Integrated analysis of the clinical consequence and associated gene expression of ALK in ALK-positive human cancers

Saifullah a,b, Toshifumi Tsukahara a,c,*

a Area of Bioscience and Biotechnology, School of Materials Science, Japan Advanced Institute of Science and Technology (JAIST), 1–1 Asahidai, Nomi City, Ishikawa, 923-1292, Japan
b Department of Molecular Therapy, National Institute of Neuroscience, National Center of Neurology and Psychiatry (NCNP), Kodaira City, Tokyo, 187-8502, Japan
c Division of Transdisciplinary Science, Japan Advanced Institute of Science and Technology (JAIST), 1–1 Asahidai, Nomi City, Ishikawa, 923-1292, Japan

ARTICLE INFO

Keywords:
ALK expression
Patient prognosis
LUAD
Co-expression
Cancers

ABSTRACT

Anaplastic lymphoma kinase (ALK) is a tyrosine kinase receptor that is genetically altered in several cancers, including NSCLC, melanoma, lymphoma, and other tumors. Although ALK is associated with various cancers, the relationship between ALK expression and patient prognosis in different cancers is poorly understood. Here, using multidimensional approaches, we revealed the correlation between ALK expression and the clinical outcomes of patients with LUAD, melanoma, OV, DLBC, AML, and BC. We analyzed ALK transcriptional expression, patient survival rate, genetic alteration, protein network, and gene and microRNA (miRNA) co-expression. Compared to that in normal tissues, higher ALK expression was found in LUAD, melanoma, and OV, which are associated with poor patient survival rates. In contrast, lower transcriptional expression was found to decrease the survival rate of patients with DLBC, AML, and BC. A total of 202 missense mutations, 17 truncating mutations, 7 fusions, and 3 in-frame mutations were identified. Further, 17 genes and 19 miRNAs were found to be exclusively co-expressed and echinoderm microtubule-associated protein-like 4 (EML4) was identified as the most positively correlated gene (log odds ratio >3). The gene ontology and signaling pathways of the genes co-expressed with ALK in these six cancers were also identified. Our findings offer a basis for ALK as a prognostic biomarker and therapeutic target in cancers, which will potentially contribute to precision oncology and assist clinicians in identifying suitable treatment options.

1. Introduction

Anaplastic lymphoma kinase (ALK) is a transmembrane tyrosine kinase receptor belonging to the insulin receptor superfamily [1]. Genetic aberrations of ALK, including gene fusions, translocations, or inversions with different gene partners, have been identified in different cancers, such as EML4-ALK oncoprotein in non-small cell lung cancer (NSCLC) [2, 3], nucleophosmin (NPM)-ALK fusion oncogene in anaplastic large cell lymphoma (ALCL) [4], truncating ALK in melanoma [5], and translocation of ALK in ovarian serous carcinoma [6]. ALK translocations also appear at low rates in diffuse large B-cell lymphoma (DLBC) [7], acute myeloid leukemia (AML) [8], breast cancer (BC) [9], and renal cell carcinomas [10]. The resultant ALK chimeraproteins comprise the C-terminal part of the entire intracellular tyrosine kinase domain and the N-terminal part of the fusion proteins, usually the di- or trimerization domain [11].

Once the fusion protein is activated by di- or trimerization, ALK plays a primary role in constitutive autophosphorylation, leading to the activation of downstream signaling and subsequent arrest of cell proliferation and growth. There are four main signaling pathways downstream of kinase-activated ALK: proto-oncogene protein p21/extracellular-signal-regulated kinase (RAS/ERK), phosphoinositol-3 kinase (PI3K)/Akt, Janus kinase/signal transducer and activator of transcription (JAK/STAT), and phospholipase Cγ (PLCγ) [12]. ALK also activates adapter and cellular proteins, including Src, IRS2, PTEN11, FAK, Shc-GRB2, GSK-3a, and FRS2, suggesting signaling as alternative pathways [13]. Furthermore, ALK is oncogenically activated by overexpression [14], point mutation [15], or truncation [16]. These discrete cellular functions and activation of downstream cascades contribute to the initiation and development of multiple malignancies.

Multi-omics data and clinical information have revolutionized the field of medicine and biology, enabling a comprehensive understanding
Figure 1. Expression profile of ALK across human normal and cancer tissues. A. Outline of ALK mRNA and protein expression in normal human tissues based on Human Protein Atlas normal tissue immunohistochemistry, Expression Atlas data, and RNA-seq expression data derived from Open Targets Genetics platform. B. RSEM (RNA-Seq by Expectation-Maximization) RNA-Seq expression profiles for each cancer and corresponding normal tissue were obtained from The Cancer Genome Atlas (TCGA) using the FireBrowse datasets. The boxes denote the median, and the 25th and 75th % dots symbolize outliers. The red boxes represent tumor tissues and the blue boxes represent corresponding normal tissues. The gray boxes signify that no normal samples exist for that disease cohort. Full form of each cancer type is presented in Table 1.
of disease genotypes and phenotypes for human health [17, 18]. One-dimensional omics data for cancer studies provide only limited information regarding the etiology of oncogenesis and cancer progression [19]. In contrast, multi-omics data can provide a greater understanding of the prognosis and predictive precision of disease phenotypes, thereby improving therapeutic and prevention measurements [20]. Prediction of cancer prognosis is an important factor of interest for clinicians, cancer patients, and healthcare professionals as it facilitates all types of decisions concerning patient care and clinical treatment [21]. Therefore, prognostic biomarkers have been employed to precisely select patient subgroups that would benefit from various treatment approaches.

In this study, we comprehensively analyzed ALK expression and its clinical significance in lung adenocarcinoma (LUAD), melanoma, ovarian carcinoma (OV), DLBC, AML, and BC patients using publicly available multi-platform datasets. We evaluated gene and mRNA expression patterns, patient prognosis, genetic alteration, protein–protein interaction network, gene co-expression, miRNA co-expression, gene ontology, and signaling pathways in human LUAD, melanoma, OV, DLBC, AML, and BC.

The combined data provide supporting evidence for the use of ALK as a prognostic biomarker and therapeutic target in ALK-positive cancers.

2. Results

2.1. Expression profile of ALK across human normal and cancer tissues

We observed ALK RNA and protein expression in various normal human tissues using the Open Target Genetics platform [22]. Both the RNA and protein expression levels of ALK were the highest in the brain. ALK expression was higher in the skin, reproductive, and lung tissues than in the blood and breast tissues (Figure 1A). The FireBrowse dataset [23] was then used to examine ALK transcript expression in multiple human cancers. The results revealed an almost similar pattern of ALK gene expression in the corresponding cancer tissues (Figure 1B and Table 1), which primarily suggests that ALK is involved in the tumorigenesis of the skin, ovary, lungs, blood, and breast cancers.

2.2. Significant variation in ALK expression in different cancer types

Next, the interactive tool Oncomine [24] was used to assess the mRNA expression of ALK in lung, skin, ovary, blood, and breast cancers. As shown in Figure 2, ALK mRNA expression was significantly higher in the lung (LUSC and LUAD), skin (CM), and ovary (OSC) and lower in the blood (DLBC and AML) and invasive breast cancer (IBC) than that observed in healthy controls. The results provide strong evidence of the upregulation of ALK in LUSC, LUAD, CM, and OSC, and its down-regulation in DLBC, AML, and IBC tissues compared to that in their corresponding normal tissues.

2.3. Prognostic investigation of ALK mRNA expression in patients with cancer

The correlation between the level of ALK expression and the survival rate of patients with lung, skin, ovarian, blood, and breast cancers was determined using the PrognoScan dataset [25] with significant Cox p-values (<0.05). Based on the analysis, ALK overexpression was negatively correlated to the survival rate of patients with LUAD (Jaco−b−00182−UM dataset), melanoma (GSE19234), and ovarian cancers (GSE9891) with increased risk association (hazard ratio [HR] > 1) (Figure 3A–C and I, Table 2). In contrast, the differential expression of ALK was associated with a higher death rate in DLBCL (E-TABM-346), AML (GSE8970), and BC (GSE9893) with lower hazard ratios (HR < 0.5) (Figure 3D–F and I). The dataset GSE8894 also revealed that lower expression of ALK decreases the survival rate of patients with NSCLC, whereas higher expression leads to lower relapse-free survival with higher risk (HR > 1) in breast cancers, as shown in Figure 3G–H and Table 2. Therefore, our findings suggest that the deregulated expression of ALK could lead to poor patient prognosis.

2.4. Evaluation of ALK alteration in associated cancers

Data on genetic alterations in the ALK gene and their relevance and expression in different cancers were obtained from the eBioPortal server [26]. Based on the data, 234 alterations were identified from 1–1620 amino acid sequences (Figure 4A and Supplementary Table S1). Of them eight are driver mutations and the remaining mutations are variant of uncertain significance (VUS). The most common type was missense mutations (202 mutations), followed by 17 truncating mutations (including nonsense, frameshift deletion, frameshift insertion, and splicing), 7 gene fusions, 5 splicing, and 3 in-frame mutations. The alteration frequency was the highest in melanoma (15.9% in 444 cases) and the lowest in AML (<1% in 200 cases) (Figure 4B). No gene alterations were found in DLBC. As depicted in Figure 4C, profiling of
mutated RNA expression shows that missense mutations are present in each cancer type (highest in melanoma), seven oncofusions, nine truncating mutations, two splicing mutations, and one in-frame mutation. In contrast, the frequency of copy number gain was the highest, which was distributed in each of the cancers, followed by copy number diploid, amplification, and deep deletion. Consequently, these results suggest that the overexpression or downregulation of ALK in LUAD, skin melanoma, ovarian cancer, AML, DLBC, and BC, respectively, is correlated with mutations and copy number alterations.

2.5. ALK protein network analysis and clinical significance in cancers

To understand the significance of the diseases and predict the association between genotype and phenotype, GeneMANIA and STRING servers using Cytoscape v3.8.2 software were used to construct a protein network. Based on the results, the predicted 20 protein partners from the GeneMANIA server were EPHB3, MDK, AMELY, PTPRB, PTPRZ1, SEC16B, EPHA1, PTPRG, SEC16A, JAK3, MMP13, FBXO24, PTN, ZC3HC1, PDLIM3, MYBPC2, PTPRJ, KRT74, MTIF2, and SHC3. The genetic alteration frequency of ALK and the 30 interacting protein partners in the six cancer types were subsequently analyzed using cBioPortal. The highest alteration frequency was observed in skin melanoma (70%), followed by LUAD (65%), IBC (58%), and OC (57%), while the lowest was observed in AML (36%). These alterations in 17 mutual proteins decreased the patient’s disease-free survival compared to the unaltered group in all six cancer types. Such findings indicate that the identified ALK and its protein partners might be associated with LUAD, SM, OC, DLBC, AML, and BC. Furthermore, genetic alterations in these proteins can affect the clinical outcomes.

Figure 2. ALK expression pattern in various cancer types. A–F. ALK expression in six cancer types derived from the Oncomine cancer microarray database. The left box plot represents ALK expression in normal tissue, while the box on the right represents cancer tissue. Statistically significant differences between normal and cancerous tissues are indicated by p < 0.05. LUSC, lung squamous cell carcinoma; LUAD, lung adenocarcinoma; CM, cutaneous melanoma; OSC, ovarian serous carcinoma CB, Centroblast; MBL, Memory B-lymphocyte; NCBL, Naive Pre-germinal Center B-lymphocyte; P cell, Plasma Cell; SCFC, Small Cleaved Follicle Center Cell; DLBCL, diffuse large B-cell lymphoma; CD-34 PP, CD34-Positive Peripheral Blood cell; AML, acute myeloid leukemia; and IBC, invasive breast carcinoma. ALK was found to be substantially upregulated in lung, skin, and ovarian cancers, and considerably downregulated in blood and breast cancers. The sample numbers are indicated in parentheses.

1 EPHB3 (Ephrin type-B receptor 3), MDK (Midkine), AMELY (Amelogenin Y-Linked), PTPRB (Receptor-type tyrosine-protein phosphatase beta), PTPRZ1 (Receptor-type tyrosine-protein phosphatase zeta), SEC16B (SEC16 Homolog B, Endoplasmic Reticulum Export Factor), EPHA1 (Ephrin Type-A Receptor 1), PTPRG (Receptor-Type Tyrosine-Protein Phosphatase Gamma), SEC16A (SEC16 Homolog A, Endoplasmic Reticulum Export Factor), JAK3 (Janus Kinase 3), MMP13 (Matrix Metallopeptidase 13), FBXO24 (F-Box Protein 24), PTN (Pleiotrophin), ZC3HC1 (Zinc Finger C3HC-Type Containing 1), PDLIM3 (PDZ And LIM Domain 3), MYBPC2 (Myosin Binding Protein C2), PTPRJ (Protein Tyrosine Phosphatase Receptor Type J), KRT74 (Keratin 74), MTIF2 (Mitochondrial Translational Initiation Factor 2), and SHC3 (SHC Adaptor Protein 3).

2 EML4 (Echinoderm Microtubule-Associated Protein-Like 4), FRS3 (Fibroblast Growth Factor Receptor Substrate 3), PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha), TNFRSF8 (Tumor Necrosis Factor Receptor Superfamily Member 8), JAK3 (Janus Kinase 3), PCLG1 (Phospholipase C Gamma 1), KRAS (Kirsten Rat Sarcoma Viral Oncogene Homolog), NPM1 (Nucleophosmin 1), HRAS (Harvey Rat Sarcoma Viral Oncogene Homolog), and JAK2 (Janus Kinase 2).
Figure 3. ALK expression and clinical prognosis in six different cancers retrieved from PrognoScan microarray cancer database. Kaplan-Meier patient survival estimate of A. lung adenocarcinoma, B. melanoma, C. ovarian cancer, D–E. blood cancer: DLBCL and AML, F. breast cancer for ALK expression, G. NSCLC, and H. breast cancer for ALK expression. The survival curve was determined as the threshold of the Cox p-value < 0.05 and p-value < 0.01. The red line denotes high expression and the blue line denotes low expression. The dotted line represents the maximum and minimum values of the average survival. HR, hazard ratio; CI, confidence interval; and n, number of patients. I. The statistically significant hazard ratio for the six different cancers was determined from Figure A–F and expressed as a forest plot.
2.6. Profiling of genes and miRNA co-occurring and co-expressed with ALK

We identified mutually exclusive co-occurrence genes from the 30 ALK protein partners identified using cBioPortal [26]. A total of 17 genes were found to exclusively co-occur with ALK based on the q-values (<0.05), which are listed in Table 3. EML4 was identified as the most mutually exclusive gene (odds ratio >3, q-value < 0.001) among all genes, whereas ZC3HC1 had the lowest exclusivity (odds ratio = 1.33, q-value = 0.036). Genetic alterations were also found to be the highest in EPHB3 (11%).

To assess co-expression, we collected the fold-change values of ALK and identified 17 genes from the microarray datasets of the Oncomine server [24]. Expression analysis revealed that all genes were either overexpressed or downregulated in LUAD, CM, OSC, DLBC, AML, and BC, as depicted in the heatmap (Figure 6A). Notably, MTIF2 and EPHB3 were overexpressed in these six cancers, whereas JAK2 and SHC3 were downregulated in all cancers, except DLBC and AML, respectively. The results indicate that ALK and the identified genes were associated with these cancers, either positively or negatively. To determine the impact of each gene expression on the expression of another gene, we used a similarity matrix analysis of the genes with their co-expression values using Morpheus [27]. As shown in Figure 6B, the expression of one gene affected that of another. This finding suggests that the identified co-expressed genes of ALK may be associated with the progression of the six cancer types included in our study.

We determined whether the miRNA were co-expressed with ALK and 17 common genes using the Enrichr server [28] based on p-value ranking (<0.05). Based on the analysis, 19 of the 382 substantially co-expressed miRNAs in humans were identified, as shown in Figure 7 (green node). These findings indicate that the identified miRNAs might contribute to LUAD, SM, OC, DLBC, AML, and BC development, along with ALK and common genes.

2.7. Gene ontologies and signaling pathway elucidation of ALK and its correlated genes

Based on the expression profile of ALK and its mutual genes, we determined gene ontological features and enrichment pathways using the Enrichr server [28] for integrative analysis of the progression of LUAD, SM, OC, DLBC, AML, and BC in humans. The resulting gene ontology (GO) terms mainly included the transmembrane receptor protein tyrosine kinase signaling pathway, protein autophosphorylation, negative regulation of cell communication (Figure 8A), protein kinase binding, protein tyrosine kinase activity, MAPK activity, growth factor receptor binding (Figure 8B), an integral component of the plasma membrane, specific granule membrane, and endosome lumen activity (Figure 8C). Finally, KEGG 2019 suggested pathways related to NSCLC, Th1, Th2, and T17 cell differentiation, pathways in cancer, axon guidance, chemokine and prolactin signaling, adherens junction, glioma, and ErbB signaling (Figure 8F). Therefore, these pathways might be involved in the tumorigenesis of LUAD, SM, OC, DLBC, AML, and BC.

3. Discussion

In this study, by using the Open Targets Genetics portal, we found that ALK is highly expressed in lung, skin, and ovary normal tissues compared with that in blood and breast tissues (Figure 1A). Further, by using FireBrowse and Oncomine, we revealed that ALK expression at the mRNA and protein levels had similar expression patterns in their respective cancer tissues, such as upregulation in LUAD, melanoma, OC and downregulation in DLBC, AML, and BC (Figure 1B and Figure 2). These findings indicate an association between ALK expression and the progression of these six cancers. This finding was consistent with the biological fact that ALK expression is upregulated in lung cancer [2, 29, 30], melanoma [31], and OC [6], and downregulated in DLBC [32], AML [33], and BC [9] tissue samples.

Prediction of a disease phenotype based on specific gene expression substantially contributes to patient care with different treatment strategies. Therefore, it is necessary to have a strong prognostic index that facilitates the treatment of LUAD, melanoma, OC, DLBC, AML, and BC. Based on the prognostic analysis results, we found that the overexpression of ALK positively correlates with the progression of LUAD,
melanoma, and OC with higher hazard risk (HR > 1), whereas the downregulation of ALK was negatively correlated with the survival rate of patients with DLBC, AML, and BC with relatively lower hazard risk (Figure 3). Previous studies on LUAD and melanoma revealed that EML4-ALK rearrangement is associated with a poorer disease-free survival rate than that for ALK-negative cancers [34, 35]. However, patients with ALK rearrangement have a better prognosis for ovarian cancer than those without ALK rearrangement [36]. Our findings are consistent with a previous study on BC showing that patients with ALK protein overexpression exhibited poor overall survival or relapse-free survival compared to those with low ALK expression [37] (Figure 3G–H). These studies, except the study on BC, presented the patient survival rate based on ALK-positive or ALK-negative samples, whereas our study revealed prognosis based on ALK expression levels. To date, no prognostic studies on DLBC and AML based on ALK expression patterns have been found. Thus, it is worthwhile to conduct further studies on the overexpression or downregulation of ALK for long-term follow-up and a detailed molecular analysis to identify ALK as a prognostic biomarker for patients with LUAD, melanoma, OC, DLBC, AML, and BC.

Figure 4. Data on genetic alteration of ALK and consequent cancers in patients derived from TCGA PanCancer Atlas database through cBioPortal. A. Schematic of the ALK protein and its associated functional domains with the physical location of mutations. A total of 234 alterations were found from 1–1620 amino acid sequences, where 8 are driver and 226 VUS mutations respectively. The highest alteration type was a missense mutation (202) and the lowest was an in-frame mutation (3). B. The alteration frequencies of ALK and their associated cancers. The Y-axis shows the alteration frequency percentage, and the X-axis represents the type of cancers resulting from the alterations, including mutation and copy number alterations (CAN). The highest alteration frequency was found in melanoma (gene-altered 15.9% in 444 cases). C. RNA seq v2 results of genetically altered ALK mRNA expression in 12 cancer studies obtained from the cBioPortal web. The mutations included missense mutations in all cancer types (green dots), seven oncofusion in LUAD and melanoma (violet dots), nine truncating (gray dots), three slicing (light brown dots), and one in-frame mutation (deep brown dot). The frequency of copy number gain was the highest and distributed in all cancer types, followed by copy number diploid, shallow deletion, amplification, and deep deletion as indicated in different colors.
Figure 5. Protein network of ALK and their clinical significance in six cancers. A–B. Predicted functional protein partners of ALK generated by considering physical interaction, co-expression, co-localization, genetic interactions, and shared protein domains from GeneMANIA and STRING web servers using Cytoscape_v3.8.2 software. C. The genetic alteration frequency of 30 gene signatures in panels A–B was generated for lung, melanoma, ovarian, DLBC, AML, and breast cancers using the cBioPortal platform. The highest alteration percentage was reported in lung cancer, followed by melanoma and breast cancer, while the lowest was in AML. D. Kaplan–Meier patient survival estimation of patients with genetically altered and unaltered ALK and 17 partner genes (Table 3) in six types of cancers generated from cBioPortal. Patients with altered forms of ALK and 17 correlated genes showed significantly less survival than those with unaltered forms.
The divergent ALK expression in the six cancers might be attributed to the many mutations and copy number alterations which were observed in this study (Figure 4 and Supplementary table S1). Cancer development is a chain of histopathological processes that are influenced by four important factors—somatically acquired genetic, transcriptomic, proteomic, and epigenetic changes [38]. Among our queried patients (N = 5939), ALK was altered in 4% of patients (248), with a somatic mutation frequency of 3.4%. Herein, missense, truncating, fusion, and in-frame mutations, splicing, copy number gain, amplification, deep deletion, shallow deletion, and diploid were identified (Figure 4). Several studies have also revealed genetic alterations, including mutations and fusion in ALK [2, 4, 39]. In contrast to our findings, Chang et al. found three recurrent hotspots (R1275Q/*/L, F1245C/L/V, and F1174 L/V) in the tyrosine kinase domain and one (R395 H/C) in the MAM domain in a population-scale cohort of different tumor samples [40]. However, we did not identify any cancer hotspots in the queried TCGA samples.

Protein–protein interactions have been used for the analysis of biological processes, disease progression, and downstream signaling, and have been targeted for drug development [41]. In the present study, we identified 30 proteins (10 from STRING and 20 from GeneMANIA) using Cytoscape software. These 30 protein signatures showed an alteration frequency of up to 65% in melanoma and lung cancers. The PIK3CA had the highest alteration frequency (22%), and HRAS was the least frequently altered gene (1.8%) (Figure 9). Similarly, Mills et al. revealed that PIK3CA has a mutational frequency of 13% in solid tumors [42]. In another study, this gene was revealed to be amplified by 33–42% in lung cancer and 15–27% in ovarian cancer [43]. Hobbs et al. also reported that HRAS is the least frequent (4%) protein relative to other isoforms, such as KRAS (85%), in human cancers [44]. These results suggest the involvement of PIK3CA, HRAS, and other partners in various cancers.

Using the cbioPortal datasets, we identified 17 genes that exclusively co-occurred with 30 genes (Table 3). Based on our analysis, EML4 was identified as the most correlated gene. Soda et al. (2007) first reported the fusion of EML4 with ALK, an oncogenic driver leading to NSCLC, which was later identified in colorectal cancer and BC [2, 45]. Other genes, such as EPHB3, PTPRJ, PTPRB, and PLCG1 in lung cancer [46, 47, 48, 49]; JAK3 in melanoma; EPHA1 in OC [50]; TNFRSF8 (CD30) in DLBC [51]; and PLCG1 and PTPRG in BC [49, 52], have been reported. However, no relationship was found for 6 of the 17 genes, including KRIT4, MTIF2, SHC3, FBXO24, MYBPC2, and ZC3H11 in the six cancers included in this study. Nonetheless, co-expression with ALK was confirmed in our co-expression analysis (Figure 6). ALK was found to be upregulated in LUAD, CM, and OSC and downregulated in DLBC, AML, and BC. This observation is consistent with those of previous studies [31, 32, 33]. Among the identified miRNAs, hsa-miR-1228, hsa-miR-600, and hsa-miR-542-5p were associated with lung cancer [15, 53, 54]; hsa-miR-3158 and hsa-miR-3145 were found in melanoma [55]; and hsa-miR-3619-3p, hsa-miR-4738, and hsa-miR-1298-5p were found in BC [56, 57, 58]. To date, 11 other miRNAs, namely hsa-miR-27a, hsa-miR-6727-5p, hsa-miR-597-5p, hsa-miR-624-5p, hsa-miR-555, hsa-miR-376a-2-5p, hsa-miR-6824-3p, hsa-miR-1255b-2-3p, hsa-miR-3912-5p, and hsa-miR-6854-3p, have not been identified in the queried cancers.

We evaluated the GO enrichment pathways of biological processes, molecular functions, and cellular components for the functional analysis of ALK and its correlated genes. Importantly, receptor tyrosine kinase (RTK) signaling activity was the most relevant in all three GO enrichment pathways (ALK and its correlated genes). Further studies are required based on gene silencing and overexpression to identify concrete pathways involved in the progression of ALK-mediated melanoma, OC, DLBC, and AML.

In summary, this study revealed the possible relationship between ALK expression and the prognosis of patients with LUAD, melanoma, OC, DLBC, AML, and BC using publicly available multiplatform bioinformatics datasets. Based on our data, the genotypic expression of ALK was substantially correlated with the clinical phenotype in the six cancers. Our study may help clinicians with precision clinical care and researchers in developing cancer therapies for ALK-positive cancers. Additional therapies targeting ALK in patients with ALK-positive melanoma, OC, DLBC, AML, and BC may increase the survival rate. However, further experimental studies are required to confirm whether ALK is a prognostic biomarker.

### 4.1. Expression profile of ALK across human normal and cancer tissues

ALK mRNA and protein expression data in normal human tissues were obtained from the Human Protein Atlas, normal tissue immunohistochemistry, Expression Atlas data, and RNA-seq expression data using the Open Targets Genetics Portal (https://genetics.opentargets.org) [22]. Gene expression patterns across normal and cancerous tissues were determined using the FireBrowse (http://firebrowse.org/) datasets [23]. The ALK query in the FireBrowse database was performed using the default settings.

**Table 3.** Significant co-occurring protein partners of the ALK gene signature were obtained using cbioPortal for cancer genomics.

| Protein partner | Alteration (%) | Log2 Odds Ratio | q-value Tendency |
|-----------------|----------------|----------------|-----------------|
| EML4            | 2.1            | -3             | <0.001          | Co-occurrence |
| KRT74           | 1.8            | 2.718          | <0.001          |
| SEC16A          | 3              | 2.611          | <0.001          |
| TNFRSF8         | 2.5            | 2.441          | <0.001          |
| MTIF2           | 2.1            | 2.357          | <0.001          |
| PTPRJ           | 2.6            | 2.312          | <0.001          |
| PTPRB           | 9              | 2.185          | <0.001          |
| PTPRG           | 4              | 2.173          | <0.001          |
| PLCG1           | 2.6            | 2.111          | <0.001          |
| EPHA1           | 4              | 1.848          | <0.001          |
| JAK2            | 4              | 1.812          | <0.001          |
| SHC3            | 1.8            | 1.722          | 0.02            |
| FBXO24          | 3              | 1.674          | 0.002           |
| JAK3            | 4              | 1.655          | 0.002           |
| MYBPC2          | 2.8            | 1.599          | 0.008           |
| EPHB3           | 11             | 1.413          | <0.001          |
| ZC3H11          | 2.6            | 1.336          | 0.036           |

Saifullah, T. Tsukahara. Heliyon 8 (2022) e09878
Figure 6. Co-expression and hierarchical clustering analyses. A. Co-expression analysis of ALK and its 17 mutually associated functional protein partner genes in six cancer types. The heat map was generated from the log2 fold-change expression value of ALK and mutually exclusive protein partner genes retrieved from the cancer microarray database (Oncomine). B. Hierarchical clustering and similarity matrix analysis of ALK and the expression of correlated genes in six cancer types. The heat map was generated from the log2 fold-change expression value of ALK and common genes retrieved from Oncomine. The outcome of the correlation was visualized using the similarity matrix and hierarchical clustering tools in the Morpheus server.
4.2. Significant ALK expression variation in different cancer types

The expression levels of the ALK gene in lung cancer (LUAD, LUSC), skin cancer (CM), ovarian cancer (OSC), blood cancer (DLBC, AML), and BC and normal tissue counterparts were assessed using the interactive Oncomine (https://www.oncomine.org) dataset [24]. The Oncomine server covers microarray data from 86,733 tumors and 12,764 normal tissues. The following parameters were considered to generate a graph from Oncomine data: p-value, <0.05; fold change, 2; gene rank, top 10%.

4.3. Prognostic investigation of ALK mRNA expression in patients with cancer

The biological relationship between ALK expression and prognosis of patients with different cancers was assessed using PrognoScan cancer microarray datasets (http://www.prognoscan.org) [25] with a Cox p-value of <0.05. PrognoScan provides an effective platform for assessing potential cancer biomarkers and therapeutic targets.

4.4. Evaluation of ALK alteration and associated cancers

TCGA PanCancer Atlas datasets (10,953 patients/10,967 samples) were used to evaluate the genetic alterations of ALK in associated cancers, and cBioPortal cancer genomics data (https://www.cbioportal.org) were used to evaluate the RNA-seq-based mRNA expression levels [26].

4.5. ALK protein network analysis and clinical significance in cancers

Protein–protein interaction analysis with the target gene was performed using GeneMANIA and STRING platforms with Cytoscape software version 3.5.2. The predicted 30 proteins were then queried using
Figure 8. Gene ontology and pathway analyses of ALK and its co-related functional protein partners. A. Biological process, B. Molecular function, C. Cellular components, D. Reactome, E. PANTHER, and F. KEGG pathways of ALK and 18 partner proteins were obtained from the Enrichr web tool. The bar chart shows the top 10 enriched terms in the chosen library, along with their corresponding p-values. Colored bars correspond to terms with significant p-values (<0.05). * indicates statistically significant values with p < 0.05.
the cBioPortal platform to identify genetic alterations and clinical significance in lung, skin, ovary, blood, and breast cancers.

4.6. Profiling of genes and miRNAs co-occurring and co-expressed with ALK

Genes exclusively co-occurring with the target ALK gene were analyzed using the cBioPortal platform based on the q-value ranking (<0.05). The co-expression value of the co-occurring genes was obtained using the Oncomine microarray datasets. The co-expression heatmap was generated using GraphPad Prism software version 8 (San Diego, CA, USA). Based on the co-expression values, the influence of one gene on the expression of another gene was evaluated using a similarity matrix and hierarchical clustering tools via the Morpheus server [27]. The miRNAs co-expressed with ALK and common genes were obtained from the Enrichr web portal (https://maayanlab.cloud/Enrichr/) [28] and analyzed using Cytoscape_v3.8.2 software.

4.7. Elucidation of gene ontologies and signaling pathways of ALK and its correlated genes

The gene ontologies and signaling pathways of ALK with co-occurring genes and relevant bar plots were derived from the Enrichr platform [28]. Enrichr is an integrative web-based gene-list enrichment analysis tool that compares different genomic datasets. The input genes were classified into biological processes, molecular functions, and cellular components according to GO terms. Similarly, signaling pathways were defined using the Reactome 2016, PANTHER 2016, and KEGG 2019. A Cox p-value of <0.05 was considered statistically significant for both GO and signaling pathway bar plots.

Figure 9. Oncoprint of ALK and common genes. The genetic alterations of ALK and 17 common genes in six cancers were evaluated using cBioPortal. The highest alteration was observed in PIK3CA (22%), followed by EPHB3 (11%), KRAS, and PTPRB (9%), while the lowest was observed in HRAS (1.7%).
5. Conclusion

In this study, to determine ALK as a potential prognostic biomarker in ALK-positive LUAD, melanoma, OC, DLBC, AML, and BC progression, we assessed ALK expression, genetic mutations, protein–protein interaction networks, correlated genes, and miRNAs, and prognostic values using various publicly available bioinformatics datasets. We conclude that both the upregulation of ALK in LUAD, melanoma, and OV, or downregulation in DLBC, AML, and BC progression negatively correlated with patient survival. Furthermore, we revealed the possible gene ontology and signaling pathways associated with ALK and its expression in the progression of these six cancers. These pathways may serve as potential targets for inhibiting the development of cancer. Collectively, these findings suggest that ALK could be an effective prognostic biomarker and a potential therapeutic target to control ALK-positive cancers. Further studies using in vitro experiments are required to validate our results.

Declarations

Author contribution statement

Saifullah: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Toshifumi Tsukahara: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Dr. Toshifumi Tsukahara was supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science [21H02067 and 18K19288].

Saifullah was supported by the scholarship from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan.
systematic identification of trait-associated genes using large-scale genetics and functional genomics, Nucleic Acids Res. 49 (2021) D1311–D1320.

23. Fire et al., Broad Institute, in: Fire, 2016, D16_01_28 Run 2016, Broad Institute of MIT and Harvard, 2016, pp. 2498–2504.

24. D.R. Rhodes, S. Kalyana-Sundaram, V. Mahavino, R. Varambally, Y. Ju, B.B. Briggs, T.R. Barrette, M.J. Annet, C. Kincead-Beal, P. Kulkarni, et al., Oncomine 3.0: genes, pathways, and networks in the collection of 18,000 cancer expression profiles, Neoplasia 9 (2007) 166–180.

25. H. Mizuno, K. Kitada, K. Nakai, A. Sari, PrognoScan: a new database for meta- analysis of the prognostic value of genes, BMC Med. Genom. 2 (2009) 18.

26. J. Gao, B.A. Akty, U. Dogruyuz, G. Desmedt, B. Grost, S.O. Nørner, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, et al., Integrative analysis of complex cancer genomes and clinical profiles using the cBioPortal, Sci. Signal. 6 (2013) pl1. pl1.

27. Morphosys: https://software.broadinstitute.org.

28. M.V. Kuleshov, M.R. Jones, A.D. Roilloud, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Research 44 (2016) W90–W97.

Sai phẫuśl, M. Sakari, T. Suzuki, S. Tano, T. Tkashurak, Effective RNA knockdown using CRISPR-cas13a and molecular targeting of the EML4-ALK transcript in H3122 lung cancer cells, Int. J. Mol. Sci. 21 (2020) 8904.

Sai phẫuśl, M. Sakari, T. Suzuki, S. Tano, T. Tkashurak, P28-11 the CRISPR-Cas13a system underlies a potential therapeutic strategy for EML4-ALK-positive lung cancer cells, Ann. Oncol. 32 (2021) S347.

29. K.J. BUSAN, R.E. VLHN, T. LAM, J.A. BUSAN, T.J. HOLMANN, R.P.M. SAW, D.C. COIT, R.A. SOOLEY, T. WIENNER, Primary and metastatic cutaneous melanomas express ALK through alternative transcriptional initiation, Am. J. Surg. Pathol. 40 (2016) 786–795.

30. V. Brune, E. Tiecci, I. Pfeifer, C. Doring, E. Sicker, C.J.M. van Noesel, W. Klapper, B. Falani, A. von Heydebreck, D. Metzler, et al., Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis, J. Exp. Med. 205 (2008) 2251–2266.

31. P.J.M. Valk, R.G.W. Verhaak, M.A. Beijen, C.A.J. Erpelinck, S.B.v.W. van Doorn-Khorzon, J.M. Hoer, B.B. Vermeerlo, M.J. Moonhoute, P.J. van der Spek, B. Lonergan, et al., Prognostically useful gene-expression profiles in acute myeloid leukemia, N. Engl. J. Med. 350 (2004) 1617–1628.

32. J. Shi, W. Gu, Y. Zhao, J. Zhu, G. Jiang, M. Bao, J. Shi, Clinicopathological and prognostic significance of ALK rearrangements in patients with surgically resected lung adenocarcinoma: a propensity score matching study, Cancer Manag. Res. 12 (2020) 589–598.

33. S. CAO, V.E. Nambudiri, Anaplastic lymphoma kinase in cutaneous malignancies, J. Invest. Dermatol. 139 (2017) 123.

34. R. BI, Q. XI, Z. XU, C. WAI, J. SHANG, H. TAN, B. CHANG, et al., ALK rearrangement: a high-frequency alteration in ovarian metastasis from lung adenocarcinoma, Diagn. Pathol. 14 (2019) 96.

35. H. OKAYAMA, T. KOHNO, Y. ISHI, Y. SHIMADA, K. SHIRAZI, R. IRAWATI, K. TAKAEDA, J. KAWADA, S. YAMAMOTO, et al., Identification of the EGFR/KRAS/ALK fusion target in lung adenocarcinomas, Cancer Res. 72 (2012) 10021–10030.

36. K.J. BUSAN, R.E. VLHN, T. LAM, J.A. BUSAN, T.J. HOLMANN, R.P.M. SAW, D.C. COIT, R.A. SOOLEY, T. WIENNER, Primary and metastatic cutaneous melanomas express ALK through alternative transcriptional initiation, Am. J. Surg. Pathol. 40 (2016) 786–795.

37. V. Brune, E. Tiecci, I. Pfeifer, C. Doring, E. Sicker, C.J.M. van Noesel, W. Klapper, B. Falani, A. von Heydebreck, D. Metzler, et al., Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis, J. Exp. Med. 205 (2008) 2251–2266.

38. P.J.M. Valk, R.G.W. Verhaak, M.A. Beijen, C.A.J. Erpelinck, S.B.v.W. van Doorn-Khorzon, J.M. Hoer, B.B. Vermeerlo, M.J. Moonhoute, P.J. van der Spek, B. Lonergan, et al., Prognostically useful gene-expression profiles in acute myeloid leukemia, N. Engl. J. Med. 350 (2004) 1617–1628.

39. J. Shi, W. Gu, Y. Zhao, J. Zhu, G. Jiang, M. Bao, J. Shi, Clinicopathological and prognostic significance of ALK rearrangement in patients with surgically resected lung adenocarcinoma: a propensity score matching study, Cancer Manag. Res. 12 (2020) 589–598.

40. S. CAO, V.E. Nambudiri, Anaplastic lymphoma kinase in cutaneous malignancies, J. Invest. Dermatol. 139 (2017) 123.

41. R. BI, Q. XI, Z. XU, C. WAI, J. SHANG, H. TAN, B. CHANG, et al., ALK rearrangement: a high-frequency alteration in ovarian metastasis from lung adenocarcinoma, Diagn. Pathol. 14 (2019) 96.

42. A.K. Siraj, S. Beg, J. Jehan, S. Prabhakaran, M. Ahmed, R. Hussain, F. Al-Dayel, A. Tullab, D. Ahmad, K.S. Al-Kaurya, ALK alteration is a frequent event in aggressive breast cancers, Breast Cancer Res. 17 (2015) 127.

43. H. OKAYAMA, T. KOHNO, Y. ISHI, Y. SHIMADA, K. SHIRAZI, R. IRAWATI, K. TAKAEDA, J. KAWADA, S. YAMAMOTO, et al., Identification of the EGFR/KRAS/ALK fusion target in lung adenocarcinomas, Cancer Res. 72 (2012) 10021–10030.

44. K.J. BUSAN, R.E. VLHN, T. LAM, J.A. BUSAN, T.J. HOLMANN, R.P.M. SAW, D.C. COIT, R.A. SOOLEY, T. WIENNER, Primary and metastatic cutaneous melanomas express ALK through alternative transcriptional initiation, Am. J. Surg. Pathol. 40 (2016) 786–795.

45. V. Brune, E. Tiecci, I. Pfeifer, C. Doring, E. Sicker, C.J.M. van Noesel, W. Klapper, B. Falani, A. von Heydebreck, D. Metzler, et al., Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis, J. Exp. Med. 205 (2008) 2251–2266.

46. P.J.M. Valk, R.G.W. Verhaak, M.A. Beijen, C.A.J. Erpelinck, S.B.v.W. van Doorn-Khorzon, J.M. Hoer, B.B. Vermeerlo, M.J. Moonhoute, P.J. van der Spek, B. Lonergan, et al., Prognostically useful gene-expression profiles in acute myeloid leukemia, N. Engl. J. Med. 350 (2004) 1617–1628.

47. J. Shi, W. Gu, Y. Zhao, J. Zhu, G. Jiang, M. Bao, J. Shi, Clinicopathological and prognostic significance of ALK rearrangement in patients with surgically resected lung adenocarcinoma: a propensity score matching study, Cancer Manag. Res. 12 (2020) 589–598.

48. S. CAO, V.E. Nambudiri, Anaplastic lymphoma kinase in cutaneous malignancies, J. Invest. Dermatol. 139 (2017) 123.

49. R. BI, Q. XI, Z. XU, C. WAI, J. SHANG, H. TAN, B. CHANG, et al., ALK rearrangement: a high-frequency alteration in ovarian metastasis from lung adenocarcinoma, Diagn. Pathol. 14 (2019) 96.

50. A.K. Siraj, S. Beg, J. Jehan, S. Prabhakaran, M. Ahmed, R. Hussain, F. Al-Dayel, A. Tullab, D. Ahmad, K.S. Al-Kaurya, ALK alteration is a frequent event in aggressive breast cancers, Breast Cancer Res. 17 (2015) 127.

51. H. OKAYAMA, T. KOHNO, Y. ISHI, Y. SHIMADA, K. SHIRAZI, R. IRAWATI, K. TAKAEDA, J. KAWADA, S. YAMAMOTO, et al., Identification of the EGFR/KRAS/ALK fusion target in lung adenocarcinomas, Cancer Res. 72 (2012) 10021–10030.
Clinicopathologic features and prognostic significance of CD30 expression in de novo diffuse large B-cell lymphoma (DLBCL): results in a homogeneous series from a single institution, Biomarkers 25 (2020) 69–75.

Receptor-type protein tyrosine phosphatases in cancer, Chin. J. Cancer 34 (2015) 61–69.

Targeting anaplastic lymphoma kinase in neuroblastoma, APMIS 127 (2019) 288–302.

Circular RNA circPSPK1A promotes non-small cell lung cancer proliferation and metastasis through miR-600/HIF-1α regulation, J. Cell. Biochem. 120 (2019) 19019–19030.

Identification of chemoresistance-associated miRNAs in breast cancer, Cancer Manag. Res. 10 (2018) 4747–4757.

Identification of new MicroRNAs in paired normal and tumor breast tissue suggests a dual role for the ERBB2/her2 gene, Cancer Res. 71 (2011) 78–86.