

# Transcriptomic Analysis Reveals Prognostic Molecular Signatures of Stage I Melanoma

Rohit Thakur<sup>1,2</sup>, Jonathan P. Laye<sup>1</sup>, Martin Lauss<sup>3</sup>, Joey Mark S. Diaz<sup>1</sup>, Sally Jane O'Shea<sup>4,5</sup>, Joanna Poźniak<sup>1,6,7</sup>, Anastasia Fila<sup>8</sup>, Mark Harland<sup>1</sup>, Joanne Gascoyne<sup>1</sup>, Juliette A. Randerson-Moor<sup>1</sup>, May Chan<sup>1</sup>, Tracey Mell<sup>1</sup>, Göran Jönsson<sup>3</sup>, D. Timothy Bishop<sup>1</sup>, Julia Newton-Bishop<sup>1</sup>, Jennifer H. Barrett<sup>1</sup>, and Jérémie Nsengimana<sup>1</sup>



## Abstract

**Purpose:** Previously identified transcriptomic signatures have been based on primary and metastatic melanomas with relatively few American Joint Committee on Cancer (AJCC) stage I tumors, given difficulties in sampling small tumors. The advent of adjuvant therapies has highlighted the need for better prognostic and predictive biomarkers, especially for AJCC stage I and stage II disease.

**Experimental Design:** A total of 687 primary melanoma transcriptomes were generated from the Leeds Melanoma Cohort (LMC). The prognostic value of existing signatures across all the AJCC stages was tested. Unsupervised clustering was performed, and the prognostic value of the resultant signature was compared with that of sentinel node biopsy (SNB) and tested as a biomarker in three published immunotherapy datasets.

**Results:** Previous Lund and The Cancer Genome Atlas signatures predicted outcome in the LMC dataset ( $P = 10^{-8}$

to  $10^{-4}$ ) but showed a significant interaction with AJCC stage ( $P = 0.04$ ) and did not predict outcome in stage I tumors ( $P = 0.3-0.7$ ). Consensus-based classification of the LMC dataset identified six classes that predicted outcome, notably in stage I disease. LMC class was a similar indicator of prognosis when compared with SNB, and it added prognostic value to the genes reported by Gerami and colleagues. One particular LMC class consistently predicted poor outcome in patients receiving immunotherapy in two of three tested datasets. Biological characterization of this class revealed high *JUN* and *AXL* expression and evidence of epithelial-to-mesenchymal transition.

**Conclusions:** A transcriptomic signature of primary melanoma was identified with prognostic value, including in stage I melanoma and in patients undergoing immunotherapy.

## Introduction

Cutaneous melanoma continues to increase in incidence worldwide. Although earlier diagnosis has been documented with correspondingly better outcomes, the rising incidence of thinner tumors means that, counterintuitively, one-fifth of

deaths now occur in patients presenting initially with early disease (1). In the United Kingdom, 91% of melanomas are diagnosed at American Joint Committee on Cancer (AJCC) stages I to II (2). Therefore, better prognostic biomarkers are needed to identify early-stage disease requiring adjuvant therapies, as well as predictive biomarkers of response to checkpoint blockade.

Previous transcriptomic analyses of cutaneous melanoma have generated gene signatures with a prognostic value independent of AJCC stage (3–7). The prognostic signature developed by Jonsson and colleagues (3) classifies metastatic melanomas into four classes (Lund 4 classes), later simplified into two classes (Lund 2 grades; ref. 4), and the signature developed by The Cancer Genome Atlas (TCGA) consortium classified melanomas into three classes (TCGA 3 classes; ref. 8). The prognostic significance of the Lund 4 class and TCGA 3 class signatures have been replicated in relatively small datasets, notably with few AJCC stage I patients (5, 9). Another transcriptomic signature based on 27 genes was developed by Gerami and colleagues (6) to classify patients with primary melanoma into tumors that were high or low risk for metastasis.

In this study, the first aim was to test the prognostic value of the Lund and TCGA signatures, as well as the gene list of signature of Gerami and colleagues (6) in a large population-based cohort of primary melanomas with a good proportion of stage I patients and extensive phenotypic annotations (Leeds Melanoma Cohort, LMC). Because the dataset was well

<sup>1</sup>University of Leeds School of Medicine, Leeds, United Kingdom. <sup>2</sup>Department of Surgical Oncology, MD Anderson Cancer Center, Houston, Texas. <sup>3</sup>Division of Oncology and Pathology, Department of Clinical Sciences Lund, Faculty of Medicine, Lund University, Lund, Sweden. <sup>4</sup>Mater Private Hospital Cork, Mahon, Cork, Ireland. <sup>5</sup>School of Medicine, University College Cork, Cork, Ireland. <sup>6</sup>Laboratory for Molecular Cancer Biology, VIB Center for Cancer Biology, KU Leuven, Leuven, Belgium. <sup>7</sup>Department of Oncology, KU Leuven, Leuven, Belgium. <sup>8</sup>Centre for Translational Research, Biomedical Research Foundation of the Academy of Athens (BRFAA), Athens, Greece.

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

J. Newton-Bishop, J.H. Barrett, and J. Nsengimana contributed equally to this article.

**Corresponding Author:** Jérémie Nsengimana, Saint James University Hospital, Clinical Sciences Building, Room 6.6, Beckett Street, Leeds LS9 7TF, United Kingdom. Phone: 44-11-3206-8974; E-mail: J.Nsengimana@leeds.ac.uk

Clin Cancer Res 2019;25:7424–35

doi: 10.1158/1078-0432.CCR-18-3659

©2019 American Association for Cancer Research.

### Translational Relevance

The introduction of adjuvant but toxic therapies for primary melanoma has highlighted the need to stratify patients based on improved prognostic and predictive biomarkers. We report a six-class transcriptomic signature generated from primary melanomas that predicted prognosis, notably in stage I disease. The signature demonstrated comparable prognostic value to that of sentinel node biopsy. When the six classes were applied to published transcriptomic datasets from patients treated with immunotherapy, one class consistently predicted poor outcome. This class was characterized by expression of *JUN* and *AXL*, both known determinants of poor therapeutic response in advanced melanoma. These findings suggest that the six-class signature should be applied to larger datasets as they become available, in order to further validate its clinical relevance as a prognostic/predictive biomarker in the adjuvant setting.

powered for discovery of novel tumor subtypes, unsupervised clustering of the tumor transcriptomes of the LMC was performed and the prognostic value of the resultant signature was compared with that of sentinel node biopsy (SNB) in analyses stratified by AJCC stage. Finally, the association between the Leeds signature and outcome was tested in published data from patients receiving immunotherapy (10–12).

## Materials and Methods

### Leeds Melanoma Cohort

As described previously (13), 2,184 patients with primary melanoma were recruited to the LMC in the period of 2000 to 2012. This was a population-ascertained cohort, which, therefore, recruited patients treated at multiple clinical centers (recruitment rate 67%). During this period, SNB was neither offered nor accepted universally. The study was ethically approved (ethical approval MREC 1/3/57, PIAG 3-09(d)/2003) and in accordance with the Declaration of Helsinki. Participants were consented to sampling of their formalin-fixed, paraffin-embedded tumor blocks that were stored in the NHS (UK National Health Service) histopathology departments of the respective hospitals. Hematoxylin and eosin-stained slides were generated and examined to facilitate subsequent sampling of the blocks using a 0.6-mm diameter tissue microarray needle as previously reported (5, 13). Prior to sampling, all the tumor blocks were reviewed, and if there was only a small amount of tumor left in the block, then the block was not sampled, lest a clinically important block be destroyed. Up to two cores were sampled from each block, and, to increase the comparability between tumors, the samples were consistently extracted from the least inflamed, least stromal regions of the invasive front of the tumor. The tumor-infiltrating lymphocytes (TILs) were scored using the classification system of Clark and colleagues (14). As previously reported (13), 703 tumor transcriptomes were profiled and in this study, 16 samples were removed in quality control leaving a cohort of 687 patients, henceforth referred to as the whole LMC dataset. The dataset contained 251 patients who had an SNB test (Supplementary Table S10), and only 16 patients are known so far to have been treated with checkpoint blockade. The LMC patients were

assigned an AJCC stage based on the AJCC staging eighth edition (15). Where patients did not have an SNB, the AJCC staging used was derived from clinical staging and pathologic examination of the wide local excision sample.

### mRNA extraction and expression data generation

Both mRNA and DNA were extracted from the tumor samples derived from cores following a previously described protocol (5, 13). The whole genome mRNA expression profiling was carried out using Illumina's DASL-HT12-v4 array. As described previously, for quality control, the mRNA was extracted from up to two cores for a number of patients (117 duplicates in total); gene expression data from only one extraction per patient was used in subsequent analyses (13). The raw transcriptomic data were extracted from the image files using GenomeStudio (Illumina Inc.) and were preprocessed as previously reported (13). Briefly, after background correction and quantile normalization (R package LUMI; ref. 16), singular value decomposition was used to remove the batch effect [R package SWAMP (refs. 13, 17)].

### Quality control in LMC

The array included 29,262 probes corresponding to 20,715 unique genes. For genes with multiple probes, the probe detected in the largest number of tumors was retained, and two additional filters were applied: genes had to be detected with  $P < 0.05$  in at least 40% of tumors and had to have a standard deviation (SD)  $> 0.40$ . This SD threshold was chosen on the basis of the overall distribution across the 20,715 genes on the log<sub>2</sub> scale. The median SD was 0.68. The data were standardized to give each gene a mean of 0 and SD of 1.

### Procedures

The LMC tumors were classified into the Lund 4 classes, Lund 2 grades, and TCGA 3 classes using the supervised nearest centroid classification (NCC) as described previously (5). All the 27 genes of the gene signature of Gerami and colleagues (6) were present in LMC dataset and were analyzed using a univariable survival analysis in the whole LMC dataset and stage I tumors. Unsupervised clustering was performed using the consensus Partitioning Around Medoids clustering method in the R-package ConsensusClusterPlus (18, 19) with Euclidean distance as the dissimilarity measure and a resampling fraction of 0.8 for both genes and samples in 1,000 iterations (Supplementary Methods).

### Statistical analysis

Cox proportional hazard models and Kaplan–Meier curves were used to test the association with survival (R-package Survival; ref. 20). The survival time was calculated from time of diagnosis to the time of last follow-up or time of death from melanoma, whichever occurred first, referred to as melanoma-specific survival (MSS). Patients with deaths caused by factors other than melanoma were censored at the time of death. ROC analysis was performed using AJCC stage pre-SNB and AJCC stage post-SNB for patients who had SNB. Clinical staging prior to SNB is described as AJCC pre-SNB. The AJCC stage post-SNB includes additional information on regional lymph node metastasis. The analysis used AJCC staging eighth edition, and MSS up to 3 years was chosen as cutoff based upon the inclusion of the majority of the deaths without loss of data as a result of censoring (Supplementary Table S11). Patients who were censored before 3 years were

not included in the analysis. The analysis was performed using R-packages pROC, plotROC, and ggplot2 (21–23).

#### Pathway enrichment analysis

The differentially expressed genes (DEG) were identified using the Significance Analysis of Microarrays (R-package SAMR) by comparing each class versus all others (24). Pathway enrichment and biological network analysis of DEGs with a  $q$  value equal to 0 were performed using ReactomeFviz in Cytoscape (25). The central nodes of the biological network were identified using a centrality measure (betweenness) in Gephi (ref. 26; Supplementary Methods).

#### Copy-number alterations

The copy-number alterations (CNA) data were generated in a subset of LMC tumors using Illumina's next-generation sequencing (NGS) platform as reported in the study by Filia and colleagues (ref. 27; Supplementary Methods). Among the 687 transcriptome-profiled patients of LMC, CNA data were available for 272 patients. The CNA were assessed in the regions spanning the genes identified as hubs in network enrichment analysis. The ratio between mean of the window read counts in the region mapping to a gene and the average read count of the 10 flanking regions around that gene was used to estimate the copy-number changes. The windows (5k) corresponding to a gene locus were identified using the R packages biomaRt (28, 29). The cutoff for calling a region amplified was chosen as a value greater than 0.4 whereas a value less than  $-0.4$  was used to identify a deletion. The 272 samples in the CNA dataset were at AJCC stages I ( $n = 80$ ), II ( $n = 147$ ), and III ( $n = 45$ ; similar distribution to the whole LMC dataset).

#### Lund validation dataset

For replication, a primary melanoma transcriptomic dataset of 223 tumors from a Lund cohort (Sweden) was used (Harbst and colleagues; ref. 4). The samples were classified using the newly generated signature by the supervised NCC approach (5). Of those 223 patients, 200 had recorded information on melanoma relapse in the follow-up time postdiagnosis and were used to test the association between patient subgroups and relapse-free survival (using Cox proportional hazard models, Kaplan–Meier curves, and log-rank test).

#### Immunotherapy datasets

Three publicly available transcriptome datasets (Hugo Cohort: GSE78220, Ulloa-Monotoya Cohort: GSE35640, and Riaz Cohort: [https://github.com/riazn/bms038\\_analysis](https://github.com/riazn/bms038_analysis)) were downloaded (10–12) and samples were quantile normalized and classified using the NCC method (Pearson correlation coefficient). The Riaz cohort was a mixture of samples from various melanoma types (cutaneous melanoma, mucosal melanoma, acral melanoma, uveal/ocular melanoma, others). In this study, the samples labeled as cutaneous melanoma were analyzed. In all the three cohorts, the association with response to immunotherapy was tested using Fisher exact test.

## Results

### Existing signatures showed no association with survival in stage I melanoma

The structure of datasets used in this study is depicted in Fig. 1. When applied to the whole LMC dataset ( $n = 687$ ),

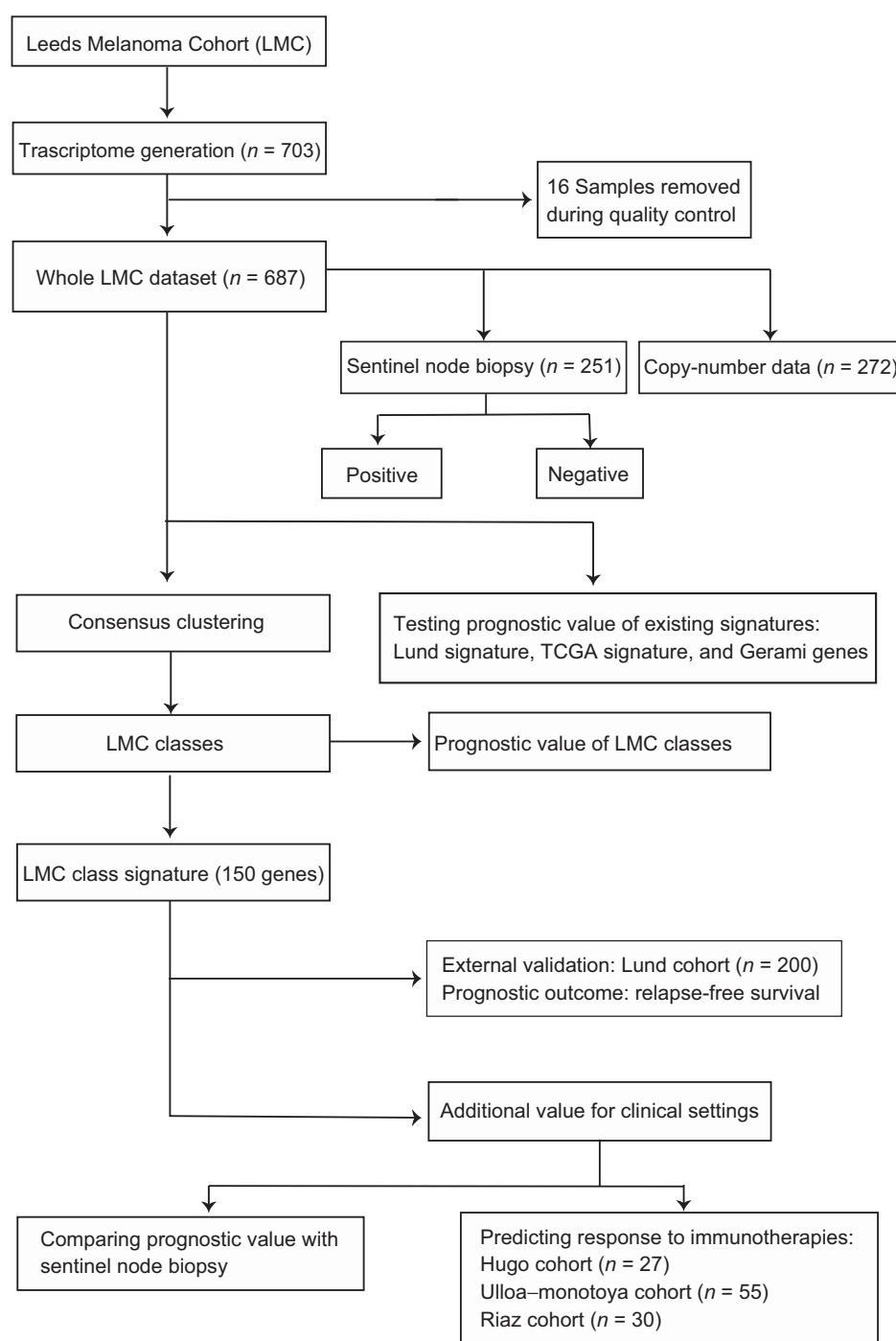
the three formerly published signatures (Lund 4 class, Lund 2 grade, and TCGA 3 class) replicated previously observed associations with MSS (Fig. 2A, C, and E). However, upon stratifying LMC patients on the basis of AJCC stage, the Lund and TCGA signatures showed no association with prognosis for LMC stage I patients (Fig. 2B, D, and F). The Lund 2-grade signature had the highest statistical power (since based on only two groups) and showed a statistically significant interaction with AJCC stage ( $P = 0.04$ , Supplementary Table S1), suggesting that the lack of association in stage I was not solely due to low sample size. Because the full details of the commercial signature of Gerami and colleagues (6) were not published, we were limited in the scope of its replication in the LMC dataset. However, analyzing the 27 Gerami genes identified 23 genes as predictors of prognosis in the whole LMC dataset (Supplementary Table S2). However, in keeping with the Lund and TCGA signatures, none of these genes showed a significant association with prognosis in stage I tumors (Supplementary Table S2).

### Generating novel LMC classes and their clinical characteristics

Consensus clustering of the LMC dataset was performed, and following additional quality control measures (Supplementary Table S3), six distinct, novel molecular classes were identified (Fig. 3A). These classes were associated with clinicopathologic variables known to have prognostic value, including tumor site ( $P = 0.03$ ), age at diagnosis ( $P = 0.03$ ), mitotic rate ( $P = 0.002$ ), ulceration ( $P = 0.01$ ), AJCC stage ( $P = 6 \times 10^{-10}$ ), TILs ( $P = 6 \times 10^{-4}$ ), and Breslow thickness ( $P = 9 \times 10^{-14}$ , Table 1). The LMC class 1 and 5 tumors tended to be thin and nonulcerated, whereas class 2 and 4 tumors were thicker. Class 3 and 6 tumors were the thickest and most frequently ulcerated. The six classes showed strong association with *BRAF* ( $P = 6 \times 10^{-5}$ ) and *NRAS* mutation status ( $P = 3 \times 10^{-4}$ ): class 1, 5, and 6 tumors were frequently *BRAF* mutated, whereas class 2, 3, and 4 tumors were frequently *NRAS* mutated (Table 1).

### LMC classes predicted prognosis in primary melanoma and in stage I subset

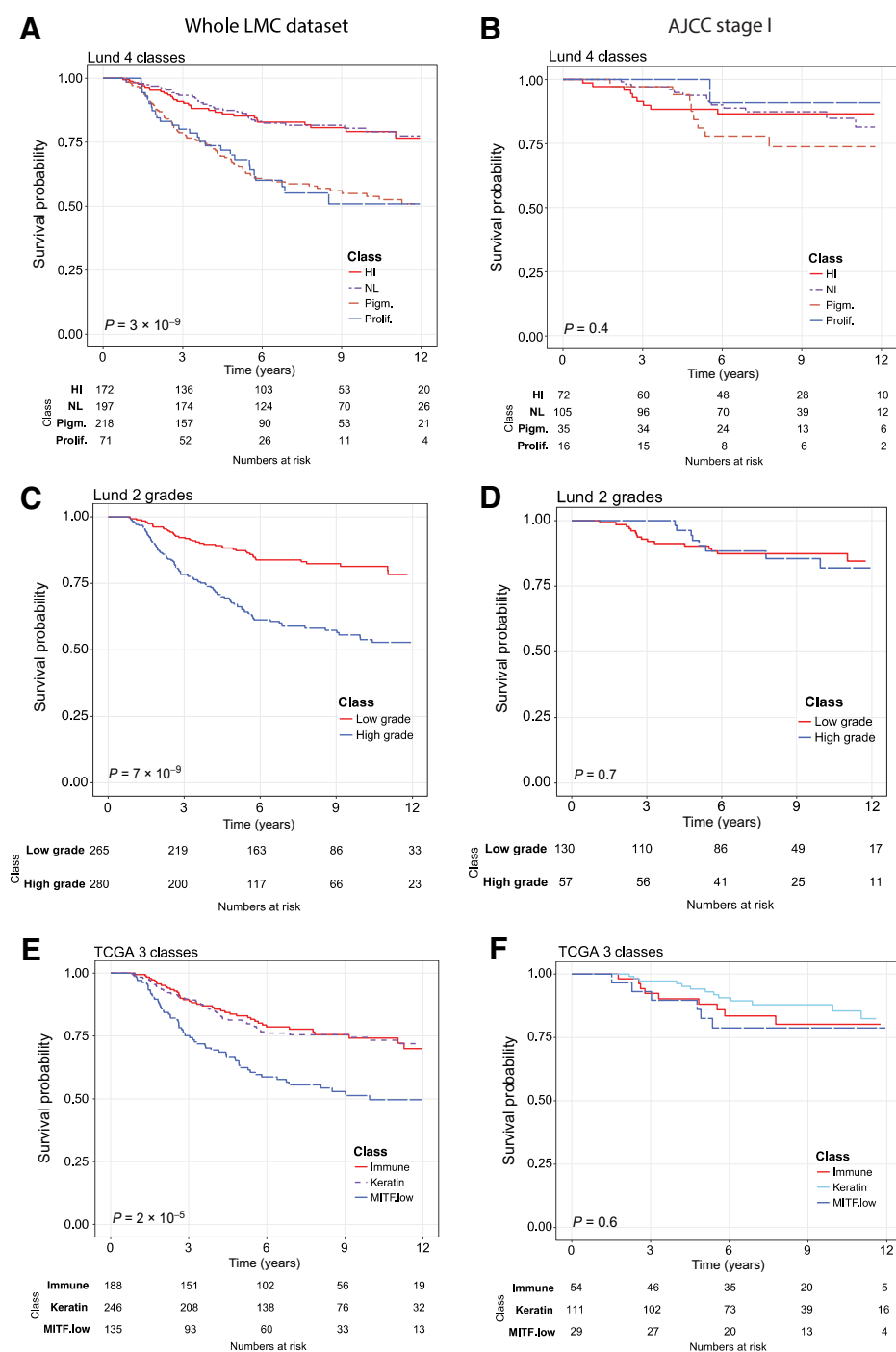
The LMC classes predicted MSS in the whole LMC dataset and notably, across AJCC stages I, II and III subsets (Fig. 3B and C, Supplementary Fig. S1). In the unadjusted analysis of the whole dataset (Fig. 3B, Supplementary Table S4), class 1 (baseline) had the best prognosis and class 2 [HR = 1.7, 95% confidence interval (CI): 0.8–3.5] and class 5 (HR = 1.5, 95% CI: 0.7–3.1) showed intermediate prognosis, whereas class 3 (HR = 5.0, 95% CI: 2.5–10.1), class 4 (HR = 2.4, 95% CI: 1.2–4.7), and class 6 (HR = 3.1, 95% CI: 1.6–6.1) had the worst prognosis. In multivariable analysis, classes 3, 4, and 6 remained significant predictors of poor prognosis after including AJCC stage, sex, age at diagnosis, mitotic rate (Supplementary Table S4), and when the AJCC stage was replaced by ulceration and Breslow thickness in the model (Supplementary Table S6). In the LMC stage I subset, class 6 (HR = 6.6, 95% CI: 1.4–31.2) significantly predicted poor prognosis in unadjusted analysis (Fig. 3C and Supplementary Table S5) and it remained significant when sex, age at diagnosis, mitotic rate, ulceration, and Breslow thickness were adjusted (HR = 9.8, 95% CI: 1.1–86.2, Supplementary Table S6). Because Gerami



**Figure 1.**  
Analysis workflow of the study.

signature was not available to us in full, we ran unsupervised clustering of the LMC dataset using the 27 Gerami genes to generate the two tumor groups analyzed by Gerami and colleagues (6), referred to as the Gerami clusters. This analysis showed that the LMC classes and Gerami clusters had independent prognostic effects in the whole LMC dataset (Supplementary Table S7); however, the Gerami clusters showed no prognostic value in stage I tumors whereas LMC class 6 remained a significant predictor in the multivariable model (Supplementary Table S8).

To validate the prognostic value of the LMC classes in an independent dataset, a 150-gene-based signature (LMC signature), generated after refining approximately 13,000 genes (Supplementary Fig. S2), was applied to the Lund dataset (4). In keeping with the observations made in the LMC dataset, class 3, class 4, and class 6 predicted worse prognosis in the Lund dataset, whereas class 1, class 2, and class 5 predicted better prognosis (Fig. 3D, Supplementary Table S9). Because the Lund dataset had only a few stage I cases ( $n = 58$ ), the prognostic value of LMC signature could not be replicated in stage I disease.

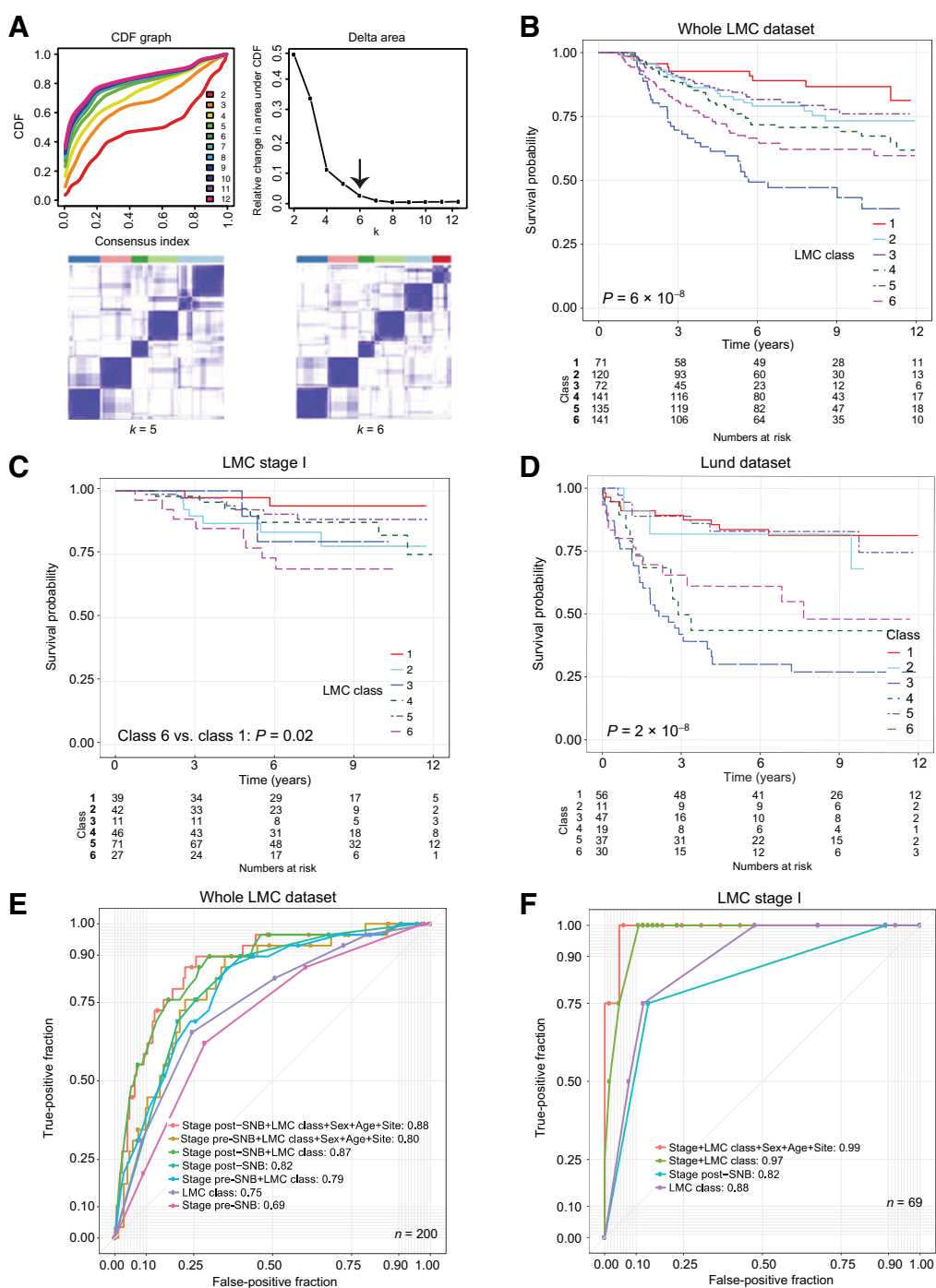


**Figure 2.** Replicating Lund and TCGA signatures using LMC dataset. Kaplan-Meier plots showing the MSS for Lund 4 classes (HI, high immune; NL, normal-like; Pigm, pigmentation; Prolif, proliferative; **A**), Lund 2 grades (low grade and high grade; **C**), and TCGA 3 classes (immune, keratin, MITF low; **E**) across the whole LMC dataset. In AJCC stage I subset, Kaplan-Meier plots showing the MSS for (B) Lund 4 classes, Lund 2 grades (D), and TCGA 3 classes (F). P values are from log-rank test. Samples that could not be classified into any of the classes were not used in survival analysis.

**LMC signature had independent prognostic value when compared with SNB**

In the dataset derived from individuals who had an SNB, the prognostic value of combined LMC class signature and pre-SNB AJCC stage was similar to that of AJCC stage with SNB (i.e. stage post-SNB; AUC 0.82 vs. 0.79,  $P = 0.7$ , Fig. 3E). Combining the LMC signature with AJCC stage post-SNB, patient's sex, age at diagnosis, and site of tumor increased the AUC to 0.88. Similarly, in the subset of patients at stage I pre-SNB, the

LMC signature alone had comparable prognostic value to AJCC stage post-SNB (AUC = 0.88 vs. 0.82,  $P = 0.1$ , Fig. 3F). In this stage I subset, addition of stage post-SNB, patient's sex, age at diagnosis, and site of tumor to the LMC signature further increased the AUC to 0.99. However, the limited sample size of stage I dataset and including so many variables clearly overfitted the model, giving near-perfect classification and illustrating that independent datasets are needed to better assess performance.



**Figure 3.** Defining LMC signature and its prognostic value. **A**, The area under the CDF and its relative change with increasing  $k$ . The delta area graph shows little variation at  $k = 6$ . Heatmap of consensus matrices at  $k = 5$  and  $6$ . The blue color indicates high consensus score and the white color indicates low consensus **(B)** Kaplan-Meier plot showing the MSS for the six classes in **(B)** the whole LMC dataset, **(C)** the LMC stage I, and **(D)** relapse-free survival in the Lund cohort ( $P$  value from log-rank test, or Wald test for two-group comparison). Seven mucosal tumors were excluded from analysis. **E**, ROC curves comparing the prognostic value of the LMC signature to that of SNB in the whole dataset. The AUCs for LMC class+ stage pre-SNB and stage post-SNB were not significantly different (DeLong test  $P = 0.7$ ). **F**, The ROC curve comparing prognostic value of LMC signature with SNB in the stage I pre-SNB group. All but 1 patient were stage IB pre-SNB; therefore, AUC for LMC signature alone was compared with stage post-SNB and the difference was not significant (DeLong test  $P = 0.7$ ). The difference in AUCs between stage post-SNB alone and LMC class + stage post-SNB was also not significant (DeLong test  $P = 0.1$ ).

Downloaded from <http://aacrjournals.org/clinccancerres/article-pdf/25/24/7424/2056500/7424.pdf> by guest on 27 August 2022

**Table 1.** The LMC classes' association with clinicohistopathologic variables

Histopathologic variables	Whole dataset <i>n</i> = 687 (%)	LMC classes						<i>P</i> <sup>a</sup>
		Class 1 ( <i>n</i> = 71)	Class 2 ( <i>n</i> = 122)	Class 3 ( <i>n</i> = 73)	Class 4 ( <i>n</i> = 143)	Class 5 ( <i>n</i> = 136)	Class 6 ( <i>n</i> = 142)	
Sex: male, <i>n</i> (%)	310 (45)	39 (55)	51 (42)	34 (47)	56 (39)	55 (40)	75 (52)	0.07
Tumor site: limbs, <i>n</i> (%)	289 (42)	37 (52)	58 (48)	26 (36)	58 (41)	64 (47)	46 (32)	0.03
Age at diagnosis (years), median (range)	58 (18–81)	59 (21–76)	59 (22–79)	60 (20–77)	58 (18–81)	53 (25–76)	59 (22–81)	0.03
Breslow thickness (mm), median (range)	2.3 (0.3–20)	1.7 (0.7–5.5)	2.1 (0.8–8.9)	3.2 (0.8–20)	2.3 (0.3–15)	1.8 (0.7–12)	3.0 (0.8–18)	$9 \times 10^{-14}$
AJCC stage, <i>n</i> (%) <sup>b</sup>								$6 \times 10^{-10}$
I	236 (35)	39 (55)	42 (35)	11 (15)	46 (33)	71 (53)	27 (19)	
II	335 (49)	26 (36)	57 (48)	46 (64)	77 (55)	45 (33)	84 (60)	
III	109 (16)	6 (9)	21 (17)	15 (21)	18 (12)	19 (14)	30 (21)	
Ulceration (present), <i>n</i> (%)	228 (33)	16 (23)	32 (26)	30 (41)	53 (37)	38 (28)	59 (42)	0.01
Mitotic rate (number of mitoses per mm <sup>2</sup> )	1 (0, 25)	0 (0, 11)	1 (0, 17)	2 (0, 25)	1 (0, 13)	1 (0, 12)	1 (0, 18)	0.002
TILs, <i>n</i> (%)								$6 \times 10^{-4}$
Absent	76 (15)	2 (4)	13 (14)	17 (32)	14 (16)	15 (16)	15 (13)	
Nonbrisk	333 (68)	30 (60)	65 (71)	32 (60)	60 (68)	63 (66)	83 (74)	
Brisk	81 (17)	18 (36)	14 (15)	4 (8)	14 (16)	17 (18)	14 (13)	
<i>BRAF</i> mutant, yes, <i>n</i> (%)	266 (47)	26 (43)	38 (30)	23 (40)	44 (36)	63 (59)	72 (61)	$6 \times 10^{-5}$
<i>NRAS</i> mutant, yes, <i>n</i> (%)	138 (25)	8 (14)	35 (34)	17 (30)	41 (34)	20 (19)	17 (15)	$3 \times 10^{-4}$

<sup>a</sup>The associations were tested using Pearson  $\chi^2$  test for categorical variables and the Kruskal-Wallis test for continuous variables. Symbol *n* is the number of samples.

<sup>b</sup>Seven patients had mucosal melanoma and, although they were classified, they were not included in survival analyses. Their AJCC stage was not reported. Each of LMC class 2 and class 4 contained two of these, whereas classes 3, 5, and 6 had one each.

### Biological overlap between the LMC and existing signatures

The six classes of LMC signature showed distinct gene expression profiles (Fig. 4A) and showed partial overlap with the existing Lund and TCGA signatures. LMC classes 1, 3, and 5 overlapped substantially with the high-immune, pigmentation, and normal-like classes of the Lund 4 classes (Fig. 4B), and with the immune, MITF low, and keratin classes of the TCGA 3-classes (Fig. 4C). In contrast, LMC classes 2, 4, and 6 represented a mixture of the Lund 4 classes and TCGA 3 classes. Gene expression pathway enrichment analysis revealed distinctive biological features of the 6 LMC classes: notably class 2 was characterized by increased WNT signaling genes and metabolic pathways; class 4 by decreased expression of immune genes, and class 6 by increased expression of cell cycle and consistent downregulation of cell metabolism pathway genes (Supplementary Table S14).

When applied to the LMC 6 classes, the Lund modules (30) revealed discrimination consistent with enriched gene pathways: LMC class 1 tumors showed higher immune module activity, and class 3 tumors showed higher cell-cycle module activity (Fig. 4D). Interestingly, class 6 tumors had relatively higher cell cycle but also immune module activity and, as expected, the immune, stroma, and interferon modules were positively correlated but they negatively correlated with cell-cycle and MITF modules (Fig. 4D). The tumor-infiltrating immune cell populations imputed for each of the LMC classes (31) were consistent with the Lund immune module, as class 1 had the highest immune cell populations and class 3 the lowest, whereas class 6 appeared to maintain an intermediate level of immune cell populations, having the second highest scores on average (Supplementary Fig. S3).

A comparison with the Consensus Immunome Clusters (CICs), previously generated in the same LMC dataset based on 380 immune genes (13), showed that the two most prognostically contrasted LMC classes (class 1 and class 3) had a near-perfect match with CIC 2 (high immune) and CIC 3 (low immune/ $\beta$ -catenin high), respectively (Supplementary Fig. S4), whereas the rest of LMC classes were a mixture of CICs. Cluster 1 had correspondingly a higher proportion of tumors with histologic evidence of brisk TILs (36% compared with 8% in class 3). Analyzing the correlation between the Gerami genes and LMC

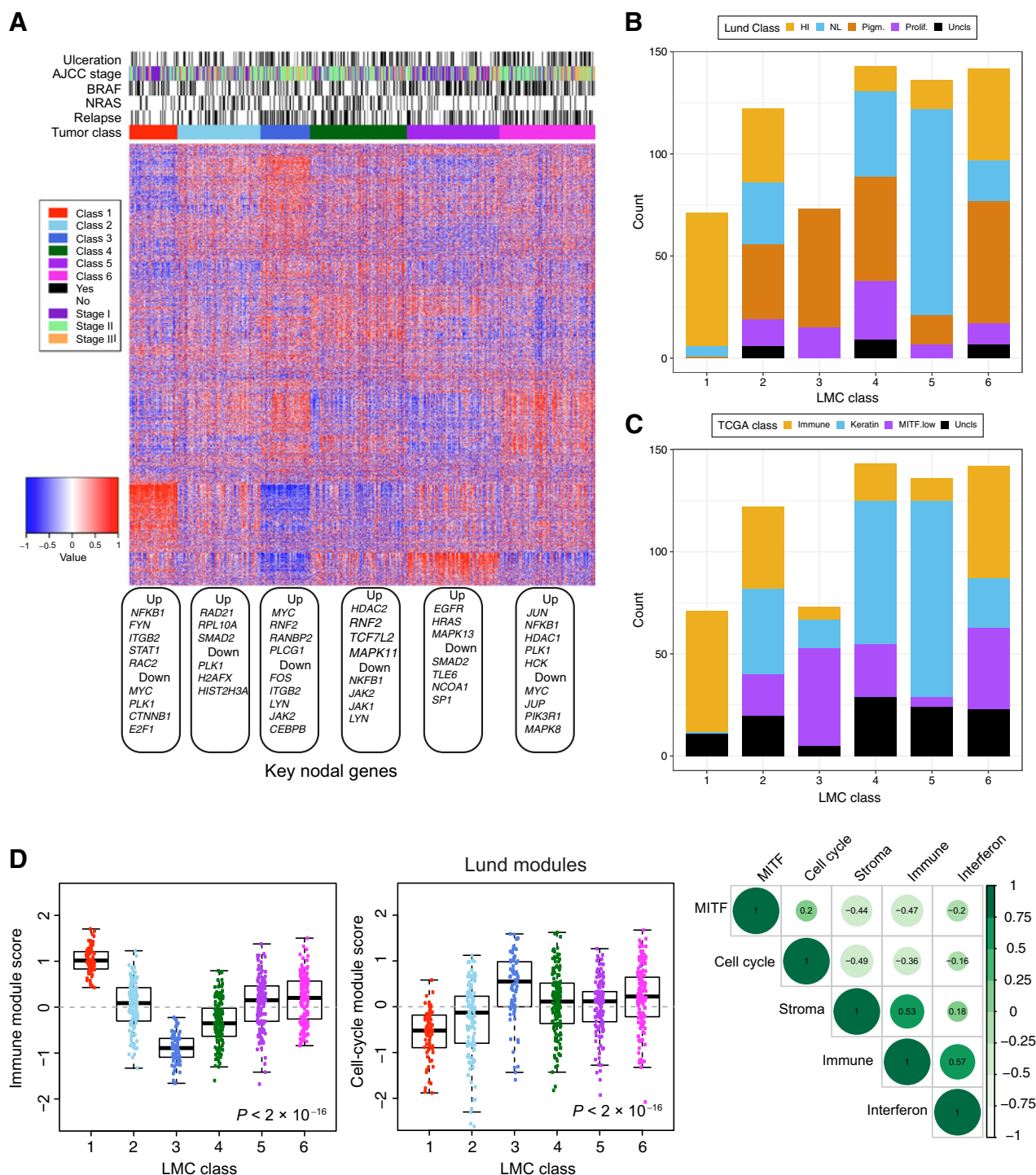
signature genes showed that the Gerami genes positively correlated with the genes upregulated in LMC class 5 tumors and negatively correlated with genes upregulated in LMC class 3 tumors (Supplementary Fig. S5). Consistent with this, Gerami clusters 1 and 2 highly overlapped with LMC classes 3 and 5, respectively (Supplementary Fig. S6).

### *JUN* as marker of poor prognosis in class 6 tumors

LMC class 6 predicted worse prognosis within AJCC stage I tumors. Further biological network analysis identified *JUN* as a key upregulated nodal gene in this class (Fig. 5A and B). The NGS-based CNA data from a subset of LMC tumors (*n* = 272) indicated that class 6 tumors were more likely to have DNA amplifications of *JUN* than other classes (*P* = 0.003, Fig. 5C, Supplementary Fig. S7). In melanoma, *JUN* has been reported to activate epithelial-to-mesenchymal transition (EMT), and accordingly a 6-gene-based (32) and 200-gene-based EMT signature (33) consistently scored higher in LMC class 6 than in all other LMC classes (Fig. 5D, Supplementary Fig. S7). A secondary key nodal gene *NFKB1* identified to be upregulated in class 6 had no copy number changes. Further examination of immunohistochemically stained sections showed that all four tumors stained from class 6 were positive for *NFKB1* protein expression, and this was similar to other LMC classes (*P* = 0.4, Supplementary Fig. S7).

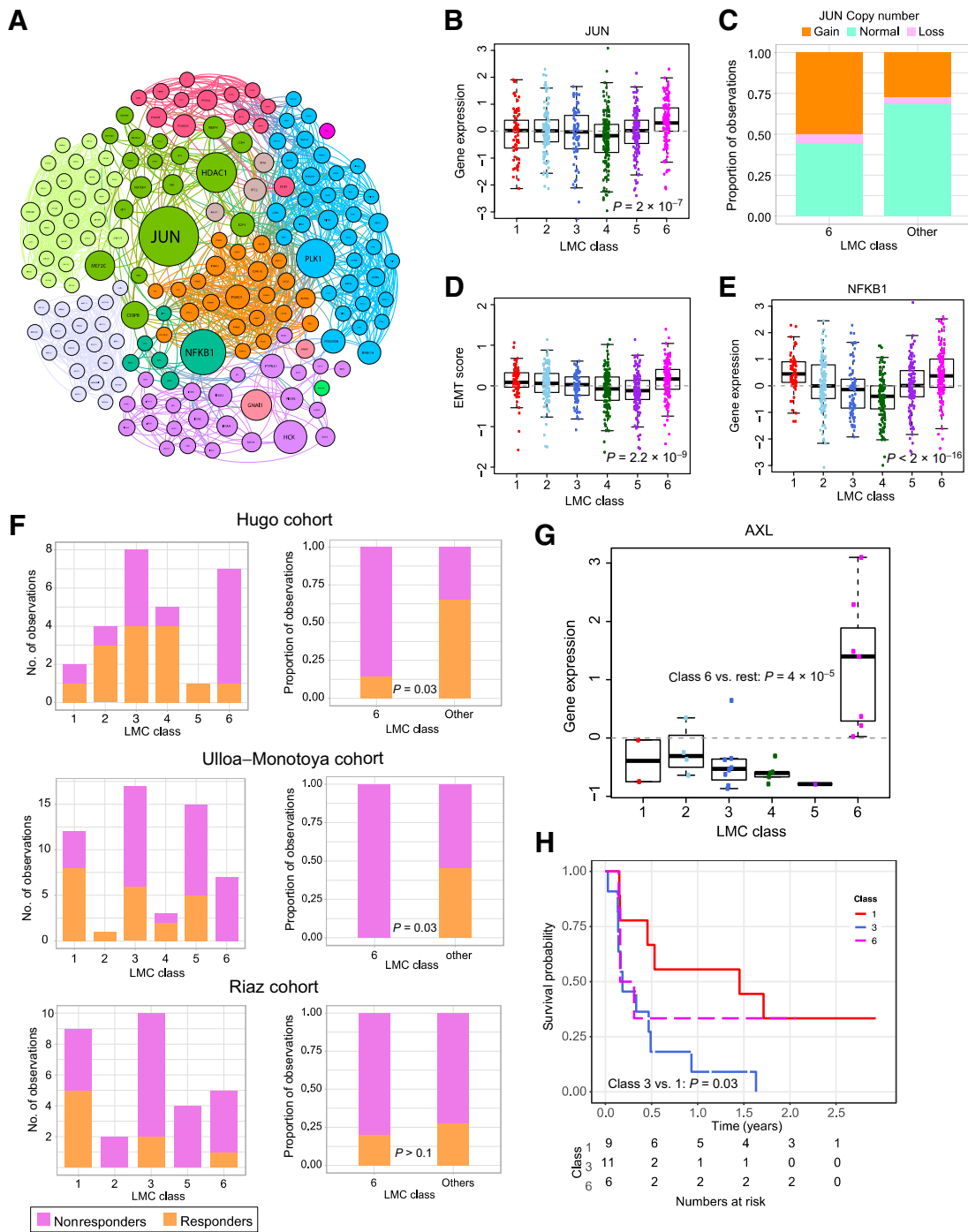
### LMC signature as a potential predictor of response to immunotherapy

The value of the LMC signature in predicting outcome in patients treated with immunotherapy was assessed in three disparate clinical trial cohorts of metastatic melanoma (Fig. 5F; refs. 10–12). In the cohort of Hugo and colleagues, tumors classified as class 6 were mainly nonresponders to PD-1 blockade in comparison to the other LMC classes (*P* = 0.03). Hugo and colleagues reported that expression of *AXL* predicts poor response to PD-1 blockade; the gene expression data revealed significantly higher *AXL* expression in class 6 tumors when compared with other classes within their cohort (Fig. 5G). Similarly, for the cohort reported by Ulloa-Montoya and colleagues, class 6 tumors showed a significantly higher proportion of nonresponders to



**Figure 4.** Biological characterization of the six LMC classes. **A**, The heatmap shows gene expression across the classes with tumor samples placed in columns and genes in rows. Blue depicts low expression and red depicts high expression. Each gene expression was standardized to mean 0 and standard deviation 1. The up- and downregulated nodal genes identified in network analyses are shown under the heatmap. The bar plot shows the overlap between the LMC classes and Lund 4 classes (HI, high immune; NL, normal-like; Pigm, pigmentation; Prolif, proliferative; **B**), and TCGA 3 classes (**C**). The samples that could not be classified into the Lund 4 classes and TCGA 3 classes were labeled here as *Uncls*. **D**, The modules (defined by a list of differentially upregulated genes) associated with melanoma-specific biological pathways as identified by the Lund group (30). Box plots of immune and cell-cycle module scores (standardized expressions) within the six LMC classes and correlation matrix of immune, cell-cycle, MITF, stroma, and interferon module scores. The module score variation across the classes was tested using the Kruskal-Wallis test.





**Figure 5.** Biological characterization of LMC class 6 and association with response to immunotherapy. **A**, Network of upregulated genes in the LMC class 6 with key genes (highest betweenness centrality) shown as large circles. Subnetworks are shown in different colors. **B**, Expression of *JUN* across the six LMC classes (*P* value from Kruskal-Wallis test). **C**, *JUN* copy-number alterations in LMC class 6 versus other classes. **D**, The six-gene-based EMT score in tumors across the six LMC classes (*P* value from Kruskal-Wallis test). **E**, The gene expression of *NFKB1* across the six LMC classes (*P* value from Kruskal-Wallis test). **F**, The LMC classes association with response to immunotherapy in three cohorts (*P* value from Fisher exact test). Patients in these cohorts were classified into the six LMC classes by the NCC method. **G**, Expression of *AXL* across the six LMC classes in the Hugo Cohort dataset (*P* value from Mann-Whitney *U* test). **H**, Kaplan-Meier plot showing survival curves of LMC class 1, class 3, and class 6 in the Riaz Cohort. Other LMC classes had less than five samples and were excluded. No., number.

Downloaded from <http://aacrjournals.org/clinccancerres/article-pdf/25/24/7424/2056500/7424.pdf> by guest on 27 August 2022

MAGE-A3 immunotherapy in comparison to other classes. The cohort reported by Riaz and colleagues was predominantly composed of nonresponders to anti-CTLA-4 further treated with PD-1 blockade but LMC classes were not convincingly predictive but class 3 predicted poor prognosis, which was consistent with the LMC dataset when compared with good prognosis class 1 (Fig. 5H).

## Discussion

In this study, transcriptome classification was performed utilizing a large population-ascertained cohort of primary melanomas, revealing classes having prognostic value in stage I disease. In stage I tumors, the LMC signature predicted outcome comparably to AJCC staging including SNB. Furthermore, evidence suggests that the signature predicted outcome in patients treated with immunotherapies.

Given the rising incidence of early-stage tumors and the cost of adjuvant therapies to health services and to patients in terms of toxicity, there is an urgent need to identify better prognostic and predictive biomarkers for early-stage disease. When previous gene signatures were applied to the LMC (3, 8), the signatures robustly predicted outcome when the dataset was analyzed as a whole but failed to do so in stage I tumors alone. Although the full Gerami signature was not available, analyzing the prognostic value of genes reported in that study (6) showed that the genes were predictive of prognosis in the whole LMC dataset but not in stage I tumors. In this work, a six-class signature (Supplementary data file) was identified, which was prognostic not only in the whole LMC dataset but also in patients diagnosed at AJCC stage I. The prognostic value of the LMC signature was validated in an independent cohort of primary melanoma built in Lund (4) although the number of stage I cases in this cohort was insufficient to allow replication of the signature's prognostic value in stage I disease.

The LMC signature showed limited overlap with the Lund and TCGA signatures. When comparing it with previously identified immunome clusters by our group (13), two LMC classes strongly overlapped with immune subgroups. The nonoverlapping classes could not be clearly discriminated using the immunome clusters suggesting that these LMC classes are driven by different genomic mechanisms. Comparison of LMC signature genes with Gerami genes indicated a biological pathway overlap as Gerami genes were found to be strongly correlated with LMC classes 3 and 5.

Although SNB is an important melanoma staging tool, the surgery is associated with morbidity (34, 35). In the LMC, SNB was observed to be of prognostic value in the whole dataset and in stage I tumors. However, the LMC signature performed just as well. Given the morbidity of SNB, it may be argued that the LMC signature should be tested in an independent study as a possible alternative to this procedure especially in stage I disease where the likelihood of a positive result is overall low and must be weighed against morbidity.

In melanoma, increased immune gene expression has been consistently shown to predict good prognosis (5, 9, 13, 36). However, a subset of tumors (LMC class 6) was observed which, despite showing immune gene expression, resulted in the patient's early death. Further biological characterization of this class identified copy-number amplifications and increased expression of *JUN*. Ramsdale and colleagues (37) have shown that *JUN* promotes an invasive cell phenotype through activation

of the EMT pathway, and a higher scoring EMT signature in LMC class 6 confirmed increased activity of the EMT pathway in this class. Riesenberger and colleagues (38) have reported that increased *JUN* expression leads to proinflammatory and stress signals that promote cytokine expression in coordination with NF- $\kappa$ B. Again, these findings are consistent with the presented transcriptomic observations of *JUN* and *NFKB1* in defining LMC class 6 (Fig. 5B and E). There was insufficient tissue to carry out immunohistochemistry for *JUN*; therefore, *JUN* protein expression in the TCGA dataset was examined and confirmed a positive correlation between *JUN* gene transcription and protein expression (Supplementary Fig. S7). Collectively, these data are indicative of copy-number gains resulting in both increased gene expression and transcriptional activity of *JUN* in LMC class 6 tumors, although further proteomic studies would be required to confirm this.

The LMC signature was associated with response to immunotherapies; specifically, class 6 associated with poor outcome in two of the three tested datasets. None of these datasets are sufficiently large to make clear inferences. It is of note that the expression of *AXL*, a known marker for immune evasion, was significantly upregulated in LMC class 6 in metastatic melanoma samples in the Hugo dataset.

The inherent strength of this study is the relatively large size of the population-ascertained cohort. A corresponding limitation is the lack of a well powered AJCC stage I dataset to allow independent replication of the signature in stage I melanoma. Another limitation of this study is that only one-third of LMC patients had an SNB, limiting the power to compare staging tests. The LMC recruitment period preceded the advent of both immunotherapy and targeted therapy, and only a very small number of the study participants have been treated with these drugs. Excluding the samples from these participants showed no modifying effect of such treatments on MSS in the LMC dataset (data not shown).

In conclusion, this study presents a novel signature with demonstrated prognostic value similar in magnitude to that of AJCC staging of melanoma but having added value in stage I melanoma. The data further confirm that AJCC stage largely captures biological variation associated with survival. The LMC class signature prognostic value was similar to that of SNB in the whole dataset (where their effects were additive) and in stage I disease. The signature predicted poor outcome in patients receiving immunotherapies and, in particular, identified high-*JUN*/high-*AXL* as a tumor phenotype with poor prognosis in early- and advanced-stage melanoma albeit in very small datasets. This signature has the potential to be trialed as a biomarker in clinical monitoring programs and may help in early identification of patients who may or may not benefit from adjuvant therapies.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Accession Number

The accession number for the microarray data reported in this article is European Genome Archive accession number: EGAS00001002922.

## Authors' Contributions

**Conception and design:** R. Thakur, J. N.-Bishop, J.H. Barrett, J. Nsengimana  
**Development of methodology:** R. Thakur, D.T. Bishop, J. N.-Bishop, J.H. Barrett  
**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** J.P. Laye, S.J. O'Shea, A. Folia, M. Harland, J. Gascoyne, M. Chan, D.T. Bishop, J. N.-Bishop, J. Nsengimana

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** R. Thakur, M. Lauss, J.M.S. Diaz, J. Poźniak, G. Jönsson, D.T. Bishop, J. N.-Bishop, J.H. Barrett

**Writing, review, and/or revision of the manuscript:** R. Thakur, J.P. Laye, M. Lauss, J.M.S. Diaz, S.J. O'Shea, J. Poźniak, M. Harland, J.A. Randerson-Moor, G. Jönsson, D.T. Bishop, J. N.-Bishop, J.H. Barrett, J. Nsengimana

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** J.P. Laye, M. Harland, J. Gascoyne, J.A. Randerson-Moor, M. Chan, T. Mell, J. N.-Bishop

**Study supervision:** J.P. Laye, M. Lauss, J. N.-Bishop, J.H. Barrett, J. Nsengimana

## Acknowledgments

We thank all the participants of the LMC study and the research nurses who conducted the recruitment.

This work was funded by Cancer Research UK C588/A19167, C8216/A6129, and C588/A10721, and NIH CA83115. R. Thakur, J.M.S. Diaz, and J. Poźniak are supported by Horizon 2020 Research and Innovation Programme no. 641458 (MELGEN). Copy-number data were generated using AICR grant 12-0023.

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

Received November 8, 2018; revised February 4, 2019; accepted August 29, 2019; published first September 12, 2019.

## References

- Lo SN, Scolyer RA, Thompson JF. Long-term survival of patients with thin (T1) cutaneous melanomas: a Breslow thickness cut point of 0.8 mm separates higher-risk and lower-risk tumors. *Ann Surg Oncol* 2018;25:894-902.
- Cancer Research UK. Melanoma skin cancer diagnosis and treatment statistics. Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer/diagnosis-and-treatment#ref-1>.
- Jonsson G, Busch C, Knappskog S, Geisler J, Miletic H, Ringnér M, et al. Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. *Clin Cancer Res* 2010;16:3356-67.
- Harbst K, Staaf J, Lauss M, Karlsson A, Mäsback A, Johansson I, et al. Molecular profiling reveals low- and high-grade forms of primary melanoma. *Clin Cancer Res* 2012;18:4026-36.
- Nsengimana J, Laye J, Filia A, Walker C, Jewell R, Van den Oord JJ, et al. Independent replication of a melanoma subtype gene signature and evaluation of its prognostic value and biological correlates in a population cohort. *Oncotarget* 2015;6:11683-93.
- Gerami P, Cook RW, Wilkinson J, Russell MC, Dhillon N, Amaria RN, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. *Clin Cancer Res* 2015;21:175-83.
- Ferris LK, Farberg AS, Middlebrook B, Johnson CE, Lassen N, Oelschlager KM, et al. Identification of high-risk cutaneous melanoma tumors is improved when combining the online American Joint Committee on Cancer Individualized Melanoma Patient Outcome Prediction Tool with a 31-gene expression profile-based classification. *J Am Acad Dermatol* 2017;76:818-25.e813.
- The Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* 2015;161:1681-96.
- Lauss M, Nsengimana J, Staaf J, Newton-Bishop J, Jonsson G. Consensus of melanoma gene expression subtypes converges on biological entities. *J Invest Dermatol* 2016;136:2502-5.
- Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 2016;165:35-44.
- Ulloa-Montoya F, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. *J Clin Oncol* 2013;31:2388-95.
- Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* 2017;171:934-49.e916.
- Nsengimana J, Laye J, Filia A, O'Shea S, Muralidhar S, Poźniak J, et al.  $\beta$ -Catenin-mediated immune evasion pathway frequently operates in primary cutaneous melanomas. *J Clin Invest* 2018;128:2048-63.
- Clark WH Jr, Elder DE, Guerry D IV, Braitman LE, Trock BJ, Schultz D, et al. Model predicting survival in stage I melanoma based on tumor progression. *J Natl Cancer Inst* 1989;81:1893-904.
- Gershenwald JE, Scolyer RA. Melanoma staging: American Joint Committee on Cancer (AJCC) 8th edition and beyond. *Ann Surg Oncol* 2018;25:2105-10.
- Du P, Kibbe WA, Lin SM. Lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008;24:1547-8.
- Lauss M, Visne I, Kriegner A, Ringner M, Jonsson G, Hoglund M. Monitoring of technical variation in quantitative high-throughput datasets. *Cancer Inform* 2013;12:193-201.
- Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010;26:1572-3.
- Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learn* 2003;52:91-118.
- Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. New York: Springer; 2000.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77.
- Sachs M. plotROC: generate useful ROC curve charts for print and interactive use. Available from: <http://sachsmc.github.io/plotROC>. R package version 2.220.
- Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer; 2016.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci* 2001;98:5116-21.
- Wu G, Dawson E, Duong A, Haw R, Stein L. ReactomeFIViz: a cytoscape app for pathway and network-based data analysis. *F1000Res* 2014;3:146.
- Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *ICWSM* 2009;8:361-2.
- Filia A, Droop A, Harland M, Thygesen H, Randerson-Moor J, Snowden H, et al. High-resolution copy number patterns from clinically relevant FFPE material. *Sci Rep* 2019;9:8908.
- Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 2005;21:3439-40.
- Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 2009;4:1184-91.
- Cirenajwis H, Ekedahl H, Lauss M, Harbst K, Carneiro A, Enoksson J, et al. Molecular stratification of metastatic melanoma using gene expression profiling: prediction of survival outcome and benefit from molecular targeted therapy. *Oncotarget* 2015;6:12297.
- Angelova M, Charoentong P, Hackl H, Fischer ML, Snajder R, Krogsdam AM, et al. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol* 2015;16:64.
- Huang RY, Kuay KT, Tan TZ, Asad M, Tang HM, Ng AH, et al. Functional relevance of a six mesenchymal gene signature in epithelial-mesenchymal transition (EMT) reversal by the triple angiokinase inhibitor, nintedanib (BIBF1120). *Oncotarget* 2015;6:22098-113.

33. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–25.
34. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27:6199–206.
35. Morton DL, Cochran AJ, Thompson JF, Elashoff R, Essner R, Glass EC, et al. Sentinel node biopsy for early-stage melanoma: accuracy and morbidity in MSLT-I, an international multicenter trial. *Ann Surg* 2005;242:302–11; discussion 311–3.
36. Lauss M, Donia M, Harbst K, Andersen R, Mitra S, Rosengren F, et al. Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat Commun* 2017;8:1738.
37. Ramsdale R, Jorissen RN, Li FZ, Al-Obaidi S, Ward T, Sheppard KE, et al. The transcription cofactor c-JUN mediates phenotype switching and BRAF inhibitor resistance in melanoma. *Sci Signal* 2015;8:ra82.
38. Riesenbergs S, Groetchen A, Siddaway R, Bald T, Reinhardt J, Smorra D, et al. MITF and c-Jun antagonism interconnects melanoma dedifferentiation with pro-inflammatory cytokine responsiveness and myeloid cell recruitment. *Nat Commun* 2015;6:8755.