



A gene expression-based single sample predictor of lung adenocarcinoma molecular subtype and prognosis

Helena Liljedahl¹ | Anna Karlsson¹ | Gudrun N. Oskarsdottir^{1,2} |
 Annette Salomonsson¹ | Hans Brunnström^{1,3} | Gigja Erlingsdottir^{4,5} |
 Mats Jönsson¹ | Sofi Isaksson¹  | Elsa Arbajian¹ | Cristian Ortiz-Villalón⁶ |
 Aziz Hussein⁷ | Bengt Bergman⁸ | Anders Vikström⁹ | Nastaran Monsef¹⁰ |
 Eva Branden^{11,12} | Hirsh Koyi^{11,12} | Luigi de Petris¹³ | Annika Patthey¹⁴ |
 Annelie F. Behndig¹⁵ | Mikael Johansson¹⁶ | Maria Planck^{1,2} | Johan Staaf¹ 

¹Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Medicon Village, Lund, Sweden

²Department of Respiratory Medicine and Allergology, Skåne University Hospital, Lund, Sweden

³Department of Pathology, Laboratory Medicine Region Skåne, Lund, Sweden

⁴Department of Pathology, Landspítali University Hospital, Reykjavik, Iceland

⁵Department of Laboratory Medicine, Department of Pathology, Skåne University Hospital, Malmö, Sweden

⁶Department of Pathology, Karolinska University Hospital, Stockholm, Sweden

⁷Department of Pathology and Cytology, Sahlgrenska University Hospital, Gothenburg, Sweden

⁸Department of Respiratory Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden

⁹Department of Pulmonary Medicine, University Hospital Linköping, Linköping, Sweden

¹⁰Department of Pathology and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

¹¹Respiratory Medicine Unit, Department of Medicine Solna and CMM, Karolinska Institute and Karolinska University Hospital Solna, Stockholm, Sweden

¹²Centre for Research and Development, Uppsala University/Region Gävleborg, Gävle, Sweden

¹³Thoracic Oncology Unit, Karolinska University Hospital and Department Oncology-Pathology, Karolinska Institute, Stockholm, Sweden

¹⁴Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden

¹⁵Department of Public Health and Clinical Medicine, Division of Medicine, Umeå University, Umeå, Sweden

¹⁶Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden

Correspondence

Johan Staaf, Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Medicon Village, SE-22381 Lund, Sweden.

Email: johan.staaf@med.lu.se

Funding information

BioCare; Cancerfonden; Fru Berta Kamprads Stiftelse; Sjöberg Foundation; Stiftelsen Jubileumsklinikens Forskningsfond mot

Abstract

Disease recurrence in surgically treated lung adenocarcinoma (AC) remains high. New approaches for risk stratification beyond tumor stage are needed. Gene expression-based AC subtypes such as the Cancer Genome Atlas Network (TCGA) terminal-respiratory unit (TRU), proximal-inflammatory (PI) and proximal-proliferative (PP) subtypes have been associated with prognosis, but show methodological limitations for robust clinical use. We aimed to derive a platform independent single sample

Abbreviations: AC, adenocarcinoma; CLAMS, Classification of Lung Adenocarcinoma Molecular Subtypes; DMFS, distant metastasis-free survival; FFPE, formalin-fixed paraffin-embedded; NCC, nearest centroid classification; NSCLC, non-small cell lung cancer; OS, overall survival; PI, proximal-inflammatory; PP, proximal-proliferative; SSP, single sample predictor; TRU, terminal-respiratory unit.

Helena Liljedahl and Anna Karlsson contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of Union for International Cancer Control.

Cancer; The National Health Services (Region Skåne/ALF)

predictor (SSP) for molecular subtype assignment and risk stratification that could function in a clinical setting. Two-class (TRU/nonTRU=SSP2) and three-class (TRU/PP/PI=SSP3) SSPs using the AIMS algorithm were trained in 1655 ACs (n = 9659 genes) from public repositories vs TCGA centroid subtypes. Validation and survival analysis were performed in 977 patients using overall survival (OS) and distant metastasis-free survival (DMFS) as endpoints. In the validation cohort, SSP2 and SSP3 showed accuracies of 0.85 and 0.81, respectively. SSPs captured relevant biology previously associated with the TCGA subtypes and were associated with prognosis. In survival analysis, OS and DMFS for cases discordantly classified between TCGA and SSP2 favored the SSP2 classification. In resected Stage I patients, SSP2 identified TRU-cases with better OS (hazard ratio [HR] = 0.30; 95% confidence interval [CI] = 0.18-0.49) and DMFS (TRU HR = 0.52; 95% CI = 0.33-0.83) independent of age, Stage IA/IB and gender. SSP2 was transformed into a NanoString nCounter assay and tested in 44 Stage I patients using RNA from formalin-fixed tissue, providing prognostic stratification (relapse-free interval, HR = 3.2; 95% CI = 1.2-8.8). In conclusion, gene expression-based SSPs can provide molecular subtype and independent prognostic information in early-stage lung ACs. SSPs may overcome critical limitations in the applicability of gene signatures in lung cancer.

KEYWORDS

gene expression, lung adenocarcinoma, molecular subtypes, prognosis, single sample predictor

1 | INTRODUCTION

Lung adenocarcinoma (AC) is the most frequent histological type of nonsmall cell lung cancer (NSCLC).¹ Compared to other NSCLC tumors, AC tumors have been associated with specific molecular and etiological traits, including a nonsmoking patient history and oncogenic driver alterations (eg, *EGFR* mutations and various fusion genes).²⁻⁵ In advanced-stage AC, immune checkpoint and tyrosine kinase inhibitors are now clinical routine. In surgically treated cases, chemotherapy remains the main adjuvant treatment option, guided by the TNM classification.⁶ Despite an overall favorable prognosis, surgically treated lung cancer is still associated with a high-risk of metastatic relapse, even for tumors of the lowest stage (Stage I).⁷ Based on the lack of significant survival benefit in Stage IA disease,⁸ adjuvant therapy was not recommended for this particular group, while Stage IB patients may receive treatment. Clearly, additional prognostic and predictive tools are needed to improve therapy decisions in surgically treated AC.

Surgically resected AC has been intensively studied using different high-throughput molecular profiling techniques. Gene expression-based studies have reported prognostic gene signatures and suggested the existence of molecular subtypes in AC.^{4,9-15} The TCGA study on AC concluded three transcriptional subtypes termed the terminal-respiratory unit (TRU), the proximal-inflammatory (PI) and the proximal-proliferative (PP) subtypes.^{4,9} These subtypes have been associated with different clinicopathological and molecular variables,

What's New?

New tools are needed in order to improve risk stratification and therapy selection in early-stage lung adenocarcinoma. Inherent differences in gene expression between adenocarcinoma subtypes could facilitate the development of such tools. The authors of this study derived platform-independent, single-sample predictors (SSP) of adenocarcinoma subtypes, based on gene expression. Derived SSPs successfully provided prognostic information in surgically treated stage I lung adenocarcinoma patients. The single-sample classifier was readily translated into assays applicable to archival tissue, indicating clinical utility. The findings highlight the clinical relevance of transcriptional signatures and gene expression predictors in lung adenocarcinoma, warranting their further investigation and development.

but also patient outcome.^{9,13} Specifically, the TRU subtype shows overrepresentation of patients with a nonsmoking history, tumors with *EGFR* mutations, tumors of lower stage, lower tumor proliferation in general, and importantly improved patient outcome.^{4,9,13} In contrast, both the PI and PP subtypes are associated with a patient smoking history and show features of often aggressive disease, including high frequencies of different nontargetable driver mutations,

higher proliferation and specific morphologic growth patterns.^{4,9,13} Recent large-scale analyses have demonstrated that, based on current treatment options (surgery with or without chemotherapy/radiotherapy), the robust prognostic power of the gene expression subtypes lies in the two-class distinction of TRU vs nonTRU samples and is mainly related to differences in expression of proliferation-associated genes.¹³

The current classification scheme for TRU, PI and PP subtypes involves classification of a new sample according to the nearest centroid classification (NCC) approach.⁹ While NCC type classifiers have been used extensively for classification of tumors (eg, the PAM50 classifier in breast cancer¹⁶), this type of classifier presents some limitations concerning the prediction of independent samples.^{13,17-19} Ideally, a single sample predictor (SSP) that does not require any preprocessing, is independent of gene expression platform and capable of predicting a single sample is desirable. In this context, predictors based on gene rules assessed on an intrasample basis have been proposed as a solution.^{17,18,20-22}

In our study, we aimed to derive SSPs of the TRU/nonTRU and TRU/PI/PP subtypes. Using machine-learning in 1655 AC cases and independent validation in 977 AC cases we developed a 36-gene SSP of the TRU/nonTRU subtypes. This SSP provided refined prognostic categorization of patients compared to the existing classification approach and also independent prognostic information in surgically treated Stage I AC. As a proof of concept, the SSP was translated into a NanoString nCounter XT assay and tested in archival tissue specimens (formalin-fixed paraffin-embedded [FFPE]) for prediction of disease relapse.

2 | MATERIALS AND METHODS

2.1 | Patient datasets

Twenty-two publicly available gene expression datasets (n = 2632 samples) were assembled. Forty-three samples overlapped between two datasets. Complete datasets (ie, all samples) were partitioned into an SSP training cohort (n = 17 datasets, n = 1655) or a validation cohort (n = 5 datasets, n = 977) (Figure 1). Partitioning was directed toward having training and validation cohorts with mixed technical platforms, and that the validation cohort should include both patients with only surgical treatment and patients with adjuvant chemotherapy, to allow relevant outcome analyses. Clinicopathological characteristics for training and validation datasets are outlined in Table 1, based on data from original studies. Mutation status (EGFR/KRAS, etc.) for these public cohorts is highly limited.

To test SSPs in FFPE, two cohorts were used. First, Fragments Per Kilobase of transcript per million mapped reads data were obtained from the GSE143486 dataset, representing 30 RNA sequenced Stage I ACs with no adjuvant therapy. Secondly, RNA from a cohort of 44 patients, diagnosed between 2006 and 2015 at different Nordic institutions, with surgically treated Stage I AC were collected. The 44 patients were selected using a “case-control” approach based on recurrence (locoregional and/or distant) to fit with the

NanoString multiplexing scheme. Twenty-three patients presented with recurrence within 5 years from diagnosis (56.5% Stage IA, 43.5% IB), and 21 were recurrence-free 5 years from diagnosis (76% Stage IA, 24% IB), thus forming two patient groups: (a) “poor” (case) and (b) “better” (control) outcome. Patients in the two groups were selected to balance gender, smoking status and original patient institution. The selection was verified through statistical testing, finding no statistically significant difference in gender, smoking status or Stage IA/IB (Fisher's exact test >0.05). Clinicopathological variables were collected from patient charts. Complete clinical mutation/gene fusion status was not available for this retrospective cohort.

2.2 | Pathology assessment of lung AC growth patterns

Assessment of histological growth patterns were assessed in patients from Karlsson et al³⁶ (GSE60644, n = 16) and Djureinovic et al³⁷ (GSE81089, n = 110) as reported in Salomonsson et al.⁴³ Briefly, all histological slides were reviewed for each case for investigation of growth patterns. The cases were graded according to predominant growth pattern: lepidic predominant (incl. minimally invasive AC) were classified as low grade, acinary and papillary predominant as medium grade and micropapillary and solid predominant or invasive mucinous AC as high grade (Reference 44 and the WHO classification from 2015).

2.3 | Preprocessing and TCGA NCC classification of gene expression data

Preprocessing of gene expression data was performed as described in Cirenajwis et al.²² Common gene symbols across the 22 datasets were extracted (n = 9659) to create a uniform expression matrix. To generate training/reference classes for the SSPs, NCC was performed for each dataset separately to assign each sample a TCGA subtype (TRU, PI or PP) based on the highest (Pearson) correlation as described elsewhere.^{9,22} In addition, a two-class constellation consisting of TRU vs nonTRU cases was also generated and used for SSP training.

2.4 | AIMS single sample classifier

The AIMS¹⁷ method was implemented using source scripts available from the GitHub repository (<https://github.com/meoyo/trainAIMS>). Training was performed on raw gene expression data (n = 9659 genes) from the 1655 training samples as outlined in Cirenajwis et al²² with the exception for the Lee et al and CLCGP datasets for which only normalized gene expression data could be obtained from public repositories. Briefly, the training cohort was trained against either the two-class (TRU/nonTRU) or three-class (TRU/PI/PP) constellations to generate the AIMS-based predictors referred to as SSP2 or SSP3, respectively. No further preprocessing of gene expression data was performed. To avoid unequal rule contribution in rule selection due to

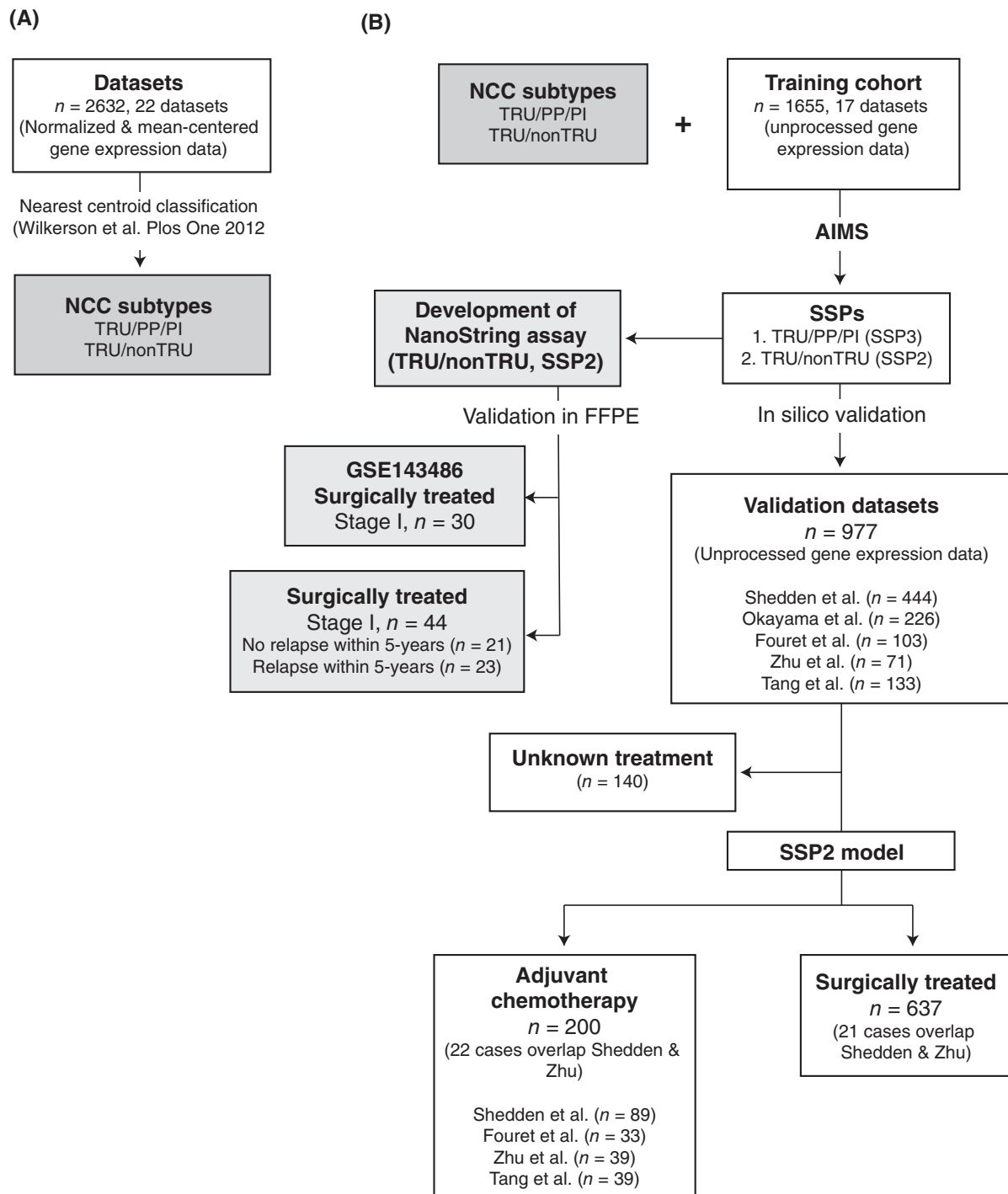


FIGURE 1 Flow-chart of study. A, Approach to derive molecular subtype training class through nearest centroid classification (NCC) of all datasets individually using the scheme reported by Wilkerson et al.⁹ For the two-class subtype approach, PP and PI subtypes were combined to a single nonTRU class. B, Training and validation scheme for deriving a two-class SSP for TRU/nonTRU (SSP2) and a three-class SSP for TRU/PI/PP subtypes (SSP3) based on the AIMS single sample method. Of the total 22 datasets included, 5 were reserved as independent validation datasets and were also used for evaluation of prognostic performance of the SSP models in both surgically treated only and adjuvantly treated patients. A patient overlap existed for the Shedden et al and Zhu et al cohorts. Patients overlapping were excluded from one cohort in survival analyses. An additional external validation of the SSP2 model was also performed in archival RNA from 44 Stage-I patients treated with surgery only, by pairing the SSP2 model with the NanoString nCounter XT technology

size differences across datasets for the merged training cohort, the AIMS algorithm applied a weighted form of rule selection provided by the R package "Rgtsp." The derived SSP2 and SSP3 models

will be available as an R package "Classification of Lung Adenocarcinoma Molecular Subtypes" (CLAMS) at Bioconductor (www.bioconductor.org).

TABLE 1 Datasets included in our study

Datasets	Total (N)	Accession	Platform	Sex: males (%)	Stage I (%)	OS	DMFS	Adj. chemo (N)	NCC status: TRU vs non TRU (%)	Cohort assignment
Chitale et al ^{23 a}	102	Chitale U133 2plus	Affymetrix	41	69	Yes	No	0	41	Training
CLCGP ^{24 b}	98	CLCGP	Illumina	48	44	Yes	Yes	0	34	Training
Bild et al ²⁵	58	GSE3141	Affymetrix	NA	45	Yes	Yes	0	36	Training
Lee et al ²⁶	63	GSE8894	Affymetrix	54	NA	No	Yes	0	38	Training
Tomida et al ²⁷	117	GSE13213	Agilent	51	68	Yes	No	0	40	Training
Hou et al ²⁸	45	GSE19188	Affymetrix	56	NA	Yes	No	0	31	Training
Lu et al ²⁹	60	GSE19804	Affymetrix	NA	58	No	No	0	38	Training
Wilkerson et al ⁹	116	GSE26939	Agilent	46	53	Yes	No	0	41	Training
Rousseaux et al ³⁰	85	GSE30219	Affymetrix	78	95	Yes	Yes	0	34	Training
Botling et al ³¹	106	GSE37745	Affymetrix	43	66	Yes	No	0	35	Training
Seo et al ³²	87	GSE40419	RNAseq	61	63	No	No	0	41	Training
Tarca et al ³³	77	GSE43580	Affymetrix	68	53	No	No	0	42	Training
Chen et al ³⁴	92	GSE46539	Illumina	17	NA	No	No	0	37	Training
Der et al ³⁵	127	GSE50081	Affymetrix	51	72	Yes	Yes	0	39	Training
Karlsson et al ³⁶	77	GSE60644	Illumina	42	88	Yes	No	0	40	Training
Djureinovic et al ³⁷	115	GSE81089	RNAseq	37	58	Yes	No	0	35	Training
TCGA ^{4 c}	230	TCGA	RNAseq	NA	NA	No	No	0	39	Training
Shedden et al ³⁸	444	Shedden	Affymetrix	50	62	Yes	Yes	89	37	Validation
Okayama et al ³⁹	226	GSE31210	Affymetrix	46	74	Yes	Yes	0	43	Validation
Fouret et al ^{40 d}	103	E-MTAB-923	Affymetrix	16	58	Yes	No	33	42	Validation
Zhu et al ^{41 e}	71	GSE14814	Affymetrix	52	59	Yes	Yes	39	35	Validation
Tang et al ⁴²	133	GSE42127	Illumina	51	67	Yes	No	39	38	Validation

^aSamples were divided into two cohorts based on the different Affymetrix platforms, U133A and U133 2plus. Only the latter subset was included in the analysis.

^bCLCGP, The Clinical Lung Cancer Genome Project (<http://www.uni-koeln.de/med-fak/clcgp/>).

^cThe Cancer Genome Atlas Network (TCGA).

^dData obtained from the "ArrayExpress" database (<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-923/>).

^ePresent dataset overlaps with Shedden et al (43 samples).

2.5 | Pathway analysis

Pathway analysis was performed using the PANTHER Classification System (<http://pantherdb.org>) and the overrepresentation test application to identify significant biological pathways covered by the SSPs. Default settings were used, and gene ontology terms with a false discovery rate adjusted Fisher's exact test $P < .05$ were considered significant.

2.6 | A NanoString SSP gene expression assay

To test the applicability of CLAMS in archival tissue (ie, RNA from FFPE tissue), a NanoString (www.nanostring.com) nCounter XT assay was designed based on the CLAMS genes for TRU/nonTRU (SSP2) prediction. RNA from FFPE tissue was extracted using the AllPrep DNA/RNA FFPE Kit (Qiagen, Hilden, Germany). A 300 ng RNA was

used in the nCounter XT CodeSet Gene Expression Assay and counts were generated on a SPRINT instrument after the manufacturer's instructions (NanoString Technologies, Seattle, WA). Four cartridges (12-sample) were run with 44 samples and 4 controls in total. Generated counts were background corrected and generated gene expression data were quality assessed as described.^{45,46} All samples passed quality thresholds.

2.7 | Statistical analysis

All statistical analyses were performed as two-sided tests using R (www.r-project.org). Classification performance (accuracy and balanced accuracy) for SSP2 and SSP3 were analyzed in each validation dataset separately. For analysis of individual gene expression across the molecular subtypes in the validation cohort ($n = 934$, no

overlapping samples), genes were ranked (from 1 to 9659) sample-wise using the function “rankGenes” provided by the R package “singscore” (version 1.5.0).⁴⁷

2.8 | Survival analysis

Survival analyses were performed using the R “survival” package (version 2.43.1) with overall survival (OS), distant metastasis-free survival (DMFS) or recurrence-free interval (for NanoString FFPE samples) as endpoints defined according to original studies (Table 1). Survival curves were compared using Kaplan-Meier estimates and the log-rank test. Hazard ratios (HR) were calculated through univariable or multivariable Cox regression using the “coxph” R function. Survival data was censored at 5 years to account for differences in follow-up time between validation datasets.

3 | RESULTS

3.1 | Clinical and molecular subtype characteristics of the patient cohort

Twenty-two reported gene expression datasets for lung AC (n = 2632 patients in total) were divided into a training cohort (n = 1655) and a validation cohort (n = 977) (Figure 1, Table 1). All datasets were classified on a per cohort basis, according to TCGA TRU/PP/PI subtypes by the NCC method. Similar proportions of subtypes were observed across the datasets (Figure 2A), despite substantial differences in, for example, distribution of tumor stage and technical platform (Table 1). This observation is an illustration of the inherent feature of NCC classification relying on gene centering across samples, which if not accounted for make sample classification cohort dependent.^{13,22} Still, TCGA NCC subtypes retained reported associated clinical features,^{4,9} including a higher proportion of Stage I tumors, never-smokers and patients with *EGFR* mutations in TRU-classified tumors in both training and validation cohorts (Supplementary Figure S1).

3.2 | Deriving SSPs for TCGA lung AC subtypes

Based on TCGA NCC subtypes for 1655 training samples and 9659 genes present across all datasets, AIMS¹⁷ was used to derive a two-class SSP for TRU vs nonTRU samples (SSP2) and a three-class SSP for the TRU/PI/PP subtypes (SSP3). SSP2 consisted of 18 gene rules (n = 36 unique genes), while SSP3 consisted of 47 gene rules per subtype (n = 141 gene rules in total, n = 259 unique genes) (Figure 2B; Supplementary Table S1). Reclassification of the 1655 training samples showed an accuracy of 0.85 for SSP2 and 0.82 for SSP3 vs TCGA NCC classifications, acknowledging the circular nature of this analysis.

Fifty-six percent (20/36) of the SSP2 model genes overlapped with the original 506 NCC genes, and 83% (15/18) of gene rules involved at least one NCC gene (Figure 2C). For SSP3, corresponding values were

56% (144/259 genes) and 79% (112/141 gene rules). Functional analysis of SSP gene rules were investigated by gene set enrichment analysis (Supplementary Table S2). For the SSP2 model, selected genes were strongly enriched for the cell cycle gene ontology process. For the SSP3 model, additional significant gene ontology processes, besides the cell cycle, included leukocyte migration and chemotaxis, extracellular matrix and structure, cell migration and localization terms.

For two cohorts (Karlsson et al³⁶ and Djureinovic et al³⁷) in the training dataset, we had access to reviewed pathology assessments of histological lung AC growth patterns (lepidic predominant, acinary and papillary predominant, and micropapillary and solid predominant) for 126 cases. Cross tabulation of SSP2 classifications vs these histological subtypes revealed that 87.5% of cases with lepidic predominant growth patterns were of the TRU subtype. For TRU cases, 14.3% had lepidic growth patterns, 71.4% acinary and papillary predominant growth patterns and 14.3% micropapillary and solid predominant growth patterns. Corresponding values for nonTRU cases were 1.3%, 63.6% and 35.1% (Fisher's exact test, $P = .001$). When excluding the small number of lepidic cases, corresponding values were 83.3% acinary and papillary predominant patterns, and 16.7% micropapillary and solid predominant patterns in TRU classified cases, and 64.5% and 35.5%, respectively, in nonTRU cases (Fisher's exact $P = .035$). These general patterns were retained also when viewing the individual datasets, although not reaching statistical significance due to lower numbers. Together, these results indicate an association of the SSP classifications with histological AC growth patterns.

3.3 | Validation of SSP2 and SSP3 as predictors of AC molecular subtype

SSP2 and SSP3 models were validated in 977 independent samples derived from five datasets analyzed by either Illumina or Affymetrix gene expression microarrays (Table 1, Figure 1). Importantly, the SSP2/SSP3 models do not rely on any preprocessing, thus new samples are classified independently based on raw data only.

Per validation dataset, the proportion of TRU classified samples for the NCC vs SSP2 model was first compared (Figure 2D). Differences were observed for specific datasets. The Okayama et al⁴⁸ dataset showed a notably higher fraction of SSP2-TRU samples than the NCC classifier, a pattern partly also present in the Fouret et al⁴⁰ and Tang et al⁴² datasets. Across all validation samples, an accuracy of 0.85 for SSP2 and 0.81 for SSP3 were observed when compared to NCC subtypes. Variations between individual datasets in accuracy were observed as outlined in Figure 2E.

3.4 | Clinical and molecular characterization of discordantly classified samples by SSP and NCC methods

In the 977-sample validation cohort, 137 cases (14%) were discordantly classified between the NCC (2-class) and SSP2 models.

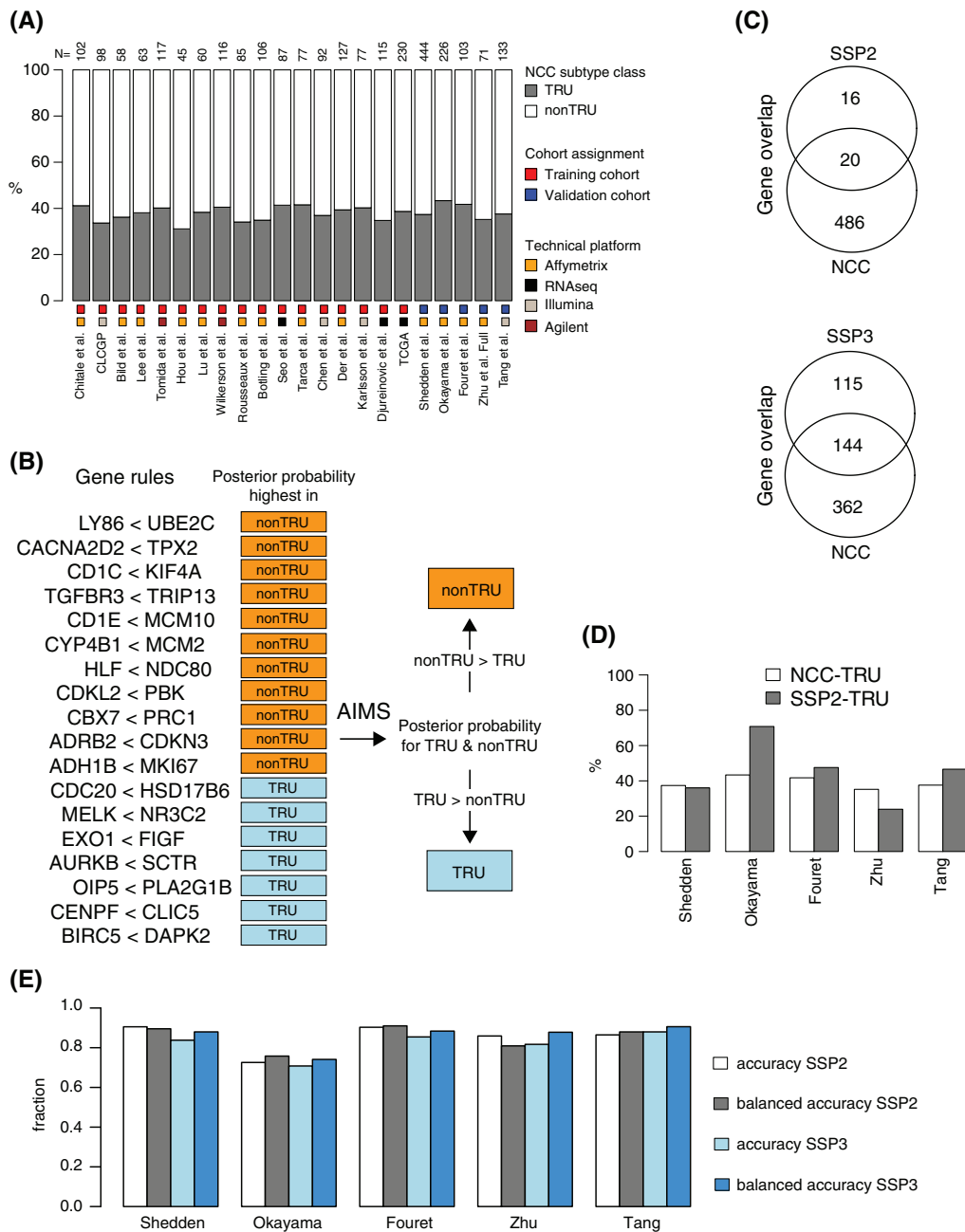


FIGURE 2 Training and validation of SSPs for prediction of molecular subtypes in lung adenocarcinoma. A, Proportion of TRU and nonTRU cases predicted by the NCC method⁹ per dataset in the study. For each dataset, assignment to training or validation cohort and technical gene expression platform is shown. Top-axis indicates dataset size. B, Schematic overview of the SSP2 classifier for TRU/nonTRU status based on training vs NCC subtype classes in the training cohort. The SSP2 classifier comprises 18 gene rules (pairs), that is, 36 genes. Gene rules are shown with indication of their highest posterior probability in the AIMS model. Based on all individual gene rule probabilities a final prediction is made. C, Overlap of genes in the SSP2 (top) and SSP3 classifiers vs the original NCC centroid genes from Wilkerson et al.⁹ D, Proportions of TRU classified cases in the five validation datasets for the NCC and SSP2 models, showing differences across datasets. E, Classification performance (accuracy and balanced accuracy) in the validation cohort for the SSP2 model vs TRU/nonTRU NCC classifier, and the SSP3 model vs the TRU/PI/PP NCC classifications

This subset of patients had a higher proportion of never-smokers (Fisher's exact test $P = .02$) and Stage II tumors ($P = .0007$) compared to concordantly classified patients (Supplementary Figure S2A). To further dissect discordant cases, all validation cases were given a label corresponding to the class assignment by the two methods (NCC_{status}-SSP_{status}). This formed four groups: (a) TRU-TRU (=concordant TRU), (b) TRU-nonTRU, (c) nonTRU-TRU and (d) nonTRU-

nonTRU (=concordant nonTRU). Analysis of the *MKI67* gene expression ranks (*Ki67*, a well-established proliferation-related gene) indicated higher proliferation in the concordant nonTRU group as compared to the concordant TRU group. For discordantly classified cases, the *MKI67* expression pattern was consistent with the SSP2 classification, meaning, for example, lower expression in nonTRU-TRU compared to TRU-nonTRU (Supplementary Figure S2B). The majority

of discordant cases had an intermediate correlation (eg, between 0 and 0.2) to the TCGA NCC centroids (TRU/PI/PP) and were weakly separated by the NCC method (Supplementary Figure S2C,D).

3.5 | Molecular subtype prediction by NCC and SSP models vs patient outcome

Based on the prognostic analyses of TCGA NCC subtypes reported by Ringner et al,¹³ outcome analysis was restricted to the TRU/nonTRU context. Patients in the validation cohort were stratified by treatment status into a prognostic group ($n = 616$ unique patients treated with surgery alone) and an adjuvant chemotherapy treated group ($n = 178$ unique patients) (Figure 1, Table 1). The four NCC_{status}-SSP_{status} groups were used to address the question of which classifier appeared more “clinically meaningful” for discordantly classified cases based on survival outcome.

In the prognostic arm, concordant TRU patients (including patients of all stages) had an improved OS compared to concordant nonTRU patients, with a 5-year survival rate of 82% vs 54%, respectively (Figure 3A). The subset of patients ($n = 104$ in total, 16.9%) with discordant NCC (2-class)/SSP2 class (TRU-nonTRU and nonTRU-TRU) showed differences in OS ($P = .0015$, log-rank test) (Figure 3B). The patient group classified as TRU by NCC and nonTRU by SSP2 (TRU-nonTRU) had a similar survival pattern as the concordant poor outcome nonTRU patient group (Table 2, Figure 3A). These patients had a significantly increased risk of death (HR = 2.90; 95% confidence interval [CI] = 1.37-6.32; $P = .005$), as compared to the concordant TRU patient group. The reverse was observed for the nonTRU-TRU patient group, which showed a survival pattern similar to that of the concordant better outcome TRU group (Table 2, Figure 3A). Thus, based on patient survival, discordantly classified patients seem to be more “accurately” classified by the SSP2 model. Similar trends were observed for DMFS, considering that this analysis included patients of all stages (Table 2, Figure 3C,D).

In the adjuvant chemotherapy patient group, survival differences between the concordant TRU and nonTRU groups were less pronounced at 5 years postsurgery, with an OS rate of 57% and 45%, respectively (Figure 3E). Herein, 12.9% ($n = 23$) had discordant subtype labels. The slightly lower discordance rate may be due to that adjuvant chemotherapy is selectively given to high-risk patients, which conceptually should more often be intrinsic nonTRU. Given the low number of discordant cases, especially for the DMFS endpoint (Figure 3F), larger datasets are needed to determine whether the findings in surgically treated patients translate to adjuvant treated discordant patients.

3.6 | Association of patient outcome with SSP2 prediction in surgically treated Stage I disease

In surgically treated Stage I patients from the validation cohort, the SSP2 predicted TRU patient group was significantly associated with a better OS (89% 5-year survival) with a hazard ratio of 0.29 (95% CI = 0.18-0.46; $P < .0001$), as compared to the nonTRU patient group

(65% 5-year survival) (Figure 4A). Of Stage IA patients with OS data, 73% were TRU-classified, while for Stage IB only 48% were TRU-classified. The better OS of TRU-classified cases remained significant also in subgroups of Stage IA or IB patients (log-rank $P = .007$ and 90% 5-year survival, and $P = .0002$ and 85% 5-year survival, respectively, and was also observed independently of age and gender in Stage IA (HR = 0.36; 95% CI = 0.17-0.75; $P = .007$) or Stage IB (HR = 0.26; 95% CI = 0.13-0.52; $P = .0001$).

In TRU-classified patients, Stage IA/IB was not associated with differences in OS (log-rank $P = .31$). SSP2 classification added independent prognostic information in a multivariable Cox regression model for OS including age, Stage IA/IB, and gender as covariates for Stage I patients (TRU HR = 0.30; 95% CI = 0.18-0.49; $P < .0001$). For DMFS, SSP2 also significantly stratified patients into better (TRU, 5-year DMFS = 79%) and worse (nonTRU, 5-year DMFS = 62%) outcome (Figure 4B), supported by univariable (TRU HR = 0.48; 95% CI = 0.31-0.74; $P = .0008$) and multivariable Cox regression analyses using age, Stage IA/IB and gender as covariates (TRU HR = 0.52; 95% CI = 0.33-0.83; $P = .006$).

3.7 | An assay based on SSP2 applicable to FFPE tissue

To test our SSP2 model in FFPE tissue, we created an R-based implementation (referred to as CLAMS) and paired it with the NanoString nCounter XT technology to create a “complete” assay. We applied this assay to FFPE RNA from 44 Stage-I ACs treated surgically only. Twenty-one of the patients were metastasis-free 5 years after diagnosis (forming a “better” prognosis group), while the remaining 23 patients had a relapse (loco-regional/distant) within 5 years (representing a “poor” outcome group).

CLAMS prediction of the 44 cases classified 39% (17/44) as TRU and 61% (27/44) as nonTRU (Supplementary Table S3, including all data). Of TRU-classified samples 88% were Stage IA, while for nonTRU 52% were Stage IA. Hierarchical clustering of raw counts for CLAMS genes across the 44 cases confirmed the TRU/nonTRU CLAMS subgroups (Figure 4C), further supported by gene rule fulfillment for predicted cases (Supplementary Figure S3). CLAMS classification was next compared to the two patient prognosis groups (no relapse/relapse, Supplementary Table S3, Figure 4D). Of the TRU classified patients, 71% were metastasis free after 5 years, in contrast to only 33% of the nonTRU classified patients (accuracy for CLAMS groups vs relapse status was 0.68). The specificity (no relapse in TRU group) was 0.71, sensitivity (relapse in nonTRU group) was 0.67 and the positive predictive value was 0.78 (relapse in poor group also being nonTRU). For discrepant cases predicted as TRU but with relapse, we observed slightly elevated expression of *MKI67* and lowered Napsin A expression (NAPSA), consistent with a more nonTRU like phenotype. Moreover, nonTRU cases without relapse showed elevated expression of *MKI67*, as compared to TRU cases, signaling why these were likely classified as nonTRU by CLAMS (Figure 4E).

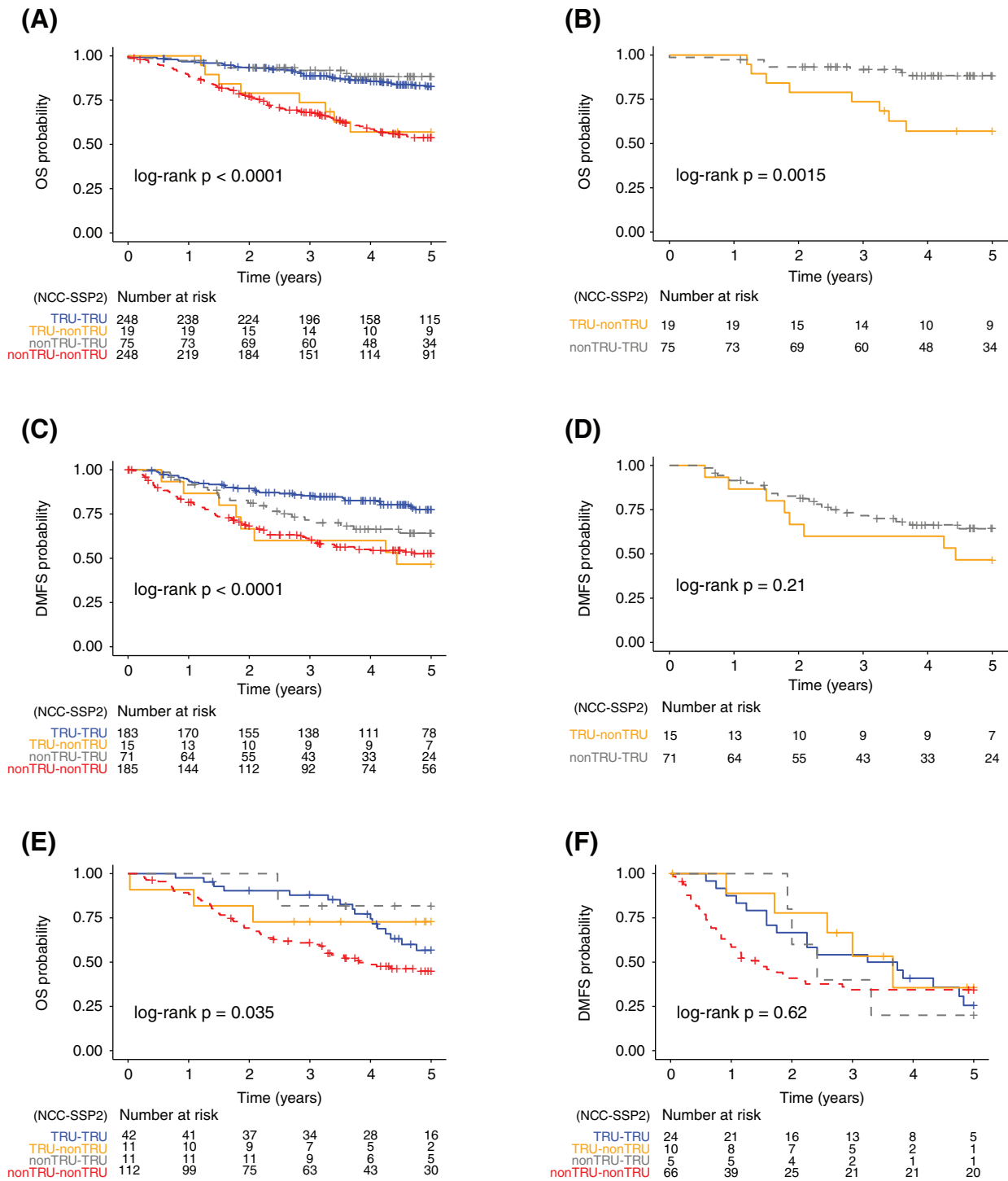


FIGURE 3 Comparison of classification methods and implication on survival outcome in lung adenocarcinoma. For details about the groups used in the Kaplan-Meier plots, see the Results section. A, Kaplan-Meier plot of OS for 590 surgically treated lung adenocarcinoma patients combined from the five validation datasets stratified by concordant or discordant NCC and SSP2 classifications. B, OS for 94 of 104 patients with discrepant SSP2/NCC classification from (A). C, DMFS for 454 surgically treated lung adenocarcinoma patients combined from the five validation datasets stratified by concordant or discordant NCC and SSP2 classifications. D, Kaplan-Meier plot of DMFS for 86 patients with discrepant SSP2/NCC classification from (C). E, Kaplan-Meier plot of OS for 176 lung adenocarcinoma patients treated with adjuvant chemotherapy combined from the five validation datasets. F, DMFS for 105 adjuvant treated lung adenocarcinoma patients combined from the five validation datasets. In all plots, P -values were calculated using the log-rank test

TABLE 2 Cox regression analysis of transcriptional subtypes in lung adenocarcinoma (surgically treated patients)

	Events (N)	Univariable analysis			Events (N)	Multivariable analysis ^a			Included confounders ^a
		HR	95% CI	P*		HR	95% CI	P*	
Overall survival ^b									
Subtypes ^c	159/590				157/586				
TRU-TRU		1.00	Ref	(<.001)		1.00	Ref	(<.001)	Stage, gender, age
TRU-nonTRU		2.9	1.37-6.32	.005		3.0	1.38-6.42	.005	
nonTRU-TRU		0.69	0.32-1.48	.3		0.76	0.35-1.65	.5	
nonTRU-nonTRU		3.3	2.30-4.84	<.001		2.9	1.92-4.23	<.001	
Stage	157/586								
I		1.00	Ref	(<.001)		1.00	Ref		
II		3.1	2.17-4.41	<.001		2.5	1.73-3.61	<.001	
III		7.2	4.66-11.10	<.001		5.4	3.46-8.41	<.001	
Gender	159/590								
Female		1.00	Ref	(.3)		1.00	Ref		
Male		1.2	0.87-1.61	.3		0.96	0.70-1.33	.8	
Age (yr)	159/590								
		1.04	1.02-1.06	(<.001)		1.03	1.01-1.05	<.001	
Distant metastasis-free survival ^d									
Subtypes ^c	146/454				145/452				
TRU-TRU		1.00	Ref	(<.001)		1.00	Ref	(<.001)	Stage, gender, age
TRU-nonTRU		3.0	1.38-6.39	.005		3.0	1.39-6.52	.005	
nonTRU-TRU		1.8	1.06-3.02	.03		1.4	0.77-2.35	.2	
nonTRU-nonTRU		2.8	1.88-4.15	<.001		2.1	1.37-3.21	<.001	
Stage	145/452								
I		1.00	Ref	(<.001)		1.00	Ref		
II		3.2	2.28-4.58	<.001		2.8	1.89-4.02	<.001	
III		3.3	1.71-6.39	<.001		3.0	1.48-5.65	.001	
Gender	146/454								
Female		1.00	Ref	(.3)		1.00	Ref		
Male		1.2	0.87-1.66	.3		1.1	0.77-1.49	.7	
Age (yr)	146/454								
		1.02	1.002-1.04	(.03)		1.02	1.01-1.05	.02	

Abbreviation: CI, confidence interval.

^aThe following confounders were included in the model: Stage (not Stage IV because of too few cases), gender and age. The confounders were selected based on their significance from the univariable analysis with $P \leq .05$ (except for gender).

^bFollow-up starts after surgical resection of the tumor lesion and ends at death by any reason (=event).

^cGroups were created based on a combination of two classifiers' outcome: TRU or nonTRU. Classifier1 (=NCC) – Classifier2 (=SSP).

^dFollow-up starts after surgical resection of the tumor lesion and ends at distant metastasis occurrence (=event).

*P-value for the pairwise comparisons were calculated using the Wald test. Overall P-values (also from the Wald test) are given within the parentheses.

Survival analysis of this selective Stage I cohort showed that CLAMS stratified patients into better and worse prognosis (log-rank $P = .02$, HR = 3.2, 95% CI = 1.2-8.8), with a 4-year relapse-free rate of 82% for the TRU group, while only 41% for the nonTRU group (Figure 4F). In multivariable analysis, CLAMS remained significant when using gender, age and Stage IA/IB as covariates (HR = 3.3, 95% CI = 1.04-10.5), and borderline nonsignificant when including also smoking status (never/smoker) in the model ($P = .06$).

In a second validation, we applied CLAMS to 30 FFPE Stage-I tumors with RNA sequencing data (GSE143486). CLAMS stratified

patients into better and worse survival (Figure 4G) and remained significant in multivariable analysis (nonTRU, HR = 8.8, 95% CI = 2.2-34) using age, gender and Stage IA/IB as covariates and OS as clinical endpoint.

4 | DISCUSSION

In the current study, we present a gene expression-based classifier of lung AC molecular subtypes applicable to single samples irrespectively

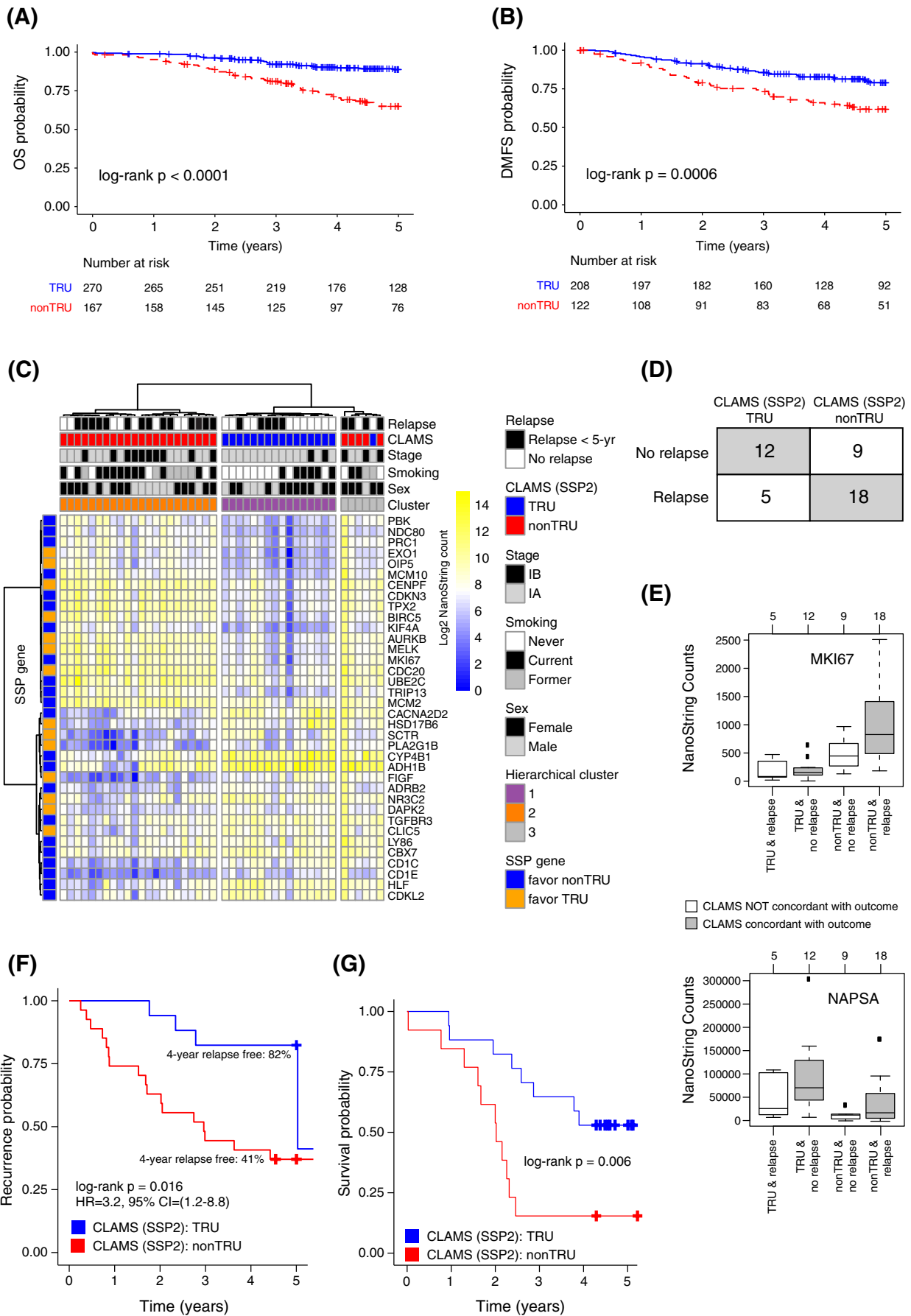


FIGURE 4 Legend on next page.

of technical platform or cohort composition. Moreover, the classifier represents an independent prognostic assessment tool for surgically treated tumors in this disease and can be translated into an assay applicable for routine clinical tissue.

The SSP2 TRU/nonTRU and the SSP3 TRU/PI/PP models both included cell proliferation as a significant biological process, in line with our previous findings of expression of proliferation-related genes representing the main prognostic component in the NCC model.¹³ Moreover, the functional analyses of SSP models demonstrate that SSP machine-learning using thousands of genes can identify biologically relevant features that can be grouped into interpretable biological processes.

Overall, we observed an accuracy of 0.85 for the SSP2 and 0.81 for the SSP3 models based on all validation samples (irrespective of disease stage). However, one has to bear in mind that the SSP models were trained on molecular assignments obtained by the NCC method, which itself comprises inherited robustness problems.¹³ Specifically, the NCC subtype training labels are not optimal for certain cohorts, due to a cohort composition different from that in which the original NCC centroids were derived. Illustrating this, notable differences in the proportion of TRU-classified samples were observed between the NCC and SSP classifiers for specific validation datasets (Figure 2D), most prominently in the Okayama et al dataset.⁴⁸ This dataset consists of 74% Stage-I tumors, and patients in this dataset have been shown to have a generally very good prognosis.¹⁵ Thus, in this context, NCC classification infers a predicted poor outcome class (ie, non-TRU) to patients with an intrinsically good prognosis due to the prerequisite of gene-centering in the NCC model. In contrast, SSP classifiers appear able to handle cohort composition bias due to their ability to classify samples truly independently. These findings illustrate the benefits of gene expression-based SSPs in a possible clinical context if there are no standardized relevant reference datasets for gene-centering available.

The proposed TRU, PI and PP subtypes have been associated with different molecular and clinicopathological characteristics,^{4,9,13} and we also demonstrate an association with histological growth patterns of lung AC. The perhaps most clinically useful feature of the expression subtypes is their association with patient outcome in early stage (operable) disease.¹³ While we find that both TRU/nonTRU NCC and SSP2 classifications are associated with patient outcome in early stage patients, actual patient outcome favors the SSP2 classification for discordantly classified cases. This observation is crucial in the context of SSP applicability, allowing a platform agnostic gene

signature to be applied to individual patients without any data preprocessing or reference cohorts. Functional analyses of SSP2 and SSP3 suggest that, in a prognostic context, these predictors provide a relative division between low- and (more) high-proliferative tumors. Application of the SSP2 model to 504 squamous cell lung carcinomas assembled from public datasets²² classified 96% of tumors as non-TRU, without any prognostic association (exploratory analysis, Supplementary Figure S4). This opposite result compared to AC is likely due to an intrinsically higher proliferation rate in squamous tumors compared to AC tumors (shown by, eg, Reference 49). This further illustrates the differences in underlying prognostic gene expression components in the histological subtypes of lung cancer, which could be one of the reasons underlying the difficulty in validating gene signatures from NSCLC studies of mixed histologies in histology specific cohorts.¹⁵

While a gene signature predictive of response to chemotherapy is highly desirable in resected lung AC, our results do not support that the proposed molecular subtypes currently seem to match that need. Instead, the current potential clinical value of our derived predictors lies in improved risk stratification of surgically treated Stage-I patients. This risk stratification could aid in identifying patient subsets for which, on a group level, additional adjuvant treatment appears less motivated (TRU-cases). For remaining patients (nonTRU), it may be argued that additional adjuvant treatment could be considered. Ideally, such claims need to be supported by randomized trial data that investigates the benefit of additional adjuvant treatment to otherwise untreated patients stratified by the molecular assay. Such studies have, to date, not been reported in lung cancer, in contrast to, for example, breast cancer.⁵⁰ To allow for the latter, robust assays applicable to degraded RNA from fixated tissue are needed. This has represented a challenge for introducing gene expression-based signatures into the clinic. As a countermeasure, focused gene expression methods, such as the NanoString nCounter technique, have been used in breast cancer (the ProSigna test). To address the requirement of analyzing degraded RNA, we paired our SSP2 model (CLAMS) with the NanoString nCounter XT technology forming a “complete” assay. In FFPE RNA from 44 Stage-I patients, we could demonstrate that the assay could recapitulate SSP2 gene rules, and that NanoString gene expression was representative of the expected subgroups. These results were supported also by application of the SSP2 model to FFPE RNA sequencing data. Together, these findings show, to the best of our knowledge for the first time in lung cancer, that *in silico* derived SSP rules can be transferred to an FFPE applicable assay. In the

FIGURE 4 SSP2 performance on surgically treated Stage-I lung adenocarcinomas. A, Kaplan-Meier plot of OS for surgically treated Stage-I patients in the validation datasets (only patients with outcome data), stratified by SSP2 classification. B, DMFS for surgically treated Stage-I patients in the validation datasets, stratified by SSP2 classification. C, Hierarchical clustering (Pearson correlation and ward.D linkage) of log2 count NanoString data for 44 FFPE Stage-I tumors using the 36 genes present in the SSP2 model through the CLAMS package. D, Confusion matrix of CLAMS prediction vs clinical status of relapse (loc-regional/distant) yes/no. E, Gene expression of MKI67 (Ki67) and NAPSA (Napsin A) across the 44 NanoString cases stratified by CLAMS prediction and clinical relapse status. Groups in gray represents agreement between TRU/non relapse and nonTRU/relapse. F, Kaplan-Meier plot of recurrence-free (loco-regional/distant) interval for the 44 NanoString cases stratified by CLAMS prediction. G, Kaplan-Meier plot of OS for 30 Stage-I tumors from GSE143486 stratified by CLAMS prediction. FFPE RNA for these samples were analyzed by RNA sequencing. In all Kaplan-Meier plots, *P*-values were calculated using the log-rank test

selected set of 44 analyzed patients (ie, not population-representative), the assay translated into a significant difference in recurrence-free interval, with a similar 4-year recurrence-free outcome as in the *in silico* validation cohort (Figure 4). While SSP performance in the FFPE cohort could desirably have been higher, one should bear in mind that, at present, there is a shortage of clinical risk stratification tools for patients with lowest disease stage subjected to curative surgery, and for whom current guidelines do not recommend adjuvant therapy. Moreover, surgical treatment may of course cure patients with high-proliferative tumors (nonTRU) that have not yet metastasized, whereas low-proliferative tumors (TRU) may have acquired metastatic potential early in their development. Such instances represent limitations for prognostic gene expression models based on surgical specimens.

In summary, we have derived platform independent single sample gene expression classifiers of proposed transcriptional subtypes in lung AC that also provides risk stratification in surgically treated Stage-I patients. Our SSP2 and SSP3 models now allow unrestricted usage of the TCGA subtypes even in highly selected lung AC cohorts, including advanced stage tumors. In malignancies such as breast cancer, gene expression signatures have now made their way into clinical practice to support clinical decision-making about adjuvant therapy. Whether a similar development will occur in early stage lung AC remains to be seen. However, robust gene expression predictors that have been translated into actual assays is an important first step in demonstrating that a similar development may be worthwhile exploring.

ACKNOWLEDGMENTS

The authors would like to acknowledge Martin Lauss at the Division of Oncology and Pathology, Lund University, Sweden, for initial support with AIMS. Financial support for this study was provided by the Swedish Cancer Society, the Sjöberg Foundation, the Fru Berta Kamprad Foundation, BioCARE a Strategic Research Program at Lund University, Stiftelsen Jubileumsklinikens Forskningsfond mot Cancer (Gustav V:s Jubilee Foundation), and The National Health Services (Region Skåne/ALF). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

CONFLICT OF INTEREST

The authors declared no potential conflicts of interest.

ETHICS STATEMENT

The study was approved by the Regional Ethical Review Board in Lund, Sweden (registration numbers: 2004/762, 2008/702, 2014/546, 2014/748, 2015/575 and 2015/831). By decision of the Ethical Review Board, specific written informed consent from included patients in this study were not required if these were not included in the ongoing LUCAS study (The Lung Cancer Study in Southern Sweden, for which written informed consent existed), as no personal data was used for this study. In accordance with the decision of the Ethical Review Board, non-LUCAS patients were

informed about the study through local advertisement in news media in the region. All experiments were conducted in agreement with patient consent and ethical review board regulations and decision.

DATA AVAILABILITY STATEMENT

Data sources of the publicly available expression data used in this study are listed in Table 1. The NanoString expression data generated in this study is presented in Supplementary Table S3. Further details are available from the corresponding author upon request.

ORCID

Sofi Isaksson  <https://orcid.org/0000-0003-0051-8507>

Johan Staaf  <https://orcid.org/0000-0001-5254-5115>

REFERENCES

1. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol*. 2015;10:1243-1260.
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *New Engl J Med*. 2008;359:1367-1380.
3. Hammerman PS, Lawrence MS, Voet D, et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489:519-525.
4. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511:543-550.
5. Swanton C, Govindan R. Clinical implications of genomic discoveries in lung cancer. *N Engl J Med*. 2016;374:1864-1873.
6. Sobin LH, Gospodarowicz MK, Wittekind C, International Union Against Cancer (UICC). *TNM Classification of Malignant Tumours*. 7th ed. Chichester, UK: Wiley-Blackwell; 2009.
7. Crino L, Weder W, van Meerbeek J, Felip E. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21(Suppl 5):v103-v115.
8. Burdett S, Pignon JP, Tierney J, et al. Adjuvant chemotherapy for resected early-stage non-small cell lung cancer. *Cochrane Database Syst Rev*. 2015;3:CD011430.
9. Wilkerson MD, Yin X, Walter V, et al. Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS One*. 2012;7:e36530.
10. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med*. 2002;8:816-824.
11. Hayes DN, Monti S, Parmigiani G, et al. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol*. 2006;24:5079-5090.
12. 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 USA*. 2001;98:13790-13795.
13. Ringner M, Jonsson G, Staaf J. Prognostic and chemotherapy predictive value of gene expression phenotypes in primary lung adenocarcinoma. *Clin Cancer Res*. 2015;22:218-229.
14. Planck M, Isaksson S, Veerla S, Staaf J. Identification of transcriptional subgroups in EGFR-mutated and EGFR/KRAS-wild type lung adenocarcinoma reveals gene signatures associated with patient outcome. *Clin Cancer Res*. 2013;19:5116-5126.

15. Ringner M, Staaf J. Consensus of gene expression phenotypes and prognostic risk predictors in primary lung adenocarcinoma. *Oncotarget*. 2016;7:52957-52973.
16. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol*. 2009;27:1160-1167.
17. Paquet ER, Hallett MT. Absolute assignment of breast cancer intrinsic molecular subtype. *J Natl Cancer Inst*. 2015;107:357.
18. Qi L, Chen L, Li Y, et al. Critical limitations of prognostic signatures based on risk scores summarized from gene expression levels: a case study for resected stage I non-small-cell lung cancer. *Brief Bioinform*. 2015;17:233-242.
19. Patil P, Bachant-Winner PO, Haibe-Kains B, Leek JT. Test set bias affects reproducibility of gene signatures. *Bioinformatics*. 2015;31:2318-2323.
20. Tan AC, Naiman DQ, Xu L, Winslow RL, Geman D. Simple decision rules for classifying human cancers from gene expression profiles. *Bioinformatics*. 2005;21:3896-3904.
21. Geman D, d'Avignon C, Naiman DQ, Winslow RL. Classifying gene expression profiles from pairwise mRNA comparisons. *Stat Appl Genet Mol Biol*. 2004;3:Article19.
22. Cirenajwis H, Lauss M, Planck M, Vallon-Christersson J, Staaf J. Performance of gene expression-based single sample predictors for assessment of clinicopathological subgroups and molecular subtypes in cancers: a case comparison study in non-small cell lung cancer. *Brief Bioinform*. 2019;21:729-740.
23. Chitale D, Gong Y, Taylor BS, et al. An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. *Oncogene*. 2009;28:2773-2783.
24. Clinical Lung Cancer Genome Project (CLCGP); Network Genomic Medicine (NGM). A genomics-based classification of human lung tumors. *Sci Transl Med*. 2013;5:209ra153.
25. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature*. 2006;439:353-357.
26. Lee ES, Son DS, Kim SH, et al. Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin Cancer Res*. 2008;14:7397-7404.
27. Tomida S, Takeuchi T, Shimada Y, et al. Relapse-related molecular signature in lung adenocarcinomas identifies patients with dismal prognosis. *J Clin Oncol*. 2009;27:2793-2799.
28. Hou J, Aerts J, den Hamer B, et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS One*. 2010;5:e10312.
29. Lu TP, Tsai MH, Lee JM, et al. Identification of a novel biomarker, SEMA5A, for non-small cell lung carcinoma in nonsmoking women. *Cancer Epidemiol Biomarkers Prev*. 2010;19:2590-2597.
30. Rousseaux S, Debernardi A, Jacquiau B, et al. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med*. 2013;5:186ra66.
31. Botling J, Edlund K, Lohr M, et al. Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis and tissue microarray validation. *Clin Cancer Res*. 2012;19:194-204.
32. Seo JS, Ju YS, Lee WC, et al. The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res*. 2012;22:2109-2119.
33. Tarca AL, Lauria M, Unger M, et al. Strengths and limitations of microarray-based phenotype prediction: lessons learned from the IMPROVER diagnostic signature challenge. *Bioinformatics*. 2013;29:2892-2899.
34. Chen KY, Hsiao CF, Chang GC, et al. Estrogen receptor gene polymorphisms and lung adenocarcinoma risk in never-smoking women. *J Thorac Oncol*. 2015;10:1413-1420.
35. Der SD, Sykes J, Pintilie M, et al. Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. *J Thorac Oncol*. 2014;9:59-64.
36. Karlsson A, Jonsson M, Lauss M, et al. Genome-wide DNA methylation analysis of lung carcinoma reveals one neuroendocrine and four adenocarcinoma epitypes associated with patient outcome. *Clin Cancer Res*. 2014;20:6127-6140.
37. Djureinovic D, Hallstrom BM, Horie M, et al. Profiling cancer testis antigens in non-small-cell lung cancer. *JCI Insight*. 2016;1:e86837.
38. Shedden K, Taylor JM, Enkemann SA, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14:822-827.
39. Okayama H, Schetter AJ, Ishigame T, et al. The expression of four genes as a prognostic classifier for stage I lung adenocarcinoma in 12 independent cohorts. *Cancer Epidemiol Biomarkers Prev*. 2014;23:2884-2894.
40. Fouret R, Laffaire J, Hofman P, et al. A comparative and integrative approach identifies ATPase family, AAA domain containing 2 as a likely driver of cell proliferation in lung adenocarcinoma. *Clin Cancer Res*. 2012;18:5606-5616.
41. Zhu CQ, Ding K, Strumpf D, et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol*. 2010;28:4417-4424.
42. Tang H, Xiao G, Behrens C, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small-cell lung cancer patients. *Clin Cancer Res*. 2013;19:1577-1586.
43. Salomonsson A, Micke P, Mattsson JSM, et al. Comprehensive analysis of RNA binding motif protein 3 (RBM3) in non-small cell lung cancer. *Cancer Med*. 2020;15:5609-5619.
44. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. 2011;6:244-285.
45. Lindquist KE, Karlsson A, Leveen P, et al. Clinical framework for next generation sequencing based analysis of treatment predictive mutations and multiplexed gene fusion detection in non-small cell lung cancer. *Oncotarget*. 2017;8:34796-34810.
46. Lira ME, Choi YL, Lim SM, et al. A single-tube multiplexed assay for detecting ALK, ROS1, and RET fusions in lung cancer. *J Mol Diagn*. 2014;16:229-243.
47. Foroutan M, Bhuva DD, Lyu R, Horan K, Cursons J, Davis MJ. Single sample scoring of molecular phenotypes. *BMC Bioinformatics*. 2018;19:404.
48. Okayama H, Kohno T, Ishii Y, et al. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res*. 2012;72:100-111.
49. Karlsson A, Brunnstrom H, Micke P, et al. Gene expression profiling of large cell lung cancer links transcriptional phenotypes to the new histological WHO 2015 classification. *J Thorac Oncol*. 2017;12:1257-1267.
50. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016;375:717-729.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Liljedahl H, Karlsson A, Oskarsdottir GN, et al. A gene expression-based single sample predictor of lung adenocarcinoma molecular subtype and prognosis. *Int. J. Cancer*. 2021;148:238-251. <https://doi.org/10.1002/ijc.33242>