Article

Plasma Next Generation Sequencing and Droplet Digital-qPCR-Based Quantification of Circulating Cell-Free RNA for Noninvasive Early Detection of Cancer

Martin Metzenmacher [1,2], **Renáta Váraljai** [3,4], **Balazs Hegedüs** [5], **Igor Cima** [4,6],
Jan Forster [4,7], **Alexander Schramm** [8], **Björn Scheffler** [4,6], **Peter A. Horn** [9],
Christoph A. Klein [10,11], **Tibor Szarvas** [12,13], **Hennig Reis** [14], **Nicola Bielefeld** [4,15],
Alexander Roesch [3,4], **Clemens Aigner** [5], **Volker Kunzmann** [16], **Marcel Wiesweg** [1,2],
Jens T. Siveke [4,15], **Martin Schuler** [1,2,4] and **Smiths S. Lueong** [4,15,*]

1   Department of Medical Oncology, West German Cancer Center, University Hospital Essen, University Duisburg-Essen, Hufelandstrasse 55, 45122 Essen, Germany; Martin.Metzenmacher@uk-essen.de (M.M.); Marcel.Wiesweg@uk-essen.de (M.W.); Martin.Schuler@uk-essen.de (M.S.)
2   Division of Thoracic Oncology, University Medicine Essen-Ruhrlandklinik, University Duisburg-Essen, Tüschener Weg 40, 45239 Essen, Germany
3   Department of Dermatology, University Hospital Essen, West German Cancer Center, University Hospital Essen, University Duisburg-Essen, Hufelandstrasse 55, 45122 Essen, Germany; Renata.Varaljai@uk-essen.de (R.V.); Alexander.Roesch@uk-essen.de (A.R.)
4   German Cancer Consortium (DKTK), Partner site University Hospital Essen, Hufelandstrasse 55, 45122 Essen, Germany; i.cima@dkfz-heidelberg.de (I.C.); j.forster@dkfz-heidelberg.de (J.F.); Bjoern.Scheffler@uk-essen.de (B.S.); Nicola.Bielefeld@uk-essen.de (N.B.); Jens.Siveke@uk-essen.de (J.T.S.)
5   Department of Thoracic Surgery, Ruhrlandklinik, University Duisburg-Essen, D-45239 Essen, Germany; Balazs.Hegedues@rlk.uk-essen.de (B.H.); Clemens.Aigner@rlk.uk-essen.de (C.A.)
6   DKFZ-Division Translational Neurooncology at the WTZ, DKTK partner site, University Hospital Essen, 45122 Essen, Germany
7   Chair for Genome Informatics, Department of Human Genetics, University Hospital Essen, University Duisburg-Essen, Hufelandstrasse 55, 45122 Essen, Germany
8   Laboratory for Molecular Oncology, Department of Medical Oncology, West German Cancer Center, University Hospital Essen, University of Duisburg-Essen, 45122 Essen, Germany; alexander.schramm@uni-due.de
9   Institute for Transfusion Medicine, University Hospital Essen, 45122 Essen, Germany; Peter.Horn@uk-essen.de
10   Experimental Medicine and Therapy Research, University of Regensburg, 93053 Regensburg, Germany; Christoph.Klein@klinik.uni-regensburg.de
11   Fraunhofer-Institute for Toxicology and Experimental Medicine, Division of Personalized Tumour Therapy, 93053 Regensburg, Germany
12   Department of Urology, West German Cancer Center, University Hospital Essen, University Duisburg-Essen, 45122 Essen, Germany; Tibor.Szarvas@uk-essen.de
13   Department of Urology, Semmelweis University, H-1085 Budapest, Hungary
14   Institute of Pathology, West German Cancer Center, University Hospital Essen, University Duisburg-Essen, 45147 Essen, Germany; Henning.Reis@uk-essen.de
15   Institute for Developmental Cancer Therapeutics, West German Cancer Center, University Hospital Essen, 45122 Essen, Germany
16   Department of Internal Medicine II, University Hospital Würzburg, 97080 Würzburg, Germany; kunzmann_v@ukw.de
*   Correspondence: smiths-sengkwawoh.lueong@uk-essen.de; Tel.: +49-(201)-723-3139

Abstract: Early detection of cancer holds high promise for reducing cancer-related mortality. Detection of circulating tumor-specific nucleic acids holds promise, but sensitivity and specificity issues remain with current technology. We studied cell-free RNA (cfRNA) in patients with non-small cell lung cancer (NSCLC; *n* = 56 stage IV, *n* = 39 stages I-III), pancreatic cancer (PDAC, *n* = 20 stage III), malignant melanoma (MM, *n* = 12 stage III-IV), urothelial bladder cancer (UBC, *n* = 22 stage II and IV), and 65 healthy controls by means of next generation sequencing (NGS) and real-time droplet digital PCR (RT-ddPCR). We identified 192 overlapping upregulated transcripts in NSCLC and PDAC by NGS, more than 90% of which were noncoding. Previously reported transcripts (e.g., HOTAIRM1) were identified. Plasma cfRNA transcript levels of POU6F2-AS2 discriminated NSCLC from healthy donors (AUC = 0.82 and 0.76 for stages IV and I–III, respectively) and significantly associated (*p* = 0.017) with the established tumor marker Cyfra 21-1. cfRNA yield and POU6F2-AS transcript abundance discriminated PDAC patients from healthy donors (AUC = 1.0). POU6F2-AS2 transcript was significantly higher in MM (*p* = 0.044). In summary, our findings support further validation of cfRNA detection by RT-ddPCR as a biomarker for early detection of solid cancers.

1. Introduction

Lung cancer (LC) is a leading cause of all cancer-related deaths in men and women worldwide, accounting for more than 14% of all newly diagnosed cancers [1,2]. Although the understanding of lung cancer pathobiology has significantly improved over the last decades, disease prognosis continue to remain poor with a five-year survival rate of only 17% and 7% for non-small cell lung cancer (NSCLC) and small cell lung cancers (SCLC), respectively [3,4]. As with other cancers, poor disease prognosis is partially attributed to late stages at diagnosis, given that there are very few early symptoms [5,6]. Indeed, about 67% of all lung cancer cases are diagnosed at a late stage, leading to high mortality rates [6]. Lung cancer treatment is stage-specific. Early stage disease (I-II UICC 8th edition) can be cured by surgical resection, while locally advanced disease requires multimodal treatment, including chemotherapy, radiotherapy, and surgery for selected cases. Recurrent metastatic disease is treated with palliative systemic therapy [7–9]. When diagnosed at an early stage, patients with NSCLC have a 5-year survival rate of about 71%, as opposed to less than 2% for patients diagnosed with stage IV disease [10]. Early diagnosis may thus improve patient outcome [11].

To this end, low-dose computed tomography (LDCT) was developed for screening in high-risk groups, with a reported sensitivity of about 93% and a 20% reduction in mortality [12,13]. Despite this, its wide application is hampered by consideration of cost effectiveness, high rates of false positives, risk of developing radiation-related cancers, and over-diagnosis [14,15]. Therefore, there is still a considerable unmet need to develop non-invasive complementary biomarkers for early detection of lung cancer.

Currently, protein- and microRNA (miRNA)-based markers are at advanced validation stages or already in clinical use [16]. Carcinogen embryonic antigen (CEA) is one of the most widely used biomarkers for cancer screening in several entities, including lung cancer. Cytokeratin-19 fragment (Cyfra21-1) is also a tumor marker for lung cancer [17,18]. However, high false positive elevated levels of these markers in benign disease and restricted sensitivity limit their use for routine screening for early diagnosis of cancer [6,18–20]. Immune-based markers are attractive. However, they show very low sensitivity. For example, one commercially available autoimmune antibodies test showed only 36% sensitivity but 91% specificity in NSCLC [21]. miRNAs have been also tested in several studies on a plethora of platforms, including microarray, next-generation sequencing (NGS), and RT-PCR. While NGS and microarray-based approaches are highly reproducible and specific, they lack the accuracy and sensitivity required for early detection, especially in the context of low copy transcripts [22].

The most attractive non-invasive solution for the identification of biomarkers for early detection of lung cancer is through simple blood test [23]. To date, very few studies have addressed the utility of cell-free RNA (cfRNA) for early diagnosis of cancer, despite previous reports highlighting the potential clinical utility of this liquid compartment. Initially identified in patients with malignant melanoma [24] and nasopharyngeal carcinoma [25], cfRNA has been reported in several other cancer entities, including breast cancer [26], colorectal cancer [27–29], follicular lymphoma [30], and hepatocellular carcinoma [31]. Recently, cfRNA has been used for non-invasive determination of gestational aged with better performance compared with standard techniques [32]. cfRNA is therefore a potential source of analyte for early detection biomarker discovery. In light of the aforementioned, we herein report on the use of cfRNA for early detection of cancer in solid cancers with an application to lung, pancreas, bladder, and skin cancers. We use a combination of NGS-based cfRNA profiling and real-time digital droplet PCR (RT-ddPCR) for the analysis of cfRNA for early diagnosis of solids cancers.

2. Results

2.1. Patients and Baseline Characteristics

In total, we analyzed cfRNA in blood samples from 153 patients with four different cancer entities and 65 healthy donors (Table 1) using total cfRNA-sequencing and RT-ddPCR. First, a pilot cohort comprising 11 stage IV NSCLC patients, 4 stage IV PDAC patients, and 4 healthy donors (cfRNA-seq) was studied. In 12 of 15 sequenced samples, more than 50 million reads were obtained, and about 40 million reads were obtained for three other samples. More than 70% of reads could be uniquely mapped in all samples, and protein-coding genes were covered by about 5% of reads in all sequenced samples (Supplementary Figure S1). This was followed by a validation cohort of 84 NSCLC, 12 MM, 20 PDAC, 22 UBC patients, and 61 healthy donors. Tissue expression of specific RNA transcripts was studied in formalin-fixed paraffin-embedded (FFPE) tissue samples from 18 NSCLC and 9 healthy lungs by RT-ddPCR.

Table 1. Baseline clinic-pathological parameters of the study population.

Variable	Discovery Cohort		Validation Cohort			
	NSCLC	**PDAC**	**NSCLC**	**PDAC**	**Melanoma**	**Bladder**
Total, *n*	11	4	84	20	12	22
Age						
Median (SD)	65 (8.0)	66 (8.7)	66 (9.2)	70 (9.6)	54 (13.5)	70 (10.3)
Gender, *n* (%)						
Male	4 (36)	2 (50)	42 (50)	11 (55)	7 (60)	18 (80)
Female	7 (64)	2 (50)	42 (50)	9 (45)	5 (40)	4 (20)
Smoking, *n* (%)						
Never smoker	-	-	5 (6)	-	-	-
Past smoker	-	-	18 (21)	-	-	-
Current smoker	-	-	14 (17)	-	-	-
Unknown	11 (100)	4 (100)	47 (56)	20 (100)	12 (100)	22 (100)
ECOG status						
0	8 (70)	-	64 (76)	-	10 (83)	-
1	3 (30)	-	16 (19)	-	2 (17)	-
2	-	-	1 (1)	-	-	-
Unknown	-	4 (100)	3 (4)	20 (100)	-	22 (100)
Stage						
Early [a]	-	1 (25)	39 (46)		-	6 (30)
Late	11 (100)	3 (75)	45 (54)	20 (100)	12 (100)	16 (70)

Table 1. *Cont.*

Variable	Discovery Cohort		Validation Cohort			
	NSCLC	PDAC	NSCLC	PDAC	Melanoma	Bladder
CEA						
<ULN	1 (1)	-	9 (11)	-	-	-
≥ULN	10 (99)	-	32 (38)	-	-	-
Not done	-	4 (100)	43 (51)	20 (100)	12 (100)	22 (100)
Cyfra21-1						
<ULN	3 (30)	-	5 (6)	-	-	-
≥ULN	8 (70)	-	36 (43)	-	-	-
Not done	-	4 (100)	43 (51)	20 (100)	12 (100)	22 (100)
CA 19.9						
<ULN	-	-	-	5 (25)	-	-
≥ULN	-	2 (50)	-	15 (75)	-	-
Not done	11 (100)	2 (50)	84 (100)	-	12 (100)	22 (100)
LDH						
<ULN	-	-	-	-	10 (83)	-
≥ULN	-	3 (75)	-	-	2 (17)	-
Not done	11 (100)	1 (25)	84 (100)	20 (100)	-	22 (100)

[a] Stage I-II-III in case of NSCLC, stage I–II in case of PDAC, melanoma and bladder; NSCLC, non-small cell lung cancer; PDAC, pancreatic ductal adenocarcinoma; SD, standard deviation; ECOG, Eastern Cooperative Oncology Group; ULN, upper limit of normal; CEA ULN: 2.6 ng/mL; Cyfra21-1 ULN: 2.1 ng/mL; CA 19.9 ULN: 37 U/mL; LDH ULN: 247 IU/L.

2.2. Epithelial Cells and Monocytes Differentiate Patient and Healthy cfRNA

Most solid tumors, including lung adenocarcinomas, mainly develop from epithelial cells. We hypothesized that, if the tumor contributes substantially to the cfRNA pool in patients, the transcript signal accounted for by epithelial cells in patients should be significantly higher in patients than in controls.

To test this, we investigated the contribution of different cell types to the cfRNA pool and if contributions from particular cell types could differentiate patient samples from healthy samples. We thus performed a cell-type signal deconvolution in which we included reference maps from 45 different cell types. As seen from the heatmap in Figure 1A, signal contribution was similar from most cell types. However, there was a significantly higher (false discovery rate = 0.04) contribution from epithelial cells in cancer patients compared with healthy controls. Similarly, monocytes contributed significantly (FDR = 0.02) higher in healthy donors compared with patients (Figure 1B,C). Indeed, in a comparison between patient and healthy samples for epithelial, endothelial, and monocytes cell signatures, we identified epithelial cells and monocytes, but not endothelial cells, to show different transcript levels in patients and controls (Figure 1B,C).
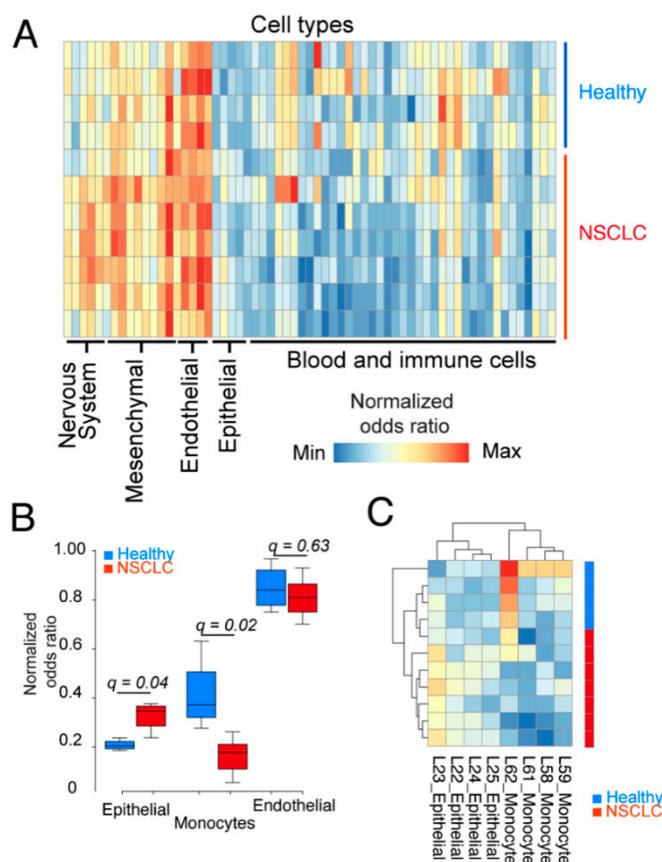
Figure 1. Epithelial cells contribute more to the cfRNA pool of lung cancer patients than controls. (**A**) Cell type deconvolution of 45 different cell types of transcriptomes derived from NSCLC patients (*n* = 7, stage IV) and healthy donors (*n* = 4). Data represents normalized odds ratios comparing number of enriched cell type-specific genes with random enrichment for each sample (rows) and cell type (columns). (**B**) Box plots represent healthy controls (blue) and NSLC patients (red) for epithelial, monocyte, and endothelial cell signatures in plasma. Data are derived from A; q-values represent the false discovery rate-adjusted *p* values. Only epithelial and monocyte signatures had a *q*-value < 0.05. (**C**) Hierarchical clustering of all epithelial and monocyte-specific signature data from (A) separate NSCLC patients.

2.3. cfRNA Sequencing Identifies Common Transcripts Upregulated in Cancer Patients

In the pilot cfRNA sequencing study, as shown in Supplementary Table S1, pairwise differential gene expression between plasma from NSCLC patients and healthy controls led to the identification of 605 upregulated transcripts (log2FC ≥ 1, FDR < 0.05). A total of 390 transcripts were selectively upregulated (log2FC ≥ 1, FDR < 0.05) in cfRNA from PDAC patients (Supplementary Table S2). A subset of 192 genes were overlappingly upregulated in plasma cfRNA derived from both PDAC and NSCLC patients (Figure 2a,b). Non-coding transcripts significantly contributed to those commonly upregulated genes, with pseudogenes, antisense transcripts, and long noncoding RNAs representing 49.2%, 19.2%, and 11.9%, respectively. Protein coding genes accounted for only 5.7% (Figure 2c).
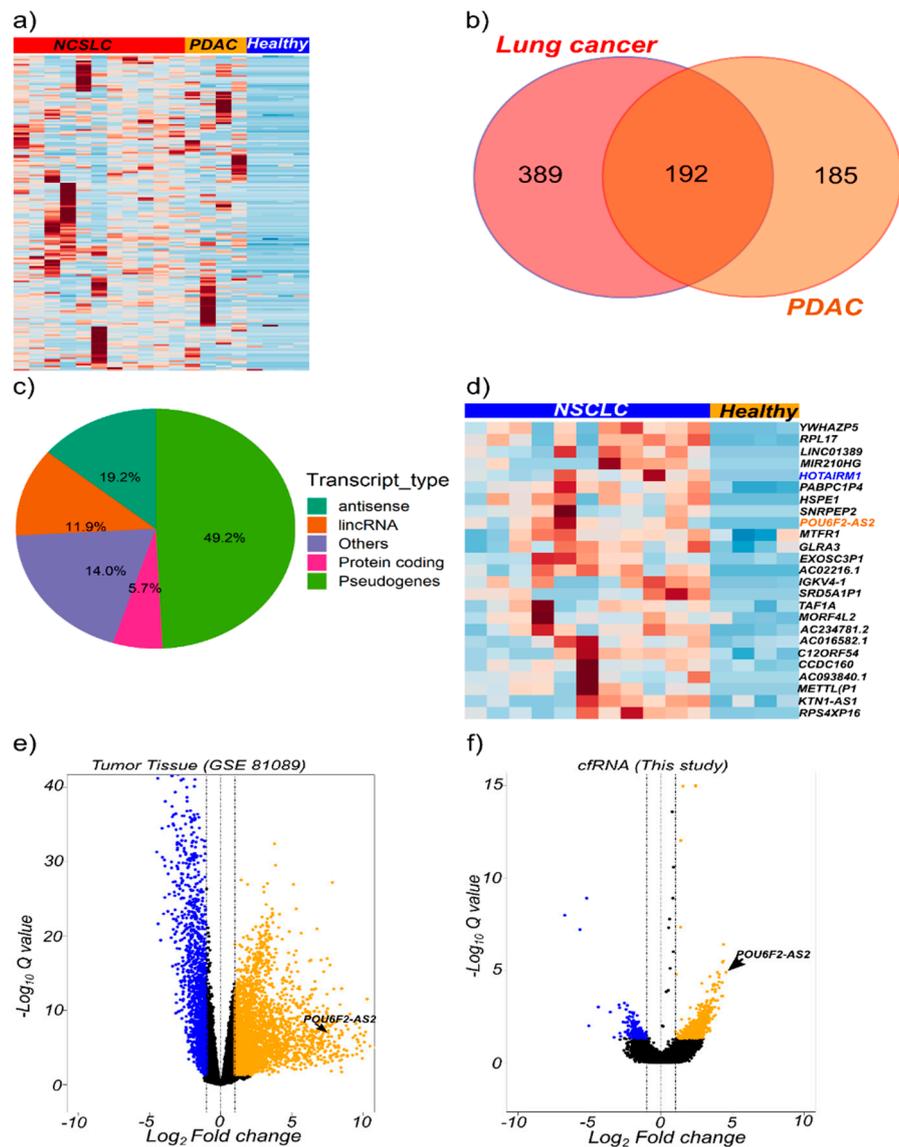
Figure 2. Total cfRNA transcriptome profiling identify potential cancer biomarkers. (**a**) Heatmap of genes upregulated in non-small cell lung cancer (NSCLC; $n = 11$ stage IV patients), and pancreatic ductal adenocarcinoma (PDAC; $n = 4$ stage IV patients) compared with healthy ($n = 4$ blood donors) plasma samples. Heatmap is plotted based on 192 common genes. (**b**) Venn diagram of the number of common upregulated genes in NSCLC and PDAC. (**c**) Pie chart showing the distribution of transcript types across the common upregulated genes. (**d**) Heatmap showing upregulated genes in NSCLC ($n = 11$) and healthy ($n = 4$) plasma samples based on the expression profiles of 24 genes differentially expressed in cfRNA and tumor-derived RNA. Candidate genes were derived from analyses of tumor/healthy tissue (GSE 81089 dataset) and RNA-Seq data from this study. (**e**) Volcano plots for differentially expressed genes in GSE 81089 dataset and (**f**) this study. Genes that passed the log$_2$ fold change ≥ 1 and FDR < 0.05 criteria are highlighted in blue (down-regulation) or yellow (up-regulation).

Gene expression data from NSCLC tumors was downloaded from the gene expression omnibus (GEO). Pair-wise differential gene expression was performed to identify genes that were upregulated in tumor tissue compared with non-tumor tissue. We identified 3060 genes to be upregulated in tumor tissue (log2FC ≥ 1 and FDR < 0.05). In order to identify tumor-relevant cfRNA transcripts, we intersected the list of upregulated genes in tumor tissue and plasma. This identified 24 upregulated transcripts (Figure 2d), including the previously reported HOTAIRM1 [33], (Supplementary Figure

S2A). HOTAIRM1 was upregulated in plasma but not in tumor tissue from NSCLC patients (Figure S2B). Conversely, the NSCLC-associated transcript MALAT1 was highly expressed in plasma from healthy donors but not NSCLC patients (Figure S2c,d).

We identified a previously reported cancer-related long non-coding RNA (lncRNA), POU6F2-AS2, to be upregulated in both lung cancer cfRNA and tumor tissue, as well as other transcripts, such as the lncRNA AC022126.1 (Figure 2d–f). LncRNA POU6F2-AS2 is known to be associated with another solid tumor, esophageal squamous cell carcinoma [34]. Hence, we selected POU6F2-AS2 and AC022126.1 to further explore the utility of detection of cfRNA transcripts as cancer biomarkers.

2.4. The lncRNAs POU6F2-AS2 and AC022126.1 Are Upregulated in Plasma from NSCLC Patients

POU6F2-AS2 was identified among other (Figure 2d) as a cancer-associated transcript by cfRNA-seq (Figure 3A). Using RT-ddPCR, we could confirm *POU6F2-AS2* as a tumor-associated transcript in plasma samples from 45 stage IV NSCLC patients and 39 stage I–III NSCLC patients (Figure 3B). There was no correlation between *POU6F2-AS2* expression and NSCLC stage (Figure 3C,D; AUC = 0.82 and 0.76; for late stage and early stage diseases, respectively) (Supplementary Figure S3A,B). Furthermore, *POU6F2-AS2* expression was similar in all stages of NSCLC and could not discriminate early stage from late stage disease (AUC = 0.42) (Supplementary Figure S2C). In contrast, the lncRNA AC022126.1 transcripts failed to discriminate cfRNA from cancer patients and healthy donors in our validation cohort (Figure 3E,F).

2.5. High POU6F2-AS2 Levels Is Associated with Higher Cyfra 21.1 Levels

Levels of the tumor marker Cyfra21-1 were significantly associated ($p = 0.017$) with higher plasma levels of *POU6F2-AS2* (Table 2). No significant association was found between our selected cfRNA transcripts and tumor stage, CEA, disease histology, smoking status, gender, and age. *POU6F2-AS2* cfRNA transcript levels were not correlated with clinical data for other cancer entities due to small sample numbers.

2.6. Serum cfRNA Concentration Correlates with POU6F2-AS2 Expression

Transcript abundance can be affected by the overall RNA yield (expressed in ng/mL) from a sample. Given that equal volumes of different samples may result in varying amounts of total cfRNA, we investigated the effect of input cfRNA amount on absolute transcript counts in plasma and serum. cfRNA yield per unit sample volume was higher in PDAC serum and could discriminate patients from healthy donors (Supplementary Figures S3D and S4A). Similarly, in PDAC sera, there was a very strong correlation ($r = 0.98$) between the transcript abundance of *POU6F2-AS* per unit volume and the total cfRNA yield (Supplementary Figure S4B). In NSCLC plasma samples, there was a weak correlation ($r = 0.33$) between total cfRNA yield and *POU6F2-AS2* transcript abundance (Supplementary Figure S4C). Similarly, there was no difference in plasma cfRNA yield between NSCLC at different stages and between patient and healthy controls (Supplementary Figure S4D).

Furthermore, the total amount of cfRNA isolated per unit volume of sample was about 4-fold lesser in plasma than in serum. This trend was observed in both healthy donor samples and PDAC patient samples (Supplementary Figure S5A,B).
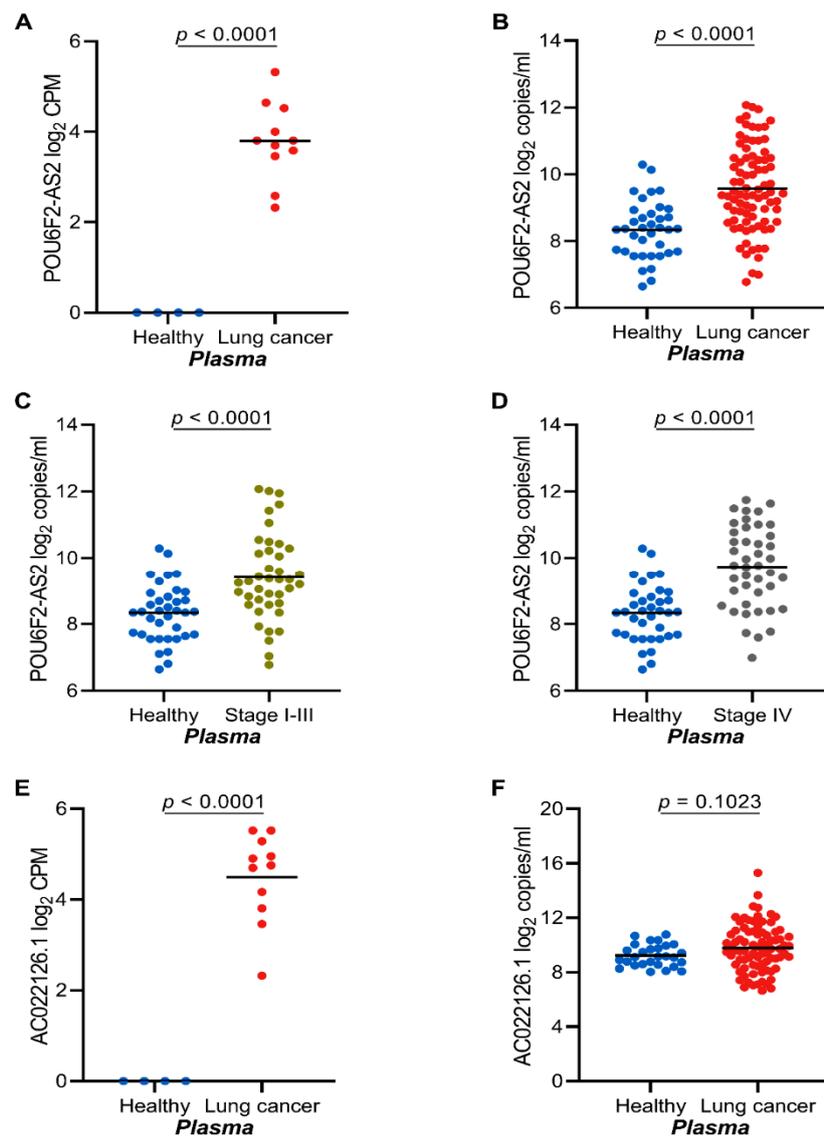
Figure 3. POU6F2-AS2 and AC022126.1 cfRNA transcripts are highly expressed in lung cancer. (**A**) Plasma expression of POU6F2-AS2 as profiled by total cfRNA sequencing in healthy donors ($n = 4$) and NSCLC ($n = 11$). Plasma expression of POU6F2-AS2 as profiled by real-time digital droplet PCR (RT-ddPCR) (**B**). in healthy donors ($n = 37$) and NSCLC ($n = 84$) patients (**C**). in healthy donors ($n = 37$) and early stage NSCLC ($n = 39$) patients (**D**). in healthy donors ($n = 37$) and late stage NSCLC ($n = 45$) patients. Plasma expression of AC022126.1 as profiled by (**E**). total cfRNA sequencing ($n = 4$ heathy donors; $n = 11$ NSCLC patients) and (**F**). by ddPCR ($n = 28$ heathy donors; $n = 84$ NSCLC patients). In (**A**–**F**) scatter dot-plots, the line indicates the mean; Student's *t*-test p values are indicated.

Table 2. Association of POU6F2-AS and AC022162.1 expression with baseline characteristics of NSCLC patients in validation cohort.

Variable	N (%)	POU6F2-AS2		AC022162.1	
	n (%)	Mean	p	Mean	p
Age					
<66 years	38 (45.2)	9.5	0.686	9.6	0.438
≥66 years	46 (54.8)	9.6		9.9	
Gender					
Male	42 (50)	9.8	0.058	9.9	0.430
Female	42 (50)	9.3		9.6	
Smoking					
Never smoker	5 (6)	9.2		9.8	
Past smoker	18 (21)	9.5	0.833	10.1	0.865
Current smoker	14 (17)	9.6		9.6	
Unknown	47 (56)	9.7		9.7	
ECOG status					
0	64 (76)	9.5		9.6	
1	16 (19)	9.9	0.517	10.3	0.355
2	1 (1)	10.9		11.4	
Unknown	3 (4)	9.7		9.6	
Histology					
Adenocarcinoma	23 (27)	9.6		9.9	
Non-squamous NSCLC	33 (39)	9.5	0.942	9.6	0.913
Squamous NSCLC	21 (25)	9.7		9.9	
Other	7 (8)	9.7		9.7	
Stage					
Early (I-II)	33 (39)	9.4	0.264	9.7	0.854
Late (III-IV)	51 (61)	9.7		9.8	
CEA					
<ULN	9 (11)	9.2		9.1	
≥ULN	32 (38)	9.8	0.371	9.8	0.467
Not done	43 (51)	9.5		9.9	
Cyfra21-1					
<ULN	5 (6)	8.2		8.3	
≥ULN	36 (43)	9.9	0.017	9.9	0.151
Not done	43 (51)	9.5		9.9	

p: student's *t*-test (2-groups) or 1-way ANOVA (3 or more groups); Mean values represent $\log_2$ RNA copies/ml plasma; CEA ULN (upper limit of normal): 2.6 ng/ML; Cyfra21-1 ULN (upper limit of normal): 2.1 ng/ML; NSCLC, non-small cell lung cancer.

2.7. The lncRNAs POU6F2-AS2 and AC022126.1 Are Upregulated in NSCLC Tissue

To investigate the tissue expression of candidate transcripts, we used RNA sequencing data from publicly available sources and RT-ddPCR validation on RNA isolated from FFPE NSCLC tumor tissue samples and adjacent normal lung tissue. In publicly available RNA sequencing data from tumor samples, the transcript levels of *POU6F2-AS2* was significantly upregulated in tumor-derived RNA as compared to non-tumor lung tissue (Figure 4A,B). Interestingly, tumor expression of *POU6F2-AS2* was independent of NSCLC stage (Figure 4C). Validation by RT-ddPCR analysis of RNA from FFPE samples revealed higher *POU6F2-AS2* transcript levels in tumor-derived RNA as compared to normal lung RNA (Figure 4D). There was no impact of NSCLC stage in terms of discrimination between tumor and normal tissue RNA (Figure 4E,F).
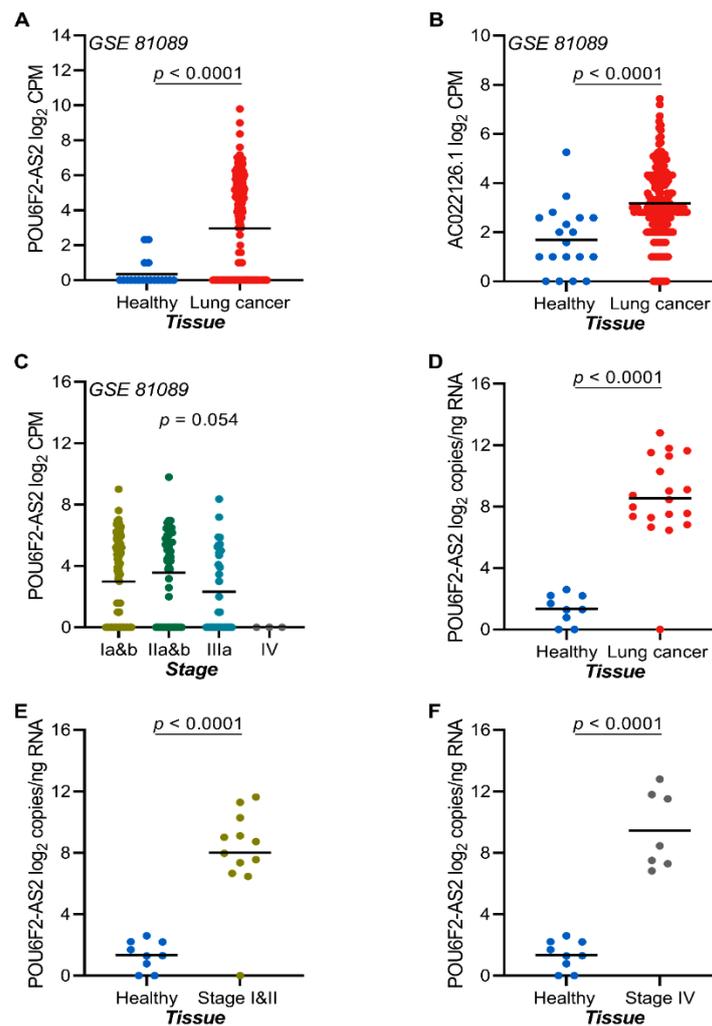
Figure 4. POU6F2-AS2 and AC022126.1 transcripts are highly expressed in lung cancer tissue and is stage independent. (**A**) Tissue expression of POU6F2-AS2 as profiled by RNA sequencing in lung cancer tissue and adjacent non-tumor lung GSE 81089 dataset). (**B**) Tissue expression of AC022126.1 as profiled by RNA sequencing in lung cancer tissue and adjacent non-tumor lung GSE 81089 dataset). (**C**) Association between POU6F2-AS2 expression and tumor stage in NSCLC tumor tissue samples from the GSE 81089 dataset (n = 199 tumor tissue samples and 19 non-tumor lung samples). Significance was tested in 1-way ANOVA. (**D**) Tissue expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors (n = 9) and NSCLC (n = 18) patients, (**E**) Tissue expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors (n = 9) and early stage NSCLC (n = 12) patients, and (**F**) Tissue expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors (n = 9) and late stage NSCLC (n = 7) patients. In (**A**–**F**) scatter dot-plots, the line indicates the mean; Student's t-test p values are indicated.

2.8. POU6F2-AS2 Is Upregulated in Additional Cancers

We next asked if *POU6F2-AS2* transcripts are exclusively upregulated in NSCLC plasma. To this end, we analyzed cfRNA isolated from serum samples of 20 additional PDAC, 22 UBC, and 22 healthy serum samples. Additionally, we analyzed cfRNA isolated from 12 MM plasma samples. In cfRNA from PDAC sera, *POU6F2-AS* transcript levels were significantly higher in patients than in healthy controls (Figure 5B). In malignant melanoma, plasma-derived cfRNA, *POU6F2-AS2* transcript levels were significantly higher in patients than controls (Figure 5C). No significant difference was observed in UBC samples and healthy controls (Figure 5D).
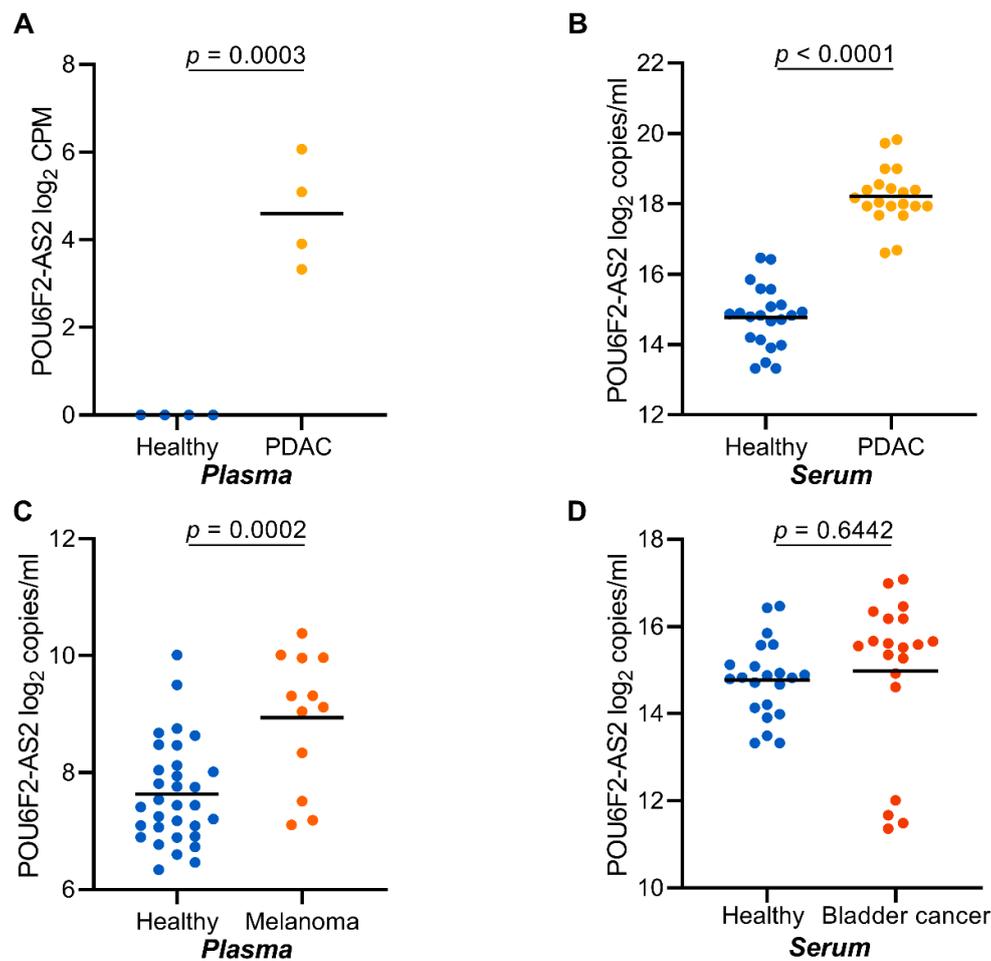
Figure 5. POU6F2-AS2 is highly expressed in cfRNA from other solid tumor entities. (**A**) Plasma expression of POU6F2-AS2 as profiled by total cfRNA sequencing in healthy donors ($n = 4$) and PDAC ($n = 4$). (**B**) Expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors' ($n = 22$) and PDAC ($n = 20$) patients' sera, (**C**) Expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors' ($n = 37$) and melanoma ($n = 12$) patients' plasma, and (**D**) Expression of POU6F2-AS2 as profiled by RT-ddPCR in healthy donors' ($n = 22$) and and bladder cancer ($n = 22$) patients' sera. In (**A**–**D**) scatter dot-plots, the line indicates the mean; Student's *t*-test *p* values are reported.

3. Discussion

Early diagnosis of cancer may improve survival outcome, as it allows for curative surgical intervention. Indeed, in lung cancer, the five-year survival rate of patient diagnosed at early stages is much higher than patients diagnosed at later stages [35,36]. Identification and validation of non-invasive biomarkers for early detection of cancer is, therefore, a considerable unmet need. Ideally, a biomarker should offer high sensitivity and sufficient specificity in a cost- and time-efficient manner to allow for its implementation into routine clinical practice [37]. Early diagnosis of lung cancer using LDCT has been shown to reduce mortality about 20% [38]. However, repeated radiation exposure, high cost, and high false positive rates represent significant challenges, and better alternatives are still needed.

Liquid biopsy-based test have been approved for detection of *EGFR* mutation ([mut]*EGFR*) in metastatic lung cancer [39]. While this represents a significant breakthrough, [mut]*EGFR* is only present in about 10–40% of all NSCLC cases [40]. Besides, the test targets metastatic patients, thus offering little or no chances for early detection for potential curative intervention. Although cfRNA was first reported to be present in several cancer entities more than a decade ago [24–31], its diagnostic

value has not been properly evaluated, especially in cancer. In spite of the fact that several legitimate concerns remain regarding the stability of cfRNA, a recent study highlighted the predictive strengths of cfRNA in pregnancy [32]. This demonstrates that cfRNA sequencing is feasible in the context of biomarker identification.

The study presented here aimed to explore a previously unexplored liquid biopsy compartment with significant opportunities for biomarker discovery. We used two independent highly sensitive technologies (next generation sequencing and droplet digital PCR) for profiling and absolute quantification of circulating-free and vesicle-encapsulated RNA in serum and plasma samples from four different solid tumor entities. Furthermore, we evaluated the expression of selected candidate transcripts in a large cohort of tumor samples and adjacent non-tumor tissue. We validated the expression pattern of a potential candidate in early and late-stage plasma samples from lung cancer patients, as well as tumor/normal tissue samples. Patient mutation status was not considered.

Unlike tissue-derived RNA, which is predominantly contributed by cells of a specific organ, cfRNA is usually contributed by several sources as blood circulates through the body. We performed transcript signal deconvolution using reference maps from 45 different cell types and identified a distinctive epithelial source of cancer-associated transcripts. Our study reveals that the majority of transcripts that are upregulated in cancer patient plasma are in fact non-coding transcripts. Indeed, some studies [41,42] have reported non-coding RNAs as liquid biopsy-based biomarkers in cancer. Very few coding transcripts were identified, most probably because of endogenous cellular degradation of these transcripts following translation by the *CNOT* complex [43]. Interestingly, we identified a previously reported lncRNA species (*HOTAIRM1*) that is associated with colorectal cancer [34,44]. Furthermore, we identified a panel of transcripts that are highly expressed in plasma from cancer patients (PDAC and lung cancers). We validated a previously reported cancer-associated transcript *POU6F2-AS2* in plasma and serum samples from lung cancer, pancreatic ductal adenocarcinoma, malignant melanoma, and bladder cancer samples. We showed that plasma and tissue transcript levels of certain transcripts correlate to a high degree, demonstrating specificity of the plasma marker for malignant tissue.

We investigated the performance of a candidate marker in early and late stage lung cancer patients. The transcript levels of *POU6F2-AS2* showed a good diagnostic performance in both early and late stage disease with an AUC of 0.82 and 0.76, respectively. This suggests that cfRNA transcript indeed may serve as a potential lung cancer biomarker. The association of tumor tissue in such discovery studies is of crucial importance, as we found out that the lung cancer associated lncRNA *MALAT1*, which is upregulated in tumor tissue [45] was upregulated in healthy plasma. Since cfRNA arises from diverse cellular sources, differential transcript contribution from different sources into the cfRNA pool may indeed affect individual transcript abundance. Hence, only transcripts with significant or exclusive tumor sources are very likely to be informative. Therefore, we performed cell-type signal deconvolution on our cfRNA sequencing data. We could show that only signals from epithelial origin could distinguish patients from healthy donors. Additionally, there was a significantly high contribution of monocytes to the cfRNA pool in healthy controls compared with patients. Furthermore, plasma total cfRNA abundance was similar in healthy donors and lung cancer patients plasma of all stages. This suggests that cfRNA abundance in plasma does not affect transcript expression patterns. On the other hand, serum derived total cfRNA abundance was significantly higher in PDAC serum samples compared with healthy serum. Although the diagnostic value of such differences was not explored here, it may represent an opportunity to be exploited.

While our study does not aim at evaluating the utility of cfRNA in disease surveillance and treatment monitoring, it does provide first-hand evidence of the usefulness of cfRNA in early diagnosis of solid tumors, especially NSCLC. Most importantly, the cross-entity validation and tumor tissue association of selected candidates strengthens the data and opens new opportunities for biomarker discovery. Finally, the fact that transcriptomic changes occur relatively faster than genomic changes

and that cellular transcriptional machinery amplifies analyte molecules improves sensitivity and also allows for detection of patients irrespective of the presence or absence of measurable genetic alterations.

4. Materials and Methods

4.1. Study Design and Patient Population

Patients analyzed in this study were prospectively recruited at the different study sites, either within clinical trials or clinical translational studies. Lung cancer patients were recruited at the outpatient unit of the Department of Medical Oncology at the West German Cancer Center, University Hospital Essen (stage IV) and at the Ruhrlandklinik (stage I–III) within the framework of the CEVIR study (https://www.transcanfp7.eu/index.php/abstract/cevir.html). Patients with either borderline or locally advanced but non-metastatic pancreatic ductal adenocarcinoma were recruited within the framework of the NEOPLAP prospective, randomized, open-label, phase II clinical trial study (https://clinicaltrials.gov/ct2/show/NCT02125136). Samples from bladder cancer patients were prospectively collected for an institutional biobank between 2008 and 2013, and samples from melanoma patients were collected within standard biobanking procedures at the Department of Dermatology of the University Hospital in Essen. Samples from healthy blood donors were obtained from blood donors at the Department of Transfusion Medicine, University Hospital Essen.

In our pilot study, we performed total cfRNA plasma sequencing on 11 NSCLC patients, 4 PDAC, and 4 healthy blood donors. In the validation study, we performed RT-ddPCR on cfRNA from 45 stage IV, as well as 39 stage I–III NSCL samples, 20 PDAC, 12 melanoma, 22 bladder cancer, and more than 61 healthy donors (serum/plasma). Additionally, RT-ddPCR was performed on a retrospective collection of 18 (unmatched to plasma) lung tumor FFPE samples and 9 adjacent non-tumor lung tissues. The local institutional review boards approved the studies at all sites, and all participants provided written informed consent for biomedical research allowing for molecular analysis on tissue and plasma samples. (NSCLC: 17-7740-BO, 14-6056-BO, PDAC: 17-7729-BO & 92/14 ff, MM: 16-7132-BO, UBC: 08-3942). All healthy controls included in the study provided consent for the use of their samples for molecular analysis, and ethical approval was obtained from the ethical committee at the Medical Faculty of the University of Essen (17-7729-BO).

4.2. Blood Sampling and Plasma/Serum Preparation

Blood samples were collected prior to treatment initiation. For plasma preparation, blood samples were collected in 7.5 mL EDTA blood tubes (ref # 01.1605.001, Sarstedt, Nümbrecht, Germany) and centrifuged at $200\times g$ for 10 min at 4 °C. The upper phase was then transferred into 2 mL tubes (cat # 296920064, Neolab, Heidelberg, Germany) and then centrifuged at $800\times g$ for 10 min at 4 °C. Finally, the supernatant was aliquoted into 2-mL tubes and then centrifuged at $16,000\times g$ for 10 min at 4 °C. The resulting plasma was transferred into new 2-mL tubes and stored at −80 °C until cfRNA extraction. When plasma preparation was not possible immediately, the samples were immediately kept at 4 °C and processed within 2 h of blood draw. Serum was prepared from blood collected in 10-mL BD Vacutainer® plus plastic tubes (cat # 366430, BD, Heidelberg, Germany). Blood samples were allowed to stand at room temperature for 30 min and then centrifuged at $2000\times g$ for 10 min at 4 °C. The resulting serum was aliquoted and stored at −80 °C until cfRNA isolation.

4.3. cfRNA Isolation and Quantification

Cell-free RNA was isolated using the Plasma/Serum RNA Purification Mini Kit (Cat. 55,000, Norgen Biotek, Canada), which isolates both cfRNA and vesicle-encapsulated RNA, following the manufacturer's instruction with slight modifications. Briefly, plasma/serum samples were allowed to thaw on ice and then 200 μL of plasma/serum cleared at $16,000 \times g$ for 2 min at 4 °C and the supernatant transferred into a new tube. The precleared plasma/serum was then combined with 3 volumes (600 μL) of Lysis buffer A, containing 0.01% β-mercaptoethanol and mixed by vortexing. To one volume of

lysate (800 µL), an equal amount of absolute ethanol was added and mixed by vortexing for 5 s. The lysate was then loaded onto a mini column and allowed to stand at room temperature for 10 min for RNA binding. The columns were then centrifuged at 6000 rpm for 3 min. The columns were then washed 3× with 400 µL of wash buffer A and dried by centrifugation at 13,000 rpm for 2 min. RNA was eluted in 25 µL of elution buffer A after incubation on the column for 15 min. The RNA concentration was measured using a Quantus fluorometer (Promega, Madison, Fitchburg WI, USA), and samples were stored at −80 °C until analyzed.

4.4. RNA Isolation from Tumor/Non-Tumor Tissue

RNA isolation was performed using the simplyRNA tissue kit (cat# AS1340, Promega Corporation), following manufacturer's instructions. Briefly, tissue pieces were cut on dry ice with a sterile scalpel blade and placed in a Precellys lysing kit tube (cat# P000912-LYSKO-A, Bertin Corp, Montigny-le Bretonneux, France). Sample homogenization buffer containing 0.02% thioglycerol was added, and the tube was placed into a precooled Precellys device. The Precellys machine was then programmed to rupture hard tissue for 2 min. After the tissue was completely disintegrated, 200 µL of lysis buffer was added to the cell suspension and then transferred into a cartridge of Maxwell SimplyRNA tissue kit (cat# AS1340, Promega Corporation), and 5 µL of DNase 1 was added to the appropriate well of the cartridge. RNA isolation was done following the installed kit-specific protocol and eluted in 60 µL of nuclease-free water. Quantification was carried out using a Quantus fluorometer (Promega), and samples were stored at −80 °C until analyses.

4.5. RNA Isolation from FFPE Sections

RNA was isolated from FFPE sections using the Maxwell® RSC RNA FFPE kit (Promega, cat# AS1440), following manufacturer's instructions. Briefly, 2.0 mm^3 of tissue sections were placed in a 1.5 mL Eppendorf tube and 300 µL of mineral oil was added and vortexed for 10 s. The samples were then heated to 80 °C for 2 min and brought to room temperature. The samples were digested and lysed and then pellet collected by centrifugation at 10,000× *g* for 20 s. The aqueous phase containing the pellet was heated at 56 °C for 15 min and then at 80 °C for 1 h and then incubated at room temperature for 30 min. Samples were DNA-digested and then loaded into the Maxwell cartridge, and RNA isolation was performed following the RSC RNA FFPE kit protocol in the fully automated system. RNA was eluted in 60 µL of nuclease-free water and stored at −80 °C until use.

4.6. In Silico Data Mining

In order to identify tumor-associated transcripts that were upregulated in plasma samples from cancer patients, we downloaded publicly available RNA sequencing data from the gene expression omnibus (GEO). A lung cancer data set GSE81089 comprising of both early and late stage cancers was downloaded [46]. The data set was composed of 199 NSCLC fresh frozen samples and 19 paired non-tumor lung tissues.

4.7. Total cfRNA Sequencing

Total RNA sequencing libraries were prepared from 450 pg of total cfRNA using the SMARTer Stranded Total RNA-Seq Kit v2—Pico Input Mammalian (cat# 634412, Takara Bio, Mountain View, CA, USA), after depletion of ribosomal RNA. RNA libraries were sequenced on an Illumina Novaseq 6000 with 100 bp paired end read. Fastq files were quality checked with the Fastqc tool, and the reads were trimmed with Trimmomatic tool. Reads were mapped to the human genome (Grch 38), and features were quantified by means of Htseq counts. Gene expression matrices with raw counts were processed with Deseq2 [47] for differential gene expression and validated with edgeR and LIMMA [48,49].

4.8. cfRNA Transcript Deconvolution

Cell-type cfRNA signal deconvolution was performed using reference maps from 45 different cell types, as previously described [50]. A complete list of all cell types included in the deconvolution is presented in Supplementary Table S3. We then used random forest ensemble classifier to rank the importance of cell type signals differentiating patient and healthy samples as previously described [51]. Hierarchical clustering and statistical test (Wilcoxon-Mann-Whitney U test) were performed in R environment.

4.9. Transcript Quantification by RT-ddPCR

Two transcripts, *POU6F2-AS2* and *AC022126.1,* both of which had no measurable transcripts in the healthy donors in the cfRNA-seq data, were further analyzed. One of these two transcripts, *POU6F2-AS2* has previously been associated with cancer pathogenesis. Circulating levels of *POU6F2-AS2* and *AC022126.1* transcripts were measured by means of RT-ddPCR using the 1-Step RT-ddPCR Advanced Kit for Probes (Bio-Rad, Hercules, CA, USA) in samples from our validation cohort, which included 45 stage IV lung cancer patients, 39 stage I–III lung cancer patients, 20 PDAC, 22 bladder cancer, and more than 65 healthy control samples. Samples from the validation cohort were both plasma and serum. All reactions were performed in duplicates using 2 μL of cfRNA from each sample. The reaction components were constituted following manufacturers instruction to a final volume of 22 μL, of which 20 μL were used for droplet generation in a QX100™/QX200™ droplet generator (Bio-Rad). RT-ddPCR reactions were performed in a C1000 Touch™ thermocycler (Bio-rad), and droplets were read in a QX100™/QX200™ droplet reader (Bio-Rad). The raw transcript concentration (copies/20 μL reaction) was used to determine the absolute transcript load per ml of plasma/serum using the formula:

$$Copies/mL = C \times EV \times 1000/ISV \times TPV$$

 where

C = copies per 20 μL reaction
EV = elution volume
ISV = cfRNA input volume
TPV = total plasma/serum volume used of cfRNA isolation

For tissue-derived samples, RT-ddPCR was performed on 4 ng of total RNA for each sample and data expressed per ng of RNA.

4.10. Statistical Analysis

Students' *t*-test was used to compare the mean of two groups, and the one-way ANOVA was used to compare three or more groups. Categorical data was analyzed by Fishers' exact test or by Chi-squared test. Strength of relationship between cfRNA copies and amounts was assessed by Person correlation. Statistical significance was set to a *p* value < 0.05. The diagnostic performance of the marker was evaluated with the ROCR package [52]. Data analysis was performed in the R version 3.6-environment and Graphpad prism version 8.3 (GraphPad Software, Inc, La Jolla, CA, USA).

5. Conclusions

We provided evidence on the utility of cfRNA for early detection of solid tumors irrespective of mutation status and identified candidate transcripts with plasma/tissue mirror images. Our findings support further validation of cfRNA detection by RT-ddPCR as a biomarker for early detection of solid cancers.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6694/12/2/353/s1, Figure S1: Total cfRNA sequencing quality control as estimated by STAR and RSEQC, Figure S2: Tissue expression

of HOTAIRM1 and MALAT1 in lung cancer tissue and non-tumor lung tissue Figure S3: Diagnostic performance of *POU6F2-AS2* in NSCLC and serum total cfRNA in PDAC, Figure S4: Association between total cfRNA concentration and *POU6F2-AS2* copies in plasma and serum, Figure S5: Total cfRNA yield is higher in serum compared with plasma, Figure S6: Transcript expression data for HOTAIRM1, MALAT1 and POU6F2-AS2 in cfRNA and the TCGA data set.

Author Contributions: M.M. investigation, resources data curation review and edition, R.V.; investigation review and edition, B.H., investigation, resources, I.C.; formal analysis. J.F.; formal analysis, A.S.; resources, B.S.; resources, P.A.H.; resources, C.A.K. resources, project administration, T.S.; investigation, resources, H.R.; investigation, resources, N.B., investigation, A.R.; resources, C.A.; resources, V.K.; resources, M.W.; writing—review and editing, J.T.S.; resources, M.S.; supervision, funding acquisition, Conceptualization, project administration, S.S.L.; Conceptualization, methodology, investigation, formal analysis, original draft preparation, project administration. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [CrossRef]

2. Herbst, R.S.; Morgensztern, D.; Boshoff, C. The biology and management of non-small cell lung cancer. *Nature* **2018**, *553*, 446–454. [CrossRef]

3. Poroyko, V.; Mirzapoiazova, T.; Nam, A.; Mambetsariev, I.; Mambetsariev, B.; Wu, X.; Husain, A.; Vokes, E.E.; Wheeler, D.L.; Salgia, R. Exosomal mirnas species in the blood of small cell and non-small cell lung cancer patients. *Oncotarget* **2018**, *9*, 19793–19806. [CrossRef]

4. Lu, S.; Kong, H.; Hou, Y.; Ge, D.; Huang, W.; Ou, J.; Yang, D.; Zhang, L.; Wu, G.; Song, Y.; et al. Two plasma microrna panels for diagnosis and subtype discrimination of lung cancer. *Lung Cancer* **2018**, *123*, 44–51. [CrossRef]

5. Gridelli, C.; Rossi, A.; Carbone, D.P.; Guarize, J.; Karachaliou, N.; Mok, T.; Petrella, F.; Spaggiari, L.; Rosell, R. Non-small-cell lung cancer. *Nat. Rev. Dis. Primers* **2015**, *1*, 15009. [CrossRef]

6. Lai, Y.; Wang, X.; Zeng, T.; Xing, S.; Dai, S.; Wang, J.; Chen, S.; Li, X.; Xie, Y.; Zhu, Y.; et al. Serum vegf levels in the early diagnosis and severity assessment of non-small cell lung cancer. *J. Cancer* **2018**, *9*, 1538–1547. [CrossRef]

7. Scagliotti, G.; Hanna, N.; Fossella, F.; Sugarman, K.; Blatter, J.; Peterson, P.; Simms, L.; Shepherd, F.A. The differential efficacy of pemetrexed according to nsclc histology: A review of two phase iii studies. *Oncology* **2009**, *14*, 253–263. [CrossRef]

8. Mok, T.S.; Wu, Y.L.; Thongprasert, S.; Yang, C.H.; Chu, D.T.; Saijo, N.; Sunpaweravong, P.; Han, B.; Margono, B.; Ichinose, Y.; et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **2009**, *361*, 947–957. [CrossRef]

9. Mitsudomi, T.; Morita, S.; Yatabe, Y.; Negoro, S.; Okamoto, I.; Tsurutani, J.; Seto, T.; Satouchi, M.; Tada, H.; Hirashima, T.; et al. Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (wjtog3405): An open label, randomised phase 3 trial. *Lancet. Oncol.* **2010**, *11*, 121–128. [CrossRef]

10. Xing, S.; Zeng, T.; Xue, N.; He, Y.; Lai, Y.Z.; Li, H.L.; Huang, Q.; Chen, S.L.; Liu, W.L. Development and validation of tumor-educated blood platelets integrin alpha 2b (itga2b) rna for diagnosis and prognosis of non-small-cell lung cancer through rna-seq. *Int. J. Biol. Sci.* **2019**, *15*, 1977–1992. [CrossRef]

11. Cassim, S.; Chepulis, L.; Keenan, R.; Kidd, J.; Firth, M.; Lawrenson, R. Patient and carer perceived barriers to early presentation and diagnosis of lung cancer: A systematic review. *Bmc Cancer* **2019**, *19*, 25. [CrossRef] [PubMed]

12. Gross, C.P. Screening with low-dose computed tomography reduced lung cancer mortality in high-risk patients. *Ann. Intern. Med.* **2011**, *155*, JC5-6. [CrossRef] [PubMed]

13. Aberle, D.R.; Adams, A.M.; Berg, C.D.; Black, W.C.; Clapp, J.D.; Fagerstrom, R.M.; Gareen, I.F.; Gatsonis, C.; Marcus, P.M.; Sicks, J.D.; et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **2011**, *365*, 395–409. [PubMed]

14. Patz, E.F.; Pinsky, P.; Gatsonis, C.; Sicks, J.D.; Kramer, B.S.; Tammemagi, M.C.; Chiles, C.; Black, W.C.; Aberle, D.R.; Writ, N.O.M. Overdiagnosis in low-dose computed tomography screening for lung cancer. *Jama Intern. Med.* **2014**, *174*, 269–274. [CrossRef]

15. Wiener, R.S.; Schwartz, L.M.; Woloshin, S.; Welch, H.G. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: An analysis of discharge records. *Ann. Intern. Med.* **2011**, *155*, 137–144. [CrossRef]

16. Loke, S.Y.; Lee, A.S.G. The future of blood-based biomarkers for the early detection of breast cancer. *Eur. J. Cancer* **2018**, *92*, 54–68. [CrossRef]

17. Ren, S.X.; Zhang, S.C.; Jiang, T.; He, Y.Y.; Ma, Z.Y.; Cai, H.R.; Xu, X.H.; Li, Y.; Cai, W.J.; Zhou, J.; et al. Early detection of lung cancer by using an autoantibody panel in chinese population. *Oncoimmunology* **2018**, *7*, e1384108. [CrossRef]

18. Molina, R.; Marrades, R.M.; Auge, J.M.; Escudero, J.M.; Vinolas, N.; Reguart, N.; Ramirez, J.; Filella, X.; Molins, L.; Agusti, A. Assessment of a combined panel of six serum tumor markers for lung cancer. *Am. J. Resp. Crit. Care* **2016**, *193*, 427–437. [CrossRef]

19. Ludwig, J.A.; Weinstein, J.N. Biomarkers in cancer staging, prognosis and treatment selection. *Nat. Rev. Cancer* **2005**, *5*, 845–856. [CrossRef]

20. Maziak, W. Carcinoembryonic antigen (CEA) levels in hookah smokers, cigarette smokers and non-smokers–a comment. *JPMA J. Pak. Med Assoc.* **2008**, *58*, 155.

21. Correa-Gallego, C.; Warshaw, A.L.; Fernandez-del Castillo, C. Fluid cea in ipmns: A useful test or the flip of a coin? *Am. J. Gastroenterol.* **2009**, *104*, 796–797. [CrossRef] [PubMed]

22. Mensah, M.; Borzi, C.; Verri, C.; Suatoni, P.; Conte, D.; Pastorino, U.; Orazio, F.; Sozzi, G.; Boeri, M. Microrna based liquid biopsy: The experience of the plasma mirna signature classifier (msc) for lung cancer screening. *J. Vis. Exp. JOVE* **2017**. [CrossRef] [PubMed]

23. Mestdagh, P.; Hartmann, N.; Baeriswyl, L.; Andreasen, D.; Bernard, N.; Chen, C.; Cheo, D.; D'Andrade, P.; DeMayo, M.; Dennis, L.; et al. Evaluation of quantitative mirna expression platforms in the microrna quality control (mirqc) study. *Nat. Methods* **2014**, *11*, 809–815. [CrossRef] [PubMed]

24. Sozzi, G.; Boeri, M.; Rossi, M.; Verri, C.; Suatoni, P.; Bravi, F.; Roz, L.; Conte, D.; Grassi, M.; Sverzellati, N.; et al. Clinical utility of a plasma-based mirna signature classifier within computed tomography lung cancer screening: A correlative mild trial study. *J. Clin. Oncol.* **2014**, *32*, 768–773. [CrossRef] [PubMed]

25. Kopreski, M.S.; Benko, F.A.; Kwak, L.W.; Gocke, C.D. Detection of tumor messenger rna in the serum of patients with malignant melanoma. *Clin. Cancer Res.* **1999**, *5*, 1961–1965. [PubMed]

26. Lo, K.W.; Lo, Y.M.D.; Leung, S.F.; Tsang, Y.S.; Chan, L.Y.S.; Johnson, P.J.; Hjelm, N.M.; Lee, J.C.K.; Huang, D.P. Analysis of cell-free epstein-barr virus-associated rna in the plasma of patients with nasopharyngeal carcinoma. *Clin. Chem.* **1999**, *45*, 1292–1294. [CrossRef]

27. Chen, X.Q.; Bonnefoi, H.; Pelte, M.F.; Lyautey, J.; Lederrey, C.; Movarekhi, S.; Schaeffer, P.; Mulcahy, H.E.; Meyer, P.; Stroun, M.; et al. Telomerase rna as a detection marker in the serum of breast cancer patients. *Clin. Cancer Res.* **2000**, *6*, 3823–3826.

28. Silva, J.M.; Dominguez, G.; Silva, J.; Garcia, J.M.; Sanchez, A.; Rodriguez, O.; Provencio, M.; Espana, P.; Bonilla, F. Detection of epithelial messenger rna in the plasma of breast cancer patients is associated with poor prognosis tumor characteristics. *Clin. Cancer Res.* **2001**, *7*, 2821–2825.

29. Chuen, S.; Wong, C.; Fong, S. Quantification of plasma beta-catenin mrna in colorectal cancer and adenoma patients. *Clin. Cancer Res.* **2004**, *10*, 1613–1617, Erratum in **2006**, *12*, 996.

30. Wong, S.C.C.; Lo, S.F.E.; Cheung, M.T.; Ng, K.O.E.; Tse, C.W.; Lai, B.S.P.; Lee, K.C.; Lo, Y.M.D. Quantification of plasma beta-catenin mrna in colorectal cancer and adenoma patients. *Clin. Cancer Res.* **2004**, *10*, 1613–1617. [CrossRef]

31. Dasi, F.; Lledo, S.; Garcia-Granero, E.; Ripoll, R.; Marugan, M.; Tormo, M.; Garcia-Conde, J.; Alino, S.F. Real-time quantification in plasma of human telomerase reverse transcriptase (htert) mrna: A simple blood test to monitor disease in cancer patients. *Lab Investig.* **2001**, *81*, 767–769. [CrossRef] [PubMed]

32. Miura, N.; Shiota, G.; Nakagawa, T.; Maeda, Y.; Sano, A.; Marumoto, A.; Kishimoto, Y.; Murawaki, Y.; Hasegawa, J. Sensitive detection of human telomerase reverse transcriptase mrna in the serum of patients with hepatocellular carcinoma. *Oncology* **2003**, *64*, 430–434. [CrossRef] [PubMed]

33. Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T. Rocr: Visualizing classifier performance in r. *Bioinformatics* **2005**, *21*, 3940–3941. [CrossRef] [PubMed]

34. Wan, L.D.; Kong, J.L.; Tang, J.L.; Wu, Y.H.; Xu, E.P.; Lai, M.D.; Zhang, H.H. Hotairm1 as a potential biomarker for diagnosis of colorectal cancer functions the role in the tumour suppressor. *J. Cell Mol. Med.* **2016**, *20*, 2036–2044. [CrossRef]

35. Liu, J.; Sun, X.; Zhu, H.; Qin, Q.; Yang, X.; Sun, X. Long noncoding rna pou6f2-as2 is associated with oesophageal squamous cell carcinoma. *J. Biochem.* **2016**, *160*, 195–204. [CrossRef]

36. Bazargan, M.; Lucas-Wright, A.; Jones, L.; Vargas, R.; Vadgama, J.V.; Evers-Manly, S.; Maxwell, A.E. Understanding perceived benefit of early cancer detection: Community-partnered research with african american women in south los angeles. *J. Women's Health* **2015**, *24*, 755–761. [CrossRef]

37. Zhu, W.Y.; Tan, L.L.; Wang, Z.Y.; Wang, S.J.; Xu, L.Y.; Yu, W.; Chen, Z.J.; Zhang, Y.K. Clinical characteristics and advantages of primary peripheral micro-sized lung adenocarcinoma over small-sized lung adenocarcinoma. *Eur. J. Cardio-Thorac. Surg.* **2016**, *49*, 1095–1102. [CrossRef]

38. Rubenstein, J.H.; Vakil, N.; Inadomi, J.M. The cost-effectiveness of biomarkers for predicting the development of oesophageal adenocarcinoma. *Aliment. Pharmacol. Ther.* **2005**, *22*, 135–146. [CrossRef]

39. Midthun, D.E. Early detection of lung cancer. *F1000Research* **2016**, *5*. [CrossRef]

40. Kwapisz, D. The first liquid biopsy test approved. Is it a new era of mutation testing for non-small cell lung cancer? *Ann. Transl. Med.* **2017**, *5*, 46. [CrossRef]

41. Pao, W.; Miller, V.; Zakowski, M.; Doherty, J.; Politi, K.; Sarkaria, I.; Singh, B.; Heelan, R.; Rusch, V.; Fulton, L.; et al. Egf receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13306–13311. [CrossRef]

42. Zeuschner, P.; Linxweiler, J.; Junker, K. Non-coding rnas as biomarkers in liquid biopsies with a special emphasis on extracellular vesicles in urological malignancies. *Expert Rev. Mol. Diagn.* **2019**, *18*, 1–17. [CrossRef]

43. Jiang, N.; Pan, J.; Fang, S.; Zhou, C.; Han, Y.; Chen, J.; Meng, X.; Jin, X.; Gong, Z. Liquid biopsy: Circulating exosomal long noncoding rnas in cancer. Clinica chimica acta. *Int. J. Clin. Chem.* **2019**, *495*, 331–337.

44. Collart, M.A. The ccr4-not complex is a key regulator of eukaryotic gene expression. *Wiley interdisciplinary reviews. RNA* **2016**, *7*, 438–454. [CrossRef]

45. Li, Z.X.; Zhu, Q.N.; Zhang, H.B.; Hu, Y.; Wang, G.; Zhu, Y.S. Malat1: A potential biomarker in cancer. *Cancer Manag. Res.* **2018**, *10*, 6757–6768. [CrossRef]

46. Ngo, T.T.M.; Moufarrej, M.N.; Rasmussen, M.L.H.; Camunas-Soler, J.; Pan, W.Y.; Okamoto, J.; Neff, N.F.; Liu, K.L.; Wong, R.J.; Downes, K.; et al. Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* **2018**, *360*, 1133–1136. [CrossRef]

47. Mezheyeuski, A.; Bergsland, C.H.; Backman, M.; Djureinovic, D.; Sjoblom, T.; Bruun, J.; Micke, P. Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.* **2018**, *244*, 421–431. [CrossRef]

48. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]

49. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]

50. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. Edger: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]

51. Cima, I.; Kong, S.L.; Sengupta, D.; Tan, I.B.; Phyo, W.M.; Lee, D.; Hu, M.; Iliescu, C.; Alexander, I.; Goh, W.L.; et al. Tumor-derived circulating endothelial cell clusters in colorectal cancer. *Sci. Transl. Med.* **2016**, *8*, 345ra389. [CrossRef] [PubMed]

52. Tan, W.J.; Cima, I.; Choudhury, Y.; Wei, X.; Lim, J.C.; Thike, A.A.; Tan, M.H.; Tan, P.H. A five-gene reverse transcription-pcr assay for pre-operative classification of breast fibroepithelial lesions. *Breast Cancer Res. BCR* **2016**, *18*, 31. [CrossRef] [PubMed]