Gupta PB, Sgroi DC et al (2005) Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell 121(3):335–348
21. Pan J, Mestas J, Burdick MD et al (2006) Stromal derived factor-1 (SDF-1/CXCL12) and CXCR4 in renal cell carcinoma metastasis. Mol Cancer 5:56
22. Ottaiano A, Franco R, Aiello Talamanca A et al (2006) Overexpression of both CXC chemokine receptor 4 and vascular endothelial growth factor proteins predicts early distant relapse in stage II-III colorectal cancer patients. Clin Cancer Res 12(9):2795–2803
23. Kato M, Kitayama J, Kazama S, Nagawa H (2003) Expression pattern of CXC chemokine receptor-4 is correlated with lymph node metastasis in human invasive ductal carcinoma. Breast Cancer Res 5(5):R144–R150
24. Cabioglu N, Yazici MS, Arun B et al (2005) CCR7 and CXCR4 as novel biomarkers predicting axillary lymph node metastasis in T1 breast cancer. Clin Cancer Res 11(16):5686–5693
25. Kim J, Takeuchi H, Lam ST et al (2005) Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. J Clin Oncol 23(12):2744–2753
26. Spano JP, Andre F, Morat L et al (2004) Chemokine receptor CXCR4 and early-stage non-small cell lung cancer: pattern of expression and correlation with outcome. Ann Oncol 15(4):613–617
27. Rempel SA, Dudas S, Ge S, Gutierrez JA (2000) Identification and localization of the cytokine SDF1 and its receptor, CXC chemokine receptor 4, to regions of necrosis and angiogenesis in human glioblastoma. Clin Cancer Res 6(1):102–111
28. Brand S, Dambacher J, Beigel F et al (2005) CXCR4 and CXCL12 are inversely expressed in colorectal cancer cells and modulate cancer cell migration, invasion and MMP-9 activation. Exp Cell Res 310(1):117–130
29. Zagzag D, Krishnamachary B, Yee H et al (2005) Stromal cell-derived factor-1alpha and CXCR4 expression in hemangioblastoma and clear cell-renal cell carcinoma: von Hippel-Lindau loss-of-function induces expression of a ligand and its receptor. Cancer Res 65(14):6178–6188
30. Salmaggi A, Gelati M, Pollo B et al (2005) CXCL12 expression is predictive of a shorter time to tumor progression in low-grade glioma: a single-institution study in 50 patients. J Neurooncol 74(3):287–293
31. Yoshitake N, Fukui H, Yamagishi H et al (2008) Expression of SDF-1 alpha and nuclear CXCR4 predicts lymph node metastasis in colorectal cancer. Br J Cancer 98(10):1682–1689
32. Ishigami S, Natsugoe S, Okumura H et al (2007) Clinical implication of CXCL12 expression in gastric cancer. Ann Surg Oncol 14(11):3154–3158
33. Uchida D, Onoue T, Tomizuka Y et al (2007) Involvement of an autocrine stromal cell derived factor-1/CXCR4 system on the distant metastasis of human oral squamous cell carcinoma. Mol Cancer Res 5(7):685–694
34. Wendt MK, Cooper AN, Dwinell MB (2008) Epigenetic silencing of CXCL12 increases the metastatic potential of mammary carcinoma cells. Oncogene 27(10):1461–1471
35. Wendt MK, Drury LJ, Vongsa RA, Dwinell MB (2008) Constitutive CXCL12 expression induces anoikis in colorectal carcinoma cells. Gastroenterology 135(2):508–517
36. Gilbert DC, Chandler I, McIntyre A et al (2009) Clinical and biological significance of CXCL12 and CXCR4 expression in adult testes and germ cell tumours of adults and adolescents. J Pathol 217(1):94–102
37. Rae JM, Johnson MD, Scheys JO, Cordero KE, Larios JM, Lippman ME (2005) GREB 1 is a critical regulator of hormone dependent breast cancer growth. Breast Cancer Res Treat 92(2):141–149
38. Frasor J, Danes JM, Komm B, Chang KC, Lyttle CR, Katzenellenbogen BS (2003) Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. Endocrinology 144(10):4562–4574
39. Pattarozzi A, Gatti M, Barbieri F et al (2008) 17beta-estradiol promotes breast cancer cell proliferation-inducing stromal cell-derived factor-1-mediated epidermal growth factor receptor

transactivation: reversal by gefitinib pretreatment. Mol Pharmacol 73(1):191–202
40. Yoshida N, Omoto Y, Inoue A et al (2004) Prediction of prognosis of estrogen receptor-positive breast cancer with combination of selected estrogen-regulated genes. Cancer Sci 95(6):496–502
41. Ceradini DJ, Kulkarni AR, Callaghan MJ et al (2004) Progenitor cell trafficking is regulated by hypoxic gradients through HIF-1 induction of SDF-1. Nat Med 10(8):858–864
42. Nakayama T, Mutsuga N, Tosato G (2007) Effect of fibroblast growth factor 2 on stromal cell-derived factor 1 production by bone marrow stromal cells and hematopoiesis. J Natl Cancer Inst 99(3):223–235
43. Yang S, Pham LK, Liao CP, Frenkel B, Reddi AH, Roy-Burman P (2008) A novel bone morphogenetic protein signaling in heterotypic cell interactions in prostate cancer. Cancer Res 68(1):198–205
44. Pils D, Pinter A, Reibenwein J et al (2007) In ovarian cancer the prognostic influence of HER2/neu is not dependent on the CXCR4/SDF-1 signalling pathway. Br J Cancer 96(3):485–491
45. Haribabu B, Richardson RM, Fisher I et al (1997) Regulation of human chemokine receptors CXCR4. Role of phosphorylation in desensitization and internalization. J Biol Chem 272(45):28726–28731
46. Fischer T, Nagel F, Jacobs S et al (2008) Reassessment of CXCR4 chemokine receptor expression in human normal and neoplastic tissues using the novel rabbit monoclonal antibody UMB-2. PLoS One 3(12):e4069

Tsuda et al. BMC Cancer 2010, 10:534
http://www.biomedcentral.com/1471-2407/10/534

BMC Cancer

RESEARCH ARTICLE Open Access

HER2 testing on core needle biopsy specimens from primary breast cancers: interobserver reproducibility and concordance with surgically resected specimens

Hitoshi Tsuda[1,4*], Masafumi Kurosumi[2], Shinobu Umemura[3], Sohei Yamamoto[4], Takayuki Kobayashi[4], Robert Yoshiyuki Osamura[3]

Abstract

Background: Accurate evaluation of human epidermal growth factor receptor type-2 (HER2) status based on core needle biopsy (CNB) specimens is mandatory for identification of patients with primary breast cancer who will benefit from primary systemic therapy with trastuzumab. The aim of the present study was to validate the application of HER2 testing with CNB specimens from primary breast cancers in terms of interobserver reproducibility and comparison with surgically resected specimens.

Methods: A total of 100 pairs of archival formalin-fixed paraffin-embedded CNB and surgically resected specimens of invasive breast carcinomas were cut into sections. All 100 paired sections were subjected to HER2 testing by immunohistochemistry (IHC) and 27 paired sections were subjected to that by fluorescence in situ hybridization (FISH), the results being evaluated by three and two observers, respectively. Interobserver agreement levels in terms of judgment and the concordance of consensus scores between CNB samples and the corresponding surgically resected specimens were estimated as the percentage agreement and κ statistic.

Results: In CNB specimens, the percentage interobserver agreement of HER2 scoring by IHC was 76% ($\kappa = 0.71$) for 3 × 3 categories (0-1+ *versus* 2+ *versus* 3+) and 90% ($\kappa = 0.80$) for 2 × 2 categories (0-2+ *versus* 3+). These levels were close to the corresponding ones for the surgically resected specimens: 80% ($\kappa = 0.77$) for 3 × 3 categories and 92% ($\kappa = 0.88$) for 2 × 2 categories. Concordance of consensus for HER2 scores determined by IHC between CNB and the corresponding surgical specimens was 87% ($\kappa = 0.77$) for 3 × 3 categories, and 94% ($\kappa = 0.83$) for 2 × 2 categories. Among the 13 tumors showing discordance in the mean IHC scores between the CNB and surgical specimens, the results of consensus for FISH results were concordant in 11. The rate of successful FISH analysis and the FISH positivity rate in cases with a HER2 IHC score of 2+ differed among specimens processed at different institutions.

Conclusion: It is mandatory to study HER2 on breast cancers, and either CNB or surgical specimen can be used.

Background

The human epidermal growth factor receptor type-2 (*HER2*) proto-oncogene (c-*erb*B-2) is amplified in 15-30% of human breast cancers, causing overexpression of its protein. *HER2* gene amplification and/or protein overexpression is an indicator of clinical tumor aggressiveness [1-3]. The efficacy of trastuzumab, a humanized anti-HER2 monoclonal antibody, against breast cancers with *HER2* gene amplification and/or protein overexpression has been established in clinical trials for patients with metastatic breast cancer or those with operable primary breast cancer as adjuvant systemic therapies [4-7]. Furthermore, as neoadjuvant therapy for patients with breast cancers showing HER2 amplification and/or

* Correspondence: hstsuda@ncc.go.jp
[1]Diagnostic Pathology Section, Clinical Laboratory Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
Full list of author information is available at the end of the article

overexpression, therapies involving a combination of trastuzumab and chemotherapy have been shown to be effective in achieving a complete pathological tumor response [8,9].

HER2 testing comprises immunohistochemistry (IHC) to examine protein overexpression and fluorescence *in situ* hybridization (FISH) to examine gene amplification. These tests are performed on tissue sections of routinely-processed formalin-fixed, paraffin-embedded tumors. High accuracy is required for these HER2 tests in order to identify patients who would benefit from trastuzumab therapy. For the test algorithm, it is generally recommended that IHC is performed first, and that FISH is added if the result of IHC is equivocal [10-12]. Studies of quality assessment have demonstrated that interobserver agreement levels are high for tumors with an IHC score of 0 or 1+, or those with a score of 3+, but that the level is generally low for those with a score of 2+ [13-16]. A higher interobserver agreement level can be achieved with FISH than with IHC, especially for tumors with an IHC score of 2+ [13,14,16]. It has also been shown that the quality of both tests is higher at institutions that perform a larger number of HER2 tests than at those where a smaller number of such tests are performed [17-19].

In recent years, core needle biopsies (CNBs) have been used for the qualitative diagnosis of breast tumors. Because of the prevalence of primary systemic therapies as a standard treatment for primary breast cancers, not only hormone receptor status but also HER2 status are generally assayed from CNB specimens to test the eligibility of patients for primary systemic therapy with trastuzumab [20]. However, it would be expected that examination of CNB specimens alone might result in a proportion of false-positive and/or false-negative results, because CNB samples represent only part of the tumor, notwithstanding the possible presence of intratumor heterogeneity [21-25]. Furthermore, because of the small volume of CNB specimens, the interobserver agreement rate of HER2 tests for CNB specimens might be lower than for surgically resected specimens.

Therefore, we examined the levels of interobserver agreement for HER2 status determination in both CNB specimens and corresponding surgically resected specimens from 100 patients with primary breast cancer who had not received primary systemic therapy. We compared the HER2 protein status determined by IHC between the CNB specimens and surgically resected specimens of the same tumor. We also compared the HER2 protein status determined by IHC, with the *HER2* gene status determined by FISH, in CNB specimens and in the corresponding surgically resected specimens. On the basis of these results, we evaluated the utility and challenges of HER2 testing within CNB specimens.

Methods

Tissue samples

We examined 100 paired samples of invasive breast carcinoma obtained by CNB and surgical resection from patients treated at Saitama Cancer Center, Ina, Saitama (50 cases), Tokai University Hospital, Isehara, Kanagawa (25 cases), and the National Defense Medical College Hospital, Tokorozawa, Saitama (25 cases), Japan. At all three institutions, HER2 testing is performed very frequently for routine diagnostics and/or for studies of quality assessment. Collaborating pathologists in the three institutes were assigned to submit almost equal number of CNB cases of each score (score 0 or 1+, 2+, and 3+), for the purpose that almost equal number of HER2-negative, equivocal and positive cases were examined in the study. However, these institutional scores were not informed to the pathologists on the central review. Therefore, the cases were not consecutive and there was some selection bias.

None of the patients with these tumors had received neoadjuvant therapy before CNB and surgical resection. At each institution, 4-μm-thick sections cut from routinely processed formalin-fixed, paraffin-embedded tissue blocks were subjected to immunohistochemistry and FISH. IHC and FISH assays were performed on parallel slides of the same CNB/surgically resected specimens. The present study was conducted with approval from the internal review board for ethical issues of the National Defense Medical College. Informed consent had been acquired from each patient for the purpose of general research use of surgically resected tissues, and the requirement for informed consent for the present study was waived by the internal review board according to the guideline of ethical issues for epidemiologic studies by the Ministry of Health, Labor and Welfare and the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Immunohistochemistry

Expression of HER2 oncoprotein was examined using HercepTest II (Dako, Glostrup, Denmark) in the Dako Japan Central Laboratory, Tokyo, Japan. Deparaffinization, antigen retrieval, and immunohistochemistry were performed to 100 CNB and 100 paired surgically resected tumor sections in accordance with the manufacturer's instructions using an Autostainer Plus (Dako) [16].

Three experienced (for >25 years) pathologists (M.K., S.U., H.T.), being blinded from institutional IHC results or the present FISH results, independently evaluated the results of IHC, and assigned a score of 0 (no staining), 1 + (weak, incomplete membrane staining in any proportion of tumor cells), 2+ (complete membrane staining that is either nonuniform or weak in intensity but with

obvious circumferential distribution in at least 10% of tumor cells, or invasive tumors show intense, complete membrane staining of 30% or fewer tumor cells), or 3+ (uniform, intense membrane staining of >30% of invasive tumor cells) in accordance with the guidelines of the American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) [10]. Scores of 0 and 1+ were categorized as IHC-negative, a score of 2+ was categorized as equivocal, and a score of 3+ was categorized as overexpression (positive). If the score assigned by three observers differed among the three categories, the majority scores were acquired as consensus judgments. When the judgments of the three observers differed from each other, the median value was acquired as the representative score.

FISH

FISH was performed on all 100 CNB specimens and on 27 surgically resected specimens for which the consensus judgment of the IHC result differed from that of the CNB result and/or was discordant among the three observers.

FISH was performed manually using a PathVysion *HER*-2 DNA probe kit (Abbott Molecular, Wiesbaden, Germany) in accordance with the manufacturer's instructions [13] at the Department of Basic Pathology, National Defense Medical College, Tokorozawa, Saitama, Japan. The slides were visualized using a Leica DMR fluorescence microscope (Leica, Cambridge, UK). Two observers (S.Y. or T.K. and H.T.), being blinded from IHC results and FISH results of the corresponding CNB or surgical resection specimen, counted the signals for *HER2* and *CEP17* on a total of 40 cancer cell nuclei. The total number of *HER2* signals was divided by the total number of *CEP17* signals on 40 nuclei, and the *HER2*/*CEP17* ratio was calculated. *HER2* gene amplification was judged as positive, equivocal, and negative if the *HER2*/*CEP17* ratio was more than 2.2, 1.8 to 2.2, and less than 1.8, respectively [10]. When a tumor was judged as equivocal, 40 additional nuclei in another tumor area on invasion were counted again by two observers. If the *HER2*/*CEP17* ratio was 2.0 or higher in the re-test, the tumor was finally judged as having *HER2* amplification.

Interobserver agreement

With regard to IHC and FISH, interobserver agreement of judgments and the concordance of consensus scores between the CNB and corresponding surgically resected specimens were estimated in terms of percentage agreement and the κ statistic. The percentage of agreement was calculated as follows:

(Number of tumors to which the three observers assigned an identical score/total number of tumors) × 100

The level of agreement was categorized as almost perfect, substantial, moderate, fair, and slight when the κ value was >0.80, >0.60-0.80, >0.40-0.60, >0.20-0.40, and 0-0.20, respectively [26,27].

Results

Interobserver agreement for IHC and FISH

For CNB specimens, the percentage interobserver agreement for HER2 scores determined by IHC was 76% for 3 × 3 categories (0-1+ vs 2+ vs 3+) and 90% for 2 × 2 categories (0, 1+ or 2+ vs 3+). In terms of the κ statistic, the interobserver agreement levels were substantial for the 3 × 3 categories (κ = 0.71, σ = 0.065) and the 2 × 2 categories (κ = 0.80, σ = 0.075). The 24 tumors for which judgment discordance arose for 3 × 3 categories are shown in Table 1. All of these tumors showed discordance in one score, and none showed a difference in 2 scores.

Table 1 Core needle biopsy tumor specimens for which interobserver disagreement arose regarding the results of HER2 immunohistochemistry

Code	Final score	IHC score			FISH *(HER2/CEP17)*
		Observer A	Observer B	Observer C	
B45	1+	1+	1+	2+	- (0.96)
B78	1+	1+	1+	2+	- (0.99)
B24	1+	2+	1+	1+	- (1.01)
B82	1+	2+	1+	1+	- (1.02)
B79	1+	2+	1+	1+	- (1.05)
B99	1+	1+	1+	2+	- (1.07)
B36	1+	1+	1+	2+	- (1.11)
B22	2+	2+	1+	2+	- (0.97)
B69	2+	2+	1+	2+	- (1.04)
B43	2+	2+	1+	2+	- (1.09)
B86	2+	2+	1+	2+	- (1.39)
B29	2+	2+	2+	1+	- (1.42)
B100	2+	2+	1+	2+	+ (2.38)
B16	2+	3+	2+	2+	+ (2.56)
B97	2+	2+	1+	2+	+ (4.56)
B52	2+	3+	2+	2+	+ (5.44)
B91	2+	2+	2+	3+	+ (6.80)
B90	2+	3+	2+	2+	+ (10.38)
B102	2+	3+	2+	2+	+ (12.62)
B50	3+	3+	3+	2+	+ (5.11)
B7	3+	3+	2+	3+	+ (8.83)
B96	3+	3+	2+	3+	+ (9.02)
B95	3+	3+	2+	3+	+ (10.92)
B27	3+	3+	3+	2+	+ (12.50)

The HER2 score was assigned according to the 2007 ASCO/CAP guideline [10]. The majority score was assigned as the final score for the tumor.
FISH, fluorescence in situ hybridization; IHC, immunohistochemistry

For surgically resected specimens, the percentage interobserver agreement for HER2 scores determined by IHC was 80% for the 3 × 3 categories and 92% for the 2 × 2 categories. In terms of the κ statistic, the interobserver agreement level was substantial for the 3 × 3 categories (κ = 0.77, σ = 0.060) and almost perfect for the 2 × 2 categories (κ = 0.88, σ = 0.051). The 20 tumors for which judgment discordance arose for 3 × 3 categories are shown in Table 2. All of these tumors showed discordance in one score, and none showed a difference in 2 scores.

The consensus score for HER2 expression determined by IHC was 0 or 1+, 2+, and 3+ in 60, 19, and 21 tumors, respectively, in CNB specimens, and in 60, 15, and 25 tumors, respectively, in surgically resected specimens.

FISH was successful in 99 of the 100 CNB specimens. By the first examinations, 96 tumors (97%) were judged to have the same score by two observers. Because three other tumors were judged differently by two observers and the average of the two judgments was within the range of equivocal, they were subjected to a re-count: for two tumors, the second judgments also differed between two observers, being positive and equivocal respectively, but the average of the judgments of *HER2/CEP17* ratio exceeded 2.20, so they were finally judged positive. For the other, the second judgments were commonly positive (Table 3). By FISH, *HER2* gene amplification was finally judged positive in 33 CNB specimens (33%) but negative in 66 (67%).

A total of 27 surgically resected tumors were subjected to FISH. Among the 25 surgical specimens for which FISH was successful, only one (no. 67) was judged differently by two observers, being positive and equivocal respectively. However, the average of the two judgments was within the range of positive (*HER2/CEP17* ratio = 2.236), the case was judged as positive without re-evaluation (Table 3). In total, *HER2* gene amplification in surgically resected specimens was positive in 12 (48%) and negative in 13 (52%).

FISH was not successful in three specimens: one CNB and two surgically resected specimens. All three of these specimens were processed in institution A.

Table 2 Surgically resected tumor specimens for which interobserver disagreement arose regarding the results of HER2 immunohistochemistry

Code	Final score	IHC score			FISH (*HER2/CEP17*)
		Observer A	Observer B	Observer C	
S54	0+	0	2+	0	- (0.92)
S17	1+	1+	1+	2+	NA
S26	1+	1+	1+	2+	- (0.68)
S78	1+	1+	1+	2+	- (0.82)
S79	1+	2+	1+	1+	- (0.89)
S76	1+	2+	1+	1+	- (0.95)
S82	1+	2+	1+	1+	- (0.97)
S77	1+	2+	1+	1+	- (1.47)
S23	2+	2+	1+	2+	- (0.47)
S22	2+	2+	1+	2+	- (0.92)
S16	2+	3+	2+	2+	- (1.03)
S84	2+	2+	1+	2+	- (1.13)
S67	2+	2+	1+	2+	+ (2.24)
S97	2+	3+	2+	2+	+ (3.89)
S103	2+	3+	2+	2+	+ (7.75)
S91	2+	3+	2+	2+	+ (7.77)
S94	3+	3+	3+	2+	+ (3.61)
S52	3+	3+	3+	2+	+ (5.09)
S42	3+	3+	3+	2+	+ (9.24)
S18	3+	3+	3+	2+	+ (12.40)

The HER2 score was assigned according to the 2007 ASCO/CAP guideline [10]. The majority score was assigned as the final score for the tumor. NA, not available because FISH was not successful.

FISH, fluorescence in situ hybridization; IHC, immunohistochemistry

Comparison of IHC and FISH test results between CNB and surgically resected specimens

Concordance in consensus HER2 IHC scores between CNB and the corresponding surgically resected specimens was 87% for 3 × 3 categories, and 94% for 2 × 2 categories (Table 4). The κ statistic indicated that their concordance was substantial for the 3 × 3 categories (κ = 0.77, σ = 0.045) and almost perfect for the 2 × 2 categories (κ = 0.83, σ = 0.038). Representative concordant cases are presented in Figure 1.

Table 3 Tumor specimens for which interobserver disagreement arose regarding the results of fluorescence in situ hybridization

CNB specimens							
Code	Final judgment	*HER2/CEP17* by 1st counts			*HER2/CEP17* by 2nd counts		
		Obs. A	Obs. B	Average	Obs. A	Obs. B	Average
B61	Amplification	2.71	1.55	2.13	2.55	1.88	2.22
B62	Amplification	2.35	1.88	2.12	2.44	2.07	2.26
B87	Amplification	1.94	2.44	2.19	3.50	2.38	2.44
Surgically resected specimens							
Code	Final judgment	*HER2/CEP17* by 1st counts			*HER2/CEP17* by 2nd counts		
		Obs. A	Obs. B	Average	Obs. A	Obs. B	Average
S67	Amplification	2.67	1.81	2.24	ND	ND	-

HER2 amplification was defined as positive, equivocal, and negative when the *HER2/CEP17* ratio was more than 2.2, between 1.8 and 2.2, and less than 1.8, respectively. For the surgically resected specimen, the average of the *HER2/CEP17* ratio (2.24) between observers (2.67 and 1.82) was adopted, because the second count was not done (ND). Obs., observer

Table 4 Concordance of consensus HER2 judgments by immunohistochemistry between core needle biopsy and corresponding surgically resected specimens

3 categories (0 or 1+ vs. 2+ vs. 3+)			
HER2 score for surgically resected specimens	Number of tumors		
	HER2 score for CNB specimens		
	0 or 1+	2+	3+
0 or 1+	56	3	1
2+	4	11	0
3+	0	5	20
% agreement = 87%, κ = 0.77, standard deviation (σ) = 0.045			

2 categories (0, 1+. or 2+ vs. 3+)		
HER2 score for surgically resected specimens	Number of tumors	
	HER2 score for CNB specimens	
	0, 1+ or 2+	3+
0, 1+, or 2+	74	1
3+	5	20
% agreement = 94%, κ = 0.83, σ = 0.038		

The judgments were performed according to the 2007 ASCO/CAP guideline.

Among the 60 tumors with a score of 0 or 1+ for CNB specimens, 56 (93%) were also judged to have a score of 0 or 1+ but the other four were scored 2+ for the surgically resected specimens. Among the 21 tumors with a score of 3+ for CNB specimens, 20 (95%) were also judged to have a score of 3+ and only one (5%) was scored 0 for the surgically resected specimens. In contrast, among the 19 tumors with a score of 2+ for CNB specimens, only 11 (58%) were also scored 2+ for the surgically resected specimens. In three and five of these 19 cases, scores of 0 or 1+ and 3+ were assigned, respectively, for the surgically resected specimens.

Overall, consensus HER2 IHC scores showed discordance between CNB and surgically resected specimens in 13 of 100 cases (Table 5). The discordance may have been attributable to intratumor heterogeneity, suboptimal processing of the specimens, and/or the borderline nature of the tumor. The borderline nature means that the state of HER2 expression was borderline between 1+ and 2+ or between 2+ and 3+, namely, the conditions that it was difficult to judge whether the entirely circumscribing membrane immunoreactivity of the HER2 was moderate (2+) or strong (3+), or whether the weak membrane HER2 immunoreactivity was entirely (2+) or incompletely (1+) circumscribing the membrane.

In at least five tumors (Nos. 23, 24, 54, 84, 92), the discordance appeared to have been due to intratumor heterogeneity. In case 54, to the surgically resected specimen, all three observers gave a score of 3+ because >30% of cancer cells showed strong membrane staining. In contrast, in the CNB specimen, the percentage of cancer cells showing strong membrane staining was around 10%, and two observers gave a score of 0, and the other gave one of 2+ (Figure 2A). In cases 23, 24, and 84, there appeared to be intratumor heterogeneity ranging between areas of moderate, and no or weak, HER2 staining (2+ vs 0 or 1+) (Figure 2B). In case 92, the HER2 score for the predominant intraductal component was uniformly 3+, but in the focus of invasive carcinoma, the immunoreaction was weaker (Figure 2C). In this case, we performed FISH on the part including both invasive carcinoma and non-invasive components.

In 4 cases (Nos. 52, 90, 94 and 102), the HER2 score was 2+ for the CNB specimen but 3+ for the surgical specimen (Figure 2D). In these cases, the membrane HER2 immunoreaction in CNB specimens was weaker and less continuous than that in surgically resected specimens. In 5 other cases (Nos. 10, 43, 52, 67 and 69), the intensity of the HER2 immunoreaction differed between the CNB and surgically resected specimens, but the immunoreaction pattern in the surgically resected specimens was uniform. Therefore, suboptimal processing of CNB specimens or long-term fixation of surgically resected specimens might have been partly responsible for the difference.

Interobserver disagreement regarding the HER2 score arose for either CNB specimens or surgically resected specimens in 11 of these 13 cases. Possible reasons for this may have been the borderline nature of the tumors, preanalytical factors, and/or inratumor heterogeneity (e.g., No. 54) (Table 5).

In 23 (92%) of the 25 tumors for which HER2 FISH was performed, consensus judgments regarding *HER2* gene status were concordant between the CNB and the surgically resected specimens, being amplified in 11 and not amplified in 12 (Table 6). In the other two cases, the consensus judgments were discordant between the CNB and the surgically resected specimens: In one (no. 67), the CNB specimen showed an IHC score of 1+ with a *HER2/CEP17* ratio of 0.84, whereas the surgically resected specimen showed an IHC score of 2+ with a *HER2/CEP17* ratio of 2.24. As mentioned above, the case was judged differently by two observers, but the average of the judgments made a score of positive (Table 3). In the other, both the CNB and surgically resected specimens showed an IHC score of 2+, and the *HER2/CEP17* ratio was 1.03 and 2.56, respectively (No. 16). For that case, there was no disagreement in HER2 score of CNB and surgically resected specimens by IHC.

Correlation between IHC and FISH results

Among the 100 CNB specimens, the percentage of *HER2* gene amplification was 95% (20 of 21), 58% (11 of 19), and 3% (2 of 60) for HER2 IHC scores of 3+, 2+, and 0 or 1+, respectively (Table 7). With regard to

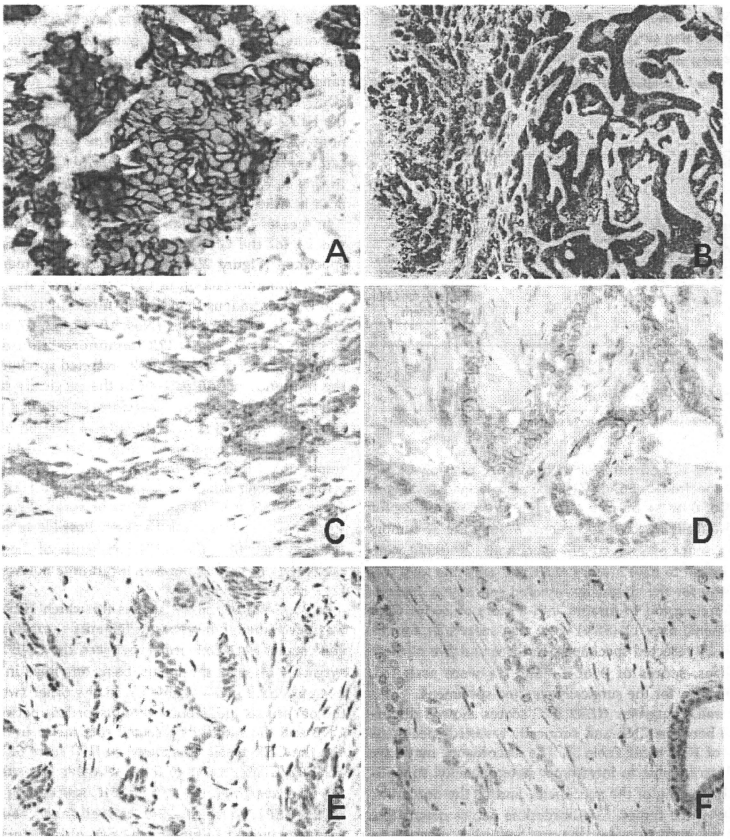

Figure 1 Cases with concordant judgment of HER2 score between CNB and surgically resected specimens. A-B. Case 3: HER2 score for both the CNB specimen (A) and the surgically resected specimen (B) was 3+. C-D. Case 22: HER2 score for both the CNB specimen (C) and the surgically resected specimen (D) was 2+. E-F. Case 2: HER2 score for both the CNB specimen (E) and the surgically resected specimen (F) was 0. Immunoperoxidase reaction, original magnification ×200.

tumors with an IHC score of 2+, the ratio of *HER2* gene amplification differed between tumor specimens from different institutions. In tumor specimens from institutions A and B, the rates of *HER2* gene amplification were 20% (1 of 5) and 33% (1 of 3), respectively, but in specimens from institute C, the rate was as high as 82% (9 of 11).

For the 25 surgically resected specimens successfully subjected to FISH, the percentage of *HER2* gene amplification was 100% (10 of 10), 43% (3 of 7), and 0% (0 of 8) for HER2 IHC scores of 3+, 2+, and 0 or 1+, respectively (Table 7). In tumor specimens from institution A, the rate of *HER2* gene amplification in IHC 2+ cases was 0% (0 of 3), whereas in specimens from institution

Table 5 13 tumors for which interobserver disagreement arose regarding the results of immunohistochemistry.

Code No.	Immunohistochemistry			FISH	
–	CNB	Surgery	Interpretation	CNB	Surgery
84	0 (0/0/0)	2+(2/1/2)	Processing, hetero	Neg	Neg
23	1+(1/1/1)	2+(2/1/2)	Heterogeneity	Neg	Neg
24	1+(2/1/1)	2+(2/2/2)	Heterogeneity	Neg	NA
67	1+(1/0/1)	2+(2/1/2)	Processing, border	Neg	Pos
10	2+(2/2/2)	1+(1/1/1)	Processing	Neg	Neg
43	2+(2/1/2)	0 (0/0/0)	Processing	Neg	Neg
69	2+(2/1/2)	1+(1/1/1)	Processing, border	Neg	Neg
52	2+(3/2/2)	3+(3/3/2)	Processing, border	Pos	Pos
90	2+(3/2/2)	3+(3/3/3)	Processing, border	Pos	Pos
92	2+(2/2/2)	3+(3/3/3)	Predominantly DCIS	Pos	Pos
94	2+(2/2/2)	3+(3/3/2)	Processing, border	Pos	Pos
102	2+(3/2/2)	3+(3/3/3)	Processing, border	Pos	Pos
54	3+(3/3/3)	0 (0/2/0)	Heterogeneity	Neg	Neg

FISH, fluorescence in situ hybridization; NA, Data were not available; neg, negative; pos, positive; parenthesis, judgments of scores by three observers (Underlines indicate interobserver disagreement.)

C, the rate was 80% (4 of 5). No cases with a score of 2+ were subjected to FISH at institution B.

For CNB specimens, 14 cases showed interobserver disagreement in HER2 IHC scores of between 1+ and 2+, and, of these, seven tumors each were finally scored as 1 + and 2+. Among these tumors, *HER2* gene amplification was detected in 0 (0%) and 2 (29%), respectively. The other 10 cases showed interobserver disagreement in HER2 IHC scores of between 2+ and 3+, and, of these, five tumors each were finally scored as 2+ and 3+. *HER2* gene amplification was detected in all of these tumors (Table 1).

For surgically resected specimens, 12 cases showed interobserver disagreement between IHC scores of 0/1+ and 2+, and eight and four tumors were finally scored as 0/1+ and 2+, respectively. Among these tumors, *HER2* gene amplification was detected in 0 (0%) and 1 (25%), respectively. The other eight showed interobserver discordance between IHC scores of 2+ and 3+, and four tumors each were finally scored as 2+ and 3+. Among these tumors, *HER2* gene amplification was detected in 3 (75%) and 4 (100%), respectively (Table 2).

In the 12 tumors showing discordance of the IHC score for HER2 expression between the CNB and surgically resected specimens, and for which the FISH assay was successful, judgments for the FISH results were concordant in 11 (92%) (Table 5).

Discussion

From the viewpoint of interobserver agreement level, the percentage agreement levels and κ values among the three observers with regard to the IHC test results were similar between the data for CNB specimens and those for surgically resected specimens. Furthermore, the present study also clarified that the results of IHC for HER2 were mostly concordant between the CNB specimen and the surgically resected specimen from an identical invasive breast cancer. These results indicated that the CNB specimens were of adequate quality for the evaluation of HER2 status by IHC, and that the IHC scores obtained for the CNB specimens were mostly representative of the HER2 IHC scores for the entire tumor specimen. Therefore, HER2 testing using IHC for CNB specimens appeared to be valid for a majority of primary breast cancers. The present results were similar to the very first studies done by Chivukula et al. that stated as CNB a better sample [25].

Previous reports have indicated that concordance of HER2 IHC scores between CNB and corresponding excisional biopsy/surgically resected specimens was 87-98.8% [21-23,28]. The introduction of FISH analysis has improved the concordance rate for HER2 status [22,25]. In the study by Apple et al., the concordance rate of 87% for IHC was improved to 92% by FISH [22]. Intratumor heterogeneity for HER2 amplification was reported to be present in 13% of tumors of an IHC score of 3+, and was especially higher in those with low-grade amplification (ratio >2.2 to <4.0) [24]. For tumors with an IHC score of 2+, the incidence of intratumor heterogeneity in HER2 scores was also relatively frequent (14%), but FISH analysis of CNB specimens almost completely resolved the issue of heterogeneous HER2 expression [25].

Some studies had suggested that the validity of IHC score 3+ in core biopsies was limited, reporting high rates of false positives (19.3%). However, Moelans et al, showed that there was only a slightly higher percentage of IHC 3+ positivity in biopsies compared to resections, and that did not reach statistical difference [29]. Their results are in line with those in the present study (25% vs 21%). In the present study, it was confirmed that concordance of the HER2 test results was higher for FISH (98%) than for IHC (87%). Among the 13 tumors that showed discordance of HER2 IHC scores between the CNB and surgically resected specimens, 11 showed concordance of the results obtained by FISH.

The disagreement in the results obtained with IHC between CNB and surgical specimens appeared to be derived from 1. intratumor heterogeneity, 2. pre-analytical factors including variations in the duration of fixation and suboptimal tissue processing, and/or 3. borderline tumor properties. To overcome the problem of intratumor heterogeneity in HER2 expression, examination of a large volume of tumor tissue appears to be necessary. To solve the problem of borderline tumor properties in terms of HER2 expression, the introduction of judgment by multiple observers and/or DNA copy analyses might be of value.

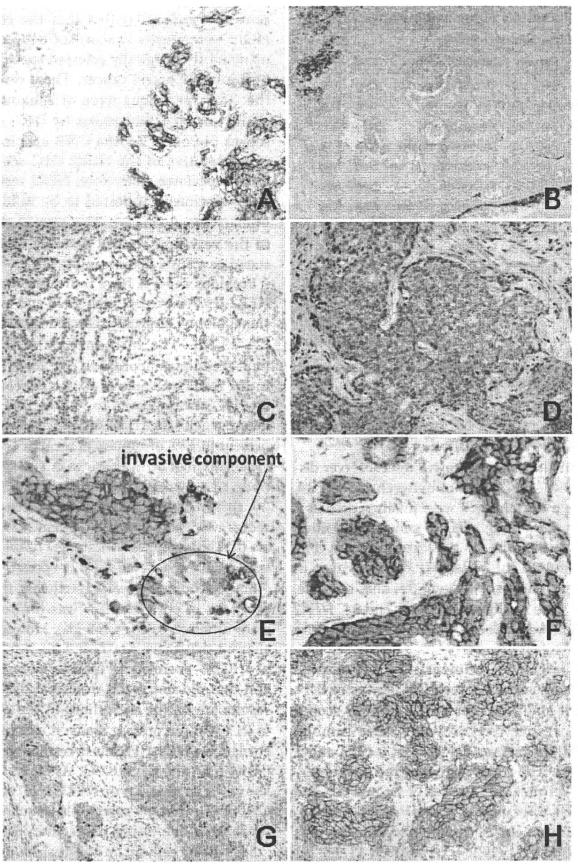

Figure 2 Cases with discordant judgment of HER2 score between CNB and surgically resected specimens. A-B. Case 54: HER2 score was 3+ for the CNB specimen (A) but 0 for the surgically resected specimen (B). The tumor had heterogeneous HER2 expression, and >30% of the area showed a strong membrane immunoreaction in the CNB specimen, whereas most of the area in the surgically resected specimen was HER2-negative. C-D. Case 84: HER2 score was 0 for the CNB specimen (C) but 2+ for the surgically resected specimen (D). This difference might have been due to suboptimal processing of CNB specimens or intratumor heterogeneity. E-F. Case 92: HER2 score was 2+ for the CNB specimen (E) but 3+ for the surgically resected specimen (F). Because the CNB specimen contained only a small amount of invasive component, the evaluation of HER2 was difficult. G-H. Case 94: HER2 score was 2+ for the CNB specimen (G) but 3+ for the surgically resected specimen (H). Processing of the CNB specimen might have been suboptimal, or the tumor may have been borderline in nature. Immunoperoxidase reaction, original magnification ×200.

Table 6 Concordance of consensus HER2 FISH results of between core needle biopsy and surgically resected specimens

HER2/CEP17 ratio for surgically resected specimens	Number of tumors		
	HER2/CEP17 ratio for CNB specimens		
	> 2.2	1.8 to 2.2	<1.8
Higher than 2.2	11	0	1
1.8 to 2.2	0	0	0
<1.8	1	0	12
% agreement = 92%			

The judgments were performed according to the 2007 ASCO/CAP guideline. For two tumors, data on CNB and/or surgically resected specimens were not available.

From the present results, the rate of *HER2* gene amplification appeared to be low when interobserver discordance was seen between IHC scores of 0/1+ and 2 +, whereas most tumors had gene amplification when there was interobserver discordance between scores of 2 + and 3+. These findings will be helpful for the decision if subsequent FISH should be performed or not. For 2 +/3+ discrepancies FISH test should always be added because the percentage of HER2 amplification is high. Counting of *HER2* gene copies by FISH, chromogenic in situ hybridization (CISH), silver-enhanced in situ hybridization (SISH), or dual-color dual-hapten in situ hybridization (DDISH) would greatly improve the concordance of HER2 status between CNB and surgically resected specimens [30-32].

Table 7 Correlation between consensus immunohistochemistry and FISH results for core needle biopsy and surgically resected specimens

Core needle biopsy specimens				
	Total	**Number of tumors (%)**		
		***HER2* gene amplification by FISH**		
		Positive	**Equivocal***	**Negative**
HER2 IHC score				
3+	21	20 (95)	0 (0)	1 (5)
2+	19	11 (58)	0 (0)	8 (42)
0 or 1+	60	2 (3)	0 (0)	58 (97)
Total	100	33 (33)	0 (0)	67 (67)
Surgically resected specimens				
	Total	Number of tumors (%)		
		HER2 gene amplification by FISH		
		Positive	Equivocal*	Negative
HER2 IHC score				
3+	10	10(100)	0 (0)	0 (0)
2+	7	3 (43)	0 (0)	4 (57)
0 or 1+	8	0 (0)	0 (0)	8 (100)
Total	25	13 (52)	0 (0)	12 (48)

*For three tumors, FISH judgment was equivocal by the first counts, but by the second counts, the judgments changed into positive (Table 3). FISH, fluorescence in situ hybridization

Another important factor to be considered is the preanalytical condition of the specimens. In the present study, FISH analysis was sometimes unsuccessful for specimens processed at institution A, whereas FISH was always successful for specimens processed at institutions B and C. Interview revealed that the suboptimal FISH results obtained at institution A were attributable to long-term (about 1 week) fixation of the specimens. It was shown that a prolonged formalin fixation could lose FISH amplification and/or yield to unsuccessful test (Hiroi S, Tsuda H et al, manuscript in preparation) [33].

On the other hand, the localization of the HER2 immunoreaction on the cancer cell membrane was not uniform in CNB specimens from HER2-positive tumors processed at institution C, whereas this was not an evident feature in specimens processed at other institutions or in surgically resected specimens. The reason for this unusual immunoreaction is unclear. It has been shown that quality assessment of HER2 tests is very important for identifying patients who are very likely to benefit from therapy with trastuzumab [10,34,35]. For quality assessment, not only improvement of the interobserver agreement level but also standardization of pre-analytical specimen preparation should be taken into consideration.

Conclusion

We have clarified that CNB specimens showed almost equal reliability to surgically resected specimens for testing of HER2 expression in terms of interobserver agreement levels and concordance with FISH results. In most of specimens with equivocal IHC results, accurate HER2 status was known determined by retesting with FISH. To further improve the reliability of HER2 tests using CNB specimens, it might be useful to sample a larger volume of tumor tissue, to conduct evaluation by multiple observers, and to take measures to improve the preanalytical conditions of the specimens.

List of abbreviations used

CNB: core needle biopsy; FISH: fluorescence in situ hybridization; HER2: human epidermal growth factor receptor type-2; IHC: immunohistochemistry