VOLUME 27 · NUMBER 17 · JUNE 10 2009

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Relapse-Related Molecular Signature in Lung Adenocarcinomas Identifies Patients With Dismal Prognosis

Shuta Tomida, Toshiyuki Takeuchi, Yukako Shimada, Chinatsu Arima, Keitaro Matsuo, Tetsuya Mitsudomi, Yasushi Yatabe, and Takashi Takahashi

From the Division of Molecular Carcinogenesis, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine; Division of Research and Development, Oncomics Co, Ltd; and Departments of Thoracic Surgery, Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital; and the Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan.

Submitted August 18, 2008; accepted December 16, 2008; published online ahead of print at www.jco.org on May 4, 2009.

Supported in part by a grant-in-aid for scientific research on priority areas, a grant-in-aid for scientific research (B), and a grant-in-aid for young scientists (B) from The Ministry of Education, Culture, Sports, Science and Technology of Japan and the Japan Society for the Promotion of Science.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Takashi Takahashi, MD, PhD, Division of Molecular Carcinogenesis, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan; e-mail: tak@med.nagoya-u.ac.jp.

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®).

0732-183X/09/2717-2793/$20.00

DOI: 10.1200/JCO.2008.19.7053

ABSTRACT

Purpose
In order to aid the development of patient-tailored therapeutics, we attempted to identify a relapse-related signature that allows selection of a group of adenocarcinoma patients with a high probability of relapse.

Patients and Methods
Whole-genome expression profiles were analyzed in 117 lung adenocarcinoma samples using microarrays consisting of 41,000 probes. A weighted voting classifier for identifying patients with a relapse-related signature was constructed with an approach that allowed no information leakage during each training step, using 10-fold cross-validation and 100 random partitioning procedures.

Results
We identified a relapse-related molecular signature represented by 82 probes (RRS-82) through genome-wide expression profiling analysis of a training set of 60 patients. The robustness of RRS-82 in the selection of patients with a high probability of relapse was then validated with a completely blinded test set of 27 adenocarcinoma patients, showing a clear association of high risk RRS-82 with very poor patient prognosis regardless of disease stage. The discriminatory power of RRS-82 was further validated using an additional independent cohort of 30 stage I patients who underwent surgery at a distinct period of time as well as with the Duke data set on a different platform. Furthermore, completely separate training and validation procedures using another data set recently reported by the Director's Challenge Consortium also successfully confirmed the predictive power of the genes comprising RRS-82.

Conclusion
RRS-82 may be useful for identifying adenocarcinoma patients at very high risk for relapse, even those with cancer in the early stage.

J Clin Oncol 27:2793-2799. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Lung cancer remains the leading cause of cancer death in industrialized countries, including Japan and the United States.[1,2] Adenocarcinomas, which account for more than 50% of non–small-cell lung cancer (NSCLC) cases, are the most frequent type of NSCLC with a heterogeneous nature in various aspects, including clinicopathologic and molecular features, and are showing an increasing trend.[3] The TNM clinical staging system has become the standard for predicting prognoses, however, the best hope for cure relies on surgical resection, which is considered as standard treatment for operable adenocarcinoma patients.[4] Nevertheless, 30% to 35% of surgically treated stage I patients eventually face relapse after the initial surgery, indicating the existence of a subgroup of patients clinically diagnosed as having early-stage disease, who actually have residual cancer cells undetectable by currently available imaging techniques used for staging.[4]

Although a number of prognostic biomarkers, such as altered expressions of oncogenes, and tumor suppressor genes have been proposed, the TNM staging system remains the standard method for predicting patient prognosis, indicating that such prediction may require information derived from the expression status of multiple genes and molecules. At the same time, the advent of microarray technology and completion of the genome project has made it possible to carry out genome-wide profiling of gene expressions.[5] These developments have provided an opportunity for establishing patient-tailored therapeutic strategies, leading to the identification of gene-expression profiles that are associated with the prognosis of individuals with lung cancer.[6-12] However, few prognostic prediction

classifiers have been validated with a sufficient number of independent cases.[12]

In this study, we report successful identification of a relapse-related molecular signature in adenocarcinomas through analysis of genome-wide expression profiles using a training set of 60 patients with lung adenocarcinomas. General applicability of the resultant classifier was successfully validated in a blind test set of 27 cases with stage I to III disease as well as with another independent cohort of 30 stage I patients. Moreover, additional validation using two data sets on a different platform further confirmed the predictive power of the genes comprising the relapse-related molecular signature.

PATIENTS AND METHODS

Patient Samples

Eighty seven lung adenocarcinoma samples from patients who underwent potential curative resection between December 1995 and August 1999 were collected at Aichi Cancer Center, Nagoya, Japan (herein referred to as data set I; online-only Appendix Table A1). An additional independent cohort of 30 adenocarcinoma samples from patients with pathologic stage (pStage) I disease were also collected at Aichi Cancer Center between February 2002 and December 2004 (herein referred to as data set II; Appendix Table A1). None of the 117 patients received adjuvant chemotherapy. General schedule of follow-up examinations was chest x-ray (every month for the first 3 months, and 3 months interval thereafter) and chest and abdominal computed tomography (CT; every year) until 5 years after surgery. Additional examinations, such as CT, bone scan, and brain magnetic resonance imaging, were also considered, if any signs of possible relapse were suspected. The median follow-up periods for patients alive at the last follow-up examination in data set I and data set II were 90 months (range, 64 to 108 months) and 64 months (range, 55 to 75 months), respectively. All tumor specimens were collected under approval from the institutional review boards of Aichi Cancer Center and Nagoya University with written informed consent from each patient.

Acquisition of Expression Profiles and Analysis of EGFR, p53, *and* K-ras *Mutations*

Double-stranded cDNA was synthesized from 500 ng of total RNA using Moloney murine leukemia virus-reverse transcriptase (Agilent Technologies, Palo Alto, CA) and poly dT primer incorporating the T7 promoter. Cy5-sample cRNA and Cy3-common reference cRNA were generated and hybridized to a Whole Human Genome oligo DNA microarray kit (G4112F, Agilent Technologies) with 41,000 distinct probes, which was scanned using an Agilent DNA microarray scanner (G2505B, Agilent Technologies), basically as described previously.[13] The mutation status of *EGFR*, *p53*, and *K-ras* was previously reported in the same set of patients.[13] All the microarray data and the pathologic and clinical data used for this study are available at Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/; accession number GSE13213). Cross-platform validation was carried out using the Duke[11] and Director's Challenge Consortium[12] data sets as detailed in the online-onlyAppendix.

Biostatistical and Bioinformatic Analyses

To identify a relapse-related signature using signals that were expressed above the background in at least 90% of samples, we used a weighted voting algorithm, in which each weight value was calculated as the signal-to-noise ratio, basically according to the detailed method that we described previously.[14] Kaplan-Meier survival curves and Cox proportional hazards model analyses (Stata, version 7.0; Stata Corp, College Station, TX) were used to analyze the relationships of the resultant relapse-related signature with overall and relapse-free survival. All statistical tests were two sided. The CLUSTER[15] program was used for average linkage hierarchical clustering of both genes and cases, and the TREEVIEW[15] program was used for display (http://rana.lbl.gov/EisenSoftware.htm).

RESULTS

Identification of Relapse-Related Signature

A schematic diagram of our strategy for constructing and validating a relapse-related signature in surgically treated lung adenocarcinoma patients is shown in Figure 1, which was formed with the intention of blocking any information leakage between the training and validation data sets. First, we divided expression profile data obtained from 87 patients into 60 training and 27 validation data sets, the latter of which was completely set aside during training. In order to identify a generic signature with clear associations with relapse in the training set of patients with lung adenocarcinomas, we selected 28 favorable samples (alive > 5 years after surgery without any evidence of relapse) and 21 fatal samples (dead in 5 years after initial surgery with evidence of relapse). The remaining 11 patients in the training set were excluded from analysis of a possible relapse-related signature, because of ambiguity related to the aggressiveness of their tumors,

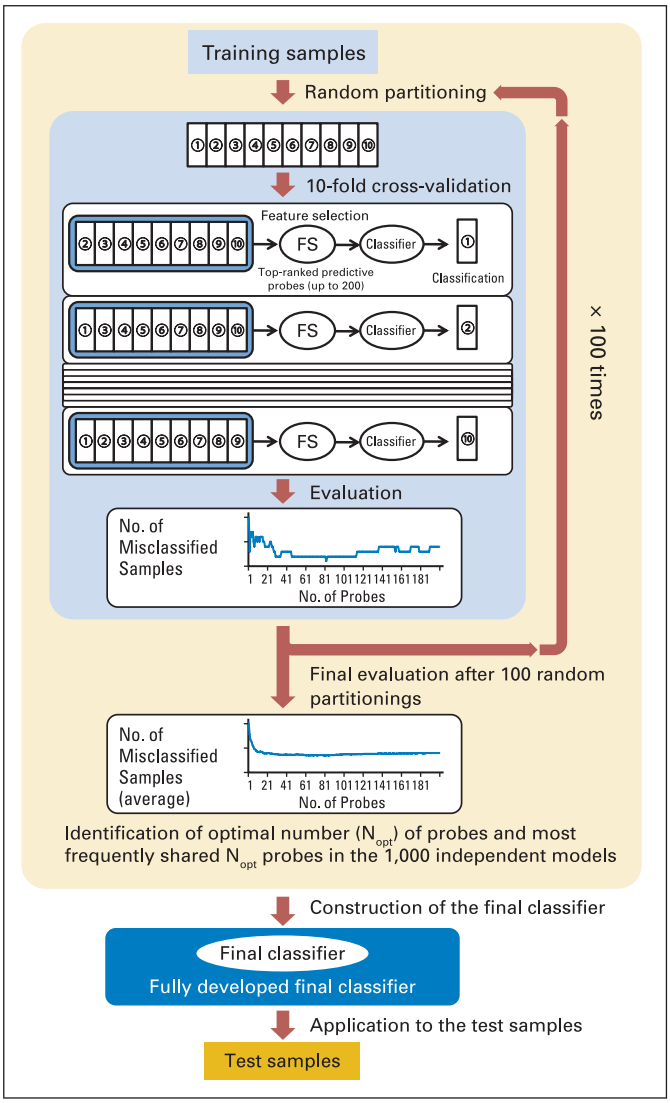

Fig 1. Schematic diagram of our training-validation stratergy for identifying replapse-related signature using 10-fold cross-validation procedures with 100 random partitions of the training data set.

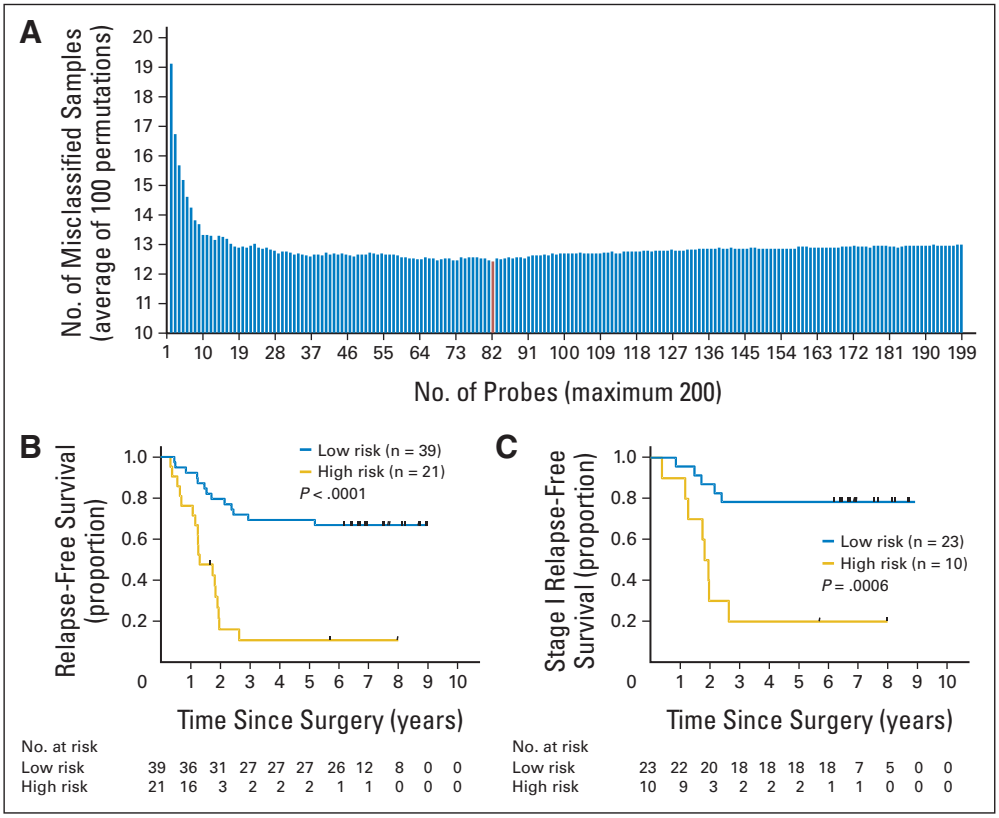

Fig 2. Results of the training procedure for identifying a relapse-related signature. (A) Results of our search for the optimum number of probes for defining a relapse-related signature. Kaplan-Meier survival curves were used to estimate survival in the training cohort. Relapse-free survival curves for patients in (B) all stages and (C) stage I.

which were five who survived for more than 5 years with some signs of relapse during follow-up, five who died of cancer after surviving for more than 5 years, and one who died within 5 years without evidence of relapse.

Of the 41,000 probes in the entire genome microarray, 23,828 passed the initial filtering criteria for selecting informative probes, and were then ranked according to a signal-to-noise metric and used to identify a relapse-related signature that could best distinguish patients who died with relapse from those cured by surgery. The learning errors for each model, to which increasing numbers of the predictive probes were applied, were calculated using 10-fold cross-validation and repeated with new randomly partitioned data sets 100 times. Thus, 1,000 independent sets consisting of up to 200 predictive probes each were selected for constructing a relapse-related signature-based classifier. As a result, 82 predictive probes were found to yield the fewest numbers of learning errors (Fig 2A), and the group of 82 probes most frequently shared among each of the 1,000 independent sets of 82 predictive probes was identified as a relapse-related signature (hereafter referred to as RRS-82; online-only Appendix Table A2). RRS-82 was able to distinguish patients with a very poor prognosis when all stages or only stage I were considered (Figs 2B and 2C for relapse-free survival and online-only Appendix Fig. A1 for overall survival). There were no associations of RRS-82 with the presence of *EGFR*, *K-ras*, or *p53* gene mutations, none of which showed any prognostic significance (Appendix Fig. A2).

Validation of RRS-82 in the Test Cohort of Data Set I

To evaluate the robustness of RRS-82, we analyzed its discriminatory power using a completely blinded data set of 27 adenocarcinomas. Results with the validation data set indicated that RRS-82 could distinguish between patients with high and low risks of recurrence and death. Relapse-free survival was significantly different between the two groups ($P = .0003$; Fig 3A), and the proportions of relapse-free patients in the high- and low-risk groups were 38% and 78%, respectively, after 2 years. In the high-risk group, the overall survival rate after surgical resection was also significantly lower than that in the low-risk group ($P = .026$; Fig 3B). It was of note that all stage I patients, who were predicted as high-risk based on RRS-82, experienced relapse within 5 years, and died during the follow-up period (Figure 3C for relapse-free survival; $P = .0008$; Fig 3D for overall survival; $P = .043$; both by log-rank test). Interestingly, Kaplan-Meier curves for both relapse-free and overall survival showed tendencies to have modest associations with pathologic disease stage ($P = .15$ for relapse-free survival and $P = .18$ for overall survival) among patients in the low-risk group but not in patients with high-risk RRS-82 (online-only Appendix Fig. A3). The presence of a high risk signature of RRS-82 was not associated with site of relapse (online-only Appendix Table A3).

Further Validation of RRS-82 With an Additional Independent Cohort of pStage I Patients

Further validation of the predictive power of RRS-82 in early-stage patients was conducted using another completely independent cohort of 30 stage I adenocarcinomas in patients who underwent surgery during a different period of time (data set II). RRS-82 was again shown capable of predicting which stage I patients were at extreme high risk (Figs 4A and 4B). In the combined validation cohort

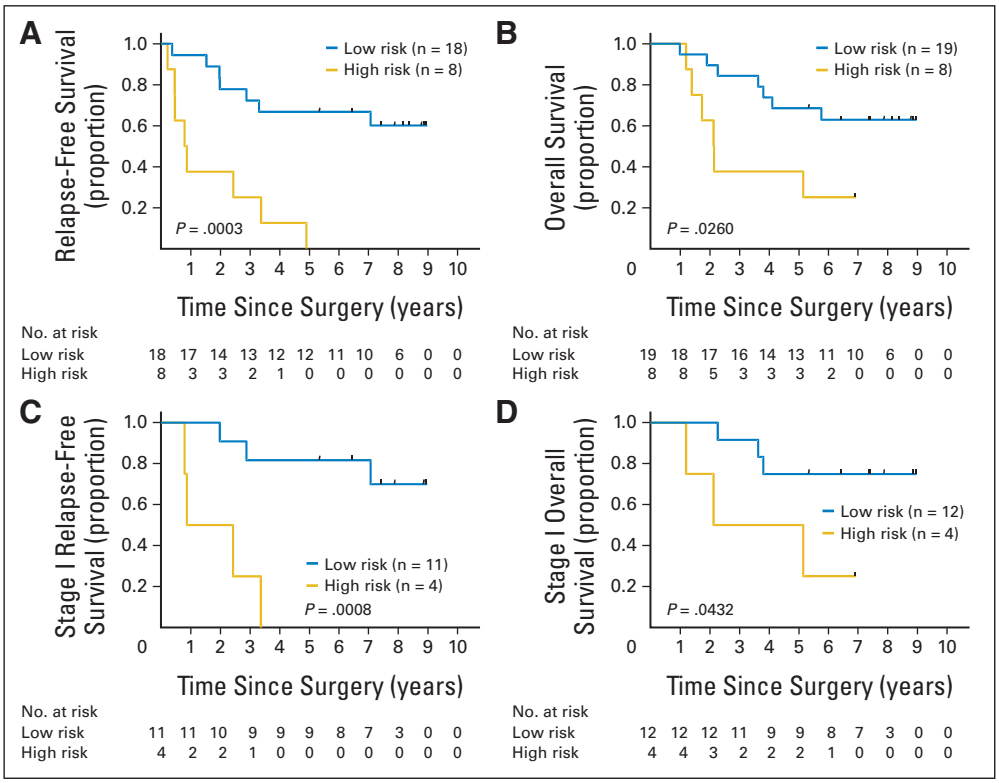

Fig 3. Validation of the RRS-82 signature with the use of completely blinded data set of 27 patients. Relapse-free survival curves for (A) all stages, (C) stage I. Overall survival curves for (B) all stages and (D) stage I.

consisting of 46 stage I cases (16 and 30 from datasets I and II, respectively), Kaplan-Meier survival curves based on RRS-82–based predictions were markedly different, showing relapse-free survival in 74% and 10% of patients with low- and high-risk signatures, respectively ($P < .0001$; Fig 4C). Overall survival was also significantly worse in the high-risk group as compared with the low-risk group ($P = .002$; Fig 4D). Data for patients in all stages are shown in online-only Appendix Figure A4. Multivariate Cox regression analysis of the combined validation data sets, in which the results of RRS-82–based predictions were considered as one of the variables, revealed that RRS-82 was highly predictive and independent of disease stage for both relapse-free survival ($P < .001$) and overall survival ($P = .005$; Table 1).

Confirmation of Predictive Capability of RRS-82 Using Two Additional Data Sets With a Different Platform

The robustness of RRS-82 for predicting survival of patients with lung adenocarcinomas was further validated using a completely independent Duke University data set of 39 lung adenocarcinomas. We conducted an unsupervised hierarchical clustering based on the expression profiles of the 46 genes, which corresponded to those constituting RRS-82 (Appendix Table A4). Thirty-nine adenocarcinomas were clearly clustered into two distinct subsets (Fig 5A), with significantly different postoperative survival results shown ($P = .028$; Fig 5B). The vast majority of genes corresponding to those related to relapse in RRS-82 showed a higher expression in patients in cluster 2, who had a poor prognosis, supporting the general applicability of RRS-82 for lung adenocarcinomas.

We further confirmed the predictive capability of the gene set constituting RRS-82 with a different approach by utilizing recently reported large training-testing, multisite data sets (Fig 5C). Using the University of Michigan data set consisting of 75 alive and 102 dead patients, we calculated each weighted value for 31 genes, which corresponded to the gene set constituting RRS-82, as the signal-to-noise ratio and then applied it to the 104 Memorial Sloan-Kettering samples, all of which had valuable information regarding relapse. The resultant RRS-82-based classifier built on the University of Michigan data set was able to predict patients at high risk in the Memorial Sloan-Kettering validation data set (Fig 5D). Taken together, these results demonstrated the predictive power of the gene set constituting RRS-82 for identifying patients at high risk for disease recurrence. Since the 31 genes in the set were selected based only on the presence of corresponding genes between the two distinct platforms, our findings suggest that potential future development of an optimally downsized classifier with sufficient predictive power based on RRS-82 is possible.

DISCUSSION

In this study, we identified a molecular signature, termed RRS-82, which was significantly associated with relapse and death in patients with adenocarcinomas of the lung. Based on the RRS-82 signature, we were able to construct a prognosis prediction classifier, which may ultimately aid in patient-tailored selection for therapeutic strategies. The robustness of the RRS-82 signature was successfully validated through application in four attempts with two independent Nagoya data sets as well as with the Duke and Director's Challenge Consortium data sets. Notably, the RRS-82–based classifier clearly distinguished patients with very poor prognosis from those with favorable outcome, including the duration of relapse-free survival, even in stage

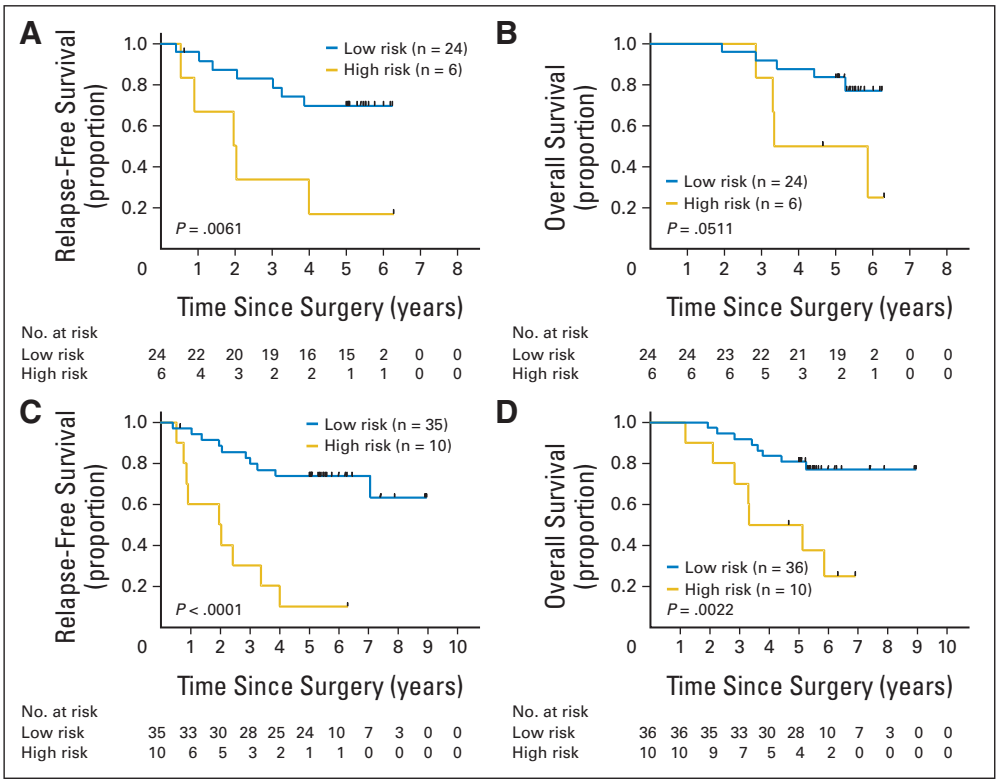

Fig 4. Independent validation of the RRS-82 signature using an additional independent cohort of 30 patients with stage I disease. (A) Relapse-free survival curves and (B) overall survival curves. Kaplan-Meier survival curves were used to estimate (C) relapse-free survival and (D) overall survival in the 46 stage I patients from data sets I and II.

I cases. These findings suggest that patients with the high-risk RRS-82 signature, who are overlooked using current diagnostic procedures for staging because of the inability of detection, are likely to have minimal residual disease. We previously reported that a 25-peak proteomic signature could also identify patients with very unfavorable outcome after surgery with curative intent at the protein level,[14] similarly to the present RRS-82 signature. Taken together, these findings support the notion that patients with very poor prognosis are certainly predictable even in stage I cases and that inclusion of molecular signature-based prognosis predictions, which take molecular and biologic characteristics manifested as signatures into consideration, may improve our capabilities for evaluating each patient with the ultimate aim of better therapeutic options.

Several studies have presented evidence supporting a model in which the propensity to metastasize reflects the predominant genetic/epigenetic state of a primary tumor, rather than the emergence of rare cells with a metastatic phenotype.[16-18] In this regard, it is interesting that disease stage at surgery appeared to have a modest tendency to affect patient outcome only in patients with a low-risk RRS-82 signature and not in those with a high-risk signature. A similar tendency was consistently observed in our previous proteomic analysis using matrix-assisted laser desorption/ionisation time of flight mass spectrometry, in which a 25-peak-based prediction model was constructed.[14] These findings therefore suggest a potential difference in biologic aggressiveness between the groups with high- and low-risk RRS-82 signatures.

Table 1. Univariate and Multivariate Cox Regression Analysis for the Combined Test Cohort (n = 57)

		Univariate			Multivariate		
Variable	Unfavorable/Favorable	Hazard Ratio	95% CI	*P*	Hazard Ratio	95% CI	*P*
Relapse-free survival (n = 56)*							
Age	> 61/≤ 61	0.68	0.32 to 1.47	.331	0.91	0.41 to 2.02	.817
Sex	Male/female	1.46	0.68 to 3.10	.329	1.19	0.54 to 2.60	.668
Stage	II-III/I	2.41	1.05 to 5.54	.038	2.00	0.84 to 4.72	.115
RRS-82	High risk/low risk	5.48	2.50 to 12.0	< .001	4.92	2.17 to 11.2	< .001
Overall survival (n = 57)							
Age	> 61/≤ 61	1.00	0.44 to 2.32	.991	1.21	0.50 to 2.91	.668
Sex	Male/female	1.61	0.70 to 3.74	.265	1.33	0.55 to 3.19	.526
Stage	II-III/I	2.56	1.04 to 6.32	.041	2.15	0.84 to 5.47	.106
RRS-82	High risk/low risk	3.68	1.58 to 8.56	.003	3.60	1.48 to 8.77	.005

*Information of relapse was not available in a single case.

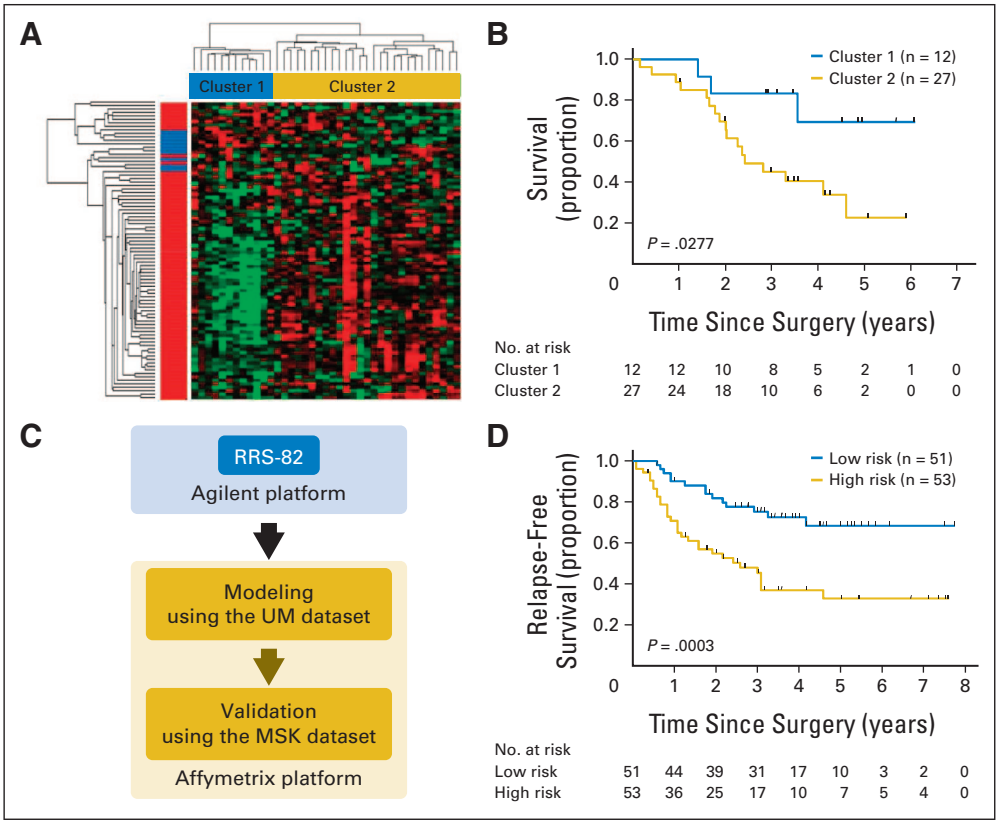

Fig 5. Results of (A) unsupervised hierarchical clustering analysis and (B) Kaplan-Meier survival curves for clusters I and II of the Duke data set. Schematic diagram showing further verification of RRS-82 constituents using another completely (C) independent training-validation data sets of 177 University of Michigan (UM) patients and 104 Memorial Sloan Kettering (MSK) patients, and (D) relapse-free survival curves for MSK patients.

The highly predictive nature of our RRS-82 signature, especially in terms of risk of relapse, may have been accomplished by our strategy used in the identification process, which paid special attention to relapse-free duration in a training cohort with high quality follow-up data. In fact, relapse within 5 years after surgery was observed in 80% and 90% of the patients with a high-risk RRS-82 signature in the training and combined validation cohorts with stage I disease, respectively. Although relapse-free survival data were not available for the 50 gene–based prediction classifier presented by the Michigan group[8] or the "A method" model by the Director's Challenge Consortium,[12] fatal outcome within 5 years after surgery was observed in approximately 55% and 60%, respectively, of those patients. In addition, the high-risk Duke metagene signature composed of nine metagene groups corresponding to 133 probes[11] was reported to correctly predicted relapse in 69% and 79% in their American College of Surgeons Oncology Group (stages I and II) and Cancer and Leukemia Group B (stages I to III) validation cohorts, respectively. Interestingly, the constituents of the RRS-82 signature do not have a significant overlap with other predictive signatures thus far reported by us and others.[8,9,11,12,19-23] Such variability among studies is commonly observed in molecular signatures for class prediction, and we suspect that it may reflect the use of different platforms for expression profiling and/or existence of distinct genes with similar predictive information, because of the presence of similarly coregulated genes that do not necessarily have similar biologic and/or biochemical properties.[24] For example, *PSMD12*, *FIP1L1*, and *UBE2V2*, included in RRS-82, are a part of the cluster six-gene set reported by the Director's Challenge Consortium, while *SMARCE1* in RRS-82 is included in the cluster 10-gene set. Additional analyses using the Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/) and Gene Ontology (http://www.geneontology.org/) databases identified only a few common pathways and networks containing predictive gene sets in such studies (examples shown in the Data Supplement). However, those results may not be surprising, since all of these studies including our own were not aimed at identifying functionally relevant gene sets or pathways associated with differences in clinical behavior such as relapse after surgery.

A number of negative results have been reported in regard to the benefits of adjuvant chemotherapy in patients with early-stage lung cancers,[25-28] although we believe that those do not preclude the potential clinical importance of molecular signature–based classification. Instead, such classification will likely add additional important information for patient-tailored evaluation of the nature of those diseases, considering that the current staging procedures, which rely on the measurement of disease spread by imaging techniques with insufficient power for detecting minute residual tumors, may be causing stage-migration of actual advanced cases into false early stages.

In conclusion, we succeeded in identifying a relapse-related molecular signature for use with patients diagnosed with adenocarcinomas, which was able to select those at extremely high risk for relapse, even in early-stage patients. In the field of breast cancer, a molecular signature–based prediction of surgically treated patients has been approved by the US Food and Drug Administration, and development of a similar useful means is urgent for lung cancer, which claims the highest number of lives each year. A future confirmatory study and clinical trial for patient-tailored adjuvant therapy with stratification according to the RRS-82 molecular signature are therefore warranted

to evaluate whether such selection may ultimately improve patient prognosis after surgery for this deadly cancer.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: Toshiyuki Takeuchi, Oncomics Co, Ltd (C) **Consultant or Advisory Role:** None **Stock Ownership:** Shuta Tomida, Oncomics Co, Ltd; Takashi Takahashi, Oncomics Co, Ltd, **Honoraria:** Tetsuya Mitsudomi, AstraZeneca Japan, Chugai Pharmaceutical, Astellas, Daiichi-Sanyo, Taiho, Eli-Lilly Japan, Kyowa hakko **Research Funding:** None **Expert Testimony:** Tetsuya Mitsudomi, AstraZeneca (U) **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Shuta Tomida, Takashi Takahashi
Financial support: Shuta Tomida, Takashi Takahashi
Administrative support: Takashi Takahashi
Provision of study materials or patients: Tetsuya Mitsudomi, Yasushi Yatabe
Collection and assembly of data: Shuta Tomida, Toshiyuki Takeuchi, Yukako Shimada
Data analysis and interpretation: Shuta Tomida, Chinatsu Arima, Keitaro Matsuo, Takashi Takahashi
Manuscript writing: Shuta Tomida, Takashi Takahashi
Final approval of manuscript: Shuta Tomida, Takashi Takahashi

REFERENCES

1. Jemal A, Siegel R, Ward E, et al: Cancer statistics, 2007. CA Cancer J Clin 57:43-66, 2007

2. McCracken M, Olsen M, Chen MS Jr, et al: Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. CA Cancer J Clin 57:190-205, 2007

3. Sun S, Schiller JH, Gazdar AF: Lung cancer in never smokers–a different disease. Nat Rev Cancer 7:778-790, 2007

4. Hoffman PC, Mauer AM, Vokes EE: Lung cancer. Lancet 355:479-485, 2000

5. Sotiriou C, Piccart MJ: Taking gene-expression profiling to the clinic: When will molecular signatures become relevant to patient care? Nat Rev Cancer 7:545-553, 2007

6. Garber ME, Troyanskaya OG, Schluens K, et al: Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci U S A 98:13784-13789, 2001

7. Bhattacharjee A, Richards WG, Staunton J, et al: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 98:13790-13795, 2001

8. Beer DG, Kardia SL, Huang CC, et al: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 8:816-824, 2002

9. Tomida S, Koshikawa K, Yatabe Y, et al: Gene expression-based, individualized outcome prediction for surgically treated lung cancer patients. Oncogene 23:5360-5370, 2004

10. Jones MH, Virtanen C, Honjoh D, et al: Two prognostically significant subtypes of high-grade lung neuroendocrine tumours independent of small-cell and large-cell neuroendocrine carcinomas identified by gene expression profiles. Lancet 363:775-781, 2004

11. Potti A, Mukherjee S, Petersen R, et al: A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. N Engl J Med 355:570-580, 2006

12. Shedden K, Taylor JM, Enkemann SA, et al: Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. Nat Med 14:822-827, 2008

13. Takeuchi T, Tomida S, Yatabe Y, et al: Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. J Clin Oncol 24:1679-1688, 2006

14. Yanagisawa K, Tomida S, Shimada Y, et al: A 25-signal proteomic signature and outcome for patients with resected non-small-cell lung cancer. J Natl Cancer Inst 99:858-867, 2007

15. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868, 1998

16. Ramaswamy S, Ross KN, Lander ES, et al: A molecular signature of metastasis in primary solid tumors. Nat Genet 33:49-54, 2003

17. Minn AJ, Gupta GP, Siegel PM, et al: Genes that mediate breast cancer metastasis to lung. Nature 436:518-524, 2005

18. Glinsky GV, Berezovska O, Glinskii AB: Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer. J Clin Invest 115:1503-1521, 2005

19. Lau SK, Boutros PC, Pintilie M, et al: Three-gene prognostic classifier for early-stage non small-cell lung cancer. J Clin Oncol 25:5562-5569, 2007

20. Chen HY, Yu SL, Chen CH, et al: A five-gene signature and clinical outcome in non-small-cell lung cancer. N Engl J Med 356:11-20, 2007

21. Lu Y, Lemon W, Liu PY, et al: A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. PLoS Med 3:e467, 2006

22. Bianchi F, Nuciforo P, Vecchi M, et al: Survival prediction of stage I lung adenocarcinomas by expression of 10 genes. J Clin Invest 117:3436-3444, 2007

23. Larsen JE, Pavey SJ, Passmore LH, et al: Gene expression signature predicts recurrence in lung adenocarcinoma. Clin Cancer Res 13:2946-2954, 2007

24. Fan C, Oh DS, Wessels L, et al: Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 355:560-569, 2006

25. Winton T, Livingston R, Johnson D, et al: Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. N Engl J Med 352:2589-2597, 2005

26. Douillard JY, Rosell R, De Lena M, et al: Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): A randomised controlled trial. Lancet Oncol 7:719-727, 2006

27. Strauss GM, Herndon JE, Maddaus MA, et al: Adjuvant chemotherapy in stage IB non-small cell lung cancer (NSCLC): Update of Cancer and Leukemia Group B (CALGB) protocol 9633. J Clin Oncol 24:365s, 2006 (suppl; abstr 7007)

28. Wakelee H, Dubey S, Gandara D: Optimal adjuvant therapy for non-small cell lung cancer–how to handle stage I disease. Oncologist 12:331-337, 2007

Acknowledgment

We thank Kiyoshi Yanagisawa at Nagoya University and Hirotaka Osada at Aichi Cancer Center Research Institute for their valuable discussion and critical reading of the manuscript.