Supplementary Online Content

Jones GD, Brandt WS, Shen R, et al. A genomic-pathologic annotated risk model to predict recurrence in early-stage lung adenocarcinoma. *JAMA Surg*. Published online December 23, 2020. doi:10.1001/jamasurg.2020.5601

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Sequencing and Analysis

MSK-IMPACT Sequencing

Patient tumors and matched normal blood specimens were analyzed in a CLIA-certified laboratory using the MSK-IMPACT assay, a capture-based next-generation sequencing (NGS) platform that detects mutations, copy-number alterations, and select rearrangements in up to 468 cancer-associated genes, as previously described.[1,2] The minimum and mean depth of coverage per sample in our study were 164x and 784x, respectively.

Gene Grouping for Biological Relevance

Grouping of genes was based on related biological function.[3] *CDK4*, *CDKN2A*, and *RB1* were grouped, as they are members of the cell cycle pathway. *PIK3CA*, *PIK3R1*, *PTEN*, and *TSC* are members of the PI3K pathway. *BRAF*, *RAF1*, *RIT1*, and *MAP2K1* are members of the RTK/RAS pathway, while *NF1* and *RASA1* are known negative regulators of this pathway. *ERBB2* and *MET* are receptor tyrosine kinases implicated in lung adenocarcinoma. *RET*, *ROS1*, and *ALK* are the most well-known and actionable fusions in lung adenocarcinoma.[3,4]

Mutational Signatures Associated with Smoking

In order to investigate mutational signatures associated with tobacco smoking, we used version 3 from the catalogue provided by COSMIC (https://cancer.sanger.ac.uk/cosmic/signatures/). Signatures were computed as described in doi: https://doi.org/10.1101/322859, and we focused on signature SBS4, which is known to have been observed in experimental systems exposed to tobacco carcinogens (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/SBS4.tt).

Statistical Analysis

The primary outcome of interest was relapse-free survival (RFS), defined as the duration between surgery and recurrence or death without recurrence. Otherwise, patients were censored at the time of last follow-up. RFS was estimated using the Kaplan-Meier approach and compared between clinicopathologic factors and genetic alterations of interest using the log-rank test, stratified by pathologic stage where appropriate. Analyses considering only genomic alterations were adjusted for multiple testing using the false discovery rate (FDR) approach. Selection as candidates for the multivariable models were conducted using the FDR-adjusted p values, where applicable. All analyses were 2-sided, and $p<0.05$ was considered statistically significant. The proportional hazards assumption was assessed using statistical tests and graphical diagnostics based on the scaled Schoenfeld residuals. There was no indication of violation of the proportional hazards assumption: the tests for each of the covariates were not significant, the global test was not statistically significant, and the graphical diagnostics did not indicate pattern with time for each covariate. Statistical analyses were conducted using Stata 13.1 (Stata Corp, College Station, Texas) and R 3.5.1 (R Development Core Team, Austria, Vienna), including the "survival," "rms," "riskRegression," and "mice" packages, downloaded in November 2018. Median follow-up duration was generated using the reverse Kaplan-Meier method.[5]

Development of the Association Models to Identify Factors Associated with RFS

In the first stage of the analysis, our goal was to identify factors associated with the primary outcome of RFS. We focused on developing a clinical-only model including relevant clinicopathologic factors. Associations between clinical factors and RFS were assessed using univariable Cox proportional hazards models, stratified by pathologic stage (except in the case of stage factor itself and pathologic tumor size) (clinicopathologic [CP]–association model). Multicollinearity was examined between lymphovascular invasion (LVI) and visceral pleural invasion (VPI). Predominant histologic subtypes were classified into one of five groups (lepidic [n=71], acinar [n=220], papillary [n=29], micropapillary [n=23], or solid [n=52]), or "other [n=26]" (which was comprised of adenocarcinoma in-situ [AIS] and minimally-invasive adenocarcinoma [MIA]). A multivariable model was constructed starting with all factors with $p<0.1$ from the univariable analyses. The relationship between each genomic factor or gene and the primary outcome was assessed using univariable Cox models, with adjustment for normalized tumor mutation burden (TMB). Analyses considering only genomic alterations were adjusted for multiple testing using the FDR approach. TMB was used as a continuous variable for all analyses. Nonlinearity of TMB was assessed with restricted cubic splines with 3 knots. Genomic factors significant on univariable analysis were then assessed in the presence of the significant factors from the CP-association model, in order to generate an adjusted hazard ratio for each factor (combined association model).

Imputation of Missing Data in the MSK Cohort

To handle missing data, multiple imputation by chained equation was conducted using the *mice* package in R 3.5.1 (R Core Team, Vienna, Austria) to derive 10 imputed data sets with a maximum of 50 iterations each.[6] This method is based on fully conditional specification, in which each incomplete variable is imputed by a separate model. All

variables considered in subsequent analyses were included in the imputation models, including survival outcomes.[7] The imputation method for primary tumor maximum standardized uptake value (14% missing) and fraction of genome altered (2% missing) were based on predictive mean modeling to ensure that imputed values were plausible. The method was (Bayesian) logistic regression for LVI (1% missing) and spread through air spaces (1% missing). The analysis model was fit on each imputed data set for the outcome of interest, and estimates from each of the 10 imputed data sets were pooled into a single set of estimates and standard errors using Rubin's rules.[8]

Integrated Risk Stratification for RFS Using PRecur

PRecur uses a statistical machine-learning approach previously described by our group to integrate genomic and clinicopathologic variables by building a collection of gradient boosting survival tree models for RFS.[9] A total of 13 clinicopathologic (see eTable 2) and 50 genomic variables (see eTable 3) were included in training the PRecur model using a boosted survival random forest method. Gradient boosting (GBM) is a machine-learning technique for regression that iteratively generates prediction models (i.e., decision trees) using a set of tuned optimal parameters. An ensemble of 1,000 survival regression trees was generated in each 5-fold cross-validation for predicting the risk of RFS using consensus prediction. The average tree-depth—or number of variables in a tree—was 2 (range 1-4), with each tree including a varying combination of clinicopathologic and genomic variables. A clinicopathologic or genomic variable's relative importance represents the contribution of that variable to the PRecur predictive model. Differences in pathologic stage between patients were accounted for by including the pathologic stage variable in the development of the model. A predicted risk score was assigned to each patient in the test set at each iteration. The predicted risk score of each patient was then averaged and rescaled between 0 and 10 (with increasing values indicating higher risk of recurrence). An optimization procedure for defining low-, intermediate-, and high-risk groups was then conducted using the maximally selected rank statistic approach to search for cutoffs that maximize the survival differences across the risk groups.[10] The *p* value for the resulting Kaplan-Meier curves was adjusted for inflation due to the cutoff optimization. Finally, at each iteration, the predictive performance of the model was recorded by assessing its concordance probability estimate (CPE) on the testing set. Of the total 426 patients, all had complete genomic data and 335 had all clinicopathologic variables required for generation of the prediction model.

Performance Measures of the Predictive Model

To assess the strength of a predictive model or risk-stratification system, a measure of discrimination must be used.[11] When survival is the outcome of interest, the most commonly applied global measure of discrimination is the concordance probability.[12] The concordance probability represents the pairwise probability of higher predicted risk equating to shorter observed survival. The concordance index (c-index)[13] and the concordance probability estimate (CPE)[12,14,15] are commonly used to estimate the concordance probability when predicted risk scores are continuous. Values closer to 1 represent better discriminatory ability. The c-index relies on the imputation of failure times, which leads to estimation bias in heavily censored cohorts such as ours. In contrast, CPE relies only on the regression parameters, and not on the observed failure and censoring times, avoiding estimation bias. As such, the CPE was chosen to assess the discrimination performance of our PRecur predictive model. For internal validation of the predictive model, CPE was derived on the basis of 200 cross-validation resamples.

External Validation of Risk Prediction Model

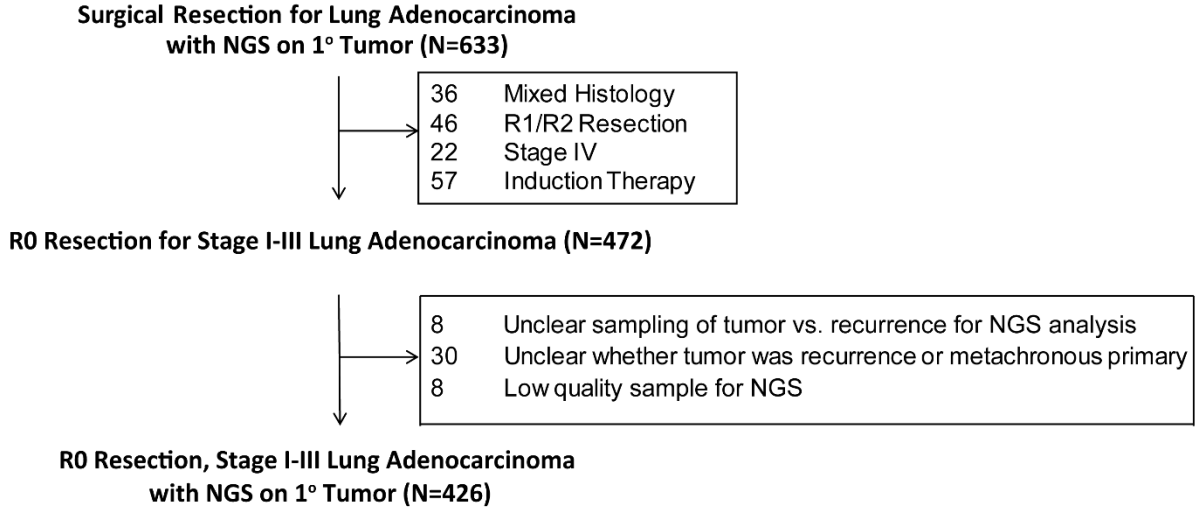
For external validation of our predictive model, we initially examined surgically resected lung adenocarcinoma data sets from The Cancer Genome Atlas (TCGA),[16] the Broad Institute[17] and TRACERx.[18] Unfortunately, in the Broad and TRACERx data sets the small numbers of patients with all data elements (n<140 in both sets) and recurrence events (n<12 in both sets), as well as differences in the distribution of tumor staging, lack of follow-up and detailed clinicopathologic variables precluded external validation studies. We therefore chose the TCGA data set (N=360), adjusted to 8th edition staging classification, for external validation. We generated an external validation model for our MSK-IMPACT early-stage cohort data set by including only the overlapping clinical covariates between ours and the TCGA data sets (age, sex, smoking status, and pathologic stage), all genes with frequency of alteration >1% (N=47, eTable 2), as well as TMB, fraction of genome altered, and whole-genome doubling. PRecur analysis was then performed identical to the complete model discussed above, with 200 cross-validations. An average predicted risk score was then assigned to each patient in the TCGA data set (N=360). Risk scores were further stratified into three groups based on the bounds of the tertiles in our ensemble training model (cuts made at tertiles correspond to risk scores of 2.5 and 5.1). Kaplan-Meier analysis was then performed to determine the difference in RFS between the 3 risk groups and compared via the log-rank test.

An interactive web application of our PRecur prediction model is available at the following link, so that users may be able to input clinicopathologic and genomic variables from other data sets:

https://axelitomartin.shinyapps.io/OncoCast-NSCLC/. The application will return a risk score for each patient calculated from the boosted survival random forest model that was trained on our dataset. All factors listed in the table below must have a designation in order for the app to run correctly. As this is a first-generation app, subsequent versions/revisions will be available at the same URL. For the convenience of the reader, for data sets that do not contain all of these factors, the model may be retrained with a unique list of factors in this external data set using the algorithm at the following link: https://github.com/AxelitoMartin/OncoCast.

Genomic Factors	Genes	Clinical Factors	Pathologic Factors
TMB	ALK	Age	Tumor_SUVmax
WGD	APC	Sex	Procedure
FGA	ARID1A	Smoker	Stage
	ARID2		Adjuvant
	ATM		LVI
	B2M		VPI
	BAP1		Lepidic
	BCOR		Acinar
	BRAF		Papillary
	BRCA2		Micropapillary
	CDK4		Solid
	CDKN2A		
	CDKN2B		
	CTNNB1		
	EGFR		
	ERBB2		
	FAT1		
	FOXA1		
	GLI1		
	GNAS		
	KEAP1		
	KIT		
	KRAS		
	MDM2		
	MED12		
	MET		
	MGA		
	MYC		
	NF1		
	NF2		
	NKX2.1		
	NTRK1		
	PIK3CA		
	PIK3R1		
	PTPRD		
	PTPRT		
	RB1		
	RBM10		
	RET		
	ROS1		
	SETD2		
	SMAD4		
	SMARCA4		
	STK11		
	TERT		
	TP53		
	U2AF1		

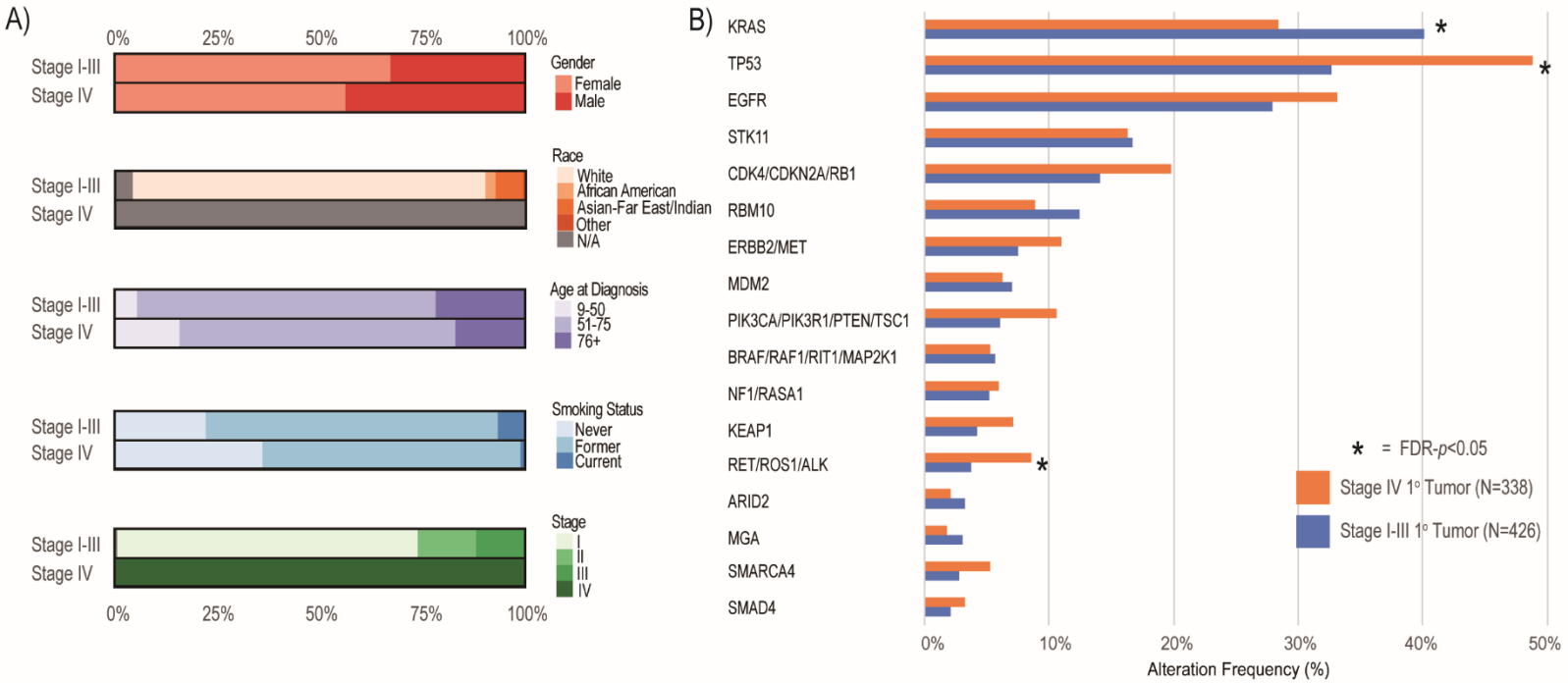
eFigure 1. CONSORT Diagram for the Stage I-III Cohort



eFigure 2. Distribution of Demographic, Clinicopathologic, and Genomic Characteristics in Early- and Late-Stage Lung Adenocarcinoma (LUAD)

(A) Demographic and clinicopathologic comparison of stage I-III (N=426) and stage IV (N=338) LUAD cohorts. (B) Percent alteration of genes present in next-generation sequencing (NGS) and grouped by biological relevance in stage I-III vs. IV LUAD.

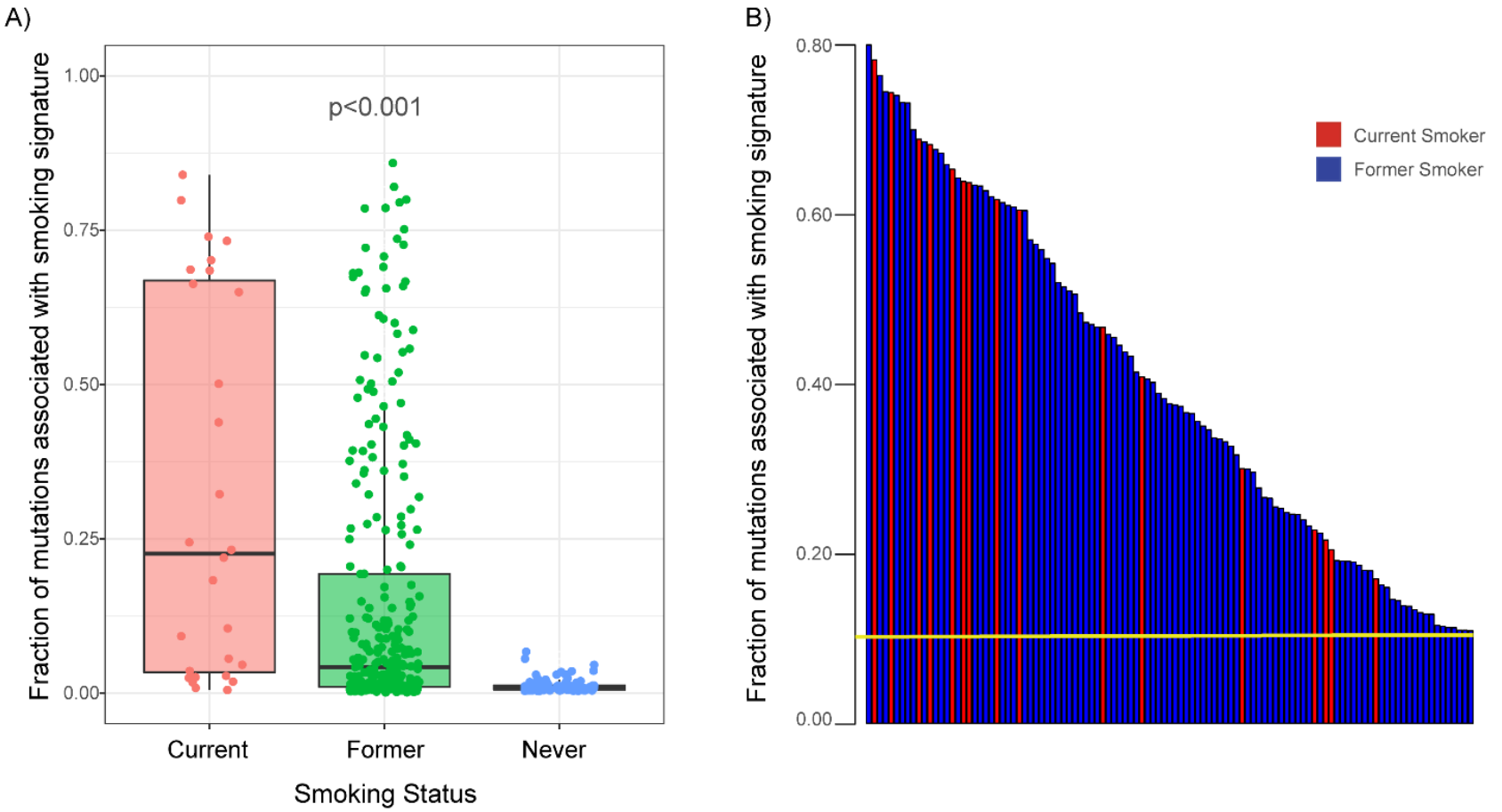
*FDR-p<0.05 for difference in alteration frequency between cohorts. Data from the stage IV cohort include NGS from primary (1º) lung tumors only—not metastatic sites.



eFigure 3. Fraction and Distribution of Mutations Associated With a Smoking Signature

A, Fraction of mutations associated with a smoking signature for current, former, and never smokers

Reported p-value based on Kruskal-Wallis test. (B) Distribution of smoking signature mutational fractions (a cutoff of 0.10 was used for fraction of mutations associated with a smoking signature).

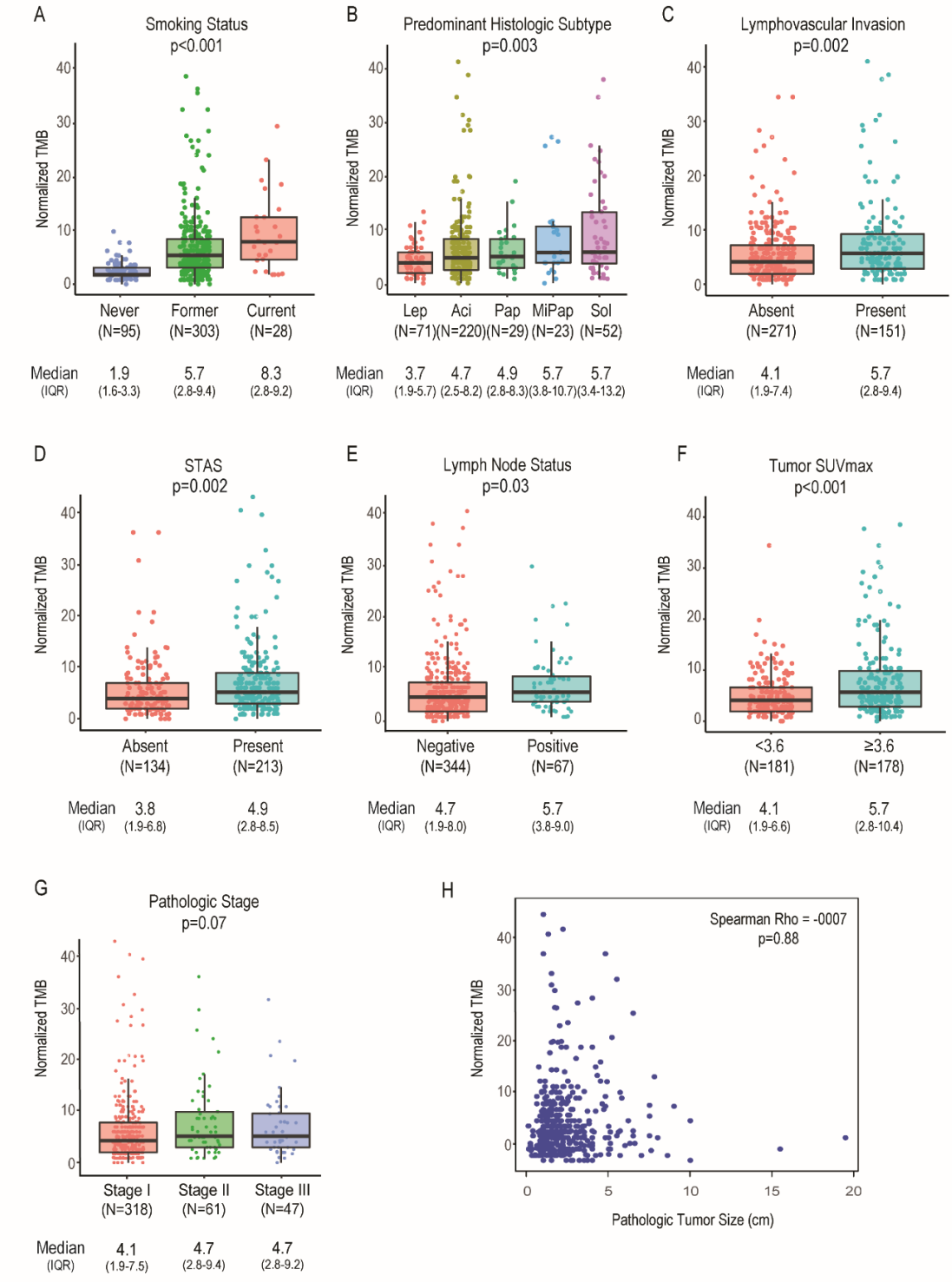


eFigure 4. Genomic Patterns of Recurrence

The relative genomic alteration profiles of primary tumors are shown in the Oncoprint for patients with (n=75) and without (n=351) recurrence, with selected clinicopathologic variables. Genes included were the level I actionable oncogenes for lung adenocarcinoma (*EGFR, BRAF, ALK*, and *ROS1*), those associated with worse relapse-free survival (*SMARCA4* and *TP53*), and *KRAS*. *FDR-p<0.05 for difference in alteration frequency between groups. FGA, fraction of genome altered; LVI, lymphovascular invasion; STAS, spread through air spaces; TMB, tumor mutation burden.

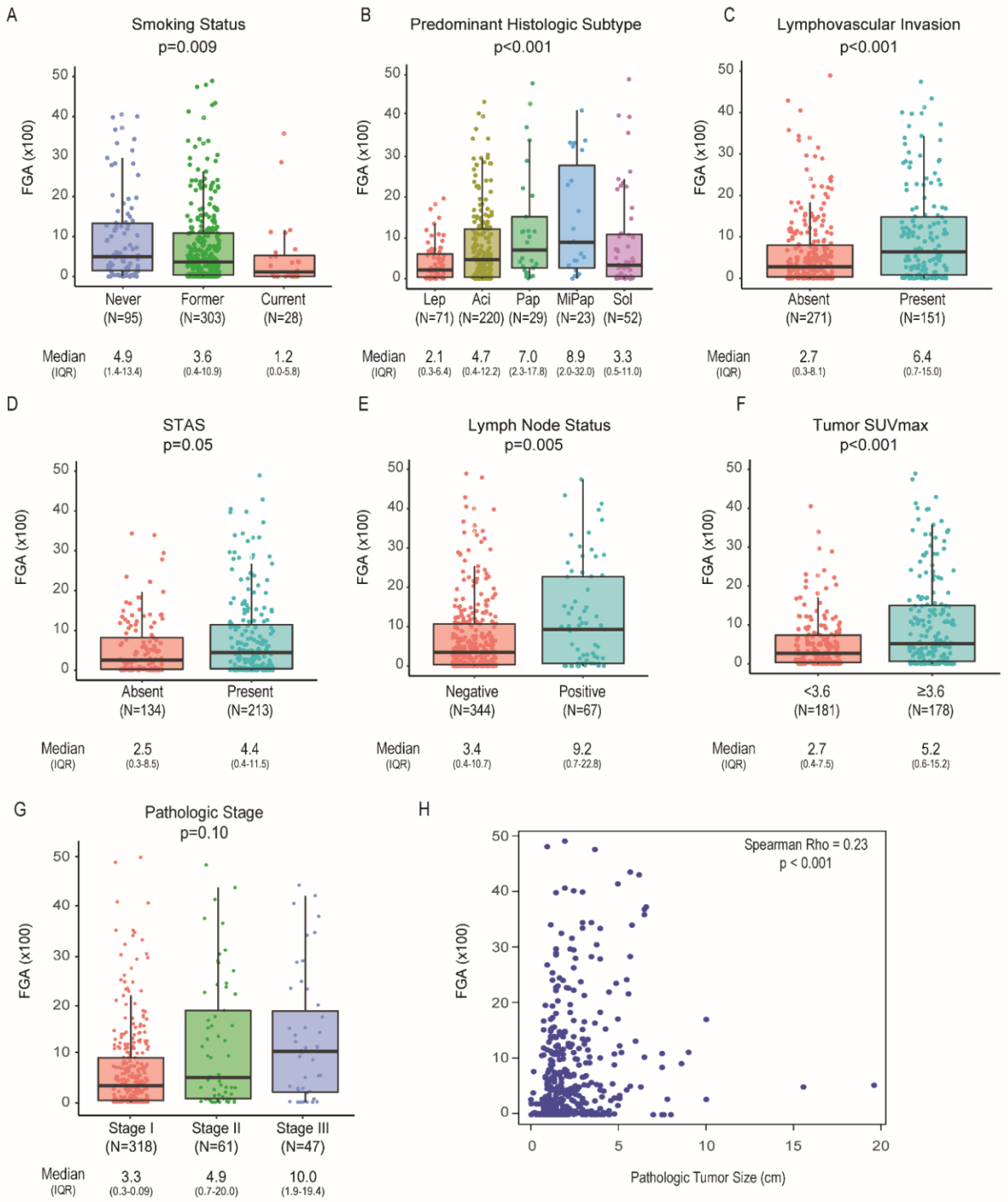
eFigure 5. Association Between Tumor Mutation Burden (TMB) and Aggressive Clinicopathologic Features

Association between tumor mutation burden (TMB) and aggressive clinicopathologic features; TMB and (A) smoking status, (B) predominant histologic subtype, (C) lymphovascular invasion, (D) spread of tumor through air spaces (STAS), (E) lymph node positivity, (F) tumor maximum standardized uptake value (SUVmax; median, 3.6), (G) pathologic stage, and (H) pathologic tumor size. The median TMB and interquartile range (IQR) are reported for each variable. Aci, acinar; lep, lepidic; mipap, micropapillary; pap, papillary; sol, solid.

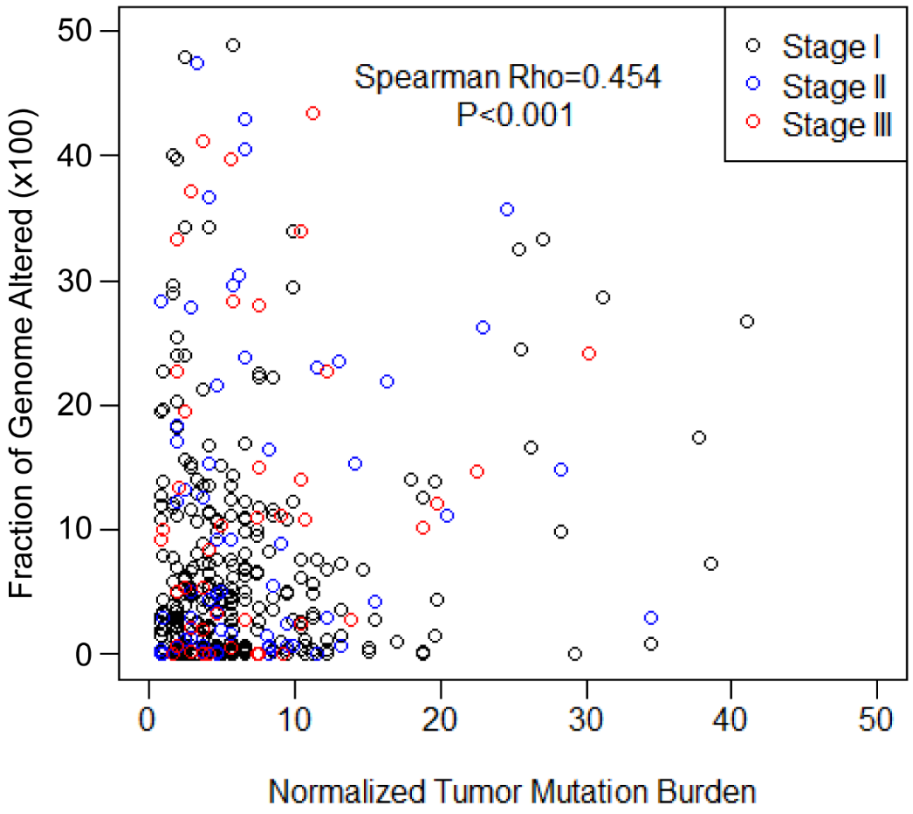


eFigure 6. Association Between Fraction of Genome Altered (FGA) and Aggressive Clinicopathologic Features

FGA and (A) smoking status, (B) predominant histologic subtype, (C) lymphovascular invasion, (D) spread of tumor through air spaces (STAS), (E) lymph node positivity, (F) tumor maximum standardized uptake value (SUVmax; median, 3.6), (G) pathologic stage, and (H) pathologic tumor size. The median FGA and interquartile range (IQR) are reported for each variable. Aci, acinar; lep, lepidic; mipap, micropapillary; pap, papillary; sol, solid.

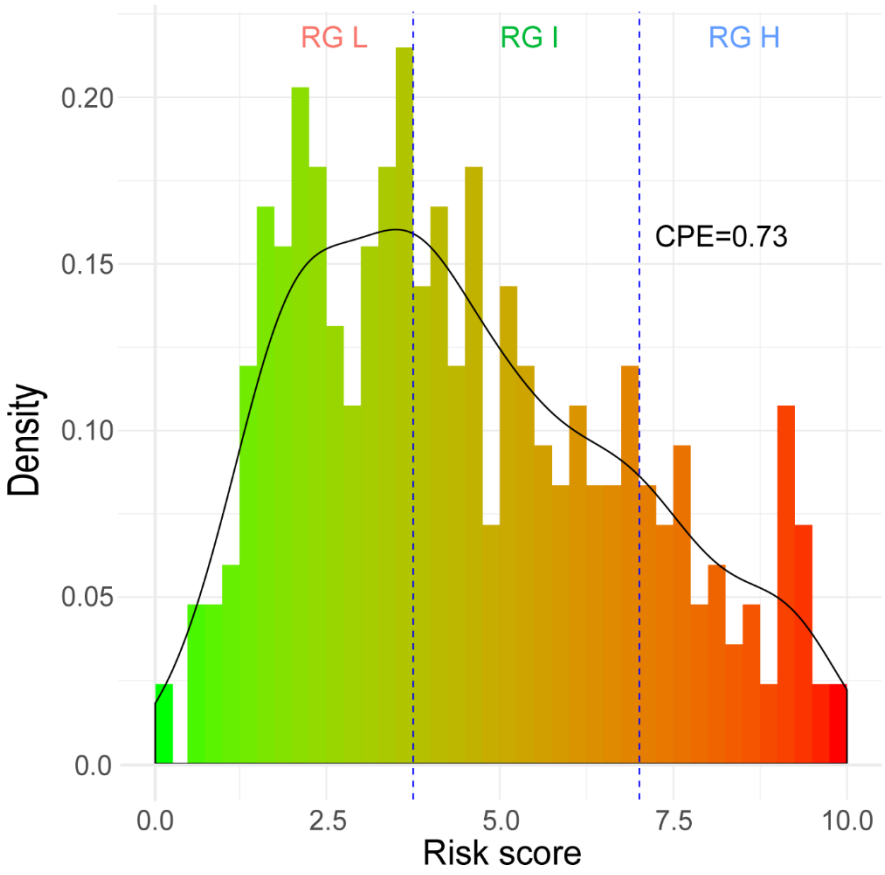


eFigure 7. Association Between Fraction of Genome Altered and Tumor Mutation Burden



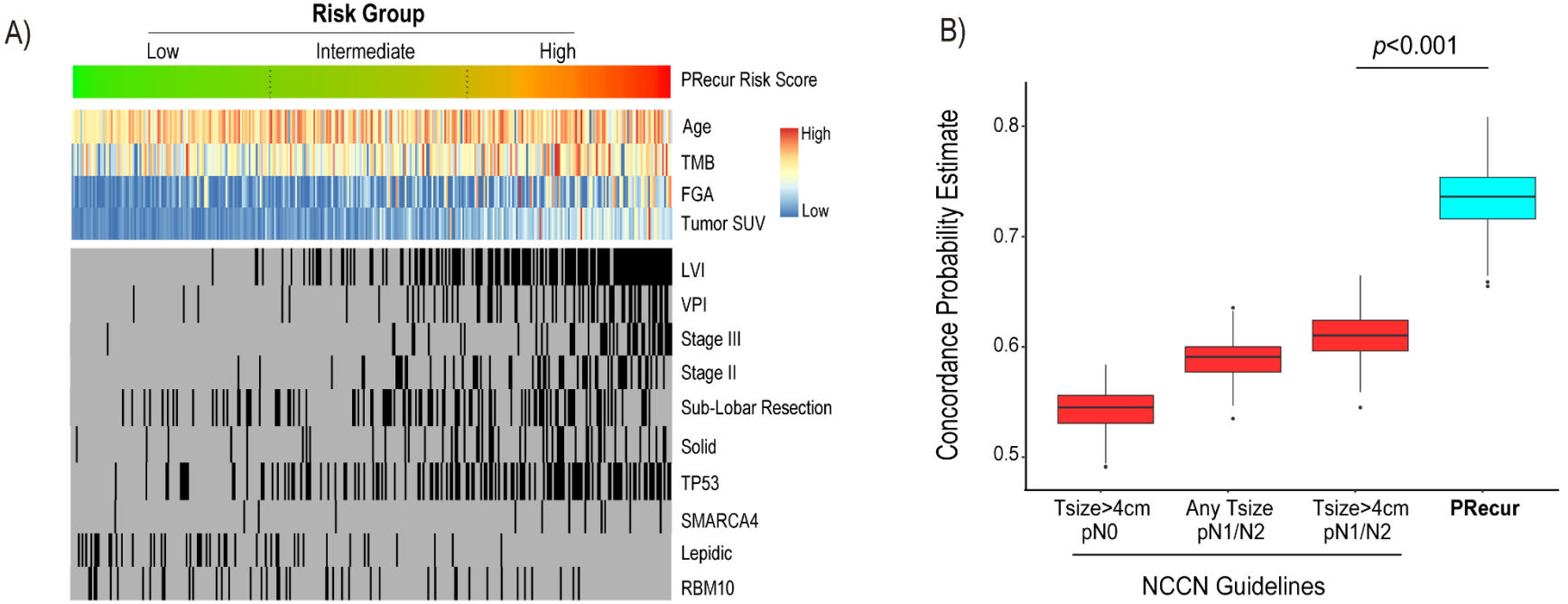
eFigure 8. Histogram of the Relapse-Free Survival (RFS) Risk Score Computed Using PRecur, Integrating Clinical and Next-Generation Sequencing Data

The risk score is scaled between 0 and 10, with a higher score indicating a higher likelihood of worse RFS. Vertical dotted lines indicate optimized risk score cutoffs using the maximally selected rank statistic approach (RGs; low [L], intermediate [I], and high [H]). Concordance probability estimate (CPE) represents the actual versus predicted time to a RFS event across the cohort (N=426).



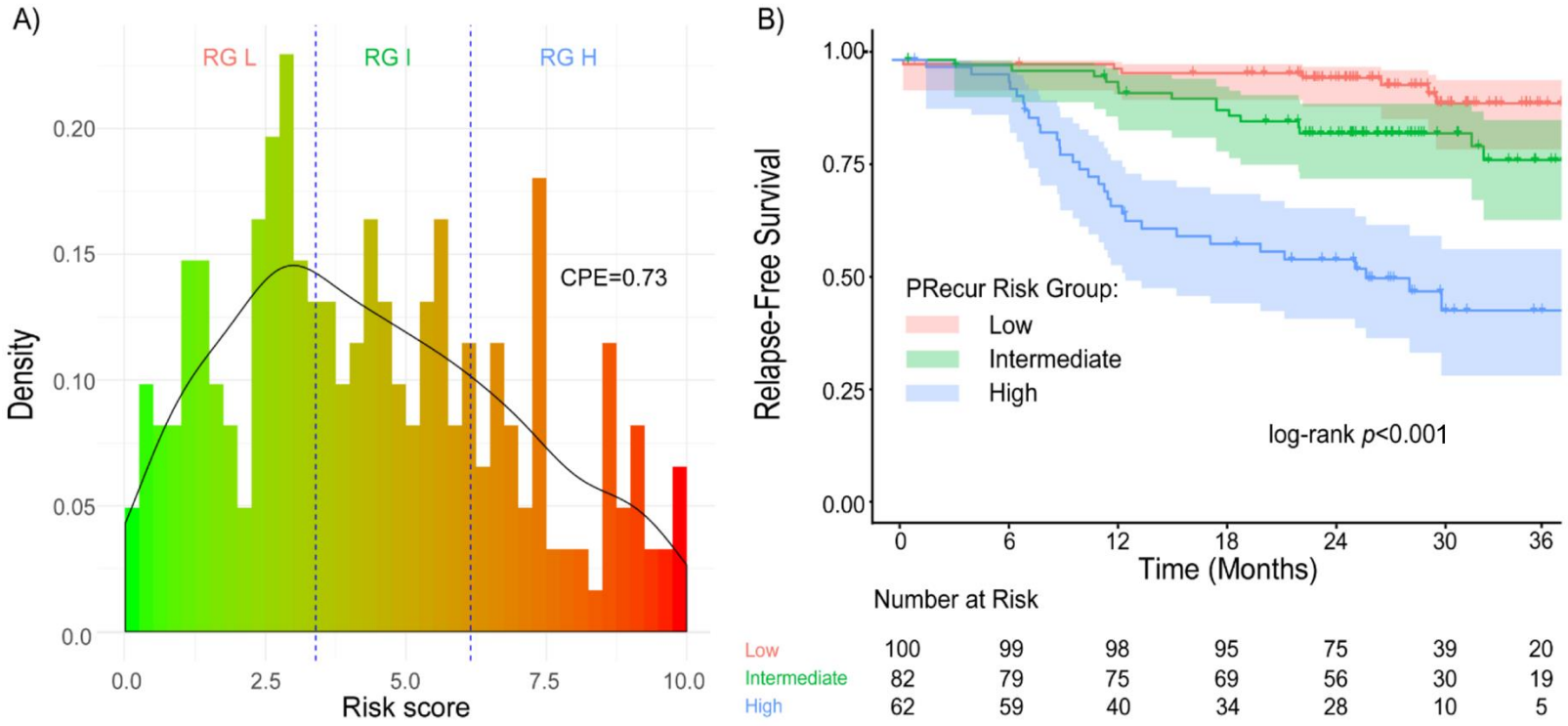
eFigure 9. PRecur Prediction Model for Relapse-Free Survival (RFS) Using Integrated Clinicopathologic and Genomic Variables for Risk Stratification

(A) Oncoprint of genomic and clinicopathologic factors sorted by risk score from low to high (left to right). (B) Boxplots showing the concordance probability estimate (CPE) for RFS prediction using TNM criteria for adjuvant therapy as recommended by the National Comprehensive Cancer Network (NCCN) (node-negative patients with tumor size >4 cm; node-positive patients with any tumor size; node-positive patients with tumor size >4 cm) versus PRecur prediction model (left to right). Reported p value based on Wilcoxon paired test. FGA, fraction of genome altered; LVI, lymphovascular invasion; SUV, standardized uptake value; TMB, tumor mutation burden; VPI, visceral pleural invasion.



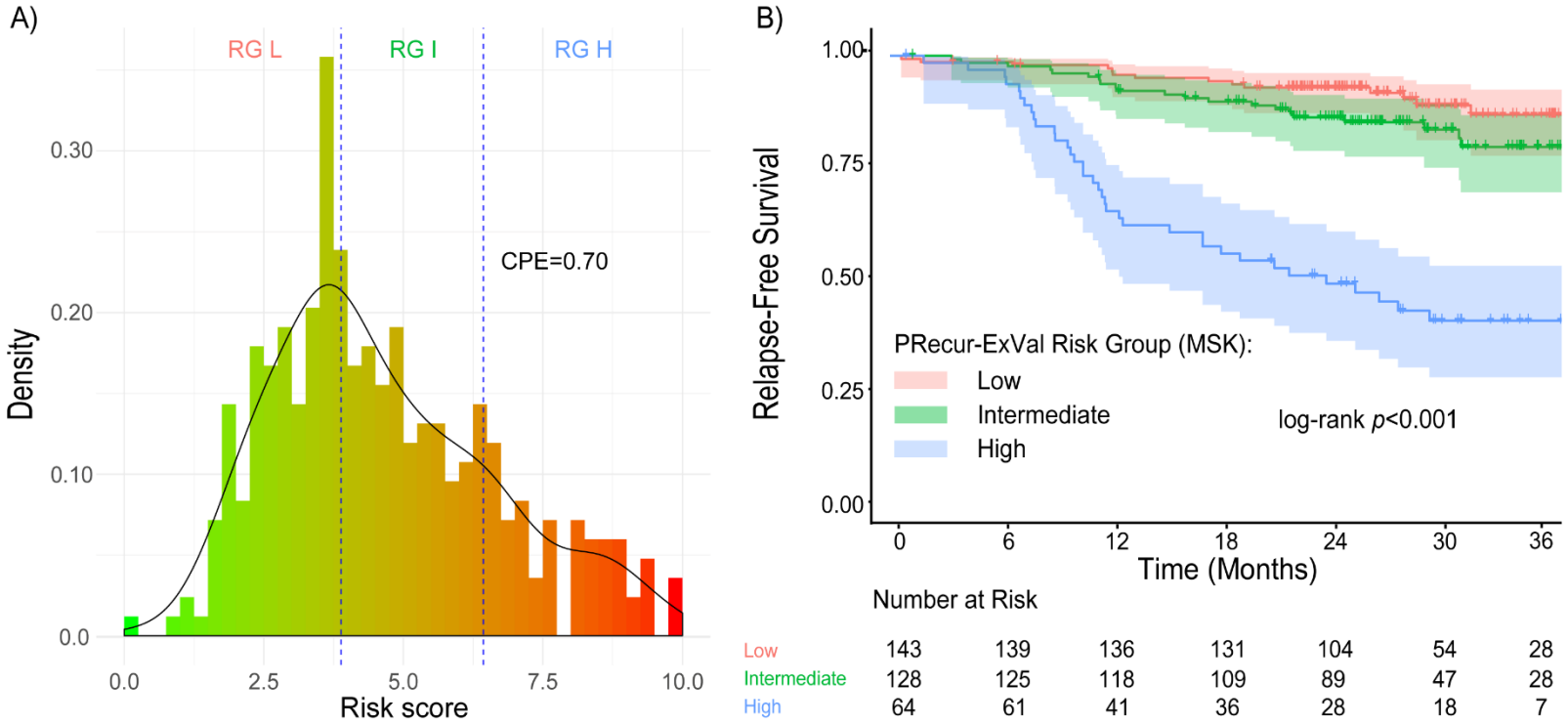
eFigure 10. PRecur Prediction Model Including Only Patients in the MSK Cohort Whose Primary Tumor Did Not Harbor a Level 1 Actionable Mutation (n = 309)

(A) Histogram of the relapse-free survival (RFS) risk score and concordance probability estimate (CPE). (B) Kaplan-Meier plot of 3-year RFS by risk group (low [L], intermediate [I], and high [H]).



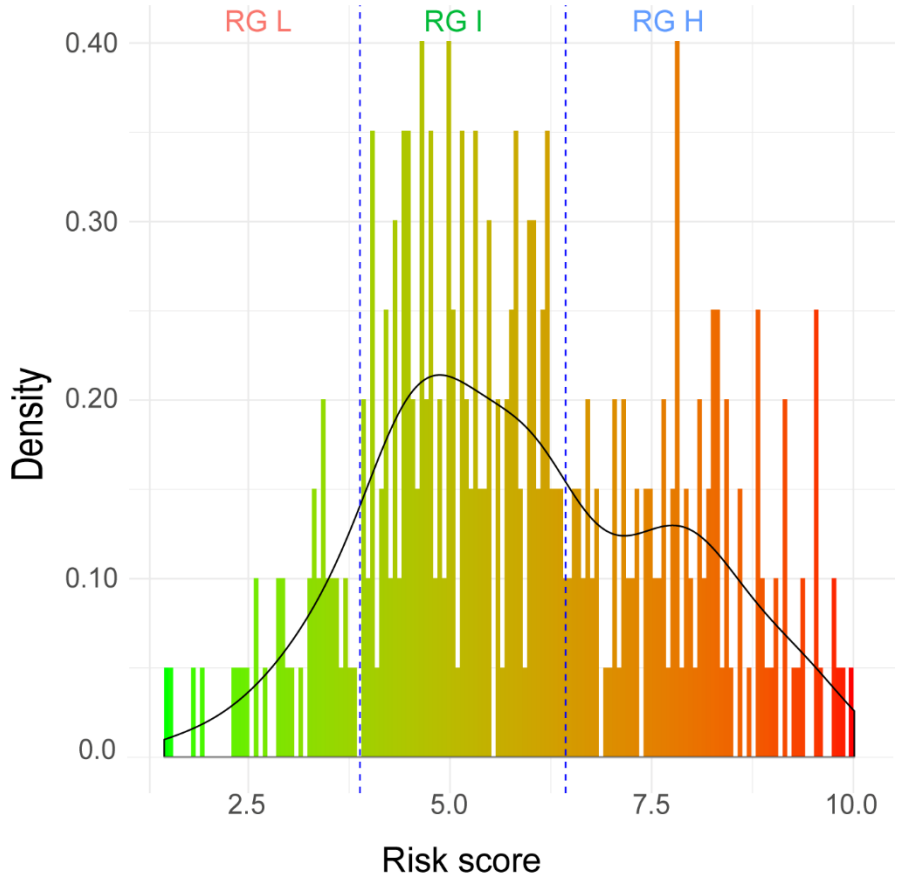
eFigure 11. PRecur-ExVal Prediction Model for the MSK Cohort (n = 426)

Including only the clinicopathologic factors available in both ours and the TCGA datasets (age, sex, smoking status, pathologic stage), in addition to all genomic variables from the full PRecur model. (A) Histogram of the relapse-free survival (RFS) risk score and concordance probability estimate (CPE). (B) Kaplan-Meier plot of 3-year RFS by risk group (low [L], intermediate [I], and high [H]).



eFigure 12. Histogram of the Relapse-Free Survival (RFS) Risk Score Computed Using PRecur-ExVal for the TCGA Data Set (n = 360)

Risk scores are scaled between 0 and 10, with a higher score indicating a higher likelihood of worse RFS. Vertical dotted lines indicate optimized risk score cutoffs using the maximally selected rank statistic approach (RGs; low [L], intermediate [I], and high [H]).



eTable 1. Summary of Patterns of Recurrence (n = 75)

Variable	Median (IQR) or No. (%)
Time to first recurrence, months	13.8 (10.2-24.5)
Recurrence confirmed via tissue biopsy	55 (73%)
General location of first recurrence, mutually exclusive[a]	
Distant	28 (37%)
Locoregional	18 (24%)
Locoregional and distant	29 (39%)
First site of distant recurrence, not mutually exclusive (n=28)	
Adrenal	4 (14%)
Bone	21 (75%)
Brain	12 (43%)
Liver	4 (14%)
Other	6 (21%)
Multiple tissue types at first recurrence	14 (19%)
Site of all distant recurrences (includes first and subsequent)	
Adrenal	6 (8%)
Bone	30 (40%)
Brain	16 (21%)
Liver	6 (8%)

[a]Distant recurrence was defined as recurrence to the opposite lung, pleura, or pleural fluid (any), any supraclavicular node, contralateral mediastinal nodes, and/or any other distant tissue or organ (American College of Chest Physicians guidelines[1]). IQR, interquartile range.

[1]Donington J, Ferguson M, Mazzone P, et al. American College of Chest Physicians and Society of Thoracic Surgeons consensus statement for evaluation and management for high-risk patients with stage I non-small cell lung cancer. Chest 2012;142:1620-35.

eTable 2. Genes in the 468-Gene MSK-IMPACT Panel

Gene Symbol	First Design	OMIM	EntrezGene	Ensembl	Approved Name
ABL1	IMPACT-341	189980	25	ENSG00000097007	c-abl oncogene 1, non-receptor tyrosine kinase
ACVR1	IMPACT-410	102576	90	ENSG00000115170	activin A receptor, type I
AGO2	IMPACT-468	606229	27161	ENSG00000123908	eukaryotic translation initiation factor 2C, 2
AKT1	IMPACT-341	164730	207	ENSG00000142208	v-akt murine thymoma viral oncogene homolog 1
AKT2	IMPACT-341	164731	208	ENSG00000105221	v-akt murine thymoma viral oncogene homolog 2
AKT3	IMPACT-341	611223	10000	ENSG00000117020	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)
ALK	IMPACT-341	105590	238	ENSG00000171094	anaplastic lymphoma receptor tyrosine kinase
ALOX12B	IMPACT-341	603741	242	ENSG00000179477	arachidonate 12-lipoxygenase, 12R type
ANKRD11	IMPACT-410	611192	29123	ENSG00000167522	ankyrin repeat domain 11
APC	IMPACT-341	611731	324	ENSG00000134982	adenomatous polyposis coli
AR	IMPACT-341	313700	367	ENSG00000169083	androgen receptor
ARAF	IMPACT-341	311010	369	ENSG00000078061	v-raf murine sarcoma 3611 viral oncogene homolog
ARID1A	IMPACT-341	603024	8289	ENSG00000117713	AT rich interactive domain 1A (SWI-like)
ARID1B	IMPACT-341	614556	57492	ENSG00000049618	AT rich interactive domain 1B (SWI1-like)
ARID2	IMPACT-341	609539	196528	ENSG00000189079	AT rich interactive domain 2 (ARID, RFX-like)
ARID5B	IMPACT-341	608538	84159	ENSG00000150347	AT rich interactive domain 5B (MRF1-like)
ASXL1	IMPACT-341	612990	171023	ENSG00000171456	additional sex combs like 1 (Drosophila)
ASXL2	IMPACT-341	612991	55252	ENSG00000143970	additional sex combs like 2 (Drosophila)
ATM	IMPACT-341	607585	472	ENSG00000149311	ataxia telangiectasia mutated
ATR	IMPACT-341	601215	545	ENSG00000175054	ataxia telangiectasia and Rad3 related
ATRX	IMPACT-341	300032	546	ENSG00000085224	alpha thalassemia/mental retardation syndrome X-linked
AURKA	IMPACT-341	603072	6790	ENSG00000087586	aurora kinase A
AURKB	IMPACT-341	604970	9212	ENSG00000178999	aurora kinase B
AXIN1	IMPACT-341	603816	8312	ENSG00000103126	axin 1
AXIN2	IMPACT-341	604025	8313	ENSG00000168646	axin 2
AXL	IMPACT-341	109135	558	ENSG00000167601	AXL receptor tyrosine kinase
B2M	IMPACT-341	109700	567	ENSG00000166710	beta-2-microglobulin
BABAM1	IMPACT-468	612766	29086	ENSG00000105393	chromosome 19 open reading frame 62
BAP1	IMPACT-341	603089	8314	ENSG00000163930	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)

BARD1	IMPACT-341	601593	580	ENSG00000138376	BRCA1 associated RING domain 1
BBC3	IMPACT-341	605854	27113	ENSG00000105327	BCL2 binding component 3
BCL10	IMPACT-410	603517	8915	ENSG00000142867	B-cell CLL/lymphoma 10
BCL2	IMPACT-341	151430	596	ENSG00000171791	B-cell CLL/lymphoma 2
BCL2L1	IMPACT-341	600039	598	ENSG00000171552	BCL2-like 1
BCL2L11	IMPACT-341	603827	10018	ENSG00000153094	BCL2-like 11 (apoptosis facilitator)
BCL6	IMPACT-341	109565	604	ENSG00000113916	B-cell CLL/lymphoma 6
BCOR	IMPACT-341	300485	54880	ENSG00000183337	BCL6 corepressor
BIRC3	IMPACT-410	601721	330	ENSG00000023445	baculoviral IAP repeat-containing 3
BLM	IMPACT-341	604610	641	ENSG00000197299	Bloom syndrome, RecQ helicase-like
BMPR1A	IMPACT-341	601299	657	ENSG00000107779	bone morphogenetic protein receptor, type IA
BRAF	IMPACT-341	164757	673	ENSG00000157764	v-raf murine sarcoma viral oncogene homolog B1
BRCA1	IMPACT-341	113705	672	ENSG00000012048	breast cancer 1, early onset
BRCA2	IMPACT-341	600185	675	ENSG00000139618	breast cancer 2, early onset
BRD4	IMPACT-341	608749	23476	ENSG00000141867	bromodomain containing 4
BRIP1	IMPACT-341	605882	83990	ENSG00000136492	BRCA1 interacting protein C-terminal helicase 1
BTK	IMPACT-341	300300	695	ENSG00000010671	Bruton agammaglobulinemia tyrosine kinase
CALR	IMPACT-410	109091	811	ENSG00000179218	calreticulin
CARD11	IMPACT-341	607210	84433	ENSG00000198286	caspase recruitment domain family, member 11
CARM1	IMPACT-468	603934	10498	ENSG00000142453	coactivator-associated arginine methyltransferase 1
CASP8	IMPACT-341	601763	841	ENSG00000064012	caspase 8, apoptosis-related cysteine peptidase
CBFB	IMPACT-341	121360	865	ENSG00000067955	core-binding factor, beta subunit
CBL	IMPACT-341	165360	867	ENSG00000110395	Cas-Br-M (murine) ecotropic retroviral transforming sequence
CCND1	IMPACT-341	168461	595	ENSG00000110092	cyclin D1
CCND2	IMPACT-341	123833	894	ENSG00000118971	cyclin D2
CCND3	IMPACT-341	123834	896	ENSG00000112576	cyclin D3
CCNE1	IMPACT-341	123837	898	ENSG00000105173	cyclin E1
CD274	IMPACT-341	605402	29126	ENSG00000120217	CD274 molecule
CD276	IMPACT-341	605715	80381	ENSG00000103855	CD276 molecule
CD79A	IMPACT-410	112205	973	ENSG00000105369	CD79a molecule, immunoglobulin-associated alpha
CD79B	IMPACT-341	147245	974	ENSG00000007312	CD79b molecule, immunoglobulin-associated beta
CDC42	IMPACT-468	116952	998	ENSG00000070831	cell division cycle 42 (GTP binding protein, 25kDa)
CDC73	IMPACT-341	607393	79577	ENSG00000134371	cell division cycle 73, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)
CDH1	IMPACT-341	192090	999	ENSG00000039068	cadherin 1, type 1, E-cadherin (epithelial)
CDK12	IMPACT-341	615514	51755	ENSG00000167258	cyclin-dependent kinase 12

CDK4	IMPACT-341	123829	1019	ENSG00000135446	cyclin-dependent kinase 4
CDK6	IMPACT-341	603368	1021	ENSG00000105810	cyclin-dependent kinase 6
CDK8	IMPACT-341	603184	1024	ENSG00000132964	cyclin-dependent kinase 8
CDKN1A	IMPACT-341	116899	1026	ENSG00000124762	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
CDKN1B	IMPACT-341	600778	1027	ENSG00000111276	cyclin-dependent kinase inhibitor 1B (p27, Kip1)
CDKN2A	IMPACT-341	600160	1029	ENSG00000147889	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)
CDKN2B	IMPACT-341	600431	1030	ENSG00000147883	cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)
CDKN2C	IMPACT-341	603369	1031	ENSG00000123080	cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)
CEBPA	IMPACT-410	116897	1050	ENSG00000245848	CCAAT/enhancer binding protein (C/EBP), alpha
CENPA	IMPACT-410	117139	1058	ENSG00000115163	centromere protein A
CHEK1	IMPACT-341	603078	1111	ENSG00000149554	CHK1 checkpoint homolog (S. pombe)
CHEK2	IMPACT-341	604373	11200	ENSG00000183765	CHK2 checkpoint homolog (S. pombe)
CIC	IMPACT-341	612082	23152	ENSG00000079432	capicua homolog (Drosophila)
CREBBP	IMPACT-341	600140	1387	ENSG00000005339	CREB binding protein
CRKL	IMPACT-341	602007	1399	ENSG00000099942	v-crk sarcoma virus CT10 oncogene homolog (avian)-like
CRLF2	IMPACT-341	400023	64109	ENSG00000205755	cytokine receptor-like factor 2
CSDE1	IMPACT-468	191510	7812	ENSG00000009307	cold shock domain containing E1, RNA-binding
CSF1R	IMPACT-341	164770	1436	ENSG00000182578	colony stimulating factor 1 receptor
CSF3R	IMPACT-410	138971	1441	ENSG00000119535	colony stimulating factor 3 receptor (granulocyte)
CTCF	IMPACT-341	604167	10664	ENSG00000102974	CCCTC-binding factor (zinc finger protein)
CTLA4	IMPACT-341	123890	1493	ENSG00000163599	cytotoxic T-lymphocyte-associated protein 4
CTNNB1	IMPACT-341	116806	1499	ENSG00000168036	catenin (cadherin-associated protein), beta 1, 88kDa
CUL3	IMPACT-341	603136	8452	ENSG00000036257	cullin 3
CXCR4	IMPACT-410	162643	7852	ENSG00000121966	chemokine (C-X-C motif) receptor 4
CYLD	IMPACT-468	605018	1540	ENSG00000083799	cylindromatosis (turban tumor syndrome)
CYSLTR2	IMPACT-468	605666	57105	ENSG00000152207	cysteinyl leukotriene receptor 2
DAXX	IMPACT-341	603186	1616	ENSG00000204209	death-domain associated protein
DCUN1D1	IMPACT-341	605905	54165	ENSG00000043093	DCN1, defective in cullin neddylation 1, domain containing 1 (S. cerevisiae)
DDR2	IMPACT-341	191311	4921	ENSG00000162733	discoidin domain receptor tyrosine kinase 2
DICER1	IMPACT-341	606241	23405	ENSG00000100697	dicer 1, ribonuclease type III
DIS3	IMPACT-341	607533	22894	ENSG00000083520	DIS3 mitotic control homolog (S. cerevisiae)

DNAJB1	IMPACT-410	604572	3337	ENSG00000132002	DnaJ (Hsp40) homolog, subfamily B, member 1
DNMT1	IMPACT-341	126375	1786	ENSG00000130816	DNA (cytosine-5-)-methyltransferase 1
DNMT3A	IMPACT-341	602769	1788	ENSG00000119772	DNA (cytosine-5-)-methyltransferase 3 alpha
DNMT3B	IMPACT-341	602900	1789	ENSG00000088305	DNA (cytosine-5-)-methyltransferase 3 beta
DOT1L	IMPACT-341	607375	84444	ENSG00000104885	DOT1-like, histone H3 methyltransferase (S. cerevisiae)
DROSHA	IMPACT-468	608828	29102	ENSG00000113360	drosha, ribonuclease type III
DUSP4	IMPACT-468	602747	1846	ENSG00000120875	dual specificity phosphatase 4
E2F3	IMPACT-341	600427	1871	ENSG00000112242	E2F transcription factor 3
EED	IMPACT-341	605984	8726	ENSG00000074266	embryonic ectoderm development
EGFL7	IMPACT-341	608582	51162	ENSG00000172889	EGF-like-domain, multiple 7
EGFR	IMPACT-341	131550	1956	ENSG00000146648	epidermal growth factor receptor
EIF1AX	IMPACT-341	300186	1964	ENSG00000173674	eukaryotic translation initiation factor 1A, X-linked
EIF4A2	IMPACT-410	601102	1974	ENSG00000156976	eukaryotic translation initiation factor 4A2
EIF4E	IMPACT-410	133440	1977	ENSG00000151247	eukaryotic translation initiation factor 4E
ELF3	IMPACT-468	602191	1999	ENSG00000163435	E74-like factor 3 (ets domain transcription factor, epithelial-specific)
EP300	IMPACT-341	602700	2033	ENSG00000100393	E1A binding protein p300
EPAS1	IMPACT-468	603349	2034	ENSG00000116016	endothelial PAS domain protein 1
EPCAM	IMPACT-341	185535	4072	ENSG00000119888	epithelial cell adhesion molecule
EPHA3	IMPACT-341	179611	2042	ENSG00000044524	EPH receptor A3
EPHA5	IMPACT-341	600004	2044	ENSG00000145242	EPH receptor A5
EPHA7	IMPACT-410	602190	2045	ENSG00000135333	EPH receptor A7
EPHB1	IMPACT-341	600600	2047	ENSG00000154928	EPH receptor B1
ERBB2	IMPACT-341	164870	2064	ENSG00000141736	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
ERBB3	IMPACT-341	190151	2065	ENSG00000065361	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
ERBB4	IMPACT-341	600543	2066	ENSG00000178568	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)
ERCC2	IMPACT-341	126340	2068	ENSG00000104884	excision repair cross-complementing rodent repair deficiency, complementation group 2
ERCC3	IMPACT-341	133510	2071	ENSG00000163161	excision repair cross-complementing rodent repair deficiency, complementation group 3 (xeroderma pigmentosum group B complementing)
ERCC4	IMPACT-341	133520	2072	ENSG00000175595	excision repair cross-complementing rodent repair deficiency, complementation group 4

ERCC5	IMPACT-341	133530	2073	ENSG00000134899	excision repair cross-complementing rodent repair deficiency, complementation group 5
ERF	IMPACT-468	611888	2077	ENSG00000105722	Ets2 repressor factor
ERG	IMPACT-341	165080	2078	ENSG00000157554	v-ets erythroblastosis virus E26 oncogene homolog (avian)
ERRFI1	IMPACT-410	608069	54206	ENSG00000116285	ERBB receptor feedback inhibitor 1
ESR1	IMPACT-341	133430	2099	ENSG00000091831	estrogen receptor 1
ETV1	IMPACT-341	600541	2115	ENSG00000006468	ets variant 1
ETV6	IMPACT-341	600618	2120	ENSG00000139083	ets variant 6
EZH1	IMPACT-468	601674	2145	ENSG00000108799	enhancer of zeste homolog 1 (Drosophila)
EZH2	IMPACT-341	601573	2146	ENSG00000106462	enhancer of zeste homolog 2 (Drosophila)
FAM123B	IMPACT-341	300647	139285	ENSG00000184675	family with sequence similarity 123B
FAM175A	IMPACT-341	611143	84142	ENSG00000163322	family with sequence similarity 175, member A
FAM46C	IMPACT-341	613952	54855	ENSG00000183508	family with sequence similarity 46, member C
FAM58A	IMPACT-468	300708	92002	ENSG00000262919	family with sequence similarity 58, member A
FANCA	IMPACT-341	607139	2175	ENSG00000187741	Fanconi anemia, complementation group A
FANCC	IMPACT-341	613899	2176	ENSG00000158169	Fanconi anemia, complementation group C
FAT1	IMPACT-341	600976	2195	ENSG00000083857	FAT tumor suppressor homolog 1 (Drosophila)
FBXW7	IMPACT-341	606278	55294	ENSG00000109670	F-box and WD repeat domain containing 7
FGF19	IMPACT-341	603891	9965	ENSG00000162344	fibroblast growth factor 19
FGF3	IMPACT-341	164950	2248	ENSG00000186895	fibroblast growth factor 3
FGF4	IMPACT-341	164980	2249	ENSG00000075388	fibroblast growth factor 4
FGFR1	IMPACT-341	136350	2260	ENSG00000077782	fibroblast growth factor receptor 1
FGFR2	IMPACT-341	176943	2263	ENSG00000066468	fibroblast growth factor receptor 2
FGFR3	IMPACT-341	134934	2261	ENSG00000068078	fibroblast growth factor receptor 3
FGFR4	IMPACT-341	134935	2264	ENSG00000160867	fibroblast growth factor receptor 4
FH	IMPACT-341	136850	2271	ENSG00000091483	fumarate hydratase
FLCN	IMPACT-341	607273	201163	ENSG00000154803	folliculin
FLT1	IMPACT-341	165070	2321	ENSG00000102755	fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)
FLT3	IMPACT-341	136351	2322	ENSG00000122025	fms-related tyrosine kinase 3
FLT4	IMPACT-341	136352	2324	ENSG00000037280	fms-related tyrosine kinase 4
FOXA1	IMPACT-341	602294	3169	ENSG00000129514	forkhead box A1
FOXL2	IMPACT-341	605597	668	ENSG00000183770	forkhead box L2
FOXO1	IMPACT-410	136533	2308	ENSG00000150907	forkhead box O1

FOXP1	IMPACT-341	605515	27086	ENSG00000114861	forkhead box P1
FUBP1	IMPACT-341	603444	8880	ENSG00000162613	far upstream element (FUSE) binding protein 1
FYN	IMPACT-410	137025	2534	ENSG00000010810	FYN oncogene related to SRC, FGR, YES
GATA1	IMPACT-341	305371	2623	ENSG00000102145	GATA binding protein 1 (globin transcription factor 1)
GATA2	IMPACT-341	137295	2624	ENSG00000179348	GATA binding protein 2
GATA3	IMPACT-341	131320	2625	ENSG00000107485	GATA binding protein 3
GLI1	IMPACT-410	165220	2735	ENSG00000111087	GLI family zinc finger 1
GNA11	IMPACT-341	139313	2767	ENSG00000088256	guanine nucleotide binding protein (G protein), alpha 11 (Gq class)
GNAQ	IMPACT-341	600998	2776	ENSG00000156052	guanine nucleotide binding protein (G protein), q polypeptide
GNAS	IMPACT-341	139320	2778	ENSG00000087460	GNAS complex locus
GPS2	IMPACT-410	601935	2874	ENSG00000132522	G protein pathway suppressor 2
GREM1	IMPACT-341	603054	26585	ENSG00000166923	gremlin 1
GRIN2A	IMPACT-341	138253	2903	ENSG00000183454	glutamate receptor, ionotropic, N-methyl D-aspartate 2A
GSK3B	IMPACT-341	605004	2932	ENSG00000082701	glycogen synthase kinase 3 beta
H3F3A	IMPACT-410	601128	3020	ENSG00000163041	H3 histone, family 3A
H3F3B	IMPACT-410	601058	3021	ENSG00000132475	H3 histone, family 3B (H3.3B)
H3F3C	IMPACT-341	616134	440093	ENSG00000188375	H3 histone, family 3C
HGF	IMPACT-341	142409	3082	ENSG00000019991	hepatocyte growth factor (hepapoietin A; scatter factor)
HIST1H1C	IMPACT-341	142710	3006	ENSG00000187837	histone cluster 1, H1c
HIST1H2BD	IMPACT-341	602799	3017	ENSG00000158373	histone cluster 1, H2bd
HIST1H3A	IMPACT-410	602810	8350	ENSG00000275714	histone cluster 1, H3a
HIST1H3B	IMPACT-341	602819	8358	ENSG00000286522	histone cluster 1, H3b
HIST1H3C	IMPACT-410	602812	8352	ENSG00000287080	histone cluster 1, H3c
HIST1H3D	IMPACT-410	602811	8351	ENSG00000197409	histone cluster 1, H3d
HIST1H3E	IMPACT-410	602813	8353	ENSG00000274750	histone cluster 1, H3e
HIST1H3F	IMPACT-410	602816	8968	ENSG00000277775	histone cluster 1, H3f
HIST1H3G	IMPACT-410	602815	8355	ENSG00000273983	histone cluster 1, H3g
HIST1H3H	IMPACT-410	602818	8357	ENSG00000278828	histone cluster 1, H3h
HIST1H3I	IMPACT-410	602814	8354	ENSG00000275379	histone cluster 1, H3i
HIST1H3J	IMPACT-410	602817	8356	ENSG00000197153	histone cluster 1, H3j
HIST2H3C	IMPACT-410	142780	126961	ENSG00000203811	histone cluster 2, H3c
HIST2H3D	IMPACT-410	NA	653604	ENSG00000183598	histone cluster 2, H3d
HIST3H3	IMPACT-410	602820	8290	ENSG00000168148	histone cluster 3, H3
HLA-A	IMPACT-410	142800	3105	ENSG00000206503	major histocompatibility complex, class I, A
HLA-B	IMPACT-468	142830	3106	ENSG00000234745	major histocompatibility complex, class I, B
HNF1A	IMPACT-341	142410	6927	ENSG00000135100	HNF1 homeobox A
HOXB13	IMPACT-410	604607	10481	ENSG00000159184	homeobox B13
HRAS	IMPACT-341	190020	3265	ENSG00000174775	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
ICOSLG	IMPACT-341	605717	23308	ENSG00000160223	inducible T-cell co-stimulator ligand
ID3	IMPACT-410	600277	3399	ENSG00000117318	inhibitor of DNA binding 3, dominant negative helix-loop-helix protein

IDH1	IMPACT-341	147700	3417	ENSG00000138413	isocitrate dehydrogenase 1 (NADP+), soluble
IDH2	IMPACT-341	147650	3418	ENSG00000182054	isocitrate dehydrogenase 2 (NADP+), mitochondrial
IFNGR1	IMPACT-341	107470	3459	ENSG00000027697	interferon gamma receptor 1
IGF1	IMPACT-341	147440	3479	ENSG00000017427	insulin-like growth factor 1 (somatomedin C)
IGF1R	IMPACT-341	147370	3480	ENSG00000140443	insulin-like growth factor 1 receptor
IGF2	IMPACT-341	147470	3481	ENSG00000167244	insulin-like growth factor 2 (somatomedin A)
IKBKE	IMPACT-341	605048	9641	ENSG00000263528	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase epsilon
IKZF1	IMPACT-341	603023	10320	ENSG00000185811	IKAROS family zinc finger 1 (Ikaros)
IL10	IMPACT-341	124092	3586	ENSG00000136634	interleukin 10
IL7R	IMPACT-341	146661	3575	ENSG00000168685	interleukin 7 receptor
INHA	IMPACT-410	147380	3623	ENSG00000123999	inhibin, alpha
INHBA	IMPACT-410	147290	3624	ENSG00000122641	inhibin, beta A
INPP4A	IMPACT-341	600916	3631	ENSG00000040933	inositol polyphosphate-4-phosphatase, type I, 107kDa
INPP4B	IMPACT-341	607494	8821	ENSG00000109452	inositol polyphosphate-4-phosphatase, type II, 105kDa
INPPL1	IMPACT-468	600829	3636	ENSG00000165458	inositol polyphosphate phosphatase-like 1
INSR	IMPACT-341	147670	3643	ENSG00000171105	insulin receptor
IRF4	IMPACT-341	601900	3662	ENSG00000137265	interferon regulatory factor 4
IRS1	IMPACT-341	147545	3667	ENSG00000169047	insulin receptor substrate 1
IRS2	IMPACT-341	600797	8660	ENSG00000185950	insulin receptor substrate 2
JAK1	IMPACT-341	147795	3716	ENSG00000162434	Janus kinase 1
JAK2	IMPACT-341	147796	3717	ENSG00000096968	Janus kinase 2
JAK3	IMPACT-341	600173	3718	ENSG00000105639	Janus kinase 3
JUN	IMPACT-341	165160	3725	ENSG00000177606	jun proto-oncogene
KDM5A	IMPACT-341	180202	5927	ENSG00000073614	lysine (K)-specific demethylase 5A
KDM5C	IMPACT-341	314690	8242	ENSG00000126012	lysine (K)-specific demethylase 5C
KDM6A	IMPACT-341	300128	7403	ENSG00000147050	lysine (K)-specific demethylase 6A
KDR	IMPACT-341	191306	3791	ENSG00000128052	kinase insert domain receptor (a type III receptor tyrosine kinase)
KEAP1	IMPACT-341	606016	9817	ENSG00000079999	kelch-like ECH-associated protein 1
KIT	IMPACT-341	164920	3815	ENSG00000157404	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
KLF4	IMPACT-341	602253	9314	ENSG00000136826	Kruppel-like factor 4 (gut)
KMT2B	IMPACT-468	606834	9757	ENSG00000272333	myeloid/lymphoid or mixed-lineage leukemia 4
KMT5A	IMPACT-468	607240	387893	ENSG00000183955	SET domain containing (lysine methyltransferase) 8
KNSTRN	IMPACT-468	614718	90417	ENSG00000128944	chromosome 15 open reading frame 23

KRAS	IMPACT-341	190070	3845	ENSG00000133703	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LATS1	IMPACT-341	603473	9113	ENSG00000131023	LATS, large tumor suppressor, homolog 1 (Drosophila)
LATS2	IMPACT-341	604861	26524	ENSG00000150457	LATS, large tumor suppressor, homolog 2 (Drosophila)
LMO1	IMPACT-341	186921	4004	ENSG00000166407	LIM domain only 1 (rhombotin 1)
LYN	IMPACT-468	165120	4067	ENSG00000254087	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog
MALT1	IMPACT-410	604860	10892	ENSG00000172175	mucosa associated lymphoid tissue lymphoma translocation gene 1
MAP2K1	IMPACT-341	176872	5604	ENSG00000169032	mitogen-activated protein kinase kinase 1
MAP2K2	IMPACT-341	601263	5605	ENSG00000126934	mitogen-activated protein kinase kinase 2
MAP2K4	IMPACT-341	601335	6416	ENSG00000065559	mitogen-activated protein kinase kinase 4
MAP3K1	IMPACT-341	600982	4214	ENSG00000095015	mitogen-activated protein kinase kinase kinase 1
MAP3K13	IMPACT-341	604915	9175	ENSG00000073803	mitogen-activated protein kinase kinase kinase 13
MAP3K14	IMPACT-410	604655	9020	ENSG00000006062	mitogen-activated protein kinase kinase kinase 14
MAPK1	IMPACT-341	176948	5594	ENSG00000100030	mitogen-activated protein kinase 1
MAPK3	IMPACT-410	601795	5595	ENSG00000102882	mitogen-activated protein kinase 3
MAPKAP1	IMPACT-468	610558	79109	ENSG00000119487	mitogen-activated protein kinase associated protein 1
MAX	IMPACT-341	154950	4149	ENSG00000125952	MYC associated factor X
MCL1	IMPACT-341	159552	4170	ENSG00000143384	myeloid cell leukemia sequence 1 (BCL2-related)
MDC1	IMPACT-341	607593	9656	ENSG00000137337	mediator of DNA-damage checkpoint 1
MDM2	IMPACT-341	164785	4193	ENSG00000135679	Mdm2 p53 binding protein homolog (mouse)
MDM4	IMPACT-341	602704	4194	ENSG00000198625	Mdm4 p53 binding protein homolog (mouse)
MED12	IMPACT-341	300188	9968	ENSG00000184634	mediator complex subunit 12
MEF2B	IMPACT-341	600661	100271849	ENSG00000213999	myocyte enhancer factor 2B
MEN1	IMPACT-341	613733	4221	ENSG00000133895	multiple endocrine neoplasia I
MET	IMPACT-341	164860	4233	ENSG00000105976	met proto-oncogene (hepatocyte growth factor receptor)
MGA	IMPACT-410	616061	23269	ENSG00000174197	MAX gene associated
MITF	IMPACT-341	156845	4286	ENSG00000187098	microphthalmia-associated transcription factor
MLH1	IMPACT-341	120436	4292	ENSG00000076242	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)
MLL	IMPACT-341	159555	4297	ENSG00000118058	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)
MLL2	IMPACT-341	602113	8085	ENSG00000167548	myeloid/lymphoid or mixed-lineage leukemia 2

MLL3	IMPACT-341	606833	58508	ENSG00000055609	myeloid/lymphoid or mixed-lineage leukemia 3
MPL	IMPACT-341	159530	4352	ENSG00000117400	myeloproliferative leukemia virus oncogene
MRE11A	IMPACT-341	600814	4361	ENSG00000020922	MRE11 meiotic recombination 11 homolog A (S. cerevisiae)
MSH2	IMPACT-341	609309	4436	ENSG00000095002	mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli)
MSH3	IMPACT-468	600887	4437	ENSG00000113318	mutS homolog 3 (E. coli)
MSH6	IMPACT-341	600678	2956	ENSG00000116062	mutS homolog 6 (E. coli)
MSI1	IMPACT-468	603328	4440	ENSG00000135097	musashi homolog 1 (Drosophila)
MSI2	IMPACT-468	607897	124540	ENSG00000153944	musashi homolog 2 (Drosophila)
MST1	IMPACT-410	142408	4485	ENSG00000173531	macrophage stimulating 1 (hepatocyte growth factor-like)
MST1R	IMPACT-410	600168	4486	ENSG00000164078	macrophage stimulating 1 receptor (c-met-related tyrosine kinase)
MTOR	IMPACT-341	601231	2475	ENSG00000198793	mechanistic target of rapamycin (serine/threonine kinase)
MUTYH	IMPACT-341	604933	4595	ENSG00000132781	mutY homolog (E. coli)
MYC	IMPACT-341	190080	4609	ENSG00000136997	v-myc myelocytomatosis viral oncogene homolog (avian)
MYCL1	IMPACT-341	164850	4610	ENSG00000116990	v-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma derived (avian)
MYCN	IMPACT-341	164840	4613	ENSG00000134323	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)
MYD88	IMPACT-341	602170	4615	ENSG00000172936	myeloid differentiation primary response gene (88)
MYOD1	IMPACT-341	159970	4654	ENSG00000129152	myogenic differentiation 1
NBN	IMPACT-341	602667	4683	ENSG00000104320	nibrin
NCOA3	IMPACT-410	601937	8202	ENSG00000124151	nuclear receptor coactivator 3
NCOR1	IMPACT-341	600849	9611	ENSG00000141027	nuclear receptor corepressor 1
NEGR1	IMPACT-410	613173	257194	ENSG00000172260	neuronal growth regulator 1
NF1	IMPACT-341	613113	4763	ENSG00000196712	neurofibromin 1
NF2	IMPACT-341	607379	4771	ENSG00000186575	neurofibromin 2 (merlin)
NFE2L2	IMPACT-341	600492	4780	ENSG00000116044	nuclear factor (erythroid-derived 2)-like 2
NFKBIA	IMPACT-410	164008	4792	ENSG00000100906	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha
NKX2-1	IMPACT-341	600635	7080	ENSG00000136352	NK2 homeobox 1
NKX3-1	IMPACT-341	602041	4824	ENSG00000167034	NK3 homeobox 1
NOTCH1	IMPACT-341	190198	4851	ENSG00000148400	notch 1
NOTCH2	IMPACT-341	600275	4853	ENSG00000134250	notch 2
NOTCH3	IMPACT-341	600276	4854	ENSG00000074181	notch 3
NOTCH4	IMPACT-341	164951	4855	ENSG00000204301	notch 4
NPM1	IMPACT-341	164040	4869	ENSG00000181163	nucleophosmin (nucleolar phosphoprotein B23, numatrin)
NRAS	IMPACT-341	164790	4893	ENSG00000213281	neuroblastoma RAS viral (v-ras) oncogene homolog
NSD1	IMPACT-341	606681	64324	ENSG00000165671	nuclear receptor binding SET domain protein 1

NTHL1	IMPACT-468	602656	4913	ENSG00000065057	nth endonuclease III-like 1 (E. coli)
NTRK1	IMPACT-341	191315	4914	ENSG00000198400	neurotrophic tyrosine kinase, receptor, type 1
NTRK2	IMPACT-341	600456	4915	ENSG00000148053	neurotrophic tyrosine kinase, receptor, type 2
NTRK3	IMPACT-341	191316	4916	ENSG00000140538	neurotrophic tyrosine kinase, receptor, type 3
NUF2	IMPACT-468	611772	83540	ENSG00000143228	NUF2, NDC80 kinetochore complex component, homolog (S. cerevisiae)
NUP93	IMPACT-410	614351	9688	ENSG00000102900	nucleoporin 93kDa
PAK1	IMPACT-341	602590	5058	ENSG00000149269	p21 protein (Cdc42/Rac)-activated kinase 1
PAK7	IMPACT-341	608038	57144	ENSG00000101349	p21 protein (Cdc42/Rac)-activated kinase 7
PALB2	IMPACT-341	610355	79728	ENSG00000083093	partner and localizer of BRCA2
PARK2	IMPACT-341	602544	5071	ENSG00000185345	parkinson protein 2, E3 ubiquitin protein ligase (parkin)
PARP1	IMPACT-341	173870	142	ENSG00000143799	poly (ADP-ribose) polymerase 1
PAX5	IMPACT-341	167414	5079	ENSG00000196092	paired box 5
PBRM1	IMPACT-341	606083	55193	ENSG00000163939	polybromo 1
PDCD1	IMPACT-341	600244	5133	ENSG00000188389	programmed cell death 1
PDCD1LG2	IMPACT-468	605723	80380	ENSG00000197646	programmed cell death 1 ligand 2
PDGFRA	IMPACT-341	173490	5156	ENSG00000134853	platelet-derived growth factor receptor, alpha polypeptide
PDGFRB	IMPACT-341	173410	5159	ENSG00000113721	platelet-derived growth factor receptor, beta polypeptide
PDPK1	IMPACT-341	605213	5170	ENSG00000140992	3-phosphoinositide dependent protein kinase-1
PGR	IMPACT-410	607311	5241	ENSG00000082175	progesterone receptor
PHOX2B	IMPACT-341	603851	8929	ENSG00000109132	paired-like homeobox 2b
PIK3C2G	IMPACT-341	609001	5288	ENSG00000139144	phosphoinositide-3-kinase, class 2, gamma polypeptide
PIK3C3	IMPACT-341	602609	5289	ENSG00000078142	phosphoinositide-3-kinase, class 3
PIK3CA	IMPACT-341	171834	5290	ENSG00000121879	phosphoinositide-3-kinase, catalytic, alpha polypeptide
PIK3CB	IMPACT-341	602925	5291	ENSG00000051382	phosphoinositide-3-kinase, catalytic, beta polypeptide
PIK3CD	IMPACT-341	602839	5293	ENSG00000171608	phosphoinositide-3-kinase, catalytic, delta polypeptide
PIK3CG	IMPACT-341	601232	5294	ENSG00000105851	phosphoinositide-3-kinase, catalytic, gamma polypeptide
PIK3R1	IMPACT-341	171833	5295	ENSG00000145675	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)
PIK3R2	IMPACT-341	603157	5296	ENSG00000105647	phosphoinositide-3-kinase, regulatory subunit 2 (beta)
PIK3R3	IMPACT-341	606076	8503	ENSG00000117461	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)
PIM1	IMPACT-341	164960	5292	ENSG00000137193	pim-1 oncogene
PLCG2	IMPACT-410	600220	5336	ENSG00000197943	phospholipase C, gamma 2 (phosphatidylinositol-specific)
PLK2	IMPACT-341	607023	10769	ENSG00000145632	polo-like kinase 2

PMAIP1	IMPACT-341	604959	5366	ENSG00000141682	phorbol-12-myristate-13-acetate-induced protein 1
PMS1	IMPACT-341	600258	5378	ENSG00000064933	PMS1 postmeiotic segregation increased 1 (S. cerevisiae)
PMS2	IMPACT-341	600259	5395	ENSG00000122512	PMS2 postmeiotic segregation increased 2 (S. cerevisiae)
PNRC1	IMPACT-341	606714	10957	ENSG00000146278	proline-rich nuclear receptor coactivator 1
POLD1	IMPACT-410	174761	5424	ENSG00000062822	polymerase (DNA directed), delta 1, catalytic subunit 125kDa
POLE	IMPACT-341	174762	5426	ENSG00000177084	polymerase (DNA directed), epsilon
PPARG	IMPACT-468	601487	5468	ENSG00000132170	peroxisome proliferator-activated receptor gamma
PPM1D	IMPACT-410	605100	8493	ENSG00000170836	protein phosphatase, Mg2+/Mn2+ dependent, 1D
PPP2R1A	IMPACT-341	605983	5518	ENSG00000105568	protein phosphatase 2, regulatory subunit A, alpha
PPP4R2	IMPACT-468	613822	151987	ENSG00000163605	protein phosphatase 4, regulatory subunit 2
PPP6C	IMPACT-410	612725	5537	ENSG00000119414	protein phosphatase 6, catalytic subunit
PRDM1	IMPACT-341	603423	639	ENSG00000057657	PR domain containing 1, with ZNF domain
PRDM14	IMPACT-468	611781	63978	ENSG00000147596	PR domain containing 14
PREX2	IMPACT-468	612139	80243	ENSG00000046889	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2
PRKAR1A	IMPACT-341	188830	5573	ENSG00000108946	protein kinase, cAMP-dependent, regulatory, type I, alpha (tissue specific extinguisher 1)
PRKCI	IMPACT-468	600539	5584	ENSG00000163558	protein kinase C, iota
PRKD1	IMPACT-468	605435	5587	ENSG00000184304	protein kinase D1
PTCH1	IMPACT-341	601309	5727	ENSG00000185920	patched 1
PTEN	IMPACT-341	601728	5728	ENSG00000171862	phosphatase and tensin homolog
PTP4A1	IMPACT-468	601585	7803	ENSG00000112245	protein tyrosine phosphatase type IVA, member 1
PTPN11	IMPACT-341	176876	5781	ENSG00000179295	protein tyrosine phosphatase, non-receptor type 11
PTPRD	IMPACT-341	601598	5789	ENSG00000153707	protein tyrosine phosphatase, receptor type, D
PTPRS	IMPACT-341	601576	5802	ENSG00000105426	protein tyrosine phosphatase, receptor type, S
PTPRT	IMPACT-341	608712	11122	ENSG00000196090	protein tyrosine phosphatase, receptor type, T
RAB35	IMPACT-410	604199	11021	ENSG00000111737	RAB35, member RAS oncogene family
RAC1	IMPACT-341	602048	5879	ENSG00000136238	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)
RAC2	IMPACT-468	602049	5880	ENSG00000128340	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)
RAD21	IMPACT-410	606462	5885	ENSG00000164754	RAD21 homolog (S. pombe)
RAD50	IMPACT-341	604040	10111	ENSG00000113522	RAD50 homolog (S. cerevisiae)

RAD51	IMPACT-341	179617	5888	ENSG00000051180	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)
RAD51C	IMPACT-341	602774	5889	ENSG00000108384	RAD51 homolog C (S. cerevisiae)
RAD51L1	IMPACT-341	602948	5890	ENSG00000182185	RAD51-like 1 (S. cerevisiae)
RAD51L3	IMPACT-341	602954	5892	ENSG00000185379	RAD51-like 3 (S. cerevisiae)
RAD52	IMPACT-341	600392	5893	ENSG00000002016	RAD52 homolog (S. cerevisiae)
RAD54L	IMPACT-341	603615	8438	ENSG00000085999	RAD54-like (S. cerevisiae)
RAF1	IMPACT-341	164760	5894	ENSG00000132155	v-raf-1 murine leukemia viral oncogene homolog 1
RARA	IMPACT-341	180240	5914	ENSG00000131759	retinoic acid receptor, alpha
RASA1	IMPACT-341	139150	5921	ENSG00000145715	RAS p21 protein activator (GTPase activating protein) 1
RB1	IMPACT-341	614041	5925	ENSG00000139687	retinoblastoma 1
RBM10	IMPACT-341	300080	8241	ENSG00000182872	RNA binding motif protein 10
RECQL	IMPACT-468	600537	5965	ENSG00000004700	RecQ protein-like (DNA helicase Q1-like)
RECQL4	IMPACT-341	603780	9401	ENSG00000160957	RecQ protein-like 4
REL	IMPACT-341	164910	5966	ENSG00000162924	v-rel reticuloendotheliosis viral oncogene homolog (avian)
RET	IMPACT-341	164761	5979	ENSG00000165731	ret proto-oncogene
RFWD2	IMPACT-341	608067	64326	ENSG00000143207	ring finger and WD repeat domain 2
RHEB	IMPACT-410	601293	6009	ENSG00000106615	Ras homolog enriched in brain
RHOA	IMPACT-341	165390	387	ENSG00000067560	ras homolog gene family, member A
RICTOR	IMPACT-341	609022	253260	ENSG00000164327	RPTOR independent companion of MTOR, complex 2
RIT1	IMPACT-341	609591	6016	ENSG00000143622	Ras-like without CAAX 1
RNF43	IMPACT-341	612482	54894	ENSG00000108375	ring finger protein 43
ROS1	IMPACT-341	165020	6098	ENSG00000047936	c-ros oncogene 1 , receptor tyrosine kinase
RPS6KA4	IMPACT-341	603606	8986	ENSG00000162302	ribosomal protein S6 kinase, 90kDa, polypeptide 4
RPS6KB2	IMPACT-341	608939	6199	ENSG00000175634	ribosomal protein S6 kinase, 70kDa, polypeptide 2
RPTOR	IMPACT-341	607130	57521	ENSG00000141564	regulatory associated protein of MTOR, complex 1
RRAGC	IMPACT-468	608267	64121	ENSG00000116954	Ras-related GTP binding C
RRAS	IMPACT-468	165090	6237	ENSG00000126458	related RAS viral (r-ras) oncogene homolog
RRAS2	IMPACT-468	600098	22800	ENSG00000133818	related RAS viral (r-ras) oncogene homolog 2
RTEL1	IMPACT-468	608833	51750	ENSG00000258366	regulator of telomere elongation helicase 1
RUNX1	IMPACT-341	151385	861	ENSG00000159216	runt-related transcription factor 1
RXRA	IMPACT-468	180245	6256	ENSG00000186350	retinoid X receptor, alpha
RYBP	IMPACT-341	607535	23429	ENSG00000163602	RING1 and YY1 binding protein
SDHA	IMPACT-341	600857	6389	ENSG00000073578	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
SDHAF2	IMPACT-341	613019	54949	ENSG00000167985	succinate dehydrogenase complex assembly factor 2

SDHB	IMPACT-341	185470	6390	ENSG00000117118	succinate dehydrogenase complex, subunit B, iron sulfur (Ip)
SDHC	IMPACT-341	602413	6391	ENSG00000143252	succinate dehydrogenase complex, subunit C, integral membrane protein, 15kDa
SDHD	IMPACT-341	602690	6392	ENSG00000204370	succinate dehydrogenase complex, subunit D, integral membrane protein
SESN1	IMPACT-468	606103	27244	ENSG00000080546	sestrin 1
SESN2	IMPACT-468	607767	83667	ENSG00000130766	sestrin 2
SESN3	IMPACT-468	607768	143686	ENSG00000149212	sestrin 3
SETD2	IMPACT-341	612778	29072	ENSG00000181555	SET domain containing 2
SF3B1	IMPACT-341	605590	23451	ENSG00000115524	splicing factor 3b, subunit 1, 155kDa
SH2B3	IMPACT-410	605093	10019	ENSG00000111252	SH2B adaptor protein 3
SH2D1A	IMPACT-341	300490	4068	ENSG00000183918	SH2 domain containing 1A
SHOC2	IMPACT-468	602775	8036	ENSG00000108061	soc-2 suppressor of clear homolog (C. elegans)
SHQ1	IMPACT-341	613663	55164	ENSG00000144736	SHQ1 homolog (S. cerevisiae)
SLX4	IMPACT-468	613278	84464	ENSG00000188827	SLX4 structure-specific endonuclease subunit homolog (S. cerevisiae)
SMAD2	IMPACT-341	601366	4087	ENSG00000175387	SMAD family member 2
SMAD3	IMPACT-341	603109	4088	ENSG00000166949	SMAD family member 3
SMAD4	IMPACT-341	600993	4089	ENSG00000141646	SMAD family member 4
SMARCA4	IMPACT-341	603254	6597	ENSG00000127616	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
SMARCB1	IMPACT-341	601607	6598	ENSG00000099956	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
SMARCD1	IMPACT-341	601735	6602	ENSG00000066117	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1
SMO	IMPACT-341	601500	6608	ENSG00000128602	smoothened homolog (Drosophila)
SMYD3	IMPACT-468	608783	64754	ENSG00000185420	SET and MYND domain containing 3
SOCS1	IMPACT-341	603597	8651	ENSG00000185338	suppressor of cytokine signaling 1
SOS1	IMPACT-468	182530	6654	ENSG00000115904	son of sevenless homolog 1 (Drosophila)
SOX17	IMPACT-341	610928	64321	ENSG00000164736	SRY (sex determining region Y)-box 17
SOX2	IMPACT-341	184429	6657	ENSG00000181449	SRY (sex determining region Y)-box 2
SOX9	IMPACT-341	608160	6662	ENSG00000125398	SRY (sex determining region Y)-box 9
SPEN	IMPACT-341	613484	23013	ENSG00000065526	spen homolog, transcriptional regulator (Drosophila)
SPOP	IMPACT-341	602650	8405	ENSG00000121067	speckle-type POZ protein

SPRED1	IMPACT-468	609291	161742	ENSG00000166068	sprouty-related, EVH1 domain containing 1
SRC	IMPACT-341	190090	6714	ENSG00000197122	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
SRSF2	IMPACT-410	600813	6427	ENSG00000161547	serine/arginine-rich splicing factor 2
STAG2	IMPACT-341	300826	10735	ENSG00000101972	stromal antigen 2
STAT3	IMPACT-410	102582	6774	ENSG00000168610	signal transducer and activator of transcription 3 (acute-phase response factor)
STAT5A	IMPACT-410	601511	6776	ENSG00000126561	signal transducer and activator of transcription 5A
STAT5B	IMPACT-410	604260	6777	ENSG00000173757	signal transducer and activator of transcription 5B
STK11	IMPACT-341	602216	6794	ENSG00000118046	serine/threonine kinase 11
STK19	IMPACT-468	604977	8859	ENSG00000204344	serine/threonine kinase 19
STK40	IMPACT-341	609437	83931	ENSG00000196182	serine/threonine kinase 40
SUFU	IMPACT-341	607035	51684	ENSG00000107882	suppressor of fused homolog (Drosophila)
SUZ12	IMPACT-341	606245	23512	ENSG00000178691	suppressor of zeste 12 homolog (Drosophila)
SYK	IMPACT-341	600085	6850	ENSG00000165025	spleen tyrosine kinase
TAP1	IMPACT-468	170260	6890	ENSG00000168394	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)
TAP2	IMPACT-468	170261	6891	ENSG00000204267	transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)
TBX3	IMPACT-341	601621	6926	ENSG00000135111	T-box 3
TCEB1	IMPACT-410	600788	6921	ENSG00000154582	transcription elongation factor B (SIII), polypeptide 1 (15kDa, elongin C)
TCF3	IMPACT-410	147141	6929	ENSG00000071564	transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
TCF7L2	IMPACT-410	602228	6934	ENSG00000148737	transcription factor 7-like 2 (T-cell specific, HMG-box)
TEK	IMPACT-468	600221	7010	ENSG00000120156	TEK tyrosine kinase, endothelial
TERT	IMPACT-341	187270	7015	ENSG00000164362	telomerase reverse transcriptase
TET1	IMPACT-341	607790	80312	ENSG00000138336	tet oncogene 1
TET2	IMPACT-341	612839	54790	ENSG00000168769	tet oncogene family member 2
TGFBR1	IMPACT-341	190181	7046	ENSG00000106799	transforming growth factor, beta receptor 1
TGFBR2	IMPACT-341	190182	7048	ENSG00000163513	transforming growth factor, beta receptor II (70/80kDa)
TMEM127	IMPACT-341	613403	55654	ENSG00000135956	transmembrane protein 127
TMPRSS2	IMPACT-341	602060	7113	ENSG00000184012	transmembrane protease, serine 2
TNFAIP3	IMPACT-341	191163	7128	ENSG00000118503	tumor necrosis factor, alpha-induced protein 3
TNFRSF14	IMPACT-341	602746	8764	ENSG00000157873	tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)
TOP1	IMPACT-341	126420	7150	ENSG00000198900	topoisomerase (DNA) I

TP53	IMPACT-341	191170	7157	ENSG00000141510	tumor protein p53
TP53BP1	IMPACT-468	605230	7158	ENSG00000067369	tumor protein p53 binding protein 1
TP63	IMPACT-341	603273	8626	ENSG00000073282	tumor protein p63
TRAF2	IMPACT-410	601895	7186	ENSG00000127191	TNF receptor-associated factor 2
TRAF7	IMPACT-341	606692	84231	ENSG00000131653	TNF receptor-associated factor 7
TSC1	IMPACT-341	605284	7248	ENSG00000165699	tuberous sclerosis 1
TSC2	IMPACT-341	191092	7249	ENSG00000103197	tuberous sclerosis 2
TSHR	IMPACT-341	603372	7253	ENSG00000165409	thyroid stimulating hormone receptor
U2AF1	IMPACT-341	191317	7307	ENSG00000160201	U2 small nuclear RNA auxiliary factor 1
UPF1	IMPACT-468	601430	5976	ENSG00000005007	UPF1 regulator of nonsense transcripts homolog (yeast)
VEGFA	IMPACT-410	192240	7422	ENSG00000112715	vascular endothelial growth factor A
VHL	IMPACT-341	608537	7428	ENSG00000134086	von Hippel-Lindau tumor suppressor
VTCN1	IMPACT-341	608162	79679	ENSG00000134258	V-set domain containing T cell activation inhibitor 1
WHSC1	IMPACT-468	602952	7468	ENSG00000109685	Wolf-Hirschhorn syndrome candidate 1
WHSC1L1	IMPACT-468	607083	54904	ENSG00000147548	Wolf-Hirschhorn syndrome candidate 1-like 1
WT1	IMPACT-341	607102	7490	ENSG00000184937	Wilms tumor 1
WWTR1	IMPACT-468	607392	25937	ENSG00000018408	WW domain containing transcription regulator 1
XIAP	IMPACT-341	300079	331	ENSG00000101966	X-linked inhibitor of apoptosis
XPO1	IMPACT-341	602559	7514	ENSG00000082898	exportin 1 (CRM1 homolog, yeast)
XRCC2	IMPACT-410	600375	7516	ENSG00000196584	X-ray repair complementing defective repair in Chinese hamster cells 2
YAP1	IMPACT-341	606608	10413	ENSG00000137693	Yes-associated protein 1
YES1	IMPACT-341	164880	7525	ENSG00000176105	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1
ZFHX3	IMPACT-410	104155	463	ENSG00000140836	zinc finger homeobox 3
ZRSR2	IMPACT-410	300028	8233	ENSG00000169249	zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2

eTable 3. Univariable and Multivariable Cox Proportional Hazards Models for Relapse-Free Survival, Using Clinicopathologic Variables*

Variable	Univariable		Multivariable	
	HR (95% CI)	***P***	**HR (95% CI)**	***P***
Age at surgery	1.01 (0.99-1.04)	0.2	-	-
Sex (vs female)	1.00 (-)	-	-	-
Male	1.31 (0.86-2.01)	0.2	-	-
Ever smoker (vs never)	1.04 (0.64-1.69)	0.9	-	-
Primary tumor SUVmax	1.06 (1.02-1.10)	0.003	-	-
Operative approach (vs minimally invasive)	1.00 (-)	-	1.00 (-)	-
Open (thoracotomy)	2.08 (1.31-3.31)	0.003	1.83 (1.14-2.94)	0.012
Resection (vs lobectomy/pneumonectomy)	1.00 (-)	-	1.00 (-)	-
Sublobar (wedge/segment)	1.86 (1.10-3.15)	0.02	1.96 (1.16-3.31)	0.013
Lymphovascular invasion present	2.61 (1.61-4.23)	<0.001	2.44 (1.48-4.05)	0.001
Spread of tumor through air spaces present	1.61 (0.76-3.40)	0.2	-	-
Visceral pleural invasion present	1.75 (1.08-2.82)	0.02	-	-
Adjuvant therapy	0.84 (0.44-1.64)	0.6	-	-
Predominant histologic subtype (vs nonsolid)	1.00 (-)	-	-	-
Solid	1.96 (1.19-3.20)	0.008	1.74 (1.05-2.89)	0.031
Pathologic stage (vs I)	1.00 (-)	-	1.00 (-)	-
II	3.28 (1.94-5.54)	<0.001	2.39 (1.28-4.45)	0.006
III	7.21 (4.46-11.68)	<0.001	3.64 (1.92-6.90)	<0.001

CI, confidence interval; HR, hazard ratio; SUVmax, maximum standardized uptake value.
*Univariable models were stratified by pathologic stage (except for stage variable).

eTable 4. Univariable and Clinicopathologic (CP)-Adjusted Multivariable Cox Model for Relapse-Free Survival

Genomic Factor/Gene	Median (IQR) or n (%)	Univariable			CP-Adjusted Multivariable	
		HR (95% CI)	***P***	**FDR-*P***	**HR (95% CI)**	***P***
TMB	4.7 (2.5-8.2)	1.02 (1.01-1.03)	<0.001	NA	1.01 (1.00-1.03)	0.074
FGA (x100)	3.8 (0.4-11.1)	1.04 (1.02-1.05)	<0.001	NA	1.03 (1.01-1.04)	0.005
WGD	92 (22%)	1.59 (1.02-2.50)	0.05	NA	-	-
ALK	5 (1.2%)	NA	NA	NA	-	-
APC	5 (1.2%)	NA	NA	NA	-	-
ARID1A	7 (1.6%)	0.79 (0.11-5.67)	0.8	0.9	-	-
ARID2	14 (3.3%)	0.32 (0.04-2.32)	0.3	0.5	-	-
ATM	10 (2.3%)	2.00 (0.73-5.46)	0.2	0.4	-	-
B2M	3 (0.7%)	NA	NA	NA	-	-
BAP1	4 (0.9%)	NA	NA	NA	-	-
BCOR	3 (0.7%)	NA	NA	NA	-	-
BRAF	20 (4.7%)	0.47 (0.11-1.89)	0.3	0.5	-	-
BRCA2	4 (0.9%)	NA	NA	NA	-	-
CDK4	25 (5.9%)	0.81 (0.30-2.21)	0.7	0.9	-	-
CDKN2A	24 (5.6%)	1.96 (0.98-3.91)	0.06	0.3	-	-
CDKN2B	8 (1.9%)	2.26 (0.71-7.17)	0.2	0.4	-	-
CTNNB1	5 (1.2%)	NA	NA	NA	-	-
EGFR	117 (27%)	0.90 (0.57-1.43)	0.7	0.9	-	-
ERBB2	20 (4.7%)	2.94 (1.47-5.88)	0.002	0.03	1.99 (0.96-4.16)	0.066
FAT1	10 (2.3%)	1.54 (0.49-4.88)	0.5	0.8	-	-
FOXA1	3 (0.7%)	NA	NA	NA	-	-
GLI1	7 (1.6%)	0.68 (0.09-4.88)	0.7	0.9	-	-
GNAS	6 (1.4%)	NA	NA	NA	-	-
KEAP1	17 (4.0%)	1.20 (0.44-3.27)	0.7	0.9	-	-
KIT	3 (0.7%)	NA	NA	NA	-	-
KRAS	169 (40%)	0.85 (0.55-1.30)	0.4	0.7	-	-
MDM2	30 (7.0%)	1.55 (0.75-3.21)	0.2	0.4	-	-
MED12	4 (0.9%)	NA	NA	NA	-	-
MET	11 (2.6%)	1.41 (0.45-4.46)	0.6	0.9	-	-
MGA	13 (3.1%)	1.44 (0.53-3.93)	0.5	0.8	-	-
MYC	12 (2.8%)	1.86 (0.68-5.08)	0.2	0.4	-	-
NF1	20 (4.7%)	1.83 (0.80-4.19)	0.2	0.4	-	-
NF2	3 (0.7%)	NA	NA	NA	-	-

NKX2-1	9 (2.1%)	2.70 (0.99-7.37)	0.05	0.3	-	-
NTRK1	4 (0.9%)	NA	NA	NA	-	-
PIK3CA	17 (4.0%)	0.96 (0.30-3.04)	0.9	0.9	-	-
PIK3R1	5 (1.2%)	NA	NA	NA	-	-
PTPRD	5 (1.2%)	NA	NA	NA	-	-
PTPRT	10 (2.3%)	2.26 (0.83-6.18)	0.1	0.4	-	-
RB1	12 (2.8%)	2.00 (0.80-4.97)	0.1	0.4	-	-
RBM10	52 (12%)	0.35 (0.13-0.94)	0.04	0.2	-	-
RET	8 (1.9%)	**	**	**	-	-
ROS1	4 (0.9%)	NA	NA	NA	-	-
SETD2	12 (2.8%)	0.80 (0.20-3.26)	0.8	0.9	-	-
SMAD4	8 (1.9%)	3.74 (1.37-10.24)	0.01	0.08	-	-
SMARCA4	12 (2.8%)	3.57 (1.55-8.19)	0.003	0.03	2.44 (1.03-5.77)	0.042
STK11	71 (17%)	0.95 (0.54-1.65)	0.8	0.9	-	-
TERT	25 (5.9%)	1.10 (0.47-2.56)	0.8	0.9	-	-
TP53	138 (32%)	2.32 (1.53-3.51)	<0.001	0.003	1.73 (1.09-2.73)	0.019
U2AF1	9 (2.1%)	0.38 (0.05-2.76)	0.3	0.5	-	-

TMB, normalized tumor mutation burden; FGA, fraction of genome altered; WGD, whole genome doubling; CI, confidence interval; HR, hazard ratio; FDR-P, False discovery rate corrected *P*-value; NA, not applicable; variable was not analyzed due to low incidence of alterations among patients with the event of interest.

*All genes with alteration frequency >1% in the MSK-IMPACT Panel are listed; however, to ensure convergence, only genes with more than 5 cases were assessed.

**Model did not converge.

eTable 5. Types of Alteration and Recurrence Rates for Patients With Alterations in Genes Associated With Relapse-Free Survival

Gene	Total Patients with Alteration No. (%)	Total Number of Alterations in Cohort (No. Mutations/CNA-type)	Total Recurrences No. (%)
SMARCA4	12/426 (2.8%)	16 Alterations (15 Mutations-11 Truncating*, 4 Missense) (1 Fusion) (0 CNA)	6/12 (50%)
TP53	138/426 (32%)	157 Alterations (155 Mutations-59 Truncating*, 92 Missense) (2 Fusions) (0 CNA)	36/138 (26%)

Truncating mutations are composed of nonsense, frameshift, and splice-site mutations. CNA, copy number alteration.

eTable 6. Number of Level 1 Actionable Alterations

Gene	Total Patients with Alteration, No. (% of Total Cohort)	Total Actionable Alterations, No. (% of Total Altered)
EGFR	117/426 (27%)	106/117 (91%)
BRAF	20/426 (4.7%)	3/20 (15%)
ALK	5/426 (1.2%)	5/5 (100%)
ROS1	4/426 (0.93%)	4/4 (100%)

eTable 7. Comparison of Proportion of Patients With Pathologic Stage I Cancer Who Recurred by TNM Risk Group

TNM Risk Group	N (%) of Total Stage I Patients	N (%) of Stage I Patients who Recurred
Low-risk (T1a-b and 0-1 aggressive factors*)	189/318 (59%)	10/28 (36%)
Intermediate-risk (T1a-b and 2-3 aggressive factors)	70/318 (22%)	8/28 (29%)
High-risk (T1c/T2a and any aggressive factors)	59/318 (19%)	10/28 (36%)

*As defined by NCCN guidelines[1]: lymphovascular invasion, visceral pleural invasion, and sublobar resection

1. National Comprehensive Cancer Network. Non-Small Cell Lung Cancer (Version 4.2020). Accessed May 28, 2020.

eTable 8. Predicted Relapse-Free Survival at 1, 2, and 3 Years by PRecur Risk Group for the MSK and TCGA Data Sets

	Predicted Relapse-Free Survival, 95%CI		
MSK Risk Group	*1-Year*	*2-Year*	*3-Year*
Low	0.98 (0.96-0.99)	0.96 (0.93-0.99)	0.93 (0.89-0.98)
Intermediate	0.90 (0.86-0.94)	0.81 (0.75-0.88)	0.72 (0.64-0.81)
High	0.70 (0.61-0.81)	0.49 (0.39-0.63)	0.33 (0.22-0.50)
TCGA Risk Group	*1-Year*	*2-Year*	*3-Year*
Low	0.90 (0.85-0.96)	0.74 (0.62-0.89)	0.62 (0.46-0.82)
Intermediate	0.88 (0.83-0.93)	0.70 (0.64-0.77)	0.57 (0.49-0.66)
High	0.83 (0.78-0.89)	0.58 (0.50-0.68)	0.42 (0.33-0.53)

eReferences.

1. Cheng DT, Mitchell TN, Zehir A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J Mol Diagn.* 2015;17(3):251-264.
2. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med.* 2017;23(6):703-713.
3. Sanchez-Vega F, Mina M, Armenia J, et al. Oncogenic signaling pathways in The Cancer Genome Atlas. *Cell.* 2018;173(2):321-337 e310.
4. Jordan EJ, Kim HR, Arcila ME, et al. Prospective comprehensive molecular characterization of lung adenocarcinomas for efficient patient matching to approved and emerging therapies. *Cancer Discov.* 2017;7(6):596-609.
5. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials.* 1996;17(4):343-346.
6. Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Soft.* 2011;45.
7. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol.* 2006;59(10):1092-1101.
8. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley and Sons; 1987.
9. Shen R, Martin A, Ni A, et al. Harnessing clinical sequencing data for survival stratification of patients with metastatic lung adenocarcinomas. *JCO Precis Oncol.* 2019;3.
10. Lausen B, Schumacher M. Maximally Selected Rank Statistics. *Biometrics.* 1992;48(1):73-85.
11. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the Performance of Prediction Models A Framework for Traditional and Novel Measures. *Epidemiology.* 2010;21(1):128-138.
12. Heller G, Mo QX. Estimating the concordance probability in a survival analysis with a discrete number of risk groups. *Lifetime Data Anal.* 2016;22(2):263-279.
13. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA.* 1982;247(18):2543-2546.
14. Gonen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika.* 2005;92(4):965-970.
15. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med.* 2002;8(8):816-824.
16. Blackstone EH. The Hazard Package (Cleveland Clinic Foundation). https://www.lerner.ccf.org/qhs/software/hazard/. Accessed Jan 3, 2020.
17. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell.* 2012;150(6):1107-1120.
18. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med.* 2017;376(22):2109-2121.