

ORIGINAL ARTICLE

Expression of long noncoding RNAs in cancer-associated fibroblasts linked to patient survival in ovarian cancer

Emily K. Colvin^{1,2}  | Viive M. Howell^{1,2} | Samuel C. Mok³ | Goli Samimi⁴ | Fatemeh Vafae⁵ 

¹Bill Walsh Translational Cancer Research Laboratory, Kolling Institute, Sydney, Australia

²Northern Clinical School, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

³Division of Surgery, Department of Gynecologic Oncology and Reproductive Medicine Research, The University of Texas MD Anderson Cancer Center, Houston, Texas

⁴Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, United States

⁵School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia

Correspondence

Fatemeh Vafae, School of Biotechnology and Biomolecular Sciences, University of New South Wales, Room 2106, Biological Sciences South Building (E26), Upper Kensington Campus, Sydney, NSW 2052, Australia.
Email: f.vafae@unsw.edu.au

Funding information

Cancer Institute NSW, Grant/Award Number: 2017/ECF007; National Institutes of Health, Grant/Award Number: R01CA142832 and R01CA169200; Proud Family Fellowship; Cancer Council NSW, Grant/Award Number: RG11-14; University of Texas MD Anderson Cancer Center, Grant/Award Number: P30CA016672 and P50CA083639

Abstract

Cancer-associated fibroblasts (CAFs) are the most abundant cell type in the tumor microenvironment and are responsible for producing the desmoplastic reaction that is a poor prognostic factor in ovarian cancer. Long non-coding RNAs (lncRNAs) have been shown to play important roles in cancer. However, very little is known about the role of lncRNAs in the tumor microenvironment. We aimed to identify lncRNAs expressed in ovarian CAFs that were associated with patient survival and used computational approaches to predict their function. Increased expression of 9 lncRNAs and decreased expression of 1 lncRNA in ovarian CAFs were found to be associated with poorer overall survival. A “guilt-by-association” approach was used to predict the function of these lncRNAs. In particular, *MIR155HG* was predicted to play a role in immune response. Further investigation revealed high *MIR155HG* expression to be associated with higher infiltrates of immune cell subsets. In conclusion, these data indicate expression on several lncRNAs in CAFs are associated with patient survival and are likely to play an important role in regulating CAF function.

KEYWORDS

biomarker, cancer-associated fibroblast, lncRNA, ovarian cancer, tumor microenvironment

Abbreviations: CAF, cancer-associated fibroblast; FDR, false discovery rate; GO, Gene Ontology; HGSOC, high-grade serous ovarian cancer; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; lncRNA, long noncoding RNA; miR, microRNA; PPI, protein-protein interaction.

Emily K. Colvin and Fatemeh Vafae contributed equally to this work.

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

© 2020 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

1 | INTRODUCTION

Epithelial ovarian cancer is the fifth leading cause of cancer death in women and the most lethal gynecological malignancy.¹ High-grade serous ovarian cancer is the most common and aggressive subtype of ovarian cancer, and despite advances in understanding the underlying genetic causes of HGSOC and improved treatment strategies, most women diagnosed with HGSOC have a poor prognosis.² Most women are diagnosed at an advanced stage and while initial response rates to chemotherapy are high, recurrence of chemoresistant disease is a significant problem.³ Continuing to improve our understanding of the factors that influence patient prognosis and response to therapy will be beneficial to improve treatment strategies and outcomes for women diagnosed with ovarian cancer.

Until recently, most studies have focused on a greater understanding of the molecular changes present in ovarian cancer cells and how these affect tumor progression and patient outcome. However, increasingly the tumor microenvironment is gaining recognition as playing a vital part in the initiation, survival, growth, and metastasis of tumors.⁴ In addition, cells within the tumor microenvironment are more genetically stable than cancer cells, which potentially reduces the likelihood of continued treatment causing the accumulation of genetic changes and subsequent development of acquired resistance to therapy.⁵ For these reasons, the tumor microenvironment is emerging as an attractive therapeutic target to treat cancer.

In ovarian cancer, the stromal proportion present in tumors can vary from 7% to 83% of tumor tissue, and patients with a higher stromal proportion have a worse overall survival.⁶ Furthermore, expression profiling of HGSOC has identified a subtype that displays a "stromal expression signature".⁷ Importantly, patients with this signature show higher levels of desmoplastic stroma and the poorest survival. Within the tumor microenvironment, the stroma contains multiple cell types such as endothelial cells, immune cells, and CAFs, which have all been shown to contribute to cancer progression. Cancer-associated fibroblasts represent the most abundant cell type in the tumor stroma and are responsible for producing the desmoplastic reaction that is a poor prognostic factor in HGSOC. Additionally, CAFs have been shown to play multiple roles in ovarian cancer to promote tumor cell proliferation, migration and invasion.⁸⁻¹¹

Gene expression profiling of several cancer types has revealed marked heterogeneity of CAFs.¹²⁻¹⁴ Therefore, distinct subtypes of CAFs could play varying roles in the tumor microenvironment and influence patient survival differently. Cancer-associated fibroblast or stromally derived prognostic gene expression signatures have been identified in several cancer types.¹⁵⁻¹⁷ Furthermore, in ovarian cancer, studies have uncovered genes differentially expressed in CAFs that are predictive or prognostic biomarkers such as *VCAN*,¹⁸ *CTGF*,¹⁹ *MFAP5*,¹⁰ *FOSB*,²⁰ *EGR1*,²⁰ and *NPPB*.⁹ However, studies investigating the DNA mutations in ovarian CAFs have concluded that somatic mutations are unlikely to contribute to gene expression changes seen in ovarian CAFs and raise the likelihood that alternative mechanisms of gene regulation occur in CAFs.²¹ A study published by Mitra et al showed that changes in the expression levels of miR-31,

miR-214, and miR-155 contribute to the reprogramming of normal fibroblasts into CAFs.²² In other cancer types, DNA methylation changes have also been shown to occur in CAFs.²³⁻²⁶ A greater understanding of how gene expression is regulated in CAFs will help to identify new stromal biomarkers and potential therapeutic targets.

Long non-coding RNAs represent another possible mechanism for regulating gene expression in CAFs. We have previously shown differences in lncRNA expression in ovarian CAFs compared to normal ovarian fibroblasts and that several of these lncRNAs might promote the prometastatic role of CAFs in ovarian cancer.²⁷ Long noncoding RNAs are noncoding RNAs greater than 200 nucleotides long that do not encode protein. Once thought to be "transcriptional noise," lncRNAs are now recognized to play crucial roles in several biological functions such as chromatin modification, transcription, and translation.²⁸ They have also been shown to play important roles in several diseases, including cancer.²⁹ Studies have identified lncRNAs involved in ovarian cancer that are also candidate prognostic biomarkers.^{30,31} However, these studies have been restricted to whole tumor specimens or cell lines and none have examined the role of lncRNAs in the tumor microenvironment. As CAFs are known to represent a heterogeneous population of cells, with varying functional capacities that could influence patient outcome, this study aimed to investigate lncRNA expression in ovarian CAFs to determine those associated with patient outcome. We then employed a network-based "guilt-by-association" approach to predict their functions. Additional analysis indicated that increased expression of one lncRNA, *MIR155HG*, was associated with significant increases in several immune cell subsets.

2 | METHODS

2.1 | Tissue specimens

Primary tumor specimens from 67 women diagnosed with HGSOC were obtained as previously described.^{19,32} All specimens were from previously untreated HGSOC patients hospitalized at the Brigham and Women's Hospital between 1990 and 2000. Patient specimens and corresponding clinical information were collected by written consent under protocols approved by the review board of the Brigham and Women's Hospital Ethics Committee. All procedures were carried out in accordance with the approved guidelines and regulations. Classification was determined according to the International Federation of Gynecology and Obstetrics standards. Survival information was not available for 5 patients whose samples were excluded.

2.2 | Microdissection, RNA isolation, amplification, and hybridization

Microdissection, RNA isolation, amplification, and hybridization to GeneChip Human Genome U133 Plus 2.0 Oligonucleotide arrays

(Affymetrix) are previously described.¹⁹ Gene expression of endothelial cell markers (*TIE-2* and *VEGFR1*) and T cell markers (*CD8* and *CD45*) were below the level of detection in our samples, indicating a lack of immune or endothelial components and enrichment for fibroblasts.¹⁹ All gene array data are available through Gene Expression Omnibus accession number GSE40595.

2.3 | Statistical data preprocessing

Data preprocessing was undertaken using R Bioconductor, “affy” package. Data were normalized and background corrected using the Robust Multi-Array Average method³³ and expression values Log2 transformed. Variations across samples were assessed using the interquartile range values for each probe, and those with interquartile range less than 1 were removed for subsequent analyses. A total of 2448 probes were identified previously to be associated with lncRNAs.³⁴ The gene symbols and titles corresponding to these probes were retrieved from GeneAnnot,³⁵ which provides revised and improved annotations of the Affymetrix Human Genome U133 Plus 2.0 probes.

2.4 | Survival analysis

The expression levels of all lncRNAs across samples were separated into low vs high expression using a fuzzy clustering algorithm, wherein each data point belongs to a cluster to some degree that is specified by a *membership* degree.³⁶ The memberships are nonnegative, and for a fixed sample, they sum to 1. For each lncRNA, the fuzzy clustering algorithm was set to identify 2 clusters of samples representing higher vs lower expression levels. Samples belonging to either of clusters with the membership degree greater than 0.7 were included, ie, the remaining samples were considered as “undetermined” and excluded from subsequent analyses, resulting in 2 distinct groups of low vs high expression for each lncRNA. Fuzzy clustering was undertaken using R “cluster” package.

Kaplan-Meier analysis and the log-rank test were used to assess the association between the expression level of each lncRNA in CAFs and the patients' overall survival. The prognostic value of each lncRNA's expression levels as well as debulking status and chemotherapy response (sensitive vs the rest) was determined with univariate Cox proportional hazard modelling, and those significantly related to survival were incorporated into a multivariate analysis. In order to overcome the multicollinearity among lncRNA expression, principle component regression analysis was performed.³⁷ Accordingly, instead of lncRNA expression profiles, the corresponding principal components were used as covariates in the multivariate Cox regression analysis. The Wald test was used to assess the statistical significance of the Cox models ($\alpha = 0.05$). All survival analyses were carried out using R “survival” package; tied event times were handled by Breslow's approximation.³⁸

2.5 | Functional prediction of lncRNAs associated with patient survival

Potential functions of prognostic CAF-expressed lncRNAs were predicted using a network-based “guilt-by-association” approach as follows.

2.5.1 | Construction of coexpressed “interactome”

A coexpression network was first constructed where nodes are the identified lncRNAs and all protein-coding genes and edges represent significant correlations, ie, $|\text{Pearson's correlation coefficient}| > 0.7$, correlation adjusted *P*-value $< 10E-6$. The gene coexpression network was then mapped on a cellular *interactome* comprising protein-protein and gene regulatory interactions. Experimentally validated human PPIs detected in more than 2 experiments were combined with highly ranked predicted PPIs (ie, FDR $> 60\%$) as predicted by kotlyar et al³⁹ to form a comprehensive PPI database. An experimentally derived gene regulatory network was secured from ORTI,⁴⁰ a comprehensive repository of mammalian transcriptional interactions. The resulting network is a *coexpressed interactome* comprising coexpressed protein-protein or gene regulatory interactions plus coexpressed lncRNA-gene associations.

Network modules were identified using “community” detection algorithms where communities are groups of nodes with dense connections internally and sparser external connections.⁴¹ Communities represent transcripts that are more likely to be involved in distinct similar biological processes and thus can be used to assign functions to lncRNAs associated with them. All network analyses were performed using R “igraph” package. Different community detection algorithms were tried out, eg, Louvain,⁴² greedy,⁴³ infomap,⁴⁴ and walktrap.⁴⁵ The Louvain algorithm found clusters with a relatively higher modularity score and thus was used to report the results.

2.5.2 | Functional enrichment analysis

Network modules containing at least 1 lncRNA underwent GO and pathway enrichment analysis using the R “enrichR” package which implements Fisher's exact test and FDR adjustment on a wide range of gene set libraries.⁴⁶ We used biological processes (GO_Biological_Process_2017b, comprising 10 125 GO terms on 13 247 genes) and KEGG pathways (KEGG_2016, comprising 293 pathways on 7010 genes) as EnrichR datasets to predict putative functions of lncRNAs. For ease of visualization, GO terms enriched by each module were summarized into representative subsets of the terms using REViGO.⁴⁷

2.6 | Validation of *MIR155HG* in stromal-enriched whole tumor specimens

To determine whether *MIR155HG* was prognostic in whole tumor samples, we used the same cohort as was used for the CIBERSORT analysis. Tothill et al⁷ have previously clustered these samples into 6 molecular

subtypes (C1-C6) using *k*-means clustering and identified that C1 is enriched by genes associated with stromal cell types, enabling us to validate *MIR155HG* prognostication in an independent cohort.

We followed our pipeline to preprocess GSE9899 raw data and to compare survival differences between patients with high and low *MIR155HG* expression using Kaplan-Meier analysis and the log-rank test in each of the subgroups. Average *MIR155HG* gene expression was also compared between the subgroups and differences were compared using the nonparametric Wilcoxon test to account for nonnormality of *MIR155HG* mean expression across samples within each subgroup.

2.7 | CIBERSORT analysis to determine immune cell infiltrates

CIBERSORT is an analytical tool designed to accurately estimate the immune cell subsets present in whole tumor samples from their gene expression profiles.⁴⁸ We used the default *LM22 signature matrix* consisting of 547 genes that accurately distinguish 22 mature human hematopoietic populations and activation states, including 7 T cell types, naïve and memory B cells, plasma cells, natural killer cells, and myeloid subsets.

Using CIBERSORT, we quantified immune cell infiltrates from a cohort of 83 previously characterized high-grade ovarian tumors⁷ described as having a stromal expression signature and an increased density of fibroblasts (C1) with data available from Gene Expression Omnibus, accession number GSE9899. Differences were compared between tumors with high and low *MIR155HG* expression. Data are expressed as absolute fractions of each immune cell type and differences between groups were measured by *t* test, with significance determined by a *P* value of less than .05.

3 | RESULTS

3.1 | Long noncoding RNA expression levels in ovarian CAFs associated with patient prognosis

Characteristics of patients and tumor samples included in this study are shown in Table 1. Microdissected CAF samples were quality

controlled for CAF composition using 14 gene markers of CAFs and fibroblasts previously reported.⁴⁹ To this end, differential expression analysis was carried out comparing CAFs vs matched microdissected epithelial tumors; a total of 7161 genes were identified to be significantly upregulated in CAFs (adjusted *P* value < .05 using moderated *t* test with FDR correction) covering 86% (12/14) of CAF markers, which indicates significant enrichment for CAF composition (ie, *P* value = 1.02e-5, Fisher's exact test). Additionally, expression profiles of marker genes in each individual CAF sample were assessed and visualized as shown in Figure S1, clearly illustrated the high expression level of CAF markers across all samples.

Kaplan-Meier survival analysis indicated increased or decreased expression levels of 10 lncRNAs in ovarian CAFs were associated with patients' overall survival (*P*-value < .05). The symbols, chromosomal locations, and titles of these 10 lncRNAs are listed in Table 2. Kaplan-Meier plots of each lncRNA identified distinct survival trends between samples in groups of high vs low expression (Figure 1). These groups were identified using a fuzzy clustering algorithm that was set to categorize samples into 2 clusters representing high vs low expression levels. The box plots beside each Kaplan-Meier plot show how samples were clustered into 2 groups of high and low expression. Fuzzy clustering assigns membership grades to each sample indicating the degree to which it belongs to each cluster. The middle gray box in each plot represents "uncategorized" and "removed" samples that did not strongly belong to either group (membership degree < 0.7). We also undertook an identical analysis on the gene expression data using microdissected epithelial tumor cells from matched patient samples. Other than *CRNDE*, none of the identified lncRNAs was significantly associated with patients' overall survival (Figure 2), confirming the CAF-specific prognostic utility of the identified lncRNAs. Additionally, CAF expression patterns of the lncRNAs were not significantly correlated with their expression in matched epithelial samples (Figure 2), corroborating CAF distinct regulatory mechanisms.

The statistics of univariate cox proportional hazards analysis are shown in Figure 3A. Expression profiles of *CRNDE*, *MALAT1*, *MEG3*, *TP73-AS1*, and *XIST*, as well as chemoresponse and tumor debulking, were statistically significant predictors of mortality in univariate analysis. The identified lncRNAs showed mutually significant

Characteristic	n = 62	Description
Age at diagnosis, years (mean ± SD)	60.94 ± 12.37	—
Stage (III/IV), grade	55/7, 3	—
Debulking (optimal/suboptimal)	49/13	Optimal debulking corresponds to <1 cm residual tumor
Site, histological type	Ovary, serous	—
Chemoresponse (R/S/R-S/Ref)	18/24/7/4	R, resistant (recurred < 6 months); S, sensitive (recurred > 6 months); R-S, resistant-sensitive (recurred at 6 months); Ref, refractory (never responded)

TABLE 1 Clinical characteristics of patients with ovarian cancer and tumor samples

TABLE 2 Long noncoding RNAs (lncRNAs) associated with overall survival in ovarian cancer patients based on Kaplan-Meier analysis (P value < .05)

Symbol	Chromosome	Title
<i>CRNDE</i>	chr16q12.2	Colorectal neoplasia differentially expressed (nonprotein coding)
<i>DANCR</i>	chr4q12	Differentiation antagonizing non-protein coding RNA
<i>LOC642852</i>	chr21q22.3	Uncharacterized LOC642852
<i>MALAT1</i>	chr11q13.1	Metastasis associated lung adenocarcinoma transcript 1 (nonprotein coding)
<i>MEG3</i>	chr14q32	Maternally expressed 3 (nonprotein coding)
<i>MGC2752</i>	chr19q13.43	Uncharacterized LOC100653267
<i>MIR155HG</i>	chr21q21.3	MIR155 host gene (nonprotein coding)
<i>NEAT1</i>	chr11q13.1	Nuclear paraspeckle assembly transcript 1 (nonprotein coding)
<i>TP73-AS1</i>	chr1p36.32	TP73 antisense RNA 1 (nonprotein coding)
<i>XIST</i>	chrXq13.2	X (inactive)-specific transcript (nonprotein coding)

correlations with each other (Figure 3B), but not with chemoresponse or tumor debulking (Figure S2). To adjust for existing collinearity among lncRNAs, the first principal component of these 10 lncRNAs (capturing 98% of variations) as well as clinical characteristics (ie, chemoresponse and debulking) were used as regressors in multivariate analysis. The first principal component of the lncRNAs (P value = .000116, HR = 0.74) and chemoresponse (P value = .000168, HR = 0.22) were significant predictors of survival in multivariate cox analysis. Debulking status approached, but did not reach statistical significance (P value = .067754, HR = 1.91).

3.2 | Prognostic lncRNAs in ovarian CAFs enriched for pathways known to be involved in CAF function

“Guilt-by-association” assigns putative functions to coding/noncoding transcripts based on genes coexpressed with them. It relies on the idea that genes with similar expression patterns across multiple samples are more likely to be coregulated, share similar functions, or are involved in similar biological processes.⁵⁰ Coexpression network analyses have been previously used to predict functions of lncRNAs.⁵¹⁻⁵³

Figure 4A depicts a schematic view of a network-based “guilt-by-association” approach followed in this work. We first constructed a coexpression network whose nodes include all protein-coding genes as well as 10 lncRNAs identified by the survival analysis; edges represent coexpression relationships ($|\text{Pearson's correlation coefficient}| > 0.7$). This network held 6791 nodes and 5 557 325 edges and shows a relatively low degree of modularity (0.18) where 91% of nodes fall into 3 gigantic clusters. A high degree of modularity, however, has often been reported in biological networks.⁵⁴ We mapped the coexpression network on a cellular interactome comprising PPIs and gene regulatory interactions to derive a coexpressed interactome. The corresponding

network held 6791 nodes and 43 545 edges whose modularity was improved to 0.47. Overall, 17 clusters (excluding singletons, ie, clusters including only 1 member) were identified; 3 clusters contained the identified lncRNAs. *DANCR*, *LOC642852*, *MALAT1*, *MEG3*, *MGC2752*, *TP73-AS1*, and *XIST* were coclustered in a relatively large module of 1711 genes. *MIR155HG* fell into a separate cluster (size = 242 genes), which is in concordance with the correlation pattern of lncRNAs visualized in Figure 2B. *CRNDE* was clustered with only 2 protein-coding genes, which is not sufficient for enrichment analysis, and *NEAT1* was identified as a singleton. *NEAT1* does not show sufficiently high correlation with any gene, indicating that the correlation-based guilt-by-association analysis cannot reveal its function and complementary analyses are required. We therefore undertook functional enrichment analysis on the 2 former clusters comprising 8 of 10 lncRNAs of interest and identified cellular processes and pathways overrepresented (adjusted P value < .05) by the corresponding 2 subnetworks as predicted by functions of the constituent lncRNAs. Figure 4B shows the clusters and summarizes representative enriched GO terms. A list of all enriched GO terms as well as overrepresented pathways and cluster composition are available in Document S1. The cluster containing *DANCR*, *LOC642852*, *MALAT1*, *MEG3*, *MGC2752*, *TP73-AS1*, and *XIST* showed an enrichment for pathways primarily involved in metabolic processes as well as autophagy and cilium assembly. The cluster containing *MIR155HG* was enriched for GO terms associated with the immune system, particularly pathways associated with T cell activation, antigen processing and presentation, leukocyte migration, and activation of an immune response. This cluster was also enriched for pathways involved in ECM organization, cell death, metabolic processes, and cytokine signaling. The KEGG pathways related to infectious diseases, immune diseases, and the immune system were also enriched, suggesting a role for *MIR155HG* in regulating the immune microenvironment.

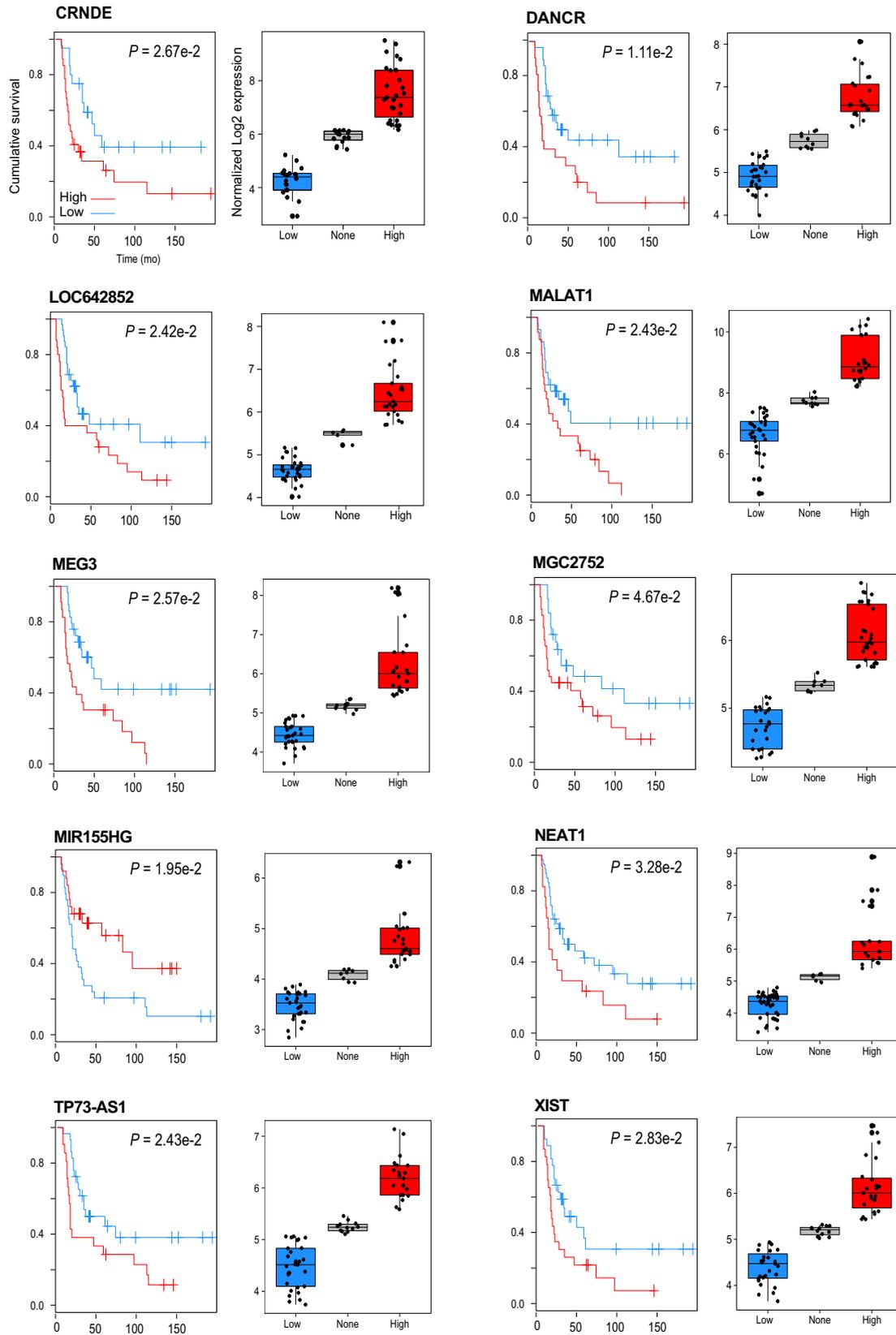


FIGURE 1 Differentially expressed long noncoding RNAs (lncRNAs) in cancer-associated fibroblasts associated with significant differences in overall survival among ovarian cancer patients. Higher expression of 9 lncRNAs was associated with shorter survival, whereas increased expression of *MIR155HG* was associated with longer survival as depicted in the Kaplan-Meier curves. Box plots show results from the fuzzy clustering algorithm that separates high and low expression into 2 distinct groups

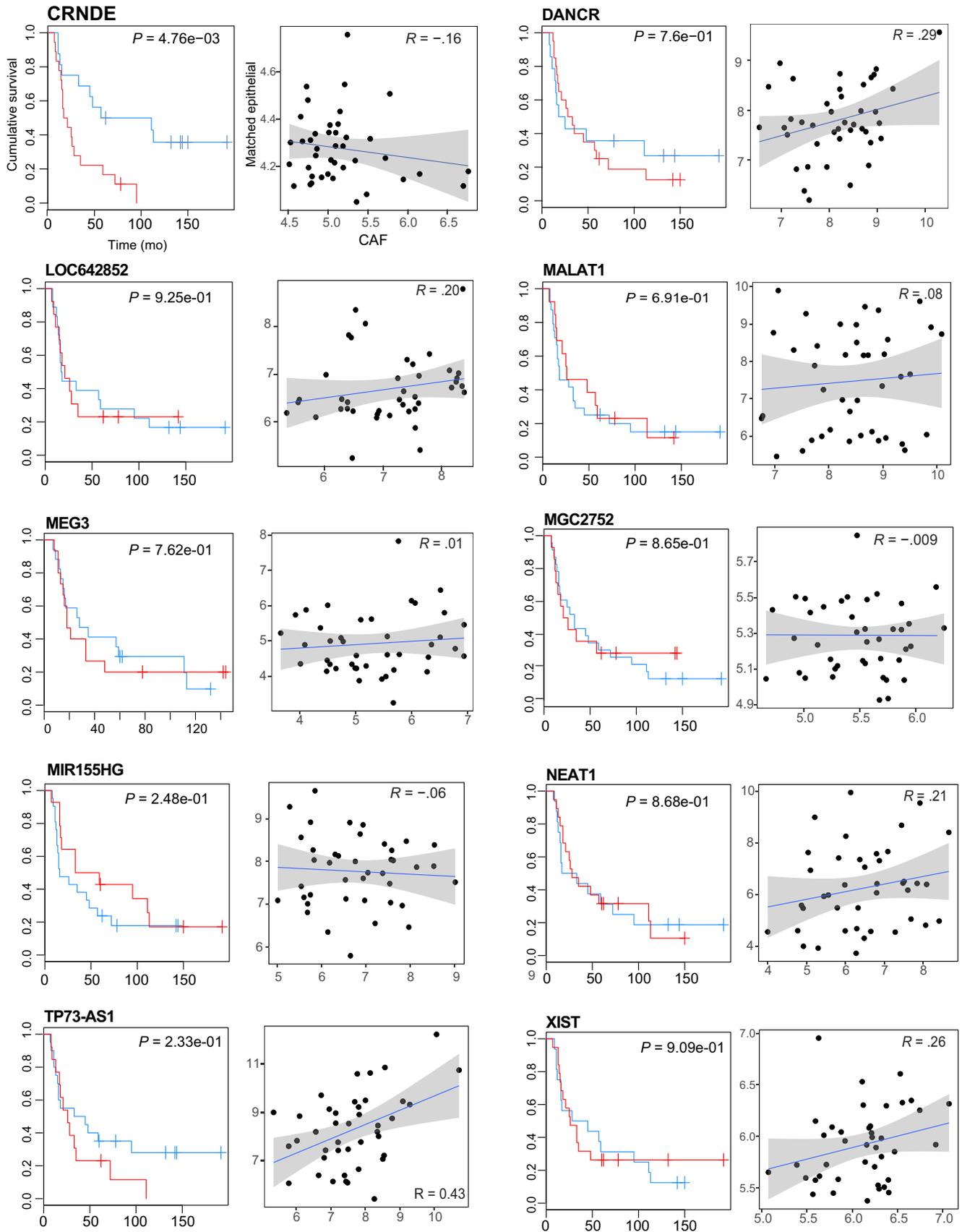


FIGURE 2 Expression of long noncoding RNAs (lncRNAs) that are prognostic in cancer-associated fibroblasts (CAFs) are not prognostic in matched tumor epithelium of ovarian cancer patients. With the exception of *CRNDE*, expression of the lncRNAs in tumor epithelium were not associated with differences in patient survival, as depicted in the Kaplan-Meier curves. Expression of lncRNAs was not correlated between microdissected CAF samples and matched tumor epithelium

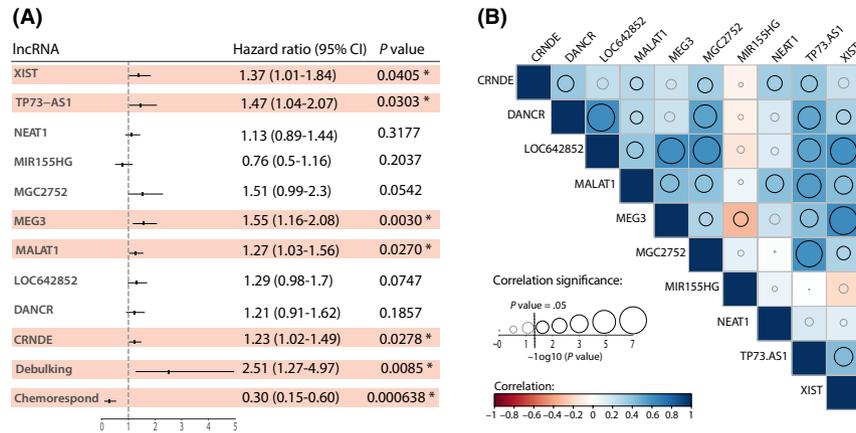


FIGURE 3 Univariate Cox proportional hazards and correlation analyses. A, Expression levels of *CRNDE*, *MALAT1*, *MEG3*, *TP73-AS1*, and *XIST* (highlighted) were significant predictors of mortality among patients with ovarian cancer as well as debulking status and chemoresponse. B, The majority of lncRNAs were significantly positively correlated with each other (blue shaded boxes, darker blue represents stronger correlations). Nonsignificant correlations are depicted by small gray circles; significant correlations are depicted by larger black circles as outlines in the correlation significance scale

3.3 | Validation of *MIR155HG* expression and prognostication in stromal-enriched whole tumor specimens

We were interested to know whether *MIR155HG* would remain prognostic using gene expression data from whole tumor samples. Therefore, we examined its prognostic value using an independent dataset.⁷ In this dataset, samples were classified into 6 subtypes (C1-C6) based on gene expression. As shown in Figure 5A, *MIR155HG* was only prognostic in the C1 subtype (log-rank test P value = .0199). This subtype is described as having a stromal expression signature and an increased density of myofibroblasts. Therefore, our observation that *MIR155HG* is only prognostic in the Cluster 1 subgroup likely reflects the higher contribution of CAFs to the observed gene expression in this subtype.

We also compared *MIR155HG* expression levels between the different subtypes. *MIR155HG* expression was significantly higher in the C1 and C2 subtypes (Figure 5B). Both these subtypes have previously been associated with higher levels of infiltrating CD3⁺ T cells, with the C1 subtype showing high levels of stromal CD3⁺ T cells and the C2 subtype showing a high level of intratumoral CD3⁺ T cells.⁷ The higher *MIR155HG* expression seen in these subtypes supports our functional prediction analysis and the CIBERSORT results showing higher CD3⁺ T cells in tumors with high *MIR155HG* expression.

3.4 | High *MIR155HG* expression associated with increase in immune cell subsets in stromal-enriched whole tumor specimens

Based on the enrichment analysis and the longer survival seen in patients with high *MIR155HG* expression, we hypothesized that

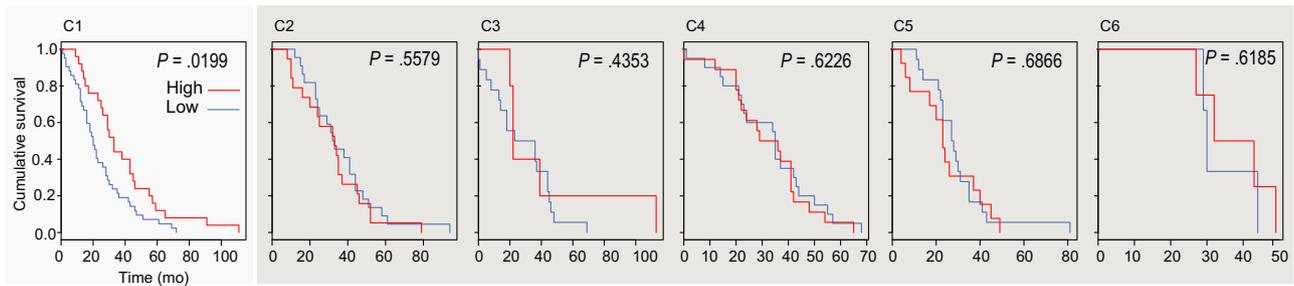
MIR155HG is associated with differences in immune cell infiltrates within the tumor. In order to investigate this further, we used CIBERSORT⁴⁸ to examine the immune infiltrates present in whole tumor specimens obtained from a cohort of 285 ovarian cancer patients,⁷ separated by their *MIR155HG* expression. As shown in Figure 6, tumors with high *MIR155HG* expression had significantly higher numbers of plasma cells, CD8⁺ T cells, CD4⁺ memory activated T cells, follicular helper T cells, gamma delta T cells, M1 macrophages, and eosinophils.

4 | DISCUSSION

Cells within the microenvironment of solid tumors are not passive bystanders in tumor progression and metastasis but play an active and essential role. In many tumor types, including ovarian, CAFs are known to influence tumor cell behavior by increasing tumor cell survival, proliferation, migration, and invasion.⁸⁻¹¹ Cancer-associated fibroblasts also interact with the other cell types present in the tumor microenvironment to promote angiogenesis and help tumor cells evade immune destruction.^{13,55} Given these findings, a greater understanding of the molecular features of CAFs and their potential role in the clinical behavior of tumors is essential when designing new therapies that target CAFs.

We recently reported that several lncRNAs are differentially expressed in ovarian CAFs compared to normal ovarian fibroblasts and that several of these lncRNAs contribute to the prometastatic phenotype of CAFs.²⁷ In the current study, we investigated whether differences in lncRNA expression in CAFs influence patient outcome in HGSOV. Given the importance of CAFs in ovarian cancer, there is a rationale for exploring the molecular aberrations present in CAFs as these could provide valuable prognostic information. We identified 10 lncRNAs with variable expression in ovarian CAFs that were

(A)



(B)

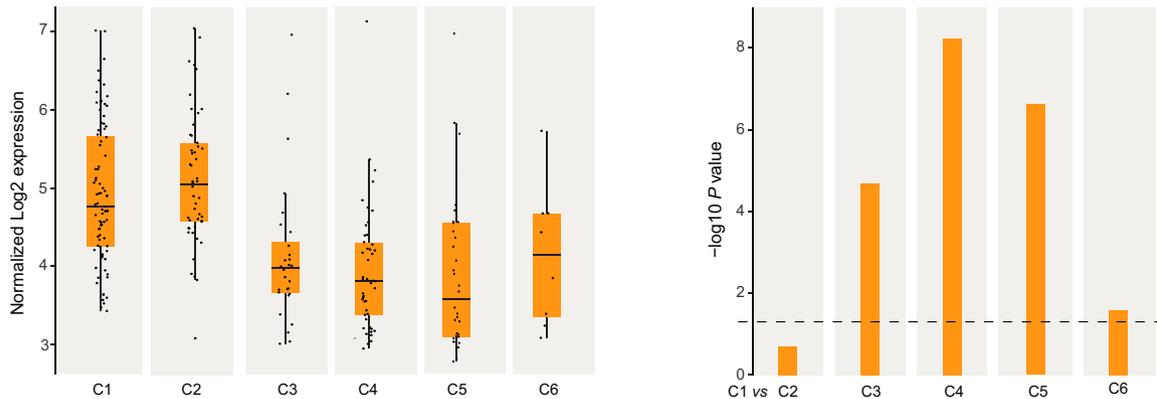


FIGURE 5 Analysis of *MIR155HG* in whole tumor specimens. A, High expression of *MIR155HG* was able to predict prolonged survival in the C1 subtype of ovarian tumors that have been shown to have a stromal gene expression signature, but was not prognostic in the other subtypes. B, C1 and C2 subsets showed significantly higher *MIR155HG* expression than the other subtypes

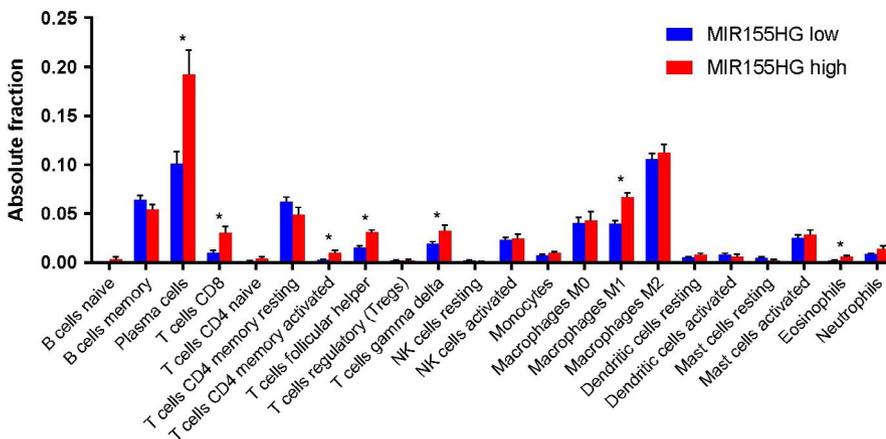


FIGURE 6 CIBERSORT immune cell infiltrate analysis. Graph displays the absolute fraction (%) ± SEM of different immune cell subsets within ovarian tumors, separated into *MIR155HG* low and high expression groups. *Significant difference (P value < .05)

associated with overall survival in HGSOC. In addition, an expression signature based on these 10 lncRNAs was an independent predictor of patient survival. Several of these lncRNAs are already known to play a role in either ovarian cancer,⁵⁶⁻⁵⁸ or other cancer types⁵⁹; however, these previous studies have only examined lncRNAs in whole tumor specimens, therefore, it is not clear whether it is expression in the tumor cells or the microenvironment that is associated with patient survival. By analyzing expression data from microdissected CAFs and matched tumor cells, we were able to show that for 9 out of 10 of our lncRNAs, differential expression in CAFs specifically,

and not tumor cells, was associated with patient survival. This suggests for the first time that these lncRNAs could play an important role in CAFs and the tumor microenvironment.

Even though several of the lncRNAs identified in this study have previously been studied in ovarian cancer or other cancers, their function in CAFs is not clear. In addition, *LOC642852* and *MGC2752* are not well characterized. In order to elucidate the potential functions of these lncRNAs in CAFs we used a network-based guilt-by-association approach. The majority of lncRNAs clustered together, suggesting they play similar roles in

CAFs. This cluster was associated with pathways involved in metabolism, autophagy, and cilium assembly. Cancer-associated fibroblasts are already known to play a role in the metabolic reprogramming of the tumor microenvironment in order to favor cancer growth and metastasis.⁶⁰ Through alterations in their metabolic activity, CAFs take on a catabolic phenotype that is then able to provide nutrients to anabolic cancer cells.⁶¹ In ovarian cancer, CAFs have been shown to have altered metabolism compared to normal ovarian fibroblasts and targeting this altered metabolism resulted in tumor regression.⁶² An essential part of tumor metabolism, the process of autophagy, is activated in CAFs as a potential mechanism to allow CAFs to provide metabolic products to feed cancer cells, and high levels of autophagy in the tumor microenvironment has been associated with cancer progression.⁶³ In ovarian cancer, autophagy could protect ovarian CAFs against oxidative stress.⁶⁴ Supporting our findings, both *MEG3* and *MALAT1* have been shown to induce autophagy in ovarian cancer.^{65,66} Cilium assembly pathways were also enriched in this cluster. Primary cilia are important for signaling between stromal cells and adjacent tumor cells and autophagy promotes cilia formation.⁶⁷ Enrichment of pathways associated with metabolism, autophagy, and cilium assembly suggests that lncRNAs belonging to this cluster might be important in creating a metabolic environment conducive to ovarian tumor growth. Higher levels of these lncRNAs could be indicative of more metabolically active and aggressive tumors, resulting in worse patient survival. However, further studies are required to validate the functional roles of these lncRNAs in ovarian CAF metabolism and autophagy.

The other cluster identified in this study contained the lncRNA associated with longer survival, *MIR155HG*. *MIR155HG* was originally identified as a proto-oncogene in B-cell lymphomas⁶⁸ and is known to regulate many immune and inflammatory processes.⁶⁹ The role of *MIR155HG* has not been well-studied in cancer; however, recent studies have shown increased expression to be associated with worse survival in glioma and pancreatic adenocarcinoma patients, but improved survival in colorectal cancer patients.⁷⁰⁻⁷³ In tumor cells, *MIR155HG* appears to have an oncogenic role and is associated with increased cell growth and decreased apoptosis.^{73,74} However, *MIR155HG* could play a different role in CAFs. Our functional prediction analysis showed the cluster containing *MIR155HG* was highly enriched for immune-related pathways involved in T cell activation and activation of an immune response. Given patients with higher expression of *MIR155HG* in CAFs survived for longer, *MIR155HG* could be an important component in the interaction between CAFs and immune cells, and might be promoting or permitting an antitumor immune response. The CIBERSORT analysis indicated high *MIR155HG* expression to be associated with increased immune cell subsets previously shown to be associated with improved survival,⁷⁵ further experiments manipulating *MIR155HG* expression in CAFs are required to determine whether or not *MIR155HG* is directly involved in the induction of an antitumor immune response. Interestingly, *MIR155HG*'s associated miRNA, miR-155, is a well-known regulator of immunity and its expression has been shown in preclinical

mouse models to be essential for mounting an antitumor immune response.^{76,77} This finding could support a role for *MIR155HG* in antitumor immunity as many miRNA host genes have been shown to have similar functions to their associated miRNA,⁷⁸ however, this requires further investigation.

The concept that lncRNAs are involved in the regulation of the immune system and more specifically in the regulation of tumor immunity is relatively recent and therefore their importance in the tumor microenvironment and their potential clinical utility is not yet well known. Given that high *MIR155HG* expression in CAFs is associated with increased patient survival as well as increased infiltration of antitumor immune cell subsets, *MIR155HG* expression could represent a useful biomarker to predict response to immunotherapy.

In summary, we have identified that variable expression of several lncRNAs in ovarian CAFs is linked to patient survival. Functional prediction using computational models highlights several potential ways these lncRNAs are regulating the ovarian tumor microenvironment to create an environment for tumors to grow and metastasize and evade immune destruction. Given the crucial role of the tumor microenvironment in cancer initiation and progression, continuing to understand the complexity of CAFs is essential to identify novel biomarkers and improved ways to therapeutically target the tumor microenvironment.

ACKNOWLEDGEMENTS

This work was supported in part by a Cancer Institute NSW Fellowship (2017/ECF007 to EKC), Proud Family Fellowship (to EKC), Cancer Council NSW project grant (RG11-14 to VMH), NIH (R01CA169200 and R01CA142832 to SCM), The University of Texas MD Anderson Cancer Center Ovarian Cancer Specialized Program of Research Excellence grant (P50CA083639 to SCM), MD Anderson Support Grant from the NIH, the US Department of Health and Human Services (P30CA016672 to SCM).

CONFLICT OF INTEREST

The authors have no conflict of interest.

ORCID

Emily K. Colvin  <https://orcid.org/0000-0002-5435-5902>

Fatemeh Vafae  <https://orcid.org/0000-0002-7521-2417>

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68(1):7-30.
2. Cannistra SA. Cancer of the ovary. *N Engl J Med.* 2004;351(24):2519-2529.
3. Bovicelli A, D'Andrilli G, Giordano A. New players in ovarian cancer. *J Cell Physiol.* 2011;226(10):2500-2504.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.
5. Joyce JA. Therapeutic targeting of the tumor microenvironment. *Cancer Cell.* 2005;7(6):513-520.
6. Labiche A, Heutte N, Herlin P, Chasle J, Gauduchon P, Elie N. Stromal compartment as a survival prognostic factor in advanced ovarian carcinoma. *Int J Gynecol Cancer.* 2010;20(1):28-33.

7. Tothill RW, Tinker AV, George J, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res*. 2008;14(16):5198-5208.
8. Lau TS, Chung TK, Cheung TH, et al. Cancer cell-derived lymphotoxin mediates reciprocal tumour-stromal interactions in human ovarian cancer by inducing CXCL11 in fibroblasts. *J Pathol*. 2014;232(1):43-56.
9. Lawrenson K, Grun B, Lee N, et al. NPPB is a novel candidate biomarker expressed by cancer-associated fibroblasts in epithelial ovarian cancer. *Int J Cancer*. 2015;136(6):1390-1401.
10. Leung CS, Yeung TL, Yip KP, et al. Calcium-dependent FAK/CREB/TNNC1 signalling mediates the effect of stromal MFAP5 on ovarian cancer metastatic potential. *Nat Commun*. 2014;5:5092.
11. Yeung TL, Leung CS, Mok SC. CAF reprogramming inhibits ovarian cancer progression. *Cell Cycle*. 2014;13(24):3783-3784.
12. Lili LN, Matyunina LV, Walker LD, Benigno BB, McDonald JF. Molecular profiling predicts the existence of two functionally distinct classes of ovarian cancer stroma. *Biomed Res Int*. 2013;2013:846387.
13. Costa A, Kieffer Y, Scholer-Dahirel A, et al. Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. *Cancer Cell*. 2018;33(3):463-479.e10.
14. Ohlund D, Handly-Santana A, Biffi G, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med*. 2017;214(3):579-596.
15. Navab R, Strumpf D, Bandarchi B, et al. Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc Natl Acad Sci U S A*. 2011;108(17):7160-7165.
16. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med*. 2008;14(5):518-527.
17. Herrera M, Islam AB, Herrera A, et al. Functional heterogeneity of cancer-associated fibroblasts from human colon tumors shows specific prognostic gene expression signature. *Clin Cancer Res*. 2013;19(21):5914-5926.
18. Ghosh S, Albitar L, LeBaron R, et al. Up-regulation of stromal versican expression in advanced stage serous ovarian cancer. *Gynecol Oncol*. 2010;119(1):114-120.
19. Moran-Jones K, Gloss BS, Murali R, et al. Connective tissue growth factor as a novel therapeutic target in high grade serous ovarian cancer. *Oncotarget*. 2015;6(42):44551-44562.
20. Kataoka F, Tsuda H, Arai T, et al. EGRI and FOSB gene expressions in cancer stroma are independent prognostic indicators for epithelial ovarian cancer receiving standard therapy. *Genes Chromosomes Cancer*. 2012;51(3):300-312.
21. Qiu W, Hu M, Sridhar A, et al. No evidence of clonal somatic genetic alterations in cancer-associated fibroblasts from human breast and ovarian carcinomas. *Nat Genet*. 2008;40(5):650-655.
22. Mitra AK, Zillhardt M, Hua Y, et al. MicroRNAs reprogram normal fibroblasts into cancer-associated fibroblasts in ovarian cancer. *Cancer Discov*. 2012;2(12):1100-1108.
23. Fiegl H, Millinger S, Goebel G, et al. Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res*. 2006;66(1):29-33.
24. Hanson JA, Gillespie JW, Grover A, et al. Gene promoter methylation in prostate tumor-associated stromal cells. *J Natl Cancer Inst*. 2006;98(4):255-261.
25. Hu M, Yao J, Cai L, et al. Distinct epigenetic changes in the stromal cells of breast cancers. *Nat Genet*. 2005;37(8):899-905.
26. Jiang L, Gonda TA, Gamble MV, et al. Global hypomethylation of genomic DNA in cancer-associated myofibroblasts. *Cancer Res*. 2008;68(23):9900-9908.
27. Vafae F, Colvin EK, Mok SC, Howell VM, Samimi G. Functional prediction of long non-coding RNAs in ovarian cancer-associated fibroblasts indicate a potential role in metastasis. *Sci Rep*. 2017;7(1):10374.
28. Geisler S, Collier J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol*. 2013;14(11):699-712.
29. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev*. 2009;23(13):1494-1504.
30. Zhou M, Wang X, Shi H, et al. Characterization of long non-coding RNA-associated ceRNA network to reveal potential prognostic lncRNA biomarkers in human ovarian cancer. *Oncotarget*. 2016;7(11):12598-12611.
31. Guo Q, Cheng Y, Liang T, et al. Comprehensive analysis of lncRNA-mRNA co-expression patterns identifies immune-associated lncRNA biomarkers in ovarian cancer malignant progression. *Sci Rep*. 2015;5:17683.
32. Bonome T, Lee JY, Park DC, et al. Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary. *Cancer Res*. 2005;65(22):10602-10612.
33. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003;4(2):249-264.
34. Zhang Z, Weaver DL, Olsen D, et al. Long non-coding RNA chromogenic in situ hybridisation signal pattern correlation with breast tumour pathology. *J Clin Pathol*. 2016;69(1):76-81.
35. Ferrari F, Bortoluzzi S, Coppe A, et al. Novel definition files for human GeneChips based on GeneAnnot. *BMC Bioinform*. 2007;8(1):1.
36. Chiu SL. Fuzzy model identification based on cluster estimation. *J Intell Fuzzy Syst*. 1994;2(3):267-278.
37. Park SH. Collinearity and optimal restrictions on regression parameters for estimating responses. *Technometrics*. 1981;23(3):289-295.
38. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *J Am Statist Assoc*. 1993;88(421):9-25.
39. Kotlyar M, Pastrello C, Pivetta F, et al. In silico prediction of physical interactions and characterization of interactome orphans. *Nat Methods*. 2015;12(1):79-84.
40. Vafae F, Krycer JR, Ma X, Burykin T, James DE, Kuncic Z. ORTI: An Open-access repository of transcriptional interactions for interrogating mammalian gene expression data. *PLoS ONE*. 2016;11(10):e0164535.
41. Girvan M, Newman ME. Community structure in social and biological networks. *Proc Natl Acad Sci U S A*. 2002;99(12):7821-7826.
42. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech-Theory E*. 2008;2008(10):P10008.
43. Clauset A, Newman ME, Moore C. Finding community structure in very large networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;70(6 Pt 2):066111.
44. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci U S A*. 2008;105(4):1118-1123.
45. Pons P, Latapy M. *Computing communities in large networks using random walks*. In Yolum P, Güngör T, Gürgeç F, Özturan C eds. Computer and Information Sciences - ISICIS 2005: 20th International Symposium, Istanbul, Turkey, October 26-28, 2005 Proceedings. Berlin, Heidelberg: Springer; 2005: 284-293.
46. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-W97.
47. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*. 2011;6(7):e21800.
48. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457.
49. Nurmik M, Ullmann P, Rodriguez F, Haan S, Letellier E. In search of definitions: Cancer-associated fibroblasts and their markers. *Int J Cancer*. 2020;146(4):895-905.

50. Signal B, Gloss BS, Dinger ME. Computational approaches for functional prediction and characterisation of long noncoding RNAs. *Trends Genet.* 2016;32(10):620-637.
51. Sun J, Shi H, Wang Z, et al. Inferring novel lncRNA-disease associations based on a random walk model of a lncRNA functional similarity network. *Mol Biosyst.* 2014;10(8):2074-2081.
52. Zhou C, York SR, Chen JY, et al. Long noncoding RNAs expressed in human hepatic stellate cells form networks with extracellular matrix proteins. *Genome Med.* 2016;8(1):31.
53. Guo X, Gao L, Liao Q, et al. Long non-coding RNAs function annotation: a global prediction method based on bi-colored networks. *Nucleic Acids Res.* 2013;41(2):e35.
54. Lorenz DM, Jeng A, Deem MW. The emergence of modularity in biological systems. *Phys Life Rev.* 2011;8(2):129-160.
55. Wei R, Lv M, Li F, et al. Human CAFs promote lymphangiogenesis in ovarian cancer via the Hh-VEGF-C signaling axis. *Oncotarget.* 2017;8(40):67315-67328.
56. Zhou Y, Xu X, Lv H, et al. The long noncoding RNA MALAT-1 is highly expressed in ovarian cancer and induces cell growth and migration. *PLoS ONE.* 2016;11(5):e0155250.
57. Chen ZJ, Zhang Z, Xie BB, Zhang HY. Clinical significance of up-regulated lncRNA NEAT1 in prognosis of ovarian cancer. *Eur Rev Med Pharmacol Sci.* 2016;20(16):3373-3377.
58. Wang X, Yang B, She Y, Ye Y. The lncRNA TP73-AS1 promotes ovarian cancer cell proliferation and metastasis via modulation of MMP2 and MMP9. *J Cell Biochem.* 2018;119(9):7790-7799.
59. Liu Y, Zhang M, Liang L, Li J, Chen YX. Over-expression of lncRNA DANCR is associated with advanced tumor progression and poor prognosis in patients with colorectal cancer. *Int J Clin Exp Pathol.* 2015;8(9):11480-11484.
60. Wu D, Zhuo L, Wang X. Metabolic reprogramming of carcinoma-associated fibroblasts and its impact on metabolic heterogeneity of tumors. *Semin Cell Dev Biol.* 2017;64:125-131.
61. Martinez-Outschoorn UE, Lisanti MP, Sotgia F. Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth. *Semin Cancer Biol.* 2014;25:47-60.
62. Yang L, Achreja A, Yeung TL, et al. Targeting stromal glutamine synthetase in tumors disrupts tumor microenvironment-regulated cancer cell growth. *Cell Metab.* 2016;24(5):685-700.
63. Capparelli C, Guido C, Whitaker-Menezes D, et al. Autophagy and senescence in cancer-associated fibroblasts metabolically supports tumor growth and metastasis via glycolysis and ketone production. *Cell Cycle.* 2012;11(12):2285-2302.
64. Wang Q, Xue L, Zhang X, Bu S, Zhu X, Lai D. Autophagy protects ovarian cancer-associated fibroblasts against oxidative stress. *Cell Cycle.* 2016;15(10):1376-1385.
65. Xiu YL, Sun KX, Chen X, et al. Upregulation of the lncRNA Meg3 induces autophagy to inhibit tumorigenesis and progression of epithelial ovarian carcinoma by regulating activity of ATG3. *Oncotarget.* 2017;8(19):31714-31725.
66. Hu J, Zhang L, Mei Z, et al. Interaction of E3 ubiquitin ligase MARCH7 with long noncoding RNA MALAT1 and autophagy-related protein ATG7 promotes autophagy and invasion in ovarian cancer. *Cell Physiol Biochem.* 2018;47(2):654-666.
67. Cao M, Zhong Q. Cilia in autophagy and cancer. *Cilia.* 2015;5:4.
68. Tam W. Identification and characterization of human BIC, a gene on chromosome 21 that encodes a noncoding RNA. *Gene.* 2001;274(1-2):157-167.
69. Elton TS, Selemo H, Elton SM, Parinandi NL. Regulation of the MIR155 host gene in physiological and pathological processes. *Gene.* 2013;532(1):1-12.
70. Wang W, Zhao Z, Yang F, et al. An immune-related lncRNA signature for patients with anaplastic gliomas. *J Neurooncol.* 2018;136(2):263-271.
71. Wu X, Wang Y, Yu T, et al. Blocking MIR155HG/miR-155 axis inhibits mesenchymal transition in glioma. *Neuro Oncol.* 2017;19(9):1195-1205.
72. Thiele JA, Hosek P, Kralovcova E, et al. lncRNAs in non-malignant tissue have prognostic value in colorectal cancer. *Int J Mol Sci.* 2018;19(9):2672.
73. Qin Y, Liu X, Pan L, Zhou R, Zhang X. Long noncoding RNA MIR155HG facilitates pancreatic cancer progression through negative regulation of miR-802. *J Cell Biochem.* 2019;120(10):17926-17934.
74. Wu W, Yu T, Wu Y, Tian W, Zhang J, Wang Y. The miR155HG/miR-185/ANXA2 loop contributes to glioblastoma growth and progression. *J Exp Clin Cancer Res.* 2019;38(1):133.
75. Santoiemma PP, Powell DJ Jr. Tumor infiltrating lymphocytes in ovarian cancer. *Cancer Biol Ther.* 2015;16(6):807-820.
76. Huffaker TB, Hu R, Runtsch MC, et al. Epistasis between microRNAs 155 and 146a during T cell-mediated antitumor immunity. *Cell Rep.* 2012;2(6):1697-1709.
77. Zonari E, Pucci F, Saini M, et al. A role for miR-155 in enabling tumor-infiltrating innate immune cells to mount effective antitumor responses in mice. *Blood.* 2013;122(2):243-252.
78. Gao X, Qiao Y, Han D, Zhang Y, Ma N. Enemy or partner: relationship between intronic microRNAs and their host genes. *IUBMB Life.* 2012;64(10):835-840.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Colvin EK, Howell VM, Mok SC, Samimi G, Vafaei F. Expression of long noncoding RNAs in cancer-associated fibroblasts linked to patient survival in ovarian cancer. *Cancer Sci.* 2020;111:1805-1817. <https://doi.org/10.1111/cas.14350>