Table 2 The methylation profiles of 8 imprinted genes by bisulphite PCR Luminex.

Cell lines	Histology	H19		GTL2		ZDBF2		PEG1		LIT1		ZAC		PEG3		SNRPN	
		BPL	COBRA	BPL	COBRA	BPL	COBRA	BPL	COBRA	BPL	COBRA	BPL	COBRA	BPL	COBRA	BPL	COBRA
		CpG (9,16)	CpG10	CpG (4,8)	CpG8	CpG (1,4,5)	CpG4	CpG15	CpG12	CpG (5,17,19)	CpG16	CpG8	CpG7	CpG20	CpG21	CpG19	CpG19
NC1	N	67.5	34.6	62.2	76.8	66.4	46.2	50.4	38.5	**42.7**	**27.4**	62.3	48.4	78.8	67.4	65.3	57.3
NC2	N	55.6	42.7	**67.3**	**82.1**	67.9	72.5	54.8	46.2	53.7	37.5	65.2	69.3	65.4	48.9	78.5	67.3
NC3	N	**65.8**	**58.3**	81.2	63.5	76.3	82.1	**45.3**	**63.9**	26.4	41.6	50.2	72.1	65.8	83.2	56.3	68.2
NC4	N	58.6	72.1	76.4	80.4	80.2	67.2	37.5	58.2	**35.8**	**54.7**	47.2	52.8	45.9	58.3	65.2	74.9
OC1	S	**10.32**	**95.4**	69.8	53.2	87.3	92.4	**83.1**	**59.2**	64.3	48.3	89.5	98.3	74.4	83.2	65.4	58.4
OC2	M	16.8	19.4	61.4	70.5	95.7	97.2	48.8	62.2	62.1	48.0	20.1	10.2	**92.1**	**89.5**	29.1	49.7
OC3	S	87.4	93.2	96.3	93.4	83.7	28.6	**98.3**	**90.4**	87.4	56.9	87.3	96.5	98.5	96.4	28.4	37.4
OC4	S	42.8	56.8	97.3	93.2	99.3	98.5	0.9	10.4	13.2	17.1	100.9	98.3	**48.7**	**47.2**	28.1	24.2
OC5	S	83.4	56.4	89.2	96.6	95.3	72.8	93.2	102.4	28.4	32.6	100.3	91.3	98.8	95.6	54.3	60.4
OC6	E	78.9	72.1	94.3	82.1	38.3	18.3	83.2	68.5	**34.2**	**56.2**	85.3	78.9	57.2	97.5	40.5	53.1
OC7	C	0.4	21.3	19.8	29.3	41.7	55.5	60.5	42.0	21.2	11.3	34.5	40.1	90.7	99.0	23.5	52.8
OC8	C	0.1	6.9	6.4	10.5	91.7	87.3	**81.8**	**88.1**	35.2	29.1	49.8	53.2	102.1	95.7	21.8	33.2
OC9	S	82.3	54.3	**37.4**	**65.2**	22.4	42.8	**23.6**	**18.5**	43.7	56.8	69.5	74.9	**68.5**	**90.2**	38.2	56.3
OC10	C	83.8	67.8	**78.4**	**65.9**	27.9	32.3	63.8	58.7	42.8	60.2	89.9	93.7	59.6	43.9	38.9	71.4
OC11	S	92.4	91.6	17.1	25.0	6.0	18.0	43.2	60.1	18.2	10.0	72.5	58.2	101.4	100.0	75.5	60.5
OC12	S	54.7	53.8	23.7	18.8	48.6	**74.3**	**68.9**	58.4	48.3	93.6	98.8	**67.4**	**39.8**	**73.4**	57.1	51.2
OC13	S	**80.3**	**66.2**	**42.3**	**40.1**	86.0	74.2	**93.4**	95.2	10.5	8.8	94.4	90.5	**102.1**	**100.0**	36.0	39.5
OC14	C	98.1	87.7	67.4	40.3	99.3	90.2	99.6	95.5	12.3	1.6	34.3	41.2	92.1	100.0	38.1	51.2
OC15	C	45.2	45.0	93.0	98.2	99.7	95.3	**92.1**	**94.8**	55.0	47.5	27.5	36.4	86.1	97.2	48.7	10.1
OC16	S	36.9	44.6	90.6	92.0	93.0	92.3	100.7	100.0	13.0	6.2	28.9	35.7	93.2	100.0	22.9	17.6
OC17	C	93.8	89.8	61.8	68.3	21.8	22.2	**37.9**	**42.5**	**48.4**	**31.8**	20.6	31.1	98.3	95.5	0.0	0.2
OC18	C	**34.4**	**46.0**	94.5	98.0	95.3	98.0	60.1	73.2	35.2	29.4	17.3	35.0	87.2	100.0	22.7	11.1
OC19	C	67	30.7	**47.8**	**62.1**	87.4	85.3	4.1	10.1	**11.2**	**0.0**	100.1	100.0	84.6	92.1	30.4	31.4
OC20	M	1.6	42.6	**58.1**	**60.3**	62.4	55.2	**73.4**	**81.4**	18.3	13.4	49.2	68.2	92.8	98.5	86.7	44.8
OC21	S	**8.0**	**24.5**	**80.2**	**87.0**	96.6	90.1	77.2	81.2	4.5	8.1	100.5	97.5	107.8	98.2	16.7	16.2
Normal ovarian surface tissues																	
NT1 (42)		57.5	76.2	43.4	51.2	37.2	65.4	51.4	38.6	**49.1**	**42.3**	53.3	41.2	42.1	40.2	521	47.8
NT2 (48)		**27.6**	**25.0**	23.6	21.6	46.1	34.5	20.9	34.4	23.1	12.2	5.1	32.3	21.3	20.6	22.1	33.3
NT3 (40)		61.5	31.8	**49.9**	**60.0**	50.4	45.0	31.0	42.8	17.1	10.9	48.8	29.4	21.0	33.6	24.3	34.8
NT4 (38)		28.9	25.0	**28.4**	**29.6**	46.1	39.3	**33.4**	**47.8**	28.8	17.7	67.4	46.1	45.6	33.8	80.5	29.0
NT5 (55)		**24.2**	**26.9**	**27.2**	**18.5**	33.2	29.0	**30.9**	**48.1**	**25.1**	**17.6**	10.7	44.5	75.5	40.0	32.3	31.3
NT6 (45)		**25.9**	**22.1**	**32.2**	**20.5**	46.4	62.5	60.0	40.4	24.1	17.3	25.2	23.4	44.2	69.0	77.4	40.9
NT7 (41)		34.6	24.6	**38.0**	**19.9**	49.9	37.4	27.3	46.2	**27.7**	**24.5**	21.0	63.1	49.7	53.7	25.6	24.5

Letters in blue and black boldface represent LOH, LOI and MOI, respectively. The Luminex value indicates the average value of DNA methylation

Table 3 Characterization of methylation profiles of the imprinted genes in DNA of ovarian cancers.

		H19	GTL2	ZDBF2	PEG1	LIT1	ZAC	PEG3	SNRPN
(A) Histology									
Normal	(n = 7)	34.3 ± 11.8	34.7 ± 9.5	44.2 ± 6.6	30.7 ± 15.1	27.9 ± 12.6	33.1 ± 5.8	42.8 ± 18.5	44.9 ± 25.3
Cancer	(n = 74)	41.7 ± 17.2	39.6 ± 18.8	43.8 ± 6.6	45.9 ± 15.5*	27.9 ± 14.1	41.1 ± 6.4	41.1 ± 14.1	42.8 ± 12.9
Serous	(n = 36)	47.6 ± 18.9	38.8 ± 22.0	42.0 ± 18.5	48.9 ± 14.7*	28.8 ± 9.8	42.7 ± 16.6	40.9 ± 14.8	39.3 ± 12.2
Mucinous	(n = 9)	36.0 ± 15.0	35.2 ± 19.2	36.8 ± 14.0	47.1 ± 20.0	34.7 ± 21.4	42.2 ± 5.9	31.2 ± 11.2	42.7 ± 17.8
Endometrioid	(n = 10)	37.9 ± 14.8	30.7 ± 14.4**	57.8 ± 19.8	45.5 ± 13.7	22.8 ± 13.0	37.5 ± 11.5	46.8 ± 15.9	41.4 ± 10.8
Clear	(n = 18)	45.6 ± 20.2	54.0 ± 20.1	39.7 ± 20.8	42.2 ± 13.8	25.5 ± 12.4	40.2 ± 16.1	45.7 ± 14.5	47.8 ± 11.1
(B) Progress (Staging)									
Localized (I, II)	(n = 29)	38.8 ± 14.9	40.8 ± 20.4	46.7 ± 19.4	47.8 ± 17.1	30.4 ± 16.1	41.0 ± 11.9	42.8 ± 17.8	44.2 ± 12.9
Advanced (III, IV)	(n = 45)	47.8 ± 19.6	41.3 ± 22.2	41.9 ± 19.2	45.9 ± 13.7	26.2 ± 10.0	41.1 ± 16.8	41.2 ± 12.8	40.9 ± 12.6
(C) Age									
Under 44 years	(n = 17)	46.7 ± 18.5	35.3 ± 21.0	42.0 ± 20.4	48.9 ± 13.5	24.2 ± 11.7	39.9 ± 12.3	44.7 ± 14.3	36.3 ± 10.7
45-55 years	(n = 29)	39.9 ± 14.7	47.2 ± 23.0	45.8 ± 17.5	46.9 ± 15.0	30.1 ± 14.8	38.9 ± 15.5	41.9 ± 15.8	45.4 ± 11.1
Over 56 years	(n = 32)	47.4 ± 21.2	38.4 ± 18.1	42.2 ± 21.1	45.1 ± 16.2	28.2 ± 10.9	44.1 ± 15.7	40.8 ± 14.9	42.5 ± 14.5

The values in the list are mean ± SD (standard deviation). Statistically significant differences between groups are presented as *P < 0.05, and **P < 0.01 by ANOVA

association between DNA methylation and LOH and/or LOI in the *IGF2/H19* and *PEG1* imprinted domains separately.

IGF2, which acts as a dominant oncogene, and *H19*, a physically and mechanistically linked gene on human chromosome 11, are reciprocally imprinted. In the paternal allele, *H19* DMR is methylated and silenced, whereas the reciprocally imprinted gene *IGF2* is transcribed. By contrast, in the maternal unmethylated allele, *H19* is expressed but *IGF2* is inactivated because of the binding of the repressor factor CTCF to the unmethylated *H19* DMR, which then prevents the *H19/IGF2* common enhancers from activating the *IGF2* promoter [19]. *IGF2* was found to have high frequencies of both LOH and LOI in HOC. *H19* was also found to have high frequencies of LOI (29.2%, 12/41). Nine of 14 cases with both *IGF2* and *H19* heterozygosity showed LOH or LOI of both genes and only one case had MOI for one of the two. Thus relaxation of *IGF2* and *H19* imprinting is frequent. In the *IGF2/H19* imprinted region, the samples with LOH and/or LOI at *H19* was more methylated than those with MOI (Figure 2A, Additional file 3: Table S2).

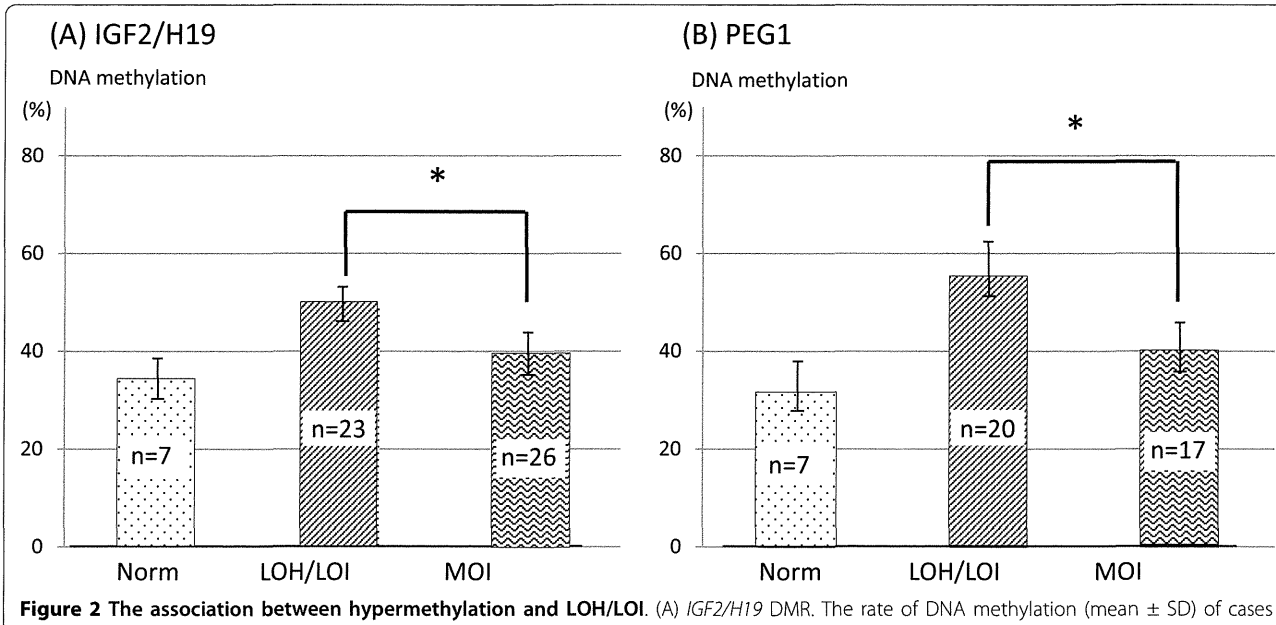

Figure 2 The association between hypermethylation and LOH/LOI. (A) *IGF2/H19* DMR. The rate of DNA methylation (mean ± SD) of cases with LOH/LOI of either *IGF2* or *H19* and MOI were compared. For this analysis, one case which *H19* was MOI and *IGF2* was LOI, was removed. (B) *PEG1* DMR. Statistically significant differences between groups are analyzed with ANOVA followed with multiple comparison method (Tukey's HSD test), presented as the *P*-value (*P < 0.01). Norm: Normal ovarian tissues.

Our results for *H19* were similar to a previously reported finding [20]. *PEG1* was reported to be a TSG and was also found to have high frequency of LOH/LOI in HOC. We also found that the samples with LOH/LOI at *PEG1* were more methylated than those with MOI with statistical significance (Figure 2B, Additional file 3: Table S2).

Discussion

Alterations in DNA methylation are the most common molecular alterations in human malignancies. Detection of the aberrant DNA methylation associated with cancer-related genes is a promising approach to improve cancer prevention, diagnosis and treatment options. Bisulphite modification is a prerequisite for most popular techniques aiming at detecting changes in methylation, but has been limited by throughput capacity. In this study, we used a high-throughput methylation detection method to analyze DNA methylation at 8 imprinted DMRs in epithelial ovarian cancer. We found that the PCR-Luminex method precisely quantified the methylation status of specific DNA regions in somatic cells and was also relatively rapid, economical and easy to use.

In the epithelial ovarian cancers, the frequency of LOI was higher than that of LOH. In particular, LOI was most frequent at *PEG1*, *IGF2* and *H19* DMRs. The frequency, extent of changes in DNA methylation and loci affected varied considerably among the samples. Generally, we found that DNA methylation at imprinted DMRs was increased in both cell lines and primary material. Importantly, we showed that gain of DNA methylation in the imprinted DMRs was apparent in tumors with LOI, especially at *PEG1* and *H19*. We also found DNA methylation changes in the absence of LOI. In other words, there were changes in DNA methylation at DMRs that were not associated with biallelic gene expression.

When we examined the clinical characteristics of the tumors, we found no significant differences in the frequency of LOI and aberrant DNA methylation between the localized early-stage and advanced-stage tumor groups. This suggested that the changes we identified occurred as a relatively early event of HOC. In general, the *PEG1* and *H19* DMRs appeared to be particularly prone to errors. This is similar to the previous findings in human sperm from subfertile men [21]. In *ZDBF2* and *GTL2* DMRs, aberrant DNA methylation occurred in HOC. As with *H19*, these DMRs are paternally methylated DMRs in somatic cells. In childhood cancers such as retinoblastoma, Wilms' tumor and osteosarcoma, changes primarily occur on the paternal allele first, followed by a second hit on the maternal allele [5,6,22]. Similarly, methylation of paternally imprinted DMR in normal somatic cells might be a first hit and cause ovarian carcinogenesis. These observations suggest a role for altered genomic imprinting in the malignant transformation process.

A previous report had demonstrated the association between the abnormal genomic imprinting of *H19* and *IGF2* expression [20]. The aberrant hypermethylation in the CTCF binding site of the *H19* gene was seen in the cases of HOC and correlated with *IGF2* LOI. Our results for *H19* were similar to those reported findings. The most frequent methylation error in HOC was seen in the *PEG1* DMR. In our previous report, we showed that demethylation of *PEG1* was present in growing oocytes from superovulated infertile women [23]. This *PEG1* DMR may be especially vulnerable to errors. LOI of PEG1 has subsequently also been implicated in the aetiology of lung adenocarcinomas, breast and colon cancer.

HOC is the leading cause of death from gynecologic malignancies because the majority of cases are not detected until the disease is well advanced. Our understanding of cancer as a clonal genetic disease has led to the identification of genetic alterations in many cancer types. However, ovarian cancer remains less well characterized. Only a few TSG genes acting in a recessive manner have been identified as somatically mutated or methylated in ovarian cancer, including *TP53* (48%) [24], *PTEN* (21% in the endometrioid subtype) [25], *RB1* (7-10%) [26], and *CDKN2A* (79% in the mucinous subtype) [27]. Biomarkers provide useful tools in screening for cancer and are now emerging as highly informative for monitoring disease status [28]. They can improve early detection and also the quality of life of patients with ovarian cancer. DNA methylation offers an additional tool that can be used in combination with other markers [29]. In addition, it has been established that DNA methylation biomarkers are present in patient serum and other body fluids [30]. To date, several methylated genes have been found to be highly prognostic for specific cancers, including those of the prostate [31], breast [32] and lung [33]. Although some methylated markers such as RASSF1A and GSTP1 have potential as prognostic indicators individually [34,35], 'methylation signature' panels could be much more informative [36] and accurate for monitoring cancer progression. Methylation patterns have previously been suggested to be tumor and stage specific [37]. Our work demonstrates that there is aberrant DNA methylation at several imprinted DMRs in HOC with changes at *PEG1* and *H19* being the most frequent and earliest alterations detected.

Conclusion

This is the first study reporting the use of PCR-Luminex for identification of prognostic panels of DNA methylation biomarkers for cancer. We believe that this approach is amenable to the classification of clinically relevant methylation patterns in a wide variety of tumors (and other pathologies) linked to the aberrant DNA methylation of imprinted DMRs. This BPL method may be sufficiently

sensitive that it can be applied to the analysis of DNA methylation in the very small number of circulating cancer cells found in blood and urine samples from patients.

Methods

Ovarian cancer cell lines and primary culture of surface epithelial cells

Twenty-one HOC cell lines were used in our study: 10 from serous adenocarcinoma (OVCAR3, CAOV3, JHOS2, HTOA, SKOV3, OV90, JHOS3, JHOS4, KF, MH), 2 from mucinous adenocarcinoma (OMC3, MCAS), 8 from clear cell adenocarcinoma (ES2, JHOC5, TOV21G, JHOC7, JHOC8, KM, HAC2, RMG) and 1 from endometrioid adenocarcinoma (TOV112D). The sources of these cells and culture methods were as described previously [38,39]. Four primary cultures of normal human ovarian surface epithelial (OSE1-4) cells were initiated from surface scrapings of normal ovaries as described [40].

Ovarian cancer tissue(s)

Seventy-four primary HOC tissues (36 serous, 9 mucinous, 10 endometrioid, 18 clear cell, and 1 other, Table 3) were obtained from patients presenting at our hospital. The mean ± standard deviation (SD) of the patients' ages for normal ovary and ovarian cancer tissues were 44.1 ± 6.6 and 52.9 ± 7.3, respectively. Seven specimens of normal ovarian surface epithelium were obtained from patients with benign non-ovarian disease. Histological diagnoses and clinical staging were performed according to the International Federation of Gynecologists and Obstetricians (FIGO) criteria. The numbers of cancer patients with localized tumors (stage I and II) and advanced tumor (stages III and IV) were 29 and 45, respectively. The samples were stained with hematoxylin and eosin to demonstrate > 85% of epithelial tumor cells. DNA and RNA were then extracted from the remaining samples [38]. DNA was also extracted from peripheral blood in matched patients. The study was performed after obtaining the patients' informed consent and with approval from the institutional ethics committee of the Tohoku University Graduate School of Medicine.

Analysis of loss of heterozygosity (LOH) and loss of imprinting (LOI)

PCR was performed on patient blood and tumor genomic DNA using the primer sequences summarized in Additional file 4: Table S3. A PCR reaction mix containing 0.5 μM of each primer set, 200 μM dNTPs, 1 × PCR buffer, and 1.25U of EX *Taq* Hot Start DNA Polymerase (Takara Bio, Tokyo, Japan) in a total volume of 20 μl was used. The following PCR program was used: 1 minute of denaturation at 94°C followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 60°C and 30 seconds at 72°C and a final extension for 5 minutes at 72°C. PCR products were digested by unique polymorphic enzymes to identify samples that were heterozygous for a single nucleotide polymorphism (SNP). For samples found to be heterozygous for a SNP, RNA was prepared from matched tumors, followed by reverse transcription-PCR (RT-PCR) and by restriction digestion [41-48]. The digested PCR products were electrophoresed on 2% agarose gel.

DNA methylation analysis

Bisulphite PCR-Luminex (BPL) methylation analysis was performed as described [18]. PCR primers sets, biotinylated at their 5'-end, were designed for gene amplification. PCR reaction mix contained 0.2 μM primer, 0.2 mM dNTPs, 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3). 3 mM $MgCl_2$, 2% dimethyl sulfoxide (DMSO), 0.625U *Taq* DNA Polymerase (Roche, Tokyo, Japan) and 100-200 ng of bisulphite treated DNA in a total volume of 25 μl. PCR conditions: 40 cycles of 95°C for 20 s/60°C for 30 s/72°C for 30 s using a GeneAmp 9700 thermal cycler (Applied Biosystems, CA, USA). Oligonucleotide probe sequences (Additional file 2: Table S1 in Ref 18) were synthesized and covalently bound to carboxylated fluorescent microbeads using ethylene dichloride (EDC). These oligonucleotide-labeled microbeads (oligobeads) were mixed together to make an oligobeads mixture of 100 oligobeads/μl and hybridized to the 5'-biotin-labeled PCR amplicons in a total volume of 50 μl per well in a 96-well plate by adding 5 μl of the appropriate oligobead mixture and 5 μl of the PCR amplicons to 40 μl of hybridization buffer. This reaction mixture was first denatured at 95°C for 2 min and then hybridized at 48°C for 30 min. After hybridization, the oligobeads were washed in 100 μl of PBS-Tween and pelleted by microcentrifugation. Pelleted oligobeads were reacted with a 70 μl aliquot of a 100 × diluted solution of SA-PE in PBS-Tween. Hybridized amplicons were labeled with SA-PE at 48°C for 15 min. Reaction outcomes were measured by the Luminex 100 flow cytometer. Methylation assays were additionally performed for each DMR using the conventional bisulphite treatment PCR methylation assay and combined bisulphite PCR restriction analysis (COBRA) as described previously [21].

Statistical analyses

Differences between groups were analysed by analysis of variance, followed by Post-hoc, Tukey's HSD test. Statistical analyses were performed using the JMP (v9.0.0, SAS Institute Japan, Tokyo, Japan). Statistically significant differences between groups are presented as *$P < 0.05$, and **$P < 0.01$. Results for BPL and COBRA were compared using Spearman's rank method and Pearson's product-moment correlation coefficients.

Additional material

Additional file 1: Table S1 The list of LOH, LOI and MOI in HOC. LOH, LOI and MOI determined using RFLP analysis of 8 imprinted genes are summarized. NC: Normal cells (NC1-4). NT: Normal ovarian tissues (NT1-7). CC: Cancer cell lines (CC1-21). CT: Cancer tissue (CT1-74).

Additional file 2: Figure S1 Validation of BPL analyses by comparison with COBRA assay. Examination of the imprinted DMRs by bisulphite PCR Luminex (BPL) and combined bisulphite PCR restriction analysis (COBRA) assay in DNA samples of ovarian cancer cell lines and normal cells. BPL: y-axis, COBRA: x-axis. The number was calculated by Spearman's rank method. *GTL2* (C), *ZDBF2* (D), *LIT1* (E), *ZAC* (F), *PEG3* (G) and *SNRPN* (H).

Additional file 3: Table S2 Sequences of PCR primers and restriction enzymes used for PCR-RFLP analysis.

Additional file 4: Table S3 Bisulphite PCR-Luminex and COBRA methylation profiles of the eight imprinted DMRs in the DNA of human ovarian cancer cells and normal ovarian tissues. Numbers in blue and black boldface indicate LOI and MOI, respectively. Luminex values indicate average methylation values at the sites tested. NC: Normal cells (OSE1, OSE2, OSE3, OSE4). CC: Cancer cell lines (CC1: OVCAR3, CC2: OMC3, CC3: CAOV3, CC4: JHOS2, CC5: HTOA, CC6: TOV112D, CC7: ES2, CC8: JHOC5, CC9: SKOV3, CC10: TOV21G, CC11: OV90, CC12: JHOS3, CC13: JHOS4, CC14: JHOC7, CC15: JHOC8, CC16: KF, CC17: KM, CC18: HAC, CC19: RMG, CC20: MCAS, CC21: MH. NT: Normal ovarian tissues (NT1-NT7). N: Normal, S: Serous adenocarcinoma, M: Mucinous adenocarcinoma, E: Endometrioid carcinoma, C: Clear cell carcinoma.

Abbreviations

BPL: Bisulphite PCR-Luminex; COBRA: Combined bisulphite PCR restriction analysis; DMR: Differentially methylated region; FIGO: International Federation of Gynecologists and Obstetricians; HOC: Human ovarian cancer; LOH: Loss of heterozygosity; LOI: Loss of imprinting; MOI: Maintenance of imprint; ND: Not determined; PCR: Polymerase chain reaction; QOL: Quality of Life; RFLP: Restriction fragment length polymorphism; RT-PCR: Reverse transcription-PCR; SNP: Single nucleotide polymorphism.

Acknowledgements

We would like to thank Yukiko Abe, Takuya Koshizaka (G&G SCIENCE Co. Ltd., Fukushima Japan) and all the members of our laboratory for technical assistance and their support and valuable suggestions. In particular, we thank Dr. R. M. John for comments on the manuscript. This work was supported by Grant-in-Aid for Scientific Research (KAKENHI) (21028003, 23013003, 23390385), Uehara Memorial Foundation and the Environment Research & Technology Development Fund (C1008) (TA).

Author details

[1]Department of Informative Genetics, Environment and Genome Research Center, Tohoku University Graduate School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan. [2]Department of BioScience, Tokyo University of Agriculture, 1-1-1 Sakuragaoka, Setagaya-ku, Tokyo 156-8502, Japan. [3]Departments of Obstetrics and Gynecology, Tohoku University Graduate, School of Medicine, Sendai, Japan. [4]Department of Development and Environmental Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan.

Authors' contributions

HH, HO, HK, NM, FS and AS performed the DNA methylation analyses and validation of the BML method. FS, SN, JS, NY carried out the product and culture of cancer cell lines and collect the tumor samples. KN did the statistical analyses. HH and TA wrote this manuscript. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 15 December 2011 Accepted: 26 March 2012
Published: 26 March 2012

References

1. Russo A, Calo V, Bruno L, Rizzo S, Bazan V, Di Fede G: **Hereditary ovarian cancer.** *Crit Rev Oncol Hematol* 2009, **69**:28-44.
2. Balch C, Fang F, Matei DE, Huang TH, Nephew KP: **Minireview: epigenetic changes in ovarian cancer.** *Endocrinology* 2009, **150**:4003-4011.
3. Knudson AG: **Genetics of human cancer.** *Annu Rev Genet* 1986, **20**:231-251.
4. Jones PA, Laird PW: **Cancer epigenetics comes of age.** *Nat Genet* 1999, **21**:163-167.
5. Dryja TP, Mukai S, Petersen R, Rapaport JM, Walton D, Yandell DW: **Parental origin of mutations of the retinoblastoma gene.** *Nature* 1989, **339**:556-558.
6. Huff V, Meadows A, Riccardi VM, Strong LC, Saunders GF: **Parental origin of de novo constitutional deletions of chromosomal band 11p13.** *Am J Hum Genet* 1990, **47**:155-160.
7. Kajii T, Ohama K: **Androgenetic origin of hydatidiform mole.** *Nature* 1977, **268**:633-634.
8. Mutter GL: **Teratoma genetics and stem cells: a review.** *Obstet Gynecol Surv* 1987, **42**:661-670.
9. Feinberg AP: **DNA methylation, genomic imprinting and cancer.** *Curr Top Microbiol Immunol* 2000, **249**:87-99.
10. Joyce JA, Schofield PN: **Genomic imprinting and cancer.** *Mol Pathol* 1998, **51**:185-190.
11. Feinberg AP: **Imprinting of a genomic domain of 11p15 and loss of imprinting in cancer: an introduction.** *Cancer Res* 1999, **59**:1743-1746.
12. Zhang L, Volinia S, Bonome T, Calin GA, Greshock J, Yang N, *et al*: **Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer.** *Proc Natl Acad Sci USA* 2008, **105**:7004-7009.
13. Kohda T, Asai A, Kuroiwa Y, Kobayashi S, Aisaka K, Nagashima G, *et al*: **Tumour suppressor activity of human imprinted gene PEG3 in a glioma cell line.** *Genes Cells* 2001, **6**:237-247.
14. Nakano S, Murakami K, Meguro M, Soejima H, Higashimoto K, Urano T, *et al*: **Expression profile of LIT1/KCNQ1OT1 and epigenetic status at the KvDMR1 in colorectal cancers.** *Cancer Sci* 2006, **97**:1147-1154.
15. Kamikihara T, Arima T, Kato K, Matsuda T, Kato H, Douchi T, *et al*: **Epigenetic silencing of the imprinted gene ZAC by DNA methylation is an early event in the progression of human ovarian cancer.** *Int J Cancer* 2005, **115**:690-700.
16. Surani MA: **Imprinting and the initiation of gene silencing in the germ line.** *Cell* 1998, **93**:309-312.
17. Koerner MV, Barlow DP: **Genomic imprinting-an epigenetic gene-regulatory model.** *Curr Opin Genet Dev* 2010, **20**:164-170.
18. Sato A, Hiura H, Okae H, Miyauchi N, Abe Y, Utsunomiya T, *et al*: **Assessing loss of imprint methylation in sperm from subfertile men using novel methylation polymerase chain reaction Luminex analysis.** *Fertil Steril* 2011, **95**:129-134.
19. Reik W, Walter J: **Genomic imprinting: parental influence on the genome.** *Nat Rev Genet* 2001, **2**:21-32.
20. Murphy SK, Huang Z, Wen Y, Spillman MA, Whitaker RS, Simel LR, *et al*: **Frequent IGF2/H19 domain epigenetic alterations and elevated IGF2 expression in epithelial ovarian cancer.** *Molecular cancer research: MCR* 2006, **4**:283-292.
21. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, *et al*: **Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients.** *Hum Mol Genet* 2007, **16**:2542-2551.
22. Toguchida J, Ishizaki K, Sasaki MS, Nakamura Y, Ikenaga M, Kato M, *et al*: **Preferential mutation of paternally derived RB gene as the initial event in sporadic osteosarcoma.** *Nature* 1989, **338**:156-158.
23. Sato A, Otsu E, Negishi H, Utsunomiya T, Arima T: **Aberrant DNA methylation of imprinted loci in superovulated oocytes.** *Hum Reprod* 2007, **22**:26-35.
24. Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P: **The IARC TP53 database: new online mutation analysis and recommendations to users.** *Hum Mutat* 2002, **19**:607-614.
25. Obata K, Morland SJ, Watson RH, Hitchcock A, Chenevix-Trench G, Thomas EJ, *et al*: **Frequent PTEN/MMAC mutations in endometrioid but not serous or mucinous epithelial ovarian tumors.** *Cancer Res* 1998, **58**:2095-2097.

26. Gorringe KL, Jacobs S, Thompson ER, Sridhar A, Qiu W, Choong DY, *et al*: **High-resolution single nucleotide polymorphism array analysis of epithelial ovarian cancer reveals numerous microdeletions and amplifications.** *Clin Cancer Res* 2007, **13**:4731-4739.
27. Milde-Langosch K, Ocon E, Becker G, Loning T: **p16/MTS1 inactivation in ovarian carcinomas: high frequency of reduced protein expression associated with hyper-methylation or mutation in endometrioid and mucinous tumors.** *Int J Cancer* 1998, **79**:61-65.
28. Ludwig JA, Weinstein JN: **Biomarkers in cancer staging, prognosis and treatment selection.** *Nat Rev Cancer* 2005, **5**:845-856.
29. Ushijima T: **Detection and interpretation of altered methylation patterns in cancer cells.** *Nat Rev Cancer* 2005, **5**:223-231.
30. Cottrell SE, Laird PW: **Sensitive detection of DNA methylation.** *Ann N Y Acad Sci* 2003, **983**:120-130.
31. Li LC, Carroll PR, Dahiya R: **Epigenetic changes in prostate cancer: implication for diagnosis and treatment.** *J Natl Cancer Inst* 2005, **97**:103-115.
32. Bae YK, Brown A, Garrett E, Bornman D, Fackler MJ, Sukumar S, *et al*: **Hypermethylation in histologically distinct classes of breast cancer.** *Clin Cancer Res* 2004, **10**:5998-6005.
33. Zochbauer-Muller S, Minna JD, Gazdar AF: **Aberrant DNA methylation in lung cancer: biological and clinical implications.** *Oncologist* 2002, **7**:451-457.
34. Henrique R, Jeronimo C: **Molecular detection of prostate cancer: a role for GSTP1 hypermethylation.** *Eur Urol* 2004, **46**:660-669.
35. Yang Q, Zage P, Kagan D, Tian Y, Seshadri R, Salwen HR, *et al*: **Association of epigenetic inactivation of RASSF1A with poor outcome in human neuroblastoma.** *Clin Cancer Res* 2004, **10**:8493-8500.
36. Esteller M, Corn PG, Baylin SB, Herman JG: **A gene hypermethylation profile of human cancer.** *Cancer Res* 2001, **61**:3225-3229.
37. Costello JF, Fruhwald MC, Smiraglia DJ, Rush LJ, Robertson GP, Gao X, *et al*: **Aberrant CpG-island methylation has non-random and tumour-type-specific patterns.** *Nat Genet* 2000, **24**:132-138.
38. Suzuki F, Akahira J, Miura I, Suzuki T, Ito K, Hayashi S, *et al*: **Loss of estrogen receptor beta isoform expression and its correlation with aberrant DNA methylation of the 5'-untranslated region in human epithelial ovarian carcinoma.** *Cancer Sci* 2008, **99**:2365-2372.
39. Ueoka Y, Kato K, Kuriaki Y, Horiuchi S, Terao Y, Nishida J, *et al*: **Hepatocyte growth factor modulates motility and invasiveness of ovarian carcinomas via Ras-mediated pathway.** *Br J Cancer* 2000, **82**:891-899.
40. Kikuchi R, Tsuda H, Kanai Y, Kasamatsu T, Sengoku K, Hirohashi S, *et al*: **Promoter hypermethylation contributes to frequent inactivation of a putative conditional tumor suppressor gene connective tissue growth factor in ovarian cancer.** *Cancer Res* 2007, **67**:7095-7105.
41. Vambergue A, Fajardy I, Dufour P, Valat AS, Vandersippe M, Fontaine P, *et al*: **No loss of genomic imprinting of IGF-II and H19 in placentas of diabetic pregnancies with fetal macrosomia.** *Growth Horm IGF Res* 2007, **17**:130-136.
42. Wylie AA, Murphy SK, Orton TC, Jirtle RL: **Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/H19 regulation.** *Genome Res* 2000, **10**:1711-1718.
43. Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, *et al*: **Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour.** *Nature* 1993, **362**:749-751.
44. Maegawa S, Yoshioka H, Itaba N, Kubota N, Nishihara S, Shirayoshi Y, *et al*: **Epigenetic silencing of PEG3 gene expression in human glioma cell lines.** *Mol Carcinog* 2001, **31**:1-9.
45. Pedersen IS, Dervan PA, Broderick D, Harrison M, Miller N, Delany E, *et al*: **Frequent loss of imprinting of PEG1/MEST in invasive breast cancer.** *Cancer Res* 1999, **59**:5449-5451.
46. Higashimoto K, Soejima H, Yatsuki H, Katsuki T, Mukai T: **An NsiI RFLP in the human long QT intronic transcript 1 (LIT1).** *J Hum Genet* 2000, **45**:96-97.
47. Mitsuya K, Meguro M, Lee MP, Katoh M, Schulz TC, Kugoh H, *et al*: **An imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids.** *Hum Mol Genet* 1999, **8**:1209-1217.
48. MacDonald HR, Wevrick R: **The necdin gene is deleted in Prader-Willi syndrome and is imprinted in human and mouse.** *Hum Mol Genet* 1997, **6**:1873-1878.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1755-8794/5/8/prepub

doi:10.1186/1755-8794-5-8
Cite this article as: Hiura *et al.*: High-throughput detection of aberrant imprint methylation in the ovarian cancer by the bisulphite PCR-Luminex method. *BMC Medical Genomics* 2012 5:8.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Clinical Cancer Research

Imaging, Diagnosis, Prognosis

High-Risk Ovarian Cancer Based on 126-Gene Expression Signature Is Uniquely Characterized by Downregulation of Antigen Presentation Pathway

Kosuke Yoshihara[1], Tatsuhiko Tsunoda[3], Daichi Shigemizu[3], Hiroyuki Fujiwara[5], Masayuki Hatae[6], Hisaya Fujiwara[7], Hideaki Masuzaki[8], Hidetaka Katabuchi[9], Yosuke Kawakami[10], Aikou Okamoto[11], Takayoshi Nogawa[15], Noriomi Matsumura[16], Yasuhiro Udagawa[17], Tsuyoshi Saito[18], Hiroaki Itamochi[19], Masashi Takano[20], Etsuko Miyagi[4], Tamotsu Sudo[21], Kimio Ushijima[22], Haruko Iwase[12], Hiroyuki Seki[23], Yasuhisa Terao[13], Takayuki Enomoto[24], Mikio Mikami[25], Kohei Akazawa[2], Hitoshi Tsuda[14], Takuya Moriya[26], Atsushi Tajima[27], Ituro Inoue[28], and Kenichi Tanaka[1] for The Japanese Serous Ovarian Cancer Study Group

Abstract

Purpose: High-grade serous ovarian cancers are heterogeneous not only in terms of clinical outcome but also at the molecular level. Our aim was to establish a novel risk classification system based on a gene expression signature for predicting overall survival, leading to suggesting novel therapeutic strategies for high-risk patients.

Experimental Design: In this large-scale cross-platform study of six microarray data sets consisting of 1,054 ovarian cancer patients, we developed a gene expression signature for predicting overall survival by applying elastic net and 10-fold cross-validation to a Japanese data set A ($n = 260$) and evaluated the signature in five other data sets. Subsequently, we investigated differences in the biological characteristics between high- and low-risk ovarian cancer groups.

Results: An elastic net analysis identified a 126-gene expression signature for predicting overall survival in patients with ovarian cancer using the Japanese data set A (multivariate analysis, $P = 4 \times 10^{-20}$). We validated its predictive ability with five other data sets using multivariate analysis (Tothill's data set, $P = 1 \times 10^{-5}$; Bonome's data set, $P = 0.0033$; Dressman's data set, $P = 0.0016$; TCGA data set, $P = 0.0027$; Japanese data set B, $P = 0.021$). Through gene ontology and pathway analyses, we identified a significant reduction in expression of immune-response–related genes, especially on the antigen presentation pathway, in high-risk ovarian cancer patients.

Conclusions: This risk classification based on the 126-gene expression signature is an accurate predictor of clinical outcome in patients with advanced stage high-grade serous ovarian cancer and has the potential to develop new therapeutic strategies for high-grade serous ovarian cancer patients. *Clin Cancer Res; 18(5); 1374–85. ©2012 AACR.*

Authors' Affiliations: [1]Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences; [2]Department of Medical Informatics and Statistics, Niigata University Graduate School of Medical and Dental Sciences, Niigata; [3]Laboratory for Medical Informatics, Center for Genomic Medicine, RIKEN Yokohama Institute; [4]Department of Obstetrics and Gynecology, Yokohama City University Hospital, Yokohama; [5]Department of Obstetrics and Gynecology, Jichi Medical University, Shimotsuke; [6]Department of Obstetrics and Gynecology, Kagoshima City Hospital, Kagoshima; [7]Department of Obstetrics and Gynecology, Hiroshima University Graduate School of Biomedical Sciences, Hiroshima; [8]Department of Obstetrics and Gynecology, School of Medicine, Nagasaki University, Nagasaki; [9]Department of Gynecology and Obstetrics, Faculty of Life Sciences, Kumamoto University, Kumamoto; [10]Department of Gynecology, National Hospital Organization, Kure Medical Center and Chugoku Cancer Center, Kure; [11]Department of Obstetrics and Gynecology, Jikei University School of Medicine; [12]Department of Gynecology, Cancer Institute Hospital; [13]Department of Obstetrics and Gynecology, Faculty of Medicine, Juntendo University; [14]Department of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tokyo; [15]Department of Gynecology, National Hospital Organization, Shikoku Cancer Center, Matsuyama; [16]Department of Gynecology and Obstetrics, Kyoto University Graduate School of Medicine, Kyoto; [17]Department of Obstetrics and Gynecology, Fujita Health University School of Medicine, Toyoake; [18]Department of Obstetrics and Gynecology, Sapporo Medical University, Sapporo; [19]Department of Obstetrics and Gynecology, Tottori University School of Medicine, Yonago; [20]Department of Obstetrics and Gynecology, National Defense Medical College, Tokorozawa; [21]Department of Gynecologic Oncology, Hyogo Cancer Center, Akashi; [22]Department of Obstetrics and Gynecology, Kurume University School of Medicine, Kurume; [23]Department of Obstetrics and Gynecology, Saitama Medical Center, Saitama Medical University, Kawagoe; [24]Departments of Obstetrics and Gynecology, Osaka University Graduate School of Medicine, Suita; [25]Department of Obstetrics and Gynecology, Tokai University School of Medicine, Isehara; [26]Department of Pathology, Kawasaki Medical School, Kurashiki; [27]Department of Human Genetics and Public Health, Institute of Health Biosciences, The University of Tokushima Graduate School, Tokushima; and [28]Division of Human Genetics, National Institute of Genetics, Mishima, Japan

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

The Japanese Ovarian Cancer Study Group listed in random order:

Mitsuaki Suzuki (Jichi Medical University); Yoshitaka Onishi, Kazunobu Sueyoshi, and Sumika Matsukida (Kagoshima City Hospital); Yoshiki Kudo

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Translational Relevance

Using large-scale microarray expression data sets (n = 1,054) by applying an elastic net method, a novel risk classification system for predicting overall survival of patients with advanced stage high-grade serous ovarian cancer based on a 126-gene expression signature was developed and successfully validated. This study has profound significance in clarifying the downregulation of human leukocyte antigen class I antigen presentation machinery that characterizes high-risk ovarian cancer. These results from comprehensive gene expression analysis using large-scale microarray data suggest that our risk classification system might have the potential to optimize treatment of high-grade serous ovarian cancer patients.

Introduction

High-grade serous ovarian cancer comprises approximately 40% of epithelial ovarian cancer cases and is the most aggressive histologic type (1–4). This type of cancer usually presents as advanced stage disease at the time of diagnosis because there are no symptoms present at the early stage and no reliable screening test for early detection (1–4). Patients with advanced stage high-grade serous ovarian cancer generally undergo primary debulking surgery followed by platinum–taxane chemotherapy. However, 30% to 40% of patients recur within 12 months after the standard treatment, and the overall 5-year survival rate remains at approximately 30% (5, 6). Clinicopathologic characteristics, such as the International Federation of Gynecology and Obstetrics (FIGO) stage, histologic grade, and debulking status after primary surgery, are clinically considered important clinical prognostic indicators of ovarian cancer but are insufficient for predicting survival time.

The development of microarray technology has provided new insights into cancer diagnosis and treatment. Large-scale microarray studies in breast cancer have succeeded in clarifying 5 molecular subtypes based on gene expression profiles and in developing genomic biomarker for predicting recurrence in early breast cancer (MammaPrint; refs. 7, 8). Thus, breast cancer treatment strategies are being stratified according to molecular characteristics. In contrast, there are no gene expression signatures with high accuracy and reproducibility for clinical diagnosis and management in patients with ovarian cancer because there is a paucity of ovarian cancer samples available for microarray analysis compared with breast cancer. Although *TP53* somatic mutation is present in almost all high-grade serous ovarian cancer and plays an important role in the pathogenesis (9, 10), high-grade serous ovarian cancer exhibits much biological and molecular heterogeneity that should be considered when developing a novel therapeutic strategy for ovarian cancer (10, 11).

In this study, we aimed to establish a novel system for predicting the prognosis of patients with advanced stage high-grade serous ovarian cancer using large-scale microarray data sets (n = 1,054; refs. 10–13), leading to an optimal treatment based on molecular characteristics (14).

Materials and Methods

Clinical samples

Three hundred Japanese patients who were diagnosed with advanced stage high-grade serous ovarian cancer between July 1997 and June 2010 were included in this study. All patients provided written informed consent for the collection of samples and subsequent analysis. Fresh-frozen samples were obtained from primary tumor tissues during debulking surgery prior to chemotherapy. All patients with advanced stage high-grade serous ovarian cancer were treated with platinum–taxane standard chemotherapy after surgery. In principle, patients were seen every 1 to 3 months for the first 2 years. Thereafter, follow-up visits had an interval of 3 to 6 months in the third to fifth year, and 6 to 12 months in the sixth to tenth year. At every follow-up visit, general physical and gynecologic examinations were carried out. CA125 serum levels were routinely determined.

Staging of the disease was assessed according to the criteria of the FIGO (15). Optimal debulking surgery was defined as less than 1 cm of gross residual disease, and suboptimal debulking surgery was defined as more than 1 cm of residual disease. Progression-free survival time was calculated as the interval from primary surgery to disease progression or recurrence. Based on the Response Evaluation Criteria in Solid Tumors (RECIST, version 1.1; 16), disease progression was defined as at least a 20% increase in the sum of the diameters of target lesions, as unequivocal progression of existing nontarget lesions, or as the appearance of one or more new lesions. Overall survival time was calculated as the interval from primary surgery to the death due to ovarian cancer.

(Hiroshima University); Hironori Tashiro (Kumamoto University); Tomoya Mizunoe (Kure Medical Center and Chugoku Cancer Center); Junzo Kigawa, Kanae Nosaka, and Hisao Ito (Tottori University); Sohei Yamamoto and Hideyuki Shimazaki (National Defense Medical College); Ken Takizawa (Cancer Institute Hospital); Kiyoko Kato and Satoru Takeda (Juntendo University); Yutaka Ueda, Yukari Miyoshi, Toshihiro Kimura, and Tadashi Kimura (Osaka University); Sosuke Adachi, Koji Nishino, Takehiro Serikawa, Tetsuro Yahata, Junko Sakurada, Go Hasegawa, and Nobutaka Kitamura (Niigata University).

Corresponding Author: Kenichi Tanaka, Niigata University Graduate School of Medical and Dental Sciences, 1-757 Asahimachi-dori, Chuo-ku, Niigata 951-8510, Japan. Phone: 81-25-227-2317; Fax: 81-25-227-0789; E-mail: tanaken@med.niigata-u.ac.jp

doi: 10.1158/1078-0432.CCR-11-2725

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

The histologic characteristics of surgically resected specimens, which were diagnosed as serous histologic type by 2 pathologists at each institute, were at first assessed on formalin-fixed and paraffin-embedded hematoxylin and eosin sections by 8 gynecologic pathologists who belong to Department of Pathology in 4 institutes (Niigata University, National Defense Medical Collage, Tottori University, Kagoshima City Hospital). Next, these sections were reviewed under central pathologic review by 2 independent gynecologic pathologists (H.T. and T.M.) with no knowledge of patients' clinical data. Histologic subtype was diagnosed based on WHO classification of ovarian tumors (17). The degree of histologic differentiation is determined according to Silverberg classification (18).

Microarray experiments

Frozen tissues containing more than 70% tumor cells upon histologic evaluation were used for RNA extraction. Total RNA was extracted from tissue samples as previously described (19). Five hundred nanograms of total RNA was converted into labeled cRNA with nucleotides coupled to a cyanine 3-CTP (Cy3; PerkinElmer) using the Quick Amp Labeling Kit, one-color (Agilent Technologies). Cy3-labeled cRNA (1.65 μg) was hybridized for 17 hours at 65°C to an Agilent Whole Human Genome Oligo Microarray (G4112F), which carries 60-mer probes to more than 40,000 human transcripts. The hybridized microarray was washed and then scanned in Cy3 channel with the Agilent DNA Microarray Scanner [model G2565CA, $n = 260$ (Japanese data set A); G2565AA, $n = 40$ (Japanese data set B)]. Signal intensity per spot was generated from the scanned image with Feature Extraction Software (version 10.1, Japanese data set A; version 9.1, Japanese data set B; Agilent Technologies) with default settings. Spots that did not pass quality control procedures were flagged as "Not Detected."

Next, we obtained Affymetrix HG-U133Plus2.0 microarray (Affymetrix) data from 10 ovarian cancer samples that had been already been analyzed by the Agilent Whole Human Genome Oligo Microarray. Ten ovarian cancer samples were randomly selected from 260 samples in the Japanese data set A. Microarray experiments were carried out according to the Affymetrix-recommended protocols. Briefly, biotinylated cRNAs were synthesized by GeneChip 3'IVT Express Kit (Affymetrix) from 250 ng total RNA according to the manufacturer's instructions. Biotinylated cRNA yield were checked with the NanoDrop ND-1000 Spectrophotometer. Following fragmentation, 10 μg of cRNA were hybridized for 16 hours at 45°C on GeneChip Human Genome U133 Plus 2.0 Array. GeneChips were washed and stained in the Affymetrix Fluidics Station 450, and scanned with GeneChip Scanner 3000 7G.

The MIAME-compliant microarray data were deposited into the Gene Expression Omnibus data repository (accession number GSE32062 and GSE32063).

Microarray data analysis

We prepared 2 our microarray data sets [Japanese data set A ($n = 260$) and B ($n = 40$)] and 4 publicly available large sample-sized ($n > 100$) microarray data sets [TCGA data set (http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp; ref. 11), Tothill's data set (GSE9891; ref. 12), Bonome's data set (GSE26712; ref. 13), and Dressman's data set (14)] to discover predictive biomarkers. Clinical information of publicly available microarray data sets was obtained from their articles and websites. From Tothill's original data set ($n = 285$), we selected 131 samples that (i) were diagnosed as advanced stage serous adenocarcinoma, (ii) were treated by platinum/taxane–based chemotherapy, and (iii) have clinical data about onset age, stage, grade, surgery, and survival time. Publicly available clinical information in TCGA was downloaded from TCGA Data portal (http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) at June 25, 2011. By the same methods above, 319 samples were selected from 562 microarray data as TCGA data set (Affymetrix HT-HG-U133A). In this study, patients who were treated by adjuvant chemotherapy including molecular-targeted agents were excluded.

On the Agilent platform, data normalization was carried out in GeneSpring GX 11.5 (Agilent Technologies) as follows: (i) threshold raw signals were set to 1.0 and (ii) 75th percentile normalization. Affymetrix microarray data were normalized and summarized with robust multiple average in GeneSpring GX 11.5 (Agilent Technologies). To compare the microarray data sets measured with 4 different platforms (Agilent Whole Human Genome Oligo Microarray, Affymetrix HG-U133A, HG-U133Plus2.0, and HT-HG-U133A), we selected genes common to all platforms based on the Entrez Gene ID and used the Median Rank Score method (20) for cross-platform normalization (Supplementary Fig. S1). Of 22,277 probes that were common among 3 Affymetrix platforms, 20,331 probes were selected to be in one to one relation between probe and gene. Using translation function based on Entrez Gene ID in GeneSpring GX 11.5, the 19,704 transcripts that matched to the 20,331 probes of the Affymetrix platform were extracted from all transcripts on the Agilent platform. Considering differences in microarray platforms, coefficient of correlation (r) in each gene between 2 microarray platforms were measured. Using 10 ovarian cancer microarray data obtained from both Agilent Whole Human Genome Oligo Microarray and Affymetrix HG-U133Plus2.0 (GSE32062), 9,141 genes with high correlation ($r > 0.8$) were extracted. After Median Rank Score analysis (20), we evaluated median value of each gene in both platforms and selected 3,553 genes with high correlation ($r > 0.8$) in which absolute value of subtracting median values between 2 platforms was less than 1. Similar analyses were conducted by 16 breast cancer microarray data (GSE17700; ref. 21) from both Affymetrix HG-U133Plus2.0 and HG-U133A. From the resulting 1,746 genes, we removed 60 that were not flagged as "Detected" in more than 90% of the Japanese data set A samples ($n = 260$), considering them to have either missing or uncertain expression signals. In addition, the data were normalized per gene in each data set by transforming the expression of each gene to obtain a mean of 0 and SD of 1 (Z-transformation) for the cross-platform study.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

We analyzed Japanese data set A as a "training set," Tothill's data set as a "test set," and the other 4 data sets as "validation sets." We applied elastic net analysis (22) with the R (23) package glmnet to identify survival-related genes for prediction of prognosis in patients with advanced stage high-grade serous ovarian cancer. Using 10-fold cross-validation, we obtained regression coefficients with optimal penalty parameter for the penalized Cox model, and calculated a prognostic index for each patient as defined by

$$\text{Prognostic index} = \sum_{i=1}^{126} \beta_i \times X_i$$

where $_i$ is the estimated regression coefficient of each gene in the Japanese data set A under elastic net ($\alpha = 0.05$) and X_i is the Z-transformed expression value of each gene. The estimated regression coefficient of each survival-related gene given by elastic net ($\alpha = 0.05$) in the Japanese data set A was also applied to calculate a prognostic index for each patient in 5 other data sets using the equation above. We classified all patients into the 2 groups (high- and low-risk groups) by the optimal cutoff value of the prognostic index in the Japanese data set A. Patients were assigned to the "high-risk" group if their prognostic index was more than or equal to cutoff value of prognostic index, whereas "low-risk" group was composed of cases with the prognostic indices that were less than cutoff value. Because risk classification divided by cutoff value 0.1517 indicated a minimum P value ($P = 1 \times 10^{-30}$) when log-rank test was used to compare differences in overall survival between high- and low-risk groups in the Japanese data set A, this value was determined as optimal cutoff value.

Both hierarchical clustering and non-negative factorization (NMF) algorithm (24) were used to assess the similarity of gene expression profiles among 126 survival-related genes. In the 2 major data sets (Japanese data set A and TCGA data set), we constructed a heat map with hierarchical clustering for both samples and genes using Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm). Correlation (centered) and complete linkage were selected on similarity metrics and clustering method, respectively. A heat map was visualized with Java TreeView (http://jtreeview.sourceforge.net/). NMF-consensus matrices averaging 50 connectivity matrices were computed at $K = 2$–7 (as the number of subclasses modeled) for 126 genes. With genes appearing along both the horizontal and vertical axes of the consensus matrices, consistency in the gene-pair clusters was visualized. We determined the optimal number of gene clusters by the cophenetic correlation coefficient that provides a scalar summary of global clustering robustness across the consensus matrix in the Japanese data set A and TCGA data set.

We conducted a volcano plot analysis to extract differentially expressed transcripts between high- and low-risk ovarian cancer groups with the 2 major data sets [Japanese data set A ($n = 260$) and TCGA data set ($n = 319$)]. When the volcano plot analysis was conducted in GeneSpring GX 11.5 (Agilent Technologies), we used gene expression data prior to Z-transformation normalization (Fig. 3).

To investigate the biological functions of gene expression signatures, we used GO Ontology Browser, embedded in GeneSpring GX11.5 (Agilent Technologies). The GO Ontology Browser was used to analyze which categories of gene ontology were statistically overrepresented among the gene list obtained. Statistical significance was determined by Fisher exact test, followed by multiple testing corrections by the Benjamini and Yekutieli false discovery rate method (25) To avoid bias of gene extraction by volcano plot analysis in GO analysis, Gene Set Enrichment Analysis (GSEA; ref. 26) was conducted with genes prior to gene selection by volcano plot analysis. Analysis settings in GSEA (software version 2.07) were as follows: (i) gene sets database: c5.all.v2.5.symbols.gmt [gene ontology], (ii) number of permutations: 1,000, (iii) collapse data set to gene symbol: true, (iv) permutation type: phenotype, (v) chip platform(s): Agilent_HumanGenome.chip or HT_HG_U133A.chip, and (vi) other settings: default.

Furthermore, we used the Core Analysis tool in the Ingenuity Pathway Analysis (IPA) system to analyze networks and pathways for a set of genes. Q value was calculated by Fisher exact test with the Benjamini and Hochberg correction (27). $Q < 0.25$ was considered as significant in GO, GSEA, and IPA.

Single-nucleotide polymorphism array experiments

We isolated genomic DNA from tumor tissues with a phenol-chloroform extraction method and from normal lymphocytes using the QIAamp DNA Blood Maxi Kit (QIAGEN). Single-nucleotide polymorphism (SNP) array experiments with the Genome-Wide Human SNP Array 6.0 (Affymetrix) were carried out at Niigata University (details in the Supplementary Methods). SNP array data were analyzed by the Partek Genomic Suite 6.5 (Partek Inc.) to investigate copy number variations in 30 genes involved in the antigen presentation pathway.

Immunoshitochemical analysis

A monoclonal anti-human CD8 antibody (M7103; 1:100; DAKO) was used for immunohistochemical staining of CD8 T lymphocyte in tumor tissues. Details of the immunohistochemistry method and sample selections are described in the Supplementary Methods.

Statistical analysis

Standard statistical tests including Pearson correlation analysis, unpaired t tests, 1-way ANOVA, Fisher exact tests, log-rank tests, and Cox proportional hazard model analysis were used to analyze the clinical data, as appropriate. Analyses of clinical data were conducted with JMP (version 8; SAS Institute) and GraphPad PRISM (version 4.0; GraphPad Software).

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Results

The distribution of the clinical variables for each microarray data set is shown in Table 1. To compare the microarray data sets measured with 4 different platforms, 1,686 genes were selected (Supplementary Fig. S1). An elastic net analysis (22) using 10-fold cross-validation on Japanese data set A (training set, $n = 260$) identified a 126-gene signature for predicting overall survival in patents with advanced stage high-grade serous ovarian cancer (Supplementary Table S1). After calculating the prognostic index for each sample from the 126-gene expression signature as reported previously (19), we divided the training set into high- and low-risk groups based on the optimal cutoff value (0.1517) of the prognostic index and verified the high predictive power of this risk classification (log-rank $P = 1 \times 10^{-30}$; Multivariate Cox $P = 4 \times 10^{-20}$; HR = 6.203; 95% CI = 4.239–9.123; Fig. 1A and Table 2).

The predictive power of the 126-gene signature was tested with Tothill's data set ($n = 131$; ref. 12). Both Kaplan–Meier survival and multivariate analysis showed that this risk classification was significantly associated with overall survival time in Tothill's data set (log-rank $P = 3 \times 10^{-6}$; Multivariate Cox $P = 1 \times 10^{-5}$; HR = 3.728; 95% CI = 2.110–6.532; Fig. 1B and Table 2). Next, we assessed the predictive power of the 126-gene expression signature on the 3 data sets in which microarray data were obtained from Affymetrix platforms and confirmed that this risk classification indicated similar results in the 3 data sets (Fig. 1C–E and Table 2). Furthermore, to exclude the influence of differences in microarray platforms, we prepared Japanese data set B from the same platform as the training set. Despite the small sample size, our risk classification showed a significant association with overall survival time in the Japanese data set B (Fig. 1F and Table 2). This risk classification showed high predictive accuracy even though

Table 1. Clinicopathologic characteristics in 6 microarray data sets

Data set	Japanese data set A	Tothill's data set[a]	Bonome's data set[a]	Dressman's data set[a]	TCGA data set[c]	Japanese data set B
Number	260	131	185	119	319	40
Age	58.2 ± 10.8	58.4 ± 9.8	62 ± 12	N/A[b]	59.5 ± 11.3	56.2 ± 9.6
Histology						
Serous	260	131	166	119	319	40
Others	0	0	19	0	0	0
Stage						
III	204	123	144	99	267	31
IV	56	8	41	20	52	9
Grade	Silverberg	Silverberg	N/A[b]	N/A[b]	N/A[b]	Silverberg
1	0	0	0	3	0	0
2	131	51	40	57	29	23
3	129	80	144	59	289	17
4	—	—	3	—	1	—
Surgery status						
Optimal	103	78	92	63	236	19
Suboptimal	157	45	93	56	83	21
Chemotherapy						
Platinum	260	131	185	119	319	40
Taxane	260	131	N/A	82	319	40
Follow-up period (mo)						
Median	42	29	38	34	31	39
Range	1–128	6–79	1–164	1–185	1–154	7–111
Number of deaths	121	60	129	69	187	22
Median survival (mo)	60	44	46	69	42	54
Microarray platform	Agilent Whole Human Genome Oligo Microarray	Affymetrix HG-U133 Plus2.0	Affymetrix HG-U133A	Affymetrix HG-U133A	Affymetix HT-HG-U133A	Agilent Whole Human Genome Oligo Microarray

[a]Clinical information in these 3 data sets were obtained from their articles and website.
[b]No available information was described as N/A.
[c]Publicly available clinical information in TCGA was downloaded from TCGA Data portal (http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) at June 25, 2011.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

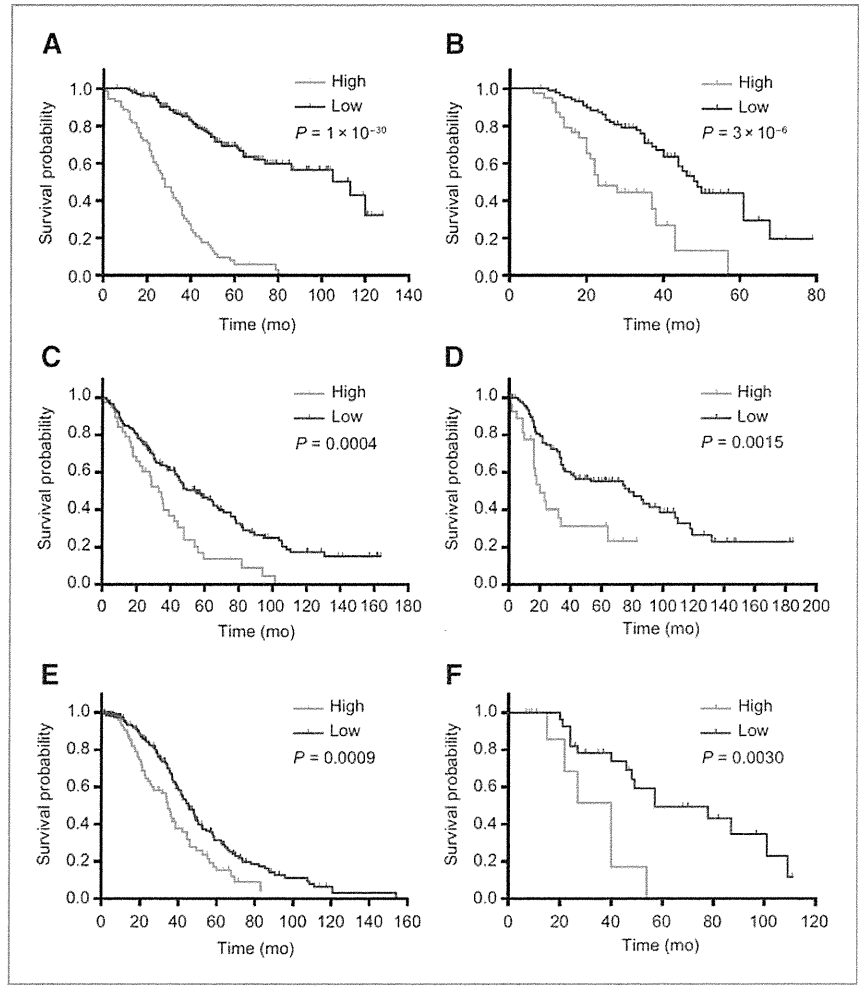

Figure 1. Kaplan–Meier survival analysis in 6 microarray data sets (A, Japanese data set A; B, Tothill's data set; C, Bonome's data set; D, Dressman's data set; E, TCGA data set; F, Japanese data set B). Based on 126-gene signature, ovarian cancer patients were divided into 2 risk groups (red, high risk; blue, low risk). *P* values correspond to the log-rank test comparing the survival curves.

overall survival was capped at 60 months (Supplementary Fig. S2; ref. 10). Moreover, the high-risk groups had shorter progression-free survival times compared with the low-risk groups in 4 microarray data sets with available progression-free survival times (Supplementary Fig. S3). In addition, we applied 193 overall survival related gene signature that have been recently reported by TCGA research networks (10) to our large data set (Japanese data set A), and 191 of 193 genes were available in our data set. High-risk ovarian cancer patients based on 191-gene signature model had significantly poor prognosis compared with low-risk patients (Supplementary Methods and Supplementary Fig. S4).

We next investigated the molecular characteristics of the 126-gene signature with both hierarchical clustering and NMF (24) analyses using the 2 major data sets (Japanese data set A and TCGA data set). Hierarchical clustering of the expression of 126 genes in the Japanese data set A ($n = 260$) resulted in 2 gene clusters (cluster 1 and 2; Fig. 2A). Cluster 1 (outlined in yellow in Fig. 2A) comprised 23 genes, 22 of which were common genes in a small cluster of NMF class assignment for $K = 2$ (Supplementary Fig. S5). Similarly, this gene cluster was seen both in hierarchical clustering and NMF analyses of TCGA data set, and the 20 genes in TCGA's cluster were consistent with these 23 genes in cluster 1 (Fig. 2B and Supplementary Fig. S5). GO analysis of these 23 genes revealed 132 significantly overrepresented GO categories; the top 20 GO categories are shown in the Supplementary Table S2. Specifically, immunity-related categories were enriched in these 23 genes and were downregulated in the high-risk groups compared with the low-risk groups. On the contrary, only 1 category [cytoplasm (GO0005737)] belonging to "cellular component" was significantly overrepresented in 103 genes of cluster 2.

To clarify the biological significance of our risk classification based on the 126-gene signature in advanced stage high-grade serous ovarian cancer, we examined differences in molecular biological characteristics between high- and low-risk groups. For subsequent analyses, all transcripts on

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Table 2. Univariate and multivariate Cox's proportional hazard model analysis of prognostic factors for overall survival

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*	HR (95% CI)	*P*
Japanese data set A				
Age	1.011 (0.993–1.029)	0.21	1.006 (0.988–1.024)	0.52
Stage IV (vs. stage III)	1.465 (0.968–2.165)	0.07	1.340 (0.884–1.984)	0.16
Optimal surgery (vs. suboptimal)	0.499 (0.334–0.730)	0.0003	0.643 (0.428–0.949)	0.026
High (vs. low)	6.823 (4.692–9.972)	2×10^{-22}	6.203 (4.239–9.123)	4×10^{-20}
Tothill's data set				
Age	1.005 (0.977–1.035)	0.73	1.006 (0.978–1.035)	0.67
Stage IV (vs. stage III)	2.264 (0.785–5.176)	0.12	2.951 (0.983–7.235)	0.053
Optimal surgery (vs. suboptimal)	0.887 (0.525–1.517)	0.66	0.910 (0.528–1.592)	0.74
High (vs. low)	3.443 (1.967–5.962)	3×10^{-5}	3.728 (2.110–6.532)	1×10^{-5}
Bonome's data set				
Optimal surgery (vs. suboptimal)	0.592 (0.414–0.840)	0.0032	0.628 (0.438–0.894)	0.0097
High (vs. low)	2.034 (1.341–3.012)	0.0011	1.897 (1.247–2.818)	0.0033
Dressman's data set				
Optimal surgery (vs. suboptimal)	0.685 (0.425–1.102)	0.12	0.602 (0.370–0.977)	0.04
High (vs. low)	2.397 (1.335–4.140)	0.0042	2.687 (1.480–4.705)	0.0016
TCGA data set				
Age	1.014 (1.001–1.028)	0.038	1.011 (0.997–1.025)	0.13
Stage IV (vs. stage III)	1.018 (0.683–1.469)	0.93	1.034 (0.692–1.499)	0.86
Optimal surgery (vs. suboptimal)	0.909 (0.668–1.254)	0.56	0.903 (0.654–1.261)	0.54
High (vs. low)	1.712 (1.234–2.345)	0.0015	1.680 (1.202–2.321)	0.0027
Japanese data set B				
Age	1.059 (1.007–1.118)	0.024	1.058 (0.999–1.131)	0.057
Stage IV (vs. stage III)	1.696 (0.542–4.501)	0.34	2.998 (0.808–10.60)	0.098
Optimal surgery (vs. suboptimal)	0.582 (0.231–1.364)	0.22	0.510 (0.175–1.413)	0.20
High (vs. low)	4.100 (1.376–11.25)	0.013	4.468 (1.265–15.49)	0.021

each platform were used to avoid the previous gene number limitations, which were for the cross-platform study. We extracted the genes differentially expressed between the 2 groups by carrying out a volcano plot analysis with the 2 major data sets (Japanese data set A and TCGA data set; Supplementary Table S3). Figure 3A shows that 1,109 and 1,381 transcripts were differentially expressed between the high- and low-risk groups in the Japanese data set A and TCGA data set, respectively. GO analysis of these transcripts indicated that 9 out of 20 top-ranked GO categories were common between both data sets and involved in the immune system (Table 3 and 4).

To exclude the influence of cutoff in fold change or *P* value in volcano plot analysis, we reevaluated differences in the biological characteristics between high- and low-risk groups using GSEA (26). Four immunity-related GO categories (immune system response, immune response, defense response, and inflammatory response) were also included in the list of 20 top-ranked GO categories when GSEA was carried out with 21,785 (Japanese data set A) and 14,023 (TCGA data set) transcripts prior to gene extraction (Supplementary Table S4).

Furthermore, possible functional relations among differentially expressed genes between high- and low-groups were investigated with pathway analysis. Of 20 top-ranked pathways notably enriched in 1,109 transcripts of the Japanese data set A and 1,381 transcripts of the TCGA data set, 14 pathways were common between both data sets (Supplementary Table S5). In particular, the antigen presentation pathway (Fig. 3B) was the most significantly overrepresented pathway in the 2 data sets (Japanese data set A, $Q = 2.0 \times 10^{-29}$; TCGA data set, $Q = 5.0 \times 10^{-14}$). In the antigen presentation pathway, 30 transcripts in this pathway were significantly downregulated in the high-risk group (Fig. 3B and Supplementary Table S6). We further examined whether molecular defects of human leukocyte antigen (HLA) class I antigen presentation machinery components were caused by structural alterations using SNP array data. Our SNP array data showed that genes of HLA class I antigen presentation machinery were not deleted (Supplementary Table S6). In the TCGA data set, only 10% of cases (24 of 242) had deletions in HLA class I genes. On the other hand, 89% of high-risk cases (55 of 62) without deletion in HLA class I genes showed significantly lower expressions in HLA class I genes compared with those in low-risk groups (Supplementary Fig. S6).

On the basis of this result, we assessed the status of tumor-infiltrating lymphocyte reflecting an immune response

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

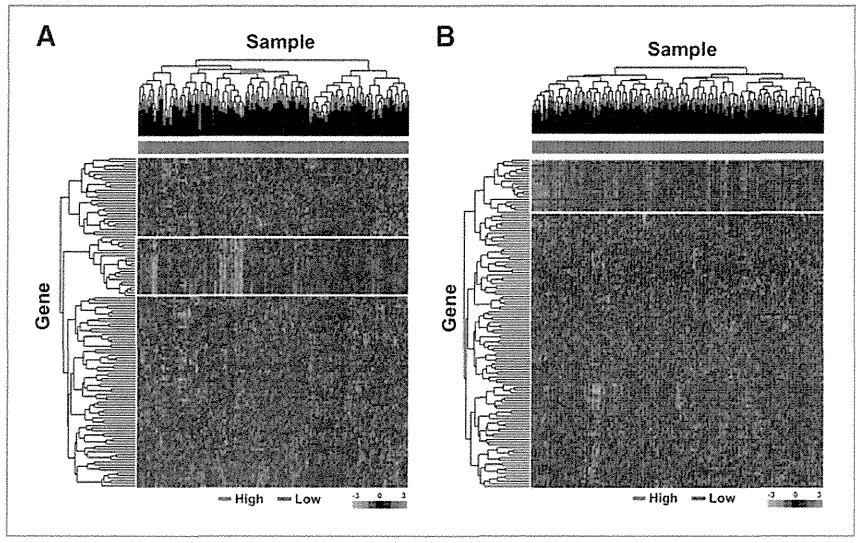

Figure 2. Molecular characteristics of 126 survival-related genes. A, heat map with a hierarchical classification of 260 patient samples with the expression profile of 126 survival-related genes in the Japanese data set A. B, heat map with a hierarchical classification of 319 patient samples with the expression profile of 126 survival-related genes in TCGA data set.

against ovarian cancer. By immunohistochemically staining for CD8-positive T lymphocytes in a subset of the Japanese data set ($n = 30$), we revealed that the number of CD8 T lymphocytes infiltrating into tumor tissues was significantly decreased in the high-risk group ($n = 10$) compared with the low-risk group ($n = 20$) and was clearly correlated with the expression level of HLA class I genes (Supplementary Fig. S7).

Table 3. Enriched GO categories among the transcripts significantly differentially expressed between high- and low-risk groups: top-ranked 20 of 534 significantly overrepresented categories ($Q < 0.25$) among 1,109 transcripts in Japanese data set A

	Genes within GO category		
GO category[a]	Number	Percentage	$-\text{Log}_{10}$ Q[b]
Immune system process (GO:0002376)	206	29.0	45.0
Immune response (GO:0006955)	172	24.2	45.0
Defense response (GO:0006952, 0002217, 0042829)	110	15.5	40.3
Response to stimulus (GO:0050896, 0051869)	251	35.4	29.2
Inflammatory response (GO:0006954)	68	9.6	26.2
Positive regulation of immune system process (GO:0002684)	40	5.6	25.2
Regulation of immune system process (GO:0002682)	52	7.3	24.5
Antigen processing and presentation (GO:0019882,0030333)	31	4.4	22.7
Signal transducer activity (GO:0004871, 0005062, 0009369, 0009370)	169	23.8	22.1
Molecular transducer activity (GO:0060089)	169	23.8	22.1
Signal transduction (GO:0007165)	224	31.5	20.8
Leukocyte activation (GO:0045321)	47	6.6	20.1
Cell activation (GO:0001775)	48	6.8	19.7
Regulation of immune response (GO:0050776)	32	4.5	18.9
Response to wounding (GO:0009611, 0002245)	70	9.9	18.7
Signal transmission (GO:0023060)	224	31.5	18.2
Signaling process (GO:0023046)	224	31.5	18.2
Regulation of cell activation (GO:0050865)	30	4.2	17.2
Positive regulation of cell activation (GO:0050867)	24	3.4	17.0
Regulation of lymphocyte activation (GO:0051249)	29	4.1	17.0

[a]Bold font denotes common categories included in top 20 lists both Japanese data set A and TCGA data set.
[b]Q value was determined by Fisher's exact test with Benjamini–Yekutieli correction.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Table 4. Enriched GO categories among the transcripts significantly differentially expressed between high- and low-risk groups: top-ranked 20 of 83 significantly overrepresented GO categories ($Q < 0.25$) among 1,381 transcripts in TCGA data set

	Genes within GO category		
GO category[a]	Number	Percentage	$-\text{Log}_{10}\ Q$[b]
Immune system process (GO:0002376)	129	20.5	20.9
Immune response (GO:0006955)	115	18.3	20.8
Defense response (GO:0006952, 0002217, 0042829)	78	12.4	10.3
Antigen processing and presentation (GO:0019882, 0030333)	25	4.0	8.4
Inflammatory response (GO:0006954)	50	7.9	8.2
Antigen processing and presentation of peptide antigen (GO:0048002)	13	2.1	6.9
Antigen processing and presentation of exogenous peptide antigen (GO:0002478)	6	1.0	5.2
MHC class I peptide loading complex (GO:0042824)	9	1.4	5.2
MHC protein complex (GO:0042611)	18	2.9	5.2
TAP complex (GO:00042825)	8	1.3	5.2
MHC protein binding (GO:0042287)	12	1.9	5.2
Antigen processing and presentation of exogenous antigen (GO:0019884)	7	1.1	5.0
Response to stimulus (GO:0050896, 0051869)	183	29.1	4.4
Antigen processing and presentation of peptide antigen via MHC class I (GO:0002474)	7	1.1	3.9
Antigen processing and presentation of peptide or polysaccharide antigen via MHC class II (GO:0002504)	12	1.9	3.9
Regulation of immune response (GO:0050776)	6	1.0	3.8
MHC class I protein binding (GO:0042288)	10	1.6	3.8
Response to wounding (GO:0009611, 0002245)	50	7.9	3.8
Positive regulation of immune response (GO:0050778)	6	1.0	3.8
Positive regulation of immune system process (GO:0002684)	6	1.0	3.8

[a]Bold font denotes common categories included in top 20 lists both Japanese data set A and TCGA data set.
[b]Q value was determined by Fisher's exact test with Benjamini–Yekutieli correction.

Discussion

In this study, we established a novel risk classification system based on the 126-gene expression signature for predicting overall survival time in patients with advanced stage high-grade serous ovarian cancer. The significant association between our risk classification and overall survival time was indicated among the 6 microarray data sets.

In expression microarray analyses, there is a well-known "curse of dimensionality" problem that the number of genes is much larger than the number of samples. The curse of dimensionality leads to a concern about the reliability of the selected genes or an overfitting phenomenon, and using a large-scale microarray data is a simple and effective method to overcome this problem. From this theory, we planned a large-scale cross-platform study using 4 publicly available

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

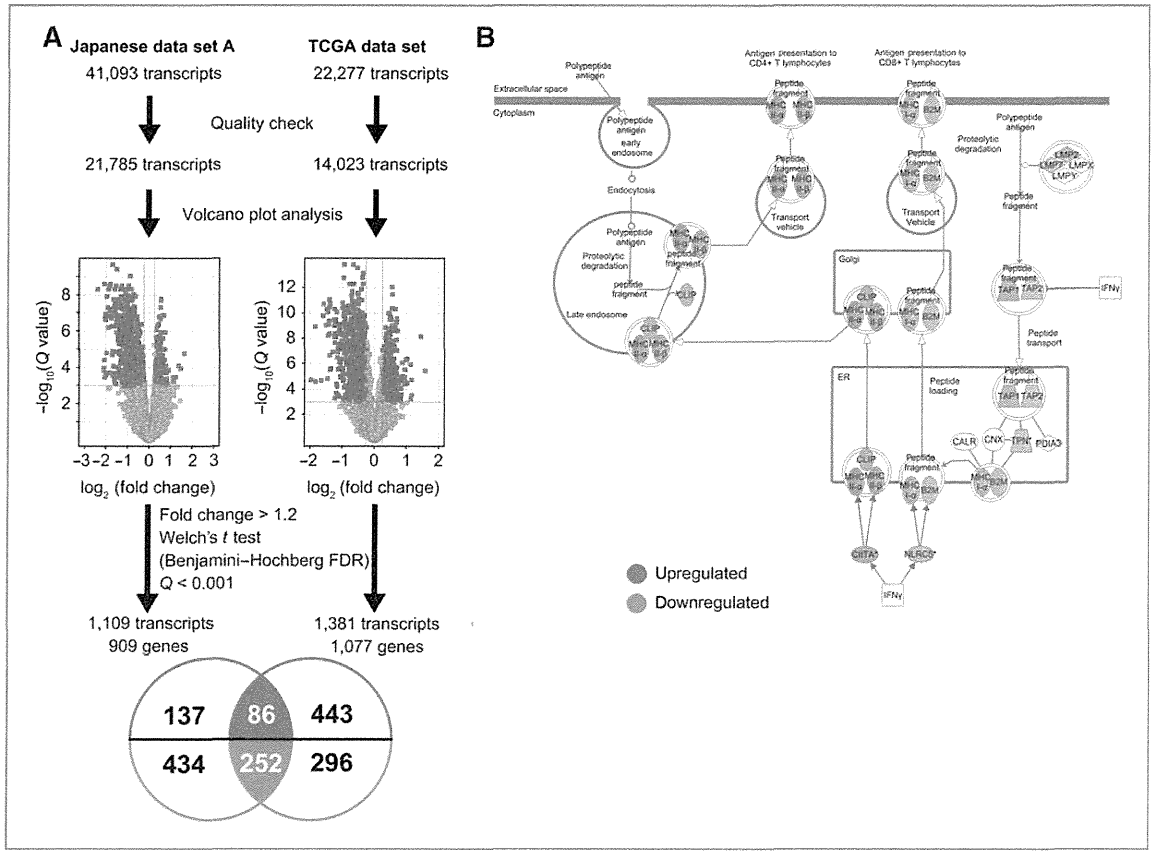

Figure 3. Molecular characteristics of differentially expressed genes between high- and low-risk groups in the 2 major data sets (Japanese data set A and TCGA data set). A, using a volcano plot analysis, 1,109 and 1,381 transcripts were extracted as differentially expressed genes between high- and low-risk groups in the Japanese data set A and TCGA data set, respectively. B, IPA showed that the antigen presentation pathway was significantly overrepresented in 1,109 Japanese data set A and 1,381 TCGA transcripts. Green indicates that gene expression was downregulated in high-risk group compared with low-risk group.

microarray data sets consisting of more than 100 samples. To carry out a cross-platform study, we took considerable care to reduce the influence of differences in microarray platforms with the Median Rank Score method and matching probes based on highly correlated expression ($r > 0.8$) between the 2 platforms. As a result, this risk classification system showed a high predictive ability in the 4 microarray data sets from different platforms. Our risk classification was suitably fitting to Japanese data set B despite the small sample size. This might reflect the same platform in the microarray experiments and similarity in clinical backgrounds between Japanese data set A and B. Intriguingly, 193 survival-related gene signature that has been reported by TCGA research networks (10) was also significantly associated with overall survival in our large data set (Supplementary Fig. S4), and 5 genes (*B4GALT5*, *CXCL9*, *ID4*, *MTRF1*, and *SLC7A11*) were common between their and our gene signatures despite different analytic processes (Supplementary Table 2). These genes might contribute strongly to developing risk classification system for high-grade serous ovarian cancer patients.

We ascertained that alterations to the immune system in cancer cells are one of the most important factors affecting survival of patients with advanced stage, high-grade serous ovarian cancer and that high-risk ovarian cancer was well characterized by the downregulation of the antigen presentation pathway at the molecular level.

The 2 large-scale subclassification studies (10, 11) based on gene expression analyses have indicated that there are at least 4 subclasses (immunoreactive, differentiated, proliferative, and mesenchymal) in high-grade serous ovarian cancer. Although there was no statistically significant difference in clinical outcome among 4 subclasses, this finding that immunity-related genes are identified in large-scale gene expression profiles despite different analytic approaches suggests that alteration of immune activity might play an important role in the molecular characterization of ovarian cancer.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

Previous findings (28–30) that the presence of tumor-infiltrating T lymphocytes is associated with long survival in ovarian cancer patients are consistent with the results from our comprehensive gene expression analysis using large-scale microarray data, suggests that the presence of immune responses to cancer cells would influence the biological phenotype of ovarian cancer. Although several antigen-specific active immunotherapy studies have been conducted, the clinical benefit of this approach has not yet been shown in large randomized-controlled trials (31, 32). Previous reports (28, 33, 34) and our data indicated that defects in HLA class I antigen presentation machinery would decrease recruitment of tumor-infiltrating lymphocytes, leading to poor prognosis in cancer patients because of a reduction in antitumor immune activity. Therefore, upregulation of HLA class I antigen presentation pathway might be one of efficient therapeutic approach for patients with high-risk serous ovarian cancer, especially when antigen-specific active immunotherapy is selected as a therapeutic strategy (35).

Defects of HLA class I gene expression occurs at the genetic, epigenetic, transcriptional, and posttranscriptional levels, and are classified into the 2 main groups: reversible and irreversible defects (36). Our data show that the frequency of deletion leading to irreversible defect of HLA class I gene expression was low in high-grade serous ovarian cancer. Only a few HLA class I gene mutations have been described thus far (36). Downregulation of antigen presentation machinery components such as *TAP1/2* or *B2M* also result in reversible defects in HLA class I molecules (37). Interestingly, the expression levels of genes in the antigen presentation pathway were positively correlated in the Japanese data set A (Supplementary Fig. S8). IFN-γ or other cytokines stimulation induces the expression of genes in the antigen presentation pathway (38). Moreover, histone deacetylase (HDAC) inhibitors increase the expression of genes in the antigen presentation pathway such as *TAP1/2* (39, 40) and *B2M* (41) in several cancer cells leading to upregulating the antigen presentation pathway. In addition, HDAC inhibitors that are promising anticancer drugs show a synergistic effect with taxane or platinum drugs, which were used as standard adjuvant chemotherapy for ovarian cancer, both *in vitro* and *in vivo* (42, 43). These genetic and epigenetic activations of the antigen presentation pathway might induce immune recognition of ovarian cancer cells and enhance antitumor immune responses.

In summary, our data suggest that this predictive biomarker based on the 126-gene signature could identify patients who should not expect long-term survival by standard treatment and that activation of the antigen presentation pathway in tumor cells is an important key in new therapeutic strategies for ovarian cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank Prof. Y. Nakamura for his help, the tissue donors and supporting medical staff for making this study possible, C. Seki and A. Yukawa for their technical assistance, T. Mizuochi for discussion, Prof. S. G. Silverberg for his enthusiastic help with the pathologic review, and K. Boroevich for helpful comments on the manuscript.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 24, 2011; revised December 9, 2011; accepted December 23, 2011; published OnlineFirst January 12, 2012.

References

1. Cannistra SA. Cancer of the ovary. N Engl J Med 2004;351:2519–29.
2. Kurman RJ, Visvanathan K, Roden R, Wu TC, Shih IeM. Early detection and treatment of ovarian cancer: shifting from early stage to minimal volume of disease based on a new model of carcinogenesis. Am J Obstet Gynecol 2008;198:351–6.
3. Levanon K, Crum C, Drapkin R. New insights into the pathogenesis of serous ovarian cancer and its clinical impact. J Clin Oncol 2008;26:5284–93.
4. Bowtell DD. The genesis and evolution of high-grade serous ovarian cancer. Nat Rev Cancer 2010;10:803–8.
5. Winter WE 3rd, Maxwell GL, Tian C, Carlson JW, Ozols RF, Rose PG, et al. Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. J Clin Oncol 2007;25:3621–7.
6. du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). Cancer 2009;115:1234–44.
7. Dowsett M, Dunbier AK. Emerging biomarkers and new understanding of traditional markers in personalized therapy for breast cancer. Clin Cancer Res 2008;14:8019–26.
8. Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol 2010;220:263–80.
9. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, et al. Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. J Pathol 2010;221:49–56.
10. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 2011;474:609–15.
11. Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. Clin Cancer Res 2008;14:5198–208.
12. Bonome T, Levine DA, Shih J, Randonovich M, Pise-Masison CA, Bogomolniy F, et al. A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer. Cancer Res 2008;68:5478–86.
13. Dressman HK, Berchuck A, Chan G, Zhai J, Bild A, Sayer R, et al. An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. J Clin Oncol 2007;25:517–25.
14. Majewski IJ, Bernards R. Taming the dragon: genomic biomarkers to individualize the treatment of cancer. Nat Med 2011;17:304–12.
15. Heintz AP, Odicino F, Maisonneuve P, Beller U, Benedet JL, Creasman WT, et al. Carcinoma of the ovary. J Epidemiol Biostat 2001;6:107–38.
16. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al.: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228–47.

Published OnlineFirst January 12, 2012; DOI:10.1158/1078-0432.CCR-11-2725

17. Tavassoli FA, Devilee P. World Health Organization Classification of Tumors. Pathology & Genetics. Tumors of the breast and female genital organs. Lyon (France): IARC Press; 2003. p. 117–45.
18. Silverberg SG. Histopathologic grading of ovarian carcinoma: a review and proposal. Int J Gynecol Pathol 2000;19:7–15.
19. Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, et al. Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets. PLoS One 2010;5:e9615.
20. Glaab E, Garibaldi J, Krasnogor N. ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. BMC Bioinformatics 2009;10:358.
21. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. J Clin Oncol 2010;28:4111–9.
22. Simon N, Friedman J, Hastie T. Regularization paths for Cox's proportional hazards model via coordinate descent. J Stat Softw 2010;39:1–22.
23. R Development Core Team. R: A language and environment for statistical computing. 2011
24. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 2004;101:4164–9.
25. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat 2001;29:1165–88.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc B 1995;57:289–300.
28. Han LY, Fletcher MS, Urbauer DL, Mueller P, Landen CN, Kamat AA, et al. HLA class I antigen processing machinery component expression and intratumoral T-Cell infiltrate as independent prognostic markers in ovarian carcinoma. Clin Cancer Res 2008;14:3372–9.
29. Leffers N, Fehrmann RS, Gooden MJ, Schulze UR, Ten Hoor KA, Hollema H, et al. Identification of genes and pathways associated with cytotoxic T lymphocyte infiltration of serous ovarian cancer. Br J Cancer 2010;103:685–92.
30. Gooden MJ, de Bock GH, Leffers N, Daemen T, Nijman HW. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. Br J Cancer 2011;105:93–103.
31. Leffers N, Daemen T, Helfrich W, Boezen HM, Cohlen BJ, Melief K, et al. Antigen-specific active immunotherapy for ovarian cancer. Cochrane Database Syst Rev 2010;1:CD007287.
32. Kandalaft LE, Powell DJ Jr, Singh N, Coukos G. Immunotherapy for ovarian cancer: what's next? J Clin Oncol 2011;29:925–33.
33. Rolland P, Deen S, Scott I, Durrant L, Spendlove I. Human leukocyte antigen class I antigen expression is an independent prognostic factor in ovarian cancer. Clin Cancer Res 2007;13:3591–6.
34. Shehata M, Mukherjee A, Deen S, Al-Attar A, Durrant LG, Chan S. Human leukocyte antigen class I expression is an independent prognostic factor in advanced ovarian cancer resistant to first-line platinum chemotherapy. Br J Cancer 2009;101:1321–8.
35. Tanaka K, Hayashi H, Hamada C, Khoury G, Jay G. Expression of major histocompatibility complex class I antigens as a strategy for the potentiation of immune recognition of tumor cells. Proc Natl Acad Sci U S A 1986;83:8723–7.
36. Garrido F, Cabrera T, Aptsiauri N. "Hard" and "soft" lesions underlying the HLA class I alterations in cancer cells: implications for immunotherapy. Int J Cancer 2010;127:249–56.
37. Khong HT, Restifo NP. Natural selection of tumor variants in the generation of "tumor escape" phenotypes. Nat Immunol 2002;3:999–1005.
38. Dunn GP, Koebel CM, Schreiber RD. Interferons, immunity and cancer immunoediting. Nat Rev Immunol 2006;6:836–48.
39. Khan AN, Gregorie CJ, Tomasi TB. Histone deacetylase inhibitors induce TAP, LMP, Tapasin genes and MHC class I antigen presentation by melanoma cells. Cancer Immunol Immunother 2008;57:647–54.
40. Setiadi AF, Omilusik K, David MD, Seipp RP, Hartikainen J, Gopaul R, et al. Epigenetic enhancement of antigen processing and presentation promotes immune recognition of tumors. Cancer Res 2008;68:9601–7.
41. Kitamura H, Torigoe T, Asanuma H, Honma I, Sato N, Tsukamoto T. Down-regulation of HLA class I antigens in prostate cancer tissues and up-regulation by histone deacetylase inhibition. J Urol 2007;178:692–6.
42. Chobanian NH, Greenberg VL, Gass JM, Desimone CP, Van Nagell JR, Zimmer SG. Histone deacetylase inhibitors enhance paclitaxel-induced cell death in ovarian cancer cell lines independent of p53 status. Anticancer Res 2004;24:539–45.
43. Qian X, LaRochelle WJ, Ara G, Wu F, Petersen KD, Thougaard A, et al. Activity of PXD101, a histone deacetylase inhibitor, in preclinical ovarian cancer studies. Mol Cancer Ther 2006;5:2086–95.

MODERN PATHOLOGY (2012) 25, 615–624

Loss of ARID1A protein expression occurs as an early event in ovarian clear-cell carcinoma development and frequently coexists with *PIK3CA* mutations

Sohei Yamamoto[1], Hitoshi Tsuda[2], Masashi Takano[3], Seiichi Tamai[4] and Osamu Matsubara[1]

[1]*Department of Basic Pathology, National Defense Medical College, Tokorozawa, Japan;* [2]*Department of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tokyo, Japan;* [3]*Department of Obstetrics and Gynecology, National Defense Medical College, Tokorozawa, Japan and* [4]*Department of Laboratory Medicine, National Defense Medical College Hospital, Tokorozawa, Japan*

ARID1A **is a recently identified tumor suppressor gene that is mutated in ~50% of ovarian clear-cell carcinomas. This mutation is associated with loss of ARID1A protein expression as assessed by immunohistochemistry. The present study aimed at determining the timing of the loss of ARID1A protein expression during the development of ovarian clear-cell carcinoma and assessing its relevance in correlation to *PIK3CA* gene mutations. A total of 42 clear-cell carcinoma cases with adjacent putative precursor lesions (endometriosis-associated carcinoma cases ($n=28$) and (clear-cell) adenofibroma-associated carcinoma cases ($n=14$)) were selected and subjected to immunohistochemical analysis for ARID1A protein expression and direct genomic DNA sequencing of exons 9 and 20 of the *PIK3CA* gene. ARID1A immunoreactivity was deficient in 17 (61%) of the 28 endometriosis-associated carcinomas and 6 (43%) of the 14 adenofibroma-associated carcinomas. Among the precursor lesions adjacent to the 23 ARID1A-deficient carcinomas, 86% of the non-atypical endometriosis (12 of 14) and 100% of the atypical endometriosis (14 of 14), benign (3 of 3), and borderline (6 of 6) clear-cell adenofibroma components were found to be ARID1A deficient. In contrast, in the 19 patients with ARID1A-intact carcinomas, all of the adjacent precursor lesions retained ARID1A expression regardless of their types and cytological atypia. Analysis of 22 solitary endometrioses and 10 endometrioses distant from ARID1A-deficient carcinomas showed that all of these lesions were diffusely immunoreactive for ARID1A. Among the 42 clear-cell carcinomas, somatic mutations of *PIK3CA* were detected in 17 (40%) tumors and majority (71%) of these were ARID1A-deficient carcinomas. These results suggest that loss of ARID1A protein expression occurs as a very early event in ovarian clear-cell carcinoma development, similar to the pattern of *PIK3CA* mutation recently reported by our group, and frequently coexists (not mutually exclusive) with *PIK3CA* mutations.**

Modern Pathology (2012) **25,** 615–624; doi:10.1038/modpathol.2011.189; published online 9 December 2011

Keywords: ARID1A; clear-cell adenofibroma; clear-cell carcinoma; endometriosis; PIK3CA

Among ovarian carcinomas, clear-cell carcinoma has been recognized as a distinct clinicopathological entity in view of its characteristic histology, frequent concurrence of endometriotic lesions, and highly chemoresistant nature, resulting in an extremely poor prognosis when surgical cytoreduction is insufficient.[1–4] Although little is known about the molecular genetic alterations underlying tumor development, a hypothesis of multistep tumorigenesis of ovarian clear-cell carcinoma, starting with histologically benign-appearing precursor lesions

Correspondence: Dr S Yamamoto, MD, Department of Basic Pathology, National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan.
E-mail: dr21001@ndmc.ac.jp
Received 19 July 2011; revised 5 September 2011; accepted 21 September 2011; published online 9 December 2011

(ie, endometriosis and benign clear-cell adenofibroma), progressing to their atypical counterparts (ie, atypical endometriosis and borderline clear-cell adenofibroma), and ultimately to clear-cell carcinoma, has been proposed and widely considered.[5–11] To establish novel therapeutic strategies for ovarian clear-cell carcinoma, it is crucial to elucidate the molecular aberrations that are characteristic of this type of carcinoma during tumor development.

Although little is known about the pathogenesis of ovarian clear-cell carcinoma, somatic mutations of the *ARID1A* (the AT-rich interactive domain 1A) and *PIK3CA* genes are the most common molecular genetic alterations identified thus far in ovarian clear-cell carcinomas.[12–15] ARID1A (also known as BAF250a), a protein encoded by the *ARID1A* gene located on chromosome 1p36, is a key component of the multi-protein SWI/SNF chromatin-remodeling complex present in all eukaryotes and is believed to confer specificity in the regulation of gene expression.[16] The chromatin-remodeling activity of SWI/SNF has an integral role in controlling gene expression and is critical in tissue development, cellular differentiation, and tumor suppression.[16–20] Inactivation of *ARID1A* is thought to enhance cell-cycle progression by potentially involving c-myc, thereby contributing to uncontrolled cellular proliferation in cancer cells.[21–23] Recently, with the use of whole-exome sequencing and transcriptome sequencing, two independent studies showed novel somatic mutations of the *ARID1A* gene in 43 and 56% of ovarian clear-cell carcinomas, suggesting a major role of *ARID1A* in the pathogenesis of this carcinoma type.[12,13] However, to date, it is unclear whether the alteration of ARID1A is an early or late event in the development and progression of ovarian clear-cell carcinoma.

The *PIK3CA* gene, encoding the catalytic subunit p110a of phosphatidylinositol-3 kinases (PI3K), is located on chromosome 3q26.3, and both somatic mutations and gene amplification of *PIK3CA* increase PI3K activity and activate the downstream Akt signaling pathway.[24–26] Somatic activating mutations of *PIK3CA* have been reported in ovarian clear-cell carcinomas with high frequencies of 32–46%.[14,15] Our recent study showed that somatic mutations of *PIK3CA* were also present in the coexisting endometriotic epithelium, adjacent to the PIK3CA-mutation-positive clear-cell carcinomas, even in those lacking cytological atypia.[15] These findings suggest that mutations of the *PIK3CA* gene occur before the development of the atypical precancerous lesions (ie, atypical endometriosis) and that *PIK3CA* mutations are very early events in the development of endometriosis-associated ovarian clear-cell carcinoma. It is therefore of interest to determine whether *PIK3CA* mutations and ARID1A could coexist or be mutually exclusive in the pathogenesis of ovarian clear-cell carcinoma.

In the present study, the timing of the alterations of ARID1A expression during the development and progression of ovarian clear-cell carcinoma was examined. For this purpose, 42 patients with clear-cell carcinoma belonging to a group of tumors with adjacent putative precursor lesions were selected. As the immunohistochemical analysis of clear-cell and endometrioid carcinomas of the ovary showed that loss of ARID1A protein expression was highly correlated with the presence of *ARID1A* mutations,[12,27] and as the *ARID1A* mutations are randomly distributed in 20 exons and most of these are the insertion/deletion type of mutations, making it very difficult to be detected in formalin-fixed paraffin-embedded specimens,[12,13] we used immunohistochemistry to assess possible alterations of ARID1A. ARID1A expression was specifically examined in the benign-appearing putative precursor lesions (ie, non-atypical endometriosis and benign clear-cell adenofibroma), their atypical forms (atypical endometriosis and borderline clear-cell adenofibroma), and the corresponding invasive carcinoma components. Moreover, to assess the possible relationships between *ARID1A* and *PIK3CA* alterations, immunoreactivity for ARID1A was compared with the mutational status of *PIK3CA*. In ovarian carcinomas, >80% of mutations in *PIK3CA* have been found in the helical (exon 9) and kinase domains (exon 20).[14,24,25] Therefore, we restricted our analyses to these regions of the gene.

Materials and methods

Cases and Histological Components Analyzed

Hematoxylin- and eosin-stained sections from a consecutive series of 90 primary clear-cell adenocarcinoma cases, retrieved from the files of the Department of Laboratory Medicine; National Defense Medical College Hospital; Japan, were histologically reviewed. On the basis of the histopathological criteria described previously,[5,7–9] of the 90 patients, 28 patients with synchronous endometriotic lesions (endometriosis-associated carcinomas) and 14 patients with adjacent clear-cell adenofibroma components (adenofibroma-associated carcinomas) were identified. All these 42 patients had undergone surgical resection between 1986 and 2007, and none had undergone chemotherapy or radiation therapy before surgery. All specimens analyzed were formalin-fixed and paraffin-embedded tissue sections. There was no overlap between the endometriosis-associated and adenofibroma-associated carcinoma cases. The diseases were clinically staged according to the International Federation of Gynecology and Obstetrics system. Of the 28 patients with endometriosis-associated carcinoma, 18 (64%) had stage I, 3 (11%) had stage II, 5 (18%) had stage III, and 2 (7%) had stage IV carcinomas. Of the 14 patients