Our findings lend support to prior work that has suggested the importance of the NER pathway in response to platinum-based chemotherapy and suggests *ERCC5* (XPG) as a novel candidate biomarker of ovarian cancer response to platinum chemotherapy. Further work on this pathway may validate XPG as a diagnostic marker and/or lead to the development of a therapeutic agent that specifically targets XPG activity for the sensitization of platinum-resistant malignancies.

PATIENTS AND METHODS

Patient Samples

All patient samples were collected at Cedars-Sinai Medical Center using protocols approved by the Cedars-Sinai institutional review board. After patients provided informed consent, fresh tumor tissue was snap-frozen in liquid nitrogen and stored in a −80°C freezer. Study participants were treated with surgical cytoreduction followed by platinum-based chemotherapy regimen and had corresponding clinical and follow-up information.

DNA was isolated from frozen tumor samples using the Qiagen DNeasy Tissue Protocol (QIAGEN, Valencia, CA). RNA was isolated from frozen tumor samples using the RNeasy kit (QIAGEN). Genomic DNA was isolated from matched peripheral-blood samples using a standard phenol chloroform extraction method. DNA and RNA quantities were measured with the Nanodrop spectrophotometer (Thermo Scientific, Wilmington, DE). RNA quality was determined through separation by capillary electrophoresis on the Agilent 2000 Bioanalyzer (Agilent Technologies, Foster City, CA). Microarray analysis was performed with high-quality RNA, defined as an RNA integrity number greater than 8.

DNA Genomic Analysis Using GeneChip Mapping 50K High-Density Oligonucleotide Array

Array experiments were performed using standard Affymetrix protocols (Affymetrix, Santa Clara, CA). For each sample, 250 ng of total genomic DNA was digested with *XbaI* restriction enzyme and ligated to common adaptors, which allowed for polymerase chain reaction (PCR) amplification of the entire genome using a single pair of primers. PCR products were digested, labeled, and hybridized to the 50K *Xba* high-density oligonucleotide microarray, which contains a marker distance of approximately 50 kb between single-nucleotide polymorphisms.

GeneChip Data Analysis

Genotype calls were processed using the Copy Number Analyzer for GeneChip software program (CNAG; University of Tokyo, Tokyo, Japan; http://www.genome.umin.ac.jp).[6] Matched peripheral-blood DNA was used as the reference for each tumor DNA sample. Each genomic region was classified as normal, amplified, deleted, or as a region of copy-neutral LOH (also known as uniparental disomy). Summary graphs were generated with STATA v8 (STATA Corp, College Station, TX) to graphically display the range and locations of each genetic abnormality.

A discrete variable was created for each genomic block affected by LOH in at least 20% of cases. In all, 106 LOH variables were defined across the tumor genome and each was tested as a predictor affecting PFS. Cox proportional hazards methods were used to determine hazard ratios (HR) for each variable. Associations were reported if the two-sided *P* values were less than .05. Multivariate Cox models were generated to control for the effects of confounding factors. As this was an exploratory analysis, corrections were not made for multiple testing.

Identification of Candidate Genes

Genomic LOH regions were found to predict prolonged PFS. These regions were hypothesized to harbor important tumor suppressor genes that mediate response to platinum-based chemotherapy. Candidate genes were identified through a direct link from the CNAG chromosome viewer to the University of California at Santa Cruz Genome Browser Build 17 (hg17) Human May 2004 assembly, based on the National Center for Biotechnology Information Build 35 (University of California at Santa Cruz, Santa Cruz, CA; http://genome.ucsc.edu/ index.html). Predicted and reference sequence genes were investigated to narrow the candidate gene list.

LOH Analysis at Candidate Gene Locus

ERCC5, a DNA repair gene from the NER pathway, was identified as a biologically plausible modulator of ovarian cancer response to platinum-based chemotherapy. The LOH status at the *ERCC5* gene locus was evaluated in 44 clinical ovarian cancer samples to determine its specific effect on ovarian cancer prognosis. All were papillary serous histology, stages IIC to IV; 10 were common to the GeneChip data set. For each sample, a dinucleotide repeat polymorphism located within an intron of the *ERCC5* gene was amplified in tumor and matched normal DNA. Previously published primer sequences[7] were lengthened to enhance the specificity of the reaction (Forward-Fam labeled 5'GCA ATG ACT CGG TAT TGG CTA AT 3'; Reverse 5' GAT GCT AAC AAG TGG GTG GAA T 3'). PCR was performed with 15-μL reactions containing 20 ng of genomic DNA, 20 pmole of each primer, 0.75 μL of AmpliTaq DNA polymerase (Applied Biosystems, Foster City, CA), and 2× dNTP (Epicenter, Madison, WI) on a GeneAmp PCR System 9700 (Applied Biosystems). PCR products were verified on 2% agarose gel and run in the University of California at Los Angeles genotyping core facility. LOH was determined in samples with a homozygous genotype in tumor DNA and a heterozygous genotype in the matched normal DNA.

Gene Expression Analysis

The significance of expression of *ERCC5* and other *ERCC* gene expression were tested using high-quality RNA extracted from 90 ovarian cancer samples. Microarray analysis was performed on the Agilent Human 1A V2 chip. Samples were labeled using the Agilent Low RNA Input Fluorescent Linear Amplification Kit. Each individual sample was labeled with cyanine-5 and characterized by comparison to a mixed reference pool labeled with cyanine-3. The mixed reference pool consisted of equal amounts of cRNA from 106 clinical samples from the tumor bank, ranging from benign to malignant, and chosen to be representative of the range of histologic pathologies occurring in the female reproductive tract. An Agilent Scanner and the Agilent Feature Extraction software version 7.5 were used to read the microarray slides and calculate gene expression values. Gene expression values of the candidate genes (*ERCC1-ERCC6, ERCC8*) were exported to STATA and linked to clinical data. The threshold for downregulation was defined as log(ratio) less than 0 with a *P* value of less than .05. The *P* values were determined according to the Agilent error model with the feature-extracted data imported into Resolver. Cox proportional hazards models, Kaplan-Meier survival curves, and the log-rank test were used to determine the effects of *ERCC* gene expression on PFS.

Quantitative PCR Validation of Gene Expression

Microarray gene expression data were validated by quantitative real-time PCR performed using the SYBR Green method (Invitrogen, Carsbad, CA) with standard curves on iCycler (Bio-Rad, Hercules, CA). RNA was converted to cDNA using the QuaniTect Reverse Transcription kit (QIAGEN). PCR reactions were performed in 96-well plates with 12.5 μL 0f SYBR Green, 0.5 μL of primer, and cDNA in a total reaction volume of 25 μL. *ERCC5* was amplified with the following primers: Forward CAGACACAGCTCCGAATTGA; Reverse TTCTGGGTTTTTCGTTTTGC. Expression of ERCC5 was normalized by 18srRNA subunit expression with the following primers: Forward CGCCGTGCCTACCATGGTGAC; Reverse CTTGGATGTGGTAGCCGTTTCTCA.

RESULTS

Study Overview and Patient Characteristics

A whole-genome analysis of the genetic changes occurring in the DNA of 20 tumors was performed to identify candidate regions that predicted improved response to treatment. Effect of LOH at the candidate gene locus was analyzed in 52 total samples. Association of RNA expression levels with prognosis was tested in 90

tumors. Patient characteristics among the three data sets are listed in Table 1. All patients were treated with initial cytoreductive surgery followed by adjuvant chemotherapy with a platinum-containing regimen. PFS was defined as the time from date of primary cytoreductive surgery to the date of first clinical evidence of recurrence. Overall survival was defined as the time from date of primary cytoreductive surgery to the date of death or censored at the date of last follow-up.

Characterization of Genomic Abnormalities in Ovarian Cancer DNA

Appendix Figure A1A (online only) is a representative example of how the visual output from the CNAG software program was converted into a color-coded graphical representation of genetic changes (amplification, deletion, or uniparental disomy). This method allowed us to summarize the genetic heterogeneity and complexity occurring over the 20 ovarian cancer samples (Appendix Fig A1B).

Table 1. Patient Characteristics at Diagnosis

Characteristic	Whole Genome Analysis (n = 20)		LOH Analysis (n = 54)		RNA Expression Analysis (n = 90)	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
Age, years						
Median	62		58		59	
Range	45-72		43-82		41-82	
PFS, months						
Median	14		20		14	
Range	0.1-39.2		2.9-112.6		2.9-192.3	
OS, months						
Median	42.9		45		38.3	
Range	11.5-90.5		5.3-116.4		5.3-192.3	
Stage						
I	0	0	0	0	5	5
II	2	10	3	6	6	7
III	14	70	37	68	63	70
IV	4	20	14	26	16	18
Grade						
2	2	10	3	6	9	10
3	18	90	51	94	81	90
Histology						
Papillary serous	20	100	54	100	81	90
Other epithelial	0	0	0	0	9	10
Tumor site						
Ovarian	15	75	43	79	77	86
Fallopian tube	0	0	2	4	2	2
Primary peritoneal	5	5	9	17	11	12
BRCA mutation status						
BRCA1 carrier	3	15	8	15	10	11
BRCA2 carrier	2	10	6	11	8	9
BRCA1 and *BRCA2*	1	5	1	2	1	1
Negative/unknown	14	70	39	72	71	79
Cytoreductive surgery						
Optimal (< 1 cm)	19	95	45	83	72	80
Suboptimal (≥ 1 cm)	1	5	8	15	10	11
Unknown	0	0	1	2	8	9
Adjuvant chemotherapy						
Carboplatin/paclitaxel	18	90	48	89	67	74
Cisplatin/cyclophosphamide	2	10	2	4	18	20
Other platinum regimen	0	0	4	7	5	6
Clinical response						
Complete response	17	85	42	78	72	80
Partial response/stable	3	15	11	20	17	19
Progression	0	0	1	2	1	1
Pathologic response						
Second look negative (pCR)	8	40	14	26	25	28
Second look positive	9	45	15	28	24	27
Not evaluated	3	15	24	46		

Abbreviations: LOH, loss of heterozygosity; PFS, progression-free survival; OS, overall survival; pCR, pathologic complete response.

A similar summary graph demonstrates the frequency of LOH in the ovarian tumor genomes (Appendix Fig A1C). The LOH data were further summarized in a frequency plot demonstrating the proportion of cases affected by LOH at each genomic locus (Appendix Fig A1D).

Predictors of Prolonged PFS

LOH blocks on chromosomes 2, 5, 13, and 22 were found to predict prolonged PFS (point estimates for HRs ranging from 0.23 to 0.36; Table 2), suggesting a survival benefit owing to loss of function of genes in the regions. Adjusted HRs were calculated for each LOH block to account for the potential confounding effects of various clinical variables. LOH of the 13q (54 to 102 MB) block was found to retain a highly significant independent association with prolonged PFS after adjustment for various confounders (Table 2).

A Candidate Gene That May Predict Response to Platinum Chemotherapy

The 48-MB region on 13q (54 to 102 MB) contains 73 genes (Appendix Table A1, online only), including *ERCC5*. We felt this gene deserved further study given its role in the NER pathway, the unique DNA repair pathway that allows cells to remove platinum adducts from DNA. We hypothesized that LOH of this 13q region leads to downregulation of *ERCC5* levels, a diminished capacity of tumor cells to recover from platinum-based chemotherapy (enhanced chemotherapy sensitivity), and prolonged PFS.

Effect of LOH of ERCC Gene Locus on Survival

The LOH status of the *ERCC5* gene locus was determined through genotyping of a dinucleotide repeat polymorphism (DRP) in 44 samples. Fourteen genotypes (32%) were noninformative because of a homozygous genotype in both normal and tumor DNA. Genotype data were available from 20 samples by GeneChip analysis. Ten samples were genotyped by both methods, resulting in 54 total samples genotyped. Among the 10 overlapping samples, six genotypes were correlated, two were noninformative on DRP analysis, and two were contradictory on the basis of the finding of LOH on GeneChip but no LOH on DRP analysis. In the four noncorrelating samples, LOH status was analyzed based on GeneChip data as a result of the self-validating nature of the genotyping of multiple single-nucleotide polymorphisms with the whole-genome approach.

Among the 54 samples genotyped, 29 (54%) had no LOH, 13 (24%) had LOH, and 12 (22%) were noninformative. LOH at the *ERCC5* locus (13 of 42) demonstrated a trend toward improved PFS (Fig 1A). A subset analysis limited to the stage IIC to IIIC papillary serous tumors (nine of 29) demonstrated a significant improvement in PFS (Fig 1B).

At the time of second-look surgery, patients with LOH at the *ERCC5* locus were significantly more likely to have a pathologic complete response (five of six; 83%) than those without LOH (two of eight; 25%; $P = .03$).

Effect of ERCC Gene Expression on Survival

Gene expression levels of the various *ERCC* genes (*ERCC1* through *ERCC6*, *ERCC8*) were analyzed for influences on survival in a data set of 90 patients. PFS (Fig 1C) and OS (Fig 1D) were both prolonged among patients with *ERCC5* downregulation in the tumors. This effect was not seen for *ERCC1*, *ERCC2*, *ERCC3*, *ERCC4*, *ERCC6*, or *ERCC8* downregulation (Appendix Figs A2 through A7, online only). Microarray gene expression analysis was validated using quantitative real-time PCR (Fig 2).

Table 3 demonstrates the results from Cox proportional hazards regression models. On univariate analysis, a beneficial impact on PFS is seen with *ERCC5* downregulation (HR = 0.44; $P = .01$), but not with downregulation of any other single *ERCC* gene. On multivariate analysis, *ERCC5* downregulation retains an independent beneficial impact on PFS (HR = 0.49; $P = .03$).

Correlation Between ERCC5 LOH and Gene Expression Levels

Forty tumors with DNA and RNA data were analyzed for correlation between LOH of the 13q (*ERCC5*) locus and *ERCC5* expression

Table 2. Genomic LOH Loci Correlating With Significantly Improved PFS

Variable	n*	Crude HR	Crude 95% CI	Crude P	Adjusted† HR	Adjusted† 95% CI	Adjusted† P	Adjusted‡ HR	Adjusted‡ 95% CI	Adjusted‡ P
Genomic LOH loci										
2p: 100,819-5,524,537	5	0.28	0.08 to 0.99	.04	0.40	0.10 to 1.55	.19	0.29	0.08 to 1.02	.06
5q: 57,385,834-63,071,109	8	0.30	0.12 to 0.88	.03	0.27	0.07 to 1.00	.05	0.31	0.10 to 0.93	.03
5q: 73,164,783-81,608,220	10	0.32	0.12 to 0.88	.03	0.35	0.11 to 1.11	.08	0.34	0.12 to 0.95	.04
5q: 81,608,220-95,277,667	9	0.36	0.13 to 1.00	.05	0.36	0.11 to 1.15	.09	0.38	0.13 to 1.07	.06
13q: 18,042,610-54,217,914	8	0.24	0.08 to 0.71	.01	0.29	0.08 to 1.06	.06	0.25	0.08 to 0.75	.01
13q: 54,217,914-102,557,751	9	0.23	0.08 to 0.65	.006	0.24	0.07 to 0.88	.03	0.24	0.08 to 0.69	.008
13q: 102,557,751-114,051,456	8	0.24	0.08 to 0.71	.01	0.29	0.08 to 1.06	.06	0.25	0.08 to 0.75	.01
22q: 26,083,853-33,362,998	8	0.29	0.10 to 0.86	.03	0.35	0.10 to 1.24	.1	0.30	0.10 to 0.92	.04
22q: 44,485,638-48,983,486	10	0.31	0.12 to 0.86	.03	0.46	0.13 to 1.57	.2	0.34	0.12 to 0.93	.04
Clinical factors										
Optimal cytoreduction		0.05	0.003 to 0.86	.04						
Stage		1.64	0.80 to 3.35	.18						
Germline *BRCA* mutation		0.55	0.28 to 1.08	.09						

Abbreviations: LOH, loss of heterozygosity; PFS, progression-free survival; HR, hazard ratio.
*n represents number of cases of 20 with LOH of the genomic region.
†Adjusted for stage, optimal cytoreduction, and germline *BCRA* mutation status.
‡Adjusted for optimal cytoreduction.

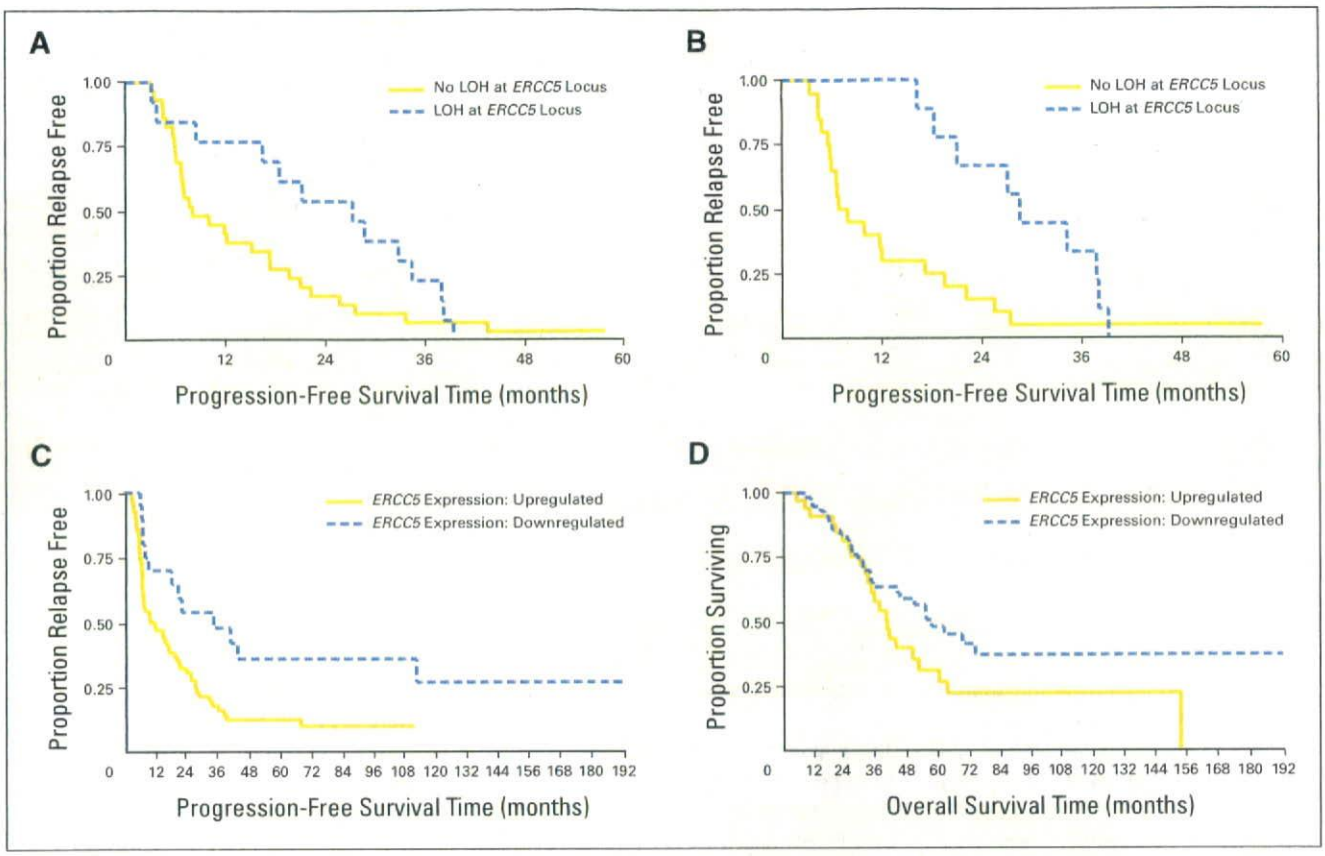

Fig 1. Survival differences based on *ERCC5* loss of heterozygosity (LOH) and gene expression. (A) Trend toward improved progression-free survival (PFS) among patients with LOH of *ERCC5* gene locus ($P = .1$) and (B) improved PFS in the subset of patients with stage IIC to IIIC disease ($P = .01$). (C) Improved PFS ($P = .004$) and (D) overall survival ($P = .08$) among patients with downregulation of *ERCC5* gene expression.

levels. Expression levels are lower in the group of tumors with LOH (mean fold change, −0.59; median −1.12) compared with the tumors without LOH (mean fold change, +0.84; median, +1.03; $P = .08$). This suggests possible biologic validity of *ERCC5* as a target gene within the 13q region.

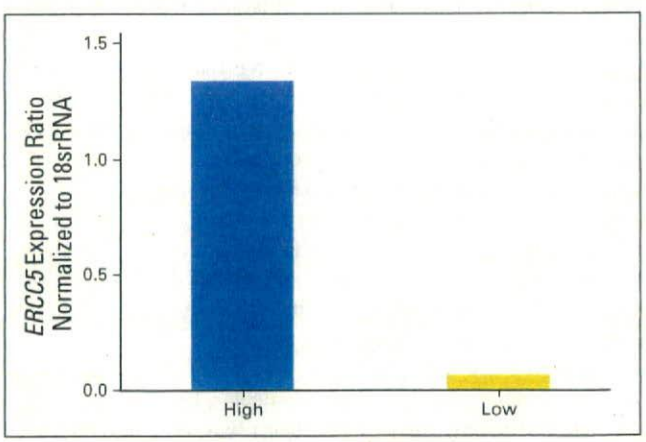

Fig 2. Quantitative polymerase chain reaction (PCR) validation of *ERCC5* gene expression. Mean differential quantitative PCR gene expression between cases with high (n = 11) and low (n = 8) *ERCC5* expression on microarray analysis ($P = .06$.).

DISCUSSION

Our study identifies *ERCC5* as a novel candidate biomarker of ovarian cancer sensitivity to platinum chemotherapy. This conclusion is supported on several different levels. We found LOH of the 13q locus containing *ERCC5* to predict prolonged PFS among platinum-treated patients with ovarian cancer. Additional genotyping also confirmed an association between LOH of the *ERCC5* gene locus and improved survival. Furthermore, downregulation of *ERCC5* mRNA expression levels also predicted prolonged PFS in an independent data set. Finally, the correlation between LOH of the *ERCC5* genomic locus with downregulation of *ERCC5* mRNA levels among the subset of tumors with integrated genomic and gene expression data suggests possible biologic plausibility of *ERCC5* being a target gene in the 13q LOH region.

Further biologic plausibility is apparent when placing these findings in the context of previous knowledge and work. Defective DNA repair pathways allow tolerance to DNA damage, permitting an accelerated rate of mutagenesis and neoplastic transformation. This characteristic turns to the disadvantage of the cancer cell when DNA damaging cancer therapies are administered, leading to an enhanced response to treatment.[8] We have found evidence to support the hypothesis that a downregulation of *ERCC5* activity leads to enhanced platinum chemotherapy sensitivity in ovarian

Table 3. Cox Proportional Hazards Regression: Impact of Clinical and *ERCC* Gene Expression Data on Progression-Free Survival

Variable	No. of Patients	No. of Relapses	HR	95% CI	P
Univariate analysis					
BRCA mutation carrier	18	15	0.63	0.36 to 1.12	.12
Optimal cytoreduction	72	56	0.53	0.30 to 0.92	.02
Late stage, III/IV	79	70	5.92	1.86 to 18.86	.003
Serous histology	81	70	2.88	0.91 to 9.16	.07
ERCC1 downregulation	12	12	0.93	0.50 to 1.74	.84
ERCC2 downregulation	23	18	0.73	0.43 to 1.25	.2
ERCC3 downregulation	16	13	0.83	0.45 to 1.51	.5
ERCC4 downregulation	0				
ERCC5 downregulation	20	13	0.44	0.24 to 0.82	.01
ERCC6 downregulation	5	4	0.71	0.26 to 1.97	.5
ERCC8 downregulation	3	3	2.17	0.67 to 6.98	.2
Multivariate analysis					
ERCC5 downregulation			0.49	0.25 to 0.92	.03
Late stage, III/IV			5.61	1.76 to 17.9	.004
Optimal cytoreduction			0.66	0.38 to 1.16	.1

Abbreviation: HR, hazard ratio.

cancer. This is provocative when considering the function of *ERCC5* in the NER pathway, the unique DNA repair pathway that repairs DNA damage caused by platinum agents.

NER recognizes and repairs bulky, helix-distorting adducts, such as those formed by cisplatin and its analogs.[9,10] A complex of proteins assembles, binds bulky DNA damage, incises the oligonucleotide fragment containing the damaged base, and fills in the resulting gap.[11,12] Platinum-resistant cells are able to more effectively remove cisplatin-DNA adducts through the action of a functional NER pathway and thus escape apoptosis.

ERCC5 (XPG) is a structure-specific endonuclease, which participates in two incision steps that are critical to the DNA repair process. XPG cleaves the damaged DNA 3′ to the damaged site, nonenzymatically participates in the 5′ incision mediated by the XPF/ERCC1 heterodimer, and stabilizes the DNA repair complex to the damaged DNA.[13-16] XPG is critical to both subpathways of NER: transcription-coupled repair (TCR), which specifically targets and repairs DNA damage on the transcribed strand of actively expressed genes, and global genomic repair (GGR), which removes DNA damage from the remaining genome.[17] TCR and GGR each have a unique mechanism for recognizing DNA damage, then progress along a common pathway that requires XPG.[11,12]

A number of studies provide evidence for the role of the NER pathway in cellular response to platinum chemotherapy, consistently demonstrating platinum-resistance with enhanced NER activity and platinum-sensitivity with diminished NER activity. Cisplatin-resistant cells have been shown to have increased levels of *XPA* mRNA,[18] overexpression of *ERCC1* or *ERCC1/XPF*,[19] increased activity of *ERCC1/XPD*,[20] and increased *XPC* and *ERCC1* levels.[21] Hypersensitivity of some cell lines may be related to reduced expression of *XPG* or *XPA*.[9] Inhibition of *ERCC1* activity with antisense oligonucleotides enhances cisplatin sensitivity in ovarian cancer cell lines.[22] Cells with deficiencies in GGR-specific factors (XPC) display normal resistance to cisplatin, whereas cells with deficiencies in TCR-specific (CSA, CSB) or common pathway proteins (XPA, XPD, XPF, XPG) are markedly hypersensitive to cisplatin.[23] The excellent response rates of testicular cancer to cisplatin may be due to a high rate of NER deficiencies.[24]

Our study adds to this body of literature, suggesting that loss of *ERCC5* function occurs naturally during the carcinogenic process of a subset of ovarian cancers and consequently leads to inherent platinum chemotherapy sensitivity. This speculation is supported by the finding of LOH of regions harboring NER genes occurring at a higher frequency in ovarian cancers (62%) than in other solid tumors, such as colon or lung cancer.[25] Ours is not the first study to report frequent LOH of the 13q locus in epithelial ovarian cancer.[26]

Our findings have important prognostic and therapeutic implications. Tumors with dysfunctional *ERCC5* expression would be predicted to demonstrate sensitivity to platinum-based therapy. *ERCC5* (XPG) may be an appropriate target for therapeutic inhibition in platinum-resistant ovarian cancers. XPG is a critical component of the rate-limiting damage recognition/excision step of NER[27] and is expressed at lower cellular protein levels than other NER factors.[28] XPG levels are correlated with cytotoxicity to cisplatin and irofulven and with cellular NER activity,[28] potentially making it an attractive therapeutic target. A recent integrated analysis of array CGH and gene expression profiling data in testicular cancers also found *ERCC5* to be both lost and downregulated.[29] Ovarian cancer and testicular cancer share the quality of platinum chemotherapy sensitivity, and data are emerging to suggest that NER dysfunction (particularly *ERCC5*) may be another shared characteristic. Further work may lead to the development of a specific XPG inhibitor that can sensitize platinum-resistant tumors to the effects of platinum chemotherapy.

Platinum drugs demonstrate activity in a wide range of tumors, including ovarian, cervical, testicular, head and neck, and non–small-cell lung cancer,[30] but their use is often limited by the development of resistance. A number of complex pathways are involved, including decreased drug uptake into the cell, increased drug inactivation, and increased DNA repair.[31] Further insights

into these mechanisms could be used to develop rational biologic therapies that target platinum resistance.[32]

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Christine S. Walsh, Dennis J. Slamon, H. Phillip Koeffler, Beth Y. Karlan

Financial support: Seishi Ogawa, Dennis J. Slamon, H. Phillip Koeffler, Beth Y. Karlan

Provision of study materials or patients: Beth Y. Karlan

Collection and assembly of data: Christine S. Walsh, Hisae Karahashi, Daniel R. Scoles, James C. Pavelka, Hang Tran, Charles Ginther, Judy Dering, Masashi Sanada, Yasuhito Nannya

Data analysis and interpretation: Christine S. Walsh, Seishi Ogawa, Carl W. Miller, Norihiko Kawamata

Manuscript writing: Christine S. Walsh

Final approval of manuscript: Christine S. Walsh, Seishi Ogawa, Hisae Karahashi, Daniel R. Scoles, James C. Pavelka, Carl W. Miller, Norihiko Kawamata, Charles Ginther, Judy Dering, Masashi Sanada, Yasuhito Nannya, Dennis J. Slamon, H. Phillip Koeffler, Beth Y. Karlan

REFERENCES

1. McGuire WP 3rd, Markman M: Primary ovarian cancer chemotherapy: Current standards of care. Br J Cancer 89:S3-S8, 2003 (suppl 3)
2. Högberg T, Glimelius B, Nygren P: A systematic overview of chemotherapy effects in ovarian cancer. Acta Oncol 40:340-360, 2001
3. Piccart MJ, Lamb H, Vermorken JB: Current and future potential roles of the platinum drugs in the treatment of ovarian cancer. Ann Oncol 12:1195-1203, 2001
4. Harries M, Kaye SB: Recent advances in the treatment of epithelial ovarian cancer. Expert Opin Investig Drugs 10:1715-1724, 2001
5. Giaccone G: Clinical perspectives on platinum resistance. Drugs 59:9-17, 2000 (suppl 4)
6. Nannya Y, Sanada M, Nakazaki K, et al: A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res 65:6071-6079, 2005
7. Samec S, Clarkson SG, Blaschak J, et al: Dinucleotide repeat polymorphism within ERCC5 gene. Hum Mol Genet 3:214, 1994
8. Kennedy RD, D'Andrea AD: DNA repair pathways in clinical practice: Lessons from pediatric cancer susceptibility syndromes. J Clin Oncol 24: 3799-3808, 2006
9. Kelland LR: Preclinical perspectives on platinum resistance. Drugs 59:1-8, 2000 (suppl 4)
10. Kartalou M, Essigmann JM: Mechanisms of resistance to cisplatin. Mutat Res 478:23-43, 2001
11. Friedberg EC: How nucleotide excision repair protects against cancer. Nat Rev Cancer 1:22-33, 2001
12. Park CJ, Choi BS: The protein shuffle: Sequential interactions among components of the human nucleotide excision repair pathway. Febs J 273:1600-1608, 2006
13. Dunand-Sauthier I, Hohl M, Thorel F, et al: The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. J Biol Chem 280:7030-7037, 2005
14. O'Donovan A, Davies AA, Moggs JG, et al: XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. Nature 371:432-435, 1994
15. Wakasugi M, Reardon JT, Sancar A: The noncatalytic function of XPG protein during dual incision in human nucleotide excision repair. J Biol Chem 272:16030-16034, 1997
16. Constantinou A, Gunz D, Evans E, et al: Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair. J Biol Chem 274:5637-5648, 1999
17. Hanawalt PC: Controlling the efficiency of excision repair. Mutat Res 485:3-13, 2001
18. States JC, Reed E: Enhanced XPA mRNA levels in cisplatin-resistant human ovarian cancer are not associated with XPA mutations or gene amplification. Cancer Lett 108:233-237, 1996
19. Ferry KV, Hamilton TC, Johnson SW: Increased nucleotide excision repair in cisplatin-resistant ovarian cancer cells: Role of ERCC1-XPF. Biochem Pharmacol 60:1305-1313, 2000
20. Yu JJ, Bicher A, Ma YK, et al: Absence of evidence for allelic loss or allelic gain for ERCC1 or for XPD in human ovarian cancer cells and tissues. Cancer Lett 151:127-132, 2000
21. Dabholkar M, Bostick-Bruton F, Weber C, et al: ERCC1 and ERCC2 expression in malignant tissues from ovarian cancer patients. J Natl Cancer Inst 84:1512-1517, 1992
22. Selvakumaran M, Pisarcik DA, Bao R, et al: Enhanced cisplatin cytotoxicity by disturbing the nucleotide excision repair pathway in ovarian cancer cell lines. Cancer Res 63:1311-1316, 2003
23. Furuta T, Ueda T, Aune G, et al: Transcription-coupled nucleotide excision repair as a determinant of cisplatin sensitivity of human cells. Cancer Res 62:4899-4902, 2002
24. Köberle B, Masters JR, Hartley JA, et al: Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. Curr Biol 9:273-276, 1999
25. Takebayashi Y, Nakayama K, Kanzaki A, et al: Loss of heterozygosity of nucleotide excision repair factors in sporadic ovarian, colon and lung carcinomas: Implication for their roles of carcinogenesis in human solid tumors. Cancer Lett 174:115-125, 2001
26. Yang-Feng TL, Li S, Han H, et al: Frequent loss of heterozygosity on chromosomes Xp and 13q in human ovarian cancer. Int J Cancer 52:575-580, 1992
27. Shivji KK, Kenny MK, Wood RD: Proliferating cell nuclear antigen is required for DNA excision repair. Cell 69:367-374, 1992
28. Koeppel F, Poindessous V, Lazar V, et al: Irofulven cytotoxicity depends on transcription-coupled nucleotide excision repair and is correlated with XPG expression in solid tumor cells. Clin Cancer Res 10:5604-5613, 2004
29. Skotheim RI, Autio R, Lind GE, et al: Novel genomic aberrations in testicular germ cell tumors by array-CGH, and associated gene expression changes. Cell Oncol 28:315-326, 2006
30. Wang D, Lippard SJ: Cellular processing of platinum anticancer drugs. Nat Rev Drug Discov 4:307-320, 2005
31. Rabik CA, Dolan ME: Molecular mechanisms of resistance and toxicity associated with platinating agents. Cancer Treat Rev 33:9-23, 2007
32. Wernyj RP, Morin PJ: Molecular mechanisms of platinum resistance: Still searching for the Achilles' heel. Drug Resist Updat 7:227-232, 2004

Acknowledgment

We thank Jenny Gross for administrative support in this research.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Cloning of genes involved in chromosomal translocations by high-resolution single nucleotide polymorphism genomic microarray

Norihiko Kawamata[a,b,c], Seishi Ogawa[b,d], Martin Zimmermann[b,e], Birte Niebuhr[f], Carol Stocking[f], Masashi Sanada[d], Kari Hemminki[g], Go Yamatomo[d], Yasuhito Nannya[d], Rolf Koehler[h], Thomas Flohr[h], Carl W. Miller[a], Jochen Harbott[i], Wolf-Dieter Ludwig[j], Martin Stanulla[e], Martin Schrappe[k], Claus R. Bartram[h,l], and H. Phillip Koeffler[a,l]

[a]Hematology/Oncology, Cedars-Sinai Medical Center/UCLA School of Medicine, Los Angeles, CA 90048; [d]Regeneration Medicine of Hematopoiesis, School of Medicine, University of Tokyo, Tokyo 113-8655, Japan; [e]Department of Pediatric Hematology and Oncology, Children's Hospital, Hannover Medical School, 30625 Hannover, Germany; [f]Molecular Pathology Heinrich-Pette-Institute, 20251 Hamburg, Germany; [g]Division of Molecular Genetic Epidemiology, German Cancer Research Center, 69120 Heidelberg, Germany; [h]Institute of Human Genetics, University of Heidelberg, 69117 Heidelberg, Germany; [i]Department of Hematology and Oncology, Center for Pediatrics, 35390 Giessen, Germany; [j]Department of Hematology, Oncology and Tumor Immunology, Robert-Rössle-Clinic at the HELIOS-Clinic Berlin-Buch, Charité, 13125 Berlin, Germany; and [k]Department of Pediatrics, University of Kiel, 69117 Kiel, Germany

Edited by Joe W. Gray, Lawrence Berkeley National Laboratory, Berkeley, CA and accepted by the Editorial Board June 13, 2008 (received for review November 21, 2007)

High-resolution single nucleotide polymorphism genomic microarray (SNP-chip) is a useful tool to define gene dosage levels over the whole genome, allowing precise detection of deletions and duplications/amplifications of chromosomes in cancer cells. We found that this new technology can also identify breakpoints of chromosomes involved in unbalanced translocations, leading to identification of fusion genes. Using this technique, we found that the *PAX5* gene was rearranged to a variety of partner genes including *ETV6*, *FOXP1*, *AUTS2*, and *C20orf112* in pediatric acute lymphoblastic leukemia (ALL). The 3′ end of the *PAX5* gene was replaced by the partner gene. The PAX5 fusion products bound to PAX5 recognition sequences as strongly as wild-type PAX5 and suppressed its transcriptional activity in a dominant-negative fashion. In human B cell leukemia cells, binding of wild-type PAX5 to a regulatory region of *BLK*, one of the direct downstream target genes of PAX5, was diminished by expression of the PAX5-fusion protein, leading to repression of *BLK*. Expression of PAX5-fusion genes in murine bone marrow cells blocked development of mature B cells. PAX5-fusion proteins may contribute to leukemogenesis by blocking differentiation of hematopoietic cells into mature B cells. SNP-chip is a powerful tool to identify fusion genes in human cancers.

chromatin immunoprecipitation | dominant negative | fusion gene | PAX5 | SNP-chip

Pediatric acute lymphoblastic leukemia (ALL) is the most common malignant disease in children (1–3). It is a genetic abnormality resulting from accumulation of mutations in tumor suppressor genes and oncogenes (1–3). Fusion genes including *ETV6/RUNX1* and *E2A/PBX1* are frequently detected in pediatric ALL (1). Deletion of the *INK4A/ARF* gene (9p21) is also a common abnormality in ALL (1). However, other genetic changes remain to be elucidated in this disease.

Identification of mutated genes in ALL has evolved with improvements in technology. A very recent approach is single nucleotide polymorphism (SNP) analysis using an array based technology (4–6) that allows identification of amplifications, deletions, and allelic imbalances, such as uniparental disomy (represents doubling of the abnormal allele due to somatic recombination or duplication, and loss of the other normal allele) (7, 8). However, SNP-chip analysis is only able to detect changes of gene dosage and is unable to identify balanced translocations, which commonly occur in ALL.

Previously, we analyzed 399 pediatric ALL cases by SNP-chip analysis and found a number of genomic abnormalities, in addition to well known common alterations (9). This technique is sensitive enough to identify genes involved in start sites of deletions/duplications. Indeed, this method allowed us to identify that the *PBX1* gene was involved in start sites of duplication of 1q23 generated by der(19)t(1;19)(q23;p13) (9). Furthermore, correlation analysis of the individual genomic abnormalities suggested the presence of der(12)t(12;21)(p13;q22) and der(21)t(12;21)(p13;q22), as well as dic(9;20)(p13;q11) (9).

In this study, we found that this new technology permitted us to identify genes involved in well known unbalanced translocations including *ETV6/RUNX1*. Further, we found previously undetected fusion genes between *PAX5* and a number of other partner genes by using this technique.

Results

Genes Involved in Unbalanced Translocations Were Identified by SNP-Chip Analysis. Because SNP-chip analysis can only detect changes of gene dosage including deletions, duplications, and amplifications (Fig. 1*A*), this technique is unable to identify balanced translocations (Figs. 1*Aii*). However, when one of a pair of reciprocally translocated chromosomes is lost, SNP-chip analysis can detect this abnormality as partial deletions of involved chromosomes (Fig. 1*Aiii*). Similarly, when one of a pair of reciprocally translocated chromosomes becomes duplicated, SNP-chip can also detect this abnormality as partial duplication of the involved chromosomes (Fig. 1*Aiv*). Furthermore, high resolution SNP-chip analysis allows us to identify the genes involved in these unbalanced translocations.

To prove that SNP-chip analysis can detect unbalanced translocations and the genes involved in these translocations, we

Author contributions: N.K., S.O., B.N., C.S., C.R.B., and H.P.K. designed research; N.K., S.O., B.N., C.S., M. Sanada, G.Y., and Y.N. performed research; N.K., S.O., and M. Sanada contributed new reagents/analytic tools; N.K., S.O., M.Z., M. Sanada, K.H., G.Y., Y.N., R.K., T.F., C.W.M., J.H., W.-D.L., M. Stanulla, M. Schrappe, C.R.B., and H.P.K. analyzed data; and N.K., S.O., M.Z., C.S., M. Stanulla, M. Schrappe, C.R.B., and H.P.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.W.G. is a guest editor invited by the Editorial Board.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EU784145, PAX5-FOXP1; EU784146, PAX5-AUTS2; EU784147, PAX5-C20orf112 short isoform; and EU784148, PAX5-C20orf112 long isoform).

[b]N.K., S.O., and M.Z. contributed equally to this work.

[c]To whom correspondence should be addressed at: Hematology/Oncology, Cedars-Sinai Medical Institute/UCLA Geffen School of Medicine, 8700 Beverly Boulevard, Los Angeles, CA 90048. E-mail: kawamatan@cshs.org.

[l]C.R.B. and H.P.K. contributed equally to this work.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711039105/DCSupplemental.

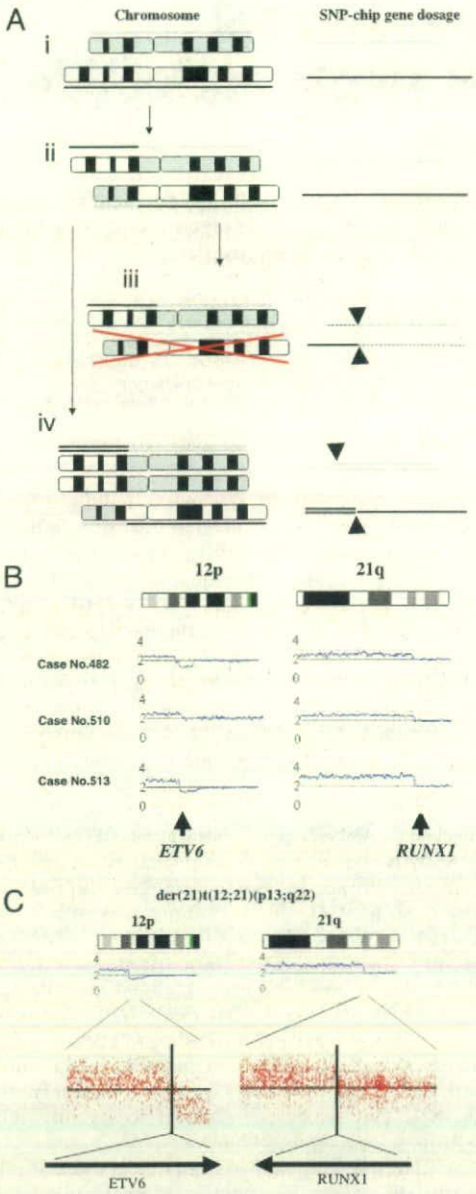

Fig. 1. SNP-chip analysis detected genes involved in unbalanced translocations. (*A*) SNP-chip analysis can identify breakpoints of translocations when one of the paired translocated chromosomes is either lost or duplicated/amplified. (*Left*) Chromosomal status. Gene dosages are indicated either above or beneath the chromosomes. (*Right*) Results of SNP-chip analysis. (*Ai*) Normal chromosomes; gene dosage is normal. (*Aii*) Reciprocal translocation; gene dosage is normal. (*Aiii*) One of the paired translocated chromosomes is lost; gene dosage is lower than normal on the left side of the upper chromosome and the right side of the lower chromosome. Arrow heads indicate the breakpoint of the translocation in each chromosome. (*Aiv*) One of the paired translocated chromosomes is duplicated; gene dosage is higher than normal on the right side of the upper chromosome and the left side of the lower chromosome. Arrow heads indicate the breakpoint of this translocation in each chromosome. (*B*) Representative cases with unbalanced translocation of der(21)t(12;21)(p13;q22). (*Left*) Start sites of duplication at 12p13 involving the *ETV6* gene. (*Right*) Start sites of duplication at 21q22 involving the *RUNX1* gene. SNP-chip data of representative cases with dup(12)(p13) and dup(21)(q22) are shown. These abnormalities were validated by FISH and/or RT-PCR (data not shown). Results of SNP-chip data were visualized by CNAG software. Lines above each chromosome show total gene dosage; level 2 indicates diploid (2N) amount of DNA, which is normal. (*C*) Magnified view of SNP-chip data. (*Upper*) Start sites of duplications at 12p13 and 21q22 are magnified. Signals of individual probe signals are shown. Vertical lines indicated the positions of start sites of duplications. (*Lower*) Genes involved in the start sites of duplications.

analyzed cases having extra copies of *ETV6/RUNX1* fusion genes generated by der(21)t(12;21)(p13;q22) (Fig. 1*B*), which were initially identified by FISH and/or RT-PCR (data not shown). SNP-chip was clearly able to identify this abnormality as duplications involving chromosome 12 and 21 (Fig. 1*B*). Further, the result of high-resolution (250k) SNP-chip clearly identified *ETV6* (12p13) and *RUNX1* (21q22) as the target genes involved in this unbalanced translocation (Fig. 1*C*).

PAX5 Gene Is Frequently Fused to Partner Genes. Our previous data showed the presence of dic(9;20)(p13;q11) in 11 cases of ALL (9), 5 of which had deletion 9p13.2-pter. These 5 cases had start sites of this deletion at 9p13.2 mapping to the *PAX5* gene (Fig. 2*A* and data not shown). This prompted us to reexamine all cases of B-ALL that had deletion of 9p [supporting information (SI) Table S1]. We found a total of 9 cases with similar start sites (9p13.2), mapping to the *PAX5* gene (Fig. 2*A* and data not shown). In 2 of these cases, simple abnormalities were detected by SNP-chip: case 514 had only del9p13.2-pter and del7q11.2-pter; case 458 had only del9p13.2-pter and dup3p13-pter (Table S1 and Fig. 2*A*). Three cases (536, 543, 572) had complex abnormalities including del9p13.2-pter and del20q11.21-qter, all with the *C20ORF112* gene within the start site of del20q (Table S1 and Fig. 2*A*). The other 2 cases (659, 767) had complex abnormalities that included *ETV6* on 12p13 (Table S1 and Fig. 2*A*).

Thus, we found four candidate partner genes fused to *PAX5* in seven cases by SNP-chip analysis; *ETV6* on 12p13 (two cases) (12), *C20orf112* on 20q11.1 (three cases), *AUTS2* on 7q11.1 (one case) and *FOXP1* on 3p13 (one case) (Fig. 2*A*). Because these translocations could lead to fusion transcripts between *PAX5* and different partner genes, the presence of the predicted fusion transcript was examined by RT-PCR using the mapping information from the SNP-chip data. RT-PCR and nucleotide sequencing data of the PCR products confirmed that the *PAX5* gene was fused to either the *ETV6* (two cases), *C20orf112* (three cases), *AUTS2* (one case), or *FOXP1* (one case) gene and transcribed into aberrant fusion messages (Fig. 2 *B* and *C*). Each fusion gene was mutually and exclusively detected in the samples studied. In one case with dic(9;20), exon 5 of *PAX5* was fused to exon 8 of *C20orf112*, and in two cases with dic(9;20), exon 8 of *PAX5* was fused to exon 3 of *C20orf112*. *PAX5/ETV6* involved exon 4 of *PAX5* and exon 3 of *ETV6*.

Cellular Localization and DNA Binding Affinity of PAX5 Fusion Products. In the *PAX5/FOXP1* fusion transcript, the amino acid coding frame of the *FOXP1* gene was not identical to that of *PAX5*, leading to a frame-shift and an early termination codon after the fusion point of these two genes (Fig. 2*D*). However, all other fusion genes were in frame and were predicted to encode chimeric proteins. Two proteins (a short and long form) with different breakpoints were predicted from the *PAX5/C20orf112* fusion genes (Fig. 2*D*).

To confirm cellular localization of PAX5-fusion proteins, we transfected vectors encoding wild-type *PAX5* and *PAX5* fusion genes (*PAX5-ETV6*, *PAX5-FOXP1*, *PAX5-C20ORF112S*, and *PAX5-C20ORF112L*) into 293T cells, fractionated the cytoplasmic and nuclear proteins, and examined the wild-type PAX5 and PAX5-fusion proteins by Western blot analysis (Fig. 2*E*). PAX5-ETV6 protein was detected in both the cytoplasm and nucleus; PAX5-FOXP1 and PAX5-C20ORF112L proteins were predom-

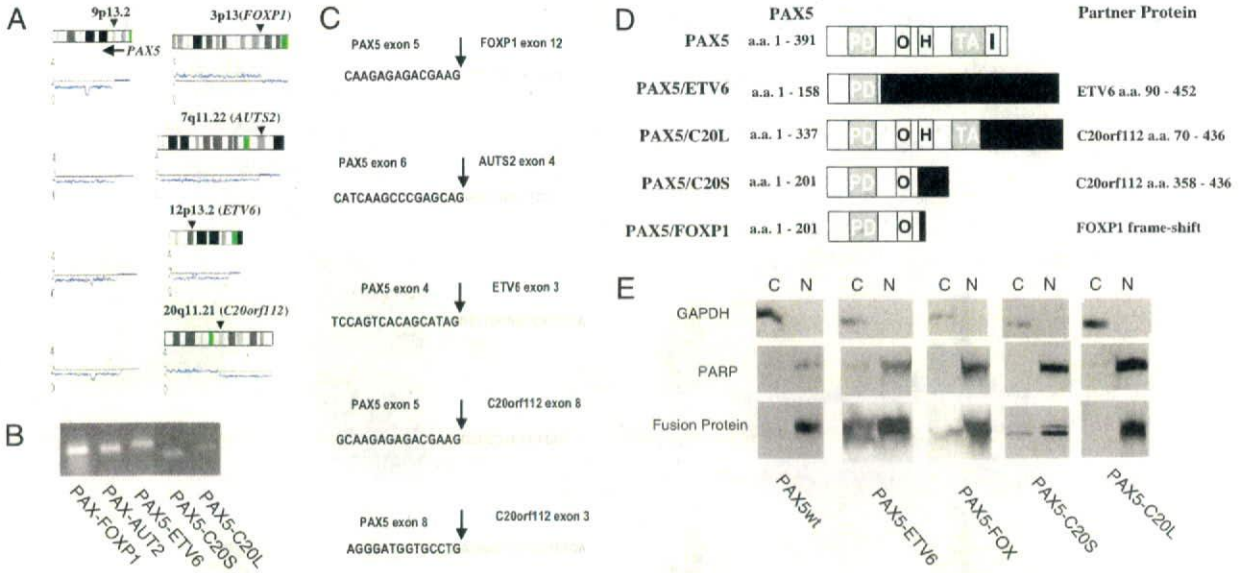

Fig. 2. *PAX5* gene is fused to partner genes. (*A*) Start sites of deletion at 9p13.2 involving the PAX5 gene. (*Left*) SNP-chip data of representative cases with 9p13.2 deletions. A vertical arrow indicates the start sites of 9p deletion that involves the *PAX5* gene. A horizontal arrow shows the direction of transcription of the *PAX5* gene. (*Right*) Chromosomal abnormalities of partner chromosomes. Arrow heads indicate the start sites of duplication or deletions. Genes involved in the start sites are shown. (*B*) Result of RT-PCR. The ALL samples suggesting the presence of PAX5 fusion genes by SNP-chip analysis were examined by RT-PCR using the primers of PAX5 and the respective partner genes. (*C*) Fusion sequences of the *PAX5* and partner genes. Joining sequences of fused transcripts are shown from the indicated exon of the fused gene. (*D*) Schematic structure of wild-type and mutant PAX5. Amino acid positions (aa) of each protein are indicated. PAX5/FOXP1 fusion construct has an early termination codon caused by a frame-shift. PD, paired domain; TA, transcription activation domain; O, octapeptide H, homeodomain-like; I, inhibitory domain. (*E*) Subcellular fractionation of PAX5-fusion proteins. pcDNA vector encoding wild-type PAX5, PAX5-ETV6, PAX5-FOXP1, PAX5-C20ORF112S, or PAX5-C20ORF112L was transfected into 293T cells. Nuclear and cytoplasmic proteins were separated and electrophoresed in the gel. Localization of PAX5-fusion proteins was examined by PAX5 N-terminal specific antibody. Purity of cytoplasmic protein was examined with anti-GAPDH antibody and purity of nuclear proteins with the anti-PARP antibody. C, cytoplasmic fraction; N, nuclear fraction.

inantly localized in the nucleus; and 20% and 80% of PAX5-C20ORF112S proteins were localized in the cytoplasm and the nucleus, respectively (Fig. 2*E*). Localization of the fusion proteins was also confirmed by immunohistochemical staining (data not shown).

Because PAX5-fusion proteins were localized in the nucleus, we analyzed DNA binding affinity of these PAX5-fusion proteins *in vitro*. DNA binding affinity of the PAX5 wild-type and fusion proteins expressed in 293T cells was analyzed by electrophoretic mobility shift assay (EMSA), and signals of probes bound to the proteins were plotted graphically (Fig. 3*A*). Binding activity of each protein in the absence of cold competitor oligonucleotide probe was regarded as 1.0, and the binding activity in the presence of cold competitor oligonucleotide probes was measured. All PAX5-fusion proteins showed similar binding activity to the PAX5 recognition sequences as the wild-type PAX5 (Fig. 3*A*).

***PAX5* Fusion Products Suppressed Transcriptional Activity of Wild-Type PAX5 in a Dominant Negative Fashion, Leading to Inhibition of B-Cell Development.** To examine the effect of PAX5-fusion proteins on transcriptional activity of wild-type PAX5, we performed a reporter gene assay using 293T cells. Cotransfection reporter gene assays using wild-type and fusion *PAX5* expression vectors along with a reporter gene driven by the murine CD19 promoter (which contains three repeats of PAX5 binding sequences) showed that the PAX5 fusion products suppressed transcriptional activity of PAX5 in a dominant-negative fashion (Fig. 3*B*). Expression of wild-type PAX5 proteins was minimally affected by coexpression of PAX5-fusion proteins (Fig. 3*C*), suggesting that PAX5-fusion proteins competed with wild-type PAX5 for the PAX5 binding sequences on the reporter gene.

Further, we transfected vectors encoding either PAX5-C20orf112S or PAX5-C20orf112L, each coexpressing the GFP marker, into Nalm 6 cells (a human B cell ALL cell line, which expresses endogenous PAX5) (data not shown). After transfection, GFP-positive cells were sorted by FACS and expression of PAX5-downstream genes was examined by semiquantitative RT-PCR (Fig. 3*D* and data not shown). We examined 10 downstream target genes (seven positively regulated direct target genes and three negatively regulated genes) of PAX5 (10–12) and found that four, including *ATP1B1*, *BLK*, *NEDD5* and *TCF7L2*, were down-regulated by induction of either PAX5-C20orf112S or PAX5-C20orf112L protein. However, expression of other reported PAX5 downstream target genes, including three positively regulated direct target genes (*IRF8*, *BST1*, *CD19*) and three negatively regulated genes (*CCR2*, *CCR5*, *NOTCH1*) were not affected by the induction of expression of the fusion proteins in these cells.

To examine the effect of PAX5 fusion protein on binding of wild-type PAX5 to the direct target gene *BLK* in the leukemic cells, we performed chromatin-immunoprecipitation (ChIP) assay using Nalm 6 cells transfected with either an empty vector or a construct encoding PAX5-C20orf112S. We used a PAX5 antibody detecting the C-terminal region of the protein, which could detect wild-type PAX5, but not PAX5-C20orf112S, as the C-terminal end of PAX5 was replaced by C20orf112S in this fusion protein. Although wild-type specific PAX5 antibody precipitated the promoter region of *BLK* after transfection of the empty vector, the amount of DNA of the *BLK* promoter region bound to wild-type PAX5 was reduced after transfection of the PAX5-C20orf112S gene (Fig. 3 *E* and *F*).

To examine the effect of PAX5-fusion proteins on B cell development in murine hematopoietic cells, we infected murine

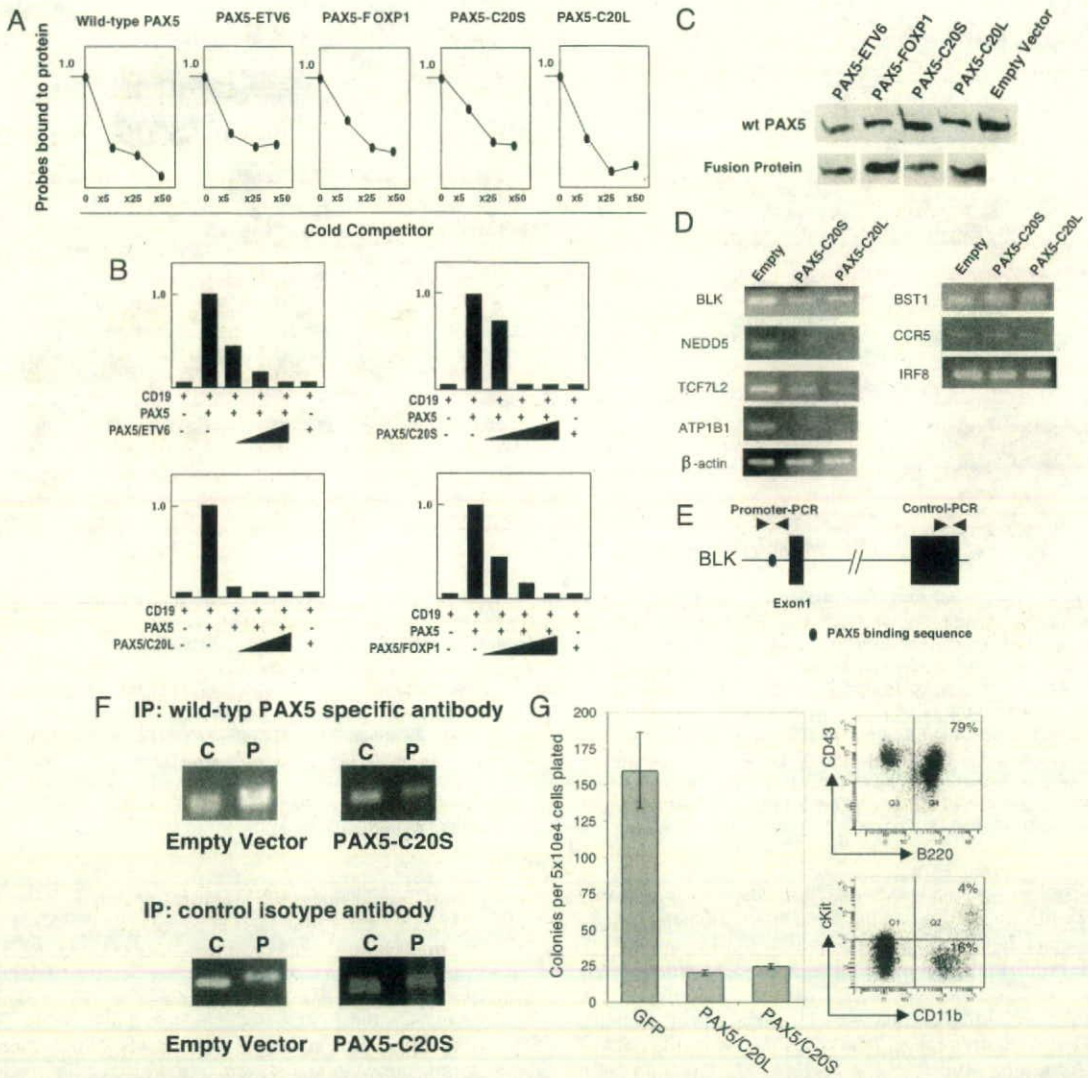

Fig. 3. *PAX5*-fusion proteins suppress transcriptional activity of PAX5 in a dominant-negative fashion and block the growth of B cells. (*A*) Result of EMSA: Wild-type PAX5 and PAX5 fusion were expressed in 293T cells, and nuclear proteins were purified. The purified nuclear proteins were mixed with radioisotope labeled double-strand oligonucleotide DNA, in either the presence or absence of cold competitor oligonucleotides (5-, 25-, and 50-fold cold competitor probes). Intensity of each shifted band indicating DNA–protein complex was measured and plotted graphically. Intensity of shifted bands in the absence of cold competitor probes was regarded as 1.0. (*B*) Reporter gene assay. Wild-type and mutant PAX5 were mixed at a various ratios (1:0, 1:0.3, 1:1, 1:3, respectively, 1 = 500 ng of construct) and transfected. Forty-eight hours later, relative activity of firefly luciferase was measured and plotted. Results represent the mean values of the three experiments. CD19, PAX5 luciferase reporter construct having PAX5 binding region of CD19 promoter; PAX5, wild-type PAX5; PAX5/ETV6, PAX5/ETV6 fusion; PAX5/C20L, long form of PAX5/C20orf112 fusion in which PAX5 exon 8 is fused to C20orf112 exon 3; PAX5/C20S, short form of PAX5/C20orf112 fusion in which PAX5 exon 5 is fused to C20orf112 exon 8; PAX5/FOXP1, PAX5/FOXP1 fusion with an early termination codon caused by a frame-shift after the site of fusion. (*C*) Results of expression of wild-type PAX5 and PAX5-fusion proteins. After cotransfection of equal amounts of vector encoding either wild-type or fusion PAX5 genes into 293T cells, the expression of respective proteins was examined by Western blot. Levels of expression of wild-type PAX5 protein were minimally affected by coexpression of the PAX5-fusion proteins. (*D*) Semiquantitative RT-PCR of downstream target genes of PAX5. Expression of PAX5 downstream target genes was examined by semiquantitative RT-PCR. Nalm 6, a human B cell ALL cell line expressing endogenous PAX5, was transfected with pMSCV-GFP (Empty), pMSCV-GFP-PAX5-C20orf112S (PAX5-C20S), or pMSCV-GFP-PAX5-C20orf112L (PAX5-C20L). GFP-positive cells were sorted and subject to semiquantitative RT-PCR. Optimal cycle numbers to semiquantify the expression of respective genes are as follows; BLK: 25 cycles; Nedd5; 25 cycles; TCF7L2: 25 cycles; ATP1B1: 25 cycles; β-actin: 22 cycles; CCR2: 25 cycles; CCR8: 30 cycles; IRF8: 30 cycles. (*E*) Structure of human *BLK* gene. Structure of *BLK* and primers used for ChIP assay within the 5′ regulatory region (Promoter-PCR) and 3′ end (Control-PCR) of the *BLK* gene is schematically shown. PAX5 binding site in the promoter region is indicated. (*F*) ChIP analysis of the PAX5 binding site in the *BLK* gene promoter. pMSCV-GFP (empty vector) or pMSCV-GFP-PAX5-C20S (PAX5-C20S) was transfected into human Nalm 6 B cell leukemia cells expressing endogenous PAX5. GFP-positive cells were subject to ChIP assay. The cells were fixed in formaldehyde solution and sonicated by ultrasound. DNA-protein complex was incubated with wild-type PAX5 specific antibody, which detected the C-terminal region of PAX5 but not the PAX5-C20orf112S protein (*Upper*). As a control, the DNA-protein complex was reacted with isotype nonspecific antibody (*Lower*). Immnoprecipitated DNA was subjected to PCR to amplify either the *BLK* promoter region containing PAX5 binding sequence (P) or, as an internal control, the 3′ end of the *BLK* gene (C). (*G*) Retrovirus infection experiments. Murine bone marrow cells were collected at 5 days after injection of 5FU. The hematopoietic cells were infected by retrovirus containing pMSCV-GFP empty vector (GFP), pMSCV-GFP-C20orf112L (PAX5/C20L), or pMSCV-GFP-C20orf112S (PAX5/C20S). GFP-positive murine hematopoietic cells were sorted and plated at 5×10^4 cells per plate in methylcellulose containing mSCF, mIL7, and hFL. At 8 days after the plating, the colony numbers were counted (*Left*; results represent means and SD of three experiments). Cell surface antigens on the GFP-positive cells infected with pMSCV-GFP (GFP) at Day 11 were examined by FACS using antibodies against CD43 and B220 (*Upper Right*), c-kit and CD11b (*Lower Right*) antibodies, to confirm the development of B cells.

hematopoietic cells from the bone marrow with retroviral vectors encoding either PAX5-C20orf112S or PAX5-C20orf112L. GFP-positive infected bone marrow cells were sorted by FACS and plated in media containing cytokines that are know to stimulate B cell differentiation (Fig. 3*F*). Murine hematopoietic cells infected with the empty vector showed abundant colonies (Fig. 3*G Left*), and 79% of the cells were B220 positive B cells (Fig. 3*G Upper Right*). In contrast, murine hematopoietic cells infected with either PAX5-C20orf112S or PAX5-C20orf112L formed very few colonies (Fig. 3*G Left*). Most of these colonies were GFP-negative (data not shown), suggesting that these PAX5-fusion proteins impaired B cell development from murine hematopoietic cells.

Discussion

In this study, we describe a paradigm for discovering fusion genes in malignancy by taking advantage of samples with unbalanced translocations and using high density SNP-chip analysis. This technique allows us to identify genes involved in translocations even if chromosomal analysis is not available, especially in solid tumors.

Steps to identify novel fusion genes using SNP chip analysis include (*i*) identify either a deletion or duplication that occurs within two genes; (*ii*) determine whether transcription of both genes is in the same direction; (*iii*) take advantage of ancillary tests such as standard chromosomal analysis or spectral karyotyping (14), which can grossly show that two chromosomes are fused; and (*iv*) design primers of candidate genes and perform RT-PCR to clone fusion genes. Rapid amplication of cDNA ends (RACE) (15) or long-distance PCR (12) also help the cloning of genes involved in translocations. In our SNP-chip data, a number of regions of segmental deletions or duplications were detected (9). Although some of them are simple deletions or duplications at the original sites of the chromosomes, the others are deletions that occurred during chromosomal translocations or when duplicated fragments were inserted into chromosomal sites other than the original region (data not shown). Therefore, data of chromosomal analysis help to define translocations, leading to identification of candidate genes in novel fusion genes.

Recently, Tomlins *et al.* found the fusion genes *TMPRSS2/ERG* and *TMPRSS2/ETV1* in prostate cancers by using expression microarray data (16). They focused on the genes *ERG* and *ETV1*, which are highly expressed in this cancer and examined levels of individual exons of these two genes (16). They found differences in expression of 5′ and 3′ regions of the genes, suggesting that these genes are fused to each other (16). In these fusion genes, the 5′ regions were replaced by the *TMPRSS2* gene, resulting in the differences in the expression of the 5′ and 3′ region of the *ERG/ETV1* genes (16). They also used SNP-chip analysis to identify these fusion genes and found a deletion of a genomic region between TMPRSS2 (21q22.3) and ERG (21q22.2), leading to fusion of these two genes (17). These new technologies, based on oligonucleotide microarrays and bioinformatics, will help to identify fusion genes in cancers.

Our study found that the *PAX5* gene was frequently fused to one of a variety of partner genes. PAX5 is a key transcription factor in the development of B cells (18, 19). We found that these PAX5 fusion proteins suppressed the function of wild-type PAX5 in a dominant-negative fashion and suppressed expression of downstream target genes of wild-type PAX5 in leukemic cells.

We found that when *PAX5* was joined to one of its fusion partner genes, its C-terminal end was replaced by one of the partner genes. Elimination of the C-terminal end of *PAX5* may play an important role in generation of a dominant negative form of mutated *PAX5*. In *in vitro* assays, PAX5-fusion proteins showed a similar affinity as wild-type PAX5 for the PAX5 recognition sequences. Although expression of several downstream targets of wild-type PAX5 was repressed by expression of PAX5-fusion proteins, others were not affected. Binding of transcription factors to DNA can be modulated by cofactors and/or neighboring transcription factors (20). Compared to PAX5, PAX5-fusion proteins may bind more strongly to some target genes and more weakly to others, depending on the contextual environment of the target genes.

Further, our data showed that PAX5-fusion protein inhibited B cell development of hematopoietic cells in a colony formation assay. This result may suggest that PAX5 fusion protein blocked differentiation of hematopoietic cells into mature B cells. PAX5-deficient mice have impairment of B cell differentiation (18). These data suggest that PAX5-fusion proteins may contribute to leukomegenesis by blocking B cell differentiation. It has been suggested that two distinct genetic abnormalities contribute to leukemogenesis in acute myelogenous leukemia (AML); one is mutations promoting cellular proliferation, for example *FLT3* or *RAS* mutations, and the other is mutations blocking differentiation, for example *PML-RARA* or *RUNX1-ETO* (21, 22). PAX5-fusion proteins may cooperate with unidentified mutations promoting cellular proliferation in the ALL cells.

Recently, Mullighan *et al.* have analyzed pediatric ALL samples by high density SNP-chips and found frequent abnormalities of *PAX5* gene (23). Their data also showed that PAX5 fusion products suppressed transcriptional activity of PAX5 in a dominant-negative fashion (23). In addition, other researcher have reported *PAX5* fusion genes, including *PAX5* fused to *ETV6* (12p13) (23, 24), *FOXP1* (3p14) (23), *ZNF521* (18q11) (23), *ELN* (7q11.23) (25), and *PML* (15q24) (26). We have found *PAX5* fused to either *ETV6*, *FOXP1*, *C20orf112* (20q11), or *AUTS2* (7q11.22).

In our study, the function of PAX5 was attenuated by the dominant-negative forms of the fusion products in B cell lineage ALL, suggesting that PAX5 behaves as a tumor suppressor in early B cells, and that impairment of its function can be associated with the development of ALL. In contrast, translocation of the *PAX5* gene to the enhancer region of the Ig heavy chain gene [t(9;14)(p13.2;q32)] or point mutations of the 5′ regulatory region of the *PAX5* gene leads to its overexpression, which is associated with B cell lineage lymphomas (27–29). Also, experimental overexpression of wild-type PAX5 can transform lymphocytes (30, 31). Therefore, an aberrant PAX5 may behave in a dominant-negative fashion at the pre-B stage of B cell development, resulting in ALL; its forced expression in a more mature B cell can lead to lymphoma. Our study showed that PAX5-fusion proteins blocked differentiation of B cells but did not transform them. B cells at different stages of differentiation may need alteration of distinct sets of pathways to transform. Why PAX5 can act as a tumor suppressor in ALL and as an oncoprotein in lymphoma is unclear. Further studies are needed to clarify the mechanism of this paradoxical phenomenon in carcinogenesis.

In summary, we identified multiple fusion genes in ALL by SNP-chip analysis, leading to the exploration of a B cell differentiation block as a contributing factor to the development of ALL. This methodology should help researchers to identify oncogenic fusion genes and explore the mechanism of tumorigenesis in other types of cancers as well.

Materials and Methods

Samples and DNA/RNA Preparation. SNP-chip was performed on 399 pediatric ALL patients consecutively enrolled in the ALL-BFM 2000 trial of the Berlin-Frankfurt-Münster (BFM) study at diagnosis and during remission (350 cases were B cell lineage ALL and 49 cases were T cell lineage ALL) (9). Detailed results of the SNP-chip analysis are published separately (9). The ALL-BFM 2000 study was approved by the local ethics committee. DNA and RNA were extracted from the ALL samples and cell lines by using standard techniques (32). Nalm 6, a human pre-B ALL cell line, was generously provided by Dr. G. Crook (Los Angeles Children's Hospital, Los Angeles, CA) and maintained in RPMI medium 1640 with 10% FBS.

SNP-Chip Analysis. SNP-chip of GeneChip Human mapping 50k array XbaI 240 and/or 250k Nsp were used for this study (Affymetrix Japan). Preparation of samples was reported previously (4, 5). The data were analyzed by CNAG program as previously described (4, 5). All 399 ALL samples and their matched control samples were analyzed by using 50K-SNP chip; selected cases with genomic abnormalities were also analyzed by using 250K SNP-chip.

RT-PCR. RT-PCR was performed by using ThermoScript RT-PCR Systems (Invitrogen) according to the manufacturer's protocol. The primers used for detection of *PAX5* fusion transcripts are listed in Table S2. Expression of PAX5 downstream target genes in Nalm 6 cells after transfection was examined by semiquantitative RT-PCR. The gene names and their primer sequences are listed in Table S3.

Reporter Gene Constructs and Expression Vectors. The PAX5 reporter gene construct with the luciferase gene and PAX5 binding region of the CD19 promoter, as well as the human PAX5 cDNA constructs, were kindly provided by Dr. M. Busslinger (Research Institute of Molecular Pathology, Vienna, Austria). PAX5-fusion constructs were generated by using PCR. All coding regions were ligated into the pcDNA3.1 vectors (Stratagene). Wild-type PAX5 cDNA was ligated into pMSCV vector (Clontech), and EGFP cDNA was ligated under the control of pGK promoter as a marker (pMSCV-GFP-wtPAX5). PAX5-C20orf112S and PAX5-C20orf112L cDNA sequences were also ligated into pMSCV-GFP vectors.

Transfection and Reporter Gene Assay. For reporter gene assays, pMSCV-GFP-wtPAX5 and pcDNA vectors encoding PAX5-fusion genes were cotransfected with the PAX5 reporter construct and pRL (*Renilla* luciferase) vector into 293T cells by using the Effecten transfection kit (Qiagen). Firefly and *Renilla* luciferase activities were measured with the Dual-Luciferase Reporter Assay System (Promega). Transfection into Nalm6 human pre-B cell ALL cell line was performed with Amaxa nucleofector. GFP-positive cells were sorted by using the MoFlo cell sorter (Dako). Detailed information about the procedure is described in *SI Text*.

Retrovirus Transduction into Murine Hematopietic Cells. Retrovirus containing pMSCV-GFP (empty), pMSCV-GFP-PAX5-C20orf112S, and pMSCV-GFP-PAX5-C20orf112L was generated. The retrovirus was transfected into murine bone marrow cells as previously reported (33). After the transfection, GFP-positive cells were sorted and plated into methylcellulose cultures (M3231; Stem Cell Technologies) as previously described (33). Surface antigens (CD43, B220, c-kit, and CD11b) of these GFP-positive cells were examined by using FACScan (Becton–Dickinson). Detailed information of the procedure is described in the *SI Text*.

Subcellular Fractionation of Proteins and EMSA. Forty-eight hours after transfection of vectors into 293T cells, the cells were subjected to subcellular fractionation with the CelLytic NuCLEAR Extraction Kit (Sigma-Aldrich). Detailed information of the procedure is described in the *SI Text*.

Purified nuclear proteins from the cells were also subjected to EMSA as previously reported (34). Detailed information of the procedure is described in the *SI Text*.

Chromatin Immunoprecipitation (ChIP) Assay. ChIP assay was performed with the Magna ChIP A kit from Millipore according to the manufacturer's protocol. pMSCV-GFP or pMSCV-GFP-C20orf112S were transfected into Nalm 6 cells as described above, and GFP-positive cells were sorted by MoFlo (Dako). Precipitated DNA was recovered and subjected to PCR to amplify the *BLK* promoter region and the 3′ end of the *BLK* gene (internal control). The primer sequences used for ChIP assay are listed in Table S4. Detailed information of the procedure is described in the *SI Text*.

ACKNOWLEDGMENTS. We thank Dr. M. Busslinger (Research Institute of Molecular Pathology, Vienna, Austria) for generously providing the CD19 promoter reporter construct and the human PAX5 cDNA. This study was supported by the Parker Hughes Fund and by grants from the National Institutes of Health. N.K. is supported by a fellowship from The Tower Cancer Research Foundation. H.P.K. holds the Mark Goodson Chair in Oncology Research at Cedars-Sinai and is a member of the Jonsson Cancer Center and the Molecular Biology Institute of the University of California, Los Angeles. This work was also supported by a grant-in-aid from the Department of Health, Welfare and Labor; by the Ministry of Education, Culture, Sports, Science and Technology (Japan); by European Union Grant FOOD-CT-2005-016320; by grants from the Deutsche Krebshilfe to Cambridge Research Biochemicals; and by the Fritz-Thyssen Foundation (C.S.). The ALL-BFM 2000 trial is supported by Grant 50-2698-Schr 1 of the Deutsche Krebshilfe.

1. Armstrong SA, Look AT (2005) Molecular genetics of acute lymphoblastic leukemia. *J Clin Oncol* 23:6306–6315.
2. Pui CH, Evans WE (2006) Treatment of acute lymphoblastic leukemia. *N Engl J Med* 354:166–178.
3. Pui CH, Relling MV, Downing JR (2004) Acute lymphoblastic leukemia. *N Engl J Med* 350:1535–1548.
4. Nannya Y, *et al.* (2005) A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res* 65:6071–6079.
5. Yamamoto G, *et al.* (2007) Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am J Hum Genet* 81:114–126.
6. Lindblad-Toh K, *et al.* (2000) Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nat Biotechnol* 18:1001–1005.
7. Raghavan M, *et al.* (2005) Genome-wide single nucleotide polymorphism analysis reveals frequent partial uniparental disomy due to somatic recombination in acute myeloid leukemias. *Cancer Res* 65:375–378.
8. Lehmann S, *et al.* (2008) Molecular allelokaryotyping of early-stage, untreated chronic lymphocytic leukemia. *Cancer* 112:1296–1305.
9. Kawamata N, *et al.* (2008) Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* 111:776–784.
10. Lin YH, Shin EJ, Campbell MJ, Niederhuber JE (1995) Transcription of the blk gene in human B lymphocytes is controlled by two promoters. *J Biol Chem* 270:25968–25975.
11. Schebesta A, *et al.* (2007) Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* 27:49–63.
12. Delogu A, *et al.* (2006) Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* 24:269–281.
13. Schröck E, Padilla-Nash H (2000) Spectral karyotyping and multicolor fluorescence in situ hybridization reveal new tumor-specific chromosomal aberrations. *Semin Hematol* 37:334–347.
14. Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts by amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85:8998–9002.
15. Cheng S, Fockler C, Barnes WM, Higuchi R (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc Natl Acad Sci USA* 91:5695–5699.
16. Tomlins SA, *et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310:644–648.
17. Perner S, *et al.* (2006) TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res* 66:8337–8341.
18. Urbanek P, Wang ZQ, Fetka I, Wagner EF, Busslinger M (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* 79:901–912.
19. Xie H, Ye M, Feng R, Graf T (2004) Stepwise reprogramming of B cells into macrophages. *Cell* 117:663–676.
20. Kumaran RI, Thakar R, Spector DL (2008) Chromatin dynamics and gene positioning. *Cell* 132:929–934.
21. Gilliland DG, Griffin JD (2002) The roles of FLT3 in hematopoiesis and leukemia. *Blood* 100:1532–1542.
22. Tenen DG (2003) Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer* 3:89–101.
23. Mullighan CG, *et al.* (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446:758–764.
24. Cazzaniga G, *et al.* (2001) The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Res* 61:4666–4670.
25. Bousquet M, *et al.* (2007) A novel PAX5-ELN fusion protein identified in B cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. *Blood* 109:3417–3423.
26. Nebral K, *et al.* (2007) Identification of PML as novel PAX5 fusion partner in childhood acute lymphoblastic leukaemia. *Br J Haematol* 139:269–274.
27. Busslinger M, Klix N, Pfeffer P, Graninger PG, Kozmik Z (1996) Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc Natl Acad Sci USA* 93:6129–6134.
28. Iida S, *et al.* (1996) The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood* 88:4110–4117.
29. Pasqualucci L, *et al.* (2007) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412:341–346.
30. Anderson K, *et al.* (2007) Ectopic expression of PAX5 promotes self renewal of biphenotypic myeloid progenitors co-expressing myeloid and B-cell lineage associated genes. *Blood* 109:3697–3705.
31. Souabni A, Jochum W, Busslinger M (2007) Oncogenic role of Pax5 in the T-lymphoid lineage upon ectopic expression from the immunoglobulin heavy-chain locus. *Blood* 109:281–289.
32. Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual.* (Cold Spring Harbor Press, Cold Spring Harbor, NY), 3rd Ed.
33. Schwieger M, *et al.* (2002) AML1-ETO inhibits maturation of multiple lymphohematopoietic lineages and induces myeloblast trnsformation in syngergy with ICSBP deficiency. *J Exp Med* 196:1227–1240.
34. Sato H, Wang D, Kudo A (2001) Dissociation of Pax-5 from KI and KII sites during kappa-chain gene rearrangement correlates with its association with the underphosphorylated form of retinoblastoma. *J Immunol* 166:6704–6710.

Vaccine 26 (2008) 4686–4696

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Application of quantitative gene expression analysis for pertussis vaccine safety control☆

Isao Hamaguchi[a,1], Jun-ichi Imai[b,1], Haruka Momose[a,1], Mika Kawamura[b,c], Takuo Mizukami[a], Seishiro Naito[a], Jun-ichi Maeyama[a], Atsuko Masumi[a], Madoka Kuramitsu[a], Kazuya Takizawa[a], Hiroshi Kato[a], Tetsuya Mizutani[d], Yoshinobu Horiuchi[e], Nobuo Nomura[f], Shinya Watanabe[b], Kazunari Yamaguchi[a,*]

[a] Department of Safety Research on Blood and Biological Products, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan
[b] Department of Clinical Informatics, Tokyo Medical and Dental University, Tokyo, Japan
[c] Medicrome, Inc., Sendagaya, Shibuya-ku, Tokyo 151-0051, Japan
[d] Department of Virology I, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan
[e] Department of Bacterial Pathogenesis and Infection Control, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashimurayama, Tokyo 208-0011, Japan
[f] Biological Information Research Center, National Institute of Advanced Industrial Science and Technology, Japan

ARTICLE INFO

Article history:
Received 19 February 2008
Received in revised form 10 June 2008
Accepted 15 June 2008
Available online 9 July 2008

Keywords:
Pertussis vaccine
Microarray
Safety test
Quantitative analysis

ABSTRACT

Although vaccines are routinely used to prevent infectious diseases, little is known about the comprehensive influences caused by vaccines. In this study, we showed, using comprehensive gene expression analysis, that pertussis vaccine affected many genes in multiple organs of vaccine-treated animals. In particular, lung was revealed to be the most suitable target to evaluate pertussis vaccine toxicity. The 13 genes identified from the analysis of vaccine-treated lung at day 1 showed a clear dendrogram corresponding to pertussis vaccine toxicity. Furthermore, quantitative analysis of these genes revealed a positive correlation between their respective expression levels and the degree of toxic effects observed in samples that had been treated with various doses of reference pertussis vaccines. The quantification of this 13 gene-set is an indicator of the vaccine toxicity-related reaction.

1. Introduction

Vaccination is among the most significant public health success stories of all time. However, like any pharmaceutical product, no vaccine is completely safe. While a large part of known vaccine adverse events is minor and self-limited, some vaccines have been associated with very rare but serious health effects. Whole-cell pertussis vaccine is composed of a suspension of inactivated *Bordetella pertussis* cells, whereas acellular pertussis vaccines contain purified, inactivated components of *B. pertussis* cells. Current acellular pertussis vaccines mostly contain pertussis toxin (PT) and filamentous hemagglutinin [1]. As with all injected vaccines, administration of pertussis vaccine may cause local reactions, such as pain, redness, or swelling in 20–40% of treated children [2]. Moderate or severe systemic events, such as fever 105 F or higher, febrile seizures and hypotonic–hyporesponsive episodes, have been reported [2]. Until now the adverse reaction of the vaccine has not been fully elucidated.

So far, fundamental to preventing safety problems with vaccines is the assurance that any vaccines for public use are made using Good Manufacturing Practices (GMP) with prerelease lot-testing for purity and potency. Furthermore, manufacturers must submit samples of each vaccine lot and results of their in-house tests for potency and purity to national control authorities before releasing them into clinical use. Conventional animal toxicity tests have been performed to detect the general toxicity of vaccines because the remaining toxicity of vaccines has the potential to cause adverse reactions. For example, the animal body weight decreasing test is the most commonly used control test to evaluate the toxicity of vaccines [3]. The body weight loss of vaccine-treated mice should be less than 10% of the whole weight for the regulation of safety of the vaccine. Although a correlation of the results from the body weight decreasing test with the vaccine's toxicity has been shown [4], a greater understanding of the molecular mechanisms involved in the development of vaccine toxicity needs to be clarified.

☆ This work was supported by Grants-in-Aid from the Ministry of Health, Labour and Welfare, Japan.

* Corresponding author.
E-mail address: kyama@nih.go.jp (K. Yamaguchi).

[1] These authors contributed equally to this work.

0264-410X/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2008.06.086

Table 1
Sequence of primers used in this study

Symbol	Primer concentration	Forward primer sequence	Reverse primer sequence
S100A8	200 nM	5′-TGCCCTCAGTTTGTGCAGAATA-3′	5′-CCAACGCAAGGAACTCTTCGA-3′
S100A9	200 nM	5′-GACCACAGGCACGGCAAAG-3′	5′-CCCAGCCCCAGAACCAAG-3′
CCL2	200 nM	5′-TTTCCACAACCACCTCAAGCA-3′	5′-TAAGGCATCACATTCCAAATCACA-3′
MMP8	200 nM	5′-AACAACCAATGCTGGAGATACGAC-3′	5′-GCATCAATTCTACAGTTTATTCCTGGG-3′
MCPT9	100 nM	5′-GTGCCAAGGCATGTCCAGAA-3′	5′-CCAGCAGTAGCCTTCCTCTCATT-3′
MMP9	50 nM	5′-AGGTGGATCCCCAGAGCGT-3′	5′-TTGGTAGTGAAAGACGTTGTGTGAG-3′
BEST5	200 nM	5′-TGTGCATGTATCTCCCTGTAGAATGA-3′	5′-TTTAATACACAGATCACGTTTCCATACACT-3′
MX2	200 nM	5′-AAGGAACATAGTGACACCAGTGAGAAG-3′	5′-GGACAGGGCCAGCTTAACCA-3′
IRF7	100 nM	5′-TGCAGCGTGAGGGTGTGTC-3′	5′-TCATCGTAGAGACTATTGGTGCTAGACA-3′
IFI27L	50 nM	5′-AAGGTCCTCGTGGGCTCTACAG-3′	5′-TTATCAGAGGAAGGAGTACAAATGCTTAC-3′
IFIT3	100 nM	5′-CCCTGAAGGATCTCATGCAGTTT-3′	5′-TCCGTTCTGCGTTTGCTGTT-3′
CYP2E1	200 nM	5′-GAAAGCGTGTGTGTGTTGGAGA-3′	5′-AACTGTGACAGGACTGAGGTCGAT-3′
NGP_predicted	100 nM	5′-GTTAATCGGGCCATAGAGGCATA-3′	5′-CAGGAGGTGGAGTGGCACTTAG-3′
GAPDH	200 nM	5′-GTGAAGCTCATTTCCTGGTATGACA-3′	5′-TTCTTACTCCTTGGAGGCCATGTAG-3′

Gene expression profiling is a unique way to characterize how cells are affected by alterations in pathologic conditions. The measurements of gene expression levels upon exposure to toxicants can be used both to provide information about the mechanism of action of the toxicant and to form a "genetic signature" for the identification of toxic products. The development of high-quality gene arrays has allowed this technology to become a standard tool in molecular toxicology. Recently, the field of toxicogenomics has validated the concept of gene expression profiles as "signatures" of toxicant classes. These signatures have effectively directed the analytical search for predictive biomarkers of toxicant effects and contributed to the understanding of the dynamic alterations in molecular mechanisms that are associated with toxic responses.

Many studies of gene expression profiles have been reported in the toxicology field. For example, Hamadeh et al. reported the patterns of gene expression in liver tissue taken from rats exposed to different chemicals [5]. Moreover, DNA microarray assays have been applied as an evaluation method for analyzing side effects of medicines [6].

Although DNA microarray analysis was used to detect the toxicity of pharmaceuticals as shown above, vaccine-related genes have not been evaluated using comprehensive gene expression analysis. The principle of nucleic acid hybridization is not new, however, microarrays have opened the way for the parallel detection and analysis of expression patterns of thousands of genes in a single experiment. Moreover, the sensitivity of microarrays allows for the detection of subtle differences that are much harder to detect with subtraction or other molecular methods. For a better understanding of the molecular toxicology regarding vaccines, the DNA microarrays are ideal methods.

In this study, we tried to identify the "genetic signatures" for the toxicants in the pertussis vaccine using DNA microarray analysis, and subsequently developed a gene expression detection system for pertussis vaccine toxicity.

2. Materials and methods

2.1. Animals

8-week-old male Wistar rats weighing 160–200 g were obtained from SLC (Tokyo, Japan). All animals were housed in rooms maintained at 23 ± 1 °C, with 50 ± 10% relative humidity, and 12-h light/dark cycles for at least 1 week prior to the test challenge.

2.2. Vaccines and toxin

Reference pertussis vaccine (reference vaccine; RE) is a lyophilized whole cell preparation of pertussis organisms treated for inactivation with formaldehyde, and has been used for National Quality Control Tests on pertussis vaccine in Japan since 1981. It was reconstituted in 12 ml of physiological saline, and 5 ml was injected intra-peritoneally (IP). From the conventional leukocyte promoting test, RE is predicted to contain the pertussis toxin at the concentration of 5 μg/ml. PV is a purified pertussis vaccine (a generous gift of The Chemo-Sero-Therapeutic Research Institute, Kumamoto, Japan) and contains inactivated pertussis toxin. Pertussis toxin (PT) (Wako chemicals, Osaka, Japan) and PV were adjusted to 5 μg/ml (PT concentration). RE, PV, and PT were injected into rats IP (5 ml/rat). Five milliliters of saline (SA) were injected as a control.

2.3. Body weight decreasing test

The rat body weight decreasing test was performed according to the Minimum Requirements of Biological Products [7]. We injected 5 ml of samples into the rat peritoneum for the treatment and rat weight was measured for 1 week after injection.

2.4. RNA preparation

Rats were sacrificed to obtain the whole lung, kidney, brain, and the lateral left lobe of the liver. Organs were immediately frozen in liquid nitrogen for storage. Thawed tissue was homogenized and

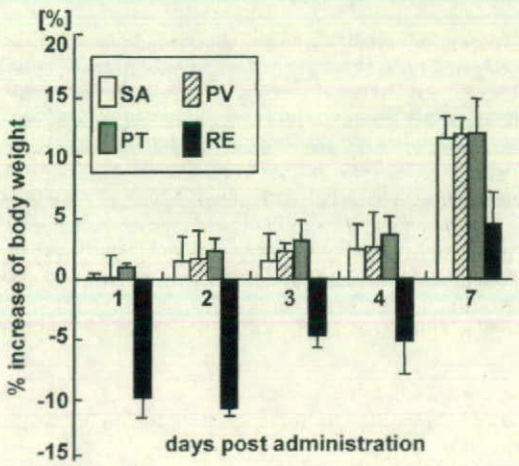

Fig. 1. Biological analysis of the vaccine-treated animals. The effects of RE, PT, PV, SA-treatment were measured using rat body weight decreasing toxicity tests. All animals were weighed at days 0, 1, 2, 3, 4 and 7. Changes in rat body weight were assessed at the percentage increase or decrease, and are indicated by the mean change ± S.D. of three independent experiments.

mixed with an ISOGEN reagent (NIPPON GENE, Tokyo, Japan). Total RNA was prepared from the lysate in accordance with the manufacturer's instructions. Poly(A)+ RNA was prepared from total RNA with a Poly(A) Purist Kit (Ambion, Austin, TX), according to the manufacturer's instructions.

2.5. Microarray preparation and expression profile acquisition

For the microarray analysis, three rats per group were treated with RE, PT, PV, and SA, and four organs from each group were analyzed on day 1–4 post-treatment. In total, 48 samples from each organ were analyzed for this experiment. A set of synthetic polynucleotides (80-mers) representing 11,468 rat transcripts and including most of the RefSeq genes deposited in the NCBI database (MicroDiagnostic, Tokyo, Japan) was arrayed on aminosilane-coated glass slides (Type I; Matsunami, Kishiwada, Japan) with a custom-made arrayer [8,9]. Poly(A)+ RNA (1.5 μg) was labeled with SuperScript II (Invitrogen, Carlsbad, CA) and Cyanine 5-dUTP for each sample or Cyanine 3-dUTP (PerkinElmer, Boston, MA) for a rat common reference RNA (MicroDiagnostic). Labeling, hybridization, and washes of microarrays were performed with a Labeling & Hybridization Kit (MicroDiagnostic) according to the manufacturer's instructions. The rat common reference RNA was purchased as a single batch and labeled as an aliquot with Cyanine-3 for a single microarray side by side with each sample labeled with Cyanine-5. Hybridization signals were measured using a GenePix 4000A scanner (Axon Instruments, Whipple Road Union City, CA) and then processed into primary expression ratios ([Cyanine 5-intensity obtained from each sample]/[Cyanine 3-intensity obtained from common reference RNA], which are indicated as 'median of ratios' in GenePix Pro 3.0 software (Axon Instruments)). Normalization was performed for the median of ratios (primary expression ratios) by multiplying normalization factors calculated for each feature on a microarray by the GenePix Pro 3.0 software.

2.6. Data analysis

Data processing and hierarchical cluster analysis were performed using Excel (Microsoft, Redmond, WA) and a MDI gene expression analysis software package (MicroDiagnostic). The primary expression ratios were converted into $\log_2$ values ($\log_2$ Cyanine-5 intensity/Cyanine-3 intensity) (designated log ratios) and compiled into a matrix (designated primary data matrix). To predict the most obvious differences obtained from cluster analysis of the primary data matrix, we extracted genes with $\log_2$ ratios over 1 or under −1 in at least one sample from the primary data matrix and subjected them to two-dimensional hierarchical cluster analysis for samples and genes. To identify genes demonstrating significant changes in expression, we extracted genes by *t*-test between SA and PT-treated samples ($p < 0.01$). Among the extracted genes, we further selected genes that exhibited differences greater than 0.75 between mean averages of log ratios for the two sample groups (SA and PT).

2.7. Histology

Vaccine-treated lungs were harvested from rats and fixed in Bouin's Solution (SIGMA, St. Louis, MO) and 4% (w/v) paraformaldehyde at 4 C for 48 h. After fixation, tissues were dehydrated through a series of graded alcohols and xylene and embedded in paraffin. Chilled paraffin blocks were cut into 4–6 μm sections, which were floated onto glass slides, dried overnight, and stained with hematoxylin and eosin (HE). Three rats per group, treated with RE, PT, PV, and SA, were analyzed on day 1 post-treatment.

Table 2
Pertussis vaccine toxicity related genes in lung

day 1	day 2	day 3	day 4
NM_053587	L18948	NM_053587	NM_053587
L18948	NM_053587	L18948	L18948
X68312	NM_022221	NM_022221	NM_022221
NM_019323	NM_022177	M57441	NM_012580
NM_053822	AF209976	NM_031530	XM_222255
M21782	NM_022524	XM_221486	NM_019296
M15402	NM_017227	XM_216379	XM_222462
NM_031055	NM_031766	NM_013154	XM_224300
Y07704	AF017638	AF062038	M15402
NM_017028	NM_012620	NM_019153	L19998
XM_215121	AF035951	NM_053822	U24441
NM_130743	XM_221486	NM_019323	NM_031315
M57441	Y07704	NM_031055	Y07704
NM_031530	XM_215208	J02585	NM_053373
X55180	XM_222462	NM_017055	NM_017183
XM_221061	NM_019296		NM_017061
AF313411	U24441		M57441
XM_220059	NM_017028		NM_031530
XM_234508	XM_215121		AB015877
NM_053769	X52711		XM_216379
NM_022221	NM_153733		XM_235354
NM_031619	NM_139186		XM_218820
D50568	NM_053373		NM_139257
J02627	NM_017061		NM_019214
XM_236646	NM_013154		NM_012620
	L31883		AF015953
	M57441		X55180
	XM_216379		XM_215121
	NM_022604		NM_017028
	L40364		X52711
	NM_031612		U02553
	NM_053687		NM_053769
	NM_019153		NM_022266
	NM_019323		NM_012862
	NM_053822		AF062038
	XM_236646		NM_031512
	NM_053929		NM_013154
	Z78279		X63369
	D50568		NM_017260
	J02627		J04035
	NM_012880		NM_019153
	J02585		AI104238
			U06436
			NM_022604
			NM_057103
			NM_053687
			NM_053822
			NM_019323
			XM_236646
			J02585
			NM_031348
			J02627
			NM_017134
			NM_012582

Gene expression patterns in SA- and PT- treated lungs were examined by microarray analysis. From the filtration with the expression ratios over 0.75 in $\log_2$ scales on average ($P < 0.01$), 25, 42, 15 and 54 genes were extracted on day 1, 2, 3 and 4, respectively.

2.8. Quantitative RT-PCR analysis

Poly(A)+ RNA was used to synthesize first-strand cDNA using a First-strand cDNA Synthesis Kit (Life Science, Inc., St. Petersburg, FL), according to the manufacturer's instructions. Expression levels of toxicity-related genes were analyzed by Real-Time PCR using a 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA) with 7500 Fast System SDS Software Version 1.3. cDNA was amplified for Real-Time PCR using SYBR Green I (Molecular Probes, Inc.) to detect the PCR products. One microliter of 6-fold diluted cDNA was used in a 20-μl final volume reaction containing 10 μl SYBR Green® PCR Master Mix (Applied Biosystems), and forward and reverse primers. The 7500 Fast System was programmed to run an initial polymerase activation step at 95 C for 10 min followed by 40 cycles of denaturation (95 C for 15 s) and extension (60 C for 1 min). Product synthesis was monitored at the end of the extension step of each cycle. The primers for each gene were shown in Table 1. Gene expression values were normalized against rat *GAPDH*.

2.9. QuantiGene Plex assays

The QuantiGene Plex assays were performed according to the procedure of QuantiGene Plex Reagent System (Panomics), which is described in more detail elsewhere [10]. Briefly, 10 μl of starting poly(A)+ RNA (50 ng) was incubated for 10 min at 65 C, then mixed with 33.3 μl of Lysis Mixture, 40 μl of Capture Buffer, 2 μl of Capture Beads, and 2 μl of the target gene-specific probe set. For the lung homogenate, 40 μl of lung lysates were mixed with 33.3 μl of Lysis Mixture, 0.2 μl of Proteinase K, 2 μl of Capture Beads, and 2 μl of the target gene-specific probe set. Probe sets were heated for 5 min prior to use. Each sample mixture was then dispensed into an individual well of a Capture Plate. The Capture Plate was sealed with foil tape and incubated at 54 C for 16–20 h. The hybridization mixture was transferred to filter plate, and the wells were washed three times with 200 μl of wash buffer. Signals for the bound target mRNA were developed by sequential hybridization with branched DNA (bDNA) amplifier, and biotin-conjugated label probe, at 48 C for 1 h each. Two washes with wash buffer were used to remove unbound material after each hybridization step. Streptavidin-conjugated phycoerythrin (SAPE) was added to the wells and incubated at room temperature for 30 min. The luminescence of each well was measured using a Luminex 100 microtiter plate luminometer (Luminex). Two replicate assays measuring RNA directly (independent sampling $n=6$ for mRNA, $n=3$ or 5 for lysate) were performed for all described experiments. The 13 targeted genes and *GAPDH* mRNA were quantified, and the ratio of the target genes to *GAPDH* mRNA was calculated.

3. Results

3.1. Vaccine-treated animals showed severe weight loss

In order to screen the general toxicity of vaccine, we analyzed the body weight of vaccine treated-animals. For this analysis, 5 ml of the vaccine was injected into the peritoneum of rats, and the weight loss of the treated animals was analyzed for 7 days. As shown in Fig. 1, the reference whole-cell vaccine (RE)-treated rats experienced severe decreases in body weight. The RE-treated rats lost

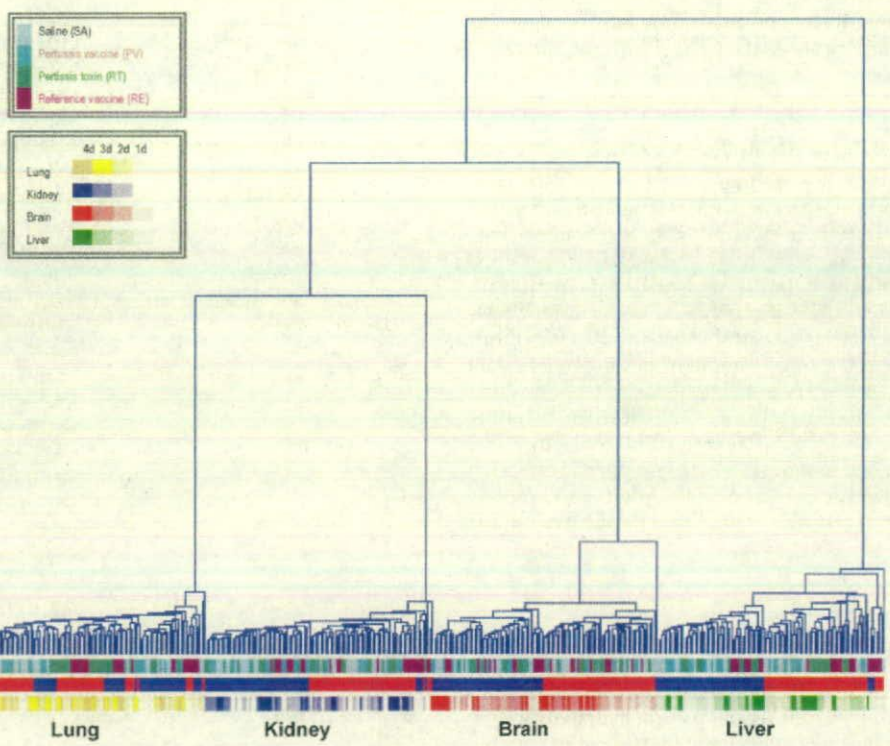

Fig. 2. Gene expression profiles obtained from RE, PT, PV and SA-treated rats. The gene expression of four organs (kidney, lung, brain and liver) was assembled in the order obtained from the results of the two-dimensional hierarchical clustering analysis. The dendrogram indicates the relationship between the samples after clustering analysis; the *y*-axis of the dendrogram depicts Euclid square distance as the dissimilarity coefficient. Color bars in pale blue, blue, green, and violet represent SA, PV, PT and RE, respectively. Blue and red bars show two independent gene profiling experiments for Exp. 1 and Exp. 2, respectively. Color bars in yellow, blue, red, and green represent lung, kidney, brain, and liver, respectively.

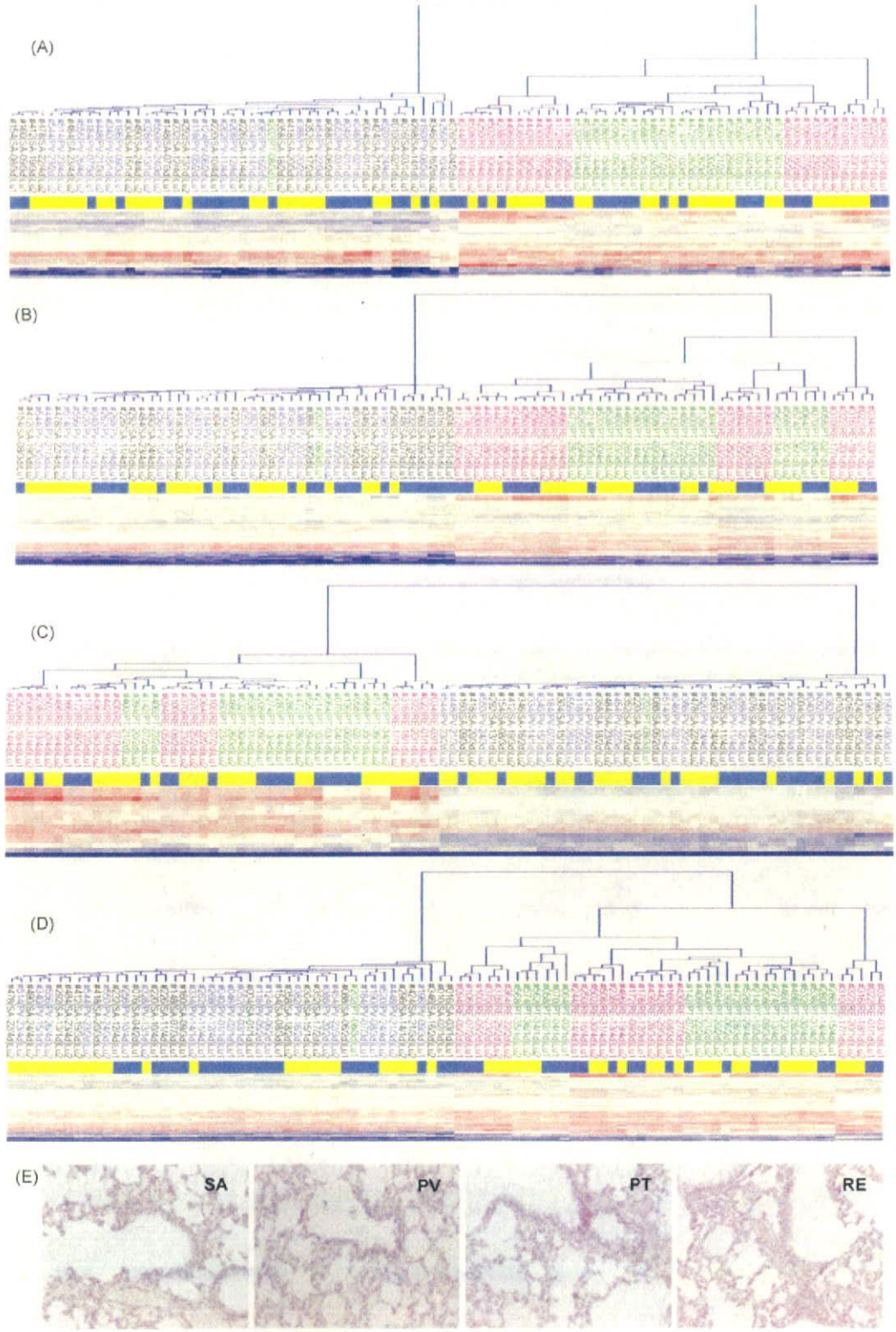

Fig. 3. Gene profiling analysis by the contrast genes between PT- and SA-treated-lungs. Specific genes, which showed large alterations between PT and SA selected by the SN method (see Section 2) included 25, 42, 15 and 54 genes at day 1 (A), day 2 (B), day 3 (C) and day 4 (D), respectively. These genes are assembled in the order obtained from the results of the two-dimensional hierarchical clustering analysis. The dendrogram for each day indicates the relationship between the samples (RE; red, PT; green, PV; blue, SA; black). Blue and yellow bars show two independent gene profiling experiments for Exp. 1 and Exp. 2, respectively. (E) The left lobes of RE, PT, PV and SA-treated lungs at day1 were stained with HE.

10% of the body weight within 1 day. Two days after treatment, the RE-treated rats started gaining weight, but it was still reduced at day 4 versus saline (SA)-treated rats. In contrast, the acellular pertussis vaccine (PV)-treated and pertussis toxin (PT)-treated rats demonstrated normal weight gain during the testing period similar to SA-treated rats.

3.2. Microarray analysis of the organ from vaccine-treated rats

To evaluate the effect of pertussis vaccine on gene expression, we prepared three rats per group in one experiment. RE-, PT-, PV-, and SA-treated animals were sacrificed to obtain the liver, lung, brain and kidney on days 1, 2, 3 and 4 post-administration. The animal experiments were performed twice with an interval of several weeks between the two experiments (designated as Exp. 1 and Exp. 2). Eventually, we obtained 96 independent tissue samples for the four time points. We labeled poly(A)+ RNA purified from the samples and a rat common reference RNA with Cyanine-5 and Cyanine-3, respectively, and hybridized to microarrays representing 11,468 transcripts. Hybridization signals were processed into expression ratios as $\log_2$ values (designated log ratios) and compiled into a matrix designated as the primary data matrix (see Section 2).

To predict the most obvious differences obtained from the clustering analysis of the primary data matrix, we extracted genes with log ratios over 1 or under −1 in at least one sample from the primary data matrix and subjected them to two-dimensional hierarchical clustering analysis for samples and genes. When the clustering

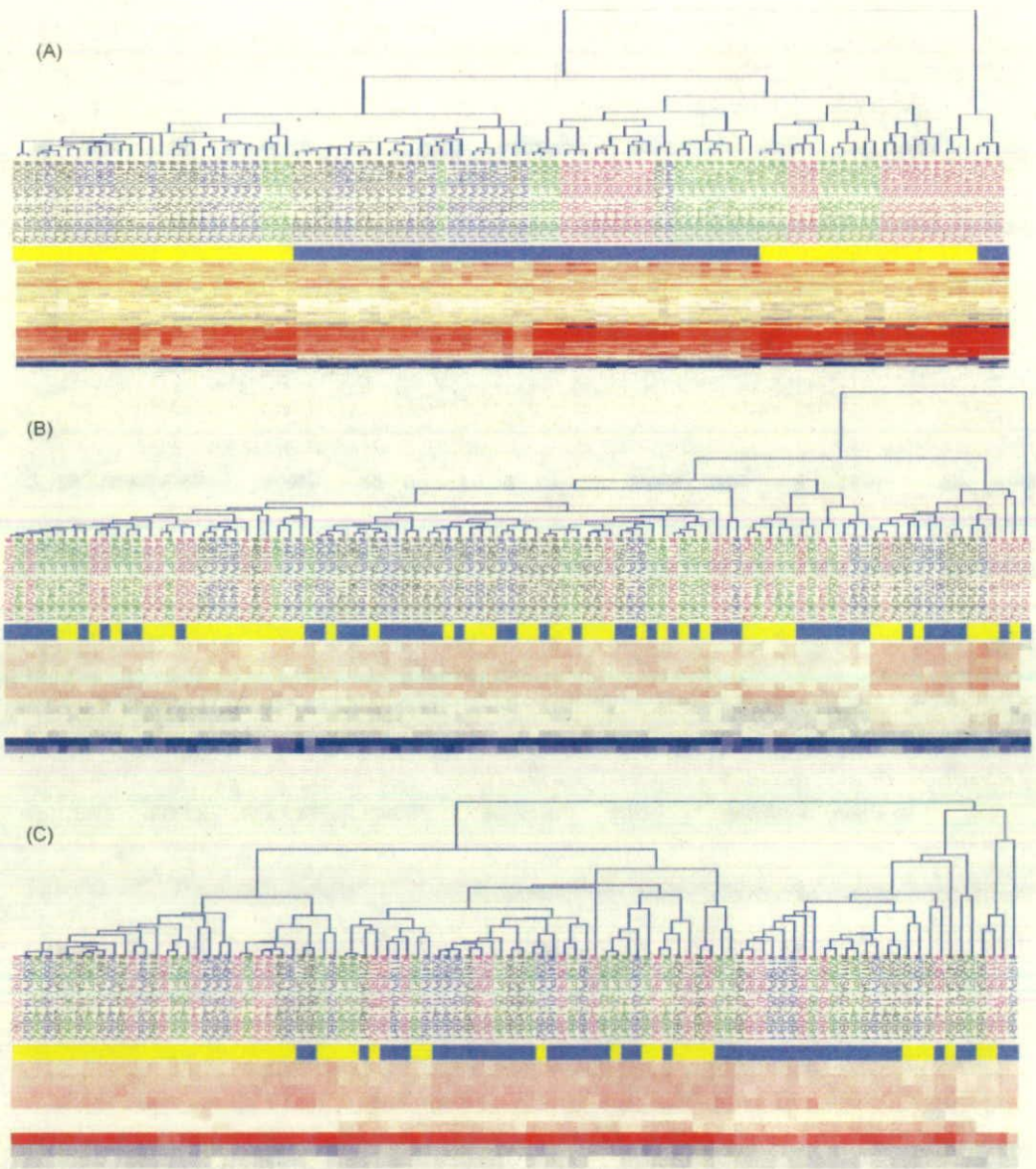

Fig. 4. Gene profiling analysis by the contrast genes between PT- and SA-treated-organs at day 1. Specific genes, which showed large alterations between PT and SA selected by the SN method (see Section 2) included 67, 14 and 9 genes from liver (A), kidney (B), and brain (C), respectively. These genes are assembled in the order obtained from the results of the two-dimensional hierarchical clustering analysis. The dendrogram indicates the relationship between the samples (RE; red, PT; green, PV; blue, SA; black). Blue and yellow bars show two independent gene profiling experiments for Exp. 1 and Exp. 2, respectively.

analysis for all samples was performed, the greatest four clusters corresponding to the individual organs were obtained as predicted. The dendrogram representing the liver, brain, lung and kidney comprised only smaller clusters (Fig. 2). The results obtained from this analysis indicate that the proceeding method of the microarray analysis is proper and by precipitating the toxicity-specific genes with statistical methods, a much sharper dendrogram will be expected.

3.3. Vaccine-treated lung had a clear dendrogram with vaccine toxicity-related genes

To analyze the effect of the pertussis vaccine and pertussis toxin, cluster analysis was performed with individual organs. We examined the gene expression patterns in SA- and PT-treated lungs. From the filtration with the expression ratios over 0.75 in $\log_2$ scale on average and the t-test ($P < 0.01$) between SA and PT-treated samples, 25, 42, 15 and 54 genes were extracted on day 1, 2, 3 and 4, respectively (Table 2). Based on these genes, the clustering analysis showed two big clusters for each day, that is, RE- and PT-treated samples established a cluster that is clearly distinct from the SA- and PV-treated samples (Fig. 3). Only one sample that was treated by PT (#202 in Fig. 3) demonstrated discordance, because of a technical mistake at the administration of the toxin. We also examined the other three organs, liver, kidney, and brain. Data from SA- and PT-treated samples were compared and filtered with the expression ratios over 0.75 in $\log_2$ scale on average and the t-test ($P < 0.01$) between SA and PT-treated samples. We extracted the marker genes on each day, and data from liver samples were reanalyzed with these genes (Fig. 4A). Unexpectedly, the resultant clusters did not necessarily correspond to the treated samples. When the expression patterns of kidney and brain were analyzed with the same procedure, the dendrograms of kidney and brain were confused as shown in Fig. 4(B and C), respectively. From these results, only the

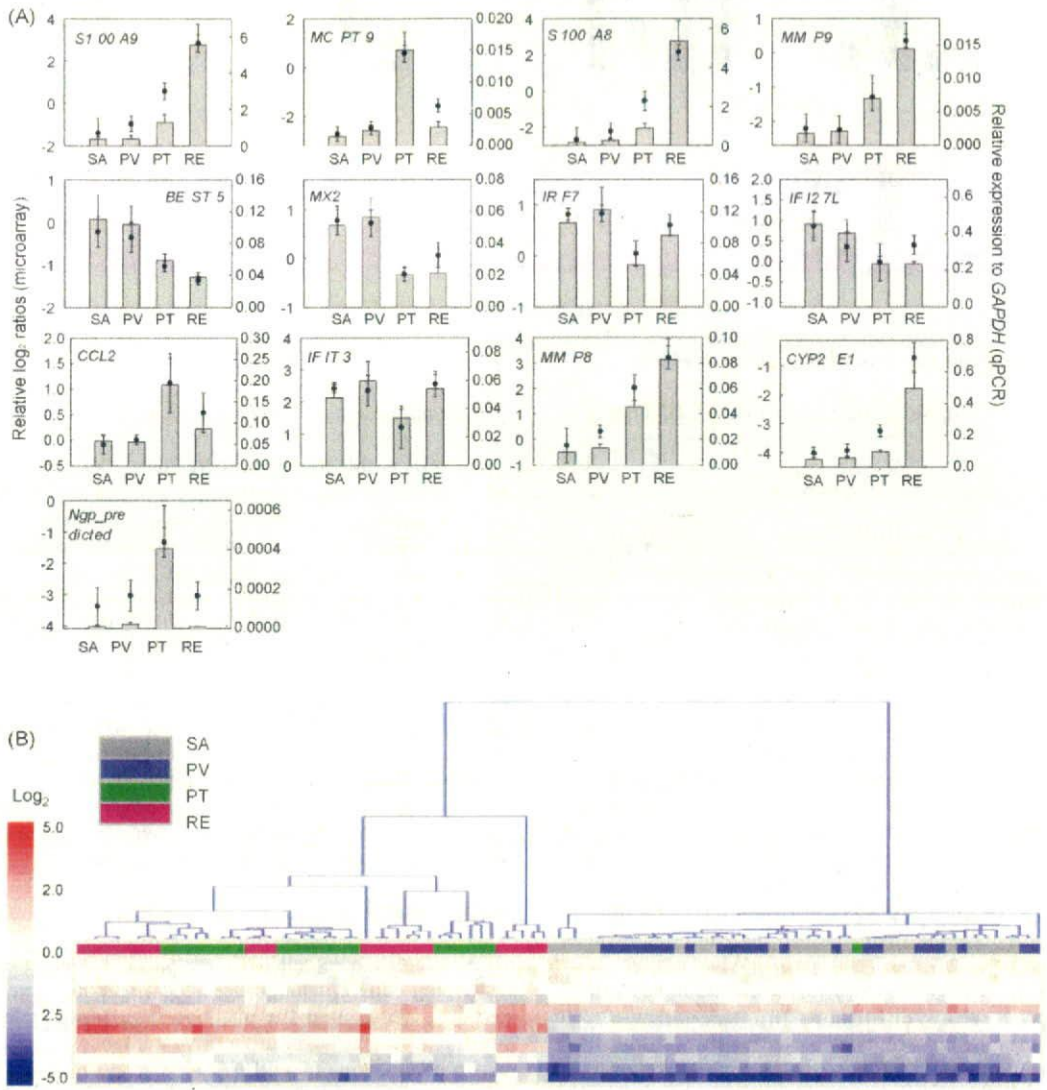

Fig. 5. Extraction of candidate genes for detection of the pertussis vaccine toxicity. (A) Expression of 13 genes from DNA microarray analysis was compared with Real-Time PCR data. Relative $\log_2$ ratios were extracted from the secondary data matrix for the 13 genes. Each symbol represents data from microarray analysis. The data of six animals in two independent experiments are shown. Expression values are indicated by the mean ratio ± S.D. Real-Time PCR analysis of the treated lung from individual animals is shown in bar graph. The data from six animals in two independent experiments are shown as the mean expression ± S.D. Quantitative PCR data was assessed relative to rat *GAPDH*. (B) Cluster analysis of day 1 samples with the selected 13 genes demonstrated a similar dendrogram to the 25 genes identified by microarray.

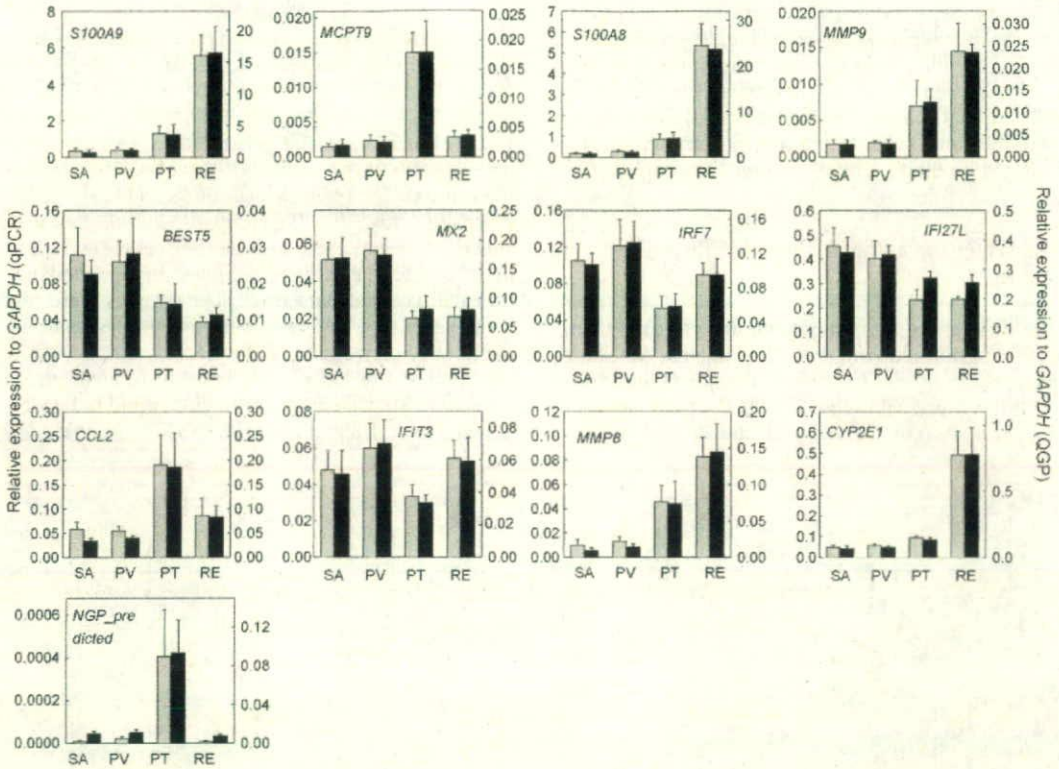

Fig. 6. Quantitative evaluation of toxicity-related genes by QuantiGene Plex. Quantification data of 13 genes with QuantiGene Plex by using poly(A)+ RNA (black bar) are compared with that of Real-Time PCR analysis (gray bar). The data from six animals in two independent experiments are shown as the mean expression ± S.D. Both Real-Time PCR data and QuantiGene Plex data were assessed relative to rat *GAPDH*.

lung is adequate to examine using gene expression-based cluster analyses.

3.4. Histological analysis of vaccine-treated lung

To analyze the effect of pertussis vaccine on the rat lung, we performed histological study on day 1 post-treatment. As shown in Fig. 3(E), SA-, PV-, PT-, and RE-treated lungs at day 1 demonstrated no significant changes in HE stained sections.

3.5. Quantification of toxicity-related genes

In the cluster analysis of lung, quite different genes were extracted on each 4 day (Table 2). Only five of the genes selected were in common, such as *S100A9* (NM_053587), *MCPT9* (NM_019323), *S100A8* (NM_053822), *CCL2* (M57441), and *MMP8* (NM_022221). These genes were thought to be strongly related to the toxicity of pertussis vaccine. Quantification analysis confirmed that 13 of the genes including *S100A9*, *MCPT9*, *S100A8*, *CCL2*, and *MMP8* at day 1 out of 25 array predicting transcripts demonstrated a good correlation between the microarray data and the quantitative gene expression data (Table 3 and Fig. 5A). When we performed the cluster analysis with these 13 genes, the cluster of RE- and PT-treated samples were clearly distinguished from SA- and PV-treated ones. Furthermore, the pattern of the dendrogram was similar to one with an array predicting 25 transcripts (Fig. 5B). In order to develop the gene expression detection system for pertussis vaccine toxicity using these 13 identified "genetic signatures", we applied the QuantiGene Plex assay with poly(A)+ RNA from the treated organs. As shown in Fig. 6, the results from the QuantiGene Plex showed good correlation with ones from Real-Time PCR analysis.

3.6. Quantitative detection of vaccine toxicity

Since it is difficult to prepare toxic vaccines for clinical use, we prepared samples containing various doses of RE vaccine (0.6–5.0 μg/ml) as a model of the accidental toxic vaccine. Three to five rats per group were analyzed at day 1 after sample administration. The lung lysates from the various sample-treated animals were assayed with QuantiGene Plex assay. As shown in Fig. 7, *S100A9*, *S100A8*, *MMP9*, *BEST5*, *MX2*, *IRF7*, *IFI27L*, *MMP8*, and *CYP2E1* were found to be suitable markers for RE vaccine toxicity. Expression of these nine genes was dependent on the dose of RE. On the other hand, *MCPT9*, *CCL2*, *IFIT3*, and *NGP_predicted* were not significantly altered with RE-treatment. These data correspond to that from poly(A)+ RNA (Fig. 6). Since these 13 "genetic signatures" can not only detect the toxicity of RE, but pertussis toxin as shown in Fig. 6, the assay system with the QuntiGene Plex should have the possibility to appropriately survey most pertussis vaccine toxicity and will become a new vaccine safety control method.

4. Discussion

Although whole-cell pertussis vaccine is effective to prevent whooping cough, it has caused local reactions such as redness, swelling, and pain at the injection site. However, little is known about the vaccine's comprehensive influences. In order to address this problem, we applied DNA microarray analysis and the quantification system of the specific gene expressions to analyze the vaccine's influence.

For DNA microarrays, we analyzed the gene expression profile in the liver, lung, kidney and brain from the treated rats, and selected the most suitable organ for the detection of vaccine toxicity. The