European Review for Medical and Pharmacological Sciences 2020; 24: 1877-1886

Identification of lung adenocarcinoma-specific exosome RNAs in peripheral blood by RNA-Seq analysis

X.-Q. LIU[1], A. TUFMAN[1], R. KIEFL[1], G.-F. LI[2], Q.-L. MA[2], R.M. HUBER[1]

[1]Department of Internal Medicine V, Division of Respiratory Medicine and Thoracic Oncology, Ludwig-Maximilians University (LMU), Thoracic Oncology Centre Munich; Comprehensive Pneumology Center Munich, Member of the German Center for Lung Research (DZL), Munich, Bavaria, Germany
[2]Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Center. Kunming, Yunnan, China

Abstract. – OBJECTIVE: Several plasma-derived exosome RNAs have been identified as key regulators in cancer development. They have been considered as potential biomarkers for a non-invasive "liquid biopsy" to diagnose and assess the progression of cancer. This study aimed to identify human lung adenocarcinoma-specific exosome RNAs in peripheral blood, while assessing the feasibility and efficiency of this recently developed deep-sequencing technology for transcriptome profiling.

PATIENTS AND METHODS: Plasma-derived exosome RNAs were isolated from 13 lung adenocarcinoma patients, 3 patients with benign lung diseases, and 15 healthy volunteers. RNA-seq analysis of ribosomal RNA-depleted total RNA was performed. RNAs differentially expressed between lung adenocarcinoma and benign lung diseases or healthy volunteers were identified, followed by GO and KEGG pathway enrichment analyses for the identification of key exosome RNAs associated with lung adenocarcinomas.

RESULTS: Significant differentially expressed RNAs, such as UDP glucuronosyltransferase family 1 member A1 (UGT1A1) and BAI1-associated protein 2 like 1 (BAIAP2L1), were identified as differentially expressed between lung adenocarcinoma patients and patients with benign lung diseases. Eight pseudogenes, including Tropomyosin 1 (Alpha) Pseudogene (LOC100129096), Prothymosin, Alpha Pseudogene 2 (PTMAP2), Cell Division Cycle 14C, Pseudogene (CDC14C), Tropomyosin 1 (Alpha) Pseudogene (LOC643634), Ferritin Heavy Chain 1 Pseudogene 2 (FTH1P2), Actin Related Protein 2/3 Complex Subunit 3 Pseudogene 3 (ARPC3P3), Ferritin Heavy Chain 1 Pseudogene 11 (FTH1P11), and Prothymosin Alpha Pseudogene 5 (PTMAP5) were identified from plasma-derived exosomes in lung adenocarcinoma patients, who were more abundant/detectable than healthy volunteers.

CONCLUSIONS: Our data indicate that plasma-derived exosome RNAs, UGT1A1, and BAIAP2L1, as well as the eight isolated pseudogenes could serve as diagnostic and prognostic biomarkers for an effective non-invasive "liquid biopsy" of lung adenocarcinomas.

Key Words:
Lung adenocarcinoma, Pseudogenes, Exosome, Biomarkers.

Introduction

Lung cancer is the predominant cause of cancer-related death worldwide[1]. Non-small cell lung cancer (NSCLC) accounts for approximately 80% of all lung cancer cases, mainly including adenocarcinomas, squamous cell carcinomas, and large cell carcinomas[2-4]. In spite of great advances in NSCLC therapy, the prognosis for NSCLC patients remains poor[5]. A previous research[6] has indicated that the use of biomarkers to guide lung cancer therapies will be beneficial for the treatment of this malignancy. However, there are no reliable diagnostic and prognostic tools for the detection and identification of changes in the disease status. Therefore, seeking biomarkers is an important component of the lung cancer research.

Exosomes are 30-100 nm diameter membrane-enclosed extracellular vesicles originating from endosomes[7]. Under normal and stressful conditions, almost all cell types can secrete exo-

Corresponding Author: Rudolf M. Huber, MD; e-mail: huber@med.uni-muenchen.de

somes, and cancer cells are confirmed to produce more exosomes than normal cells from the same organ[8]. Moreover, exosomes are found to carry various biological molecules, such as lipids, proteins, and nucleic acid species, including mRNAs, miRNA, lncRNAs, rRNAs, genomic DNA, and pseudogenes[8,9]. Increasing evidence has shown that those exosomal-derived molecules play a crucial role in intercellular signal transduction, immunoregulation, and tumorigenesis. Also, these molecules, present in exosomes in bodily fluids, are considered promising biomarkers for the diagnosis and prognosis of many diseases[9,10], including a variety of cancers[11-13]. Authors[14,15] have shown that several plasma-derived exosome cargoes, such as exosome RNAs and proteins, are overexpressed in NSCLC and have been suggested as potential biomarkers for a non-invasive "liquid biopsy". Noteworthy, exosomal LRG1 has been identified as a potential urinary biomarker for detecting NSCLC[16]. Meanwhile, a number of techniques have been used and are under development for exosome genetic cargo analysis, such as quantitative Real Time Polymerase Chain Reaction (qRT-PCR), microarray analyses, and next-generation sequencing. Due to the power of RNA sequencing, which involves a combination of identification and quantification in a single high-throughput sequencing assay, the adoption of this technology has spread quickly.

In the present study, our findings provide a theoretical basis for the development of novel diagnostic and predictive biomarkers for lung adenocarcinomas.

Patients and Methods

Patient Samples and Exosome RNAs Isolation

A total of 13 lung adenocarcinoma cancer patients with newly diagnosed, untreated, histologically confirmed lung adenocarcinoma according to the World Health Organization classification were enrolled. Three patients with benign lung diseases and 15 healthy volunteers were recruited as controls. All the patients and healthy volunteer blood samples came from The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Center, China. This investigation was approved by the local Ethics Committee and all participants gave informed consent. About 10 ml blood samples were collected in Ethylene Diamine Tetraacetic Acid (EDTA)-anticoagulated blood collection tubes (BD Vacutainer, Plymouth, UK). After centrifugation, plasma samples were then collected and saved at –80° C. Plasma-derived exosome RNAs were isolated using the exoRNeasy Serum Plasma Kit (cat. no. 77064; Qiagen GmbH, Hilden, Germany), which was designed for the rapid purification of total vesicular RNAs, including non-coding RNA, mRNA, miRNA, and other small RNAs from serum or plasma. Sterility was maintained and all procedures were performed according to the protocols recommended by the manufacturers.

Construction of the cDNA Library and High-Throughput Sequencing

The rRNA removal reagents in the Ribo-Zero Gold Kit (Epicentre Biotechnologies, Madison, WI, USA) were used for the rRNA removal. Using the ScriptSeq™ mRNA-Seq Library Preparation Kit (Epicentre, Biotechnologies, Madison, WI, USA), the rRNA-depleted RNA was used for the construction of cDNA libraries. The cDNA libraries were then sequenced on Illumina® GAIIx and HiSeq 2000 platforms.

Preprocessing and Quality Control of Sequencing Data

For the retrieval of the clean reads, preprocessing was performed to remove the adapter sequences and the reads with a low quality of < 16 were removed. The quality control of clean reads was then conducted using FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). Moreover, RNA-seq intrinsic biases, such as nucleotide composition bias and GC bias, were assessed.

Sequence Alignment and Transcriptome Assembly

Genomic sequences of various organisms and their annotation information were deposited in the UCSC Genome Browser database (http://genome.ucsc.edu)[17]. Thus, the clean reads were mapped to the human reference genome (hg19) in the UCSC database using Tophat2.1.0 (http://tophat.cbcb.umd.edu)[18] with default parameters. According to the results of the sequence alignment, transcript assembly was then performed using StringTie[19]. The transcript assemblies were further merged using the Cuffmerge tool[20]. Cuffmerge results were compared with the Ensembl database (http://ensemblgenomes.org/) using Cuffcompare[20] to find known and novel RNAs.

Bioinformatics Analysis

Differentially expressed RNAs (including mRNAs and non-coding RNAs) were identified in lung adenocarcinoma patient plasma-derived exosomes using the DESeq2 package[21]. The cut-off values were an absolute fold change ≥ 2, with $p \leq 0.05$.

To understand the biological functions of differentially expressed RNAs, functional enrichment analyses were performed. Gene Ontology (GO, http://www.geneontology.org/)[22] contains a set of terms for the unification of the biology of large-scale gene lists. GO terms mainly include cellular component (CC), molecular function (MF), and biological process (BP). Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.ad.jp/kegg/)[23] pathway analysis is used to annotate the molecular interaction networks of large-scale gene lists. Therefore, GO and KEGG pathway enrichment analyses for the functional annotation of differentially expressed RNAs were carried out using clusterProfiler v3.2.1[24]. $p < 0.05$ was defined as the cut-off value.

Results

Using the DESeq2 package, the top 10 RNAs differentially regulated between lung adenocarcinoma patients and lung benign diseases were UDP glucuronosyltransferase family 1 member A1 (UGT1A1), BAI1 associated protein 2 like 1 (BAIAP2L1), calcium voltage-gated channel subunit alpha1 H (CACNA1H), carboxylesterase 1 (CES1), matrix Gla protein (MGP), DnaJ heat shock protein family (Hsp40) member C22 (DNAJC22), alanyl aminopeptidase, membrane (ANPEP), RNU1-2 RNA.U1 small nuclear 2, 3-hydroxyacyl-CoA dehydratase 2 (PTPLB), and LOC101928932. In addition, the top 10 RNA differentially expressed between lung adenocarcinoma patients and healthy controls were tropomyosin 1 (alpha) pseudogene (LOC100129096), prothymosin, alpha pseudogene 2 (PTMAP2), ubiquinol-cytochrome C reductase hinge protein like (UQCRHL), small EDRK-rich factor 2 (SERF2), cell division cycle 14C, pseudogene (CDC14C), tropomyosin 1 (alpha) pseudogene (LOC643634), ferritin heavy chain 1 pseudogene 2 (FTH1P2), actin-related protein 2/3 complex subunit 3 pseudogene 3 (ARPC3P3), ferritin heavy chain 1 pseudogene 11 (FTH1P11), and prothymosin alpha pseudogene 5 (PTMAP5). Notably, LOC100129096, PTMAP2, CDC14C, LOC643634, FTH1P2, ARPC3P3, FTH1P11, and PTMAP5 were pseudogenes.

Differentially expressed RNAs between lung adenocarcinoma and patients with benign lung diseases were significantly enriched in the GO BP functions, associated with steroid metabolic processes, protein activation cascades, and the acute inflammatory response. While, GO CC functions related to blood microparticles, the cytoplasmic membrane-bounded vesicle lumen, and vesicle lumen, and GO MF functions were associated with monooxygenase activity, steroid hydroxylase activity, and iron ion binding (Table I, Figure 1).

Furthermore, these differentially expressed RNAs fell into related KEGG pathways, such as complement and coagulation cascades, metabolism of xenobiotics by cytochrome P450, and retinol metabolism (Table II, Figure 2).

GO BP functions associated with differentially expressed RNAs between lung adenocarcinoma cancer patients and healthy controls were SRP (signal recognition particle)- membrane dependent co-translational protein targeting membrane, protein targeting to ER (Endoplasmic Reticulum), and co-translational protein targeting to the membrane. GO CC functions were cytosolic ribosome, ribosomal subunit, and ribosome. While GO MF functions were structural constituents of the ribosomes, protein binding involved in the cell-cell adhesion, and protein binding involved in the cell adhesion (Table III, Figure 3).

Also, these differentially expressed RNAs belonged to pathways associated with ribosomes, platelet activation, and allograft rejection (Table IV, Figure 4).

Discussion

RNA-Seq (RNA sequencing), also called whole transcriptome shotgun sequencing[25], uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment[26,27]. It can be used to detect transcripts that correspond to existing genomic sequences and are relative to DNA microarrays, it has a very low background signal because DNA sequences can be unambiguously mapped to unique regions of the genome[27]. However, there are numerous variations among RNA-seq protocols, and it is still a challenge to conduct standardized RNA-seq studies.

Table I. The top 10 GO terms associated with genes differentially expressed between lung adenocarcinoma and lung benign diseases patients.

Term	Description	Count	*p*
BP			
GO:0008202	Steroid metabolic process	41	3.13E-26
GO:0072376	Protein activation cascade	23	1.17E-18
GO:0002526	Acute inflammatory response	23	1.31E-16
GO:0044283	Small molecule biosynthetic process	39	1.93E-15
GO:1901615	Organic hydroxy compound metabolic process	36	2.15E-15
GO:0002576	Platelet degranulation	19	4.34E-14
GO:0042632	Cholesterol homeostasis	16	5.63E-14
GO:0055092	Sterol homeostasis	16	5.63E-14
GO:0016125	Sterol metabolic process	20	3.28E-13
GO:0055088	Lipid homeostasis	19	5.99E-13
CC			
GO:0072562	Blood microparticle	33	5.16E-28
GO:0060205	Cytoplasmic membrane-bounded vesicle lumen	22	4.65E-18
GO:0031983	Vesicle lumen	22	5.77E-18
GO:0034774	Secretory granule lumen	19	3.20E-16
GO:0031093	Platelet alpha granule lumen	15	1.51E-14
GO:0031091	Platelet alpha granule	15	1.73E-12
GO:0042627	Chylomicron	7	6.45E-10
GO:0034358	Plasma lipoprotein particle	9	1.21E-08
GO:1990777	Lipoprotein particle	9	1.21E-08
GO:0032994	Protein-lipid complex	9	1.97E-08
MF			
GO:0004497	Monooxygenase activity	17	7.76E-13
GO:0008395	Steroid hydroxylase activity	11	6.13E-12
GO:0005506	Iron ion binding	20	9.47E-12
GO:0020037	Heme binding	17	8.54E-11
GO:0005539	Glycosaminoglycan binding	21	1.02E-10
GO:0046906	Tetrapyrrole binding	17	2.59E-10
GO:0019825	Oxygen binding	11	4.42E-10
GO:0016705	Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	18	4.50E-10
GO:0016709	Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NAD(P)H as one donor, and incorporation of one atom of oxygen	10	2.33E-09
GO:0008201	Heparin binding	17	3.05E-09

GO: Gene Ontology; BP: biological progress; MF: molecular function; CC: cellular component, which represent the name of GO terms; Count represents the number of enriched genes.

UGT1A1 is located in the long arm of chromosome 2. Growing evidence confirms UGT1A1 polymorphisms are associated with irinotecan-induced toxicities in patients with lung cancer[28,29]. Irinotecan-containing regimens are actively used in patients with lung cancer[30,31]. Moreover, Nakamura et al[32] suggested that two genotypes, UGT1A1*6 and UGT1A1*27 might be used as promising predictors of grade 4 neutropenia following irinotecan-based chemotherapy in NSCLC patients. Yamamoto et al[33] also indicated that the influence of UGT1A1*28 and *6 polymorphisms could predict irinotecan-related hematological toxicity in elderly patients with advanced NSCLC. These data imply that UGT1A1 polymorphisms may be used as a biomarker for the prediction of the risk of severe toxicities during irinotecan therapy for NSCLC. Considering our results, we provide further evidence that abnormal expression of UGT1A1 may regulate the malignant progression of NSCLC and speculate that dysregulation of UGT1A1 may serve as a biomarker for predicting treatment outcomes and prognosis of this disease.

BAIAP2L1 is located on chromosome 7q21.3-q22.1. In a previous research, BAIAP2L1 has been shown to regulate the deformation of the plasma membrane and actin cytoskel-

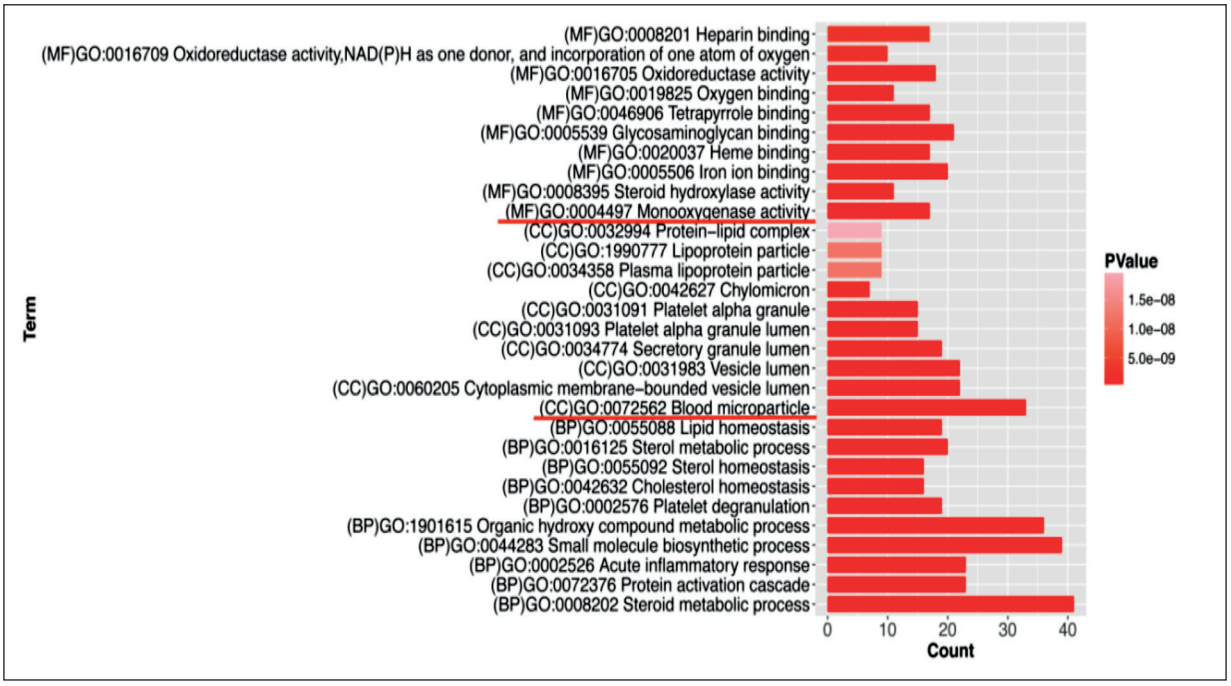

Figure 1. The top 10 GO terms associated with genes differentially expressed between lung adenocarcinoma and lung benign diseases patients.

eton remodeling, which play a central role in cell migration[34]. In tumorigenesis, BAIAP2L1 is up-regulated in ovarian cancer tissues and its upregulation promotes cell proliferation and inhibits apoptosis[35]. Wang et al[36] demonstrated that BAIAP2L1 could increase hepatocellular carcinoma cell proliferation via activation of the EGFR-ERK pathway. Moreover, BAIAP2L1 expression is correlated with the levels of C-reactive protein (CRP) in rheumatoid arthritis fibroblast-like synovial cells[37]. CRP levels could be used for predicting survival in patients with urological cancers[38]. Notably, the identification of an FGF (fibroblast growth factor) receptor 3 (FGFR3)-BAIAP2L1 fusion gene may aid in selecting lung cancer patients for FGFR-targeted therapy[39]. Collectively, we believe that BAIAP2L1 may play a key role in cancer progression and treatment. Based on our results, we further speculate that exosome BAIAP2L1 may be involved in the malignant progression of NSCLC and may also be a promising biomarker to predict the treatment outcome and prognosis of this disease.

Table II. The top 10 KEGG pathways associated with genes differentially expressed between lung adenocarcinoma and lung benign diseases patients.

Term	Description	Count	*p*
hsa04610	Complement and coagulation cascades	24	3.29E-20
hsa00980	Metabolism of xenobiotics by cytochrome P450	15	2.96E-10
hsa00830	Retinol metabolism	14	5.06E-10
hsa00982	Drug metabolism - cytochrome P450	14	1.44E-09
hsa00140	Steroid hormone biosynthesis	12	1.49E-08
hsa04976	Bile secretion	13	1.76E-08
hsa05204	Chemical carcinogenesis	13	1.06E-07
hsa05150	Staphylococcus aureus infection	11	1.07E-07
hsa00120	Primary bile acid biosynthesis	6	2.32E-06
hsa00983	Drug metabolism - other enzymes	8	1.61E-05

KEGG: Kyoto Encyclopedia of Genes and Genomes. Description represents the name 8of KEGG pathways; Count represents the number of enriched genes.

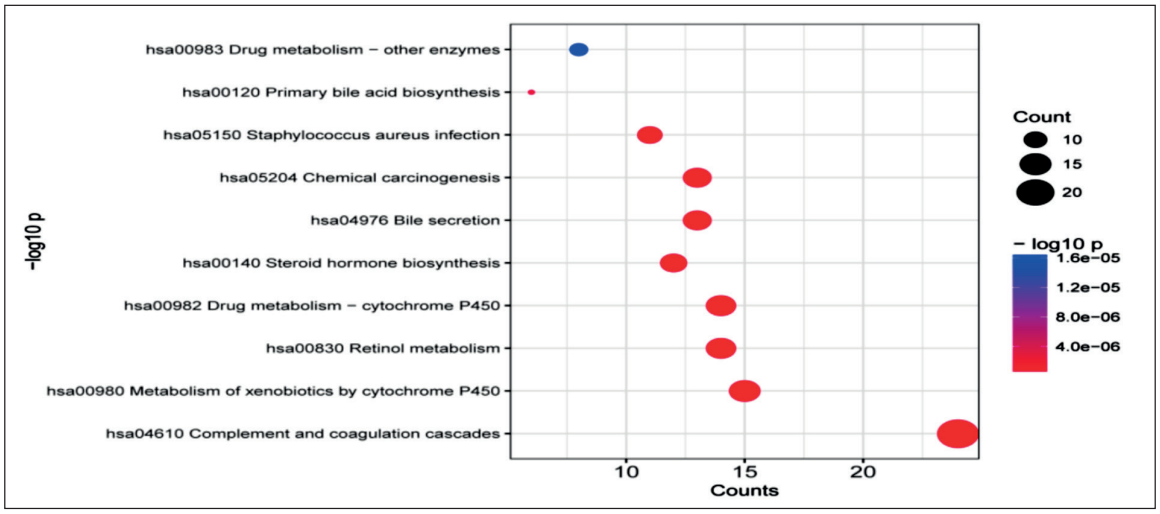

Figure 2. The top 10 KEGG pathways associated with genes differentially expressed between lung adenocarcinoma and lung benign diseases patients. KEGG: Kyoto

Table III. The top 10 GO terms associated with genes differentially expressed between lung adenocarcinoma patients and healthy controls.

Term	Description	Count	*p*
BP			
GO:0006614	SRP-dependent co-translational protein targeting to membrane	77	8.90E-41
GO:0045047	Protein targeting to ER	78	2.55E-38
GO:0006613	Co-translational protein targeting to membrane	78	9.47E-38
GO:0072599	Establishment of protein localization to endoplasmic reticulum	78	3.99E-36
GO:0006364	rRNA processing	136	2.16E-34
GO:0016072	rRNA metabolic process	138	2.24E-34
GO:0070972	Protein localization to endoplasmic reticulum	84	2.98E-33
GO:0042254	Ribosome biogenesis	154	6.52E-33
GO:0000184	Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	82	1.61E-32
GO:0022613	Ribonucleoprotein complex biogenesis	193	4.52E-31
CC			
GO:0022626	Cytosolic ribosome	83	2.88E-38
GO:0044391	Ribosomal subunit	103	1.03E-37
GO:0005840	Ribosome	120	4.64E-32
GO:0044445	Cytosolic part	112	5.07E-27
GO:0022625	Cytosolic large ribosomal subunit	48	5.67E-25
GO:0005925	Focal adhesion	156	4.61E-24
GO:0005924	Cell-substrate adherens junction	156	1.18E-23
GO:0030055	Cell-substrate junction	157	1.35E-23
GO:0015934	Large ribosomal subunit	59	2.40E-21
GO:0015935	Small ribosomal subunit	44	9.42E-18
MF			
GO:0003735	Structural constituent of ribosome	110	5.67E-28
GO:0098632	Protein binding involved in cell-cell adhesion	102	1.08E-10
GO:0098631	Protein binding involved in cell adhesion	102	3.18E-10
GO:0098641	Cadherin binding involved in cell-cell adhesion	98	3.56E-10
GO:0019843	rRNA binding	33	4.31E-10
GO:0045296	Cadherin binding	100	1.96E-09
GO:0050839	Cell adhesion molecule binding	136	4.95E-09
GO:0003924	GTPase activity	75	3.95E-07
GO:0003779	Actin binding	111	2.06E-06
GO:0023026	MHC class II protein complex binding	11	1.80E-05

GO: Gene Ontology; BP: biological progress; MF: molecular function; CC: cellular component, which represent the name of GO terms; Count represents the number of enriched genes.

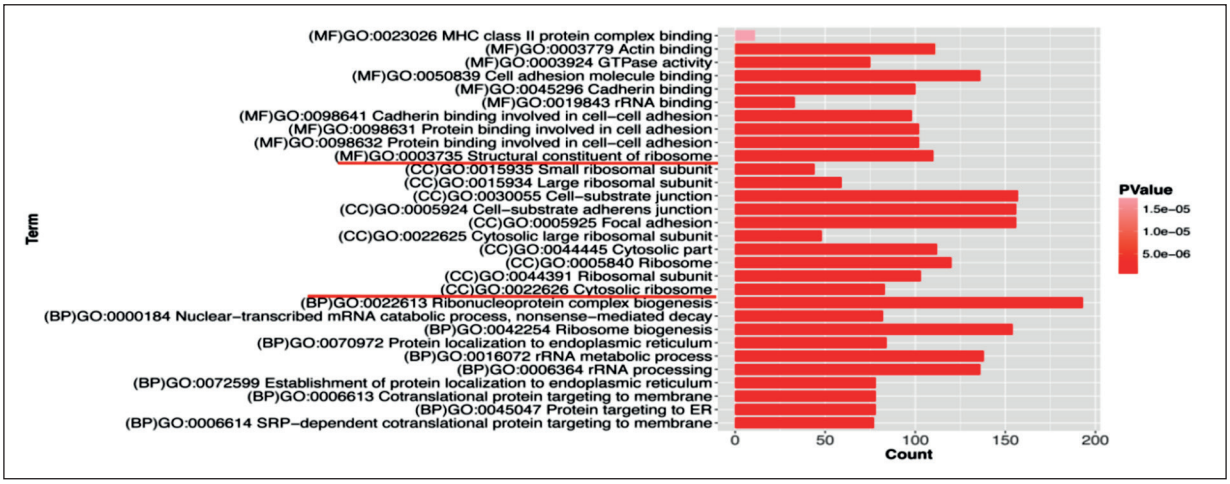

Figure 3. The top 10 GO terms associated with genes differentially expressed between lung adenocarcinoma patients and healthy controls.

Compared with healthy controls, eight of the top 10 differentially expressed RNAs identified from lung adenocarcinoma patients were pseudogenes, including LOC100129096, PTMAP2, CDC14C, LOC643634, FTH1P2, ARPC3P3, FTH1P11, and PTMAP5. Pseudogenes are generally noncoding. In the past decades, pseudogenes have been treated as nonfunctional "junk" sequences within genomic DNA originally derived from protein-coding genes. However, these genes have a high sequence similarity with their parental protein-coding genes, thus generating the capability for sequence-specific regulation. Thanks to the rapid development of next-generation sequencing technology, our understanding of pseudogene transcription and function has increased. Pseudogenes have been newly discovered as key players in the development of human cancers[40,41]. Increasing evidence[42-44] has revealed that several pseudogenes may contribute to the development of lung cancer. Moreover, CDC14C is down-regulated in large cell carcinomas in lung cancer[45]. FTH1P2 and FTH1P11 are up-regulated by TUSC2-erlotinib and auranofin-TUSC2-erlotinib combination treatments in NSCLC patients[46], suggesting that FTH1P2 and FTH1P11 may play a key role in regulating drug sensitivity. However, the roles of other identified pseudogenes in cancer biology and treatment have not been investigated. Given the rising understanding of pseudogenes in cancer development, we speculate that the identified pseudogenes from plasma-derived exosomes may be involved in tumorigenesis and may also serve

Table IV. The top 10 KEGG pathways associated with genes differentially expressed between lung adenocarcinoma patients and healthy controls.

Term	Description	Count	*p*
hsa03010	Ribosome	96	1.29E-28
hsa04611	Platelet activation	51	3.24E-07
hsa05330	Allograft rejection	21	4.86E-06
hsa05416	Viral myocarditis	28	7.10E-06
hsa04940	Type I diabetes mellitus	22	1.52E-05
hsa05166	HTLV-I infection	82	3.05E-05
hsa05012	Parkinson's disease	51	4.27E-05
hsa04612	Antigen processing and presentation	32	4.76E-05
hsa05340	Primary immunodeficiency	19	5.44E-05
hsa04510	Focal adhesion	66	6.34E-05

KEGG: Kyoto Encyclopedia of Genes and Genomes. Count represents the number of enriched genes.

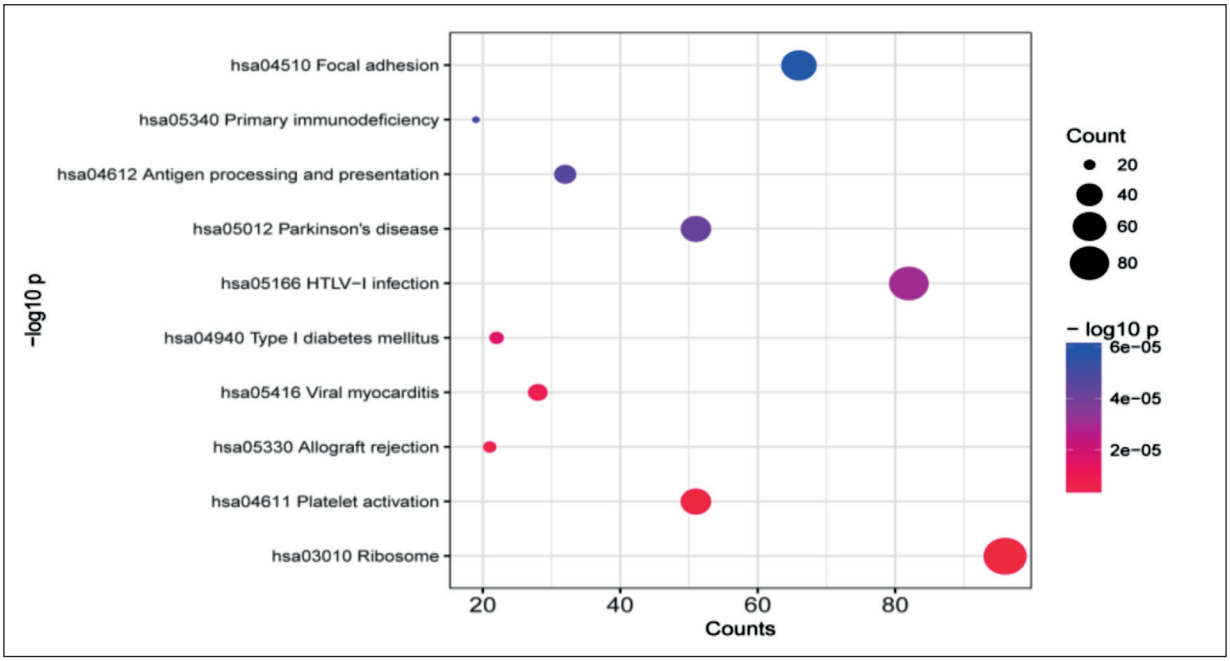

Figure 4. The top 10 KEGG pathways associated with genes differentially expressed between lung adenocarcinoma patients and healthy controls.

as biomarkers for lung adenocarcinoma diagnosis and could be used to monitor treatment. The roles of pseudogenes in NSCLC tumorigenesis and diagnosis still require investigation.

Conclusions

This study showed the advantages of isolating exosome RNAs from the peripheral blood of lung cancer patients and evaluating them using the RNA-Seq analysis. Isolated plasma-derived exosome RNAs could serve as novel candidate diagnostic and prognostic biomarkers for an effective non-invasive "liquid biopsy" of lung adenocarcinomas. In future studies, the candidate exosome biomarkers should be further validated and evaluated for prognosis using quantitative Real Time Polymerase Chain Reaction in a follow-up in a cohort of ~30 NSCLC patients. Furthermore, the role of isolated plasma-derived exosome RNAs in the tumorigenesis of lung adenocarcinomas should be investigated.

Conflict of Interest

The Authors declare that they have no conflict of interests.

References

1) Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015; 65: 87-108.
2) Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Dilling TJ, Dobelbower MC, Govindan R, Grannis FW Jr, Horn L, Jahan TM, Komaki R, Krug LM, Lackner RP, Lanuti M, Lilenbaum R, Lin J, Loo BW Jr1, Martins R, Otterson GA, Patel JD, Pisters KM, Reckamp K, Riely GJ, Rohren E, Schild SE, Shapiro TA, Swanson SJ, Tauer K, Yang SC, Gregory K, Hughes M; National comprehensive cancer network. Non-small cell lung cancer, Version 6.2015. J Natl Compr Canc Netw 2015; 13: 515-524.
3) Li N, Feng XB, Tan Q, Luo P, Jing W, Zhu M, Liang C, Tu J, Ning Y. Identification of circulating long noncoding RNA Linc00152 as a novel biomarker for diagnosis and monitoring of non-small-cell lung cancer. Dis Markers 2017; 2017: 7439698.
4) Weber DG, Johnen G, Casjens S, Bryk O, Pesch B, Jöckel KH, Kollmeier J, Brüning T. Evaluation of long noncoding RNA MALAT1 as a candidate blood-based biomarker for the diagnosis of non-small cell lung cancer. BMC Res Notes 2013; 6: 518.
5) Maeda T, Ueoka H, Tabata M, Kiura K, Shibayama T, Gemba K, Takigawa N, Hiraki A, Katayama H, Harada M. Prognostic factors in advanced non-small cell

lung cancer: elevated serum levels of neuron specific enolase indicate poor prognosis. Jpn J Clin Oncol 2000; 30: 534-541.

6) Reck M, Heigener DF, Mok T, Soria JC, Rabe KF. Management of non-small-cell lung cancer: recent developments. Lancet 2013; 382: 709-719.

7) Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol 2013; 200: 373-383.

8) Srivastava A, Filant J, Moxley KM, Sood A, McMeekin S, Ramesh R. Exosomes: a role for naturally occurring nanovesicles in cancer growth, diagnosis and treatment. Curr Gene Ther 2015; 15: 182-192.

9) Miranda KC, Bond DT, McKee M, Skog J, Păunescu TG, Da Silva N, Brown D, Russo LM. Nucleic acids within urinary exosomes/microvesicles are potential biomarkers for renal disease. Kidney Int 2010; 78: 191-199.

10) Boukouris S, Mathivanan S. Exosomes in bodily fluids are a highly stable resource of disease biomarkers. Proteomics Clin Appl 2015; 9: 358-367.

11) Alegre E, Zubiri L, Perez-Gracia JL, Gonzalez-Cao M, Soria L, Martin-Algarra S, Gonzalez A. Circulating melanoma exosomes as diagnostic and prognosis biomarkers. Clin Chim Acta 2016; 454: 28-32.

12) Taylor DD, Gercel-Taylor C. MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. Gynecol Oncol 2008; 110: 13-21.

13) Chiara M. Exosomes: new biomarkers for targeted cancer therapy. Springer International Publishing; 2017.

14) Reclusa P, Taverna S, Pucci M, Durendez E, Calabuig S, Manca P, Serrano MJ, Sober L, Pauwels P, Russo A, Rolfo C. Exosomes as diagnostic and predictive biomarkers in lung cancer. J Thorac Dis 2017; 9: S1373-S1382.

15) Sandfeld-Paulsen B, Jakobsen KR, Baek R, Folkersen BH, Rasmussen TR, Meldgaard P, Varming K, Jørgensen MM, Sorensen BS. Exosomal proteins as diagnostic biomarkers in lung cancer. J Thorac Oncol 2016; 11: 1701-1710.

16) Li Y, Zhang Y, Qiu F, Qiu Z. Proteomic identification of exosomal LRG1: a potential urinary biomarker for detecting NSCLC. Electrophoresis 2011; 32: 1976-1983.

17) Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, Harte RA, Heitner S, Hinrichs AS, Learned K, Lee BT, Li CH, Raney BJ, Rhead B, Rosenbloom KR, Sloan CA, Speir ML, Zweig AS, Haussler D, Kuhn RM, Kent WJ. The UCSC Genome Browser database: 2014 update. Nucleic Acids Res 2014; 42: D764-770.

18) Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009; 25: 1105-1111.

19) Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 2015; 33: 290-295.

20) Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol 2013; 31: 46-53.

21) Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.

22) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-29.

23) Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000; 28: 27-30.

24) Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.

25) Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques 2008; 45: 81-94.

26) Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Ther 2012; 22: 271-274.

27) Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009; 10: 57-63.

28) Chen X, Liu L, Guo Z, Liang W, He J, Huang L, Deng Q, Tang H, Pan H, Guo M, Liu Y, He Q, He J. UGT1A1 polymorphisms with irinotecan-induced toxicities and treatment outcome in Asians with lung cancer: a meta-analysis. Cancer Chemother Pharmacol 2017; 79: 1109-1117.

29) Fukuda M, Okumura M, Iwakiri T, Arimori K, Honda T, Kobayashi K, Senju H, Takemoto S, Ikeda T, Yamaguchi H, Nakatomi K, Matsuo N, Mukae H, Ashizawa K. Relationship between UGT1A1*27 and UGT1A1*7 polymorphisms and irinotecan-related toxicities in patients with lung cancer. Thorac Cancer 2018; 9: 51-58.

30) Rutledge MR, Waddell JA, Solimando DA, Jr. Irinotecan and carboplatin (IC) regimen for small-cell lung cancer and non-small-cell lung cancer. Hosp Pharm 2014; 49: 603-611.

31) Jeong J, Cho BC, Sohn JH, Choi HJ, Kim SH, Lee YJ, Jung MK, Shin SJ, Park MS, Kim SK, Chang J, Kim JH. Belotecan for relapsing small-cell lung cancer patients initially treated with an irinotecan-containing chemotherapy: a phase II trial. Lung Cancer 2010; 70: 77-81.

32) Nakamura Y, Soda H, Oka M, Kinoshita A, Fukuda M, Fukuda M, Takatani H, Nagashima S, Soejima Y, Kasai T, Nakatomi K, Masuda N, Tsukamoto K, Kohno S. Randomized phase II trial of irinotecan with pacl-

itaxel or gemcitabine for non-small cell lung cancer: association of UGT1A1*6 and UGT1A1*27 with severe neutropenia. J Thorac Oncol 2011; 6: 121-127.

33) Yamamoto N, Takahashi T, Kunikane H, Masuda N, Eguchi K, Shibuya M, Takeda Y, Isobe H, Ogura T, Yokoyama A, Watanabe K. Phase I/II pharmacokinetic and pharmacogenomic study of UGT1A1 polymorphism in elderly patients with advanced non-small cell lung cancer treated with irinotecan. Clin Pharmacol Ther 2009; 85: 149-154.

34) Chen G, Li T, Zhang L, Yi M, Chen F, Wang Z, Zhang X. Src-stimulated IRTKS phosphorylation enhances cell migration. FEBS Lett 2011; 585: 2972-2978.

35) Chao A, Tsai CL, Jung SM, Chuang WC, Kao C, Hsu A, Chen SH, Lin CY, Lee YC, Lee YS, Wang TH, Wang HS, Lai CH. BAI1-associated protein 2-like 1 (BAIAP2L1) is a potential biomarker in ovarian cancer. PLoS One 2015; 10: e0133081.

36) Wang YP, Huang LY, Sun WM, Zhang ZZ, Fang JZ, Wei BF, Wu BH, Han ZG. Insulin receptor tyrosine kinase substrate activates EGFR/ERK signalling pathway and promotes cell proliferation of hepatocellular carcinoma. Cancer Lett 2013; 337: 96-106.

37) Galligan CL, Baig E, Bykerk V, Keystone EC, Fish EN. Distinctive gene expression signatures in rheumatoid arthritis synovial tissue fibroblast cells: correlates with disease activity. Genes Immun 2007; 8: 480-491.

38) Saito K, Kihara K. C-reactive protein as a biomarker for urological cancers. Nat Rev Urol 2011; 8: 659-666.

39) Nakanishi Y, Akiyama N, Tsukaguchi T, Fujii T, Satoh Y, Ishii N, Aoki M. Mechanism of oncogenic signal activation by the novel fusion kinase FGFR3-BAIAP2L1. Mol Cancer Ther 2015; 14: 704-712.

40) Poliseno L. Pseudogenes: newly discovered players in human cancer. Sci Signal 2012; 5: re5.

41) Poliseno L, Marranci A, Pandolfi PP. Pseudogenes in human cancer. Front Med (Lausanne) 2015; 2: 68.

42) Tutar Y, Özgür A, Tutar E, Tutar L, Pulliero A, Izzotti A. Regulation of oncogenic genes by MicroRNAs and pseudogenes in human lung cancer. Biomed Pharmacother 2016; 83: 1182-1190.

43) Wei Y, Chang Z, Wu C, Zhu Y, Li K, Xu Y. Identification of potential cancer-related pseudogenes in lung adenocarcinoma based on ceRNA hypothesis. Oncotarget 2017; 8: 59036-59047.

44) Sun M, Nie FQ, Zang C, Wang Y, Hou J, Wei C, Li W, He X, Lu KH. The pseudogene DUXAP8 promotes non-small-cell lung cancer cell proliferation and invasion by epigenetically silencing EGR1 and RHOB. Mol Ther 2017; 25: 739-751.

45) Wang J, Zhang Y, Wan H, Dong N, Bao L, Sun X, Xu M, Wang X. Global analyses of subtype- or stage-specific genes on chromosome 7 in patients with lung cancer. Cancer Metastasis Rev 2015; 34: 333-345.

46) Xiaobo C, Majidi M, Feng M, Shao R, Wang J, Zhao Y, Baladandayuthapani V, Song J, Fang B, Ji L, Mehran R, Roth JA. TUSC2(FUS1)-erlotinib induced vulnerabilities in epidermal growth factor receptor(EGFR) wildtype non-small cell lung cancer(NSCLC) targeted by the repurposed drug auranofin. Sci Rep 2016; 6: 35741.